Ancestral Lineages in Mutation Selection Equilibria with Moving Optimum

Raphaël Forien 1 · Jimmy Garnier 2 · Florian Patout 1

Received: 17 November 2021 / Accepted: 29 June 2022 / Published online: 27 July 2022
© The Author(s), under exclusive licence to Society for Mathematical Biology 2022

Abstract
Many populations can somehow adapt to rapid environmental changes. To understand this fast evolution, we investigate the genealogy of individuals inside those populations. More precisely, we use a deterministic model to describe the phenotypic density of a population under selection when the fitness optimum moves at constant speed. We study the inside dynamics of this population using the neutral fractions approach. We then define a Markov process characterizing the distribution of ancestral phenotypic lineages inside the equilibrium. This construction yields qualitative as well as quantitative properties on the phenotype of typical ancestors. In particular, we show that in asexual populations typical ancestors of present individuals carried traits much closer to the fitness optimum than most individuals alive at the same time. We also investigate more deeply the asymptotic regime of small mutation effects. In this regime, we obtain an explicit formula for the typical ancestral lineage using the description of the solutions of the Hamilton Jacobi equation as a minimizer of an optimization problem. In addition, we compare our deterministic results on lineages with the lineages of stochastic models.

Keywords Integro-differential operator · Moving environment · Ancestral lineages
1 Introduction

1.1 Model Description

Numerous studies have reported rapid evolution in populations facing brutal environmental changes, such as climate change, habitat alteration or drug treatment (Bradshaw and Holzapfel 2006; Hairston et al. 2005; Hoffmann and Sgro 2011; Parmesan 2006). These populations are under sustained pressure to adapt to changing environments. For instance, pathogens like influenza continuously adapt to evade their host’s immune system, causing chronic infections despite heavy immune responses (Bedford et al. 2011). Such rapid evolution is associated with specific genealogies and leaves distinctive footprints in the genetic structure of populations. In particular, in asexual populations continuously adapting to changing environment, lineages trace back to a small pool of highly fit ancestors (Neher and Hallatschek 2013). Although many theoretical studies have focused on the statistical properties of genealogies using coalescent models (Berestycki 2012; Brunet et al. 2007a; Kingman 1982), little is known about the lineages dynamics (Calvez et al. 2021).

Over the past few decades, important theoretical progress has been made in predicting phenotypic evolution in a changing environment (reviewed in Kopp and Matuszewski 2014). Since the pioneer work of Burger and Lynch (1995), Lynch et al. (1991) and Lynch and Lande (1993), most theoretical approaches to adaptation in a changing environment rely on a quantitative phenotypic trait subject to stabilizing selection around some optimal phenotype, whose value is shifted continuously through time.

In this work, we investigate the genealogy of a population adapting to a changing environment. Our aim is to discover which traits contribute to the adaptation of future generations. We first focus on the dynamics of lineages using a deterministic model. Adapting the inside dynamics methods of Garnier et al. (2012) and Roques et al. (2012), we track the traits of the progeny of different neutral fractions of the total population adapting to the changing environment. Then, using duality between partial differential equations and stochastic processes, we look backwards in time to investigate the ancestors of the population. To this end, we define a stochastic ancestral process that describes the traits of the past ancestors of an individual sampled uniformly from those with a given trait in the present population. Our analysis goes back and forth between a forward deterministic description of the population adapting to a changing environment with the dynamics of neutral fractions inside, and a backwards description of ancestral lineages in the form of stochastic processes. We consider a population characterized by a phenotypic trait \( x \in \mathbb{R} \). Its trait density \( f(t, x) \) changes as a result of mutations and natural selection. Individuals in this population give birth at a constant rate \( \beta \). Their offspring inherit the parents’ trait with possibly a variation due to mutations described by the mutation kernel \( K_\sigma \). This yields the reproduction operator \( \mathcal{B} \):

\[
\mathcal{B}(f)(x) := (K_\sigma * f)(x) = \frac{1}{\sigma} \int_{\mathbb{R}} K \left( \frac{x - x'}{\sigma} \right) f(x') \, dx'
\]  (1)
This convolution term accounts for an asexual reproduction of individuals. We assume that $K$ satisfies:

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

so that the parameter $\sigma^2$ corresponds to the variance of the mutation kernel. We assume that selection acts through the intrinsic mortality rate $d(t, \cdot)$. Environmental change is modeled by assuming that this function is of the form:

$$d(t, x) = \mu(x - ct),$$

for some $c > 0$ which measures the speed of environmental change. We assume that $\mu$ is a convex function such that

$$\lim_{|x| \to +\infty} \mu(x) = +\infty. \quad (3)$$

Without loss of generality, we suppose that $\mu$ admits a (global) minimum at $x = 0$ (so that the optimal trait is always $ct$). Finally, we assume that the death rate due to competition is given by $\gamma \int_{\mathbb{R}} f(t, z)dz$, a non-local logistic competition term where $\gamma > 0$ is the competition factor.

Overall, the evolution of $f$ is described by the following integro-differential equation (IDE):

$$\partial_t f(t, x) = \beta B(f(t, \cdot))(x) - \left( \mu(x - ct) + \gamma \int_{\mathbb{R}} f(t, x')dx' \right) f(t, x),$$

for $t > 0, \ x \in \mathbb{R}. \quad (4)$

In order to describe a population keeping pace with the changing environment, we look for traveling pulse solutions of (4) which are special solutions of the form:

$$f(t, x) = F(x - ct), \quad (5)$$

where the constant profile $F$ describes a “mutation-selection equilibrium,” in the frame moving at speed $c$. It satisfies

$$\gamma \int_{\mathbb{R}} F(z')dz' F(z) - c \partial_z F(z) = \beta B(F)(z) - \mu(z) F(z), \quad z \in \mathbb{R}, \quad (6)$$

where $z$ represents the variable in the moving frame at speed $c$,

$$z := x - ct.$$

The phenotypic trait $z$ in the moving frame corresponds to the difference between the trait $x$ and the moving optimal trait $ct$. The existence of a traveling pulse solution of
(6) can be derived from the existence, stated in Coville and Hamel (2019), of a spectral pair \((\lambda, F)\) solving
\[
\lambda F(z) - c\partial_z F(z) = \beta B(F)(z) - \mu(z) F(z), \quad z \in \mathbb{R}.
\] (7)

In the following, we always assume that \(\lambda > 0\). In this case, since (7) is a linear equation, we can choose \(F > 0\) such that
\[
\lambda = \gamma \int_{\mathbb{R}} F(z) \, dz,
\] (8)
yielding a solution to (6). The profile \(F\) is bell shaped and centered around a value \(z^*\). This dominant trait \(z^*\) is negative due to the advection term in (6) and represents the lag of adaptation in a changing environment (see Fig. 1a).

### 1.2 Mathematical Results

In order to describe the ancestral lineages in the population, we first track the offspring of individuals. To do so, we investigate the inside dynamics of the traveling pulse \(F\) using the approach of neutral fractions (see Fig. 1 for a schematic representation). The idea is that individuals are labeled and transmit their label to their offspring. Since individuals only differ by their label and their trait, each label \(k \in \mathbb{N}\) corresponds to a neutral fraction of density \(\nu^k\) inside the population with density \(F\). Initially, we assume
\[
F(z) = \sum_{k \geq 1} \nu^k(0, z), \quad \text{for all } z \in \mathbb{R}.
\] (9)

Their dynamics in the moving frame is described by
\[
\begin{aligned}
\partial_t \nu^k(t, z) - c\partial_z \nu^k(t, z) &+ \left(\mu(z) + \gamma \int_{\mathbb{R}} F(z') \, dz'\right) \nu^k(t, z) \\
&= \beta B(\nu^k(t, \cdot))(z), \quad \text{for } t > 0, \ z \in \mathbb{R} \\
\nu^k(0, z) &= \nu^k_0(z), \ z \in \mathbb{R}.
\end{aligned}
\] (10)
In particular, mutation and selection act on fractions the same way they do on the entire population $F$ (solution of (6)). Note that by linearity, the sum $\sum_{k \geq 1} v_k(t, z + ct)$ verifies (4) with $f(0, x) = \sum_{k \geq 1} v_k(0, x)$ (fixed frame). Therefore, with the corresponding initial assumptions (9), we see that the neutral fractions are, at all times, a subdivision of the equilibrium profile:

$$\forall t \geq 0, \forall x \in \mathbb{R} : \quad F(x - ct) = \sum_{k \geq 1} v_k(t, x - ct).$$

This approach has been introduced in the context of reaction diffusion equations to understand the evolution of diversity inside traveling wave solutions (in their context, the front solution plays a similar role to our equilibrium profile $F$: (Garnier et al. 2012; Hallatschek and Nelson 2008, 2010; Roques et al. 2012).

Here, the dynamics of neutral fractions describes the evolution of the progeny of different subgroups of individuals in the population, depending on the initial distribution of the subgroups. As we shall see below, the distribution of the neutral fractions at any given point $x$ of the trait space can be deduced from the initial distribution of neutral fractions and the distribution of the ancestors of the individuals carrying the trait $x$.

Combining this forward approach with the characterization of stochastic processes, we are able to define a backward ancestral process $Y_s$, in the moving frame, for each individual with trait $z$ at time $t$. More precisely, the ancestral process at time $s$ aims to describe the trait of the ancestor (alive at time $t - s$) of an individual sampled at random from those carrying trait $z$ at time $t$. In this light, the backward process is related to the forward model of neutral fractions through the following relationship for all $(t, z) \in \mathbb{R}_+ \times \mathbb{R}$ and $k \in \mathbb{N}$:

$$\frac{v_k(t, z)}{F(z)} = \mathbb{E}_z \left[ \frac{v_k(s, Y_{t-s})}{F(Y_{t-s})} \right], \quad 0 \leq s \leq t.$$  

This equality means that the probability of sampling an individual of type $k$ from those with trait $z$ at time $t$ (left-hand side) is equal to the probability of drawing an individual of type $k$ among individuals of trait $y$ in the past, at the time $t - s$ (right-hand side). Rigorously, this relationship stems from the following result.

**Theorem 1.1 (Ancestral process)** Let $A$ denote the linear operator defined by

$$A \psi := \beta \left[ K_\sigma * (F \psi) - (K_\sigma * F) \psi \right] + c \partial_y \psi,$$

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

1. The operator $A$ generates a Feller semigroup, $(M_s, s \geq 0)$, acting on the set of real continuous and functions vanishing at infinity, $C_{b,0}(\mathbb{R})$. We call the Markov process associated with the semigroup $(M_s, s \geq 0)$, denoted by $(Y_s, s \geq 0)$, the ancestral process.
The semigroup \((M_s, s \geq 0)\) is such that, for all \((t, z) \in \mathbb{R}_+ \times \mathbb{R}\), and all \(k \in \mathbb{N}\):

\[
\nu^k(t, z) = F(z)M_{t-s}\left(\frac{\nu^k(s, \cdot)}{F}\right)(z), \quad \forall 0 \leq s \leq t.
\]

In other words, it is the moment dual of \(\nu^k/F\), in the sense of stochastic processes, see (12).

The generator of \((Y_s, s \geq 0)\) comprises two parts: a jump part and a drift part. The process \((Y_s, s \geq 0)\) jumps at rate \(\beta F(Y_s)K_\sigma\) to a location \(z\) whose density is given by

\[
K_\sigma(Y_s - z)F(z)
\]

Moreover, between jumps, \(Y_s\) drifts to the right at constant speed \(c\).

Note that the initial time of the ancestral process \(Y_s\) ("\(s = 0\)"), corresponds to any time \(t \in \mathbb{R}\). The trajectory of ancestors is independent of the time \(t\) at which we sample individuals in the population \(F\) because the population is at equilibrium in the moving frame.

Next, we establish some properties of lineages which correspond to trajectories of the ancestral process. We first start with the long-time behavior of the ancestral process as \(s \to +\infty\) which informs us on the trait distribution of the ancestors far back in the past.

**Proposition 1.2** (Common ancestors distribution) When \(s \to \infty\), the ancestral process \(Y_s\) converges in law towards a random variable \(Y_\infty\), which admits the following density:

\[
F\phi,
\]

where \(\phi\) is the non-negative stationary solution of the dual problem:

\[
(\gamma \int_{\mathbb{R}} F) \phi + c \partial_z \phi = \beta (K_\sigma \ast \phi - \phi) + (\beta - \mu) \phi, \quad \text{with} \quad \int_{\mathbb{R}} F(y')\phi(y')dy' = 1.
\]

Proposition 1.2 hinges on an intermediary result stated in Cloez and Gabriel (2020). The description of the density of \(Y_\infty\) relies on the long-time behavior analysis of the neutral fractions. It also provides interesting properties on the traveling pulse \(F\). Indeed, we prove that any fraction inside the traveling pulse converges to a positive proportion of the traveling pulse as time goes to \(\infty\). Thus, the traveling pulse is pushed in the sense that any fraction inside this pulse pushes it forward (Garnier et al. 2012; Roques et al. 2012).

However, the common ancestor distribution is not uniform, some traits are more represented among ancestors than in the overall population. In particular, we show that ancestors are more likely to have a trait close to the optimal trait \(z = 0\). More precisely, we show that \(Y_\infty\) satisfies these additional properties:
Corollary 1.3 Under the additional assumption that selection is symmetric: $\mu(x) = \mu(|x|)$, the distribution of $Y_\infty$ admits the following density:

$$z \mapsto \frac{F(z)F(-z)}{\int F(y')F(-y')dy'}.$$

In particular,

$$\mathbb{E}[\psi(-Y_\infty)] = \mathbb{E}[\psi(Y_\infty)], \text{ for any } \psi \in \mathcal{C}_b(\mathbb{R}) \text{ and } \mathbb{E}[Y_\infty] = 0.$$

Corollary 1.3 states that on average the ancestors of the present population share the optimal trait with respect to selection. However, the population density at the optimal trait $F(0)$ is low because the dominant trait of the population lags behind the optimal trait at a distance $z^*$ (see Fig. 1a). Thus, we learn an interesting feature of the adaptation phenomenon of our model: the ancestors of typical individuals in the present population were far from typical within the population at the time. The existence of (potentially very few) optimally fit individuals is thus crucial to the survival of the whole population since they will be the (most likely) ancestors of the next generations.

To further understand the trajectories of the ancestral process $Y_s$, we focus on the asymptotic regime of small mutations, that is $\sigma \to 0$ with a time of order $s/\sigma$. Since the time is scaled by $1/\sigma$, we also need to rescale the speed of change $c$:

$$c = \sigma c'.$$  \hspace{1cm} (16)

We thus let $F_\sigma$ denote the solution to (6) with $c$ replaced by $\sigma c'$, and $(\mathcal{M}_{\sigma,s}, s \geq 0)$ the associated lineage semigroup. Under this asymptotic regime the equilibrium profile $F_\sigma$ concentrates around the dominant trait $z^*$, determined by selection. More precisely, using the Hop–Cole transform $U_\sigma = \sigma \log(F_\sigma)$, we can see that $U_\sigma$ converges toward an asymptotic profile $U$, Lorz et al. (2011). This asymptotic profile satisfies an Hamilton–Jacobi equation of the form:

$$\lambda + c'\partial_z U(z) + \mu(z) = \beta + \beta H(\partial_z U(z)),$$

with $H(p) := \int_{\mathbb{R}} K(y) \exp(yp)dy - 1$. \hspace{1cm} (17)

The Hamiltonian $H$ is the trace of the mutation kernel at this scale.

Under this asymptotic regime and with the help of the theory of representation of solution of Hamilton–Jacobi equations, we are able to fully characterize the ancestral process $Y_s$.

Proposition 1.4 (Small mutations regime) Let $\mathcal{M}_s^\sigma := \mathcal{M}_{\sigma,s/\sigma}$. Then, for all times $s \in \mathbb{R}_+$, $z \in \mathbb{R}$ and $\psi \in \mathcal{C}_b(\mathbb{R})$,

$$\mathcal{M}_s^\sigma \psi(z) \xrightarrow{\sigma \to 0} \psi(\Gamma_z(s)), \text{ locally uniformly in time},$$ \hspace{1cm} (18)
and, for any $\varepsilon > 0$ and $T > 0$,

$$\lim_{\sigma \to 0} \mathbb{P}_\sigma \left( \sup_{s \in [0,T]} |Y_{s/\sigma} - \Gamma_z(s)| > \varepsilon \left| Y_0 = z \right. \right) = 0,$$

(19)

where $\mathbb{P}_\sigma$ denotes the law of $(Y_s, s \geq 0)$. The limit process $\Gamma_z$ corresponds to the solution of the following ODE:

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

(20)

where the function $U$ and the Hamiltonian $H$ satisfy (17).

The limit $-\sigma \log(F_\sigma) \to U$ holds locally uniformly, and the existence of $U$ results from Lorz et al. (2011). In particular, the authors show that $U$ is a $C^1(\mathbb{R})$ function, solution of the Hamilton Jacobi Eq. (17) (see Appendix A.1 for a detailed derivation of this equation). Under this particular regime, the ancestral process becomes a deterministic process defined by $\Gamma_z$. This situation was expected because the cumulative variance in the ancestral process due to mutations is proportional to $\sigma^2(t/\sigma) = \sigma t$ in the rescaled time scale, and thus vanishes for any fixed $t$.

This asymptotic regime has been widely studied in evolutionary contexts since the pioneer work of Diekmann et al. (2005). This regime provides a good approximation when either the effects of mutations are small or when selection is strong, and describes how in this regime a population concentrates around one or several traits. This approach is connected to large deviations theory (Champagnat et al. 2019), and it generally involves Hamilton-Jacobi equations (Barles et al. 2009; Lorz et al. 2011). Here, we show that the asymptotic equations keep some trace of the history of ancestors, and thus provide insights on the trajectories of the ancestral process that are the typical lineages. To the best of our knowledge, this aspect is new. In Sect. 5.1, we detail how the Hamilton–Jacobi equation provides heuristics about typical lineages.

1.3 Related Works

Our results, stated for the non-local birth operator $\mathcal{B}$ defined by (1), can be extended to the case where mutations are modeled by a Laplace operator ($\mathcal{B} = \Delta$). Some specific additional results are provided in Sect. 4. In particular, the trajectories of ancestral lineages follow an explicit SDE, see Proposition 4.1.

The same diffusion process was obtained in Calvez et al. (2021) as the limit, as the number of individuals tends to infinity, of the trajectory of the traits along an individual’s ancestral lineage in an individual-based stochastic population model. More precisely, the authors considered a stochastic model describing the evolution of a population of individuals following the rules described in Sect. 1.1, with a finite number of individuals $N$ tending to infinity. They showed, using the historical process associated with their population model and a branching process approximation along with techniques initiated by Marguet (2019), that the lineage of an individual sampled uniformly from those alive at some time $T$ converges in distribution to the trajectory of a
solution to the SDE (38), while at the same time the renormalized population process converges to a measure with density given by $F$, solution of (34). Their result can be seen as a microscopic justification of our genealogical interpretation of the duality relation (12), in the case of a diffusive mutation operator.

Mathematically, a large number of models describe the adaptation of populations in a steady environment, (see, e.g., Diekmann et al. 2005; Gil et al. 2019). In the case of a linearly varying environment, there exist reaction–diffusion models where a favorable region moves at a certain speed, Berestycki et al. (2009) and Berestycki and Fang (2018), but it does not describe an adaptation phenomenon contrary to Alfaro et al. (2017). Periodically fluctuating environments have also been studied, see Figueroa Iglesias and Mirrahimi (2018, 2019) and Lorenzi et al. (2015). Recently, Roques et al. (2020) proposed a methodology to deal with general changing environments (linear, oscillating or stochastic) in the case of quadratic selection and a diffusive mutation operator.

Our work has a close kinship with Bouin et al. (2020), where (mostly) formal analytical features measuring the dynamics of adaptation are obtained, for integro-differential models close to (4). The methodology is based upon asymptotic expansions in the same regime as Proposition 1.4. It also encapsulates the case where the operator $B$ describes sexual reproduction via the infinitesimal operator, see Calvez et al. (2019) and Patout (2020) for rigorous asymptotic treatment (without environmental change) of this operator.

1.4 Ancestral Lineages: An IDE Point of View

The fundamental solution associated with $A$, the generator of the ancestral process $Y_s$, solves the following linear IDE:

$$
\begin{align*}
-\partial_s w^z(s, y) &= \beta \int_{\mathbb{R}} K(h) \frac{F(y + \sigma h)}{F(y)} \left( w^z(s, y + \sigma h) - w^z(s, y) \right) dh \\
+c \partial_y w^z(s, y), s \geq 0 \\
w^z(t, y) &= \delta(z - y).
\end{align*}
$$

(21)

This can be interpreted as an IDE describing the dynamics of the trait $y$ of the ancestors whose descendants have reached the trait $z$ in the population, while $t$ is the arbitrary time of the equilibrium where the population is sampled (and the past "begins").

Formally, the discrete fractions label $k$ in (10) is replaced by a continuum of neutral alleles. Each neutral fraction then corresponds to the progeny of a single ancestral trait in the population.

Indeed, let us define for any $y \in \mathbb{R}$, $\nu^y$ as follows:

$$
\nu^y(t - s, z) = F(y)w^z(s, y) \quad \text{for any } 0 \leq s \leq t, \text{ and } z \in \mathbb{R},
$$

where $w^z$ is the fundamental solution in (21). Let $L$ be the linear operator defined by

$$
L(\psi) = \beta \left( K_{\sigma} \ast \psi - \psi \right) + c \partial_z \psi - \left( \mu - \mu \right) \psi,
$$

(22)
where (integrating (6) to obtain the last equality)

\[ \bar{\mu} = \beta - \lambda = \int_{\mathbb{R}} \mu(z) \frac{F(z)}{\int_{\mathbb{R}} F(z') \, dz'} \, dz. \]

A simple computation, detailed in Sect. 2, guarantees that \( \nu^y \) solves:

\[
\begin{aligned}
\partial_t \nu^y(t, z) &= \mathcal{L}(\nu^y(t, \cdot))(z), \\
\nu^y(0, z) &= \delta(z - y) F(z).
\end{aligned}
\]

(23)

This corresponds to Eq. (10) satisfied by the neutral fractions:

\[ \forall k \in \mathbb{N}, \quad \partial_s \nu^k = \mathcal{L}(\nu^k). \]

(24)

As a consequence, \( \nu^y \) can be seen as a neutral fraction, with a neutral label that singles out one trait in the population at \( t = 0 \). Therefore, the (stochastic, backwards) ancestral process defined in Theorem 1.1 corresponds to a continuum of (deterministic, forward) neutral fractions. Numerical simulations highlight this in Fig. 10, with more details provided in Appendix B.4.

We finally propose an interpretation in terms of partial differential equations of Proposition 1.4. The fundamental solution \( w^z_\sigma \), introduced in (21), converges in the sense of distributions, when \( \sigma \to 0 \), up to the acceleration of time, towards \( w^0_0 \) the solution of:

\[
\begin{aligned}
\partial_s w^z_0(s, y) &= -\left( c' - \beta \partial_p H(\partial_y U(y)) \right) \partial_y w^z_0(s, y), \\
w^z_0(0, y) &= \delta(z - y).
\end{aligned}
\]

(25)

The integral flow of the transport Eq. (25) coincides with the ODE (20) solved by \( \Gamma \). In Sect. 5.1, we recover this formula (established in (53)), directly from an Hamilton–Jacobi equation (rigorously derived in Lorz et al. 2011). It has somehow its origin in the field of the Weak-KAM theory, independently of neutral fractions. Therefore, thanks to \( \Gamma \), our work links the Hamilton–Jacobi framework and the ancestral process, in the regime \( \sigma \to 0 \).

The rest of this article is organized as follows. In Sect. 2, we prove Theorem 1.1, using classical semigroup theory. Then, Sect. 3 is devoted to the proof of Proposition 1.2. In Sect. 4, we consider the case of a diffusive mutation operator, for which explicit computations are possible. Finally, we prove the Proposition 1.4 in Sect. 5. We also provide a discussion on how to obtain the ODE from heuristics on the Hamilton–Jacobi equation and a comparison of theoretical results with stochastic and deterministic simulations. In the appendices, we provide additional discussions on the link between lineages and the methodology introduced in Bouin et al. (2020) under the regime of small mutations. Finally, we detail the numerical methods to keep track of lineages with individual-based simulations and the comparison with the deterministic model.
2 Link Between Lineages and Fractions: Proof of Theorem 1.1

(1) From the definition of the operator \( \mathcal{A} \) in (13), we obtain the following equivalent expression (see also (21)):

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

(26)

Since \( F \) is a bounded function, \( K \) is a probability kernel, and \( 1/F \) is locally bounded, the first term in (26) is the generator of a pure jump Markov process with locally bounded jump rate. The second term corresponds to an additional transport at speed \( c \), so the existence of a Feller semigroup acting on \( C_{b,0} \) whose generator is given by \( \mathcal{A} \) is straightforward (see, for instance, Section 4.2 of Ethier and Kurtz (2009)).

(2) To conclude, we need to prove the relationship (14). Note that \( \mathcal{A} \) and \( L \) satisfy

\[
\mathcal{A}\psi = \frac{L(F\psi)}{F}, \quad \text{for all } \psi \in C^1_b(\mathbb{R}).
\]

(27)

To see this, use (22) to write

\[
\frac{1}{F} L(F\psi) = \frac{\beta}{F} \left( K_\sigma * (F\psi) - F\psi \right) + \frac{1}{F} c \partial_z (F\psi) - \frac{1}{F} \left( \mu - \bar{\mu} \right) \psi F.
\]

(28)

Multiplying each side of the equality \( LF = 0 \) by \( \psi/F \), one obtains:

\[
\frac{\psi}{F} c \partial_z (F) - \frac{1}{F} \left( \mu - \bar{\mu} \right) \psi F - \frac{1}{F} \beta \psi F = -\frac{\psi}{F} \beta K_\sigma * F.
\]

Plugging this into (28),

\[
\frac{1}{F} L(F\psi) = \frac{\beta}{F} \left( K_\sigma * (F\psi) - \psi K_\sigma * F \right) + c \partial_z \psi = \mathcal{A}\psi.
\]

(29)

This proves (27).

Then, for all \((t, z) \in \mathbb{R}_+ \times \mathbb{R}, 0 \leq s \leq t \) and \( k \in \mathbb{N} \), we compute:

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

(29)

We plug the identity (27) for the first term, and (24) for the second. This yields

\[
\frac{d}{ds} \mathcal{M}_{t-s} \left( \frac{u^k(s, \cdot)}{F} \right)(z) = -\mathcal{M}_{t-s} \left[ \frac{L(u^k(s, \cdot))}{F} \right](z) + \mathcal{M}_{t-s} \left[ \frac{L(u^k(s, \cdot))}{F} \right](z).
\]
Since $M_0 = Id$, by evaluating in $s = t$ we have shown that for all $k \in \mathbb{N}$ and $(t, z) \in \mathbb{R}_+ \times \mathbb{R}$,

$$\frac{u^k(t, z)}{F(z)} = M_{t-s} \left( \frac{u^k(s, \cdot)}{F} \right)(z), \quad 0 \leq s \leq t. \quad (30)$$

The identity (14) is now established.

### 3 Long Time Behavior of Lineages

We now study of the long time behavior of the ancestral process $Y_s$. We can observe from the duality relationship (12) that the time of ancestors is “backwards” compared to the time $t$ of the equilibrium $F$ (and the fractions, per (11)). The regime $s \to +\infty$ corresponds to the study of the most ancient ancestor.

#### 3.1 Long Time Behavior of Neutral Fractions

**Proof of Proposition 1.2** Let $\psi \in C^1_b(\mathbb{R})$, then $\nu_0 := F\psi$ is also a function of $C^1_b(\mathbb{R})$. We then consider a neutral fraction $\nu$ with initial condition $\nu_0$. With Eqs. (10) and (11), this means

$$\begin{align*}
\begin{cases}
\partial_t \nu(t, z) = \mathcal{L}(\nu(t, \cdot))(z) & \text{for } t > 0, \quad z \in \mathbb{R} \\
\nu(0, z) = \nu_0(z), \quad z \in \mathbb{R},
\end{cases}
\end{align*} \quad \text{(31)}$$

and $\mathcal{L}$ is defined by (22). The property (12) applied at $s = t$ yields

$$\frac{\nu(t, z)}{F(z)} = \mathbb{E}_z \left[ \frac{\nu_0(Y_t)}{F(Y_t)} \right] = \mathbb{E}_z [\psi(Y_t)]. \quad (32)$$

From Cloez and Gabriel (2020, Theorem 2.1) where the long time behavior of solutions of (31) is studied, we obtain

$$v(t, \cdot) \xrightarrow{\mathcal{L}_{loc}^\infty}{\mathcal{P}[\nu_0]} F, \quad \text{with } \mathcal{P}[\nu_0] \colon = \frac{\int_{\mathbb{R}} \nu_0(z) \varphi(z) dz}{\int_{\mathbb{R}} F(z') \varphi(z') dz'}, \quad (33)$$

where $\varphi$ is defined as the solution of (15). As a result, we have

$$\mathcal{P}[\nu_0] = \frac{\int_{\mathbb{R}} F(z) \psi(z) \varphi(z) dz}{\int_{\mathbb{R}} F(z') \varphi(z') dz'} : = \mathbb{E}_z [\psi(Y_{\infty})].$$
Therefore, with (32), we have shown that for any $\psi \in C^1_b(\mathbb{R})$,

$$\mathbb{E}_z[\psi(Y_t)] \underset{t \to \infty}{\longrightarrow} \mathbb{E}_z[\psi(Y_\infty)].$$

□

This proof hinges on the recent results of Cloez and Gabriel (2020), and we believe it can be deduced as well from the general semigroup analysis of growth fragmentation equations presented in Mischler and Scher (2016). In any case, the linearity of the operator $L$ is crucial to the argument. Indeed, Eq. (33) states the convergence toward the projection on the dual eigenspace generated by $\varphi$ solution of (12) as well as $L^* (\varphi) = 0$.

Moreover, the asymptotic proportion of the fractions coincides with the heuristics proposed in Sect. 1.4. Indeed, choosing $\psi_0 (z) = \delta(z - y)F(z)$ in (31) (as in (23)), and applying our convergence theorem (33), we recover ’$p[y]$’, which corresponds exactly to the asymptotic density stated in Proposition 1.2.

**Proof of Corollary 1.3** The key is to find out that there exists an explicit link between the solution of the dual problem $\varphi$ and the original one (7). Thanks to the symmetry of $\mu$, the function $z \mapsto F(-z)$ solves the dual problem (15). Therefore, the function $F\varphi$ is even, and $Y_\infty$ admits an even density given in Proposition 1.2. Corollary 1.3 immediately follows. □

### 3.2 Numerical Simulations: Dynamics of the Ancestral Process

In this section, we present some numerical simulations illustrating the previous results. Figure 1 shows the initial and ‘final’ distribution of fractions solving (10). Here, each fraction corresponds to a given interval of traits among the population, and each color corresponds to a neutral fraction. Thus, the cumulative density of the neutral fractions stays equal to the entire population represented by the stationary density profile $F$ moving at constant speed (per (11)). The dominant fraction inside the population after a long time is the ‘blue’ fraction, which corresponds to the few individuals optimally adapted initially. Thus, we see that the blue fraction contributes the most to the adaptation of the population to the changing environment.

We now turn to the numerical simulations of the ancestral process, introduced in Theorem 1.1. In Fig. 2a, we represent the dynamics of the distribution of the ancestral process $Y_s$ starting at $s = 0$ from a single trait $z = z^*$ corresponding to the dominant trait in the population, that is $\max(F) = F(z^*)$. We can observe that the ancestral process distribution starts from a Dirac mass at $z = z^*$ when $s = 0$ (top), and then, it flattens over $z$ and its mean gradually shifts towards the optimal trait $z = 0$ as $s$ increases. Eventually, it reaches a stationary distribution which corresponds to the explicit expression of Proposition 1.2 (see blue curve in Fig. 2b). This asymptotic density represents the proportion of ancestors of phenotype $y$ in the population, asymptotically as $s \to \infty$. We recover that the asymptotic density is even with a maximum at $z = 0$, the optimal trait. Although the density at the optimal trait is very low (see dashed curve in Fig. 2b), most common ancestors, which are ancestors when $s \to \infty$, have a trait close to this optimal trait (see blue curve in Fig. 2b).
Fig. 2 Ancestral Process distribution over time in panel a and asymptotic ancestral distribution $Y_\infty$ (blue curve) compared with the population density $F$ (dotted black curve) in panel b. In panel a the black dotted line corresponds to the dominant trait $z^\ast$ inside the population and the red dotted curve represents the optimal trait $z = 0$ (Color figure online)

4 The Diffusive Approximation

When the mutational variance $\sigma^2$ is small, the convolution operator $B$ defined in (1) can be approximated by:

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

In the diffusive approximation regime, the profile $F_d$ satisfies the following equation:

$$0 = \beta \frac{\sigma^2}{2} \partial_z^2 F_d(z) + c \partial_z F_d(z) + \left( \mu(z) - \gamma \int_{\mathbb{R}} F_d(z')dz' \right) F_d(z). \quad (34)$$

Therefore, the neutral fractions operator $L$ defined by (22) becomes:

$$L_d(\nu) = \beta \frac{\sigma^2}{2} \partial_z^2 \nu + c \partial_z \nu - \left( \mu - \bar{\mu} \right) \nu \quad \text{with} \quad \bar{\mu} = \gamma \int_{\mathbb{R}} F_d(z)dz.$$

$$= \int_{\mathbb{R}} \mu(z) \frac{F_d(z)}{\int_{\mathbb{R}} F_d(z')dz'}dz. \quad (35)$$

With this new model, we can define the ancestral lineage under the diffusive approximation and we can state the following properties.

Proposition 4.1 (Ancestral process under the diffusive approximation)
The diffusive ancestral process \( (Y^d_s, s \geq 0) \) associated with (35) admits the following generator:

\[
A_d \psi = \beta \frac{\sigma^2}{2} \partial_y^2 \psi + \left( \beta \frac{\sigma^2}{2} \frac{\partial_y F_d}{F_d} + c \right) \partial_y \psi.
\]  

(b) **Diffusive common ancestor distribution.** When \( s \to \infty \), the limit process \( Y^d_\infty \) admits a density given by

\[
\frac{\left( F_d(y)e^{cy/2} \right)^2}{\int_{\mathbb{R}} \left( F_d(y')e^{cy'/2} \right)^2 dy'}.
\]

(c) In addition, the above density admits a local maximum at \( y = 0 \) for \( \sigma \) small enough.

In this case, Eq. (36) implies that the ancestral process \( (Y^d_s, s \geq 0) \) solves the following SDE:

\[
dY_s = \left( \beta \frac{\sigma^2}{2} \frac{\partial_y F_d(Y_s)}{F_d(Y_s)} + c \right) ds + \sqrt{\beta \sigma^2} dB_s,
\]

where \( (B_s, s \geq 0) \) is standard Brownian motion. A similar statement is made in Calvez et al. (2021, Theorem 1.1) in the case of quadratic selection (\( F_d \) is then explicit, see below).

**Proof of Proposition 4.1**

(a) Formula (36) is obtained by plugging expression (35) in (27), yielding

\[
A_d \psi = \frac{1}{F_d} \mathcal{L}_d(F_d \psi) = \beta \frac{\sigma^2}{2} \left( \frac{1}{F_d} \partial_y^2 (F_d \psi) - \frac{1}{F_d} \psi \partial_y^2 F_d \right) + c \partial_y \psi.
\]

Then, after simplifications using (34), we recover (36).

(b) To establish (37), we introduce the dual problem, similar to (15):

\[
\mathcal{L}_d^*(\varphi) = \beta \frac{\sigma^2}{2} \partial_y^2 \varphi - c \partial_y \varphi - \left( \mu - \overline{\mu} \right) \varphi.
\]

In order to prove the convergence of \( Y_s \), we must use similar result to (33), with \( \mathcal{L}_d \). This is a classical statement, see, for instance, (Garnier et al. 2012) and references therein (the convergence now holds in \( L^2 \)). Moreover, the density in (37) is slightly more explicit than in Proposition 1.2, because there exists a special relationship between the primal and dual problem. Indeed, computations show that if \( \mathcal{L}_d^*(\varphi_d) = 0 \), then

\[
\varphi_d(y) = F_d(y)e^{2cy/(\beta \sigma^2)}.
\]
This explains (37), using the formula of Proposition 1.2. (c) The function defined for each \( z \) by \( \tilde{F}(y) := F_d e^{cy/(\beta\sigma^2)} \) is even, since \( \mu \) is an even function by hypothesis and \( \tilde{F} \) is the solution of:

\[
\beta \frac{\sigma^2}{2} \frac{\partial^2}{\partial z^2} \tilde{F} - \left( \mu - \bar{\mu} + \frac{c^2}{2\beta\sigma^2} \right) \tilde{F} = 0. \tag{39}
\]

Therefore, \( y \mapsto \left( F_d(y) e^{cy/(\beta\sigma^2)} \right)^2 \) is also an even function. As a consequence, \( Y_\infty \) is symmetric, by (37). In particular, \( \tilde{F} \) admits a (local) extrema at \( z = 0 \). To obtain more information on the local shape of the density, we must investigate the sign of \( \partial^2_y \tilde{F}(0) \), since \( \tilde{F} \) is the numerator in (37). First, we find that \( \mu(0) - \bar{\mu} < 0 \).

Back to (39), if \( \sigma \) is sufficiently small, we get \( \partial^2_y \tilde{F}(0) > 0 \), and therefore, \( Y_\infty \) admits a local maximum at 0, as claimed in Proposition 4.1.

\[\square\]

**Quadratic Selection**

Here, we tackle the special case where \( \mu(z) = z^2/2 \), which verifies our general hypotheses in (3). In this case, the profile \( F_g \) is known to be Gaussian, which allows us to illustrate our previous qualitative comments (Burger and Lynch 1995; Kopp and Matuszewski 2014). More precisely, the profile \( F_g \) solves

\[
\lambda_g F_g(z) - c \partial_z F_g(z) + \frac{z^2}{2} F_g(z) = \beta F_g(z) + \frac{\beta \sigma^2}{2} \partial^2_z F_g(z), \text{ for } z \in \mathbb{R}, \tag{40}
\]

whose solution is given by

\[
F_g(z) = \frac{\lambda_g}{\sqrt{2\pi \sigma \sqrt{\beta}}} \exp \left( -\frac{1}{2\sigma \sqrt{\beta}} \left( z + \frac{c}{\sigma \sqrt{\beta}} \right)^2 \right), \quad \lambda_g = \beta - \frac{c^2}{2\beta\sigma^2} - \frac{\sigma \sqrt{\beta}}{2}. \tag{41}
\]

Up to a constant, \( F_g \) is indeed a Gaussian distribution centered around an optimum proportional to \( c \), lagging behind the optimal trait, and with a variance proportional to \( \sigma \) (instead of \( \sigma^2 \) for \( K_\sigma \)). In the eigenvalue \( \lambda_g \), we recognize the lag load: \( c^2/2\beta\sigma^2 \), that is the cost to keep pace with the changing environment, and the mutation load: \( \sigma \sqrt{\beta}/2 \) (Burger and Lynch 1995). In addition, the speed of change \( c \) must be small enough for the population to persist (\( \lambda > 0 \) if and only if \( c < \sigma \beta \sqrt{2} \sqrt{1 - \sigma/(2\sqrt{\beta})} \)) (Bürger 2000). In particular, \( c \) must be of the order of \( \sigma \), just as in the regime of Proposition 1.4.
Moreover, from (37), we get the following explicit formula for the asymptotic ancestral distribution $Y^\infty_g$:

$$
\frac{(F_g \varphi_g)(y)}{\int_{\mathbb{R}} (F_g \varphi_g)(y') \, dy'} = \frac{1}{\sqrt{\pi \sigma \sqrt{\beta}}} \exp\left(-\frac{y^2}{\sigma \sqrt{\beta}}\right).
$$

This is a Gaussian distribution centered at $y = 0$ and with variance $\sigma \sqrt{\beta} / 2$. The variance of the asymptotic ancestral distribution is half that of the population and coincides with the mutation load.

In addition, under the diffusive approximation, we can characterize the entire ancestral process $(Y_s, s \geq 0)$ by combining the expression of $F_g$ stated in (41) and the SDE equation solved by $Y_s$. The ancestral process is an Ornstein–Uhlenbeck process. This characterization can also be found in Calvez et al. (2021, Theorem 1.1) (with $\beta = 1$, (41) corresponds to Calvez et al. 2021, Proposition 2.3). Therefore, for all $s \geq 0$, $Y_s$ follows a Gaussian distribution with mean and variance given by

$$
\forall s \geq 0, \quad \mathbb{E}_z(Y_s) = ze^{-\sigma \sqrt{\beta} s} \quad \text{and} \quad \text{Var}_z(Y_s) = \sigma \sqrt{\beta} \left(1 - e^{-2 \sigma \sqrt{\beta} s}\right).
$$

The variance does not depend on the reference point $z$, while the mean of the ancestral distribution converges to 0 exponentially fast, at a rate $\sigma \sqrt{\beta}$. Building on this, we conjecture that for general symmetric selection function $\mu$, the mean $\mathbb{E}_z(Y_s)$ should converge exponentially fast to 0 at a rate $\sigma^2 \beta / \text{Var}(F_d)$, where $\text{Var}(F_d)$ corresponds to the variance of the profile $F_d$ defined by (34), we refer to Appendix A.4 for further details.

### 5 Small Mutations Regime

In this section, we tackle the asymptotic regime where mutations have very small effects (the trait of a mutant offspring is very close to that of its parent). First, we explain how we can deduce the heuristic formula (20) for the typical lineage from the following Hamilton–Jacobi equation obtained as the limit when $\sigma \to 0$ of the equation satisfied by $F$:

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

where $U = \lim_{\sigma \to 0} \sigma \ln(F)$ and $H$ is the Hamiltonian defined in Proposition 1.4. The rigorous convergence when $\sigma \to 0$, in the sense of viscosity solutions, is established in Barles et al. (2009) and Lorz et al. (2011). In Appendix A.1, we explain how to derive (formally) the Hamilton–Jacobi equation from (7) when $\sigma \to 0$. 

\[\text{Springer}\]
5.1 Hamilton–Jacobi Equation and the Typical Lineage

Before proving Proposition 1.4, we start with some heuristic arguments linking the Hamilton–Jacobi Eq. (43) and the typical lineage using a dual point of view. Let us introduce the Lagrangian function $L$ associated with $K$, corresponding to the Legendre transform of the Hamiltonian $H$:

$$L(v) := \max_{p \in \mathbb{R}} \left( pv - H(p) \right).$$  \hspace{1cm} (44)

Using this function, we can write the solutions of (43) with this variational representation:

$$U(z) = \inf_{\gamma \text{ s.t. } \gamma(0)=z} \int_0^{+\infty} \left[ \beta L \left( \frac{-\dot{\gamma}(s) + c'}{\beta} \right) - \beta + \mu(\gamma(s)) + \lambda \right] ds. \hspace{1cm} (45)$$

We refer to Barles and Roquejoffre (2006) for the origin of this formula, which stems from the Weak-KAM theory. The infimum is taken over all functions $\gamma \in C^1(\mathbb{R}_+)$ that reach the phenotype $z$ at time 0. These functions $\gamma$ are phenotypic paths, among them any optimal trajectory minimizes a cost. The expression of this functional shows the combined cost of mutations through $L$ (at speed $\dot{\gamma} + c'$) and selection through $\mu$. The birth rate $\beta$ plays a role opposite to natural selection, while $\lambda$ is the term that balances the expression, just as in (43).

From a genealogical point of view, the phenotypic path can be seen as an ancestral lineage. As a result, the special lineage that minimizes the cost should correspond to a ‘typical’ ancestral lineage of the individuals of trait $z$ at time 0.

For the time being, let $\Gamma$ be such a minimizing trajectory. Then, as a byproduct of the Weak-KAM theory, one can show that the optimal trajectory $\Gamma$ of the variational problem (45) is the solution of the following ordinary differential equation:

$$\dot{\Gamma}(s) = c' - \beta \partial_p H \left( \partial_z U(\Gamma(s)) \right), \quad s > 0$$

$$\Gamma(0) = z. \hspace{1cm} (46)$$

This result comes from the Hamiltonian/Lagrangian structure of (45) and more precisely from the study of the characteristics of this Hamilton Jacobi equation (see, for instance, Hairer et al. 2006). This characterization of the optimal trajectory will help us gain qualitative insights on the typical lineage and will improve our numerical computation of $\Gamma$ (Hairer et al. 2006).

The ODE (46) coincides with the limit Eq. (20) of the ancestral process in Proposition 1.4. This is the key point of this part, and as a result, we can state qualitative properties of the ancestral process studying $\Gamma$. If $\Gamma$ satisfies (46), then we have

$$U(z) = \int_0^{+\infty} \left[ \beta L \left( \frac{-\dot{\Gamma}(s) + c'}{\beta} \right) - \beta + \mu(\Gamma(s)) + \lambda \right] ds. \hspace{1cm} (47)$$
Ancestral Lineages in Mutation Selection Equilibria

Formally, one should expect \( \dot{\Gamma}(s) / \Gamma_1(s) \) to converge to 0 when \( s \to +\infty \). Otherwise, \( \Gamma \) and thus the mortality rate \( \mu(\Gamma) \) would become arbitrarily large (see (3)). Thus, the trajectory \( \Gamma \) would not be a minimizer, as \( U(z) \) could take the value \(+\infty\). As a result, \( \Gamma \) converges when \( s \to +\infty \) and its limit is necessary 0 to minimize the selection function \( \mu \):

\[
\lim_{s \to +\infty} \Gamma(s) = 0.
\]

The typical lineage traces back to an ancestor with the optimal trait 0 as we already observed in Corollary 1.3; however, here it is in a much stronger sense. We prove the limit (48) in Appendix A.3, using convexity methods and qualitative properties extracted from the Hamilton–Jacobi equation beyond lineages (see Appendix A.2).

5.2 Proof of Proposition 1.4

We aim to prove the convergence of the semigroup \( \mathcal{M}^\sigma \) as \( \sigma \to 0 \). From classical results on semigroups, we only need to prove the convergence of its generator \( A^\sigma \) (Kallenberg 2006, Theorem 17.25). We prove the convergence of the generator on \( C_2^b(\mathbb{R}_+ \times \mathbb{R}) \) which is actually a core for the generator \( A^\sigma \) (Bansaye et al. 2019; Cloez and Gabriel 2020).

Now, let \( \psi \in C_2^b(\mathbb{R}_+ \times \mathbb{R}) \) and \( A^\sigma \) be the generator corresponding to the semigroup \( \mathcal{M}^\sigma \). Then, we have

\[
A^\sigma \psi(y) = \frac{\beta}{\sigma} \int_{\mathbb{R}} K(h) \frac{F(y + \sigma h)}{F(y)} \left( \psi(y + \sigma h) - \psi(y) \right) dh + \frac{c}{\sigma} \partial_y \psi(y), \quad y \in \mathbb{R}.
\]

Let \( U_\sigma \) be defined by

\[
U_\sigma(z) = -\sigma \log(F_\sigma(z)) \quad \text{for all} \quad z \in \mathbb{R}.
\]

Then, the generator \( A^\sigma \) is written as

\[
A^\sigma \psi(y) = \frac{\beta}{\sigma} \int_{\mathbb{R}} K(h) \exp \left( -\frac{U_\sigma(y + \sigma h) - U_\sigma(y)}{\sigma} \right) \left( \psi(y + \sigma h) - \psi(y) \right) dh + \frac{c}{\sigma} \partial_y \psi(y).
\]

Using a Taylor expansion, for all \( y \in \mathbb{R} \), there exists \( \tilde{y} \) such that \( |y - \tilde{y}| \leq \sigma h \) and

\[
\psi(y + \sigma h) - \psi(y) = \sigma h \partial_y \psi(y) + \frac{\sigma^2 h^2}{2} \partial^2_y \psi(\tilde{y}).
\]

Plugging this expression into the definition of the generator \( A^\sigma \), we end up with

\[
A^\sigma \psi(y) = \beta \int_{\mathbb{R}} K(h) \exp \left( -\frac{U_\sigma(y + \sigma h) - U_\sigma(y)}{\sigma} \right) h \partial_y \psi(y) dh.
\]
\[ + \sigma \beta \int_{\mathbb{R}} K(h) \exp \left( - \frac{U_\sigma(y + \sigma h) - U_\sigma(y)}{\sigma} \right) \frac{h^2}{2} \partial_y^2 \psi(y) \, dh + \frac{c}{\sigma} \partial_y \psi(y) \]

(50)

Thanks to Barles et al. (2009) and Lorz et al. (2011), we know that \( U_\sigma \) converges locally uniformly towards \( U \) as in (17). Moreover, \( U \) solves the problem (43), in the sense of viscosity solutions, and we have the following Lipschitz uniform bound:

\[ \| U'_\sigma \|_\infty < \eta, \quad (51) \]

for all \( \sigma > 0 \) small enough, where \( \eta \) is such that

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

As a result, we show that the first integral term of (50) converges and the second integral vanishes, pointwise, for any \( \psi \in C^2_b(\mathbb{R}_+ \times \mathbb{R}) \). Using the rescaled speed \( c' = c/\sigma \), we obtain

\[ \mathcal{A}^\sigma \psi(y) \xrightarrow{\sigma \to 0} -\beta \int_{\mathbb{R}} K(h) \exp \left( h U'(y) \right) h \partial_y \psi(y) \, dh + c' \partial_y \psi(y) := \mathcal{A}^0 \psi(y). \]

(52)

Then, using Theorem 17.25 in Kallenberg (2006), we conclude that the semigroup \( \mathcal{M}^\sigma_s \) associated to \( \mathcal{A}^\sigma \) converges to a semigroup \( \mathcal{M}^0_s \) associated with the asymptotic operator \( \mathcal{A}^0 \).

To conclude the proof of Proposition 1.4, we need to characterize \( \mathcal{M}^0 \). Let \( \mathcal{V} \) be defined by

\[ \mathcal{V}(z) = c' - \beta \partial_p H(U'(z)), \quad \text{for all } z \in \mathbb{R}, \]

and \( \Gamma \) be its corresponding integral flow:

\[ \begin{cases} \partial_s \Gamma(t, s, z) = \mathcal{V}(\Gamma(t, s, z)), \\ \Gamma(t, t, z) = z. \end{cases} \]

(53)

The expression for the derivative of the Hamiltonian as defined in (17) is:

\[ \partial_p H(p) = \int_{\mathbb{R}} K(y) \exp(y p) p \, dy. \]

Now for any test function \( \psi \), we define \( \theta(s, z) = \mathcal{M}^0_s \psi(z) \). Using the expression of \( \mathcal{A}^0 \), we see that \( \theta \) solves the following equation:

\[ \begin{cases} \partial_s \theta(s, z) = \mathcal{V}(z) \partial_z \theta(s, z), & s > 0, z \in \mathbb{R}, \\ \theta(0, z) = \psi(z). \end{cases} \]
Classically, this advection equation with non-constant velocity field admits an “explicit” solution, by following the characteristics, which correspond in our case to the flow $\Gamma$ defined by (53):

$$\theta(s, z) = \psi(\Gamma(0, s, z)).$$  \hspace{1cm} (54)

From the definition of $\Gamma_z$ stated in (20) and the expression (53) of $\Gamma$, we observe that $\Gamma(0, s, z) = \Gamma_z(s)$. From the expression of $\theta$ in (54), we prove the convergence of (18) in Proposition 1.4.

We now turn to the proof of (19). We claim that the family of Markov processes $(Y_\sigma^s, s \geq 0)$ indexed by the parameter $\sigma \in [0, 1]$ is tight for the Skorokhod topology. To prove this, note that

$$Y_\sigma^s = Y_0^s + V_\sigma^s + M_\sigma^s,$$

where

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

and $(M_\sigma^s, s \geq 0)$ is a local martingale with predictable variation

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

Using (51), we then see that

$$|V_\sigma^s - V_\sigma^{s'}| \leq \left( c' + \beta \int \mathbb{R} K(h) \exp(\|U_\sigma'\|_\infty |h|) |h| \, dh \right) |s' - s|,$$

$$|\langle M_\sigma^s \rangle_s| \leq \sigma s \beta \int \mathbb{R} K(h) \exp(\|U_\sigma'\|_\infty |h|) h^2 \, dh.$$

Since $\|U_\sigma'\|_\infty < \eta$ as in (2), there exists a constant $C > 0$, independent of $\sigma$, such that

$$\int \mathbb{R} K(h) \exp(\|U_\sigma'\|_\infty |h|) |h| \wedge |h|^2 \, dh \leq C,$$

where $a \wedge b = \min(a, b)$. As a result, for any $\varepsilon > 0$,

$$\lim_{\delta \to 0} \limsup_{\sigma \to 0} P\left( w_\delta(Y^\sigma) > \varepsilon \right) = 0,$$

where $w_\delta(Y)$ denotes the modulus of continuity of $Y$, i.e.,

$$w_\delta(Y) = \sup_{s, t \in [0, T] \atop |t-s| \leq \delta} |Y_s - Y_t|.$$
Moreover, for each fixed \( s \geq 0 \), \( Y_s^{\sigma} \) converges in distribution to the deterministic limit \( \Gamma_z(s) \), this convergence holds also for finite-dimensional marginals of \( (Y_s^{\sigma}, s \geq 0) \). This shows that the family of processes \( (Y_s^{\sigma}, s \geq 0) \) is C-tight (Billingsley 2013). This yields the convergence in distribution of \( (Y_s^{\sigma}, s \geq 0) \) in the uniform topology to the deterministic process \( (\Gamma_z(s), s \geq 0) \). Finally, since the limit is deterministic, the convergence holds in probability. Hence, (19) is proven which concludes the proof of Proposition 1.4.

6 Numerical Simulations

The aim of this section is to compare our ancestral process defined from the deterministic model with the ancestral lineages of classical Individual-Based Model (IBM) taken from Champagnat et al. (2006, 2007). Moreover, we assess the accuracy of our approximation formula (diffusive approximation and Hamilton–Jacobi approximation) with respect to the mean of the ancestral process \( Y_s \) and the mean of the lineages of the IBM model.

6.1 The Stochastic Model

We consider a stochastic IBM model where each individual is characterized by its trait. They reproduce and die at rates that may depend on their traits (see Appendix B.1 for more details) and on the total population size. A logistic competition term keeps the population size finite, and when the strength of competition tends to zero, the population size tends to infinity and the (renormalized) stochastic model converges to the deterministic model (4) per (Champagnat et al. 2006) (see Fig. 6 in Appendix B.1). Therefore, for large populations, we may expect the lineages of the IBM model and the trajectories of the ancestral process \( Y_s \) associated with the deterministic model (4) to behave similarly. In order to compare the genealogy of the individuals in the stochastic model and the ancestral process, each individual carries a label which encodes its genealogy, that is the trait of its parents, together with its current trait (see Appendix B.1 for more details). With this definition of the genealogy, we compare the dynamics of lineages obtained through the IBM model with the ancestral process distribution for the general model (4), and the lineage trajectory \( \Gamma_z \) obtained in the asymptotic regime of small mutations (\( \sigma \to 0 \)).

From a numerical point of view, we look at the lineages of individuals with the dominant trait of the population \( z = z^* \) in order to sample initially as many individuals as possible for the IBM model. The population size of the stochastic model is around 20,000, while the number of individuals carrying the dominant trait is around 1000 (see Fig. 3).

6.2 Ancestral Process and Stochastic Lineages

First of all, we have verified that the dynamics of the lineages in both models coincide (see Fig. 10 and movie Fig. 11 in Appendix B.4). Thus, our ancestral process captures
Ancestral Lineages in Mutation Selection Equilibria… Page 23 of 43

93

Fig. 3 Dynamics of the ancestral lineages starting from the dominant trait \( z^* \) for the IBM model and the deterministic PDE model. The \( y \)-axis represents the phenotype in the moving frame. The blue background represents the distribution of the ancestral process \( Y_s \) starting from \( z^* \) at \( s = 0 \), and the gray curve represents the typical lineage \( \Gamma_{z^*} \) defined by (20). The colored lines starting from \( z = z^* \) correspond to the lineages of the IBM model with \( N = 2 \times 10^4 \) individuals. The black dashed lines represent the mean of the lineages. The blue dash-dotted line represents the optimal trait \( z = 0 \), and the green dash-dotted line the dominant trait \( z^* \). For the simulation \( \sigma = 0.1 \) and \( c = 0.2 \) (Color figure online)

The distribution of the ancestral lineages induced by the stochastic model. In particular, the mean of the ancestral process \( \mathbb{E}_z(Y_s) \) and the mean of the stochastic ancestral lineages are close for each replicate (see Fig. 10).

We compare the stochastic model with the Hamilton Jacobi approximation obtained in the limit of small mutations (\( \sigma \to 0 \)). To solve (20) numerically, we coupled this equation with (43) to obtain a system of ODE. Although \( \sigma \) is not 0 in the stochastic model, we observe in Fig. 3 that the trajectory \( \Gamma_{z^*} \) follows the trajectory of the mean of the stochastic ancestral lineages. We further show in the appendix that the Hamilton Jacobi approximation as well as the diffusive approximation provide a good estimate of the mean of the ancestral process for both deterministic and stochastic model (see Figs. 10 and 12).

However, going far away in the past in Fig. 3, the mean of the stochastic ancestral lineages fluctuates a lot due to the small number of ancestors at this time. In that time regime, it makes more sense to compare averages over a large number of IBM replicates, which we do in Fig. 4, see also Fig. 6. As already mentioned, the deterministic model is relevant when the size of the population is large enough (see Fig. 10). However, per Fig. 4, our ancestral process fully captures the average behavior of the mean over many replicates of the IBM model even if the size of the population is small.

7 Discussion

In the context of adaptation to a changing environment, we proposed a method to track lineages using a deterministic mathematical model of mutation and selection.
More precisely, we are able to define the ancestral process \( (Y_s, s \geq 0) \) describing the trajectory of the ancestral traits of an individual sampled uniformly from those with a given trait in the present population. Our results show that every trait is represented among the ancestors, but that ancestral traits are strongly biased toward the fitness optimum (which is shifting linearly). Furthermore, combining the asymptotic regime of small mutation (\( \sigma \to 0 \)) with the Lagrangian structure of the solution to the Hamilton–Jacobi equation, we provide a good approximation of the ancestral lineages as the solution to an ODE, see (20) in Proposition 1.4.

Some modeling choices made when writing Eq. (4) may seem somewhat arbitrary and limiting the scope of the present study. The mathematical analysis performed here can nonetheless be made fairly general. In fact, the same analysis can be carried out for any density profile \( f(t, \cdot) \) satisfying an equation of the form:

\[
\frac{\partial}{\partial t} f(t, x) = \mathcal{B} f(t, \cdot)(x) + r(f(t, \cdot)) f(t, x),
\]

where \( \mathcal{B}^* \) generates a Feller semigroup and which admits a non-negative stationary solution. In particular, we could include a dependence of the birth rate on the trait, more general competition terms, etc. In fact, we already make fairly weak assumptions on the shape of the function \( \mu \), contrary to, for instance, Fisher’s Geometric Model (FGM) which assumes an explicit quadratic relationship between phenotype and fitness (see Gil et al. 2019; Martin and Roques 2016 for related works on the FGM model). In addition, our framework deals with a general form of mutations. Although it can lead to the “diffusive approximation” accounting for small mutations (Kimura 1964) (see Sect. 4), the non-local operator \( \mathcal{B} \) defined by the mutation kernel \( K \) can describe general distributions of mutation effects. We show that the ancestral process truly depends on this mutation kernel (see (21)). In particular, we show that as the effect of mutations vanishes (\( \sigma \to 0 \)), the lineages retain a trace of the whole mutation kernel.
through the Hamiltonian in (20) (while the diffusive approximation only keeps track of the variance of the mutation kernel through the parameter $\sigma^2$).

### Beyond our Model

We show that the fittest individuals in the genealogy drive the adaptation in an asexual population facing changing environment, a feature already observed in other evolutionary context (Hermisson et al. 2002; Neher and Hallatschek 2013; Rouzine and Coffin 2007). Beyond this qualitative information, we quantify this phenomenon in specific contexts (diffusive approximation and small mutation regime). In particular, in the regime of small mutations, we show that the structure of Hamilton–Jacobi equations carries some information on the genealogies resulting from adaptation to this changing environment. This opens a broad range of applications for the method presented in this work, since this Hamilton–Jacobi approach has been used extensively in different ecological models, and, in particular, when mutations may have a large effect, for instance, with a kernel $K$ that does not satisfy the assumption (2) of being exponentially bounded. Recently, Bouin et al. (2018) and Mirrahimi (2020) found the (non-stationary) Hamilton–Jacobi equation when $\sigma = 0$, in the case of a broad range of fat tailed kernels. However, in that case, the dual Lagrangian point of view of the equation is no longer valid, since the Hamiltonian $H$, in (17), is no longer well defined. However, our investigation of genealogies, based on neutral fractions, can still apply, and would provide a description of ancestral lineages in this context. From a mathematical point of view, we mainly use the linearity of $B$ and spectral results about the linearized operator $L$ around the equilibrium $F$. These properties hold true in many models, see Mischler and Scher (2016).

In our model, we assume that the environmental change equally affects all the individuals in the population. However, habitats may differ between locations. Recently, many works have focused on the interaction of two populations living in two different habitats, e.g., Hamel et al. (2020) and Mirrahimi and Gandon (2020). Each habitat favors a different optimal trait. Individuals can move between habitats and are subject to natural selection. In this scenario, migration tends to shift the trait distribution in each habitat toward the optimal trait of the other habitat. In particular, polymorphism can appear in a habitat provided migration is strong enough. Investigating ancestral lineages in this context could yield valuable insights on this phenomenon. Recently, Garnier and Lafontaine (2020) have extended the notion of neutral fractions to metapopulation model. Thus, using our notion of ancestral process based on neutral fractions can be an efficient tool to investigate the genealogy of a metapopulation located in different areas.

An other important issue in spatial ecology is the evolution of populations undergoing range expansions. Over the last decades, several theoretical works have focused on models of adaptation with continuous space and trait variable, such as the “cane toad equation” to tackle these issues (Bénichou et al. 2012; Bouin et al. 2017). In expanding population, the fittest individuals are at the front of the propagation range, a phenomenon called ‘spatial sorting’. Up to our knowledge, the mechanisms underlying this phenomenon are poorly understood. Recently, Calvez et al. (2018) showed that
a non-local competition term can slow down the acceleration of the front, conversely to previous formal intuitive results of Bouin et al. (2012). Thus, the investigation of the genealogy of individuals at the leading edge of the front, using our methodology, might provide new insights on this issue.

In a related work, Etheridge and Penington (2020) studied a spatial Moran process modeling an expanding population with a strong Allee effect (corresponding to a bistable reaction–diffusion equation). Using neutral markers, they showed that the genealogy of individuals sampled near the position of the traveling front is asymptotically close to the classical Kingman coalescent. This result is in stark contrast with classical results on pulled waves (as in the stochastic Fisher-KPP equation) where the genealogy of individuals sampled at the leading edge of the front is believed to follow the Bolthausen–Sznitman coalescent, in which multiple lineages can merge at the same time (Berestycki et al. 2013; Brunet et al. 2007b; Desai et al. 2013; Neher and Hallatschek 2013). As it happens, in order to obtain their result, Etheridge and Penington (2020) use the fact that ancestral lineages of individuals sampled near the front approach a diffusion process which admits a stationary distribution in the moving frame centered on the front position, a fact that is reminiscent of what we obtain here.

In the present paper, we focus on asexual population, while many species reproduce sexually. When this is the case, each individual has two parents and the reproduction can be described using the infinitesimal model (Barton et al. 2017; Fisher 1918; Turelli 2017; Turelli and Barton 1994) which takes the following form in our framework:

$$B(f)(x) := \frac{1}{\sigma \sqrt{\pi}} \int_{\mathbb{R}^2} \exp \left[ -\frac{1}{\sigma^2} \left( x - \frac{x_1 + x_2}{2} \right)^2 \right] f(x_1) \frac{f(x_2)}{\int_{\mathbb{R}} f(x_2') \, dx_2'} \, dx_1 \, dx_2.$$  (55)

This operator states that individuals of trait $x_1$ and $x_2$ give birth to an individual whose trait $x$ is drawn from a Gaussian distribution centered at the mean of the traits of its parents, $(x_1 + x_2)/2$, and with variance $\sigma^2$. Thus, in this context, the genealogy of individuals becomes a binary tree which tracks the whole pedigree of each individual. The analysis of ancestral lineages thus becomes more intricate, since the number of genealogical ancestors can quickly reach the size of the whole population. However, recent asymptotic studies (Calvez et al. 2019; Patout 2020) have laid the groundwork for the adaptation of the neutral fractions framework to this setting.

**Stochastic Framework, Coalescence**

As already mentioned, the deterministic models (4) and (34) can be obtained as large population limits of stochastic individual-based models described in Appendix B.1 (Champagnat et al. 2006, 2007). In this stochastic context, the neutral fractions approach has also been used to understand how lineages coalesce back in time. For instance, recently (Billiard et al. 2015) studied neutral markers in the background of a trait-substitution sequence in the adaptive dynamics regime, see also Etheridge and Penington (2020) for genealogies in bistable traveling wave.
Although our deterministic model cannot track the microscopic aspect of the lineages because the coalescence events become increasingly rare as the size of the population tends to infinity, we can still provide some heuristics on the timescale of coalescence in the genealogy of individuals. For instance, Fig. 3 suggests that lineage are less likely to coalesce before they reach a trait close to the optimal trait. Moreover, in our setting, the growth rate of individuals is maximal near the moving optimum and becomes negative far ahead of the optimum, so that lineages do not escape in the tip of the front (as they do in pulled fronts). This means that the traveling pulse of (34) is “pushed” and should behave similarly to bistable waves, as studied in Etheridge and Penington (2020). As a result, most coalescence events should take place near the optimum, and no single individual is likely to quickly produce a large progeny (of a size comparable to the total population). We can thus conjecture that as in Etheridge and Penington (2020), the genealogy of a finite sample of individuals follows Kingman’s coalescent.

In addition, using some of our results, we can provide heuristics about the mean coalescence time of two lineages $T_2$. First, our analytic approximations (42) and (73) on the dynamics of the mean of the ancestral lineage provide a good approximation on the characteristic time $T_0$ before which two lineages reach the optimal value 0. From our diffusive approximation, we get

$$T_0 \approx \frac{1}{\sigma \sqrt{\beta}}$$

In the case of more general operators, based on the approximation formula (71) in the small variance regime, we expect

$$T_0 \approx \frac{1}{\text{Var}(F)},$$

where $\text{Var}(F)$ is the variance of the phenotypic distribution at equilibrium. After this time delay $T_0$, the coalescent time $T_2$ seems to follow an exponential distribution (see Fig. 8). However, the parameter of the exponential time depends on the speed of the changing environment. If $c$ is small ($0 < c \leq \sigma$), our simulations suggest an exponential rate $T_0$, while if $c$ is large ($c > \sigma$), the exponential rate seems to depend on the distribution of the ancestral process and $N$ size scale of the population (see Fig. 8). When the changing speed $c$ is small, the amount of individuals at the optimal trait $ct$ is large compared to the size of the population. Thus, using the heuristic of Etheridge and Penington (2020) and Neher and Hallatschek (2013) with our Proposition 1.2, we expect that the exponential rate does not truly depend on the population size $N$ and it should be $T_0$. However, when the speed increases, the population size at the optimal trait is low compared to the size of the population; thus, the exponential rate should depend on it. Using the heuristic of Etheridge and Penington (2020), we suggest that the exponential rate should be

$$\frac{1}{2\sqrt{N}} \int_{\mathbb{R}} \frac{1}{F(z)} \frac{(F(z)\varphi(z))^2}{\left(\int_{\mathbb{R}} F(z)\varphi(z)dz\right)^2}dz.$$
These arguments are all in the preliminary stage of a larger work and need to be further investigated.

Acknowledgements The authors acknowledge Vincent Calvez for introducing them to the problem, helping to formulate it and for his support during this work. We also warmly thank Jérôme Coville and Lionel Roques for their helpful comments. 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” and ANR-16-CE02-0009 “GLOBNETS”). R. F. was supported in part by the Chaire Modélisation Mathématiques et Biodiversité (École Polytechnique, Muséum national d’Histoire naturelle, Fondation de l’École Polytechnique, VEOLIA Environnement).

Appendices

A. The Regime of Small Variance: Further Developments

A.1. Derivation of the Hamilton-Jacobi Equation

In this section, we explain formally how to derive a Hamilton–Jacobi equation from the pulse characterization \((7)\), in the regime of small mutations.

From the work of Bouin et al. (2020), we know that the regime of small mutation effects can be described by the following scaling factor:

\[
\varepsilon := \sqrt{\frac{\sigma^2 \alpha}{\beta}}, \quad \text{with} \quad \alpha := \mu''(0) > 0.
\]  

When \(\varepsilon \to 0\) our Eq. \((7)\) converges to a Hamilton–Jacobi equation which allows us to use the rigorous results of Lorz et al. (2011). This limiting regime captures the weak selection regime, in the sense that either the variance vanishes, or the selection is weak compared to birth:

\[
\sigma \ll 1 \quad \text{or} \quad \frac{\alpha}{\beta} \ll 1.
\]

We show here how the pulse \(F\) concentrates around a mean trait value, \(z^*\), when \(\sigma \to 0\).

Recall that \(F\) is a solution of the integro-differential equation \((7)\), which we rewrite here:

\[
\lambda\sigma F_\sigma(z) - cF'_\sigma(z) + \mu(z)F_\sigma(z) = \frac{\beta}{\sigma} \int_{\mathbb{R}} K\left(\frac{z - z'}{\sigma}\right) F_\sigma(z') \, dz'.
\]  

When the variance of the mutation kernel vanishes, that is \(\sigma \to 0\), we may expect the function \(F\) to concentrate around a specific trait, guided by selection, see Barles et al. (2009) and Lorz et al. (2011) for instance. To capture this phenomenon, we use
the following Hopf–Cole transform, identical to (17):

\[ F_{\sigma}(z) = \exp \left( -\frac{U_{\sigma}(z)}{\sigma} \right). \]  

(58)

The scaled quantity \( U_{\sigma} \) then satisfies the following equation:

\[ \lambda_{\sigma} + \frac{c}{\sigma} U'_{\sigma}(z) + \mu(z) = \frac{\beta}{\sigma} \int_{\mathbb{R}} K \left( \frac{z - z'}{\sigma} \right) \exp \left( -\frac{U_{\sigma}(z) - U_{\sigma}(z')}{\sigma} \right) dz'. \]  

(59)

To avoid degeneration of terms in the equations when \( \sigma \to 0 \),
we rescale the speed of environmental change by \( \sigma \), as in (16):

\[ c := \sigma c'. \]  

(60)

In order to find the limit equation, we use the following Taylor expansion of the exponential term inside the integral:

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

Plugging this approximation into the integral term of (59), with an affine change of variable, we obtain formally when \( \sigma \to 0 \):

\[ \lambda_0 + c' U'_0(z) + \mu(z) = \beta \int_{\mathbb{R}} K(y) \exp \left( y U'_0(z) \right) dy. \]

In the following, we omit the index 0, as in (17). We then recover Eq. (43). Let us observe that the limit equation is well defined thanks to assumption (2) on the exponential decay of \( K \). It also guarantees the finiteness of the Hamiltonian \( H \) in (17) for all \( p \in \mathbb{R} \).

**A.2. Lagrangian Point of View: Qualitative Formulas**

The Hamilton–Jacobi Eq. (43) provides analytical formula as well as qualitative behavior on \( U \) and \( \lambda \), which are expected to be a good approximation when \( \sigma \) is small. Let \( \mu_0 := \mu(0) \) be the minimum mortality rate.

We first start with the following formula on \( \lambda \):

\[ \lambda = \beta - \mu_0 - \beta L \left( \frac{c'}{\beta} \right). \]  

(61)

where \( L \) is the Lagrangian associated with the Hamiltonian \( H \) and related to the mutation kernel \( K \), see (44). The quantity \( \beta - \mu_0 \) corresponds to the intrinsic growth rate (fitness) of the population, while the additional quantity \( \beta L(c'/\beta) \) measures the lag load induced by the changing environment.
A short argument for (61) consists in assuming that the asymptotic behavior of $\Gamma$ stated in (48) holds true. Then, plugging this into (47), it prescribes the value of $\lambda$ such that $U$ does not take infinite values:

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

Since $\mu(0) = \mu_0$, this formula coincides with (61).

The proof of this analytical formula relies on convex analysis methods (see Bouin et al. 2020). First, the function $p \mapsto c' p - \beta H(p)$ admits a maximum value: $\beta L(c'/\beta)$. Indeed, the functions $H$ and $L$ have reciprocal derivatives functions: $\partial_p H \circ \partial_v L = Id$. Adding this maximum value on each side of the Hamilton–Jacobi Eq. (43), we obtain

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

(62)

We claim that on the left-hand side of this relationship, the term between brackets must vanish. Otherwise, it would lead to a contradiction as follows.

Indeed, the function $C : p \mapsto \beta H(p) - c' p + \beta L(c'/\beta)$ is convex, nonnegative and reaches zero from the properties of the Hamiltonian $H$ and the Lagrangian $L$. If the term between brackets does not vanish in (62), the function $z \mapsto C \circ U'$ takes only (strictly) positive values. As a consequence, $U'$ only takes values in one of the two branches of the convex function $C$, and on each of these branches, $C$ is invertible. Therefore, for each $z \in \mathbb{R}$, we can invert the relationship (62) to deduce the value taken by $U'(z)$. From (3), $\mu$ has the same infinite value at $\pm \infty$. Inverting (62) for $z = \pm \infty$ yields

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

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. Therefore, the term between brackets in (62) vanishes, exactly as in the desired formula (61).

As a side, $z = z^*$ leads to a formula that dictates the position of the dominant trait $z^*$:

$$\lambda + \mu(z^*) = \beta.$$  

Combined with (61), we find

$$\mu(z^*) = \mu_0 + \beta L\left(\frac{c'}{\beta}\right).$$

(63)

As a matter of fact, those formulas are consistent with those in Bouin et al. (2020),

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

Springer
They further show how to obtain more accurate expansions, up to an arbitrary order (in $\varepsilon$ defined in (56)), and compute explicitly the following corrector terms.

In addition, we know from (8) that $\lambda$ is a measure of the size of the population. Thus, this size should remain positive which gives us a critical threshold for the speed $c'$ so that the population does not go extinct:

$$c_{max} = \sigma \beta L^{-1} \left( \frac{\beta - \mu_0}{\beta} \right).$$

Moreover, we can check by integrating (7) that

$$\lambda = \int_{\mathbb{R}} \frac{\mu(x) F(x)}{\int_{\mathbb{R}} F(x') dx'} dx$$

which corresponds to the mean fitness of the population, or its mean intrinsic rate of increase.

### A.3. Long Time Behavior of $\Gamma_z$

In this part, we show a somehow stronger version of Proposition 1.2: the typical ancestral lineages $\Gamma_z$ converges toward 0, the optimal trait, when $s$ goes to infinity. In the regime of small variance, this result concerns the ODE (20).

Let $z_0$ be a steady state of (20). Then,

$$0 = -c' + \beta \partial_p H(U'(z_0)). \quad (64)$$

Let $p_0 := U'(z_0)$. Then, from (64), $p_0$ is a critical point of the function $p \mapsto -c' p + \beta H(p)$. Since this is a convex function, we deduce that

$$p_0 \in \arg\min_p \left( -c' p + \beta H(p) \right).$$

Therefore, by definition of the Lagrangian function $L$ in (44)

$$-\beta L \left( \frac{c'}{\beta} \right) = -c' p_0 + \beta H(p_0). \quad (65)$$

On the other hand, by evaluating the Hamilton Jacobi Eq. (43) in $z = z_0$, one finds

$$\mu(z_0) + \lambda = \beta + \beta H(p_0) - c' p_0.$$

Plugging in (65)

$$\mu(z_0) = \beta - \lambda + \beta L \left( \frac{c'}{\beta} \right). \quad (66)$$
Finally, thanks to the formula (61) for \( \lambda \), we get

\[
\mu(z_0) = \mu_0. \tag{67}
\]

According to our assumptions on \( \mu \) stated in (3), 0 is the only global point of minimum of the convex function \( \mu \), therefore \( z_0 = 0 \). We conclude that 0 is the unique steady state of the ODE (20). Moreover, it is established in Lorz et al. (2011) that under the assumptions (2), \( U \) is a convex function. Therefore, the flow in (20) is an increasing function, and it is straightforward that \( \Gamma_z \) converges to this unique steady state of (20):

\[
\Gamma_z(s) \xrightarrow{s \to \infty} 0.
\]

**A.4. Approximation of the Mean of the Ancestral Lineages**

Working further on the Hamilton Jacobi Eq. (43), we can make an analytic approximation of \( \Gamma_z \), the typical lineage. Let us first differentiate (43) with respect to \( z \), and then divide by \( U'' \) on each side. We obtain for all \( z \in \mathbb{R} \),

\[
-c' + \beta \partial_p H(U'(z)) = \frac{\mu'(z)}{U''(z)}. \tag{68}
\]

Now, our idea is to link \( U''(z) \) with the variance of \( F \). Using results from Bouin et al. (2020), we obtain the following approximation of the variance of \( F \) at the leading order in \( \sigma \):

\[
\text{Var}(F) = \frac{\sigma}{U''(z^*)} + o(\sigma), \tag{69}
\]

where \( z^* \) is the dominant trait in our population. This approximation comes from a Taylor expansion (with Laplace’s method) of the integrals defining the variance:

\[
\text{Var}(F) = \frac{\int_{\mathbb{R}} (z - z^*)^2 \exp \left( -\frac{U(z)}{\sigma} \right) \, dz}{\int_{\mathbb{R}} \exp \left( -\frac{U(z)}{\sigma} \right) \, dz}.
\]

In addition, we make the following rough approximation, valid if \( z \) is close to \( z^* \):

\[
U''(z) \approx U''(z^*).\]

Plugging these approximations into (68), we find that

\[
-c' + \beta \partial_p H(U'(z)) \approx \frac{\mu'(z)\text{Var}(F)}{\sigma} \quad \text{for } z \text{ close to } z^*.
\]

The ODE (20) satisfied by \( \Gamma_z \) then becomes:

\[
\dot{\Gamma}_z(s) = -\frac{\text{Var}(F)}{\sigma} \mu'(\Gamma_z(s)),
\]

\[
\Gamma_z(0) = z. \tag{70}
\]
In particular, if the selection function is quadratic, $\mu(z) = \mu_0 + z^2/2$, the solution of (70) is precisely

$$\Gamma_z(s) = z \exp\left(-\frac{\text{Var}(F)}{\sigma} s\right).$$

(71)

Moreover, on this case, we can derive an explicit formula for the variance from the HJ equation.

More precisely, we compute $U'''(z^*)$ from (43) which provides the following formula from (69):

$$\text{Var}(F) = -\frac{c'\sigma}{\mu'(z^*)}.$$  

Since $\mu$ is quadratic and $z^*$ satisfies (63), we obtain

$$\text{Var}(F) = \frac{c'\sigma}{\sqrt{2\beta} \sqrt{L\left(c'/\beta\right)}}.$$  

(72)

Finally, we get the following approximation for the mean trajectories of the ancestral process:

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

(73)

This formula is explicit, since it only depends on the mutation kernel through the Lagrangian $L$. However, this approximation only applies to the case of a quadratic selection and for traits close to the dominant trait. But we can check from numerical simulations, that this approximation is quite robust (see Fig. 12).

B. Numerical Methods

B.1. Ancestral Lineages in the Stochastic Model

We detail in this section the numerical simulations how we deal with the simulations of the individual based model and the lineages. The 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 $1 \leq i \leq N(t)$. It has the trait $z_i \in \mathbb{R}$. The first event for this individual is one of the following:

- **Birth of a descendant**: it happens at an uniform rate among individuals $r^i_B = \beta$.
- **Death of the individual $i$**: it is decomposed in two separate events:
(a) **Death by selection** The individual may die because its phenotype is ill-adapted in the phenotype landscape. This happens at the rate:

\[ r_{Ds}^i = \mu (z_i - ct). \]

(b) **Death by competition** Alternatively, an individual may die because of the density dependence in the population, at a rate that depends on the total size of population at time \( t \), and on 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:

\[ dt \sim \min_{1 \leq i \leq N(t)} E(\beta (r_B^i + r_{Ds}^i + r_{Ddd})). \] (B.1)

By the memory loss property of the exponential distribution, \( dt \) also can be drawn from an exponential law which rate is the sum of the rates of all the independent events:

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

**Update of the population:** Once the next event is decided, according to the law (B.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 to the operator \( B \) in (1):

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

We repeat all the steps until reaching the final time of simulation. Numerically, this model has a very high computational cost, because it needs a relatively high number of individuals to approach the deterministic model given by (4). As a consequence, we performed the simulations using an approximating model, by first fixing \( dt \) to a small but deterministic value. Then, for each individual, we draw a time of birth following the law \( E(\beta) \) and a time of death following the law \( E(\mu(z_i) + N(t)/N) \). 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 time interval, individuals cannot reproduce more than once.

Finally, to follow the lineage of individuals, we create a huge matrix at the start of the simulation. We will stock the lineage of every individual in this matrix, filling it progressively. Every time an individual appears, its lineage is similar to the one of
Fig. 5  Keeping track of the lineages: a schematic representation of the ancestral lineages algorithm described in Appendix B.1 (Color figure online)

Fig. 6  Population density profile $F$ in the moving frame obtained from deterministic model (4) (plain green curve) and from the IBM model with a number of individuals of order $2 \cdot 10^4$ (blue histogram). The selection function $\mu$ is quadratic with minimum at $z = 0$ (red dashed curve). The value $z^*$ corresponds to the mean phenotypic trait of the deterministic density (Color figure online)

its parent, translated by one generation. The numerical procedure works as described in Fig. 5, where different columns correspond to different individuals, and each line corresponds to a new generation.

This algorithm led to Figs. 6 and 4 with the following parameters:

$$\alpha = 2, \quad \beta = 2, \quad \sigma = 0.1, \quad N = 20000, \quad c = 0.2.$$
Fig. 7  Mean (left) and variance (right) of the ancestral process \( Y_s \) for different set of parameters: red curves correspond to the model (12) and the black dashed curves correspond to the IBM model averaged over 50 replicates and the gray regions correspond to the 5% and 95% confident interval. On the left panel, the horizontal blue lines are the optimal trait in the moving frame, 0. The green lines are 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 1.2 (Color figure online)

B.2. Mean Traits Along the Lineages

We extend the simulations of Fig. 4 to different scenarios corresponding to various mutational variances \( \sigma \) and speeds \( c \). More precisely, for two different couples of mutational variance and speed, we compare the mean and the variance of the stochastic lineages with the mean and the variance of the ancestral process defined by our PDE model. We show that our deterministic model provides agreed with the individual based model. In addition, we see that the variance of the ancestral process increases with the mutational variance \( \sigma \), while it slightly increases with the speed \( c \) (Fig. 7).

B.3. Coalescent Time

We first investigate the time \( T_2 \) before two individuals lineages meet in the past. More precisely, for each individual in the population, we compute the minimal time such
Fig. 8 Distribution of the time $T_2$ to the most recent common ancestor of two individuals for different speed $c \ (	ext{a}) c = 0.1$ and $\text{b} c = 0.2$ and typical size $N$ of the population. The inset shows the complementary of the cumulative distribution function $P(T_2 > t)$ in the semilog coordinates. An exponential $\exp(-t/T_0)$ is indicated as a black dashed line. Different line style and color corresponds to $N = 10^4$ (plain blue), $N = 10^5$ (dashed red) and $N = 10^6$ (dot-dashed orange). The mutational variance is fixed to $\sigma^2 = 0.1$ and $\alpha = \beta = 2$ (Color figure online)

Fig. 9 Distribution of the time $\tilde{T}_2$ to the most recent common ancestor of two individuals sampled randomly in the population for different speed $c$ and typical size $N$ of the population. The inset shows the cumulative distribution function $P(T_2 \leq t)$. Different line style and color corresponds to $N = 10^4$ (plain blue), $N = 10^5$ (dashed red) and $N = 10^6$ (dot-dashed orange). The mutational variance is fixed to $\sigma^2 = 0.1$ and $\alpha = \beta = 2$ (Color figure online)

that its lineage coalesces with an other lineage in the population. From Fig. 8, we can observe that after a time delay $T_0 = 1/(\sigma \sqrt{\alpha \beta})$, the time $T_2$ is exponentially distributed. The exponential distribution of $T_2$ is apparent from the inset in Fig. 8 which shows the complementary of the cumulative distribution function $P(T_2 > t)$.

The time $T_2$ is different from the time $\tilde{T}_2$ before the lineages of two individuals sampled randomly in the population coalesce (see Fig. 9).
B.4. The PDE Approach of Sect. 1.4

In this section, we compare the realizations of the IBM model with the ancestral process \( Y_s \), obtained by the simulation of the fundamental solution defined by (21). As mentioned in Sect. 1.4, the distribution \( w^z(s, y) \) of the ancestral process \( Y_s \) starting from a trait \( z \) is given by

\[
w^z(s, y) = \frac{\nu^y(-s, y)}{F(z)}
\]

where \( \nu^z \) satisfies

\[
\begin{aligned}
\frac{\partial}{\partial t} \nu^y(t, z) &= \mathcal{L}(\nu^y(t, \cdot))(z), \\
\nu^y(0, z) &= \delta(z - y) F(z).
\end{aligned}
\] (B.2)

In order to solve numerically this equation, we replace the Dirac \( \delta(z - y) \) with the characteristic function of a small interval around \( y \) of the form \([y - dy, y + dy]\) as pictured in Fig. 1. More precisely, we solve Eq. (B.2) on a finite interval of the form \([z_{min}, z_{max}]\) with \( z_{min} = -z_{max} = -10 \), and we add Dirichlet boundary condition. For the initial conditions, we evenly decompose the interval \([z_{min}, z_{max}]\) into \( n = 200 \) intervals \( I_k \) of size \( dy = (z_{max} - z_{min})/n \) and centered around \( y_k \). The numerical solution of \( \nu^k \) satisfying (B.2) starting from \( \nu^0_k(z) = 1_{I_k}(z) \) is obtain using a semi-explicit Euler scheme. The advantage of the decomposition is that \( \sum_k \nu^k(t, z) = F(z) \) for all \( t > 0 \) and \( z \in [z_{min}, z_{max}] \). Then, the numerical distribution \( w^k(s) \) of \( Y_s \) starting from the trait \( z_i \) is given by \( w^k(s, y_k) = \frac{\nu^k(-s, z_i)}{F(z_i)} \) (see video and Fig. 10). In our simulation, we look at the particular point \( z^* \) which corresponds to the dominant trait of the population, that is the mean of \( F \).

As expected, we observe a good fit between the distribution of the ancestral lineage of the IBM model and the distribution of the ancestral process (see Fig. 10 and video Fig. 11). In particular, the means coincide and each ancestral lineage of the IBM model lies in the region where the ancestral process is the most likely to be (blue region in Fig. 1).

B.5. Deterministic Approximation of the Ancestral Lineage

In this section, we aim to compare the diffusive approximation stated in Sect. 4 and the Hamilton–Jacobi approximation stated in Sect. 5.1 with the ancestral lineage defined in Theorem 1.1. We compute numerically the mean of the ancestral lineage \( E_z(Y_s) \) using the fraction approach detailed in the above section, and we compare it to the formula (42) and to the solution of the ODE (46) (Fig. 12).
Fig. 10  Ancestral lineages over time of individuals with the dominant trait $z^*$: distribution of the ancestral process $Y_s$ (blue region) and the mean of the ancestral lineages of one realization of the IBM model (dashed black curve). The plain black curve and blue curve correspond to the confidence interval at 5% around the mean for the IBM lineages (black) and the ancestral process (blue). The mean of the ancestral lineages corresponds to the dashed black curve, and the mean of the ancestral process $E_z(Y_s)$ is represented by the plain red curve. The gray and green dot-dashed curves represent, respectively, the optimal and the dominant trait (Color figure online)
Fig. 11 Ancestral lineages distribution over time of individuals with the dominant trait $z^*$. DynlineagePDE-vsIBM.mp4 (Color figure online)

Fig. 12 Mean trait of the ancestral lineage along time (plain curve), diffusive approximation (circle marked curve) and the Hamilton Jacobi approximation (diamond marked curve) in the semilog representation. The color corresponds to different parameters: blue curve ($\sigma = 0.05, c = 0.025$), red curve ($\sigma = 0.05, c = 0.05$), orange curve ($\sigma = 0.1, c = 0.05$), purple curve ($\sigma = 0.1, c = 0.1$) (Color figure online)

References

Alfaro M, Berestycki H, Raoul G (2017) The effect of climate shift on a species submitted to dispersion, evolution, growth, and nonlocal competition. SIAM J Math Anal 49(1):562–596
Bansaye V, Cloez B, Gabriel P (2019) Ergodic behavior of non-conservative semigroups via generalized Doeblin’s conditions. Acta Appl Math, pp 1–44
Barles G, Roquejoffre J-M (2006) Ergodic type problems and large time behaviour of unbounded solutions of Hamilton-Jacobi equations. Commun Partial Differ Equ 31(8):1209–1225
Barles G, Mirrahimi S, Perthame B (2009) Concentration in Lotka-Volterra parabolic or integral equations: a general convergence result. Methods Appl Anal 16(3):321–340
Barton NH, Etheridge AM, Véber A (2017) The infinitesimal model. Theor Popul Biol 118:50–73
Bedford T, Cobey S, Pascual M (2011) Strength and tempo of selection revealed in viral gene genealogies. BMC Evol Biol 220(11)
Ancestral Lineages in Mutation Selection Equilibria... Page 41 of 43

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

Berestycki N (2012) Recent progress in coalescent theory. arXiv:math.PR/0909.3985

Berestycki N, Fang J (2018) Forced waves of the fisher-KPP equation in a shifting environment. J Differ Equ 264(3):2157–2183

Berestycki H, Diekmann O, Nagelkerke CJ, Zegeling PA (2009) Can a species keep pace with a shifting climate? Bull Math Biol 71(2):399–429

Berestycki J, Berestycki N, Schweinsberg J (2013) The genealogy of branching Brownian motion with absorption. Ann Probabil 41(2):527–618

Billiard S, Ferrière R, Méladé S, Tran VC (2015) Stochastic dynamics of adaptive trait and neutral marker driven by eco-evolutionary feedbacks. J Math Biol 71(5):1211–1242

Billingsley P (2013) Convergence of probability measures. Wiley, New York

Bouin E, Calvez V, Meunier N, Mirrahimi S, Perthame B, Raoul G, 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, Henderson C, Ryzhik L (2017) Super-linear spreading in local and non-local cane toads equations. Journal de mathématiques Pures et Appliquées 108(5):724–750

Bouin E, Garnier J, Henderson C, Patout F (2018) Thin front limit of an integro-differential fisher-KPP equation with fat-tailed kernels. SIAM J Math Anal 50(3):3365–3394

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

Bradshaw WE, Holzapfel CM (2006) Evolutionary response to rapid climate change. Science 312(5779):1477–1478

Brunet E, Derrida B, Mueller AH, Munier S (2007) Effect of selection on ancestry: an exactly soluble case and its phenomenological generalization. Phys Rev E 76:041104

Brunet E, Derrida B, Mueller AH, Munier S (2007a) Dynamics of lineages in adaptation to a gradual environmental change. Phys Rev E Stat Nonlinear Soft Matter Phys, 76

Bürger R (2000) The mathematical theory of selection, recombination, and mutation. Wiley series in mathematical & computational biology. Wiley, New York

Burger R, Lynch M (1995) Evolution and extinction in a changing environment: a quantitative-genetic analysis. Evolution 49(1):151–163

Calvez V, Garnier J, 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, Dumont T (2018) Non-local competition slows down front acceleration during dispersal evolution. arXiv:1810.07634

Calvez V, Henry B, Méladé S, Tran VC (2021) Dynamics of lineages in adaptation to a gradual environmental change. arXiv:2104.10427

Champagnat N, Henry B et al (2019) A probabilistic approach to Dirac concentration in nonlocal models of adaptation with several resources. Ann Appl Probab 29(4):2175–2216

Champagnat N, Ferrière R, Méladé S (2006) Unifying evolutionary dynamics: from individual stochastic processes to macroscopic models. Theor Popul Biol 69(3):297–321

Champagnat N, Ferrière R, Méladé S (2007) Individual-based probabilistic models of adaptive evolution and various scaling approximations. In: Seminar on stochastic analysis, random fields and applications V, pp 75–113. Springer

Cloez B, Gabriel P (2020) On an irreducibility type condition for the ergodicity of nonconservative semi-groups. Comptes Rendus. Mathématique 358(6):733–742

Coville J, Hamel F (2019) On generalized principal eigenvalues of nonlocal operators with drift. Nonlinear Anal 193:111569

Desai MM, Walczak AM, Fisher DS (2013) Genetic diversity and the structure of genealogies in rapidly adapting populations. Genetics 193(2):565–585

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

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

Ethier SN, Kurtz TG (2009) Markov processes: characterization and convergence, vol 282. Wiley, New York
Figueroa Iglesias S, Mirrahimi S (2018) Long time evolutionary dynamics of phenotypically structured populations in time-periodic environments. SIAM J Math Anal 50(5):5537–5568
Figueroa Iglesias S, Mirrahimi S (2019) Selection and mutation in a shifting and fluctuating environment. HAL Preprint 02320525
Fisher RA (1918) The correlation between relatives on the supposition of mendelian inheritance. Trans R Soc Edinb 52:399–433
Garnier J, Giletti T, Hamel F, Roques L (2012) Inside dynamics of pulled and pushed fronts. Journal de mathématiques pures et appliquées 98(4):428–449
Garnier J, Lafontaine P (2020) Dispersal and good habitat quality promote neutral genetic diversity in metapopulations. arXiv preprint
Gil M-E, Hamel F, Martin G, Roques L (2019) Dynamics of fitness distributions in the presence of a phenotypic optimum: an integro-differential approach. Nonlinearity 32:3485
Hairer E, Lubich C, Wanner G (2006) Geometric numerical integration: structure-preserving algorithms for ordinary differential equations, vol 31. Springer, Berlin
Hairson NG, Ellner S, Geber MA, Yoshida T, Fox J (2005) Rapid evolution and the convergence of ecological and evolutionary time. Ecol Lett 8:1114–1127
Hallatschek O, Nelson DR (2008) Gene surfing in expanding populations. Theor Popul Biol 73(1):158–170
Hallatschek O, Nelson DR (2010) Life at the front of an expanding population. Evolut Int J Org Evolut 64(1):193–206
Hamel F, Lavigne F, Roques L (2020) Adaptation in a heterogeneous environment. I: Persistence versus extinction. arXiv:2005.09869
Hermisson J, Redner O, Wagner H, Baake E (2002) Mutation-selection balance: ancestry, load, and maximum principle. Theor Popul Biol 62(1):9–46
Hoffmann AA, Sgro CM (2011) Climate change and evolutionary adaptation. Nature 470(7335):479
Kallenberg O (2006) Foundations of modern probability. Springer, Berlin
Kimura M (1964) Diffusion models in population genetics. J Appl Probab 1(2):177–232
Kingman JFC (1982) On the genealogy of large populations. J Appl Probab 19:27–43
Kopp M, Matuszewski S (2014) Rapid evolution of quantitative traits: theoretical perspectives. Evolut Appl 7(1):169–191
Lorenzi T, Chisholm RH, Desvillettes L, Hughes BD (2015) Dissecting the dynamics of epigenetic changes in phenotype-structured populations exposed to fluctuating environments. J Theor Biol 386:166–176
Lorz A, Mirrahimi S, Perthame B (2011) Dirac mass dynamics in multidimensional nonlocal parabolic equations. Commun Partial Differ Equ 36(6):1071–1098
Lynch M, Gabriel W, Wood AM (1991) Adaptive and demographic responses of plankton populations to environmental change. Limnol Oceanogr 36:1301–1312
Lynch M, Lande R (1993) Evolution and extinction in response to environmental change. Sinauer Assoc
Marguet A (2019) Uniform sampling in a structured branching population. Bernoulli 25(4A):2649–2695
Martin G, Roques L (2016) The non-stationary dynamics of fitness distributions: asexual model with epistasis and standing variation. Genetics 204(4):1541–1558
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, Gandon S (2020) Evolution of specialization in heterogeneous environments: equilibrium between selection, mutation and migration. Genetics 214(2):479–491
Mischler S, Scher J (2016) Spectral analysis of semigroups and growth-fragmentation equations. In: Annales de l’Institut Henri Poincare (C) Non Linear Analysis, vol 33, pp 849–898. Elsevier
Neher RA, Hallatschek O (2013) Genealogies of rapidly adapting populations. Proc Natl Acad Sci 110(2):437–442
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
Roques L, Garnier J, Hamel F, Klein EK (2012) Allee effect promotes diversity in traveling waves of colonization. Proc Natl Acad Sci 109(23):8828–8833
Roques L, Patout F, Bonnefon O, Martin G (2020) Adaptation in general temporally changing environments. SIAM J Appl Math 80(6):2420–2447
Rouzine I, Coffin J (2007) Highly fit ancestors of a partly sexual haploid population. Theor Popul Biol 71(2):239–250
Turelli M (2017) Commentary: Fisher’s infinitesimal model: a story for the ages. Theor Popul Biol 118:46–49
Turelli M, Barton NH (1994) Genetic and statistical analyses of strong selection on polygenic traits: what, me normal? Genetics 138(3):913–941

Publisher’s Note  Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.