

Dynamics of concentration in a population structured by age and a phenotypic trait with mutations. Convergence of the corrector

Samuel Nordmann, Benoît Perthame

▶ **To cite this version:**

Samuel Nordmann, Benoît Perthame. Dynamics of concentration in a population structured by age and a phenotypic trait with mutations. Convergence of the corrector. Journal of Differential Equations, 2021, 10.1016/j.jde.2021.04.029 . hal-02437168

HAL Id: hal-02437168
https://hal.science/hal-02437168

Submitted on 13 Jan 2020

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

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

Dynamics of concentration in a population structured by age and a phenotypic trait with mutations. Convergence of the corrector

Samuel Nordmann* Benoît Perthame†‡

January 13, 2020

Abstract

We study an equation structured by age and a phenotypic trait describing the growth process of a population subject to aging, competition between individuals, and mutations. This leads to a renewal equation which occurs in many evolutionary biology problems. We aim to describe precisely the asymptotic behavior of the solution, to infer properties that illustrate the concentration and adaptive dynamics of such a population. This work is a continuation of [38] where the case without mutations is considered. When mutations are taken into account, it is necessary to control the corrector which is the main novelty of the present paper.

Our approach consists in defining, by the Hopf transform, a Hamilton-Jacobi equation with an effective Hamiltonian as in homogenization problems. Its solution carries the singular part of the limiting density (typically Dirac masses) and the corrector defines the weights. The main new result of this paper is to prove that the corrector is uniformly bounded, using only the global Lipschitz and semi-convexity estimates for the viscosity solution of the Hamilton-Jacobi equation. We also establish the limiting equation satisfied by the corrector. To the best of our knowledge, this is the first example where such bounds can be proved in such a context.

Key-words: Adaptive evolution; Asymptotic behaviour; Dirac concentration; Hamilton-Jacobi equations; Mathematical biology; Renewal equation; Viscosity solutions; correctors.

AMS Class. No: 35B40, 35F21, 35Q92, 49L25.

1 Introduction

1.1 Main results

We study a mathematical model describing the growth process of a population structured by age and a phenotypic trait, subject to aging, competition between individuals and mutations. Our goal is to describe the asymptotic behavior of the solution, in particular the selection of the fittest traits and the adaptative dynamics of such traits. For $\varepsilon > 0$, we choose $m_\varepsilon(t,x,y)$ to represent the population density of individuals who, at time $t \geq 0$, have age $x \geq 0$ and a quantitative phenotipic trait $y \in \mathbb{R}^n$, solution of a renewal type equation

$$\begin{cases} \varepsilon \partial_t m_\varepsilon + \partial_x \left[A(x,y) m_\varepsilon \right] + \left(\rho_\varepsilon(t) + d(x,y) \right) m_\varepsilon = 0, \\ A(0,y) m_\varepsilon\left(t,0,y\right) = \frac{1}{\varepsilon^n} \int_{\mathbb{R}^n} \int_{\mathbb{R}_+} M(\frac{y'-y}{\varepsilon}) b(x',y') m_\varepsilon(t,x',y') dx' dy', \\ \rho_\varepsilon(t) = \int_{\mathbb{R}_+} \int_{\mathbb{R}^n} m_\varepsilon(t,x,y) dx dy, \\ m_\varepsilon(t=0,x,y) = m_\varepsilon^0(x,y) > 0. \end{cases} \tag{1}$$

*École des Hautes Études en Sciences Sociales, CAMS, 54 boulevard Raspail, 75006 Paris, France. Email: samuel.nordmann@ehess.fr

†Sorbonne Université, CNRS, Université de Paris, Inria, Laboratoire Jacques-Louis Lions, 75005 Paris, France. Email: benoit.perthame@upmc.fr

‡**Acknowledgment.** BP has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 740623)

The parameter $\varepsilon > 0$ stands for a hyperbolic rescaling $(t, y) \leftrightarrow (\varepsilon^{-1}t, \varepsilon^{-1}y)$ in the mutation term. Our main concern is to study the asymptotics of m_ε when $\varepsilon \to 0$.

This work is the continuation of the study begun in [38], where the model without mutations is studied and where we proved that there is a measure μ (typically a Dirac mass) and a bounded profile $\mathcal{Q}$ such that

$$m_\varepsilon(t, x, y) \rightharpoonup \mathcal{Q}(t, x, y)\, \mu(t, y),$$

in other words, the asymptotic singularity is carried in the variable y only. In the present work, the mutation term in the second line of (1) adds a significant difficulty because the profile $\mathcal{Q}$ turns out to be strongly related to the limit of the corrector in the spectral problem defining the underlying effective Hamiltonian, i.e., the *effective fitness* in our context, arising in a Hamilton-Jacobi equation which defines the singular part.

While the classical approach consists in studying the asymptotics of $v_\varepsilon(t, x, y) := \varepsilon \ln(m_\varepsilon(t, x, y))$, here, and following [38, 40], we define a priori some kind of variable separation setting

$$m_\varepsilon(t, x, y) := p_\varepsilon(t, x, y) e^{\frac{U_\varepsilon(t,y) - \int_0^t \rho_\varepsilon(\cdot)}{\varepsilon}}$$

where the exponential term carries the singular part of the limiting population density and U_ε is defined autonomously through a Hamilton-Jacobi equation (see (7)) involving the *effective fitness* Λ (defined in (6)). The main finding of this paper is the proof that, with our choice of U_ε, the corrector p_ε is uniformly bounded. We also show that p_ε converges to a multiple of the principal eigenfunction $Q(t, x, y)$ of the formal limiting operator (defined in (6)). It justifies that the ansatz U_ε is appropriate and carries all the information on the singular behavior of m_ε when $\varepsilon \to 0$. The formal idea of the method is explained with more details in section 1.3 and consists in keeping as simple as possible the Hamilton-Jacobi equation which defines U_ε while all the functional analytic difficulties are carried by the corrector p_ε which satisfies a linear equation.

The first step is to define U_ε and study its properties.

Theorem 1.1 (Convergence of U_ε) *Under the assumptions of section 1.5, when $\varepsilon \to 0$, the ansatz $U_\varepsilon(t, y)$ is well defined through (7) and converges in $W^{1,r}_{loc}$, $1 \le r < \infty$, to some $U \in W^{1,\infty}_{loc}$ which is semi-convex and is a viscosity solution of the Hamilton-Jacobi equation (1.3).*

The second step is to prove that the corrector is uniformly bounded and converges.

Theorem 1.2 (Convergence of the corrector) *Under the assumptions of section 1.5, for any fixed $T > 0$, $p_\varepsilon(t, x, y)$ and $\int_{x>0} p_\varepsilon(t, x, y)dx$ are bounded from above and below, uniformly in $\varepsilon > 0$, $t \in [0, T]$, $y \in \mathbb{R}^n$.*

In addition, for $t \in [0, T]$, $x \in [0, \overline{x}]$, $y \in \mathbb{R}^n$, up to extraction of a subsequence $\varepsilon \to 0$, p_ε converges in L^∞-weak-$\star$ to $\gamma(t, y)Q(t, x, y)$, where Q is the eigenfunction defined in (6) and $\gamma \in L^\infty$ formally satisfies (30).

With these informations, it is standard that the *concentration points* of the population density $m_\varepsilon(t, x, y)$ are carried by the set

$$\mathcal{S} := \left\{ t \ge 0,\ y \in \mathbb{R}^n : U(t, y) = \sup_{y' \in \mathbb{R}^n} U(t, y') \right\}. \tag{2}$$

Theorem 1.3 (concentration) *Under the assumptions of section 1.5, when $\varepsilon \to 0$*

1. The total population ρ_ε converges weakly to some positive $\rho \in L^\infty$ and

$$\forall t > 0, \quad \int_0^t \rho = \sup_{y \in \mathbb{R}^n} U(t, y).$$

2. The population m_ε vanishes locally uniformly outside the set $\mathcal{S}$.

3. Under further assumptions on the initial conditions, and for small times $t \in [0, T]$, we have

$$\mathcal{S} = \{(t, \bar{y}(t))\},$$

where $\bar{y}(t)$ follows the Canonical Equation

$$\begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}t}\bar{y}(t) = \left(D^2_y U(t, \bar{y}(t))\right)^{-1} \cdot \nabla_y \Lambda(\bar{y}(t), 1) + \partial_\eta \Lambda(\bar{y}(t), 1) \displaystyle\int_{\mathbb{R}^n} M(z) z dz, \\ \bar{y}(0) = \bar{y}^0. \end{cases}$$

Notice that we establish convergence of the full families U_ε and ρ_ε, without using the famous uniqueness result of [14], because of the simple dependency on the unknown ρ_ε in our setting.

We point out that the main restriction on the coefficients is the assumption that the transport in x outwards the support of $b(\cdot, y)$ occurs in finite time, see (10). The other assumptions, detailed in section 1.5, are formulated directly on the limiting eigenproblem and are quite general.

1.2 The model

Some possible biological interpretations of the model (1) are as follows. The function $A(x, y)$ is the speed of aging of individuals with age x and phenotypic trait y. The total size of the population at time t is denoted with $\rho_\varepsilon(t)$. In the mortality term, $d(x, y) > 0$ represents intrinsic death rate and the nonlocal term $\rho_\varepsilon(t)$ represents competition. The condition at the boundary $x = 0$ describes the birth of newborns that happens with rate $b(x, y) > 0$ and with the probability kernel M for mutations.

The terminology of *renewal equation* comes from this boundary condition. It is related to the McKendrick-von Foerster equation which is only structured in age (see [39] for a study of the linear equation). This model has been extended with other structuring variables, as size [32, 37], DNA content, maturation, etc., in the context of cell divisions [21], or proliferative and quiescent states of tumor cells [2, 26]. Space structured problems have also been extensively studied [27, 34, 35, 40].

To keep the model (1) quite general, we have allowed the progression speed A to depend on x. Thus, although the variable x is referred to as *age*, it can represent other biological quantities that evolve throughout the individual lifespan such as, for instance, the size of individuals, a physiological age, a parasite load, etc.

The rescaling parameter $\varepsilon > 0$ comes from a hyperbolic rescaling of t and y. Accordingly, the dynamics are considered in two different time scales. The first one is the individual lifetime scale εt, i.e., the characteristic time for the population to reach the dynamical equilibrium for a fixed y. The second one is the evolutionary time scale t, corresponding to the evolution of the population distribution with respect to the variable y. Formally, at the limit when $\varepsilon \to 0$, the time scales are completely separated. This rescaling is a classical way to give a continuous formulation of the adaptive evolution of a phenotypically structured population (see [16, 19, 20, 30]). Note that the mutation kernel is supposed to be *thin-tailed*, i.e., it decreases faster than any exponential. A fat-tailed kernel needs a different rescaling, see [10].

From the modelization point of view, Theorem 1.3 is a form of mathematical formulation of Natural Selection and Evolution. On an ecological time scale, only the phenotype $\bar{y}(t)$ which maximizes the ecological fitness $U(t, \cdot)$ can survive. On an evolutionary time scale, we observe the dynamics of $\bar{y}(t)$. Similar results as Theorem 1.3 have been obtained for various models with parabolic equations [5, 7, 30] and integrodifferential equations [6, 18, 29]. More generally, convergence to positive measures in selection-mutation models has been studied by many authors [1, 11, 12]. The special case of age-structured populations are also considered in [13, 31, 41].

1.3 A formal presentation of the method

To analyze the singular perturbation problem at hand, the usual approach relies on the WKB change of unknown ([7, 20]), which consists in the change of variable $m_\varepsilon(t, x, y) = e^{\frac{v_\varepsilon(t,x,y)}{\varepsilon}}$. In the context of concentrations, this form is motivated by the heuristics that a Dirac mass is nothing but a narrow Gaussian. Indeed, in a weak sense $(\pi\varepsilon)^{\frac{-n}{2}} e^{-\frac{\|y-\bar{y}\|^2}{\varepsilon}} \rightharpoonup \delta_{\bar{y}=y}$ as $\epsilon \to 0$. This approach has been extensively used in works on a similar issue, see for instance [4, 8, 17, 24]. With this change of unknown, at the limit $\varepsilon \to 0$, the function $v_\varepsilon(t, x, y)$ satisfies a constrained Hamilton-Jacobi equation, on which estimates can be difficult to prove because it carries all the difficulties in the asymptotic analysis, concentration effect and profile defined by the corrector. For that reason, the perturbed test function method has been invented [22] and widely used, which avoids computing the corrector.

Here, we propose a variant of the method. The principle can be viewed as a Taylor expansion $v_\varepsilon(t, x, y) = v^1_\varepsilon(t, y) + \varepsilon v^2_\varepsilon(t, x, y)$, to choose only some convenient terms to define the Hamilton-Jacobi equation for v^1_ε, and then to prove that the corrector v^2_ε is bounded. With a slight rewriting, we proceed to the change of variable

$$m_\varepsilon(t, x, y) = p_\varepsilon(t, x, y) \exp \frac{U_\varepsilon(t, y) - \int_0^t \rho_\varepsilon(s)ds}{\varepsilon} \tag{3}$$

where $U_\varepsilon(t,y)$ is defined *ad hoc* through a standard Hamilton-Jacobi equation for which classical regularity properties can be proved. Then,the new and difficult step is to prove estimates on the corrector $p_\varepsilon(t,x,y)$. Note that p_ε satisfies a linear equation rather than a constrained Hamilton-Jacobi equation: it makes thus possible to use classical comparison principles and ideas issued from the General Relative Entropy method (see [33]).

We are now left with the task of finding a good candidate for $U_\varepsilon(t,y)$ and formally identifying $p_\varepsilon(t,x,y)$. Injecting (3) in (1), we find

$$\begin{cases} \varepsilon\partial_t p_\varepsilon + \partial_x\left[A(x,y)p_\varepsilon\right] + d(x,y)p_\varepsilon + \partial_t U_\varepsilon(t,y)p_\varepsilon = 0, \\ A(0,y)p_\varepsilon\left(t,0,y\right) = \frac{1}{\varepsilon^n}\int_{\mathbb{R}^n}\int_{\mathbb{R}_+} M(\frac{y'-y}{\varepsilon})e^{\frac{U_\varepsilon(t,y')-U_\varepsilon(t,y)}{\varepsilon}}b(x',y')p_\varepsilon(t,x',y')dx'dy'. \end{cases} \tag{4}$$

With the change of variable $z = \frac{y'-y}{\varepsilon}$, the renewal term becomes

$$A(0,y)p_\varepsilon(t,0,y) = \int_{\mathbb{R}^n}\int_{\mathbb{R}_+} M(z)e^{\frac{U_\varepsilon(t,y+\varepsilon z)-U_\varepsilon(t,y)}{\varepsilon}}b(x',y+\varepsilon z)p_\varepsilon(t,x',y+\varepsilon z)dx'dz.$$

When ε is small, we can formally approximate

$$A(0,y)p_\varepsilon(t,0,y) \approx \eta_\varepsilon(t,y)\int_{\mathbb{R}_+} b(x',y)p_\varepsilon(t,x',y)dx', \tag{5}$$

where

$$\eta_\varepsilon(t,y) := \int_{\mathbb{R}^n} M(z)e^{\frac{U_\varepsilon(t,y+\varepsilon z)-U_\varepsilon(t,y)}{\varepsilon}}dz.$$

Then, formally putting $\varepsilon\partial_t p_\varepsilon = O(\varepsilon)$ in the first line of (4), we end up with the following approximate problem

$$\begin{cases} \partial_x\left[A(x,y)p_\varepsilon\right] + d(x,y)p_\varepsilon + \partial_t U_\varepsilon(t,y)p_\varepsilon = O(\varepsilon), \\ A(0,y)p_\varepsilon\left(t,0,y\right) = \eta_\varepsilon(t,y)\int_{\mathbb{R}_+} b(x',y)p(t,x',y)dx' + O(\varepsilon). \end{cases}$$

Considering $\eta_\varepsilon(t,y)$ as a parameter, we introduce the following eigenproblem: for fixed $(y,\eta) \in \mathbb{R}^n \times (0,+\infty)$, find $(\Lambda(y,\eta), Q(x,y,\eta))$, solution of

$$\begin{cases} \partial_x\left[A(x,y)Q\right] + d(x,y)Q - \Lambda(y,\eta)Q = 0, \quad \forall x > 0, \\ A(0,y)Q(0,y,\eta) = \eta, \\ Q > 0, \quad \int_{\mathbb{R}_+} b(x,y)Q(x,y,\eta)dx = 1. \end{cases} \tag{6}$$

The third line corresponds to a normalization of the eigenfunction which is convenient for future calculations. Formally, Λ corresponds to the *effective fitness*, and Q to the age profile at equilibrium in an environment characterized by the parameters (y,η).

In an attempt to indentify p_ε with Q, this formal approach suggests to define U_ε as a solution of the Hamilton-Jacobi equation

$$\begin{cases} \partial_t U_\varepsilon(t,y) = -\Lambda\left(y, \int_{\mathbb{R}^n} M(z)e^{\frac{U_\varepsilon(t,y+\varepsilon z)-U_\varepsilon(t,y)}{\varepsilon}}dz\right) & \forall t \geq 0,\ y \in \mathbb{R}^n, \\ U_\varepsilon(0,y) = U_\varepsilon^0(y) & \forall y \in \mathbb{R}^n, \end{cases} \tag{7}$$

for some initial conditions U_ε^0.

We stress out that, when ε vanishes, the full term $\exp\frac{U_\varepsilon(t,y)-\int_0^t \rho_\varepsilon(s)ds}{\varepsilon}$ represents a bounded measure, e.g., a Dirac mass as in the example of Gaussian concentration, and thus

$$\sup_{y\in\mathbb{R}^n} U_\varepsilon(t,y) - \int_0^t \rho_\varepsilon(s)ds = 0, \qquad \forall t \geq 0,$$

which explains the first formula in Theorem 1.3.

However, because of the non-local term η_ε, proving uniform in $\varepsilon > 0$ estimates on U_ε is quite technical, a fact that can also be seen on the limiting equation when $\varepsilon \to 0$. Taking for granted that U_ε converges to a function U locally uniformly, U turns out to be a viscosity solution of the Hamilton-Jacobi equation

$$\begin{cases} \partial_t U(t,y) = H(y, \nabla_y U) & \forall t \geq 0,\ \forall y \in \mathbb{R}^n, \\ U(0,y) = U^0(y) & \forall y \in \mathbb{R}^n. \end{cases} \tag{8}$$

with the Hamiltonian $H(y,p)$ defined by

$$H(y,p) := -\Lambda\left(y, \int_{\mathbb{R}^n} M(z) e^{p\cdot z} dz\right).$$

From this equation, it is classical to prove uniform a priori bounds on $\partial_t U$ (indeed, $\partial_t U$ satisfies a transport equation). We deduce that $H(y, \nabla_y U)$ is bounded, then that

$$\eta(t,y) := \int_{\mathbb{R}^n} M(z) e^{\nabla_y U(t,y)\cdot z} dz$$

is bounded. Besides, since $p \mapsto H(y,p)$ is convex, we deduce that U is semi-convex, namely in space-time, the Hessian $D^2 U$ is bounded from below. Since $\partial_t U$ is bounded, we have $\partial_t^2 U \in L^1_{loc}$. Then, using $\partial_t^2 U = -\partial_t \eta\, \partial_\eta \Lambda$, we infer that $\partial_t \eta \in L^1_{loc}$. In the sequel, all these estimates are proved on U_ε, uniformly in $\varepsilon > 0$. These are the classical (and sharp) general estimates for Hamilton-Jacobi equations with convex Hamiltonians.

With these optimal estimates on U_ε in hands, we can bound the corrector p_ε. We set

$$Q_\varepsilon(t,x,y) := Q(x,y,\eta_\varepsilon(t,y)), \qquad \Lambda_\varepsilon(t,y) := \Lambda(y, \eta_\varepsilon(t,y))$$

and find that

$$\begin{cases} \varepsilon \partial_t Q_\varepsilon + \partial_x \left[A(x,y) Q_\varepsilon\right] + d(x,y) Q_\varepsilon + \partial_t U_\varepsilon(t,y) Q_\varepsilon = \varepsilon \partial_t Q_\varepsilon, \\ A(0,y) Q_\varepsilon\,(t,0,y) = \frac{1}{\varepsilon^n} \int_{\mathbb{R}^n} \int_{\mathbb{R}_+} M(\frac{y'-y}{\varepsilon}) e^{\frac{U_\varepsilon(t,y') - U_\varepsilon(t,y)}{\varepsilon}} b(x',y') Q_\varepsilon(t,x',y') dx' dy'. \end{cases}$$

Note that the boundary term at $x = 0$ is obtained using the definition of $\eta_\varepsilon(t,y)$ and the normalization $\int_{\mathbb{R}_+} b(x,y) Q(x,y,\eta) dx = 1$ for all (y,η).

The right hand side of the first line can be controlled with the available a priori bounds on U_ε. Indeed, integrating equation (6) to obtain AQ, we compute

$$\partial_t Q_\varepsilon = \partial_t \eta_\varepsilon(t,y) \partial_\eta Q(x,y,\eta_\varepsilon(t,y)) = \partial_t \eta_\varepsilon \left(\frac{1}{\eta_\varepsilon} + \partial_\eta \Lambda(y,\eta_\varepsilon) \int_0^x \frac{1}{A(\cdot,y)}\right) Q_\varepsilon.$$

Except from this term, we see that Q_ε and p_ε satisfy the same equation (4), which is linear and admits a comparison principle. Under assumptions on the initial conditions, we deduce that p_ε is bounded from above and below by multiples of Q_ε. Passing to the limit in the equation, we prove that p_ε converges weakly to a multiple of Q. It justifies our approach, especially the approximation (5), and formally proves Theorem 1.2.

Now, having in hand that U_ε converges (Theorem 1.1) and that p_ε is uniformly bounded (Theorem 1.2), Theorem 1.3 can be understood and formally justified as follows. On the one hand, the saturation term "$\rho_\varepsilon(t)$" in (1) implies the total population ρ_ε to be bounded, uniformly in $\varepsilon > 0$. On the other hand, from (3), the asymptotics of $m_\varepsilon(t,\cdot,\cdot)$ when ε vanishes are driven by the points y where $U(t,\cdot)$ is maximal. In other words, when $\varepsilon \to 0$, m_ε vanishes outside the set $\mathcal{S}$ defined in (2). Then, we can study the evolutionary dynamics through the dynamics of the (unique) critical point of $U(t,\cdot)$.

1.4 Outline of the paper

Section 2 is devoted to the definition of the eigenelements (Q, Λ), the definition of the ansatz U_ε, the statement of a priori estimates, and the asymptotics of U_ε when $\varepsilon \to 0$. The proofs are postponed to section 5. In section 3, we study the corrector p_ε and prove Theorem 1.2. Next, we study the asymptotics of the population density, and we prove Theorem 1.3 in section 4. Finally, some longer or more technical proofs are gathered in section 5.

1.5 Assumptions

Most of our assumptions are formulated directly on the solution $(\Lambda(y,\eta), Q(x,y,\eta))$ of the limiting eigenproblem (6) and, therefore, may seem abstract to the reader. However, we think that, besides being quite general, this formulation gives a better insight into the nature of our assumptions and the spirit of our approach.

1.5.1 Example

Before stating the general assumptions, we first give for the reader's convenience a concrete set of assumptions on the coefficients A, b, d which are sufficient to fulfil the general assumptions. Note that, besides, we need the initial conditions to be "well prepared," which is not detailed in this example.

To avoid any difficulty when $|y| \to +\infty$, we can assume for example that A, b and d have a compact dependence on y. Namely, if ψ is a globally smooth diffeomorphism from $\mathbb{R}^n$ into the unit ball, we assume $A(x,y) := A_\star(x, \psi(y))$, $b(x,y) := b_\star(x,\psi(y))$, $d(x,y) := d_\star(x,\psi(y))$, where $A_\star$, $b_\star$ and $d_\star$ are defined on the closed unit ball. This way, the y space $\mathbb{R}^n$ is compactified. Then, the coefficients can be chosen to fullfill the following conditions, for all $x \geq 0$ and uniformly in $y \in \overline{B}_1$,

$$\begin{cases} A_\star(\cdot,\cdot),\ b_\star(x,\cdot) \geq 0,\ d_\star(x,\cdot) \geq 0 \text{ are } C^1, \\ A_\star(x,y) \geq \underline{A} > 0, \qquad \dfrac{1}{A_\star(\cdot,y)} \text{ is integrable on the support of } b_\star(\cdot,y), \\ b_\star(x,y) \leq K e^{Kx}, \qquad \underline{\eta} b_\star(x,y) - d_\star(x,y) \geq \underline{r}, \\ M(\cdot) \text{ is a Gaussian probability kernel,} \end{cases}$$

for some constants $K > 0$, $\underline{r} > 0$ and where $\underline{\eta}$ is determined from the initial conditions, see (16).

These assumptions can be substantially generalized. For instance, using the formula (obtained by integrating the first line in (6))

$$\Lambda(y,\eta) = \frac{\int_{\mathbb{R}_+} (d(x,y) - \eta b(x,y))\, Q(x,y,\eta) dx}{\int_{\mathbb{R}_+} Q(x,y,\eta) dx},$$

the relation between $b_\star$ and $d_\star$ is only used to ensure the inequality

$$\Lambda(y,\eta) \leq -\underline{r} < 0, \quad \forall \eta \geq \underline{\eta}$$

which is sufficient in the sequel, see (16).

1.5.2 Assumptions on the coefficients

We assume that for some $\underline{A} > 0$

$$b \geq 0, \quad d \geq 0, \quad A \geq \underline{A} > 0 \text{ are continuous functions,} \tag{9}$$

$$A(\cdot,\cdot),\ b(x,\cdot),\ d(x,\cdot) \text{ are } C^1,$$

$$d(\cdot,y),\ A(\cdot,y) \text{ are bounded from above and below in some interval of } \mathbb{R}_+, \text{ uniformly in } y. \tag{10}$$

which can be viewed as a non-degeneracy condition. Regarding the mutation kernel, we assume that

$$M(\cdot) \text{ is a probability kernel and vanishes faster than any exponential.}$$

For instance, $M(\cdot)$ can be a Gaussian distribution or have a compact support. Note that the case without mutations corresponds to $M = \delta_0$ and has been already treated in [38].

In addition, we assume

$$\exists K > 0 \text{ such that } \sup_{y\in\mathbb{R}^n} \int_0^{\overline{x}} \frac{1}{A(x',y)} dx' \leq, K \tag{11}$$

where $\overline{x}$ defines the largest support of $b(x,y)$,

$$\overline{x} := \sup\{x \geq 0 : \exists y \in \mathbb{R}^n,\ b(x,y) > 0\} \in [0, +\infty].$$

This assumption means that the transport outwards the support of b occurs in finite time (it can be seen on the characteristics). It is a necessary and sufficient condition for the ratio $\frac{\partial_\eta Q}{Q}$ to be bounded on $[0, \overline{x}]$, which is used in section 3.1 to prove an estimate on p_ε. Our approach would also work if the support of $A(\cdot, y)$ is compact, but we omit this case for simplicity.

We also need the eigenvalue Λ of (6) to be well defined and differentiable (w.r.t η). As we see it in Proposition 2.1,defining,

$$F(y,\lambda) := \int_{\mathbb{R}_+} \frac{b(x,y)}{A(x,y)} \exp\left(\int_0^x \frac{\lambda - d(x',y)}{A(x',y)} dx'\right) dx, \qquad \lambda \in \mathbb{R}, \tag{12}$$

the eigenvalue $\Lambda(y,\eta)$ is defined through the relation

$$F(y, \Lambda(y,\eta)) = \frac{1}{\eta}.$$

We assume there exists $\overline{\Lambda} < 0$ such that

$$F(y,\overline{\Lambda}),\ \partial_\lambda F(y,\overline{\Lambda}) < +\infty \quad \text{for all } y \in \mathbb{R}^n.$$

This assumption is not very restrictive, it is satisfied if $b(x,y) \leq K' e^{Kx}$ choosing $\overline{\Lambda} \leq -K$.

1.5.3 Assumptions on the initial conditions.

We need the population $m_\varepsilon(t,x,y)$ to be "well prepared" for concentration. We write $m^0_\varepsilon(x,y) = p^0_\varepsilon(x,y) e^{\frac{U^0_\varepsilon(y)}{\varepsilon}}$ according to (3), and we assume that

U^0_ε smoothly converges to a function U^0 when ε vanishes,

$$\exists k^0 > 0 \text{ such that } \forall \varepsilon > 0,\ \forall y \in \mathbb{R}^n, \quad |\nabla_y U^0_\varepsilon(y)| \leq k^0, \tag{13}$$

$$\exists C > 0 \text{ such that } \forall \varepsilon > 0,\ \forall y \in \mathbb{R}^n, \quad \partial^2_{y_i} U^0_\varepsilon(y) \geq -C, \tag{14}$$

$$\underline{J}^0 \leq \int_{\mathbb{R}^n} e^{\frac{U^0_\varepsilon(y)}{\varepsilon}} dy \leq \overline{J}^0, \text{ for some } \underline{J}^0, \overline{J}^0 > 0, \tag{15}$$

and p^0_ε such that, for some $\underline{\gamma}^0, \overline{\gamma}^0 > 0$:

$$\underline{\gamma}^0 \leq \frac{p^0_\varepsilon(x,y)}{Q(x,y,\eta^0_\varepsilon(y))} \leq \overline{\gamma}^0, \tag{16}$$

where Q is defined through (6) and we define

$$\eta^0_\varepsilon(y) := \int_{z\in\mathbb{R}^n} M(z) e^{\frac{U^0_\varepsilon(y+\varepsilon z) - U^0_\varepsilon(y)}{\varepsilon}} dz, \qquad \Lambda^0_\varepsilon(y) := \Lambda(y, \eta^0_\varepsilon(y)). \tag{17}$$

Another condition is also required on η^0_ε, see (16). Note that (13) and (14) ensures that $\rho^0_\varepsilon := \iint_{\mathbb{R}_+\times\mathbb{R}^n} m^0_\varepsilon$ is uniformly bounded. Note also that assumption (13) implies $\sup_{\mathbb{R}^n} U^0 = 0$.

1.5.4 Assumptions on the distribution of phenotypes

The following assumptions only deal with the dependence of the coefficients on $y \in \mathbb{R}^n$. In particular, if all the coefficients have a compact dependence on y, then all the following assumptions are automatically satisfied.

First, we need a condition on the initial data, namely that, from (15), $\eta_\varepsilon^0(y)$ is bounded and $\Lambda_\varepsilon^0(y)$ is bounded and negative. More precisely, in accordance to section 1.5.2, we assume that there are two negative constants $\underline{\Lambda} \leq \overline{\Lambda} < 0$ and two positive constants $0 < \underline{\eta} \leq \overline{\eta}$, such that for all $y \in \mathbb{R}^n$,

$$\frac{1}{\overline{\eta}} \leq F(y, \underline{\Lambda}) \leq F(y, \Lambda_\varepsilon^0(y)) := \frac{1}{\eta_\varepsilon^0(y)} \leq F(y, \overline{\Lambda}) \leq \frac{1}{\underline{\eta}}. \tag{18}$$

This assumption implies $\underline{\eta} \leq \eta_\varepsilon^0(y) \leq \overline{\eta}$ and $\underline{\Lambda} \leq \Lambda_\varepsilon^0(y) \leq \overline{\Lambda}$ (since $\lambda \mapsto F(y, \lambda)$ is increasing). We will see in Corollary 2.3 that those two inequalities hold for all times, namely $\underline{\eta} \leq \eta_\varepsilon(t, y) \leq \overline{\eta}$ and $\underline{\Lambda} \leq \Lambda(y, \eta_\varepsilon(t, y)) \leq \overline{\Lambda}$. Note that with the notations

$$\frac{1}{\underline{\eta}(y)} := F(y, \overline{\Lambda}), \quad \frac{1}{\overline{\eta}(y)} := F(y, \underline{\Lambda}), \tag{19}$$

assumption (16) can be written

$$\underline{\eta} \leq \underline{\eta}(y) \leq \eta_\varepsilon^0(y) \leq \overline{\eta}(y) \leq \overline{\eta}.$$

We stress out that this assumption implies $-\Lambda(y, \eta_\varepsilon(t, y)) > 0$. It means that every phenotype has a positive fitness, and is thus able to survive in absence of other phenotypes. This assumption is somehow restrictive, but it is not irrealistic, and allows to avoid some technicalities.

Finally, the next assumptions deal with the derivatives of Λ. We need $\partial_\eta \Lambda$, $\nabla_y \Lambda$ and $\nabla_y \partial_\eta \Lambda$ to be bounded, and Λ to be semi-convex. According to (23), we assume that there exists two constants $l, L > 0$ such that for all $y \in \mathbb{R}^n$, $\lambda \in [\underline{\Lambda}, \overline{\Lambda}]$,

$$l \leq \partial_\lambda F(y, \lambda) \leq L, \tag{20}$$

$$|\nabla_y F(y, \lambda)|, |\nabla_y \partial_\lambda F(y, \lambda)| \leq L, \tag{21}$$

$$\forall i \in \{1, \dots, n\}, \quad \partial_{y_i}^2 F(y, \lambda) \geq -L. \tag{22}$$

2 Definition and properties of U_ε

To make sense to the above heuristic, we first give a rigorous definition of the eigenelements (Λ, Q), which only uses classical arguments. Then, we define U_ε, formally introduced in section 1.3, and state some a priori estimates. We use those results to derive estimates on η_ε and Λ_ε. Note that the study of U_ε is autonomous and can be done separately from the analysis of the corrector. Finally, we pass to the limit, as $\varepsilon \to 0$, in the quantities U_ε, η_ε and Λ_ε to recover a viscosity solution of the Hamilton-Jacobi equation. The longer or more technical proofs are postponed to section 5.

2.1 The eigenproblem and effective Hamiltonian

We consider the limiting problem (6) and prove the existence of the eigenelements (Λ, Q), along with some properties. The proof only uses elementary arguments and is postponned to section 5.1.

Proposition 2.1 *Under the assumptions of section 1.5, for fixed $y \in \mathbb{R}^n$ and $\eta \in (\underline{\eta}(y), \overline{\eta}(y))$ (from (1.5.4)), there exists a unique couple $(\Lambda(y, \eta), Q(x, y, \eta))$ which satisfies (6).*

Moreover, with F defined in (11), $\underline{\Lambda}$ and $\overline{\Lambda}$ defined in (16), the eigenvalue $\Lambda(y, \eta)$ is continuously differentiable and it holds, for all $y \in \mathbb{R}^n$, $\eta \in (\underline{\eta}(y), \overline{\eta}(y))$,

$$F(y, \Lambda(y, \eta)) = \frac{1}{\eta}, \tag{23}$$

$$\underline{\Lambda} \leq \Lambda(y, \eta) \leq \overline{\Lambda} < 0, \tag{24}$$

$$Q(x, y, \eta) = \eta \frac{1}{A(x, y)} \exp\left(\int_0^x \frac{\Lambda(y, \eta) - d(x', y)}{A(x', y)} dx' \right). \tag{25}$$

PROOF See section 5.1. ∎

Differentiating the relation (20), we immediately obtain the derivatives of Λ

$$\partial_\eta \Lambda(y,\eta) = \frac{-1}{\eta^2 \partial_\lambda F(y,\Lambda(y,\eta))} \leq \frac{-1}{\eta^2 L}, \qquad \nabla_y \Lambda(y,\eta) = \frac{-\nabla_y F(y,\Lambda(y,\eta))}{\partial_\lambda F(y,\Lambda(y,\eta))}. \tag{26}$$

In particular, the property $\partial_\eta \Lambda < 0$, turns out to be fundamental in the following section.

We also have some kind of concavity property for Λ. For all $y \in \mathbb{R}^n$, $\eta \in (\underline{\eta}(y), \overline{\eta}(y))$, we have

$$\partial_\eta^2 \Lambda(y,\eta) + \frac{\partial_\eta \Lambda(y,\eta)}{\eta} \leq 0. \tag{27}$$

PROOF We use the short notations $F := F(y,\Lambda(y,\eta))$ (where F is defined in (11)) and $\Lambda := \Lambda(y,\eta)$. Differentiationg the first relation in (23) with respect to η, we find

$$\partial_\lambda^2 F(\partial_\eta \Lambda)^2 + \partial_\lambda F \partial_\eta^2 \Lambda = \frac{2}{\eta^3}, \qquad \partial_\eta^2 \Lambda = \frac{2}{\eta^3 \partial_\lambda F} - \frac{1}{\eta^4 (\partial_\lambda F)^3} \partial_\lambda^2 F.$$

Combining it with the expression of $\partial_\eta \Lambda(y,\eta)$ in (23) and (20), we deduce

$$\partial_\eta^2 \Lambda + \frac{\partial_\eta \Lambda}{\eta} = \frac{1}{(\partial_\lambda F)^3 \eta^3} \left((\partial_\lambda F)^2 - F \partial_\lambda^2 F \right).$$

Using the Cauchy-Schwarz inequality and the definition of F in (11), we have

$$(\partial_\lambda F)^2 - F \partial_\lambda^2 F \leq 0$$

and the proof of (24) is completed. ∎

2.2 Construction of U_ε and a priori estimates

We give a rigourous definition of U_ε, formally introduced in (7).

Proposition 2.2 *Under the assumptions of section 1.5, for all $\varepsilon > 0$ there exists a solution $U_\varepsilon(t,y)$ of (7). In addition, it satisfies, with $\underline{\Lambda}, \overline{\Lambda} < 0$ defined in (16),*

$$-\overline{\Lambda} \leq \partial_t U_\varepsilon(t,y) \leq -\underline{\Lambda}, \qquad \forall \varepsilon > 0,\ t \geq 0\ y \in \mathbb{R}^n.$$

PROOF See section 5.2. ∎

In fact, U_ε is the unique solution of (7) with a locally bounded time derivative.

As a direct consequence of Proposition 2.2, we deduce the following useful corollary.

Corollary 2.3 *With $\underline{\Lambda}$, $\overline{\Lambda}$, $\underline{\eta}$, $\overline{\eta}$ defined in (16), and setting*

$$\eta_\varepsilon(t,y) := \int_{\mathbb{R}^n} M(z) e^{\frac{U_\varepsilon(t,y+\varepsilon z) - U_\varepsilon(t,y)}{\varepsilon}} dz, \qquad \Lambda_\varepsilon(t,y) := \Lambda(y, \eta_\varepsilon(t,y)), \tag{28}$$

we have

$$\underline{\eta} \leq \eta_\varepsilon(t,y) \leq \overline{\eta}, \qquad \underline{\Lambda} \leq \Lambda_\varepsilon(t,y) \leq \overline{\Lambda} < 0. \tag{29}$$

PROOF Simply use Proposition 2.2, $\Lambda_\varepsilon = -\partial_t U_\varepsilon$ and assumption (16). ∎

2.3 Further estimates

Proposition 2.4 *Under the assumptions of section 1.5, with k^0 defined in (1.5.3), $L, l > 0$ in (17) and $\underline{\eta}$ in (26), we have*

$$|\nabla_y U_\varepsilon(t,y)| \leq k^0 + \frac{L}{l\underline{\eta}^2} t, \qquad \forall \varepsilon > 0,\ t \geq 0,\ y \in \mathbb{R}^n.$$

PROOF See section 5.3. ∎

Note that the coefficient $\frac{L}{l\underline{\eta}^2}$ comes from a bound on $|\nabla_y \Lambda(y, \eta_\varepsilon)|$. We will see that, at the limit when $\varepsilon \to 0$, we can prove Lipschitz continuity globally in time.

We also need the following control of second order derivatives.

Proposition 2.5 (Semi-convexity) *The function U_ε is semi-convex in (t,y), that is, for all the $\nu \in \mathbb{S} := \{(t,y) \in \mathbb{R}^{n+1} : t^2 + |y|^2 = 1\}$, $\partial^2_{\nu\nu} U_\varepsilon$ are bounded from below, uniformly in $\varepsilon > 0$, $y \in \mathbb{R}^n$, locally uniformly in $t \geq 0$. Therefore U_ε belongs to $W^{2,1}_{loc}$ in (t,y), uniformly in $\varepsilon > 0$.*

PROOF The idea is to use that the Hamiltonian has properties closely related to convexity, namely (24) and to use the Lipschitz bounds. See section 5.4. ∎

The following corollary is essential when studying the corrector in section 3.

Corollary 2.6 *We have $\eta_\varepsilon \in W^{1,1}_{loc}$, uniformly in $\varepsilon > 0$. In addition, denoting $\overline{\eta}(t) := \sup_{y \in \mathbb{R}^n} \eta_\varepsilon(t,y)$, we have*

$$\int_0^T |\partial_t \overline{\eta}_\varepsilon(t)|\, dt \text{ is bounded uniformly in } \varepsilon > 0, \qquad \forall T \geq 0.$$

PROOF See section 5.5. ∎

2.4 Asymptotics

With these regularity properties, we are ready to establish the asymptotics of U_ε when ε vanishes.

Proposition 2.7 *Under the assumptions of section 1.5, when ε vanishes, U_ε converges locally uniformly (and in $W^{1,r}_{loc}$, $1 \leq r < \infty$) to a function $U(t,y) \in W^{1,\infty}_{loc}$ which is a semi-convex viscosity solution of*

$$\begin{cases} \partial_t U(t,y) = -\Lambda\left(y, \displaystyle\int_{\mathbb{R}^n} M(z) e^{\nabla_y U(t,y)\cdot z} dz\right), & \forall t > 0,\ \forall y \in \mathbb{R}^n, \\ U(0,y) = U^0(y), & \forall y \in \mathbb{R}^n. \end{cases} \tag{30}$$

Moreover, under the assumption that M is not degenerate (i.e., $M(\cdot) > 0$ in a neighborhood of 0), the function $U(t,y)$ is globally Lipschitzian.

PROOF See section 5.6 and section 5.7. ∎

We also point out that, from Proposition 4.3 in [38],

$$p \mapsto -\Lambda\left(y, \int_{\mathbb{R}^n} M(z) e^{p\cdot z} dz\right) \text{ is a convex mapping, } \forall y \in \mathbb{R}^n.$$

This class of Hamiltonian has been widely studied, and numerous results on regularity as well as representation formula are available [9, 25].

As a direct consequence of the L^r_{loc} convergence of ∇U_ε to ∇U, we have the following corollary.

Corollary 2.8 *When $\varepsilon \to 0$, η_ε converges in L^r_{loc}, $1 \leq r < \infty$, to*

$$\eta(t,y) := \int_{\mathbb{R}^n} M(z) e^{\nabla_y U \cdot z} dz.$$

Consequently, $\Lambda_\varepsilon(t,y) := \Lambda(t, \eta_\varepsilon(t,y))$ converges to $\Lambda(y, \eta(t,y))$ and $Q_\varepsilon(t,x,y) := Q(x,y,\eta_\varepsilon(t,y))$ to $Q(x,y,\eta(t,y))$ in L^r_{loc}, $1 \leq r < \infty$.

3 Asymptotics of the corrector - Proof of Theorem 1.2

We now turn to the main new results of the paper. These are boundedness from above and below and the asymptotics of the corrector $p_\varepsilon(t,x,y)$ defined through the factorisation (3), according to the definition of U_ε in (7).

3.1 Estimates on p_ε

This section is devoted to the proof of the first statement of Theorem 1.2, that is, $p_\varepsilon(t,x,y)$ and $\int_{x>0} p_\varepsilon(t,x,y)$ are bounded uniformly in $\varepsilon>0$, $x\geq 0$, $y\in\mathbb{R}^n$, locally uniformly in $t\geq 0$.

Our first result states a control of p_ε, from above and below, for $x\in[0,\overline{x}]$ (with $\overline{x}$ from (10)).

Lemma 3.1 *Under the assumptions of section 1.5, for any fixed $T>0$, there exists two constants $\underline{\gamma},\overline{\gamma}>0$ such that*

$$\underline{\gamma}\; Q_\varepsilon(t,x,y)\leq p_\varepsilon(t,x,y)\leq \overline{\gamma}\; Q_\varepsilon(t,x,y),$$

for all $\varepsilon>0$, $t\in[0,T]$, $y\in\mathbb{R}^n$, $x\in[0,\overline{x}]$, where

$$Q_\varepsilon(t,x,y):=Q(x,y,\eta_\varepsilon(t,y)).$$

Proof The function $p_\varepsilon(t,x,y)$ satisfies the following equation, for $\varepsilon>0,\ t>0,\ x>0,\ y\in\mathbb{R}^n$,

$$\begin{cases}\varepsilon\partial_t p_\varepsilon+\partial_x\left[A(x,y)p_\varepsilon\right]+\left(d(x,y)-\Lambda_\varepsilon\right)p_\varepsilon=0,\\ A(0,y)p_\varepsilon(t,0,y)=\displaystyle\iint\limits_{x>0,\ z\in\mathbb{R}^n} M(z)e^{\frac{U_\varepsilon(t,y+\varepsilon z)-U_\varepsilon(t,y)}{\epsilon}}b(x,y+\varepsilon z)p_\varepsilon(t,x,y+\varepsilon z)dxdz.\end{cases}$$

and Q_ε satisfies, for (t,y) as parameters, the equation in the variable x

$$\begin{cases}\partial_x[A(x,y)Q_\varepsilon]-(d(x,y)-\Lambda_\varepsilon(t,y))Q_\varepsilon=0,\\ A(0,y)Q_\varepsilon(t,0,y)=\eta_\varepsilon(t,y).\end{cases}$$

Setting

$$\gamma_\varepsilon(t,x,y):=\frac{p_\varepsilon(t,x,y)}{Q_\varepsilon(t,x,y)},$$

we have

$$\begin{cases}\partial_t\gamma_\varepsilon+\dfrac{A(x,y)}{\varepsilon}\partial_x\gamma_\varepsilon=-\dfrac{\partial_t Q_\varepsilon}{Q_\varepsilon}\gamma_\varepsilon,\\ \gamma_\varepsilon(t,0,y)=\displaystyle\iint\limits_{x>0,\ z\in\mathbb{R}^n} J_\varepsilon(t,x,y,z)\gamma_\varepsilon(t,x,y+\varepsilon z)dz,\end{cases}\tag{31}$$

where

$$J_\varepsilon(t,x,y,z):=\frac{1}{\eta_\varepsilon}M(z)e^{\frac{U_\varepsilon(t,y+\varepsilon z)-U_\varepsilon(t,y)}{\epsilon}}b(x,y+\varepsilon z)Q_\varepsilon(t,x,y+\varepsilon z).\tag{32}$$

Our goal is to infer some bounds on γ_ε.

First, from the definition of η_ε and the normalization $\int_{x>0}b(x,y)Q(x,y,\eta)dx=1$, we see that J_ε is a probability kernel,

$$\iint_{x>0,\ z\in\mathbb{R}^n} J_\varepsilon(t,x,y,z)dxdz=1,\qquad \forall t\geq 0,\ y\in\mathbb{R}^n.$$

We need to estimate $\frac{\partial_t Q_\varepsilon}{Q_\varepsilon}$. We compute, using the representation formula (22),

$$\begin{aligned}\frac{\partial_t Q_\varepsilon}{Q_\varepsilon}(t,x,y)&=\partial_t\eta_\varepsilon(t,y)\frac{\partial_\eta Q}{Q}(x,y,\eta_\varepsilon(t,y))\\ &=\partial_t\eta_\varepsilon(t,y)\left(\frac{1}{\eta_\varepsilon(t,y)}+\partial_\eta\Lambda(y,\eta_\varepsilon(t,y))\int_0^x\frac{1}{A(x',y)}dx'\right).\end{aligned}$$

Because η_ε is bounded from below, see (26), and $\partial_\eta \Lambda$ is bounded, see (23) and assumption (17), we have

$$\left|\frac{\partial_t Q_\varepsilon}{Q_\varepsilon}(t,x,y)\right| \le K|\partial_t \eta_\varepsilon(t,y)| \left(1 + \int_0^x \frac{1}{A(x',y)}dx'\right),$$

for some constant $K > 0$. Then, using assumption (10), we have, for $x \in [0, \overline{x}]$,

$$\left|\frac{\partial_t Q_\varepsilon}{Q_\varepsilon}(t,x,y)\right| \le K|\partial_t \eta_\varepsilon(t,y)|$$

for some constant still denoted by K. Setting $\overline{\eta}_\varepsilon(t) := \sup_{y\in\mathbb{R}^n} \eta_\varepsilon(t,y)$, we define

$$\underline{\gamma}_\varepsilon(t,x,y) := \gamma_\varepsilon(t,x,y) \exp\left(-K\int_0^t |\partial_t \overline{\eta}_\varepsilon(t')|dt'\right),$$

so that $\underline{\gamma}_\varepsilon$ is a subsolution to (28), namely

$$\begin{cases} \partial_t \underline{\gamma}_\varepsilon + \dfrac{A(x,y)}{\varepsilon}\partial_x \underline{\gamma}_\varepsilon \le 0 \\ \underline{\gamma}_\varepsilon(t,0,y) \le \displaystyle\iint_{x>0,\ z\in\mathbb{R}^n} J_\varepsilon(t,x,y,z)\underline{\gamma}_\varepsilon(t,x,y+\varepsilon z)dzdx \end{cases}$$

(in fact, equality holds in the second line). From the comparison principle, we deduce

$$\underline{\gamma}_\varepsilon(t,x,y) \le \sup_{\substack{x\in[0,\overline{x}]\\ y\in\mathbb{R}^n}} \underline{\gamma}_\varepsilon(0,x,y) \le \overline{\gamma}^0,$$

where $\overline{\gamma}^0$ comes from assumption (14). This gives a control from above on $\gamma_\varepsilon(t,x,y)$ which implies

$$p_\varepsilon(t,x,y) \le \overline{\gamma}^0 Q_\varepsilon(t,x,y) \exp\left(K\int_0^t |\overline{\partial}_t \eta_\varepsilon(t')|dt'\right).$$

By Corollary 2.6, $\int_0^t |\partial_t \overline{\eta}_\varepsilon|$ is bounded uniformly in $\varepsilon > 0$, therefore, for some constant $\overline{\gamma}$,

$$p_\varepsilon(t,x,y) \le \overline{\gamma} Q_\varepsilon(t,x,y).$$

Identically, we infer the bound from below, and the proof of Lemma 3.1 is completed. ∎

With the previous lemma in hand, we now estimate p_ε for all $x \ge 0$ (which is useless if $\overline{x} = +\infty$). We set

$$\overline{Q}(x,y) := \overline{\gamma}\frac{\overline{\eta}}{\underline{\eta}(y)}Q(x,y,\underline{\eta}(y)) = \overline{\gamma}\frac{\overline{\eta}}{A(x,y)} \exp\left(\int_0^x \frac{\overline{\Lambda} - d(x',y)}{A(x',y)}dx'\right),$$
$$\underline{Q}(x,y) := \underline{\gamma}\frac{\underline{\eta}}{\overline{\eta}(y)}Q(x,y,\overline{\eta}(y)) = \underline{\gamma}\frac{\underline{\eta}}{A(x,y)} \exp\left(\int_0^x \frac{\underline{\Lambda} - d(x',y)}{A(x',y)}dx'\right),$$

where $\underline{\gamma}, \overline{\gamma}$ are given by the previous lemma.

Lemma 3.2 *Under the same condition as in the previous lemma, we have*

$$\underline{Q}(x,y) \le p_\varepsilon(t,x,y) \le \overline{Q}(x,y), \qquad \forall \varepsilon > 0,\ t \in [0,T],\ x \ge 0,\ y \in \mathbb{R}^n.$$

Note that, when restricted to $[0,\overline{x}]$ these bounds are weaker than in Lemma 3.1 since $\frac{1}{\underline{\gamma}}\underline{Q} \le Q_\varepsilon \le \frac{1}{\overline{\gamma}}\overline{Q}$.

PROOF From Lemma 3.1, we deduce

$$\underline{\gamma}\ \underline{\eta} \leq A(0,y)p_\varepsilon(t,0,y) \leq \overline{\gamma}\ \overline{\eta}.$$

Hence, on the one hand, we have

$$\begin{cases} \varepsilon\partial_t p_\varepsilon + \partial_x\left[A(x,y)p_\varepsilon\right] + \left(d(x,y) - \overline{\Lambda}\right)p_\varepsilon \leq 0, \\ A(0,y)p_\varepsilon(t,0,y) \leq \overline{\gamma}\ \overline{\eta}, \end{cases}$$

for $\varepsilon > 0,\ t \in [0,T],\ x > 0,\ y \in \mathbb{R}^n$. On the other hand,

$$\begin{cases} \varepsilon\partial_t \overline{Q} + \partial_x\left[A(x,y)\overline{Q}\right] + \left(d(x,y) - \overline{\Lambda}\right)\overline{Q} = 0, \\ A(0,y)\overline{Q}(0,y) = \overline{\gamma}\ \overline{\eta}. \end{cases}$$

From the comparison principle, we deduce $p_\varepsilon \leq \overline{Q}$. The lower bound can be proved similarily. ∎

We are now ready to prove our first main result

Proposition 3.3 (Uniform estimates on p_ε in $L^1 \cap L^\infty$) *Under the assumptions of section 1.5 and for any fixed $T > 0$, $p_\varepsilon(t,x,y)$ and $\int_{x>0} p_\varepsilon(t,x,y)dx$ are bounded from above and below, uniformly in $\varepsilon > 0$, $t \in [0,T]$, $x \geq 0$, $y \in \mathbb{R}^n$.*

PROOF The first point is deduced from Lemma 3.2 and $\overline{Q} \leq \overline{\gamma}\frac{\overline{\eta}}{\underline{A}}$, with $\underline{A}$ from assumption (8).

To prove the second point, we only need to estimate the integrals of $\underline{Q}$ and $\overline{Q}$. Recalling $\overline{\Lambda} < 0$, we compute

$$\begin{aligned} \int_0^{+\infty} \overline{Q}(x,y)dx &\leq \int_0^{+\infty} \overline{\gamma}\frac{\overline{\eta}}{A(x,y)} \exp\left(\int_0^x \frac{\overline{\Lambda}}{A(x',y)}dx'\right) \\ &= \left[\overline{\gamma}\frac{\overline{\eta}}{\overline{\Lambda}} \exp\left(\int_0^x \frac{\overline{\Lambda}}{A(x',y)}dx'\right)\right]_{x=0}^{+\infty} \\ &= \overline{\gamma}\frac{\overline{\eta}}{-\overline{\Lambda}}\left[1 - \exp\left(\int_0^{+\infty} \frac{\overline{\Lambda}}{A(x',y)}dx'\right)\right] \leq \overline{\gamma}\frac{\overline{\eta}}{-\overline{\Lambda}}, \end{aligned}$$

which proves the bound from above.

For the other inequality, we use the non-degeneracy assumption (9) which implies $\int_{\mathbb{R}_+} Q(x,y,\eta)dx > \alpha$, for some $\alpha > 0$. ∎

3.2 Asymptotics of p_ε

We complete the proof of Theorem 1.2 by showing that, when $\varepsilon \to 0$, the mutations affect the equilibrium distribution $Q(x,y,\eta)$ only through a multiplicative factor $\gamma(t,y)$,

Proposition 3.4 (Convergence of the corrector) *For $t \in [0,T]$ (with $T > 0$ fixed), $x \in [0,\overline{x}]$, $y \in \mathbb{R}^n$, and up to extraction of a subsequence when $\varepsilon \to 0$, p_ε converges in L^∞ weak-$\star$ to $\gamma(t,y)Q(x,y,\eta(t,y))$ with $\gamma(t,y) \in L^\infty$ which formally satisfies the equation*

$$\begin{cases} \partial_t\gamma + \partial_\eta\Lambda\left(\displaystyle\int_{z\in\mathbb{R}^n} M(z)e^{\nabla U\cdot z}zdz\right)\cdot\nabla_y\gamma + \partial_t\eta\frac{\partial_\eta^2\Lambda}{2\partial_\eta\Lambda}\gamma = 0, \\ \gamma(t=0) = \gamma^0. \end{cases} \tag{33}$$

The difficulty in stating rigorously equation (30) is that $\partial_t\eta$ and $\int_{z\in\mathbb{R}^n} M(z)e^{\nabla U\cdot z}zdz$ are nothing more than bounded measures which is not smooth enough since γ is just L^∞. Nevertheless, we can prove establish the convergence of the full sequence p_ε to γQ if T is small enough: we use the regularity of the initial condition to rigourously pass to the limiting equation (30), and then use a standard uniqueness result on this equation.

Note also that, according to Theorem 1.3, the population concentrates on $\mathcal{S}$ where $U(t,y)$ achieved its maximum. From [7], U is differentiable at points of $\mathcal{S}$ and one has $\partial_t U = \nabla_y U = 0$ and $\eta = 1$. Thus p_ε converges to a multiple of $Q(x,y,1)$ on $\mathcal{S}$. In addition, if $M(\cdot)$ is even, then the drift term vanishes and the equation can be written

$$\partial_t[\partial_\eta \Lambda(y,\eta(t,y))^{1/2}\gamma] = 0.$$

To prove Proposition 3.4, and establish (30) we use the dual problem associated with (6); for fixed $(y,\eta) \in \mathbb{R}^n \times (0,+\infty)$ consider $\Phi(x,y,\eta)$ the unique solution of

$$\begin{cases} A(x,y)\partial_x \Phi + [\Lambda(y,\eta) - d(x,y)]\,\Phi = -\eta b(x,y)\Phi(0,y,\eta), \quad \forall x > 0, \\ \displaystyle\int_{x>0} Q(x,y,\eta)\Phi(x,y,\eta)dx = 1. \end{cases} \tag{34}$$

We can solve this ordinary differential equation and find

$$\Phi(x,y,\eta) = -\partial_\eta \Lambda(y,\eta)\eta \int_x^{+\infty} \frac{b(x',y)}{A(x',y)} \exp\left(\int_x^{x'} \frac{\Lambda(y,\eta) - d(x'',y)}{A(x'',y)} dx''\right) dx'.$$

Also, multiplying equation (31) by $\partial_\eta Q$ and integrating by parts, we get

$$\Phi(0,y,\eta) = -\partial_\eta \Lambda(y,\eta), \tag{35}$$

and, multiplying equation (31) by $\partial^2_{\eta\eta} Q$, we end up with

$$\int_0^{+\infty} Q(x,y,\eta)\partial_\eta \Phi(x,y,\eta)dx = -\frac{\partial^2_\eta \Lambda}{2\partial_\eta \Lambda}. \tag{36}$$

Proof We fix $T > 0$ and, throughout the proof, choose $t \in [0,T]$, $x \in [0,\overline{x}]$, $y \in \mathbb{R}^n$. We define $\gamma_\varepsilon := \frac{p_\varepsilon}{Q_\varepsilon}$ and recall equation (28). From Lemma 3.1, we know that γ_ε is bounded and thus converges (up to extraction of a subsequence) to some γ in L^∞ weak-$\star$. Passing to the limit in (28), we deduce $\partial_x \gamma \equiv 0$ (in the sense of distributions), and thus γ only depends on t and y. Since Q_ε (strongly) converges to some Q (see Corollary 2.8), we deduce that $p_\varepsilon(t,x,y)$ converges (up to extraction of a subsequence) to $\gamma(t,y)Q(t,x,y)$ in L^∞ weak-$\star$.

We are now left with the task of identifying γ. To do so, we set

$$E_\varepsilon(t,y) := \int_0^{+\infty} \gamma_\varepsilon(t,x,y)Q_\varepsilon(t,x,y)\Phi_\varepsilon(t,x,y)dx, \qquad \Phi_\varepsilon(t,x,y) := \Phi(x,y,\eta_\varepsilon(t,y)).$$

Note that $\Phi_\varepsilon(t,x,y) \equiv 0$ for $x > \overline{x}$ and thus $E_\varepsilon(t,y) := \int_0^{\overline{x}} \gamma_\varepsilon Q_\varepsilon \Phi_\varepsilon$. From (6), (28) and (31), we can write

$$\begin{aligned}
\varepsilon\partial_t E_\varepsilon &= \int_{x>0} \varepsilon\partial_t\gamma_\varepsilon\, Q_\varepsilon\Phi_\varepsilon + \varepsilon\int_{x>0} \gamma_\varepsilon\partial_t[Q_\varepsilon\Phi_\varepsilon] \\
&= -\int_{x>0} A\partial_x\gamma_\varepsilon\, Q_\varepsilon\Phi_\varepsilon - \varepsilon\int_{x>0}\gamma_\varepsilon\partial_t Q_\varepsilon\Phi_\varepsilon + \varepsilon\int_{x>0}\gamma_\varepsilon\partial_t[Q_\varepsilon\Phi_\varepsilon] \\
&= [A\gamma_\varepsilon Q_\varepsilon\Phi_\varepsilon](x=0) + \int_{x>0}\gamma_\varepsilon\partial_x[AQ_\varepsilon\Phi_\varepsilon] + \varepsilon\int_{x>0}\gamma_\varepsilon Q_\varepsilon\partial_t\Phi_\varepsilon
\end{aligned}$$

(from $\partial_x[AQ_\varepsilon\Phi_\varepsilon] = -\eta_\varepsilon b Q_\varepsilon\Phi_\varepsilon(x=0)$, and with the probability kernel J_ε defined in (29))

$$= \eta_\varepsilon\Phi_\varepsilon(x=0) \iint_{x>0,\ z\in\mathbb{R}^n} J_\varepsilon(t,x,y,z)\left(\gamma_\varepsilon(t,x,y+\varepsilon z) - \gamma_\varepsilon(t,x,y)\right)dzdx + \varepsilon\int_{x>0}\gamma_\varepsilon Q_\varepsilon\partial_t\Phi_\varepsilon$$

Recalling Corollary 2.8, we know that Φ_ε converges to some Φ when $\varepsilon \to 0$. Dividing by ε and passing to the limit $\varepsilon \to 0$ (after extracting a subsequence, in the sense of distributions), we deduce

$$\partial_t\left[\gamma\int_{x>0} Q\Phi\right] = \partial_t\gamma = \Phi(t,0,y)\left(\int_{z\in\mathbb{R}^n} M(z)e^{\nabla U\cdot z}z\right)\cdot\nabla_y\gamma + \gamma\int_{x>0} Q\partial_t\Phi. \tag{37}$$

Injecting $\partial_t \Phi_\varepsilon(t,x,y) = \partial_t \eta_\varepsilon \partial_\eta \Phi(x,y,\eta_\varepsilon)$ and (32)-(33) in (34), we end up with equation (30). From classical uniqueness results, we deduce that γ_ε converges weakly to γ for the whole sequence $\varepsilon \to 0$ (and not for an extracted subsequence). ∎

4 Concentration of the population density - proof of Theorem 1.3

We now conclude on the consequences of our study on U_ε and p_ε with the concentration effect for the population density m_ε itself.

4.1 Selection of the fittest phenotypes

The following result states that the total population ρ_ε is uniformly bounded and converges when $\varepsilon \to 0$. Recalling (3), the two first statements of Theorem 1.3 are direct consequences of the following proposition and the uniform L^1 estimate on p_ε (Proposition 3.3).

Proposition 4.1 *Under the assumptions of section 1.5,*
1. There exist two positive constants $\underline{\rho}, \overline{\rho} > 0$ such that

$$\underline{\rho} \le \rho_\varepsilon(t) \le \overline{\rho}, \quad \forall \varepsilon > 0,\ t \ge 0.$$

In addition, ρ_ε converges to some ρ in the L^∞-weak* topology.
2. The integral $\int_{\mathbb{R}^n} e^{\frac{U_\varepsilon(t,y) - \int_0^t \rho_\varepsilon}{\varepsilon}} dy$ is bounded away from 0, uniformly in $\varepsilon > 0$, $t \ge 0$.
Consequently, when ε vanishes, we have

$$\int_0^t \rho = \sup_{y \in \mathbb{R}^n} U(t,y), \quad \forall t \ge 0. \tag{38}$$

In addition, with $\mathcal{S}$ defined in (2), we have

$$\mathcal{S} := \left\{ t \ge 0,\ y \in \mathbb{R}^n : U(t,y) = \sup_{\mathbb{R}^n} U(t,\cdot) \right\} = \left\{ t \ge 0,\ y \in \mathbb{R}^n : U(t,y) = \int_0^t \rho \right\}.$$

Remark 4.2 *T,he proof of Proposition 4.1 becomes much simpler if we assume that there are $\underline{r}, \overline{r} > 0$ such that*

$$\underline{r} \le b(x,y) - d(x,y) \le \overline{r}, \quad \forall x \ge 0,\ y \in \mathbb{R}^n,$$

Indeed, integrating (1) *and using an integration by parts, we obtain*

$$\begin{aligned} \varepsilon \frac{\mathrm{d}}{\mathrm{d}t} \rho_\varepsilon(t) &= \iint_{\mathbb{R}^n \times \mathbb{R}_+} \left[\left(\frac{1}{\varepsilon^n} \int_{\mathbb{R}^n} M(\frac{y'-y}{\varepsilon}) dy \right) b(x,y') - d(x,y') \right] m_\varepsilon(t,x,y') dx dy' - \rho_\varepsilon^2(t) \\ &\le \overline{r} \rho_\varepsilon(t) - \rho_\varepsilon^2(t). \end{aligned}$$

which implies $0 \le \rho_\varepsilon(t) \le \max\left(\overline{r}, \rho_\varepsilon^0\right)$ and provides us with an a priori upper bound on ρ_ε. With the same method, we also infer a positive lower bound on ρ_ε.

Then, using the uniform L^1 estimate on p_ε from Proposition 3.3, we directly deduce (35).

Proof (of Proposition 4.1) We recall that

$$\rho_\varepsilon(t) = \iint_{x,y} p_\varepsilon(t,x,y) e^{\frac{U_\varepsilon(t,y) - \int_0^t \rho_\varepsilon}{\varepsilon}} dx dy. \tag{39}$$

Multiplying by $e^{\int_0^t \frac{\rho_\varepsilon}{\varepsilon}}$ we have

$$\rho_\varepsilon(t) e^{\frac{\int_0^t \rho_\varepsilon}{\varepsilon}} = \int_{\mathbb{R}^n} e^{\frac{U_\varepsilon(t,y)}{\varepsilon}} \int_{\mathbb{R}_+} p_\varepsilon(t,x,y) dx dy,$$

and integrating over $(0,t)$, we deduce

$$\varepsilon\left(e^{\frac{\int_0^t \rho_\varepsilon}{\varepsilon}}-1\right)=\int_0^t\int_{\mathbb{R}^n} e^{\frac{U_\varepsilon(t,y)}{\varepsilon}}\int_{\mathbb{R}_+} p_\varepsilon(t,x,y)dxdydt.$$

From $0<-\overline{\Lambda}\le\partial_t U_\varepsilon\le-\underline{\Lambda}$ (Proposition 2.2) and the $L^1(dx)$ estimate on p_ε (Proposition 3.3), we have

$$\begin{aligned}\varepsilon\left(e^{\frac{\int_0^t \rho_\varepsilon}{\varepsilon}}-1\right)&\ge \underline{I}\int_0^t\int_{\mathbb{R}^n}\frac{\varepsilon}{-\underline{\Lambda}}\frac{\partial_t U_\varepsilon(t,y)}{\varepsilon}e^{\frac{U_\varepsilon(t,y)}{\varepsilon}}dy\ dt\\&\ge\frac{\varepsilon\underline{I}}{-\underline{\Lambda}}\int_{\mathbb{R}^n} e^{\frac{U_\varepsilon(t,y)}{\varepsilon}}-e^{\frac{U_\varepsilon^0(y)}{\varepsilon}}dy.\end{aligned}$$

Dividing by $\varepsilon e^{\frac{\int_0^t\rho_\varepsilon}{\varepsilon}}$ on both sides we find

$$\left(1-e^{-\frac{\int_0^t\rho_\varepsilon}{\varepsilon}}\right)\ge\frac{\underline{I}}{-\underline{\Lambda}}\int_{\mathbb{R}^n}\left(e^{\frac{U_\varepsilon(t,y)}{\varepsilon}}-e^{\frac{U_\varepsilon^0(y)}{\varepsilon}}\right)e^{-\frac{\int_0^t\rho_\varepsilon}{\varepsilon}}dy,$$

that we rewrite, with $u_\varepsilon(t,y)=U_\varepsilon(t,y)-\int_0^t\rho_\varepsilon$

$$\int_{\mathbb{R}^n}e^{\frac{u_\varepsilon(t,y)}{\varepsilon}}dy\le\frac{-\underline{\Lambda}}{\underline{I}}\left(1-e^{-\frac{\int_0^t\rho_\varepsilon}{\varepsilon}}\right)+e^{-\frac{\int_0^t\rho_\varepsilon}{\varepsilon}}\int_{\mathbb{R}^n}e^{\frac{U_\varepsilon^0(y)}{\varepsilon}}dy.$$

Then, from assumption (13), we deduce

$$\int_{\mathbb{R}^n}e^{\frac{u_\varepsilon(t,y)}{\varepsilon}}dy\le\frac{-\underline{\Lambda}}{\underline{I}}\left(1-e^{-\frac{\int_0^t\rho_\varepsilon}{\varepsilon}}\right)+\overline{J}^0e^{-\frac{\int_0^t\rho_\varepsilon}{\varepsilon}},$$

and

$$\int_{\mathbb{R}^n}e^{\frac{u_\varepsilon(t,y)}{\varepsilon}}dy\le\overline{K}:=\max\left(\frac{-\underline{\Lambda}}{\underline{I}},\overline{J}^0\right).$$

Identically, we infer

$$\int_{\mathbb{R}^n}e^{\frac{u_\varepsilon(t,y)}{\varepsilon}}dy\ge\underline{K}:=\min\left(\frac{-\overline{\Lambda}}{\overline{I}},\underline{J}^0\right)>0.$$

Now, from the definition of ρ_ε (recalled in (36)) and the L^1 estimate on p_ε (Proposition 3.3) we find

$$\underline{I}\ \underline{K}\le\rho_\varepsilon(t)\le\overline{I}\ \overline{K}.$$

Since $\rho_\varepsilon(t)$ is uniformly bounded, there exists a sequence $\varepsilon_k\to0$ such that ρ_{ε_k} converges to some ρ in the $\mathcal{L}^\infty$-weak* topology when $k\to+\infty$. Since, in addition, U_ε converges locally uniformly to U (Proposition 2.7), we deduce that u_{ε_k} converges locally uniformly to some u.

Now, from $\underline{K}\le\int_{\mathbb{R}^n}e^{\frac{u_\varepsilon(t,\cdot)}{\varepsilon}}\le\overline{K}$, at the limit $k\to+\infty$, we have

$$\forall t>0,\ \sup_{y\in\mathbb{R}^n}u(t,y)=0.$$

We deduce

$$\int_0^t\rho=\sup_{y\in\mathbb{R}^n}U(t,y).$$

Therefore, $\int_0^t\rho$ does not depend on the extracted subsequence, and the convergence occurs for the whole sequence $\varepsilon\to0$, which achieves the proof. ∎

Notice that the bound on $\rho(s)$ can be made more precise in the Cesaro sense

$$\forall t>0,\quad-\overline{\Lambda}\le\frac{1}{t}\int_0^t\rho(s)ds\le-\underline{\Lambda}$$

since, from Proposition 2.2 we know that $\int_0^t\rho=\sup_{\mathbb{R}^n}U(t,\cdot)\le\sup_{\mathbb{R}^n}U^0(\cdot)-\underline{\Lambda}t=-\underline{\Lambda}t$ and the bound from below is similar.

4.2 Adaptive dynamics

We now prove the third statement of Theorem 1.3. We need further assumptions on the initial conditions,

$$\exists!\ \bar{y}^0_\varepsilon \in \mathbb{R}^n, \quad U^0_\varepsilon(\bar{y}^0_\varepsilon) = \max_{y\in\mathbb{R}^n} U^0_\varepsilon(y) = 0, \tag{40}$$

$$\bar{y}^0_\varepsilon \text{ converges to some } \bar{y}^0 \in \mathbb{R}^n \text{ when } \varepsilon \text{ vanishes,}$$

$$\nabla^2_y U^0(\bar{y}^0) < 0. \tag{41}$$

Proposition 4.3 *Under the assumptions of section 1.5 and* (37)–(38)*, for a short time interval* $[0,T]$*, there exists a unique* $\bar{y}(t) \in \mathbb{R}^n$ *on which* $U(t,\cdot)$ *reaches its maximum. Moreover,* $t \mapsto \bar{y}(t) \in \mathcal{C}^1$ *and satisfies the Canonical Equation*

$$\begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}t}\bar{y}(t) = \left(\nabla^2_y U(t,\bar{y}(t))\right)^{-1} \cdot \nabla_y \Lambda(\bar{y}(t),1) + \partial_\eta \Lambda(\bar{y}(t),1) \displaystyle\int_{\mathbb{R}^n} M(z) z dz, \\ \bar{y}(0) = \bar{y}^0. \end{cases} \tag{42}$$

.

Note that (39) features a drift term $\partial_\eta \Lambda \int_{\mathbb{R}^n} M(z) z dz$. If the mutation kernel $M(\cdot)$ is even, this term vanishes and we recover the classical Canonical Equation. Let us also recall that uniqueness and regularity of a unique concentration point (monomorphism) is a hard questions with few progresses, see [7, 14, 36].

Proof Since U (defined in Proposition 2.7) satisfies (27) with smooth initial datum, it is uniformly C^2 in the y variable for short times $t \in [0,T]$, $T > 0$ (this can be proved with the method of the characteristics, see Chapter 3.2 in [23]).

Consider such a time interval $[0,T]$, and $V \subset \mathbb{R}^n$ a neighborhood of y^0. We are interested in the solutions $(t,\bar{y}) \in [0,T] \times V$ of

$$\nabla_y U(t,\bar{y}) = 0. \tag{43}$$

From (37) we know that at initial time there exists a unique solution $\bar{y}^0$ of (40). Besides, $\nabla^2_y U^0(\bar{y}^0)$ is invertible. From the implicit functions theorem, there exists a unique $\bar{y}(t) \in \mathbb{R}^n$ satisfying (40), for t in a certain time interval still denoted $[0,T]$. We can again choose a smaller T to ensure that $\bar{y}(t)$ remains in V.

From (38), we can also choose T and V small enough to guarantee, $\forall t \in [0,T]$,

$$\bar{y}(t) \in V, \quad U(t,\cdot) \text{ is striclty concave in } V,$$
$$\max_{y\in V} U(t,y) = \max_{y\in\mathbb{R}^n} U(t,y).$$

Hence, for all $t \in [0,T]$, the solution $\bar{y}(t)$ of (40) must satisfy

$$U(t,\bar{y}(t)) = \max_{y\in V} U(t,y) = \max_{y\in\mathbb{R}^n} U(t,y) = \int_0^t \rho, \tag{44}$$

which proves the first part of the proposition and that $t \mapsto \bar{y}(t)$ is $\mathcal{C}^1$.

For the Canonical Equation, we differentiate (40) with respect to t, and noting that

$$\eta(t,\bar{y}(t)) = \int_{\mathbb{R}^n} M(z) e^{\nabla_y U(t,\bar{y}(t))\cdot z} dz = 1,$$

we obtain

$$\begin{aligned} 0 &= \frac{\mathrm{d}}{\mathrm{d}t}\left[\nabla_y U(t,\bar{y}(t))\right] \\ &= -\nabla_y \Lambda(\bar{y}(t),1) - \nabla^2_y U(t,\bar{y}(t)) \cdot \partial_\eta \Lambda(\bar{y}(t),1) \int_{\mathbb{R}^n} M(z) z dz + \nabla^2_y U(t,\bar{y}(t)) \cdot \dot{\bar{y}}(t), \end{aligned}$$

and (39) follows, for $t \in [0,T]$. ∎

Corollary 4.4 *Under the same assumptions, $t \mapsto \rho(t) \in C^1([0,T])$ and for all $t \in [0,T]$,*

$$\rho(t) = -\Lambda(\bar{y}(t), 1)$$

and

$$\frac{\mathrm{d}}{\mathrm{d}t}\rho(t) = -\nabla_y \Lambda \cdot \nabla_y^2 U \cdot \nabla_y \Lambda - \partial_\eta \Lambda \left(\int_{\mathbb{R}^n} M(z) z dz \right) \cdot \left(\nabla_y^2 U \right)^2 \cdot \nabla_y \Lambda.$$

where $\nabla_y^2 U$ is evaluated in $(t, \bar{y}(t))$, and the derivatives of Λ in $(\bar{y}(t), 1)$.

In particular, if $M(\cdot)$ is even, then $\frac{\mathrm{d}}{\mathrm{d}t}\rho(t) \geq 0$ and $-\overline{\Lambda} \leq \rho \leq -\underline{\Lambda}$.

Proof Follows from $\partial_t U(t, \bar{y}(t)) = \rho(t)$ and $\frac{\mathrm{d}}{\mathrm{d}t}\left[\partial_t u(t, \bar{y}(t))\right] = \frac{\mathrm{d}}{\mathrm{d}t}\rho(t)$. ∎

Remark 4.5 *The limitation of Proposition 4.3 to a short time interval is merely due to three independant phenomena. First, the possible loss of concavity, or apparition of singularities for U, coming from the Hamilton-Jacobi equation (27). Secondly, the possible "jump" of the point where U reaches its maximum, contradicting $\max_{y \in V} U(t,y) = \max_{y \in \mathbb{R}^n} U(t,y)$ in (41). Finally, the possible blow-up in finite time of $\bar{y}(t)$ from the dynamics of the Canonical Equation (39). Regarding the last point, we point out that Λ can sometimes be used as a Lyapunov function. Indeed, we have*

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}t}\left[\Lambda(\bar{y}(t), 1)\right] = \nabla_y \Lambda(\bar{y}(t), 1) \cdot \dot{\bar{y}}(t) \\
&= \nabla_y \Lambda(\bar{y}(t), 1) \cdot \nabla_y^2 U^{-1} \cdot \nabla_y \Lambda(\bar{y}(t), 1) + \partial_\eta \Lambda(\bar{y}(t), 1) \nabla_y \Lambda(\bar{y}(t), 1) \cdot \int_{\mathbb{R}^n} M(z) z dz \\
&\leq \partial_\eta \Lambda(\bar{y}(t), 1) \nabla_y \Lambda(\bar{y}(t), 1) \cdot \int_{\mathbb{R}^n} M(z) z dz.
\end{aligned}$$

In particular, if $M(\cdot)$ is even, then $\frac{\mathrm{d}}{\mathrm{d}t}\left[\Lambda(\bar{y}(t), 1)\right] \leq 0$. Thus, if $\bar{y}^0$ belongs to a "well" of Λ, then $\bar{y}(t)$ remains "trapped", which prevents from an blow-up in finite time. It also implies, at least formally, that $\bar{y}(t)$ converges to a local minimum of $\Lambda(\cdot, 1)$ when $t \to +\infty$.

5 Construction, estimates, and asymptotics of U_ε - Proof of Theorem 1.1

5.1 The eigenproblem - proof of Proposition 2.1

An immediate calculation on (6) gives the explicit solution (22) for Q in terms of Λ. Multiplying by $b(x,y)$ and integrating in x, we obtain formula (20). Next, from (16) we have, for $y \in \mathbb{R}^n$, $\forall \eta \in (\underline{\eta}(y), \overline{\eta}(y))$

$$F(y, \underline{\Lambda}) \leq \frac{1}{\eta} \leq F(y, \overline{\Lambda}).$$

As $\partial_\lambda F > 0$, we conclude the existence and uniqueness of $\Lambda(y,\eta)$ as the unique solution of (20). Now, using (22), we obtain existence and uniqueness of Q.

Finally, the bounds (21) follow from $\partial_\lambda F > 0$ and

$$F(y, \underline{\Lambda}) \leq F(y, \Lambda(y,\eta)) \leq F(y, \overline{\Lambda}).$$

For later purpose, we also neeed the following stronger version of (24).

Lemma 5.1 *There exists $\delta > 0$ such that, for all $y \in \mathbb{R}^n$, $\eta \in (\underline{\eta}(y), \overline{\eta}(y))$, we have*

$$\partial_\eta^2 \Lambda(y,\eta) + \frac{\partial_\eta \Lambda(y,\eta)}{\eta} \leq -\delta.$$

PROOF From the proof of (24), we have

$$\partial_\eta^2\Lambda + \frac{\partial_\eta\Lambda}{\eta} = \frac{1}{(\partial_\lambda F)^3\eta^3}\left((\partial_\lambda F)^2 - F\partial_\lambda^2 F\right).$$

Since $\frac{1}{(\partial_\lambda F)^3\eta^3} \geq \frac{1}{L^3\overline{\eta}^3}$, our goal is to show that

$$(\partial_\lambda F)^2 - F\partial_\lambda^2 F \leq -\delta,$$

for all $y \in \mathbb{R}^n$, $\eta \in (\underline{\eta}(y), \overline{\eta}(y))$.

We set, for all $x \geq 0$, $y \in \mathbb{R}^n$, $\lambda \in \mathbb{R}$,

$$f(x,y,\lambda) := \frac{b(x,y)}{A(x,y)}\exp\left(\int_0^x \frac{\lambda - d(x',y)}{A(x',y)}dx'\right).$$

According to (11), we have $F(y,\lambda) = \int_{\mathbb{R}_+} f(x,y,\lambda)dx$. We also define the probability measure $\tilde{f}(x,y,\lambda) := \frac{f(x,y,\lambda)}{F(y,\lambda)}$. Now, setting $\mathcal{A}(x,y) := \int_0^x \frac{1}{A(x',y)}dx'$, we have

$$\partial_\lambda F(y,\lambda) = \int_{\mathbb{R}_+} \mathcal{A}(x,y)f(x,y,\lambda)dx, \quad \partial_\lambda^2 F(y,\lambda) = \int_{\mathbb{R}_+} \mathcal{A}(x,y)^2 f(x,y,\lambda)dx.$$

It gives,

$$(\partial_\lambda F)^2 - F\partial_\lambda^2 F = -F^2\int_{\mathbb{R}_+}\left(\mathcal{A}(x,y) - \int_{\mathbb{R}_+}\mathcal{A}\tilde{f}dx\right)^2\tilde{f}dx$$

(the function are evaluated on y and $\lambda = \Lambda(y,\eta)$). We have $F^2 \leq \frac{1}{\underline{\eta}^2}$, and assumption (9) implies that the above term is negative uniformly in $y \in \mathbb{R}^n$, $\eta \in (\underline{\eta}(y), \overline{\eta}(y))$. ∎

5.2 Construction of U_ε - proof of Proposition 2.2

We present a proof of existence based on a regularization argument. For the constrained Hamilton-Jacobi equation, a fixed point method has also been proposed in [28]. Our proof is divided into three parts. First, we construct U_ε on a truncated problem. Then, we prove a uniform a priori estimate on $\partial_t U_\varepsilon$, which allows finally to remove the truncation.

Extending Λ. Proposition 2.1 defines $\Lambda(y,\eta)$ only for $y \in \mathbb{R}^n$ and $\eta \in \big(\underline{\eta}(y), \overline{\eta}(y)\big)$. We first need to artificially extend $\Lambda(y,\eta)$ for $\eta \in (0,+\infty)$. For $y \in \mathbb{R}^n$, we set

$$\tilde{\Lambda}(y,\eta) = \begin{cases} \underline{\Lambda} - \underline{B_y}(\eta) & \text{if } \eta < \underline{\eta}(y), \\ \Lambda(y,\eta) & \text{if } \underline{\eta}(y) \leq \eta \leq \overline{\eta}(y), \\ \overline{\Lambda} + \overline{B_y}(\eta) & \text{if } \eta > \overline{\eta}(y), \end{cases}$$

where $\underline{B_y}$ and $\overline{B_y}$ are chosen to be positive, increasing, bounded by 1, and such that $\tilde{\Lambda}$ is smooth. Note that the extension of Λ is completely arbitrary, but we will show, a posteriori, that $\eta_\varepsilon \in \big(\underline{\eta}(y), \overline{\eta}(y)\big)$. We consider the following problem

$$\begin{cases} \partial_t\tilde{U}_\varepsilon(t,y) = -\tilde{\Lambda}\left(y, \displaystyle\int_{\mathbb{R}^n} M(z)e^{\frac{\tilde{U}_\varepsilon(t,y+\varepsilon z)-\tilde{U}_\varepsilon(t,y)}{\varepsilon}}dz\right), & \forall t \geq 0,\ \forall y \in \mathbb{R}^n, \\ \tilde{U}_\varepsilon(0,y) = u_\varepsilon^0(y), & \forall y \in \mathbb{R}^n. \end{cases} \tag{45}$$

Solution for the truncated problem. For a fixed $R > 0$, we consider a truncation function $\phi_R : \mathbb{R} \to \mathbb{R}$ which is smooth, increasing and satisfies the following conditions:

- $\phi_R(r) = r$ for $r \in [-\frac{R}{2}, \frac{R}{2}]$,
- $\phi_R(r) = R$ for $r \geq 2R$,
- $\phi_R(r) = -R$ for $r \leq -2R$,
- $\phi_R' \geq 0$ is uniformly bounded.

For $\varepsilon > 0$, we consider the Cauchy problem

$$\begin{cases} \partial_t \tilde{U}_\varepsilon^R(t,y) = \phi_R\left(-\tilde{\Lambda}\left(y, \displaystyle\int_{\mathbb{R}^n} M(z) e^{\frac{\tilde{U}_\varepsilon^R(t,y+\varepsilon z) - \tilde{U}_\varepsilon^R(t,y)}{\varepsilon}} dz\right)\right), \\ \tilde{U}_\varepsilon^R(0,\cdot) = U_\varepsilon^0. \end{cases} \tag{46}$$

for which the classical Cauchy-Lipschitz theorem provides existence and uniqueness of a solution $\tilde{U}_\varepsilon^R$, defined globally in time.

Estimate on the time derivative.

Lemma 5.2 *We have, with $\partial_t U_\varepsilon^0 := -\Lambda\left(y, \eta_\varepsilon^0(y)\right)$ and η_ε^0 is defined in* (15)*,*

$$\inf_{y\in\mathbb{R}^n} \partial_t U_\varepsilon^0(y) \leq \partial_t \tilde{U}_\varepsilon^R(t,y) \leq \sup_{y\in\mathbb{R}^n} \partial_t U_\varepsilon^0(y), \quad \forall \varepsilon > 0,\ t > 0,\ y \in \mathbb{R}^n.$$

The full proof of this statement, which is technical, can be found in [38] (proof of Proposition 4.7, Appendix D). We give the formal idea of the method.

Let us fix $\varepsilon > 0$, $R > 0$ and set $V(t,y) := \partial_t \tilde{U}_\varepsilon^R(t,y)$. Differentiating (43) with respect to t, we obtain

$$\partial_t V(t,y) = \int_{\mathbb{R}^n} K(t,y,z)\left(\frac{V(t,y+\varepsilon z) - V(t,y)}{\varepsilon}\right) dz,$$

where $K(t,y,z) := -\phi_R'\ \partial_\eta \tilde{\Lambda}\ M(z) e^{\frac{\tilde{U}_\varepsilon^R(t,y+\varepsilon z) - \tilde{U}_\varepsilon^R(t,y)}{\varepsilon}}$. Since $\partial_\eta \Lambda < 0$, we have $K \geq 0$. Then, if for some $t > 0$, $V(t,\cdot)$ reaches its maximum at $\bar{y} \in \mathbb{R}^n$, we obtain the inequality

$$\partial_t V(t,\bar{y}) = \int_{\mathbb{R}^n} K(t,\bar{y},z)\left(\frac{V(t,\bar{y}+\varepsilon z) - V(t,\bar{y})}{\varepsilon}\right) dz \leq 0.$$

Formally, it shows that the maximum value of V is decreasing with time, that is,

$$\sup_y V(t,y) \leq \sup_y V(0,y) = \sup_y \partial_t U_\varepsilon^0.$$

With the same method we show $\inf_y V \geq \inf_y \partial_t U_\varepsilon^0$, which conclude the proof of Lemma 5.2.

Hereafter, from assumption (16) and $\partial_\lambda F > 0$, we have

$$-\overline{\Lambda} \leq \partial_t U_\varepsilon^0(y) \leq -\underline{\Lambda}.$$

Using Lemma 5.2, we infer

$$-\overline{\Lambda} \leq \phi_R\left(-\tilde{\Lambda}\left(y, \int_{\mathbb{R}^n} M(z) e^{\frac{\tilde{U}_\varepsilon^R(t,y+\varepsilon z) - \tilde{U}_\varepsilon^R(t,y)}{\epsilon}}\right)\right) \leq -\underline{\Lambda}. \tag{47}$$

Removing the truncation. From (44) and the choice of ϕ_R, for R large enough, we have

$$-\phi_R\left(\tilde{\Lambda}\left(y,\int_{\mathbb{R}^n} M(z)e^{\frac{\tilde{U}_\varepsilon^R(t,y+\varepsilon z)-\tilde{U}_\varepsilon^R(t,y)}{\epsilon}}\right)\right)=\tilde{\Lambda}\left(y,\int_{\mathbb{R}^n} M(z)e^{\frac{\tilde{U}_\varepsilon^R(t,y+\varepsilon z)-\tilde{U}_\varepsilon^R(t,y)}{\epsilon}}\right).$$

Besides, since $\partial_\eta\tilde{\Lambda}<0$, we have

$$\underline{\eta}(y)\le\int_{\mathbb{R}^n} M(z)e^{\frac{\tilde{U}_\varepsilon^R(t,y+\varepsilon z)-\tilde{U}_\varepsilon^R(t,y)}{\epsilon}}\le\overline{\eta}(y),$$

for all R large enough, $\varepsilon>0$, $t\ge 0$, $y\in\mathbb{R}^n$. Thus, from the definition of $\tilde{\Lambda}$ in (42), we have

$$-\tilde{\Lambda}\left(y,\int_{\mathbb{R}^n} M(z)e^{\frac{\tilde{U}_\varepsilon^R(t,y+\varepsilon z)-\tilde{U}_\varepsilon^R(t,y)}{\varepsilon}}dz\right)=-\Lambda\left(y,\int_{\mathbb{R}^n} M(z)e^{\frac{\tilde{U}_\varepsilon^R(t,y+\varepsilon z)-\tilde{U}_\varepsilon^R(t,y)}{\varepsilon}}dz\right),$$

that is, $\tilde{U}_\varepsilon^R$ is a solution of (7). The proof is thereby achieved.

5.3 A priori Lipschitz estimate - proof of Proposition 2.4

We follow the same idea as for Lemma 5.2. However, there are some technical difficulties. First, we have to deal with a "source term" $\nabla_y\Lambda\left(y,\eta_\varepsilon(t,y)\right)$ which is bounded by a constant $\frac{L}{l\underline{\eta}^2}$, using (17), (18), (23) and (26). In addition, we first need to prove the estimate on a truncated function, then to remove the truncation.

We fix $i\in\{1,\dots,n\}$, $T>0$, and we set

$$W_\varepsilon(t,y):=\partial_{y_i}U_\varepsilon(t,y).$$

Differentiating (7), we obtain

$$\begin{aligned}&\partial_t W_\varepsilon(t,y)\\&=-\partial_{y_i}\Lambda(y,\eta_\varepsilon)-\partial_\eta\Lambda(y,\eta_\varepsilon)\left(\int M(z)e^{\frac{U_\varepsilon(t,y+\varepsilon z)-U_\varepsilon(t,y)}{\varepsilon}}\left[\frac{W_\varepsilon(t,y+\varepsilon z)-W_\varepsilon(t,y)}{\varepsilon}\right]dz\right)\\&:=\mathcal{F}(t,y,W_\varepsilon(t,\cdot)).\end{aligned}$$

We formally define a truncated problem, for $R>0$, and its solution W_ε^R satisfying

$$W_\varepsilon^R(t,y)=\phi_R\left(\partial_{y_i}U_\varepsilon^0(y)+\int_0^t\mathcal{F}(s,y,W_\varepsilon^R(s,\cdot))ds\right),\tag{48}$$

where $\mathcal{F}$ is defined above and ϕ_R is a truncation function as in (43). We can prove existence and uniqueness of a global solution of (45) by a direct application of the Cauchy-Lipschitz theorem.

We set $\bar{W}_\varepsilon^R:=W_\varepsilon^R-Ct$ with $C:=\frac{L}{l\underline{\eta}^2}$. Our goal is to show

$$\forall t\in[0,T],\forall y\in\mathbb{R}^n,\quad\bar{W}_\varepsilon^R(t,y)\le\sup\partial_{y_i}U_\varepsilon^0.\tag{49}$$

By contradiction, assume (46) does not hold, i.e., there exist $y_0\in\mathbb{R}^n,t_0\in[0,T]$ such that

$$\bar{W}_\varepsilon^R(t_0,y_0)-\sup\partial_{y_i}U_\varepsilon^0>0.\tag{50}$$

For $\beta>0$, $\alpha>0$ small enough, $t\in[0,T],y\in\mathbb{R}^n$ we introduce

$$\varphi_{\alpha,\beta}(t,y):=\bar{W}_\varepsilon^R(t,y)-\alpha t-\beta|y-y_0|.$$

As $\bar{W}_\varepsilon^R$ is bounded, $\varphi_{\alpha,\beta}$ reaches its maximum on $[0,T]\times\mathbb{R}^n$ at a point $(\bar{t},\bar{y})$. We have

$$\forall z\in\mathbb{R}^n,\ \varphi_{\alpha,\beta}(\bar{t},\bar{y}+\varepsilon z)\le\varphi_{\alpha,\beta}(\bar{t},\bar{y}).$$

Then, we obtain the inequality

$$\forall z \in \mathbb{R}^n,\ \frac{\bar{W}_\varepsilon^R(\bar{t},\bar{y}+\varepsilon z)-\bar{W}_\varepsilon^R(\bar{t},\bar{y})}{\varepsilon} \le \beta \frac{|\bar{y}+z-y_0|-|\bar{y}-y_0|}{\varepsilon} \le \beta|z|.$$

We choose α small enough so that $\varphi_{\alpha,\beta}(t_0,y_0) > \varphi_{\alpha,\beta}(0,y_0) = \partial_{y_i} U_\varepsilon^0(y_0)$, which is possible thanks to (47). It implies $\bar{t} > 0$. Hence $\partial_t \varphi_{\alpha,\beta}(\bar{t},\bar{y}) \ge 0$, i.e. $\partial_t \bar{W}^R(\bar{t},\bar{y}) \ge \alpha$ (if $\bar{t} = T$, then ∂_t stands for the left derivative). Differentiating (45) at $(\bar{t},\bar{y})$, we have

$$\begin{aligned}
\alpha &\le \partial_t \bar{W}_\varepsilon^R(\bar{t},\bar{y}) \\
&\le -\sup \phi_R' \times \partial_{y_i}\Lambda(y,\eta_\varepsilon(t,y)) - C \\
&\quad + \sup \phi_R' \times (-\partial_\eta \Lambda(y,\eta)) \int M(z) e^{\frac{U_\varepsilon(\bar{t},\bar{y}+\varepsilon z)-U(\bar{t},\bar{y})}{\varepsilon}} \left[\frac{W_\varepsilon^R(\bar{t},\bar{y}+\varepsilon z)-W_\varepsilon^R(\bar{t},\bar{y})}{\varepsilon}\right] dz.
\end{aligned}$$

Now, from $|\partial_{y_i}\Lambda(y,\eta_\varepsilon(t,y))| \le \frac{L}{l\underline{\eta}^2} = C$ and $0 \le -\partial_\eta \Lambda(y,\eta) \le \frac{L}{\underline{\eta}^2}$, we have

$$\begin{aligned}
\alpha &\le \frac{L}{\underline{\eta}^2}\left(\int M(z) e^{\frac{U_\varepsilon(\bar{t},\bar{y}+\varepsilon z)-U(\bar{t},\bar{y})}{\varepsilon}}|z|\mathrm{d}z\right) \times \beta \\
&\le \frac{L}{\underline{\eta}^2}\left(\int M(z) e^{\frac{-2\Lambda T}{\varepsilon}+k_0|z|}|z|\mathrm{d}z\right) \times \beta.
\end{aligned}$$

Then, passing to the limit $\beta \to 0$ we obtain $\alpha \le 0$: contradiction. Thus, we have

$$\bar{W}_\varepsilon^R \le \sup|\partial_{y_i} U_\varepsilon^0| = k^0.$$

We proceed similarily to obtain the reverse inequality $\bar{W}_\varepsilon^R \ge k^0$. We have, for all $R > 0,\ \varepsilon > 0,\ t \in [0,T],\ y \in \mathbb{R}^n$

$$|W_\varepsilon^R(t,y)| \le k^0 + Ct.$$

Finally, the bound on W_ε^R is uniform in R so we can remove the truncation, as detailed in section 5.2. Thus, $W_\varepsilon^R = W_\varepsilon$ for R large enough and

$$|\partial_{y_i} U_\varepsilon(t,y)| \le k^0 + Ct.$$

5.4 Semi-convexity - proof of Proposition 2.5

For convex Hamiltonian, the semi-convexity of the solution is a classical matter, [3, 15]. Here, we have to deal with a nonlocal operator which features a difference rather than a gradient

Semi-convexity in t. For shorter formulas, we need some notations

$$V_\varepsilon := \partial_t^2 U_\varepsilon, \qquad J_\varepsilon(t,y,z) := M(z) e^{\frac{U_\varepsilon(t,y+\varepsilon z)-U_\varepsilon(t,y)}{\varepsilon}}.$$

We begin with two results that are used later and express some properties usually connected to the convexity of the Hamiltonian. Firstly we observe that

Lemma 5.3 *For all $\varepsilon > 0$, $t > 0$, $y \in \mathbb{R}^n$, we have*

$$(\partial_t \eta_\varepsilon)^2 \le \eta_\varepsilon \int_{\mathbb{R}^n} \left(\frac{\partial_t U_\varepsilon(t,y+\varepsilon z)-\partial_t U_\varepsilon(t,y)}{\varepsilon}\right)^2 J_\varepsilon dz.$$

Proof Use Jensen's inequality and the definition of η_ε in (25). ∎

Next, we prove that

$$\partial_t V_\varepsilon \ge -\partial_\eta \Lambda(y,\eta_\varepsilon) \int_{\mathbb{R}^n} \frac{V_\varepsilon(t,y+\varepsilon z)-V_\varepsilon(t,y)}{\varepsilon} J_\varepsilon dz. \tag{51}$$

PROOF Differentiating (7) twice, we find

$$\begin{aligned}\partial_t V_\varepsilon =& -\partial_\eta\Lambda \int_{\mathbb{R}^n} \frac{V_\varepsilon(t,y+\varepsilon z)-V_\varepsilon(t,y)}{\varepsilon} J_\varepsilon dz \\ & -\partial_\eta^2\Lambda\left(\partial_t\eta_\varepsilon\right)^2 - \partial_\eta\Lambda\int_{\mathbb{R}^n}\left(\frac{\partial_t U_\varepsilon(t,y+\varepsilon z)-\partial_t U_\varepsilon(t,y)}{\varepsilon}\right)^2 J_\varepsilon dz.\end{aligned} \tag{52}$$

Next, combining Lemma 5.3 with (24), we find

$$\partial_\eta^2\Lambda(y,\eta_\varepsilon)\left(\partial_t\eta_\varepsilon\right)^2 + \partial_\eta\Lambda(y,\eta_\varepsilon)\int_{\mathbb{R}^n}\left(\frac{\partial_t U_\varepsilon(t,y+\varepsilon z)-\partial_t U_\varepsilon(t,y)}{\varepsilon}\right)^2 J_\varepsilon dz \le 0,$$

Using the above inequality and (49), we find (48). ∎

From inequality (48), we deduce

Lemma 5.4 *V_ε is uniformly bounded from below and more precisely, with $V_\varepsilon^0(y) := V_\varepsilon(t=0,y)$, we have*

$$V_\varepsilon(t,y) \ge \inf_{y\in\mathbb{R}^n} V_\varepsilon^0(y) > -\infty, \quad \forall\varepsilon>0,\ \forall t>0,\ \forall y\in\mathbb{R}^n. \tag{53}$$

The proof follows closely the method of section 5.3. The formal idea is the following. If, for some $t>0$, $V_\varepsilon(t,\cdot)$ reaches its minimum at $\bar y\in\mathbb{R}^n$, from (48) we obtain $\partial_t V_\varepsilon(t,\bar y)\ge 0$. Formally, it shows that the minimum value of V_ε is increasing with time, that is, $\inf_y V_\varepsilon(t,y)\ge\inf_y V_\varepsilon(0,y)$. Then, we conclude with the fact that $\inf_y V_\varepsilon(0,y)$ is bounded, uniformly in $\varepsilon>0$.

PROOF Differentiating (55) in t, we obtain

$$V_\varepsilon(t,y) = -\partial_\eta\Lambda(y,\eta_\varepsilon)\int_{\mathbb{R}^n}\frac{\Lambda_\varepsilon(t,y+\varepsilon z)-\Lambda_\varepsilon(t,y)}{\varepsilon}J_\varepsilon dz. \tag{54}$$

In particular, our assumptions imply $\inf_{y\in\mathbb{R}^n} V_\varepsilon^0 > -\infty$ uniformly in $\varepsilon>0$, thus (50) implies that V_ε is bounded from below, uniformly in ε.

We prove (50) by contradiction. We assume that there exists $(T,y_0)\in(0,+\infty)\times\mathbb{R}^n$ such that

$$V_\varepsilon(T,y_0) - \inf_{y\in\mathbb{R}^n} V_\varepsilon^0(y) < 0. \tag{55}$$

For $\beta>0$, $\alpha>0$ small and for $t\in[0,T], y\in\mathbb{R}^n$, we also introduce

$$\varphi_{\alpha,\beta}(t,y) := V_\varepsilon(t,y) + \alpha t + \beta|y-y_0|.$$

From (51) and for a fixed $\varepsilon>0$, we have $V_\varepsilon(t,y)$ is bounded from below uniformly in $t\in[0,T]$, $y\in\mathbb{R}^n$. Therefore, $\varphi_{\alpha,\beta}$ goes to $+\infty$ as $|y|\to+\infty$ and reaches its minimum on $[0,T]\times\mathbb{R}^n$ at a point $(\bar t,\bar y)$. We have

$$\varphi_{\alpha,\beta}(\bar t,\bar y+\varepsilon z)\ge\varphi_{\alpha,\beta}(\bar t,\bar y),\quad\forall z\in\mathbb{R}^n,$$

thus

$$\frac{V_\varepsilon(\bar t,\bar y+\varepsilon z)-V_\varepsilon(\bar t,\bar y)}{\varepsilon}\ge\beta\frac{|\bar y-y_0|-|\bar y-y_0+\varepsilon z|}{\varepsilon}\ge-\beta|z|,\quad\forall z\in\mathbb{R}^n. \tag{56}$$

We choose α small enough to ensure $\varphi_{\alpha,\beta}(T,y_0)<\varphi_{\alpha,\beta}(0,y_0)$, which is possible thanks to assumption (52). It implies $\bar t>0$. Hence $\partial_t\varphi_{\alpha,\beta}(\bar t,\bar y)\le 0$, that is $\partial_t V_\varepsilon(\bar t,\bar y)\le-\alpha$ (if $\bar t=T$ then $\partial_t V_\varepsilon^R(\bar t,\bar y)$ stands for the left-derivative). From (48) at $(\bar t,\bar y)$, using (53),

$$\begin{aligned}-\alpha\ge\partial_t V_\varepsilon(\bar t,\bar y) &= -\partial_\eta\Lambda\int_{\mathbb{R}^n} M(z)e^{\frac{U_\varepsilon(t,y+\varepsilon z)-U_\varepsilon(t,y)}{\varepsilon}}\frac{V_\varepsilon(t,y+\varepsilon z)-V_\varepsilon(t,y)}{\varepsilon}dz\\ &\ge\beta\inf\left[-\partial_\eta\Lambda\right]\int M(z)e^{\frac{U_\varepsilon(\bar t,\bar y+\varepsilon z)-U_\varepsilon(\bar t,\bar y)}{\varepsilon}}|z|dz\\ &\ge\beta\frac{1}{L\overline{\eta}}\int M(z)e^{\left(k_0+\frac{L}{l\underline{\eta}^2}T\right)|z|}|z|dz\end{aligned}$$

where, in the last step, we used Proposition 2.4 and $-\partial_\eta\Lambda\ge\frac{1}{L\overline{\eta}}$. As β goes to 0, we obtain $\alpha\le 0$, which is absurd. The proof is thereby achieved. ∎

Semi-convexity in y. Let us show how the method can be adapted to prove that $\partial_i^2 U_\varepsilon$ (the second derivative w.r.t. y_i) is bounded from below. We set $W_\varepsilon := \partial_i^2 U_\varepsilon$. Differentiating (7) twice, we find

$$\begin{aligned}\partial_t W_\varepsilon =& -\partial_\eta \Lambda \int_{\mathbb{R}^n} \frac{W_\varepsilon(t,y+\varepsilon z)-W_\varepsilon(t,y)}{\varepsilon} J_\varepsilon dz \\ & -\partial_\eta^2 \Lambda \left(\partial_i \eta_\varepsilon\right)^2 - \partial_\eta \Lambda \int_{\mathbb{R}^n} \left(\frac{\partial_i U_\varepsilon(t,y+\varepsilon z)-\partial_i U_\varepsilon(t,y)}{\varepsilon}\right)^2 J_\varepsilon dz \\ & -\partial_i^2 \Lambda - 2\partial_{i,\eta}^2 \Lambda \partial_i \eta_\varepsilon.\end{aligned} \tag{57}$$

In contrast with the equation (49) on $\partial_t^2 U_\varepsilon$, we need to deal with a source term and a linear term in the last line.

For any constant $K>0$, Young's inequality implies

$$-2\partial_{i,\eta}^2 \Lambda \partial_i \eta_\varepsilon \geq -K^2 |\partial_{i,\eta}^2 \Lambda|^2 - \frac{1}{K^2}(\partial_i \eta_\varepsilon)^2.$$

Applying this inequality and Lemma 5.3 (replacing ∂_t by ∂_i) in (54), we obtain

$$\begin{aligned}\partial_t W_\varepsilon \geq & -\partial_\eta \Lambda \int_{\mathbb{R}^n} \frac{W_\varepsilon(t,y+\varepsilon z)-W_\varepsilon(t,y)}{\varepsilon} J_\varepsilon dz \\ & -\left(\partial_\eta^2 \Lambda + \frac{\partial_\eta \Lambda}{\eta_\varepsilon} + \frac{1}{K^2}\right)(\partial_i \eta_\varepsilon)^2 \\ & -\partial_i^2 \Lambda - K^2 |\partial_{i,\eta}^2 \Lambda|^2.\end{aligned}$$

Using lemma Lemma 5.1, and choosing $K \geq \frac{1}{\sqrt{\delta}}$, we have

$$\begin{aligned}\partial_t W_\varepsilon \geq & -\partial_\eta \Lambda \int_{\mathbb{R}^n} \frac{W_\varepsilon(t,y+\varepsilon z)-W_\varepsilon(t,y)}{\varepsilon} J_\varepsilon dz \\ & -\partial_i^2 \Lambda - K^2 |\partial_{i,\eta}^2 \Lambda|^2.\end{aligned}$$

From assumptions (18)-(19), the source term $-\partial_i^2 \Lambda - K^2 |\partial_{i,\eta}^2 \Lambda|^2$ is bounded from below by some constant $-K' < 0$. We end up with

$$\partial_t W_\varepsilon \geq -\partial_\eta \Lambda \int_{\mathbb{R}^n} \frac{W_\varepsilon(t,y+\varepsilon z)-W_\varepsilon(t,y)}{\varepsilon} J_\varepsilon dz - K'.$$

Then, applying the same method as in the proof of Lemma 5.4 (see also the proof of Proposition 2.2), we show

$$W_\varepsilon(t,y) \geq \inf_{y\in\mathbb{R}^n} W_\varepsilon(t=0,y) - K't, \quad \forall \varepsilon>0,\ t\geq 0,\ y\in\mathbb{R}^n.$$

Finally, we conclude the lower bound since $W_\varepsilon(t=0,y) \geq -C$ (from assumption (12)), we have

$$W_\varepsilon(t,y) \geq -C - K't, \quad \forall \varepsilon>0,\ t\geq 0,\ y\in\mathbb{R}^n.$$

Other directional derivatives and conclusion. Lower bounds on other second order derivatives in directions $\nu \in \mathbb{S} := \{(t,y)\in\mathbb{R}^{n+1} : t^2+|y|^2=1\}$ can be obtained by a slight adaptation of the previous steps show. We deduce that U_ε is semi-convex and that ∇U_ε is uniformly in BV_{loc} (see Proposition 1.1.3 and Theorem 2.3.1 in [15]). We obtain that U_ε is uniformly bounded in $W^{2,1}_{loc}$.

5.5 Proof of Corollary 2.6

We recall the definition of η_ε in (25), and note that differentiating (7), we obtain

$$\partial_t^2 U_\varepsilon = -\partial_t \eta_\varepsilon(t,y)\partial_\eta \Lambda(t,\eta_\varepsilon(t,y)), \qquad \partial_{t,x_i}^2 U_\varepsilon = -\partial_{x_i}\eta_\varepsilon(t,y)\partial_\eta\Lambda(t,\eta_\varepsilon(t,y)). \tag{58}$$

Using that U_ε is uniformly bounded in $W^{2,1}$ (Proposition 2.5) and that $-\partial_\eta\Lambda$ is positively bounded (23), we deduce that η_ε is bounded in $W^{1,1}$, which proves the first part of Corollary 2.6.

Let us now fix $T > 0$ and prove the second part. We recall that U_ε is semi-concave (Proposition 2.5). Thus, there exists a constant $K > 0$ such that $\partial_t \eta_\varepsilon(t,y) \geq -K$ for all $\varepsilon > 0$, $(t,y) \in [0,T] \times \mathbb{R}^n$. Denoting $\overline{\eta}(t) := \sup_{y \in \mathbb{R}^n} \eta(t,y)$, we deduce $\partial_t \overline{\eta}_\varepsilon(t) \geq -K$ (indeed, the mapping $\theta_y : t \mapsto \eta_\varepsilon(t,y) + Kt$ is nondecreasing, and so is $\sup_{y \in \mathbb{R}^n} \theta_y$). The last inequality should be understood in the sense of the distributions.

We deduce

$$\int_0^T |\partial_t \overline{\eta}_\varepsilon(t)| dt = \int_0^T \partial_t \overline{\eta}_\varepsilon(t) dt + 2 \int_0^T (\partial_t \overline{\eta}_\varepsilon(t))^- dt \leq \overline{\eta}_\varepsilon(t) + 2Kt \leq \overline{\eta} + 2Kt,$$

where $^-$ denotes the negative part, and the last inequality comes from (26). The proof is complete.

5.6 Asymptotics of U_ε - proof of Proposition 2.7

Extraction of a subsequence From the a priori estimate of Proposition 2.2 and Ascoli's theorem, we know that U_ε converges locally uniformly to some U, up to extraction of a subsequence. Incidentally, this convergence also occurs in $W^{1,1}$, from the $W^{2,1}$ estimate in Proposition 2.4 and a classical compact embedding. In addition, we know from Proposition 2.2 and Proposition 2.4 that U_ε is locally Lipschitz continuous, uniformly in $\varepsilon > 0$: for all $t \geq t' \geq 0$ and $y, y' \in \mathbb{R}^n$

$$|U(t,y) - U(t',y')| \leq -\underline{\Lambda}(t-t') + \left(k^0 + \frac{L}{l\underline{\eta}^2} t\right) |y - y'|.$$

Thus, the convergence occurs in $W^{1,r}_{loc}$, $1 \leq r < \infty$. Also notice that Proposition 2.5 implies that U is semi-convex, uniformly in $y \in \mathbb{R}^n$, locally uniformly in t.

Viscosity solution We are going to show that U is a viscosity solution of (27), i.e., U satisfies

$$\partial_t U = H(y, \nabla_y U), \quad U(0,y) = U^0(y), \tag{59}$$

with

$$H(y,p) := -\Lambda\left(y, \int_{\mathbb{R}^n} M(z) e^{p \cdot z} dz\right).$$

The proof is adapted from classical stability results for viscosity solutions of Hamilton-Jacobi equations (see [3]). However, this case is not completely standard because of the nonlocal term $\int_{\mathbb{R}^n} M(z) e^{\frac{U_\varepsilon(t,y+\varepsilon z) - U_\varepsilon(t,y)}{\varepsilon}} dz$.

Lemma 5.5 *The function U is a viscosity solution of* (56) *in $(0,\infty) \times \mathbb{R}^n$. Also, for all $T > 0$, the viscosity inequalities stand for $t \in (0,T]$.*

Proof We are going to prove that U is a subsolution of (56). Let us consider a test function φ and a point (t_0, y_0) such that $U - \varphi$ reaches a global maximum at (t_0, y_0). From classical results, there exists $(t_\varepsilon, y_\varepsilon)$ such that

$$\begin{cases} (t_\varepsilon, y_\varepsilon) \underset{\varepsilon \to 0}{\longrightarrow} (t_0, y_0), \\ \max\limits_{t,y} U_\varepsilon - \varphi = (U_\varepsilon - \varphi)(t_\varepsilon, y_\varepsilon). \end{cases}$$

For all $z \in \mathbb{R}^n$, $\varphi(t_\varepsilon, y_\varepsilon + \varepsilon z) - U_\varepsilon(t_\varepsilon, y_\varepsilon + \varepsilon z) \geq \varphi(t_\varepsilon, y_\varepsilon) - U_\varepsilon(t_\varepsilon, y_\varepsilon)$, thus we have

$$\frac{\varphi(t_\varepsilon, y_\varepsilon + \varepsilon z) - \varphi(t_\varepsilon, y_\varepsilon)}{\varepsilon} \geq \frac{U_\varepsilon(t_\varepsilon, y_\varepsilon + \varepsilon z) - U_\varepsilon(t_\varepsilon, y_\varepsilon)}{\varepsilon}.$$

Since $\partial_\eta \Lambda < 0$, equation (56) gives

$$\partial_t \varphi(t_\varepsilon, y_\varepsilon) = -\Lambda\left(y_\varepsilon, \int_{\mathbb{R}^n} M(z) e^{\frac{U_\varepsilon(t_\varepsilon, y_\varepsilon + \varepsilon z) - U_\varepsilon(t_\varepsilon, y_\varepsilon)}{\varepsilon}} dz\right) \leq -\Lambda\left(y_\varepsilon, \int_{\mathbb{R}^n} M(z) e^{\frac{\varphi(t_\varepsilon, y_\varepsilon + \varepsilon z) - \varphi(t_\varepsilon, y_\varepsilon)}{\varepsilon}} dz\right).$$

As ε goes to 0,

$$\partial_t \varphi(t_0, y_0) \leq -\Lambda\left(y_0, \int_{\mathbb{R}^n} M(z) e^{\nabla_y \varphi(t_0, y_0) \cdot z}\right) = H(y_0, \nabla_y \varphi(t_0, y_0)),$$

then U is a viscosity subsolution of (56). With the same method, we prove that U is also a viscosity supersolution. It completes the first part of the proof. The second part of the statement is a well-known result, and proof can be found in [3]. ∎

Uniqueness We point out that the Hamiltonian H is Lipschitz continuous in the y variable. We introduce a truncated Hamiltonian

$$\tilde{H}(y,p) := \begin{cases} H(y,p) & \text{if } \displaystyle\int_{\mathbb{R}^n} M(z)e^{p\cdot z}dz \in [\underline{\eta},\overline{\eta}], \\ 0 & \text{otherwise.} \end{cases}$$

Since $\underline{\eta} \le \eta_\varepsilon(t,y) \le \overline{\eta}$ (from (26)), we have

$$\partial_t U = \tilde{H}(y, \nabla_y U).$$

For this equation, a classical uniqueness result is in order (see e.g [3, 38]). We deduce that U_ε converges to U for the whole sequence $\varepsilon \to 0$ (and not for an extracted subsequence).

5.7 A posteriori Lipschitz estimate - proof of the global Lipschitz regularity in Proposition 2.7

From Proposition 2.2 and Proposition 2.4, we know that U is Lipschitz, globally in t and locally in y. Our goal is to show that U is globally Lipschitz continuous, i.e., that there exists a constant $C > 0$ such that

$$\forall t \ge 0,\ \forall (y,y') \in (\mathbb{R}^n)^2, \quad U(t,y) - U(t,y') \le C|y-y'|. \tag{60}$$

Let us fix $t \ge 0$ and $(y,y') \in (\mathbb{R}^n)^2$.

With η_ε defined in (25) and the bound $\eta_\varepsilon \le \overline{\eta}$ from (26), we have, for all $\varepsilon > 0$, $z \in \mathbb{R}^n$

$$\int_{\mathbb{R}^n} M(z) e^{\frac{U_\varepsilon(t,y+\varepsilon z)-U(t,y)}{\varepsilon}} dz \le \eta_\varepsilon(t,y) \le \overline{\eta}.$$

From the assumption that $M(\cdot)$ is not degenerate, we deduce that, for some $r_0 > 0$, and for all $z \in \mathbb{R}^n$ such that $|z| = r_0$, then

$$\frac{U_\varepsilon(t,y+\varepsilon z) - U_\varepsilon(t,y)}{\varepsilon} \le C$$

for some constant C (independant of ε, t, y and z). Then, chosing z and ε such that $y - y' = \varepsilon z$, we have $U_\varepsilon(t,y) - U_\varepsilon(t,y') \le C|y-y'|$. As $\varepsilon \to 0$, we prove the goal (57).

References

[1] A. S. Ackleh, B. G. Fitzpatrick, and H. R. Thieme. Rate distributions and survival of the fittest: a formulation on the space of measures. *Discrete Contin. Dyn. Syst. Ser. B*, 5(4):917–928, 2005.

[2] M. Adimy, F. Crauste, and S. Ruan. A mathematical study of the hematopoiesis process with applications to chronic myelogenous leukemia. *SIAM J. Appl. Math.*, 65(4), 2005.

[3] G. Barles. *Solutions de viscosité des équations de Hamilton-Jacobi.* Springer-Verlag Berlin Heidelberg, 1994.

[4] G. Barles, L. C. Evans, and P. E. Souganidis. Wavefront propagation for reaction diffusion systems of PDE. *Duke Math. J.*, 61(3):835–858, 1990.

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

[6] G. Barles and B. Perthame. Concentrations and constrained Hamilton-Jacobi equations arising in adpative dynamics. *Contemporary Mathematics*, 439:57, 2007.

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

[8] G. Barles and P. E. Souganidis. Front propagation for reaction-diffusion equations arising in combustion theory. *Asymptot. Anal.*, 14(3):277–292, 1997.

[9] S. Bianchini and D. Tonon. SBV regularity for Hamilton-Jacobi equations with Hamiltonian depending on (t,x). *SIAM J. Math. Anal*, 44(3):2179–2203, 2012.

[10] E. Bouin, J. Garnier, C. Henderson, and F. Patout. Thin front limit of an integro-differential Fisher-KPP equation with fat-tailed kernels. *ArXiv preprint*, 2018.

[11] J.-E. Busse, P. Gwiazda, and A. Marciniak-Czochra. Mass concentration in a nonlocal model of clonal selection. *Journal of Mathematical Biology*, 73(4):1001–1033, oct 2016.

[12] À. Calsina and J. M. Palmada. Steady states of a selection-mutation model for an age structured population. *Journal of Mathematical Analysis and Applications*, 400(2):386–395, apr 2013.

[13] V. Calvez, P. Gabriel, and Á. Mateos González. Limiting Hamilton-Jacobi equation for the large scale asymptotics of a subdiffusion jump-renewal equation. *ArXiv preprint*, 2016.

[14] V. Calvez and K.-Y. Lam. Uniqueness of the viscosity solution of a constrained Hamilton-Jacobi equation. working paper or preprint, Sept. 2018.

[15] P. Cannarsa and C. Sinestrari. *Semiconcave Functions, Hamilton-Jacobi Equations, and Optimal Control*. Birkhäuser, Basel, 2004.

[16] N. Champagnat, R. Ferrière, and G. Ben Arous. The canonical equation of adaptive dynamics: A mathematical view. *Selection*, 2(1):73–83, 2002.

[17] N. Champagnat and P.-E. Jabin. The evolutionary limit for models of populations interacting competitively via several resources. *Journal of Differential Equations*, 251(1):176–195, jul 2011.

[18] L. Desvillettes, P.-E. Jabin, S. Mischler, and G. Raoul. On selection dynamics for continuous structured populations. *Commun. Math. Sci.*, 6(3):729–747, 2008.

[19] O. Diekmann. A beginner's guide to adaptive dynamics. *Banach Center publications*, 63:47–86, 2004.

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

[21] M. Doumic and P. Gabriel. Eigenelements of a general aggregation-fragmentation model. *Math. Models Methods Appl. Sci.*, 20(5):757–783, 2010.

[22] L. C. Evans. The perturbed test function method for viscosity solutions of nonlinear PDE. *Proc. Roy. Soc. Edinburgh Sect. A*, 111(3-4):359–375, 1989.

[23] L. C. Evans. *Partial differential equations*, volume 19 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 1998.

[24] L. C. Evans and P. E. Souganidis. A PDE approach to geometric optics for certain semilinear parabolic equations. *Indiana Univ. Math. J.*, 38(1):141–172, 1989.

[25] W. H. Fleming and R. W. Rishel. *Deterministic and Stochastic Optimal Control*. Springer, New York, 1975.

[26] M. Gyllenberg and G. F. Webb. A nonlinear structured population model of tumor growth with quiescence. *J. Math. Biol.*, 28(6):671–694, 1990.

[27] P.-E. Jabin and R. S. Schram. Selection-Mutation dynamics with spatial dependence. *arXive Preprint*, 2016.

[28] Y. Kim. Well-posedness for constrained Hamilton-Jacobi equations. *Acta Appl Math.*, 2019.

[29] T. Lorenzi, A. Lorz, and G. Restori. Asymptotic Dynamics in Populations Structured by Sensitivity to Global Warming and Habitat Shrinking. *Acta Appl Math.*, 131(1):49–67, jun 2014.

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

[31] S. Méléard and V. C. Tran. Trait Substitution Sequence process and Canonical Equation for age-structured populations. *Journal of Mathematical Biology*, 58(6):881–921, jun 2009.

[32] J. A. J. Metz and O. Diekmann. *The dynamics of physiologically structured populations*, volume 68 of *Lecture Notes in Biomathematics*. Springer-Verlag, Berlin, 1986.

[33] P. Michel, S. Mischler, and B. Perthame. General relative entropy inequality: an illustration on growth models. *Journal de Mathématiques Pures et Appliquées*, 84(9):1235–1260, 2005.

[34] S. Mirrahimi. Adaptation and migration of a population between patches. *Discrete and Continuous Dynamical System - B*, 18(3):753–768, 2013.

[35] S. Mirrahimi and B. Perthame. Asymptotic analysis of a selection model with space. *Journal de Mathématiques Pures et Appliquées*, 104(6):1108–1118, 2015.

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

[37] S. Mischler, B. Perthame, and L. Ryzhik. Stability in a nonlinear population maturation model. *Math. Models Methods Appl. Sci.*, 12(12):1751–1772, 2002.

[38] S. Nordmann, B. Perthame, and C. Taing. Dynamics of Concentration in a Population Model Structured by Age and a Phenotypical Trait. *Acta Appl Math.*, 155(1):197–225, jun 2018.

[39] B. Perthame. *Transport equations in biology*. Die Deutsche Bibliothek, Birkhäuser Verlag, Basel, frontiers edition, 2007.

[40] B. Perthame and P. E. Souganidis. Rare Mutations Limit of a Steady State Dispersal Evolution Model. *Math. Model. Nat. Phenom.*, 11(4):154–166, 2016.

[41] V. C. Tran. Large population limit and time behaviour of a stochastic particle model describing an age-structured population. *ESAIM: Probability and Statistics*, 12:345–386, apr 2008.