DIFFUSIVE LIMIT TO A SELECTION-MUTATION EQUATION WITH SMALL MUTATION FORMULATED ON THE SPACE OF MEASURES

AZMY S. ACKLEH*
Department of Mathematics
University of Louisiana at Lafayette
Lafayette, Louisiana 70504, USA

NICOLAS SAINTIER
Departamento de Matemática
Facultad de Ciencias Exactas y Naturales
Universidad de Buenos Aires
(1428) Pabellón I - Ciudad Universitaria
Buenos Aires - Argentina

(Communicated by Peter Hinow)

Abstract. In this paper we consider a selection-mutation model with an advection term formulated on the space of finite signed measures on $\mathbb{R}^d$. The selection-mutation kernel is described by a family of measures which allows the study of continuous and discrete kernels under the same setting. We rescale the selection-mutation kernel to obtain a diffusively rescaled selection-mutation model. We prove that if the rescaled selection-mutation kernel converges to a pure selection kernel then the solution of the diffusively rescaled model converges to a solution of an advection-diffusion equation.

1. Introduction. Selection-mutation equations with continuous trait values have been extensively studied in the literature (e.g. [1, 2, 8, 9, 10, 11, 12, 13, 17, 18]). Such equations have been formulated on several state spaces including the space of integrable functions (e.g., [10, 11, 12]) and the space of finite signed measures (e.g., [1, 13, 17]). For example, in [10] and [11] the authors consider selection-mutation equations in $L^1$ describing the density of individuals with respect to a continuous phenotypic evolutionary trait. They establish the existence of steady states and study their asymptotic behavior as the mutation rate goes to zero. Papers with selection-mutation models formulated on the space of measures are fewer. We mention for instance [1] where the authors consider a selection-mutation model in $\mathcal{M}_b(Q)$, the space of finite signed measures on a compact metric space $Q$, and study the long term behavior of the solution. They show that for small mutation when the trait space is discrete there exists an equilibrium measure that attracts all solutions with nonnegative not identically zero initial measure/condition. In [13] the authors...
consider a selection-mutation model formulated on the space of measures and study existence-uniqueness of nonnegative solutions to this model.

In this paper, we consider a selection-mutation equation on $\mathcal{M}_b(\mathbb{R}^d)$ (the space of finite signed measures on $\mathbb{R}^d$), modeling the evolution of a population structured by a trait $x \in \mathbb{R}^d$. Denoting by $\mu_t$ the trait distribution at time $t$, $\mu_t$ evolves according to

$$\partial_t \mu_t + \nabla \cdot (v[\mu_t] \mu_t) = \int_{\mathbb{R}^d} K(z, \cdot) d\mu_t(z) - b\mu_t,$$

with $K : \mathbb{R}^d \to \mathcal{M}_{b,+}(\mathbb{R}^d)$ (the nonnegative cone of $\mathcal{M}_b(\mathbb{R}^d)$) and $b(x) = \int_{\mathbb{R}^d} K(x, dy)$. Here, the kernel $K(x, dy)$ models the trait distribution of the offspring of an individual with trait $x$ and $b(x)$ is the birth rate of an individual with trait $x$. Considering such a general $K(x, dy)$ is important as it allows us to treat in a unified way continuous and discrete selection-mutation kernels and offspring distributions. The function $v$ is a density-dependent vector-field and the advection term $\nabla \cdot (v[\mu_t] \mu_t)$ models the individuals actively adapting to their environment seeking phenotype changes that makes them fitter. These changes can be seen as stress-induced epi-mutations (e.g., [5]). Such advection terms are also referred to in the literature as fast evolution with respect to the structure variable [16, 28].

We are interested in studying the limiting behavior of the solution to equations of type (1) when mutation becomes small, that is, when $K$ is suitably rescaled so as to converge to a pure selection kernel. This question has already been addressed in several selection-mutation models in different settings. For example in [28][Chapter 9] the author considers the model

$$\partial_t n(t, x) = \int_{\mathbb{R}^d} \hat{K}(z, x - z)n(t, z)dz - b(x)n(t, x), \quad b(x) = \int_{\mathbb{R}^d} \hat{K}(x, y)dy$$

on the state space $C([0, T], L^1(\mathbb{R}^d))$ where $\hat{K}(x, y) \geq 0$ models the mutations. He then rescales this equation using diffusive rescaling as follows:

$$\partial_t n_\varepsilon(t, x) = \frac{1}{\varepsilon^2} \left( \int_{\mathbb{R}^d} \varepsilon^{-d} \hat{K}_\varepsilon(z, (x - z)/\varepsilon)n_\varepsilon(t, z)dz - b(x)n_\varepsilon(t, x) \right).$$

The family of kernels $\hat{K}_\varepsilon$ allows for a slight asymmetry of the offspring distribution namely

$$\int_{\mathbb{R}^d} y\hat{K}_\varepsilon(x, y)dy = \varepsilon U(x) \quad \text{where} \quad U \in L^\infty(\mathbb{R}^d).$$

Under some regularity assumptions on $\hat{K}_\varepsilon$ the solution $n_\varepsilon$ is shown to converge, up to a subsequence, weakly in $L^2([0, T] \times \mathbb{R}^d)$ to $n$, the solution of the convection-diffusion equation

$$\partial_t n + \nabla \cdot (Un) = \frac{1}{2} \sum_{i,j=1}^d \partial_{x_i} B_{ij} n, \quad B_{ij}(x) = \lim_{\varepsilon \to 0} \int_{\mathbb{R}^d} y_i y_j \hat{K}_\varepsilon(x, y)dy.$$

A different rescaling, namely the hyperbolic rescaling

$$\partial_t n_\varepsilon(t, x) = \frac{1}{\varepsilon} \left( \int_{\mathbb{R}^d} \varepsilon^{-d} \hat{K}(z, (x - z)/\varepsilon)n(t, z)dz - b(x)n_\varepsilon(t, x) \right)$$

leads (in the limit as $\varepsilon \to 0$) to the transport equation

$$\partial_t n + \nabla \cdot (Un) = 0.$$
Similar results (but from a different perspective) can be found in the book [6] where non-local equations of the form

\[ \partial_t u(t, x) = (J * u)(t, x) - u(t, x) = \int_{\mathbb{R}^d} J(x - z)(u(t, z) - u(t, x)) \, dz \]

are considered. Here, \( J \) is a smooth radial nonnegative kernel with integral one. By considering the rescaled kernel \( J_{\varepsilon}(x) = \varepsilon^{-d} J(x/\varepsilon) \), it is proved that the solution to the rescaled equation converges to the solution of the heat equation and an estimate of the rate of convergence is provided. Many other variants of this equation are also considered in [6].

In the aforementioned works, the state spaces are Lebesgue or Sobolev function spaces. A similar result in the state space of measures can be found in [15] where the authors investigate the model

\[ \partial_t \xi_t(x) = [(1 - \eta(x)) b(x, V * \xi_t(x)) - d(x, U * \xi_t(x))] \xi_t(x) + \int_{\mathbb{R}^d} M(z, x - z) \eta(z) b(z, V * \xi_t(z)) \xi_t(z) \, dz \]

with initial condition \( \xi_{t=0} = \xi_0 \in \mathcal{M}_b(\mathbb{R}^d) \). The solution \( \xi_t \) is obtained as the limit of the law of a rescaled probabilistic particles system where each particle represents an animal and the location of the particle is the phenotypic characterization (trait) of the animal. Here, \( b \) and \( d \) are the birth rate and death rate, \( \ast \) denotes the convolution operator, which means that the interaction kernels \( U \) and \( V \) give the weight of each individual when interacting with a focal individual as a function of how phenotypically different they are, \( \eta \) is the mutation probability and \( M(x, y) \) is the mutation kernel. They rescale the mutation kernel \( M(x, y) \) assuming it is the density of a random variable with mean zero and variance-covariance matrix \( \Sigma(x)/\kappa^2 \), e.g., a Gaussian variable. They also accelerate the birth and death processes at a rate proportional to \( \kappa^2 \) considering \( b_\kappa(x, \zeta) = \kappa^2 r(x) + b(x, \zeta) \) and \( d_\kappa(x, \zeta) = \kappa^2 r(x) + d(x, \zeta) \). Assuming \( \eta < 1 \), they prove among other results that the law of the particle system converges as \( \kappa \to +\infty \) to the unique solution of the diffusion equation

\[ \partial_t \xi_t(x) = [b(x, V * \xi_t(x)) - d(x, U * \xi_t(x))] \xi_t(x) + \frac{1}{2} \Delta (\sigma^2 r \mu_\xi)(x). \]

Following those results we rescale (1) considering

\[ \partial_t \mu_t + \nabla \cdot (v[\mu_t^\varepsilon]) = \frac{1}{\varepsilon^2} \left( \int_{\mathbb{R}^d} K_{\varepsilon}(z, \cdot) \, d\mu_t^\varepsilon(z) - b \mu_t^\varepsilon \right). \quad (2) \]

We will explain in detail below the appropriate way to rescale our general measure kernel \( K(x, dy) \). The main result of this paper states that as \( \varepsilon \to 0 \) the solution \( \mu_t^\varepsilon \) of (2) converges in \( C([0, T], \mathcal{M}_{b_+}(\mathbb{R}^d)) \), \( T > 0 \), to a solution of the diffusion equation

\[ \partial_t \mu_t + \nabla \cdot (v[\mu_t^\varepsilon]) = \frac{1}{2} \sum_{i,j=1}^{d} \partial_{ij} \{d_{ij}(\cdot, \mu_t) \mu_t \} \quad (3) \]

where the diffusion coefficients \( d_{ij} \) are given by

\[ d_{ij}(x, \mu) = \int_{\mathbb{R}^d} (y_i - x_i)(y_j - x_j) \, K(x, dy) \quad 1 \leq i, j \leq d. \]
Our model and proofs are flexible enough to allow for several generalizations including asymmetric offspring distributions, hyperbolic rescaling and systems of equations.

This paper is organized as follows: in Section 2 we formulate a selection-mutation model on the space of finite signed measures similar to (1). We present the rescaled equation of type (2) and the main result concerning the convergence of \( \mu_t \) to \( \mu \) (the solution of the limiting equation of the form (3)) is presented in Section 6. A final section is devoted to some extensions of our model (asymmetric distribution, hyperbolic rescaling and systems of equations).

2. The model and the main result. We begin by modeling the selection-mutation kernel. To this end, consider a kernel \( K : \mathbb{R}^d \to \mathcal{M}_{b,+}(\mathbb{R}^d) \) where for any \( x \in \mathbb{R}^d \), \( K(x, dy) \) is a nonnegative finite measure over \( \mathbb{R}^d \) which models the distribution of offspring of an animal with trait value \( x \). Let \( b(x) = \int_{\mathbb{R}^d} K(x, dy) \) be the total offspring (birth rate) of an individual with trait \( x \). Note that \( K(x, dy) \) is a probability measure when \( b(x) = 1 \).

We point out that modeling the distribution of offspring \( K(x, dy) \) by a general measure has the advantage of allowing the treatment of singular and absolutely continuous distributions simultaneously. For example, following [8], we can choose

\[
K(x, dy) = b(x)(1-m(x))\delta_x + b(x)\beta(y-x)dy,
\]

where \( \beta \in L^1(\mathbb{R}^d) \) is nonnegative and \( \beta(y-x)dy \) is the probability density that an animal with trait \( x \) gives rise to an offspring with trait \( y \), and \( m(x) \in [0,1] \) is the proportion of offsprings with mutations (i.e., with different trait values).

Consider the equation

\[
\partial_t \mu_t + \nabla \cdot (v[\mu_t] \mu_t) = \int_{\mathbb{R}^d} K(z, \cdot) d\mu_t(z) - b\mu_t,
\]

where \( v : \mathcal{M}_b(\mathbb{R}^d) \to W^{1,\infty}(\mathbb{R}^d) \) (the space of bounded Lipschitz continuous functions) is a density-dependent vector-field. For example, a common choice of \( v \) that depends on \( \mu \) through some weighted mean of \( \mu \) is of the form

\[
v[\mu](x) = V \left( x, \int_{\mathbb{R}^d} \eta(y)d\mu(y) \right)
\]

for given maps \( V : \mathbb{R}^d \times \mathbb{R} \to \mathbb{R}^d \) and \( \eta : \mathbb{R}^d \to \mathbb{R} \).

We are interested in rescaling the reproduction kernel \( K \) so that the size of the mutations goes to 0 as the rescaling parameter \( \varepsilon \) goes to 0. For instance if \( K(x, dy) = \beta(y-x)dy \) we let \( \beta_{\varepsilon}(z) = \varepsilon^{-d} \beta(z/\varepsilon) \) and then rescale \( K \) considering \( K_{\varepsilon}(x, dy) = \beta_{\varepsilon}(y-x)dy \). To do the same for a general kernel, it is useful to think that the probability measure \( \beta_{\varepsilon}(y-x)dy \) converges to the Dirac mass \( \delta_x \) as \( \varepsilon \to 0 \) and, from a more abstract point of view, that the curve \( \varepsilon \in [0,1] \to \beta_{\varepsilon}(y-x)dy \) connects \( \delta_x \) to \( \beta(y-x)dy \) in \( P(\mathbb{R}^d) \), the space of probability measures. Optimal transport theory provides us with a simple way to implement this idea when replacing \( \beta(y-x)dy \) by an arbitrary transition kernel \( K(x, dy) \). Consider the constant map \( T_x(y) = x, y \in \mathbb{R}^d \). Notice that \( T_x \mu \), the push-forward of the
bounded measure \( \mu \) by \( T_x \), is simply \( \mu(\mathbb{R}^d)^{\delta_x} \). We then consider the new kernel \( K_\varepsilon(x, dy) \), \( \varepsilon \in [0, 1] \), defined by
\[
K_\varepsilon(x, dy) = (T_x + \varepsilon(Id - T_x))zK(x, dy)
\] (where \( Id \) is the identity map). This means that for any test-function \( \phi : \mathbb{R}^d \to \mathbb{R} \),
\[
(K_\varepsilon(x, dy), \phi) = \int_{\mathbb{R}^d} \phi(x + \varepsilon(y - x)) K(x, dy).
\]
(8)

Taking \( \phi \equiv 1 \), we observe that the measure \( K_\varepsilon(x, dy) \) has the same total mass as \( K(x, dy) \), namely \( b(x) \). Thus, for each \( x \), the map \( \varepsilon \in [0, 1] \to K_\varepsilon(x, dy) \) is a curve in \( \mathcal{M}_{b,+}(\mathbb{R}^d) \) connecting \( b(x) \delta_x \) to \( K(x, dy) \). When \( \varepsilon \equiv 1 \), so that \( K(x, dy) \) is a probability measure in \( P(\mathbb{R}^d) \) for each \( x \), it is known that this curve is indeed the shortest path (or McCann interpolant) connecting these two measures in the Wasserstein space \( (P(\mathbb{R}^d), W_2) \), where \( W_2 \) is the Monge-Kantorovich distance (see e.g. [29]).

Notice that in the smooth case where \( K(x, dy) = \beta(y - x)dy \) and \( \beta \in L^1(\mathbb{R}^d) \), we have for any test-function \( \phi \) that
\[
(K_\varepsilon(x, dy), \phi) = \int_{\mathbb{R}^d} \phi(x + \varepsilon(y - x)) K(x, dy) = \int_{\mathbb{R}^d} \phi(x + \varepsilon(y - x))\beta(y - x) dy = \int_{\mathbb{R}^d} \phi(z)\varepsilon^{-d}\beta((z - x)/\varepsilon) dz.
\]
So \( K_\varepsilon(x, dy) = \beta_\varepsilon(y - x) dy \) and we recover the smooth case rescaling kernel.

Rescaling the kernel \( K \) by considering \( K_\varepsilon \) defined in (8) and at the same time increasing the mutation rate leads to the equation
\[
\partial_t \mu_t + \nabla \cdot (\psi[\mu_t] \mu_t) = \frac{1}{\varepsilon^2} \left( \int_{\mathbb{R}^d} K_\varepsilon(z, \cdot) d\mu_t(z) - b \mu_t \right).
\]
(9)

Inspired by model (9), in this paper we study the following more general equation with a nonlinear source term \( \bar{N}(t, \cdot, \mu) \) and a nonlinear selection-mutation kernel \( K(x, \mu, dy) \):
\[
\partial_t \mu_t + \nabla \cdot (\psi[\mu_t] \mu_t) = \bar{N}(t, \cdot, \mu_t) \mu_t + \frac{1}{\varepsilon^2} \left( \int_{\mathbb{R}^d} K_\varepsilon(z, \mu_t, \cdot) d\mu_t(z) - b(\cdot, \mu_t) \mu_t \right) \quad (10)
\]
with a given nonnegative initial condition \( \mu_0 \in \mathcal{M}_{b,+}(\mathbb{R}^d) \), and where \( b(x, \mu) = \int_{\mathbb{R}^d} K(x, \mu, dy) \). Introducing a nonlinearity with respect to \( \mu \) in the kernel \( K \) and hence in the birth rate \( b \) allows to model competition between individuals due to limiting resources (e.g., [1, 2]). Furthermore, the added term \( \bar{N}(t, \cdot, \mu_t) \mu_t \) can be used to model a loss term due to mortality (e.g., [1, 2]) and/or a source term due to external forces (e.g., seeds blown by wind in a population of plants).

For \( \varepsilon > 0 \), denote by \( \mu_t^\varepsilon \) the solution to equation (10). In Theorem 5.3 below we extend the results obtained in [3] to show that given a nonnegative initial condition \( \mu_0 \), then equation (10) has a unique nonnegative solution \( \mu_t^\varepsilon \) defined for any \( t \geq 0 \). Our main goal here is to study the asymptotic behavior of \( \mu_t^\varepsilon \), the solution of (10), as \( \varepsilon \to 0 \).

The need to consider nonnegative initial conditions \( \mu_0 \), and thus nonnegative solutions \( \mu_t^\varepsilon \), is motivated both from the biological and the mathematical point of view. From a biological perspective \( \mu_t^\varepsilon \) models the distribution of the population in the trait space and thus must be a nonnegative measure. From a mathematical point of view, the non-negativity of \( \mu_t^\varepsilon \) implies that the total variation norm of \( \mu_t^\varepsilon \)
where the diffusion coefficients is\( \| \mu_\varepsilon \|_{TV} = \int_{\mathbb{R}^d} d\mu_\varepsilon \) which is useful to obtain the bound (30) below. Moreover, it allows us to use the norm \( \| \cdot \|_{BL,3} \) defined in (18) below which is more suitable to our problem than the usual Bounded Lipschitz (BL) norm and at the same time has the same convergence properties as the BL norm on the positive cone of the measures.

We will show that up to a subsequence, as \( \varepsilon \to 0 \), the solution \( \mu_\varepsilon \) of (10) converges to \( \mu_t \) a solution of the nonlinear convection-diffusion equation

\[
\begin{aligned}
\partial_t \mu_t + \nabla \cdot (v|\mu_t|) \mu_t &= N(t, \cdot, \mu_t) \mu_t + \frac{1}{2} \sum_{i,j=1}^d \partial_{ij}(d_{ij}(\cdot, \mu_t)) \\
\mu_{t=0} &= \mu_0
\end{aligned}
\]  

(11)

where the diffusion coefficients \( d_{ij} \) are given by

\[
d_{ij}(x, \mu) = \int_{\mathbb{R}^d} (y_i - x_i)(y_j - x_j) K(x, \mu, dy) \quad 1 \leq i, j \leq d.
\]  

(12)

To our knowledge, existence of solutions for a parabolic but possibly degenerate equation like (11) has not been established in the literature. Furthermore, this limiting equation provides a connection between two types of approaches that have been used to model small mutation including using an integral operator (e.g., [1]) and a diffusion operator (e.g., [20]).

The matrix \( (d_{ij}(x, \mu))_{ij} \) can be seen as a covariance matrix. Indeed, let \( \tilde{K}(x, \mu, dy) := \tau_{-z}^\varepsilon K(x, \mu, dy) \) where \( \tau_{-z}(y) = y - z \). Thus, \( \int_{\mathbb{R}^d} \phi(y) \tilde{K}(x, \mu, dy) = \int_{\mathbb{R}^d} \phi(y - x) \tilde{K}(x, \mu, dy) \) for any \( \phi \). In particular, we can rewrite \( d_{ij} \) as

\[
d_{ij}(x, \mu) = \int_{\mathbb{R}^d} y_i y_j \tilde{K}(x, \mu, dy) \quad 1 \leq i, j \leq d.
\]

Recall that \( K(x, \mu, dy) \) is a nonnegative finite measure with total mass \( b(x, \mu) \). Therefore, so is \( \tilde{K}(x, \mu, dy) \). Moreover, the assumption (K4) below implies that \( \int_{\mathbb{R}^d} y \tilde{K}(x, \mu, dy) = 0 \). Thus, the matrix \( (d_{ij}(x, \mu)/b(x, \mu))_{ij} \) is the covariance matrix of the probability measure \( \tilde{K}(x, \mu, dy)/b(x, \mu) \). In particular, it follows that \( (d_{ij}(x, \mu))_{ij} \) is a symmetric nonnegative matrix. The non-negativity can be seen directly by noticing that for any \( z \in \mathbb{R}^d \),

\[
\sum_{i,j=1}^d d_{ij}(x, \mu) z_i z_j = \int_{\mathbb{R}^d} \sum_{i,j=1}^d (y_i z_i)(y_j z_j) \tilde{K}(x, \mu, dy) = \int_{\mathbb{R}^d} (\sum_{k=1}^d y_k z_k)^2 \tilde{K}(x, \mu, dy) \geq 0.
\]

Before stating the assumptions on the model parameters, we need to recall some known measure-theoretical facts. Unless specified otherwise the space \( \mathcal{M}_b(\mathbb{R}^d) \) of bounded Borel measures on \( \mathbb{R}^d \) is always endowed with the BL norm defined by

\[
\| \mu \|_{BL} = \sup \left\{ \int_{\mathbb{R}^d} \phi d\mu : \phi \in W^{1,\infty}(\mathbb{R}^d), \| \phi \|_\infty + Lip(\phi) \leq 1 \right\}.
\]

Here, \( W^{1,\infty}(\mathbb{R}^d) \) denotes the space of bounded Lipschitz continuous functions, i.e., \( \phi \in W^{1,\infty}(\mathbb{R}^d) \) if \( \phi \) is bounded and there exists \( C > 0 \) such that

\[
|\phi(x) - \phi(y)| \leq C|x - y|, \text{ for any } x, y \in \mathbb{R}^d.
\]

We denote by \( Lip(\phi) \) the least admissible constant \( C \) and we let \( \| \phi \|_{W^{1,\infty}} = \| \phi \|_\infty + Lip(\phi) \).
Other topologies can be considered on \( M_b(\mathbb{R}^d) \) like the weak* convergence and the total-variation norm. However, the BL norm seems to be the most useful when dealing with transport equations. We recall a few facts concerning the relationship between these three topologies. Let \( \mu_n, \mu \in M_b(\mathbb{R}^d) \). If \( \mu_n \to \mu \) weak* (in the sense that \((\mu_n, \phi) \to (\mu, \phi)\) for any \( \phi \in C_b(\mathbb{R}^d) \)) then \( \|\mu_n - \mu\|_{BL} \to 0 \) but in general the converse is false except when \( \mu_n, \mu \geq 0 \) (see Thm 6 and 8 in [19]). Also, notice that a converging sequence \( (\mu_n)_n \) in the BL norm may not be tight nor bounded in total variation. As an example, consider \( \mu_n = \sqrt{n}(\delta_{\alpha+1/n} - \delta_{\alpha}) \). Then \( \|\mu_n\|_{BL} = \sqrt{n}/n \to 0 \) but \( (\mu_n)_n \) is neither tight nor TV-bounded since \( \|\mu_n\|_{TV} = 2\sqrt{n} \). Notice that \( M_b(\mathbb{R}^d) \) is not complete under the BL norm (see Thm 4.8 in [22] for a complete treatment of the completeness under the BL norm) but that a subset of the form

\[
M_{b,R}(\mathbb{R}^d) := \{ \mu \in M_b(\mathbb{R}^d) : ||\mu||_{TV} \leq R \},
\]

i.e., the ball in \( M_b(\mathbb{R}^d) \) of total variation radius \( R \), is complete (see e.g. Thm 2.7 in [23]).

We denote by \( C([0, T], M_b(\mathbb{R}^d)) \) the space of continuous functions from \([0, T] \) to \((M_b(\mathbb{R}^d), \|\cdot\|_{BL}) \) endowed with the usual sup-norm. Next we state our assumptions on the model functions \( v, N, K \) and \( b \) appearing in (10). We start with the vector field \( v \) and assume that

\[
v : \mu \in M_b(\mathbb{R}^d) \to v[\mu] \in W^{1,\infty}(\mathbb{R}^d)
\]
satisfies

(V1) \( v : M_b(\mathbb{R}^d) \to L^\infty(\mathbb{R}^d) \) is continuous and for any \( R > 0 \) there exists \( L_{v,R} > 0 \) such that

\[
\|v[\mu] - v[\bar{\mu}]\|_\infty \leq L_{v,R} ||\mu - \bar{\mu}||_{BL} \quad \text{for any } \mu, \bar{\mu} \in M_{b,R}(\mathbb{R}^d),
\]

(V2) there exists \( C_v > 0 \) such that

\[
\text{Lip}(v[\mu]) \leq C_v \quad \text{for any } \mu \in M_b(\mathbb{R}^d).
\]

Concerning the source term \( N \), we assume that

\[
N : (t, x, \mu) \in \mathbb{R}_+ \times \mathbb{R}^d \times M_b(\mathbb{R}^d) \to N(t, x, \mu) \in \mathbb{R}
\]
is continuous in \((t, x, \mu)\) and

(N1) for any \( R > 0 \), there exist \( L_{N,R} > 0 \) such that for any \( t \in \mathbb{R} \), any \( \mu, \bar{\mu} \in M_{b,R}(\mathbb{R}^d) \) and any \( x \in \mathbb{R}^d \),

\[
|N(t, x, \mu) - N(t, x, \bar{\mu})| \leq L_{N,R} ||\mu - \bar{\mu}||_{BL},
\]

(N2) there exists \( C_N > 0 \) such that

\[
||N(t, \cdot, \mu)||_{W^{1,\infty}} \leq C_N \quad \text{for any } \mu \in M_b(\mathbb{R}^d) \text{ and any } t \geq 0.
\]

Finally, concerning the reproduction kernel \( K \) we assume that

\[
K : (x, \mu) \in \mathbb{R}^d \times M_b(\mathbb{R}^d) \to K(x, \mu, dy) \in M_b(\mathbb{R}^d)
\]
satisfies

(K0) \( K(x, \mu, dy) \geq 0 \) whenever \( \mu \geq 0 \),

(K1) for any \( \phi : \mathbb{R}^d \to \mathbb{R} \) measurable and bounded and any \( \mu \in M_b(\mathbb{R}^d) \), the function \( x \to \int_{\mathbb{R}^d} \phi(y)K(x, \mu, dy) \) is measurable,

(K2) the function \( b : (x, \mu) \in \mathbb{R}^d \times M_b(\mathbb{R}^d) \to \int_{\mathbb{R}^d} K(x, \mu, dy) \in \mathbb{R} \) is continuous and satisfies
(i) there exists $C_b > 0$ such that $\|b(\cdot, \mu)\|_{W_1^1} \leq C_b$ for any $\mu \in \mathcal{M}_b(\mathbb{R}^d)$,
(ii) for any $R > 0$ there exists $L_{b,R} > 0$ such that
$$|b(x, \mu) - b(x, \tilde{\mu})| \leq L_{b,R} \|\mu - \tilde{\mu}\|_{BL}$$
for any $\mu, \tilde{\mu} \in \mathcal{M}_{b,R}(\mathbb{R}^d)$ and any $x \in \mathbb{R}^d$,
(K3) for any $\phi \in W^{1,\infty}(\mathbb{R}^d)$, the function $x \to (K(x, \mu, \cdot), \phi)$ is Lipschitz with
$$\sup_{\mu \in \mathcal{M}_b(\mathbb{R}^d), \|\phi\|_{W^{1,\infty}} \leq 1} \text{Lip}(x \to (K(x, \mu, \cdot), \phi)) < \infty,$$
(K4) for any $x \in \mathbb{R}^d$ and any $\mu \in \mathcal{M}_b(\mathbb{R}^d)$,
$$\int_{\mathbb{R}^d} (y - x) K(x, \mu, dy) = 0,$$
(K5) for any $R > 0$ there exists $C_{K,R} > 0$ such that
$$\int_{\mathbb{R}^d} |y - x|^2 K(x, \mu, dy) \leq C_{K,R}$$
for any $x \in \mathbb{R}^d$ and $\mu \in \mathcal{M}_{b,R}(\mathbb{R}^d)$,
(K6) if $\|\mu_n - \mu\|_{BL} \to 0$ then $\|K(x, \mu_n, \cdot) - K(x, \mu, \cdot)\|_{BL} \to 0$ for any $x \in \mathbb{R}^d$, and
for any $R > 0$ there exists $L_{K,R} > 0$ such that
$$\|K(x, \tilde{\mu}, \cdot) - K(x, \mu, \cdot)\|_{BL} \leq L_{K,R} \|\tilde{\mu} - \mu\|_{BL}$$
for any $x \in \mathbb{R}^d$ and any $\mu, \tilde{\mu} \in \mathcal{M}_{b,R}(\mathbb{R}^d)$,
(K7) the function $d_{ij}(x, \mu) = \int_{\mathbb{R}^d} (y_i - x_i)(y_j - x_j) K(x, \mu, dy)$ satisfies
(i) for any $R > 0$ there exists $L_{d,R} > 0$ such that
$$\|d_{ij}(\cdot, \mu)\|_{W^{1,\infty}} \leq L_{d,R}$$
for any $x \in \mathbb{R}^d$ and any $\mu \in \mathcal{M}_{b,R}(\mathbb{R}^d)$.
(ii) for any $x \in \mathbb{R}^d$ and for any nonnegative TV-bounded sequence $(\mu_n)_n$ converging in the BL norm to some $\mu$, we have $d_{ij}(x, \mu_n) \to d_{ij}(x, \mu)$.
(K8) There exists $C_K > 0$ such that
$$\|K(x, \mu, \cdot)\|_{TV} \leq C_K$$
for any $x \in \mathbb{R}^d$ and any $\mu \in \mathcal{M}_b(\mathbb{R}^d)$.

An example of admissible vector fields $v[\mu]$ can be obtained through weighted means of $\mu$ such as
$$v[\mu](x) = V \left( x, \int_{\mathbb{R}^d} \eta(y) d\mu(y) \right)$$
for given maps $V : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d$ and $\eta : \mathbb{R}^d \to \mathbb{R}$. Indeed, if we assume that $V$ and $\eta$ are both bounded and globally Lipschitz then for any $\mu, \tilde{\mu} \in \mathcal{M}_b(\mathbb{R}^d)$,
$$\text{Lip}(v[\mu]) \leq \text{Lip}(V) \quad \text{and} \quad \|v[\mu] - v[\tilde{\mu}]\|_{\infty} \leq \text{Lip}(V) \|\eta\|_{W^{1,\infty}} \|\mu - \tilde{\mu}\|_{BL}$$
so that assumptions (V1) and (V2) hold.

We can build examples of admissible source terms $\bar{N}$ in the same spirit considering
$$\bar{N}(t, x, \mu) = A \left( t, x, \int_{\mathbb{R}^d} \eta(y) d\mu(y) \right)$$
for given bounded and globally Lipschitz maps $A : \mathbb{R}_+ \times \mathbb{R}^d \times \mathcal{M}_b(\mathbb{R}^d) \to \mathbb{R}$ and $\eta : \mathbb{R}^d \to \mathbb{R}$. Then $\bar{N}$ is continuous and satisfies assumptions (N1) and (N2) since
$$|\bar{N}(t, x, \mu) - \bar{N}(t, x, \tilde{\mu})| \leq \text{Lip}(A) \|\eta\|_{W^{1,\infty}} \|\mu - \tilde{\mu}\|_{BL}$$
and
$$\|\bar{N}(t, \cdot, \mu)\|_{W^{1,\infty}} \leq \|A\|_{W^{1,\infty}}.$$

Since the general formulation of the selection-mutation kernel $K$ (and its rescaled version $K_\epsilon$) that we consider in this paper is important and novel, we will devote
Theorem 2.1. Assume that \( K \), kernels, the next section to provide several examples of biologically relevant and admissible 

kernels, that satisfy the assumptions above.

Now we are in a position to state our main result:

Theorem 2.1. Assume that \( v: \mathcal{M}_b(\mathbb{R}^d) \rightarrow W^{1,\infty}(\mathbb{R}^d) \) satisfies (V1)-(V2), that \( \bar{N}: \mathbb{R}_+ \times \mathbb{R}^d \times \mathcal{M}_b(\mathbb{R}^d) \rightarrow \mathbb{R} \) is continuous and satisfies (N1)-(N2), and that \( K: \mathbb{R}^d \times \mathcal{M}_b(\mathbb{R}^d) \rightarrow \mathcal{M}_b(\mathbb{R}^d) \) satisfies (K0)-(K8). Then, for any nonnegative initial condition \( \mu_0 \in \mathcal{M}_b(\mathbb{R}^d) \) such that \( \int_{\mathbb{R}^d} |x| \, d\mu_0(x) < \infty \), there exists a solution \( \mu \in C([0, \infty), \mathcal{M}_{b,+}(\mathbb{R}^d)) \) to

\[
\begin{align*}
\frac{\partial}{\partial t} \mu_t + \nabla \cdot (v(\mu_t) \mu_t) &= \bar{N}(t, \cdot, \mu_t) \mu_t + \frac{1}{2} \sum_{i,j=1}^{d} \partial_{ij} (d_{ij}(\cdot, \mu_t) \mu_t) \\
\mu_{t=0} &= \mu_0
\end{align*}
\]

where the diffusion coefficients \( d_{ij} \) are given by

\[
d_{ij}(x, \mu) = \int_{\mathbb{R}^d} (y_i - x_i)(y_j - x_j) K(x, \mu, dy) \quad 1 \leq i, j \leq d.
\]

This solution is obtained as the limit as \( \varepsilon \rightarrow 0 \) (up to a subsequence) of the unique solution \( \mu^\varepsilon \) of equation (10), namely

\[
\begin{align*}
\frac{\partial}{\partial t} \mu^\varepsilon_t + \nabla \cdot (v\mu^\varepsilon_t) \mu^\varepsilon_t &= \bar{N}(t, \cdot, \mu^\varepsilon_t) \mu^\varepsilon_t + \frac{1}{2} \sum_{i,j=1}^{d} \partial_{ij} (d_{ij}(\cdot, \mu^\varepsilon_t) \mu^\varepsilon_t) \\
\mu^\varepsilon_{t=0} &= \mu_0
\end{align*}
\]

in the sense that \( \mu^\varepsilon \rightarrow \mu \) in \( C([0, T], \mathcal{M}_b(\mathbb{R}^d)) \) for any \( T > 0 \). Moreover for any \( T > 0 \) there exist \( C_T > 0 \) such that

\[
\int_{\mathbb{R}^d} (1 + |x|) \, d\mu_t \leq C_T \quad 0 \leq t \leq T.
\]

We point out that the existence a unique solution \( \mu^\varepsilon \) of equation (10) stated in Theorem 2.1 will be established in Theorem 5.3 below.

3. Examples of admissible selection-mutation kernels. To demonstrate the usefulness of the general formulation of the kernel \( K(x, \mu, dy) \) that we consider here, we present several examples of selection-mutation kernels from the literature that can be unified under this general formulation.

- A pure selection kernel: a simple example is \( K_S(x, \mu, dy) = \delta_x \) which models the case where mutation never occurs, i.e., the offspring inherits exactly the trait of its parent. In this case \( b = 1, d_{ij} = 0 \) and the assumptions (K0)-(K8) trivially hold. Such a mutation kernel has been considered in [2].

- A continuous mutation distribution: for this kernel the offspring’s trait is assumed to be distributed according to a continuous probability distribution \( \beta \), namely,

\[
K_{CM}(x, \mu, dy) = \beta(x, \mu, y - x)dy
\]

where \( \beta(x, \mu, \cdot) \) is a nonnegative integrable function (so that \( K(x, \mu, dy) \in \mathcal{M}_{b,+}(\mathbb{R}^d) \)). This type of kernel has been studied, for example, in [10]. Here,

\[
d_{ij}(x, \mu) = \int_{\mathbb{R}^d} z_i z_j \beta(x, \mu, z)dz.
\]
Notice in particular that assumption (K7) is satisfied if we assume that $d_{ij}$ is continuous in $\mu$ for any $x \in \mathbb{R}^d$, and if for any $z \in \mathbb{R}^d$ and any $\mu \in \mathcal{M}_b(\mathbb{R}^d)$, the function $x \in \mathbb{R}^d \rightarrow \beta(x, \mu, z)$ is Lipschitz and satisfies for any $R > 0$,

$$
\sup_{\mu \in \mathcal{M}_b(\mathbb{R}^d)} \int_{\mathbb{R}^d} |z|^2 \|\beta(\cdot, \mu, z)\|_{W^{1,\infty}} \, dz < \infty.
$$

Indeed,

$$
\sup_{\mu \in \mathcal{M}_b(\mathbb{R}^d)} \|d_{ij}(\cdot, \mu)\|_{W^{1,\infty}} \leq \sup_{\mu \in \mathcal{M}_b(\mathbb{R}^d)} \int_{\mathbb{R}^d} |z|^2 \|\beta(\cdot, \mu, z)\|_{W^{1,\infty}} \, dz < \infty.
$$

- **A discrete mutation distribution:** for this kernel the offspring’s trait is assumed to be distributed according to a discrete probability distribution, namely,

$$
K_{DM}(x, \mu, dy) = \sum_{k=1}^{N} \alpha_k(x, \mu) \delta_{x+h_k(x, \mu)}
$$

where $N \in \mathbb{N}$, $\alpha_1, \ldots, \alpha_N \in [0,1]$ with $\sum_{k=1}^{N} \alpha_k = 1$, and $h_1(x, \mu), \ldots, h_N(x, \mu) \in \mathbb{R}^d$. This reproduction kernel models the situation where an offspring of a parent with trait $x$ will have trait $x+h_k(x, \mu)$ with probability $\alpha_k(x, \mu) \in [0,1]$. A similar discrete selection-mutation kernel has been considered, for example, in [4].

- **A mixed mutation distribution:** we can combine the previous cases to obtain a reproduction kernel of the form

$$
K(x, \mu, dy) = (1 - a(x, \mu) - b(x, \mu)) \delta_x + a(x, \mu) K_{CM}(x, \mu, dy) + b(x, \mu) K_{DM}(x, \mu, dy)
$$

where $a, b \in [0,1]$ are the probabilities of choosing $K_{CM}$ and $K_{DM}$, respectively. Of course this is only an example, since our model allows a priori any kind of measure (for instance a measure supported on some sub-manifold like a sphere for example).

A simple particular case is worth mentioning which corresponds to the case where $K$ is homogeneous in the trait space, i.e., $K$ depends on $x$ only through a translation at $x$ of a given measure $K(\mu, dy)$. To treat this case, consider a nonnegative bounded measure $K(\mu, dy)$. Let $\tau_x : y \in \mathbb{R}^d \rightarrow y + x \in \mathbb{R}^d$. Define $K(x, \mu, dy) := \tau_x^* K(\mu, dy)$ so that $\int_{\mathbb{R}^d} \phi(y) K(x, \mu, dy) = \int_{\mathbb{R}^d} \phi(y + x) K(\mu, dy)$ for any $\phi$. Here $K(\mu, dy)$ is the trait distribution of an offspring with parent’s trait $0$ and $K(x, \mu, dy)$ is obtained centering $K(\mu, dy)$ at $x$.

Notice that $K_S(x, dy) = \delta_x$ is space homogeneous (take $K(\mu, dy) = \delta_0$). Space homogeneous versions of $K_{CM}$ and $K_{DM}$ are

$$
K_{CM}(x, \mu, dy) = \beta(\mu, y-x) dy
$$

obtained from $\beta(\mu, y) dy$, and

$$
K_{DM}(x, \mu, dy) = \sum_{k=1}^{N} \alpha_k(\mu) \delta_{x+h_k(\mu)}
$$

obtained from $\sum_{k=1}^{N} \alpha_k(\mu) \delta_{h_k(\mu)}$. 
In the case of a space homogeneous kernel, it is easily seen that (K1) and (K3) are satisfied. Towards (K1) note that if $\nu \in M_{b,+}(\mathbb{R}^d)$ then for any $\phi : \mathbb{R}^d \to \mathbb{R}$ measurable and bounded and any $\mu \in M_b(\mathbb{R}^d)$ we have $\int_{\mathbb{R}^d \times \mathbb{R}^d} |\phi(y+x)|K(\mu, dy)\nu(dx) < \infty$ and a simple application of Fubini’s Theorem results in $x \to \int_{\mathbb{R}^d} \phi(y + x) K(\mu, dy)$ being measurable.

As for (K3), since for any $\phi \in W^{1,\infty}(\mathbb{R}^d)$, $\|\phi\|_{W^{1,\infty}} \leq 1$, we have $(K(x, \mu, \cdot), \phi) = \int \phi(y + x) K(\mu, dy)$ so that

$$
|K(x, \mu, \cdot) - (K(\tilde{x}, \mu, \cdot), \phi)| \leq \int_{\mathbb{R}^d} |\phi(y + x) - \phi(y + \tilde{x})|K(\mu, dy) \\
\leq |x - \tilde{x}|\|K(\mu, dy)\|_{TV} \leq C_K|x - \tilde{x}|,
$$

where we used (K8) in the last inequality. Thus, the function $x \to (K(x, \mu, \cdot), \phi)$ is Lipschitz with Lipschitz constant less than $C_K$ for any $\|\phi\|_{W^{1,\infty}} \leq 1$ and any $\mu \in M_b(\mathbb{R}^d)$.

The remaining assumptions can be simplified as follows by dropping the dependence on $x$.

(K0') $K(\mu, dy) \geq 0$ whenever $\mu \geq 0$,

(K2') the function $b : \mu \in M_b(\mathbb{R}^d) \to \int_{\mathbb{R}^d} K(\mu, dy) \in \mathbb{R}$ is continuous and for any $R > 0$ there exists $L_{b,R} > 0$ such that

$$
|b(\mu) - b(\tilde{\mu})| \leq L_{b,R}\|\mu - \tilde{\mu}\|_{BL} \quad \text{for any } \mu, \tilde{\mu} \in M_{b,R}(\mathbb{R}^d),
$$

(K4') $\int_{\mathbb{R}^d} y K(\mu, dy) = 0$,

(K5') for any $R > 0$ there exists $C_{K,R} > 0$ such that

$$
\int_{\mathbb{R}^d} |y|^3 K(\mu, dy) \leq C_{K,R} \quad \text{for any } \mu \in M_{b,R}(\mathbb{R}^d),
$$

(K6') if $\|\mu_n - \mu\|_{BL} \to 0$ then $\|K(\mu_n, \cdot) - K(\mu, \cdot)\|_{BL} \to 0$, and for any $R > 0$ there exists $L_{K,R} > 0$ such that

$$
\|K(\tilde{\mu}, \cdot) - K(\mu, \cdot)\|_{BL} \leq L_{K,R}\|\tilde{\mu} - \mu\|_{BL} \quad \text{for any } \tilde{\mu}, \mu \in M_{b,R}(\mathbb{R}^d),
$$

(K7') the function $d_{ij}(\mu) = \int_{\mathbb{R}^d} y_i y_j K(\mu, dy)$ is such that $d_{ij}(\mu_n) \to d_{ij}(\mu)$ for any nonnegative TV-bounded sequence $(\mu_n)_n$ converging in the BL norm to $\mu$.

(K8') There exists $C_K > 0$ such that

$$
\|K(\mu, \cdot)\|_{TV} \leq C_K \quad \text{for any } \mu \in M_b(\mathbb{R}^d).
$$

We can make a further simplification assuming that $K$ is space homogeneous and also independent of $\mu$, i.e., $K(x, \mu, dy) = K(x, dy) = \tau_{x}zK(dy)$ for some nonnegative $K(dy) \in M_{b,+}(\mathbb{R}^d)$. It is easily verified that all the assumptions hold with the exception of (K4) and (K5) that can be simplified to

(K4'') $\int_{\mathbb{R}^d} y K(dy) = 0$

(K5'') $\int_{\mathbb{R}^d} |y|^3 K(dy) < +\infty$.

For instance in the case of a continuous kernel $K_{CM}(dy) = \beta(y)dy$ with $\beta \in L^1(\mathbb{R}^d)$, $\beta \geq 0$, assumptions (K4'') and (K5'') reduce to

$$
\int_{\mathbb{R}^d} y \beta(y) \, dy = 0 \quad \text{and} \quad \int_{\mathbb{R}^d} |y|^3 \beta(y) \, dy < \infty.
$$

(15)
In the case of a discrete mutation kernel $K_{DM}(dy) = \sum_{k=1}^{N} \alpha_k \delta_{h_k}$ with $\alpha_1, \ldots, \alpha_N \in [0, 1]$, $\alpha_1 + \cdots + \alpha_N = 1$ and $h_1, \ldots, h_N \in \mathbb{R}^d$, (K5') is void and (K4') is $\sum_{k=1}^{N} \alpha_k h_k = 0$.

Using space homogenous kernels, we can introduce easily some space inhomogeneity and measure-dependence by considering a convex combination like

$$K(x, \mu, dy) = (1 - a_1(x, \mu) - a_2(x, \mu)) \delta_x + a_1(x, \mu) \beta(y - x)dy + a_2(x, \mu) \sum_{k=1}^{N} \alpha_k \delta_{x+h_k}.\quad (16)$$

We then have the following direct corollary of Theorem 2.1:

**Corollary 3.1.** Assume that $a_1, a_2 : \mathbb{R}^d \times \mathcal{M}_b(\mathbb{R}^d) \to [0, 1]$ satisfy $a_1 + a_2 \leq 1$, and for $k = 1, 2$,

(i) there exists $C_a > 0$ such that $\|a_k(\cdot, \mu)\|_{W^{1,\infty}} \leq C_a$ for any $\mu \in \mathcal{M}_b(\mathbb{R}^d)$,

(ii) for any $R > 0$ there exists $L_{a,R} > 0$ such that

$$|a_k(x, \mu) - a_k(x, \tilde{\mu})| \leq L_{a,R} \|\mu - \tilde{\mu}\|_{BL} \quad \text{for any } \mu, \tilde{\mu} \in \mathcal{M}_b(\mathbb{R}^d),$$

(iii) $a_k$ is continuous in $\mu$ for any fixed $x \in \mathbb{R}^d$.

Assume also that $\beta \in L^1(\mathbb{R}^d)$ is nonnegative and satisfies (15) and that $\alpha_1, \ldots, \alpha_N \geq 0$ and $h_1, \ldots, h_N \in \mathbb{R}^d$ satisfy $\sum_{k=1}^{N} \alpha_k h_k = 0$.

Then, for any nonnegative initial condition $\mu_0 \in \mathcal{M}_b(\mathbb{R}^d)$ such that $\int_{\mathbb{R}^d} |x| d\mu_0(x) < \infty$, there exists a solution $\mu \in C([0, +\infty), \mathcal{M}_b(\mathbb{R}^d))$ to (11) with diffusion coefficients

$$d_{ij}(x, \mu) = a_1(x, \mu) \int_{\mathbb{R}^d} z_i z_j \beta(z) dz + a_2(x, \mu) \sum_{k=1}^{N} h_k^i h_k^j$$

(where $h_k = (h_k^1, \ldots, h_k^d)$). This solution is obtained as the limit as $\varepsilon \to 0$ (up to a subsequence) of the solutions $\mu_\varepsilon^\beta$ of (10) in the sense that $\mu_\varepsilon^\beta \to \mu$ in $C([0, T], \mathcal{M}_b(\mathbb{R}^d))$ for any $T > 0$.

**Proof.** Notice that

$$b(x, \mu) = \int_{\mathbb{R}^d} K(x, \mu, dy) = (1 - a_1(x, \mu) - a_2(x, \mu)) + a_1(x, \mu) \int_{\mathbb{R}^d} \beta(y - x) dy + a_2(x, \mu) \sum_{k=1}^{N} \alpha_k.$$ 

In view of the assumptions on $a_1$ and $a_2$, $b$ satisfies (K2) and $d_{ij}$ satisfies (K7). The assumptions (K3) and (K6) clearly hold, while (K4) and (K5) are satisfied by the assumptions made on $\beta$ and $\alpha_1, \ldots, \alpha_N$.

As we already saw in examples (13) and (14) of admissible vector fields $v[\mu]$ and source terms $N(t, x, \mu)$, we can verify that admissible functions $a_k$, $k = 1, 2$, are given by

$$a_k(x, \mu) := A_k \left( x, \int_{\mathbb{R}^d} \eta_k(y) d\mu(y) \right) \quad (17)$$

for some bounded and globally Lipschitz functions $\eta_k : \mathbb{R}^d \to \mathbb{R}$ and $A_k : \mathbb{R}^d \times \mathbb{R} \to [0, +\infty)$ such $A_1(x, t) + A_2(x, t) \leq 1$ for any $(x, t) \in \mathbb{R}^d \times \mathbb{R}$.
Except for the last section which contains comments on extension of our results, the rest of the paper is devoted to the proof of Theorem 2.1. In particular, in Section 4 we establish some technical results necessary for the proofs. Specifically, we introduce a variant of the BL norm denoted by \( \| \cdot \|_{BL,I} \) and show that this norm is equivalent to the BL norm on the cone of nonnegative measures. We also establish the existence of \( \int_{[0,1]} K(x, \mu, dy) d\mu \) as a Bochner integral. We then proceed in section 5 with the existence and uniqueness of the solution \( \mu_t \) of equation (10). We finally present the proof of Theorem 2.1 in section 6.

4. Preliminary technical results. We begin this section by introducing a variant of the BL norm and showing that it is equivalent to the BL norm on the cone of nonnegative measures.

4.1. An equivalent norm on nonnegative measures. We will need the following variant of the BL norm: given an integer \( I \geq 1 \) we consider the norm

\[
\| \mu \|_{BL,I} = \sup \left\{ \int_{\mathbb{R}^d} \phi \, d\mu : \phi \in C^I(\mathbb{R}^d), \| D^\alpha \phi \|_\infty \leq 1, \ 0 \leq |\alpha| \leq I \right\}. \tag{18}
\]

This kind of norm has been considered in [21] but taking \( |\alpha| \geq 1 \).

**Proposition 4.1.** Let \( \mu_n, \mu \in \mathcal{M}_{b,+}(\mathbb{R}^d) \). Then

\[
\| \mu_n - \mu \|_{BL} \to 0 \quad \text{if and only if} \quad \| \mu_n - \mu \|_{BL,I} \to 0. \tag{19}
\]

As a consequence, a function from \([0,T]\) to \((\mathcal{M}_{b,+}(\mathbb{R}^d), \| \cdot \|_{BL})\) is continuous if and only if it is continuous from \([0,T]\) to \((\mathcal{M}_{b,+}(\mathbb{R}^d), \| \cdot \|_{BL,I})\). Moreover, consider \( \mu^\varepsilon, \mu \in C([0,T], \mathcal{M}_{b,+}(\mathbb{R}^d)) \). Then, convergence as \( \varepsilon \to 0 \) of \( \mu^\varepsilon \) to \( \mu \) in \( C([0,T], (\mathcal{M}_{b,+}(\mathbb{R}^d), \| \cdot \|_{BL})) \) is equivalent to convergence in \( C([0,T], (\mathcal{M}_{b,+}(\mathbb{R}^d), \| \cdot \|_{BL,I})) \).

The proof of (19) is inspired by [19][proof of Thm 8].

**Proof.** We first prove (19). Since \( \| \mu \|_{BL,I} \leq \| \mu \|_{BL} \) the \( \implies \) implication is obvious. Let us assume that \( \| \mu_n - \mu \|_{BL,I} \to 0 \). We have to prove that \( \mu_n \to \mu \) in the BL norm. Since \( \mu_n, \mu \geq 0 \), this is equivalent to proving that \( \mu_n \to \mu \) weak*. According to Portmanteau Theorem (see e.g., Thm 2.1. in [7]) this amount to showing that \( \liminf_n \mu_n(U) \geq \mu(U) \) for any open subset \( U \subset \mathbb{R}^d \). Let \( F_m = \{ x \in U : \text{dist}(x, \mathbb{R}^d \setminus U) \geq 1/m, |x| \leq m \}, m \in \mathbb{N} \). Notice that the sets \( F_m \) are compact and they form an increasing sequence with \( \cup_m F_m = U \). Thus \( \mu(F_m) \uparrow \mu(U) \) so that, given \( \varepsilon > 0 \), we can fix \( m \) such that \( \mu(F_m) \geq \mu(U) - \varepsilon \). Let \( \phi \in C_c^{\infty}(U) \) such that \( 0 \leq \phi \leq 1 \) and \( \phi = 1 \) on \( F_m \). Here \( \phi \) is obtained as the convolution of the characteristic function of a neighborhood of \( F_m \) contained in \( U \) and the standard mollifiers (see, e.g., Thm 1.4.1 in [25]). We have \( \| D^\alpha \phi \|_\infty \leq c_k m^k \) for \( |\alpha| \leq k \). Then, \( C_1 \phi \) is an admissible test function for \( \| \cdot \|_{BL,I} \) for some \( C_1 > 0 \). Thus,

\[
\liminf_n \mu_n(U) \geq \liminf_n \int \phi \, d\mu_n = \int \phi \, d\mu \geq \mu(F_m) \geq \mu(U) - \varepsilon.
\]

It follows that \( \liminf_n \mu_n(U) \geq \mu(U) \).

We now consider \( \mu^\varepsilon, \mu \in C([0,T], \mathcal{M}_b(\mathbb{R}^d)) \). Since \( \| \cdot \|_{BL,I} \leq \| \cdot \|_{BL} \), convergence of \( \mu^\varepsilon \) to \( \mu \) in \( C([0,T], (\mathcal{M}_b(\mathbb{R}^d), \| \cdot \|_{BL})) \) implies convergence in \( C([0,T], (\mathcal{M}_b(\mathbb{R}^d), \| \cdot \|_{BL,I})) \). Let us now assume that the convergence holds in \( C([0,T], (\mathcal{M}_b(\mathbb{R}^d), \| \cdot \|_{BL,I})) \) but not in \( C([0,T], (\mathcal{M}_b(\mathbb{R}^d), \| \cdot \|_{BL})) \). Then there exist \( \delta > 0 \), \( (t_k)_k \subset \)
Since negative measure $\rho C$, sure. Finally, for any bounded measurable function $\mu$, $N$, and $\epsilon$ such that $\mu^\epsilon \to \mu$ in $C([0,T], (M_b(R^d), ||BL,1||))$. Thus, $||\mu^\epsilon - \mu||_{BL} \to 0$ by (19).

4.2. Existence of $\int_{R^d} K(x,\mu,dy) d\mu$ as a Bochner integral. Consider a mutation kernel $K : (x,\mu) \in R^d \times M_b(R^d) \to K(x,\mu,dy) \in M_b(R^d)$ and fix a non-negative measure $\mu \in M_b(R^d)$. Using results from $[20]$[Appendix C] we will argue that the integral $\int_{R^d} K(x,\mu,dy) d\mu(x)$ exists as a Bochner integral in $M_b(R^d)$. Let $p(x) := K(x,\mu,dy)$ so that $p : R^d \to M_b(R^d)$. From assumption (K1) it follows that the map $x \in R^d \to p(x) \phi$ is measurable. Thus according to [20][Appendix C], $p$ is Bochner measurable as a map with values in $M_b(R^d)$ (the completion of $M_b(R^d)$ in the BL norm). Using (K8), the map $x \to ||p(x)||_{TV}$ is in $L^1(R^d, \mu)$. It follows from $[20]$[Appendix C] that the integral $\int_{R^d} K(x,\mu,dy) d\mu$ exists as a Bochner integral in $M_b(R^d)$. Furthermore, its total variation is less or equal to $\int_{R^d} ||K(x,\mu,dy)||_{TV} d\mu(x)$, and thus $\int_{R^d} K(x,\mu,dy) d\mu$ is in fact a bounded measure. Finally, for any bounded measurable function $\phi : R^d \to R$, $\left( \int_{R^d} K(x,\mu,dy) d\mu, \phi \right) = \int_{R^d} \left( \int_{R^d} \phi(y) K(x,\mu,dy) \right) d\mu.$

Notice that the above still hold true for a signed finite measure $\mu$ by writing $\mu = \mu^+ - \mu^-$ and defining $\int_{R^d} K(x,\mu,dy) d\mu := \int_{R^d} K(x,\mu,dy) d\mu^+ - \int_{R^d} K(x,\mu,dy) d\mu^-.$

5. Global nonnegative solution to a nonlinear transport equation and existence of $\mu^\epsilon$. Consider the following nonlinear transport equation in $R^d$: $\partial_t \mu_t + \nabla \cdot (v[\mu_t] \mu_t) = N(t,\mu_t).$ (21)

Assume that the vector-field $v : M_b(R^d) \to W^{1,\infty}(R^d)$ satisfies (V1) and (V2) above. Furthermore, assume that the source term $N : R_+ \times M_b(R^d) \to M_b(R^d)$ is continuous in $(t,\mu)$ (recall that we endow $M_b(R^d)$ with the BL norm). We also assume that there exists $C_N > 0$, and for any $R > 0$ there exists $L_{N,R} > 0$ such that the following is satisfied:

(N1') $\|N(t,\mu) - N(t,\bar{\mu})\|_{BL} \leq L_{N,R} \|\mu - \bar{\mu}\|_{BL}$ for any $t \in R$ and any $\mu, \bar{\mu} \in M_{b,R}(R^d),$

(N2') $\|N(t,\mu)\|_{TV} \leq C_N (1 + \|\mu\|_{TV})$ for any $t \in R$ and $\mu \in M_b(R^d)$.

The following result is from [3][Cor. 7.3]:

**Theorem 5.1.** Assume that the vector-field $v : M_b(R^d) \to W^{1,\infty}(R^d)$ satisfies (V1) and (V2) above, and that the source term $N : R_+ \times M_b(R^d) \to M_b(R^d)$ is continuous and satisfies (N1')-(N2') above. Then for any initial condition $\mu_0 \in [0,T]$, and $\epsilon_k \to 0$ such that $\|\mu^{\epsilon_k} - \mu_k\|_{BL} \geq \delta$. We can assume without loss of generality that there exists $t := \lim k t_k$. Then $\delta \leq \|\mu^{\epsilon_k} - \mu_k\|_{BL} + \|\mu_k - \mu_t\|_{BL}$.

Since $\|\mu_k - \mu_t\|_{BL,1} \to 0$, we have $\|\mu^{\epsilon_k} - \mu_t\|_{BL} \to 0$ by (19). Thus for large $k$, $\|\mu^{\epsilon_k} - \mu_t\|_{BL} \geq \delta/2$. (20)

On the other hand $\|\mu^{\epsilon_k} - \mu_t\|_{BL,1} \leq \|\mu^{\epsilon_k} - \mu_t\|_{BL,1} + \|\mu_k - \mu_t\|_{BL,1} \leq \max_{0 \leq s \leq T} \|\mu^{\epsilon_k} - \mu_s\|_{BL,1} + o(1) \to 0$ since $\mu^\epsilon \to \mu$ in $C([0,T], (M_b(R^d), ||BL,1||))$. Thus, $\|\mu^{\epsilon_k} - \mu_t\|_{BL} \to 0$ by (19) which contradicts (20).
$\mathcal{M}_b(\mathbb{R}^d)$ there exists a unique solution $\mu \in C([0, +\infty), \mathcal{M}_b(\mathbb{R}^d))$ to equation (21), namely

$$\partial_t \mu_t + \nabla \cdot (v[\mu_t] \mu_t) = \mathcal{N}(t, \mu_t).$$

Moreover for any $T > 0$, there exists $R_T > 0$ such that $\|\mu_t\|_{TV} \leq R_T$ for $t \in [0, T]$.

In order to obtain a nonnegative solution when $\mu_0 \in \mathcal{M}_{b,+}(\mathbb{R}^d)$ we first consider an equation of the form

$$\partial_t \mu_t + \nabla \cdot (v(t, \cdot) \mu_t) = c(t, \cdot) \mu_t + B(\mu_t) \quad t \in [0, T]$$

(22)

with initial condition $\mu_0 \in \mathcal{M}_b(\mathbb{R}^d)$, where $v : [0, T] \times \mathbb{R}^d \to \mathbb{R}^d$, $c : [0, T] \times \mathbb{R}^d \to \mathbb{R}$ and $B : \mathcal{M}_b(\mathbb{R}^d) \to \mathcal{M}_b(\mathbb{R}^d)$. We assume that

(H1) $v$ is continuous bounded in $(t, x)$ and globally Lipschitz in $x$ uniformly in $t$, i.e., there exists $C > 0$ such that $|v(t, x) - v(t, x')| \leq C|x - x'|$ for any $t \in [0, T]$ and any $x, x' \in \mathbb{R}^d$.

(H2) $c$ is continuous and there exists $C > 0$ such that $\|c(t, \cdot)\|_{W^{1, \infty}} \leq C$ for any $t \in [0, T]$.

(H3) $B$ is continuous, $B(\mu) \geq 0$ if $\mu \geq 0$, there exists $C > 0$ such that $\|B(\mu)\|_{TV} \leq C(1 + \|\mu\|_{TV})$ for any $\mu \in \mathcal{M}_b(\mathbb{R}^d)$, and for any $R > 0$ there exists $C_R > 0$ such that $\|B(\mu) - B(\tilde{\mu})\|_{BL} \leq C_R \|\mu - \tilde{\mu}\|_{BL}$ for any $\mu, \tilde{\mu} \in \mathcal{M}_{b,R}(\mathbb{R}^d)$.

**Proposition 5.1.** Assume that $v : [0, T] \times \mathbb{R}^d \to \mathbb{R}^d$, $c : [0, T] \times \mathbb{R}^d \to \mathbb{R}$ and $B : \mathcal{M}_b(\mathbb{R}^d) \to \mathcal{M}_b(\mathbb{R}^d)$ satisfy assumptions (H1)-(H3) above. Then equation (22) has a unique solution $\mu \in C([0, T], \mathcal{M}_b(\mathbb{R}^d))$. Moreover, the solution $\mu_t$ is nonnegative if $\mu_0 \geq 0$.

**Proof.** With the above assumptions it is easily seen from [3][Cor. 7.3] and [3][Prop. 5.1] that (22) has a unique solution $\tilde{\mu} \in C([0, T], \mathcal{M}_b(\mathbb{R}^d))$. Let us verify that this solution is nonnegative if $\mu_0 \geq 0$. To this end, we denote by $T_{s,t}$ the flow of $v$ (which is well-defined for any $s, t \in [0, T]$) and consider

$$h(s, t, x) = \exp \left( \int_s^t c(\tau, T_{\tau,s}(x)) \, d\tau \right).$$

Notice that

$$h(s, t, x) \leq e^{\|c\|_{\infty} T}, \quad |h(s, t', x) - h(s, t, x)| \leq C|t - t'|, \quad \text{and}$$

$$|h(s, t, x') - h(s, t, x)| \leq C|x - x'|.$$  

(23)

Then, for $\mu \in C([0, T], \mathcal{M}_b(\mathbb{R}^d))$ we define $\Gamma(\mu) \in C([0, T], \mathcal{M}_b(\mathbb{R}^d))$ by

$$\Gamma(\mu) := h(0, 0, \cdot)(T_{0,t}^\mu \mu_0) + \int_0^t h(s, t, \cdot)(T_{s,t}^\mu B(\mu_s)) \, ds.$$ 

This means that for any $\phi \in C_b(\mathbb{R}^d)$,

$$\Gamma(\mu) \phi = \int_{\mathbb{R}^d} \phi(T_{0,t}(x)) h(0, t, T_{0,t}(x)) \, d\mu_0(x)$$

$$+ \int_0^t \int_{\mathbb{R}^d} \phi(T_{s,t}(x)) h(s, t, T_{s,t}(x)) \, d(B(\mu_s))(x) \, ds.$$
and thus
\[
(\Gamma(\mu), \phi) = \int_{\mathbb{R}^d} \phi(T_0(t,x)) \exp \left( \int_0^t c(\tau, T_0(\tau,x)) d\tau \right) d\mu_0(x) + \int_0^t \int_{\mathbb{R}^d} \phi(T_s(x)) \exp \left( \int_s^t c(\tau, T_s(\tau,x)) d\tau \right) d(B(\mu_s))(x) ds.
\]

Notice that if \( \mu \in C((0,T), \mathcal{M}_b(\mathbb{R}^d)) \) is a fixed-point of \( \Gamma \) in the sense that \( \mu_t = \Gamma(\mu)_t \) for any \( t \in [0,T] \), then \( \mu \) is a solution of (22) and thus is equal to \( \bar{\mu} \), the unique solution to (22).

We look for a nonnegative fixed-point of \( \Gamma \)
\( X_S = \{ \mu \in C([0, S], \mathcal{M}_b(\mathbb{R}^d)) : \mu|_{t=0} = \mu_0, ||\mu_t||_{TV} \leq 2 ||\mu_0||_{TV}, \mu_t \geq 0, t \in [0, S] \} \)
for some \( S \), where \( X_S \) is endowed with the standard norm \( \sup_{t \in [0, S]} ||\mu_t||_{BL} \). Notice that \( X_S \) is complete as a closed subset of \( C((0,S], \mathcal{M}_b, R(\mathbb{R}^d)) \) with \( R = 2 ||\mu_0||_{TV} \). Moreover, since \( B(\mu) \geq 0 \) for \( \mu \geq 0 \) we have that \( \Gamma(\mu)|_{t \geq 0} \geq 0 \) if \( \mu_t \geq 0 \). We can then prove in a standard way that for \( S \) small enough, depending only on \( ||\mu_0||_{TV} \), \( \Gamma(X) \subset X \) and that \( \Gamma \) is a strict contraction. Thus \( \Gamma \) has a unique fixed-point \( \mu \) in \( X_S \) which is therefore equal to \( \bar{\mu} \). We can then extend the fixed-point \( \mu \) to a maximal time interval \( [0, S^*) \). If \( S^* < T \) then \( \lim_{t \to S^*} ||\mu_t||_{TV} = +\infty \). But since \( \mu_t = \bar{\mu} \) on \( [0, S^*) \) and we know that \( ||\mu_t||_{TV} \) is bounded on \( [0, T] \), we obtain a contradiction. Thus \( \bar{\mu} \), the unique solution to (22), is nonnegative.

Combining Theorem 5.1 and Proposition 5.1 we can obtain nonnegative global solution to the following equation:
\[
\partial_t \mu_t + \nabla \cdot (v|\mu_t|\mu_t) = \tilde{N}(t, \cdot, \mu_t)\mu_t + B(\mu_t) \quad t > 0 \tag{24}
\]
with initial condition \( \mu_0 \in \mathcal{M}_b(\mathbb{R}^d) \).

**Theorem 5.2.** Assume that \( v : \mathcal{M}_b(\mathbb{R}^d) \to W^{1,\infty}(\mathbb{R}^d) \) satisfies (V1)-(V2), \( \tilde{N} : [0, +\infty) \times \mathbb{R}^d \times \mathcal{M}_b(\mathbb{R}^d) \to \mathbb{R} \) is continuous and satisfies (N1)-(N2), and \( B : \mathcal{M}_b(\mathbb{R}^d) \to \mathcal{M}_b(\mathbb{R}^d) \) satisfies (H3). Then equation (24) has a unique solution \( \mu \in C([0, +\infty), \mathcal{M}_b(\mathbb{R}^d)) \) for any initial condition \( \mu_0 \in \mathcal{M}_b(\mathbb{R}^d) \). Furthermore, if \( \mu_0 \geq 0 \) then \( \mu_t \geq 0 \) for any \( t > 0 \).

**Proof.** The source term \( N(t, \cdot, \mu) := \tilde{N}(t, \cdot, \mu)\mu + B(\mu) \) is continuous in \( (t, \mu) \) and satisfies (N1') and (N2') from Theorem (5.1) above. Indeed, this is obvious for \( B \) from assumption (H3). Concerning \( \tilde{N}(t, \cdot, \mu) \), we clearly have (N2') in view of (N2). Concerning (N1'), for any \( \mu, \tilde{\mu} \in \mathcal{M}_b, R(\mathbb{R}^d) \) and any \( \phi \in W^{1,\infty}(\mathbb{R}^d) \) with \( ||\phi||_{W^{1,\infty}} \leq 1 \) we have
\[
(\tilde{N}(t, \cdot, \mu) - \tilde{N}(t, \cdot, \tilde{\mu}), \phi)
= \int_{\mathbb{R}^d} \phi(x)[\tilde{N}(t, x, \mu) - \tilde{N}(t, x, \tilde{\mu})] d\mu(x) + \int_{\mathbb{R}^d} \phi(x)\tilde{N}(t, x, \tilde{\mu}) d(\mu - \tilde{\mu})(x)
\leq ||\tilde{N}(t, \cdot, \mu) - \tilde{N}(t, \cdot, \tilde{\mu})||_{TV} ||\mu||_{TV} + ||\phi\tilde{N}(t, \cdot, \tilde{\mu})||_{W^{1,\infty}} ||\mu - \tilde{\mu}||_{BL}.
\]
In view of (N1)-(N2) we get
\[
(\tilde{N}(t, \cdot, \mu) - \tilde{N}(t, \cdot, \tilde{\mu}), \phi) \leq (L_{\tilde{N}, R} + C_{\tilde{N}})||\mu - \tilde{\mu}||_{BL}
\]
which yields (N1') after taking the supremum over all such \( \phi \). We thus obtain the existence of a unique solution \( \mu \in C((0, +\infty), \mathcal{M}_b(\mathbb{R}^d)) \) to equation (24).
We define \( v(t,x) := v[\mu_t](x) \) and \( c(t,x) := \tilde{N}(t,x,\mu_t) \). For a given \( T > 0 \), we know from Theorem 5.1 that there exists \( R_T > 0 \) such that \( \|\mu_t\|_{TV} \leq R_T \) for \( t \in [0,T] \). Then it follows from (V1)-(V2) that \( v \) satisfies (H1) and from (N1)-(N2) that \( c \) satisfies (H2). Thus, if \( \mu_0 \geq 0 \) we deduce from Proposition 5.1 that \( \mu_t \geq 0 \) for \( t \in [0,T] \). Since \( T > 0 \) is arbitrary we obtain that \( \mu_t \geq 0 \) for any \( t \geq 0 \).

We can now easily obtain the existence and uniqueness of \( \mu^\varepsilon \).

**Theorem 5.3.** Let (V1)-(V2), (N1)-(N2) and (K0),(K1),(K2),(K3),(K6) and (K8) hold. For any \( \varepsilon > 0 \) and any initial condition \( \mu_0 \in \mathcal{M}_b(\mathbb{R}^d) \) equation \((10)\) has a unique solution \( \mu^\varepsilon \in C([0,\infty),\mathcal{M}_b(\mathbb{R}^d)) \) and for any \( T > 0 \) there exists \( C_{T,\varepsilon} > 0 \) such that

\[
\|\mu_t^\varepsilon\|_{TV} \leq C_{T,\varepsilon} \quad \text{for any } t \in [0,T].
\]

Moreover, if \( \mu_0 \geq 0 \) then \( \mu_t^\varepsilon \geq 0 \) for any \( t \geq 0 \).

**Proof.** Since \( v \) satisfies (V1)-(V2) and \( \tilde{N}(t,x,\mu) \) is continuous and satisfies (N1)-(N2), we only need to check in view of Theorem 5.2 that (i) \( b(x,\mu) \) is continuous and satisfies (N1)-(N2), and that (ii) \( B(\mu) := \frac{1}{\varepsilon^2} \int_{\mathbb{R}^d} K\varepsilon(x,\mu,\cdot) d\mu(x) \) satisfies (H3).

Concerning \( b \) this is obvious in view of assumption (K2). As for \( B \), first notice that \( B(\mu) \geq 0 \) if \( \mu \geq 0 \) since \( K(x,\mu,dy) \geq 0 \) according to (K0). Moreover, for any \( \phi \) with \( \|\phi\|_\infty \leq 1 \), by recalling definition \((8)\) we obtain

\[
(\varepsilon^2 B(\mu),\phi) = \int_{\mathbb{R}^d} \left( \int_{\mathbb{R}^d} \phi(x + \varepsilon(y-x)) K(x,\mu,dy) \right) d\mu(x)
\leq \int_{\mathbb{R}^d} \|K(x,\mu,\cdot)\|_{TV} d\mu(x)
\leq C_K \|\mu\|_{TV}
\]

where we used assumption (K8) in the last equality. Thus, \( \|B(\mu)\|_{TV} \leq C_K \|\mu\|_{TV} \).

It remains to check that \( B \) is continuous and locally Lipschitz. For any \( \mu,\tilde{\mu} \in \mathcal{M}_b(\mathbb{R}^d) \) and any test-function \( \phi \) such that \( \|\phi\|_\infty + \text{Lip}(\phi) \leq 1 \) we have

\[
\varepsilon^2 (B(\mu) - B(\tilde{\mu}),\phi) = \left( \int_{\mathbb{R}^d} K\varepsilon(x,\mu,dy) d\mu(x) - \int_{\mathbb{R}^d} K\varepsilon(x,\tilde{\mu},dy) d\tilde{\mu}(x),\phi \right)
= \int_{\mathbb{R}^d} (K\varepsilon(x,\mu,dy) - K\varepsilon(x,\tilde{\mu},dy),\phi) d\mu(x)
+ \int_{\mathbb{R}^d} (K\varepsilon(x,\tilde{\mu},dy),\phi) d(\mu - \tilde{\mu})(x)
\leq C_\varepsilon \int_{\mathbb{R}^d} \|K(x,\mu,\cdot)\% 2BL d\mu(x) + C_\varepsilon \|\mu - \tilde{\mu}\|_{BL}
\]

where we used assumption (K3) in the last equality. Thus

\[
\|B(\mu) - B(\tilde{\mu})\|_{BL}
\leq \frac{1}{\varepsilon^2} \left( C_\varepsilon \int_{\mathbb{R}^d} \|K(x,\mu,\cdot) - K(x,\tilde{\mu},\cdot)\|_{BL} d\mu(x) + C_\varepsilon \|\mu - \tilde{\mu}\|_{BL} \right).
\]

Notice that for any \( x \in \mathbb{R}^d \), \( \|K(x,\mu,\cdot) - K(x,\tilde{\mu},\cdot)\|_{BL} \to 0 \) when \( \tilde{\mu} \to \mu \) by (K6) and also that \( \|K(x,\mu,\cdot) - K(x,\tilde{\mu},\cdot)\|_{BL} \leq \|K(x,\mu,\cdot)\|_{TV} + \|K(x,\tilde{\mu},\cdot)\|_{TV} \leq 2C_K \) by (K8). Thus, the integral on the right-hand side converges to 0 as \( \tilde{\mu} \to \mu \) by the Dominated Convergence Theorem. It follows that \( B \) is continuous. Moreover, if
\(\mu, \tilde{\mu} \in \mathcal{M}_{b,R}(\mathbb{R}^d)\) for some \(R > 0\) then from (26) and (K6) we get that there exists \(C_{R,c} > 0\) such that \(\|B(\mu) - B(\tilde{\mu})\|_{BL} \leq C_{R,c}\|\mu - \tilde{\mu}\|_{BL}\). Thus, \(B\) satisfies (H3).

Hence, for any \(\phi \in C^1_0(\mathbb{R}^d)\) and any \(t \geq 0\), we have

\[
\int_{\mathbb{R}^d} \phi \, d\mu_{t} = \int_{\mathbb{R}^d} \phi \, d\mu_{0} + \int_{0}^{t} \int_{\mathbb{R}^d} v[\mu_{s}^{\varepsilon}] \nabla \phi \, d\mu_{s} \, ds + \int_{0}^{t} \int_{\mathbb{R}^d} \tilde{N}(s, x, \mu_{s}^{\varepsilon}) \phi \, d\mu_{s} \, ds + \int_{0}^{t} (\tilde{N}_{\varepsilon}(\mu_{s}^{\varepsilon}), \phi) \, ds
\]  

(27)

where \(\tilde{N}_{\varepsilon}(\mu) = \frac{1}{\varepsilon^2} \left( \int_{\mathbb{R}^d} K_{\varepsilon}(x, \mu, dy) \, d\mu(x) - b(\cdot, \mu)\mu \right)\). A priori the test-function \(\phi\) must have compact support. However, as a direct consequence of (30) we can easily show that the weak formulation (27) holds for bounded functions \(\phi\) with bounded first derivative without assuming that they have compact support:

**Proposition 5.2.** The weak formulation (27) holds for any bounded \(\phi \in C^1(\mathbb{R}^d)\) with bounded first derivative.

**Proof.** Take a bounded \(\phi \in C^1(\mathbb{R}^d)\) with bounded first derivative. Given \(R > 0\) we truncate \(\phi\) considering \(\phi_R(x) = \phi(x) \rho(|x|/R)\) where \(\rho : [0, +\infty) \rightarrow [0, 1]\) is smooth, supported in \([0, 2]\) and \(\rho = 1\) in \([0, 1]\). Since \(\phi_R\) is \(C^1\) with compact support we can use it as a test-function in (27):

\[
\int_{\mathbb{R}^d} \phi(x) \rho(|x|/R) \, d\mu_{t} - \int_{\mathbb{R}^d} \phi(x) \rho(|x|/R) \, d\mu_{0} = \frac{1}{R} \int_{0}^{t} \int_{\mathbb{R}^d} \phi(x) \rho'(|x|/R) \frac{x}{|x|} v[\mu_{s}^{\varepsilon}] \, d\mu_{s} \, ds + \int_{0}^{t} \int_{\mathbb{R}^d} \rho(|x|/R) v[\mu_{s}^{\varepsilon}] \nabla \phi \, d\mu_{s} \, ds + \int_{0}^{t} \int_{\mathbb{R}^d} \tilde{N}(s, x, \mu_{s}^{\varepsilon}) \phi(x) \rho(|x|/R) \, d\mu_{s} \, ds + \int_{0}^{t} (\tilde{N}_{\varepsilon}(\mu_{s}^{\varepsilon}), \phi(\cdot) \rho(|\cdot|/R)) \, ds
\]  

(28)

We now pass to the limit \(R \rightarrow +\infty\) in each integral to drop \(\rho\). We pass to the limit in the two integrals on the left-hand side by Monotone Convergence. Using (30) and (31) we bound the first integral on the right-hand side by

\[
\frac{1}{R} \|\phi\|_{\infty} \|\rho'\|_{\infty} \int_{0}^{t} \|v[\mu_{s}^{\varepsilon}]\|_{\infty} \mu_{s}^{\varepsilon}(\mathbb{R}^d) \, ds \leq C/R \rightarrow 0.
\]

Concerning the second integral, notice that for any \(s\) we have \(|\rho(|x|/R) v[\mu_{s}^{\varepsilon}] \nabla \phi| \leq C\). Thus, the Dominated Convergence Theorem implies that \(\int_{\mathbb{R}^d} \rho(|x|/R) v[\mu_{s}^{\varepsilon}] \nabla \phi \, d\mu_{s} \rightarrow \int_{\mathbb{R}^d} v[\mu_{s}^{\varepsilon}] \nabla \phi \, d\mu_{s}\). Moreover, \(|\int_{\mathbb{R}^d} \rho(|x|/R) v[\mu_{s}^{\varepsilon}] \nabla \phi \, d\mu_{s}| \leq C \mu_{s}^{\varepsilon}(\mathbb{R}^d) \leq C\) so that we can pass to the limit in the second integral once again applying the Dominated Convergence Theorem. We treat the third and fourth integral terms on the right-hand side in the same way recalling that \(|\tilde{N}(s, x, \mu_{s}^{\varepsilon})| \leq C_N\) by (N2) and \(|K(x, \mu_{s}^{\varepsilon}, \cdot)|_{TV} \leq C_K\) for any \(x \in \mathbb{R}^d\) and any \(s \in [0, t]\) by (K8). We conclude that \(\phi\) satisfies (27).

**6. Proof of Theorem 2.1.** We denote by \(\mu_{s}^{\varepsilon} \in C([0, +\infty), \mathcal{M}_{b}(\mathbb{R}^d))\) the unique solution to equation (10) as given by Theorem 5.3. Thus, for any \(\phi \in C^1_0(\mathbb{R}^d)\) and any \(t \geq 0\) there holds
\[
\int_{\mathbb{R}^d} \phi \, d\mu_t^\varepsilon = \int_{\mathbb{R}^d} \phi \, d\mu_0 + \int_0^t \int_{\mathbb{R}^d} v[\mu_s^\varepsilon] \nabla \phi \, d\mu_s^\varepsilon \, ds \\
+ \int_0^t \int_{\mathbb{R}^d} \tilde{N}(s, x, \mu_s^\varepsilon) \phi \, d\mu_s^\varepsilon \, ds + \int_0^t (\tilde{N}_\varepsilon(\mu_s^\varepsilon), \phi) \, ds
\]
\] (29)

where \( \tilde{N}_\varepsilon(\mu) = \frac{1}{\varepsilon} \left( \int_{\mathbb{R}^d} K_\varepsilon(x, \mu, dy) \, d\mu(x) - b(\cdot, \mu) \mu \right) \).

From now on we assume that the initial condition \( \mu_0 \) is a nonnegative measure. In view of Theorem 5.3 it follows that \( \mu_t^\varepsilon \geq 0 \) for any \( t \geq 0 \).

We divide the proof into several steps.

**Step 6.1.** For any \( T > 0 \) there exists \( C_T > 0 \) independent of \( \varepsilon \) such that
\[
\|\mu_t^\varepsilon\|_{TV} \leq C_T \quad \text{for any } t \in [0, T].
\]

**Proof.** Since \( \mu_t^\varepsilon \geq 0 \) we have \( \|\mu_t^\varepsilon\|_{TV} = \int_{\mathbb{R}^d} d\mu_t^\varepsilon \). In view of Proposition 5.2 we can take \( \phi \equiv 1 \) in (29). Recalling that \( (\tilde{N}_\varepsilon(\mu), 1) = 0 \) by definition of \( b \), and using assumption (N2), we obtain
\[
\|\mu_t^\varepsilon\|_{TV} = \int_{\mathbb{R}^d} d\mu_t^\varepsilon = \int_{\mathbb{R}^d} d\mu_0 + \int_0^t \int_{\mathbb{R}^d} \tilde{N}(s, x, \mu_s^\varepsilon) \, ds \leq \|\mu_0\|_{TV} + C_N \int_0^t \|\mu_s^\varepsilon\|_{TV} \, ds.
\]

Using Gronwall lemma we deduce
\[
\|\mu_t^\varepsilon\|_{TV} \leq \|\mu_0\|_{TV} e^{C_T t} \leq \|\mu_0\|_{TV} e^{C_N T} =: C_T.
\]

\( \square \)

Notice that in view of (30) and assumptions (V1) and (V2), the vector fields \( v^\varepsilon(t, x) := v[\mu_t^\varepsilon](x) \) are bounded in \( W^{1,\infty}(\mathbb{R}^d) \):
\[
\|v^\varepsilon\|_{W^{1,\infty}} \leq C.
\] (31)

This fact will be used below.

**Step 6.2.** For any \( T > 0 \) there exists \( C_T > 0 \) such that for any bounded \( \phi \in C^3(\mathbb{R}^d) \) with all derivatives up to order 3 bounded and for any \( t \in [0, T] \), there holds
\[
\frac{1}{\varepsilon^2} \left( \int_{\mathbb{R}^d} K_\varepsilon(x, \mu_t^\varepsilon, dy) \, d\mu_t^\varepsilon(x) - b(\cdot, \mu_t^\varepsilon) \mu_t^\varepsilon, \phi \right) \\
- \frac{1}{2} \int_{\mathbb{R}^d} \sum_{i,j=1}^d \partial_{ij} \phi(x) d_{ij}(x, \mu_t^\varepsilon) \, d\mu_t^\varepsilon(x) \leq C_T \varepsilon \|D^3 \phi\|_{\infty}
\] (32)

where
\[
d_{ij}(x, \mu) = \int_{\mathbb{R}^d} (y_i - x_i)(y_j - x_j) K(x, \mu, dy).
\]

**Proof.** Recalling the definition of \( b \) and of \( K_\varepsilon \) given in (8) we can write
\[
\left( \int_{\mathbb{R}^d} K_\varepsilon(x, \mu_t^\varepsilon, dy) \, d\mu_t^\varepsilon(x) - b(\cdot, \mu_t^\varepsilon) \mu_t^\varepsilon, \phi \right) \\
= \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} (\phi(y) - \phi(x)) K_\varepsilon(x, \mu_t^\varepsilon, dy) \, d\mu_t^\varepsilon(x) \\
= \int_{\mathbb{R}^d} \left( \int_{\mathbb{R}^d} \phi(x + \varepsilon(y - x)) - \phi(x) \right) K(x, \mu_t^\varepsilon, dy) \, d\mu_t^\varepsilon(x).
\]
Performing a Taylor expansion we then obtain
\[
\left( \int_{\mathbb{R}^d} K_{\varepsilon}(x, \mu_\varepsilon^\tau, dy) \, d\mu_\varepsilon^\tau(x) - b(\cdot, \mu_\varepsilon^\tau) \mu_\varepsilon^\tau, \phi \right)
\]
\[
= \int_{\mathbb{R}^d} \left( \int_{\mathbb{R}^d} (\varepsilon \nabla \phi(x)(y-x) + \varepsilon^2 \frac{(y-x)^T D^2 \phi(x + \theta \varepsilon(y-x))(y-x))}{} \right) K(x, \mu_\varepsilon^\tau, dy) \, d\mu_\varepsilon^\tau(x)
\]
where \( \theta = \theta(x, y, \varepsilon) \in (0, 1) \). From (K4) we have \( \int_{\mathbb{R}^d}(y-x) \, K(x, \mu_\varepsilon^\tau, dy) = 0 \) for any \( x \in \mathbb{R}^d \), thus the first integral term on the right-hand side of (33) of order \( \varepsilon \) equals to zero. Hence, we get
\[
\frac{1}{\varepsilon} \int_{\mathbb{R}^d} K_{\varepsilon}(x, \mu_\varepsilon^\tau, dy) \, d\mu_\varepsilon^\tau(x) - b(\cdot, \mu_\varepsilon^\tau) \mu_\varepsilon^\tau, \phi
\]
\[
= \frac{1}{\varepsilon} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} (y-x)^T D^2 \phi(x + \varepsilon(y-x))(y-x) K(x, \mu_\varepsilon^\tau, dy) \, d\mu_\varepsilon^\tau(x).
\]
Thus, the left-hand side of (32) is bounded by
\[
\frac{1}{\varepsilon} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} [D^2 \phi(x + \varepsilon(y-x)) - D^2 \phi(x)][y-x]^2 K(x, \mu_\varepsilon^\tau, dy) \, d\mu_\varepsilon^\tau(x)
\]
\[
\leq \frac{\varepsilon}{2} \|D^3 \phi\|_\infty \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} (y-x)^3 K(x, \mu_\varepsilon^\tau, dy) \, d\mu_\varepsilon^\tau(x).
\]
Thus we obtain the result by invoking (30) and assumption (K5). \( \square \)

**Step 6.3.** For any \( T > 0 \) there exists a positive constant \( C_T \) independent of \( \varepsilon \) such that
\[
\int_{\mathbb{R}^d} |x| \, d\mu_\varepsilon^\tau \leq C_T \text{ for any } t \in [0, T]. \tag{34}
\]

**Proof.** We denote by \(| \cdot |_\delta\) the function \(| \cdot | \) smoothed near 0 in such a way that \(| \cdot |_\delta \leq | \cdot | \). Given \( R > 0 \) we then truncate \(| \cdot |_\delta\) considering \( \psi_{\delta,R}(x) = |x|_\delta \rho(|x|/R) \) where \( \rho : [0, +\infty) \to [0, 1] \) is smooth and compactly supported we can use it as a test-function in (29):
\[
\int_{\mathbb{R}^d} \psi_{\delta,R} \, d\mu_\varepsilon^\tau = \int_{\mathbb{R}^d} \psi_{\delta,R} \, d\mu_0 + \int_0^t \int_{\mathbb{R}^d} \tilde{N}(s, x, \mu_\varepsilon^s) \psi_{\delta,R} \, d\mu_\varepsilon^s \, ds
\]
\[
+ \int_0^t \int_{\mathbb{R}^d} v(\mu_\varepsilon^s) \nabla \psi_{\delta,R} \, d\mu_\varepsilon^s \, ds + \int_0^t (\tilde{N}(s, \mu_\varepsilon^s), \psi_{\delta,R}) \, ds. \tag{35}
\]
Using assumption (N2), \( 0 \leq \psi_{\delta,R}(x) \leq |x|_\delta \) and \( \int_{\mathbb{R}^d} |x| \, d\mu_0(x) < \infty \) we can bound the first two integrals on the right-hand side by
\[
\int_{\mathbb{R}^d} |x|_\delta \, d\mu_0 + C_N \int_0^t \left( \int_{\mathbb{R}^d} \psi_{\delta,R} \, d\mu_\varepsilon^s \right) \, ds \leq C + C_N \int_0^t \left( \int_{\mathbb{R}^d} \psi_{\delta,R} \, d\mu_\varepsilon^s \right) \, ds.
\]
Recalling that from (31) \(|v(\mu_\varepsilon^s)|_\delta \in C\) (because of (V1) and (30)), \(|\mu_\varepsilon^{\tau}\|_{TV} \leq C\) by (30) and \(|\nabla \psi_{\delta,R}(x)| \leq C_\delta\), we can bound the integral involving \( \nabla \psi_{\delta,R} \) by \( C_\delta T \).
It remains to bound the last integral on the right-hand side of (35). Using (32) we have

$$|\langle \tilde{N}_e(\mu_s^\varepsilon), \psi_{R,\delta} \rangle| \leq \frac{1}{2} \int_{\mathbb{R}^d} \sum_{i,j=1}^d |\partial_{ij} \psi_{R,\delta}(x)||d_{ij}(x, \mu_s^\varepsilon)| \, d\mu_t^\varepsilon(x) + C_T \varepsilon \|D^3 \psi_{R,\delta}\|_\infty.$$  

It follows from assumptions (K7) and (30) that

$$|d_{ij}(x, \mu_t^\varepsilon)| \leq C$$

for any $t \in [0, T]$ and any $\varepsilon$.  

We then obtain

$$|\langle \tilde{N}_e(\mu_s^\varepsilon), \psi_R \rangle| \leq C_{\delta,T}.$$  

Thus,

$$\int_{\mathbb{R}^d} \psi_{R,\delta} \, d\mu_t^\varepsilon \leq C_{\delta,T} + C_N \int_0^t \left( \int_{\mathbb{R}^d} \psi_{R,\delta} \, d\mu_s^\varepsilon \right) \, ds.$$  

Gronwall lemma gives

$$\int_{\mathbb{R}^d} \psi_{R,\delta} \, d\mu_t^\varepsilon \leq C_{\delta,T} e^{C_{\delta,T}}.$$  

Letting $R \to +\infty$ and using the Monotone Convergence theorem we obtain

$$\int_{\mathbb{R}^d} |x| \, d\mu_t^\varepsilon \leq C_{\delta,T}.$$  

Since $|x| \leq |x|_{\delta}$ we deduce the claim.

**Step 6.4.** The sequence $\mu^\varepsilon : [0, T] \to M_b(\mathbb{R}^d)$ is uniformly equicontinuous, i.e., there exists $C_T > 0$ such that for any $\varepsilon \in (0, 1]$ and any $t, t' \in [0, T],

$$\|\mu_t^\varepsilon - \mu_{t'}^\varepsilon\|_{BL,3} \leq C_T|t - t'|,$$

where the norm $\|\cdot\|_{BL,3}$ is defined in (18) with $I = 3$.

**Proof.** Without loss of generality, we assume that $t' > t$ and take a bounded test-function $\phi \in C^3(\mathbb{R}^d)$ whose derivatives up to order three are bounded. Then according to Proposition 5.2,

$$\langle \mu_t^\varepsilon - \mu_{t'}^\varepsilon, \phi \rangle = \int_t^{t'} \int_{\mathbb{R}^d} \tilde{N}(s, x, \mu_s^\varepsilon) \phi(x) \, d\mu_s^\varepsilon(x) \, ds + \int_t^{t'} \int_{\mathbb{R}^d} v[\mu_s^\varepsilon] \nabla \phi \, d\mu_s^\varepsilon(x) \, ds$$

$$+ \int_t^{t'} \langle \tilde{N}_e(\mu_s^\varepsilon), \phi \rangle \, ds.$$  

Recalling that $\|v[\mu_s^\varepsilon]\|_\infty \leq C$ from (31), the TV-bound (30) and using (32) and (36) we have

$$|\langle \mu_t^\varepsilon - \mu_{t'}^\varepsilon, \phi \rangle| \leq \tilde{C}_T \left( \||\phi||_\infty + \||\nabla \phi||_\infty + \varepsilon \|D^3 \phi\|_\infty \right)|t - t'|$$

$$+ \frac{1}{2} \sum_{i,j=1}^d \int_t^{t'} \int_{\mathbb{R}^d} |\partial_{ij} \phi(x)||d_{ij}(x, \mu_s^\varepsilon)| \, d\mu_s^\varepsilon(x)$$

$$\leq \tilde{C}_T|t - t'| \sum_{0 \leq |\alpha| \leq 3} \|D^\alpha \phi\|_\infty.$$  

Now, we obtain the result by taking the supremum over all such $\phi$ and recalling the definition (18) of the norm $\|\cdot\|_{BL,3}$.  

\[ \square \]
Notice that a sequence of measures \((\mu_n)_n\) satisfying \(\int_{\mathbb{R}^d} |x| \, d|\mu_n|(x) \leq C\) for some constant \(C\) is tight since \(\int_{|x| \geq R} |x| \, d|\mu_n| \leq C/R\). It follows that a set of the form

\[
K = \left\{ \mu \in M_{b,+}(\mathbb{R}^d) : \int_{\mathbb{R}^d} (1 + |x|) \, d\mu \leq C \right\}
\]

for some constant \(C > 0\), is compact under the BL norm (see [19]) and thus also compact under the norm \(\|\cdot\|_{BL,3}\) in view of Proposition 4.1. According to (30) and (34), and recalling that \(\mu^\varepsilon \geq 0\), we know that the sequence \(\mu^\varepsilon\) belongs to \(C([0, T], K)\). Moreover, it is uniformly equicontinuous in view of (37). It then follows from Arzela-Ascoli theorem that up to a subsequence there exists \(\mu := \lim_{\varepsilon \to 0} \mu^\varepsilon\) in \(C([0, T], K)\). Notice that the convergence holds not only in \(C([0, T], (\mathcal{M}_b(\mathbb{R}^d), \|\cdot\|_{BL,3}))\) but also in \(C([0, T], (\mathcal{M}_b(\mathbb{R}^d), \|\cdot\|_{BL}))\) by Proposition 4.1.

The last Step of the proof is

**Step 6.5.** \(\mu_t\) is a solution of

\[
\partial_t \mu_t + \nabla \cdot (v(\mu_t) \mu_t) = \dot{N}(t, \cdot, \mu_t) \mu_t + \frac{1}{2} \sum_{i,j=1}^d \partial_{ij}(\dot{d}_{ij}(\cdot, \mu_t) \mu_t) . \tag{38}
\]

**Proof.** According to (29) and (32) we have for any \(T > 0\), any \(t \in [0, T]\), any \(\varepsilon > 0\) and any \(\phi \in C_c^\infty(\mathbb{R}^d)\), that

\[
\int_{\mathbb{R}^d} \phi \, d\mu^\varepsilon_t = \int_{\mathbb{R}^d} \phi \, d\mu_0 + \int_0^t \int_{\mathbb{R}^d} \dot{N}(s, x, \mu_s^\varepsilon) \phi(x) \, d\mu_s^\varepsilon \, ds + \int_0^t \int_{\mathbb{R}^d} \nabla \phi \cdot v(\mu_s^\varepsilon) \, d\mu_s^\varepsilon \, ds \\
+ \frac{1}{2} \sum_{i,j=1}^d \int_0^t \int_{\mathbb{R}^d} d_{ij}(x, \mu_s^\varepsilon) \partial_{ij} \phi(x) \, d\mu_s^\varepsilon \, ds + O(\varepsilon)\|\phi(3)\|_{\infty} . \tag{39}
\]

Let us first check that for any \(s \in [0, T]\),

\[
\lim_{\varepsilon \to 0} \int_{\mathbb{R}^d} d_{ij}(x, \mu_s^\varepsilon) \partial_{ij} \phi(x) \, d\mu_s^\varepsilon(x) = \int_{\mathbb{R}^d} d_{ij}(x, \mu_s) \partial_{ij} \phi(x) \, d\mu_s(x) . \tag{40}
\]

Indeed

\[
\int_{\mathbb{R}^d} d_{ij}(x, \mu_s^\varepsilon) \partial_{ij} \phi(x) \, d\mu_s^\varepsilon(x) - \int_{\mathbb{R}^d} d_{ij}(x, \mu_s) \partial_{ij} \phi(x) \, d\mu_s(x) \\
= \int_{\mathbb{R}^d} d_{ij}(x, \mu_s^\varepsilon) \partial_{ij} \phi(x) \, d(\mu_s^\varepsilon - \mu_s)(x) \\
+ \int_{\mathbb{R}^d} (d_{ij}(x, \mu_s^\varepsilon) - d_{ij}(x, \mu_s)) \partial_{ij} \phi(x) \, d\mu_s(x) \\
=: \mathcal{Y}_1 + \mathcal{Y}_2 .
\]

Using assumption (K7) with TV-bound (30) we have

\[
|\mathcal{Y}_1| \leq \|\mu_s^\varepsilon - \mu_s\|_{BL} \|d_{ij}(\cdot, \mu_s^\varepsilon) \partial_{ij} \phi\|_{W^{1,\infty}} \leq C \|\mu_s^\varepsilon - \mu_s\|_{BL} \to 0 .
\]

Moreover since \(\|d_{ij}(\cdot, \mu_s^\varepsilon)\|_{\infty}, \|d_{ij}(\cdot, \mu_s)\|_{\infty} \leq C\) by (K7) and (30), and \(d_{ij}(x, \mu_s^\varepsilon) \to d_{ij}(x, \mu_s)\) as \(\varepsilon \to 0\) for any \(x \in \mathbb{R}^d\), we have \(\mathcal{Y}_2 \to 0\) by the Dominated Convergence Theorem. This proves (40).

Now using (40) and the bound \(\int_{\mathbb{R}^d} d_{ij}(x, \mu_s^\varepsilon) \partial_{ij} \phi(x) \, d\mu_s^\varepsilon(x) \leq C_T\) for any \(s \in [0, T]\), we can pass to the limit in the fourth integral on the right-hand side using
the Dominated Convergence Theorem to obtain
\[
\lim_{\varepsilon \to 0} \int_0^t \int_{\mathbb{R}^d} d_{ij}(x, \mu^\varepsilon_s) \partial_{ij} \phi(x) \, d\mu^\varepsilon_s(x) \, ds = \int_0^t \int_{\mathbb{R}^d} d_{ij}(x, \mu_s) \partial_{ij} \phi(x) \, d\mu_s(x) \, ds.
\]

We prove in the same way that
\[
\lim_{\varepsilon \to 0} \int_0^t \int_{\mathbb{R}^d} \tilde{N}(s, x, \mu^\varepsilon_s) \phi(x) \, d\mu^\varepsilon_s(x) \, ds = \int_0^t \int_{\mathbb{R}^d} \tilde{N}(s, x, \mu_s) \phi(x) \, d\mu_s(x) \, ds.
\]

We now pass to the limit in the third integral term on the right-hand side of (39). We first write it as
\[
\int_0^t \int_{\mathbb{R}^d} \nabla \phi(v[\mu^\varepsilon_s] - v[\mu_s]) \, d\mu^\varepsilon_s(x) \, ds + \int_0^t \int_{\mathbb{R}^d} \nabla \phi(v[\mu_s]) \, d\mu^\varepsilon_s(x) \, ds =: \mathcal{Y}_3 + \mathcal{Y}_4.
\]

We bound \(\mathcal{Y}_3\) using assumption (V1) by
\[
\|\nabla \phi\|_{L^1(\mathbb{R}^d)} \int_0^t \int_{\mathbb{R}^d} \|\mu^\varepsilon_s - \mu_s\|_{BL} \, d\mu^\varepsilon_s(x) \, ds \leq \|\nabla \phi\|_{L^1(\mathbb{R}^d)} \max_{0 \leq s \leq T} \|\mu^\varepsilon_s - \mu_s\|_{BL} \|\mu^\varepsilon_s\|_{TV} T \leq C \max_{0 \leq s \leq T} \|\mu^\varepsilon_s - \mu_s\|_{BL}
\]

which converges to 0 since \(\mu^\varepsilon \to \mu\) in \(C([0, T], \mathcal{M}_b(\mathbb{R}^d))\). Moreover, \(\mathcal{Y}_4\) converges to \(\int_0^t \int_{\mathbb{R}^d} \nabla \phi(v[\mu_s]) \, d\mu^\varepsilon_s(x) \, ds\) by the Dominated Convergence Theorem since 
\[
\int_{\mathbb{R}^d} \nabla \phi(v[\mu_s]) \, d\mu^\varepsilon_s(x) \leq C.
\]

Thus
\[
\lim_{\varepsilon \to 0} \int_0^t \int_{\mathbb{R}^d} \nabla \phi(v[\mu^\varepsilon_s]) \, d\mu^\varepsilon_s(x) \, ds = \int_0^t \int_{\mathbb{R}^d} \nabla \phi(v[\mu_s]) \, d\mu_s(x) \, ds.
\]

Hence, we can pass to the limit in (39) to obtain
\[
\int_{\mathbb{R}^d} \phi \, d\mu_t = \int_{\mathbb{R}^d} \phi \, d\mu_0 + \int_0^t \int_{\mathbb{R}^d} \nabla \phi v[\mu_s] \, d\mu_s(x) \, ds + \frac{1}{2} \sum_{i,j=1}^d \int_0^t \int_{\mathbb{R}^d} d_{ij} \partial_{ij} \phi \, d\mu_s(x) \, ds
\]

which is the weak form of (38).

7. Remarks and extensions. In this paper we formulated a selection-mutation model on the space of finite signed measures and we modeled the selection-mutation kernel as a family of finite nonnegative measures in the trait parameter \(x\). This allows to simultaneously treat discrete and continuous kernels under the same theory. We established the convergence of a nonlinear diffusively rescaled selection-mutation model with a nonlinear transport and source terms to a nonlinear reaction-advection-diffusion equation. To end this paper, below we provide remarks on possible extensions to the theory we presented here.

7.1. Asymmetric distribution. We can slightly generalize the results above by considering instead of \(K\) a sequence of kernels \(K^\varepsilon(x, \mu, \cdot) \in \mathcal{M}_b(\mathbb{R}^d), \varepsilon > 0,\) satisfying basically the same assumptions as \(K\) with estimates to be uniform in \(\varepsilon\) but relaxing the assumption (K4) allowing for slightly asymmetric distribution:

(K0\(\varepsilon\)) \(K^\varepsilon(x, \mu, dy) \geq 0\) whenever \(\mu \geq 0,\)

(K1\(\varepsilon\)) For any \(\phi : \mathbb{R}^d \to \mathbb{R}\) measurable and bounded and any \(\mu \in \mathcal{M}_b(\mathbb{R}^d)\) and \(\varepsilon > 0,\) the function \(x \to \int_{\mathbb{R}^d} \phi(y)K^\varepsilon(x, \mu, dy)\) is measurable.

(K2\(\varepsilon\)) the function \(b^\varepsilon : (x, \mu) \in \mathbb{R}^d \times \mathcal{M}_b(\mathbb{R}^d) \to \int_{\mathbb{R}^d} K^\varepsilon(x, \mu, dy) \in \mathbb{R}\) is continuous and satisfies
(i) there exists $C_b > 0$ such that $\|b^\varepsilon(\cdot, \mu)\|_{W^{1, \infty}} \leq C_b$ for any $\mu \in \mathcal{M}_b(\mathbb{R}^d)$ and any $\varepsilon > 0$,

(ii) for any $R > 0$ there exists $L_{b,R} > 0$ such that
\[
|b^\varepsilon(x, \mu) - b^\varepsilon(x, \tilde{\mu})| \leq L_{b,R} \|\mu - \tilde{\mu}\|_{BL} \quad \text{for any } \mu, \tilde{\mu} \in \mathcal{M}_{b,R}(\mathbb{R}^d) \text{ and any } \varepsilon > 0,
\]

(K3') For any $\phi \in W^{1, \infty}(\mathbb{R}^d)$, the function $x \to (K^\varepsilon(x, \cdot, \cdot), \phi)$ is Lipschitz with
\[
\sup_{\mu \in \mathcal{M}_b(\mathbb{R}^d), \|\phi\|_{W^{1, \infty}} \leq 1} L_{ip}(x \to (K^\varepsilon(x, \cdot, \cdot), \phi)) \leq C \quad \text{where } C \text{ is independent of } \varepsilon.
\]

(K4') the functions $V^\varepsilon(x, \mu) := \frac{1}{\varepsilon} \int_{\mathbb{R}^d} (y - x)^3 K^\varepsilon(x, \mu, dy)$ satisfy
- for any $R > 0$ there exists $C_{V,R} > 0$ such that
\[
\|V^\varepsilon(\cdot, \mu)\|_{W^{1, \infty}} \leq C_{V,R} \quad \text{for any } \mu \in \mathcal{M}_{b,R}(\mathbb{R}^d) \text{ and any } \varepsilon > 0,
\]
- there exists $V(x, \mu)$ such that for any $x \in \mathbb{R}^d$ and any nonnegative TV-bounded sequence $(\mu_\varepsilon)_\varepsilon$ converging in the BL norm to some $\mu$, we have $V^\varepsilon(x, \mu_\varepsilon) \to V(x, \mu)$.

(K5') for any $R > 0$ there exists $C_{K,R} > 0$ such that
\[
\sup_{x \in \mathbb{R}^d} \int_{\mathbb{R}^d} |y - x|^3 K^\varepsilon(x, \mu, dy) \leq C_{K,R} \quad \text{for any } \mu \in \mathcal{M}_{b,R}(\mathbb{R}^d) \text{ and any } \varepsilon > 0.
\]

(K6') for any $R > 0$ there exists $L_{K,R} > 0$ such that
\[
\|K^\varepsilon(x, \tilde{\mu}, \cdot) - K^\varepsilon(x, \mu, \cdot)\|_{BL} \leq L_{K,R} \|\tilde{\mu} - \mu\|_{BL}
\]
for any $x \in \mathbb{R}^d$, any $\mu, \tilde{\mu} \in \mathcal{M}_{b,R}(\mathbb{R}^d)$ and any $\varepsilon > 0$.

(K7') the functions $d_{ij}^\varepsilon(x, \mu) = \int_{\mathbb{R}^d} (y_i - x_i)(y_j - x_j) K^\varepsilon(x, \mu, dy)$ satisfy
- for any $R > 0$ there exists $L_{d,R} > 0$ such that
\[
\|d_{ij}^\varepsilon(\cdot, \mu)\|_{W^{1, \infty}} \leq L_{d,R} \quad \text{for any } \mu \in \mathcal{M}_{b,R}(\mathbb{R}^d) \text{ and any } \varepsilon > 0.
\]
- there exists $d_{ij}(x, \mu)$ such that for any $x \in \mathbb{R}^d$ and any nonnegative TV-bounded sequence $(\mu_\varepsilon)_\varepsilon$ converging in the BL norm to some $\mu$, we have $d_{ij}^\varepsilon(x, \mu_\varepsilon) \to d_{ij}(x, \mu)$.

(K8') There exists $C_K > 0$ such that
\[
\|K^\varepsilon(x, \cdot, \cdot)\|_{TV} \leq C_K \quad \text{for any } x \in \mathbb{R}^d, \text{ any } \mu \in \mathcal{M}_b(\mathbb{R}^d) \text{ and any } \varepsilon > 0.
\]

The main difference with the setting of the previous section lies in assumption (K4) where we do not assume anymore that $\int_{\mathbb{R}^d} (y - x) K(x, dy) = 0$, i.e., the offspring is distributed symmetrically around $x$, but instead that this distribution could be slightly asymmetric with an asymmetry of order $\varepsilon$.

We rescale as before the kernels $K^\varepsilon(x, \mu, dy)$ defining $K^\varepsilon(x, \mu, dy)$ by
\[
(K^\varepsilon(x, \mu, dy), \phi) = \int_{\mathbb{R}^d} \phi(x + \varepsilon(y - x)) K^\varepsilon(x, \mu, dy).
\]

We then obtain as before a unique nonnegative global solution $\mu^\varepsilon_t$ to
\[
\partial_t \mu_t + \nabla \cdot (V[\mu_t] \mu_t) = \tilde{N}(t, \cdot, \mu_t) \mu_t + \frac{1}{\varepsilon^2} \left( \int_{\mathbb{R}^d} K^\varepsilon(z, \cdot) d\mu_t(z) - b^\varepsilon \mu_t \right). \tag{41}
\]

We still assume that the vector field $v$ satisfies (V1)-(V2) and that the source term $\tilde{N}$ satisfies (N1)-(N2).

The asymmetry of the reproduction kernel $K^\varepsilon$ quantified by $V^\varepsilon$ in assumption (K4') is reflected by an additional transport term $\nabla \cdot (V \mu_t)$ in the limit equation:
Theorem 7.1. For any nonnegative initial condition $\mu_0 \in M_b(\mathbb{R}^d)$ such that $\int_{\mathbb{R}^d} |x| \, d\mu_0 < \infty$, the $\mu^\varepsilon_t$ converge up to a subsequence as $\varepsilon \to 0$ in $C([0,T], M_b(\mathbb{R}^d))$ for any $T > 0$ to some $\mu \in C([0, +\infty), M_b(\mathbb{R}^d))$ which solves

$$\partial_t \mu_t + \nabla \cdot (v[\mu_t] \mu_t) + \nabla \cdot (V(\cdot, \mu_t) \mu_t) = \tilde{N}(t, \cdot, \mu_t) \mu_t + \frac{1}{2} \partial_{ij} (d_{ij}(\cdot, \mu_t) \mu_t)$$

(42)

with initial condition $\mu_{t=0} = \mu_0$ and where $V$ and $d_{ij}$ are defined in $(K\xi^\varepsilon)$ and $(K\eta^\varepsilon)$ respectively.

Proof. The proof is an adaptation of the proof of Theorem 2.1. We only highlight the main changes. Step 6.1 remains unchanged. Step 6.2 must be slightly modified replacing (32) by

$$\left| \frac{1}{\varepsilon^2} \left( \int_{\mathbb{R}^d} K^\varepsilon(x, \mu_t^\varepsilon, dy) \, d\mu_t^\varepsilon(x) - b^\varepsilon(\cdot, \mu_t^\varepsilon) \mu_t^\varepsilon, \phi \right) 
- \int_{\mathbb{R}^d} V^\varepsilon(x, \mu_t^\varepsilon) \nabla \phi \, d\mu_t^\varepsilon(x) - \frac{1}{2} \int_{\mathbb{R}^d} \partial_{ij} \phi(x) d_{ij}^\varepsilon(x, \mu_t^\varepsilon) \, d\mu_t^\varepsilon(x) \right|$$

\leq C_T \varepsilon \|D^3 \phi\|_\infty.

(43)

Indeed, the Taylor expansion (33) can be written now as

$$\frac{1}{\varepsilon^2} \left( \int_{\mathbb{R}^d} K^\varepsilon(x, \mu_t^\varepsilon, dy) \, d\mu_t^\varepsilon(x) - b^\varepsilon(\cdot, \mu_t^\varepsilon) \mu_t^\varepsilon, \phi \right)
= \int_{\mathbb{R}^d} \left( \int_{\mathbb{R}^d} \left( \frac{1}{\varepsilon} \nabla \phi(y) (y - x)
+ \frac{1}{2} (D^2 \phi(x + \theta \varepsilon (y - x)) (y - x, y - x) \right) K^\varepsilon(x, \mu_t^\varepsilon, dy) \right) \, d\mu_t^\varepsilon(x)
= \int_{\mathbb{R}^d} \nabla \phi(x) V^\varepsilon(x, \mu_t^\varepsilon) \, d\mu_t^\varepsilon(x)
+ \frac{1}{2} \int_{\mathbb{R}^d} \left( \int_{\mathbb{R}^d} (D^2 \phi(x + \theta \varepsilon (y - x)) (y - x, y - x) \right) K^\varepsilon(x, \mu_t^\varepsilon, dy) \right) \, d\mu_t^\varepsilon(x).$$

Step 6.3 and Step 6.4 remain unchanged since $|d_{ij}^\varepsilon(x, \mu_t^\varepsilon)|$ is bounded uniformly in $\varepsilon$ by assumption $(K7^\varepsilon)$.

The other change is in Step 6.5 of the proof when proving that $\mu_t := \lim_{\varepsilon \to 0} \mu_t^\varepsilon$ solves equation (42). Indeed we have to pass to the limit in the additional term and prove that

$$\int_0^t \int_{\mathbb{R}^d} \nabla \phi(x) V_t^\varepsilon(x, \mu_s^\varepsilon) \, d\mu_s^\varepsilon(x) \, ds \to \int_0^t \int_{\mathbb{R}^d} \nabla \phi(x) V(x, \mu_s) \, d\mu_s(x) \, ds.$$  \hspace{1cm} (44)

To this end, we write

$$\int_0^t \int_{\mathbb{R}^d} \nabla \phi(x) V_t^\varepsilon(x, \mu_s^\varepsilon) \, d\mu_s^\varepsilon(x) \, ds = \int_0^t \int_{\mathbb{R}^d} \nabla \phi(x) V_t^\varepsilon(x, \mu_s) \, d\mu_s(x) \, ds
- \int_0^t \int_{\mathbb{R}^d} \nabla \phi(x) V(x, \mu_s) \, d\mu_s(x) \, ds
- \int_0^t \int_{\mathbb{R}^d} \nabla \phi(x) (V_t^\varepsilon(x, \mu_s) - V(x, \mu_s)) \, d\mu_s(x) \, ds$$

$$=: \mathcal{Y}_5 + \mathcal{Y}_6.$$
We have for any \( s \in [0,T] \) that
\[
\left| \int_{\mathbb{R}^d} \nabla \phi(x) V^\varepsilon(x, \mu^\varepsilon_s) d(\mu^\varepsilon_s - \mu_s)_s(x) \right| \leq \|\mu^\varepsilon_s - \mu_s\|_{BL} \|\nabla \phi V^\varepsilon(\cdot, \mu^\varepsilon_s)\|_{W^{1,\infty}}
\]
\[
\leq C\|\mu^\varepsilon_s - \mu_s\|_{BL} \to 0
\]
in view of assumption (K4\(^\varepsilon\)). We deduce that \( Y_5 \to 0 \) by the Dominated Convergence Theorem. Moreover,
\[
\int_{\mathbb{R}^d} \nabla \phi(x)(V^\varepsilon(x, \mu^\varepsilon_s) - V(x, \mu_s)) d\mu_s(x) \to 0
\]
by the Dominated Convergence Theorem in view of assumption (K4\(^\varepsilon\)). We then deduce that \( Y_6 \to 0 \) again by the Dominated Convergence Theorem.

We also have to pass to the limit in the diffusion term and prove that for any \( s \in [0,T] \),
\[
\lim_{\varepsilon \to 0} \int_{\mathbb{R}^d} d^\varepsilon_{ij}(x, \mu^\varepsilon_s) \partial_{ij} \phi(x) d\mu^\varepsilon_s(x) = \int_{\mathbb{R}^d} d_{ij}(x, \mu_s) \partial_{ij} \phi(x) d\mu_s(x).
\]
(45)
This can be done exactly in the same way as (44) using (K7\(^\varepsilon\)).

As an example of admissible kernels \( K^\varepsilon \) consider
\[
K^\varepsilon(x, \mu, dy) := K(x, \mu, dy) + \varepsilon \tilde{K}(x, \mu, dy)
\]
where \( K \) satisfies (K0)-(K8) and \( \tilde{K} \) satisfies (K0)-(K8) except (K4). Then \( K^\varepsilon \) satisfies (K0\(^\varepsilon\))-(K8\(^\varepsilon\)) except (K4\(^\varepsilon\)). Concerning (K4\(^\varepsilon\)) we have
\[
V^\varepsilon(x, \mu) = V(x, \mu) = \int_{\mathbb{R}^d} (y - x) \tilde{K}(x, \mu, dy)
\]
and we must add assumptions on \( \tilde{K} \) such that \( V \) satisfies (K4\(^\varepsilon\)). Concerning (K7\(^\varepsilon\)) notice that
\[
d^\varepsilon_{ij}(x, \mu) = d_{ij}(x, \mu) + \varepsilon \int_{\mathbb{R}^d} (y_i - x_i)(y_j - x_j) \tilde{K}(x, \mu, dy),
\]
where
\[
d_{ij}(x, \mu) = \int_{\mathbb{R}^d} (y_i - x_i)(y_j - x_j) K(x, \mu, dy).
\]
Then (K7\(^\varepsilon\)) is satisfied. Indeed if \((\mu_\varepsilon)_\varepsilon\) is a nonnegative TV-bounded sequence converging in the BL norm to some \( \mu \) then for any \( x \in \mathbb{R}^d \),
\[
|d^\varepsilon_{ij}(x, \mu_\varepsilon) - d_{ij}(x, \mu)| \leq |d_{ij}(x, \mu_\varepsilon) - d_{ij}(x, \mu)| + \varepsilon \int_{\mathbb{R}^d} |y - x|^2 \tilde{K}(x, \mu_\varepsilon, dy)
\]
\[
\leq o(1) + C\varepsilon \int_{\mathbb{R}^d} (1 + |y - x|^3) \tilde{K}(x, \mu_\varepsilon, dy)
\]
\[
\leq o(1) + C\varepsilon
\]
since \( K \) satisfies (K7) and \( \tilde{K} \) satisfies (K5) and (K8).
7.2. Slower mutation rate: Hyperbolic rescaling. We still consider variable kernels $K^\varepsilon$ as in subsection 7.1. Instead of accelerating the mutation rate by $1/\varepsilon^2$, thus obtaining a diffusion equation in the limit, we only accelerate it by $1/\varepsilon$ (hyperbolic rescaling), i.e., we consider instead the equation
\[
\partial_t \mu^\varepsilon_t + \nabla \cdot (v[\mu^\varepsilon_t] \mu^\varepsilon_t) = \bar{N}(t, \cdot ; \mu_t) \mu_t + \frac{1}{\varepsilon} \left( \int K^\varepsilon(x, z) d\mu^\varepsilon_t(z) - b^\varepsilon \mu^\varepsilon_t \right).
\]

We still assume that the vector field $v$ satisfies (V1)-(V2) and that the source term $\bar{N}$ satisfies (N1)-(N2). Concerning $K^\varepsilon$, we assume (K0$^\varepsilon$)-(K6$^\varepsilon$) (we do not need (K7$^\varepsilon$) anymore) of subsection 7.1 but we replace (K4$^\varepsilon$) and (K5$^\varepsilon$) by

(K4$^\varepsilon$) the functions $V^\varepsilon(x, \mu) := \int_{\mathbb{R}^d} (y - x) K^\varepsilon(x, \mu, dy)$ satisfy
(i) for any $R > 0$ there exists $C_{V,R} > 0$ such that
\[
\|V^\varepsilon(\cdot, \mu)\|_{W^{1,\infty}} \leq C_{V,R} \quad \text{for any } \mu \in \mathcal{M}_{b,R}(\mathbb{R}^d) \text{ and any } \varepsilon > 0,
\]
(ii) there exists $V(x, \mu)$ such that for any nonnegative TV-bounded sequence $\mu_\varepsilon \to \mu$ in the BL norm and any $x \in \mathbb{R}^d$, $V^\varepsilon(x, \mu_\varepsilon) \to V(x, \mu)$.

(K5$^\varepsilon$) for any $R > 0$ there exists $C_{K,R} > 0$ such that
\[
\sup_{x \in \mathbb{R}^d} \int_{\mathbb{R}^d} |y - x|^2 K^\varepsilon(x, \mu, dy) \leq C_{K,R} \quad \text{for any } \mu \in \mathcal{M}_{b,R}(\mathbb{R}^d) \text{ and any } \varepsilon > 0.
\]

The slower mutation rate implies that in the Taylor expansion (33) only the first order term matters. We thus have to replace (32) in Step 6.2 by
\[
\left| \frac{1}{\varepsilon} \left( \int_{\mathbb{R}^d} K^\varepsilon(x, z, \mu^\varepsilon_t, dy) d\mu^\varepsilon_t(z) - b^\varepsilon(\cdot, \mu^\varepsilon_t) \mu^\varepsilon_t \right) - \int_{\mathbb{R}^d} V^\varepsilon \nabla \phi d\mu^\varepsilon_t \right| \leq C_T \varepsilon \|D^2 \phi\|_{\infty}.
\]

The other steps of the proof remain unchanged. We then obtain

Theorem 7.2. For any nonnegative initial condition $\mu_0 \in \mathcal{M}_b(\mathbb{R}^d)$ such that $\int_{\mathbb{R}^d} |x| \, d\mu_0 < \infty$, and for any $T > 0$, $\mu^\varepsilon_T$ converges as $\varepsilon \to 0$ in $C([0,T], \mathcal{M}_b(\mathbb{R}^d))$ to the unique solution $\mu \in C([0, +\infty), \mathcal{M}_b(\mathbb{R}^d))$ of the differential equation
\[
\partial_t \mu_t + \nabla \cdot (v[\mu_t] \mu_t) + \nabla \cdot (V \mu_t) = \bar{N}(t, \cdot ; \mu_t) \mu_t,
\]
with initial condition $\mu_{t=0} = \mu_0$.

Notice that the whole sequence $\mu^\varepsilon_t$ converges since equation (46) has a unique solution (see [14]).

7.3. Case of a system. We can adapt the result and the proof in a straightforward way to the case of a system like

\[
\begin{align*}
\partial_t \mu^\varepsilon_{t,1} + \nabla \cdot (v[\mu^\varepsilon_{t,1}] \mu^\varepsilon_{t,1}) &= \bar{N}^1(t, \cdot ; \mu^\varepsilon_t) \mu^\varepsilon_{t,1} + \frac{1}{\varepsilon^2} \left( \int_{\mathbb{R}^d} K^\varepsilon_1(z, \mu^\varepsilon_t, \cdot) d\mu^\varepsilon_{t,1}(z) - b^1(\cdot, \mu^\varepsilon_t) \mu^\varepsilon_{t,1} \right) \\
\partial_t \mu^\varepsilon_{t,2} + \nabla \cdot (v[\mu^\varepsilon_{t,2}] \mu^\varepsilon_{t,2}) &= \bar{N}^2(t, \cdot ; \mu^\varepsilon_t) \mu^\varepsilon_{t,2} + \frac{1}{\varepsilon^2} \left( \int_{\mathbb{R}^d} K^\varepsilon_2(z, \mu^\varepsilon_t, \cdot) d\mu^\varepsilon_{t,2}(z) - b^2(\cdot, \mu^\varepsilon_t) \mu^\varepsilon_{t,2} \right) \\
\mu^\varepsilon_{t,1} \Big|_{t=0} &= \mu^0_1, \quad \mu^\varepsilon_{t,2} \Big|_{t=0} = \mu^0_2
\end{align*}
\]

where $\mu^\varepsilon_t = (\mu^\varepsilon_{t,1}, \mu^\varepsilon_{t,2})$ and $b^k(x, \mu) = \int_{\mathbb{R}^d} K^k(x, \mu, dy)$, $k = 1, 2$.

Assume that for any $k = 1, 2$, the vector fields $v^k$ satisfy (V1)-(V2), the source term $\bar{N}^k$ satisfies (N1)-(N2), and the kernel $K^k$ satisfies (K0) through (K8), where $\mu = (\mu^1, \mu^2) \in \mathcal{M}_b(\mathbb{R}^d) \times \mathcal{M}_b(\mathbb{R}^d)$. 


First the existence and uniqueness result stated in Theorem 5.1 holds since the result in [3][Cor. 7.3] deals with systems. We then deduce existence and uniqueness of a global nonnegative solution \( \mu_\varepsilon \) to the system (47) as in Theorem 5.3. Finally, Step 6.1 to Step 6.5 hold with exactly the same proof. We thus obtain Theorem 2.1 for the system (47) namely

**Theorem 7.3.** For any nonnegative initial condition \( \mu_0^1, \mu_0^2 \in M_b(\mathbb{R}^d) \) such that \( \int_{\mathbb{R}^d} |x| d\mu_0^k(x) < \infty \), \( k = 1, 2 \), there exists a solution \( \mu \in C([0, +\infty), M_b(\mathbb{R}^d) \times M_b(\mathbb{R}^d)) \) to the system

\[
\begin{aligned}
\partial_t \mu^1_1 + \nabla \cdot (v^1 [\mu^1_1] \mu^1_1) &= \mathbb{N}^1(t, \cdot, \mu^1_1) \mu^1_1 + \frac{1}{2} \sum_{i,j=1}^d \partial_{ij}(d_{ij}^1(\cdot, \mu^1_1) \mu^1_1) \\
\partial_t \mu^2_2 + \nabla \cdot (v^2 [\mu^2] \mu^2_2) &= \mathbb{N}^2(t, \cdot, \mu^2_2) \mu^2_2 + \frac{1}{2} \sum_{i,j=1}^d \partial_{ij}(d_{ij}^2(\cdot, \mu^2_2)) \\
\mu^1_{|t=0} &= \mu^1_0, \quad \mu^2_{|t=0} = \mu^2_0
\end{aligned}
\]  

with
\[
d_{ij}^k(x, \mu) = \int_{\mathbb{R}^d} (y - x_i)(y - x_j) K^k(x, \mu, dy), \quad 1 \leq i, j \leq d, \quad k = 1, 2
\]

for any \( x \in \mathbb{R}^d \) and any \( \mu = (\mu^1, \mu^2) \in M_b(\mathbb{R}^d) \times M_b(\mathbb{R}^d) \).

This solution is obtained as the limit as \( \varepsilon \to 0 \) (up to a subsequence) of the unique solution \( \mu^k_\varepsilon \) of the system (47) in the sense that \( \mu^k_\varepsilon \to \mu^k \) in \( C([0, T], M_b(\mathbb{R}^d)) \) for any \( T > 0 \) and \( k = 1, 2 \). Moreover for any \( T > 0 \) there exist \( C_T > 0 \) such that
\[
\int_{\mathbb{R}^d} (1 + |x|) d\mu_t^k \leq C_T \quad 0 \leq t \leq T, \quad k = 1, 2
\]

**Acknowledgments.** The authors would like to thank two anonymous referees for useful suggestions that led to an improved version of the manuscript.

**REFERENCES**

[1] A. S. Ackleh, J. Cleveland and H. Thieme, Population dynamics under selection and mutation: Long-time behavior of differential equations in measure spaces, *Journal of Differential Equations*, **261** (2016), 1472–1505.

[2] 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 and Continuous Dynamical Systems Series B*, **5** (2005), 917–928.

[3] A. S. Ackleh and N. Saintier, Well-posedness for a system of transport and diffusion equations in measure spaces, *Journal of Mathematical Analysis and Applications*, in revision.

[4] A. S. Ackleh, B. L. Ma and P. L. Salceanu, Persistence and global stability in a selection-mutation size-structured model, *Journal of Biological Dynamics*, **5** (2011), 436–453.

[5] L. Almeida, R. H. Chisholm, J. Clairambault, T. Lorenzi, A. Lorz and C. Poucho, Why is evolution important in cancer and what mathematics should be used to treat cancer? Focus on drug resistance, *Trends in Biomathematics: Modeling, Optimization and Computational Problems, Springer*, Cham, (2018), 107–120.

[6] F. Andreu-Vailló, J. M. Mazón, J. D. Rossi and J. J. Toledo-Melero, Nonlocal Diffusion Problems, Mathematical Surveys and Monographs, 165, American Mathematical Society, Providence, RI, Real Sociedad Matemática Española, Madrid, 2010.

[7] P. Billingsley, *Convergence of Probability Measures*, Second edition, Wiley Series in Probability and Statistics: Probability and Statistics, A Wiley-Interscience Publication, John Wiley & Sons, Inc., New York, 1999.

[8] R. Burger and I. M. Bomze, Stationary distributions under mutation-selection balance: Structure and properties, *Advances in Applied Probability*, **28** (1996), 227–251.

[9] Á. Calsina and S. Cuadrado, Small mutation rate and evolutionarily stable strategies in infinite dimensional adaptive dynamics, *Journal of Mathematical Biology*, **48** (2004), 135–159.
Á. Calsina, S. Cuadrado, L. Desvillettes and G. Raoul, Asymptotics of steady states of a selection mutation equation for small mutation rate, Proc. Roy. Soc. Edinburgh Sect. A, 143 ((2013), 1123–1146.

Á. Calsina, S. Cuadrado, L. Desvillettes and G. Raoul, Asymptotic profile in selection-mutation equations: Gauss versus Cauchy distributions, Journal of Mathematical Analysis and Applications, 444 (2016), 1515–1541.

Á. Calsina and J. M. Palmada, Steady states of a selection-mutation model for an age structured population, J. Math. Anal. Appl., 400 (2013), 386–395.

J. A. Cañizo, J. A. Carrillo and S. Cuadrado, Measure solutions for some models in population dynamics, Acta Applicandae Mathematicae, 123 (2013), 141–156.

J. A. Cañizo, J. A. Carrillo and J. Rosado, A well-posedness theory in measures for some kinetic models of collective motion, Mathematical Models and Methods in Applied Sciences, 21 (2011), 515–539.

R. H. Chisholma, T. Lorenzi and J. Clairambault, Cell population heterogeneity and evolution towards drug resistance in cancer: Biological and mathematical assessment, theoretical, treatment optimisation, Biochimica et Biophysica Acta, 1860 (2016).

J. Cleveland and A. S. Ackleh, Evolutionary game theory on measure spaces: Well-Posedness, Nonlinear Anal. Real World Appl., 14 (2013), 785–797.

R. Cressman and J. Hofbauer, Measure dynamics on a one-dimensional continuous trait space: Theoretical foundations for adaptive dynamics, Theoretical Population Biology, 67 (2005), 47–59.

R. M. Dudley, Convergence of Baire measures, Studia Mathematica, 27 (1966), 251–268.

J. H. M. Evers, S. C. Hille and A. Muntean, Mild solutions to a measure-valued mass evolution problem with flux boundary conditions, Journal of Differential Equations, 259 (2015), 1068–1097.

G. Gabetta, G. Toscani and B. Wennberg, Metrics for probability distributions and the trend to equilibrium for solutions of the Boltzmann equation, Journal of Statistical Physics, 81 (1995), 901–934.

P. Gwiazda, A. Marciniak Czochra and H. R. Thieme, Measures under the flat norm as ordered normed vector space, Positivity, 22 (2018), 105–138.

P. Gwiazda, T. Lorenz and A. Marciniak Czochra, A nonlinear structured population model: Lipschitz continuity of measure-valued solutions with respect to model ingredients, Journal of Differential Equations, 248 (2010), 2703–2735.

S. C. Hille, T. Szarek, D. T. H. Worm and M. A. Ziemlanska, On a Schur-like property for spaces of measures, http://arxiv.org/pdf/1703.00677.pdf.

L. Hormander, The Analysis of Linear Partial Differential Operators. I. Distribution Theory and Fourier Analysis, Second edition, Grundlehren der Mathematischen Wissenschaften, 256. Springer-Verlag, Berlin, 1990.

P. Magal and G. F. Webb, Mutation, selection, and recombination in a model of phenotype evolution, Discrete and Continuous Dynamical Systems, 6 (2000), 221–236.

H. G. Othmer and T. Hillen, The diffusion limit of transport equations. II: Chemotaxis equations, SIAM Journal on Applied Mathematics, 62 (2002), 1222–1250.

B. Perthame, Parabolic Equations in Biology: Growth, Reaction, Movement and Diffusion, Lecture Notes on Mathematical Modelling in the Life Sciences, Springer, Cham, 2015.

C. Villani, Topics in Optimal Transportation, Graduate Studies in Mathematics, 58. American Mathematical Society, Providence, RI, 2003.

Received September 2019; revised February 2020.

E-mail address: azmy.ackleh@louisiana.edu
E-mail address: nsaintie@dm.uba.ar