Ancestral lineages in mutation-selection equilibria with moving optimum

Raphaël Forien, Jimmy Garnier, Florian Patout

ANCESTRAL LINEAGES IN MUTATION SELECTION EQUILIBRIA WITH MOVING OPTIMUM

RAPHAËL FORIEN, JIMMY GARNIER, AND FLORIAN PATOUT

Abstract. We investigate the evolutionary dynamics of a population structured in phenotype, subjected to trait dependent selection with a linearly moving optimum and an asexual mode of reproduction. Our model consists of a non-local and non-linear parabolic PDE. Our main goal is to measure the history of traits when the population stays around an equilibrium. We define an ancestral process based on the idea of *neutral fractions*. It allows us to derive quantitative information upon the evolution of diversity in the population along time. First, we study the long-time asymptotics of the ancestral process. We show that the very few fittest individuals drive adaptation. We then tackle the adaptive dynamics regime, where the effect of mutations is asymptotically small. In this limit, we provide an interpretation for the minimizer of some related optimization problem, an Hamilton Jacobi equation, as the *typical* ancestral lineage. We check the theoretical results against individual based simulations.

1. Introduction

We are interested in studying phenotypic lineages inside a population that stays at a mutation selection equilibrium, while also keeping pace with an environmental change. First, we detail the ecological model. The population is structured by a one–dimensional phenotypic trait, denoted by $x \in \mathbb{R}$. The density of individuals with trait x is denoted $f(t,x)$ at any time $t > 0$. We assume that the density function is solution to the following integro-differential equation:

(1.1)
$$\begin{cases} \partial_t f(t,x) + \left(\mu(x-ct) + (\beta - \mu_0)\displaystyle\int_{\mathbb{R}} f(t,x')dx'\right) f(t,x) = \beta \mathcal{B}(f(t,\cdot))(x) \ \text{ for } \ t>0,\, x \in \mathbb{R}, \\ f(0,x) = f_0(x), \quad x \in \mathbb{R}. \end{cases}$$

The operator $\mathcal{B}$ describes the apparition of new individuals via reproduction (birth) events occurring at rate β. It takes the following shape:

$$\mathcal{B}(g)(x) = \frac{1}{\sigma}\int_{\mathbb{R}} K\left(\frac{x-x'}{\sigma}\right) g(x')\,dx', \tag{1.2}$$

with $\sigma > 0$. K is a symmetric normalized probability density kernel, that encodes the shape of the deviation (for instance due to mutations) between the trait of an offspring and that of its parent, while σ^2 measures the variance of mutations. This operator can describe asexual reproduction.

The population is subject to trait dependent selection through an intrinsic mortality rate $\mu(x)$, that admits a minimal value at an *optimal trait*. The function μ is decomposed as follows:

$$\mu(x) = \mu_0 + m(x),$$

where μ_0 is a common mortality rate shared by all phenotypes, and $m(x)$ is an increment of mortality which models the load of an non-optimal phenotype. Environmental changes are modeled through a linear drift of the optimal trait at speed c. Mal-adaptation of a given individual with

Date: November 6, 2020.

trait x is thus defined as the difference between its phenotypic value x and the optimal value for the trait at time t. It will be denoted by $z \in \mathbb{R}$:

$$z := x - ct.$$

Finally, the non-linear term $\left((\beta - \mu_0) \int_{\mathbb{R}} f(t, x')dx'\right) f(t, x)$ accounts for density dependence in the population. We provide precise mathematical assumptions later, in Assumption 1.1.

We expect that under the assumption $\beta > \mu_0$, the population at the optimal trait will not go extinct in the absence of environmental changes. In our model, the interplay between selection, through mortality of fitter individuals close to a (moving) optimum, and traits diversification, through mutation, should balance out. We expect this balance to create a mutation–selection equilibrium of the phenotypical distribution. In that scenario, the dominant phenotypic trait of the population always lags behind the moving optimum, since adaptation is not instantaneous, creating a *lag*, as pictured in Figure 1.

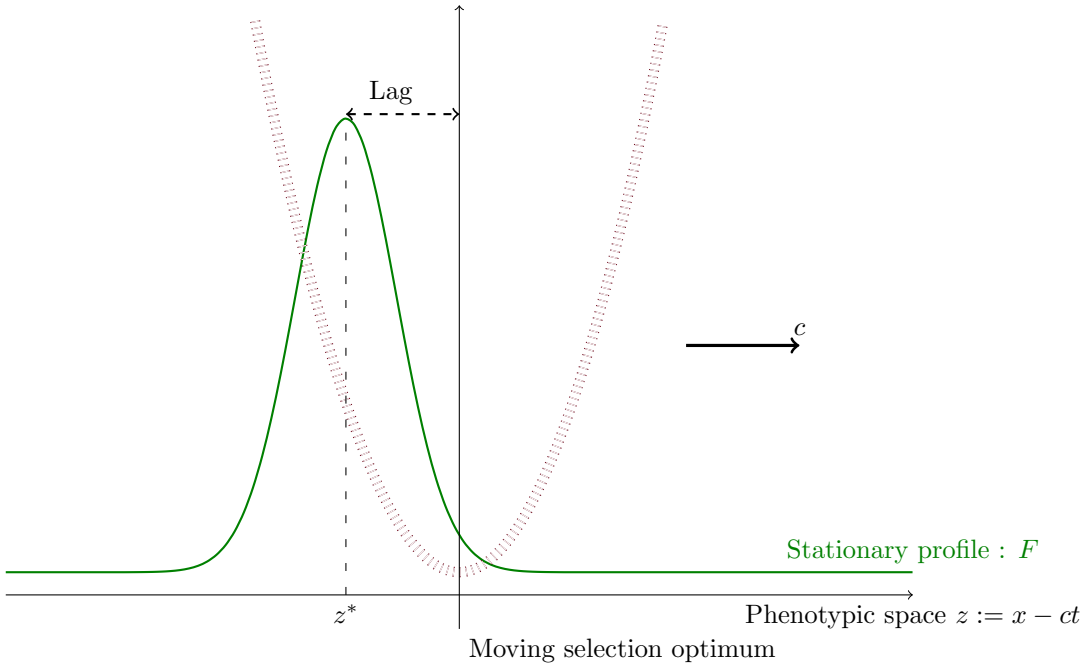


FIGURE 1. Shape of the equilibrium profile F solution of (1.1) in the *moving frame* at speed c (green plain line). In dashed red, a representation of an admissible selection function μ. The shape of equilibrium is not assumed *a priori*, this is obtained with a numerical resolution of (1.1). There phenotypic lag is defined by the position of the dominant trait in the population, z^*, and the optimal trait relatively to selection.

In this paper, we investigate the evolution of phenotypic lineages, *i.e.* the trajectories of the trait of an individual's ancestors across generations. We are not interested in the transient dynamics of (1.1), and any bias induced by the initial distribution $f(0, z)$. In other words, we want to characterize the dynamics inside an equilibrium of (1.1), and link it with ancestral lineages. Our method consists in looking at the dynamics of descendants issued from a small group of individuals. Let us assume that the individuals are labeled and that they transmit their label to their offspring even if the offspring differs in trait. Somehow, those label can be seen as neutral genes that are promoted or demoted because they are close to a gene under selection, creating a phenomenon of

genetic hitchhiking, or genetic draft, see Barton (2000). Mathematically, this approach has been introduced in the context of reaction diffusion equations by Hallatschek and Nelson (2008, 2010), and in Garnier et al. (2012); Roques et al. (2012) to understand the inside dynamic of the traveling wave solution. They were able, thanks to this methodology, to establish a robust dichotomy between two different possible behaviors for the inside dynamic of the front, by defining *pulled* or *pushed* fronts, a generalization of a concept of Stokes (1976). The distinction is important, since pushed fronts preserve the neutral diversity, as opposed to the pulled ones, where the neutral diversity is lost.

In the referential moving at the same speed as the front, the profile of a traveling wave is constant. In our case, its shape is no longer a front but rather a pulse, see Figure 1, and there are no previous works studying its inside dynamics. We assume that the equilibrium is made of several components, v^k, that we will call *neutral fractions*. By studying the evolution of these components, we will gain an understanding of the evolution of diversity inside the equilibrium. Namely, the fractions verify, at $t = 0$:

$$f_0(x) = \sum_{k \geqslant 1} v_0^k(x), \text{ with } v_0^k \geqslant 0 \text{ for all } k \geqslant 1.$$

We make the seminal assumption that every fraction only differs by their label, while mutation and selection act the same way they do on the entire population f. The density of each fraction verifies an equation of the form:

(1.3)
$$\begin{cases} \partial_t v^k(t,x) + \Big(\mu(x - ct) + (\beta - \mu_0) \displaystyle\int_{\mathbb{R}} f(t,x')\, dx'\Big) v^k(t,x) = \beta \mathcal{B}(v^k(t,\cdot))(x)\,, \text{ for } t > 0,\ x \in \mathbb{R} \\ v^k(0,x) = v_0^k(x),\ x \in \mathbb{R}. \end{cases}$$

Note that by linearity, the sum over k of the fraction densities v^k verifies (1.1). Our results are twofold.

- First we establish the time asymptotics of the neutral fraction v^k. We will therefore know, in the long time, which trait(s) contribute to the neutral diversity inside the equilibrium.
- Next, we will establish a partial differential equation upon the density of ancestors that reach a given trait z at a time t (the *lineages*). In the regime where the variance σ is small, we will find an explicit solution of the limit problem. This formula coïncides with heuristics based upon adaptive dynamics methods and Hamilton-Jacobi theory for the typical lineage among the population, which provides an innovative link between two popular approaches to theoretical evolutionary theory.

A companion paper in the field of probability theory deals with the issues raised here, Henry et al. (2020). They focus on the branching measure-valued population process underlying equation (1.1). It is well known that, in the large population limit, this individual-based process converges to the integro-differential equation, Fournier et al. (2004); Champagnat et al. (2006, 2007). Similarly to our current paper, they capture the dynamics of lineages. Interestingly, as a first step of their analysis, they consider the same linear equation as in (1.3). First, they deal with the case where the integral operator $\mathcal{B}$ is replaced by a Laplace (diffusive) operator, which we tackle in Section 3.2, and then proceed to follow the same methodology with a model with a probability kernel. This work is loosely based upon an article of Marguet (2019), where the author studies the dynamics of trait of a "typical" individual in branching processes.

A number of studies have investigated the change of phenotype in response to environmental changes, such as in climate change or drug resistance, see Parmesan (2006); Hoffmann and Sgro (2011); Collot et al. (2018). Mathematically, a very large number of models describe the adaptation of population in a *steady* environment, for instance, Diekmann et al. (2005); Alfaro and Carles (2014); Champagnat et al. (2006); Gil et al. (2019). In the case of a linearly varying environment,

there exists reaction diffusion models where a favorable region moves at a certain speed, Berestycki et al. (2009); Berestycki and Fang (2018), but it does not describe an adaptation phenomenon contrary to Alfaro et al. (2017), which considers a population structured in trait and space. In a still ongoing work, Bouin et al. (2020) obtain analytical features measuring the dynamics of adaptation for an integro-differential model in the same vein than (1.1). Their methodology is based upon quantitative genetics models, see Diekmann et al. (2005); Perthame and Barles (2008); Lorz et al. (2011), but also encapsulates the case where the operator $\mathcal{B}$ describes sexual reproduction via the infinitesimal operator, see Calvez et al. (2019); Patout (2020) for an adaptive dynamics description (without environmental change) of this operator. The case of a periodically fluctuating environment has also been studied, see Lorenzi et al. (2015); Figueroa Iglesias and Mirrahimi (2018, 2019); Carrère and Nadin (2020). Recently, Roques et al. (2020) proposed a methodology to deal with general changing environments (linear, oscillating or stochastic for instance) in the case of quadratic selection and a diffusion operator.

Adaptation also has profound consequences on the neutral genetic diversity and genealogies. In particular, in rapidly adapting populations, selective sweeps have a strong effect on genealogies Smith and Haigh (1974); Kaplan et al. (1989); Barton (1998); Barton and Etheridge (2004); Billiard et al. (2015). Some theoretical studies based on fitness stochastic model have investigated these effects on population expanding their range. They found that the genealogies of individuals at the front are not described by the classical Kingman coalescent but by a special type of multiple-merger coalescent, the Bolthausen-Sznitzman coalescent Brunet et al. (2007); Neher and Hallatschek (2013); Berestycki et al. (2013); Desai et al. (2013), although see also Etheridge and Penington (2020) where a Kingman coalescent is obtained when including a strong Allee effect in the range expansion. Moreover, in the context of soft sweeps which corresponds to selective sweeps that originate from multiple genomic backgrounds, Brunet et al. (2007); Neher and Hallatschek (2013); Desai et al. (2013) have shown that most of individuals trace back to one of two or more ancestral parents on which the selected mutation arose. In the context of adaptation to changing environment we could ask what are the traits of the ancestors of the individuals with the most common trait? However, little is known about the effect of adaptation due to changing environment on the neutral genetic diversity and the corresponding genealogies. In this paper, we aim to tackle this question using new mathematical tools.

We will use a seminal spectral result of Coville and Hamel (2019), where the authors tackle the existence of a spectral pair (λ, F) solution of

$$\lambda F(z) - c\partial_z F(z) + \mu(z)F(z) = \beta\mathcal{B}(F)(z)\,, \tag{1.4}$$

This provides an equilibrium profile, since $F(z-ct)$ is a stationary solution of (1.1) in the *moving frame*, and the eigenvalue is given by a simple formula:

$$\lambda = (\beta - \mu_0)\int_{\mathbb{R}} F(z')dz'. \tag{1.5}$$

This eigenvalue plays an important role in our analysis. On the one hand, thanks to (1.5), it delivers information about the size of population at equilibrium in the presence of a changing environment. Moreover, one should note that a formal integration of (1.4) yields

$$\lambda = \int_{\mathbb{R}} (\beta - \mu(z))\frac{F(z)}{\int_{\mathbb{R}} F(z')dz'}dz := \beta - \bar{\mu}, \tag{1.6}$$

where the mean mortality rate $\overline{\mu}$ caused by selection is defined by

$$\overline{\mu} = \int_{\mathbb{R}} \mu(z)\frac{F(z)}{\int_{\mathbb{R}} F(z')dz'}dz.$$

As a consequence of (1.6), λ can also be interpreted as a measure of the mean "fitness" of the population, or its mean intrinsic rate of increase, where $\beta - \mu(z)$ is the contribution to population growth rate of an individual with maladaptation z.

An extended analysis of (λ, F) is provided in Bouin et al. (2020), in a specific regime of adaptive dynamics. In particular, they obtain analytical formula for the critical speed of environmental change above which extinction is predicted, that is the case $\lambda < 0$.

The results of Coville and Hamel (2019) hold for general (bounded) potential functions μ but compactly supported kernels K. However, in the case of a confining potential, for instance quadratic, the result can be extended to more general kernels. As a matter of fact, recently, Cloez and Gabriel (2019) have extended the same existence result to this case, via a semigroup method inspired by irreducible aperiodic Markov chains. They further show exponential convergence for the semigroup, an important result in our study. We therefore make the following assumptions, to stay in the scope of those previous studies:

Assumptions 1.1.

- ▷ *$\beta > \mu_0$, so that the eigenvalue defined by (1.5) is positive.*
- ▷ *m is a convex function, such that $m(x) = m(|x|)$, and*

$$\lim_{|x|\to+\infty} m(x) = +\infty. \tag{1.7}$$

  Without loss of generality we suppose that it admits a (global) minimum at $x = 0$, such that $m(0) = 0$.
- ▷ *K is a standardized, positive, symmetric and thin-tailed probability kernel:*

$$\int_{\mathbb{R}} K(y)dy = \int_{\mathbb{R}} y^2 K(y)dy = 1, \qquad \exists \eta > 0 \ s.t. \ \int_{\mathbb{R}} K(y)e^{\eta|y|}dy < +\infty.$$

- ▷ *From now on, we assume, until the end of this article, that the solution of (1.1) has reached its equilibrium, which is*

$$f(t, z) = F(z) \text{ for all times } t \geqslant 0, \tag{1.8}$$

  and F is a non negative function in $\mathcal{C}^1_b(\mathbb{R})$.

Therefore, we investigate properties of the equilibrium F defined by (1.4) whose existence was investigated in the earlier works we mentioned: Cloez and Gabriel (2019); Coville and Hamel (2019).

We can now state our first result. Using the neutral fractions defined in (1.3), we are able to prove the existence of the following ancestral process inside the equilibrium F.

Theorem 1.2 (Ancestral process)**.**

(1) There exists a Feller semigroup, $(\mathcal{M}_s, s \geqslant 0)$, defined on the set of real continuous bounded functions, $\mathcal{C}_b(\mathbb{R})$, such that, for all $(t, z) \in \mathbb{R}_+ \times \mathbb{R}$, and all $k \in \mathbb{N}$:

$$\upsilon^k(t, z) = F(z)\mathcal{M}_{t-s}\left(\frac{\upsilon^k(s, \cdot)}{F}\right)(z), \quad \forall\, 0 \leqslant s \leqslant t. \tag{1.9}$$

The ancestral *Markov process $(Y_s, s \geqslant 0)$, associated to the semigroup $(\mathcal{M}_s, s \geqslant 0)$, is thus the moment dual of υ^k/F, in the sense of stochastic processes, see (1.11).*

(2) The generator of the ancestral process, $\mathcal{A}$, is given by:

$$\mathcal{A}\psi := \frac{\beta}{F}\Big[K * (F\psi) - (K * F)\psi\Big] + c\partial_z\psi. \tag{1.10}$$

for all $\psi \in \mathcal{C}^1_b(\mathbb{R}_+ \times \mathbb{R})$.

Remark 1.3.

- ▷ *Note that the initial time of the ancestral process Y_s ("$s = 0$"), can be any $t \in \mathbb{R}$. This means that the trajectory of ancestors is independent of the time at which we sample individuals in the population.*
- ▷ *By definition of Y_s, it also verifies for all $(t, z) \in \mathbb{R}_+ \times \mathbb{R}$, for all $k \in \mathbb{N}$:*

$$\frac{\upsilon^k(t,z)}{F(z)} = \mathbb{E}_z\left[\frac{\upsilon^k(s, Y_{t-s})}{F(Y_{t-s})}\right],\ 0 \leqslant s \leqslant t. \tag{1.11}$$

In this seminal relation lies the reason we refer to Y_s as an ancestral process. On the left of the equality is measured the probability that an individual drawn from those of trait z at time t is of type k. This is equal, on average, to draw an individual of type k in the past*, at the time $t-s$, but with a trait biased by the process Y. It measures the relative contributions of ancestors living at time s among the individuals with trait z at time t. Therefore Y is the representation of the phenotypic history of each type inside the population.*

- ▷ *A straightforward consequence of our point (2) is that w, the fundamental solution associated to $\mathcal{A}$, solves the following linear PDE:*

$$\begin{cases} \partial_s w(s,y,z) &= \beta \displaystyle\int_{\mathbb{R}} K(h)\frac{F(z+\sigma h)}{F(z)}\Big(w(s,y,z+\sigma h) - w(s,y,z)\Big)dh + c\partial_z w(s,y,z), \\ w(0,y,z) &= \delta(z-y). \end{cases} \tag{1.12}$$

This can be interpreted as a PDE describing the dynamics in the lineages of the trait y of ancestors whose descendants eventually reach a trait z in the population.

- ▷ *From (1.12), we can read explicitly that the ancestral process Y_s is a jump process, which realizes a jump of size σh at the rate $K(h)\frac{F(z+\sigma h)}{F(z)}$, coupled with a linear drift at speed c.*

Remarkably, the ancestral process Y_s does not depend on k, since the fractions are *neutral*. A major motivation behind the new formulation (1.9) is to be able to replace the discrete neutral fractions label k in (1.3) by a continuum of neutral alleles to fully apprehend the dynamics of trait. This is achieved if all the initial data are Dirac masses, indexed by $y \in \mathbb{R}$, and can be made rigorous in the framework of measure-valued equations. The underlying assumption is that each initial fraction in (1.3) corresponds to a single trait in the population, different for each fraction. In the following heuristic remark, we detail, via the fundamental solution w, this link between the ancestral process of Theorem 1.2 and Dirac initialized neutral fractions.

Remark 1.4. *Let $\mathcal{L}$ be the following linear operator:*

$$\mathcal{L}(\psi) = \beta\Big(K * \psi - \psi\Big) + c\partial_z\psi - \Big(\mu - \overline{\mu}\Big)\psi. \tag{1.13}$$

By construction, $\mathcal{L}$ is the differential operator that acts on the fractions in (1.3). Those are equivalently defined as the solution of the following Cauchy problem:

$$\begin{cases} \partial_t \upsilon^k(t,z) = \mathcal{L}(\upsilon^k(t,\cdot))(z) \text{ for } t > 0,\ z \in \mathbb{R} \\ \upsilon^k(0,z) = \upsilon_0^k(z),\ z \in \mathbb{R}. \end{cases} \tag{1.14}$$

Next, define for any $y \in \mathbb{R}$, υ^y as:

$$\upsilon^y(t-s,z) = F(z)w(s,y,z) \text{ for any } 0 \leqslant s \leqslant t, \text{ and } z \in \mathbb{R}, \tag{1.15}$$

where w is the fundamental solution associated to $\mathcal{A}$, the generator of the ancestral process, see (1.12). Then, one can show, via a simple computation detailed in Section 2, that υ^y solves the

following Cauchy problem:

$$\begin{cases} \partial_s v^y(s,z) &= \mathcal{L}(v^y)(s,\cdot)(z), \\ v^y(0,z) &= \delta(z-y)F(y). \end{cases} \tag{1.16}$$

This problem must be compared with (1.14). As a consequence, one can interpret v^y as a neutral fraction (both solve the same evolution equation involving the operator $\mathcal{L}$), with an initial data that samples a single trait in the population. In that sense, v^y corresponds to the phenotypic dynasty of the trait y in the population. This leads to one of the main message of this article: the ancestral process defined in Theorem 1.2 admits a correspondence with an understanding of the equilibrium in terms of sharp neutral fractions.

As a result of Theorem 1.2, we establish the long time behaviour of the lineages.

Proposition 1.5. *(Long time asymptotics)*
When $s \to \infty$, the ancestral process Y_s converges in law towards a random variable Y_∞, which admits the following density:

$$\frac{F\varphi}{\int F(y')\varphi(y')dy'},$$

where φ is the non-negative stationary solution of the dual problem:

$$\mathcal{L}^*(\varphi) = \beta\Big(K * \varphi - \varphi\Big) - c\partial_z\varphi - \Big(\mu - \overline{\mu}\Big)\varphi = 0. \tag{1.17}$$

Proposition 1.5 hinges on Proposition 3.1, which is an adaptation of previous results adapted to the case of our model. As a matter of fact, time asymptotic result were sufficient, in those previous studies that used neutral fractions in the field of reaction diffusion equations, to gather pertinent information. In Garnier et al. (2012), it is precisely a dichotomy in the large time behavior of fractions that proves to be the adequate criteria to discriminate between pulled and pushed fonts. In our case, not every trait z present in the equilibrium contributes to the distribution of ancestors the same way. Their contribution is given by the distribution of Y_∞. We have the more quantitative following result.

Corollary 1.6. *The distribution of Y_∞ is symmetric: for any $\psi \in \mathcal{C}_b(\mathbb{R})$,*

$$\mathbb{E}\left[\psi(-Y_\infty)\right] = \mathbb{E}\left[\psi(Y_\infty)\right],$$

in particular,

$$\mathbb{E}\left[Y_\infty\right] = 0. \tag{1.18}$$

Moreover, the law of Y_∞ is absolutely continuous with respect to the Lebesgue measure and its density admits a local extremum at 0.

Corollary 1.6 shows that, if we look far enough in the past, the ancestor of any individual in the population should have on average the trait 0 (in the moving frame), corresponding to the optimal trait with respect to selection. In particular, the common ancestors of individuals with the dominant trait have, on average, the optimal trait. This shows a striking non trivial dynamics inside the equilibrium. As seen on Figure 1, the density of individuals at the optimal trait 0 is very low in the equilibrium, because of the lag in adaptation. Therefore, we learn from the analysis an interesting feature of adaptation of our model: the persistence of the equilibrium is critically linked with the existence of mutants with the optimal trait, in order to be able to follow the linear drift of the environment.

We next focus on the regime of adaptive dynamics. It considers that mutations have small effects, and that time is accelerated to be able to look at the cumulative effect of those mutations. It this case, this corresponds to $\sigma \to 0$, with a rescaled time of order t/σ. We show that in this regime our

ancestral process Y_s takes an "explicit" form. Before stating our result, we define the Hamiltonian corresponding to the kernel of mutations K:

$$H(p) := \int_{\mathbb{R}} K(y)\exp(yp)dy - 1. \tag{1.19}$$

We also introduce the following rescaling limit of F, that holds locally uniformly, and is the result of Lorz et al. (2011):

$$U := \lim_{\sigma\to 0}(-\sigma \log F). \tag{1.20}$$

U is a $\mathcal{C}^1(\mathbb{R})$ function, solution of an Hamilton Jacobi equation, see section 4.1 for further details.

Proposition 1.7 (adaptive dynamics regime)**.**
We suppose that the speed of the moving optimum is scaled as follows:

$$c' = \frac{c}{\sigma}.$$

We define $\mathcal{M}^\sigma_s := \mathcal{M}_{s/\sigma}$. *Then, for all times* $s \in \mathbb{R}_+$, $z \in \mathbb{R}$ *and* $\psi \in \mathcal{C}_b(\mathbb{R})$,

$$\mathcal{M}^\sigma_s\psi(z) \xrightarrow[\sigma\to 0]{} \psi(\Gamma_z(s)), \text{ locally uniformly in time,} \tag{1.21}$$

and, for any $\varepsilon > 0$ *and* $T > 0$,

$$\lim_{\sigma\to 0}\mathbb{P}\left(\sup_{s\in[0,T]}|Y_s - \Gamma_z(s)| > \varepsilon \,\middle|\, Y_0 = z\right) = 0, \tag{1.22}$$

where Γ_z *is the solution of the following ODE:*

$$\begin{cases} \dot{\Gamma}_z(s) &= c' - \beta\partial_p H(\partial_z U(\Gamma(s))), \quad s > 0 \\ \Gamma_z(0) &= z. \end{cases} \tag{1.23}$$

H is the Hamiltonian of (1.19) and U the limit function such that the limit (1.20) holds.

The ancestral process in this limit regime is completely deterministic, which is to be expected since the source of randomness is the kernel K whose variance vanishes.

Remark 1.8. *We can propose a different interpretation of Proposition 1.7, in terms of partial differential equations. Alternatively, it means that the fundamental solution w_σ, introduced in (1.12), converges in the sense of distributions, when $\sigma \to 0$, up to the acceleration of time, towards w_0 the solution of this transport equation:*

$$\begin{cases} \partial_s w_0(s,y,z) &= -\Big(c' - \beta\partial_z H(\partial_z U(z))\Big)\partial_z w_0(s,y,z), \\ w_0(0,y,z) &= \delta(z-y). \end{cases} \tag{1.24}$$

The integral flow of the transport equation (1.24), which arises from the study of neutral fraction is exactly given by the ODE (1.23). Strikingly, it coincides precisely with an heuristic formula for the 'typical lineage' inside an equilibrium described by an Hamilton-Jacobi equation. We refer to the details around equations (4.11) and (4.14) in Section 4.1, where we explain how this formula has its origin in the field of the Weak-KAM theory, independently of neutral fractions. Therefore, we justify this interpretation, as $\sigma \to 0$, since Γ is indeed the typical lineage in the sense of (1.22).

The regime $\sigma \to 0$ has been widely studied following its introduction in the context of evolution by Diekmann et al. (2005). It is connected to large deviations theory, see Champagnat et al. (2019). It typically manages to describe how a population concentrates around one or many traits, see for instance Barles et al. (2009); Lorz et al. (2011), in the regime of rare mutations. The method usually implies that time is accelerated, as in Proposition 1.7, to be able to look at the effects of the (rare) mutations on the population. In fact, this kind of asymptotics has already been studied for (1.1) in the aforementioned literature, and the existence and convergence towards U in (1.20) is

a direct consequence. The Proposition 1.7 is a new bond with the adaptive dynamics literature. It establishes that the Hamilton-Jacobi structure delivers very rich information upon the way solutions concentrate, keeping even a trace of ancestral lineages. To the best of our knowledge, it was not shown before this article.

A detailed analysis of this specific regime, and what relevant ecological information can be deduced from it, is carried out in the ongoing work of Bouin et al. (2020). The rest of this article is organized as follows

- ▷ First, we provide the proof of Theorem 1.2, and we illustrate the link between fractions and the ancestral lineages on numerical simulations of equations (1.3) and (1.16).
- ▷ Next, we detail the proof of Proposition 1.5. We also provide numerous numerical simulations and we compare the theoretical long-time asymptotics with the simulations. Section 3.2 is devoted to the simpler diffusive approximation, in which explicit computations are possible. Corollary 1.6 is then proved in Section 3.1.
- ▷ We finally tackle in Section 4 the specific adaptive dynamics regime, where we prove Proposition 1.7 and we further discuss its implications. It includes a detailed discussion on the adaptive dynamics methodology for our model, inspired by Bouin et al. (2020). We finally compare the theoretical results with individual based simulations keeping track of the lineages.

Acknowledgements

Authors are very grateful for the many helpful comments of Jérôme Coville. This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 639638) and from the French Agence Nationale de la Recherche (ANR-18-CE45-0019 "RESISTE").

2. Link between lineages and fractions: proof of Theorem 1.2

The existence of the Feller semigroup holds true because the generator $\mathcal{A}$ defined, for all $\psi \in \mathcal{C}_b^1(\mathbb{R}_+ \times \mathbb{R})$, by

$$\mathcal{A}\psi := \frac{\beta}{F}\Big[K * (F\psi) - (K * F)\psi\Big] + c\partial_z\psi,$$

verifies all the hypotheses of the Hille-Yosida Theorem.

a) First, $\mathcal{A}$ is defined upon the set $\mathcal{C}_b^1(\mathbb{R})$, a dense part of $\mathcal{C}_b(\mathbb{R})$.

b) Next, the operator $\mathcal{A}$ verifies the maximum principle. This is clear from the expression of $\mathcal{A}$, and is maybe even clearer on the following equivalent expression of $\mathcal{A}$ which highlights the structure of a jump process coupled with transport:

$$\mathcal{A}\psi(z) = \beta \int_{\mathbb{R}} K(h)\frac{F(z+\sigma h)}{F(z)}\Big(\psi(z+\sigma h) - \psi(z)\Big)dh + c\,\partial_z\psi(z), \quad z \in \mathbb{R}. \tag{2.1}$$

c) Finally, one needs to check that there exists $\theta \in \mathbb{R}$ such that for any g in a dense subset of $\mathcal{C}_b(\mathbb{R})$, the equation

$$\theta\psi - \mathcal{A}\psi = g \tag{2.2}$$

admits a solution $\psi \in \mathcal{C}_b^1(\mathbb{R})$. To check this, the simplest way is to notice that the generator $\mathcal{A}$ can be rewritten as follows, for all $\psi \in \mathcal{C}_b^1(\mathbb{R})$:

$$\mathcal{A}\psi = \frac{\mathcal{L}(F\psi)}{F}, \tag{2.3}$$

where $\mathcal{L}$ is defined in (1.13) as the differential operator acting upon fraction v^k for every k. This is a consequence of the following computation, starting by:

$$\frac{1}{F}\mathcal{L}(F\psi) = \frac{\beta}{F}\Big(K*(F\psi) - F\psi\Big) + \frac{1}{F}c\,\partial_z(F\psi) - \frac{1}{F}\Big(\mu - \bar{\mu}\Big)\psi F. \tag{2.4}$$

In addition, F is a stationary profile, *i.e.*

$$\mathcal{L}(F) = 0.$$

Thus, by multiplying each side of this equality by ψ, one obtains:

$$\frac{\psi}{F}c\,\partial_z(F) - \frac{1}{F}\Big(\mu - \bar{\mu}\Big)\psi F - \frac{1}{F}\beta\psi F = -\frac{\psi}{F}\beta K * F. \tag{2.5}$$

Plugging (2.5) into (2.4),

$$\frac{1}{F}\mathcal{L}(F\psi) = \frac{\beta}{F}\Big(K*(F\psi) - \psi K * F\Big) + c\,\partial_z\psi = \mathcal{A}\psi,$$

and therefore the identity (2.3) is justified. We deduce that the existence of solution of (2.2) is a direct consequence of the similar statement for $\mathcal{L}$, and a justification of this standard result can be found in Bansaye et al. (2019).

Therefore, thanks to points $a), b)$ and $c)$, the Hille-Yosida Theorem (Ethier and Kurtz, 2009, Chapter 4, Theorem 2.2), guarantees that there exists a strongly continuous semigroup $(\mathcal{M}_s, s \geqslant 0)$ such that for every $\psi \in \mathcal{C}_b(\mathbb{R})$ and $s > 0$:

$$\frac{d}{ds}\mathcal{M}_s\psi = \mathcal{M}_s\mathcal{A}\psi, \qquad \text{and } \mathcal{M}_0\psi = \psi.$$

To conclude the proof, we now check that $\mathcal{M}$, the semi group associated to $\mathcal{A}$ is associated to the ancestral process defined by (1.9). To that end, for all $(t,z) \in \mathbb{R}_+ \times \mathbb{R}$, $0 \leqslant s \leqslant t$ and $k \in \mathbb{N}$, we compute:

$$\frac{d}{ds}\mathcal{M}_{t-s}\left(\frac{v^k(s,\cdot)}{F}\right)(z) = -\mathcal{M}_{t-s}\left[\mathcal{A}\left(\frac{v^k(s,\cdot)}{F}\right)\right](z) + \mathcal{M}_{t-s}\left[\frac{\partial_s v^k(s,\cdot)}{F}\right](z). \tag{2.6}$$

We use the relationship (2.3) between $\mathcal{A}$ and $\mathcal{L}$ for the first term. For the second, we recall that :

$$\partial_s v^k = \mathcal{L}(v^k).$$

Plugging this into (2.6),

$$\begin{aligned}\frac{d}{ds}\mathcal{M}_{t-s}\left(\frac{v^k(s,\cdot)}{F}\right)(z) &= -\mathcal{M}_{t-s}\left[\frac{\mathcal{L}\left(v^k(s,\cdot)\right)}{F}\right](z) + \mathcal{M}_{t-s}\left[\frac{\mathcal{L}(v^k(s,\cdot))}{F}\right](z),\\ &= 0.\end{aligned}$$

Since $\mathcal{M}_0 = Id$, by evaluating in $s = t$ we have shown that for all $k \in \mathbb{N}$ and $(t,z) \in \mathbb{R}_+ \times \mathbb{R}$,

$$\frac{v^k(t,z)}{F(z)} = \mathcal{M}_{t-s}\left(\frac{v^k(s,\cdot)}{F}\right)(z), \quad 0 \leqslant s \leqslant t. \tag{2.7}$$

The identity (1.9) is established. As a matter of fact, we can also straightforwardly define the Markov process $(Y_s, s \geqslant 0)$ associated to $(\mathcal{M}_s, s \geqslant 0)$, see for instance (Ethier and Kurtz, 2009, Chapter 4, Theorem 2.7). As a direct consequence of (1.9), the dual relationship of (1.11) is verified. □

3. Long time asymptotics of lineages

We now turn towards the study of the long time asympotics of the ancestral process Y_s. Note that, as it can be seen on the duality relationship (1.11), the time of ancestors and of the population is reversed : the regime $s \to +\infty$ corresponds to studying the most ancient ancestor, *"backwards"* in terms of the time t of the equilibrium.

3.1. Long time asymptotics of fractions.

Our statement Proposition 1.5 is based upon a similar result concerning the neutral fractions, Proposition 3.1:

Proposition 3.1 (Cloez and Gabriel (2019))**.**
For any k, let $v_0^k(z) \in \mathcal{C}_b^1(\mathbb{R})$ be the initial data. Then, the neutral fraction v^k that solves the Cauchy problem (1.3) converges, when $s \to \infty$ towards a proportion of the total population F. This proportion only depends on the initial data $v_0^k(z)$, as follows:

$$v^k(s,\cdot) \xrightarrow[s\to\infty]{L^1} p[v_0^k]F, \quad \text{with} \quad p[v_0^k] := \frac{\displaystyle\int_{\mathbb{R}} v_0^k(z)\varphi(z)dz}{\displaystyle\int_{\mathbb{R}} F(z')\varphi(z')dz'}, \tag{3.1}$$

where φ is defined as the non-negative solution of the following dual stationary problem:

$$0 = \beta\Big(K * \varphi - \varphi\Big) - c\,\partial_z\varphi - \Big(\mu - \overline{\mu}\Big)\varphi.$$

The proof of this result hinges entirely on the recent results of (Cloez and Gabriel, 2019, Theorem 2.1), and we believe it can be deduced as well from the general semigroup analysis of growth fragmentation equations presented in Mischler and Scher (2016).

We can observe that every fraction contributes to the equilibrium asymptotically. This means that the soliton generated by the Cauchy problem (1.1) is pushed in the sense of Garnier et al. (2012). This was expected because the heterogeneity of $\mu(z)$ produces on the leading and rear edges of the pulse two unfavourable zones where the mortality is high. We know from Garnier and Lewis (2016), that any mechanisms that constrain the propagation tend to produce pushed travelling wave solutions.

Proof of Proposition 1.5.
Consider a function $\psi \in \mathcal{C}_b^1(\mathbb{R})$, and define $\widetilde{\psi} = F\psi$, also a function of $\mathcal{C}_b^1(\mathbb{R})$. Next we consider a neutral fraction v, in the sense of (1.3), initiated with $\widetilde{\psi}$. This means that

$$\begin{cases} \partial_t v(t,z) = \mathcal{L}(v(t,\cdot))(z) \text{ for } t>0,\ z\in\mathbb{R} \\ v(0,z) = \widetilde{\psi}(z),\ z\in\mathbb{R}. \end{cases}$$

One applies the semigroup property (1.11) and takes $s = t$. This yields the following dual relationship:

$$\frac{v(t,z)}{F(z)} = \mathbb{E}_z\left[\frac{\widetilde{\psi}(Y_t)}{F(Y_t)}\right] = \mathbb{E}_z\left[\psi(Y_t)\right]. \tag{3.2}$$

Moreover, with Proposition 3.1, the left hand side converges, when $t \to \infty$ towards $p[\widetilde{\psi}]$:

$$p[\widetilde{\psi}] = \frac{\displaystyle\int_{\mathbb{R}} \widetilde{\psi}(z)\varphi(z)dz}{\displaystyle\int_{\mathbb{R}} F(z')\varphi(z')dz'} = \frac{\displaystyle\int_{\mathbb{R}} F(z)\psi(z)\varphi(z)dz}{\displaystyle\int_{\mathbb{R}} F(z')\varphi(z')dz'},$$

where φ verifies $\mathcal{L}^*(\varphi) = 0$. Note that, by definition of Y_∞, we have that

$$p[\widetilde{\psi}] = \mathbb{E}_z\left[\psi(Y_\infty)\right].$$

Therefore, with (3.2), we have shown that for any $\psi \in \mathcal{C}_b^1(\mathbb{R})$,

$$\mathbb{E}_z[\psi(Y_t)] \xrightarrow[t\to\infty]{} \mathbb{E}_z\left[\psi(Y_\infty)\right].$$

□

For the sake of consistency, one may observe that the density found for the fractions in Proposition 3.1 is consistent with the heuristics of Dirac initiated fractions, made in Remark 1.4. Consider formally $\upsilon_0^k(z) = \delta(z-y)F(z)$ and apply to it the asymptotic result of the fractions stated in Proposition 3.1. One recovers the formula for the density of Proposition 1.5.

Proof of Corollary 1.6.
To show Corollary 1.6, we adapt the arguments of the diffusive approximation. We find out that there exists again an explicit link between the solution of the dual problem and the original one. Notice that $\varphi(z) := F(-z)$ satisfies

$$\mathcal{L}^*(\varphi) = 0,$$

with $\mathcal{L}$ defined back in (1.17). The reason is that μ is an *even* function. Therefore, **the function $F\varphi$ is even**, and Y_∞ admits an even density. Corollary 1.6 immediately follows. □

3.2. **An illuminating example, the diffusive approximation.**
To illustrate the scope of Proposition 1.5, we take an example where explicit computations are possible, the diffusive approximation. The method we will use in that context will be enlightening to tackle the Corollary 1.6 in Section 3.1. It is well known that if the mutational variance σ^2 is small, the convolution operator $\mathcal{B}$ defined in (1.2) can be approximated by:

$$\mathcal{B}(F)(z) \approx \beta F(z) + \beta\frac{\sigma^2}{2}\partial_z^2 F(z).$$

For clarity, we adopt in this section the same notations as previously. Therefore, in the present section only, the operator $\mathcal{L}$, defined in (1.13), becomes:

$$\mathcal{L}(v) = \beta\frac{\sigma^2}{2}\partial_z^2 v + c\partial_z v - \Big(\mu - \overline{\mu}\Big)v \tag{3.3}$$

This corresponds to the evolution of fractions inside an equilibrium F, the stationary profile, that verifies $\mathcal{L}(F) = 0$. We sum up the results for this approximation in the following proposition.

Proposition 3.2 (The case of the diffusive approximation)**.**
The ancestral process $(Y_s, s \geqslant 0)$ associated to the model of the diffusive approximation (3.3) admits the following generator:

$$\mathcal{A}\psi = \beta\frac{\sigma^2}{2}\partial_z^2\psi + \left(\beta\frac{\sigma^2}{2}\frac{\partial_z F}{F} + c\right)\partial_z\psi. \tag{3.4}$$

When $s \to \infty$, the limit process Y_∞ admits a density given by

$$\frac{\left(F(y)e^{cy/(\beta\sigma^2)}\right)^2}{\displaystyle\int_{\mathbb{R}}\left(F(y')e^{cy'/(\beta\sigma^2)}\right)^2 dy'}.$$

In addition Y_∞ admits a local maximum at $z = 0$ for σ small enough.

Classically, a similar result to Proposition 3.1 can be established in the case of the diffusive approximation, and Proposition 3.2 hinges on that to get the similar results of Theorem 1.2 on the existence of the ancestral process. As a matter of fact, this shows that mathematically, our results can be extended to all operators $\mathcal{L}$ such that asymptotic results have been established, and we refer to Section 5 for further discussion.

Applying our ancestral result of Theorem 1.2, we deduce that the generator $\mathcal{A}$ of the ancestral process verifies

$$\mathcal{A}\psi = \frac{1}{F}\mathcal{L}(F\psi) = \beta\frac{\sigma^2}{2}\left(\frac{1}{F}\partial_z^2(F\psi) - \frac{1}{F}\psi\partial_z^2 F\right) + c\partial_z\psi.$$

After simplification, using $\mathcal{L}(F) = 0$, one finds that the expression for $\mathcal{A}$ becomes

$$\mathcal{A}\psi = \beta\frac{\sigma^2}{2}\partial_z^2\psi + \left(\beta\frac{\sigma^2}{2}\frac{\partial_z F}{F} + c\right)\partial_z\psi,$$

as in (3.4). We thus see that, in this diffusive approximation, the ancestral process $(Y_s, s \geqslant 0)$ solves a stochastic differential equation of the form

$$dY_s = \left(\beta\frac{\sigma^2}{2}\frac{\partial_z F(Y_s)}{F(Y_s)} + c\right)ds + \sqrt{\beta\sigma^2}dB_s, \tag{3.5}$$

where $(B_s, s \geqslant 0)$ is standard Brownian motion.

Similarly to its definition in (1.17), let φ be the solution of the dual problem:

$$\mathcal{L}^*(\varphi) = \beta\frac{\sigma^2}{2}\partial_z^2\varphi - c\partial_z\varphi - \Big(\mu - \overline{\mu}\Big)\varphi = 0.$$

It turns out that, there exists an explicit link between F and φ. Computations show that

$$\varphi(z) = F(z)e^{2cz/(\beta\sigma^2)}.$$

By applying the formula of Proposition 1.5, it means that the asymptotic distribution of ancestors Y_∞ admits an 'explicit' density:

$$\frac{\Big(F(y)e^{cy/(\beta\sigma^2)}\Big)^2}{\displaystyle\int_{\mathbb{R}}\Big(F(y')e^{cy'/(\beta\sigma^2)}\Big)^2 dy'}.$$

The function defined for each z as $\widetilde{F}(z) := Fe^{cz/(\beta\sigma^2)}$ is even, since μ is an even function by hypothesis and $\widetilde{F}$ is the (unique) solution of:

$$\beta\frac{\sigma^2}{2}\partial_z^2\widetilde{F} - \Big(\mu - \overline{\mu} + \frac{c^2}{2\beta\sigma^2}\Big)\widetilde{F} = 0. \tag{3.6}$$

Therefore $y \mapsto \Big(F(y)e^{cy/(\beta\sigma^2)}\Big)^2$ is also even, and as a consequence Y_∞ is symmetric. As a result, $\widetilde{F}$ admits a (local) extrema at $z = 0$. To obtain more information we must investigate the sign of $\partial_z^2\widetilde{F}(0)$. Going back to (3.6), we first notice that $\mu_0 - \bar{\mu} < 0$. Therefore, if σ is sufficiently small, we find that $\partial_z^2\widetilde{F}(0) > 0$, and therefore Y_∞ admits a local maximum at 0, as claimed in Proposition 3.2.

The quadratic selection. We can even consider a simpler model, where $m(z) = z^2/2$, instead of a general convex selection function. In that case, it is well known that explicit Gaussian solutions of the stationary equation exist. More precisely, the solution of

$$\lambda F(z) - c\partial_z F(z) + \frac{z^2}{2}F(z) = \beta F(z) + \frac{\beta\sigma^2}{2}\partial_z^2 F(z), \text{ for } z \in \mathbb{R}, \tag{3.7}$$

is given by

$$F(z) = \frac{\lambda}{\sqrt{2\pi\sigma\sqrt{\beta}}} \exp\left(-\frac{1}{2\sigma\sqrt{\beta}}\left(z + \frac{c}{\sigma\sqrt{\beta}}\right)^2\right), \quad \lambda = \beta - \frac{c^2}{2\beta\sigma^2} - \frac{\sigma\sqrt{\beta}}{2}. \tag{3.8}$$

Up to a constant, F is a Gaussian distribution centered around an optimum proportional to c, that lags behind the optimal trait, and a variance proportional to σ (instead of σ^2). In the eigenvalue λ, we recognize the *lag load*: $c^2/2\beta\sigma^2$, that is the weight to keep pace with the environment, and the *mutation load*: $\sigma\sqrt{\beta}/2$. In addition, we see that the speed of change c must be small enough for the population to persist ($\lambda > 0$ if $c \leqslant \sigma\beta\sqrt{2}\sqrt{1 - \sigma/(2\sqrt{\beta})}$). In particular, c must be of order σ, just as in the adaptive dynamics regime of Proposition 1.7.

Moreover, we show that Y_∞, the asymptotic (as $s \to \infty$) ancestral distribution of Y_s, is a Gaussian distribution centered at $z = 0$ and with variance $\sigma\sqrt{\beta}/2$, since its density is given by:

$$(F\varphi)(z) = \frac{1}{\sqrt{\pi\sigma\sqrt{\beta}}} \exp\left(-\frac{z^2}{\sigma\sqrt{\beta}}\right).$$

The variance of the ancestral distribution is $\sigma\sqrt{\beta}/2$. Noticeably, it is reduced compared to the variance of the trait distribution at equilibrium. This variance is also equal to the mutation load.

In fact, in this setting, the process $(Y_s, s \geqslant 0)$ is an Ornstein-Uhlenbeck process, as can be seen by substituting (3.8) in (3.5). We can thus also derive an explicit formula for the trajectories along time of the mean and the variance of the ancestral distribution Y_s. We have:

$$\mathbb{E}_z(Y_s) = ze^{-\sigma\sqrt{\beta}s} \quad \text{and} \quad \mathrm{Var}_z(Y_s) = \frac{\sigma\sqrt{\beta}}{2}\left(1 - e^{-2\sigma\sqrt{\beta}s}\right). \tag{3.9}$$

We can observe that the variance does not depend on the reference point z and eventually converges as $s \to \infty$ to $\sigma\sqrt{\beta}/2$, the mutation load. Moreover, we see that the mean of the ancestral distribution converges to 0 exponentially fast, at a rate $\sigma\sqrt{\beta}$. Building on this, we can conjecture that, if the selection function is no longer quadratic, as in (3.3), the convergence rate of the mean $\mathbb{E}_z(Y_s)$ is given, this time around by $\sigma^2\beta/\mathrm{Var}(F)$ where $\mathrm{Var}(F)$ is the variance of the trait distribution of the population at equilibrium.

3.3. Transient and asymptotic dynamics of the ancestral process: numerical insights.

We first present some detailed numerical pictures of our results. In Figure 2, we show the asymptotic density of the ancestral lineage Y_∞. Thanks to Proposition 3.1, one knows that it also represents the proportion $p[y]$ of ancestors of phenotype y in the population, asymptotically as $s \to \infty$. As expected by our results, it is an even function with a maximum at the phenotype 0. It is striking to notice that despite the very low density F around 0, most ancestors have a trait close to this optimal trait.

As we explain in Remark 1.4, an interpretation of the ancestral process is to assume a continuous number of neutral fractions v^y, indexed by a parameter $y \in \mathbb{R}$, and to start with the initial data $\delta(z-y)F(y)$. Numerically, we simulate the problem (1.16), for a large number of initial phenotype $y \in \mathbb{R}$, corresponding to a sharp discretisation of the trait space. Figure 3 represents the cumulative (over y) densities of the fraction v^y we use for our simulation. We observe that, as expected, the sum over all fractions v^y is equal to F, the total population.

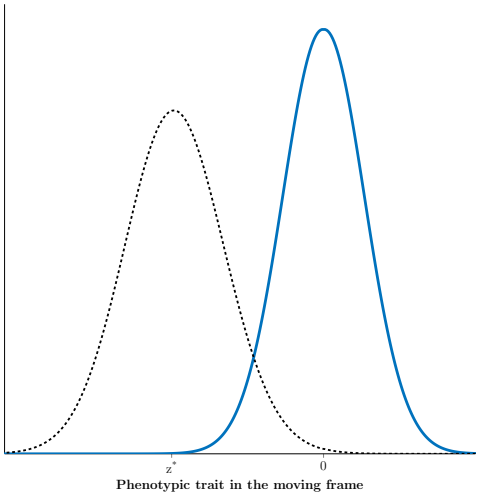


FIGURE 2. In solid blue: density of Y_∞, and in dashed black: the equilibrium F.

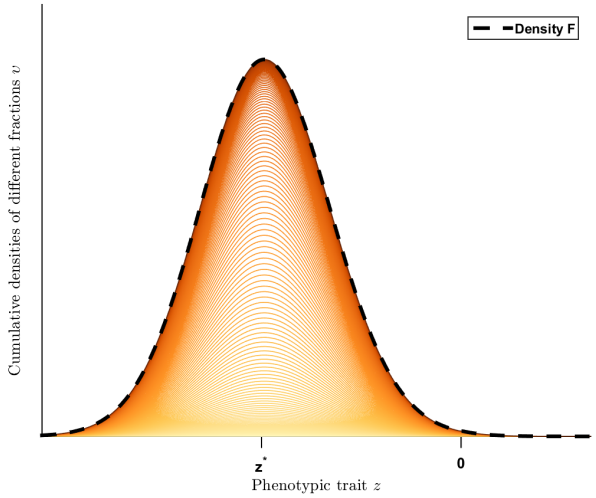


FIGURE 3. Initial cumulative densities of the fractions υ^y, $s = 0$.

Thanks the simulations of all those fractions, we can track the evolution of lineages at all times in the population, and not only settle for the asymptotics of Proposition 1.5 and Figure 2. Figure 4 shows the evolution of the distribution of the ancestral lineage at different times s. The evolution depicted by Figure 4 is explained as follows. Initially ($s = 0$), the fraction consists of a Dirac mass at z^*. This means that we sample individuals in the population with the trait z^*. Then, as $s \to \infty$ we go back in time and Figure 4 pictures the distribution of the ancestors with trait z at different fixed times. As s increases to infinity, we see that the distribution gradually shifts towards the right, until eventually it reaches the stationary state. The thick line we observe is therefore another representation of the asymptotic density of ancestors displayed in Figure 2.

To better grasp the evolution of lineages ending at trait z along time, we look at the mean trait of the ancestors $\mathbb{E}_z(Y_s)$ and the variance of the ancestors' trait $\mathrm{Var}_z(Y_s)$ as a function of s. We can guess from Figure 4 and the diffusive approximation, that the variance grows until it reaches

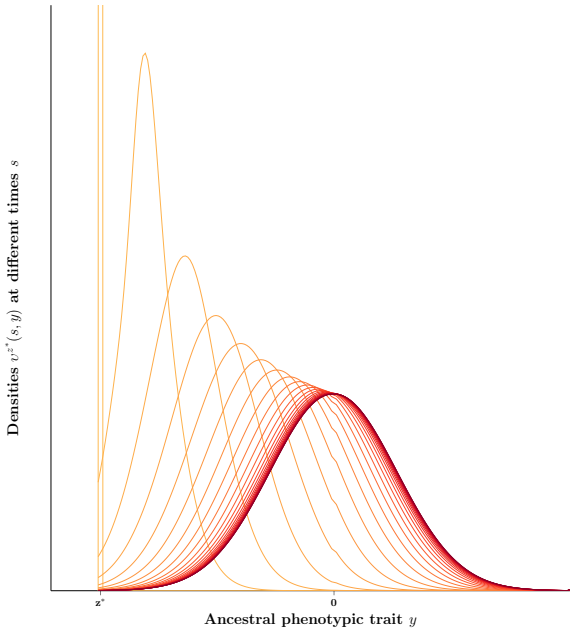


FIGURE 4. Snapshots of the density of Y_s for different times s, with initial sample at the trait z^*. Lighter colors refer to smaller times.

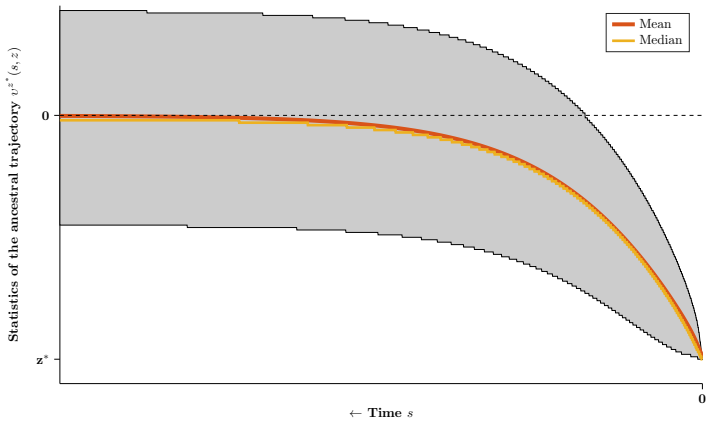


FIGURE 5. Evolution of the mean and median of $v^{z^*}(s,z)$ along time. The grey area corresponds to the region between the 5% and the 95% quantile of the distribution. Time s increases towards the left, to insist that it is "backwards", and to be consistent with individual based simulations and the adaptive dynamics regime of further sections, see Figure 7.

its stationary value, while the mean converges to 0. This is confirmed by the Figure 5. We also displayed along the 5% and the 95% quantile of the distribution, to get a grasp of the width of the ancestral process.

4. Adaptive dynamics regime

We finally tackle in this section the adaptive dynamics regime. In a first part, largely inspired by the ongoing work of Bouin et al. (2020), we explain how to obtain quantitative results measuring the impact of environmental change on the fitness of individuals or the shape of the distribution,

for instance. Their idea is to introduce a small scaling parameter:

$$\varepsilon := \sqrt{\sigma^2 \frac{\beta}{\alpha}}, \text{ with } \alpha := \partial_z^2 m(0) > 0. \tag{4.1}$$

and then be able to characterize the limit $\varepsilon \to 0$. They monitor a *weak selection regime*, in the sense that either the variance σ^2 is small, or the selection is weak compared to birth:

$$\frac{\alpha}{\beta} \ll 1.$$

In both cases the parameter ε is small. We recall that α is a way to measure the strength of selection around the global optimum of μ at the trait 0, see Assumption 1.1.

We will explain how in the regime of $\sigma \ll 1$, we can get intuitive results about lineages. Then, we will prove the Proposition 1.7 and show how it covers those previous heuristics. We will finally provide numerical simulations of individual based models keeping track of the lineages to illustrate our previous results.

4.1. Hamilton Jacobi limit for (1.1) and heuristics about lineages.

We will briefly explain how, when $\sigma \to 0$, the equilibrium F concentrates around a mean trait value, z^*, with a standing variance. Recall that F is a solution of the integro-differential equation (1.4), which we rewrite here:

$$\lambda F(z) - c\partial_z F(z) + \mu(z)F(z) = \frac{\beta}{\sigma}\int_{\mathbb{R}} K\left(\frac{z-z'}{\sigma}\right) F(z')\, dz'. \tag{4.2}$$

Here, to keep things simple, we will not introduce the parameter ε as in Bouin et al. (2020). We keep our biological parameters introduced in (1.1) and study a less general regime, the small variance limit, that is $\sigma \to 0$. When that happens, one expects concentration around a specific trait, guided by selection, see Barles et al. (2009); Lorz et al. (2011) for instance. This motivates the following logarithmic transform:

$$F(z) = \exp\left(-\frac{U(z)}{\sigma}\right). \tag{4.3}$$

Plugging (4.3) into (4.2), yields the following equality :

$$\lambda + \frac{c}{\sigma}\partial_z U(z) + \mu(z) = \frac{\beta}{\sigma}\int_{\mathbb{R}} K\left(\frac{z-z'}{\sigma}\right) \exp\left(\frac{U(z)-U(z')}{\sigma}\right) dz'. \tag{4.4}$$

To obtain a finite limit when $\sigma \to 0$, it seems clear that the speed of environmental change must be scaled as σ, this is why we make the following assumption, as in Proposition 1.7:

$$c := \sigma c'. \tag{4.5}$$

We finally make the single approximation of this computation, for the integral term of (4.4):

$$\forall z' \in \mathbb{R}, \quad \exp\left(\frac{U(z)-U(z')}{\sigma}\right) \approx \exp\left(-\frac{z-z'}{\sigma}\partial_z U(z)\right).$$

Therefore by an affine change of variable, (4.4) is expressed as:

$$\lambda + c'\partial_z U(z) + \mu(z) = \beta\int_{\mathbb{R}} K(y)\exp\Big(y\partial_z U(z)\Big)\, dy.$$

Note that assumptions made in Assumption 1.1, on the exponential decay of K, make possible to define the Hamiltonian H as in (1.19) for all $p \in \mathbb{R}$:

$$H(p) := \int_{\mathbb{R}} K(y)\exp(yp)dy - 1.$$

Finally, we have established the following Hamilton-Jacobi equation, for $\sigma = 0$:

$$\lambda + c'\partial_z U(z) + \mu(z) = \beta + \beta H(\partial_z U(z)). \tag{4.6}$$

Note that the convergence when $\sigma \to 0$ of U_σ solution of (4.4) towards U_0 solution of the Hamilton Jacobi equation (4.6), in the sense of viscosity solutions, is established in Barles et al. (2009); Lorz et al. (2011).

Introduce the Lagrangian function L corresponding to K, the Legendre transform of H:

$$L(v) := \max_{p\in\mathbb{R}} \Big(pv - H(p)\Big). \tag{4.7}$$

The asymptotic $\sigma \to 0$ yields explicit formulas, expected to be true when σ is small. For instance, Bouin et al. (2020) show how to get a first order (in ε defined in (4.1)) approximation for the growth rate λ :

$$\lambda \approx \beta - \mu_0 - \beta L\left(\frac{c'}{\beta}\right) + O(\varepsilon). \tag{4.8}$$

Thanks to this formula, they can compute a threshold on c so that the population does not go extinct, corresponding to the limit case $\lambda < 0$. See (4.20) and what follows for a full justification. Moreover, by taking $z = z^*$ in (4.6), we obtain an equation that dictates the position of the dominant trait z^*:

$$m(z^*) = \beta L\left(\frac{c'}{\beta}\right), \tag{4.9}$$

Beyond those formulas, we are interested, in our case, in the dual representation of U, the solution of (4.6), see Barles and Roquejoffre (2006) for a justification of this formula, that stems from the Weak-KAM theory:

$$U(z) = \inf_{\gamma \text{ s.t. } \gamma(0)=z} \int_{-\infty}^{0} \left[L\left(\frac{\dot\gamma(s) + c'}{\beta}\right) - \beta + \mu(\gamma(s)) + \lambda\right] ds. \tag{4.10}$$

The infimum is taken over all functions $\gamma \in \mathcal{C}^1(\mathbb{R}_-)$ that reach the phenotype z at time 0, when going backward in time. This constitutes a reformulation of the spectral problem (4.6) into an equivalent variational problem. Along a phenotypic path γ, an optimal trajectory must minimize the cost. This cost comes from the combined weight of mutations through L (at speed $\dot\gamma + c'$) and selection through μ. The birth rate β plays an opposite role of selection, while λ is the term that balances the expression, just as in (4.6).

In terms of ancestral lineages, an interpretation of the formula is to consider the minimizing trajectory of (4.10) as the typical phenotype of ancestors inside the equilibrium that eventually give the phenotype z at a given end time, here 0.

Let Γ be such a minimizing trajectory. The knowledge of the function Γ for any end point z is somehow richer than the one of the profile U since it accounts, backwards in time, for the full dynamics of the equilibrium from time $-\infty$ to 0. Given the supposed existence of Γ, or of a trajectory that is arbitrarily close to optimum, we can make the following formal analysis, starting by evaluating the energy of (4.10) along Γ:

$$U(z) = \int_{-\infty}^{0} \left[L\left(\frac{\dot\Gamma(s) + c'}{\beta}\right) - \beta + \mu(\Gamma(s)) + \lambda\right] ds. \tag{4.11}$$

One should expect $\dot\Gamma(s)$ to converge to 0 when $s \to -\infty$. Otherwise, the mortality $\mu(\Gamma)$ would become arbitrarily large and so the trajectory would not be a minimizer since $U_0(z)$ could take the value $+\infty$. Formally, this implies that if Γ converges when $s \to -\infty$, then necessarily it is towards

the value 0, in order to minimize the selection function μ. With this formal argument we expect that :

$$\Gamma(s) \xrightarrow[s\to-\infty]{} 0.$$

This limit has to be put in parallel with Corollary 1.6. We discover here once again that asymptotically in time the ancestors are expected to come, in majority, from the optimal trait 0. Assuming this limit holds true, it prescribes the value of λ such that the integrand in (4.11) vanishes, since otherwise U would be infinite:

$$L\left(\frac{c'}{\beta}\right) - \beta + \mu(0) + \lambda = 0. \tag{4.12}$$

Since $\mu(0) = \mu_0$, and $c = c'\sigma$ was the rescaling of the original speed of adaptation to the variance of mutations, the formula coincides with (4.8). The term $L\left(\frac{c}{\sigma\beta}\right)$ measures the constraint of having to keep pace with the environment evolving at speed c, while $\mu_0 - \beta$ is the fitness cost for the optimally fitted asymptotic individuals of the lineage.

To conclude the remarks around the variational (4.10), one can, by means of consistency consistency with the rest of this article, we can rewrite it, up to a slight abuse of notation as

$$U(z) = \inf_{\gamma \text{ s.t. } \gamma(0)=z} \int_0^{+\infty} \left[L\left(\frac{-\dot\gamma(s) + c'}{\beta}\right) - \beta + \mu(\gamma(s)) + \lambda \right] ds. \tag{4.13}$$

Then, as a byproduct of the Weak-KAM theory, on can show that the optimal trajectory Γ of the variational problem (4.13) is the solution of an Ordinary Differential Equation, for instance Hairer et al. (2006):

$$\begin{aligned} &\dot\Gamma(s) = c' - \beta\, \partial_p H\Big(\partial_z U\big(\Gamma(s)\big)\Big), \quad s > 0 \\ &\Gamma(0) = z. \end{aligned} \tag{4.14}$$

This result comes from the Hamiltonian/Lagrangian structure of (4.10), and more precisely from working on the characteristics of this Hamilton Jacobi equation. It is not simpler to solve (4.14), since U itself depends on Γ, but it gives information on the whole behavior of Γ. It will prove to be nonetheless a useful equation to simulate Γ, if one is externally handed the profile U, for instance numerically.

Going back to the initial motivations stated at the beginning of this article, a key point of our analysis is that this equation on Γ can be recovered from our ancestral process. In other words, equations (1.23) and (4.14) coincide. In the next part, we will establish this link, proving Proposition 1.7. Finally, to further assess our interpretation of Γ as a typical lineage, we will run individual based stochastic simulations and compare them with both the neutral fractions and the theoretical formula for Γ given by (1.23).

4.2. **Proof of Proposition 1.7.**

We will use a classical result, stated for instance in (Kallenberg, 2006, Theorem 17.25) to link the convergence of the generator $\mathcal{A}$ and the convergence of the semi group $\mathcal{M}$. We consider $\psi \in \mathcal{C}_b^2(\mathbb{R}_+ \times \mathbb{R})$, as it is a classical result that it constitutes a core for $\mathcal{A}$, Cloez and Gabriel (2019); Bansaye et al. (2019). Then, by defining $\mathcal{A}^\sigma$ the generator corresponding to the semigroup $\mathcal{M}_{s/\sigma}$, for any $s \geqslant 0$, we get,

$$\mathcal{A}^\sigma \psi(z) = \frac{\beta}{\sigma} \int_{\mathbb{R}} K(h) \frac{F(z+\sigma h)}{F(z)} \Big(\psi(z+\sigma h) - \psi(z)\Big) dh + \frac{c}{\sigma} \partial_z \psi(z), \quad z \in \mathbb{R}.$$

We introduce the function U_σ as in (4.3). Then the previous equation writes

$$\mathcal{A}^\sigma\psi(z) = \frac{\beta}{\sigma}\int_{\mathbb{R}} K(h)\exp\left(-\frac{U_\sigma(z+\sigma h)-U_\sigma(z)}{\sigma}\right)\Big(\psi(z+\sigma h)-\psi(z)\Big)dh + \frac{c}{\sigma}\partial_z\psi(z).$$

By a Taylor expansion,

$$\psi(z+\sigma h)-\psi(z) = \sigma h\partial_z\psi(z) + \frac{\sigma^2h^2}{2}\partial_z^2\psi(\widetilde{z}),$$

for some $z \leqslant \widetilde{z} \leqslant z+\ \sigma h$. Therefore,

$$\begin{aligned}\mathcal{A}^\sigma\psi(z) = \beta\int_{\mathbb{R}} K(h)\exp\left(-\frac{U_\sigma(z+\sigma h)-U_\sigma(z)}{\sigma}\right)h\partial_z\psi(z)dh \\ + \sigma\beta\int_{\mathbb{R}} K(h)\exp\left(-\frac{U_\sigma(z+\sigma h)-U_\sigma(z)}{\sigma}\right)\frac{h^2}{2}\partial_z^2\psi(\widetilde{z})dh + \frac{c}{\sigma}\partial_z\psi(z).\end{aligned} \tag{4.15}$$

Thanks to Barles et al. (2009) and Lorz et al. (2011), we know that U_σ converges locally uniformly towards U, and that U solves the problem (4.6), in the sense of viscosity solutions. Moreover, they show that the following Lipschitz uniform bound holds true:

$$\|\ \partial_z U_\sigma\|_\infty < \eta, \tag{4.16}$$

with η defined in Assumption 1.1. Following the same steps, with viscosity technique, one can show that the first term of (4.15) converges and the second vanishes. With the rescaling of the speed $c = c'\sigma$, one finally gets

$$\mathcal{A}^\sigma\psi(z) \xrightarrow[\sigma\to 0]{} -\beta\int_{\mathbb{R}} K(h)\exp\Big(h\partial_z U(z)\Big)h\partial_z\psi(z)dh + c'\partial_z\psi(z) := \mathcal{A}^0\psi(z). \tag{4.17}$$

Therefore, thanks to this convergence of generators and (Kallenberg, 2006, Theorem 17.25), we conclude that $\mathcal{M}_s^\sigma$, the semigroup associated to $\mathcal{A}^\sigma$ converges to a semigroup $\mathcal{M}_s^0$ associated to the asymptotic operator $\mathcal{A}^0$. To get the precise convergence of Proposition 1.7, we need to be more explicit about $\mathcal{M}^0$.

Let $\mathcal{V}$ be defined by:

$$\mathcal{V}(z) = c' - \beta\partial_p H(\partial_z U(z)), \qquad \text{for all } z\in\mathbb{R},$$

and let Γ be defined as the corresponding integral flow:

$$\begin{cases} \partial_t\Gamma(t,s,z) = \mathcal{V}(\Gamma(t,s,z)), \\ \Gamma(s,s,z) = z. \end{cases} \tag{4.18}$$

We recall the expression of the derivative of the Hamiltonian defined in (1.19):

$$\partial_p H(p) = \int_{\mathbb{R}} yK(y)\exp(yp)dy.$$

One can notice that this term appears in the definition of $\mathcal{A}_0$, in (4.17), within a transport structure. More precisely, for any test function ψ, let $\theta(s,z) = \mathcal{M}_s^0\psi(z)$. Given the expression of $\mathcal{A}^0$, θ then solves by definition the following equation:

$$\begin{cases} \partial_s\theta(s,z) = \mathcal{V}(z)\partial_z\theta(s,z), \quad s>0, z\in\mathbb{R}, \\ \theta(0,z) = \psi(z). \end{cases}$$

Classically, this advection equation with non constant velocity field admits an "explicit" solution, based on the knowledge of the integral flow Γ in (4.18):

$$\theta(s,z) = \psi(\Gamma(s,0,z)). \tag{4.19}$$

Notice that the integral flow Γ, defined in (4.18), coincides by definition with the solution Γ_z of the ODE (1.23) in Proposition 1.7: $\Gamma(s,0,z) = \Gamma_z(s)$. Since $\theta(s,z) = \mathcal{M}_s^0\psi(z)$, the formula (4.19) proves the part of Proposition 1.7 about the convergence of the semigroup.

Furthermore, we claim that the family of Markov processes $(Y_s, s \geqslant 0)$ indexed by the parameter $\sigma \in [0,1]$ is tight for the Skorokhod topology. To prove this, note that

$$Y_s = Y_0 + V_s + M_s,$$

where

$$V_s = c's + \beta \int_0^s \int_{\mathbb{R}} K(h) \frac{F(Y_r + \sigma h)}{F(Y_r)} h \, dh dr$$

and $(M_s, s \geqslant 0)$ is a local martingale with predictable variation

$$\langle M \rangle_s = \sigma\beta \int_0^s \int_{\mathbb{R}} K(h) \frac{F(Y_r + \sigma h)}{F(Y_r)} h^2 dh dr.$$

Using (4.16), we then see that

$$|V_{s'} - V_s| \leqslant \left(c' + \beta \int_{\mathbb{R}} K(h) \exp(\|\partial_z U_\sigma\|_\infty |h|) \, |h| dh \right) |s' - s|,$$

$$|\langle M \rangle_{s'} - \langle M \rangle_s| \leqslant \sigma\beta \int_{\mathbb{R}} K(h) \exp(\|\partial_z U_\sigma\|_\infty |h|) \, h^2 dh \, |s' - s|.$$

Since $\|\partial_z U_\sigma\|_\infty < \eta$ where η is such that

$$\int_{\mathbb{R}} K(h) e^{\eta|h|} dh < +\infty,$$

there exists a constant $C > 0$, independent of σ, such that

$$\int_{\mathbb{R}} K(h) \exp(\|\partial_z U_\sigma\|_\infty |h|) \, |h| \wedge |h|^2 dh \leqslant C.$$

This shows that $(Y_s, s \geqslant 0)$ satisfies the Aldous-Rebolledo criterion for tightness of stochastic processes, see Aldous (1978) and Rebolledo (1980). Since, for each fixed $s \geqslant 0$, Y_s converges in distribution to the deterministic value $\Gamma_z(s)$, this convergence holds also for finite-dimensional marginals of $(Y_s, s \geqslant 0)$. Together with tightness, this yields the convergence in distribution of $(Y_s, s \geqslant 0)$ in the Skorokhod topology to the deterministic process $(\Gamma_z(s), s \geqslant 0)$. Given that the limit is continuous, the convergence also holds in the uniform topology, moreover, since the limit is deterministic, the convergence also holds in probability, hence (1.22). □

Long time behavior of the ancestral process for $\sigma = 0$.
Let us now look at the long time behavior of Y_s^0, the asymptotic ancestral process when $\sigma = 0$. According to our result Y_s^0 is deterministic, given by Γ_z, and therefore we must investigate the limit of Γ_z when $s \to \infty$. We will show independently, that as expected from Proposition 1.5, it converges towards 0, the optimal trait.

To do so, we first look at the stationary state(s) of the ODE (1.23) solved by Γ_z. If 0 is a steady state of this ODE, then necessarily, we should have

$$0 = -c' + \beta \partial_p H(U'(0)).$$

This means that $U'(0)$ is a critical point of the function $p \mapsto -\frac{c'}{\beta} p + H(p)$. From the convexity of H and the definition of the Lagragian function L in (4.7), we deduce that $U'(0)$ should satisfy:

$$-\beta L\left(\frac{c'}{\beta}\right) = -c'U'(0) + \beta H(U'(0)).$$

Plugging this into the problem (4.6) satisfied by U, we conclude that 0 is a steady state of (1.23) implies

$$\lambda = \beta - \mu_0 - L\left(\frac{c'}{\beta}\right). \tag{4.20}$$

Conversely, let us assume that (4.20) holds true. We know, see (4.6) for instance, that (λ, U) satisfies the following equation

$$m(z) + \lambda - \beta + \mu_0 = \beta H(U'(z)) - c'U'(z). \tag{4.21}$$

By evaluating this equation at $z = 0$, we get

$$\lambda = \beta - \mu_0 + \beta H(U'(0)) - c'U'(0). \tag{4.22}$$

Then, by using our assumption made in (4.20) on the expression of λ, we obtain that necessarily

$$-\beta L\left(\frac{c'}{\beta}\right) = -c'U'(0) + \beta H(U'(0)).$$

By definition of the Lagrangian function, see (4.7), this implies that

$$U'(0) = \arg\max_p \left(\frac{c'}{\beta}p - H(p)\right).$$

Therefore, we get by differentiating with respect to p that

$$0 = c' - \beta\partial_p H(U'(0)).$$

Finally, we can conclude that 0 is a steady state of (4.14) if and only if $\lambda = \beta - \mu_0 - \beta L(c'/\beta)$.

Let us now show that this formula for λ is true, that is

$$\lambda = \beta - \mu_0 - L\left(\frac{c'}{\beta}\right). \tag{4.23}$$

We will use convex analysis arguments, see (4.12) and below for a heuristic about this formula. First, the function $p \mapsto c'p - \beta H(p)$ admits a maximum value denoted $\beta L(c'/\beta)$. Adding this value on each side of the equation (4.21), we obtain

$$m(z) + \left[\lambda - \beta + \mu_0 + L\left(\frac{c'}{\beta}\right)\right] = \beta H(U'(z)) - c'U'(z) + L\left(\frac{c'}{\beta}\right). \tag{4.24}$$

On the right hand side, the function $p \mapsto \beta H(p) - c'p + \beta L(c'/\beta)$ is convex, nonnegative and reaches zero from the properties of the Hamiltonian H and the Lagrangian L. For the left hand side of (4.24), this means that the term between brackets must vanish, which gives the desired formula (4.23). Otherwise, the function $z \mapsto \beta H(U')(z) - c'U'(z) + L\left(\frac{c'}{\beta}\right)$ takes only (strictly) positive values. Therefore, U' only takes values in one of the two branches of the convex function $p \mapsto \beta H(p) - c'p + L\left(\frac{c'}{\beta}\right)$. On each of these branches, this function is invertible, and therefore, for each $z \in \bar{\mathbb{R}}$, we can invert the relationship (4.24) to deduce the value taken $U'(z)$. By symmetry of m, this shows that U' is an even function. In turn, this means that

$$\lim_{z\to+\infty} U'(z) = \lim_{z\to-\infty} U'(z).$$

This is in contradiction with the assumption $U(\pm\infty) = +\infty$, or equivalently, $F(\pm\infty) = 0$, *i.e.* the population density vanishes at infinity.

Coming back to our initial problem, we have proved that 0 was a stationary state of (1.23). By convexity of H, $p_0 = U'(0)$ is the only root of the function $\beta H - c'Id$. Moreover, it is established in Lorz et al. (2011) that U, the solution of (4.6) is concave. Therefore U' is invertible, and the equation $p_0 = U'(z)$ has at most a single solution. We conclude that 0 is the unique steady

steady state of the ODE (1.23). With the same monotony arguments, it is straightforward that, Γ converges to the unique steady state of the ODE it solves, and therefore, $\Gamma(s) \xrightarrow[s\to\infty]{} 0$.

Approximation of the mean of the ancestral lineages.
Other information can be gathered from the Hamilton Jacobi equation (4.6), in the form of an analytical approximation of Γ, the typical lineage. Let us first differentiate (4.6) with respect to z, and then divide by $\partial_z^2 U$ on each side. We obtain for all $z \in \mathbb{R}$,

$$-c' + \beta\, \partial_p H\big(\partial_z U(z)\big) = \frac{m'(z)}{\partial_z^2 U(z)}. \tag{4.25}$$

Now, our idea is to link $\partial_z^2 U(z)$ to the variance of the equilibrium F. First, from Bouin et al. (2020), we have the following approximation at the leading order in σ for the variance:

$$\mathrm{Var}(F) = \frac{\sigma}{\partial_z^2 U(z^*)} + o(\sigma), \tag{4.26}$$

where we recall that z^* is the dominant trait in our population. This originates from a Taylor expansion with Laplace's method of the integrals defining the variance (see Wong (2001) for instance):

$$\mathrm{Var}(F) = \frac{\displaystyle\int_{\mathbb{R}} (z-z^*)^2 \exp\left(-\frac{U(z)}{\sigma}\right) dz}{\displaystyle\int_{\mathbb{R}} \exp\left(-\frac{U(z)}{\sigma}\right) dz}.$$

In addition, we make the following rough approximation, valid if z is close to z^*: $\partial_z^2 U(z) \approx \partial_z^2 U(z^*)$. Plugging this and (4.26) into (4.25), we find that

$$-c' + \beta\, \partial_p H\big(\partial_z U(z)\big) \approx \frac{m'(z)\mathrm{Var}(F)}{\sigma} \quad \text{for} \quad z \ \text{ close to } \ z^*.$$

We eventually plug this into the ODE for Γ_z, (1.23), to find the following approximation

$$\begin{aligned} \dot\Gamma_z(s) &= -\frac{\mathrm{Var}(F)}{\sigma} m'\big(\Gamma_z(s)\big), \\ \Gamma_z(0) &= z. \end{aligned} \tag{4.27}$$

In particular, **if the selection function is quadratic**, $m(z) = z^2/2$, then the solution of this ODE is

$$\Gamma_z(s) = z \exp\left(-\frac{\mathrm{Var}(F)}{\sigma}\, s\right). \tag{4.28}$$

In fact, working on the Hamilton Jacobi equation, we can derive a completely explicit expression for Γ, thanks to a formula for the variance:

$$\mathrm{Var}(F) = \frac{c'\sigma}{\sqrt{2\beta}\sqrt{L\left(\dfrac{c'}{\beta}\right)}}. \tag{4.29}$$

To get this equality, we first use (4.6) to compute the second derivative of U at z^*, and then plug it into (4.26), to find, as in Bouin et al. (2020), that

$$\mathrm{Var}(F) = -\frac{c'\sigma}{m'(z^*)}. \tag{4.30}$$

Note that this remains true for any selection function m that satisfies our hypotheses. We also have another equation that dictates the position of the dominant trait z^*, stated in (4.9). In the

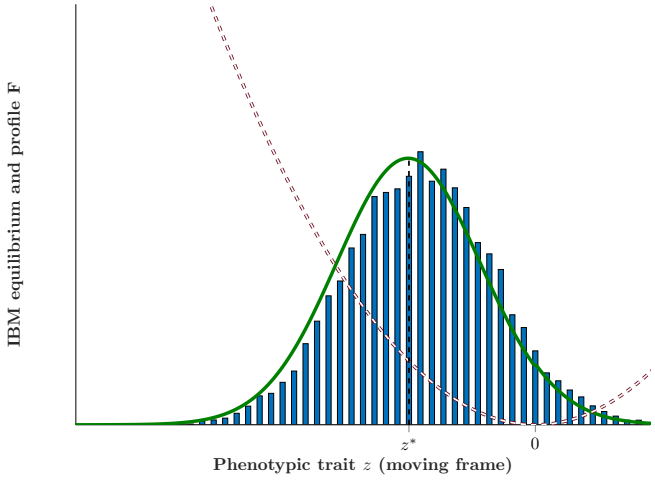


FIGURE 6. In continuous green is the solution F of (1.4), and in blue the histogram of one simulation of the IBM, with about 20000 individuals, see Section 6 for more numerical details. In dashed red, the corresponding selection function.

particular case of the quadratic selection, we can use it to compute z^*, and then thanks to (4.30), we find (4.29).

By plugging (4.29) into (4.28), we find the following approximation for the mean trajectories of the ancestral process, when the selection function is quadratic but the mutation kernel is general:

$$\Gamma_z(s) = z \exp\left(- \frac{c' \, s}{\sqrt{2\beta}\sqrt{L\left(c'/\beta\right)}} \right). \tag{4.31}$$

This formula is completely explicit, since it only depends on the mutation kernel through the Lagrangian L. However it is an approximation only valid for the case of quadratic selection. This approximation turns out to be quite robust in our simulations, see the discussion in the next section.

4.3. **Numerical methods and simulations.**

Thanks to Proposition 1.7 we have justified the interpretation of the Lagrangian structure introduced in section 4.1. We have also established an equation on the typical lineage in this regime, Γ, solution of (1.23). The aim of this section is first to look at the accuracy of our approximation in a regime which might be far away from small variance. And secondly, we compare our deterministic model and our approximation with a classical individual-based model Champagnat et al. (2006, 2007).

The population is described by a vector of individuals, each labeled by its trait. Each individual has a birth and a death clock, that depends of its phenotypic trait. The size of population must be large enough to be considered infinite, and even when at equilibrium, it continues to fluctuate around its deterministic value. In Figure 6, we provide a comparison between the result of one simulation of the Individual Based Model, and the corresponding deterministic equilibrium. We observe a good fit between the equilibrium F obtained by a simulation of (1.1) and the distribution of the IBM, pictured by an histogram. We further expect that simulations of the IBM prove to be a good match for the lineages of individuals in the equilibrium, even in the regime of Proposition 1.7. Therefore, we also keep track of the lineages inside the IBM, via a numerical procedure described in Section 6.

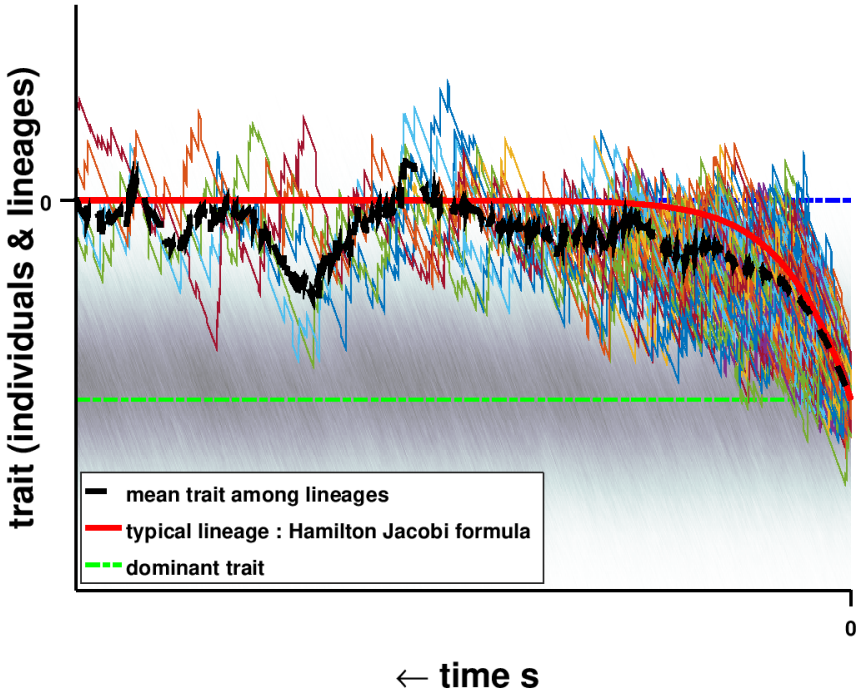


FIGURE 7. The $y-$axis represents the phenotype z in the moving frame ($z = x - ct$). The grey background is the (stationary) distribution of individuals according to F, as simulated by the Individual Based Model. Each line with a different color represent the lineage of an individual along time, there are many individuals when the population is sampled at $s = 0$, on the right hand side (sample among all individuals that have the dominant trait, pictured in light green), and fewer after coalescence and a large time s in the lineage (on the left hand side). The equilibrium has about 20000 individuals, see section 6 for more details.

This leads to the main numerical result of this section, depicted in Figure 7. A simulated profile F, gives in turn an approximated function U solution of (4.10). We are then able to solve numerically (1.23), and we plot the theoretical lineage (in red) on Figure 7. We compare that to the stochastic lineages that were obtained by one simulation of the IBM. Time of the simulation increases to the right, but, of course, the time of the lineages is reversed. Therefore, the initial condition of the process Y_s is on the right of the figure, individuals were sampled at the dominant trait. Going left, as the time corresponding to the ancestors increase, we observe a similar behavior between the mean (in black) of the lineages, and the theoretical one in red. One observes more noise in the simulated lineages after a certain time, when there are fewer individuals. This is consistent with the foundation (coalescence) effect for our model we explained by crefcor explicit asymp, since asymptotically in time the most likely ancestor has the optimal trait, with very few optimally fitted individuals.

We selected z^* as the starting point of the lineages, in order to sample initially as many lineages as possible. From around $20,000$ individuals inside equilibrium, there were around 1,000 initial lineages, and based on Figure 7, no more than a handful of different common ancestors going back in time, which explains the eventual difference we observe between the Hamilton-Jacobi formula and the simulation.

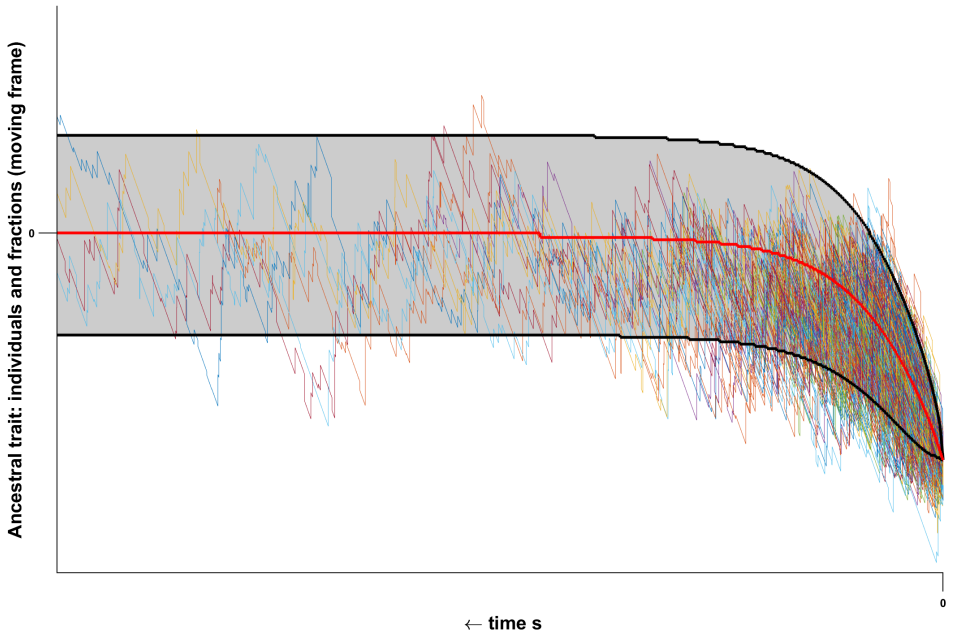


FIGURE 8. The red line is the mean of the ancestral process Y_s, obtained by simulations of the PDE (1.16). The black lines are the 5% and the 95% quantile of the distribution of Y_s. Every other line with a different color represent the lineage of an individual in a simulation of the Individual Based Model, as in Figure 7. About 20000 individuals, see section 6 for more details.

Another factor of discrepancy is that σ is not 0 in the simulations, contrary to the regime where Proposition 1.7 is valid. However, even if σ is not really small, we have a good fit of the deterministic model and its approximation to the stochastic model, we provide a variety of parameters in REF FIG. Finally, our analysis cannot capture the coalescence events in the genealogical tree arising in the individual based simulations, contrary to what is done in Henry et al. (2020). The reason for this is that we only consider the deterministic large population limit of the process, and that coalescence events vanish in this limit.

We provide a comparison between the numerical results of Section 3.3, *i.e.* a PDE approach of the ancestral process Y_s and the results of the IBM. This is pictured in Figure 8. Once again we only show one realization of the lineages obtained by the IBM, superposed with the statistics of simulations of (1.16), which were already shown in Figure 5, with the same parameters. We observe a good fit between our PDE description of Y_s, and the distribution of lineages of the individual-based model along time.

To conclude, we provide a discussion around other sets of parameters in the Figure 10, with a focus around the mean and the variance of the ancestral process Y_s.

5. Discussion

We proposed a methodology to track genealogies inside a mathematical model of mutation and selection. The ecological limitations of our study originate from the underlying choices of modeling in the equation (1.1). In that regard, we believe that the assumptions made are fairly general, at least from a mathematical point of view. The birth rate is assumed constant and independent of

phenotype, contrary to the mortality rate (selection). From a modeling perspective, this means that the trait appears through the Malthusian growth rate M, defined as the difference between the birth and death rate: $M(z) = \beta(z) - \mu(z)$. Therefore up to a change of function μ, we can always assume that the birth rate is constant and that selection acts only on mortality. We do not assume an explicit shape for the selection function μ, contrary to, for instance, the celebrated Fisher's Geometric Model (FGM). This model assumes an explicit quadratic relationship between phenotype and fitness. Recently it was used in a PDE setting not far from ours with local or non-local effect of mutations, Martin and Roques (2016); Gil et al. (2019), to derive explicit formulas for the full trajectory of mean fitness. We do not consider the so called 'diffusive approximation', Kimura (1964), which would consist in $\mathcal{B}(f) = f + \Delta f$ in (1.2). Instead, this case can be encapsulated in our analysis as we detail in Section 3.2. More generally, we study a non-local convolution operator, which can encapsulate many kernels and different scenarios. This is even more noticeable in the regime of adaptive dynamics, *i.e.* when $\sigma \to 0$, as in Proposition 1.7 and Section 4. This is traditionally the hypothesis made to justify the diffusive approximation. However, we observe that there is still a (nonlocal) trace of the kernel in this regime, *e.g.* directly on the formula for the typical lineage in this regime, (1.23). This is also clear on the formula for the eigenvalue λ in this regime, see (4.23). This highlights a qualitative difference between our model and the diffusive approximation, even in the small variance regime, see also the numerical results in the Figure 10.

Keeping in mind those modeling caveats, our methodology is relatively robust. Using the tools that we deployed, we believe that it is possible to investigate the inside dynamics of a broad range of equilibria. Our study has confirmed that this kind of analysis provides crucial information. In our case, we learned that the source of genetic diversity in the equilibrium lies in the individuals at the phenotype optimum. This means that persistence of the equilibrium depends on a scarce number of individuals, because of the lag load. This phenomenon has been also observed for the genealogies in model of continual adaptation in asexual populations (Nordborg, 1997; Hermisson et al., 2002; Rouzine and Coffin, 2007; Neher and Hallatschek, 2013). We showed that our results were compatible with the popular adaptive dynamics regime. In that case, we established an explicit formula, valid when $\sigma = 0$. In particular, it showed that the structure of Hamilton Jacobi equations was rich enough to deliver information upon the genealogies during adaptation. This opens a broad range of different applications for our method, since this 'Hamilton-Jacobi approach' has been used widely for different adaptation models. As a matter of fact, before appearing in Diekmann et al. (2005) in an evolutionary context, it was first introduced to study the propagation speed of reaction diffusion equations, by Freidlin (1985); Evans and Souganidis (1989). It would be fascinating to use the methodology of neutral fractions in that context, as in Garnier et al. (2012), to recover a dichotomy on pulled and pushed fronts directly from the Hamilton-Jacobi equation in the regime of vanishing viscosity. Going back to ecological contexts, it was recently proposed to model mutations with a kernel K as in (1.2), but which does not satisfy the Assumption 1.1 of being *exponentially bounded*. In particular Méléard and Mirrahimi (2015) found the (non-stationary) Hamilton-Jacobi equation when $\sigma = 0$, in the case of the fractional Laplacian: $\mathcal{B}(f) = (-\Delta)^{\alpha}(f)$. This was later expanded to a broad range of 'fat tailed' kernels by Bouin et al. (2018); Mirrahimi (2020). However, in that case, the 'dual' Lagrangian point of view of the equation is no longer valid, and as a matter of fact the Hamiltonian H, in (1.19), is no longer well defined. Our investigation of genealogies, based on neutral fractions, can still take place, and it would lead an insight that is not accessible from heuristics, even in the regime $\sigma \to 0$. In addition, from a mathematical point of view, our analysis encapsulates very general operators $\mathcal{B}$. In fact, we mostly use the linearity of $\mathcal{B}$ and some spectral results about the linearized operator $\mathcal{L}$, that hold true quite broadly, see Mischler and Scher (2016).

A missing feature in our model is spatial heterogeneity. Recently, a lot of work has focused on the interaction of two populations living in two different habitats, see Mirrahimi and Gandon (2020);

Hamel et al. (2020). Each patch has a different optimum and it is possible to travel from one patch to the other at a given migration rate. A common observation in both works is that positive migration rates shift the distribution towards the optimum of the other patch, influenced by migrants from the other patch. In a certain regime of parameters, polymorphism in the population may appear. Knowing the genealogies of each population and the patch of ancestors could better explain this kind of phenomenon. Recently, Garnier and Lafontaine (2020) have extended the notion of neutral fraction to metapopulation model, thus we believe that the neutral fractions toolbox can be used to describe the genealogy of a metapopulation located in different areas. More generally, we would like to look into models of adaptation with continuous space and trait variable, such as the 'cane toad equation', Bénichou et al. (2012); Bouin et al. (2017). The dynamics of the propagation front appears to be driven by optimally fitted individuals, a phenomenon called 'spatial sorting', but few is known about the mechanisms of this effect, to the extent that recently, it was discovered that with a non local competition term, the acceleration of the front is slowed down, Calvez et al. (2018), contrary to what was first formally investigated in Bouin et al. (2012). Our methodology to study inside dynamics could shed some light on this model.

Finally, a far reaching perspective is to expand our study of genealogies to sexual populations. To that end, an appropriate operator in our context is given by the infinitesimal model. It consists in, with the same notations of (1.2):

$$\mathcal{B}(f)(x) := \frac{1}{\sigma\sqrt{\pi}} \iint_{\mathbb{R}^2} \exp\left[-\frac{1}{\sigma^2}\left(x - \frac{x_1+x_2}{2}\right)^2\right] f(x_1) \frac{f(x_2)}{\int_{\mathbb{R}} f(x_2')\,dx_2'}\,dx_1 dx_2. \tag{5.1}$$

This operator describes how an offspring with trait x appears in the population. It stems from a first parent with trait x_1, and a second parent x_2, chosen uniformly in the population. Then, the trait of the offspring is drawn from a Gaussian normal law centered around the mean of the trait of the parents: $(z_1 + z_2)/2$, with variance σ^2. It is known in the literature as a bridge between Mendelian genetics and a statistical approach to inheritance, Turelli (2017). Recently it was derived from a microscopic point of view in Barton et al. (2017), and asymptotically studied in a regime of adaptive dynamics, Calvez et al. (2019); Patout (2020), which lays the groundwork for a study with neutral fractions. Of course, for each individual in a population with a sexual mode of reproduction, a genealogy consists in a tree of size 2^N, where N is the number of generations. This shows that, compared to this article, more elaborate tools must be introduced to determine the typical phenotype of ancestors. Mathematically, the operator (5.1) is no longer linear, which complicates the definition of neutral fractions as in (1.3). One must introduce a linear operator to study fractions, and we propose:

$$\mathcal{B}_F(v)(x) := \frac{1}{\sigma\sqrt{\pi}} \iint_{\mathbb{R}^2} \exp\left[-\frac{1}{\sigma^2}\left(x - \frac{x_1+x_2}{2}\right)^2\right] v(x_1) \frac{F(x_2)}{\int_{\mathbb{R}} F(x_2')\,dx_2'}\,dx_1 dx_2. \tag{5.2}$$

This operator has all the necessary qualitative properties of $\mathcal{L}$ defined in (1.16), and we could investigate lineages similarly to (1.3).

All the population models considered in the present paper can be derived as large population limits of stochastic individual-based models. Hence it is natural to ask whether neutral fractions can be investigated directly in the stochastic framework, in order to gain more insight into the dynamics of small (or moderately large) populations. One could in particular hope to retrieve some knowledge on the timescale of coalescence in the genealogy of a sample of individuals from the population. For instance, looking at Figure 7, one may ask how far back in the past one has to look in order to find a common ancestor of all the individuals occupying some spatial region at present. Such questions are very hard to answer in the deterministic large population limit framework, since in this regime the coalescence events become increasingly rare.

However, using some of our results, we can provide heuristics about the mean time of coalescent of two lineages, T_2. First, our analytical approximations (3.9) and (4.31) on the dynamics of the mean the ancestral lineage provides a good approximation on the characteristic time T_c before which two lineages reach the optimal value 0. From our diffusive approximation we get

$$T_c \approx \frac{1}{\sigma\sqrt{\beta}}$$

In the case of more general operators, based on the approximation formula (4.28) in the small variance regime, we get

$$T_c \approx \frac{1}{\mathrm{Var}(F)},$$

where $\mathrm{Var}(F)$ is the variance of the phenotypic distribution at equilibrium. Thus the time of coalescence should be rescaled by T_c the delay before coalescence occurs. Then, once the ancestral lineage are close to the optimal, the ancestral lineage have reached the stationnary distribution of Y_∞. Using the heuristic of Etheridge and Penington (2020) and Neher and Hallatschek (2013), combined with our Proposition 1.5, one could expect that the rescaled time of coalescence $T_2 - T_c$ should be approximately exponential, with parameter

$$\int_{\mathbb{R}} \frac{1}{NF(z)}(F(z)\varphi(z))^2 dz.$$

We can also compute this parameter using either our diffusive approximation or our Hamilton-Jacobi approximation. These argument are all formal for now, and need to be further investigated.

On the other hand, studying neutral fractions in stochastic population models turns out to be mathematically challenging and poses some serious analytical difficulties. An example of a recent work in this direction is Billiard et al. (2015), who study neutral markers in the background of a trait-substitution sequence in the adaptive dynamics regime.

More recently, Etheridge and Penington (2020) studied a spatial Moran process mimicking an expanding population with a strong Allee effect (corresponding to a bistable reaction-diffusion equation), and showed, using neutral markers, that the genealogy of a sample of individuals taken at the front of the wave follows the classical Kingman coalescent with a scaling parameter given in terms of the coefficients of the equation. This result is in stark contrast with what is known about pulled waves (as in the stochastic Fisher-KPP equation) where the genealogy of a sample of individuals at the front follows a Bolthausen-Sznitman coalescent Berestycki et al. (2013). In our moving optimum setting, there are reasons to believe that the genealogy of a sample of individuals at the front of the population follows a Kingman coalescent, as in Etheridge and Penington (2020). The reason for this is that individuals all descend from ancestors whose trait was close to the optimum, and if an individual gets too much ahead of the front, it will find itself at a disadvantage compared to individuals closer to the optimum. Thus the inside dynamics in our model has a "pushed" nature, and it is easy to check that the integrability conditions required in Etheridge and Penington (2020) are satisfied in our case. Making this argument rigorous is of course an entirely different matter, and will be the subject of future work.

References

Aldous, D. (1978). Stopping times and tightness. *The Annals of Probability*, 6(2):335–340.

Alfaro, M., Berestycki, H., and Raoul, G. (2017). The effect of climate shift on a species submitted to dispersion, evolution, growth, and nonlocal competition. *SIAM Journal on Mathematical Analysis*, 49(1):562–596.

Alfaro, M. and Carles, R. (2014). Explicit solutions for replicator-mutator equations: Extinction versus acceleration. *SIAM Journal on Applied Mathematics*, 74(6):1919–1934.

Bansaye, V., Cloez, B., and Gabriel, P. (2019). Ergodic behavior of non-conservative semigroups via generalized doeblin's conditions. *Acta Applicandae Mathematicae*, pages 1–44.

Barles, G., Mirrahimi, S., and Perthame, B. (2009). Concentration in Lotka-Volterra parabolic or integral equations: a general convergence result. *Methods and Applications of Analysis*, 16(3):321–340.

Barles, G. and Roquejoffre, J.-M. (2006). Ergodic type problems and large time behaviour of unbounded solutions of hamilton–jacobi equations. *Communications in Partial Differential Equations*, 31(8):1209–1225.

Barton, N., Etheridge, A., and Véber, A. (2017). The infinitesimal model: Definition, derivation, and implications. *Theoretical population biology*, 118:50–73.

Barton, N. H. (1998). The effect of hitch-hiking on neutral genealogies. *Genetical Research*, 72(2):123–133.

Barton, N. H. (2000). Genetic hitchhiking. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 355(1403):1553–1562.

Barton, N. H. and Etheridge, A. M. (2004). The effect of selection on genealogies. *Genetics*, 166(2):1115–1131.

Bénichou, O., Calvez, V., Meunier, N., and Voituriez, R. (2012). Front acceleration by dynamic selection in fisher population waves. *Physical Review E*, 86(4):041908.

Berestycki, H., Diekmann, O., Nagelkerke, C. J., and Zegeling, P. A. (2009). Can a species keep pace with a shifting climate? *Bull Math Biol*, 71(2):399–429.

Berestycki, H. and Fang, J. (2018). Forced waves of the fisher–KPP equation in a shifting environment. *Journal of Differential Equations*, 264(3):2157–2183.

Berestycki, J., Berestycki, N., and Schweinsberg, J. (2013). The genealogy of branching Brownian motion with absorption. *The Annals of Probability*, 41(2):527–618.

Billiard, S., Ferrière, R., Méléard, S., and Tran, V. C. (2015). Stochastic dynamics of adaptive trait and neutral marker driven by eco-evolutionary feedbacks. *Journal of Mathematical Biology*, 71(5):1211–1242.

Bouin, E., Bourgeron, T., Calvez, V., Cotto, O., Garnier, J., Lepoutre, T., and Ronce, O. (2020). Equilibria of quantitative genetics models beyond the gaussian approximation i: Maladaptation to a changing environment. *In preparation*.

Bouin, E., Calvez, V., Meunier, N., Mirrahimi, S., Perthame, B., Raoul, G., and Voituriez, R. (2012). Invasion fronts with variable motility: phenotype selection, spatial sorting and wave acceleration. *Comptes Rendus Mathematique*, 350(15-16):761–766.

Bouin, E., Garnier, J., Henderson, C., and Patout, F. (2018). Thin front limit of an integro-differential fisher-kpp equation with fat-tailed kernels. *SIAM Journal on Mathematical Analysis*, 50(3):3365–3394.

Bouin, E., Henderson, C., and Ryzhik, L. (2017). Super-linear spreading in local and non-local cane toads equations. *Journal de mathématiques Pures et Appliquées*, 108(5):724–750.

Brunet, E., Derrida, B., Mueller, A. H., and Munier, S. (2007). Effect of selection on ancestry: An exactly soluble case and its phenomenological generalization. *preprint*, 76(4):041104.

Calvez, V., Garnier, J., and Patout, F. (2019). Asymptotic analysis of a quantitative genetics model with nonlinear integral operator. *Journal de l'École polytechnique — Mathématiques*, 6:537–579.

Calvez, V., Henderson, C., Mirrahimi, S., Turanova, O., and Dumont, T. (2018). Non-local competition slows down front acceleration during dispersal evolution. *arXiv preprint arXiv:1810.07634*.

Carrère, C. and Nadin, G. (2020). Influence of mutations in phenotypically-structured populations in time periodic environment. *Discrete & Continuous Dynamical Systems-B*, 22(11).

Champagnat, N., Ferrière, R., and Méléard, S. (2006). Unifying evolutionary dynamics: from individual stochastic processes to macroscopic models. *Theoretical population biology*, 69(3):297–321.

Champagnat, N., Ferrière, R., and Méléard, S. (2007). Individual-based probabilistic models of adaptive evolution and various scaling approximations. In *Seminar on Stochastic Analysis, Random Fields and Applications V*, pages 75–113. Springer.

Champagnat, N., Henry, B., et al. (2019). A probabilistic approach to dirac concentration in nonlocal models of adaptation with several resources. *The Annals of Applied Probability*, 29(4):2175–2216.

Cloez, B. and Gabriel, P. (2019). On an irreducibility type condition for the ergodicity of nonconservative semigroups. *arXiv preprint arXiv:1909.07363*.

Collot, D., Nidelet, T., Ramsayer, J., Martin, O. C., Méléard, S., Dillmann, C., Sicard, D., and Legrand, J. (2018). Feedback between environment and traits under selection in a seasonal environment: consequences for experimental evolution. *Proceedings of the Royal Society B: Biological Sciences*, 285(1876):20180284.

Coville, J. and Hamel, F. (2019). On generalized principal eigenvalues of nonlocal operators witha drift. *Nonlinear Analysis*, page 111569.

Desai, M. M., Walczak, A. M., and Fisher, D. S. (2013). Genetic diversity and the structure of genealogies in rapidly adapting populations. *Genetics*, 193(2):565–585.

Diekmann, O., Jabin, P.-E., Mischler, S., and Perthame, B. (2005). The dynamics of adaptation: an illuminating example and a Hamilton-Jacobi approach. *Theoretical Population Biology*, 67(4):257–271.

Etheridge, A. and Penington, S. (2020). Genealogies in bistable waves. *arXiv:2009.03841 [math]*, 2009.03841.

Ethier, S. N. and Kurtz, T. G. (2009). *Markov processes: characterization and convergence*, volume 282. John Wiley & Sons.

Evans, L. and Souganidis, P. (1989). A PDE approach to geometric optics for certain semilinear parabolic equations. *Indiana University mathematics journal*, 38(1):141–172.

Figueroa Iglesias, S. and Mirrahimi, S. (2018). Long time evolutionary dynamics of phenotypically structured populations in time-periodic environments. *SIAM Journal on Mathematical Analysis*, 50(5):5537–5568.

Figueroa Iglesias, S. and Mirrahimi, S. (2019). Selection and mutation in a shifting and fluctuating environment. *HAL Preprint 02320525*.

Fournier, N., Méléard, S., et al. (2004). A microscopic probabilistic description of a locally regulated population and macroscopic approximations. *The Annals of Applied Probability*, 14(4):1880–1919.

Freidlin, M. (1985). Limit theorems for large deviations and reaction-diffusion equations. *The Annals of Probability*, pages 639–675.

Garnier, J., Giletti, T., Hamel, F., and Roques, L. (2012). Inside dynamics of pulled and pushed fronts. *Journal de mathématiques pures et appliquées*, 98(4):428–449.

Garnier, J. and Lafontaine, P. (2020). Dispersal and good habitat quality promote neutral genetic diversity in metapopulations. *arXiv preprint*.

Garnier, J. and Lewis, M. A. (2016). Expansion under climate change: the genetic consequences. *Bulletin of mathematical biology*, 78(11):2165–2185.

Gil, M.-E., Hamel, F., Martin, G., and Roques, L. (2019). Dynamics of fitness distributions in the presence of a phenotypic optimum: an integro-differential approach. *Nonlinearity*, 32.

Hairer, E., Lubich, C., and Wanner, G. (2006). *Geometric numerical integration: structure-preserving algorithms for ordinary differential equations*, volume 31. Springer Science & Business Media.

Hallatschek, O. and Nelson, D. R. (2008). Gene surfing in expanding populations. *Theoretical population biology*, 73(1):158–170.

Hallatschek, O. and Nelson, D. R. (2010). Life at the front of an expanding population. *Evolution: International Journal of Organic Evolution*, 64(1):193–206.

Hamel, F., Lavigne, F., and Roques, L. (2020). Adaptation in a heterogeneous environment. i: Persistence versus extinction. *arXiv preprint arXiv:2005.09869*.

Henry, B., , Méléard, S., and Tran, V. C. (2020). Dynamics of lineages in climate change adaptation. *In preparation, private communication*.

Hermisson, J., Redner, O., Wagner, H., and Baake, E. (2002). Mutation–selection balance: Ancestry, load, and maximum principle. *Theoretical Population Biology*, 62(1):9 – 46.

Hoffmann, A. A. and Sgro, C. M. (2011). Climate change and evolutionary adaptation. *Nature*, 470(7335):479.

Kallenberg, O. (2006). *Foundations of modern probability*. Springer Science & Business Media.

Kaplan, N. L., Hudson, R. R., and Langley, C. H. (1989). The "hitchhiking effect" revisited. *Genetics*, 123(4):887–899.

Kimura, M. (1964). Diffusion models in population genetics. *Journal of Applied Probability*, 1(2):177–232.

Lorenzi, T., Chisholm, R. H., Desvillettes, L., and Hughes, B. D. (2015). Dissecting the dynamics of epigenetic changes in phenotype-structured populations exposed to fluctuating environments. *Journal of Theoretical Biology*, 386:166–176.

Lorz, A., Mirrahimi, S., and Perthame, B. (2011). Dirac mass dynamics in multidimensional nonlocal parabolic equations. *Communications in Partial Differential Equations*, 36(6):1071–1098.

Marguet, A. (2019). Uniform sampling in a structured branching population. *Bernoulli*, 25(4A):2649–2695.

Martin, G. and Roques, L. (2016). The non-stationary dynamics of fitness distributions: Asexual model with epistasis and standing variation. *Genetics*, 204(4):1541–1558.

Méléard, S. and Mirrahimi, S. (2015). Singular limits for reaction-diffusion equations with fractional Laplacian and local or nonlocal nonlinearity. *Communications in Partial Differential Equations*, 40(5):957–993.

Mirrahimi, S. (2020). Singular limits for models of selection and mutations with heavy-tailed mutation distribution. *Journal de Mathématiques Pures et Appliquées*, 134:179–203.

Mirrahimi, S. and Gandon, S. (2020). Evolution of specialization in heterogeneous environments: equilibrium between selection, mutation and migration. *Genetics*, 214(2):479–491.

Mischler, S. and Scher, J. (2016). Spectral analysis of semigroups and growth-fragmentation equations. In *Annales de l'Institut Henri Poincare (C) Non Linear Analysis*, volume 33, pages 849–898. Elsevier.

Neher, R. A. and Hallatschek, O. (2013). Genealogies of rapidly adapting populations. *Proceedings of the National Academy of Sciences*, 110(2):437–442.

Nordborg, M. (1997). Structured coalescent processes on different time scales. *Genetics*, 146(4):1501–1514.

Parmesan, C. (2006). Ecological and evolutionary responses to recent climate change. *Annu. Rev. Ecol. Evol. Syst.*, 37:637–669.

Patout, F. (2020). The cauchy problem for the infinitesimal model in the regime of small variance. 2001.04682.

Perthame, B. and Barles, G. (2008). Dirac concentrations in lotka-volterra parabolic pdes. *Indiana University Mathematics Journal*, pages 3275–3301.

Rebolledo, R. (1980). Sur l'existence de solutions à certains problèmes de semi-martingales.

Roques, L., Garnier, J., Hamel, F., and Klein, E. K. (2012). Allee effect promotes diversity in traveling waves of colonization. *Proceedings of the National Academy of Sciences*, 109(23):8828–8833.

Roques, L., Patout, F., Bonnefon, O., and Martin, G. (2020). Adaptation in general temporally changing environments. 2002.09542.

Rouzine, I. and Coffin, J. (2007). Highly fit ancestors of a partly sexual haploid population. *Theoretical Population Biology*, 71(2):239 – 250.

Smith, J. M. and Haigh, J. (1974). The hitch-hiking effect of a favourable gene. *Genetical Research*, 23(1):23–35.

Stokes, A. (1976). On two types of moving front in quasilinear diffusion. *Mathematical Biosciences*, 31(3-4):307–315.

Turelli, M. (2017). Commentary: Fisher's infinitesimal model: A story for the ages. *Theoretical Population Biology*, 118:46–49.

Wong, R. (2001). *Asymptotic approximations of integrals*. SIAM.

6. Appendix

We detail in this Appendix how we deal with the simulations of the individual based model and the lineages. Population at each time t is made of a number $N(t)$ of alive individuals. Let us consider an individual, denoted i, that is alive at time t, with $1 \leqslant i \leqslant N(t)$. It has the trait $z_i \in \mathbb{R}$. The outcome for this individual is one of the following :

- ▷ **Birth of a descendant:** happens with an uniform rate among individuals $r_B^i = \beta$.
- ▷ **Death of the individual i:** the possibility of dying is decomposed in two separate events:
  - (a) **Death by selection** The individual may die because its phenotype is really ill-adapted in the phenotype landscape. This happens at the rate:
$$r_{Ds}^i = \mu(z_i - ct).$$
  - (b) **Death by density dependence** Alternatively, an individual may die because of the density dependence in the population, at a rate that is prescribed by the size of population at time t and the carrying capacity N
$$r_{Ddd} = \sum_{j=1}^{N(t)} \frac{1}{N} = \frac{N(t)}{N}.$$

Next event : incrementation of the time step The time step is the smallest time for all individuals to go through one of the previous steps. Thanks to the Markov property, each event occurs following an exponential law of parameter dictated by its rate r.

$$dt \sim \min_{1 \leqslant i \leqslant N(t)} \mathcal{E}(r_B^i + r_{Ds}^i + r_{Ddd}). \tag{6.1}$$

By the property of "absence of memory" of the exponential law, dt also can be drawn from an exponential law which rate is the sum of the rates of all the independent events:

$$dt \sim \mathcal{E}\left(\sum_i (r_B^i + r_{Ds}^i + r_{Ddd})\right).$$

Alternatively, one can present the individual based process as the following point measure (empirical density):

$$\nu_t^K(dz) := \frac{1}{N} \sum_{j=1}^{N_t^N} \delta_{z_i(t)}(dz),$$

and describe its evolution through its actions on measurable bounded functions by the generator of ν_t^K.

Actualization of the population: Once the next event is decided, according to the law (6.1), the population at time $t + dt$ is deduced by either adding the individual that was born ($N(t + dt) = N(t) + 1$) or subtracting the one that died ($N(t + dt) = N(t) - 1$). In the case of a

birth event, the trait of the offspring is drawn according as prescribed by the operator $\mathcal{B}$ in (1.2), as:

$$z_{\text{offspring}} = z_{\text{parent}} + \sigma dK.$$

We repeat all the steps until reaching the desired final time of simulation.

Numerically, this model has a very high computational cost, because it needs a relatively high number of individuals to best fit the deterministic model given by (1.1). As a consequence, we performed the simulations using an approximated model, by first fixing dt to a small but deterministic value. Then, for each individual, we draw a time of birth following the law $\mathcal{E}(\beta)$ and a time of death following the law $\mathcal{E}(\mu(z_i) + N(t)/N)$. Then we simply count which individuals led to a reproduction event and which died on the time-window $[t, t + dt]$. This amounts to the supposition that on this interval of time, individuals cannot reproduce more than once. This algorithm led to the figure 6.

Finally, let's explain how we follow the lineage of individuals. We create a huge matrix at the beginning which is where we will stock the lineage of every individual along the simulation. Every time an individual appears, its lineage is the one of its parent, translated by one generation. The numerical procedure works as described in Figure 9, where each line corresponds to a generation for each individual, without distinction of the time of the birth event, which has to be recorded separately.

This procedure lead to the Figure 7. The following parameters were used :

$$\alpha = 2, \quad \beta = 2, \quad \sigma = 0.1, \quad N = 20000, \quad c = 0.2\,.$$

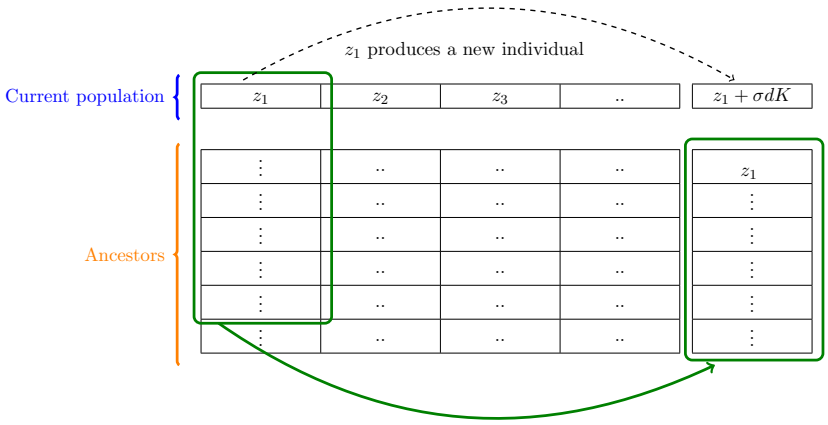


FIGURE 9. Keeping track of the lineages : a tentative explicative drawing

To conclude, let us discuss the effects of a different set of parameters compared to the one we previously mentioned. We look at the effect of the speed c and the variance σ, as they are the two key parameters that must be well adapted in the adaptive dynamics regime we described. Our results are presented in Figure 10. On the left panel, we present the evolution along time of the mean of the ancestral process Y_s, and on the right panel the evolution of its variance. In red are the deterministic simulations. The continuous red is obtained by the simulations of the non-local PDE (1.16), as in the figures of Section 3.3. It is superposed, in dotted red with the formula for Γ of (4.28). We recall this is obtained by an approximation, with quadratic selection of the Hamilton Jacobi formula (1.23) of Proposition 1.7. Finally, we plot in dashed red the mean

of lineages obtained in the diffusive approximation, which corresponds to the model described in (3.3). In black, is the average trajectory, over 50 replicates of the IBM, of the mean (respectively the variance) of ancestors among lineages. The grey shadow envelop corresponds to the confidence interval of the mean (respectively the variance) from 5% to 95% given by 50 replicates of the IBM.

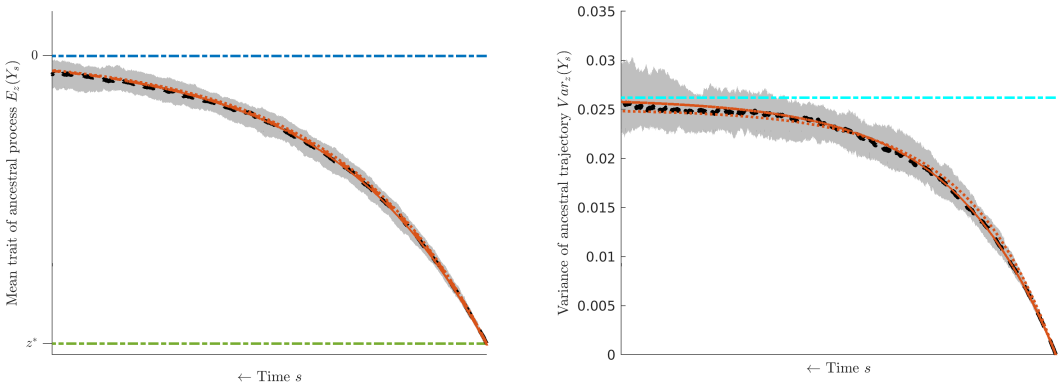
Fist, one observes that the dotted red line which represents the approximation of the Hamilton Jacobi formula is barely visible, and almost completely overshadowed by the continuous red line depicting the result of the PDE. This is a little surprising, since we are not considering the specific regime where the Hamilton-Jacobi approximation is supposed to hold, which consists in $c, \sigma \ll 1$ for the adaptive dynamics and $z \sim z^*$ for the approximation.

As expected, compared to the case of one realization of the IBM in Figures 7 and 8, there is a better match for all times between the averaged IBM simulations and the PDE simulations. We still observe a better match for small times s. However, the confidence interval for the variance is considerably larger for all parameters compared to the one for the mean. We observe more stochasticity for the evolution of the variance than for the evolution of the mean among the replicates of the IBM. As a matter of fact, the diffusive approximation paints also a different picture more clearly in the evolution of the variance. Predictably, the non local model consistently presents an higher variance than the diffusive approximation, even in the regime $\sigma \to 0$.

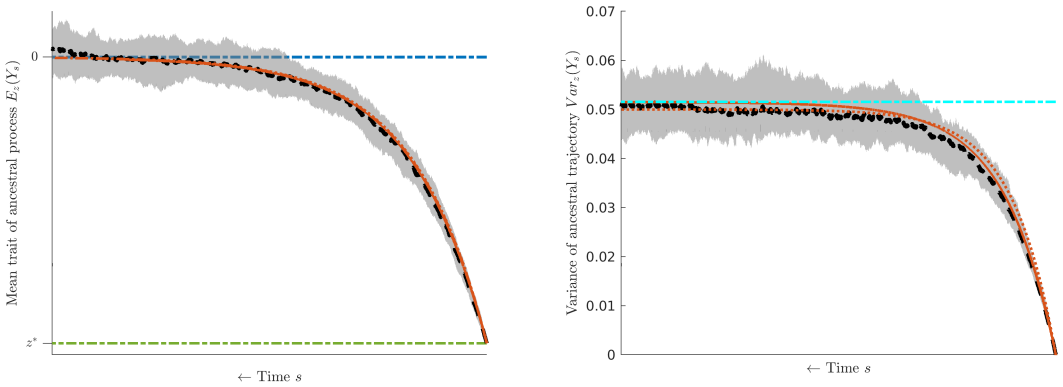
BioSP, INRAE, 84914, Avignon France
Email address: raphel.forien@inrae.fr

LAMA, UMR 5127 CNRS & Univ. Savoie Mont-Blanc, Chambéry, France
Email address: jimmy.garnier@univ-smb.fr

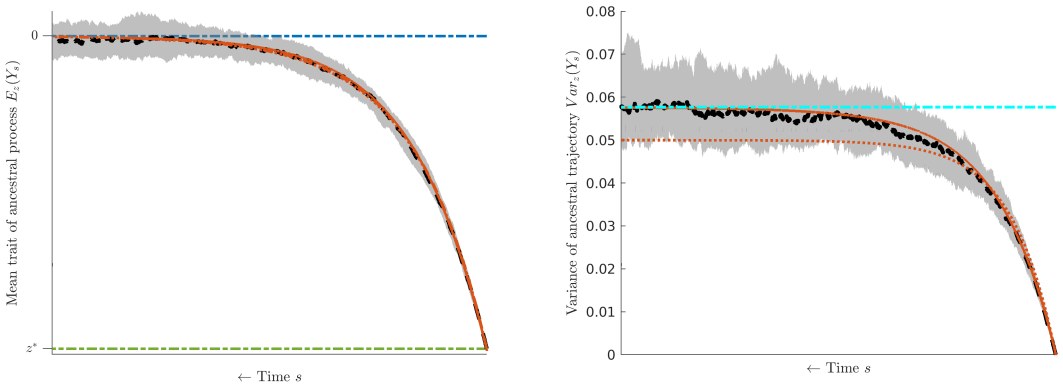
BioSP, INRAE, 84914, Avignon France
Email address: florian.patout@inrae.fr



(a) $\sigma = 0.05$ and $c = 0.025$



(b) $\sigma = 0.1$ and $c = 0.1$



(c) $\sigma = 0.1$ and $c = 0.05$

FIGURE 10. **Mean (left) and variance (right) of the ancestral process Y_s for different set of parameters** On the left panel, the horizontal blue line is the optimal trait in the moving frame, 0, and the green line is the dominant trait of the equilibrium F, denoted z^*. On the right panel, the horizontal cyan line corresponds to the asymptotic variance given by the deterministic model, see Proposition 1.5. In all cases, the black dashed line is the average trajectory, over 50 replicates of the IBM, of the mean (respectively the variance) of ancestors among lineages. All other parameters are taken as previously.