MULTISCALE ECO-EVOLUTIONARY MODELS: FROM INDIVIDUALS TO POPULATIONS

NICOLAS CHAMPAGNAT, SYLVIE MÉLÉARD, AND VIET CHI TRAN

ABSTRACT

Motivated by recent biological experiments, we emphasize the effects of small and random populations in various biological/medical contexts related to evolution such as invasion of mutant cells or emergence of antibiotic resistances. Our main mathematical challenge is to quantify such effects on macroscopic approximations. The individual behaviors are described by the mean of stochastic multiscale models. The latter, in the limit of large population and according to the assumptions on mutation size and frequency, converge to different macroscopic equations. Sufficiently rare mutations yield a timescale separation between competition and mutation. In that case, the stochastic measure-valued process at the mutation timescale converges to a jump process which describes the successive invasions of successful mutants. The gene transfer can drastically affect the evolutionary outcomes. For faster mutation timescales, numerical simulations indicate that these models exhibit as cyclic behaviors. Mathematically, population sizes and times are considered on a log-scale to keep track of small subpopulations that have negligible sizes compared with the size of the resident population. Explicit criteria on the model parameters are given to characterize the possible evolutionary outcomes. The impact of these time and size scales on macroscopic approximations is also investigated, leading to a new class of Hamilton–Jacobi equations with state constraint boundary conditions.

MATHEMATICS SUBJECT CLASSIFICATION 2020

Primary 92D25; Secondary 92D15, 60K35, 60F99

KEYWORDS

Multiscale models, stochastic individual-based models, measure-valued Markov processes, long time behaviour, large population approximations, branching processes, partial differential equations

INTERNATIONAL CONGRESS
OF MATHEMATICIANS
2022 JULY 6–14

Proc. Int. Cong. Math. 2022, Vol. 7, pp. 5656–5678
DOI 10.4171/ICM2022/24

1. INTRODUCTION AND PRESENTATION OF THE INDIVIDUAL-BASED MODEL

Since Darwin's revolutionary work on evolution and natural selection [18], many mathematicians have worked on modeling his theories. Different schools of thought have developed, involving different classes of mathematical objects. Ecological models of structured population dynamics usually rely on deterministic models in large populations, such as dynamical systems (as the famous Lotka–Volterra system) and partial differential equations. Population genetics are more interested in random fluctuations of gene frequencies in small populations (like in the Wright–Fisher model) and therefore make extensive use of probabilistic tools. A few decades ago, eco-evolutionary models emerged, seeking to link these two approaches. Our work is placed in this framework. Our point of view consists in focusing on stochastic individual behaviors, taking into account demographic parameters (birth and death rates), evolutionary parameters (mutations, gene transfer), and ecological parameters (interactions between individuals), all these parameters depending on the genetic or phenotypic characteristics of the individual. This point of view is strongly reinforced by the ability of biologists to obtain more and more individual data, for example, for bacteria, thanks to single cell microscopes or microfluidic techniques. The notion of individual variability took a long time to emerge, especially for the biology of microorganisms, and it was not until the 2000s that biologists began to take it into account [24, 39].

There are three main sources of randomness in eco-evolutionary mechanisms which happen at different time and size scales: at the molecular level (errors in DNA replication or genetic information exchanges), at the individual level (division time, life span, contacts, access to resources), and at a macroscopic level (environmental variations). Mathematically, it is very exciting that all the parameters we have mentioned have their own scales, which can be different according to the species considered and also can vary according to the environment. Depending on these scales, the mathematical models and the associated mathematical questions can be of different nature and challenging, and open new fields of investigation.

We consider bacteria or cell populations. The ability of an individual to survive or divide depends on phenotypic or genetic parameters whose quantitative expression (real or vectorial) is called a trait. The evolution of the trait distribution results from different main mechanisms. The heredity is the vertical transmission of the ancestral trait to offspring, except when a mutation occurs. Mutations generate trait variability in the population. The selection process takes place at two levels. The variability in traits allows an individual with a higher probability of survival or a better ability to reproduce to create a subpopulation of offspring that will invade the population (genetic selection). In addition, selection also favors those individuals best able to survive in competition with others (ecological selection). Although their reproduction is asexual, bacteria or cells can also horizontally exchange genetic information during their life. Horizontal gene transfer is obtained by direct contact between cells, either by the transfer of small parts of chromosomal DNA or by the transfer of plasmids, small circular double-stranded DNA structures which can be very costly for the cell in terms of energy used. Gene transfer plays an essential role in the evolution, maintenance, transmission of virulence and antibiotic resistance.

Our goal in this paper is to show the richness of models, mathematical questions and theorems that can emerge from these eco-evolutionary dynamics and from the understanding of their long-term evolution. One is faced with the fundamental question: how to describe and quantify the successive invasions of favorable mutants? All our constructions will be based on the stochastic behavior of the individuals from which we will derive different macroscopic approximations depending on the parameter assumptions.

The seminal papers concerning eco-evolutionary modeling are based on game theory and dynamical systems, see Hofbauer–Sigmund [28], Marrow–Law–Cannings [33], Metz et al. [35,36]. Then more general models for structured populations have been introduced based either on partial differential equations, see, for example, the founding papers of Diekmann [21], Diekmann–Jabin–Mischler–Perthame [22], Barles–Mirrahimi–Perthame [3], Desvillettes–Jabin–Mischler–Raoul [19], or on stochastic individual-based models as in the theoretical biological papers by Dieckmann–Law [20], Bolker–Paccala [9], or in the rigorous mathematical papers by Fournier–Méléard [25], Champagnat–Ferrière–Méléard [13], Champagnat [11], Champagnat–Méléard [15]. Models including horizontal transfer have been proposed in the literature based on the seminal contribution of Anderson and May on host-pathogen deterministic population dynamics [1] (see also Levin et al. [30,40]) or on a population genetics framework without ecological concern (see [4,38,41]).

The basis of our approach is a stochastic individual-based model: it is a pure jump point measure-valued process in continuous time, weighted by the carrying capacity K of the system (order of magnitude of the population size), whose jump events are births with or without mutation, transfers, and deaths. The jump rates depend on the trait value of each individual, on the total population and for some of them on K. From this basic process, one can derive different approximations following the main biological assumptions of the adaptive biology. The population size is assumed to be large ($K \to \infty$), but we will also need to keep track of small populations. Mutations are rare (p_K tends to 0), but not necessarily from the population standpoint, depending on whether Kp_K tends to 0 or not. Mutation steps in the trait space may be considered small or not. The population process will be considered on different time scales: of order 1, of order $\frac{1}{Kp_K}$, or of order $\log K$.

After introducing in Section 2 the individual-based model scaled by the carrying capacity K, we will study in Section 3 large population limits on finite time intervals when K tends to infinity, using ideas developed in [25]. The stochastic process is shown to converge to the unique solution of a nonlinear integro-differential equation (see also Billiard et al. [5,6] for models with horizontal transfer). In the case where the trait support is composed of two values, the equation reduces to a nonstandard two-dimensional dynamical system whose long-time behavior is studied. In Section 4, we analyze the invasion probability and time to fixation of an initially rare mutant population. In this case, the stochastic behavior of the mutant population is fundamental and needs to be combined with the deterministic approximation of the resident population size. In Section 5 we assume that mutations are rare at the population scale to imply a separation between the competition and mutation time scales, following ideas of [11,13,15]. Under an "invasion implies fixation" assumption, a pure jump (single support) measure-valued process is derived from the population process at the mutation time scale. When the mutation steps tend to 0, a limiting differential equation for the support dynamics is also derived in a longer time scale. These results are illustrated by simulations of a simple model in Section 6. Depending on the transfer rate, we obtain dramatically different behaviors, ranging from expected evolution toward the optimal trait, to extinction (evolutionary suicide). When the individual mutation rate is small, but not from the population standpoint, intermediary values of transfer rates lead to surprising cyclic behaviors related to reemergence of traits. To capture these phenomena, we consider in Section 7 the small populations of order K^{β_K} for $0 < \beta_K \leq 1$ that can be observed in the long time scale $\log K$. We study the asymptotic dynamics of the exponents $(\beta_K(t), t \geq 0)$ and analyze the first reemergence of the optimal trait. In Section 8, under the additional assumption that the individual mutations are small, we establish in a simple framework that the stochastic discrete exponent process converges to the viscosity solution of a Hamilton–Jacobi equation with state constraint boundary conditions, allowing us to fill the gap between the stochastic [11,15] and deterministic [3,22] approaches of Dirac concentration in adaptive dynamics. In the coming years, we hope to generalize this result in a much more general framework.

Notation. The set E being a Polish space, the Skorohod space $\mathbb{D}([0,T],E)$ is the functional space of right-continuous and left-limited functions from $[0,T]$ to E. It is endowed by the Skorohod topology (cf. Billingsley [7]) which makes it a Polish space.

2. A GENERAL STOCHASTIC INDIVIDUAL-BASED MODEL FOR VERTICAL AND HORIZONTAL TRAIT TRANSMISSION

2.1. The model

The population dynamics is described by a stochastic system of interacting individuals (cf. [12,13,25]). The individuals are characterized by a quantitative parameter x, called trait, belonging to a compact subset $\mathcal{X}$ of $\mathbb{R}^d$, which summarizes the phenotypic or genotypic information of each individual. The trait determines the demographic rates. It is inherited from parent to offspring, except when a mutation occurs, in which case the trait of the offspring takes a new value. It can also be transmitted by horizontal transfer from an individual to another one. The demographic and ecological rates are scaled by the *carrying capacity* K which is taken as a measure of the "system size" (resource limitation, living area, initial number of individuals). We will derive macroscopic behaviors for the population by letting K tend to infinity with the appropriate scaling $\frac{1}{K}$ for individuals' weight.

At each time t, the population state at time t is described by the point measure

$$\nu_t^K(dx) = \frac{1}{K}\sum_{i=1}^{N_t^K} \delta_{X_i(t)}(dx), \quad N_t^K = K\int \nu_t^K(dx),$$

where $X_i(t)$ is the trait of the ith individual living at t, individuals being ranked according to the lexicographic order of their trait values. Recall that notation δ_x means the Dirac measure at x. Later we will denote indifferently, for a measurable bounded function f on $\mathbb{R}^d$, $\langle \nu_t^K, f\rangle = \int_{\mathbb{R}^d} f(x)\nu_t^K(dx) = \sum_{i=1}^{N_t^K} f(X_i(t))/K$.

The right-continuous and left-limited measure-valued process $(\nu_t^K, t \geq 0)$ is a Markov process whose transitions are described as follows. An individual with trait x gives birth to a new individual with rate $b(x)$. With probability $1 - p_K$, the new individual carries the trait x and with probability p_K, there is a mutation on the trait. The trait z of the new individual is chosen according to the probability distribution $m(x, dz)$. An individual with trait x dies with intrinsic death rate $d(x)$ and from the competition with any other individual alive at the same time. If the competitor has the trait y, the competition death rate is $\frac{C(x,y)}{K}$, leading for a population $\nu = \frac{1}{K}\sum_{i=1}^n \delta_{x_i}$ to a total individual death rate $d(x) + \frac{1}{K}\sum_{i=1}^n C(x, x_i) = d(x) + C * \nu(x)$. Horizontal transfers can occur from individuals x to y, or vice versa. In a population ν, an individual with trait x chooses a partner with trait y at rate $\frac{1}{K}\frac{\tau(x,y)}{\langle \nu,1\rangle}$. After transfer, (x, y) becomes (x, x).

2.2. Generator

We denote by $\mathcal{M}_K$ the set of point measures on $\mathcal{X}$ weighted by $1/K$ and by $\mathcal{M}_F$ the set of finite measures on $\mathcal{X}$. The generator of the process $(\nu_t^K)_{t\geq 0}$ is given for measurable bounded functions F on $\mathcal{M}_K$ and $\nu = \frac{1}{K}\sum_{i=1}^n \delta_{x_i}$ by

$$\begin{aligned}
&\sum_{i=1}^n b(x_i)\biggl((1-p_K)\biggl(F\biggl(\nu + \frac{1}{K}\delta_{x_i}\biggr) - F(\nu)\biggr) \\
&\quad + p_K \int_{\mathcal{X}} \biggl(F\biggl(\nu + \frac{1}{K}\delta_z\biggr) - F(\nu)\biggr) m(x_i, dz)\biggr) \\
&\quad + \sum_{i=1}^n \bigl(d(x_i) + C * \nu(x_i)\bigr)\biggl(F\biggl(\nu - \frac{1}{K}\delta_{x_i}\biggr) - F(\nu)\biggr) \\
&\quad + \sum_{i,j=1}^n \frac{\tau(x_i, x_j)}{K\langle \nu, 1\rangle}\biggl(F\biggl(\nu + \frac{1}{K}\delta_{x_i} - \frac{1}{K}\delta_{x_j}\biggr) - F(\nu)\biggr).
\end{aligned}$$

It is standard to construct the measure-valued process ν^K as the solution of a stochastic differential equation driven by Poisson point measures and to derive the following moment and martingale properties (see, for example, [25] or Bansaye–Méléard [2]).

Theorem 2.1. *Under the previous assumptions and assuming also that for some* $p \geq 2$, $\mathbb{E}(\langle \nu_0^K, 1\rangle^p) < \infty$, *the following properties hold. For a bounded measurable function* f *on* $\mathcal{X}$,

$$\begin{aligned}
\int f(x)\nu_t^K(dx) = \int f(x)\nu_0^K(dx) + M_t^{K,f} \\
&+ \int_0^t \int_{\mathcal{X}} \biggl\{\bigl((1-p_K)b(x) - d(x) - C * \nu_s^K(x)\bigr) f(x) \\
&+ p_K b(x) \int_{\mathcal{X}} f(z) m(x, dz) \\
&+ \int_{\mathcal{X}} \frac{\tau(x, y)}{\langle \nu_s^K, 1\rangle}\bigl(f(x) - f(y)\bigr)\nu_s^K(dy)\biggr\}\nu_s^K(dx)ds,
\end{aligned}$$

where $M^{K,f}$ is a right-continuous and left-limited square-integrable martingale starting from 0 *with quadratic variation*

$$\begin{aligned}\langle M^{K,f}\rangle_t = \frac{1}{K}\int_0^t \int_{\mathcal{X}} \Big\{ & \big((1-p_K)b(x) + d(x) + C * \nu_s^K(x)\big) f^2(x) \\ & + p_K b(x) \int_{\mathcal{X}} f^2(z) m(x,dz) \\ & + \int_{\mathcal{X}} \frac{\tau(x,y)}{\langle \nu_s^K, 1\rangle} \big(f(x)-f(y)\big)^2 \nu_s^K(dy)\Big\} \nu_s^K(dx)ds.\end{aligned}$$

3. LARGE POPULATION LIMIT AND RARE MUTATION IN THE ECOLOGICAL TIME-SCALE

3.1. A deterministic approximation

Assuming that p_K converges to p when K tends to infinity, we derive a macroscopic approximation of the population process on any finite time interval.

Assumptions (H). (i) When $K \to +\infty$, the stochastic initial point measures ν_0^K converge in probability (and for the weak topology) to the deterministic measure $u_0 \in \mathcal{M}_F(\mathcal{X})$ and $\sup_K \mathbb{E}(\langle \nu_0^K, 1\rangle^3) < +\infty$.

(ii) The functions b, d, C, and τ are continuous. The intrinsic growth rate of the subpopulation of trait x is denoted by $r(x) = b(x) - d(x)$. For any $x, y \in \mathcal{X}$, we also assume $r(x) > 0$, $C(x,y) > 0$. It means that, in absence of competition, the subpopulation with trait x has a tendency to grow and the regulation of the population size comes from the competition pressure.

Proposition 3.1. *Assume (H) and that $p_K \to p$ when K tends to infinity. Then, for $T > 0$ and when $K \to \infty$, the sequence $(\nu^K)_{K\geq 1}$ converges in probability in $\mathbb{D}([0,T], \mathcal{M}_F(\mathcal{X}))$ to the deterministic function $u \in \mathcal{C}([0,T], \mathcal{M}_F(\mathcal{X}))$, the unique weak measure-solution of*

$$\begin{aligned}\partial_t u(t,x) = & \big(r(x) - C * u(t,x)\big)u(t,x) + p\int_{\mathcal{X}} b(y)m(y,x)u(t,y)dy \\ & + \frac{u(t,x)}{\|u(t,\cdot)\|_1}\int_{\mathcal{X}} \alpha(x,y)u(t,y)dy, \qquad (3.1)\end{aligned}$$

with $C * u(t,x) = \int C(x,y)u(t,y)dy$ and $\alpha(x,y) = \tau(x,y) - \tau(y,x)$.

The proof is standard and consists of a tightness and uniqueness argument, see [2,25] or [6] for details. Let us note that the horizontal transfer acts on the dynamics (3.1) through the "horizontal flux" rate α which quantifies the asymmetry between transfers and can be positive as well as negative (or zero in the case of perfectly symmetrical transfer). Nevertheless, the fully stochastic population process depends not only on α but also on τ itself. Let us mention that, to the best of our knowledge, the long-time behavior of a solution of (3.1) is unknown, except in the case without transfer studied by Desvillettes et al. [19]. The existence of steady-states for some similar equations has been studied in Hinow et al. [27] and Magal–Raoul [32].

3.2. Particular cases when $p = 0$

Standard biological observations lead us to assume small individual mutation rate,

$$\lim_{K\to\infty} p_K = 0. \tag{3.2}$$

Under this assumption, the mutational term in (3.1) disappears, meaning that mutation events are too rare to be observed at the demographic/ecological timescale (of births, deaths, and interaction). In the particular case when the support of the initial measure u_0 is a single point x, i.e., $u_0 = n_x(0)\delta_x$, $n_x(0) \in \mathbb{R}_+$, the support of the measure u_t is $\{x\}$ for all $t > 0$ and $u_t = n_x(t)\delta_x$. From (3.1), we deduce that $n_x(t)$ is the solution of the logistic equation

$$n'_x(t) = n_x(t)\big(r(x) - C(x,x)n_x(t)\big).$$

This equation has a unique stable equilibrium

$$\bar{n}_x = \frac{r(x)}{C(x,x)}. \tag{3.3}$$

Similarly, in the case when the support of u_0 is composed of two points x and y, i.e., $u_0 = n_x(0)\delta_x + n_y(0)\delta_y$, $n_x(0), n_y(0) \in \mathbb{R}_+$, the support of the measure u_t is $\{x, y\}$ for all $t > 0$ and $u_t = n_x(t)\delta_x + n_y(t)\delta_y$, and $(n_x(t), n_y(t))$ is the solution of the dynamical system

$$\begin{aligned} \frac{dn_x}{dt} &= \left(r(x) - C(x,x)n_x - C(x,y)n_y + \frac{\alpha(x,y)}{(n_x+n_y)}n_y\right)n_x, \\ \frac{dn_y}{dt} &= \left(r(y) - C(y,x)n_x - C(y,y)n_y - \frac{\alpha(x,y)}{(n_x+n_y)}n_x\right)n_y. \end{aligned} \tag{3.4}$$

This system can be seen as a perturbation of a competitive Lotka–Volterra system, but presents more possible limit behaviors (but no cycles, see [5] for a detailed study). It is easy to see that trait y will invade a resident population of trait x and get fixed if and only if

$$r(y) - r(x) + \alpha(y,x) > 0. \tag{3.5}$$

In particular, the horizontal transfer can revert the outcome of the dynamical system without transfer, provided that $|\alpha(y,x)| > |r(y) - r(x)|$ and $\operatorname{sign}(\alpha(y,x)) = -\operatorname{sign}(r(y) - r(x))$, where $\operatorname{sign}(x) = 1$ if $x > 0$; 0 if $x = 0$; -1 if $x < 0$.

The situation is even simpler if the function C is constant. The system becomes

$$\begin{aligned} \frac{dn}{dt} &= n\big(qr(x) + (1-q)r(y) - Cn\big), \\ \frac{dq}{dt} &= q(1-q)\big(r(y) - r(x) + \alpha(y,x)\big), \end{aligned}$$

where $n = n_x + n_y$ and $q = n_x/(n_x + n_y)$. There are only two equilibria for the second equation, $q = 0$ and $q = 1$, corresponding to the equilibria $(\frac{r(x)}{C}, 1)$ and $(\frac{r(y)}{C}, 0)$, respectively. This illustrates an important assumption, called the 'invasion implies fixation" principle (IIF).

Assumption (IIF). Given any $x \in \mathcal{X}$ and Lebesgue almost any $y \in \mathcal{X}$, either $(\bar{n}_x, 0)$ is a stable steady state of (3.4), or $(\bar{n}_x, 0)$ and $(0, \bar{n}_y)$ are respectively unstable and stable steady states, and any solution of (3.4) with an initial state in $(\mathbb{R}_+^*)^2$ converges to $(0, \bar{n}_y)$ when $t \to \infty$.

Biologically speaking, this means that the ecological coefficients impede the coexistence of two traits (which is biologically accepted when there is only one type of resource, see [14]).

4. RARE MUTATIONS – FIXATION PROBABILITY

For this section, we refer to [11,13,15] for rigorous proofs.

Let us now assume (3.2) and that the resident population is uniquely composed of individuals with trait x and near its size equilibrium, i.e., when K is large, the population size $N^{x,K}$ is then close to the equilibrium $K\bar{n}^x$. Let us now investigate the fate of a newly mutated individual with trait y in this resident population, as observed in Figure 1. When the mutant appears, it begins to develop (by heredity) a small population with trait y whose size is initially negligible. During this first phase, the number $N^{y,K}$ of individuals with trait y is very small with respect to $N^{x,K}$. Its dynamics can be approximated by a linear birth and death stochastic process, at least until it reaches the threshold ηK, for a given small $\eta > 0$. The transfer $x \to y$ acts as a birth term and the transfer $y \to x$ as a death term. Therefore, the growth rate of an individual with trait y for this first phase is approximately given by

$$S(y;x) = r(y) - C(y,x)\bar{n}_x + \alpha(y,x) = r(y) - C(y,x)\frac{r(x)}{C(x,x)} + \alpha(y,x). \qquad (4.1)$$

The quantity $S(y;x)$ is called invasion fitness of trait y in the resident population of trait x. Note that S is not symmetric and null on the diagonal; for C constant, it is given by (3.5). When K tends to infinity, the probability for the process $N^{y,K}$ to reach ηK (for some $\eta > 0$) is approximately the survival probability of the underlying linear birth and death process, i.e., the positive part of the growth rate $S(y;x)$ divided by the birth rate $b(y) + \tau(y,x)$,

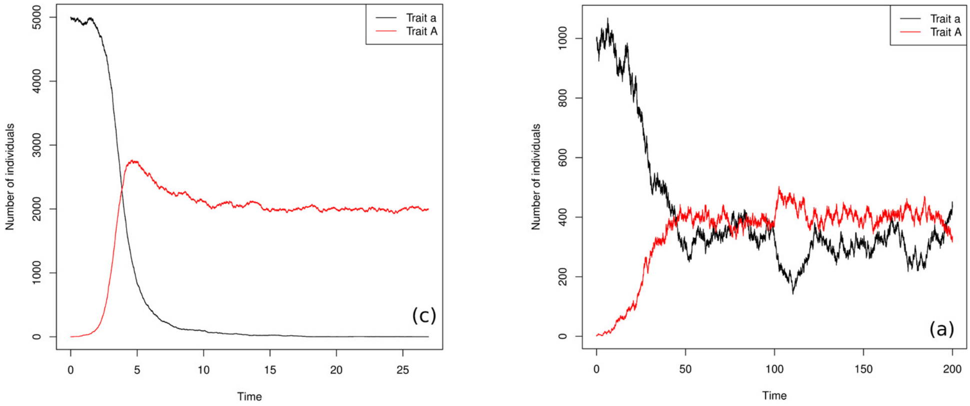


FIGURE 1

Invasion and fixation or polymorphic persistence of a deleterious mutation for unilateral transfer rate: (left) $C \equiv 1$, $b(y) = 0.5$, $b(x) = 1$, $d(x) = d(y) = 0$, $K = 5\,000$, $\alpha(y,x) = \tau(y,x) = 0.7$; (right) $C(y,x) = C(x,x) = 2$, $C(y,y) = 4$, $C(x,y) = 1$, $b(y) = 0.8$, $b(x) = 1$, $d(x) = d(y) = 0$, $K = 1\,000$, $\alpha(y,x) = \tau(y,x) = 0.5$.

namely

$$P(y;x) = \frac{[r(y) - C(y,x)\bar{n}^x + \alpha(y,x)]_+}{b(y) + \tau(y,x)}. \tag{4.2}$$

In particular, invasion is impossible if $S(y;x) \le 0$.

Let us assume that $S(y;x) > 0$. Then, the duration of the first phase (growth of the y-population from 1 to ηK individuals) is of order $\log K / S(y;x)$. It can be proved rigorously but, to be convinced of this, it is enough to notice that if t is the time elapsed from the appearance of the single mutant individual with trait y to threshold ηK, then $\mathbb{E}(N_t^{y,K}) \approx e^{S(y;x)t} = \eta K$, and $t = \log K / S(y;x)$. Then the second phase begins, where the processes $(N^{x,K}, N^{y,K})$ stay close to the dynamical system (3.4) with nonnegligible initial data η. Under Assumption (IIF), the trait y invades the population and the x-population size decreases to $N_t^{x,K} < \eta K$ in a duration of order of magnitude 1. Should the latter happen, the third phase begins and $N^{x,K}$ can be approximated by a subcritical linear birth and death process, until y is fixed and x is lost. In this case, the transfer $y \to x$ acts as a birth term and the transfer $x \to y$ as a death term. The duration of this third phase behaves as $\log K/(d-b)$ when $K \to \infty$ (see [34, SECTION 5.5.3, P. 190] for precise computation) where $b = b(x) + \tau(x,y)$, $d = d(x) + \frac{C(x,y)r(y)}{C(y,y)} + \tau(y,x)$. Summing up, the fixation time of an initially rare trait y going to fixation is of order

$$T_{\text{fix}} = \log K \left(\frac{1}{S(y;x)} + \frac{1}{|S(x;y)|} \right) + o(\log K). \tag{4.3}$$

5. VERY RARE MUTATIONS IN AN EVOLUTIONARY TIME SCALE

We wish to rigorously define and quantify the evolutionary process describing the successive invasions of successful mutants under hypothesis (3.2). In Section 3, mutations are not seen in the limit. *To observe the dynamical impact of mutations, we have to wait for a longer time than* $O(1)$. Depending on the rate of convergence of p_K to 0, different timescales will be considered in the next sections.

We assume here that not only $p_K \to 0$ but also $K\,p_K \to 0$, meaning that both individual and population mutation rates are small. We will consider the behavior of the population process at the very long time scale $\frac{1}{K\,p_K}$. Moreover, we will assume that

$$\forall\, V > 0, \quad \log K \ll \frac{1}{K\,p_K} \ll e^{VK}. \tag{5.1}$$

This assumption leads to a separation of time scales between competition phases and mutation arrivals. Indeed, by (4.3), mutations are rare enough so that the selection has time to eliminate deleterious traits or to fix advantageous traits before the arrival of a new mutant.

5.1. Trait substitution sequence

Let us study the convergence of the process $(\nu^K_{./(K p_K)})_{K \ge 1}$ when K tends to infinity, under the assumption (5.1). By simplicity we assume the *invasion implies fixation* (IIF) principle. This implies that, for a monomorphic ancestral population, the dynamics at the time scale $t/(Kp_K)$ can be approximated by a jump process over singleton measures on $\mathcal{X}$ whose mass at any time is at equilibrium. More precisely, we have

Theorem 5.1. *Assume* (H), (5.1), *and* (IIF). *Suppose the initial conditions are* $\nu_0^K(dx) = N_0^K \delta_{x_0}(dx)$ *with* $x_0 \in \mathcal{X}$, $\lim_{K\to\infty} N_0^K = \bar{n}_{x_0}$ *in probability, and* $\sup_{K\in\mathbb{N}^*} \mathbb{E}((N_0^K)^3) < +\infty$.

Then, the sequence of processes $(\nu^K_{\cdot/(Kp_K)})_{K\geq 1}$ *converges in law (for finite-dimensional distributions) to the* $\mathcal{M}_F(\mathcal{X})$*-valued process* $(V_t(dx) = \bar{n}_{Y_t}\delta_{Y_t}(dx), t \geq 0)$ *where* $(Y_t)_{t\geq 0}$ *is a pure jump process on* $\mathcal{X}$*, started at* x_0*, with the jump measure from* x *to* y *being*

$$b(x)\bar{n}_x P(y;x)m(x,dy) \tag{5.2}$$

and $P(y;x)$ *being defined in* (4.2).

The jump process $(Y_t, t \geq 0)$ (with $Y_0 = x_0$) describes the support of $(V_t, t \geq 0)$. It has been heuristically introduced in **[35]** and rigorously studied in **[11]**, in the case without transfer. It is often called the trait substitution sequence (TSS). Theorem 5.1 can be generalized when the assumption (IIF) is not satisfied, see **[15]**.

Main ideas for the proof of Theorem 5.1. The proof is a direct adaptation of **[11]**. The birth and death rates of the resident x and mutant y are

$$b(x) + \frac{\tau(x,y)N^{y,K}}{N^K}, \quad d(x) + C(x,x)N^{x,K} + C(x,y)N^{y,K} + \frac{\tau(y,x)N^{y,K}}{N^K},$$
$$b(y) + \frac{\tau(y,x)N^{x,K}}{N^K}, \quad d(y) + C(y,x)N^{x,K} + C(y,y)N^{y,K} + \frac{\tau(x,y)N^{x,K}}{N^K}.$$

The proof consists in combining (5.1), the results in Section 4, and the Markov property. Let us fix $\eta > 0$. At $t = 0$, the population is monomorphic with trait x_0 and satisfies the assumptions of Theorem 5.1. As long as no mutation occurs, the population stays monomorphic with trait x_0 and, for t and K large enough, the density process $\langle \nu_t^K, \mathbf{1}_{x_0}\rangle$ belongs to the η-neighborhood of $\bar{n}_{x_0}$ with large probability (cf. Proposition 3.1). From the large deviations principle (see Freidlin–Wentzell **[26]**), one deduces that the time taken by the density process in absence of mutations to leave the η-neighborhood of $\bar{n}_{x_0}$ is larger than $\exp(VK)$, for some $V > 0$, with high probability. Hence assumption (5.1) ensures that the approximation of the population process by $\bar{n}^{x_0}\delta_{x_0}$ stays valid until the first mutation occurrence.

The invasion dynamics of a mutant with trait y in the resident population has been studied in Section 4. If $S(y;x_0) > 0$, the process $N^{y,K}$ is supercritical, and therefore, for large K, the probability for the mutant population's density to attain η is close to the probability $P(y;x_0)$. After this threshold and thanks to Assumption (IIF), the density process $(\langle \nu^K_{\frac{t}{Kp_K}}, \mathbf{1}_{x_0}\rangle, \langle \nu^K_{\frac{t}{Kp_K}}, \mathbf{1}_y\rangle)$ will attain, when K tends to infinity, an η-neighborhood of the unique stable equilibrium $(0, \bar{n}_y)$ of (3.4) and will stabilize around this equilibrium. We have shown in Section 4 that the time elapsed between the occurrence of the mutant and the final stabilization is given by (4.3). Hence, if $\log K \ll \frac{1}{Kp_K}$, with a large probability this phase of competition–stabilization will be complete before the occurrence of the next mutation. Using Markovian arguments, we reiterate the reasoning after each mutation event. Therefore, the population process on the time-scale t/Kp_K only keeps in the limit the successive stationary states corresponding to successive advantageous mutations. If the process belongs to an η-neighborhood of $\bar{n}_x$, the mutation rate from an individual with trait x is close to $Kp_K b(x)\bar{n}_x$. At the time scale $\frac{t}{Kp_K}$, it becomes $b(x)\bar{n}_x$. The limiting process is a pure jump process $(V_t, t \geq 0)$ whose jump measure from a state $\bar{n}_x\delta_x$ is $b(x)\bar{n}_x P(y;x)m(x,dy)$. ■

Example 5.2. Let us consider a simple model with trait $x \in [0,4]$, C being constant, and $b(x) = 4 - x$, $d \equiv 1$, $\tau(x,y) = \tau e^{x-y}$. Then $S(x+h;x) = -h + \tau(e^h - e^{-h})$ and, for $\tau > 1/2$, it is positive if and only if $h > 0$. Thus the evolution with transfer is directed towards larger and larger traits, decreasing the growth rate until possible extinction. For τ small enough, $S(x+h;x) < 0$ for $h > 0$ so that a mutant of trait $x + h$ with $h > 0$ would disappear at the TSS scale. In this case, evolution drives the population to smaller and smaller traits until trait 0. The evolution for intermediary τ's is an open challenging question.

5.2. Canonical equation of the adaptive dynamics

Let us now assume that the mutation effects are very small: the mutation distribution m_σ depends on a parameter $\sigma > 0$ as follows:

$$\int g(z)m_\sigma(x,dz) = \int g(x+\sigma h)m_1(x,dh),$$

where m_1 is a reference symmetric measure with finite variance. Then the generator of the TSS Y^σ (which now depends on the parameter σ) is given by

$$L^\sigma g(x) = \int \big(g(x+\sigma h) - g(x)\big)b(x)\bar{n}_x \frac{[S(x+\sigma h;x)]_+}{b(x+\sigma h) + \tau(x+\sigma h, x)\bar{n}_x} m_1(x,dh).$$

For smooth S and since $S(x;x) = 0$, we have when σ tends to 0,

$$L^\sigma g(x) \sim \sigma^2 \frac{1}{2} g'(x)\bar{n}_x \partial_1 S(x;x) \int h^2 m_1(x,dh).$$

Let us observe that $\sigma \to 0$ makes the dynamics stop at this time scale. To observe a nontrivial behavior, we have to wait a longer time of order of magnitude $1/\sigma^2$.

Standard tightness and identification arguments allow showing the convergence in probability in $\mathbb{D}([0,T],\mathcal{X})$ of the process $(Y^\sigma_{t/\sigma^2}, t \in [0,T])$ to the deterministic function $(x(t), t \in [0,T])$, solving the equation

$$x'(t) = \frac{1}{2}\bar{n}_{x(t)}\partial_1 S\big(x(t);x(t)\big) \int h^2 m_1\big(x(t),dh\big), \tag{5.3}$$

the so-called *canonical equation of adaptive dynamics* introduced in [20] (cf. [15] for a rigorous proof). Note also that there is another candidate for the canonical equation obtained from partial differential equation arguments related to Hamilton–Jacobi equations [22,31,37].

Let us come back to Example 5.2 introduced previously. We assume that $m_1(x,dh)$ is a symmetric measure keeping the trait in [0, 4], i.e., with support in $[-x, 4-x]$. In this case, $\bar{n}_x = \frac{3-x}{C}$ and the canonical equation is given by

$$x'(t) = \frac{3 - x(t)}{C}(2\tau - 1)\int h^2 m_1\big(x(t),dh\big),$$

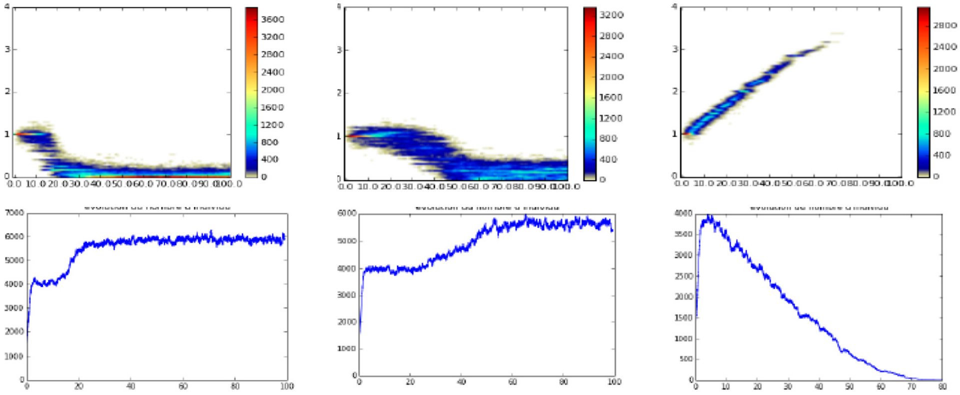
since $r'(x) = -1$ and $\partial_1\tau(x,x) = -\partial_2\tau(x,x) = \tau$. Then for $\tau > 1/2$, the trait support is an increasing function, the population size $\bar{n}_{x(t)}$ is decreasing to 0, and therefore evolution drives the population to an evolutionary suicide. Conversely, for $\tau < 1/2$, evolution leads to the optimal null trait (which maximizes the growth rate).

6. SIMULATIONS – CASE OF FREQUENCY-DEPENDENCE

(Simulations due to the Master students Lucie Desfontaines and Stéphane Krystal)

We focus on the special case of unilateral transfer, that is, $\tau(x,y) = \tau \mathbf{1}_{x>y}$, which is relevant for plasmids transfer. The next simulations are concerned with Example 5.2, with $C \equiv 0.5$, $p = 0.03$, and $m_\sigma(x,h)dh = \mathcal{N}(0,\sigma^2)$, conditioned on $x + h \in [0,4]$, with $\sigma = 0.1$. The initial state is composed of $K = 1\,000$ individuals with trait 1. Thus the corresponding population size at equilibrium is $1\,000 \times \frac{b(1)-d(1)}{C} = 4\,000$ individuals.

The constant τ will be the varying parameter. Figure 2(a) shows the evolution dynamics when $\tau = 0$. The evolution drives the population to its optimal trait 0 corresponding to a size at equilibrium equal to $1\,000 \times \frac{b(0)-d(0)}{C} = 6\,000$ individuals. The case $\tau = 0.2$ in Figure 2(b) shows a scenario similar to the case $\tau = 0$, although the evolution to optimal trait 0 takes a longer time. Conversely, when $\tau = 1$ (Figure 2(c)), the transfer drives the traits to larger and larger values, corresponding to lower and lower population sizes until extinction (evolutionary suicide). These simulations correspond to the theoretical study of the previous section. Let us now consider the intermediary value $\tau = 0.7$ (Figure 3). The evolution exhibits different patterns. In the first picture, high transfer converts at first individuals to larger traits and at the same time the population decreases. At some point, the population size is so small that the transfer does not play a role anymore leading to the brutal resurgence of a quasiinvisible strain, issued from a few individuals with small traits (and then with larger growth rate). We observe cyclic resurgences driving the mean trait towards



(a) (b) (c)

FIGURE 2

(a) $\tau = 0$; (b) $\tau = 0.2$ – almost no modification; (c) $\tau = 1$ – evolutionary suicide. Time in abscissa. First line, trait evolution; second line, size evolution.

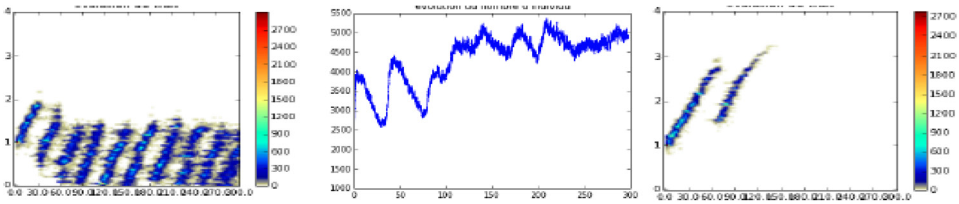


FIGURE 3
$\tau = 0.7$ – stepwise evolution with the trait evolution (left), and population size (center). Another pattern with extinction (right).

the optimal trait 0. In the last picture, we observe extinction of the population: the remaining individuals with smaller traits allow for a single resurgence of a new strain, but the traits of the individuals alive are too large to allow for survival.

7. STOCHASTIC ANALYSIS OF EMERGENCE OF EVOLUTIONARY CYCLIC BEHAVIOR – A SIMPLE MODEL

From now on, we are interested in the mathematical understanding of the previous simulations. In the latter, the chosen mutation probability p was small, but not the population mutation rate Kp, so (5.1) was not satisfied. We have to consider different time and size scales than the previous ones to capture the surprising resurgence behaviors. This part is largely inspired from Champagnat–Méléard–Tran [17].

7.1. A trait-discretized model

From now on, we consider a model inspired by Example 5.2 with a discrete trait space of mesh $\delta > 0$: $\mathcal{X} = [0, 4] \cap \delta\mathbb{N} = \{0, \delta, \dots, L\delta\}$ where $L = \lfloor 4/\delta \rfloor$. We choose $b(x) = 4 - x$, $\tau(x, y) = \tau \mathbf{1}_{x>y}$, $d(\cdot) \equiv 1$ and $C(\cdot, \cdot) \equiv C$. Therefore, $\bar{n}_x = \frac{3-x}{C}$ and the invasion fitness of a mutant individual of trait y in the population of resident trait x and size $K\bar{n}_x$ is

$$S(y; x) = x - y + \tau \mathbf{1}_{x<y} - \tau \mathbf{1}_{x>y} = x - y + \tau\,\mathrm{sign}(y - x). \tag{7.1}$$

We also define the fitness of an individual of trait y in a negligible population (of size $o(K)$) with dominant trait x to be

$$\hat{S}(y; x) = 3 - y + \tau\,\mathrm{sign}(y - x). \tag{7.2}$$

Indeed, the competition part is negligible in that case and vanishes at the limit when $K \to \infty$.

We assume that

$$p_K = K^{-\alpha} \quad \text{with } \alpha \in (0, 1), \tag{7.3}$$

and when a mutation occurs from an individual with trait $\ell\delta$, the new offspring carries the mutant trait $(\ell + 1)\delta$ (the mutations are directed to the right). The total mutation rate in a population with size of order K is thus equal to $K^{1-\alpha}$ and then goes to infinity with K. We are very far from the situation described in [6, 11, 15] where (5.1) was satisfied. Here, small populations of size order K^{β}, $\beta < 1$ can have a nonnegligible contribution to evolution by mutational events, and we need to take into account all subpopulations with size of order K^{β}.

The population is described by the vector $(N_0^K(t), \dots, N_\ell^K(t), \dots, N_L^K(t))$, where $N_\ell^K(t)$ is the number of individuals of trait $x = \ell\delta$ at time t. The total population size N_t^K is now $N_t^K = \sum_{\ell=0}^{L} N_\ell^K(t)$. Our study of the (evolutionary) long-time dynamics of the process is based on a fine analysis of the size order, as power of K, of each subpopulation. These powers of K evolve on the timescale $\log K$, as can be easily seen in the case of branching processes (see Lemma 7.1). We thus define $\beta_\ell^K(t)$ for $0 \le \ell \le L$ such that

$$N_\ell^K(t \log K) = K^{\beta_\ell^K(t)} - 1, \quad \text{i.e., } \beta_\ell^K(t) = \frac{\log(1 + N_\ell^K(t \log K))}{\log K}. \tag{7.4}$$

We assume that $N^K(0) = (\lfloor \frac{3K}{C} \rfloor, \lfloor K^{1-\alpha} \rfloor, \dots, \lfloor K^{1-\ell\alpha} \rfloor, \dots, \lfloor K^{1-\lfloor 1/\alpha \rfloor \alpha} \rfloor, 0, \dots, 0)$. Then trait $x = 0$ is initially resident, with density $3/C$. With this initial condition, we have

$$\beta_\ell^K(0) \xrightarrow[K \to +\infty]{} (1 - \ell\alpha)\mathbf{1}_{0 \le \ell < \frac{1}{\alpha}}. \tag{7.5}$$

The main result of this section will give the asymptotic dynamics of $\beta^K(t) = (\beta_0^K(t), \dots, \beta_L^K(t))$ for $t \ge 0$ when $K \to +\infty$. We show that the limit is a piecewise affine continuous function, which can be described along successive phases determined by their resident or dominant traits. When the latter trait changes, the fitnesses governing the slopes are modified. Moreover, inside each phase, other changes of slopes are possible due to a delicate balance between mutations, transfer, and growth of subpopulations. We will deduce from the asymptotic dynamics of $\beta^K(t)$ explicit criteria for some of the evolutionary outcomes observed in Section 6 (Theorem 7.5).

Such an approach based on the behavior of the exponents β_K at the time scale $\log K$ has also been used in Durrett–Mayberry [23] for constant population size or pure birth process, with directional mutations and increasing fitness parameter, in Bovier et al. [10] for a density-dependent model where the evolution crosses the fitness valley constituted of unfit traits, in Blath et al. [8] for models with dormancy. In a deterministic setting with similar scales, we also refer to Kraut–Bovier [29]. In our case, the dynamics is far more complex due to the trade-off between larger birth rates for small trait values and transfer to higher traits, leading to diverse evolutionary outcomes. As a consequence, we need to consider cases where the dynamics of a given trait is completely driven by immigrations (see Lemma 7.2). This complexifies a lot the analysis.

7.2. Some enlightening lemmas

Before stating the main result (Theorem 7.3) which can be difficult to read and understand, we state two lemmas whose proof can be found in the Appendix of [17]. These lemmas are interesting by themselves.

(i) Assume first that a mutant with trait y appears in a resident population with trait x such that $y < x$. Then the dynamics of the initial (small) y-subpopulation size behaves as a linear birth and death process with birth rate approximated by $4-y$ and death rate by $1+\frac{CN^{x,K}(t)}{K}+\tau$. We are thus led to study the following process.

Lemma 7.1. *Let us consider a linear birth and death process* $(Z_t^K, t\geq 0)$, *i.e., a binary branching process, with individual birth rate* $b\geq 0$, *individual death rate* $d\geq 0$, *and initial value* $Z_0^K=K^\beta$ *with* $\beta>0$.

The process $(\frac{\log(1+Z^K_{s\log K})}{\log K}, s\in[0,T])$ *converges in probability in* $L^\infty([0,T])$ *for all* $T>0$ *to* $((\beta+rs)\vee 0, s\in[0,T])$ *when* K *tends to infinity, with* $r=b-d$.

In addition, if $b<d$, *for all* $s>\beta/r$, *then* $\lim_{K\to+\infty}\mathbb{P}(Z^K_{s\log K}=0)=1$.

The limit can be understood from $\mathbb{E}(Z^K(t))=K^\beta e^{rt}$. The proof of Lemma 7.1 uses the martingale property of $(e^{-rt}Z_t^K)_{t\geq 0}$. The proof is easy for $r\geq 0$ and more technical in the case $r<0$, necessitating to control the extinction events after a certain time.

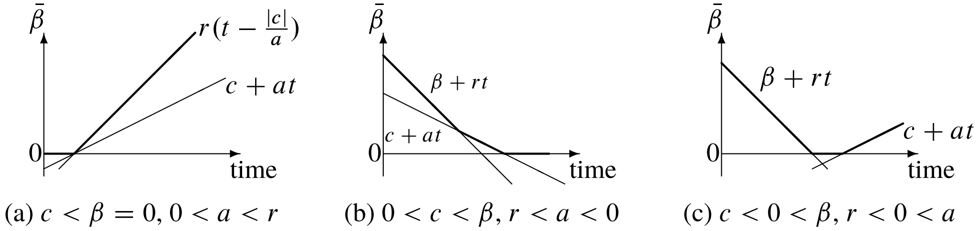
(ii) Assume now that a mutant with trait $y=x+\delta$ appears in a resident population with trait x. Then the dynamics of the initial (small) y-subpopulation size behaves as a linear birth and death process with birth rate approximated by $4-y+\tau$ and death rate by $1+\frac{CN^{x,K}(t)}{K}$. But in addition, trait y may receive a contribution from x at time t due to mutations at total rate $N^{x,K}(t)K^{-\alpha}$. By Lemma 7.1, we know that $N^{x,K}(s\log K)\approx K^{c+as}$ for constant $a,c\in\mathbb{R}$. This justifies the following lemma.

Lemma 7.2. *Let us consider a linear birth and death process with immigration* $(Z_t^K, t\geq 0)$, *with individual birth rate* $b\geq 0$, *individual death rate* $d\geq 0$, *initial value* $Z_0^K=K^\beta$ *with* $\beta>0$, *and immigration rate at time* t *given by* $K^c e^{at}$, *with* $a,c\in\mathbb{R}$.

The process $(\frac{\log(1+Z^K_{s\log K})}{\log K}, s\in[0,T])$ *converges when* K *tends to infinity in probability in* $L^\infty([0,T])$ *for all* $T>0$ *to a continuous deterministic function* $\bar\beta(s)$.

For $c\leq\beta$ *and* $\beta>0$, $\bar\beta(s)=(\beta+rs)\vee(c+as)\vee 0$. *For* $\beta=0$, $c<0$ *and* $a>0$, $\bar\beta(s)=((r\vee a)(s-|c|/a))\vee 0$. *For* $\beta=0$, $c<0$, *and* $a\leq 0$, $\bar\beta(s)=0$. *The other cases are immediate (see* [17]*).*

This convergence is illustrated in Figure 4.



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

FIGURE 4

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

7.3. Dynamics of the exponents

Let us come back to the asymptotic dynamics of $\beta^K(t) = (\beta_0^K(t), \dots, \beta_L^K(t))$ for $t \geq 0$ when $K \to +\infty$, which are characterized in the next result by a succession of deterministic time intervals $[s_{k-1}, s_k]$, $k \geq 1$, called phases and delimited by changes of resident or dominant traits. The latter are unique except at times s_k and are denoted by $\ell_k^* \delta$, $k \geq 1$. This asymptotic result holds until a time T_0, which guarantees that there is ambiguity neither on these traits nor on the extinct subpopulations at the phase transitions. We will not give the exact (and technical) definition of T_0 and refer to [17].

Theorem 7.3. *Assume (7.3) with $\alpha \in (0,1)$, $\delta \in (0,4)$, and (7.5).*

(i) *For $0 < T \leq T_0$, the sequence $(\beta^K(t), t \in [0,T])$ converges in probability in $\mathbb{D}([0,T],[0,1]^{L+1})$ to a deterministic piecewise affine continuous function $(\beta(t) = (\beta_0(t), \dots, \beta_L(t)), t \in [0,T])$, such that $\beta_\ell(0) = (1 - \ell\alpha)\mathbf{1}_{0 \leq \ell < \frac{1}{\alpha}}$. The functions β are parameterized by α, δ, and τ defined as follows.*

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

$$s_k = \inf\bigl\{t > s_{k-1} : \exists \ell \neq \ell_k^*,\ \beta_\ell(t) = \beta_{\ell_k^*}(t)\bigr\}. \tag{7.6}$$

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

$$\ell_{k+1}^* = \arg\max_{\ell \neq \ell_k^*} \beta_\ell(s_k), \tag{7.7}$$

if the argmax is unique. In the other cases, we stop the induction.

(iii) *The functions β_ℓ are defined, for all $t \in [s_{k-1}, s_k]$ and $\ell \in \{0, \dots, L\}$, by*

$$\beta_\ell(t) = \begin{cases} [\mathbb{1}_{\beta_0(s_{k-1})>0}(\beta_0(s_{k-1}) + \int_{s_{k-1}}^t \tilde{S}_{s,k}(0; \ell_k^* \delta)\, ds)] \vee 0, & \text{if } \ell = 0, \\ (\beta_\ell(s_{k-1}) + \int_{t_{\ell-1,k} \wedge t}^t \tilde{S}_{s,k}(\ell\delta; \ell_k^* \delta)\, ds) & \\ \qquad \vee (\beta_{\ell-1}(t) - \alpha) \vee 0, & \text{otherwise,} \end{cases} \tag{7.8}$$

where, for all traits x, y, $\tilde{S}_{t,k}(y;x) = \mathbb{1}_{\beta_{\ell_k^*}(t)=1} S(y;x) + \mathbb{1}_{\beta_{\ell_k^*}(t)<1} \hat{S}(y;x)$ and where

$$t_{\ell-1,k} = \begin{cases} \inf\{t \geq s_{k-1},\ \beta_{\ell-1}(t) = \alpha\}, & \text{if } \beta_\ell(s_{k-1}) = 0, \\ s_{k-1}, & \text{otherwise.} \end{cases} \tag{7.9}$$

In addition, for all ℓ and all $a < b$ such that the time interval $[a,b]$ is included in the interior of the zero-set of β_ℓ, the event $\{N_\ell^K(t \log K) = 0, \forall t \in [a,b]\}$ has a probability converging to 1 as K tends to infinity.

Simulations are shown in Figure 5 for various parameter values.

Roughly speaking, slope changes of the exponents $((\beta_0(t), \dots, \beta_L(t)), t \in [0,T])$ can take place at the times when a new exponent reaches 1 and there is a change of the resident

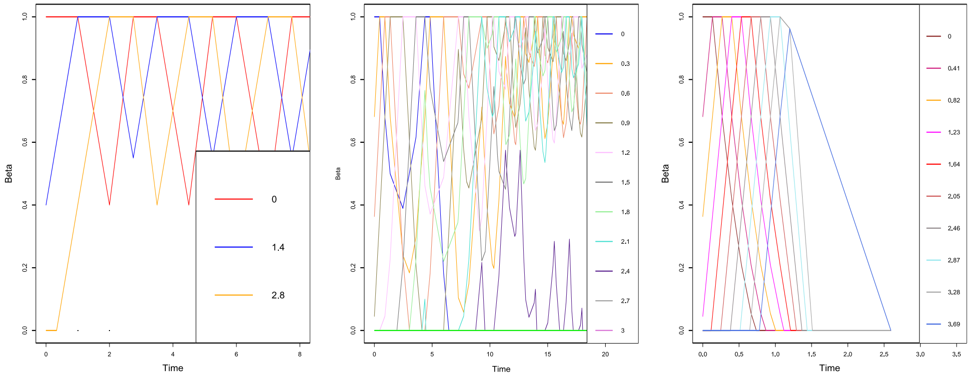


FIGURE 5

Exponents $\beta_\ell(t)$ as functions of time: (left) $\delta = 1.4$, $\alpha = 0.6$, $\tau = 2$. We see a periodic behavior showing reemergence of the fittest traits; (center) $\delta = 0.3$, $\alpha = 1/\pi$, $\tau = 1$. A cyclic but aperiodic behavior is observed; (right) $\delta = 0.41$, $\alpha = 1/\pi$, $\tau = 2.8$. The population is directly driven to evolutionary suicide.

trait, when a new exponent reaches 0 and there is extinction of the trait, and when the slope of an exponent formerly directed by its fitness becomes directed by incoming mutations.

Remark 7.4. (i) By the definition of s_k and ℓ^*_{k+1}, $\max_\ell \beta_\ell(t) = \beta_{\ell^*_k}(t)$ for $t \in [s_{k-1}, s_k)$.

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

Main ideas of the proof. We need to consider in the sequel two different situations: either there is a single trait x with population size of order K, called *resident* trait, or the total population size is $o(K)$. We explain the proof for simplicity assuming that there is always a resident trait. Theorem 7.3 is obtained by a fine comparison of the size of each subpopulation defined by a given trait value with carefully chosen branching processes with immigration. The stochastic dynamics consists in a succession of steps, composed of long phases $[\sigma_k^K \log K, \theta_k^K \log K]$ for $k \geq 1$ (with $\sigma_1^K = 0$) followed by short intermediate phases $[\theta_k^K \log K, \sigma_{k+1}^K \log K]$, where the stopping time θ_k^K is defined as the first time when the resident population size exits a neighborhood of its equilibrium density, or when the other subpopulations stop to be negligible with respect to the resident population. In each long phase, there is a single resident trait. Short intermediate phases correspond to the replacement of the resident trait, where two subpopulations are of maximal order. We prove that θ_k^K converges in probability to s_k, $k \geq 1$. In the limit, intermediate steps vanish on the time scale $\log K$. The proof proceeds by induction on k until some step k_0 where one of the three following events occurs: the exponents of three traits become maximal simultaneously, extinction, or the exponent of some trait vanishes at the same time as a change of resident population.

We then stop the induction and set $T_0 = s_{k_0}$ in the first and third cases, or $T_0 = +\infty$ in the second case.

To control the exponents $\beta_\ell^K(t)$, we proceed by a double induction, first on the steps, and second, inside each step, on the traits $\ell\delta$, for $\ell = 0$ to $\ell = L$. The exponents are approximately piecewise affine. Changes of slopes may happen when a new trait emerges, when a trait dies or when the dynamics of a trait becomes driven by incoming mutations. We use Lemma 7.2. During intermediate phases, we use comparisons with dynamical systems, described in Section 3. ■

7.4. Reemergence of trait 0

Recall that we work with birth, death and transfer rates presented in Section 7.1. In Figure 5, we have exhibited different evolutionary dynamics (reemergence of a trait, cyclic behavior, local extinction, evolutionary suicide). By reemergence of a trait $\ell\delta$, we mean that $\beta_\ell(s) = 1$ on some nonempty time interval $[t_1, t_2]$, then $\beta_\ell(s) < 1$ on some nonempty interval (t_2, t_3), and then $\beta_\ell(s) = 1$ again on some nonempty interval $[t_3, t_4]$. We would like to predict the evolutionary outcome as a function of parameters α, δ, τ. There are so many situations that we are not able to fully characterize the outcomes (see [17] for a detailed study in the case of three traits). Therefore, we focus on the beginning of the dynamics until either global extinction or reemergence of one trait occurs.

The resurgence of trait 0 is a prerequisite for a cyclic dynamics as those observed in Figure 5. We assume here that $\delta < 4/3$ (so that the cardinal of $\mathcal{X}$ is $L + 1 \geq 4$) and only consider the case $\delta < \tau < 3$. Computing the fitness functions, one can observe that for the first phases, $s_k = \frac{k\alpha}{\tau-\delta}$, and the trait $k\delta$ is resident on $[s_k, s_{k+1})$ ($\beta_k(s) = 1$) and for all $s \in [s_k, s_{k+1})$,

$$\beta_0(s) = 1 - \frac{\alpha(k-1)}{\tau-\delta}\left(\tau - \frac{k}{2}\delta\right) - (\tau - k\delta)(s - s_k).$$

This formula stays valid until either $\beta_0(s) = 0$ (loss of 0), or $\beta_0(s) = 1$ for some $s > s_1$ (reemergence of 0), or when the population size becomes $o(K)$. The slope of the function $\beta_0(s)$ becomes positive at time $s_{\tilde{k}}$, where $\tilde{k} := \lceil \frac{\tau}{\delta} \rceil$. Hence its minimal value is equal to

$$m_0 = \beta_0(s_{\tilde{k}}) = 1 - \frac{\alpha(\tilde{k}-1)}{\tau-\delta}\left(\tau - \frac{\tilde{k}}{2}\delta\right). \tag{7.10}$$

If the latter is positive, β_0 reaches 1 again in phase $[s_{\bar{k}}, s_{\bar{k}+1})$, where $\bar{k} = \lfloor 2\frac{\tau}{\delta} \rfloor$, at time

$$\bar{s} := s_{\bar{k}} + \frac{\alpha(\bar{k}-1)}{\tau-\delta}\,\frac{\tau - \frac{\bar{k}}{2}\delta}{\bar{k}\delta - \tau} = s_{\lfloor 2\frac{\tau}{\delta}\rfloor} + \frac{\alpha(\lfloor 2\frac{\tau}{\delta}\rfloor - 1)}{\tau-\delta}\,\frac{\tau - \frac{\lfloor 2\frac{\tau}{\delta}\rfloor}{2}\delta}{\lfloor 2\frac{\tau}{\delta}\rfloor\delta - \tau}. \tag{7.11}$$

The previous calculations give the intuition for the following theorem (see the proof in [17]).

Theorem 7.5. *Assuming $\delta < \tau < 3$, $\delta < 4/3$ and, under the assumptions of Theorem 7.3,*

(a) *If $m_0 > 0$ and $\bar{k}\delta < 3$, then the first reemerging trait is 0 and the maximal exponent is always 1 until this reemergence time;*

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

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

Biologically, case (b) corresponds to evolutionary suicide. In cases (a) and (c), very few individuals with small traits remain, which are able to reinitiate a population of size of order K (reemergence) after the resident trait becomes too large. In these cases, one can expect successive reemergences. However, we do not know if there exists a limit cycle for the dynamics. Case (c) means that the total population is $o(K)$ on some time interval, before reemergence occurs after populations with too large traits become small enough.

It seems very difficult to go further with probabilistic tools. Another approach could consist in obtaining a macroscopic approximation of the exponents β^K in a trait continuum in terms of Hamilton–Jacobi equations and then using the tools of analysis.

8. MACROSCOPIC HAMILTON–JACOBI APPROXIMATION OF THE EXPONENTS

This part is a collaboration in progress with S. Mirrahimi [16]. We will give the ideas of our ongoing results, in particular a partial result concerning the simple case of stochastic supercritical birth–death–mutation process without transfer and competition. We assume that trait x belongs to the continuum $[0,1]$. Starting from a finite population, our goal is to recover, by a direct scaling, the Hamilton–Jacobi equation that has been introduced in [3,22]. For this, we consider a discretization of the trait space $[0,1]$ with step $\delta_K \to 0$, scale the mutation steps by a factor $1/\log K$ (small mutation steps), and assume that the initial population sizes are of the order of K^{β_0} for an exponent β_0 that can depend on the trait. More precisely, the population is composed of individuals with traits belonging to the discrete space $\mathcal{X}_K := \{i\delta_K : i \in \{0, 1, \dots, [\frac{1}{\delta_K}]\}\}$. The number of individuals with trait $i\delta_K$ is described by the stochastic process $(N_i^K(t), t \ge 0)$. As in the previous sections, an individual with trait $x \in \mathcal{X}_K$ gives birth to a new individual with same trait x at rate $b(x)$, dies at rate $d(x)$, but we assume that, for all $y \in \mathcal{X}_K$, it gives birth to a mutant individual with trait y at rate

$$p(x)\delta_K \log K m\big(\log K(x-y)\big).$$

Assumption 8.1. (i) The functions b, d, and p are nonnegative C^1-functions defined on $[0,1]$ such that, for all $x \in [0,1]$, $b(x) > d(x)$.

(ii) The function m is nonnegative, continuous, defined on $\mathbb{R}$, satisfies $\int_{\mathbb{R}} m(y)\,dy = 1$. It has exponential moments of any order and behaves as the Gaussian kernel $m(h) = \frac{1}{\sqrt{2\pi}\sigma} e^{-h^2/2\sigma^2}$ at infinity.

(iii) There exists $a > 0$ such that, for all $K \in \mathbb{N}$ and all $i \in \{0, 1, \dots, [\frac{1}{\delta_K}]\}$, $N_i^K(0) \ge K^a$.

(iv) There exists $a_2 < a$ such that $K^{-a_2/4} << \delta_K << 1/\log(K)$. Then, for $h_K := \delta_K \log K$, we have $\lim_{K\to+\infty} h_K = 0$.

Note that points 1 and 3 of Assumption 8.1 impede the subpopulations to be extinct. Note also that, for all $x \in (0,1)$, the total mutation rate from an individual with trait $x_K = i_K \delta_K$ with $i_K = [x/\delta_K]$, converges as $K \to +\infty$ to

$$\lim_{K\to+\infty} p(x_K) \sum_{j=0}^{[\frac{1}{\delta_K}]} h_K m\big(h_K(i_k - j)\big) = p(x) \int_{\mathbb{R}} m(y)\,dy = p(x).$$

Defining the exponents $\beta_i^K(t)$ as in (7.4), we introduce their interpolations: for all $x \in [0,1]$ and $K \geq 1$, let i be such that $x \in [i\delta_K, (i+1)\delta_K)$ and define

$$\tilde{\beta}^K(t,x) = \beta_i^K(t)\left(1 - \frac{x}{\delta_K} + i\right) + \beta_{i+1}^K(t)\left(\frac{x}{\delta_K} - i\right).$$

The sequence of processes $(\tilde{\beta}^K)_{K\geq 1}$ belongs to $\mathbb{D}([0,T], \mathcal{C}([0,1],\mathbb{R}))$, where $\mathcal{C}([0,1],\mathbb{R})$ is endowed with the topology of uniform convergence.

Theorem 8.2. *We assume that Assumptions* 8.1 *hold, and that the sequence* $(\tilde{\beta}^K(0,\cdot))$ *converges in probability on* $\mathcal{C}([0,1],\mathbb{R})$ *to a deterministic function* $\beta_0(\cdot)$ *and that there exists a constant* A *such that*

$$\lim_{K\to+\infty} \mathbb{P}\big(L_0^K > A\big) = 0, \quad \textit{where } L_0^K := \sup_{i\neq j} \frac{|\beta_i^K(0) - \beta_j^K(0)|}{\delta_K |i-j|}.$$

Then $\tilde{\beta}^K$ *converges in probability in* $\mathbb{D}([0,T], \mathcal{C}([0,1],\mathbb{R}))$ *to the unique viscosity solution* β *of the Hamilton–Jacobi equation with state constraint boundary conditions*

$$\begin{cases} \dfrac{\partial}{\partial t}\beta(t,x) = b(x) - d(x) + p(x) \displaystyle\int_{\mathbb{R}} m(h) e^{h\partial_x \beta(t,x)} dh, & (t,x) \in \mathbb{R}_+ \times (0,1), \\ \beta(0,x) = \beta_0(x), & x \in [0,1]. \end{cases} \tag{8.1}$$

More precisely, β *is a viscosity supersolution of* (8.1) *in* $(0,+\infty) \times (0,1)$ *and a viscosity subsolution in* $(0,+\infty) \times [0,1]$.

Usually, the analytical proof of such concentration results is based on the maximum principle (see [3]) which does not hold in this stochastic framework. To prove the tightness of the sequence $\tilde{\beta}^K$, a technical and delicate point consists in showing that the increments $(\beta_{i+1}^K(t) - \beta_i^K(t))/\delta_K$ are bounded uniformly in time for K large enough. These increments are semimartingales, and we easily obtain their Doob–Meyer decomposition. The martingale part is proved to be small for large K. The maximum principle is used to control the finite variation part, with an ω-by-ω argument. Once the tightness is obtained, we have to identify the limiting values of $\tilde{\beta}^K$, which only charge deterministic and continuous trajectories. We identify the limiting paths as viscosity solutions of the Hamilton–Jacobi equation (8.1).

FUNDING

This work was partially supported by the Chair “Modélisation Mathématique et Biodiversité” of Veolia Environnement-Ecole Polytechnique-Museum National d’Histoire Naturelle-Fondation X and by Labex Bézout (ANR-10-LABX-58).

REFERENCES

[1] R. Anderson and R. May, Population biology of infectious diseases: Part I. *Nature* **280** (1979), 361–367.

[2] V. Bansaye and S. Méléard, *Stochastic models for structured populations. Scaling limits and long time behavior*. MBI Lecture Ser. 1.4, Springer, 2015.

[3] G. Barles, S. Mirrahimi, and B. Perthame, Concentration in Lotka–Volterra parabolic equations: a general convergence result. *Methods Appl. Anal.* **16** (2009), 321–340.

[4] F. Baumdicker and P. Pfaffelhuber, The infinitely many genes model with horizontal gene transfer. *Electron. J. Probab.* **19** (2014), 1–27.

[5] S. Billiard, P. Collet, R. Ferrière, S. Méléard, and V. Tran, The effect of competition and horizontal trait inheritance on invasion, fixation and polymorphism. *J. Theoret. Biol.* **411** (2016), 48–58.

[6] S. Billiard, P. Collet, R. Ferrière, S. Méléard, and V. Tran, Stochastic dynamics for adaptation and evolution of microorganisms. In *Proceedings of 7th European Congress of Mathematics*, edited by V. Mehrmann and M. Skutella, pp. 527–552, European Mathematical Society, 2018.

[7] P. Billingsley, *Convergence of probability measures*. John Wiley & Sons, New York, 1968.

[8] J. Blath, T. Paul, and A. Tobias, A Stochastic Adaptive Dynamics Model for Bacterial Populations with Mutation, Dormancy and Transfer. 2021, arXiv:2105.09228.

[9] B. Bolker and S. Pacala, Using moment equations to understand stochastically driven spatial pattern formation in ecological systems. *Theor. Popul. Biol.* **52** (1997), 179–197.

[10] A. Bovier, L. Coquille, and C. Smadi, Crossing a fitness valley as a metastable transition in a stochastic population model. *Ann. Appl. Probab.* **29** (2019), no. 6, 3541–3589.

[11] N. Champagnat, A microscopic interpretation for adaptive dynamics trait substitution sequence models. *Stochastic Process. Appl.* **116** (2006), 1127–1160.

[12] N. Champagnat, R. Ferrière, and S. Méléard, Individual-based probabilistic models of adaptive evolution and various scaling approximations. In *Proceedings of the 5th seminar on stochastic analysis, random fields and applications*, Probab. Prog. Ser., Birkhäuser, Ascona, Suisse, 2006.

[13] N. Champagnat, R. Ferrière, and S. Méléard, Unifying evolutionary dynamics: from individual stochastic processes to macroscopic models via timescale separation. *Theor. Popul. Biol.* **69** (2006), 297–321.

[14] N. Champagnat, P.-E. Jabin, and S. Méléard, Adaptive dynamics in a stochastic multi-resources chemostat model. *J. Math. Pures Appl.* **101** (2014), no. 6, 755–788.

[15] N. Champagnat and S. Méléard, Polymorphic evolution sequence and evolutionary branching. *Probab. Theory Related Fields* **151** (2011), no. 1–2, 45–94.

[16] N. Champagnat, S. Méléard, S. Mirrahimi, and V. C. Tran, Filling the gap between individual-based evolutionary models and Hamilton–Jacobi equations. 2022, in preparation.

[17] N. Champagnat, S. Méléard, and V. C. Tran, Stochastic analysis of emergence of evolutionary cyclic behaviour in population dynamics with transfer. *Ann. Appl. Probab.* **31** (2021), no. 4, 1820–1867.

[18] C. Darwin, *On the origin of species: A facsimile of the first edition*. Harvard University Press, 1964.

[19] L. Desvillettes, P. E. Jabin, S. Mischler, and G. Raoul, On selection dynamics for continuous structured populations. *Commun. Math. Sci.* **6** (2008), no. 3, 729–747.

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

[21] O. Diekmann, A beginner's guide to adaptive dynamics. *Banach Center Publ.* **63** (2003), 47–86.

[22] O. Diekmann, P.-E. Jabin, S. Mischler, and B. Perthame, The dynamics of adaptation: an illuminating example and a Hamilton–Jacobi approach. *Theor. Popul. Biol.* **67** (2005), 257–271.

[23] R. Durrett and J. Mayberry, Travelling waves of selective sweeps. *Ann. Appl. Probab.* **21** (2011), no. 2, 699–744.

[24] M. B. Elowitz, A. J. Levine, E. D. Siggia, and P. S. Swain, Stochastic Gene Expression in a Single Cell. *Science* **297** (2002), no. 5584, 1183–1186.

[25] N. Fournier and S. Méléard, A microscopic probabilistic description of a locally regulated population and macroscopic approximations. *Ann. Appl. Probab.* **14** (2004), no. 4, 1880–1919.

[26] M. I. Freidlin and A. Ventzell, *Random Perturbations of Dynamical Systems*. Springer, Berlin, 1984.

[27] P. Hinow, F. Le Foll, P. Magal, and G. Webb, Analysis of a model for transfer phenomena in biological populations. *SIAM J. Appl. Math.* **70** (2009), 40–62.

[28] J. Hofbauer and R. Sigmund, Adaptive dynamics and evolutionary stability. *Appl. Math. Lett.* **3** (1990), 75–79.

[29] A. Kraut and A. Bovier, From adaptive dynamics to adaptive walks. *J. Math. Biol.* **79** (2019), 1699–1747.

[30] B. Levin, F. Stewart, and V. Rice, Kinetics of conjugative plasmid transmission: fit of a simple mass action model. *Plasmid* **2** (1979), 247–260.

[31] A. Lorz, S. Mirrahimi, and B. Perthame, Dirac mass dynamics in multidimensional nonlocal parabolic equations. *Comm. Partial Differential Equations* **36** (2011), no. 6, 1071–1098.

[32] P. Magal and G. Raoul, Dynamics of a kinetic model describing protein exchanges in a cell population. 2015, arXiv:1511.02665.

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

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

[35] J. Metz, S. Geritz, G. Meszéna, F. Jacobs, and J. V. Heerwaarden, Adaptative dynamics, a geometrical study of the consequences of nearly faithful reproduction. In *Stochastic and Spatial Structures of Dynamical Systems*, edited by S. J. Van Strien and S. M. Verduyn Lunel, Konink. Nederl. Akad. Wetensch. Verh. Afd. Natuurk. Eerste Reeks 45, North-Holland, Amsterdam, pp. 183–231, 1996.

[36] J. Metz, R. Nisbet, and S. Geritz, How should we define 'fitness' for general ecological scenarios? *Trends Ecol. Evol.* **7** (1992), 198–202.

[37] S. Mirrahimi, B. Perthame, and J. Y. Wakano, Evolution of species trait through resource competition. *J. Math. Biol.* **64** (2011), no. 7, 1189–1223.

[38] A. Novozhilov, G. Karev, and E. Koonin, Mathematical modeling of evolution of horizontally transferred genes. *Mol. Biol. Evol.* **22** (2005), 1721–1732.

[39] A. Raj and A. van Oudenaarden, Nature, nurture, or chance: stochastic gene expression and its consequences. *Cell* **135** (2008), no. 2, 216–226.

[40] F. Stewart and B. Levin, The population biology of bacterial plasmids: A priori conditions for the existence of conjugationally transmitted factors. *Genetics* **87** (1977), 209–228.

[41] S. J. Tazzyman and S. Bonhoeffer, Fixation probability of mobile elements such as plasmids. *Theor. Popul. Biol.* **90** (2013), 49–55.

NICOLAS CHAMPAGNAT

Université de Lorraine, CNRS, Inria, IECL, F-54000 Nancy, France, nicolas.champagnat@inria.fr

SYLVIE MÉLÉARD

Institut Universitaire de France and Ecole polytechnique, CNRS, Institut polytechnique de Paris, route de Saclay, 91128 Palaiseau Cedex-France, sylvie.meleard@polytechnique.edu

VIET CHI TRAN

LAMA, Univ Gustave Eiffel, Univ Paris Est Creteil, CNRS, F-77454 Marne-la-Vallée, France, chi.tran@univ-eiffel.fr