MEASURE SOLUTIONS TO A SYSTEM OF CONTINUITY EQUATIONS DRIVEN BY NEWTONIAN NONLOCAL INTERACTIONS

José Antonio Carrillo

Department of Mathematics, Imperial College London
London SW7 2AZ, United Kingdom

Marco Di Francesco

DISIM - Department of Information Engineering, Computer Science and Mathematics, University of L'Aquila
Via Vetoio 1 (Coppito) 67100 L'Aquila (AQ) - Italy

Antonio Esposito

DISIM - Department of Information Engineering, Computer Science and Mathematics, University of L'Aquila
Via Vetoio 1 (Coppito) 67100 L'Aquila (AQ) - Italy

Simone Fagioli

DISIM - Department of Information Engineering, Computer Science and Mathematics, University of L'Aquila
Via Vetoio 1 (Coppito) 67100 L'Aquila (AQ) - Italy

Markus Schmidtchen

Department of Mathematics, Imperial College London
London SW7 2AZ, United Kingdom

(Communicated by the associate editor name)

Abstract. We prove global-in-time existence and uniqueness of measure solutions of a nonlocal interaction system of two species in one spatial dimension. For initial data including atomic parts we provide a notion of gradient-flow solutions in terms of the pseudo-inverses of the corresponding cumulative distribution functions, for which the system can be stated as a gradient flow on the Hilbert space $L^2(0,1)^2$ according to the classical theory by Brézis. For absolutely continuous initial data we construct solutions using a minimising movement scheme in the set of probability measures. In addition we show that the scheme preserves finiteness of the L^m-norms for all $m \in [1,+\infty]$ and of the second moments. We then provide a characterisation of equilibria and prove that they are achieved (up to time subsequences) in the large time asymptotics. We conclude the paper constructing two examples of non-uniqueness of measure solutions emanating from the same (atomic) initial datum, showing that the notion of gradient flow solution is necessary to single out a unique measure solution.

1991 *Mathematics Subject Classification.* Primary: 45K05, 35A15; Secondary: 92D25, 35L45, 35L80;

Key words and phrases. systems of aggregation equations; Newtonian potentials; uniqueness of solutions; gradient flows; long time asymptotics.

* Corresponding author: José A. Carrillo.

1. Introduction

In this work we consider a particular instance of the following nonlocal interaction system for the evolution of two probability measures ρ and η on the whole real line

$$\begin{cases} \dfrac{\partial \rho}{\partial t} = \dfrac{\partial}{\partial x}\big(\rho H_1' \star \rho + \rho K_1' \star \eta\big), \\ \dfrac{\partial \eta}{\partial t} = \dfrac{\partial}{\partial x}\big(\eta H_2' \star \eta + \eta K_2' \star \rho\big). \end{cases} \tag{1}$$

Here H_1, H_2 model the way any two agents of the same species interact with one another (so-called *self-interaction* potentials, or *intraspecific* interaction potentials). Respectively, K_1, K_2 are called the *cross-interaction* potentials, or *interspecific* interaction potentials, as they describe the interaction between any two agents of opposing species.

This model can be easily understood as a natural extension of the well-known aggregation equation (cf. [6, 9, 25, 51, 55]) to two species. This link was first established in the paper [36] as the continuous counterpart of a system of ordinary differential equations. More precisely, for $M, N \in \mathbb{N}$ suppose $(x_i)_{i=1}^N$ and $(y_i)_{i=1}^M$ denote the locations of agents of two different species, each of them with masses $\frac{1}{N}$ and $\frac{1}{M}$ respectively. Then, assuming the velocity of any agent is given as an average of the forces exhibited by all other agents upon that agent, one gets

$$\dot{x}_i = -\frac{1}{N}\sum_{j\neq i} H_1'(x_i - x_j) - \frac{1}{M}\sum_{j} K_1'(x_i - y_j), \quad i = 1, ..., N,$$
$$\dot{y}_i = -\frac{1}{M}\sum_{j\neq i} H_2'(y_i - y_j) - \frac{1}{N}\sum_{j} K_2'(y_i - x_j), \quad i = 1, ..., M.$$

The choice of the interaction potentials depends on the application or the phenomena of interest. In particular in mathematical biology contexts the potentials are often assumed to be radial, i.e. $W(x) = w_W(|x|)$, for $W \in \{H_i, K_i \,|\, i = 1, 2\}$, i.e. they only depend on the relative distance between any two agents. An interaction potential is said to be *attractive* if $w_W'(|x|) > 0$ and *repulsive* if $w_W'(|x|) < 0$. The existence theory developed in [36] covers the case of C^1-potentials H_i, K_i, $i = 1, 2$ (with suitable growth conditions at infinity) and provides a semigroup defined in the space of probability measures with finite second moment equipped with the Wasserstein distance, in the spirit of [1]. More specifically, the JKO scheme, [43], can be adopted in the special case $K_1 = K_2 = K$. Then Eq. (1) can actually be seen as the gradient flow of the functional

$$\mathcal{F}(\rho, \eta) := \frac{1}{2}\int_{\mathbb{R}} H_1 \star \rho \, \mathrm{d}\rho + \frac{1}{2}\int_{\mathbb{R}} H_2 \star \eta \, \mathrm{d}\eta + \int_{\mathbb{R}} K \star \eta \, \mathrm{d}\rho. \tag{2}$$

In this case a slightly lower regularity needs to be required on the potentials H_1, H_2, K as long as all of them are convex up to a quadratic perturbation. Thus, uniqueness can be proven via the notion of λ-convexity along geodesics of $\mathcal{F}$, see [50, 1].

Common interaction potentials for the one species case include power laws $W(x) = |x|^p/p$, as for instance in the case of granular media models, cf. [5, 56]. Another possible choice is a combination of power laws of the form $W(x) = |x|^a/a - |x|^b/b$, for $-d < b < a$ where d is the space dimension. These potentials, featuring short-range repulsion and long-range attraction, are typically chosen in the context of swarming models, cf. [4, 7, 24, 21, 28, 41, 42, 44]. Other typical choices include characteristic functions of sets or Morse potentials

$$W(x) = -c_a \exp(-|x|/l_a) + c_r \exp(-|x|/l_r),$$

or their regularised versions $W_p(x) = -c_a \exp(-|x|^p/l_a) + c_r \exp(-|x|^p/l_r)$, where c_a, c_r and l_a, l_r denote the interaction strength and radius of the attractive (resp. repulsive) part and $p \geq 2$, cf. [28, 29, 38]. These potentials display a decaying interaction strength, *e.g.* accounting for biological limitations of visual, acoustic or olfactory sense. The asymptotic behaviour of solutions to one single equation where the repulsion is modelled by non-linear diffusion and the attraction by non-local forces has also received lots of attention in terms of qualitative properties, stationary states and metastability, see [17, 19, 20, 22, 40, 23] and the references therein, as well as its two-species counterparts, cf. e.g. [35, 21, 27, 18], and references therein.
As set out earlier we shall study a particular instance of the above system where all interactions are modelled by Newtonian potentials. More precisely, by setting $N(x) := |x|$, we consider repulsive Newtonian intraspecific interactions and attractive Newtonian interspecific interactions, i.e. we will deal with the system

$$\begin{cases} \partial_t \rho = \partial_x(-\rho N' \star \rho + \rho N' \star \eta), \\ \partial_t \eta = \partial_x(-\eta N' \star \eta + \eta N' \star \rho). \end{cases} \tag{3}$$

Following (2), there is a natural functional that can be associated to system (3), namely

$$\mathcal{F}(\rho, \eta) := -\frac{1}{2} \int_{\mathbb{R}} N \star \rho \, \mathrm{d}\rho - \frac{1}{2} \int_{\mathbb{R}} N \star \eta \, \mathrm{d}\eta + \int_{\mathbb{R}} N \star \eta \, \mathrm{d}\rho. \tag{4}$$

We mention at this stage that this choice of the functional does not fit the set of assumptions in [36], in that the (repulsive) intraspecific parts of $\mathcal{F}$ are not defined through convex potential (up to a quadratic perturbation).
The corresponding equation for one species has been attracting a lot of interest. In [6] and [9], the authors provide an L^∞ and an L^p-theory for the aggregation equation $\partial_t u + \operatorname{div}(uv) = 0$, $v = -\nabla K \star u$, with initial data in $\mathcal{P}_2(\mathbb{R}^d) \cap L^p(\mathbb{R}^d)$, where $d \geq 2$ and $\mathcal{P}_2(\mathbb{R}^d)$ denotes the set of probability measures with bounded second order moments. They consider radially symmetric kernels whose singularity is of order $|x|^\alpha$, $\alpha > 2 - d$, at the origin. In particular, the authors prove local well-posedness in $\mathcal{P}_2(\mathbb{R}^d) \cap L^p(\mathbb{R}^d)$ for $p > p_s$, where $p_s = p_s(d, \alpha)$. Moreover, when $K(x) = |x|$, the exponent $p_s = \frac{d}{d-1}$ is sharp since for any $p < \frac{d}{d-1}$ the solution instantaneously concentrates mass at the origin for initial data in $\mathcal{P}_2(\mathbb{R}^d) \cap L^p(\mathbb{R}^d)$. Global well-posedness of solutions with initial data in $\mathcal{P}_2(\mathbb{R}^d)$ was proven in [25] for a general class of potentials including in particular $K(x) = |x|$. The gradient flow structure was crucial to show a unique continuation after blow-up of solutions to the aggregation equation. Let us also mention [8] that provides a well-posedness theory of compactly supported $L^1 \cap L^\infty$-solutions for the Newtonian potentials in $d \geq 2$. The gradient flow structure introduced in [25] in the particular case of $K(x) = |x|$ in one dimension was further developed in [12], where the authors prove the equivalence of the Wasserstein gradient flow for

$$\partial_t \rho = \partial_x(\rho \partial_x W \star \rho), \qquad x \in \mathbb{R}, \quad t > 0,$$

with $W(x) = -|x|$ or $W(x) = |x|$, and the notion of entropy solution of a scalar nonlinear conservation law of Burgers-type

$$\partial_t F + \partial_x g(F) = 0, \qquad x \in \mathbb{R}, \quad t > 0,$$

where

$$g(F) = F^2 - F \qquad \text{or} \qquad g(F) = -F^2 + F,$$

in the repulsive or attractive case respectively, being $F(t,x) = \int_{-\infty}^{x} \rho(t,x)\,\mathrm{d}x$. Such a result is relevant in particular in the repulsive case, as it shows that all point particles initial data evolve into an L^1-density on $t \in (0,+\infty)$ as a simple consequence of the uniqueness of entropy solutions to the corresponding scalar conservation law, [45]. More precisely, a point particle $\rho_0 = \delta_0$ in the repulsive case corresponds to the initial datum $F_0 = \mathbf{1}_{[0,+\infty)}$ for the equation $F_t + (F^2 - F)_x = 0$, and the discontinuity can be resolved (in a weak solution sense) either via a stationary Heaviside function or through a *rarefaction wave* with time-decaying slope connecting the two states 0 and 1. As the flux $g(F) = F^2 - F$ is convex, the latter is the only admissible solution in the entropy sense (see e.g. [32]). Therefore, the equivalence result in [12] implies that the distributional derivative $\rho = \partial_x F$ is the only gradient flow solution to the repulsive aggregation equation $\rho_t = -\partial_x(\rho\partial_x(|x| \star \rho))$. Notice that such a solution satisfies $\rho(t,\cdot) \in L^\infty(\mathbb{R})$ for all $t > 0$, whereas the initial condition ρ_0 is an atomic measure.

The occurrence of such a *smoothing* effect in the one-species repulsive case suggests that similar phenomena may be observed in the two-species case, at least in one space dimension. Understanding such an issue is one of the purposes of this work. However, the equivalence to the 2×2 system of conservation laws

$$\text{(5)} \qquad \begin{cases} \partial_t F + 2(F - G)\partial_x F = 0, \\ \partial_t G + 2(G - F)\partial_x G = 0, \end{cases} \qquad F(t,x) = \int_{-\infty}^{x} \rho(t,y)\,\mathrm{d}y, \qquad G(t,x) = \int_{-\infty}^{x} \eta(t,y)\,\mathrm{d}y,$$

does not provide any useful insights in this case, as we shall discuss in detail in Section 5. We would like to stress at this stage that the persistence of an atomic part for one of the two species in (3) would make the definition of measure solutions rather difficult, as the velocity fields are given by convolutions of the solution with a discontinuous function. In the (F,G) version (5) this corresponds to the impossibility of e.g. multiplying a discontinuous function $F - G$ by an atomic measure $\partial_x F$. On the other hand, we shall see that the equivalence to $L^2(0,1)^2$-gradient flows in the pseudo-inverse formalism (see e.g. [13, 30, 12]) provides a natural way to state a suitable notion of solution for (3) with measure initial data giving rise to a *unique* solution for all times. As for the L^p-regularity of solutions, we mention here that a system similar to (3) with the addition of linear diffusion in both components was studied in the context of semiconductor device modelling, see e.g. [3] in which solutions are shown to maintain the finiteness of the L^p-norms for $p \in (1,+\infty)$. As we will show in our paper, the same holds in the one-dimensional diffusion free case (3). However, when initial data feature an atomic part, the attractive part in the functional $\mathcal{F}$ may inhibit solutions to instantaneously become L^1-densities. This may happen for instance when the two species share an atomic part at the same position initially, see Section 5 below.

Another interesting issue related to (3) is its asymptotic behaviour for large times. The one species case features total collapse, i.e. the formation of one point particle in a finite time in the attractive case with all the mass of the system, see [25], and large time decay to zero in the repulsive case (as a consequence of the results in [12] and on classical results on the large time behaviour for scalar conservation laws, see e.g. [48]). The asymptotic behaviour in the 2×2 case (3) is much more complex. Finite time concentration for smooth initial data cannot happen in the view of the non-expansiveness of the L^p-norms proven in the present work. Similarly the case for the large time decay to zero is impossible, as we will construct explicit solutions featuring a steady atomic part for all times. We shall prove that the ω-limit in a suitable topology for (3) is a subset of $\{\rho = \eta\}$, which also coincides with the minimising set of the corresponding functional $\mathcal{F}$. The rest of this paper is organised as follows:

- Section 2 contains preliminary concepts on gradient flows in Wasserstein spaces and about the one-dimensional case in particular.
- Section 3 deals with the existence and uniqueness of solutions. We first prove it in Subsection 3.1, for the notion of solution introduced in Definition 7. In Subsection 3.2 we consider the case of densities as initial conditions, more precisely in $L^m(\mathbb{R})$ for some $m \in (1, +\infty]$, and we show that a suitable notion of gradient flow solution in the Wasserstein sense (see Definition 9) can be achieved via the Jordan-Kinderlehrer-Otto scheme, which also allows to prove that the L^m-regularity is maintained. In addition a uniform-in-time control of the second moment is obtained. Moreover, we prove that our solutions also satisfy Definition 7 given the additional regularity. All the results on the absolutely continuous case are collected in Theorem 12.
- Section 4 contains a detailed study of the steady states for (3), as well as of the minimisers of (4). A characterisation of the steady states is provided in Proposition 8. A consequent result concerning the asymptotic behaviour is provided in Theorem 14.
- Section 5 describes two relevant examples of atomic initial data. In both cases, non uniqueness of weak measure solutions is shown, and the relevant gradient flow solution is singled out as well. These two examples allow to conclude interesting properties related with the occurrence or not of the smoothing effect (or lack thereof) of initial atomic parts, see Remarks 11 and 12. The link with the hyperbolic system (5) is described in detail in Subsection 5.3 leading to several open problems.

2. Preliminaries

This section is devoted to setting up the framework to show existence and uniqueness of solutions to the system

$$\begin{cases} \partial_t \rho = -\partial_x(\rho N' \star \rho) + \partial_x(\rho N' \star \eta), \\ \partial_t \eta = -\partial_x(\eta N' \star \eta) + \partial_x(\eta N' \star \rho), \end{cases} \tag{6a}$$

with Newtonian interactions, $N(x) = |x|$. Throughout the paper, $\rho = \rho(t)$ and $\eta = \eta(t)$ will be considered as time dependent curves with values on the set $\mathcal{P}(\mathbb{R})$ of probability measures on $\mathbb{R}$. System (6a) is equipped with initial data

$$\rho(0) = \rho_0, \qquad \text{and} \qquad \eta(0) = \eta_0, \tag{6b}$$

for some $\rho_0, \eta_0 \in \mathcal{P}(\mathbb{R})$. Moreover, we write $\mathcal{P}_2(\mathbb{R})$ to denote the set of probability measures with finite second moment, i.e.

$$\mathcal{P}_2(\mathbb{R}) = \{\mu \in \mathcal{P}(\mathbb{R}) \,|\, m_2(\mu) < +\infty\}, \text{ where } m_2(\mu) = \int_{\mathbb{R}} |x|^2 \, \mathrm{d}\mu(x).$$

In the following we shall use the symbol $\mathcal{P}_2^a(\mathbb{R})$ referring to elements of $\mathcal{P}_2(\mathbb{R})$ which are absolutely continuous with respect to the Lebesgue measure. Consider a measure $\mu \in \mathcal{P}(\mathbb{R})$ and a Borel map $T : \mathbb{R} \to \mathbb{R}$. We denote by $\nu = T_\# \mu$ the push-forward of μ through T, defined by

$$\nu(A) = \mu(T^{-1}(A))$$

for all Borel sets $A \subset \mathbb{R}$. We refer to T as the transport map pushing μ to ν. Next let us equip the set $\mathcal{P}_2(\mathbb{R})$ with the 2-Wasserstein distance. For any measures $\mu, \nu \in \mathcal{P}_2(\mathbb{R})$ it is defined as

$$W_2(\mu,\nu) = \left(\inf_{\gamma \in \Gamma(\mu,\nu)} \int_{\mathbb{R}^2} |x-y|^2 \, \mathrm{d}\gamma(x,y) \right)^{1/2}, \tag{7}$$

where $\Gamma(\mu,\nu)$ is the class of transport plans between μ and ν, that is,

$$\Gamma(\mu,\nu) := \{\gamma \in \mathcal{P}(\mathbb{R}^2) \,|\, \pi^1_{\#}\gamma = \mu, \, \pi^2_{\#}\gamma = \nu\},$$

where $\pi^i : \mathbb{R}\times\mathbb{R} \to \mathbb{R}$, $i = 1, 2$, denotes the projection operator on the i^{th} component of the product space $\mathbb{R}^2$. Setting $\Gamma_0(\mu,\nu)$ as the class of optimal plans, i.e. minimisers of (7), the Wasserstein distance becomes

$$W_2^2(\mu,\nu) = \int_{\mathbb{R}^2} |x-y|^2 \, \mathrm{d}\gamma(x,y),$$

for any $\gamma \in \Gamma_0(\mu,\nu)$. The set $\mathcal{P}_2(\mathbb{R})$ equipped with the 2-Wasserstein metric is a complete metric space which can be seen as a length space, see for instance [1, 54, 57, 58].

Remark 1. *Given two measures $\mu, \nu \in \mathcal{P}_2(\mathbb{R})$, by using the elementary inequality $|y|^2 \le 2|x|^2 + 2|x-y|^2$ and the above definition of the 2-Wasserstein distance, one can easily show the inequality*

$$m_2(\nu) \le 2m_2(\mu) + 2W_2^2(\mu,\nu),$$

which will be used later on.

Next we introduce the notion of the Fréchet sub-differential in the space of probability measures.

Definition 1 (Fréchet sub-differential in $\mathcal{P}_2(\mathbb{R})$)**.** *Let $\phi : \mathcal{P}_2(\mathbb{R}) \to (-\infty, +\infty]$ be a proper and lower semicontinuous functional, and let $\mu \in D(\phi) := \{\mu \in \mathcal{P}_2(\mathbb{R}) \,|\, \phi(\mu) < \infty\}$. We say that $v \in L^2(\mathbb{R};\mu)$ belongs to the* Fréchet sub-differential *at μ, denoted by $\partial\phi(\mu)$, if*

$$\phi(\nu) - \phi(\mu) \ge \inf_{\gamma \in \Gamma_0(\mu,\nu)} \int_{\mathbb{R}\times\mathbb{R}} v(x)(y-x) \, \mathrm{d}\gamma(x,y) + o(W_2(\mu,\nu)).$$

Moreover, if $\partial\phi(\mu) \neq \emptyset$ we denote by $\partial^0\phi(\mu)$ the element of minimal $L^2(\mathbb{R};\mu)$-norm in $\partial\phi(\mu)$.

This definition will play a crucial role when introducing the notion of gradient flow solutions to system (6) later on, cf. Section 3.2.

A curve $\mu : [0,1] \to \mathcal{P}_2(\mathbb{R})$ is a *constant speed geodesic* if $W_2(\mu(s),\mu(t)) = (t-s)W_2(\mu(0),\mu(1))$ for all $0 \le s \le t \le 1$. Due to [1, Theorem 7.2.2], a constant speed geodesic connecting μ and ν can be written as

$$\gamma_t = \left((1-t)\pi^1 + t\pi^2\right)_{\#} \gamma,$$

where $\gamma \in \Gamma_0(\mu,\nu)$ and thus $\mu = \gamma_0$ and $\nu = \gamma_1$. In the literature γ_t is also known as *McCann interpolation*, cf. [50]. Next, we introduce a modified notion of convexity, the so-called λ-geodesic convexity, which is of paramount importance in the study of gradient flows in the metric space $\mathcal{P}_2(\mathbb{R})$.

Definition 2 (λ-geodesic convexity)**.** *Let $\lambda \in \mathbb{R}$. A functional $\phi : \mathcal{P}_2(\mathbb{R}) \to (-\infty, +\infty]$ is said to be* λ-geodesically convex *in $\mathcal{P}_2(\mathbb{R})$ if for every $\mu, \nu \in \mathcal{P}_2(\mathbb{R})$ there exists $\gamma \in \Gamma_0(\mu,\nu)$ such that*

$$\phi(\gamma_t) \le (1-t)\phi(\mu) + t\phi(\nu) - \frac{\lambda}{2} t(1-t) W_2^2(\mu,\nu),$$

for any $t \in [0,1]$.

It is necessary to recall that the λ-geodesic convexity is strictly linked with the concept of *k-flow*.

Definition 3 (k-flow)**.** *A semigroup* $S_\phi : [0, +\infty] \times \mathcal{P}_2(\mathbb{R}) \to \mathcal{P}_2(\mathbb{R})$ *is a* k*-flow for a functional* $\phi : \mathcal{P}_2(\mathbb{R}) \to (-\infty, \infty]$ *with respect to* W_2 *if, for an arbitrary* $\mu \in \mathcal{P}_2(\mathbb{R})$*, the curve* $t \mapsto S^t_\phi \mu$ *is absolutely continuous on* $[0, +\infty[$ *and satisfies the evolution variational inequality (**E.V.I.**)*

$$\frac{1}{2}\frac{\mathrm{d}^+}{\mathrm{d}t} W_2^2(S^t_\phi \mu, \tilde{\mu}) + \frac{k}{2} W_2^2(S^t_\phi \mu, \tilde{\mu}) \le \phi(\tilde{\mu}) - \phi(S^t_\phi \mu) \tag{8}$$

for all $t > 0$*, with respect to any reference measure* $\tilde{\mu} \in \mathcal{P}_2(\mathbb{R})$ *such that* $\phi(\tilde{\mu}) < \infty$*.*

As already mentioned, the previous concepts of k-flow and λ-convexity are closely intertwined. Indeed, a $\lambda-$convex functional possesses a uniquely determined $k-$flow for $k \ge \lambda$, and if a functional possesses a $k-$flow, then it is $\lambda-$convex with $\lambda \ge k$, cf. Refs. [1, 37, 49], for further details. The notion of k-flow will be of great help in the use of the flow interchange technique, cf. Subsection 3.2.

Now, let $\mu_t \in AC([0, +\infty); \mathcal{P}_2(\mathbb{R}))$ be an absolutely continuous curve in $\mathcal{P}_2(\mathbb{R})$. We can define the metric derivative of μ_t as

$$|\mu'_t|(t) := \limsup_{h \to 0} \frac{W_2(\mu_{t+h}, \mu_t)}{|h|},$$

which is well-defined almost everywhere since μ_t is an absolutely continuous curve, cf. [1, 54].

As system (6) describes the evolution of two interacting species, it is necessary to work on the product space $\mathcal{P}_2(\mathbb{R}) \times \mathcal{P}_2(\mathbb{R})$ equipped with the 2-Wasserstein product distance, which is defined in the canonical way

$$\mathcal{W}_2^2(\gamma, \tilde{\gamma}) = W_2^2(\rho, \tilde{\rho}) + W_2^2(\eta, \tilde{\eta}),$$

for all $\gamma = (\rho, \eta), \tilde{\gamma} = (\tilde{\rho}, \tilde{\eta})$ belonging to $\mathcal{P}_2(\mathbb{R}) \times \mathcal{P}_2(\mathbb{R})$. Now, let us introduce another crucial tool in our paper. For a given $\mu \in \mathcal{P}_2(\mathbb{R})$ its cumulative distribution function is given by

$$F_\mu(x) = \mu((-\infty, x]). \tag{9}$$

Since F_μ is a non-decreasing, right continuous function such that

$$\lim_{x \to -\infty} F_\mu(x) = 0, \quad \text{and} \quad \lim_{x \to +\infty} F_\mu(x) = 1,$$

we may define the pseudo-inverse function X_μ associated to F_μ, by

$$X_\mu(s) := \inf_{x \in \mathbb{R}} \{F_\mu(x) > s\}, \tag{10}$$

for any $s \in (0, 1)$. It is easy to see that X_μ is right-continuous and non-decreasing as well. Having introduced the pseudo-inverse, let us now recall some of its important properties. First we notice that it is possible to pass from X_μ to F_μ as follows

$$F_\mu(x) = \int_0^1 \mathbf{1}_{(-\infty, x]}(X_\mu(s))\, \mathrm{d}s = |\{X_\mu(s) \le x\}|. \tag{11}$$

For any probability measure $\mu \in \mathcal{P}_2(\mathbb{R})$ and the pseudo-inverse , X_μ, associated to it, we have

$$\int_{\mathbb{R}} f(x)\, \mathrm{d}\mu(x) = \int_0^1 f(X_\mu(s))\, \mathrm{d}s, \tag{12}$$

for every bounded continuous function f. Moreover, for $\mu,\nu\in\mathcal{P}_2(\mathbb{R})$, the Hoeffding-Fréchet theorem [53, Section 3.1] allows us to represent the 2-Wasserstein distance, $W_2(\mu,\nu)$, in terms of the associated pseudo-inverse functions as

$$W_2^2(\mu,\nu)=\int_0^1 |X_\mu(s)-X_\nu(s)|^2\,\mathrm{d}s, \tag{13}$$

since the optimal plan is given by $(X_\mu(s)\otimes X_\nu(s))_\#\mathcal{L}$, where $\mathcal{L}$ is the Lebesgue measure on the interval $[0,1]$, cf. also [57, 30]. We have seen that for every $\mu\in\mathcal{P}_2(\mathbb{R})$ we can construct a non-decreasing X_μ according to (10), and by the change of variables formula (12) we also know that X_μ is square integrable. We now recall that this mapping is indeed a distance-preserving bijection between the space of probability measure with finite second moments and the convex cone of non-decreasing L^2-functions

$$\mathcal{C}:=\{f\in L^2(0,1)\,|\,f\text{ is non-decreasing}\}\subset L^2(0,1). \tag{14}$$

Proposition 1 ($\mu\mapsto X_\mu$ is an isometry)**.** *The map*

$$\Psi:\mathcal{P}_2(\mathbb{R})\ni\mu\mapsto X_\mu\in\mathcal{C}, \tag{15}$$

mapping probability measures onto the convex cone of non-decreasing L^2-functions is an isometry.

Let us introduce the notion of sub-differential for functions in $L^2(0,1)^2$.

Definition 4 (Fréchet sub-differential in $L^2(0,1)^2$)**.** *For a given proper and lower semi-continuous functional $\mathfrak{F}$ on $L^2(0,1)^2$ we say that $Z=(X,Y)\in L^2(0,1)^2$ belongs to the sub-differential of $\mathfrak{F}$ at $\tilde Z=(\tilde X,\tilde Y)\in L^2(0,1)^2$ if and only if*

$$\mathfrak{F}(R)-\mathfrak{F}(\tilde Z)\ge\int_0^1\Big[X(s)(R_1(s)-\tilde X(s))+Y(s)(R_2(s)-\tilde Y(s))\Big]\,\mathrm{d}s+o(\|R-\tilde Z\|),$$

as $\|R-\tilde Z\|\to 0$, with the notation $R=(R_1,R_2)\in L^2(0,1)^2$. The sub-differential of $\mathfrak{F}$ at $\tilde Z$ is denoted by $\partial\mathfrak{F}(\tilde Z)$, and if $\partial\mathfrak{F}(\tilde Z)\neq\emptyset$ then we denote by $\partial^0\mathfrak{F}(\tilde Z)$ the element of minimal L^2-norm of $\partial\mathfrak{F}(\tilde Z)$.

Remark 2 (Mass normalisation)**.** *In the most general situation possible, the two species, ρ and η, have different masses $M_\rho,M_\eta>0$. The change of variables*

$$\widetilde{\rho}=\frac{1}{M_\rho}\rho,\qquad\text{and}\qquad\widetilde{\eta}=\frac{1}{M_\eta}\eta,$$

allows to rewrite system (6a) (by dropping the tildes) as

$$\begin{cases}\partial_t\rho=-M_\rho\partial_x(\rho N'\star\rho)+M_\eta\partial_x(\rho N'\star\eta),\\ \partial_t\eta=-M_\eta\partial_x(\eta N'\star\eta)+M_\rho\partial_x(\eta N'\star\rho),\end{cases}$$

and the gradient flow structure in the product Wasserstein metric $\mathcal{W}_2$ would be lost because the two interspecific potentials are different. However, this problem can be overcome by using a weighted version of the $\mathcal{W}_2$ product distance of the form

$$\mathcal{W}_2^2((\rho,\eta),(\widetilde{\rho},\widetilde{\eta}))=W_2^2(\rho,\widetilde{\rho})+\frac{M_\eta}{M_\rho}W_2^2(\eta,\widetilde{\eta}),$$

as done in [36]. As these multiplying constants M_ρ,M_η do not bring significant technical difficulties (while making the notation much heavier), for the sake of convenience we shall assume throughout the whole paper that $M_\rho=M_\eta=1$, unless specified otherwise.

3. Existence and Uniqueness

In this section we provide the mathematical theory for system (6). In the first subsection we will deal with the case of general probability measures in $\mathcal{P}_2(\mathbb{R}) \times \mathcal{P}_2(\mathbb{R})$ as initial conditions. In the second subsection we shall restrict ourselves to the case of measures that are absolutely continuous with respect to the Lebesgue measure. In the former we will provide a notion of solutions that is linked to the concept of gradient flows in Hilbert space *a-la* Brézis [15], working with the pseudo-inverse formulation of the problem. In the latter, a better regularity can be achieved and the theory is developed in the framework of gradient flows in Wasserstein space, [1]. Before entering the details, let us recall the definition of *interaction energy functional* $\mathcal{F}$ in (4): for all $(\rho, \eta) \in \mathcal{P}_2(\mathbb{R}) \times \mathcal{P}_2(\mathbb{R})$ we set

$$\mathcal{F}(\rho, \eta) := -\frac{1}{2}\int_{\mathbb{R}} N \star \rho \, \mathrm{d}\rho - \frac{1}{2}\int_{\mathbb{R}} N \star \eta \, \mathrm{d}\eta + \int_{\mathbb{R}} N \star \eta \, \mathrm{d}\rho,$$

which is well defined due to the control on the second order moment.

3.1. **General Measures Initial Data.**

In this first subsection we will use the concept of L^2-gradient flow by studying system (1) in terms of the pseudo-inverse functions X_ρ and X_η defined in Section 2. Throughout the rest of this section we set $X := X_\rho$ and $Y := X_\eta$ to simplify the notation. Hence, system (6) (formally) becomes

$$\begin{cases} \dfrac{\partial X}{\partial t} = \displaystyle\int_0^1 \mathrm{sign}(X(z) - X(\xi))\, \mathrm{d}\xi - \int_0^1 \mathrm{sign}(X(z) - Y(\xi))\, \mathrm{d}\xi, \\ \dfrac{\partial Y}{\partial t} = \displaystyle\int_0^1 \mathrm{sign}(Y(z) - Y(\xi))\, \mathrm{d}\xi - \int_0^1 \mathrm{sign}(Y(z) - X(\xi))\, \mathrm{d}\xi, \end{cases} \tag{16}$$

for $s \in (0, 1)$ and $t \geq 0$, cf. [11, 30, 47] for similar computations. In order to give a meaning to the above system in the case of μ or η having atoms, we use the convention $\mathrm{sign}(0) = 0$. In terms of the pseudo inverses X and Y, the functional $\mathcal{F}(\rho, \eta)$ becomes

$$\begin{aligned} \mathcal{F}(\rho, \eta) = \mathfrak{F}(X, Y) = &- \frac{1}{2}\int_0^1 \int_0^1 |X(z) - X(\xi)| \, \mathrm{d}z \, \mathrm{d}\xi - \frac{1}{2}\int_0^1 \int_0^1 |Y(z) - Y(\xi)| \, \mathrm{d}z \, \mathrm{d}\xi \\ &+ \int_0^1 \int_0^1 |X(z) - Y(\xi)| \, \mathrm{d}z \, \mathrm{d}\xi. \end{aligned} \tag{17}$$

In the remainder of this section we shall see that (16) is the $L^2 \times L^2$-gradient flow associated to (an extended version of) the energy functional (17).

Remark 3. *Later on in the paper we need to distinguish between the self-interaction part of $\mathfrak{F}$ and its cross-interaction part. Thus, let us rewrite $\mathfrak{F}$ as*

$$\mathfrak{F}(X, Y) = S(X) + S(Y) + K(X, Y),$$

where S is the energy functional arising from the self-interactions and K is associated to the cross-interaction.

Following the procedure of [12], since we are dealing with distribution of particles, we have to ensure that the flow remains in the set $\mathcal{C} \times \mathcal{C}$, with $\mathcal{C}$ defined in (14), see also [10, 13]. Hence, for $X \in L^2(0, 1)$ given, we introduce the *indicator function* of $\mathcal{C}$, defined by,

$$\mathcal{I}_{\mathcal{C}}(X) = \begin{cases} 0, & \text{if } X \in \mathcal{C}, \\ +\infty, & \text{otherwise.} \end{cases} \tag{18}$$

Thus we consider the extended functional

$$\bar{\mathfrak{F}}(X,Y) = \mathfrak{F}(X,Y) + \mathcal{I}_C(X) + \mathcal{I}_C(Y). \tag{19}$$

In [12, Proposition 2.8] the authors proved that the self-interaction part of $\mathfrak{F}$, i.e. S, is actually linear when restricted to $\mathcal{C}$. Let us recall this result in the next proposition.

Proposition 2. *Let $X \in \mathcal{C}$. Then*

$$S(X) = \int_0^1 (1-2z)X(z)\,\mathrm{d}z.$$

As a trivial consequence of Proposition 2 we have the following result.

Proposition 3. *The functional $\tilde{\bar{\mathfrak{F}}}$ is convex on $L^2(0,1)^2$.*

Proof. The proof is trivial since K, the cross-interaction part of $\bar{\mathfrak{F}}$, is convex due to the convexity of the Newtonian potential N. Moreover, from [12, Proposition 2.9] we argue that the remaining part is convex. □

We now present the definition of L^2-gradient flow solutions to system (16).

Definition 5. *Let $(X_0, Y_0) \in \mathcal{C}\times\mathcal{C}$. An absolutely continuous curve $(X(t,\cdot), Y(t,\cdot)) \in L^2(0,1)^2$, $t \geq 0$, is a gradient flow for the functional $\bar{\mathfrak{F}}$ if $Z(t) := (X(t), Y(t))$ is a Lipschitz function on $[0,+\infty)$, i.e., $\frac{dZ}{dt} \in L^\infty(0,+\infty; L^2(0,1)^2)$ (in the sense of distributions) and if it satisfies the sub-differential inclusion*

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix} X(t,\cdot) \\ Y(t,\cdot)\end{pmatrix} \in -\partial\bar{\mathfrak{F}}\big[\big(X(t,\cdot), Y(t,\cdot)\big)\big] \tag{20}$$

for every $t > 0$ with $(X(0,\cdot), Y(0,\cdot)) = (X_0(\cdot), Y_0(\cdot))$.

We observe that the assumption $(X_0, Y_0) \in \mathcal{C}\times\mathcal{C}$ is natural as X_0 (respectively Y_0) is the pseudo-inverse of the cumulative distribution of the initial measure ρ_0 (respectively η_0). We also observe that this assumption easily implies $\partial\bar{\mathfrak{F}}\big[\big(X_0, Y_0\big)\big] \neq \emptyset$.

Remark 4. *The gradient flow notion defined in Definition 5 is taken from the book* [15, Theorem 3.1]. *Actually, in* [15, Theorem 3.1] *the following extra condition is required, namely*

$$\left\| \frac{\mathrm{d}Z}{\mathrm{d}t} \right\|_{L^\infty((0,+\infty);L^2(0,1)^2)} \leq \left\| \partial^0 \bar{\mathfrak{F}}\big[\big(X_0, Y_0\big)\big] \right\|_{L^2(0,1)^2}. \tag{21}$$

According to [15, Theorem 3.1] *a solution in the sense of Definition 5 in conjunction with* (21) *directly verifies the following properties:*

(1) Z admits a right derivative for every $t \in [0,+\infty)$ and

$$\frac{d^+Z}{dt}(t) = -\partial^0\bar{\mathfrak{F}}\left[Z(t)\right],$$

for every $t \in [0,+\infty)$;

(2) the function $t \mapsto \partial^0\bar{\mathfrak{F}}\left[Z(t)\right]$ is right continuous and the function $t \mapsto \left\|\partial^0\bar{\mathfrak{F}}\left[Z(t)\right]\right\|_{L^2(0,1)^2}$ is non-increasing;

(3) if $Z_{1,t} := (X_1(t,\cdot), Y_1(t,\cdot))$ and $Z_{2,t} := (X_2(t,\cdot), Y_2(t,\cdot))$ are two solutions to system (20), then there holds

$$\|Z_{1,t} - Z_{2,t}\|_{L^2\times L^2} \le \|Z_{1,0} - Z_{2,0}\|_{L^2\times L^2},$$

for all $t \ge 0$.

On the other hand, [39, Section 9.6, Theorem 3] *shows that condition (21) can be avoided in order to prove uniqueness. In fact, the estimate (21) can be proven as a consequence of the properties stated in Definition 5.*

Since we know that $\bar{\mathfrak{F}}$ is a proper, lower semi-continuous, and convex functional on the Hilbert space $L^2(0,1)^2$, it easy to show that $\partial\bar{\mathfrak{F}}$ is a maximal monotone operator. Thus we can apply the theory of Brézis [15, Theorem 3.1] combined with [39, Section 9.6, Theorem 3] in order to prove existence and uniqueness of an absolutely continuous curve satisfying the differential inclusion above.

Theorem 6. *Let $(X_0, Y_0) \in \mathcal{C}\times\mathcal{C}$. There exists a unique solution $(X(t,\cdot), Y(t,\cdot))$ in the sense of Definition 5 with initial datum (X_0, Y_0).*

Now, let us go back to system (6) and state our definition of solution.

Definition 7. *Let $\gamma_0 = (\rho_0, \eta_0) \in \mathcal{P}_2(\mathbb{R})\times\mathcal{P}_2(\mathbb{R})$. An absolutely continuous curve $\gamma(t) = (\rho(t), \eta(t)) : [0,T] \to \mathcal{P}_2(\mathbb{R})\times\mathcal{P}_2(\mathbb{R})$ is a* gradient flow *solution to system (6) if the pseudo-inverses $(X(t,\cdot), Y(t,\cdot)) \in \mathcal{C}\times\mathcal{C}$ of the space cumulative distribution functions associated to $(\rho(t,\cdot), \eta(t,\cdot))$ are a solution to system (16) in the sense of Definition 5 with initial datum $(X_0, Y_0) = (X_{\rho_0}, Y_{\eta_0})$.*

According to Definition 7 the following theorem is then a consequence of the isometry (15) and Theorem 6.

Theorem 8. *Let $\gamma_0 = (\rho_0, \eta_0) \in \mathcal{P}_2(\mathbb{R})\times\mathcal{P}_2(\mathbb{R})$. There exists a unique solution to the system (6) in the sense of Definition 7.*

So far we assumed that the link between (6) and (16) is somewhat natural and we just referred to similar situations in the literature. However, the theory developed in this subsection would be somewhat meaningless if we did not show that the concept of solution in Definition 7 extends a more classical notion of solution for (6). The following subsection is dedicated to establishing exactly this link.

3.2. Absolutely Continuous Initial Data.

In this subsection we consider the case of densities as initial data. Following the approach of [1] combined with the results from [11, 12, 25, 26, 36], we pose system (6) as the gradient flow of the interaction energy functional (4) (that we recall here for the reader's convenience)

$$\mathcal{F}(\rho,\eta) = -\frac{1}{2}\int_{\mathbb{R}} N\star\rho \,\mathrm{d}\rho - \frac{1}{2}\int_{\mathbb{R}} N\star\eta \,\mathrm{d}\eta + \int_{\mathbb{R}} N\star\eta \,\mathrm{d}\rho,$$

for all $(\rho,\eta) \in \mathcal{P}_2^a(\mathbb{R})\times\mathcal{P}_2^a(\mathbb{R})$, and $N(x) = |x|$, for all $x \in \mathbb{R}$.

Definition 9. *Given any $\gamma_0 = (\rho_0, \eta_0) \in \mathcal{P}_2^a(\mathbb{R})\times\mathcal{P}_2^a(\mathbb{R})$, an absolutely continuous curve $\gamma(t) = (\rho(t), \eta(t)) : [0,T] \to \mathcal{P}_2^a(\mathbb{R})\times\mathcal{P}_2^a(\mathbb{R})$ is a gradient flow for $\mathcal{F}$ if $\rho(t)$ and $\eta(t)$ solve the following system in the distributional sense*

$$\begin{cases}\partial_t\rho(t) + \partial_x(\rho(t)v_1(t)) = 0,\\ \partial_t\eta(t) + \partial_x(\eta(t)v_2(t)) = 0,\end{cases} \tag{22}$$

with initial datum γ_0 and the velocity field $v(t) = (v_1(t), v_2(t))$ such that

$$v_i(t) = -\left(\partial^0 \mathcal{F}[\gamma(t)]\right)_i$$

for $i = 1, 2$ and

$$\|v(t)\|_{L^2(\gamma(t))} = |\gamma'|(t),$$

for a.e. $t > 0$.

Note that it is easy to check that the element of minimal norm in $\partial N(x)$ is given by

$$\partial^0 N(x) = \begin{cases} \mathrm{sign}(x), & x \neq 0, \\ 0, & x = 0. \end{cases}$$

Using the results obtained in [12, 25, 26, 36], we easily get the following proposition.

Proposition 4. *The functional $\mathcal{F}$ is λ-geodesically convex on $\mathcal{P}_2^a(\mathbb{R}) \times \mathcal{P}_2^a(\mathbb{R})$ for all $\lambda \leq 0$. Moreover, for all $\rho, \eta \in \mathcal{P}_2^a(\mathbb{R})$ the vector field*

$$\partial^0 \mathcal{F}[\rho, \eta] = \begin{pmatrix} -\partial^0 N \star \rho + \partial^0 N \star \eta \\ -\partial^0 N \star \eta + \partial^0 N \star \rho \end{pmatrix}, \tag{23}$$

is the unique element of the minimal Fréchet sub-differential of $\mathcal{F}$, where

$$\partial^0 N \star \rho(x) = \int_{\{x \neq y\}} \mathrm{sign}(x-y)\, \mathrm{d}\rho(y), \ \text{and} \ \ \partial^0 N \star \eta(x) = \int_{\{x \neq y\}} \mathrm{sign}(x-y)\, \mathrm{d}\eta(y).$$

Proof. The geodesic convexity of $\mathcal{F}$ on $\mathcal{P}_2^a(\mathbb{R}) \times \mathcal{P}_2^a(\mathbb{R})$ is the consequence of two observations. First, the cross-interaction part is geodesically convex as the interspecific interaction potential is given by $N(x) = |x|$, a convex function. Second, the geodesic convexity of the intraspecific self-interaction part can be proven using a nice monotonicity property of the transport map between two measures in $\mathcal{P}_2^a(\mathbb{R})$ in one dimension. Indeed, in [26, Lemma 1.4] the authors prove that, given $\mu, \nu \in \mathcal{P}_2^a(\mathbb{R})$, the transport map $T = T_\mu^\nu$ is essentially non-decreasing, i.e. it is non-decreasing except on a μ-null set. Hence, we can prove $\mathcal{S}[\mu] := -\frac{1}{2}\int_{\mathbb{R}} N \star \mu \, \mathrm{d}\mu$ is geodesically convex. More precisely, if T is the optimal transport map between μ and ν, then $g_t = ((1-t)\mathrm{id} + tT)_{\#}\mu$ is the geodesic connecting μ and ν. In particular, by using the mentioned monotonicity property for T, we have

$$\begin{aligned}
\mathcal{S}[g_t] &= -\frac{1}{2} \iint_{\mathbb{R}^2} |x-y| \, dg_t(y)\, dg_t(x) \\
&= -\frac{1}{2} \iint_{\mathbb{R}^2} |(1-t)(x-y) + t(T(x) - T(y))| \, d\mu(y)\, d\mu(x) \\
&= -\frac{1}{2}(1-t) \iint_{\mathbb{R}^2} |x-y| \, d\mu(y)\, d\mu(x) - \frac{1}{2} \iint_{\mathbb{R}^2} |T(x) - T(y)| \, d\mu(y)\, d\mu(x) \\
&= (1-t) S[\mu] + t S[\nu],
\end{aligned}$$

whence we get geodesic convexity for S, and for $\mathcal{F}$ as well. Obviously, 0-geodesic convexity implies λ-geodesic convexity for any $\lambda \leq 0$. In order to prove formula (23), let us notice the functional can be written as

$$\mathcal{F}[\rho, \eta] = \mathcal{S}[\rho] + \mathcal{S}[\eta] + \mathcal{K}[\rho, \eta],$$

where $\mathcal{S}$ has been introduced above and $\mathcal{K}[\rho,\eta] := \int_{\mathbb{R}} N \star \eta \,\mathrm{d}\rho$. Now, thanks to [12, Theorem 5.1] and [11, Proposition 4.3.3] we know $\partial^0 S[\rho] = -\partial^0 N \star \rho$, while [36, Proposition 3.1] yields

$$\partial^0 \mathcal{K}[\rho,\eta] = \begin{pmatrix} \partial^0 N \star \eta \\ \partial^0 N \star \rho \end{pmatrix}.$$

In particular, it is easy to check the vector field

$$\begin{pmatrix} -\partial^0 N \star \rho + \partial^0 N \star \eta \\ -\partial^0 N \star \eta + \partial^0 N \star \rho \end{pmatrix}$$

is an element of the Fréchet sub-differential of $\mathcal{F}$, and it is the unique one of minimal L^2-norm by arguing as in [25, Proposition 2.2]. □

Remark 5. *We highlight that in the presence of atomic parts for ρ or η the sub-differential may be empty, as shown in* [12].

Recall that, for $\mu, \nu \in \mathcal{P}_2^a(\mathbb{R}) \times \mathcal{P}_2^a(\mathbb{R})$, the *slope* of a functional $\mathcal{F}$ on $\mathcal{P}_2^a(\mathbb{R}) \times \mathcal{P}_2^a(\mathbb{R})$ is defined as

$$|\partial \mathcal{F}|[\mu] := \limsup_{\nu \to \mu} \frac{(\mathcal{F}(\mu) - \mathcal{F}(\nu))^+}{\mathcal{W}_2(\nu,\mu)},$$

and it can be written as

$$|\partial \mathcal{F}|[\mu] = \min\{\|\nu\|_{L^2(\mu)} \,|\, \nu \in \partial \mathcal{F}(\mu)\},$$

under certain conditions, cf. [1, Chapter 10].

Definition 10. *An absolutely continuous curve $\gamma(t) : [0,T] \to \mathcal{P}_2^a(\mathbb{R}) \times \mathcal{P}_2^a(\mathbb{R})$ is a* curve of maximal slope *for the functional $\mathcal{F}$ if the map $t \mapsto \mathcal{F}(\gamma(t))$ is an absolutely continuous function and the following inequality holds*

$$\mathcal{F}(\gamma(s)) - \mathcal{F}(\gamma(t)) \geq \frac{1}{2} \int_s^t \left[|\gamma'|^2(\tau) + |\partial \mathcal{F}|[\gamma(\tau)]^2 \right] \,\mathrm{d}\tau,$$

for all $0 \leq s \leq t \leq T$.

In order to construct a solution to system (6) in the sense of Definition 9 we follow the strategy proposed in [1] and used in [25, 36]. First, we prove the existence of a curve of maximal slope by means of the so-called "*Minimizing Movement Scheme*", cf. [1, 34], or *Jordan-Kinderlehrer-Otto scheme*, cf. [43]. Then, we prove the limit curve of the scheme is absolutely continuous w.r.t. the Lebesgue measure provided the initial datum is in $L^m(\mathbb{R}) \times L^m(\mathbb{R})$ for some $m > 1$.
Let $\tau > 0$ be a fixed time step and let $\gamma_0 = (\rho_0, \eta_0) \in \mathcal{P}_2^a(\mathbb{R}) \times \mathcal{P}_2^a(\mathbb{R})$ be a fixed initial datum such that $\mathcal{F}(\gamma_0) < +\infty$. We define a sequence $\{\gamma_\tau^n\}_{n\in\mathbb{N}} = \{(\rho_\tau^n, \eta_\tau^n)\}_{n\in\mathbb{N}}$ recursively. We set $\gamma_\tau^0 = \gamma_0$ and, for a given $\gamma_\tau^n \in \mathcal{P}_2(\mathbb{R}) \times \mathcal{P}_2(\mathbb{R})$ with $n \geq 0$, we choose γ_τ^{n+1} as

$$\gamma_\tau^{n+1} \in \operatorname{argmin}_{\gamma \in \mathcal{P}_2(\mathbb{R})^2} \left\{ \frac{1}{2\tau} \mathcal{W}_2^2(\gamma_\tau^n, \gamma) + \mathcal{F}(\gamma) \right\}. \tag{24}$$

Note that (24) is well-posed arguing as in [25, Lemma 2.3 and Proposition 2.5] for each component. Next we define the piecewise constant interpolation of the sequence $\{\gamma_\tau^n\}_{n\in\mathbb{N}}$. Let $T > 0$ be fixed and let $N := \left[\frac{T}{\tau}\right]$. We set

$$\gamma_\tau(t) = \gamma_\tau^n,$$

for $t \in ((n-1)\tau, n\tau]$. We proceed by showing that the family $\{\gamma_\tau\}_{\tau>0}$ admits a limiting curve and conclude by identifying this limit as a distributional solution to system (6). The proof in

Proposition 5 follows a, by now, classical argument of [1, Chapter 3], with only minor issues related to some moment estimates in our case. We present it here for the reader's convenience.

Proposition 5 (Narrow compactness)**.** *There exists an absolutely continuous curve $\gamma : [0,T] \to \mathcal{P}_2(\mathbb{R})^2$ such that the family of piecewise constant interpolations, $\{\gamma_\tau\}_{\tau>0}$ admits a subsequence $\{\gamma_k\}_{k\in\mathbb{N}} := \{\gamma_{\tau_k}\}_{k\in\mathbb{N}}$ which converges narrowly to γ uniformly in $t \in [0,T]$ as $k \to +\infty$.*

Proof. Consider two consecutive iterations, γ_τ^n and γ_τ^{n+1}, of the JKO scheme, Eq. (24). By the optimality of γ_τ^{n+1} we obtain

$$\frac{1}{2\tau}\mathcal{W}_2^2(\gamma_\tau^n, \gamma_\tau^{n+1}) \leq \mathcal{F}(\gamma_\tau^n) - \mathcal{F}(\gamma_\tau^{n+1}), \tag{25}$$

which implies

$$\mathcal{F}(\gamma_\tau^n) \leq \mathcal{F}(\gamma_0), \tag{26}$$

for all $n \in \mathbb{N}$. Summing over k from m to n with $m < n$ we obtain the following telescopic sum

$$\frac{1}{2\tau}\sum_{k=m}^{n}\mathcal{W}_2^2(\gamma_\tau^k, \gamma_\tau^{k+1}) \leq \mathcal{F}(\gamma_\tau^m) - \mathcal{F}(\gamma_\tau^{n+1}). \tag{27}$$

Now, let us consider $t \in ((n-1)\tau, n\tau]$; using the estimate (27) we obtain

$$\mathcal{W}_2^2(\gamma_\tau(0), \gamma_\tau(t)) \leq 2n\tau\mathcal{F}(\gamma_0) - 2n\tau\mathcal{F}(\gamma_\tau^n). \tag{28}$$

Using the Remark 1 after Hölder and (weighted) Young inequalities, it is possible to obtain a bound from below for $\mathcal{F}(\gamma_\tau^n)$, which gives in combination with (28) the following estimate

$$\mathcal{W}_2^2(\gamma_\tau(0), \gamma_\tau(t)) \leq 4T\mathcal{F}(\gamma_0) + m_2(\gamma_0) + 16T^2 =: C(\gamma_0, T). \tag{29}$$

From estimate (29) and the inequality in Remark 1 we can deduce the second moment of $\gamma_\tau(t)$ is uniformly bounded on compact time intervals. Moreover, as a consequence of estimates (27) and (29), using once again the bound from below for $\mathcal{F}(\gamma_\tau^n)$ we get

$$\sum_{k=m}^{n-1}\mathcal{W}_2^2(\gamma_\tau^k, \gamma_\tau^{k+1}) \leq \tau\bar{C}(\gamma_0, T). \tag{30}$$

Now, let us consider $0 \leq s < t$ such that $s \in ((m-1)\tau, m\tau]$ and $t \in ((n-1)\tau, n\tau]$. It easy to check that $|n-m| < \frac{|t-s|}{\tau} + 1$. By means of the Cauchy-Schwartz inequality and (30) we obtain $\frac{1}{2}$-Hölder equi-continuity for γ_τ (up to a negligible error of order $\sqrt{\tau}$) since

$$\begin{aligned}\mathcal{W}_2(\gamma_\tau(s), \gamma_\tau(t)) &\leq \sum_{k=m}^{n-1}\mathcal{W}_2(\gamma_\tau^k, \gamma_\tau^{k+1}) \leq \left(\sum_{k=m}^{n-1}\mathcal{W}_2^2(\gamma_\tau^k, \gamma_\tau^{k+1})\right)^{\frac{1}{2}} |n-m|^{\frac{1}{2}} \\ &\leq c\left(\sqrt{|t-s|} + \sqrt{\tau}\right),\end{aligned} \tag{31}$$

where c is a positive constant. The refined version of the Ascoli-Arzelà theorem yields the narrow compactness, cf. [1, Proposition 3.3.1]. This completes the proof. □

Remark 6 (Extension of solutions)**.** *From Proposition 5 we can construct a curve $\gamma : [0,T] \to \mathcal{P}_2(\mathbb{R})^2$ for any $T > 0$. Thus we may extend the solution up to the point where the second order moments of the solution become unbounded. By construction this is not possible and can only* *happen at $T = +\infty$. As a consequence Proposition 5 proves the existence of a limiting curve $\gamma : [0, \infty) \to \mathcal{P}_2(\mathbb{R})^2$.*

As already mentioned, we can prove more refined estimates for the solution γ to system (6). Indeed, we show that, starting with initial data $\rho_0, \eta_0 \in \mathcal{P}_2(\mathbb{R}) \cap L^m(\mathbb{R})$, for $m \in (1, +\infty]$, the solution keeps this regularity for every $t \geq 0$ and it has second order moments uniformly bounded in time. These properties can be establish by using the "flow interchange" technique developed by Otto in [52], and Matthes, McCann and Savaré in [49]. This technique is based on the idea that *the dissipation of one functional along the gradient flow of another functional equals the dissipation of the second functional along the gradient flow of the first one.* In this spirit, the "Evolution Variational Inequality" (**E.V.I.**) linked with the auxiliary gradient flow is crucial in order to obtain useful refined estimates (see for instance [35, 37]). The connection between gradient flows and evolutionary PDEs of diffusion type shown in [1, 43, 52, 54] allows us to consider the (decoupled) system

$$\begin{cases} \partial_t u_1 = \partial_{xx} u_1^m + \varepsilon \partial_{xx} u_1, \\ \partial_t u_2 = \partial_{xx} u_2^m + \varepsilon \partial_{xx} u_2, \end{cases} \tag{32a}$$

as the gradient flow of the functional

$$\begin{aligned} \mathcal{E}(u_1, u_2) = & \frac{1}{m-1} \int_{\mathbb{R}} [u_1(x)^m + u_2(x)^m] \, \mathrm{d}x \\ & + \varepsilon \int_{\mathbb{R}} [u_1(x) \log u_1(x) + u_2(x) \log u_2(x)] \, \mathrm{d}x, \end{aligned} \tag{32b}$$

as well as the following system

$$\begin{cases} \partial_t u_1 = \partial_x (2x u_1) + \varepsilon \partial_{xx} u_1, \\ \partial_t u_2 = \partial_x (2x u_2) + \varepsilon \partial_{xx} u_2, \end{cases} \tag{33a}$$

which can be seen as the gradient flow of the functional

$$\begin{aligned} \mathcal{G}(u_1, u_2) = & \int_{\mathbb{R}} |x|^2 (u_1(x) + u_2(x)) \, \mathrm{d}x \\ & + \varepsilon \int_{\mathbb{R}} [u_1(x) \log u_1(x) + u_2(x) \log u_2(x)] \, \mathrm{d}x, \end{aligned} \tag{33b}$$

for $\varepsilon > 0$ and $m \in (1, \infty)$, with respect to the product 2-Wasserstein distance $\mathcal{W}_2$. We shall employ the flow interchange strategy twice taking as auxiliary functional

(1) $\mathcal{E}$ to get L^m-regularity ($m > 1$) for the solution γ,
(2) $\mathcal{G}$ in order to obtain a uniform bound in time for the second order moments of γ.

For the reader's convenience we shall sometimes use the symbol $\mathcal{A}$ to denote either $\mathcal{G}$ or $\mathcal{E}$. The functional $\mathcal{A} \in \{\mathcal{G}, \mathcal{E}\}$ possess a 0-flow given by the semigroup and $\boldsymbol{S}_{\mathcal{A}} = (\boldsymbol{S}_{\mathcal{A}}^1, \boldsymbol{S}_{\mathcal{A}}^2)$, see for instance [33]. In particular, by setting

$$\boldsymbol{S}_{\mathcal{A}}^{1,t}(\nu_1) := u_1(t, \cdot), \quad \text{and} \quad \boldsymbol{S}_{\mathcal{A}}^{2,t}(\nu_2) := u_2(t, \cdot),$$

we have $u(t, \cdot) = (u_1(t, \cdot), u_2(t, \cdot))$ is the unique classical solution at time t of system (32a) (respectively (33a)) coupled with an initial value (ν_1, ν_2) at $t = 0$ in case $\mathcal{A} = \mathcal{E}$ ($\mathcal{A} = \mathcal{G}$, respectively).

Remark 7. *As in* [43, Proposition 4.1]*, we know that the log-entropy*

$$\mathcal{H}(\rho) = \int_{\mathbb{R}^d} \rho(x) \log \rho(x) \, \mathrm{d}x,$$

is bounded from below in terms of the second moment $m_2(\rho)$, i.e.,

$$\mathcal{H}(\rho) \geq -C(m_2(\rho)+1)^\beta,$$

for every $\rho \in \mathcal{P}_2^a(\mathbb{R}^d)$, $\beta \in (\frac{d}{d+2}, 1)$, and $C < +\infty$, depending only on the space dimension d. We are going to use this inequality in order to have a uniform bound from below for the entropic part in (32b) *and* (33b)*.*

For every $\gamma = (\rho, \eta) \in \mathcal{P}_2^a(\mathbb{R})^2$, let us define the dissipation of $\mathcal{F}$ along $\boldsymbol{S}_\mathcal{A}$ by

$$\boldsymbol{D}_\mathcal{A}\mathcal{F}(\gamma) := \limsup_{s \downarrow 0} \frac{\mathcal{F}(\gamma) - \mathcal{F}(\boldsymbol{S}^s_\mathcal{A}\gamma)}{s},$$

where $\mathcal{A} \in \{\mathcal{G}, \mathcal{E}\}$. We prove the following proposition.

Proposition 6. *Let $m \in (1, +\infty)$ and let $\gamma_0 = (\rho_0, \eta_0) \in (\mathcal{P}_2^a(\mathbb{R}) \cap L^m(\mathbb{R}))^2$ be such that $\mathcal{E}(\gamma_0) < +\infty$. The piecewise constant interpolation $\gamma_\tau = (\rho_\tau, \eta_\tau)$ satisfies*

$$\|\rho_\tau\|_{L^\infty(0,+\infty;L^m(\mathbb{R}))} + \|\eta_\tau\|_{L^\infty(0,+\infty;L^m(\mathbb{R}))} \leq \|\rho_0\|_{L^m(\mathbb{R})} + \|\eta_0\|_{L^m(\mathbb{R})} \tag{34}$$

Moreover, the limit curve γ belongs to $L^\infty(0, +\infty; L^m(\mathbb{R}))^2$. In fact, this property can be extended to the cases $m = +\infty$.

Proof. As an easy consequence of the definition of the sequence $\{\gamma_\tau^n\}_{n\in\mathbb{N}}$, for all $s > 0$ we have that

$$\frac{1}{2\tau}\mathcal{W}_2^2(\gamma_\tau^{n+1}, \gamma_\tau^n) + \mathcal{F}(\gamma_\tau^{n+1}) \leq \frac{1}{2\tau}\mathcal{W}_2^2(\boldsymbol{S}^s_\mathcal{E}\gamma_\tau^{n+1}, \gamma_\tau^n) + \mathcal{F}(\boldsymbol{S}^s_\mathcal{E}\gamma_\tau^{n+1}).$$

Dividing by $s > 0$ and passing to the $\limsup$ as $s \downarrow 0$ we get

$$\tau \boldsymbol{D}_\mathcal{E}\mathcal{F}(\gamma_\tau^{n+1}) \leq \frac{1}{2}\frac{d^+}{dt}\left(\mathcal{W}_2^2(\boldsymbol{S}^t_\mathcal{E}\gamma_\tau^{n+1}, \gamma_\tau^n)\right)\Big|_{t=0} \overset{\textbf{(E.V.I.)}}{\leq} \mathcal{E}(\gamma_\tau^n) - \mathcal{E}(\gamma_\tau^{n+1}), \tag{35}$$

where in the last inequality the well-known connection between displacement convexity and the **E.V.I.** is crucial, see e.g. [33]. Now, concerning the left-hand side of (35), we notice that

$$\begin{aligned} \boldsymbol{D}_\mathcal{E}\mathcal{F}(\gamma_\tau^{n+1}) &= \limsup_{s\downarrow 0} \frac{\mathcal{F}(\gamma_\tau^{n+1}) - \mathcal{F}(\boldsymbol{S}^s_\mathcal{E}\gamma_\tau^{n+1})}{s} \\ &= \limsup_{s\downarrow 0} \int_0^1 \left(-\frac{d}{dz}\Big|_{z=st} \mathcal{F}(\boldsymbol{S}^z_\mathcal{E}\gamma_\tau^{n+1})\right) dt. \end{aligned} \tag{36}$$

Hence, let us focus on the time derivative inside the above integral. Keep in mind that $\boldsymbol{S}^t_\mathcal{E}\gamma_\tau^{n+1}$ is the solution to the decoupled system of nonlinear parabolic equations with strictly positive coefficients, system (32a). Then, using the C^∞-regularity of $\boldsymbol{S}^t_\mathcal{E}\gamma_\tau^{n+1}$ we may infer

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\mathcal{F}(\boldsymbol{S}^t_\mathcal{E}\gamma_\tau^{n+1}) = &- \int_\mathbb{R} \left(\boldsymbol{S}^{1,t}_\mathcal{E}\rho_\tau^{n+1} - \boldsymbol{S}^{2,t}_\mathcal{E}\eta_\tau^{n+1}\right)\left([\boldsymbol{S}^{1,t}_\mathcal{E}\rho_\tau^{n+1}]^m - [\boldsymbol{S}^{2,t}_\mathcal{E}\eta_\tau^{n+1}]^m\right) \mathrm{d}x \\ &- \varepsilon \int_\mathbb{R} \left(\boldsymbol{S}^{1,t}_\mathcal{E}\rho_\tau^{n+1} - \boldsymbol{S}^{2,t}_\mathcal{E}\eta_\tau^{n+1}\right)^2 \mathrm{d}x, \end{aligned} \tag{37}$$

where the terms at infinity in the integration by parts vanish due to the rapid decay of the solution to a nondegenerate diffusion equation [46]. (37) yields

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{F}(\boldsymbol{S}^t_{\mathcal{E}}\gamma^{n+1}_\tau) \le 0. \tag{38}$$

Combining (38) with (36) and (35) we obtain

$$0 \le \tau \boldsymbol{D}_{\mathcal{E}}\mathcal{F}(\gamma^{n+1}_\tau) \le \mathcal{E}(\gamma^n_\tau) - \mathcal{E}(\gamma^{n+1}_\tau),$$

whence

$$\mathcal{E}(\gamma^n_\tau) \le \mathcal{E}(\gamma_0), \tag{39}$$

for all $n \in \mathbb{N}$. By Remark 7 we control the log-entropic part of $\mathcal{E}$ and we deduce that, as $\varepsilon \downarrow 0$,

$$\int_{\mathbb{R}} [\rho^n_\tau(x)]^m + [\eta^n_\tau(x)]^m \, \mathrm{d}x \le \int_{\mathbb{R}} [\rho_0(x)]^m + [\eta_0(x)]^m \, \mathrm{d}x, \tag{40}$$

whence

$$\int_{\mathbb{R}} [\rho_\tau(t,x)]^m + [\eta_\tau(t,x)]^m \, \mathrm{d}x \le \int_{\mathbb{R}} [\rho_0(x)]^m + [\eta_0(x)]^m \, \mathrm{d}x, \tag{41}$$

for every $t \ge 0$. This proves estimate (34) for $m \in (1,\infty)$. In order to extend the estimate to the case $m = +\infty$, we observe that

$$\begin{aligned}
\|\rho_\tau(t,\cdot)\|_{L^\infty(\mathbb{R})} + \|\eta_\tau(t,\cdot)\|_{L^\infty(\mathbb{R})} &\le \limsup_{m\to+\infty} \left[\|\rho_\tau(t,\cdot)\|_{L^m(\mathbb{R})} + \|\eta_\tau(t,\cdot)\|_{L^m(\mathbb{R})}\right] \\
&\le \limsup_{m\to+\infty} \left[\|\rho_0\|_{L^m(\mathbb{R})} + \|\eta_0\|_{L^m(\mathbb{R})}\right] \\
&\le \limsup_{m\to+\infty} \left[\|\rho_0\|^{\frac{m-1}{m}}_{L^\infty(\mathbb{R})}\|\rho_0\|^{\frac{1}{m}}_{L^1(\mathbb{R})} + \|\eta_0\|^{\frac{m-1}{m}}_{L^\infty(\mathbb{R})}\|\eta_0\|^{\frac{1}{m}}_{L^1(\mathbb{R})}\right] \\
&= \|\rho_0\|_{L^\infty(\mathbb{R})} + \|\eta_0\|_{L^\infty(\mathbb{R})}.
\end{aligned}$$

We conclude that, for all $T>0$, the subsequence $\{\gamma_{\tau_k}\}_{k\in\mathbb{N}}$ obtained from Proposition 5 is uniformly bounded in $L^\infty([0,T];L^m(\mathbb{R}))^2$. By Banach-Alaoglu's Theorem, in case m is finite, there exists a subsequence $(\tau'_k) \subset (\tau_k)$ such that $\{\gamma_{\tau'_k}\}_{k\in\mathbb{N}}$ converges in the weak $L^m_{x,t}$ topology to some limit $\gamma' \in L^m([0,T]\times\mathbb{R})^2$. In the case of $m = +\infty$ the above subsequence exists in the weak-$\star$ topology of $L^\infty([0,T]\times\mathbb{R})$. Due to Proposition 5 the limit γ' coincides with γ on $[0,T]$. By a simple weak lower semi-continuity argument we deduce that γ inherits the same estimates as the approximating sequence γ_{τ_k}. Since T was arbitrary we conclude the proof. □

Lemma 11. *Let $\gamma_0 = (\rho_0,\eta_0) \in \mathcal{P}^a_2(\mathbb{R}) \times \mathcal{P}^a_2(\mathbb{R})$ be such that $\mathcal{G}(\gamma_0) < +\infty$. The piecewise constant interpolation $\gamma_\tau = (\rho_\tau,\eta_\tau)$ satisfies*

$$\int_{\mathbb{R}} |x|^2[\rho_\tau(t,x) + \eta_\tau(t,x)] \, \mathrm{d}x \le \int_{\mathbb{R}} |x|^2[\rho_0(x) + \eta_0(x)] \, \mathrm{d}x, \tag{42}$$

for every $t \ge 0$.In addition, the limiting curve γ has uniformly bounded second order moments in time.

Proof. Arguing as in the proof of Proposition 6, from the scheme (24) we easily get

$$\tau \boldsymbol{D}_{\mathcal{G}}\mathcal{F}(\gamma^{n+1}_\tau) \le \frac{1}{2}\frac{\mathrm{d}^+}{\mathrm{d}t}\left(\mathcal{W}^2_2(\boldsymbol{S}^t_{\mathcal{G}}\gamma^{n+1}_\tau,\gamma^n_\tau)\right)\Big|_{t=0} \overset{\textbf{\textit{(E.V.I.)}}}{\le} \mathcal{G}(\gamma^n_\tau) - \mathcal{G}(\gamma^{n+1}_\tau), \tag{43}$$

where, again, the left-hand side of (43) can be rewritten as

$$\boldsymbol{D}_{\mathcal{G}}\mathcal{F}(\gamma_\tau^{n+1}) = \limsup_{s\downarrow 0}\int_0^1 \left(-\frac{\mathrm{d}}{\mathrm{d}z}\Big|_{z=st}\mathcal{F}(\boldsymbol{S}_{\mathcal{G}}^{z}\gamma_\tau^{n+1})\right)\,\mathrm{d}t. \tag{44}$$

Since $\boldsymbol{S}_{\mathcal{G}}^{t}\gamma_\tau^{n+1}$ is the solution to system (33a), which is a decoupled system of linear Fokker-Planck equations, we use its C^∞-regularity to obtain

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{F}(\boldsymbol{S}_{\mathcal{G}}^{t}\gamma_\tau^{n+1}) = &-\int_{\mathbb{R}}|N'\star(\boldsymbol{S}_{\mathcal{G}}^{1,t}\rho_\tau^{n+1}-\boldsymbol{S}_{\mathcal{G}}^{2,t}\eta_\tau^{n+1})|^2\,\mathrm{d}x\\
&-\varepsilon\int_{\mathbb{R}}(\boldsymbol{S}_{\mathcal{G}}^{1,t}\rho_\tau^{n+1}-\boldsymbol{S}_{\mathcal{G}}^{2,t}\eta_\tau^{n+1})^2\,\mathrm{d}x\\
&+\frac{1}{2}\left[N\star\left(\boldsymbol{S}_{\mathcal{G}}^{1,t}\rho_\tau^{n+1}-\boldsymbol{S}_{\mathcal{G}}^{2,t}\eta_\tau^{n+1}\right)\,N'\star\left(\boldsymbol{S}_{\mathcal{G}}^{1,t}\rho_\tau^{n+1}-\boldsymbol{S}_{\mathcal{G}}^{2,t}\eta_\tau^{n+1}\right)\right]_{x=-\infty}^{x=+\infty}\\
&+\varepsilon\left[N'\star\left(\boldsymbol{S}_{\mathcal{G}}^{1,t}\rho_\tau^{n+1}-\boldsymbol{S}_{\mathcal{G}}^{2,t}\eta_\tau^{n+1}\right)\,\left(\boldsymbol{S}_{\mathcal{G}}^{1,t}\rho_\tau^{n+1}-\boldsymbol{S}_{\mathcal{G}}^{2,t}\eta_\tau^{n+1}\right)\right]_{x=-\infty}^{x=+\infty}\\
&-\left[2xN\star\left(\boldsymbol{S}_{\mathcal{G}}^{1,t}\rho_\tau^{n+1}-\boldsymbol{S}_{\mathcal{G}}^{2,t}\eta_\tau^{n+1}\right)\,\left(\boldsymbol{S}_{\mathcal{G}}^{1,t}\rho_\tau^{n+1}-\boldsymbol{S}_{\mathcal{G}}^{2,t}\eta_\tau^{n+1}\right)\right]_{x=-\infty}^{x=+\infty}\\
&-\varepsilon\left[N\star\left(\boldsymbol{S}_{\mathcal{G}}^{1,t}\rho_\tau^{n+1}-\boldsymbol{S}_{\mathcal{G}}^{2,t}\eta_\tau^{n+1}\right)\,\partial_x\left(\boldsymbol{S}_{\mathcal{G}}^{1,t}\rho_\tau^{n+1}-\boldsymbol{S}_{\mathcal{G}}^{2,t}\eta_\tau^{n+1}\right)\right]_{x=-\infty}^{x=+\infty}.\end{aligned} \tag{45}$$

Let us consider the boundary terms individually. For convenience we set $\rho^t = \boldsymbol{S}_{\mathcal{G}}^{1,t}\rho_\tau^{n+1}$, $\eta^t = \boldsymbol{S}_{\mathcal{G}}^{2,t}\eta_\tau^{n+1}$, and $\kappa := \rho^t - \eta^t$ in order to simplify the notation. Concerning the first term, we have

(46)
$$\begin{aligned}\left|N\star\kappa\,N'\star\kappa\right| &= \left|\int_{\mathbb{R}}|x-y|\kappa(y)\,\mathrm{d}y\int_{\mathbb{R}}\mathrm{sign}(x-y)\kappa(y)\,\mathrm{d}y\right|\\
&= \left|\int_{\mathbb{R}}(x-y)\kappa(y)\,\mathrm{d}y + 2\int_x^{+\infty}(y-x)\kappa(y)\,\mathrm{d}y\right|\times\left|\int_{\mathbb{R}}\kappa(y)\,\mathrm{d}y - 2\int_x^{+\infty}\kappa(y)\,\mathrm{d}y\right|\\
&= \left|-\int_{\mathbb{R}}y\kappa(y)\,\mathrm{d}y - 2x\int_x^{+\infty}\kappa(y)\,\mathrm{d}y + 2\int_x^{+\infty}y\kappa(y)\,\mathrm{d}y\right|\times\left|-2\int_x^{+\infty}\kappa(y)\,\mathrm{d}y\right|\\
&\le 2\left(2+2m_2(\rho^t)+2m_2(\eta^t)+\frac{2}{|x|}(m_2(\rho^t)+m_2(\eta^t))\right)\frac{2\left[m_2(\rho^t)+m_2(\eta^t)\right]}{|x|^2},\end{aligned}$$

which vanishes as $|x|\to+\infty$. Regarding the second term we note that

$$\begin{aligned}\left|N'\star\kappa\,\kappa\right| &= \left|\int_{\mathbb{R}}\mathrm{sign}(x-y)\kappa(y)\kappa(x)\,\mathrm{d}y\right|\\
&\le 2(\rho^t(x)+\eta^t(x))\end{aligned} \tag{47}$$

which vanishes as $|x| \to +\infty$ because ρ^t and η^t are solutions of linear Fokker-Planck equations decaying exponentially at infinity. Now, for the same reason, there holds

$$\begin{aligned}|2xN \star \kappa \cdot \kappa| &= \left| \int_{\mathbb{R}} 2x|x-y|\kappa(y)\kappa(x)\,\mathrm{d}y \right| \\ &\leq 2\int_{\mathbb{R}} |x|^2(\rho^t(y)+\eta^t(y))(\rho^t(x)+\eta^t(x))\,\mathrm{d}y \\ &\quad + 2\int_{\mathbb{R}} |x||y|(\rho^t(y)+\eta^t(y))(\rho^t(x)+\eta^t(x))\,\mathrm{d}y \\ &\leq 4|x|^2(\rho^t(x)+\eta^t(x)) + (4+m_2(\rho^t)+m_2(\eta^t))|x|(\rho^t(x)+\eta^t(x)),\end{aligned} \tag{48}$$

vanishes $|x| \to +\infty$. As for the last boundary term we get

$$\begin{aligned}|N \star \kappa\, \partial_x \kappa| &= \left| \int_{\mathbb{R}} |x-y|\kappa(y)\partial_x\kappa(x)\,\mathrm{d}y \right| \\ &\leq \int_{\mathbb{R}} |x|(\rho^t(y)+\eta^t(y))(|\partial_x\rho^t(x)|+|\partial_x\eta^t(x)|)\,\mathrm{d}y \\ &\quad + \int_{\mathbb{R}} |y|(\rho^t(y)+\eta^t(y))(|\partial_x\rho^t(x)|+|\partial_x\eta^t(x)|)\,\mathrm{d}y \\ &\leq 2|x|(|\partial_x\rho^t(x)|+|\partial_x\eta^t(x)|) + (2+m_2(\rho^t)+m_2(\eta^t))(|\partial_x\rho^t(x)|+|\partial_x\eta^t(x)|),\end{aligned} \tag{49}$$

which, again, goes to 0 as $|x| \to +\infty$. Using (46), (47), (48), and (49) in (45) we deduce that

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{F}(\boldsymbol{S}^t_{\mathcal{G}}\gamma^{n+1}_\tau) \leq 0, \tag{50}$$

which, in combination with (44) and (43), gives

$$\mathcal{G}(\gamma^n_\tau) \leq \mathcal{G}(\gamma_0).$$

Hence, taking into account Remark 7, letting $\varepsilon \to 0^+$ we get

$$\int_{\mathbb{R}} |x|^2[\rho_\tau(t,x)+\eta_\tau(t,x)]\,\mathrm{d}x \leq \int_{\mathbb{R}} |x|^2[\rho_0(x)+\eta_0(x)]\,\mathrm{d}x, \tag{51}$$

for every $t \geq 0$, i.e. estimate (42). By similar considerations to the ones at the end of the proof of Proposition 6 and from Proposition 5 we conclude that γ has second order moments uniformly bounded in time. □

Remark 8 (Preservation of absolute continuity)**.** *A natural question is to ask as to whether absolute continuity of solutions is kept provided that the initial data satify $\rho_0, \eta_0 \in L^1(\mathbb{R}) \cap \mathcal{P}_2(\mathbb{R})$. The answer to this question is positive. In fact, $\rho_0, \eta_0 \in L^1(\mathbb{R}) \cap \mathcal{P}_2(\mathbb{R})$ implies the existence of two nonnegative, superlinear, and convex functions Φ_1, Φ_2 with $\Phi_i(0) = 0$, for $i = 1, 2$, satisfying $\Phi_1(\rho_0), \Phi_2(\eta_0) \in L^1(\mathbb{R})$. It is easy to check that, individually, the two Φ_i are geodesically convex as they satisfy the McCann condition trivially in one dimension. Setting $\Phi := \sup_{i=1,2}(\Phi_i)$, we readily verify that this function satisfies the McCann condition and is therefore geodesically convex. Moreover, by this choice, $\Phi(\rho_0), \Phi(\eta_0) \in L^1(\mathbb{R})$. Applying the flow interchange argument as above for the extended functional*

$$\mathcal{E}(\nu_1,\nu_2) = \int_{\mathbb{R}} \Phi(\nu_1) + \Phi(\nu_2)\,\mathrm{d}x + \epsilon\mathcal{H}(\nu_1) + \epsilon\mathcal{H}(\nu_2),$$

yields uniform bounds of the form

$$\int_{\mathbb{R}} \Phi(\rho_\tau^{n+1}) + \Phi(\eta_\tau^{n+1}) \, \mathrm{d}x \leq \int_{\mathbb{R}} \Phi(\rho_0) + \Phi(\eta_0) \, \mathrm{d}x.$$

The uniform control of superlinear function Φ composed with the two species then yields uniform bound on $(\rho_\tau^n)_{\tau>0,n\in\mathbb{N}}, (\eta_\tau^n)_{\tau>0,n\in\mathbb{N}}$ in $L^1(\mathbb{R})$. Together with the uniform control of the second order moments we may invoke the Dunford-Pettis theorem to obtain weak compactness in $L^1(\mathbb{R})$.

The final step in this procedure is to prove the curve γ obtained in Proposition 5 is a curve of maximal slope for $\mathcal{F}$, arguing as in [25, 36]. Since curves of maximal slope coincide with gradient flows, see [1, Theorem 11.1.3], we actually have that γ is a gradient flow solution to (6) in the sense of Definition 9. Let us summarise the procedure for the sake of completeness. We denote by $\tilde{\gamma}_\tau$ the *De Giorgi variational interpolation* (cf. [1, Definition 3.2.1]), i.e. any interpolation $\tilde{\gamma}_\tau : [0, +\infty) \to \mathcal{P}_2(\mathbb{R})^2$ of the discrete values $\{\gamma_\tau^n\}_{n\in\mathbb{N}}$ defined through scheme (24) such that

$$\tilde{\gamma}_\tau(t) = \tilde{\gamma}_\tau((n-1)\tau + \delta) \in \operatorname{argmin}_{\gamma\in\mathcal{P}_2(\mathbb{R})^2} \left\{ \frac{1}{2\delta} \mathcal{W}_2^2(\gamma_\tau^{n-1}, \gamma) + \mathcal{F}(\gamma) \right\}, \tag{52}$$

if $t = (n-1)\tau + \delta \in ((n-1)\tau, n\tau]$. Now, from [1, Theorem 3.14, Lemma 3.2.2] it is possible to get the following energy inequality

$$\mathcal{F}(\gamma_0) \geq \frac{1}{2} \int_0^T \|v_k(t)\|^2_{L^2(\gamma_k(t))} \, \mathrm{d}t + \frac{1}{2} \int_0^T |\partial \mathcal{F}|^2 [\tilde{\gamma}_k(t)] \, \mathrm{d}t + \mathcal{F}(\gamma_k(T)), \tag{53}$$

where the pair (γ_k, v_k) is the solution of the continuity equation $\partial_t \gamma_k(t) + \operatorname{div}(v_k(t)\gamma_k(t)) = 0$ in the sense of distributions. Here γ_k is the subsequence from Proposition 5 and $v_k(t)$ is the unique velocity field with minimal $L^2(\gamma_k(t))$-norm (see Remark below), and $\tilde{\gamma}_k := \tilde{\gamma}_{\tau_k}$ is defined by (52).

Remark 9 (Absolutely continuous curve and the continuity equation)**.** *Thanks to Proposition 5 we know γ_k is an absolutely continuous curve, therefore we can identify its tangent vectors with the velocity fields $v_k(t)$ such that the continuity equation $\partial_t \gamma_k(t) + div(v_k(t)\gamma_k(t)) = 0$ is satisfied in a distributional sense, according to* [1, Theorem 8.3.1]*. Furthermore,* [1, Proposition 8.4.5] *asserts there is only one $v_k(t)$ with minimal $L^2(\gamma_k(t))$-norm, equal to the metric derivative of $\gamma_k(t)$ for a.e. t.*

Up to a subsequence both the interpolations γ_τ and $\tilde{\gamma}_\tau$ narrowly converge to γ in view of Proposition 5. Proving that γ is a curve of maximal slope in the sense of Definition 10 is then a consequence of the lower semi-continuity of the slope and the energy inequality (53) retracing [25, Lemma 2.7 and Theorem 2.8]. Thanks to [1, Theorem 11.1.3] we actually have that γ is a gradient flow solution to (6) in the sense of Definition 9. At this stage, the uniqueness of the gradient flow solutions in the sense of Definition 9 follows from the geodesic convexity of $\mathcal{F}$ proven in Proposition 4, relying on [1, Theorem 11.1.4]. More precisely, given two gradient flow solutions $\gamma_1(t)$ and $\gamma_2(t)$ in the sense of Definition 9, we obtain the stability property

$$\mathcal{W}_2(\gamma_1(t), \gamma_2(t)) \leq \mathcal{W}_2(\gamma_1(0), \gamma_2(0)), \tag{54}$$

for all $t \geq 0$. In addition, the unique gradient flow solution satisfies the Evolution Variational Inequality (**E.V.I.**):

$$\frac{1}{2} \frac{\mathrm{d}}{\mathrm{d}t} \mathcal{W}_2^2(\gamma(t), \bar{\gamma}) \leq \mathcal{F}(\bar{\gamma}) - \mathcal{F}(\gamma(t)) \tag{55}$$

for almost all $t > 0$ and all $\bar{\gamma} \in \mathcal{P}(\mathbb{R})^2$.

The property (55) can actually be used to show a stronger property, namely that γ is a gradient flow in the sense of Definition 7, which implies uniqueness in the weaker notion of solution defined in Definition 7 in view of Theorem 8. We prove this statement in the following Theorem, which also collects all the estimates proven in this subsection.

Theorem 12. *Let $m \in (1, +\infty]$. Let $\rho_0, \eta_0 \in \mathcal{P}_2(\mathbb{R}) \cap L^m(\mathbb{R})$. Then, there exists a unique $\gamma = (\rho, \eta) \in L^\infty([0, +\infty); \mathcal{P}_2(\mathbb{R})^2 \cap L^m(\mathbb{R})^2)$ solving (6) in the sense of Definition 9. Moreover, γ is the unique solution to (6) in the sense of Definition 7 as well. Finally, we have the properties*

$$\|\rho(t,\cdot)\|_{L^m(\mathbb{R})} + \|\eta(t,\cdot)\|_{L^m(\mathbb{R})} \le \|\rho_0\|_{L^m(\mathbb{R})} + \|\eta_0\|_{L^m(\mathbb{R})},$$

$$\int_{\mathbb{R}} |x|^2 \rho(t,x)\,\mathrm{d}x + \int_{\mathbb{R}} |x|^2 \eta(t,x)\,\mathrm{d}x \le \int_{\mathbb{R}} |x|^2 \rho_0(x)\,\mathrm{d}x + \int_{\mathbb{R}} |x|^2 \eta_0(x)\,\mathrm{d}x.$$

Proof. All the statements have been proven earlier in this subsection, in particular in Proposition 6 and Lemma 11. We only need to prove that γ is a solution to (6) in the sense of Definition 7. Recall the **E.V.I.** (55) for a general $\overline{\gamma} \in \mathcal{P}_2(\mathbb{R})^2$. Integrating the inequality (55) on the time interval $[s, t]$ and dividing by $t - s$ we obtain

$$\frac{1}{2(t-s)} \left[\mathcal{W}_2^2(\gamma(t), \bar{\gamma}) - \mathcal{W}_2^2(\gamma(s), \bar{\gamma})\right] \le \frac{1}{t-s} \int_s^t \left[\mathcal{F}(\bar{\gamma}) - \mathcal{F}(\gamma(t'))\right] \mathrm{d}t'.$$

Consider now the pseudo-inverse variables X_ρ, X_η, $\overline{X}_\rho$, and $\overline{X}_\eta$ of ρ, η, $\overline{\rho}$, and $\overline{\eta}$ respectively, where $\gamma = (\rho, \eta)$ and $\overline{\gamma} = (\overline{\rho}, \overline{\eta})$. The formulas (12) and (13) applied to our case imply

$$\frac{1}{2(t-s)} \left[\|X_\rho(t) - \overline{X}_\rho\|_{L^2}^2 + \|X_\eta(t) - \overline{X}_\eta\|_{L^2}^2 - \|X_\rho(s) - \overline{X}_\rho\|_{L^2}^2 + \|X_\eta(s) - \overline{X}_\eta\|_{L^2}^2\right]$$
$$\le \frac{1}{t-s} \int_s^t \left[\mathfrak{F}(\overline{X}_\rho, \overline{X}_\eta) - \mathfrak{F}(X_\rho(t'), X_\eta(t'))\right] \mathrm{d}t'.$$

We notice that the indicator function has been considered as zero in $\mathfrak{F}$ as all the involved variables are in the cone $\mathcal{C}$. By absolute continuity in time of the curve $t \mapsto (X_\rho(t), X_\eta(t))$, recalling the expression of $\mathfrak{F}$, we can let $s \uparrow t$ and obtain

$$(\dot{X}_\rho(t), X_\rho(t) - \overline{X}_\rho)_{L^2(0,1)} + (\dot{X}_\eta(t), X_\eta(t) - \overline{X}_\eta)_{L^2(0,1)} \le \mathfrak{F}(\overline{X}_\rho, \overline{X}_\eta) - \mathfrak{F}(X_\rho(t), X_\eta(t)),$$

which, since $(\overline{X}_\rho, \overline{X}_\eta)$ was arbitrary, is equivalent to state that $Z(t) = (X_\rho(t), X_\eta(t))$ satisfies

$$-\dot{Z}(t) \in \partial\mathfrak{F}[Z(t)].$$

The proof will be completed once we show that $Z(t)$ is a Lipschitz curve. Recall the estimate (28) at the level of the JKO scheme, which can be rewritten as

$$\mathcal{W}_2^2(\gamma_\tau(0), \gamma_\tau(h)) \le 2(h+\tau)[\mathcal{F}(\gamma_0) - \mathcal{F}(\gamma_\tau(h))],$$

for all $h > 0$. Sending $\tau \downarrow 0$ and using the fact that the functional $\mathcal{F}$ is continuous w.r.t. $\mathcal{W}_2$, we get

$$\frac{1}{h^2}\mathcal{W}_2^2(\gamma_0, \gamma(h)) \le \frac{2}{h}[\mathcal{F}(\gamma_0) - \mathcal{F}(\gamma(h))].$$

In the pseudo-inverse formalism the above estimate reads

$$\frac{1}{h^2}\left[\|X_\rho(h) - X_0\|_{L^2}^2 + \|X_\eta(h) - Y_0\|_{L^2}^2\right] \le \frac{2}{h}[\mathfrak{F}(X_0, Y_0) - \mathfrak{F}(X_\rho(h), X_\eta(h))],$$

where X_0 and Y_0 are the pseudo-inverses corresponding to ρ_0 and η_0, respectively. The definition of sub-differential, $\partial\mathfrak{F}$, then implies

$$\lim_{h\downarrow 0}\frac{1}{h^2}\left[\|X_\rho(h)-X_0\|_{L^2}^2+\|X_\eta(h)-Y_0\|_{L^2}^2\right]\le\lim_{h\downarrow 0}\frac{2}{h}\left[(\tilde X_\rho,X_0-X_\rho(h))_{L^2}+(\tilde X_\eta,Y_0-X_\eta(h))_{L^2}\right],$$

for all $(\tilde X_\rho,\tilde X_\eta)\in\partial\mathfrak{F}[(X_0,Y_0)]$. Hence, we easily get that $\dot Z(0)$ is bounded in $L_2(0,1)^2$ by the estimate

$$\|\dot Z(0)\|_{L_2^2}\le 2\|\partial^0\mathfrak{F}[(X_0,Y_0)]\|_{L_2^2}.$$

Finally, the stability property (54) implies for all $t,h>0$

$$\|Z(t+h)-Z(t)\|_{L_2^2}\le\|Z(h)-Z(0)\|_{L_2^2}.$$

Upon dividing by h and letting $h\downarrow 0$ we get

$$\|\dot Z(t)\|_{L_2^2}\le\|\dot Z(0)\|_{L_2^2},$$

which gives the desired regularity for $Z(t)$. □

4. Steady states and minimisers of the energy

In what follows we shall study the energy (4) associated to system (3). First we shall see that the energy can be written in a completely symmetric way which then allows us to show its boundedness from below by zero.

$$\begin{aligned}\mathcal{F}(\rho,\eta)&=-\frac12\int_{\mathbb{R}}\rho N\star\rho\,\mathrm{d}x-\frac12\int_{\mathbb{R}}\eta N\star\eta\,\mathrm{d}x+\int_{\mathbb{R}}\rho N\star\eta\,\mathrm{d}x\\&=-\frac12\iint_{\mathbb{R}^2}N(x-y)\big[\rho(x)\rho(y)-2\rho(x)\eta(y)+\eta(x)\eta(y)\big]\,\mathrm{d}y\,\mathrm{d}x\\&=-\frac12\iint_{\mathbb{R}^2}N(x-y)\big[\rho(x)[\rho(y)-\eta(y)]-[\rho(x)+\eta(x)]\eta(y)\big]\,\mathrm{d}y\,\mathrm{d}x.\end{aligned}$$

As the kernel, $N(x)=|x|$, is symmetric we may swap the roles of x and y in the second term in the integral to obtain

$$\begin{aligned}\mathcal{F}(\rho,\eta)&=-\frac12\iint_{\mathbb{R}^2}N(x-y)\big[\rho(x)[\rho(y)-\eta(y)]-[\rho(y)+\eta(y)]\eta(x)\big]\,\mathrm{d}y\,\mathrm{d}x\\&=-\frac12\int_{\mathbb{R}}(\rho-\eta)N\star(\rho-\eta)\,\mathrm{d}x,\end{aligned}$$

hence the energy does not depend on the individual densities but merely on their difference. By abuse of notation we shall write

$$\mathcal{F}(\kappa):=-\frac12\int_{\mathbb{R}}\kappa N\star\kappa\,\mathrm{d}x,$$

for $\kappa\in L^1((1+|x|^2)\,\mathrm{d}x)$ with zero mean. We introduce the set of L^1-functions with finite second order moments with zero mean

$$L_0^1:=\left\{f\in L^1\big((1+|x|^2)\,\mathrm{d}x\big)\;\middle|\;\int_{\mathbb{R}}f\,\mathrm{d}x=0\right\},$$

in order to formulate the following Proposition establishing the boundedness of the energy functional.

Proposition 7 (Characterisation of energy minimisers – 1)**.** *There holds*

$$\mathcal{F}(\kappa) \geq 0,$$

for any $\kappa \in L^1_0\big((1+|x|^2)\,\mathrm{d}x\big)$*. Moreover* $\mathcal{F}=0$ *if and only if* $\kappa = 0$ *almost everywhere.*

Proof. Let $\kappa \in L^1\big((1+|x|^2)\,\mathrm{d}x\big)$ be arbitrary. It is well-known that

$$\kappa = \delta \star \kappa = \frac{1}{2} N'' \star \kappa,$$

where the last equality holds due to the fact that $N/2$ is the fundamental solution of the Laplace equation in one dimension. Thus we may write

$$\begin{aligned}
\mathcal{F}(\kappa) &= -\frac{1}{2}\int_{\mathbb{R}} \kappa N \star \kappa \,\mathrm{d}x \\
&= -\frac{1}{2}\int_{\mathbb{R}} \delta \star \kappa N \star \kappa \,\mathrm{d}x \\
&= -\frac{1}{4}\int_{\mathbb{R}} N'' \star \kappa N \star \kappa \,\mathrm{d}x \\
&= -\left[\frac{1}{4} N' \star \kappa \, N \star \kappa\right]_{x=-\infty}^{x=+\infty} + \frac{1}{4}\int_{\mathbb{R}} |N' \star \kappa|^2 \,\mathrm{d}x,
\end{aligned}$$

by an integration by parts. Arguing as in the proof of Proposition 11, the boundary term vanishes. Hence we conclude

$$\mathcal{F}(\kappa) = \frac{1}{4}\int_{\mathbb{R}} \big|N' \star \kappa\big|^2 \,\mathrm{d}x \geq 0.$$

Clearly, equality holds if and only if $N' \star \kappa = 0$. Differentiating this expression once yields the second assertion and concludes the proof. □

Notice that the energy is not bounded from below in case of different masses for ρ and η as specified in Remark 10.

Proposition 8 (Energy minimisers are steady states)**.** *A pair* $(\rho,\eta) \in \mathcal{P}_2^a(\mathbb{R}) \times \mathcal{P}_2^a(\mathbb{R})$ *is an energy minimiser if and only if it is a steady states of system* (6)*.*

Proof. In view of Proposition 7 we know that (ρ,η) is an energy minimiser if and only if $\rho = \eta$ almost everywhere. Hence, Proposition 4 gives $v_i = -\left(\partial^0 \mathcal{F}[\rho,\eta]\right)_i = 0$ for $i = 1,2$ on $\mathrm{supp}(\rho)$, whence we conclude that (ρ,η) is a stationary state. Now, assume that (ρ,η) is a stationary state of system (6). From the energy inequality (53) we get

$$\frac{1}{2}\int_0^T \|v(t)\|^2_{L^2(\gamma)}\,\mathrm{d}t + \frac{1}{2}\int_0^T |\partial\mathcal{F}|^2[\gamma(t)]\,\mathrm{d}t \leq 0,$$

which implies $|\partial\mathcal{F}|^2[\gamma] = 0$. Thus, in view of Proposition 4 we obtain $\rho = \eta$ almost everywhere. As a consequence of Proposition 7 we conclude (ρ,η) is a minimiser of the energy $\mathcal{F}$. □

We conclude this section by providing a characterisation for the ω-limit set of a solution to system (6). For the sake of completeness, let us recall the definition of ω-limit set according to [31, Definition 9.1.5].

Definition 13. *Let (X,d) be a complete metric space and consider a dynamical system $\{S_t\}_{t\geq 0}$. For $x \in X$ the set*

$$\omega(x) = \{y \in X \,|\, \exists\, t_n \to \infty \text{ s.t. } S_{t_n}[x] \to y, \text{ as } n \to \infty\}$$

is called $\omega-$limit set of x.

Now, let us state the following Theorem.

Theorem 14. *Let $\gamma = (\rho, \eta)$ be the solution to system (6) with initial datum $\gamma_0 = (\rho_0, \eta_0) \in (\mathcal{P}_2^a(\mathbb{R}) \cap L^m(\mathbb{R}))^2$. Then*

$$\omega(\gamma) \subseteq \{(\rho,\eta) \in (\mathcal{P}_2^a(\mathbb{R}) \cap L^m(\mathbb{R}))^2 | \rho = \eta \ a.e.\}.$$

Proof. Since $\gamma_0 = (\rho_0, \eta_0) \in (\mathcal{P}_2^a(\mathbb{R}) \cap L^m(\mathbb{R}))^2$ from Proposition 6 we know

$$\|\gamma\|_{L^\infty(0,+\infty;L^m(\mathbb{R}))^2} \leq \|\rho_0\|_{L^m(\mathbb{R})} + \|\eta_0\|_{L^m(\mathbb{R})},$$

whence

$$\|\gamma(t)\|_{L^m(\mathbb{R})^2} \leq \|\rho_0\|_{L^m(\mathbb{R})} + \|\eta_0\|_{L^m(\mathbb{R})},$$

for a.e. $t > 0$. Then we can consider an unbounded, increasing sequence $\{t_n\}_{n\in\mathbb{N}}$ such that $t_n \to +\infty$ and $\gamma(t_n) \rightharpoonup \tilde{\gamma}$ weakly in L^m, as $n \to +\infty$, where $\tilde{\gamma} = (\tilde{\rho}, \tilde{\eta})$. According to [2, Theorem 5.3], since γ is a gradient flow solution to system (6) we have

$$-\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{F}(\gamma(t)) = \|v_t(t)\|_{L^2}^2 = |\partial\mathcal{F}|^2[\gamma(t)],$$

for a. e. $t > 0$, whence

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{F}(\gamma(t)) = -|\partial\mathcal{F}|^2[\gamma(t)] \leq 0.$$

Now, if we integrate in a general time interval $(t_n, t_n + 1)$, we have

$$\int_{t_n}^{t_n+1} \left(-\frac{d}{ds}\mathcal{F}(\gamma(s))\right) \,\mathrm{d}s = \int_0^1 \left(-\frac{d}{ds}\mathcal{F}(\gamma(s+t_n))\right) \,\mathrm{d}s = \int_0^1 |\partial\mathcal{F}|^2[\gamma(s+t_n)] \,\mathrm{d}s,$$

which gives, passing to the lim inf as $n \to +\infty$,

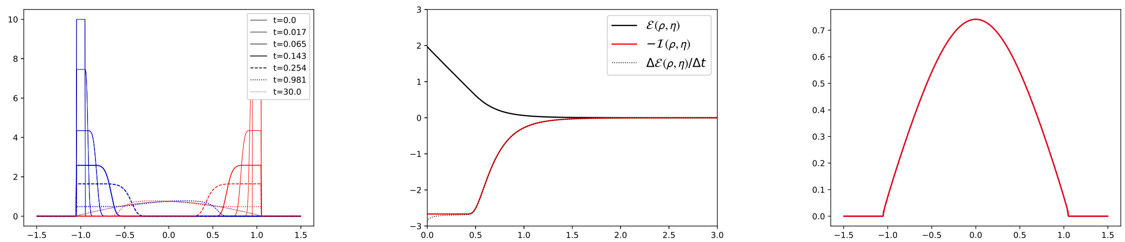
$$0 = \liminf_{n\to+\infty} \int_0^1 |\partial\mathcal{F}|^2[\gamma(s+t_n)] \,\mathrm{d}s \geq \int_0^1 |\partial\mathcal{F}|^2[\tilde{\gamma}] \,\mathrm{d}s = |\partial\mathcal{F}|^2[\tilde{\gamma}], \tag{56}$$

by means of the lower semi-continuity of the slope already used in Subsection 3.2 (cf. [25, Lemma 2.7]). Hence, as a trivial consequence of Eq. (56) we get $|\partial\mathcal{F}|^2[\tilde{\gamma}] = 0$, which, according to Proposition 4, implies that $\partial^0 N \star (\tilde{\rho} - \tilde{\eta}) = 0$ almost everywhere. Thus by differentiating we obtain the result. □

Remark 10. *We specify that in case ρ and η have masses $M_\rho \neq M_\eta$ the results in this section are no longer valid as the energy $\mathcal{F}$ is no longer bounded from below. In fact, let us assume for instance $\eta = \beta\rho$ which implies $M_\eta = \beta M_\rho$. Hence the energy becomes*

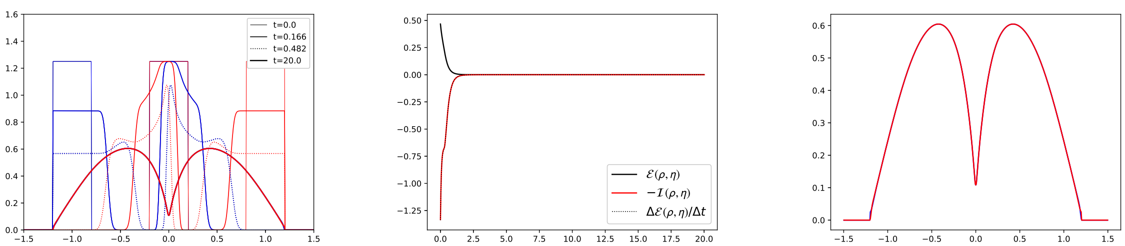
$$\mathcal{F}(\rho,\eta) = -\frac{1}{2}M_\rho(\beta^2 - 1)^2 \int_{\mathbb{R}} N \star \rho \, d\rho.$$

By a simple rescaling argument the energy is shown to be unbounded from below.



(a) Evolution. (b) Energy dissipation. (c) Steady state.

FIGURE 1. This example has two separated indicator functions as initial data. In the left graph we see the evolution of system (6) to the stationary state (right graph). In the middle we see the energy (black) of the solution and its dissipation (red). The dotted line is the numerical time derivative of the energy. It matches well with the analytically obtained dissipation.



(a) Evolution. (b) Energy dissipation. (c) Steady state.

FIGURE 2. We choose partially overlapping initial data and observe, as before that mixing occurs. The graph on the left displays the evolution of both densities at different time instances, while the rightmost graph displays the stationary state with identical densities. The graph in the middle shows the energy decay along the solution and the numerical dissipation and the analytical dissipation agree well.

5. Initial Data with Atomic Part, Non-Uniqueness & link with Hyperbolic Systems

In this section we study the evolution of measure valued initial data consisting of atomic parts. We will consider two peculiar examples: first we handle the case of two distinct Dirac deltas as initial condition, namely $\rho_0 = \delta_{-1}$ and $\eta_0 = \delta_1$, and then we will consider the case $\rho_0 = \delta_0$ and $\eta_0 = m\delta_0 + (1-m)\delta_1$, for some $m \geq 0$. In both cases we will provide two candidate weak solutions for system (3) and we will show that only one of them can be selected in the spirit of Definition 7. As the notion of gradient flow in measure valued solution setting is essentially formulated in the pseudo-inverse formalism through Definition 5, in this section we shall work directly with the pseudo-inverse variables.

We start by providing an explicit expression for elements in the sub-differential of $\mathfrak{F}$ in both the examples we shall consider.

Proposition 9. *Let $m \in [0,1]$. Let $(X,Y) \in \mathcal{C} \times \mathcal{C}$ be such that $X = Y$ on $[0,m)$ and*

$$\sup_{z \in [m,1]} X(z) < \inf_{z \in [m,1]} Y(z).$$

Then,

$$\begin{aligned} K(X,Y) = &\int_0^1 [(2z-1)\mathbf{1}_{[0,m)}(z) + (2m-1)\mathbf{1}_{[m,1]}(z)] X(z) \; \mathrm{d}z \\ &+ \int_0^1 [(2z-1)\mathbf{1}_{[0,m)}(z) + \mathbf{1}_{[m,1]}(z)] Y(z) \; \mathrm{d}z, \end{aligned}$$

and in particular the functional becomes

$$\begin{aligned} \mathfrak{F}(X,Y) = &\int_0^1 [1 - 2z + (2z-1)\mathbf{1}_{[0,m)}(z) + (2m-1)\mathbf{1}_{[m,1]}(z)] X(z) \; \mathrm{d}z \\ &+ \int_0^1 [1 - 2z + (2z-1)\mathbf{1}_{[0,m)}(z) + \mathbf{1}_{[m,1]}(z)] Y(z) \; \mathrm{d}z, \end{aligned} \tag{57}$$

In addition, let $X, Y \in L^2(0,1)$. Then, $\partial\bar{\mathfrak{F}}[(X,Y)] \neq \emptyset$ if and only if $(X,Y) \in \mathcal{C} \times \mathcal{C}$. In that case, if $(f_1, f_2) \in \partial\bar{\mathfrak{F}}[(X,Y)]$ then

$$f_1(z) = \begin{cases} 0, & z \in [0,m), \\ 2(m-z), & z \in [m,1], \end{cases} \qquad \text{and} \qquad f_2(z) = \begin{cases} 0, & z \in [0,m), \\ 2 - 2z, & z \in [m,1], \end{cases}.$$

Proof. Letting $(X,Y) \in \mathcal{C} \times \mathcal{C}$ as in the statement we have $\mathcal{I}_{\mathcal{C}}(X) = \mathcal{I}_{\mathcal{C}}(Y) = 0$ and

$$\begin{aligned} K(X,Y) &= \int_0^1 \int_0^1 |Y(\xi) - X(z)| \; \mathrm{d}z \, d\xi \\ &= \int_0^m \int_0^m |Y(\xi) - X(z)| \; \mathrm{d}z \; \mathrm{d}\xi + \int_0^m \int_m^1 |Y(\xi) - X(z)| \; \mathrm{d}z \; \mathrm{d}\xi \\ &\quad + \int_m^1 \int_0^m |Y(\xi) - X(z)| \; \mathrm{d}z \; \mathrm{d}\xi + \int_m^1 \int_m^1 |Y(\xi) - X(z)| \; \mathrm{d}z \; \mathrm{d}\xi \\ &=: I_1 + I_2 + I_3 + I_4. \end{aligned}$$

First, let us compute I_1 taking into account that $X = Y$ on $[0,m)$:

$$\begin{aligned} I_1 &= \int_0^m \int_0^m |Y(\xi) - X(z)| \; \mathrm{d}z \; \mathrm{d}\xi \\ &= \int_0^m \int_0^m |X(\xi) - X(z)| \; \mathrm{d}z \; \mathrm{d}\xi \\ &= \int\int_{[0,m]^2 \cap \{X(\xi) \geq X(z)\}} (X(\xi) - X(z)) \; \mathrm{d}z \; \mathrm{d}\xi - \int\int_{[0,m]^2 \cap \{X(\xi) \leq X(z)\}} (X(\xi) - X(z)) \; \mathrm{d}z \; \mathrm{d}\xi \\ &= 2 \int\int_{\{X(\xi) \geq X(z)\}} (X(\xi) - X(z)) \; \mathrm{d}z \; \mathrm{d}\xi. \end{aligned}$$

Since X is non-decreasing, we have

$$\{X(\xi) \geq X(z)\} = \{\xi \geq z\} \cup \{\xi \leq z \leq S(\xi)\}, \quad \textit{where} \quad S(\xi) = \sup\{z \in [0,1] \,|\, X(z) = X(\xi)\},$$

and therefore we get

$$\begin{aligned} I_1 =&2\int\int_{[0,m]^2\cap\{X(\xi)\geq X(z)\}}(X(\xi)-X(z))\ \mathrm{d}z\ \mathrm{d}\xi \\ =&2\int\int_{[0,m]^2\cap\{\xi\geq z\}}(X(\xi)-X(z))\ \mathrm{d}z\ \mathrm{d}\xi \\ =&\int_0^m (2z-m)(X(z)+Y(z))\ \mathrm{d}z. \end{aligned}$$

Concerning the other integrals, we easily obtain

$$\begin{aligned} I_2 =&m\int_m^1 Y(z)\ \mathrm{d}z-(1-m)\int_0^m X(z)\ \mathrm{d}z, \\ I_3 =&m\int_m^1 X(z)\ \mathrm{d}z-(1-m)\int_0^m Y(z)\ \mathrm{d}z, \\ I_4 =&(1-m)\int_m^1 (Y(z)-X(z))\ \mathrm{d}z. \end{aligned}$$

Summing up all the contributions we have

$$\begin{aligned} K(X,Y) =& \int_0^1 [(2z-1)\mathbf{1}_{[0,m)}(z)+(2m-1)\mathbf{1}_{[m,1]}(z)]X(z)\ \mathrm{d}z \\ &+\int_0^1 [(2z-1)\mathbf{1}_{[0,m)}(z)+\mathbf{1}_{[m,1]}(z)]Y(z)\ \mathrm{d}z. \end{aligned}$$

Then (57) follows as a direct consequence. Now, let us characterise the sub-differential of $\bar{\mathfrak{F}}$. Assume without loss of generality that $X\notin\mathcal{C}$ and $Y\in\mathcal{C}$, which implies $\mathcal{I}_{\mathcal{C}}(X)=+\infty$ and $\mathcal{I}_{\mathcal{C}}(Y)=0$. If $(\tilde{X}_1,\tilde{X}_2)\in\partial\bar{\mathfrak{F}}[(X,Y)]$ we would have

$$\begin{aligned} &\mathfrak{F}(R_1,R_2)+\mathcal{I}_{\mathcal{C}}(R_1)+\mathcal{I}_{\mathcal{C}}(R_2)-\mathfrak{F}(X,Y) \\ &-\int_0^1 \tilde{X}_1(z)(R_1(z)-X(z))+\tilde{X}_2(z)(R_2(z)-Y(z))\ \mathrm{d}z+o(\|(X,Y)-(R_1,R_2)\|)\geq\mathcal{I}_{\mathcal{C}}(X), \end{aligned} \tag{58}$$

for all $(R_1,R_2)\in L^2(0,1)^2$ with $\|(X,Y)-(R_1,R_2)\|\to 0$, and in particular for all $(R_1,R_2)\in\mathcal{C}\times\mathcal{C}$. In the latter case we obviously have a contradiction because the left-hand side is finite while the right-hand side is infinite and therefore $\partial\bar{\mathfrak{F}}[(X,Y)]=\emptyset$. Let $(X,Y)\in\mathcal{C}\times\mathcal{C}$ and $(R_1,R_2)\in L^2(0,1)^2$. Now we have to consider two cases:

(1) $(R_1,R_2)\notin\mathcal{C}\times\mathcal{C}$;
(2) $(R_1,R_2)\in\mathcal{C}\times\mathcal{C}$.

In the first case the definition of sub-differential is trivially satisfied, whereas in the second one we get

$$\begin{aligned} \mathfrak{F}(R_1,R_2)-\mathfrak{F}(X,Y) =& \int_0^1 [1-2z+(2z-1)\mathbf{1}_{[0,m)}(z)+(2m-1)\mathbf{1}_{[m,1]}(z)]\,[R_1(z)-X(z)]\ \mathrm{d}z \\ &+\int_0^1 [1-2z+(2z-1)\mathbf{1}_{[0,m)}(z)+\mathbf{1}_{[m,1]}(z)]\,[R_2(z)-Y(z)]\ \mathrm{d}z, \end{aligned}$$

which concludes the proof. □

We remark in particular that under the assumptions in the previous proposition the sub-differential of $\mathfrak{F}$ is *single-valued* on $\mathcal{C}\times\mathcal{C}$. Note that the case $\sup X < \inf Y$ $(m=0)$ is included in the previous proposition, and the functional becomes

$$\mathfrak{F}(X,Y) = 2\int_0^1 (1-z)Y(z)\,\mathrm{d}z - 2\int_0^1 zX(z)\,\mathrm{d}z. \tag{59}$$

5.1. **The case of two distinct deltas as initial condition.** Let us consider the first example, with

$$\rho_0 = \delta_{-1} \qquad \text{and} \qquad \eta_0 = \delta_1. \tag{60}$$

At the level of weak (measure) solutions, both

$$\rho(t,x) = \frac{1}{2t}\mathbf{1}_{[-1,-1+2t]}(x), \qquad \text{and} \qquad \eta(t,x) = \frac{1}{2t}\mathbf{1}_{[1-2t,1]}, \tag{61}$$

and $(\tilde\rho,\tilde\eta)$ given by

$$\tilde\rho(t,x) = \delta_{t-1}, \quad \text{and} \quad \tilde\eta(t,x) = \delta_{1-t}, \tag{62}$$

satisfy system (6) in the weak sense on $[0,1/2]\times\mathbb{R}$ and equal (60) at $t=0$, see considerations in Subsection 5.1.1. The corresponding pseudo-inverse functions (X,Y), given by

$$X(t,z) = -1+2zt, \qquad \text{and} \qquad Y(t,z) = 1+t(2z-2), \tag{63}$$

as well as $(\tilde X,\tilde Y)$, given by

$$\tilde X(t,z) = -1+t, \qquad \text{and} \qquad \tilde Y(t,z) = 1-t, \tag{64}$$

satisfy system (16) in the strong sense on $[0,1/2]\times[0,1]$, see Subsection 5.1.2. Working in the context of pseudo-inverses we can show, in Subsection 5.1.3 that actually only the time derivative of the pseudo-inverse in (63) is an element of the sub-differential.

5.1.1. *Weak measure solutions.* In order to prove that (ρ,η) in (61) is a weak solution of the system we begin by simplifying the velocity term. To this end we compute the convolution with the Heaviside function

$$\begin{aligned}(N'\star\rho)(t,x) &= \int_{-\infty}^{+\infty} \operatorname{sign}(x-y)\rho(t,y)\,\mathrm{d}y \\ &= \begin{cases} -1, & \text{if } x\le -1,\\ \dfrac{x+1-t}{t}, & \text{if } x\in(-1,-1+2t],\\ 1, & \text{else.}\end{cases}\end{aligned}$$

Similarly, for the convolution with the second species we obtain

$$\begin{aligned}(N'\star\eta)(t,x) &= \int_{-\infty}^{+\infty} \operatorname{sign}(x-y)\eta(t,y)\,\mathrm{d}y \\ &= \begin{cases} -1, & \text{if } x\le 1-2t,\\ \dfrac{x-1+t}{t}, & \text{if } x\in(1-2t,1],\\ 1, & \text{else.}\end{cases}\end{aligned}$$

We claim that ρ, η as defined above are weak solutions to system (6). Here we only check that ρ is a weak solution as the computation for the second species is done in an analogous way. Now, let $\phi \in C_c^\infty$ and consider the weak formulation

$$\int_0^{1/2} \int_{-\infty}^{\infty} \rho \Big[\phi_t - N' \star (\eta - \rho)\phi_x \Big] \,\mathrm{d}x\,\mathrm{d}t =: I_t + I_x,$$

where

$$I_t := \int_0^{1/2} \int_{-\infty}^{\infty} \rho\phi_t(t,x)\,\mathrm{d}x\,\mathrm{d}t,$$

and

$$I_x := -\int_0^{1/2} \int_{-\infty}^{\infty} \rho N' \star (\eta - \rho)\phi_x(t,x)\,\mathrm{d}x\,\mathrm{d}t.$$

Let us begin by simplifying the time related term. By changing the order of integration it is easy to see that

$$\begin{aligned} I_t &= \int_0^{1/2} \int_{-\infty}^{\infty} \rho\phi_t(t,x)\,\mathrm{d}x\,\mathrm{d}t = \int_0^{1/2} \int_{-1}^{-1+2t} \frac{1}{2t}\phi_t(t,x)\,\mathrm{d}t\,\mathrm{d}x \\ &= \int_{-1}^{0} \int_{\frac{x+1}{2}}^{1/2} \frac{1}{2t^2}\phi(t,x)\,\mathrm{d}t\,\mathrm{d}x + \int_{-1}^{0} \left[\frac{1}{2t}\phi(t,x)\right]_{t=\frac{x+1}{2}}^{1/2} \mathrm{d}x, \end{aligned}$$

by an integration by parts. Hence, switching the order of integration another time and simplifying the boundary term we obtain

$$I_t = \int_0^{1/2} \int_{-1}^{-1+2t} \frac{1}{2t^2}\phi(t,x)\,\mathrm{d}x\,\mathrm{d}t - \int_{-1}^{0} \frac{1}{x+1}\phi\left(\frac{x+1}{2}, x\right)\,\mathrm{d}x.$$

A change of variables $x + 1 = 2t$ finally yields

$$I_t = \int_0^{1/2} \int_{-1}^{-1+2t} \frac{1}{2t^2}\phi(t,x)\,\mathrm{d}x\,\mathrm{d}t - \int_0^{1/2} \frac{1}{t}\phi\left(t, 2t-1\right)\,\mathrm{d}t. \tag{65}$$

Next we shall address the term space related term. We observe

$$\begin{aligned} I_x &= -\int_0^{1/2} \int_{-1}^{-1+2t} \frac{1}{2t}\left[-1 - \frac{x+1-t}{t}\right]\phi_x(t,x)\,\mathrm{d}x\,\mathrm{d}t \\ &= \int_0^{1/2} \int_{-1}^{-1+2t} \frac{x+1}{2t^2}\phi_x(t,x)\,\mathrm{d}x\,\mathrm{d}t. \end{aligned}$$

An integration by parts yields

$$\begin{aligned} I_x &= \int_0^{1/2} \int_{-1}^{-1+2t} \frac{x+1}{2t^2}\phi_x(t,x)\,\mathrm{d}x\,\mathrm{d}t \\ &= -\int_0^{1/2} \int_{-1}^{-1+2t} \frac{1}{2t^2}\phi(t,x)\,\mathrm{d}x\,\mathrm{d}t + \int_0^{1/2} \left[\frac{x+1}{2t^2}\phi(t,x)\right]_{x=-1}^{-1+2t} \mathrm{d}t \\ &= -\int_0^{1/2} \int_{-1}^{-1+2t} \frac{1}{2t^2}\phi(t,x)\,\mathrm{d}x\,\mathrm{d}t + \int_0^{1/2} \frac{1}{t}\phi(t, 2t-1)\,\mathrm{d}t. \end{aligned}$$

Upon adding up I_t and I_x we observe

$$I_t + I_x = 0,$$

i.e. ρ is a weak solution of the first equation in system (6). Similarly it can be shown that η is a weak solution to the second equation in (6).

Next we show that $(\tilde{\rho}, \tilde{\eta})$ is also a weak solution. As before we compute the terms including the convolutions first. It is easy to check that

$$(N' \star \tilde{\rho})(x) = \operatorname{sign}(x - t + 1), \qquad \text{and} \qquad (N' \star \tilde{\eta})(x) = \operatorname{sign}(x + t - 1),$$

for all $x \in \mathbb{R}$, thus the velocity is given by $u := N' \star (\tilde{\eta} - \tilde{\rho}) = \operatorname{sign}(x + t - 1) - \operatorname{sign}(x - t + 1)$. Let us consider a test function $\phi \in C_c^\infty$ in order to check the weak formulation as follows:

$$\begin{aligned}
\int_0^1 \int_{\mathbb{R}} \tilde{\rho}[\phi_t - u\phi_x] \, \mathrm{d}x \, \mathrm{d}t &= \int_0^1 \phi_t(t, t-1) - u(t, t-1)\phi_x(t, t-1) \, \mathrm{d}t \\
&= \int_0^1 \phi_t(t, t-1) + \phi_x(t, t-1) \, \mathrm{d}t \\
&= \int_0^1 \frac{\mathrm{d}}{\mathrm{d}t} \left(\phi(t, t-1) \right) \, \mathrm{d}t \\
&= 0.
\end{aligned}$$

Arguing similarly for $\tilde{\eta}$ we have $(\tilde{\rho}, \tilde{\eta})$ is a weak solution to system (6) with initial data $\rho_0 = \delta_{-1}$ and $\eta_0 = \delta_1$ as well.

5.1.2. *Strong solutions in the pseudo-inverse formalism.* Next, let us show that (X, Y) defined in (63) is the solution to system (16) in the strong sense. Using that $t < 1/2$ there holds

$$\begin{aligned}
&\int_0^1 \operatorname{sign}(X(t,z) - X(t,\xi)) \, \mathrm{d}\xi - \int_0^1 \operatorname{sign}(X(t,z) - Y(t,\xi) \, \mathrm{d}\xi \\
&= \int_0^1 \operatorname{sign}(2t(z - \xi)) \, \mathrm{d}\xi - \int_0^1 \operatorname{sign}(-2 + 2t(z - \xi + 1)) \, \mathrm{d}\xi \\
&= \int_0^z \mathrm{d}\xi - \int_z^1 \mathrm{d}\xi + 1 \\
&= 2z \\
&= \frac{\partial}{\partial t} X(t,z),
\end{aligned}$$

and

$$\begin{aligned}
&\int_0^1 \operatorname{sign}(Y(t,z)-Y(t,\xi))\,\mathrm{d}\xi - \int_0^1 \operatorname{sign}(Y(t,z)-X(t,\xi)\,\mathrm{d}\xi \\
&= \int_0^1 \operatorname{sign}(2t(z-\xi))\,\mathrm{d}\xi - \int_0^1 \operatorname{sign}(2+2t(z-\xi-1))\,\mathrm{d}\xi \\
&= \int_0^z \mathrm{d}\xi - \int_z^1 \mathrm{d}\xi - 1 \\
&= 2z-2 \\
&= \frac{\partial}{\partial t} Y(t,z).
\end{aligned}$$

As for the second pair of pseudo–inverses (64) we observe

$$\begin{aligned}
&\int_0^1 \operatorname{sign}(\tilde{X}(t,z)-\tilde{X}(t,\xi))\,\mathrm{d}\xi - \int_0^1 \operatorname{sign}(\tilde{X}(t,z)-\tilde{Y}(t,\xi)\,\mathrm{d}\xi \\
&= \int_0^1 \operatorname{sign}(0)\,\mathrm{d}\xi - \int_0^1 \operatorname{sign}(2(t-1))\,\mathrm{d}\xi \\
&= 1 \\
&= \frac{\partial}{\partial t} \tilde{X}(t,z),
\end{aligned}$$

and similarly the equation for $\tilde{Y}$ is satisfied.

5.1.3. *Characterisation of the sub-differential.* We need to check the differential inclusion. According to Proposition 9 we have

$$\frac{\partial}{\partial t}\begin{pmatrix} X(t,z) \\ Y(t,z) \end{pmatrix} = \begin{pmatrix} 2z \\ 2z-2 \end{pmatrix} \in -\partial\bar{\mathfrak{F}}\big[\big(X(t,z),Y(t,z)\big)\big],$$

as we claimed. However, there holds $\frac{\partial}{\partial t}\tilde{X}(t,z) = 1 \neq 2z$, which shows the pair $(\tilde{X},\tilde{Y})$ defined in (64) is not a solution to system (16) in the sense of Definition 5 for the given initial data.

Remark 11. *In this example, both species are initially concentrated at one point, but there is no overlap between them. Therefore, only the intraspecific energies are affected at "singular points", i.e. at points in which the convolution kernel is not smooth. In fact, the interspecific energy is not effected by the Lipschitz point at the origin. In this sense, one expects the qualitative behaviour of this system to be essentially the same as in the one species case, see [12]. More precisely, both species get immediately absolutely continuous w.r.t. to the Lebesgue measure. The attractive cross-interaction energy makes the two patches get closer to each other until they eventually merge. In conclusion, the existence of two distinct measure solutions in this example is not a distinctive feature of the two species system, but rather an extension of a property holding in the one species case.*

5.2. **The case of two overlapping deltas as initial condition.** Let $0 \le m < 1$ be given and initialise system (6) as follows

$$\rho_0 = \delta_0, \quad \text{and} \quad \eta_0 = m\delta_0 + (1-m)\delta_1. \tag{66}$$

Then the pair (ρ,η) given by

$$\rho(t,x) = m\delta_0 + \frac{1}{2t}\mathbf{1}_{[0,2(1-m)t]}(x), \qquad \eta(t,x) = m\delta_0 + \frac{1}{2t}\mathbf{1}_{[1-2(1-m)t,1]}(x),$$

is a weak solution to system (6) on $[0,T)\times[0,1]$, with $T := \frac{1}{4(1-m)}$, as well as the pair $(\tilde\rho,\tilde\eta)$ given by

$$\tilde\rho(t,x) = m\delta_0 + (1-m)\delta_{(1-m)t}, \quad \text{and} \quad \tilde\eta(t,x) = m\delta_0 + (1-m)\delta_{1-(1-m)t}.$$

Moreover, for $t\in[0,T)$, the associated pseudo-inverse functions

$$X(t,z) = 2t(z-m)\mathbf{1}_{[m,1]}(z), \qquad Y(t,z) = \big(1-2t(1-z)\big)\mathbf{1}_{[m,1]}(z),$$

and

$$\tilde X(t,z) = (1-m)t\mathbf{1}_{[m,1]}(z), \qquad \tilde Y(t,z) = \big(1-(1-m)t\big)\mathbf{1}_{[m,1]}(z).$$

are both strong strong solutions to system (16), but only (X,Y) is the gradient flow solution in the sense of Definition 5.

5.2.1. *Weak measure solutions.* Let us start by verifying that (ρ,η) is a weak solution to system (6) on $[0,T)\times[0,1]$. Next we compute the vector field for ρ on $x\in[0,1/2]$. For the self-interactions we get

$$\begin{aligned}(N'\star\rho)(t,x) &= \int_{\mathbb{R}} \operatorname{sign}(x-y)\rho(t,y)\,\mathrm{d}y\\ &= m\operatorname{sign}(x) + \frac{1}{2t}\int_0^{2(1-m)t}\operatorname{sign}(x-y)\,\mathrm{d}y\\ &= m\operatorname{sign}(x) + \frac{2x-2(1-m)t}{2t}\\ &= m\operatorname{sign}(x) + \frac{x}{t} + m - 1,\end{aligned}$$

whereas, for the cross-interactions, we get

$$\begin{aligned}(N'\star\eta)(t,x) &= m\operatorname{sign}(x) - \int_{1-2(1-m)t}^1 \frac{1}{2t}\,\mathrm{d}y\\ &= m\operatorname{sign}(x) + m - 1.\end{aligned}$$

Hence, the velocity on $[0,1/2]$ is given by

$$u = N'\star(\eta-\rho) = -\frac{x}{t}.$$

We shall now verify that ρ and η are weak solutions. It is easy to see that

$$\begin{aligned}I &= \int_0^T\int_{\mathbb{R}} \rho(t,x)\Big[\phi_t(t,x) - u\phi_x(t,x)\Big]\,\mathrm{d}x\,\mathrm{d}t\\ &= m\int_0^T \phi_t(t,0) - u(t,0)\phi_x(t,0)\,\mathrm{d}t\\ &\quad + \underbrace{\int_0^T\int_0^{2(1-m)t}\frac{1}{2t}\phi_t(t,x)\,\mathrm{d}x\,\mathrm{d}t}_{=:I_1} + \underbrace{\int_0^T\int_0^{2(1-m)t}\frac{x}{2t^2}\phi_x(t,x)\,\mathrm{d}x\,\mathrm{d}t}_{=:I_2}.\end{aligned}$$

We will treat each term individually. Note that $u(t,0)=0$ for all $t>0$. Together with the fact that ϕ is compactly supported and the application of the fundamental theorem the first term vanishes and it remains to treat the terms I_1, I_2. Using Fubini's theorem and an integration by parts we may write

$$\begin{aligned}
I_1 &= \int_0^T \int_0^{2(1-m)t} \frac{1}{2t}\phi_t(t,x)\,\mathrm{d}x\,\mathrm{d}t \\
&= \int_0^{2(1-m)T} \int_{x/(2(1-m))}^{T} \frac{1}{2t}\phi_t(t,x)\,\mathrm{d}t\,\mathrm{d}x \\
&= \int_0^{2(1-m)T} \int_{x/(2(1-m))}^{T} \frac{1}{2t^2}\phi(t,x)\,\mathrm{d}t\,\mathrm{d}x - \int_0^{2(1-m)T} \frac{1-m}{x}\phi\left(\frac{x}{2(1-m)},x\right)\mathrm{d}x \\
&= \int_0^{2(1-m)T} \int_{x/(2(1-m))}^{T} \frac{1}{2t^2}\phi(t,x)\,\mathrm{d}t\,\mathrm{d}x - \int_0^{T} \frac{1-m}{t}\phi(t,2(1-m)t)\,\mathrm{d}t.
\end{aligned}$$

As for the second term a simple integration by parts yields

$$\begin{aligned}
I_2 &= \int_0^T \int_0^{2(1-m)t} \frac{x}{2t^2}\phi_x(t,x)\,\mathrm{d}x\,\mathrm{d}t \\
&= -\int_0^T \int_0^{2(1-m)t} \frac{1}{2t^2}\phi(t,x)\,\mathrm{d}x\,\mathrm{d}t + \int_0^T \frac{1-m}{t}\phi(t,2(1-m)t)\,\mathrm{d}t.
\end{aligned}$$

Thus we get

$$I = I_1 + I_2 = 0.$$

Now, we need to check that $(\tilde\rho,\tilde\eta)$ satisfies the weak formulation. Here we only check the statement for $\tilde\rho$ as the second species is shown analogously. Again we compute the velocity field for $\tilde\rho$.

$$\begin{aligned}
u(t,x) &= N' \star (\tilde\eta - \tilde\rho)(t,x) \\
&= (1-m)\big(\mathrm{sign}(x-1+(1-m)t) - \mathrm{sign}(x-(1-m)t)\big).
\end{aligned}$$

Note that $u(t,0)=0$ and $u(t,(1-m)t) = -(1-m)$. Thus there holds

$$\begin{aligned}
&\int_0^T \int_{\mathbb{R}} \tilde\rho\Big[\phi_t(t,x) - u(t,x)\phi_x(t,x)\Big]\,\mathrm{d}x\,\mathrm{d}t \\
&= m\int_0^T \phi_t(t,0) + 0\phi_x(t,0)\,\mathrm{d}t \\
&\quad + (1-m)\int_0^T \phi_t(t,(1-m)t) + (1-m)\phi_x(t,(1-m)t)\,\mathrm{d}t \\
&= \int_0^T m\frac{\mathrm{d}}{\mathrm{d}t}\phi(t,0) + (1-m)\frac{\mathrm{d}}{\mathrm{d}t}\phi(t,(1-m)t)\,\mathrm{d}t \\
&= 0,
\end{aligned}$$

457 by the fundamental theorem of calculus.

5.2.2. *Strong solutions in the pseudo-inverse formalism.* Let us now consider the associated pseudo-inverse functions X and Y, given by

$$X(t,z) = 2t(z-m)\mathbf{1}_{[m,1]}(z), \quad \text{and} \quad Y(t,z) = \big(1-2t(1-z)\big)\mathbf{1}_{[m,1]}(z),$$

for $z \in [0,1]$ and $0 \le t < T$. (X,Y) is a solution to system (16) in a strong sense, in fact

$$\begin{aligned}
&\int_0^1 \operatorname{sign}(X(z)-X(\xi))\,\mathrm{d}\xi - \int_0^1 \operatorname{sign}(X(z)-Y(\xi))\,\mathrm{d}\xi \\
&= \int_0^1 \operatorname{sign}\Big(2t(z-m)\mathbf{1}_{[m,1]}(z) - 2t(\xi-m)\mathbf{1}_{[m,1]}(\xi)\Big)\,\mathrm{d}\xi \\
&\quad - \int_0^1 \operatorname{sign}\Big(2t(z-m)\mathbf{1}_{[m,1]}(z) - \big(1-2t(1-\xi)\big)\mathbf{1}_{[m,1]}(\xi)\Big)\,\mathrm{d}\xi \\
&= (m-1)\mathbf{1}_{[0,m]}(z) + (2z-1)\mathbf{1}_{[m,1]}(z) + (1-m)\mathbf{1}_{[0,m]}(z) + (1-2m)\mathbf{1}_{[m,1]}(z) \\
&= 2(z-m)\mathbf{1}_{[m,1]}(z) \\
&= \frac{\partial}{\partial t}X(t,z),
\end{aligned}$$

and

$$\begin{aligned}
&\int_0^1 \operatorname{sign}(Y(z)-Y(\xi))\,\mathrm{d}\xi - \int_0^1 \operatorname{sign}(Y(z)-X(\xi))\,\mathrm{d}\xi \\
&= \int_0^1 \operatorname{sign}\Big(\big(1-2t(1-z)\big)\mathbf{1}_{[m,1]}(z) - \big(1-2t(1-\xi)\big)\mathbf{1}_{[m,1]}(\xi)\Big)\,\mathrm{d}\xi \\
&\quad - \int_0^1 \operatorname{sign}\Big(\big(1-2t(1-z)\big)\mathbf{1}_{[m,1]}(z) - 2t(\xi-m)\mathbf{1}_{[m,1]}(\xi)\Big)\,\mathrm{d}\xi \\
&= (m-1)\mathbf{1}_{[0,m]}(z) + (2z-1)\mathbf{1}_{[m,1]}(z) + (1-m)\mathbf{1}_{[0,m]}(z) - \mathbf{1}_{[m,1]}(z) \\
&= 2(z-1)\mathbf{1}_{[m,1]}(z) \\
&= \frac{\partial}{\partial t}Y(t,z),
\end{aligned}$$

as we claimed. Moreover, the pair $(\tilde{X},\tilde{Y})$ of pseudo-inverses associated to the moving Diracs, i.e.,

$$\tilde{X}(t,z) = (1-m)t\mathbf{1}_{[m,1]}(z), \qquad \tilde{Y}(t,z) = \big(1-(1-m)t\big)\mathbf{1}_{[m,1]}(z),$$

is another strong solution to system (16), since

$$\begin{aligned}
&\int_0^1 \operatorname{sign}(X(z)-X(\xi))\,\mathrm{d}\xi - \int_0^1 \operatorname{sign}(X(z)-Y(\xi))\,\mathrm{d}\xi \\
&= \int_0^1 \operatorname{sign}((1-m)t\mathbf{1}_{[m,1]}(z) - (1-m)t\mathbf{1}_{[m,1]}(\xi))\,\mathrm{d}\xi \\
&\quad - \int_0^1 \operatorname{sign}((1-m)t\mathbf{1}_{[m,1]}(z) - (1-(1-m)t)\mathbf{1}_{[m,1]}(\xi))\,\mathrm{d}\xi \\
&= (1-m)\mathbf{1}_{[m,1]}(z) \\
&= \frac{\partial}{\partial t}\tilde{X}(t,z),
\end{aligned}$$

and, repeating the same computation for $\tilde{Y}(t,z)$, we have that $\frac{\partial}{\partial t}\tilde{Y}(t,z) = -(1-m)\mathbf{1}_{[m,1]}(z)$.

5.2.3. *Characterisation of the sub-differential.* We notice that both solutions satisfy the assumptions of Proposition 9. Since for $(t,z) \in [0,T) \times [0,m)$

$$\frac{\partial}{\partial t}\begin{pmatrix} X(t,z) \\ Y(t,z) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

and, for $(t,z) \in [0,T) \times [m,1]$,

$$\frac{\partial}{\partial t}\begin{pmatrix} X(t,z) \\ Y(t,z) \end{pmatrix} = \begin{pmatrix} 2(z-m) \\ 2(z-1) \end{pmatrix},$$

we have $\frac{\partial}{\partial t}\begin{pmatrix} X(t,z) \\ Y(t,z) \end{pmatrix} \in -\partial\bar{\mathfrak{F}}\big[\big(X(t,z), Y(t,z)\big)\big]$, so that we can affirm (X,Y) is a gradient flow solution to system (16). In conclusion, $(\tilde{X},\tilde{Y})$ is not a gradient flow solution since $\frac{\partial}{\partial t}\tilde{X}(t,z) = 1-m \neq 2(z-m)$, as we claimed.

Remark 12. *Unlike the case of two separate Dirac deltas, the phenomenon arising in this example is indeed a distinctive feature of the two species case. This time the inter-specific energy is indeed affected at the singular point, since both species are present at the same position initially. The common mass m at the point zero is driven both by a self-repulsion and by a cross-attraction effect annihilating each other and producing no movement at all as a result. The extra mass of ρ is instead only driven by self-repulsion, and therefore it gets smoothed. At the point* 1*, only the smoothing effect occurs, as there is no singular cross-interaction. There is a significant aspect in this solution:* the gradient flow solution maintains a bit of its initial atomic part*, which never happens in the one species case.*

5.3. **Link with hyperbolic systems.** In this section we want to highlight the link between system (6) and a particular nonlinear 2×2 system of conservation laws in one space dimension, see [14, 32]. Indeed, considering the cumulative distribution functions F and G of ρ and η respectively (as defined in (9)), we can rewrite system (6) as

$$\begin{cases} \partial_t F + 2(F-G)\partial_x F = 0, \\ \partial_t G + 2(G-F)\partial_x G = 0, \end{cases} \tag{67}$$

or in the equivalent matrix form

$$\begin{pmatrix} \partial_t F \\ \partial_t G \end{pmatrix} + \begin{pmatrix} 2(F-G) & 0 \\ 0 & 2(G-F) \end{pmatrix} \cdot \begin{pmatrix} \partial_x F \\ \partial_x G \end{pmatrix} = 0.$$

We stress that the initial condition F_0, G_0 for F and G are non-decreasing and achieving values in $[0,1]$, with $F_0(-\infty) = G_0(-\infty) = 0$ and $F_0(+\infty) = G_0(+\infty) = 1$. System (67) is hyperbolic, though not strictly, as the eigenvalues are

$$\lambda_1(F,G) = 2(F-G)$$
$$\lambda_2(F,G) = 2(G-F),$$

and $\lambda_1 = \lambda_2$ on the diagonal $F = G$. Moreover, system (67) is *nonconservative*, in the sense that there exists no flux function $\boldsymbol{f} : [0,1] \times [0,1] \to \mathbb{R}^2$ such that (67) can be written as

$$\partial_t U + \partial_x(\boldsymbol{f}(U)) = 0, \qquad U = (F,G). \tag{68}$$

To our knowledge, no general theories on hyperbolic systems currently allow to define a notion of entropy solution for such a system due to the lack of strict hyperbolicity. In particular, there is no canonical way to define a suitable Riemann solver.
The link between (67) and (6) can be easily established at the level of weak solutions for (67) with sufficiently smooth initial data, which will correspond to weak solutions for (6). Such a link is slightly more tricky at the level of discontinuous solutions. On the other hand, the use of the Evolution Variational Inequality for (6) seems to be like a natural way to characterise a solution for (67) as well. This task will be performed in a future work. In this subsection we will just display the gradient flow solutions found in the previous subsections at the level of the hyperbolic system (67), as relevant examples of solutions of Cauchy problems which can be solved via the composition of two Riemann problems.
Let us start by considering the Cauchy problem

$$\begin{cases}\partial_t F + 2(F-G)\partial_x F = 0\\ \partial_t G + 2(G-F)\partial_x G = 0\end{cases}, \quad F_0 = \begin{cases}0 & x<-1\\ 1 & x\geq -1\end{cases}, \quad G_0 = \begin{cases}0 & x<1\\ 1 & x\geq 1\end{cases}, \tag{69}$$

which correspond to the initial condition for (6)

$$\rho_0 = \delta_{-1}, \quad \text{and} \quad \eta_0 = \delta_1.$$

As anticipated, we construct the solution (only for short time) by solving two separate Riemann problems, i.e.

$$\begin{cases}\partial_t F + 2(F-G)\partial_x F = 0\\ \partial_t G + 2(G-F)\partial_x G = 0\end{cases}, \quad F_0 = \begin{cases}0 & x<-1\\ 1 & x\geq -1\end{cases}, \quad G_0 = \begin{cases}0 & x<-1\\ 0 & x\geq -1\end{cases}, \tag{70}$$

and

$$\begin{cases}\partial_t F + 2(F-G)\partial_x F = 0\\ \partial_t G + 2(G-F)\partial_x G = 0\end{cases}, \quad F_0 = \begin{cases}1 & x<1\\ 1 & x\geq 1\end{cases}, \quad G_0 = \begin{cases}0 & x<1\\ 1 & x\geq 1\end{cases}. \tag{71}$$

On the basis of our previous results, we know the solution at the level of pseudo-inverses functions, i.e. for $z \in [0,1]$ and for t small enough,

$$X(t,z) = 2tz - 1, \qquad \text{and} \qquad Y(t,z) = 2t(z-1)+1.$$

Computing the corresponding cumulative distributions F and G, our candidate solution to problem (70) for t small enough is given by

$$F(t,x) = \begin{cases}0 & x \leq -1,\\ \dfrac{x+1}{2t} & -1 \leq x \leq 2t-1,\\ 1 & x \geq 2t-1,\end{cases}$$

$$G(t,x) = 0 \qquad \text{for} \quad x < 1,$$

whereas for problem (71) we have

$$F(t,x) = 1 \qquad \text{for} \quad x \geq 2t-1,$$
$$G(t,x) = \begin{cases} 0 & 0 \leq x < -2t+1, \\ 1 + \dfrac{x-1}{2t} & -2t+1 \leq x \leq 1 \\ 1 & x > 1. \end{cases}$$

The composition of these two solutions for short times is represented in Figure 3.

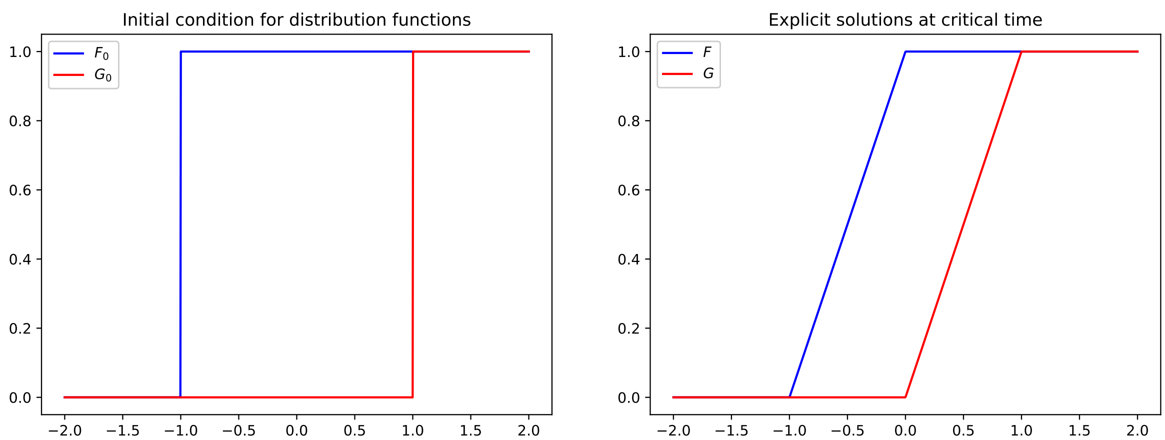


FIGURE 3. Initial (left) and exact solution (right) at time $t = 0.5$ for the case of two distinct Dirac deltas at the level of distribution functions.

Let us now consider as initial datum the cumulative distribution functions of

$$\rho_0 = \delta_0, \quad \text{and} \quad \eta_0 = m\delta_0 + (1-m)\delta_1,$$

as in (66). As before we have to deal with two different Riemann problems, i.e.

$$\begin{cases} \partial_t F + 2(F-G)\partial_x F = 0 \\ \partial_t G + 2(G-F)\partial_x G = 0, \end{cases} \qquad F_0 = \begin{cases} 0 & x < 0 \\ 1 & x \geq 0 \end{cases}, \quad G_0 = \begin{cases} 0 & x < 0 \\ m & x \geq 0 \end{cases}. \tag{72}$$

and

$$\begin{cases} \partial_t F + 2(F-G)\partial_x F = 0 \\ \partial_t G + 2(G-F)\partial_x G = 0, \end{cases} \qquad F_0 = \begin{cases} 1 & 0 < x < 1 \\ 1 & x \geq 1 \end{cases}, \quad G_0 = \begin{cases} m & x < 1 \\ 1 & x \geq 1 \end{cases}. \tag{73}$$

Going back to the results of the previous subsection, the gradient flow solution for the pseudo-inverse system (16) is given by

$$X(t,z) = \begin{cases} 0 & 0 \leq z < m \\ 2t(z-m) & m \leq z \leq 1, \end{cases} \qquad Y(t,z) = \begin{cases} 0 & 0 \leq z < m \\ 2t(z-1)+1 & m \leq z \leq 1, \end{cases}$$

and then, for t small enough, the candidate solution to problem (72) is given by

$$F(t,x) = \begin{cases} 0 & x < 0, \\ m + \dfrac{x}{2t} & 0 \le x \le 2(1-m)t, \\ 1 & x > 2(1-m)t, \end{cases}$$

$$G(t,x) = \begin{cases} 0 & x < 0, \\ m & x \ge 0, \end{cases}$$

whereas for problem (73) we have

$$F(t,x) = 1 \qquad \text{for} \quad x > 2(1-m)t,$$

$$G(t,x) = \begin{cases} m & 0 \le x < 2(m-1)t + 1, \\ 1 + \dfrac{x-1}{2t} & 2(m-1)t + 1 \le x \le 1 \\ 1 & x > 1, \end{cases}$$

see Figure 4.

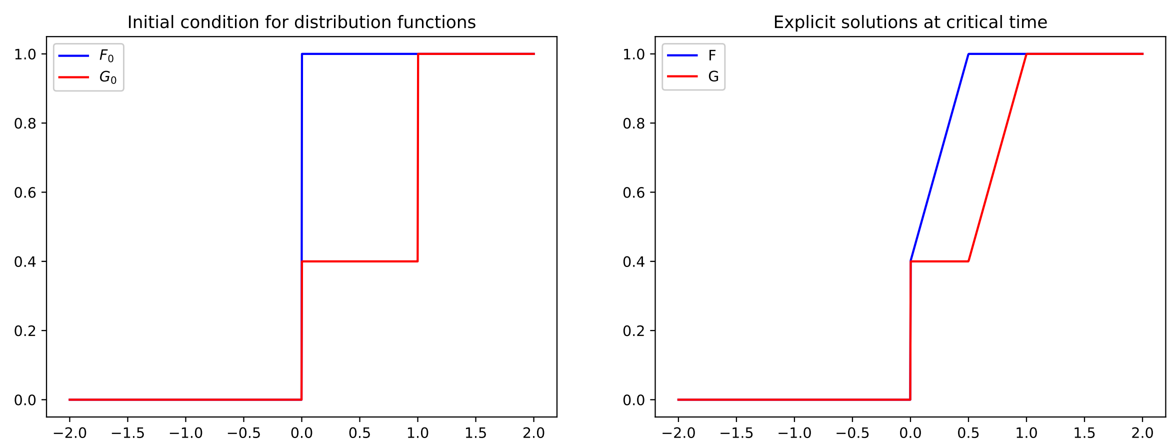


FIGURE 4. Initial (left) and exact solution (right) at time $t = 1/(4(1-m))$ with $m = 0.4$ for the case of two partially overlapping deltas at the level of distribution functions.

ACKNOWLEDGEMENTS

JAC was partially supported by the EPSRC grant number EP/P031587/1. The authors acknowledge discussions with D. Amadori and F. Ancona on the link with hyperbolic systems. MDF, AE, and SF acknowledge support by the EU-funded Erasmus Mundus programme "MathMods - Mathematical models in engineering: theory, methods, and applications" at the University of L'Aquila, by the Italian GNAMPA mini-project "Analisi di modelli matematici della fisica, della biologia e delle scienze sociali", and by the local fund of the University of L'Aquila 'DP-LAND (Deterministic Particles for Local And Nonlocal Dynamics).

References

[1] L. Ambrosio, N. Gigli, and G. Savaré. *Gradient flows in metric spaces and in the space of probability measures.* Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, second edition, 2008.

[2] L. Ambrosio and G. Savaré. Gradient flows of probability measures. *Handbook of differential equations: evolutionary equations*, 3:1–136, 2007.

[3] A. Arnold, P. Markowich, and G. Toscani. On large time asymptotics for drift-diffusion-Poisson systems. In *Proceedings of the Fifth International Workshop on Mathematical Aspects of Fluid and Plasma Dynamics (Maui, HI, 1998)*, volume 29, pages 571–581, 2000.

[4] D. Balagué, J. A. Carrillo, T. Laurent, and G. Raoul. Dimensionality of local minimizers of the interaction energy. *Arch. Ration. Mech. Anal.*, 209(3):1055–1088, 2013.

[5] D. Benedetto, E. Caglioti, and M. Pulvirenti. A kinetic equation for granular media. *RAIRO-Modélisation mathématique et analyse numérique*, 31(5):615–641, 1997.

[6] A. L. Bertozzi, J. A. Carrillo, and T. Laurent. Blow-up in multidimensional aggregation equations with mildly singular interaction kernels. *Nonlinearity*, 22(3):683–710, 2009.

[7] A. L. Bertozzi, T. Kolokolnikov, H. Sun, D. Uminsky, and J. von Brecht. Ring patterns and their bifurcations in a nonlocal model of biological swarms. *Commun. Math. Sci.*, 13(4):955–985, 2015.

[8] A. L. Bertozzi, T. Laurent, and F. Léger. Aggregation and spreading via the newtonian potential: the dynamics of patch solutions. *Mathematical Models and Methods in Applied Sciences*, 22(supp01):1140005, 2012.

[9] A. L. Bertozzi, T. Laurent, and J. Rosado. Lp theory for the multidimensional aggregation equation. *Communications on Pure and Applied Mathematics*, 64(1):45–83, 2011.

[10] F. Bolley, Y. Brenier, and G. Loeper. Contractive metrics for scalar conservation laws. *J. Hyperbolic Differ. Equ.*, 2(1):91–107, 2005.

[11] G. A. Bonaschi. Gradient flows driven by a non-smooth repulsive interaction potential. *Master's thesis, University of Pavia, Italy, arXiv:1310.3677*, 2011.

[12] G. A. Bonaschi, J. A. Carrillo, M. Di Francesco, and M. A. Peletier. Equivalence of gradient flows and entropy solutions for singular nonlocal interaction equations in 1d. *ESAIM Control, Optimisation and Calculus of Variations*, 21(2):414–441, 2015.

[13] Y. Brenier. L^2 formulation of multidimensional scalar conservation laws. *Arch. Ration. Mech. Anal.*, 193(1):1–19, 2009.

[14] A. Bressan. *Hyperbolic systems of conservation laws. The one-dimensional Cauchy problem.* Oxford Lecture Series in Mathematics and its Applications. Oxford University Press, Oxford, 20 edition, 2000.

[15] H. Brézis. *Opérateurs maximaux monotones et semi-groupes de contractions dans les espaces de Hilbert.* North-Holland Publishing Co., Amsterdam, 1973.

[16] M. Burger, M. Di Francesco, and M. Franek. Stationary states of quadratic diffusion equations with long-range attraction. *Communications in Mathematical Sciences*, 11(3):709–738, 2013.

[17] M. Burger, R. Fetecau, and Y. Huang. Stationary states and asymptotic behavior of aggregation models with nonlinear local repulsion. *SIAM J. Appl. Dyn. Syst.*, 13(1):397–424, 2014.

[18] M. Burger, M. Francesco, S. Fagioli, and A. Stevens. Sorting phenomena in a mathematical model for two mutually attracting/repelling species. *SIAM Journal on Mathematical Analysis*, 50(3):3210–3250, 2018.

[19] V. Calvez, J. A. Carrillo, and F. Hoffmann. Equilibria of homogeneous functionals in the fair-competition regime. *Nonlinear Anal.*, 159:85–128, 2017.

[20] V. Calvez, J. A. Carrillo, and F. Hoffmann. The geometry of diffusing and self-attracting particles in a one-dimensional fair-competition regime. 2186:1–71, 2017.

[21] J. Carrillo, Y. Huang, and M. Schmidtchen. Zoology of a nonlocal cross-diffusion model for two species. *SIAM Journal on Applied Mathematics*, 78(2):1078–1104, 2018.

[22] J. A. Carrillo, A. Chertock, and Y. Huang. A finite-volume method for nonlinear nonlocal equations with a gradient flow structure. *Communications in Computational Physics*, 17(01):233–258, 2015.

[23] J. A. Carrillo, K. Craig, and Y. Yao. Aggregation-diffusion equations: dynamics, asymptotics, and singular limits. *arXiv preprint arXiv:1810.03634*, 2018.

[24] J. A. Carrillo, M. G. Delgadino, and A. Mellet. Regularity of local minimizers of the interaction energy via obstacle problems. *Comm. Math. Phys.*, 343(3):747–781, 2016.

[25] J. A. Carrillo, M. Di Francesco, A. Figalli, T. Laurent, and D. Slepcev. Global-in-time weak measure solutions and finite-time aggregation for nonlocal interaction equations. *Duke Math. J.*, 156(2):229–271, 2011.

[26] J. A. Carrillo, L. C. Ferreira, and J. C. Precioso. A mass-transportation approach to a one dimensional fluid mechanics model with nonlocal velocity. *Advances in Mathematics*, 231(1):306–327, 2012.

[27] J. A. Carrillo, F. Filbet, and M. Schmidtchen. Convergence of a Finite Volume Scheme for a System of Interacting Species with Cross-Diffusion. *ArXiv e-prints*, Apr. 2018.

[28] J. A. Carrillo, Y. Huang, and S. Martin. Explicit flock solutions for Quasi-Morse potentials. *European J. Appl. Math.*, 25(5):553–578, 2014.

[29] J. A. Carrillo, S. Martin, and V. Panferov. A new interaction potential for swarming models. *Phys. D*, 260:112–126, 2013.

[30] J. A. Carrillo and G. Toscani. Wasserstein metric and large–time asymptotics of nonlinear diffusion equations. *New Trends in Mathematical Physics, (In Honour of the Salvatore Rionero 70th Birthday)*, pages 234–244, 2005.

[31] T. Cazenave and A. Haraux. *An introduction to semilinear evolution equations*. Oxford Lecture Series in Mathematics and its Applications, 13. Oxford University Press, New York, 1998.

[32] C. M. Dafermos. *Hyperbolic conservation laws in continuum physics. Fourth edition*. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]. Springer-Verlag, Berlin, 325 edition, 2016.

[33] S. Daneri and G. Savaré. Eulerian calculus for the displacement convexity in the wasserstein distance. *SIAM J. Math. Anal.*, 40(3):1104—1122, 2008.

[34] E. De Giorgi. New problems on minimizing movements. *Boundary Value Problems for PDE and Applications, C. Baiocchi and J. L. Lions eds., Masson*, pages 81—-98, 1993.

[35] M. Di Francesco, A. Esposito, and S. Fagioli. Nonlinear degenerate cross-diffusion systems with nonlocal interaction. *Nonlinear Analysis*, 169:94–117, 2018.

[36] M. Di Francesco and S. Fagioli. Measure solutions for nonlocal interaction pdes with two species. *Nonlinearity*, 26:2777–2808, 2013.

[37] M. Di Francesco and D. Matthes. Curves of steepest descent are entropy solutions for a class of degenerate convection-diffusion equations. *Calc. Var. Partial Differential Equations*, 50(1–2):199–230, 2014.

[38] M. R. D'Orsogna, Y.-L. Chuang, A. L. Bertozzi, and L. S. Chayes. Self-propelled particles with soft-core interactions: patterns, stability, and collapse. *Physical review letters*, 96(10):104302, 2006.

[39] L. C. Evans. *Partial Differential Equations*, volume 19 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, Rhode Island, 1998.

[40] J. H. M. Evers and T. Kolokolnikov. Metastable states for an aggregation model with noise. *SIAM J. Appl. Dyn. Syst.*, 15(4):2213–2226, 2016.

[41] R. C. Fetecau and Y. Huang. Equilibria of biological aggregations with nonlocal repulsive-attractive interactions. *Phys. D*, 260:49–64, 2013.

[42] R. C. Fetecau, Y. Huang, and T. Kolokolnikov. Swarm dynamics and equilibria for a nonlocal aggregation model. *Nonlinearity*, 24(10):2681–2716, 2011.

[43] R. Jordan, D. Kinderlehrer, and F. Otto. The variational formulation of the fokker–planck equation. *SIAM journal on mathematical analysis*, 29(1):1–17, 1998.

[44] T. Kolokolnikov, H. Sun, D. Uminsky, and A. L. Bertozzi. Stability of ring patterns arising from two-dimensional particle interactions. *Phys. Rev. E*, 84:015203, Jul 2011.

[45] S. Kružhkov. First order quasilinear equations with several independent variables. *Mat. Sb. (N.S.)*, 81 (123):228–255, 1970.

[46] O. A. Ladyženskaja, V. A. Solonnikov, and N. N. Ural'ceva. *Linear and quasilinear equations of parabolic type*. Translated from the Russian by S. Smith. Translations of Mathematical Monographs, Vol. 23. American Mathematical Society, Providence, R.I., 1967.

[47] H. Li and G. Toscani. Long-time asymptotics of kinetic models of granular flows. *Arch. Ration. Mech. Anal.*, 172(3):407—-428, 2004.

[48] T.-P. Liu and M. Pierre. Source solutions and asymptotic behaviour in conservations laws. *J. Differential Equations*, 51:419–441, 1984.

[49] D. Matthes, R. McCann, and G. Savaré. A family of fourth order equations of gradient flow type. *Comm. P.D.E.*, 34(11):1352–1397, 2009.

[50] R. J. McCann. A convexity principle for interacting gases. *Advances in mathematics*, 128(1):153–179, 1997.

[51] A. Mogilner and L. Edelstein-Keshet. A non-local model for a swarm. *Journal of Mathematical Biology*, 38(6):534–570, 1999.

[52] F. Otto. The geometry of dissipative evolution equations: the porous medium equation. *Comm. Partial Differential Equations*, 26(1–2):101–174, 2001.
[53] S. T. Rachev and L. Rüschendorf. *Mass Transportation Problems*, volume I of *Probability and Its Applications*. Springer, New York, 1998.
[54] F. Santambrogio. *Optimal Transport for Applied Mathematicians*, volume 86 of *Progress in Nonlinear Differential Equations and Their Applications*. Birkhäuser Verlag, Basel, 2015.
[55] C. M. Topaz, A. L. Bertozzi, and M. A. Lewis. A nonlocal continuum model for biological aggregation. *Bull. Math. Biol.*, 68(7):1601–1623, 2006.
[56] G. Toscani. One-dimensional kinetic models of granular flows. *ESAIM: Modélisation Mathématique et Analyse Numérique*, 34(6):1277–1291, 2000.
[57] C. Villani. *Topics in optimal transportation*, volume 58 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2003.
[58] C. Villani. *Optimal transport : old and new*. Grundlehren der mathematischen Wissenschaften. Springer, Berlin, 2009.

Email address: carrillo@imperial.ac.uk
Email address: marco.difrancesco@univaq.it
Email address: antonio.esposito3@graduate.univaq.it
Email address: simone.fagioli@univaq.it
Email address: m.schmidtchen15@imperial.ac.uk