Evolution of discrete populations
and the canonical diffusion of adaptive dynamics

By Nicolas Champagnat[*] and Amaury Lambert[†]

Abstract

The biological theory of adaptive dynamics proposes a description of the long-term evolution of a structured asexual population. It is based on the assumptions of large population, rare mutations and small mutation steps, that lead to a deterministic ODE describing the evolution of the dominant type, called the 'canonical equation of adaptive dynamics'. Here, in order to include the effect of stochasticity (genetic drift), we consider self-regulated randomly fluctuating populations subject to mutation, so that the number of coexisting types may fluctuate. We apply a limit of rare mutations to these populations, 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. Then we apply a limit of small mutation steps (weak selection) to this jump process, that leads to a diffusion process that we call the 'canonical diffusion of adaptive dynamics', in which genetic drift is combined with directional selection driven by the gradient of the fixation probability, also interpreted as an invasion fitness. Finally, we study in detail the particular case of multitype logistic branching populations and seek explicit formulae for the invasion fitness of a mutant deviating slightly from the resident type. In particular, second-order terms of the fixation probability are products of functions of the initial mutant frequency, times functions of the initial total population size, called the invasibility coefficients of the resident by increased fertility, defence, aggressiveness, isolation, or survival.

Running head. Adaptive dynamics for finite populations.[1]

1 Introduction

Consider a multitype population where each individual of type x gives birth independently at rate $b(x)$ to an individual of type x (clonal reproduction) and dies either naturally at rate $d(x)$ or by *competition*. Deaths by competition are distributed as follows. Each individual of type y points to each

individual of type x an independent exponential clock of parameter $c(x, y)$. Death by competition of an individual of type x occurs as soon as a clock pointing to her rings. The vector of subpopulation sizes is distributed as the so-called *multitype logistic branching process* [25].

In this model, the population size cannot go to infinity and eternal coexistence of two or more types is impossible, so the number of types decreases until it reaches 1 at some random time T called the *fixation time*. For a two-type population starting with n individuals of type x and m individuals of type y, we denote by $u_{n,m}(x, y)$ the probability that y fixes, i.e. that y is the surviving type at time T. We refer to *neutrality* as the case when $x = y$. After time T the population is said to be *monomorphic*. If the natural death rate d of the surviving type is not zero, the population eventually becomes extinct, otherwise the size process is positive-recurrent.

The previous observations also apply to the population dynamics considered in this paper, that generalize the logistic branching process to a multitype setting with mutation and more general interactions, and that we call *GL-populations*. Types live in a subset $\mathcal{X}$ of $\mathbb{R}^k$ and, in a pure x-type population of size n, any individual gives birth to a new individual at rate $b(x, n)$ and dies at rate $d(x, n)$. We assume that $d(\cdot, 1) \equiv 0$ (i.e. extinction is impossible) and that, uniformly in x, $b(x, \cdot)$ is bounded, and $d(x, \cdot)$ bounded from below by some positive power of the total population size. Then, when there is no mutation, the size of a pure x-type population is positive-recurrent, as for the logistic branching process when natural death rates vanish, and converges in distribution to some random variable $\xi(x)$. However, in general, reproduction is not clonal. Specifically, each time an individual of type x gives birth, the type of the newborn individual is x with probability $1 - \mu(x)$ and is $x + h$ with probability $\mu(x)M(x, dh)$, where $\mu(x)$ is the *mutation probability* and $M(x, \cdot)$ is a probability measure on $\mathcal{X} - x$ called the *mutation kernel* or *mutation step law*. In this setting, the loss of diversity is counteracted by the occurrence of mutations, and the number of types can also increase.

The evolution of structured populations (with or without the presence of mutations) has long been studied, and in numerous different models. The renowned field that takes into account the complexity of the *genetic structure* — x is a *genotype* — is called *population genetics* [6, 23, 17, 13]. The emerging field where the emphasis is put on the structure of *ecological interactions* — x is a *phenotypic trait*, such as body size — is called *adaptive dynamics* [19, 28, 30]. The link between both is still unclear, but see [24, 33].

Let T_n denote the n-th time when the population becomes monomorphic, and V_n the then surviving type. The (possibly finite) sequence (V_n) is of obvious interest to evolutionary biologists and is called the *trait substitution sequence*. It can be defined in much more general contexts, as soon as eternal coexistence of two or more types is not permitted by the model. It was in adaptive dynamics that this sequence was invented and studied [31, 3], under two additional assumptions. First, the biologically motivated *assumption of rare mutations* guarantees that, in the timescale of mutations (speeding up time), the widths of time intervals during which the population is polymorphic vanish, so that there is one and only one type surviving at any time t. To prevent the population from rapidly becoming extinct in the new timescale, one also has to rescale population sizes, thereby making the *assumption of large populations*.

Subsequently, the trait substitution sequence (TSS) is a Markov jump process on $\mathcal{X}$ whose semi-

group is shown [3] to depend especially on the *invasion fitnesses* (as defined in [30]) $f(x, y)$, $x, y \in \mathcal{X}$, where $f(x, y)$ is the *expected growth rate* of a single individual of type y — the mutant — entering a monomorphic population of type x 'at equilibrium' — the residents. The evolution of the population can be described by this fitness function because of the large population assumption, which implies that deleterious mutants — those with negative invasion fitness — can never invade. Thus, evolution proceeds by *successive invasions of advantageous mutant types* replacing the resident one (selective sweeps [29, 10, 11]), which can be summarized by the jump process of fixed types (the TSS), also called phenotypic traits, or simply *traits*, in the adaptive dynamics setting (e.g. size, age at maturity or rate of food intake).

The TSS has been a powerful tool for understanding various evolutionary phenomena, such as evolutionary branching (evolution from a monomorphic population to a stably polymorphic one [31] that may lead to speciation [7]) and is the basis for other biological models, such as the 'canonical equation of adaptive dynamics' [8, 4]. This last phrase refers to the following ODE, describing the evolution of a one-dimensional trait x, obtained from the TSS in the limit of *small mutations*

$$\frac{dx}{dt} = \frac{1}{2}\sigma(x)^2 \mu(x)\bar{n}(x)\frac{\partial}{\partial y}f(x, x),\tag{1}$$

where $\sigma(x)^2$ stands for the variance of the mutation step law, $\bar{n}(x)$ for the equilibrium size of a pure x-type population, and $f(x, y)$ for the aforementioned invasion fitness. Note how only advantageous types get fixed (the trait follows the fitness gradient).

However, it is well-known that slightly deleterious types can be fixed in finite populations. This phenomenon is known under the name of *genetic drift* (more generally, this term refers to allelic fluctuations which have stochastic causes). Depending on the strength of genetic drift, selection is said to be *strong* (genetic drift has negligible effects) or *weak*. Adaptive dynamics models usually assume infinite populations subject to deterministic dynamics, where only strong selection can be observed. As a consequence, the first goal of this paper is to consider *population finiteness* and *weak selection* [22], and continue using the bottom-up approach of adaptive dynamics; that is, model (macroscopic) evolution from (microscopic) populations [5]. In particular, we allow the population sizes to fluctuate randomly through time, in contrast to classical works in both population genetics (where the population size traditionally remains constant, see [26]) and adaptive dynamics (where population sizes are infinite).

Thus, we consider GL-populations in which the mutation probabilities μ are replaced with $\gamma\mu$, and time t with t/γ (mutation timescale). We prove the existence as γ vanishes (rare mutations) of the limiting TSS and characterize its law (Theorem 3.2). Recall that $u_{n,m}(x, y)$ denotes the fixation probability of type y in a two-type population starting with n individuals of type x and m individuals of type y.

Theorem 1.1 *The process* $(S_t^\gamma; t \geq 0)$ *defined as*

$$S_t^\gamma = \sum_{n \geq 0} V_n \mathbf{1}_{\{T_n \leq t/\gamma < T_{n+1}\}}$$

3

converges in law for the Skorohod topology on $\mathbb{D}(\mathbb{R}_+, \mathcal{X})$ as $\gamma \to 0$ to a pure-jump Markov process $(S_t; t \geq 0)$, whose jumping rates $q(x, dh)$ from x to $x + h$ are given by

$$q(x, dh) = \beta(x)\chi(x, x + h)M(x, dh),$$

where

$$\beta(x) = \mu(x)\mathbb{E}(\xi(x)b(x, \xi(x))),$$

and

$$\chi(x, y) = \sum_{n \geq 1} \frac{nb(x, n)\mathbb{P}(\xi(x) = n)}{\mathbb{E}(\xi(x)b(x, \xi(x)))}u_{n,1}(x, y). \tag{2}$$

Observe that $\beta(x)$ is the *mean mutant production rate* of a stationary x-type population and that $\chi(x, y)$ is the *probability of fixation* of a single (mutant) y-type mutant entering a pure (resident) x-type population with $b(x, \cdot)$-*size-biased stationary size*. In particular, $\chi(x, y)$ *is the random analogue of the usual invasion fitness*. More precisely, with the notations of (1), $\beta(x)$ is analogous to $\mu(x)b(x)\bar{n}_x$ and $\chi(y, x)$ is the exact counterpart of the expression $[f(y, x)]_+/b(y)$ obtained in [3] for the probability of fixation of a y-type mutant in an equilibrium x-type population.

The second goal of this paper is to take the limit of small mutations on the TSS defined in the last theorem. Namely, assume for simplicity that $\mathcal{X} = \mathbb{R}^k$, and denote by Z^ϵ the process S modified by replacing $M(x, \cdot)$ with its image by the contraction ϵ Id and time t with t/ϵ^2. Under some technical assumptions which ensure that $\chi(x, \cdot)$ is differentiable, and assuming that mutation steps have zero expectation ($\int hM(x, dh) = 0$), we get the following theorem (Theorem 4.2), where $\sigma(x)$ is the symmetric square root matrix of the covariance matrix of $M(x, \cdot)$ and ∇_2 denotes the gradient taken w.r.t. the second variable.

Theorem 1.2 *The process Z^ϵ converges in law for the Skorohod topology on $\mathbb{D}(\mathbb{R}_+, \mathbb{R}^k)$ to the diffusion process $(Z_t; t \geq 0)$ unique solution to the stochastic differential equation*

$$dZ_t = \beta(Z_t)\sigma^2(Z_t) \cdot \nabla_2\chi(Z_t, Z_t)dt + \sqrt{\beta(Z_t)\chi(Z_t, Z_t)}\sigma(Z_t) \cdot dB_t \tag{3}$$

where B is a standard k-dimensional Brownian motion.

In words, we obtain a diffusion model for the evolution of a trait grounded on microscopic realistic population dynamics. The diffusion term embodies *genetic drift*. It is proportional to the mean mutant production rate $\beta(x)$, to the neutral fixation probability $\chi(x, x)$ and to the covariance matrix of $M(x, \cdot)$. The deterministic term embodies *directional selection*, and is the exact counterpart of the ODE (1).

This equation, that we christen 'canonical *diffusion* of adaptive dynamics' involves the stochastic invasion fitness χ and its gradient with respect to the second variable at neutrality. As seen in the definition (2) of χ, the fitness gradient only depends on the behaviour of the fixation probabilities near neutrality for two interacting types (see also [26]), and on the stationary distribution $\xi(x)$. The explicit computation of these quantities is possible in the multitype logistic branching case, which we study in detail.

First consider a pure x-type logistic branching population with dynamical parameters $(b(x), c(x, x), d(x))$. When $d = 0$, the population size is positive-recurrent and the r.v. ξ distributed according to the stationary probability is a Poisson variable of parameter $\theta := b/c$ conditioned on being nonzero. This yields

$$\mathbb{E}(\xi(x)) = \frac{\theta(x)}{1 - e^{-\theta(x)}}$$

and

$$\chi(x, x) = \frac{e^{-\theta(x)} - 1 + \theta(x)}{\theta(x)^2} .$$

Second, we characterize the two-type logistic branching process by a birth vector B, a competition matrix C, and a death vector D

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

where $\mathbf{1}$ is a matrix full of ones with dimensions *ad hoc* and $\lambda, \delta, \alpha, \varepsilon, \sigma$ are the selection coefficients of the mutant respectively associated to *fertility, defence, aggressiveness, isolation* and *survival*. In the following theorem (Theorem 5.1), 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 the *invasibility coefficient* of the resident population.

Theorem 1.3 *As a function of the multidimensional selection coefficient $s = (\lambda, \delta, \alpha, \varepsilon, \sigma)'$, the probability u is differentiable, and in a neighbourhood of $s = 0$ (selective neutrality),*

$$u = p + v'.s + o(s),$$

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

$$v^\iota_{n,m} = p(1-p) g^\iota_{n+m} \qquad \iota \neq \varepsilon,$$

$$v^\varepsilon_{n,m} = p(1-p)(1-2p) g^\varepsilon_{n+m},$$

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

As a result, we get

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

where a_λ, a_δ and a_α are called *adaptive slopes* in terms of resp. fertility, defence and aggressiveness. Explicit formulae are provided for the adaptive slopes as well, that can be plugged into (3). These results allow to get deeper insight into the canonical diffusion of adaptive dynamics when the microscopic interactions are those of multitype logistic branching processes.

The paper is organized as follows. First, our stochastic individual-based model is described in the next section (Section 2). Next, we state precisely the convergence results to the TSS in finite populations (Section 3), and to the canonical diffusion of adaptive dynamics (Section 4). In Section 5, we stick to the logistic branching case and obtain explicit formulae for the derivative of the fixation probability, the invasibility coefficients and the adaptive slopes. Finally (Section 6), we give the detailed proofs of the convergence results of Sections 3 and 4.

2 Model

2.1 Preliminaries

Recall from the Introduction that a monotype (binary) logistic branching process with dynamical parameters (b, c, d) is a Markov chain in continuous time $(X_t; t \geq 0)$ with nonnegative integer values and transition rates

$$
q_{ij} = \begin{cases}
bi & \text{if} & j = i + 1 \\
ci(i - 1) + di & \text{if} & j = i - 1 \\
-i(b + c(i - 1) + d) & \text{if} & j = i \\
0 & \text{otherwise.}
\end{cases}
$$

We know from [25] that ∞ is an entrance boundary for X and that $\mathbb{E}_\infty(\tau) < \infty$, for τ the hitting time of, say, 1. If $d \neq 0$, then the process goes extinct a.s., and if $d = 0$, it is positive-recurrent and converges in distribution to a r.v. ξ, where ξ is a Poisson variable of parameter $\theta := b/c$ conditioned on being nonzero

$$
\mathbb{P}(\xi = i) = \frac{e^{-\theta}}{1 - e^{-\theta}} \frac{\theta^i}{i!} \qquad i \geq 1. \tag{4}
$$

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

For fixed $\alpha > 0$, one can generalize the interaction in the previous model to obtain the so-called α-logistic branching process with dynamical parameters (b, c, d) by modifying the transition rates as

$$
q_{ij}^\alpha = \begin{cases}
bi & \text{if} & j = i + 1 \\
ci(i - 1)^\alpha + di & \text{if} & j = i - 1 \\
-i(b + c(i - 1)^\alpha + d) & \text{if} & j = i \\
0 & \text{otherwise.}
\end{cases}
$$

As before, it is easy to check that, when $d \neq 0$, the process goes extinct a.s., and when $d = 0$, it is positive-recurrent and converges in distribution to a r.v. $\xi^{(\alpha)}$ which law can be explicitly computed as

$$
\mathbb{P}(\xi^{(\alpha)} = i) = C \frac{\theta^i}{i((i - 1)!)^\alpha} \qquad i \geq 1, \tag{5}
$$

where $\theta = b/c$ and C scales the sum of these terms to 1.

2.2 GL-populations

In this subsection, we define the general populations considered in this paper, that we call GL-populations, and give their basic properties. These populations are structured (multitype) populations with mutation. Their dynamics are those of eternal birth-and-death processes with birth rates at most linear and death rates of order at least $1 + \alpha$ in the total population size, hence their name (generalized logistic).

2.2.1 Definition

At any time t, the population is composed of a finite number $N(t)$ of individuals characterized by their phenotypic traits, or simply *traits*, $x_1(t), \ldots, x_{N(t)}(t)$ 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 σ-field induced by the Borel σ-field on $\mathcal{X} \subset \mathbb{R}^k$ as follows: let φ denote the application mapping any element $\sum_{i=1}^k \delta_{x_i}$ of $\mathcal{M}$ to the k-tuple $(x_{\pi(1)}, \ldots, x_{\pi(k)})$ where the permutation π of $\{1, \ldots, k\}$ is chosen such that this vector is ranked in, say, the lexicographical order. Then, this function φ is a bijection from $\mathcal{M}$ to the set of lexicographically ordered vectors of $\cup_{i=0}^k \mathcal{X}^k$. The Lebesgue σ-field on this set therefore provides a σ-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 $\Gamma \subset \mathcal{X}$.

Let us consider a general structured birth-and-death process with mutation whose present state is given by the point measure ν.

- $b(x, \nu)$ is the rate of birth from an individual of type x in a population in state ν

- $d(x, \nu)$ is the rate of death of an individual of type x in a population in state ν

- $\gamma\mu(x)$ 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. In Section 3, we will be interested in the limit of rare mutations ($\gamma \to 0$).

- $M(x, dh)$ is the law of the trait difference $h = y - x$ 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 $\mathcal{X} - x := \{y - x : y \in \mathcal{X}\} \subset \mathbb{R}^k$. We assume that $M(x, dh)$ has a density on $\mathbb{R}^k$ which is uniformly bounded in $x \in \mathcal{X}$ by some integrable function $\bar{M}(h)$.

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

$$
\begin{aligned}
L_\gamma \varphi(\nu) = {} & \int_{\mathcal{X}} [\varphi(\nu + \delta_x) - \varphi(\nu)] \, (1 - \gamma\mu(x)) \, b(x, \nu)\nu(dx) \\
& + \int_{\mathcal{X}} \int_{\mathbb{R}^k} [\varphi(\nu + \delta_{x+h}) - \varphi(\nu)]\gamma\mu(x)b(x,\nu)M(x, dh)\nu(dx) \\
& + \int_{\mathcal{X}} [\varphi(\nu - \delta_x) - \varphi(\nu)]d(x, \nu)\nu(dx). \quad\quad (6)
\end{aligned}
$$

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 ν_t on the parameter γ with the notation ν_t^γ.

Definition 2.1 *We say that the structured birth-and-death process with mutation defined previously is a GL-population, if*

- *there is $\bar{b}$ such that for any ν and x, $0 < b(x, \nu) \leq \bar{b}$*

- *there are $\underline{c}$ and $\alpha > 0$ such that for any ν and x, $\underline{c} \, (\langle \nu, \mathbf{1}\rangle - 1)^\alpha \leq d(x, \nu)$*

- *if $\nu = \delta_x$, then $d(x, \nu) = 0$.*

The classical structured logistic branching process [9, 15, 25, 5] is a *GL*-population with

$$
b(x, \nu) = b(x)
$$

and

$$
d(x, \nu) = \int_{\mathcal{X}} c(x, y)(\nu - \delta_x)(dy).
$$

Then, the assumption above translates as $b(\cdot) \leq \bar{b}$, $\alpha = 1$ and $c(\cdot, \cdot) \geq \underline{c}$.

2.2.2 Basic properties

First, the total population size $N_t = \langle \nu_t, \mathbf{1}\rangle$ of a *GL*-population is dominated by a scalar α-logistic branching process with parameters $(\bar{b}, \underline{c}, 0)$, so $(\nu_t; t \geq 0)$ has infinite lifetime.

Second, when there is only one individual in the population ($\nu_t = \delta_x$ for some $x \in \mathcal{X}$), the death rate equals 0, so that extinction is impossible.

Third, for a *GL*-population with two types x and y and **no** mutation ($\mu \equiv 0$), $\nu_t = X_t\delta_x + Y_t\delta_y$, where $(X_t, Y_t : t \geq 0)$ is a bivariate Markov chain. We refer to (selective) *neutrality* as the case when $x = y$. Because of the previous domination, for the Markov chain (X, Y), the union of the axes

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

is accessible and absorbing, and its complementary set is transient. So $\mathbb{P}(T < \infty) = 1$, where $T := T_{\Omega_1} \wedge T_{\Omega_2}$, and for any subset Γ of $\mathbb{N}^2$, T_Γ denotes the first hitting time of Γ by (X, Y). Also

notice 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 y) the event $\{T_{\Omega_2} < T_{\Omega_1}\}$. The probability of fixation will be denoted by $u_{n,m}(x, y)$

$$u_{n,m}(x, y) := \mathbb{P}(T_{\Omega_2} < T_{\Omega_1} \mid X_0 = n, Y_0 = m).$$

More generally, we have the following result.

Proposition 2.2 *Consider a GL-population $(\nu_t; t \geq 0)$ with **no** mutation. For any initial condition, the fixation time T*

$$T := \inf\{t \geq 0 : |Supp(\nu)| = 1\}$$

is finite a.s. and from time T, the population remains monomorphic with type, say, x. Then conditional on T and x, the post-T size process $(N(t); t \geq T)$ is positive-recurrent and converges in distribution to a random integer $\xi(x)$ such that $\sup_x \mathbb{E}(\xi(x)^n) < \infty$ for any n.

Proof. This result follows from the domination of the total population size by a α-logistic branching process $(Z_t; t \geq 0)$ with dynamical parameters $(\bar{b}, \underline{c}, 0)$, and from the fact that such a process is positive recurrent with stationary distribution given by (5). Then, the total population size hits 1 in a.s. finite time (greater than or equal to T), and the population size returns then to 1 in a time bounded by the time of return to 1 of Z, which has finite expectation as a consequence of its positive-recurrence. So the post-T size process N is positive-recurrent as well and its stationary distribution is dominated by the law (5), which has finite moments. $\square$

More generally, the domination of the population size $\langle \nu_t^\gamma, \mathbf{1} \rangle$ by a monotype α-logistic branching process with dynamical parameters $(\bar{b}, \underline{c}, 0)$ for any $\gamma \in (0, 1)$, allows us to prove the following long time bound for the moments of ν^γ.

Proposition 2.3 *Fix $p \geq 1$ and pick a positive C. There is a constant C' such that, for any $\gamma \in (0, 1)$,*

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

Proof. With the notation of the previous proof, 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

$$\frac{d}{dt}\mathbb{E}(Z_t^p) = \sum_{k \geq 1} k^p \frac{dp_t^k}{dt} = \sum_{k \geq 1} k^p [\bar{b}(k-1)p_t^{k-1} + \underline{c}(k+1)k^\alpha p_t^{k+1} - k(\bar{b} + \underline{c}(k-1)^\alpha)p_t^k]$$

$$= \sum_{k \geq 1} \left[\bar{b}\left(\left(1 + \frac{1}{k}\right)^p - 1\right) + \underline{c}(k-1)^\alpha \left(\left(1 - \frac{1}{k}\right)^p - 1\right) \right] k^{p+1} p_t^k.$$

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

$$\bar{b}((1 + 1/k)^p - 1) + \underline{c}(k-1)^\alpha((1 - 1/k)^p - 1) \leq -\bar{b}[3 - 2(1 - 1/k)^p - (1 + 1/k)^p]$$

9

which is equivalent to $-\bar{b}p/k$. Then, enlarging k_0 if necessary, we obtain

$$\frac{d}{dt}\mathbb{E}(Z_t^p) \leq \sum_{k=1}^{k_0} \bar{b}(2^p - 1)k_0^p - \sum_{k \geq k_0+1} \frac{\bar{b}p}{2}k^p p_t^k$$

$$\leq K - \frac{\bar{b}p}{2}\mathbb{E}(Z_t^p),$$

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

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

which gives the required uniform bound. $\qquad\square$

3 The trait substitution sequence in finite populations

In this section, we consider the GL-population ν^γ of Section 2. Our goal is to apply to this process a limit of rare mutations ($\gamma \to 0$) while keeping the population size finite, in order to describe the evolutionary process on the mutation timescale t/γ as a 'trait substitution sequence' (TSS, [31, 3]) where evolution proceeds by successive fixations of mutant types.

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 r.v. Y if and only if $\mathbb{E}(f(X_n)) \to \mathbb{E}(f(Y))$ as $n \to \infty$ for any bounded measurable real function f. Let us also make a slight abuse of notation by writing $b(x, n)$ instead of $b(x, n\delta_x)$ for the birth rate in a monomorphic population (as in the Introduction).

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

Theorem 3.1 *For any $0 < t_1 < \ldots < t_n$, the n-tuple $(\nu_{t_1/\gamma}^\gamma, \ldots, \nu_{t_n/\gamma}^\gamma)$ converges strongly in law to $(\zeta_{t_1}, \ldots, \zeta_{t_n})$ where $\zeta_{t_i} = N_{t_i}\delta_{S_{t_i}}$ and*

(1) *$(S_t; t \geq 0)$ is a Markov jump process on $\mathcal{X}$ with initial value $S_0 = x$ and whose jumping rates $q(x, dh)$ from x to $x + h$ are given by*

$$q(x, dh) = \beta(x)\chi(x, x + h)M(x, dh), \tag{7}$$

where

$$\beta(x) = \mu(x)\mathbb{E}(\xi(x)b(x, \xi(x))), \tag{8}$$

and

$$\chi(x, y) = \sum_{n \geq 1} \frac{nb(x, n)\mathbb{P}(\xi(x) = n)}{\mathbb{E}(\xi(x)b(x, \xi(x)))}u_{n,1}(x, y). \tag{9}$$

(2) *Conditional on $(S_{t_1}, \ldots, S_{t_n}) = (x_1, \ldots, x_n)$, the N_{t_i} are independent and respectively distributed as $\xi(x_i)$.*

10

Observe that $\beta(x)$ in (8) can be seen as the *mean mutant production rate* of a stationary x-type population, and that $\chi(x,y)$ is the *probability of fixation* of a single y-type mutant entering a pure x-type population with $b(x,\cdot)$-*size-biased stationary size*. In particular, $\chi(x,y)$ is the random analogue of the traditional invasion fitness, in the sense proposed by Metz et al. [30].

This result shows that, in the limit of rare mutations, on the mutation timescale, the population is always monomorphic and that the dominant trait of the population evolves as a jump process over the trait space, where a jump corresponds to the appearance **and** fixation of a mutant type.

Let us denote by τ_n the n-th mutation time ($\tau_0 = 0$), by ρ_n the first time after time τ_n when the population becomes monomorphic, and by V_n the single trait value surviving at time ρ_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 3.1, namely the convergence of the support of the measure $\nu^\gamma_{\cdot/\gamma}$ to the process S.

Theorem 3.2 *The process $(S_t^\gamma; t \geq 0)$ defined as*

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

converges in law for the Skorohod topology on $\mathbb{D}(\mathbb{R}_+, \mathcal{X})$ as $\gamma \to 0$ to the process $(S_t; t \geq 0)$ with initial state $S_0 = x$ characterized by (7).

Observe that such a convergence for the measure $\nu^\gamma_{\cdot/\gamma}$ cannot hold because the population size N_t in Theorem 3.1 is not a càdlàg process.

The proofs of the two preceding theorems are put to Section 6.2.

4   The canonical diffusion of adaptive dynamics

4.1   Notation and assumptions

For any integer r, we denote by $\mathcal{C}_b^r$ the set of r times differentiable functions with (image space *ad hoc* and) all derivatives bounded.

For a two-type GL-population, we will make a slight abuse of notation by writing $b(x,y,n(x),n(y))$ instead of $b(x, n(x)\delta_x + n(y)\delta_y)$ and the analogue notation for the death rate.

If $\|\cdot\|$ is the L^1 norm, for any $i = (n,m)$ and j such that $\|j - i\| = 1$, then let π_{ij} stand for the transition probability of the embedded Markov chain associated to the two-type GL-population without mutation. For example if $j - i = (1,0)$,

$$\pi_{ij}(x,y) = \frac{nb(x,y,n,m)}{nb(x,y,n,m) + mb(y,x,m,n) + nd(x,y,n,m) + md(y,x,m,n)} \ . \tag{10}$$

From now on, we make the following additional assumptions.

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

- for all n, m the functions $b(\cdot, \cdot, n, m)$ and $d(\cdot, \cdot, n, m)$ are in $\mathcal{C}_b^2$

- there are constants C_1, C_2 such that for any $x, y \in \mathcal{X}$, for any $i = (n, m)$ and j such that $\|j - i\| = 1$,
$$\|\nabla_2 \pi_{ij}(x, y)\| \leq C_1, \qquad \|H_2 \pi_{ij}(x, y)\| \leq C_2, \tag{11}$$
where ∇_2 is the gradient and H_2 the Hessian matrix taken w.r.t. the second variable

- the mutation kernels $M(x, \cdot)$ satisfy

  - for any $x \in \mathbb{R}^k$, $M(x, \cdot)$ has 0 expectation, i.e. $\int_{\mathbb{R}^k} h M(x, dh) = 0$
  - 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 $\int_{\mathbb{R}^k} (s'h)^2 M(x, dh) \geq C\|s\|^2$ 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.

4.2 Differentiability of the probability of fixation

In the following theorem, we state the existence of the partial derivatives of $u_{n,m}(x, y)$, and show that these derivatives are always sublinear in the initial condition. We also give a uniform bound for the second-order derivatives of the fixation probability.

Theorem 4.1 (a) *The fixation probability $y \mapsto u_{n,m}(x, y)$ is differentiable and its derivative $v_{n,m}(x, y)$ satisfies*
$$\sup_{n,m,x,y} \frac{\|v_{n,m}(x, y)\|}{n + m} < +\infty. \tag{12}$$

(b) *In addition, $y \mapsto u_{n,m}(x, y)$ is in $\mathcal{C}^2$, and its second-order derivatives are bounded by $c(n + m)^2$, where c does not depend on x and y.*

This theorem will be proved in Subsection 6.3.

It is easy to see from the last theorem and from the fact that $\xi(x)$ is stochastically dominated by a random variable with law (5) (with $b = \bar{b}$ and $c = \underline{c}$) that the function χ defined in (9) is in $\mathcal{C}_b^2$.

4.3 The canonical diffusion of adaptive dynamics

Here, we want to apply a limit of small mutation steps (weak selection) to the TSS S defined in the previous section, in order to obtain the equivalent of the canonical equation of adaptive dynamics, but in finite populations [8, 4].

The limit of small jumps is obtained by introducing a parameter $\epsilon > 0$ and replacing the mutation kernels $M(x, \cdot)$ with their image by the application $h \mapsto \epsilon h$. Of course, this scaling of jumps' sizes

has to be combined with a scaling of time in order to observe a non-trivial limit. This leads to the following generator for the rescaled TSS $(Z_t^\epsilon; t \geq 0)$ (as will appear futher below, the factor $1/\epsilon^2$ is the right timescaling)

$$A_\epsilon \varphi(x) = \frac{1}{\epsilon^2} \int_{\mathbb{R}^k} (\varphi(x + \epsilon h) - \varphi(x)) \beta(x) \chi(x, x + \epsilon h) M(x, dh). \quad (13)$$

Fix a function φ in $\mathcal{C}_b^3$. For any $x, h \in \mathbb{R}^k$ and $\epsilon > 0$, there exists $0 \leq \epsilon_1 \leq \epsilon$ such that

$$\varphi(x + \epsilon h) - \varphi(x) = \epsilon h' \nabla \varphi(x) + \frac{\epsilon^2}{2} h' H \varphi(x + \epsilon_1 h) h$$

where $H\varphi(y)$ denotes the Hessian matrix of φ at y, and there exists $0 \leq \epsilon_2 \leq \epsilon$ such that

$$\chi(x, x + \epsilon h) = \chi(x, x) + \epsilon h' \nabla_2 \chi(x, x + \epsilon_2 h),$$

where $\nabla_2 \chi$ is the gradient of χ w.r.t. the second variable. Therefore, using the fact that $H\varphi$ and $\nabla_2 \chi$ are bounded Lipschitz functions, it takes only elementary computations to prove that

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

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 $\varphi \in \mathcal{C}_b^3$, that $A_\epsilon \varphi$ converges uniformly to the function $A_0 \varphi$ defined as

$$A_0 \varphi(x) = \int_{\mathbb{R}^k} (h' \nabla \varphi(x)) \beta(x) (h' \nabla_2 \chi(x, x)) M(x, dh) + \frac{1}{2} \int_{\mathbb{R}^k} (h' H \varphi(x) h) \beta(x) \chi(x, x) M(x, dh). \quad (14)$$

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.

Theorem 4.2 *If the family $(Z_0^\epsilon)_{\epsilon > 0}$ has bounded first-order moments and converges in law as $\epsilon \to 0$ to a random variable Z_0, then the process Z^ϵ generated by (13) with initial state Z_0^ϵ 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*

$$dZ_t = \beta(Z_t) \sigma^2(Z_t) \cdot \nabla_2 \chi(Z_t, Z_t) dt + \sqrt{\beta(Z_t) \chi(Z_t, Z_t)} \sigma(Z_t) \cdot dB_t \quad (15)$$

where B is a standard k-dimensional Brownian motion.

Theorem 4.2 will be proved in Subsection 6.4.

Note that, by an elementary martingale (or exchangeability) argument, the neutral fixation probability $u_{n,m}(x, x)$ equals $m/(n + m)$, so that

$$\chi(x, x) = \sum_{n \geq 1} \frac{n b(x, n) \mathbb{P}(\xi(x) = n)}{\mathbb{E}(\xi(x) b(x, \xi(x)))} \frac{1}{n + 1} = \frac{1}{\mathbb{E}(\xi(x) b(x, \xi(x)))} \mathbb{E}\left(\frac{\xi(x) b(x, \xi(x))}{1 + \xi(x)}\right).$$

13

Remark 4.3 *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^ϵ to be well-defined, one needs to assume that scaling the mutation law $M(x, dh)$ cannot drive Z^ϵ out of $\mathcal{X}$. This is true for example, when $\epsilon \leq 1$, 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 (see [16]).*

Remark 4.4 *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_ϵ does not vanish, so that the correct timescaling is $1/\epsilon$ (instead of $1/\epsilon^2$), and the TSS process Z^ϵ can be shown to converge to the solution of the deterministic ODE*

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

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 4.2 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 as tools for evolutionary biology, but mainly to describe the fluctuations of allelic frequencies (see, among many others, [14, 13, 21]). In sparser works [27, 18], a diffusion can describe the evolution of the dominant or mean value of a quantitative trait in a population. Our process provides such a model, but it is grounded on a microscopic precise modelling of the population dynamics, in a realistic way. In particular, the population size is not fixed and may fluctuate randomly through time but remains finite because of the density-dependence.

The diffusion part in (15) gives the strength of genetic drift: its square is proportional to the mutation rate $\beta(x)$, the neutral fixation probability $\chi(x, x)$, and the covariance matrix of the mutation step law $M(x, dh)$. As for the deterministic drift part, observe the similarity with the standard canonical equation of adaptive dynamics (1). This term 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 χ, which appears in the deterministic part of (15) in the shape of its gradient with respect to the second variable, in a similar way as in (1). 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.

5 The logistic branching case

In this section, we restrict our attention to the structured logistic branching process described in the Introduction and in Section 2. Since we seek explicit expressions for the drift and diffusion parts of the canonical diffusion, we are only concerned with two-type populations with no mutation.

5.1 Preliminaries

Since mutations are absent, we will usually call the first type (1) resident (or wild type) and the second type (2) mutant. Then 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.

Formally, the two-type logistic branching process is a bivariate integer-valued continuous-time Markov process $(X_t, Y_t; t \geq 0)$ with rate matrix $Q = (q_{ij}; i \in \mathbb{N}^2, j \in \mathbb{N}^2)$, where

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

and where the total jumping rate r_{nm} is the sum of the four jumping rates from (n, m).

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

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

To comply with the GL-population framework, we will always consider that $c_{11} c_{12} c_{21} c_{22} \neq 0$. Assuming that the presence of a mutant form does not modify the dynamical characteristics (b, c, d) of the resident, we may focus on deviations from the neutral case so as to express the dynamical parameters 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* λ, δ, α, ε, σ, 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** (λ, as the usual letter standing for growth rate in discrete-time deterministic models) : positive λ means increased mutant birth rate

2. **defence** capacity (δ, as in defence) : positive δ means reduced competition sensitivity of mutant individuals w.r.t. the total population size

3. **aggressiveness** (α, as in aggressive, or attack) : positive α means raised competition pressure exerted from any mutant individual onto the rest of the population

4. **isolation** (ε, as in exclusion) : positive ε means lighter cross-competition between different morphs, that would lead, if harsher, to the exclusion of the less abundant one

5. **survival** (σ, as in survival) : positive σ means reduced mutant death rate.

Under neutrality, an elementary martingale argument shows that the fixation probability equals the initial mutant frequency $p := m/(m+n)$, that is,

$$u = p.$$

The goal of the following theorem is to unveil the dependence of u upon λ, δ, α, ε, σ, when they slightly deviate from 0. It is proved in Subsections 5.3 and 5.4.

Theorem 5.1 *As a function of the multidimensional selection coefficient $s = (\lambda, \delta, \alpha, \varepsilon, \sigma)'$, the probability u is differentiable, and in a neighbourhood of $s = 0$ (selective neutrality),*

$$u = p + v'.s + o(s), \tag{16}$$

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

$$v^\iota_{n,m} \;\;=\;\; p\,(1-p)\,g^\iota_{n+m} \qquad\qquad \iota \neq \varepsilon,$$

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

where 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 the invasibility coefficients.

The invasibility coefficients of a pure resident population are interesting to study, as they provide insight as to how the fixation probability deviates from p as the selection coefficients of the mutant deviate from 0. Their name is due to the fact that they only depend on the resident's characteristics and are multipliers of the mutant's selection coefficients [26]. In the next subsection, we apply this result to the canonical diffusion of adaptive dynamics. The remainder of the section is then devoted to the proof of the theorem and the study of the invasibility coefficients.

5.2   The canonical diffusion for logistic branching populations

Consider a logistic branching population satisfying the conditions given in Section 2.2.1. Recall that $b(x)$ is the birth rate of x-type individuals, and $c(x,y)$ is the competition rate felt by x-type individuals from y-type individuals. Assume as in Section 4.1 that $b(\cdot)$ and $c(\cdot,\cdot)$ are in $\mathcal{C}^2_b$. Since $c(\cdot,\cdot) \geq \underline{c}$, it is then elementary to check that (11) holds. By (4), the invasion fitness is given by

$$\chi(x,y) = \sum_{n \geq 1} \frac{n\mathbb{P}(\xi(x) = n)}{\mathbb{E}(\xi(x))} u_{n,1}(x,y) = \sum_{n \geq 1} e^{-\theta(x)} \frac{\theta(x)^{n-1}}{(n-1)!} u_{n,1}(x,y),$$

where $\theta(x) = b(x)/c(x,x)$.

Observe 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 ($i = 1, 2$).

Recall Theorem 4.2 and the canonical diffusion of adaptive dynamics

$$dZ_t = \beta(Z_t)\sigma^2(Z_t) \cdot \nabla_2\chi(Z_t, Z_t)dt + \sqrt{\beta(Z_t)\chi(Z_t, Z_t)}\sigma(Z_t) \cdot dB_t \tag{17}$$

where B is a standard k-dimensional Brownian motion. From Theorem 5.1 and elementary computations, we get

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

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

$$a_\iota(x) = \sum_{n=1}^{\infty} v_{n,1}^\iota(x)\frac{\theta(x)^{n-1}}{(n-1)!} = \sum_{n=1}^{\infty} \frac{n\theta(x)^{n-1}}{(n+1)^2(n-1)!}g_{n+1}^\iota(x), \tag{18}$$

and g^λ, g^δ, g^α are the invasibility coefficients in terms of resp. fertility, defence, aggressiveness. Since the coefficients a_ι appear as factors of the gradients of the microscopic parameters b and c in the deterministic part of the canonical diffusion (17), we call them the *adaptive slopes*. Explicit formulae for invasibility coefficients and adaptive slopes are given in Subsection 5.4.

Remark 5.2 *Observe that ε-invasibilities do not appear in this computation, because of the symmetry between resident and mutant types in the competition kernel. One could include ε-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 could 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 to whether x is the resident or not.

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 step law $M(x, \cdot)$. Then, thanks to forthcoming Proposition 5.14 about adaptive slopes (18), 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.

5.3 Fixation probability

Here, we go back to the two-type population with no mutation and characterize the fixation probability thanks to a discrete harmonic equation (corresponding to Kolmogorov forward equations). Then we prove Theorem 5.1.

Proposition 5.3 *The fixation probability $u_{n,m}$ is the unique bounded solution to*

$$\begin{cases} (\Delta u)_{n,m} = 0 & \text{for } (n,m) \notin \Omega_1 \cup \Omega_2 \\ u_{n,m} = 0 & \text{for } (n,m) \in \Omega_1 \\ u_{n,m} = 1 & \text{for } (n,m) \in \Omega_2, \end{cases} \tag{19}$$

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

$$(\Delta w)_{n,m} = r_{n,m} w_{n,m} - b_1 n w_{n+1,m} - b_2 m w_{n,m+1}$$
$$- n\big(c_{11}(n-1) + c_{12}m + d_1\big)w_{n-1,m} - m\big(c_{21}n + c_{22}(m-1) + d_2\big)w_{n,m-1}. \tag{20}$$

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 (19) follows from the Markov property at the first jump time of (X, Y), and the uniqueness relies on Lemma 5.4 below.

Lemma 5.4 *Consider a subset Γ 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 Γ, the equation*

$$\begin{cases} (\Delta h)_{n,m} = 0 & \text{for } (n,m) \notin \Gamma \\ h(n,m) = f(n,m) & \text{for } (n,m) \in \Gamma \end{cases} \tag{21}$$

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

Proof. It suffices to prove that (21) 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 2.3, $\sup_{t \geq 0} \mathbb{E}_{n,m}(X_t + Y_t)) < +\infty$, $(h(X_t, Y_t))_{t \geq 0}$ is actually a uniformly integrable martingale. Applying the stopping theorem at time T_Γ, 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. □

Proof of Theorem 5.1. As seen in Proposition 5.3, the Kolmogorov forward equations translate into a discrete harmonic equation satisfied by u with boundary condition 1 on Ω_1, and 0 on Ω_2, written as $(\Delta u)_{n,m} = 0$, where Δ is defined in (20). Combining (16) and (20), and identifying

second-order terms, we get

$$(\Delta_0 v^\iota)_{n,m} = \begin{cases} \dfrac{nm}{(n+m)(n+m+1)} & \text{if} \quad \iota = \lambda \\[3mm] \dfrac{nm}{n+m} & \text{if} \quad \iota = \delta \\[3mm] \dfrac{nm}{(n+m)(n+m-1)} & \text{if} \quad \iota = \alpha, \sigma \\[3mm] \dfrac{nm\,(n-m)}{(n+m)(n+m-1)} & \text{if} \quad \iota = \varepsilon, \end{cases} \tag{22}$$

where Δ_0 corresponds to the neutral case of Δ: for any doubly indexed w,

$$(\Delta_0 w)_{n,m} = (n+m)\Big[b + c(n+m-1) + d\Big]w_{n,m} - bn\, w_{n+1,m} - bm\, w_{n,m+1}$$
$$- n\Big[c(n+m-1) + d\Big]w_{n-1,m} - m\Big[c(n+m-1) + d\Big]w_{n,m-1} \qquad n, m \geq 0. \tag{23}$$

We know from Theorem 4.1 that the vector $v = (v^\lambda, v^\delta, v^\alpha, v^\varepsilon, v^\sigma)'$ is sublinear in (n, m), that is, $(\|v_{n,m}\|/(n+m))_{n,m}$ is bounded. *Since the r.h.s. in (22) are all sublinear*, Lemma 5.4 ensures that *v is the unique sublinear vector in (n, m) solving (22).*

Thanks to this uniqueness result, it is sufficient to show that there are solutions of (22) of the following form

$$v^\iota_{n,m} = \begin{cases} \dfrac{nm}{n+m}\, u^\iota_{n+m} & \text{if} \quad \iota = \lambda, \delta, \alpha, \sigma \\[3mm] \dfrac{nm\,(n-m)}{n+m}\, u^\iota_{n+m} & \text{if} \quad \iota = \varepsilon, \end{cases} \tag{24}$$

where for $\iota \neq \varepsilon$, u^ι is a *bounded* real sequence indexed by $\mathbb{N} - \{0, 1\}$ (u^ι_1 has no effect on the values of $v^\iota_{1,0}$ and $v^\iota_{0,1}$), and u^ε is a real sequence indexed by $\mathbb{N} - \{0, 1, 2\}$ (u^ε_1 and u^ε_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

$$g^\iota_n = \begin{cases} n\, u^\iota_n & \text{if} \quad \iota = \lambda, \delta, \alpha, \sigma, \quad n \geq 2 \\[3mm] n^2\, u^\iota_n & \text{if} \quad \iota = \varepsilon, \qquad\qquad n \geq 3. \end{cases} \tag{25}$$

In this setting, (22) holds iff

$$\begin{cases} (Lu^\lambda)_n = \dfrac{1}{n(n+1)} \quad \text{and} \quad (Lu^\delta)_n = \dfrac{1}{n} \quad n \geq 2, \\[3mm] (Lu^\alpha)_n = (Lu^\sigma)_n = \dfrac{1}{n(n-1)} \qquad\qquad n \geq 2, \\[3mm] (L'u^\varepsilon)_n = \dfrac{1}{n(n-1)} \qquad\qquad\qquad n \geq 3, \end{cases} \tag{26}$$

19

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

$$
\begin{cases}
(Lw)_n &= -b\dfrac{n+2}{n+1}w_{n+1} + \Big[b + c(n-1) + d\Big]w_n - (n-2)\Big(c + \dfrac{d}{n-1}\Big)w_{n-1} \quad n \geq 2 \\[4mm]
(L'w)_n &= -b\dfrac{n+3}{n+1}w_{n+1} + \Big[b + c(n-1) + d\Big]w_n - (n-3)\Big(c + \dfrac{d}{n-1}\Big)w_{n-1} \quad n \geq 3.
\end{cases}
\tag{27}
$$

The following lemma ends the proof.

Lemma 5.5 (existence) *There are solutions $u^\lambda, u^\delta, u^\alpha, u^\varepsilon, u^\sigma$ of (26) such that $u^\lambda, u^\delta, u^\alpha, u^\sigma$ and $(nu_n^\varepsilon)_n$ are bounded.*

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

5.4 Invasibility coefficients and adaptive slopes

In this subsection, we give the explicit formulae for the invasibility coefficients g^ι of Theorem 5.1 and the adaptive slopes a_ι of (18).

5.4.1 Preliminary results

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

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

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

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

Then it is elementary to check that for $k \geq 1$,

$$
Le^{(k)} = -\frac{b}{k}e^{(1)} + \frac{d}{k}e^{(-1)} - b\frac{k-1}{k}e^{(k+1)} + (b - (k+1)c + d)e^{(k)} + (k+1)(c - \frac{d}{k})e^{(k-1)},
\tag{28}
$$

and that

$$
Le^{(-1)} = -2be^{(0)} + be^{(1)} + be^{(-1)} + (c+d)\delta^{(2)}.
\tag{29}
$$

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

$$
L'e^{(k)} = -\frac{2b}{k}e^{(1)} + \frac{2d}{k}e^{(-1)} - b\frac{k-2}{k}e^{(k+1)} + (b - (k+1)c + d)e^{(k)} + (k+2)(c - \frac{d}{k})e^{(k-1)},
\tag{30}
$$

and also

$$
L'e^{(-2)} = -2(b+d)e^{(-1)} + be^{(1)} + (b+c+d)e^{(-2)} + (c + \frac{d}{2})\delta^{(3)}.
\tag{31}
$$

20

Next observe that (26) can be written in the form

$$\begin{cases} Lu^\lambda = e^{(0)} - e^{(1)} \ \ \text{and} \ \ Lu^\delta = e^{(0)}, \\\\ Lu^\alpha = Lu^\sigma = e^{(-1)} - e^{(0)}, \\\\ L'u^\varepsilon = e^{(-1)} - e^{(0)}, \end{cases} \tag{32}$$

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, all individuals are assigned distinct labels. We denote by $\mathbf{P}_n$ the law of the logistic branching process (b, c, d) starting from n individuals distinctly labelled at time 0, where the label 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

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

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 5.6 *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 $L'D^{(3)} = \delta^{(3)}$.

Moreover, the sequences $q^{(k)}$ satisfy the following property.

Lemma 5.7 *For any $k \geq 2$, $(q_n^{(k)})_n$ has a nonzero limit $q_\infty^{(k)}$ as $n \to \infty$.*

Proofs of Lemmas 5.6 and 5.7 are put to Subsection 5.5.1.

The next (sub)subsections are devoted to results related to invasibility coefficients and adaptive slopes, which prove Lemma 5.5. In Propositions 5.8, 5.9, 5.11 and 5.13, we display the solutions of (26) such that u^ι ($\iota \neq \varepsilon$) and $(nu_n^\varepsilon)_n$ are bounded (therefore proving Lemma 5.5). We also specify the behaviour of each invasibility coefficient as the population size grows to infinity. Proofs of Propositions 5.11 and 5.13 are to be found in Subsection 5.5.2. Proposition 5.14 gives explicit expressions for the adaptive slopes.

5.4.2 Results for the λ-invasibility

Here, we must find a bounded sequence u^λ in $\mathcal{L}_2$ such that $Lu^\lambda = e^{(0)} - e^{(1)}$.

Recall from Lemma 5.6 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 (29),

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

and by (28)

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

we can readily state the following

Proposition 5.8 (fertility) *The sequence u^λ defined as*

$$u^\lambda = -\frac{d}{2bc}e^{(-1)} + \frac{d(c+d)}{2bc}D^{(2)} + \frac{1}{2c}e^{(1)}, \tag{33}$$

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

$$g_n^\lambda = -\frac{dn}{2bc(n-1)} + \frac{d(c+d)}{2bc\kappa}\frac{nq_n^{(2)}}{n-1} + \frac{n}{2c(n+1)} \qquad n \geq 2. \tag{34}$$

In particular,

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

5.4.3 Results for the α and σ-invasibilities

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

Proposition 5.9 (aggressiveness, survival) *The sequences u^α and u^σ defined as*

$$u^\alpha = u^\sigma = \frac{2c-d}{2bc}e^{(-1)} - \frac{(2c-d)(c+d)}{2bc}D^{(2)} + \frac{1}{2c}e^{(1)}, \tag{35}$$

*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*

$$g_n^\alpha = g_n^\sigma = \frac{(2c-d)n}{2bc(n-1)} - \frac{(2c-d)(c+d)}{2bc\kappa}\frac{nq_n^{(2)}}{n-1} + \frac{n}{2c(n+1)} \qquad n \geq 2. \tag{36}$$

In particular,

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

5.4.4 Results for the δ-invasibility

For δ and ε-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^δ in $\mathcal{L}_2$ such that $Lu^\delta = e^{(0)}$.

Lemma 5.10 *Let Φ be the sequence of $\mathcal{L}_2$ defined recursively as $\Phi_2 = 1$ and*

$$c(n+2)\Phi_{n+1} + [b - c(n+1)]\Phi_n - b\frac{n-2}{n-1}\Phi_{n-1} = 0. \tag{37}$$

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

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

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

Proposition 5.11 (defence capacity) *Define the sequence ϕ 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^δ 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^δ associated to **defence capacity** $(g_n^\delta = nu_n^\delta)$ is given by*

$$g_n^\delta = \sum_{k \geq 1} \frac{n\phi_k}{n+k} \qquad n \geq 2. \tag{38}$$

In particular,

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

The proofs of these results are given in Subsection 5.5.

5.4.5 Results for the ε-invasibility

Recall from Lemma 5.6 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 (δ-invasibility), we assume that the resident species has no natural death rate, that is, $d = 0$.

Here, we must find a sequence u^ε 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 5.12 *Let Ψ be the sequence of $\mathcal{L}_3$ defined recursively as $\Psi_3 = 1$ and*

$$c(n+3)\Psi_{n+1} + [b - c(n+1)]\Psi_n - b\frac{n-3}{n-1}\Psi_{n-1} = 0. \tag{39}$$

Then the sequence $(n^2\Psi_n)_n$ converges to a nonzero finite limit Ψ_∞, 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 5.13 (isolation) *Define the sequence ψ of $\mathcal{L}_3$ as*

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

Then, with

$$\psi_{-2} = -\frac{1}{b(\theta+3)}$$

$$\psi_{-1} = \frac{\theta+1}{b(\theta+3)}$$

$$\psi_1 = \frac{2\theta}{3c(\theta+3)}$$

$$\psi_2 = \frac{\Sigma}{c\Psi_\infty} - \frac{2\theta+3}{3c(\theta+3)},$$

the sequence u^ε 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^ε associated to **isolation** $(g_n^\varepsilon = n^2 u_n^\varepsilon)$ is given by*

$$g_n^\varepsilon = \sum_{k\geq 1} \frac{n^2\psi_k}{n+k} + \frac{1}{\kappa'\theta(\theta+3)}\frac{n^2 q_n^{(3)}}{(n-1)(n-2)} \qquad n \geq 3. \tag{40}$$

In particular,

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

The proofs of these results are given in the Subsection 5.5.

5.4.6 Adaptive slopes

We will denote by $q_n^{(2)}(x)$, $\kappa(x)$ and $\phi_n(x)$ the quantities appearing in Propositions 5.8, 5.9 and Proposition 5.11 when the resident population has trait x (and dynamical characteristics $(b(x), c(x,x), 0)$).

The adaptive slopes appearing in the deterministic part of the canonical diffusion (15) can be expressed as follows.

Proposition 5.14 *The coefficients a_ι 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*

$$a_\lambda(x) = \frac{e^{\theta(x)}(\theta(x)^2 - 3\theta(x) + 4) - \theta(x) - 4}{2b(x)\theta(x)^2} \tag{41}$$

$$a_\alpha(x) = \frac{e^{\theta(x)}(\theta(x)^2 - \theta(x) + 2) - \theta(x) - 2}{2b(x)\theta(x)^2} - \frac{1}{\kappa(x)\theta(x)} \sum_{n=1}^{\infty} \frac{n q_{n+1}^{(2)}(x)\theta(x)^{n-1}}{(n+1)!} \tag{42}$$

$$a_\delta(x) = \sum_{k \geq 1} \frac{\phi_k(x)}{\theta(x)^{k+2}} \int_0^{\theta(x)} u^{k-1}(e^u(u^2 - u + 1) - 1)du \tag{43}$$

$$= \frac{1}{\theta(x)^3} \int_0^{\theta(x)} (e^u(u^2 - u + 1) - 1)\pi_x\left(\frac{u}{\theta(x)}\right) du \tag{44}$$

where for $k \geq 1$,

$$\int_0^{\theta} u^{k-1}(e^u(u^2 - u + 1) - 1)du = e^{\theta}\left(\theta^{k+1} - (k+2)\theta^k + (k+1)^2(k-1)! \sum_{i=0}^{k-1} \frac{(-1)^i\theta^{k-i-1}}{(k-i-1)!}\right)$$
$$- (-1)^{k-1}(k+1)^2(k-1)! - \frac{\theta^k}{k} \tag{45}$$

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

$$u^2(1-u)\pi_x''(u) + u(\theta(x)u(1-u) + 2 - 3u)\pi_x'(u) - 2\pi_x(u) + \frac{\theta(x)}{b(x)} = 0. \tag{46}$$

Proof. It follows from Proposition 5.8 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 5.9 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 (41) and (42).

For the δ-invasibility, using Proposition 5.11 and switching the two sums, we get

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

25

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 (43). (44) follows from switching the sum and the integral in (43) which is standard since $\phi_n(x) = O(1/n)$.

Equation (45) 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, (46) 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

$$0 = \sum_{n \geq 3}(n+1)(n-2)\phi_n(x)u^n + \theta(x)\sum_{n \geq 2}(n-1)\phi_n(x)u^{n+1}$$

$$- \sum_{n \geq 2}(n+1)(n-1)\phi_n(x)u^{n+1} - \theta(x)\sum_{n \geq 2}(n-1)\phi_n(x)u^{n+2}$$

$$= \frac{d}{du}\left(u^4\frac{d}{du}\left(\frac{\pi_x(u) - \phi_1(x) - \phi_2(x)u}{u}\right)\right) + \theta(x)u^3\pi_x'(u) - u\frac{d}{du}\left(u^3\pi_x'(u)\right) - \theta(x)u^4\pi_x'(u)$$

which finally gives (46). $\qquad\square$

5.5 Proofs

5.5.1 Proofs of Lemmas 5.6 and 5.7

Proof of Lemma 5.6. 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 $1, 2, \ldots, n$. Recall individuals of all types are exchangeable (because in this setting, the discrete operators Δ_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,

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

and once again by exchangeability,

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

26

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 $n \geq 3$. The proof is completed by checking that $(Lv)_2 = \kappa \neq 0$.

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

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

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

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

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

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

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 $(L'v)_3 = \kappa' \neq 0$. $\qquad \square$

Proof of Lemma 5.7. From [25], we know that ∞ 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

$$\exists k_0 \geq 2, \quad q_\infty^{(k_0)} = 0 \quad \Rightarrow \quad \mathbf{P}_\infty(N_{0+} \leq k_0) = 1. \tag{47}$$

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 (47). 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 $\mathbf{P}_\infty(N_{0+} \leq k_0) = 1$. $\qquad\square$

5.5.2 Proofs for δ and ε-invasibilities

Proof of Lemma 5.10. The proof stems immediately from the two following claims. Claim 1 will also be helpful in the proof of Proposition 5.11. $\qquad\square$

Claim 1. Assume that $(n\Phi_n)_n$ converges to a finite limit Φ_∞, where $(\Phi_n)_n$ is defined in (37). 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

$$\beta_n := (n + 1)\Phi_n \quad \text{and} \quad \gamma_n := (n - 1)\Phi_n \qquad n \geq 2, \tag{48}$$

so that

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

28

and, thanks to (37),

$$\beta_{n+1} - \beta_n = -\frac{\theta}{n-1}(\gamma_n - \gamma_{n-1}) \qquad n \geq 2, \tag{49}$$

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 (28) and (29), and by continuity of linear operators,

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

which ends the proof. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad$ $\square$

Proof of Claim 2. We split this proof into the four following steps (recall (48))

(i) if $(n\Phi_n)_n$ converges to a finite limit Φ_∞, 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 Φ_∞ 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 (26) and (27) show that $\Delta_0 w = 0$. The contradiction comes with Lemma 5.4, 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 Δ_0.

(ii) First observe that (37) reads

$$\beta_{n+1} = \left(1 - \frac{\theta}{n+1}\right)\beta_n + \theta \frac{n-2}{n(n-1)}\beta_{n-1}, \tag{50}$$

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 β_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 β 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 (49) 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). $\qquad\square$

Proof of Proposition 5.11. Thanks to Claim 1 above, since $u^\delta = \phi_1 e^{(1)} + (c\Phi_\infty)^{-1} W$,

$$\begin{aligned}
L u^\delta &= \phi_1 L e^{(1)} + (c\Phi_\infty)^{-1} L W \\
&= (2c)^{-1}(-2c e^{(1)} + 2c e^{(0)}) + (c\Phi_\infty)^{-1} c\Phi_\infty e^{(1)} \\
&= e^{(0)}.
\end{aligned}$$

The boundedness of u^δ is straightforward. To get the equivalent of g_n^δ 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

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

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

Proof of Lemma 5.12. Since proofs for the isolation ε are quite similar to those done for the defence capacity δ, we will often sketch them. The proof of Lemma 5.12 stems immediately from the following two claims. $\qquad\square$

Claim 1. Assume that $(n^2 \Psi_n)_n$ converges to a finite limit Ψ_∞, where $(\Psi_n)_n$ is defined in (39). Then

31

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

$$\beta_n := (n+2)(n+1)\Psi_n \quad \text{and} \quad \gamma_n := (n-2)(n-1)\Psi_n \qquad n \geq 3, \tag{51}$$

so that

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

and, thanks to (39),

$$\beta_{n+1} - \beta_n = -\frac{\theta(n+2)}{(n-1)(n-2)}(\gamma_n - \gamma_{n-1}) \qquad n \geq 3, \tag{52}$$

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

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

which ends the proof of (i). With the help of (30) and (31), (ii) can be proved easily mimicking what was done for the δ-invasibility. $\qquad \square$

Proof of Claim 2. We proceed just as for the defence capacity. First, we prove that if Ψ_∞ exists, it cannot be 0. Indeed, consider Σ 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 5.4.

Next recall the sequences β and γ defined in (51). Thanks to (39),

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

which proves that β_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 (52) 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. $\qquad\square$

Proof of Proposition 5.13. 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

$$u^\varepsilon = -(c\Psi_\infty)^{-1}(Z - \Sigma e^{(2)}) + V + \frac{1}{\theta(\theta+3)}D^{(3)}. \tag{53}$$

By an elementary computation relying on (30) and (31), 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^ε as $n \to \infty$, first recall (53) 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 5.11. 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 $g_n^\varepsilon = n^2 u_n^\varepsilon$. $\hspace{2cm}$ $\square$

6 Proof of the convergence to the TSS and the canonical diffusion

6.1 Preliminary result

We start this section by stating and proving a technical proposition about the two type particular case without mutation ($\mu \equiv 0$) of the GL-population of Section 2.2. In the case where $\nu_0 = X_0 \delta_x + Y_0 \delta_y$, for any $t > 0$, $\nu_t = X_t \delta_x + Y_t \delta_y$. Then, the Markov process $(X_t, Y_t; t \geq 0)$ satisfies the following properties.

Proposition 6.1 (a) *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 constant* C *independent of* x *and* y *such that*

$$
\mathbb{E}_{n,m}(\hat{T}) < C(n+m) \quad and \quad \mathbb{E}_{n,m}(\hat{T}^2) < C(n+m)^2. \tag{54}
$$

(b) $\mathbb{E}_{n,m}(X_T^2 + Y_T^2) < C(n+m)^2$ *for some constant* C *independent of* x *and* y.

(c) *With the notation of Section 4.1, when the birth rates* $b(x, y, n, m)$ *are multiplied by a positive constant* a*, the fixation probabilities* $u_{n,m}$ *are continuous as a function of* a.

Proof. **(a)** 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

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

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 $E_k(\hat{S}) \leq Ck$ and $E_k(\hat{S}^2) \leq Ck^2$ for some $C > 0$.

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 τ be the first hitting time of 0 by $\tilde{U}$. For any $k \geq 0$, one can compute explicitly (see e.g. Norris [32]) that

$$\tilde{E}_k(\tau) = 3k, \quad \tilde{E}_k(\tau^2) = 3k(3k+8) \quad \text{and} \quad \tilde{E}_k(\exp(\rho\tau)) = \exp(\sigma_\rho k)$$

for $0 \leq \rho \leq \ln(9/8)/2$ with $\exp(\sigma_\rho) = \left(1 - \sqrt{1 - 8\exp(2\rho)/9}\right)3\exp(-\rho)/2$.

Now, let k_0 be large enough to have $\underline{c}(k_0 - 1)^\alpha > 2\bar{b}$. First, observe that any excursion of $\hat{Z}$ above $k_0 + 1$ is stochastically dominated by an excursion of the random walk $\tilde{U}$ above $k_0 + 1$. Second, let

$$\hat{S}' := \inf\{j \geq k_0 : \hat{Z}_j = 1, \hat{Z}_{j-1} = 2, \cdots, \hat{Z}_{j-k_0} = k_0 + 1\}.$$

Obviously, $\hat{S} \leq \hat{S}'$. Moreover, for any n such that $\hat{Z}_n = k_0 + 1$, $\hat{S}' = n + k_0$ with probability $\beta := p_{k_0+1,k_0} \cdots p_{2,1} > 0$, and otherwise, $\hat{Z}_{n+k_0} \leq 2k_0 + 1$. Therefore, under P_k, $\hat{S}$ is dominated by

$$\tau_0 + \sum_{i=1}^{G}(k_0 + \tau_i)$$

where $G, \tau_0, \tau_1, \ldots$ are independent, τ_0 has the law of the first hitting time of $k_0 + 1$ under $\tilde{P}_{k \vee (k_0+1)}$, G is a geometric r.v. with parameter β, and the τ_i's are i.i.d. r.v. distributed as the first hitting time of $k_0 + 1$ under $\tilde{P}_{2k_0+1}$.

Therefore, if ρ is small enough to have $(1 - \beta)\exp(k_0(\rho + \sigma_\rho)) < 1$ (observe that $\sigma_\rho \to 0$ when $\rho \to 0$), then

$$E_k(e^{\rho\hat{S}}) \leq \tilde{E}_{(k-k_0-1)\vee 0}(e^{\rho\tau}) \left(\sum_{j \geq 1} \beta(1-\beta)^{j-1} \left(e^{\rho k_0}\tilde{E}_{k_0}(e^{\rho\tau})\right)^j\right)$$

$$\leq \frac{\beta}{1-\beta}e^{\mu_\rho(k-k_0)} \sum_{j \geq 1} \left[(1-\beta)e^{k_0(\rho+\sigma_\rho)}\right]^j.$$

Therefore, there exists $C, \rho, \sigma > 0$ such that

$$E_k(e^{\rho\hat{S}}) \leq Ce^{\sigma k}. \tag{55}$$

Moreover, 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

$$E_k(\hat{S}^2) \leq 2E_{k_0}(\hat{S}^2) + 2\tilde{E}_{k-k_0}(\tau^2) = 2E_{k_0}(\hat{S}^2) + 6(k - k_0)(3(k - k_0) + 8).$$

Since $E_{k_0}(\hat{S}^2) < \infty$ by (55), this ends the proof of (54).

(b)   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 $\mathbb{E}_{n,m}(X_T^2 + Y_T^2) \leq 2\mathbb{E}_{n,m}((n + m + \hat{T})^2)$ and the required bound follows from (a).

(c) Let us only denote the dependence of the fixation probability in a by $u_{n,m}(a)$, and let us denote by $\pi_{ij}(a)$ the transitions (10) of the Markov chain $(\hat{X}, \hat{Y})$ when $b(x, y, n, m)$ and $b(y, x, n, m)$ are replaced by $ab(x, y, n, m)$ and $ab(y, x, n, m)$. It suffices to prove that $u_{n,m}(a)$ is continuous at $a = 1$.

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

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

so it is sufficient to prove that the previous series (in k) is uniformly convergent for a in some neighborhood of 1.

Fix $i = (n, m)$ and $j \in \mathbb{N}^2$ such that $\|i - j\| = 1$, where $\|\cdot\|$ denotes the L^1 norm in $\mathbb{Z}^2$. It is elementary to check that $\pi_{ij}(a)/\pi_{ij}(1) \leq (a \vee 1)/(1 + (a-1)c_{n,m})$, where $c_{n,m} = B_{n,m}/(B_{n,m} + D_{n,m})$ with $B_{n,m} = nb(x, y, n, m) + mb(y, x, m, n)$ and $D_{n,m} = nd(x, y, n, m) + md(y, x, m, n)$. Because of the bounds we assumed on b and d (Definition 2.1), the $c_{n,m}$ are bounded by some constant C. Then, if $|a - 1|C < 1/2$, we finally get $\pi_{ij}(a) \leq 2(a \vee 1)\pi_{ij}(1)$, which implies the required uniform convergence. $\square$

6.2 Proofs of Theorems 3.1 and 3.2

Recall from Section 3 the definition the stopping times τ_n (n-th mutation time) and ρ_n (first time after τ_n when the population becomes monomorphic) and of the random variables V_n (the surviving trait at time ρ_n). Recall also the notation $b(x, n) \equiv b(x, n\delta_x)$ for the birth rate of an individual of trait x in a monomorphic population made of n individuals of trait x.

The proof relies on the following three lemmas. The first one states that there is no accumulation of mutations on the timescale t/γ. The second one gives the limiting laws of $\gamma\tau_1$ and of the population size at time τ_1. The last one gives the behaviour of ρ_0 and V_0 when the initial population is dimorphic.

Lemma 6.2 *Fix $C, \eta > 0$. There is $\varepsilon > 0$ such that for any $\gamma \in (0, 1)$,*

$$\mathbb{E}(\langle \nu_0^\gamma, \mathbf{1}\rangle) \leq C \quad \Rightarrow \quad \forall t \geq 0, \quad \mathbb{P}\left(\exists n \in \mathbb{N}^* : \frac{t}{\gamma} \leq \tau_n \leq \frac{t+\varepsilon}{\gamma}\right) < \eta. \tag{56}$$

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

$$\mathbb{E}(\langle \nu_0^\gamma, \mathbf{1}\rangle) \leq C \quad \Rightarrow \quad \mathbb{P}(\tau_n \leq t/\gamma) < \eta. \tag{57}$$

Lemma 6.3 *Assume $\nu_0^\gamma = n\delta_x$ where $\mu(x) > 0$.*

(a) *As $\gamma \to 0$, the pair $(\gamma\tau_1, \langle \nu_{\tau_1-}^\gamma, \mathbf{1}\rangle)$ 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 (8) and the law of N is obtained by $b(x, \cdot)$-size-biasing $\xi(x)$*

$$\mathbb{P}(N = k) = \frac{kb(x, k)\mathbb{P}(\xi(x) = k)}{\mathbb{E}(\xi(x)b(x, \xi(x)))}. \tag{58}$$

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

Lemma 6.4 *Assume $\nu_0^{\gamma} = n\delta_x + \delta_y$ (with $y \neq x$). Then*

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

(b) *$\sup_{\gamma \in (0,1)} \mathbb{E}^{\gamma}\left(\langle \nu_{\rho_0}^{\gamma}, \mathbf{1} \rangle^2 \mathbf{1}_{\{\rho_0 < \tau_1\}}\right) < \infty$.*

(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 6.2. Fix $C > 0$ and assume $\mathbb{E}(\langle \nu_0^{\gamma}, \mathbf{1} \rangle) \leq C$. By Proposition 2.3, 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 (56) 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 Poisson point process with intensity $\gamma \bar{b} \langle \nu_s, \mathbf{1} \rangle ds$. 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 (56).

Similarly, for $t \geq 0$, $\mathbb{P}(M_{t/\gamma} \geq n) \leq t\bar{b}C'/n$, which implies (57). □

Proof of Lemma 6.3. Fix $\gamma \in (0, 1)$ and assume that $\nu_0^{\gamma} = n\delta_x$. Recall that τ_1 is the first mutation time. Notice that $(\nu_t^{\gamma}; t < \tau_1)$ is distributed as $(X_t^{\gamma} \delta_x; t < \tau)$, where X^{γ} is a birth-and-death process with initial state $X_0^{\gamma} = n$ and transition rates $(1 - \gamma\mu(x))ib(x, i)$ from i to $i + 1$ and $id(x, i)$ from i to $i - 1$, and τ is the first point of a Poisson point process with inhomogeneous intensity $g^{\gamma}(X_t^{\gamma}) := \gamma\mu(x)X_t^{\gamma} b(x, X_t^{\gamma})$ (depending on X^{γ} solely through its intensity). Therefore, for any bounded function $f : \mathbb{N}^* \to \mathbb{R}$ and for any $t \geq 0$,

$$\mathbb{E}^{\gamma}(f(\langle \nu_{t/\gamma}^{\gamma}, \mathbf{1} \rangle); \gamma\tau_1 > t) = \mathbb{E}\left[f(X_{t/\gamma}^{\gamma}) \exp\left(-\int_0^{t/\gamma} g^{\gamma}(X_s^{\gamma}) ds \right) \right], \tag{59}$$

and

$$\mathbb{E}^{\gamma}(f(\langle \nu_{\tau_1 -}^{\gamma}, \mathbf{1} \rangle); \gamma\tau_1 \leq t) = \mathbb{E}\left(\int_0^{t/\gamma} f(X_s^{\gamma}) g(X_s^{\gamma}) e^{-\int_0^s g(X_u^{\gamma}) du} ds \right),$$

which yields, after a change of variable,

$$\mathbb{E}^{\gamma}(f(\langle \nu_{\tau_1 -}^{\gamma}, \mathbf{1} \rangle) \mathbf{1}_{\{\gamma\tau_1 \leq t\}}) = \mu(x) \int_0^t \mathbb{E}\left(f(X_{s/\gamma}^{\gamma}) X_{s/\gamma}^{\gamma} b(x, X_{s/\gamma}^{\gamma}) e^{-\gamma\mu(x) \int_0^{s/\gamma} X_u^{\gamma} b(x, X_u^{\gamma}) du} \right) ds. \tag{60}$$

Now, since the individual birth rates of X^{γ} decrease with γ, all the processes X^{γ} can be constructed on a same space in such a way that, for $0 \leq \gamma \leq \gamma' \leq 1$ and $t \geq 0$, $X_t^{\gamma'} \leq X_t^{\gamma} \leq X_t^0$.

To compute the limit of (60) when $\gamma \to 0$, let us first prove that

$$\lim_{\gamma \to 0} \left| \int_0^t \mathbb{E} \left(g_1(X_{s/\gamma}^\gamma) e^{-\gamma \int_0^{s/\gamma} g_2(X_u^\gamma) du} \right) ds - \int_0^t \mathbb{E} \left(g_1(X_{s/\gamma}^0) e^{-\gamma \int_0^{s/\gamma} g_2(X_u^0) du} \right) ds \right| = 0 \qquad (61)$$

for any functions g_1 and g_2 on $\mathbb{N}$ such that $|g_1(x)| \leq Cx$ and $0 \leq g_2(x) \leq Cx$. For $M > 0$, this quantity is bounded by

$$\int_0^t \mathbb{E}|g_1(X_{s/\gamma}^\gamma) - g_1(X_{s/\gamma}^0)| ds + C \int_0^t \gamma \int_0^{s/\gamma} \mathbb{E}(X_{s/\gamma}^\gamma |g_2(X_u^\gamma) - g_2(X_u^0)|) du ds$$

$$\leq 2C\gamma \int_0^{t/\gamma} \mathbb{E}(X_u^0; X_u^0 \neq X_u^\gamma) du + 2C^2 \int_0^t \gamma \int_0^{s/\gamma} \mathbb{E}(X_{s/\gamma}^\gamma X_u^0; X_u^0 \neq X_u^\gamma) du ds$$

$$\leq +2C\gamma \int_0^{t/\gamma} [\mathbb{E}(X_u^0; X_u^0 > M) + M\mathbb{P}(X_u^0 \leq M, X_u^0 \neq X_u^\gamma)] du$$

$$+ 2C^2 \int_0^t \gamma \int_0^{s/\gamma} [\mathbb{E}(X_{s/\gamma}^\gamma X_u^0; X_u^0 > M) + M\mathbb{E}(X_{s/\gamma}^\gamma; X_u^0 \leq M, X_u^0 \neq X_u^\gamma)] du ds.$$

Therefore, by Proposition 2.3, using Cauchy-Schwartz inequality to bound the term involving $\mathbb{E}(X_{s/\gamma}^\gamma X_u^0; X_u^0 > M)$ and distinguishing between $X_{s/\gamma}^\gamma \leq M$ and $X_{s/\gamma}^\gamma > M$ in the term involving $\mathbb{E}(X_{s/\gamma}^\gamma; X_u^0 \leq M, X_u^0 \neq X_u^\gamma)$, it is sufficient to prove that

$$\lim_{\gamma \to 0} \gamma \int_0^{t/\gamma} \mathbb{P}(X_u^0 \leq M, X_u^0 \neq X_u^\gamma) du = 0$$

or equivalently that γ times the expected time length between 0 and t/γ where $X_u^0 \leq M$ and $X_u^0 \neq X_u^\gamma$ goes to 0.

Since the difference between the birth rates of X_u^0 and X_u^γ when $X_u^0 = X_u^\gamma$ is less than $\gamma \bar{b} X_u^0$, any time when the two processes can start to differ belongs to the set of times when a point of a Poisson point process on $\mathbb{R}_+^2$ with intensity $\gamma \bar{b} du ds$ independent of X^0 lies below the curve $(t, X_t^0)_{t \geq 0}$. For each of these points, the time length where the two processes differ is dominated by the first hitting time of 1 by X^0 (at this time, the two processes are necessarily equal). Since, moreover, we only have to consider the time intervals where $X_u^0 \leq M$, all these time lengths are dominated by independent realizations of the hitting time of 1 by X^0 starting from M.

Let n_γ denote the number of points of the previous Poisson point process below X^0 before time t/γ and let (R_k) be a sequence of r.v. independent of n_γ distributed as the hitting time of 1 by X^0 starting from M. Then

$$\gamma \int_0^{t/\gamma} \mathbb{P}(X_u^0 \leq M, X_u^0 \neq X_u^\gamma) du \leq \gamma \mathbb{E}\left(\sum_{k=1}^{n_\gamma} R_k \right) = \gamma \mathbb{E}(R_1) \mathbb{E}(n_\gamma) = \gamma \mathbb{E}(R_1) \bar{b} \gamma \int_0^{t/\gamma} \mathbb{E}(X_u^0) du$$

which goes to 0 as $\gamma \to 0$ by Proposition 2.3.

Therefore, (61) is proved. Combined with (60), we get

$$\lim_{\gamma \to 0} \mathbb{E}^\gamma(f(\langle \nu_{\tau_1-}^\gamma, \mathbf{1} \rangle) \mathbf{1}_{\{\gamma \tau_1 \leq t\}}) = \lim_{\gamma \to 0} \mu(x) \int_0^t \mathbb{E}\left(f(X_{s/\gamma}^0) X_{s/\gamma}^0 b(x, X_{s/\gamma}^0) e^{-\gamma \mu(x) \int_0^{s/\gamma} X_u^0 b(x, X_u^0) du} \right) ds.$$

38

Since X^0 is exactly the positive-recurrent Markov chain mentioned at the end of Proposition 2.2, apply the ergodic theorem to get that $\gamma \int_0^{s/\gamma} X_u^0 b(x, X_u^0) du \to s\mathbb{E}(\xi(x)b(x, \xi(x)))$ a.s. as $\gamma \to 0$. Since $\sup_{t \geq 0} \mathbb{E}((X_t^0)^2) < \infty$ and by Lebesgue's theorem, we finally get

$$\lim_{\gamma \to 0} \mathbb{E}^\gamma(f(\langle \nu_{\tau_1-}^\gamma, \mathbf{1}\rangle)\mathbf{1}_{\{\gamma\tau_1 \leq t\}}) = \mu(x) \int_0^t \mathbb{E}\left(f(\xi(x))\xi(x)b(x, \xi(x))e^{-\mu(x)\mathbb{E}[\xi(x)b(x,\xi(x))]s}\right)$$

$$= \frac{\mathbb{E}(f(\xi(x))\xi(x)b(x, \xi(x)))}{\mathbb{E}(\xi(x)b(x, \xi(x)))} \int_0^t \beta(x)e^{-\beta(x)s}ds,$$

which completes the proof of Lemma 6.3(a).

Lemma 6.3(b) can be obtained by taking $f(x) = x^p \wedge K$ in (60), then letting first K go to infinity and next t to infinity. Then, we get that

$$\mathbb{E}_{n\delta_x}^\gamma(\langle \nu_{\tau_1}, \mathbf{1}\rangle^p) \leq \bar{b} \int_0^\infty \mathbb{E}\left[(X_{s/\gamma}^\gamma)^{p+1} \exp\left(-\gamma\mu(x)\int_0^{s/\gamma} X_u^\gamma b(x, X_u^\gamma)du\right)\right] ds$$

$$\leq \bar{b} \int_0^\infty \mathbb{E}\left[(X_{s/\gamma}^\gamma)^{p+1} \exp\left(-\gamma\mu(x)b(x, 1)L_{s/\gamma}\right)\right] ds$$

where $L_t = \int_0^t \mathbf{1}_{\{X_u^0=1\}}du$. By the ergodic theorem for X^0, L_t/t converges a.s. as $t \to \infty$ to a positive nonrandom limit l. However, we need a finer result to conclude. Fix $\lambda > 0$. Distinguishing between $L_{s/\gamma} \leq \lambda s/\gamma$ and $L_{s/\gamma} > \lambda s/\gamma$ and using Cauchy-Schwartz inequality and Proposition 2.3, one can bound from above the last displayed integral by a constant times

$$\int_0^\infty \left[\mathbb{P}(L_{s/\gamma} \leq \lambda s/\gamma)^{1/2} + \exp(-\mu(x)b(x, 1)\lambda s)\right] ds.$$

Therefore, it suffices to prove that there exist $\lambda, \lambda', C > 0$ such that $\mathbb{P}(L_t \leq \lambda t) \leq Ce^{-\lambda' t}$ for any $t \geq 0$.

Now, define recursively $t_0 = 0$, and for $i \geq 1$, $s_i = \inf\{s \geq t_{i-1} : X_s^0 = 1\}$, $t_i = \inf\{t \geq s_i : X_s^0 = 2\}$. Then for $i \geq 1$, set $T_i := t_i - s_i$ and $S_i := s_i - t_{i-1}$. By the strong Markov property, all these r.v. are independent, and more specifically $(T_i)_{i \geq 1}$ are i.i.d. exponential r.v. with parameter $b(x, 1)$ (the jump rate of X^0 from state 1), $(S_i)_{i \geq 2}$ are i.i.d. r.v. distributed as the hitting time of 1 by the process X^0 started at 2, whereas S_1 is the hitting time of 1 by the process X^0 started at n (remind that n has been defined by $\nu_0 = n\delta_x$). Then, for any $\rho, \sigma > 0$, using Chebyshev's exponential inequality to

get the last line,

$$\mathbb{P}(L_t \leq \lambda t) \leq \mathbb{P}\left(\exists k \geq 1 : \sum_{i=1}^{k} T_i \leq \lambda t \text{ and } \sum_{i=1}^{k+1}(S_i + T_i) \geq t\right)$$

$$\leq \mathbb{P}\left(\exists k \geq 1 : \sum_{i=1}^{k} T_i \leq \lambda t \text{ and } T_{k+1} + \sum_{i=1}^{k+1} S_i \geq (1-\lambda)t\right)$$

$$\leq \sum_{k=1}^{\infty} \mathbb{P}\left(\sum_{i=1}^{k} T_i \leq \lambda t\right) \mathbb{P}\left(T_{k+1} + \sum_{i=1}^{k+1} S_i \geq (1-\lambda)t\right)$$

$$\leq \mathbb{E}(e^{\sigma(T_1+S_1)}) e^{[\rho\lambda - \sigma(1-\lambda)]t} \sum_{k=1}^{\infty} \left[\mathbb{E}(e^{-\rho T_1})\mathbb{E}(e^{\sigma S_2})\right]^k.$$

Observe that $\mathbb{E}(\exp(\sigma T_1)) = b(x,1)/(b(x,1) - \sigma)$ if $\sigma < b(x,1)$. Therefore, if we can prove that there exists $\sigma > 0$ such that $\mathbb{E}(\exp(\sigma S_1)) < \infty$ (and thus $\mathbb{E}(\exp(\sigma S_2)) < \infty$), then ρ can be chosen large enough to have $\mathbb{E}(\exp(-\rho T_1))\mathbb{E}(\exp(\sigma S_2)) < 1$, and next $\lambda > 0$ can be chosen small enough to have $\rho\lambda - \sigma(1 - \lambda) < 0$. This would end the proof of Lemma 6.3(b).

Therefore, it only remains to prove that there exists $\sigma > 0$ such that $\mathbb{E}(\exp(\sigma S_1)) < \infty$. This can be done as follows. Let $\hat{X}^0$ be the discrete-time Markov chain associated with X^0 and let U_k be the holding time of X^0 in the state $\hat{X}_k^0$. Let also $\hat{S}_1$ be the first integer k such that $\hat{X}_k^0 = 1$. Then, for any k, U_k is dominated by an exponential r.v. with parameter $\kappa := b(x,1) \vee \underline{c}$, which is a lower bound for the jump rates of X^0. Therefore, if $\sigma < \kappa$ and $(R_i)_{i \geq 1}$ denote i.i.d exponential r.v. with parameter κ,

$$\mathbb{E}(e^{\sigma S_1}) \leq \sum_{k=0}^{\infty} \mathbb{E}(\exp(\sigma \sum_{i=1}^{k} R_i))\mathbb{P}(\hat{S}_1 = k)$$

$$= \mathbb{E}\left[\left(\frac{\kappa}{\kappa - \sigma}\right)^{\hat{S}_1}\right],$$

which has already been proved to be finite for small enough σ in (55). $\qquad\square$

Proof of Lemma 6.4. Before the first mutation, $\nu_t^\gamma = X_t^\gamma \delta_x + Y_t^\gamma \delta_y$ where (X_t^γ, Y_t^γ) is a two-type GL-population with transition rates

$$\begin{cases} (1 - \gamma\mu(x))nb(x,y,n,m) & \text{from } (n,m) \text{ to } (n+1,m) \\ (1 - \gamma\mu(y))mb(y,x,m,n) & \text{from } (n,m) \text{ to } (n,m+1) \\ nd(x,y,n,m) & \text{from } (n,m) \text{ to } (n-1,m) \\ mb(y,x,m,n) & \text{from } (n,m) \text{ to } (n,m-1) \\ 0 & \text{otherwise.} \end{cases}$$

On the event $\{\tau_1 > \rho_0\}$, $V_0 = y$ if and only if there is fixation in this process, $V_0 = x$ otherwise, and ρ_0 equals the fixation time T (see Section 2.2.2).

Now, by Lemma 6.2, 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 Proposition 6.1(c) that (c) holds. Finally, (b) follows from the observation that

$$\mathbb{E}_{n\delta_x + \delta_y} \left(\langle \nu_{\rho_0}^\gamma, \mathbf{1} \rangle^2 \mathbf{1}_{\{\rho_0 < \tau_1\}} \right) \leq \mathbb{E}_{n,1}((X_T^\gamma)^2 + (Y_T^\gamma)^2)$$

and from Proposition 6.1(b). $\qquad\square$

Proof of Theorem 3.1. Observe that the generator A of the process S, defined in (7), can be written as

$$A\varphi(x) = \int_{\mathcal{X}} (\varphi(x + h) - \varphi(x))\beta(x)\kappa(x, dh), \tag{62}$$

where $\beta(x)$ has been defined in (8) and where $\kappa(x, dh)$ is the **probability** measure on $\mathcal{X} - x$ defined by

$$\kappa(x, dh) = \sum_{n=1}^{\infty} u_{n,1}(x, x+h) \frac{nb(x,n)\mathbb{P}(\xi(x) = n)}{\mathbb{E}(\xi(x)b(x, \xi(x)))} M(x, dh)$$
$$+ \left(\int_{\mathbb{R}^k} \sum_{n=1}^{\infty} (1 - u_{n,1}(x, x+y)) \frac{nb(x,n)\mathbb{P}(\xi(x) = n)}{\mathbb{E}(\xi(x)b(x, \xi(x)))} M(x, dy) \right) \delta_0(dh). \tag{63}$$

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

$$S_t = U \circ P \left(\int_0^t \beta(S_s)ds \right)$$

is a Markov process with infinitesimal generator (62) [12, Chap. 6]. Let $(J_n)_{n \geq 1}$ denote the sequence of jump times of the Poisson process P and define $(T_n)_{n \geq 1}$ by $J_n = \int_0^{T_n} \beta(S_s)ds$ or $T_n = \infty$ if $\int_0^{\infty} \beta(S_s)ds < J_n$. Observe that any jump of the process S occurs at some time T_n, but that all T_n may not be effective jump times for S, because of the Dirac mass at 0 appearing in (63). As will appear below, the sequence (T_n) can be interpreted as the sequence of mutation times in the limit process. Whether an effective jump occurs at time T_n or not then corresponds to the fixation or extinction of the mutant.

Let $\mathbf{P}_x$ denote the law of ζ_t, as defined in the theorem, conditional on $\mathrm{Supp}(\zeta_0) = S_0 = x$. Fix $t > 0$, $m \in \mathbb{N}^*$, $x \in \mathcal{X}$ and a measurable subset Γ of $\mathcal{X}$. Under $\mathbf{P}_x$, T_1 and S_{T_1} are independent, T_1 is an exponential random variable with parameter $\beta(x)$, and S_{T_1} has law $\kappa(x, \cdot)$. Therefore, for any $n \geq 1$, applying the strong Markov property to the process S at time T_1 in the second line,

$$\mathbf{P}_x(T_n \leq t < T_{n+1}, \exists z \in \Gamma : \zeta_t = m\delta_z) = \mathbf{E}_x \left(\mathbf{1}_{\{S_t \in \Gamma\}} \mathbf{1}_{\{T_n \leq t < T_{n+1}\}} \mathbb{P}(\xi(S_t) = m) \right)$$
$$= \int_0^t \beta(x)e^{-\beta(x)s} \int_{\mathbb{R}^k} \mathbf{P}_{x+h}(T_{n-1} \leq t - s < T_n, \exists z \in \Gamma : \zeta_{t-s} = m\delta_z)\kappa(x, dh)ds. \tag{64}$$

Moreover,

$$\mathbf{P}_x(0 \leq t < T_1, \exists z \in \Gamma : \zeta_t = m\delta_z) = \mathbf{1}_{\{x \in \Gamma\}} e^{-\beta(x)t} \, \mathbb{P}(\xi(x) = m). \tag{65}$$

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

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

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

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

Lemma 6.5 *For any $x \in \mathcal{X}$, $m, k \geq 1$, $n \geq 0$, $t > 0$ and any measurable subset Γ 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*

$$p_0(t, \Gamma, m, x) = \mathbf{1}_{\{x \in \Gamma\}} e^{-\beta(x)t} \, \mathbb{P}(\xi(x) = m) \tag{66}$$

$$and \quad \forall n \geq 1, \quad p_n(t, \Gamma, m, x) = \int_0^t \beta(x) e^{-\beta(x)s} \int_{\mathbb{R}^k} p_{n-1}(t - s, \Gamma, m, x + h) \kappa(x, dh) ds. \tag{67}$$

Comparing (64) and (65) with (66) and (67), this lemma implies that $p_n(t, \Gamma, m, x) = \mathbf{P}_x(T_n \leq t < T_{n+1}, \exists z \in \Gamma : \zeta_t = m\delta_z)$.

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 2.3, $\sup_{\gamma \in (0,1)} \sup_{t \geq 0} \mathbb{E}(\langle \nu_t^\gamma, \mathbf{1} \rangle^p) < +\infty$ and

$$\left| \mathbb{P}_{\nu_0^\gamma}^\gamma(\exists z \in \Gamma : \nu_{t/\gamma}^\gamma = m\delta_z) - \mathbf{P}_x(\exists z \in \Gamma : \zeta_t = m\delta_z) \right|$$

$$\leq 2\mathbb{P}(N_0^\gamma > M) + \sum_{k=1}^M \left| \sum_{n=0}^\infty (p_n^\gamma(t, \Gamma, m, k\delta_x) - p_n(t, \Gamma, m, x)) \right| \mathbb{P}(N_0^\gamma = k).$$

Because of Lemma 6.2 (57), the quantity inside the absolute value in the r.h.s. of this equation vanishes as $\gamma \to 0$. Thus,

$$\lim_{\gamma \to 0} \mathbb{P}_{\nu_0^\gamma}^\gamma(\exists z \in \Gamma : \nu_{t/\gamma}^\gamma = m\delta_z) = \mathbf{P}_x(\exists z \in \Gamma : \zeta_t = m\delta_z). \tag{68}$$

Taking $\Gamma = \mathcal{X}$ and summing this relation over $m \in \mathbb{N}^*$, Fatou's lemma yields

$$\lim_{\gamma \to 0} \mathbb{P}_{\nu_0^\gamma}^\gamma(\mathrm{Supp}(\nu_{t/\gamma}^\gamma) \text{ is a singleton}) = 1. \tag{69}$$

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 (68) and (69) that

$$\lim_{\gamma \to 0} \mathbb{E}_{\nu_0^\gamma}^\gamma \left(f(\nu_{t/\gamma}^\gamma) \right) = \lim_{\gamma \to 0} \mathbb{E}_{\nu_0^\gamma}^\gamma \left(\hat{f}(\mathrm{Supp}(\nu_{t/\gamma}^\gamma)); \mathrm{Supp}(\nu_{t/\gamma}^\gamma) \text{ is a singleton and } \langle \nu_{t/\gamma}^\gamma, \mathbf{1} \rangle = m \right)$$

$$= \mathbf{E}_x(\hat{f}(S_t); N_t = m) = \mathbf{E}_x(f(\zeta_t)) \quad (70)$$

This equality generalizes to any bounded measurable $f : \mathcal{M} \to \mathbb{R}$ using once again that $\sup_{\gamma \in (0,1)} \mathbb{E}(\langle \nu_{t/\gamma}^\gamma, \mathbf{1} \rangle^p) < +\infty$. This completes the proof of Theorem 3.1 for one-dimensional distributions.

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

Proof of Lemma 6.5. Recall $\beta(x) = \mu(x)\mathbb{E}(\xi(x)b(x,\xi(x)))$. The result is trivial when the mutation rate $\mu(x)$ is 0. Let us assume that $\mu(x) \neq 0$. 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 (59) that

$$p_0^\gamma(t, \Gamma, m, k\delta_x) = \mathbf{1}_{\{x \in \Gamma\}} \mathbb{P}_{k\delta_x}^\gamma \left(\langle \nu_{t/\gamma}^\gamma, \mathbf{1} \rangle = m, \, \gamma\tau_1 > t \right)$$

$$= \mathbf{1}_{\{x \in \Gamma\}} \mathbb{E} \left[\mathbf{1}_{\{X_{t/\gamma}^\gamma = m\}} \exp\left(-\gamma\mu(x) \int_0^{t/\gamma} X_{s-}^\gamma b(x, X_{s-}^\gamma) ds \right) \right]$$

where X^γ has been defined in the proof of Lemma 6.3. Using (61) and the ergodic theorem for positive-recurrent Markov chains, we get

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

which entails (66).

Then, fix $n \geq 1$ and assume that $\lim_{\gamma \to 0} p_{n-1}^\gamma(t, \Gamma, m, k\delta_x) = p_n(t, \Gamma, m, x)$ for any $t > 0, \Gamma \subset \mathcal{X}, x \in \mathcal{X}$ and $m, k \geq 1$. Applying the strong Markov property to the process ν^γ at time τ_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 $\nu_{\tau_1-}^\gamma$, we get

$$p_n^\gamma(t, \Gamma, m, k\delta_x) = \int_{\mathbb{R}^k} \mathbb{E}_{k\delta_x}^\gamma \left[\mathbf{1}_{\{\gamma\tau_1 \leq t\}} p_{n-1}^\gamma(t - \gamma\tau_1, \Gamma, m, \langle \nu_{\tau_1-}^\gamma, \mathbf{1}\rangle \delta_x + \delta_{x+h}) \right] M(x, dh). \tag{71}$$

Now, we want to apply the strong Markov property to ν^γ at time ρ_0 to compute the quantity $p_{n-1}^\gamma(s, \Gamma, m, l\delta_x + \delta_y)$ 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

$$p_{n-1}^\gamma(s, \Gamma, m, l\delta_x + \delta_y) = \mathbb{E}_{l\delta_x + \delta_y}^\gamma \left[\mathbf{1}_{\{\rho_0 < \tau_1, \, \langle \nu_{\rho_0}^\gamma, \mathbf{1}\rangle \leq K, \, V_0 = x\}} p_{n-1}^\gamma(s - \gamma\rho_0, \Gamma, m, \langle \nu_{\rho_0}^\gamma, \mathbf{1}\rangle \delta_x) \right]$$

$$+ \mathbb{E}_{l\delta_x + \delta_y}^\gamma \left[\mathbf{1}_{\{\rho_0 < \tau_1, \, \langle \nu_{\rho_0}^\gamma, \mathbf{1}\rangle \leq K, \, V_0 = y\}} p_{n-1}^\gamma(s - \gamma\rho_0, \Gamma, m, \langle \nu_{\rho_0}^\gamma, \mathbf{1}\rangle \delta_y) \right]$$

$$+ \mathbb{P}_{l\delta_x + \delta_y}^\gamma(\{\rho_0 \geq \tau_1\} \cap E) + \mathbb{P}_{l\delta_x + \delta_y}^\gamma(\{\rho_0 < \tau_1\} \cap \{\langle \nu_{\rho_0}^\gamma, \mathbf{1}\rangle > K\} \cap E)$$

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. vanishes as $\gamma \to 0$ because of Lemma 6.4(a) and the last term vanishes as $K \to +\infty$ uniformly for $\gamma \in (0, 1)$ because of Lemma 6.4(b).

As for the first two terms, assume that $p_{n-1}^\gamma(t, \Gamma, m, k\delta_x)$ converges to $p_{n-1}(t, \Gamma, m, x)$ as in the statement of Lemma 6.5. As a consequence of Lemma 6.2 (56), for any $t > 0$, the function $s \mapsto p_{n-1}^\gamma(s, \Gamma, m, k\delta_x)$ is uniformly continuous on $[0, t]$, uniformly in γ. Combining this observation with Lemma 6.4(c),

$$\lim_{\gamma \to 0} p_{n-1}^\gamma(s, \Gamma, m, l\delta_x + \delta_y) = u_{l,1}(x, y) p_{n-1}(s, \Gamma, m, y) + (1 - u_{l,1}(x, y)) p_{n-1}(s, \Gamma, m, x). \tag{72}$$

This uniform continuity argument also applies to $s \mapsto p_{n-1}^\gamma(s, \Gamma, m, l\delta_x + \delta_y)$, so that the convergence in (72) is uniform in $s \in [0, t]$ and $l \in \{1, \ldots, L\}$, for fixed $L \geq 1$. Therefore, distinguishing between $\langle \nu_{\tau_1-}^\gamma, \mathbf{1} \rangle \leq L$ and $\langle \nu_{\tau_1-}^\gamma, \mathbf{1} \rangle > L$, we can combine Lemma 6.3(a) and (b) to get

$$\lim_{\gamma \to 0} \mathbb{E}_{k\delta_x}^\gamma \left[\mathbf{1}_{\{\gamma\tau_1 \leq t\}} p_{n-1}^\gamma(t - \gamma\tau_1, \Gamma, m, \langle \nu_{\tau_1-}^\gamma, \mathbf{1} \rangle \delta_x + \delta_y) \right] = \int_0^t ds\, \beta(x) e^{-\beta(x)s} \times$$

$$\times \sum_{l=1}^\infty \frac{lb(x,l)\mathbb{P}(\xi(x) = l)}{\mathbb{E}(\xi(x)b(x, \xi(x)))} \left[u_{l,1}(x,y) p_{n-1}(t - s, \Gamma, m, y) + (1 - u_{l,1}(x,y)) p_{n-1}(t - s, \Gamma, m, x) \right].$$

Finally, using Lebesgue's theorem, this limit applies inside the integral in (71), which gives exactly (67) and ends the proof of Lemma 6.5. $\qquad \square$

Proof of Theorem 3.2. Since the limiting law of the process $(S_t^\gamma; t \geq 0)$ is characterized by its finite-dimensional distributions, obtained in Theorem 3.1, we only have to show the tightness of their laws. Fix $T > 0$. By Ascoli's theorem for càdlàg functions (see e.g. [2]), we have to show that, for any $\varepsilon, \eta > 0$, there is $\delta > 0$ such that

$$\limsup_{\gamma \to 0} \mathbb{P}(\omega'(S^\gamma, \delta) > \eta) < \varepsilon \tag{73}$$

where the modulus of continuity ω' 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 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 ω is the usual modulus of continuity

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

Now, fix $\varepsilon > 0$ and, by Lemma 6.2 (57), choose N such that $\mathbb{P}^\gamma(\gamma\tau_N \leq T) \leq \varepsilon/2$ for any $\gamma \leq 1$. For any $n \leq N$,
$$\mathbb{P}^\gamma(\rho_{n+1} - \rho_n < \delta/\gamma) \leq \mathbb{P}^\gamma(\rho_n - \tau_n > \delta/\gamma) + \mathbb{P}^\gamma(\tau_{n+1} - \tau_n < 2\delta/\gamma).$$

By Lemma 6.2 (56) and Lemma 6.3(b), one can choose δ such that the second term is bounded by $\varepsilon/2N$ uniformly in γ. Then, by Lemma 6.4(a) and Lemma 6.3(b), the first term goes to 0 when $\gamma \to 0$. Therefore, for any $\varepsilon > 0$, there exists $\delta > 0$ such that

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

which implies (73). $\qquad \square$

6.3   Proof of Theorem 4.1

Since $v_{n,m} \equiv 0$ on $\Omega_1 \cup \Omega_2$, we will always assume $n, m \geq 1$. Recall the definition (10) of the transition probabilities $\pi_{ij}(x, y)$ $(i, j \in \mathbb{N}^2)$ of the embedded Markov chain $(\hat{X}_n, \hat{Y}_n; n \geq 0)$ associated to the

(x, y)-type GL-population $(X_t, Y_t; t \geq 0)$ without mutation. Recall also the notation $\pi_{i_1,\ldots,i_k}(x, y)$ for the product $\pi_{i_1 i_2}(x, y)\pi_{i_2 i_3}(x, y)\ldots\pi_{i_{k-1} i_k}(x, y)$ and $S_{(n,m)\to\Gamma}$ for the set of paths linking (n, m) to a subset Γ of $\mathbb{N}^2$ without hitting $\Omega_1 \cup \Omega_2$ before Γ, i.e. the set of all k-tuples $(i_1, i_2, \ldots, i_k)$ for all $k \geq 1$ such that $i_1 = (n, m)$, $i_2, \ldots, i_{k-1} \in \mathbb{N}^2 \setminus (\Omega_1 \cup \Omega_2)$ and $i_k \in \Gamma$.

Now,
$$u_{n,m}(x, y) = \sum_{k\geq 2,(i_1,\ldots,i_k)\in S_{(n,m)\to\Omega_1}} \pi_{i_1,\ldots,i_k}(x, y) \; ,$$

so if we prove that
$$R_{n,m}(x, y) := \sum_{k\geq 2,(i_1,\ldots,i_k)\in S_{(n,m)\to\Omega_1}} \left|\frac{\partial \pi_{i_1,\ldots,i_k}}{\partial y}(x, y)\right|$$

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

$$R_{n,m}(x, y) \leq C_1 \sum_{k\geq 2,(i_1,\ldots,i_k)\in S_{(n,m)\to\Omega_1}} \sum_{l=1}^{k-1} \pi_{i_1,\ldots,i_l}(x, y)\pi_{i_{l+1},\ldots,i_k}(x, y),$$

where C_1 is defined in Assumption (11) and $\pi_{i,i}(x, y) = 1$ by convention. Next, with $\|\cdot\|$ denoting the L^1-norm in $\mathbb{Z}^2$,

$$R_{n,m}(x, y) \leq C_1 \sum_{l\geq 1} \sum_{(n',m')\in(\mathbb{N}^*)^2} \sum_{(i_1,\ldots,i_l)\in S_{(n,m)\to(n',m')}} \pi_{i_1,\ldots,i_l}(x, y)\times$$
$$\times \sum_{\|\varepsilon\|=1} \sum_{k'\geq 0,(j_1,\ldots,j_{k'})\in S_{(n',m')+\varepsilon\to\Omega_1}} \pi_{j_1,\ldots,j_{k'}}(x, y)$$

with the convention that $k' = 0$ if $(n', m') + \varepsilon \in \Omega_1$, so that, if $\hat{T}_\Gamma$ denotes the first hitting of $\Gamma \subset \mathbb{N}^2$ by $(\hat{X}, \hat{Y})$,

$$R_{n,m}(x, y) \leq C_1 \sum_{l\geq 1} \sum_{(n',m')\in(\mathbb{N}^*)^2} \sum_{(i_1,\ldots,i_l)\in S_{(n,m)\to(n',m')}} \pi_{i_1,\ldots,i_l}(x, y) \sum_{\|\varepsilon\|=1} \mathbb{P}_{(n',m')+\varepsilon}(\hat{T}_{\Omega_1} < \hat{T}_{\Omega_2})$$
$$\leq 4C_1 \sum_{l\geq 1} \mathbb{P}_{n,m}(\hat{T} > l)$$
$$= 4C_1 \mathbb{E}_{n,m}(\hat{T} - 1),$$

The proof of (12) is completed thanks to Proposition 6.1(a).

(b) Following exactly the same method as above, we obtain from (54) that

$$\sum_{(i_1,\ldots,i_k)\in S_{(n,m)\to\Omega_1}}\left|\frac{\partial^2 \pi_{i_1,\ldots,i_k}}{\partial y^2}(x,y)\right|$$

$$\leq 4C_1^2\sum_{l\geq 1}\sum_{(n',m')\in(\mathbb{N}^*)^2}\sum_{(i_1,\ldots,i_l)\in S_{(n,m)\to(n',m')}}\pi_{i_1,\ldots,i_l}(x,y)\sum_{\|\varepsilon\|=1}\mathbb{E}_{(n',m')+\varepsilon}(\hat{T}-1)$$

$$+4C_2\mathbb{E}_{n,m}(\hat{T}-1)$$

$$\leq 16C_1^2 C\sum_{l\geq 1}\sum_{(n',m')\in(\mathbb{N}^*)^2}\sum_{(i_1,\ldots,i_l)\in S_{(n,m)\to(n',m')}}\pi_{i_1,\ldots,i_l}(x,y)(n'+m')+4C_2 C(n+m)$$

$$\leq 16C_1^2 C\sum_{l\geq 1}(n+m+l)\mathbb{P}_{n,m}(\hat{T}>l)+4C_2 C(n+m)$$

$$\leq 16C_1^2 C\mathbb{E}_{n,m}\left((\hat{T}-1)(n+m+\hat{T}/2)\right)+4C_2 C(n+m),$$

and the result follows again from Proposition 6.1(a). $\qquad\square$

6.4   Proof of Theorem 4.2

We will use the classical method of tightness and martingale problem formulation to prove this theorem (e.g. [20]). We divide the proof in three steps.

Step 1. Uniqueness of the limit process.   Strong existence and uniqueness for the SDE (15) follow standardly from the Lipschitz-continuity of its coefficients.

Step 2. Tightness of the family of laws of Z^ϵ.   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 β and χ 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_ϵ is the infinitesimal generator of the Markov process Z^ϵ

$$Z_t^\epsilon = Z_0^\epsilon + \epsilon\int_0^t\int_0^1\int_{\mathbb{R}^k} h\mathbf{1}_{\left\{u\leq\frac{\beta(Z_{s-}^\epsilon)}{\bar{\beta}}\frac{\chi(Z_{s-}^\epsilon, Z_{s-}^\epsilon+\epsilon h)}{\bar{\chi}}\frac{m(Z_{s-}^\epsilon, h)}{\bar{M}(h)}\right\}}N_\epsilon(dh, du, ds)$$

Since β and χ are bounded, a process generated by A_ϵ is unique in law (e.g. [12]), and this construction characterizes the law of the process Z^ϵ appearing in the statement of Theorem 4.2. 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^ϵ 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

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

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

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 measures $(\mathbf{P}_\epsilon)_{\epsilon>0}$ on $\mathbb{D}(\mathbb{R}_+, \mathbb{R}^k)$. Fix $\delta, \epsilon > 0$ and let τ and τ' be two stopping times such that $\tau < \tau' < \tau + \delta$. Since $|\chi(x, x + \epsilon h) - \chi(x, x)| \le \epsilon K\|h\|$ for some constant K, $\|\hat{Z}_{\tau'}^\epsilon - \hat{Z}_\tau^\epsilon\| \le \delta\bar{\beta}KM_2$, 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,

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

Therefore, for any $\eta > 0$,

$$\mathbf{P}_\epsilon(\|Z_{\tau'}^\epsilon - Z_\tau^\epsilon\| \ge \eta) \le \mathbf{P}_\epsilon(\|\hat{Z}_{\tau'}^\epsilon - \hat{Z}_\tau^\epsilon\| \ge \frac{\eta}{2}) + \mathbf{P}_\epsilon(\|\tilde{Z}_{\tau'}^\epsilon - \tilde{Z}_\tau^\epsilon\| \ge \frac{\eta}{2}) \le \mathbf{1}_{\{2\delta\bar{\beta}KM_2 \ge \eta\}} + \frac{4\delta\bar{\beta}\bar{\chi}M_2}{\eta^2}$$

which converges to 0 as $\delta \to 0$, uniformly w.r.t. the choice of τ and τ'. This is the first part of Aldous' criterion. For the second part, we have to prove the tightness of $(Z_t^\epsilon)_{\epsilon>0}$ for any $t \ge 0$. 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 $(\mathbf{P}_\epsilon)_{\epsilon>0}$ as $\epsilon \to 0$ on $\mathbb{D}(\mathbb{R}_+, \mathbb{R}^k)$, endowed with the canonical filtration $\mathcal{F}_t$. Since the martingale problem for (15) 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 β and χ 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 $0 \le t_1 < \ldots < t_p \le s$ for some $p \ge 1$,

and a continuous bounded function $q : (\mathbb{R}^k)^p \to \mathbb{R}$. We can write

$$\left| \mathbf{E}_0 \left\{ q(w_{t_1}, \ldots, w_{t_p}) \left[\varphi(w_t) - \varphi(w_s) - \int_s^t A_0 \varphi(w_u) du \right] \right\} \right|$$

$$\leq \left| \mathbf{E}_\epsilon \left\{ q(w_{t_1}, \ldots, w_{t_p}) \left[\varphi(w_t) - \varphi(w_s) - \int_s^t A_\epsilon \varphi(w_u) du \right] \right\} \right|$$

$$+ \left| \mathbf{E}_\epsilon \left\{ q(w_{t_1}, \ldots, w_{t_p}) \int_s^t (A_\epsilon \varphi(w_u) - A_0 \varphi(w_u)) du \right\} \right|$$

$$+ \left| \mathbf{E}_0 \left\{ q(w_{t_1}, \ldots, w_{t_p}) \left[\varphi(w_t) - \varphi(w_s) - \int_s^t A_0 \varphi(w_u) du \right] \right\} \right.$$

$$\left. - \mathbf{E}_\epsilon \left\{ q(w_{t_1}, \ldots, w_{t_p}) \left[\varphi(w_t) - \varphi(w_s) - \int_s^t A_0 \varphi(w_u) du \right] \right\} \right|.$$

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 (14), the second term converges to 0 as $\epsilon \to 0$. The third term also goes to 0 as $\epsilon \to 0$ because $\mathbf{P}_\epsilon$ converges to $\mathbf{P}_0$. Finally, since the l.h.s. does not depend on ϵ, 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^φ is a $\mathbf{P}_0$-martingale. This result can easily be extended to any $\mathcal{C}^2$ function φ by a standard truncation technique, which completes the proof of Theorem 4.2. $\quad\square$

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.* **69** 297-321.

[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] DIECKMANN, U., LAW, R. 2000. Relaxation projections and the method of moments. In: Dieckmann, U., Law, R., Metz, J.A.J. (Eds.), The Geometry of Ecological Interactions: Symplifying Spatial Complexity. Cambridge University Press, Cambridge, pp. 412–455.

[10] DURRETT, R., SCHWEINSBERG, J. 2004. Approximating selective sweeps. *Theor. Pop. Biol.* **66** 129-138.

[11] ETHERIDGE, A., PFAFFELHUBER, P. WAKOLBINGER, A. 2006. An approximate sampling formula under genetic hitchhiking. *Ann. Appl. Probab.* (in press)

[12] ETHIER, S. N., KURTZ, T. G. 1986. *Markov Processes, Characterization And Convergence.* John Wiley & Sons, New York.

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

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

[15] FOURNIER, N., MÉLÉARD, S. 2004. A microscopic probabilistic description of a locally regulated population and macroscopic approximations. *Ann. Appl. Probab.* **14**, 1880–1919.

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

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

[18] HANSEN, T.F. 1997. Stabilizing selection and the comparative analysis of adaptation. *Evolution* **51** 1341-1351.

[19] HOFBAUER, J., SIGMUND, R. 1990. Adaptive dynamics and evolutionary stability. *Appl. Math. Letters* **3** 75-79.

[20] JACOD, J., SHIRYAEV, A. N. 1987. *Limit Theorems for Stochastic Processes.* Springer-Verlag, Berlin.

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

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

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

[24] KISDI, E., GYLLENBERG, M. 2005. Adaptive dynamics and the paradigm of diversity. *J. Evol. Biol.* **18** 1170-1173.

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

[26] LAMBERT, A. 2006. Probability of fixation under weak selection: a branching process unifying approach. *Theor. Popul. Biol.* **69** 419-441.

[27] LANDE, R. 1976. Natural selection and random genetic drift in phenotypic evolution. *Evolution* **30** 314-334.

[28] 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.

[29] MAYNARD SMITH, J., HAIGH, J. 1974. The hitch-hicking effect of a favorable gene. *Genet. Res.* **23** 23-35.

[30] 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.

[31] 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.

[32] NORRIS, J.R. 1997. *Markov Chains.* Cambridge U. Press, Cambridge.

[33] SPENCER, H.G., FELDMAN, M.W. 2005. Adaptive dynamics, game theory and evolutionary population genetics. *J. Evol. Biol.* **18** 1191-1193.

NICOLAS CHAMPAGNAT
WEIERSTRASS INSTITUTE FOR APPLIED ANALYSIS AND STOCHASTICS
MOHRENSTRASSE 39
10117 BERLIN, GERMANY
E-MAIL: champagn@wias-berlin.de
URL: http://www.eleves.ens.fr/home/champagn

AMAURY LAMBERT
UNIT OF MATHEMATICAL EVOLUTIONARY BIOLOGY
UMR 7625 LABORATOIRE D'ÉCOLOGIE
ÉCOLE NORMALE SUPÉRIEURE
46 RUE D'ULM
F-75230 PARIS CEDEX 05, FRANCE;
E-MAIL: amaury.lambert@ens.fr
URL: http://www.biologie.ens.fr/ecologie/ecoevolution/lambert/index.en.html

UMR 7625 LABORATOIRE D'ÉCOLOGIE
UNIVERSITÉ PIERRE ET MARIE CURIE-PARIS6
7 QUAI SAINT BERNARD
F-75252 PARIS CEDEX 05, FRANCE