

PHENOTYPIC DIVERSITY AND POPULATION GROWTH IN A FLUCTUATING ENVIRONMENT

CLÉMENT DOMBRY,* *Université de Poitiers*

CHRISTIAN MAZZA,** *Université de Fribourg*

VINCENT BANSAYE,*** *École Polytechnique*

Abstract

Organisms adapt to fluctuating environments by regulating their dynamics, and by adjusting their phenotypes to environmental changes. We model population growth using multitype branching processes in random environments, where the offspring distribution of some organism having trait $t \in \mathcal{T}$ in environment $e \in \mathcal{E}$ is given by some (fixed) distribution $\Upsilon_{t,e}$ on \mathbb{N} . Then, the phenotypes are attributed using a distribution (strategy) $\pi_{t,e}$ on the trait space \mathcal{T} . We look for the optimal strategy $\pi_{t,e}$, $t \in \mathcal{T}$, $e \in \mathcal{E}$, maximizing the net growth rate or Lyapounov exponent, and characterize the set of optimal strategies. This is considered for various models of interest in biology: hereditary versus nonhereditary strategies and strategies involving or not involving a sensing mechanism. Our main results are obtained in the setting of nonhereditary strategies: thanks to a reduction to simple branching processes in a random environment, we derive an exact expression for the net growth rate and a characterization of optimal strategies. We also focus on typical genealogies, that is, we consider the problem of finding the typical lineage of a randomly chosen organism.

Keywords: Branching process in a random environment; phenotypic diversity; Lyapounov exponent; optimal strategy; extinction; typical genealogy

2010 Mathematics Subject Classification: Primary 60J80

Secondary 60K37; 62D25

1. Introduction

Organisms adapt to fluctuating environments by regulating their intrinsic dynamics, and adjust their phenotypes or traits to the random environment. Observations reveal that most cell populations are heterogeneous, that is, are composed of various phenotypes. Recently, several authors proposed models to explain this heterogeneity; see [9], [10], and [13]. In these papers, continuous-time stochastic processes representing the evolution of the different populations are constructed. They can be seen as birth processes coupled by migration in between them, where the birth and migration rates depend on a fluctuating environment.

We present more precisely the models considered by Kussel and Leibler [10]: the trait space \mathcal{T} is finite and the environment space \mathcal{E} is such that $\mathcal{E} = \mathcal{T}$, the idea being that trait i has the fastest growth rate in environment $e = i$. The migration rate matrix $H^{(k)} = (H_{ij}^{(k)})_{i,j \in \mathcal{T}}$ gives

Received 27 May 2010; revision received 17 January 2011.

* Postal address: Laboratoire de Mathématiques et Applications, Téléport 2-BP30179, Boulevard Pierre et Marie Curie, 86962 Futuroscope Chasseneuil Cedex, France. Email address: clement.dombry@math.univ-poitiers.fr

** Postal address: Département de Mathématique, Université de Fribourg, Chemin du Musée 23, CH-1700 Fribourg, Switzerland. Email address: christian.mazza@unifr.ch

*** Postal address: CMAP, Ecole Polytechnique, Route de Saclay, 91128 Palaiseau Cedex, France.

Email address: bansaye@polytechnique.edu

the switching rates $H_{ij}^{(k)}$ from phenotype i to j in environment $e = k$. The changing environment is modeled as an alternating renewal process. The migration rates $H_{ij}^{(k)}$ define possible strategies to overcome the uncertain future. The authors considered two basic strategies: *stochastic switching* and *responsive switching*. In the first case, organisms switch their phenotypes independently of the running environment, they do not use sensors, so the migration rates do not depend on the environment and can be written as $H_{ij}^{(k)} = H_{ij}$. Responsive switching assumes strong use of sensors, at a certain cost; if the environment evolves slowly, a good way to overcome the environmental changes is when organisms always opt for the fittest phenotype in the running environment, that is, for the trait having the largest growth rate; in this case, the migration rates are given by $H_{ij}^{(k)} = H_m$ when $j = k$ and $H_{ij}^{(k)} = 0$ when $j \neq k$. In both cases, denoting by $X_i(t)$ the mean number of organisms having phenotype i at time t , the vector $X(t) = (X_i(t))_{i \in \mathcal{T}}$ is a solution of the random differential equation $dX(t)/dt = A_{\mathcal{E}(t)}X(t)$, where $A_{\mathcal{E}(t)}$ denotes the rate matrix composed of the various growth and migration rates in environment $\mathcal{E}(t)$ at time t . The authors then estimated the related Lyapounov exponent $\ln(N(t))/t$ as $t \rightarrow \infty$, where $N(t) = \sum_{i \in \mathcal{E}} X_i(t)$ is the average total population size. The computations were performed under the assumption that the environment changes slowly, so equilibrium is reached during each phase; see [10] for more details.

A similar model was also considered by Thattai [13] for two phenotypes and two environments. Monte Carlo simulations indicate that responsive switching is sometimes not the optimal strategy. The average proportion of cells having the fastest growth rate is maximized for strategies using a sensing mechanism and allowing migration to the unfavorable state. This leads to heterogeneous populations and provides a rational explanation for population diversification. These models were then studied analytically by Gander *et al.* [6]. Modifications of these models were also considered more recently by Visco *et al.* [14], who modeled catastrophic events.

In the present work, we treat similar problems using discrete-time multitype branching processes in random environments, which offer alternatives to the above birth with migration processes. The advantage of multitype branching processes is that we can separate more clearly the birth and migration phases. We can, for example, treat in this way populations composed of organisms where birth occurs in a very precise period. When this is not the case, discrete-time modeling still provides relevant information for population growth. We consider the problem of finding optimal strategies that maximize the random Lyapounov exponent. Our mathematical approach allows us to deal with continuous trait and environment spaces \mathcal{T} and \mathcal{E} , illustrated on an explicitly solvable model originally developed by Haccou and Iwasa [9]. For the attention of scientists not familiar with advanced mathematics, we first illustrate some results and provide examples of when both \mathcal{T} and \mathcal{E} are finite.

1.1. Results

Assume that $\mathcal{T} = \{t_1, \dots, t_q\}$ and that $\mathcal{E} = \{e_1, \dots, e_r\}$. The process of interest $(Z_n)_{n \geq 0}$ is written as a random vector $Z_n = (Z_n^t)_{t \in \mathcal{T}}$, where Z_n^t denotes the number of organisms having trait $t \in \mathcal{T}$ at generation n . The transition between generations n and $n + 1$ is modeled as a two-step procedure. First, for given $t \in \mathcal{T}$, each of the $i = 1, \dots, Z_n^t$ organisms having trait t gives birth to a random number of descendants given by a random variable $\xi_{i,n}^{t,\omega_n}$, where the index $\omega_n \in \mathcal{E}$ models the environmental state at generation n . In the second phase, each of the $j = 1, \dots, \xi_{i,n}^{t,\omega_n}$ new individuals is assigned to a random trait $\tau_{i,j,n}^{t,\omega_n} \in \mathcal{T}$. As usual in such processes, we assume that all of these random variables are independent, with independent and identically distributed (i.i.d.) $\xi_{i,n}^{t,\omega_n}$ and $\tau_{i,j,n}^{t,\omega_n}$. We further assume that the environmental process $(\omega_n)_{n \geq 0}$ is fixed, and suppose it to be the realization of a stationary (i.e. for all $i \in \mathbb{N}$,

(w_0, w_1, \dots) is distributed as (w_i, w_{i+1}, \dots) and ergodic process (i.e. for all bounded Borel functions f , $P(f(w_0, w_1, \dots) = f(w_1, w_2, \dots)) = 1$ implies that $f(w_0, w_1, \dots)$ is constant) taking values in \mathcal{E} . Let $\Upsilon_{t,e}$ and $\pi_{t,e}$ be distributions of the $\xi_{i,n}^{t,e}$ and $\tau_{i,j,n}^{t,e}$ for a trait $t \in \mathcal{T}$ in environment $e \in \mathcal{E}$. We will often use the first moment $m_{t,e}$ of $\Upsilon_{t,e}$, giving the average number of descendants for an organism having trait t in environment e . Optimal strategies will be considered for fixed distributions $\Upsilon_{t,e}$, that is, we will maximize the population growth rate as a function of the distributions $\pi_{t,e}$. In Sections 3, 4, and 5 we consider extinction and optimal growth questions. Section 6 deals with typical genealogies, that is, we consider the question of finding the typical lineage of an individual chosen randomly at generation n . We also illustrate these notions both for finite spaces \mathcal{T} and \mathcal{E} and for continuous spaces.

As stated in the previous section, Kussel and Leibler [10] distinguished between stochastic and responsive switching. In the first case, the distributions $\pi_{t,e}$ on \mathcal{T} depend on t but not on e , so we can write $\pi_{t,e} \equiv \pi_t$. Concerning the second family of strategies, $\pi_{t,e}$ depends on e but not on t . We will also distinguish between several natural situations, namely between strategies involving or not involving a sensing mechanism, and strategies $\pi_{t,e}$ depending or not depending on t , called hereditary and nonhereditary in what follows. We illustrate some results when both \mathcal{T} and \mathcal{E} are finite, and for

1. nonhereditary strategies with no-sensing mechanism, where $\pi_{t,e} \equiv p$, for some probability measure on \mathcal{T} ,
2. nonhereditary strategies using a sensing mechanism, of the form $\pi_{t,e} = p_e$, where $\bar{p} = (p_e)_{e \in \mathcal{E}}$ is a family of probability measures on \mathcal{T} . This second family contains responsive switching strategies.

We will describe the optimal set for nonhereditary strategies, that is, we will characterize the set of distributions $(\pi_{t,e})$ yielding the fastest growth rate. We assume that the random environment $(\omega_n)_{n \geq 0}$ is a stationary process taking values in \mathcal{E} .

1.1.1. *The nonhereditary with no-sensing mechanism case.* We first prove in Proposition 1 below that

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log E_\omega[|Z_n|] = \gamma(p),$$

where the Lyapounov exponent $\gamma(p)$ is given by

$$\gamma(p) = E[\log m_{p,\omega_0}]$$

and $m_{p,e}$ is the first moment or average value of the mean distribution

$$\Upsilon_{p,e} = \int_{\mathcal{T}} \Upsilon_{t,e} p(dt).$$

Let $\mathcal{P}(\mathcal{T})$ be the set of probability measures on \mathcal{T} , and let $\gamma^* = \sup\{\gamma(p); p \in \mathcal{P}(\mathcal{T})\}$ be the maximal Lyapounov exponent. We denote by \mathcal{P}^* the subset of $\mathcal{P}(\mathcal{T})$ containing the distributions p maximizing the Lyapounov exponent. We show in Proposition 3 below that $p \in \mathcal{P}^*$ if and only if

$$\int_{\mathcal{E}} \frac{m_{t,e}}{m_{p,e}} \nu_1(de) \leq 1 \quad \text{for all } t \in \mathcal{T},$$

where ν_1 is the law of the stationary random environment. A strategy is called *pure* if it is concentrated on a single $t \in \mathcal{T}$, that is, takes the form $p = \delta_t$. It then follows that a pure

strategy $p = \delta_t$ is optimal if and only if

$$\int_{\mathcal{E}} \frac{m_{t',e}}{m_{t,e}} v_1(de) \leq 1 \quad \text{for all } t' \in \mathcal{T}.$$

When $\mathcal{T} = \{t_1, \dots, t_q\}$ and $\mathcal{E} = \{e_1, \dots, e_r\}$, with $q \geq r$, we also prove that there is a unique maximizer in \mathcal{P}^* which is supported by a set containing at most p elements.

As a further illustration, we consider the simplest case when $p = q = 2$.

1. $p^* = \delta_{t_2}$ and $\gamma^* = v_1(e_1) \log(m_{t_2,e_1}) + v_1(e_2) \log(m_{t_2,e_2})$ when $v_1(e_1)m_{t_1,e_1}/m_{t_2,e_1} + v_1(e_2)m_{t_1,e_2}/m_{t_2,e_2} \leq 1$.
2. $p^* = \delta_{t_1}$ and $\gamma^* = v_1(e_1) \log(m_{t_1,e_1}) + v_1(e_2) \log(m_{t_1,e_2})$ when $v_1(e_1)m_{t_2,e_1}/m_{t_1,e_1} + v_1(e_2)m_{t_2,e_2}/m_{t_1,e_2} \leq 1$.
3. Otherwise p^* is given by

$$p^* = \left(\frac{v_1(e_1)m_{t_2,e_2}}{m_{t_2,e_2} - m_{t_1,e_2}} + \frac{v_1(e_2)m_{t_2,e_1}}{m_{t_2,e_1} - m_{t_1,e_1}} \right) \delta_{t_1} + \left(\frac{v_1(e_1)m_{t_1,e_2}}{m_{t_1,e_2} - m_{t_2,e_2}} + \frac{v_1(e_2)m_{t_1,e_1}}{m_{t_1,e_1} - m_{t_2,e_1}} \right) \delta_{t_2}$$

with

$$\begin{aligned} \gamma^* = & \log |m_{t_1,e_1}m_{t_2,e_2} - m_{t_1,e_2}m_{t_2,e_1}| - v_1(e_1) \log |m_{t_2,e_2} - m_{t_1,e_2}| \\ & - v_1(e_2) \log |m_{t_2,e_1} - m_{t_1,e_1}| + v_1(e_1) \log(v_1(e_1)) + v_1(e_2) \log(v_1(e_2)). \end{aligned}$$

As a numerical example, consider the case in which $v_1(e_1) = v_1(e_2) = \frac{1}{2}$ and

$$m_{t_1,e_1} = 1.5, \quad m_{t_2,e_1} = 0.6, \quad m_{t_1,e_2} = 0.6, \quad m_{t_2,e_2} = 1.5. \tag{1}$$

The two environments are equiprobable, with type t_1 better fitted to environment e_1 , and a symmetry in the model. In this case, the Lyapounov exponent for the population following a pure strategy is given by

$$\gamma(\delta_{t_1}) = \gamma(\delta_{t_2}) = 0.5(\log(1.5) + \log(0.6)) \approx -0.053.$$

This implies that both types are subcritical and that the corresponding homogeneous populations go extinct almost surely. The optimal strategy is then $p^* = (0.5, 0.5)$ and the corresponding Lyapounov exponent is

$$\gamma^* = \log(0.5(1.5 + 0.6)) \approx 0.049.$$

With this optimal strategy, $\gamma^* > 0$ and the population explodes with positive probability almost surely. This is one of the simplest examples when polymorphism is a necessary condition for survival (see Section 3.3).

1.1.2. *The nonhereditary with sensing mechanism case.* Assume that $\pi_{t,e} = p_e$ for some strategy $\bar{p} = (p_e)_{e \in \mathcal{E}}$. It turns out that the relevant piece of environment is given by the pair process $\omega^{(2)} = ((\omega_{n-1}, \omega_n))_{n \geq 1} \in \mathcal{E}^2$, of stationary measure $\nu_2(de_1, de_2)$. Consider the average distribution

$$\Upsilon_{\bar{p},(e_1,e_2)} = \int_{\mathcal{T}} \Upsilon_{t,e_2} p_{e_1}(dt),$$

of expected value

$$m_{p_{e_1},e_2} = \int_{\mathcal{T}} m_{t,e_2} p_{e_1}(dt).$$

When

$$\gamma(\bar{p}) = \int_{\mathcal{E}^2} \log(m_{p_{e_1}, e_2}) v_2(de_1, de_2)$$

exists, we prove in Proposition 7 below that

$$\lim_{n \rightarrow \infty} n^{-1} \log E_\omega[|Z_n|] = \gamma(\bar{p}) \quad \text{almost surely (a.s.).}$$

Let γ^{**} be the optimal growth rate, that is, the supremum of $\gamma(\bar{p})$ among all possible families \bar{p} , and denote by \mathcal{P}^{**} the related set of optimal strategies. Let $v_{e_1}(de_2)$ be the conditional distribution of ω_2 given that $\omega_1 = e_1$. We prove in Proposition 9 below that \bar{p} is optimal if and only if

$$\int_{\mathcal{E}} \frac{m_{t, e_2}}{m_{p_{e_1}, e_2}} v_{e_1}(de_2) \leq 1 \quad \text{for all } t \in \mathcal{T},$$

$v_1(de_1)$ -a.s. An interesting consequence is that there is no gain to be expected using a sensing mechanism when the random environment has some independence property: if ω_1 and ω_2 are independent, then $\gamma^* = \gamma^{**}$. The information we can gather on the present environmental state using sensors does not help when dealing with future events.

Next, we develop further the simplest case when $\mathcal{E} = \{e_1, e_2\}$ and $\mathcal{T} = \{t_1, t_2\}$. We suppose that the distribution for the environment v is a Markov chain with transitions

$$\begin{aligned} v(\omega_{n+1} = e_1 \mid \omega_n = e_1) &= 1 - q_1, & v(\omega_{n+1} = e_2 \mid \omega_n = e_1) &= q_1, \\ v(\omega_{n+1} = e_1 \mid \omega_n = e_2) &= q_2, & v(\omega_{n+1} = e_2 \mid \omega_n = e_2) &= 1 - q_2, \end{aligned}$$

where $q_i \in (0, 1)$ denotes the probability that the environment switches in the next step when it is currently in state e_i . The sequence ω is then ergodic and its stationary distribution is

$$v_1 = \frac{q_2}{q_1 + q_2} \delta_{e_1} + \frac{q_1}{q_1 + q_2} \delta_{e_2}.$$

We suppose further that the Markov chain is started from the steady state, i.e. ω_0 has distribution v_1 , so the sequence ω is stationary. Then, the conditional distributions of ω_1 given ω_0 is precisely given by the transition of the Markov chain. The optimal strategy with sensing is then such that

1. $p_{e_1}^{**} = \delta_{t_2}$ when $(1 - q_1)m_{t_1, e_1}/m_{t_2, e_1} + q_1m_{t_1, e_2}/m_{t_2, e_2} \leq 1$,
2. $p_{e_1}^{**} = \delta_{t_1}$ when $(1 - q_1)m_{t_2, e_1}/m_{t_1, e_1} + q_1m_{t_2, e_2}/m_{t_1, e_2} \leq 1$,
3. and otherwise

$$p_{e_1}^{**} = \left(\frac{(1 - q_1)m_{t_2, e_2}}{m_{t_2, e_2} - m_{t_1, e_2}} + \frac{q_1m_{t_2, e_1}}{m_{t_2, e_1} - m_{t_1, e_1}} \right) \delta_{t_1} + \left(\frac{(1 - q_1)m_{t_1, e_2}}{m_{t_1, e_2} - m_{t_2, e_2}} + \frac{q_1m_{t_1, e_1}}{m_{t_1, e_1} - m_{t_2, e_1}} \right) \delta_{t_2}.$$

Similar equations hold for $p_{e_2}^{**}$. Rather than giving a general formula, we consider the numerical case when $q_1 = q_2 = q$ so that the stationary distribution is given by $v_1(e_1) = v_1(e_2) = \frac{1}{2}$, and we use the same values for the number of offspring as in (1). Using the above results, the optimal strategy p^{**} is given by

$$\begin{aligned} p_{e_1}^{**} &= \delta_{t_1} \quad \text{and} \quad p_{e_2}^{**} = \delta_{t_2} \quad \text{if } q \leq \frac{2}{7}, \\ p_{e_1}^{**} &= \delta_{t_2} \quad \text{and} \quad p_{e_2}^{**} = \delta_{t_1} \quad \text{if } q \geq \frac{5}{7}, \\ p_{e_1}^{**} &= \frac{5 - 7q}{3} \delta_{t_1} + \frac{7q - 2}{3} \delta_{t_2} \quad \text{and} \quad p_{e_2}^{**} = \frac{7q - 2}{3} \delta_{t_1} + \frac{5 - 7q}{3} \delta_{t_2} \quad \text{if } \frac{2}{7} \leq q \leq \frac{5}{7}. \end{aligned}$$

We see three different environmental regimes, corresponding to low, intermediate, and high frequency switching rates. When the environment fluctuates slowly, with $q \leq \frac{2}{7}$, the optimal strategy is pure and corresponds to what Kussel and Leibler [10] called responsive switching. In the intermediate regime, the optimal strategy is a mixture of pure strategies. In the high frequency regime where the environment fluctuates quickly ($q \geq \frac{5}{7}$), the optimal strategy is a pure strategy, where organisms being in the favorable state are pushed to the unfavorable state.

We deduce the optimal growth rate

$$\gamma^{**} = \begin{cases} \log \frac{3}{2} - q \log \frac{5}{2} & \text{if } q \leq \frac{2}{7}, \\ q \log q + (1 - q) \log(1 - q) + \log \frac{21}{10} & \text{if } \frac{2}{7} \leq q \leq \frac{5}{7}, \\ \log \frac{3}{5} - q \log \frac{5}{2} & \text{if } q \geq \frac{5}{7}. \end{cases}$$

When $q = \frac{1}{2}$, the sequence ω is an independent sequence, so $\gamma^{**}(\frac{1}{2}) = \gamma^*(\frac{1}{2}) \approx 0.049$, the optimal strategy is a strategy without sensing. When $q \approx 1$ or $q \approx 0$, the environment in the next step is very likely to stay the same or, respectively, to switch, so we can determine with high probability which type will be fitted in the next generation. Indeed, we observe that

$$\lim_{q \rightarrow 0} \gamma^{**}(q) = \lim_{q \rightarrow 1} \gamma^{**}(q) = \log \frac{3}{2} \approx 0.176.$$

2. Definitions of the multitype branching process in a random environment and of optimal strategies

First, we recall that ω_n represents the environment at time n , Z_n represents the trait-structured population, Z_n^t denotes the number of individuals with trait t , and $|Z_n|$ denotes the total number of individuals at time n . Moreover, $\xi_{i,n}^{t,\omega_n}$ gives the size of the offspring of the i th individual with trait t at time n in environment ω_n , and $\tau_{i,j,n}^{t,\omega_n}$ gives the trait of the j th descendant of this individual.

Let us now give the formal definition of the process. Let (Ω, \mathcal{F}, P) be a probability space, let \mathcal{T} be a metric space, and let \mathcal{E} be a Polish space. For each pair $(t, e) \in \mathcal{T} \times \mathcal{E}$, let $\Upsilon_{t,e}$ and $\pi_{t,e}$ be distributions on \mathbb{N} and \mathcal{T} , respectively, and suppose that $\Upsilon_{t,e}$ has a finite first moment denoted by $m_{t,e}$. Let $\omega = (\omega_n)_{n \geq 0}$ be an \mathcal{E} -valued stationary ergodic random process with distribution ν on $\mathcal{E}^{\mathbb{N}}$.

Denoting by $\mathbb{N}^{\mathcal{T}}$ the set of almost null \mathbb{N} -valued functions on \mathcal{T} and by $(1_t)_{t \in \mathcal{T}}$ the canonical basis, the stochastic process $Z_n = (Z_n^t : t \in \mathcal{T})$ with values in $\mathbb{N}^{\mathcal{T}}$ can be defined as

$$Z_0 = N_0, \quad Z_{n+1} = \sum_{t \in \mathcal{T}} \sum_{i=1}^{Z_n^t} \sum_{j=1}^{\xi_{i,n}^{t,\omega_n}} 1_{\tau_{i,j,n}^{t,\omega_n}}, \quad n \geq 0, \tag{2}$$

where

- N_0 is an $\mathbb{N}^{\mathcal{T}}$ -valued random variable giving the population at time 0,
- $\{\xi_{i,n}; i \geq 1, n \geq 0\}$ and $\{\tau_{i,j,n}; i \geq 1, j \geq 1, n \geq 0\}$ are infinite arrays of i.i.d. random variables with values in $\mathbb{N}^{\mathcal{T} \times \mathcal{E}}$ and $\mathcal{T}^{\mathcal{T} \times \mathcal{E}}$, respectively, and common distributions

$$\Upsilon = \bigotimes_{(t,e) \in \mathcal{T} \times \mathcal{E}} \Upsilon_{t,e} \quad \text{and} \quad \pi = \bigotimes_{(t,e) \in \mathcal{T} \times \mathcal{E}} \pi_{t,e},$$

respectively,

- N_0 , $\{\xi_{i,n}; i \geq 1, n \geq 0\}$, $\{\tau_{i,j,n}; i \geq 1, j \geq 1, n \geq 0\}$, and $\omega = (\omega_n)_{n \geq 0}$ are independent.

Thus, Υ corresponds to the offspring distribution and π corresponds to the trait distribution of the offspring. Both distributions a priori depend on the trait of the parent and on the environment. We set $|Z_n| = \sum_{i \in \mathcal{J}} Z_n^i$, $n \geq 0$ (an empty sum is equal to 0).

Our attention will be focused on the role of the trait distribution π and how it affects the evolution of the population. The offspring distribution Υ will hence be considered as given and fixed, whereas the trait distribution π will be considered as a parameter. The intuitive idea is that (i) the local fitness of an individual is determined by its trait and its environment, and corresponds to its mean number of descendants; and (ii) the intergenerational variability of the traits has to be tuned so as to maximize the global fitness of the whole population.

The trait distribution π can be seen as the strategy used by the population to maximize its growth. We measure the performance of the strategy π by the long-term growth of the corresponding population $(Z_n)_{n \geq 0}$: let

$$\gamma(\pi) = \lim_{n \rightarrow \infty} n^{-1} \log E_{\omega, N_0}[|Z_n|].$$

Here E_{ω, N_0} denotes the conditional expectation given the environment and the initial population. As will be discussed in the sequel, fairly general assumptions ensure that this limit exists and does not depend on the environment and initial population. Note that it is important here to consider the quenched model (i.e. conditionally on the environment) and not the averaged model (i.e. averaged over all possible environments): the criterion $\gamma(\pi)$ measures the population growth rate in a typical environment. An averaged criterion would be biased by unlikely environments where the population grows unusually fast.

Several mechanisms can induce the intergenerational variability of the traits and leads to different assumptions on the trait distribution π .

3. The nonhereditary case with no-sensing mechanism

According to the approach set out in the previous paragraph, we begin by studying the simplest case of nonhereditary traits in the absence of a sensing mechanism, meaning that the trait distribution of the offspring depends neither on the trait of the parent nor on the environment. We thus suppose that Z_n evolves according to model (2) with $\pi_{i,e} \equiv p$ for some distribution p on \mathcal{T} , and let $\pi = \pi(p)$ be the corresponding trait distribution.

3.1. Reduction to a simple branching process in a random environment

In some sense, the nonhereditary assumption makes the structure of the population trivial: the trait distribution is given by p whatever the past evolution of the process. In mathematical terms, we take advantage of some stochastic independence.

Lemma 1. *For any $n \geq 1$, the population structure is conditionally independent of the past population process given the size of the population, i.e.*

$$Z_n \Big| |Z_n| \prod (Z_0, \dots, Z_{n-1}) \Big| |Z_n|.$$

Proof. It is easily seen from the assumptions on model (2) and from the nonhereditary assumption $\pi = \pi(p)$ that the distribution of Z_n given (Z_0, \dots, Z_{n-1}) and $|Z_n|$ is equal to the distribution of $\sum_{1 \leq i \leq |Z_n|} 1_{\tau_i}$, where τ_i is an i.i.d. sequence with distribution p . This distribution does not depend on (Z_0, \dots, Z_{n-1}) , proving the conditional independence.

It is worth noting that the result also holds for the quenched model (i.e. conditionally on the environment ω). Lemma 1 implies that the distribution of the population process $(Z_n)_{n \geq 1}$ is easily recovered from the size process $(|Z_n|)_{n \geq 1}$. This latter process turns out to be a simple branching process in a random environment (BPRE) and this allows us to compute the performance $\gamma(p)$ of the strategy p .

Proposition 1. *The size process $(|Z_n|)_{n \geq 1}$ is a simple BPRE with offspring distribution in environment e given by*

$$\Upsilon_{p,e} = \int_{\mathcal{T}} \Upsilon_{t,e} p(dt), \quad e \in \mathcal{E}.$$

Conditionally on ω , the expected population size at time n is

$$E_\omega[|Z_n|] = E_{\omega_0}[|Z_1|] \prod_{k=1}^{n-1} m_{p,\omega_k},$$

where $m_{p,e}$ is the first moment of $\Upsilon_{p,e}$. If $\gamma(p) = E[\log m_{p,\omega_0}]$ exists then

$$\lim_{n \rightarrow \infty} n^{-1} \log E_\omega[|Z_n|] = \gamma(p) \quad \text{a.s.}$$

Proof. According to Lemma 1, given $(|Z_1|, \dots, |Z_n|)$, the population Z_n has the same distribution as $\sum_{1 \leq i \leq |Z_n|} 1_{\tau_i}$, where τ_i is an i.i.d. sequence with distribution p . Intuitively, the i th individual has type τ_i chosen randomly on \mathcal{T} with distribution p . The size of the next generation $|Z_{n+1}|$ is then $\sum_{1 \leq i \leq |Z_n|} \xi_{n,i}^{\tau_i, \omega_n}$, where $\xi_{n,i}^{\tau_i, \omega_n}$ is the offspring of the i th individual of type τ_i in environment e_n . From this two-step procedure, that is, the random choice of the type t and the reproduction with random offspring in environment e , we obtain the effective offspring distribution $\Upsilon_{p,e}$ in environment e : it is the mixture of the offspring distributions $\Upsilon_{t,e}$, with mixing distribution p . This proves the branching property for $(|Z_n|)_{n \geq 1}$. The other properties follow: the branching property implies the recursive formula

$$E_\omega[|Z_{n+1}|] = E_{(\omega_0, \dots, \omega_{n-1})}[|Z_n|] m_{p,\omega_n},$$

and the formula for $E_\omega[|Z_n|]$ follows. Taking the logarithm, we have

$$n^{-1} \log E_\omega[|Z_n|] = n^{-1} \log E_{\omega_0}[|Z_1|] + n^{-1} \sum_{k=1}^{n-1} \log m_{p,\omega_k},$$

which converges to $\gamma(p)$ a.s. according to the ergodic theorem and the integrability assumptions. This completes the proof.

For the sake of simplicity, we suppose in what follows that the initial population consists of a single individual of random trait with distribution p . The whole process $(|Z_n|)_{n \geq 0}$ is then a simple BPRE with offspring distribution Υ_p and initial condition $|Z_n| = 1$. Simple branching processes in a random environment have been introduced by Smith and Wilkinson [11] and Athreya and Karlin [2], [3], and have been studied rather intensively since then [1], [7], [12]. We recall here some important results concerning the asymptotic behavior of such processes (see, e.g. the classification theorem in [12], which states conditions under which either the population becomes extinct or explodes at a geometric rate).

To avoid the trivial case of a constant population, we suppose that Υ_{p,ω_0} is not a.s. equal to δ_1 . We suppose also that $\gamma(p) = E[\log m_{p,\omega_0}]$ exists and is finite. We say that extinction

occurs if $Z_n \rightarrow 0$ as $n \rightarrow \infty$ (in which case the sequence vanishes eventually), and that the population survives otherwise. Let $q(\omega) = P_\omega(Z_n \rightarrow 0)$ be the probability of extinction given the environment.

Theorem 1. (Classification theorem.) 1. *In both the subcritical case, $\gamma(p) < 0$, and the critical case, $\gamma(p) = 0$, the population becomes extinct a.s., i.e. $P(q(\omega) = 1) = 1$.*

2. *In the supercritical case, $\gamma(p) > 0$, if, furthermore, $E[-\log(1 - \Upsilon_{p,\omega_0}(\{0\}))] < \infty$ then the population can survive with positive probability in almost every environment, i.e. $P(q(\omega) < 1) = 1$. Furthermore, conditionally on nonextinction, the population explodes at a geometric rate:*

$$\lim_{n \rightarrow \infty} n^{-1} \log Z_n = \gamma(p) \quad \text{a.s. on } \{Z_n > 0 \text{ for all } n \geq 0\}.$$

3.2. Optimal strategies

We now focus on optimal strategies, i.e. what choice of the distribution p allows for the fastest growth of the population? As discussed in Section 2, the performance of the strategy $\pi = \pi(p)$ is measured by the Lyapounov exponent $\gamma(p)$. In the nonhereditary case with no sensing, an explicit formula for $\gamma(p)$ has been derived in Proposition 1. The questions which naturally arise are those of how to determine the supremum of $\gamma(p)$ for p varying in $\mathcal{P}(\mathcal{T})$, the space of probability measures on \mathcal{T} , and the set of optimal strategies p^* in the case when this supremum is reached. Let $\gamma^* = \sup\{\gamma(p); p \in \mathcal{P}(\mathcal{T})\}$ be the optimal growth rate, and let $\mathcal{P}^* = \{p^* \in \mathcal{P}(\mathcal{T}); \gamma(p^*) = \gamma^*\}$ be the set of optimal strategies. A strategy p is called pure if $p = \delta_t$ for some $t \in \mathcal{T}$ (i.e. all individuals in the population have the same trait t) or mixed otherwise. We further ask whether the optimal strategies are pure or mixed.

We now give conditions under which the existence of optimal strategies are ensured. We suppose that

(C1) there is some $M > 0$ such that $m_{t,e} \leq M$ for all $(t, e) \in \mathcal{T} \times \mathcal{E}$,

(C2) for all $e \in \mathcal{E}$, the mapping $t \mapsto m_{t,e}$ is continuous on \mathcal{T} ,

(C3) for all $e \in \mathcal{E}$ and $\varepsilon > 0$, there is a compact set $K \subset \mathcal{T}$ such that $m_{t,e} \leq \varepsilon$ for all $t \in \mathcal{T} \setminus K$.

Condition (C1) is rather relevant from the biological point of view since an individual could hardly have an arbitrarily high number of offspring in a fixed amount of time. Conditions (C2) and (C3) are related with the topology of \mathcal{T} : individuals with close traits are supposed to have approximately the same behavior, and traits close to infinity are supposed to have a poor fitness.

Proposition 2. *Under conditions (C1)–(C3), optimal strategies exist and form a closed convex set, that is, the set \mathcal{P}^* is nonempty, closed, and convex in $\mathcal{P}(\mathcal{T})$ endowed with the topology of weak convergence.*

If, furthermore, the family $\mathcal{M} = \{e \mapsto m_{t,e}; t \in \mathcal{T}\} \subset L^\infty(\mathcal{E}, \nu_1)$ is linearly independent then the optimal strategy is unique.

Proof. First, we prove that the map $\gamma: \mathcal{P}(\mathcal{T}) \rightarrow \mathbb{R}$ defined by

$$\gamma(p) = \int_{\mathcal{E}} \log(m_{p,e}) \nu_1(de) \quad \text{with} \quad m_{p,e} = \int_{\mathcal{T}} m_{t,e} p(dt) \tag{3}$$

is concave and upper semicontinuous on $\mathcal{P}(\mathcal{T})$ with respect to the topology of weak convergence. Let p_n converge weakly to p in $\mathcal{P}(\mathcal{T})$. Using conditions (C1) and (C2), we see

that $m_{p_n,e} \rightarrow m_{p,e}$ for all $e \in \mathcal{E}$. Defining $f_n(e) = \log(M) - \log(m_{p_n,e})$ to be nonnegative measurable functions on \mathcal{E} and applying Fatou's lemma, we obtain

$$\int_{\mathcal{E}} \liminf f_n(e) \nu_1(de) \leq \liminf \int_{\mathcal{E}} f_n(e) \nu_1(de),$$

or, equivalently,

$$\limsup \int_{\mathcal{E}} \log(m_{p_n,e}) \nu_1(de) \leq \int_{\mathcal{E}} \log(m_{p,e}) \nu_1(de).$$

Hence, $\limsup \gamma(p_n) \leq \gamma(p)$, and this gives the upper semicontinuity of the application γ . The concavity of the application γ on $\mathcal{P}(\mathcal{T})$ is a direct consequence of the concavity of the logarithm and the linearity of integration: first, for $p_1, p_2 \in \mathcal{P}(\mathcal{T})$, $\lambda \in [0, 1]$, and $e \in \mathcal{E}$, we have

$$\log(m_{\lambda p_1 + (1-\lambda)p_2,e}) = \log(\lambda m_{p_1,e} + (1-\lambda)m_{p_2,e}) \geq \lambda \log(m_{p_1,e}) + (1-\lambda) \log(m_{p_2,e}).$$

This implies that $\gamma(\lambda p_1 + (1-\lambda)p_2) \geq \lambda \gamma(p_1) + (1-\lambda)\gamma(p_2)$.

Secondly, we prove that \mathcal{P}^* is nonempty, closed, and convex. The closeness and convexity properties are straightforward since \mathcal{P}^* can be seen as the level set $\{p \in \mathcal{P}(\mathcal{T}); \gamma(p) \geq p^*\}$ of the concave upper semicontinuous application γ . It remains to check nonemptiness, for which we use compactness arguments. In the case when \mathcal{T} is a compact space, $\mathcal{P}(\mathcal{T})$ is also compact with respect to the weak topology and the upper semicontinuous map γ reaches its maximum on $\mathcal{P}(\mathcal{T})$ so that \mathcal{P}^* is nonempty. In the case when \mathcal{T} is noncompact, we consider its compactification $\hat{\mathcal{T}} = \mathcal{T} \cup \{\infty\}$. We extend the definition of m by $m_{\infty,e} = 0$ for all $e \in \mathcal{E}$. Condition (C3) ensures that this extension is continuous on $\hat{\mathcal{T}}$. Then the map γ extends on the compact space $\mathcal{P}(\hat{\mathcal{T}})$, is upper semicontinuous, and, hence, reaches its maximum at some point $p^* \in \mathcal{P}(\hat{\mathcal{T}})$. It remains to check that $p^*(\{\infty\}) = 0$ so that p^* can be seen as an element of $\mathcal{P}(\mathcal{T})$. It is straightforward since $p^*(\{\infty\}) > 0$ would imply that $\gamma(p^*) = -\infty$.

Lastly, we prove the uniqueness in the case when the family \mathcal{M} is linearly independent. Let p_1^* and p_2^* be two points where γ reaches its maximum. Using the strict concavity of the logarithm, we see that necessarily $m_{p_1^*,e} = m_{p_2^*,e} \nu_1(de)$ -almost everywhere. Using the linear independence, this in turn implies that $p_1^* = p_2^*$.

The following characterization of optimal strategies can be useful.

Proposition 3. *A strategy p is optimal if and only if*

$$\int_{\mathcal{E}} \frac{m_{t,e}}{m_{p,e}} \nu_1(de) \leq 1 \quad \text{for all } t \in \mathcal{T}.$$

Proof. The strategy p is optimal if and only if

$$\gamma((1-\varepsilon)p + \varepsilon p') - \gamma(p) \leq 0 \quad \text{for all } p' \text{ and } 0 < \varepsilon < 1.$$

Using concavity and differentiability, this is equivalent to

$$\left. \frac{d}{d\varepsilon} [\gamma(p + \varepsilon(p' - p))] \right|_{\varepsilon=0} \leq 0,$$

which can be rewritten as

$$\left. \frac{d}{d\varepsilon} \left[\int_{\mathcal{E}} \log(m_{p,e} + \varepsilon(m_{p',e} - m_{p,e})) \nu_1(de) \right] \right|_{\varepsilon=0} = \int_{\mathcal{E}} \frac{m_{p',e} - m_{p,e}}{m_{p,e}} \nu_1(de) \leq 0.$$

Thus, a necessary and sufficient condition is

$$\int_{\mathcal{E}} \frac{m_{p',e}}{m_{p,e}} \nu_1(de) \leq 1 \quad \text{for all } p' \in \mathcal{P}(\mathcal{T}).$$

It is easily seen that this is equivalent to testing the condition for $p' = \delta_t$, $t \in \mathcal{T}$, and this completes the proof.

As a direct application of Proposition 3, we can now answer the question of whether there is a pure optimal strategy.

Corollary 1. *The pure strategy $p = \delta_t$ is optimal if and only if*

$$\int_{\mathcal{E}} \frac{m_{t',e}}{m_{t,e}} \nu_1(de) \leq 1 \quad \text{for all } t' \in \mathcal{T}.$$

This gives us a simple criterion for the existence of optimal pure strategies.

3.3. Extinction

According to Theorem 1, on the one hand, the population becomes extinct in the case when $\gamma(p) \leq 0$ and, on the other hand (under additional technical conditions), the population survives and explodes with positive probability when $\gamma(p) > 0$. For the population's survival, strategies p with $\gamma(p) > 0$ are of the highest importance. This leads to the definition of the set \mathcal{S} of strategies that allow the population to survive:

$$\mathcal{S} = \{p \in \mathcal{P}(\mathcal{T}); \gamma(p) > 0\}.$$

The following property holds.

Proposition 4. *If $\gamma^* \leq 0$, the set \mathcal{S} is empty. If $\gamma^* > 0$, \mathcal{S} is a nonempty convex set containing \mathcal{P}^* .*

Proof. From the definition of γ^* , $\gamma(p) \leq \gamma^*$ for all strategies p . Hence, \mathcal{S} is empty if $\gamma^* \leq 0$. If $\gamma^* > 0$ then every optimal strategy $p^* \in \mathcal{P}^*$ satisfies $\gamma(p^*) = \gamma^* > 0$ and, hence, $\mathcal{P}^* \subset \mathcal{S}$. The map $p \mapsto \gamma(p)$ is a concave function, so the level set $\mathcal{S} = \{p \in \mathcal{P}(\mathcal{T}); \gamma(p) > 0\}$ is convex. This completes the proof.

In some cases, a striking phenomenon may happen: no pure strategy can allow for survival, i.e. every homogeneous population with a single trait t suffers from extinction; however, some mixed strategies may prevent extinction, i.e. some polymorphic populations may survive forever. In this case, we should say that polymorphism is a necessary condition for survival. This phenomenon occurs when $\gamma(\delta_t) \leq 0$ for all $t \in \mathcal{T}$ whereas $\gamma^* > 0$.

The intuitive idea is the following: extinction occurs when the environment is bad for almost all the individuals in the population and, hence, a diversification of the traits in the population should imply a smaller number of environments that are bad for almost all individuals. This can be seen as a consequence of the concavity property: suppose that, for all traits t , $\gamma(\delta_t) \equiv \gamma$, i.e. homogeneous populations have the same growth rate no matter the trait t . Then, for any strategy p , $\gamma(p) \geq \int_{\mathcal{T}} \gamma(t)p(dt) = \gamma$, i.e. any polymorphic population has a better growth rate than any homogeneous population. See the examples below for further illustration of this phenomenon.

3.4. Example: finite-dimensional case

We consider the case when the vectorial space spanned by the family $\mathcal{M} = \{e \mapsto m_{t,e}; t \in \mathcal{T}\}$ in $L^\infty(\mathcal{E}, \nu_1)$ has finite dimension denoted by d . This occurs in particular as soon as one of the following occurs.

- The environment space is finite: $\mathcal{E} = \{e_1, \dots, e_r\}$ and ν_1 is a discrete measure such that $\nu_1(e_i) > 0$ for all $i \in \{1, \dots, p\}$. In this case $L^\infty(\mathcal{E}, \nu_1)$ is of dimension p and $d \leq p$.
- The trait is finite: $\mathcal{T} = \{t_1, \dots, t_q\}$. Then $d \leq q$ and conditions (C1)–(C3) are automatically satisfied.

For convenience, we further require that the following condition holds.

- (H) For any pairwise distinct $t_1, \dots, t_d \in \mathcal{T}$, the family of functions $\{e \mapsto m_{t_i,e}; 1 \leq i \leq d\}$ is linearly independent in $L^\infty(\mathcal{E}, \nu_1)$.

This condition will be convenient because it ensures the uniqueness of the optimal strategy.

Proposition 5. *Suppose that $\text{span}\mathcal{M}$ has dimension d , and that conditions (C1)–(C3) and (H) hold. Then*

- *there exists a unique optimal strategy $p^* \in \mathcal{P}(\mathcal{T})$,*
- *p^* is a discrete probability on \mathcal{T} supported by at most d different types.*

Proof. Existence of an optimal strategy is a consequence of Proposition 2. However, it is worth noting that in this finite-dimensional case, it is a consequence of standard analysis. We denote by $\text{Conv}(\mathcal{M})$ the closed convex hull of \mathcal{M} in $L^\infty(\mathcal{E}, \nu_1)$ which can be seen as a closed convex set in a d -dimensional vector space. Note that $\text{Conv}(\mathcal{M})$ is equal to the set of functions $\{m_p: e \mapsto m_{p,e}; p \in \mathcal{P}(\mathcal{T})\}$, where $m_{p,e}$ is defined in (3). Introduce the application $\Theta: \text{Conv}(\mathcal{M}) \rightarrow [-\infty, \infty)$ defined by $\Theta(m) = \int_{\mathcal{E}} \log(m(e)) \nu_1(de)$ (with the convention that $\log(u) = -\infty$ if $u \leq 0$). With this notation, we easily see that $\gamma(p) = \Theta(m_p)$ and maximizing γ on $\mathcal{P}(\mathcal{T})$ is equivalent to maximizing Θ on $\text{Conv}(\mathcal{M})$. Next we observe that the function Θ is upper semicontinuous and strictly concave, so it reaches its maximum at a unique point m^* on the compact set $\text{Conv}(\mathcal{M})$. Furthermore, the gradient of Θ never vanishes, so the extremum must be reached on a boundary point of $\text{Conv}(\mathcal{M})$ and $m^* \in \partial\text{Conv}(\mathcal{M})$. The boundary of $\text{Conv}(\mathcal{M})$ consists of extremal types (type t such that $m_t: e \mapsto m_{t,e}$ is an extremal point of the convex set $\text{Conv}(\mathcal{M})$) and of k -dimensional faces determined by the convex hull of $k + 1$ extremal types, with $0 \leq k \leq n$. If m^* belongs to such a k -dimensional face then $m^* = m_{p^*}$, where p^* is a discrete probability on \mathcal{T} supported by the $k + 1$ corresponding extremal types, with $k + 1 \leq d$. The uniqueness property then follows from assumption (H). Indeed, under this assumption, a k -dimensional face determined by $k + 1$ extremal types does not contain any other point of \mathcal{M} , so the decomposition in barycentric coordinates is unique. This completes the proof.

It is worth noting that the application Θ always has a unique minimizer m^* in the above proof. Condition (H) ensures that there is a unique strategy p^* such that $m_{p^*} = m^*$. In the absence of this condition, there might be several mixing distributions p such that $m_p = m^*$ and then several optimal strategies.

An appealing particular case is the case when both $\mathcal{E} = \{e_1, \dots, e_r\}$ and $\mathcal{T} = \{t_1, \dots, t_q\}$ are finite with $q \geq p$. In this case $L^\infty(\mathcal{E}, \nu_1)$ is of dimension $d = p$, and a generic configuration will always satisfy assumption (H) (in the sense that an exact linear relation between the types

is very unlikely from a biological point of view). Then, according to Proposition 5, there exists a unique optimal strategy mixing at most p different types. This shows that the number of types involved in the optimal strategy is less than the number of different environments. This is reminiscent of a rule in ecology which states that the number of species in an ecosystem is bounded above by the numbers of different niche in the sense that two species cannot occupy the same niche for a long time (competitive exclusion principle). From a practical point of view, in this finite settings, the optimal strategy can be computed using standard numerical convex optimization (see [5] for instance). For more illustrations in this setting, see Section 1.1.1.

3.5. Example: Gaussian distributions

This example is due to Haccou and Iwasa [9]. The environment space is the set of real numbers $\mathcal{E} = \mathbb{R}$ and the environment ω is assumed to be a Gaussian stationary ergodic sequence with stationary distribution $\nu_1 = \mathcal{N}(\mu, \sigma_2^2)$, the Gaussian distribution with mean $\mu \in \mathbb{R}$ and variance $\sigma_2^2 > 0$. The dynamic for the environment is irrelevant in the no-sensing case. The trait space is the set of real numbers $\mathcal{T} = \mathbb{R}$, and the mean offspring number of an individual of trait t in environment e has the Gaussian form

$$m_{t,e} = \frac{C}{\sqrt{2\pi\sigma_1^2}} \exp\left(-\frac{(t-e)^2}{2\sigma_1^2}\right)$$

for some parameters $C > 0$ and $\sigma_1^2 > 0$. In environment e , individuals with trait value $t = e$ have the largest mean offspring number. Note that conditions (C1)–(C3) are fulfilled so that Proposition 2 holds: optimal strategies exist. Furthermore, the family $\mathcal{M} = \{e \mapsto m_{t,e}; t \in \mathcal{T}\}$ is linearly independent so that unicity holds: there is a unique optimal strategy p^* depending a priori on μ, σ_1, σ_2 , and C .

First we easily compute the fitness of a pure strategy:

$$\gamma(\delta_t) = \int_{\mathcal{E}} \log(m_{t,e})\nu_1(de) = \log C - \frac{1}{2} \log(2\pi\sigma_1^2) - \frac{(\mu - t)^2 + \sigma_2^2}{2\sigma_1^2}.$$

The optimal pure strategy is equal to δ_μ . Then, according to Corollary 1, we determine whether this pure strategy is optimal in the set of all mixed strategies. For this we compute

$$\int_{\mathcal{E}} \frac{m_{t,e}}{m_{\mu,e}} \nu_1(de) = \int_{\mathcal{E}} \exp\left(\frac{\mu^2 - t^2 + 2e(t - \mu)}{2\sigma_1^2}\right) \nu_1(de) = \exp\left(\frac{(t - \mu)^2(\sigma_2^2 - \sigma_1^2)}{2\sigma_1^4}\right)$$

and check that it is less than 1 when $\sigma_2 \leq \sigma_1$. Hence, if the fluctuations of the environment are small ($\sigma_2 \leq \sigma_1$) then the pure strategy δ_μ is optimal.

Gaussian strategies $p = \mathcal{N}(\mu_p, \sigma_p^2)$ yield easy computations: recalling that a Gaussian mixture of Gaussian distributions is again Gaussian, we compute

$$m_{p,e} = \int_{\mathcal{T}} m_{t,e} p(dt) = \frac{C}{\sqrt{2\pi(\sigma_1^2 + \sigma_p^2)}} \exp\left(-\frac{(\mu_p - e)^2}{2(\sigma_1^2 + \sigma_p^2)}\right),$$

and then

$$\gamma(p) = \int_{\mathcal{E}} \log(m_{t,e})\nu_1(de) = \log C - \frac{1}{2} \log(2\pi(\sigma_1^2 + \sigma_p^2)) - \frac{(\mu - \mu_p)^2 + \sigma_2^2}{2(\sigma_1^2 + \sigma_p^2)}.$$

We find that $\gamma(p)$ is maximal (among Gaussian strategies) for $\mu_p = \mu$ and

$$\sigma_p^2 = \begin{cases} 0 & \text{if } \sigma_2^2 \leq \sigma_1^2, \\ \sigma_2^2 - \sigma_1^2 & \text{if } \sigma_2^2 > \sigma_1^2. \end{cases}$$

We have seen that this is the optimal strategy in the case when $\sigma_2^2 \leq \sigma_1^2$. This is still the case when $\sigma_2^2 > \sigma_1^2$: indeed, we compute, for $t \in \mathcal{T}$ and $p^* = \mathcal{N}(\mu, \sigma_2^2 - \sigma_1^2)$,

$$\int_{\mathcal{E}} \frac{m_{t,e}}{m_{p^*,e}} \nu_1(de) = \int_{\mathcal{E}} \frac{1}{C} m_{t,e} de = 1,$$

and this characterizes the optimal strategy according to Proposition 3.

We have so far proved the following result.

Proposition 6. *The optimal strategy is*

$$p^* = \begin{cases} \delta_\mu & \text{if } \sigma_2^2 \leq \sigma_1^2, \\ \mathcal{N}(\mu, \sigma_2^2 - \sigma_1^2) & \text{if } \sigma_2^2 \geq \sigma_1^2, \end{cases} \tag{4}$$

and the optimal growth rate is

$$\gamma^* = \begin{cases} \log C - \frac{1}{2} \log(2\pi\sigma_1^2) - \frac{\sigma_2^2}{2\sigma_1^2} & \text{if } \sigma_2^2 \leq \sigma_1^2, \\ \log C - \frac{1}{2} \log(2\pi\sigma_2^2) - \frac{1}{2} & \text{if } \sigma_2^2 \geq \sigma_1^2. \end{cases} \tag{5}$$

An interesting quantity is the relative gain of the best mixed strategy over the best pure strategy: it gives an indication of the strength of the selection pressure on mixed as opposed to pure strategies. From the above computation, defining $\chi = \sigma_2^2/\sigma_1^2$, we obtain

$$\gamma(p^*) - \gamma(\delta_\mu) = \begin{cases} 0 & \text{if } \chi \leq 1, \\ \frac{1}{2}(\chi - 1 - \log \chi) & \text{if } \chi \geq 1. \end{cases}$$

This is the log-ratio of the expected long-term growth rates of individuals playing the different types of strategy. It is nondecreasing with respect to χ : when the environmental variance σ_2^2 is large compared to σ_1^2 , there is a strong advantage in playing a mixed strategy.

Next we illustrate the phenomenon discussed in Section 3.3 when polymorphism is a necessary condition for survival. This happens when the optimal pure strategy δ_μ leads to almost-sure extinction of the population ($\gamma(\delta_\mu) \leq 0$), whereas the optimal mixed strategy p^* allows the population to survive ($\gamma^* > 0$). This phenomenon occurs when $\sigma_2 > \sigma_1$ for every C belonging to the following nonempty interval:

$$\frac{1}{2} \log(2\pi\sigma_2^2) + \frac{1}{2} < \log C \leq \frac{1}{2} \log(2\pi\sigma_1^2) + \frac{\sigma_2^2}{2\sigma_1^2}.$$

4. The nonhereditary case with sensing mechanism

We now study the case of nonhereditary traits when some sensing mechanism is available, that is, the trait distribution of the offspring does not depend on the trait of the parent, but does depend on the environment because the individuals get some information about the environment they evolve in and are able to suitably adapt the trait distribution of their offspring. We thus suppose in this section that Z_n evolves according to model (2) with $\pi_{t,e} \equiv p_e$ for some family $\bar{p} = (p_e)_{e \in \mathcal{E}}$ of distributions on \mathcal{T} , and let $\pi = \pi(\bar{p})$ be the corresponding product distribution.

4.1. Reduction to a simple BPRE

In a similar way as in the nonhereditary case with no sensing, the nonhereditary assumption makes the structure of the population very simple: the trait distribution at time n is given by the strategy the parents followed at time $n - 1$ in environment ω_{n-1} , that is, $p_{\omega_{n-1}}$. This is expressed in the following lemma.

Lemma 2. *For any $n \geq 1$, the population structure is conditionally independent of the past population process given the size of the population and the environment, i.e.*

$$Z_n \Big| (|Z_n|, \omega_{n-1}) \perp\!\!\!\perp (Z_0, \dots, Z_{n-1}) \Big| (|Z_n|, \omega_{n-1}).$$

Proof. It is easily seen from the assumptions on the model (2) and from the nonhereditary assumption $\pi = \pi(\bar{p})$ that the distribution of Z_n given (Z_0, \dots, Z_{n-1}) , $|Z_n|$, and ω_{n-1} is equal to the distribution of $\sum_{i=1}^{|Z_n|} 1_{\tau_i}$, where τ_i is an i.i.d. sequence with distribution $p_{\omega_{n-1}}$. This distribution does not depend on (Z_0, \dots, Z_{n-1}) , proving the conditional independence.

Lemma 2 implies that, given the environmental sequence ω , the distribution of the population process $(Z_n)_{n \geq 1}$ is easily recovered from the size process $(|Z_n|)_{n \geq 1}$. Once again, this latter process turns out to be a simple BPRE and this allows us to compute the performance $\gamma(\bar{p})$ of the strategy \bar{p} . Let $\omega^{(2)} = ((\omega_{n-1}, \omega_n))_{n \geq 1}$ denote the pair environment with values in \mathcal{E}^2 . It is also stationary and ergodic, and we denote by ν_2 its stationary distribution, which is the distribution of the pair (ω_1, ω_2) .

Proposition 7. *The size process $(|Z_n|)_{n \geq 1}$ is a simple branching process in environment $\omega^{(2)}$ with offspring distribution*

$$\Upsilon_{\bar{p}, (e_1, e_2)} = \int_{\mathcal{T}} \Upsilon_{t, e_2} p_{e_1}(dt).$$

Conditionally on ω , the expected population size at time n is

$$E_{\omega}[|Z_n|] = E_{\omega_0}[|Z_1|] \prod_{k=1}^{n-1} m_{p_{\omega_{k-1}, \omega_k}},$$

where $m_{p_{e_1, e_2}} = \int_{\mathcal{T}} m_{t, e_2} p_{e_1}(dt)$ is the first moment of $\Upsilon_{\bar{p}, (e_1, e_2)}$. Suppose that

$$\gamma(\bar{p}) = \int_{\mathcal{E}^2} \log(m_{p_{e_1, e_2}}) \nu_2(de_1, de_2)$$

exists. Then

$$\lim_{n \rightarrow \infty} n^{-1} \log E_{\omega}[|Z_n|] = \gamma(\bar{p}) \quad a.s.$$

Proof. According to Lemma 2, given $(|Z_1|, \dots, |Z_n|)$ and ω_{n-1} , the population Z_n has the same distribution as $\sum_{i=1}^{|Z_n|} 1_{\tau_i}$, where τ_i is an i.i.d. sequence with distribution $p_{\omega_{n-1}}$. Intuitively, the i th individual has type τ_i chosen randomly on \mathcal{T} with distribution $p_{\omega_{n-1}}$. The size of the next generation $|Z_{n+1}|$ is then $\sum_{i=1}^{|Z_n|} \xi_{n,i}^{\tau_i, \omega_n}$, where $\xi_{n,i}^{\tau_i, \omega_n}$ is the offspring of the i th individual of type τ_i in environment ω_n . From this two-step procedure, that is, the random choice of the traits t according to $p_{\omega_{n-1}}$ and the reproduction with random offspring in environment ω_n , we obtain the effective offspring distribution $\Upsilon_{p_{\omega_{n-1}, \omega_n}}$ in environment $\omega_n^{(2)} = (\omega_{n-1}, \omega_n)$: it is the mixture of the offspring distributions Υ_{t, ω_n} , with mixing distribution $p_{\omega_{n-1}}$. Other properties are proved as in Proposition 1.

Note that the classification theorem, Theorem 1, applies in this case as well and gives criteria for the extinction or explosion of the population.

4.2. Optimal strategies

We now describe the set of optimal strategies when sensing is allowed. It is also interesting to evaluate the gain between the optimal strategies with and without sensing mechanisms. Optimality when sensing mechanisms are allowed will be denoted with a double asterisk, whereas we keep a single asterisk for optimality without sensing. Let $\gamma^{**} = \sup\{\gamma(\bar{p}); \bar{p} \in \mathcal{P}(\mathcal{T})^\mathcal{E}\}$ be the optimal growth rate when sensing is allowed, and let $\mathcal{P}^{**} = \{\bar{p} \in \mathcal{P}(\mathcal{T})^\mathcal{E}; \gamma(\bar{p}) = \gamma^{**}\}$ be the set of optimal strategies. Let $\nu_{e_1}(de_2)$ be the conditional distribution of ω_2 given $\omega_1 = e_1$, so $\nu_2(de_1, de_2) = \nu_1(de_1)\nu_{e_1}(de_2)$. Note that conditional distributions are well defined since \mathcal{E} is assumed to be a Polish space. From the previous section, the set of optimal strategies without sensing is denoted by \mathcal{P}^* . It depends implicitly on the environment distribution ν_1 and we write $\mathcal{P}^*(\nu_1)$ to emphasize this dependence. Suppose that the assumptions of Proposition 2 hold. Then optimal strategies with sensing are related to optimal strategies without sensing in the following way.

Proposition 8. *A strategy \bar{p} is optimal if and only if*

$$p_{e_1} \in \mathcal{P}^*(\nu_{e_1}) \quad \nu_1(de_1)\text{-almost everywhere.}$$

*If conditions (C1)–(C2) hold then optimal strategies exist and form a closed convex set, i.e. \mathcal{P}^{**} is nonempty, closed, and convex.*

*If the family $\mathcal{M} = \{e \mapsto m_{t,e}; t \in \mathcal{T}\}$ is linearly independent, there is a unique optimal strategy \bar{p}^{**} .*

Proof. Using the explicit formula for $\gamma(\bar{p})$ given in Proposition 7 and conditional probabilities, we compute

$$\begin{aligned} \gamma(\bar{p}) &= \int_{\mathcal{E}^2} \log(m_{p_{e_1},e_2})\nu_2(de_1, de_2) \\ &= \int_{\mathcal{E}} \nu_1(de_1) \int_{\mathcal{E}} \log(m_{p_{e_1},e_2})\nu_{e_1}(de_2) \\ &= \int_{\mathcal{E}} \gamma(p_{e_1}, \nu_{e_1})\nu_1(de_1), \end{aligned}$$

where $\gamma(p_{e_1}, \nu_{e_1}) = \int_{\mathcal{E}} \log(m_{p_{e_1},e_2})\nu_{e_1}(de_2)$ is the growth rate associated with strategy p_{e_1} in environment ν_{e_1} . Hence, $\gamma(\bar{p})$ is optimal if we choose p_{e_1} so that $\gamma(p_{e_1}, \nu_{e_1})$ is maximal, i.e. $p_{e_1} \in \mathcal{P}^*(\nu_{e_1})$ a.s. The properties of \mathcal{P}^{**} and the uniqueness follow similarly as in the \mathcal{P}^* case (cf. Proposition 2).

Then, the characterization of optimal strategies in the no-sensing case given in Proposition 3 directly extends to strategies with sensing as follows.

Proposition 9. *A strategy with sensing $\bar{p} \in \mathcal{P}(\mathcal{T})^\mathcal{E}$ is optimal if and only if*

$$\int_{\mathcal{E}} \frac{m_{t,e_2}}{m_{p_{e_1},e_2}}\nu_{e_1}(de_2) \leq 1 \quad \text{for all } t \in \mathcal{T} \text{ and } \nu_1(de_1)\text{-almost everywhere.}$$

Proof. The characterization follows directly from Proposition 3 and Proposition 8.

The corollary below is interesting because it states that no gain has to be expected from sensing mechanisms if the environment has some independence property.

Corollary 2. *Suppose that ω_1 and ω_2 are independent (i.e. $\nu_2 = \nu_1 \otimes \nu_1$). Then the optimal growth rates with and without sensing are equal, i.e. $\gamma^* = \gamma^{**}$.*

Proof. In this product case, the conditional distributions are trivial, i.e. $\nu_{e_1} \equiv \nu_1$ a.s. Hence, an optimal strategy is such that $p_{e_1}^{**} \in \mathcal{P}^*(\nu_1)$ and, hence, $\gamma(p_{e_1}^{**}, \nu_1) = \gamma^*$. Integrating with respect to $\nu_1(de_1)$, we obtain

$$\gamma^{**} = \int_{\mathcal{E}} \gamma(p_{e_1}^{**}, \nu_1) \nu_1(de_1) = \gamma^*.$$

This proves the result.

Note that this result is rather intuitive: the sensing mechanism provides some information to the individual about the current environment state, but, from the independence property, this is not useful for inference to the future environment state; hence, the information is useless for deciding which traits will be well suited to the next environment.

4.3. Example: finite-dimensional case continued

The results for optimal strategies without sensing developed in Section 3.4 together with Proposition 8 allow us to easily deduce the following properties for optimal strategies with sensing in the finite-dimensional case.

Proposition 10. *Suppose that $\text{span } \mathcal{M}$ has dimension d and that conditions (C1)–(C3) and (H) hold. Then there exists a unique optimal strategy with sensing $\bar{p}^{**} \in \mathcal{P}(\mathcal{T})^{\mathcal{E}}$ such that*

$$p^{**}(e) = p^*(\nu_e),$$

where $p^*(\nu_e)$ is the optimal strategy without sensing when the environment has marginal distribution ν_e .

Recall further from Proposition 5 that $p^*(\nu_e)$ is a discrete probability measure on \mathcal{T} with at most d extremal types (which may depend on e). See Section 1.1.2 for more examples.

4.4. Example: Haccou and Iwasa’s example continued

Here we continue the example by Haccou and Iwasa given in Section 3.5. Recall that the environment is given by a Gaussian stationary ergodic sequence $\omega = (\omega_n)_{n \geq 0}$ with stationary distribution $\nu_1 = \mathcal{N}(\mu, \sigma_2^2)$. Let $\rho \in (-1, 1)$ be the pair correlation $\rho = \text{corr}(\omega_0, \omega_1)$. The case in which $\rho = 0$ corresponds to independent environments, i.e. $\nu_2 = \nu_1 \otimes \nu_1$. Otherwise dependence holds and standard Gaussian computations give the conditional distribution $\nu_{e_1} = \mathcal{N}(\mu + \rho(e_1 - \mu), (1 - \rho^2)\sigma_2^2)$. A particular realization of such a sequence is the Ornstein–Uhlenbeck sequence defined by

$$\omega_0 = \mu + \sigma_2 \eta_0, \quad \omega_n = \mu + \rho(\omega_{n-1} - \mu) + \sqrt{1 - \rho^2} \sigma_2 \eta_n, \quad n \geq 1,$$

for i.i.d. standard normal innovations $(\eta_n)_{n \geq 0}$ (Gaussian white noise).

We showed in Section 3.5 that the optimal strategy without sensing is Gaussian (possibly degenerate) and we gave explicit formulae for the parameters (4) and for the corresponding growth rate (5). Using Proposition 8, we now deduce the optimal strategy when sensing is allowed.

Proposition 11. *There exists a unique optimal strategy with sensing, which is denoted by \bar{p}^{**} and satisfies $\nu_1(de)$ -almost everywhere:*

$$p_e^{**} = \begin{cases} \delta_{\mu + \rho(e - \mu)} & \text{if } (1 - \rho^2)\sigma_2^2 \leq \sigma_1^2, \\ \mathcal{N}(\mu + \rho(e - \mu), (1 - \rho^2)\sigma_2^2 - \sigma_1^2) & \text{if } (1 - \rho^2)\sigma_2^2 \geq \sigma_1^2. \end{cases}$$

The corresponding optimal growth rate is given by

$$\gamma^{**} = \begin{cases} \log C - \frac{1}{2} \log(2\pi\sigma_1^2) - \frac{(1-\rho^2)\sigma_2^2}{2\sigma_1^2} & \text{if } (1-\rho^2)\sigma_2^2 \leq \sigma_1^2, \\ \log C - \frac{1}{2} \log(2\pi(1-\rho^2)\sigma_2^2) - \frac{1}{2} & \text{if } (1-\rho^2)\sigma_2^2 \geq \sigma_1^2. \end{cases}$$

Finally, we can evaluate the relative gain of strategies with sensing over strategies without sensing: it gives an indication of the benefit that can be expected from sensing mechanisms. Let $\chi = \sigma_2^2/\sigma_1^2$. Then

$$\gamma^{**} - \gamma^* = \begin{cases} \frac{1}{2}\rho^2\chi & \text{if } \chi \leq 1, \\ \frac{1}{2} \log \chi - \frac{1}{2}(1-\rho^2)\chi + \frac{1}{2} & \text{if } 1 \leq \chi \leq (1-\rho^2)^{-1}, \\ -\frac{1}{2} \log(1-\rho^2) & \text{if } \chi \geq (1-\rho^2)^{-1}. \end{cases}$$

It is worth noting that this is an increasing function of the square correlation ρ^2 : this indicates that the more correlated the random environment, the more useful the sensing mechanisms. The intuitive idea is that higher correlations allow for more accurate prevision for the next environment and, hence, for a better-fitted offspring trait distribution in the environment to come.

5. The hereditary case

In the hereditary case, the trait distribution $\pi_{t,e} \in \mathcal{P}(\mathcal{T})$ may depend on the trait of the parents. This dependency makes the study of the Lyapounov exponent $\gamma = \gamma(\pi)$ much more difficult because no reduction to a simple branching process in a random environment is available. We have no explicit formula for $\gamma(\pi)$ in this case and determining the optimal strategy π^* and the optimal growth rate $\gamma(\pi^*)$ might be very challenging.

Nevertheless, we propose an interesting representation of the ‘finite time’ growth rate

$$\gamma_n(\omega, \pi_0) = n^{-1} \log E_{\omega, \pi_0}[|Z_n|]$$

in environment ω with the initial population consisting of a single individual of random trait with distribution π_0 . The representation is in terms of a functional of the \mathcal{T} -valued Markov chain $T = (T_n)_{n \geq 0}$ in environment ω such that

$$\begin{aligned} P_{\pi_0, \omega}(T_0 \in \cdot) &= \pi_0(\cdot), \\ P_{\pi_0, \omega}(T_k \in \cdot \mid T_0 = t_0, \dots, T_{k-1} = t_{k-1}) &= \pi_{t_{k-1}, \omega_{k-1}}(\cdot), \quad 1 \leq k \leq n. \end{aligned} \tag{6}$$

The Markov chain T_n in random environment ω is time heterogeneous because the transitions depend on time n through the value of the environment ω_n . However, in the no-sensing case when the transitions $\pi_{t,e} \equiv \pi_t$ do not depend on e , the Markov chain T is time homogeneous. The result is given as follows.

Proposition 12. *The finite time growth rate in environment ω with the initial population consisting of a single individual whose trait is distributed according to π_0 is given by*

$$\gamma_n(\omega, \pi_0) = n^{-1} \log E_{\omega, \pi_0}[M_n(T, \omega)],$$

where

$$M_n(T, \omega) = \prod_{k=0}^{n-1} m_{T_k, \omega_k}.$$

For the sake of clarity and conciseness, this proposition will be proved together with Theorem 2 below. To the best of the authors' knowledge, there is no simple way to deal with the asymptotic behavior of γ_n and this issue might be challenging. We provide a lower bound for the growth rate $\gamma_n(\omega, \pi_0)$ that might be more tractable. The lower bound for $\gamma_n(\omega, \pi_0)$ is obtained using Jensen's inequality (with the concave function \log):

$$\gamma_n(\omega, \pi_0) = n^{-1} \log E_{\omega, \pi_0}[M_n(T, \omega)] \geq n^{-1} E_{\omega, \pi_0}[\log M_n(T, \omega)].$$

The advantage here is that the lower bound

$$n^{-1} E_{\omega, \pi_0}[\log M_n(T, \omega)] = n^{-1} \sum_{k=0}^{n-1} E_{\omega, \pi_0}[\log m_{T_k, \omega_k}]$$

can be written as an additive functional of the environment ω and we can then use the ergodic theorem to control the convergence. Note that this technique can be refined using changes of measure. Let $\tilde{\pi}_{t,e}$ be any kernel family, and denote by \tilde{T} the \mathcal{T} -valued Markov chain in environment ω starting from distribution π_0 and with transitions given by $\tilde{\pi}_{t,e}$ (given by equations similar to (6)). We suppose that $\pi_{t,e}$ is absolutely continuous with respect to $\tilde{\pi}_{t,e}$, i.e. $\pi_{t,e}(dt') = f_{t,e}(t')\tilde{\pi}_{t,e}(dt')$, where $f_{t,e}$ stands for the density of $\pi_{t,e}$ with respect to $\tilde{\pi}_{t,e}$. Then, using changes of measure, we have

$$E_{\omega, \pi_0}[M_n(T, \omega)] = E_{\omega, \pi_0}[\tilde{M}_n(\tilde{T}, \omega)],$$

where

$$\tilde{M}_n(\tilde{T}, e) = \prod_{k=0}^{n-1} m_{\tilde{T}_k, \omega_k} f_{\tilde{T}_k, \omega_k}(\tilde{T}_{k+1}).$$

Using this, the lower bound becomes

$$\gamma_n(\omega, \pi_0) = n^{-1} E_{\omega, \pi_0}[\log \tilde{M}_n(\tilde{T}, \omega)] \geq n^{-1} \sum_{k=0}^{n-1} E_{\omega, \pi_0}[\log(m_{\tilde{T}_k, \omega_k} f_{\tilde{T}_k, \omega_k}(\tilde{T}_{k+1}))].$$

6. Typical genealogies: a mean field approach

6.1. Convergence of the typical genealogy in the infinite population limit

As explained by Baake and Georgii in [4] and [8], the evolution of a branching population can be studied from two possible perspectives: either forward or backward in time. So far, we have focused on the first point of view and mainly studied the growth rate of the population after a large number of generations. By way of contrast, the backward or retrospective aspect of the population concerns the lineages extending back into the past from the presently living individuals and asks for the characteristics of the ancestors along such lineages. We now turn to this second perspective and wonder what is the typical lineage or genealogy (backward in time) of an individual chosen at random in the n th generation.

Some definitions are needed here. The correct formalism to keep track of the genealogy is the formalism of labeled rooted trees and forests, where the trees stand for the descendance of each

ancestor represented by a root, and labels keep track of the traits of the individuals. However, we keep this formalism to its minimum. Let \mathcal{G}_n denote the population at the n th generation. To each individual $g \in \mathcal{G}_n$, we associate its lineage or genealogy $\ell(g) = (t_0, \dots, t_n) \in \mathcal{T}^{n+1}$ with the interpretation that t_n is the trait of g and t_k is the trait of his ancestor in the k th generation \mathcal{G}_k , $0 \leq k \leq n - 1$. The typical genealogy is defined as the genealogy of an individual chosen at random in the n th generation. This obviously requires the n th generation to be nonempty, in which case we adopt the convention that the typical genealogy is \emptyset . The distribution of the typical genealogy is given by

$$\pi_n = \begin{cases} \frac{1}{\text{card}\mathcal{G}_n} \sum_{g \in \mathcal{G}_n} \delta_{\ell(g)} & \text{if } \mathcal{G}_n \neq \emptyset, \\ \delta_{\emptyset} & \text{if } \mathcal{G}_n = \emptyset. \end{cases}$$

This is a random measure on $\mathcal{T}^{n+1} \cup \{\emptyset\}$. We denote by $P_{N,\pi_0,\omega}$ the probability measure corresponding to a population evolving in environment ω , starting a time 0 from N individuals with i.i.d. traits with distribution π_0 . In the following, we focus on the typical genealogy in the infinite population limit $N \rightarrow \infty$. Recall the definition of the Markov chain $T = (T_n)_{n \geq 0}$ given in (6). We have the following mean field result.

Theorem 2. *Under the probability $P_{N,\pi_0,\omega}$, the typical genealogy distribution π_n a.s. weakly converges as $N \rightarrow \infty$ to the distribution $\hat{\pi}_{n,\pi_0,\omega}$, defined by*

$$\hat{\pi}_{n,\pi_0,\omega}(A) = \frac{E_{\omega,\pi_0}[M_n(T, \omega)\mathbf{1}_{(T_0,\dots,T_n) \in A}]}{E_{\omega,\pi_0}[M_n(T, \omega)]}, \quad A \subset \mathcal{T}^{n+1}.$$

Recall from Proposition 12 that $E_{\omega,\pi_0}[M_n(T, \omega)] = \exp(n\gamma_n(\pi_0, \omega))$ is the mean number of individuals in the n th generation of a population evolving in environment ω and starting from a single individual of random trait with distribution π_0 .

Proofs of Theorem 2 and Proposition 12. Let $A = A_0 \times \dots \times A_n$ be a product subset of \mathcal{T}^{n+1} . The number of individuals in the n th generation with genealogy in A is

$$\mathcal{N}_n(A) = \sum_{g \in \mathcal{G}_n} \delta_{\ell(g)}(A),$$

and $\mathcal{N}_n(\mathcal{T}^{n+1})$ denotes the total number of individuals in \mathcal{G}_n . We can view \mathcal{N}_n as a nonnormalized measure and π_n as the probability measure associated with \mathcal{N}_n through the relation

$$\pi_n = \frac{1}{\mathcal{N}_n(\mathcal{T}^{n+1})} \mathbf{1}_{\{\mathcal{N}_n(\mathcal{T}^{n+1}) > 0\}} \mathcal{N}_n + \mathbf{1}_{\{\mathcal{N}_n(\mathcal{T}^{n+1}) = 0\}} \delta_{\emptyset}.$$

From the branching property, the distribution of $\mathcal{N}_n(A)$ under $P_{N,\pi_0,\omega}$ is equal to the sum of N independent copies $\sum_{i=1}^N \mathcal{N}_n^{(i)}(A)$ under $P_{1,\pi_0,\omega}^{\otimes N}$. As a consequence of the weak law of large numbers, the distribution of $(1/N)\mathcal{N}_n(A)$ under $P_{N,\pi_0,\omega}$ weakly converges to $E_{1,\pi_0,\omega}[\mathcal{N}_n(A)]$. The results also hold for $A = \mathcal{T}^{n+1}$, and taking the quotient, we see that, under $P_{N,\pi_0,\omega}$, $\pi_n(A)$ weakly converges to

$$\hat{\pi}_{n,\pi_0,\omega}(A) = \frac{E_{1,\pi_0,\omega}[\mathcal{N}_n(A)]}{E_{1,\pi_0,\omega}[\mathcal{N}_n(\mathcal{T}^{n+1})]},$$

provided that the denominator is nonzero. Note that

$$\gamma_n(\omega, \pi_0) = n^{-1} \log E_{1,\pi_0,\omega}[\mathcal{N}_n(\mathcal{T}^{n+1})].$$

Theorem 2 and Proposition 12 (letting $A = \mathcal{T}^n$) are then a consequence of

$$\begin{aligned} E_{1,\pi_0,\omega}(\mathcal{N}_n(A)) &= \int_A \pi_0(dt_0) \prod_{k=0}^{n-1} m_{t_k,\omega_k} \pi_{t_k,\omega_k}(dt_{k+1}) \\ &= E_{\omega,\pi_0}[M_n(T, \omega) \mathbf{1}_{(T_0,\dots,T_n) \in A}]. \end{aligned}$$

This last relation is proved by induction. Individuals in the n th generation with genealogy in A are the offspring with traits in A_n of individuals in the $(n - 1)$ th generation with genealogy in $A_0 \times \dots \times A_{n-1}$. This yields

$$\begin{aligned} E_{1,\pi_0,\omega}[\mathcal{N}_n(A)] &= \int_{A_{n-1} \times A_n} E_{1,\pi_0,\omega}[\mathcal{N}_{n-1}(A_0 \times \dots \times A_{n-2} \times dt_{n-1})] \\ &\quad \times m_{t_{n-1},\omega_{n-1}} \pi_{t_{n-1},\omega_{n-1}}(dt_n). \end{aligned}$$

For $n = 0$, $E_{1,\pi_0,\omega}[\mathcal{N}_0(A_0)] = \pi_0(A)$. It remains to note that $E_{1,\pi_0,\omega}[\mathcal{N}_n(\mathcal{T}^{n+1})] \neq 0$, and this implies that the mass of \emptyset vanishes in the limit.

6.2. The typical genealogy in the nonhereditary case

When the traits are nonhereditary, the trait distributions $\pi_{t,e}$ do not depend on t and the mean field typical genealogy distribution $\hat{\pi}_{n,\omega,\pi_0}$ has a very simple form. The following proposition is given in the context of a population with a sensing mechanism $\pi = \pi((p_e)_{e \in \mathcal{E}})$. The no-sensing case corresponds to the particular case when $p_e \equiv p$.

Proposition 13. *In the nonhereditary case, the typical genealogy distribution $\hat{\pi}_{n,\omega,\pi_0}$ is the product measure on \mathcal{T}^{n+1} defined by*

$$\hat{\pi}_{n,\omega,\pi_0}(dt_0, \dots, dt_n) = \bigotimes_{i=0}^n \hat{\pi}_{i,\omega_i}(dt_i),$$

where

$$\begin{aligned} \hat{\pi}_{0,\omega_0}(dt_0) &= \frac{m_{t_0,\omega_0}}{m_{\pi_0,\omega_0}} \pi_0(dt_0), \\ \hat{\pi}_{i,\omega_i}(dt_i) &= \frac{m_{t_i,\omega_i}}{m_{p_{\omega_{i-1}},\omega_i}} p_{\omega_{i-1}}(dt_i), \quad 1 \leq i \leq n - 1, \\ \hat{\pi}_{n,\omega_n}(dt_n) &= p_{\omega_{n-1}}(dt_n). \end{aligned}$$

Proof. In the nonhereditary case, the Markov chain in a random environment ω defined by (6) is simple because, conditionally on the environment ω , the random variables T_0, \dots, T_n are independent with T_0 distributed as π_0 and, for $1 \leq i \leq n$, T_i distributed according to $p_{\omega_{i-1}}$. Let $A = A_0 \times \dots \times A_n \subset \mathcal{T}^{n+1}$. Using independence and Theorem 2, we compute

$$\begin{aligned} \hat{\pi}_{n,\pi_0,\omega}(A) &= \frac{E_{\omega,\pi_0}[M_n(T, \omega) \mathbf{1}_{(T_0,\dots,T_n) \in A}]}{E_{\omega,\pi_0}[M_n(T, \omega)]} \\ &= \frac{E_{\omega,\pi_0}[m_{T_0,\omega_0} \mathbf{1}_{\{T_0 \in A_0\}}] \times \dots \times E_{\omega,\pi_0}[m_{T_{n-1},\omega_{n-1}} \mathbf{1}_{\{T_{n-1} \in A_{n-1}\}}] E_{\omega,\pi_0}[\mathbf{1}_{\{T_n \in A_n\}}]}{E_{\omega,\pi_0}[m_{T_0,\omega_0}] \times \dots \times E_{\omega,\pi_0}[m_{T_{n-1},\omega_{n-1}}]}. \end{aligned}$$

This proves the independence property and gives the marginal distributions

$$\hat{\pi}_i(A_i) = \frac{E_{\omega,\pi_0}[m_{T_i,\omega_i} \mathbf{1}_{\{T_i \in A_i\}}]}{E_{\omega,\pi_0}[m_{T_i,\omega_i}]}, \quad 1 \leq i \leq n - 1.$$

The interpretation of the above proposition is the following. The mean field typical genealogy in environment ω consists of independent traits $\hat{T}_0, \dots, \hat{T}_n$, where

- the distribution of \hat{T}_0 is a biased version of $\pi_0(dt_0)$ with bias function equal to m_{t_0, ω_0} , the mean number of offspring of an individual of type t_0 ,
- the distribution of \hat{T}_i is a biased version of $p_{\omega_i}(dt_i)$ with bias function equal to m_{t_i, ω_i} ,
- the distribution of \hat{T}_n is $p_{\omega_n}(dt_n)$ (there is no bias because the offspring of the last generation \mathcal{G}_n is not involved, since we only consider the population until time n).

6.3. The typical genealogy in a hereditary case: Haccou and Isawa’s example continued

Recall that $\mathcal{T} = \mathcal{E} = \mathbb{R}$, ω is a Gaussian stationary ergodic sequence with stationary distribution $\mathcal{N}(\mu, \sigma_2^2)$, and $m_{t,e}$ is given by

$$m_{t,e} = \frac{C}{\sqrt{2\pi\sigma_1^2}} \exp\left(-\frac{(t-e)^2}{2\sigma_1^2}\right).$$

It remains to give the trait distributions $\pi_{t,e}$ and π_0 . To make the model explicitly solvable, we require the Markov chain $T = (T_0, \dots, T_n)$ in environment ω to be multivariate Gaussian. This means that the transitions $\pi_{t,e}$ are of the form $\pi_{t,e} = \mathcal{N}(\alpha_e t + \beta_e, \theta_e)$ with

$$E[T_1 \mid \omega_0 = e, T_0 = t] = \alpha_e + \beta_e t \quad \text{and} \quad \text{var}[T_1 \mid \omega_0 = e, T_0 = t] = \theta_e^2.$$

Let $\pi_0 = \mathcal{N}(\mu_0, s_0^2)$ be the initial distribution. Alternatively, in environment ω we have the representation

$$T_0 = \mu_0 + s_0 N_0, \quad T_{k+1} = \alpha_{\omega_k} + \beta_{\omega_k} T_k + \theta_{\omega_k} N_{k+1}, \quad 0 \leq k \leq n-1,$$

where N_0, \dots, N_n are independent standard normal variables. We introduce the $(n+1) \times 1$ vectors

$$T = \begin{pmatrix} T_0 \\ T_1 \\ \vdots \\ T_n \end{pmatrix}, \quad N = \begin{pmatrix} N_0 \\ N_1 \\ \vdots \\ N_n \end{pmatrix}, \quad A_\omega = \begin{pmatrix} \mu_0 \\ \alpha_{\omega_0} \\ \vdots \\ \alpha_{\omega_{n-1}} \end{pmatrix},$$

and the $(n+1) \times (n+1)$ matrices

$$C_\omega = \begin{pmatrix} 0 & 0 & \dots & 0 & 0 \\ \beta_{\omega_0} & 0 & \dots & 0 & 0 \\ 0 & \beta_{\omega_1} & 0 & 0 & 0 \\ 0 & 0 & \ddots & 0 & 0 \\ 0 & 0 & \dots & \beta_{\omega_{n-1}} & 0 \end{pmatrix}, \quad S_\omega = \begin{pmatrix} s_0^2 & 0 & \dots & 0 & 0 \\ 0 & \theta_{\omega_0}^2 & \dots & 0 & 0 \\ 0 & 0 & \theta_{\omega_1}^2 & 0 & 0 \\ 0 & 0 & \dots & \ddots & 0 \\ 0 & 0 & \dots & 0 & \theta_{\omega_{n-1}}^2 \end{pmatrix}.$$

With such notation, the recursive relation turns into

$$T = A_\omega + C_\omega T + S_\omega N,$$

and this yields

$$T = (\text{Id} - C_\omega)^{-1}(A_\omega + S_\omega N).$$

We deduce the mean μ and covariance matrix Σ for the Gaussian vector T :

$$\mu = (\text{Id} - C_\omega)^{-1} A_\omega \quad \text{and} \quad \Sigma = (\text{Id} - C_\omega)^{-1} S_\omega (\text{Id} - C_\omega^\top)^{-1}.$$

Suppose that S_ω is invertible. Then T has density

$$f_T(t) = (2\pi)^{-(n+1)/2} \det(\Sigma)^{-1/2} \exp\left(-\frac{1}{2}(t - \mu)^\top \Sigma^{-1}(t - \mu)\right).$$

Denote by $\hat{T} = (\hat{T}_0, \dots, \hat{T}_n)$ a random vector with distribution $\hat{\pi}_{n,\omega,\pi_0}$. According to Theorem 2, \hat{T} is a biased version of T and has a density given by

$$\begin{aligned} f_{\hat{T}}(t) &= e^{-n\gamma_n(\omega,\pi_0)} M_n(t, \omega) f_T(t) \\ &= e^{-n\gamma_n(\omega,\pi_0)} C^n \sigma_1^{-n} (2\pi)^{-n-1/2} \det(\Sigma)^{-1/2} \\ &\quad \times \exp\left(-\frac{1}{2}(t - \mu)^\top \Sigma^{-1}(t - \mu) - \frac{1}{2\sigma_1^2} \sum_{i=0}^{n-1} (\omega_i - t_i)^2\right). \end{aligned}$$

Introducing $J_{n,1}$ as the diagonal matrix $J_{n,1} = (\delta_{0 \leq i=j \leq n-1})_{0 \leq i,j \leq n}$ and $V_{n,\omega}$ as the column vector $V_{n,\omega} = (\omega_0, \dots, \omega_{n-1}, 0)^\top$, the last exponential factor becomes

$$\exp\left(-\frac{1}{2}t^\top (\Sigma^{-1} + \sigma_1^{-2} J_{n,1})t + (\Sigma^{-1}\mu + \sigma_1^{-2} V_{n,\omega})^\top t - \frac{1}{2}\mu^\top \Sigma^{-1}\mu - \frac{1}{2}\sigma_1^{-2} V_{n,\omega}^\top V_{n,\omega}\right).$$

We recognize that $f_{\hat{T}}$ is a multivariate Gaussian density with mean $\hat{\mu}$ and covariance matrix $\hat{\Sigma}$ of the form

$$f_{\hat{T}}(t) = (2\pi)^{-(n+1)/2} \det(\hat{\Sigma})^{-1/2} \exp\left(-\frac{1}{2}(t - \hat{\mu})^\top \hat{\Sigma}^{-1}(t - \hat{\mu})\right).$$

Comparing both expressions, we obtain after simplification

$$\hat{\Sigma} = (\Sigma^{-1} + \sigma_1^{-2} J_{n,1})^{-1}, \tag{7}$$

$$\hat{\mu} = \hat{\Sigma}(\Sigma^{-1}\mu + \sigma_1^{-2} V_{n,\omega}) = (\Sigma^{-1} + \sigma_1^{-2} J_{n,1})^{-1}(\Sigma^{-1}\mu + \sigma_1^{-2} V_{n,\omega}), \tag{8}$$

$$\begin{aligned} \gamma_n(\omega, \pi_0) &= \log \frac{C}{\sqrt{2\pi}\sigma_1} - \frac{1}{2n} \log \det(I_{n+1} + \sigma_1^{-2} J_{n,1} \Sigma) \\ &\quad + \frac{1}{2n} (\hat{\mu}^\top \hat{\Sigma}^{-1} \hat{\mu} - \mu^\top \Sigma \mu - \sigma_1^{-2} V_{n,\omega}^\top V_{n,\omega}). \end{aligned} \tag{9}$$

These computations prove the following result.

Proposition 14. *In the Haccou and Isawa model with Markov Gaussian environment, the typical genealogy distribution $\hat{\pi}_{n,\omega,\pi_0}$ is the Gaussian distribution on \mathbb{R}^{n+1} with mean $\hat{\mu}$ given by (8) and covariance matrix $\hat{\Sigma}$ given by (7). Furthermore, the finite time growth rate $\gamma_n(\omega, \pi_0)$ is given by (9).*

References

- [1] AFANASYEV, V. I., GEIGER, J., KERSTING, G. AND VATUTIN, V. A. (2005). Criticality for branching processes in random environment. *Ann. Prob.* **33**, 645–673.
- [2] ATHREYA, K. B. AND KARLIN, S. (1971). On branching processes with random environments. I. Extinction probabilities. *Ann. Math. Statist.* **42**, 1499–1520.
- [3] ATHREYA, K. B. AND KARLIN, S. (1971). Branching processes with random environment. II. Limit theorems. *Ann. Math. Statist.* **42**, 1843–1858.

- [4] BAAKE, E. AND GEORGII, H.-O. (2007). Mutation, selection, and ancestry in branching models: a variational approach. *J. Math. Biol.* **54**, 257–303.
- [5] BONNANS, J. F., GILBERT, J. C., LEMARÉCHAL, C. AND SAGASTIZÁBAL, C. A. (2006). *Numerical Optimization*, 2nd edn. Springer, Berlin.
- [6] GANDER, M. J., MAZZA, C. AND RUMMLER, H. (2007). Stochastic gene expression in switching environments. *J. Math. Biol.* **55**, 249–269.
- [7] GEIGER, J., KERSTING, G. AND VATUTIN, V. A. (2003). Limit theorems for subcritical branching processes in random environment. *Ann. Inst. H. Poincaré Prob. Statist.* **39**, 593–620.
- [8] GEORGII, H.-O. AND BAAKE, E. (2003). Supercritical multitype branching processes: the ancestral types of typical individuals. *Adv. Appl. Prob.* **35**, 1090–1110.
- [9] HACCOU, P. AND IWASA, Y. (1995). Optimal mixed strategies in stochastic environments. *Theoret. Pop. Biol.* **47**, 212–243.
- [10] KUSSEL, E. AND LEIBLER, S. (2005). Phenotypic diversity, population growth and information in fluctuating environments. *Science* **309**, 2075–2078.
- [11] SMITH, W. L. AND WILKINSON, W. E. (1969). On branching processes in random environments. *Ann. Math. Statist.* **40**, 814–827.
- [12] TANNY, D. (1977). Limit theorems for branching processes in a random environment. *Ann. Prob.* **5**, 100–116.
- [13] THATTAI, M. AND VAN OUDENAARDEN, A. (2004). Stochastic gene expression in fluctuating environments. *Genetics* **167**, 523–530.
- [14] VISCO, P., ALLEN, R. J., MAJUMDA, S. N. AND EVANS M. R. (2010). Switching and growth for microbial populations in catastrophic responsive environments. Preprint. Available at <http://arxiv.org/abs/0908.1351v2>.