Probab. Theory Relat. Fields (2011) 151:191–231
DOI 10.1007/s00440-010-0297-4

Quasi-stationary distributions for structured birth and death processes with mutations

Pierre Collet · Servet Martínez · Sylvie Méléard · Jaime San Martín

Received: 17 March 2009 / Revised: 1 April 2010 / Published online: 5 May 2010

Abstract We study the probabilistic evolution of a birth and death continuous time measure-valued process with mutations and ecological interactions. The individuals are characterized by (phenotypic) traits that take values in a compact metric space. Each individual can die or generate a new individual. The birth and death rates may depend on the environment through the action of the whole population. The offspring can have the same trait or can mutate to a randomly distributed trait. We assume that the population will be extinct almost surely. Our goal is the study, in this infinite dimensional framework, of the quasi-stationary distributions of the process conditioned on non-extinction. We first show the existence of quasi-stationary distributions. This result is based on an abstract theorem proving the existence of finite eigenmeasures for some positive operators. We then consider a population with constant birth and death rates per individual and prove that there exists a unique quasi-stationary distribution with maximal exponential decay rate. The proof of uniqueness is based on an absolute continuity property with respect to a reference measure.

P. Collet (✉)
CNRS Physique Théorique, Ecole Polytechnique, 91128 Palaiseau Cedex, France
e-mail: collet@cpht.polytechnique.fr

S. Martínez · J. San Martín
Departamento Ingeniería Matemática, Centro Modelamiento Matemático,
Universidad de Chile, UMI 2807 CNRS, Casilla 170-3, Correo 3, Santiago, Chile
e-mail: smartine@dim.uchile.cl

J. San Martín
e-mail: jsanmart@dim.uchile.cl

S. Méléard
Ecole Polytechnique, CMAP, 91128 Palaiseau Cedex, France
e-mail: meleard@cmap.polytechnique.fr

Keywords Quasi-stationary distribution · Birth–death process · Population dynamics · Measured valued Markov processes

Mathematics Subject Classification (2000) Primary 92D25; Secondary 60K35 · 60J70 · 60J80

1 Introduction and main results

1.1 Introduction

We consider a general discrete model describing a structured population with a microscopic individual-based and stochastic point of view. The dynamics takes into account all reproduction and death events. Each individual is characterized by an heritable quantitative parameter, usually called *trait*, which can for example be the expression of its genotype or phenotype. During the reproduction process, mutations of the trait can occur, implying some variability in the trait space. Moreover, the individuals can die. In the general model, the individual reproduction and death rates, as well as the mutation distribution, depend on the trait of the individual and on the whole population. In particular, cooperation or competition between individuals in this population are taken into account.

To precise these notions let us introduce some notation. For a topological space A we denote by $\mathcal{B}(A)$ the Borel σ-field, by $\mathcal{P}(A)$ the set of probability measures on $(A, \mathcal{B}(A))$ and by $\mathcal{M}(A)$ the set of (positive) measures on $(A, \mathcal{B}(A))$.

In our model the set of traits $\mathbb{T}$ is a compact metric space with metric d. For convenience we assume diameter$(\mathbb{T}) = 1$. The structured population is described by a finite point measure on $\mathbb{T}$. Thus, the state space, denoted by $\mathcal{A}$, is the set of all finite point measures which is a subset of $\mathcal{M}(\mathbb{T})$.

In all what follows, the set $\mathcal{A}$ will be endowed with the Prokhorov metric which makes it a Polish space (complete separable metric space). This metric induces the weak convergence topology for which $\mathcal{A}$ is closed in the set of finite positive measures (See for example [7, Chapter 7] and "Appendix").

A configuration $\eta \in \mathcal{A}$ can also be described by $(\eta_y : y \in \mathbb{T})$ with $\eta_y \in \mathbb{Z}_+ = \{0, 1, \ldots\}$, where only a finite subset of elements $y \in \mathbb{T}$ satisfy $\eta_y > 0$. The finite set of present traits (i.e., traits of alive individuals) is denoted by

$$\{\eta\} := \{y \in \mathbb{T} : \eta_y > 0\}$$

and it is called the support of η, its cardinality is denoted by $\#\eta$. We denote by $\|\eta\| = \sum_{y \in \{\eta\}} \eta_y$ the total number of individuals in η. The void configuration is denoted by $\eta = 0$, hence $\#0 = \|0\| = 0$ and we define $\mathcal{A}^{-0} := \mathcal{A} \setminus \{0\}$ the set of nonempty configurations with the induced topology.

The structured population dynamics is given by an individual-based model, taking into account each (clonal or mutation) birth and death events.

The clonal birth rate, the mutation birth rate and the death rate per individual with trait y in a population $\eta \in \mathcal{A}$, are denoted respectively by $b_y(\eta)$, $m_y(\eta)$ and $\lambda_y(\eta)$. The total reproduction rate for an individual with trait $y \in \{\eta\}$ is equal to $b_y(\eta) + m_y(\eta)$. We assume $\lambda_y(0) = b_y(0) = m_y(0) = 0$ for all $y \in \mathbb{T}$, which is natural for population dynamics. In what follows we assume that the functions

$$\lambda_y(\eta), b_y(\eta), m_y(\eta) : \mathbb{T} \times \mathcal{A}^{-0} \to \mathbb{R}_+ \quad \text{are continuous and strictly positive.} \quad (1)$$

We notice that $\lambda_y(\eta), b_y(\eta), m_y(\eta)$ do not necessarily vanish when $\eta_y = 0$. This is not a problem because in the transition rates what matters are the quantities $\eta_y \lambda_y(\eta), \eta_y b_y(\eta), \eta_y m_y(\eta)$.

Let σ be a fixed non-atomic probability measure on $(\mathbb{T}, \mathcal{B}(\mathbb{T}))$, whose support coincides with $\mathbb{T}$. The density location function of the mutations is $g : \mathbb{T} \times \mathbb{T} :\to \mathbb{R}_+$, $(y, z) \to g(y, z)$, where $g(y, \cdot)$ is the probability density of the trait of the new mutated individual born from y. It satisfies

$$\int_{\mathbb{T}} g(y, z) d\sigma(z) = 1 \qquad \text{for all} \quad y \in \mathbb{T}.$$

We assume that the function g is jointly continuous. To simplify notations we express the mutation part using the location kernel $G(\eta, z) : \mathcal{A} \times \mathbb{T} \to \mathbb{R}^+$ given by

$$\forall \eta \in \mathcal{A} \forall z \in \mathbb{T}, \quad G(\eta, z) = \sum_{y \in \{\eta\}} \eta_y m_y(\eta) g(y, z).$$

Note that the ratio $G(\eta, z) d\sigma(z) / \int_{\mathbb{T}} G(\eta, z') d\sigma(z')$ is the probability that, given there is a mutation from η, the new trait is located at z. The function G is continuous on $\mathcal{A} \times \mathbb{T}$ under condition (1) (see Lemma 2.6 below).

We define a continuous time pure jump Markov process $Y = (Y_t)$ taking values on $\mathcal{A}$ using the above transition rates.

We denote by $Q : \mathcal{A} \times \mathcal{B}(\mathcal{A}) \to \mathbb{R}_+$, $(\eta, B) \to Q(\eta, B)$, the kernel of measure jump rates given by

$$Q(\eta, B) = \sum_{y \in \{\eta\}, \eta + \delta_y \in B} \eta_y b_y(\eta) + \sum_{y \in \{\eta\}, \eta - \delta_y \in B} \eta_y \lambda_y(\eta) + \int_{\eta + \delta_z \in B} G(\eta, z) d\sigma(z).$$

To simplify notation, sometimes we use $Q(\eta, \eta') = Q(\eta, \{\eta'\})$.

The total jump rate $Q(\eta) = Q(\eta, \mathcal{A})$ at configuration η is always finite and given by

$$Q(\eta) = \sum_{y \in \{\eta\}} (\eta_y b_y(\eta) + \eta_y \lambda_y(\eta) + \eta_y m_y(\eta))$$

The construction of this process Y with càdlàg trajectories is the canonical one. Namely, assume that the process starts from $Y_0 = \eta$. Then, after an exponential time of parameter $Q(\eta)$, the process jumps to $\eta + \delta_y$ for $y \in \{\eta\}$ with probability $\eta_y b_y(\eta)/Q(\eta)$, or to $\eta - \delta_y$ for $y \in \{\eta\}$ with probability $\eta_y \lambda_y(\eta)/Q(\eta)$, or with probability density $G(\eta, z)/Q(\eta)$ (w.r.t. σ) to a point $\eta + \delta_z$ for $z \in \mathbb{T} \setminus \{\eta\}$. The process restarts independently at the new configuration.

The process Y can have explosions. We will avoid explosion assuming throughout the paper that

$$\sup_{\eta \in \mathcal{A}} \sup_{y \in \{\eta\}} (b_y(\eta) + m_y(\eta)) < \infty. \tag{2}$$

This condition also guarantees the existence of the process $(Y_t : t \geq 0)$ as the unique solution of a stochastic differential equation driven by Poisson point measures. This is proved in Sect. 2 following [11,5]. In order to control the martingale properties of (Y_t), we will assume there exist $p \geq 1$ and $c > 0$ such that

$$\forall \eta \in \mathcal{A} : \sup_{y \in \mathbb{T}} \lambda_y(\eta) \leq c \|\eta\|^p. \tag{3}$$

Since $Q(0) = 0$, the void configuration is an absorbing state for the process Y. We denote by

$$\mathscr{T}_0 = \inf\{t \geq 0 : Y_t = 0\}$$

the extinction time. We will make later on further assumptions ensuring that the process almost surely extincts when starting from any initial configuration:

$$\forall \eta \in \mathcal{A} : \mathbb{P}_\eta(\mathscr{T}_0 < \infty) = 1.$$

So, in our setting we assume that competition between individuals, often due to the sharing of limited amount of resources, leads the discrete population to extinction with probability 1. Nevertheless, the extinction time $\mathscr{T}_0$ can be very large compared to the typical lifetime of individuals, and for some species one can observe fluctuations of the population size for large amounts of time before extinction ([18,17]). To capture this phenomenon, we work with the notion of quasi-stationary measure, that is the class of probability measures that are invariant under the conditioning to non-extinction. This notion has been extensively studied since the pioneering work of Yaglom for the branching process in [23] and the classification of killed processes introduced by Vere-Jones in [22]. The description of quasi stationary distributions (q.s.d.) for finite state Markov chains was done in [6]. For countable Markov chains the infinitesimal description of q.s.d. on countable spaces was studied in [16] and [21] among others, and the more general existence result in the countable case was shown in [10]. For one-dimensional diffusions there is the pioneering work of Mandl [13] further developed in [4,14,19] and for bounded regions one can see [15] among others. For models of population dynamics and demography see [2,12] and [3].

Let us recall the definition of a quasi-stationary distribution.

Definition 1.1 A probability measure ν supported by the set of nonempty configurations $\mathcal{A}^{-0}$ is said to be a q.s.d. if

$$\forall B \in \mathcal{B}(\mathcal{A}^{-0}) : \mathbb{P}_\nu(Y_t \in B | \mathcal{T}_0 > t) = \nu(B),$$

where $\mathcal{B}(\mathcal{A}^{-0})$ is the class of Borel sets of $\mathcal{A}^{-0}$ and where as usual we put $\mathbb{P}_\nu = \int_{\mathcal{A}^{-0}} \mathbb{P}_\eta d\nu(\eta)$.

When starting from a q.s.d. ν, the absorption at the state 0 is exponentially distributed (for instance see [10]). Indeed, by the Markov property, the q.s.d. equality $\mathbb{P}_\nu(Y_t \in d\eta, \mathcal{T}_0 > t) = \nu(d\eta)\mathbb{P}_\nu(\mathcal{T}_0 > t)$ gives

$$\begin{aligned}\mathbb{P}_\nu(\mathcal{T}_0 > t + s) &= \int_{\mathcal{A}^{-0}} \mathbb{P}_\nu(Y_t \in d\eta, \mathcal{T}_0 > t + s) = \mathbb{P}_\nu(\mathcal{T}_0 > t) \int_{\mathcal{A}^{-0}} \nu(d\eta)\mathbb{P}_\eta(\mathcal{T}_0 > s)\\ &= \mathbb{P}_\nu(\mathcal{T}_0 > t)\mathbb{P}_\nu(\mathcal{T}_0 > s).\end{aligned}$$

Hence there exists $\theta(\nu) \geq 0$, the exponential decay rate (of absorption), such that

$$\forall t \geq 0 : \mathbb{P}_\nu(\mathcal{T}_0 > t) = e^{-\theta(\nu)t}. \tag{4}$$

In nontrivial situations as ours, $0 < \mathbb{P}_\nu(\mathcal{T}_0 > t) < 1$ (for $t > 0$), then $0 < \theta(\nu) < \infty$. We deduce that for all $\theta < \theta(\nu)$, $\mathbb{E}_\nu(e^{\theta \mathcal{T}_0}) < \infty$. So, for all $\theta < \theta(\nu)$, ν almost surely in η it holds: $\mathbb{E}_\eta(e^{\theta \mathcal{T}_0}) < \infty$. Then, a necessary condition for the existence of a q.s.d. is exponential absorption at 0, that is

$$\exists \eta \in \mathcal{A}^{-0}, \quad \exists \theta > 0, \quad \mathbb{E}_\eta(e^{\theta \mathcal{T}_0}) < \infty. \tag{5}$$

This condition is equivalent to

$$\exists \eta \in \mathcal{A}^{-0}, \quad \exists \theta' > 0, \quad \exists C > 0, \quad \mathbb{P}_\eta(\mathcal{T}_0 > t) \leq C e^{-\theta' t}, \qquad \forall t \geq 0.$$

1.2 The main results

Let us introduce the global quantities

$$\lambda_* = \liminf_{k \to \infty} \inf_{\substack{\eta \in \mathcal{A}^{-0} \\ \|\eta\| = k}} \inf_{y \in \{\eta\}} \lambda_y(\eta).$$

$$\Gamma^* = \limsup_{k \to \infty} \sup_{\substack{\eta \in \mathcal{A} \\ \|\eta\| = k}} \sup_{y \in \{\eta\}} (b_y(\eta) + m_y(\eta)).$$

Theorem 1.2 *Under the assumptions* (2), (3), *and*

$$\Gamma^* < \lambda_* \tag{6}$$

there exists a q.s.d. ν such that $\|Y_0\|$ has exponential moments, that is there exists $a > 0$ such that

$$\int e^{a\|\eta\|} \nu(d\eta) < \infty.$$

For all $k \geq 1$

$$\theta(\nu) = -\log\left(\frac{\int \|Y_1(\eta)\|^k \nu(d\eta)}{\int \|\eta\|^k \nu(d\eta)}\right). \tag{7}$$

Notice that under the hypothesis (6) and continuity of $\lambda_y(\eta)$ [see (1)] we obtain

$$\inf_{\eta \in \mathcal{A}^{-0}} \inf_{y \in \{\eta\}} \lambda_y(\eta) > 0. \tag{8}$$

Theorem 1.2 is proved in Sect. 4. This proof is based on an intermediate abstract theorem that shows existence of a finite eigenmeasures for some positive operators (Theorem 4.2). Note that our result contains the case where the total birth rate per individual is bounded and the death rate per individual is bounded away from 0 and grows polynomially for large populations.

In Sect. 5, we will introduce a natural σ-finite measure μ and show that absolute continuity with respect to μ is preserved by the process. We study the Lebesgue decomposition of a q.s.d. with respect to μ.

In Sect. 6 we will study the uniform case, which is given by

$$\forall y \in \mathbb{T}, \eta \in \mathcal{A}^{-0} : \lambda_y(\eta) = \lambda, b_y(\eta) = b(1-\rho), \quad m_y(\eta) = b\rho,$$

where λ, b and ρ are positive numbers with $\rho < 1$. The property (6) reads $\lambda > b$. In this case Theorem 1.2 ensures the existence of a q.s.d. with exponential decay rate $\lambda - b$. We will prove that this q.s.d. is the unique one with this decay rate, under the (irreducibility) condition

$$\sigma \otimes \sigma \left\{(y, z) \in \mathbb{T}^2 : g_{k_0}(y, z) = 0\right\} = 0, \tag{9}$$

for some fixed integer $k_0 \geq 1$. Here g_k denotes the iterated mutation kernel defined recursively by $g_1 = g$ and for $k \geq 1$

$$g_{k+1}(x, y) = \int_{\mathbb{T}} g(x, z) g_k(z, y) d\sigma(z).$$

Furthermore, given the weights of the configuration, the locations of the traits under this q.s.d. are absolutely continuous with respect to σ.

Theorem 1.3 *In the uniform case assume that* $\lambda > b$ *and* (9). *Then there is a unique q.s.d.* ν *on* $\mathcal{A}^{-0}$, *associated with the exponential decay rate* $\theta = \lambda - b$. *Moreover* ν *satisfies the absolute continuity property,*

$$\nu(\vec{\eta} \in \bullet | \overline{\eta}) \ll \sigma^{\otimes \# \eta}(\bullet).$$

In this statement, $\vec{\eta}$ denotes the ordered sequence of the elements of the support $\{\eta\}$, (the compact metric space $(\mathbb{T}, d)$ being ordered in a measurable way, see Sect. 2.1), and

$$\overline{\eta} = (\eta_y : y \in \{\eta\}) \tag{10}$$

is the associated sequence of strictly positive weights ordered accordingly.

Remark 1.4 It is an open problem to exhibit, in the general case, a q.s.d. with a non trivial singular part or to prove that this could not happen.

2 Poisson construction, martingale and Feller properties

Recall that (1) and (2) are assumed. We now give a pathwise construction of the process Y. As a preliminary result, we introduce an equivalent representation of finite point measures as finite sequences of ordered elements.

2.1 Representation of the finite point measures

Since $(\mathbb{T}, d)$ is a compact metric space there exists a countable basis of open sets $(\mathcal{U}_i : i \in \mathbb{N} = \{1, 2, \ldots\})$, that we fix once for all. The representation

$$\mathcal{R} : \mathbb{T} \to \{0, 1\}^{\mathbb{N}}, \quad z \to \mathcal{R}(z) = (c_i : i \in \mathbb{N}) \text{ with } c_i = \mathbf{1}(z \in \mathcal{U}_i)$$

is an injective measurable mapping, where the set $\{0, 1\}^{\mathbb{N}}$ is endowed with the product $\sigma-$field. On $\{0, 1\}^{\mathbb{N}}$ we consider the lexicographical order $\leq_l$ which induces the following order on $\mathbb{T} : z \preceq z' \Leftrightarrow \mathcal{R}(z) \leq_l \mathcal{R}(z')$. This order relation is measurable.

The support $\{\eta\}$ of a configuration η can be ordered by $\preceq$ and represented by the tuple $\vec{\eta} = (y_1, \ldots, y_{\#\eta})$ and the associated discrete structure is $\overline{\eta} = (\overline{\eta}(k) := \eta_{y_k} : k \in \{1, \ldots, \#\eta\})$. Let us define $S_0(\eta) = 0$ and $S_k(\eta) = \sum_{l=1}^{k} \eta_{y_l}$ for $k \in \{1, \ldots, \#\eta\}$. Observe that $S_{\#\eta}(\eta) = \|\eta\|$. It is convenient to add an extra topologically isolated point ∂ to $\mathbb{T}$. Now we can introduce the functions $H^i : \mathcal{A} \mapsto \mathbb{T} \cup \{\partial\}$ given by $H^0(\eta) = \partial$ for all $\eta \in \mathcal{A}$ and for $i \geq 1$

$$H^i(\eta) = \begin{cases} y_k & \text{if } i \in (S_{k-1}(\eta), S_k(\eta)] \text{ for } k \leq \#\eta \\ \partial & \text{otherwise} \end{cases}.$$

The functions H^i are measurable. We extend the functions b, λ and m to ∂ by putting $b_\partial(\eta) = \lambda_\partial(\eta) = m_\partial(\eta) = 0$ for all $\eta \in \mathcal{A}$. The functions H^i will be used to enumerate in a measurable way all the individuals in the population.

2.2 Pathwise Poisson construction

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space in which there are defined two independent Poisson point measures:

- $M_1(ds, di, dz, d\theta)$ is a Poisson point measure on $[0, \infty) \times \mathbb{N} \times \mathbb{T} \times \mathbb{R}^+$, with intensity measure $ds(\sum_{k\geq 1} \delta_k(di))d\sigma(z)d\theta$ (the birth Poisson measure).
- $M_2(ds, di, d\theta)$ is a Poisson point measure on $[0, \infty) \times \mathbb{N} \times \mathbb{R}^+$, with the intensity measure $ds(\sum_{k\geq 1} \delta_k(di))d\theta$ (the death Poisson measure).

We denote $(\mathcal{F}_t : t \geq 0)$ the canonical filtration generated by these processes.

We define the process $(Y_t : t \geq 0)$ as a $(\mathcal{F}_t : t \geq 0)$-adapted stochastic process such that almost surely and for all $t \geq 0$,

$$
\begin{aligned}
Y_t = Y_0 + \int\limits_{[0,t]\times\mathbb{N}\times\mathbb{T}\times\mathbb{R}^+} & \mathbf{1}_{\{i\leq\|Y_{s-}\|\}} \Big\{ \delta_{H^i(Y_{s-})} \mathbf{1}_{\left\{\theta\leq b_{H^i(Y_{s-})}(Y_{s-})\right\}} \\
& + \delta_z \, \mathbf{1}_{\left\{b_{H^i(Y_{s-})}(Y_{s-})\leq\theta\leq b_{H^i(Y_{s-})}(Y_{s-})+m_{H^i(Y_{s-})}(Y_{s-})g(H^i(Y_{s-}),z)\right\}} \Big\} \\
& \times M_1(ds, di, dz, d\theta) \\
& - \int\limits_{[0,t]\times\mathbb{N}\times\mathbb{R}^+} \delta_{H^i(Y_{s-})} \, \mathbf{1}_{\{i\leq\|Y_{s-}\|\}} \mathbf{1}_{\left\{\theta\leq\lambda_{H^i(Y_{s-})}(Y_{s-})\right\}} M_2(ds, di, d\theta). \qquad (11)
\end{aligned}
$$

The existence of such process is proved in [11], as well as its uniqueness in law. Its jump rates are those given in Sect. 1.1 so this process has the same law as the process introduced in Sect. 1.1. In particular, the law of Y does not depend on the choice of the functions H^i neither on the order defined in Sect. 2.1.

Observe that the process $\|Y\|$ is dominated everywhere by the birth process Z solution of

$$
Z_t = \|Y_0\| + \int\limits_{[0,t]\times\mathbb{N}\times\mathbb{T}\times\mathbb{R}^+} \mathbf{1}_{\{i\leq\|Z_{s-}\|\}} \mathbf{1}_{\{\theta\leq C^*\}} M_1(ds, di, dz, d\theta), \qquad (12)
$$

where

$$
C^* = \sup_{\eta\in\mathcal{A}} \sup_{y\in\{\eta\}} (b_y(\eta) + m_y(\eta)). \qquad (13)
$$

This means that almost surely $\|Y_t\| \leq Z_t$.

We will use below the standard notation,

$$
\forall K \in \mathbb{N} : T_K^W = \inf\{t \geq 0 : W \geq K\}, \qquad (14)
$$

for a real process W. We shall use the notation T_K^Y instead of $T_K^{\|Y\|}$. The following results are consequences of the above domination.

Proposition 2.1 *For any $\eta \in \mathcal{A}$, any $p \geq 1$, there exist two positive constants c_p and b_p such that for all $t_0 > 0$*

$$\mathbb{E}_\eta\big(\sup_{t\in[0,t_0]} \|Y_t\|^p\big) \leq c_p\, e^{b_p t_0} < \infty.$$

We refer to [11] for the proof.

Lemma 2.2 *For any $t \geq 0$ and any $\eta \in \mathcal{A}^{-0}$, there is a number $c = c(t, \|\eta\|) \in (0, 1)$ such that,*

$$\forall K > \|\eta\| : \mathbb{P}_\eta\left(T_K^Y \leq t\right) \leq \mathbb{P}_{\|\eta\|}\left(T_K^Z \leq t\right) \leq c^{-1} e^{-cK}. \tag{15}$$

Proof Since $\|\eta\| < K$ then $T_K^Y \geq T_K^Z$ almost surely. Therefore, for any $t \geq 0$, and any $\eta \in \mathcal{A}^{-0}$

$$\mathbb{P}_\eta\left(T_K^Y \leq t\right) \leq \mathbb{P}_{\|\eta\|}\left(T_K^Z \leq t\right).$$

For a pure birth process (see for example [9]) we have

$$\mathbb{P}_{\|\eta\|}\left(T_K^Z \leq t\right) = \mathbb{P}_{\|\eta\|}\left(Z_t \geq K\right) = \sum_{m=K}^{\infty} \binom{m-1}{\|\eta\|-1} e^{-C^*\|\eta\| t}\left(1 - e^{-C^* t}\right)^{m-\|\eta\|},$$

where C^* was defined in (13). The result follows at once from this estimate. □

Let us introduce the following notation for any integer $m \geq 1$

$$\overline{b}(m) = \sup_{0<\|\eta\|\leq m} \frac{1}{m} \sum_{y\in\{\eta\}} \eta_y (b_y(\eta) + m_y(\eta))$$

and

$$\underline{\lambda}(m) = \inf_{\|\eta\|=m} \inf_{y\in\{\eta\}} \lambda_y(\eta).$$

Note that from (2), continuity of the rates, and compactness of $\mathbb{T}$, we have $\sup_{m\geq 1} \overline{b}(m) < \infty$. Moreover, if (8) is satisfied, we have $\inf_m \underline{\lambda}(m) > 0$. More precisely we have the following result.

Lemma 2.3 *Under the hypothesis of Theorem 1.2, we have*

$$\limsup_{m\to\infty} \overline{b}(m) \leq \Gamma^* < \lambda_* = \liminf_{m\to\infty} \underline{\lambda}(m). \tag{16}$$

Proof By definition of λ_* and (6) we need only to prove the first inequality. By definition of Γ^*, for any $\epsilon > 0$, there exists an integer $K(\epsilon)$ such that for any $k > K(\epsilon)$ we have

$$\sup_{\|\eta\|=k} \sup_{y\in\{\eta\}} (b_y(\eta) + m_y(\eta)) \le \Gamma^* + \epsilon.$$

Therefore for $m \ge K(\epsilon)$

$$\begin{aligned}\overline{b}(m) \le & \sup_{0<\|\eta\|\le K(\epsilon)} \frac{1}{m} \sum_{y\in\{\eta\}} \eta_y (b_y(\eta) + m_y(\eta)) \\ & + \sup_{K(\epsilon)\le\|\eta\|\le m} \frac{1}{m} \sum_{y\in\{\eta\}} \eta_y (b_y(\eta) + m_y(\eta)).\end{aligned}$$

This implies

$$\limsup_{m\to\infty} \overline{b}(m) \le \Gamma^* + \epsilon.$$

Since this holds for any $\epsilon > 0$ the result follows. □

In the next result we shall need the use of $X = (X_t : t \ge 0)$, a birth and death process on $\mathbb{Z}_+ = \{0, 1, \ldots\}$ absorbed at 0, with individual birth rates $(\overline{b}(m) : m \ge 1)$ and individual death rates $(\underline{\lambda}(m) : m \ge 1)$. This means $\mathbb{P}(X_{t+h} = m + 1 | X_t = m) = m\overline{b}(m)h + o(h)$ and $\mathbb{P}(X_{t+h} = m - 1 | X_t = m) = m\underline{\lambda}(m)h + o(h)$, for $m \ge 1$. We will prove that this birth and death chain X is exponentially absorbed. On the other hand it is easy to see that irreducibility implies that the condition of exponential absorption of a birth and death process does not depend on the initial state, and that the exponential absorption is uniform, namely

$$\exists\theta > 0, \quad \mathbb{E}_n\left(e^{\theta \mathcal{T}_0^X}\right) < \infty, \quad \forall n > 0,$$

where $\mathcal{T}_0^X = \inf\{t > 0 : X_t = 0\}$.

Lemma 2.4 *Under the hypothesis of Theorem 1.2 the process $\|Y\|$ is exponentially absorbed. Moreover, we have uniform absorption in the sense that*

$$\exists\theta > 0, \quad \mathbb{E}_\eta\left(e^{\theta \mathcal{T}_0}\right) < \infty, \qquad \forall \eta \in \mathcal{A}^{-0}.$$

Proof We will show that $\|Y\|$ is dominated by X, and X is exponentially absorbed.

We introduce a coupling on the subset $\mathcal{J}$ of $\mathcal{A} \times \mathbb{N}$ defined by

$$\mathcal{J} = \{(\eta, m) \in \mathcal{A} \times \mathbb{N} : \|\eta\| \le m\}.$$

The coupled process is defined by the nonzero rates

$$
\begin{aligned}
J(\eta,m;\eta+\delta_y,m+1) &= \eta_y b_y(\eta), \quad y\in\{\eta\},\\
J(\eta,m;\eta+\delta_z,m+1)d\sigma(z) &= G(\eta,z)d\sigma(z), \quad z\notin\{\eta\},\\
J(\eta,m;\eta,m+1) &= m\overline{b}(m) - \sum_{y\in\{\eta\}} \eta_y(b_y(\eta)+m_y(\eta)),\\
J(\eta,m;\eta-\delta_y,m-1) &= \underline{\lambda}(m)\eta_y \mathbf{1}_{\|\eta\|=m}, \quad y\in\{\eta\},\\
J(\eta,m;\eta-\delta_y,m) &= \eta_y\big(\lambda_y(\eta)-\underline{\lambda}(m)\mathbf{1}_{\|\eta\|=m}\big), \quad y\in\{\eta\},\\
J(\eta,m;\eta,m-1) &= m\underline{\lambda}(m)\mathbf{1}_{\|\eta\|<m},\\
J(0,m;0,m+1) &= m\overline{b}(m),\\
J(0,m;0,m-1) &= m\underline{\lambda}(m).
\end{aligned}
$$

It is immediate to check that the marginals of this process have, respectively, the laws of Y and X. On the other hand when the coupled process starts from $\mathscr{J}$ it remains in $\mathscr{J}$ forever, so the domination follows.

From the coupling the lemma will be established as soon we prove X is exponentially absorbed.

In the sequel Exp$[q]$ denotes the distribution of an exponential random variable with mean $1/q$. Also we put $\overline{q}_i = \overline{b}_i + \underline{\lambda}_i$ for all $i \geq 1$.

We claim that starting from any k the time of first visit to $k-1$ has exponential moment, that is there exists $\theta_k > 0$ such that $\mathbb{E}_k(e^{\theta \mathscr{T}^X_{k-1}}) < \infty$ for all $\theta \leq \theta_k$, where $\mathscr{T}^X_{k-1} = \inf\{t>0 : X_t = k-1\}$. This proves that $\mathscr{T}^X_0$ has exponential moment, starting from any k because $\mathbb{E}_k(e^{\theta \mathscr{T}^X_0}) = \prod_{j=k}^{1} \mathbb{E}_j(e^{\theta \mathscr{T}^X_{j-1}})$ which is finite for all $0<\theta\leq \min\{\theta_j : j=1,\ldots,k\}$.

The condition (16), namely $\limsup\limits_{m\to\infty} \overline{b}(m) < \liminf\limits_{m\to\infty} \underline{\lambda}(m)$, implies there exists i_0 such that for $i \geq i_0$

$$0 < \overline{b}^*_i < \underline{\lambda}_{*i} \quad \text{where } \overline{b}^*_i = \sup_{l\geq i} \overline{b}_l, \quad \underline{\lambda}_{*i} = \inf_{l\geq i} \underline{\lambda}_l.$$

From the results in [21], there exists $\theta_i \geq \underline{\lambda}_{*i} - \overline{b}^*_i > 0$ such that

$$\mathbb{E}_{i+1}(e^{\theta_i \mathscr{T}^X_i}) < \infty.$$

Therefore the claim is proved for $i \geq i_0$. The desired property will be proved by induction on $k = i_0, i_0-1, \ldots, 1$.

Assume that for $k \leq i_0$ there exists θ_k such that $\mathbb{E}_{k+1}(e^{\theta_k \mathscr{T}^X_k}) < \infty$. We shall prove, from this fact, the existence of $\theta_{k-1} > 0$ such that $\mathbb{E}_k(e^{\theta_{k-1}\mathscr{T}^X_{k-1}}) < \infty$. This will be done by studying the sequence of times of successive visits to k, starting from k, before to visit $k-1$. Let $(E_m : m \geq 1)$ be a sequence of i.i.d. random variables with common distribution Exp$[\overline{q}_k]$. These random variables represent the time spent at k in each successive visit.

On the other hand consider $(L_m : m \geq 1)$ be a sequence of i.i.d. random variables, where the common distribution is the one of the hitting time at k, under $\mathbb{P}_{k+1}$. We also take them independent from (E_m). The last ingredient we need is M, a random variable geometrically distributed with parameter $0 < \epsilon = \frac{\overline{b}_k}{\overline{q}_k} < 1$, that is, $\mathbb{P}(M = m) = (1-\epsilon)\epsilon^m$ for all $m \in \mathbb{Z}_+$. M represents the number of excursions from k to k, previous to visit $k-1$ for the first time. We also take M independent from $(E_m)_m, (L_m)_m$.

It is straightforward to show that under $\mathbb{P}_k$

$$\mathscr{T}^X_{k-1} \sim \sum_{m=1}^{M} E_m + \sum_{m=1}^{M-1} L_m \leq \sum_{m=1}^{M} (E_m + L_m).$$

Hence we obtain for any $\theta > 0$

$$\mathbb{E}_k\left(e^{\theta \mathscr{T}^X_{k-1}}\right) \leq \sum_{m=1}^{\infty} \left(\mathbb{E}\left(e^{\theta(E_1+L_1)}\right)\right)^m (1-\epsilon)\epsilon^m.$$

We notice that for $0 \leq \theta < \theta_k \wedge \overline{q}_k$ one has $\mathbb{E}(e^{\theta(E_1+L_1)}) < \infty$. From the monotone convergence Theorem we have

$$\lim_{\theta \downarrow 0} \mathbb{E}(e^{\theta(E_1+L_1)}) = 1.$$

Therefore, there exists $\theta_{k-1} > 0$ such that for all $0 \leq \theta \leq \theta_{k-1}$

$$\mathbb{E}(e^{\theta(E_1+L_1)}) < \frac{1}{\epsilon}$$

and then $\mathbb{E}_k(e^{\theta_{k-1}\mathscr{T}^X_{k-1}}) < \infty$. As previously mentioned the result follows by induction. □

2.3 Martingale properties

The process Y is Markovian and we describe its infinitesimal generator L, in a weak form, using related martingales. Given $f : \mathcal{A} \to \mathbb{R}$, a measurable and locally bounded function with $f(0) = 0$, we define Lf as

$$\begin{aligned} Lf(\eta) = &\sum_{y \in \{\eta\}} \eta_y b_y(\eta)(f(\eta + \delta_y) - f(\eta)) \\ &+ \sum_{y \in \{\eta\}} \eta_y m_y(\eta) \int_{\mathbb{T}} (f(\eta + \delta_z) - f(\eta)) g(y, z) d\sigma(z) \\ &+ \sum_{y \in \{\eta\}} \eta_y \lambda_y(\eta)(f(\eta - \delta_y) - f(\eta)). \end{aligned} \tag{17}$$

Proposition 2.5 *Assume (2) and (3). Let $f : \mathbb{R}_+ \times \mathcal{A} \to \mathbb{R}$ be a measurable function such that for any $\xi \in \mathcal{A}$ the marginal function $f(\cdot, \xi)$ is continuously differentiable. We assume $f(\cdot, 0) = 0$ and we take $Y_0 = \eta$.*

(i) *If for any $t_0 \geq 0$ we have for some finite $C(t_0)$ and any η*

$$\sup_{0 \leq t \leq t_0} |f(t, \eta)| + |\partial_t f(t, \eta)| \leq C(t_0)(1 + \|\eta\|^{p-1}),$$

then

$$\mathscr{M}_t^f =: f(t, Y_t) - f(0, \eta) - \int_0^t (\partial_s f(s, Y_s) + Lf(Y_s))ds$$

is a càdlàg $(\mathcal{F}_t : t \geq 0)$-martingale.

(ii) *If the functions $f, \partial_s f$ are assumed to be continuous, or more generally locally bounded, then $\mathscr{M}^f$ is a local martingale and the process $(\mathscr{M}^f_{T_N^Y \wedge t} : t \geq 0)$ is a martingale.*

Proof Let us first prove (i) for f with compact support. Using (11) we get for all $t \geq 0$

$$f(t, Y_t) - f(0, \eta) - \int_0^t \partial_s f(s, Y_s)ds = \sum_{s \leq t} (f(s, Y_{s-} + (Y_s - Y_{s-})) - f(s, Y_{s-}))$$

$$= \int_{[0,t]\times\mathbb{N}\times\mathbb{T}\times\mathbb{R}^+} \mathbf{1}_{\{i \leq \|Y_{s-}\|\}} \Big\{ \Big(f\big(s, Y_{s-} + \delta_{H^i(Y_{s-})}\big)$$
$$- f(s, Y_{s-})\Big) \mathbf{1}_{\left\{\theta \leq b_{H^i(Y_{s-})}(Y_{s-}) g(H^i(Y_{s-}), z)\right\}}$$
$$+ (f(s, Y_{s-} + \delta_z) - f(s, Y_{s-}))\, \mathbf{1}_{\left\{\theta \leq m_{H^i(Y_{s-})}(Y_{s-}) g(H^i(Y_{s-}), z)\right\}}\Big\} M_1(ds, di, dz, d\theta)$$
$$+ \int_{[0,t]\times\mathbb{N}\times\mathbb{R}^+} \Big(f(s, Y_{s-} - \delta_{H^i(Y_{s-})}) - f(s, Y_{s-})\Big)\, \mathbf{1}_{\left\{i \leq \|Y_{s-}\|, \theta \leq \lambda_{H^i(Y_{s-})}(Y_{s-})\right\}}$$
$$\times M_2(ds, di, d\theta),$$

where both integrals belong to $L^1(\mathbb{P}_\eta)$. Compensating each Poisson measure, using Fubini's Theorem, and the fact that $\int_{\mathbb{T}} g(y, z)d\sigma(z) = 1$, we obtain

$$f(t, Y_t) - f(0, \eta) - \int_0^t (\partial_s f(s, Y_s) + Lf(Y_s))ds$$

is a martingale. The rest of the Proposition is proved by localization arguments, justified by Proposition 2.1. ◻

2.4 Feller property of the semi-group

Let us start with a general smoothness result needed later on.

Lemma 2.6 *Let $F : \mathcal{A}\times\mathbb{T}\times\mathbb{T} \to \mathbb{R}$ be a continuous function on $\mathbb{T}$. Then the function $\hat{F}$ defined on $\mathcal{A}\times\mathbb{T}$ by*

$$\hat{F}(\eta, z) = \sum_{y\in\{\eta\}} F(\eta, y, z)\eta_y,$$

is continuous.

Proof Let $\eta, \tilde{\eta} \in \mathcal{A}$ and $z, z' \in \mathbb{T}$. Thus

$$\begin{aligned}|\hat{F}(\eta, z) - \hat{F}(\tilde{\eta}, z')| \le &\sum_{y\in\{\eta\}} |F(\eta, y, z) - F(\eta, y, z')|\eta_y \\ &+ \sum_{y\in\{\eta\}} |F(\eta, y, z') - F(\tilde{\eta}, y, z')|\eta_y \\ &+ \left|\sum_{y\in\{\eta\}\cup\{\tilde{\eta}\}} (\eta_y - \tilde{\eta}_y)F(\tilde{\eta}, y, z')\right|.\end{aligned}$$

Since $\mathbb{T}$ is a compact set, it is immediate that the two first terms are small if z is close to z' and η close to $\tilde{\eta}$. If $\tilde{\eta}$ is in a small enough neighborhood of η, these two atomic measures have the same weights, and the corresponding traits are close. In particular, $\tilde{\eta}$ belongs to a compact set, and the smallness of the last term follows by the equicontinuity and boundedness of F on compact sets. □

Let $\mathcal{N} = (\mathcal{N}_t : t \ge 0)$, where $\mathcal{N}_t$ is the number of jumps of the process Y up to time t. We shall prove by induction the following result.

Lemma 2.7 *Assume that $f : \mathbb{R}_+ \times \mathcal{A} \to \mathbb{R}$ is a bounded continuous function. Then for all $m \ge 0$*

$$(t, \eta) \to \mathbb{E}_\eta(f(t, Y_t), \mathcal{N}_t = m)$$

is a continuous function.

Proof Denote by $j = \|\eta\|$ and $n = m + j + 1$. Observe that because of (1)

$$\forall k \ge 1,\ Q_+(k) = \sup\{Q(\eta) : \|\eta\| \le k\} < \infty. \tag{18}$$

We first prove the continuity on time. For this purpose, we assume that $0 \le u \le t \le t_0$ where t, u are close and t_0 is fixed. From

$$f(t, Y_t) = f(t, Y_u)\mathbf{1}_{\mathcal{N}_t=\mathcal{N}_u} + f(t, Y_t)\mathbf{1}_{\mathcal{N}_t\neq\mathcal{N}_u}$$

we find [recall the notation (18)],

$$\begin{aligned}&|\mathbb{E}_\eta(f(t,Y_t),\mathcal{N}_t=m)-\mathbb{E}_\eta(f(u,Y_u),\mathcal{N}_u=m)|\\&\quad\le\sup_{\|\xi\|\le j+m}|f(t,\xi)-f(u,\xi)|+\|f\|Q_+(m+j)((t-u)+o(t-u)),\end{aligned}$$

where $\|f\|$ is the uniform norm of f. Since the set $\{\xi : \|\xi\| \le j+m\}$ is compact, it follows from the uniform continuity of f on compact sets that the first term on the r.h.s. is small if $|t-u|$ is small. Hence the result follows.

So in what follows we consider that $t=u$ and we prove continuity on η. We will do it by induction on m. In the case $m=0$, we have $\mathbb{E}_\eta(f(t,Y_t),\mathcal{N}_t=0)=f(t,\eta)e^{-Q(\eta)t}$ which is clearly continuous on η. Now we prove the induction step, so we assume that the statement holds for m and all continuous functions f. We have

$$\begin{aligned}\mathbb{E}_\eta(f(t,Y_t),\mathcal{N}_t=m+1)=\int_0^t&(A_1(\eta,m,t-s)+A_2(\eta,m,t-s)\\&+A_3(\eta,m,t-s))e^{-Q(\eta)s}ds,\end{aligned}$$

where

$$\begin{aligned}A_1(\eta,m,t-s)&=\sum_{y\in\eta}\eta_y b_y(\eta)\mathbb{E}_{\eta+\delta_y}(f(t-s,Y_{t-s}),\mathcal{N}_{t-s}=m)\\A_2(\eta,m,t-s)&=\sum_{y\in\eta}\eta_y\lambda_y(\eta)\mathbb{E}_{\eta-\delta_y}(f(t-s,Y_{t-s}),\mathcal{N}_{t-s}=m)\\A_3(\eta,m,t-s)&=\int_{\mathbb{T}}\mathbb{E}_{\eta+\delta_z}(f(t-s,Y_{t-s}),\mathcal{N}_{t-s}=m)G(\eta,z)\sigma(dz).\end{aligned}$$

From the inductive assumption and Lemma 2.6, it is immediate that the functions A_1, A_2 and A_3 are continuous in (t,η). We conclude by the dominated convergence Theorem since f is bounded. □

Proposition 2.8 *Let $f:\mathbb{R}_+\times\mathcal{A}\to\mathbb{R}$ be a bounded continuous function. Then*

$$(t,\eta)\to\mathbb{E}_\eta(f(t,Y_t))$$

is a continuous bounded function. In other words, the process (Y_t) has the Feller property.

Proof It is easy to verify that if the process starting from η has $M>\|\eta\|$ jumps up to time $t>0$ then there are at least $(M-\|\eta\|)/2$ birth or mutation events up to time t. Hence, $Z_t\ge(M+\|\eta\|)/2$, where Z is the pure birth process defined in (12), starting from $Z_0=\|\eta\|$. Using (15), we obtain that for each $\eta\in\mathcal{A}$, $t>0$, there exists

$1 > a = a(t, \|\eta\|) > 0$ such that for any positive integer $M > \|\eta\|$,

$$\mathbb{P}_\eta \left(T_M^{\mathcal{N}} \le t \right) = \mathbb{P}_\eta (\mathcal{N}_t \ge M) \le \mathbb{P}_{\|\eta\|} (Z_t \ge (M + \|\eta\|)/2)$$
$$= \mathbb{P}_{\|\eta\|} \left(T^Z_{(M+\|\eta\|)/2} \le t \right) \le a^{-1} e^{-aM}.$$

Assume that η' is close to η (such that $\|\eta\| = \|\eta'\|$) and consider u, t close and smaller than t_0 fixed. Then for $K > \|\eta\|$ we have

$$|\mathbb{E}_\eta(f(t, Y_t)) - \mathbb{E}_{\eta'}(f(u, Y_u))| \le 2\|f\|a^{-1}e^{-aK} + \sum_{m=0}^{K} |\mathbb{E}_\eta(f(t, Y_t), \mathcal{N}_t = m) - \mathbb{E}_{\eta'}(f(u, Y_u), \mathcal{N}_u = m)|.$$

The result follows by taking a large K and using Lemma 2.7. □

3 Quasi-stationary distributions

3.1 The process killed at 0

Let us recall that the state 0 is absorbing for the population process Y. We have moreover assumed in (4) that the population goes almost surely to extinction, that is $\mathbb{P}(\mathcal{T}_0 < \infty) = 1$. This is in particular true if $\lambda_* > \Gamma^*$. Our aim is the study of existence and possibly uniqueness of a q.s.d. ν, which is a probability measure on $\mathcal{A}^{-0}$ satisfying $\mathbb{P}_\nu(Y_t \in \bullet | \mathcal{T}_0 > t) = \nu(\bullet)$. Let us now give some preliminary results for q.s.d.

Since by condition (6), the process Y is almost surely but not immediately absorbed then its exponential decay rate satisfies $0 < \theta(\nu) < \infty$. Since 0 is absorbing it holds $\mathbb{P}_\nu(Y_t \in B) = \mathbb{P}_\nu(Y_t \in B, \mathcal{T}_0 > t)$ for $B \in \mathcal{B}(\mathcal{A}^{-0})$. So, the q.s.d. equation can be written as,

$$\forall B \in \mathcal{B}(\mathcal{A}^{-0}), \quad \nu(B) = e^{\theta(\nu)t} \mathbb{P}_\nu(Y_t \in B). \tag{19}$$

Recall that a necessary condition for the existence of a q.s.d. is exponential absorption at 0 [see (5)].

Let $(P_t : t \ge 0)$ be the semi-group of the process before killing at 0, acting on the set $\mathcal{C}_b(\mathcal{A}^{-0})$ of real continuous bounded functions defined on $\mathcal{A}^{-0}$:

$$\forall \eta \in \mathcal{A}^{-0}, \forall f \in \mathcal{C}_b(\mathcal{A}^{-0}) : (P_t f)(\eta) = \mathbb{E}_\eta(f(Y_t), \mathcal{T}_0 > t).$$

Let us observe that for any continuous and bounded function $h : \mathcal{A} \to \mathbb{R}$ and for any $\eta \in \mathcal{A}^{-0}$, we have

$$\mathbb{E}_\eta(h(Y_t)) = \mathbb{E}_\eta(h(Y_t), \mathcal{T}_0 > t) + h(0)\mathbb{P}_\eta(\mathcal{T}_0 \le t).$$

In particular, if $h(0) = 0$, we get $\mathbb{E}_\eta(h(Y_t)) = \mathbb{E}_\eta(h(Y_t), \mathscr{T}_0 > t)$.

We denote by $P_t^\dagger$ the adjoint semi-group acting on the space of positive measures $\mathcal{M}(\mathcal{A}^{-0})$. It is defined trough the formula

$$P_t^\dagger v(f) = v(P_t f),$$

for any $v \in \mathcal{M}(\mathcal{A}^{-0})$ and any positive measurable function f (or equivalently for all $f \in \mathcal{C}_b(\mathcal{A}^{-0})$).

From relation (19) we get that a probability measure ν is a q.s.d. if and only if there exists $\theta > 0$ such that for all $t \geq 0$

$$\nu(P_t f) = e^{-\theta t} \nu(f),$$

holds for all positive measurable function f. Then ν is a q.s.d. with exponential decay rate θ if and only if it verifies

$$\forall t \geq 0 : P_t^\dagger \nu = e^{-\theta t} \nu. \tag{20}$$

3.2 Some properties of q.s.d

Let us show that the existence of a q.s.d. will be proved if for a fixed strictly positive time, the eigenmeasure equation (20) is satisfied.

Lemma 3.1 *Assume there exists $\tilde{\nu} \in \mathcal{P}(\mathcal{A}^{-0})$ and $\beta > 0$ such that $P_1^\dagger \tilde{\nu} = \beta \tilde{\nu}$. Then $\beta < 1$ and there exists ν a q.s.d. with exponential decay rate $\theta := -\log \beta > 0$.*

Proof First note that for all $\eta \in \mathcal{A}^{-0}$ we have $\mathbb{P}_\eta(\mathscr{T}_0 > 1) < 1$. Hence $\beta = P_1^\dagger \tilde{\nu}(\mathcal{A}^{-0}) = \mathbb{P}_{\tilde{\nu}}(\mathscr{T}_0 > 1) < 1$, so $\theta := -\log \beta > 0$. We must show the existence of $\nu \in \mathcal{P}(\mathcal{A}^{-0})$ such that $P_t^\dagger \nu = e^{-\theta t} \nu$ for all $t \geq 0$. Consider,

$$\nu = \int_0^1 e^{\theta s} P_s^\dagger \tilde{\nu} ds.$$

For $t \in (0, 1)$ we have

$$\begin{aligned}
P_t^\dagger \nu &= \int_0^1 e^{\theta s} P_{t+s}^\dagger \tilde{\nu} ds = \int_0^{1-t} e^{\theta s} P_{t+s}^\dagger \tilde{\nu} ds + \int_{1-t}^1 e^{\theta s} P_{t+s}^\dagger \tilde{\nu} ds \\
&= \int_t^1 e^{\theta(u-t)} P_u^\dagger \tilde{\nu} du + \int_1^{1+t} e^{\theta(u-t)} P_u^\dagger \tilde{\nu} du \\
&= e^{-\theta t} \int_t^1 e^{\theta u} P_u^\dagger \tilde{\nu} du + e^{-\theta t} \int_0^t e^{\theta u} e^{\theta} P_u^\dagger P_1^\dagger \tilde{\nu} du = e^{-\theta t} \nu.
\end{aligned}$$

For $t \geq 1$ we write $t = n + r$ with $0 \leq r < 1$ and $n \in \mathbb{N}$. We have

$$P_t^\dagger \nu = P_r^\dagger P_n^\dagger \nu = \beta^n P_r^\dagger \nu = e^{-n\theta} e^{-r\theta} \nu = e^{-\theta t} \nu.$$

□

Note that

$$\theta(\nu) = \lim_{t \to 0^+} \frac{1}{t}(1 - \mathbb{P}_\nu(\mathscr{T}_0 > t)) = \lim_{t \to 0^+} \frac{\mathbb{P}_\nu(\mathscr{T}_0 \leq t)}{t}.$$

In the next result we give an explicit expression for the exponential decay rate associated to a q.s.d. We will use the identification between the trait $y \in \mathbb{T}$ and δ_y the singleton configuration that gives unit weight to y.

Lemma 3.2 *If $\nu \in \mathcal{P}(\mathcal{A}^{-0})$ is a q.s.d. then its exponential decay rate $\theta(\nu)$ satisfies*

$$\theta(\nu) = \int_{\|\eta\|=1} Q(\eta, 0) \nu(d\eta),$$

where for $\eta = \delta_y$ we have $Q(\eta, 0) = \lambda_y(\eta)$.

Proof We will denote by $\tau = \inf\{t > 0 : Y_t \neq Y_0\}$ the time of the first jump of the process Y. For $i = 1, 2$ we introduce

$$a_i(t) = \sup\{\mathbb{P}_\xi(\mathscr{T}_0 \leq t) : \|\xi\| = i\}.$$

Consider η such that $\|\eta\| \geq 3$, then the strong Markov property implies

$$\mathbb{P}_\eta(\mathscr{T}_0 \leq t) = \mathbb{E}_\eta(T_2 < t, \mathbb{E}_{Y_{T_2}}(\mathscr{T}_0 \leq t - T_2)) \leq \mathbb{E}_\eta(T_2 < t, \mathbb{E}_{Y_{T_2}}(\mathscr{T}_0 \leq t)) \leq a_2(t).$$

When $\|\eta\| = 2$, conditioning on the first jump and using the previous estimate we obtain

$$\begin{aligned}
\mathbb{P}_\eta(\mathscr{T}_0 \leq t) &= \mathbb{E}_\eta(\tau < t, \mathbb{E}_{Y_\tau}(\mathscr{T}_0 \leq t - \tau)) \\
&= \mathbb{E}_\eta(\tau < t, \|Y_\tau\| = 1, \mathbb{E}_{Y_\tau}(\mathscr{T}_0 \leq t - \tau)) \\
&\quad + \mathbb{E}_\eta(\tau < t, \|Y_\tau\| = 3, \mathbb{E}_{Y_\tau}(\mathscr{T}_0 \leq t - \tau)) \\
&\leq \mathbb{P}_\eta(\tau < t)(a_1(t) + a_2(t)) = (1 - e^{-Q(\eta)t})(a_1(t) + a_2(t)) \\
&\leq (1 - e^{-C_2 t})(a_1(t) + a_2(t)),
\end{aligned}$$

where $C_2 = \max\{Q(\xi) : \|\xi\| = 2\} < \infty$. Hence we deduce that

$$a_2(t) \leq \left(e^{C_2 t} - 1\right) a_1(t).$$

On the other hand, if $\|\eta\| = 1$ we have

$$\mathbb{P}_\eta(\mathcal{T}_0 \le t) \le \mathbb{P}_\eta(\tau \le t) = 1 - e^{-Q(\eta)t} \le 1 - e^{-C_1 t},$$

where $C_1 = \max\{Q(\xi) : \|\xi\| = 1\} < \infty$. In particular we deduce that $a_1(t) = O(t)$. Collecting all these facts we deduce

$$a_2(t) = O(t^2).$$

Now, if $\|\eta\| = 1$ we get

$$\begin{aligned}\mathbb{P}_\eta(\tau \le t)\frac{Q(\eta,0)}{Q(\eta)} &\le \mathbb{P}_\eta(\mathcal{T}_0 \le t)\\ &= \mathbb{P}_\eta(\tau = \mathcal{T}_0 \le t) + \mathbb{E}_\eta(\tau < t, \|Y_\tau\| = 2, \mathbb{E}_{Y_\tau}(\mathcal{T}_0 \le t - \tau))\\ &\le \mathbb{P}_\eta(\tau \le t)\frac{Q(\eta,0)}{Q(\eta)} + a_2(t).\end{aligned}$$

Given that $\mathbb{P}_\eta(\tau \le t) = 1 - e^{-Q(\eta)t}$ we deduce

$$\lim_{t\to 0}\frac{\mathbb{P}_\eta(\mathcal{T}_0 \le t)}{t} = \begin{cases} Q(\eta,0) & \text{if } \|\eta\| = 1\\ 0 & \text{otherwise}\end{cases},$$

and the uniform bound: for all $\eta \in \mathcal{A}^{-0}$ and all $t > 0$

$$\frac{\mathbb{P}_\eta(\mathcal{T}_0 \le t)}{t} \le \frac{(e^{C_2 t} - 1)(1 - e^{-C_1 t})}{t} + \frac{1 - e^{-C_1 t}}{t}.$$

Finally from the q.s.d. equation we have $\int \mathbb{P}_\eta(\mathcal{T}_0 > t)d\nu(\eta) = e^{-\theta(\nu)t}$ and then $\int \mathbb{P}_\eta(\mathcal{T}_0 \le t)d\nu(\eta) = 1 - e^{-\theta(\nu)t}$. The result follows from the Dominated Convergence Theorem. □

4 Proof of the existence of q.s.d

In this section we give a proof of Theorem 1.2. The proof is based upon a more general result, Theorem 4.2, which shows that for a class of positive linear operators defined in some Banach spaces, whose elements are real functions with domain in a Polish space, there exist finite eigenmeasures. We show Theorem 1.2 in Sect. 4.2. For this purpose we construct the appropriate Banach spaces and the operator, in order that the eigenmeasure given by Theorem 4.2 is a q.s.d. of the original problem.

4.1 An abstract result

In this paragraph, $(\mathcal{X}, d)$ is a Polish metric space. We will denote by $\mathcal{C}_b(\mathcal{X})$ the set of bounded continuous functions on $\mathcal{X}$. This set becomes Banach space when equipped with the supremum norm. We will also make the following hypothesis.
Hypothesis $\mathscr{H}$: There exists a continuous real function φ_2 on $(\mathcal{X}, d)$ such that

$\mathscr{H}_1$ $\varphi_2 \geq 1$.
$\mathscr{H}_2$ For any $u \geq 0$, the set $\varphi_2^{-1}([0, u])$ is compact.

It follows from $\mathscr{H}_2$ that if $(\mathcal{X}, d)$ is not compact, there is a sequence $(x_j : j \in \mathbb{N})$ in $\mathcal{X}$ such that $\lim_{j\to\infty} \varphi_2(x_j) = \infty$.

Before stating the main result of this section we state and prove a lemma which will be useful later on.

Lemma 4.1 *Let v be a continuous nonnegative linear form on $\mathcal{C}_b(\mathcal{X})$. Assume there is a positive number K such that for any function $\psi \in \mathcal{C}_b(\mathcal{X})$ satisfying $0 \leq \psi \leq \varphi_2$, we have*

$$v(\psi) \leq K.$$

Then there exists a positive measure ν on $\mathcal{X}$ such that for any function $f \in \mathcal{C}_b(\mathcal{X})$

$$v(f) = \int f d\nu.$$

Proof Let $\mathcal{C}_0(\mathcal{X})$ be the set of continuous functions vanishing at infinity. Let ϖ be a real continuous non-increasing nonnegative function on $\mathbb{R}^+$. Assume that $\varpi = 1$ on the interval $[0, 1]$ and $\varpi = 0$ on $[2, \infty)$. For any integer m, let v_m be the continuous positive linear form defined on $\mathcal{C}_0(\mathcal{X})$ by

$$v_m(f) = v(\varpi(\varphi_2/m)f).$$

This linear form has support in the set $\varphi_2^{-1}([0, 2m])$ in the sense that it vanishes on those functions which vanish on this set. Note also that $\varphi_2^{-1}([0, 2m])$ is compact by hypothesis $\mathscr{H}_2$. Therefore it can be identified with a nonnegative measure ν_m on $\mathcal{X}$, namely for any $f \in \mathcal{C}_b(\mathcal{X})$ we have

$$v_m(f) = \int f d\nu_m.$$

We now prove that this sequence of measures is tight. Let $u > 0$ and define the set

$$K_u = \varphi_2^{-1}([0, u]).$$

Again, by hypothesis $\mathscr{H}_2$, for any $u > 0$ this is a compact set. We now observe that $\mathbf{1}_{K_u^c} \le 1 - \varpi(2\,\varphi_2/u)$. Therefore,

$$\begin{aligned}\nu_m\left(K_u^c\right) &\le \nu_m(1-\varpi(2\varphi_2/u)) = v_m(1-\varpi(2\varphi_2/u))\\ &= v(\varpi(\varphi_2/m)(1-\varpi(2\,\varphi_2/u))).\end{aligned}$$

We now use the fact that the function $\varphi_2\varpi(\varphi_2/m)(1-\varpi(2\,\varphi_2/u))$ is in $\mathcal{C}_b(\mathcal{X})$ and satisfies

$$\frac{u}{2}\varpi(\varphi_2/m)(1-\varpi(2\,\varphi_2/u)) \le \varpi(\varphi_2/m)(1-\varpi(2\varphi_2/u))\varphi_2 \le \varphi_2$$

to obtain from the hypotheses of the lemma that

$$v(\varpi(\varphi_2/m)(1-\varpi(2\varphi_2/u))) \le \frac{2}{u}v(\varpi(\varphi_2/m)(1-\varpi(2\varphi_2/u))\varphi_2) \le \frac{2K}{u}.$$

In other words, for any $u > 0$ we have for any integer m

$$\nu_m\left(K_u^c\right) \le \frac{2K}{u}.$$

The sequence of measures ν_m is therefore tight, and we denote by ν an accumulation point which is a nonnegative measure on $\mathcal{X}$. We now prove that for any $f \in \mathcal{C}_b(\mathcal{X})$ we have $\nu(f) = v(f)$. For this purpose, we write

$$v(f) = v(\varpi(\varphi_2/m)f) + v((1-\varpi(\varphi_2/m))f).$$

We now use the inequality

$$\varphi_2 \ge (1-\varpi(\varphi_2/m))\varphi_2 \ge m(1-\varpi(\varphi_2/m)),$$

to conclude using the hypothesis of the lemma (since $(1-\varpi(\varphi_2/m))\varphi_2 \in \mathcal{C}_b(\mathcal{X})$) that

$$|v((1-\varpi(\varphi_2/m))f)| \le v((1-\varpi(\varphi_2/m))|f|) \le \|f\|v(1-\varpi(\varphi_2/m)) \le \frac{K\|f\|}{m}.$$

In other words, we have for any $f \in \mathcal{C}_b(\mathcal{X})$

$$|v(f) - \nu_m(f)| \le \frac{K\|f\|}{m}.$$

From the tightness bound, we have for any $f \in \mathcal{C}_b(\mathcal{X})$

$$\lim_{m\to\infty} \nu_m(f) = \nu(f),$$

see for example [1], and therefore $\nu(f) = v(f)$ which completes the proof of the lemma. □

We now state the general result.

Theorem 4.2 *Assume hypothesis $\mathscr{H}$. Let S be a bounded positive linear operator on $\mathcal{C}_b(\mathcal{X})$. Assume that there exist three constants $c_1 > \gamma > 0$ and $D > 0$ such that*

$$S1 \geq c_1$$

and for any $\psi \in \mathcal{C}_b(\mathcal{X})$ with $0 \leq \psi \leq \varphi_2$

$$S\psi \leq \gamma\varphi_2 + D.$$

Then there is a probability measure ν on $\mathcal{X}$ such that $S^\dagger\nu = \beta\nu$, with $\beta = \nu(S1) > 0$ ($S^\dagger$ being the adjoint operator of S). Moreover, we have $\nu(\varphi_2) < \infty$.

Proof In the dual space $\mathcal{C}_b(\mathcal{X})^*$, we define for any real $K > 0$ the convex set $\mathscr{K}_K$ given by

$$\mathscr{K}_K = \left\{ v \in \mathcal{C}_b(\mathcal{X})^* : v \geq 0, v(1) = 1, \sup_{\psi \in \mathcal{C}_b(\mathcal{X}), 0 \leq \psi \leq \varphi_2} v(\psi) \leq K \right\},$$

Note that by Lemma 4.1, the elements of $\mathscr{K}_K$ are positive measures.

We observe that for any K large enough the set $\mathscr{K}_K$ is non empty. It suffices to consider a Dirac measure δ_x on a point $x \in \mathcal{X}$ and to take $K \geq \varphi_2(x)$. Since for any $K \geq 0$, $\mathscr{K}_K$ is an intersection of weak* closed subsets, it is closed in the weak* topology. Since it is contained in the ball of $\mathcal{C}_b(\mathcal{X})^*$ of radius K, it is compact in the weak* topology (see for example [24], Theorem 1 and Corollary, in Appendix to chapter V).

We now introduce the non-linear operator $\mathscr{U}$ having domain $\mathscr{K}_K$ and defined by

$$\mathscr{U}(v) = \frac{S^\dagger v}{v(S1)}.$$

Note that since $S1 > c_1 > 0$, we have $v(S1) \geq c_1 v(1)$ and this operator $\mathscr{U}$ is well defined on $\mathscr{K}_K$. We have obviously $\mathscr{U}(v)(1) = 1$. We now prove that $\mathscr{U}$ maps $\mathscr{K}_K$ into itself. Let $\psi \in \mathcal{C}_b(\mathcal{X})$ with $0 \leq \psi \leq \varphi_2$. Since

$$S\psi \leq \gamma\varphi_2 + D$$

and obviously

$$0 \leq S\psi \leq \gamma\frac{\|S\psi\|}{\gamma}$$

we get

$$0 \le S\psi \le \gamma(\varphi_2 \wedge (\|S\psi\|/\gamma)) + D.$$

Therefore since the function $\psi' = \varphi_2 \wedge (\|S\psi\|/\gamma)$ satisfies $\psi' \in \mathcal{C}_b(\mathcal{X})$ and $0 \le \psi' \le \varphi_2$. We conclude that for $v \in \mathscr{K}_K$

$$\mathscr{U}(v)(\psi) \le \frac{\gamma v(\psi') + D}{c_1}.$$

From the bound $v(\psi') \le K$ we get

$$\mathscr{U}(v)(\psi) \le \frac{\gamma v(\psi') + D}{c_1} \le \frac{\gamma}{c_1} K + \frac{D}{c_1} \le K$$

if $K > D/(c_1 - \gamma)$. Therefore, for any K large enough, the set $\mathscr{K}_K$ is non empty and mapped into itself by $\mathscr{U}$.

It is easy to show that $\mathscr{U}$ is continuous on $\mathscr{K}_K$ in the weak* topology. This follows at once from the continuity of the operator S. We can now apply Tychonov's fixed point theorem (see [20] or [8]) to deduce that $\mathscr{U}$ has a fixed point. This implies that there is a point $\nu \in \mathscr{K}_K$ such that $S^\dagger \nu = v(S1)\nu$. In particular we have $\nu(\varphi_2) < \infty$ and this concludes the proof of the Theorem. □

4.2 Proof of Theorem 1.2

We assume that the hypotheses of Theorem 1.2 hold. The proof of this result is based on Theorem 4.2 and Lemma 3.1. Here $S = P_1$, where P_t is the semigroup of Y acting on $\mathcal{C}_b(\mathcal{A}^{-0})$. Then

$$Sf(\eta) = P_1 f(\eta) = \mathbb{E}_\eta(f(Y_1), \mathscr{T}_0 > 1), \quad \eta \in \mathcal{A}^{-0}.$$

Here the Polish metric space $(\mathcal{X}, d)$ of the previous paragraph will be $(\mathcal{A}^{-0}, d_P)$ (where d_P is the Prokhorov metric), and so the function φ_2, of hypothesis $\mathscr{H}$, will be defined on the set of nonempty configurations.

We recall the elementary formula valid for any continuous and bounded function f on $\mathcal{A}$ and any $t \ge 0$

$$\mathbb{E}_\eta(f(Y_t)) = \mathbb{E}_\eta(f(Y_t), \mathscr{T}_0 > t) + f(0)\mathbb{P}_\eta(\mathscr{T}_0 \le t).$$

We start with the following bounds.

Lemma 4.3 *Let* $\overline{\lambda}_1 = \sup_{\eta:\|\eta\|=1} \sup_{y\in\eta} \lambda_y(\eta) < \infty$. *Then, the generator L defined in* (17), *verifies*

$$-\overline{\lambda}_1 \le L\mathbf{1}_{\mathcal{A}^{-0}} \le 0,$$

and

$$\forall t \geq 0 : e^{-\overline{\lambda}_1 t} \leq P_t \mathbf{1} \leq \mathbf{1}.$$

Proof The proof follows at once from Lemma 2.5 and a computation of $L\mathbf{1}_{\mathcal{A}^{-0}}$ [see formula (17)]. □

To define the function φ_2, we will need the two following results.

Lemma 4.4 *Let* $0 < v < u < \infty$ *be two real numbers. The differential equation*

$$\frac{da}{dt} = u(1 - e^{-a}) + v(1 - e^{a}) \tag{21}$$

has two fixed points $a = 0$ *and* $a = \log(u/v)$. *The trajectory of any initial condition* $a_0 \in (0, \log(u/v))$ *is increasing in time and converges to* $\log(u/v)$.

Proof Left to the reader. □

Lemma 4.5 *Assume* (6). *Let u and v be two real numbers such that*

$$\Gamma^* < v < u < \lambda_*.$$

Let $a(t)$ *be the solution of (21) with initial condition* $a_0 \in (0, \log(u/v))$. *Then there exists a constant* $C > 0$ *such that for any* $t \geq 0$

$$\mathbb{E}_\eta \left(e^{a(t)\|Y_t\|}, \mathcal{T}_0 > t \right) \leq e^{a_0\|\eta\|} + Ct.$$

Proof We introduce the function

$$f(t, \eta) = e^{a(t)\|\eta\|} \mathbf{1}_{\mathcal{A}^{-0}}(\eta),$$

and for any integer N we denote by f^N the function

$$f^N(t, \eta) = f(t, \eta) \mathbf{1}_{\|\eta\| \leq N}.$$

Note that $f^N(t, \eta)$ is continuous with compact support $\{\eta : \|\eta\| \leq N\}$.

Using Proposition 2.5 (iii) we get

$$f^N\left(t \wedge T_M^Y, Y_{t\wedge T_M^Y}\right) = f^N(0, Y_0) + \int_0^{t\wedge T_M^Y} \left(\partial_s f^N(s, Y_s) + L f^N(s, Y_s)\right) ds + \mathcal{M}^{f^N}_{t\wedge T_M^Y},$$

where $\mathscr{M}^{f^N}$ is a martingale and T_M^Y is the stopping time defined in (14). Then we obtain

$$\mathbb{E}_\eta\left(f^N\left(t\wedge T_M^Y, Y_{t\wedge T_M^Y}\right)\right) = f^N(0, Y_0) + \mathbb{E}_\eta\left(\int_0^{t\wedge T_M^Y}\left(\partial_s f^N(s, Y_s) + Lf^N(s, Y_s)\right)ds\right).$$

Observe that if $N > M$ and $s \le T_M^Y$ we have $f^N(s, Y_s) = f(s, Y_s)$. Let N tend to infinity to get

$$\mathbb{E}_\eta\left(f\left(t\wedge T_M^Y, Y_{t\wedge T_M^Y}\right)\right) = f(0, Y_0) + \mathbb{E}_\eta\left(\int_0^{t\wedge T_M^Y}(\partial_s f(s, Y_s) + Lf(s, Y_s))\,ds\right). \quad (22)$$

We now compute $L\varphi^a(\eta)$ using (17), where $\varphi^a(\eta) = e^{a\|\eta\|}\mathbf{1}_{\mathcal{A}^{-0}}(\eta)$. For $\eta \in \mathcal{A}^{-0}$ we have

$$L\varphi^a(\eta) = \sum_{y\in\{\eta\}}\eta_y(b_y(\eta) + m_y(\eta))(e^a - 1)e^{a\|\eta\|} + \sum_{y\in\{\eta\}}\eta_y\lambda_y(\eta)(e^{-a} - 1)e^{a\|\eta\|} - \lambda_y(\eta)\mathbf{1}_{\|\eta\|=1}.$$

Using hypothesis (6), we derive that there is a constant $C > 0$ such that for any $\eta \in \mathcal{A}^{-0}$ and for any $a \in [0, \log(u/v)]$

$$L\varphi^a(\eta) \le (u(e^{-a} - 1) + v(e^a - 1))\|\eta\|e^{a\|\eta\|} + C.$$

Therefore

$$\begin{aligned}&\partial_s f(s, Y_s) + Lf(s, Y_s)\\ &\quad = (u(1 - e^{-a(s)}) + v(1 - e^{a(s)}))\|Y_s\|\varphi^{a(s)}(Y_s) + L\varphi^{a(s)}(Y_s) \le C.\end{aligned}$$

This implies from (22)

$$\mathbb{E}_\eta\left(f\left(t, Y_{t\wedge T_M^Y}\right)\right) \le f(0, Y_0) + Ct.$$

Letting M tend to infinity and by using Fatou's Lemma we obtain,

$$\mathbb{E}_\eta(f(t, Y_t)) \le f(0, Y_0) + Ct.$$

The result follows from the definition of f. □

We now choose once for all two real numbers u and v such that

$$\Gamma^* < v < u < \lambda_*.$$

We take for φ_2 the function defined on $\mathcal{A}^{-0}$ by

$$\varphi_2(\eta) = \varphi^{a(1)}(\eta) = e^{a(1)\|\eta\|},$$

for a solution of (21) with initial condition $a_0 \in (0, \log(u/v))$. The operator $S = P_1$ is positive and maps continuously $\mathcal{C}_b(\mathcal{A}^{-0})$ into itself. We must now show that S, and φ_2 satisfy the hypothesis of Theorem 4.2.

Lemma 4.6 (i) *The hypotheses $\mathscr{H}_1$ and $\mathscr{H}_2$ are satisfied.*
(ii) $S1 \geq c_1 > 0$, *with* $c_1 = e^{-\overline{\lambda}_1}$.
(iii) *For any $\gamma > 0$, there is a constant $D = D(\gamma) > 0$ such that for any $\psi \in \mathcal{C}_b(\mathcal{A}^{-0})$ with $0 \leq \psi \leq \varphi_2$*

$$S\psi \leq \gamma\varphi_2 + D.$$

Proof The hypotheses $\mathscr{H}_1$ and $\mathscr{H}_2$ are easy to check using the Feller property of P_1 (see Proposition 2.8).

(ii) follows at once from Lemma 4.3. We now prove (iii).

Let $\psi \in \mathcal{C}_b(\mathcal{X})$ with $0 \leq \psi \leq \varphi_2$. We have from Lemma 4.5

$$P_1\psi(\eta) = \mathbb{E}_\eta(\psi(Y_1),\ \mathscr{T}_0 > 1) \leq \mathbb{E}_\eta(\varphi_2(Y_1),\ \mathscr{T}_0 > 1) \leq e^{a_0\|\eta\|} + C.$$

Since $a(1) > a_0$ by Lemma 4.4, for any $1 > \gamma > 0$ there is an integer m_γ such that for any $m \geq m_\gamma$ we have

$$e^{a_0\, m} \leq \gamma e^{a(1)m}.$$

Therefore, for any η we have

$$P_1\psi(\eta) \leq e^{a_0\|\eta\|} + C \leq \gamma e^{a(1)\|\eta\|} + e^{a_0\, m_\gamma} + C.$$

In other words, we have proved (iii) with the constant $D = e^{a_0 m_\gamma} + C$. □

Theorem 1.2 follows immediately from the previous Lemma, Theorem 4.2 and Lemma 3.1. Since $\nu(\varphi_2) < \infty$ we get that under ν, $\|Y_0\|$ has exponential moment and therefore for all $k \geq 1$ the integrability of $\|\eta\|^k$ with respect to ν. From the q.s.d. condition we have

$$\mathbb{E}_\nu(\|Y_1\|^k) = \mathbb{E}_\nu(\|Y_1\|^k, \mathscr{T}_0 > 1) = e^{-\theta(\nu)} \int \|\eta\|^k \nu(d\eta),$$

which implies (7).

5 The process and absolute continuity

In this section we introduce a natural $\sigma-$finite measure μ. We will show that the process Y preserves the absolute continuity with respect to μ and that when the process starts from any point measure after any positive time the absolutely continuous part of the marginal distribution does not vanish.

5.1 The measures

We will denote by $\widehat{\mathbb{T}^k}$ the set of all strictly increasing k-tuples in $\mathbb{T}$ with respect to the order $\preceq$ (see Sect. 2.1). So, for $\eta \in \mathcal{A}$, its ordered support $\vec{\eta} = (y_1, \ldots, y_{\#\eta})$ belongs to $\widehat{\mathbb{T}^{\#\eta}}$. The discrete structure $\overline{\eta} = (\eta_{y_1}, \ldots, \eta_{y_{\#\eta}})$ is an element in $\mathbb{N}^{\#\eta}$ and the set of all discrete structures is denoted by

$$\Sigma(\mathbb{N}) = \bigcup_{n \in \mathbb{Z}_+} \mathbb{N}^n.$$

Here $\mathbb{N}^0$ contains a unique element denoted by 0 and it is the discrete structure of the void configuration $\eta = 0$. A generic element of $\Sigma(\mathbb{N})$ will be denoted by $\mathbf{q}$. Moreover for each $\mathbf{q} \in \Sigma(\mathbb{N})$ we put $\#\mathbf{q} = k$ if $\mathbf{q} \in \mathbb{N}^k$. We put

$$\mathcal{A}_{\mathbf{q}} = \{\eta \in \mathcal{A} : \overline{\eta} = \mathbf{q}\} \ \text{ for } \mathbf{q} \in \Sigma(\mathbb{N}),$$

and for $B \subseteq \mathcal{A}$,

$$B_{\mathbf{q}} = \{\eta \in B : \overline{\eta} = \mathbf{q}\} \quad \text{for } \mathbf{q} \in \Sigma(\mathbb{N}).$$

In the sequel for $\mathbf{q} \in \mathbb{N}^k$ and $C \subseteq \widehat{\mathbb{T}^k}$ we denote

$$\{\mathbf{q}\} \times C := \{\eta \in \mathcal{A} : \overline{\eta} = \mathbf{q}, \vec{\eta} \in C\}.$$

We denote by $\mathcal{M}_f(\mathcal{A})$ the set of measures on $(\mathcal{A}, \mathcal{B}(\mathcal{A}))$ that give finite weight to all sets $\mathcal{A}_{\mathbf{q}}$. By $\mathcal{M}_f(\Sigma(\mathbb{N}))$ we mean the set of measures on $\Sigma(\mathbb{N})$ giving finite weight to all the subsets $\mathbb{N}^k$, and $\mathcal{M}_f(\mathbb{N})$ denotes the measures on $\mathbb{N}$ giving finite weight to all its points. Every measure $v \in \mathcal{M}_f(\mathcal{A})$ defines a measure $\overline{v} \in \mathcal{M}_f(\Sigma(\mathbb{N}))$ by

$$\overline{v}(\mathbf{q}) = v(\mathcal{A}_{\mathbf{q}})\,. \tag{23}$$

Also v defines a set of conditional measures $v_{\mathbf{q}} \in \mathcal{M}(\widehat{\mathbb{T}^{\#\mathbf{q}}})$ by

$$v_{\mathbf{q}}(\bullet) = \begin{cases} 0 & \text{if } \overline{v}(\mathbf{q}) = 0, \\ v(\eta \in \mathcal{A} : \overline{\eta} = \mathbf{q}, \vec{\eta} \in \bullet)/\overline{v}(\mathbf{q}) & \text{otherwise.} \end{cases}$$

Then $v_{\mathbf{q}} \in \mathcal{P}(\widehat{\mathbb{T}^{\#\mathbf{q}}})$ is a probability measure if $\overline{v}(\mathbf{q}) > 0$. In the case that $v \in \mathcal{P}(\mathcal{A})$ we have

$$v_{\mathbf{q}}(\bullet) = v(\vec{\eta} \in \bullet | \overline{\eta} = \mathbf{q}).$$

Conversely a probability measure $v \in \mathcal{P}(\mathcal{A})$ is given by a probability measure $\overline{v} \in \mathcal{P}(\Sigma(\mathbb{N}))$ and the family of conditional measures $(v_{\overline{\eta}} \in \mathcal{P}(\mathcal{A}_{\overline{\eta}}))$ so that

$$v(B) = \sum_{\mathbf{q} \in \Sigma(\mathbb{N})} \overline{v}(\mathbf{q}) v_{\mathbf{q}}(B_{\mathbf{q}}), \quad B \in \mathcal{B}(\mathcal{A}^{-0}).$$

Let $\varphi : \mathcal{A} \to \mathbb{R}$ be a function. Observe that its restriction to $\mathcal{A}_{\mathbf{q}}$ can be identified with a function $\varphi\big|_{\mathcal{A}_{\mathbf{q}}}$ with domain in $\widehat{\mathbb{T}^{\#\mathbf{q}}}$ by the formula $\varphi\big|_{\mathcal{A}_{\mathbf{q}}}(\vec{\eta}) = \varphi(\eta)$. Let $\varphi : \mathcal{A} \to \mathbb{R}$ be a $v-$integrable function, we have

$$\int_{\mathcal{A}} \varphi(\eta) dv(\eta) = \sum_{\vec{q} \in \Sigma(\mathbb{N})} \overline{v}(\mathbf{q}) \int_{\widehat{\mathbb{T}^{\#\mathbf{q}}}} \varphi\big|_{\mathcal{A}_{\mathbf{q}}}(\vec{y}) dv_{\mathbf{q}}(\vec{y}).$$

Now we define the measure μ by,

$$\mu(\mathbb{N}^0 \times \mathbb{T}^0) = \mu(\{0\}) = 1 \text{ and } \mu|_{\mathbb{N}^k \times \widehat{\mathbb{T}^k}} = \ell^k \times \sigma^k \text{ for } k \geq 1,$$

where ℓ^k is the point measure on $\mathbb{N}^k$ that gives a unit mass to every point, and σ^k is the restriction to $\widehat{\mathbb{T}^k}$ of the product measure $\sigma^{\otimes k}$ defined on $\mathbb{T}^k$. Note that $v \in \mathcal{M}_f(\mathcal{A})$ satisfies

$$v \ll \mu \Leftrightarrow \left(\forall \mathbf{q} \in \Sigma(\mathbb{N}) : v_{\mathbf{q}} \ll \sigma^{\#\mathbf{q}}\right).$$

Hence, if $v \in \mathcal{P}(\mathcal{A})$ is such that $v \ll \mu$, then v is of the form

$$v(\eta \in \mathcal{A} : \overline{\eta} = \vec{q}, \vec{\eta} \in d\vec{y}) = \overline{v}(\mathbf{q}) \varphi_{\mathbf{q}}(\vec{y}) d\sigma^{\#\mathbf{q}}(\vec{y}).$$

where $\overline{v} \in \mathcal{P}(\Sigma(\mathbb{N}))$ and for each fixed $\mathbf{q} \in \Sigma(\mathbb{N})$, $\varphi_{\mathbf{q}}(\vec{\bullet})$ is a density function in $\widehat{\mathbb{T}^{\#\mathbf{q}}}$ with respect to $\sigma^{\#\mathbf{q}}$.

5.2 Absolute continuity is preserved

Proposition 5.1 *The process Y preserves the absolutely continuity with respect to μ, that is*

$$\forall v \in \mathcal{P}(\mathcal{A}), v \ll \mu \Rightarrow \forall t > 0 \; \mathbb{P}_v(Y_t \in \bullet) \ll \mu(\bullet).$$

Proof Let us define the jump time sequence,

$$\tau_0 = 0 \text{ and } \tau_n = \inf\{t > \tau_{n-1} : Y_t \neq Y_{\tau_{n-1}}\} \quad \text{for } n \geq 1.$$

In particular $\tau = \tau_1$ is the time of the first jump. Remark that the sequence τ_n tends almost surely to infinity since the process Y has no explosions. We have

$$\mathbb{P}_v(Y_t \in \bullet) = \sum_{n\geq 0} \mathbb{P}_v(Y_t \in \bullet, \tau_n \leq t < \tau_{n+1}).$$

Since the sum of absolutely continuous measures is also absolutely continuous it suffices to prove that

$$\mathbb{P}_v(Y_t \in \bullet, \tau_n \leq t < \tau_{n+1}) \ll \mu \quad \text{for all } n \geq 0.$$

First, let us show the case $n = 0$. For $B \in \mathcal{B}(\mathcal{A})$

$$\mathbb{P}_v(Y_t \in B, t < \tau) = v(B)$$

and the property holds trivially in this case.

Now, for $n \geq 1$ we have

$$\mathbb{P}_v(Y_t \in \bullet, \tau_n \leq t < \tau_{n+1}) \leq \mathbb{P}_v(Y_{\tau_n} \in \bullet).$$

So, in order to prove that $\mathbb{P}_v(Y_t \in \bullet, \tau_n \leq t < \tau_{n+1})$ is absolutely continuous it is enough to prove this property for $\mathbb{P}_v(Y_{\tau_n} \in \bullet)$. On the other hand, the strong Markov property shows, for $n \geq 1$,

$$\mathbb{P}_v(Y_{\tau_n} \in \bullet) = \mathbb{E}_v(\mathbb{P}_{Y_\tau}(Y_{\tau_{n-1}} \in \bullet)).$$

A recurrence argument shows the result as soon as we prove $\mathbb{P}_v(Y_\tau \in \bullet)$ is absolutely continuous. For proving the absolute continuity, it is enough to consider a fixed $\mathbf{p} \in \Sigma(\mathbb{N})$ with $\#\mathbf{p} = k$, $C \in \mathcal{B}(\widehat{\mathbb{T}^k})$ such that $\sigma^k(C) = 0$ and prove that

$$\mathbb{P}_v(Y_\tau \in \{\mathbf{p}\} \times C) = 0.$$

To develop a formula that will show this property it will be useful to introduce the following relation on $\Sigma(\mathbb{N})$. This relation gives the allowed transitions from $\mathbf{q}$ to $\mathbf{p}$.

We put

$$\mathbf{q} \xrightarrow[b]{i} \mathbf{p} \Leftrightarrow \#\mathbf{q} = k, \qquad \begin{cases} q_j = p_j & j \neq i \\ q_i = p_i - 1 \\ q_i \geq 1 \end{cases}$$

$$\mathbf{q} \xrightarrow[m]{i} \mathbf{p} \Leftrightarrow \#\mathbf{q} = k-1, \quad \begin{cases} q_j = p_j & 1 \leq j \leq i-1 \\ q_j = p_{j+1} & i \leq j \leq k-1 \\ p_i = 1 \end{cases}$$

$$\mathbf{q} \xrightarrow[1,d]{i} \mathbf{p} \Leftrightarrow \#\mathbf{q} = k, \qquad \begin{cases} q_j = p_j & j \neq i \\ q_i = p_i + 1 \\ p_i \geq 1 \end{cases}$$

$$\mathbf{q} \xrightarrow[0,d]{i} \mathbf{p} \Leftrightarrow \#\mathbf{q} = k+1, \quad \begin{cases} q_j = p_j & 1 \leq j \leq i-1 \\ q_{j+1} = p_j & i \leq j \leq k \\ q_i = 1 \end{cases}$$

With this notation we obtain

$$\begin{aligned}
\mathbb{P}_v(Y_\tau \in \{\mathbf{p}\} \times C) &= \int_{\mathcal{A}} dv(\eta) \int_{\{\mathbf{p}\}\times C} \frac{1}{Q(\eta)} Q(\eta, d\eta') \\
&= \sum_{i=1}^{k} \sum_{\mathbf{q}:\mathbf{q} \xrightarrow[b]{i} \mathbf{p}} \bar{v}(\mathbf{q}) \int_{\vec{y}\in C} \frac{q_i b_{y_i}\left(\sum_{j=1}^k q_j \delta_{y_j}\right)}{Q\left(\sum_{j=1}^k q_j \delta_{y_j}\right)} \varphi_{\mathbf{q}}(\vec{y}) d\sigma^k(\vec{y}) \\
&+ \sum_{i=1}^{k} \sum_{\mathbf{q}:\mathbf{q} \xrightarrow[1,d]{i} \mathbf{p}} \bar{v}(\mathbf{q}) \int_{\vec{y}\in C} \frac{q_i \lambda_{y_i}\left(\sum_{j=1}^k q_j \delta_{y_j}\right)}{Q\left(\sum_{j=1}^k q_j \delta_{y_j}\right)} \varphi_{\mathbf{q}}(\vec{y}) d\sigma^k(\vec{y}) \\
&+ \sum_{i=1}^{k+1} \sum_{\mathbf{q}:\mathbf{q} \xrightarrow[0,d]{i} \mathbf{p}} \bar{v}(\mathbf{q}) \int_{\vec{y}\in C} \left(\int_{\{z: y_{i-1} \preceq z \preceq y_i\}} \frac{\lambda_z\left(\sum_{j=1}^k p_j \delta_{y_j} + \delta_z\right)}{Q\left(\sum_{j=1}^k p_j \delta_{y_j} + \delta_z\right)} \varphi_{\mathbf{q}}(\vec{y}^{+i,z}) d\sigma(z) \right) d\sigma^k(\vec{y}) \\
&+ \sum_{i=1}^{k} \sum_{\mathbf{q}:\mathbf{q} \xrightarrow[m]{i} \mathbf{p}} \bar{v}(\mathbf{q}) \int_{\vec{y}\in C} \frac{G\left(\sum_{j=1}^k p_j \delta_{y_j} - \delta_{y_i}, y_i\right)}{Q\left(\sum_{j=1}^k p_j \delta_{y_j} - \delta_{y_i}\right)} \varphi_{\mathbf{q}}(\vec{y}^{-i}) d\sigma^k(\vec{y}).
\end{aligned}$$

where $\vec{y}^{+i,z} = (y_1, \ldots, y_{i-1}, z, y_i, \ldots, y_k)$ and $\vec{y}^{-i} = (y_1, \ldots, y_{i-1}, y_{i+1}, \ldots, y_k)$ when $\vec{y} = (y_1, \ldots, y_k) \in \widehat{\mathbb{T}^k}$. We take by convention $\vec{y}^{+1,z} = (z, y_1, \ldots, y_k)$ and $\vec{y}^{+(k+1),z} = (y_1, \ldots, y_k, z)$. When $\vec{y} \in \widehat{\mathbb{T}^k}$ the restriction $y_0 \preceq z \preceq y_1$ is interpreted as $z \preceq y_1$ and similarly $y_k \preceq z \preceq y_{k+1}$ as $y_k \preceq z$. We conclude that $\mathbb{P}_v(Y_\tau \in \{\mathbf{q}\} \times C) = 0$ whenever $\mu(\{\mathbf{q}\} \times C) = 0$, and the result follows. □

5.3 Evolution after the first mutation

We want to study the absolute continuity with respect to μ of the law of Y_t initially distributed according to a general measure v. To this aim we will introduce the first mutation time. Note that a mutant individual has a different trait from those of its parent, so the time of first mutation is almost surely

$$\chi = \inf\{t \geq 0 : \{Y_t\} \not\subseteq \{Y_0\}\}.$$

When χ is finite we have $\{Y_\chi\} \neq \emptyset$, so $(\chi < \infty) \Rightarrow (\chi < \mathcal{T}_0)$.

Now, let us consider the first time where the traits of the initial configuration disappear,

$$\kappa = \inf\{t \geq 0 : \{Y_t\} \cap \{Y_0\} = \emptyset\},$$

and for a fixed η, the first time where the traits of η disappear

$$\kappa^\eta = \inf\{t \geq 0 : \{Y_t\} \cap \{\eta\} = \emptyset\}.$$

When $\{\eta\} \cap \{Y_0\} = \emptyset$, then κ^η=0. We have $\kappa \neq \chi$ except when $\kappa = \chi = \infty$. Obviously $\kappa \leq \mathcal{T}_0$. Moreover

$$(\chi > \kappa) \Leftrightarrow (\infty = \chi > \kappa) \Leftrightarrow (\chi > \kappa = \mathcal{T}_0), \text{ and } (\kappa < \mathcal{T}_0) \Leftrightarrow (\chi < \kappa < \mathcal{T}_0).$$

Also note that $(\kappa < \chi) \cap (\kappa \leq t) \subseteq (\kappa = \mathcal{T}_0 \leq t)$.

Since $\mathbb{P}_\eta(Y_t \in \bullet, \mathcal{T}_0 \leq t) = \delta_0(\bullet)\mathbb{P}_\eta(\mathcal{T}_0 \leq t)$ is concentrated on $\eta = 0$, then

$$\mathbb{P}_\eta(Y_t \in \bullet, \kappa < \chi, \kappa \leq t) = \delta_0(\bullet)\mathbb{P}_\eta(\kappa < \chi, \kappa \leq t).$$

The unique nontrivial cases are the following two ones.

Proposition 5.2 *Let $\eta \in \mathcal{A}^{-0}$ and $t \geq 0$, we have*:

(i) $\mathbb{P}_\eta(Y_t \in \bullet, \chi < \kappa \leq t < \mathcal{T}_0)$ *is absolutely continuous with respect to μ and it is concentrated in $\mathcal{A}^{-0}$;*

(ii) $\mathbb{P}_\eta(Y_t \in \bullet, t < \kappa)$ *is singular with respect to μ.*

Proof Let us show (i). From the Markov property we have,

$$\begin{aligned}\mathbb{P}_\eta(Y_t \in \bullet, \chi < \kappa \leq t) &= \sum_{\xi:\emptyset\neq\{\xi\}\subseteq\{\eta\}} \mathbb{P}_\eta(\chi < \kappa \leq t, Y_{\chi^-} = \xi, Y_t \in \bullet)\\ &= \sum_{\xi:\emptyset\neq\{\xi\}\subseteq\{\eta\}} \sum_{y\in\{\xi\}} \left(\frac{\xi_y m_y(\xi)}{\sum\limits_{y'\in\{\xi\}} \xi_{y'} m_{y'}(\xi)}\right) \int_0^t \mathbb{P}_\eta(\chi \in ds, Y_{s^-} = \xi)\\ &\quad\times \int_{\mathbb{T}\setminus\{\xi\}} g(y,z)\mathbb{P}_{\xi+z}(Y_{t-s} \in \bullet, \kappa^\xi \leq t-s)d\sigma(z).\end{aligned}$$

Hence, it is sufficient to show that for every $u > 0$, $\eta \in \mathcal{A}^{-0}$ and $y \in \{\eta\}$, it holds

$$\int_{\mathbb{T}\setminus\{\eta\}} \mathbb{P}_{\eta+z}(Y_u \in \bullet, \kappa^\eta \le u)\, g(y,z) d\sigma(z) \ll \mu(\bullet).$$

By using $\int_{\mathbb{T}\setminus\{\eta\}} \mathbb{P}_{\eta+z}(\{Y_u\} \cap \{\eta\} \neq \emptyset, \kappa^\eta \le u) g(y,z) d\sigma(z) = 0$, and since the measure σ is non-atomic, a similar proof to the one showing Proposition 5.1 works and proves the result. Indeed, for each $t > 0$, the singular part with respect to μ of $\mathbb{P}_\eta(Y_t \in \cdot)$ is a measure on the set of atomic measures with support contained in $\{\eta\}$ (corresponding to death or clonal events from individuals initially alive).

Let us show (ii). Let $\{\eta\} \subset \mathbb{T}$ be the finite set of initial traits and put $k = \#\eta$. Consider the Borel set $B = \{\xi \in \mathcal{A}^{-0} : \{\xi\} \cap \{\eta\} \neq \emptyset\}$ and define $B_{l,n} = \{\xi \in \mathcal{A}^{-0} : \#\xi = n, |\{\xi\} \cap \{\eta\}| = l\}$ for $n \in \mathbb{N}, l = 1, \ldots, n \wedge k$. We have $B = \bigcup_{n \in \mathbb{N}, l \in \{1,\ldots,n\wedge k\}} B_{l,n}$. Since σ is non-atomic we have $\mu(B_{l,n}) = 0$ for all $l \in \{1, \ldots, n \wedge k\}$. On the other hand, from the definition of κ we have $\mathbb{P}_\eta(Y_t \in B, t < \kappa) = 1$, and the result follows. □

Let $v \in \mathcal{P}(\mathcal{A})$. We denote by v^t the distribution of Y_t when the distribution of Y_0 is v, that is

$$v^t(B) = \mathbb{P}_v(Y_t \in B), \qquad B \in \mathcal{B}(\mathcal{A}), \quad t \ge 0.$$

We denote by $v = v^{\text{ac}} + v^{\text{si}}$ the Lebesgue decomposition of v into its absolutely continuous part $v^{\text{ac}} \ll \mu$ and its singular part v^{si} with respect to μ. For v^t this decomposition is written as $v^t = v^{t,\text{ac}} + v^{t,\text{si}}$. As usual, δ_η is the Dirac measure at $\eta \in \mathcal{A}$, so δ^t_η denotes the measure $\delta^t_\eta(\bullet) = \mathbb{P}_\eta(Y_t \in \bullet)$. We will denote by $d_P(\eta, \eta')$ the Prokhorov distance between η and η'.

Proposition 5.3 *The process Y verifies*:

(i) *For all $t > 0$ and all $\eta \in \mathcal{A}^{-0}$ we have $\delta^{t,ac}_\eta(\mathcal{A}^{-0}) > 0$;*

(ii) *For all $t > 0$ and all $v \in \mathcal{P}(\mathcal{A}^{-0})$ it holds $v^{t,ac} \ge \int_{\mathcal{A}^{-0}} \delta^{t,ac}_\eta v(d\eta) > 0$;*

(iii) *For all $\eta \in \mathcal{A}^{-0}$ and $\epsilon > 0$, the following relation holds*

$$\forall t > 0, \quad \delta^{t,ac}_\eta(B(\xi, \epsilon)) > 0,$$

where $B(\xi, \epsilon) = \{\xi' \in \mathcal{A} : d_P(\xi, \xi') < \epsilon\}$. In particular the closed support of $\delta^{t,ac}_\eta$ is $\mathcal{A}^{-0}$, for all $t > 0$.

Proof It suffices to show (i) and (iii). Let us show the first part. Fix $t \ge 0$ and $\eta \in \mathcal{A}^{-0}$. We claim that $\mathbb{P}_\eta(\chi < \kappa \le t < \mathscr{T}_0) > 0$. In fact, it suffices to consider the event where a mutation occurs at the first jump and after it all the initial traits disappear before t and these changes are the unique ones before t. This event has strictly positive probability, so the claim is proved. Proposition 5.2 (i) gives $\mathbb{P}_\eta(Y_t \in \bullet, \chi < \kappa \le t < \mathscr{T}_0) \ll \mu$ and we deduce,

$$\delta^{t,\text{ac}}_\eta(\mathcal{A}^{-0}) \ge \mathbb{P}_\eta(\chi < \kappa \le t < \mathscr{T}_0) > 0.$$

Then (i) holds.

The proof of (iii) is entirely similar to the proof of (i) but we need a previous remark. From condition (9), the set

$$D = \left\{ y \in \mathbb{T} : \sigma\left(\left\{ z \in \mathbb{T} : g_{k_0}(y, z) > 0 \right\} \right) = 1 \right\}$$

verifies $\sigma(D) = 1$.

Now, let $\vec{\eta} = (y_1, \ldots, y_k)$ and let $\xi \in \mathcal{A}^{-0}$ with $\vec{\xi} = (z_1, \ldots, z_l)$ and $\overline{\xi} = (p_1, \ldots, p_l)$. Consider the following event: a mutation occurs at the first jump, from the initial configuration η, to $\eta + \delta_{y'}$ with $y' \in D$. This is possible because $\sigma(D) = 1$ and so $\sigma(\{z \in D : g(y_j, z) > 0\}) > 0$ for all j. Consider then the event where l blocs of k_0 births with mutations occur in order to arrive to traits $z'_i \in B(z_i, \epsilon) \cap D$ for $i = 1, \ldots, l$. We must now produce the discrete structure $\overline{\xi}$. For this purpose, for each trait z'_i we consider the trajectories having $p_i - 1$ clonal births (when $p_i > 1$). Finally we consider the event where all the traits different from $z'_1, \ldots, z'_l$ disappear. These events occur before t and these changes are the unique ones that happen before t. A straightforward argument shows that the measure $\delta_\eta^{t,\mathrm{ac}}$ of this event is strictly positive. The claim is proved. □

5.4 Decomposition of q.s.d

Let us study the Lebesgue decomposition of a q.s.d. with respect to μ.

Proposition 5.4 *Let ν be a q.s.d. on $\mathcal{A}^{-0}$. Then,*

- (i) $\nu^{ac} \neq 0$;
- (ii) *The closed support of ν^{ac} is $\mathcal{A}^{-0}$;*
- (iii) *If $\nu^{si} \neq 0$, the probability measure $\nu^{*si} := \nu^{si}/\nu^{si}(\mathcal{A}^{-0})$ satisfies*

$$\mathbb{P}_{\nu^{*si}}(Y_t \in B) = e^{-\theta(\nu)t} \nu^{*si}(B) \quad \forall B \in \mathcal{B}(B^{si}), t \geq 0, \tag{24}$$

where $B^{si} \in \mathcal{B}(\mathcal{A}^{-0})$ is a measurable set such that $\mu(B^{si}) = 0$ and $\nu^{si}(B^{si}) = \nu^{si}(\mathcal{A}^{-0})$.

Proof We first note that the existence of the set $B^{\mathrm{si}} \in \mathcal{B}(\mathcal{A}^{-0})$ satisfying $\mu(B^{\mathrm{si}}) = 0$ and $\nu^{\mathrm{si}}(B^{\mathrm{si}}) = \nu^{\mathrm{si}}(\mathcal{A}^{-0})$ is ensured by the Radon–Nikodym decomposition theorem. Define $H := \mathcal{A}^{-0} \setminus B^{\mathrm{si}}$. Let us show that

$$\forall t > 0, \quad \forall \eta \in \mathcal{A}^{-0} : \delta_\eta^t(H) > 0. \tag{25}$$

Since $\mu(B^{\mathrm{si}}) = 0$ and $\delta_\eta^{t,\mathrm{ac}} \ll \mu$ we have $\delta_\eta^{t,\mathrm{ac}}(B^{\mathrm{si}}) = 0$. Then $\delta_\eta^{t,\mathrm{ac}}(H) = \delta_\eta^{t,\mathrm{ac}}(\mathcal{A}^{-0})$. By Proposition 5.3, $\delta_\eta^{t,\mathrm{ac}}(\mathcal{A}^{-0}) > 0$ for all $t > 0$ and all $\eta \in \mathcal{A}^{-0}$. So

$$\delta_\eta^t(H) \geq \delta_\eta^{t,\mathrm{ac}}(H) = \delta_\eta^{t,\mathrm{ac}}(\mathcal{A}^{-0}) > 0,$$

and the assertion (25) holds.

Now we prove part (i). We can assume $\nu^{\mathrm{si}} \neq 0$, if not the result is trivial. From (25) we get,

$$\nu^t(H) = \int_{\mathcal{A}^{-0}} \delta_\eta^t(H)\nu(d\eta) > 0.$$

On the other hand, from relation (19) we obtain $\nu(H) = e^{\theta(\nu)t}\nu^t(H) > 0$. Since $\nu^{\mathrm{si}}(H) = 0$, we necessarily have $\nu^{\mathrm{ac}}(H) = \nu(H) > 0$, so (i) holds.

Similar arguments as above and the use of (iii) in Proposition 5.3, show (ii).

Let us show (iii). Let $\nu^{*\mathrm{ac}} := \nu^{\mathrm{ac}}/\nu^{\mathrm{ac}}(\mathcal{A}^{-0})$. For every $B \subseteq B^{\mathrm{si}}$, $B \in \mathcal{B}(\mathcal{A}^{-0})$, we have

$$\nu(B) = e^{\theta(\nu)t}\left(\nu^{\mathrm{ac}}(\mathcal{A}^{-0})\mathbb{P}_{\nu^{*\mathrm{ac}}}(Y_t \in B) + \nu^{\mathrm{si}}(\mathcal{A}^{-0})\mathbb{P}_{\nu^{*\mathrm{si}}}(Y_t \in B)\right). \tag{26}$$

By Proposition 5.1, Y preserves absolute continuity with respect to μ, so $\mathbb{P}_{\nu^{*\mathrm{ac}}}(Y_t \in \bullet) \ll \mu$. Since $\mu(B^{\mathrm{si}}) = 0$ we get $\mathbb{P}_{\nu^{*\mathrm{ac}}}(Y_t \in B^{\mathrm{si}}) = 0$. By evaluating (26) at $t = 0$ and since $B \in \mathcal{B}(B^{\mathrm{si}})$ we find $\nu^{*\mathrm{si}}(B) = \nu(B)/\nu(B^{\mathrm{si}})$. By putting all these elements together we obtain relation (24). □

6 The uniform case

6.1 The model

In this section, we assume that the individual jump rates satisfy,

$$\lambda_y(\eta) = \lambda, \quad b_y(\eta) = b(1-\rho), \quad m_y(\eta) = b\rho, \qquad \forall y \in \{\eta\},$$

λ, b and ρ are positive numbers with $\rho < 1$. Recall that $g : \mathbb{T} \times \mathbb{T} \to \mathbb{R}_+$ is a jointly continuous nonnegative function satisfying $\int_{\mathbb{T}} g(y, c)d\sigma(c) = 1$ for all $y \in \mathbb{T}$ and the condition (9).

We observe that in this case the process of the total number of individuals $\|Y\| = (\|Y_t\| : t \geq 0)$ is a Markov process and that $Y_t = 0 \Leftrightarrow \|Y_t\| = 0$, which means that the time of absorption at 0 of the processes Y and $\|Y\|$ is the same (note that although the 0's have a different meaning, they are identified).

Now, in [21] it is shown that there exists a q.s.d. for the process $\|Y\|$ killed at 0 if and only if $\lambda > b$. In addition, the extremal exponential decay rate of $\|Y\|$, defined by $\sup\{\theta(\nu) : \nu \ q.s.d.\}$, is equal to $\lambda - b$ and there exists a unique (extremal) q.s.d. $\zeta^{\mathbf{e}}$ for $\|Y\|$ with this exponential decay rate $\lambda - b$, given by

$$\zeta^{\mathbf{e}}(k) = \left(\frac{b}{\lambda}\right)^{k-1}\left(1 - \frac{b}{\lambda}\right), \quad k \geq 1. \tag{27}$$

When ν is a q.s.d. for Y with exponential decay rate $\theta(\nu)$ then the probability vector $\zeta = (\zeta(k) : k \in \mathbb{N})$ given by

$$\zeta(k) = \nu(\{\eta \in \mathcal{A} : \|\eta\| = k\}), \quad k \in \mathbb{N},$$

is a q.s.d. with exponential decay rate $\theta = \theta(\nu)$, associated with the linear birth and death process $\|Y\|$. Hence a necessary condition for the existence of a q.s.d. for the process Y is $\lambda > b$. We also deduce that all quasi-stationary probability measures $\tilde{\nu}$ of Y with exponential decay rate $\lambda - b$ are such that $\tilde{\nu}(\{\eta \in \mathcal{A} : \|\eta\| = k\}) = \zeta^{\mathbf{e}}(k)$, so by (27) we get

$$\tilde{\nu}(\varphi_1) < \infty, \quad \text{where } \varphi_1(\eta) = \|\eta\|.$$

Now, we know from Theorem 1.2 that there exists a q.s.d. ν with exponential decay rate $\theta(\nu) = -\log(\nu(P_1(1)))$. Moreover, it is immediate to show that φ_1 satisfies $L\varphi_1 = -(\lambda - b)\varphi_1$. Then, from Proposition 2.5 we get $P_1\varphi_1 = e^{-(\lambda-b)}\varphi_1$. Hence, if ν is a q.s.d. satisfying $\nu(\varphi_1) < \infty$ (this is the case for those provided by Theorem 1.2) its exponential decay rate is

$$\theta = \lambda - b.$$

Let us now consider the semi-group R_t given by,

$$R_t(\varphi)(\eta) = e^{\theta t} P_t(\varphi)(\eta) = e^{\theta t}\mathbb{E}_\eta\big(\varphi(Y_t)\mathbf{1}_{\mathcal{T}_0 > t}\big), \quad t \geq 0.$$

The function φ_1 satisfies $R_t\varphi_1 = \varphi_1$.

Proposition 6.1 *Every q.s.d. ν with exponential decay rate $\theta = \lambda - b$ is absolutely continuous with respect to μ.*

Proof Let ν be a q.s.d. which is not absolutely continuous. Then we can write the Lebesgue decomposition

$$\nu = f\mu + \xi \quad \text{that is } \nu(B) = \int_B f d\mu + \xi(B), B \in \mathcal{B}(E),$$

where f is a nonnegative $\mu-$integrable function and ξ is a singular measure with respect to μ.

From now on we denote by $R_t^\dagger$ the dual action of R_t on the set of measures defined by $(R_t^\dagger v)(\varphi) = v(R_t\varphi)$ for every measure $v \in \mathcal{M}_f(\mathcal{A})$ and any positive measurable function φ. Since ν is a q.s.d. it is invariant by the adjoint semi-group $R_t^\dagger$, that is $R_t^\dagger \nu = \nu$, then

$$f\mu + \xi = \nu = R_t^\dagger(\nu) = R_t^\dagger(f\mu) + R_t^\dagger(\xi).$$

On the other hand, it follows from Proposition 5.1 that $R_t^\dagger(f\mu) \ll \mu$. Therefore

$$R_t^\dagger(f\mu) \le f\mu.$$

Since φ_1 is ν integrable it must also be $f d\mu$ integrable. From the relation $R_t(\varphi_1) = \varphi_1$ we get,

$$\int \varphi_1 f d\mu = \int \varphi_1 d R_t^\dagger(f\mu),$$

and since φ_1 is strictly positive, we conclude $R_t^\dagger(f\mu) = f\mu$. This implies $R_t^\dagger(\xi) = \xi$. However by Proposition 5.3 (ii) (with $v = \xi$), $R_t^\dagger(\xi)$ cannot be completely singular with respect to μ unless ξ vanishes. This concludes the proof of the proposition. □

Let us now turn to the study of uniqueness. Recall that the we denote by (g_k) the sequence of kernels defined recursively by $g_1(x, y) = g(x, y)$ and

$$g_{k+1}(x, y) = \int_{\mathbb{T}} g(x, z) g_k(z, y) d\sigma(z).$$

Lemma 6.2 *Assume condition* (9): *that is* $\sigma \otimes \sigma(\{g_{k_0} = 0\}) = 0$, *for* $k_0 \ge 1$. *Let* $\mathcal{A}_{(1,1)} = \{\eta \in \mathcal{A} : \mathbf{q} = (1,1)\}$. *Then for any q.s.d.* ν *with exponential decay rate* $\theta = \lambda - b$ *and for any Borel set* B *with* $\mu(\mathcal{A}_{(1,1)} \cap B) > 0$, *we have* $\nu(\mathcal{A}_{(1,1)} \cap B) > 0$.

Proof Let us consider a q.s.d. ν with exponential decay rate $\theta = \lambda - b$. Lemma 3.2 implies that the restriction $\nu_{(1)}$ of ν to $\{\eta \in \mathcal{A} : \|\eta\| = 1\}$, does not vanish. On the other hand by the previous result, it is absolutely continuous with respect to σ. Then,

$$d\nu_{(1)} = f_1 d\sigma$$

for some nonnegative function f_1 that does not vanish σ almost surely.

For any function f in $\mathcal{C}_b(\mathcal{A})$ such that $f(0) = 0$ and with compact support, that is $f(\eta) = 0$ for all $\|\eta\|$ large enough, it follows from $\nu(P_t f) = \exp(-\theta t)\nu(f)$ that

$$\nu(Lf\mathbf{1}_{\mathcal{A}^{-0}}) = -\theta\nu(f).$$

Since this is true for any such function we get, using notations introduced in (10) and (23),

$$- \theta\overline{\nu}(\overline{\eta}) d\nu_{\overline{\eta}}(\vec{\eta}) = b(1-\rho) \sum_{y:\eta_y > 1} \big(\eta_y - 1\big)\overline{\nu}(\overline{\eta}^{-y}) d\nu_{\overline{\eta}^{-y}}(\vec{\eta})$$
$$+\lambda \sum_{y \in \{\eta\}} \big(\eta_y + 1\big)\, \overline{\nu}(\overline{\eta}^{+y}) d\nu_{\overline{\eta}^{+y}}(\vec{\eta}) + \lambda \int_{z \in \mathbb{T}\setminus\{\eta\}} \overline{\nu}(\overline{\eta}^{+z}) d\nu_{\overline{\eta}^{+z}}(\vec{\eta}^{+z})$$

$$+b\rho \sum_{y:\eta_y=1} \overline{\nu}(\overline{\eta}^{-y}) \sum_{y'\in\{\eta\}\setminus\{y\}} \eta_{y'} g(y',y) d\nu_{\overline{\eta}^{-y}}(\vec{\eta}^{-y}) d\sigma(y)$$
$$-(\lambda+b) \sum_{y\in\{\eta\}} \eta_y \overline{\nu}(\overline{\eta}) d\nu_{\overline{\eta}}(\vec{\eta}).$$

In this formula if $\overline{\eta} = (q_1, \ldots, q_k)$, $\vec{\eta} = (y_1, \ldots, y_k)$ then $\overline{\eta}^{-y}$, $\vec{\eta}^{-y}$, for $y \in \{\eta\}$, and $\overline{\eta}^{+z}$, $\vec{\eta}^{+z}$, for $z \notin \{\eta\}$, are defined as

$$\overline{\eta}^{-y} = \begin{cases} (q_1, \ldots, q_{i-1}, q_i - 1, q_{i+1}, \ldots, q_k) & \text{if } y_i = y, q_i > 1 \\ (q_1, \ldots, q_{i-1}, q_{i+1}, \ldots, q_k) & \text{if } y_i = y, q_i = 1 \end{cases}$$
$$\vec{\eta}^{-y} = \begin{cases} \vec{\eta} & \text{if } y_i = y, q_i > 1 \\ (y_1, \ldots, y_{i-1}, y_{i+1}, \ldots, y_k) & \text{if } y_i = y, q_i = 1 \end{cases}$$
$$\overline{\eta}^{+z} = (q_1, \ldots, q_i, 1, q_{i+1}, \ldots, q_k) \quad \text{if } y_i < z < y_{i+1}$$
$$\vec{\eta}^{+z} = (y_1, \ldots, y_i, z, y_{i+1}, \ldots, y_k) \quad \text{if } y_i < z < y_{i+1}$$

Since $\theta > 0$, we get

$$(\lambda+b) \sum_{y\in\{\eta\}} \eta_y \overline{\nu}(\overline{\eta}) d\nu_{\overline{\eta}}(\vec{\eta}) \geq \lambda \int_{z\in\mathbb{T}\setminus\{\eta\}} \overline{\nu}(\overline{\eta}^{+z}) d\nu_{\overline{\eta}^{+z}}(\vec{\eta}^{+z})$$
$$+b\rho \sum_{y:\eta_y=1} \overline{\nu}(\overline{\eta}^{-y}) \sum_{y'\in\{\eta\}\setminus\{y\}} \eta_{y'} g(y',y) d\nu_{\overline{\eta}^{-y}}(\vec{\eta}^{-y}) d\sigma(y). \tag{28}$$

For any integer $m \geq 1$, we denote by $(1^m) \in \Sigma(\mathbb{N})$ the discrete configuration

$$(1^m) = \Big(\underbrace{1, \ldots, 1}_{m \text{ times}} \Big)$$

From inequality (28) (keeping only the second term on the right hand side), we get for any integer $m \geq 2$ and some constant $C_m > 0$

$$d\nu_{(1^m)}(y_1, \ldots, y_m) \geq C_m g(y_{m-1}, y_m) d\sigma(y_m) d\nu_{(1^{m-1})}(y_1, \ldots, y_{m-1}).$$

Using this inequality recursively, we obtain for any integer $m \geq 2$

$$d\nu_{(1^m)}(y_1, \ldots, y_m) \geq f_1(y_1) d\sigma(y_1) \prod_{j=1}^{m-1} C_{j+1} g(y_j, y_{j+1}) d\sigma(y_{j+1}). \tag{29}$$

We also derive from inequality (28), using only the first term on the right hand side, that for any integer $m \geq 1$, there exists a constant $C'_m > 0$ such that

$$d\nu_{(1^m)}(y_1, \dots, y_m) \geq C'_m \int\limits_{z \in \mathbb{T} \setminus \{y_1, \dots, y_m\}} d\nu_{(1^{m+1})}(y_1, z, y_2, \dots, y_m).$$

Using this inequality recursively, we get for any integer $m > 2$

$$\begin{aligned} &d\nu_{(1^2)}(x_1, x_2) \\ &\quad \geq \prod_{j=2}^{m-1} C'_j \int\limits_{\mathbb{T} \setminus \{x_1, x_2\}} \int\limits_{\mathbb{T} \setminus \{x_1, z_1, x_2\}} \dots \int\limits_{\mathbb{T} \setminus \{x_1, z_1, \dots, z_{m-2}, x_2\}} d\nu_{(1^m)}(x_1, z_1, \dots, z_{m-2}, x_2). \end{aligned}$$

Combining this inequality with (29), we get for any $m > 2$, and for some constant $C''_m > 0$

$$d\nu_{(1^2)}(x_1, x_2) \geq C''_m f_1(x_1) g_{m-1}(x_1, x_2) d\sigma(x_1) d\sigma(x_2). \tag{30}$$

It follows directly from (28) that the same inequality holds for $m = 2$.

We also derive from inequality (28) that for some constant $C_1 > 0$ we have

$$\begin{aligned} d\nu_{(1)}(x) &\geq C_1 \int\limits_{\mathbb{T} \setminus \{x\}} d\nu_{(1^2)}(z, x) \\ &\geq C_1 \, C''_m d\sigma(x) \int\limits_{\mathbb{T}} f_1(z) g_{m-1}(z, x) d\sigma(z). \end{aligned}$$

where for the last inequality, which holds for any $m > 2$, we use (30). This implies immediately that for σ almost every x

$$f_1(x) \geq C_1 C''_m \int\limits_{\mathbb{T}} f_1(z) g_{m-1}(z, x) d\sigma(z).$$

We now choose $m = k_0 + 1$ (k_0 is given in (9)) therefore $g_{m-1}(z, x) > 0, \sigma \otimes \sigma$ almost surely, and from the above inequality we get $f_1 > 0, \sigma$ almost surely. Therefore, $f_1(x_1) g_{m-1}(x_1, x_2) > 0, \sigma \otimes \sigma$ almost surely, and the result follows. □

Proposition 6.3 *There is a unique q.s.d. associated with the exponential decay rate $\theta = \lambda - b$.*

Proof Let ν and ν' be two different q.s.d. with the exponential decay rate θ. We can write the Lebesgue decomposition

$$\nu' = f\nu + \xi$$

with f a nonnegative measurable function and ξ a singular measure with respect to ν. Assume $\xi \neq 0$. Applying $R_t^\dagger$ we get

$$\nu' = f\nu + \xi = R_t^\dagger(f\nu) + R_t^\dagger(\xi).$$

If f is bounded, since ν is a q.s.d., we have $R_t^\dagger(f\nu) \ll \nu$. In the general case, the same result holds by approximating f by an increasing sequence of nonnegative functions. Therefore, we must have

$$R_t^\dagger(f\nu) \leq f\nu.$$

Integrating the function φ_1 as before, we conclude that $R_t^\dagger(f\nu) = f\nu$, and therefore $R_t^\dagger(\xi) = \xi$.

Then, we have two q.s.d. ν and ξ with exponential decay rate $\theta = \lambda - b$, which are mutually singular. We claim that this is excluded by Lemma 6.2. Indeed let B be a measurable subset such that $\xi(B) = \nu(B^c) = 0$. Then $\nu(B \cap \mathcal{A}_{(11)}) = \nu(\mathcal{A}_{(11)}) > 0$. Since $\nu \ll \mu$ we get $\mu(B) \geq \mu(B \cap \mathcal{A}_{(11)}) > 0$. From Lemma 6.2 we deduce $\xi(B \cap \mathcal{A}_{(11)}) > 0$ which is a contradiction. Namely $\xi = 0$ and we conclude that $\nu' = f\nu$. Let us now show that $f \equiv 1$, which will yield $\nu' = \nu$ and so will conclude the uniqueness result.

Multiplying if necessary by a positive constant, we can assume than ν and ν' are probability measures. In particular, the integral of f with respect to ν is equal to 1. The measure $\nu - \nu' = (1 - f)\nu$ is also invariant by the semi-group $R_t^\dagger$. Therefore, denoting by $(1 - f)^+$ the positive part of $1 - f$, we have

$$R_t^\dagger((1-f)^+\nu) = (1-f)^+\nu + R_t^\dagger((f-1)^+\nu) - (f-1)^+\nu.$$

Since $R_t^\dagger((f-1)^+\nu)$ is a nonnegative measure, this implies the relations

$$R_t^\dagger((1-f)^+\nu)(\bullet \cap \{f \leq 1\}) \geq (1-f)^+\nu(\bullet \cap \{f \leq 1\})$$
$$R_t^\dagger((1-f)^+\nu)(\bullet \cap \{f > 1\}) \geq 0 = (1-f)^+\nu(\bullet \cap \{f > 1\}),$$

showing that

$$R_t^\dagger((1-f)^+\nu) \geq (1-f)^+\nu.$$

Since φ_1 is invariant by R_t we get

$$0 = \int \varphi_1 d\left(R_t^\dagger((1-f)^+\nu) - (1-f)^+\nu\right).$$

Using that φ_1 is strictly positive we deduce

$$R_t^\dagger((1-f)^+\nu) = (1-f)^+\nu.$$

We also get

$$R_t^\dagger((f-1)^+\nu) = (f-1)^+\nu.$$

If f is not almost surely equal to one, since its integral with respect to ν is equal to one, the two measures $(1-f)^+\nu$ and $(f-1)^+\nu$ are two non-trivial nonnegative measures, invariant by R_t. Moreover, they are mutually singular, but we saw above that this is impossible. This contradiction ensures that $f = 1$ almost surely and completes the proof of the proposition. □

The proof of Theorem 1.3 is now complete, it follows from Propositions 6.1 and 6.3.

Acknowledgments The authors are grateful to two anonymous referees for their comments, suggestions and corrections that greatly helped to improve the presentation of this article. The authors acknowledge the partial support given by the Millennium Nucleus Information and Randomness P04-069-F, CMM FONDAP and CMM BASAL projects. S. Martínez thanks Guggenheim Fellowship and the hospitality of Ecole Polytechnique, Palaiseau, and Sylvie Méléard the ECOS-CONICYT project.

References

1. Billingsley, P.: Convergence of probability measures. Wiley, New York (1968)
2. Cattiaux, P., Collet, P., Lambert, A., Martinez, S., Méléard, S., San Martín, J.: Quasi-stationary distributions and diffusions models in population dynamics. Ann. Probab. **37**(5), 1926–1969 (2009)
3. Cattiaux, P., Méléard, S.: Competitive or weak cooperative stochastic Lotka–Volterra systems conditioned on non-extinction. J. Math. Biol. **6**, 797–829 (2010)
4. Collet, P., Martínez, S., San Martín, J.: Asymptotic laws for one-dimensional diffusions conditioned to non-absorption. Ann. Probab. **23**(3), 1300–1314 (1995)
5. Champagnat, N., Ferrière, R., Méléard, S.: Unifying evolutionary dynamics: from individual stochastic processes to macroscopic models. Theoret. Pop. Biol. **69**, 297–321 (2006)
6. Darroch, J.N., Seneta, E.: On quasi-stationary distributions in absorbing discrete-time finite Markov chains. J. Appl. Prob. **2**, 88–100 (1965)
7. Daley, D.J., Vere-Jones, D.: An Introduction to the theory of point processes. In: Springer Series in Statistics. Springer, New York (1988)
8. Dunford, N., Schwartz, J.: Linear Operators, vol. I. Interscience (1958)
9. Feller, W.: An Introduction to probability Theory and its Applications. Wiley, New York (1968)
10. Ferrari, P.A., Kesten, H., Martinez, S., Picco, P.: Existence of quasi-stationary distributions. A renewal dynamical approach. Ann. Probab. **23**(2), 501–521 (1995)
11. Fournier, N., Méléard, S.: A microscopic probabilistic description of a locally regulated population and macroscopic approximations. Ann. Appl. Probab. **14**(4), 1880–1919 (2004)
12. Gosselin, F.: Asymptotic behavior of absorbing Markov chains conditional on nonabsorption for applications in conservation biology. Ann. Appl. Probab. **11**(1), 261–284 (2001)
13. Mandl, P.: Spectral theory of semi-groups connected with diffusion processes and its applications. Czech. Math. J. **11**, 558–569 (1961)
14. Martínez, S., San Martín, J.: Classification of killed one-dimensional diffusions. Ann. Probab. **32**(1), 530–552 (2004)
15. Pinsky, R.: On the convergence of diffusion processes conditioned to remain in bounded region for large time to limiting positive recurrent diffusion processes. Ann. Probab. **13**, 363–378 (1985)
16. Nair, M.G., Pollett, P.: On the relationship between $\mu-$invariant measures and quasi-stationary distributions for continuous-time Markov chains. Adv. Appl. Probab. **25**(1), 82–102 (1993)
17. Renault, O., Ferrière, R., Porter, J.: The quasi-stationary route to extinction. Private communication
18. Steinsaltz, D., Evans, S.: Markov mortality models: implications of quasi-stationarity and varying initial distributions. Theor. Popul. Biol. **65**, 319–337 (2004)
19. Steinsaltz, D., Evans, S.: Quasi-stationary distributions for one-dimensional diffusions with killing. Trans. Am. Math. Soc. **359**(3), 1285–1324 (2007)

20. Tychonov, A.: Ein Fixpunktsatz. Math. Ann. **117**, 767–776 (1935)
21. Van Doorn, E.: Quasi-stationary distributions and convergence to quasi-stationary of birth-death processes. Adv. Appl. Probab. **23**(4), 683–700 (1991)
22. Vere-Jones, D.: Geometric ergodicity in denumerable Markov chains. Quart. J. Math. Oxf. (2) **13**, 7–28 (1962)
23. Yaglom, A.M.: Certain limit theorems of the theory of branching processes (in Russian). Dokl. Akad. Nauk. SSSR **56**, 795–798 (1947)
24. Yosida, K.: Functional Analysis. Springer, Berlin (1968)