A moment-based approach for the analysis of the infinitesimal model in the regime of small variance

J. Guerand[*]   M. Hillairet[†]   S. Mirrahimi[‡]

January 25, 2024

Abstract

We provide an asymptotic analysis of a nonlinear integro-differential equation describing the evolutionary dynamics of a population which reproduces sexually and which is subject to selection and competition. The sexual reproduction is modeled via a nonlinear integral term, known as the "infinitesimal model". We consider a regime of small segregational variance, where a parameter in the infinitesimal operator, which measures the deviation between the trait of the offspring and the mean parental trait, is small. We prove that, in this regime, the phenotypic distribution remains close to a Gaussian profile with a fixed small variance and we characterize the dynamics of the mean phenotypic trait via an ordinary differential equation. While similar properties were already proved for a closely related model using a Hopf-Cole transformation and perturbative analysis techniques, we provide an alternative proof which simplifies considerably the analysis. Our method relies on a direct study of the dynamics of the moments of the phenotypic distribution and a contraction property of the Wasserstein distance. Such a simplification is mandatory to tackle analytically more comprehensive models with for instance a spatial structure.

Keywords : Infinitesimal model, structured population, quantitative genetics, integro-differential equations, singular limits.
2020 M.S.C. : 35B40, 35Q92, 92D15, 47G20.

1 Introduction

1.1 Model and question

The purpose of this article is to provide an asymptotic analysis for small ε of the following equation

$$\begin{cases} \varepsilon^2 \partial_t n_\varepsilon = r T_\varepsilon[n_\varepsilon] - (m + \kappa \rho_\varepsilon(t)) \, n_\varepsilon, \\ \rho_\varepsilon(t) = \int_{\mathbb{R}} n_\varepsilon(t, y) dy, \\ n_\varepsilon(0, x) = n_{\varepsilon,0}(x), \end{cases} \tag{1}$$

where

$$T_\varepsilon[n_\varepsilon](x) = \int_{\mathbb{R}} \int_{\mathbb{R}} \Gamma_\varepsilon \left(x - \frac{(y + y')}{2} \right) n_\varepsilon(t, y) \frac{n_\varepsilon(t, y')}{\rho_\varepsilon(t)} dy dy',$$

---

[*]Institut Montpelliérain Alexander Grothendieck, Univ. Montpellier, CNRS, Montpellier, France; E-mail: `jessica.guerand@umontpellier.fr`

[†]Institut Montpelliérain Alexander Grothendieck, Univ. Montpellier, CNRS, Montpellier, France; E-mail: `matthieu.hillairet@umontpellier.fr`

[‡]Corresponding author: Sepideh Mirrahimi, Institut Montpelliérain Alexander Grothendieck, Univ. Montpellier, CNRS, Montpellier, France; E-mail: `sepideh.mirrahimi@umontpellier.fr`

$$\Gamma_\varepsilon(x) = \frac{1}{\varepsilon\sqrt{\pi}} \exp\left(-\frac{x^2}{\varepsilon^2}\right).$$

This equation describes the evolutionary dynamics of a phenotypically structured population, which reproduces sexually and which is subject to selection and competition. The function $n_\varepsilon(t,x)$ stands for the density of a population characterized by a phenotypic trait x and $m(x)$ represents a trait-dependent mortality rate. The population is also subject to a mortality rate due to uniform competition between individuals, represented by the term $\kappa\rho_\varepsilon$, where ρ_ε is the total size of the population. The parameter $r > 0$ scales reproduction in the population.

The term T_ε describes the sexual reproduction and is based on the assumption that the trait of an offspring has a normal distribution with a fixed variance, centered around the mean parental trait. This reproduction model, which is usually referred to as the "infinitesimal model", was first suggested by R. A. Fisher in [11] and it has been widely used in the biological literature [3, 18, 31]. Its validity is under the assumption that the trait x is coded by a large number of alleles which have small and additive effects (see [2] for a recent rigorous derivation). The parameter ε has been introduced in the problem above to consider a small segregational variance, that is the deviation between the trait of an offspring from the mean parental trait is small. This assumption is closely related to the so-called "weak selection regime" in Quantitative Genetics [4, 31].

We will prove that, under appropriate assumptions and for ε small, the phenotypic density n_ε remains close to a Gaussian distribution with variance ε^2 and with a mean value which varies according to an ordinary differential equation. Note that in many biological articles, when considering the infinitesimal model, it is assumed that the phenotypic distribution is a Gaussian with constant variance (see for instance [16, 14, 30, 15]). Our result justifies this assumption, considering a homogeneous environment and a small segregational variance. Such a result, considering a closely related model without a competition term and with different assumptions, was already provided in [24] using a perturbative analysis via a Hopf-Cole transformation, that is

$$n_\varepsilon(t,x) = \frac{1}{\sqrt{2\pi}\varepsilon} \exp\left(\frac{u_\varepsilon(t,x)}{\varepsilon^2}\right).$$

Here, we provide an alternative proof and we show that a rather straightforward analysis of the moments, together with a contraction property of the Wasserstein distance, leads to the result. We believe that this approach could be more easily adapted to study more complex models considering spatial or temporal heterogeneities of the environment as in [8].

1.2 Assumptions and notations

We assume that the mortality rate $m \in C^\infty(\mathbb{R})$ enjoys the following properties for some $L > 0$:

(H0) $m(x) \geq 0$ with $\inf_{x \in \mathbb{R}} m(x) = 0$,

(H1) $m'(0) = 0$ and $A_0 \leq m''(x)$ for all $x \in (-L, L)$, for some constant $A_0 > 0$,

(H2) $\max_{x \in [-L,L]} m(x) < r$,

(H3) $|m''(x)| \leq A_m(1 + |x|^p)$ for all $x \in \mathbb{R}$ and some exponent $p \in \mathbb{N}^*$ and constant $A_m \in (0, \infty)$.

In assumption (H0), we fix that the infimum value of m is 0. Indeed, in (1) we can always substract a constant to m up to multiply n_ε with a suitable time-exponential. Assumption (H1) states that the trait $x = 0$ is a non-degenerate local minimum point of the mortality rate m. Assumption (H2) ensures that the value of m around this minimum point is not too far from its global minimum. Assumption (H3) imposes a technical limitation on the growth at infinity. These assumptions entail

in particular that:

$$
\begin{cases}
A_0|x| \le |m'(x)|, & A_0\dfrac{|x|^2}{2} \le m(x)-m(0), & \forall\, x \in (-L, L), \\[2mm]
|m'(x)| \le A_m\left(|x| + \dfrac{|x|^{p+1}}{p+1}\right), & m(x)-m(0) \le A_m\left(\dfrac{|x|^2}{2} + \dfrac{|x|^{p+2}}{(p+1)(p+2)}\right), & \forall\, x \in \mathbb{R}.
\end{cases}
\tag{2}
$$

We comment on the global-in-time well-posedness of (1) under assumption (H0)-(H3) below. Given a solution n_ε to (1), let's denote

$$
q_\varepsilon(t, x) = \frac{n_\varepsilon(t, x)}{\rho_\varepsilon(t)}.
\tag{3}
$$

One can verify that q_ε solves the following equation

$$
\varepsilon^2 \partial_t q_\varepsilon = r\left(\widetilde{T}_\varepsilon[q_\varepsilon] - q_\varepsilon\right) - \left(m - \int_{\mathbb{R}} m q_\varepsilon dx\right) q_\varepsilon,
\tag{4}
$$

with

$$
\widetilde{T}_\varepsilon[q_\varepsilon](x) = \int_{\mathbb{R}} \int_{\mathbb{R}} \Gamma_\varepsilon\left(x - \frac{(y + y')}{2}\right) q_\varepsilon(t, y) q_\varepsilon(t, y') dy\, dy'.
$$

We remark that (4) preserves probability densities. We next introduce the following notations corresponding to the moments of the phenotypic distribution:

$$
M_{\varepsilon,1}(t) = \int_{\mathbb{R}} x q_\varepsilon(t, x) \mathrm{d}x,
$$
$$
M^c_{\varepsilon,k}(t) = \int_{\mathbb{R}} (x - M_{\varepsilon,1}(t))^k q_\varepsilon(t, x) \mathrm{d}x, \quad M^{|c|}_{\varepsilon,k}(t) = \int_{\mathbb{R}} |x - M_{\varepsilon,1}(t)|^k q_\varepsilon(t, x) \mathrm{d}x \quad \forall\, k \in \mathbb{N}.
\tag{5}
$$

We remark that the above convention entails in particular $M^c_{\varepsilon,0} = 1$ and $M^c_{\varepsilon,1} = 0$ in the cases $k = 0, 1$. We finally define

$$
g_\varepsilon(t, x) = \frac{1}{\sqrt{2\pi}\varepsilon} \exp\left(-\frac{(x - \bar{Z}_\varepsilon(t))^2}{2\varepsilon^2}\right)
\tag{6}
$$

with $\bar{Z}_\varepsilon$ the solution to:

$$
\begin{cases}
\dot{\bar{Z}}_\varepsilon(t) = -m'(\bar{Z}_\varepsilon(t)), \\
\bar{Z}_\varepsilon(0) = M_{\varepsilon,1}(0).
\end{cases}
\tag{7}
$$

We will make the following assumptions guaranteeing a well-prepared initial condition $q_{\varepsilon,0}$ (or equivalently $n_{\varepsilon,0}$):

$$
M_{\varepsilon,1}(0) \in (-L, L), \qquad \varepsilon^{-2k_0} M^c_{\varepsilon,2k_0}(0) \le C_1,
\tag{H4}
$$

with $C_1 > 0$ fixed and k_0 large enough such that

$$
k_0 > 3 + \lceil p/2 \rceil, \qquad r\left(1 - \frac{2}{4^{k_0}}\right) - \max_{x \in [-L, L]} m(x) =: \eta > 0.
\tag{8}
$$

Both parameters C_1 and k_0 shall be chosen independent of ε. To address the limiting behavior when $\varepsilon \to 0$ we also assume that

$$
M_{\varepsilon,1}(0) = x_0 + O(\varepsilon), \qquad x_0 \in (-L, L),
\tag{H5}
$$

$$
\rho_m \le \rho_\varepsilon(0) \le \rho_M, \qquad \text{for some positive constants } \rho_m \text{ and } \rho_M.
\tag{H6}
$$

We point out that Assumption (H5) implies formally that, for all $t > 0$ and as $\varepsilon \to 0$, $\bar{Z}_\varepsilon(t) \to \bar{Z}(t)$, with

$$
\begin{cases}
\dot{\bar{Z}}(t) = -m'(\bar{Z}(t)), \\
\bar{Z}(0) = x_0.
\end{cases}
\tag{9}
$$

In Section 4, we give a quantitative description of this convergence.

1.3 Main results

We first provide a preliminary result guaranteeing the well-posedness of the Cauchy problem.

Proposition 1 *Let $\ell \in \mathbb{N}^*$, $\ell \geq 2$, and assume that m satisfies (H0) and (H3) and $n_0 \in L^1(\mathbb{R}) \setminus \{0\}$ is nonnegative and has more than 2ℓ bounded moments. Given $\varepsilon > 0$, there exists a unique global solution n_ε to equation (1) with 2ℓ bounded moments. This solution is nonnegative and all moments of n_ε of order less than 2ℓ are well defined and continuous on $[0, \infty)$.*

In the case where m is considered bounded, the proposition above is a direct consequence of the Cauchy-Lipschitz Theorem. In the case of unbounded m the result can be obtained via a standard truncation procedure. We provide the main elements of the proof in Appendix A in which we make precise the definition of solution. We show in particular that the j-th moment of n is also C^1 in time if $j < 2\ell - (p + 3)$.

From now on we assume that n_ε is a solution to (1) with an arbitrary large number of bounded moments. We remark that q_ε is then a solution to (4) with the same regularity/integrability. We next state our main result.

Theorem 2 *Assume (H1)–(H4).*
(i) Let $\delta \in (0, 1)$. There exists a constant K depending on the initial condition, r and m such that, for any ε sufficiently small, there holds

$$\sup_{[0, +\infty)} W_1(q_\varepsilon(t, \cdot), g_\varepsilon(t, \cdot)) \leq K \left(W_1(q_{\varepsilon,0}, g_{\varepsilon,0}) + \varepsilon^{1-\delta} \right). \tag{10}$$

with $q_{\varepsilon,0}$ and $g_{\varepsilon,0}$ the respective values at time $t = 0$ of q_ε and g_ε.
(ii) Assume additionally (H5)–(H6) and let $\delta \in (0, 1)$, and $\beta \in (1, 2)$. Then, there exists a constant K_ρ such that, for any ε sufficiently small, we have

$$|\rho_\varepsilon(t) - \rho(t)| \leq K_\rho \varepsilon^{1-\delta}, \qquad \rho(t) = \frac{r - m(\bar{Z}(t))}{\kappa}, \quad \text{for all } t \in [\varepsilon^\beta, +\infty) \tag{11}$$

where $\bar{Z}$ solves (9). Moreover, as $\varepsilon \to 0$, n_ε converges in $C\left((0, \infty); \mathcal{M}^1(\mathbb{R})\right)$ to a measure n, which is given by

$$n(t, x) = \rho(t) \partial(x - \bar{Z}(t)).$$

Here we denote $\partial(x - X_0)$ the Dirac mass centered in X_0. We use the (non-standard) symbol ∂ in order to avoid confusion with the parameter δ. In this statement, we denote also W_1 the wasserstein distance for the L^1 distance. The proof of the theorem above relies on some estimates on the moments of the phenotypic distribution and a contraction property of the Wasserstein distance satisfied by the reproduction operator $\widetilde{T}_\varepsilon$. We state below our estimates on the moments of the phenotypic distribution.

Theorem 3 *Assume (H1)–(H4) and $\delta \in (0, 1)$. For any ε sufficiently small, there exist large enough positive constants K_0, K_1 and K_2 depending on the parameters C_1, r, L such that:*

$$M_{\varepsilon, 2k_0}^c(t) \leq K_2 \varepsilon^{2k_0}, \tag{12}$$

$$\left| M_{\varepsilon, 2}^c(t) - \varepsilon^2 \right| \leq K_1 \varepsilon^2 \left(\varepsilon^{1-\delta} + \exp(-rt/2\varepsilon^2) \right), \tag{13}$$

$$\left| M_{\varepsilon, 1}(t) - \bar{Z}_\varepsilon(t) \right| \leq K_0 \varepsilon^{1-\delta}. \tag{14}$$

for any $t \in \mathbb{R}^+$.

We emphasize that the parameters K_0, K_1, K_2 are independent of ε, δ. In the course of the proof, we shall choose first K_1 depending on data C_1 (see assumption (H4)), then K_2 depending on K_1, C_1 and finally K_0 depending on K_1, L. We restrict ε according to K_2, K_0 and L. The estimates on the moments are obtained via direct computations of the corresponding equations and using the

Taylor expansion of the mortality rate m around the mean phenotypic trait. The main ingredients are the presence of some dissipation terms, due to an algebra satisfied by T_ε, and the fact that the central moments of the phenotypic distribution have simple approximations when the segregational variance ε is small.

One may wonder whether the method suggested in this article could be adapted to study models with an asexual reproduction term. In Section 5 we will explain why the dissipation terms are lacking in this latter case and our method cannot be adapted to such models straightforwardly.

1.4 State of the art

The study of integro-differential equations involving the infinitesimal model has gained increasing attention recently from the mathematical community. Several works have used the contraction property of the Wasserstein distance to study models which involve the infinitesimal operator [29, 28] (see also [19] for a related work on a model of protein exchanges in a cell population). In [29], G. Raoul studied a model describing the evolutionary dynamics of a population in an environment with continuous spatial structure, under a small selection assumption. He justified the Gaussian assumption and the so-called Kirkpatrick and Barton model, *i.e.* a system of equations describing the dynamics of the size of the population and its mean phenotypic trait, as functions of time and space variables [14] (see also [20] for an earlier formal derivation of this model). Later in [28], G. Raoul also studied a model with homogeneous environment, still in a small selection regime, and obtained that the solution of this problem remains close to a Gaussian distribution with fixed variance, while the mean of the distribution varies according to an ordinary differential equation similar to (9). He also showed that in long time the solution converges to the steady solution of the problem. This work is different from ours in several ways. First, in [28] the selection is considered to act on the reproduction and not the mortality term. In other words the mortality rate m is supposed to be constant, while a trait dependent reproduction rate $r(\cdot)$ is taken into account in the reproduction term T. More importantly, this reproduction rate is assumed to be constant outside of a compact set. Second, [28] does not consider a competition term as in (1) but considers a renormalized equation closely related to (4) such that the total population size remains constant equal to 1. Finally, in [28] the segregational variance is assumed to be of order 1, while in our work it is of order ε^2, but the selection term is supposed to be small, that is $r(x) = 1 + \varepsilon a(x)$, with $a(x)$ compactly supported. One can make a change of variable in our model, that is $\widetilde{n}_\varepsilon(t, x) = \varepsilon n_\varepsilon(\varepsilon^2 t, \varepsilon x)$ and $\widetilde{\rho}_\varepsilon(t) = \rho(\varepsilon^2 t)$, to bring the segregational variance equal to 1, in order to compare our model with the one studied in [28]. Then (1) becomes

$$\partial_t \widetilde{n}_\varepsilon = T_1[\widetilde{n}_\varepsilon] - (m(\varepsilon \cdot) + \kappa \widetilde{\rho}_\varepsilon(t)) \widetilde{n}_\varepsilon. \tag{15}$$

Comparing $m(\varepsilon \cdot)$ and $r(x) = 1 + \varepsilon a(x)$ we realize that these models have different natures. The most important difference in the outcome of these models is that, since a is compactly supported, the work of [28] only allows for an order 1 evolution of the mean phenotypic trait of the population, while our model allows for variations of order $1/\varepsilon$, after the change of variable leading to (15). Our work is indeed in the weak selection regime only in the sense that the evolutionary dynamics are slower than the demographic dynamics but it still allows, contrary to [28], important changes of the mean phenotypic trait in long time.

The scaling that we consider is indeed the one studied in [5, 24] (see also [8] for the analysis of a model with a spatial structure and where the analysis of the infinitesimal operator is partially formal). This scaling is inspired from previous works on selection-mutation models with asexual reproduction under the assumption of small mutational variance [10, 25, 1, 17]. The main ingredient of this approach is a Hopf-Cole transformation that leads to a Hamilton-Jacobi equation with constraint. In the case of the infinitesimal operator, a similar scaling was suggested in [5] and an asymptotic analysis of a similar model to (1), without the competition term, was provided in [5] to characterize steady solutions and in [24] to study the time-dependent problem. Although these studies are also based on the Hopf-Cole transformation, they are very different from the case of

asexual reproduction. These analyses rely on a perturbative analysis and do not involve Hamilton-Jacobi equations. Our work allows to provide an alternative approach to the one in [24] that relies on a direct study of the dynamics of the moments rather than a Hopf-Cole transformation. Note also that our assumptions are different from those of [24]. In [24] a function $M(t, x)$ was defined as follows (rewriting the definition with our notations)

$$M(t, x) = r + m(x) - m(\overline{Z}(t)) - m'(\overline{Z}(t))(x - \overline{Z}(t)).$$

Some technical assumptions on the growth of M were made in [24] and it was additionally assumed that

$$\inf_{(t,x)\in\mathbb{R}_+\times\mathbb{R}} M(t, x) > 0. \tag{16}$$

This assumption may be compared to Assumption (H2). Assumption (H2) together with the first statement of Assumption (H4) can indeed be replaced by the following

$$0 < r + m(x) - m(\overline{Z}(t)), \qquad \text{for all } (t, x) \in \mathbb{R}_+ \times \mathbb{R}. \tag{17}$$

This assumption is less restrictive, compared to (16), on the choice of initial condition when the mortality rate m is non-convex and it has for instance several local minima. Note for instance that when the function m has two global minima $x_1 < x_2$, Assumption (16) does not allow for initial conditions which are concentrated around a point close but to the right of x_1, while Assumption (17) allows for initial conditions which are concentrated around a point located in the convexity zones around x_1 or x_2. For convex growth rates m, Assumption (16) is less restrictive since it always holds true. Note however from (11) that to have a positive asymptotic population size $\rho(t)$ one should have $r - m(\overline{Z}(t)) > 0$, which trivially implies that $r + m(x) - m(\overline{Z}(t)) > 0$. Assumption (17) is not hence restrictive from the modeling point of view. Finally it is worth mentioning that in [5] the authors proved the existence of steady solutions which concentrate around local minimum points x_* which satisfy the following condition

$$0 < r + m(x) - m(x_*), \qquad \text{for all } x \in \mathbb{R}.$$

More recent works have also studied the long time behavior of closely related models, but considering discrete time, without the assumption of weak selection [6, 7]. Other models of adaptive evolution of populations with sexual reproduction and quantitative traits have been studied in [12, 26, 9]. An asymptotic analysis of a selection-mutation model with an asymmetric reproduction term, considering for instance traits that are mostly inherited from the female, has been provided in [26]. In [9] a model describing the adaptation of quantitative alleles at two loci in a haploid sexually reproducing population has been studied. In [12] a non-expanding transport distance has been introduced in a model with sexual reproduction considering quantitative traits, but with constant birth and death rates.

1.5   Plan of the paper

In Section 2 we provide the estimates on the moments of the phenotypic distribution and prove Theorem 3. In Section 3 and Section 4 we prove respectively Theorem 2-(i) and (ii). In Section 5 we provide a comparison between our work and the study of related models with an asexual reproduction term. Finally the main ingredients of the proof of Proposition 1 are given in Appendix A.

In all computations we use the symbol C for a harmless constant that may vary between lines. We use labelled constants to keep track of the important parameters.

2 Estimates on the moments of the phenotypic distribution; the proof of Theorem 3

In the whole section, q_ε is a solution to (4) with an arbitrary large number of bounded moments so that we may assume that all the necessary moments to our analysis are C^1 in time (see Appendix A). We split the approach of Theorem 3 into two steps. Firstly, we compute *a priori* ODEs satisfied by moments of q_ε. Secondly, we apply a continuation argument to these ODEs to showing that we can find K_0, K_1, K_2 depending only on the data of the problem such that (12)-(13)-(14) hold globally in time for arbitrary $\delta < 1$ provided ε is sufficiently small.

Before providing our analysis of the moments of the phenotypic distribution, we first introduce some notations and ingredients of the proof. We denote $r^f[X]$ the (normalized) remainder in the first order Taylor-Lagrange expansion in $X \in \mathbb{R}$ of a C^2 function f. Namely, we set:

$$r^f[X](x) = \int_0^1 (1 - \sigma)f''(X + \sigma(x - X))\mathrm{d}\sigma,$$

so that

$$f(x) = f(X) + (x - X)f'(X) + (x - X)^2 r^f[X](x). \tag{18}$$

In case of $f = m$, thanks to (H3), we have

$$|r^m[X](x)| \le CA_m(1 + |X|^p + |x - X|^p) \quad \forall (x, X) \in \mathbb{R}^2. \tag{19}$$

Finally, a key-quantity below is:

$$I_{\varepsilon,m}(t) := \int_{\mathbb{R}} m(x)q_\varepsilon(t, x)\mathrm{d}x.$$

Using the Taylor expansion $m(x) = m(M_{\varepsilon,1}(t)) + m'(M_{\varepsilon,1}(t))(x - M_{\varepsilon,1}(t)) + r^m[M_{\varepsilon,1}(t)](x)(x - M_{\varepsilon,1}(t))^2$ and recalling that $M_{\varepsilon,1}(t)$ is the first moment of $q_\varepsilon(t, \cdot)$, we conclude that

$$I_{\varepsilon,m}(t) = m(M_{\varepsilon,1}(t)) + \int_{\mathbb{R}} (x - M_{\varepsilon,1}(t))^2 r^m[M_{\varepsilon,1}(t)](x)q_\varepsilon(t, x)\mathrm{d}x. \tag{20}$$

The inequality (19) entails

$$|I_{\varepsilon,m}(t) - m(M_{\varepsilon,1}(t))| \le CA_m \left((1 + |M_{\varepsilon,1}(t)|^p)M^c_{\varepsilon,2}(t) + M^{|c|}_{\varepsilon,2+p}(t) \right). \tag{21}$$

2.1 Time-evolution of the mean

Firstly, we multiply (4) by x and integrate. Using that the first moment of $\Gamma_\varepsilon(x)$ vanishes, we obtain:

$$\varepsilon^2 \dot{M}_{\varepsilon,1} = - \left(\int_{\mathbb{R}} m(x)x q_\varepsilon(\cdot, x)\mathrm{d}x - I_{\varepsilon,m}M_{\varepsilon,1} \right).$$

We compute the right-hand side RHS of this identity by using taylor expansions. We write $\Psi(x) = xm(x)$ and expand Ψ and m around $X = M_{\varepsilon,1}$ with a taylor formula of order 1. For the first term, we note that $\Psi(x) = m(M_{\varepsilon,1})M_{\varepsilon,1} + \Psi'(M_{\varepsilon,1})(x - M_{\varepsilon,1}) + r^\Psi[M_{\varepsilon,1}](x)(x - M_{\varepsilon,1})^2$. Since $M_{\varepsilon,1}$ is the first moment of q_ε, we obtain

$$\int_{\mathbb{R}} m(x)x q_\varepsilon(\cdot, x)\mathrm{d}x = M_{\varepsilon,1}m(M_{\varepsilon,1}) + \int_{\mathbb{R}} (x - M_{\varepsilon,1})^2 r^\Psi[M_{\varepsilon,1}](x)q_\varepsilon(\cdot, x)\mathrm{d}x.$$

For the second term of the RHS, we recall (20) and we have finally:

$$RHS = - \int_{\mathbb{R}} \left[r^\Psi[M_{\varepsilon,1}](x) - M_{\varepsilon,1}r^m[M_{\varepsilon,1}](x) \right] (x - M_{\varepsilon,1})^2 q_\varepsilon(\cdot, x)\mathrm{d}x.$$

Noticing that $\Psi''(x) = xm''(x) + 2m'(x)$ we conclude that:

$$\varepsilon^2 \dot{M}_{\varepsilon,1} + m'(M_{\varepsilon,1})M_{\varepsilon,2}^c = -F_1 \tag{22}$$

where:

$$F_1 = \int_{\mathbb{R}} \left[\int_0^1 (1-\sigma)\left(m''(M_{\varepsilon,1} + \sigma(x - M_{\varepsilon,1}))\sigma(x - M_{\varepsilon,1})^3 \right. \right.$$
$$\left. \left. + 2(m'(M_{\varepsilon,1} + \sigma(x - M_{\varepsilon,1})) - m'(M_{\varepsilon,1}))(x - M_{\varepsilon,1})^2 \right) \mathrm{d}\sigma \right] q_\varepsilon(\cdot,x)\mathrm{d}x$$

At this point, we apply (H3) to obtain:

$$|m''(M_{\varepsilon,1} + \sigma(x - M_{\varepsilon,1}))| \le CA_m \left(1 + |M_{\varepsilon,1}|^p + |x - M_{\varepsilon,1}|^p\right),$$
$$|(m'(M_{\varepsilon,1} + \sigma(x - M_{\varepsilon,1})) - m'(M_{\varepsilon,1}))| \le CA_m|x - M_{\varepsilon,1}| \left(1 + |M_{\varepsilon,1}|^p + |x - M_{\varepsilon,1}|^p\right).$$

Introducing this control in the definition of F_1 we infer:

$$|F_1| \le CA_m \left((1 + |M_{\varepsilon,1}|^p)M_{\varepsilon,3}^{|c|} + M_{\varepsilon,3+p}^{|c|} \right). \tag{23}$$

2.2 Time-evolution for the second moment

Concerning the second centered moment, we compute

$$\varepsilon^2 \dot{M}_{\varepsilon,2}^c = \int_{\mathbb{R}} (x - M_{\varepsilon,1})^2 \varepsilon^2 \partial_t q_\varepsilon(\cdot,x)dx - 2\varepsilon^2 \dot{M}_{\varepsilon,1} \int_{\mathbb{R}} (x - M_{\varepsilon,1})q_\varepsilon(\cdot,x)dx$$

$$= r \int_{\mathbb{R}} (x - M_{\varepsilon,1})^2 (\widetilde{T}_\varepsilon[q_\varepsilon](x) - q_\varepsilon(\cdot,x))dx$$

$$- \left(\int_{\mathbb{R}} m(x)(x - M_{\varepsilon,1})^2 q_\varepsilon(\cdot,x)dx - I_{\varepsilon,m} \int_{\mathbb{R}} (x - M_{\varepsilon,1})^2 q_\varepsilon(\cdot,x)dx \right)$$

$$= r \left(\int_{\mathbb{R}} (x - M_{\varepsilon,1})^2 \widetilde{T}_\varepsilon[q_\varepsilon](x)dx - M_{\varepsilon,2}^c \right) - \left(\int_{\mathbb{R}} m(x)(x - M_{\varepsilon,1})^2 q_\varepsilon(\cdot,x)dx - I_{\varepsilon,m} M_{\varepsilon,2}^c \right).$$

We go now into the details of the right-hand side. For the first term, we compute:

$$\int_{\mathbb{R}} (x - M_{\varepsilon,1})^2 \widetilde{T}_\varepsilon[q_\varepsilon](x)dx$$

$$= \int_{\mathbb{R}} \int_{\mathbb{R}} \int_{\mathbb{R}} \left(\left(x - \frac{y + y'}{2} \right) + \left(\frac{y + y'}{2} - M_{\varepsilon,1} \right) \right)^2 \Gamma_\varepsilon \left(x - \frac{y + y'}{2} \right) q_\varepsilon(\cdot,y)q_\varepsilon(\cdot,y')\mathrm{d}y\mathrm{d}y'\mathrm{d}x$$

$$= \frac{\varepsilon^2}{2} + \int_{\mathbb{R}} \int_{\mathbb{R}} \left(\frac{(y - M_{\varepsilon,1}) + (y' - M_{\varepsilon,1})}{2} \right)^2 q_\varepsilon(\cdot,y)q_\varepsilon(\cdot,y')\mathrm{d}y\mathrm{d}y'$$

$$= \frac{\varepsilon^2}{2} + \frac{M_{\varepsilon,2}^c}{2}.$$

Eventually, we obtain:

$$\varepsilon^2 \dot{M}_{\varepsilon,2}^c = r \left[\frac{\varepsilon^2}{2} - \frac{M_{\varepsilon,2}^c}{2} \right] - \left(\int_{\mathbb{R}} m(x)(x - M_{\varepsilon,1})^2 q_\varepsilon(\cdot,x)dx - I_{\varepsilon,m} M_{\varepsilon,2}^c \right)$$

We proceed with the same expansion trick as above to compute the last term in parenthesis. We have:

$$\int_{\mathbb{R}} m(x)(x - M_{\varepsilon,1})^2 q_\varepsilon(\cdot,x)dx = m(M_{\varepsilon,1})M_{\varepsilon,2}^c + m'(M_{\varepsilon,1})M_{\varepsilon,3}^c$$

$$+ \int_{\mathbb{R}} \int_0^1 (1-\sigma)m''(M_{\varepsilon,1} + \sigma(x - M_{\varepsilon,1}))(x - M_{\varepsilon,1})^4 \mathrm{d}\sigma\, q_\varepsilon(\cdot,x)\mathrm{d}x,$$

8

that we combine with (20) to yield:

$$\left(\int_{\mathbb{R}} m(x)(x-M_{\varepsilon,1})^2 q_\varepsilon(\cdot,x)dx - I_{\varepsilon,m}\,M_{\varepsilon,2}^c\right) = m'(M_{\varepsilon,1})M_{\varepsilon,3}^c$$

$$+ \int_{\mathbb{R}}\int_0^1 (1-\sigma)m''(M_{\varepsilon,1}+\sigma(x-M_{\varepsilon,1}))((x-M_{\varepsilon,1})^4 - (x-M_{\varepsilon,1})^2 M_{\varepsilon,2}^c)\mathrm{d}\sigma\, q_\varepsilon(\cdot,x)\mathrm{d}x,$$

and finally:

$$\varepsilon^2 \dot{M}_{\varepsilon,2}^c + \frac{r}{2}M_{\varepsilon,2}^c = r\frac{\varepsilon^2}{2} + F_2, \tag{24}$$

where F_2 corresponds to the right-hand side of this latter identity and is controlled thanks to (H3) by:

$$|F_2| \leq CA_m\left[(1+|M_{\varepsilon,1}|^p)\left(|M_{\varepsilon,2}^c|^2 + M_{\varepsilon,4}^c\right)\right.$$

$$\left. +(|M_{\varepsilon,1}|+|M_{\varepsilon,1}|^{p+1})M_{\varepsilon,3}^{|c|} + M_{\varepsilon,4+p}^{|c|} + M_{\varepsilon,2}^c M_{\varepsilon,2+p}^{|c|}\right]. \tag{25}$$

2.3 Time-evolution for higher-order moments

Finally, we compute a time-evolution equation for $M_{\varepsilon,2k}^c$ with k sufficiently large. Below, we shall apply this computation with $k=k_0$ fixed by (8). Similarly to the case of the second order moment, we derive:

$$\varepsilon^2 \dot{M}_{\varepsilon,2k}^c = \int_{\mathbb{R}}(x-M_{\varepsilon,1})^{2k}\varepsilon^2\partial_t q_\varepsilon(\cdot,x)\,dx - 2k\varepsilon^2\dot{M}_{\varepsilon,1}\int_{\mathbb{R}}(x-M_{\varepsilon,1})^{2k-1}q_\varepsilon(\cdot,x)\,dx$$

$$= r\left(\int_{\mathbb{R}}(x-M_{\varepsilon,1})^{2k}\widetilde{T}_\varepsilon[q_\varepsilon](x)dx - M_{\varepsilon,2k}^c\right) - \int_{\mathbb{R}} m(x)(x-M_{\varepsilon,1})^{2k}q_\varepsilon(\cdot,x)\,dx + I_{\varepsilon,m}M_{\varepsilon,2k}^c$$

$$+ 2k(F_1 + m'(M_{\varepsilon,1})M_{\varepsilon,2}^c)M_{\varepsilon,2k-1}^c$$

$$\leq r\left[\int_{\mathbb{R}}(x-M_{\varepsilon,1})^{2k}\widetilde{T}_\varepsilon[q_\varepsilon](x)dx - M_{\varepsilon,2k}^c\right] + m(M_{\varepsilon,1})M_{\varepsilon,2k}^c + (I_{\varepsilon,m} - m(M_{\varepsilon,1}))M_{\varepsilon,2k}^c$$

$$+ 2k(F_1 + m'(M_{\varepsilon,1})M_{\varepsilon,2}^c)M_{\varepsilon,2k-1}^c.$$

Here we have used the fact that $m(x)(x-M_{\varepsilon,1})^{2k} \geq 0$. Concerning the first term in the bracket multiplied by r we have actually an algebra based on polynomial expansions. To compute this term, we note that all odd moments of Γ_ε vanish. By a scaling argument, we have also that the $2l$ moments of Γ_ε read $\sigma_l\varepsilon^{2l}$ with a series of constants $(\sigma_l)_{l\in\mathbb{N}^*}$ independent of $\varepsilon > 0$ such that $\sigma_1 = 1/2$. We now compute

$$\int(x-M_{\varepsilon,1})^{2k}\widetilde{T}_\varepsilon[q_\varepsilon](x)dx =$$

$$= \int_{\mathbb{R}}\int_{\mathbb{R}}\int_{\mathbb{R}}\left(\left(x-\frac{y+y'}{2}\right)+\left(\frac{y+y'}{2}-M_{\varepsilon,1}\right)\right)^{2k}\Gamma_\varepsilon\left(x-\frac{y+y'}{2}\right)q_\varepsilon(\cdot,y)q_\varepsilon(\cdot,y')\mathrm{d}y\mathrm{d}y'\mathrm{d}x$$

$$= \sum_{l=0}^{k}\sum_{j=0}^{2l}\sigma_{k-l}\varepsilon^{2(k-l)}\frac{1}{4^l}\binom{2k}{2l}\binom{2l}{j}\int_{\mathbb{R}}\int_{\mathbb{R}}(y-M_{\varepsilon,1})^{2l-j}(y'-M_{\varepsilon,1})^j q_\varepsilon(\cdot,y)q_\varepsilon(\cdot,y')\mathrm{d}y\mathrm{d}y'$$

$$= \frac{2}{4^k}M_{\varepsilon,2k}^c + \sum_{l=0}^{k-1}\sum_{j=0}^{2l}\sigma_{k-l}\varepsilon^{2(k-l)}\frac{1}{4^l}\binom{2k}{2l}\binom{2l}{j}M_{\varepsilon,2l-j}^c M_{\varepsilon,j}^c + \sum_{j=2}^{2k-2}\frac{1}{4^k}\binom{2k}{j}M_{\varepsilon,2k-j}^c M_{\varepsilon,j}^c,$$

where we use the convention that $M_{\varepsilon,0}^c = 1$ and $M_{\varepsilon,1}^c = 0$. Eventually, we obtain:

$$\varepsilon^2 \dot{M}_{\varepsilon,2k}^c + \left[r\left(1-\frac{2}{4^k}\right) - m(M_{\varepsilon,1})\right]M_{\varepsilon,2k}^c \leq F_{2k} \tag{26}$$

9

where

$$F_{2k} \leq 2k \left(|F_1| + CA_m(|M_{\varepsilon,1}| + |M_{\varepsilon,1}|^{p+1})M_{\varepsilon,2}^c \right) |M_{\varepsilon,2k-1}^c| \quad \text{[Time direvative of } \dot{M}_{\varepsilon,1}]$$

$$+ CA_m M_{\varepsilon,2k}^c \left((1 + |M_{\varepsilon,1}|^p)M_{\varepsilon,2}^c + M_{\varepsilon,2+p}^{|c|} \right) \quad \text{[Remainder of } I_{\varepsilon,m}]$$

$$+ \sum_{l=0}^{k-1} \sum_{j=0}^{2l} \sigma_{k-l} \varepsilon^{2(k-l)} \frac{1}{4^l} \binom{2k}{2l} \binom{2l}{j} M_{\varepsilon,2l-j}^c M_{\varepsilon,j}^c$$

$$+ \sum_{j=2}^{2k-2} \frac{1}{4^k} \binom{2k}{j} M_{\varepsilon,2k-j}^c M_{\varepsilon,j}^c.$$

2.4 The proof of Theorem 3

We are now in position to prove Theorem 3. We first recall the function $\bar{Z}_\varepsilon$ given by (7). Due to the local strict-convexity assumption on m, we have $|m'(\bar{Z}_\varepsilon)| \geq A_0|\bar{Z}_\varepsilon|$ as long as $Z_\varepsilon \in (-L, L)$. In particular, thanks to Assumption (H4), we have $\bar{Z}_\varepsilon(0) \in (-L, L)$. Hence, we have the following decay

$$|\bar{Z}_\varepsilon(t)| \leq \exp(-A_0 t)|\bar{Z}_\varepsilon(0)|.$$

We note also here that we have the following classical result. Assume that $Y \geq 0$ satisfies:

$$\varepsilon^2 \dot{Y} + kY \leq F$$

with $F \in L^\infty(0, T)$. Then there holds:

$$|Y(t)| \leq \frac{\|F\|_{L^\infty(0,T)}}{k} + Y(0)e^{\frac{-kt}{\varepsilon^2}}, \quad \forall t \in [0, T]. \tag{27}$$

Combining this property with the estimates obtained previously, we will prove Theorem 3. For this, given $\varepsilon > 0$ and K_0, K_1, K_2, (to be chosen below) we set:

$$T_\varepsilon := \sup\{t \in [0, \infty) \text{ s.t. (12)–(14) hold on } [0, t]\}.$$

Our aim is to find a K_0, K_1, K_2 so that for small ε we have $T_\varepsilon = +\infty$. We show that such a choice is possible with the following proposition:

Proposition 4 *Assume (H1)–(H4) are in force. We can find K_0, K_1 and K_2 large enough, such that for $\delta \in (0, 1)$ and arbitrary small ε depending on δ there holds:*

i) $T_\varepsilon > 0$

ii) if $T_\varepsilon \in (0, \infty)$, we have

$$M_{\varepsilon,2k_0}^c(t) < K_2 \varepsilon^{2k_0} \tag{28}$$

$$\left| M_{\varepsilon,2}^c(t) - \varepsilon^2 \right| < K_1 \varepsilon^2 \left(\varepsilon^{1-\delta} + \exp(-rt/2\varepsilon^2) \right) \tag{29}$$

$$\left| M_{\varepsilon,1}(t) - \bar{Z}_\varepsilon(t) \right| < K_0 \varepsilon^{1-\delta}. \tag{30}$$

on $[0, T_\varepsilon]$.

We notice that the inequalities in (12)–(14) are $\leq$ while in (28)–(30) they are $<$. With the above proposition, we may conclude by a standard contradiction argument that (12)-(13)-(14) hold on $[0, \infty)$ since $M_{\varepsilon,1}$, $M_{\varepsilon,2}^c$ and $M_{\varepsilon,2k_0}^c$ are time-continuous.

Proof. Thanks to assumption (H4), item $i)$ only amounts to choosing K_0, K_1, K_2 sufficiently large w.r.t. initial data. Indeed, by the second part of (H4), we have $M_{\varepsilon,2k_0}^c(0) \leq C_1 \varepsilon^{2k_0}$ so that

we have (12) on some initial time-interval whenever $K_2 > C_1$. Then, by a consequence of Hölder inequality, we have

$$|M^c_{\varepsilon,2}(0) - \varepsilon^2| \leq (C_1^{\frac{1}{k_0}} + 1)\varepsilon^2.$$

Since the right-hand side of (13) is larger than $K_1\varepsilon^2$ intially, we conclude that (13) holds on some initial time-interval whenever $K_1 > (C_1^{\frac{1}{k_0}} + 1)$. Finally, $M_{\varepsilon,1}(0) - \bar{Z}_\varepsilon(0) = 0$ so that (14) is also true for t small and K_0 arbitrary.

We focus now on the proof of item $ii)$. For the proof, we assume that K_0, K_1, K_2 are constructed and we explain the restrictions that they must satisfy so that our computations hold true. We emphasize that we shall choose first K_1 depending on data C_1, then K_2 depending on K_1, C_1 and finally K_0 depending on K_1, L. We restrict ε according to K_2, K_0 and L. We note first that with (13)–(14) we have directly, when ε is sufficiently small:

$$|M_{\varepsilon,1}(t)| \leq L, \qquad |M^c_{\varepsilon,2}(t)| \leq 2K_1\varepsilon^2 \quad \text{on } [0, T_\varepsilon]. \tag{31}$$

Fixing from now on that $K_2 > 2K_1$ the control (12)–(13) also entails thanks to a consequence of Hölder inequality that

$$M^{|c|}_{\varepsilon,l} \leq K_2^{\frac{l-2}{2k_0-2}}(2K_1)^{\frac{2k_0-l}{2k_0-2}}\varepsilon^l \leq K_2\varepsilon^l, \quad \forall l \in \{2,\ldots,2k_0-1\} \quad \text{on } [0, T_\varepsilon]. \tag{32}$$

We start now with the computation of $M^c_{\varepsilon,2}$ by noting that (24) rewrites:

$$\begin{cases} \dot{M}^c_{\varepsilon,2} + \dfrac{r}{2\varepsilon^2}M^c_{\varepsilon,2} = \dfrac{r}{2} + \dfrac{1}{\varepsilon^2}F_2 \\ M^c_{\varepsilon,2}(0) = \alpha_{\varepsilon,0}\varepsilon^2, \end{cases} \tag{33}$$

where, $\alpha_{\varepsilon,0} \leq C_1^{\frac{1}{k_0}}$ and

$$|F_2| \leq CA_m \left[(1 + |M_{\varepsilon,1}|^p)\left(|M^c_{\varepsilon,2}|^2 + M^c_{\varepsilon,4}\right) + (|M_{\varepsilon,1}| + |M_{\varepsilon,1}|^{p+1})M^{|c|}_{\varepsilon,3} + M^{|c|}_{\varepsilon,4+p} + M^c_{\varepsilon,2}M^{|c|}_{\varepsilon,2+p}\right].$$

Plugging the controls (31)–(32) into this identity we obtain that, on $(0, T_\varepsilon)$, we have:

$$|F_2| \leq CA_m \left[(1 + L^p)\left(K_2^2\varepsilon^4 + K_2\varepsilon^4\right) + (L + L^{p+1})K_2\varepsilon^3 + K_2\varepsilon^{4+p} + K_2^2\varepsilon^{4+p}\right]$$
$$\leq CA_m(1 + L^{p+1})K_2^2\varepsilon^3.$$

By direct integration of (33) we conclude that:

$$M^c_{\varepsilon,2}(t) = \alpha_{\varepsilon,0}\varepsilon^2 \exp\left[-\frac{r}{2\varepsilon^2}t\right] + \int_0^t \exp\left[\frac{r}{2\varepsilon^2}(\sigma - t)\right]\left[\frac{r}{2} + \frac{1}{\varepsilon^2}F_2(\sigma)\right]\mathrm{d}\sigma$$

$$= \varepsilon^2 + \varepsilon^2(\alpha_{\varepsilon,0} - 1)\exp\left[-\frac{r}{2\varepsilon^2}t\right] + \frac{1}{\varepsilon^2}\int_0^t \exp\left[\frac{r}{2\varepsilon^2}(\sigma - t)\right]F_2(\sigma)\mathrm{d}\sigma.$$

We also note that the bound above on $F_2(t)$ leads to

$$\left|\frac{1}{\varepsilon^2}\int_0^t \exp\left[\frac{r}{2\varepsilon^2}(\sigma - t)\right]F_2(\sigma)\mathrm{d}\sigma\right| \leq \frac{2}{r}CA_m(1 + L^{p+1})K_2^2\varepsilon^3$$

$$\leq \frac{K_1}{2}\varepsilon^{3-\delta},$$

when ε is sufficiently small. We obtain finally that, provided ε is sufficiently small

$$|M^c_{\varepsilon,2}(t) - \varepsilon^2| \leq \varepsilon^2(C_1 + 1)\exp\left[-\frac{r}{2\varepsilon^2}t\right] + \frac{K_1}{2}\varepsilon^{3-\delta},$$

and we have (29) if K_1 is chosen sufficiently large with respect to C_1.

We can then rewrite (26) thanks to (8) and (31):

$$\varepsilon^2 \dot{M}^c_{\varepsilon,2k_0} + \eta M^c_{\varepsilon,2k_0} \leq F_{2k_0}, \tag{34}$$

where we recall that:

$$
\begin{aligned}
F_{2k_0} \leq & 2k_0 \left(|F_1| + CA_m(|M_{\varepsilon,1}| + |M_{\varepsilon,1}|^{p+1})M^c_{\varepsilon,2} \right) |M^c_{\varepsilon,2k_0-1}| \\
& + CA_m M^c_{\varepsilon,2k_0} \left((1 + |M_{\varepsilon,1}|^p)M^c_{\varepsilon,2} + M^{|c|}_{\varepsilon,2+p} + M^2_{\varepsilon,1} + M^{p+2}_{\varepsilon,1} \right) \\
& + \sum_{l=0}^{k_0-1} \sum_{j=0}^{2l} \sigma_{k_0-l} \varepsilon^{2(k_0-l)} \frac{1}{4^l} \binom{2k_0}{2l}\binom{2l}{j} M^c_{\varepsilon,2l-j} M^c_{\varepsilon,j} + \sum_{j=2}^{2k_0-2} \frac{1}{4^{k_0}} \binom{2k_0}{j} M^c_{\varepsilon,2k_0-j} M^c_{\varepsilon,j},
\end{aligned}
$$

with

$$|F_1| \leq CA_m((1 + |M_{\varepsilon,1}|^p)M^{|c|}_{\varepsilon,3} + M^{|c|}_{\varepsilon,3+p}).$$

Introducing again Assumption (H4), (31) and (32) we obtain that:

$$
\begin{aligned}
2k_0 \left(|F_1| + CA_m(|M_{\varepsilon,1}| + |M_{\varepsilon,1}|^{p+1})M^c_{\varepsilon,2} \right) |M^c_{\varepsilon,2k_0-1}| \\
+ CA_m M^c_{2k_0} \left((1 + |M_{\varepsilon,1}|^p)M^c_{\varepsilon,2} + M^{|c|}_{\varepsilon,2+p} \right) \leq CA_m(1 + L^{p+1})K_2^2 \varepsilon^{2k_0+1}.
\end{aligned}
$$

Choosing ε sufficiently small, we can then make the left-hand side of this inequality smaller than $\alpha K_2 \varepsilon^{2k_0}$ with α arbitrary small. Concerning the last line of F_{2k_0} we use the first interpolation inequalities in (32) (since this quantity only involves moments of order less than $2k_0 - 1$) to yield (we recall that we keep only larger powers of K_1, K_2 since they are assumed larger than 1):

$$
\begin{aligned}
\sum_{j=2}^{2k_0-2} \frac{1}{4^{k_0}} \binom{2k_0}{j} M^c_{\varepsilon,2k_0-j} M^c_{\varepsilon,j} + \sum_{l=0}^{k_0-1} \sum_{j=0}^{2l} \sigma_{k_0-l} \varepsilon^{2(k_0-l)} \frac{1}{4^l} \binom{2k_0}{2l}\binom{2l}{j} M^c_{\varepsilon,2l-j} M^{\varepsilon,c}_j \\
\leq C_{k_0} K_1^{\frac{2k_0}{k_0-1}} K_2^{\frac{k_0-2}{k_0-1}} \varepsilon^{2k_0}
\end{aligned}
$$

We point out that the constant C_{k_0} depends on k_0 also through the explicit quantities $\sigma_1, \ldots, \sigma_{k_0}$. Finally, we can choose K_2 large with respect to K_1 to conclude that:

$$F_{2k_0} \leq \frac{\eta}{2} K_2 \varepsilon^{2k_0} \text{ on } (0, T_\varepsilon).$$

Integrating (34) and using Assumption (H4) entails that:

$$M^c_{\varepsilon,2k}(t) \leq C_1 \varepsilon^{2k_0} + \frac{K_2}{2} \varepsilon^{2k_0},$$

and we have (28) when K_2 is again sufficiently large with respect to C_1.

We end up with the computation of $M_{\varepsilon,1}$. We can rewrite (22) as follows

$$\dot{M}_{\varepsilon,1} + m'(M_{\varepsilon,1}) = \tilde{F}_1(t)$$

where thanks to (12), (13) and (31):

$$|\tilde{F}_1(t)| \leq \frac{|F_1|}{\varepsilon^2} + \frac{1}{\varepsilon^2} |m'(M_{\varepsilon,1})| \, |M^c_{\varepsilon,2} - \varepsilon^2| \leq CA_m(1 + L^{p+1})K_1 \left(\varepsilon^{1-\delta} + \exp(-rt/2\varepsilon^2) \right).$$

Hence $N_1 = M_{\varepsilon,1} - \bar{Z}_\varepsilon$ satisfies

$$\dot{N}_1 + m''(\bar{Z}_\varepsilon)N_1 = \tilde{F}_1 - (m'(M_{\varepsilon,1}) - m'(\bar{Z}_\varepsilon) - m''(\bar{Z}_\varepsilon)N_1), \qquad N_1(0) = 0.$$

Moreover, we can write $m'(M_{\varepsilon,1}) - m'(\bar{Z}_\varepsilon) = m''(Z_\theta)N_1$ for some Z_θ between $\min(M_{\varepsilon,1}, \bar{Z}_\varepsilon)$ and $\max(M_{\varepsilon,1}, \bar{Z}_\varepsilon)$ and use again a finite difference theorem to compute $m''(Z_\theta) - m''(\bar{Z}_\varepsilon)$. This entails thanks to (14):

$$|m'(M_{\varepsilon,1}) - m'(\bar{Z}_\varepsilon) - m''(\bar{Z}_\varepsilon)N_1| \leq C_m K_0 \varepsilon^{1-\delta}|N_1|.$$

with a constant C_m depending on m'''. Consequently, we have by a standard energy estimate

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[\frac{|N_1|^2}{2}\right] + A_0|N_1|^2 \leq (\tilde{F}_1 + C_m K_0 \varepsilon^{1-}|N_1|)N_1.$$

Eventually, we obain that, for ε sufficiently small:

$$\frac{|N_1(t)|^2}{2} \leq \frac{C}{A_0}\int_0^t \exp(-A_0(t-s))|\tilde{F}_1(s)|^2\mathrm{d}s$$

$$\leq \frac{CA_m^2(1+L^{p+1})^2 K_1^2}{A_0}\left(\int_0^t \exp(-A_0(t-s))\varepsilon^{2(1-\delta)}\mathrm{d}s + \int_0^t \exp(-rs/\varepsilon^2)\mathrm{d}s\right)$$

$$\leq \frac{CA_m^2(1+L^{p+1})^2 K_1^2}{A_0}\left(\frac{\varepsilon^{2(1-\delta)}}{A_0} + \frac{\varepsilon^2}{r}\right)$$

We obtain (14) provided K_0 is chosen sufficiently large with respect to K_1, A_0 and L. This concludes the proof. $\square$

3 Wasserstein estimates; The proof of Theorem 2-(i).

In this section, we assume that q_ε is a global solution to (4) with sufficiently large bounded moments. We assume that the conclusions of Theorem 3 hold true and we recall that notations $g_\varepsilon, \bar{Z}_\varepsilon$ are defined in the introduction. Again, we use throughout the section the symbol C for a harmless constant that may vary between lines. We reserve the symbol K for a constant that depend only on L, K_0, K_1 and K_2. This constant may also vary between lines.

We first remark that g_ε solves

$$\varepsilon^2 \partial_t g_\varepsilon(t,x) = r(\tilde{T}[g_\varepsilon](t,x) - g_\varepsilon(t,x)) - (x - \bar{Z}_\varepsilon(t))m'(\bar{Z}_\varepsilon(t))g_\varepsilon(t,x). \tag{35}$$

We will use below that the last term writes:

$$(x - \bar{Z}_\varepsilon)m'(\bar{Z}_\varepsilon) = m(x) - m(\bar{Z}_\varepsilon) - r^m[\bar{Z}_\varepsilon](x)(x - \bar{Z}_\varepsilon)^2$$

where:

$$|r^m[\bar{Z}_\varepsilon](x)| \leq CA_m\left(1 + |\bar{Z}_\varepsilon|^p + |x - \bar{Z}_\varepsilon|^p\right).$$

We also define

$$\bar{g}_\varepsilon(t,x) = \frac{1}{\sqrt{2\pi}\varepsilon}\exp\left(-\frac{(x - M_{\varepsilon,1}(t))^2}{2\varepsilon^2}\right) \tag{36}$$

with $M_{\varepsilon,1}$ given by (5).

We propose now to derive some duality estimates for the difference $(q_\varepsilon - g_\varepsilon)$ by computing:

$$I_\phi(t) = \int_{\mathbb{R}} \phi(x)(q_\varepsilon(t,x) - g_\varepsilon(t,x))dx,$$

for arbitrary lipschitz bounded (possiblity smooth) test-function such that $\phi(0) = 0$. Multiplying (4) and (35) by ϕ and taking the difference, we get

$$\varepsilon^2 \dot{I}_\phi = rT_r - rI_\phi - T_s, \tag{37}$$

where

$$T_r = \int_{\mathbb{R}} \phi(\tilde{T}_\varepsilon[q_\varepsilon] - \tilde{T}_\varepsilon[g_\varepsilon]),$$

$$T_s = \left(\int_{\mathbb{R}} mq_\varepsilon\phi - \int_{\mathbb{R}} mg_\varepsilon\phi\right) - \left[\left(\int_{\mathbb{R}} mq_\varepsilon\right)\left(\int_{\mathbb{R}} q_\varepsilon\phi\right) - m(\bar{Z}_\varepsilon)\left(\int_{\mathbb{R}} g_\varepsilon\phi\right)\right]$$

$$- \int_{\mathbb{R}} (x - \bar{Z}_\varepsilon)^2 r^m[\bar{Z}_\varepsilon](x)g_\varepsilon(\cdot, x)\phi(x)dx$$

$$= T_s^{(a)} - T_s^{(b)} + T_s^{(c)}.$$

We proceed by controlling now T_r and T_s. Concerning T_r, we write

$$T_r = \int_{\mathbb{R}} \phi(\tilde{T}_\varepsilon[q_\varepsilon] - \tilde{T}_\varepsilon[\bar{g}_\varepsilon]) + \int_{\mathbb{R}} \phi(\tilde{T}_\varepsilon[\bar{g}_\varepsilon] - \tilde{T}_\varepsilon[g_\varepsilon]).$$

For the first term in T_r we use Tanaka's estimate [29, Appendix 4.2] and obtain

$$\left|\int_{\mathbb{R}} \phi(\tilde{T}_\varepsilon[q_\varepsilon] - \tilde{T}_\varepsilon[\bar{g}_\varepsilon])\right| \leq \|\phi'\|_{L^\infty} W_1(\tilde{T}_\varepsilon[q_\varepsilon], \tilde{T}_\varepsilon[\bar{g}_\varepsilon]) \leq \|\phi'\|_{L^\infty} W_2(\tilde{T}_\varepsilon[q_\varepsilon], \tilde{T}_\varepsilon[\bar{g}_\varepsilon])$$

$$\leq \frac{\|\phi'\|_{L^\infty}}{\sqrt{2}} W_2(q_\varepsilon, \bar{g}_\varepsilon) \leq \frac{\|\phi'\|_{L^\infty}}{\sqrt{2}}\left(W_2(q_\varepsilon, g_\varepsilon) + W_2(g_\varepsilon, \bar{g}_\varepsilon)\right),$$

where:

$$W_2(g_\varepsilon, \bar{g}_\varepsilon) \leq \left(\int_{\mathbb{R}}\int_{\mathbb{R}} |x_1 - x_2|^2 g_\varepsilon(x_1)\bar{g}_\varepsilon(x_2)dx_1dx_2\right)^{\frac{1}{2}}$$

$$\leq C\left(\int_{\mathbb{R}}\int_{\mathbb{R}} |(x_1 - \bar{Z}_\varepsilon) - (x_2 - M_{\varepsilon,1})|^2 g_\varepsilon(\cdot, x_1)\bar{g}_\varepsilon(\cdot, x_2)dx_1dx_2\right)^{\frac{1}{2}} + C|M_{\varepsilon,1} - \bar{Z}_\varepsilon|$$

$$\leq C\left(2\int_{\mathbb{R}}\int_{\mathbb{R}} (|x_1 - \bar{Z}_\varepsilon|^2 + |x_2 - M_{\varepsilon,1}|^2)g_\varepsilon(\cdot, x_1)\bar{g}_\varepsilon(\cdot, x_2)dx_1dx_2\right)^{\frac{1}{2}} + C|M_{\varepsilon,1} - \bar{Z}_\varepsilon|$$

$$\leq C\varepsilon + C|M_{\varepsilon,1} - \bar{Z}_\varepsilon|.$$

Furthermore, for any π in the transference plane between q_ε and g_ε, we have

$$\int_{\mathbb{R}}\int_{\mathbb{R}} |x_1 - x_2|^2 d\pi(x_1, x_2)$$

$$\leq \left(\int_{\mathbb{R}}\int_{\mathbb{R}} |x_1 - x_2|d\pi(x_1, x_2)\right)^{\frac{2}{3}}\left(\int_{\mathbb{R}}\int_{\mathbb{R}} |x_1 - x_2|^4 d\pi(x_1, x_2)\right)^{\frac{1}{3}}$$

$$\leq C\left(\int_{\mathbb{R}}\int_{\mathbb{R}} |x_1 - x_2|d\pi(x_1, x_2)\right)^{\frac{2}{3}}\left(\left[\int_{\mathbb{R}}\int_{\mathbb{R}} (|x_1 - M_{\varepsilon,1}|^4 + |x_2 - \bar{Z}_\varepsilon|^4)d\pi(x_1, x_2)\right]^{\frac{1}{3}}\right.$$

$$\left. + |M_{\varepsilon,1} - \bar{Z}_\varepsilon|^{\frac{4}{3}}\right)$$

$$\leq c_0^2\left(\int_{\mathbb{R}}\int_{\mathbb{R}} |x_1 - x_2|d\pi(x_1, x_2)\right)^2 + C\left(\frac{M_{\varepsilon,4}^c + \varepsilon^4 + |M_{\varepsilon,1} - \bar{Z}_\varepsilon|^4}{c_0^2}\right)^{\frac{1}{2}}$$

where c_0 can be chosen arbitrarily small. With a classical inf argument, we obtain

$$W_2(q_\varepsilon, g_\varepsilon)^2 \leq c_0^2 W_1(q_\varepsilon, g_\varepsilon)^2 + C\left(\frac{M_{\varepsilon,4}^c + \varepsilon^4 + |M_{\varepsilon,1} - \bar{Z}_\varepsilon|^4}{c_0^2}\right)^{\frac{1}{2}}.$$

It follows that

$$W_2(q_\varepsilon, g_\varepsilon) \le c_0 W_1(q_\varepsilon, g_\varepsilon) + C \left(\frac{M_{\varepsilon,4}^c + \varepsilon^4 + |M_{\varepsilon,1} - \bar{Z}_\varepsilon|^4}{c_0^2} \right)^{\frac{1}{4}}.$$

and

$$\left| \int \phi(\tilde{T}_\varepsilon[q_\varepsilon] - \tilde{T}_\varepsilon[\bar{g}_\varepsilon]) \right| \le \|\phi'\|_{L^\infty} \left(c_0 W_1(q_\varepsilon, g_\varepsilon) + C \left(\frac{\varepsilon + |M_{\varepsilon,1} - \bar{Z}_\varepsilon| + (M_{\varepsilon,4}^c)^{\frac{1}{4}}}{c_0^{\frac{1}{2}}} \right) \right)$$

To control the second term in T_r, we can simply use

$$\tilde{T}_\varepsilon[\bar{g}_\varepsilon] = \bar{g}_\varepsilon, \qquad \tilde{T}_\varepsilon[g_\varepsilon] = g_\varepsilon.$$

We then obtain with a similar decomposition of $W_1(\bar{g}_\varepsilon, g_\varepsilon)$ as previously:

$$\left| \int_{\mathbb{R}} \phi(\tilde{T}_\varepsilon[\bar{g}_\varepsilon] - \tilde{T}_\varepsilon[g_\varepsilon]) \right| = \left| \int_{\mathbb{R}} \phi(\bar{g}_\varepsilon - g_\varepsilon) \right| \le \|\phi'\|_{L^\infty} W_1(\bar{g}_\varepsilon, g_\varepsilon) \le C(\varepsilon + |M_{\varepsilon,1} - \bar{Z}_\varepsilon|)$$

and

$$|T_r| \le \|\phi'\|_{L^\infty} \left[c_0 W_1(q_\varepsilon, g_\varepsilon) + C \left(\frac{\varepsilon + |M_{\varepsilon,1} - \bar{Z}_\varepsilon| + (M_{\varepsilon,4}^c)^{\frac{1}{4}}}{c_0^{\frac{1}{2}}} \right) \right].$$

Introducing the result of Theorem 3 we conclude that:

$$|T_r| \le \|\phi'\| \left[c_0 W_1(q_\varepsilon, g_\varepsilon) + \frac{K}{c_0^{1/2}} \varepsilon^{1-\delta} \right] \tag{38}$$

where K is a constant that depend only on L, K_0, K_1 and K_2.

We turn to estimating T_s. We use below without mention the controls induced by Theorem 3. Concerning $T_s^{(c)}$, we apply (H3) to bound m'' and we have:

$$|T_s^{(c)}| \le C A_m \int_{\mathbb{R}} (1 + |\bar{Z}_\varepsilon|^p + |x - \bar{Z}_\varepsilon|^p)|x - \bar{Z}_\varepsilon|^2 (|\phi(\bar{Z}_\varepsilon)| + |\phi(x) - \phi(\bar{Z}_\varepsilon)|) g_\varepsilon(\cdot, x) dx$$

$$\le C A_m \|\phi'\|_{L^\infty} \left((1 + |\bar{Z}_\varepsilon|^p) \int_{\mathbb{R}} |x - \bar{Z}_\varepsilon|^3 g_\varepsilon(x) dx + (1 + |\bar{Z}_\varepsilon|^p)|\bar{Z}_\varepsilon| \int_{\mathbb{R}} |x - \bar{Z}_\varepsilon|^2 g_\varepsilon(\cdot, x) dx \right.$$

$$\left. + |\bar{Z}_\varepsilon| \int_{\mathbb{R}} |x - \bar{Z}_\varepsilon|^{p+2} g_\varepsilon(\cdot, x) dx + \int_{\mathbb{R}} |x - \bar{Z}_\varepsilon|^{p+3} g_\varepsilon(\cdot, x) dx \right)$$

$$\le C A_m \|\phi'\|_{L^\infty} (1 + |\bar{Z}_\varepsilon|^{p+1}) \varepsilon^2.$$

Eventually, we conclude that:

$$|T_s^{(c)}| \le K \|\phi'\|_{L^\infty} \varepsilon^2. \tag{39}$$

Concerning $T_s^{(b)}$, we have

$$\left| T_s^{(b)} - m(\bar{Z}_\varepsilon) \int_{\mathbb{R}} (q_\varepsilon - g_\varepsilon) \phi \right| \le \left| \int_{\mathbb{R}} (m - m(\bar{Z}_\varepsilon)) q_\varepsilon \right| \left| \int_{\mathbb{R}} q_\varepsilon \phi \right|$$

where, applying Taylor expansion of m in $\bar{Z}_\varepsilon$:

$$\left| \int_{\mathbb{R}} (m - m(\bar{Z}_\varepsilon)) q_\varepsilon \right| \le \int |r^m[\bar{Z}_\varepsilon](x)|(x - \bar{Z}_\varepsilon)^2 q_\varepsilon(\cdot, x) dx + |m'(\bar{Z}_\varepsilon)||M_{\varepsilon,1} - \bar{Z}_\varepsilon|$$

$$\le C A_m \left(\int_{\mathbb{R}} \left((1 + |\bar{Z}_\varepsilon|^p)|x - \bar{Z}_\varepsilon|^2 + |x - \bar{Z}_\varepsilon|^{p+2} \right) q_\varepsilon(\cdot, x) dx \right) + K|M_{\varepsilon,1} - \bar{Z}_\varepsilon|$$

$$\le C A_m \left((1 + |\bar{Z}_\varepsilon|^p)(M_{\varepsilon,2}^c + |M_{\varepsilon,1} - \bar{Z}_\varepsilon|^2) + M_{\varepsilon,2+p}^{|c|} + |M_{\varepsilon,1} - \bar{Z}_\varepsilon|^{2+p} \right)$$

$$+ K|M_{\varepsilon,1} - \bar{Z}_\varepsilon|,$$

and, with a standard Jensen inequality:

$$\left|\int_{\mathbb{R}} q_\varepsilon \phi\right| \le \|\phi'\|_{L^\infty} \left(\int_{\mathbb{R}} q_\varepsilon(\cdot, x)|x|^2 dx\right)^{\frac{1}{2}} \le \|\phi'\|_{L^\infty} \left|M_{\varepsilon,2}^c + M_{\varepsilon,1}^2\right|^{\frac{1}{2}} \le K\|\phi'\|_{L^\infty}.$$

This entails :

$$\left|T_s^{(b)} - m(\bar{Z}_\varepsilon)I_\phi\right| \le K\varepsilon^{1-\delta}\|\phi'\|_{L^\infty}. \tag{40}$$

Finally we compute the third term, that is

$$T_s^{(a)} = \int_{\mathbb{R}} m(q_\varepsilon - g_\varepsilon)\phi = \int_{\mathbb{R}} mq_\varepsilon \phi - \int_{\mathbb{R}} mg_\varepsilon \phi.$$

We use again Taylor expansion

$$\left|\int_{\mathbb{R}} mq_\varepsilon \phi - m(M_{\varepsilon,1}) \int_{\mathbb{R}} q_\varepsilon \phi\right|$$
$$= \left|\int_{\mathbb{R}} r^m[M_{\varepsilon,1}](x)(x - M_{\varepsilon,1})^2 q_\varepsilon(\cdot, x)\phi(x)\mathrm{d}x + m'(M_{\varepsilon,1}) \int_{\mathbb{R}} (x - M_{\varepsilon,1})q_\varepsilon(\cdot, x)\phi(x)\mathrm{d}x\right|$$
$$\le K\varepsilon\|\phi'\|_{L^\infty}.$$

Proceeding similarly with g_ε centered in $\bar{Z}_\varepsilon$ we end up with

$$\left|T_s^{(a)} - (m(M_{\varepsilon,1}) - m(\bar{Z}_\varepsilon)) \int q_\varepsilon \phi - m(\bar{Z}_\varepsilon)I_\phi\right| \le K\varepsilon\|\phi'\|_{L^\infty}. \tag{41}$$

Combining (39)-(40)-(41), we obtain

$$|T_s| \le K\varepsilon^{1-\delta}\|\phi'\|_{L^\infty}.$$

Eventually, we conclude that

$$\varepsilon^2 \dot{I}_\phi + r I_\phi = F$$

where

$$|F| \le \|\phi'\|_{L^\infty} \left[rc_0 W_1(q_\varepsilon, g_\varepsilon) + \frac{K}{c_0^{1/2}}\varepsilon^{1-\delta}\right].$$

Applying the property (27), we conclude that

$$|I_\phi(t)| \le I_\phi(0) + \frac{1}{r}\sup_{(0,t)}|F(t)|$$

$$\le \|\phi'\|_{L^\infty} \left(W_1(q_{\varepsilon,0}, g_{\varepsilon,0}) + c_0 \sup_{(0,t)} W_1(q_\varepsilon(t, \cdot), g_\varepsilon(t, \cdot)) + \frac{K}{rc_0^{1/2}}\varepsilon^{1-\delta}\right).$$

Applying a sup argument on $\phi \in C_b^1(\mathbb{R})$ such that $\|\phi'\|_{L^\infty} \le 1$ and then on $t \in [0, T]$ (for fixed $T < \infty$) we obtain that

$$\sup_{[0,T]} W_1(q_\varepsilon(t, \cdot), g_\varepsilon(t, \cdot)) \le W_1(q_{\varepsilon,0}, g_{\varepsilon,0}) + c_0 \sup_{(0,T)} W_1(q_\varepsilon(t, \cdot), g_\varepsilon(t, \cdot)) + \frac{K}{rc_0^{1/2}}\varepsilon^{1-\delta}$$

and thus, since c_0 can be chosen arbitrary small, we obtain

$$\sup_{[0,T]} W_1(q_\varepsilon(t, \cdot), g_\varepsilon(t, \cdot)) \le K\left(W_1(q_{\varepsilon,0}, g_{\varepsilon,0}) + \varepsilon^{1-\delta}\right).$$

This ends the proof of (i).

4 Approximation of the population size ρ_ε; the proof of Theorem 2-(ii)

In this section we provide the proof of Theorem 2-(ii), that is the approximation of the total population size ρ_ε and the convergence of n_ε as $\varepsilon \to 0$.

(a) **The approximation of ρ_ε.** We integrate the first line of (1) with respect to x to obtain

$$\varepsilon^2 \frac{d}{dt}\rho_\varepsilon = \rho_\varepsilon \left(r - \int_\mathbb{R} m(y)q_\varepsilon(t,y)dy - \kappa\rho_\varepsilon \right). \tag{42}$$

We use this equation to prove the second statement of Theorem 2 in two steps.

Step 1. We first prove that

$$\int_\mathbb{R} m(y)q_\varepsilon(t,y)dy = m\big(\overline{Z}(t)\big) + O(\varepsilon^{1-\delta}). \tag{43}$$

We use the Taylor expansion (18) for $f = m$ at $z = M_{\varepsilon,1}$ to find

$$\begin{aligned}
\int_\mathbb{R} m(y)q_\varepsilon(t,y)dy =& m(M_{\varepsilon,1}(t)) \int_R q_\varepsilon(t,y)dy + m'(M_{\varepsilon,1}(t)) \int_\mathbb{R} (y - M_{\varepsilon,1}(t))q_\varepsilon(t,y)dy \\
&+ \int_\mathbb{R} (y - M_{\varepsilon,1}(t))^2 r^m[M_{\varepsilon,1}(t)](y)q_\varepsilon(t,y)dy \\
=& m(M_{\varepsilon,1}(t)) + \int_\mathbb{R} (y - M_{\varepsilon,1}(t))^2 r^m[M_{\varepsilon,1}(t)](y)q_\varepsilon(t,y)dy.
\end{aligned}$$

Combining this equality with (19) we obtain

$$\left| \int_\mathbb{R} m(y)q_\varepsilon(\cdot,y)dy - m(M_{\varepsilon,1}) \right| \le C \int_\mathbb{R} (y - M_{\varepsilon,1})^2 \big(1 + |M_{\varepsilon,1}|^p + |y - M_{\varepsilon,1}|^p\big)q_\varepsilon(t,y)dy.$$

Note from (H4), (7) and (14) that, for ε small enough,

$$|M_{\varepsilon,1}(t)| \le L, \qquad \text{for all } t \in \mathbb{R}^+.$$

We deduce that

$$\begin{aligned}
\left| \int_\mathbb{R} m(y)q_\varepsilon(t,y)dy - m(M_{\varepsilon,1}(t)) \right| \le& C(1 + L^p) \int_\mathbb{R} (y - M_{\varepsilon,1}(t))^2 q_\varepsilon(t,y)dy \\
&+ \int_\mathbb{R} (y - M_{\varepsilon,1}(t))^{2+p} q_\varepsilon(t,y)dy.
\end{aligned}$$

We next use (13) and (12) to obtain that

$$\left| \int_\mathbb{R} m(y)q_\varepsilon(t,y)dy - m(M_{\varepsilon,1}(t)) \right| \le C\varepsilon^2,$$

for ε small enough and up to changing the constant C. In order to prove (43) we next compare $M_{\varepsilon,1}(t)$ with $\overline{Z}(t)$. To this end we first recall from (14) that

$$|M_{\varepsilon,1}(t) - \overline{Z}_\varepsilon(t)| \le K_0 \varepsilon^{1-\delta}.$$

Next we subtract (9) from (7) to find that

$$\frac{d}{dt}\big(\overline{Z}_\varepsilon(t) - \overline{Z}(t)\big) = -m'(\overline{Z}_\varepsilon(t)) + m'(\overline{Z}(t)) = -m''(\theta(t))(\overline{Z}_\varepsilon(t) - \overline{Z}(t)),$$

with
$$\theta(t) \in [\min(\overline{Z}_\varepsilon(t), \overline{Z}(t)), \max(\overline{Z}_\varepsilon(t), \overline{Z}(t))] \subset [-L, L].$$

From this equality together with Assumption (H1) we deduce that
$$|\overline{Z}_\varepsilon(t) - \overline{Z}(t)| \leq e^{-A_0 t} |\overline{Z}_\varepsilon(0) - \overline{Z}(0)| \leq C\varepsilon e^{-A_0 t}.$$

Combining the inequalities above we obtain that
$$|m(M_{\varepsilon,1}(t)) - m(\overline{Z}(t))| = O(\varepsilon^{1-\delta}),$$

and hence
$$\sup_{[0,T]} \Big| \int_{\mathbb{R}} m(y) q_\varepsilon(t, y) dy - m(\overline{Z}(t)) \Big| = O(\varepsilon^{1-\delta}).$$

Step 2. We next prove (11). To this end, we first notice, thanks to (42) and (43), that we have
$$\varepsilon^2 \frac{d}{dt} \rho_\varepsilon = \rho_\varepsilon \big(r - m(\overline{Z}) - \kappa \rho_\varepsilon + O(\varepsilon^{1-\delta})\big).$$

We then define
$$J_\varepsilon := \frac{1}{\rho_\varepsilon}.$$

Replacing this in the equation on ρ_ε we obtain the following linear equation
$$\varepsilon^2 \frac{d}{dt} J_\varepsilon = \kappa - J_\varepsilon \big(r - m(\overline{Z}) + O(\varepsilon^{1-\delta})\big).$$

We solve this equation to obtain
$$J_\varepsilon(t) = J_\varepsilon(0) e^{-\frac{1}{\varepsilon^2} \int_0^t \big(r - m(\overline{Z}(s)) + O(\varepsilon^{1-\delta})\big) ds} + \frac{\kappa}{\varepsilon^2} \int_0^t e^{-\frac{1}{\varepsilon^2} \int_\tau^t \big(r - m(\overline{Z}(s)) + O(\varepsilon^{1-\delta})\big) ds} d\tau. \qquad (44)$$

We next use (H5), (7) and (H2) to obtain that there exists a positive constant a_0 such that, for ε small enough,
$$r - m(\overline{Z}(s)) + O(\varepsilon^{1-\delta}) \geq a_0, \qquad \text{for all } s \in \mathbb{R}^+.$$

We deduce that
$$J_\varepsilon(0) e^{-\frac{1}{\varepsilon^2} \int_0^t \big(r - m(\overline{Z}(s)) + O(\varepsilon^{1-\delta})\big) ds} \leq \frac{1}{\rho_\varepsilon(0)} e^{-\frac{a_0 t}{\varepsilon^2}}.$$

To control the second term in the r.h.s. of (44) we let $t \geq \varepsilon^\beta$, with $\beta \in (1, 2)$, and we write
$$\frac{\kappa}{\varepsilon^2} \int_0^t e^{-\frac{1}{\varepsilon^2} \int_\tau^t \big(r - m(\overline{Z}(s)) + O(\varepsilon^{1-\delta})\big) ds} d\tau = \frac{\kappa}{\varepsilon^2} \int_0^{t-\varepsilon^\beta} e^{-\frac{1}{\varepsilon^2} \int_\tau^t \big(r - m(\overline{Z}(s)) + O(\varepsilon^{1-\delta})\big) ds} d\tau$$
$$+ \frac{\kappa}{\varepsilon^2} \int_{t-\varepsilon^\beta}^t e^{-\frac{1}{\varepsilon^2} \int_\tau^t \big(r - m(\overline{Z}(s)) + O(\varepsilon^{1-\delta})\big) ds} d\tau$$
$$= J_{\varepsilon,1} + J_{\varepsilon,2}.$$

We control each of these terms separately. To control the first term we write
$$J_{\varepsilon,1} = \frac{\kappa}{\varepsilon^2} \int_0^{t-\varepsilon^\beta} e^{-\frac{1}{\varepsilon^2} \int_\tau^t \big(r - m(\overline{Z}(s)) + O(\varepsilon^{1-\delta})\big) ds} d\tau \leq \frac{\kappa}{\varepsilon^2} \int_0^{t-\varepsilon^\beta} e^{-\frac{1}{\varepsilon^2} a_0(t-\tau)} d\tau$$
$$\leq \frac{\kappa}{a_0} \big(e^{-\frac{a_0}{\varepsilon^{2-\beta}}} - e^{-\frac{a_0 t}{\varepsilon^2}}\big) \leq \frac{\kappa}{a_0} e^{-\frac{a_0}{\varepsilon^{2-\beta}}}.$$

18

To control $J_{\varepsilon,2}$ we use the fact that

$$|m(\overline{Z}(s)) - m(\overline{Z}(t))| \leq C\varepsilon^{\beta}, \qquad \text{for all } s \in [t - \varepsilon^{\beta}, t].$$

Up to changing the constant C, we deduce that, since $\beta > 1$:

$$J_{\varepsilon,2}(t) = \frac{\kappa}{\varepsilon^2} \int_{t-\varepsilon^{\beta}}^{t} e^{-\frac{1}{\varepsilon^2} \int_{\tau}^{t} \left(r - m(\overline{Z}(s)) + O(\varepsilon^{1-\delta})\right) ds} d\tau$$

$$\leq \frac{\kappa}{\varepsilon^2} \int_{t-\varepsilon^{\beta}}^{t} e^{-\frac{1}{\varepsilon^2} \left(r - m(\overline{Z}(t)) - C\varepsilon^{1-\delta}\right)(t-\tau)} d\tau$$

$$\leq \frac{\kappa}{r - m(\overline{Z}(t)) - C\varepsilon^{1-\delta}}.$$

We similarly find

$$J_{\varepsilon,2}(t) = \frac{\kappa}{\varepsilon^2} \int_{t-\varepsilon^{\beta}}^{t} e^{-\frac{1}{\varepsilon^2} \int_{\tau}^{t} \left(r - m(\overline{Z}(s)) + O(\varepsilon^{1-\delta})\right) ds} d\tau$$

$$\geq \frac{\kappa}{\varepsilon^2} \int_{t-\varepsilon^{\beta}}^{t} e^{-\frac{1}{\varepsilon^2} \left(r - m(\overline{Z}(t)) + C\varepsilon^{1-\delta}\right)(t-\tau)} d\tau$$

$$= \frac{\kappa}{r - m(\overline{Z}(t)) + C\varepsilon^{1-\delta}} \left(1 - e^{-\frac{r - m(\overline{Z}(t)) + C\varepsilon^{1-\delta}}{\varepsilon^{2-\beta}}}\right).$$

Using again (H5), (7) and (H2) , when ε is chosen small enough, we have

$$r - m(\overline{Z}(t)) + C\varepsilon^{1-\delta} \geq a_0/2,$$

which implies that

$$J_{\varepsilon,2}(t) \geq \frac{\kappa}{r - m(\overline{Z}(t)) + C\varepsilon^{1-\delta}} \left(1 - e^{-\frac{a_0}{2\varepsilon^{2-\beta}}}\right),$$

and hence

$$\left| J_{\varepsilon,2}(t) - \frac{\kappa}{r - m(\overline{Z}(t))} \right| = O(\varepsilon^{1-\delta}).$$

Combining the inequalities above we obtain that

$$\left| J_{\varepsilon}(t) - \frac{\kappa}{r - m(\overline{Z}(t))} \right| \leq O(\varepsilon^{1-\delta}) + \left(\frac{1}{\rho_{\varepsilon}(0)} + \frac{\kappa}{a_0}\right) e^{-\frac{a_0}{\varepsilon^{2-\beta}}}, \qquad \text{for all } t \geq \varepsilon^{\beta}.$$

We deduce that, for all $t \geq \varepsilon^{\beta}$,

$$\left| \rho_{\varepsilon}(t) - \frac{r - m(\overline{Z}(t))}{\kappa} \right| \leq \rho_{\varepsilon}(t) \left(\frac{r - m(\overline{Z}(t))}{\kappa}\right) \left(O(\varepsilon^{1-\delta}) + \left(\frac{1}{\rho_{\varepsilon}(0)} + \frac{\kappa}{a_0}\right) e^{-\frac{a_0}{\varepsilon^{2-\beta}}}\right).$$

We next apply the maximum principle to (42) and use Assumption (H6) to obtain that

$$\rho_{\varepsilon}(t) \leq \max(\rho_M, r), \qquad \text{for all } t \geq 0.$$

We also recall from Assumption (H6) that $\rho_{\varepsilon}(0) \geq \rho_m > 0$. We conclude that

$$\left| \rho_{\varepsilon}(t) - \frac{r - m(\overline{Z}(t))}{\kappa} \right| = O(\varepsilon^{1-\delta}), \qquad \text{for all } t \geq \varepsilon^{\beta}.$$

(b) **The convergence of n_{ε} as $\varepsilon \to 0$.** From the definition of q_{ε} and thanks to the convergence of ρ_{ε} which was established in the paragraph above, it is enough to prove that q_{ε} converges in $C((0,\infty); \mathcal{M}^1(\mathbb{R}))$ to $\partial(x - \overline{Z}(t))$. We next notice from the definition of g_{ε} that it converges in

$C((0,\infty);\mathcal{M}^1(\mathbb{R}))$ to $\partial(x - \overline{Z}(t))$. The idea is then to use the estimate (10) to prove that q_ε and g_ε have the same limit.

Let $\phi : \mathbb{R} \to \mathbb{R}$ be a Lipschitz continuous function. We compute, for all $t \in \mathbb{R}^+$ and using (10),

$$\left| \int_{\mathbb{R}} \phi(x)(q_\varepsilon(t,x) - g_\varepsilon(t,x))dx \right| \leq \|\phi'\|_{L^\infty} W_1(q_\varepsilon, g_\varepsilon) \leq K\|\phi'\|_{L^\infty} \left(W_1(q_{\varepsilon,0}, g_{\varepsilon,0}) + \varepsilon^{1-\delta} \right). \quad (45)$$

We next estimate $W_1(q_{\varepsilon,0}, g_{\varepsilon,0})$. To this end, we will use the following triangular inequality

$$W_1(q_{\varepsilon,0}, g_{\varepsilon,0}) \leq W_1(q_{\varepsilon,0}, \partial(\cdot - x_0)) + W_1(g_{\varepsilon,0}, \partial(\cdot - x_0)).$$

Again from the definition of g_ε, it is immediate that the second term in the r.h.s. of the inequality above tends to 0, as $\varepsilon \to 0$. We prove the same property for the first term. We first compute, using the Cauchy-Schwartz inequality,

$$W_1(q_{\varepsilon,0}, \partial(\cdot - M_{\varepsilon,1}(0))) \leq \int |y - M_{\varepsilon,1}(0)|q_{\varepsilon,0}(y)dy \leq \left(\int |y - M_{\varepsilon,1}(0)|^2 q_{\varepsilon,0}(y)dy \right)^{1/2} = \sqrt{M_{\varepsilon,2}(0)}.$$

We then use Assumption (H4) to obtain that

$$W_1(q_{\varepsilon,0}, \partial(\cdot - M_{\varepsilon,1}(0))) \to 0, \qquad \text{as } \varepsilon \to 0.$$

Combining this property with Assumption (H6) we deduce that

$$W_1(q_{\varepsilon,0}, \partial(\cdot - x_0)) \to 0, \qquad \text{as } \varepsilon \to 0.$$

and hence

$$W_1(q_{\varepsilon,0}, g_{\varepsilon,0}) \to 0, \qquad \text{as } \varepsilon \to 0.$$

This concludes the proof thanks to (45).

5 A comparison with a model with an asexual reproduction term

A closely related model, considering an asexual reproduction and including mutations, is written

$$\begin{cases} \varepsilon\partial_t n_\varepsilon = pM_\varepsilon[n_\varepsilon] + (r - m(x) - \kappa\rho_\varepsilon(t)) \, n_\varepsilon, \\ \rho_\varepsilon(t) = \int_{\mathbb{R}} n_\varepsilon(t,y)dy, \\ n_\varepsilon(0,x) = n_{\varepsilon,0}(x). \end{cases} \quad (46)$$

where

$$M_\varepsilon[n_\varepsilon](x) = \int G_\varepsilon(x - y)\big(n_\varepsilon(t,y) - n_\varepsilon(t,x)\big)dy, \qquad G_\varepsilon(x) = \frac{1}{\varepsilon}G(\frac{x}{\varepsilon}).$$

A typical example of G_ε is a normal distribution with variance of order ε^2, for instance Γ_ε. Here, since each offspring has a single parent, contrary to model (1) the reproduction is modeled via linear terms. The parameter r corresponds to the clonal birth rate of individuals of trait x and $m(x)$ represents a trait-dependent death rate. The term M_ε models the mutations which arise with rate p. The function G_ε corresponds to the mutation law. The small parameter ε is introduced to rescale the mutation size, such that we consider small mutational effects. It is shown in previous works [25, 1, 17] and under suitable assumptions that, as $\varepsilon \to 0$, the phenotypic density n_ε has a similar behavior to the one observed in Theorem 2 in the case of model (1). It indeed converges to a measure n, with

$$n(t) = \rho(t)\partial(x - \overline{x}(t)).$$

Under suitable assumptions [17, 23], the dynamics of $\bar{x}(t)$ can also be described via an ordinary differential equation

$$\frac{d}{dt}\bar{x}(t) = -A(t)m'(\bar{x}(t)).$$

This equation can be compared to (9), with the difference that here the coefficient in front of the gradient of the growth rate, $A(t)$, is time dependent. This dependency with respect to time comes from the fact that the dominant term of the variance of the phenotypic distribution corresponding to (46) varies with time, unlike to the model (1) (see (13)). The dynamics of this coefficient $A(t)$ is rather complex and can be obtained via a Hamilton-Jacobi equation with constraint. The usual method to study (46) is indeed via an approach based on Hamilton-Jacobi equations, the first step being to perform a Hopf-Cole transformation

$$n_\varepsilon(t, x) = \exp\left(\frac{u_\varepsilon(t, x)}{\varepsilon}\right).$$

It can be shown [25, 1] that u_ε converges, as $\varepsilon \to 0$, to a continuous function u, which solves the following Hamilton-Jacobi equation in the viscosity sense

$$\begin{cases} \partial_t u(t, x) = p \int_{\mathbb{R}} G(h) e^{\frac{\partial}{\partial x} u(t,x)h} dh + r - m(x) - \kappa\rho(t), & (t, x) \in \mathbb{R}^+ \times \mathbb{R}, \\ \max_{x \in \mathbb{R}} u(t, x) = 0, & x \in \mathbb{R}, \\ u(0, x) = u_0(x), \end{cases} \quad (47)$$

where ρ is the limit of ρ_ε as $\varepsilon \to 0$. The function u relates to the phenotypic density n via the following relation

$$\text{supp } n(t, \cdot) \subset \{x \,|\, u(t, x) = 0\}.$$

Under some global concavity assumptions [17, 23] one can prove that the set above has a single point $\bar{x}(t)$ and hence

$$n(t, x) = \rho(t)\partial(x - \bar{x}(t)).$$

The coefficient $A(t)$ is then given by the following formula

$$A(t) = -\frac{1}{\frac{\partial^2}{\partial x^2} u(t, \bar{x}(t))}.$$

One can wonder whether the moment-based approach introduced in this article can allow to obtain similar results in the case of model (46) providing in this way a direct proof, rather than using Hamilton-Jacobi equations. We will explain below why the complexity of the dynamics in the model with asexual reproduction does not allow to apply the direct method presented above.

In order to keep the computations more straightforward and to highlight the main arguments we will focus on the following particular death rate

$$m(x) = sx^2.$$

We also assume that G is symmetric, that is

$$G(x) = G(-x), \qquad \text{for all } x \in \mathbb{R}.$$

We then define q_ε similarly to Section 1:

$$q_\varepsilon(t, x) = \frac{n_\varepsilon(t, x)}{\rho_\varepsilon(t)}.$$

Replacing this in (46) we obtain

$$\begin{cases} \varepsilon\partial_t q_\varepsilon = pM_\varepsilon[q_\varepsilon] + \left(-m(x) + \int_{\mathbb{R}} m(y)q_\varepsilon(\cdot, y)dy\right) q_\varepsilon, \\ q_\varepsilon(0, x) = \frac{n_{\varepsilon,0}(x)}{\rho_\varepsilon(0)}. \end{cases} \quad (48)$$

21

We also define $M_{\varepsilon,i}$ and $M_{\varepsilon,i}^c$ similarly to Section 1. We next derive the equations on the first and the second order moments $M_{\varepsilon,1}$ and $M_{\varepsilon,2}^c$, similarly to Section 2.

To derive an equation on $M_{\varepsilon,1}$, we first compute

$$\int_{\mathbb{R}} x M_{\varepsilon}[q_{\varepsilon}](x) dx = \int_{\mathbb{R}} \int_{\mathbb{R}} x \big(q_{\varepsilon}(\cdot, x + \varepsilon h) - q_{\varepsilon}(\cdot, x) \big) G(h) dh dx$$

$$= \int_{\mathbb{R}} \int_{\mathbb{R}} (x + \varepsilon h) q_{\varepsilon}(\cdot, x + \varepsilon h) G(h) dh dx - \int_{\mathbb{R}} \int_{\mathbb{R}} x q_{\varepsilon}(\cdot, x) G(h) dh dx$$

$$- \varepsilon \int_{\mathbb{R}} \int_{\mathbb{R}} h q_{\varepsilon}(\cdot, x + \varepsilon h) G(h) dh dx$$

$$= - \varepsilon \int_{\mathbb{R}} h G(h) dh = 0.$$

Note that here we have used the fact that G is symmetric. We next multiply (48) by x and integrate with respect to x to obtain

$$\varepsilon \dot{M}_{\varepsilon,1} = -s \int_{\mathbb{R}} x^3 q_{\varepsilon}(\cdot, x) dx + s \int_{\mathbb{R}} x^2 q_{\varepsilon}(\cdot, x) dx\, M_{\varepsilon,1}$$

$$= -s \int_{\mathbb{R}} (x - M_{\varepsilon,1} + M_{\varepsilon,1})^3 q_{\varepsilon}(\cdot, x) dx + s \int_{\mathbb{R}} (x - M_{\varepsilon,1} + M_{\varepsilon,1})^2 q_{\varepsilon}(\cdot, x) dx\, M_{\varepsilon,1}$$

$$= - 2s M_{\varepsilon,1} M_{\varepsilon,2}^c - s M_{\varepsilon,3}^c$$

$$= - m'(M_{\varepsilon,1}) M_{\varepsilon,2}^c - s M_{\varepsilon,3}^c.$$

Therefore, $M_{\varepsilon,1}$ solves an equation with similar structure to (22), that is

$$\varepsilon \dot{M}_{\varepsilon,1} + m'(M_{\varepsilon,1}) M_{\varepsilon,2}^c = F_1 = -s M_{\varepsilon,3}^c,$$

where the r.h.s. is a priori of upper order term in ε. The equation on $M_{\varepsilon,1}$ is not indeed the one that leads to difficulty.

We next compute the equation solved by $M_{\varepsilon,2}^c$. To this end, we first compute

$$\int_{\mathbb{R}} (x - M_{\varepsilon,1})^2 M_{\varepsilon}[q_{\varepsilon}](x) dx = \int_{\mathbb{R}} \int_{\mathbb{R}} (x - M_{\varepsilon,1})^2 \big(q_{\varepsilon}(\cdot, x + \varepsilon h) - q_{\varepsilon}(\cdot, x) \big) G(h) dh dx$$

$$= \int_{\mathbb{R}} \int_{\mathbb{R}} x^2 \big(q_{\varepsilon}(\cdot, x + \varepsilon h) - q_{\varepsilon}(\cdot, x) \big) G(h) dh dx$$

$$= \int_{\mathbb{R}} \int_{\mathbb{R}} (x + \varepsilon h - \varepsilon h)^2 q_{\varepsilon}(\cdot, x + \varepsilon h) G(h) dh dx - \int_{\mathbb{R}} \int_{\mathbb{R}} x^2 q_{\varepsilon}(\cdot, x) G(h) dh dx$$

$$= - 2\varepsilon M_{\varepsilon,1} \int_{\mathbb{R}} h G(h) dh + \varepsilon^2 \int_{\mathbb{R}} h^2 G(h) dh$$

$$= \varepsilon^2 \sigma,$$

where σ denotes the variance of the distribution G. We next multiply (48) by $(x - M_{\varepsilon,1})^2$ and integrate with respect to x to obtain

$$\varepsilon \dot{M}_{\varepsilon,2}^c = p \sigma \varepsilon^2 - s \int_{\mathbb{R}} (x - M_{\varepsilon,1})^2 x^2 q_{\varepsilon}(\cdot, x) dx + s \int_{\mathbb{R}} x^2 q_{\varepsilon}(\cdot, x) dx\, M_{\varepsilon,2}^c$$

$$= p \sigma \varepsilon^2 - s \int_{\mathbb{R}} (x - M_{\varepsilon,1})^2 (x - M_{\varepsilon,1} + M_{\varepsilon,1})^2 q_{\varepsilon}(\cdot, x) dx$$

$$+ s \int_{\mathbb{R}} (x - M_{\varepsilon,1} + M_{\varepsilon,1})^2 q_{\varepsilon}(\cdot, x) dx\, M_{\varepsilon,2}^c.$$

We deduce that

$$\varepsilon \dot{M}_{\varepsilon,2}^c = p \sigma \varepsilon^2 - s M_{\varepsilon,4}^c - 2s M_{\varepsilon,3}^c M_{\varepsilon,1} + s (M_{\varepsilon,2}^c)^2. \tag{49}$$

22

Here the structure of the equation is different from (24). The dissipative term that allowed us to close the equation is not present here (*i.e.* the equation is not written in the form $\varepsilon \dot{M}^c_{\varepsilon,2} + CM^c_{\varepsilon,2} = F$ with $C \in (0,\infty)$). Moreover, all the terms on the r.h.s. of (49) may potentially be of the same order and could equally contribute to the dynamics of $M^c_{\varepsilon,2}$ (see [22, 21, 13]). We hence do not expect to be able to close the equations of the moments at this stage. The dynamics of moments of order 3 and 4 are indeed driven by the solution to the Hamilton-Jacobi equation (47) [21, 13].

A The Cauchy problem; the proof of Proposition 1

In this section, we focus on the Cauchy theory analyzed on Proposition 1. This issue is tackled previously in [29], when the trait dependency is only included in the reproduction kernel T_ε, or in [27] on bounded domains. For completeness, we explain here how the arguments may be adapted to our framework. We shall in particular detail the issues related to the boundedness of m. This rate is classically assumed quadratic and thus unbounded while this makes the analysis surprisingly more difficult.

In the whole section we fix $\varepsilon = 1$ since this has no influence on our computations and we drop all ε in notations. Also, we focus on the case of a smooth mortality rate m satisfying (H0) and (H3) (the other assumptions have no influence here). The outline of the appendix is as follows. Firstly, we provide a definition of solution to (1) and analyze its properties. Then, we show that existence of solutions in case m is bounded reduces to a classical application of Cauchy-Lipschitz theorem and tackle the existence of solutions in case m is unbounded. We conclude by obtaining uniqueness of solutions with a sufficiently large number of bounded moments.

A.1 What is a solution to (1) ?

From now on $k \in (1,\infty)$ is fixed. We denote then $w_k(x) = (1+|x|)^k$ and

$$X_k := \left\{ n \in L^1(\mathbb{R}) \text{ s.t. } \int_{\mathbb{R}} w_k(x)|n(x)|\mathrm{d}x < \infty \right\}.$$

This is a Banach space endowed with the norm:

$$\|n\|_{X_k} := \int_{\mathbb{R}} w_k(x)|n(x)|\mathrm{d}x.$$

Inside X_k we denote:

$$\mathcal{U}_k := \left\{ n \in X_k \text{ s.t. } \int_{\mathbb{R}} n(x)\mathrm{d}x \neq 0 \right\}.$$

Clearly $\mathcal{U}_k$ is an open subset of X_k. We denote also below $C_b(\mathcal{O})$ the set of bounded continuous functions on $\mathcal{O}$ endowed with its classical L^∞-norm. We recall the definition of the mapping T:

$$T[n(t,\cdot)](x) = \int_{\mathbb{R}} \int_{\mathbb{R}} \Gamma\left(x - \frac{(y+y')}{2}\right) n(t,y) \frac{n(t,y')}{\rho(t)} dy dy'.$$

For the following analysis, we define solutions to (1) as follows

Definition 5 *Given $n_0 \in X_k \cap C_b(\mathbb{R})$ and $T_0 > 0$, we say that n is a solution to (1) on $(0,T_0)$ if*

- *$n \in C_b([0,T_0] \times \mathbb{R}) \cap L^\infty((0,T_0); X_k)$ with $n(t,\cdot) \in \mathcal{U}_k$ for all $t \in (0,T_0)$,*
- *for all $x \in \mathbb{R}$ the restriction $n(\cdot,x)$ solves (1) in $\mathcal{D}'((0,T_0))$.*

We say that n is a global solution if its restriction on $[0,T_0]$ is a solution on $(0,T_0)$ for arbitrary $T_0 > 0$.

We point out that we do not assume *a priori* that $n_0 \geq 0$ here. However, we shall restrict in the second part of the analysis to such initial data. A straightforward interpolation argument entails also that a solution n satisfies $n \in C([0, T_0]; X_{\tilde{k}})$ for arbitrary $\tilde{k} < k$. We also point out that the introduction of distributions $\mathcal{D}'((0, T_0))$ is just for convenience here. Indeed, under the (satisfied) condition that $n \in C([0, T_0]; L^1(\mathbb{R}))$ has non-zero integral, we observe that $t \mapsto T[n(t, \cdot)](x) \in C([0, T_0])$ so that equation $n(\cdot, x) \in C^1([0, T_0])$ and (1) holds in a classical sense.

To analyze the Cauchy problem, we will rely on the following remark

Lemma 6 *The mapping $T : \mathcal{U}_k \cap C_b(\mathbb{R}) \to X_k \cap C_b(\mathbb{R})$ is positive and locally lipschitz. More precisely,*

i) given $(n, \tilde{n}) \in (\mathcal{U}_k \cap C_b(\mathbb{R}))^2$, there exists a constant K depending increasingly on :

$$\max \left(\frac{\|n\|_{X_k}}{\left| \int_{\mathbb{R}} n(z) \mathrm{d}z \right|}, \frac{\|\tilde{n}\|_{X_k}}{\left| \int_{\mathbb{R}} \tilde{n}(z) \mathrm{d}z \right|} \right)$$

such that:

$$\|T[n] - T[\tilde{n}]\|_{X_k \cap C_b(\mathbb{R})} \leq K \|n - \tilde{n}\|_{X_k},$$

ii) if $n \geq 0$ we have $T[n] \geq 0$.

Proof. Property ii) follows from the fact that all the quantities in the integrals are nonnegative. For property i), we first get:

$$\|T[n] - T[\tilde{n}]\|_{X_k} \leq \int_{\mathbb{R}} \int_{\mathbb{R}} \int_{\mathbb{R}} \Gamma \left(x - \frac{(y + y')}{2} \right) w_k(x) \left| n(y) \frac{n(y')}{\int_{\mathbb{R}} n(z) \mathrm{d}z} - \tilde{n}(y) \frac{\tilde{n}(y')}{\int_{\mathbb{R}} \tilde{n}(z) \mathrm{d}z} \right| \mathrm{d}y \mathrm{d}y' \mathrm{d}x.$$

We next notice that

$$w_k(x) \leq C_k \left(w_k \left(x - \frac{(y + y')}{2} \right) + w_k(y) + w_k(y') \right),$$

with C_k a constant depending on k. Moreover, we have

$$n(y) \frac{n(y')}{\int_{\mathbb{R}} n(z) \mathrm{d}z} - \tilde{n}(y) \frac{\tilde{n}(y')}{\int_{\mathbb{R}} \tilde{n}(z) \mathrm{d}z} = \frac{n(y)}{\int_{\mathbb{R}} n(z) \mathrm{d}z} (n(y') - \tilde{n}(y')) + \frac{\tilde{n}(y')}{\int_{\mathbb{R}} \tilde{n}(z) \mathrm{d}z} (n(y) - \tilde{n}(y))$$
$$+ \frac{n(y)}{\int_{\mathbb{R}} n(z) \mathrm{d}z} \frac{\tilde{n}(y')}{\int_{\mathbb{R}} \tilde{n}(z) \mathrm{d}z} \int_{\mathbb{R}} (n(z) - \tilde{n}(z)) \mathrm{d}z$$

We deduce, since $w_k \Gamma \in L^1(\mathbb{R})$, that there is a constant $C_{k,\Gamma}$ depending further on Γ for which:

$$\|T[n] - T[\tilde{n}]\|_{X_k} \leq C_{k,\Gamma} \int_{\mathbb{R}} \int_{\mathbb{R}} (w_k(y) + w_k(y')) \left| \frac{n(y)}{\int_{\mathbb{R}} n(z) \mathrm{d}z} (n(y') - \tilde{n}(y')) \right| \mathrm{d}y \mathrm{d}y'$$
$$+ C_{k,\Gamma} \int_{\mathbb{R}} \int_{\mathbb{R}} (w_k(y) + w_k(y')) \left| \frac{\tilde{n}(y')}{\int_{\mathbb{R}} \tilde{n}(z) \mathrm{d}z} (n(y) - \tilde{n}(y)) \right| \mathrm{d}y \mathrm{d}y'$$
$$+ C_{k,\Gamma} \int_{\mathbb{R}} \int_{\mathbb{R}} (w_k(y) + w_k(y')) \left| \frac{n(y)}{\int_{\mathbb{R}} n(z) \mathrm{d}z} \frac{\tilde{n}(y')}{\int_{\mathbb{R}} \tilde{n}(z) \mathrm{d}z} \int_{\mathbb{R}} (n(z) - \tilde{n}(z)) \mathrm{d}z \right| \mathrm{d}y \mathrm{d}y'$$
$$\leq C_{k,\Gamma} \left(\frac{\|n\|_{X_k}}{\left| \int_{\mathbb{R}} n(z) \mathrm{d}z \right|} + \frac{\|\tilde{n}\|_{X_k}}{\left| \int_{\mathbb{R}} \tilde{n}(z) \mathrm{d}z \right|} + \frac{\|n\|_{X_k}}{\left| \int_{\mathbb{R}} n(z) \mathrm{d}z \right|} \frac{\|\tilde{n}\|_{X_k}}{\left| \int_{\mathbb{R}} \tilde{n}(z) \mathrm{d}z \right|} \right) \|n - \tilde{n}\|_{X_k}$$
$$\leq C_{k,\Gamma} \frac{\|n\|_{X_k}}{\left| \int_{\mathbb{R}} n(z) \mathrm{d}z \right|} \frac{\|\tilde{n}\|_{X_k}}{\left| \int_{\mathbb{R}} \tilde{n}(z) \mathrm{d}z \right|} \|n - \tilde{n}\|_{X_k}.$$

24

We obtain with similar and even simpler computations that there is a constant $C_{k,\Gamma}$ for which:

$$\|T[n] - T[\tilde{n}]\|_{C_b(\mathbb{R})} \leq C_{k,\Gamma} \frac{\|n\|_{X_k}}{\left|\int_{\mathbb{R}} n(z)\mathrm{d}z\right|} \frac{\|\tilde{n}\|_{X_k}}{\left|\int_{\mathbb{R}} \tilde{n}(z)\mathrm{d}z\right|} \|n - \tilde{n}\|_{X_k}.$$

$\square$

This lemma entails that the mapping T extends to a continuous mapping $C([0,T_0];X_k) \cap C_b([0,T_0] \times \mathbb{R})) \to C([0,T_0];X_k) \cap C_b([0,T_0] \times \mathbb{R})$. We point out also that a straight forward adaptation of the above proof entails the following pointwise bound for nonnegative $n \in \mathcal{U}_k$:

$$|T[n](x)| \leq \frac{C_k}{w_k(x)} \|n\|_{X_k} \tag{50}$$

where C_k is a constant independent of n. In particular, if n is a nonnegative solution to (1) on $(0, T_0)$ with initial data $n_0 \in \mathcal{U}_k$ we have that $\partial_t n(t,x) \in C([0,T_0])$ for all $x \in \mathbb{R}$ and combining the latter inequality with (2) we infer that:

$$|\partial_t n(t,x)| \leq \frac{C_k}{w_k(x)} \|n(t,\cdot)\|_{X_k} + (A_m w_{p+2}(x) + \kappa\rho)n(t,x). \tag{51}$$

Given $j < k - (p+3)$, classical parameter integral arguments entail in particular that

$$M_j(t) := \int_{\mathbb{R}} x^j n(t,x)\mathrm{d}x \in C^1([0,T_0])$$

with:

$$\dot{M}_j(t) = \int_{\mathbb{R}} x^j \partial_t n(t,x)\mathrm{d}x \quad \forall\, t \in [0,T_0].$$

We use without mention this regularity property in the main body of the article.

A.2   Existence of solutions

In case m is bounded, the mapping

$$n \mapsto \left(m + \kappa \int_{\mathbb{R}} n\right) n$$

is clearly locally lipschitz on $\mathcal{U}_k$. Combining this remark with Lemma 6, we obtain via a direct application of the Cauchy-Lipschitz theorem that, given $n_0 \in \mathcal{U}_k$. there exists T_0 depending decreasingly on

$$\|n_0\|_{X_k} \left(1 + \frac{1}{\int_{\mathbb{R}} n_0}\right)$$

such that for arbitrary $\tilde{T}_0 < T_0$ there is a unique solution n to (1) on $(0, \tilde{T}_0)$. We point out that the Cauchy-Lipschitz theorem yields also the further regularity $n \in C^1([0,\tilde{T}_0); X_k)$. Furthermore, it entails the existence of a unique non-extendable solution defined on $[0, \tilde{T}_0)$ with the blow-up alternative:

- either $\tilde{T}_0 = +\infty$
- or $\tilde{T}_0 < \infty$ and

$$\limsup_{t \to \tilde{T}_0} \frac{1}{\int_{\mathbb{R}} n(t,x)\mathrm{d}x} + \|n(t,\cdot)\|_{X_k} = +\infty$$

25

A.2.1 Quantitative estimates and global existence result

To yield global existence of solution we now restrict to non-negative initial data. We prove *a priori* estimates satisfied by solutions to (1). We point out that we derive estimates satisfied by any solution to (1) according to Definition 5 not necessarily the ones obtained through the Cauchy-Lipschitz theorem above. Moreover, we derive estimates that are valid for mortality rates m that possibly diverge at infinity. In the remainder of this subsection, we assume that $n_0 \in \mathcal{U}_k$, for some $k > 0$, is nonnegative. We note that n_0 has non-zero integral so that it may not vanish identically on $\mathbb{R}$. We assume that n is a solution to (1) on $(0, T)$ according to Definition 5.

Firstly, we show that positivity propagates in the solution:

Lemma 7 *We assume that $n_0(x) \geq 0$ for all $x \in \mathbb{R}$. There holds:*

$$n(t, x) \geq 0, \quad \forall\, (t, x) \in [0, T] \times \mathbb{R}.$$

Proof. We show that n remains non-negative on a small time interval $[0, T_0]$. The result extends to $[0, T]$ with a standard connectedness argument. Indeed, our argument propagates positivity as long as $n(t, \cdot)$ does not vanish identically. However, if $n(t, \cdot) \equiv 0$ then the solution blows up in t. We split the proof into three steps.

Step 1. We obtain first a general property for the operator T on $L^1(\mathbb{R})$. Namely, given $f \in L^1(\mathbb{R}) \cap C_b(\mathbb{R})$ with positive integral, we have clearly via standard parameter-integral results that $T[f] \in C_b(\mathbb{R})$. We prove that:

$$\|(T[f])_-\|_{L^\infty(\mathbb{R})} \leq \frac{4\|f\|_{L^1(\mathbb{R})}}{\int_{\mathbb{R}} f(x)\mathrm{d}x}\|(f)_-\|_{L^\infty(\mathbb{R})}, \tag{52}$$

where $(\cdot)_-$ denotes the negative part of continuous functions. Indeed, if we denote:

$$D_- := \{x \in \mathbb{R} \text{ s.t. } f(x) < 0\} \qquad D_+ := \{x \in \mathbb{R} \text{ s.t. } f(x) \geq 0\}$$

we remark that $f(y)f(y') \geq 0$ whenever $(y, y') \in D_- \times D_- \cup D_+ \times D_+$. Therefore, we have:

$$
\begin{aligned}
(T[f])_-(x) &\leq -\int_{D_- \times D_+} \Gamma\left(x - \frac{y + y'}{2}\right)\frac{f(y)f(y')}{\int_{\mathbb{R}} f(z)\mathrm{d}z}\mathrm{d}y\mathrm{d}y' + sym.\\
&\leq \frac{1}{\int_{\mathbb{R}} f(z)\mathrm{d}z}\left(\int_{D_- \times D_+} \Gamma\left(x - \frac{y + y'}{2}\right)|f(y)||f(y')|\mathrm{d}y\mathrm{d}y + sym.\right)\\
&\leq \frac{1}{\int_{\mathbb{R}} f(z)\mathrm{d}z}\left(\int_{D_- \times \mathbb{R}} \Gamma\left(x - \frac{y + y'}{2}\right)\|(f)_-\|_{L^\infty(\mathbb{R})}|f(y')|\mathrm{d}y\mathrm{d}y + sym.\right)
\end{aligned}
$$

We conclude then by recalling that:

$$\int_{\mathbb{R}} \Gamma\left(x - \frac{y + y'}{2}\right)\mathrm{d}y = 2.$$

Step 2. We consider now n_0 and its associated solution n. Thanks to the remark after **Definition 5**, we note that $n \in C([0, T]; L^1(\mathbb{R}))$ with $n(0, \cdot) = n_0$. In particular, we can construct a small time T_0 such that:

$$\|n(t, \cdot)\|_{L^1(\mathbb{R})} \leq 2\|n_0\|_{L^1(\mathbb{R})} \qquad 0 < \frac{\int_{\mathbb{R}} n_0(z)\mathrm{d}z}{2} \leq \int_{\mathbb{R}} n(t, z)\mathrm{d}z \qquad \forall\, t \in [0, T_0]. \tag{53}$$

We choose $\Lambda_0 > 32r$. Since $n_0 \geq 0$ we note that we have in particular:

$$\Lambda_0 > 32r \frac{\|n_0\|_{L^1(\mathbb{R})}}{\int_\mathbb{R} n_0(z) \mathrm{d}z}.$$

We construct then:

$$\tilde{n}(t,x) = n(t,x) \exp(-\Lambda_0 t) \qquad \forall t \in [0,T].$$

Like n, we have that $\tilde{n} \in C_b([0,T] \times \mathbb{R}) \cap C([0,T]; L^1(\mathbb{R}))$ and that $\tilde{n}(\cdot, x) \in C^1([0,T])$ for all $x \in \mathbb{R}$. In particular, for arbitrary $x \in \mathbb{R}$, there holds:

$$\partial_t \tilde{n}(t,x) = (\partial_t n(t,x) - \Lambda_0 n(t,x)) \exp(-\Lambda_0 t).$$

Replacing $\partial_t n$ and arguing that $T[\cdot]$ is 1-homogeneous, we obtain that:

$$\partial_t \tilde{n}(t,x) = rT[\tilde{n}(t,\cdot)](x) - (m(x) + \kappa\rho(t) + \Lambda_0)\tilde{n}(t,x) \text{ on } [0,T]. \tag{54}$$

Step 3. We assume now that n becomes strictly negative on $[0,T_0]$ or equivalently that $\tilde{n}$ takes strictly negative values on $[0,T_0]$. Since $\tilde{n} \in C_b([0,T_0] \times \mathbb{R})$, we have:

$$\max_{[0,T_0] \times \mathbb{R}} (\tilde{n})_-(t,x) = \delta > 0,$$

and we can find $(t_0, x_0) \in (0, T_0) \times \mathbb{R}$ so that $\tilde{n}(t_0, x_0) \leq -\delta/2$. Since $\tilde{n}(\cdot, x_0) \in C^1([0, T_0])$ satifies $\tilde{n}(0, x_0) = n_0(x_0) \geq 0$ we can then construct a time $t_- \in (0, T_0)$ such that $\tilde{n}(t, x_0) \geq \tilde{n}(t_-, x_0) = -\delta/2$ for all $t \leq t_-$. In praticular, there holds :

$$\partial_t \tilde{n}(t_-, x_0) \leq 0.$$

Replacing with (54) and recalling that $\tilde{n}(t_-, x_0) = -\delta/2$ we infer that:

$$r(T[\tilde{n}])_-(t_-, x_0) \geq (m + \kappa\rho + \Lambda_0)\frac{\delta}{2} \geq \frac{\Lambda_0}{2}\delta > 16r\delta.$$

However, we can apply (52) to $f = \tilde{n}(t_-, \cdot)$ and we obtain:

$$\begin{aligned}
r(T[\tilde{n}])_-(t_-, x_0) &\leq \frac{4r\|\tilde{n}(t_-, \cdot)\|_{L^1(\mathbb{R})}}{\int_\mathbb{R} \tilde{n}(t_-, z)\mathrm{d}z}\|(\tilde{n})_-(t_-, \cdot)\|_{L^\infty(\mathbb{R})} \\
&\leq \frac{4r\|n(t_-, \cdot)\|_{L^1(\mathbb{R})}}{\int_\mathbb{R} n(t_-, z)\mathrm{d}z}\|(\tilde{n})_-\|_{L^\infty((0,T_0) \times \mathbb{R})} \\
&\leq 16r\frac{\|n_0\|_{L^1(\mathbb{R})}}{\int_\mathbb{R} n_0(z)\mathrm{d}z}\delta = 16r\delta.
\end{aligned}$$

Since $r\delta > 0$ we obtain a contradiction. Hence n may not take strictly negative values on $[0, T_0]$. $\square$

With this positivity result at-hand we show now quantitative bounds satisfied by the solution.

Lemma 8 *We have:*

- *the pointwise bounds:*

$$n(t,x) \geq n_0(t,x) \exp\left[-\left(m(x) + \kappa\max\left(\int_\mathbb{R} n_0(z)\mathrm{d}z, \frac{r}{\kappa}\right)\right)t\right], \tag{55}$$

$$n(t,x) \leq n_0(x) + r\|\Gamma\|_{L^\infty(\mathbb{R})}\max\left(\int_\mathbb{R} n_0(z)\mathrm{d}z, \frac{r}{\kappa}\right)t, \tag{56}$$

27

- *the averaged bounds:*

$$\int_{\mathbb{R}} n(t,x) \le \max\left(\int_{\mathbb{R}} n_0(z)\mathrm{d}z, \frac{r}{\kappa}\right), \tag{57}$$

$$\|n(t,\cdot)\|_{X_k} \le \|n_0\|_{X_k} \exp(C_k t), \tag{58}$$

where C_k is a constant depending on the kernel Γ, r and k only.

Proof. We recall that n is a solution of

$$\partial_t n = rT[n] - (m + \kappa\rho(t))\, n.$$

In this lemma we assume that n is non-negative. For the bound (57), we first take the average in the equation,

$$\partial_t \int_{\mathbb{R}} n(t,x)\mathrm{d}x = r\int_{\mathbb{R}} T[n](x)\mathrm{d}x - \int_{\mathbb{R}} m(x)n(t,x)\mathrm{d}x - \kappa\left(\int_{\mathbb{R}} n(t,x)\mathrm{d}x\right)^2.$$

Integrating in time between 0 and t gives

$$\int_{\mathbb{R}} n(t,x)\mathrm{d}x - \int_{\mathbb{R}} n_0(x)\mathrm{d}x \le r\int_{\mathbb{R}} n(t,x)\mathrm{d}x - \int_{\mathbb{R}} m(x)n(t,x)\mathrm{d}x - \kappa\left(\int_{\mathbb{R}} n(t,x)\mathrm{d}x\right)^2,$$

so that

$$\int_{\mathbb{R}} (1 + m(x))n(t,x)\mathrm{d}x - \int_{\mathbb{R}} n_0(x)\mathrm{d}x \le r\int_{\mathbb{R}} n(t,x)\mathrm{d}x - \kappa\left(\int_{\mathbb{R}} n(t,x)\mathrm{d}x\right)^2.$$

Either $\int_{\mathbb{R}} n(t,x)\mathrm{d}x \le \int_{\mathbb{R}} n_0(x)\mathrm{d}x$ or $\int_{\mathbb{R}}(1+m(x))n(t,x)\mathrm{d}x \ge \int_{\mathbb{R}} n(t,x)\mathrm{d}x \ge \int_{\mathbb{R}} n_0(x)\mathrm{d}x \ge 0$ so that $\int_{\mathbb{R}} n(t,x)\mathrm{d}x \le \dfrac{r}{\kappa}$ which gives (57). For the bound (55), we use ii) of Lemma 6 so that $T[n] \ge 0$ and using (57)

$$\partial_t n \ge -(m + \kappa\rho(t))n \ge -\left(m + \kappa\max\left(\int_{\mathbb{R}} n_0(x)\mathrm{d}x, \frac{r}{\kappa}\right)\right)n.$$

We deduce

$$\partial_t\left(n\exp\left(\left(m + \kappa\max\left(\int_{\mathbb{R}} n_0(x)\mathrm{d}x, \frac{r}{\kappa}\right)\right)t\right)\right) \ge 0$$

so

$$n(t,x) \ge n_0(x)\exp\left(-\left(m + \kappa\max\left(\int_{\mathbb{R}} n_0(x)\mathrm{d}x, \frac{r}{\kappa}\right)\right)t\right).$$

For the bound (56), using that $n \ge 0$ and (57) we get

$$\partial_t n \le rT[n] \le r\|\Gamma\|_{L^\infty(\mathbb{R})}\rho(t) \le r\|\Gamma\|_{L^\infty(\mathbb{R})}\max\left(\int_{\mathbb{R}} n_0(x)\mathrm{d}x, \frac{r}{\kappa}\right)$$

which gives (56). For the bound (58), for a constant C_k to be fixed, we compute

$$\partial_t\left(\exp(-C_k t)\|n(t,\cdot)\|_{X_k}\right) = \int_{\mathbb{R}} w_k(x)\exp(-C_k t)\partial_t n(t,x)\mathrm{d}x - C_k\exp(-C_k t)\|n(t,\cdot)\|_{X_k}$$

$$\le \exp(-C_k t)r\int_{\mathbb{R}} w_k(x)T[n](t,x)\mathrm{d}x - C_k\exp(-C_k t)\|n(t,\cdot)\|_{X_k}$$

We bound thanks to the change of variable $X = x - \dfrac{y + y'}{2}$ and the inequality $1 \le w_k(X + \dfrac{y+y'}{2}) \le w_k(X) + w_k(y) + w_k(y')$,

$$\int_{\mathbb{R}} w_k(x)T[n](t,x)\mathrm{d}x \le \int_{\mathbb{R}}\int_{\mathbb{R}}\int_{\mathbb{R}} w_k(X + \frac{y+y'}{2})\Gamma(X)\, n(t,y)\frac{n(t,y')}{\rho(t)}\mathrm{d}y\mathrm{d}y'\mathrm{d}X$$

$$\le 3\|w_k\Gamma\|_{L^1(\mathbb{R})}\|n(t,\cdot)\|_{X_k}.$$

We then deduce, choosing $C_k := 3r\|w_k\Gamma\|_{L^1(\mathbb{R})}$,

$$\partial_t \left(\exp(-C_k t)\|n(t,\cdot)\|_{X_k}\right) \leq 3r\|w_k\Gamma\|_\infty \|n(t,\cdot)\|_{X_k}\exp(-C_k t) - C_k\|n(t,\cdot)\|_{X_k}\exp(-C_k t)$$
$$\leq 0,$$

which gives the result. $\qquad\square$

We point out that the pointwise bound (55) turns into a bound from below for the integral of $n(t,\cdot)$. Indeed, in case $n_0 \in \mathcal{U}_k \cap C_b(\mathbb{R})$ is nonnegative, there exists a compact interval I such that

$$\int_I n_0(x)\mathrm{d}x > 0.$$

We have then:

$$\int_\mathbb{R} n(t,x)\mathrm{d}x \geq \int_I n(t,x)\mathrm{d}x \geq \exp\left[-\left(\|m\|_{L^\infty(I)} + \kappa\max\left(\int_\mathbb{R} n_0(x)\mathrm{d}x, \frac{r}{\kappa}\right)\right)t\right]\int_I n_0(x)\mathrm{d}x > 0.$$

In case m is globally bounded, we have then that the second item of the blow-up alternative never occurs for a nonnegative initial data in $\mathcal{U}_k$ so that solutions are global.

A.2.2 Existence result for unbounded m.

We proceed with extending the existence result to a possibly unbounded m, we apply a compactness argument. From now on, the initial data $n_0 \in \mathcal{U}_k \cap C_b(\mathbb{R})$ is fixed. Firstly, we truncate m and apply the previous construction yielding a sequence of approximate global solutions $(n_p)_{p\in\mathbb{N}}$. We can then use the *a priori* bounds constructed in the latter lemma. These bounds enable to extract a subsequence converging in $L^\infty((0,T_0)\times\mathbb{R}) - w*$ (whatever $T_0 > 0$) but is insufficient to obtain that the limit is continuous and satisfies (1). For this purpose we provide the following a priori estimate:

Lemma 9 *Let R and T_0 positive and assume that n is a solution to (1) on $(0,T_0)$. Then, we have for any $t \in [0,T_0]$:*

$$\sup_{(x,\tilde{x})\in[-R,R]^2} |n(t,x) - n(t,\tilde{x})| \leq \sup_{(x,\tilde{x})\in[-R,R]^2} |n_0(x) - n_0(\tilde{x})| + C_0|x - \tilde{x}|\left(1 + \|m\|_{C^1([-R,R])}\right),$$

where C_0 depends only on initial data, T_0, r and κ.

Proof. Let us define $\delta(t,x,\tilde{x}) = n(t,x) - n(t,\tilde{x})$ so that

$$\partial_t\delta = r(T[n](t,x) - T[n](t,\tilde{x})) - (m(x) - m(\tilde{x}))n(t,x) - (m(\tilde{x}) + \kappa\rho)\delta.$$

By Duhamel formula we get

$$\delta(t,x,\tilde{x}) = \delta(0,x,\tilde{x})\exp\left(-m(\tilde{x})t - \int_0^t \kappa\rho(s)\mathrm{d}s\right)$$
$$+ \int_0^t (r(T[n](s,x) - T[n](s,\tilde{x})) - (m(x) - m(\tilde{x}))n(s,\tilde{x}))\exp\left(-m(\tilde{x})(t-s) - \int_s^t \kappa\rho(\tau)\mathrm{d}\tau\right)\mathrm{d}s.$$

We then deduce for $(x,\tilde{x}) \in [-R,R]^2$ that,

$$|\delta(t,x,\tilde{x})| \leq |\delta(0,x,\tilde{x})| + \int_0^t r|T[n](s,x) - T[n](s,\tilde{x})|\mathrm{d}s + |x-\tilde{x}|\|m\|_{C^1([-R,R])}\|n\|_{L^\infty}t.$$

Using that $x \to \exp(-x^2)$ is Lipschitz on $\mathbb{R}$, we get $|T[n](s,x) - T[n](s,\tilde{x})| \leq C|x-\tilde{x}|\rho(s)$ and we deduce the result with (56) and (57). $\qquad\square$

We emphasize that the above *a priori* estimate holds only on bounded intervals of $\mathbb{R}$ and is independent of the behavior of m at infinity. We can then apply this estimate to the approximate solutions (n_p) yielding a local uniform bound. Since we have also an *a priori* estimate (51) on $(\partial_t n_p)$ we infer that the sequence of approximate solutions (n_p) is also compact in $C([0, T_0] \times [-R, R])$ for arbitrary positive T_0, R. Eventually, we combine the subsequent local strong convergence of a subsequence of approximate solutions with the uniform bounds in $L^\infty((0, T_0); X_k)$ to yield that the limit lies in $L^\infty(0, T_0; X_k)$ also and that this very subsequence converges also in $C([0, T_0]; X_{\tilde{k}})$ for arbitrary $\tilde{k} < k$. In particular we have convergence in $C_{loc}([0, \infty), L^1(\mathbb{R}))$ and we can pass to the limit in the equations satisfied by approximate solutions. This entails that the limit is a solution to (1).

A.3 Uniqueness of solutions

To end up this section, we prove that, given $n_0 \in \mathcal{U}_k \cap C_b(\mathbb{R})$, the solution we constructed above in case m is not necessarily bounded is the unique solution in the sense of Definition 5. We restrict to the case $k > 3$.

For this, we denote n the solution constructed above and $\tilde{n}$ a possible other solution with the same initial data $n_0 \in \mathcal{U}_k \cap C_b(\mathbb{R})$. We note that both solutions satisfy the conclusions of Lemma 8. We pick then $k_0 \in (3, k)$ and define

$$\Delta_0(t) = \int_{\mathbb{R}} |n(t, x) - \tilde{n}(t, x)|^2 w_{k_0}(x) \mathrm{d}x.$$

Since n and $\tilde{n}$ are bounded and continuous with values in $X_{\tilde{k}}$ whatever $\tilde{k} \in (k_0, k)$ we infer that $\Delta_0 \in C([0, \infty))$ and vanishes initially. The key remark in the proof is then the following Gronwall inequality:

Lemma 10 *Given $T_0 > 0$ there exists a constant C_0 that depends only on initial data n_0, k_0, T_0, r, κ, m and Γ for which:*

$$\Delta_0(\tau) - \Delta_0(\sigma) \leq C_0 \int_\tau^\sigma \Delta_0(s) \mathrm{d}s.$$

Since Δ_0 is continuous, we conclude by applying the Gronwall lemma that Δ_0 vanishes identically on $[0, T_0]$ and then globally. We end up this appendix by a proof of **Lemma 10**.

Proof. Let $\delta(t, x) := n(t, x) - \tilde{n}(t, x)$ and define $\Delta_0^R(t) = \int_{-R}^R |n(t, x) - \tilde{n}(t, x)|^2 w_{k_0}(x) \mathrm{d}x$. Noticing that δ satisfies

$$\partial_t \delta(t, x) = r(T[n] - T[\tilde{n}])(t, x) - (m + \kappa\rho(t))\delta(t, x) - \kappa(\rho(t) - \tilde{\rho}(t))\tilde{n}(t, x),$$

we multiply this equation by δw_{k_0} and integrate on $[\sigma, \tau] \times [-R, R]$ and get

$$\frac{1}{2}(\Delta_0^R(\tau) - \Delta_0^R(\sigma)) + \int_\sigma^\tau \int_{-R}^R (m + \kappa\rho(s))\delta^2(s, x)w_{k_0}(x)\mathrm{d}x\mathrm{d}s = I_1 + I_2$$

where

$$I_1 = -\kappa \int_\sigma^\tau \int_{-R}^R (\rho(s) - \tilde{\rho}(s))\tilde{n}(s, x)\delta(s, x)w_{k_0}(x)\mathrm{d}x\mathrm{d}s$$

and

$$I_2 = r \int_\sigma^\tau \int_{-R}^R (T[n] - T[\tilde{n}])(s, x)\delta(s, x)w_{k_0}(x)\mathrm{d}x\mathrm{d}s.$$

Using that $\int_\sigma^\tau \int_{-R}^R (m + \kappa\rho(s))\delta^2(s, x)w_{k_0}(x)\mathrm{d}x\mathrm{d}s \geq 0$ we deduce

$$\Delta_0^R(\tau) - \Delta_0^R(\sigma) \leq 2(I_1 + I_2). \tag{59}$$

For I_1, we bound $|\rho - \tilde{\rho}|$ using a Cauchy-Schwarz inequality and the fact that $\frac{1}{w_{k_0}} \in L^1(\mathbb{R})$,

$$|\rho(s) - \tilde{\rho}(s)| \le \int_{\mathbb{R}} |n(s,x) - \tilde{n}(s,x)| \frac{\sqrt{w_{k_0}(x)}}{\sqrt{w_{k_0}(x)}} \, \mathrm{d}x$$

$$\le \left(\int_{\mathbb{R}} |n(s,x) - \tilde{n}(s,x)|^2 w_{k_0}(x) \mathrm{d}x \right)^{1/2} \left(\int_{\mathbb{R}} \frac{1}{w_{k_0}(x)} \mathrm{d}x \right)^{1/2}$$

$$\le \Delta_0(s)^{1/2} \left(\int_{\mathbb{R}} \frac{1}{w_{k_0}(x)} \mathrm{d}x \right)^{1/2}$$

$$\le C_{k_0} \Delta_0(s)^{1/2},$$

and we also get thanks to Cauchy-Schwarz inequality and using the bounds (56), (58),

$$\int_{-R}^{R} \tilde{n}(s,x) \delta(s,x) w_{k_0}(x) \mathrm{d}x \le \left(\int_{\mathbb{R}} \tilde{n}(s,x)^2 w_{k_0}(x) \mathrm{d}x \right)^{1/2} \left(\int_{\mathbb{R}} \delta(s,x)^2 w_{k_0}(x) \mathrm{d}x \right)^{1/2}$$

$$\le \|n(s,\cdot)\|_{L^\infty(\mathbb{R})}^{1/2} \|n(s,\cdot)\|_{X_k}^{1/2} \Delta_0^{1/2}(s)$$

$$\le C_{k_0,n_0,\Gamma,r,\kappa} \Delta_0(s)^{1/2},$$

so that

$$|I_1| \le C_{k_0,n_0,\Gamma,r,\kappa} \int_{\sigma}^{\tau} \Delta_0(s) \mathrm{d}s. \tag{60}$$

For I_2, we first have by a standard Cauchy-Schwarz inequality:

$$I_2 \le r \int_{\sigma}^{\tau} \left(\int_{\mathbb{R}} |T[n] - T[\tilde{n}]|^2 w_{k_0} \right)^{\frac{1}{2}} \Delta_0^{\frac{1}{2}}.$$

However, introducing $\ell \in (1, k_0 - 1)$ to be fixed later on, recalling Lemma 6 with (50) and remarking that $\|\cdot\|_{X_\ell} \le \|\cdot\|_{X_k}$, we infer that:

$$\int_{\mathbb{R}} |T[n] - T[\tilde{n}]|^2 w_{k_0} \le \|w_\ell(T[n] - T[\tilde{n}])\|_{C_b(\mathbb{R})} \|T[n] - T[\tilde{n}]\|_{X_{k_0 - \ell}}$$

$$\le K \left(\frac{\|n(s,\cdot)\|_{X_k}}{|\int_{\mathbb{R}} n(s,z) \mathrm{d}z|} \frac{\|\tilde{n}(s,\cdot)\|_{X_k}}{|\int_{\mathbb{R}} \tilde{n}(s,z) \mathrm{d}z|} \right) \left(\|n\|_{X_k} + \|\tilde{n}\|_{X_k} \right) \|n - \tilde{n}\|_{X_{k_0 - \ell}}.$$

Then, by Lemma 8 there is a constant C_0 depending only on initial data n_0, r, κ, Γ, m and T_0 so that:

$$\sup_{s \in (0,T_0)} K \left(\frac{\|n(s,\cdot)\|_{X_k}}{|\int_{\mathbb{R}} n(s,z) \mathrm{d}z|} \frac{\|\tilde{n}(s,\cdot)\|_{X_k}}{|\int_{\mathbb{R}} \tilde{n}(s,z) \mathrm{d}z|} \right) \left(\|n\|_{X_k} + \|\tilde{n}\|_{X_k} \right) \le C_0,$$

using that there exists a bounded interval I such that $\int_{\mathbb{R}} n_0(z) \mathrm{d}z > 0$ so that m is bounded and

$$\int_{\mathbb{R}} n(s,z) \mathrm{d}z \ge \int_I n(s,z) \mathrm{d}z > \int_I n_0(z) \exp\left[-\left(m(z) + \kappa \max\left(\int_{\mathbb{R}} n_0(x) \mathrm{d}x, \frac{r}{\kappa} \right) \right) T_0 \right] \mathrm{d}z$$

$$> C_0 \int_{\mathbb{R}} n_0(z) \mathrm{d}z.$$

Then a Cauchy-Schwarz inequality entails that, for arbitrary $\mu > 0$ there is C_μ such that:

$$\|n - \tilde{n}\|_{X_{k_0 - \ell}} \le C_\mu \left(\int_{\mathbb{R}} |n - \tilde{n}|^2 w_{2(k_0 - \ell) + (1+\mu)} \right)^{\frac{1}{2}}$$

Since $k_0 > 3$ we can choose ℓ and μ so that $2(k_0 - \ell) + (1 + \mu) = k_0$ and we obtain finally that:

$$I_2 \leq C_0 \int_\sigma^\tau \Delta_0(s)\mathrm{d}s. \tag{61}$$

with a constant C_0 depending on the chosen ℓ, ε, initial data and T. We conclude combining (59), (60), (61) and sending R to $+\infty$. $\quad\square$

References

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

[2] N. H. Barton, A. M. Etheridge, and A. Veber. The infinitesimal model: definition, derivation and implications. *Theor. Pop. Biol.*, 118:50–73, 2017.

[3] M. G. Bulmer. *The mathematical theory of quantitative genetics.* Oxford, UK: Clarendon Press, 1980.

[4] R. Bürger. *The Mathematical theory of selection, recombination and mutation.* Wiley, New-York, 2000.

[5] V. Calvez, J. Garnier, and F. Patout. A quantitative genetics model with sexual mode of reproduction in the regime of small variance. *Journal de l'École polytechnique — Mathématiques*, 6:537–579, 2019.

[6] V. Calvez, T. Lepoutre, and D. Poyato. Ergodicity of the fisher infinitesimal model with quadratic selection. Preprint.

[7] V. Calvez, D. Poyato, and F. Santambrogio. Uniform contractivity of the fisher infinitesimal model with strongly convex selection. Preprint.

[8] L. Dekens. Evolutionary dynamics of complex traits in sexual populations in a strongly heterogeneous environment: how normal? *Journal of Mathematical Biology*, 84(3):15, 2022.

[9] L. Dekens and S. Mirrahimi. Dynamics of Dirac concentrations in the evolution of quantitative alleles with sexual reproduction. *Nonlinearity*, 35(11):5781, 2022.

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

[11] R. A. Fisher. Xv.-the correlation between relatives on the supposition of mendelian inheritance. *Earth and Environmental Science Transactions of the Royal Society of Edinburgh*, 52(2):399–433, 1919.

[12] N. Fournier, and B. Perthame. A Nonexpanding Transport Distance for Some Structured Equations. *SIAM Journal on Mathematical Analysis*, 53(6):6847–6872, 2021.

[13] C. Guinet, S. Mirrahimi and J.-M. Roquejoffre. Approximation of solutions of selection-mutation models and error estimates. *In preparation.*

[14] M. Kirkpatrick and N. H. Barton. Evolution of a species' range. *Amer. Nat.*, 150(1):1–23, 1997.

[15] M. Kopp and S. Matuszewski. Rapid evolution of quantitative traits: theoretical perspectives. *Evolutionary Applications*, 7(1):169–191, 2014.

[16] R. Lande and S. Shannon. The role of genetic variation in adaptation and population persistence in a changing environment. *Evolution*, 50(1):434–437, 1996.

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

[18] M. Lynch and B. Walsh. *Genetics and Analysis of Quantitative Traits*. Sinauer Associates, Inc., 1998.

[19] P. Magal and G. Raoul. Dynamics of a kinetic model describing protein exchanges in a cell population. *ArXiv:1511.02665*, 2015.

[20] S. Mirrahimi and G. Raoul. Population structured by a space variable and a phenotypical trait. *Theoretical Population Biology*, 84:87–103, 2013.

[21] S. Mirrahimi. A Hamilton-Jacobi approach to characterize the evolutionary equilibria in heterogeneous environments. *Mathematical Models and Methods in Applied Sciences*, 27(13):2425–2460, 2017.

[22] S. Mirrahimi and J.-M. Roquejoffre. Uniqueness in a class of Hamilton-Jacobi equations with constraints. *C. R. Math. Acad. Sci. Paris*, 353:489–494, 2015.

[23] S. Mirrahimi and J.-M. Roquejoffre. A class of Hamilton-Jacobi equations with constraint: Uniqueness and constructive approach. *Journal of Differential Equations*, 260(5):4717–4738, 2016.

[24] F. Patout. The cauchy problem for the infinitesimal model in the regime of small variance. *Analysis & PDE*, To appear.

[25] B. Perthame and G. Barles. Dirac concentrations in Lotka-Volterra parabolic PDEs. *Indiana Univ. Math. J.*, 57(7):3275–3301, 2008.

[26] B. Perthame and M. Strugarek and C. Taing Selection–mutation dynamics with asymmetrical reproduction kernels. *Nonlinear Analysis*, 222:112947, 2022.

[27] C. Prevost. *Applications of partial differential equations and their numerical simulations of population dynamics.* PhD. Université d'Orléans. 2004

[28] G. Raoul. Exponential convergence to a steady-state for a population genetics model with sexual reproduction and selection. *Preprint.*

[29] G. Raoul. Macroscopic limit from a structured population model to the kirkpatrick-barton model. *Preprint.*

[30] O. Ronce and M. Kirkpatrick. When sources become sinks: migration meltdown in heterogeneous habitats. *Evolution*, 55(8):1520–1531, 2001.

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