HAL open science

Stochastic analysis of emergence of evolutionary cyclic behavior in population dynamics with transfer

Nicolas Champagnat, Sylvie Méléard, Viet Chi Tran

To cite this version:

Nicolas Champagnat, Sylvie Méléard, Viet Chi Tran. Stochastic analysis of emergence of evolutionary cyclic behavior in population dynamics with transfer. 2019. hal-01974289v1

HAL Id: hal-01974289

https://inria.hal.science/hal-01974289v1

Preprint submitted on 8 Jan 2019 (v1), last revised 1 Feb 2021 (v3)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution 4.0 International License

Stochastic analysis of emergence of evolutionary cyclic behavior in population dynamics with transfer

Nicolas Champagnat[*], Sylvie Méléard[†], Viet Chi Tran[‡]

January 8, 2019

Abstract

Horizontal gene transfer consists in exchanging genetic materials between microorganisms during their lives. This is a major mechanism of bacterial evolution and is believed to be of main importance in antibiotics resistance. We consider a stochastic model for the evolution of a discrete population structured by a trait taking finitely many values, with density-dependent competition. Traits are vertically inherited unless a mutation occurs, and can also be horizontally transferred by unilateral conjugation with frequency dependent rate. Our goal is to analyze the trade-off between natural evolution to higher birth rates and transfer, which drives the population towards lower birth rates. Simulations show that evolutionary outcomes include evolutionary suicide or cyclic re-emergence of small populations with well-adapted traits. We focus on a parameter scaling where individual mutations are rare but the global mutation rate tends to infinity. This implies that negligible sub-populations may have a strong contribution to evolution. Our main result quantifies the asymptotic dynamics of subpopulation sizes on a logarithmic scale. We characterize the possible evolutionary outcomes with explicit criteria on the model parameters. An important ingredient for the proofs lies in comparisons of the stochastic process describing the population with logistic or branching birth and death processes with immigration. For the latter processes, we derive several results of independent interest.

Keywords: horizontal gene transfer, bacterial conjugation, stochastic individual-based models, long time behavior, large population approximation, coupling, branching processes with immigration, logistic competition.

MSC 2000 subject classification: 92D25, 92D15, 60J80, 60K35, 60F99.

1 Introduction and presentation of the model

Bacterial evolution understanding is fundamental in biology, medicine and industry. The ability of a bacterium to survive and reproduce depends on its genes, and evolution mainly results from

[*]Université de Lorraine, CNRS, Inria, IECL, F-54000 Nancy, France; E-mail: `nicolas.champagnat@inria.fr`

[†]Ecole Polytechnique, CNRS, UMR 7641 - CMAP, route de Saclay, 91128 Palaiseau Cedex-France; E-mail: `sylvie.meleard@polytechnique.edu`

[‡]Univ. Lille, CNRS, UMR 8524 - Laboratoire P. Painlevé, Département de Mathématiques, F-59000 Lille, France; E-mail: `chi.tran@univ-lille.fr`

the following basic mechanisms: heredity (also called vertical transmission); mutation which generates variability of the traits; selection which results from the interactions between individuals and their environment; exchange of genetic information between non-parental individuals during their lifetimes (also called horizontal gene transfer (HGT)), see for example [15], [13]. In many biological situations, these mechanisms drive the population to different evolutionary outcomes. The resulting effects may have a key role in the transmission of an epidemic, in the development of antibiotic resistances, in epigenetics or for the bacterial degradation of novel compounds such as human-created pesticides (e.g. [16], [12], [11]). In [1] and [2], the authors introduced an individual-based stochastic process for the trade-off between competition, transfer and advantageous mutations from which they derived some macroscopic approximations. They proved that the whole population is driven (by transfer) to evolutionary suicide, under the assumption that mutations are very rare. However, simulations show much richer evolutionary behaviors when mutation events are more frequent.

We propose below a toy model to capture these phenomena. The simulations of Figure 1.1 show that, depending on the transfer rate, we can observe either dominance of the trait with higher birth rate, or a cyclic phenomenon, or evolutionary suicide.

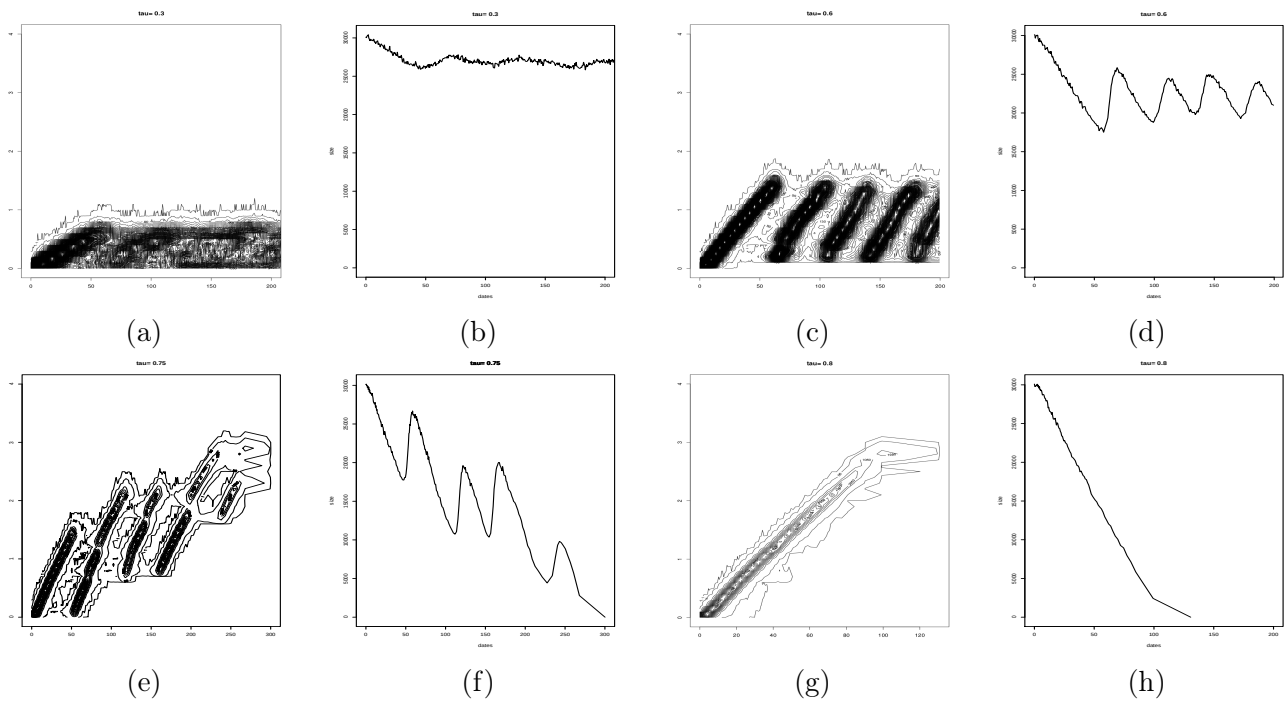


Figure 1.1: *Simulations of eco-evolutionary dynamics with unilateral frequency dependent trait transfer. The evolution of the trait distribution is pictured on the left columns, the evolution of the population sizes* $(N_0^K(t), \ldots, N_L^K(t))$ *is on the right columns. In all the simulations,* $K = 10000$ *and* $\delta = 0.1$ *and* $\alpha = 0.5$*. (a)-(b):* $\tau = 0.3$*. Smaller transfer rates are considered on longer time windows. (c)-(d):* $\tau = 0.6$*. We see cyclic re-emergences of the fittest traits. (e)-(f):* $\tau = 0.75$*. Re-emergence still occurs, but the higher transfer rate drives the trait distribution towards higher and less fit trait values. On stochastic simulations, this can lead to extinction or (temporary) cyclic behaviour. (g)-(h):* $\tau = 0.8$*. An evolutionary suicide takes place.*

We consider a stochastic discrete population of individuals characterized by some trait. The population evolves in continuous time through births (with or without mutation), deaths and transfer of traits between couples of individuals. More precisely, we study a continuous time Markov pure jump process, scaled by some integer parameter K, often called carrying capacity. The initial population size is of the order of K for large K. The trait space is the grid of mesh $\delta > 0$ of $[0,4]$: $\mathcal{X} = [0,4] \cap \delta\mathbb{N} = \{0, \delta, \dots, L\delta\}$ where $L = \lfloor 4/\delta \rfloor$. The population is described by the vector

$$\big(N_0^K(t), \dots, N_\ell^K(t), \dots, N_L^K(t)\big)$$

where $N_\ell^K(t)$ is the number of individuals of trait $x = \ell\delta$ at time t. We define the total population size N_t^K as

$$N_t^K = \sum_{\ell=0}^{L} N_\ell^K(t).$$

Let us now describe the dynamics of the population process.

- An individual with trait $x = \ell\delta$ in the population gives birth to another individual with rate $b(x) = 4 - x$. With probability

$$p_K = K^{-\alpha} \quad \text{with} \quad \alpha \in (0,1) \ , \tag{1.1}$$

  a mutation occurs and the new offspring carries the mutant trait $(\ell+1)\delta$.
  With probability $1 - p_K = 1 - K^{-\alpha}$, the new individual inherits the ancestral trait.

- An individual with trait x transfers its trait to a given individual of trait y in a population of total size N at rate

$$\tau(x, y, N) = \frac{\tau}{N} \mathbb{1}_{x>y},$$

  for some parameter $\tau > 0$.

- The individuals compete to survive (to share resources or territories). An individual with trait $x = \ell\delta$ in the population of total size N dies with natural death rate $d_K(x, N) = 1 + CN/K$.

The trait x may be interpreted in the biological setting as a phenotypic value quantifying either the pathogenic strength of bacteria or their antibiotic resistance [2]. It may be assumed to be related to the quantity of plasmids held by a bacterium. The birth rate favors small values of x with optimum at $x = 0$. It may represent a reproductive cost due to plasmids quantity. When a transfer happens, the quantity of plasmids in the recipient bacterium increases. The recipient bacterium receives the donor trait (this is called *conjugation* in the biological setting [1]). The transfer rate is assumed frequency-dependent, as observed by biologists for large populations [10]. The density-dependence in the death rate is uniform over the trait space. Because of the factor $1/K$, competition is governed only by traits with population size of order K. Therefore, density-dependence disappears when the total population size is negligible with respect to K.

Let us note that under Assumption (1.1), the total mutation rate in a population with size of order K, is equal to $K^{1-\alpha}$ and then goes to infinity with K. We are very far from the situation described in many papers as [4, 6, 2] where the authors explore the assumptions of the adaptive dynamics theory. In that cases, the total mutation rate Kp_K is assumed to satisfy $\log K \ll 1/(Kp_K) \ll$

e^{CK}, leading to a time scale separation between demographic and mutational events. Here, small populations of size order $K^\beta, \beta < 1$ can have a non negligible contribution to evolution by mutational events and we need to take into account all subpopulations with size of order K^β.

We need to consider in the sequel two different situations: either there is a single trait x with population size of order K, called *resident* trait, or the total population size is $o(K)$. In this last case, a trait with larger population size is called *dominant* trait.

When the trait x is the unique resident trait, it is well known [9] that, when K tends to infinity, the total population size can be approximated by $K\,n(t)$ where $n(.)$ solves the ODE

$$\dot{n}(t) = n(t)(3 - x - Cn(t)),$$

whose unique positive stable equilibrium is given by

$$\overline{n}(x) = \frac{(3-x) \vee 0}{C}. \tag{1.2}$$

The invasion fitness of a mutant individual of trait y in this resident population is then given by

$$S(y;x) = b(y) - d_K(y, K\overline{n}(x)) + \tau \mathbb{1}_{x<y} - \tau \mathbb{1}_{x>y} = x - y + \tau\,\mathrm{sign}(y-x), \tag{1.3}$$

where $\mathrm{sign}(x) = 1$ if $x > 0$; 0 if $x = 0$; -1 if $x < 0$. Indeed, the total transfer rate from x to y is given by $K\overline{n}(x)\,\frac{\tau}{(K\overline{n}(x)+1)}\mathbb{1}_{y>x} \sim \tau\,\mathbb{1}_{y>x}$ when $K \to +\infty$ and similarly from y to x. Note that $S(x;x) = 0$ and, for all traits x, y, $S(y;x) = -S(x;y)$ (see Figure 1.2). This implies in particular that there is no long-term coexistence of two resident traits. We also define the fitness of an individual of trait y in a negligible population (of size $o(K)$) with dominant trait x. In this case, the density-dependence vanishes, and since $N_x^K(t)/N_t^K \approx 1$, the fitness is now

$$\widehat{S}(y;x) = 3 - y + \tau \mathbb{1}_{x<y} - \tau \mathbb{1}_{x>y}. \tag{1.4}$$

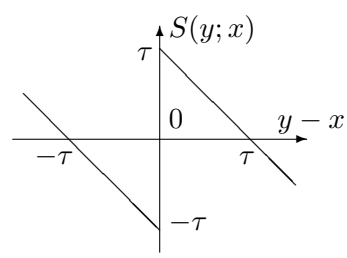


Figure 1.2: *Fitness function $S(y;x)$. In the absence of transfer ($\tau = 0$), evolution favors traits y smaller than x (their fitness is positive). The introduction of a positive transfer reverts this evolutive trend: $S(y;x) > 0$ if $x < y < x + \tau$. Note that $S(y;x) > 0$ also for $y < x - \tau$, which explains possible re-emergence of sufficiently smaller traits.*

Our study of the evolutionary dynamics of the model is based on a fine analysis of the size order, as power of K, of each subpopulation corresponding to the different trait compartments.

These powers of K evolve on the timescale $\log K$, as can be easily seen in the case of branching processes (see Lemma A.1). We thus define $\beta_\ell^K(t)$ for $0 \le \ell \le L$ such that

$$N_\ell^K(t \log K) = K^{\beta_\ell^K(t)} - 1, \quad \text{i.e.} \quad \beta_\ell^K(t) = \frac{\log(1 + N_\ell^K(t \log K))}{\log K}. \tag{1.5}$$

We assume that the trait $x = 0$ is initially resident, with density $3/C$. The initial condition is given by

$$N^K(0) = \big(\lfloor \frac{3K}{C} \rfloor, \lfloor K^{1-\alpha} \rfloor, \dots, \lfloor K^{1-\ell\alpha} \rfloor, \dots, \lfloor K^{1-\lfloor 1/\alpha \rfloor \alpha} \rfloor, 0, \dots, 0\big).$$

Due to mutations, this is close (in logarithmic scale) to the population state reached instantaneously on the time scale $\log K$ when the initial population is only composed of $\lfloor \frac{3K}{C} \rfloor$ individuals with trait 0. This can be understood from Lemma B.4 (with $\beta = 0$ and $c = 1 - \alpha$ for trait δ). With this initial condition, we have

$$\beta_\ell^K(0) \xrightarrow[K \to +\infty]{} (1 - \ell\alpha) \mathbb{1}_{0 \le \ell < \frac{1}{\alpha}}. \tag{1.6}$$

Our main result (Theorem 2.1) gives the asymptotic dynamics of $\beta^K(t) = (\beta_0^K(t), \dots, \beta_L^K(t))$ for $t \ge 0$ when $K \to +\infty$. We show that the limit is a piecewise affine continuous function, which can be described along successive phases determined by their resident or dominant traits. When the latter trait changes, the fitnesses governing the slopes are modified. Moreover, inside each phase, other changes of slopes are possible due to a delicate balance between mutations, transfer and growth of subpopulations. Our ambition is to cover all the possible cases: local extinctions, re-emergence of subpopulations, changes of slopes due to mutation and selection, dynamics when the total population size is $o(K)$, total extinction of the population... We deduce from the asymptotic dynamics of $\beta^K(t)$ explicit criteria for the occurrence of the different evolutionary outcomes observed in Figure 1.1 (Theorem 2.5). We provide a detailed study of the case of three traits in Section 3.

Such approach has already been used in [8, 3]. Durrett and Mayberry [8] consider constant population size or pure birth (Yule process) models, with directional mutations and increasing fitness parameter. They obtain travelling waves of selective sweeps. Bovier, Coquille and Smadi [3] consider a model with density-dependence but without transfer, with a single trait with positive fitness separated from the initial trait by unfit traits. They obtain bounds on the time needed to cross the fitness valley.

In our case, the dynamics is far more complex due to the trade-off between larger birth rates for small trait values and transfer to higher traits, leading to diverse evolutionary outcomes, including cyclic dynamics or evolutionary suicide. As a consequence, we need to consider cases where the dynamics of a given trait is completely driven by immigrations due to mutations from the resident trait, with time inhomogeneous immigration rates (see Theorem B.5). This complexifies a lot the analysis.

We first state in Section 2 our main results. First, we give our general result on the convergence of the exponents β_ℓ^K (Theorem 2.1) in Section 2.1. Then, we give general criteria on the parameters τ, δ and α for re-emergence of trait 0 and evolutionary suicide in Section 2.2 (Theorem 2.5). We then study in details the limit process in the case of three traits ($L = 2$) in Section 3. The proofs of Theorems 2.1 and 2.5 are given in Sections 4 and 5. Useful lemmas

on branching processes and branching processes with immigration are given respectively in Appendices A and B. Technical lemmas on birth and death processes with logistic competition and transfer are given in Appendix C. We conclude in Appendix D with the algorithmic construction of the limit of the exponents β_ℓ^K, used to perform simulations.

2 Main results

We state the principal results of the paper. Their proofs are given in Sections 4 and 5.

2.1 Asymptotic dynamics

The next result characterizes the asymptotic dynamics of $\beta^K(t) = (\beta_0^K(t), \dots, \beta_L^K(t))$ (when $K \to +\infty$) by a succession of deterministic time intervals $[s_{k-1}, s_k], k \geq 1$, called phases and delimited by changes of resident or dominant traits. The latter are unique except at times s_k and are denoted by $\ell_k^* \delta, k \geq 1$. This asymptotic result holds until a time T_0, which guarantees that there is neither ambiguity on these traits (Point (a) below) nor on the extinct subpopulations at the phase transitions (Point (c) below).

Theorem 2.1 *Assume that $\alpha \in (0,1)$, $\delta \in (0,4)$, $3/\delta \notin \mathbb{N}, \frac{\tau \pm 3}{\delta} \notin \mathbb{N}$ and (1.6) hold true.*

(i) *For all $T > 0$, the sequence $(\beta^K(t), t \in [0, T \wedge T_0])$ converges in probability in $\mathbb{D}([0, T \wedge T_0], [0,1]^L)$ to a deterministic piecewise affine continuous function $(\beta(t) = (\beta_1(t), \dots, \beta_L(t)), t \in [0, T \wedge T_0])$, such that $\beta_\ell(0) = (1 - \ell\alpha)\mathbb{1}_{0 \leq \ell < \frac{1}{\alpha}}$. The functions β and T_0 are parameterized by α, δ and τ defined as follows.*

(ii) *There exists an increasing nonnegative sequence $(s_k)_{k \geq 0}$ and a sequence $(\ell_k^*)_{k \geq 1}$ in $\{0, \dots, L\}$ defined inductively as follows: $s_0 = 0$, $\ell_1^* = 0$, and, for all $k \geq 1$, assuming that $s_{k-1} < T_0$ and ℓ_k^* have been constructed and that $\beta(s_{k-1}) \neq 0$, we can construct $s_k > s_{k-1}$ as follows*

$$s_k = \inf\{t > s_{k-1} : \exists \ell \neq \ell_k^*, \beta_\ell(t) = \beta_{\ell_k^*}(t)\}. \tag{2.1}$$

We can then decide whether we continue the induction after time s_k (i.e. $T_0 > s_k$) or not as follows:

(a) *if $\beta_{\ell_k^*}(s_k) > 0$, we set*

$$\ell_{k+1}^* = \arg\max_{\ell \neq \ell_k^*} \beta_\ell(s_k) \tag{2.2}$$

if the argmax is unique, or otherwise we set $T_0 = s_k$ and we stop the induction;

(b) *if $\beta_{\ell_k^*}(s_k) = 0$, we set $s_{k+1} = T_0 = +\infty$ and $\beta(t) = 0$ for all $t \geq s_k$;*

(c) *if in one of the previous cases, we have for some $\ell \neq \ell_k^*$, $\beta_\ell(s_k) = 0$ and $\beta_\ell(s_k - \varepsilon) > 0$ for all $\varepsilon > 0$ small enough, then we also set $T_0 = s_k$ and stop the induction; otherwise, the induction proceeds to the next step.*

(iii) *In (ii), the functions β_ℓ are defined, for all $t \in [s_{k-1}, s_k]$, by*

$$\beta_0(t) = \left[\mathbb{1}_{\beta_0(s_{k-1}) > 0} \left(\beta_0(s_{k-1}) + \int_{s_{k-1}}^t \widetilde{S}_{s,k}(0; \ell_k^* \delta)\, ds \right) \right] \vee 0 \tag{2.3}$$

and, for all $\ell \in \{1, \dots, L\}$,

$$\beta_\ell(t) = \left(\beta_\ell(s_{k-1}) + \int_{t_{\ell-1,k} \wedge t}^{t} \widetilde{S}_{s,k}(\ell\delta; \ell_k^*\delta)\, ds \right) \vee (\beta_{\ell-1}(t) - \alpha) \vee 0, \tag{2.4}$$

where, for all traits x, y,

$$\widetilde{S}_{t,k}(y; x) = \mathbb{1}_{\beta_{\ell_k^*}(t)=1}\, S(y; x) + \mathbb{1}_{\beta_{\ell_k^*}(t)<1}\, \widehat{S}(y; x) \tag{2.5}$$

and where

$$t_{\ell-1,k} = \begin{cases} \inf\{t \geq s_{k-1},\ \beta_{\ell-1}(t) = \alpha\}, & \text{if } \beta_\ell(s_{k-1}) = 0, \\ s_{k-1}, & \text{otherwise.} \end{cases} \tag{2.6}$$

In addition, for all ℓ *and all* $a < b < T_0$ *such that the time interval* $[a, b]$ *is included in the interior of the zero-set of* β_ℓ*, the event* $\{N_\ell^K(t \log K) = 0, \forall t \in [a, b]\}$ *has a probability converging to one as* K *tends to infinity.*

Remark 2.2 *1. It follows from the definition of* s_k *and* ℓ_{k+1}^* *that* $\max_\ell \beta_\ell(t) = \beta_{\ell_k^*}(t)$ *for all* $t \in [s_{k-1}, s_k)$.

2. In (2.5)*, when* $\beta_{\ell_k^*}(t) = 1$ *for some* $t \in (s_{k-1}, s_k)$*, there is a single resident trait* $\ell_k^*\delta$ *with population of the order of* K *and the function* S *defined in* (1.3) *is used. In the case where* $\beta_{\ell_k^*}(t) < 1$*, there is a single dominant trait and the total population size is of order* $o(K)$ *and the fitness function is* $\widehat{S}$ *defined in* (1.4)*. During each phase, the function* $\widetilde{S}_{t,k}$ *is actually constant, equal to* S *or* $\widehat{S}$ *as above, except when a dominant population becomes resident in the same phase. In the first case, for all* $t \in [s_{k-1}, s_k)$*, Eq.* (2.3) *and* (2.4) *take the simpler form*

$$\beta_0(t) = \begin{cases} \left[\mathbb{1}_{\beta_0(s_{k-1})>0}\ (\beta_0(s_{k-1}) + S(\ell\delta; \ell_k^*\delta)(t - s_{k-1}))\right] \vee 0 & \text{if } \beta_{\ell_k^*}(s_{k-1}) = 1, \\ \left[\mathbb{1}_{\beta_0(s_{k-1})>0}\ \left(\beta_0(s_{k-1}) + \widehat{S}(\ell\delta; \ell_k^*\delta)(t - s_{k-1})\right)\right] \vee 0 & \text{if } \beta_{\ell_k^*}(s_{k-1}) < 1 \end{cases}$$

and, for all $\ell \in \{1, \dots, L\}$,

$$\beta_\ell(t) = \begin{cases} (\beta_\ell(s_{k-1}) + S(\ell\delta; \ell_k^*\delta)(t - t_{\ell-1,k})_+) \vee (\beta_{\ell-1}(t) - \alpha) \vee 0 & \text{if } \beta_{\ell_k^*}(s_{k-1}) = 1, \\ \left(\beta_\ell(s_{k-1}) + \widehat{S}(\ell\delta; \ell_k^*\delta)(t - t_{\ell-1,k})_+\right) \vee (\beta_{\ell-1}(t) - \alpha) \vee 0 & \text{if } \beta_{\ell_k^*}(s_{k-1}) < 1. \end{cases} \tag{2.7}$$

Otherwise, $\widetilde{S}_{t,k}$ *switches from* $\widehat{S}$ *to* S *at the first time where* $\max \beta_\ell(t) = \beta_{\ell_k^*}(t) = 1$*. Therefore, since* $S(\ell_k^*\delta, \ell_k^*\delta) = 0$*, we obtain in all cases*

$$\beta_{\ell_k^*}(t) = \begin{cases} 1 & \text{if } \beta_{\ell_k^*}(s_{k-1}) = 1, \\ \left[\left(\beta_{\ell_k^*}(s_{k-1}) + \widehat{S}(\ell_k^*\delta; \ell_k^*\delta)\,(t - s_{k-1})\right) \wedge 1\right] \vee 0 & \text{if } \beta_{\ell_k^*}(s_{k-1}) < 1. \end{cases}$$

3. It follows from the last formula that $\max_\ell \beta_\ell(t) \leq 1$ *for all* $t \in [0, T_0]$.

4. When $\beta_\ell(s_{k-1}) = 0$*, the time* $t_{\ell-1,k}$ *corresponds to the first time where the incoming mutation rate in subpopulation* $\ell\delta$ *becomes significant.*

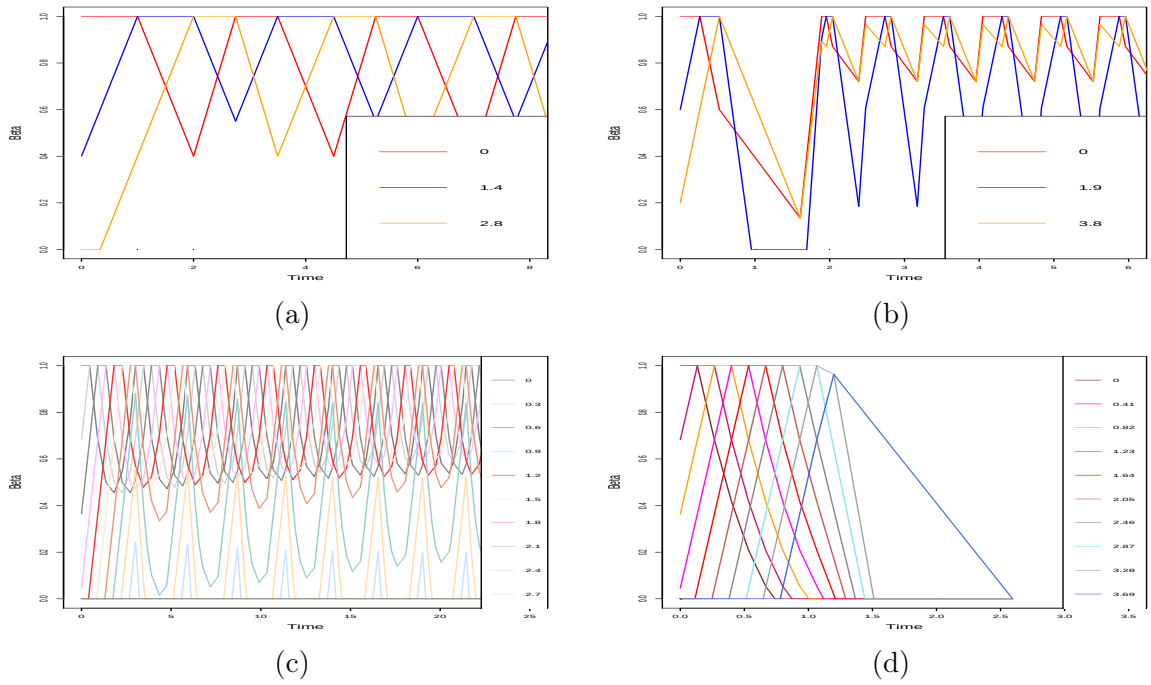


Figure 2.1: *Exponents $\beta_\ell(t)$ as functions of time. (a): $\delta = 1.4$, $\alpha = 0.6$, $\tau = 2$. We see a periodic behavior showing re-emergences of the fittest traits. (b): $\delta = 1.9$, $\alpha = 0.4$, $\tau = 3.43$. When the trait 2δ becomes dominant, the population size is of order $o(K)$. We see a re-emergence of trait 0 after a phase of apparent macroscopic extinction (i.e. a total population size $o(K)$). Although the trait δ goes extinct while 2δ is dominant, it is recreated by mutations from trait 0. (c): $\delta = 0.3$, $\alpha = 1/\pi$, $\tau = 1$. A cyclic but non-periodic behaviour is observed. (d): $\delta = 0.41$, $\alpha = 1/\pi$, $\tau = 2.8$. The population is directly driven to evolutionary suicide.*

We cannot ensure that $T_0 = +\infty$ for almost all parameters α, δ and τ. However we have not encountered any case where $T_0 < +\infty$ in the simulations. In the sequel, we exhibit large sets of parameters where $T_0 = +\infty$ in the case of three traits (Section 3). We also prove in Theorem 2.5 that, for any sets of parameters, T_0 is larger than the time of extinction or of first re-emergence.

Note that the previous result keeps track of populations of size K^β for $0 < \beta \leq 1$, but not of populations of smaller order, which go fast to extinction on the time scale $\log K$.

The next theorem gives a characterization of β as solution of a dynamical system.

Corollary 2.3 *Under the assumptions of Theorem 2.1, we set*

$$\ell^*(t) = \sum_{k\geq 1} \ell_k^* \mathbb{1}_{[s_{k-1},s_k[}(t) \text{ and } \widetilde{S}_t(y;x) = \mathbb{1}_{\beta_{\ell^*(t)}(t)=1}\, S(y;x) + \mathbb{1}_{\beta_{\ell^*(t)}(t)<1}\, \widehat{S}(y;x).$$

The function $\beta(t)$ is right-differentiable on $[0, T_0)$ and satisfies

$$\dot{\beta}_\ell(t) = \Sigma_\ell(t) \mathbb{1}_{\beta_\ell(t)>0 \text{ or } (\beta_\ell(t)=0 \text{ and } \beta_{\ell-1}(t)=\alpha)} \tag{2.8}$$

where Σ_ℓ is defined recursively by $\Sigma_0(t) = \widetilde{S}_t(0, \delta\ell^*(t))$ *and* $\forall \ell \geq 1$

$$\Sigma_\ell(t) \quad = \quad \begin{cases} \widetilde{S}_t(\ell\delta; \ell^*(t)\delta) \vee \Sigma_{\ell-1}(t) & \text{if } \beta_\ell(t) = \beta_{\ell-1}(t) - \alpha \\ \widetilde{S}_t(\ell\delta; \ell^*(t)\delta) & \text{if } \beta_\ell(t) > \beta_{\ell-1}(t) - \alpha. \end{cases} \tag{2.9}$$

Remark 2.4 *One may wonder if the ODE* (2.8) *characterizes the function β. For this we first need to characterize $\ell^*(t)$ as an explicit function of $\beta(t)$. One would like to define it as $\ell^*(t) = \arg\max_{0\leq\ell\leq L} \beta_\ell(t)$ and take it right-continuous. This is correct if there is a single argmax. Otherwise, there are by definition of T_0 only two choices ℓ and ℓ' and there is a single admissible choice in the sense that the corresponding affine solution to* (2.8) *on $[t, t+\varepsilon]$ satisfies $\ell^*(s) = \arg\max_{0\leq\ell\leq L} \beta_\ell(s)$ locally for $s \in (t, t+\varepsilon)$ for $\varepsilon > 0$ small enough. Indeed if $\max_{0\leq\ell\leq L} \beta_\ell(t) = 1$ and since $S(\ell'\delta, \ell\delta) = -S(\ell\delta, \ell'\delta)$, one of the two fitnesses is positive, for example $S(\ell\delta, \ell'\delta)$. If one takes the wrong choice $\ell^*(t) = \ell'$, then $\Sigma_\ell(t) = S(\ell\delta, \ell'\delta) > 0$, hence the solution of* (2.8) *gives $\beta_\ell(s) > 1$ for $s > t$ locally. If $\max_{0\leq\ell\leq L} \beta_\ell(t) < 1$, a similar argument with $\widehat{S}$ consists in choosing the trait with higher invasion fitness.*

Therefore, (2.8) *can be expressed as an autonomous ODE and there is a unique admissible solution. Generalizations of our result to models with different birth, death and transfer rates, can be obtained by changing accordingly the fitness function in this ODE.*

Simulations are shown in Figure 2.1 for various parameter values. The times s_k correspond to changes of resident or dominant populations. However, we observe several changes of slopes between these times. The computation of these successive times called t_k is given in Theorem D.1 in Appendix D.

2.2 Re-emergence of trait 0

In Figure 2.1, we have exhibited different evolutionary dynamics (re-emergence of a trait, cyclic behavior, local extinction, evolutionary suicide). By re-emergence of a trait $\ell\delta$, we mean that $\beta_\ell(s) = 1$ on some non-empty time interval $[t_1, t_2]$, then $\beta_\ell(s) < 1$ on some non-empty interval (t_2, t_3) and then $\beta_\ell(s) = 1$ again on some non-empty interval $[t_3, t_4]$. We would like to predict the evolutionary outcome in function of the parameters α, δ, τ. As detailed for three traits ($L = 2$) in the next section, there are so many situations that we are not able to fully characterize the outcomes. Therefore, we focus on the beginning of the dynamics until either global extinction or re-emergence of one trait. The resurgence of trait 0 is a prerequisite for a cyclic dynamics as those observed in Figures 1.1 (c).

We assume that $\delta < 4/3$ (so that $L \geq 3$) and only consider the case $\delta < \tau < 3$. Let

$$\widetilde{k} := \lceil \frac{\tau}{\delta} \rceil \quad \text{and} \quad \bar{k} = \lfloor 2\frac{\tau}{\delta} \rfloor.$$

We will see in the proof of the next result that, for the first phases,

$$s_k := \frac{k\alpha}{\tau - \delta},$$

the trait $k\delta$ is resident on $[s_k, s_{k+1})$ ($\beta_k(s) = 1$) and for all $s \in [s_k, s_{k+1})$,

$$\beta_\ell(s) = \begin{cases} [1 - (\ell - k)\alpha + (\tau - \delta)(s - s_k)] \vee 0 & \text{if } k < \ell \leq L, \\ 1 - \frac{\alpha(k-\ell-1)}{\tau-\delta}\left(\tau - \frac{k-\ell}{2}\delta\right) - (\tau - (k-\ell)\delta)(s - s_k) & \text{if } 0 \leq \ell < k. \end{cases}$$

These formulas stay valid until either $\beta_0(s) = 0$ (loss of 0), or $\beta_0(s) = 1$ for some $s > s_1$ (re-emergence of 0), or $\ell_k^* \delta > 3$, where ℓ_k^* has been defined in (2.2) (the population size becomes $o(K)$). The function $\beta_0(s)$ in the previous equation is piecewise affine and its slope becomes positive at time s_{k^*}. Hence its minimal value is equal to

$$m_0 = \beta_0(s_{\widetilde{k}}) = 1 - \frac{\alpha(\widetilde{k}-1)}{\tau - \delta}\left(\tau - \frac{\widetilde{k}}{2}\delta\right).$$

Provided the latter is positive, β_0 reaches 1 again in phase $[s_{\bar{k}}, s_{\bar{k}+1})$ at time

$$\bar{\tau} := s_{\bar{k}} + \frac{\alpha(\bar{k}-1)}{\tau - \delta}\,\frac{\tau - \frac{\bar{k}}{2}\delta}{\bar{k}\delta - \tau} = s_{\lfloor 2\frac{\tau}{\delta}\rfloor} + \frac{\alpha(\lfloor 2\frac{\tau}{\delta}\rfloor - 1)}{\tau - \delta}\,\frac{\tau - \frac{\lfloor 2\frac{\tau}{\delta}\rfloor}{2}\delta}{\lfloor 2\frac{\tau}{\delta}\rfloor\delta - \tau}. \tag{2.10}$$

Theorem 2.5 *Assume $\delta < \tau < 3$, $\delta < 4/3$ and under the assumptions of Theorem 2.1,*

(a) *If $m_0 > 0$ and $\bar{k}\delta < 3$, then the first re-emerging trait is 0 and the maximal exponent is always 1 until this re-emergence time.*

(b) *If $m_0 < 0$, the trait 0 gets lost before its re-emergence and there is global extinction of the population before the re-emergence of any trait*

(c) *If $m_0 > 0$ and $\bar{k}\delta > 3$, there is re-emergence of some trait $\ell\delta < 3$ and, for some time t before the time of first re-emergence, $\max_{1\le\ell\le L} \beta_\ell(t) < 1$.*

Biologically, Case (b) corresponds to evolutionary suicide. In Cases (a) and (c), very few individuals with small traits remain, which are able to "reborn" the first re-emerging population after the resident or dominant trait becomes too large. In these cases, one can expect successive re-emergences. However, we don't know if there exists a limit cycle for the dynamics. Case (c) means that the total population is $o(K)$ on some time interval, before re-emergence occurs after populations with too large traits become small enough.

Heuristically, using the approximation that $\widetilde{k} \approx \tau/\delta$, we obtain that $m_0 \approx 1 - \frac{\alpha\tau}{2\delta}$. Hence, we have $m_0 > 0$ (re-emergence) provided $\tau \lesssim 2\delta/\alpha$ and extinction otherwise. Transfer rates higher than $2\delta/\alpha$ favor extinction because the population is pushed to higher trait values. Small values of δ or high values of α give more time for extinction of the small subpopulations. Note that, for $m_0 > 0$, the condition $\bar{k}\delta < 3$ is roughly $\tau < 3/2$. Hence, for transfer rates smaller than 3/2, 0 re-emerges first, while other traits can re-emerge before 0 otherwise.

3 Case of three traits

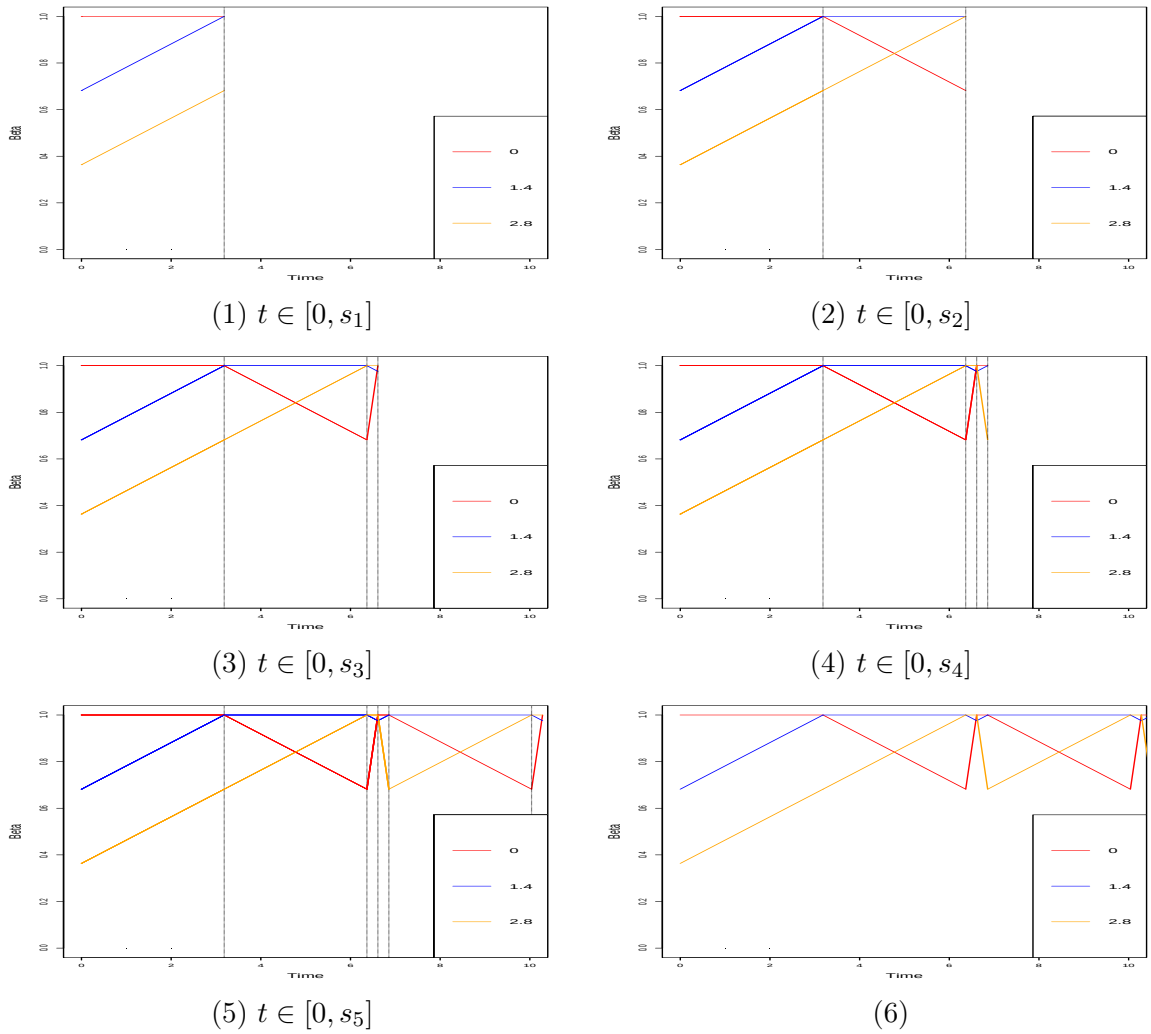
Before proving our main results, let us illustrate the limit exponents $\beta(t)$ in the case of three traits. Let us consider $\delta > 0$ such that $2\delta < 4 < 3\delta$, so that the possible traits are 0, δ and 2δ. A simulation is shown step by step in Figure 3.1, which we will now explain.

The initial condition is

$$\beta(0) = (\beta_0(0),\, \beta_1(0),\, \beta_2(0)) = (1, 1-\alpha, (1-2\alpha)\vee 0). \tag{3.1}$$

The fitnesses given in (1.3) are

$$S(0;0) = 0, \qquad S(\delta;0) = \tau - \delta, \qquad S(2\delta;0) = \tau - 2\delta.$$



(1) $t \in [0, s_1]$ (2) $t \in [0, s_2]$

(3) $t \in [0, s_3]$ (4) $t \in [0, s_4]$

(5) $t \in [0, s_5]$ (6)

Figure 3.1: *Construction step by step of the exponents* $\beta_0(t)$, $\beta_1(t)$ *and* $\beta_2(t)$, *as function of time in the case where* $\delta < \tau < 2\delta < 3 < 4 < 3\delta$. *Here,* $\delta = 1.4$, $\alpha = \frac{1}{\pi}$, $\tau = 1.5$. *The traits are* 0, 1.4 *and* 2.8.

3.1 Case 1: $\tau < \delta$

In this case, neither the traits δ nor the trait 2δ are advantageous and these populations survive only thanks to the mutations from the trait 0 to δ and from the trait δ to 2δ. The exponents remain constant and $\forall t \geq 0,\ \beta(t) = \beta(0)$.

3.2 Case 2: $\delta < \tau < 2\delta$

Following Theorem 2.1, we shall decompose the dynamics of $\beta(t)$ into successive phases corresponding to the time intervals $[s_{k-1}, s_k]$.

Phase 1: time interval $[0, s_1]$**.** In this Case 2, $S(0;0) = 0$, $S(\delta;0) > 0$ and $S(2\delta;0) < 0$. While the resident population remains the population with trait 0, the population with trait δ

has positive fitness and its growth is described for $t \in [0, s_1]$ by the exponent:

$$\beta_1(t) = \big[(1-\alpha) + (\tau-\delta)t\big] \vee (1-\alpha) \vee 0 = (1-\alpha) + (\tau-\delta)t.$$

The bracket corresponds to the intrinsic growth associated to the fitness $S(\delta; 0)$, the term $1-\alpha$ is the contribution of mutations from the population of trait 0 and is here smaller than the term with the bracket.
The population with trait 2δ has negative fitness and

$$\beta_2(t) = \big[(1-2\alpha) - (2\delta-\tau)t\big] \vee \big[(1-2\alpha) + (\tau-\delta)t\big] \vee 0 = [(1-2\alpha) + (\tau-\delta)t] \vee 0.$$

As for the trait δ, the first bracket corresponds to the intrinsic growth with a negative slope $\tau - 2\delta < 0$, while the second bracket corresponds to the contribution of mutations from the population with trait δ.
It is clear that $\beta_2(t) < \beta_1(t) \leq 1$. Hence the first phase stops when $\beta_1(t) = 1$, for

$$s_1 = \frac{\alpha}{\tau-\delta}.$$

The first phase is illustrated in Fig. 3.1(1).

Phase 2: time interval $[s_1, s_2]$. At time s_1, the populations with traits 0 and δ both have sizes of order K: more precisely, the exponents are

$$\beta_0(s_1) = 1, \qquad \beta_1(s_1) = 1, \qquad \beta_2(s_1) = 1-\alpha.$$

Because $S(0;\delta) < 0$ and $S(\delta;0) > 0$, the new resident population with trait δ replaces the population with trait 0 whose exponent decreases after time s_1. The size of the population with trait δ remains close to $(3-\delta)K/C$, i.e. $\beta_1(t) = 1$, during the whole Phase 2, and using (1.3):

$$S(0;\delta) = \delta - \tau < 0, \qquad S(\delta;\delta) = 0, \qquad S(2\delta;\delta) = \tau - \delta > 0.$$

Thus, the decrease of the population with trait 0 is described by $\beta_0(s_1+t) = [1-(\tau-\delta)\,t] \vee 0$ (recall that no mutant can have trait 0). The population with trait 2δ has a positive fitness (first bracket in the following equation) and benefits from mutations coming from the trait δ (second bracket):

$$\beta_2(s_1+t) = \big[(1-\alpha) + (\tau-\delta)\,t\big] \vee \big[1-\alpha\big] \vee 0 = (1-\alpha) + (\tau-\delta)\,t.$$

This second phase stops when $\beta_2(t) = 1$, at time

$$s_2 = s_1 + \frac{\alpha}{\tau-\delta} = \frac{2\alpha}{\tau-\delta}.$$

We check that $\beta_0(s_1+t) = 1-(\tau-\delta)\,t, \quad \forall t \in [s_2 - s_1]$. This phase is illustrated in Fig. 3.1(2).

3.2.1 Case 2(a): $2\delta < 3$

In this case, trait 2δ can survive on its own, i.e. its equilibrium population size $\frac{3-2\delta}{C}$ is positive.

Phase 3: time interval $[s_2, s_3]$. Because $S(\delta; 2\delta) < 0$ and $S(2\delta;\delta) > 0$, we have at time s_2 a replacement of the resident population with trait δ by the population with trait 2δ which becomes the new resident population, i.e. $\beta_2(t) = 1$. At time s_2, the exponents are:

$$\beta_0(s_2) = 1-\alpha, \qquad \beta_1(s_2) = 1, \qquad \beta_2(s_2) = 1. \tag{3.2}$$

The population size of trait 2δ is close to $(3-2\delta)K/C$ so that the fitnesses are:

$$S(0;2\delta) = 2\delta - \tau > 0, \qquad S(\delta;2\delta) = \delta - \tau < 0, \qquad S(2\delta;2\delta) = 0.$$

The population with trait 0 increases with the exponent $\beta_0(s_2+t) = (1-\alpha) + (2\delta-\tau)\,t$.

The trait δ has negative fitness but benefits from mutations coming from the trait 0:

$$\beta_1(s_2+t) = \big[1-(\tau-\delta)\,t\big] \vee \big[(1-2\alpha)+(2\delta-\tau)\,t\big] \vee 0. \tag{3.3}$$

This third phase is illustrated in Fig. 3.1(3). This phase stops when $\beta_0(t)=1$, i.e. at time

$$s_3 = s_2 + \frac{\alpha}{2\delta-\tau} = \frac{2\alpha}{\tau-\delta} + \frac{\alpha}{2\delta-\tau}.$$

We also have $\beta_1(s_3) = \left(1-\alpha\,\frac{\tau-\delta}{2\delta-\tau}\right) \vee (1-\alpha)$.

We have to distinguish two cases for Phase 4, depending on the value of $\beta_1(s_3)$.

Phase 4, case 2(a)(i): time interval $[s_3, s_4]$ under the assumption $\tau-\delta < 2\delta-\tau$. Then,

$$\beta_0(s_3) = 1, \quad \beta_1(s_3) = 1-\alpha\,\frac{\tau-\delta}{2\delta-\tau}, \quad \beta_2(s_3) = 1.$$

The new resident population is the one with trait 0 and the fitnesses are the same as in Phase 1, but the initial conditions are different. We obtain as above the exponents $\beta_0(s_3+t)=1$,

$$\beta_1(s_3+t) = \big[1-\frac{\tau-\delta}{2\delta-\tau}\alpha + (\tau-\delta)\,t\big] \vee \big[1-\alpha\big] \vee 0 = 1-\frac{\tau-\delta}{2\delta-\tau}\alpha + (\tau-\delta)\,t.$$

and

$$\beta_2(s_3+t) = \big[1-(2\delta-\tau)\,t\big] \vee \big[1-\frac{\delta\alpha}{2\delta-\tau} + (\tau-\delta)\,t\big] \vee 0,$$

as illustrated in Fig. 3.1 (4). The phase stops when $\beta_1(t)=1$ at time

$$s_4 = s_3 + \frac{\alpha}{2\delta-\tau} = \frac{2\alpha}{\tau-\delta} + \frac{2\alpha}{2\delta-\tau}.$$

We check that $\beta_2(s_3+t) = 1-(2\delta-\tau)\,t, \quad \forall t \le s_4 - s_3$ and hence

$$\beta_0(s_4) = 1, \qquad \beta_1(s_4) = 1, \qquad \beta_2(s_4) = 1-\alpha.$$

We recognize the initial condition of Phase 2. Therefore, the system behaves periodically (as in Figure 2.1(a)) starting from time s_1, with period

$$s_4 - s_1 = \frac{\alpha}{\tau-\delta} + \frac{2\alpha}{2\delta-\tau}.$$

Phase 4, case 2(a)(ii): time interval $[s_3, s_4]$ under the assumption $\tau-\delta > 2\delta-\tau$. In this case,

$$\beta_0(s_3) = 1, \quad \beta_1(s_3) = 1-\alpha, \quad \beta_2(s_3) = 1.$$

In this case, we obtain $\beta_0(s_3+t)=1,\ \beta_1(s_3+t)=1-\alpha+(\tau-\delta)\,t,$

$$\beta_2(s_3+t)=1-(2\delta-\tau)\,t \quad \text{and} \quad s_4=s_3+\frac{\alpha}{\tau-\delta}=\frac{3\alpha}{\tau-\delta}+\frac{\alpha}{2\delta-\tau}.$$

Phase 5, case 2(a)(ii): time interval $[s_4, s_5]$. We have

$$\beta_0(s_4)=1, \qquad \beta_1(s_4)=1, \qquad \beta_2(s_4)=1-\alpha\,\frac{2\delta-\tau}{\tau-\delta}.$$

Proceeding as above, we obtain

$$s_5=s_4+\alpha\,\frac{2\delta-\tau}{(\tau-\delta)^2}=\frac{3\alpha}{\tau-\delta}+\frac{\alpha}{2\delta-\tau}+\alpha\,\frac{2\delta-\tau}{(\tau-\delta)^2}$$

and for all $t\in[0, s_5-s_4]$, $\beta_1(s_4+t)=1$,

$$\beta_0(s_4+t)=1-(\tau-\delta)t \quad \text{and} \quad \beta_2(s_4+t)=1-\alpha\,\frac{2\delta-\tau}{\tau-\delta}+(\tau-\delta)\,t.$$

Phase 6, case 2(a)(ii): time interval $[s_5, s_6]$. We have

$$\beta_0(s_5)=1-\alpha\,\frac{2\delta-\tau}{\tau-\delta}, \qquad \beta_1(s_5)=1, \qquad \beta_2(s_5)=1.$$

We obtain

$$s_6=s_5+\frac{\alpha}{\tau-\delta}=\frac{4\alpha}{\tau-\delta}+\frac{\alpha}{2\delta-\tau}+\alpha\,\frac{2\delta-\tau}{(\tau-\delta)^2}$$

and for all $t\in[0, s_6-s_5]$, $\beta_2(s_5+t)=1$,

$$\beta_0(s_5+t)=1-\alpha\,\frac{2\delta-\tau}{\tau-\delta}+(2\delta-\tau)\,t \quad \text{and} \quad \beta_1(s_5+t)=1-(\tau-\delta)\,t.$$

Hence $\beta_0(s_6)=1,\ \beta_1(s_6)=1-\alpha,\ \beta_2(s_6)=1$. We recognize the initial condition as in Phase 4, case 2(a)(ii). Therefore, the system behaves periodically starting from time s_3, with period

$$s_6-s_3=\frac{2\alpha}{\tau-\delta}+\alpha\,\frac{2\delta-\tau}{(\tau-\delta)^2}.$$

3.2.2 Case 2(b): $2\delta>3$

Phase 3: time interval $[s_2, s_3]$. In this case, the trait 2δ, which replaces the former resident trait δ at time s_2, cannot survive alone and becomes dominant. So $\beta_2(s_2+t)$ does not remain equal to 1, but decreases with slope $\widehat{S}(2\delta;2\delta)$. Recall that, at time s_2, the exponents are given by (3.2). The fitnesses now become:

$$\widehat{S}(0;2\delta)=3-\tau, \qquad \widehat{S}(\delta;2\delta)=3-\delta-\tau<0, \qquad \widehat{S}(2\delta;2\delta)=3-2\delta<0.$$

Note that we do not distinguish yet on the sign of $\widehat{S}(0;2\delta)$, which may be either positive or negative in this case. We obtain

$$\beta_0(s_2+t)=[1-\alpha+(3-\tau)\,t]\vee 0, \quad \beta_1(s_2+t)=[1-(\tau+\delta-3)\,t]\vee[1-2\alpha+(3-\tau)\,t]\vee 0$$

and $\beta_2(s_2+t) = 1-(2\delta-3)\,t$ until either $\beta_0(t)=\beta_2(t)>0$, which corresponds to a change of dominant population with exponent smaller than 1, or $\beta_2(t)=0$, which corresponds to the extinction of the whole population. Note that we cannot have $\beta_1(t)=\beta_2(t)$ before $\beta_0(t)=\beta_2(t)$ since $\tau+\delta-3>2\delta-3$. One can easily check that $\beta_2(t)$ hits 0 before crossing the curve $\beta_0(t)$ if and only if $\frac{2\delta-\tau}{2\delta-3}<\alpha$. Of course, this cannot occur if $\tau<3$, since in this case $\widehat{S}(0;2\delta)>0$.

Phase 3, case 2(b)(i): $\frac{2\delta-\tau}{2\delta-3}<\alpha$. In this case, the whole population gets extinct at time

$$s_3 = s_2 + \frac{1}{2\delta-3} = \frac{2\alpha}{\tau-\delta} + \frac{1}{2\delta-3}.$$

Phase 3, case 2(b)(ii): $\frac{2\delta-\tau}{2\delta-3}>\alpha$. Trait 0 becomes dominant and replaces trait 2δ at time

$$s_3 = s_2 + \frac{\alpha}{2\delta-\tau} = \frac{2\alpha}{\tau-\delta} + \frac{\alpha}{2\delta-\tau}.$$

We obtain the exponents $\beta_0(s_3)=\beta_2(s_3)=1-\alpha\,\frac{2\delta-3}{2\delta-\tau}\in(0,1)$ and

$$\beta_1(s_3) = \left[1-\alpha\,\frac{\delta+\tau-3}{2\delta-\tau}\right] \vee \left[1-\alpha\,\frac{4\delta-\tau-3}{2\delta-\tau}\right] \vee 0.$$

Phase 4, case 2(b)(ii). We obtain the new fitnesses

$$\widehat{S}(0;0)=3, \qquad \widehat{S}(\delta;0)=3+\tau-\delta>3, \qquad S(2\delta;0)=3-2\delta+\tau\in(0,3).$$

To compute $\beta_1(s_3+t)$, one needs to distinguish whether $\beta_1(s_3)>0$ or $\beta_1(s_3)=0$. In the last case, one needs to wait until $\beta_0(s_3+t)=\alpha$ before β_1 starts to increase with slope $3+\tau-\delta$. The phase stops either when $\beta_0(s_3+t)=1$ (re-emergence of trait 0, which becomes resident once again) or $\beta_0(t)=\beta_1(t)<1$ (change of dominant trait). A delicate case study shows that the first case occurs if $\tau-\delta<3/2$ when $\beta_1(s_3)>0$, or if $\tau-\delta<3\alpha/(1-\alpha)$ when $\beta_1(s_3)=0$, and the second case if $\tau-\delta>3/2$ when $\beta_1(s_3)>0$, or if $\tau-\delta>3\alpha/(1-\alpha)$ when $\beta_1(s_3)=0$. In the first case, one needs to proceed with similar computations as in the first phases. In the second case, either trait δ re-emerges first (i.e. becomes resident again), or trait 2δ becomes dominant once again. Explicit computations of the subsequent dynamics are very lengthy. This case is illustrated by Figure 2.1(b).

3.3 Case 3: $2\delta<\tau$

We proceed similarly as in Case 2.

Phase 1: time interval $[0,s_1]$. This phase is the same as in Case 2. The resident trait is 0 and the fitnesses of the three traits are given by

$$S(0;0)=0, \quad S(\delta;0)=\tau-\delta>0, \quad S(2\delta;0)=\tau-2\delta>0.$$

We obtain, for all $t\in[0,s_1]$, $\beta_0(t)=1$,

$$\beta_1(t)=1-\alpha+(\tau-\delta)\,t \quad \text{and} \quad \beta_2(t)=1-2\alpha+(\tau-\delta)\,t$$

where $s_1 = \frac{\alpha}{\tau-\delta}$. Thus $\beta_0(s_1) = \beta_1(s_1) = 1$ and $\beta_2(s_1) = 1-\alpha$.

Phase 2: time interval $[s_1, s_2]$. This phase is also the same as in Case 2. The resident trait is δ and the fitnesses are given by

$$S(0;\delta) = -(\tau-\delta) < 0, \quad S(\delta;\delta) = 0, \quad S(2\delta;\delta) = \tau-\delta > 0.$$

We obtain, for all $t \in [0, s_2 - s_1]$, $\beta_1(s_1+t) = 1$,

$$\beta_0(s_1+t) = 1-(\tau-\delta)\,t \quad \text{and} \quad \beta_2(s_1+t) = 1-\alpha+(\tau-\delta)\,t,$$

where $s_2 = s_1 + \frac{\alpha}{\tau-\delta}$. Thus $\beta_1(s_2) = \beta_2(s_2) = 1$ and $\beta_0(s_2) = 1-\alpha$.

Again, we have to separate the cases $2\delta < 3$ and $2\delta > 3$.

3.3.1 Case 3(a): $2\delta < 3$

Phase 3: time interval $[s_2, +\infty)$. The resident trait is 2δ and the fitnesses are given by

$$S(0;2\delta) = -(\tau-2\delta) < 0, \quad S(\delta;2\delta) = -(\tau-\delta) < 0, \quad S(2\delta;2\delta) = 0.$$

We obtain, for all $t \geq 0$, $\beta_2(s_2+t) = 1$, $\beta_0(s_2+t) = [1-\alpha-(\tau-2\delta)\,t] \vee 0$ and

$$\beta_1(s_2+t) = [1-(\tau-\delta)\,t] \vee [1-2\alpha-(\tau-2\delta)\,t] \vee 0.$$

Therefore, 2δ remains the resident trait forever.

3.3.2 Case 3(b): $2\delta > 3$

Time interval $[s_2, +\infty)$. The resident trait 2δ cannot survive by itself. Hence, the fitnesses are now given by

$$\widehat{S}(0;2\delta) = 3-\tau < 0, \quad \widehat{S}(\delta;2\delta) = 3-\delta-\tau < 0, \quad \widehat{S}(2\delta;2\delta) = 3-2\delta < 0.$$

Since $\widehat{S}(0;2\delta) < \widehat{S}(2\delta;2\delta)$ and $\widehat{S}(\delta;2\delta) < \widehat{S}(2\delta;2\delta)$, Phase 3 will end when $\beta_2(t)$ hits 0, i.e. when the population gets extinct. Hence, for all $t \geq 0$, $\beta_0(s_2+t) = [1-\alpha+(3-\tau)\,t] \vee 0$, $\beta_1(s_2+t) = [1-(\tau+\delta-3)\,t] \vee [1-2\alpha+(3-\tau)\,t] \vee 0$ and $\beta_2(s_2+t) = [1-(2\delta-3)\,t] \vee 0$, and the extinction time is $s_3 = s_2 + \frac{1}{2\delta-3}$.

4 Proof of Theorem 2.1

4.1 Main ideas of the proof

We start from the stochastic birth and death process with mutation, competition and transfer, $(N_0^K(t), \ldots, N_L^K(t))$. Our goal is to study the limit behaviour of the vector $(\beta_0^K(t), \ldots, \beta_L^K(t))$ defined in (1.5).

Theorem 2.1 will be obtained by a fine comparison of the size of each subpopulation defined by a given trait value with carefully chosen branching processes with immigration. The stochastic dynamics consists in a succession of phases $[\sigma_k^K \log K, \theta_k^K \log K]$ followed by intermediate steps $[\theta_k^K \log K, \sigma_{k+1}^K \log K]$. We will prove that θ_k^K converges in probability to s_k, $k \geq 1$. In each phase

there is a single resident or dominant subpopulation. When another trait reaches a comparable size, the intermediate step starts and ends after the replacement of the resident or dominant trait. In the limit, intermediate steps vanish on the time scale $\log K$.

To control the exponents $\beta_\ell^K(t)$, we proceed by a double induction, first on the phases, and inside each phase, on the traits $\ell\delta$, for $\ell = 0$ to $\ell = L$. The exponents are approximately piecewise affine. Changes of slopes may happen when a new trait emerges, when a trait dies or when the dynamics of a trait becomes driven by incoming mutations. We use asymptotic results on branching processes with immigration detailed in Appendix B to control the sizes of the non-dominant subpopulations. The main result used inside phases is Theorem B.5.
During intermediate steps, two subpopulations are of maximal order. We use comparisons with dynamical systems, see Lemmas C.2 and C.3.

Let us now describe in more details how the stopping times θ_k^K are constructed. At each phase, this time may have two different definitions depending on whether the future emerging trait is dominant or resident. For this, it is convenient to consider θ_*^K and $\widetilde{\theta}_*^K$ where the index k is omitted. Consider a phase, starting at time $\sigma_*^K \log K$ with the largest subpopulation of trait $\ell^*\delta$. Two cases may occur: either $\ell^*\delta$ is a resident trait with population size close to its equilibrium $(3-\ell^*\delta)K/C$, or $\ell^*\delta$ is a dominant trait with population size $o(K)$.

In the first case, for the k-th step, we have $\sigma_*^K = \sigma_k^K \to s_{k-1}$ in probability when $K \to +\infty$ and $\ell^* = \ell_k^*$. Other population sizes are negligible with respect to K. Given parameters $\varepsilon_* > 0$ and given $m > 0$ to be fixed later, we introduce:

$$\theta_*^K = \inf\Big\{ t \geq \sigma_*^K : N_{\ell^*}^K(t\log K) \notin \Big[\big(\frac{3-\ell^*\delta}{C} - 3\varepsilon_*\big)K, \big(\frac{3-\ell^*\delta}{C} + 3\varepsilon_*\big)K\Big] \\ \text{or } \sum_{\ell\neq\ell^*} N_\ell^K(t\log K) \geq m\varepsilon_* K \Big\}. \quad (4.1)$$

On the time interval $[\sigma_*^K \log K, \theta_*^K \log K]$, we proceed by induction on the traits. For $\ell \in \{1, \dots L\}$, having proved the convergence of $\beta_0^K(t), \dots \beta_{\ell-1}^K(t)$, to $\beta_0(t), \dots \beta_{\ell-1}(t)$, we will bound from below and above the population size $N_\ell^K(t)$ for $\ell \neq \ell^*$ by branching processes with immigration on each interval where $\beta_{\ell-1}$ is affine. Either $\theta_*^K \to +\infty$ when $K \to +\infty$ with probability tending to 1 and in the limit there is no further change of the resident population. Or, θ_*^K has a finite limit s_* in probability ($s_* = s_k$ for the k-th phase). In this case, we will show that at time θ_*^K, only two traits have sizes of order K and are then competing with each other, or $s_* = T_0$ and we stop the analysis. In the first case, there is a transition step, with duration of order 1, leading with large probability to the replacement of the resident population by the new trait.

In the second case, for the k-th phase, the time σ_*^K is chosen such that $\sigma_*^K = \sigma_k^K \to s_{k-1} + s$ for some small $s > 0$ in probability when $K \to +\infty$. We proceed similarly as before, replacing θ_*^K by

$$\widetilde{\theta}_*^K = \inf\Big\{ t \geq \sigma_*^K : \frac{\log(1 + N_{\ell^*}^K(t\log K))}{\log K} \notin \Big[\beta_{\ell^*}(t) - \varepsilon_*, \beta_{\ell^*}(t) + \varepsilon_*,\Big] \\ \text{or } \sum_{\ell\neq\ell^*} N_\ell^K(t\log K) \geq m\varepsilon_* N_{\ell^*}^K(t\log K) \Big\}. \quad (4.2)$$

4.2 Phase 1

Let us begin the induction on the phases and start with Phase 1. We study the stochastic process between times 0 and $s_1 \log K$. For the chosen initial condition (1.6), $\beta_\ell(0) = (1-\ell\alpha)_+$ and the resident trait is 0. We distinguish two cases: either $\tau < \delta$ or $\delta < \tau$.

4.2.1 Case $\tau < \delta$: a single phase

Using formulas (2.3) and (2.4) recursively, we obtain

$$\begin{aligned}
&\beta_0(t) = 1 && t_{0,1} = 0\\
&\beta_1(t) = \big(1-\alpha+(\tau-\delta)t\big) \vee (1-\alpha) \vee 0 = 1-\alpha, && t_{1,1} = 0\\
&\quad\vdots && \quad\vdots\\
&\beta_\ell(t) = \big(1-\ell\alpha+(\tau-\ell\delta)t\big) \vee (1-\ell\alpha) \vee 0 = (1-\ell\alpha)_+, && t_{\ell,1} = 0.
\end{aligned}$$

From (2.1), we obtain $s_1 = +\infty$, and $\beta(t)$ is constant. This is due to the fact that all non-resident traits have negative fitness and their size is kept of constant order due to mutations from the resident trait 0.

In the sequel, we denote by $BP_K(b,d,\beta)$ the distribution of the branching process, with individual birth rate $b \geq 0$, individual death rate $d \geq 0$ and initial value $Z_0^K = \lfloor K^\beta - 1 \rfloor \in \mathbb{N}$. We refer to Appendix A for the classical properties of $BP_K(b,d,\beta)$ which will be used to obtain the following results.
We also denote by $BPI_K(b,d,a,c,\beta)$ the distribution of a branching process with immigration, with individual birth rate $b \geq 0$, individual death rate $d \geq 0$, immigration rate $K^c e^{as}$ at time $s \geq 0$ with $a, c \in \mathbb{R}$ and same initial value. We refer to Appendix B.
Similarly, $LBDI_K(b,d,C,\gamma)$ is the distribution of a one-dimensional logistic birth and death process (see Appendix C) with individual birth rate $b \geq 0$ and individual death rate $d + Cn/K$ when the population size is n and immigration at predictable rate $\gamma(t) \geq 0$ at time t.

We introduce a parameter $\varepsilon_0 > 0$ and we choose K large enough so that $N_0^K(0) \in \Big[\big(\frac{3}{C} - \varepsilon_0\big)K, \big(\frac{3}{C} + \varepsilon_0\big)K\Big]$. Consider θ_*^K in (4.1) with $\ell^* = 0$, $\varepsilon_* = \varepsilon_0$ and $\sigma_*^K = 0$.

In order to couple the population with branching processes with immigration, we start with computing the arrival and death rates in the subpopulation of trait $\ell\delta$, for $\ell \geq 0$. For K large enough, arrivals in this population due to reproduction of trait $\ell\delta$ or transfer occur at time $t \leq \theta_*^K \wedge T$ at rate $N_\ell^K(t\log K)\left[(4-\ell\delta)(1-K^{-\alpha}) + \tau\,\frac{\sum_{\ell'=0}^{\ell-1} N_{\ell'}^K(t\log K)}{\sum_{\ell'=0}^{L} N_{\ell'}^K(t\log K)}\right]$ satisfying

$$\begin{aligned}
&N_\ell^K(t\log K)\left[4-\ell\delta-\varepsilon_0+\tau\frac{3-3C\varepsilon_0}{3+C(3+m)\varepsilon_0}\right]\\
&\quad\leq N_\ell^K(t\log K)\left[(4-\ell\delta)(1-K^{-\alpha}) + \tau\,\frac{\sum_{\ell'=0}^{\ell-1} N_{\ell'}^K(t\log K)}{\sum_{\ell'=0}^{L} N_{\ell'}^K(t\log K)}\right] \leq N_\ell^K(t\log K)\left[4-\ell\delta+\tau\right].
\end{aligned}$$

Arrivals due to incoming mutations from trait $(\ell-1)\delta$ occur at time $t \leq \theta_*^K \wedge T$ at rate $N_{\ell-1}^K(t\log K)(4-(\ell-1)\delta)K^{-\alpha}$.

Deaths occur at rate $N_\ell^K(t\log K)\left[1+\frac{C}{K}\sum_{\ell'=1}^L N_{\ell'}^K(t\log K)\right]$ satisfying

$$N_\ell^K(t\log K)\,(4-3C\varepsilon_0)\le N_\ell^K(t\log K)\left[1+\frac{C}{K}\sum_{\ell'=1}^{L}N_{\ell'}^K(t\log K)\right]\le N_\ell^K(t\log K)\,(4+C(3+m)\varepsilon_0)\,.$$

Step 1 Let us prove by induction on $\ell\ge 0$ the following bounds on the mutation rates: for all $t\le\theta_*^K\wedge T$, with probability converging to 1 as $K\to+\infty$,

$$K^{\beta_\ell(t)-\alpha-(\ell+1)\varepsilon_0}\le N_\ell^K(t\log K)(4-\ell\delta)K^{-\alpha}\le K^{\beta_\ell(t)-\alpha+(\ell+1)\varepsilon_0}.$$

For $\ell=0$, by definition of θ_*^K,

$$K^{1-\alpha-\varepsilon_0}\le N_0^K(t\log K)\,4K^{-\alpha}\le K^{1-\alpha+\varepsilon_0}\tag{4.3}$$

is clear. To proceed to $\ell=1$, we use standard coupling arguments to obtain

$$Z_{1,1}^K(t\log K)\le N_1^K(t\log K)\le Z_{1,2}^K(t\log K),\quad\forall t\le\theta_*^K\wedge T,\tag{4.4}$$

where $Z_{1,1}^K$ is a $BPI_K(4-\delta+\tau-\overline{C}\varepsilon_0,4+\overline{C}\varepsilon_0,0,1-\alpha-\varepsilon_0,1-\alpha-\varepsilon_0)$ and $Z_{1,2}^K$ is a $BPI_K(4-\delta+\tau,4-\overline{C}\varepsilon_0,0,1-\alpha+\varepsilon_0,1-\alpha+\varepsilon_0)$, with

$$\overline{C}=1+(1\vee\tau)C(6+m).$$

Note that the addition of ε_0 in the coefficient β of the upper-bounding branching process ensures that $c\le\beta$, so that we can use Theorem B.5(i) to deduce that as K tends to infinity, since $\tau-\delta<0$,

$$\frac{\log(1+Z_{1,1}^K(t\log K))}{\log K}\longrightarrow\big[1-\alpha-\varepsilon_0+(\tau-\delta-2\overline{C}\varepsilon_0)t\big]\vee[1-\alpha-\varepsilon_0]\ge 1-\alpha-\varepsilon_0=\beta_1(t)-\varepsilon_0$$

and

$$\frac{\log(1+Z_{1,2}^K(t\log K))}{\log K}\longrightarrow\big[1-\alpha+\varepsilon_0+(\tau-\delta+\overline{C}\varepsilon_0)t\big]\vee[1-\alpha+\varepsilon_0]\le 1-\alpha+\varepsilon_0=\beta_1(t)+\varepsilon_0,$$

in $L^\infty([0,T])$ and provided that ε_0 is small enough.

Assume now that the induction hypothesis is true for $\ell-1\ge 1$ and let us prove it for ℓ. Either $\ell\le 1+\lfloor 1/\alpha\rfloor$ and we have for all $t\le\theta_*^K\wedge T$, with probability converging to 1 as $K\to+\infty$,

$$K^{1-(\ell-1)\alpha-\alpha-\ell\varepsilon_0}\le N_{\ell-1}^K(t\log K)(4-(\ell-1)\delta)K^{-\alpha}\le K^{1-(\ell-1)\alpha-\alpha+\ell\varepsilon_0},\tag{4.5}$$

or $\ell>1+\lfloor 1/\alpha\rfloor$ and $N_{\ell-1}^K(t\log K)(4-(\ell-1)\delta)K^{-\alpha}=0$. Hence, with probability converging to 1,

$$Z_{\ell,1}^K(t\log K)\le N_\ell^K(t\log K)\le Z_{\ell,2}^K(t\log K),\quad\forall t\le\theta_*^K\wedge T,\tag{4.6}$$

where $Z_{\ell,1}^K$ is a $BPI_K(4-\ell\delta+\tau-\overline{C}\varepsilon_0,4+\overline{C}\varepsilon_0,0,(1-(\ell-1)\alpha)_+-\alpha-\ell\varepsilon_0,(1-\ell\alpha-\ell\varepsilon_0)_+)$ and $Z_{\ell,2}^K$ is a $BPI_K(4-\ell\delta+\tau,4-\overline{C}\varepsilon_0,0,(1-(\ell-1)\alpha)_+-\alpha+\ell\varepsilon_0,(1-\ell\alpha+\ell\varepsilon_0)_+)$. Note that when $\ell>1+\lfloor 1/\alpha\rfloor$, the lower bound has nonzero but negligible immigration, hence the

comparison is true only on the event where there is no immigration in $Z^K_{\ell,1}$ on $[0, T\log K]$, which has probability converging to 1 (see Lemma B.7).

If $\ell \leq \lfloor 1/\alpha \rfloor$, we use Theorem B.5 (i) to deduce that as K tends to infinity, uniformly for $t \in [0,T]$,

$$\frac{\log(1 + Z^K_{\ell,1}(t\log K))}{\log K} \longrightarrow \big[1 - \ell\alpha - \ell\varepsilon_0 + (\tau - \ell\delta - 2\overline{C}\varepsilon_0)t\big] \vee [1 - \ell\alpha - \ell\varepsilon_0] \geq \beta_\ell(t) - \ell\varepsilon_0,$$

and

$$\frac{\log(1 + Z^K_{\ell,2}(t\log K))}{\log K} \longrightarrow \big[1 - \ell\alpha + \ell\varepsilon_0 + (\tau - \ell\delta + \overline{C}\varepsilon_0)t\big] \vee [1 - \ell\alpha + \ell\varepsilon_0] \leq \beta_\ell(t) + \ell\varepsilon_0.$$

If $\lfloor 1/\alpha \rfloor + 1 \leq \ell \leq L$, we use Theorem B.5 (iii) and (B.15) (assuming that ε_0 is small enough) to prove that $\log(1+Z^K_{\ell,i}(t\log K))/\log K$ converges to 0 in $L^\infty([0,T])$, and N^K_ℓ gets extinct with high probability for $t \leq \theta^K_* \wedge T$.

We deduce that, with probability converging to 1, (4.5) is satisfied with $\ell - 1$ replaced by ℓ. This completes the proof of (4.5) by induction. In particular, on the time interval $[0, \theta^K_* \wedge T]$, $\log(1 + N^K_\ell(t\log K))/\log K$ converges to $\beta_\ell(t) = (1 - \ell\alpha)_+$.

Step 2 It remains to prove that $\theta^K_* > T$ with probability converging to 1. Using the previous steps and recalling that $\beta_\ell(t) = \beta_\ell(0) = (1 - \ell\alpha)_+$, we have, with probability converging to 1, that $\sup_{t\in[0,2T]} |\beta^K_\ell(t) - \beta_\ell(t)| < \alpha/2$, and thus, for all $t \leq \theta^K_* \wedge 2T$,

$$\sum_{\ell=1}^{L} N^K_\ell(t\log K) \leq K^{\max_{1\leq\ell\leq L,\, t\in[0,2T]}\, \beta_\ell(t) + \frac{\alpha}{2}} = K^{1-\frac{\alpha}{2}}. \tag{4.7}$$

Hence, we have for all $t \leq \theta^K_* \wedge 2T$ that

$$Z^K_{0,1}(t\log K) \leq N^K_0(t\log K) \leq Z^K_{0,2}(t\log K), \tag{4.8}$$

where $Z^K_{0,1}$ is a $LBDI_K(4(1-K^{-\alpha}), 1+CK^{-\alpha/2}, C, 0)$ and $Z^K_{0,2}$ is a $LBDI_K(4, 1, C, 0)$. Applying Lemma C.1(i) to the processes $Z^K_{0,i}$, $i = 1, 2$, we deduce that $\theta^K_* > T$ with probability converging to 1 when $K \to +\infty$.

4.2.2 Case $\tau > \delta$: Phase 1

We already computed in Lemma 5.1 (5.4) the functions $\beta_\ell(t)$ and the duration

$$s_1 = \frac{\alpha}{\tau - \delta}$$

of the first phase. Recall that, on $[0, s_1]$, $\beta_\ell(t) = (1 - \ell\alpha + (\tau - \delta)t)_+$. Notice that $t_{\lfloor 1/\alpha \rfloor,1} < s_1 < t_{\ell,1}$ for all $\ell > \lfloor 1/\alpha \rfloor$, where $t_{\ell,k}$ are defined in (2.6), so that, on $[0, s_1]$,

- for $\ell < \lfloor 1/\alpha \rfloor + 1$, $\beta_\ell(t)$ is an affine function with slopes $\tau - \delta$ except for the resident population;

- for $\ell > \lfloor 1/\alpha \rfloor + 1$, $\beta_\ell(t) = 0$;
- for $\ell = \lfloor 1/\alpha \rfloor + 1$, there is a change of slope at time $t_{\lfloor 1/\alpha \rfloor, 1}$.

We define the same stopping time θ_*^K as above. The comparisons of arrival and death rates of $N_\ell^K(t \log K)$ are the same as in Section 4.2.1. We proceed by induction over $\ell \geq 0$ as above.

For $\ell = 0$, (4.3) remains true by definition of θ_*^K. To proceed to $\ell = 1$, we observe that (4.4) remains valid and Theorem B.5(i) gives as before that as K tends to infinity, since $\tau - \delta > 0$,

$$\frac{\log(1 + Z_{1,1}^K(t \log K))}{\log K} \longrightarrow \big[1 - \alpha - \varepsilon_0 + (\tau - \delta - 2\overline{C}\varepsilon_0)t\big] \vee [1 - \alpha - \varepsilon_0] \geq \beta_1(t) - 2\overline{C}\varepsilon_0 T,$$

and

$$\frac{\log(1 + Z_{1,2}^K(t \log K))}{\log K} \longrightarrow \big[1 - \alpha + \varepsilon_0 + (\tau - \delta + \overline{C}\varepsilon_0)t\big] \vee [1 - \alpha + \varepsilon_0] \leq \beta_1(t) + \varepsilon_0(1 + \overline{C}T),$$

in $L^\infty([0, s_1 \wedge T])$ and provided that ε_0 is small enough.

For $2 \leq \ell \leq 1 + \lfloor 1/\alpha \rfloor$, the induction relation (4.5) is modified as follows. For all $\ell \geq 1$, let

$$C_\ell = \ell + 2\overline{C}T. \tag{4.9}$$

We assume that, for all $t \leq \theta_*^K \wedge s_1 \wedge T$,

$$\beta_{\ell-1}(t) - \alpha - C_{\ell-1}\varepsilon_0 \leq \frac{\log(N_{\ell-1}^K(t \log K)(4 - (\ell-1)\delta)K^{-\alpha})}{\log K} \leq \beta_{\ell-1}(t) - \alpha + C_{\ell-1}\varepsilon_0. \tag{4.10}$$

Hence, with probability converging to 1,

$$Z_{\ell,1}^K(t \log K) \leq N_\ell^K(t \log K) \leq Z_{\ell,2}^K(t \log K), \quad \forall t \leq \theta_*^K \wedge s_1 \wedge T, \tag{4.11}$$

where $Z_{\ell,1}^K$ is a $BPI_K(4 - \ell\delta + \tau - \overline{C}\varepsilon_0, 4 + \overline{C}\varepsilon_0, \tau - \delta, 1 - \ell\alpha - C_{\ell-1}\varepsilon_0, (1 - \ell\alpha - C_{\ell-1}\varepsilon_0)_+)$ and $Z_{\ell,2}^K$ is a $BPI_K(4 - \ell\delta + \tau, 4 - \overline{C}\varepsilon_0, \tau - \delta, 1 - \ell\alpha + C_{\ell-1}\varepsilon_0, (1 - \ell\alpha + C_{\ell-1}\varepsilon_0)_+)$.

Note that, in this Phase 1, $\beta_{\ell-1}(t)$ is affine on $[0, s_1]$, so we can apply Theorem B.5 on the whole interval $[0, s_1]$. If $\ell \leq \lfloor 1/\alpha \rfloor$, we apply Theorem B.5 (i) and if $\ell = 1 + \lfloor 1/\alpha \rfloor$, $\beta_\ell(0) = 0$ so we apply Theorem B.5 (ii), using that $a = \tau - \delta > r = \tau - \ell\delta + \overline{C}\varepsilon_0$ for ε_0 small enough. We deduce that, in both cases, as K tends to infinity, for all $t \leq \theta_*^K \wedge s_1 \wedge T$,

$$(\beta_\ell(t) - C_{\ell-1}\varepsilon_0)_+ \leq \frac{\log(1 + N_\ell^K(t \log K))}{\log K} \leq (1 - \ell\alpha - C_{\ell-1}\varepsilon_0 + (\tau - \delta)t)_+ \leq \beta_\ell(t) + C_{\ell-1}\varepsilon_0. \tag{4.12}$$

Hence, we have proved (4.10) for $\ell + 1$ and the induction step is complete. We also deduce from (B.15) that $N_{1+\lfloor 1/\alpha \rfloor}^K(t \log K) = 0$ for all t in a closed interval of $[0, s_1]$ included in the complement of the support of $(1 - \ell\alpha - C_{\ell-1}\varepsilon_0 + (\tau - \delta)t)_+$.

For $\ell = 2 + \lfloor 1/\alpha \rfloor$, because $\beta_{1+\lfloor 1/\alpha \rfloor}$ has a change of slope at time

$$t_{\lfloor 1/\alpha \rfloor, 1} = \frac{\alpha - (1 - \lfloor 1/\alpha \rfloor \alpha)}{\tau - \delta},$$

the bounds (4.10) on the immigration rate do not allow to apply Theorem B.5 on the whole interval $[0, s_1]$. Instead, we first apply them successively on the intervals $[0, t_{\lfloor 1/\alpha \rfloor,1}]$ and $[t_{\lfloor 1/\alpha \rfloor,1}, s_1]$. On the first interval, we have the bounds

$$Z^K_{\ell,1}(t\log K) \le N^K_\ell(t\log K) \le Z^K_{\ell,2}(t\log K), \quad \forall t \le \theta^K_* \wedge t_{\lfloor 1/\alpha \rfloor,1} \wedge T,$$

where $Z^K_{\ell,1}$ is a $BPI_K(4-\ell\delta+\tau-\overline{C}\varepsilon_0, 4+\overline{C}\varepsilon_0, 0, -\alpha-C_{\ell-1}\varepsilon_0, 0)$ and $Z^K_{\ell,2}$ is a $BPI_K(4-\ell\delta+\tau, 4-\overline{C}\varepsilon_0, 0, -\alpha+C_{\ell-1}\varepsilon_0, 0)$. We apply Theorem B.5 (iii) to deduce that, with probability converging to 1, $N^K_\ell(t\log K) = 0$ for all $t \in [0, t_{\lfloor 1/\alpha \rfloor,1} \wedge \theta^K_* \wedge T]$.
On the second interval, we obtain similar bounds with $Z^K_{\ell,1}$ a $BPI_K(4-\ell\delta+\tau-\overline{C}\varepsilon_0, 4+\overline{C}\varepsilon_0, \tau-\delta, -\alpha-C_{\ell-1}\varepsilon_0, 0)$ and $Z^K_{\ell,2}$ is a $BPI_K(4-\ell\delta+\tau, 4-\overline{C}\varepsilon_0, \tau-\delta, -\alpha+C_{\ell-1}\varepsilon_0, 0)$. Because $(\tau-\delta)(s_1-t_{\lfloor 1/\alpha \rfloor,1}) < \alpha$, we apply Theorem B.5 (ii) to deduce that, with probability converging to 1, $N^K_\ell(t\log K) = 0$ for all $t \in [t_{\lfloor 1/\alpha \rfloor,1} \wedge \theta^K_* \wedge T, s_1 \wedge \theta^K_* \wedge T]$.

For $\ell > 2+\lfloor 1/\alpha \rfloor$, we proceed similarly to prove by induction that, with probability converging to 1, $N^K_\ell(t\log K) = 0$ for all $t \le s_1 \wedge \theta^K_* \wedge T$.

Using the comparison argument with logistic birth-death processes (4.8) on the interval $[0, \theta^*_K \wedge (s_1-\eta) \wedge T]$, we can prove as in the previous section that, for all $\eta > 0$, $\theta^K_* > (s_1-\eta)\wedge T$ with probability converging to 1.

To conclude, since ε_0 in (4.12) is arbitrary, we have proved that, for all $\eta > 0$, $\log(1 + N^K_\ell(t\log K))/\log K$ converges to $\beta_\ell(t)$ in $L^\infty([0, (s_1-\eta)\wedge T])$.

4.2.3 Case $\tau > \delta$: Intermediate Step 1

The goal of the intermediate step is to prove that, on a time interval $[\theta^K_* \log K, \theta^K_* \log K + T(\varepsilon, m)]$ with $T(\varepsilon, m)$ to be defined below, the two traits $\ell^*_1\delta = 0$ and $\ell^*_2\delta = \delta$ are of size-order K and population with trait 0 is declining below a small threshold while population with trait δ reaches a neighborhood of its equilibrium $K\bar{n}(\delta) = \frac{(3-\delta)K}{C}$.

Let us first show that $\theta^K_* < s_1+\eta$ with probability converging to one, for any $\eta > 0$. For this, we observe that the bounds of (4.10) and (4.11) are actually true for all $t \le \theta^K_* \wedge T$ provided we use $(1-\ell\alpha+(\tau-\delta)t)_+$ instead of $\beta_\ell(t)$ inside (4.10), since $\beta_\ell(t)$ is constructed only on $[0, s_1]$ in Phase 1. This gives with high probability for all $t \le \theta^K_* \wedge T$

$$(1-\ell\alpha+(\tau-\delta)t)_+ - C\ell\varepsilon_0 \le \frac{\log(1+N^K_\ell(t\log K))}{\log K} \le (1-\ell\alpha+(\tau-\delta)t)_+ + C\ell\varepsilon_0. \quad (4.13)$$

In particular, if $\theta^K_* \ge s_1+\eta$ with positive probability, we would obtain, with high probability on this event

$$\frac{\log(1+N^K_1((s_1+\eta)\log K))}{\log K} \ge 1-\alpha-C_1\varepsilon_0+(\tau-\delta)(s_1+\eta),$$

which is larger than 1 provided ε_0 is small enough. This would contradict the assumption that $\theta^K_* > s_1+\eta$. Hence, $\lim_{K\to+\infty}\theta^K_* = s_1$ in probability.

In the previous phase, we used bounds on $N^K_0(t\log K)$ until time $s_1-\eta$. We now need to extend them until θ^K_*. In this case, (4.7) is not true anymore. Therefore, assuming K

large enough to get $4K^{-\alpha} < Cm\varepsilon_0$, we couple $N_0^K(t)$ with two logistic processes $Z_{0,1}^K$ and $Z_{0,2}^K$ of respective law $LBDI_K(4,1,C,0)$ and $LBDI_K(4 - Cm\varepsilon_0, 1 + \frac{\tau m\varepsilon_0}{3/C - 3\varepsilon_0} + Cm\varepsilon_0, C, 0)$. The equilibrium density of $Z_{0,1}^K$ is $3/C$ but the one of $Z_{0,2}^K$ is

$$\bar{z}_{0,2} := \frac{3}{C} - \varepsilon_0 \left(2m + \frac{\tau m}{3 - 3\varepsilon_0 C}\right).$$

We choose m small enough to have

$$0 < 2m + \frac{\tau m}{3 - 3\varepsilon_0 C} < \frac{1}{3}. \tag{4.14}$$

Then, observing that $Z_{0,2}^K(0) \in \Big[\big(\bar{z}_{0,2} - 4\varepsilon_0/3\big)K, \big(\bar{z}_{0,2} + 4\varepsilon_0/3\big)K\Big]$ we can apply Lemma C.1(i) to $Z_{0,2}^K$ with $\varepsilon = 4\varepsilon_0/3$ to deduce that

$$\lim_{K\to\infty} \mathbb{P}\Big(\forall t \in [0, s_1 + \eta], \frac{Z_{0,2}^K(t\log K)}{K} \le \frac{3}{C} + 3\varepsilon_0\Big) = 1.$$

Applying similarly Lemma C.1(i) to $Z_{0,1}^K$, we obtain

$$\lim_{K\to\infty} \mathbb{P}\Big(\frac{N_0^K(\theta_*^K \log K)}{K} \in [\frac{3}{C} - 3\varepsilon_0, \frac{3}{C} + 3\varepsilon_0]\Big) = 1.$$

Hence, by construction of θ_*^K, we have with probability converging to 1

$$\sum_{\ell=1}^{L} N_\ell^K(\theta_*^K \log K) \ge m\varepsilon_0 K.$$

Remark 4.1 *The constraint* (4.14) *on the constant $m > 0$ ensures that, when trait* 0 *is resident, with high probability, at time θ_*^K, a second trait emerges. Similar constraints can be defined for any other resident trait $\ell\delta < 3$ and we can choose m such that all the constraints are satisfied. In all the proof, we assume that such $m > 0$ is chosen.*

Using Formula (4.13) and the convergence of θ_*^K to s_1, we show immediately that, with probability converging to 1,

$$\sum_{\ell=2}^{L} N_\ell^K(\theta_*^K \log K) \le K^{1-\alpha/2}. \tag{4.15}$$

Hence $N_1^K(\theta_*^K \log K) \ge m\varepsilon_0 K/2$ for K large enough.

It is now possible to use Markov property by conditioning at time $\theta_*^K \log K$. We distinguish whether the emerging trait δ can survive on its own ($\delta < 3$) or not ($\delta > 3$).

Case $\tau > \delta$: Intermediate step 1, case $\delta < 3$ For $\delta < 3$, we first claim that there exists $s > 0$ small enough such that

$$\sum_{\ell=2}^{L} N_\ell^K(t\log K) \le K^{1-\alpha/4}, \quad \forall\, t \in [\theta_*^K, \theta_*^K + s].$$

This can be obtained from the continuity argument of Lemma B.9. Then, we can apply Lemma C.3(i) with

$$b_0^K(t) = 4(1-K^{-\alpha}), \quad b_1^K(t) = (4-\delta)(1-K^{-\alpha}),$$
$$d_0^K(t) = d_1^K(t) = 1 + \left(\frac{C}{K} + \frac{\tau}{\sum_{\ell=0}^L N_\ell^K(t)}\right)\sum_{\ell=2}^L N_\ell^K(t),$$
$$\tau^K(t) = \tau\frac{N_0^K(t)+N_1^K(t)}{\sum_{\ell=0}^L N_\ell^K(t)}, \quad \gamma_0^K(t) = 0, \quad \gamma_1^K(t) = 4K^{-\alpha}N_0^K(t),$$

which converge respectively to $b_0 = 4$, $b_1 = 4-\delta$, $d_0 = d_1 = 1$, τ, 0 and 0. Note that Point (i) of Lemma C.3 applies here since $r_0 = b_0 - d_0 = 3 > 0$, $r_1 = b_1 - d_1 = 3-\delta > 0$ and $S = \tau - \delta > 0$. We obtain that there exists a finite time $T(m,\varepsilon_0)$ such that with large probability,

$$N_0^K\big(\theta_*^K \log K + T(m,\varepsilon_0)\big) \le m\varepsilon_0 K \quad \text{and} \quad N_1^K\big(\theta_*^K \log K + T(m,\varepsilon_0)\big) \in \Big[\frac{3-\delta}{C} - 3\varepsilon_0, \frac{3-\delta}{C} + 3\varepsilon_0\Big].$$

Hence, the first intermediate phase ends at the time

$$\sigma_1^K \log K = \theta_*^K \log K + T(m,\varepsilon_0)$$

where $\sigma_1^K \to s_1$ in probability. Using (4.13), the continuity argument of Lemma B.9 and the fact that $(\tau-\delta)s_1 = \alpha$, the population state at time σ_1 satisfies with probability converging to 1

$$K^{1-\varepsilon_0} \le N_0^K(\sigma_1^K \log K) \le m\varepsilon_0 K, \quad \frac{N_1^K(\sigma_1^K \log K)}{K} \in \Big[\frac{3-\delta}{C} - \varepsilon_1, \frac{3-\delta}{C} + \varepsilon_1\Big],$$
$$\frac{\log(1+N_\ell^K(\sigma_1^K \log K))}{\log K} \in \Big[1-(\ell-1)\alpha - \varepsilon_1, 1-(\ell-1)\alpha+\varepsilon_1\Big] = [\beta_\ell(s_1)-\varepsilon_1, \beta_\ell(s_1)+\varepsilon_1],$$
$$\forall\, 2 \le \ell \le 1 + \lfloor 1/\alpha \rfloor,$$
$$N_\ell^K(\sigma_1^K \log K) = 0, \quad \forall\, 2 + \lfloor 1/\alpha \rfloor \le \ell \le L.$$

where

$$\varepsilon_1 = (C_L \vee 3)\varepsilon_0. \tag{4.16}$$

We are then ready to proceed to phase 2 (see Section 4.3).

Case $\tau > \delta$: Intermediate step 1, case $\delta > 3$ When $\delta > 3$, we need to apply Lemma C.3(iii) instead of (i) since $r_1 = 3-\delta < 0$. Using (4.13) and the continuity argument of Lemma B.9 as above, we obtain for all $s > 0$ small enough, with probability converging to 1,

$$\frac{\log(1+N_\ell^K((\theta_*^K+s)\log K))}{\log K} \in \Big[1-(\ell-1)\alpha-\varepsilon_1, 1-(\ell-1)\alpha+\varepsilon_1\Big], \quad \forall\, 2 \le \ell \le 1+\lfloor 1/\alpha \rfloor,$$
$$N_\ell^K((\theta_*^K+s)\log K)) = 0, \quad \forall\, 2+\lfloor 1/\alpha \rfloor \le \ell \le L,$$

with ε_1 defined in (4.16). Without loss of generality, we can take $s > 0$ small enough to apply Lemma C.3(iii).

We define the end of the first intermediate phase as

$$\widetilde{\sigma}_1^K = \theta_*^K + s$$

and

$$K^{1-\varepsilon_1} \le N_0^K(\widetilde{\sigma}_1^K \log K) \le K^{-s\rho} N_1^K(\widetilde{\sigma}_1^K), \quad N_1^K(\widetilde{\sigma}_1^K \log K) \le K^{1-s\rho}.$$

4.3 Phase $k+1$

Recall that when $\tau < \delta$ there is only one phase. So now, we only consider $\tau > \delta$.

Assume that both Phase $k \geq 1$ and intermediate step k are completed and that $s_k < T_0$ (otherwise, we stop the induction at the end of phase $[s_{k-1}, s_k]$). Two cases may occur: either the emerging trait becomes resident during Phase $k+1$, or it becomes only a dominant trait. The latter case can occur when the trait is fit ($\ell^*_{k+1}\delta < 3$) but its population size is small, or when it is unfit ($\ell^*_{k+1}\delta > 3$).

We proceed as in Phase 1 by induction over $\ell \in \{0, \dots, L\}$. For each ℓ, we couple N^K_ℓ with branching processes with immigration on each time interval included in $[s_k, s_{k+1}]$ where $\beta_{\ell-1}$ is affine.

In Phase 1, we took care to give bounds involving explicit constants C_ℓ for sake of precision. From now on, we use a constant C_* which may change from line to line.

4.3.1 Phase $k+1$, case 1

We assume that, $\max_{0\leq\ell\leq L} \beta_\ell(s_k+s) = 1$ and $\ell^*_{k+1}\delta < 3$, and this maximum is attained only for ℓ^*_k and ℓ^*_{k+1}, since $s_k < T_0$.

The induction assumption in this case is as follows: suppose that, for all $\varepsilon_k > 0$ small enough, we have constructed σ^K_k such that $\sigma^K_k \log K$ is a stopping time and

- σ^K_k converges in probability to s_k;
- $\frac{N^K_{\ell^*_{k+1}}(\sigma^K_k \log K)}{K} \in \left[\frac{3-\delta\ell^*_{k+1}}{C} - \varepsilon_k, \frac{3-\delta\ell^*_{k+1}}{C} + \varepsilon_k\right]$;
- $K^{1-\varepsilon_k} \leq N^K_{\ell^*_k}(\sigma^K_k \log K) \leq \frac{m}{2}\varepsilon_k K$;
- for all $\ell \neq \ell^*_k, \ell^*_{k+1}$, either $N^K_\ell(\sigma^K_k \log K) = 0$ if $\beta_\ell(s_k) = 0$, or otherwise

$$\beta_\ell(s_k) - \varepsilon_k \leq \frac{\log(1 + N^K_\ell(\sigma^K_k \log K))}{\log K} \leq \beta_\ell(s_k) + \varepsilon_k < 1.$$

We now give the proof of the induction step. Let us define θ^K_* as in (4.1) with $\ell^* = \ell^*_{k+1}$, $\varepsilon_* = \varepsilon_k$ and $\sigma^K_* = \sigma^K_k$.

Observe first that, by definition of θ^K_*,

$$1 - \varepsilon_* \leq \frac{\log(1 + N^K_{\ell^*}(t \log K))}{\log K} \leq 1 + \varepsilon_*, \quad \forall t \in [\sigma^K_*, \theta^K_*].$$

Our goal is to obtain bounds on N^K_ℓ by induction on $\ell \neq \ell^*$.

Induction initialization: If $\ell^* = 0$, we start the induction at $\ell = 1$; otherwise, we start it at $\ell = 0$.

In the first case, using Markov property at time σ^K_* which converges to s_k, we can proceed exactly as in Phase 1 to prove that, with high probability for all $t \in [\sigma^K_*, \theta^K_* \wedge T]$,

$$\beta_1(s_k) + (\tau - \delta)(t - s_k) - C_*\varepsilon_* \leq \frac{\log(N^K_1(t \log K))}{\log K} \leq \beta_1(s_k) + (\tau - \delta)(t - s_k) + C_*\varepsilon_*.$$

Note that, for $s_k \le t \le s_{k+1}$, $\beta_1(s_k) + (\tau - \delta)(t - s_k) = \beta_1(t)$, so that we recover bounds of the form (4.10).

In the second case, since there is no incoming mutation for trait 0, we can bound the process $N_0^K(\sigma_*^K + t)$ with branching processes with constant parameters, and, if $\beta_0(s_k) > 0$, we deduce from Lemma A.1 that

$$\beta_0(s_k) + (\ell^*\delta - \tau)(t - s_k) - C_*\varepsilon_* \le \frac{\log(N_0^K(t\log K))}{\log K} \le \beta_0(s_k) + (\ell^*\delta - \tau)(t - s_k) + C_*\varepsilon_*,$$

and that $N_0^K(t\log K) = 0$ for all $t \ge \sigma_*^K$ if $\beta_0(s_k) = 0$ (recall that no incoming mutation can enter the sub-population with trait 0).

Induction step: Assume that we have proved that, with high probability for all $t \in [\sigma_*^K, \theta_*^K \wedge s_{k+1} \wedge T]$,

$$\beta_{\ell-1}(t) - C_*\varepsilon_* \le \frac{\log(N_{\ell-1}^K(t\log K))}{\log K} \le \beta_{\ell-1}(t) + C_*\varepsilon_*.$$

and that $N_{\ell-1}^K(t\log K) = 0$ for all $t \in [\sigma_*^K, \theta_*^K \wedge s_{k+1} \wedge T]$ such that $\beta_{\ell-1}(s) = 0$ for all $s \in [(t - \varepsilon_*) \vee s_k, t + \varepsilon^*]$. Our goal is to prove that this holds true with ℓ replaced by $\ell + 1$. We split the interval $[s_k, s_{k+1}]$ into subintervals where $\beta_{\ell-1}$ is affine. We proceed inductively on each of these subintervals by coupling with branching processes with immigration as in Phase 1. Let us detail the computation for the first subinterval, say $[s_k, t_1]$. On this interval, we introduce a and c such that

$$\beta_{\ell-1}(t) = c + \alpha + a\ (t - s_k), \quad \forall t \in [s_k, t_1],$$

to be coherent with the notations of Appendix B. We can then construct as in Phase 1 branching processes with immigration bounding from above and below $N_\ell^K(t\log K)$ for all $t \in [\sigma_*^K, t_1 \wedge \tau_*^K \wedge T]$ with distributions $BPI(4 - \ell\delta + \tau\mathbb{1}_{\ell>\ell^*} \pm C_*\varepsilon_*, 4 - \ell^*\delta + \tau\mathbb{1}_{\ell<\ell^*} \mp C_*\varepsilon_*, a, c \pm C_*\varepsilon_*, \beta_\ell(s_k) \pm C_*\varepsilon_*)$ if $\beta_\ell(s_k) > 0$, or $BPI(4 - \ell\delta + \tau\mathbb{1}_{\ell>\ell^*} \pm C_*\varepsilon_*, 4 - \ell^*\delta + \tau\mathbb{1}_{\ell<\ell^*} \mp C_*\varepsilon_*, a, c \pm C_*\varepsilon_*, 0)$ otherwise. We shall consider several cases.

Case (a) Assume that $\beta_\ell(s_k) > 0$. Then, applying Theorem B.5 (i) to the bounding processes, we deduce that, with high probability on the time interval $[\sigma_*^K, \theta_*^K \wedge t_1 \wedge T]$,

$$\begin{aligned}&(\beta_\ell(s_k) + S(\ell\delta; \ell^*\delta)\,(t - s_k) - C_*\varepsilon_*) \vee (\beta_{\ell-1}(t) - \alpha - C_*\varepsilon_*) \vee 0\\ &\quad \le \frac{\log(N_\ell^K(t\log K))}{\log K} \le (\beta_\ell(s_k) + S(\ell\delta; \ell^*\delta)\,(t - s_k) + C_*\varepsilon_*) \vee (\beta_{\ell-1}(t) - \alpha + C_*\varepsilon_*) \vee 0.\end{aligned}$$

Case (b) Assume that $\beta_\ell(s_k) = 0$ and the time $t_{\ell-1,k}$ defined in (2.6) satisfies $t_{\ell-1,k} < t_1$. This last inequality implies that $c < 0$ and then $t_{\ell-1,k} = s_k - c/a$. Applying Theorem B.5 (ii) to the bounding processes, we deduce that, with high probability on the time interval $[\sigma_*^K, \theta_*^K \wedge t_1 \wedge T]$,

$$\begin{aligned}&[(S(\ell\delta; \ell^*\delta) - C_*\varepsilon_*) \vee a] \vee \left(t - s_k + \frac{c - C_*\varepsilon_*}{a}\right) \vee 0\\ &\qquad \le \frac{\log(N_\ell^K(t\log K))}{\log K} \le [(S(\ell\delta; \ell^*\delta) + C_*\varepsilon_*) \vee a] \vee \left(t - s_k + \frac{c + C_*\varepsilon_*}{a}\right) \vee 0.\end{aligned}$$

The bound can be rewritten as

$$[S(\ell\delta; \ell^*\delta) \vee a](t - t_{\ell-1,k})_+ - C_*\varepsilon_* \le \frac{\log(N_\ell^K(t\log K))}{\log K} \le [S(\ell\delta; \ell^*\delta) \vee a](t - t_{\ell-1,k})_+ + C_*\varepsilon_*.$$

Case (c) Assume that $\beta_\ell(s_k) = 0$ and $t_{\ell-1,k} > t_1$. Then, applying Theorem B.5 (ii) or (iii) to the bounding processes, we deduce that, with high probability on the time interval $[\sigma_*^K, \theta_*^K \wedge t_1 \wedge T]$, $N_\ell^K(t \log K) = 0$.

Summing up all these cases, it appears that we can extend $\beta_\ell(t)$ on the interval $[s_k, t_1]$ as in (2.4) to obtain, in all cases,

$$\beta_\ell(t) - C_*\varepsilon_* \leq \frac{\log(N_\ell^K(t \log K))}{\log K} \leq \beta_\ell(t) + C_*\varepsilon_*. \tag{4.17}$$

We proceed similarly for the other subintervals and deduce that (4.17) holds true with high probability on the time interval $[\sigma_*^K, \theta_*^K \wedge s_{k+1} \wedge T]$. This finishes the induction on $\ell \neq \ell^*$.

As in Phase 1, we can control $N_{\ell^*}^K(t)$ by logistic branching processes to show that $\theta_*^K \geq (s_{k+1} - \eta) \wedge T$ with probability converging to 1 for all $\eta > 0$.

It also follows from (B.15) that $N_\ell^K(t \ \log K) = 0$ with high probability on close subintervals of $\text{int}\{\beta_\ell = 0\} \cap [s_k, s_{k+1})$.

4.3.2 Intermediate step $k+1$, case 1

As in Intermediate step 1, we first extend the inequalities (4.17) to $[\sigma^K *, \theta_*^K \wedge T]$. These inequalities involve the $\beta_\ell(t)$ that are constructed in Phase $k+1$ only until s_{k+1}. Explicit expressions of $\beta_\ell(t)$, that we do not develop here, can be obtained using the recursion in (2.4). Using these formulas instead of $\beta_\ell(t)$ in (4.17), we obtain inequalities valid on $[\sigma^K *, \theta_*^K \wedge T]$, with probability converging to 1.

At time s_{k+1}, there exists at least one $\ell_{k+2}^* \neq \ell_{k+1}^*$ such that $\beta_{\ell_{k+2}^*}(s_{k+1}) = 1$ and $S(\ell_{k+2}^*\delta, \ell_{k+1}^*\delta) > 0$. Proceeding by contradiction as in Intermediate step 1, this allows us to prove that

$$\lim_{K \to +\infty} \theta_*^K = s_{k+1}.$$

This ends the proof of Theorem 2.1 if $s_{k+1} = T_0$. If $s_{k+1} < T_0$, as in Intermediate step 1, with high probability, for some $\kappa > 0$,

$$\sum_{\ell \notin \{\ell_{k+1}^*, \ell_{k+2}^*\}} N_\ell^K(\theta_*^K \log K) \leq K^{1-\kappa},$$

and $N_{\ell_{k+2}^*}^K(\theta_*^K \log K) \geq m\varepsilon_* K/2$.

Distinguishing as in Section 4.2.3 whether the emerging trait $\ell_{k+2}^*\delta$ is above 3 or not, we can define a time $\sigma_{k+1}^K \log K$ or $\widetilde{\sigma}_{k+1}^K \log K$ satisfying the recursion properties stated at the beginning of Section 4.3.1. The fact that the populations such that $\beta_\ell(s_{k+1}) = 0$ are actually extinct at time $\sigma_{k+1}^K \log K$ follows from (B.15) and from the fact that, by definition of T_0, we also have $\beta_\ell(s_{k+1} - \varepsilon) = 0$ for $\varepsilon > 0$ small enough.

This ends the Phase $k+1$ case 1 and its intermediate step.

4.3.3 Phase $k+1$, case 2

We assume here that $s_k < T_0$ and $\max_{0\le\ell\le L}\beta_\ell(s_k) < 1$ or $\ell^*_{k+1}\delta > 3$ and the maximum is attained only for ℓ^*_k and ℓ^*_{k+1}.

Here, the induction assumption is as follows: assume that, for all $\varepsilon_k > 0$ small enough, for all $s > 0$ small enough, we have constructed a stopping time $\widetilde{\sigma}^K_k \log K$ such that

- $\widetilde{\sigma}^K_k$ converges in probability to $s_k + s$;
- for all $\ell \neq \ell^*_{k+1}$, either $N^K_\ell(\sigma^K_k \log K) = 0$ if $\beta_\ell(s_k+s) = 0$, or otherwise

$$\beta_\ell(s_k+s) - \varepsilon_k \le \frac{\log(1+N^K_\ell(\sigma^K_k\log K))}{\log K} \le \beta_\ell(s_k+s) + \varepsilon_k < \beta_{\ell^*_{k+1}}(s_k+s) < 1.$$

Let us define $\widetilde{\theta}^K_*$ as in (4.2) with $\ell^* = \ell^*_{k+1}$, $\varepsilon_* = \varepsilon_k$ and $\sigma^K_* = \widetilde{\sigma}^K_k$. As in the Phase $k+1$, case 1, our goal is to obtain bounds on N^K_ℓ, by an induction on ℓ. Either the trait $\ell^*\delta$ remains dominant during the whole phase ($\sup_{t\in[s_k,s_{k+1}]}\beta_{\ell^*}(t) < 1$, see case (a) below) or it becomes resident during the phase (case (b)).

Case (a) In this case, there is no density dependence since the whole population is of size order $o(K)$. We can proceed exactly as in Phase $k+1$, case 1, with the use of the fitness $\widehat{S}$ instead of S.

With high probability for all $t \in [\sigma^K_*, \widetilde{\theta}^K_* \wedge s_{k+1} \wedge T]$, for all $\ell \in \{0,\dots L\}$,

$$\beta_\ell(t) - C_*\varepsilon_* \le \frac{\log(N^K_\ell(t\log K))}{\log K} \le \beta_\ell(t) + C_*\varepsilon_* \tag{4.18}$$

and that $N^K_\ell(t\log K) = 0$ for all $t \in [\sigma^K_*, \widetilde{\theta}^K_* \wedge s_{k+1} \wedge T]$ such that $\beta_{\ell-1}(s) = 0$ for all $s \in [(t-\varepsilon_*)\vee s_k, t+\varepsilon^*]$.

Case (b) Let us denote by t_1 the first time at which $\beta_{\ell^*}(t) = 1$. We introduce:

$$\widehat{\theta}^K_* = \inf\big\{t \ge \sigma^K_* \ :\ N^K_{\ell^*}(t\log K) \ge m\varepsilon_* K\big\}.$$

On the time interval $[\sigma^K_*, \widehat{\theta}^K_* \wedge \widetilde{\theta}^K_* \wedge t_1 \wedge T]$, we proceed as in the case (a) to deduce that (4.18) holds with high probability on this time interval. As in case (a), the exponents β_ℓ are defined with $\widehat{S}$ for all $s \in [s_k, t_1]$. Extending β_ℓ on $[s_k, t_1]$ like this is consistent with (2.4) where $\widetilde{S}_{s,k} = \widehat{S}$ (see (2.5)).

Proceeding as in Section 4.3.2, we can prove that, with high probability, $\widehat{\theta}^K_* < \widetilde{\theta}^K_*$ and

$$\lim_{K\to+\infty}\widehat{\theta}^K_* = t_1.$$

Using Lemma C.1(ii), there exists $T(\varepsilon_*)$ such that with high probability

$$\frac{N^K_{\ell^*}(\widehat{\theta}^K_*\log K + T(\varepsilon_*))}{K} \in \big[\frac{3-\ell^*\delta}{C} - 3\varepsilon_*, \frac{3-\ell^*\delta}{C} + 3\varepsilon_*\big].$$

By Lemma B.9, at the time $\widehat{\theta}_*^K \log K + T(\varepsilon_*)$, we have for all $\ell \neq \ell^*$,

$$\frac{\log(1 + N_\ell^L(\widehat{\theta}_*^K \log K + T(\varepsilon_*)))}{\log K} \in \big[\beta_\ell(t_1) - \varepsilon_*, \beta_\ell(t_1) + \varepsilon_*\big].$$

Applying the Markov property at this time $\widehat{\theta}_*^K \log K + T(\varepsilon_*)$ and proceeding as in Phase $k+1$, case 1, we obtain (4.18) where β_ℓ now evolves with the fitness S instead of $\widehat{S}$. This is consistent with (2.4) where $\widetilde{S}_{s,k} = S$ on $[t_1, s_{k+1}]$. This enlightens the introduction of $\widetilde{S}$ in (2.5). This case is the only one where $\widetilde{S}_{t,k}$ is not constant on the phase $[s_k, s_{k+1}]$.

4.3.4 Intermediate step $k+1$, case 2

In case (b), a new trait emerges in a resident population. This is treated in Section 4.3.2.
In case (a), either there is extinction of the population at time $s_{k+1} < +\infty$ (note that extinction can occur only in this case), or there is a change of dominant population at time $s_{k+1} < +\infty$.

Extinction In this case, either $s_{k+1} = T_0$, and we can conclude the induction as in Section 4.3.2, or $s_{k+1} < T_0$. This means that, for all $\ell \neq \ell^*$, $\beta_\ell(t_1 - \eta) = 0$ for some $\eta > 0$, so after time $t_1 - \eta$, $N_\ell^K(t) = 0$ with high probability. The population is then composed only of individuals with trait $\ell^*\delta$ and can be dominated by a subcritical branching process. Lemma A.1 proves the extinction of the population.

Emergence of a new dominant population The emergence occurs in a population of size $o(K)$. As in Section 4.3.2, we show that $\lim_{K\to+\infty} \widetilde{\theta}_*^K = s_{k+1}$ in probability and that at time $\widetilde{\theta}_*^K$, with high probability, for some $\kappa > 0$,

$$\sum_{\ell \notin \{\ell_{k+1}^*, \ell_{k+2}^*\}} N_\ell^K(\widetilde{\theta}_*^K \log K) \leq K^{-\kappa} N_{\ell_{k+1}^*}^K(\widetilde{\theta}_*^K \log K),$$

and $N_{\ell_{k+2}^*}^K(\theta_*^K \log K) \geq m\varepsilon_* N_{\ell_{k+1}^*}^K(\widetilde{\theta}_*^K \log K)/2$.

Depending on the sign of $S = \widehat{S}(\ell_{k+1}^*\delta, \ell_{k+2}^*\delta) - \widehat{S}(\ell_{k+2}^*\delta, \ell_{k+1}^*\delta)$, we can use Lemma C.2(i) or (ii). We can then define a time $\widetilde{\sigma}_{k+1}^K$ satisfying the recursion properties stated in Section 4.3.

This ends the Phase $k+1$ and its intermediate step and finishes the proof.

5 Proof of Theorem 2.5

Recall that we assume $\tau > \delta$. We first give a lemma describing the dynamics of the exponents before the first re-emergence of trait 0 or the first time when a trait in $(3, 4)$ becomes dominant.

Lemma 5.1 *Under the assumptions of Theorem 2.5 and with the definition* (2.10) *of* $\bar{\tau}$,

(a) *if* $m_0 > 0$, *we have*

$$s_k = \frac{k\alpha}{\tau - \delta} \quad \textit{for all} \quad k\delta \leq \bar{k}\delta \wedge 3.$$

For all $s \le s_{\lceil 3/\delta \rceil} \wedge \bar{\tau}$, *let* $k\delta < 3$ *be such that* $s \in [s_k, s_{k+1}]$, *then, if* $k \le \bar{k} - 1$, *we have*

$$\beta_\ell(s) = \begin{cases} 1 & \text{if } \ell = k, \\ [1 - (\ell - k)\alpha + (\tau - \delta)(s - s_k)] \vee 0 & \text{if } k < \ell \le L, \\ 1 - \frac{\alpha(k-\ell-1)}{\tau-\delta}\left(\tau - \frac{k-\ell}{2}\delta\right) - (\tau - (k-\ell)\delta)(s - s_k) & \text{if } 0 \le \ell < k, \end{cases} \tag{5.1}$$

and if $k = \bar{k}$, *we have*

$$\beta_\ell(s) = \begin{cases} 1 & \text{if } \ell = \bar{k}, \\ \left[1 - (\ell - \bar{k})\alpha + (\tau - \delta)(s - s_{\bar{k}})\right] \vee 0 & \text{if } \bar{k} < \ell \le L, \\ \left[1 - \frac{\alpha(\bar{k}-\ell-1)}{\tau-\delta}\left(\tau - \frac{\bar{k}-\ell}{2}\delta\right) - (\tau - (\bar{k}-\ell)\delta)(s - s_{\bar{k}})\right] & \\ \quad \vee \left[1 - \ell\alpha - \frac{\alpha(\bar{k}-1)}{\tau-\delta}\left(\tau - \frac{\bar{k}}{2}\delta\right) - (\tau - \bar{k}\delta)(s - s_{\bar{k}})\right] & \text{if } 0 \le \ell < \bar{k}; \end{cases} \tag{5.2}$$

(b) *if* $m_0 < 0$, *we have*

$$s_k = \frac{k\alpha}{\tau - \delta} \quad \text{for all} \quad k\delta < 3.$$

For all $s \le s_{\lceil 3/\delta \rceil}$, *let* $k\delta < 3$ *be such that* $s \in [s_k, s_{k+1}]$, *then*

$$\beta_\ell(s) = \begin{cases} 1 & \text{if } \ell = k, \\ [1 - (\ell - k)\alpha + (\tau - \delta)(s - s_k)] \vee 0 & \text{if } k < \ell \le L, \\ \left[1 - \frac{\alpha(k-\ell-1)}{\tau-\delta}\left(\tau - \frac{k-\ell}{2}\delta\right) - (\tau - (k-\ell)\delta)(s - s_k)\right] \vee 0 & \text{if } 0 \vee (k - \widetilde{k} + 1) \le \ell < k, \\ 0 & \text{if } 0 \le \ell \le k - \widetilde{k}. \end{cases} \tag{5.3}$$

Proof We apply Theorem 2.1 and proceed by induction on $k \ge 0$. The respective fitnesses of the traits 0 to $L\delta$ at time 0 are $S(0;0) = 0$, $S(\ell\delta;0) = \tau - \ell\delta$, for all $1 \le \ell \le L$. We use formula (2.7) recursively from $\ell = 1$ to L to prove that, for $\ell \le \lfloor \frac{1}{\alpha} \rfloor$

$$\beta_\ell(t) = \big(1 - \ell\alpha + (\tau - \ell\delta)t\big) \vee \big(1 - \ell\alpha + (\tau - \delta)t\big) \vee 0 = 1 - \ell\alpha + (\tau - \delta)t,$$

and $t_{\ell,1} = 0$ when $\ell < \lfloor 1/\alpha \rfloor$. When $\ell = \lfloor \frac{1}{\alpha} \rfloor$, we have $\beta_{\lfloor 1/\alpha \rfloor}(0) \in (0, \alpha)$ and we deduce that $t_{\lfloor 1/\alpha \rfloor,1} = \frac{\alpha(1+\lfloor 1/\alpha \rfloor)-1}{\tau-\delta} > 0$.

For $\ell = \lfloor \frac{1}{\alpha} \rfloor + 1$, we use (2.4) to prove

$$\beta_{\lfloor 1/\alpha \rfloor+1}(t) = \Big((\tau - (\lfloor \frac{1}{\alpha} \rfloor + 1)\delta)(t - t_{\lfloor 1/\alpha \rfloor,1})\Big) \vee \Big(1 - (\lfloor \frac{1}{\alpha} \rfloor + 1)\alpha + (\tau - \delta)t\Big) \vee 0 = \Big(1 - (\lfloor \frac{1}{\alpha} \rfloor + 1)\alpha + (\tau - \delta)t\Big)_+.$$

Similar computation gives the general formula for all $\ell \in \{1, \dots L\}$:

$$\beta_\ell(t) = \big(1 - \ell\alpha + (\tau - \delta)t\big)_+ \,; \qquad t_{\ell,1} = \begin{cases} 0 & \text{if } \ell < \lfloor 1/\alpha \rfloor \\ \frac{\alpha(\ell+1)-1}{\tau-\delta} & \text{otherwise.} \end{cases} \tag{5.4}$$

We see that $\beta_1(t)$ is the first exponent to reach 1, so $s_1 = \frac{\alpha}{\tau-\delta}$. This completes the initialization of the induction in both cases (a) and (b).

Proof of (a) Assume that $m_0 > 0$ and that we proved (5.1) until time s_k for some $k \in \{0, 1, \dots, (\lceil 3/\delta \rceil - 1) \wedge \bar{k}\}$, and let us prove that (5.1) is valid until time $s_{k+1} \wedge \bar{\tau}$ if $k < \bar{k}$, or that (5.2) is valid until time $\bar{\tau} < s_{\bar{k}+1}$ if $k = \bar{k}$. At time s_k, the new resident trait $k\delta < 3$ replaces the former resident trait $(k-1)\delta$ and hence, the values of the fitnesses are given by

$$S(\ell\delta; k\delta) = \tau - (\ell - k)\delta \text{ if } \ell > k,\ S(k\delta; k\delta) = 0,\ S(\ell\delta; k\delta) = (k - \ell)\delta - \tau \text{ if } \ell < k. \tag{5.5}$$

Since $S(\ell\delta; k\delta) < \tau - \delta = S((k+1)\delta; k\delta)$ for all $\ell > k+1$, applying Theorem 2.1, we obtain for all $s \geq s_k$,

$$\beta_\ell(s) = [1 - (\ell - k)\alpha + (\tau - \delta)(s - s_k)] \vee 0, \quad \forall \ell > k,$$

until the next change of resident population. Since obviously $\beta_k(s) = 1$ until this time, we have proved the first two lines in (5.1) (or of (5.2) if $k = \bar{k}$). For the last lines, we obtain for all $s \geq s_k$ and all $\ell < k$

$$\begin{aligned}\beta_\ell(s) = [\beta_\ell(s_k) - [\tau - (k-\ell)\delta](s - s_k)] \vee [\beta_{\ell-1}(s_k) - \alpha - [\tau - (k - \ell + 1)\delta](s - s_k)] \vee \dots \\ \vee [\beta_0(s_k) - \ell\alpha - [\tau - k\delta](s - s_k)] \vee 0,\end{aligned} \tag{5.6}$$

until the next change of resident trait. Now, all the terms in the right-hand side except the last one lie between two consecutive terms of the sequence $\left(1 - \frac{\alpha(n-1)}{\tau-\delta}\left(\tau - \frac{n}{2}\delta\right)\right)_{n\geq 0}$, so they are all positive since we assumed $m_0 > 0$. Hence

$$\begin{aligned}\beta_\ell(s) = [\beta_\ell(s_k) - [\tau - (k-\ell)\delta](s - s_k)] \vee [\beta_{\ell-1}(s_k) - \alpha - [\tau - (k - \ell + 1)\delta](s - s_k)] \vee \dots \\ \vee [\beta_0(s_k) - \ell\alpha - [\tau - k\delta](s - s_k)]\,.\end{aligned} \tag{5.7}$$

In addition, for all $1 \leq \ell < k$, the function

$$s \in [s_k, s_{k+1} \wedge \bar{\tau}] \mapsto \beta_\ell(s_k) - [\tau - (k-\ell)\delta](s - s_k) - \beta_{\ell-1}(s_k) + [\tau - (k-\ell+1)\delta](s - s_k) + \alpha \tag{5.8}$$

is affine and hence takes values between its values at times s_k and s_{k+1}, i.e. between

$$\frac{\alpha}{\tau - \delta}\left(2\tau - (k - \ell + 1)\delta\right) \quad \text{and} \quad \frac{\alpha}{\tau - \delta}\left(2\tau - (k - \ell + 2)\delta\right).$$

In the case where $k < \bar{k}$, we have $1 \leq k - \ell \leq k - 1 \leq \bar{k} - 2 = \lfloor 2\frac{\tau}{\delta} \rfloor - 2$, so both terms above are nonnegative, and the function (5.8) is positive for all $s - s_k \in [0, \frac{\alpha}{\tau-\delta}]$. This means that the maximum of two consecutive terms in the right-hand side of (5.7) is always the first term, therefore

$$\beta_\ell(s) = \beta_\ell(s_k) - [\tau - (k-\ell)\delta](s - s_k), \quad \forall 0 \leq \ell < k,$$

until the next change of resident population. Since $k - \ell + 1 \leq \bar{k}$ for all $0 \leq \ell < k$, we have

$$\beta_\ell(s_k) - [\tau - (k-\ell)\delta]\frac{\alpha}{\tau - \delta} = 1 - \frac{\alpha(k-\ell)}{\tau - \delta}\left(\tau - \frac{k - \ell + 1}{2}\delta\right) < 1,$$

so the first exponent β_ℓ for $\ell \neq k$ reaching 1 after time s_k is $\ell = k+1$ and the next change of resident population occurs at time s_{k+1}. Therefore, we have proved (5.1).

Let us now consider the case where $k = \bar{k}$. The previous argument shows that the maximum of two consecutive terms in the right-hand side of (5.7) is always the first term, except possibly for the last two terms. Therefore, in all cases,

$$\beta_\ell(s) = \left[\beta_\ell(s_{\bar{k}}) - [\tau - (\bar{k} - \ell)\delta](s - s_{\bar{k}})\right] \vee \left[\beta_0(s_{\bar{k}}) - \ell\alpha - [\tau - \bar{k}\delta](s - s_{\bar{k}})\right], \quad \forall 0 \leq \ell < \bar{k},$$

until the next change of resident population. In this case, the first exponent to reach 1 is β_0, at time $\bar{\tau}$, since $\bar{\tau} - s_{\bar{k}} < \frac{\alpha}{\tau-\delta}$. This concludes the proof of (5.2).

Proof of (b) Assume now that $m_0 < 0$. This means that the sequence $\left(1 - \frac{\alpha(n-1)}{\tau-\delta}\left(\tau - \frac{n}{2}\delta\right)\right)_{n\geq 0}$ becomes negative at some index $k^{**} \leq \widetilde{k}$ and decreases until index $n = \widetilde{k}$.

Assume that we proved (5.3) until time s_k for some $k \in \{0, 1, \ldots, \lceil 3/\delta \rceil - 1\}$, and let us prove that (5.3) is valid until time s_{k+1}. The fitnesses are the same as in (5.5), and the first computations of case (a) apply similarly to prove the first two lines of (5.3) until the next change of resident population.

In order to prove the last two lines of (5.3), we need to modify (5.6) accordingly. To this aim, let us observe that, when the exponent $\beta_\ell(s_k)$ of trait $\ell\delta$ is 0, it cannot increase (even if its fitness is positive) unless some mutant individuals gets born from trait $(\ell-1)\delta$, which only occurs at times s such that $\beta_{\ell-1}(s) \geq \alpha$. Therefore, (5.6) needs to be modified as $\beta_\ell(s) = 0$ if $0 \leq \ell \leq k - k^{**}$, for all $s \geq s_k$ until the next change of resident trait, or (5.6) remains true for $0 \vee (k - k^{**} + 1) \leq \ell \leq k - 2$. As in case (a), we can prove that the maximum between two successive terms in the previous expression (except the last one, 0) is always reached by the first one, hence

$$\beta_\ell(s) = [\beta_\ell(s_k) - [\tau - (k-\ell)\delta](s - s_k)] \vee 0.$$

In view of the expression for $\beta_\ell(s)$, the next change of resident population occurs when $\beta_{k+1}(s)$ hits 1, at time $s_{k+1} = s_k + \frac{\alpha}{\tau-\delta}$. This ends the proof of (5.3). □

Proof [Proof of Theorem 2.5] Let us first prove (a). We assume $m_0 > 0$ and $\bar{k}\delta < 3$. Then, it follows from Lemma 5.1 (a) that there is re-emergence of trait 0 at time $\bar{\tau}$.

Let us now consider the case (b). Assume $m_0 < 0$. We introduce the first index such that the trait becomes larger than 3 by $\widehat{k} = \lceil \frac{3}{\delta} \rceil$.

Note that the integer k^{**} defined in the last proof satisfies $k^{**} \leq \widetilde{k} \leq \widehat{k}$. Hence, Lemma 5.1 (b) implies that traits $0, \delta, \cdots, (\widehat{k} - k^{**})\delta$, get lost before their re-emergence and before a trait larger than 3 becomes dominant. Note that they remain extinct forever since mutations only produce individuals with larger traits. Our goal is to prove that, after time $s_{\widehat{k}}$, the other traits get progressively lost until the global extinction of the population.

At time $s_{\widehat{k}}$, the dominant trait becomes $\widehat{k}\delta > 3$ and the fitnesses are given by

$$\widehat{S}(\ell\delta; \widehat{k}\delta) = 3 - \ell\delta - \tau \text{ if } \ell < \widehat{k}; \quad \widehat{S}(\widehat{k}\delta; \widehat{k}\delta) = 3 - \widehat{k}\delta; \quad \widehat{S}(\ell\delta; \widehat{k}\delta) = 3 - \ell\delta + \tau \text{ if } \ell > \widehat{k}. \tag{5.9}$$

Now, for $\widehat{k} - k^{**} < \ell < \widehat{k}$, $\widehat{S}(\ell\delta; \widehat{k}\delta) < S(\ell\delta; \widehat{k}\delta) < 0$ (recall from Lemma 5.1 (b) that $k^{**} \leq \widetilde{k}$). Therefore, all positive exponents $\beta_\ell(s)$ for $\ell < \widehat{k}$ keep on decreasing until they hit 0 (which means that trait $\ell\delta$ is lost) or until the next change of dominant population. For all $\ell > \widehat{k}$, $\beta_\ell(s_{\widehat{k}}) = [1 - (\ell - \widehat{k})\delta]_+$ and β_ℓ has slope $\widehat{S}((\widehat{k}+1)\delta; \widehat{k}\delta) = \widehat{S}(\widehat{k}\delta; \widehat{k}\delta) + \tau - \delta$, therefore the next change of dominant population occurs when $\beta_{\widehat{k}+1}(s) = \beta_{\widehat{k}}(s)$ at time $s_{\widehat{k}+1}$. At this time, by definition of k^{**},

$$\beta_{\widehat{k}-k^{**}+1}(s_{\widehat{k}+1}) \leq \left[\beta_{\widehat{k}-k^{**}+1}(s_{\widehat{k}}) + ((k^{**}-1)\delta - \tau)\frac{\alpha}{\tau-\delta}\right] \vee 0 = 0,$$

so trait $(\widehat{k} - k^{**} + 1)\delta$ is lost. For all $\widehat{k} - \ell^* + 1 < \ell < \widehat{k} + 1$,

$$\begin{aligned}\beta_\ell(s_{\widehat{k}+1}) \leq & \left[\beta_\ell(s_{\widehat{k}}) + (3 - \ell\delta - \tau)\frac{\alpha}{\tau-\delta}\right] \vee \ldots \\ & \vee \left[\beta_{\widehat{k}-k^{**}+2}(s_{\widehat{k}}) - (\ell + k^{**} - \widehat{k} - 2)\alpha + (3 - (\widehat{k} - k^{**} + 2)\delta - \tau)\frac{\alpha}{\tau-\delta}\right] \vee 0\end{aligned}$$

$$
\begin{aligned}
&\leq \left[\beta_\ell(s_{\widehat{k}}) - (\tau - (\widehat{k} - \ell)\delta)\frac{\alpha}{\tau - \delta}\right] \vee \dots \\
&\quad \vee \left[\beta_{\widehat{k}-k^{**}+2}(s_{\widehat{k}}) - (\ell + \ell^* - \widehat{k} - 2)\alpha - (\tau - (k^{**} - 2)\delta)\frac{\alpha}{\tau - \delta}\right] \vee 0 \\
&= \left[\beta_\ell(s_{\widehat{k}}) - (\tau - (\widehat{k} - \ell)\delta)\frac{\alpha}{\tau - \delta}\right] \vee 0,
\end{aligned}
$$

where the last equality was proved in the proof of Lemma 5.1 (b). Hence, we can proceed inductively to prove that, for all $1 \leq k \leq L - \widehat{k}$, at time $s_{\widehat{k}+k} = (\widehat{k}+k)\frac{\alpha}{\tau-\delta}$, $\max_{0\leq\ell\leq L} \beta_\ell(s_{\widehat{k}+k}) = \beta_{\widehat{k}+k}(s_{\widehat{k}+k}) > 0$, $\beta_\ell(s_{\widehat{k}+k}) = 0$ for all $\ell \leq \widehat{k}+k-k^{**}$, all exponents $\beta_\ell(s)$ with $\widehat{k}+k-k^{**} < \ell \leq \widehat{k}+k$ decrease. In addition,

- either $\beta_{\widehat{k}+k}(s_{\widehat{k}+k}) > \alpha$ and $\widehat{k}+k < L$, and then $\beta_{\widehat{k}+k+\ell}(s_{\widehat{k}+k}) = [\beta_{\widehat{k}+k}(s_{\widehat{k}+k}) - \ell\alpha] \vee 0$ for all $1 \leq \ell \leq L - \widehat{k} - k$, the next time of change of dominant population is $s_{\widehat{k}+k+1} = (\widehat{k}+k+1)\frac{\alpha}{\tau-\delta}$;
- or $\beta_{\widehat{k}+k}(s_{\widehat{k}+k}) < \alpha$ or $\widehat{k}+k = L$, and then $\beta_{\widehat{k}+k+\ell}(s_{\widehat{k}+k}) = 0$ for all $\ell \geq 1$, so every exponent keeps on decreasing until $\beta_{\widehat{k}+k}(s) = 0$ and there is extinction of the population.

The second case will necessarily occur after a finite number of steps, so extinction of the population occurs before the re-emergence of any trait. This ends the proof of (b).

To prove (c), we assume $m_0 > 0$ and $\bar{k} \geq \widehat{k}$. By Lemma 5.1, in both cases, $\beta_0(s_{\widehat{k}}) \in (0, 1)$ and trait $\widehat{k}\delta$ becomes dominant at time $s_{\widehat{k}}$. Since $\widehat{k}\delta > 3$, the dynamics of the exponents β_ℓ after time $s_{\widehat{k}}$ is governed by the fitnesses given in (5.9).

Since $\widehat{S}(\widehat{k}\delta; \widehat{k}\delta) < 0$, $\max_{0\leq\ell\leq L} \beta_\ell(s) < 1$ for all $s > s_{\widehat{k}}$ until the next change of dominant trait. Observe that the fitness of trait 0, $\widehat{S}(0; \widehat{k}\delta) = 3 - \tau$ is positive as long as $\widehat{k}\delta$ is the dominant trait, and for any other dominant trait $\ell\delta$, we have

$$
\widehat{S}(0; \ell\delta) = \begin{cases} 3 - \tau & \text{if } \ell \geq 1, \\ 3 & \text{if } \ell = 0. \end{cases}
$$

So, as long as $\max_{0\leq\ell\leq L} \beta_\ell(s) < 1$, $\beta_0(s)$ is increasing. This implies that there is necessarily re-emergence of some trait before the extinction of the population. Since in addition $\widehat{S}(\ell\delta; \ell\delta) = 3 - \ell\delta < 0$ for all $\ell \geq \widehat{k}$, the first re-emerging trait cannot be any of the traits $\ell\delta$ for $\ell \geq \widehat{k}$. This ends the proof of (c). □

A Branching process in continuous time

The goal of this section and the next two ones is to give general results on specific branching or birth-and-death one and two-dimensional processes.

Here, we consider a single population $(Z_t^K, t \geq 0)$, following a linear birth and death process, i.e. a branching process, with individual birth rate $b \geq 0$, individual death rate $d \geq 0$ and initial value $Z_0^K = \lfloor K^\beta - 1 \rfloor \in \mathbb{N}$. The definition of Z_0^K means that the population is extinct initially when $\beta = 0$, and that $Z_0^K \sim K^\beta$ when $K \to +\infty$ if $\beta > 0$. We denote the law of Z^K by $BP_K(b, d, \beta)$. In the sequel, we set $r = b - d$.

Lemma A.1 *Let Z^K follow the law $BP_K(b,d,\beta)$, where $b,d \geq 0$ and $\beta > 0$. Then the process $\left(\frac{\log(1+Z^K_{s\log K})}{\log K}, s \geq 0\right)$ converges in probability in $L^\infty([0,T])$ for all $T > 0$ to $((\beta + rs) \vee 0, s \geq 0)$ when K tends to infinity.*

In addition, if $b < d$, for all $t > \beta/r$,

$$\lim_{K\to+\infty} \mathbb{P}\left(Z^K_{t\log K} = 0\right) = 1. \tag{A.1}$$

Proof We divide the proof in several steps. We borrow ideas from [8].

Step 1. Construction of a martingale. The process writes

$$Z^K_t = K^\beta + M^K_t + \int_0^t r\, Z^K_s ds,$$

where M^K is a square integrable martingale with quadratic variation $\langle M^K\rangle_t = \int_0^t (b+d) Z^K_s ds$. Taking the expectation, we obtain $\mathbb{E}(Z^K_t) = K^\beta + \int_0^t r\, \mathbb{E}(Z^K_s) ds$, which gives that

$$\mathbb{E}(Z^K_t) = K^\beta e^{rt}. \tag{A.2}$$

Using Itô's formula, $1 + e^{-rt} Z^K_t = 1 + K^\beta + \widehat{M}^K_t$, where $\widehat{M}^K_t = \int_0^t e^{-rs} dM^K_s$ is a square integrable martingale with quadratic variation process

$$\langle \widehat{M}^K\rangle_t = \int_0^t e^{-2rs}(b+d) Z^K_s ds.$$

Using Doob inequality, for any $0 < \eta < \beta$,

$$\begin{aligned}
\mathbb{P}\big(\sup_{t\leq T\log K} \left|e^{-rt} Z^K_t - K^\beta\right| \geq K^\eta\big) &= \mathbb{P}\big(\sup_{t\leq T\log K} |\widehat{M}^K_t| \geq K^\eta\big) \\
&\leq 4K^{-2\eta}\mathbb{E}\big(\langle M^K\rangle_{T\log K}\big) \\
&= 4K^{-2\eta}\int_0^{T\log K} (b+d) e^{-2rs} K^\beta e^{rs} ds \\
&= 4K^{\beta-2\eta}(b+d)\begin{cases} \frac{1}{r}\left(1 - K^{-rT}\right) & \text{if } b \neq d, \\ T\log K & \text{if } b = d. \end{cases}
\end{aligned} \tag{A.3}$$

Step 2. Case $r > 0$. Fix $T > 0$ and $\eta = 2\beta/3$. On the set

$$\Omega^K_1 = \big\{\sup_{t\leq T\log K} \left|e^{-rt} Z^K_t - K^\beta\right| \leq K^{\frac{2\beta}{3}}\big\},$$

whose probability tends to 1, we have:

$$\begin{aligned}
&\sup_{t\leq T}\left|\frac{\log(1+Z^K_{t\log K})}{\log K} - \big(\beta + rt\big)\right| \\
&\qquad = \sup_{t\leq T}\left|\frac{\log(1+Z^K_{t\log K})}{\log K} - \frac{\log(1+K^{\beta+rt})}{\log K} + \frac{\log(1+K^{\beta+rt})}{\log K} - \frac{\log(K^{\beta+rt})}{\log K}\right|
\end{aligned}$$

$$
\begin{aligned}
&\leq \sup_{t\leq T}\left[\frac{1}{\log K}\left|\log\left(\frac{1+Z^K_{t\log K}}{1+K^{\beta+rt}}\right)\right|+\frac{\log\left(1+K^{-\beta-rt}\right)}{\log K}\right]\\
&\leq \sup_{t\leq T}\left[\frac{1}{\log K}\left|\log\left(\frac{1+Z^K_{t\log K}\vee K^{\beta+rt}}{1+Z^K_{t\log K}\wedge K^{\beta+rt}}\right)\right|+\frac{K^{-\beta-rt}}{\log K}\right]\\
&= \sup_{t\leq T}\left[\frac{1}{\log K}\log\left(1+\frac{\left|Z^K_{t\log K}-K^{\beta+rt}\right|}{1+Z^K_{t\log K}\wedge K^{\beta+rt}}\right)+\frac{K^{-\beta-rt}}{\log K}\right],\\
&\leq \sup_{t\leq T}\left[\frac{1}{\log K}\frac{|Z^K_{t\log K}-K^{\beta+rt}|}{1+K^{\beta+rt}-K^{\frac{2\beta}{3}+rt}}+\frac{K^{-\beta-rt}}{\log K}\right]\\
&\leq \frac{K^{\frac{2\beta}{3}}}{\log K}\sup_{t\leq T}\frac{K^{rt}}{K^{\beta+rt}-K^{\frac{2\beta}{3}+rt}}+\frac{K^{-\beta}}{\log K}\\
&\leq \frac{2K^{-\frac{\beta}{3}}+K^{-\beta}}{\log K}\quad\text{(for } K \text{ large enough)}\ ,
\end{aligned}
$$

which converges to zero when $K\to+\infty$.

A simple adaptation of the previous argument gives the same conclusion for $b=d$.

Step 3. Case $r<0$. Since the function $t\mapsto\beta+rt$ vanishes at time $\beta/|r|$, we need to consider three phases. We fix $\varepsilon\in(0,\beta)$ and set $T_\varepsilon=\frac{\beta-\varepsilon}{|r|}$ and $\eta=\beta-\varepsilon/3$. First we prove that, before time $T_\varepsilon\log K$ the population size remains large enough to use the same argument as in Step 2. Second, the population gets extinct with high probability between time $T_\varepsilon\log K$ and $(T_\varepsilon+2\varepsilon/|r|)\log K$ and third, in order to obtain the convergence for the L^∞ norm, we prove that the supremum of the process on the this time interval remains of the order of ε. Since ε can be chosen arbitrarily small, the result follows.

Step 3(i). On the set

$$
\Omega_2^K=\big\{\sup_{t\leq T_\varepsilon\log K}\left|e^{-rt}Z^K_t-K^\beta\right|\leq K^{\beta-\frac{\varepsilon}{3}}\big\},
$$

whose probability tend to 1, we have:

$$
\begin{aligned}
\sup_{t\leq T_\varepsilon}\left|\frac{\log(1+Z^K_{t\log K})}{\log K}-\big(\beta+rt\big)\right| &\leq \sup_{t\leq T_\varepsilon}\left[\frac{1}{\log K}\frac{|Z^K_{t\log K}-K^{\beta+rt}|}{1+\left(K^{\beta+rt}-K^{\frac{2\beta}{3}+rt}\right)_+}+\frac{K^{-\beta-rt}}{\log K}\right]\\
&\leq \frac{K^{\beta-\frac{\varepsilon}{3}}}{\log K}\sup_{t\leq T_\varepsilon}\frac{K^{rt}}{1+\left(K^{\beta+rt}-K^{\frac{2\beta}{3}+rt}\right)_+}+\frac{K^{-\varepsilon}}{\log K}\\
&\leq \frac{2K^{-\frac{\varepsilon}{3}}+K^{-\varepsilon}}{\log K}\quad\text{(for } K \text{ large enough)}\ ,
\end{aligned}
$$

which also converges to zero when $K\to+\infty$.

Step 3(ii). It follows from the last step that, with probability converging to 1, $Z^K_{T_\varepsilon \log K} \leq 2K^\varepsilon$. Remind from [14, Section 5.4.5, p. 180] that, if T_{ext} denotes the extinction time of a $BP(b,d,1)$,

$$\mathbb{P}(T_{\text{ext}} > t) = \frac{re^{rt}}{be^{rt} - d}. \tag{A.4}$$

Hence, for a $BP(b,d,2K^\varepsilon)$ branching process,

$$\mathbb{P}(T_{\text{ext}} > t) = 1 - \left(1 - \frac{re^{rt}}{be^{rt} - d}\right)^{2K^\varepsilon}.$$

Thus, for $t = \frac{2\varepsilon}{|r|}\log K$,

$$\mathbb{P}\left(T_{\text{ext}} > \frac{2\varepsilon}{|r|}\log K\right) \sim 2\,\frac{|r|}{d}\,K^{-\varepsilon}$$

as $K \to +\infty$. Since this goes to 0 when $K \to +\infty$, we have completed the proof of (A.1).

Step 3(iii). Since the last two steps were true for any value of $\varepsilon > 0$, in order to complete the proof, it is enough to check that

$$\sup_{s \in [T_\varepsilon, T_\varepsilon + 2\varepsilon/|r|]} \frac{\log(1 + Z^K_{s\log K})}{\log K} \leq \frac{2d}{|r|}\,\varepsilon.$$

For this, we observe that the maximal size of the families stemming from each individual alive at time $T_\varepsilon \log K$ on the time interval $\log K \times [T_\varepsilon, T_\varepsilon + 2\varepsilon/|r|]$ is bounded independently by the size at time $\frac{2\varepsilon \log K}{|r|}$ of a Yule process with birth rate b, i.e. a geometric random variable G_i with expectation $K^{2b\varepsilon/|r|}$. Hence, with probability converging to 1,

$$\sup_{t \in [T_\varepsilon, T_\varepsilon + 2\varepsilon/|r|]} Z^K_{t\log K} \leq \sum_{i=1}^{2K^\varepsilon} G_i \leq K^{\frac{2d}{|r|}\varepsilon}.$$

The proof is completed. □

B Branching process with immigration

Our goal in this section is to extend the arguments of the previous section to include immigration.

We consider a linear birth and death process with immigration $(Z^K_t, t \geq 0)$ with law $BPI_K(b,d,a,c,\beta)$, where $Z^K_0 = \lfloor K^\beta - 1 \rfloor$ with $\beta \geq 0$, $b \geq 0$ is the individual birth rate, $d \geq 0$ the individual death rate and $K^c e^{as}$ is the immigration rate at time $s \geq 0$ with $a, c \in \mathbb{R}$.

With these notations, the generator of Z^K at time $t \geq 0$ is given for bounded measurable functions f by

$$L_t f(n) = (bn + K^c e^{at})\,(f(n+1) - f(n)) + dn\,(f(n-1) - f(n)).$$

Recall that $r = b - d$. We start with a first result in the case where the population does not go extinct (Theorem B.1) and a series of lemmas. Our general result on branching processes with immigration is Theorem B.5, whose proof combines similar ideas as for Lemma A.1.

Theorem B.1 (Large population) *Assume that $c \leq \beta$, $\beta > 0$. Then, for all $T > 0$ such that*

$$\inf_{t\in[0,T]} (\beta + rt) \vee (c + at) > 0, \tag{B.1}$$

the process $\left(\frac{\log(1+Z^K_{s\log K})}{\log K}, 0 \leq s \leq T\right)$ *converges when K tends to infinity in probability in $L^\infty([0,T])$ to* $\Big((\beta + rs) \vee (c + as), s \in [0,T]\Big)$.

In the proofs of the main results of [8] and [3], similar asymptotic results on branching processes with immigration were proved. Our framework is more general since we consider cases where immigration is time-dependent ($a \neq 0$, contrary to [3]) and where growth can be driven by immigration ($a > r$, contrary to [8]). The results of these references are based on Doob's inequality. In our general case, this is not sufficient: we also need to make use of the maximal inequality for supermartingales [7, Ch. VI, p. 72] in the case $a > r$.

To motivate this result, let us first compute the expectation and variance of a $BPI_K(b, d, a, c, \beta)$ process. Note that this result is valid also if $c > \beta$.

Lemma B.2 *Assume that $(Z^K_t, t \geq 0)$ is a $BPI_K(b, d, a, c, \beta)$ process. Then, for all $t \geq 0$,*

$$x^K_t = \mathbb{E}(Z^K_t) = \begin{cases} \left(K^\beta - 1 + \frac{K^c}{r-a}\right) e^{rt} - \frac{K^c e^{at}}{r-a} & \text{if } r \neq a, \\ e^{rt}(K^\beta - 1 + K^c t) & \text{if } r = a \end{cases} \tag{B.2}$$

and

$$\mathrm{Var}(Z^K_t) = \begin{cases} (b+d)\left(K^\beta - 1 + \frac{K^c}{r-a}\right)\frac{e^{2rt}-e^{rt}}{r} + K^c\left(1 - \frac{b+d}{r-a}\right)\frac{e^{at}-e^{2rt}}{a-2r} & \text{if } r \neq a, \\ (b+d)(K^\beta - 1)\frac{e^{2rt}-e^{rt}}{r} + K^c\frac{e^{2rt}-e^{rt}}{r} + (b+d)K^c\frac{e^{2rt}-e^{rt}-rte^{rt}}{r^2} & \text{if } r = a \neq 0, \\ (K^c + 2b(K^\beta - 1))t + bK^c t^2 & \text{if } r = a = 0 \quad (b = d), \end{cases} \tag{B.3}$$

where, in the first line, by convention $\frac{e^{2rt}-e^{rt}}{r} = t$ if $r = 0$ and $\frac{e^{at}-e^{2rt}}{a-2r} = t$ if $a = 2r$.

Proof The semimartingale decomposition of Z^K is given by

$$Z^K_t = Z^K_0 + \int_0^t \left(rZ^K_s + K^c e^{as}\right)ds + M^K_t, \tag{B.4}$$

where M^K is a square integrable martingale with quadratic variation

$$\langle M^K\rangle_t = \int_0^t \left((b+d)Z^K_s + K^c e^{as}\right)ds. \tag{B.5}$$

By a usual argument, the expectation $x(t) = \mathbb{E}(Z^K_t)$ solves the linear equation

$$\dot{x}(t) = rx(t) + K^c e^{at} \ ; \ x(0) = K^\beta - 1$$

whose solution is given by (B.2).

Itô's formula applied to $(Z^K_t)^2$ yields that $u(t) = \mathbb{E}((Z^K_t)^2)$ is solution to

$$\dot{u}(t) = 2ru(t) + (2K^c e^{at} + b + d)\, x(t) + K^c e^{at} \ ; \ u(0) = (K^\beta - 1)^2.$$

Straightforward computation in each case gives (B.3). □

We deduce from the last lemma that

$$x^K_{t\log K} = \begin{cases} K^{rt}(K^\beta - 1) + \frac{K^{c+rt} - K^{c+at}}{r-a} & \text{if } r \neq a, \\ K^{rt}(K^\beta - 1) + K^{c+rt} t \log K & \text{if } r = a, \end{cases}$$

so that,

$$\frac{\log(1 + x^K_{t\log K})}{\log K} \sim \begin{cases} (\beta + rt) \vee (c + rt) \vee (c + at) & \text{if } \beta > 0 \\ (c + rt) \vee (c + at) & \text{if } \beta = 0 \end{cases}$$

as $K \to +\infty$. Note that the expression (B.3) in the case $r \neq a$ can be written as

$$\mathrm{Var}(Z^K_{t\log K}) = (b+d)(K^\beta - 1)\varphi_t(r) + K^c \varphi_t(a) + (b+d)K^c \frac{\varphi_t(r) - \varphi_t(a)}{r-a} \tag{B.6}$$

where

$$\varphi_t(x) = \frac{K^{xt} - K^{2rt}}{x - 2r}.$$

Since φ_t is nonnegative and nondecreasing for all $t \geq 0$, the three terms in the right-hand side of (B.6) are positive, so there is no compensation between these terms and hence

$$\mathrm{Var}(Z^K_{t\log K}) \sim \begin{cases} CK^{\beta + (2r \vee r)t} \vee K^{c + (2r \vee r \vee a)t} & \text{if } \beta > 0, \\ CK^{c + (2r \vee r \vee a)t} & \text{if } \beta = 0. \end{cases}$$

for a constant C that may depend on t but remains uniformly bounded and bounded away from 0 for bounded values of $t \in \mathbb{R}_+$.

In several cases, we can compare the process with its expectation because the standard deviation $\mathrm{Var}(Z^K_{t\log K})^{1/2}$ is negligible compared with $x^K_{t\log K}$. One can easily check that this is always the case when $t > 0$ and $\beta > 0$ and $r \geq 0$ or when $\beta \geq 0$, $r \leq 0$ and $(\beta + rt) \vee (c + (a \vee r)t) > 0$ or when $\beta = 0$, $r > 0$ and $c > 0$.

In particular, Theorem B.1 corresponds to cases 1) and part of 2). Its proof is based on martingales inequalities. We start by defining a martingale and computing its quadratic variation.

Lemma B.3 *The process*

$$\widetilde{M}^K_t := e^{-rt}\big(Z^K_t - x^K_t\big) \tag{B.7}$$

is a martingale whose predictable quadratic variation process satisfies

$$\begin{aligned} \mathbb{E}\big(\langle \widetilde{M}^K \rangle_{T\log K}\big) &= K^c \frac{K^{(a-2r)T} - 1}{a - 2r} + (b+d)(K^\beta - 1)\frac{1 - K^{-rT}}{r} \\ &\quad + (b+d) \times \begin{cases} \frac{K^c}{a-r}\left(\frac{K^{(a-2r)T} - 1}{a-2r} - \frac{1 - K^{-rT}}{r}\right) & \textit{if } \ r \neq a \\ K^c \frac{1 - K^{-rT}}{r^2} - K^c \frac{T\log(K) K^{-rT}}{r} & \textit{if } \ r = a. \end{cases} \end{aligned}$$

Proof Recall the semi-martingale decomposition of $(Z^K_t)_{t \in \mathbb{R}_+}$ in (B.4) where the bracket of the martingale part is given in (B.5). It follows that $\widetilde{M}^K_t$ is a martingale with predictable quadratic variation process

$$\langle \widetilde{M}^K \rangle_t = \int_0^t e^{-2rs}\big((b+d)Z^K_s + K^c e^{as}\big)ds. \tag{B.8}$$

Using (B.2), we have that

$$\int_0^{T\log K} e^{-2rs}x_s^K\ ds = \begin{cases} (K^\beta-1)\frac{1-K^{-rT}}{r} + \frac{K^c}{a-r}\left(\frac{K^{(a-2r)T}-1}{a-2r} - \frac{1-K^{-rT}}{r}\right) & \text{if } r\neq a \\ (K^\beta-1)\frac{1-K^{-rT}}{r} + K^c\frac{1-K^{-rT}}{r^2} - K^c\frac{T\log(K)K^{-rT}}{r} & \text{if } r=a, \end{cases}$$

and Lemma B.3 follows. □

Proof of Theorem B.1

Let us denote $\bar\beta_t = (\beta+rt)\vee(c+at)$. The proof extends ideas used to prove Lemma A.1. We give the proof in the case where $r\neq a$, $r\neq 0$ and $a\neq 2r$. The extension to the other cases can be easily deduced using comparison arguments.

We shall use two different martingale inequalities applied to $\widetilde{M}_t^K$. The first one is Doob's inequality: for any $0<\eta<\beta$,

$$\begin{aligned}\mathbb{P}\big(\sup_{t\le T\log K}\big|e^{-rt}(Z_t^K-x_t^K)\big|\ge K^\eta\big) =& \mathbb{P}\big(\sup_{t\le T\log K}|\widetilde{M}_t^K|\ge K^\eta\big)\\ \le& 4K^{-2\eta}\mathbb{E}\big(\langle M^K\rangle_{T\log K}\big)\\ \le& CK^{-2\eta}\times\begin{cases}K^{\beta\vee(\beta-rT)\vee(c+(a-2r)T)} & \text{if } r\neq a\\ K^{\beta\vee(\beta-rT)}\log K & \text{if } r=a,\end{cases}\end{aligned} \tag{B.9}$$

where we used the fact that $c\le\beta$ in the last inequality. When $a>r$, we also apply the maximal inequality of [7, Ch. VI, p. 72] to the supermartingale $e^{-(a-r)t}\widetilde{M}_t^K$:

$$\begin{aligned}K^\eta\,\mathbb{P}\big(\sup_{t\le T\log K}e^{-at}|Z_t^K-x_t^K|\ge K^\eta\big) &= K^\eta\,\mathbb{P}\big(\sup_{t\le T\log K}e^{-(a-r)t}|\widetilde{M}_t^K|\ge K^\eta\big)\\ &\le 3\sup_{t\le T\log K}e^{-(a-r)t}\left(\mathbb{E}\langle\widetilde{M}_t^K\rangle\right)^{1/2}\end{aligned}$$

Using Lemma B.3, we deduce

$$\mathbb{P}\big(\sup_{t\le T\log K}e^{-at}|Z_t^K-x_t^K|\ge K^\eta\big)\le 3CK^{-\eta}\sup_{t\le T}K^{-(a-r)t+\frac{\beta\vee(\beta-rt)\vee(c+(a-2r)t)}{2}} \tag{B.10}$$

Step 1: Doob's inequality. Fix T satisfying (B.1). We consider first the case where we can find η such that

$$\frac{\beta\vee(\beta-rT)\vee(c+(a-2r)T)}{2}<\eta<\beta. \tag{B.11}$$

Then, it follows from (B.9) that the probability of the event Ω_1^K converges to 1, where

$$\Omega_1^K=\big\{\sup_{t\le T\log K}\big|e^{-rt}(Z_t^K-x_t^K)\big|\le K^\eta\big\}.$$

On this event, a similar computation as in Step 2 of Lemma A.1 entails

$$\sup_{t\le T}\left|\frac{\log(1+Z_{t\log K}^K)}{\log K}-\bar\beta_t\right|\le\sup_{t\le T}\left[\frac{1}{\log K}\frac{K^{-rt}|Z_{t\log K}^K-x_{t\log K}^K|}{K^{-rt}\left(x_{t\log K}^K\wedge Z_{t\log K}^K\right)}+\frac{\log\left(K^{-\bar\beta_t}(1+x_{t\log K}^K)\right)}{\log K}\right].$$

Because of Lemma B.2, we observe that $\bar{\beta}_t$ has been chosen such that, for all $t \leq T$,

$$C^{-1}K^{\bar{\beta}_t} \leq x^K_{t\log K} \leq CK^{\bar{\beta}_t} \tag{B.12}$$

for some constant $C > 0$. Hence,

$$\log\left(K^{-\bar{\beta}_t}(1 + x^K_{t\log K})\right) \leq \log(C + K^{-\bar{\beta}_t}) \leq C',$$

where the last inequality follows from the assumption that $\inf_{t\leq T} \bar{\beta}_t > 0$.

Hence, on the event Ω^K_1,

$$\begin{aligned}\sup_{t\leq T}\left|\frac{\log(1 + Z^K_{t\log K})}{\log K} - \bar{\beta}_t\right| &\leq \left[\frac{K^\eta}{\log K}\sup_{t\leq T}\frac{K^{rt}}{\left(x^K_{t\log K} - K^{\eta+rt}\right)_+} + \frac{C'}{\log K}\right] \\ &\leq \left[\frac{K^\eta}{\log K}\sup_{t\leq T}\frac{2K^{rt}}{x^K_{t\log K}} + \frac{C'}{\log K}\right],\end{aligned}$$

where we used (B.12) and the fact that $K^{\eta+rt} = o(x^K_{t\log K})$, since $\eta < \beta$. Hence,

$$\begin{aligned}\sup_{t\leq T}\left|\frac{\log(1 + Z^K_{t\log K})}{\log K} - \bar{\beta}_t\right| &\leq C\left[\frac{K^\eta}{\log K}K^{-[\beta\vee\inf_{t\leq T}(c+(a-r)t)]} + \frac{1}{\log K}\right] \\ &\leq C\left[\frac{K^{\eta-\beta}}{\log K} + \frac{1}{\log K}\right].\end{aligned} \tag{B.13}$$

Case 1(a): $r \geq 0$ and $a \leq 2r$. In this case, the constraint (B.11) becomes $\beta/2 < \eta < \beta$, hence we can choose $\eta = 3\beta/4$ for any value of T (note that (B.1) is always satisfied here). Hence (B.13) implies that $\sup_{t\leq T}\left|\frac{\log(1+Z^K_{t\log K})}{\log K} - \bar{\beta}_t\right|$ converges to 0 on the event Ω^K_1.

Case 1(b): $r < 0$ and $a \leq r$. In this case, Assumption (B.1) is equivalent to $\inf_{t\leq T}(\beta + rt) > 0$, i.e. $T < \beta/|r|$. Since $c + (a - 2r)T \leq \beta - rT + (a - r)T \leq \beta - rT < 2\beta$, it is possible to find η such that

$$\frac{\beta \vee (\beta - rT) \vee (c - (r \wedge (2r - a))T)}{2} = \frac{\beta - rT}{2} < \eta < \beta$$

and (B.13) allows again to conclude.

Case 1(c): $r \geq 0$ and $a > 2r$. In this case, (B.11) is satisfied provided T is such that $c + (a - 2r)T < 2\beta$, i.e. $T < T^* := \frac{2\beta-c}{a-2r}$. Let us observe that the two lines $\beta + rt$ and $c + at$ intersect at time $t^* = \frac{\beta-c}{a-r}$, and in our case, $T^* > t^*$ since $a\beta > rc$. Therefore, we can apply the computation (B.13) to $T \in (t^*, T^*)$ to obtain the convergence of $\sup_{t\leq T}\left|\frac{\log(1+Z^K_{t\log K})}{\log K} - \bar{\beta}_t\right|$ to 0. We explain below (in step 3) how to conclude for any value of T satisfying (B.1).

Case 1(d): $r < 0$, $a > r$ and $c + a\beta/|r| \leq 0$. In this case, $\beta + rt \geq c + at$ for all $t \leq \beta/|r|$ and hence Assumption (B.1) is satisfied if and only if $T < \beta/|r|$. For such T, (B.11) is satisfied since $c + (a - 2r)T < c + (a - 2r)\frac{\beta}{|r|} = 2\beta + c + a\frac{\beta}{|r|} \leq 2\beta$.

Case 1(e): $r<0$, $a>r$ and $c+a\beta/|r|>0$. In this case, Condition (B.11) is satisfied provided $T<T^*=\frac{\beta}{|r|}\wedge\frac{2\beta-c}{a-2r}$. Since $a\beta>rc$, we actually have $T^*=\frac{2\beta-c}{a-2r}$. In addition, if we define $t^*=\frac{\beta-c}{a-r}$ the first time where the line $\beta+rt$ intersects the line $c+at$, we can see exactly as in Case 1(c) that $T^*>t^*$. Hence, we can apply the computation (B.13) to $T\in(t^*,T^*)$ to obtain the convergence of $\sup_{t\le T}\left|\frac{\log(1+Z^K_{t\log K})}{\log K}-\bar\beta_t\right|$ to 0. We explain below (in step 3) how to conclude for any value of T satisfying (B.1).

Step 2: maximal inequality. We restrict here to the case $\beta=c$ and $a>r$. In this case,

$$\begin{aligned}\sup_{t\le T}\frac{(\beta-2(a-r)t)\vee(\beta-(2a-r)t)\vee(c-at)}{2}&=\sup_{t\le T}\frac{(\beta-2(a-r)t)\vee(\beta-at)}{2}\\&=\sup_{t\le T}\frac{\beta-at}{2}=\frac{\beta\vee(\beta-aT)}{2},\end{aligned}$$

where the second equality comes from the fact that the maximum of $\beta-2(a-r)t$ is attained for $t=0$, and the function $\beta-at$ takes the same value at time $t=0$. Assuming

$$\frac{\beta\vee(\beta-aT)}{2}<\eta<\beta, \tag{B.14}$$

it follows from (B.10) that the probability of the event Ω_2^K converges to 1, where

$$\Omega_2^K=\big\{\sup_{t\le T\log K}\big|e^{-at}(Z_t^K-x_t^K)\big|\le K^\eta\big\}.$$

On this event, a similar computation as in Case 1 entails

$$\begin{aligned}\sup_{t\le T}\left|\frac{\log(1+Z^K_{t\log K})}{\log K}-\bar\beta_t\right|&\le\left[\frac{K^\eta}{\log K}\sup_{t\le T}\frac{K^{at}}{\left(x^K_{t\log K}-K^{\eta+at}\right)_+}+\frac{C'}{\log K}\right]\\&\le\left[\frac{K^\eta}{\log K}\sup_{t\le T}\frac{2K^{at}}{x^K_{t\log K}}+\frac{C'}{\log K}\right]\le C\left[\frac{K^{\eta-\beta}}{\log K}+\frac{1}{\log K}\right],\end{aligned}$$

where we used the fact that $K^{\eta+at}=o(K^{\beta+at})$, since $\eta<\beta$ and that $x^K_{t\log K}\ge CK^{\beta+(a\vee r)t}=CK^{\beta+at}$ since $c=\beta$.

Case 2(a): $c=\beta$, $a>r$ and $a\ge0$. In this case, we can choose any $\eta\in(\beta/2,\beta)$ and deduce the convergence of $\sup_{t\le T}\left|\frac{\log(1+Z^K_{t\log K})}{\log K}-\bar\beta_t\right|$ to 0.

Case 2(b): $c=\beta$, $a>r$ and $a<0$. In this case, Assumption (B.1) is equivalent to $\inf_{t\le T}(\beta+(a\vee r)t)=\inf_{t\le T}(\beta+at)>0$, which is satisfied if and only if $T<\beta/|a|$. Now, for such T, one can find η satisfying (B.14), so we can again conclude.

Step 3: gluing parts together. The remaining cases not covered by the previous steps are

- $r\ge0$, $a>2r$, $c<\beta$ and $T\ge T^*=\frac{2\beta-c}{a-2r}$;
- $r<0$, $a>r$, $c<\beta$, $c+\frac{a\beta}{|r|}$ and $T\ge T^*=\frac{2\beta-c}{a-2r}$.

In both cases, we recall that $T^* > t^*$, where t^* is the first time where $\beta + rt$ crosses $c + at$. So we can fix $T_1 \in (t^*, T^*)$ and apply Case 1(c) or Case 1(e) of Step 1 to obtain the convergence in probability in $L^\infty([0, T_1])$ of $\log(1 + Z^K_{t \log K})/\log K$ to $\bar{\beta}_t = (\beta + rt) \vee (c + at)$. In particular, for all $\varepsilon > 0$, on an event Ω^K_3 with probability converging to 1, $K^{c+aT_1-\varepsilon} \leq Z^K_{T_1 \log K} \leq K^{c+aT_1+\varepsilon}$ for K large enough.

Now, on Ω^K_3, standard coupling arguments show that, for all $t \geq 0$, $\hat{Z}^K_t \leq Z^K_{T_1 \log K + t} \leq \bar{Z}^K_t$, where $\hat{Z}^K$ is a $BPI_K(b, d, a, c+aT_1-\varepsilon, c+aT_1-\varepsilon)$ and $\bar{Z}^K$ is a $BPI_K(b, d, a, c+aT_1+\varepsilon, c+aT_1+\varepsilon)$. Hence, we can apply Case 2(a) or 2(b) of Step 2 to $\hat{Z}^K$ and $\bar{Z}^K$ to obtain the convergence of $\log(1 + \hat{Z}^K_{t \log K})/\log K$ to $c - \varepsilon + a(T_1 + t)$ and of $\log(1 + \bar{Z}^K_{t \log K})/\log K$ to $c + \varepsilon + a(T_1 + t)$ on $[0, T - T_1]$. Note that, in the case where $a < 0$, Assumption (B.1) requires that $T < c/|a|$. We can choose $\varepsilon > 0$ small enough to have $T < (c - \varepsilon)/|a|$. In this case, we have $T - T_1 < \frac{c+aT_1-\varepsilon}{|a|}$, so we can indeed apply Case 2(b) to $\hat{Z}^K$ on the time interval $[0, T - T_1]$.

Since $\varepsilon > 0$ is arbitrary, Markov's property allows to conclude. □

Our next goal is to extend Theorem B.1 to the case where $\beta = 0$ or $\beta > 0$ without assuming (B.1) nor $c \leq \beta$.

We first consider the case where $c > \beta$ in Lemma B.4. It shows that the population instantaneously (on the time scale $\log K$) reaches a population size of order K^c. We can use Markov's property at time $\varepsilon \log K$ and comparison techniques in a similar way as in Step 3 of Theorem B.1 to reduce the problem to the case $c \leq \beta$.

Lemma B.4 (Initial growth for strong immigration) *Assume that $\beta < c$. Then, for all $\varepsilon > 0$ and all $\bar{a} > |r| \vee |a|$,*

$$\lim_{K \to +\infty} \mathbb{P}\left(Z^K_{\varepsilon \log K} \in [K^{c-\bar{a}\varepsilon}, K^{c+\bar{a}\varepsilon}]\right) = 1.$$

Proof We give the proof in the case where $r \neq a$, $r \neq 0$ and $a \neq 2r$. The extension to the other cases is straightforward. Using Lemma B.2, there exists a positive constant C such that

$$C^{-1} K^{c-\varepsilon(|r|\vee|a|)} \leq \mathbb{E}(Z^K_{\varepsilon \log K}) \leq C K^{c+\varepsilon(|r|\vee|a|)}$$

and $\mathrm{Var}(Z^K_{\varepsilon \log K}) \leq C K^{c+(2|r|\vee|a|)\varepsilon}$. The result follows from Chebyshev's inequality. □

Let us now state our general result with $c \leq \beta$.

Theorem B.5 *Let $(Z^K_t, t \geq 0)$ be a $BPI_K(b, d, a, c, \beta)$ process with $c \leq \beta$ and assume either $\beta > 0$ or $c \neq 0$. The process $\left(\frac{\log(1+Z^K_{t \log K})}{\log K}, t > 0\right)$ converges when K tends to infinity in probability in $L^\infty([0, T])$ for all $T > 0$ to the continuous deterministic function $\bar{\beta}$ given by*

(i) *if $\beta > 0$, $\bar{\beta} : t \mapsto (\beta + rt) \vee (c + at) \vee 0$;*

(ii) *if $\beta = 0$, $c < 0$ and $a > 0$, $\bar{\beta} : t \mapsto ((r \vee a)(t - |c|/a)) \vee 0$;*

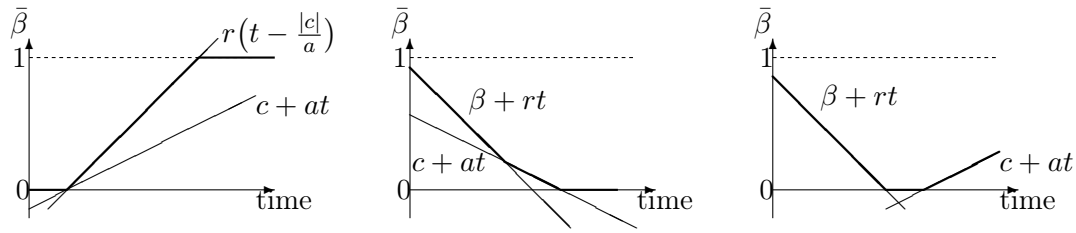
(iii) *if $\beta = 0$, $c < 0$ and $a \leq 0$, $\bar{\beta} : t \mapsto 0$.*

Note that, in case (i), when $r \geq 0$, one can remove $\vee 0$ from the definition of $\bar{\beta}_t$.

In addition, in the case where $c \neq 0$ or $a \neq 0$, for all compact interval $I \subset \mathbb{R}_+$ which does not intersect the support of $\bar{\beta}$,

$$\lim_{K \to +\infty} \mathbb{P}\left(Z^K_{t \log K} = 0, \ \forall t \in I\right) = 1. \tag{B.15}$$

The case $\beta = c = 0$ could be deduced from the previous result using comparison argument, but is not useful here.



(a): $c < \beta = 0$, $0 < a < r$ (b): $0 < c < \beta$, $r < a < 0$ (c): $c < 0 < \beta$, $r < 0 < a$

Figure B.1: *Illustration of Theorem B.5. (a): Initially $\bar\beta = 0$, but thanks to immigration, the population is revived. Once this happens, the growth rate r being larger than a, immigration have a negligible effect after time $|c|/a$. (b): After time $(\beta - c)/(a - r)$, the dynamics is driven by mutation before getting extinct when $\bar\beta_t = 0$. (c): We observe a local extinction before the population is revived thanks to incoming mutations.*

Remark B.6 *Note that point (i) in Theorem B.5 means that, at least for small $t > 0$, either $\beta = c$ and then $\overline{\beta}(t) = \beta + (r \vee a)t$ or $\beta > c$ and then $\overline{\beta}(t) = \beta + rt$. This explains Equation* (2.9) *in Corollary 2.3.*

Proof The proof combines Theorem B.1 with a series of lemmas and extensive use of Markov's property. The proofs of the lemmas are given at the end of the section.

Proof of (iii). Theorem B.5 (iii) follows directly from the next lemma. Note that it also proves that (B.15) holds true in case (iii) for all $I \subset [0, T]$.

Lemma B.7 (Non emergence of any new population) *Assume that $\beta = 0$ and $c < 0$. Let us consider $T > 0$ such that $-|c| + aT < 0$. Then:*

$$\lim_{K \to +\infty} \mathbb{P}\Big(Z_t^K = 0, \forall t \leq T \log K \Big) = 1. \tag{B.16}$$

Proof of (ii). We need to combine Theorem B.1 and Lemma B.7 with the next two lemmas.

Lemma B.8 (Emergence of a new population) *Assume that $\beta = 0$, that $c = -\varepsilon$ with $\varepsilon > 0$ and that $a > 0$ (so that the immigration rate starts being positive at time $t_0 = \frac{\varepsilon}{a} \log K$). Then, for all $\eta > (1 \vee \frac{2r}{a})\varepsilon$,*

$$\lim_{K \to +\infty} \mathbb{P}\Big(K^{\varepsilon/2} - 1 \leq Z^K_{\frac{2\varepsilon}{a} \log K} \leq K^{\eta} - 1 \Big) = 1. \tag{B.17}$$

The second lemma is valid for all values of $c \leq \beta$, r and a. It gives uniform estimates on the modulus of continuity of $\log(1 + Z^K_{t \log K})/\log K$.

Lemma B.9 (continuity of the exponent) *Let $(Z_t^K, t \geq 0)$ be a $BPI_K(b, d, a, c, \beta)$ with $c \leq \beta$. Then, there exists a constant $\bar c = \bar c(b, d, a)$ such that, for all $\varepsilon > 0$,*

$$\lim_{K \to +\infty} \mathbb{P}\Big(\forall t \in [0, \varepsilon \log K],\ K^{\beta - \bar c \varepsilon} - 1 \leq Z_t^K \leq K^{\beta + \bar c \varepsilon} - 1 \Big) = 1.$$

We proceed as follows. Fix $\varepsilon > 0$ and apply Lemma B.7 on the time interval $[0, t_1]$ with $t_1 = |c|/a - \varepsilon$. We deduce that

$$\frac{\log(1 + Z^K_{t \log K})}{\log K} = 0, \quad \forall t \in [0, t_1] \tag{B.18}$$

on an event Ω^K_1 with probability converging to 1 and, since $\bar{\beta}_t = 0$ if and only if $t \leq |c|/a$, (B.15) is proved in case (ii). Applying Markov's property at time $t_1 \log K$ on Ω^K_1. Since $Z^K_{t_1 \log K} = 0$ and $c + at_1 < 0$, we can apply Lemma B.8 to deduce that

$$\frac{\log(1 + Z^K_{(t_1+\delta\varepsilon) \log K})}{\log K} \in (\underline{c}\varepsilon, \bar{c}\varepsilon)$$

with probability converging to 1 for constants $\delta, \underline{c} > 0$ and $\bar{c} < \infty$ independent of ε and K. In addition, Lemma B.9 implies that, with probability converging to 1,

$$\sup_{t \in [t_1, t_1+\delta\varepsilon]} \frac{\log(1 + Z^K_{t \log K})}{\log K} \leq \bar{c}'\varepsilon \tag{B.19}$$

for a constant $\bar{c}' > \bar{c}$ independent of ε and K.

Using the comparison trick of Step 3 of the proof of Theorem B.1 after time $(t_1 + \delta) \log K$, we can then apply Theorem B.1 to prove that, with probability converging to 1, for all $t \in [t_1 + \delta\varepsilon, T]$,

$$\begin{aligned} \underline{c}\varepsilon + (r \vee a)(t - t_1 - \delta\varepsilon) &\leq \liminf_{K \to +\infty} \frac{\log(1 + Z^K_{t \log K})}{\log K} \\ &\leq \limsup_{K \to +\infty} \frac{\log(1 + Z^K_{t \log K})}{\log K} \leq \bar{c}\varepsilon + (r \vee a)(t - t_1 - \delta\varepsilon). \end{aligned} \tag{B.20}$$

We conclude combining (B.18), (B.19) and (B.20) and letting $\varepsilon \to 0$.

Proof of (i). Note that, when $r \geq 0$, Point (i) has been already proved in Theorem B.1. Similarly, if $r < 0$, $a \geq 0$ and $c + a\beta/|r| > 0$, Point (i) also follows directly from Theorem B.1. In these two cases, (B.15) is also trivial. We divide our study of the remaining cases in four.

Case (a): $r < 0$, $a < 0$ or $r < 0$, $a = 0$ and $c < 0$. In this case, we combine the previous lemmas with the next one, similarly to the proof of (ii). Note that we need to use Lemma B.10 (i) if $c + a\beta/|r| < 0$, Lemma B.10 (ii) if $c + a\beta/|r| > 0$. If $c + a\beta/|r| = 0$, we use a comparison between Z^K and a $BPI_K(b, d, a, c - \varepsilon, \beta)$ and a $BPI_K(b, d, a, c + \varepsilon, \beta \vee (c + \varepsilon))$, for which the previous cases apply, and let $\varepsilon \to 0$.

Lemma B.10 (Extinction) *Assume that $r < 0$.*

(i) *Asume also that $c < 0$ and $c + a\beta/|r| < 0$. Then for all $\eta > 0$ small enough,*

$$\lim_{K \to +\infty} \mathbb{P}\Big(\forall t \in \Big[\big(\frac{\beta}{|r|} + \eta\big) \log K, \big(\frac{\beta}{|r|} + 2\eta\big) \log K\Big],\ Z^K_t = 0\Big) = 1. \tag{B.21}$$

(ii) *Assume also that $c + a\beta/|r| > 0$ and $a < 0$. Then for all $\eta > 0$ and $T > \eta$,*

$$\lim_{K \to +\infty} \mathbb{P}\Big(\forall t \in \Big[\big(\frac{c}{|a|} + \eta\big) \log K, \big(\frac{c}{|a|} + T\big) \log K\Big],\ Z^K_t = 0\Big) = 1. \tag{B.22}$$

Case (b): $r < 0$, $a > 0$, $c < 0$ and $\frac{\beta}{|r|} < \frac{|c|}{a}$. This corresponds to Figure B.1 (c). In this case, we combine as above Theorem B.1, Lemma B.10, Lemma B.7 and Lemma B.8.

Case (c): $r < 0$, $a > 0$, $c < 0$ and $\frac{\beta}{|r|} = \frac{|c|}{a}$. This can be treated using comparisons between Z_t^K and $BPI_K(b, d, a, c - \varepsilon, \beta)$ and $BPI_K(b, d, a, c + \varepsilon, \beta \vee (c + \varepsilon))$ and letting $\varepsilon \to 0$.

Case (d): $r < 0$, $c = a = 0$. This can be treated using comparisons between Z_t^K and $BPI_K(b, d, 0, -\varepsilon, \beta)$ and $BPI_K(b, d, 0, \varepsilon, \beta \vee \varepsilon)$ and letting $\varepsilon \to 0$. □

Proof of Lemma B.7. Since the rate of immigration is upper bounded by $K^{c\vee(c+aT)}$ on $[0, T \log K]$, the probability that a migrant arrives during this time interval is upper bounded by $TK^{c\vee(c+aT)} \log K$ which tends to 0 when $K \to +\infty$. The lemma is proved. □

Proof of Lemma B.8. The number of immigrant families which arrived during the time interval $[0, \frac{2\varepsilon}{a} \log K]$ and which survived up to time $\frac{2\varepsilon}{a} \log K$ is, by thinning, a Poisson random variable with parameter

$$\lambda = \int_0^{\frac{2\varepsilon}{a} \log K} K^{-\varepsilon} e^{at} \frac{rK^{\frac{2r\varepsilon}{a}} e^{-rt}}{bK^{\frac{2r\varepsilon}{a}} e^{-rt} - d} dt.$$

This formula is obtained by (A.4), where the probability of keeping a family immigrated at time t is the fraction in the above expression.
In the case where $r \geq 0$,

$$\lambda \geq \frac{r}{b} \int_0^{\frac{2\varepsilon}{a} \log K} K^{-\varepsilon} e^{at} dt \geq \frac{rK^{\varepsilon}}{2ab}.$$

In the case where $r < 0$,

$$\lambda \geq \frac{d-b}{d} K^{\frac{2r\varepsilon}{a} - \varepsilon} \int_0^{\frac{2\varepsilon}{a} \log K} e^{(a+|r|)t} dt \geq \frac{|r| K^{\varepsilon}}{2d(a + |r|)}.$$

Therefore,

$$\lim_{K \to +\infty} \mathbb{P}\left(Z^K_{\frac{2\varepsilon}{a} \log K} \geq K^{\varepsilon/2} \right) = 1.$$

For the upper bound, it follows from (B.2) that

$$\mathbb{E}\left(Z^K_{\frac{2\varepsilon}{a} \log K} \right) \leq \begin{cases} \frac{K^{\varepsilon}}{a-r} & \text{if } r < a, \\ \frac{4\varepsilon}{a} K^{\varepsilon} \log K & \text{if } r = a, \\ \frac{K^{\frac{2r}{a}\varepsilon}}{r-a} & \text{if } r > a. \end{cases}$$

Therefore, (B.17) follows from Markov's inequality and the choice $\eta > (1 \vee \frac{2r}{a})\varepsilon$. □

Proof of Lemma B.9. The number of immigrant families which arrive during the time interval $[0, \varepsilon \log K]$ and which survive up to time $\varepsilon \log K$ is a Poisson random variable with parameter

$$\lambda \leq \int_0^{\varepsilon \log K} K^c e^{at} dt \leq K^{\beta} \frac{K^{(|a|+1/2)\varepsilon}}{|a| + 1/2}.$$

The maximal size of each families (immigrant or present at time 0) on the time interval $[0, \varepsilon \log K]$ is bounded by the size at time $\varepsilon \log K$ of a Yule process with birth rate b, i.e. a geometric random variable G_i with parameter $K^{b\varepsilon}$, independently for each immigrant families. Hence, with probability converging to 1,

$$\sup_{t\in[0,\varepsilon \log K]} Z_t^K \leq \sum_{i=1}^{K^{\beta+(|a|+3/4)\varepsilon}} G_i \leq K^{\beta+(|a|+1+b)\varepsilon} \tag{B.23}$$

For the lower bound, we observe that Z_t^K is bigger than a linear pure death process $BP(0, d, \beta)$. For each of the $\lfloor K^\beta - 1 \rfloor$ initial individuals, the probability of survival up to time $\varepsilon \log K$ is $K^{-d\varepsilon}$, hence, with probability converging to 1, $\inf_{t\in[0,\varepsilon \log K]} Z_t^K \geq K^{\beta-2d\varepsilon}$. Hence the lemma is proved with $\bar{c} = (2d) \vee (|a| + b + 1)$. □

Proof of Lemma B.10. We first prove (i). Using the same argument as in the proof of Lemma B.7, we can prove that the probability of the event Γ that a migrant arrives during the time interval $[0, T \log K]$ converges to 0. Therefore, on the complementary event Γ^c and on this time interval, the process Z^K is a $BP(b, d, K^\beta)$ and, using (A.4), its extinction time T_{ext} satisfies, for all $t \leq T \log K$,

$$\mathbb{P}(T_{\text{ext}} > t; \Gamma^c) \leq 1 - \left(1 - \frac{re^{rt}}{be^{rt} - d}\right)^{K^\beta}.$$

Thus, for $t = \eta \log K$ with $\eta > \beta/|r|$, there exists a constant C such that

$$\mathbb{P}(T_{\text{ext}} > \eta \log K) \leq \mathbb{P}(\Gamma) + C\, K^{-|r|\eta+\beta} \xrightarrow[K\to+\infty]{} 0.$$

Now, let us prove (ii). Using Markov's property and Theorem B.1, we can assume without loss of generality that $c \leq \beta < \eta|r|/4$. Note that assuming $c + a\beta/|r| > 0$ implies that $a > r$. In addition, we can prove as above that, with a probability converging to 1, there is no new immigrant arriving in the population between times $(c/|a| + \eta/2) \log K$ and $(c/|a| + T) \log K$. Hence we only need to check that all the families that descend from each of the individuals initially present in the population and of all immigrants which arrive before time $(c/|a| + \eta/2) \log K$ are extinct before time $(c/|a| + \eta) \log K$. Using (A.4), each of these families has a probability to survive longer than a time $(\eta/2) \log K$ which is smaller than $2K^{-\eta|r|/2}$ (for K large enough). Since the number of these families is equal to K^β plus a Poisson random variable of parameter $K^c \int_0^{(c/|a|+\eta/2)\log K} e^{as} ds \leq K^c/a$, we deduce that it is smaller that $K^{\eta r/3}$ with probability converging to 1. Hence, on this event, the probability that at least one of these families survives up to time $(c/|a| + \eta) \log K$ is smaller than

$$\left(1 - 2K^{-\eta|r|/2}\right)^{K^{\eta r/3}} \xrightarrow{K\to 0} 0.$$

This concludes the proof of Lemma B.10. □

C Logistic birth and death process with immigration

C.1 One-dimensional case

We consider here a one-dimensional logistic birth and death process with individual birth rate b and individual death rate $d+Ck/K$ when the population size is k and immigration at predictable rate $\gamma(t) \geq 0$ at time t. We denote by $LBDI_K(b,d,C,\gamma)$ the law of this process. We consider a specific initial condition in the next result. Recall that $r = b - d$.

Lemma C.1 *Assume that $(Z_t^K, t \geq 0)$ follows the law $LBDI_K(b,d,C,\gamma)$ with $b > d$. Let $T > 0$ and assume that $\gamma(t) \leq K^{1-\alpha}$ for all $t \in [0, T\log K]$ for some $\alpha > 0$.*

(i) *If $\frac{Z_0^K}{K} \in \left[\frac{r}{C} - \varepsilon, \frac{r}{C} + \varepsilon\right]$ for some $\varepsilon > 0$, then*

$$\lim_{K\to+\infty} \mathbb{P}\left(\forall t \in [0, T\log K],\ \frac{Z_t^K}{K} \in \left[\frac{r}{C} - 2\varepsilon, \frac{r}{C} + 2\varepsilon\right]\right) = 1.$$

(ii) *For all $\varepsilon, \varepsilon' > 0$, there exists $T(\varepsilon,\varepsilon') < +\infty$ such that for all initial condition $\frac{Z_0^K}{K} \geq \varepsilon$ we have that*

$$\lim_{K\to+\infty} \mathbb{P}\left(\frac{Z_{T(\varepsilon,\varepsilon')}^K}{K} \in \left[\frac{r}{C} - 2\varepsilon', \frac{r}{C} + 2\varepsilon'\right]\right) = 1.$$

Proof This result is related to the problem of exit from a domain of [Freidlin-Wentzell] and can be proved with standard arguments as in [4, 6]. The only difficulty comes from the additional immigration rate (smaller than K^{1-a} for all $t \in [0, T\log K]$), which is negligible with respect to the reproduction rate (of order K), but this can be handled for example adapting the proof of [5, Prop. 4.2]. □

C.2 Two-dimensional case

C.2.1 Transfer birth-death process with immigration

We consider a two-dimensional transfer process with immigration $(Y_t^K, Z_t^K)_{t\geq 0}$, with transition rates from $(n,m) \in \mathbb{N}^2$ to

$$\begin{aligned}
(n+1,m) &\text{ with rate } nb_1^K(\omega,t) + \gamma_1^K(\omega,t),\\
(n-1,m) &\text{ with rate } nd_1^K(\omega,t),\\
(n,m+1) &\text{ with rate } mb_2^K(\omega,t) + \gamma_2^K(\omega,t),\\
(n,m-1) &\text{ with rate } md_2^K(\omega,t),\\
(n-1,m+1) &\text{ with rate } \tau^K(\omega,t)\frac{nm}{n+m},
\end{aligned}$$

with $C > 0$ and predictable rates $\tau^K, b_i^K, d_i^K, \gamma_i^K : \Omega \times \mathbb{R}_+ \to \mathbb{R}_+$. Remark that transfer only occurs from population Y^K to population Z^K.
We denote by $TBDI_K(b_1^K, d_1^K, b_2^K, d_2^K, \tau^K, \gamma_1^K, \gamma_2^K)$ the law of (Y^K, Z^K).

Lemma C.2 *Assume that* $(Y_t^K, Z_t^K)_{t\geq 0}$ *follows the law* $TBDI_K(b_1^K, d_1^K, b_2^K, d_2^K, \tau^K, \gamma_1^K, \gamma_2^K)$ *and that there exist constants* $b_1, d_1, b_2, d_2, \tau > 0$ *such that for some* $s > 0$,

$$\sup_{t\in[0,s\log K]} \|b_1^K(t) - b_1\| + \|b_2^K(t) - b_2\| + \|d_1^K(t) - d_1\| + \|d_2^K(t) - d_2\| + \|\tau^K(t) - \tau\| + \|\gamma_1^K(t)\| + \|\gamma_2^K(t)\| \xrightarrow[K\to+\infty]{} 0 \quad \text{(C.1)}$$

in probability. Let $S := r_2 - r_1 + \tau$.

(i) *Assume* $S > 0$, *that* $Y_0^K \geq K^\beta$ *for some* $\beta > 0$ *and that* $\eta Y_0^K < Z_0^K < Y_0^K$ *for some* $\eta > 0$. *Then, there exists* $T = T(\eta) < \infty$ *and* $\rho > 0$ *such that for* $s > 0$ *small enough,*

$$\lim_{K\to+\infty} \mathbb{P}\left(Y_{s\log K}^K \leq K^{-s\rho} Z_{s\log K}^K\right) = 1. \quad \text{(C.2)}$$

(ii) *Assume* $S < 0$, *that* $Z_0^K \geq K^\beta$ *for some* $\beta > 0$ *and that* $\eta Z_0^K < Y_0^K < Z_0^K$ *for some* $\eta > 0$. *Then, there exists* $T = T(\eta) < \infty$ *and* $\rho > 0$ *such that for* $s > 0$ *small enough,*

$$\lim_{K\to+\infty} \mathbb{P}\left(Z_{s\log K}^K \leq K^{-s\rho} Y_{s\log K}^K\right) = 1. \quad \text{(C.3)}$$

Lemma C.2 (i) says that, if the population Y^K is initially dominant and if the population Z^K has a positive fitness and is initially not negligible, after a short time on the time-scale $\log K$, Y^K becomes negligible with respect to Z^K. For the proof, we refer to Lemma C.3 (iii).

C.2.2 Logistic transfer birth-death process with immigration

We consider a two-dimensional logistic transfer process with immigration $(Y_t^K, Z_t^K)_{t\geq 0}$, with transition rates from $(n, m) \in \mathbb{N}^2$ to

$$\begin{aligned}
(n+1, m) &\text{ with rate } n b_1^K(\omega, t) + \gamma_1^K(\omega, t),\\
(n-1, m) &\text{ with rate } n\left[d_1^K(\omega, t) + \frac{C}{K}(n+m)\right],\\
(n, m+1) &\text{ with rate } m b_2^K(\omega, t) + \gamma_2^K(\omega, t),\\
(n, m-1) &\text{ with rate } m\left[d_2^K(\omega, t) + \frac{C}{K}(n+m)\right],\\
(n-1, m+1) &\text{ with rate } \tau^K(\omega, t)\frac{nm}{n+m},
\end{aligned}$$

with $C > 0$ and predictable rates $\tau^K, b_i^K, d_i^K, \gamma_i^K : \Omega \times \mathbb{R}_+ \to \mathbb{R}_+$. Remark that transfer only occurs from population Y^K to population Z^K.
We denote by $LTBDI_K(b_1^K, d_1^K, b_2^K, d_2^K, C, \tau^K, \gamma_1^K, \gamma_2^K)$ the law of (Y^K, Z^K).

Lemma C.3 (Competition) *Assume that* $(Y_t^K, Z_t^K)_{t\geq 0}$ *follows the law* $LTBDI_K(b_1^K, d_1^K, b_2^K, d_2^K, C, \tau^K, \gamma_1^K, \gamma_2^K)$ *and that there exist constants* $b_1, d_1, b_2, d_2, \tau > 0$ *such that for some* $s > 0$, *the convergence* (C.1) *is satisfied.*

(i) *Assume that* $r_1 := b_1 - d_1 > 0$, $r_2 := b_2 - d_2 > 0$, $S := r_2 - r_1 + \tau > 0$ *and* $\frac{Y_0^K}{K} \in \left[\frac{r_1}{C} - \varepsilon, \frac{r_1}{C} + \varepsilon\right]$ *for some* $\varepsilon > 0$ *and* $\frac{Z_0^K}{K} \geq m\varepsilon$ *for some* $m > 0$. *Then, for all* $\varepsilon' > 0$, *there exists* $T = T(m, \varepsilon, \varepsilon') < \infty$ *such that*

$$\lim_{K \to +\infty} \mathbb{P}\left(Y_T^K \leq \varepsilon' K,\ \frac{Z_T^K}{K} \in \left[\frac{r_2}{C} - \varepsilon', \frac{r_2}{C} + \varepsilon'\right]\right) = 1.$$

(ii) *Assume that* $r_1 > 0$, $r_2 > 0$, $S < 0$ *and* $\frac{Y_0^K}{K} \geq m\varepsilon$ *for some* $\varepsilon, m > 0$ *and* $\frac{Z_0^K}{K} \in \left[\frac{r_2}{C} - \varepsilon, \frac{r_2}{C} + \varepsilon\right]$. *Then, for all* $\varepsilon' > 0$ *there exists* $T = T(m, \varepsilon, \varepsilon') < \infty$ *such that*

$$\lim_{K \to +\infty} \mathbb{P}\left(\frac{Y_T^K}{K} \in \left[\frac{r_1}{C} - \varepsilon', \frac{r_1}{C} + \varepsilon'\right],\ Z_T^K \leq \varepsilon' K,\right) = 1.$$

(iii) *Assume that* $r_1 > 0$, $r_2 < 0$, $S > 0$ *and* $\frac{Y_0^K}{K} \in \left[\frac{r_1}{C} - \varepsilon, \frac{r_1}{C} + \varepsilon\right]$ *for some* $\varepsilon > 0$ *and* $\frac{Z_0^K}{K} \geq m\varepsilon$ *for some* $m > 0$. *Then there exists a constant* $\rho > 0$ *such that, for all* $s > 0$ *small enough,*

$$\lim_{K \to +\infty} \mathbb{P}\left(Y_{s \log K}^K \leq K^{-s\rho} Z_{s \log K}^K,\ Z_{s \log K}^K \leq K^{1-s\rho},\ \right) = 1. \tag{C.4}$$

In Lemma C.3 (iii), the initial population Y^K is resident and after invasion, the population Z^K becomes dominant but not resident, as can be seen from the exponent $1 - s\rho$ in (C.4).

Proof Let us first prove (i). By Condition (C.1), the proofs of [1] or [9, Ch. 11] can be easily adapted to prove that, when $\frac{1}{K}(Y_0^K, Z_0^K)$ converges in probability to $(y_0, z_0) \in (0, \infty)^2$, the process (Y^K, Z^K) converges in probability in $L^\infty_{\text{loc}}(\mathbb{R}_+)$ to the solution $(y(t), z(t))$ of the dynamical system

$$\begin{cases} \dot{y} &= y\left(r_1 - C(y+z)\right) - \tau \frac{yz}{y+z} \\ \dot{z} &= z\left(r_2 - C(y+z)\right) + \tau \frac{yz}{y+z}. \end{cases} \tag{C.5}$$

In our frequency-dependent case with constant competition, invasion implies fixation [1, Section 3.3.1], so that, for all $(y_0, z_0) \in (0, \infty)^2$, $(y(t), z(t))$ converges when $t \to +\infty$ to $(0, r_2/C)$ (since $S > 0$). The lemma follows from this result as in [4, Thm. 3(b)].

The proof of (ii) follows the same arguments.

Let us turn to the proof of (iii). First, we remark that, for $y(t)$, $z(t)$ solution to (C.5), $p(t) = y(t)/z(t)$ solves $\dot{p} = -Sp$ when (y, z) solves (C.5). Hence, if $y(0) \in \left[\frac{r_1}{C} - \varepsilon, \frac{r_1}{C} + \varepsilon\right]$ and $z(0) \geq m\varepsilon$,

$$p(t) \leq \frac{r_1/C + \varepsilon}{m\varepsilon} e^{-St}.$$

In particular, there exists t_0 such that, for all $t \geq t_0$, $y(t) \leq \frac{|r_2|}{2} z(t)$. Therefore, $\dot{z}/z \leq r_2 + |r_2|/2 = r_2/2 < 0$ for all $t \geq t_0$. Therefore, $(y(t), z(t))$ converges to $(0, 0)$ for $t \to +\infty$. In addition, for any $\eta > 0$, there exists t_η large enough such that the solutions at time t_η of the dynamical system issued from any initial condition in $\left[\frac{r_1}{C} - \varepsilon, \frac{r_1}{C} + \varepsilon\right] \times [m\varepsilon, \infty)$, belongs to a compact subset of $\mathcal{C}_\eta = \{(y, z) \in B(0, \eta); 0 < y/z < \eta\}$.

We deduce that there exists a compact subset of $\mathcal{C}_{1.5\eta}$ containing $(Y_{t_\eta}^K/K, Z_{t_\eta}^K/K)$ with a probability close to 1 for K large enough. In particular, there exists a constant $\kappa > 0$ such that

$$\mathbb{P}\left(Y_{t_\eta}^K > \kappa K,\ Z_{t_\eta}^K > \kappa K\right) \xrightarrow[K \to +\infty]{} 1 \tag{C.6}$$

After time t_η, we can construct a coupling between the stochastic population processes as follows: for all t smaller than $T^{\text{exit}}_{2\eta}$ the first exit time of $\text{Adh}(K\mathcal{C}_{2\eta})$ by (Y^K, Z^K),

$$A^0_t \leq Y^K_{t+t_\eta} \leq A^1_t + I^1_t, \quad A^2_t \leq Z^K_{t+t_\eta} \leq A^3_t + I^3_t,$$

where $A^0_0 = A^1_0 = Y^K_{t_\eta}$, $A^2_0 = A^3_0 = Z^K_{t_\eta}$, A^0 has law $BP(b_1, d_1+\tau+4C\eta)$, A^1 has law $BP(b_1, d_1+\tau/(1+2\eta))$, A^2 has law $BP(b_2, d_2+4C\eta)$, A^3 has law $BP(b_2+2\tau\eta/(1+2\eta), d_2)$, and the random variables I^1_t and I^3_t count, in Y^K and Z^K respectively, the number of individuals alive at time $t+t_\eta$ born from immigrant individuals which arrived in the population after time t_η. Note that all theses processes are not necessarily independent.

Using domination of immigrant population by Yule processes as in the proof of Lemma B.9, we can prove as in (B.23) (with $\beta = 0$) that, for $s > 0$ small enough,

$$\lim_{K\to+\infty} \mathbb{P}\left(\forall t \leq s\log K,\ I^1_t + I^3_t \leq K^{1-\alpha/2}\right) = 1. \tag{C.7}$$

Let us first prove that $T^{\text{exit}}_{2\eta} > s\log K$ with probability converging to 1. We first need to check that $(Y^K_t, Z^K_t) \in B(0, 2\eta)$ for all $t \leq s\log K$ with probability converging to 1. Because of (C.7), it is enough to prove that

$$\mathbb{P}_{\lfloor 1.5\eta K\rfloor}(T_{ext}(A^1) < T_{2\eta K}(A^1)) \xrightarrow[K\to+\infty]{} 1$$

and similarly for A^3, where $T_{\text{ext}}(A^i)$ is the extinction time of A^i and $T_M(A^i)$ is the first time t such that $A^i_t \geq M$. It is classical to prove that, for $A \sim BP(b, d)$ with $b < d$,

$$\mathbb{P}_k(T_{\text{ext}}(A) < T_n(A)) = \frac{\left(\frac{d}{b}\right)^n - \left(\frac{d}{b}\right)^k}{\left(\frac{d}{b}\right)^n - 1},$$

so the result follows.

The second steps consists in proving that $Y^K_t \leq 2\eta Z^K_t$ for all $t \in [t_\eta, t_\eta + s\log K]$. Recall that $\kappa K \leq Y^K_{t_\eta} \leq 1.5\eta Z^K_{t_\eta}$ with high probability. Using (C.6), we can apply Lemma A.1 to obtain that, for $s > 0$ small enough and all $\varepsilon' > 0$, for all $t \leq s\log K$,

$$\begin{aligned} A^1_t &\leq (1+\varepsilon')A^1_0 e^{(r_1-\tau/(1+2\eta))t} \\ A^2_t &\geq (1-\varepsilon')A^2_0 e^{(r_2-4C\eta)t}. \end{aligned} \tag{C.8}$$

We choose $\eta > 0$ and $\varepsilon' > 0$ small enough so that $r_1 - \tau/(1+2\eta) < r_2 - 4C\eta$ and $1.5(1+\varepsilon') < 2(1-\varepsilon')$. Hence, with probability converging to 1, for all $t \leq s\log K$,

$$Y^K_{t+t_\eta} \leq A^1_t \leq 2(1-\varepsilon')\eta A^2_0 e^{(r_1-\tau/(1+2\eta))t} \leq 2\eta Z^K_{t+t_\eta}.$$

So we have proved that $T^{\text{exit}}_{2\eta} > s\log K$ with probability converging to 1.

In addition, it follows from (C.8) that

$$Y^K_{t+t_\eta} \leq 2\eta \exp\left((r_1 - r_2 - \tau/(1+2\eta) + 4C\eta)t\right) Z^K_{t+t_\eta}.$$

Introducing $0 < \rho < -(r_1 - r_2 - \tau/(1+2\eta) + 4C\eta)$ and choosing $t = s\log K - t_\eta$ and $\eta < \varepsilon/2$, we deduce (C.4). □

D Algorithmic construction of the slopes of $\beta(t)$

We defined in Theorem 2.1 the times s_k of change of resident or dominant populations. For algorithmic purpose , it is useful to characterize all the times t_k where the functions $\beta_\ell(t)$ change their slopes. The successive slopes are given in the next result which stems from Corollary 2.3. The next formula (D.2) is explained after the statement of the theorem. We define

$$\Sigma^0_\ell(t) := \Sigma_\ell(t)\mathbb{1}_{\beta_\ell(t)>0 \text{ or } (\beta_\ell(t)=0 \text{ and } \beta_{\ell-1}(t)=\alpha)} \tag{D.1}$$

Theorem D.1 *Under the same assumptions as in Theorem 2.1, the limit $\beta(t)$ is continuous and piecewise affine and can be constructed recursively as follows: assume that, for some $k \geq 0$, we have constructed times $0 = t_0 < t_1 < \ldots < t_k < T_0$ and integers $\ell'_0, \ldots, \ell'_{k-1}$ such that $(\beta(t), t \in [t_i, t_{i+1}])$ is affine and $\ell^*(t) = \ell'_i$ for all $t \in [t_i, t_{i+1})$ for all $0 \leq i \leq k-1$. Recall that, since $t_k < T_0$, we have Card$(Argmax_{0\leq\ell\leq L}\beta_\ell(t_i)) \in \{1, 2\}$ for all $0 \leq i \leq k$.*

Then

$$\begin{aligned}
t_{k+1} = t_k + \Bigg(&\inf\left\{\frac{\beta_{\ell'_k}(t_k) - \beta_\ell(t_k)}{\Sigma^0_\ell(t_k+) - \widetilde{S}_{t_k+}(\ell'_k\delta; \ell'_k\delta)}; \ell \neq \ell'_k \text{ s.t. } \Sigma^0_\ell(t_k+) > \widetilde{S}_{t_k+}(\ell'_k\delta; \ell'_k\delta)\right\} \\
&\wedge \inf\left\{\frac{\beta_\ell(t_k)}{-\Sigma^0_\ell(t_k+)}; \ell \text{ s.t. } \beta_\ell(t_k) > 0 \text{ and } \Sigma^0_\ell(t_k+) < 0\right\} \\
&\wedge \inf\Bigg\{\frac{\beta_\ell(t_k) - \beta_{\ell-1}(t_k) + \alpha}{\Sigma^0_{\ell-1}(t_k+) - \widetilde{S}_{t_k+}(\ell\delta, \ell'_k\delta)\mathbb{1}_{\beta_\ell(t_k)>0}}; \ell \neq \ell'_k \text{ s.t. } \beta_\ell(t_k) > \beta_{\ell-1}(t_k) - \alpha \\
&\qquad\qquad \text{and } \Sigma^0_{\ell-1}(t_k+) - \widetilde{S}_{t_k+}(\ell\delta, \ell'_k\delta)\mathbb{1}_{\beta_\ell(t_k)>0} > 0\Bigg\} \\
&\wedge \inf\left\{\frac{1 - \beta_{\ell'_k}(t_k)}{\Sigma^0_{\ell'_k}(t_k+)}; \ \Sigma^0_{\ell'_k}(t_k+) > 0 \text{ and } \beta_{\ell'_k}(t_k) < 1\right\}\Bigg)
\end{aligned} \tag{D.2}$$

and, for all $t \in [0, t_{k+1} - t_k]$,

$$\beta_\ell(t + t_k) = \beta_\ell(t_k) + \Sigma^0_\ell(t_k+)\, t, \quad \forall 0 \leq \ell \leq L. \tag{D.3}$$

It follows from Corollary 2.3 that

$$\widetilde{S}_{t_k+} = \mathbb{1}_{\{\beta_{\ell'_k}(t_k)=1;\, \ell'_k\delta<3\}}\, S(y; x) + \mathbb{1}_{\{\beta_{\ell'_k}(t_k)<1 \text{ or } \ell'_k\delta>3\}}\, \widehat{S}(y; x).$$

and $\Sigma^0_\ell(t_k+)$ is defined from $\Sigma_\ell(t_k+)$ as in (D.1), and $\Sigma_\ell(t_k+)$ is defined using (2.9) replacing $\widetilde{S}_{t_k}$ by $\widetilde{S}_{t_k+}$.

Note that the slope in (D.3) only depends on the vector $\beta(t_k)$. Hence these formula make possible exact numerical simulation of $\beta(t)$. In the definition (D.2) of t_{k+1},

1. the first infimum corresponds to the first time where another component of $\beta(t)$ intersects $\beta_{\ell'_k}(t)$ (change of resident or dominant population),
2. the second one to the first time where a component of $\beta(t)$ hits 0 (extinction of a subpopulation),

3. the third one to the first time where immigration due to mutations becomes dominant in the dynamics of one coordinate of $\beta(t)$,

4. the fourth one to the first time where $\beta_{\ell'_k}(t)$ hits 1 when $\beta_{\ell'_k}(t_k) < 1$ (transition from dominant to resident for trait $\ell'_k\delta$).

The smallest of the four infima in (D.2) gives the nature of the next event.

The third infimum involves only traits $\ell\delta$ whose initial dynamics (after time t_k) is not driven by mutations, i.e. such that $\beta_{\ell-1}(t_k) < \beta_\ell(t_k) + \alpha$. In this case,

$$\Sigma^0_\ell(t_k+) = \widetilde{S}_{t_k+}(\ell\delta, \ell^*_k\delta)\mathbb{1}_{\beta_\ell(t_k)>0}.$$

When $\Sigma^0_{\ell-1}(t_k+) > \widetilde{S}_{t_k+}(\ell\delta, \ell^*_k\delta)\mathbb{1}_{\beta_\ell(t_k)>0}$, the time where the curve $\beta_\ell(t_k)+t\,\widetilde{S}_{t_k+}(\ell\delta, \ell^*_k\delta)\mathbb{1}_{\beta_\ell(t_k)>0}$ crosses $\beta_{\ell-1}(t_k) + t\,\Sigma^0_{\ell-1}(t_k+) - \alpha$ is then given by $\frac{\beta_\ell(t_k)-\beta_{\ell-1}(t_k)+\alpha}{\Sigma^0_{\ell-1}(t_k+)-\widetilde{S}_{t_k+}(\ell\delta,\ell^*_k\delta)\mathbb{1}_{\beta_\ell(t_k)>0}}$.

Funding

N.C., S.M. and V.C.T. have been supported by the Chair "Modélisation Mathématique et Biodiversité" of Veolia Environnement-Ecole Polytechnique-Museum National d'Histoire Naturelle-Fondation X. V.C.T. also acknowledges support from ANR CADENCE (ANR-16-CE32-0007).

References

[1] S. Billiard, P. Collet, R. Ferrière, S. Méléard, V.C. Tran. The effect of competition and horizontal trait inheritance on invasion, fixation and polymorphism. *Journal of Theoretical Biology*, 411:48–58, 2016.

[2] S. Billiard, P. Collet, R. Ferrière, S. Méléard, V.C. Tran. Stochastic dynamics for adaptation and evolution of microorganisms. *Journal of the European Mathematical Society*, pages 527–552, 2018. special issue for the Proceedings ECM2016.

[3] A. Bovier, L. Coquille, C. Smadi. Crossing a fitness valley as a metastable transition in a stochastic population model. *ArXiv e-print 1801.06473*, 2018.

[4] N. Champagnat. A microscopic interpretation for adaptative dynamics trait substitution sequence models. *Stochastic Processes and their Applications*, 116:1127–1160, 2006.

[5] N. Champagnat, P.-E. Jabin, S. Méléard. Adaptive dynamics in a stochastic multi-resources chemostat model. *Journal de Mathématiques Pures et Appliquées*, 101(6):755–788, 2014.

[6] N. Champagnat and S. Méléard. Polymorphic evolution sequence and evolutionary branching. *Probability Theory and Related Fields*, 151(1-2):45–94, 2011.

[7] C. Dellacherie and P.-A. Meyer. *Probabilités et potentiel.* Collection Enseignement des Sciences. Hermann, 1975.

[8] R. Durrett and J. Mayberry. Travelling waves of selective sweeps. *Annals of Applied Probability*, 21(2):699–744, 2011.

[9] S.N. Ethier and T.G. Kurtz. *Markov Processus, Characterization and Convergence.* John Wiley & Sons, New York, 1986.

[10] R. Fernandez Lopez. private communication. 2014.

[11] M. Getino, D. Sanabria-Rios, R. Fernandez-Lopez, J. Sanchez-Lopez, A. Fernandez, N. Carballeira, F. de la Cruza. Synthetic fatty acids prevent plasmid-mediated horizontal gene transfer. mBio 6,e01032–15, 2015.

[12] S.M. Ginty, L. Lehmann, S. Brown, D. Rankin.The interplay between relatedness and horizontal gene transfer drives the evolution of plasmid-carried public goods. Proceedings R. Soc.B 208 0400, 2013.

[13] P. Keeling, J. Palmer. Horizontal gene transfer in eukaryotic evolutio. Nat. Rev. Genet. 9, 605–618, 2008.

[14] S. Méléard. *Modèles aléatoires en Ecologie et Evolution*. Springer, 2016.

[15] H. Ochman, J. Lawrence, E. Groisman. Lateral gene transfer and the nature of bacterial innovatio. Nature 405, 299–304, 2000.

[16] F. Stewart, B. Levin. The population biology of bacterial plasmids: a priori conditions for the existence of conjugationally transmitted factors. Genetics 87, 209–228, 1977.