Ergodicity of the Fisher infinitesimal model with quadratic selection

Vincent Calvez, Thomas Lepoutre, David Poyato

ERGODICITY OF THE FISHER INFINITESIMAL MODEL WITH QUADRATIC SELECTION

VINCENT CALVEZ, THOMAS LEPOUTRE, AND DAVID POYATO

Abstract. We study the convergence towards a unique equilibrium distribution of the solutions to a time-discrete model with non-overlapping generations arising in quantitative genetics. The model describes the dynamics of a phenotypic distribution with respect to a multi-dimensional trait, which is shaped by selection and Fisher's infinitesimal model of sexual reproduction. We extend some previous works devoted to the time-continuous analogues, that followed a perturbative approach in the regime of weak selection, by exploiting the contractivity of the infinitesimal model operator in the Wasserstein metric. Here, we tackle the case of quadratic selection by a global approach. We establish uniqueness of the equilibrium distribution and exponential convergence of the renormalized profile. Our technique relies on an accurate control of the propagation of information across the large binary trees of ancestors (the pedigree chart), and reveals an ergodicity property, meaning that the shape of the initial datum is quickly forgotten across generations. We combine this information with appropriate estimates for the emergence of Gaussian tails and propagation of quadratic and exponential moments to derive quantitative convergence rates. Our result can be interpreted as a generalization of the Krein-Rutman theorem in a genuinely non-linear, and non-monotone setting.

1. Introduction

Fisher's infinitesimal model (also known as the *polygenic model*) is a widely used statistical model in quantitative genetics initially proposed by R. Fisher [15]. It assumes that the genetic component of a quantitative phenotypical trait is affected by an infinite number of loci with infinitesimal and additive allelic effects and claims that the genetic component of descendants' traits is normally distributed around the mean value of parents' traits, with a constant (genetic) variance across generations. This model allowed reconcilling Mendelian inheritance and the continuous trait variations documented by F. Galton via a Central Limit Theorem. More specifically, by taking limits when the number of underlying loci tends to infinity on a model with Mendelian inheritance, N. Barton, A. Etheridge and A. Véber [5] recently proved rigorously the validity of Fisher's infinitesimal model under various evolutionary processes (*e.g.*, natural selection).

In this paper we study a time-discrete evolution problem for the distribution of a phenotypical trait $x \in \mathbb{R}^d$ in a population undergoing sexual reproduction and the effect of natural selection. Specifically, starting at any initial configuration $F_0 \in \mathcal{M}_+(\mathbb{R}^d)$ of trait distribution, we analyze the long term dynamics of trait distributions $\{F_n\}_{n\in\mathbb{N}}$ across successive generations $n \in \mathbb{N}$, which solves the following recursion

$$F_n = \mathcal{T}[F_{n-1}], \tag{1.1}$$

for any $n \in \mathbb{N}$. The operator $\mathcal{T}$ encodes the balanced effect of sexual reproduction in the population (under the infinitesimal model) and natural selection. Specifically, $\mathcal{T}$ is defined by

$$\mathcal{T}[F] := e^{-m}\mathcal{B}[F], \tag{1.2}$$

for any trait distribution $F \in \mathcal{M}_+(\mathbb{R}^d)$. On the one hand, $m = m(x) \geq 0$ is called the *selection function* and represents the mortality effect of a trait-dependent natural selection on the population, so that e^{-m} stands for the survival probability in the next generation. On the other hand, the operator $\mathcal{B}$ is chosen

Date: January 11, 2024.

2020 *Mathematics Subject Classification.* 35B40; 35P30; 35Q92; 47G20; 92D15.

Key words and phrases. Integro-differential equations, asymptotic behavior, nonlinear spectral theory, quantitative genetics.

Acknowledgment. VC and DP have received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement No 865711). DP has also received founding from the European Union's Horizon Europe research and innovation program under the Marie Skłodowska-Curie grant agreement No 101064402, and partially from the State Research Agency (SRA) of the Spanish Ministry of Science and Innovation and European Regional Development Fund (ERDF), project PID2022-137228OB-I00, and by Modeling Nature Research Unit, project QUAL21-011.

to be *Fisher's infinitesimal operator* and it takes the form

$$\mathcal{B}[F](x) := \int_{\mathbb{R}^{2d}} G\left(x - \frac{x_1+x_2}{2}\right) \frac{F(x_1)F(x_2)}{\int_{\mathbb{R}^d} F(x')\,dx'}\,dx_1\,dx_2, \quad x \in \mathbb{R}^d, \tag{1.3}$$

for any trait distribution $F \in \mathcal{M}_+(\mathbb{R}^d)$. Here $G = G(x)$ is a probability density (the *mixing kernel*) and the factor $G(x - \frac{x_1+x_2}{2})$ represents the transition probability that two given individuals with trait values $x_1, x_2 \in \mathbb{R}^d$ will mate and yield a descendant with trait value $x \in \mathbb{R}^d$. In other words the resulting trait distributes around the mean value $\frac{x_1+x_2}{2}$ of parents' trait with law G. By definition, $\mathcal{B}[F]$ quantifies the number of births after all possible matches of any couple of individuals according to the trait distribution F. Altogether, $\mathcal{T}[F]$ quantifies the amount of offspring of a population distributed according to F having resisted the effect of selection.

The above sexual reproduction operator $\mathcal{B}$ has recently pulled the attention of both the applied and more theoretical communities, *cf.* [5, 10, 18, 19, 20, 21]. In this paper we shall restrict to Gaussian mixing kernel and quadratic selection function, *i.e.*,

$$G(x) := \frac{1}{(2\pi)^{d/2}} e^{-\frac{|x|^2}{2}}, \quad x \in \mathbb{R}^d, \tag{1.4}$$

$$m(x) := \frac{\alpha}{2}|x|^2, \qquad x \in \mathbb{R}^d, \tag{1.5}$$

where $\alpha \in \mathbb{R}_+$ is a fixed parameter. Thereby, the trait of offsprings is normally distributed around the mean value of the trait of parents by assumption (1.4), thus reducing to the standard infinitesimal model when the assumptions of the Central Limit Theorem are met [5]. For simplicity, we set a Gaussian G with unit variance, but any value of the genetic variance could also be considered (see nondimensionalization in Appendix A).

Before introducing our results, we shall relate the previous time-discrete problem (1.1) to analogous time-continuous quantitative genetics models of evolutionary dynamics that have been studied in the literature. Meanwhile, we will anticipate the major difficulties that can be faced when analyzing the long-time dynamics of (1.1). To this end, we consider, the following type of integro-differential equations for the evolutionary dynamics of a trait distribution $f(t, x)$:

$$\begin{cases} \partial_t f = -m(x)f + \mathcal{R}[f], & t \geq 0,\, x \in \mathbb{R}^d, \\ f(0,x) = f_0(x), & x \in \mathbb{R}^d, \end{cases} \tag{1.6}$$

where, $m = m(x)$ is again the mortality rate and $\mathcal{R} = \mathcal{R}[f]$ is the reproduction operator. Hence, $\mathcal{R}[f](x)$ determines the amount of births with trait value $x \in \mathbb{R}^d$ per unit of time. As for (1.1), the resulting dynamics of the population becomes a consequence of a balance between selection encoded by the trait-dependent mortality and the diversity generated by the growth term $\mathcal{R}$ across generations.

Many studies consider a *linear* reproduction operator $\mathcal{R}$, associated with a probability density $K = K(x)$ characterizing the mutational effects at birth, of the form

$$\mathcal{R}[f](x) := \int_{\mathbb{R}^d} K(x-y)f(t,y)\,dy, \quad x \in \mathbb{R}^d. \tag{1.7}$$

The factor $K(x - y)$ determines the probability that an individual with trait value $y \in \mathbb{R}^d$ produces a descendant with trait value $x \in \mathbb{R}^d$ (possibly deviating from y). This class of linear reproduction operators is well-suited for an asexual mode of reproduction. This includes parabolic equations in the limit of small variance of K. In particular, we refer to a series of works about the long-time asymptotics in the regime of small variance, initiated in [3, 4, 14], including an additional density-dependent competition term that makes the analysis non standard (see *e.g.* [11] and references therein for the well-posedness of the constrained Hamilton-Jacobi equation derived in the limit).

Recently, inspired by the infinitesimal model, the case of *sexual reproduction* has been addressed by invoking the preceding nonlinear version $\mathcal{B}$ in (1.3) as reproduction operator $\mathcal{R} = \mathcal{B}$. Several asymptotic regimes have been addressed: large reproduction rate [18, 20, 21], small variance asymptotics [10, 19]. In the latter case, the limiting problem keeps the non-local nature of the problem, being of a finite-difference type, rather than a Hamilton-Jacobi PDE. Before we continue the discussion about the state-of-the-art, let us emphasize that there is no restriction on the parameter α in the present work.

Since both $f \mapsto mf$ and $f \mapsto \mathcal{B}[f]$ are 1-homogeneous operators, we can seek for special steady solutions of (1.3) through the following ansatz:

$$f(t,x) = e^{\lambda t}F(x), \quad (t,x) \in \mathbb{R}_+ \times \mathbb{R}^d, \tag{1.8}$$

The parameter $\lambda \in \mathbb{R}$ represents the rate at which the number of individuals grows (if $\lambda \geq 0$) or decreases (if $\lambda \leq 0$), and $F = F(x) \geq 0$ is an unknown probability density. By imposing such an ansatz on (1.6)-(1.3), the following generalized spectral problem arises for the couple (λ, F):

$$\begin{cases} \lambda F(x) + m(x)F(x) = \mathcal{B}[F](x), \quad x \in \mathbb{R}^d, \\ \int_{\mathbb{R}^d} F(x')\,dx' = 1. \end{cases} \tag{1.9}$$

Note that the operator $\mathcal{B}$ is genuinely non-linear so that methods based on the Krein-Rutman theory or maximum principles cannot be applied straightforwardly, see [6, 13] and the references therein for the linear case. Further, usual extensions of the Krein-Rutman theory to 1-homogeneous operators [16] cannot be applied neither because $\mathcal{B}$ is not monotone. To date, the main strategy behind the existence of solutions of (1.9) relies on a suitable application of Schauder fixed-point theorem to the operator $F \mapsto (\lambda + m)^{-1}\mathcal{B}[F]$ over an appropriate cone of $L^1(\mathbb{R}^d)$ that is conserved by the nonlinear operator, see [8]. However, uniqueness cannot be achieved by this method. Moreover, it has been proven in [10, Corollary 1.5] that several equilibrium states (λ, F) can co-exist in the presence of multiple local minima of m (provided the variance of G is small enough). That is, the generalized eigenproblem (1.9) does not admit a unique positive eigenfunction, in contrast with general conclusions of the Krein-Rutman theory.

Recently, G. Raoul addressed the long-time dynamics of (1.6) in 1D, with $\mathcal{R} = \mathcal{B}$ and a trait dependent fecundity rate [21]. He obtained local uniqueness and exponential relaxation under the assumption of weak and localized (compactly supported) selection effects. To this end, he controlled locally in space the Wasserstein distance between the solution and the stationary state, using the uniform contraction property of $\mathcal{B}$ in the space of probability measures sharing the same center of mass. Unfortunately, the Wasserstein metric is not fully compatible with multiplicative operators, such as trait-dependent fecundity (see the discussion in [21, Section 3.4] and Section 2 below). This can be circumvented under the additional assumption that the trait density is locally uniformly bounded below, following [7]. Obviously, this cannot hold globally for integrable densities, hence G. Raoul developed estimates of the distribution's tail to complete the contraction estimates. Also, a lot of attention has to be paid to the dynamics of the center of the distribution which is essentially driven by selection. Indeed, in the case of flat selection ($m \equiv 0$), the problem is invariant by translation, so that local uniqueness cannot hold. To conclude this discussion, let us emphasize that global uniqueness and the asymptotic behavior of generic solutions to the evolution problem (1.6)-(1.3) is still open. Below, we provide a first result in this direction, for the time-discrete problem (1.1), though.

We remark that the time-discrete version (1.1) that we propose in this paper can be partially regarded as a discretization in time of the above time-continuous problem (1.6)-(1.3), with non-overlapping generations (see Appendix A for further details). As for the time-continuous problem, we could seek special solutions to the time-discrete problem (1.1) in the following form

$$F_n(x) = \lambda^n F(x), \quad (n, x) \in \mathbb{N} \times \mathbb{R}^d. \tag{1.10}$$

Again, $\lambda \in \mathbb{R}_+^*$ is the rate of growth (if $\lambda \geq 1$) or decrease (if $\lambda \leq 1$) of individuals, and $F = F(x)$ is an unknown probability density. This yields the following generalized eigenproblem for the couple (λ, F):

$$\begin{cases} \lambda F(x) = \mathcal{T}[F](x), \quad x \in \mathbb{R}^d, \\ \int_{\mathbb{R}^d} F(x')\,dx' = 1. \end{cases} \tag{1.11}$$

In this paper we aim to address the following questions:

- **(Q1)** Does the eigenproblem (1.11) have a unique solution $(\boldsymbol{\lambda}_\alpha, \boldsymbol{F}_\alpha)$, with $\boldsymbol{\lambda}_\alpha \in \mathbb{R}$ and $\boldsymbol{F}_\alpha$ being a probability measure, for each $\alpha \in \mathbb{R}_+^*$?
- **(Q2)** Consider any generic initial datum $F_0 \in \mathcal{M}_+(\mathbb{R}^d)$ and its associated solution $\{F_n\}_{n\in\mathbb{N}}$ of the time-discrete problem (1.1). Do the renormalized profiles $F_n/\|F_n\|_{L^1(\mathbb{R}^d)}$ converge to the unique steady profile $\boldsymbol{F}_\alpha$ solving (1.11) when $n \to \infty$?

We shall prove that the answer to both questions is affirmative. It stands to reason that similar existence and local uniqueness results like in [8, 10] could be extended to the new eigenproblem (1.11) by applying Schauder and Banach fixed point theorems. Nevertheless, in this paper we introduce a novel method that unravels an ergodicity property of the operator $\mathcal{T}$, leading to quantitative estimates for the relaxation of profiles $\{F_n\}_{n\in\mathbb{N}}$ towards $\boldsymbol{F}_\alpha$. First, we prove that an explicit (Gaussian) solution to the eigenproblem (1.11) exists. Second, by computing n iterations of the operator $\mathcal{T}$, that is $F_n = \mathcal{T}^n[F_0]$, we notice that information of F_n at the trait value x is propagated from the initial datum F_0 across 2^n ancestors over a binary tree with height n and rooted at x (the pedigree chart). Interestingly, an appropriate reformulation of $\mathcal{T}^n$ in the case of Gaussian mixing (1.4) and quadratic selection (1.5) shows that the dependence of the solution $\{F_n\}_{n\in\mathbb{N}}$ on the initial datum F_0 is rapidly lost across the different

levels of the tree. More specifically, a strong convergence of generic solutions $\{F_n\}$ of the time-discrete problem (1.1) towards the steady Gaussian profile $\boldsymbol{F}_\alpha$ solving (1.11) is achieved locally with respect to x. Third, we prove an appropriate propagation of quadratic and exponential moments, leading to uniform tightness of the family $\{F_n\}_{n\in\mathbb{N}}$. Finally, we glue all the information together and conclude the final global convergence result in question **(Q2)** in relative entropy. We refer to Section 2 for a more detailed sketch of our strategy of proof. Specifically, we obtain our main result:

Theorem 1.1. *Assume that $\alpha \in \mathbb{R}_+^*$ and set any initial non-negative measure $F_0 \in \mathcal{M}_+(\mathbb{R}^d)$. The solution $\{F_n\}_{n\in\mathbb{N}}$ to the time-discrete problem (1.1) verifies that the growth rates $\|F_n\|_{L^1(\mathbb{R}^d)}/\|F_{n-1}\|_{L^1(\mathbb{R}^d)}$ relax towards $\boldsymbol{\lambda}_\alpha$ and the normalized profiles $F_n/\|F_n\|_{L^1(\mathbb{R}^d)}$ relax towards $\boldsymbol{F}_\alpha$ with*

$$\boldsymbol{\lambda}_\alpha := \left(1+\alpha\left(1+\frac{\boldsymbol{\sigma}_\alpha^2}{2}\right)\right)^{-\frac{d}{2}}, \quad \boldsymbol{F}_\alpha := G_{0,\boldsymbol{\sigma}_\alpha^2}, \tag{1.12}$$

and the variance $\boldsymbol{\sigma}_\alpha^2 \in \mathbb{R}_+^*$ is the unique positive root of the equation

$$\frac{1}{\boldsymbol{\sigma}_\alpha^2} = \alpha + \frac{1}{1+\frac{\boldsymbol{\sigma}_\alpha^2}{2}}, \quad \textit{i.e.,} \quad \boldsymbol{\sigma}_\alpha^2 = \frac{\sqrt{(1+2\alpha)^2+8\alpha}-(1+2\alpha)}{2\alpha}. \tag{1.13}$$

Specifically, for any $\varepsilon \in \mathbb{R}_+^*$ there exists a sufficiently large $C_\varepsilon \in \mathbb{R}_+^*$ such that

$$\mathcal{D}_{\mathrm{KL}}\left(\frac{F_n}{\|F_n\|_{L^1(\mathbb{R}^d)}}\middle\|\boldsymbol{F}_\alpha\right) \le C_\varepsilon((2\,\boldsymbol{k}_\alpha)^2+\varepsilon)^n,$$
$$\left|\frac{\|F_n\|_{L^1(\mathbb{R}^d)}}{\|F_{n-1}\|_{L^1(\mathbb{R}^d)}} - \boldsymbol{\lambda}_\alpha\right| \le C_\varepsilon(2\,\boldsymbol{k}_\alpha+\varepsilon)^n,$$

for any $n\in\mathbb{N}$, where $\mathcal{D}_{\mathrm{KL}}$ is the Kullback-Leibler divergence (or relative entropy), i.e.,

$$\mathcal{D}_{\mathrm{KL}}(P\|Q) := \int_{\mathbb{R}^d} P(x)\log\left(\frac{P(x)}{Q(x)}\right)dx, \tag{1.14}$$

for any $P,Q \in L^1_+(\mathbb{R}^d)\cap\mathcal{P}(\mathbb{R}^d)$, and the coefficient $\boldsymbol{k}_\alpha \in (0,\frac{1}{2})$ reads

$$\boldsymbol{k}_\alpha = \frac{\boldsymbol{\sigma}_\alpha^2}{2+\boldsymbol{\sigma}_\alpha^2}, \quad \textit{i.e.,} \quad \frac{(3+2\alpha)-\sqrt{(1+2\alpha)^2+8\alpha}}{4}. \tag{1.15}$$

Note that any solution (λ, F) to the eigenproblem (1.11) yields a solution to the time-discrete problem (1.1) through the ansatz (1.10). Then, question **(Q1)** will readily follow from question **(Q2)** and, in particular, the unique solution $(\boldsymbol{\lambda}_\alpha, \boldsymbol{F}_\alpha)$ to (1.11) is Gaussian.

Corollary 1.2. *Assume that $\alpha \in \mathbb{R}_+^*$, then there exists a unique solution $(\boldsymbol{\lambda}_\alpha, \boldsymbol{F}_\alpha)$ to the eigenproblem (1.11), with $\boldsymbol{F}_\alpha$ being a probability measure, namely given by (1.12).*

Remark 1.3 (About the assumption of quadratic selection)**.** *By the assumption of a quadratic selection function, we can henceforth push extensively explicit computations of the iterated operator. The latter consists in recursive multiplication and convolution by Gaussian functions. This enables capturing the essence of the relaxation phenomenon, and having a precise description of the behaviour at infinity, which crucially helps to localize the convergence argument. As a by-product, we are able to consider very general initial condition F_0. Two of the authors, together with F. Santambrogio, have obtained similar results when the selection function m is more generally assumed to be strongly convex* [12]*. However, they imposed stringent conditions on the initial datum F_0, that is, it should behave at infinity as the equilibrium profile $\mathbf{F}_\alpha$ in a very strong sense. It would be of interest to merge the two works, having a general (strongly convex) selection function, and a general initial datum. This is left for future work.*

Remark 1.4 (About the choice of metric)**.** *The convergence of the profiles has been quantified in Kullback-Leibler divergence* (1.14) *in Theorem 1.1, in contrast with the results in* [20, 21]*, where the quadratic Wasserstein distance was used for perturbative regimes of the case without selection. We anticipate that there are several compelling reasons for such a change of metric in our non-perturbative setting:*

(i) **(Quadratic Wasserstein distance)** *When $\alpha = 0$, the operator $\mathcal{T}$ reduces to $\mathcal{B}$, which is non-expansive with respect to the quadratic Wasserstein distance, and indeed contractive over distributions with common center of mass, cf. Section 2.1. However, when $\alpha > 0$, the multiplicative operator leads to an operator $\mathcal{T}$ which is not even Lipschitz continuous with respect to the quadratic Wasserstein distance, cf Section 2.2. Therefore, the quadratic Wasserstein distance seems to be unadapted to scenarios where reproduction and selection operate together.*

(ii) **(Log-Lipschitz norm)** *As we show in Section 2.3, the operator $\mathcal{T}$ is non-expansive in the log-Lipschits norm $\|\nabla \log \frac{F}{\boldsymbol{F}_\alpha}\|_{L^\infty(\mathbb{R}^d)}$ for all $\alpha \geq 0$, and indeed it is contractive if $\alpha > 0$. However, this contraction only gives actual information when the initial datum F_0 has identical tails to the Gaussian density $\boldsymbol{F}_\alpha$ so that initially the log-Lipschitz norm is finite.*

(iii) **(Relative entropy distance)** *Note that the above log-Lipschit norm amounts to the natural L^∞ version of the relative Fisher information $\|\nabla \log \frac{F}{\boldsymbol{F}_\alpha}\|_{L^2(\mathbb{R}^d,F)}$. By the log-Soboled inequality, contraction of the log-Lipschits norm readily implies decay of the Kullback-Leibler divergence, which is a more standard metric in relative entropy arguments. However, we emphasize that our use of the Kullback-Leibler divergence is not only aesthetic, but we actually need it in order to go beyond the above structural constraint on the tails of the initial data. Specifically, for generic initial data we need to prove an appropriate shaping of tails over time, which cannot be expressed using uniform norms, but only in an averaged sense (compatible with the relative entropy).*

Altogether justifies that the whilst the quadratic Wasserstein distance is useful in perturbative regimes, the use of alternative norms is necessary to quantify contraction in purely non-perturbative settings.

Remark 1.5 (About the positivity of α)**.** *Our form of ergodicity, as measured in Theorem 1.1, it breaks down when $\alpha = 0$, simply because the operator is invariant by translation in that case, and it admits a one-parameter family $\boldsymbol{F}_{\alpha=0}(\cdot - \mu)$ with $\mu \in \mathbb{R}^d$ of fixed Gaussian probability densities. Nevertheless, as mentioned in item (i) above, when $\alpha = 0$ and we restrict to centered initial data, there is convergence to the right centered Gaussian probability density $\boldsymbol{F}_{\alpha=0}$ with respect to the Wasserstein distance. It is an open problem to make both approaches meet for $\alpha = 0$, and prove contraction in norms stronger than the Wasserstein distance, but comparable to the log-Lipschitz norm, in this subclass of initial data.*

The rest of the paper is organized as follows. In Section 2 we discuss about the generic incompatibility of the Wasserstein distance with a multiplicative operator and we provide a brief outlook of the strategy of our proof. In Section 3 we provide some necessary notation and we introduce the special class of Gaussian solutions of both problems (1.1) and (1.11), which will inspire some parts of the paper. Section 4 is devoted to introduce some main properties of $\mathcal{T}$ regarding the emergence of Gaussian behavior (in the large) from generic initial data, and a suitable propagation of quadratic and exponential moments across generations. In Section 5 we reformulate (1.1) via a high-dimensional integral operator propagating ancestors' information across the different levels of the pedigree chart, which will be the cornerstone to study the long-term dynamics. In Section 6 we prove our main results, namely, Theorem 1.1 and Corollary 1.2. Section 7 contains some numerical experiments that illustrate the results in this paper. In Section 8 we provide some conclusions and perspectives. Finally, Appendix A contains the adimensionalization of the problem, and the relationship between (1.1) and the previous time-continuous analogues in the literature is discussed. A full list of the main notations in the paper is presented in Table 1.

2. Motivation and strategy of the proof of Theorem 1.1

In this section, we discuss the incompatibility of the quadratic Wasserstein distance to quantify directly contractivity under the joint effect of the reproduction operator $\mathcal{B}$ in (1.3) and a generic multiplicative selection e^{-m}. Although a small perturbation of the case of flat selection ($m \equiv 0$) could still be considered via a perturbative argument (see discussion above, [21] and also [20]), our novel approach is able to tackle a purely non-perturbative setting. We end the section by briefly discussing the strategy of our proof.

2.1. Some properties of the sexual reproduction operator.

We start by recalling some of the main properties of the sexual reproduction operator $\mathcal{B}$. Since the fecundity rate has been normalized to 1 (see Appendix A), then $\mathcal{B}$ preserves the mass and center of mass, namely,

$$\|\mathcal{B}[F]\|_{L^1(\mathbb{R}^d)} = \|F\|_{\mathcal{M}(\mathbb{R}^d)}, \quad \int_{\mathbb{R}^d} x\, \mathcal{B}[F](x)\, dx = \int_{\mathbb{R}^d} x\, F(dx), \tag{2.1}$$

for any $F \in \mathcal{M}_+(\mathbb{R}^d)$. Furthermore, it is contractive in the space of probability measures with a common center of mass, endowed with the quadratic Wasserstein metric. As discussed above, this property has been used fruitfully by G. Raoul (*cf.* [21]) to analyse the long term behavior of the time-continuous problem in the regime of weak (and compactly supported) selection acting on fecundity. For the sake of clarity, we recall this fact and its proof below:

Lemma 2.1. *Assume that $F_1, F_2 \in \mathcal{P}_2(\mathbb{R}^d)$ and they have the same center of mass. Then,*

$$\mathcal{W}_2^2(\mathcal{B}[F_1], \mathcal{B}[F_2]) \leq \frac{1}{2}\, \mathcal{W}_2^2(F_1, F_2)\,.$$

Notation	Meaning	Page
$\mathcal{M}(\mathbb{R}^d)$, $\mathcal{M}_+(\mathbb{R}^d)$	Signed and non-negative finite Radon measures	4
$\mathcal{P}(\mathbb{R}^d)$, $\mathcal{P}_2(\mathbb{R}^d)$	Probability measures (with finite 2nd order moment)	5
$P \otimes Q \in \mathcal{M}(\mathbb{R}^{2d})$	Tensor product of the finite measures $P, Q \in \mathcal{M}(\mathbb{R}^d)$	43
$\mathcal{D}_{\mathrm{KL}}(P\|Q)$	Kullback-Leibler divergence between $P, Q \in \mathcal{P}(\mathbb{R}^d)$	4
$\mathcal{W}_2(P,Q)$	Quadratic Wasserstein distance between $P, Q \in \mathcal{P}_2(\mathbb{R}^d)$	6
G_{μ,σ^2}	Gaussian with mean $\mu \in \mathbb{R}^d$ and covariance $\sigma^2 I_d \in \mathbb{R}^{d\times d}$	13
$G_{\mu,\Sigma}$	Gaussian with mean $\mu \in \mathbb{R}^d$ and covariance $\Sigma \in \mathbb{R}^{d\times d}$	33
$\mathcal{N}(\mu,\Sigma)$	Normal with mean $\mu \in \mathbb{R}^d$ and covariance $\Sigma \in \mathbb{R}^{d\times d}$	33
$\mathbb{E}[X]$	Expectation of a random variable X	34
$G(x)$	Gaussian mixing kernel G_{0,I_d}	2
$m(x)$	Quadratic selection function $\frac{\alpha}{2}\|x\|^2$	2
$\mathcal{B}[F]$	Fisher's infinitesimal operator	2
$\mathcal{T}[F]$	Selection-reproduction operator	1
$\mathcal{M}[F]$	Normalized multiplicative operator by e^{-m}	8
$\mathcal{S}[F]$	Scaled selection-reproduction operator	21
$\mathcal{E}_n[F]$	High-dimensional integral operators	35
$\{F_n\}_{n\in\mathbb{N}}$	Solution to the time-evolution equation (1.1)	1
$(\boldsymbol{\lambda}_\alpha, \boldsymbol{F}_\alpha)$	Gaussian solution to the non-linear eigenproblem (1.11)	4
$\boldsymbol{\sigma}^2_\alpha$	Variance of Gaussian eigenfunction $\boldsymbol{F}_\alpha$	4
$\bar{F} = \frac{e^m F}{\boldsymbol{F}_{\alpha=0}}$	Normalized profile associated to $F \in \mathcal{M}_+(\mathbb{R}^d)$	26
T^n, T^n_*, $\widehat{\mathsf{T}}^n$	Perfect binary tree, rootless tree and leafless tree	11
L^n_m, L^n	Level m and leaves (level n) of the tree	11
$i1, i2$	Parents of a node $i \in \widehat{\mathsf{T}}^n$ of the tree	11
$\mathbf{x}_n = (x_i)_{i\in\mathsf{T}^n_*}$ $\mathbf{y}_n = (y_i)_{i\in\mathsf{T}^n_*}$	Variables indexed by the rootless tree	13
$\mathbf{z}_n = (z_j)_{j\in\mathsf{L}^n}$	Variables indexed by leafs	13
$\|\mathbf{x}\| = \left(\sum_{i=1}^n \|x_i\|^2\right)^{1/2}$	ℓ_2 sum of Euclidean norms of $\mathbf{x} = (x_1,\dots,x_n) \in \mathbb{R}^{dn}$	36
$\Phi^j_n(x;\mathbf{y}_n)$	Lineage map from leaf $j \in \mathsf{L}^n$ to root value $x \in \mathbb{R}^d$	30
$u \otimes v \in \mathbb{R}^{d\times d}$	Kronecker product $(u_i v_j)_{1\le i,j\le N}$ of vectors $u, v \in \mathbb{R}^d$	34
k_n, κ_n	Sequences of coefficients in the change of variables	28
$(2\,\boldsymbol{k}_\alpha)^2$	Convergence rate of $\mathcal{D}_{\mathrm{KL}}$	4
$2\,\boldsymbol{k}_\alpha$	Convergence rate of the log-Lipschitz norm	8
$\boldsymbol{r}_\alpha$	Relaxation rates of variances recursion	15

TABLE 1. List of notations

Proof. Recall the following dual characterizations of the quadratic Wasserstein distance (*cf.* [1, 24]):

$$\begin{aligned}\mathcal{W}_2^2(F_1,F_2) &= \inf\left\{\int_{\mathbb{R}^{2d}} |x-y|^2\ \gamma(dx,dy) :\ \gamma\in\mathcal{P}(\mathbb{R}^d\times\mathbb{R}^d),\ \pi_{1\#}\gamma = F_1,\ \pi_{2\#}\gamma = F_2\right\}\\ &= \sup\left\{\int_{\mathbb{R}^d}\phi_1\,dF_1 + \int_{\mathbb{R}^d}\phi_2\,dF_2 :\ |\phi_1(x)+\phi_2(y)| \le |x-y|^2 \quad \forall\, x,y\in\mathbb{R}^d\right\},\end{aligned} \tag{2.2}$$

where $\pi_i : \mathbb{R}^d\times\mathbb{R}^d \to \mathbb{R}^d$ is the projection onto the i-th component for $i = 1, 2$. Taking any couple (ϕ_1,ϕ_2) as above and using the specific form of the operator $\mathcal{B}$ we obtain:

$$\begin{aligned}&\int_{\mathbb{R}^d}\mathcal{B}[F_1](x)\,\phi_1(x)\,dx + \int_{\mathbb{R}^d}\mathcal{B}[F_2](y)\,\phi_2(y)\,dy\\ &= \int_{\mathbb{R}^{3d}} G\left(x-\frac{x_1+x_2}{2}\right)\phi_1(x)\,F_1(dx_1)\,F_1(dx_2)\,dx + \int_{\mathbb{R}^{3d}} G\left(y-\frac{y_1+y_2}{2}\right)\phi_2(y)\,F_2(dy_1)\,F_2(dy_2)\,dy\\ &= \int_{\mathbb{R}^{3d}} G(z)\,\phi_1\left(z+\frac{x_1+x_2}{2}\right)F_1(dx_1)\,F_1(dx_2)\,dz + \int_{\mathbb{R}^{3d}} G(z)\,\phi_2\left(z+\frac{y_1+y_2}{2}\right)F_2(dy_1)\,F_2(dy_2)\,dz.\end{aligned}$$

Consider any transference plan $\gamma \in \mathcal{P}(\mathbb{R}^d \times \mathbb{R}^d)$ between (F_1, F_2) as above. Specifically, we have that $\pi_{1\#}\gamma = F_1$ and $\pi_{2\#}\gamma = F_2$. Then, we can gather both integrals as follows:

$$\int_{\mathbb{R}^d} \mathcal{B}[F_1](x)\,\phi_1(x)\,dx + \int_{\mathbb{R}^d} \mathcal{B}[F_2](y)\,\phi_2(y)\,dy \\ = \int_{\mathbb{R}^{5d}} G(z)\left(\phi_1\left(z + \frac{x_1+x_2}{2}\right) + \phi_2\left(z + \frac{y_1+y_2}{2}\right)\right)\gamma(dx_1, dy_1)\,\gamma(dx_2, dy_2)\,dz.$$

Since the condition $|\phi_1(x) + \phi_2(y)| \le |x-y|^2$ is verified for all $x, y \in \mathbb{R}^d$, then we find

$$\begin{aligned}\int_{\mathbb{R}^d} \mathcal{B}[F_1](x)\,\phi_1(x)\,dx + \int_{\mathbb{R}^d} \mathcal{B}[F_2](y)\,\phi_2(y)\,dy \\ &\le \frac{1}{4}\int_{\mathbb{R}^{5d}} G(z)\,|(x_1+x_2)-(y_1+y_2)|^2\,\gamma(dx_1, dy_1)\,\gamma(dx_2, dy_2)\,dz \\ &= \frac{1}{4}\int_{\mathbb{R}^{4d}} |(x_1+x_2)-(y_1+y_2)|^2\,\gamma(dx_1, dy_1)\,\gamma(dx_2, dy_2) \\ &= \frac{1}{4}\int_{\mathbb{R}^{2d}} |x_1-y_1|^2\,\gamma(dx_1, dy_1) + \frac{1}{4}\int_{\mathbb{R}^{2d}} |x_2-y_2|^2\,\gamma(dx_2, dy_2),\end{aligned}$$

where in the third line we have used that G is a probability density and in the last line we have used the crucial fact that F_1 and F_2 share the same center of mass in order to cancel the cross-terms (otherwise the estimate would boil down to a non-expansiveness estimate). Taking supremum over (ϕ_1, ϕ_2) and infimum over γ and using the dual characterizations (2.2) yields the result. □

Note that the fact that both F_1 and F_2 have the same center of mass has been crucially used to cancel the crossed term. Otherwise, by the Cauchy–Schwarz inequality we would merely obtain non-expansiveness:

$$\mathcal{W}_2(\mathcal{B}[F_1], \mathcal{B}[F_2]) \le \mathcal{W}_2(F_1, F_2). \tag{2.3}$$

Using the contractivity property of Lemma 2.1 along with the conservation of mass and center of mass in (2.1) yields the long term dynamics of (1.1) in the special case $\alpha = 0$.

Corollary 2.2. *Set any initial datum $F_0 \in \mathcal{M}_+(\mathbb{R}^d)$ such that $\int_{\mathbb{R}^d} |x|^2 F_0(dx) < \infty$ and consider the solution $\{F_n\}_{n\in\mathbb{N}}$ to the time-discrete problem* (1.1) *with $\alpha = 0$,* i.e., *$F_n = \mathcal{B}[F_{n-1}]$ for all $n \in \mathbb{N}$. Then,*

$$\frac{\|F_n\|_{L^1(\mathbb{R}^d)}}{\|F_{n-1}\|_{L^1(\mathbb{R}^d)}} = 1, \quad \mathcal{W}_2\left(\frac{F_n}{\|F_n\|_{L^1(\mathbb{R}^d)}}, G_{\mu_0,2}\right) \lesssim \frac{1}{2^{n/2}},$$

for every $n \in \mathbb{N}$, where $\mu_0 := \int_{\mathbb{R}^d} x\,F_0(dx)$. In particular, the set of stationary distributions under $\mathcal{B}$ is $\{\boldsymbol{F}_{\alpha=0}(\cdot - \mu) : \mu \in \mathbb{R}^d\}$, where $\boldsymbol{F}_{\alpha=0} = G_{0,2}$ will denote here on the Gaussian centered at the origin with variance equals 2 in agreement with the notation (1.12) *in Theorem 1.1.*

The preceding result might be regarded as the counterpart of Theorem 1.1 for $\alpha = 0$. Indeed, by Talagrand transportation inequality for a Gaussian measure [23] we obtain the relation

$$\mathcal{W}_2^2\left(\frac{F_n}{\|F_n\|_{L^1(\mathbb{R}^d)}}, G_{\mu_0,2}\right) \le 4\,\mathcal{D}_{\mathrm{KL}}\left(\frac{F_n}{\|F_n\|_{L^1(\mathbb{R}^d)}}\middle\| G_{\mu_0,2}\right).$$

However, Theorem 1.1 does not hold when $\alpha = 0$, as mentioned in Remark 1.5, due to two fundamental reasons. First, when $\alpha = 0$ there is translation invariance and therefore, for generic $F_0 \in \mathcal{M}_+(\mathbb{R}^d)$ one cannot expect that the normalized profiles $F_n/\|F_n\|_{L^1(\mathbb{R}^d)}$ always converge to the Gaussian $\boldsymbol{F}_{\alpha=0}$ centered at the origin (contrarily to what happens when $\alpha > 0$). Otherwise, the centers of mass must get shifted toward the origin, thus breaking the translation invariance. Indeed, as mentioned in Corollary 2.2, the equilibria are not unique when $\alpha = 0$ (contrarily to the case $\alpha > 0$), and the center of mass of the resulting equilibria must stay equal to the initial one. Second, even if we set the initial center of mass at the origin so that we kill the translation invariance, our method of proof of Theorem 1.1 leads to estimates that blow up as $\alpha \to 0$ since we have $2\,\boldsymbol{k}_{\alpha=0} = 1$.

2.2. **Incompatibility of Wasserstein metric with multiplicative operators.** Note that by definition (1.2), our operator $\mathcal{T}$ is the composition of the sexual resproduction operator $\mathcal{B}$ with the multiplicative operator by the survival probability e^{-m}. Then, it might be natural to study perturbations of the previous Lemma 2.1 including the following conservative multiplicative operator.

Definition 2.3 (Normalization of multiplicative operator)**.**

$$\mathcal{M}[F] := \frac{e^{-m}F}{\|e^{-m}F\|_{\mathcal{M}(\mathbb{R}^d)}}, \quad F \in \mathcal{M}_+(\mathbb{R}^d) \setminus \{0\}.$$

However, the latter is not Lipschitz continuous with respect to the quadratic Wasserstein metric. Hence the composition of $\mathcal{B}$ and $\mathcal{M}$ is not expected to be contractive, even in the case of weak selection, without any additional restriction. We illustrate such a Lipschitz discontinuity of $\mathcal{M}$ in the following example.

Example 2.4. *Suppose that the $m \in C^1_+(\mathbb{R})$ is radially symmetric around the origin and set $F_1, F_2 \in \mathcal{P}_2(\mathbb{R})$ as the sum of two Dirac masses, F_1 being symmetric, and F_2 nearly symmetric. More precisely,*

$$F_1 = \frac{1}{2}\delta_{-h} + \frac{1}{2}\delta_h, \quad F_2 = \frac{1}{2}\delta_{-h+\varepsilon} + \frac{1}{2}\delta_{h+\varepsilon},$$

where $h > 0$ is fixed so that $m'(h) \neq 0$ and $\varepsilon > 0$ is small. Note that

$$\mathcal{M}[F_1] = F_1, \quad \mathcal{M}[F_2] = (1 - p_\varepsilon)\delta_{-h+\varepsilon} + p_\varepsilon \delta_{h+\varepsilon},$$

where $p_\varepsilon \in (0,1)$ is given by

$$p_\varepsilon = \frac{e^{-m(h+\varepsilon)}}{e^{-m(-h+\varepsilon)} + e^{-m(h+\varepsilon)}}.$$

On the one hand, we have $\mathcal{W}_2(F_1, F_2) = \varepsilon$ because F_2 is simply deduced from F_1 by a translation of size ε. On the other hand, assuming $m'(h) > 0$ for simplicity (a similar argument holds if $m'(h) < 0$) and taking $\varepsilon > 0$ sufficiently small we obtain $p_\varepsilon < \frac{1}{2}$ and

$$\mathcal{W}_2(\mathcal{M}[F_1], \mathcal{M}[F_2]) = \sqrt{\varepsilon^2 + 4h(h-\varepsilon)\left(\frac{1}{2} - p_\varepsilon\right)} \sim \varepsilon + h\, m'(h)^{1/2}\varepsilon^{1/2},$$

as $\varepsilon \to 0$ by symmetry and the mean value theorem. In a sense, the leading order term corresponds to the cost of moving a piece of mass $\frac{1}{2} - p_\varepsilon$ from $-h + \varepsilon$ to h in order to equilibrate the Dirac masses in the transport plan. Since $m'(h) \neq 0$ this leads to the Lipschitz-discontinuity of the operator $\mathcal{M}$.

In [21], this case is ruled out by assuming that the densities are uniformly bounded below on compact intervals. In our case, we will avoid relying on the above quantification in Lemma 2.1 for the contractivity of $\mathcal{B}$ and will propose a different strategy that we discuss in the sequel.

2.3. **Log-Lipschitz contraction estimate.** In this paper, we explore an alternative approach to derive some suitable contraction of the operator $\mathcal{T}$ in appropriate norms in the regime of strong, but quadratic, selection. As anticipated in Remark 1.4, our computations suggest using a log-Lipschitz norm, which measures the uniform deviation of tails of any profile F relative to the Gaussian tails of $\boldsymbol{F}_\alpha$. More specifically, we obtain the following result for any $\alpha \geq 0$ (but only useful when $\alpha > 0$).

Lemma 2.5 (Log-Lipschitz estimate)**.** *Let m be the quadratic selection function in* (1.5) *and consider any value $\alpha \geq 0$. Then, the following estimate holds true*

$$\left\| \nabla \log \frac{\mathcal{T}[F]}{\boldsymbol{F}_\alpha} \right\|_{L^\infty(\mathbb{R}^d)} \leq 2\,\boldsymbol{k}_\alpha \left\| \nabla \log \frac{F}{\boldsymbol{F}_\alpha} \right\|_{L^\infty(\mathbb{R}^d)},$$

for any $F \in L^1_+(\mathbb{R}^d) \cap C^1(\mathbb{R}^d)$ such that $\|\nabla \log \frac{F}{\boldsymbol{F}_\alpha}\|_{L^\infty(\mathbb{R}^d)} < \infty$. The coefficient $\boldsymbol{k}_\alpha$ above is given by the numerical value (1.15) *in Theorem 1.1.*

When applied to $F = F_n$ for any solution $\{F_n\}_{n\in\mathbb{N}}$ of the time-evolution problem (1.1), the above inequality in Lemma 2.5 allows propagating a control of growth/decrease of the log-Lipschitz norm under the flow of the equation. In particular, when $\alpha > 0$, we have $2\,\boldsymbol{k}_\alpha < 1$ (*cf.* Figure 1), and therefore we obtain an actual contractivity estimate. There is one main drawback though: typically $\nabla \log \frac{F_n}{\boldsymbol{F}_\alpha}$ are genuinely unbounded unless F_0 and $\boldsymbol{F}_\alpha$ have the same Gaussian decay for large x (which is much too restrictive). A major point of our work will precisely be to circumvent this unbounded factors.

However, when $\alpha = 0$ we have $2\,\boldsymbol{k}_\alpha = 1$, and we simply obtain non-expansiveness. We emphasize that in the derivation of such a rough estimate, we do not use carefully the preservation of the center of mass. This results in a non-expansiveness estimate, which is conceptually no better than the previous non-expansiveness property (2.3) in the quadratic Wasserstein metric, obtained as in Lemma 2.1 without assuming that the centers of mass are the same. For this reason, our method of proof will finally not yield satisfying results in the case $\alpha = 0$, but we find this idea illuminating to address the case of a non-trivial effect of selection when $\alpha > 0$. A refinement of Lemma 2.5 leading to real contractivity would require tackling more carefully the center of mass. However, as mentioned in Remark 1.5, we will not address this in the current paper, and we refer to Section 8 for some perspectives and future works in this line.

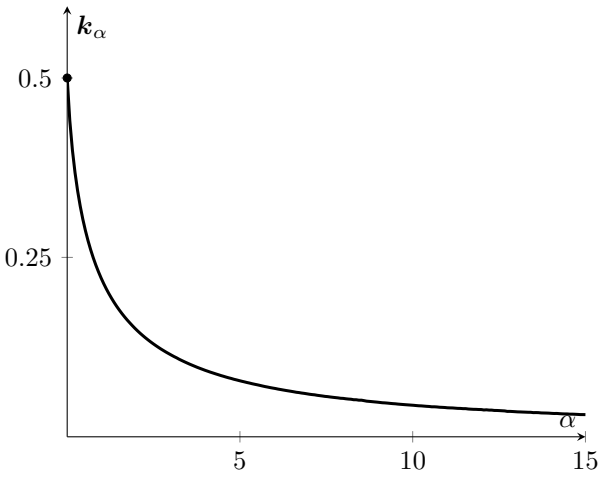


FIGURE 1. Parameter $\boldsymbol{k}_\alpha$ against parameter α.

Proof of Lemma 2.5. Our starting point resides in the following normalization of the operator $\mathcal{T}$:

$$\frac{\mathcal{T}[F](x)}{\boldsymbol{F}_\alpha(x)} = \frac{\boldsymbol{\lambda}_\alpha}{\int_{\mathbb{R}^d} \frac{F(x')}{\boldsymbol{F}_\alpha(x')}\,\boldsymbol{F}_\alpha(x')\,dx'} \iint_{\mathbb{R}^{2d}} \mathbf{P}(x;x_1,x_2)\frac{F(x_1)}{\boldsymbol{F}_\alpha(x_1)}\frac{F(x_2)}{\boldsymbol{F}_\alpha(x_2)}\,dx_1\,dx_2, \tag{2.4}$$

for all $x \in \mathbb{R}^d$, where we have

$$\mathbf{P}(x;x_1,x_2) = \frac{e^{\frac{1}{2}(\frac{1}{\boldsymbol{\sigma}_\alpha^2}-\alpha)|x|^2}}{\boldsymbol{\lambda}_\alpha(2\pi\,\boldsymbol{\sigma}_\alpha)^d} \exp\left[-\frac{1}{2}\left|x-\frac{x_1+x_2}{2}\right|^2 - \frac{1}{2\,\boldsymbol{\sigma}_\alpha^2}(|x_1|^2+|x_2|^2)\right]. \tag{2.5}$$

Above we have exploited the explicit Gaussian shape of $\boldsymbol{F}_\alpha$ to find $\mathbf{P}$ explicitly. Note that $\mathbf{P}$ consists in a one-step Markov transition kernel representing the probability that parental traits (x_1,x_2) lead to a descendant trait x. Indeed, $\iint_{\mathbb{R}^{2d}} \mathbf{P}(x;x_1,x_2)\,dx_1\,dx_2 = 1$ for all $x \in \mathbb{R}^d$ because $(\boldsymbol{\lambda}_\alpha, \boldsymbol{F}_\alpha)$ solves the non-linear eigenproblem (1.11). We remark that the quadratic form in the exponential in (2.5) reaches its maximum value at $(x_1,x_2) = (kx,kx)$ for $k = \boldsymbol{\sigma}_\alpha^2/(2+\boldsymbol{\sigma}_\alpha^2)$ and by definition (1.15) of $\boldsymbol{k}_\alpha$, we infer $k = \boldsymbol{k}_\alpha$. This motivates using the change of variables

$$x_1 = \boldsymbol{k}_\alpha\,x + y_1, \quad x_2 = \boldsymbol{k}_\alpha\,x + y_2, \tag{2.6}$$

which appropriately centers the quadratic form at the minimum. Specifically, we obtain

$$\begin{aligned}
&\frac{1}{2}\left|x-\frac{x_1+x_2}{2}\right|^2 - \frac{1}{2\,\boldsymbol{\sigma}_\alpha^2}(|x_1|^2+|x_2|^2)\\
&= -\frac{1}{2}\left|\frac{y_1+y_2}{2}\right|^2 - \frac{1}{2\,\boldsymbol{\sigma}_\alpha^2}(|y_1|^2+|y_2|^2) - \frac{1}{2}\left((1-\boldsymbol{k}_\alpha)^2 + \frac{2\,\boldsymbol{k}_\alpha^2}{\boldsymbol{\sigma}_\alpha^2}\right)|x|^2 + \frac{1}{2}\left((1-\boldsymbol{k}_\alpha) - \frac{2\,\boldsymbol{k}_\alpha}{\boldsymbol{\sigma}_\alpha^2}\right)x\cdot(y_1+y_2),\\
&= -\frac{1}{2}\left|\frac{y_1+y_2}{2}\right|^2 - \frac{1}{2\,\boldsymbol{\sigma}_\alpha^2}(|y_1|^2+|y_2|^2) - \frac{1}{2}\left(\frac{1}{\boldsymbol{\sigma}_\alpha^2}-\alpha\right)|x|^2,
\end{aligned}$$

where in the second line we have noticed that the coefficient of the crossed term $x\cdot(y_1+y_2)$ vanishes due to the relation $\boldsymbol{k}_\alpha = \frac{\boldsymbol{\sigma}_\alpha^2}{2+\boldsymbol{\sigma}_\alpha^2}$, and the coefficient of the $|x|^2$ factor can be reformulated as

$$(1-\boldsymbol{k}_\alpha)^2 + \frac{2\,\boldsymbol{k}_\alpha^2}{\boldsymbol{\sigma}_\alpha^2} = \frac{2}{2+\boldsymbol{\sigma}_\alpha^2} = \frac{1}{1+\frac{\boldsymbol{\sigma}_\alpha^2}{2}} = \frac{1}{\boldsymbol{\sigma}_\alpha^2} - \alpha,$$

thanks to the implicit equation (1.13) satisfied by $\boldsymbol{\sigma}_\alpha^2$. Therefore, (2.4)-(2.5) transform into

$$\frac{\mathcal{T}[F](x)}{\boldsymbol{F}_\alpha(x)} = \frac{\boldsymbol{\lambda}_\alpha}{\int_{\mathbb{R}^d} \frac{F(x')}{\boldsymbol{F}_\alpha(x')}\,\boldsymbol{F}_\alpha(x')\,dx'} \iint_{\mathbb{R}^{2d}} \widetilde{\mathbf{P}}(y_1,y_2)\frac{F(\boldsymbol{k}_\alpha\,x+y_1)}{\boldsymbol{F}_\alpha(\boldsymbol{k}_\alpha\,x+y_1)}\frac{F(\boldsymbol{k}_\alpha\,x+y_2)}{\boldsymbol{F}_\alpha(\boldsymbol{k}_\alpha\,x+y_2)}\,dy_1\,dy_2, \tag{2.7}$$

$$\widetilde{\mathbf{P}}(y_1,y_2) = \frac{1}{\boldsymbol{\lambda}_\alpha(2\pi\,\boldsymbol{\sigma}_\alpha)^d} \exp\left[-\frac{1}{2}\left|\frac{y_1+y_2}{2}\right|^2 - \frac{1}{2\,\boldsymbol{\sigma}_\alpha^2}(|y_1|^2+|y_2|^2)\right]. \tag{2.8}$$

We remark that the new Markov transition kernel $\widetilde{\mathbf{P}} = \widetilde{\mathbf{P}}(y_1,y_2)$ does not depend on x thanks to the explicit cancellation of the $|x|^2$ dependent factors. At this level, we start to observe the ergodicity

phenomenon since the dependence of x on the right hand side of (2.7) has shrunk by a factor $\boldsymbol{k}_\alpha$. A possible strategy to see if there is a quantitative degradation of the dependence on x is to take logarithmic derivatives and try to relate the log-Lipschitz norms of $\mathcal{T}[F]$ and F. Specifically, we have

$$\nabla \log \frac{\mathcal{T}[F](x)}{\boldsymbol{F}_\alpha(x)} = \boldsymbol{k}_\alpha \int_{\mathbb{R}^{2d}} \left(\nabla \log \frac{F(\boldsymbol{k}_\alpha\, x + y_1)}{\boldsymbol{F}_\alpha(\boldsymbol{k}_\alpha\, x + y_1)} + \nabla \log \frac{F(\boldsymbol{k}_\alpha\, x + y_2)}{\boldsymbol{F}_\alpha(\boldsymbol{k}_\alpha\, x + y_2)} \right) \nu(x; dy_1, dy_2), \tag{2.9}$$

where the x-dependent measures $\nu(x; dy_1, dy_2)$ on the variables (y_1, y_2) have the following density with respect to the Lebesgue measure:

$$\frac{\nu(x; dy_1, dy_2)}{dy_1\, dy_2} = \frac{\widetilde{\mathbf{P}}(y_1, y_2) \dfrac{F(\boldsymbol{k}_\alpha\, x + y_1)}{\boldsymbol{F}_\alpha(\boldsymbol{k}_\alpha\, x + y_1)} \dfrac{F(\boldsymbol{k}_\alpha\, x + y_2)}{\boldsymbol{F}_\alpha(\boldsymbol{k}_\alpha\, x + y_2)}}{\displaystyle\iint_{\mathbb{R}^{2d}} \widetilde{\mathbf{P}}(y_1', y_2') \frac{F(\boldsymbol{k}_\alpha\, x + y_1')}{\boldsymbol{F}_\alpha(\boldsymbol{k}_\alpha\, x + y_1')} \frac{F(\boldsymbol{k}_\alpha\, x + y_2')}{\boldsymbol{F}_\alpha(\boldsymbol{k}_\alpha\, x + y_2')}\, dy_1'\, dy_2'}.$$

Since the integrands of (2.9) are uniformly bounded by our assumptions, we end the proof by taking L^∞ bounds and using that ν are probability measures on the variables (y_1, y_2). □

2.4. **Brief description of our strategy.** As advanced before, our strategy is based on a finer understanding of the iterations of $\mathcal{T}$ across generations. Specifically, it relies on a suitable reformulation of the solutions $\{F_n\}_{n\in\mathbb{N}}$ solving the recursion (1.1) for the special quadratic selection function m in (1.5). At this stage we intentionally keep notation simple and intuitive, since our goal is to briefly present the main strategy. However, a rigorous approach with more descriptive notation for the trees of ancestors, which arise from the recursion, is developed in detail in Section 5.

The first step consists in choosing the appropriate normalization extending the previous normalization $F_n/\boldsymbol{F}_\alpha$ in the proof of Lemma 2.5. In this paper, we have opted for

$$\bar{F}_n = \frac{e^m F_n}{\boldsymbol{F}_{\alpha=0}},$$

(*cf.* Definition 5.1) but there is some freedom here. In particular, the term e^m is not mandatory, but it is convenient as such for an easier sorting of the various terms. Indeed, we may typically consider any normalization $F_n/G_{0,\sigma^2}$ with σ^2 larger but arbitrarily close to $\boldsymbol{\sigma}_\alpha^2$ (being $\boldsymbol{\sigma}_\alpha^2$ the variance (1.13) of the expected equilibrium $\boldsymbol{F}_\alpha$). In other words, we could take any "close-to-optimal" normalization as compared to the "optimal" normalization used in the strategy in Section 2.3. However, for general selection functions m we do not know the expected equilibrium. To account for a more robust viewpoint, we follow a more robust approach so that we do not "optimize" the normalization.

In a nutshell: we iterate the operator $\mathcal{T}$ in the recursion for F_n up to the initial datum and we obtain (modulo a multiplicative constant) a formula of the following type

$$F_n(x) \propto \int_{\mathbb{R}^{(2^{n+1}-2)d}} \mathbf{P}_n(x; \{x_i\}) \prod_j \bar{F}_0(x_j)\, d\{x_i\},$$

for an explicit n-step transition kernel $\mathbf{P}_n$ with Gaussian shape:

$$\mathbf{P}_n(x; \{x_i\}) := \exp(-\mathbf{Q}_n(x; \{x_i\})).$$

Here, the index i spans the $2^{n+1} - 2$ members of the genealogical tree with n generations of ancestors from the root (excluded) to the leaves, whilst j only spans the 2^n members at the leaves. In addition, $\mathbf{Q}_n = \mathbf{Q}_n(x; \{x_i\})$ is a quadratic form taking as arguments all traits x and $\{x_i\}$ (*i.e.*, all the $2^{n+1} - 1$ ancestors in the family chart including the root x). As compared to Section 2.3, we have iterated n times, which raises the dimension of the quadratic form to $2^{n+1} - 1$. In addition, the change of normalization leads to additional quadratic contributions $\frac{\gamma}{2}|x_i|^2$ (with $\gamma = \frac{\boldsymbol{\sigma}_\alpha^2}{2+\boldsymbol{\sigma}_\alpha^2}$)). Therefore, the minimizers of $\mathbf{Q}_n$ get shifted (*cf.* Section 5). As a result of successive backwards changes of variables from leaves to the root very much in the spirit of (2.6), the following alternative expression is obtained:

$$F_n(x) \propto \widetilde{\boldsymbol{F}}_n(x) \int_{\mathbb{R}^{(2^{n+1}-2)d}} \widetilde{\boldsymbol{P}}_n(\{y_i\}) \prod_j \bar{F}_0(\Phi_n^j(x; \{y_i\}))\, d\{y_i\}\,. \tag{2.10}$$

Above, for every leaf j the maps Φ_n^j defined by

$$\Phi_n^j(x; \{y_i\}) := \kappa_n x + \Lambda^j(\{y_i\}),$$

where Λ_n^j are affine transformations with respect to the variables $\{y_i\}$. We shall call them the lineage maps (*cf.* Definition 5.6), since they contain precise information of the effect of the leaf j of the genealogical

tree on the resulting trait x. In addition, $\boldsymbol{P}_n$ are n-step transition kernel with Gaussian shape:

$$\widetilde{\boldsymbol{P}}_n(\{y_i\}) = \exp(-\widetilde{\mathbf{Q}}_n(\{y_i\})),$$

for a x-independent quadratic form $\widetilde{\mathbf{Q}}_n$ taking as arguments all traits $\{y_i\}$ (root excluded). In this case, the $|x|^2$ remainders coming from the changes of variables on $\mathbf{P}_n$ do not simplify due to our different normalization, and they contribute with an explicit Gaussian factor $\widetilde{\boldsymbol{F}}_n(x)$. Then, the dependency upon x is split into two parts: the explicit term $\widetilde{\boldsymbol{F}}_n(x)$ outside the integral (which is proven to converge towards $\boldsymbol{F}_\alpha$ by construction) and the contribution $\kappa_n x$ of the lineage maps $\Phi_n^j(x;\{y_i\})$ inside $\bar{F}_0$ for an appropriate sequence $\{\kappa_n\}_{n\in\mathbb{N}} \subseteq \mathbb{R}_+$. The major observation here is that $\kappa_n \ll 2^{-n}$ when $\alpha > 0$ because of the strong shift towards the origin under selection. Otherwise, for $\alpha = 0$ we only have $\kappa_n = 2^{-n}$.

Formula (2.10) for $\alpha > 0$ then suggests a strong form of ergodicity, reminiscent of the contraction of the log-Lipschitz norm in Lemma 2.5, where the dependency on x fades as $n \to \infty$. Indeed, there are still 2^n terms in the product indexed by j, but the contraction parameter κ_n appears to decay fast enough to compensate them. To make this argument quantitative, let us differentiate (2.10) again to obtain

$$\nabla \log F_n(x) = \nabla \log \widetilde{\boldsymbol{F}}_n(x) + \kappa_n \sum_j \int_{\mathbb{R}^{(2^{n+1}-2)d}} \nabla \log \bar{F}_0(\Phi_n^j(x;\{y_i\}))\, \nu_n(x; d\{y_i\})\,, \tag{2.11}$$

where the x-dependent probability measures $\nu_n(x; d\{y_i\})$ on the variables $\{y_i\}$ have the following density with respect to the Lebesgue measure:

$$\frac{\nu_n(x; d\{y_i\})}{d\{y_i\}} = \frac{\widetilde{\boldsymbol{P}}_n(\{y_i\}) \prod_j \bar{F}_0(\Phi_n^j(x;\{y_i\}))}{\int_{\mathbb{R}^{(2^{n+1}-2)d}} \widetilde{\boldsymbol{P}}_n(\{y_i'\}) \prod_j \bar{F}_0(\Phi_n^j(x;\{y_i'\}))\, d\{y_i'\}}.$$

We remark that a naive repetition of the strategy in Lemma 2.5 under the additional assumption that $\nabla \log \bar{F}_0 \in L^\infty(\mathbb{R}^d)$ would immediately retrieve an exponential decay of the second factor in (2.11) when $\alpha > 0$ (or only a uniform bound if $\alpha = 0$) since the large sum over j would be bounded by $2^n \kappa_n \|\nabla \log \bar{F}_0\|_{L^\infty(\mathbb{R}^d)}$. However, as mentioned before, there is a strong drawback: the factors $\nabla \log \bar{F}_0$ are not bounded, not even after long enough times. In the current formal argument, we have not discussed about the precise shape of the linear components Λ_n^j of the lineage maps Φ_n^j, since it was not relevant thus far. However, we anticipate that to overcome the aforementioned complication, we shall require more precise estimates for large values of $\{y_i\}$ in the high-dimensional integral, which are guaranteed by the strong enough decay of κ_n, and which will be essential in the rigorous proof in Section 6.

More specifically, and this explains why our approach moves from one-step contraction estimates (in the spirit of Section 2.3) to ergodicity results, note that the same reformulation of the iteration as above could be done exactly $n - k$ times up to an advanced enough time step k. By doing so, we can grasp on some natural Gaussian shaping under selection, which we expect to lead to unbounded $\nabla \log \bar{F}_k$, but growing sublinearly at infinity. Hence, we require precise compensations in the high-dimensional integral (2.11) which we find by deriving a suitable control of moments of ν_n. By doing so, one has to irremediably move from uniform estimates to averaged estimates, and being able to propagate them for large times. This requires a thorough and highly technical analysis, which becomes the main objective of this paper.

For an easier readability, and to guide the reader along the various steps of the proof, we provide an overall map of it in Figure 2, which allows interconnecting the main fundamental results.

3. Preliminaries

In this part, we collect some necessary preliminary tools and results that will be used later on.

3.1. **Perfect binary trees.** In this paper, we shall systematically use indices i that do not range on discrete intervals $\{1,\dots,n\}$ but rather on the vertices of a specific type of trees, which are called *perfect binary trees*, see [22, 25]. First, we introduce the notion of *binary trees*, which we present using their *universal address system*, see [22, Section 11.3.3]. Namely, a binary tree T consists in a finite subset of words $\mathsf{T} \subseteq \mathsf{W}_2 := \cup_{k=0}^\infty \{1,2\}^k$ with letters in the alphabet $\{1,2\}$ that verifies:

(i) $\emptyset \in \mathsf{T}$.

(ii) If $i1 \in \mathsf{T}$ or $i2 \in \mathsf{T}$ for some word $i \in \mathsf{W}_2$, then $i \in \mathsf{T}$.

This implies that the root is the empty word , and T is stable under chopping letters on the right of its words. Given a word $i \in \mathsf{T}$, we denote its length (number of letters) or *height* in the tree by $|i|$. In particular, $|\emptyset| = 0$. A binary tree T is said to be *perfect* if, in addition, the following properties hold:

(iii) Given a word $i \in \mathsf{W}_2$, then $i1 \in \mathsf{T}$ if, and only if, $i2 \in \mathsf{T}$.

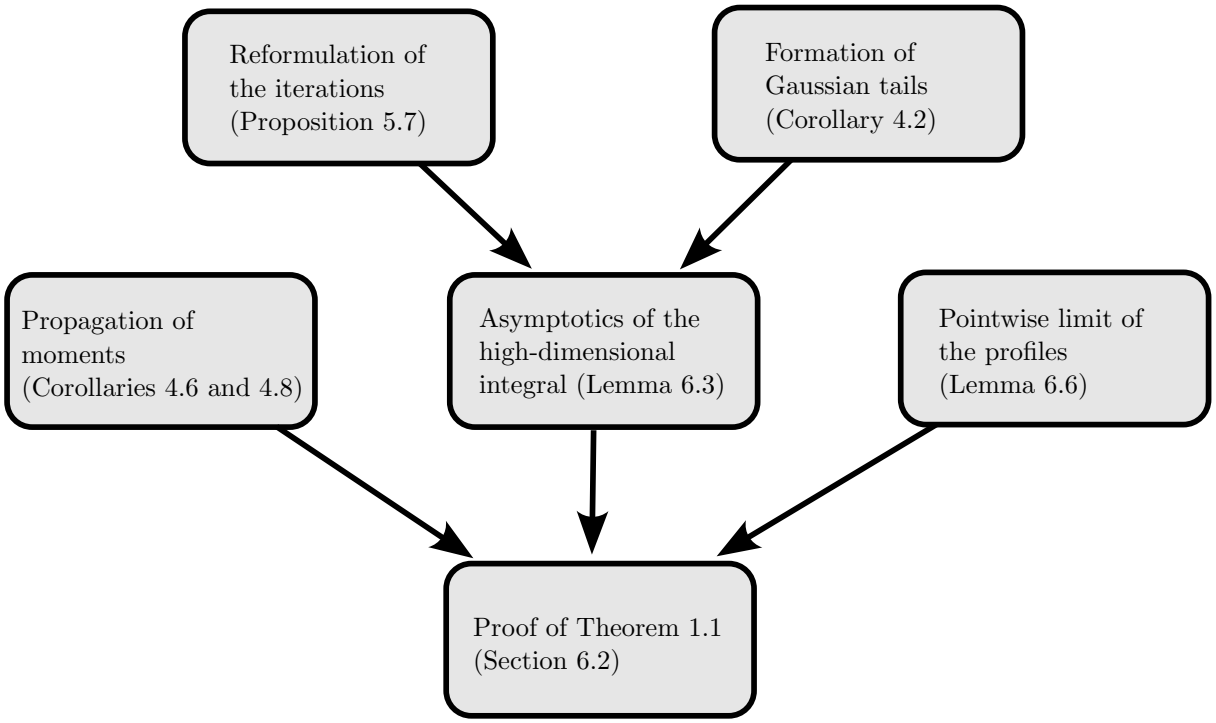


FIGURE 2. Overall map of the proof of Theorem 1.1

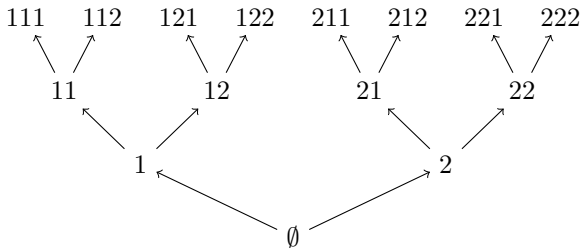


FIGURE 3. Perfect binary tree T^3

(iv) There exists $n \in \mathbb{N}$ such that $|i| \leq n$ for every $i \in \mathsf{T}$ and

$$\#\{i \in \mathsf{T} : |i| = n\} = 2^n.$$

This implies that, except for the root, words in T appear in couples $\{i1, i2\}$ and paths in the tree arising from the root achieve the same maximal height n. The above four conditions determine a unique tree, which we call the *perfect binary tree with height* $n \in \mathbb{N}$ and we denote by T^n. See Figure 3 for a graphical representation of the perfect binary tree T^3 of height 3. Binary trees are often used to describe the different generations of offsprings after a given individual. In our case, we shall make a reverse use of trees. Namely, a binary tree will represent the *family tree* or *pedigree chart* of an individual, consisting of the different generations of *ancestors* of such an individual. For this purpose, we shall establish the following terminology:

(i) (**Leaves**) We say that a vertex $i \in \mathsf{T}^n$ is a leaf of the tree if $|i| = n$. The set of all leaves of T^n will be denoted by

$$\mathsf{L}^n := \{i \in \mathsf{T}^n : |i| = n\}.$$

(ii) (**Levels**) We say that a vertex $i \in \mathsf{T}^n$ is on the m-th level of the tree if $|i| = m$ for some $1 \leq m \leq n$. The set of all vertices on the m-th level will be denoted

$$\mathsf{L}^n_m := \{i \in \mathsf{T}^n : |i| = m\}.$$

In particular, note that $\mathsf{L}^n_0 = \{\emptyset\}$ is the root and $\mathsf{L}^n_n = \mathsf{L}^n$ are the leaves. For simplicity of notation, we denote the set of vertices of the root-free and leaves-free tree respectively by

$$\mathsf{T}^n_* := \mathsf{T}^n \setminus \{\emptyset\}, \quad \widehat{\mathsf{T}}^n := \mathsf{T}^n \setminus \mathsf{L}^n.$$

(iii) (**Child**) Given a vertex $i \in \mathsf{T}_*^n$, then there exists a unique word $j \in \mathsf{W}_2$ such that $i = j1$ or $i = j2$. Such a vertex j is called the *child* of i and will be denoted by $\mathsf{c}(i) = j$.

(iv) (**Parents**) Given a vertex $\in \widehat{\mathsf{T}}^n$, we define the *parents* of i as the subset of all vertices that have the same common child i, that is,

$$\mathsf{P}(i) := \{i1, i2\}.$$

(v) (**Mate**) Given a vertex $i \in \mathsf{T}_*^n$, we define the *mate* of i and we denote it $\mathsf{m}(i)$ as the only other vertex in T^n that has the same child as i, *i.e.*,

$$\{i, \mathsf{m}(i)\} = \{\mathsf{c}(i)1, \mathsf{c}(i)2\}.$$

(vi) (**Tree order**) Given two vertices $i, j \in \mathsf{T}^n$, we say that $i \leq j$ if the associated words are so ordered according to the lexicographical order of the set of words W_2 with two letters $\{1, 2\}$.

(vii) (**Highest common descendant**) Given two vertices $i, j \in \widehat{\mathsf{T}}^n$, we define the *highest common descendant* of i and j, and we denote it by $i \wedge j$, as

$$i \wedge j := \max\{l \in \mathsf{T}^n : l \leq i, l \leq j\},$$

where the maximum is considered with respect to above tree (lexicographic) order.

Remark 3.1 (Tree-indexed variables)**.** *Given $n \in \mathbb{N}$, we shall identify $\mathbb{R}^{2(2^n-1)d} \equiv (\mathbb{R}^d)^{\mathsf{T}_*^n}$ and $\mathbb{R}^{2^n d} \equiv (\mathbb{R}^d)^{\mathsf{L}^n}$. Specifically, vectors $\mathbf{x}_n, \mathbf{y}_n \in \mathbb{R}^{2(2^n-1)d}$ and $\mathbf{z}_n \in \mathbb{R}^{2^n d}$ will be regarded as indexed families*

$$\mathbf{x}_n = (x_i)_{i \in \mathsf{T}_*^n}, \quad \mathbf{y}_n = (y_i)_{i \in \mathsf{T}_*^n}, \quad \mathbf{z}_n = (z_j)_{j \in \mathsf{L}^n}, \tag{3.1}$$

where $x_i, y_i \in \mathbb{R}^d$ and $z_j \in \mathbb{R}^d$ for each $i \in \mathsf{T}_*^n$ and $j \in \mathsf{L}^n$.

3.2. **Gaussian solutions.** In this part, we compute particular solutions of the time-discrete problem (1.1) and the associated eigenproblem (1.11). As we advanced before, we shall exploit the explicit algebraic structure imposed by G (given by the Gaussian mixing kernel (1.4)) and m (given by the quadratic selection function (1.5)). Namely, explicit Gaussian solution will be obtained. We recall first the following stability property of Gaussians under convolutions.

Lemma 3.2 (Stability of Gaussians)**.** *The following relation holds true*

$$G_{\mu_1,\sigma_1^2} * G_{\mu_2,\sigma_2^2} = G_{\mu_1+\mu_2,\sigma_1^2+\sigma_2^2}, \tag{3.2}$$

for any couple of means $\mu_1, \mu_2 \in \mathbb{R}^d$ and variances $\sigma_1^2, \sigma_2^2 > 0$.

Using the above result, we can obtain the following explicit evaluation of the operator $\mathcal{T}$ in (1.2) over the class of Gaussian functions.

Lemma 3.3 (Evaluation on Gaussians)**.** *Consider any $\mu \in \mathbb{R}^d$ and $\sigma^2 \in \mathbb{R}_+^*$, then*

$$\mathcal{T}[G_{\mu,\sigma^2}] = m_* \, G_{\mu_*,\sigma_*^2},$$

where the parameters m_*, μ_* and σ_*^2 are given by:

$$m_* = \frac{e^{-\frac{1}{2}\frac{\alpha|\mu|^2}{1+\alpha(1+\frac{\sigma^2}{2})}}}{(1+\alpha(1+\frac{\sigma^2}{2}))^{d/2}}, \quad \mu_* = \frac{\mu}{1+\alpha(1+\frac{\sigma^2}{2})}, \quad \sigma_*^2 = \frac{1+\frac{\sigma^2}{2}}{1+\alpha(1+\frac{\sigma^2}{2})}. \tag{3.3}$$

Proof. On the one hand, note that by Lemma 3.2

$$\mathcal{B}[G_{\mu,\sigma^2}](x) = \left(G\left(\frac{\cdot}{2}\right) * G_{\mu,\sigma^2} * G_{\mu,\sigma^2}\right)(2x) = G_{\mu,1+\frac{\sigma^2}{2}}(x),$$

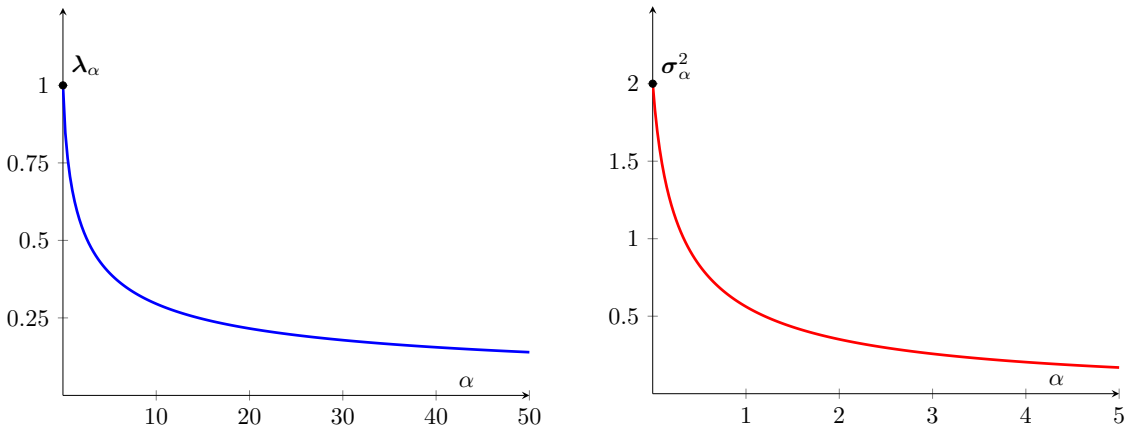
for each $x \in \mathbb{R}^d$. Therefore, by definition of $\mathcal{T}$ in (1.2) we obtain

$$\mathcal{T}[G_{\mu,\sigma^2}](x) = e^{-\frac{\alpha}{2}|x|^2} G_{\mu,1+\frac{\sigma^2}{2}}(x) = \frac{1}{(2\pi(1+\frac{\sigma^2}{2}))^{d/2}} \exp\left(-\frac{\alpha}{2}|x|^2 - \frac{1}{2}\frac{1}{1+\frac{\sigma^2}{2}}|x-\mu|^2\right),$$

for each $x \in \mathbb{R}^d$. By completing the square inside the exponential, we conclude our result. □

Consequently, the following explicit Gaussian solution of the eigenproblem (1.11) is found.

Proposition 3.4 (Gaussian solution of the eigenproblem)**.** *Assume that $\alpha \in \mathbb{R}_+^*$, then the eigenproblem* (1.11) *has a unique Gaussian solution $(\boldsymbol{\lambda}_\alpha, \boldsymbol{F}_\alpha)$, determined by the relation* (1.12) *in Theorem 1.1.*



(a) The malthusian exponent $\boldsymbol{\lambda}_\alpha$ against parameter α for $d = 1$.

(b) The equilibrium variance $\boldsymbol{\sigma}^2_\alpha$ against parameter α.

FIGURE 4. Plot of eigenvalue $\boldsymbol{\lambda}_\alpha$ and variance at equilibrium $\boldsymbol{\sigma}^2_\alpha$ against α for $d = 1$. The case $\alpha = 0$ corresponds to absence of selection, which is conservative ($\boldsymbol{\lambda}_{\alpha=0} = 1$), and the variance at equilibrium $\boldsymbol{\sigma}^2_{\alpha=0}$ is twice the variance of the mixing kernel in $\mathcal{B}$.

Proof. We look for $\lambda \in \mathbb{R}^*_+$, $\mu \in \mathbb{R}^d$ and $\sigma^2 \in \mathbb{R}^*_+$ such that $\lambda G_{\mu,\sigma^2} = \mathcal{T}[G_{\mu,\sigma^2}]$. By Lemma 3.3 and bearing in mind parameters m_*, μ_* and σ^2_* in (3.3) we obtain that (λ, μ, σ^2) must solve the equations:

$$\lambda = \frac{e^{-\frac{1}{2}\frac{\alpha|\mu|^2}{1+\alpha(1+\frac{\sigma^2}{2})}}}{(1+\alpha(1+\frac{\sigma^2}{2}))^{d/2}}, \quad \mu = \frac{\mu}{1+\alpha(1+\frac{\sigma^2}{2})}, \quad \sigma^2 = \frac{1+\frac{\sigma^2}{2}}{1+\alpha(1+\frac{\sigma^2}{2})}.$$

Hence, the only solution is given by $\mu = 0$ and λ and σ^2 are determined by (1.12) and (1.13). □

Remark 3.5 (Dependency on selection)**.** *The eigenvalue $\boldsymbol{\lambda}_\alpha \in \mathbb{R}^*_+$ and the variance $\boldsymbol{\sigma}^2_\alpha \in \mathbb{R}^*_+$ determined by the relations* (1.12) *and* (1.13) *for the the unique Gaussian solution $(\boldsymbol{\lambda}_\alpha, \boldsymbol{F}_\alpha)$ to the eigenproblem* (1.11) *are monotonically decreasing with the selection coefficient α. Namely, the larger α, the smaller $\boldsymbol{\sigma}^2_\alpha$ and $\boldsymbol{\lambda}_\alpha$. Indeed, we obtain (see also Figure 4)*

$$\begin{aligned} \alpha \nearrow \infty \quad &\Longrightarrow \quad \boldsymbol{\sigma}^2_\alpha \searrow 0, \quad \boldsymbol{\lambda}_\alpha \searrow 0 \\ \alpha \searrow 0 \quad &\Longrightarrow \quad \boldsymbol{\sigma}^2_\alpha \nearrow 2, \quad \boldsymbol{\lambda}_\alpha \nearrow 1. \end{aligned}$$

Therefore, if selection is strong, then $\boldsymbol{F}_\alpha$ is very concentrated around the origin and, if selection ceases, then $\boldsymbol{F}_\alpha$ is twice as spread as the mixing kernel G in (1.4)*, a famous result in quantitative genetics, see* e.g. [9]*. Note that $\boldsymbol{\lambda}_\alpha < 1$ for any $\alpha \in \mathbb{R}^*_+$ and, consequently, the special steady solutions $F_n = \boldsymbol{\lambda}^n_\alpha \boldsymbol{F}_\alpha$ coming from ansatz* (1.10) *always get extinct for large n. This is an artificial consequence of our parameter reduction in Appendix A. In particular, if we maintain parameter β in* (A.4) *coming from the ratio between birth and mortality rates, then the above special steady solution only extincts if $\beta < (1 + \alpha(1 + \frac{\boldsymbol{\sigma}^2_\alpha}{2}))$.*

Remark 3.6 (Eigenproblem with $\alpha = 0$)**.** *The same ideas as in Proposition 3.4 yield the Gaussian solutions to the eigenproblem* (1.11) *in the absence of selection (*i.e.*, $\alpha = 0$). However, there is no longer uniqueness due to the translation invariance of $\mathcal{B}$. Indeed, we obtain the Gaussian solutions:*

$$\lambda = \boldsymbol{\lambda}_{\alpha=0} = 1, \quad F = \boldsymbol{F}_{\alpha=0}(\cdot - \mu) = G_{\mu,2},$$

for any $\mu \in \mathbb{R}^d$. Indeed, it is a consequence of the contraction property stated in Section 2.1 that these are all possible generic solutions (not only Gaussian). Specifically, by the conservation of mass (2.1) *we obtain that the only possible eigenvalue is $\lambda = 1$. Then, the eigenproblem* (1.11) *reduces to a fixed point equation for $\mathcal{B}$. We can then conclude by Corollary 2.2.*

We emphasize that the above presents a crucial difference of behavior between the nonlinear problem (1.11) with sexual reproduction and the analogous linear version with asexual reproduction operator (1.7). Namely, whilst in the former case there exists nontrivial solutions in the absence of selection, in the latter case it can be seen that such a nontrivial solution does not exists (*e.g.*, through Fourier arguments). This suggests that whilst in the linear problem both reproduction and selection stay balanced, in the nonlinear problem reproduction appears to dominates selection structurally.

Finally, we note that by iteration of the operator $\mathcal{T}$ and Lemma 3.3 we can compute the explicit form of solutions of the time-discrete problem (1.1) issued at Gaussian initial data. In fact, the Gaussian structure is preserved along generations, although mass, mean and variance are modified.

Proposition 3.7 (Gaussian solutions of the time-discrete problem)**.** *Consider any $m_0 \in \mathbb{R}_+^*$, $\mu_0 \in \mathbb{R}^d$ and $\sigma_0^2 \in \mathbb{R}_+$, and set the Gaussian initial datum $F_0 = m_0\, G_{\mu_0,\sigma_0^2}$. Hence, the solution $\{F_n\}_{n\in\mathbb{N}}$ to the time-discrete problem* (1.1) *takes the form*

$$F_n = m_n\, G_{\mu_n,\sigma_n^2}, \tag{3.4}$$

for $n \in \mathbb{N}$, where the parameters $m_n \in \mathbb{R}_+^*$, $\mu_n \in \mathbb{R}^d$ and $\sigma_n^2 \in \mathbb{R}_+^*$ are governed by the recursions:

$$m_n = m_{n-1}\frac{e^{-\frac{1}{2}\frac{\alpha|\mu_{n-1}|^2}{1+\alpha\left(1+\frac{\sigma_{n-1}^2}{2}\right)}}}{(1+\alpha(1+\frac{\sigma_{n-1}^2}{2}))^{d/2}}, \quad \mu_n = \frac{\mu_{n-1}}{1+\alpha(1+\frac{\sigma_{n-1}^2}{2})}, \quad \frac{1}{\sigma_n^2} = \alpha + \frac{1}{1+\frac{\sigma_{n-1}^2}{2}}. \tag{3.5}$$

Lemma 3.8 (Relaxation of variance)**.** *Assume that $\alpha \in \mathbb{R}_+^*$, consider any $\sigma_0^2 \in \mathbb{R}_+^*$ and define the sequence $\{\sigma_n^2\}_{n\in\mathbb{N}}$ by recursion according to the third recursion in* (3.5)*,* i.e.

$$\frac{1}{\sigma_n^2} = \alpha + \frac{1}{1+\frac{\sigma_{n-1}^2}{2}}, \tag{3.6}$$

for every $n \in \mathbb{N}$. Hence, if $\sigma_0^2 > \boldsymbol{\sigma}_\alpha^2$ then $\sigma_n \searrow \boldsymbol{\sigma}_\alpha^2$ as $n \to \infty$ and, if $\sigma_0^2 < \boldsymbol{\sigma}_\alpha^2$ then $\sigma_n \nearrow \boldsymbol{\sigma}_\alpha^2$ as $n \to \infty$, where $\boldsymbol{\sigma}_\alpha^2$ is given by (1.13)*. In addition, we obtain the convergence rates*

$$|\sigma_n^2 - \boldsymbol{\sigma}_\alpha^2| \le C_v\, \boldsymbol{r}_\alpha^n, \tag{3.7}$$

for every $n \in \mathbb{N}$, where the constant $C_v \in \mathbb{R}_+$ depends on σ_0^2 and α (we obtain that $C_v = 0$ if $\sigma_0^2 = \boldsymbol{\sigma}_\alpha^2$), and the ratio $\boldsymbol{r}_\alpha \in (0, \frac{1}{2})$ satisfies $\boldsymbol{r}_\alpha = 2\,\boldsymbol{k}_\alpha^2$ and is given explicitly by

$$\boldsymbol{r}_\alpha := \frac{8}{\left((2\alpha+3)+\sqrt{(2\alpha+1)^2+8\alpha}\right)^2}. \tag{3.8}$$

Proof. • Step 1: Monotone convergence of $\{\sigma_n^2\}_{n\in\mathbb{N}}$.
Consider the function $f : \mathbb{R}_+ \longrightarrow \mathbb{R}$ given by

$$f(x) = \alpha + \frac{1}{1+\frac{1}{2x}}, \quad x \in \mathbb{R}_+,$$

and, for simplicity of notation, define

$$x_* := (\boldsymbol{\sigma}_\alpha^2)^{-1}, \quad x_n := (\sigma_n^2)^{-1}, \quad n \in \mathbb{N}. \tag{3.9}$$

Notice that x_* is the unique fixed point of f in $\mathbb{R}_+$ and $\{x_n\}_{n\in\mathbb{N}}$ determines the fixed-point iteration of the map f issued at $x_0 = (\sigma_0^2)^{-1}$, that is, $x_n = f(x_{n-1}), \quad n \in \mathbb{N}$. Our goal is to show that $\{x_n\}_{n\in\mathbb{N}}$ converges towards x_* and find convergence rates. By direct computation we obtain $f'(x) = 2/(1+2x)^2$, which is above 1 near $x = 0$, and therefore f is not contractive. Hence, we cannot apply the usual Banach contraction principle and a different argument is provided. First, note that $x \in \mathbb{R}_+ \mapsto f(x)$ is strictly increasing and $x \in \mathbb{R}_+^* \mapsto \frac{f(x)}{x}$ is strictly decreasing. Then, we obtain that

$$\begin{aligned} &x < f(x) < x_*, && \text{if } 0 \le x < x_*,\\ &f(x) = x_*, && \text{if } x = x_*,\\ &x_* < f(x) < x, && \text{if } x > x_*, \end{aligned}$$

which implies the aforementioned monotonicity properties of $\{x_n\}_{n\in\mathbb{N}}$ and, in addition,

$$\min\{x_*, x_0\} \le x_n \le \max\{x_*, x_0\}, \tag{3.10}$$

for any $n \in \mathbb{N}$. In particular, this yields the monotonic convergence of $\{x_n\}_{n\in\mathbb{N}}$ towards x_*, or equivalently, the claimed monotonic convergence of $\{\sigma_n^2\}_{n\in\mathbb{N}}$ towards $\boldsymbol{\sigma}_\alpha^2$ as $n \to \infty$.

• Step 2: Convergence rates.
Second, to compute the convergence rates we shall use the special algebraic structure of function f, which is the restriction to $\mathbb{R}_+$ of the Möbius function $M : \mathbb{R} \setminus \{-\frac{1}{2}\} \longrightarrow \mathbb{R}$ given by

$$M(x) := \frac{2(\alpha+1)x+\alpha}{2x+1}, \quad x \in \mathbb{R} \setminus \left\{-\frac{1}{2}\right\}.$$

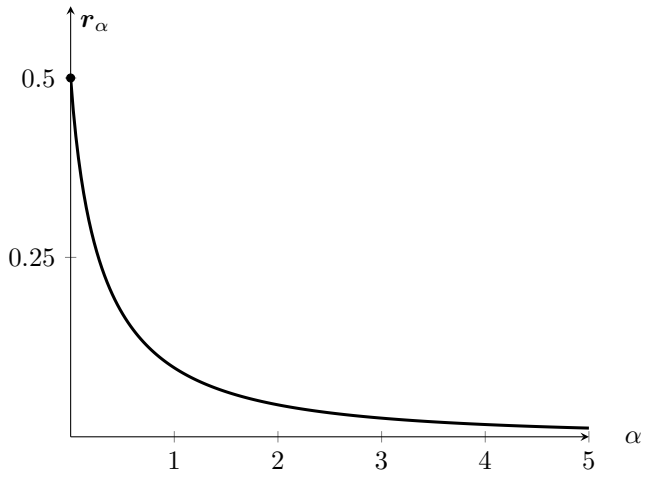


FIGURE 5. The ratio $\boldsymbol{r}_\alpha$ against parameter α.

Note that M has two fixed points:

$$x_\pm := \frac{2\alpha+1 \pm \sqrt{(2\alpha+1)^2+8\alpha}}{4},$$

where we note that $x_+ = x_* \in \mathbb{R}_+^*$ and $x_- \in \mathbb{R}^-$. Then, by definition of $\{x_n\}_{n\in\mathbb{N}}$ in (3.9) we obtain

$$x_n - x_+ = \frac{2(\alpha+1)x_{n-1}+\alpha}{2x_{n-1}+1} - \frac{2(\alpha+1)x_+ + \alpha}{2x_+ + 1} = \frac{2}{(2x_{n-1}+1)(2x_+ +1)}(x_{n-1}-x_+). \tag{3.11}$$

Since x_0 (thus x_{n-1} for initial time steps n) can be chosen arbitrary close to 0, then, the best a priori control that we can have on the prefactor in the right hand side of (3.11) is

$$0 \leq \frac{2}{(2x_{n-1}+1)(2x_+ +1)} \leq \frac{2\,\boldsymbol{\sigma}_\alpha^2}{2+\boldsymbol{\sigma}_\alpha^2} = 2\,\boldsymbol{k}_\alpha,$$

which gives true contraction in (3.11) when $\alpha > 0$, since $2\,\boldsymbol{k}_\alpha < 1$. Nevertheless, such a rate is non-optimal, and we show an alternative approach to achieve a sharper one. Specifically, note that

$$x_n - x_- = \frac{2(\alpha+1)x_{n-1}+\alpha}{2x_{n-1}+1} - \frac{2(\alpha+1)x_- + \alpha}{2x_- + 1} = \frac{2}{(2x_{n-1}+1)(2x_- +1)}(x_{n-1}-x_-). \tag{3.12}$$

By dividing (3.11) by (3.12) and iterating the identity we obtain

$$\frac{x_n - x_+}{x_n - x_-} = \left(\frac{2x_- +1}{2x_+ +1}\right)^n \frac{x_0 - x_+}{x_0 - x_-}, \tag{3.13}$$

for every $n \in \mathbb{N}$. In fact, the basis can be restated as follows

$$\frac{2x_- +1}{2x_+ +1} = \frac{(2\alpha+3)-\sqrt{(2\alpha+1)^2+8\alpha}}{(2\alpha+3)+\sqrt{(2\alpha+1)^2+8\alpha}} = \frac{8}{\left((2\alpha+3)+\sqrt{(2\alpha+1)^2+8\alpha}\right)^2} = \boldsymbol{r}_\alpha < \frac{1}{2}.$$

Therefore, using the uniform control (3.10) and (3.13), we conclude that

$$|x_n - x_*| = \frac{|x_n - x_-|}{|x_0 - x_-|}|x_0 - x_*|\,\boldsymbol{r}_\alpha^n \leq \frac{\max\{x_*, x_0\} - x_-}{\min\{x_*, x_0\} - x_-}|x_0 - x_*|\,\boldsymbol{r}_\alpha^n, \tag{3.14}$$

for every $n \in \mathbb{N}$, thus leading to an improved rate since $\boldsymbol{r}_\alpha = 2\,\boldsymbol{k}_\alpha^2 < 2\,\boldsymbol{k}_\alpha$. Finally, notice that

$$|\sigma_n^2 - \boldsymbol{\sigma}_\alpha^2| = \sigma_n^2\,\boldsymbol{\sigma}_\alpha^2\,|x_n - x_*| \leq \max\{\boldsymbol{\sigma}_\alpha^2, \sigma_0^2\}^2\,|x_n - x_*|, \tag{3.15}$$

for every $n \in \mathbb{N}$. Hence, joining (3.14), (3.15) along with (3.10) ends the proof. □

In particular, for Gaussian initial data we recover an explicit particular case of the general relaxation result in Theorem 1.1.

Corollary 3.9 (Relaxation of Gaussian solutions)**.** *Assume that $\alpha \in \mathbb{R}_+^*$, consider any $m_0 \in \mathbb{R}_+$, $\mu_0 \in \mathbb{R}^d$ and $\sigma_0^2 \in \mathbb{R}_+$, and set the Gaussian initial datum $F_0 = m_0\,G_{\mu_0,\sigma_0}$. Hence, the Gaussian solution $\{F_n\}_{n\in\mathbb{N}}$ to the time-discrete problem* (1.1) *verifies that the growth rates $\|F_n\|_{L^1(\mathbb{R}^d)}/\|F_{n-1}\|_{L^1(\mathbb{R}^d)}$ relax towards*

$\boldsymbol{\lambda}_\alpha$ and the normalized profiles $F_n/\|F_n\|_{L^1(\mathbb{R}^d)}$ relax towards $\boldsymbol{F}_\alpha$, where $(\boldsymbol{\lambda}_\alpha, \boldsymbol{F}_\alpha)$ is the unique Gaussian solution (1.12) *to the eigenproblem* (1.11) *as proved in Proposition 3.4. Specifically, for any $\varepsilon \in \mathbb{R}_+^*$*

$$\mathcal{D}_{\mathrm{KL}}\left(\frac{F_n}{\|F_n\|_{L^1(\mathbb{R}^d)}}\middle\| \boldsymbol{F}_\alpha\right) \leq C_{\mu,\varepsilon}(\boldsymbol{\lambda}_\alpha^{4/d}+\varepsilon)^n + C_v\,\boldsymbol{r}_\alpha^n,$$
$$\left|\frac{\|F_n\|_{L^1(\mathbb{R}^d)}}{\|F_{n-1}\|_{L^1(\mathbb{R}^d)}} - \boldsymbol{\lambda}_\alpha\right| \leq C_{\mu,\varepsilon}(\boldsymbol{\lambda}_\alpha^{4/d}+\varepsilon)^n + C_v\,\boldsymbol{r}_\alpha^n,$$

where $\mathcal{D}_{\mathrm{KL}}$ is the Kullback-Leibler divergence (or relative entropy), $\boldsymbol{r}_\alpha$ is given by (3.8) *in Lemma 3.8, $C_{\mu,\varepsilon} \in \mathbb{R}_+$ depends on α, μ_0, σ_0^2 and ε, and $C_v \in \mathbb{R}_+$ depends only α and σ_0^2. In fact, $C_{\mu,\varepsilon} = 0$ if $\mu_0 = 0$ and $C_v = 0$ if $\sigma_0^2 = \boldsymbol{\sigma}_\alpha^2$.*

Again, in this case the computations of the Kullback-Leibler divergence becomes explicit, and it is based on the general formula below for the divergence between two Gaussian density functions:

$$\mathcal{D}_{\mathrm{KL}}(G_{\mu_1,\Sigma_1}\|G_{\mu_2,\Sigma_2}) = \frac{1}{2}(\mu_2-\mu_1)^\top\Sigma_2^{-1}(\mu_2-\mu_1) + \frac{1}{2}\mathrm{trace}(\Sigma_1\Sigma_2^{-1}) - \frac{d}{2} + \frac{1}{2}\log\det(\Sigma_2\Sigma_1^{-1}), \quad (3.16)$$

for every $\mu_1, \mu_2 \in \mathbb{R}^d$ and any positive non-singular matrices $\Sigma_1, \Sigma_2 \in \mathbb{R}^{d\times d}$.

Proof of Corollary 3.9. First, note that when $\alpha \in \mathbb{R}_+^*$, then the parameters m_n, μ_n and σ_n^2 in (3.5) verify

$$\lim_{n\to\infty}\frac{m_n}{m_{n-1}} = \boldsymbol{\lambda}_\alpha, \quad \lim_{n\to\infty}\mu_n = 0, \quad \lim_{n\to\infty}\sigma_n^2 = \boldsymbol{\sigma}_\alpha^2\,.$$

In the sequel, we quantify the rates of convergence. Since σ_n^2 has already been studied in Lemma 3.8 we shall focus only on μ_n and m_n. Let us fix any arbitrarily small $\varepsilon \in \mathbb{R}_+^*$. On the one hand note that

$$\frac{|\mu_n|}{|\mu_{n-1}|} - \boldsymbol{\lambda}_\alpha^{2/d} = \frac{1}{1+\alpha\big(1+\frac{\sigma_{n-1}^2}{2}\big)} - \frac{1}{1+\alpha\big(1+\frac{\boldsymbol{\sigma}_\alpha^2}{2}\big)},$$

for every $n \in \mathbb{N}$. By the mean value theorem and Lemma 3.8 we obtain

$$\left|\frac{|\mu_n|}{|\mu_{n-1}|} - \boldsymbol{\lambda}_\alpha^{2/d}\right| \leq C_v\,\boldsymbol{r}_\alpha^n,$$

for an appropriate $C_v \in \mathbb{R}_+^*$ depending on α and σ_0^2. Since $\boldsymbol{\lambda}_\alpha \in (0,1)$ then d'Alembert's ratio test readily shows that μ_n relaxes to zero as a geometric sequence. Indeed, note that

$$|\mu_n| = \left(\frac{|\mu_n|}{|\mu_{n-1}|} - \boldsymbol{\lambda}_\alpha^{2/d}\right)|\mu_{n-1}| + \boldsymbol{\lambda}_\alpha^{2/d}\,|\mu_{n-1}| \leq (\boldsymbol{\lambda}_\alpha^{2/d}+C_v\,\boldsymbol{r}_\alpha^n)|\mu_{n-1}|,$$

for any $n \in \mathbb{N}$. By an inductive argument, this yields

$$|\mu_n| \leq \prod_{k=1}^n(\boldsymbol{\lambda}_\alpha^{2/d}+C_v\,\boldsymbol{r}_\alpha^n)|\mu_0| \lesssim C_\varepsilon(\boldsymbol{\lambda}_\alpha^{2/d}+\varepsilon)^n, \quad (3.17)$$

for sufficiently large $C_{\mu,\varepsilon} \in \mathbb{R}^d$, where we have used that $\boldsymbol{r}_\alpha \in (0,1)$ to absorb $C_v\,\boldsymbol{r}_\alpha^n$ in an ε-small term. On the other hand, note that

$$\begin{aligned}\frac{m_n}{m_{n-1}} - \boldsymbol{\lambda}_\alpha &= \frac{e^{-\frac{1}{2}\frac{\alpha|\mu_{n-1}|^2}{1+\alpha\left(1+\frac{\sigma_{n-1}^2}{2}\right)}}}{(1+\alpha(1+\frac{\sigma_{n-1}^2}{2}))^{d/2}} - \frac{1}{(1+\alpha(1+\frac{\boldsymbol{\sigma}_\alpha^2}{2}))^{d/2}}\\ &= \frac{e^{-\frac{1}{2}\frac{\alpha|\mu_{n-1}|^2}{1+\alpha\left(1+\frac{\sigma_{n-1}^2}{2}\right)}}-1}{(1+\alpha(1+\frac{\sigma_{n-1}^2}{2}))^{d/2}} + \left(\frac{1}{(1+\alpha(1+\frac{\sigma_{n-1}^2}{2}))^{d/2}} - \frac{1}{(1+\alpha(1+\frac{\boldsymbol{\sigma}_\alpha^2}{2}))^{d/2}}\right).\end{aligned}$$

By the mean value theorem,

$$\left|\frac{m_n}{m_{n-1}} - \boldsymbol{\lambda}_\alpha\right| \leq \frac{\alpha|\mu_{n-1}|^2}{2(1+\alpha)^{\frac{d}{2}+1}} + \frac{d\alpha}{4}\frac{|\sigma_{n-1}^2 - \boldsymbol{\sigma}_\alpha^2\,|}{(1+\alpha)^{\frac{d}{2}+1}}.$$

Therefore, using contraction of mean (3.17) and contraction of variance (3.7) in Lemma 3.8 entail the rate of convergence for the m_n/m_{n-1}. Finally, given that both $F_n/\|F_n\|_{L^1(\mathbb{R}^d)}$ and $\boldsymbol{F}_\alpha$ are Gaussian functions, the relative entropy can be computed explicitly through (3.16). Specifically,

$$\mathcal{D}_{\mathrm{KL}}\left(\frac{F_n}{\|F_n\|_{L^1(\mathbb{R}^d)}}\middle\|\boldsymbol{F}_\alpha\right) = \frac{|\mu_n|^2}{2\,\boldsymbol{\sigma}_\alpha^2} + \frac{d}{2}\left(\frac{\sigma_n^2}{\boldsymbol{\sigma}_\alpha^2}-1\right) - \frac{d}{2}\log\left(\frac{\sigma_n^2}{\boldsymbol{\sigma}_\alpha^2}\right).$$

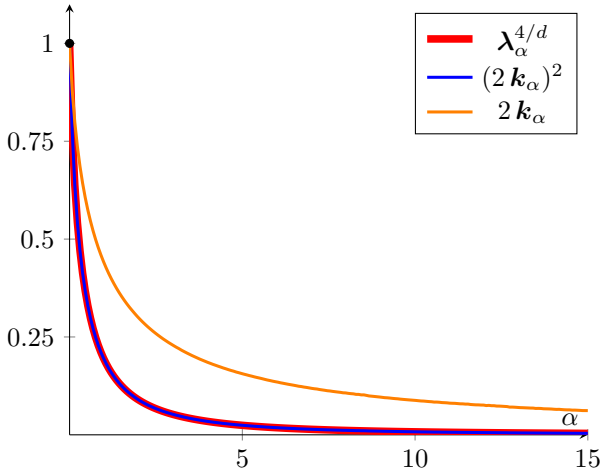


FIGURE 6. Comparison of convergence rates in Theorem 1.1 and Corollary 3.9

Hence, using again (3.17) and (3.7) concludes our result. □

The proof of Theorem 1.1 for generic non-Gaussian initial data $F_0 \in \mathcal{M}_+(\mathbb{R}^d)$ will be the core of this paper and we postpone it to Section 6.

Remark 3.10 (Optimality of convergence rates)**.** *As illustrated in Figure 6, the convergence rate of the normalized profiles obtained in our main Theorem 1.1 (blue line) is sharp, compared to the explicit one found in Corollary 3.9 (red line) for the special class of Gaussian solutions. In fact, one can easily check that the identity* $\lambda_\alpha^{4/d} = (2\,k_\alpha)^2$ *holds. However, there is a mismatch between the convergence rate of the rate of growth of mass in Theorem 1.1 (orange line) and the sharp rate for Gaussian solutions found in Corollary 3.9 (again red line). On the one hand, it stands to reason that the relative entropy has a certain quadratic structure, whilst the rate of growth of mass is related to* L^1 *norms, and then it is not quadratic. We will see that more clearly in the proof of Theorem 1.1 in Section 6, where we use explicitly the following relation*

$$\left| \frac{\|F_n\|_{L^1(\mathbb{R}^d)}}{\|F_{n-1}\|_{L^1(\mathbb{R}^d)}} - \lambda_\alpha \right| \lesssim \sqrt{\mathcal{D}_{\mathrm{KL}}\left(\frac{F_n}{\|F_n\|_{L^1(\mathbb{R}^d)}} \middle\| \boldsymbol{F}_\alpha \right)},$$

based on Pinsker's inequality. However, a certain quadratic structure is still present in the rate of convergence of m_n/m_{n-1} *in Corollary 3.9, which may become explicit for generic solutions to* (1.1) *if one finds the hidden cancellations. However, for simplicity we do not address this technical detail here.*

4. Some properties of the operator $\mathcal{T}$

First, note that for any $F \in \mathcal{M}_+(\mathbb{R}^d)$ we obtain that $\mathcal{T}[F] \in W^{k,1}(\mathbb{R}^d) \cap W^{k,\infty}(\mathbb{R}^d)$ for each $k \in \mathbb{N}$ thanks to the fact that F has finite mass and the Gaussian mixing kernel G in (1.4) is smooth and has bounded and integrable derivatives of any order. In particular, solutions $\{F_n\}_{n\in\mathbb{N}}$ to the time-discrete problem (1.1) become instantaneously smooth after the first generation for any generic initial datum $F_0 \in \mathcal{M}_+(\mathbb{R}^d)$. In the sequel, we will derive a quantitative control on the emergence of Gaussian tails for $\mathcal{T}[F]$. In addition, we will quantify the propagation of quadratic and exponential moments for the normalized profile $\mathcal{T}[F]/\|\mathcal{T}[F]\|_{L^1(\mathbb{R}^d)}$. Both *a priori* estimates will be required later in Section 6.

4.1. Emergence of Gaussian tails.

Lemma 4.1 (Emergence of tails)**.** *Assume that* $\alpha \in \mathbb{R}_+^*$*, consider* $F \in \mathcal{M}_+(\mathbb{R}^d)$ *and set* $\overline{\sigma}^2, \underline{\sigma}^2 \in \mathbb{R}_+^*$ *by*

$$\overline{\sigma}^2 := \frac{1}{\alpha}, \quad \underline{\sigma}^2 \in \left(0, \frac{1}{1+\alpha}\right).$$

Then, the following properties are fulfilled:

(i) **(Upper control of tails)** *There exists* $C = C(\alpha, F) > 0$ *such that*

$$\mathcal{T}[F](x) \le C\, G_{0,\overline{\sigma}^2}(x), \tag{4.1}$$

for every $x \in \mathbb{R}^d$*.*

(ii) **(Lower control of tails)** *There exists $c = c(\alpha, \underline{\sigma}^2, F) > 0$ such that*

$$\mathcal{T}[F](x) \geq c\, G_{0,\underline{\sigma}^2}(x), \tag{4.2}$$

for every $x \in \mathbb{R}^d$.

(iii) **(Control of log-derivative)** *Assume that F is absolutely continuous with respect to the Lebesgue measure and it has Gaussian tail,* i.e.*, there exists $\sigma^2 \in \mathbb{R}_+^*$ and $C' \in \mathbb{R}_+^*$ such that $F(x) \leq C' G_{0,\sigma^2}$ for every $x \in \mathbb{R}^d$. Then, there exists $C'' = C''(\alpha, \underline{\sigma}^2, \sigma^2, C', F) > 0$ such that*

$$|\nabla \log \mathcal{T}[F](x)| \leq C''(1 + |x|), \tag{4.3}$$

for every $x \in \mathbb{R}^d$.

Proof. First, notice that $F \in \mathcal{M}_+(\mathbb{R}^d)$ is any generic distribution but $G \in L^\infty(\mathbb{R}^d)$. Then, $\mathcal{B}[F] \in L^\infty(\mathbb{R}^d)$ and, by definition of $\mathcal{T}$ in (1.2), we obtain

$$\mathcal{T}[F](x) \leq \|\mathcal{B}[F]\|_{L^\infty(\mathbb{R}^d)} e^{-\frac{\alpha}{2}|x|^2} = \|\mathcal{B}[F]\|_{L^\infty(\mathbb{R}^d)} \left(2\pi\alpha^{-1}\right)^{d/2} G_{0,\overline{\sigma}^2}(x),$$

for every $x \in \mathbb{R}^d$. Hence, (4.1) holds for appropriate $C > 0$. Second, note that

$$\begin{aligned}
\mathcal{T}[F](x) &= \frac{e^{-\frac{\alpha}{2}|x|^2}}{(2\pi)^{d/2}\|F\|_{\mathcal{M}_+(\mathbb{R}^d)}} \int_{\mathbb{R}^{2d}} \exp\left(-\frac{1}{2}\left|x - \frac{x_1+x_2}{2}\right|^2\right) F(dx_1)F(dx_2) \\
&= \frac{e^{-\frac{\alpha}{2}|x|^2}}{(2\pi)^{d/2}\|F\|_{\mathcal{M}_+(\mathbb{R}^d)}} \int_{\mathbb{R}^{2d}} \exp\left(-\frac{1}{2}|x|^2 - \frac{1}{8}|x_1+x_2|^2 + \frac{1}{2}x\cdot(x_1+x_2)\right) F(dx_1)F(dx_2) \\
&\geq \frac{e^{-\frac{\alpha}{2}|x|^2}}{(2\pi)^{d/2}\|F\|_{\mathcal{M}_+(\mathbb{R}^d)}} \int_{\mathbb{R}^{2d}} \exp\left(-\frac{2+\varepsilon}{4}|x|^2 - \frac{2+\varepsilon}{8\varepsilon}|x_1+x_2|^2\right) F(dx_1)F(dx_2) \\
&\geq \frac{e^{-\frac{\alpha}{2}|x|^2}}{(2\pi)^{d/2}\|F\|_{\mathcal{M}_+(\mathbb{R}^d)}} \int_{\mathbb{R}^{2d}} \exp\left(-\frac{2+\varepsilon}{4}|x|^2 - \frac{2+\varepsilon}{4\varepsilon}|x_1|^2 - \frac{2+\varepsilon}{4\varepsilon}|x_2|^2\right) F(dx_1)F(dx_2),
\end{aligned}$$

for every $\varepsilon > 0$, where in the third and last lines we have used Cauchy–Schwarz's and Young's inequalities. Consequently, we obtain that

$$\mathcal{T}[F](x) \geq (2\pi)^{d/2} \left(\frac{2\varepsilon}{2+\varepsilon}\right)^d \frac{\|G_{0,\frac{2\varepsilon}{2+\varepsilon}} F\|^2_{\mathcal{M}_+(\mathbb{R}^d)}}{\|F\|_{\mathcal{M}_+(\mathbb{R}^d)}} e^{-\frac{\alpha}{2}|x|^2} e^{-\frac{2+\varepsilon}{4}|x|^2},$$

for every $x \in \mathbb{R}^d$. Taking $\varepsilon > 0$ small enough and an appropriate constant $c > 0$ yields (4.2). Finally, taking derivatives on (1.2) we obtain

$$\begin{aligned}
\nabla\mathcal{T}[F](x) &= -\alpha\, x\, \mathcal{T}[F](x) - \frac{e^{-\frac{\alpha}{2}|x|^2}}{\|F\|_{L^1(\mathbb{R}^d)}} \int_{\mathbb{R}^{2d}} \left(x - \frac{x_1+x_2}{2}\right) G\left(x - \frac{x_1+x_2}{2}\right) F(dx_1)\, F(dx_2) \\
&= -(1+\alpha)\, x\, \mathcal{T}[F](x) - \frac{e^{-\frac{\alpha}{2}|x|^2}}{\|F\|_{L^1(\mathbb{R}^d)}} \int_{\mathbb{R}^{2d}} \left(\frac{x_1+x_2}{2}\right) G\left(x - \frac{x_1+x_2}{2}\right) F(dx_1)\, F(dx_2).
\end{aligned}$$

Therefore, dividing by $\mathcal{T}[F](x)$ and taking norms yield the following estimate

$$|\nabla \log \mathcal{T}[F](x)| \leq (1+\alpha)|x| + \frac{1}{\sqrt{2}\|F\|_{L^1(\mathbb{R}^d)}} \frac{e^{-\frac{\alpha}{2}|x|^2}}{\mathcal{T}[F](x)} \int_{\mathbb{R}^{2d}} |(x_1,x_2)|\, G\left(x - \frac{x_1+x_2}{2}\right) F(dx_1)F(dx_2), \tag{4.4}$$

for every $x \in \mathbb{R}^d$. Consider any $R > 0$ and split the integral in the right hand side of (4.4) into the subsets $\{(x_1,x_2) \in \mathbb{R}^{2d} : |(x_1,x_2)| \leq \sqrt{2}R|x|\}$ and $\{(x_1,x_2) \in \mathbb{R}^{2d} : |(x_1,x_2)| > \sqrt{2}R|x|\}$. For the first subset, we use the definition (1.2) of the operator $\mathcal{T}$. For the second subset, we use the uniform bound

of G along with the assumptions on F and the preceding lower bound (4.2) of $\mathcal{T}[F]$. Then, we obtain

$$\begin{aligned}
|\nabla \log \mathcal{T}[F](x)| &\leq (1+\alpha+R)|x| + \frac{C^2 \underline{\sigma}^d e^{-\frac{\alpha}{2}|x|^2} e^{\frac{1}{2\underline{\sigma}^2}|x|^2}}{\sqrt{2}c\|F\|_{L^1(\mathbb{R}^d)}} \int_{|(x_1,x_2)|>\sqrt{2}R|x|} |(x_1,x_2)| e^{-\frac{1}{2\sigma^2}|(x_1,x_2)|^2}\, dx_1\, dx_2 \\
&= (1+\alpha+R)|x| + \frac{C^2 \underline{\sigma}^d e^{-\frac{\alpha}{2}|x|^2} e^{\frac{1}{2\underline{\sigma}^2}|x|^2}}{\sqrt{2}c\|F\|_{L^1(\mathbb{R}^d)}} \int_{\sqrt{2}R|x|}^{+\infty} r^{2d} e^{-\frac{1}{2\sigma^2}r^2}\, dr \\
&\leq (1+\alpha+R)|x| + \frac{C^2 \underline{\sigma}^d e^{-\frac{\alpha}{2}|x|^2} e^{\frac{1}{2\underline{\sigma}^2}|x|^2}}{\sqrt{2}c\|F\|_{L^1(\mathbb{R}^d)}} \left(\frac{8d-4}{e}\right)^{d-\frac{1}{2}} \sigma^{2d-1} \int_{\sqrt{2}R|x|}^{+\infty} r e^{-\frac{1}{4\sigma^2}r^2}\, dr \\
&= (1+\alpha+R)|x| + \frac{\sqrt{2}C^2 \underline{\sigma}^d \sigma^{2d+1}}{c\|F\|_{L^1(\mathbb{R}^d)}} \left(\frac{8d-4}{e}\right)^{d-\frac{1}{2}} e^{-\frac{1}{2}\left(\alpha+\frac{R^2}{\sigma^2}-\frac{1}{\underline{\sigma}^2}\right)|x|^2},
\end{aligned}$$

for every $x \in \mathbb{R}^d$, where in the third line we have used the inequality

$$r^{2d-1} \leq \left(\frac{2d-1}{e}\right)^{d-\frac{1}{2}} \left(\frac{\sigma}{\varepsilon}\right)^{2d-1} \exp\left(\frac{\varepsilon}{2\sigma^2} r^2\right),$$

for every $r > 0$ and $\varepsilon > 0$ in the particular case of $\varepsilon = 1/2$. Taking $R > 0$ large enough ends the proof. □

The previous result can be iterated to obtain similar properties for solutions $\{F_n\}_{n\in\mathbb{N}}$ to the time-discrete problem (1.1) issued at a generic initial datum $F_0 \in \mathcal{M}_+(\mathbb{R}^d)$. In particular, we note that Gaussian tails emerge instantaneously after the first generation $n = 1$ and linear growth of the log-derivative is guaranteed after the second generation $n = 2$.

Corollary 4.2 (Propagation of tails)**.** *Assume that $\alpha \in \mathbb{R}_+^*$, consider the solution $\{F_n\}_{n\in\mathbb{N}}$ of (1.1) issued at a generic initial datum $F_0 \in \mathcal{M}_+(\mathbb{R}^d)$ and set $\overline{\sigma}_1^2, \underline{\sigma}_1^2 \in \mathbb{R}_+^*$ by*

$$\overline{\sigma}_1^2 := \frac{1}{\alpha}, \quad \underline{\sigma}_1^2 \in \left(0, \frac{1}{1+\alpha}\right).$$

Define the associated sequences of variances $\{\overline{\sigma}_n^2\}_{n\in\mathbb{N}}$ and $\{\underline{\sigma}_n^2\}_{n\in\mathbb{N}}$ by the following recursive relations

$$\frac{1}{\overline{\sigma}_{n+1}^2} := \alpha + \frac{1}{1+\frac{\overline{\sigma}_n^2}{2}}, \quad \frac{1}{\underline{\sigma}_{n+1}^2} := \alpha + \frac{1}{1+\frac{\underline{\sigma}_n^2}{2}}, \tag{4.5}$$

for every $n \in \mathbb{N}$. Then, the following properties are fulfilled:

(i) **(Upper control of tails of F_n)** *There exist $C_n = C_n(\alpha, F_0) > 0$ such that*

$$F_n(x) \leq C_n\, G_{0,\overline{\sigma}_n^2}(x), \tag{4.6}$$

for each $x \in \mathbb{R}^d$ and every $n \in \mathbb{N}$.

(ii) **(Lower control of tails of F_n)** *There exist $c_n = c_n(\alpha, \underline{\sigma}_1^2, F_0) > 0$ such that*

$$F_n(x) \geq c_n\, G_{0,\underline{\sigma}_n^2}(x), \tag{4.7}$$

for each $x \in \mathbb{R}^d$ and every $n \in \mathbb{N}$.

(iii) **(Control of log-derivative of F_n)** *There exists $D_n = D_n(\alpha, \underline{\sigma}_1^2, F_0)$ such that*

$$|\nabla \log F_n(x)| \leq D_n(1+|x|), \tag{4.8}$$

for each $x \in \mathbb{R}^d$ and every $n \geq 2$.

Proof. Note that once (4.6) and (4.7) are proved, we can readily apply (4.3) in Lemma 4.1 to $F = F_{n-1}$ with $n \geq 2$ (which satisfies the required upper control by a Gaussian function) and we recover the estimate (4.8) for the log-derivative of $F_n = \mathcal{T}[F_{n-1}]$. Then, we just focus on proving estimates (4.6) and (4.7) through an inductive argument. First, note that Lemma 4.1 applied to $F = F_0$ yields (4.6) and (4.7) with $n = 1$. Let us assume that (4.6) and (4.7) hold for some $n \in \mathbb{N}$ and let us prove it for $n+1$. Specifically, using the induction hypothesis we obtain that

$$\begin{aligned}
F_{n+1}(x) &= e^{-\frac{\alpha}{2}|x|^2} \mathcal{B}[F_n](x) \\
&\leq \frac{C_n^2}{\|F_n\|_{L^1(\mathbb{R}^d)}} e^{-\frac{\alpha}{2}|x|^2} \left(G\left(\frac{\cdot}{2}\right) * G_{0,\overline{\sigma}_n^2} * G_{0,\overline{\sigma}_n^2}\right)(2x) \\
&= \frac{C_n^2}{\|F_n\|_{L^1(\mathbb{R}^d)}} e^{-\frac{\alpha}{2}|x|^2} G_{0,1+\frac{\overline{\sigma}_n^2}{2}}(x) \leq C_{n+1} G_{0,\overline{\sigma}_{n+1}^2}(x),
\end{aligned}$$

for every $x \in \mathbb{R}^d$ and appropriate C_{n+1}, where we have used Lemma 3.2 for the stability of Gaussian under convolutions and the definition (4.5) of $\overline{\sigma}_{n+1}^2$. This proves (4.6) for $n+1$ and a similar argument yields the lower estimate (4.7). □

By Lemma 3.8 we note that the sequence of variances $\{\overline{\sigma}_n^2\}_{n\in\mathbb{N}}$ and $\{\underline{\sigma}_n^2\}_{n\in\mathbb{N}}$ relax towards the asymptotic value $\boldsymbol{\sigma}_\alpha^2$ with a geometric convergence rate. This is consistent with our main result in Theorem 1.1 and Corollary 1.2. In fact, this suggests that any solution $\{F_n\}_{n\in\mathbb{N}}$ to the time-discrete problem (1.1) must relax towards the asymptotic profile $\boldsymbol{F}_\alpha$. In addition, recall that for any solution (λ, F) of the eigenproblem (1.11) we recover a particular solution of the time-discrete problem (1.1) via the ansatz (1.10), *i.e.*, $F_n = \lambda^n F$. Hence, we expect that $(\boldsymbol{\lambda}_\alpha, \boldsymbol{F}_\alpha)$ must indeed be the unique solution to the eigenvalue problem (1.11). However, we are still far from characterizing the full long-term dynamics for the solution $\{F_n\}_{n\in\mathbb{N}}$ in Theorem 1.1 and the uniqueness result in Corollary 1.2. Namely, the above coefficients C_n and c_n contains crucial information about the balance of mass and their values in Corollary 4.2 are not necessarily optimal. Indeed, note that they are given by explicit recursive formulas

$$C_{n+1} = (1-\alpha\overline{\sigma}_n^2)^{d/2}\frac{C_n^2}{\|F_n\|_{L^1(\mathbb{R}^d)}}, \quad c_{n+1} = (1-\alpha\overline{\sigma}_n^2)^{d/2}\frac{c_n^2}{\|F_n\|_{L^1(\mathbb{R}^d)}},$$

but they require further knowledge about the behavior of $\|F_n\|_{L^1(\mathbb{R}^d)}$. In the next paragraph, we overcome this important issue of scaling factor by focusing on renormalized profiles, for which we prove uniform propagation of moments.

4.2. **Propagation of quadratic and exponential moments.** In this section we analyze the propagation of quadratic and exponential moments along the normalized profiles $F_n/\|F_n\|_{L^1(\mathbb{R}^d)}$, where $\{F_n\}_{n\in\mathbb{N}}$ is a solution to (1.1). As we anticipated in Section 2.4, we need to move from uniform estimates of the high-dimensional integral (2.11) to estimates in an average sense. This will be precisely the point in proof of the main Theorem 1.1 in Section 6.2 where a suitable control of moments will be crucial, as sketched in the overall map in Figure 2. Before addressing the case of generic initial data $F_0 \in \mathcal{M}_+(\mathbb{R}^d)$, we illustrate the particular explicit case of Gaussian initial data $F_0 = m_0\, G_{\mu_0,\sigma_0^2}$.

Remark 4.3 (Gaussian case)**.** *Consider any $m_0 \in \mathbb{R}_+$, $\mu_0 \in \mathbb{R}^d$ and $\sigma_0^2 \in \mathbb{R}_+$, set $\theta \in \mathbb{R}_+^*$ with $\theta < \frac{1}{2\max\{\sigma_0^2, \boldsymbol{\sigma}_\alpha^2\}}$ and consider the solution $\{F_n\}_{n\in\mathbb{N}}$ to the time-discrete problem* (1.1) *with initial datum $F_0 = m_0 G_{\mu_0,\sigma_0^2}$. Then, by direct calculation we obtain*

$$\int_{\mathbb{R}^d} |x|^2 \frac{F_n(x)}{\|F_n\|_{L^1(\mathbb{R}^d)}}\, dx = |\mu_n|^2 + \sigma_n^2,$$

$$\int_{\mathbb{R}^d} e^{\theta|x|^2} \frac{F_n(x)}{\|F_n\|_{L^1(\mathbb{R}^d)}}\, dx = \frac{e^{\frac{\theta}{1-2\theta\sigma_n^2}|\mu_n|^2}}{(1-2\theta\sigma_n^2)^{d/2}},$$

Here, $\{\mu_n\}_{n\in\mathbb{N}}$ and $\{\sigma_n^2\}_{n\in\mathbb{N}}$ are the mean and variance of $\{F_n\}_{n\in\mathbb{N}}$ according to formula (3.5) *in Proposition 3.7. As studied in the proof of Corollary 3.9 we have the uniform bounds*

$$|\mu_n| \le |\mu_0| \quad \textit{and} \quad \min\{\sigma_0^2, \boldsymbol{\sigma}_\alpha^2\} \le \sigma_n^2 \le \max\{\sigma_0^2, \boldsymbol{\sigma}_\alpha^2\},$$

for every $n \in \mathbb{N}$. Then, we obtain that both the quadratic and exponential moments are uniformly bounded.

In the sequel, we explore the case of generic initial data $F_0 \in \mathcal{M}_+(\mathbb{R})$. For simplicity of notation, we define the following conservative operator associated with $\mathcal{T}$.

Definition 4.4 (Normalization of $\mathcal{T}$)**.**

$$\mathcal{S}[F] := \frac{\mathcal{T}[F]}{\|\mathcal{T}[F]\|_{L^1(\mathbb{R}^d)}}, \quad F \in \mathcal{M}_+(\mathbb{R}^d) \setminus \{0\}.$$

Therefore, we obtain the following control of quadratic moments under $\mathcal{S}$.

Lemma 4.5 (Control of quadratic moments)**.** *Assume that $\alpha \in \mathbb{R}_+^*$ and set any parameter $\eta \in \left(\frac{1}{2(1+\alpha)^2}, 1\right)$. Then, there exists a constant $M = M(\alpha,\eta) \in \mathbb{R}_+^*$ such that*

$$\int_{\mathbb{R}^d} |x|^2 \mathcal{S}[F](x)\, dx \le M + \eta \int_{\mathbb{R}^d} |x|^2 \frac{F(dx)}{\|F\|_{\mathcal{M}(\mathbb{R}^d)}},$$

for any measure $F \in \mathcal{M}_+(\mathbb{R}^d)$.

Before we proceed with the proof, let us comment on some apparent difficulties. For this discussion, we assume normalized $F \in \mathcal{P}(\mathbb{R}^d)$ without loss of generality, as we are only concerned with the shape of profiles here. Note that the conservative operator $\mathcal{S}$ in Definition 4.4 associated with $\mathcal{T}$ is actually the composition of the conservative operator $\mathcal{B}$, and the conservative multiplicative operator $\mathcal{M}$ in Definition 2.3, namely, $\mathcal{S} = \mathcal{M} \circ \mathcal{B}$. The former is invariant by translation, whereas the latter is not. However, we have the following inequality for the operator $\mathcal{M}$:

$$\int_{\mathbb{R}^d} |x|^2 \mathcal{M}[F](x)\, dx \leq \int_{\mathbb{R}^d} |x|^2 F(dx)\,. \tag{4.9}$$

This inequality is intuitively clear: by applying the selection operator $\mathcal{M}$, the quadratic moment (that is, the Wasserstein distance to the Dirac distribution at the origin) is non-increasing. In fact, the proof is a straightforward consequence of the following identity:

$$\int_{\mathbb{R}^d} |x|^2 \mathcal{M}[F](x)\, dx - \int_{\mathbb{R}^d} |x|^2 F(dx) = \frac{1}{2} \int_{\mathbb{R}^{2d}} \left(|x|^2 - |y|^2\right) \left(e^{-m(x)} - e^{-m(y)}\right) \frac{F(dx)F(dy)}{\int_{\mathbb{R}^d} e^{-m(z)} F(dz)}\,.$$

Note that the right-hand-side of this equality is indeed non-positive, provided that m is radially non-decreasing as it is the case for our quadratic choice (1.5). A similar estimate can also be tested on the operator $\mathcal{B}$. Specifically, by a simple change of variable the following identity holds true:

$$\int_{\mathbb{R}^d} |x|^2 \mathcal{B}[F](x)\, dx = \operatorname{Var} G + \frac{1}{2} \int_{\mathbb{R}^d} |x|^2\, F(dx) + \frac{1}{2} \left| \int_{\mathbb{R}^d} x\, F(dx) \right|^2 . \tag{4.10}$$

Then, it might seem natural to combine these two relationships in order to control the quadratic moments of the composition $\mathcal{S} = \mathcal{M} \circ \mathcal{B}$. Suppose, for instance, that the center of mass in the last term of (4.10) can be bounded *a priori* uniformly on F. Then, the combination of (4.9) and (4.10) yields the contraction property in Lemma 4.5 due to the reduction by half in the size of the quadratic moment of F. Of course, without such a uniform control on the center of mass, the last two terms in (4.10) would contribute together (by Jensen's inequality) with a merely non-expansive factor $\int_{\mathbb{R}^d} |x|^2\, F(dx)$ that is not enough for our purpose (we will iterate Lemma 4.5 on the F_n later). This *a priori* estimate of the center of mass should then be derived mainly from the selection component $\mathcal{M}$, as the reproduction component $\mathcal{B}$ is invariant by translation. However, it cannot be derived solely from $\mathcal{M}$, because the latter can have a dramatic effect on nearly symmetric distributions. This can be seen on the same configuration as in Example 2.4 above: for a sum of nearly symmetrical Dirac masses, the center of mass is close to zero before selection, and close to one of the endpoints after selection (when h is large). Nevertheless, such a configuration is destroyed by $\mathcal{B}$, so that the combination of $\mathcal{M}$ and $\mathcal{B}$ is crucial to guarantee the uniform boundedness of the quadratic moments as shown in the argument below.

Proof of Lemma 4.5. Set any $F \in \mathcal{M}_+(\mathbb{R}^d)$. First, note that

$$\int_{\mathbb{R}^d} |x|^2 \mathcal{S}[F](x)\, dx = \frac{\displaystyle\int_{\mathbb{R}^{3d}} |x|^2 e^{-m(x)} G\left(x - \frac{x_1+x_2}{2}\right) \frac{F(dx_1)}{\|F\|_{\mathcal{M}(\mathbb{R}^d)}} \frac{F(dx_2)}{\|F\|_{\mathcal{M}(\mathbb{R}^d)}} dx}{\displaystyle\int_{\mathbb{R}^{3d}} e^{-m(x)} G\left(x - \frac{x_1+x_2}{2}\right) \frac{F(dx_1)}{\|F\|_{\mathcal{M}(\mathbb{R}^d)}} \frac{F(dx_2)}{\|F\|_{\mathcal{M}(\mathbb{R}^d)}}\, dx}.$$

Since G is a Gaussian function given by (1.4) and m is a quadratic function given by (1.5), then we can compute explicitly the integrals with respect to x in the numerator and denominator. Specifically, for fixed $x_1, x_2 \in \mathbb{R}^d$ we define $\bar{x} := \frac{x_1+x_2}{2}$ and we obtain

$$\int_{\mathbb{R}^d} |x|^2 e^{-m(x)} G(x - \bar{x})\, dx = \frac{1}{(1+\alpha)^{d/2}} \exp\left(-\frac{1}{2}\frac{\alpha}{1+\alpha}|\bar{x}|^2\right) \left(\frac{1}{1+\alpha} + \frac{|\bar{x}|^2}{(1+\alpha)^2}\right), \tag{4.11}$$

$$\int_{\mathbb{R}^d} e^{-m(x)} G(x - \bar{x})\, dx = \frac{1}{(1+\alpha)^{d/2}} \exp\left(-\frac{1}{2}\frac{\alpha}{1+\alpha}|\bar{x}|^2\right). \tag{4.12}$$

Therefore, we can write alternatively

$$\int_{\mathbb{R}^d} |x|^2 \mathcal{S}[F](x)\, dx = \frac{1}{1+\alpha} + \frac{1}{(1+\alpha)^2} \frac{\displaystyle\int_{\mathbb{R}^{2d}} |\bar{x}|^2 \exp\left(-\frac{1}{2}\frac{\alpha}{1+\alpha}|\bar{x}|^2\right) \frac{F(dx_1)}{\|F\|_{\mathcal{M}(\mathbb{R}^d)}} \frac{F(dx_2)}{\|F\|_{\mathcal{M}(\mathbb{R}^d)}}}{\displaystyle\int_{\mathbb{R}^{2d}} \exp\left(-\frac{1}{2}\frac{\alpha}{1+\alpha}|\bar{x}|^2\right) \frac{F(dx_1)}{\|F\|_{\mathcal{M}(\mathbb{R}^d)}} \frac{F(dx_2)}{\|F\|_{\mathcal{M}(\mathbb{R}^d)}}}. \tag{4.13}$$

Our ultimate goal is to bound the second term in (4.13). To this end, we provide a lower bound for the denominator and an upper bound for the numerator in two separate steps.

• Step 1: Consider an arbitrary $R_1 \in \mathbb{R}_+^*$ to be determined later. Then, we obtain

$$\begin{aligned}
&\int_{\mathbb{R}^{2d}} \exp\left(-\frac{1}{2}\frac{\alpha}{1+\alpha}|\bar{x}|^2\right) \frac{F(dx_1)}{\|F\|_{\mathcal{M}(\mathbb{R}^d)}} \frac{F(dx_2)}{\|F\|_{\mathcal{M}(\mathbb{R}^d)}} \\
&\quad \geq \int_{|\bar{x}|\leq R_1} \exp\left(-\frac{1}{2}\frac{\alpha}{1+\alpha}|\bar{x}|^2\right) \frac{F(x_1)}{\|F\|_{L^1(\mathbb{R}^d)}} \frac{F(x_2)}{\|F\|_{L^1(\mathbb{R}^d)}}\, dx_1\, dx_2 \\
&\quad \geq \exp\left(-\frac{1}{2}\frac{\alpha}{1+\alpha}R_1^2\right)\left(1 - \int_{|\bar{x}|>R_1} \frac{F(dx_1)}{\|F\|_{\mathcal{M}(\mathbb{R}^d)}} \frac{F(dx_2)}{\|F\|_{\mathcal{M}(\mathbb{R}^d)}}\right) \\
&\quad \geq \exp\left(-\frac{1}{2}\frac{\alpha}{1+\alpha}R_1^2\right)\left(1 - \frac{1}{2R_1^2}\int_{\mathbb{R}^d} |x|^2 \frac{F(dx)}{\|F\|_{\mathcal{M}(\mathbb{R}^d)}}\right),
\end{aligned} \tag{4.14}$$

where in last line we have used Young's and Chebyshev's inequalities.

• Step 2: Consider an arbitrary $R_2 \in \mathbb{R}_+^*$ to be determined later. Then, we obtain

$$\int_{\mathbb{R}^{2d}} |\bar{x}|^2 \exp\left(-\frac{1}{2}\frac{\alpha}{1+\alpha}|\bar{x}|^2\right) \frac{F(dx_1)}{\|F\|_{\mathcal{M}(\mathbb{R}^d)}} \frac{F(dx_2)}{\|F\|_{\mathcal{M}(\mathbb{R}^d)}} = I_1 + I_2,$$

where each term reads

$$\begin{aligned}
I_1 &:= \int_{|\bar{x}|\leq R_2} |\bar{x}|^2 \exp\left(-\frac{1}{2}\frac{\alpha}{1+\alpha}|\bar{x}|^2\right) \frac{F(dx_1)}{\|F\|_{\mathcal{M}(\mathbb{R}^d)}} \frac{F(dx_2)}{\|F\|_{\mathcal{M}(\mathbb{R}^d)}}, \\
I_2 &:= \int_{|\bar{x}|>R_2} |\bar{x}|^2 \exp\left(-\frac{1}{2}\frac{\alpha}{1+\alpha}|\bar{x}|^2\right) \frac{F(dx_1)}{\|F\|_{\mathcal{M}(\mathbb{R}^d)}} \frac{F(dx_2)}{\|F\|_{\mathcal{M}(\mathbb{R}^d)}}.
\end{aligned}$$

On the one hand, the first term can be readily estimated by

$$I_1 \leq R_2^2 \int_{\mathbb{R}^{2d}} \exp\left(-\frac{1}{2}\frac{\alpha}{1+\alpha}|\bar{x}|^2\right) \frac{F(dx_1)}{\|F\|_{\mathcal{M}(\mathbb{R}^d)}} \frac{F(dx_2)}{\|F\|_{\mathcal{M}(\mathbb{R}^d)}}. \tag{4.15}$$

On the other hand, the second term can be estimated by

$$\begin{aligned}
I_2 &\leq \frac{2}{e\gamma}\frac{1+\alpha}{\alpha} \int_{|\bar{x}|>R_2} \exp\left(-\frac{1-\gamma}{2}\frac{\alpha}{1+\alpha}|\bar{x}|^2\right) \frac{F(dx_1)}{\|F\|_{\mathcal{M}(\mathbb{R}^d)}} \frac{F(dx_2)}{\|F\|_{\mathcal{M}(\mathbb{R}^d)}} \\
&\leq \frac{2}{e\gamma}\frac{1+\alpha}{\alpha} \exp\left(-\frac{1-\gamma}{2}\frac{\alpha}{1+\alpha}R_2^2\right),
\end{aligned} \tag{4.16}$$

for any $\gamma \in (0,1)$, where in the first line we have used that $e^{\theta s} \geq \theta e s$ with $s = |\bar{x}|^2$ and $\theta = \frac{\gamma}{2}\frac{\alpha}{1+\alpha}$, and in the second line we have exploited that $F/\|F\|_{\mathcal{M}(\mathbb{R}^d)}$ are probability measures, thus normalized.

• Step 3: Putting (4.14), (4.15) and (4.16) into (4.13) yields

$$\int_{\mathbb{R}^d} |x|^2 \mathcal{S}[F](x)\, dx \leq \frac{1}{1+\alpha} + \frac{R_2^2}{(1+\alpha)^2} + \frac{2}{e\gamma}\frac{1}{\alpha(1+\alpha)} \frac{\exp\left(\frac{1}{2}\frac{\alpha}{1+\alpha}R_1^2\right)\exp\left(-\frac{1-\gamma}{2}\frac{\alpha}{1+\alpha}R_2^2\right)}{1 - \frac{1}{2R_1^2}\int_{\mathbb{R}^d} |x|^2 \frac{F(dx)}{\|F\|_{\mathcal{M}(\mathbb{R}^d)}}},$$

for any $R_1, R_2 \in \mathbb{R}_+^*$ and $\gamma \in (0,1)$. We set the values

$$R_1^2 = \frac{1}{2(1-\delta)}\int_{\mathbb{R}^d} |x|^2 \frac{F(dx)}{\|F\|_{\mathcal{M}(\mathbb{R}^d)}}, \qquad R_2^2 = \frac{R_1^2}{1-\gamma},$$

with $\gamma, \delta \in (0,1)$. Then, we obtain

$$\int_{\mathbb{R}^d} |x|^2 \mathcal{S}[F](x)\, dx \leq \frac{1}{1+\alpha} + \frac{2}{e\gamma\delta}\frac{1}{\alpha(1+\alpha)} + \frac{1}{2(1-\delta)(1-\gamma)(1+\alpha)^2}\int_{\mathbb{R}^d} |x|^2 \frac{F(dx)}{\|F\|_{\mathcal{M}(\mathbb{R}^d)}},$$

thus ending the proof by the arbitrariness of $\gamma, \delta \in (0,1)$. □

By iteration, we then obtain the following contractivity of the variance of generic solutions of (1.1).

Corollary 4.6 (Propagation of quadratic moments)**.** *Assume that $\alpha \in \mathbb{R}_+^*$, set any parameter $\eta \in \left(\frac{1}{2(1+\alpha)^2}, 1\right)$ and consider the solution $\{F_n\}_{n\in\mathbb{N}}$ of (1.1) issued at a generic initial datum $F_0 \in \mathcal{M}_+(\mathbb{R}^d)$ with $\int_{\mathbb{R}^d} |x|^2 F_0(dx) < \infty$. Then, there exists $M = M(\alpha, \eta)$ such that*

$$\int_{\mathbb{R}^d} |x|^2 \frac{F_n(x)}{\|F_n\|_{L^1(\mathbb{R}^d)}}\, dx \leq \frac{M}{1-\eta} + \eta^n \int_{\mathbb{R}^d} |x|^2 \frac{F_0(dx)}{\|F_0\|_{\mathcal{M}(\mathbb{R}^d)}},$$

for any $n \in \mathbb{N}$.

Let us note that, on the one hand the factor η^n makes the dependence on the variance of the initial datum F_0 negligible as the amount of generations increases. On the other hand, the constant M does not depend on the chosen initial profile F_0. This is compatible with our ergodicity property in Theorem 1.1.

Lemma 4.7 (Control of exponential moments)**.** *Assume that $\alpha \in \mathbb{R}_+^*$, set $\theta \in \mathbb{R}_+^*$ so that $\theta < \frac{\alpha}{2}$ and any parameter $\chi > \frac{\alpha}{1+\alpha}\frac{\theta}{2(1+\alpha-2\theta)(\alpha-2\theta)}$. Then, there exists a constant $C = C(\alpha, \theta, \chi)$ such that*

$$\int_{\mathbb{R}^d} e^{\theta|x|^2} \mathcal{S}[F](x)\, dx \leq C \left\{ 1 + \exp\left(\chi \int_{\mathbb{R}^d} |x|^2 \frac{F(dx)}{\|F\|_{\mathcal{M}(\mathbb{R}^d)}} \right) \right\},$$

for any measure $F \in \mathcal{M}_+(\mathbb{R}^d)$.

Proof. Since G is a Gaussian function given by (1.4) and m is a quadratic function given by (1.5), then we can compute explicitly the integrals with respect to x in the numerator and denominator like formulas (4.11)-(4.12) in the proof of Lemma 4.5 and we obtain

$$\int_{\mathbb{R}^d} e^{\theta|x|^2} \mathcal{S}[F](x)\, dx = \left(\frac{1+\alpha}{1+\alpha-2\theta} \right)^{d/2} \frac{\displaystyle\int_{\mathbb{R}^{2d}} \exp\left(-\frac{1}{2}\frac{\alpha-2\theta}{1+\alpha-2\theta}|\bar{x}|^2 \right) \frac{F(dx_1)}{\|F\|_{\mathcal{M}(\mathbb{R}^d)}} \frac{F(dx_2)}{\|F\|_{\mathcal{M}(\mathbb{R}^d)}}}{\displaystyle\int_{\mathbb{R}^{2d}} \exp\left(-\frac{1}{2}\frac{\alpha}{1+\alpha}|\bar{x}|^2 \right) \frac{F(dx_1)}{\|F\|_{\mathcal{M}(\mathbb{R}^d)}} \frac{F(dx_2)}{\|F\|_{\mathcal{M}(\mathbb{R}^d)}}}, \quad (4.17)$$

where $\bar{x} = \frac{x_1+x_2}{2}$. Again, we provide bounds for the numerator and denominator in (4.17). Regarding the denominator, we recover estimate (4.14) for any $R_1 \in \mathbb{R}_+^*$. Thus, we focus on the bound of the numerator

$$\int_{\mathbb{R}^{2d}} \exp\left(-\frac{1}{2}\frac{\alpha-2\theta}{1+\alpha-2\theta}|\bar{x}|^2 \right) \frac{F(dx_1)}{\|F\|_{\mathcal{M}(\mathbb{R}^d)}} \frac{F(dx_2)}{\|F\|_{\mathcal{M}(\mathbb{R}^d)}} = I_1 + I_2,$$

where each term reads

$$I_1 := \int_{|\bar{x}|\leq R_2} \exp\left(-\frac{1}{2}\frac{\alpha-2\theta}{1+\alpha-2\theta}|\bar{x}|^2 \right) \frac{F(dx_1)}{\|F\|_{\mathcal{M}(\mathbb{R}^d)}} \frac{F(dx_2)}{\|F\|_{\mathcal{M}(\mathbb{R}^d)}},$$
$$I_2 := \int_{|\bar{x}|> R_2} \exp\left(-\frac{1}{2}\frac{\alpha-2\theta}{1+\alpha-2\theta}|\bar{x}|^2 \right) \frac{F(dx_1)}{\|F\|_{\mathcal{M}(\mathbb{R}^d)}} \frac{F(dx_2)}{\|F\|_{\mathcal{M}(\mathbb{R}^d)}}.$$

On the one hand, note that

$$I_1 \leq \exp\left(\frac{\theta R_2^2}{(1+\alpha-2\theta)^2} \right) \int_{\mathbb{R}^{2d}} \exp\left(-\frac{1}{2}\frac{\alpha}{1+\alpha}|\bar{x}|^2 \right) \frac{F(dx_1)}{\|F\|_{\mathcal{M}(\mathbb{R}^d)}} \frac{F(dx_2)}{\|F\|_{\mathcal{M}(\mathbb{R}^d)}}, \quad (4.18)$$

where we have used the mean value theorem applied to the function $r \in \mathbb{R}_+^* \mapsto \frac{r}{1+r}$ to derive the relation

$$\frac{\alpha}{1+\alpha} - \frac{\alpha-2\theta}{1+\alpha-2\theta} \leq \frac{2\theta}{(1+\alpha-2\theta)^2}.$$

On the other hand,

$$I_2 \leq \exp\left(-\frac{1}{2}\frac{\alpha-2\theta}{1+\alpha-2\theta} R_2^2 \right). \quad (4.19)$$

Putting (4.14), (4.18) and (4.19) into (4.17) we obtain

$$\int_{\mathbb{R}^d} e^{\theta|x|^2} \mathcal{S}[F](x)\, dx \leq \left(\frac{1+\alpha}{1+\alpha-2\theta} \right)^{d/2} \left\{ \exp\left(\frac{\theta R_2^2}{(1+\alpha-2\theta)^2} \right) + \frac{\exp\left(-\frac{1}{2}\frac{\alpha-2\theta}{1+\alpha-2\theta} R_2^2 \right) \exp\left(\frac{1}{2}\frac{\alpha}{1+\alpha} R_1^2 \right)}{1 - \frac{1}{2R_1^2}\int_{\mathbb{R}^d} |x|^2 \frac{F(dx)}{\|F\|_{\mathcal{M}(\mathbb{R}^d)}}} \right\},$$

for any $R_1, R_2 \in \mathbb{R}_+^*$. As in the proof of Lemma 4.5 we choose them as follows

$$R_1^2 = \frac{1}{2(1-\delta)} \int_{\mathbb{R}^d} |x|^2 \frac{F(dx)}{\|F\|_{\mathcal{M}(\mathbb{R}^d)}}, \qquad R_2^2 = \frac{\alpha}{1+\alpha}\frac{1+\alpha-2\theta}{\alpha-2\theta} R_1^2,$$

where $\delta \in (0,1)$. Then, we obtain

$$\int_{\mathbb{R}^d} e^{\theta|x|^2} \mathcal{S}[F](x)\, dx$$
$$\leq \left(\frac{1+\alpha}{1+\alpha-2\theta} \right)^{d/2} \left\{ \frac{1}{\delta} + \exp\left(\frac{\alpha}{1+\alpha}\frac{\theta}{(1+\alpha-2\theta)(\alpha-2\theta)}\frac{1}{2(1-\delta)} \int_{\mathbb{R}^d} |x|^2 \frac{F(dx)}{\|F\|_{\mathcal{M}(\mathbb{R}^d)}} \right) \right\},$$

and we end the proof by arbitrariness of $\delta \in (0,1)$. □

Therefore, small exponential moments can be controlled by quadratic moments. We can then couple Lemma 4.7 and Corollary 4.6 to obtain the propagation of exponential moments in the time-discrete problem (1.1).

Corollary 4.8 (Propagation of exponential moments)**.** *Assume that $\alpha \in \mathbb{R}_+^*$, set $\theta \in \mathbb{R}_+^*$ with $\theta < \frac{\alpha}{2}$ and any parameter $\eta \in \left(\frac{1}{2(1+\alpha)^2}, 1\right)$. Consider the solution $\{F_n\}_{n\in\mathbb{N}}$ of (1.1) issued at a generic initial datum $F_0 \in \mathcal{M}_+(\mathbb{R}^d)$ with $\int_{\mathbb{R}^d} |x|^2 F_0(dx) < \infty$. Then, there exist $C = C(\alpha,\theta,\eta)$ and $C' = C'(\alpha,\theta,\eta)$ such that*

$$\int_{\mathbb{R}^d} e^{\theta|x|^2} \frac{F_n(x)}{\|F_n\|_{L^1(\mathbb{R}^d)}}\, dx \leq C\left\{1 + \exp\left(C'\left(1 + \eta^{n-1}\int_{\mathbb{R}^d} |x|^2 \frac{F_0(dx)}{\|F_0\|_{\mathcal{M}(\mathbb{R}^d)}}\right)\right)\right\},$$

for any $n \in \mathbb{N}$.

5. Reformulating the recursion

In this section we find an appropriate reformulation of the recursion for the solutions $\{F_n\}_{n\in\mathbb{N}}$ of the time-discrete problem (1.1), as already discussed in Section 2.4. Note that by iterating $\mathcal{T}$ n-times we find that F_n depends on the initial datum $F_0 \in \mathcal{M}_+(\mathbb{R}^d)$ via a high-dimensional integral parametrized by variables indexed on the binary tree T_*^n (representing the traits of the ancestors in the pedigree chart). Specifically, for any initial datum $F_0 \in \mathcal{M}_+(\mathbb{R}^d)$ we obtain:

$\diamond$ *Iteration 1*:
By the explicit definition of $\mathcal{T}$ in (1.2) we readily obtain

$$F_1(x) = \frac{1}{\|F_0\|_{\mathcal{M}(\mathbb{R}^d)}} \int_{\mathbb{R}^{2d}} e^{-m(x)} G\left(x - \frac{x_1+x_2}{2}\right) F_0(x_1)F_0(x_2)\, d\mathbf{x}_1, \tag{5.1}$$

where we denote $\mathbf{x}_1 = (x_1, x_2) \in \mathbb{R}^{2d}$.

$\diamond$ *Iteration 2*:
By (5.1) with F_0 replaced by F_1 and using $F_1 = \mathcal{T}[F_0]$ in the right hand side we get

$$\begin{aligned} F_2(x) = \frac{1}{\|F_1\|_{L^1(\mathbb{R}^d)}\|F_0\|^2_{\mathcal{M}(\mathbb{R}^d)}} \int_{\mathbb{R}^{6d}} e^{-(m(x)+m(x_1)+m(x_2))} G\left(x - \frac{x_1+x_2}{2}\right) \\ \times G\left(x_1 - \frac{x_{11}+x_{12}}{2}\right) G\left(x_2 - \frac{x_{21}+x_{22}}{2}\right) F_0(x_{11})F_0(x_{12})F_0(x_{21})F_0(x_{22})\, d\mathbf{x}_2, \end{aligned} \tag{5.2}$$

where we denote $\mathbf{x}_2 = (x_1, x_2, x_{11}, x_{12}, x_{21}, x_{22}) \in \mathbb{R}^{6d}$.

$\diamond$ *Iteration $n \in \mathbb{N}$*:
By a clear inductive process, we can iterate the operator $\mathcal{T}$ as many times $n \in \mathbb{N}$ as needed to recover an explicit dependence of F_n on the initial datum F_0. Of course, this generates a high-dimensional integral involving $2(2^n - 1)$ variables that we can label along the vertices of the perfect binary tree T^n according to the *universal address system* notation in Section 3.1. Namely, we obtain

$$\begin{aligned} F_n(x) = \frac{1}{\prod_{m=0}^{n-1} \|F_m\|^{2^{n-1-m}}_{L^1(\mathbb{R}^d)}} \\ \times \int_{\mathbb{R}^{2(2^n-1)d}} \exp\left(-\sum_{i\in\widehat{\mathsf{T}}^n} m(x_i)\right) \prod_{i\in\widehat{\mathsf{T}}^n} G\left(x_i - \frac{x_{i1}+x_{i2}}{2}\right) \prod_{i\in\mathsf{L}^n} F_0(x_i)\, d\mathbf{x}_n, \end{aligned} \tag{5.3}$$

where we have used the tree-indexed notation $\mathbf{x}_n = (x_i)_{i\in\mathsf{T}_*^n} \in \mathbb{R}^{2(2^n-1)d}$ in Remark 3.1 and $x_\emptyset := x$.

The ultimate goal of this section is to find a suitable change of variables as discussed in Section 2.4. We shall see that the effect of such a change of variable on the above high-dimensional integral in (5.3) will reveal that actual dependence of $\{F_n\}_{n\in\mathbb{N}}$ on the shape of F_0. We will see that such a dependence is indeed weak and the initial datum is rapidly "forgotten" across the different levels of the tree T^n. This ergodicity property will be crucial in our analysis and will be exploited later in Section 6 to derive our main result in Theorem 1.1.

5.1. **A change of variables across the binary tree.** The goal of this section is to appropriately reformulate the high-dimensional integral (5.3) by exploiting the special form of the Gaussian mixing kernel G in (1.4) and the quadratic selection function m in (1.5). Namely, we shall derive a suitable change of variables according to a n-step backwards process starting at the leaf variables x_j with $j \in \mathsf{L}^n$ and ending at the root variable $x_\emptyset = x$. More specifically, at each level m in the tree we shall change the reference frame of the variables x_i indexed with $i \in \mathsf{L}^n_m$ so that the quadratic form involving those variables in the exponential of (5.3) gets appropriately centered at its minimum. By doing so we find that the minimum is located at a contracted value of the variable $x_{\mathsf{c}(i)} \in L^n_{m-1}$ at the level $m-1$ below, which is indexed by its children $\mathsf{c}(i)$ in the binary tree (recall notation in Section 3.1). By appropriately propagating the backwards process across the tree, and by tracking the accumulated contraction of variables, we shall see that the eventual dependence on the root variable x is weakly gradually forgotten as $n \to \infty$. This suggests an apparent form of ergodicity, which we will be crucial later in Section 6.

First, for the sake of clarity, we illustrate our method in the simpler case $n = 2$. Later, we address the general case with $n > 2$ driven by the same strategy. As proposed in Section 2.4, we define the following choice for the rescaled profiles that will be used along the change of variables.

Definition 5.1 (Rescaled distributions)**.** *Consider any $F \in \mathcal{M}_+(\mathbb{R}^d)$. Then, we define*

$$\bar{F}(dx) := e^{m(x)} \frac{F(dx)}{\boldsymbol{F}_{\alpha=0}(x)}, \quad x \in \mathbb{R}^d.$$

Here, we remind that $\boldsymbol{F}_{\alpha=0} = G_{0,2}$ is the Gaussian eigenfunction corresponding to (1.12) with $\alpha = 0$. We also recall that there is some freedom in the normalization, and in particular the term e^m is not mandatory but it is convenient for an easier sorting of the various terms, as we anticipated in Section 2.4. By substituting (1.4) and (1.5) into (5.1) and writting the integral in the right hand side in terms of the rescaled function $\bar{F}_0$ in Definition 5.1, we obtain that

$$\begin{aligned}
(e^m F_2)(x) = {} & \frac{1}{(2\pi)^{3d/2}} \frac{1}{(4\pi)^{2d}} \frac{1}{\|F_1\|_{L^1(\mathbb{R}^d)} \|F_0\|^2_{\mathcal{M}(\mathbb{R}^d)}} \\
\times \int_{\mathbb{R}^{6d}} & \exp\left(-\frac{\alpha}{2}\big(|x_1|^2 + |x_2|^2\big) - \frac{1}{2}\left| x - \frac{x_1 + x_2}{2} \right|^2 \right) \\
& \times \exp\left(-\frac{\alpha_0}{2}\big(|x_{11}|^2 + |x_{12}|^2 + |x_{21}|^2 + |x_{22}|^2\big) - \frac{1}{2}\left| x_1 - \frac{x_{11} + x_{12}}{2} \right|^2 - \frac{1}{2}\left| x_2 - \frac{x_{21} + x_{22}}{2} \right|^2 \right) \\
& \times \bar{F}_0(x_{11}) \bar{F}_0(x_{12}) \bar{F}_0(x_{21}) \bar{F}_0(x_{22})\, d\mathbf{x}_2.
\end{aligned} \tag{5.4}$$

Here, the parameter $\alpha_0 \in \mathbb{R}^*_+$ has been defined by

$$\alpha_0 := \frac{1}{2} + \alpha,$$

so that the quadratic terms in the fist term of the third line of (5.4) correct the rescaling $\bar{F}_0$ of F_0. The new formulation will be obtained by an appropriate change of variables, so that the quadratic forms inside the above exponential get appropriately centered. To do so, notice that this formula involves as many variables x_i as indices i in the perfect binary tree T^2. Indeed, we have sorted the different terms in (5.4) in such a way that the second line only involves variables x_1 and x_2 in the first level L^2_1 of the tree, whilst the third line contains the terms involving the variables x_{11}, x_{12}, x_{21} and x_{22} in the second level of the tree L^2_2 (*i.e.*, the leaves L^2). We will divide the method into two steps. First, we will address the change of the variables indexed by the leaves L^2. Second, we will perform the change of variables indexed by the first level of the tree L^2_1.

• Step 1: Change of variables for x_{11}, x_{12}, x_{21} and x_{22}.
Consider some coefficient $k_1 > 0$ to be determined later and define the change of variables $x_{11} \to y_{11}$, $x_{12} \to y_{12}$, $x_{21} \to y_{21}$ and $x_{22} \to y_{22}$ given by

$$\begin{aligned}
x_{11} &= k_1 x_1 + y_{11}, & x_{12} &= k_1 x_1 + y_{12}, \\
x_{21} &= k_1 x_2 + y_{21}, & x_{22} &= k_1 x_2 + y_{12}.
\end{aligned}$$

On the one hand, using the change of variables in the terms inside the exponential of (5.4) which involve x_{11} and x_{12}, we obtain

$$\begin{aligned}\frac{\alpha_0}{2}|x_{11}|^2+\frac{\alpha_0}{2}|x_{12}|^2+\frac{1}{2}\left|x_1-\frac{x_{11}+x_{12}}{2}\right|^2&\\ =\frac{\alpha_0}{2}|k_1x_1+y_{11}|^2+\frac{\alpha_0}{2}|k_1x_1+y_{12}|^2+\frac{1}{2}\left|(1-k_1)x_1-\frac{y_{11}+y_{12}}{2}\right|^2&\\ =\frac{1}{2}\left(2\alpha_0k_1^2+(1-k_1)^2\right)|x_1|^2+\frac{\alpha_0}{2}|y_{11}|^2+\frac{\alpha_0}{2}|y_{12}|^2+\frac{1}{8}|y_{11}+y_{12}|^2&\\ +\left(\alpha_0k_1-\frac{1-k_1}{2}\right)x_1\cdot y_{11}+\left(\alpha_0k_1-\frac{1-k_1}{2}\right)x_1\cdot y_{12}.&\end{aligned}$$

On the other hand, using the change of variables in the terms which involve x_{21} and x_{22}, we get

$$\begin{aligned}\frac{\alpha_0}{2}|x_{21}|^2+\frac{\alpha_0}{2}|x_{22}|^2+\frac{1}{2}\left|x_2-\frac{x_{21}+x_{22}}{2}\right|^2&\\ =\frac{\alpha_0}{2}|k_1x_2+y_{21}|^2+\frac{\alpha_0}{2}|k_1x_2+y_{22}|^2+\frac{1}{2}\left|(1-k_1)x_2-\frac{y_{21}+y_{22}}{2}\right|^2&\\ =\frac{1}{2}\left(2\alpha_0k_1^2+(1-k_1)^2\right)|x_2|^2+\frac{\alpha_0}{2}|y_{21}|^2+\frac{\alpha_0}{2}|y_{22}|^2+\frac{1}{8}|y_{21}+y_{22}|^2&\\ +\left(\alpha_0k_1-\frac{1-k_1}{2}\right)x_2\cdot y_{21}+\left(\alpha_0k_1-\frac{1-k_1}{2}\right)x_2\cdot y_{22}.&\end{aligned}$$

Notice that one can eliminate the crossed terms in both expressions by choosing

$$k_1:=\frac{1}{1+2\alpha_0}=\frac{1}{2(1+\alpha)}.$$

In that case, adding both terms we obtain that

$$\begin{aligned}\frac{\alpha_0}{2}|x_{11}|^2+\frac{\alpha_0}{2}|x_{12}|^2+\frac{1}{2}\left|x_1-\frac{x_{11}+x_{12}}{2}\right|^2+\frac{\alpha_0}{2}|x_{21}|^2+\frac{\alpha_0}{2}|x_{22}|^2+\frac{1}{2}\left|x_2-\frac{x_{21}+x_{22}}{2}\right|^2&\\ =\frac{1}{2}(1-k_1)|x_1|^2+\frac{1}{4k_1}|y_{11}|^2+\frac{1}{4k_1}|y_{12}|^2-\frac{1}{8}|y_{11}-y_{12}|^2&\\ +\frac{1}{2}(1-k_1)|x_2|^2+\frac{1}{4k_1}|y_{21}|^2+\frac{1}{4k_1}|y_{22}|^2-\frac{1}{8}|y_{21}-y_{22}|^2.&\end{aligned}$$

Putting everything together into (5.4) yields

$$\begin{aligned}(e^mF_2)(x)=\frac{1}{(2\pi)^{3d/2}}\frac{1}{(4\pi)^{2d}}\frac{1}{\|F_1\|_{L^1(\mathbb{R}^d)}\|F_0\|^2_{\mathcal{M}(\mathbb{R}^d)}}&\\ \times\int_{\mathbb{R}^{6d}}\exp\left(-\frac{\alpha_1}{2}\left(|x_1|^2+|x_2|^2\right)-\frac{1}{2}\left|x-\frac{x_1+x_2}{2}\right|^2\right)&\\ \times\exp\left(-\frac{1}{4k_1}\left(|y_{11}|^2+|y_{12}|^2+|y_{21}|^2+|y_{22}|^2\right)+\frac{1}{8}\left(|y_{11}-y_{12}|^2+|y_{21}-y_{22}|^2\right)\right)&\\ \times\bar{F}_0(k_1x_1+y_{11})\bar{F}_0(k_1x_1+y_{12})\bar{F}_0(k_1x_2+y_{21})\bar{F}_0(k_1x_2+y_{22})\,dx_1\,dx_2\,dy_{11}\,dy_{12}\,dy_{21}\,dy_{22},&\end{aligned}\tag{5.5}$$

where the parameter $\alpha_1\in\mathbb{R}_+^*$ has been defined by

$$\alpha_1:=1-k_1+\alpha,$$

in order to recombine the initial terms in the second line of (5.4) involving variables indexed by the first level L_1^2 with the new remainders that have appeared from the previous step.

• Step 2: Change of variables for x_1 and x_2.
Consider some coefficient $k_2>0$ to be determined later and define the change of variables $x_1\to y_1$ and $x_2\to y_2$ given by

$$x_1=k_2x+y_1,\quad x_2=k_2x+y_2.$$

This time, using the change of variables in the terms of the exponential in the second line of (5.5) yields

$$\begin{aligned}\frac{\alpha_1}{2}|x_1|^2+\frac{\alpha_1}{2}|x_2|^2+\frac{1}{2}\left|x-\frac{x_1+x_2}{2}\right|^2&\\ =\frac{\alpha_1}{2}|k_2x+y_1|^2+\frac{\alpha_1}{2}|k_2x+y_2|^2+\frac{1}{2}\left|(1-k_2)x-\frac{y_1+y_2}{2}\right|^2&\\ =\frac{1}{2}\left(2\alpha_1k_2^2+(1-k_2)^2\right)|x|^2+\frac{\alpha_1}{2}|y_1|^2+\frac{\alpha_1}{2}|y_2|^2+\frac{1}{8}|y_1+y_2|^2&\\ +\left(\alpha_1k_2-\frac{1-k_2}{2}\right)x\cdot y_1+\left(\alpha_1k_2-\frac{1-k_2}{2}\right)x\cdot y_2.&\end{aligned}$$

Again, we can cancel the crossed term by choosing

$$k_2:=\frac{1}{1+2\alpha_1}=\frac{1}{3+2\alpha-2k_1}.$$

Namely, we obtain that

$$\begin{aligned}\frac{\alpha_1}{2}|x_1|^2+\frac{\alpha_1}{2}|x_2|^2+\frac{1}{2}\left|x-\frac{x_1+x_2}{2}\right|^2&\\ =\frac{1}{2}(1-k_2)|x|^2+\frac{1}{4k_2}|y_1|^2+\frac{1}{4k_2}|y_2|^2-\frac{1}{8}|y_1-y_2|^2.&\end{aligned}$$

Then, putting everything together into (5.5) yields

$$\begin{aligned}F_2(x)=&\frac{1}{(2\pi)^{3d/2}}\frac{1}{(4\pi)^{2d}}\frac{1}{\|F_1\|_{L^1(\mathbb{R}^d)}\|F_0\|^2_{\mathcal{M}(\mathbb{R}^d)}}e^{-\frac{1+\alpha-k_2}{2}|x|^2}\\ &\times\int_{\mathbb{R}^{6d}}\exp\left(-\frac{1}{4k_2}\big(|y_1|^2+|y_2|^2\big)+\frac{1}{8}|y_1-y_2|^2\right)\\ &\quad\times\exp\left(-\frac{1}{4k_1}\big(|y_{11}|^2+|y_{12}|^2+|y_{21}|^2+|y_{22}|^2\big)+\frac{1}{8}\big(|y_{11}-y_{12}|^2+|y_{21}-y_{22}|^2\big)\right)\\ &\quad\times\bar F_0(k_1k_2x+k_1y_1+y_{11})\bar F_0(k_1k_2x+k_1y_1+y_{12})\\ &\quad\times\bar F_0(k_1k_2x+k_1y_2+y_{21})\bar F_0(k_1k_2x+k_1y_2+y_{22})\,d\mathbf{y}_2,\end{aligned}\tag{5.6}$$

where we denote again $\mathbf{y}_2=(y_1,y_2,y_{11},y_{12},y_{21},y_{22})\in\mathbb{R}^{6d}$.

Before stating the main result for general $n\in\mathbb{N}$, we collect some natural notation according to the preceding computations, that will be useful here on. First, we define the following sequences of coefficients.

Definition 5.2 (Coefficients)**.** *Consider any $\alpha\in\mathbb{R}_+$.*

(1) The coefficients $\{k_n\}_{n\in\mathbb{N}}$ are defined by the recursive formula

$$\begin{aligned}k_1&:=\frac{1}{2(1+\alpha)},\\ k_n&:=\frac{1}{3+2\alpha-2k_{n-1}},\quad \text{for } n\geq 2.\end{aligned}\tag{5.7}$$

(2) The coefficients $\{\kappa_n\}_{n\in\mathbb{N}}$ are defined by the recursive formula

$$\begin{aligned}\kappa_0&:=1,\\ \kappa_n&:=k_1\cdots k_n,\quad \text{for } n\geq 1.\end{aligned}\tag{5.8}$$

We note that for $n=2$ the above coefficients k_1,k_2 reduce to those appearing in the previous reformulation (5.6) of the recursion. Since it will be used later, we study the asymptotic behavior of such sequences of coefficients as $n\to\infty$.

Lemma 5.3 (Asymptotic behavior of the coefficients)**.** *Consider the sequences $\{k_n\}_{n\in\mathbb{N}}$ and $\{\kappa_n\}_{n\in\mathbb{N}}$ in Definition 5.2. Then, the following properties hold true:*

(i) **(Coefficients k_n)** *The coefficients $\{k_n\}_{n\in\mathbb{N}}$ are positive numbers and $k_n\searrow \boldsymbol{k}_\alpha$ as $n\to\infty$, where $\boldsymbol{k}_\alpha\in\mathbb{R}_+^*$ is the smallest root of the equation*

$$\frac{1}{3+2\alpha-2\,\boldsymbol{k}_\alpha}=\boldsymbol{k}_\alpha\,.\tag{5.9}$$

Specifically, $\boldsymbol{k}_\alpha$ *is given explicitly by formula* (1.15) *and it is related to the coefficient* $\boldsymbol{r}_\alpha$ *in formula* (3.8) *of Lemma 3.8 by* $\boldsymbol{r}_\alpha = 2\,\boldsymbol{k}_\alpha^2$. *In addition,*

$$k_n - \boldsymbol{k}_\alpha \leq C\,\boldsymbol{r}_\alpha^{n-1}, \tag{5.10}$$

for every $n \in \mathbb{N}$, *where* $\boldsymbol{r}_\alpha$ *is and* $C \in \mathbb{R}_+^*$ *depends only on* α.

(ii) **(Coefficients** κ_n**)** *The coefficients* $\{\kappa_n\}_{n\in\mathbb{N}}$ *are positive numbers and decay geometrically with maximal rate* $\boldsymbol{k}_\alpha$. *Specifically, for every* $\varepsilon \in \mathbb{R}_+^*$ *there exists* $C_\varepsilon \in \mathbb{R}_+^*$ *such that*

$$\kappa_n \leq C_\varepsilon(\boldsymbol{k}_\alpha + \varepsilon)^n, \tag{5.11}$$

for every $n \in \mathbb{N}$.

Proof. On the one hand, the properties of $\{\kappa_n\}_{n\in\mathbb{N}}$ readily follow from those of $\{k_n\}_{n\in\mathbb{N}}$ and the relation (5.8) in Definition 5.2. Specifically, for any given $q \in \mathbb{N}$ and any $n \geq q$ we get the decomposition

$$\kappa_n = \left(\prod_{m=0}^{q} k_m\right)\left(\prod_{m=q+1}^{n} k_m\right) \leq k_1^q k_{q+1}^{n-q} = \frac{k_1^q}{k_{q+1}^q} k_{q+1}^n \leq \frac{k_1^q}{k_{q+1}^q}(C\,\boldsymbol{r}_\alpha^{q+1} + \boldsymbol{k}_\alpha)^n.$$

Therefore, taking q sufficiently large so that $C\,\boldsymbol{r}_\alpha^{q+1} \leq \varepsilon$ we conclude (5.11). We then focus on the properties of $\{k_n\}_{n\in\mathbb{N}}$ and we use a similar strategy like in Lemma 3.8.

• Step 1: Monotone convergence of $\{k_n\}_{n\in\mathbb{N}}$.
Define the following function

$$f(x) := \frac{1}{3 + 2\alpha - 2x}, \quad x \in \big(-\infty, \frac{3}{2} + \alpha\big).$$

Then, by (5.7) $\{k_n\}_{n\in\mathbb{N}}$ obeys the following recursive relation

$$k_n = f(k_{n-1}), \quad n > 1.$$

Let us consider the fixed points $x_- < x_+$ of f, i.e,

$$x_\pm = \frac{(3 + 2\alpha) \pm \sqrt{(1 + 2\alpha)^2 + 8\alpha}}{4}.$$

By inspection, it is clear that $0 < \boldsymbol{k}_\alpha = x_- < x_+ < \frac{3}{2} + \alpha$ and

$$x_- < f(x) < x, \quad \text{if } x \in (x_-, x_+),$$

where $\boldsymbol{k}_\alpha$ is given in (1.15). Since $k_1 \in (x_-, x_+)$, then we conclude that $\{k_n\}_{n\in\mathbb{N}}$ is a well defined, positive and monotonically decreasing sequence contained in the compact interval $[x_-, x_+]$. Therefore, it must converge towards some limit ℓ, which is a solution of $f(x) = x$ *i.e.*, $\ell \in \{x_-, x_+\}$. Since the full sequence $\{k_n\}_{n\in\mathbb{N}}$ is below x_+ and decreasing, we then conclude that $\ell = x_- = \boldsymbol{k}_\alpha$. In particular, we obtain

$$\boldsymbol{k}_\alpha < k_n \leq k_1, \tag{5.12}$$

for any $n \in \mathbb{N}$.

• Step 2: Convergence rates.
As for Lemma 3.8, we shall use the special algebraic structure of the function f, which is the restriction to $\big(-\infty, \frac{3}{2} + \alpha\big)$ of the Möbius function $M : \mathbb{R} \setminus \{\frac{3}{2} + \alpha\} \longrightarrow \mathbb{R}$ given by

$$M(x) := \frac{1}{3 + 2\alpha - 2x}, \quad x \in \mathbb{R} \setminus \left\{\frac{3}{2} + \alpha\right\}.$$

By definition of $\{k_n\}_{n\in\mathbb{N}}$ in (5.7) we obtain

$$k_n - x_+ = \frac{1}{3 + 2\alpha - 2k_{n-1}} - \frac{1}{3 + 2\alpha - 2x_+} = \frac{2}{(3 + 2\alpha - 2k_{n-1})(3 + 2\alpha - 2x_+)}(k_{n-1} - x_+).$$

Similarly, we obtain

$$k_n - x_- = \frac{1}{3 + 2\alpha - 2k_{n-1}} - \frac{1}{3 + 2\alpha - 2x_-} = \frac{2}{(3 + 2\alpha - 2k_{n-1})(3 + 2\alpha - 2x_-)}(k_{n-1} - x_-).$$

By dividing both expressions and iterating such an identity we get

$$\frac{k_n - x_-}{k_n - x_+} = \left(\frac{3 + 2\alpha - 2x_+}{3 + 2\alpha - 2x_-}\right)^{n-1} \frac{k_1 - x_-}{k_1 - x_+}, \tag{5.13}$$

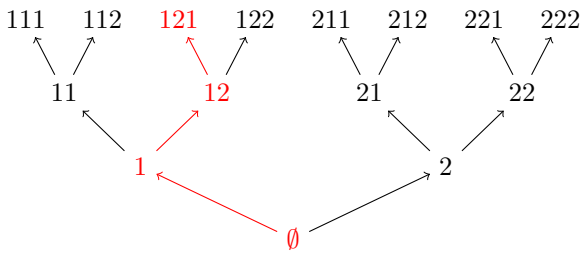


FIGURE 7. Unique path joining the leaf $j = 121$ to the root in the perfect binary tree T^3. The corresponding lineage map takes the form $\Phi_n^{121}(x; \mathbf{y}_3) = \kappa_3 x + \kappa_2 y_1 + \kappa_1 y_{12} + y_{121}$.

for every $n \in \mathbb{N}$. In fact, the basis can be restated as follows

$$\frac{3+2\alpha-2x_+}{3+2\alpha-2x_-} = \frac{3+2\alpha-\sqrt{(1+2\alpha)^2+8\alpha}}{3+2\alpha+\sqrt{(1+2\alpha)^2+8\alpha}} = \frac{8}{\left(3+2\alpha+\sqrt{(1+2\alpha)^2+8\alpha}\right)^2} \equiv \boldsymbol{r}_\alpha < 1,$$

where $\boldsymbol{r}_\alpha$ is determined by (3.8) in Lemma 3.8. Therefore, using (5.12) and (5.13) we conclude that

$$k_n - \boldsymbol{k}_\alpha = \frac{k_1 - x_-}{x_+ - k_1}(x_+ - k_n)\, \boldsymbol{r}_\alpha^{n-1} \leq \frac{k_1 - x_-}{x_+ - k_1}(x_+ - x_-)\, \boldsymbol{r}_\alpha^{n-1},$$

for any $n \in \mathbb{N}$, thus ending the proof. □

Definition 5.4 (Quadratic forms)**.** *Consider the sequence $\{k_n\}_{n\in\mathbb{N}}$ in Definition 5.2. Then, we define the quadratic form $Q_n = Q_n(\mathbf{y}_n)$ by*

$$Q_n(\mathbf{y}_n) := \sum_{m=0}^{n-1} \sum_{i \in \mathsf{L}_m^n} \left(\frac{1}{4k_{n-m}} (|y_{i1}|^2 + |y_{i2}|^2) - \frac{1}{8} |y_{i1} - y_{i2}|^2 \right), \quad \mathbf{y}_n \in \mathbb{R}^{2(2^n-1)d}, \tag{5.14}$$

where we are using the tree-indexed notation $\mathbf{y}_n = (y_i)_{i \in \mathsf{T}_*^n} \in \mathbb{R}^{2(2^n-1)d}$ in Remark 3.1.

Again, note that when $n = 2$ the above quadratic form $Q_2(\mathbf{y}_2)$ reduces to the one inside the exponential of (5.6). We now show the following uniform control of the quadratic forms Q_n as $n \to \infty$.

Lemma 5.5 (Uniform positive definite quadratic forms)**.** *Consider the quadratic form $Q_n = Q_n(\mathbf{y}_n)$ in Definition 5.4 and define the couple of coefficients $\beta_{\min}, \beta_{\max} \in \mathbb{R}_+^*$ by*

$$\beta_{\min} := \frac{1+2\alpha}{4}, \quad \beta_{\max} := \frac{1}{4\,\boldsymbol{k}_\alpha},$$

where $\boldsymbol{k}_\alpha$ is given by formula (1.15). Then, we obtain that

$$\beta_{\min} \| \mathbf{y}_n \|^2 \leq Q_n(\mathbf{y}_n) \leq \beta_{\max} \| \mathbf{y}_n \|^2,$$

where we are using the tree-indexed notation $\mathbf{y}_n = (y_i)_{i \in \mathsf{T}_*^n} \in \mathbb{R}^{2(2^n-1)d}$ in Remark 3.1.

Proof. By virtue of Young's inequality, we obtain the following lower and upper bound for Q_n

$$\sum_{m=0}^{n-1} \sum_{i \in \mathsf{L}_m^n} \frac{1}{4} \left(\frac{1}{k_{n-m}} - 1 \right) (|y_{i1}|^2 + |y_{i2}|^2) \leq Q_n(\mathbf{y}_n) \leq \sum_{m=0}^{n-1} \sum_{i \in \mathsf{L}_m^n} \frac{1}{4} \frac{1}{k_{n-m}} (|y_{i1}|^2 + |y_{i2}|^2),$$

for every $\mathbf{y}_n = (y_i)_{i \in \mathsf{T}_*^n} \in \mathbb{R}^{2(2^n-1)d}$. Then, the result follows from Lemma 5.3, which guarantees the uniform control $\boldsymbol{k}_\alpha < k_{m-n} \leq k_1$ for every $m = 0, \dots, n-1$. □

Definition 5.6 (Lineage maps)**.** *We define the lineage maps $\Phi_n^j = \Phi_n^j(x; \mathbf{y}_n)$ associated to the leave $j \in \mathsf{L}^n$ (see Figure 7) as follows*

$$\Phi_n^j(x; \mathbf{y}_n) := \kappa_n x + \sum_{m=0}^{n-1} \kappa_m y_{\mathsf{c}^m(j)}, \quad \mathbf{y}_n \in \mathbb{R}^{2(2^n-1)d}, \tag{5.15}$$

where $x \in \mathbb{R}^d$ represents the root value, $\mathbf{y}_n = (y_i)_{i \in \mathsf{T}_*^n} \in \mathbb{R}^{2(2^n-1)d}$ is represented according to the tree-indexed notation in Remark 3.1, $\{\kappa_n\}_{n\in\mathbb{N}}$ is given in Definition 5.2, and $\mathsf{c}^m = \mathsf{c} \circ \cdots \circ \mathsf{c}$ is the m times iterated map $\mathsf{c} : \mathsf{T}_*^n \to \widehat{\mathsf{T}}^n$ which, to any vertex $i \in \mathsf{T}_*^n$, it associates its child $\mathsf{c}(i) \in \widehat{\mathsf{T}}^n$ (cf. Section 3.1).

We are now ready to state the main result of this section extending formula (5.6) to any $n \in \mathbb{N}$. The starting point is again formula (5.3). In the particular case of Gaussian mixing kernel (1.4) and quadratic selection function (1.5), it takes the explicit form

$$\begin{aligned}(e^m F_n)(x) = &\frac{1}{(2\pi)^{(2^n-1)d/2}} \frac{1}{(4\pi)^{2^{n-1}d}} \frac{1}{\prod_{m=0}^{n-1} \|F_m\|_{L^1(\mathbb{R}^d)}^{2^{n-1-m}}} \\ &\times \int_{\mathbb{R}^{2(2^n-1)d}} \exp\left(- \sum_{m=0}^{n-2} \sum_{i \in \mathsf{L}_m^n} \left[\frac{\alpha}{2}(|x_{i1}|^2 + |x_{i2}|^2) + \frac{1}{2} \left| x_i - \frac{x_{i1}+x_{i2}}{2} \right|^2 \right] \right) \\ &\qquad \times \exp\left(- \sum_{i \in \mathsf{L}_{n-1}^n} \left[\frac{1}{2}(\alpha + \frac{1}{2})(|x_{i1}|^2 + |x_{i2}|^2) + \frac{1}{2} \left| x_i - \frac{x_{i1}+x_{i2}}{2} \right|^2 \right] \right) \\ &\qquad \times \prod_{j \in \mathsf{L}^n} \bar{F}_0(x_j)\, d\mathbf{x}_n,\end{aligned} \tag{5.16}$$

for any $x \in \mathbb{R}^d$ and $n \in \mathbb{N}$. Above, we have used the tree-indexed notation $\mathbf{x}_n = (x_i)_{i \in \mathsf{T}_*^n} \in \mathbb{R}^{2(2^n-1)d}$ in Remark 3.1, we have set $x_\emptyset := x$ and we have considered the rescaled distribution $\bar{F}_0$ according to Definition 5.1. Again, notice that the exponential terms in the fist term of the third line of (5.16) have been introduced to appropriately correct the rescaled distribution $\bar{F}_0$ of F_0. As a consequence, the factor $(4\pi)^{-2^{n-1}d}$ arises from the normalization by $\boldsymbol{F}_{\alpha=0}(x_j)$ whilst the factor $(2\pi)^{-(2^n-1)d/2}$ comes from the repeated products of the Gaussian mixing kernel G. The main result then reads as follows

Proposition 5.7 (Reformulation of the iterations I)**.** *Assume that $\alpha \in \mathbb{R}_+$ and set any initial datum $F_0 \in \mathcal{M}_+(\mathbb{R}^d)$. Hence, the solution $\{F_n\}_{n\in\mathbb{N}}$ to the time-discrete problem* (1.1) *admits the following form*

$$F_n(x) = \frac{e^{-\frac{1+\alpha-k_n}{2}|x|^2}}{(4\pi)^{2^{n-1}d}(2\pi)^{(2^n-1)d/2}} \frac{1}{\prod_{m=0}^{n-1} \|F_m\|_{L^1(\mathbb{R}^d)}^{2^{n-1-m}}} \int_{\mathbb{R}^{2(2^n-1)d}} e^{-Q_n(\mathbf{y}_n)} \prod_{j \in \mathsf{L}^n} \bar{F}_0(\Phi_n^j(x;\mathbf{y}_n))\, d\mathbf{y}_n, \tag{5.17}$$

for every $x \in \mathbb{R}^d$ and $n \in \mathbb{N}$, where we are using the tree-indexed notation $\mathbf{y}_n = (y_i)_{i \in \mathsf{T}_*^n} \in \mathbb{R}^{2(2^n-1)d}$ in Remark 3.1 along with further notation from Definitions 5.1, 5.2, 5.4 and 5.6.

As we anticipated for the particular case $n = 2$ at the beginning of this section, the proof of Proposition 5.7 will be based on the following change of variables

$$x_{i1} = k_{n-m} x_i + y_{i1}, \quad x_{i2} = k_{n-m} x_i + y_{i2},$$

for every $i \in \mathsf{L}_m^n$, which we will apply in a backwards way starting at indices $i \in \mathsf{L}_{n-1}^n$ and ending at $i = \emptyset$. Again, the objective of such a change of variables is to appropriately move the reference frame so that the above quadratic forms $\frac{1}{2}\left| x_i - \frac{x_{i1}+x_{i2}}{2} \right|^2$ inside the exponential of (5.16) get appropriately centered and no crossed terms remain when the selection parts $\frac{\alpha}{2}(|x_{i1}|^2 + |x_{i2}|^2)$ are taken into account. For simplicity, along the proof we shall restrict to tracking how the various terms inside the exponentials in (5.16) get modified after the change of variables.

Proof of Proposition 5.7.

• Step 1: Change of variables $x_{i1} \to y_{i1}$ and $x_{i2} \to y_{i2}$ for $i \in \mathsf{L}_{n-1}^n$.
We consider the change of variables given by

$$x_{i1} = k_1 x_i + y_{i1}, \quad x_{i2} = k_1 x_i + y_{i2},$$

where k_1 is given by Definition 5.2. The collection of all the terms for $i \in \mathsf{L}^n_{n-1}$ in (5.16) then read

$$\begin{aligned}
&\frac{\alpha_0}{2}|x_{i1}|^2 + \frac{\alpha_0}{2}|x_{i2}|^2 + \frac{1}{2}\left|x_i - \frac{x_{i1}+x_{i2}}{2}\right|^2 \\
&\qquad = \frac{\alpha_0}{2}|k_1 x_i + y_{i1}|^2 + \frac{\alpha_0}{2}|k_1 x_i + y_{i2}|^2 + \frac{1}{2}\left|(1-k_1)x_i - \frac{y_{i1}+y_{i2}}{2}\right|^2 \\
&\qquad = \frac{1}{2}\left(2\alpha_0 k_1^2 + (1-k_1)^2\right)|x_i|^2 + \frac{\alpha_0}{2}|y_{i1}|^2 + \frac{\alpha_0}{2}|y_{i2}|^2 + \frac{1}{8}|y_{i1}+y_{i2}|^2 \\
&\qquad\qquad + \left(\alpha_0 k_1 - \frac{1-k_1}{2}\right) x_i \cdot y_{i1} + \left(\alpha_0 k_1 - \frac{1-k_1}{2}\right) x_i \cdot y_{i2} \\
&\qquad = \frac{1}{2}\left(2\alpha_0 k_1^2 + (1-k_1)^2\right)|x_i|^2 + \frac{\alpha_0}{2}|y_{i1}|^2 + \frac{\alpha_0}{2}|y_{i2}|^2 + \frac{1}{8}|y_{i1}+y_{i2}|^2 \\
&\qquad = \frac{1}{2}(1-k_1)|x_i|^2 + \frac{1}{4k_1}|y_{i1}|^2 + \frac{1}{4k_1}|y_{i2}|^2 - \frac{1}{8}|y_{i1}-y_{i2}|^2,
\end{aligned}$$

where we have defined the coefficient $\alpha_0 \in \mathbb{R}_+^*$ by

$$\alpha_0 := \alpha + \frac{1}{2},$$

to recombine terms in the second and third lines of (5.16), and we have used the following relation between k_1 and α_0 in order to cancel the crossed terms, *i.e.*,

$$k_1 = \frac{1}{1+2\alpha_0}.$$

• Step 2: Change of variables $x_{i1} \to y_{i1}$ and $x_{i2} \to y_{i2}$ for $i \in \mathsf{L}^n_{n-2}$.
Again, we consider the change of variables given by

$$x_{i1} = k_2 x_i + y_{i1}, \quad x_{i2} = k_2 x_i + y_{i2},$$

where k_2 is given by Definition 5.2. Putting together the terms for $i \in \mathsf{L}^n_{n-2}$ in (5.16) and the above x_i-dependent remainder in the above expression in Step 1 yields the following term under the above change of variables

$$\begin{aligned}
&\frac{\alpha_1}{2}|x_{i1}|^2 + \frac{\alpha_1}{2}|x_{i2}|^2 + \frac{1}{2}\left|x_i - \frac{x_{i1}+x_{i2}}{2}\right|^2 \\
&\qquad = \frac{\alpha_1}{2}|k_2 x_i + y_{i1}|^2 + \frac{\alpha_1}{2}|k_2 x_i + y_{i2}|^2 + \frac{1}{2}\left|(1-k_2)x_i - \frac{y_{i1}+y_{i2}}{2}\right|^2 \\
&\qquad = \frac{1}{2}\left(2\alpha_1 k_2^2 + (1-k_2)^2\right)|x_i|^2 + \frac{\alpha_1}{2}|y_{i1}|^2 + \frac{\alpha_1}{2}|y_{i2}|^2 + \frac{1}{8}|y_{i1}+y_{i2}|^2 \\
&\qquad\qquad + \left(\alpha_1 k_2 - \frac{1-k_2}{2}\right) x_i \cdot y_{i1} + \left(\alpha_1 k_2 - \frac{1-k_2}{2}\right) x_i \cdot y_{i2} \\
&\qquad = \frac{1}{2}\left(2\alpha_1 k_2^2 + (1-k_2)^2\right)|x_i|^2 + \frac{\alpha_1}{2}|y_{i1}|^2 + \frac{\alpha_1}{2}|y_{i2}|^2 + \frac{1}{8}|y_{i1}+y_{i2}|^2 \\
&\qquad = \frac{1}{2}(1-k_2)|x_i|^2 + \frac{1}{4k_2}|y_{i1}|^2 + \frac{1}{4k_2}|y_{i2}|^2 - \frac{1}{8}|y_{i1}-y_{i2}|^2,
\end{aligned}$$

where we have defined again coefficient the $\alpha_1 \in \mathbb{R}_+^*$ by

$$\alpha_1 := 1 + \alpha - k_1,$$

in order to absorb the above-mentioned x_i-dependent remainder. In addition, note that we have used again the following relation between k_2 and α_1 to cancel the crossed term, *i.e.*,

$$k_2 = \frac{1}{1+2\alpha_1}.$$

• Step 3: Change of variables $x_{i1} \to y_{i1}$ and $x_{i2} \to y_{i2}$ for $i \in \mathsf{L}^n_{n-3}$.
We consider the change of variables given by

$$\begin{aligned}
x_{i1} &= k_3 x_i + y_{i1}, \\
x_{i2} &= k_3 x_i + y_{i2},
\end{aligned}$$

where k_3 is given by Definition 5.2. Putting together the terms for $i \in \mathsf{L}^n_{n-3}$ in (5.16) and the above x_i-dependent remainder in the above expression in STEP 2 yields the following term under the above change of variables

$$\begin{aligned}
&\frac{\alpha_2}{2}|x_{i1}|^2 + \frac{\alpha_2}{2}|x_{i2}|^2 + \frac{1}{2}\left|x_i - \frac{x_{i1}+x_{i2}}{2}\right|^2 \\
&\qquad = \frac{\alpha_2}{2}|k_3 x_i + y_{i1}|^2 + \frac{\alpha_2}{2}|k_3 x_i + y_{i2}|^2 + \frac{1}{2}\left|(1-k_3)x_i - \frac{y_{i1}+y_{i2}}{2}\right|^2 \\
&\qquad = \frac{1}{2}\left(2\alpha_2 k_3^2 + (1-k_3)^2\right)|x_i|^2 + \frac{\alpha_2}{2}|y_{i1}|^2 + \frac{\alpha_2}{2}|y_{i2}|^2 + \frac{1}{8}|y_{i1}+y_{i2}|^2 \\
&\qquad\qquad + \left(\alpha_2 k_3 - \frac{1-k_3}{2}\right) x_i \cdot y_{i1} + \left(\alpha_2 k_3 - \frac{1-k_3}{2}\right) x_i \cdot y_{i2} \\
&\qquad = \frac{1}{2}\left(2\alpha_2 k_3^2 + (1-k_3)^2\right)|x_i|^2 + \frac{\alpha_2}{2}|y_{i1}|^2 + \frac{\alpha_2}{2}|y_{i2}|^2 + \frac{1}{8}|y_{i1}+y_{i2}|^2 \\
&\qquad = \frac{1}{2}(1-k_3)|x_i|^2 + \frac{1}{4k_3}|y_{i1}|^2 + \frac{1}{4k_3}|y_{i2}|^2 - \frac{1}{8}|y_{i1}-y_{i2}|^2,
\end{aligned}$$

where we have defined again coefficient the $\alpha_2 \in \mathbb{R}^*_+$

$$\alpha_2 := 1 + \alpha - k_2,$$

in order to absorb the above-mentioned x_i-dependent remainder. In addition, note that we have used again the following relation between k_3 and α_2 to cancel the crossed term, *i.e.*,

$$k_3 = \frac{1}{1+2\alpha_2}.$$

Following a similar recursive process, we readily identify the quadratic form Q_n in Definition 5.4 in the exponential of the integrand in the final expression (5.17).

• STEP 4: Identifying the lineage maps Φ^j_n and the exponential factor.
On the one hand, the hierarchy of changes of variables immediately leads to

$$x_j = \Phi^j_n(x; \mathbf{y}_n),$$

for any leaf $j \in \mathsf{L}^n$, thanks to Definition 5.2 for the coefficients $\{\kappa_n\}_{n\in\mathbb{N}}$ and Definition 5.6 for the lineage maps Φ^j_n. On the other hand, since the Jacobian determinant of each change of variables is 1, no further factors appear during the iterative process except for the exponential x-dependent remainder $e^{-\frac{1+\alpha-k_n}{2}|x|^2}$ of the last step of the recurrence, which is not absorbed in the quadratic form Q_n. □

5.2. **Probabilistic reinterpretation.** For general purposes, in this part we reformulate the result in Proposition 5.7 in appropriate probabilistic terms. More precisely, this reformulation will not only provide shorter formulas, but it will also allow identifying a problematic key point when studying the asymptotics of the high-dimensional integral (5.17) in Section 6, namely, the presence of non-negligible correlations between the various factors indexed with indexes over leaves $j \in \mathsf{L}^n$ of the tree, which do not dissipate even for long time $n \to \infty$. We shall use the following notation.

Definition 5.8 (High-dimensional normal distributions)**.** *Consider any integer $n \in \mathbb{N}$ and*

$$\Sigma^m_n := \begin{pmatrix} \dfrac{(2-k_{n-m})k_{n-m}}{1-k_{n-m}} I_d & -\dfrac{k^2_{n-m}}{1-k_{n-m}} I_d \\ -\dfrac{k^2_{n-m}}{1-k_{n-m}} I_d & \dfrac{(2-k_{n-m})k_{n-m}}{1-k_{n-m}} I_d \end{pmatrix}, \tag{5.18}$$

for every $m = 0, \ldots, n-1$. Then, we will define the random vector $\mathbf{Y}_n = (Y_i)_{i\in\mathsf{T}^n_*}$ distributed according to the following multivariate normal distribution

$$G_n(\mathbf{y}_n) := \prod_{m=0}^{n-1} \prod_{i\in\mathsf{L}^n_m} \frac{1}{(2\pi)^d \sqrt{\det(\Sigma^m_n)}} \exp\left(-\frac{1}{2}(y^T_{i1}, y^T_{i2})(\Sigma^m_n)^{-1}\begin{pmatrix} y_{i1} \\ y_{i2}\end{pmatrix}\right), \quad \mathbf{y}_n \in \mathbb{R}^{2(2^n-1)d}, \tag{5.19}$$

where again we are using the tree-indexed notation $\mathbf{y}_n = (y_i)_{i\in\mathsf{T}^n_*} \in \mathbb{R}^{2(2^n-1)d}$ in Remark 3.1.

Notice that for every index $i \in \mathsf{L}^n_m$ with $m = 0, \ldots, n-1$ the pair (Y_{i1}, Y_{i2}) is normally distributed according to $\mathcal{N}(0, \Sigma^m_n)$ and Y_{i1} and Y_{i2} are certainly correlated. Nevertheless, the vector (Y_{i1}, Y_{i2}) is independent on any other random vector Y_j for indices $j \neq i1$ and $j \neq i2$. We are now ready to introduce the following probabilistic reformulation of Proposition 5.7.

Proposition 5.9 (Reformulation of the iterations II)**.** *Assume that $\alpha \in \mathbb{R}_+^*$ and set any initial datum $F_0 \in \mathcal{M}_+(\mathbb{R}^d)$. Hence, the solution $\{F_n\}_{n\in\mathbb{N}}$ to the time-discrete problem* (1.1) *admits the following form*

$$F_n(x) = \frac{e^{-\frac{1+\alpha-k_n}{2}|x|^2}}{(2\pi)^{d/2}2^{2^{n-1}d}} \left(\prod_{m=0}^{n-1} \frac{\left(\frac{4k_{n-m}^2}{1-k_{n-m}}\right)^{2^{m-1}d}}{\|F_m\|_{L^1(\mathbb{R}^d)}^{2^{n-1-m}}} \right) \mathbb{E}\left[\prod_{j\in\mathsf{L}^n} \bar{F}_0(\Phi_n^j(x;\mathbf{Y}_n)) \right], \tag{5.20}$$

for every $x \in \mathbb{R}^d$ and $n \in \mathbb{N}$, where we are using the high-dimensional random vector $\mathbf{Y}_n = (Y_i)_{i\in\mathsf{T}_*^n}$ in Definition 5.8 along with further notation from Definitions 5.1, 5.2, 5.4 and 5.6.

Proof. By Proposition 5.7 we obtain

$$\begin{aligned} F_n(x) &= \frac{e^{-\frac{1+\alpha-k_n}{2}|x|^2}}{(4\pi)^{2^{n-1}d}(2\pi)^{(2^n-1)d/2}} \frac{1}{\prod_{m=0}^{n-1}\|F_m\|_{L^1(\mathbb{R}^d)}^{2^{n-1-m}}} \int_{\mathbb{R}^{2(2^n-1)d}} e^{-Q_n(\mathbf{y}_n)} \prod_{j\in\mathsf{L}^n} \bar{F}_0(\Phi_n^j(x;\mathbf{y}_n))\, d\mathbf{y}_n \\ &= \frac{(2\pi)^{(2^n-1)d/2}}{(4\pi)^{2^{n-1}d}} e^{-\frac{1+\alpha-k_n}{2}|x|^2} \left(\prod_{m=0}^{n-1} \frac{\left(\frac{4k_{n-m}^2}{1-k_{n-m}}\right)^{2^{m-1}d}}{\|F_m\|_{L^1(\mathbb{R}^d)}^{2^{n-1-m}}} \right) \int_{\mathbb{R}^{2(2^n-1)d}} G_n(\mathbf{y}_n) \prod_{j\in\mathsf{L}^n} \bar{F}_0(\Phi_n^j(x;\mathbf{y}_n))\, d\mathbf{y}_n, \end{aligned}$$

for every $x \in \mathbb{R}^d$ and each $n \in \mathbb{N}$, where in the last line we have used that the inverse and determinant of the covariance matrices Σ_n^m in (5.18) take the form

$$(\Sigma_n^m)^{-1} := \begin{pmatrix} \frac{2-k_{n-m}}{4k_{n-m}} I_d & \frac{1}{4} I_d \\ \frac{1}{4} I_d & \frac{2-k_{n-m}}{4k_{n-m}} I_d \end{pmatrix}, \quad \det(\Sigma_n^m) = \left(\frac{4k_{n-m}^2}{1-k_{n-m}}\right)^d.$$

Then, the result follows by applying the law of the unconscious statistician (LOTUS) to relate the preceding integral to the expectation of the random variable $\prod_{j\in\mathsf{L}^n} \bar{F}_0(\Phi_n^j(x;\mathbf{Y}_n))$. □

As illustrated in the overall map in Figure 2, the reformulation of the iterations in Propositions 5.7 and 5.9 will be fundamental in order to characterize the long-time behavior of the solution $\{F_n\}_{n\in\mathbb{N}}$ to the time-discrete problem (1.1). More specifically, we need to unravel the asymptotic behavior of the high-dimensional integral encoded in the expectation term (5.20). As discussed in Sections 2.3 and 2.4, one is able to control it in the log-Lipschitz norm under stringent constraints on the initial data, which are not fully satisfactory. In view of the structure of the expectation term, alternatively one may expect to be able to interchange expectations with products and control each factor separately. However, as we show below, this naive idea is bound to fail since the involved random variables are correlated, and they do not uncorrelate even in the limit $n \to \infty$. This complicated structure imposes severe problems when handling formula (5.20), and one needs more powerfull ideas, which we introduce later in Lemma 6.3.

Remark 5.10 (Non-negligible correlations)**.** *Since $\mathbf{Y}_n$ in Definition 5.8 is normally distributed and the lineage map $\Phi_n^j(x;\mathbf{y}_n)$ in Definition 5.6 are affine transformations of $\mathbf{y}_n$, then all $(\Phi_n^j(x;\mathbf{Y}_n))_{j\in\mathsf{L}^n}$ are also normally distributed. Unfortunately, the components of $(\Phi_n^j(x;\mathbf{Y}_n))_{j\in\mathsf{L}^n}$ are correlated since, in particular, all the components $(i1, i2)$ are correlated in view of the structure* (5.18) *of their covariance matrices. Specifically, a straightforward computation shows that the covariance matrix*

$$K_{ij} = \mathbb{E}\left[(\Phi_n^i(x;\mathbf{Y}_n) - \mathbb{E}\,\Phi_n^i(x;\mathbf{Y}_n)) \otimes (\Phi_n^j(x;\mathbf{Y}_n) - \mathbb{E}\,\Phi_n^j(x;\mathbf{Y}_n)) \right],$$

between any couple $(i,j) \in \mathsf{L}^n \times \mathsf{L}^n$ of components (see Table 1 for the notation $\otimes$ of the Kronecker product) takes the following explicit form

$$K_{ij} = \begin{cases} \left(\displaystyle\sum_{q=0}^{n-1} \kappa_q^2 \frac{(2-k_{q+1})k_{q+1}}{1-k_{q+1}} \right) I_d, & \text{if} \quad |i\wedge j| = n, \\ \left(-\kappa_{n-|i\wedge j|-1}^2 \dfrac{k_{n-|i\wedge j|}^2}{1-k_{n-|i\wedge j|}} + \displaystyle\sum_{q=n-|i\wedge j|}^{n-1} \kappa_q^2 \frac{(2-k_{q+1})k_{q+1}}{1-k_{q+1}} \right) I_d, & \text{if} \quad 0 < |i\wedge j| < n, \\ \left(-\kappa_{n-1}^2 \dfrac{k_n^2}{1-k_n} \right) I_d, & \text{if} \quad |i\wedge j| = 0. \end{cases}$$

*Here, $|i\wedge j|$ represents the level of the highest common descendant $i\wedge j$ of i and j (*cf. *Section 3.1). Note that correlations only disappears for leaves (i,j) such that $n - |i\wedge j| \to \infty$, that is, when the length of the*

path from i (or j) to the highest common descendant $i \wedge j$ diverges as the height n of the tree increases. Unfortunately, correlations are non-negligible in all other cases even for $n \to \infty$.

6. Uniqueness of equilibria and quantitative ergodicity property

In this part we shall prove our main results, namely, Theorem 1.1 and Corollary 1.2. First, we will derive a local convergence result as $n \to \infty$ of the profiles $\{F_n\}_{n\in\mathbb{N}}$ solving (1.1) with generic initial datum $F_0 \in \mathcal{M}_+(\mathbb{R}^d)$ towards the Gaussian profile $\boldsymbol{F}_\alpha$ in (1.12). This weaker local analogue of Theorem 1.1 will be enough to derive the uniqueness of the solution $(\boldsymbol{\lambda}_\alpha, \boldsymbol{F}_\alpha)$ of the eigenproblem (1.11) in Corollary 1.2. Second, we will extend our local result into our main global result in Theorem 1.1 by quantifying the relaxation of the growth rate $\|F_n\|_{L^1(\mathbb{R}^d)}/\|F_{n-1}\|_{L^1(\mathbb{R}^d)}$ and the normalized profiles $F_n/\|F_n\|_{L^1(\mathbb{R}^d)}$ towards the unique solution $(\boldsymbol{\lambda}_\alpha, \boldsymbol{F}_\alpha)$ of the eigenproblem (1.11). Our strategy will exploit the high-dimensional iterative formula (5.20) in Proposition 5.9.

6.1. Local convergence result.

In this part we shall prove our local convergence result. To this end, we will show that the dependency of the expectation term in (5.20) on the variable x decays to zero as $n \to \infty$ uniformly over compact sets. To simplify notation, we shall define the following operator.

Definition 6.1 (Expectation operators)**.** *Under the notation in Definitions 5.6 and 5.8, we define*

$$\mathcal{E}_n[F](x) := \mathbb{E}\left[\prod_{j\in \mathsf{L}^n} F(\Phi_n^j(x;\mathbf{Y}_n))\right], \quad x \in \mathbb{R}^d, \tag{6.1}$$

for any $n \in \mathbb{N}$ and any non-negative measurable function $F : \mathbb{R}^d \to \mathbb{R}$.

Note that $\mathcal{E}_n[F](x)$ possibly takes extended values in $(-\infty, +\infty]$ if the expectation does not exist. However, if F is a bounded function (and it will be often the case), we obtain that $\mathcal{E}_n[F](x)$ always exists and is finite. The following technical lemma for the tail of the incomplete Gamma function will be required later to determine the asymptotic behavior of the above operator $\mathcal{E}_n$.

Lemma 6.2. *Consider the (upper) incomplete Gamma function with parameter $a \in \mathbb{R}_+^*$,* i.e.*,*

$$\Gamma(a, x) := \int_x^{+\infty} s^{a-1} e^{-s}\, ds, \quad x \in \mathbb{R}_+^*. \tag{6.2}$$

Then, the following estimate holds true

$$\Gamma(a+1, x) \le 2e^{-x/2}, \tag{6.3}$$

for every $x, a \in \mathbb{R}_+^*$ verifying the constraints

$$x > e, \quad 2a < \frac{x}{\log(x)}. \tag{6.4}$$

Proof. Let us restate the incomplete Gamma function as follows

$$\Gamma(a+1, x) = \int_x^{+\infty} \exp(-\psi_a(s)s)\, ds,$$

where $\psi_a(s)$ is the function

$$\psi_a(s) := 1 - a\frac{\log(s)}{s}, \quad s > 0.$$

Notice that ψ_a is monotonically increasing for $s \in [e, \infty)$. Hence, we obtain that

$$\psi_a(s) \ge \psi_a(x) \ge \frac{1}{2},$$

for every $s \ge x$, where we have used (6.4) on each of the above inequalities. Consequently, we get

$$\Gamma(a+1, x) \le \int_x^{+\infty} e^{-\frac{s}{2}}\, ds,$$

and this ends the proof. □

Therefore, we obtain the following weak dependency of $\mathcal{E}_n[F]$ on the variable x as $n \to \infty$.

Lemma 6.3 (Asymptotics of $\mathcal{E}_n$)**.** *Assume that $\alpha \in \mathbb{R}_+^*$ and consider any $F \in C^1(\mathbb{R}^d)$ such that*

$$\begin{aligned} 0 < F(x) &\leq C, \\ |\nabla \log F(x)| &\leq D(1+|x|), \end{aligned} \tag{6.5}$$

for every $x \in \mathbb{R}^d$ and appropriate $C, D \in \mathbb{R}_+^*$. Then, for every $\varepsilon \in \mathbb{R}_+^*$, there exists a large enough constant $C_\varepsilon \in \mathbb{R}_+^*$ such that the following property holds true

$$|\nabla \log \mathcal{E}_n[F](x)| \leq C_\varepsilon \left((2\,\boldsymbol{k}_\alpha + \varepsilon)^n + (\boldsymbol{r}_\alpha + \varepsilon)^n |x| + e^{-\frac{(2+\varepsilon)^n}{\varepsilon}} e^{C_\varepsilon (2\,\boldsymbol{r}_\alpha + \varepsilon)^n |x|^2} \right), \tag{6.6}$$

for every $x \in \mathbb{R}^d$ and any $n \in \mathbb{N}$, where $\mathcal{E}_n$ is the operator in Definition 6.1.

Proof. Before proving the main result, note that the hypothesis (6.5) guarantee that F is bounded below by a Gaussian profile. Indeed, by the fundamental theorem of calculus we obtain the following relation

$$\log F(x) = \log F(0) + \int_0^1 \nabla (\log F)(\theta x) \cdot x \, d\theta,$$

for any $x \in \mathbb{R}^d$. Using the linear growth assumption in (6.5) for the log-derivative $\nabla \log F$ we obtain that

$$\log F(x) \geq \log F(0) - D|x| \int_0^1 (1 + \theta |x|) \, d\theta \geq \log F(0) - \frac{D}{2} - D|x|^2,$$

for any $x \in \mathbb{R}^d$, where in the last inequality we have integrated with respect to θ and we have used Young's inequality to interpolate $|x|$ by $|x|^2$. Therefore, we obtain the following Gaussian lower bound

$$F(x) \geq c e^{-\beta |x|^2}, \tag{6.7}$$

for each $x \in \mathbb{R}^d$ and some $c, \beta \in \mathbb{R}_+^*$ depending only on $F(0)$ and D.

For simplicity of notation we define

$$E_n := \mathcal{E}_n[F],$$

for any $n \in \mathbb{N}$. In the sequel, we shall study the behavior of $\nabla \log E_n$ as $n \to \infty$.

• Step 1: Control of $\sum_{j \in \mathsf{L}^n} |\Phi_n^j(x; \mathbf{y}_n)|$.

By definition (5.15) of the lineage map Φ_n^j, we obtain that

$$\sum_{j \in \mathsf{L}^n} |\Phi_n^j(x; \mathbf{y}_n)| = \sum_{j \in \mathsf{L}^n} \left| \kappa_n x + \sum_{m=0}^{n-1} \kappa_m y_{\mathsf{c}^m(j)} \right| \leq 2^n \kappa_n |x| + \sum_{m=0}^{n-1} 2^m \kappa_m \sum_{i \in \mathsf{L}_{n-m}^n} |y_i|,$$

where we are using the tree-indexed notation $\mathbf{y}_n = (y_i)_{i \in \mathsf{T}_*^n} \in \mathbb{R}^{2(2^n-1)d}$ in Remark 3.1 and $\mathsf{c} : \mathsf{T}_*^n \longrightarrow \widehat{\mathsf{T}}^n$ is defined in Section 3.1. Note that in the last inequality we have used that each index of the level L_{n-m}^n in the tree contributes to the amount of 2^m leaves. Then, we obtain

$$\begin{aligned} \sum_{j \in \mathsf{L}^n} \left| \Phi_n^j(x; \mathbf{y}_n) \right| &\leq 2^n \kappa_n |x| + \left(\sum_{m=0}^{n-1} \sum_{i \in \mathsf{L}_{n-m}^n} (2^m \kappa_m)^2 \right)^{1/2} \left(\sum_{m=0}^{n-1} \sum_{i \in \mathsf{L}_{n-m}^n} |y_i|^2 \right)^{1/2} \\ &= 2^n \kappa_n |x| + \left(\sum_{m=0}^{n-1} 2^{n-m} (2^m \kappa_m)^2 \right)^{1/2} \| \mathbf{y}_n \| \\ &\leq 2^n \kappa_n |x| + \sqrt{2}\, 2^{n/2} \| \mathbf{y}_n \|, \end{aligned} \tag{6.8}$$

for any $\mathbf{y}_n \in \mathbb{R}^{2(2^n-1)d}$, where in the first line we have used the Cauchy–Schwarz inequality on the joint sum over indices m and i, in the second line we have used the definition $\| \mathbf{y}_n \| = (\sum_{i \in \mathsf{T}_*^n} |y_i|^2)^{1/2}$ of the ℓ_2 norm (*cf.* Table 1), and in the last line we have used that

$$\sum_{m=0}^{n-1} 2^{n-m} (2^m \kappa_m)^2 \leq 2^n \sum_{m=0}^{n-1} 2^m k_1^{2m} \leq 2^n \sum_{m=0}^{\infty} 2^m k_1^{2m} = \frac{2^n}{1 - 2k_1^2} \leq 2^{n+1}.$$

Above, we have used that $k_n \leq k_1$ for all $n \in \mathbb{N}$ by virtue of Lemma 5.3(i) and also that

$$\frac{1}{1 - 2k_1^2} = \frac{1 + 4\alpha + 2\alpha^2}{2(1+\alpha)^2} = \frac{2(1+\alpha^2)}{(1+\alpha)^2 + 2\alpha + \alpha^2} < 2.$$

• Step 2: Derivative of $\log E_n$.
Taking derivatives on (6.1), using the second property in (6.5) along with (6.8) entails

$$\begin{aligned}|\nabla \log E_n(x)| &\le \kappa_n \frac{\mathbb{E}\left[\left(\sum_{j\in\mathsf{L}^n} |\nabla \log F(\Phi_n^j(x;\mathbf{Y}_n))|\right)\prod_{j\in\mathsf{L}^n} F(\Phi_n^j(x;\mathbf{Y}_n))\right]}{\mathbb{E}\left[\prod_{j\in\mathsf{L}^n} F(\Phi_n^j(x;\mathbf{Y}_n))\right]}\\ &\lesssim \kappa_n \frac{\mathbb{E}\left[\left(\sum_{j\in\mathsf{L}^n} (1+|\Phi_n^j(x;\mathbf{Y}_n)|)\right)\prod_{j\in\mathsf{L}^n} F(\Phi_n^j(x;\mathbf{Y}_n))\right]}{\mathbb{E}\left[\prod_{j\in\mathsf{L}^n} F(\Phi_n^j(x;\mathbf{Y}_n))\right]}\\ &\lesssim 2^n\kappa_n + 2^n\kappa_n^2|x| + 2^{n/2}\kappa_n R_n(x),\end{aligned} \tag{6.9}$$

for every $x\in\mathbb{R}^d$, where the remainder R_n reads

$$R_n(x) := \frac{\mathbb{E}\left[\|\mathbf{Y}_n\|\prod_{j\in\mathsf{L}^n} F(\Phi_n^j(x;\mathbf{Y}_n))\right]}{\mathbb{E}\left[\prod_{j\in\mathsf{L}^n} F(\Phi_n^j(x;\mathbf{Y}_n))\right]},\quad x\in\mathbb{R}^d.$$

This estimation provides a first insight about the ergodicity property. Indeed the correlations, as measured by κ_n, decreases faster than 2^{-n} (*cf.* Lemma 5.3), meaning that the contribution $2^n\kappa_n$ decays fast to zero. It is also the case of the other contributions, locally in x, but the last one, namely $2^{n/2}\kappa_n R_n(x)$, requires more work. Consider any $\sqrt{2} < q < \frac{1}{\sqrt{2}\,\boldsymbol{k}_\alpha}$, which exists because $2\,\boldsymbol{k}_\alpha < 1$ for $\alpha\in\mathbb{R}_+^*$, and split $R_n(x) = R_n^1(x) + R_n^2(x)$ in terms of the functions

$$R_n^1(x) := \frac{1}{Z_n(x)}\int_{\|\mathbf{y}_n\|\le q^n} \|\mathbf{y}_n\| \prod_{j\in\mathsf{L}^n} F(\Phi_n^j(x;\mathbf{y}_n)) e^{-Q_n(\mathbf{y}_n)}\,d\mathbf{y}_n, \tag{6.10}$$

$$R_n^2(x) := \frac{1}{Z_n(x)}\int_{\|\mathbf{y}_n\|> q^n} \|\mathbf{y}_n\| \prod_{j\in\mathsf{L}^n} F(\Phi_n^j(x;\mathbf{y}_n)) e^{-Q_n(\mathbf{y}_n)}\,d\mathbf{y}_n, \tag{6.11}$$

and the normalization factor

$$Z_n(x) := \int_{\mathbb{R}^{2(2^n-1)d}} \prod_{j\in\mathsf{L}^n} F(\Phi_n^j(x;\mathbf{y}_n)) e^{-Q_n(\mathbf{y}_n)}\,d\mathbf{y}_n. \tag{6.12}$$

• Step 3: Lower bound for Z_n.
Using the above lower bound by a Gaussian in (6.7) along with the upper bound for Q_n in Lemma 5.5 implies the following estimate of Z_n in (6.12):

$$\begin{aligned}Z_n(x) &\ge c^{2^n}\int_{\mathbb{R}^{2(2^n-1)d}} \exp\left(-\beta\sum_{j\in\mathsf{L}^n}|\Phi_n^j(x;\mathbf{y}_n)|^2\right)\exp(-\beta_{\max}\|\mathbf{y}_n\|^2)\,d\mathbf{y}_n\\ &\ge c^{2^n}e^{-2^{2n+1}\kappa_n^2\beta\|x\|^2}\int_{\mathbb{R}^{2(2^n-1)d}}\exp\left(-(2^{n+2}\beta+\beta_{\max})\|\mathbf{y}_n\|^2\right)d\mathbf{y}_n\\ &= c^{2^n}e^{-2^{2n+1}\kappa_n^2\beta|x|^2}\left|\mathbb{S}^{2(2^n-1)d-1}\right|\int_0^{+\infty} r^{2(2^n-1)d-1}\exp\left(-(2^{n+2}\beta+\beta_{\max})r^2\right)dr\\ &= \frac{c^{2^n}e^{-2^{2n+1}\kappa_n^2\beta|x|^2}}{2(2^{n+2}\beta+\beta_{\max})^{(2^n-1)d}}\left|\mathbb{S}^{2(2^n-1)d-1}\right|\int_0^{+\infty} s^{(2^n-1)d-1}e^{-s}\,ds\\ &= \frac{c^{2^n}e^{-2^{2n+1}\kappa_n^2\beta|x|^2}}{2(2^{n+2}\beta+\beta_{\max})^{(2^n-1)d}}\left|\mathbb{S}^{2(2^n-1)d-1}\right|\Gamma((2^n-1)d),\end{aligned}$$

where in the last step we have used the Gamma function, that is defined by

$$\Gamma(a) := \int_0^{+\infty} s^{a-1}e^{-s}\,ds,\quad a>0.$$

Recall that area of the hypersphere is $|\mathbb{S}^{d-1}| = \frac{2\pi^{d/2}}{\Gamma\left(\frac{d}{2}\right)}$. Then, we obtain

$$Z_n(x) \ge \frac{c^{2^n}\pi^{(2^n-1)d}}{(2^{n+2}\beta+\beta_{\max})^{(2^n-1)d}}e^{-2^{2n+1}\kappa_n^2\beta|x|^2}, \tag{6.13}$$

for every $x\in\mathbb{R}^d$.

• Step 4: Upper bound for R_n^1 and R_n^2.
On the one hand, notice that

$$R_n^1(x) \leq q^n, \tag{6.14}$$

for all $x \in \mathbb{R}^d$ by definition of R_n^1 in (6.10). On the other hand, using the lower bound of Q_n in Lemma 5.5 we can estimate R_n^2 in (6.11) by

$$\begin{aligned} R_n^2(x) &\leq \frac{C^{2^n}}{q^n Z_n(x)} \int_{\|\mathbf{y}_n\| \geq q^n} \|\mathbf{y}_n\|^2 e^{-\beta_{\min}\|\mathbf{y}_n\|^2} \, d\mathbf{y}_n \\ &= \frac{C^{2^n}}{q^n Z_n(x)} \left|\mathbb{S}^{2(2^n-1)d-1}\right| \int_{q^n}^{+\infty} r^{2(2^n-1)d+1} e^{-\beta_{\min} r^2} \, dr \\ &= \frac{C^{2^n}}{q^n Z_n(x)} \left|\mathbb{S}^{2(2^n-1)d-1}\right| \frac{1}{2\beta_{\min}^{(2^n-1)d+1}} \int_{\beta_{\min} q^{2n}}^{+\infty} s^{(2^n-1)d} e^{-s} \, ds \\ &= \frac{C^{2^n} \pi^{(2^n-1)d}}{q^n \beta_{\min}^{(2^n-1)d+1} Z_n(x)} \frac{\Gamma\left((2^n-1)d+1, \beta_{\min} q^{2n}\right)}{\Gamma((2^n-1)d)}, \end{aligned} \tag{6.15}$$

for all $x \in \mathbb{R}^d$, where we have used again the incomplete Gamma function (6.2) in Lemma 6.2. Putting the lower estimate (6.13) for Z_n into (6.15) yields

$$R_n^2(x) \lesssim e^{2^{2n+1}\kappa_n^2 \beta |x|^2} \frac{C^{2^n} (2^{n+2}\beta + \beta_{\max})^{(2^n-1)d}}{q^n c^{2^n} \beta_{\min}^{(2^n-1)d}} \frac{\Gamma\left((2^n-1)d+1, \beta_{\min} q^{2n}\right)}{\Gamma((2^n-1)d)}, \tag{6.16}$$

for all $x \in \mathbb{R}^d$. Now, take $a = (2^n-1)d$ and $x = \beta_{\min} q^{2n}$ and notice that the constraints (6.4) in Lemma 6.2 hold true for large enough $n \in \mathbb{N}$ as long as we choose $q > \sqrt{2}$. Hence, we obtain the estimate

$$\Gamma\left((2^n-1)d+1, \beta_{\min} q^{2n}\right) \leq 2 e^{-\frac{\beta_{\min}}{2} q^{2n}},$$

for all $n \geq n_* = n_*(\beta_{\min}, d, q)$. Using Stirling's formula for the remaining Gamma function in (6.16), we conclude the estimate

$$R_n^2(x) \lesssim e^{2^{2n+1}\kappa_n^2 \beta |x|^2} \frac{C^{2^n} (2^{n+2}\beta + \beta_{\max})^{(2^n-1)d}}{q^n c^{2^n} \beta_{\min}^{(2^n-1)d}} \frac{e^{-\frac{\beta_{\min}}{2} q^{2n}}}{(2^n-1)^{1/2} \left(\frac{(2^n-1)d}{e}\right)^{(2^n-1)d}}, \tag{6.17}$$

for every $x \in \mathbb{R}^d$ and each $n \geq n_*$.

• Step 5: Final conclusion.
Let us put the preceding estimates (6.14) and (6.17) into (6.9) to achieve

$$|\nabla \log E_n(x)| \lesssim 2^n \kappa_n + 2^{n/2} q^n \kappa_n + 2^n \kappa_n^2 |x| + \frac{M^{2^n}}{e^{\frac{\beta_{\min}}{2} q^{2n}}} e^{2^{2n+1}\kappa_n^2 \beta |x|^2}, \tag{6.18}$$

for every $x \in \mathbb{R}^d$ and each $n \geq n_*$, where $M = M(C, c, \beta, \beta_{\min}, \beta_{\max}, d, q)$ is a universal constant that we have introduced to absorb all the double exponential factors of the form a^{2^n} with $a > 0$ in the fourth term along with any further lower order factor. On the one hand, choosing $q > \sqrt{2}$ sufficiently close to $\sqrt{2}$ and using (5.11) in Lemma 5.3 we obtain the following asymptotics for the factors in the first three terms of (6.18):

$$\begin{aligned} 2^n \kappa_n &\leq C_\varepsilon (2\,\boldsymbol{k}_\alpha + \varepsilon)^n, \\ 2^{n/2} q^n \kappa_n &\leq C_\varepsilon (2\,\boldsymbol{k}_\alpha + \varepsilon)^n, \\ 2^n \kappa_n^2 &\leq C_\varepsilon (2(\boldsymbol{k}_\alpha + \varepsilon)^2)^n \leq C_\varepsilon (\boldsymbol{r}_\alpha + \varepsilon)^n, \end{aligned}$$

for arbitrarily small $\varepsilon \in \mathbb{R}_+^*$ and sufficiently large $C_\varepsilon \in \mathbb{R}_+^*$. Note that in the last identity we have used the relation $\boldsymbol{r}_\alpha = 2\,\boldsymbol{k}_\alpha^2$. On the other hand, since q has been taken larger than $\sqrt{2}$ and arbitrarily close to $\sqrt{2}$, then the double exponential in the denominator of the last term in (6.18) kills the double exponential on the numerator independently on the explicit value of M. Therefore, we obtain

$$\frac{M^{2^n}}{e^{\frac{\beta_{\min}}{2} q^{2n}}} \leq C_\varepsilon \, e^{-\frac{(2+\varepsilon)^n}{\varepsilon}},$$

by appropriately increasing C_ε if necessary. Similarly, we infer the following control for the remaining x-dependent exponential factor in the fourth term of (6.18)

$$e^{2^{2n+1}\kappa_n^2 \beta |x|^2} \leq e^{C_\varepsilon (4(\boldsymbol{k}_\alpha + \varepsilon)^2)^n |x|^2} \leq e^{C_\varepsilon (2\,\boldsymbol{r}_\alpha + \varepsilon)^n |x|^2},$$

which yields the announced estimate (6.6). □

Remark 6.4. *Under the assumptions in Lemma 6.3, notice that the above result implies in particular that for every $\varepsilon \in \mathbb{R}_+^*$ and $R \in \mathbb{R}_+^*$, there exists a large enough constant $C_{\varepsilon,R} \in \mathbb{R}_+^*$ such that*

$$\sup_{|x|\leq R} |\nabla \log \mathcal{E}_n[F](x)| \leq C_{\varepsilon,R}\,(2\,\boldsymbol{k}_\alpha)^n,$$

for any $n \in \mathbb{N}$. In particular, $\nabla \log \mathcal{E}_n[F]$ converge to zero as $n \to \infty$ uniformly on compact sets.

The above observation is the cornerstone to prove the following local version of our main Theorem 1.1.

Corollary 6.5 (Local convergence result)**.** *Assume that $\alpha \in \mathbb{R}_+^*$, set any initial datum $F_0 \in \mathcal{M}_+(\mathbb{R}^d)$ and consider the solution $\{F_n\}_{n\in\mathbb{N}}$ to the time-discrete problem* (1.1)*. Then, for every $\varepsilon \in \mathbb{R}_+^*$ and $R \in \mathbb{R}_+^*$, there exists a large enough constant $C_{\varepsilon,R} \in \mathbb{R}_+^*$ such that*

$$\sup_{|x|\leq R} \left| \frac{F_n(x)}{\|F_n\|_{L^1(\mathbb{R}^d)}} - \boldsymbol{F}_\alpha(x) \right| \leq C_{\varepsilon,R}\,(2\,\boldsymbol{k}_\alpha)^n, \tag{6.19}$$

for any $n \in \mathbb{N}$, where $\boldsymbol{F}_\alpha$ is the Gaussian profile (1.12)*. In particular, $F_n/\|F_n\|_{L^1(\mathbb{R}^d)}$ converge to $\boldsymbol{F}_\alpha$ as $n \to \infty$ uniformly on compact sets.*

In the proof of Corollary 6.5 we will require a careful control of the asymptotic behavior of the ratios $F_n(0)/\|F_n\|_{L^1(\mathbb{R}^d)}$ as $n \to \infty$, which we provide below. First, note that if $\{F_n\}_{n\in\mathbb{N}}$ is the solution (3.4) of the time-discrete problem (1.1) issued at a Gaussian initial datum F_0 like in Proposition 3.7, we have

$$\frac{F_n(0)}{\|F_n\|_{L^1(\mathbb{R}^d)}} = G_{\mu_n,\sigma_n^2}(0) \to \boldsymbol{F}_\alpha(0) = \frac{1}{(2\pi\,\boldsymbol{\sigma}_\alpha^2)^{d/2}},$$

as $n \to \infty$ explicitly, where μ_n and σ_n^2 are determined by the recursive relations (3.5) and we have used that $\mu_n \to 0$ and $\sigma_n^2 \to \boldsymbol{\sigma}_\alpha^2$ according to Lemma 3.8. In the following lemma, we provide a similar quantitative result with convergence rates for generic initial data $F_0 \in \mathcal{M}_+(\mathbb{R}^d)$.

Lemma 6.6. *Assume that $\alpha \in \mathbb{R}_+^*$, set any initial datum $F_0 \in \mathcal{M}_+(\mathbb{R}^d)$ and consider the solution $\{F_n\}_{n\in\mathbb{N}}$ to the time-discrete problem* (1.1)*. For any $\varepsilon \in \mathbb{R}_+^*$ there exists $C_\varepsilon \in \mathbb{R}_+^*$ sufficiently large with*

$$\left| \frac{F_n(0)}{\|F_n\|_{L^1(\mathbb{R}^d)}} - \boldsymbol{F}_\alpha(0) \right| \leq C_\varepsilon (2\,\boldsymbol{k}_\alpha + \varepsilon)^n, \tag{6.20}$$

for any $n \in \mathbb{N}$.

Proof. Let us fix any $\varepsilon \in \mathbb{R}_+^*$ arbitrarily small so that $2\,\boldsymbol{k}_\alpha + \varepsilon < 1$. Again, this is possible because $2\,\boldsymbol{k}_\alpha < 1$ for $\alpha \in \mathbb{R}_+^*$. For simplicity of notation, we shall define the following sequence of coefficients:

$$c_n := \frac{F_n(0)}{\|F_n\|_{L^1(\mathbb{R}^d)}}, \quad n \in \mathbb{N}.$$

Our goal then reduces to studying the asymptotic behavior of $\{c_n\}_{n\in\mathbb{N}}$ and obtaining quantitative convergence rates. Note that by formula (5.20) in Proposition 5.9 for the recursion, we obtain

$$\frac{1}{c_n} = \frac{\int_{\mathbb{R}^d} e^{-\frac{1+\alpha-k_n}{2}|x|^2} \mathcal{E}_n[\bar{F}_0](x)\,dx}{\mathcal{E}_n[\bar{F}_0](0)} = \int_{|x|\leq R_n} e^{-\frac{1+\alpha-k_n}{2}|x|^2} \frac{\mathcal{E}_n[\bar{F}_0](x)}{\mathcal{E}_n[\bar{F}_0](0)}\,dx + \frac{1}{c_n}\int_{|x|>R_n} \frac{F_n(x)}{\|F_n\|_{L^1(\mathbb{R}^d)}}\,dx,$$

where we set the sequence of radii $\{R_n\}_{n\in\mathbb{N}}$ as follows

$$R_n := \frac{1}{(2\,\boldsymbol{k}_\alpha)^{\frac{n}{2m+1}}}, \quad n \in \mathbb{N}, \tag{6.21}$$

for a fixed value $m \in \mathbb{N}$, which we take large enough so that

$$(2\,\boldsymbol{k}_\alpha)^{\frac{2m}{2m+1}} \leq 2\,\boldsymbol{k}_\alpha + \frac{\varepsilon}{2}. \tag{6.22}$$

Whilst our choice of R_n is not justified at first glance, we claim that it has been taken as to minimize the decay rate on the error terms $\widetilde{E}_n$ and E_n below. By solving the implicit equation on c_n we infer

$$c_n = \frac{1-\tilde{E}_n}{(2\pi\,\boldsymbol{\sigma}_\alpha^2)^{d/2} + E_n}, \tag{6.23}$$

for every $n \in \mathbb{N}$, where the error terms $\tilde{E}_n$ and E_n take the form

$$\begin{aligned} \tilde{E}_n &:= \int_{|x|>R_n} \frac{F_n(x)}{\|F_n\|_{L^1(\mathbb{R}^d)}}\, dx, \\ E_n &:= \int_{|x|\leq R_n} e^{-\frac{1+\alpha-k_n}{2}|x|^2} \frac{\mathcal{E}_n[\bar{F}_0](x)}{\mathcal{E}_n[\bar{F}_0](0)}\, dx - (2\pi\, \boldsymbol{\sigma}_\alpha^2)^{d/2}, \end{aligned} \tag{6.24}$$

Additionally, we can split the second error term as $E_n = E_{n,1} - E_{n,2} + E_{n,3}$ with

$$\begin{aligned} E_{n,1} &:= \left(\frac{2\pi}{1+\alpha-k_n}\right)^{d/2} - \left(\frac{2\pi}{1+\alpha-\boldsymbol{k}_\alpha}\right)^{d/2}, \\ E_{n,2} &:= \int_{|x|>R_n} e^{-\frac{1+\alpha-k_n}{2}|x|^2}\, dx, \\ E_{n,3} &:= \int_{|x|\leq R_n} e^{-\frac{1+\alpha-k_n}{2}|x|^2} \left(\frac{\mathcal{E}_n[\bar{F}_0](x)}{\mathcal{E}_n[\bar{F}_0](0)} - 1\right) dx, \end{aligned} \tag{6.25}$$

where we have used the relationship $\boldsymbol{\sigma}_\alpha^2(1+\alpha-\boldsymbol{k}_\alpha) = 1$, which results from (1.13) and (1.15), and therefore the decomposition of E_n becomes clear since $E_{n,1}$ can be reformulated as

$$E_{n,1} = \int_{\mathbb{R}^d} e^{-\frac{1+\alpha-k_n}{2}|x|^2}\, dx - (2\pi\, \boldsymbol{\sigma}_\alpha^2)^{d/2}.$$

On the one hand, given any fixed value $\theta \in \mathbb{R}^+$ with $\theta < \frac{\alpha}{2}$ (for instance $\theta = \frac{\alpha}{4}$) the propagation of exponential moments in Corollary 4.8 implies that

$$E_\theta := \sup_{n\in\mathbb{N}} \int_{\mathbb{R}^d} e^{\theta|x|^2} \frac{F_n(x)}{\|F_n\|_{L^1(\mathbb{R}^d)}}\, dx < \infty.$$

By expanding the exponential in power series, one obtains in particular uniformly bounded moments of any order, and in particular, of order $2m$, namely

$$M_m := \sup_{n\in\mathbb{N}} \int_{\mathbb{R}^d} |x|^{2m} \frac{F_n(x)}{\|F_n\|_{L^1(\mathbb{R}^d)}}\, dx \leq \frac{E_\theta\, m!}{\theta^m} < \infty.$$

Hence, we infer the following control on the first error term in (6.24)

$$\tilde{E}_n \leq \frac{1}{R_n^{2m}} \int_{\mathbb{R}^d} |x|^{2m} \frac{F_n(x)}{\|F_n\|_{L^1(\mathbb{R}^d)}}\, dx \lesssim \frac{1}{R_n^{2m}}, \tag{6.26}$$

for any $n \in \mathbb{N}$. On the other hand, by Lemma 5.3 we have that $E_{n,1}$ in (6.25) can be bounded by

$$|E_{n,1}| \lesssim k_n - \boldsymbol{k}_\alpha \lesssim \boldsymbol{r}_\alpha^n\,. \tag{6.27}$$

By direct calculation we also obtain

$$E_{n,2} = |\mathbb{S}^{d-1}| \int_{R_n}^{+\infty} r^{d-1} e^{-\frac{1+\alpha-k_n}{2}r^2}\, dr \lesssim \Gamma\left(\frac{d}{2}, \frac{1+\alpha-k_n}{2} R_n^2\right) \lesssim \exp\left(-\frac{1+\alpha-k_n}{4(2\,\boldsymbol{k}_\alpha)^{\frac{2n}{2m+1}}}\right), \tag{6.28}$$

where $\Gamma(a,x)$ is the incomplete Gamma function (6.2) and we have used estimate (6.3) in Lemma 6.2. Note that the constraint (6.4) is trivially satisfied since $R_n \to \infty$ by our choice (6.21). Finally, we control the error term $E_{n,3}$ in (6.25) under the addition assumption that $\bar{F}_0$ satisfies the hypothesis (6.5) in Lemma 6.3. Whilst this is not always true, we show at the end of the proof that we can always assume so without loss of generality by replacing the argument on $\bar{F}_0$ by an advance enough time step $\bar{F}_m$ so that selection has properly shaped the Gaussian tails. Under this condition, Lemma 6.3 implies that given any $\varepsilon' \in \mathbb{R}_+^*$ there is $C_{\varepsilon'} \in \mathbb{R}_+^*$ so that, for $|x| \leq R_n$, we have

$$\begin{aligned} &\left|\int_0^1 (\nabla \log \mathcal{E}_n)[\bar{F}_0](\theta x)\cdot x\, d\theta\right| \\ &\lesssim (2\,\boldsymbol{k}_\alpha + \varepsilon')^n R_n + (\boldsymbol{r}_\alpha + \varepsilon')^n R_n^2 + e^{-\frac{(2+\varepsilon')^n}{\varepsilon'}} e^{C_{\varepsilon'}(2\,\boldsymbol{r}_\alpha+\varepsilon')^n R_n^2} R_n \\ &\lesssim \left(\frac{2\,\boldsymbol{k}_\alpha}{(2\,\boldsymbol{k}_\alpha)^{\frac{1}{2m+1}}} + \frac{\varepsilon}{2}\right)^n + \left(\frac{\boldsymbol{r}_\alpha}{(2\,\boldsymbol{k}_\alpha)^{\frac{2}{2m+1}}} + \frac{\varepsilon}{2}\right)^n + e^{-\frac{(2+\varepsilon')^n}{\varepsilon'}} \exp\left[C_{\varepsilon'}\left(\frac{2\,\boldsymbol{r}_\alpha}{(2\,\boldsymbol{k}_\alpha)^{\frac{2}{2m+1}}} + \frac{\varepsilon}{2}\right)^n\right] \frac{1}{(2\,\boldsymbol{k}_\alpha)^{\frac{n}{2m+1}}} \\ &\lesssim \left((2\,\boldsymbol{k}_\alpha)^{\frac{2m}{2m+1}} + \frac{\varepsilon}{2}\right)^n \lesssim (2\,\boldsymbol{k}_\alpha + \varepsilon)^n\,. \end{aligned}$$

In the second inequality above we have taken ε' small enough compared with ε. In the third inequality we have used the relation $\boldsymbol{r}_\alpha = 2\,\boldsymbol{k}_\alpha^2$, which guarantees that all the contributions in the third line are

dominated by the first term $((2\,\boldsymbol{k}_\alpha)^{\frac{2m}{2m+1}}+\frac{\varepsilon}{2})^n$. Finally, in the last inequality we have used our choice of $m\in\mathbb{N}$ large enough so that (6.22) holds. Therefore, the mean value theorem implies

$$\left|\frac{\mathcal{E}_n[\bar{F}_0](x)}{\mathcal{E}_n[\bar{F}_0](0)}-1\right| = \left|\exp\left(\int_0^1(\nabla\log\mathcal{E}_n[\bar{F}_0])(\theta x)\cdot x\,d\theta\right)-1\right| \lesssim (2\,\boldsymbol{k}_\alpha+\varepsilon)^n.$$

Hence, $E_{n,3}$ in (6.25) can be controlled by

$$|E_{n,3}| \leq \int_{|x|\leq R_n} e^{-\frac{1+\alpha-k_n}{2}|x|^2}\left|\frac{\mathcal{E}_n[\bar{F}_0](x)}{\mathcal{E}_n[\bar{F}_0](0)}-1\right|dx \lesssim (2\,\boldsymbol{k}_\alpha+\varepsilon)^n. \tag{6.29}$$

Putting (6.27), (6.28) and (6.29) together yields

$$|E_n| \lesssim (2\,\boldsymbol{k}_\alpha+\varepsilon)^n. \tag{6.30}$$

Thereby, using the mean value theorem on (6.23) along with the bounds (6.26) and (6.30) concludes that

$$\left|c_n-\frac{1}{(2\pi\,\boldsymbol{\sigma}_\alpha^2)^{d/2}}\right| \lesssim \tilde{E}_n+|E_n| \lesssim (2\,\boldsymbol{k}_\alpha+\varepsilon)^n.$$

To end the proof, we show that we can always assume that the hypothesis (6.5) in Lemma 6.3 is satisfied without loss of generality. Indeed, note that after two iterations $\bar{F}_2$ already satisfies $(6.5)_2$ thanks to Corollary 4.2, but only two iterations are not enough to guarantee $(6.5)_1$, in general. Nevertheless, applying Corollary 4.2 leads to the following upper bound on the normalized profiles:

$$\bar{F}_m(x) \leq C_m\, e^{\frac{1}{2}\left(\alpha+\frac{1}{2}-\frac{1}{\overline{\sigma}_m^2}\right)|x|^2},$$

for appropriate $C_m\in\mathbb{R}_+^*$, with variances $\{\overline{\sigma}_m^2\}_{m\in\mathbb{N}}$ defined by the recurrence (4.5), *i.e.*,

$$\frac{1}{\overline{\sigma}_{m+1}^2} = \alpha+\frac{1}{1+\frac{\overline{\sigma}_m^2}{2}},\quad m\in\mathbb{N},$$

and with initial datum $\overline{\sigma}_1^2=1/\alpha$. The boundedness condition $(6.5)_2$ is then satisfied by $\bar{F}_m$ if we can show that the prefactor $\alpha+\frac{1}{2}-\frac{1}{\overline{\sigma}_m^2}$ in the above exponential bound becomes non-positive for large enough m. This is where the precise choice of normalization in Definition 5.1 plays a role. Specifically, recall that so defined $\overline{\sigma}_m^2\to\boldsymbol{\sigma}_\alpha^2$ and we have precise relaxation rates by Lemma 3.8. Hence,

$$\alpha+\frac{1}{2}-\frac{1}{\overline{\sigma}_m^2} = \left(\alpha+\frac{1}{2}-\frac{1}{\boldsymbol{\sigma}_\alpha^2}\right)+\left(\frac{1}{\boldsymbol{\sigma}_\alpha^2}-\frac{1}{\overline{\sigma}_m^2}\right) \leq \left(\alpha+\frac{1}{2}-\frac{1}{\boldsymbol{\sigma}_\alpha^2}\right)+C_v\,\boldsymbol{r}_\alpha^m,$$

for any $m\in\mathbb{N}$. Since we chose a normalization $\bar{F}_n$ of the profiles by a Gaussian G_{0,σ^2} with variance σ^2 strictly larger that the variance $\boldsymbol{\sigma}_\alpha^2$ of the equilibrium $\boldsymbol{F}_\alpha$ (more specifically $\frac{1}{\sigma^2}=\frac{1}{\sigma^2}=\alpha+\frac{1}{2}$), we can guarantee that the right hand side above is non-positive if $m\geq n_\alpha$ for a sufficiently large $n_\alpha\geq 2$ (depending only on α). This justifies that $\bar{F}_{n_\alpha}$ satisfies both conditions in (6.5).

□

We are now in position to prove the above local convergence result in Corollary 6.5

Proof of Corollary 6.5. By formula (5.20) in Proposition 5.9 and Definition 6.1 we obtain

$$\nabla\log F_n(x) = -(1+\alpha-k_n)x+\nabla\log\mathcal{E}_n[\bar{F}_0](x),$$

for any $x\in\mathbb{R}^d$ and every $n\in\mathbb{N}$. Using the mean value theorem we then achieve

$$\begin{aligned}\log\frac{F_n(x)}{F_n(0)}-\log\frac{\boldsymbol{F}_\alpha(x)}{\boldsymbol{F}_\alpha(0)} &= \int_0^1(\nabla\log F_n)(\theta x)\cdot x\,d\theta+\frac{1}{2\,\boldsymbol{\sigma}_\alpha^2}|x|^2\\ &= -\frac{1+\alpha-k_n}{2}|x|+\int_0^1\nabla\log\mathcal{E}_n[\bar{F}_0](\theta x)\cdot x\,d\theta+\frac{1}{2\,\boldsymbol{\sigma}_\alpha^2}|x|^2\\ &= \frac{k_n-\boldsymbol{k}_\alpha}{2}|x|^2+\int_0^1\nabla\log\mathcal{E}_n[\bar{F}_0](\theta x)\cdot x\,d\theta,\end{aligned} \tag{6.31}$$

where we have used the relation $1/\boldsymbol{\sigma}_\alpha^2=1+\alpha-\boldsymbol{k}_\alpha$. On the one hand, the first term in the right hand side of (6.31) converges to zero thanks to Lemma 5.3. On the other hand, for the second term we shall apply Lemma 6.3. To do so we need to guarantee again that $\bar{F}_0$ verifies the hypothesis (6.5) of such a lemma. Recall that we can always assume the condition (6.5) without loss of generality (*cf.* last step in the proof of Lemma 6.6). Therefore, we can apply Remark 6.4 with $F=\bar{F}_0$ and obtain that $\nabla\log\mathcal{E}_n[\bar{F}_0]$

converges to zero uniformly over compact sets with explicit convergence rates. Putting it into (6.31) we obtain more explicitly

$$\sup_{|x|\leq R}\left|\log\frac{F_n(x)}{F_n(0)}-\log\frac{\boldsymbol{F}_\alpha(x)}{\boldsymbol{F}_\alpha(0)}\right|\leq C_{\varepsilon,R}\,(2\,\boldsymbol{k}_\alpha)^n,$$

for any $R\in\mathbb{R}_+^*$, any $n\in\mathbb{N}$ and a large enough $C_{\varepsilon,R}\in\mathbb{R}_+^*$. This, together with the above control on the asymptotics of $\|F_n\|_{L^1(\mathbb{R}^d)}/F_n(0)$ in Lemma 6.6, allow proving (6.19). □

In particular, note that the above Corollary 6.5 is enough to prove the uniqueness of solutions to the eigenproblem (1.11), as stated in Corollary 1.2.

Proof of Corollary 1.2. Let (λ,F) be any solution of the eigenproblem (1.11). Then, the ansatz (1.10) defines a solution $F_n=\lambda^n F$ of the time-discrete problem (1.1). By Corollary 6.5 we obtain that

$$\lim_{n\to\infty}\sup_{|x|\leq R}\left|\frac{\lambda^n F(x)}{\lambda^n}-\boldsymbol{F}_\alpha(x)\right|=0,$$

for any $R\in\mathbb{R}_+^*$. Hence, $F\equiv\boldsymbol{F}_\alpha$, and thus $\lambda=\boldsymbol{\lambda}_\alpha$. □

Whilst the above local convergence result is enough to identify asymptotically the profile $\boldsymbol{F}_\alpha$, a global result with quantitative convergence rates is still missing. In particular, note that the constants $C_{\varepsilon,R}$ above blow up when $R\to\infty$ as we see explicitly in Lemma 6.3. A second drawback of this we are unable to characterize the long-time behavior of the mass $\|F_n\|_{L^1(\mathbb{R}^n)}$ in terms of the eigenvalue $\boldsymbol{\lambda}_\alpha$ via this method. In the following section, we give an answer to both questions by better exploiting the previous fundamental Lemma 6.3 and using the propagation of quadratic and exponential moments in Section 4.2.

6.2. **Global convergence result.** We are now ready to prove our main result. Let us emphasize that our final convergence result in Theorem 1.1 is presented using the Kullback-Leibler divergence, which is a very different metric from the log-Lipschitz type metrics used in the previous Section 6.1 for the local convergence results. As anticipated in Remark 1.4, this decision does not obey aesthetic reasons only, but we actually need to move from uniform norms (like the log-Lipschitz norm) to averaged norms (like the Kullback-Leibler divergence) in order to address the deficiency encountered in Lemma 6.3. Specifically, recall that for generic initial data $F_0\in\mathcal{M}_+(\mathbb{R}^d)$, the log-Lipschitz norms of the high-dimensional integral $\mathcal{E}_n[\bar{F}_0]$ cannot be controlled uniformly due to additional exponentially growing terms.

Proof of Theorem 1.1.

• Step 1: Convergence of the profiles $F_n/\|F_n\|_{L^1(\mathbb{R}^d)}$.
Since $\boldsymbol{F}_\alpha$ is Gaussian (thus strongly log-concave), then the logarithmic-Sobolev inequality holds true and therefore we obtain the following control of the relative entropy by the relative Fisher information

$$\mathcal{D}_{\mathrm{KL}}\left(\frac{F_n}{\|F_n\|}\middle\|\boldsymbol{F}_\alpha\right)\leq\frac{\boldsymbol{\sigma}_\alpha^2}{2}\int_{\mathbb{R}^d}\left|\nabla\log\left(\frac{F_n(x)}{\boldsymbol{F}_\alpha(x)}\right)\right|^2\frac{F_n(x)}{\|F_n\|_{L^1(\mathbb{R}^d)}}\,dx,$$

see Corollary 5.7.2 and Section 9.3.1 in [2] for details. By the reformulation of F_n in formula (5.20) of Proposition 5.9 and using the notation in Definition 6.1 we have that

$$\nabla\log\left(\frac{F_n(x)}{\boldsymbol{F}_\alpha(x)}\right)=(k_n-\boldsymbol{k}_\alpha)\,x+\nabla\log\mathcal{E}_n[\bar{F}_0](x).$$

Therefore, we obtain the following upper bound

$$\mathcal{D}_{\mathrm{KL}}\left(\frac{F_n}{\|F_n\|}\middle\|\boldsymbol{F}_\alpha\right)\lesssim D_{n,1}+D_{n,2},\tag{6.32}$$

where each factor takes the form

$$D_{n,1}:=|k_n-\boldsymbol{k}_\alpha|^2\int_{\mathbb{R}^d}|x|^2\frac{F_n(x)}{\|F_n\|_{L^1(\mathbb{R}^d)}}\,dx,$$
$$D_{n,2}:=\int_{\mathbb{R}^d}\left|\nabla\log\mathcal{E}_n[\bar{F}_0](x)\right|^2\frac{F_n(x)}{\|F_n\|_{L^1(\mathbb{R}^d)}}\,dx.$$

On the one hand, for $D_{n,1}$ we can use the propagation of quadratic moments in Corollary 4.6 together with the explicit relaxation rates of $\{k_n\}_{n\in\mathbb{N}}$ to $\boldsymbol{k}_\alpha$ in Lemma 5.3 to show that

$$D_{n,1}\lesssim\boldsymbol{r}_\alpha^{2n}.\tag{6.33}$$

On the other hand, for $D_{n,2}$ we shall refine the argument in the proof of Corollary 6.5. Specifically, recall again that we can assume that $\bar{F}_0$ satisfies the hypothesis (6.5) without loss of generality (*cf.* proof of Corollary 6.5). Then, applying Lemma 6.3 in order to control $\nabla \log \mathcal{E}_n[\bar{F}_0]$ with fixed $\varepsilon \in \mathbb{R}_+^*$ yields

$$D_{n,2} \lesssim \left((2\,\boldsymbol{k}_\alpha)^2 + \varepsilon\right)^n \int_{\mathbb{R}^d} \left(1 + |x|^2 + e^{2\,C_\varepsilon (2\,\boldsymbol{r}_\alpha + \varepsilon)^n |x|^2}\right) \frac{F_n(x)}{\|F_n\|_{L^1(\mathbb{R}^d)}}\,dx,$$

for some sufficiently large $C_\varepsilon \in \mathbb{R}_+^*$. Above we have used that $(2\,\boldsymbol{k}_\alpha)^n$ is the decay rate that controls all the others in formula (6.6). Take $n \geq n_\varepsilon$ sufficiently large such that

$$\theta_\varepsilon := \sup_{n \geq n_\varepsilon} 2\,C_\varepsilon (2\,\boldsymbol{r}_\alpha + \varepsilon)^n < \frac{\alpha}{2}.$$

Then, Corollaries 4.6 and 4.8 imply that the above quadratic and exponential moments are uniformly bounded with respect to n for $n \geq n_\varepsilon$. Therefore, we have

$$D_{n,2} \lesssim \left((2\,\boldsymbol{k}_\alpha)^2 + \varepsilon\right)^n. \tag{6.34}$$

Finally, putting the estimates (6.33) and (6.34) into the split (6.32), and noticing that $\boldsymbol{r}_\alpha^2 < (2\,\boldsymbol{k}_\alpha)^2$ thanks to the relation $\boldsymbol{r}_\alpha = 2\,\boldsymbol{k}_\alpha^2$, ends this part of the proof.

• Step 2: Convergence of the growth rates $\|F_n\|_{L^1(\mathbb{R}^d)}/\|F_{n-1}\|_{L^1(\mathbb{R}^d)}$.
For simplicity of notation we shall define the following sequence of coefficients:

$$\lambda_n := \frac{\|F_n\|_{L^1(\mathbb{R}^d)}}{\|F_{n-1}\|_{L^1(\mathbb{R}^d)}}, \quad n \in \mathbb{N}.$$

Our goal then reduces to studying the asymptotic behavior of $\{\lambda_n\}_{n\in\mathbb{N}}$ and obtaining quantitative convergence rates. By definition of operator $\mathcal{T}$ in (1.2) we obtain that

$$\lambda_n = \int_{\mathbb{R}^d} e^{-m(x)} \int_{\mathbb{R}^{2d}} G\left(x - \frac{x_1 + x_2}{2}\right) \frac{F_{n-1}(x_1)}{\|F_{n-1}\|_{L^1(\mathbb{R}^d)}} \frac{F_{n-1}(x_2)}{\|F_{n-1}\|_{L^1(\mathbb{R}^d)}}\,dx_1\,dx_2\,dx.$$

By direct computation of the integral with respect to x as it was done in (4.11) we find that

$$\lambda_n = \int_{\mathbb{R}^{2d}} H(x_1, x_2) \frac{F_{n-1}(x_1)}{\|F_{n-1}\|_{L^1(\mathbb{R}^d)}} \frac{F_{n-1}(x_2)}{\|F_{n-1}\|_{L^1(\mathbb{R}^d)}}\,dx_1\,dx_2, \tag{6.35}$$

where the function $H(x_1, x_2)$ takes the form

$$H(x_1, x_2) := \frac{1}{(1+\alpha)^{d/2}} \exp\left(-\frac{\alpha}{(1+\alpha)} \left|\frac{x_1 + x_2}{2}\right|^2\right), \quad (x_1, x_2) \in \mathbb{R}^d. \tag{6.36}$$

In addition, recalling the explicit form of the eigen-pair $(\boldsymbol{\lambda}_\alpha, \boldsymbol{F}_\alpha)$ in (1.12) implies

$$\boldsymbol{\lambda}_\alpha = \int_{\mathbb{R}^{2d}} H(x_1, x_2)\,\boldsymbol{F}_\alpha(x_1)\,\boldsymbol{F}_\alpha(x_2)\,dx_1\,dx_2. \tag{6.37}$$

Therefore, taking the difference of (6.35) and (6.37) and noticing that H in (6.36) is bounded yields

$$|\lambda_n - \boldsymbol{\lambda}_\alpha| \lesssim \left\| \frac{F_n}{\|F_n\|_{L^1(\mathbb{R}^d)}} \otimes \frac{F_n}{\|F_n\|_{L^1(\mathbb{R}^d)}} - \boldsymbol{F}_\alpha \otimes \boldsymbol{F}_\alpha \right\|_{L^1(\mathbb{R}^{2d})}, \tag{6.38}$$

for any $n \in \mathbb{N}$, where $P \otimes Q$ denotes the Kronecker product of two measures $P, Q \in \mathcal{M}(\mathbb{R}^d)$, namely

$$\int_{\mathbb{R}^{2d}} \varphi(x,y)\,(P \otimes Q)(dx, dy) = \int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} \varphi(x,y)\,P(dx)\right) Q(dy),$$

for all $\varphi \in C_c(\mathbb{R}^d)$ (*cf.* Table 1). We do not have a direct convergence result of $F_n/\|F_n\|_{L^1(\mathbb{R}^d)}$ in L^1 norms, but we do have convergence of the Kullback-Leibler divergence thanks to Step 1. By Pinsker's inequality, the latter metric controls the former, namely

$$\begin{aligned}
&\left\| \frac{F_n}{\|F_n\|_{L^1(\mathbb{R}^d)}} \otimes \frac{F_n}{\|F_n\|_{L^1(\mathbb{R}^d)}} - \boldsymbol{F}_\alpha \otimes \boldsymbol{F}_\alpha \right\|_{L^1(\mathbb{R}^{2d})} \\
&\qquad \leq \frac{1}{\sqrt{2}} \sqrt{\mathcal{D}_{\mathrm{KL}}\left(\frac{F_n}{\|F_n\|_{L^1(\mathbb{R}^d)}} \otimes \frac{F_n}{\|F_n\|_{L^1(\mathbb{R}^d)}} \middle\| \boldsymbol{F}_\alpha \otimes \boldsymbol{F}_\alpha \right)} \\
&\qquad = \frac{1}{\sqrt{2}} \sqrt{2\,\mathcal{D}_{\mathrm{KL}}\left(\frac{F_n}{\|F_n\|_{L^1(\mathbb{R}^d)}} \otimes \middle\| \boldsymbol{F}_\alpha \right)},
\end{aligned} \tag{6.39}$$

where in last step we have used the tensorization property of the Kullback-Leilber divergence. The result then follows from (6.38)-(6.39) and the above explicit convergence rates of the normalized profiles $F_n/\|F_n\|_{L^1(\mathbb{R}^d)}$ in STEP 1. □

7. Numerical experiments

For our numerical simulation, we have restricted to one-dimensional traits (*i.e.*, $d = 1$) and we have considered a step function of the following form as initial datum:

$$F_0 = \frac{1}{Z}(30\,\mathbb{1}_{[-7,-3]} + 20\,\mathbb{1}_{[7.5,12.5]} + 50\,\mathbb{1}_{[30,40]} + 30\,\mathbb{1}_{[52,5,57,5]}), \tag{7.1}$$

where $Z \in \mathbb{R}_+^*$ is a normalizing factor so that $F_0 \in \mathcal{P}(\mathbb{R})$. Let $\{F_n\}_{n\in\mathbb{N}}$ be the solution to the time-discrete problem (1.1) starting at F_0. Our numerical simulation is performed with PYTHON on the finite computation domain $[-15, 60]$ for the trait variable x, which contains the support of the previous initial datum F_0 and all the mass of each F_n (except for a negligible Gaussian tail). In our simulation, we set $\Delta x = 0.001$ as the step for the discretization of our computational domain. In particular, with such a step we compute numerical integrals with respect to x according to the left rectangle rule.

In the following, we explore numerically two different scenarios: *weak selection* and *strong selection*. Specifically, we obtain numerical approximations for the profiles F_n in both regimes and we illustrate numerically the convergence of the growth rates $\|F_n\|_{L^1(\mathbb{R}^d)}/\|F_{n-1}\|_{L^1(\mathbb{R}^d)}$ towards the eigenvalue $\boldsymbol{\lambda}_\alpha$, and the relaxation of the normalized profiles $F_n/\|F_n\|_{L^1(\mathbb{R}^d)}$ towards the eigenfunction $\boldsymbol{F}_\alpha$. As a consequence, we derive numerical approximations of the convergence rates to be compared with the theoretical results in this paper. More specifically, we note that the theoretical convergence rates in Theorem 1.1 are sharp, except a mismatch for the rates of growth of mass, which was discussed previously in Remark 3.10 (see also Figure 6). Indeed, we actually attain numerically the same convergence rates as in Corollary 3.9, which were sharp for Gaussian initial data.

7.1. **Weak selection.** In this part, we consider a small value $\alpha = 0.015$. This leads to the following numerical values of the features of the equilibrium:

$$\boldsymbol{\lambda}_\alpha \approx 0.9857, \qquad \boldsymbol{\sigma}_\alpha^2 \approx 1.8897.$$

characterizing the eigenpair $(\boldsymbol{\lambda}_\alpha, \boldsymbol{F}_\alpha)$, according to (1.12) and (1.13). Note that since the selection parameter has been taken very small, then the eigenvalue and the variance of the eigenfunction are close to those at linkage equilibrium, namely, $\boldsymbol{\lambda}_{\alpha=0} = 1$ and $\boldsymbol{\sigma}^2_{\alpha=0} = 2$ (see Remark 3.5 and Figure 4).

In Figure 8 we observe the relaxation of the normalized profiles $F_n/\|F_n\|_{L^1(\mathbb{R}^d)}$ towards the eigenfunction $\boldsymbol{F}_\alpha$ along the time iterations $n = 0, 1, 2, 3, 4, 7, 150$. We remark on the strong contraction of the variance during the first few iterations leading to a well-identified Gaussian-shaped profile at time $n = 3$. At time $n = 7$ the variance of the profile is already close to that of the equilibrium $\boldsymbol{F}_\alpha$, but its mean is still substantially shifted to the right. Thereafter, we observe the gradual motion of the profiles towards the left at a much lower speed. After 150 iterations, the approximate mean and variance of $F_n/\|F_n\|_{L^1(\mathbb{R}^d)}$ are given by ≈ 0.0474 and ≈ 1.8897 respectively, where the latter agrees with the above exact value $\boldsymbol{\sigma}^2_\alpha$ up to 5 digits.

In Figure 9 we represent the convergence of the growth rates $\|F_n\|_{L^1(\mathbb{R}^d)}/\|F_{n-1}\|_{L^1(\mathbb{R}^d)}$ towards the eigenvalue $\boldsymbol{\lambda}_\alpha$. After 150 iterations we obtain that the growth rate takes the approximate value ≈ 0.9857, which again agrees with the exact value $\boldsymbol{\lambda}_\alpha$ above up to 5 digits.

In Figure 10 we have represented the errors

$$\varepsilon_n^{\mathrm{prof}} := \mathcal{D}_{\mathrm{KL}}\left(\frac{F_n}{\|F_n\|_{L^1(\mathbb{R}^d)}}\,\middle\|\,\boldsymbol{F}_\alpha\right), \quad \varepsilon_n^{\mathrm{mass}} := \left|\frac{\|F_n\|_{L^1(\mathbb{R}^d)}}{\|F_{n-1}\|_{L^1(\mathbb{R}^d)}} - \boldsymbol{\lambda}_\alpha\right|,$$

in semi-log plots so that the horizontal axis appears in the natural scale and represents each time iteration, and the vertical axis contains the logarithm of the errors. As observed, both plots reduce to essentially straight lines, suggesting exponential relaxation. Interestingly, the numerical rate of convergence in the Kullback-Leibler divergence measured in Figure 10a is approximately 0.9441. It coincides numerically with the rate of relaxation among the subclass of Gaussian solutions (see Corollary 3.9), namely, $\boldsymbol{\lambda}_\alpha^4 \approx 0.9441$, which in turns is identical to the theoretical rate $(2\,\boldsymbol{k}_\alpha)^2$ obtained in Theorem 1.1. These results are in perfect agreement with the other convergence results in Figure 10b: the numerical rate of convergence in the rate of growth of mass is approximately 0.9442, which is close to the one among the class of Gaussian solutions $\boldsymbol{\lambda}_\alpha^4 \approx 0.9441$, in contrast with the theoretical upper bound obtained in Theorem 1.1, namely $2\,\boldsymbol{k}_\alpha \approx 0.9717$. Moreover, the numerical rate of convergence of the variance of the normalized profiles $F_n/\|F_n\|_{L^1(\mathbb{R}^d)}$ is much faster: it is approximately $\approx 0.4721 < 0.5$, again close to

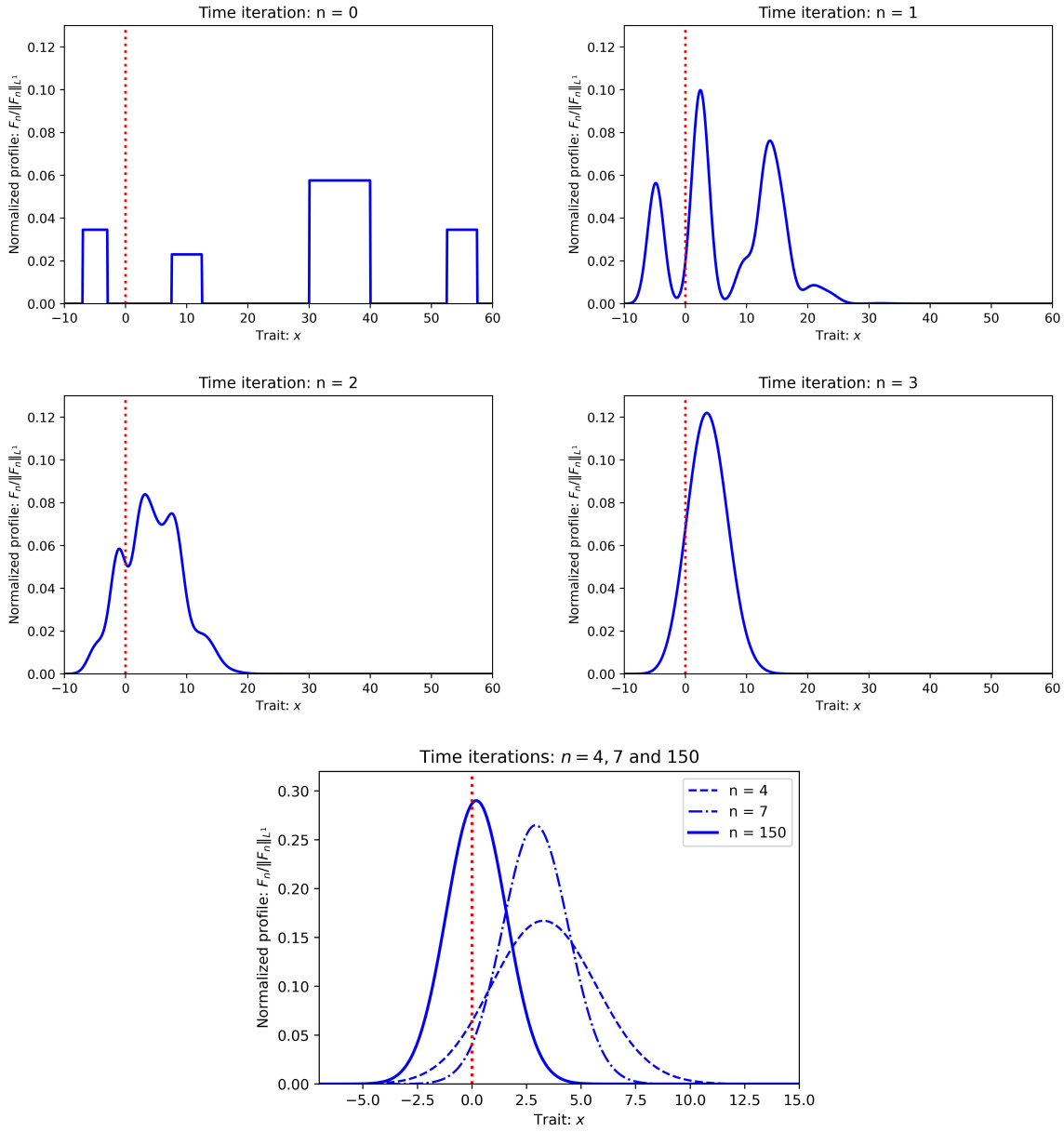


FIGURE 8. Relaxation of the normalized profiles $F_n/\|F_n\|_{L^1(\mathbb{R}^d)}$ towards the eigenfunction $\boldsymbol{F}_\alpha$ along the time iterations $n = 0, 1, 2, 3, 4, 7$ and 150 for a step function (7.1) as initial datum and weak selection parameter $\alpha = 0.015$. The vertical dotted line represents the location of the mean of the equilibrium profile $\boldsymbol{F}_\alpha$.

the expected value for Gaussian solutions, namely, $\boldsymbol{r}_\alpha \approx 0.472$. These numerical results illustrate the two-step process arising in the relaxation dynamics: convergence towards a Gaussian profile occurs much faster than relaxation of the mean towards the origin due to (weak) selection. Alternatively speaking, the equilibrium variance builds up much faster than the center of the distribution gets to the origin.

7.2. **Strong selection.** In this part, we consider a larger value of α. We opted for $\alpha = 0.4$. Indeed, if α is taken too large then there is no visual difference between the solution and the equilibrium configuration after one single iteration. This time, the numerical values of the features of the equilibrium are:

$$\boldsymbol{\lambda}_\alpha \approx 0.7944, \quad \boldsymbol{\sigma}_\alpha^2 \approx 0.9221.$$

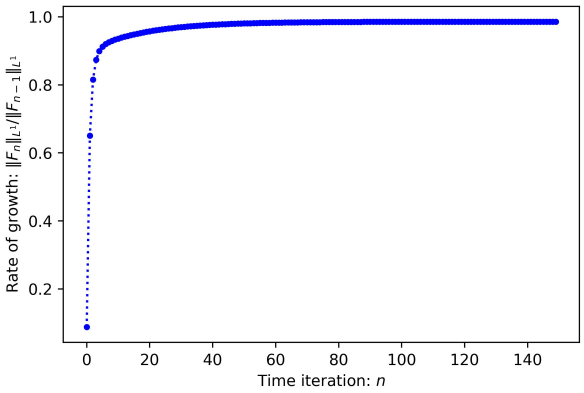
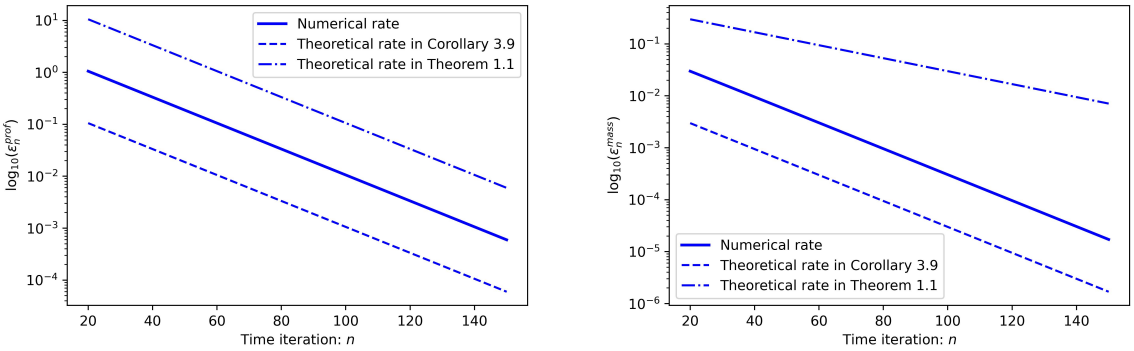


FIGURE 9. Relaxation of the growth rates $\|F_n\|_{L^1(\mathbb{R}^d)}/\|F_{n-1}\|_{L^1(\mathbb{R}^d)}$ towards the eigenvalue $\boldsymbol{\lambda}_\alpha$ along time iterations $0 \leq n \leq 150$ for a step function (7.1) as initial datum and weak selection parameter $\alpha = 0.015$.



(a) (b)

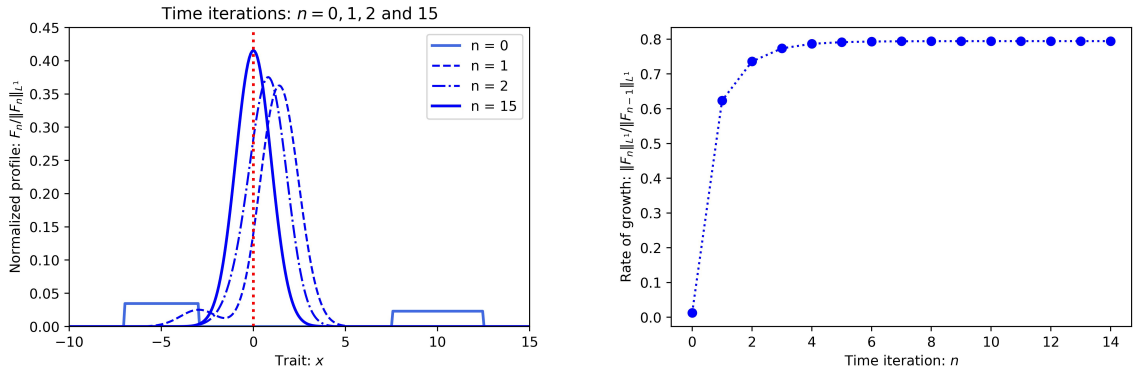
FIGURE 10. Numerical computation of the rates of convergence of the normalized profiles and the growth rates: (a) semi-log plot of the errors $\varepsilon_n^{\mathrm{prof}} := \mathcal{D}_{\mathrm{KL}}\left(F_n/\|F_n\|_{L^1} \,\|\, \boldsymbol{F}_\alpha\right)$, and (b) semi-log plot of the errors $\varepsilon_n^{\mathrm{mass}} := \left|\|F_n\|_{L^1}/\|F_{n-1}\|_{L^1} - \boldsymbol{\lambda}_\alpha\right|$.

The latter substantially differs from the value at linkage equilibrium (*i.e.*, $\boldsymbol{\sigma}^2_{\alpha=0} = 2$). In Figure 11a we note that the solution resembles a Gaussian distribution after a couple of iterations. After 15 iterations we obtain that the normalized profile $F_n/\|F_n\|_{L^1(\mathbb{R}^d)}$ has approximate mean and variance respectively given by ≈ 0.0017 and ≈ 0.9221. In Figure 11b we observe the convergence of the growth rates $\|F_n\|_{L^1(\mathbb{R}^d)}/\|F_{n-1}\|_{L^1(\mathbb{R}^d)}$ towards the eigenvalue $\boldsymbol{\lambda}_\alpha$. After 15 iterations the growth rate becomes ≈ 0.7944, which again agrees with the exact value $\boldsymbol{\lambda}_\alpha$ above up to 5 digits.

Similar computations as in the previous Figure 10 allow finding numerical approximations for the rate of convergence of the normalized profiles and the growth rates. Specifically, we obtain an approximation 0.3966 for the rate of convergence of the Kullback-Leibler divergence. Once again, such a value is close to the rate of relaxation among the subclass of Gaussian solutions, namely, $\boldsymbol{\lambda}^4_\alpha \approx 0.3983$, which in turns agrees with the theoretical rate $(2\,\boldsymbol{k}_\alpha)^2$ obtained in Theorem 1.1. Similarly, the numerical rate of convergence of the growth rates is approximately 0.3901, which is close to the one among the class of Gaussian solutions $\boldsymbol{\lambda}^4_\alpha \approx 0.3983$, in contrast with the upper bound $2\,\boldsymbol{k}_\alpha \approx 0.6311$ in Theorem 1.1.

8. Conclusions and perspectives

In this paper, we have proven asynchronous exponential growth in a quantitative genetics model for the evolution of the distribution of traits in a population governed by sexual reproduction and multiplicative effect of selection. Our model assumes time-discrete non-overlapping generations, which rule out an eventual mixing with previous generations of ancestors. In addition, our non-linear sexual reproduction operator is set in agreement with Fisher's infinitesimal model, and we have chosen selection to act on



(a) (b)

FIGURE 11. (a) Zoom near the origin of the relaxation of the normalized profiles $F_n/\|F_n\|_{L^1(\mathbb{R}^d)}$ towards the eigenfunction $\boldsymbol{F}_\alpha$ along the time iterations $n = 0, 1, 2$ and 15 for the step function (7.1) as initial datum and strong selection parameter $\alpha = 0.4$. The vertical dotted line represents the location of the mean of the equilibrium profile $\boldsymbol{F}_\alpha$. (b) Relaxation of the growth rates $\|F_n\|_{L^1(\mathbb{R}^d)}/\|F_{n-1}\|_{L^1(\mathbb{R}^d)}$ towards the eigenvalue $\boldsymbol{\lambda}_\alpha$.

the survival probability of individuals. Our main result provides quantitative convergence rates of the renormalized distributions towards a unique stationary profile. Indeed, rates are exponential, which can be interpreted loosely as a spectral gap in this non-linear context.

It is noticeable that the sexual reproduction operator is contractive under the Wasserstein distance (see Lemma 2.1). However, the generic incompatibility of multiplicative selection with transport distances becomes an apparent obstruction to the use of a direct perturbative approach in the regime of weak selection (see Example 2.4). This obstacle was circumvented by G. RAOUL [21], assuming that selection is restricted to a compact support. We follow a different route in a purely non-perturbative setting. To this end, we restrict to the specific choice of a quadratic selection function of the trait, for the sake of simplicity. Our alternative approach relies on an appropriate study of the propagation of information along large binary trees of ancestors, by a suitable reformulation of high-dimensional integrals, inspired by the changes of variables performed in [10] in the regime of small variance. This reveals an ergodicity property where the exact shape of the initial distribution is quickly forgotten across generations.

We remark that the above heuristic arguments indicate that the quadratic Wasserstein metric is not fully appropriate for dealing with the problem at hand. However, we could not identify yet a proper metric that extends the contraction property of the neutral case to the quadratic selection case.

Several perspectives are envisaged. First, there is an apparent price to pay with our method, in which we fully exploit the Gaussian structure induced by quadratic selection in order to perform tractable computations within the high-dimensional integrals. We believe though that the restriction to quadratic selection might be overcome in future works to allow for more general selection functions. Second, as studied in [10], the case of multiple minima on the selection function leads to non-uniqueness of stable equilibria. Then, uncovering the hidden metastability and quantifying the relaxation towards a specific equilibrium is of great interest for its applications in quantitative genetics. Finally, a major problem is to transcend non-overlapping generations and tackle the full time-continuous model as presented in previous literature.

APPENDIX A. NONDIMENSIONALIZATION AND DERIVATION OF THE TIME-DISCRETE VERSION

In this section, we shall nondimensionalize the time-continuous model (1.6)-(1.3) with Gaussian mixing kernel G and quadratic selection function m. By time discretization on the Duhamel formulation, we will derive the time-discrete version (1.1) which has been central in this paper. Indeed, we shall show that we can reduce parameters into only one, namely, parameter $\alpha \in \mathbb{R}_+$ in the quadratic selection function (1.5). Bearing in mind all the biological parameters and dimensions, our evolution problem reads:

$$\begin{cases} \partial_t f = -r_m\, m(x) f + r_b\, \mathcal{B}_{\sigma^2}[f], & t \geq 0,\ x \in \mathbb{R}^d \\ f(0,x) = f_0(x), & x \in \mathbb{R}^d. \end{cases} \tag{A.1}$$

Here, $r_m, r_b \in \mathbb{R}_+$ represent the mortality and birth rates and have frequency units, whilst the dimensionless corrections of m and $\mathcal{B}$ take the form

$$m(x) = \frac{|x|^2}{\sigma_m^2}, \qquad \mathcal{B}_{\sigma^2}[f](t,x) = \int_{\mathbb{R}^{2d}} G_{0,\sigma^2}\left(x - \frac{x_1+x_2}{2}\right) \frac{f(t,x_1)f(t,x_2)}{\int_{\mathbb{R}^d} f(t,x')\,dx'}\,dx_1\,dx_2,$$

for each $x \in \mathbb{R}^d$. Here, G_{0,σ^2} represents the Gaussian centered at 0 with variance σ^2, and σ_m and σ have the same units as the quantitative trait. On the one hand, σ_m can be regarded as a characteristic unit quantifying the effective range of selection. Namely, if $|x| \geq \sigma_m$ then $m(x) \geq 1$. On the other hand, σ^2 is the genetic variance and $2\sigma^2$ represents the variance at linkage equilibrium often denoted by $\sigma_{LE}^2 := 2\sigma^2$ (*cf.* [9]). We nondimensionalize our system by appropriately scaling our variables as follows

$$\widehat{t} = \frac{t}{T}, \quad \widehat{x} = \frac{x}{L}, \quad \widehat{f}(\widehat{t},\widehat{x}) = L^d f(t,x).$$

Here, T and L are characteristic units for time and trait. For simplicity, we set them as follows in terms of the parameters of the system:

$$T := \frac{1}{r_m}, \quad L := \sigma.$$

In other words, we scale time according to the mortality rate r_m and the trait variable according to the genetic variance σ^2. Dropping hats for simplicity yields the following dimensionless form of the equation

$$\begin{cases} \partial_t f = -\frac{\alpha}{2}|x|^2 f + \beta\mathcal{B}[f], & t \geq 0,\ x \in \mathbb{R}^d, \\ f(0,x) = f_0(x), & x \in \mathbb{R}^d, \end{cases} \tag{A.2}$$

where $\mathcal{B} = \mathcal{B}_1$ involves again unit genetic variance thanks to our scaling assumption. This reduces the amount of free parameters to only two of them $\alpha, \beta \in \mathbb{R}_+$ defined as follows:

$$\alpha := \frac{\sigma_{LE}^2}{\sigma_m^2}, \quad \beta := \frac{r_b}{r_m}.$$

In the sequel, we justify our time-discrete version (1.11). Specifically, let us integrate equation $(A.2)_1$ for $t \in [t_n, t_{n+1}]$ where $\{t_n\}_{n\in\mathbb{N}} \subseteq \mathbb{R}_+$ is any increasing sequence. Then, we recover Duhamel's formula

$$f(t_n,x) = e^{-\frac{\alpha}{2}|x|^2(t_n - t_{n-1})} f(t_{n-1},x) + \beta \int_{t_{n-1}}^{t_n} e^{-\frac{\alpha}{2}|x|^2(t_n - s)} \mathcal{B}[f(s,\cdot)](x)\,ds,$$

for any $n \in \mathbb{N}$ and each $x \in \mathbb{R}^d$. Using the rectangle rule for the integral in the right hand side leads to the approximations $F_n(x) \simeq f(t_n,x)$ of the the above time-continuous problem for $f = f(t,x)$:

$$F_n(x) = e^{-\frac{\alpha}{2}|x|^2} F_{n-1}(x) + \beta e^{-\frac{\alpha}{2}|x|^2} \mathcal{B}[F_{n-1}](x), \tag{A.3}$$

for any $n \in \mathbb{N}$ and each $x \in \mathbb{R}^d$. Here, we have set unitary time steps $t_n - t_{n-1} = 1$ for simplicity and $t_0 = 0$. We emphasize that the right hand side of (A.3) consists of two different terms. On the one hand, the first term describes the amount of old individuals from the previously generation $n-1$ having resisted the effect of selection in generation n. On the other hand, the second term computes the amount of offspring conceived by individuals in the previous generation $n-1$, where an eventual decline due to selection has also been taken into account. Let us emphasize that in (A.3) individuals from generation $n-1$ can merge (after breeding) with their offspring to form next generation n. This is typical in most biological populations, which have *overlapping generations* (*e.g.*, some multivoltine flies like Drosophila melanogaster). However, some other systems have *non-overlapping generations* so that the full generation $n-1$ gets extinct after breeding and only their offspring survive at generation n (*e.g.*, some univoltine insects like the Dawson's burrowing bee). We refer to [17] for a discussion on the various types of voltinism and the role of diapause. In that case, by canceling the first term (A.3) reduces to

$$F_n(x) = \beta e^{-\frac{\alpha}{2}|x|^2} \mathcal{B}[F_{n-1}](x), \tag{A.4}$$

for any $n \in \mathbb{N}$ and each $x \in \mathbb{R}$. We remark that two different parameters α and β are still present in (A.4). However, given that $F \mapsto e^{-\frac{\alpha}{2}|x|^2}\mathcal{B}[F]$ is a 1-homogeneous operator we can kill the parameter β by rescaling the trait distributions F_n. Specifically, define $\widehat{F}_n = \beta^{-n} F_n$ and note that (A.4) reduces to our time-discrete problem (1.1) with only one parameter α.

References

1. L. Ambrosio, N. Gigli, and G. Savaré, *Gradient flows in metric spaces and in the space of probability measures*, Birkhäuser, Basel, 2008.
2. D. Bakry, I. Gentil, and M. Ledoux, *Analysis and Geometry of Markov Diffusion Operators*, A Series of Comprehensive Studies in Mathematics, vol. 348, Springer, Cham, 2014.
3. G. Barles and S. Mirrahimi, *Concentrations and constrained Hamilton-Jacobi equations arising in adaptive dynamics*, Recent Developments in Nonlinear Partial Differential Equations (D. Danielli, ed.), Contemp. Math, vol. 439, American Mathematical Society, Providence, RI, 2007, pp. 57–68.
4. G. Barles, S. Mirrahimi, and B. Perthame, *Concentration in Lotka-Volterra parabolic or integral equations: a general convergence result*, Methods App. Anal. **16** (2009), no. 3, 321–340.
5. N. H. Barton, A. M. Etheridge, and A. Véber, *The infinitesimal model: Definition, derivation, and implications*, Theoret. Population Biol. **118** (2017), 50–73.
6. H. Berestycki, J. Coville, and H.-H. Vo, *Persistence criteria for populations with non-local dispersion*, J. Math. Biol. **72** (2016), 1693–1745.
7. F. Bolley, I. Gentil, and A. Guillin, *Convergence to equilibrium in Wasserstein distance for Fokker-Planck equations*, J. Funct. Anal. **263** (2012), no. 8, 2430–2457.
8. T. Bourgeron, V. Calvez, J. Garnier, and L. Lepoutre, *Existence of recombination-selection equilibria for sexual populations*, 2017, arXiv:1703.09078.
9. M. G. Bulmer, *The mathematical theory of quantitative genetics*, Clarendon Press, Oxford, 1980.
10. V. Calvez, J. Garnier, and F. Patout, *Asymptotic analysis of a quantitative genetics model with nonlinear integral operator*, J. Éc. polytech. Math. **6** (2019), 537–579.
11. V. Calvez and K.-Y. Lam, *Uniqueness of the viscosity solution of a constrained Hamilton-Jacobi equation*, Calc. Var. **59** (2020), no. 163.
12. V. Calvez, D. Poyato, and F. Santambrogio, *Uniform contractivity of the Fisher infinitesimal model with strongly convex selection*, 2023, arXiv:2302.12063.
13. J. Coville, F. Li, and X. Wang, *On eigenvalue problems arising from nonlocal diffusion models*, Discrete Contin. Dyn. Syst. Ser. A **37** (2017), no. 2, 879–903.
14. O. Diekmann, P.-E. Jabin, S. Mischler, and B. Perthame, *The dynamics of adaptation: an ilumination example and a Hamilton-Jacobi approach*, Theoret. Population Biol. **67** (2005), no. 4, 25—271.
15. R. A. Fisher, *The correlation between relatives on the supposition of mendelian inheritance*, Trans. Roy. Soc. Edinburgh **52** (1918), 399–433.
16. R. Mahadevan, *A note on a non-linear Krein-Rutman theorem*, Nonlinear Anal. Theory Methods Appl. **67** (2007), no. 11, 3084–3090.
17. S. Masaki, *Geographic variation of diapause in insects*, Bull. Fac. Agric. Hirosaki Univ. **7** (1961), 66–98.
18. S. Mirrahimi and G. Raoul, *Dynamics of sexual populations structured by a space variable and a phenotypical trait*, Theoret. Population Biol. **84** (2013), 87–103.
19. F. Patout, *The cauchy problem for the infinitesimal model in the regime of small variance*, Anal. PDE **16** (2023), no. 6, 1289–1350.
20. G. Raoul, *Macroscopic limit from a structured population model to the Kirkpatrick-Barton model*, 2017, arXiv:1706.04094.
21. ———, *Exponential convergence to a steady-state for a population genetics model with sexual reproduction and selection*, 2021, arXiv:2104.06089.
22. K. Rosen, *Discrete Mathematics and Its Applications*, McGraw-Hill Science, New York, 2011.
23. M. Talagrand, *Transportation cost for Gaussian and other measures*, Geom. Funct. Anal. **6** (1996), 587–600.
24. C. Villani, *Optimal transport: old and new*, Grundlehren der mathematischen Wissenschaften, vol. 338, Springer, Berlin, Heidelberg, 2009.
25. S. G. Williamson, *Combinatorics for computer science*, Computer Science Press, Rockville, Maryland, 1985.

Institut Camille Jordan (ICJ), UMR 5208 CNRS & Université Claude Bernard Lyon 1, 69100 Villeurbanne, France

Email address: vincent.calvez@math.cnrs.fr

Univ Lyon, Inria, Université Claude Bernard Lyon 1, CNRS UMR5208, Institut Camille Jordan, F-69603 Villeurbanne, France

Email address: thomas.lepoutre@inria.fr

Departamento de Matemática Aplicada and Research Unit "Modeling Nature" (MNat), Facultad de Ciencias, Universidad de Granada, 18071 Granada, Spain

Email address: davidpoyato@ugr.es