Individual-based probabilistic models of adaptive evolution and various scaling approximations

Nicolas Champagnat[1,2], Régis Ferrière[1,3], Sylvie Méléard[1,2]

January 26, 2006

Abstract

We are interested in modelling Darwinian evolution, resulting from the interplay of phenotypic variation and natural selection through ecological interactions. Our models are rooted in the microscopic, stochastic description of a population of discrete individuals characterized by one or several adaptive traits. The population is modelled as a stochastic point process whose generator captures the probabilistic dynamics over continuous time of birth, mutation, and death, as influenced by each individual's trait values, and interactions between individuals. An offspring usually inherits the trait values of her progenitor, except when a mutation causes the offspring to take an instantaneous mutation step at birth to new trait values. We look for tractable large population approximations. By combining various scalings on population size, birth and death rates, mutation rate, mutation step, or time, a single microscopic model is shown to lead to contrasting macroscopic limits, of different nature: deterministic, in the form of ordinary, integro-, or partial differential equations, or probabilistic, like stochastic partial differential equations or superprocesses. In the limit of rare mutations, we show that a possible approximation is a jump process, justifying rigorously the so-called trait substitution sequence. We thus unify different points of view concerning mutation-selection evolutionary models.

Key-words: Darwinian evolution, birth-death-mutation-competition point process, mutation-selection dynamics, nonlinear integro-differential equations, nonlinear partial differential equations, nonlinear superprocesses, fitness, adaptive dynamics, trait substitution sequence.

1 Introduction

In this paper, we are interested in modelling the dynamics of populations as driven by the interplay of phenotypic variation and natural selection operating through ecological interactions, i.e. Darwinian evolution. The fundamental property of living systems is the propensity of each individual to create and to select the diversity. This feature requires to

---

[1]Laboratoire d'Ecologie, Equipe Eco-Evolution Mathématique, Ecole Normale Supérieure, 46 rue d'Ulm, 75230 Paris cedex 05, France

[2]Equipe MODAL'X Université Paris 10, 200 avenue de la République, 92001 Nanterre Cedex, France

[3]Department of Ecology and Evolutionary Biology, University of Arizona, Tucson AZ 85721, USA

focus on the stochastic dynamics of each individual in the population. The study of such evolutionary-ecological models is very complicated, and several approximations have been proposed. Firstly, Bolker and Pacala [2] and Dieckmann and Law [11] have introduced the moment equations of the distribution of traits in the population and studied different moment closure heuristics. Secondly, various nonlinear macroscopic models (integro-differential equations, partial differential equations, superprocesses) have been proposed without microscopic justification. Finally, the emerging field of adaptive dynamics have proposed a new class of macroscopic models on the evolutionary time scale, defined as jump processes and ordinary differential equations (trait substitution sequences, Metz et al. [22], canonical equation of adaptive dynamics, Dieckmann and Law [10]). In all these cases and from a biological point of view, the pathway from microscopic to macroscopic models deserves a firm mathematical pavement, at least to clarify the significance of the implicit biological assumptions underlying the choice of a particular model.

In this work, we unify several macroscopic approximations by recovering them from a single microscopic model. In particular, we point out the importance of large population assumptions and that the nature of the approximation strongly depends on the combination of various scalings of the biological parameters (birth and death rates, mutation rate, mutation step and time).

This paper starts (Section 2) with the microscopic description of a population of discrete individuals, whose phenotypes are described by a vector of trait values. The population is modelled as a stochastic Markov point process whose generator captures the probabilistic dynamics over continuous time of birth, mutation and death, as influenced by each individual's trait values and interactions between individuals. The adaptive nature of a trait implies that an offspring usually inherits the trait values of her progenitor, except when a mutation occurs. In this case, the offspring makes an instantaneous mutation step at birth to new trait values. We will refer to the state space parameterized by adaptive traits as the trait space, and will often (slightly abusively) call trait the actual trait value. This process is defined as the solution of a stochastic differential equation driven by point Poisson measures (Section 2.1). In Section 2.2, we give an algorithmic construction of the population point process and propose some simulations, for various parameters, of an asymmetrical example developed in Kisdi [18]. Next, we prove that the point population process is a measure-valued semimartingale and compute its characteristics (Section 2.3). Then we look for tractable approximations, following different mathematical paths. Our first approach (Section 3) aims at deriving deterministic equations to describe the moments of trajectories of the point process, i.e. the statistics of a large number of independent realizations of the process. We explain the difficult hierarchy between these equations coming from competition kernels and preventing, even in the simple mean-field case, decorrelations and tractable moment closure. The alternative approach involves renormalizations of the point process based on a large population limit. The measure-valued martingale properties of the renormalized point process allow us to show that, according to different scalings of birth, death and mutation rates, one obtains qualitatively different limiting partial differential equations and the appearance or not of some demographic stochasticity. We show in Section 4.1 that by itself, the large-population limit leads to a deterministic, nonlinear integro-differential equation. Then, in Section 4.2.1, we combine the large-population limit with an acceleration of birth (hence mutation) and death according to small mutation steps. That yields either a deterministic nonlinear reaction-diffusion model, or a

stochastic measure-valued process (depending on the acceleration rate of the birth-and-death process). If now this acceleration of birth and death is combined with a limit of rare mutations, the large-population limit yields a nonlinear integro-differential equation either deterministic or stochastic, depending here again on the speed of the scaling of the birth-and-death process, as described in Section 4.2.2.

In Section 5, we model a time scale separation between ecological events (fast births and deaths) and evolution (rare mutations), for an initially monomorphic population. The competition between individuals takes place on the short time scale. In a large population limit, this leads on the mutation time scale to a jump process over the trait space, where the population stays monomorphic at any time. Thereby we provide a rigorous justification to the notion of trait substitution sequence introduced by Metz et al. [21].

2 Population point process

Even if the evolution manifests itself as a global change in the state of a population, its basic mechanisms, mutation and selection, operate at the level of individuals. Consequently, we model the evolving population as a stochastic interacting individual system, where each individual is characterized by a vector of phenotypic trait values. The trait space $\mathcal{X}$ is assumed to be a closed subset of $\mathbb{R}^d$, for some $d \geq 1$.

We will denote by $M_F(\mathcal{X})$ the set of finite non-negative measures on $\mathcal{X}$. Let also $\mathcal{M}$ be the subset of $M_F(\mathcal{X})$ consisting of all finite point measures:

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

Here and below, δ_x denotes the Dirac mass at x. For any $m \in M_F(\mathcal{X})$, any measurable function f on $\mathcal{X}$, we set $\langle m, f \rangle = \int_{\mathcal{X}} f dm$.

We aim to study the stochastic process ν_t, taking its values in $\mathcal{M}$, and describing the distribution of individuals and traits at time t. We define

$$\nu_t = \sum_{i=1}^{I(t)} \delta_{X_t^i}, \tag{2.1}$$

$I(t) \in \mathbb{N}$ standing for the number of individuals alive at time t, and $X_t^1, ..., X_t^{I(t)}$ describing the individuals' traits (in $\mathcal{X}$).

For a population $\nu = \sum_{i=1}^{I} \delta_{x^i}$, and a trait $x \in \mathcal{X}$, we define the birth rate $b(x, V * \nu(x)) = b(x, \sum_{i=1}^{I} V(x - x^i))$ and the death rate $d(x, U * \nu(x)) = d(x, \sum_{i=1}^{I} U(x - x^i))$ of individuals with trait x; V and U denote the interaction kernels affecting respectively reproduction and mortality. Let $\mu(x)$ and $M(x, z)dz$ be respectively the probability that an offspring produced by an individual with trait x carries a mutated trait and the law of this mutant trait.

Thus, the population evolution can be roughly summarized as follows. The initial population is characterized by a (possibly random) counting measure $\nu_0 \in \mathcal{M}$ at time 0, and any individual with trait x at time t has two independent random exponentially distributed "clocks": a birth clock with parameter $b(x, V * \nu_t(x))$, and a death clock with

parameter $d(x, U * \nu_t(x))$. If the death clock of an individual rings, this individual dies and disappears. If the birth clock of an individual with trait x rings, this individual produces an offspring. With probability $1 - \mu(x)$ the offspring carries the same trait x; with probability $\mu(x)$ the trait is mutated. If a mutation occurs, the mutated offspring instantly acquires a new trait z, picked randomly according to the mutation step measure $M(x, z)dz$. When one of these events occurs, all individual's clock are reset to 0.

Thus we are looking for a $\mathcal{M}$-valued Markov process $(\nu_t)_{t \geq 0}$ with infinitesimal generator L, defined for real bounded functions ϕ by

$$
\begin{aligned}
L\phi(\nu) = & \sum_{i=1}^{I} b(x^i, V * \nu(x^i))(1 - \mu(x^i))(\phi(\nu + \delta_{x^i}) - \phi(\nu)) \\
& + \sum_{i=1}^{I} b(x^i, V * \nu(x^i))\mu(x^i) \int_{\mathcal{X}} (\phi(\nu + \delta_z) - \phi(\nu))M(x^i, z)dz \\
& + \sum_{i=1}^{I} d(x^i, U * \nu(x^i))(\phi(\nu - \delta_{x^i}) - \phi(\nu)).
\end{aligned}
\tag{2.2}
$$

The first term of (2.2) captures the effect on the population of birth without mutation; the second term that of birth with mutation, and the last term that of death. The density-dependence makes all terms nonlinear.

2.1 Process construction

Let us justify the existence of a Markov process admitting L as infinitesimal generator. The explicit construction of $(\nu_t)_{t \geq 0}$ also yields three side benefits: providing a rigorous and efficient algorithm for numerical simulations (given hereafter), laying the mathematical basis to derive the moment equations of the process (Section 3), and establishing a general method that will be used to derive some large population limits (Sections 4 and 5).

We make the biologically natural assumption that the trait dependency of birth parameters is "bounded", and at most linear for the death rate. Specifically, we assume

Assumptions (H):

There exist constants $\bar{b}$, $\bar{d}$, $\bar{U}$, $\bar{V}$ and C and a probability density function $\bar{M}$ on $\mathbb{R}^d$ such that for each $\nu = \sum_{i=1}^{I} \delta_{x^i}$ and for $x, z \in \mathcal{X}$,

$$
b(x, V * \nu(x)) \leq \bar{b}, \quad d(x, U * \nu(x)) \leq \bar{d}(1 + I),
$$

$$
U(x) \leq \bar{U}, \quad V(x) \leq \bar{V},
$$

$$
M(x, z) \leq C\bar{M}(z - x).
$$

These assumptions ensure that there exists a constant $\bar{C}$, such that the total event rate, for a population counting measure $\nu = \sum_{i=1}^{I} \delta_{x^i}$, obtained as the sum of all event rates, is bounded by $\bar{C}I(1 + I)$.

Let us now give a pathwise description of the population process $(\nu_t)_{t \geq 0}$. We introduce the following notation.

Notation 2.1 *Let $\mathbb{N}^* = \mathbb{N} \backslash \{0\}$. Let $H = (H^1, ..., H^k, ...) : \mathcal{M} \mapsto (\mathbb{R}^d)^{\mathbb{N}^*}$ be defined by $H\left(\sum_{i=1}^{n} \delta_{x_i}\right) = (x_{\sigma(1)}, ..., x_{\sigma(n)}, 0, ..., 0, ...)$, where $x_{\sigma(1)} \preccurlyeq ... \preccurlyeq x_{\sigma(n)}$, for some arbitrary order $\preccurlyeq$ on $\mathbb{R}^d$ (for example the lexicographic order).*

This function H allows us to overcome the following (purely notational) problem. Choosing a trait uniformly among all traits in a population $\nu \in \mathcal{M}$ consists in choosing i uniformly in $\{1, ..., \langle \nu, 1 \rangle\}$, and then in choosing the individual *number i* (from the arbitrary order point of view). The trait value of such an individual is thus $H^i(\nu)$.

We now introduce the probabilistic objects we will need.

Definition 2.2 *Let $(\Omega, \mathcal{F}, P)$ be a (sufficiently large) probability space. On this space, we consider the following four independent random elements:*

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

(ii) *independent Poisson point measures $M_1(ds, di, d\theta)$, and $M_3(ds, di, d\theta)$ on $[0, \infty) \times \mathbb{N}^* \times \mathbb{R}^+$, with the same intensity measure $ds \left(\sum_{k \geq 1} \delta_k(di) \right) d\theta$ (the "clonal" birth and the death Poisson measures),*

(iii) *a Poisson point measure $M_2(ds, di, dz, d\theta)$ on $[0, \infty) \times \mathbb{N}^* \times \mathcal{X} \times \mathbb{R}^+$, with intensity measure $ds \left(\sum_{k \geq 1} \delta_k(di) \right) dz d\theta$ (the mutation Poisson measure).*

Let us denote by $(\mathcal{F}_t)_{t \geq 0}$ the canonical filtration generated by these processes.

We finally define the population process in terms of these stochastic objects.

Definition 2.3 *Assume (H). A $(\mathcal{F}_t)_{t \geq 0}$-adapted stochastic process $\nu = (\nu_t)_{t \geq 0}$ is called a population process if a.s., for all $t \geq 0$,*

$$
\begin{aligned}
\nu_t = \nu_0 &+ \int_{[0,t] \times \mathbb{N}^* \times \mathbb{R}^+} \delta_{H^i(\nu_{s-})} \mathbf{1}_{\{i \leq \langle \nu_{s-}, 1 \rangle\}} \\
&\qquad \mathbf{1}_{\{\theta \leq b(H^i(\nu_{s-}), V * \nu_{s-}(H^i(\nu_{s-})))(1 - \mu(H^i(\nu_{s-})))\}} M_1(ds, di, d\theta) \\
&+ \int_{[0,t] \times \mathbb{N}^* \times \mathcal{X} \times \mathbb{R}^+} \delta_z \mathbf{1}_{\{i \leq \langle \nu_{s-}, 1 \rangle\}} \\
&\qquad \mathbf{1}_{\{\theta \leq b(H^i(\nu_{s-}), V * \nu_{s-}(H^i(\nu_{s-}))) \mu(H^i(\nu_{s-})) M(H^i(\nu_{s-}), z)\}} M_2(ds, di, dz, d\theta) \\
&- \int_{[0,t] \times \mathbb{N}^* \times \mathbb{R}^+} \delta_{H^i(\nu_{s-})} \mathbf{1}_{\{i \leq \langle \nu_{s-}, 1 \rangle\}} \mathbf{1}_{\{\theta \leq d(H^i(\nu_{s-}), U * \nu_{s-}(H^i(\nu_{s-})))\}} M_3(ds, di, d\theta) \qquad (2.3)
\end{aligned}
$$

Let us now show that if ν solves (2.3), then ν follows the Markovian dynamics we are interested in.

Proposition 2.4 *Assume (H) and consider a solution $(\nu_t)_{t \geq 0}$ of Eq. (2.3) such that $E(\sup_{t \leq T} \langle \nu_t, \mathbf{1} \rangle^2) < +\infty$, $\forall T > 0$. Then $(\nu_t)_{t \geq 0}$ is a Markov process. Its infinitesimal generator L is defined for all bounded and measurable maps $\phi : \mathcal{M} \mapsto \mathbb{R}$, all $\nu \in \mathcal{M}$, by (2.2). In particular, the law of $(\nu_t)_{t \geq 0}$ does not depend on the chosen order $\preccurlyeq$.*

Proof The fact that $(\nu_t)_{t \geq 0}$ is a Markov process is classical. Let us now consider a function ϕ as in the statement. With our notation, $\nu_0 = \sum_{i=1}^{\langle \nu_0, 1 \rangle} \delta_{H^i(\nu_0)}$. A simple

computation, using the fact that a.s., $\phi(\nu_t) = \phi(\nu_0) + \sum_{s \le t}(\phi(\nu_{s-} + (\nu_s - \nu_{s-})) - \phi(\nu_{s-}))$, shows that

$$\phi(\nu_t) = \phi(\nu_0) + \int_{[0,t] \times \mathbb{N}^* \times \mathbb{R}^+} \left(\phi(\nu_{s-} + \delta_{H^i(\nu_{s-})}) - \phi(\nu_{s-})\right) \mathbf{1}_{\{i \le \langle \nu_{s-}, 1 \rangle\}}$$

$$\mathbf{1}_{\{\theta \le b(H^i(\nu_{s-}), V*\nu_{s-}(H^i(\nu_{s-})))(1 - \mu(H^i(\nu_{s-})))\}} M_1(ds, di, d\theta)$$

$$+ \int_{[0,t] \times \mathbb{N}^* \times \mathcal{X} \times \mathbb{R}^+} \left(\phi(\nu_{s-} + \delta_z) - \phi(\nu_{s-})\right) \mathbf{1}_{\{i \le \langle \nu_{s-}, 1 \rangle\}}$$

$$\mathbf{1}_{\{\theta \le b(H^i(\nu_{s-}), V*\nu_{s-}(H^i(\nu_{s-})))\mu(H^i(\nu_{s-}))M(H^i(\nu_{s-}),z)\}} M_2(ds, di, dz, d\theta)$$

$$+ \int_{[0,t] \times \mathbb{N}^* \times \mathbb{R}^+} \left(\phi(\nu_{s-} - \delta_{H^i(\nu_{s-})}) - \phi(\nu_{s-})\right) \mathbf{1}_{\{i \le \langle \nu_{s-}, 1 \rangle\}}$$

$$\mathbf{1}_{\{\theta \le d(H^i(\nu_{s-}), U*\nu_{s-}(H^i(\nu_{s-})))\}} M_3(ds, di, d\theta).$$

Taking expectations, we obtain

$$E(\phi(\nu_t)) = E(\phi(\nu_0))$$

$$+ \int_0^t E\Big(\sum_{i=1}^{\langle \nu_s, 1 \rangle} \Big\{ \left(\phi(\nu_s + \delta_{H^i(\nu_s)}) - \phi(\nu_s)\right) b(H^i(\nu_s), V*\nu_s(H^i(\nu_s)))(1 - \mu(H^i(\nu_s)))$$

$$+ \int_{\mathcal{X}} \left(\phi(\nu_s + \delta_z) - \phi(\nu_s)\right) b(H^i(\nu_s), V*\nu_s(H^i(\nu_s)))\mu(H^i(\nu_s))M(H^i(\nu_s),z)dz$$

$$+ \left(\phi(\nu_s - \delta_{H^i(\nu_s)}) - \phi(\nu_s)\right) d(H^i(\nu_s), U*\nu_s(H^i(\nu_s))) \Big\} \Big) ds$$

Differentiating this expression at $t = 0$ leads to (2.2). $\qquad\square$

Let us show existence and moment properties for the population process.

Theorem 2.5 (i) *Assume (H) and that $E(\langle \nu_0, 1 \rangle) < \infty$. Then the process $(\nu_t)_{t \ge 0}$ defined by Definition 2.3 is well defined on $\mathbb{R}_+$.*

(ii) *If furthermore for some $p \ge 1$, $E(\langle \nu_0, 1 \rangle^p) < \infty$, then for any $T < \infty$,*

$$E(\sup_{t \in [0,T]} \langle \nu_t, 1 \rangle^p) < \infty. \tag{2.4}$$

Proof We first prove (ii). Consider the process $(\nu_t)_{t \ge 0}$. We introduce for each n the stopping time $\tau_n = \inf\{t \ge 0, \langle \nu_t, 1 \rangle \ge n\}$. Then a simple computation using Assumption (H) shows that, dropping the non-positive death terms,

$$\sup_{s \in [0, t \wedge \tau_n]} \langle \nu_s, 1 \rangle^p \le \langle \nu_0, 1 \rangle^p + \int_{[0, t \wedge \tau_n] \times \mathbb{N}^* \times \mathbb{R}^+} \left((\langle \nu_{s-}, 1 \rangle + 1)^p - \langle \nu_{s-}, 1 \rangle^p\right) \mathbf{1}_{\{i \le \langle \nu_{s-}, 1 \rangle\}}$$

$$\mathbf{1}_{\{\theta \le b(H^i(\nu_{s-}), V*\nu_{s-}(H^i(\nu_{s-})))(1 - \mu(H^i(\nu_{s-})))\}} M_1(ds, di, d\theta)$$

$$+ \int_{[0,t] \times \mathbb{N}^* \times \mathcal{X} \times \mathbb{R}^+} \left((\langle \nu_{s-}, 1 \rangle + 1)^p - \langle \nu_{s-}, 1 \rangle^p\right) \mathbf{1}_{\{i \le \langle \nu_{s-}, 1 \rangle\}}$$

$$\mathbf{1}_{\{\theta \le b(H^i(\nu_{s-}), V*\nu_{s-}(H^i(\nu_{s-})))\mu(H^i(\nu_{s-}))M(H^i(\nu_{s-}),z)\}} M_2(ds, di, dz, d\theta).$$

6

Using the inequality $(1+x)^p - x^p \leq C_p(1 + x^{p-1})$ and taking expectations, we thus obtain, the value of C_p changing from line to line,

$$E(\sup_{s \in [0, t \wedge \tau_n]} \langle \nu_s, 1 \rangle^p) \leq C_p \left(1 + E \left(\int_0^{t \wedge \tau_n} \bar{b} \left(\langle \nu_{s-}, 1 \rangle + \langle \nu_{s-}, 1 \rangle^p \right) ds \right) \right)$$
$$\leq C_p \left(1 + E \left(\int_0^t \left(1 + \langle \nu_{s \wedge \tau_n}, 1 \rangle^p \right) ds \right) \right).$$

The Gronwall Lemma allows us to conclude that for any $T < \infty$, there exists a constant $C_{p,T}$, not depending on n, such that

$$E(\sup_{t \in [0, T \wedge \tau_n]} \langle \nu_t, 1 \rangle^p) \leq C_{p,T}. \tag{2.5}$$

First, we deduce that τ_n tends a.s. to infinity. Indeed, if not, one may find a $T_0 < \infty$ such that $\epsilon_{T_0} = P(\sup_n \tau_n < T_0) > 0$. This would imply that $E\left(\sup_{t \in [0, T_0 \wedge \tau_n]} \langle \nu_t, 1 \rangle^p \right) \geq \epsilon_{T_0} n^p$ for all n, which contradicts (2.5). We may let n go to infinity in (2.5) thanks to the Fatou Lemma. This leads to (2.4).

Point (i) is a consequence of Point (ii). Indeed, one builds the solution $(\nu_t)_{t \geq 0}$ step by step. One only has to check that the sequence of jump instants T_n goes a.s. to infinity as n tends to infinity. But this follows from (2.4) with $p = 1$. $\qquad \square$

2.2 Examples and simulations

Let us remark that Assumption (H) is satisfied in the case where

$$b(x, V * \nu(x)) = b(x), \quad d(x, U * \nu(x)) = d(x) + \alpha(x) \int_{\mathcal{X}} U(x - y) \nu(dy), \tag{2.6}$$

where b, d and α are bounded functions.

In the case where moreover, $\mu \equiv 1$, this individual-based model can also be interpreted as a model of "spatially structured population", where the trait is viewed as a spatial location and the mutation at each birth event is viewed as dispersal. This kind of models have been introduced by Bolker and Pacala ([2, 3]) and Law et al. ([19]), and mathematically studied by Fournier and Méléard [15]. The case $U \equiv 1$ corresponds to a density-dependence in the total population size.

We will consider later the particular set of parameters for the logistic interaction model, taken from Kisdi [18] and corresponding to a model of asymmetrical competition:

$$\bar{\mathcal{X}} = [0, 4], \quad d(x) = 0, \quad \alpha(x) = 1, \quad \mu(x) = \mu,$$
$$b(x) = 4 - x, \quad U(x - y) = \frac{2}{K} \left(1 - \frac{1}{1 + 1, 2 \exp(-4(x - y))} \right) \tag{2.7}$$

and $M(x, z)dz$ is a Gaussian law with mean x and variance σ^2 conditionned to the fact that the mutant stays in $[0, 4]$. As we will see in Section 4, the constant K scaling the strength of competition also scales the population size (when the initial population size is proportional

7

to K). In this model, the trait x can be interpreted as body size. Equation (2.7) means that body size influences the birth rate negatively, and creates asymmetrical competition reflected in the sigmoid shape of U (being larger is competitively advantageous).

Let us give an algorithmic construction for the population process (in the general case), simulating the size $I(t)$ of the population, and the trait vector $\mathbf{X}_t$ of all individuals alive at time t.

At time $t = 0$, the initial population ν_0 contains $I(0)$ individuals and the corresponding trait vector is $\mathbf{X}_0 = (X_0^i)_{1 \leq i \leq I(0)}$. We introduce the following sequences of independent random variables, which will drive the algorithm.

- The type of birth or death events will be selected according to the values of a sequence of random variables $(W_k)_{k \in \mathbb{N}^*}$ with uniform law on $[0, 1]$.

- The times at which events may be realized will be described using a sequence of random variables $(\tau_k)_{k \in \mathbb{N}}$ with exponential law with parameter $\bar{C}$.

- The mutation steps will be driven by a sequence of random variables $(Z_k)_{k \in \mathbb{N}}$ with law $\bar{M}(z)dz$.

We set $T_0 = 0$ and construct the process inductively for $k \geq 1$ as follows.

At step $k-1$, the number of individuals is I_{k-1}, and the trait vector of these individuals is $\mathbf{X}_{T_{k-1}}$.

Let $T_k = T_{k-1} + \dfrac{\tau_k}{I_{k-1}(I_{k-1}+1)}$. Notice that $\dfrac{\tau_k}{I_{k-1}(I_{k-1}+1)}$ represents the time between jumps for I_{k-1} individuals, and $\bar{C}(I_{k-1}+1)$ gives an upper bound on the total event rate for each individual.

At time T_k, one chooses an individual $i_k = i$ uniformly at random among the I_{k-1} alive in the time interval $[T_{k-1}, T_k)$; its trait is $X_{T_{k-1}}^i$. (If $I_{k-1} = 0$ then $\nu_t = 0$ for all $t \geq T_{k-1}$.)

- If $0 \leq W_k \leq \dfrac{d(X_{T_{k-1}}^i, \sum_{j=1}^{I_{k-1}} U(X_{T_{k-1}}^i - X_{T_{k-1}}^j))}{\bar{C}(I_{k-1}+1)} = W_1^i(\mathbf{X}_{T_{k-1}})$, then the chosen individual dies, and $I_k = I_{k-1} - 1$.

- If $W_1^i(\mathbf{X}_{T_{k-1}}) < W_k \leq W_2^i(\mathbf{X}_{T_{k-1}})$, where

$$W_2^i(\mathbf{X}_{T_{k-1}}) = W_1^i(\mathbf{X}_{T_{k-1}}) + \frac{[1-\mu(X_{T_{k-1}}^i)]b(X_{T_{k-1}}^i, \sum_{j=1}^{I_{k-1}} V(X_{T_{k-1}}^i - X_{T_{k-1}}^j))}{\bar{C}(I_{k-1}+1)},$$

then the chosen individual gives birth to an offspring with trait $X_{T_{k-1}}^i$, and $I_k = I_{k-1} + 1$.

- If $W_2^i(\mathbf{X}_{T_{k-1}}) < W_k \leq W_3^i(\mathbf{X}_{T_{k-1}}, Z_k)$, where

$$W_3^i(\mathbf{X}_{T_{k-1}}, Z_k) = W_2^i(\mathbf{X}_{T_{k-1}}) +$$
$$\frac{\mu(X_{T_{k-1}}^i)b(X_{T_{k-1}}^i, \sum_{j=1}^{I_{k-1}} V(X_{T_{k-1}}^i - X_{T_{k-1}}^j))M(X_{T_{k-1}}^i, X_{T_{k-1}}^i + Z_k)}{\bar{C}\bar{M}(Z_k)(I_{k-1}+1)},$$

then the chosen individual gives birth to a mutant offspring with trait $X_{T_{k-1}}^i + Z_k$, and $I_k = I_{k-1} + 1$.

8

- If $W_k > W_3^i(\mathbf{X}_{T_{k-1}}, Z_k)$, nothing happens, and $I_k = I_{k-1}$.

Then, at any time $t \geq 0$, the number of individuals is defined by $I(t) = \sum_{k \geq 0} 1_{\{T_k \leq t < T_{k+1}\}} I_k$ and the population process is obtained as $\nu_t = \sum_{k \geq 0} 1_{\{T_k \leq t < T_{k+1}\}} \sum_{i=1}^{I_k} \delta_{X_{T_k}^i}$.

The simulation of Kisdi's example (2.7) can be carried out following this algorithm. We can show a very wide variety of qualitative behavior according to the value of the parameters σ, μ and K.

In the following figures, the upper part gives the distribution of the traits in the population at any time, using a grey scale code for the number of individuals holding a given trait. The lower part of the simulation represents the dynamics of the total size $I(t)$ of the population.

These simulations will serve to illustrate the different mathematical scalings described in Sections 4 and 5. Let us observe for the moment the qualitative differences between the cases where K is large (Fig. 1 (c)), in which a wide population density evolves regularly (see Section 4.1) and where μ is small (Fig. 1 (d)), in which the population trait evolves according to a jump process (see Section 5.1).

The simulations of Fig. 2 involve an acceleration of the birth and death processes (see Section 4.2) as

$$b(x, \zeta) = K^\eta + b(x) \quad \text{and} \quad d(x, \zeta) = K^\eta + d(x) + \alpha(x)\zeta.$$

There is a noticeable qualitative difference between Fig. 2 (a) and (b), where $\eta = 1/2$, and Fig. 2 (c) and (d), where $\eta = 1$. In the latter, we observe strong fluctuations in the population size and a finely branched structure of the evolutionnary pattern, revealing a new form of stochasticity in the large population approximation.

More discussions about these simulations are given in [7], especially about the branching pattern of some of them.

2.3   Martingale Properties

We finally give some martingale properties of the process $(\nu_t)_{t \geq 0}$, which are the key point of our approach.

Theorem 2.6 *Assume (H), and that for some $p \geq 2$, $E\left(\langle \nu_0, 1 \rangle^p\right) < \infty$.*

(i) *For all measurable functions ϕ from $\mathcal{M}$ into $\mathbb{R}$ such that for some constant C, for all $\nu \in \mathcal{M}$, $|\phi(\nu)| + |L\phi(\nu)| \leq C(1 + \langle \nu, 1 \rangle^p)$, the process*

$$\phi(\nu_t) - \phi(\nu_0) - \int_0^t L\phi(\nu_s)ds \tag{2.8}$$

is a càdlàg $(\mathcal{F}_t)_{t \geq 0}$-martingale starting from 0.

(ii) *Point (i) applies to any function $\phi(\nu) = \langle \nu, f \rangle^q$, with $0 \leq q \leq p-1$ and with f bounded and measurable on $\mathcal{X}$.*

9

(a) $\mu = 0.03$, $K = 100$, $\sigma = 0.1$.

(b) $\mu = 0.03$, $K = 3000$, $\sigma = 0.1$.

(c) $\mu = 0.03$, $K = 100000$, $\sigma = 0.1$.

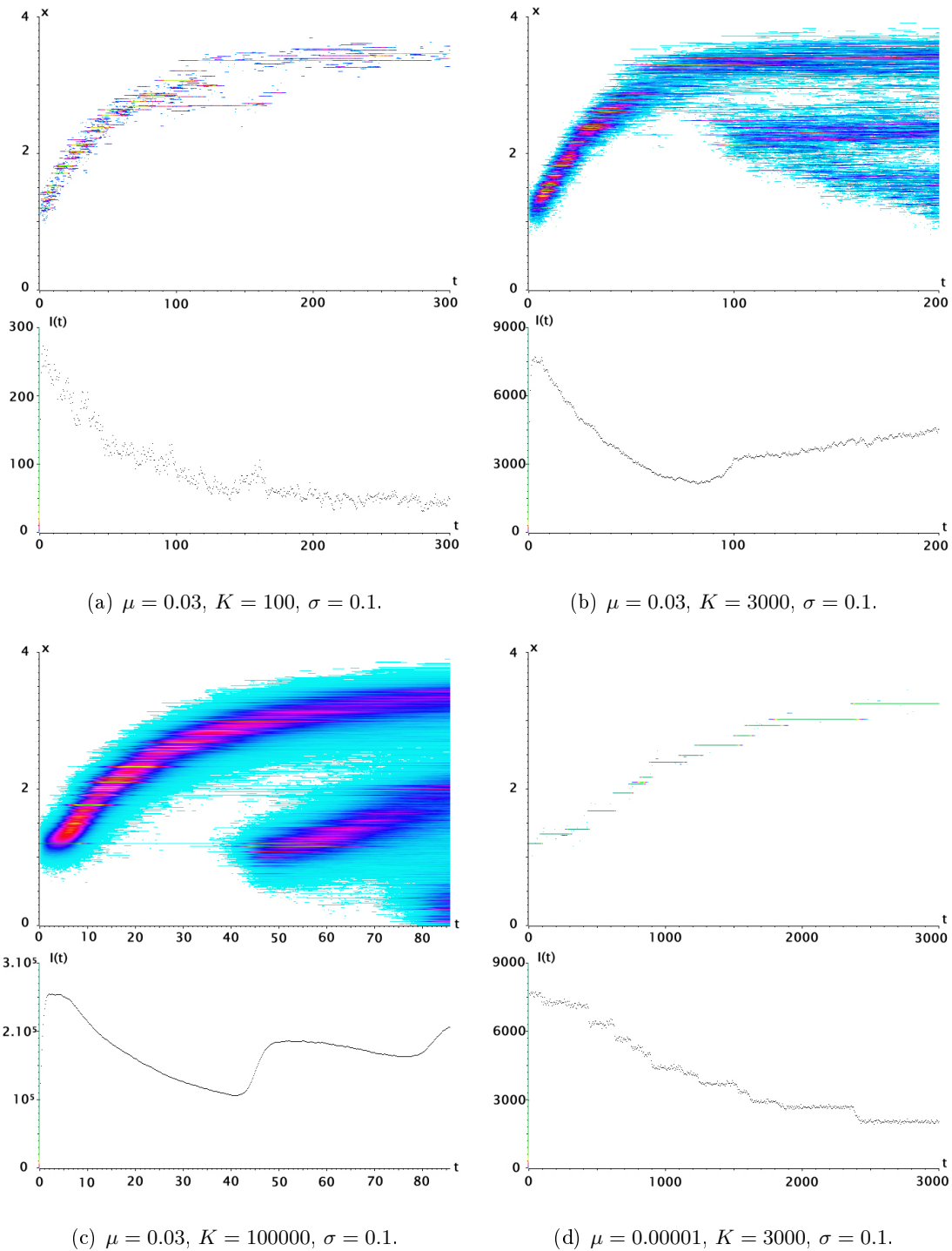(d) $\mu = 0.00001$, $K = 3000$, $\sigma = 0.1$.

Figure 1: Numerical simulations of trait distributions (upper panels, darker is higher frequency) and population size (lower panels). The initial population is monomorphic with trait value 1.2 and contains K individuals. (a–c) Qualitative effect of increasing system size (measured by parameter K). (d) Large parameter K and very small mutation probability (μ).

(a) $\mu = 0.3$, $K = 10000$, $\sigma = 0.3/K^{\eta/2}$, $\eta = 0.5$.

(b) $\mu = 0.1/K^{\eta}$, $K = 10000$, $\sigma = 0.1$, $\eta = 0.5$.

(c) $\mu = 0.3$, $K = 10000$, $\sigma = 0.3/K^{\eta/2}$, $\eta = 1$.

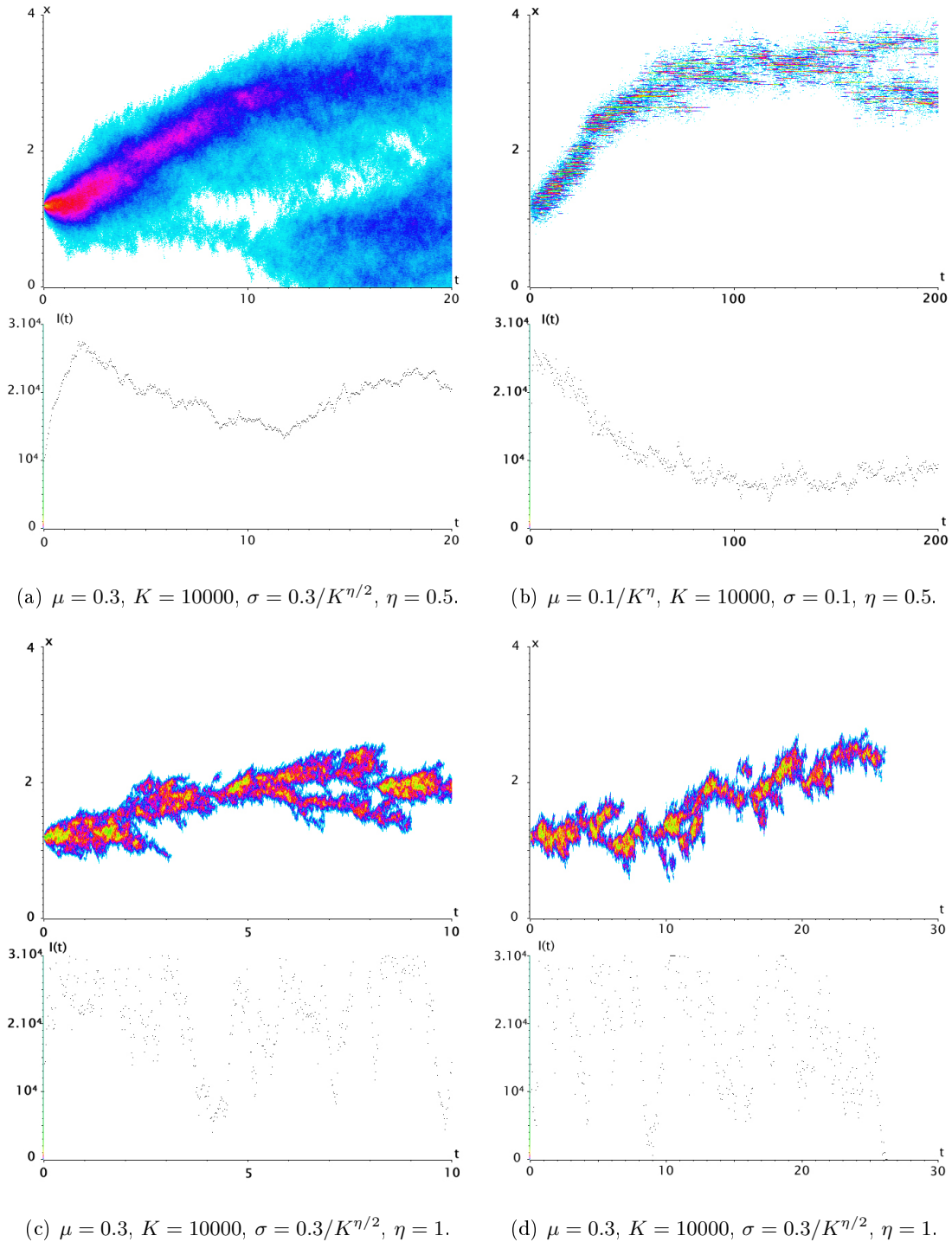(d) $\mu = 0.3$, $K = 10000$, $\sigma = 0.3/K^{\eta/2}$, $\eta = 1$.

Figure 2: Numerical simulations of trait distribution (upper panels, darker is higher frequency) and population size (lower panels) for accelerated birth and death and concurrently increased parameter K. Parameter η (between 0 and 1) relates the acceleration of demographic turnover and the increase of system size K. (a) Rescaling mutation step. (b) Rescaling mutation probability. (c–d) Rescaling mutation step in the limit case $\eta = 1$; two samples for the same population. The initial population is monomorphic with trait value 1.2 and contains K individuals.

(iii) *For such a function f, the process*

$$M_t^f = \langle \nu_t, f \rangle - \langle \nu_0, f \rangle - \int_0^t \int_{\mathcal{X}} \left\{ \left((1 - \mu(x))b(x, V * \nu_s(x)) - d(x, U * \nu_s(x)) \right) f(x) \right.$$

$$\left. + \mu(x)b(x, V * \nu_s(x)) \int_{\mathcal{X}} f(z)M(x, z)dz \right\} \nu_s(dx)ds \tag{2.9}$$

is a càdlàg square integrable martingale starting from 0 with quadratic variation

$$\langle M^f \rangle_t = \int_0^t \int_{\mathcal{X}} \left\{ \left((1 - \mu(x))b(x, V * \nu_s(x)) - d(x, U * \nu_s(x)) \right) f^2(x) \right.$$

$$\left. + \mu(x)b(x, V * \nu_s(x)) \int_{\mathcal{X}} f^2(z)M(x, z)dz \right\} \nu_s(dx)ds. \tag{2.10}$$

Proof   First of all, note that point (i) is immediate thanks to Proposition 2.4 and (2.4). Point (ii) follows from a straightforward computation using (2.2). To prove (iii), we first assume that $E\left(\langle \nu_0, 1 \rangle^3 \right) < \infty$. We apply (i) with $\phi(\nu) = \langle \nu, f \rangle$. This yields that M^f is a martingale. To compute its bracket, we first apply (i) with $\phi(\nu) = \langle \nu, f \rangle^2$ and obtain that

$$\langle \nu_t, f \rangle^2 - \langle \nu_0, f \rangle^2 - \int_0^t \int_{\mathcal{X}} \left\{ \left((1 - \mu(x))b(x, V * \nu_s(x))(f^2(x) + 2f(x) \langle \nu_s, f \rangle) \right. \right.$$

$$\left. + d(x, U * \nu_s(x))(f^2(x) - 2f(x) \langle \nu_s, f \rangle) \right)$$

$$\left. + \mu(x)b(x, V * \nu_s(x)) \int_{\mathcal{X}} (f^2(z) + 2f(z) \langle \nu_s, f \rangle)M(x, z)dz \right\} \nu_s(dx)ds \tag{2.11}$$

is a martingale. In another hand, we apply the Itô formula to compute $\langle \nu_t, f \rangle^2$ from (2.9). We deduce that

$$\langle \nu_t, f \rangle^2 - \langle \nu_0, f \rangle^2 - \int_0^t 2 \langle \nu_s, f \rangle \int_{\mathcal{X}} \left\{ \left((1 - \mu(x))b(x, V * \nu_s(x)) - d(x, U * \nu_s(x)) \right) f(x) \right.$$

$$\left. + \mu(x)b(x, V * \nu_s(x)) \int_{\mathcal{X}} f(z)M(x, z)dz \right\} \nu_s(dx)ds - \langle M^f \rangle_t \tag{2.12}$$

is a martingale. Comparing (2.11) and (2.12) leads to (2.10). The extension to the case where only $E\left(\langle \nu_0, 1 \rangle^2 \right) < \infty$ is straightforward, since even in this case, $E(\langle M^f \rangle_t) < \infty$ thanks to (2.4) with $p = 2$. $\qquad \square$

3   Moment equations

Moment equations have been proposed by Bolker and Pacala ([2, 3]) and Dieckmann and Law ([11]) as handy analytical models for spatially structured populations.

The philosophy of moment equations is germane to the principle of Monte-Carlo methods: computing the mean path of the point process from a large number of independent realizations. (Another approach, as we shall see in Section 4, is to model the behavior of a single trajectory when it is the initial number of individuals which is made large).

Let us define the deterministic measure $E(\nu)$ associated with a random measure ν by $\int_{\mathcal{X}} \varphi(x)E(\nu)(dx) = E(\int_{\mathcal{X}} \varphi(x)\nu(dx))$. Taking expectations in (2.9), we obtain some formula for $\int_{\mathcal{X}} \varphi(x)E(\nu)(dx)$ involving the expectations of integrals with respect to $\nu(dx)$ or to $\nu(dx)\nu(dy)$. Nevertheless, this equation is very intricate and presents an unresolved hierarchy of nonlinearities. Writing an equation for $E(\nu(dx)\nu(dy))$ could be possible but will involve integrals with respect to $\nu(dx)\nu(dy)\nu(dz)$ and so on. Whether this approach may eventually help describe the population dynamics in the trait space is still unclear.

Let us consider the case of spatially structured population (see Section 2.2) where $d(x,\zeta) = d(x) + \alpha(x)\zeta$, $b(x,\zeta) = b(x)$ and $\mu(x) = 1$. Let $N(t) = E(I(t))$ where $I(t)$ is the number of individuals at time t. Taking expectations on (2.9) with $\varphi \equiv 1$ yields:

$$N(t) = N(0) + \int_0^t E\left(\int_{\mathcal{X}}(b(x) - d(x))\nu_s(dx) - \int_{\mathcal{X}\times\mathcal{X}} \alpha(x)U(x - y)\nu_s(dx)\nu_s(dy)\right)ds. \quad (3.1)$$

In the specific case where b, d and α are independent of (the spatial location) x, (cf. [19]), (3.1) recasts into

$$\dot{N} = (b - d)N - \alpha E\left(\int_{\mathcal{X}\times\mathcal{X}} U(x - y)\nu_t(dx)\nu_t(dy)\right). \quad (3.2)$$

Even in the specific mean-field case where $U = 1$, we get

$$\dot{N} = (b - d)N - \alpha E\left(\int_{\mathcal{X}\times\mathcal{X}} \nu_t(dx)\nu_t(dy)\right). \quad (3.3)$$

The quadratic term corresponding to spatial correlations can not be simplified and (3.3) allows us to precisely identify the mathematical issues raised by the problem of moment closure. In Section 4.1, we will see that one needs the additional large population hypothesis to decorrelate the quadratic term and to recover the well-known logistic equation.

Nevertheless, even if we are not able to produce a closed equation satisfied by $E(\nu)$, we are able to show, in the general case, the following qualitative important property concerning the absolute continuity of the expectation of ν_t.

Proposition 3.1 *Assume (H), that $E(\langle\nu_0, 1\rangle) < \infty$ and that $E(\nu_0)$ is absolutely continuous with respect to the Lebesgue measure. Then for all $t \geq 0$, $E(\nu_t)$ is absolutely continuous with respect to the Lebesgue measure.*

Remark 3.2 *This implies in particular that, when the initial trait distribution $E(\nu_0)$ has no singularity w.r.t. the Lebesgue measure, these singularities, such as Dirac masses, can only appear in the limit of infinite time.*

Proof Consider a Borel set A of $\mathbb{R}^d$ with Lebesgue measure zero. Consider also, for each $n \geq 1$, the stopping time $\tau_n = \inf\{t \geq 0, \langle\nu_t, 1\rangle \geq n\}$. A simple computation allows us to obtain, for all $t \geq 0$, all $n \geq 1$,

$$E\left(\langle\nu_{t\wedge\tau_n}, \mathbf{1}_A\rangle\right) \leq E(\langle\nu_0, \mathbf{1}_A\rangle) + \bar{b}\, E\left(\int_0^{t\wedge\tau_n} \int_{\mathcal{X}} \mathbf{1}_A(x)\nu_s(dx)ds\right)$$
$$+ \bar{b}\, E\left(\int_0^{t\wedge\tau_n} \int_{\mathcal{X}} \left(\int_{\mathcal{X}} \mathbf{1}_A(z)M(x,z)dz\right)\nu_s(dx)ds\right).$$

13

By assumption, the first term on the RHS is zero. The third term is also zero, since for any $x \in \mathcal{X}$, $\int_{\mathcal{X}} \mathbf{1}_A(z)M(x,z)dz = 0$. By Gronwall's lemma, we conclude that for each n, $E(\langle \nu_{t \wedge \tau_n}, \mathbf{1}_A \rangle)$ is zero. Thanks to (2.4) with $p = 1$, τ_n a.s. grows to infinity with n, which concludes the proof. $\qquad \square$

4   Large-population renormalizations of the individual-based process

The moment equation approach outlined above is based on the idea of averaging a large number of independent realizations of the population process initiated with a finite number of individuals. If K scales the initial number of individuals, the alternative approach consists in studying the exact process by letting that system size become very large and making some appropriate renormalizations. Several types of approximations can then be derived, depending on these renormalizations.

For any K, let the set of parameters U_K, V_K, b_K, d_K, M_K, μ_K satisfy the Assumption (H). Let ν_t^K be the counting measure of the population at time t. We define the measure-valued Markov process $(X_t^K)_{t \geq 0}$ by

$$X_t^K = \frac{1}{K} \nu_t^K.$$

As the system size K goes to infinity, we need to assume the

Assumption (H1): The parameters U_K, V_K, b_K, d_K, M_K and μ_K are all continuous, $\zeta \mapsto b(x, \zeta)$ and $\zeta \mapsto d(x, \zeta)$ are Lipschitz for any $x \in \mathcal{X}$, and

$$U_K(x) = U(x)/K, \quad V_K(x) = V(x)/K.$$

A biological interpretation of this renormalization is that larger systems are made up of smaller individuals, which may be a consequence of a fixed amount of available resources to be partitioned among individuals. Thus, the biomass of each interacting individual scales as $1/K$, which may imply that the interaction effect of the global population on a focal individual is of order 1. Parameter K may also be interpreted as scaling the resources available, so that the renormalization of U_K and V_K reflects the decrease of competition for resources.

The generator $\tilde{L}^K$ of $(\nu_t^K)_{t \geq 0}$ is given by (2.2), with parameters U_K, V_K, b_K, d_K, M_K, μ_K. The generator L^K of $(X_t^K)_{t \geq 0}$ is obtained by writing, for any measurable function ϕ from $M_F(\mathcal{X})$ into $\mathbb{R}$ and any $\nu \in M_F(\mathcal{X})$,

$$L^K \phi(\nu) = \partial_t E_\nu(\phi(X_t^K))_{t=0} = \partial_t E_{K\nu}(\phi(\nu_t^K/K))_{t=0} = \tilde{L}^K \phi^K(K\nu)$$

where $\phi^K(\mu) = \phi(\mu/K)$. Then we get

$$
\begin{aligned}
L^K \phi(\nu) = {}& K \int_{\mathcal{X}} b_K(x, V * \nu(x))(1 - \mu_K(x))(\phi(\nu + \frac{1}{K}\delta_x) - \phi(\nu))\nu(dx) \\
& + K \int_{\mathcal{X}} \int_{\mathcal{X}} b_K(x, V * \nu(x))\mu_K(x)(\phi(\nu + \frac{1}{K}\delta_z) - \phi(\nu))M_K(x,z)dz\nu(dx) \\
& + K \int_{\mathcal{X}} d_K(x, U * \nu(x))(\phi(\nu - \frac{1}{K}\delta_x) - \phi(\nu))\nu(dx).
\end{aligned}
\tag{4.1}
$$

By a similar proof as the one of Section 2.3, we may summarize the moment and martingale properties of X^K.

Proposition 4.1 *Assume that for some $p \geq 2$, $E(\langle X_0^K, 1 \rangle^p) < +\infty$.*

(1) *For any $T > 0$, $E(\sup_{t \in [0,T]} \langle X_t^K, 1 \rangle^p) < +\infty$.*

(2) *For any bounded and measurable functions ϕ on M_F such that $|\phi(\nu)| + |L^K \phi(\nu)| \leq C(1+ < \nu, 1 >^p)$, the process $\phi(X_t^K) - \phi(X_0^K) - \int_0^t L^K \phi(X_s^K)ds$ is a càdlàg martingale.*

(3) *For each measurable bounded function f, the process*

$$m_t^{K,f} = \langle X_t^K, f \rangle - \langle X_0^K, f \rangle$$
$$- \int_0^t \int_{\mathcal{X}} (b_K(x, V * X_s^K(x)) - d_K(x, U * X_s^K(x)))f(x)X_s^K(dx)ds$$
$$- \int_0^t \int_{\mathcal{X}} \mu_K(x)b_K(x, V * X_s^K(x)\left(\int_{\mathcal{X}} f(z)M_K(x,z)dz - f(x) \right)X_s^K(dx)ds$$

is a square integrable martingale with quadratic variation

$$\langle m^{K,f} \rangle_t = \frac{1}{K}\left\{ \int_0^t \int_{\mathcal{X}} \mu_K(x)b_K(x, V * X_s^K(x))\left(\int_{\mathcal{X}} f^2(z)M_K(x,z)dz - f^2(x) \right)X_s^K(dx)ds \right.$$
$$\left. + \int_0^t \int_{\mathcal{X}} (b_K(x, V * X_s^K(x)) + d_K(x, U * X_s^K(x)))f^2(x)X_s^K(dx)ds \right\} \quad (4.2)$$

The search of tractable limits for the semimartingales $\langle X^K, f \rangle$ yields the different choices of scalings of the parameters developed in this section. In particular, we obtain the deterministic or stochastic nature of the approximation by studying the quadratic variation of the martingale term, given in (4.2).

4.1   Large-population limit

We assume here that $b_K = b$, $d_K = d$, $\mu_K = \mu$, $M_K = M$.

Theorem 4.2 *Assume Assumptions (H) and (H1). Assume moreover that the initial conditions X_0^K converge in law and for the weak topology on $M_F(\mathcal{X})$ as K increases, to a finite deterministic measure ξ_0, and that $\sup_K E(\langle X_0^K, 1 \rangle^3) < +\infty$.*
Then for any $T > 0$, the process $(X_t^K)_{t \geq 0}$ converges in law, in the Skorohod space $\mathbb{D}([0,T], M_F(\mathcal{X}))$, as K goes to infinity, to the unique deterministic continuous function $\xi \in C([0,T], M_F(\mathcal{X}))$ satisfying for any bounded $f : \mathcal{X} \to \mathbb{R}$

$$\langle \xi_t, f \rangle = \langle \xi_0, f \rangle + \int_0^t \int_{\mathcal{X}} f(x)[(1 - \mu(x))b(x, V * \xi_s(x)) - d((x, U * \xi_s(x))]\xi_s(dx)ds$$
$$+ \int_0^t \int_{\mathcal{X}} \mu(x)b(x, V * \xi_s(x)) \left(\int_{\mathcal{X}} f(z)M(x,z)dz \right) \xi_s(dx)ds \quad (4.3)$$

The proof of Theorem 4.2 is let to the reader. It can be adapted from the proofs of Theorem 4.3 and 4.5 below, or obtained as a generalization of Theorem 5.3 in [15]. This result is illustrated by the simulations of Figs. 1 (a)–(c).

Main Examples:

(1) **A density case.** Following similar arguments as in the proof of Proposition 3.1, one shows that if the initial condition ξ_0 has a density w.r.t. Lebesgue measure, then the same property holds for the finite measure ξ_t, which is then solution of the functional equation:

$$\partial_t \xi_t(x) = [(1 - \mu(x))b(x, V * \xi_t(x)) - d(x, U * \xi_t(x))] \xi_t(x)$$
$$+ \int_{\mathbb{R}^d} M(y, x)\mu(y)b(y, V * \xi_t(y))\xi_t(y)dy \qquad (4.4)$$

for all $x \in \mathcal{X}$ and $t \geq 0$. Desvillettes et al. [9] suggest to refer to ξ_t as the population number density; then the quantity $n_t = \int_{\mathcal{X}} \xi_t(x)dx$ can be interpreted as the total population density over the whole trait space.

(2) **The mean field case.** As for moment equations (cf. Section 3), the case of spatially structured populations with constant rates b, d, α is meaningful. In this context, (4.3) leads to the following equation on n_t:

$$\partial_t n_t = (b - d)n_t - \alpha \int_{\mathcal{X} \times \mathcal{X}} U(x - y)\xi_t(dx)\xi_t(dy). \qquad (4.5)$$

With the assumption $U \equiv 1$, we recover the classical mean-field logistic equation of population growth:
$$\partial_t n_t = (b - d)n_t - \alpha n_t^2.$$

Comparing (4.5) with the first-moment equation (3.3) obtained previously stresses out the "decorrelative" effect of the large system size renormalization (only in case $U \equiv 1$). In (3.3), the correction term capturing the effect of spatial correlations in the population remains, even if one assumes $U \equiv 1$.

(3) **Monomorphic and dimorphic cases without mutation.** We assume here that the population evolves without mutation (parameter $\mu = 0$); then the population traits are the initial ones.

(a) **Monomorphic case:** only trait x is present in the population at time $t = 0$. Thus, we can write $X_0^K = n_0^K(x)\delta_x$, and then $X_t^K = n_t^K(x)\delta_x$ for any time t. Theorem 4.2 recasts in this case into $n_t^K(x) \to n_t(x)$ with $\xi_t = n_t(x)\delta_x$, and (4.3) writes
$$\frac{d}{dt}n_t(x) = n_t(x)\big(b(x, V(0)n_t(x)) - d(x, U(0)n_t(x))\big), \qquad (4.6)$$

(b) **Dimorphic case:** when the population contains two traits x and y, i.e. when $X_0^K = n_0^K(x)\delta_x + n_0^K(y)\delta_y$, we can define in a similar way $n_t(x)$ and $n_t(y)$ for any t as before, such that $\xi_t = n_t(x)\delta_x + n_t(y)\delta_y$ satisfies (4.3), which recasts into the following system of coupled ordinary differential equations:

16

$$\frac{d}{dt}n_t(x) = n_t(x)\big(b(x, V(0)n_t(x) + V(x-y)n_t(y)) - d(x, U(0)n_t(x) + U(x-y)n_t(y))\big)$$
$$\frac{d}{dt}n_t(y) = n_t(y)\big(b(y, V(0)n_t(y) + V(y-x)n_t(x)) - d(y, U(0)n_t(y) + U(y-x)n_t(x))\big). \tag{4.7}$$

4.2 Large-population limit with accelerated births and deaths

We consider here an alternative limit of a large population, combined with accelerated birth and death. This may be useful to investigate the qualitative differences of evolutionary dynamics across populations with allometric demographies (larger populations made up of smaller individuals who reproduce and die faster, See [5], [8]).

Here, we assume that $\mathcal{X} = \mathbb{R}^d$. Let us denote by M_F the space $M_F(\mathbb{R}^d)$. We consider the acceleration of birth and death processes at a rate proportional to K^η while preserving the demographic balance. That is, the birth and death rates scale with system size according to

Assumption (H2):

$$b_K(x, \zeta) = K^\eta r(x) + b(x, \zeta), \quad d_K(x, \zeta) = K^\eta r(x) + d(x, \zeta).$$

The allometric effect (smaller individuals reproduce and die faster) is parameterized by the function r, positive and bounded over $\mathbb{R}^d$, and the constant η. A detailed discussion of the biological meaning of these parameters in terms of allometry and life-history scalings can be found in [7]. As in Section 4.1, the interaction kernels V and U are renormalized by K. Using similar arguments as in Section 4.1, the process $X^K = \frac{1}{K}\nu^K$ is now a Markov process with generator

$$L^K\phi(\nu) = K\int_{\mathbb{R}^d}(K^\eta r(x) + b(x, V * \nu(x)))(1 - \mu_K(x))(\phi(\nu + \frac{1}{K}\delta_x) - \phi(\nu))\nu(dx)$$

$$+ K\int_{\mathbb{R}^d}(K^\eta r(x) + b(x, V * \nu(x)))\mu_K(x)\int_{\mathbb{R}^d}(\phi(\nu + \frac{1}{K}\delta_z) - \phi(\nu))M_K(x, z)dz\nu(dx)$$

$$+ K\int_{\mathbb{R}^d}(K^\eta r(x) + d(x, U * \nu(x)))(\phi(\nu - \frac{1}{K}\delta_x) - \phi(\nu))\nu(dx).$$

As before, for any measurable functions ϕ on M_F such that $|\phi(\nu)| + |L^K\phi(\nu)| \leq C(1 + \langle\nu, 1\rangle^3)$, the process

$$\phi(X_t^K) - \phi(X_0^K) - \int_0^t L^K\phi(X_s^K)ds \tag{4.8}$$

is a martingale. In particular, for each measurable bounded function f, we obtain

$$M_t^{K,f} = \langle X_t^K, f\rangle - \langle X_0^K, f\rangle$$

$$- \int_0^t \int_{\mathbb{R}^d}(b(x, V * X_s^K(x)) - d(x, U * X_s^K(x)))f(x)X_s^K(dx)ds \tag{4.9}$$

$$- \int_0^t \int_{\mathbb{R}^d}\mu_K(x)(K^\eta r(x) + b(x, V * X_s^K(x)))\left(\int_{\mathbb{R}^d}f(z)M_K(x, z)dz - f(x)\right)X_s^K(dx)ds,$$

17

is a square integrable martingale with quadratic variation

$$\langle M^{K,f} \rangle_t = \frac{1}{K} \Bigg\{ \int_0^t \int_{\mathbb{R}^d} (2K^\eta r(x) + b(x, V * X_s^K(x)) + d(x, U * X_s^K(x))) f^2(x) X_s^K(dx) ds$$

$$+ \int_0^t \int_{\mathbb{R}^d} \mu_K(x)(K^\eta r(x) + b(x, V * X_s^K(x))) \Big(\int_{\mathbb{R}^d} f^2(z) M_K(x,z) dz - f^2(x) \Big) X_s^K(dx) ds \Bigg\}. \tag{4.10}$$

Two interesting cases will be considered hereafter, in which the variance effect $\mu_K M_K$ is of order $1/K^\eta$. That will ensure the deterministic part in (4.9) to converge. In the large-population renormalization (Section 4.1), the quadratic variation of the martingale part was of order $1/K$. Here, it is of order $K^\eta \times 1/K$. This quadratic variation will thus stay finite provided that $\eta \in (0, 1]$, in which case tractable limits will result. Moreover, this limit will be zero if $\eta < 1$ and nonzero if $\eta = 1$, which will lead to deterministic or random limit models.

4.2.1 Accelerated mutation and small mutation steps

We consider here that the mutation rate is fixed, so that mutations are accelerated as a consequence of accelerating birth. We assume

Assumptions (H3):

(1) $\mu_K = \mu$.

(2) The mutation step density $M_K(x, z)$ is the density of a random variable with mean x, variance-covariance matrix $\Sigma(x)/K^\eta$ (where $\Sigma(x) = (\Sigma_{ij}(x))_{1 \le i,j \le d}$) and with third moment of order $1/K^{\eta+\varepsilon}$ uniformly in x ($\varepsilon > 0$). (Thus, as K goes to infinity, mutant traits become more concentrated around their progenitors').

(3) $\sqrt{\Sigma}$ denoting the symmetrical square root matrix of Σ, the function $\sqrt{\Sigma r \mu}$ is Lipschitz continuous.

The main example is when the mutation step density is taken as the density of a vector of independent Gaussian variables with mean x and variance $\sigma^2(x)/K^\eta$:

$$M_K(x, z) = \left(\frac{K^\eta}{2\pi\sigma^2(x)} \right)^{d/2} \exp[-K^\eta |z - x|^2 / 2\sigma^2(x)] \tag{4.11}$$

where $\sigma^2(x)$ is positive and bounded over $\mathbb{R}^d$.

Then the convergence results of this section can be stated as follows.

Theorem 4.3 (1) *Assume (H), (H1), (H2), (H3) and $0 < \eta < 1$. Assume also that the initial conditions X_0^K converge in law and for the weak topology on M_F as K increases, to a finite deterministic measure ξ_0, and that*

$$\sup_K E(\langle X_0^K, 1 \rangle^3) < +\infty. \tag{4.12}$$

18

Then, for each $T > 0$, the sequence of processes (X^K) belonging to $\mathbb{D}([0,T], M_F)$ converges (in law) to the unique deterministic function $(\xi_t)_{t \geq 0} \in C([0,T], M_F)$ satisfying: for each function $f \in C_b^2(\mathbb{R}^d)$,

$$\langle \xi_t, f \rangle = \langle \xi_0, f \rangle + \int_0^t \int_{\mathbb{R}^d} (b(x, V * \xi_s(x)) - d(x, U * \xi_s(x))) f(x) \xi_s(dx) ds$$

$$+ \int_0^t \int_{\mathbb{R}^d} \frac{1}{2} \mu(x) r(x) \sum_{1 \leq i,j \leq d} \Sigma_{ij}(x) \partial_{ij}^2 f(x) \xi_s(dx) ds, \qquad (4.13)$$

where $\partial_{ij}^2 f$ denotes the second-order partial derivative of f with respect to x_i and x_j ($x = (x_1, \ldots, x_d)$).

(2) *Assume moreover that there exists $c > 0$ such that $r(x)\mu(x)s^*\Sigma(x)s \geq c\|s\|^2$ for any x and s in $\mathbb{R}^d$. Then for each $t > 0$, the measure ξ_t has a density with respect to Lebesgue measure.*

Remark 4.4 *In case (2), Eq. (4.13) may be written as*

$$\partial_t \xi_t(x) = \left(b(x, V * \xi_t(x)) - d(x, U * \xi_t(x)) \right) \xi_t(x) + \frac{1}{2} \sum_{1 \leq i,j \leq d} \partial_{ij}^2 (r\mu\Sigma_{ij}\xi_t)(x). \qquad (4.14)$$

Observe that, for the example (4.11), this equation writes

$$\partial_t \xi_t(x) = \left(b(x, V * \xi_t(x)) - d(x, U * \xi_t(x)) \right) \xi_t(x) + \frac{1}{2} \Delta(\sigma^2 r\mu\xi_t)(x). \qquad (4.15)$$

Therefore, Eq. (4.15) generalizes the Fisher reaction-diffusion equation known from classical population genetics (see e.g. [4]).

Theorem 4.5 *Assume (H), (H1), (H2), (H3) and $\eta = 1$. Assume also that the initial conditions X_0^K converge in law and for the weak topology on M_F as K increases, to a finite (possibly random) measure X_0, and that $\sup_K E(\langle X_0^K, 1 \rangle^3) < +\infty$.*

Then, for each $T > 0$, the sequence of processes (X^K) converges in law in $\mathbb{D}([0,T], M_F)$ to the unique (in law) continuous superprocess $X \in C([0,T], M_F)$, defined by the following conditions:

$$\sup_{t \in [0,T]} E\left(\langle X_t, 1 \rangle^3 \right) < \infty, \qquad (4.16)$$

and for any $f \in C_b^2(\mathbb{R}^d)$,

$$\bar{M}_t^f = \langle X_t, f \rangle - \langle X_0, f \rangle - \frac{1}{2} \int_0^t \int_{\mathbb{R}^d} \mu(x) r(x) \sum_{1 \leq i,j \leq d} \Sigma_{ij}(x) \partial_{ij}^2 f(x) X_s(dx) ds$$

$$- \int_0^t \int_{\mathbb{R}^d} f(x) \left(b(x, V * X_s(x)) - d(x, U * X_s(x)) \right) X_s(dx) ds \qquad (4.17)$$

is a continuous martingale with quadratic variation

$$\langle \bar{M}^f \rangle_t = 2 \int_0^t \int_{\mathbb{R}^d} r(x) f^2(x) X_s(dx) ds. \qquad (4.18)$$

Remark 4.6 (1) *The limiting measure-valued process X appears as a generalization of the one proposed by Etheridge [12] to model spatially structured populations.*

(2) *The conditions characterizing the process X above can be formally rewritten as*

$$\partial_t X_t(x) = \left(b(x, V * X_t(x)) - d(x, U * X_t(x)) \right) X_t(x) + \frac{1}{2} \sum_{1 \le i,j \le d} \partial_{ij}^2 (r\mu \Sigma_{ij} X_t)(x) + \dot{M}_t$$

where $\dot{M}_t$ is a random fluctuation term, which reflects the demographic stochasticity of this fast birth-and-death process, that is, faster than the accelerated birth-and-death process which led to the deterministic reaction-diffusion approximation (4.15).

(3) *As developed in Step 1 of the proof of Theorem 4.5 below, a Girsanov's theorem relates the law of X_t and the one of a standard super-Brownian motion, which leads to conjecture that a density for X_t exists only when $d = 1$, as for the super-Brownian motion.*

These two theorems are illustrated by the simulations of Figs. 2 (a), (c) and (d).

Proof of Theorem 4.3   We divide the proof in several steps. Let us fix $T > 0$.

Step 1   Let us first show the uniqueness for a solution of the equation (4.13).

To this aim, we define the evolution equation associated with (4.13). It is easy to prove that if ξ is a solution of (4.13) satisfying $\sup_{t \in [0,T]} \langle \xi_t, 1 \rangle < \infty$, then for each test function $\psi_t(x) = \psi(t, x) \in C_b^{1,2}(\mathbb{R}_+ \times \mathbb{R}^d)$, one has

$$\langle \xi_t, \psi_t \rangle = \langle \xi_0, \psi_0 \rangle + \int_0^t \int_{\mathbb{R}^d} (b(x, V * \xi_s(x)) - d(x, U * \xi_s(x))) \psi(s, x) \xi_s(dx) ds$$

$$+ \int_0^t \int_{\mathbb{R}^d} (\partial_s \psi(s, x) + \frac{1}{2} r(x) \mu(x) \sum_{i,j} \Sigma_{ij}(x) \partial_{ij}^2 \psi_s(x)) \xi_s(dx) ds. \qquad (4.19)$$

Now, since the function $\sqrt{\Sigma r \mu}$ is Lipschitz continuous, we may define the transition semigroup (P_t) whith infinitesimal generator $f \mapsto \frac{1}{2} r\mu \sum_{i,j} \Sigma_{ij} \partial_{ij}^2 f$. Then, for each function $f \in C_b^2(\mathbb{R}^d)$ and fixed $t > 0$, to choose $\psi(s, x) = P_{t-s} f(x)$ yields

$$\langle \xi_t, f \rangle = \langle \xi_0, P_t f \rangle + \int_0^t \int_{\mathbb{R}^d} (b(x, V * \xi_s(x)) - d(x, U * \xi_s(x))) P_{t-s} f(x) \xi_s(dx) ds, \quad (4.20)$$

since $\partial_s \psi(s, x) + \frac{1}{2} r(x) \mu(x) \sum_{i,j} \Sigma_{ij}(x) \partial_{ij}^2 \psi_s(x) = 0$ for this choice.

We now prove the uniqueness of a solution of (4.20).

Let us consider two solutions $(\xi_t)_{t \ge 0}$ and $(\bar{\xi}_t)_{t \ge 0}$ of (4.20) satisfying $\sup_{t \in [0,T]} \langle \xi_t + \bar{\xi}_t, 1 \rangle = A_T < +\infty$. We consider the variation norm defined for μ_1 and μ_2 in M_F by

$$||\mu_1 - \mu_2|| = \sup_{f \in L^\infty(\mathbb{R}^d), ||f||_\infty \le 1} |\langle \mu_1 - \mu_2, f \rangle|. \qquad (4.21)$$

Then, we consider some bounded and measurable function f defined on $\mathcal{X}$ such that $||f||_\infty \le 1$ and obtain

$$
\begin{aligned}
|\langle \xi_t - \bar{\xi}_t, f\rangle| &\le \int_0^t \left| \int_{\mathbb{R}^d} [\xi_s(dx) - \bar{\xi}_s(dx)] \left(b(x, V * \xi_s(x)) - d(x, U * \xi_s(x)) \right) P_{t-s} f(x) \right| ds \\
&+ \int_0^t \left| \int_{\mathbb{R}^d} \bar{\xi}_s(dx)(b(x, V * \xi_s(x)) - b(x, V * \bar{\xi}_s(x))) P_{t-s} f(x) \right| ds \\
&+ \int_0^t \left| \int_{\mathbb{R}^d} \bar{\xi}_s(dx)(d(x, U * \xi_s(x)) - d(x, U * \bar{\xi}_s(x))) P_{t-s} f(x) \right| ds. \qquad (4.22)
\end{aligned}
$$

Since $||f||_\infty \le 1$, then $||P_{t-s}f||_\infty \le 1$ and for all $x \in \mathbb{R}^d$,

$$
|(b(x, V * \xi_s(x)) - d(x, U * \xi_s(x))) P_{t-s} f(x)| \le \bar{b} + \bar{d}(1 + \bar{U} A_T).
$$

Moreover, b and d are Lipschitz continuous in their second variable with respective constants K_b and K_d. Thus we obtain from (4.22) that

$$
|\langle \xi_t - \bar{\xi}_t, f\rangle| \le \left[\bar{b} + \bar{d}(1 + \bar{U} A_T) + K_b A_T \bar{V} + K_d A_T \bar{U} \right] \int_0^t ||\xi_s - \bar{\xi}_s|| ds. \qquad (4.23)
$$

Taking the supremum over all functions f such that $||f||_\infty \le 1$, and using the Gronwall Lemma, we finally deduce that for all $t \le T$, $||\xi_t - \bar{\xi}_t|| = 0$. Uniqueness holds.

Step 2 Next, we would like to obtain some moment estimates. First, we check that for all $T < \infty$,
$$
\sup_K \sup_{t \in [0,T]} E\big(\langle X_t^K, 1\rangle^3\big) < \infty. \qquad (4.24)
$$
To this end, we use (4.8) with $\phi(\nu) = \langle \nu, 1\rangle^3$. (To be completely rigorous, one should first use $\phi(\nu) = \langle \nu, 1\rangle^3 \wedge A$, make A tend to infinity). Taking expectation, we obtain that for all $t \ge 0$, all K,

$$
\begin{aligned}
E\left(\langle X_t^K, 1\rangle^3\right) &= E\left(\langle X_0^K, 1\rangle^3\right) \\
&+ \int_0^t E\bigg(\int_{\mathbb{R}^d} \bigg([K^{\eta+1} r(x) + K b(x, V * X_s^K(x))] \left\{ [\langle X_s^K, 1\rangle + \frac{1}{K}]^3 - \langle X_s^K, 1\rangle^3 \right\} \\
&\left\{ K^{\eta+1} r(x) + K d(x, U * X_s^K(x)) \right\} \left\{ [\langle X_s^K, 1\rangle - \frac{1}{K}]^3 - \langle X_s^K, 1\rangle^3 \right\} \bigg) X_s^K(dx) \bigg) ds.
\end{aligned}
$$

Dropping the non-positive death term involving d, we get

$$
\begin{aligned}
E\left(\langle X_t^K, 1\rangle^3\right) &\le E\left(\langle X_0^K, 1\rangle^3\right) \\
&+ \int_0^t E\bigg(\int_{\mathbb{R}^d} \bigg(K^{\eta+1} r(x) \left\{ [\langle X_s^K, 1\rangle + \frac{1}{K}]^3 + [\langle X_s^K, 1\rangle - \frac{1}{K}]^3 - 2\langle X_s^K, 1\rangle^3 \right\} \\
&+ K b(x, V * X_s^K(x)) \left\{ [\langle X_s^K, 1\rangle + \frac{1}{K}]^3 - \langle X_s^K, 1\rangle^3 \right\} \bigg) X_s^K(dx) \bigg) ds.
\end{aligned}
$$

But for all $x \ge 0$, all $\epsilon \in (0,1]$, $(x+\epsilon)^3 - x^3 \le 6\epsilon(1+x^2)$ and $|(x+\epsilon)^3 + (x-\epsilon)^3 - 2x^3| = 6\epsilon^2 x$. We finally obtain

$$
E\left(\langle X_t^K, 1\rangle^3\right) \le E\left(\langle X_0^K, 1\rangle^3\right) + C \int_0^t E\left(\langle X_s^K, 1\rangle + \langle X_s^K, 1\rangle^2 + \langle X_s^K, 1\rangle^3\right) ds.
$$

21

Assumption (4.12) and the Gronwall Lemma allows us to conclude that (4.24) holds. Next, we wish to check that

$$\sup_K E\big(\sup_{t\in[0,T]} \langle X_t^K, 1\rangle^2\big) < \infty. \qquad (4.25)$$

Applying (4.9) with $f \equiv 1$, we obtain

$$\langle X_t^K, 1\rangle = \langle X_0^K, 1\rangle + \int_0^t \int_{\mathcal{X}} \left(b(x, V * X_s^K(x)) - d(x, U * X_s^K(x)) \right) X_s^K(dx) ds + m_t^{K,1}.$$

Hence

$$\sup_{s\in[0,t]} \langle X_s^K, 1\rangle^2 \le C\left(\langle X_0^K, 1\rangle^2 + \bar{b} \int_0^t \langle X_s^K, 1\rangle^2 ds + \sup_{s\in[0,t]} |M_s^{K,1}|^2 \right).$$

Thanks to (4.12), the Doob inequality and the Gronwall Lemma, there exists a constant C_t not depending on K such that

$$E\big(\sup_{s\in[0,t]} \langle X_s^K, 1\rangle^2\big) \le C_t \left(1 + E\left(\langle M^{K,1}\rangle_t \right) \right).$$

Using now (4.10), we obtain, for some other constant C_t not depending on K,

$$E\left(\langle M^{K,1}\rangle_t \right) \le C \int_0^t \left(E\left(\langle X_s^K, 1\rangle + \langle X_s^K, 1\rangle^2 \right) \right) ds \le C_t$$

thanks to (4.24). This concludes the proof of (4.25).

Step 3   We first endow M_F with the vague topology, the extension to the weak topology being handled in Step 6 below. To show the tightness of the sequence of laws $Q^K = \mathcal{L}(X^K)$ in $\mathcal{P}(\mathbb{D}([0,T], M_F))$, it suffices, following Roelly [23], to show that for any continuous bounded function f on $\mathbb{R}^d$, the sequence of laws of the processes $\langle X^K, f\rangle$ is tight in $\mathbb{D}([0,T], \mathbb{R})$. To this end, we use the Aldous criterion [1] and the Rebolledo criterion (see [17]). We have to show that

$$\sup_K E\big(\sup_{t\in[0,T]} |\langle X_t^K, f\rangle|\big) < \infty, \qquad (4.26)$$

and the tightness respectively of the laws of the predictable quadratic variation of the martingale part and of the drift part of the semimartingales $\langle X^K, f\rangle$.
Since f is bounded, (4.26) is a consequence of (4.25): let us thus consider a couple (S, S') of stopping times satisfying a.s. $0 \le S \le S' \le S + \delta \le T$. Using (4.10) and (4.25), we get for constants C, C'

$$E\left(\langle M^{K,f}\rangle_{S'} - \langle M^{K,f}\rangle_S \right) \le CE\left(\int_S^{S+\delta} \left(\langle X_s^K, 1\rangle + \langle X_s^K, 1\rangle^2 \right) ds \right) \le C'\delta.$$

In a similar way, the expectation of the finite variation part of $\langle X_{S'}^K, f\rangle - \langle X_S^K, f\rangle$ is bounded by $C'\delta$.
    Hence, the sequence $Q^K = \mathcal{L}(X^K)$ is tight.

22

Step 4 Let us now denote by Q the limiting law of a subsequence of Q^K. We still denote this subsequence by Q^K. Let $X = (X_t)_{t \geq 0}$ a process with law Q. We remark that by construction, almost surely,

$$\sup_{t \in [0,T]} \sup_{f \in L^\infty(\mathbb{R}^d), \|f\|_\infty \leq 1} |\langle X_t^K, f \rangle - \langle X_{t-}^K, f \rangle| \leq 1/K.$$

This implies that the process X is a.s. strongly continuous.

Step 5 The time $T > 0$ is fixed. Let us now check that almost surely, the process X is the unique solution of (4.13). Thanks to (4.25), it satisfies $\sup_{t \in [0,T]} \langle X_t, 1 \rangle < +\infty$ a.s., for each T. We fix now a function $f \in C_b^3(\mathbb{R}^d)$ (the extension of (4.13) to any function f in C_b^2 is not hard) and some $t \leq T$.
For $\nu \in C([0,T], M_F)$, denote by

$$\Psi_t^1(\nu) = \langle \nu_t, f \rangle - \langle \nu_0, f \rangle - \int_0^t \int_{\mathbb{R}^d} (b(x, V * \nu_s(x)) - d(x, U * \nu_s(x))) f(x) \nu_s(dx) ds,$$

$$\Psi_t^2(\nu) = -\int_0^t \int_{\mathbb{R}^d} \frac{1}{2} \mu(x) r(x) \sum_{i,j} \Sigma_{ij}(x) \partial_{ij}^2 f(x) \nu_s(dx) ds. \tag{4.27}$$

We have to show that

$$E_Q \left(|\Psi_t^1(X) + \Psi_t^2(X)| \right) = 0. \tag{4.28}$$

By (4.9), we know that for each K,

$$M_t^{K,f} = \Psi_t^1(X^K) + \Psi_t^{2,K}(X^K),$$

where

$$\Psi_t^{2,K}(X^K) = -\int_0^t \int_{\mathbb{R}^d} \mu(x)(K^\eta r(x) + b(x, V * X_s^K(x)))$$

$$\left(\int_{\mathbb{R}^d} f(z) M_K(x,z) dz - f(x) \right) X_s^K(dx) ds. \tag{4.29}$$

Moreover, (4.25) implies that for each K,

$$E \left(|M_t^{K,f}|^2 \right) = E \left(\langle M^{K,f} \rangle_t \right) \leq \frac{C_f K^\eta}{K} E \left(\int_0^t \{ \langle X_s^K, 1 \rangle + \langle X_s^K, 1 \rangle^2 \} ds \right) \leq \frac{C_{f,T} K^\eta}{K}, \tag{4.30}$$

which goes to 0 as K tends to infinity, since $0 < \eta < 1$. Therefore,

$$\lim_K E(|\Psi_t^1(X^K) + \Psi_t^{2,K}(X^K)|) = 0.$$

Since X is a.s. strongly continuous, since $f \in C_b^3(\mathbb{R}^d)$ and thanks to the continuity of the parameters, the functions Ψ_t^1 and Ψ_t^2 are a.s. continuous at X. Furthermore, for any $\nu \in \mathbb{D}([0,T], M_F)$,

$$|\Psi_t^1(\nu) + \Psi_t^2(\nu)| \leq C_{f,T} \sup_{s \in [0,T]} \left(1 + \langle \nu_s, 1 \rangle^2 \right). \tag{4.31}$$

23

Hence using (4.24), we see that the sequence $(\Psi^1_t(X^K) + \Psi^2_t(X^K))_K$ is uniformly integrable, and thus

$$\lim_K E\left(|\Psi^1_t(X^K) + \Psi^2_t(X^K)|\right) = E\left(|\Psi^1_t(X) + \Psi^2_t(X)|\right). \tag{4.32}$$

We have now to deal with $\Psi^{2,K}_t(X^K) - \Psi^2_t(X^K)$. The convergence of this term is due to the fact that the measure $M_K(x,z)dz$ has mean x, variance $\Sigma(x)/K^\eta$, and third moment bounded by $C/K^{\eta+\varepsilon}$ ($\varepsilon > 0$) uniformly in x. Indeed, if $Hf(x)$ denotes the Hessian matrix of f at x,

$$\int_{\mathbb{R}^d} f(z) M_K(x,z)dz$$
$$= \int_{\mathbb{R}^d} \left(f(x) + (z-x) \cdot \nabla f(x) + \frac{1}{2}(z-x)^* Hf(x)(z-x) + O((z-x)^3) \right) M_K(x,z)dz$$
$$= f(x) + \frac{1}{2} \sum_{i,j} \frac{\Sigma_{ij}(x)}{K^\eta} \partial^2_{ij} f(x) + o(\frac{1}{K^\eta}). \tag{4.33}$$

where $K^\eta o(\frac{1}{K^\eta})$ tends to 0 uniformly in x (since f is in C^3_b), as K tends to infinity. Then,

$$\Psi^{2,K}_t(X^K) = -\int_0^t \int_{\mathbb{R}^d} \mu(x)(K^\eta r(x) + b(x, V * X^K_s(x))) \times$$
$$\times \left(\frac{1}{2} \sum_{i,j} \frac{\Sigma_{ij}(x)}{K^\eta} \partial^2_{ij} f(x) + o(\frac{1}{K^\eta}) \right) X^K_s(dx)ds,$$

and

$$|\Psi^{2,K}_t(X^K) - \Psi^2_t(X^K)| \leq C_f \left(\sup_{s \leq T} < X^K_s, 1 > \right) \left(\frac{1}{K^\eta} + K^\eta o(\frac{1}{K^\eta}) \right).$$

Using (4.25), we conclude the proof of (4.28).

Step 6   The previous steps imply that $(X^K)_K$ converges to ξ in $\mathbb{D}([0,T], M_F)$, where M_F is endowed with the vague topology. To extend the result to the case where M_F is endowed with the weak topology, we use a criterion proved in Méléard and Roelly [20]: since the limiting process is continuous, it suffices to prove that the sequence $(\langle X^K, 1 \rangle)$ converges to $\langle \xi, 1 \rangle$ in law, in $\mathbb{D}([0,T], \mathbb{R})$. One may of course apply Step 5 with $f \equiv 1$, which concludes the proof of (1).

(2) Let us now assume the non-degeneracy property $r(x)\mu(x)s^* \Sigma(x)s \geq c\|s\|^2 > 0$ for each $x \in \mathbb{R}^d, s \in \mathbb{R}^d$. That implies that for each time $t > 0$, the transition semigroup $P_t(x, dy)$ introduced in Step 1 of this proof has for each x a density function $p_t(x,y)$ with respect to the Lebesgue measure. Then if we come back to the evolution equation (4.20), we can write

$$\int_{\mathbb{R}^d} f(x)\xi_t(dx) = \int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} f(y)p_t(x,y)dy \right) \xi_0(dx)$$
$$+ \int_0^t \int_{\mathbb{R}^d} (b(x, V * \xi_s(x)) - d(x, U * \xi_s(x))) \left(\int_{\mathbb{R}^d} f(y)p_{t-s}(x,y)dy \right) \xi_s(dx)ds.$$

Using the fact that the parameters are bounded, that $\sup_{t \le T} \langle \xi_t, 1 \rangle < +\infty$ and that f is bounded, we can apply Fubini's theorem and deduce that

$$\int_{\mathbb{R}^d} f(x) \xi_t(dx) = \int_{\mathbb{R}^d} H_t(y) f(y) dy$$

with $H \in L^\infty([0,T], L^1(\mathbb{R}^d))$, which implies that ξ_t has a density with respect to the Lebesgue measure for each time $t \le T$.

Equation (4.14) is then the dual form of (4.13). $\qquad\square$

Proof of Theorem 4.5 We will use a similar method as the one of the previous theorem. Steps 2, 3, 4 and 6 of this proof can be achieved exactly in the same way. Therefore, we only have to prove the uniqueness (in law) of the solution to the martingale problem (4.16)–(4.18) (Step 1), and that any accumulation point of the sequence of laws of X^K is solution to (4.16)–(4.18) (Step 5).

Step 1 This uniqueness result is well-known for the super-Brownian process (defined by a similar martingale problem, but with $b = d = 0$, $r = \mu = 1$ and $\Sigma = \mathrm{Id}$, cf. [23]). Following [12], we may use the version of Dawson's Girsanov transform obtained in Evans and Perkins [14] (Theorem 2.3), to deduce the uniqueness in our situation, provided the condition

$$E \left(\int_0^t \int_{\mathbb{R}^d} [b(x, V * X_s(x)) - d(x, U * X_s(x))]^2 X_s(dx) ds \right) < +\infty$$

is satisfied. This is easily obtained from the assumption that $\sup_{t \in [0,T]} E[\langle X_t, 1 \rangle^3] < \infty$ since the coefficients are bounded.

Step 5 Let us identify the limit. Let us call $Q^K = \mathcal{L}(X^K)$ and denote by Q a limiting value of the tight sequence Q^K, and by $X = (X_t)_{t \ge 0}$ a process with law Q. Because of Step 4, X belongs a.s. to $C([0,T], M_F)$. We have to show that X satisfies the conditions (4.16), (4.17) and (4.18). First note that (4.16) is straightforward from (4.25). Then, we show that for any function f in $C_b^3(\mathbb{R}^d)$, the process $\bar{M}_t^f$ defined by (4.17) is a martingale (the extension to every function in C_b^2 is not hard). We consider $0 \le s_1 \le ... \le s_n < s < t$, some continuous bounded maps $\phi_1, ... \phi_n$ on M_F, and our aim is to prove that, if the function Ψ from $\mathbb{D}([0,T], M_F)$ into $\mathbb{R}$ is defined by

$$\Psi(\nu) = \phi_1(\nu_{s_1})...\phi_n(\nu_{s_n}) \Big\{ \langle \nu_t, f \rangle - \langle \nu_s, f \rangle$$

$$- \int_s^t \int_{\mathbb{R}^d} \left(\frac{1}{2} \mu(x) r(x) \sum_{i,j} \Sigma_{ij} \partial_{ij}^2 f(x) + f(x) \left[b(x, V * \nu_u(x)) - d(x, U * \nu_u(x)) \right] \right) \nu_u(dx) du \Big\},$$

$$(4.34)$$

then

$$E\left(\Psi(X) \right) = 0. \qquad (4.35)$$

It follows from (4.9) that

$$0 = E \left(\phi_1(X_{s_1}^K)...\phi_n(X_{s_n}^K) \left\{ M_t^{K,f} - M_s^{K,f} \right\} \right) = E\left(\Psi(X^K) \right) - A_K, \qquad (4.36)$$

where A_K is defined by

$$A_K = E\Big(\phi_1(X_{s_1}^K)...\phi_n(X_{s_n}^K) \int_s^t \int_{\mathbb{R}^d} \mu(x)\Big\{b(x, V * X_u^K(x))\Big[\int_{\mathbb{R}^d} (f(z) - f(x))M_K(x,z)dz\Big]$$

$$+ r(x)K\Big[\int_{\mathbb{R}^d} (f(z) - f(x) - \sum_{i,j} \frac{\Sigma_{ij}(x)}{2K}\partial_{ij}^2 f(x))M_K(x,z)dz\Big]\Big\}X_u^K(dx)du\Big).$$

It turns out from (4.33) that A_K tends to zero as K grows to infinity, and using (4.25), that the sequence $(|\Psi(X^K)|)_K$ is uniformly integrable, so

$$\lim_K E\left(|\Psi(X^K)|\right) = E_Q\left(|\Psi(X)|\right). \tag{4.37}$$

Collecting the previous results allows us to conclude that (4.35) holds, and thus $\bar{M}^f$ is a martingale.
We finally have to show that the bracket of $\bar{M}^f$ is given by (4.18). To this end, we first check that

$$\bar{N}_t^f = \langle X_t, f\rangle^2 - \langle X_0, f\rangle^2 - \int_0^t \int_{\mathbb{R}^d} 2r(x)f^2(x)X_s(dx)ds$$

$$- 2\int_0^t \langle X_s, f\rangle \int_{\mathbb{R}^d} f(x)\left[b(x, V * X_s(x)) - d(x, U * X_s(x))\right]X_s(dx)ds$$

$$- \int_0^t \langle X_s, f\rangle \int_{\mathbb{R}^d} \mu(x)r(x)\sum_{i,j}\Sigma_{ij}(x)\partial_{ij}^2 f(x)X_s(dx)ds \tag{4.38}$$

is a martingale. This can be done exactly as for $\bar{M}_t^f$, using the semimartingale decomposition of $\langle X_t^K, f\rangle^2$, given by (4.8) with $\phi(\nu) = \langle \nu, f\rangle^2$. In another hand, Itô's formula implies that

$$\langle X_t, f\rangle^2 - \langle X_0, f\rangle^2 - \langle \bar{M}^f\rangle_t - \int_0^t \langle X_s, f\rangle \int_{\mathbb{R}^d} r(x)\mu(x)\sum_{i,j}\Sigma_{ij}(x)\partial_{ij}^2 f(x)X_s(dx)ds$$

$$- 2\int_0^t \langle X_s, f\rangle \int_{\mathbb{R}^d} f(x)\left[b(x, V * X_s(x)) - d(x, U * X_s(x))\right]X_s(dx)ds$$

is a martingale. Comparing this formula with (4.38), we obtain (4.18). $\qquad\square$

4.2.2 Rare mutations

In this case, the mutation step density M is fixed and the mutation rate is decelerated proportionally to $1/K^\eta$:

Assumption (H4):

$$M_K = M, \quad \mu_K = \frac{\mu}{K^\eta}.$$

Thus only births without mutation are accelerated.

As in Section 4.2.1, we obtain deterministic or random limits, according to the value of $\eta \in (0, 1]$.

26

Theorem 4.7 (1) *Assume (H), (H1), (H2), (H4) and $0 < \eta < 1$. Assume also that the initial conditions X_0^K converge in law and for the weak topology on M_F as K increases, to a finite deterministic measure ξ_0, and that $\sup_K E(\langle X_0^K, 1 \rangle^3) < +\infty$.*

Then, for each $T > 0$, the sequence of processes (X^K) belonging to $\mathbb{D}([0,T], M_F)$ converges (in law) to the unique deterministic function $(\xi_t)_{t \geq 0} \in C([0,T], M_F)$ weak solution of the deterministic nonlinear integro-differential equation:

$$\partial_t \xi_t(x) = [b(x, V * \xi_t(x)) - d(x, U * \xi_t(x))]\xi_t(x) + \int_{\mathbb{R}^d} M(y,x)\mu(y)r(y)\xi_t(y)dy$$
$$- \mu(x)r(x)\xi_t(x). \quad (4.39)$$

(2) *Assume now $\eta = 1$ and that X_0^K converge in law to X_0. Then, for each $T > 0$, the sequence of processes (X^K) converges in law in $\mathbb{D}([0,T], M_F)$ to the unique (in law) continuous superprocess $X \in C([0,T], M_F)$, defined by the following conditions:*

$$\sup_{t \in [0,T]} E\left(\langle X_t, 1 \rangle^3\right) < \infty,$$

and for any $f \in C_b^2(\mathbb{R}^d)$,

$$\bar{M}_t^f = \langle X_t, f \rangle - \langle X_0, f \rangle - \int_0^t \int_{\mathbb{R}^d} \mu(x)r(x) \int_{\mathbb{R}^d} M(x,z)(f(z) - f(x))dz X_s(dx)ds$$
$$- \int_0^t \int_{\mathbb{R}^d} f(x)\left(b(x, V * X_s(x)) - d(x, U * X_s(x))\right) X_s(dx)ds$$

is a continuous martingale with quadratic variation

$$\langle \bar{M}^f \rangle_t = 2 \int_0^t \int_{\mathbb{R}^d} r(x)f^2(x)X_s(dx)ds.$$

In a SPDE formalism, one can write the last limit as formal solution of the equation

$$\partial_t X_t(x) = [b(x, V * X_t(x)) - d(x, U * X_t(x))]X_t(x) + \int_{\mathbb{R}^d} M(y,x)\mu(y)r(y)X_t(dy) + \dot{M}$$
$$- \mu(x)r(x)X_t(x), \quad (4.40)$$

where $\dot{M}$ is a random fluctuation term.

The proof of Theorem 4.7 is similar to proofs of Theorems 4.3 and 4.5 and we leave it to the reader. Theorem 4.7 (1) is illustrated in the simulation of Fig. 2 (b).

5   Rare mutation renormalization of the monomorphic process and adaptive dynamics

In the previous section, Eqs. (4.39) and (4.40) have been obtained at the population growth time scale (ecological time scale), under an assumption of rare mutation. Here, we are interested in the behavior of the population process at the evolutionary time scale, when mutations are extremely rare, as illustrated by the simulation of Fig. 1 (d). We hence

recover rigorously the stochastic "trait substitution sequence" jump process of adaptive dynamics (Metz et al. [22]) when the initial condition is monomorphic. The biological idea behind such a scaling of the population process is that selection has sufficient time between two mutations to eliminate all disadvantaged traits, so that the population remains monomorphic on the evolutionary timescale. Then the evolution proceeds by successive invasions of mutant traits, replacing the resident trait from which the mutant trait is born. These invasions occur on an infinitesimal timescale with respect to the mutation timescale. Our result emphasizes how the mutation scaling should compare to the system size (K) in order to obtain the correct time scale separation between the "mutant-invasions" (taking place on a short time scale) and the mutations (evolutionary time scale).

5.1   Statement of the result

We consider here a limit of rare mutations combined with the large population limit of Section 4.1 (Assumption (H1) and $b_K = b$, $d_K = d$ and $M_K = M$). We assume
   Assumptions (H5):
   (i) $\mu_K(x) = u_K \mu(x)$.
   (ii) For any constant $C > 0$,

$$e^{-CK} \ll u_K \ll \frac{1}{K \log K} \tag{5.1}$$

(thus $u_K \to 0$ when $K \to +\infty$), or, equivalently, for any C and $t > 0$,

$$\log K \ll \frac{t}{K u_K} \ll e^{CK}. \tag{5.2}$$

   (iii) For any $x \in \mathcal{X}$, $\zeta \mapsto b(x, \zeta)$ and $\zeta \mapsto d(x, \zeta)$ are positive functions, non-increasing and increasing respectively, satisfying

$$\forall x \in \mathcal{X}, \ b(x, 0) - d(x, 0) > 0,$$
$$\lim_{\zeta \to +\infty} \inf_{x \in \mathcal{X}} d(x, \zeta) = +\infty. \tag{5.3}$$

   (iv) There exists a constant $\underline{U} > 0$ such that $U(h) \geq \underline{U}$ for any $h \in \mathbb{R}^d$.

   Assumption (H5)-(i) entails the rare mutation asymptotic, and (H5)-(ii) gives the correct scaling between the mutation probability and the system size in order to obtain the correct time scale separation. Observe that (H5)-(ii) implies that $K u_K \to 0$ when $K \to +\infty$, so that the timescale $t/K u_K$, which corresponds to the timescale of mutations (the population size is proportional to K, and each birth event produces a mutant with a probability proportional to u_K, which gives a total mutation rate in the population proportional to $K u_K$) is a long timescale. Our result gives the behavior of the population process on this long timescale.

   Assumptions (H5)-(iii) and (iv) will allow us to bound the population size on the mutation timescale, and to study the behavior of the population when it is monomorphic or dimorphic between two (rare) mutation events. Specifically, the monotonicity properties of b and d in Assumption (H5)-(iii) ensures, for any $x \in \mathcal{X}$, the existence of a unique non-trivial stable equilibrium $\bar{n}(x)$ for the monomorphic logistic equation (4.6) of Example 3

in Section 4.1. Moreover, since $b(x, V(0)u) - d(x, U(0)u) > 0$ for any $u < \bar{n}(x)$ and $b(x, V(0)u) - d(x, U(0)u) < 0$ for any $u > \bar{n}(x)$, any solution to (4.6) with positive initial condition converges to $\bar{n}(x)$.

Concerning the dimorphic logistic equations (4.7), an elementary linear analysis of the equilibrium $(\bar{n}(x), 0)$ gives that it is stable if $f(y, x) < 0$ and unstable if $f(y, x) > 0$, where the function

$$f(y, x) = b(y, V(y - x)\bar{n}(x)) - d(y, U(y - x)\bar{n}(x)) \tag{5.4}$$

is known as the "fitness function" ([21, 22]), which gives a measure of the selective advantage of a mutant individual with trait y in a monomorphic population of trait x at equilibrium. Similarly, the stability of the equilibrium $(0, \bar{n}(y))$ is governed by the sign of $f(x, y)$.

In order to ensure that, when the invasion of a mutant trait is possible, then this invasion will end with the extinction of the resident trait, we will need the following additional assumption:

Assumptions (H6):

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

(i) either $f(y, x) < 0$ (so that $(\bar{n}(x), 0)$ is stable),

(ii) or $f(y, x) > 0$, $f(x, y) < 0$ and any solution to (4.7) with initial condition with positive coordinates in a given neighborhood of $(\bar{n}(x), 0)$ converges to $(0, \bar{n}(y))$.

In the case of linear logistic density-dependence introduced in Section 2.2 ($b(x, \zeta) = b(x)$ and $d(x, \zeta) = d(x) + \alpha(x)\zeta$), the equilibrium monomorphic density $\bar{n}(x)$ writes $(b(x) - d(x))/\alpha(x)U(0)$ and the condition (H6)-(ii) is actually equivalent to $f(y, x) > 0$ and $f(x, y) < 0$ (see [6]).

Our convergence result writes

Theorem 5.1 *Assume (H), (H1), (H5) and (H6). Given $x \in \mathcal{X}$, $\gamma > 0$ and a sequence of $\mathbb{N}$-valued random variables $(\gamma_K)_{K \in \mathbb{N}}$, such that γ_K/K is bounded in $\mathbb{L}^1$ and converges in law to γ, consider the process $(X_t^K, t \geq 0)$ of Section 4 generated by (4.1) with initial state $\frac{\gamma_K}{K}\delta_x$. Then, for any $n \geq 1$, $\varepsilon > 0$ and $0 < t_1 < t_2 < \ldots < t_n < \infty$, and for any measurable subsets $\Gamma_1, \ldots, \Gamma_n$ of $\mathcal{X}$,*

$$\lim_{K \to +\infty} P\big(\forall i \in \{1, \ldots, n\}, \ \exists x_i \in \Gamma_i \ : \ \mathrm{Supp}(X_{t_i/Ku_K}^K) = \{x_i\}$$

$$\text{and } |\langle X_{t_i/Ku_K}^K, \mathbf{1}\rangle - \bar{n}(x_i)| < \varepsilon\big) = P(\forall i \in \{1, \ldots, n\}, \ Y_{t_i} \in \Gamma_i) \tag{5.5}$$

where for any $\nu \in M_F(\mathcal{X})$, $\mathrm{Supp}(\nu)$ is the support of ν and $(Y_t, t \geq 0)$ is a Markov jump process with initial state x generated by

$$A\varphi(x) = \int_{\mathbb{R}^d} (\varphi(y) - \varphi(x))g(y, x)M(x, y)dy \tag{5.6}$$

where

$$g(y, x) = \mu(x)b(x, V(0)\bar{n}(x))\bar{n}(x)\frac{[f(y, x)]_+}{b(y, V(y - x)\bar{n}(x))} \tag{5.7}$$

and $[\cdot]_+$ denotes the positive part.

Corollary 5.2 *With the same notations and assumptions as in Theorem 5.1, assuming moreover that γ_K/K is bounded in $\mathbb{L}^q$ for some $q > 1$, the process $(X^K_{t/Ku_K}, t \geq 0)$ converges when $K \to +\infty$, in the sense of the finite dimensional distributions for the topology on $M_F(\mathcal{X})$ induced by the functions $\nu \mapsto \langle \nu, f \rangle$ with f bounded and measurable on $\mathcal{X}$, to the process $(Z_t, t \geq 0)$ defined by*

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

This corollary follows from the following long time moment estimates.

Lemma 5.3 *Under (H), (H1), (H5)(iii) (5.3) and (iv), and if $\sup_{K \geq 1} E(\langle X^K_0, \mathbf{1}\rangle^q) < +\infty$ for some $q \geq 1$, then*

$$\sup_{K \geq 1} \sup_{t \geq 0} E\big(\langle X^K_t, \mathbf{1}\rangle^q\big) < +\infty,$$

and therefore, if $q > 1$, the family of random variables $\{\langle X^K_t, \mathbf{1}\rangle\}_{\{K \geq 1,\, t \geq 0\}}$ is uniformly integrable.

Proof of Lemma 5.3   Observe that, if we replace $b(x, V * \nu)$ by $\bar{b}$ and $d(x, U * \nu)$ by $g(U\langle\nu, \mathbf{1}\rangle)$ where $g(\zeta) := \inf_{x \in \mathcal{X}} d(x, \zeta)$ in the indicator functions of each terms of the construction (2.3) of the process X^K_t, we can stochastically dominate the population size $\langle X^K_t, \mathbf{1}\rangle$ by a birth and death Markov process $(Z^K_t)_{t \geq 0}$ with initial state $Z^K_0 = \langle X^K_0, \mathbf{1}\rangle$ and transition rates

$$\begin{array}{ll} i\bar{b} & \text{from } i/K \text{ to } (i+1)/K, \\ ig(\underline{U}\frac{i}{K}) & \text{from } i/K \text{ to } (i-1)/K. \end{array}$$

Therefore, it suffices to prove that $\sup_{K \geq 0} \sup_{t \geq 0} E((Z^K_t)^q) < +\infty$.

Let us define $p^k_t = P(Z^K_t = k/K)$. Then

$$\frac{d}{dt}E((Z^K_t)^q) = \sum_{k \geq 1} \left(\frac{k}{K}\right)^q \frac{dp^k_t}{dt}$$

$$= \frac{1}{K^q}\sum_{k \geq 1} k^q \left[\bar{b}(k-1)p^{k-1}_t + (k+1)g\left(\underline{U}\frac{k+1}{K}\right)p^{k+1}_t - k\left(\bar{b} + g\left(\underline{U}\frac{k}{K}\right)\right)p^k_t\right]$$

$$= \frac{1}{K^q}\sum_{k \geq 1}\left[\bar{b}\left(\left(1+\frac{1}{k}\right)^q - 1\right) + g\left(\underline{U}\frac{k}{K}\right)\left(\left(1-\frac{1}{k}\right)^q - 1\right)\right]k^{q+1}p^k_t.$$

Now, by (H5) (iii) (5.3), $g(\alpha) \to +\infty$ when $\alpha \to +\infty$, so there exists α_0 such that, for any $\alpha \geq \alpha_0$, $g(\underline{U}\alpha) \geq 2\bar{b}$. Therefore, for $k \geq K\alpha_0$, $\bar{b}((1+1/k)^q - 1) + g(\underline{U}k/K)((1-1/k)^q - 1) \leq -\bar{b}[3 - 2(1-1/k)^q - (1+1/k)^q]$, the RHS term being equivalent to $-\bar{b}q/k$. Therefore, enlarging α_0 if necessary and using in the first inequality the facts that $(1+\alpha)^q - 1 \leq \alpha(2^q - 1)$ and $(1-\alpha)^q - 1 \leq 0$ for any $\alpha \in [0,1]$, we can write

$$\frac{d}{dt}E((Z^K_t)^q) \leq \sum_{k=1}^{\lceil K\alpha_0\rceil - 1} \bar{b}(2^q - 1)\left(\frac{k}{K}\right)^q p^k_t - \sum_{k \geq \lceil K\alpha_0\rceil}\frac{\bar{b}q}{2}\left(\frac{k}{K}\right)^q p^k_t$$

$$\leq \sum_{k=1}^{\lceil K\alpha_0\rceil - 1}\bar{b}(q/2 + 2^q - 1)\alpha^q_0\, p^k_t - \frac{\bar{b}q}{2}E((Z^K_t)^q) \leq \frac{\bar{b}q}{2}[C - E((Z^K_t)^q)],$$

30

where $C = (1 + 2(2^q - 1)/q)\alpha_0^q$. This differential inequality solves as

$$E((Z_t^K)^q) \leq C + [E((Z_0^K)^q) - C]e^{-\bar{b}qt/2},$$

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

Proof of Corollary 5.2 Let Γ be a measurable subset of $\mathcal{X}$. Let us prove that

$$\lim_{K \to +\infty} E\big[\langle X_{t/Ku_K}^K, \mathbf{1}_\Gamma \rangle\big] = E\big[\bar{n}(Y_t)\mathbf{1}_{Y_t \in \Gamma}\big]. \tag{5.8}$$

By (H5)-(iii)-(5.3), there exists $\zeta_0 > 0$ such that for any $\zeta > \zeta_0$ and $x \in \mathcal{X}$, $d(x, \zeta) > \bar{b}$. Therefore, by (H5)-(iv), for any $x \in \mathcal{X}$, $\bar{n}(x) \in [0, \zeta_0/\underline{U}]$. Fix $\varepsilon > 0$, and write $[0, \zeta_0/\underline{U}] \subset \cup_{i=1}^p I_i$, where p is the integer part of $\zeta_0/(\underline{U}\varepsilon)$, and $I_i = [(i-1)\varepsilon, i\varepsilon[$. Define $\Gamma_i = \{x \in \mathcal{X} : \bar{n}(x) \in I_i\}$ for $1 \leq i \leq p$, and apply (5.5) to the sets $\Gamma \cap \Gamma_1, \ldots, \Gamma \cap \Gamma_p$ with $n = 1$, $t_1 = t$ and the constant ε above. Then, by Lemma 5.3, for some constant $C > 0$ and for sufficiently large K,

$$\begin{aligned}
\limsup_{K \to +\infty} E\big[\langle X_{t/Ku_K}^K, \mathbf{1}_\Gamma \rangle\big] &\leq \limsup_{K \to +\infty} E\big[\langle X_{t/Ku_K}^K, \mathbf{1}_\Gamma \rangle \mathbf{1}_{\langle X_{t/Ku_K}^K, \mathbf{1} \rangle \leq C}\big] + \varepsilon \\
&\leq \sum_{i=1}^p \limsup_{K \to +\infty} E\big[\langle X_{t/Ku_K}^K, \mathbf{1}_{\Gamma \cap \Gamma_i} \rangle \mathbf{1}_{\langle X_{t/Ku_K}^K, \mathbf{1} \rangle \leq C}\big] + \varepsilon \\
&\leq \sum_{i=1}^p (i+1)\varepsilon P(Y_t \in \Gamma \cap \Gamma_i) + \varepsilon \\
&\leq \sum_{i=1}^p \big(E\big[\bar{n}(Y_t)\mathbf{1}_{Y_t \in \Gamma \cap \Gamma_i}\big] + 2\varepsilon P(Y_t \in \Gamma_i)\big) + \varepsilon \\
&\leq E\big[\bar{n}(Y_t)\mathbf{1}_{Y_t \in \Gamma}\big] + 3\varepsilon.
\end{aligned}$$

A similar estimate for the *lim inf* ends the proof of (5.8), which implies the convergence of one-dimensional laws for the required topology.

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

Observe that the fact that the limit process is not right-continuous prevents the possibility to obtain a convergence for the Skorohod topology on $\mathbb{D}([0, T], M_F(\mathcal{X}))$.

5.2 Idea of the proof

Theorem 5.1 can be proved in a similar way as in Champagnat [6]. Let us give an idea of the method in order to explain the assumptions, the various parameters appearing in Theorem 5.1 and the tools involved in the proof. It is based on two ingredients: the study of a monomorphic population before the first mutation, and the study of the invasion of a single mutant individual in this population.

1) The first part obtains from large deviation results for the convergence of X_t^K to $n_t(x)\delta_x$ when the initial population is monomorphic with trait x, where $n_t(x)$ satisfies (4.6).

Any positive solution to (4.6) converges to $\bar{n}(x)$ when $t \to +\infty$, and hence reaches a given neighborhood of $\bar{n}(x)$ in finite time, i.e. on an infinitesimal time scale with respect to the mutation time scale. Large deviations theory allows us to show that the exit time of $\langle X_t^K, \mathbf{1} \rangle$ from this neighborhood behaves as $\exp(KC)$ for some $C > 0$ (problem of exit from a domain, Freidlin and Wentzell [16]). Thanks to the right part of Assumption (5.2), we can prove that, with high probability, $\langle X_t^K, \mathbf{1} \rangle$ is close to $\bar{n}(x)$ when the first mutation occurs. Therefore, the total mutation rate is close to $u_K \mu(x) K \bar{n}(x) b(x, V(0)\bar{n}(x))$ and so, on the mutation time scale t/Ku_K, the rate of mutation is close to $\bar{n}(x)\mu(x)b(x, V(0)\bar{n}(x))$, which explain the left part of the RHS of (5.7). This argument can be made rigorous using stochastic domination results similar to the one used at the beginning of the proof of Lemma 5.3, and leads to the following result:

Lemma 5.4 *Let τ_1 denote the first mutation time and $\mathbf{P}_{X_0^K}^K$ the law of X^K with initial state X_0^K. Given $x \in \mathcal{X}$ and a sequence of integers $(z_K)_{K \geq 1}$ such that $z_K/K \to z > 0$,*

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

$$\lim_{K \to +\infty} \mathbf{P}_{\frac{z_K}{K}\delta_x}^K \left(\tau_1 > \log K, \ \sup_{t \in [\log K, \tau_1]} |\langle X_t^K, \mathbf{1} \rangle - \bar{n}(x)| > \varepsilon \right) = 0 \qquad (5.9)$$

and

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

In particular, under $\mathbf{P}_{\frac{z_K}{K}\delta_x}^K$, $X_{\log K}^K \to \bar{n}(x)\delta_x$ and $X_{\tau_1-}^K \to \bar{n}(x)\delta_x$ in probability.

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

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

where $\beta(x) = \mu(x)\bar{n}(x)b(x, V(0)\bar{n}(x))$.

2) The study of the invasion of a mutant individual with trait y can be divided in three steps represented in Fig. 3.

Firstly, the invasion of the mutant (between 0 and t_1 in Fig. 3) can be defined as the growth of the mutant density $\langle X_t^K, \mathbf{1}_{\{y\}} \rangle$ from $1/K$ (one individual) to a fixed small level ε (εK individuals). As long as the mutant density is small, the dynamics of the resident density $\langle X_t^K, \mathbf{1}_{\{x\}} \rangle$ is close to the one it followed before the mutation, so it is close to $\bar{n}(x)$ with high probability. Therefore, between 0 and t_1, the birth and death rates of an individual with trait y are close to $b(y, V(y-x)\bar{n}(x))$ and $d(y, U(y-x)\bar{n}(x))$ respectively. Therefore, the number of mutant individuals is close to a binary branching process with the parameters above. When $K \to +\infty$, the probability that such a branching process reaches level εK is close to its survival probability, which writes $[f(y,x)]_+/b(y, V(y-x)\bar{n}(x))$. This gives the second part of the RHS of (5.7).

Secondly, once the invasion succeeded (which is possible only if $f(y,x) > 0$), the dynamics of the densities of traits x and y are close to the solution to the dimorphic logistic equation (4.7) with initial state $(\bar{n}(x), \varepsilon)$, represented in dotted curves between t_1 and t_2 in Fig. 3. Because of Assumption (H6), the resident density can be proved to reach level ε with high probability (at time t_2 in Fig. 3).
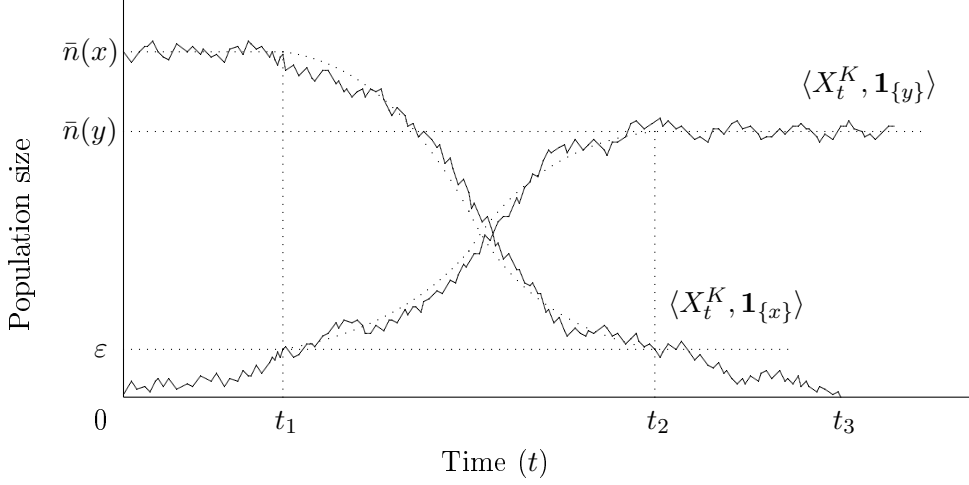
Figure 3: The three steps of the invasion and fixation of a mutant trait y in a monomorphic population with trait x. Plain curves represent the resident and mutant densities $\langle X_t^K, \mathbf{1}_{\{x\}} \rangle$ and $\langle X_t^K, \mathbf{1}_{\{y\}} \rangle$, respectively. Dotted curves represent the solution of Eq. (4.7) with initial state $n_0(x) = \bar{n}(x)$ and $n_0(y) = \varepsilon$.

Finally, a similar argument as in the first step above allows us to prove that the resident population density $\langle X_t^K, \mathbf{1}_{\{x\}} \rangle$ follows approximately a binary branching process with birth rate $b(y, V(x-y)\bar{n}(y))$ and death rate $d(y, U(x-y)\bar{n}(y))$. Since $f(x, y) < 0$ by Assumption (H6), this is a sub-critical branching process, and therefore, the resident trait x disappears in finite time t_3 with high probability.

We can show, using results on branching processes, that t_1 and $t_3 - t_2$ are of order $\log K$, whereas $t_2 - t_1$ depends only on ε. Therefore, the left part of (5.2) ensures that the three steps of the invasion are completed before the next mutation, with high probability. The previous heuristics can be made rigorous using further comparison results, and leads to the following result.

Lemma 5.5 *Assume that the initial population is made of individuals with traits x and y satisfying assumption (H6) (i) or (ii). Let θ_0 denote the first time when the population gets monomorphic, and V_0 the remaining trait. Let $(z_K)_{K \geq 1}$ be a sequence of integers such that $z_K/K \to \bar{n}(x)$. Then,*

$$\lim_{K \to +\infty} \mathbf{P}^K_{\frac{z_K}{K}\delta_x + \frac{1}{K}\delta_y}(V_0 = y) = \frac{[f(y,x)]_+}{b(y, V(y-x)\bar{n}(x))}, \tag{5.10}$$

$$\lim_{K \to +\infty} \mathbf{P}^K_{\frac{z_K}{K}\delta_x + \frac{1}{K}\delta_y}(V_0 = x) = 1 - \frac{[f(y,x)]_+}{b(y, V(y-x)\bar{n}(x))}, \tag{5.11}$$

$$\forall \eta > 0, \quad \lim_{K \to +\infty} \mathbf{P}^K_{\frac{z_K}{K}\delta_x + \frac{1}{K}\delta_y}\left(\theta_0 > \frac{\eta}{Ku_K} \wedge \tau_1\right) = 0 \tag{5.12}$$

$$and \quad \forall \varepsilon > 0, \quad \lim_{K \to +\infty} \mathbf{P}^K_{\frac{z_K}{K}\delta_x + \frac{1}{K}\delta_y}\left(|\langle X_{\theta_0}^K, \mathbf{1}\rangle - \bar{n}(V_0)| < \varepsilon\right) = 1, \tag{5.13}$$

where $f(y,x)$ has been defined in (5.4).

33

Once these lemmas are proved, the proof can be completed by observing that the generator A of the process $(Y_t, t \geq 0)$ of Theorem 5.1 can be written as

$$A\varphi(x) = \int_{\mathbb{R}^l} (\varphi(y) - \varphi(x))\beta(x)\kappa(x, dy), \tag{5.14}$$

where $\beta(x)$ has been defined in Lemma 5.4 and the probability measure $\kappa(x, dh)$ is defined by

$$\kappa(x, dy) = \left(1 - \int_{\mathbb{R}^l} \frac{[f(z, x)]_+}{b(z, V(z-x)\bar{n}(x))} M(x, z)dz\right)\delta_x(dy)$$
$$+ \frac{[f(y, x)]_+}{b(y, V(y-x)\bar{n}(x))} M(x, y)dy. \tag{5.15}$$

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

$$Y_t := M\left(N\left(\int_0^t \beta(Y_s)ds\right)\right)$$

is a Markov process with infinitesimal generator (5.14) (cf. [13] chapter 6).

Let P_x denote its law, and define $(S_n)_{n\geq 1}$ by $T_n = \int_0^{S_n} \beta(Y_s)ds$. Observe that any jump of the process Y occurs at some time S_n, but that all S_n may not be effective jump times for Y, because of the Dirac mass at x appearing in (5.15).

Fix $t > 0$, $x \in \mathcal{X}$ and a measurable subset Γ of $\mathcal{X}$. Under P_x, S_1 and Y_{S_1} are independent, S_1 is an exponential random variable with parameter $\beta(x)$, and Y_{S_1} has law $\kappa(x, \cdot)$. Therefore, for any $n \geq 1$, the strong Markov property applied to Y at time S_1 yields

$$P_x(S_n \leq t < S_{n+1},\ Y_t \in \Gamma)$$
$$= \int_0^t \beta(x)e^{-\beta(x)s} \int_{\mathbb{R}^l} \mathbf{P}_y(S_{n-1} \leq t - s < S_n,\ Y_{t-s} \in \Gamma)\kappa(x, dy)ds \tag{5.16}$$

and

$$P_x(0 \leq t < S_1,\ Y_t \in \Gamma) = \mathbf{1}_{\{x \in \Gamma\}}e^{-\beta(x)t}. \tag{5.17}$$

Using the Markov property at time τ_1 and Lemmas 5.4 and 5.5, we can prove that, when we replace S_n by the n-th mutation time of X^K_{t/Ku_K} and Y_t by the support of X^K_{t/Ku_K} (when it is a singleton) in the LHS of (5.16) and (5.17), the same relations hold in the limit $K \to +\infty$. Therefore, Theorem 5.1 is proved for one-dimensional time marginals. A similar method generalizes to finite dimensional laws.

References

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

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

[3] Bolker, B.M., Pacala, S.W.: Spatial moment equations for plant competition: Understanding spatial strategies and the advantages of short dispersal. Am. Nat. **153**, 575–602 (1999).

[4] Bürger, R.: The Mathematical Theory of Selection, Recombination, and Mutation. John Wiley & Sons, Chichester, (2000).

[5] Calder III, W.A.: Size, Function and Life History, Harvard University Press, Cambridge (1984).

[6] Champagnat, N.: A microscopic interpretation for adaptive dynamics trait substitution sequence models. Preprint MODALX 04/20, University of Paris X, (2004), to appear in Stoch. Process. Appl.

[7] Champagnat, N., Ferrière, R., Méléard, S.: Unifying evolutionary dynamics: From individual stochastic processes to macroscopic models. To appear in Theoretical Population Biology (2006).

[8] Charnov, E.L.: Life History Invariants. Oxford University Press, Oxford (1993).

[9] Desvillettes, L., Prévost, C., Ferrière, R.: Infinite dimensional reaction-diffusion for evolutionary population dynamics. Preprint CMLA, Ecole Normale Superieure de Cachan (2004).

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

[11] Dieckmann, U., Law, R.: Relaxation projections and the method of moments. Pages 412-455 in The Geometry of Ecological Interactions: Symplifying Spatial Complexity (U Dieckmann, R. Law, J.A.J. Metz, editors). Cambridge University Press, Cambridge (2000).

[12] Etheridge, A.: Survival and extinction in a locally regulated population. Ann. Appl. Probab. **14**, 188–214 (2004).

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

[14] Evans, S.N., Perkins, E.A.: Measure-valued branching diffusions with singular interactions. Canad. J. Math. **46**, 120–168 (1994).

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

[16] Freidlin, M.I., Wentzel, A.D.: Random Perturbations of Dynamical Systems. Springer-Verlag, Berlin (1984).

[17] Joffe, A.,Métivier, M.: Weak convergence of sequences of semimartingales with applications to multitype branching processes. Adv. Appl. Probab. **18**, 20–65 (1986).

[18] Kisdi, E.: Evolutionary branching under asymmetric competition. J. Theor. Biol. **197**, 149–162 (1999).

[19] Law, R., Murrell, D.J., Dieckmann, U.: Population growth in space and time: Spatial logistic equations. Ecology **84**, 252–262 (2003).

[20] Méléard, S., Roelly, S.: Sur les convergences étroite ou vague de processus à valeurs mesures. C. R. Acad. Sci. Paris Sér. I Math. **317**, 785–788 (1993).

[21] Metz J.A.J., Nisbet, R.M., Geritz, S.A.H.: How should we define fitness for general ecological scenarios. Trends Ecol. Evol. **7**, 198–202 (1992).

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

[23] Roelly-Coppoletta, S.: A criterion of convergence of measure-valued processes: application to measure branching processes. Stoch. Stoch. Rep. **17**, 43–65 (1986).