



HAL
open science

Dynamics of lineages in adaptation to a gradual environmental change

Vincent Calvez, Benoît Henry, Sylvie Méléard, Viet Chi Tran

► **To cite this version:**

Vincent Calvez, Benoît Henry, Sylvie Méléard, Viet Chi Tran. Dynamics of lineages in adaptation to a gradual environmental change. *Annales Henri Lebesgue*, 2022, 5, pp.729-777. 10.5802/ahl.135 . hal-03205023

HAL Id: hal-03205023

<https://hal.science/hal-03205023v1>

Submitted on 22 Apr 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Dynamics of lineages in adaptation to a gradual environmental change

Vincent Calvez*, Benoît Henry†, Sylvie Méléard‡, Viet Chi Tran§

April 21, 2021

Abstract

We investigate a simple quantitative genetics model subject to a gradual environmental change from the viewpoint of the phylogenies of the living individuals. We aim to understand better how the past traits of their ancestors are shaped by the adaptation to the varying environment. The individuals are characterized by a one-dimensional trait. The dynamics -births and deaths- depend on a time-changing mortality rate that shifts the optimal trait to the right at constant speed. The population size is regulated by a nonlinear non-local logistic competition term. The macroscopic behaviour can be described by a PDE that admits a unique positive stationary solution. In the stationary regime, the population can persist, but with a lag in the trait distribution due to the environmental change. For the microscopic (individual-based) stochastic process, the evolution of the lineages can be traced back using the historical process, that is, a measure-valued process on the set of continuous real functions of time. Assuming stationarity of the trait distribution, we describe the limiting distribution, in large populations, of the path of an individual drawn at random at a given time T . Freezing the non-linearity due to competition allows the use of a many-to-one identity together with Feynman-Kac's formula. This path, in reversed time, remains close to a simple Ornstein-Uhlenbeck process. It shows how the lagged bulk of the present population stems from ancestors once optimal in trait but still in the tail of the trait distribution in which they lived.

Keywords: stochastic individual-based model, spine of birth-death process, many-to-one formula, ancestral path, historical process, genealogy, phylogeny, resilience.

MSC 2000 subject classification: 92D25, 92D15, 60J80, 60K35, 60F99.

*ICJ, UMR 5208 CNRS & Université Claude Bernard Lyon 1, Lyon France; E-mail: vincent.calvez@math.cnrs.fr

†IMT Lille Douai, Institut Mines-Télécom, Univ. Lille, F-59000 Lille, France; E-mail: benoit.henry@imt-lille-douai.fr

‡Ecole Polytechnique, CNRS, UMR 7641 - CMAP, route de Saclay, 91128 Palaiseau Cedex-France; E-mail: sylvie.meleard@polytechnique.edu

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

Acknowledgements: The authors thank Pierre-Louis Lions for his proof of the uniqueness of the stationary distribution. This work has been supported by the Chair “Modélisation Mathématique et Biodiversité” of Veolia Environnement-Ecole Polytechnique-Museum National d’Histoire Naturelle-Fondation X. V.C.T. also acknowledges support from Labex Bézout (ANR-10-LABX-58). 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 865711).

1 Introduction

An increasing number of studies have demonstrated rapid phenotypic changes in invasive or natural populations that are subject to environmental changes, such as climate change or habitat alteration due for instance to human activities [10, 65, 44, 45, 14, 2, 73]. Such fast evolution may result in adaptation for these populations facing changing environment, as observed in evolutionary experiment [36, 37, 20, 39]. Important theoretical progress has been made to predict phenotypic evolution in a changing environment since the pioneer works of [57, 58, 13, 54], see [52] for a review.

A major prediction of these models is that when the optimal phenotype changes linearly with time, the phenotypical distribution of individuals is moving at the same speed to keep pace of the change, but is lagged behind the optimum. The equilibrium value of the lag depends on the rate of the change, on the genetic variance and on the strength of selection. Above a critical rate of change of the optimal phenotype with time, this lag becomes so large that the fitness of the population falls below the value that allows its persistence and the population is doomed to extinction. In the case of small enough lag such that the population persists under constant adaptation, we study, in the present paper, the genealogies of its individuals. Our aim is to understand how individual dynamics build the macroscopic adaptation of the population *via* the quantitative description of the typical ancestral lineage.

We consider a population dynamics where individuals are characterized by a trait $x \in \mathbb{R}$ and that give birth and die in continuous time. During their life the individual trait variations are modelled by a diffusion operator with variance σ^2 . The environment in which the population lives is shifted at constant speed $\sigma c > 0$ that drives the adaptation of the population. The constant $c > 0$ can be interpreted as the speed of environmental change. Two complementary descriptions of the dynamics are opted for: (i) a macroscopic, deterministic, description of the phenotypic density, and (ii) a microscopic, stochastic, description of the individual phenotypes. We believe that the later is better suited for the analysis of the lineages.

The macroscopic dynamics of the phenotypic density in the moving environment is described by the following partial differential equation (PDE) for the density with respect

to the one-dimensional trait:

$$\partial_t u(t, x) = \frac{\sigma^2}{2} \partial_{xx}^2 u(t, x) + \left(1 - \frac{1}{2}(x - \sigma ct)^2 - \int_{\mathbb{R}} u(t, y) dy \right) u(t, x). \quad (1)$$

Here, $u(t, x)$ denotes the density of population at time t and trait $x \in \mathbb{R}$. Due to the environmental change, the optimal trait with regards to the growth rate $1 - (x - \sigma ct)^2/2$ is $x_t^* = \sigma ct$. The choice of the scaling of the speed is discussed after Theorem 1.1 below. The nonlinear term involving the total mass of the population accounts for the mean field competition between individuals at time t .

It is well known (cf. [15, 16, 34]) that equation (1) can be derived from a stochastic system describing the random individual dynamics. More precisely, we consider the following branching-diffusion process with interaction. An individual, at trait $x \in \mathbb{R}$ at time $t \geq 0$, gives birth to a new individual at the same trait with rate 1. Each individual dies with rate $(x - \sigma ct)^2/2 + N_t/K$, where N_t is the total population size at time t and K is the carrying capacity of the system. The natural death rate $(x - \sigma ct)^2$ reflects the gradual environmental change, as in the PDE. The term N_t/K in the death rate corresponds to density-dependent competition. Changes in the trait during the lives of individuals are driven by independent Brownian motions, accounting for infinitesimal changes of the phenotypes. It is standard to rigorously prove that the empirical measure on the individual traits weighted by $1/K$ satisfies a semi-martingale decomposition, which is a stochastic equation analogous to (1) (and given later), and that it converges weakly to the solution of the PDE when K tends to infinity (provided the initial conditions are scaled suitably).

Due to the environmental change, the behavior of the population is naturally observed in the moving frame. In what follows, we will always work in this setting, defining the density in the moving frame as $f(t, x) = u(t, x + \sigma ct)$. As such, we obtain an additional transport term in the PDE (associated with a drifted Brownian motion in the individual-based model):

$$\partial_t f(t, x) = \frac{\sigma^2}{2} \partial_{xx}^2 f(t, x) + \sigma c \partial_x f(t, x) + \left(1 - \frac{x^2}{2} - \int_{\mathbb{R}} f(t, y) dy \right) f(t, x). \quad (2)$$

The unique positive stationary state of this equation can easily be computed. It can exist if, and only if, $1 - \sigma/2 - c^2/2 > 0$, which is the persistence condition on the speed c . Under this condition, the stationary state is a weighted Gaussian density centered on $-c$, with variance σ , hereafter denoted by F (see Section 2.2). The shift by c relative to the fitness optimum at $x = 0$ can be interpreted as a lag in the process of adapting to a moving environment. Indeed individuals try keeping pace of the gradual change, so that they can never be optimal in average. This maladaptation can be measured by the shift c which is associated with a load in the fitness of value $c^2/2$.

Additionally, this model predicts that the population collapses when the speed of environmental change is above a certain threshold $c^* = (2 - \sigma)^{1/2}$. Here, we consider that the speed c is below c^* , as already mentioned above.

Our purpose here is to provide more insight on this phenomenon by studying the trait ancestry of the individuals at a given time T , *i.e.* the sequence of traits of their ancestors in the past. We assume that, in the moving frame, the population dynamics are nearly stationary, starting from the equilibrium F . In particular, the solution $f(t, x)$ of the deterministic PDE (2) remains constant in time, equal to this equilibrium. Consequently, the stochastic process will stay close to this equilibrium on finite time intervals in the regime of large population. In the stationary regime the dynamics of the PDE is trivial but the dynamics of the lineages are not, as can be seen on numerical simulations of the individual-based models (cf. Fig. 1, both in the original variables, and in the moving frame).

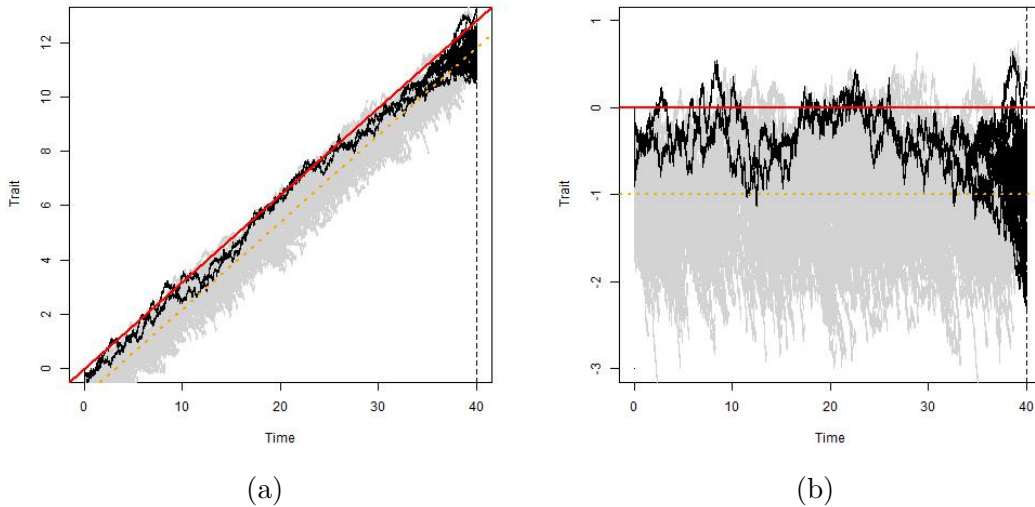


Figure 1: Ancestral lineages of the present population. To an individual of trait x living at time T , its lineage corresponds to the function that associates with each time $t < T$ the trait of this individual if it was already born, or else the trait of its closest ancestor at that time (its parent if the latter was born, otherwise its grand-parent etc.). The traits in the population (ordinate) are shown with respect to time (abscissa). The extinct lineages are in gray, whereas the lineages of the living particles are in black. As can be seen, the trait distribution is nearly stationary (gray background on the right image), whereas the lineages follow an Ornstein-Uhlenbeck process (see our main result, Theorem 1.1). (a): fixed frame. (b): moving frame. The parameters are $c = 1$, $\sigma = 0.32$, $K = 250$.

We observe the following pattern: a stabilized cloud of points representing the stationary state and solid lines representing the lineages, highlighting the response to environment. One observes that the individuals alive at the final observation time are all coming from past individuals whose traits were far from being representative in the past distribution but who were better fitted.

More precisely, we will describe the approximate dynamical lineage of a fixed individual sampled uniformly in a large population at a time $T > 0$. We will show the following

Theorem 1.1 stating that in backward time, these trajectories are asymptotically (when the carrying capacity K tends to infinity), Ornstein-Uhlenbeck processes.

Theorem 1.1. *In the moving framework, assuming the trait distribution of the population stationary, the backward in time process describing the lineage of an individual sampled in the living population at time $T > 0$ converges, when $K \rightarrow +\infty$, to the following time homogeneous Ornstein-Uhlenbeck process driving the ancestral trajectories around 0, according to the equation*

$$d\widehat{Y}_s = -\sigma\widehat{Y}_s ds + \sigma dW_s, \quad (3)$$

for a Brownian motion W .

This result is made more precise in Theorem 4.10. A similar conclusion was derived for a similar model, independently of this work, by another approach in [33]. The latter analysis remains on a macroscopic level and follows the tracking of neutral fractions in the PDE, as initiated in [69].

It is an immediate observation that the Ornstein-Uhlenbeck process is independent of the speed c . This is indeed due to our choice of scaling the speed of change c by the standard deviation σ in (1) and (2). This is to say that the speed of change is measured relatively to how many units of standard mutational deviation are shifted per time unit. With this scaling, the lag load $c^2/2$ is independent of the mutational variance rate σ^2 . In particular it does not vanish as the mutational variance goes to zero.

Although we cannot handle the long time asymptotics with our methodology, we can still notice that the stationary distribution of the backward Ornstein-Uhlenbeck process is another Gaussian distribution centered at the origin, with variance $\sigma/2$. Hence, individuals sampled at time T come from ancestors that were close to being optimal in the past, but not representative in the distribution at that time (see also Fig. 1).

Notice that in the extreme case of a vanishing variance $\sigma^2 \rightarrow 0$, a simple long time scaling $s' = \sigma s$ in the SDE (3) makes it close to the deterministic ODE $d\widehat{Y}_{s'} = -\widehat{Y}_{s'} ds'$. The solution of the latter equation converges in long time to 0.

This study shows how important it is in ecology or agriculture to preserve the trait diversity, as the subpopulation with the majority trait may not be the one ensuring the survival of the species in case of environmental shift. In cancer therapy or for understanding antibiotic resistances, our results show that the eradication of such majority trait with gradual effects of drugs or antibiotics may not be enough to fight against the persistence of tumors or bacterial strains. We also refer to [39] for similar consideration in experimental evolution.

The proof of Theorem 1.1 is now sketched. Our approach mixes here two points of view based on the stochastic individual-based model: on the one hand, the spinal approach as developed for branching diffusion in [3, 41, 42, 60, 61] and on the other hand, the historical processes, as introduced in Dawson and Perkins [21, 66, 67] and Dynkin [28], and then developed in Méléard and Tran [62] (with a correction, see [50, 74]).

The historical process taken at a time t describes the history of each individual in the population stopped at time t . Since the individual death rate depends on the total size of the population, the historical process cannot be reduced to an accumulation of independent trajectories. Nevertheless, assuming that the initial condition converges to the stationary solution of (2), an important step in our approach is to replace (up to a negligible error that we can control) the nonlinearity in the stochastic population process (the total number of individuals) by the mass of the stationary distribution. The birth-and-death process with diffusion becomes a branching-diffusion process and computation becomes rather easier. By coupling techniques, we can therefore capture the dynamics of the historical process using reasoning proper to branching-diffusion processes and we can easily prove in this context formulas based on the so-called many-to-one formulas describing the distribution of the ancestry (in forward time) of a typical individual in the population living at time T , as it is done in a general context in [61] with a more complicate proof (since more general). Furthermore, the coupling also allows to justify the use of the well known spinal theory to obtain the law of an individual chosen uniformly at random at time T .

The process that we obtain involves the expectation of the number $m(t, x)$ of individuals at time t issued from one individual with trait x . This quantity is obtained as expectation of an additive functional of a drifted Brownian motion and can be explicitly computed by tricky arguments based on Girsanov transform and inspired by [32]. Note that this computation allows to obtain the explicit value of the solution of

$$\begin{aligned} \partial_t m(t, x) &= \frac{\sigma^2}{2} \partial_{xx}^2 m(t, x) - \sigma c \partial_x m(t, x) + \left(1 - \frac{x^2}{2} - \|F\|_1\right) m(t, x) \\ m_0(x) &= 1. \end{aligned} \quad (4)$$

The many-to-one formula allows to characterize the forward lineage dynamics as obtained from an auxiliary non-homogeneous Markov process. In this specific case, we obtain the exact trajectory of the trait lineage and prove that they are Gaussian at any time. The last step consists in using the results by Hausmann and Pardoux [43] on time reversed diffusion processes. We prove that the time reversed paths are Ornstein-Uhlenbeck processes attracted by 0 as stated by Theorem 1.1. This proves how the genealogical tree is strongly unbalanced in our case, as observed in Fig. 1.

For alternative points of view, let us mention that there has been a large literature related to our work. First, there has been a considerable amount of studies dealing with simple models of waves advancing a fitness landscape in an asexual reproducing population, starting from the seminal papers [75, 49], see also [70] for a similar model with a nonlinear diffusion operator. In these models, the trait is the fitness itself (*i.e.* the growth rate per capita) centered by its average on the population, so that new mutants can outcompete the resident population if their fitness is higher than the mean. We also refer to analogous studies in the absence of deleterious mutations, by [23] (including experimental evidence supporting the theory), [4] in the context of oncogenesis, and [64] for a review article. Mathematical results in this direction were obtained in [27], then [71]. Several authors

also investigated the structure of the genealogies in stochastic models, exhibiting coalescent structures. For coalescent processes in modelling genealogies for populations without competition or interaction non-linearities, we refer to [8] for a review. For directed selection, when the population at latter stages is issued from individuals at the tip of the wave, strongly asymmetric genealogical trees arise (see [12, 11, 24, 63, 72, 7]). In [55], the genealogies in an adaptive dynamics time scale are described with a forward-backward coalescent. For structured populations with competition, other approaches include the look-down processes [25, 26, 30] or the tree-valued descriptions as in [5, 38, 51]. Let us emphasize that, here, we focus on typical lineages rather than coalescent analysis. This is left for a future work.

Finally, let us cite other mathematical contributions with spatial displacement and competition local in space (contrary to (1) where it is global in trait) [56, 68, 6, 1]. However, these studies focus on the ability of the species to keep pace of the climate change, *i.e.* the conditions of persistence for the species, rather than on lineages dynamics.

In Section 2, we introduce and study the individual-based stochastic measure-valued process underlying the PDE (2). The stationary solution of this PDE, which will play a central role in what follows, is also carefully detailed. The stochastic processes associated with (2) are non-linear because of the competition term. However, when we are close to the equilibrium, a coupling with a linear birth-death process (with a time-varying growth rate) is possible. This coupling holds for the trait distribution at a given time T but also for the historical picture, *i.e.* for the ancestral paths of the individuals alive at T . This is explained in Section 3. For the linear birth-death process, we can apply a Feynman-Kac formula. This, together with fine stochastic calculus techniques, allows us to compute the exact solution of (4). In Section 4, we use a many-to-one formula together with the expression of $m_t(x)$ obtained previously and the coupling of historical processes to obtain the approximating stochastic differential equation (SDE) satisfied by the ancestral path of an individual chosen at random in the population at a given time T . This SDE is non-homogeneous in time but its time-reverse SDE is a simple time-homogeneous Ornstein-Uhlenbeck process.

2 The partial differential equation and the population process in the moving framework

2.1 The underlying measure-valued stochastic process

As explained in the introduction, we are interested in the dynamics of the population density in the moving framework. We have seen that it is given by (2). This equation is well posed. Existence of a weak solution will be obtained from the study of the underlying stochastic process and uniqueness by use of the associate mild equation.

Let us introduce the stochastic process associated with Equation (2). On a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, we consider a random process $(Z_t^K)_{t \in \mathbb{R}_+}$ with values in the set of point

measures on \mathbb{R} , and defined by

$$Z_t^K = \frac{1}{K} \sum_{i \in V_t^K} \delta_{X^i(t)}, \quad (5)$$

where V_t^K is the set of labels of individuals alive at time t and where $X^i(t)$ denotes the position of the i -th individual at time t . Individual labels can be chosen in the Ulam-Harris-Neveu set $\mathcal{I} = \cup_{n \in \mathbb{N}} \mathbb{N}^n$ (e.g. see [35]) where offspring labels are obtained by concatenating the label of their parent with their ranks among their siblings. Note that the size N_t^K of the population at time t satisfies $N_t^K = |V_t^K| = K \langle Z_t^K, 1 \rangle$, where the brackets are the notation for the integral of the constant function equal to 1 with respect to the measure $Z_t^K(dx)$. More generally, for a finite measure μ and a positive measurable function φ , $\langle \mu, \varphi \rangle = \int_{\mathbb{R}} \varphi(x) \mu(dx)$ denotes the integral of φ with respect to μ .

In the sequel, we will denote by $\mathcal{M}_F(\mathbb{R})$ the set of finite measures on \mathbb{R} equipped with the topology of weak convergence. The process Z^K belongs to $\mathbb{D}(\mathbb{R}_+, \mathcal{M}_F(\mathbb{R}))$, the space of left-limited and right-continuous processes with values in $\mathcal{M}_F(\mathbb{R})$, that we equip with the Skorokhod topology (see e.g. [9]).

When times vary, the process $(Z_t^K)_{t \in \mathbb{R}_+}$ defines a Markov process whose transitions are as follows. For an individual at position x in the population of N individuals, its birth rate is 1 and its death rate is $x^2/2 + (N-1)/K$. Between the jumps, the positions $X^i(t)$ behave as a drifted Brownian motions $\sigma B_t - c \sigma t$ started at their positions after the jump. All individual births and deaths events and the diffusions between jumps are independent but the interaction between individuals to survive is modeled at the individual level by the additional death rate $(N-1)/K$.

Following Champagnat-Méléard [17], we can construct the process Z^K as the unique solution of a stochastic differential equation driven by a Poisson point measure and Brownian motions indexed by \mathcal{I} . (see Appendix A.1). From this representation, and using stochastic calculus for diffusions with jumps (e.g. [46]), we can derive the following moment estimates, proved in Appendix B:

Lemma 2.1. *We assume that the initial condition Z_0^K satisfies for $\epsilon > 0$ that:*

$$\sup_{K \in \mathbb{N}^*} \mathbb{E}(\langle Z_0^K, 1 \rangle^{2+\epsilon}) < +\infty \quad \text{and} \quad \sup_{K \in \mathbb{N}^*} \mathbb{E}(\langle Z_0^K, x^4 \rangle^{1+\epsilon}) < +\infty. \quad (6)$$

Then, for any $T > 0$, we have

$$\sup_{K \in \mathbb{N}^*} \mathbb{E} \left(\sup_{t \in [0, T]} \langle Z_t^K, 1 \rangle^{2+\epsilon} \right) < +\infty \quad \text{and} \quad \sup_{K \in \mathbb{N}^*} \mathbb{E} \left(\sup_{t \in [0, T]} \langle Z_t^K, x^2 \rangle^{1+\epsilon/2} \right) < +\infty. \quad (7)$$

It is also standard to write the semi-martingale decomposition of the process $(\langle Z_t^K, \varphi \rangle)_{t \in \mathbb{R}_+}$

for a function $\varphi \in C_b^2(\mathbb{R})$, under the assumption (6):

$$\begin{aligned} \langle Z_t^K, \varphi \rangle &= \langle Z_0^K, \varphi \rangle \\ &+ \int_0^t \int_{\mathbb{R}} \left\{ \left(1 - \frac{1}{2}x^2 - \langle Z_s^K, 1 \rangle \right) \varphi(x) - \sigma c \varphi'(x) + \frac{\sigma^2}{2} \varphi''(x) \right\} Z_s^K(dx) ds + M_t^{K,\varphi}, \end{aligned} \quad (8)$$

where the process $M^{K,\varphi}$ is a square integrable martingale with predictable quadratic variation process given by

$$\langle M^{K,\varphi} \rangle_t = \frac{1}{K} \int_0^t \int_{\mathbb{R}} \left\{ \left(1 + \frac{x^2}{2} + \langle Z_s^K, 1 \rangle \right) \varphi^2(x) + \sigma^2 (\varphi')^2(x) \right\} Z_s^K(dx) ds. \quad (9)$$

In the next section we will need a mild version of this equation. To do that, we introduce the semigroup $(P_t)_{t \in \mathbb{R}_+}$ of the process $\sigma B_t - c\sigma t$ and we define, for a fixed $t > 0$ and for $\varphi \in C_b^2(\mathbb{R})$,

$$\psi(s, x) = P_{t-s}\varphi(x). \quad (10)$$

Using the trajectorial representation of Z_t^K (cf. Appendix A.1) and integrating these functions, we show in Appendix A.3 that:

$$\langle Z_t^K, \varphi \rangle = \langle Z_0^K, P_t\varphi \rangle + \int_0^t \int_{\mathbb{R}} \left(1 - \frac{x^2}{2} - \langle Z_s^K, 1 \rangle \right) P_{t-s}\varphi(x) Z_s^K(dx) ds + \mathcal{M}_t^{K,\varphi}, \quad (11)$$

where $\mathcal{M}_t^{K,\varphi}$ is a square integrable martingale computed explicitly in Appendix A.3.

Theorem 2.2. *Let us assume that the initial condition $(Z_0^K(dx))_K$ satisfies (6) and that $(Z_0^K(dx))_K$ converges in probability (weakly as measures) to the deterministic finite measure $\xi_0(dx)$. Let $T > 0$ be given. The sequence of processes $(Z_t^K)_{t \in [0, T]}$ converges in probability and in \mathbb{L}^2 , in $\mathbb{D}([0, T], \mathcal{M}_F(\mathbb{R}))$ to a deterministic continuous function $(\xi_t, t \leq T)$ of $C([0, T], \mathcal{M}_F(\mathbb{R}))$, satisfying for each $t > 0$ that $\langle \xi_t, 1 + x^2 \rangle < +\infty$ which is the unique solution of the weak equation: $\forall \varphi \in C_b^2(\mathbb{R})$,*

$$\langle \xi_t, \varphi \rangle = \langle \xi_0, \varphi \rangle + \int_0^t \int_{\mathbb{R}} \left\{ \left(1 - \frac{1}{2}x^2 - \langle \xi_s, 1 \rangle \right) \varphi(x) - \sigma c \varphi'(x) + \frac{\sigma^2}{2} \varphi''(x) \right\} \xi_s(dx) ds. \quad (12)$$

More precisely:

$$\lim_{K \rightarrow \infty} \mathbb{E}(\sup_{t \leq T} |\langle Z_t^K, \varphi \rangle - \langle \xi_t, \varphi \rangle|^2) = 0. \quad (13)$$

Moreover, for any $t > 0$, the measure ξ_t is absolutely continuous with respect to Lebesgue measure and its density $f(t, x)$ is solution of (2) issued from ξ_0 .

Proof We break the proof into several steps.

Step 1: Let us first prove the uniqueness of ξ solution of (12). For a test function $\psi \in C_b^{1,2}(\mathbb{R}_+ \times \mathbb{R})$ of s and x , we have by standard arguments that:

$$\langle \xi_t, \psi(t, \cdot) \rangle = \langle \xi_0, \psi(0, \cdot) \rangle + \int_0^t \int_{\mathbb{R}} \left\{ \partial_s \psi(s, x) + \left(1 - \frac{1}{2}x^2 - \langle \xi_s, 1 \rangle \right) \psi(s, x) - \sigma c \partial_x \psi(s, x) + \frac{\sigma^2}{2} \partial_{xx}^2 \psi(s, x) \right\} \xi_s(dx) ds$$

Now, we define for a fixed $t > 0$, for $\varphi \in C_b^2(\mathbb{R})$, the $C_b^{1,2}(\mathbb{R})$ function ψ^t by

$$\psi^t(s, x) = \mathbb{E}_x \left(\varphi(Y_{t-s}) \exp \left(- \int_0^{t-s} \frac{Y_u^2}{2} du \right) \right),$$

where Y is the drifted Brownian motion $dY_t = \sigma(dB_t - cdt)$. Then

$$\begin{aligned} \partial_s(\psi^t)(s, x) + \left(1 - \frac{1}{2}x^2 - \langle \xi_s, 1 \rangle \right) \psi^t(s, x) - \sigma c \partial_x(\psi^t)(s, x) + \frac{\sigma^2}{2} \partial_{xx}^2(\psi^t)(s, x) \\ = (1 - \langle \xi_s, 1 \rangle) \psi^t(s, x), \end{aligned} \quad (14)$$

since $\mathbb{E}_x(\varphi(Y_{t-s}))$ is solution of the backward ‘‘heat’’ equation. Noting that $\psi^t(t, x) = \varphi(x)$, and coming back to (14) with this function, we obtain

$$\langle \xi_t, \varphi \rangle = \langle \xi_0, \psi^t(0, \cdot) \rangle + \int_0^t \int_{\mathbb{R}} \left(1 - \langle \xi_s, 1 \rangle \right) \psi^t(s, x) \xi_s(dx) ds.$$

Notice that if $\|\varphi\|_{\infty} \leq 1$, then $\|\psi^t(s, \cdot)\|_{\infty} \leq 1$. By a Gronwall argument we easily prove (see for example Fournier and Méléard [34]) that two solutions in $\mathcal{C}([0, T], \mathcal{M}_F(\mathbb{R}))$ of this equation started with the same initial condition coincide.

Since the transition semi-group (P_t) of the process (Y_t) is absolutely continuous with respect to Lebesgue measure for any $t > 0$, we also deduce by using Fubini’s theorem that the same property holds for ξ_t . Then we write

$$\xi_t(dx) = f(t, x) dx$$

and the function f is the unique weak solution of (2) issued from ξ_0 .

Step 2: The proof of the convergence is obtained by a compactness-identification-uniqueness argument and the tightness is deduced from the uniform moments obtained in Lemma 2.1. It is postponed in Appendix.

Step 3: Since the sequence of processes is proved to converge in law in $\mathbb{D}([0, T], \mathcal{M}_F(\mathbb{R}))$ to a deterministic function, it also converges in probability. The limit is continuous in time and thus the convergence is also a uniform convergence (see [9, p. 124]). Then we have proved that for any $T > 0$, for any continuous and bounded function φ , for any $\varepsilon > 0$,

$$\lim_{K \rightarrow \infty} \mathbb{P}(\sup_{t \leq T} |\langle Z_t^K, \varphi \rangle - \langle \xi_t, \varphi \rangle| > \varepsilon) = 0.$$

Moreover, uniform moment estimates yield uniform integrability and then we also have (13). \square

2.2 A unique positive stationary distribution

For Equation (2), computation is simple and the existence and explicit value of a stationary state are easy to obtain. Uniqueness is more delicate. In the next sections, we will be interested in considering as initial condition ξ_0 the stationary state of Equation (2).

Proposition 2.3. *There exists a unique non zero positive stationary distribution of (2) if and only if*

$$\frac{c^2}{2} + \frac{\sigma}{2} < 1. \quad (15)$$

In this case, the equilibrium is given by

$$F(x) = \frac{\lambda}{\sqrt{2\pi\sigma}} \exp\left(-\frac{(x+c)^2}{2\sigma}\right), \quad (16)$$

with

$$\|F\|_1 = \lambda = 1 - \frac{c^2}{2} - \frac{\sigma}{2}. \quad (17)$$

Let us note that under Condition (15), the population will persist in long time and its long time density admits a mode in $-c$. This value differs from the optimal trait 0, which can be interpreted as a lag in the adaptation to environmental change. See Fig. 2. Indeed, in long time, the solution u of (1) behaves as $F(x - c\sigma t)$ optimal at $-c + c\sigma t$.

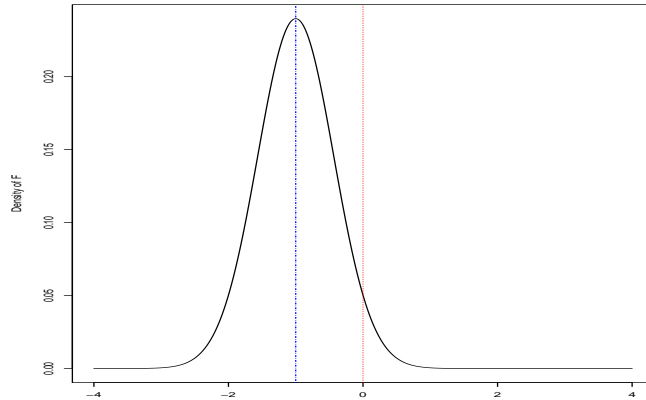


Figure 2: Density of the stationary solution F of (2). The mode of this density is $-c = -1$ here, and the variance of F is $\sigma = 0.32$, so that $\lambda = 0.34$.

Proof The announced proposition can be obtained from general results in the literature, as the ones of Cloez and Gabriel [19]. We give here a simple proof. Deriving twice the

function F defined in (16) and replacing in (2) proves that

$$\frac{\sigma^2}{2}F'' + c\sigma F' + \left(1 - \frac{1}{2}x^2 - \int_{\mathbb{R}} F(y)dy\right)F = 0, \quad (18)$$

if and only if $\lambda = 1 - \frac{c^2}{2} - \frac{\sigma}{2}$. Then under this condition, F is a stationary state.

Let us define the operator A on $C_b^2(\mathbb{R})$ by

$$A\phi = -\frac{\sigma^2}{2}\phi'' - c\sigma\phi' + \frac{1}{2}x^2\phi,$$

for $\phi \in C_b^2(\mathbb{R}, \mathbb{R})$. Then F is solution of $AF = \alpha F$, with $\alpha = 1 - \int F$. Let us also notice that if \tilde{F} is defined as F but replacing in (16) c by $-c$, then $A^*\tilde{F} = \alpha\tilde{F}$. Let us now consider (μ, φ) a solution of $A\varphi = \mu\varphi$ for a positive function φ satisfying $\|\varphi\|_1 = \lambda$. Then we have

$$\int A\varphi\tilde{F} - \int A^*\tilde{F}\varphi = 0 = \int (\alpha - \mu)\varphi\tilde{F},$$

with positive functions φ, \tilde{F} and then $\mu = \alpha$.

Let us now prove that $\varphi = F$. Straightforward computation with $\psi = \sqrt{F}\varphi$ yields

$$A\psi - \alpha\psi = -\frac{\sigma^2}{8} \frac{(\varphi'F - \varphi F')^2}{\psi^3} \leq 0.$$

Further, we note that

$$\int A\psi\tilde{F} = \int \psi A^*\tilde{F} = \alpha \int \psi\tilde{F}.$$

Then, if $D = -\frac{\sigma^2}{8} \frac{(\varphi'F - \varphi F')^2}{\psi^3}$, then $\int D\tilde{F} = 0$, with $D\tilde{F} \leq 0$ and we deduce that $D = 0$ (since D is continuous). Then we obtain $\varphi'F = \varphi F'$ and finally that $(\sqrt{\frac{\varphi}{F}})' = 0$, which implies that φ and F are proportional. Since they are both positive with the same \mathbb{L}^1 norm, they are equal. □

The next corollary is then an obvious consequence of (13).

Corollary 2.4. *Let us assume that the initial measures Z_0^K converge weakly to $F(x)dx$ when K tends to infinity, then for any continuous and bounded function φ*

$$\lim_{K \rightarrow \infty} \mathbb{E}(\sup_{t \leq T} |\langle Z_t^K, \varphi \rangle - \langle F, \varphi \rangle|^2) = 0. \quad (19)$$

3 Feynman-Kac approach for an auxiliary branching-diffusion process

3.1 Coupling of the process Z^K with a branching-diffusion process

Let us assume in all what follows that the initial measures Z_0^K weakly converge to $F(x)dx$ when K tends to infinity as in the Assumption of Corollary 2.4.

As explained in introduction, we are interested in capturing the genealogies of our particle system. Recall that the ancestral lineage or past history of an individual living at time T consists in the succession of ancestral traits: it is obtained by the concatenation of the (diffusive) paths of this individual with the path of their parent before their birth, then with the path of their grand-parent before the birth of their parent etc. To sum up, the lineage of an individual alive at time T is the path that associates with each time $t \leq T$ the trait of its most recent ancestor at this time. Because of the interactions between individuals, the shape of the lineages of living individuals reflects the competition terms in the past, with lineages that might be extinct. Thus, obtaining an equation describing the ancestry of a “typical individual” chosen at random in the population at T is difficult to obtain. See for example the developments of Perkins [66] but with assumptions that exclude logistic competition or see the attempts in [62]. Corollary 2.4 suggests us to replace the interaction logistic term $\langle Z_t^K, 1 \rangle$ by the constant $\|F\|_1 = \int F(x)dx$. The new process is a much more tractable branching particle system.

Therefore we couple Z^K with an auxiliary measure-valued process $(\tilde{Z}_t^K)_{t \geq 0}$, started from the same initial condition Z_0^K and with the same transitions, except that the logistic term is frozen at $\lambda = \|F\|_1$ (see Appendix A.1).

For the auxiliary process, (8) becomes, for any $\varphi \in C_b^2(\mathbb{R})$,

$$\langle \tilde{Z}_t^K, \varphi \rangle = \langle Z_0^K, \varphi \rangle + \int_0^t \int_{\mathbb{R}} \left\{ \left(1 - \frac{1}{2}x^2 - \lambda \right) \varphi(x) - c\sigma\varphi'(x) + \frac{\sigma^2}{2}\varphi''(x) \right\} \tilde{Z}_s^K(dx) ds + \tilde{M}_t^{K,\varphi}, \quad (20)$$

where $\tilde{M}^{K,\varphi}$ is a square integrable martingale with predictable quadratic variation

$$\langle \tilde{M}^{K,\varphi} \rangle_t = \frac{1}{K} \int_0^t \int_{\mathbb{R}} \left\{ \left(1 + \frac{x^2}{2} + \lambda \right) \varphi^2(x) + \sigma^2(\varphi')^2(x) \right\} \tilde{Z}_s^K(dx) ds. \quad (21)$$

Let us remark that with the same arguments as in Theorem 2.2, we can prove that for any $T > 0$, the measure-valued process \tilde{Z}^K converges in $\mathbb{D}([0, T], \mathcal{M}_F(\mathbb{R}))$ uniformly and in probability to the unique weak solution $(\tilde{\xi}_t, t \geq 0)$ of

$$\partial_t \tilde{\xi}(t, x) = \frac{\sigma^2}{2} \Delta \tilde{\xi}(t, x) + c\sigma \partial_x \tilde{\xi}(t, x) + \left(1 - \frac{1}{2}x^2 - \lambda \right) \tilde{\xi}(t, x). \quad (22)$$

starting from the initial data $\tilde{\xi}(0, x) = F(x)$. By an analogous argument as previously, the measure $\tilde{\xi}_t$ has a density for any $t > 0$ whose uniqueness is classical. Further F is also its unique positive stationary distribution (with given norm). We also have a similar convergence as in (19): for any continuous and bounded function φ ,

$$\lim_{K \rightarrow \infty} \mathbb{E}(\sup_{t \leq T} |\langle \tilde{Z}_t^K, \varphi \rangle - \langle F, \varphi \rangle|^2) = 0. \quad (23)$$

As an immediate corollary, we can couple the process Z^K and the branching-diffusion process \tilde{Z}^K , using that $|\langle \tilde{Z}_t^K, \varphi \rangle - \langle Z_t^K, \varphi \rangle| \leq |\langle \tilde{Z}_t^K, \varphi \rangle - \langle F, \varphi \rangle| + |\langle Z_t^K, \varphi \rangle - \langle F, \varphi \rangle|$.

Proposition 3.1. *Assume that the initial conditions $(Z_0^K)_K$ satisfy (6) and that $Z_0^K \xrightarrow[K \rightarrow \infty]{w} F$. Then for any continuous and bounded function φ ,*

$$\lim_{K \rightarrow +\infty} \mathbb{E}(\sup_{t \leq T} |\langle Z_t^K, \varphi \rangle - \langle \tilde{Z}_t^K, \varphi \rangle|^2) = 0.$$

We can now work with the process $(\tilde{Z}_t^K)_{t \geq 0}$. The main improvement with this process is that the nonlinearity has been tackled. Therefore the process satisfies the branching property and we are authorized to use some classical tools for these processes.

3.2 Coupling of the historical processes

Until now, we described the evolving distributions of the trait, but the individual dimension is lost when the population becomes large. In the sequel, we will also investigate the large population dynamics of the historical processes, which at a time t describes the trait ancestry of individuals alive at that time. Recall the definition of the lineage of an individual given in Section 2.1. For an individual $i \in V_T^K$, let us define their lineage X^i . In (5), the labels are taken in the Ulam-Harris-Neveu set \mathcal{I} and we can define by \preceq the usual partial order on \mathcal{I} : $j \preceq i$ means that $j \in \mathcal{I}$ is the ancestor of $i \in \mathcal{I}$ i.e. that there exists $k \in \mathcal{I}$ such that $i = (j, k)$, the concatenation of the labels j and k . If the individual $i \in V_T^K$ was living at time t , then $X^i(t)$ still denotes the position of i at t . But if the individual i was not born at time t , then, $X^i(t) = X^j(t)$ where $j \prec i$ is the most recent ancestor of i living at t . Since an offspring inherits their parent's trait at birth, and since the trait evolves continuously according to a diffusion during an individual's life, such lineage is a continuous function on \mathbb{R} . The path is extended after time t by the trait value at time t , so that this continuous function can be defined from \mathbb{R} to \mathbb{R} (and not from $[0, t]$ to \mathbb{R}).

Here, we will adopt the approach developed in Dawson and Perkins [21] or Méléard and Tran [62]. Let us define the historical process H^K as the following càdlàg process with values in $\mathcal{M}_F(\mathcal{C}(\mathbb{R}_+, \mathbb{R}))$:

$$H_t^K = \frac{1}{K} \sum_{i \in V_t^K} \delta_{X_{s \wedge t}^i}, \quad (24)$$

where $(X_{s \wedge t}^i, s \in \mathbb{R}_+)$ is the lineage of the individual $i \in V_t^K$. To investigate the asymptotic behavior of this process, we introduce (as in Dawson [22] or Etheridge [29]) the class of test functions on paths of the form: $\forall y \in \mathcal{C}(\mathbb{R}_+, \mathbb{R})$,

$$\varphi(y) = \prod_{j=1}^m g_j(y_{t_j}), \quad (25)$$

for $m \in \mathbb{N}^*$, $0 = t_0 \leq t_1 < \dots < t_m$ and $\forall j \in \{1, \dots, m\}$, $g_j \in \mathcal{C}_b^2(\mathbb{R}, \mathbb{R})$. As proved in [22], this class is convergence determining.

If y is a continuous path stopped at time t (as the trajectories chosen according $H_t^K(dy)$), then

$$\varphi(y) = \prod_{j=1}^m g_j(y_{t_j \wedge t}) = \sum_{k=0}^{m-1} \mathbb{1}_{[t_k, t_{k+1})}(t) \left(\prod_{j=1}^k g_j(y_{t_j}) \prod_{j=k+1}^m g_j(y_t) \right).$$

and, we introduce:

$$\tilde{D}\varphi(t, y) = \sum_{k=0}^{m-1} \mathbb{1}_{[t_k, t_{k+1})}(t) \left(\prod_{j=1}^k g_j(y_{t_j}) \left(\prod_{j=k+1}^m g_j \right)'(y_t) \right) \quad (26)$$

and

$$\tilde{\Delta}\varphi(t, y) = \sum_{k=0}^{m-1} \mathbb{1}_{[t_k, t_{k+1})}(t) \left(\prod_{j=1}^k g_j(y_{t_j}) \Delta \left(\prod_{j=k+1}^m g_j \right)(y_t) \right) \quad (27)$$

With this notation, the next lemma is obtained by a direct adaptation of the results in [17] and can be founded in Appendix:

Lemma 3.2. *Assume that*

$$\sup_{K \in \mathbb{N}^*} \mathbb{E}(\langle H_0^K, 1 \rangle^{2+\varepsilon} + \langle H_0^K(dy), y_0^2 \rangle^{1+\varepsilon}) = \sup_{K \in \mathbb{N}^*} \mathbb{E}(\langle Z_0^K, 1 \rangle^{2+\varepsilon} + \langle Z_0^K, x^2 \rangle^{1+\varepsilon}) < +\infty.$$

For φ defined in (25),

$$\begin{aligned} \langle H_t^K, \varphi \rangle &= \langle H_0^K, \varphi \rangle + \int_0^t \int_{\mathcal{C}(\mathbb{R}_+, \mathbb{R})} \left(\frac{\sigma^2}{2} \tilde{\Delta}\varphi(s, y) - \sigma c \tilde{D}\varphi(s, y) \right. \\ &\quad \left. + \left(1 - \frac{y_s^2}{2} - \langle H_s^K, 1 \rangle \right) \varphi(y) \right) H_s^K(dy) ds + \mathcal{M}_t^{K, \varphi}, \end{aligned} \quad (28)$$

where $\mathcal{M}_t^{K, \varphi}$ is a square integrable martingale with predictable quadratic variation process:

$$\langle \mathcal{M}^{K, \varphi} \rangle_t = \frac{1}{K} \int_0^t \int_{\mathcal{C}(\mathbb{R}_+, \mathbb{R})} \left(\left(1 + \frac{y_s^2}{2} + \langle H_s^K, 1 \rangle \right) \varphi^2(s, y) + \sigma^2 (\tilde{D}\varphi(s, y))^2 \right) H_s^K(dy) ds. \quad (29)$$

We extend here the mild formula (11). For $t > 0$ fixed, for $m \in \mathbb{N}^*$, $0 \leq t_1 < \dots < t_m$ and $\forall j \in \{1, \dots, m\}$, $g_j \in C_b^2(\mathbb{R}, \mathbb{R})$, we define for $0 \leq s < t$ and $y \in \mathcal{C}(\mathbb{R}_+, \mathbb{R})$ a generalized version of the semigroup as

$$\psi^t(s, y) = \sum_{k=0}^{m-1} \mathbb{1}_{[t_k, t_{k+1})}(s) \left(\prod_{j=1}^k g_j(y_{t_j}) S^{t_{k+1} \wedge t} \left(\prod_{j=k+1}^m g_j \right)(s, y_s) \right), \quad (30)$$

where $S^t(g)(s, x) = \mathbb{E}_x(g(\tilde{Y}_{t-s}) \exp(-\int_0^{t-s} \frac{\tilde{Y}_u^2}{2} du))$ and $\tilde{Y}_t = x + \sigma(B_t - ct)$. Note that $\psi^t(t, y) = \varphi(y)$ and that

$$-\frac{y_s^2}{2} \psi^t(s, y) + \partial_s \psi^t(s, y) + \frac{\sigma^2}{2} \tilde{\Delta}_y \psi^t(s, y) - \sigma c \tilde{D}_y \psi^t(s, y) = 0.$$

The next lemma follows from this property and from Appendix A.2 (see (77) and (78)).

Lemma 3.3. *Under the same assumptions as in Lemma 3.2,*

$$\langle H_t^K, \varphi \rangle = \langle H_0^K(dy), \psi^t(0, y) \rangle + \int_0^t \int_{\mathcal{C}(\mathbb{R}_+, \mathbb{R})} (1 - \langle H_s^K, 1 \rangle) \psi^t(s, y) H_s^K(dy) ds + R_{t,t}^K \quad (31)$$

where for $R_{u,t}^K$ is defined for $u \leq t$ by:

$$\begin{aligned} R_{u,t}^K &= \frac{1}{K} \int_0^u \sum_{i \in V_s^K} \sigma \partial_x \psi^t(s, X_{(\cdot, \wedge s)}^i) dB_s^i + \frac{1}{K} \int_0^u \int_{\mathcal{I}} \int_{\mathbb{R}_+} \mathbb{1}_{\{i \in V_{s-}^K\}} \left[\psi^t(s, X_{(\cdot, \wedge s)}^i) \mathbb{1}_{\{\theta \leq 1\}} \right. \\ &\quad \left. - \psi^t(s, X_{(\cdot, \wedge s)}^i) \mathbb{1}_{\{1 < \theta \leq 1 + \frac{(X_s^i)^2}{2} + \langle H_{s-}^K, 1 \rangle\}} \right] \tilde{N}(ds, di, d\theta). \end{aligned} \quad (32)$$

For any $T > 0$, there exists a positive constant C_T such that for all $0 \leq s \leq t \leq T$

$$\mathbb{E}(\sup_{u \leq t} (R_{u,t}^K)^2) < \frac{C_T}{K}. \quad (33)$$

Note that the process $t \mapsto R_{t,t}^K$ is not a local martingale, as seen in the proof.

Proof Using Lemma 2.1 and noticing that $\langle H_s^K, 1 \rangle = \langle Z_s^K, 1 \rangle$ and $\langle H_s^K, y_s^2 \rangle = \langle Z_s^K, x^2 \rangle$, we have for any $T > 0$ that

$$\sup_{K \geq 1} \mathbb{E} \left(\sup_{t \in [0, T]} \langle H_t^K, 1 \rangle^{2+\varepsilon} + \langle H_t^K(dy), y_t^2 \rangle^{1+\varepsilon} \right) < +\infty. \quad (34)$$

Using (77) in appendix, we can write (31)-(32). The process $t \mapsto R_{t,t}^K$ is not a martingale, but the process $u \mapsto R_{u,t}^K$, defined for $u \leq t$, is a martingale. Then we can apply Doob's inequality and write

$$\begin{aligned} &\mathbb{E}(\sup_{u \leq t} (R_{u,t}^K)^2) \\ &\leq \frac{1}{K} \mathbb{E} \left(\int_0^t \int_{\mathcal{C}(\mathbb{R}_+, \mathbb{R})} \left(\left(1 + \frac{y_s^2}{2} + \langle H_s^K, 1 \rangle \right) (\psi^t)^2(s, y) + \sigma^2 (\partial_x \psi^t(s, y))^2 \right) H_s^K(dy) ds \right). \end{aligned}$$

The function φ defining ψ^t being bounded, we can conclude with (34). \square

As in the previous section, we can freeze the nonlinearity in the competition term to λ and couple the historical process H^K with the historical process \tilde{H}^K associated with the process \tilde{Z}^K (this coupling can be done using the same Poisson point measures, Brownian motions and initial condition as Z^K and \tilde{Z}^K , see Appendix A.2).

Proposition 3.4. *Assume that (6) hold and that $Z_0^K \xrightarrow[K \rightarrow \infty]{w} F$. Then for any continuous and bounded function φ of the form (25),*

$$\lim_{K \rightarrow +\infty} \mathbb{E}(\sup_{t \leq T} |\langle H_t^K, \varphi \rangle - \langle \tilde{H}_t^K, \varphi \rangle|^2) = 0.$$

Proof Using Appendix A.2, we have for \tilde{H}^K a similar decomposition as (31) with $\langle H_t^K, 1 \rangle$ replaced by λ . We now use the mild equation (31) for H^K and the analogous equation for \tilde{H}^K , involving a term $\tilde{R}_{t,t}^K$ similar to (32) but with again $\langle H_t^K, 1 \rangle$ replaced by λ . Recall that both processes are built on the same probability space with the same initial condition. For $M > 0$, let us introduce the stopping times:

$$\begin{aligned} \tau_M^K &= \inf \{t \in \mathbb{R}_+, \langle H_t^K, 1 \rangle = \langle Z_t^K, 1 \rangle > M\}, \\ \text{and } \tilde{\tau}_M^K &= \inf \{t \in \mathbb{R}_+, \langle \tilde{H}_t^K, 1 \rangle = \langle \tilde{Z}_t^K, 1 \rangle > M\}. \end{aligned} \quad (35)$$

Let $\varepsilon > 0$ be fixed. Because the processes $(\langle Z_t^K, 1 \rangle)_{t \in \mathbb{R}_+}$ and $(\langle \tilde{Z}_t^K, 1 \rangle)_{t \in \mathbb{R}_+}$ converge to deterministic continuous and bounded processes, there exists $M = M(\varepsilon)$, independent from K , such that

$$\mathbb{P}(\tau_M^K \wedge \tilde{\tau}_M^K \leq T) < \varepsilon.$$

Then, using (31) for a bounded cylindrical test-function ψ^t as in (25):

$$\begin{aligned} & \left| \langle H_{t \wedge \tau_M^K \wedge \tilde{\tau}_M^K}^K, \varphi \rangle - \langle \tilde{H}_{t \wedge \tau_M^K \wedge \tilde{\tau}_M^K}^K, \varphi \rangle \right| \\ \leq & \left| \int_0^{t \wedge \tau_M^K \wedge \tilde{\tau}_M^K} \int_{\mathcal{C}(\mathbb{R}_+, \mathbb{R})} (1 - \langle H_s^K, 1 \rangle) \psi^t(s, y) H_s^K(dy) ds \right. \\ & \left. - \int_0^{t \wedge \tau_M^K \wedge \tilde{\tau}_M^K} \int_{\mathcal{C}(\mathbb{R}_+, \mathbb{R})} (1 - \lambda) \psi^t(s, y) \tilde{H}_s^K(dy) ds \right| + |R_{t \wedge \tau_M^K \wedge \tilde{\tau}_M^K, t}^K| + |\tilde{R}_{t \wedge \tau_M^K \wedge \tilde{\tau}_M^K, t}^K| \\ \leq & |1 - \lambda| \int_0^t \sup_{u \leq s \wedge \tau_M^K \wedge \tilde{\tau}_M^K} |\langle H_u^K(dy) - \tilde{H}_u^K(dy), \psi^t(u, y) \rangle| ds \\ & + \int_0^{t \wedge \tau_M^K \wedge \tilde{\tau}_M^K} |\langle Z_s^K, 1 \rangle - \lambda| \times |\langle H_s^K(dy), \psi^t(s, y) \rangle| ds + \sup_{s \leq t} |R_{s \wedge \tau_M^K \wedge \tilde{\tau}_M^K, t}^K| + \sup_{s \leq t} |\tilde{R}_{s \wedge \tau_M^K \wedge \tilde{\tau}_M^K, t}^K| \\ \leq & |1 - \lambda| \|\varphi\|_\infty \int_0^t \sup_{u \leq s} \|H_{u \wedge \tau_M^K \wedge \tilde{\tau}_M^K}^K - \tilde{H}_{u \wedge \tau_M^K \wedge \tilde{\tau}_M^K}^K\|_{TV} ds \\ & + T \|\varphi\|_\infty M \sup_{s \leq T} |\langle Z_s^K, 1 \rangle - \lambda| + \sup_{s \leq t} |R_{s \wedge \tau_M^K \wedge \tilde{\tau}_M^K, t}^K| + \sup_{s \leq t} |\tilde{R}_{s \wedge \tau_M^K \wedge \tilde{\tau}_M^K, t}^K|, \end{aligned}$$

where $\|\cdot\|_{TV}$ denotes the norm in total variation. Taking the supremum with respect to φ of the form (25) and with norm $\|\varphi\|_\infty \leq 1$ in the left hand side, and then taking the expectation, we have by (33):

$$\begin{aligned} \mathbb{E} \left(\sup_{s \leq t} \|H_{s \wedge \tau_M^K \wedge \tilde{\tau}_M^K}^K - \tilde{H}_{s \wedge \tau_M^K \wedge \tilde{\tau}_M^K}^K\|_{TV} \right) & \leq |1 - \lambda| \int_0^t \mathbb{E} \left(\sup_{u \leq s} \|H_{u \wedge \tau_M^K \wedge \tilde{\tau}_M^K}^K - \tilde{H}_{u \wedge \tau_M^K \wedge \tilde{\tau}_M^K}^K\|_{TV} \right) ds \\ & + TM \mathbb{E} \left(\sup_{s \leq T} |\langle Z_s^K, 1 \rangle - \lambda| \right) + 2 \sqrt{\frac{C_T}{K}}. \end{aligned}$$

Using Gronwall's lemma:

$$\mathbb{E}\left(\sup_{s \leq t} \|H_{s \wedge \tau_M^K \wedge \tilde{\tau}_M^K}^K - \tilde{H}_{s \wedge \tau_M^K \wedge \tilde{\tau}_M^K}^K\|_{TV}\right) \leq \left(TM \mathbb{E}\left(\sup_{s \leq T} |\langle Z_s^K, 1 \rangle - \lambda|\right) + 2\sqrt{\frac{C_T}{K}}\right)e^{1-\lambda|T}.$$

Then, note that

$$\begin{aligned} & \mathbb{E}\left(\sup_{t \leq T} \|H_t^K - \tilde{H}_t^K\|_{TV}\right) \leq \mathbb{E}\left(\sup_{t \leq T} \|H_{t \wedge \tau_M^K \wedge \tilde{\tau}_M^K}^K - \tilde{H}_{t \wedge \tau_M^K \wedge \tilde{\tau}_M^K}^K\|_{TV}\right) \\ & + \sqrt{\mathbb{P}(\tau_M^K \wedge \tilde{\tau}_M^K \leq T)} \sqrt{2\mathbb{E}\left(\sup_{t \leq T} \langle Z_t^K, 1 \rangle^2 + \langle \tilde{Z}_t^K, 1 \rangle^2\right)} \\ & \leq \left(TM \mathbb{E}\left(\sup_{t \leq T} |\langle Z_t^K, 1 \rangle - \lambda|\right) + 2\sqrt{\frac{C_T}{K}}\right)e^{1-\lambda|T} + \sqrt{\varepsilon} \sqrt{2 \sup_{K \geq 1} \mathbb{E}\left(\sup_{t \leq T} \langle Z_t^K, 1 \rangle^2 + \langle \tilde{Z}_t^K, 1 \rangle^2\right)}, \end{aligned} \quad (36)$$

for $M > M_\varepsilon$. Then, choosing K sufficiently large, the first term in the right hand side is upper bounded by a constant times ε , by (23). This, with (7), concludes the proof. \square

3.3 Feynman-Kac approach for the law of the branching-diffusion process

As the process $(\tilde{Z}_t^K, t \leq T)$ is a branching process without interaction, the genealogies started from the initial individuals evolve independently from each other, with the same law. It follows that

$$\mathbb{E}\left[\langle \tilde{Z}_t^K, \varphi \rangle\right] = \int_{\mathbb{R}} \mathbb{E}_{\delta_x}\left[\langle \tilde{Z}_t, \varphi \rangle\right] Z_0^K(dx).$$

where \tilde{Z} is a branching process satisfying Equation (20) started from $\tilde{Z}_0 = \delta_x$. For the reasons mentioned above, we consider from this point a particle system starting from a single particle with trait x .

The formulas/theory used below come from [53, 59] further developed for instance in [3, 18, 41, 61]. Here we give original and simpler proofs.

Lemma 3.5. *Let φ in $C_b(\mathbb{R})$. Then, for any positive time t , for any $x \in \mathbb{R}$, we have*

$$\mathbb{E}_{\delta_x}\left[\langle \tilde{Z}_t, \varphi \rangle\right] = \mathbb{E}_x\left[\exp\left(\int_0^t \left(1 - \frac{1}{2}X_s^2 - \lambda\right) ds\right) \varphi(X_t)\right], \quad (37)$$

where X is the drifted Brownian motion

$$dX_t = \sigma(dB_t - c dt). \quad (38)$$

Proof Let us give a very simple proof based on Itô's formula.

Let us first note that the measure $\nu_t(dy) = \mathbb{E}_{\delta_x}\left(\tilde{Z}_t(dy)\right)$ defined for any φ in $C_b(\mathbb{R})$ by

$$\langle \nu_t, \varphi \rangle = \mathbb{E}_{\delta_x} \left[\langle \tilde{Z}_t, \varphi \rangle \right]$$

is the unique weak solution of

$$\begin{cases} \partial_t \nu_t = \Delta \nu_t + \sigma c \partial_x \nu_t + \left(1 - \frac{x^2}{2} - \lambda\right) \nu_t \\ \nu_0 = \delta_x \end{cases} . \quad (39)$$

Indeed, it is enough to take expectation in (20). Uniqueness of such a solution has been proved in Theorem 2.2.

Let us now show that the right hand side term of (37) also satisfies (39). Uniqueness will yield the result.

Let φ in $C_b^2(\mathbb{R})$ and apply Itô's formula to the semimartingale

$$\exp \left(\int_0^t \left(1 - \frac{1}{2} X_s^2 - \lambda\right) ds \right) \varphi(X_t).$$

We have

$$\begin{aligned} & \exp \left(\int_0^t \left(1 - \frac{1}{2} X_s^2 - \lambda\right) ds \right) \varphi(X_t) = \varphi(X_0) + \int_0^t \exp \left(\int_0^s \left(1 - \frac{1}{2} X_u^2 - \lambda\right) du \right) \sigma \varphi'(X_s) dB_s \\ & + \int_0^t \exp \left(\int_0^s \left(1 - \frac{1}{2} X_u^2 - \lambda\right) du \right) \left\{ \left(1 - \frac{1}{2} X_s^2 - \lambda\right) \varphi(X_s) + \frac{\sigma^2}{2} \varphi''(X_s) - c \sigma \varphi'(X_s) \right\} ds. \end{aligned}$$

Taking the expectation, we obtain that

$$\begin{aligned} \mathbb{E}_x \left[\exp \left(\int_0^t \left(1 - \frac{1}{2} X_s^2 - \lambda\right) ds \right) \varphi(X_t) \right] &= \varphi(x) + \mathbb{E}_x \left[\int_0^t \exp \left(\int_0^s \left(1 - \frac{1}{2} X_u^2 - \lambda\right) du \right) \right. \\ & \left. \times \left\{ \left(1 - \frac{1}{2} X_s^2 - \lambda\right) \varphi(X_s) + \frac{\sigma^2}{2} \varphi''(X_s) - c \sigma \varphi'(X_s) \right\} ds \right]. \quad (40) \end{aligned}$$

If we define the measure μ_t for any test function $\varphi \in C_b^2(\mathbb{R})$ by

$$\langle \mu_t, \varphi \rangle = \mathbb{E}_x \left[\exp \left(\int_0^t \left(1 - \frac{1}{2} X_s^2 - \lambda\right) ds \right) \varphi(X_t) \right],$$

we obtain from (40) that

$$\langle \mu_t, \varphi \rangle = \langle \delta_x, \varphi \rangle + \int_0^t \left\langle \mu_s, \left(1 - \frac{x^2}{2} - \lambda\right) \varphi(x) + \frac{\sigma^2}{2} \varphi''(x) - c \sigma \varphi'(x) \right\rangle ds.$$

This proves that the flow $(\mu_t, t \geq 0)$ is a weak solution of (39) and the conclusion follows by uniqueness. \square

Corollary 3.6. *Let us define for any $t \geq 0$ and $x \in \mathbb{R}$ the expectation of the number of individuals at time t in the branching process \tilde{Z}_t started from one individual with trait x ,*

$$m_t(x) = \mathbb{E}_{\delta_x}(\langle \tilde{Z}_t, 1 \rangle). \quad (41)$$

Then we have

$$m_t(x) = \mathbb{E}_x \left[\exp \left(\int_0^t \left(1 - \frac{1}{2} X_s^2 - \lambda \right) ds \right) \right]. \quad (42)$$

We deduce that the function $(t, x) \mapsto m_t(x)$ belongs to $C_b^{1,\infty}([0, T] \times \mathbb{R})$.

Proof Equation (42) is obvious by applying Lemma 3.5 to $\varphi = 1$.

Since the process X is a drifted Brownian motion, we can write

$$m_t(x) = \mathbb{E}_0 \left[\exp \left(\int_0^t \left(1 - \frac{1}{2} (x + X_s)^2 - \lambda \right) ds \right) \right].$$

Lebesgue's Theorem allows us to conclude. \square

Remark: From (42) and Feynman-Kac formula (see [31, 47, 48]), we deduce that the function $(m_t(x), x \in \mathbb{R}, t \geq 0)$ is the unique strong solution of

$$\begin{cases} \partial_t m = \frac{\sigma^2}{2} \partial_{xx} m - \sigma c \partial_x m + \left(1 - \frac{x^2}{2} - \lambda \right) m \\ m_0(x) = 1. \end{cases} \quad (43)$$

Let us also note that (43) and the stationarity of F (see Eq. (18)) imply that $t \rightarrow \int m_t(x) F(x) dx$ is constant and then

$$\int_{\mathbb{R}} m_t(x) F(x) dx = \int_{\mathbb{R}} F(x) dx = \lambda. \quad (44)$$

Our aim is now to generalize (37) to trajectories. In what follows, for a time T , we label individuals by $i \in \tilde{V}_T$ where \tilde{V}_T denotes the set of individuals alive at time T and started from one individual with trait x at time 0. For a time $t < T$, we will introduce the notation $\tilde{X}^i(t)$ to denote the historical lineage of the individual $i \in \tilde{V}_T$ at time t .

Lemma 3.7. *Let φ in $C_b(\mathbb{R})$. Then, for any positive times t and T such that $t \leq T$, for any $x \in \mathbb{R}$, we have*

$$\mathbb{E}_{\delta_x} \left[\sum_{i \in \tilde{V}_T} \varphi(\tilde{X}_t^i) \right] = \mathbb{E}_x \left[\exp \left(\int_0^T \left(1 - \frac{1}{2} X_s^2 - \lambda \right) ds \right) \varphi(X_t) \right], \quad (45)$$

where X is the drifted Brownian defined in Lemma 3.5.

Proof The case where $t = T$ results from Lemma 3.5. To obtain formula (45) from (37) when $t < T$, one can proceed as follows. For every individual i alive at time T , and for every $t \leq T$, there exists a unique $j \in \mathcal{I}$ such that (j, t) belongs to the ancestral path of i . Thus, we have

$$\sum_{i \in \tilde{V}_T} \varphi(\tilde{X}_t^i) = \sum_{i \in \tilde{V}_T} \sum_{j \in \tilde{V}_t} \mathbb{1}_{(j,t) \preceq (i,T)} \varphi(\tilde{X}_t^j),$$

where $\tilde{X}_t^i = \tilde{X}_t^j$ since $(j, t) \preceq (i, T)$. Thus,

$$\sum_{i \in \tilde{V}_T} \sum_{j \in \tilde{V}_t} \mathbb{1}_{j \preceq i} \varphi(\tilde{X}_t^i) = \sum_{j \in \tilde{V}_t} \varphi(\tilde{X}_t^j) \sum_{i \in \tilde{V}_T} \mathbb{1}_{j \preceq i} = \sum_{j \in \tilde{V}_t} \varphi(\tilde{X}_t^j) \tilde{N}(j)_{t,T},$$

where $\tilde{N}(j)_{t,T}$ denotes the number of descendants at time T of an individual j alive at time t (with the convention that $\tilde{N}(j)_{t,T} = 0$ if $j \notin \tilde{V}_t$, i.e. if j does not exist at time t). Thus, denoting by $(\mathcal{F}_t, t \in \mathbb{R}_+)$ the natural filtration associated with \tilde{Z} , we have that

$$\begin{aligned} \mathbb{E} \left[\sum_{i \in \tilde{V}_T} \varphi(\tilde{X}_t^i) \middle| \mathcal{F}_t \right] &= \sum_{j \in \tilde{V}_t} \varphi(\tilde{X}_t^j) \mathbb{E} \left[\tilde{N}(j)_{t,T} \middle| \mathcal{F}_t \right] = \sum_{j \in \tilde{V}_t} \varphi(\tilde{X}_t^j) m_{T-t}(\tilde{X}_t^j) \\ &= \langle \tilde{Z}_t, \varphi m_{T-t} \rangle, \end{aligned} \quad (46)$$

where m_t has been defined in Corollary 3.6.

We now apply Lemma 3.5 to (46) for the function $x \mapsto \varphi(x) m_{T-t}(x)$. That gives

$$\mathbb{E}_{\delta_x} \left[\langle \tilde{Z}_t, \varphi m_{T-t} \rangle \right] = \mathbb{E}_x \left[\exp \left(\int_0^t \left(1 - \frac{1}{2} X_s^2 - \lambda \right) ds \right) \varphi(X_t) m_{T-t}(X_t) \right].$$

Then, from the expression of m_t given in (42) and one obtains by the Markov property that

$$\mathbb{E}_{\delta_x} \left[\langle \tilde{Z}_t, \varphi m_{T-t} \rangle \right] = \mathbb{E}_x \left[\exp \left(\int_0^T \left(1 - \frac{1}{2} X_s^2 - \lambda \right) ds \right) \varphi(X_t) \right].$$

That concludes the proof. \square

We are now interested in trajectorial extension of the previous formulae.

Proposition 3.8. *Let φ in $\mathcal{C}_b(\mathbb{R}^n, \mathbb{R})$. Then, for any positive times $t_1 < t_2 < \dots < t_n < T$, for any $x \in \mathbb{R}$, we have*

$$\mathbb{E}_{\delta_x} \left[\sum_{i \in \tilde{V}_T} \varphi(\tilde{X}_{t_1}^i, \dots, \tilde{X}_{t_n}^i) \right] = \mathbb{E}_x \left[\exp \left(\int_0^T \left(1 - \frac{1}{2} X_s^2 - \lambda \right) ds \right) \varphi(X_{t_1}, \dots, X_{t_n}) \right], \quad (47)$$

where X is the drifted Brownian defined in Lemma 3.5.

Proof By usual arguments, it is enough to prove the result for product functions

$$\varphi(x_1, \dots, x_n) = \prod_{i=1}^n \varphi_i(x_i).$$

The proof of this result follows the same lines as the proof of Lemma 3.7 conditioning first by \mathcal{F}_{t_1} , and then by \mathcal{F}_{t_2} and so on. We leave the remaining of the proof to the reader. \square

Let us recall that \tilde{H} is the historical process associated with \tilde{Z} .

Lemma 3.9. *We have that for $T > 0$, $\Phi : C([0, T], \mathbb{R}) \rightarrow \mathbb{R}$ a continuous and bounded function and $x \in \mathbb{R}$:*

$$\begin{aligned} \mathbb{E}_{\delta_x} \left[\langle \tilde{H}_T, \Phi \rangle \right] &= \mathbb{E}_{\delta_x} \left[\sum_{i \in \tilde{V}_T} \Phi(\tilde{X}_s^i, s \leq T) \right] \\ &= \mathbb{E}_x \left[\exp \left(\int_0^T \left(1 - \frac{1}{2} X_s^2 - \lambda \right) ds \right) \Phi(X_s, s \leq T) \right], \end{aligned} \quad (48)$$

where X is the drifted Brownian motion defined in Lemma 3.5.

Proof Let us consider the linear interpolation $I_n : \mathbb{R}^n \rightarrow C([0, T])$ such that, for all $j \in \{0, \dots, n-1\}$, for all $t \in [jT/n, (j+1)T/n)$ and $(u_1, \dots, u_n) \in \mathbb{R}^{n+1}$,

$$I_n(u_0, \dots, u_n)(t) = (u_{j+1} - u_j) \frac{n}{T} \left(t - \frac{j}{n} T \right) + u_j.$$

Thus, we have for $x \in C([0, T], \mathbb{R})$,

$$\|I_n(x(0), \dots, x(jT/n), \dots, x(T)) - x\|_\infty \leq 2 \sum_{j=0}^n \mathbf{1}_{t \in [jT/n, (j+1)T/n)} \omega(x, T/n) = 2 \omega(x, T/n),$$

where $\omega(x, \cdot)$ is the modulus of continuity of x . Thus, the functions $I_n(x(0), \dots, x(jT/n), \dots, x(T))$ converge uniformly on $[0, T]$ to x , as n tends to infinity. The result then follows from Lebesgue's theorem, Lemma 3.8 and the continuity of x . \square

3.4 Computation of $m_t(x)$

In this Brownian framework, one can explicitly compute $m_t(x)$ from (42) by a method adapted from Fitzsimmons Pitman and Yor [32].

Proposition 3.10. *For any $x \in \mathbb{R}$ and $t \in [0, T]$, we have*

$$m_t(x) = \sqrt{1 + \tanh(\sigma t)} \exp \left(- \frac{(x + e^{-\sigma t} c)^2}{2\sigma} (1 + \tanh(\sigma t)) + \frac{(x + c)^2}{2\sigma} \right). \quad (49)$$

Proof Recall that $X_t = \sigma B_t - c\sigma t$ (see (38)). By notational simplicity we assume that the Brownian motion B starts from x , then $X_0 = \sigma x$ and we will compute $m_t(\sigma x)$. Equation (42) gives

$$\begin{aligned} m_t(\sigma x) &= \mathbb{E}_{\delta_{\sigma x}} [\langle \tilde{Z}_t, 1 \rangle] = \mathbb{E}_x \left[\exp \left(\int_0^t \left(1 - \frac{\sigma^2}{2} (B_s - cs)^2 - \lambda \right) ds \right) \right] \\ &= e^{(1-\lambda)t} \mathbb{E}_x \left[\exp \left(-\frac{\sigma^2}{2} \int_0^t (B_s - cs)^2 ds \right) \right] =: e^{(1-\lambda)t} I. \end{aligned} \quad (50)$$

Let us compute explicitly I , the expectation appearing in the right hand side of (50). Recall that our probability space is endowed with the probability measure \mathbb{P} . Let $(\mathcal{F}_t^B)_{t \geq 0}$ be the filtration of the Brownian motion B and define the new probability \mathbb{Q} by

$$\frac{d\mathbb{Q}}{d\mathbb{P}} \Big| \mathcal{F}_t^B = \exp(cB_t - \frac{c^2}{2}t - cx)$$

using Girsanov theorem to kill the drift. Under \mathbb{Q} , $W_t = B_t - ct$ is a Brownian motion. Hence,

$$\begin{aligned} I &= \mathbb{E}_x^{\mathbb{Q}} \left[\exp \left(-\frac{\sigma^2}{2} \int_0^t (B_s - cs)^2 ds \right) \exp(-cB_t + \frac{c^2}{2}t + cx) \right] \\ &= \mathbb{E}_x^{\mathbb{Q}} \left[\exp \left(-\frac{\sigma^2}{2} \int_0^t W_s^2 ds \right) \exp(-cW_t - \frac{c^2}{2}t + cx) \right] \\ &= e^{cx - \frac{c^2}{2}t} \mathbb{E}_x^{\mathbb{Q}} \left[\exp \left(-\frac{\sigma^2}{2} \int_0^t W_s^2 ds - cW_t \right) \right]. \end{aligned} \quad (51)$$

Now, we want to compute the expectation in this last term. We use that $M_t = \frac{\sigma}{2}(W_t^2 - t) - \frac{\sigma x^2}{2} = \sigma \int_0^t W_s dW_s$ is a martingale with

$$\langle M \rangle_t = \sigma^2 \int_0^t W_s^2 ds.$$

Let $(\mathcal{F}_t^W)_{t \geq 0}$ be the filtration of W and set:

$$\frac{d\mathbb{Q}'}{d\mathbb{Q}} \Big| \mathcal{F}_t^W = \exp \left(\frac{1}{2} \sigma W_t^2 - \frac{1}{2} \sigma t - \frac{\sigma^2}{2} \int_0^t W_s^2 ds \right) e^{-\frac{\sigma x^2}{2}}.$$

We have

$$\mathbb{E}_x^{\mathbb{Q}} \left[\exp \left(-\frac{\sigma^2}{2} \int_0^t W_s^2 ds - cW_t \right) \right] = \mathbb{E}_x^{\mathbb{Q}'} \left[\exp \left(-\frac{1}{2} \sigma W_t^2 - cW_t \right) \right] e^{\frac{\sigma x^2}{2} + \frac{\sigma t}{2}}. \quad (52)$$

On the other hand, we have that under \mathbb{Q}' ,

$$W'_t = W_t - \sigma \int_0^t W_s ds$$

is a \mathbb{Q}' -Brownian motion, and

$$W_t = \sigma \int_0^t W_s ds + W'_t$$

is an Ornstein-Uhlenbeck process. Hence, for a given $t \geq 0$, W_t under \mathbb{Q}' is distributed as

$$\mathcal{N}\left(e^{\sigma t}x, \frac{1}{2\sigma}(e^{2\sigma t} - 1)\right).$$

This allows us to compute (52). For Y a standard Gaussian random variable $\mathcal{N}(0, 1)$,

$$\Psi(u, v) := \mathbb{E}\left[e^{uY^2 + vY}\right] = \frac{\exp\left(\frac{v^2}{2(1-2u)}\right)}{\sqrt{1-2u}},$$

when $u < \frac{1}{2}$. For Y any $\mathcal{N}(m, \delta^2)$ random variable,

$$\mathbb{E}\left[e^{aY^2 + bY}\right] = e^{(bm + m^2a)}\Psi(a\delta^2, \delta(2am + b)),$$

when $a\delta^2 < 1/2$.

Here we have $m = e^{\sigma t}x$, $\delta^2 = (e^{2\sigma t} - 1)/(2\sigma)$, $a = -\sigma/2$ and $b = -c$. Thus,

$$\begin{cases} a\delta^2 = -\frac{1}{4}(e^{2\sigma t} - 1) \\ \delta(2am + b) = \sqrt{\frac{(e^{2\sigma t} - 1)}{2\sigma}}(-\sigma x e^{\sigma t} - c), \end{cases}$$

and applying the above computation yields for the expectation in the r.h.s. of (52):

$$\begin{aligned} \mathbb{E}_x^{\mathbb{Q}'} \left[\exp\left(-\frac{1}{2}\sigma W_t^2 - cW_t\right) \right] &= e^{(-cxe^{\sigma t} - \frac{\sigma}{2}x^2e^{2\sigma t})} \frac{1}{\sqrt{\frac{1}{2} + \frac{1}{2}e^{2\sigma t}}} \exp\left(\frac{\frac{(e^{2\sigma t} - 1)}{2\sigma}(\sigma x e^{\sigma t} + c)^2}{1 + e^{2\sigma t}}\right) \\ &= \sqrt{\frac{2}{1 + e^{2\sigma t}}} e^{-\frac{1}{2\sigma}(\sigma x e^{\sigma t} + c)^2} e^{\frac{c^2}{2\sigma}} \exp\left(\frac{(\sigma x e^{\sigma t} + c)^2}{2\sigma} \tanh(\sigma t)\right) \\ &= \sqrt{\frac{2}{1 + e^{2\sigma t}}} e^{\frac{c^2}{2\sigma}} \exp\left(-\frac{(\sigma x e^{\sigma t} + c)^2}{2\sigma}(1 - \tanh(\sigma t))\right) \quad (53) \end{aligned}$$

where the second line has been obtained by using that

$$\exp\left(-\frac{1}{2\sigma}(\sigma^2 x^2 e^{2\sigma t} + 2cx\sigma e^{\sigma t} + c^2 - c^2)\right) = \exp\left(-\frac{1}{2\sigma}((\sigma x e^{\sigma t} + c)^2 - c^2)\right),$$

and that

$$\frac{e^{2\sigma t} - 1}{e^{2\sigma t} + 1} = \tanh(\sigma t).$$

Gathering (53) with (51) and (52), we obtain that:

$$\begin{aligned} I &= e^{cx - \frac{c^2}{2}t} e^{\frac{\sigma x^2}{2} + \frac{\sigma t}{2}} e^{\frac{c^2}{2\sigma}} \sqrt{\frac{2}{1 + e^{2\sigma t}}} \exp\left(-\frac{(\sigma x e^{\sigma t} + c)^2}{2\sigma} (1 - \tanh(\sigma t))\right) \\ &= e^{-\frac{c^2 t}{2} + \frac{\sigma t}{2}} \sqrt{\frac{2}{1 + e^{2\sigma t}}} \exp\left(\frac{(\sigma x + c)^2}{2\sigma} - \frac{(\sigma x e^{\sigma t} + c)^2}{2\sigma} (1 - \tanh(\sigma t))\right). \end{aligned}$$

Plugging this result into (50) gives:

$$\begin{aligned} m_t(\sigma x) &= e^{(1 - \frac{1}{2}c^2 + \frac{\sigma}{2} - \lambda)t} \frac{\sqrt{2}}{\sqrt{1 + e^{2\sigma t}}} \exp\left(\frac{(\sigma x + c)^2}{2\sigma} - \frac{(\sigma x e^{\sigma t} + c)^2}{2\sigma} (1 - \tanh(\sigma t))\right) \\ &= e^{(1 - \frac{1}{2}c^2 - \frac{\sigma}{2} - \lambda)t} \frac{\sqrt{2}}{\sqrt{1 + e^{-2\sigma t}}} \exp\left(\frac{(\sigma x + c)^2}{2\sigma} - \frac{(\sigma x e^{\sigma t} + c)^2}{2\sigma} (1 - \tanh(\sigma t))\right) \\ &= \sqrt{1 + \tanh(\sigma t)} \exp\left(\frac{(\sigma x + c)^2}{2\sigma} - \frac{(\sigma x e^{\sigma t} + c)^2}{2\sigma} (1 - \tanh(\sigma t))\right), \end{aligned}$$

where we use (17) and $2/(1 + e^{-2x}) = 1 + \tanh(x)$ for the third equality. Replacing in the above expression x with x/σ yields the announced expression for $m_t(x)$. \square

4 A spinal approach - The typical trajectory

4.1 The spinal process

Recall that the stochastic population process is assumed starting from the stationary distribution F . We want to characterize the behavior of the ancestral path of an individual uniformly sampled at time T . To this aim, the spinal approach [40, 41, 3] consists in considering the trajectory of a “typical” individual in the population whose behavior summarizes the behavior of the entire population. The next theorem, will allow to describe the trait process along the spine and can be found in [61] in a more general context. To make the paper easy to read, a proof in our context is given in Appendix B.

Theorem 4.1. *Recall that $m_t(x) = \mathbb{E}_x(\tilde{N}_t)$ has been defined in (42). For $T > 0$, $x \in \mathbb{R}$ and Φ a continuous bounded function on $\mathcal{C}([0, T], \mathbb{R})$, we have*

$$\mathbb{E}_{\delta_x} \left[\sum_{i \in \tilde{V}_T} \Phi(\tilde{X}_s^i, s \leq T) \right] = m_T(x) \mathbb{E}_x [\Phi(Y_s, s \leq T)], \quad (54)$$

where Y is an inhomogeneous Markov process (depending on t) and with infinitesimal generator at time t given for $\phi \in C_b^2(\mathbb{R})$ by

$$\mathcal{G}_t \phi(x) = \frac{L(m_{T-t} \phi)(x) - \phi(x) L m_{T-t}(x)}{m_{T-t}(x)}, \quad (55)$$

where L is the infinitesimal generator of the process $(X_t, t \geq 0)$ associated to \tilde{Z} and defined in (38).

Note that the law of the spinal process is biased by the population size at each time, described by the function m , which makes the process inhomogeneous. This highlights the form of the generator given in (54).

The next proposition allows to relate (54) to the distribution of an individual chosen uniformly at random among the population alive at time t (i.e. in the empirical distribution) when the population is large.

Proposition 4.2. *For any Φ continuous and bounded,*

$$\lim_{K \rightarrow +\infty} \mathbb{E} \left[\frac{1}{\tilde{N}_T^K} \sum_{i \in \tilde{V}_T^K} \Phi(\tilde{X}_s^i, s \leq T) \mid \tilde{X}_0^i = x, \forall i \in \tilde{V}_0^K \right] = \frac{\mathbb{E}_{\delta_x} \left[\sum_{i \in \tilde{V}_T} \Phi(\tilde{X}_s^i, s \leq T) \right]}{m_T(x)},$$

where $\tilde{N}_T^K = K \langle \tilde{Z}_T^K, 1 \rangle$ and \tilde{V}_T^K is the set of individuals alive at time T .

Proof By the branching property, the trees started from each of the K individuals with trait x are independent with same law. Then by the law of large numbers, the two sequences $(\frac{\tilde{N}_T^K}{K})_K$ and $(\frac{\sum_{i \in \tilde{V}_T^K} \Phi(\tilde{X}_s^i, s \leq T)}{K})_K$ converge almost surely respectively to $m_T(x)$ and $\mathbb{E}_{\delta_x} \left[\sum_{i \in \tilde{V}_T} \Phi(\tilde{X}_s^i, s \leq T) \right]$. The result follows. \square

Notice that a corollary of Theorem 4.1 and Proposition 4.2 is that:

Corollary 4.3. *When (Z_0^K) satisfies (6) and converges weakly and in probability to the measure ξ_0 when $K \rightarrow +\infty$,*

$$\begin{aligned} \lim_{K \rightarrow +\infty} \mathbb{E}_{Z_0^K} \left[\frac{\langle \tilde{H}_T^K, \Phi \rangle}{\langle \tilde{H}_T^K, 1 \rangle} \right] &= \lim_{K \rightarrow +\infty} \mathbb{E} \left[\frac{1}{\tilde{N}_T^K} \sum_{i \in \tilde{V}_T^K} \Phi(\tilde{X}_s^i, s \leq T) \mid \sum_{i \in \tilde{V}_0^K} \delta_{\tilde{X}_0^i} = K Z_0^K \right] \\ &= \frac{\int_{\mathbb{R}} m_T(x) \mathbb{E}_x [\Phi(Y_s, s \leq T)] \xi_0(dx)}{\int_{\mathbb{R}} m_T(x) \xi_0(dx)}, \end{aligned} \quad (56)$$

where Y is the process with generator (55).

The explicit computation of $m_t(x)$ yields the generator of Y :

Proposition 4.4. *The generator of the spine Y describing in forward time the path of particle chosen at random in \tilde{V}_T^K is given for $x \in \mathbb{R}$ and $0 \leq t \leq T$ by*

$$\mathcal{G}_t f(x) = \frac{\sigma^2}{2} f''(x) - \sigma x \tanh(\sigma(T-t)) f'(x) - \frac{\sigma c}{\cosh(\sigma(T-t))} f'(x). \quad (57)$$

Proof We have

$$\begin{aligned} m_{T-t}(x)\mathcal{G}_t f(x) &= \frac{\sigma^2}{2}(m_{T-t}f)''(x) - c\sigma(m_{T-t}f)'(x) - f\left(\frac{\sigma^2}{2}m_{T-t}'' - c\sigma m_{T-t}'\right)(x) \\ &= \frac{\sigma^2}{2}(2m_{T-t}'(x)f'(x) + m_{T-t}(x)f''(x)) - c\sigma m_{T-t}(x)f'(x). \end{aligned}$$

Since, by a derivation of (49),

$$\begin{aligned} \partial_x m_t(x) &= \frac{1}{\sigma} \left(-(x + ce^{-\sigma t})(1 + \tanh(\sigma t)) + x + c \right) m_t(x) \\ &= \frac{1}{\sigma} \left(-x \tanh(\sigma t) - c(e^{-\sigma t}(1 + \tanh(\sigma t)) - 1) \right) m_t(x), \end{aligned}$$

we obtain that for $x \in \mathbb{R}$ and $0 \leq t \leq T$,

$$\begin{aligned} \mathcal{G}f(x) &= \frac{\sigma^2}{2}f''(x) + \sigma \left(-x \tanh(\sigma(T-t)) - c \left(\frac{1}{\cosh(\sigma(T-t))} - 1 \right) \right) f'(x) - \sigma c f'(x) \\ &= \frac{\sigma^2}{2}f''(x) - \sigma x \tanh(\sigma(T-t))f'(x) - \sigma c \frac{1}{\cosh(\sigma(T-t))}f'(x). \end{aligned}$$

□

Let us highlight that

$$\begin{cases} \mathcal{G}f \simeq \frac{\sigma^2}{2}f'' - \sigma x f', & \text{for } |T-t| \gg \frac{1}{\sigma} \\ \mathcal{G}f \simeq \frac{\sigma^2}{2}f'' - \sigma c f', & \text{for } |T-t| \ll \frac{1}{\sigma}. \end{cases}$$

We have two regimes depending on the distance between t and the final observation time T . For a small t and large T , the generator is close to the one of an Ornstein-Uhlenbeck process fluctuating around 0 and for t close to T , the generator is close to the one of the drifted Brownian motion Y driving the population to the neighborhood of $-c$, as observed in the simulations.

At this point, we can give the law of the history of an uniformly sampled individual when the initial condition was F . From the explicit value of the generator given in (57), one can deduce that there exists a Brownian motion $(B_t)_t$ independent of Y_0 such that for $0 \leq t \leq T$, the process Y satisfies the stochastic differential equation:

$$dY_t = -\sigma \tanh(\sigma(T-t))Y_t dt - \frac{\sigma c}{\cosh(\sigma(T-t))}dt + \sigma dB_t. \quad (58)$$

Proposition 4.5. *The Markov process Y with generator (57) is a Gaussian process which can be expressed explicitly for $0 \leq t \leq T$:*

$$\begin{aligned} Y_t &= \frac{\cosh(\sigma(T-t))}{\cosh(\sigma T)}Y_0 + c \cosh(\sigma(T-t))(\tanh(\sigma(T-t)) - \tanh(\sigma T)) \\ &\quad + \sigma \cosh(\sigma(T-t)) \int_0^t \frac{dB_s}{\cosh(\sigma(T-s))}. \end{aligned} \quad (59)$$

Proof Because (58) looks like an Ornstein-Uhlenbeck process, we define for all $t \leq T$,

$$Z_t = e^{\int_0^t \sigma \tanh(\sigma(T-s)) ds} Y_t. \quad (60)$$

Applying Itô's formula to (60), one obtains that:

$$Z_t = Y_0 - \int_0^t \frac{\sigma c}{\cosh(\sigma(T-s))} e^{\int_0^s \sigma \tanh(\sigma(T-u)) du} ds + \int_0^t \sigma e^{\int_0^s \sigma \tanh(\sigma(T-u)) du} dB_s. \quad (61)$$

Since the primitive of $\tanh(x)$ is $\log \cosh(x)$, we obtain that:

$$\int_0^t \sigma \tanh(\sigma(T-s)) ds = \left[-\log \cosh(\sigma(T-s)) \right]_0^t = \log \left(\frac{\cosh(\sigma T)}{\cosh(\sigma(T-t))} \right). \quad (62)$$

Thus, it is possible to rewrite (60) as

$$Z_t = \frac{\cosh(\sigma T)}{\cosh(\sigma(T-t))} Y_t, \quad (63)$$

and we obtain moreover from (61) that

$$\begin{aligned} Z_t &= Y_0 - \int_0^t \frac{\sigma c \cosh(\sigma T)}{\cosh^2(\sigma(T-s))} ds + \int_0^t \frac{\sigma \cosh(\sigma T)}{\cosh(\sigma(T-s))} dB_s \\ &= Y_0 + c \cosh(\sigma T) \left[\tanh(\sigma(T-s)) \right]_0^t + \sigma \cosh(\sigma T) \int_0^t \frac{dB_s}{\cosh(\sigma(T-s))}, \end{aligned} \quad (64)$$

by using that a primitive of $1/\cosh^2(x)$ is $\tanh(x)$. Equations (63) and (64) give the announced result. \square

4.2 Return to the initial population process

The spinal process Y obtained in Theorems 4.1 and with generator given in (57) is associated to the auxiliary branching-diffusion process \tilde{Z}^K . We have now to prove that it is close to its analogous for the initial population process Z^K . By Corollary 3.1, we know that when we start from the stationary measure, these two processes are uniformly (in time) close when K is large, at least on a finite time interval. From this fact, we can obtain a similar result for Z^K as the one enounced for \tilde{Z}^K in Corollary 4.3.

Proposition 4.6. *When (Z_0^K) satisfies (6) and converges weakly and in probability to the stationary measure F defined in (16), then for any Φ continuous and bounded,*

$$\begin{aligned} \lim_{K \rightarrow +\infty} \mathbb{E}_{Z_0^K} \left[\frac{\langle H_T^K, \Phi \rangle}{\langle H_T^K, 1 \rangle} \right] &= \lim_{K \rightarrow +\infty} \mathbb{E} \left[\frac{1}{\tilde{N}_T^K} \sum_{i \in \tilde{V}_T^K} \Phi(\tilde{X}_s^i, s \leq T) \mid \sum_{i \in \tilde{V}_0^K} \delta_{\tilde{X}_0^i} = K Z_0^K \right] \\ &= \int_{\mathbb{R}} m_T(x) \mathbb{E}_x [\Phi(Y_s, s \leq T)] \frac{F(dx)}{\lambda}. \end{aligned} \quad (65)$$

Proof The right expression is obtained from Corollary 4.3 with $\xi_0 = F$ and by using (44). Let us consider $\varepsilon > 0$. It is possible to find a cylindrical test-function φ of the form (25) such that:

$$\sup_{y \in \mathcal{C}(\mathbb{R}, \mathbb{R})} |\Phi(y) - \varphi(y)| \leq \varepsilon. \quad (66)$$

By (56), the right hand side of (65) is the limit when K tends to infinity of

$$\mathbb{E} \left[\frac{1}{\tilde{N}_T^K} \sum_{u \in \tilde{V}_T^K} \Phi(\tilde{X}_s^u, s \leq T) \mid \sum_{i \in \tilde{V}_0^K} \delta_{\tilde{X}_0^i} = K Z_0^K \right] = \mathbb{E}_{Z_0^K} \left[\frac{\langle \tilde{H}_T^K, \Phi \rangle}{\langle \tilde{H}_T^K, 1 \rangle} \right]. \quad (67)$$

To prove the proposition, it is hence sufficient to prove that the left hand side of (65) and (67) have the same limit. For this, we write:

$$\begin{aligned} \left| \frac{\langle H_T^K, \Phi \rangle}{\langle H_T^K, 1 \rangle} - \frac{\langle \tilde{H}_T^K, \Phi \rangle}{\langle \tilde{H}_T^K, 1 \rangle} \right| &\leq \left| \frac{\langle H_T^K, \Phi \rangle}{\langle H_T^K, 1 \rangle} - \frac{\langle H_T^K, \varphi \rangle}{\langle H_T^K, 1 \rangle} \right| + \left| \frac{\langle H_T^K, \varphi \rangle}{\langle H_T^K, 1 \rangle} - \frac{\langle \tilde{H}_T^K, \varphi \rangle}{\langle \tilde{H}_T^K, 1 \rangle} \right| \\ &\quad + \left| \frac{\langle \tilde{H}_T^K, \varphi \rangle}{\langle \tilde{H}_T^K, 1 \rangle} - \frac{\langle \tilde{H}_T^K, \Phi \rangle}{\langle \tilde{H}_T^K, 1 \rangle} \right|. \end{aligned} \quad (68)$$

Notice that each of the fraction is upper-bounded by $\|\Phi\|_\infty$ or $\|\varphi\|_\infty$ (with the convention $0/0 = 0$) so that each of the terms in the right hand side is bounded. For the first term on the right hand side of (68), we have by (66):

$$\begin{aligned} \mathbb{E}_{Z_0^K} \left[\left| \frac{\langle H_T^K, \Phi \rangle}{\langle H_T^K, 1 \rangle} - \frac{\langle H_T^K, \varphi \rangle}{\langle H_T^K, 1 \rangle} \right| \right] &\leq \mathbb{E}_{Z_0^K} \left[\frac{1}{\langle H_T^K, 1 \rangle} |\langle H_T^K, \Phi - \varphi \rangle| \right] \\ &\leq \|\Phi - \varphi\|_\infty \leq \varepsilon. \end{aligned}$$

Proceeding similarly, we can show that the third term is also upper bounded by ε . For the second term, let us first introduce the following stopping times, for $1 > \eta > 0$:

$$\tau_\eta^K = \inf \{ t \in \mathbb{R}_+, \langle H_t^K, 1 \rangle \notin (\eta, 1/\eta) \}, \quad \tilde{\tau}_\eta^K = \inf \{ t \in \mathbb{R}_+, \langle \tilde{H}_t^K, 1 \rangle \notin (\eta, 1/\eta) \}.$$

Because the processes $(\langle H_t^K, 1 \rangle)_{t \in \mathbb{R}_+} = (\langle Z_t^K, 1 \rangle)_{t \in \mathbb{R}_+}$ and $(\langle \tilde{H}_t^K, 1 \rangle)_{t \in \mathbb{R}_+}$ converge to λ (see Proposition 3.4, Corollary 2.4 and Equation (44)), we have that for η small enough:

$$\lim_{K \rightarrow +\infty} \mathbb{P}(\tau_\eta^K \leq T) = \lim_{K \rightarrow +\infty} \mathbb{P}(\tilde{\tau}_\eta^K \leq T) = 0.$$

Thus, it is possible to choose η such that both probabilities are smaller than ε . Then,

$$\begin{aligned} \left| \frac{\langle H_T^K, \varphi \rangle}{\langle H_T^K, 1 \rangle} - \frac{\langle \tilde{H}_T^K, \varphi \rangle}{\langle \tilde{H}_T^K, 1 \rangle} \right| &\leq |\langle \tilde{H}_T^K, \varphi \rangle| \times \left| \frac{1}{\langle \tilde{H}_T^K, 1 \rangle} - \frac{\mathbb{1}_{\tau_\eta^K > T} \mathbb{1}_{\tilde{\tau}_\eta^K > T}}{\langle \xi_T, 1 \rangle} \right| \\ &\quad + \left| \frac{\mathbb{1}_{\tau_\eta^K > T} \mathbb{1}_{\tilde{\tau}_\eta^K > T}}{\langle \xi_T, 1 \rangle} \right| \times |\langle \tilde{H}_T^K, \varphi \rangle - \langle H_T^K, \varphi \rangle| \\ &\quad + |\langle H_T^K, \varphi \rangle| \times \left| \frac{1}{\langle H_T^K, 1 \rangle} - \frac{\mathbb{1}_{\tau_\eta^K > T} \mathbb{1}_{\tilde{\tau}_\eta^K > T}}{\langle \xi_T, 1 \rangle} \right|. \end{aligned}$$

For the first term in the right hand side, we use that:

$$|\langle \tilde{H}_T^K, \varphi \rangle| \times \left| \frac{1}{\langle \tilde{H}_t^K, 1 \rangle} - \frac{\mathbb{1}_{\tau_\eta^K > T} \mathbb{1}_{\tilde{\tau}_\eta^K > T}}{\langle \xi_T, 1 \rangle} \right| \leq \mathbb{1}_{\tau_\eta^K \wedge \tilde{\tau}_\eta^K \leq T} \|\varphi\|_\infty + \frac{1}{\eta^2 \langle \xi_T, 1 \rangle} \left| \langle H_T^K, 1 \rangle - \langle \xi_T, 1 \rangle \right|,$$

and taking the expectation:

$$\mathbb{E}_{Z_0^K} \left[\left| \langle \tilde{H}_T^K, \varphi \rangle \right| \times \left| \frac{1}{\langle \tilde{H}_t^K, 1 \rangle} - \frac{\mathbb{1}_{\tau_\eta^K > T} \mathbb{1}_{\tilde{\tau}_\eta^K > T}}{\langle \xi_T, 1 \rangle} \right| \right] \leq 2\|\varphi\|_\infty \varepsilon + \frac{1}{\eta^2 \langle \xi_T, 1 \rangle} \mathbb{E} \left[\left| \langle \tilde{H}_T^K, 1 \rangle - \langle \xi_T, 1 \rangle \right| \right].$$

A similar upper-bound can be obtained for the third term. Gathering the latter bounds:

$$\begin{aligned} \mathbb{E}_{Z_0^K} \left[\left| \frac{\langle H_T^K, \varphi \rangle}{\langle H_T^K, 1 \rangle} - \frac{\langle \tilde{H}_T^K, \varphi \rangle}{\langle \tilde{H}_T^K, 1 \rangle} \right| \right] &\leq 4\|\varphi\|_\infty \varepsilon + \frac{1}{\eta^2 \langle \xi_T, 1 \rangle} \mathbb{E} \left[\left| \langle \tilde{H}_T^K, 1 \rangle - \langle \xi_T, 1 \rangle \right| + \left| \langle H_T^K, 1 \rangle - \langle \xi_T, 1 \rangle \right| \right] \\ &\quad + \frac{1}{\langle \xi_T, 1 \rangle} \mathbb{E} \left[\left| \langle \tilde{H}_T^K, \varphi \rangle - \langle H_T^K, \varphi \rangle \right| \right]. \end{aligned}$$

We can now conclude with Corollary 2.4 and Proposition 3.4. \square

4.3 The spinal time reversed equation

Our purpose in this section is to recover the trait ancestor of an individual sampled in F at time T , that is, in the population at time T when the initial condition is the stationary solution F . For this, we need to reverse the time in the equation of the spinal process, and we will use to this purpose a result by Haussmann and Pardoux [43]. Their formula to reverse the diffusion (59) requires the computation of the density of Y_t for every time $t \geq 0$.

First, notice that:

Proposition 4.7. *The approximating (for $K \rightarrow +\infty$) distribution at time 0 of a trait chosen uniformly in the population at time T according to the stationary measure F comes from a biased initial condition, $\frac{1}{\lambda} m_T(x) F(x)$ and not $\frac{1}{\lambda} F(x)$. Conditionally to $Y_T \rightsquigarrow F$,*

$$Y_0 \sim \mathcal{N} \left(-ce^{-\sigma T}, \frac{\sigma}{1 + \tanh(\sigma T)} \right). \quad (69)$$

Proof Applying (65) with $\Phi(x_s, s \leq T) = f(x_0)$, we obtain that the random variable Y_0 has the distribution $\frac{1}{\lambda} m_T(x) F(x) dx$. Computing this measure yields that Y_0 has a Gaussian law whose expectation and variance are respectively $-ce^{-\sigma T}$ and $\sigma/(1 + \tanh(\sigma T))$. \square

Remember that F is a Gaussian distribution centered in $-c$. The biased distribution $\frac{1}{\lambda} m_T F$ describes the traits at time 0 of the individuals producing individuals alive at T . When T is large, its support is in the tail of the distribution F .

We are now able to compute the density of Y_t , using Proposition 4.5.

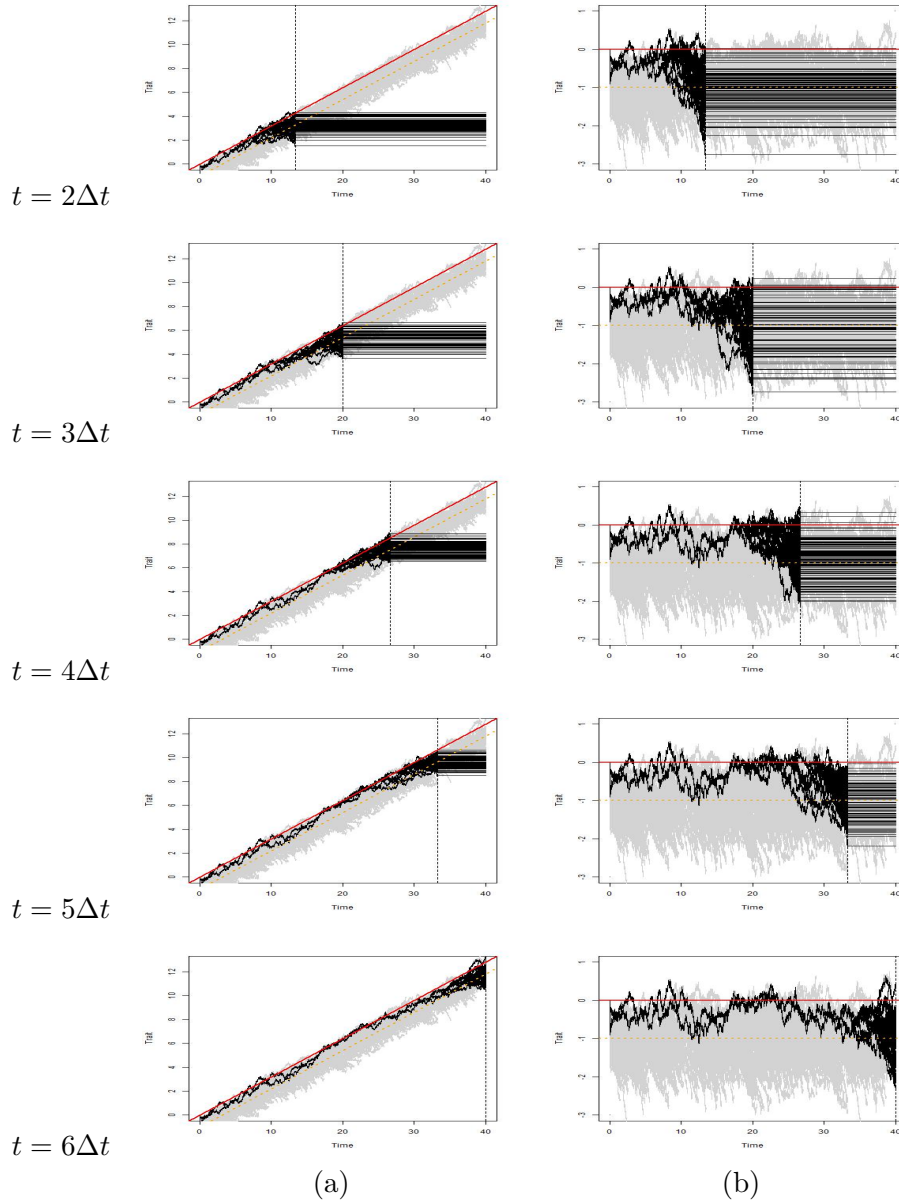


Figure 3: Evolution of the ancestral lineages of the present population, for various times $t = k\Delta t$ with $\Delta t = 20/3$ and $k \in \{2, \dots, 6\}$. The traits in the population (ordinate) are shown with respect to time (abscissa). The extinct lineages are in gray, whereas the lineages of the living particles are in block. (a): fixed framework. (b): mobile framework.

Proposition 4.8. For any $0 \leq t \leq T$, the random variable Y_t is a normal variable with law

$$Y_t \sim \mathcal{N}\left(-ce^{-\sigma(T-t)}, \frac{\sigma}{1 + \tanh(\sigma(T-t))}\right)$$

whose density $p(t, x)$ is given by

$$\partial_x \log p(t, x) = -\frac{x + ce^{-\sigma(T-t)}}{\sigma} (1 + \tanh(\sigma(T-t))).$$

Proof From (59) and (69), we deduce that Y_t has a normal law.

$$\begin{aligned} \mathbb{E}[Y_t] &= -ce^{-\sigma T} \frac{\cosh(\sigma(T-t))}{\cosh(\sigma T)} + c \cosh(\sigma(T-t)) \{ \tanh(\sigma(T-t)) - \tanh(\sigma T) \} \\ &= c \cosh(\sigma(T-t)) \left(-\frac{e^{-\sigma T}}{\cosh(\sigma T)} - \tanh(\sigma T) + \tanh(\sigma(T-t)) \right). \end{aligned}$$

Since

$$\tanh(\sigma T) - 1 = \frac{e^{\sigma T} - e^{-\sigma T} - e^{\sigma T} - e^{-\sigma T}}{e^{\sigma T} + e^{-\sigma T}} = -\frac{e^{-\sigma T}}{\cosh(\sigma T)}, \quad (70)$$

we have

$$\begin{aligned} \mathbb{E}[Y_t] &= c \cosh(\sigma(T-t)) (\tanh(\sigma(T-t)) - 1) \\ &= -c \cosh(\sigma(T-t)) \left(\frac{e^{-\sigma(T-t)}}{\cosh(\sigma(T-t))} \right) = -ce^{-\sigma(T-t)}. \end{aligned}$$

Additionally,

$$\begin{aligned} \text{var}(Y_t) &= \cosh(\sigma(T-t))^2 \left(\frac{\sigma}{\cosh^2(\sigma T) (1 + \tanh(\sigma T))} + \sigma^2 \int_0^t \frac{1}{\cosh^2(\sigma(T-s))} ds \right) \\ &= \cosh(\sigma(T-t))^2 \left(\frac{\sigma e^{-\sigma T}}{\cosh(\sigma T)} + \sigma (\tanh(\sigma T) - \tanh(\sigma(T-t))) \right) \\ &= \cosh(\sigma(T-t))^2 \sigma (1 - \tanh(\sigma(T-t))) \\ &= \sigma \cosh(\sigma(T-t)) e^{-\sigma(T-t)} = \frac{\sigma}{1 + \tanh(\sigma(T-t))}. \end{aligned}$$

The result follows. \square

We are now able to obtain the time reversed equation giving the trajectory leading from the trait of a “typical” individual living at time T in the stationary distribution F , to its ancestor.

Proposition 4.9. *The time reversed process of the spinal process Y is the time homogeneous Ornstein-Uhlenbeck process driving the ancestral trajectories around 0, satisfying the equation*

$$d\hat{Y}_s = -\sigma \hat{Y}_s ds + \sigma dW_s, \quad (71)$$

for a Brownian motion W .

Proof To reverse time in the equation (58), we apply an explicit formula given in [43]. The reverse process will be a diffusion process with the same diffusion coefficient σ and with a new drift term

$$b^r(t, x) = -b(T - t, x) + \sigma^2 \partial_x \log p(T - t, x),$$

where $p(t, \cdot)$ is the density of Y_t and $b(t, x)$ is the drift term in (58):

$$b(t, x) = -\sigma \tanh(\sigma(T - t))x - \frac{\sigma c}{\cosh(\sigma(T - t))}.$$

We obtain

$$\begin{aligned} b^r(t, x) &= \sigma \tanh(\sigma t)x + \frac{\sigma c}{\cosh(\sigma t)} - \sigma x (\tanh(\sigma t)) - \sigma x - \sigma c e^{-\sigma t} (1 + \tanh(\sigma t)) \\ &= -\sigma x, \end{aligned}$$

by using (70). The reverse process is then a very simple time homogeneous Ornstein-Uhlenbeck process driving the ancestral trajectories around 0, satisfying Equation (71) for a Brownian motion W . \square

As a consequence of Propositions 4.6 and 4.9, we can now summarize our results in the following.

Theorem 4.10. *Let U_K be a random variable whose conditional distribution with respect to H_T^K is uniform on V_T^K and consider the processes $(\widehat{Y}_s^K)_{0 \leq s \leq T}$ defined by*

$$\widehat{Y}_s^K = X_{T-s}^{U_K}, \quad \forall s \in [0, T].$$

Then, under the hypotheses of Proposition 4.6, the processes \widehat{Y}^K converges, as K goes to infinity, weakly to \widehat{Y} in $\mathcal{C}([0, T], \mathbb{R})$.

References

- [1] M. Alfaro, H. Berestycki, and G. Raoul. The Effect of Climate Shift on a Species Submitted to Dispersion, Evolution, Growth, and Nonlocal Competition. *SIAM Journal on Mathematical Analysis*, 49(1):562–596, Jan. 2017.
- [2] J. T. Anderson, A. M. Panetta, and T. Mitchell-Olds. Evolutionary and Ecological Responses to Anthropogenic Climate Change: Update on Anthropogenic Climate Change. *Plant Physiology*, 160(4):1728–1740, Dec. 2012. Publisher: American Society of Plant Biologists Section: UPDATES - FOCUS ISSUE.
- [3] V. Bansaye, J.-F. Delmas, L. Marsalle, and V. C. Tran. Limit theorems for Markov processes indexed by continuous time Galton–Watson trees. *The Annals of Applied Probability*, 21(6):2263–2314, Dec. 2011. Publisher: Institute of Mathematical Statistics.
- [4] N. Beerenwinkel, T. Antal, D. Dingli, A. Traulsen, K. W. Kinzler, V. E. Velculescu, B. Vogelstein, and M. A. Nowak. Genetic Progression and the Waiting Time to Cancer. *PLOS Computational Biology*, 3(11):e225, Nov. 2007. Publisher: Public Library of Science.
- [5] A. B. Benítez, S. Gufler, S. Kliem, V. Tran, and A. Wakolbinger. Evolving genealogies for branching populations under selection and competition. in preparation, 2020.

- [6] H. Berestycki, O. Diekmann, C. J. Nagelkerke, and P. A. Zegeling. Can a Species Keep Pace with a Shifting Climate? *Bulletin of Mathematical Biology*, 71(2):399–429, Feb. 2009.
- [7] J. Berestycki, N. Berestycki, and J. Schweinsberg. The genealogy of branching Brownian motion with absorption. *Annals of Probability*, 41(2):527–618, 2013.
- [8] N. Berestycki. Recent progress in coalescent theory. page 193.
- [9] P. Billingsley. *Convergence of Probability Measures*. Wiley Series in Probability and Statistics. John Wiley & Sons, Inc., Hoboken, NJ, USA, July 1999.
- [10] W. E. Bradshaw and C. M. Holzapfel. Evolutionary response to rapid climate change. *Science*, 312(5779):1477–1478, 2006.
- [11] E. Brunet and B. Derrida. Genealogies in simple models of evolution. *Journal of Statistical Mechanics: Theory and Experiment*, 2013(01):P01006, Jan. 2013. Publisher: IOP Publishing.
- [12] E. Brunet, B. Derrida, A. H. Mueller, and S. Munier. Effect of selection on ancestry: an exactly soluble case and its phenomenological generalization. *Physical Review E*, 76(4):041104, Oct. 2007. arXiv: 0704.3389.
- [13] R. Burger and M. Lynch. Evolution and Extinction in a Changing Environment: A Quantitative-Genetic Analysis. *Evolution*, 49(1):151–163, 1995.
- [14] M. T. Burrows, D. S. Schoeman, L. B. Buckley, P. Moore, E. S. Poloczanska, K. M. Brander, C. Brown, J. F. Bruno, C. M. Duarte, B. S. Halpern, J. Holding, C. V. Kappel, W. Kiessling, M. I. O’Connor, J. M. Pandolfi, C. Parmesan, F. B. Schwing, W. J. Sydeman, and A. J. Richardson. The Pace of Shifting Climate in Marine and Terrestrial Ecosystems. *Science*, 334(6056):652–655, Nov. 2011.
- [15] N. Champagnat. A microscopic interpretation for adaptive dynamics trait substitution sequence models. *Stochastic Processes and their Applications*, 116(8):1127–1160, Aug. 2006.
- [16] N. Champagnat, R. Ferrière, and S. Méléard. From individual stochastic processes to macroscopic models in adaptive dynamics. *Stochastic Models*, 24:2–44, 2008.
- [17] N. Champagnat and S. Méléard. Invasion and adaptive evolution for individual-based spatially structured populations. *Journal of Mathematical Biology*, 55:147–188, 2007.
- [18] B. Cloez. Limit theorems for some branching measure-valued processes. *Advances in Applied Probability*, 49(2):549–580, June 2017.
- [19] B. Cloez and P. Gabriel. On an irreducibility type condition for the ergodicity of nonconservative semigroups. *Comptes Rendus. Mathématique*, 358(6):733–742, 2020.
- [20] D. Collot, T. Nidelet, J. Ramsayer, O. C. Martin, S. Méléard, C. Dillmann, D. Sicard, and J. Legrand. Feedback between environment and traits under selection in a seasonal environment: consequences for experimental evolution. *Proceedings of the Royal Society of London B: Biological Sciences*, 285(1876), 2018.
- [21] D. Dawson and E. Perkins. *Historical Processes*, volume 93. American Mathematical Society, Memoirs of the American Mathematical Society edition, 1991.
- [22] D. A. Dawson. Measure-valued Markov processes. In Springer, editor, *Ecole d’Eté de probabilités de Saint-Flour XXI*, volume 1541 of *Lectures Notes in Math.*, pages 1–260, New York, 1993.
- [23] M. M. Desai, D. S. Fisher, and A. W. Murray. The speed of evolution and maintenance of variation in asexual populations. *Current biology: CB*, 17(5):385–394, Mar. 2007.
- [24] M. M. Desai, A. M. Walczak, and D. S. Fisher. Genetic Diversity and the Structure of Genealogies in Rapidly Adapting Populations. *Genetics*, 193(2):565–585, Feb. 2013.
- [25] P. Donnelly and T. Kurtz. A countable representation of the Fleming-Viot measure-valued diffusion. *Annals of Probability*, 24:698–742, 1996.
- [26] P. Donnelly and T. Kurtz. Particle representations for measure-valued population models. *Annals of Probability*, 27(1):166–205, 1999.

- [27] R. Durrett and J. Mayberry. Traveling waves of selective sweeps. *The Annals of Applied Probability*, 21(2):699–744, Apr. 2011. Publisher: Institute of Mathematical Statistics.
- [28] E. Dynkin. Branching particle systems and superprocesses. *Annals of Probability*, 19:1157–1194, 1991.
- [29] A. Etheridge. *An introduction to superprocesses*, volume 20 of *University Lecture Series*. American Mathematical Society, Providence, 2000.
- [30] A. Etheridge and T. Kurtz. Genealogical constructions of population models. *Annals of Probability*, 47(4):1827–1910, 2019.
- [31] R. P. Feynman. Space-Time Approach to Non-Relativistic Quantum Mechanics. *Reviews of Modern Physics*, 20(2):367–387, Apr. 1948. Publisher: American Physical Society.
- [32] P. Fitzsimmons, J. Pitman, and M. Yor. Markovian bridges: construction, palm interpretation, and splicing. In *Seminar on Stochastic Processes, 1992*, pages 101–134. Springer, 1993.
- [33] R. Forien, J. Garnier, and F. Patout. Ancestral lineages in mutation selection equilibria with moving optimum. arXiv:2011.05192.
- [34] N. Fournier and S. Méléard. A microscopic probabilistic description of a locally regulated population and macroscopic approximations. *Ann. Appl. Probab.*, 14(4):1880–1919, 2004.
- [35] J.-F. L. Gall. *Random trees and applications*. Probability Surveys, 2006.
- [36] A. Gonzalez, O. Ronce, R. Ferrière, and M. Hochberg. Evolutionary rescue: An emerging focus at the intersection between ecology and evolution. 368:20120404, 2013.
- [37] F. A. Gorter, M. M. G. Aarts, B. J. Zwaan, and J. A. G. M. de Visser. Dynamics of adaptation in experimental yeast populations exposed to gradual and abrupt change in heavy metal concentration. *Am. Nat.*, 187(1):110–119, 2016.
- [38] A. Greven, P. Pfaffelhuber, and A. Winter. Convergence in distribution of random metric measure spaces (λ -coalescent measure trees). *Probability Theory and Related Fields*, 145(1):285–322, 2009.
- [39] T. S. Guzella, S. Dey, I. M. Chelo, A. Pino-Querido, V. F. Pereira, S. R. Proulx, and H. Teotónio. Slower environmental change hinders adaptation from standing genetic variation. *PLOS Genetics*, 14(11):e1007731, Nov. 2018. Publisher: Public Library of Science.
- [40] R. Hardy and S. Harris. A new formulation of the spine approach to branching diffusions. 2006. preprint <http://arxiv.org/abs/math.PR/0611054>.
- [41] R. Hardy and S. Harris. A spine approach to branching diffusions with applications to L^p -convergence of martingales. In Springer, editor, *Séminaire de Probabilités*, volume XLII of *Lectures Notes in Math.*, pages 281–330, Berlin, 2009.
- [42] S. Harris and M. Roberts. The many-to-few lemma and multiple spines. 2017.
- [43] U. Haussmann and E. Pardoux. Time reversal of diffusions. *The Annals of Probability*, 14(4):1188–1205, 1986.
- [44] A. P. Hendry, T. J. Farrugia, and M. T. Kinnison. Human influences on rates of phenotypic change in wild animal populations. *Molecular Ecology*, 17(1):20–29, 2008.
- [45] A. A. Hoffmann and C. M. Sgrò. Climate change and evolutionary adaptation. *Nature*, 470(7335):479–485, 2011.
- [46] N. Ikeda and S. Watanabe. *Stochastic Differential Equations and Diffusion Processes*, volume 24. North-Holland Publishing Company, 1989. Second Edition.
- [47] M. Kac. On distributions of certain Wiener functionals. *Transactions of the American Mathematical Society*, 65(1):1–13, 1949.
- [48] M. Kac. On Some Connections between Probability Theory and Differential and Integral Equations. *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, pages 189–215, Jan. 1951. Publisher: University of California Press.

- [49] D. A. Kessler, H. Levine, D. Ridgway, and L. Tsimring. Evolution on a smooth landscape. *Journal of Statistical Physics*, 87(3-4):519–544, May 1997.
- [50] S. Kliem. A compact containment result for nonlinear historical superprocess approximations for population models with trait-dependence. *Electronic Journal of Probability*, 19(97):1–13, 2014.
- [51] S. Kliem and A. Winter. Evolving phylogenies of trait-dependent branching with mutation and competition. part i: Existence. *Stochastic Processes and their Applications*, 129(12):4837–4877, December 2019.
- [52] M. Kopp and S. Matuszewski. Rapid evolution of quantitative traits: theoretical perspectives. *Evolutionary Applications*, 7(1):169–191, 2013.
- [53] T. Kurtz, R. Lyons, R. Pemantle, and Y. Peres. A conceptual proof of the kesten-stigum theorem for multi-type branching processes. In *Classical and modern branching processes*, pages 181–185. Springer, 1997.
- [54] R. Lande and S. Shannon. The role of genetic variation in adaptation and population persistence in a changing environment. *Evolution*, 50(1):434–437, 1996.
- [55] C. Lepers, S. Billiard, M. Porte, S. Méléard, and V. Tran. Inference with selection, varying population size and evolving population structure: Application of abc to a forward-backward. *Heredity*, 2020. published online.
- [56] S. R. Loarie, P. B. Duffy, H. Hamilton, G. P. Asner, C. B. Field, and D. D. Ackerly. The velocity of climate change. *Nature*, 462(7276):1052–1055, Dec. 2009.
- [57] M. Lynch, W. Gabriel, and A. M. Wood. Adaptive and demographic responses of plankton populations to environmental change. *Limnology and Oceanography*, 36:1301–1312, 1991.
- [58] M. Lynch and R. Lande. *Evolution and extinction in response to environmental change*. Sinauer Assoc. 1993.
- [59] R. Lyons, R. Pemantle, and Y. Peres. Conceptual proofs of $l \log l$ criteria for mean behavior of branching processes. *The Annals of Probability*, pages 1125–1138, 1995.
- [60] A. Marguet. A law of large numbers for branching markov processes by the ergodicity of ancestral lineages. *ESAIM: Probability and Statistics*, 23:638–661, 2019.
- [61] A. Marguet. Uniform sampling in a structured branching population. *Bernoulli*, 25(4A):2649–2695, 2019.
- [62] S. Méléard and V. Tran. Nonlinear historical superprocess approximations for population models with past dependence. *Electronic Journal of Probability*, 17(47), 2012.
- [63] R. A. Neher and O. Hallatschek. Genealogies of rapidly adapting populations. *Proceedings of the National Academy of Sciences*, 110(2):437–442, Jan. 2013.
- [64] S.-C. Park, D. Simon, and J. Krug. The Speed of Evolution in Large Asexual Populations. *Journal of Statistical Physics*, 138(1):381–410, Feb. 2010.
- [65] C. Parmesan. Evolutionary and ecological responses to recent climate change. *Annu. Rev. Ecol. Evol.*, 37(8):637–669, 2006.
- [66] E. Perkins. *On the Martingale Problem for Interactive Measure-Valued Branching Diffusions*, volume 115(549). American Mathematical Society, Memoirs of the American Mathematical Society edition, May 1995.
- [67] E. A. Perkins. Dawson-Watanabe superprocesses and measure-valued diffusions. In Springer, editor, *Ecole d’Eté de probabilités de Saint-Flour*, volume 1781 of *Lectures Notes in Math.*, pages 125–329, New York, 1993.
- [68] A. B. Potapov and M. A. Lewis. Climate and competition: The effect of moving range boundaries on habitat invasibility. *Bulletin of Mathematical Biology*, 66(5):975–1008, Sept. 2004.

- [69] L. Roques, J. Garnier, F. Hamel, and E. Klein. Allee effect promotes diversity in traveling waves of colonization. *PNAS*, 109(23):8828–8833, 2012.
- [70] I. M. Rouzine, J. Wakeley, and J. M. Coffin. The solitary wave of asexual evolution. *Proceedings of the National Academy of Sciences*, 100(2):587–592, Jan. 2003. Publisher: National Academy of Sciences Section: Biological Sciences.
- [71] J. Schweinsberg. Rigorous results for a population model with selection I: evolution of the fitness distribution. *Electronic Journal of Probability*, 22(none):1–94, Jan. 2017. Publisher: Institute of Mathematical Statistics and Bernoulli Society.
- [72] J. Schweinsberg. Rigorous results for a population model with selection II: genealogy of the population. *Electronic Journal of Probability*, 22(none):1–54, Jan. 2017. Publisher: Institute of Mathematical Statistics and Bernoulli Society.
- [73] K. S. Sheldon. Climate Change in the Tropics: Ecological and Evolutionary Responses at Low Latitudes. *Annual Review of Ecology, Evolution, and Systematics*, 50(1):303–333, Nov. 2019.
- [74] V. Tran. *Une ballade en forêts aléatoires. Théorèmes limites pour des populations structurées et leurs généalogies, étude probabiliste et statistique de modèles SIR en épidémiologie, contributions à la géométrie aléatoire.* Habilitation à diriger des recherches, Université de Lille 1, 11 2014. <http://tel.archives-ouvertes.fr/tel-01087229>.
- [75] L. S. Tsimring, H. Levine, and D. A. Kessler. RNA Virus Evolution via a Fitness-Space Model. *Physical Review Letters*, 76(23):4440–4443, June 1996. Publisher: American Physical Society.

A SDEs for the stochastic birth-death particle system and the historical particle system

A.1 Pathwise representation of the population process

We recall here the pathwise representation of our measure-valued processes, as solution of stochastic differential equations driven by independent Poisson point measures and Brownian motions. We refer to [17] and [62] for more details.

To model the random occurrence of birth and death events, let us consider a Poisson point process $N(ds, di, d\theta)$ on $\mathbb{R}_+ \times \mathcal{I} \times \mathbb{R}_+$, with intensity measure $ds \otimes n(di) \otimes d\theta$, where $n(di)$ is the counting measure on the set of labels $\mathcal{I} = \bigcup_{n \in \mathbb{N}} \mathbb{N}^n$.

We also introduce a family of independent standard Brownian motions $(B^i, i \in \mathcal{I})$ indexed by \mathcal{I} that will drive the particle motions.

The atoms of the Poisson point process determine birth and death events. These events modify the set of individuals alive, V_t^K . Between these events, the position of a particle alive, say i , is modelled by a drifted diffusion

$$dX_t^i = \sigma(dB_t^i - c dt). \quad (72)$$

Let us consider a test function $f \in C_b^{1,2}(\mathbb{R}_+ \times \mathbb{R})$. We will use the notation $f_t(x) = f(t, x)$. Between two jump times, the set of living individuals is fixed and we can apply Itô's formula to the diffusion processes (72) related to the individuals i alive. At a jump time τ , if we have a birth of individual i , a new offspring appears at the same position and the process

increases from $f_\tau(X_\tau^i)$. If we have a death of individual i , the process decreases of $f_\tau(X_\tau^i)$. Then the measure-valued population process Z^K acts on the test function f as:

$$\begin{aligned}
\langle Z_t^K, f_t \rangle &= \langle Z_0^K, f_0 \rangle \\
&+ \frac{1}{K} \int_0^t \int_{\mathcal{I}} \int_{\mathbb{R}_+} \mathbb{1}_{\{i \in V_{s-}^K\}} \left[f_s(X_s^i) \mathbb{1}_{\{\theta \leq 1\}} - f_s(X_s^i) \mathbb{1}_{\{1 < \theta \leq 1 + \frac{(X_s^i)^2}{2} + \langle Z_{s-}^K, 1 \rangle\}} \right] N(ds, di, d\theta) \\
&+ \frac{1}{K} \sum_{i \in \mathcal{I}} \int_0^t \mathbb{1}_{\{i \in V_s^K\}} \sigma \partial_x f_s(X_s^i) dB_s^i \\
&+ \frac{1}{K} \sum_{i \in \mathcal{I}} \int_0^t \mathbb{1}_{\{i \in V_s^K\}} \left(\partial_s f_s(X_s^i) - c \sigma \partial_x f_s(X_s^i) + \frac{\sigma^2}{2} \partial_{xx}^2 f_s(X_s^i) \right) ds, \tag{73}
\end{aligned}$$

and where the set of living individuals is changing as follows.

- $V_0^K = \{1, \dots, K\}$ and $|V_0^K| = K \langle Z_0^K, 1 \rangle$.
- For each atom (s, i, θ) of N such that $i \in V_{s-}^K$ and $\theta \leq 1$, there is a new birth by individual i , and the label of the new offspring is $j = (i, k)$ where k is the rank of the new individual among the daughters of i .
- For each atom (s, i, θ) of N such that $i \in V_{s-}^K$ and $1 < \theta \leq 1 + (X_s^i)^2/2 + |V_{s-}^K|/K$, there is a death and the label i is removed from V_{s-}^K .

Introducing the compensated martingale measure of the Poisson point measure, we obtain that

$$\begin{aligned}
\langle Z_t^K, f_t \rangle &= \langle Z_0^K, f_0 \rangle + M_t^{K, \varphi} + \int_0^t \int_{\mathbb{R}} \left\{ \left(1 - \frac{1}{2} x^2 - \langle Z_s^K, 1 \rangle \right) f_s(x) \right. \\
&\quad \left. + \partial_s f_s(x) - \sigma c \partial_x f_s(x) + \frac{\sigma^2}{2} \partial_{xx}^2 f_s(x) \right\} Z_s^K(dx) ds, \tag{74}
\end{aligned}$$

where the process $M^{K, \varphi}$ is a square integrable martingale with quadratic variation process given by

$$\langle M^{K, \varphi} \rangle_t = \frac{1}{K} \int_0^t \int_{\mathbb{R}} \left\{ \left(1 + \frac{x^2}{2} + \langle Z_s^{c, K}, 1 \rangle \right) f_s^2(x) + \sigma^2 (\partial_x f_s)^2(x) \right\} Z_s^K(dx) ds. \tag{75}$$

A.2 Pathwise representation of the historical population process

Let us consider test functions φ defined on $\mathbb{R}_+ \times \mathcal{C}(\mathbb{R}_+, \mathbb{R})$ with a similar form as in (25), i.e. for any $s, y \in \mathbb{R}_+ \times \mathcal{C}(\mathbb{R}_+, \mathbb{R})$,

$$\varphi(s, y) = \varphi_s(y) = \prod_{j=1}^m g_j(s, y_{t_j}),$$

for $m \in \mathbb{N}^*$, $0 \leq t_1 < \dots < t_m$ and $\forall j \in \{1, \dots, m\}$, $g_j \in C_b^{1,2}(\mathbb{R}_+ \times \mathbb{R}, \mathbb{R})$. Note that

$$\varphi(s, y_{\cdot \wedge s}) = \prod_{j=1}^m g_j(s, y_{t_j \wedge s}) = \sum_{k=0}^{m-1} \mathbb{1}_{[t_k, t_{k+1})}(s) \left(\prod_{j=1}^k g_j(s, y_{t_j}) \prod_{j=k+1}^m g_j(s, y_s) \right).$$

It is possible to write a stochastic differential equation for the historical process H^K defined in (24) that is driven by the same Poisson point measures and Brownian motion as the process Z^K . With the notation (26) and (27) introduced in Section 3.2, we have

$$\begin{aligned} \langle H_t^K, \varphi_t \rangle &= \langle H^K, \varphi_0 \rangle + \frac{1}{K} \sum_{i \in \mathcal{I}} \int_0^t \mathbb{1}_{\{i \in V_s^K\}} \sigma \tilde{D} \varphi_s(X_s^i) dB_s^i \\ &+ \frac{1}{K} \int_0^t \int_{\mathcal{I}} \int_{\mathbb{R}_+} \mathbb{1}_{\{i \in V_s^K\}} \left[\varphi_s(X_{(\cdot \wedge s)}^i) \mathbb{1}_{\{\theta \leq 1\}} - \varphi_s(X_{(\cdot \wedge s)}^i) \mathbb{1}_{\{1 < \theta \leq 1 + \frac{(X_s^i)^2}{2} + \langle H_{s-}^K, 1 \rangle\}} \right] N(ds, di, d\theta) \\ &+ \frac{1}{K} \int_0^t \sum_{i \in V_s^K} \left(\partial_s \varphi_s(X_{(\cdot \wedge s)}^i) - c \sigma \tilde{D} \varphi_s(X_{(\cdot \wedge s)}^i) + \frac{\sigma^2}{2} \tilde{\Delta} \varphi_s(X_{(\cdot \wedge s)}^i) \right) ds. \end{aligned} \quad (76)$$

Then introducing the compensated martingales measures associated with the Poisson point processes, we obtain that

$$\begin{aligned} \langle H_t^K, \varphi_t \rangle &= \langle H^K, \varphi_0, \cdot \rangle + \frac{1}{K} \int_0^t \sum_{i \in V_s^K} \sigma \tilde{D} \varphi_s(X_{(\cdot \wedge s)}^i) dB_s^i \\ &+ \frac{1}{K} \int_0^t \int_{\mathcal{I}} \int_{\mathbb{R}_+} \mathbb{1}_{\{i \in V_s^K\}} \left[\varphi_s(X_{(\cdot \wedge s)}^i) \mathbb{1}_{\{\theta \leq 1\}} - \varphi_s(X_{(\cdot \wedge s)}^i) \mathbb{1}_{\{1 < \theta \leq 1 + \frac{(X_s^i)^2}{2} + \langle H_{s-}^K, 1 \rangle\}} \right] \tilde{N}(ds, di, d\theta) \\ &+ \frac{1}{K} \int_0^t \sum_{i \in V_s^K} \left(1 - \frac{(X_s^i)^2}{2} + \langle H_{s-}^K, 1 \rangle \right) \varphi_s(X_{(\cdot \wedge s)}^i) ds \\ &+ \frac{1}{K} \int_0^t \sum_{i \in V_s^K} \left(\partial_s \varphi_s(X_{(\cdot \wedge s)}^i) - c \sigma \tilde{D} \varphi_s(X_{(\cdot \wedge s)}^i) + \frac{\sigma^2}{2} \tilde{\Delta} \varphi_s(X_{(\cdot \wedge s)}^i) \right) ds \\ &= \langle H^K, \varphi_0, \cdot \rangle + \mathcal{M}_t^K(\varphi) \\ &+ \int_0^t \int_{\mathcal{C}(\mathbb{R}_+, \mathbb{R})} \left(\left(1 - \frac{y_s^2}{2} - \langle H_s^K, 1 \rangle \right) \varphi(s, y) + \partial_s \varphi_s(y) + \frac{\sigma^2}{2} \tilde{\Delta} \varphi(s, y) - \sigma c \tilde{D} \varphi(s, y) \right) H_s^K(dy) ds. \end{aligned} \quad (77)$$

The process is a square integrable local martingale with quadratic variation

$$\langle \mathcal{M}^K(\varphi) \rangle_t = \frac{1}{K} \int_0^t \int_{\mathcal{C}(\mathbb{R}_+, \mathbb{R})} \left(\left(1 + \frac{y_s^2}{2} + \langle H_s^K, 1 \rangle \right) \varphi(s, y) + \sigma^2 (\tilde{D} \varphi(s, y))^2 \right) H_s^K(dy) ds. \quad (78)$$

A.3 Stochastic mild equation

Recall that $(P_t)_{t \geq 0}$ is the semi-group defined in (10). For a fixed $t > 0$ and a test function $\varphi \in \mathcal{C}_b^2(\mathbb{R})$, choosing $f(s, x) = P_{t-s}\varphi(x)$, we obtain from (73) a mild stochastic equation:

$$\begin{aligned}
& \langle Z_t^K, \varphi \rangle = \langle Z_0^K, P_t \varphi \rangle \\
& + \frac{1}{K} \int_0^t \int_{\mathcal{I}} \int_{\mathbb{R}_+} \mathbb{1}_{\{i \in V_{s-}^K\}} \left[P_{t-s}\varphi(X_s^i) \mathbb{1}_{\{\theta \leq 1\}} - P_{t-s}\varphi(X_s^i) \mathbb{1}_{\{1 < \theta \leq 1 + \frac{(X_s^i)^2}{2} + \langle Z_{s-}^K, 1 \rangle\}} \right] N(ds, di, d\theta) \\
& + \frac{1}{K} \sum_{i \in \mathcal{I}} \int_0^t \mathbb{1}_{\{i \in V_s^K\}} \sigma \partial_x P_{t-s}\varphi(X_s^i) dB_s^i \\
& = \langle Z_0^K, P_t \varphi \rangle + \int_0^t \int_{\mathbb{R}} \left(1 - \frac{x^2}{2} - \langle Z_s^K, 1 \rangle \right) P_{t-s}\varphi(x) Z_s^K(dx) ds + \mathcal{M}_t^{K,\varphi}
\end{aligned} \tag{79}$$

where $\mathcal{M}_t^{K,\varphi}$ is the following square integrable martingale:

$$\begin{aligned}
\mathcal{M}_t^{K,\varphi} &= \frac{1}{K} \int_0^t \int_{\mathcal{I}} \int_{\mathbb{R}_+} \mathbb{1}_{\{i \in V_{s-}^K\}} \left[P_{t-s}\varphi(X_s^i) \mathbb{1}_{\{\theta \leq 1\}} \right. \\
& \quad \left. - P_{t-s}\varphi(X_s^i) \mathbb{1}_{\{1 < \theta \leq 1 + \frac{(X_s^i)^2}{2} + \langle Z_{s-}^K, 1 \rangle\}} \right] (N(ds, di, d\theta) - ds \otimes n(di) \otimes d\theta) \\
& + \frac{1}{K} \sum_{i \in \mathcal{I}} \int_0^t \mathbb{1}_{\{i \in V_s^K\}} \sigma \partial_x P_{t-s}\varphi(X_s^i) dB_s^i.
\end{aligned} \tag{80}$$

The predictable quadratic variation of $\mathcal{M}_t^{K,\varphi}$ is

$$\begin{aligned}
\langle \mathcal{M}^{K,\varphi} \rangle_t &= \frac{1}{K} \int_0^t \left(1 + \frac{x^2}{2} + \langle Z_s^K, 1 \rangle \right) (P_{t-s}\varphi(x))^2 Z_s^K(dx) ds \\
& + \frac{1}{K} \int_0^t \sigma^2 \langle Z_s^K, (\partial_x P_{t-s}\varphi)^2 \rangle ds.
\end{aligned} \tag{81}$$

B Moment estimates for Z^K : proof of Lemma 2.1

We prove a more precise form of Lemma 2.1.

Lemma B.1. *We assume that the initial condition Z_0^K satisfies for $\epsilon > 0$ that:*

$$\sup_{K \in \mathbb{N}^*} \mathbb{E}(\langle Z_0^K, 1 \rangle^{2+\epsilon}) < +\infty \quad \text{and} \quad \sup_{K \in \mathbb{N}^*} \mathbb{E}(\langle Z_0^K, x^2 \rangle^{1+\epsilon}) < +\infty. \tag{82}$$

Then, for any $T > 0$, we have

$$\sup_{K \in \mathbb{N}^*} \mathbb{E} \left(\sup_{t \in [0, T]} \langle Z_t^K, 1 \rangle^{2+\epsilon} \right) < +\infty \quad \text{and} \quad \sup_{K \in \mathbb{N}^*} \sup_{t \in [0, T]} \mathbb{E}(\langle Z_t^K, x^2 \rangle^{1+\epsilon}) < +\infty. \tag{83}$$

Under the additional assumption that:

$$\sup_{K \in \mathbb{N}^*} \mathbb{E}(\langle Z_0^K, x^4 \rangle^{1+2\epsilon}) < +\infty, \quad (84)$$

we also have that:

$$\sup_{K \in \mathbb{N}^*} \mathbb{E} \left(\sup_{t \in [0, T]} \langle Z_t^K, x^2 \rangle^{1+\epsilon} \right) < +\infty. \quad (85)$$

Using classical computation (see e.g. [34]), several moment estimates can be derived under Assumption (82). Recall that $T > 0$ and assume (82), i.e. that the initial condition Z_0^K satisfies for $\epsilon > 0$ that:

$$\sup_{K \geq 1} \mathbb{E}(\langle Z_0^K, 1 \rangle^{2+\epsilon}) < +\infty \quad \text{and} \quad \sup_{K \in \mathbb{N}^*} \mathbb{E}(\langle Z_0^K, x^2 \rangle^{1+\epsilon}) < +\infty.$$

Step 1: Let us introduce the stopping time, for $M > 0$ and for $K \geq 1$:

$$\tau_M^K = \inf \{ t \geq 0, \langle Z_t^K, 1 \rangle^{2+\epsilon} > M \quad \text{or} \quad \langle Z_t^K, x^2 \rangle^{1+\epsilon} > M \}. \quad (86)$$

Choosing the test function $\varphi \equiv 1$ and neglecting the natural death term of rate $x^2/2$ gives in (74):

$$\langle Z_{t \wedge \tau_M^K}^K, 1 \rangle \leq \langle Z_0^K, 1 \rangle + \int_0^{t \wedge \tau_M^K} (\langle Z_s^K, 1 \rangle - \langle Z_s^K, 1 \rangle^2) ds + M_{t \wedge \tau_M^K}^{K,1}.$$

Taking the expectation and using the convexity of $x \mapsto x^2$, it follows that

$$\begin{aligned} \mathbb{E}(\langle Z_t^K, 1 \rangle) &\leq \mathbb{E}(\langle Z_0^K, 1 \rangle) + \int_0^t \left[\mathbb{E}(\langle Z_{s \wedge \tau_M^K}^K, 1 \rangle) - \mathbb{E}(\langle Z_s^K, 1 \rangle^2) \right] ds \\ &\leq \frac{\mathbb{E}(\langle Z_0^K, 1 \rangle)}{\mathbb{E}(\langle Z_0^K, 1 \rangle) + (1 - \mathbb{E}(\langle Z_0^K, 1 \rangle))e^{-t}}, \end{aligned}$$

since we recognize the logistic equation. Because the upper-bound does not depend on M , a direct consequence is that τ_M^K tends a.s. to infinity when $M \rightarrow +\infty$ and that:

$$\sup_{t \in \mathbb{R}_+} \mathbb{E}(\langle Z_t^K, 1 \rangle) < +\infty. \quad (87)$$

Step 2: Now, choosing the test function $\varphi(x) = 1$, using Itô's formula (see e.g. [46, p.66]) and neglecting the death terms:

$$\begin{aligned} \langle Z_{t \wedge \tau_M^K}^K, 1 \rangle^{2+\epsilon} &\leq \langle Z_0^K, 1 \rangle^{2+\epsilon} + \int_0^{t \wedge \tau_M^K} \int_{\mathcal{I}} \int_{\mathbb{R}} \mathbb{1}_{i \in V_{s-}^K} \left((\langle Z_{s-}^K, 1 \rangle + \frac{1}{K})^{2+\epsilon} - \langle Z_{s-}^K, 1 \rangle^{2+\epsilon} \right) \mathbb{1}_{\theta \leq 1} N(ds, di, d\theta) \\ &\leq \langle Z_0^K, 1 \rangle^{2+\epsilon} + \int_0^{t \wedge \tau_M^K} \int_{\mathcal{I}} \int_{\mathbb{R}} \mathbb{1}_{i \in V_{s-}^K} \frac{C}{K} \langle Z_{s-}^K, 1 \rangle^{1+\epsilon} \mathbb{1}_{\theta \leq 1} N(ds, di, d\theta), \end{aligned}$$

for a constant $C > 0$ and using that

$$\left(x + \frac{1}{K}\right)^{2+\epsilon} - x^{2+\epsilon} = x^{2+\epsilon} \left[\exp\left((2+\epsilon) \ln\left(1 + \frac{1}{Kx}\right)\right) - 1 \right] \stackrel{K \rightarrow +\infty}{\sim} x^{2+\epsilon} \times \frac{2+\epsilon}{xK} = \frac{(2+\epsilon)x^{1+\epsilon}}{K}. \quad (88)$$

Introducing the supremum in the right hand side, then in the left hand side and taking the expectation provides that:

$$\mathbb{E}\left(\sup_{s \leq t} \langle Z_{s \wedge \tau_M^K}^K, 1 \rangle^{2+\epsilon}\right) \leq \mathbb{E}(\langle Z_0^K, 1 \rangle^{2+\epsilon}) + C \int_0^t \mathbb{E}\left(\sup_{u \leq s} \langle Z_u^K, 1 \rangle^{2+\epsilon}\right) ds,$$

from which we obtain by Gronwall's lemma that:

$$\mathbb{E}\left(\sup_{s \leq t} \langle Z_{s \wedge \tau_M^K}^K, 1 \rangle^{2+\epsilon}\right) \leq \mathbb{E}(\langle Z_0^K, 1 \rangle^{2+\epsilon}) \exp(Ct),$$

where the upper bound does not depend on M nor on K . Then, letting $M \rightarrow +\infty$ provides the first estimate of (83).

Notice that a similar computation would have yielded that:

$$\sup_{K \in \mathbb{N}^*} \mathbb{E}\left(\sup_{s \leq t} \langle Z_s^K, 1 \rangle^{1+\epsilon}\right) < +\infty. \quad (89)$$

Step 3: Let us now consider the test function $\varphi(x) = x^2$. Using Itô's formula and neglecting the death terms, we obtain from (74):

$$\begin{aligned} & \langle Z_{t \wedge \tau_M^K}^K, x^2 \rangle^{1+\epsilon} \leq \langle Z_0^K, x^2 \rangle^{1+\epsilon} \\ & + \int_0^{t \wedge \tau_M^K} \int_{\mathcal{I}} \int_{\mathbb{R}_+} \mathbb{1}_{\{i \in V_{s-}^K\}} \mathbb{1}_{\{\theta \leq 1\}} \left(\langle Z_{s-}^K, x^2 \rangle + \frac{(X_s^i)^2}{K} \right)^{1+\epsilon} - \langle Z_{s-}^K, x^2 \rangle^{1+\epsilon} N(ds, di, d\theta) \\ & + \frac{1}{K} \sum_{i \in \mathcal{I}} \int_0^{t \wedge \tau_M^K} \mathbb{1}_{\{i \in V_s^K\}} 2\sigma(1+\epsilon) \langle Z_s^K, x^2 \rangle^\epsilon X_s^i dB_s^i + \frac{1}{K^2} \sum_{i \in \mathcal{I}} \int_0^{t \wedge \tau_M^K} \mathbb{1}_{\{i \in V_s^K\}} \frac{\epsilon(1+\epsilon)}{\langle Z_s^K, x^2 \rangle^{1-\epsilon}} 2\sigma^2 (X_s^i)^2 ds \\ & + \frac{1}{K} \sum_{i \in \mathcal{I}} \int_0^{t \wedge \tau_M^K} \mathbb{1}_{\{i \in V_s^K\}} (1+\epsilon) \langle Z_s^K, x^2 \rangle^\epsilon (\sigma^2 - 2c\sigma X_s^i) ds \end{aligned}$$

$$\begin{aligned} & \langle Z_{t \wedge \tau_M^K}^K, x^2 \rangle^{1+\epsilon} \leq \langle Z_0^K, x^2 \rangle^{1+\epsilon} + \frac{1}{K} \sum_{i \in \mathcal{I}} \int_0^{t \wedge \tau_M^K} \mathbb{1}_{\{i \in V_s^K\}} 2\sigma(1+\epsilon) \langle Z_s^K, x^2 \rangle^\epsilon X_s^i dB_s^i \\ & + \int_0^{t \wedge \tau_M^K} \int_{\mathcal{I}} \int_{\mathbb{R}_+} \mathbb{1}_{\{i \in V_{s-}^K\}} \mathbb{1}_{\{\theta \leq 1\}} \langle Z_{s-}^K, x^2 \rangle^{1+\epsilon} \left(\left(1 + \frac{(X_s^i)^2}{K \langle Z_{s-}^K, x^2 \rangle}\right)^{1+\epsilon} - 1 \right) N(ds, di, d\theta) \\ & + \epsilon(1+\epsilon) \frac{2}{K} \sigma^2 \int_0^t \langle Z_{s \wedge \tau_M^K}^K, x^2 \rangle^\epsilon ds + (1+\epsilon) \int_0^{t \wedge \tau_M^K} \langle Z_s^K, x^2 \rangle^\epsilon \times (\sigma^2 \langle Z_s^K, 1 \rangle - 2c\sigma \langle Z_s^K, x \rangle) ds \end{aligned} \quad (90)$$

First, because $x \leq 1 + x^2$, we have that:

$$\begin{aligned} |\langle Z_s^k, x^2 \rangle^\epsilon \times (\sigma^2 \langle Z_s^K, 1 \rangle - 2c\sigma \langle Z_s^K, x \rangle)| &\leq 2c\sigma \langle Z_s^K, x^2 \rangle^{1+\epsilon} + (\sigma^2 + 2c\sigma) \langle Z_s^K, x^2 \rangle^\epsilon \langle Z_s^K, 1 \rangle \\ &\leq (\sigma^2 + 4c\sigma) \langle Z_s^K, x^2 \rangle^{1+\epsilon} + (\sigma^2 + 2c\sigma) \langle Z_s^K, 1 \rangle^{1+\epsilon}. \end{aligned}$$

Then, notice that a computation similar to (88) gives that for a constant $C > 0$ sufficiently large,

$$\langle Z_{s-}^K, x^2 \rangle^{1+\epsilon} \left(\left(1 + \frac{(X_s^i)^2}{K \langle Z_{s-}^K, x^2 \rangle} \right)^{1+\epsilon} - 1 \right) \leq \langle Z_{s-}^K, x^2 \rangle^\epsilon \frac{C(1+\epsilon)}{K} (X_s^i)^2.$$

Gathering these results in (90):

$$\begin{aligned} \langle Z_{t \wedge \tau_M^K}^K, x^2 \rangle^{1+\epsilon} &\leq \langle Z_0^K, x^2 \rangle^{1+\epsilon} + 2\sigma^2 \epsilon (1+\epsilon)t + (1+\epsilon)(C + \sigma^2 + 4c\sigma) \int_0^t \langle Z_{s \wedge \tau_M^K}^K, x^2 \rangle^{1+\epsilon} ds \\ &\quad + (\sigma^2 + 2c\sigma)T \sup_{s \leq T} \langle Z_s^K, 1 \rangle^{1+\epsilon} + M_{t \wedge \tau_M^K}^K \end{aligned} \quad (91)$$

where $(M_{t \wedge \tau_M^K}^K)_{t \geq 0}$ is a square integrable martingale. Taking the expectation, using Gronwall's lemma and (89) implies that:

$$\begin{aligned} \sup_{t \in [0, T]} \mathbb{E}(\langle Z_{t \wedge \tau_M^K}^K, x^2 \rangle^{1+\epsilon}) &\leq \sup_{K \in \mathbb{N}^*} \left(\mathbb{E}(\langle Z_0^K, x^2 \rangle^{1+\epsilon}) + 2\sigma^2 \epsilon (1+\epsilon)T + (\sigma^2 + 2c\sigma)T \sup_{s \leq T} \langle Z_s^K, 1 \rangle^{1+\epsilon} \right) \times \\ &\quad \times \exp(T(1+\epsilon)(C + \sigma^2 + 4c\sigma)). \end{aligned} \quad (92)$$

Because the right hand side does not depend on M for K , we obtain:

$$\sup_{K \in \mathbb{N}^*} \sup_{t \in [0, T]} \mathbb{E}(\langle Z_t^K, x^2 \rangle^{1+\epsilon}) < +\infty. \quad (93)$$

A similar computation yields that under the additional assumption (84), we also have:

$$\sup_{K \in \mathbb{N}^*} \sup_{t \in [0, T]} \mathbb{E}(\langle Z_t^K, x^4 \rangle^{1+\epsilon}) < +\infty. \quad (94)$$

Step 4: Now, let us take the supremum in (91):

$$\begin{aligned} \sup_{s \leq t} \langle Z_{s \wedge \tau_M^K}^K, x^2 \rangle^{1+\epsilon} &\leq \langle Z_0^K, x^2 \rangle^{1+\epsilon} + 2\sigma^2 \epsilon (1+\epsilon)t + (1+\epsilon)(C + \sigma^2 + 4c\sigma) \int_0^t \sup_{u \leq s} \langle Z_{u \wedge \tau_M^K}^K, x^2 \rangle^{1+\epsilon} ds \\ &\quad + (\sigma^2 + 2c\sigma)T \sup_{s \leq T} \langle Z_s^K, 1 \rangle^{1+\epsilon} + \sup_{s \leq t} M_{s \wedge \tau_M^K}^K \end{aligned} \quad (95)$$

The bracket of the martingale $(M_{t \wedge \tau_M^K}^K)_{t \geq 0}$ is:

$$\begin{aligned}
\langle M^K \rangle_{t \wedge \tau_M^K} &= \frac{4\sigma^2(1+\epsilon)^2}{K} \int_0^{t \wedge \tau_M^K} \langle Z_s^K, x^2 \rangle^{1+2\epsilon} ds \\
&\quad + \int_0^{t \wedge \tau_M^K} \sum_{i \in V_s^K} \left(\langle Z_{s-}^K, x^2 \rangle + \frac{(X_s^i)^2}{K} \right)^{1+\epsilon} - \langle Z_{s-}^K, x^2 \rangle^{1+\epsilon} ds \\
&\leq \int_0^{t \wedge \tau_M^K} \left(\frac{4\sigma^2(1+\epsilon)^2}{K} \langle Z_s^K, x^2 \rangle^{1+2\epsilon} + \frac{C^2(1+\epsilon)^2}{K} \langle Z_s^K, x^2 \rangle^{2\epsilon} \langle Z_s^K, x^4 \rangle \right) ds \\
&\leq \int_0^{t \wedge \tau_M^K} \left(\left(\frac{4\sigma^2(1+\epsilon)^2}{K} + \frac{C^2(1+\epsilon)^2}{K} \right) \langle Z_s^K, x^2 \rangle^{1+2\epsilon} + \frac{C^2(1+\epsilon)^2}{K} \langle Z_s^K, x^4 \rangle^{1+2\epsilon} \right) ds.
\end{aligned} \tag{96}$$

Thus, using Doob's lemma:

$$\begin{aligned}
\mathbb{E} \left(\sup_{s \leq t} M_s^K \right) &\leq 4\mathbb{E} \left(\int_0^{t \wedge \tau_M^K} \left(\left(\frac{4\sigma^2(1+\epsilon)^2}{K} + \frac{C^2(1+\epsilon)^2}{K} \right) \langle Z_s^K, x^2 \rangle^{1+2\epsilon} + \frac{C^2(1+\epsilon)^2}{K} \langle Z_s^K, x^4 \rangle^{1+2\epsilon} \right) ds \right) \\
&\leq \frac{C(T)}{K},
\end{aligned} \tag{97}$$

by (93) and (94). Now, taking the expectation in (95), and using Gronwall's inequality with (89) and (97) yields that

$$\sup_{K \in \mathbb{N}^*} \mathbb{E} \left(\sup_{t \in [0, T]} \langle Z_t^K, x^2 \rangle^{1+\epsilon} \right) < +\infty. \tag{98}$$

C Proof of Theorem 4.1

Proof We denote here $h(x) = 1 - \frac{x^2}{2} - \lambda$. Notice that the proof here holds for any function h that is upper bounded (but not necessarily lower bounded).

For $x \in \mathbb{R}$, $T > 0$ and $t \leq T$, let us define the following measure for a test function Φ continuous and bounded on $C([0, T], \mathbb{R})$, where Y is the diffusion process defined in (38):

$$\langle \mu_t^{T,x}, \Phi \rangle = \frac{\mathbb{E}_x \left(\exp \left(\int_0^T h(X_s) ds \right) \Phi(X_s, s \leq t) \right)}{\mathbb{E}_x \left(\exp \left(\int_0^T h(X_s) ds \right) \right)}. \tag{99}$$

Let us prove that under $\mu_t^{T,x}$, the canonical process is an inhomogeneous Markov process with infinitesimal generator (55).

Denoting $\mathbb{E}^{\mu^{T,x}}$ the expectation under $\mu^{T,x}$, we have that, for some real numbers t and u s.t. $t \geq u \geq 0$,

$$\mathbb{E}^{\mu^{T,x}} \left[f(X_t) \middle| \mathcal{F}_u \right] = \frac{\mathbb{E} \left[f(X_t) \exp \left(\int_0^T h(X_s) ds \right) \middle| \mathcal{F}_u \right]}{\mathbb{E} \left[\exp \left(\int_0^T h(X_s) ds \right) \middle| \mathcal{F}_u \right]}. \quad (100)$$

Markov property for X under \mathbb{P}_x and Formula (42) entail that

$$\mathbb{E} \left[\exp \left(\int_0^T h(X_s) ds \right) \middle| \mathcal{F}_u \right] = m_{T-u}(X_u) \exp \left(\int_0^u h(X_s) ds \right). \quad (101)$$

We also have

$$\begin{aligned} & \mathbb{E} \left[f(X_t) \exp \left(\int_0^T h(X_s) ds \right) \middle| \mathcal{F}_u \right] \\ &= \mathbb{E} \left[f(X_t) \exp \left(\int_0^T h(X_s) ds \right) \middle| \mathcal{F}_t \middle| \mathcal{F}_u \right] \\ &= \exp \left(\int_0^u h(X_s) ds \right) \mathbb{E} \left[f(X_t) \exp \left(\int_u^t h(X_s) ds \right) m_{T-t}(X_t) \middle| \mathcal{F}_u \right]. \end{aligned} \quad (102)$$

Now, as f and m are smooth (cf. Corollary 3.6) and X is a Markov process with generator L (see Lemma 3.5), we have for any $0 \leq u \leq t$,

$$f(X_t) m_{T-t}(X_t) = f(X_u) m_{T-u}(X_u) + \int_u^t (L(f m_{T-s})(X_s) - f(X_s) \partial_t m_{T-s}(X_s)) ds + M_t, \quad (103)$$

where M is some \mathbb{P} -martingale started at 0. Thus, applying Itô's formula, we get

$$\begin{aligned} f(X_t) m_{T-t}(X_t) \exp \left(\int_u^t h(X_s) ds \right) &= f(X_u) m_{T-u}(X_u) \\ &+ \int_u^t \left(f(X_s) \partial_t m_{T-s}(X_s) + L(f m_{T-s})(X_s) - f(X_s) m_{T-s}(X_s) h(X_s) \right) \exp \left(\int_u^s h(X_v) dv \right) ds \\ &+ \int_u^t \exp \left(\int_u^s h(X_v) dv \right) dM_s. \end{aligned}$$

Using (43) gives

$$\begin{aligned}
f(X_t)m_{T-t}(X_t) \exp\left(\int_u^t h(X_s)ds\right) &= f(X_u)m_{T-u}(X_u) \\
&+ \int_u^t \left(L(fm_{T-s})(X_s) - f(X_s)Lm_{T-s}(X_s)\right) \exp\left(\int_u^s h(X_v) dv\right) ds \\
&+ \int_u^t \exp\left(\int_u^s h(X_v) dv\right) dM_s.
\end{aligned}$$

Using (101) and notation (55), we finally obtain

$$\begin{aligned}
&\mathbb{E}\left[f(X_t)m_{T-t}(X_t) \exp\left(\int_0^t h(X_s)ds\right) \middle| \mathcal{F}_u\right] \\
&= f(X_u)m_{T-u}(X_u) \exp\left(\int_0^u h(X_s) ds\right) \\
&\quad + \mathbb{E}\left[\int_u^t \mathcal{G}_s f(X_s)m_{T-s}(X_s) \exp\left(\int_0^s h(X_v) dv\right) ds \middle| \mathcal{F}_u\right] \\
&= f(X_u)\mathbb{E}\left[\exp\left(\int_0^T h(X_s) ds\right) \middle| \mathcal{F}_u\right] \\
&\quad + \mathbb{E}\left[\int_u^t \mathcal{G}_s f(X_s)\mathbb{E}\left[\exp\left(\int_0^T h(X_v) dv\right) \middle| \mathcal{F}_s\right] ds \middle| \mathcal{F}_u\right].
\end{aligned}$$

Thus, using (100) and (102), we have

$$\mathbb{E}^{\mu^{T,x}}\left[f(X_t) \middle| \mathcal{F}_u\right] = f(X_u) + \mathbb{E}^{\mu^{T,x}}\left[\int_u^t \mathcal{G}_s f(X_s) \middle| \mathcal{F}_u\right].$$

This ends the proof. □