Weierstraß-Institut
für Angewandte Analysis und Stochastik

im Forschungsverbund Berlin e.V.

Preprint ISSN 0946 – 8633

# Discrete logistic branching populations and the canonical diffusion of adaptive dynamics

Nicolas Champagnat[1], Amaury Lambert[2]

submitted: 23. Januar 2006

[1] Weierstraß-Institut
für Angewandte Analysis
und Stochastik
Mohrenstraße 39
10117 Berlin Germany
E-Mail: champagn@wias-berlin.de

[2] Unit of Mathematical Evolutionary Biology
UMR 7625 Laboratoire d'Écologie
École Normale Supérieure
46 rue d'Ulm
75230 Paris Cedex 05, France

UMR 7625 Laboratoire d'Écologie
Université Pierre et Marie Curie-Paris 6
7 quai Saint Bernard
75252 Paris Cedex 05, France
E-mail: amaury.lambert@ens.fr

No. 1095
Berlin 2006

---

2000 *Mathematics Subject Classification.* 60J80, 60J25, 60J70, 60J75, 60J85, 60K35, 92D10, 92D15, 92D25, 92D40.

*Key words and phrases.* Logistic branching process, population dynamics, density-dependence, competition, fixation probability, genetic drift, weak selection, adaptive dynamics, invasion fitness, timescale separation, trait substitution sequence, diffusion approximation; harmonic equations, convergence of measure-valued processes.

Research supported part by the DFG in the Dutch-German Bilateral Rersearch Group "Mathematics of Random Spatial Models from Physics and Biology".

Edited by
Weierstraß-Institut für Angewandte Analysis und Stochastik (WIAS)
Mohrenstraße 39
10117 Berlin
Germany

Fax: + 49 30 2044975
E-Mail: preprint@wias-berlin.de
World Wide Web: http://www.wias-berlin.de/

## Abstract

The biological theory of adaptive dynamics proposes a description of the long-time evolution of an asexual population, based on the assumptions of large population, rare mutations and small mutation steps, that lead to a deterministic ODE, called 'canonical equation of adaptive dynamics'. However, in order to include the effect of genetic drift in this description, we have to apply a limit of weak selection to a finite stochastically fluctuating discrete population subject to competition in the logistic branching fashion. We start with the study of the particular case of two competing subpopulations (resident and mutant) and seek explicit first-order formulae for the probability of fixation of the mutant, also interpreted as the mutant's fitness, in the vicinity of neutrality. In particular, the first-order term is a linear combination of products of functions of the initial mutant frequency times functions of the initial total population size, called invasibility coefficients (fertility, defence, aggressiveness, isolation, survival). Then we apply a limit of rare mutations to a population subject to mutation, birth and competition where the number of coexisting types may fluctuate, while keeping the population size finite. This leads to a jump process, the so-called 'trait substitution sequence', where evolution proceeds by successive invasions and fixations of mutant types. Finally, we apply a limit of weak selection (small mutation steps) to this jump process, that leads to a diffusion process of evolution, called 'canonical diffusion of adaptive dynamics', in which genetic drift is combined with directional selection driven by the fitness gradient.

# 1 Introduction

The three main forces in evolutionary biology are mutation, selection and drift. Mutations allow the number of coexisting types in a structured population to increase; on the other hand, *Darwinian selection* eliminates the deleterious types, thus 'fixing' the most beneficial one; however, in finite populations (but they all are in the real world), deleterious types can sometimes be fixed by chance, which is then called *genetic drift*. When selection is *weak* (types differ only slightly), its action can be counteracted by genetic drift, whereas strong selection triggers almost certain fixation of beneficial types, known as *selective sweep*.

The field where these three forces are studied with taking into account the complexity of the genetic substrate is called *population genetics* [6, 21, 15, 12]. The (recent) field where they are studied with taking into account the structure of ecological interactions is called *adaptive dynamics* [17, 26, 27]. The link between both is still unclear, but see [22, 30].

One of the fundamental models of adaptive dynamics is the 'trait substitution sequence' (TSS) [28, 3], which is based on the following biological heuristics. Sufficient time is given to selection to eliminate unlucky types between two mutations, so that, on the mutation timescale, only one type survives at a time. In this model, evolution proceeds by successive invasions of mutant types replacing the resident one (selective sweeps [9, 10]), and can be described as a stochastic jump process over the space of types, also called *traits* (e.g. size, age at maturity or rate of food intake).

The TSS has revealed a powerful tool for understanding various evolutionary phenomena, such as evolutionary branching (evolution from a monomorphic population to a polymorphic one [28] that may lead to speciation [7]) and is the basis for other biological models, such as the 'canonical equation of adaptive dynamics' [8, 4]. The TSS is based on a combination of a limit of *rare mutations* and a limit of *large populations*, leading to a timescale separation between the mutation events and the birth and death events [3].

Most mathematical models for population dynamics used in the study of evolution assume either constant population sizes (population genetics) or infinite populations (adaptive dynamics). This has two shortcomings: in population genetics, fluctuations of population sizes are neglected, which is rather unrealistic, especially as new types get fixed; in adaptive dynamics, (even slightly) deleterious types are never fixed, impeding the modelling of genetic drift. As a consequence, the direction of evolution is partly deterministic, and a small jump renormalization of the TSS model [8, 4] leads to a totally deterministic ODE, the one known as the canonical equation of adaptive dynamics.

Here, we want to use the bottom-up approach of adaptive dynamics, that is, model (macroscopic) evolution from (microscopic) populations [5], but we also want to keep the two main ingredients of genetic drift, that is, *population finiteness* and *weak selection* [20] and to allow the population size to fluctuate randomly through time. Thus, we consider the minimal model of population dynamics featuring stochastic birth and death as well as regulatory competition, that is, the (multitype) logistic branching process [23]. More precisely, we consider a finite population with $n$ types, where individuals of type $j$ give birth at constant rate $b_j$, die a constant rate $d_j$, and kill each individual of type $i$ at constant rate $c_{ij}$. This quadratic density-dependence prevents the process from going to infinity. Then, the first time $T$ when the population is monotype is finite a.s., and the remaining type is then said to have undergone *fixation*. If in addition all natural death rates vanish ($d_i = 0$ for any $i$), then the logistic branching process never dies out, and we are left with a positive-recurrent process with a nice stationary probability (see next section).

We start with the particular case of two interacting types, one resident (wild type) and one mutant, and study the behaviour of the fixation probability of a mutant type in a resident population when its dynamical characteristics deviate slightly from those of the wild type (weak selection, see also [24]). In Section 2, this model and some of its properties are displayed, other quantities are defined and a rather precise outline of the paper is given. Fixation probabilities are characterized in Section 3. In Section 4, we seek for explicit first-order formulae for the fixation probability of the mutant in the vicinity of neutrality. Specifically, a pure-resident population is characterized by the dynamical parameters $(b_1, c_{11}, d_1)$, and when the mutant type is present, we can write $b_2 = b_1 + \lambda$, $d_2 = d_1 - \sigma$, $c_{12} = c_{11} + \alpha - \varepsilon$, $c_{21} = c_{11} - \delta - \varepsilon$ and $c_{22} = c_{11} - \delta + \alpha$, where $\lambda, \delta, \alpha, \varepsilon, \sigma$ are the selection coefficients of the mutant respectively associated to fertility, defence, aggressiveness, isolation and survival. Then, we prove that each partial derivative of the fixation probability w.r.t. any selection coefficient factorizes as a function of the initial mutant frequency $p$ (either $p(1-p)$ or $p(1-p)(1-2p)$) times a function of the initial total population size, called an invasibility coefficient of the resident.

In Section 5, we consider the multitype logistic branching process described above and add mutations to the model. Each individual, upon giving birth, begets a mutant daughter with a (small) probability, whose type is chosen according to some mutation kernel. First, our goal is to apply to this process a limit of rare mutations and to describe the evolutionary process on the mutation timescale, in order to obtain the equivalent of the TSS in this finite population setting. In particular, this jump process over the trait space allows evolution in any direction. Then, we apply to this process a limit of small mutation steps. Because of the population finiteness, genetic drift carries over to the limit, which gives rise to a diffusion process over the trait space that we call the 'canonical diffusion of adaptive dynamics'.

These results are proved in Section 6.

## 2 Model, outline and preliminary results

From now on, $\mathbb{N}^\star$ denotes the set of positive integers, and $\mathbb{N}$ denotes $\mathbb{N}^\star \cup \{0\}$.

In this section, we consider a two-type discrete population undergoing binary birth-death-competition events, in the logistic branching fashion [23], where the first type (1) is called **resident** (or wild type) and the second type (2) **mutant**.

For $i, j \in \{1, 2\}$, $b_i$ (resp. $d_i$) is the birth (resp. death) rate of type $i$, and $c_{ij}$ is the competition rate felt **by** an individual of type $i$ **from** an individual of type $j$. More precisely, individuals of type $j$ give birth independently to a single individual of type $j$ at rate $b_j$, die at rate $d_j$, or kill any other fixed individual of type $i$ at rate $c_{ij}$.

More formally, we consider a bivariate integer-valued continuous-time Markov process $(X_t, Y_t; t \geq 0)$ with rate matrix $Q = (q_{kl}; k \in \mathbb{N}^2, l \in \mathbb{N}^2)$, where

$$q_{kl} = \begin{cases} b_1 n & \text{if} \quad k = (n,m) \quad \text{and} \quad l = (n+1, m) \\ b_2 m & \text{if} \quad k = (n,m) \quad \text{and} \quad l = (n, m+1) \\ c_{11} n(n-1) + c_{12} nm + d_1 n & \text{if} \quad k = (n,m) \quad \text{and} \quad l = (n-1, m) \\ c_{21} mn + c_{22} m(m-1) + d_2 m & \text{if} \quad k = (n,m) \quad \text{and} \quad l = (n, m-1) \\ -r_{nm} & \text{if} \quad k = (n,m) \quad \text{and} \quad l = (n,m) \\ 0 & \qquad \text{otherwise,} \end{cases}$$

and where the jumping rate $r_{nm}$ is

$$r_{nm} = n(b_1 + c_{11}(n-1) + c_{12} m + d_1) + m(b_2 + c_{21} n + c_{22}(m-1) + d_2). \tag{1}$$

The law of this process conditioned on fixed initial state $(n, m)$ will be denoted by $\mathbb{P}_{n,m}$. Let $B$ denote the birth vector, $C$ the competition matrix, and $D$ the death vector

$$B = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}, \qquad C = \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix}, \qquad D = \begin{pmatrix} d_1 \\ d_2 \end{pmatrix}.$$

This definition of the two-type logistic process with dynamical characteristics $(B, C, D)$ could readily be extended to $n$-type populations, by considering a $n$-dimensional birth (resp. death) vector and a $n \times n$-dimensional competition matrix. However, in this paper, our ultimate goal is to model the sequential arrivals of mutants and their subsequent extinction or fixation, on different time scales, so that the simultaneous occurrence of three or more distinct morphs will be shown to have negligible probability.

Notice that $(X,Y)$ is dominated by $(Z_1,Z_2)$, where $Z_i$, $i=1,2$, is a scalar logistic branching process with dynamical characteristics $(b_i, c_{ii}, d_i)$, and $Z_1$, $Z_2$ are independent. Since we will always assume that $c_{11}c_{22}\neq 0$, the total population size $X+Y$ cannot go to infinity [23].

As a remark, we point out that we only consider binary splittings for the sake of simplicity and the interest for explicit formulae. Apart from explicit formulae, the last statement and all others in the present paper remain true even if assuming that, when giving birth, an individual of any of the two types can give birth to a random number of offspring, where their (common) distribution $Z$ merely has to satisfy $\mathbb{E}(\log(Z))<\infty$.

Now the union of the axes

$$\Omega_1 := \mathbb{N}\times\{0\} \text{ and } \Omega_2 := \{0\}\times\mathbb{N}$$

is accessible, except in the case when $d_1=d_2=c_{12}=c_{21}=0$, and absorbing. Since we will always assume below that $d_1+d_2+c_{12}+c_{21}\neq 0$, $\mathbb{N}^\star\times\mathbb{N}^\star$ is transient, and

$$\mathbb{P}(T<\infty)=1,$$

where

$$T := T_{\Omega_1}\wedge T_{\Omega_2},$$

and for any subset $\Gamma$ of $\mathbb{N}^2$, $T_\Gamma$ denotes the first hitting time of $\Gamma$ by $(X,Y)$. Also notice that the origin is not accessible from $\mathbb{N}^\star\times\mathbb{N}^\star$, so that for any $(n,m)\neq(0,0)$, $\mathbb{P}_{n,m}(T_{\Omega_1}=T_{\Omega_2})=0$.

Then we call **fixation** (of the mutant) the event $\{T_{\Omega_2}<T_{\Omega_1}\}$, and **extinction** (of the mutant) the event $\{T_{\Omega_1}<T_{\Omega_2}\}$.

The probability of fixation will be denoted by $u$, and the initial frequency of the mutant by $p$

$$u := \mathbb{P}(T_{\Omega_2}<T_{\Omega_1}),\qquad p := \frac{m}{m+n}.$$

If $d_2=0$, then $(0,0)$ is not accessible from $\Omega_2$, and a straightforward application of the strong Markov property at $T_{\Omega_2}$ and Theorem 2.2 in [23] show that, on the event $\{T_{\Omega_2}<T_{\Omega_1}\}$, $Y$ is positive-recurrent and converges in distribution to $\xi$, where $\xi$ is a Poisson variable of parameter $\theta := b_2/c_{22}$ conditioned on being nonzero

$$\mathbb{P}(\xi=i)=\frac{\mathrm{e}^{-\theta}}{1-\mathrm{e}^{-\theta}}\frac{\theta^i}{i!}\qquad i\geq 1.$$

Notice that $\mathbb{E}(\xi)=\theta/(1-\exp(-\theta))$, so that $\theta$ can roughly be taken as the mean population size ('carrying capacity') of the stationary logistic branching process.

Since a similar statement for $X$ holds on the extinction of the mutant if $d_1=0$, a nice feature of the case when $d_1$ and $d_2$ both equal 0 is that $(X,Y)$ converges in distribution towards $(0,\xi^{(2)})$ with probability $u$ and towards $(\xi^{(1)},0)$ with probability $1-u$, where $\xi^{(i)}$ is a Poisson variable of parameter $b_i/c_{ii}$ conditioned on being nonzero.

We refer to **selective neutrality** as the case of *exchangeability* between individuals of both types, that is,

$$B=b\mathbf{1},\qquad C=c\mathbf{1},\qquad D=d\mathbf{1},$$

where **1** is a matrix with all components equal to 1 and dimensions *ad hoc.* Assuming that the presence of a mutant form does not modify the dynamical characteristics $(b_1, c_{11}, d_1)$ of the resident, we may focus on deviations from the neutral case so as to express the two-type characteristics as

$$B = b\mathbf{1} + \begin{pmatrix} 0 \\ \lambda \end{pmatrix}, \qquad C = c\mathbf{1} - \begin{pmatrix} 0 & 0 \\ \delta & \delta \end{pmatrix} + \begin{pmatrix} 0 & \alpha \\ 0 & \alpha \end{pmatrix} - \begin{pmatrix} 0 & \varepsilon \\ \varepsilon & 0 \end{pmatrix}, \qquad D = d\mathbf{1} - \begin{pmatrix} 0 \\ \sigma \end{pmatrix}.$$

In words, deviations from the neutral case are a linear combination of five **fundamental** (additive) **selection coefficients** $\lambda$, $\delta$, $\alpha$, $\varepsilon$, $\sigma$, that are chosen to be positive when they confer an advantage to the mutant. In the sequel, we will see that it is indeed convenient to assess deviations to the neutral case with the help of selection coefficients in terms of

1. **fertility** ($\lambda$, as the usual letter standing for growth rate in discrete-time deterministic models) : positive $\lambda$ means increased mutant birth rate
2. **defence** capacity ($\delta$, as in defence) : positive $\delta$ means reduced competition sensitivity of mutant individuals w.r.t. the total population size
3. **aggressiveness** ($\alpha$, as in aggressive, or attack) : positive $\alpha$ means raised competition pressure exerted from any mutant individual onto the rest of the population
4. **isolation** ($\varepsilon$, as in exclusion) : positive $\varepsilon$ means lighter cross-competition between different morphs, that would lead, if harsher, to the exclusion of the less abundant one
5. **survival** ($\sigma$, as in survival) : positive $\sigma$ means reduced mutant death rate.

**Under neutrality**, an elementary martingale argument shows that the fixation probability equals the initial mutant frequency, that is,

$$u = p.$$

The first goal of this paper is to unveil the dependence of $u$ upon $\lambda$, $\delta$, $\alpha$, $\varepsilon$, $\sigma$, when they deviate from 0.

In Section 3, we study the fixation probability $u$ as a function of the parameters of the model and the initial condition $(n, m)$. We prove that $u$ is differentiable with respect to the components of $B$, $C$ and $D$, and that $(n+m)^{-1}\nabla u_{n,m}$, is bounded (for each given triple $B$, $C$, $D$). The Kolmogorov forward equations translate into a discrete harmonic equation of the type $\Delta u(n, m) = 0$ outside $\Omega_1 \cup \Omega_2$, with boundary condition 0 on $\Omega_1$, and 1 on $\Omega_2$, where $\Delta$ is a linear operator on doubly indexed sequences. Then, a uniqueness result allows us to characterize the partial derivatives of $u$ (in the dynamical parameters) in the vicinity of neutrality, as the only sublinear solutions $v$ of harmonic equations of the type $\Delta v(n, m) = f(n, m)$ outside $\Omega_1 \cup \Omega_2$, where $f$ is sublinear, and has zero boundary condition on $\Omega_1 \cup \Omega_2$.
These derivatives are interesting to study, as they provide insight as to how does the fixation probability deviate from $p$ as the selection coefficients of the mutant deviate from 0. The vicinity of neutrality is called *weak selection*, as opposed to strong selection, in which case fixation events are called selective sweeps (rapid fixation). In Section 4, entitled 'Weak selection', we thus write

$$u = p + \mathbf{v}'.\mathbf{s} + o(\mathbf{s}), \tag{2}$$

where $\mathbf{s}$ is the multidimensional selection coefficient (fertility, defence, aggressiveness, isolation, survival) and $\mathbf{v}$ is the corresponding gradient of the fixation probability

$$\mathbf{s} := \begin{pmatrix} \lambda \\ \delta \\ \alpha \\ \varepsilon \\ \sigma \end{pmatrix}, \qquad \mathbf{v} := \begin{pmatrix} v^\lambda \\ v^\delta \\ v^\alpha \\ v^\varepsilon \\ v^\sigma \end{pmatrix}.$$

In particular, $\mathbf{v}$ only depends on $b, c, d$, and $n, m$. It is then proved that for $\iota = \lambda, \alpha, \delta, \sigma$, the gradient coefficient $v^\iota$ is of the form

$$v^\iota = \frac{nm}{(n+m)^2}\, g^\iota_{n+m} = p\,(1-p)\; g^\iota_{n+m},$$

and that $v_\varepsilon$ is of the form

$$v^\varepsilon = \frac{nm(n-m)}{(n+m)^3}\, g^\varepsilon_{n+m} = p\,(1-p)\,(1-2p)\; g^\varepsilon_{n+m},$$

where the $g$'s only depend on the resident's characteristics $b, c, d$, and on the **total initial population size** $n+m$. For that reason, and because they are multipliers of the mutant's selection coefficients, we call them the **invasibility coefficients** of the resident, as in [24]. The remainder of Section 4 is devoted to the study of the invasibility coefficients.

Finally, Section 5 implements these results in models of adaptive dynamics where mutants appear on large timescales and are either fixed or eliminated on shorter timescales, until the next mutant arises. This allows to follow the value of the resident trait $(b, c, d)$ as evolutionary time goes by in two fashions. First, we consider the process on the mutation timescale in the limit of rare mutations (Subsection 5.2). On this new timescale, if $d$ was positive, extinction would occur with high probability before any mutation event, so we will assume that populations have zero natural death rate ($d = 0$). The limiting process is a Markov jump process over the trait space, the so-called *trait substitution sequence* (TSS). Second, we apply a limit of small jumps to the TSS by rescaling the mutation kernel and time accordingly (Subsection 5.3). This yields a diffusion process over the trait space, where the deterministic term accounts for *directional selection* driven by the gradient of the fixation probability, and the diffusion term accounts for *genetic drift*. In reference to the so-called 'canonical equation of adaptive dynamics', we give it the name of *canonical diffusion of adaptive dynamics.*

We end this section by stating and proving a technical proposition.

**Proposition 2.1** *The bivariate logistic branching process* $(X_t, Y_t; t \geq 0)$ *satisfies the following properties*

(a) *If, for some* $p \geq 1$, $\mathbb{E}(X_0^p + Y_0^p) < +\infty$, *then* $\sup_{t\geq 0} \mathbb{E}(X_t^p + Y_t^p) < +\infty$

(b) *Let* $(\hat{X}_n, \hat{Y}_n; n \in \mathbb{N})$ *denote the discrete-time Markov chain associated with* $(X_t, Y_t; t \geq 0)$*, and* $\hat{T}$ *denote the first hitting time of* $\Omega_1 \cup \Omega_2$ *by* $(\hat{X}, \hat{Y})$*. There is some positive* $a$ *such that*

$$\mathbb{E}_{n,m}(\hat{T}) < a(n+m) \quad \textit{and} \quad \mathbb{E}_{n,m}(\hat{T}^2) < a(n+m)^2. \tag{3}$$

*Moreover, for fixed* $\bar{b}, \underline{c}, \underline{d} > 0$*, (3) holds for some* $a$ *that can be chosen uniformly for all parameters such that* $b_1 \vee b_2 \leq \bar{b}$*,* $c_{11} \wedge c_{22} \geq \underline{c}$ *and* $\max(d_1, d_2, c_{12}, c_{21}) \geq \underline{d}$*.*

(c) *For any* $p \in [1,2]$ *and* $n, m \in \mathbb{N}$, $\mathbb{E}_{n,m}(X_T^p + Y_T^p) < K$, *where the constant* $K$ *can be chosen uniformly for all parameters such that* $b_1 \vee b_2 \leq \bar{b}$, $c_{11} \wedge c_{22} \geq \underline{c}$ *and* $\max(d_1, d_2, c_{12}, c_{21}) \geq \underline{d}$ *for fixed* $\bar{b}, \underline{c}, \underline{d} > 0$.

**Proof.** **(a)** Since $(X, Y)$ is dominated by $(Z, Z')$, where $Z$ and $Z'$ are independent scalar logistic branching processes with common dynamical characteristics $(b, c, 0)$ where $b := b_1 \vee b_2$ and $c := c_{11} \wedge c_{22}$, it suffices to show that $\sup_{t \geq 0} \mathbb{E}(Z_t^p) < +\infty$.

Let us define $p_t^k = \mathbb{P}(Z_t = k)$. The backward Kolmogorov equation reads

$$\begin{aligned}\frac{d}{dt}\mathbb{E}(Z_t^p) &= \sum_{k \geq 1} k^p \frac{dp_t^k}{dt} = \sum_{k \geq 1} k^p [b(k-1)p_t^{k-1} + c(k+1)kp_t^{k+1} - k(b + c(k-1))p_t^k] \\ &= \sum_{k \geq 1} \left[ b\left(\left(1 + \frac{1}{k}\right)^p - 1\right) + c(k-1)\left(\left(1 - \frac{1}{k}\right)^p - 1\right)\right] k^{p+1} p_t^k.\end{aligned}$$

Now, for any $k > k_0$, where $k_0 := [2b/c] + 1$, $c(k-1) \geq 2b$. Therefore, for $k > k_0$,

$$b((1 + 1/k)^p - 1) + c(k-1)((1 - 1/k)^p - 1) \leq -b[3 - 2(1 - 1/k)^p - (1 + 1/k)^p]$$

which is equivalent to $-bp/k$. Then, enlarging $k_0$ if necessary, we obtain

$$\begin{aligned}\frac{d}{dt}\mathbb{E}(Z_t^p) &\leq \sum_{k=1}^{k_0} b(2^p - 1)k_0^p - \sum_{k \geq k_0 + 1} \frac{bp}{2} k^p p_t^k \\ &\leq K - \frac{bp}{2}\mathbb{E}(Z_t^p),\end{aligned}$$

where the constant $K$ depends solely on $k_0$. This differential inequality yields

$$\mathbb{E}(Z_t^p) \leq \frac{2K}{bp} + \left(\mathbb{E}(Z_0^p) - \frac{2K}{bp}\right) e^{-bpt/2},$$

which gives the required uniform bound.

**(b)** Let $\Pi = (\pi_{ij})_{i,j \in \mathbb{N}^2}$ denote the transition matrix of $(\hat{X}, \hat{Y})$:

$$\pi_{ij} = \begin{cases} b_1 n / r_{nm} & \text{if } i = (n,m) \text{ and } j = (n+1, m) \\ b_2 m / r_{nm} & \text{if } i = (n,m) \text{ and } j = (n, m+1) \\ \dfrac{n(c_{11}(n-1) + c_{12}m + d_1)}{r_{nm}} & \text{if } i = (n,m) \text{ and } j = (n-1, m) \\ \dfrac{m(c_{21}n + c_{22}(m-1) + d_2)}{r_{nm}} & \text{if } i = (n,m) \text{ and } j = (n, m-1) \\ 0 & \text{otherwise,} \end{cases} \tag{4}$$

where $r_{nm}$ has been defined in (1).

Since $d_1 + d_2 + c_{12} + c_{21} \neq 0$, it follows from the facts that $n(n-1) + m(m-1) \geq (n+m)^2/2 - (n+m)$ for any $n, m \geq 0$ and $\pi_{(2,0),(1,0)}, \pi_{(1,1),(0,1)}, \pi_{(1,1),(1,0)}, \pi_{(0,2),(0,1)} \geq d/(2b+d)$ where $d := \max(d_1, d_2, c_{12}, c_{21}) \wedge c$, that the process $(\hat{X}_n + \hat{Y}_n; n \in \mathbb{N})$ is dominated by the Markov chain $(\hat{Z}_n; n \in \mathbb{N})$ in $\mathbb{N}^*$ with initial state $k = X_0 + Y_0$ and transition probabilities

$$p_{ij} = \begin{cases} b/[b + c(i/2 - 1)] & \text{if } i \geq 3 \text{ and } j = i + 1 \\ c(i/2 - 1)/[b + c(i/2 - 1)] & \text{if } i \geq 3 \text{ and } j = i - 1 \\ 2b/(2b + d) & \text{if } i = 2 \text{ and } j = 3 \\ d/(2b + d) & \text{if } i = 2 \text{ and } j = 1 \\ 0 & \text{otherwise.} \end{cases}$$

Let us denote by $P_k$ its law. Therefore, $\hat{T}$ is dominated by $\hat{S} := \inf\{n \geq 0, \hat{Z}_n = 1\}$ and it suffices to prove that $E_k(\hat{S}) \leq ak$ and $E_k(\hat{S}^2) \leq ak^2$ for some $a > 0$.

Let $(\tilde{U}_n; n \geq 0)$ be the discrete-time random walk on $\mathbb{Z}$ with right transition probability $1/3$ and left transition probability $2/3$. The law of $\tilde{U}$ conditional on $\tilde{U}_0 = k$ is denoted by $\tilde{P}_k$. Let $\tau$ be the first hitting time of 0 by $\tilde{U}$. For any $k \geq 0$, one can compute explicitly (see e.g. Norris [29]) that

$$\tilde{E}_k(\tau) = 3k \qquad \text{and} \qquad \tilde{E}_k(\tau^2) = 3k(3k+8).$$

Now, let $k_0$ be large enough to have $c(k_0/2 - 1) > 2b$. Then observe that any excursion of $\hat{Z}$ above $k_0 + 1$ is stochastically dominated by an excursion of the random walk $\tilde{U}$ above $k_0 + 1$. On the one hand, it is a simple exercise to check that this domination entails $E_{k_0}(\hat{S}^2) < +\infty$ (and $E_{k_0}(\hat{S}) < +\infty$). On the other hand, this domination also implies that for any $k \geq k_0$,

$$E_k(\hat{S}) \leq E_{k_0}(\hat{S}) + \tilde{E}_{k-k_0}(\tau) = E_{k_0}(\hat{S}) + 3(k - k_0).$$

and

$$E_k(\hat{S}^2) \leq 2E_{k_0}(\hat{S}^2) + 2\tilde{E}_{k-k_0}(\tau^2) = 2E_{k_0}(\hat{S}^2) + 3(k-k_0)(3(k-k_0)+8),$$

which ends the proof of (3).

**(c)** With the same notation as above, since $\hat{T}$ is the number of jumps of the process $(X, Y)$ that occurred on the time interval $[0, T]$, $X_T \leq X_0 + \hat{T}$ and $Y_T \leq Y_0 + \hat{T}$. Hence $\mathbb{E}_{n,m}(X_T^p + Y_T^p) \leq 2\mathbb{E}_{n,m}((n + m + \hat{T})^p)$ which is finite by (3). □

## 3 Fixation probability

### 3.1 A discrete harmonic equation

Recall that the fixation probability is the probability that the bivariate logistic branching process hits axis $\Omega_1$ before axis $\Omega_2$. Here, we characterize it thanks to a discrete harmonic equation (corresponding to Kolmogorov forward equations).

**Proposition 3.1** *The fixation probability $u_{n,m}$ is the unique bounded solution to*

$$\begin{cases} (\Delta u)_{n,m} = 0 & \text{for } (n,m) \notin \Omega_1 \cup \Omega_2 \\ u_{n,m} = 0 & \text{for } (n,m) \in \Omega_1 \\ u_{n,m} = 1 & \text{for } (n,m) \in \Omega_2, \end{cases} \tag{5}$$

*where $\Delta$ is the harmonic (its coefficients sum to zero) operator defined for any doubly indexed sequence $w$ as*

$$\begin{aligned}(\Delta w)_{n,m} = {} & \big[b_1 n + b_2 m + n\big(c_{11}(n-1) + c_{12} m + d_1\big) + m\big(c_{21} n + c_{22}(m-1) + d_2\big)\big] w_{n,m} \\ & - b_1 n w_{n+1,m} - b_2 m w_{n,m+1} - n\big(c_{11}(n-1) + c_{12} m + d_1\big) w_{n-1,m} \\ & - m\big(c_{21} n + c_{22}(m-1) + d_2\big) w_{n,m-1}. \end{aligned} \tag{6}$$

Note that in the previously displayed equation, whenever a term is not defined, the multiplying coefficient is zero.

The fact that $u_{n,m}$ satisfies (5) follows from the Markov property at the first jump time of $(X,Y)$, and the uniqueness relies on Lemma 3.2 below.

**Lemma 3.2** *Fix* $p \geq 1$ *and consider a subset* $\Gamma$ *of* $\mathbb{N}^2$ *such that* $T_\Gamma < +\infty$ $\mathbb{P}_{n,m}$*-a.s. for any* $n, m \geq 0$. *Then, for any function* $f : \Gamma \to \mathbb{R}$ *such that,* $|f(n,m)|/(n+m+1)$ *is bounded on* $\Gamma$, *the equation*

$$\begin{cases} (\Delta h)_{n,m} = 0 & \text{for } (n,m) \notin \Gamma \\ h(n,m) = f(n,m) & \text{for } (n,m) \in \Gamma \end{cases} \tag{7}$$

*admits at most one solution* $h$ *such that* $|h(n,m)|/(n+m+1)$ *is bounded.*

**Proof.** It suffices to prove that (7) with $f \equiv 0$ admit $h \equiv 0$ as unique sublinear solution. Let $h$ be such a function and fix $n, m \geq 0$. Then $(h(X_{t\wedge T_\Gamma}, Y_{t\wedge T_\Gamma}); t \geq 0)$ is a $\mathbb{P}_{n,m}$-semi-martingale for $t \leq T_\Gamma$. Since, by Proposition 2.1 (a), $\sup_{t\geq 0} \mathbb{E}_{n,m}(X_t + Y_t)) < +\infty$, $(h(X_t, Y_t))_{t\geq 0}$ is actually a uniformly integrable martingale. Applying the stopping theorem at time $T_\Gamma$, we get

$$0 = \mathbb{E}_{n,m}(h(X_{T_\Gamma}, Y_{T_\Gamma})\mathbf{1}_{T_\Gamma<+\infty}) = \mathbb{E}_{n,m}(h(X_0, Y_0)) = h(n,m),$$

which completes the proof. □

## 3.2 Differentiability of the fixation probability

Here, we prove the existence of the partial derivatives of the fixation probability with respect to the components of $B$, $C$ and $D$, and show that these derivatives are always sublinear in the initial condition. We also give some bounds for the higher-order derivatives of the fixation probability, which will be useful in Section 5.3.

**Theorem 3.3** **(a)** *The fixation probability* $u_{n,m}$, *as a function of* $b_i, c_{ij}, d_i$ $(i, j = 1, 2)$, *is differentiable on the set* $\{c_{11}c_{22} \neq 0, c_{12} + c_{21} + d_1 + d_2 \neq 0\}$, *and its differential* $v_{n,m}$ *satisfies*

$$\sup_{(n,m)\in\mathbb{N}^*\times\mathbb{N}^*} \frac{\|v_{n,m}\|}{n+m} < +\infty. \tag{8}$$

*Moreover, for any* $\bar{b}, \underline{c}, \underline{d} > 0$, *(8) holds uniformly on the set of dynamical parameters* $\{b_1 \vee b_2 \leq \bar{b},\ c_{11} \wedge c_{22} \geq \underline{c},\ \max(d_1, d_2, c_{12}, c_{21}) \geq \underline{d}\}$.

**(b)** *In addition,* $u_{n,m}$ *is* $\mathcal{C}^2$ *on* $\{c_{11}c_{22} \neq 0, c_{12} + c_{21} + d_1 + d_2 \neq 0\}$, *and its second-order derivatives are bounded by some constant times* $(n+m)^2$ *uniformly on any set of parameters of the form given above.*

**Proof.** Since $v_{n,m} \equiv 0$ on $\Omega_1 \cup \Omega_2$, we will always assume $n, m \geq 1$.

Recall the expression (4) for the transition probabilities of the discrete-time Markov chain $(\hat{X},\hat{Y})$ associated with $(X,Y)$. It is then elementary to get

$$\frac{\partial \pi_{ij}}{\partial b_1} = \begin{cases} n(1-\pi_{ij})/r_{nm} & \text{if } i=(n,m) \text{ and } j=(n+1,m) \\ -n\pi_{ij}/r_{nm} & \text{if } i=(n,m) \text{ and } j=(n,m+1),\ (n-1,m) \text{ or } (n,m-1) \\ 0 & \text{otherwise,} \end{cases} \tag{9}$$

as well as similar results for $b_2$, $d_1$, $d_2$. Now for $c_{11}$

$$\frac{\partial \pi_{ij}}{\partial c_{11}} = \begin{cases} n(n-1)(1-\pi_{ij})/r_{nm} & \text{if } i=(n,m) \text{ and } j=(n-1,m) \\ -n(n-1)\pi_{ij}/r_{nm} & \text{if } i=(n,m) \text{ and } j=(n,m+1),\ (n+1,m) \text{ or } (n,m-1) \\ 0 & \text{otherwise,} \end{cases} \tag{10}$$

and a similar result for $c_{22}$. Finally for $c_{12}$

$$\frac{\partial \pi_{ij}}{\partial c_{12}} = \begin{cases} nm(1-\pi_{ij})/r_{nm} & \text{if } i=(n,m) \text{ and } j=(n-1,m) \\ -nm\pi_{ij}/r_{nm} & \text{if } i=(n,m) \text{ and } j=(n,m+1),\ (n+1,m) \text{ or } (n,m-1) \\ 0 & \text{otherwise,} \end{cases} \tag{11}$$

along with a similar result for $c_{21}$. As a consequence, for any $i,j$ and any parameter $\beta \in \{b_h, c_{hl}, d_h : h,l=1,2\}$,

$$\left|\frac{\partial \pi_{ij}}{\partial \beta}\right| \leq \frac{\max(n^2,m^2)}{r_{nm}}\ .$$

Next notice that $r_{11} \geq c_{12}+c_{21}+d_1+d_2$ and $r_{nm} \geq c_{11}n(n-1)+c_{22}m(m-1)$, so that $r_{nm} \geq c_{11}n(n-1)$ and $r_{nm} \geq c_{22}m(m-1)$. Combine these observations to the last displayed equation, and get the existence of some $K$ such that for any $i,j$ and any $\beta$ in the set of parameters such that $c_{11} \wedge c_{22} \geq \underline{c}$ and $\max(d_1,d_2,c_{12},c_{21}) \geq \underline{d}$,

$$\left|\frac{\partial \pi_{ij}}{\partial \beta}\right| \leq K \leq 2(\underline{c} \wedge \underline{d})^{-1}.$$

We will use the notation $\pi_{i_1,\dots,i_k}$ for the product $\pi_{i_1 i_2}\pi_{i_2 i_3}\dots\pi_{i_{k-1}i_k}$ and $S_{(n,m)\to\Gamma}$ for the set of paths linking $(n,m)$ to a subset $\Gamma$ of $\mathbb{N}^2$ without hitting $\Omega_1 \cup \Omega_2$ before $\Gamma$, i.e. the set of all $k$-tuples $(i_1,i_2,\dots,i_k)$ for all $k \geq 1$ such that $i_1=(n,m)$, $i_2,\dots,i_{k-1} \in \mathbb{N}^2 \backslash (\Omega_1 \cup \Omega_2)$ and $i_k \in \Gamma$. Now,

$$u_{n,m} = \sum_{k\geq 2,(i_1,\dots,i_k)\in S_{(n,m)\to\Omega_1}} \pi_{i_1,\dots,i_k}\ ,$$

so if we prove that for any parameter $\beta$,

$$R_{n,m} := \sum_{k\geq 2,(i_1,\dots,i_k)\in S_{(n,m)\to\Omega_1}} \left|\frac{\partial \pi_{i_1,\dots,i_k}}{\partial \beta}\right|$$

is finite, we get the differentiability of $u_{n,m}$ and the inequality $|\partial u_{n,m}/\partial\beta| \leq R_{n,m}$. Observe that

$$R_{n,m} \leq K \sum_{k\geq 2,(i_1,\dots,i_k)\in S_{(n,m)\to\Omega_1}} \sum_{l=1}^{k-1} \pi_{i_1,\dots,i_l}\pi_{i_{l+1},\dots,i_k},$$

where $\pi_{i,i}=1$ by convention. Next, with $\|\cdot\|$ denoting the $L^1$-norm in $\mathbb{Z}^2$,

$$R_{n,m} \leq K \sum_{l\geq 1} \sum_{(n',m')\in(\mathbb{N}^*)^2} \sum_{(i_1,\dots,i_l)\in S_{(n,m)\to(n',m')}} \pi_{i_1,\dots,i_l} \sum_{\|\varepsilon\|=1} \sum_{k'\geq 0,(j_1,\dots,j_{k'})\in S_{(n',m')+\varepsilon\to\Omega_1}} \pi_{j_1,\dots,j_{k'}}$$

with the convention that $k' = 0$ if $(n', m') + \varepsilon \in \Omega_1$, so that

$$\begin{aligned}
R_{n,m} &\leq K \sum_{l\geq 1} \sum_{(n',m')\in(\mathbb{N}^*)^2} \sum_{(i_1,\dots,i_l)\in S_{(n,m)\to(n',m')}} \pi_{i_1,\dots,i_l} \sum_{\|\varepsilon\|=1} \mathbb{P}_{(n',m')+\varepsilon}(\hat{T}_{\Omega_1} < \hat{T}_{\Omega_2}) \\
&\leq 4K \sum_{l\geq 1} \mathbb{P}_{n,m}(\hat{T} > l) \\
&= 4K\mathbb{E}_{n,m}(\hat{T} - 1),
\end{aligned}$$

The proof of (8) is completed thanks to Proposition 2.1 (b). The uniformity with respect to the dynamical parameters follows trivially from the uniformity in Proposition 2.1 (b), and the uniform choice of $K$.

**(b)** Following exactly the same method as above, we obtain for any parameters $\beta$ and $\gamma$,

$$\begin{aligned}
&\sum_{(i_1,\dots,i_k)\in S_{(n,m)\to\Omega_1}} \left| \frac{\partial^2 \pi_{i_1,\dots,i_k}}{\partial\beta\partial\gamma} \right| \\
&\qquad \leq 4K^2 \sum_{l\geq 1} \sum_{(n',m')\in(\mathbb{N}^*)^2} \sum_{(i_1,\dots,i_l)\in S_{(n,m)\to(n',m')}} \pi_{i_1,\dots,i_l} \sum_{\|\varepsilon\|=1} \mathbb{E}_{(n',m')+\varepsilon}(\hat{T} - 1) \\
&\qquad \leq 16aK^2 \sum_{l\geq 1} \sum_{(n',m')\in(\mathbb{N}^*)^2} \sum_{(i_1,\dots,i_l)\in S_{(n,m)\to(n',m')}} \pi_{i_1,\dots,i_l}(n' + m') \\
&\qquad \leq 16aK^2 \sum_{l\geq 1} (n + m + l)\mathbb{P}_{n,m}(\hat{T} > l) \\
&\qquad \leq 16aK^2 \mathbb{E}_{n,m}\left((\hat{T} - 1)(n + m + \hat{T}/2)\right),
\end{aligned}$$

and the result again follows from Proposition 2.1 (b). □

## 4 Weak selection

### 4.1 Definition of the invasibility coefficients

From Section 2 recall the definition of selective neutrality, and that under neutrality, the fixation probability $u$ of the mutant is equal to its initial frequency $p$ in the population. Now $u$ is a function of the resident's characteristics $b, c, d$ (resp. birth, competition and death rates), the additive selection coefficients of the mutant (deviations from neutrality) $\lambda, \delta, \alpha, \varepsilon, \sigma$ (resp. fertility, defence, aggressiveness, isolation, survival), and the initial composition of the population $n, m$ (resp. numbers of residents and mutants). As stated in the previous section (Theorem 3.3), $u$ is differentiable in all non-integer parameters, so that in particular, $u$ is differentiable in $\mathbf{s} := (\lambda, \delta, \alpha, \varepsilon, \sigma)'$ on a neighbourhood of the origin, that is, in the vicinity of neutrality. Vicinity of neutrality is usually referred to as **weak selection**.

**Theorem 4.1** *As a function of the multidimensional selection coefficient* $\mathbf{s} = (\lambda, \delta, \alpha, \varepsilon, \sigma)'$, *the probability* $u$ *is differentiable, and in a neighbourhood of* $\mathbf{s} = \mathbf{0}$ *(selective neutrality),*

$$u = p + \mathbf{v}'.\mathbf{s} + o(\mathbf{s}), \tag{12}$$

*where the (weak) selection gradient* $\mathbf{v} = (v^\lambda, v^\delta, v^\alpha, v^\varepsilon, v^\sigma)'$ *can be expressed as*

$$\begin{aligned} v^\iota_{n,m} &= p\,(1-p)\; g^\iota_{n+m} \qquad\qquad \iota \neq \varepsilon, \\ v^\varepsilon_{n,m} &= p\,(1-p)\,(1-2p)\; g^\varepsilon_{n+m} \end{aligned}$$

*And the g's depend solely on the resident's characteristics* $b, c, d$*, and on the total initial population size* $n+m$*. They are called invasibility coefficients.*

We prove this theorem herafter. As seen in Proposition 3.1, the Kolmogorov forward equations translate into a discrete harmonic equation satisfied by $u$ with boundary condition 1 on $\Omega_1$, and 0 on $\Omega_2$, written as $(\Delta u)_{n,m} = 0$, where $\Delta$ is defined in (6). Combining (12) and (6), and identifying second-order terms, we get

$$(\Delta_0 v^\iota)_{n,m} = \begin{cases} \dfrac{nm}{(n+m)(n+m+1)} & \text{if} \quad \iota = \lambda \\[2ex] \dfrac{nm}{n+m} & \text{if} \quad \iota = \delta \\[2ex] \dfrac{nm}{(n+m)(n+m-1)} & \text{if} \quad \iota = \alpha, \sigma \\[2ex] \dfrac{nm\,(n-m)}{(n+m)(n+m-1)} & \text{if} \quad \iota = \varepsilon, \end{cases} \tag{13}$$

where $\Delta_0$ corresponds to the neutral case of $\Delta$: for any doubly indexed $w$,

$$\begin{aligned} (\Delta_0 w)_{n,m} &= (n+m)\Big[b + c(n+m-1) + d\Big] w_{n,m} - bn\, w_{n+1,m} - bm\, w_{n,m+1} \\ &\quad - n\Big[c(n+m-1)+d\Big] w_{n-1,m} - m\Big[c(n+m-1)+d\Big] w_{n,m-1} \qquad n, m \geq 0. \end{aligned} \tag{14}$$

We know from Theorem 3.3 that the vector $\mathbf{v} = (v^\lambda, v^\delta, v^\alpha, v^\varepsilon, v^\sigma)'$ is sublinear in $(n,m)$, that is, $(\|v_{n,m}\|/(n+m))_{n,m}$ is bounded. **Since the r.h.s. in (13) are all sublinear**, Lemma 3.2 ensures that $\mathbf{v}$ *is the unique sublinear vector in* $(n,m)$ *solving (13)* (if two such sequences existed, apply Lemma 3.2 to their difference).

Thanks to this uniqueness result, it is sufficient to show that there are solutions of (13) of the following form

$$v^\iota_{n,m} = \begin{cases} \dfrac{nm}{n+m}\, u^\iota_{n+m} & \text{if} \quad \iota = \lambda, \delta, \alpha, \sigma \\[2ex] \dfrac{nm\,(n-m)}{n+m}\, u^\iota_{n+m} & \text{if} \quad \iota = \varepsilon, \end{cases} \tag{15}$$

where for $\iota \neq \varepsilon$, $u^\iota$ is a *bounded* real sequence indexed by $\mathbb{N} - \{0,1\}$ ($u^\iota_1$ has no effect on the values of $v^\iota_{1,0}$ and $v^\iota_{0,1}$), and $u^\varepsilon$ is a real sequence indexed by $\mathbb{N} - \{0,1,2\}$ ($u^\varepsilon_1$ and $u^\varepsilon_2$ have no effect on the corresponding values of $v^\varepsilon_{n,m}$) such that $(nu^\varepsilon_n)_n$ is bounded. The proof will then end up by writing

$$g^\iota_n = \begin{cases} n\, u^\iota_n & \text{if} \quad \iota = \lambda, \delta, \alpha, \sigma, \quad n \geq 2 \\[1ex] n^2\, u^\iota_n & \text{if} \quad \iota = \varepsilon, \qquad\qquad n \geq 3. \end{cases} \tag{16}$$

In this setting, (13) holds iff

$$\begin{cases} (Lu^\lambda)_n = \dfrac{1}{n(n+1)} \quad \text{and} \quad (Lu^\delta)_n = \dfrac{1}{n} & n \geq 2, \\ \\ (Lu^\alpha)_n = (Lu^\sigma)_n = \dfrac{1}{n(n-1)} & n \geq 2, \\ \\ (L'u^\varepsilon)_n = \dfrac{1}{n(n-1)} & n \geq 3, \end{cases} \tag{17}$$

where $L$ (resp. $L'$) is the endomorphism of the vector space $\mathcal{L}_2$ (resp. $\mathcal{L}_3$) of real sequences indexed by $\mathbb{N} - \{0,1\}$ (resp. by $\mathbb{N} - \{0,1,2\}$) defined as

$$\begin{cases} (Lw)_n &= -b\dfrac{n+2}{n+1}\, w_{n+1} + \Big[b + c(n-1) + d\Big] w_n - (n-2)\left(c + \dfrac{d}{n-1}\right) w_{n-1} \quad n \geq 2 \\ \\ (L'w)_n &= -b\dfrac{n+3}{n+1}\, w_{n+1} + \Big[b + c(n-1) + d\Big] w_n - (n-3)\left(c + \dfrac{d}{n-1}\right) w_{n-1} \quad n \geq 3. \end{cases} \tag{18}$$

The following lemma ends the proof.

**Lemma 4.2 (existence)** *There are solutions $u^\lambda, u^\delta, u^\alpha, u^\varepsilon, u^\sigma$ of (17) such that $u^\lambda, u^\delta, u^\alpha, u^\sigma$ and $(nu^\varepsilon_n)_n$ are bounded.*

This lemma will be proved in the following subsection, by actually displaying explicit expressions for these solutions.

## 4.2 Properties of invasibility coefficients

### 4.2.1 Preliminary results

For $k \geq -2$, let $e^{(k)}$ be the sequence defined for $n \geq 2$ (3 if $k = -2$) by

$$e^{(k)}_n = \frac{1}{n+k},$$

and for $k = 2, 3$, let $\delta^{(k)}$ denote the Dirac mass at $k$

$$\delta^{(k)}_n = \begin{cases} 1 & \text{if } n = k \\ 0 & \text{otherwise.} \end{cases}$$

Then it is elementary to check (but has to be done carefully) that for $k \geq 1$,

$$Le^{(k)} = -\frac{b}{k}e^{(1)} + \frac{d}{k}e^{(-1)} - b\frac{k-1}{k}e^{(k+1)} + (b - (k+1)c + d)e^{(k)} + (k+1)(c - \frac{d}{k})e^{(k-1)}, \tag{19}$$

and that

$$Le^{(-1)} = -2be^{(0)} + be^{(1)} + be^{(-1)} + (c+d)\delta^{(2)}. \tag{20}$$

Likewise, for any $k \geq 1$ and for $k = -1$,

$$L'e^{(k)} = -\frac{2b}{k}e^{(1)} + \frac{2d}{k}e^{(-1)} - b\frac{k-2}{k}e^{(k+1)} + (b-(k+1)c+d)e^{(k)} + (k+2)(c-\frac{d}{k})e^{(k-1)}, \quad (21)$$

and also

$$L'e^{(-2)} = -2(b+d)e^{(-1)} + be^{(1)} + (b+c+d)e^{(-2)} + (c+\frac{d}{2})\delta^{(3)}. \quad (22)$$

Next observe that (17) can be written in the form

$$\begin{cases} Lu^{\lambda} = e^{(0)} - e^{(1)} \text{ and } Lu^{\delta} = e^{(0)}, \\ Lu^{\alpha} = Lu^{\sigma} = e^{(-1)} - e^{(0)}, \\ L'u^{\varepsilon} = e^{(-1)} - e^{(0)}, \end{cases} \quad (23)$$

so it is likely that the $u$'s can be expressed as linear combinations of the $e^{(k)}$'s. Actually, we will show they can be expressed as such linear combinations, with a potential extra additive term whose image by $L$ (resp. $L'$) is proportional to $\delta^{(2)}$ (resp. $\delta^{(3)}$). So we end these preliminaries with displaying two sequences: one in $\mathcal{L}_2$ whose image by $L$ is $\delta^{(2)}$, and one in $\mathcal{L}_3$ whose image by $L'$ is $\delta^{(3)}$.

Assume that, at time 0, instead of being assigned one of two different types (resident or mutant), all individuals are assigned distinct types. We denote by $\mathbf{P}_n$ the law of the (binary) logistic branching process $(b, c, d)$ starting from $n$ individuals distinctly labelled at time 0, where the type of an individual is transmitted to its offspring. In other words, under $\mathbf{P}$, we keep track of the whole descendance of each ancestral individual.
Then for k=2,3, let $T_k$ denote the first time when the total population size (i.e., the unlabelled process) is $k$. Finally, we define

$$q_n^{(k)} := \mathbf{P}_n(\text{ at time } T_k, \text{ the } k \text{ living individuals have } k \text{ distinct types }).$$

In the tree terminology, $q^{(k)}$ is the probability that all individuals in the first surviving $k$-tuple have different ancestors at time 0. In particular, $q_k^{(k)} = 1$.

**Lemma 4.3** *Let $D^{(2)} \in \mathcal{L}_2$ and $D^{(3)} \in \mathcal{L}_3$ be the sequences defined as*

$$D_n^{(2)} = \frac{q_n^{(2)}}{\kappa(n-1)} \qquad n \geq 2,$$

$$D_n^{(3)} = \frac{q_n^{(3)}}{\kappa'(n-1)(n-2)} \qquad n \geq 3,$$

*where*

$$\kappa = b\left(1 - \frac{2q_3^{(2)}}{3}\right) + c + d, \qquad \kappa' = \frac{b}{2}\left(1 - \frac{q_4^{(3)}}{2}\right) + c + \frac{d}{2}.$$

*Then $LD^{(2)} = \delta^{(2)}$ and $L'D^{(3)} = \delta^{(3)}$.*

**Proof.** It is quite elementary to check the result by standard applications of the Markov property under $\mathbf{P}$, but we prefer to give a more conceptual proof. We start with $D^{(2)}$. Under $\mathbb{P}$, we only keep track of two types at time $t$, i.e. the number $X_t$ of residents, and the number $Y_t$ of mutants, whereas under $\mathbf{P}_n$, there are $n$ types at time 0, say $1, 2, \dots, n$. Recall individuals of all types are exchangeable (because in this setting, the discrete operators $\Delta_0$ and $L$ are associated to selective neutrality). Set

$$w_{n,m} := \mathbb{P}_{n,m}(X_{T_2} = Y_{T_2} = 1).$$

Now by exchangeability,

$$\begin{aligned} w_{n,m} &= \sum_{i=1}^{n} \sum_{j=n+1}^{n+m} \mathbf{P}_{n+m}(\text{ at } T_2, \text{ type } i \text{ and type } j \text{ have one representative each }) \\ &= nm\mathbf{P}_{n+m}(\text{ at } T_2, \text{ type 1 and type 2 have one representative each }), \end{aligned}$$

and once again by exchangeability,

$$\begin{aligned} q_n^{(2)} &= \sum_{1\le i<j\le n} \mathbf{P}_n(\text{ at } T_2, \text{ type } i \text{ and type } j \text{ have one representative each }) \\ &= \binom{n}{2} \mathbf{P}_n(\text{ at } T_2, \text{ type 1 and type 2 have one representative each }). \end{aligned}$$

As a consequence,

$$w_{n,m} = \frac{2nm}{(m+n)(m+n-1)}\, q_{m+n}^{(2)}.$$

Observe that by definition $w$ is harmonic (in the sense that $\Delta_0 w = 0$) on the complementary of $\Omega_1 \cup \Omega_2 \cup \{(1,1)\}$. Then as in the previous subsection, with $v_n = q_n^{(2)}/(n-1)$, we get that $(Lv)_n = 0$ for any $n \ge 3$. The proof is completed by checking that $(Lv)_2 = \kappa \ne 0$.

As to $D^{(3)}$, set

$$w_{n,m} := \mathbb{E}_{n,m}(X_{T_3}Y_{T_3}(X_{T_3}-Y_{T_3})) = 2\mathbb{P}_{n,m}((X_{T_3}, Y_{T_3}) = (2,1)) - 2\mathbb{P}_{n,m}((X_{T_3}, Y_{T_3}) = (1,2)).$$

Now by exchangeability

$$\begin{aligned} &\mathbb{P}_{n,m}((X_{T_3}, Y_{T_3}) = (1,2)) \\ &= \sum_{i=1}^{n} \sum_{j=n+1}^{n+m} \mathbf{P}_{n+m}(\text{ at } T_3, \text{ type } i \text{ has one representative and type } j \text{ has two }) \\ &+ \sum_{1\le i\le n} \sum_{n+1\le j<k\le n+m} \mathbf{P}_{n+m}(\text{ at } T_3, \text{ types } i,\ j \text{ and } k \text{ have one representative each }) \\ &= nm\mathbf{P}_{n+m}(\text{ at } T_3, \text{ type 1 has one representative and type 2 has two }) \\ &+ n\frac{m(m-1)}{2}\mathbf{P}_{n+m}(\text{ at } T_3, \text{ types 1, 2 and 3 have one representative each }) \end{aligned}$$

Since $\mathbb{P}_{n,m}((X_{T_3}, Y_{T_3}) = (1,2)) = \mathbb{P}_{m,n}((X_{T_3}, Y_{T_3}) = (2,1))$, the corresponding first terms in the difference cancel out, and we are left with

$$w_{n,m} = nm(n-m)\mathbf{P}_{n+m}(\text{ at } T_3, \text{ types 1, 2 and 3 have one representative each }).$$

But again we get an expression involving the last displayed probability as

$$q_n^{(3)} = \begin{pmatrix} n \\ 3 \end{pmatrix} \mathbf{P}_n(\text{ at } T_3,\text{ types 1, 2 and 3 have one representative each }),$$

so that

$$w_{n,m} = \frac{6nm(n-m)}{(m+n)(m+n-1)(m+n-2)}\, q_{m+n}^{(3)}.$$

This time $w$ is harmonic on the complementary of $\Omega_1 \cup \Omega_2 \cup \{(1,1),(1,2),(2,1)\}$. Then with $v_n = q_n^{(3)}/(n-1)(n-2)$, we get that $(L'v)_n = 0$ for any $n \geq 4$. The proof is completed by checking that $(L'v)_3 = \kappa' \neq 0$. $\square$

**Lemma 4.4** *For any $k \geq 2$, $(q_n^{(k)})_n$ has a nonzero limit $q_\infty^{(k)}$ as $n \to \infty$.*

**Proof.** From [23], we know that $\infty$ is an entrance boundary for the probabilities $\mathbf{P}_n$, $n \geq 1$, so that $\mathbf{P}_\infty$ and $q_\infty^{(k)}$ are properly defined for any $k \geq 1$. At time $t$, we denote by $Z_t$ the number of living individuals and by $N_t$ the number of types represented. Obviously, under $\mathbf{P}_\infty$, $Z_t \to \infty$ as $t \to 0+$. As to $N$, since it is a nonincreasing function of time, it has a right-limit $N_{0+} \leq \infty$ at $t = 0$. Next we want to show that

$$\exists k_0 \geq 2, \quad q_\infty^{(k_0)} = 0 \quad \Rightarrow \quad \mathbf{P}_\infty(N_{0+} \leq k_0) = 1. \tag{24}$$

This will end the proof of the Lemma. Indeed, $N_{0+} \leq k_0$ means that, under $\mathbf{P}_\infty$, there are at most $k_0$ individuals whose total descendance at any time $t$ is $Z_t$. Then, conditional on these individuals, $Z$ would be dominated by a binary logistic branching process starting at $k_0$, which contradicts the fact that $Z_{0+} = +\infty$. Conclude by summing over all possible $k_0$-tuples.

Now, we prove (24). Assume there is $k_0 \geq 2$ such that $q_\infty^{(k_0)} = 0$. Since for $k \geq k_0$, $q_n^{(k_0)} > q_n^{(k)} q_k^{(k_0)}$, we get that $q_\infty^{(k)} = 0$ for all $k \geq k_0$. Recall that $T_j$ is the first hitting time of $j$ by $Z$. For $n \geq j \geq k \geq k_0$,

$$\begin{aligned} q_n^{(k)} &> \mathbf{P}_n(N_{T_j} = k, N_{T_k} = k) \\ &> \mathbf{P}_n(N_{T_j} = k) C(j,k), \end{aligned}$$

where $C(j,k)$ is the probability that conditional on $T_j = k$, after picking $k$ representative individuals at $T_j$ (one for each type), the first $j - k$ events after $T_j$ are the deaths of all non-representative individuals. Because this probability only depends on $j$ and $k$, we get that $\mathbf{P}_\infty(N_{T_j} = k) = 0$ for all $j \geq k \geq k_0$. As a consequence,

$$\mathbf{P}_\infty(N_{T_j} \leq k_0) = 1.$$

But under $\mathbf{P}_\infty$, $\lim_{j\to\infty} T_j = 0$ a.s., so that $\mathbf{P}_\infty(N_{0+} \leq k_0) = 1$. $\square$

The proof of Lemma 4.2 is split up into the following (sub)subsections, each of which being dedicated to one (or two) invasibility coefficient(s). In Propositions 4.5, 4.6, 4.8 and 4.10, we display the solutions of (17) such that $u^\iota$ ($\iota \neq \varepsilon$) and $(nu_n^\varepsilon)_n$ are bounded (therefore proving Lemma 4.2). We also specify the behaviour of each invasibility coefficient as the population size grows to infinity.

### 4.2.2 Results for the $\lambda$-invasibility

Here, we must find a bounded sequence $u^\lambda$ in $\mathcal{L}_2$ such that $Lu^\lambda = e^{(0)} - e^{(1)}$.
Recall from Lemma 4.3 that $D^{(2)} \in \mathcal{L}_2$ is a sequence such that $LD^{(2)} = \delta^{(2)}$, and

$$D_n^{(2)} = \frac{q_n^{(2)}}{\kappa(n-1)} \qquad n \geq 2,$$

where $q_n^{(2)}$ is the probability that the first surviving pair in the (labelled) logistic branching process $(b, c, d)$ have two distinct ancestors in the initial $n$-tuple.
Since, by (20),

$$Le^{(-1)} = -2be^{(0)} + be^{(1)} + be^{(-1)} + (c+d)\delta^{(2)},$$

and by (19)

$$Le^{(1)} = de^{(-1)} + (-2c+d)e^{(1)} + 2(c-d)e^{(0)},$$

we can readily state the following

**Proposition 4.5 (fertility)** *The sequence $u^\lambda$ defined as*

$$u^\lambda = -\frac{d}{2bc}e^{(-1)} + \frac{d(c+d)}{2bc}D^{(2)} + \frac{1}{2c}e^{(1)}, \tag{25}$$

*is a bounded sequence of $\mathcal{L}_2$ such that $Lu^\lambda = e^{(0)} - e^{(1)}$. Then the invasibility coefficient $g^\lambda$ associated to **fertility** ($g_n^\lambda = nu_n^\lambda$) is given by*

$$g_n^\lambda = -\frac{dn}{2bc(n-1)} + \frac{d(c+d)}{2bc\kappa}\frac{nq_n^{(2)}}{n-1} + \frac{n}{2c(n+1)} \qquad n \geq 2. \tag{26}$$

*In particular,*

$$\lim_{n\to\infty} g_n^\lambda = \frac{b - d + d(c+d)q_\infty^{(2)}/\kappa}{2bc}$$

### 4.2.3 Results for the $\alpha$ and $\sigma$-invasibilities

Here, we must find bounded sequences $u^\alpha$ and $u^\sigma$ in $\mathcal{L}_2$ such that $Lu^\alpha = Lu^\sigma = e^{(-1)} - e^{(0)}$. Exactly in the same way as for the $\lambda$-invasibility coefficient, we can readily make the needed statement.

**Proposition 4.6 (aggressiveness, survival)** *The sequences $u^\alpha$ and $u^\sigma$ defined as*

$$u^\alpha = u^\sigma = \frac{2c-d}{2bc}e^{(-1)} - \frac{(2c-d)(c+d)}{2bc}D^{(2)} + \frac{1}{2c}e^{(1)}, \tag{27}$$

*are bounded sequences of $\mathcal{L}_2$ such that $Lu^\alpha = Lu^\sigma = e^{(-1)} - e^{(0)}$. Then the invasibility coefficients associated to **aggressiveness** ($g_n^\alpha = nu_n^\alpha$) and **survival**, ($g_n^\sigma = nu_n^\sigma$) are given by*

$$g_n^\alpha = g_n^\sigma = \frac{(2c-d)n}{2bc(n-1)} - \frac{(2c-d)(c+d)}{2bc\kappa}\frac{nq_n^{(2)}}{n-1} + \frac{n}{2c(n+1)} \qquad n \geq 2. \tag{28}$$

*In particular,*

$$\lim_{n\to\infty} g_n^\alpha = \lim_{n\to\infty} g_n^\sigma = \frac{b + 2c - d - (2c-d)(c+d)q_\infty^{(2)}/\kappa}{2bc}$$

### 4.2.4 Results for the $\delta$-invasibility

For $\delta$ and $\varepsilon$-invasibility coefficients, the task is mathematically more challenging. A side-effect is that we only obtain fine results in the case when the resident species has no natural death rate. This shortcoming is not very disturbing, however, because we are especially interested in precisely those populations with stationary behaviour (which are those needed for applications to adaptive dynamics). From now on, we assume that $d = 0$.
Recall that we must find a bounded sequence $u^\delta$ in $\mathcal{L}_2$ such that $Lu^\delta = e^{(0)}$.

Set $\theta := b/c$.

**Lemma 4.7** *Let $\Phi$ be the sequence of $\mathcal{L}_2$ defined recursively as $\Phi_2 = 1$ and*

$$c(n+2)\Phi_{n+1} + [b - c(n+1)]\,\Phi_n - b\frac{n-2}{n-1}\Phi_{n-1} = 0. \tag{29}$$

*Then the sequence $(n\Phi_n)_n$ converges to a nonzero finite limit $\Phi_\infty$, and the (thus well-defined) sum*

$$S := \sum_{n\geq 2} n^{-1}\Phi_n$$

*has $3c - bS = c\Phi_\infty$.*

**Proposition 4.8 (defence capacity)** *Define the sequence $\phi$ of $\mathcal{L}_2$ as*

$$\phi_n := \Phi_n / c\Phi_\infty \qquad n \geq 2.$$

*Then, with $\phi_1 := 1/2c$, the sequence $u^\delta$ of $\mathcal{L}_2$ defined as*

$$u^\delta := \sum_{k\geq 1} \phi_k e^{(k)}$$

*is a bounded sequence such that $Lu^\delta = e^{(0)}$. The invasibility coefficient $g^\delta$ associated to* ***defence*** *capacity ($g_n^\delta = nu_n^\delta$) is given by*

$$g_n^\delta = \sum_{k\geq 1} \frac{n\phi_k}{n+k} \qquad n \geq 2. \tag{30}$$

*In particular,*

$$g_n^\delta \sim \frac{1}{c}\ln(n) \qquad \text{as } n \to \infty.$$

The proof of Lemma 4.7 stems immediately from the two following claims. Claim 1 will also be helpful in the proof of Proposition 4.8.

**Claim 1.** Assume that $(n\Phi_n)_n$ converges to a finite limit $\Phi_\infty$, where $(\Phi_n)_n$ is defined in (29). Then the real number $S := \sum_{n\geq 2} n^{-1}\Phi_n$ and the sequence $W := \sum_{k\geq 2} \Phi_k e^{(k)}$ of $\mathcal{L}_2$ are well-defined, and

(i) $c\Phi_\infty = 3c - bS$

(ii) $LW = c\Phi_\infty e^{(1)}$.

**Claim 2.** The sequence $(n\Phi_n)_n$ converges to a nonzero finite limit.

**Proof of Claim 1.** To prove (i), let

$$\beta_n := (n+1)\Phi_n \quad \text{and} \quad \gamma_n := (n-1)\Phi_n \qquad n \geq 2, \tag{31}$$

so that

$$\lim_{n\to\infty} \beta_n = \lim_{n\to\infty} \gamma_n = \Phi_\infty,$$

and, thanks to (29),

$$\beta_{n+1} - \beta_n = -\frac{\theta}{n-1}(\gamma_n - \gamma_{n-1}) \qquad n \geq 2, \tag{32}$$

with $\gamma_1 = 0$. As a consequence, by Abel's transform, we get

$$\begin{aligned} S &= \sum_{n\geq 2} \gamma_n \left(\frac{1}{n-1} - \frac{1}{n}\right) = \sum_{n\geq 2} \frac{\gamma_n - \gamma_{n-1}}{n-1} \\ &= -\theta^{-1} \sum_{n\geq 2} (\beta_{n+1} - \beta_n) = -\theta^{-1}(\Phi_\infty - \beta_2) \\ &= (3 - \Phi_\infty)/\theta. \end{aligned}$$

As for (ii), thanks to (19) and (20), and by continuity of linear operators,

$$\begin{aligned} LW &= \lim_{l\to\infty} \sum_{k=2}^{l} \Phi_k L e^{(k)} \\ &= \lim_{l\to\infty} \sum_{k=2}^{l} \Phi_k \left( -\frac{b}{k} e^{(1)} - b\frac{k-1}{k} e^{(k+1)} + [b - (k+1)c] e^{(k)} + (k+1)c e^{(k-1)} \right) \\ &= \lim_{l\to\infty} \left\{ -b \left( \sum_{k=2}^{l} k^{-1} \Phi_k \right) e^{(1)} - b \sum_{i=2}^{l+1} \frac{i-2}{i-1} \Phi_{i-1} e^{(i)} \right. \\ &\qquad \left. + \sum_{k=2}^{l} [b - (k+1)c] \Phi_k e^{(k)} + \sum_{j=1}^{l-1} c(j+2) \Phi_{j+1} e^{(j)} \right\} \\ &= -bSe^{(1)} + \lim_{l\to\infty} \left\{ \sum_{k=2}^{l-1} \left( -b\frac{k-2}{k-1} \Phi_{k-1} + [b - (k+1)c] \Phi_k + c(k+2) \Phi_{k+1} \right) e^{(k)} \right. \\ &\qquad \left. -b\frac{l-2}{l-1} \Phi_{l-1} e^{(l)} - b\frac{l-1}{l} \Phi_l e^{(l+1)} + [b - (l+1)c] \Phi_l + 3c\Phi_2 e^{(1)} \right\} \\ &= (3c - bS) e^{(1)}, \end{aligned}$$

which ends the proof. □

**Proof of Claim 2.** We split this proof into the four following steps (recall (31))

(i) if $(n\Phi_n)_n$ converges to a finite limit $\Phi_\infty$, then $\Phi_\infty \neq 0$

(ii) $(\beta_n)_n$ has constant sign for large $n$

(iii) $(\beta_n)_n$ is bounded

(iv) $(\beta_n)_n$ converges.

Since we are only interested in the asymptotic properties of the sequences $(\Phi_n)_n$, $(\beta_n)_n$ and $(\gamma_n)_n$, we will implicitly assume throughout this proof that $\theta/(n+1) < 1$, that is, $n \geq \theta$.

(i) If $\Phi_\infty$ exists, then thanks to Claim 1, we can define $W = \sum_{k\geq 2} \Phi_k e^{(k)}$ and the doubly indexed sequence $w$ as

$$w_{n,m} = \frac{nm}{n+m} W_{n+m} \qquad (n,m) \in \mathbb{N} \times \mathbb{N} \setminus (0,0).$$

Because $W$ is bounded, $w$ is sublinear. Assume $\Phi_\infty = 0$. Then by Claim 1, $LW = 0$, and the same calculations as those yielding (17) and (18) show that $\Delta_0 w = 0$. The contradiction comes with Lemma 3.2, which implies that the null sequence is the only sublinear doubly indexed sequence which vanishes on $\Omega_1 \cup \Omega_2$ and is in the kernel of $\Delta_0$.

(ii) First observe that (29) reads

$$\beta_{n+1} = \left(1 - \frac{\theta}{n+1}\right) \beta_n + \theta \frac{n-2}{n(n-1)} \beta_{n-1}, \tag{33}$$

so if there is $n_0 \geq \theta$ such that $\beta_{n_0}\beta_{n_0-1} \geq 0$, then a straightforward induction shows that $(\beta_n)_{n\geq n_0}$ never changes sign. Now, we prove that if no such $n_0$ exists, then $(\beta_n)_n$ converges to 0, which contradicts (i). Indeed, assume that for all $n \geq \theta$, $\beta_n\beta_{n-1} < 0$. Then for any $n \geq \theta + 1$, if $\beta_{n-1} < 0$, then $\beta_n > 0$ and $\beta_{n+1} < 0$, so that

$$\left(1 - \frac{\theta}{n+1}\right) \beta_n < -\theta \frac{n-2}{n(n-1)} \beta_{n-1},$$

which can be written as

$$\left(1 - \frac{\theta}{n+1}\right) |\beta_n| < \theta \frac{n-2}{n(n-1)} |\beta_{n-1}|,$$

and we would get the same inequality if $\beta_{n-1} > 0$. This would imply that $|\beta_n|/|\beta_{n-1}|$ would vanish as $n$ grows, and so would $\beta_n$.

(iii) Without loss of generality, we can assume thanks to (ii) that there is $n_0$ such that $\beta_n \geq 0$ for all $n \geq n_0$ (otherwise change $\beta$ for $-\beta$). Next, we prove that for all $n \geq n_0$, $\beta_{n+1} < \max(\beta_n, \beta_{n-1})$. It is then elementary to see that $(\beta_n)_n$ is bounded. First check that

$$\beta_{n+1} - \beta_n = -\frac{\theta}{n-1}(\beta_n - \beta_{n-1}) - 2\theta \frac{\beta_{n-1}}{n(n-1)(n+1)},$$

so for any $n \geq n_0$,

$$\beta_{n+1} - \beta_n \leq -\frac{\theta}{n-1}(\beta_n - \beta_{n-1}).$$

In particular, if $\beta_{n+1} \geq \beta_n$, then $\beta_n \leq \beta_{n-1}$, and

$$|\beta_{n+1} - \beta_n| \leq \frac{\theta}{n-1} |\beta_n - \beta_{n-1}| < |\beta_n - \beta_{n-1}|,$$

which reads $\beta_{n+1} - \beta_n < -\beta_n + \beta_{n-1}$, that is, $\beta_{n+1} < \beta_{n-1}$. As a conclusion, $\beta_{n+1} < \beta_n$ or $\beta_{n+1} < \beta_{n-1}$.

(iv) By (32) and Abel's transform, we get

$$\begin{aligned}\beta_{n+1} - \beta_2 &= -\theta \sum_{k=2}^{n} \frac{\gamma_k - \gamma_{k-1}}{k-1} \\ &= -\theta \frac{\gamma_n}{n-1} - \theta \sum_{k=2}^{n-1} \frac{\gamma_k}{k(k-1)},\end{aligned}$$

and the r.h.s. converges, because $(\gamma_n)_n$ is bounded, thanks to (iii). □

**Proof of Proposition 4.8.** Thanks to Claim 1 above, since $u^\delta = \phi_1 e^{(1)} + (c\Phi_\infty)^{-1} W$,

$$\begin{aligned}Lu^\delta &= \phi_1 Le^{(1)} + (c\Phi_\infty)^{-1} LW \\ &= (2c)^{-1}(-2ce^{(1)} + 2ce^{(0)}) + (c\Phi_\infty)^{-1} c\Phi_\infty e^{(1)} \\ &= e^{(0)}.\end{aligned}$$

The boundedness of $u^\delta$ is straightforward. To get the equivalent of $g_n^\delta$ as $n \to \infty$, it is sufficient to prove that $nW_n \sim \Phi_\infty \ln(n)$. First, starting over from the proof of Claim 2 (iv) above, get that

$$\Phi_\infty - \beta_2 = -\theta \sum_{k \geq 2} \frac{\gamma_k}{k(k-1)},$$

so that

$$\beta_{n+1} - \Phi_\infty = -\theta \frac{\gamma_n}{n-1} + \theta \sum_{k \geq n} \frac{\gamma_k}{k(k-1)},$$

whch implies that $\beta_n - \Phi_\infty = o\left(n^{-1}\right)$. Next, writing $\rho_k := k\Phi_k$, we get

$$\begin{aligned}nW_n &= \sum_{k \geq 2} \frac{n\rho_k}{k(n+k)} = \sum_{k \geq 2} \rho_k \left(\frac{1}{k} - \frac{1}{n+k}\right) \\ &= \lim_{l \to \infty} \left\{ \sum_{k=2}^{n+1} \frac{\rho_k}{k} + \sum_{k=2}^{l} \frac{\rho_{n+k} - \rho_k}{n+k} - \sum_{k=l+1}^{l+n} \frac{\rho_k}{k} \right\} \\ &= \sum_{k=2}^{n+1} \frac{\rho_k - \Phi_\infty}{k} + \sum_{k=2}^{n+1} \frac{\Phi_\infty}{k} + \sum_{k \geq 2} \frac{\rho_{n+k} - \rho_k}{n+k} \\ &= \Phi_\infty \ln(n) + O\left(1\right),\end{aligned}$$

where the last equation comes from the fact that $\rho_k = \Phi_\infty + O\left(k^{-1}\right)$ as $k \to \infty$. □

### 4.2.5 Results for the $\varepsilon$-invasibility

Recall from Lemma 4.3 that $D^{(3)}$ is a sequence in $\mathcal{L}_3$ such that $LD^{(3)} = \delta^{(3)}$, and

$$D_n^{(3)} = \frac{q_n^{(3)}}{\kappa'(n-1)(n-2)} \qquad n \geq 3,$$

where $q_n^{(3)}$ is the probability that the first surviving triple in the (labelled) logistic branching process $(b, c, d)$ have three distinct ancestors in the initial $n$-tuple. Now as in the previous problem ($\delta$-invasibility), we assume that the resident species has no natural death rate, that is, $d = 0$.

Here, we must find a sequence $u^\varepsilon$ in $\mathcal{L}_3$ such that $(nu_n^\varepsilon)_n$ is bounded, and $L'u^\varepsilon = e^{(-1)} - e^{(0)}$. Recall $\theta = b/c$.

**Lemma 4.9** *Let $\Psi$ be the sequence of $\mathcal{L}_3$ defined recursively as $\Psi_3 = 1$ and*

$$c(n+3)\Psi_{n+1} + [b - c(n+1)]\,\Psi_n - b\frac{n-3}{n-1}\Psi_{n-1} = 0. \tag{34}$$

*Then the sequence $(n^2\Psi_n)_n$ converges to a nonzero finite limit $\Psi_\infty$, and the (thus well-defined) sums*

$$S := \sum_{n\geq 3} n^{-1}\Psi_n \text{ and } \Sigma := \sum_{n\geq 3} \Psi_n$$

*have*

$$\Sigma + 2\theta S = \Psi_\infty + (\theta - 3)\Sigma = 5.$$

**Proposition 4.10 (isolation)** *Define the sequence $\psi$ of $\mathcal{L}_3$ as*

$$\psi_n := -\Psi_n/c\Psi_\infty \qquad n \geq 3.$$

*Then, with*

$$\begin{aligned}
\psi_{-2} &= -\frac{1}{b(\theta+3)}\\
\psi_{-1} &= \frac{\theta+1}{b(\theta+3)}\\
\psi_1 &= \frac{2\theta}{3c(\theta+3)}\\
\psi_2 &= \frac{\Sigma}{c\Psi_\infty} - \frac{2\theta+3}{3c(\theta+3)},
\end{aligned}$$

*the sequence $u^\varepsilon$ of $\mathcal{L}_3$ defined as*

$$u^\varepsilon := \sum_{k\geq -2, k\neq 0} \psi_k e^{(k)} + \frac{1}{\theta(\theta+3)} D^{(3)}$$

*is such that $(nu_n^\varepsilon)_n$ is bounded and $L'u^\varepsilon = e^{(-1)} - e^{(0)}$. Then the invasibility coefficient $g^\varepsilon$ associated to* ***isolation*** *($g_n^\varepsilon = n^2 u_n^\varepsilon$) is given by*

$$g_n^\varepsilon = \sum_{k\geq 1} \frac{n^2\psi_k}{n+k} + \frac{1}{\kappa'\theta(\theta+3)} \frac{n^2 q_n^{(3)}}{(n-1)(n-2)} \qquad n \geq 3. \tag{35}$$

*In particular,*

$$g_n^\varepsilon \sim \frac{1}{c}\ln(n) \qquad \text{as } n \to \infty.$$

Since proofs of these statements are quite similar to those done for the $\delta$-invasibility, we will often sketch them. First, we prove the following two claims.

**Claim 1.** Assume that $(n^2\Psi_n)_n$ converges to a finite limit $\Psi_\infty$, where $(\Psi_n)_n$ is defined in (34). Then the real numbers $S := \sum_{n\geq 3} n^{-1}\Psi_n$ and $\Sigma := \sum_{n\geq 3}\Psi_n$, as well as the sequence $Z := \sum_{k\geq 3}\Psi_k e^{(k)}$ of $\mathcal{L}_3$ are well-defined, and

(i) $\Sigma + 2\theta S = \Psi_\infty + (\theta - 3)\Sigma = 5$.

(ii) $L'(Z - \Sigma e^{(2)}) = c\Psi_\infty(e^{(2)} - e^{(1)})$.

**Claim 2.** The sequence $(n^2\Psi_n)_n$ converges to a nonzero finite limit.

**Proof of Claim 1.** To prove (i), let

$$\beta_n := (n+2)(n+1)\Psi_n \quad \text{and} \quad \gamma_n := (n-2)(n-1)\Psi_n \qquad n \geq 3, \tag{36}$$

so that

$$\lim_{n\to\infty}\beta_n = \lim_{n\to\infty}\gamma_n = \Psi_\infty,$$

and, thanks to (34),

$$\beta_{n+1} - \beta_n = -\frac{\theta(n+2)}{(n-1)(n-2)}(\gamma_n - \gamma_{n-1}) \qquad n \geq 3, \tag{37}$$

with $\gamma_2 = 0$. As a consequence, by two applications of Abel's transform, we get

$$\begin{aligned}
\Sigma &= \sum_{n\geq 3}\Psi_n = \sum_{n\geq 3}\beta_n\left(\frac{1}{n+1} - \frac{1}{n+2}\right)\\
&= \frac{\beta_3}{4} + \sum_{n\geq 3}\frac{\beta_{n+1}-\beta_n}{n+2} = 5\Psi_3 - \theta\sum_{n\geq 3}\frac{\gamma_n - \gamma_{n-1}}{(n-1)(n-2)}\\
&= 5 - \theta\sum_{n\geq 3}\frac{2\gamma_n}{n(n-1)(n-2)}\\
&= 5 - 2\theta S.
\end{aligned}$$

On the other hand, the same type of arguments as above show that

$$\begin{aligned}
\Sigma &= \sum_{n\geq 3}\frac{\gamma_n}{(n-2)(n-1)} = \sum_{n\geq 3}\frac{\gamma_n - \gamma_{n-1}}{n-2}\\
&= -\theta^{-1}\sum_{n\geq 3}\frac{n-1}{n+2}(\beta_{n+1} - \beta_n)\\
&= -\theta^{-1}\left\{-\frac{\beta_3}{4} + \Psi_\infty + \sum_{n\geq 3}\left(\frac{n-2}{n+1}\beta_n - \frac{n-1}{n}\beta_n\right)\right\}\\
&= -\theta^{-1}\left(-5 + \Psi_\infty - 3\Sigma\right),
\end{aligned}$$

which ends the proof of (i). With the help of (21) and (22), (ii) can be proved easily mimicking what was done for the $\delta$-invasibility. □

**Proof of Claim 2.** We proceed just as for the defence capacity. First, we prove that if $\Psi_\infty$ exists, it cannot be 0. Indeed, consider $\Sigma$ and $Z \in \mathcal{L}_3$ defined in Claim 1, and further define

$$z_{n,m} = \frac{nm(n-m)}{n+m}\left(Z_{n+m} - \frac{\Sigma}{n+m+2}\right) \qquad (n,m) \in \mathbb{N}^\star \times \mathbb{N}^\star.$$

Because $(nZ_n)_n$ is bounded, $z$ is sublinear. If $\Psi_\infty = 0$, then thanks to Claim 1, we would get $\Delta_0 z = 0$, but this would contradict Lemma 3.2.
Next recall the sequences $\beta$ and $\gamma$ defined in (36). Thanks to (34),

$$\beta_{n+1} = \left(1 - \frac{\theta}{n+1}\right)\beta_n + \theta\frac{(n+2)(n-3)}{n(n-1)(n+1)}\beta_{n-1},$$

which proves that $\beta_n$ has constant sign for large $n$, otherwise it would converge to 0 (and then $\Psi_\infty = 0$, which would contradict (i)). Therefore we can assume that $\beta_n \geq 0$ for large $n$ without loss of generality. Since

$$\beta_{n+1} - \beta_n = -\frac{\theta}{n+1}(\beta_n - \beta_{n-1}) - 6\theta\frac{\beta_{n-1}}{n(n-1)(n+1)},$$

then for any sufficiently large $n$, $0 \leq \beta_{n+1} < \max(\beta_n, \beta_{n-1})$, so that $(\beta_n)_n$ is bounded, and so is $(\gamma_n)_n$. Use (37) to show that

$$\beta_{n+1} - \beta_3 = -\theta\frac{n+2}{(n-1)(n-2)}\gamma_n - \theta\sum_{k=3}^{n-1}\frac{k+6}{k(k-1)(k-2)}\gamma_k,$$

and conclude that $(\beta_n)_n$ is convergent. $\square$

**Proof of Proposition 4.10.** Recall $Z$ defined in Claim 1 and set $\varphi_2 := -(2\theta+3)/3c(\theta+3)$, as well as $V \in \mathcal{L}_3$

$$V := \psi_{-2}e^{(-2)} + \psi_{-1}e^{(-1)} + \psi_1 e^{(1)} + \varphi_2 e^{(2)},$$

so that

$$u^\varepsilon = -(c\Psi_\infty)^{-1}(Z - \Sigma e^{(2)}) + V + \frac{1}{\theta(\theta+3)}D^{(3)}. \tag{38}$$

By an elementary computation relying on (21) and (22), get

$$L'V = -\frac{1}{\theta(\theta+3)}\delta^{(3)} + e^{(-1)} - e^{(0)} - e^{(1)} + e^{(2)},$$

and conclude, thanks to Claim 1, that $L'u^\varepsilon = e^{(-1)} - e^{(0)}$.
To get the equivalent of $g_n^\varepsilon$ as $n \to \infty$, first recall (38) and observe that

$$\psi_{-2} + \psi_{-1} + \psi_1 + \varphi_2 = 0,$$

so that $(n^2 V_n)_n$ converges. Next consider $nZ_n$

$$\begin{aligned} nZ_n &= \sum_{k\geq 3}\frac{n\Psi_k}{n+k} = \sum_{k\geq 3}k\Psi_k\left(\frac{1}{k} - \frac{1}{n+k}\right) \\ &= \Sigma - \sum_{k\geq 3}\frac{k\Psi_k}{n+k} \\ &= \Sigma - \Psi_\infty\frac{\ln(n)}{n} + o\left(\frac{\ln(n)}{n}\right), \end{aligned}$$

by a similar method as in the proof of Proposition 4.8. As a consequence,

$$nu_n^\varepsilon = -(c\Psi_\infty)^{-1}\left(nZ_n - \Sigma\frac{n}{n+2}\right) + nV_n + \frac{1}{\theta(\theta+3)}nD_n^{(3)} = \frac{\ln(n)}{cn} + o\left(\frac{\ln(n)}{n}\right),$$

which ends the proof, since $g_n^\varepsilon = n^2 u_n^\varepsilon$. □

# 5 Adaptive dynamics in finite populations

In this section, we consider a stochastic model of evolution in a finite population generalizing the previous one, where the number of coexisting types and the corresponding population sizes may fluctuate. First, our goal is to apply to this process a limit of rare mutations while keeping the population size finite, in order to describe the evolutionary process on the mutation timescale as a 'trait substitution sequence' (TSS, [28, 3]) where evolution proceeds by successive fixations of mutant types. Second, we want to apply to this TSS a limit of small mutation steps (weak selection) in the way leading to the 'canonical equation of adaptive dynamics' [8, 4]. Because of the combination of population finiteness and weak selection, the genetic drift carries over to the limiting process. This is a diffusion process of evolution over the trait space that we call 'canonical diffusion of adaptive dynamics', and is grounded on a realistic individual-based description of the evolutionary dynamics, in which the population size may fluctuate because of interactions of the logistic branching type.

However, in finite stochastic populations, making mutations rare eventually leads to the extinction of the population before any mutation occurs. Therefore, in order to apply our first limit, we have to consider a model where the extinction is impossible, in a similar way as in the case where $d_1 = d_2 = 0$ in the two-type logistic branching model of the previous sections.

## 5.1 Description of the process

At any time $t$, the population is composed of a finite number $N(t)$ of individuals characterized by their (phenotypic) traits $x_1(t), \dots, x_{N(t)}(t)$ belonging to a given trait space $\mathcal{X}$, assumed to be a closed subset of $\mathbb{R}^k$ for some $k \geq 1$. The population state at time $t$ will be represented by the counting measure on $\mathcal{X}$

$$\nu_t = \sum_{i=1}^{N(t)} \delta_{x_i(t)}.$$

Let us denote by $\mathcal{M}$ the set of finite counting measures on $\mathcal{X}$, endowed with the $\sigma$-field induced by the Borel $\sigma$-field on $\mathcal{X} \subset \mathbb{R}^d$ as follows: let $\varphi$ denote the application mapping any element $\sum_{i=1}^k \delta_{x_i}$ of $\mathcal{M}$ to the $k$-tuple $(x_{\pi(1)}, \dots, x_{\pi(k)})$ where the permutation $\pi$ of $\{1, \dots, k\}$ is chosen such that this vector is ranked in, say, the lexicographical order. Then, this function $\varphi$ is a bijection from $\mathcal{M}$ to the set of lexicographically ordered vectors of $\cup_{i=0}^k \mathcal{X}^k$. The Lebesgue $\sigma$-field on this set therefore provides a $\sigma$-field on $\mathcal{M}$.

For any $\nu \in \mathcal{M}$ and any measurable function $f : \mathcal{X} \to \mathbb{R}$, we will use the notation $\langle \nu, f \rangle$ for $\int f(x)\nu(dx)$. Observe that $N(t) = \langle \nu_t, \mathbf{1} \rangle$ and that $\langle \nu_t, \mathbf{1}_\Gamma \rangle$ is the number of individuals at time $t$ with trait value in $\Gamma \subset \mathcal{X}$.

The Markovian dynamics of the population are governed by the following parameters:

- $b(x)$ is the birth rate of an individual with trait $x$.
- $c(x,y)$ is the rate of competition felt **by** an individual with trait $x$ **from** an individual with trait $y$. Specifically, in a population with state $\nu = \sum_{i=1}^{N} \delta_{x_i}$, the death rate of an individual with trait $x$ (where $x$ is one of the $x_i$) is given by

$$\sum_{i=1}^{N} c(x,x_i) - c(x,x) = \int_{\mathcal{X}} c(x,y)(\nu(dy) - \delta_x(dy)). \tag{39}$$

- $\gamma\mu(x)$ is the probability that a birth from an individual with trait $x$ produces a mutant individual, where $\mu(x) \in [0,1]$ and where $\gamma \in (0,1)$ is a parameter scaling the frequence of mutations. We are interested in the limit $\gamma \to 0$.
- $M(x,dh)$ is the law of the trait step $h = y - x$ between a mutant individual with trait $y$ born from an individual with trait $x$. Since the mutant trait $y = x + h$ must belong to $\mathcal{X}$, this measure has its support in $\mathcal{X} - x := \{y - x : y \in \mathcal{X}\} \subset \mathbb{R}^k$.

In other words, the infinitesimal generator of the $\mathcal{M}$-valued Markov jump process $(\nu_t^\gamma)_{t\geq 0}$ is given by

$$\begin{aligned} L_\gamma\varphi(\nu) = &\int_{\mathcal{X}} [\varphi(\nu + \delta_x) - \varphi(\nu)]\,(1 - \gamma\mu(x))\,b(x)\nu(dx) \\ &+ \int_{\mathcal{X}}\int_{\mathbb{R}^k} [\varphi(\nu + \delta_{x+h}) - \varphi(\nu)]\gamma\mu(x)b(x)M(x,dh)\nu(dx) \\ &+ \int_{\mathcal{X}} [\varphi(\nu - \delta_x) - \varphi(\nu)]\left(\int_{\mathcal{X}} c(x,y)\nu(dy) - c(x,x)\right)\nu(dx). \end{aligned} \tag{40}$$

In this equation, each integral with respect to $\nu \in \mathcal{M}$ corresponds to a sum over all individuals in the population. The first term describes the birth events without mutation, the second term the birth events with mutation, and the third term describes the death events (by competition). We will denote by $\mathbb{P}^\gamma$ the law of this process (or $\mathbb{P}^\gamma_{\nu_0}$ when the initial condition has to be specified). When necessary, we will denote the dependence of $\nu_t$ on the parameter $\gamma$ with the notation $\nu_t^\gamma$.

We make the following assumptions:

- There are positive constants $\bar{b}$, $\underline{c}$ and $\bar{c}$ such that $b(\cdot) \leq \bar{b}$ and $\underline{c} \leq c(\cdot,\cdot) \leq \bar{c}$.
- $M(x,dh)$ has a density on $\mathbb{R}^k$ which is uniformly bounded in $x \in \mathcal{X}$ by some integrable function $\bar{M}(h)$.

For any $\gamma \in (0,1)$, the population size $\langle \nu_t^\gamma, \mathbf{1}\rangle$ is dominated by a scalar logistic branching process with dynamical parameters $(\bar{b}, 0, \underline{c})$, so that one can prove exactly as for Proposition 2.1(a) the following result.

**Proposition 5.1** *Fix $p \geq 1$ and pick a positive $C$. There is a constant $C'$ such that, for any $\gamma \in (0,1)$,*

$$\mathbb{E}(\langle \nu_0^\gamma, \mathbf{1}\rangle^p) \leq C \quad \Rightarrow \quad \sup_{t\geq 0} \mathbb{E}(\langle \nu_t^\gamma, \mathbf{1}\rangle^p) \leq C'.$$

Observe that, if there is no mutation ($\mu \equiv 0$) and $\nu_0 = X_0\delta_x + Y_0\delta_y$ (dimorphic population), then, for any $t \geq 0$, $\nu_t = X_t\delta_x + Y_t\delta_y$, where $(X_t, Y_t)$ is the two-type logistic branching process of Section 2 with

$$B = \begin{pmatrix} b(x) \\ b(y) \end{pmatrix}, \qquad C = \begin{pmatrix} c(x,x) & c(x,y) \\ c(y,x) & c(y,y) \end{pmatrix} \qquad \text{and} \qquad D = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Accordingly, we will denote by $u_{n,m}(x,y)$ the probability of fixation of the mutant type $y$. More generally, if at time $t$ the population is composed of $n$ different traits, the process follows a $n$-type logistic branching process until the next mutation.

Let us also point out that, when there is only one individual in the population ($\nu_t = \delta_x$ for some $x \in \mathcal{X}$), the death rate (39) equals 0, so that extinction is impossible. Then, as observed in the beginning of Section 2, if $\mu \equiv 0$ and if the initial population is monomorphic ($\nu_0 = n\delta_x$ for some $n \in \mathbb{N}^*$ and $x \in \mathcal{X}$), the number of individuals $\langle \nu_t, \mathbf{1} \rangle$ is a positive recurrent Markov chain converging in distribution to $\xi_{\theta(x)}$, where $\xi_{\theta(x)}$ is a Poisson random variable of parameter $\theta(x) := b(x)/c(x,x)$ conditioned on being nonzero

$$\mathbb{P}(\xi_{\theta(x)} = i) = \frac{e^{-\theta(x)}}{1 - e^{-\theta(x)}} \frac{\theta(x)^i}{i!}, \quad i \geq 1. \tag{41}$$

Recall that $\mathbb{E}(\xi_\theta) = \theta/(1 - \exp(-\theta))$.

### 5.2 Limit of rare mutations and the trait substitution sequence in finite population

Here, we study the behaviour of the process $\nu^\gamma$ in the limit of rare mutations ($\gamma \to 0$) on the mutation timescale $t/\gamma$.

Let us introduce the following strong form of convergence in law. We will say that a sequence of random variables $(X_n)$ **converges strongly in law** to a random variable $Y$ if and only if $\mathbb{E}(f(X_n)) \to \mathbb{E}(f(Y))$ when $n \to \infty$ for any bounded measurable real function $f$.

Fix $x \in \mathcal{X}$. For $\gamma \in (0,1)$, let $\nu_0^\gamma = N_0^\gamma \delta_x$ where the $\mathbb{N}^*$-valued random variables $N_0^\gamma$ satisfy $\sup_{\gamma \in (0,1)} \mathbb{E}((N_0^\gamma)^p) < \infty$ for some $p > 1$.

**Theorem 5.2** *For any positive $0 < t_1 < \ldots < t_n$, the $n$-tuple $(\nu^\gamma_{t_1/\gamma}, \ldots, \nu^\gamma_{t_n/\gamma})$ converges strongly in law to $(\zeta_{t_1}, \ldots, \zeta_{t_n})$ where $\zeta_{t_i} = N_{t_i}\delta_{Z_{t_i}}$ such that*

**(1)** *$(Z_t; t \geq 0)$ is a Markov jump process on $\mathcal{X}$ with initial value $Z_0 = x$ and with infinitesimal generator*

$$A\varphi(x) = \int_{\mathcal{X}} (\varphi(x+h) - \varphi(x))\mu(x)b(x)\frac{\theta(x)}{1 - e^{-\theta(x)}} \sum_{n=1}^{+\infty} u_{n,1}(x, x+h)\frac{e^{-\theta(x)}\theta(x)^{n-1}}{(n-1)!} M(x, dh) \tag{42}$$

**(2)** *Conditional on $(Z_{t_1}, \ldots, Z_{t_n}) = (z_1, \ldots, z_n)$, the $N_{t_i}$ are independent and respectively distributed as $\xi_{\theta(z_i)}$.*

Observe that the TSS generator (45) may be written as

$$A\varphi(x) = \int_{\mathcal{X}} (\varphi(x+h) - \varphi(x))\beta(x)\chi(x, x+h)M(x, dh)$$

where

$$\beta(x) := \mu(x)b(x)\frac{\theta(x)}{1 - e^{-\theta(x)}} \tag{43}$$

is the production rate (on the timescale $t/\gamma$) of mutants in a stationary $x$-type population and

$$\chi(x, y) := e^{-\theta(x)} \sum_{n=1}^{+\infty} u_{n,1}(x, y)\frac{\theta(x)^{n-1}}{(n-1)!}M(x, dh) \tag{44}$$

will be interpreted as the fixation probability of a mutant individual with trait $y$ in a stationary $x$-type population.

This result shows that, in the limit of rare mutations, on the mutation timescale, the population is always monomorphic and that the trait of the population evolves as a jump process, where a jump corresponds to the appearance and fixation of a mutant subpopulation. The process $Z$ is the TSS in our model (case of finite populations).

Let us denote by $\tau_n$ the $n$-th mutation time ($\tau_0 = 0$), by $\rho_n$ the first time after time $\tau_n$ when the population gets monomorphic, and by $V_n$ the single trait value surviving at time $\rho_n$ ($\rho_0 = 0$ if the initial population is monomorphic). With this notation, we can state the following result, adressing the main biological issue of Theorem 5.2, namely the convergence of the support of the measure $\nu^{\gamma}_{\cdot/\gamma}$ to the process $Z$.

**Theorem 5.3** *The process* $(S^{\gamma}_t; t \geq 0)$ *defined as*

$$S^{\gamma}_t = \sum_{n=0}^{\infty} V_n \mathbf{1}_{\{\rho_n \leq t/\gamma < \rho_{n+1}\}}$$

*converges in law for the Skorohod topology on* $\mathbb{D}(\mathbb{R}_+, \mathcal{X})$ *when* $\gamma \to 0$ *to the process* $(Z_t; t \geq 0)$ *with initial state* $Z_0 = x$ *and with infinitesimal generator (42).*

Observe that such a convergence for the measure $\nu^{\gamma}_{\cdot/\gamma}$ cannot hold because the population size $N_t$ in Theorem 5.2 is not a cÃ dlÃ g process.

The proofs of the two preceding theorems are put to Subsection 6.1.

## 5.3 Limit of small mutation steps and the 'canonical diffusion of adaptive dynamics' in finite populations

Here, we want to apply a limit of small jumps to the TSS process with generator (42) of the previous subsection, in order to obtain the equivalent of the canonical equation of adaptive dynamics in finite populations. Let $\mathcal{C}^k_b$ be the set of $k$ times differentiable functions with bounded $i$-th derivatives, $1 \leq i \leq k$. We will need in this section the following additional assumptions:

- $\mathcal{X} = \mathbb{R}^k$ for simplicity

- $b$ and $c$ are in $\mathcal{C}_b^2$
- the mutation kernels $M(x,\cdot)$ satisfy
  - for any $x\in\mathbb{R}^k$, $M(x,\cdot)$ has 0 expectation, i.e. $\int_{\mathbb{R}^k} hM(x,dh)=0$
  - the covariance matrix of $M(x,\cdot)$ has Lipschitz entries and is uniformly elliptic in $x$, i.e. there is a positive constant $C$ such that $\int_{\mathbb{R}^k}(s'h)^2M(x,dh)\geq C\|s\|^2$ for any $s\in\mathbb{R}^k$.
  - the third order moments of $M(x,\cdot)$ are uniformly bounded in $x$.

Recall that there is a symmetric matrix $\sigma(x)$ such that $\sigma(x)\sigma(x)'=\sigma(x)^2$ is the covariance matrix of $M(x,\cdot)$ which is called its square root. Then its uniform ellipticity ensures that $\sigma(x)$ has also Lipschitz entries in $x$.

The limit of small jumps is obtained by introducing a parameter $\epsilon>0$ and replacing the mutation law $M(x,dh)$ by its image by the application $h\mapsto\epsilon h$, $M(x,dh/\epsilon)$. Of course, this scaling of the size of the jumps has to be combined with a scaling of time in order to observe a non-trivial limit. This leads (after a change of variable) to the following generator of the rescaled TSS $(Z_t^\epsilon;t\geq 0)$:

$$A_\epsilon\varphi(x)=\frac{1}{\epsilon^2}\int_{\mathbb{R}^k}(\varphi(x+\epsilon h)-\varphi(x))\beta(x)\chi(x,x+\epsilon h)M(x,dh) \tag{45}$$

where the mutant production rate in a stationary $x$-type population, $\beta(x)$, is defined in (43), and where the fixation probability of a mutant of trait $y$ appearing in a stationary $x$-type population, $\chi(x,y)$, is defined in (44). The factor $1/\epsilon^2$ in (45) corresponds to the correct timescaling, as will appear below.

We will use the notation $v^\iota(x)$ and $g^\iota(x)$ ($\iota=\lambda,\delta,\alpha,\varepsilon,\sigma$) for, respectively, the first-order derivatives of the fixation probability in the vicinity of neutrality (all individuals have trait $x$), and the corresponding invasibility coefficients. Similarly, we will denote by $q_n^{(2)}(x)$, $\kappa(x)$ and $\phi_n(x)$ the quantities appearing in Sections 4.5 and 4.6 for $\lambda$, $\alpha$ and $\sigma$-invasibilities, and in Section 4.8 for the $\delta$-invasibility.

It follows from Theorem 3.3 that the function $\chi$ is $\mathcal{C}_b^2$. Observe also that

$$\frac{d}{dy}\Big|_{y=x}\begin{pmatrix} c(x,x) & c(x,y)\\ c(y,x) & c(y,y)\end{pmatrix}=\nabla_1 c(x,x)\begin{pmatrix}0&0\\1&1\end{pmatrix}+\nabla_2 c(x,x)\begin{pmatrix}0&1\\0&1\end{pmatrix}$$

where, for bivariate $f$, $\nabla_i f$ is the gradient of $f$ w.r.t. the $i$-th variable ($i=1,2$). Then, by Theorem 4.1,

$$\chi(x,x)=e^{-\theta(x)}\sum_{n\geq 1}\frac{\theta(x)^{n-1}}{(n+1)(n-1)!}=\frac{e^{-\theta(x)}-1+\theta(x)}{\theta(x)^2}$$

and

$$\nabla_2\chi(x,x)=e^{-\theta(x)}(a_\lambda(x)\nabla b(x)-a_\delta(x)\nabla_1 c(x,x)+a_\alpha(x)\nabla_2 c(x,x))$$

where, for $\iota=\lambda,\delta,\alpha$,

$$a_\iota(x)=\sum_{n=1}^{\infty}v_{n,1}^\iota(x)\frac{\theta(x)^{n-1}}{(n-1)!}=\sum_{n=1}^{\infty}\frac{ng_{n+1}^\iota(x)\theta(x)^{n-1}}{(n+1)^2(n-1)!}.$$

**Remark 5.4** *Observe that $\varepsilon$-invasibilities do not appear in this computation, because of the symmetry between resident and mutant types in the competition kernel. One could include $\varepsilon$-invasibilities in the formula of $\nabla_2\chi(x,x)$ by assuming a competition matrix of the form*

$$\begin{pmatrix} c_1(x,x) & c_1(x,y) \\ c_2(y,x) & c_2(y,y) \end{pmatrix}$$

*for some functions $c_1$ and $c_2$ coinciding on the diagonal. Such an asymmetry between resident and mutant would not be unrealistic biologically and can be explained by the resident constructing its own niche. This ecological adaptation of the resident to its medium would then result in a difference in the competition felt by $x$ from $y$ according whether $x$ is the resident or not.*

Now, fix a function $\varphi$ in $\mathcal{C}^3_b$. For any $x,h\in\mathbb{R}^k$ and $\epsilon>0$, there exists $0\leq\epsilon'\leq\epsilon$ such that

$$\varphi(x+\epsilon h)-\varphi(x)=\epsilon h'\nabla\varphi(x)+\frac{\epsilon^2}{2}h'H\varphi(x+\epsilon' h)h$$

where $H\varphi(y)$ denotes the Hessian matrix of $\varphi$ at $y$, and there exists $0\leq\epsilon''\leq\epsilon$ such that

$$\chi(x,x+\epsilon h)=\chi(x,x)+\epsilon h'\nabla_2\chi(x,x+\epsilon'' h).$$

Therefore, using the fact that $H\varphi$ and $\nabla_2 g$ are bounded Lipschitz functions, it takes only elementary computations to prove that

$$\begin{aligned}(\varphi(x+\epsilon h)-\varphi(x))\chi(x,x+\epsilon h)&=\epsilon(h'\nabla\varphi(x))\chi(x,x)+\epsilon^2(h'\nabla\varphi(x))(h'\nabla_2\chi(x,x))\\&\quad+\frac{\epsilon^2}{2}(h'H\varphi(x)h)\chi(x,x)+O(\epsilon^3\|h\|^3)\end{aligned}$$

where the $O(\epsilon^3\|h\|^3)$ is uniform in $x\in\mathbb{R}^k$. Now, since the mutation kernel has zero expectation,

$$\int_{\mathbb{R}^k}(h'\nabla\varphi(x))\chi(x,x)M(x,dh)=0.$$

Combining these results, thanks to boundedness of the third-order moments of the mutation kernel, we easily obtain for any $\varphi\in\mathcal{C}^3_b$, that $A_\epsilon\varphi$ converges uniformly to the function $A_0\varphi$ defined as

$$\begin{aligned}A_0\varphi(x)=&\int_{\mathbb{R}^k}(h'\nabla\varphi(x))\beta(x)e^{-\theta(x)}h'(a_\lambda(x)\nabla b(x)-a_\delta(x)\nabla_1 c(x,x)+a_\alpha(x)\nabla_2 c(x,x))M(x,dh)\\&+\frac{1}{2}\int_{\mathbb{R}^k}(h'H\varphi(x)h)\beta(x)\frac{e^{-\theta(x)}-1+\theta(x)}{\theta(x)^2}M(x,dh).\end{aligned}\tag{46}$$

In view of this, the following theorem is natural. Recall that $\sigma(x)$ is the symmetric square root matrix of the covariance matrix of $M(x,\cdot)$, which is Lipschitz in $x$, and that $\chi(z,z)=(\exp(-\theta(z))-1+\theta(z))/\theta(z)^2$.

**Theorem 5.5** *If the family $(Z^\epsilon_0)_{\epsilon>0}$ has bounded first-order moments and converges in law when $\epsilon\to 0$ to a random variable $Z_0$, then the process $Z^\epsilon$ converges in law for the Skorohod topology on $\mathbb{D}(\mathbb{R}_+,\mathbb{R}^k)$ to the diffusion process $(Z_t;t\geq 0)$ with initial state $Z_0$ unique solution to the stochastic differential equation*

$$dZ_t=r(Z_t)dt+\sqrt{\beta(Z_t)\chi(Z_t,Z_t)}\sigma(Z_t)dB_t\tag{47}$$

*where $B$ is a standard $k$-dimensional Brownian motion and*

$$r(z)=\beta(z)e^{-\theta(z)}\int_{\mathbb{R}^k}M(z,dh)\big(h'\big[a_\lambda(z)\nabla b(z)-a_\delta(z)\nabla_1c(z,z)+a_\alpha(z)\nabla_2c(z,z)\big]\big)h.$$

**Remark 5.6** *In the case where $\mathcal{X}\neq\mathbb{R}^k$, this result is still valid, apart from the following technical difficulties. First, for the process $Z^\epsilon$ to be well-defined, one needs to assume that scaling the mutation law $M(x,dh)$ cannot drive $Z^\epsilon$ out of $\mathcal{X}$. This is true for example, when $\varepsilon\leq 1$, if $\mathcal{X}$ is convex, or if Supp$(M(x,dh))$ is convex for any $x\in\mathcal{X}$. Second, uniqueness in law has to hold for the diffusion with generator $A_0$. For example, one can ensure the existence of a Lipschitz factorization $\sigma(x)\sigma(x)'$ of the covariance matrix of $M(x,\cdot)$ by assuming that the function from $\mathcal{X}$ to the set of nonnegative symmetric matrices mapping $x$ to the covariance matrix of $M(x,\cdot)$ can be extended to $\mathbb{R}^k$ in a $\mathcal{C}^2$ fashion [14].*

**Remark 5.7** *In the case where the mutation step law $M(x,\cdot)$ has non-zero expectation, the calculation above shows that the first-order term in $A_\epsilon$ does not vanish, so that the correct timescaling is $1/\epsilon$ (instead of $1/\epsilon^2$), and the TSS process $Z^\epsilon$ can be shown to converge to the solution of the deterministic ODE*

$$\frac{dz}{dt}=\beta(z)\chi(z,z)\int_{\mathbb{R}^k}hM(z,dh).$$

*In this case, the main force driving evolution is the mutation bias. The mutation rate $\beta(x)$ and the fixation probability $\chi(x,x)$ only affect the speed of evolution.*

Theorem 5.5 gives the equivalent of the canonical equation of adaptive dynamics [8, 4] when the population is finite. It is no longer a deterministic ODE, but a diffusion process, in which the genetic drift remains present, as a consequence of the population finiteness and of the asymptotic of weak selection ($\epsilon\to 0$). Diffusion processes have long been used to describe evolution in biology (see, among many others, [13, 12, 19, 25, 16]). Our process provides such a model, which is grounded on a microscopic precise desription of the population dynamics, in a realistic way. In particular, because of the quadratic density-dependence, the population size is not fixed and may fluctuate.

The diffusion part in (47) gives the strength of the genetic drift, which square is proportional to the mutation rate $\beta(x)$ and to the covariance matrix of the mutation step law $M(x,dh)$. The drift part gives the expression of the deterministic strength driving evolution, which is often related in macroscopic evolutionary models to a fitness gradient. In our case, the fitness is given by the function $\chi$. Indeed, $\chi(x,y)$ is the fitness of a mutant trait $y$ in a stationary resident population with trait $x$ (in the sense proposed by Metz et al. [27]), and it appears in the deterministic part of (47) as the gradient of $\chi$ with respect to the second variable, in a similar way as in the standard canonical equation of adaptive dynamics [8, 4]. Therefore, the 'hill-climbing' process of evolution occurs here, as in the classical models of adaptive dynamics, in a fitness landscape $y\mapsto\chi(x,y)$ that depends on the current state $x$ of the population. Observe also that the deterministic drift term in (47), as in the canonical equation of adaptive dynamics, is proportional to the mutation rate $\beta(x)$, the covariance matrix of the mutation step law $M(x,dh)$ and the gradient of the fixation probability.

The coefficients appearing in the deterministic part of (47) can be expressed as follows.

**Proposition 5.8** *The coefficients $a_\iota$ for $\iota = \lambda, \delta, \alpha$ can be expressed in terms of the microscopic parameters $b(x), \theta(x), q_n^{(2)}(x), \kappa(x)$ and $\phi_n(x)$ as*

$$a_\lambda(x) = \frac{e^{\theta(x)}(\theta(x)^2 - 3\theta(x) + 4) - \theta(x) - 4}{2b(x)\theta(x)^2} \tag{48}$$

$$a_\alpha(x) = \frac{e^{\theta(x)}(\theta(x)^2 - \theta(x) + 2) - \theta(x) - 2}{2b(x)\theta(x)^2} - \frac{1}{\kappa(x)\theta(x)} \sum_{n=1}^{\infty} \frac{n q_{n+1}^{(2)}(x)\theta(x)^{n-1}}{(n+1)!} \tag{49}$$

$$a_\delta(x) = \sum_{k \geq 1} \frac{\phi_k(x)}{\theta(x)^{k+2}} \int_0^{\theta(x)} u^{k-1}(e^u(u^2 - u + 1) - 1)du \tag{50}$$

$$= \frac{1}{\theta(x)^3} \int_0^{\theta(x)} (e^u(u^2 - u + 1) - 1)\pi_x\left(\frac{u}{\theta(x)}\right) du \tag{51}$$

*where for $k \geq 1$,*

$$\int_0^\theta u^{k-1}(e^u(u^2-u+1)-1)du = e^\theta \left(\theta^{k+1} - (k+2)\theta^k + (k+1)^2(k-1)! \sum_{i=0}^{k-1} \frac{(-1)^i \theta^{k-i-1}}{(k-i-1)!}\right) - (-1)^{k-1}(k+1)^2(k-1)! - \frac{\theta^k}{k} \tag{52}$$

*and for any $v \in [0, 1)$, $\pi_x(v) := \sum_{k\geq 1} \phi_k(x)v^{k-1}$. Moreover, $\pi_x$ is solution on $[0, 1)$ to*

$$u^2(1-u)\pi_x''(u) + u(\theta(x)u(1-u) + 2 - 3u)\pi_x'(u) - 2\pi_x(u) + \frac{\theta(x)}{b(x)} = 0. \tag{53}$$

The proofs of Theorem 5.5 and Proposition 5.8 can be found in Subsection 6.2.

**Example.** Let us consider a one-dimensional trait $x \in \mathbb{R}$ in a population undergoing symmetric competition $c(x, y) = c(|x-y|)$. This type of competition kernel has been considered in numerous earlier works, see e.g. [7]. As a consequence, $\partial c/\partial x(x, x) = \partial c/\partial y(x, x) = 0$. We may and will assume that $c(0) = 1$. We still denote by $\sigma(x)$ the standard deviation of the mutation kernel $M(x, \cdot)$. Then, thanks to Theorem 5.5 and Proposition 5.8, the canonical diffusion of adaptive dynamics is given by

$$dZ_t = r(Z_t)dt + \sigma(Z_t)\mu(Z_t)^{1/2}\left(\frac{b(Z_t)}{1 - e^{-b(Z_t)}} - 1\right)^{1/2} dB_t$$

where

$$r(x) = \frac{\mu(x)\sigma(x)^2}{2}\left(1 + \frac{4}{b(x)} + \frac{b(x) - 4}{1 - e^{-b(x)}}\right) b'(x).$$

In forthcoming work, this diffusion and other examples will be investigated.

# 6 Proofs

## 6.1 Proofs of Theorems 5.2 and 5.3

These proofs rely on the following three lemmas. The first one states that there is no accumulation of mutations on the timescale $t/\gamma$. The second one gives the limiting laws of

$\gamma\tau_1$ and of the population size at time $\tau_1$. The last one gives the behaviour of $\rho_0$ and $V_0$ when the initial population is monomorphic.

**Lemma 6.1** *Fix $C, \eta > 0$. There is $\varepsilon > 0$ such that for any $\gamma \in (0,1)$,*

$$\mathbb{E}(\langle \nu_0^\gamma, \mathbf{1}\rangle) \leq C \quad \Rightarrow \quad \forall t \geq 0, \quad \mathbb{P}\left(\exists n \in \mathbb{N}^* : \frac{t}{\gamma} \leq \tau_n \leq \frac{t+\varepsilon}{\gamma}\right) < \eta. \tag{54}$$

*Moreover, for any $\eta > 0$ and $t \geq 0$, there exists $n \in \mathbb{N}^*$ such that, for any $\gamma \in (0,1)$,*

$$\mathbb{E}(\langle \nu_0^\gamma, \mathbf{1}\rangle) \leq C \quad \Rightarrow \quad \mathbb{P}(\tau_n \leq t/\gamma) < \eta. \tag{55}$$

**Lemma 6.2** *Assume $\nu_0^\gamma = n\delta_x$.*

(a) *The couple $(\gamma\tau_1, \langle \nu_{\tau_1-}^\gamma, \mathbf{1}\rangle)$ converges in law to a couple of independent random variables $(T, N)$ where $T$ is an exponentially distributed random variable with parameter $\beta(x)$ defined in (43) and the law of $N$ is the following size-biased version of $\xi_{\theta(x)}$*

$$\mathbb{P}(N = k) = \frac{k\mathbb{P}(\xi_{\theta(x)} = k)}{\mathbb{E}(\xi_{\theta(x)})} = e^{-\theta(x)} \frac{\theta(x)^{k-1}}{(k-1)!}, \tag{56}$$

*which is the law of a shifted Poisson random variable of parameter $\theta(x)$.*

(b) *For any $p \geq 1$, $\sup_{\gamma\in(0,1)} \mathbb{E}_{n\delta_x}^\gamma(\langle \nu_{\tau_1}, \mathbf{1}\rangle^p) < \infty$.*

**Lemma 6.3** *Assume $\nu_0^\gamma = n\delta_x + \delta_y$ (with $y \neq x$). Then*

(a) *$\gamma\rho_0 \to 0$ in probability and $\mathbb{P}^\gamma(\rho_0 < \tau_1) \to 1$ when $\gamma \to 0$.*

(b) *For any $p \in [1,2]$, $\sup_{\gamma\in(0,1)} \mathbb{E}_{n\delta_x+\delta_y}\left(\langle \nu_{\rho_0}^\gamma, \mathbf{1}\rangle^p \mathbf{1}_{\{\rho_0<\tau_1\}}\right) < \infty$.*

(c) *$\lim_{\gamma\to 0} \mathbb{P}^\gamma(V_0 = y) = 1 - \lim_{\gamma\to 0} \mathbb{P}^\gamma(V_0 = x) = u_{n,1}(x,y)$.*

**Proof of Lemma 6.1.** Fix $C > 0$ and assume $\mathbb{E}(\langle \nu_0^\gamma, \mathbf{1}\rangle) \leq C$. By Proposition 5.1, there exists a constant $C'$ such that $\mathbb{E}(\langle \nu_t^\gamma, \mathbf{1}\rangle) \leq C'$ for any $t \geq 0$ and $\gamma > 0$. Therefore, it is sufficient to show Lemma 6.1 for $t = 0$.

Now, when the total population size is $n$, the total mutation rate in the population is bounded by $\gamma\bar{b}n$, so that the number of mutations $M_t$ between times 0 and $t$ is dominated by a point process with intensity $\gamma\bar{b}\langle \nu_s, \mathbf{1}\rangle ds$. More precisely (using for example the coupling argument of Theorem 2 in [3]) $M_t$ can be shown to be smaller than $\int_0^t \int_{\mathbb{N}^*} \mathbf{1}_{\{i \leq \langle \nu_{s-}, \mathbf{1}\rangle\}} P(di, ds)$, where $P(di, ds)$ is a Poisson point process on $\mathbb{N}^* \times \mathbb{R}_+$ with intensity measure $\gamma\bar{b}\sum_{k\geq 1} \delta_k(di)\, ds$. Therefore,

$$\mathbb{P}(M_{\varepsilon/\gamma} \geq 1) \leq \mathbb{E}(M_{\varepsilon/\gamma}) \leq \gamma\bar{b} \int_0^{\varepsilon/\gamma} \mathbb{E}(\langle \nu_s^\gamma, \mathbf{1}\rangle) ds \leq \varepsilon\bar{b}C',$$

which concludes the proof of (54).

Similarly, for $t \geq 0$, $\mathbb{P}(M_{t/\gamma} \geq n) \leq t\bar{b}C'/n$, which implies (55). □

**Proof of Lemma 6.2.** Fix $\gamma\in(0,1)$ and assume that $\nu^\gamma_0=n\delta_x$. The number of mutations issued from individuals of type $x$ before time $t$ can be constructed as $\int_0^t\int_{\mathbb{N}^*}\mathbf{1}_{\{i\leq Y^\gamma_{s-}\}}P^\gamma(di,ds)$ where $Y^\gamma_t=\langle\nu^\gamma_t,\mathbf{1}_{\{x\}}\rangle$ and $P^\gamma(di,ds)$ is a Poisson point process on $\mathbb{N}^*\times\mathbb{R}_+$ with intensity measure $\gamma\mu(x)b(x)\sum_{k\geq 1}\delta_k(di)\,ds$ independent of $(Y^\gamma_t;t\geq 0)$. On another hand, on the event $\{t<\tau_1\}$, $Y^\gamma_t=X^\gamma_t$, where $X^\gamma$ is a scalar logistic branching process with dynamical parameters $((1-\gamma\mu(x))b(x),c(x,x),0)$ and with initial state $X^\gamma_0=n$. Therefore, for any bounded function $f:\mathbb{N}^*\to\mathbb{R}$ and for any $t\geq 0$,

$$\mathbb{E}^\gamma(f(\langle\nu^\gamma_{t/\gamma},\mathbf{1}\rangle);\gamma\tau_1>t)=\mathbb{E}\left[f(X^\gamma_{t/\gamma})\exp\left(-\gamma\mu(x)b(x)\int_0^{t/\gamma}X^\gamma_{s-}ds\right)\right]. \tag{57}$$

Now, fix $n\geq 1$, $t>0$ and a bounded function $f:\mathbb{N}^*\to\mathbb{R}$. On the one hand, since $\langle\nu^\gamma_t,\mathbf{1}\rangle$ is a piecewise constant process,

$$\mathbb{E}^\gamma(f(\langle\nu^\gamma_{\tau_1-},\mathbf{1}\rangle)\mathbf{1}_{\{\gamma\tau_1\leq t\}})=\lim_{n\to+\infty}\mathbb{E}^\gamma\left(\sum_{k=0}^{[2^nt]}f(\langle\nu^\gamma_{k2^{-n}/\gamma},\mathbf{1}\rangle)\mathbf{1}_{\{2^{-n}k<\gamma\tau_1\leq 2^{-n}(k+1)\}}\right)$$

and on the other hand, by (57),

$$\begin{aligned}
&\lim_{n\to+\infty}\sum_{k=0}^{[2^nt]}\mathbb{E}^\gamma\left[f(\langle\nu^\gamma_{k2^{-n}/\gamma},\mathbf{1}\rangle)\mathbf{1}_{\{2^{-n}k<\gamma\tau_1\leq 2^{-n}(k+1)\}}\right]\\
&\quad=\lim_{n\to+\infty}\sum_{k=0}^{[2^nt]}\mathbb{E}\left[f(X^\gamma_{k2^{-n}/\gamma})\left(e^{-\gamma\mu(x)b(x)\int_0^{k2^{-n}/\gamma}X^\gamma_s ds}-e^{-\gamma\mu(x)b(x)\int_0^{(k+1)2^{-n}/\gamma}X^\gamma_s ds}\right)\right]\\
&\quad=\mu(x)b(x)\lim_{n\to+\infty}2^{-n}\sum_{k=0}^{[2^nt]}\mathbb{E}\left[f(X^\gamma_{k2^{-n}/\gamma})X^\gamma_{k2^{-n}/\gamma}e^{-\gamma\mu(x)b(x)\int_0^{k2^{-n}/\gamma}X^\gamma_s ds}\right]
\end{aligned}\tag{58}$$

where the last equality follows from Lebesgue's theorem and the facts that $|1-e^{-u}-u|\leq u^2/2$, that $\sup_{t\geq 0}\mathbb{E}((X^\gamma_t)^2)\leq C<\infty$ and that

$$\left|\mathbb{E}\left(\int_t^{t+\varepsilon}X^\gamma_s ds-\varepsilon X^\gamma_t\right)\right|\leq\int_t^{t+\varepsilon}\mathbb{E}(|X^\gamma_s-X^\gamma_t|)ds=o(\varepsilon)$$

uniformly in $t\geq 0$. This last equation is an immediate consequence of the fact that, for any $s<t$ and $M>0$,

$$\begin{aligned}
\mathbb{E}(|X^\gamma_s-X^\gamma_t|)&\leq\mathbb{E}(X^\gamma_s+X^\gamma_t;X^\gamma_s\neq X^\gamma_t)\leq\mathbb{E}(X^\gamma_s;X^\gamma_s>M)+\mathbb{E}(X^\gamma_t;X^\gamma_t>M)\\
&\quad+2M\mathbb{P}(\text{no jump of the process }X^\gamma\text{ occurred on the time interval }[s,t])\\
&\leq\frac{2C}{M}+2M\left(1-e^{-M(b(x)+c(x,x)(M-1))(t-s)}\right).
\end{aligned}$$

Define $\varphi(t):=\mathbb{E}[f(X^\gamma_t)X^\gamma_t\exp(-\gamma\mu(x)b(x)\int_0^t X^\gamma_s ds)]$. Distinguishing as above between the cases where $X^\gamma_s>M$, $X^\gamma_t>M$, $X^\gamma_s=X^\gamma_t$ and $X^\gamma_s\neq X^\gamma_t$, one can easily prove that $\varphi(s)\to\varphi(t)$ when $s\to t$, i.e. that $\varphi$ is continuous. Thus, the Riemann sum in the right hand side of (58) converges to the corresponding integral, and for any $\gamma\in(0,1)$,

$$\mathbb{E}^\gamma(f(\langle\nu^\gamma_{\tau_1-},\mathbf{1}\rangle)\mathbf{1}_{\{\gamma\tau_1\leq t\}})=\mu(x)b(x)\int_0^t\mathbb{E}\left(f(X^\gamma_{s/\gamma})X^\gamma_{s/\gamma}e^{-\gamma\mu(x)b(x)\int_0^{s/\gamma}X^\gamma_u du}\right)ds. \tag{59}$$

Now, since the individual birth rate of the logistic branching process $X^\gamma$ is decreasing with respect to $\gamma$, all the processes $X^\gamma$ can be constructed on a same space in such a way that, for any $0 \leq \gamma_1 \leq \gamma_2 \leq 1$ and $t \geq 0$, $X_t^{\gamma_2} \leq X_t^{\gamma_1}$. Therefore, assuming that $f$ is a nonnegative and nondecreasing function, for any $0 \leq \gamma \leq \gamma_0 \leq 1$,

$$\begin{aligned}\mathbb{E}\left(f(X^{\gamma_0}_{s/\gamma})X^{\gamma_0}_{s/\gamma}e^{-\gamma\mu(x)b(x)\int_0^{s/\gamma}X^0_u du}\right) &\leq \mathbb{E}\left(f(X^{\gamma}_{s/\gamma})X^{\gamma}_{s/\gamma}e^{-\gamma\mu(x)b(x)\int_0^{s/\gamma}X^\gamma_u du}\right)\\ &\leq \mathbb{E}\left(f(X^{0}_{s/\gamma})X^{0}_{s/\gamma}e^{-\gamma\mu(x)b(x)\int_0^{s/\gamma}X^{\gamma_0}_u du}\right).\end{aligned}$$

For any $\gamma \in (0,1)$, let us denote by $\xi^{(\gamma)}$ the random variable $\xi_{(1-\gamma\mu(x))\theta(x)}$. By the ergodic theorem for positive recurrent Markov chains, $\gamma\int_0^{s/\gamma}X_u^0 du \to s\mathbb{E}(\xi^{(0)})$ when $\gamma \to 0$. Then, since $\sup_{t\geq 0}\mathbb{E}((X_t^{\gamma_0})^2) < \infty$, for any $s > 0$,

$$\lim_{\gamma\to 0}\mathbb{E}\left(f(X^{\gamma_0}_{s/\gamma})X^{\gamma_0}_{s/\gamma}e^{-\gamma\mu(x)b(x)\int_0^{s/\gamma}X^0_u du}\right) = \mathbb{E}\left(f(\xi^{(\gamma_0)})\xi^{(\gamma_0)}e^{-\mu(x)b(x)\mathbb{E}(\xi^{(0)})s}\right)$$

and, by Lebesgue's theorem,

$$\limsup_{\gamma\to 0}\ \mathbb{E}^\gamma(f(\langle\nu^\gamma_{\tau_1-},\mathbf{1}\rangle)\mathbf{1}_{\{\gamma\tau_1\leq t\}}) \leq \frac{\mathbb{E}(f(\xi^{(\gamma_0)})\xi^{(\gamma_0)})}{\mathbb{E}(\xi^{(0)})}\left(1-e^{-\mu(x)b(x)\mathbb{E}(\xi^{(0)})t}\right).$$

Similarly,

$$\liminf_{\gamma\to 0}\ \mathbb{E}^\gamma(f(\langle\nu^\gamma_{\tau_1-},\mathbf{1}\rangle)\mathbf{1}_{\{\gamma\tau_1\leq t\}}) \geq \frac{\mathbb{E}(f(\xi^{(0)})\xi^{(0)})}{\mathbb{E}(\xi^{(\gamma_0)})}\left(1-e^{-\mu(x)b(x)\mathbb{E}(\xi^{(\gamma_0)})t}\right).$$

Letting $\gamma_0$ decrease to 0 finally yields

$$\lim_{\gamma\to 0}\ \mathbb{E}^\gamma(f(\langle\nu^\gamma_{\tau_1-},\mathbf{1}\rangle)\mathbf{1}_{\{\gamma\tau_1\leq t\}}) = \frac{\mathbb{E}(f(\xi^{(0)})\xi^{(0)})}{\mathbb{E}(\xi^{(0)})}\left(1-e^{-\mu(x)b(x)\mathbb{E}(\xi^{(0)})t}\right). \qquad (60)$$

The same method applies to the case of nonincreasing nonnegative functions $f$, which completes the proof of Lemma 6.2(a).

Lemma 6.2(b) can be obtained by taking $f(x) = x^p \wedge K$ in (59), then letting first $K$ go to infinity and next $t$ to infinity (using the fact that $\int_0^{s/\gamma}X_u^\gamma du \geq s/\gamma$ since $X_u^\gamma \geq 1$ for any $u \geq 0$), and finally using Proposition 5.1. □

**Proof of Lemma 6.3.** Before the first mutation, $\nu_t^\gamma = X_t^\gamma\delta_x + Y_t^\gamma\delta_y$ where $(X_t^\gamma, Y_t^\gamma)$ is a two-type logistic branching process with dynamical parameters

$$B = \begin{pmatrix}(1-\gamma\mu(x))b(x)\\(1-\gamma\mu(y))b(y)\end{pmatrix},\qquad C = \begin{pmatrix}c(x,x) & c(x,y)\\ c(y,x) & c(y,y)\end{pmatrix},\qquad D = \begin{pmatrix}0\\0\end{pmatrix}.$$

On the event $\{\tau_1 > \rho_0\}$, $V_0 = y$ if and only if there is fixation in the process $(X^\gamma, Y^\gamma)$, $V_0 = x$ otherwise, and $\rho_0 = T$.

Now, by Lemma 6.1, for any $\eta > 0$, there exists $\varepsilon > 0$ such that $\mathbb{P}(\tau_1 > \varepsilon/\gamma) \geq 1-\eta$. Since $\mathbb{P}_{n,1}(T < \infty) = 1$, this implies easily (a). It is then elementary to deduce from the continuity of the fixation probability with respect to the components of the matrix $B$ that (c) holds. Finally, (b) follows from the observation that

$$\mathbb{E}_{n\delta_x+\delta_y}\left(\langle\nu^\gamma_{\rho_0},\mathbf{1}\rangle^p\mathbf{1}_{\{\rho_0<\tau_1\}}\right) \leq \mathbb{E}_{n,1}((X_T^\gamma)^p + (Y_T^\gamma)^p)$$

and from Proposition 2.1(c). □

**Proof of Theorem 5.2.** Observe that the generator $A$ of the process $Z$, defined in (42) can be written as

$$A\varphi(x) = \int_{\mathcal{X}} (\varphi(x+h) - \varphi(x))\beta(x)\kappa(x,dh), \tag{61}$$

where $\beta(x)$ has been defined in (43) and where $\kappa(x,dh)$ is the **probability** measure on $\mathcal{X} - x$ defined by

$$\kappa(x,dh) = e^{-\theta(x)} \sum_{n=1}^{\infty} u_{n,1}(x,x+h)\frac{\theta(x)^{n-1}}{(n-1)!}M(x,dh) \\ + e^{-\theta(x)} \left( \int_{\mathbb{R}^k} \sum_{n=1}^{\infty} (1 - u_{n,1}(x,x+y))\frac{\theta(x)^{n-1}}{(n-1)!}M(x,dy) \right) \delta_0(dh). \tag{62}$$

This means that the TSS model $Z$ with initial state $x$ can be constructed as follows: let $(U(k), k = 0,1,2,\dots)$ be a Markov chain in $\mathcal{X}$ with initial state $x$ and with transition kernel $\kappa(x,dh)$, and let $(P(t), t \geq 0)$ be an independent standard Poisson process. Then, the process $(Z_t, t \geq 0)$ defined by

$$Z_t := U \circ P\left(\int_0^t \beta(Z_s)ds\right)$$

is a Markov process with infinitesimal generator (62) [11, Chap. 6]. Let $(T_n)_{n\geq 1}$ denote the sequence of jump times of the Poisson process $P$ and define $(S_n)_{n\geq 1}$ by $T_n = \int_0^{S_n} \beta(Z_s)ds$ or $S_n = \infty$ if $\int_0^\infty \beta(Z_s)ds < T_n$. Observe that any jump of the process $Z$ occurs at some time $S_n$, but that all $S_n$ may not be effective jump times for $Z$, because of the Dirac mass at 0 appearing in (62). As will appear below, the sequence $(S_n)$ can be interpreted as the sequence of mutation times in the limit process. Whether an effective jump occurs at time $S_n$ or not then corresponds to the fixation or extinction of the mutant.

Let $\mathbf{P}_x$ denote the law of $\zeta_t$ conditional on $\mathrm{Supp}(\zeta_0) = Z_0 = x$. Fix $t > 0$, $m \in \mathbb{N}^*$, $x \in \mathcal{X}$ and a measurable subset $\Gamma$ of $\mathcal{X}$. Under $\mathbf{P}_x$, $S_1$ and $Z_{S_1}$ are independent, $S_1$ is an exponential random variable with parameter $\beta(x)$, and $Z_{S_1}$ has law $\kappa(x,\cdot)$. Therefore, for any $n \geq 1$, applying the strong Markov property to the process $Z$ at time $S_1$ in the second line,

$$\mathbf{P}_x(S_n \leq t < S_{n+1},\ \exists z \in \Gamma : \zeta_t = m\delta_z) = \mathbf{E}_x\left(\mathbf{1}_{\{Z_t\in\Gamma\}}\mathbf{1}_{\{S_n\leq t<S_{n+1}\}}\frac{e^{-\theta(Z_t)}}{1-e^{-\theta(Z_t)}}\frac{\theta(Z_t)^m}{m!}\right) \\ = \int_0^t \beta(x)e^{-\beta(x)s}\int_{\mathbb{R}^k} \mathbf{P}_{x+h}(S_{n-1} \leq t-s < S_n,\ \exists z \in \Gamma : \zeta_{t-s} = m\delta_z)\kappa(x,dh)ds. \tag{63}$$

Moreover,

$$\mathbf{P}_x(0 \leq t < S_1,\ \exists z \in \Gamma : \zeta_t = m\delta_z) = \mathbf{1}_{\{x\in\Gamma\}}e^{-\beta(x)t}\frac{e^{-\theta(x)}}{1-e^{-\theta(x)}}\frac{\theta(x)^m}{m!}. \tag{64}$$

These two relations characterize the one-dimensional laws of the process $\zeta$. The idea of our proof is to show that the same relations hold when we replace $S_n$ by $\tau_n$ and the support of $\zeta_t$ by the support of $\nu^\gamma_{t/\gamma}$ (when it is a singleton) in the limit $\gamma \to 0$.

More precisely, let us define for any $\nu_0 \in \mathcal{M}$ and $n \in \mathbb{N}$

$$p_n^\gamma(t,\Gamma,m,\nu_0) := \mathbb{P}^\gamma_{\nu_0}\left(\rho_n \leq \frac{t}{\gamma} < \tau_{n+1},\ \exists z \in \Gamma : \nu_{t/\gamma} = m\delta_z\right)$$

We will prove the following lemma after the end of this proof.

**Lemma 6.4** *For any $x \in \mathcal{X}$, $m, k \geq 1$, $n \geq 0$, $t > 0$ and any measurable subset $\Gamma$ of $\mathcal{X}$, $p_n(t,\Gamma,m,x) := \lim_{\gamma\to 0} p_n^\gamma(t,\Gamma,m,k\delta_x)$ exists, is independent of $k$ and satisfies*

$$p_0(t,\Gamma,m,x) = \mathbf{1}_{\{x\in\Gamma\}} e^{-\beta(x)t} \frac{e^{-\theta(x)}}{1-e^{-\theta(x)}} \frac{\theta(x)^m}{m!} \tag{65}$$

$$\textit{and} \quad \forall n \geq 1, \quad p_n(t,\Gamma,m,x) = \int_0^t \beta(x) e^{-\beta(x)s} \int_{\mathbb{R}^k} p_{n-1}(t-s,\Gamma,m,x+h)\kappa(x,dh)ds. \tag{66}$$

Comparing (63) and (64) with (65) and (66), this lemma implies that $p_n(t,\Gamma,m,x) = \mathbf{P}_x(S_n \leq t < S_{n+1},\ \exists z \in \Gamma : \zeta_t = m\delta_z)$.

Recall that $\nu_0^\gamma = N_0^\gamma \delta_x$ with $\sup_{\gamma\in(0,1)} \mathbb{E}((N_0^\gamma)^p) < \infty$ for some $p > 1$. By Proposition 5.1, $\sup_{\gamma\in(0,1)} \sup_{t\geq 0} \mathbb{E}(\langle \nu_t^\gamma, \mathbf{1}\rangle^p) < +\infty$ and

$$\begin{aligned}\Big|\mathbb{P}^\gamma_{\nu_0^\gamma}(\exists z \in \Gamma : \nu^\gamma_{t/\gamma} = m\delta_z) - \mathbf{P}_x(\exists z \in \Gamma : \zeta_t = m\delta_z)\Big| \\ \leq 2\mathbb{P}(N_0^\gamma > M) + \sum_{k=1}^M \left|\sum_{n=0}^\infty (p_n^\gamma(t,\Gamma,m,k\delta_x) - p_n(t,\Gamma,m,x))\right| \mathbb{P}(N_0^\gamma = k).\end{aligned}$$

Because of Lemma 6.1 (55), the quantity inside the absolute value in the r.h.s. of this equation converges to 0 when $\gamma \to 0$. Thus,

$$\lim_{\gamma\to 0} \mathbb{P}^\gamma_{\nu_0^\gamma}(\exists z \in \Gamma : \nu^\gamma_{t/\gamma} = m\delta_z) = \mathbf{P}_x(\exists z \in \Gamma : \zeta_t = m\delta_z). \tag{67}$$

Summing this relation over $m \in \mathbb{N}^*$ and taking $\Gamma = \mathcal{X}$ implies that

$$\lim_{\gamma\to 0} \mathbb{P}^\gamma_{\nu_0^\gamma}(\mathrm{Supp}(\nu^\gamma_{t/\gamma}) \text{ is a singleton}) = 1. \tag{68}$$

Now, consider a bounded measurable $f : \mathcal{M} \to \mathbb{R}$ such that $f(\nu) = 0$ if $\langle \nu, \mathbf{1}\rangle \neq m \in \mathbb{N}^*$ and define the function $\hat{f} : \mathcal{X} \to \mathbb{R}$ by $\hat{f}(x) = f(m\delta_x)$. Then, it follows from (67) and (68) that

$$\begin{aligned}\lim_{\gamma\to 0} \mathbb{E}^\gamma_{\nu_0^\gamma}\left(f(\nu^\gamma_{t/\gamma})\right) &= \lim_{\gamma\to 0} \mathbb{E}^\gamma_{\nu_0^\gamma}\left(\hat{f}(\mathrm{Supp}(\nu^\gamma_{t/\gamma})); \mathrm{Supp}(\nu^\gamma_{t/\gamma}) \text{ is a singleton and } \langle \nu^\gamma_{t/\gamma}, \mathbf{1}\rangle = m\right) \\ &= \mathbf{E}_x(\hat{f}(Z_t); N_t = m) = \mathbf{E}_x(f(\zeta_t)) \end{aligned} \tag{69}$$

This equality generalizes to any bounded measurable $f : \mathcal{M} \to \mathbb{R}$ using once again that $\sup_{\gamma\in(0,1)} \mathbb{E}(\langle \nu_t^\gamma, \mathbf{1}\rangle^p) < +\infty$. This completes the proof of Theorem 5.2 for one-dimensional distributions.

The extension to finite dimensional marginals can be proved exactly in the same fashion.
□

**Proof of Lemma 6.4.** We will prove this lemma by induction over $n \geq 0$.

Fix $x \in \mathcal{X}$, $m, k \geq 1$ and $t > 0$. First, we have already proved in (57) that

$$\begin{aligned} p_0^\gamma(t, \Gamma, m, k\delta_x) &= \mathbf{1}_{\{x \in \Gamma\}} \mathbb{P}^\gamma_{k\delta_x}\left(\langle \nu^\gamma_{t/\gamma}, \mathbf{1}\rangle = m,\ \gamma\tau_1 > t\right) \\ &= \mathbf{1}_{\{x \in \Gamma\}} \mathbb{E}\left[\mathbf{1}_{\{X^\gamma_{t/\gamma} = m\}} \exp\left(-\gamma\mu(x)b(x)\int_0^{t/\gamma} X^\gamma_{s-}\, ds\right)\right] \end{aligned}$$

where the process $X^\gamma$ is a scalar logistic branching process with dynamical parameters $((1-\gamma\mu(x))b(x), c(x,x), 0)$ and with $X^\gamma_0 = n$. Using the method that led us to (60), we get

$$\lim_{\gamma \to 0} p_0^\gamma(t, \Gamma, m, k\delta_x) = \mathbf{1}_{\{x \in \Gamma\}} \mathbb{P}(\xi_{\theta(x)} = m) \exp(-\mu(x)b(x)\mathbb{E}(\xi_{\theta(x)})t),$$

which entails (65).

Then, fix $n \geq 1$. Applying the strong Markov property to the process $\nu^\gamma$ at time $\tau_1$, and using the fact that the mutant trait at this time is $x + U$ where $U$ has law $M(x, dh)$ and is independent of $\nu^\gamma_{\tau_1-}$, we get

$$p_n^\gamma(t, \Gamma, m, k\delta_x) = \int_{\mathbb{R}^k} \mathbb{E}^\gamma_{k\delta_x}\left[\mathbf{1}_{\{\gamma\tau_1 \leq t\}} p^\gamma_{n-1}(t - \gamma\tau_1, \Gamma, m, \langle \nu^\gamma_{\tau_1-}, \mathbf{1}\rangle\delta_x + \delta_{x+h})\right] M(x, dh). \tag{70}$$

Now, we want to apply the strong Markov property to $\nu^\gamma$ at time $\rho_0$ to compute the quantity $p^\gamma_{n-1}(s, \Gamma, m, l\delta_x + \delta_y)$ appearing inside the expectation in the last formula. For $K > 0$, distinguishing between the cases where $\rho_0 > \tau_1$, $\langle \nu_{\rho_0}, \mathbf{1}\rangle > K$, $V_0 = x$ and $V_0 = y$ yields

$$\begin{aligned} p^\gamma_{n-1}(s, \Gamma, m, l\delta_x + \delta_y) = \mathbb{E}^\gamma_{l\delta_x + \delta_y}\Big[&\mathbf{1}_{\{\rho_0 < \tau_1,\ \langle \nu^\gamma_{\rho_0}, \mathbf{1}\rangle \leq K,\ V_0 = x\}} p^\gamma_{n-1}(s - \gamma\rho_0, \Gamma, m, \langle \nu^\gamma_{\rho_0}, \mathbf{1}\rangle\delta_x) \\ &+ \mathbf{1}_{\{\rho_0 < \tau_1,\ \langle \nu^\gamma_{\rho_0}, \mathbf{1}\rangle \leq K,\ V_0 = y\}} p^\gamma_{n-1}(s - \gamma\rho_0, \Gamma, m, \langle \nu^\gamma_{\rho_0}, \mathbf{1}\rangle\delta_y)\Big] \\ &+ \mathbb{P}^\gamma_{l\delta_x + \delta_y}(\{\rho_0 \geq \tau_1\} \cap E) + \mathbb{P}^\gamma_{l\delta_x + \delta_y}(\{\rho_0 < \tau_1\} \cap \{\langle \nu^\gamma_{\rho_0}, \mathbf{1}\rangle > K\} \cap E) \end{aligned}$$

where

$$E = \{\rho_{n-1} \leq s/\gamma < \tau_n,\ \exists z \in \Gamma : \nu_{s/\gamma} = m\delta_z\}.$$

The third term of the r.h.s. converges to 0 when $\gamma \to 0$ because of Lemma 6.3(a) and the last term converges to 0 when $K \to +\infty$ uniformly for $\gamma \in (0, 1)$ because of Lemma 6.3(b).

Now, assume that $p^\gamma_{n-1}(t, \Gamma, m, k\delta_x)$ converges to $p_{n-1}(t, \Gamma, m, x)$ as in the statement of Lemma 6.4. As a consequence of Lemma 6.1 (54), for any $t > 0$, the function $s \mapsto p^\gamma_{n-1}(s, \Gamma, m, k\delta_x)$ is uniformly continuous on $[0, t]$. Combining this observation with Lemma 6.3(c),

$$\lim_{\gamma \to 0} p^\gamma_{n-1}(s, \Gamma, m, l\delta_x + \delta_y) = u_{l,1}(x, y)p_{n-1}(s, \Gamma, m, y) + (1 - u_{l,1}(x, y))p_{n-1}(s, \Gamma, m, x). \tag{71}$$

This uniform continuity argument also applies to $s \mapsto p^\gamma_{n-1}(s, \Gamma, m, l\delta_x + \delta_y)$, so that the convergence in (71) is uniform in $s \in [0, t]$ and $l \in \{1, \ldots, L\}$, for fixed $L \geq 1$. Therefore, we can combine Lemma 6.2(a) and (b) as above to get

$$\begin{aligned} &\lim_{\gamma \to 0} \mathbb{E}^\gamma_{k\delta_x}\left[\mathbf{1}_{\{\gamma\tau_1 \leq t\}} p^\gamma_{n-1}(t - \gamma\tau_1, \Gamma, m, \langle \nu^\gamma_{\tau_1-}, \mathbf{1}\rangle\delta_x + \delta_y)\right] \\ &= \int_0^t ds\beta(x)e^{-\beta(x)s} \sum_{l=1}^{\infty} e^{-\theta(x)} \frac{\theta(x)^{l-1}}{(l-1)!} \left[u_{l,1}(x, y)p_{n-1}(t - s, \Gamma, m, y) + (1 - u_{l,1}(x, y))p_{n-1}(t - s, \Gamma, m, x)\right]. \end{aligned}$$

Finally, using Lebesgue's theorem, this limit applies inside the integral in (70), which gives exactly (66) and ends the proof of Lemma 6.4. □

**Proof of Theorem 5.3.** Since the limiting law of the process $(S^\gamma_t; t \geq 0)$ is characterized by its finite dimensional distributions, obtained in Theorem 5.2, we only have to show the tightness of their laws. Fix $T > 0$. By Ascoli's theorem for cÃ dlÃ g functions (see e.g. [2]), we have to show that, for any $\varepsilon, \eta > 0$, there is $\delta > 0$ such that

$$\limsup_{\gamma \to 0} \mathbb{P}(\omega'(S^\gamma, \delta) > \eta) < \varepsilon \tag{72}$$

where the modulus of continuity $\omega'$ is defined as

$$\omega'(f, \delta) = \inf\{\max_{0 \leq i \leq r-1} \omega(f, [t_i, t_{i+1}))\}$$

where the infimum is taken over all the finite partitions $0 = t_0 < t_1 < \ldots < t_r = T$ of $[0, T]$ such that $t_{i+1} - t_i > \delta$ for any $0 \leq i \leq r - 1$, and where $\omega$ is the usual modulus of continuity

$$\omega(f, I) = \sup_{s,t \in I} \|f(t) - f(s)\|.$$

Now, for any $n \geq 0$,

$$\mathbb{P}^\gamma(\rho_{n+1} - \rho_n < \delta) \leq \mathbb{P}^\gamma(\rho_n - \tau_n > \delta) + \mathbb{P}^\gamma(\tau_{n+1} - \tau_n < 2\delta).$$

When $\gamma \to 0$, the first term goes to 0 by Lemma 6.3(a) and Lemma 6.2(b), and the lim sup of the second term is smaller than $1 - \exp(-\underline{\beta}\delta)$ by Lemma 6.2(a) and (b), where $\underline{\beta} := \inf_{x \in \mathcal{X}} \beta(x) > 0$. Therefore, by Lemma 6.1(b), for any $\varepsilon > 0$, there exists $\delta > 0$ such that

$$\limsup \mathbb{P}^\gamma(\exists n \geq 0 : \rho_{n+1} - \rho_n < \delta \text{ and } \rho_{n+1} \leq T) \leq \varepsilon,$$

which implies (72). □

## 6.2 Proofs of Theorem 5.5 and Proposition 5.8

**Proof of Theorem 5.5.** We will use the classical method of tightness and martingale problem formulation to prove this theorem (e.g. [18]). We divide the proof in three steps.

*Step 1. Uniqueness of the limit process.* Strong existence and uniqueness for the SDE (47) follow standardly from the Lipschitz-continuity of its coefficients.

*Step 2. Tightness of the family of laws of $Z^\epsilon$.* For any $\epsilon > 0$, let $N_\epsilon(dh, du, dt)$ be a Poisson point process on $\mathbb{R}^k \times [0, 1] \times \mathbb{R}_+$ with intensity measure $q_\epsilon(dh, du, dt) = \bar{M}(h)dh\bar{\beta}\bar{\chi}du dt/\epsilon^2$, where $\bar{\beta}$ and $\bar{\chi}$ are constants bounding the functions $b$ and $\chi$ from above, respectively, and $\bar{M}$ has been defined as the integrable function bounding the density $m(x, \cdot)$ of $M(x, \cdot)$ for any $x \in \mathbb{R}^k$. Then it is straightforward to check that $A_\epsilon$ is the infinitesimal generator of the Markov process $Z^\epsilon$

$$Z^\epsilon_t = Z^\epsilon_0 + \epsilon \int_0^t \int_0^1 \int_{\mathbb{R}^k} h \mathbf{1}_{\left\{u \leq \frac{\beta(Z^\epsilon_{s-})}{\bar{\beta}} \frac{\chi(Z^\epsilon_{s-}, Z^\epsilon_{s-} + \epsilon h)}{\bar{\chi}} \frac{m(Z^\epsilon_{s-}, h)}{\bar{M}(h)}\right\}} N_\epsilon(dh, du, ds)$$

Since $\beta$ and $\chi$ are bounded, a process generated by $A_\epsilon$ is unique in law (e.g. [11]), and this construction characterizes the law of the process $Z^\epsilon$ appearing in the statement of Theorem 5.5. Let us denote this law by $\mathbf{P}_\epsilon$.

Observe that, if we denote by $\tilde{N}_\epsilon$ the compensated Poisson measure $N_\epsilon - q_\epsilon$, $Z_t^\epsilon$ can be decomposed as $Z_0^\epsilon + \tilde{Z}_t^\epsilon + \hat{Z}_t^\epsilon$, where

$$\tilde{Z}_t^\epsilon = \epsilon \int_0^t \int_0^1 \int_{\mathbb{R}^k} h \mathbf{1}_{\left\{u \leq \frac{\beta(Z_{s-}^\epsilon)}{\bar{\beta}} \frac{\chi(Z_{s-}^\epsilon, Z_{s-}^\epsilon + \epsilon h)}{\bar{\chi}} \frac{m(Z_{s-}^\epsilon, h)}{\bar{M}(h)}\right\}} \tilde{N}_\epsilon(dh, du, ds)$$

and

$$\begin{aligned}\hat{Z}_t^\epsilon &= \frac{1}{\epsilon} \int_0^t \int_{\mathbb{R}^k} h \beta(Z_{s-}^\epsilon) \chi(Z_{s-}^\epsilon, Z_{s-}^\epsilon + \epsilon h) M(Z_{s-}^\epsilon, dh) ds \\ &= \frac{1}{\epsilon} \int_0^t \int_{\mathbb{R}^k} h \beta(Z_{s-}^\epsilon) [\chi(Z_{s-}^\epsilon, Z_{s-}^\epsilon + \epsilon h) - \chi(Z_{s-}^\epsilon, Z_{s-}^\epsilon)] M(Z_{s-}^\epsilon, dh) ds\end{aligned}$$

where the last equality follows from the fact that the mutation step law $M(x, \cdot)$ has 0 expectation.

We will use Aldous' criterion [1] to prove the tightness of the family of probability measure $(\mathbf{P}_\epsilon)_{\epsilon > 0}$ on $\mathbb{D}(\mathbb{R}_+, \mathbb{R}^k)$. Fix $\delta, \epsilon > 0$ and let $\tau$ and $\tau'$ be two stopping times such that $\tau < \tau' < \tau + \delta$. Since $|\chi(x, x + \epsilon h) - \chi(x, x)| \leq \epsilon K \|h\|$ for some constant $K$, $\|\hat{Z}_{\tau'}^\epsilon - \hat{Z}_\tau^\epsilon\| \leq \delta \bar{\beta} K M_2$, where $M_2 = \sup_x \int \|h\|^2 M(x, dh)$, which is finite by assumption. By standard results on stochastic integrals with respect to Poisson point measures,

$$\begin{aligned}\mathbf{E}_\epsilon(\|\tilde{Z}_{\tau'}^\epsilon - \tilde{Z}_\tau^\epsilon\|^2) &= \mathbf{E}_\epsilon \left( \int_\tau^{\tau'} \int_0^1 \int_{\mathbb{R}^k} \epsilon^2 \|h\|^2 \mathbf{1}_{\left\{u \leq \frac{\beta(Z_{s-}^\epsilon)}{\bar{\beta}} \frac{\chi(Z_{s-}^\epsilon, Z_{s-}^\epsilon + \epsilon h)}{\bar{\chi}} \frac{m(Z_{s-}^\epsilon, h)}{\bar{M}(h)}\right\}} q_\epsilon(dh, du, ds) \right) \\ &\leq \delta \bar{\beta} \bar{\chi} M_2.\end{aligned}$$

Therefore, for any $\eta > 0$,

$$\mathbf{P}_\epsilon(\|Z_{\tau'}^\epsilon - Z_\tau^\epsilon\| \geq \eta) \leq \mathbf{P}_\epsilon(\|\hat{Z}_{\tau'}^\epsilon - \hat{Z}_\tau^\epsilon\| \geq \frac{\eta}{2}) + \mathbf{P}_\epsilon(\|\tilde{Z}_{\tau'}^\epsilon - \tilde{Z}_\tau^\epsilon\| \geq \frac{\eta}{2}) \leq \mathbf{1}_{\{2\delta \bar{\beta} K M_2 \geq \eta\}} + \frac{4 \delta \bar{\beta} \bar{\chi} M_2}{\eta^2}$$

which converges to 0 when $\delta \to 0$. This gives the first part of Aldous' criterion. For the second part, we have to prove the tightness of $(Z_t^\epsilon)_{\epsilon > 0}$ for any $t \geq 0$. Similar computations as above prove easily that $(\tilde{Z}_t^\epsilon)_{\epsilon > 0}$ and $(\hat{Z}_t^\epsilon)_{\epsilon > 0}$ are tight, and the tightness of $(Z_0^\epsilon)_{\epsilon > 0}$ follows from the assumption that it is bounded in $L^1$.

*Step 3. Martingale problem.* Let $\mathbf{P}_0$ be an accumulation point of $(\mathbf{P}_\epsilon)_{\epsilon > 0}$ when $\epsilon \to 0$ on $\mathbb{D}(\mathbb{R}_+, \mathbb{R}^k)$, endowed with the canonical filtration $\mathcal{F}_t$. Since the martingale problem for (47) is well-posed, it suffices to show that, for any $\varphi \in \mathcal{C}^2(\mathcal{X})$, under $\mathbf{P}_0$, the process

$$M_t^\varphi(w) = \varphi(w_t) - \varphi(w_0) - \int_0^t A_0 \varphi(w_s) ds.$$

on $\mathbb{D}(\mathbb{R}_+, \mathcal{X})$ is a local $\mathcal{F}_t$-martingale. We already know that under $\mathbf{P}_\epsilon$,

$$M_t^{\epsilon, \varphi}(w) = \varphi(w_t) - \varphi(w_0) - \int_0^t A_\epsilon \varphi(w_s) ds$$

is a local $\mathcal{F}_t$-martingale. Since $\beta$ and $\chi$ are bounded, this is a square-integrable martingale as soon as $\varphi \in \mathcal{C}_b^3$.

Fix $\varphi \in \mathcal{C}_b^3$, $s > 0$ and $t > s$, and consider $p$ real numbers $0 \leq t_1 < \ldots < t_p \leq s$ for some $p \geq 1$, and a continuous bounded function $q : (\mathbb{R}^k)^p \to \mathbb{R}$. We can write

$$\begin{aligned}
&\left|\mathbf{E}_0\left\{q(w_{t_1},\ldots,w_{t_p})\left[\varphi(w_t)-\varphi(w_s)-\int_s^t A_0\varphi(w_u)du\right]\right\}\right| \\
&\qquad\leq \left|\mathbf{E}_\epsilon\left\{q(w_{t_1},\ldots,w_{t_p})\left[\varphi(w_t)-\varphi(w_s)-\int_s^t A_\epsilon\varphi(w_u)du\right]\right\}\right| \\
&\qquad+\left|\mathbf{E}_\epsilon\left\{q(w_{t_1},\ldots,w_{t_p})\int_s^t (A_\epsilon\varphi(w_u)-A_0\varphi(w_u))du\right\}\right| \\
&\qquad+\left|\mathbf{E}_0\left\{q(w_{t_1},\ldots,w_{t_p})\left[\varphi(w_t)-\varphi(w_s)-\int_s^t A_0\varphi(w_u)du\right]\right\}\right. \\
&\qquad\quad\left.-\mathbf{E}_\epsilon\left\{q(w_{t_1},\ldots,w_{t_p})\left[\varphi(w_t)-\varphi(w_s)-\int_s^t A_0\varphi(w_u)du\right]\right\}\right|.
\end{aligned}$$

The first term of the r.h.s. is 0 since $M^{\epsilon,\varphi}$ is a $\mathbf{P}_\epsilon$-martingale. Because of the uniform convergence of generators (46), the second term converges to 0 when $\epsilon \to 0$. The third term also goes to 0 when $\epsilon \to 0$ since $\mathbf{P}_\epsilon$ converges to $\mathbf{P}_0$. Finally, since the l.h.s. does not depend on $\epsilon$, it is 0.

A classical use of the monotone class theorem allows to extend this equality to all $\mathcal{F}_s$-measurable bounded functions $q$, so $M^\varphi$ is a $\mathbf{P}_0$-martingale. This result can easily be extended to any $\mathcal{C}^2$ function $\varphi$ by a standard truncation technique, which completes the proof of Theorem 5.5. □

**Proof of Proposition 5.8.** It follows from Proposition 4.5 that

$$a_\lambda(x) = \sum_{n=1}^{\infty} \frac{n\theta(x)^{n-1}}{2c(x,x)(n+2)(n+1)(n-1)!}$$

and from Proposition 4.6 that

$$a_\alpha(x) = \sum_{n=1}^{\infty} \frac{\theta(x)^{n-1}}{b(x)(n+1)(n-1)!} + \sum_{n=1}^{\infty} \frac{n\theta(x)^{n-1}}{2c(x,x)(n+2)(n+1)(n-1)!} - \sum_{n=1}^{\infty} \frac{q_{n+1}^{(2)}(x)\theta(x)^{n-1}}{\theta(x)\kappa(x)(n+1)(n-1)!}.$$

Elementary calculations then give (48) and (49).

For the $\delta$-invasibility, using Proposition 4.8 and switching the two sums, we get

$$a_\delta(x) = \sum_{k\geq 1} \phi_k(x) \sum_{n\geq 1} \frac{n^2\theta(x)^{n-1}}{(n+k+1)(n+1)!}.$$

The following observation

$$\sum_{n\geq 1} \frac{n^2 u^{n-1}}{(n+1)!} = \frac{d}{du}\left(u\frac{d}{du}\sum_{n\geq 1}\frac{u^n}{(n+1)!}\right) = \frac{d}{du}\left(u\frac{d}{du}\left(\frac{e^u-1}{u}\right)\right) = \frac{e^u(u^2-u+1)-1}{u^2}$$

yields (50). (51) follows from switching the sum and the integral in (50) which is standard since $\phi_n(x) = O(1/n)$.

Equation (52) can be checked using the fact that $e^u(u^k - ku^{k-1} + k(k-1)u^{k-2} + \ldots + (-1)^k k!)$ is a primitive of $e^u u^k$.

Finally, (53) can be deduced from the facts that $\phi_1(x) = 1/2c(x,x) = \theta(x)/2b(x)$ and

$$\forall n \geq 2, \quad (n+2)\phi_{n+1}(x) + (\theta(x) - n - 1)\phi_n(x) - \theta(x)\frac{n-2}{n-1}\phi_{n-1}(x) = 0.$$

Multiplying these equations by $(n-1)x^{n+1}$ and summing over $n \geq 2$ yields

$$\begin{aligned} 0 &= \sum_{n\geq 3}(n+1)(n-2)\phi_n(x)u^n + \theta(x)\sum_{n\geq 2}(n-1)\phi_n(x)u^{n+1} \\ &\qquad - \sum_{n\geq 2}(n+1)(n-1)\phi_n(x)u^{n+1} - \theta(x)\sum_{n\geq 2}(n-1)\phi_n(x)u^{n+2} \\ &= \frac{d}{du}\left(u^4\frac{d}{du}\left(\frac{\pi_x(u) - \phi_1(x) - \phi_2(x)u}{u}\right)\right) + \theta(x)u^3\pi'_x(u) - u\frac{d}{du}\left(u^3\pi'_x(u)\right) - \theta(x)u^4\pi'_x(u) \end{aligned}$$

which finally gives (53). □

## References

[1] Aldous, D. 1978. Stopping times and tightness. *Ann. Probab.* **6** 335–340.

[2] Billingsley, P. 1968. *Convergence of Probability Measures.* John Wiley & Sons.

[3] Champagnat, N. 2006. A microscopic interpretation for adaptive dynamics trait substitution sequence models. *Stoch. Proc. Appl.* (in press)

[4] Champagnat, N., Ferrière, R., Ben Arous, G. 2001. The canonical equation of adaptive dynamics: A mathematical view. *Selection* **2** 71-81.

[5] Champagnat, N., Ferrière, R., Méléard S. 2006. Unifying evolutionary dynamics: from individual stochastic processes to macroscopic models via timescale separation. *Theor. Popul. Biol.* (in press)

[6] Crow, J. F., Kimura, M. 1970. *An Introduction to Population Genetics Theory.* Harper & Row, New York.

[7] Dieckmann, U., Doebeli, M. 1999. On the origin of species by sympatric speciation. *Nature* **400** 354-357.

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

[9] Durrett, R., Schweinsberg, J. 2004. Approximating selective sweeps. *Theor. Pop. Biol.* **66** 129-138.

[10] Etheridge, A., Pfaffelhuber, P. Wakolbinger, A. 2006. An approximate sampling formula under genetic hitchhiking. *Ann. Appl. Probab.* (in press)

[11] Ethier, S. N., Kurtz, T. G. 1986. *Markov Processes, Characterization And Convergence.* John Wiley & Sons, New York.

[12] Ewens, W.J. 2005. *Mathematical Population Genetics.* 2nd edition, Springer-Verlag, Berlin.

[13] Feller, W. 1951. Diffusion processes in genetics. *Proc. Second Berkeley Symp. Math. Statist. Prob.* 227-246.

[14] Freidlin, M.I. 1968. On the factorization of non-negative definite matrices. *Theory Probab. Appl.* **13** 354-356.

[15] Gillespie, J.H. 1998. *Population Genetics: A Concise Guide.* Johns Hopkins U. Press, Baltimore, MD.

[16] Hansen, T.F. 1997. Stabilizing selection and the comparative analysis of adaptation. *Evolution* **51** 1341Â–1351.

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

[18] Jacod, J., Shiryaev, A. N. 1987. *Limit Theorems for Stochastic Processes.* Springer-Verlag, Berlin.

[19] Kimura, M. 1964. Diffusion models in population genetics. *J. Appl. Prob.* **1** 177-232.

[20] Kimura, M. 1983. *The neutral theory of molecular evolution.* Cambridge U. Press, Cambridge.

[21] Kimura, M. 1994. *Population Genetics, Molecular Evolution, and the Neutral Theory: Selected Papers.* U. of Chicago Press, Chicago.

[22] Kisdi, E., Gyllenberg, M. 2005. Adaptive dynamics and the paradigm of diversity. *J. Evol. Biol.* **18** 1170Â–1173.

[23] Lambert, A. 2005. The branching process with logistic growth. *Ann. Appl. Prob.* **15** 1506-1535.

[24] Lambert, A. 2006. Probability of fixation under weak selection: a branching process unifying approach. *Theor. Popul. Biol.* (in press)

[25] Lande, R. 1976. Natural selection and random genetic drift in phenotypic evolution. *Evolution* **30** 314Â–334.

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

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

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

[29] Norris, J.R. 1997. *Markov Chains*. Cambridge U. Press, Cambridge.

[30] Spencer, H.G., Feldman, M.W. 2005. Adaptive dynamics, game theory and evolutionary population genetics. *J. Evol. Biol.* **18** 1191Â–1193.