A Microscopic Interpretation for Adaptive Dynamics Trait Substitution Sequence Models

Nicolas Champagnat*

October 29, 2004

Abstract

We consider an interacting particle Markov process for Darwinian evolution in an asexual population, involving a linear birth rate, a density-dependent logistic death rate, and a probability μ of mutation at each birth event. We introduce a renormalization parameter K scaling the size of the population, which leads, when $K \to +\infty$ and $\mu \equiv 0$, to a deterministic dynamics for the density of individuals holding a given trait. By combining in a non-standard way the limits of large population ($K \to +\infty$) and of small mutations ($\mu \to 0$), we prove that a time scale separation between the birth and death events and the mutation events occurs and that the interacting particle microscopic process converges for finite dimensional distributions to the biological model of evolution known as the "monomorphic trait substitution sequence" model of adaptive dynamics [14, 4], which describes the Darwinian evolution in an asexual population as a Markov jump process in the trait space.

Keywords: measure-valued process; interacting particle system; adaptive dynamics; finite dimensional distributions convergence; time scale separation; stochastic domination; branching processes; large deviations.

1 Introduction

We will study in this article the link between two biological models of Darwinian evolution in an asexual population. The first one is a system of interacting particles modeling evolution at the *individual* level, referred below as the *microscopic model*, and which has been already studied in Dieckmann [3], Fournier and Méléard [8] and Ferrière *et al* [7]. The second one models the evolution at the *population* level as a jump Markov process in the space of traits characterizing individuals, called "trait substitution sequence", and referred below as the *TSS model*. The TSS model belongs to the recent biological theory of evolution called *adaptive dynamics* (Hofbauer and Sigmund [10], Marrow *et al.* [12] and Metz *et al.* [13]), and has been introduced in 1996 by Metz *et al.* [14] and Dieckmann and Law [4], and mathematically studied in Champagnat [2]. This model

*Equipe MODAL'X, Bât. G, Université Paris 10 – Nanterre, 200 avenue de la République, 92001 Nanterre Cedex, France

Mathematical Eco-Evolutionary Theory Group, Laboratoire d'Ecologie, Ecole Normale Supérieure, 46 rue d'Ulm, 75230 Paris Cedex 05, France

e-mail: `champagn@clipper.ens.fr`

and its generalizations have revealed a powerful tool for understanding various evolutionary phenomena, such as polymorphism (stable coexistence of different traits) or evolutionary branching (evolution of a monomorphic population to a polymorphic one). The heuristic leading to the TSS model [14, 4] is actually based on the biological assumptions of large population and rare mutations, and on another assumption stating that no two different types of individuals can co-exist on a long time scale: the competition eliminates one of them. We propose to prove in this article a convergence result of the microscopic model to this TSS model when the parameters are normalized in a non-standard way, leading to a time scale separation. This limit combines a *large population* asymptotic with a *rare mutations* asymptotic. Such a result will provide a mathematical justification of the TSS model and of the biological heuristic on which it is based.

Let us describe the microscopic model: in a population, Darwinian evolution acts on a set of phenotypes, or *traits*, characterizing each individual's ability to survive and reproduce (e.g. body size, rate of food intake, age at maturity). We will consider a finite number of quantitative traits in an asexual population (clonal reproduction), and we will assume that the trait space $\mathcal{X}$ is a compact subset of $\mathbb{R}^l$ ($l \geq 1$).

The microscopic model involves the three basic mechanisms of Darwinian evolution: *heredity*, which transmits traits to new offsprings, *mutation*, driving a variation in the trait values in the population, and *selection* between these different trait values, caused by the competition for limited resources or area.

For any $x, y \in \mathcal{X}$, we introduce the following biological parameters

$b(x) \in \mathbb{R}_+^*$ is the rate of birth from an individual holding trait x.

$d(x) \in \mathbb{R}_+^*$ is the rate of "natural" death for an individual holding trait x.

$\alpha(x, y) \in \mathbb{R}_+^*$ is competition kernel representing the pressure felt by an individual holding trait x from an individual holding trait y.

$\mu(x) \in [0, 1]$ is the probability that a mutation occurs in a birth from an individual with trait x.

$m(x, dh)$ is the law of $h = y - x$, where the mutant trait y is born from an individual with trait x. It is a probability measure on $\mathbb{R}^l$, and since y must belong to the trait space $\mathcal{X}$, the support of $m(x, \cdot)$ is a subset of

$$\mathcal{X} - x = \{y - x : y \in \mathcal{X}\}.$$

$K \in \mathbb{N}$ is a parameter rescaling the competition kernel $\alpha(\cdot, \cdot)$. Biologically, K may be linked to the ressources or area available, and is related to the biological concept of "carrying capacity". As will appear later, this parameter is linked to the size of the population: large K means a large population (provided that the initial condition is proportional to K).

$u_K \in [0, 1]$ is a parameter depending on K rescaling the probability of mutation $\mu(\cdot)$. Small u_K means rare mutations.

Let us also introduce the following notations, used throughout this paper :

$$\bar{n}_x = \frac{b(x) - d(x)}{\alpha(x, x)}, \tag{1}$$

$$\beta(x) = \mu(x) b(x) \bar{n}_x \tag{2}$$

$$\text{and} \quad f(y, x) = b(y) - d(y) - \alpha(y, x) \bar{n}_x. \tag{3}$$

We consider, at any time $t \geq 0$, a finite number N_t of individuals, each of them holding a trait value in $\mathcal{X}$. Let us denote by $x_1, \dots, x_{N_t}$ the trait values of these individuals. The state of the population, rescaled by K, at time $t \geq 0$ can be represented by the finite point measure on $\mathcal{X}$

$$\nu_t^K = \frac{1}{K} \sum_{i=1}^{N_t} \delta_{x_i},$$

where δ_x is the Dirac measure at x. Let $\mathcal{M}_F$ denote the set of finite nonnegative measures on $\mathcal{X}$, and define

$$\mathcal{M}^K = \left\{ \frac{1}{K} \sum_{i=1}^{n} \delta_{x_i} : n \geq 0,\ x_1, \dots, x_n \in \mathcal{X} \right\},$$

An individual holding trait x in the population ν_t^K gives birth to another individual with rate $b(x)$ and dies with rate $d(x) + \int \alpha(x,y)\nu_t^K(dy) = d(x) + (\sum_{i=1}^{N_t} \alpha(x, x_i))/K$.

A newborn holds the same trait value as its progenitor's with probability $1 - u_K\mu(x)$, and with probability $u_K\mu(x)$, the newborn is a mutant which trait value y is chosen according to $y = x + h$, where h is a random variable with law $m(x, dh)$.

In other words, the process $(\nu_t^K, t \geq 0)$ is a $\mathcal{M}^K$-valued Markov process with infinitesimal generator defined for any bounded measurable functions ϕ from $\mathcal{M}^K$ to $\mathbb{R}$ by

$$\begin{aligned}
L^K\phi(\nu) = &\int_{\mathcal{X}} \left(\phi\left(\nu + \frac{\delta_x}{K}\right) - \phi(\nu) \right) (1 - u_K\mu(x))b(x)K\nu(dx) \\
&+ \int_{\mathcal{X}} \int_{\mathbb{R}^l} \left(\phi\left(\nu + \frac{\delta_{x+h}}{K}\right) - \phi(\nu) \right) u_K\mu(x)b(x)m(x,dh)K\nu(dx) \\
&+ \int_{\mathcal{X}} \left(\phi\left(\nu - \frac{\delta_x}{K}\right) - \phi(\nu) \right) \left(d(x) + \int_{\mathcal{X}} \alpha(x,y)\nu(dy) \right) K\nu(dx). \quad (4)
\end{aligned}$$

The first term (linear) describes the births without mutation, the second term (linear) describes the births with mutation, and the third term (non-linear) describes the deaths by oldness or competition. This logistic density-dependence models the competition in the population, and hence drives the selection process.

Let us denote by (A) the following three assumptions

(A1) b, d and α are measurable functions, and there exists $\bar{b}, \bar{d}, \bar{\alpha} < +\infty$ such that

$$b(\cdot) \leq \bar{b}, \quad d(\cdot) \leq \bar{d} \quad \text{and} \quad \alpha(\cdot,\cdot) \leq \bar{\alpha}.$$

(A2) $m(x, dh)$ is absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}^l$ with density $m(x,h)$, and there exists a function $m : \mathbb{R}^l \to \mathbb{R}_+$ such that $m(x,h) \leq m(h)$ for any $x \in \mathcal{X}$ and $h \in \mathbb{R}^l$, and $\int m(h)dh < \infty$.

(A3) $\mu(x) > 0$ and $b(x) - d(x) > 0$ for any $x \in \mathcal{X}$, and there exists $\underline{\alpha} > 0$ such that

$$\underline{\alpha} \leq \alpha(\cdot,\cdot).$$

For fixed K, under (A1) and (A2) and assuming that $\mathbf{E}(\langle \nu_0^K, \mathbf{1}\rangle) < \infty$ (where $\langle \nu, f\rangle$ denotes the integral of the measurable function f with respect to the measure ν), the existence and uniqueness in law of a process with infinitesimal generator L^K has been proved by Fournier and Méléard [8].

In this model, the biological assumption of large population corresponds to the limit $K \to +\infty$, and the assumption of rare mutations to $u_K \to 0$. In order to give a precise formulation of the biological assumption on the impossibility of coexistence of two different traits, let us define:

Definition 1

(a) *For any $K \geq 1$, $b, d, c \geq 0$ and for any $\mathbb{N}/K$-valued random variable z, we will denote by $\mathbf{P}^K(b, d, c, z)$ the law of the $\mathbb{N}/K$-valued Markov jump process with initial state z and with transition rates*

$$\begin{array}{ll} ib & \text{from } i/K \text{ to } (i+1)/K, \\ i(d + ci/K) & \text{from } i/K \text{ to } (i-1)/K. \end{array}$$

(b) *For any $K \geq 1$, $b_k, d_k, c_{kl} \geq 0$ with $k, l \in \{1, 2\}$, and for any $\mathbb{N}/K$-valued random variables z_1 and z_2, we will denote by*

$$\mathbf{Q}^K(b_1, b_2, d_1, d_2, c_{11}, c_{12}, c_{21}, c_{22}, z_1, z_2)$$

the law of the $(\mathbb{N}/K)^2$-valued Markov jump process with initial state (z_1, z_2) and with transition rates

$$\begin{array}{ll} ib_1 & \text{from } (i/K, j/K) \text{ to } ((i+1)/K, j/K), \\ jb_2 & \text{from } (i/K, j/K) \text{ to } (i/K, (j+1)/K), \\ i(d_1 + c_{11}i/K + c_{12}j/K) & \text{from } (i/K, j/K) \text{ to } ((i-1)/K, j/K), \\ j(d_2 + c_{21}i/K + c_{22}j/K) & \text{from } (i/K, j/K) \text{ to } (i/K, (j-1)/K). \end{array}$$

Fix x and y in $\mathcal{X}$. The proof of the following two results can be found in chap. 11 of Ethier and Kurtz [6].

Proposition 1

(a) *Assume $\mu \equiv 0$ and $\nu_0^K = N_x^K(0)\delta_x$. Then, for any $t \geq 0$, $\nu_t^K = N_x^K(t)\delta_x$, where N_x^K has the law $\mathbf{P}^K(b(x), d(x), \alpha(x, x), N_x^K(0))$. Assume $N_x^K(0) \to n_x(0)$ in probability when $K \to +\infty$. Then, the sequence (N_x^K) converges in probability on $[0, T]$ for the uniform norm to the deterministic function n_x with initial condition $n_x(0)$ solution to*

$$\dot{n}_x = (b(x) - d(x) - \alpha(x, x)n_x)n_x. \tag{5}$$

(b) *Assume $\mu \equiv 0$ and $\nu_0^K = N_x^K(0)\delta_x + N_y^K(0)\delta_y$. Then, for any $t \geq 0$, $\nu_t^K = N_x^K(t)\delta_x + N_y^K(t)\delta_y$, where (N_x^K, N_y^K) has the law*

$$\mathbf{Q}^K(b(x), b(y), d(x), d(y), \alpha(x, x), \alpha(x, y), \alpha(y, x), \alpha(y, y), N_x^K(0), N_y^K(0)).$$

Assume $N_x^K(0) \to n_x(0)$ and $N_y^K(0) \to n_y(0)$ in probability when $K \to +\infty$. Then, (N_x^K, N_y^K) converges in probability when $K \to +\infty$ on $[0, T]$ for the uniform norm to the deterministic function (n_x, n_y) with initial condition $(n_x(0), n_y(0))$ solution to

$$\begin{cases} \dot{n}_x = (b(x) - d(x) - \alpha(x, x)n_x - \alpha(x, y)n_y)n_x \\ \dot{n}_y = (b(y) - d(y) - \alpha(y, x)n_x - \alpha(y, y)n_y)n_y. \end{cases} \tag{6}$$

Note that (5) has two steady states, 0, unstable, and $\bar{n}_x$, defined in (1), stable. System (6) has at least three steady states, $(0,0)$, unstable, $(\bar{n}_x, 0)$ and $(0, \bar{n}_y)$.

Let us introduce the following assumption :

(B) Given any $x \in \mathcal{X}$, Lebesgue almost any $y \in \mathcal{X}$ satisfies one of the two following conditions:

$$\text{either} \quad (b(y) - d(y))\alpha(x,x) - (b(x) - d(x))\alpha(y,x) < 0, \tag{7}$$

$$\text{or} \quad \begin{cases} (b(y) - d(y))\alpha(x,x) - (b(x) - d(x))\alpha(y,x) > 0, \\ (b(x) - d(x))\alpha(y,y) - (b(y) - d(y))\alpha(x,y) < 0. \end{cases} \tag{8}$$

Assumption (B) is the mathematical formulation of the impossibility of co-existence of two different traits. Actually, an elementary analysis of system (6) shows that (cf. e.g. Istas [11] p. 25–27) :

Proposition 2 *If x and y satisfy (7), then $(\bar{n}_x, 0)$ is a stable steady state of (6). If x and y satisfy (8), then $(\bar{n}_x, 0)$ is an unstable steady state of (6), $(0, \bar{n}_y)$ is stable, and any solution to (6) with initial state in $(\mathbb{R}_+^*)^2$ converges to $(0, \bar{n}_y)$ when $t \to +\infty$.*

The TSS model is a Markov jump process in $\mathcal{X}$ with infinitesimal generator given, for any bounded measurable function φ from $\mathcal{X}$ to $\mathbb{R}$, by

$$A\varphi(x) = \int_{\mathbb{R}^l} (\varphi(x+h) - \varphi(x))\beta(x)\frac{[f(x+h,x)]_+}{b(x+h)} m(x,h)dh, \tag{9}$$

where $[a]_+$ denotes the positive part of $a \in \mathbb{R}$, and where $\beta(x)$ and $f(y,x)$ are defined in (2) and (3). The existence and uniqueness in law of a process generated by A holds as soon as $\beta(x)[f(y,x)]_+/b(y)$ is bounded (see e.g. Ethier and Kurtz [6]), which is true under assumption (A) ($[f(y,x)]_+/b(y) \leq 1$).

Our main result writes:

Theorem 1 *Assume (A) and (B). Fix a sequence $(u_K)_{K\in\mathbb{N}}$ in $[0,1]^{\mathbb{N}}$ such that*

$$\forall V > 0, \quad \exp(-VK) \ll u_K \ll \frac{1}{K \log K} \tag{10}$$

(where $f(K) \ll g(K)$ means that $f(K)/g(K) \to 0$ when $K \to \infty$). Fix also $x \in \mathcal{X}$, $\gamma > 0$ and a sequence of $\mathbb{N}$-valued random variables bounded in $\mathbb{L}^1$, $(\gamma_K)_{K\in\mathbb{N}}$, such that γ_K/K converges in law to γ. Consider the process $(\nu_t^K, t \geq 0)$ generated by (4) with initial state $\frac{\gamma_K}{K}\delta_x$. Then, for any $n \geq 1$, $\varepsilon > 0$ and $0 < t_1 < t_2 < \ldots < t_n < \infty$, and for any measurable subsets $\Gamma_1, \ldots, \Gamma_n$ of $\mathcal{X}$,

$$\lim_{K\to+\infty} \mathbf{P}\big(\forall i \in \{1,\ldots,n\},\ \mathrm{Supp}(\nu^K_{t_i/Ku_K}) \text{ is a singleton } \{x_i\},\ x_i \in \Gamma_i \\ \text{and } |\langle \nu^K_{t_i/Ku_K}, \mathbf{1}\rangle - \bar{n}_{x_i}| < \varepsilon\big) = \mathbf{P}(\forall i \in \{1,\ldots,n\},\ X_{t_i} \in \Gamma_i) \tag{11}$$

where for any $\nu \in \mathcal{M}_F$, $\mathrm{Supp}(\nu)$ is the support of ν and $(X_t, t \geq 0)$ is the TSS process generated by (9) with initial state x.

Corollary 1 *With the same notations and assumptions than in Theorem 1, when $K \to +\infty$, the process $(\nu^K_{t/Ku_K}, t \geq 0)$ converges in the sense of the finite dimensional distributions, for the topology on $\mathcal{M}_F$ induced by the functions $\nu \mapsto \langle \nu, f \rangle$ with f bounded and measurable on $\mathcal{X}$, to the process $(Y_t, t \geq 0)$ defined by*

$$Y_t = \begin{cases} \gamma\delta_x & \text{if } t = 0 \\ \bar{n}_{X_t}\delta_{X_t} & \text{if } t > 0. \end{cases}$$

Proof of Corollary 1 Let Γ be a measurable subset of $\mathcal{X}$. Let us prove that

$$\lim_{K\to+\infty} \mathbf{E}(\langle \nu^K_{t/Ku_K}, \mathbf{1}_\Gamma \rangle) = \mathbf{E}(\bar{n}_{X_t}\mathbf{1}_{X_t\in\Gamma}). \tag{12}$$

Fix $\varepsilon > 0$, and observe that $\bar{n}_x \in [0, \bar{b}/\underline{\alpha}]$. Write $[0, \bar{b}/\underline{\alpha}] \subset \cup_{i=1}^p I_i$, where p is the integer part of $\bar{b}/\varepsilon\underline{\alpha}$, and $I_i = [(i-1)\varepsilon, i\varepsilon[$. Define $\Gamma_i = \{x \in \mathcal{X} : \bar{n}_x \in I_i\}$ for $1 \leq i \leq p$, and apply (11) to the sets $\Gamma \cap \Gamma_1, \ldots, \Gamma \cap \Gamma_p$ with $n = 1$, $t_1 = t$ and the constant ε above. Then, for sufficiently large K,

$$\begin{aligned}
\limsup_{K\to+\infty} \mathbf{E}(\langle \nu^K_{t/Ku_K}, \mathbf{1}_\Gamma \rangle) &\leq \sum_{i=1}^p \limsup_{K\to+\infty} \mathbf{E}(\langle \nu^K_{t/Ku_K}, \mathbf{1}_{\Gamma\cap\Gamma_i} \rangle) \\
&\leq \sum_{i=1}^p (i+1)\varepsilon \mathbf{P}(X_t \in \Gamma \cap \Gamma_i) \\
&\leq \sum_{i=1}^p \big(\mathbf{E}(\bar{n}_{X_t}\mathbf{1}_{X_t\in\Gamma\cap\Gamma_i}) + 2\varepsilon\mathbf{P}(X_t \in \Gamma_i)\big) \\
&\leq \mathbf{E}(\bar{n}_{X_t}\mathbf{1}_{X_t\in\Gamma}) + 2\varepsilon.
\end{aligned}$$

A similar estimate for the lim inf ends the proof of (12), which implies the convergence of one-dimensional laws for the required topology.

The same method gives easily the required limit when we consider a finite number of times $t_1, \ldots, t_n$. □

Remarks 1

- *The time scale $1/Ku_K$ of Theorem 1 is the time scale of the mutation events for the process ν^K. Assumption (10) is the condition leading to the time scales separation between the mutation events and the birth and death events. The limit (11) means that, when this time scales separation occurs, the population is monomorphic at any time (*i.e. *composed of individuals holding the* same *trait value) with high probability, and that the transition periods from a resident trait to a mutant one are infinitesimal on this mutation time scale.*

- *It is not possible to obtain the convergence in law for the Skorohod topology on $\mathbb{D}([0,T], \mathcal{M}_F)$ because of the right discontinuity of the process Y at time 0^+.*

This result is different from usual time scale separation results (averaging principle, cf. Freidlin and Wentzell [9] and Skorohod *et al.* [16]), because no assumption of ergodicity has been made, and because we have to combine two limits simultaneously. Original methods are necessary to prove Theorem 1.

Our proof is based on two ingredients : first, when $\mu \equiv 0$ and ν^K_0 is monomorphic with trait x, we have seen in Proposition 1 (a) the convergence of

ν^K to $n(t)\delta_x$, where $n(t)$ is solution to (5). Any solution to this equation with positive initial condition converges for large time to $\bar{n}_x$. The *large deviations* estimates for this convergence will allow us to show that the time during which the stochastic process stays in a neighborhood of its limit (problem of exit from domain [9]) is of the order of $\exp(KV)$ with $V > 0$.

Now, when u_K is small, the process ν^K with a monomorphic initial condition with trait x is near to the same process with $\mu \equiv 0$, as long as no mutation occurs. Therefore, the left inequality of (10) will allow us to prove that, with high probability, ν^K is near to $\bar{n}_x\delta_x$ when the first mutation occurs.

The second ingredient of our proof is the study of the invasion of a mutant trait y that has just appeared in a monomorphic population with trait x. This invasion can be divided in three steps :

- Firstly, as long as the mutant population size $\langle \nu_t^K, \mathbf{1}_{\{y\}}\rangle$ (initially equal to $1/K$) is smaller than a fixed small $\varepsilon > 0$, the resident dynamics is very close to what it was before the mutation, so $\langle \nu_t^K, \mathbf{1}_{\{x\}}\rangle$ stays close to $\bar{n}_x$. Then, the death rate of a mutant individual is close to the constant $d(y) + \alpha(y,x)\bar{n}_x$. Since its birth rate is constant, equal to $b(y)$, we can approximate the mutant dynamics by a binary branching process. Therefore, the probability that $\langle \nu_t^K, \mathbf{1}_{\{y\}}\rangle$ reaches ε is approximately equal to the probability that this branching process reaches εK, which converges when $K \to +\infty$ to its probability of non-extinction $[f(y,x)]_+/b(y)$.

- Secondly, once $\langle \nu_t^K, \mathbf{1}_{\{y\}}\rangle$ has reached ε, by Proposition 1 (b), for large K, ν^K is close to the solution to (6) with initial state $(\bar{n}_x, \varepsilon)$ with high probability. We will show that Proposition 2 implies that this solution reaches the ε-neighborhood of $(0, \bar{n}_y)$ in finite time.

- Finally, once $\langle \nu_t^K, \mathbf{1}_{\{y\}}\rangle$ is close to $\bar{n}_y$ and $\langle \nu_t^K, \mathbf{1}_{\{x\}}\rangle$ is small, $K\langle \nu_t^K, \mathbf{1}_{\{x\}}\rangle$ can be approximated, in a similar way than in the first step, by a binary branching process, which is subcritical and hence gets extinct a.s. in finite time.

We will see in sections 2.2 and 2.3 that the time needed to complete the first and third steps is proportional to $\log K$, whereas the time needed for the second step is bounded. Therefore, since the time between two mutations is of the order of $1/Ku_K$, the right inequality in (10) will allow us to prove that, with high probability, the three steps above are completed before a new mutation occurs.

Section 2 will provide the large deviations and branching process results needed to make formal the preceding heuristics. We will also prove several comparison results between $\langle \nu_t^K, \mathbf{1}\rangle$ and the birth and death processes of Definition 1. In section 3, the proof of Theorem 1 is achieved by computing, for any t, the limit law of ν^K_{t/Ku_K} according to the random number of mutations having occured between 0 and t/Ku_K.

Notations

- $\lceil a \rceil$ denotes the first integer greater or equal to a, and $\lfloor a \rfloor$ denotes the integer part of a.

- For any $K \geq 1$ and $\nu \in \mathcal{M}^K$, we will denote by $\mathbf{P}_\nu^K$ the law of the process ν^K generated by (4) with initial state ν, and by $\mathbf{E}_\nu^K$ the expectation with respect to $\mathbf{P}_\nu^K$.

- The convergence in probability of finite dimensional random variables will be denoted by $\xrightarrow{\mathcal{P}}$.
- We will denote by $\mathcal{L}(Z)$ the law of the stochastic process $(Z_t, t \geq 0)$.
- We will denote by $\preceq$ the following stochastic domination relation: if $\mathbf{Q}_1$ and $\mathbf{Q}_2$ are the laws of $\mathbb{R}$-valued processes, we will write $\mathbf{Q}_1 \preceq \mathbf{Q}_2$ if we can construct on the same probability space $(\Omega, \mathcal{F}, \mathbf{P})$ two processes X^1 and X^2 such that $\mathcal{L}(X^i) = \mathbf{Q}_i$ $(i = 1, 2)$ and $\forall t \geq 0$, $\forall \omega \in \Omega$, $X^1_t(\omega) \leq X^2_t(\omega)$.
- Finally, if X^1 and X^2 are two random processes and T is a stopping time for X^1, we will write $X^1 \preceq X^2$ for $t \leq T$ if we can construct a process $\hat{X}^2$ on the probability space on which X^1 is constructed, such that $\mathcal{L}(\hat{X}^2) = \mathcal{L}(X^2)$ and $\forall t \leq T$, $\forall \omega \in \Omega$, $X^1_t(\omega) \leq \hat{X}^2_t(\omega)$.

2 Birth and death processes

We will collect in this section various results about the birth and death processes that appeared in Definition 1.

2.1 Comparison results

The following theorem gives various stochastic domination results.

Theorem 2

(a) *Assume (A). For any $K \geq 1$ and any $\mathbb{L}^1$ initial condition ν^K_0 of the process ν^K,*

$$\mathcal{L}(\langle \nu^K, \mathbf{1} \rangle) \preceq \mathbf{P}^K(2\bar{b}, 0, \underline{\alpha}, \langle \nu^K_0, \mathbf{1} \rangle).$$

(b) *With the same assumptions than in (a), let A^K_t denote the number of mutations occuring in ν^K between times 0 and t, and let $a, a_1, a_2 \geq 0$. Then, for $t \leq \inf\{s \geq 0 : \langle \nu^K_s, \mathbf{1} \rangle \geq a\}$,*

$$A^K \preceq B^K,$$

where B^K is a Poisson process with parameter $K u_K a \bar{b}$.

If moreover $\nu^K_0 = \langle \nu^K_0, \mathbf{1} \rangle \delta_x$, define $\tau_1 = \inf\{t \geq 0 : A^K_t = 1\}$ (the first mutation time). Then, for $t \leq \tau_1 \wedge \inf\{s \geq 0 : \langle \nu^K_s, \mathbf{1} \rangle \notin [a_1, a_2]\}$,

$$B^K \preceq A^K \preceq C^K,$$

where B^K and C^K are Poisson process with parameter, respectively, $K u_K a_1 \mu(x) b(x)$ and $K u_K a_2 \mu(x) b(x)$.

(c) *Fix $K \geq 1$ and take b, d, c, z as in Definition 1 (a). Then, for any $\varepsilon_1, \varepsilon_2, \varepsilon_3 \geq 0$ and any $\mathbb{N}/K$-valued random variable ε_4,*

$$\mathbf{P}^K(b, d + \varepsilon_2, c + \varepsilon_3, z) \preceq \mathbf{P}^K(b + \varepsilon_1, d, c, z + \varepsilon_4).$$

(d) *Let (Z^1, Z^2) be a stochastic process with law*

$$\mathbf{Q}^K(b_1, b_2, d_1, d_2, c_{11}, c_{12}, c_{21}, c_{22}, z_1, z_2)$$

where the parameters are as in Definition 1 (b). Fix $a > 0$ and define $T = \inf\{t \geq 0, Z^2 \geq a\}$. Then, there exists two processes M^1 and M^2 such that, for $t \leq T$,

$$M^1 \preceq Z^1 \preceq M^2,$$
$$\text{where} \quad \mathcal{L}(M^1) = \mathbf{P}^K(b_1, d_1 + ac_{12}, c_{11}, z_1)$$
$$\text{and} \quad \mathcal{L}(M^2) = \mathbf{P}^K(b_1, d_1, c_{11}, z_1).$$

(e) *Take (Z^1, Z^2) as above, fix $0 \leq a_1 < a_2$ and $a > 0$, and define $T = \inf\{t \geq 0, Z^1 \notin [a_1, a_2]$ or $Z^2 \geq a\}$. Then, there exists M^1 and M^2 such that, for $t \leq T$,*

$$M^1 \preceq Z^2 \preceq M^2,$$
$$\text{where} \quad \mathcal{L}(M^1) = \mathbf{P}^K(b_2, d_2 + a_2c_{21} + ac_{22}, 0, z_2)$$
$$\text{and} \quad \mathcal{L}(M^2) = \mathbf{P}^K(b_2, d_2 + a_1c_{21}, 0, z_2).$$

Remark 2 *Point (a) explains why it is necessary to combine simultaneously the limits $K \to +\infty$ and $u_K \to 0$ in order to obtain the TSS process in Theorem 1. The limit $K \to +\infty$ taken alone leads to a deterministic dynamics (cf. Fournier and Méléard [8]), so making the rare mutations limit afterwards cannot lead to a stochastic process. Conversely, taking the limit of rare mutations without making the population larger would lead to an immediate extinction of the population in the mutations time scale, because the stochastic domination of Theorem 2 (a) is independent of u_K and $\mu(\cdot)$, and because a process Z with law $\mathbf{P}^K(2\bar{b}, 0, \underline{\alpha}, \gamma_K/K)$ gets a.s. extinct in finite time for any $K \geq 1$.*

Proof of (a) We will use the construction of the process ν^K given by Fournier and Méléard [8]: let $(\Omega, \mathcal{F}, \mathbf{P})$ be a sufficiently large probability space, and consider on this space the following five independent random objects:

(i) a $\mathcal{M}^K$-valued random variable ν_0^K (the initial distribution),

(ii) a Poisson point measure $N_1(ds, di, dv)$ on $[0, \infty[\times \mathbb{N} \times [0, 1]$ with intensity measure $q_1(ds, di, dv) = \bar{b} ds \sum_{k\geq 1} \delta_k(di)dv$ (the birth without mutation Poisson point measure),

(iii) a Poisson point measure $N_2(ds, di, dh, dv)$ on $[0, \infty[\times \mathbb{N} \times \mathbb{R}^l \times [0, 1]$ with intensity measure $q_2(ds, di, dh, dv) = \bar{b} ds \sum_{k\geq 1} \delta_k(di)m(h)dhdv$ (the birth with mutation Poisson point measure),

(iv) a Poisson point measure $N_3(ds, di, dv)$ on $[0, \infty[\times \mathbb{N} \times [0, 1]$ with intensity measure $q_3(ds, di, dv) = \bar{d} ds \sum_{k\geq 1} \delta_k(di)dv$ (the natural death Poisson point measure),

(v) a Poisson point measure $N_4(ds, di, dj, dv)$ on $[0, \infty[\times \mathbb{N} \times \mathbb{N} \times [0, 1]$ with intensity measure $q_4(ds, di, dj, dv) = (\bar{\alpha}/K) ds \sum_{k\geq 1} \delta_k(di) \sum_{m\geq 1} \delta_m(dj)dv$ (the competition death Poisson point measure).

We will also need the following function, solving the purely notational problem of associating a number to each individual in the population: for any $K \geq 1$, let $H=(H^1,\ldots,H^k,\ldots)$ be the map from $\mathcal{M}^K$ into $(\mathbb{R}^l)^{\mathbb{N}}$ defined by

$$H\left(\frac{1}{K}\sum_{i=1}^n \delta_{x_i}\right)=(x_{\sigma(1)},\ldots,x_{\sigma(n)},0,\ldots,0,\ldots),$$

where $x_{\sigma(1)}\preccurlyeq\ldots\preccurlyeq x_{\sigma(n)}$ for the lexicographic order $\preccurlyeq$ on $\mathbb{R}^d$. For convenience, we have omitted in our notation the dependence of H and H^i on K.

Then a process ν^K with generator L^K and initial state ν_0^K can be constructed as follows: for any $t\geq 0$,

$$\begin{aligned}
\nu_t^K=\nu_0^K &+\int_0^T\int_{\mathbb{N}}\int_0^1 \mathbf{1}_{\{i\leq K\langle\nu_{s-}^K,\mathbf{1}\rangle\}}\frac{\delta_{H^i(\nu_{s-}^K)}}{K} \\
&\qquad \mathbf{1}_{\left\{v\leq\frac{[1-u_K\mu(H^i(\nu_{s-}^K))]b(H^i(\nu_{s-}^K))}{\bar b}\right\}}N_1(ds,di,dv) \\
&+\int_0^T\int_{\mathbb{N}}\int_{\mathbb{R}^l}\int_0^1 \mathbf{1}_{\{i\leq K\langle\nu_{s-}^K,\mathbf{1}\rangle\}}\frac{\delta_{H^i(\nu_{s-}^K)+h}}{K} \\
&\qquad \mathbf{1}_{\left\{v\leq\frac{u_K\mu(H^i(\nu_{s-}^K))b(H^i(\nu_{s-}^K))}{\bar b}\frac{m(H^i(\nu_{s-}^K),h)}{\bar m(h)}\right\}}N_2(ds,di,dh,dv) \\
&-\int_0^T\int_{\mathbb{N}}\int_0^1 \mathbf{1}_{\{i\leq K\langle\nu_{s-}^K,\mathbf{1}\rangle\}}\frac{\delta_{H^i(\nu_{s-}^K)}}{K}\mathbf{1}_{\left\{v\leq\frac{d(H^i(\nu_{s-}^K))}{\bar d}\right\}}N_3(ds,di,dv) \\
&-\int_0^T\int_{\mathbb{N}}\int_{\mathbb{N}}\int_0^1 \mathbf{1}_{\{i\leq K\langle\nu_{s-}^K,\mathbf{1}\rangle\}}\mathbf{1}_{\{j\leq K\langle\nu_{s-}^K,\mathbf{1}\rangle\}}\frac{\delta_{H^i(\nu_{s-}^K)}}{K} \\
&\qquad \mathbf{1}_{\left\{v\leq\frac{\alpha(H^i(\nu_{s-}^K),H^j(\nu_{s-}^K))}{\bar\alpha}\right\}}N_4(ds,di,dj,dv).
\end{aligned}\tag{13}$$

Although this formula is quite complicated, the principle is very simple: for each type of event, the corresponding Poisson point process jumps faster than ν^K has to. We decide whether a jump of the process ν^K occurs by comparing v to a quantity related to the rates of the various events. The indicator functions involving i and j simply ensures that the i^{th} and j^{th} individuals are alive in the population (since $K\langle\nu_t^K,\mathbf{1}\rangle$ is the number of individuals in the population at time t).

Under (A1), (A2) and the assumption that $\mathbf{E}(\langle\nu_0^K,\mathbf{1}\rangle)<\infty$, Fournier and Méléard [8] prove the existence and uniqueness of the solution to (13), and that this solution is a Markov process with infinitesimal generator (4).

Then, the $\mathbb{N}/K$-valued Markov process Z^K defined by

$$\begin{aligned}
Z_t^K=\langle\nu_0^K,\mathbf{1}\rangle+\frac{1}{K}\int_0^t\int_{\mathbb{N}}\mathbf{1}_{\{i\leq KZ_{s-}^K\}}\Bigg(\int_0^1 N_1(ds,di,dv) \\
+\int_{\mathbb{R}^l}\int_0^1 N_2(ds,di,dh,dv)-\int_{\mathbb{N}}\int_0^1\mathbf{1}_{\{j\leq KZ_{s-}^K,\ v\leq\underline{\alpha}/\bar\alpha\}}N_4(ds,di,dj,dv)\Bigg)
\end{aligned}\tag{14}$$

can be easily proved to satisfy $\mathcal{L}(Z^K)=\mathbf{P}^K(2\bar b,0,\underline{\alpha},\langle\nu_0^K,\mathbf{1}\rangle)$. Moreover, if for some $\omega\in\Omega$, and for some $t\geq 0$, $Z_t^K(\omega)=\langle\nu_t^K(\omega),\mathbf{1}\rangle$, let

$$T^K=\inf\{s\geq t, Z_s^K(\omega)\neq\langle\nu_s^K(\omega),\mathbf{1}\rangle\}.$$

Then, the comparison of (13) and (14) yields that, on the time interval $[t,T^K]$, any birth time (with or without mutation) for ν^K is also a birth time for Z^K, and any death time for Z^K is also a death time for ν^K. Consequently, $Z^K_{T^K}(\omega)$ is necessarily greater than $\langle \nu^K_{T^K}(\omega), \mathbf{1}\rangle$, which implies the required domination result. $\square$

Proof of (b) With the same notations as above,

$$A_t^K := \int_0^t \int_{\mathbb{N}} \int_{\mathbb{R}^l} \int_0^1 \mathbf{1}_{\{i \leq K\langle \nu^K_{s-}, \mathbf{1}\rangle\}} \times \\ \times \mathbf{1}_{\left\{v \leq \frac{u_K \mu(H^i(\nu^K_{s-})) b(H^i(\nu^K_{s-}))}{\bar{b}} \frac{m(H^i(\nu^K_{s-}), h)}{m(h)}\right\}} N_2(ds, di, dh, dv).$$

Therefore, for $t \leq \inf\{s \geq 0 : \langle \nu^K_s, \mathbf{1}\rangle \geq a\}$,

$$A_t^K \leq \int_0^t \int_{\mathbb{N}} \int_{\mathbb{R}^l} \int_0^1 \mathbf{1}_{\{i \leq Ka\}} \mathbf{1}_{\{v \leq u_K\}} N_2(ds, di, dh, dv). \tag{15}$$

Since the intensity measure of N_2 is

$$q_2(ds, di, dh, dv) = \bar{b} ds \sum_{k \geq 1} \delta_k(di) m(h) dh dv, \tag{16}$$

the right-hand side of (15) is a Poisson process with parameter $K u_K a \bar{b}$.

In the case where $\nu^K_0 = \langle \nu^K_0, \mathbf{1}\rangle \delta_x$, as long as $t < \tau_1$, $\nu^K_t = \langle \nu^K_t, \mathbf{1}\rangle \delta_x$, therefore, for $t \leq \tau_1 \wedge \inf\{s \geq 0 : \langle \nu^K_s, \mathbf{1}\rangle \notin [a_1, a_2]\}$,

$$\int_0^t \int_{\mathbb{N}} \int_{\mathbb{R}^l} \int_0^1 \mathbf{1}_{\{i \leq K a_1\}} \mathbf{1}_{\left\{v \leq \frac{u_K \mu(x) b(x)}{\bar{b}} \frac{m(x,h)}{m(h)}\right\}} N_2(ds, di, dh, dv) \leq A_t^K \\ \leq \int_0^t \int_{\mathbb{N}} \int_{\mathbb{R}^l} \int_0^1 \mathbf{1}_{\{i \leq K a_2\}} \mathbf{1}_{\left\{v \leq \frac{u_K \mu(x) b(x)}{\bar{b}} \frac{m(x,h)}{m(h)}\right\}} N_2(ds, di, dh, dv).$$

By (16), the left-hand side and the right-hand side of this inequality is are Poisson processes with parameters $K u_K a_1 \mu(x) b(x)$ and $K u_K a_2 \mu(x) b(x)$, respectively. $\square$

Proof of (c) (d) and (e) The proofs of (c), (d) and (e) are very similar. We will only prove in detail (e).

Consider, on a sufficiently rich probability space $(\Omega, \mathcal{F}, \mathbf{P})$ the random variables z_1 and z_2 as in the statement of Theorem 2, and the following independent random objects:

(i) two Poisson point measures $\hat{N}^k_1(ds, di, dv)$ $(k = 1, 2)$ on $[0, \infty[\times \mathbb{N} \times [0, 1]$ with intensity measures $q^k_1(ds, di, dv) = b_k ds \sum_{n \geq 1} \delta_n(di) dv$ $(k = 1, 2)$,

(ii) two Poisson point measures $\hat{N}^k_2(ds, di, dv)$ $(k = 1, 2)$ on $[0, \infty[\times \mathbb{N} \times [0, 1]$ with intensity measures $q^k_2(ds, di, dv) = d_k ds \sum_{n \geq 1} \delta_n(di) dv$ $(k = 1, 2)$,

(iii) two Poisson point measures $\hat{N}^k_3(ds, di, dj, dv)$ $(k = 1, 2)$ on $[0, \infty[\times \mathbb{N} \times \mathbb{N} \times [0, 1]$ with intensity measures

$$q^k_3(ds, di, dj, dv) = (c_{k1}/K) ds \sum_{n \geq 1} \delta_n(di) \sum_{m \geq 1} \delta_m(dj) dv \quad (k = 1, 2).$$

(iv) two Poisson point measures $\hat{N}_4^k(ds, di, dj, dv)$ $(k = 1, 2)$ on $[0, \infty[\times\mathbb{N}\times\mathbb{N}\times[0, 1]$ with intensity measures

$$q_4^k(ds, di, dj, dv) = (c_{k2}/K)ds\sum_{n\geq 1}\delta_n(di)\sum_{m\geq 1}\delta_m(dj)dv \quad (k = 1, 2).$$

The processes Z^1 and Z^2 can be constructed on Ω as follows: for any $t \geq 0$, and for $k = 1, 2$,

$$\begin{aligned} Z_t^k = z_k + \frac{1}{K}\int_0^t\int_{\mathbb{N}}\mathbf{1}_{\{i\leq KZ_{s-}^k\}}\bigg(\int_0^1 \hat{N}_1^k(ds, di, dv) - \int_0^1 \hat{N}_2^k(ds, di, dh, dv) \\ - \int_{\mathbb{N}}\int_0^1\Big(\mathbf{1}_{\{j\leq KZ_{s-}^1\}}\hat{N}_3^k(ds, di, dj, dv) + \mathbf{1}_{\{j\leq KZ_{s-}^2\}}\hat{N}_4^k(ds, di, dj, dv)\Big)\bigg). \end{aligned}$$

Then, we can define on Ω the processes M^1 and M^2 by, for any $t \geq 0$,

$$\begin{aligned} M_t^1 = z_2 + \frac{1}{K}\int_0^t\int_{\mathbb{N}}\mathbf{1}_{\{i\leq KM_{s-}^1\}}\bigg(\int_0^1 \hat{N}_1^2(ds, di, dv) - \int_0^1 \hat{N}_2^2(ds, di, dh, dv) \\ - \int_{\mathbb{N}}\int_0^1\Big(\mathbf{1}_{\{j\leq Ka_2\}}\hat{N}_3^2(ds, di, dj, dv) + \mathbf{1}_{\{j\leq Ka\}}\hat{N}_4^2(ds, di, dj, dv)\Big)\bigg) \end{aligned}$$

and

$$\begin{aligned} M_t^2 = z_2 + \frac{1}{K}\int_0^t\int_{\mathbb{N}}\mathbf{1}_{\{i\leq KM_{s-}^2\}}\bigg(\int_0^1 \hat{N}_1^2(ds, di, dv) - \int_0^1 \hat{N}_2^2(ds, di, dh, dv) \\ - \int_{\mathbb{N}}\int_0^1\mathbf{1}_{\{j\leq Ka_1\}}\hat{N}_3^2(ds, di, dj, dv)\bigg). \end{aligned}$$

A comparison between the birth and death events of Z^2, M^1 and M^2 in a similar way than in the proof of (a) proves that $M_t^1(\omega) \leq Z_t^2(\omega) \leq M_t^2(\omega)$ for any $t \leq \inf\{t \geq 0, Z^1 \notin [a_1, a_2] \text{ or } Z^2 \geq a\}$ and for any $\omega \in \Omega$. $\Box$

2.2 Problem of exit from a domain

Let us give some results on $\mathbf{P}^K(b, d, c, z)$ when $c > 0$. Points (a) and (b) of the following theorem strengthen Proposition 1, and point (c) studies the problem of exit from a domain.

Theorem 3

(a) *Let $c, T > 0$ and $b, d \geq 0$, let C be a compact subset of $\mathbb{R}_+^*$, and write $\mathbf{P}_z^K = \mathbf{P}^K(b, d, c, z)$ for $z \in \mathbb{N}/K$. Let ϕ_z denote the solution to*

$$\dot{\phi} = (b - d - c\phi)\phi \tag{17}$$

with initial condition $\phi_z(0) = z$. Then

$$r := \inf_{z\in C}\ \inf_{0\leq t\leq T}|\phi_z(t)| > 0 \tag{18}$$

and

$$R := \sup_{z\in C}\ \sup_{0\leq t\leq T}|\phi_z(t)| < +\infty. \tag{19}$$

Then, for any $\delta < r$, *(with the convention* $\sup\emptyset = 0$*)*

$$\lim_{K\to+\infty} \sup_{z\in C} \mathbf{P}_z^K\left(\sup_{0\le t\le T} |w_t - \phi_z(t)| \ge \delta\right) = 0, \tag{20}$$

where w_t *is the canonical process on* $\mathbb{D}(\mathbb{R}_+, \mathbb{R})$.

(b) *Let* $T, c_{ij} > 0$ *and* $b_i, d_i \ge 0$ $(i, j \in \{1, 2\})$, *let* C *be a compact subset of* $(\mathbb{R}_+^*)^2$, *and write* $\mathbf{Q}^K_{z_1,z_2} = \mathbf{Q}^K(b_1, b_2, d_1, d_2, c_{11}, c_{12}, c_{21}, c_{22}, z_1, z_2)$ *for* z_1 *and* z_2 *in* $\mathbb{N}/K$. *Let* $\phi_{z_1,z_2} = (\phi^1_{z_1,z_2}, \phi^2_{z_1,z_2})$ *denote the solution to*

$$\dot{\phi}^1 = (b_1 - d_1 - c_{11}\phi^1 - c_{12}\phi^2)\phi^1$$
$$\dot{\phi}^2 = (b_2 - d_2 - c_{21}\phi^1 - c_{22}\phi^2)\phi^2$$

with initial conditions $\phi^1_{z_1,z_2}(0) = z_1$ *and* $\phi^2_{z_1,z_2}(0) = z_2$. *Then*

$$r := \inf_{z\in C} \inf_{0\le t\le T} \|\phi_{z_1,z_2}(t)\| > 0 \tag{21}$$

and

$$\sup_{z\in C} \sup_{0\le t\le T} \|\phi_{z_1,z_2}(t)\| < +\infty.$$

Then, for any $\delta < r$,

$$\lim_{K\to+\infty} \sup_{z\in C} \mathbf{Q}^K_{z_1,z_2}(\sup_{0\le t\le T} \|\hat{w}_t - \phi_{z_1,z_2}(t)\| \ge \delta) = 0,$$

where $\hat{w}_t = (\hat{w}^1_t, \hat{w}^2_t)$ *is the canonical process on* $\mathbb{D}(\mathbb{R}_+, \mathbb{R}^2)$.

(c) *Let* $b, c > 0$ *and* $0 \le d < b$. *Observe that* $(b-d)/c$ *is the unique stable steady state of (17). Fix* $0 < \eta_1 < (b-d)/c$ *and* $\eta_2 > 0$, *and define on* $\mathbb{D}(\mathbb{R}_+, \mathbb{R})$

$$T^K = \inf\left\{t \ge 0 : w_t \notin \left[\frac{b-d}{c} - \eta_1, \frac{b-d}{c} + \eta_2\right]\right\}.$$

Then, there exists $V > 0$ *such that, for any compact subset* C *of* $](b-d)/c - \eta_1, (b-d)/c + \eta_2[$,

$$\lim_{K\to+\infty} \sup_{z\in C} \mathbf{P}_z^K(T^K < e^{KV}) = 0. \tag{22}$$

Proof of (a) and (b) Observe that any solution to (17) with positive initial condition is bounded, since $\dot{\phi} < 0$ as soon as $\phi > (b-d)/c$. This implies easily (19).

Moreover, since a solution to (17) can be written as

$$\phi(t) = \phi(0)\exp\left(\int_0^t (b - d - c\phi(s))ds\right),$$

it cannot reach 0 in finite time if $\phi(0) > 0$. Therefore, (18) follows from the continuity of the flow, which is a classical consequence of the fact that $z \mapsto (b - d - cz)z$ is locally Lipschitz and of Gronwall's Lemma (cf. e.g. Queffélec and Zuily [15] p. 356).

Finally, (20) is a consequence of large deviations estimates for the sequence of laws $(\mathbf{P}_z^K)_{K\geq 1}$. As can be seen in Theorem 10.2.6 in chapter 10 of Dupuis and Ellis [5], a large deviations principle on $[0,T]$ with a good rate function I_T holds for $\mathbb{Z}/K$-valued Markov jump processes with transition rates

$$\begin{array}{ll} Kp(i/K) & \text{from } i/K \text{ to } (i+1)/K, \\ Kq(i/K) & \text{from } i/K \text{ to } (i-1)/K, \end{array}$$

where p and q are functions defined on $\mathbb{R}$ and with positive values, bounded, Lipschitz and uniformly bounded away from 0. The rate function I_T writes

$$I_T(\phi) = \begin{cases} \displaystyle\int_0^T L(\phi(t),\dot\phi(t))dt & \text{if } \phi \text{ is absol. cont. on } [0,T] \\ +\infty & \text{otherwise,} \end{cases} \tag{23}$$

where $L(y,z) = 0$ if $z = p(y) - q(y)$ and $L(y,z) > 0$ otherwise. Therefore, $I_T(\phi) = 0$ if and only if ϕ is absolutely continuous and

$$\dot\phi = p(\phi) - q(\phi). \tag{24}$$

Moreover, this large deviation is *uniform* with respect to the initial condition. This means that, if $\mathbf{R}_z^K$ denotes the law of this process with initial condition z, for any compact set $C \subset \mathbb{R}$, for any closed set F and any open set G of $\mathbb{D}([0,T],\mathbb{R})$,

$$\liminf_{K\to+\infty} \frac{1}{K} \log \inf_{z\in C} \mathbf{R}_z^K(G) \geq -\sup_{z\in C} \inf_{\psi\in G,\ \psi(0)=z} I_T(\psi) \tag{25}$$

$$\text{and} \quad \limsup_{K\to+\infty} \frac{1}{K} \log \sup_{z\in C} \mathbf{R}_z^K(F) \leq - \inf_{\psi\in F,\ \psi(0)\in C} I_T(\psi). \tag{26}$$

Our birth and death process does not satisfy these asumptions. However, if we define

$$p(z) = b\chi(z) \quad \text{and} \quad q(z) = d\chi(z) + c\chi(z)^2,$$
$$\text{where} \quad \chi(z) = z \text{ if } z \in [r-\delta, R+\delta];\ r-\delta \text{ if } z < r-\delta;\ R+\delta \text{ if } z > R+\delta,$$

then $\mathbf{R}_z^K = \mathbf{P}_z^K$ on the time interval $[0,\tau]$, where $\tau = \inf\{t \geq 0, w_t \notin [r-\delta, R+\delta]\}$, and p and q satisfy the assumptions above. Therefore, by (26),

$$\limsup_{K\to+\infty} \frac{1}{K} \log \sup_{z\in C} \mathbf{P}_z^K\left(\sup_{0\leq t\leq T} |w_t - \phi_z(t)| \geq \delta \right) \leq -\inf_{\psi\in F^\delta} I_T(\psi), \quad \text{where}$$
$$F^\delta := \left\{ \psi \in \mathbb{D}([0,T],\mathbb{R}) : \psi(0) \in C \text{ and } \exists t \in [0,T],\ |\psi(t) - \phi_{\psi(0)}(t)| \geq \delta \right\}$$

By the continuity of the flow of (24), the set F^δ is closed. Since I_T is a good rate function, the infimum of I_T over this set is attained at some function belonging to F^δ, which cannot be a solution to (24), and thus non-zero. This ends the proof of (20).

The proof of (b) can be made in a very similar way, using the large deviations estimates for two-dimensional jump processes of Theorem 10.2.6 in chapter 10 of Dupuis and Ellis [5]. $\square$

Proof of (c) Define the function χ on $\mathbb{R}$ by $\chi(z) = z$ if $z \in [(b-d)/c-\eta_1,(b-d)/c+\eta_2]$, $\chi(z) = (b-d)/c-\eta_1$ for $z < (b-d)/c-\eta_1$ and $\chi(z) = (b-d)/c+\eta_2$ for $z > (b-d)/c-\eta_2$. As in the proof of (a), we can construct from the functions $p(z) = b\chi(z)$ and $q(z) = d\chi(z) + c\chi(z)^2$ a family of laws $(\mathbf{R}^K_z)$ such that $\mathbf{R}^K_z = \mathbf{P}^K_z$ on the time interval $[0, T^K]$, and such that (25) and (26) hold for the good rate function I_T defined in (23).

Observe that any solution to (24) are monotonic and converge to $(b-d)/c$ when $t \to +\infty$. Therefore, the following estimates for the time of exit from an attracting domain are classical (cf. Freidlin and Wentzell [9], chapter 5, section 4): there exists $\bar{V} \geq 0$ such that, for any $\delta > 0$,

$$\lim_{K\to+\infty} \inf_{z\in C} \mathbf{R}^K_z \left(e^{K(\bar{V}-\delta)} < T^K < e^{K(\bar{V}+\delta)} \right) = 1,$$

which implies (22) if we can prove that $\bar{V} > 0$.

The constant $\bar{V}$ is obtained as follows (see [9] pp. 108–109): for any $y, z \in \mathbb{R}$, define

$$V(y,z) := \inf_{t>0,\ \varphi(0)=y,\ \varphi(t)=z} I_t(\varphi).$$

Then

$$\bar{V} := V\left(\frac{b-d}{c}, \frac{b-d}{c} - \eta_1\right) \wedge V\left(\frac{b-d}{c}, \frac{b-d}{c} + \eta_2\right).$$

Now, Theorem 5.4.3. of [9] states that, for any $y, z \in \mathbb{R}$, the infimum defining $V(y, z)$ is attained at some function ϕ linking y to z, in the sense that, either there exists an absolutely continuous function ϕ defined on $[0, T]$ for some $T > 0$ such that $\phi(0) = y$, $\phi(T) = z$ and $V(y, z) = I_T(\phi) = \int_0^T L(\phi(t), \dot{\phi}(t))dt$, or there exists an absolutely continuous function ϕ defined on $]-\infty, T]$ for some $T > -\infty$ such that $\lim_{t\to-\infty} \phi(t) = y$, $\phi(T) = z$ and $V(y, z) = \int_{-\infty}^T L(\phi(t), \dot{\phi}(t))dt$.

Since any solution to (24) is decreasing as long as it stays in $[(b-d)/c, +\infty[$, a function ϕ defined on $[0, T]$ or $]-\infty, T]$ linking $(b-d)/c$ to $(b-d)/c+\eta_2$ cannot be a solution to (24), and thus $V((b-d)/c, (b-d)/c+\eta_2) > 0$. Similarly, $V((b-d)/c, (b-d)/c-\eta_1) > 0$, and so $\bar{V} > 0$, which concludes the proof of Theorem 3. □

2.3 Some results on branching processes

Observe that, when $c = 0$, $\mathbf{P}^K(b, d, 0, z)$ is the law of a binary branching process divided by K. Let us give some results on these processes.

Theorem 4 *Let $b, d > 0$. As in Theorem 3, define, for any $K \geq 1$ and any $z \in \mathbb{N}/K$, $\mathbf{P}^K_z = \mathbf{P}^K(b, d, 0, z)$. Define also, for any $z \in \mathbb{R}$, on $\mathbb{D}(\mathbb{R}_+, \mathbb{R})$, the stopping time*

$$T_z = \inf\{t \geq 0 : w_t = z\}.$$

Finally, let $(t_K)_{K\geq1}$ be a sequence of positive numbers such that $\log K \ll t_K$.

(a) *If $b < d$ (sub-critical case), for any $\varepsilon > 0$,*

$$\lim_{K\to+\infty} \mathbf{P}^K_{1/K}(T_0 \leq t_K \wedge T_{\lceil\varepsilon K\rceil/K}) = 1, \tag{27}$$

$$\textit{and} \quad \lim_{K\to+\infty} \mathbf{P}^K_{\lfloor\varepsilon K\rfloor/K}(T_0 \leq t_K) = 1. \tag{28}$$

Moreover, for any $K \geq 1$, $k \geq 1$ and $n \geq 1$,

$$\mathbf{P}^K_{n/K}(T_{kn/K} \leq T_0) \leq \frac{1}{k}. \tag{29}$$

(b) *If $b > d$ (super-critical case), for any $\varepsilon > 0$,*

$$\lim_{K\to+\infty} \mathbf{P}^K_{1/K}(T_0 \leq t_K \wedge T_{\lceil \varepsilon K\rceil/K}) = \frac{d}{b} \tag{30}$$

$$\textit{and} \quad \lim_{K\to+\infty} \mathbf{P}^K_{1/K}(T_{\lceil \varepsilon K\rceil/K} \leq t_K) = 1 - \frac{d}{b}. \tag{31}$$

Proof Let us denote by $\mathbf{Q}_n$ the law of the binary branching process with initial state $n \in \mathbb{N}$, with individual birth rate b and individual death rate d. Then (27), (28), (29), (30) and (31) rewrite respectively

$$\lim_{K\to+\infty} \mathbf{Q}_1(T_0 \leq t_K \wedge T_{\lceil \varepsilon K\rceil}) = 1, \tag{32}$$

$$\lim_{K\to+\infty} \mathbf{Q}_{\lfloor \varepsilon K\rfloor}(T_0 \leq t_K) = 1, \tag{33}$$

$$\mathbf{Q}_n(T_{kn} \leq T_0) \leq \frac{1}{k}, \tag{34}$$

$$\lim_{K\to+\infty} \mathbf{Q}_1(T_0 \leq t_K \wedge T_{\lceil \varepsilon K\rceil}) = \frac{d}{b} \tag{35}$$

$$\text{and} \quad \lim_{K\to+\infty} \mathbf{Q}_1(T_{\lceil \varepsilon K\rceil} \leq t_K) = 1 - \frac{d}{b}. \tag{36}$$

The limit (33) follows easily from the distribution of the extinction time for binary branching processes when $b \neq d$ (cf. Athreya and Ney [1] p. 109): for any $t \geq 0$ and $n \in \mathbb{N}$,

$$\mathbf{Q}_n(T_0 \leq t) = \left(\frac{d\left(1 - e^{-(b-d)t}\right)}{b - de^{-(b-d)t}} \right)^n. \tag{37}$$

It is known that there is no accumulation of jumps for branching processes. Therefore, under $\mathbf{Q}_1$, when $K \to +\infty$, $T_{\lceil \varepsilon K\rceil} \to +\infty$ a.s. , and thus $\mathbf{Q}_1(T_0 \leq T_{\lceil \varepsilon K\rceil},\ T_0 < \infty) \to \mathbf{Q}_1(T_0 < \infty)$. Therefore, (32) and (35) follow easily from (37).

The inequality (34) follows from the fact that, if $(Z_t, t \geq 0)$ is a process with law $\mathbf{Q}_n$, $(Z_t \exp(-(b-d)t), t \geq 0)$ is a martingale (cf. [1] p. 111). Then, Doob's stopping theorem applied to the stopping time $T_0 \wedge T_{kn}$ yields,

$$\mathbf{E}_n(kne^{(d-b)T_{kn}} \mathbf{1}_{\{T_{kn} < T_0\}}) = n,$$

where $\mathbf{E}_n$ is the expectation with respect to $\mathbf{Q}_n$. Therefore, when $b < d$, $kn\mathbf{Q}_n(T_{kn} < T_0) \leq n$, and the proof of (34) is completed.

The limit (36) follows from the fact that, if $(Z_t, t \geq 0)$ is a process with law $\mathbf{Q}_1$, the martingale $(Z_t \exp(-(b-d)t), t \geq 0)$ converges a.s. when $t \to +\infty$ to a random variable W, where $W = 0$ on the event $\{T_0 < \infty\}$ and $W > 0$ on the event $\{T_0 = \infty\}$ (cf. [1] p. 112). Hence, on the event $\{T_0 = \infty\}$, when $b > d$,

$$\log Z_t/t \to b - d > 0,$$

and so, for almost any $\omega \in \{T_0 = \infty\}$, there exists $S(\omega) < \infty$ such that for any $t \geq S(\omega)$,

$$Z_t \geq \exp((b-d)t/2).$$

Therefore, since $\log K \ll t_K$, for any $\varepsilon > 0$, $\mathbf{Q}_1(T_0 = \infty,\ T_{\lceil \varepsilon K \rceil} \geq t_K) \to 0$ when $K \to +\infty$. Then, (36) follows from the fact that $\mathbf{Q}_1(T_0 = \infty) = 1 - d/b$, which is an immediate consequence of (37). $\square$

3 Proof of Theorem 1

Let us assume, without loss of generality, that ν^K is constructed by (13) on a sufficiently rich probability space $(\Omega, \mathcal{F}, \mathbf{P})$.

Let us introduce the following sequences of stopping times: for all $n \geq 1$, let τ_n be the first mutation time after time τ_{n-1}, with $\tau_0 = 0$ (*i.e.* τ_n is the n^{th} mutation time), and for any $n \geq 0$, let θ_n be the first time after τ_n when the population gets monomorphic. Observe that $\theta_0 = 0$ if the initial population is monomorphic. For any $n \geq 1$, define the random variable U_n as the new trait value appearing at the mutation time τ_n, and, when $\theta_n < \infty$, define V_n by $\mathrm{Supp}(\nu^K_{\theta_n}) = \{V_n\}$. When $\theta_n = +\infty$, define $V_n = +\infty$.

Our proof of Theorem 1 is based on the following two lemmas. The first lemma proves that there is no accumulation of mutations on the time scale of Theorem 1, and studies the asymptotic behavior of τ_1 starting from a monomorphic population, when $K \to +\infty$.

Lemma 1
(a) *Assume that the initial condition of ν^K satisfies $\sup_K \mathbf{E}(\langle \nu^K_0, \mathbf{1}\rangle) < +\infty$. Then, for any $\eta > 0$, there exists $\varepsilon > 0$ such that, for any $t > 0$,*

$$\limsup_{K\to+\infty} \mathbf{P}^K_{\nu^K_0}\left(\exists n \geq 0 : \frac{t}{Ku_K} \leq \tau_n \leq \frac{t+\varepsilon}{Ku_K}\right) < \eta. \tag{38}$$

Let $x \in \mathcal{X}$ and let $(z_K)_{K\geq 1}$ be a sequence of integers such that $z_K/K \to z > 0$.

(b) *For any $\varepsilon > 0$,*

$$\lim_{K\to+\infty} \mathbf{P}^K_{\frac{z_K}{K}\delta_x}\left(\tau_1 > \log K,\ \sup_{t\in[\log K, \tau_1]} |\langle \nu^K_t, \mathbf{1}\rangle - \bar{n}_x| > \varepsilon\right) = 0. \tag{39}$$

Observe that, by (a) with $t = 0$, since $\log K \ll 1/Ku_K$,

$$\lim_{K\to+\infty} \mathbf{P}^K_{\frac{z_K}{K}\delta_x}(\tau_1 < \log K) = 0.$$

In particular, under $\mathbf{P}^K_{\frac{z_K}{K}\delta_x}$, $\nu^K_{\log K} \xrightarrow{\mathcal{P}} \bar{n}_x\delta_x$ and $\nu^K_{\tau_1 -} \xrightarrow{\mathcal{P}} \bar{n}_x\delta_x$.

If, moreover, $z = \bar{n}_x$, then, for any $\varepsilon > 0$,

$$\lim_{K\to+\infty} \mathbf{P}^K_{\frac{z_K}{K}\delta_x}\left(\sup_{t\in[0,\tau_1]} |\langle \nu^K_t, \mathbf{1}\rangle - \bar{n}_x| > \varepsilon\right) = 0. \tag{40}$$

(c) *For any $t > 0$,*

$$\lim_{K\to+\infty} \mathbf{P}^K_{\frac{z_K}{K}\delta_x}\left(\tau_1 > \frac{t}{Ku_K}\right) = \exp(-\beta(x)t),$$

where $\beta(\cdot)$ has been defined in (2).

The second lemma studies the asymptotic behavior of θ_0 and V_0 starting from a dimorphic population, when $K \to +\infty$.

Lemma 2 *Fix $x, y \in \mathcal{X}$ satisfying (7) or (8), and let $(z_K)_{K\geq 1}$ be a sequence of integers such that $z_K/K \to \bar{n}_x$. Then,*

$$\lim_{K\to+\infty} \mathbf{P}^K_{\frac{z_K}{K}\delta_x+\frac{1}{K}\delta_y}(V_0 = y) = \frac{[f(y,x)]_+}{b(y)}, \tag{41}$$

$$\lim_{K\to+\infty} \mathbf{P}^K_{\frac{z_K}{K}\delta_x+\frac{1}{K}\delta_y}(V_0 = x) = 1 - \frac{[f(y,x)]_+}{b(y)}, \tag{42}$$

$$\forall \eta > 0, \quad \lim_{K\to+\infty} \mathbf{P}^K_{\frac{z_K}{K}\delta_x+\frac{1}{K}\delta_y}\left(\theta_0 > \frac{\eta}{Ku_K} \wedge \tau_1\right) = 0 \tag{43}$$

$$\text{and} \quad \forall \varepsilon > 0, \quad \lim_{K\to+\infty} \mathbf{P}^K_{\frac{z_K}{K}\delta_x+\frac{1}{K}\delta_y}\left(|\langle \nu^K_{\theta_0}, \mathbf{1}\rangle - \bar{n}_{V_0}| < \varepsilon\right) = 1, \tag{44}$$

where $f(y,x)$ has been defined in (3).

Observe that (43) implies in particular that

$$\lim_{K\to+\infty} \mathbf{P}^K_{\frac{z_K}{K}\delta_x+\frac{1}{K}\delta_y}(\theta_0 < \tau_1) = 1.$$

The proofs of these lemmas are postponed at the end of this section.

Proof of Theorem 1 Observe that the generator A, defined in (9), of the TSS process $(X_t, t \geq 0)$ of Theorem 1 can be written as

$$A\varphi(x) = \int_{\mathbb{R}^l} (\varphi(x+h) - \varphi(x))\beta(x)\kappa(x, dh), \tag{45}$$

where the probability measure $\kappa(x, dh)$ is defined by

$$\kappa(x, dh) = \left(1 - \int_{\mathbb{R}^l} \frac{[f(x+v,x)]_+}{b(x+v)} m(x,v)dv\right)\delta_0(dh) + \frac{[f(x+h,x)]_+}{b(x+h)} m(x,h)dh. \tag{46}$$

This means that the TSS model X with initial state x can be constructed as follows: let $(Z(k), k = 0, 1, 2, \ldots)$ be a Markov chain in $\mathcal{X}$ with initial state x and with transition kernel $\kappa(x, dh)$, and let $(N(t), t \geq 0)$ be and independent standard Poisson process. Let also $(T_n)_{n\geq 1}$ denote the sequence of jump times of the Poisson process N. Then, the process $(X_t, t \geq 0)$ defined by

$$X_t := Z\left(N\left(\int_0^t \beta(X_s)ds\right)\right)$$

is a Markov process with infinitesimal generator (45) (cf. [6] chapter 6).

Let $\mathbf{P}_x$ denote its law, and define $(S_n)_{n\geq 1}$ by $T_n = \int_0^{S_n} \beta(X_s)ds$. By (A1) and (A3), $\beta(\cdot) > 0$, and so S_n is finite for any $n \geq 1$. Observe that any jump of the process X occurs at some time S_n, but that all S_n may not be effective jump times for X, because of the Dirac mass at 0 appearing in (46).

Fix $t > 0$, $x \in \mathcal{X}$ and a measurable subset Γ of $\mathcal{X}$. Under $\mathbf{P}_x$, S_1 and X_{S_1} are independent, S_1 is an exponential random variable with parameter $\beta(x)$, and

$X_{S_1} - x$ has law $\kappa(x,\cdot)$. Therefore, for any $n \geq 1$, the strong Markov property applied to X at time S_1 yields

$$\mathbf{P}_x(S_n \leq t < S_{n+1},\ X_t \in \Gamma)$$
$$= \int_0^t \beta(x)e^{-\beta(x)s} \int_{\mathbb{R}^l} \mathbf{P}_{x+h}(S_{n-1} \leq t-s < S_n,\ X_{t-s} \in \Gamma)\kappa(x,dh)ds. \quad (47)$$

Moreover,

$$\mathbf{P}_x(0 \leq t < S_1,\ X_t \in \Gamma) = \mathbf{1}_{\{x\in\Gamma\}} e^{-\beta(x)t}. \quad (48)$$

The idea of our proof of Theorem 1 is to show that the same relations hold when we replace S_n by τ_n and X_t by the support of ν^K_{t/Ku_K} (when it is a singleton) and when $K \to +\infty$.

More precisely, fix $x \in \mathcal{X}$, $t > 0$ and a measurable subset Γ of $\mathcal{X}$, and observe that

$$\Big\{\mathrm{Supp}(\nu^K_{t/Ku_K}) \text{ is a singleton } \{y\},\ y \in \Gamma \text{ and } |\langle \nu^K_{t/Ku_K}, \mathbf{1}\rangle - \bar{n}_y| < \varepsilon\Big\}$$
$$= \bigcup_{n\geq 0} A^K_n(t,\Gamma,\varepsilon), \quad (49)$$

where

$$A^K_n(t,\Gamma,\varepsilon) := \left\{\theta_n \leq \frac{t}{Ku_K} < \tau_{n+1},\ V_n \in \Gamma,\ |\langle \nu^K_{t/Ku_K}, \mathbf{1}\rangle - \bar{n}_{V_n}| < \varepsilon\right\}.$$

Let us define, for any $z \in \mathbb{N}$ and $n \geq 0$,

$$p^K_n(t,x,\Gamma,\varepsilon,z) := \mathbf{P}^K_{\frac{z}{K}\delta_x}\Bigg(\theta_n \leq \frac{t}{Ku_K} < \tau_{n+1},\ V_n \in \Gamma$$
$$\text{and} \sup_{s\in[\theta_n,\tau_{n+1}]} |\langle \nu^K_s, \mathbf{1}\rangle - \bar{n}_{V_n}| < \varepsilon\Bigg),$$

and define also,

$$q^K_0(t,x,\Gamma,\varepsilon,z) := \mathbf{P}^K_{\frac{z}{K}\delta_x}\left(\frac{t}{Ku_K} < \tau_1,\ V_0 \in \Gamma,\ \sup_{s\in[\log K,\tau_1]} |\langle \nu^K_s, \mathbf{1}\rangle - \bar{n}_{V_0}| < \varepsilon\right)$$
$$= \mathbf{1}_{\{x\in\Gamma\}} \mathbf{P}^K_{\frac{z}{K}\delta_x}\left(\frac{t}{Ku_K} < \tau_1,\ \sup_{s\in[\log K,\tau_1]} |\langle \nu^K_s, \mathbf{1}\rangle - \bar{n}_x| < \varepsilon\right).$$

Let us also extend these definitions to $\varepsilon = \infty$ by suppressing the condition involving the supremum of $|\langle \nu^K, \mathbf{1}\rangle - \bar{n}_{V_n}|$.

Then

Lemma 3

(a) *For any $x \in \mathcal{X}$, $n \geq 1$, $t > 0$, $\varepsilon \in]0,\infty]$ and for any sequence of integers (z_K) such that $z_K/K \to z > 0$, $p_n(t,x,\Gamma) := \lim_{K\to+\infty} p^K_n(t,x,\Gamma,\varepsilon,z_K)$ exists, and is independent of (z_K), $z > 0$ and ε.*

Similarly, $p_0(t,x,\Gamma) := \lim_{K\to+\infty} q^K_0(t,x,\Gamma,\varepsilon,z_K)$ exists, and is independent of (z_K), $z > 0$ and ε, and, if $z = \bar{n}_x$, $\lim_{K\to+\infty} p^K_0(t,x,\Gamma,\varepsilon,z_K)$ exists and is also equal to $p_0(t,x,\Gamma)$.

Finally, if we assume that (z_K) is a sequence of $\mathbb{N}$-valued random variables such that z_K/K converge in probability to a deterministic $z > 0$, then the limits above hold in probability (with the same restriction that z has to be equal to $\bar{n}_x$ for p_0^K).

(b) *The functions $p_n(t, x, \Gamma)$ are continuous with respect to t and measurable with respect to x, and satisfy*

$$p_0(t, x, \Gamma) = \mathbf{1}_{\{x \in \Gamma\}} e^{-\beta(x)t} \quad \text{and} \quad \forall n \geq 0,$$
$$p_{n+1}(t, x, \Gamma) = \int_0^t \beta(x) e^{-\beta(x)s} \int_{\mathbb{R}^l} p_n(t - s, x + h, \Gamma) \kappa(x, dh) ds.$$

Let us postpone the proof of this lemma after the proof of Theorem 1.

Observe that, because of (47) and (48), Lemma 3 (b) implies that $\mathbf{P}_x(S_n \leq t < S_{n+1},\ X_t \in \Gamma) = p_n(t, x, \Gamma)$.

Now, let $\tilde{\mathbf{P}}_\nu^K$ denote the law of the process ν^K with *random* initial state ν. Since ν^K is Markov, $\tilde{\mathbf{P}}^K_{\gamma_K/K\delta_x} = \mathbf{E}[\mathbf{P}^K_{\gamma_K(\omega)/K\delta_x}]$. By (49),

$$\tilde{\mathbf{P}}^K_{\frac{\gamma_K}{K}\delta_x}\Big(\text{Supp}(\nu^K_{t/Ku_K}) \text{ is a singleton } \{y\},\ y \in \Gamma \\ \text{and } |\langle \nu^K_{t/Ku_K}, \mathbf{1}\rangle - \bar{n}_y| < \varepsilon\Big) = \sum_{n \geq 0} \tilde{\mathbf{P}}^K_{\frac{\gamma_K}{K}\delta_x}(A_n^K(t, \Gamma, \varepsilon)),$$

where (γ_K) is the sequence of $\mathbb{N}$-valued random variables of Theorem 1.

For any $K \geq 1$ and $n \geq 1$,

$$p_n^K(t, x, \Gamma, \varepsilon, \gamma_K) \leq \mathbf{P}^K_{\frac{\gamma_K}{K}\delta_x}(A_n^K(t, \Gamma, \varepsilon)) \leq p_n^K(t, x, \Gamma, \infty, \gamma_K),$$
$$\text{and} \quad q_0^K(t, x, \Gamma, \varepsilon, \gamma_K) \leq \mathbf{P}^K_{\frac{\gamma_K}{K}\delta_x}(A_n^K(t, \Gamma, \varepsilon)) \leq p_n^K(t, x, \Gamma, \infty, \gamma_K),$$

so, by Lemma 3 (a), for any $n \geq 0$, $\mathbf{P}^K_{\gamma_K/K\delta_x}(A_n^K(t, \Gamma, \varepsilon)) \xrightarrow{\mathcal{P}} p_n(t, x, \Gamma)$, and therefore, $\lim_{K\to+\infty} \tilde{\mathbf{P}}^K_{\gamma_K/K\delta_x}(A_n^K(t, \Gamma, \varepsilon)) = p_n(t, x, \Gamma)$.

Now, by (49), for any $K \geq 1$,

$$\sum_{n=0}^{+\infty} \tilde{\mathbf{P}}^K_{\frac{\gamma_K}{K}\delta_x}(A_n^K(t, \Gamma, \varepsilon)) \leq 1,$$

so, by the dominated convergence theorem,

$$\lim_{K\to+\infty} \tilde{\mathbf{P}}^K_{\frac{\gamma_K}{K}\delta_x}\Big(\text{Supp}(\nu^K_{t/Ku_K}) \text{ is a singleton } \{y\},\ y \in \Gamma \\ \text{and } |\langle \nu^K_{t/Ku_K}, \mathbf{1}\rangle - \bar{n}_y| < \varepsilon\Big) = \sum_{n \geq 0} p_n(t, x, \Gamma) = \mathbf{P}_x(X_t \in \Gamma),$$

which is (11) in the case of a single time t.

In order to complete the proof of Theorem 1, we have to generalize this limit to any sequence of times $0 < t_1 < \ldots < t_n$.

We will specify the method only in the case of two times $0 < t_1 < t_2$. It can be easily generalized to a sequence of n times. We introduce for any integers

$0 \leq n_1 \leq n_2$ the probabilities

$$\begin{aligned}
&p^K_{n_1,n_2}(t_1,t_2,x,\Gamma_1,\Gamma_2,\varepsilon,z) \\
&:= \mathbf{P}^K_{\frac{z}{K}\delta_x}\Bigg(\theta_{n_1} \leq \frac{t_1}{Ku_K} < \tau_{n_1+1},\ V_{n_1}\in\Gamma_1,\ \sup_{s\in[\theta_{n_1},\tau_{n_1+1}]}|\langle\nu^K_s,\mathbf{1}\rangle - \bar{n}_{V_{n_1}}| < \varepsilon, \\
&\qquad \theta_{n_2} \leq \frac{t_2}{Ku_K} < \tau_{n_2+1},\ V_{n_2}\in\Gamma_2 \text{ and } \sup_{s\in[\theta_{n_2},\tau_{n_2+1}]}|\langle\nu^K_s,\mathbf{1}\rangle - \bar{n}_{V_{n_2}}| < \varepsilon\Bigg),
\end{aligned}$$

and

$$\begin{aligned}
&q^K_{0,n_2}(t_1,t_2,x,\Gamma_1,\Gamma_2,\varepsilon,z) \\
&\qquad := \mathbf{1}_{\{x\in\Gamma_1\}}\mathbf{P}^K_{\frac{z}{K}\delta_x}\Bigg(\frac{t_1}{Ku_K} < \tau_1,\ \sup_{s\in[\log K,\tau_1]}|\langle\nu^K_s,\mathbf{1}\rangle - \bar{n}_x| < \varepsilon, \\
&\qquad \theta_{n_2} \leq \frac{t_2}{Ku_K} < \tau_{n_2+1},\ V_{n_2}\in\Gamma_2 \text{ and } \sup_{s\in[\theta_{n_2},\tau_{n_2+1}]}|\langle\nu^K_s,\mathbf{1}\rangle - \bar{n}_{V_{n_2}}| < \varepsilon\Bigg).
\end{aligned}$$

Then, we can use a calculation very similar to the proof of Lemma 3 to prove that, as $K \to +\infty$, $p^K_{n_1,n_2}(t_1,t_2,x,\Gamma_1,\Gamma_2,\varepsilon,z_K)$ converges to a limit $p_{n_1,n_2}(t_1,t_2,x,\Gamma_1,\Gamma_2)$ independent of $\varepsilon \in]0,\infty]$, z_K and the limit $z > 0$ of z_K/K (with the restriction that z has to be equal to $\bar{n}_x$ if $n_1 = 0$), and that $\lim q^K_{0,n_2}(t_1,t_2,x,\Gamma_1,\Gamma_2,\varepsilon,z) = p_{0,n_2}(t_1,t_2,x,\Gamma_1,\Gamma_2)$, where

$$\begin{cases}
p_{0,n_2}(t_1,t_2,x,\Gamma_1,\Gamma_2) = \mathbf{1}_{\{x\in\Gamma_1\}}e^{-\beta(x)t_1}p_{n_2}(t_2-t_1,x,\Gamma_2); \\
p_{n_1+1,n_2+1}(t_1,t_2,x,\Gamma_1,\Gamma_2) \\
\qquad = \displaystyle\int_0^{t_1}\beta(x)e^{-\beta(x)s}\int_{\mathbb{R}^l}p_{n_1,n_2}(t_1-s,t_2-s,x+h,\Gamma_1,\Gamma_2)\kappa(x,dh)ds.
\end{cases}$$

As above, we obtain equation (11) for $n = 2$ by observing that the same relation holds for the TSS process X.

This completes the proof of Theorem 1. □

Proof of Lemma 3 First, let us prove that the convergence property of $p^K_n(t,x,\Gamma,\varepsilon,z_K)$ when $z_K \in \mathbb{N}$ in Lemma 3 (a) implies the convergence in probability of these quantities when z_K are random variables. Actually, if (z_K) is a sequence of random variables such that $z_K/K \xrightarrow{\mathcal{P}} z$, by Skorohod's Theorem, we can construct on an auxiliary probability space $\hat{\Omega}$ a sequence of random variables $(\hat{z}_K)$ such that $\mathcal{L}(\hat{z}_K) = \mathcal{L}(z_K)$ and $\hat{z}_K(\hat{\omega})/K \to z$ for any $\hat{\omega} \in \hat{\Omega}$. Then, $\lim p^K_n(t,x,\Gamma,\varepsilon,\hat{z}_K(\hat{\omega})) = p_n(t,x,\Gamma)$ for any $\hat{\omega} \in \hat{\Omega}$, which implies that $p^K_n(t,x,\Gamma,\varepsilon,z_K) \xrightarrow{\mathcal{P}} p_n(t,x,\Gamma)$. The same method applies to $q^K_0(t,x,\Gamma,\varepsilon,z_k)$.

We will prove Lemma 3 (a) and (b) by induction over $n \geq 0$.

First, when $t > 0$, it follows from the fact that $t/Ku_K > \log K$ for sufficiently large K, and from Lemma 1 (b) and (c), that

$$\lim_{K\to+\infty} q^K_0(t,x,\Gamma,\varepsilon,z_K) = \mathbf{1}_{\{x\in\Gamma\}}e^{-\beta(x)t},$$

and that, if $z = \bar{n}_x$,

$$\lim_{K\to+\infty} p^K_0(t,x,\Gamma,\varepsilon,z_K) = \mathbf{1}_{\{x\in\Gamma\}}e^{-\beta(x)t}.$$

Then, fix $n \geq 0$ and assume that Lemma 3 (a) holds for n. We intend to prove the convergence of $p^K_{n+1}(t,x,\Gamma,\varepsilon,z_K)$ to $p_{n+1}(t,x,\Gamma)$ such that

$$p_{n+1}(t,x,\Gamma) = \int_0^t \beta(x)e^{-\beta(x)s}\int_{\mathbb{R}^l} p_n(t-s,x+h,\Gamma)\kappa(x,dh)ds,$$

by applying the strong Markov property at time τ_1, in a similar way than when we obtained (47). However, the convergence of $p^K_n(t,x,\Gamma,\varepsilon,z_K)$ to $p_n(t,x,\Gamma)$ only holds for *non-random* t. Therefore, we will divide the time interval $[0,t]$ in a finite number of small intervals and use the Markov property at time τ_1 when τ_1 is in each of these intervals. Moreover, we will also use the Markov property at time θ_1 and, in order to be able to apply Lemma 2 (which holds for a *non-random* mutant trait y) after this time, we will use the fact that U_1 is independent of τ_1 and $\nu^K_{\tau_1-}$ and that $U_1 - x$ is a random variable with law $m(x,h)dh$.

Following this program, we can bound $p^K_{n+1}(t,x,\Gamma,\varepsilon,z_K)$ from above as follows: fix $\eta > 0$; using Lemma 1 (a) in the first inequality, for sufficiently large $k \geq 0$ and $K \geq 1$,

$$\begin{aligned}
p^K_{n+1}(t,x,\Gamma,\varepsilon,z_K) &\leq \mathbf{P}^K_{\frac{z_K}{K}\delta_x}\left(\theta_{n+1} \leq \frac{t}{Ku_K},\ \tau_{n+2} > \frac{t+2/2^k}{Ku_K},\ V_{n+1}\in\Gamma\right) + \eta \\
&\leq \sum_{i=0}^{\lceil t2^k\rceil -1} \mathbf{P}^K_{\frac{z_K}{K}\delta_x}\left(\frac{i}{2^kKu_K} \leq \tau_1 \leq \frac{i+1}{2^kKu_K},\ \theta_{n+1} \leq \frac{t}{Ku_K},\right. \\
&\qquad \left.\tau_{n+2} > \frac{t+2/2^k}{Ku_K} \text{ and } V_{n+1}\in\Gamma\right) + \eta \\
&\leq \sum_{i=0}^{\lceil t2^k\rceil -1} \mathbf{E}^K_{\frac{z_K}{K}\delta_x}\Bigg[\mathbf{1}_{\left\{\frac{i}{2^kKu_K}\leq\tau_1\leq\frac{i+1}{2^kKu_K}\right\}}\mathbf{P}^K_{\nu^K_{\tau_1-}+\frac{1}{K}\delta_{U_1}}\left(\theta_n \leq \frac{t-i/2^k}{Ku_K},\right. \\
&\qquad \left.\tau_{n+1} > \frac{t-(i-1)/2^k}{Ku_K} \text{ and } V_n\in\Gamma\right)\Bigg] + \eta \\
&\leq \sum_{i=0}^{\lceil t2^k\rceil -1} \mathbf{E}^K_{\frac{z_K}{K}\delta_x}\Bigg[\mathbf{1}_{\left\{\frac{i}{2^kKu_K}\leq\tau_1\leq\frac{i+1}{2^kKu_K}\right\}}\int_{\mathbb{R}^l}\mathbf{E}^K_{\nu^K_{\tau_1-}+\frac{1}{K}\delta_{x+h}}\Bigg(\mathbf{1}_{\left\{\theta_0\geq\frac{1}{2^kKu_K}\wedge\tau_1\right\}} \\
&\qquad + \mathbf{1}_{\left\{\theta_0<\frac{1}{2^kKu_K}\wedge\tau_1\right\}}\mathbf{P}^K_{\nu^K_{\theta_0}}\left(\theta_n\leq\frac{t-i/2^k}{Ku_K}<\tau_{n+1},\ V_n\in\Gamma\right)\Bigg)m(x,h)dh\Bigg] + \eta. \\
&\leq \sum_{i=0}^{\lceil t2^k\rceil -1} \mathbf{E}^K_{\frac{z_K}{K}\delta_x}\Bigg[\mathbf{1}_{\left\{\frac{i}{2^kKu_K}\leq\tau_1\leq\frac{i+1}{2^kKu_K}\right\}}\int_{\mathbb{R}^l}\mathbf{E}^K_{\nu^K_{\tau_1-}+\frac{1}{K}\delta_{x+h}}\Bigg(\mathbf{1}_{\left\{\theta_0\geq\frac{1}{2^kKu_K}\wedge\tau_1\right\}} \\
&\qquad + \mathbf{1}_{\left\{\theta_0<\frac{1}{2^kKu_K}\wedge\tau_1\right\}}p^K_n(t-i/2^k,V_0,\Gamma,\infty,K\langle\nu^K_{\theta_0},\mathbf{1}\rangle)\Bigg)m(x,h)dh\Bigg] + \eta. \qquad (50)
\end{aligned}$$

Now, since $\nu^K_{\tau_1-} = \langle\nu^K_{\tau_1-},\mathbf{1}\rangle\delta_x$, under $\mathbf{P}^K_{\nu^K_{\tau_1-}+\frac{1}{K}\delta_{x+h}}$, on the event $\{\theta_0 < \tau_1\}$,

$$\begin{aligned}
p^K_n(t-i/2^k,V_0,\Gamma,\infty,K\langle\nu^K_{\theta_0},\mathbf{1}\rangle) &= \mathbf{1}_{\{V_0=x\}}p^K_n(t-i/2^k,x,\Gamma,\infty,K\langle\nu^K_{\theta_0},\mathbf{1}\rangle) \\
&\quad + \mathbf{1}_{\{V_0=x+h\}}p^K_n(t-i/2^k,x+h,\Gamma,\infty,K\langle\nu^K_{\theta_0},\mathbf{1}\rangle). \qquad (51)
\end{aligned}$$

By Lemma 1 (b), $\nu^K_{\tau_1-} \xrightarrow{\mathcal{P}} \bar{n}_x\delta_x$ under $\mathbf{P}_{\frac{z_K}{K}\delta_x}$, so we can use Skorohod's Theorem to construct random variables $\hat{N}_K$ on an auxiliary probability space $\hat{\Omega}$ with the same law that $\langle \nu^K_{\tau_1-}, \mathbf{1}\rangle$ and converging to $\bar{n}_x$ for any $\hat{\omega} \in \hat{\Omega}$.

Fix $\hat{\omega} \in \hat{\Omega}$. Under $\mathbf{P}^K_{\hat{N}_K(\hat{\omega})\delta_x + \frac{1}{K}\delta_{x+h}}$, define

$$Z_1^K = \langle \nu^K_{\theta_0}, \mathbf{1}\rangle \mathbf{1}_{\{V_0=x,\ \theta_0<\tau_1\}} + \frac{\lceil K\bar{n}_x\rceil}{K}\mathbf{1}_{\{V_0\neq x\}\cup\{\theta_0\geq\tau_1\}}.$$

It follows from Lemma 2 (43) and (44), and from assumption (B) that, for Lebesgue almost every h, $Z_1^K \xrightarrow{\mathcal{P}} \bar{n}_x$, so the induction assumption yields that, under $\mathbf{P}^K_{\hat{N}_K(\hat{\omega})\delta_x + \frac{1}{K}\delta_{x+h}}$, when $K \to +\infty$,

$$p_n^K(t - i/2^k, x, \Gamma, \infty, KZ_1^K) \xrightarrow{\mathcal{P}} p_n(t - i/2^k, x, \Gamma).$$

Now, given two sequences of uniformly bounded random variables $(X_K)_{K\geq 1}$ and $(Y_K)_{K\geq 0}$ such that X_K and Y_K are defined on the same probability space for any $K \geq 1$, and such that, when $K \to +\infty$, X_K converges in probability to a constant C and $\lim_K \mathbf{E}(Y_K)$ exists, it is easy to prove that

$$\lim_{K\to+\infty} \mathbf{E}(X_K Y_K) = C \lim_{K\to+\infty} \mathbf{E}(Y_K). \tag{52}$$

Applying this with $X_K = p_n^K(t-i/2^k, x, \Gamma, \infty, KZ_1^K)$ and $Y_K = \mathbf{1}_{\{V_0=x,\ \theta_0<\tau_1\}}$, by Lemma 2 (42) and (43) and assumption (B), for Lebesgue almost any h, and for any $\hat{\omega} \in \hat{\Omega}$,

$$\lim_{K\to+\infty} \mathbf{E}^K_{\hat{N}_K(\hat{\omega})\delta_x + \frac{1}{K}\delta_{x+h}} \left(\mathbf{1}_{\{V_0=x,\ \theta_0<\tau_1\}} p_n^K(t - i/2^k, x, \Gamma, \infty, K\langle \nu^K_{\theta_0}, \mathbf{1}\rangle)\right)$$
$$= \left(1 - \frac{[f(x+h,x)]_+}{b(x+h)}\right) p_n(t - i/2^k, x, \Gamma).$$

Finally, we obtain that, for Lebesgue almost any h, under $\mathbf{P}^K_{\frac{z_K}{K}\delta_x}$,

$$\mathbf{E}^K_{\nu^K_{\tau_1-} + \frac{1}{K}\delta_{x+h}} \left(\mathbf{1}_{\{V_0=x,\ \theta_0<\tau_1\}} p_n^K(t - i/2^k, x, \Gamma, \infty, K\langle \nu^K_{\theta_0}, \mathbf{1}\rangle)\right)$$
$$\xrightarrow{\mathcal{P}} \left(1 - \frac{[f(x+h,x)]_+}{b(x+h)}\right) p_n(t - i/2^k, x, \Gamma). \tag{53}$$

Similarly, we can use Lemma 2 (41) and the random variable

$$Z_2^K = \langle \nu^K_{\theta_0}, \mathbf{1}\rangle \mathbf{1}_{\{V_0=x+h,\ \theta_0<\tau_1\}} + \bar{n}_{x+h}\mathbf{1}_{\{V_0\neq x+h\}\cup\{\theta_0\geq\tau_1\}}$$

to prove that, for Lebesgue almost any h, under $\mathbf{P}^K_{\frac{z_K}{K}\delta_x}$,

$$\mathbf{E}^K_{\nu^K_{\tau_1-} + \frac{1}{K}\delta_{x+h}} \left(\mathbf{1}_{\{V_0=x+h,\ \theta_0<\tau_1\}} p_n^K(t - i/2^k, x+h, \Gamma, \infty, K\langle \nu^K_{\theta_0}, \mathbf{1}\rangle)\right)$$
$$\xrightarrow{\mathcal{P}} \frac{[f(x+h,x)]_+}{b(x+h)} p_n(t - i/2^k, x+h, \Gamma). \tag{54}$$

Moreover, by Lemma 2 (43), for Lebesgue almost any h, under $\mathbf{P}^K_{(z_K/K)\delta_x}$,

$$\mathbf{P}^K_{\nu^K_{\tau_1-} + \frac{1}{K}\delta_{x+h}} \left(\theta_0 \geq \frac{1}{2^k K u_K} \wedge \tau_1\right) \xrightarrow{\mathcal{P}} 0. \tag{55}$$

Collecting these results together, applying (52) again, it follows from Lemma 1 (c) and (51) that, for Lebesgue almost any h,

$$\lim_{K\to+\infty}\mathbf{E}^K_{\frac{z_K}{K}\delta_x}\Bigg[\mathbf{1}_{\left\{\frac{i}{2^kKu_K}\leq\tau_1\leq\frac{i+1}{2^kKu_K}\right\}}\mathbf{E}^K_{\nu^K_{\tau_1-}+\frac{1}{K}\delta_{x+h}}\Bigg(\mathbf{1}_{\left\{\theta_0\geq\frac{1}{2^kKu_K}\wedge\tau_1\right\}}$$

$$+\mathbf{1}_{\left\{\theta_0<\frac{1}{2^kKu_K}\wedge\tau_1\right\}}p^K_n(t-i/2^k,V_0,\Gamma,\infty,K\langle\nu^K_{\theta_0},\mathbf{1}\rangle)\Bigg)\Bigg]$$

$$=\left(e^{-\beta(x)\frac{i}{2^k}}-e^{-\beta(x)\frac{i+1}{2^k}}\right)\Bigg[\frac{[f(x+h,x)]_+}{b(x+h)}p_n(t-i/2^k,x+h,\Gamma)$$

$$+\left(1-\frac{[f(x+h,x)]_+}{b(x+h)}\right)p_n(t-i/2^k,x,\Gamma)\Bigg].$$

Finally, taking the integral of both sides with respect to $m(x,h)dh$, the dominated convergence theorem and (50) yield

$$\limsup_{K\to+\infty}\ p^K_{n+1}(x,t,\Gamma,\varepsilon,z_K)$$

$$\leq\sum_{i=0}^{\lceil t2^k\rceil-1}\left(e^{-\beta(x)\frac{i}{2^k}}-e^{-\beta(x)\frac{i+1}{2^k}}\right)\int_{\mathbb{R}^l}p_n(t-i/2^k,x+h,\Gamma)\kappa(x,dh)+\eta.$$

Taking the limit $k\to+\infty$ first and then $\eta\to0$, it follows from the fact that

$$e^{-\beta(x)i/2^k}-e^{-\beta(x)(i+1)/2^k}=e^{-\beta(x)i/2^k}(\beta(x)/2^k+O(1/2^{2k}))$$

and from the convergence of Riemann sums that

$$\limsup_{K\to+\infty}\ p^K_{n+1}(x,t,\Gamma,\varepsilon,z_K)\leq\int_0^t\beta(x)e^{-\beta(x)s}\int_{\mathbb{R}^l}p_n(t-s,x+h,\Gamma)\kappa(x,dh)ds.$$

Using the same method than for (50), we can give a lower bound for p^K_n as follows: for any $\eta>0$, for sufficiently large $k\geq0$ and $K\geq1$,

$$p^K_{n+1}(t,x,\Gamma,\varepsilon,z_K)\geq\mathbf{P}^K_{\frac{z_K}{K}\delta_x}\Bigg(\theta_{n+1}\leq\frac{t}{Ku_K},\ \tau_{n+2}>\frac{t-2/2^k}{Ku_K},\ V_{n+1}\in\Gamma$$

$$\text{and}\sup_{s\in[\theta_{n+1},\tau_{n+2}]}|\langle\nu^K_s,\mathbf{1}\rangle-\bar{n}_{V_{n+1}}|<\varepsilon\Bigg)-\eta$$

$$\geq\sum_{i=0}^{\lfloor t2^k\rfloor-3}\mathbf{E}^K_{\frac{z_K}{K}\delta_x}\Bigg[\mathbf{1}_{\left\{\frac{i}{2^kKu_K}\leq\tau_1\leq\frac{i+1}{2^kKu_K}\right\}}\mathbf{P}^K_{\nu^K_{\tau_1-}+\frac{1}{K}\delta_{U_1}}\Bigg(\theta_n\leq\frac{t-(i+1)/2^k}{Ku_K},$$

$$\tau_{n+1}>\frac{t-(i+2)/2^k}{Ku_K},\ V_n\in\Gamma\text{ and}\sup_{s\in[\theta_n,\tau_{n+1}]}|\langle\nu^K_s,\mathbf{1}\rangle-\bar{n}_{V_n}|<\varepsilon\Bigg)\Bigg]-\eta$$

$$\geq\sum_{i=0}^{\lfloor t2^k\rfloor-3}\mathbf{E}^K_{\frac{z_K}{K}\delta_x}\Bigg[\mathbf{1}_{\left\{\frac{i}{2^kKu_K}\leq\tau_1\leq\frac{i+1}{2^kKu_K}\right\}}\int_{\mathbb{R}^l}\mathbf{E}^K_{\nu^K_{\tau_1-}+\frac{1}{K}\delta_{x+h}}\Bigg(\mathbf{1}_{\left\{\theta_0<\frac{1}{2^kKu_K}\wedge\tau_1\right\}}$$

$$p^K_n(t-(i+2)/2^k,V_0,\Gamma,\varepsilon,K\langle\nu^K_{\theta_0},\mathbf{1}\rangle)\Bigg)m(x,h)dh\Bigg]-\eta.$$

Then, using the same method as above, letting $K \to +\infty$, then $k \to +\infty$ and finally $\eta \to 0$,

$$\liminf_{K\to+\infty} p^K_{n+1}(x,t,\Gamma,\varepsilon,z_K) \geq \int_0^t \beta(x)e^{-\beta(x)s}\int_{\mathbb{R}^l} p_n(t-s,x+h,\Gamma)\kappa(x,dh)ds,$$

which completes the proof of Lemma 3 by induction. $\square$

Proof of Lemma 1 (a) Fix $\eta > 0$. By Theorem 2 (a) and (c), for any $K \geq 1$,

$$\langle \nu^K, \mathbf{1}\rangle \preceq Z^K,$$
$$\text{where} \quad \mathcal{L}(Z^K) = \mathbf{P}^K(2\bar{b}, 0, \underline{\alpha}, \langle \nu^K_0, \mathbf{1}\rangle + 1).$$

Since $\sup_K \mathbf{E}(\langle \nu^K_0, \mathbf{1}\rangle) < +\infty$, we can choose $M < +\infty$ such that

$$\sup_{K\geq 1} \mathbf{P}(\langle \nu^K_0, \mathbf{1}\rangle + 1 > M) < \eta/3.$$

Then, apply Theorem 3 (c) to $\mathbf{P}^K(2\bar{b}, 0, \underline{\alpha}, \langle \nu^K_0, \mathbf{1}\rangle + 1)$ with $C = [1, M]$, $\eta_2 = M$ and η_1 such that $0 < 2\bar{b}/\underline{\alpha} - \eta_1 < 1/2$: there exists $V > 0$ such that

$$\limsup_{K\to+\infty} \mathbf{P}(T^K < e^{KV}) < \eta/3, \tag{56}$$
$$\text{where} \quad T^K = \inf\{t \geq 0, Z^K_t \notin [1/2, M + 2\bar{b}/\underline{\alpha}]\}.$$

Fix $t, \varepsilon > 0$. Since, for $s \leq T^K$, $\langle \nu^K_s, \mathbf{1}\rangle \leq M + 2\bar{b}/\underline{\alpha}$, if we apply Theorem 2 (b) to the process $(\nu^K_{s+(t/Ku_K)} - \nu^K_{t/Ku_K},\ s \geq 0)$, we obtain, for $s \leq T^K - t/Ku_K$,

$$A^K_{t/Ku_K+s} - A^K_{t/Ku_K} \preceq B^K_s,$$

where A^K_s is the number of mutations occuring between 0 and s, and where B^K is a Poisson process with parameter $Ku_K\bar{b}(M + 2\bar{b}/\underline{\alpha})$. Therefore, combining (56) with the fact that $1/Ku_K \ll e^{KV}$ for sufficiently large K, we obtain that, for sufficiently large K

$$\begin{aligned}\mathbf{P}(A^K_{(t+\varepsilon)/Ku_K} - A^K_{t/Ku_K} \geq 1) &\leq \mathbf{P}(B^K_{\varepsilon/Ku_K} \geq 1) + 2\eta/3 \\ &= 1 - \exp(-\bar{b}(M + 2\bar{b}/\underline{\alpha})\varepsilon) + 2\eta/3,\end{aligned}$$

which can be made smaller than η if ε is sufficiently small. This ends the proof of (38). $\square$

Proof of Lemma 1 (b) Fix $\varepsilon > 0$. It follows from the formula (13) for ν^K that, for $t < \tau_1$, under $\mathbf{P}^K_{\frac{z_K}{K}\delta_x}$,

$$\nu^K_t = Z^K_t \delta_x,$$
$$\text{where} \quad \mathcal{L}(Z^K) = \mathbf{P}^K((1 - u_K\mu(x))b(x), d(x), \alpha(x,x), z_K/K).$$

Therefore, by Theorem 2 (c), for sufficiently large K such that $u_K < \varepsilon$ and for $t \leq \tau_1$,

$$Z^{K,1} \preceq \langle \nu^K, \mathbf{1}\rangle \preceq Z^{K,2}, \tag{57}$$
$$\text{where} \quad \mathcal{L}(Z^{K,1}) = \mathbf{P}^K((1-\varepsilon)b(x), d(x), \alpha(x,x), z_K/K)$$
$$\text{and} \quad \mathcal{L}(Z^{K,2}) = \mathbf{P}^K(b(x), d(x), \alpha(x,x), z_K/K).$$

Now, let ϕ_y^1, resp. ϕ_y^2, be the solution to

$$\begin{aligned}
&\dot{\phi} = ((1-\varepsilon)b(x) - d(x) - \alpha(x,x)\phi)\phi, \\
\text{resp.} \quad &\dot{\phi} = (b(x) - d(x) - \alpha(x,x)\phi)\phi,
\end{aligned}$$

with initial state y, and observe that, for any $y > 0$, when $t \to +\infty$, $\phi_y^1(t) \to e^1 := \bar{n}_x - \varepsilon b(x)/\alpha(x,x)$ and $\phi_y^2(t) \to e^2 := \bar{n}_x$.

Define, for any $y > 0$, $t_\varepsilon^{i,y}$ the first time such that $\forall s \geq t_\varepsilon^{i,y}$, $\phi_y^i(s) \in [e^i - \varepsilon, e^i + \varepsilon]$ $(i = 1, 2)$. Because of the continuity of the flow of these ODEs (see [15] p. 356),

$$t_\varepsilon^i := \sup_{y \in [z/2, 2z]} t_\varepsilon^{i,y} < +\infty.$$

Let us apply Theorem 3 (a) to $Z^{K,1}$ and $Z^{K,2}$ on $[0, t_\varepsilon]$, where $t_\varepsilon = t_\varepsilon^1 \vee t_\varepsilon^2$: since $z_K/K \to z$, for sufficiently small $\delta > 0$, and for $i = 1, 2$,

$$\lim_{K \to +\infty} \mathbf{P}\left(\sup_{0 \leq t \leq t_\varepsilon} |Z_t^{K,i} - \phi_{z_K/K}^i(t)| > \delta \right) = 0.$$

If we choose $\delta < \varepsilon$, we obtain, for $i = 1, 2$,

$$\lim_{K \to +\infty} \mathbf{P}(|Z_{t_\varepsilon}^{K,i} - e^i| < 2\varepsilon) = 1,$$

and so, because of the expression of e^i, for $i = 1, 2$,

$$\lim_{K \to +\infty} \mathbf{P}(|Z_{t_\varepsilon}^{K,i} - \bar{n}_x| < M\varepsilon) = 1, \tag{58}$$

where $M = 2 + b(x)/\alpha(x,x)$.

Now, assuming ε sufficiently small for $(M+1)\varepsilon < \bar{n}_x$, define the stopping times

$$T_\varepsilon^{K,i} = \inf\{t \geq t_\varepsilon : |Z_t^{K,i} - \bar{n}_x| > (M+1)\varepsilon\}$$

for $i = 1, 2$, and $T_\varepsilon^K = T_\varepsilon^{K,1} \wedge T_\varepsilon^{K,2}$.

For any $z \in \mathbb{N}/K$, define also

$$\mathbf{P}_z^{K,1} := \mathbf{P}^K((1-\varepsilon)b(x), d(x), \alpha(x,x), z).$$

Then, applying Theorem 3 (c) to $\mathbf{P}_z^{K,1}$ with $C = [\bar{n}_x - M\varepsilon, \bar{n}_x + M\varepsilon]$, $\eta_1 = 3\varepsilon$ and $\eta_2 = (2M-1)\varepsilon$, there exists $V_1 > 0$ such that

$$\lim_{K \to +\infty} \inf_{z \in C} \mathbf{P}_z^{K,1}(\hat{T}_\varepsilon > e^{KV_1}) = 1, \tag{59}$$

$$\text{where} \quad \hat{T}_\varepsilon = \inf\{t \geq 0 : |w_t - \bar{n}_x| > (M+1)\varepsilon\}.$$

Since, by the Markov property, for any $K \geq 1$,

$$\mathbf{P}(T_\varepsilon^{K,1} > e^{KV_1} + t_\varepsilon) = \mathbf{E}\left(\mathbf{P}_{Z_{t_\varepsilon}^{K,1}}^{K,1}(\hat{T}_\varepsilon > e^{KV_1}) \right),$$

it follows from (58) that

$$\lim_{K \to +\infty} \mathbf{P}(T_\varepsilon^{K,1} > e^{KV_1} + t_\varepsilon) = 1.$$

Similarly, applying Theorem 3 (c) to $\mathbf{P}^K(b(x), d(x), \alpha(x,x), y)$ with $C = [\bar{n}_x - M\varepsilon, \bar{n}_x + M\varepsilon]$ and $\eta_1 = \eta_2 = (M+1)\varepsilon$, there exists $V_2 > 0$ such that

$$\lim_{K\to+\infty} \mathbf{P}(T_\varepsilon^{K,2} > e^{KV_2} + t_\varepsilon) = 1$$

for $i = 1, 2$, and for some constants $V_1, V_2 > 0$.

Therefore,

$$\lim_{K\to+\infty} \mathbf{P}(T_\varepsilon^K > e^{KV}) = 1 \tag{60}$$

where $V := V_1 \wedge V_2$.

Now, because of (57),

$$\forall t \in [t_\varepsilon, T_\varepsilon^K \wedge \tau_1], \quad |\langle \nu_s^K, \mathbf{1}\rangle - \bar{n}_x| < (M+1)\varepsilon. \tag{61}$$

Therefore, since $\log K > t_\varepsilon$ for sufficiently large K, in order to complete the proof of (39), it suffices to show that

$$\lim_{K\to+\infty} \mathbf{P}(\tau_1 < T_\varepsilon^K) = 1. \tag{62}$$

If we denote by A_t^K the number of mutations occuring between t_ε and $t + t_\varepsilon$, by Theorem 2 (b), for t such that $t_\varepsilon + t \le T_\varepsilon^K \wedge \tau_1$,

$$B^K \preceq A^K,$$

where B^K is a Poisson process with parameter $Ku_K(\bar{n}_x - (M+1)\varepsilon)\mu(x)b(x)$.

Therefore, if we denote by S^K the first time when $B_t^K = 1$, on the event $\{t_\varepsilon + S^K < T_\varepsilon^K\}$,

$$\tau_1 \le t_\varepsilon + S^K.$$

Since $\exp(-KV) \ll Ku_K$, $\lim_K \mathbf{P}(t_\varepsilon + S^K < e^{KV}) = 1$, and hence, by (60),

$$\lim_{K\to+\infty} \mathbf{P}(t_\varepsilon + S^K < T_\varepsilon^K) = 1,$$

which implies (62).

In the case where $z_K/K \to \bar{n}_x$, using (59) as above, we obtain easily

$$\lim_{K\to+\infty} \mathbf{P}(S_\varepsilon^K > e^{KV}) = 1,$$

$$\text{where} \quad S_\varepsilon^K = \inf\{t \ge 0 : |Z_t^{K,i} - \bar{n}_x| > (M+1)\varepsilon,\ i = 1, 2\}.$$

Then, the proof of (40) can be completed using a method similar to the one we used above. □

Proof of Lemma 1 (c) Fix $t > 0$ and $\varepsilon > 0$. Take K large enough for $\log K < t/Ku_K$. The Markov property at time $\log K$ for ν^K yields

$$\begin{aligned}
\mathbf{P}^K_{\frac{z_K}{K}\delta_x}&\left(\tau_1 > \frac{t}{Ku_K},\ \sup_{t\in[\log K, \tau_1]} |\langle \nu_t^K, \mathbf{1}\rangle - \bar{n}_x| < \varepsilon\right) \\
&= \mathbf{E}^K_{\frac{z_K}{K}\delta_x}\Bigg[\mathbf{1}_{\{\tau_1 > \log K\}} \mathbf{P}^K_{\nu^K_{\log K}}\Bigg(\tau_1 > \frac{t}{Ku_K} - \log K, \\
&\qquad \sup_{t\in[0,\tau_1]} |\langle \nu_t^K, \mathbf{1}\rangle - \bar{n}_x| < \varepsilon\Bigg)\Bigg]. \qquad (63)
\end{aligned}$$

For any initial condition $\nu_0^K = \langle \nu_0^K, \mathbf{1}\rangle \delta_x$ of ν^K, by Theorem 2 (b), the number A_t^K of mutations of ν^K between 0 and t satisfies, for any $t \leq \tau_1$ such that $\sup_{s\in[0,t]} |\langle \nu_s^K, \mathbf{1}\rangle - \bar{n}_x| < \varepsilon$,

$$B^K \preceq A^K \preceq C^K,$$

where B_t^K and C_t^K are Poisson processes with respective parameters $Ku_K(\bar{n}_x - \varepsilon)\mu(x)b(x)$ and $Ku_K(\bar{n}_x + \varepsilon)\mu(x)b(x)$.

Therefore, on the event $\{\sup_{s\in[0,\tau_1]} |\langle \nu_s^K, \mathbf{1}\rangle - \bar{n}_x| < \varepsilon\}$, $S^K \leq \tau_1 \leq T^K$, where T^K is the first time when $B_t^K = 1$, and S^K the first time when $C_t^K = 1$.

Now, by Lemma 1 (b), under $\mathbf{P}^K_{(z_K/K)\delta_x}$, $\nu^K_{\log K} \xrightarrow{\mathcal{P}} \bar{n}_x\delta_x$, so, by Skorohod's Theorem, we can construct $\hat{N}^K$ with the same law as $\langle \nu^K_{\log K}, \mathbf{1}\rangle$ on an auxiliary probability space $\hat{\Omega}$ such that $\hat{N}^K(\hat{\omega}) \to \bar{n}_x$ for any $\hat{\omega} \in \hat{\Omega}$. Fix $\hat{\omega} \in \hat{\Omega}$. Then, by Lemma 1 (b),

$$\lim_{K\to+\infty} \mathbf{P}^K_{\hat{N}(\hat{\omega})\delta_x}\left(\sup_{t\in[0,\tau_1]} |\langle \nu_t^K, \mathbf{1}\rangle - \bar{n}_x| < \varepsilon\right) = 1,$$

and so,

$$\begin{aligned}
&\limsup_{K\to+\infty} \mathbf{P}^K_{\hat{N}(\hat{\omega})\delta_x}\left(\tau_1 > \frac{t}{Ku_K} - \log K,\ \sup_{t\in[0,\tau_1]} |\langle \nu_t^K, \mathbf{1}\rangle - \bar{n}_x| < \varepsilon\right) \\
&\qquad \leq \limsup_{K\to+\infty} \mathbf{P}^K_{\hat{N}(\hat{\omega})\delta_x}\left(T^K > \frac{t}{Ku_K} - \log K\right) = \exp(-t(\bar{n}_x - \varepsilon)\mu(x)b(x)).
\end{aligned}$$

Therefore, under $\mathbf{P}^K_{(z_K/K)\delta_x}$,

$$\begin{aligned}
\limsup_{K\to+\infty} \mathbf{P}^K_{\nu^K_{\log K}}&\left(\tau_1 > \frac{t}{Ku_K} - \log K,\ \sup_{t\in[0,\tau_1]} |\langle \nu_t^K, \mathbf{1}\rangle - \bar{n}_x| < \varepsilon\right) \\
&\qquad\qquad \leq \exp(-t(\bar{n}_x - \varepsilon)\mu(x)b(x))
\end{aligned}$$

in probability (where $\limsup X_n \leq a$ in probability means that, for any $\eta > 0$, $\mathbf{P}(X_n > a + \eta) \to 0$).

Similarly, under $\mathbf{P}^K_{(z_K/K)\delta_x}$,

$$\begin{aligned}
\liminf_{K\to+\infty} \mathbf{P}^K_{\nu^K_{\log K}}&\left(\tau_1 > \frac{t}{Ku_K} - \log K,\ \sup_{t\in[0,\tau_1]} |\langle \nu_t^K, \mathbf{1}\rangle - \bar{n}_x| < \varepsilon\right) \\
&\qquad\qquad \geq \exp(-t(\bar{n}_x + \varepsilon)\mu(x)b(x))
\end{aligned}$$

in probability.

Now, by Lemma 1 (a) and (b),

$$\begin{aligned}
&\lim_{K\to+\infty} \mathbf{P}^K_{\frac{z_K}{K}\delta_x}(\tau_1 > \log K) = 1 \\
\text{and}\quad &\lim_{K\to+\infty} \mathbf{P}^K_{\frac{z_K}{K}\delta_x}\left(\sup_{t\in[\log K,\tau_1]} |\langle \nu_t^K, \mathbf{1}\rangle - \bar{n}_x| < \varepsilon\right) = 1.
\end{aligned}$$

So, using results similar to (52), it follows from (63) that

$$\begin{aligned}
&\limsup_{K\to+\infty} \mathbf{P}^K_{\frac{z_K}{K}\delta_x}\left(\tau_1 > \frac{t}{Ku_K}\right) \leq \exp(-t(\bar{n}_x - \varepsilon)\mu(x)b(x)) \\
\text{and}\quad &\liminf_{K\to+\infty} \mathbf{P}^K_{\frac{z_K}{K}\delta_x}\left(\tau_1 > \frac{t}{Ku_K}\right) \geq \exp(-t(\bar{n}_x + \varepsilon)\mu(x)b(x)).
\end{aligned}$$

Since this holds for any $\varepsilon > 0$, we have completed the proof of Lemma 1 (c). $\square$

Proof of Lemma 2 The proof of this lemma follows the three steps of the invasion of a mutant described at the end of the introduction.

Fix $\eta > 0$, $\varepsilon_0 > 0$ and $0 < \varepsilon < \varepsilon_0$. By Lemma 1 (a), there exists a constant $\rho > 0$ that we can assume smaller than η, such that, for sufficiently large K,

$$\mathbf{P}^K_{\frac{z_K}{K}\delta_x+\frac{1}{K}\delta_y}\left(\tau_1 < \frac{\rho}{Ku_K}\right) < \varepsilon. \tag{64}$$

Observe that, under $\mathbf{P}^K_{\frac{z_K}{K}\delta_x+\frac{1}{K}\delta_y}$, for $t \leq \tau_1$,

$$\begin{aligned}\mathcal{L}((\langle\nu^K,\mathbf{1}_{\{x\}}\rangle,\langle\nu^K,\mathbf{1}_{\{y\}}\rangle)) = \mathbf{Q}^K((1-u_K\mu(x))b(x),(1-u_K\mu(y))b(y),\\ d(x),d(y),\alpha(x,x),\alpha(x,y),\alpha(y,x),\alpha(y,y),z_K/K,1/K).\end{aligned}$$

Fix K large enough for $u_K < \varepsilon$. Define

$$S^K_\varepsilon := \inf\{s \geq 0 : \langle\nu^K_s,\mathbf{1}_{\{y\}}\rangle \geq \varepsilon\}$$

By Theorem 2 (c) and (d), for $t < \tau_1 \wedge S^K_\varepsilon$,

$$\begin{aligned} &Z^{K,1} \preceq \langle\nu^K,\mathbf{1}_{\{x\}}\rangle \preceq Z^{K,2}, \\ \text{where}\quad &\mathcal{L}(Z^{K,1}) = \mathbf{P}^K((1-\varepsilon)b(x),d(x)+\varepsilon\alpha(x,y),\alpha(x,x),z_K/K) \\ \text{and}\quad &\mathcal{L}(Z^{K,2}) = \mathbf{P}^K(b(x),d(x),\alpha(x,x),z_K/K).\end{aligned} \tag{65}$$

Using exactly the same method than led us to (60), we can deduce from Theorem 3 (c) that there exists $V > 0$ such that

$$\begin{aligned}&\lim_{K\to+\infty}\mathbf{P}(R^K_\varepsilon > e^{KV}) = 1, \\ \text{where}\quad &R^K_\varepsilon = \inf\{t \geq 0 : |Z^{K,i}_t - \bar{n}_x| > M\varepsilon,\ i=1,2\},\end{aligned} \tag{66}$$

with $M = 3 + (b(x)+\alpha(x,y))/\alpha(x,x)$.

Now, observe that, by (65),

$$\forall t \leq \tau_1 \wedge S^K_\varepsilon \wedge R^K_\varepsilon,\quad \langle\nu^K_t,\mathbf{1}_{\{x\}}\rangle \in [\bar{n}_x - M\varepsilon, \bar{n}_x + M\varepsilon].$$

Therefore, by Theorem 2 (b) and (d), for $t \leq \tau_1 \wedge S^K_\varepsilon \wedge R^K_\varepsilon$

$$\begin{aligned} &Z^{K,3}_t \preceq \langle\nu^K_t,\mathbf{1}_{\{y\}}\rangle \preceq Z^{K,4}_t, \quad\text{where} \\ &\mathcal{L}(Z^{K,3}) = \mathbf{P}^K((1-\varepsilon)b(y),d(y)+(\bar{n}_x+M\varepsilon)\alpha(y,x)+\varepsilon\alpha(y,y),0,1/K) \\ \text{and}\quad &\mathcal{L}(Z^{K,4}) = \mathbf{P}^K(b(y),d(y)+(\bar{n}_x-M\varepsilon)\alpha(y,x),0,1/K).\end{aligned} \tag{67}$$

Define, for any $K \geq 1$, $n \in \mathbb{N}$ and $i \in \{3,4\}$, the stopping time

$$T^{K,i}_{n/K} = \inf\{t \geq 0 : Z^{K,i}_t = n/K\}.$$

Observe that, if $S^K_\varepsilon < \tau_1 \wedge R^K_\varepsilon$,

$$T^{K,4}_{\lceil\varepsilon K\rceil/K} \leq S^K_\varepsilon \leq T^{K,3}_{\lceil\varepsilon K\rceil/K} \tag{68}$$

and that, if $T_0^{K,4} < T^{K,4}_{\lceil \varepsilon K\rceil/K} \wedge \tau_1 \wedge R^K_\varepsilon$,

$$\theta_0 \leq T_0^{K,4}.$$

If $Z^{K,4}$ is sub-critical, apply Theorem 4 (27), and if $Z^{K,4}$ is super-critical, apply Theorem 4 (30) (the critical case can be excluded by slightly changing the value of ε). Since $\log K \ll 1/Ku_K$, we obtain

$$\begin{aligned}\lim_{K\to+\infty} \mathbf{P}\left(T_0^{K,4} \leq \frac{\rho}{Ku_K} \wedge T^{K,4}_{\lceil \varepsilon K\rceil/K}\right) \\ = \frac{d(y)+(\bar n_x - M\varepsilon)\alpha(y,x)}{b(y)} \wedge 1 \geq 1 - \frac{[f(y,x)]_+}{b(y)} - \frac{\alpha(y,x)}{b(y)}M\varepsilon. \quad (69)\end{aligned}$$

Combining (64), (66), (67) and (69), and using the facts that $\rho < \eta$, $\varepsilon < \varepsilon_0$ and $\exp(KV) > \rho/Ku_K$ for sufficiently large K, we obtain, taking K larger if necessary,

$$\begin{aligned}\mathbf{P}\bigg(\theta_0 < \tau_1 \wedge \frac{\eta}{Ku_K},\ V_0 = x \text{ and } |\langle \nu^K_{\theta_0}, \mathbf{1}\rangle - \bar n_x| < M\varepsilon_0\bigg) \\ \geq \mathbf{P}\left(\theta_0 < \tau_1 \wedge S^K_\varepsilon \wedge R^K_\varepsilon \wedge \frac{\rho}{Ku_K} \text{ and } V_0 = x\right) \\ \geq \mathbf{P}\left(T_0^{K,4} < \tau_1 \wedge T^{K,4}_{\lceil \varepsilon K\rceil/K} \wedge R^K_\varepsilon \wedge \frac{\rho}{Ku_K}\right) \\ \geq 1 - \frac{[f(y,x)]_+}{b(y)} - \left(\frac{\alpha(y,x)}{b(y)}M + 3\right)\varepsilon. \quad (70)\end{aligned}$$

This ends the proof of Lemma 2 in the case where $f(y,x) \leq 0$.

Let us assume that $f(y,x) > 0$, *i.e.* that $b(y) - d(y) - \bar n_x \alpha(y,x) > 0$. If we choose $\varepsilon > 0$ sufficiently small, then $Z^{K,3}$ is super-critical. By Theorem 4 (31),

$$\begin{aligned}\lim_{K\to+\infty} \mathbf{P}\left(T^{K,3}_{\lceil \varepsilon K\rceil/K} < \frac{\rho}{3Ku_K}\right) \\ = \frac{(1-\varepsilon)b(y) - d(y) - (\bar n_x + M\varepsilon)\alpha(y,x) - \varepsilon\alpha(y,y)}{(1-\varepsilon)b(y)} \\ \geq \frac{f(y,x)}{(1-\varepsilon)b(y)} - \varepsilon\frac{b(y)+M\alpha(y,x)+\alpha(y,y)}{(1-\varepsilon)b(y)}.\end{aligned}$$

Therefore, by (66) and (64), assuming (without loss of generality) that $\varepsilon < 1/2$, for sufficiently large K,

$$\mathbf{P}\left(T^{K,3}_{\lceil \varepsilon K\rceil/K} < \tau_1 \wedge R^K_\varepsilon \wedge \frac{\rho}{3Ku_K}\right) \geq \frac{f(y,x)}{(1-\varepsilon)b(y)} - M'\varepsilon,$$

where $M' := 2(b(y)+M\alpha(y,x)+\alpha(y,y))/b(y)+3$. Then, it follows from (68) that

$$\mathbf{P}\left(S^K_\varepsilon < \tau_1 \wedge R^K_\varepsilon \wedge \frac{\rho}{3Ku_K}\right) \geq \frac{f(y,x)}{(1-\varepsilon)b(y)} - M'\varepsilon. \quad (71)$$

Observe that, on the event $\{S^K_\varepsilon < \tau_1 \wedge R^K_\varepsilon \wedge (\rho/3Ku_K)\}$,

$$\langle \nu^K_{S^K_\varepsilon}, \mathbf{1}_{\{y\}}\rangle = \lceil \varepsilon K\rceil/K \quad \text{and} \quad |\langle \nu^K_{S^K_\varepsilon}, \mathbf{1}_{\{x\}}\rangle - \bar n_x| < M\varepsilon. \quad (72)$$

Now, since we have assumed $f(y,x) > 0$, x and y satisfy (8) and, by Proposition 2, any solution to (6) with initial state in the compact set $[\bar{n}_x - M\varepsilon, \bar{n}_x + M\varepsilon] \times [\varepsilon/2, 2\varepsilon]$ converges to $(0, \bar{n}_y)$ when $t \to +\infty$. As in the proof of Lemma 1 (b), because of the continuity of the flow of system (6), we can find $t_\varepsilon < +\infty$ large enough such that any of these solutions do not leave the set $[0, \varepsilon^2/2] \times [\bar{n}_y - \varepsilon/2, \bar{n}_y + \varepsilon/2]$ after time t_ε.

Apply Theorem 3 (b) on $[0, t_\varepsilon]$, with $C = [\bar{n}_x - M\varepsilon, \bar{n}_x + M\varepsilon] \times [\varepsilon/2, 2\varepsilon]$ and with a constant $\delta < \varepsilon^2/2 \wedge r$, where r is defined in (21) (with $T = t_\varepsilon$). Then, with the notations of Theorem 3 (b), because of (71) and (72), the Markov property at time S_ε^K yields

$$\liminf_{K\to+\infty} \mathbf{P}\bigg(S_\varepsilon^K < \tau_1 \wedge R_\varepsilon^K \wedge \frac{\rho}{3Ku_K}, \\ \sup_{S_\varepsilon^K \le s \le S_\varepsilon^K + t_\varepsilon} \big\| \big(\langle \nu_s^K, \mathbf{1}_{\{x\}} \rangle, \langle \nu_s^K, \mathbf{1}_{\{y\}} \rangle \big) - \phi_{\langle \nu^K_{S_\varepsilon^K}, \mathbf{1}_{\{x\}} \rangle, \langle \nu^K_{S_\varepsilon^K}, \mathbf{1}_{\{y\}} \rangle}(s) \big\| \le \delta \bigg) \\ \ge \frac{f(y,x)}{(1-\varepsilon)b(y)} - M'\varepsilon. \quad (73)$$

Now, observe that, since $\delta < r$, on the event

$$\bigg\{ S_\varepsilon^K < \tau_1 \wedge R_\varepsilon^K, \\ \sup_{S_\varepsilon^K \le s \le S_\varepsilon^K + t_\varepsilon} \big\| \big(\langle \nu_s^K, \mathbf{1}_{\{x\}} \rangle, \langle \nu_s^K, \mathbf{1}_{\{y\}} \rangle \big) - \phi_{\langle \nu^K_{S_\varepsilon^K}, \mathbf{1}_{\{x\}} \rangle, \langle \nu^K_{S_\varepsilon^K}, \mathbf{1}_{\{y\}} \rangle}(s) \big\| \le \delta \bigg\},$$

for any $t \in [S_\varepsilon^K, S_\varepsilon^K + t_\varepsilon]$, $\langle \nu_t^K, \mathbf{1}_{\{x\}} \rangle \ge r - \delta > 0$ and $\langle \nu_t^K, \mathbf{1}_{\{y\}} \rangle \ge r - \delta > 0$, and thus

$$\theta_0 > S_\varepsilon^K + t_\varepsilon.$$

Therefore, since $\delta < \varepsilon^2/2 < \varepsilon/2$, by (64) and (73), for sufficiently large K,

$$\mathbf{P}\bigg(S_\varepsilon^K < R_\varepsilon^K \wedge \frac{\rho}{3Ku_K},\ \tau_1 > \frac{\rho}{3Ku_K} + t_\varepsilon,\ \theta_0 > S_\varepsilon^K + t_\varepsilon, \\ \langle \nu^K_{S_\varepsilon^K + t_\varepsilon}, \mathbf{1}_{\{x\}} \rangle < \varepsilon^2 \text{ and } \langle \nu^K_{S_\varepsilon^K + t_\varepsilon}, \mathbf{1}_{\{y\}} \rangle \in [\bar{n}_y - \varepsilon, \bar{n}_y + \varepsilon] \bigg) \\ \ge \frac{f(y,x)}{(1-\varepsilon)b(y)} - (M'+2)\varepsilon. \quad (74)$$

Now, we will compare $\langle \nu^K, \mathbf{1}_{\{x\}} \rangle$ with a branching process after time $S_\varepsilon^K + t_\varepsilon$ in order to prove that trait x gets extinct with a very high probability. We will use a method very similar to the one we used in the beginning of the proof of Lemma 2. First, on the event inside the probability in (74), $\langle \nu^K_{S_\varepsilon^K + t_\varepsilon}, \mathbf{1}_{\{x\}} \rangle < \varepsilon^2$. In order to prove that the population with trait x stays small after $S_\varepsilon^K + t_\varepsilon$, since $\varepsilon^2 < \varepsilon$ ($\varepsilon < 1/2$), let us define the stopping time

$$\hat{S}_\varepsilon^K = \inf\{ t \ge S_\varepsilon^K + t_\varepsilon : \langle \nu_t^K, \mathbf{1}_{\{x\}} \rangle > \varepsilon \}.$$

Using Theorem 2 (c) and (d) again, we see that, on the event

$$F^{K,\varepsilon} := \big\{ \langle \nu^K_{S_\varepsilon^K + t_\varepsilon}, \mathbf{1}_{\{x\}} \rangle < \varepsilon^2,\ \langle \nu^K_{S_\varepsilon^K + t_\varepsilon}, \mathbf{1}_{\{y\}} \rangle \in [\bar{n}_y - \varepsilon, \bar{n}_y + \varepsilon] \big\},$$

for any $t \geq 0$ such that $S_\varepsilon^K + t_\varepsilon + t \leq \hat{S}_\varepsilon^K \wedge \tau_1$,

$$Z_t^{K,5} \preceq \langle \nu^K_{S_\varepsilon^K+t_\varepsilon+t}, \mathbf{1}_{\{y\}}\rangle \preceq Z_t^{K,6},$$
$$\text{where} \quad \mathcal{L}(Z^{K,5}) = \mathbf{P}^K((1-\varepsilon)b(y), d(y)+\varepsilon\alpha(y,x), \alpha(y,y), \lfloor(\bar{n}_y-\varepsilon)K\rfloor/K)$$
$$\text{and} \quad \mathcal{L}(Z^{K,6}) = \mathbf{P}^K(b(y), d(y), \alpha(y,y), \lceil(\bar{n}_y+\varepsilon)K\rceil/K).$$

We can apply Theorem 3 (c) to $Z^{K,5}$ and $Z^{K,6}$ as above to obtain a constant $V' > 0$ such that

$$\lim_{K\to+\infty} \mathbf{P}(\hat{R}_\varepsilon^K > e^{KV'}) = 1, \tag{75}$$
$$\text{where} \quad \hat{R}_\varepsilon^K = \inf\{t \geq 0 : |Z_t^{K,i} - \bar{n}_y| > M''\varepsilon, i = 5, 6\},$$

with $M'' = 3 + (b(y) + \alpha(y,x))/\alpha(y,y)$.

Observe that, on the event $F^{K,\varepsilon}$, for any $t \leq \hat{R}_\varepsilon^K$ such that $S_\varepsilon^K + t_\varepsilon + t \leq \hat{S}_\varepsilon^K \wedge \tau_1$,

$$|\langle \nu^K_{S_\varepsilon^K+t_\varepsilon+t}, \mathbf{1}_{\{y\}}\rangle - \bar{n}_y| \leq M''\varepsilon,$$

and so, by Theorem 2 (c) and (e), on $F^{K,\varepsilon}$ and for t as above,

$$\langle \nu^K_{S_\varepsilon^K+t_\varepsilon+t}, \mathbf{1}_{\{x\}}\rangle \preceq Z_t^{K,7}$$
$$\text{where} \quad \mathcal{L}(Z^{K,7}) = \mathbf{P}^K(b(x), d(x) + (\bar{n}_y - M''\varepsilon)\alpha(x,y), 0, \lceil\varepsilon^2 K\rceil/K).$$

Now, since x and y satisfy (8), $b(x) - d(x) - \bar{n}_y\alpha(x,y) < 0$, and thus $Z^{K,7}$ is sub-critical for sufficiently small ε. Fix such an $\varepsilon > 0$ and define for any $n \geq 0$

$$\hat{T}^K_{n/K} = \inf\{t \geq 0 : Z_t^{K,7} = n/K\}.$$

If $\hat{T}^K_{\lceil\varepsilon K\rceil/K} \leq \hat{R}_\varepsilon^K$ and $S_\varepsilon^K + t_\varepsilon + \hat{T}^K_{\lceil\varepsilon K\rceil/K} \leq \tau_1$, then

$$\hat{S}_\varepsilon^K \geq S_\varepsilon^K + t_\varepsilon + \hat{T}^K_{\lceil\varepsilon K\rceil/K}$$

and if $\hat{T}_0^K \leq \hat{R}_\varepsilon^K$ and $S_\varepsilon^K + t_\varepsilon + \hat{T}_0^K \leq \hat{S}_\varepsilon^K \wedge \tau_1$, then

$$\theta_0 \leq \hat{T}_0^K.$$

Moreover, by Theorem 4 (28) and (29), for sufficiently large K,

$$\mathbf{P}\left(\hat{T}_0^K \leq \frac{\rho}{3Ku_K}\right) \geq 1 - \varepsilon$$
$$\text{and} \quad \mathbf{P}(\hat{T}^K_{\lceil K\varepsilon\rceil/K} \leq \hat{T}_0^K) \leq 2\varepsilon.$$

Combining the last two inequalities with (64), (74) and (75), and reminding that $\rho < \eta$ and $\varepsilon < \varepsilon_0$, we finally obtain, for sufficiently large K,

$$\begin{aligned}
&\mathbf{P}\left(\theta_0 < \tau_1 \wedge \frac{\eta}{Ku_K},\ V_0 = y \text{ and } |\langle \nu^K_{\theta_0}, \mathbf{1}\rangle - \bar{n}_y| < M''\varepsilon_0\right) \\
&\geq \mathbf{P}\Bigg(S_\varepsilon^K < R_\varepsilon^K \wedge \frac{\rho}{3Ku_K},\ \theta_0 > S_\varepsilon^K + t_\varepsilon,\ \tau_1 > \frac{2\rho}{3Ku_K} + t_\varepsilon,\ \langle \nu^K_{S_\varepsilon^K+t_\varepsilon}, \mathbf{1}_{\{x\}}\rangle < \varepsilon^2, \\
&\qquad \langle \nu^K_{S_\varepsilon^K+t_\varepsilon}, \mathbf{1}_{\{y\}}\rangle \in [\bar{n}_y - \varepsilon, \bar{n}_y + \varepsilon],\ \hat{T}_0^K < \frac{\rho}{3Ku_K} \wedge \hat{T}^K_{\lceil K\varepsilon\rceil/K} \text{ and } \hat{R}_\varepsilon^K > \frac{\rho}{Ku_K}\Bigg) \\
&\geq \frac{f(y,x)}{(1-\varepsilon)b(y)} - (M'+7)\varepsilon.
\end{aligned}$$

Adding this inequality with (70), we obtain

$$\mathbf{P}\left(\theta_0 < \tau_1 \wedge \frac{\eta}{Ku_K}\right) \geq 1 - \frac{\varepsilon}{1-\varepsilon}\frac{f(y,x)}{b(y)} - \left(M\frac{\alpha(y,x)}{b(y)} + M' + 10\right)\varepsilon \geq 1 - M'''\varepsilon,$$

where $M''' = 2f(y,x)/b(y) + M\alpha(y,x)/b(y) + M' + 10$ (remind that $\varepsilon < 1/2$), which implies (43), and

$$\mathbf{P}\big(|\langle \nu^K_{\theta_0}, \mathbf{1}\rangle - \bar{n}_{V_0}| < (M \vee M'')\varepsilon_0\big) \geq 1 - M'''\varepsilon,$$

which implies (44).

Therefore,

$$\mathbf{P}(V_0 = x) \geq 1 - \frac{f(y,x)}{b(y)} - 2M'''\varepsilon \quad \text{and} \quad \mathbf{P}(V_0 = y) \geq \frac{f(y,x)}{(1-\varepsilon)b(y)} - 2M'''\varepsilon.$$

Since $\mathbf{P}(V_0 = x) \leq 1 - \mathbf{P}(V_0 = y)$, we finally obtain (41) and (42). □

References

[1] Athreya, K. B. and Ney, P. E. *Branching Processes.* Springer-Verlag, Berlin (1972).

[2] Champagnat, N. *Convergence and existence for polymorphic adaptive dynamics jump and degenerate diffusion models.* Preprint Université Paris X – Nanterre No. 03/7 (2003).

[3] Dieckmann, U. *Coevolutionary Dynamics of Stochastic Replicator Systems.* Master Thesis (internal report Jül-3018), University of Jülich (1995).

[4] Dieckmann, U. and Law, R. *The dynamical theory of coevolution: a derivation from stochastic ecological processes.* J. Math. Biol. **34**, 579-612 (1996).

[5] Dupuis, P. and Ellis, R. S. *A Weak Convergence Approach to the Theory of Large Deviations.* Wiley, New York (1997).

[6] Ethier, S. N. and Kurtz, T. G. *Markov Processes, characterization and convergence.* John Wiley & Sons, New York (1986).

[7] Ferrière, R., Méléard, S., Fournier, N. and Champagnat, N. *The mathematics of Darwinian evolution: from stochastic individual processes to adaptive dynamics.* Preprint Université Paris 10 – Nanterre No. 04/1 (2004).

[8] Fournier, N. and Méléard, S. *A microscopic probabilistic description of a locally regulated population and macroscopic approximations.* To appear in Ann. Appl. Probab. (2004).

[9] Freidlin, M. I. and Wentzell, A. D. *Random Perturbations of Dynamical Systems.* Springer-Verlag, Berlin (1984).

[10] Hofbauer, J. and Sigmund, R. *Adaptive dynamics and evolutionary stability.* Appl. Math. Letters **3**, 75-79 (1990).

[11] Istas, J. *Introduction aux modélisations mathématiques pour les sciences du vivant.* Springer (2000).

[12] Marrow, P., Law, R. and Cannings, C. *The coevolution of predator-prey interactions: ESSs and Red Queen dynamics.* Proc. R. Soc. Lond. B **250**, 133-141 (1992).

[13] Metz, J. A. J., Nisbet, R. M. and Geritz, S. A. H. *How should we define 'fitness' for general ecological scenarios?* Trends in Ecology and Evolution **7**, 198-202 (1992).

[14] Metz, J. A. J., Geritz, S. A. H., Meszéna, G., Jacobs, F. A. J. and van Heerwaarden, J. S. *Adaptive Dynamics, a geometrical study of the consequences of nearly faithful reproduction.* In S.J. van Strien & S.M. Verduyn Lunel (ed.), *Stochastic and Spatial Structures of Dynamical Systems.* North Holland, Amsterdam, 183-231 (1996).

[15] Queffélec, H. and Zuily, G. *Éléments d'Analyse.* Masson, Paris (1995).

[16] Skorohod, A. V., Hoppensteadt, F. C. and Salehi, H. *Random Perturbation Methods.* Springer-Verlag (2002).