ON NONLINEAR CROSS-DIFFUSION SYSTEMS: AN OPTIMAL TRANSPORT APPROACH

INWON KIM AND ALPÁR RICHÁRD MÉSZÁROS

ABSTRACT. We study a nonlinear, degenerate cross-diffusion model which involves two densities with two different drift velocities. A general framework is introduced based on its gradient flow structure in Wasserstein space to derive a notion of discrete-time solutions. Its continuum limit, due to the possible mixing of the densities, only solves a weaker version of the original system. In one space dimension, we find a stable initial configuration which allows the densities to be segregated. This leads to the evolution of a stable interface between the two densities, and to a stronger convergence result to the continuum limit. In particular derivation of a standard weak solution to the system is available. We also study the incompressible limit of the system, which addresses transport under a height constraint on the total density. In one space dimension we show that the problem leads to a two-phase Hele-Shaw type flow.

1. INTRODUCTION

Let Ω be a bounded domain in $\mathbb{R}^d$ with C^1 boundary, and let $T > 0$ and $m > 1$ be given constants. In this paper we study a gradient flow formulation of the following system in $[0, T] \times \Omega$:

$$(\mathrm{PME}_m) \qquad \begin{cases} \partial_t \rho^1 - \nabla \cdot \left((\nabla p + \nabla \Phi_1) \rho^1 \right) = 0; \\ \partial_t \rho^2 - \nabla \cdot \left((\nabla p + \nabla \Phi_2) \rho^2 \right) = 0, \end{cases}$$

where $\Phi_1, \Phi_2 : \Omega \to \mathbb{R}$ are given and the common diffusion term is generated by the pressure variable

$$(1.1) \qquad p := \frac{m}{m-1} (\rho^1 + \rho^2)^{m-1}.$$

In this article the system is subject to no flux condition on $[0, T] \times \partial\Omega$ and is equipped with initial nonnegative densities $\rho_0^1, \rho_0^2 \in L^1(\Omega)$.

Formally (PME_m) can be seen as the gradient flow in Wasserstein (product) space of the free energy

$$(1.2) \qquad (\rho^1, \rho^2) \mapsto \int_\Omega \frac{1}{m-1} (\rho^1 + \rho^2)^m \, \mathrm{d}x + \int_\Omega \Phi_1 \rho^1 \, \mathrm{d}x + \int_\Omega \Phi_2 \rho^2 \, \mathrm{d}x.$$

Staying at the formal level, in the incompressible limit as $m \to +\infty$ where the first term in free energy turns into the constraint $\rho^1 + \rho^2 \le 1$, the corresponding system for the limiting density pair $(\rho^{1,\infty}, \rho^{2,\infty})$ is

$$(\mathrm{PME}_\infty) \qquad \begin{cases} \partial_t \rho^{1,\infty} - \nabla \cdot \left((\nabla p^\infty + \nabla \Phi_1) \rho^{1,\infty} \right) = 0; \\ \partial_t \rho^{2,\infty} - \nabla \cdot \left((\nabla p^\infty + \nabla \Phi_2) \rho^{2,\infty} \right) = 0, \end{cases}$$

where the pressure p^∞ is supported in the region $\{\rho^{1,\infty} + \rho^{2,\infty} = 1\}$. When the densities $\rho^{i,\infty}$'s are characteristic functions with separate supports, the problem corresponds to a two-phase Hele-Shaw type flow with drifts.

Our goal in this paper is to study the problems (PME_m) and (PME_∞) in the context of the aforementioned gradient flow, and verify the above heuristics. More precisely we will formulate the problem in terms of the discrete-time gradient flow (i.e. JKO or minimizing movement scheme) of the aforementioned free energy (1.2), posed in the product space equipped with the 2-Wasserstein metric. Then we will study the solutions of this discrete scheme as the time step goes to zero. We will show that the limiting pair of densities (ρ^1, ρ^2) satisfies a set of transport equation that will reduce to (PME_m) under a stronger convergence assumption (for the precise statement we refer to Theorem 3.8). To strengthen this result, it

seems necessary to consider "stable" initial configurations which avoids mixing: see the discussion below. It turns out that in one dimension, in the setting of stable initial configurations which avoids mixing, a stronger convergence result holds and as a consequence the continuum limit densities satisfy (PME_m) in the standard weak sense. Below is a summary of our main results: precise statements are contained in the quoted theorems.

The main results of the paper are obtained in one space dimension. Here we assume that the density with stronger drift in x direction sits on the right side on the x-axis, i.e.,

$$(1.3) \qquad -\partial_x \Phi_1 \geq -\partial_x \Phi_2 \quad \text{and} \quad x_1 \geq x_2 \text{ for } x_i \in \{\rho_0^i > 0\}, i = 1, 2.$$

Under the assumption (1.3) the following theorems hold:

Theorem 1.1 (Segregation of solutions: Proposition 4.1, Theorem 4.2). *For given $m \in (1, \infty]$ and $\tau > 0$, let $(\rho^{1,\tau}, \rho^{2,\tau})$ be the time-discrete solutions given by the minimizing movement scheme with time step size $\tau > 0$, as given in (MM_m). Then, the pair stays ordered for all times $t > 0$, i.e. $\rho_t^{1,\tau}$ is supported to the right of the support of $\rho_t^{2,\tau}$ for all $t > 0$. Moreover, as $\tau \downarrow 0$ each density $\rho^{i,\tau}$ converges weakly in $L^{2m-1}([0,T] \times \Omega)$, along a subsequence, to $\rho^{i,m}$ for $i = 1, 2$. Also, along a subsequence $\rho^{i,\tau}$ converges pointwise a.e. to $\rho^{i,m}$ for $i = 1, 2$. The pair of limiting densities $(\rho^{1,m}, \rho^{2,m})$ solves (PME_m) in the weak sense and it stays ordered for all times.*

Theorem 1.2 (Convergence of weak solutions as $m \to \infty$: Theorem 4.7). *Let $(\rho^{1,m}, \rho^{2,m})$ be as given above for given $m > 1$. Then as $m \to \infty$ and along a subsequence, the density pairs converge weakly in $L^p([0,T] \times \Omega)$ for any $1 < p < \infty$ to $(\rho^{1,\infty}, \rho^{2,\infty})$, which is a weak solution of (PME_∞).*

Theorem 1.3 (Patch solutions for (PME_∞): Proposition 4.9). *Let $m = \infty$ and suppose, in addition to (1.3), that $\partial_{xx}^2 \Phi_i \geq 0$ and $\rho_0^i = \chi_{(a^i(0), b^i(0))}$. Then $\rho^{i,\infty}$ remains a patch for all $t > 0$, i.e. $\rho_t^{i,\infty} = \chi_{(a^i(t), b^i(t))}$, $i = 1, 2$. The density pair in this case is a solution to a two-phase Hele-Shaw type flow with drifts.*

Remark 1.1. *For the linear diffusion $m = 1$ the logarithmic entropy $\int \rho \ln \rho \, \mathrm{d}x$ replaces $\frac{1}{m-1} \int \rho^m \, \mathrm{d}x$ in the free energy, resulting in slightly different, however mostly parallel, analysis. Let us point out that in this case the sum of the densities is always positive, and thus they will always form an interface between each other in the event of segregation.*

Let us discuss now the existing results from the literature that are relevant to our work. The single density version of the system (PME_∞) has been introduced in [30] in the gradient flow setting, the free boundary characterization and its links to (PME_m) has been studied in [1]. These and similar systems received a lot of attention in the past a few years (see for instance [31, 32, 33] and the references therein). These models are strongly related to the so-called Hele-Shaw models, as we can see in [1] (for other references we direct the reader to [23, 24, 35, 39] and to the references therein).

Cross diffusion systems arise naturally from mathematical biology. These appear either as systems of reaction-diffusion equations (as in [28, 22] for instance) or systems of advection-diffusion equations (as in [7, 6, 16, 45] for instance). These systems appear also in fluid mechanics, such as the thin film approximation of the Muskat problem studied in [20, 27]. Beside the PDE approach for degenerate parabolic systems (used in most of the above references), more recently the optimal transport and gradient flow theories have been adopted to study these systems. For a non-exhaustive list of the fast-growing references in this direction we refer to [5, 8, 9, 10, 12, 13, 14, 15, 16, 18, 25, 26, 27].

Most of the aforementioned papers concern systems including separate diffusion terms for the evolution of each densities. Such feature enables tracking of separate densities in the evolution. This is in contrast to our case where recovering separate densities out of the dynamic system appears to be out of reach unless the densities are guaranteed to be segregated. There are very few results available for systems without separate diffusions, we mention here a few particular papers in this direction. In [6] the authors study the well-posedness of the system (PME_m) in one space dimension when $m = 2$, $\Phi_1 = \Phi_2 \equiv 0$ within the class of segregated solutions. This is a particular case of our model where the total density satisfies a degenerate parabolic equation, from which one can construct a segregated solution. In [38] Otto studies in one dimension the case $m = \infty$ with gravity potentials $\Phi_i(x) = C_i x$ and with full saturation, i.e. with

the condition $\rho_1 + \rho_2 = 1$. There the mixing profile of one density correponds to an entropy solution of a Burger's type equation. This interesting description of mixing phenomena remains open to be extended beyond the specific setting given in the paper. Lastly in the recent paper [9] the authors study existence and segregation properties of one dimensional stationary solutions for systems of similar form to ours, when $m = 2$ and the drift is generated by interaction energies.

Main difficulties and ingredients

As mentioned above, our main challenge lies in the fact that the densities may mix into each other during the evolution, which indeed happens with the "unstable" initial configurations where the densities are initially positioned in the opposite order to the equilibrium solution (see the discussion in Section 2.5). Such situation indicates low regularity of each densities, hindering the system from being well-posed. Indeed, in general we are only able to obtain strong convergence on the sum of the two density variables in the continuum limit, as we will see in Theorem 3.8. Naturally this reasoning leads to the question of whether one can formulate a "stable" initial configuration to obtain a stronger result. This question, while under investigation by the authors, stands open beyond the one dimensional result for segregated solutions, Theorems 1.1, 1.2, 1.3.

In terms of gradient flows, the challenge lies in the lack of available estimates or convexity properties. Though not surprising in the context of the above discussion, this is an interesting contrast to the single species case, where for instance stability of the discrete gradient flow solutions based on λ-convexity properties played an important role in the analysis. For us the higher order space regularity estimates in the JKO scheme are available only for functions of the sum of the two densities. For similar models considered in [25] and [27] this difficulty was overruled by presence of separate diffusions, or "separate entropies" of the form $\varepsilon(\int f(\rho^1) + g(\rho^2) \, dx)$ in the free energy (1.2), however estimates obtained here do not carry through as $\varepsilon \downarrow 0$. Let us also mention that the flow interchange technique introduced in [29], which has been quite successful to analyze some non-convex gradient flow systems such as in [19] or [27], does not appear to be applicable to our system. Thus here we derive all our estimates relying only on the first order optimality conditions satisfied by the discrete in time minimizers in the JKO scheme (see for instance the proof of Theorem 2.4). This procedure is rather natural yet appears to be unexploited in the literature for similar models.

Structure of the paper

In Section 2 the discrete-time scheme for the gradient flow is introduced, set in the W_2-product space. In Section 2.4 the properties of discrete-time minimizers are studied. Here we observe that while the total density $\rho^1 + \rho^2$ is relatively regular (Lipschitz continuous), separate densities may be segregated and discontinuous. The segregation of densities with respect to the ordering properties of their potentials are more obvious in Section 2.5, where one discusses the equilibrium solutions. Such segregation and ordering property suggests that fingering and mixing is inevitable for densities starting from "unstable" initial configurations, to position themselves into the stationary profile.

In Section 3 we analyze the continuum limit of discrete-time solutions by studying their convergence modes as the time step size is sent to zero. We show that the limit solution satisfies a system of transport equations which can be interpreted as a generalized solutions for the system (PME_m). We also introduce the standard notion of weak solution for our systems and show that the continuum limit satisfies this notion when pointwise convergence holds for separate densities. It remains an open question whether the densities indeed converge pointwise, i.e. whether we can track down the position of each density in the evolution of the problem in general framework or in general dimension.

Section 4 is devoted to the analysis in one space dimension, where we consider stable initial configurations that line up with the strength order of the drift potentials. In this setting we are able to guarantee that solutions stay segregated with an evolving interface between them. As a consequence it follows that pointwise convergence holds for each densities, which in turn yields the existence of weak solutions for the system (PME_m). The continuum solutions of (PME_m) are then shown to converge as m tends to infinity to a weak solution of (PME_∞) along a subsequence. Furthermore when the drift is compressive (or incompressible), we show that patch solutions appear, yielding a solution to the two-phase Hele-Shaw flow.

Finally, in the Appendices A and B we recall some results from the theory of optimal transport and a refined version of Aubin-Lions lemma respectively.

2. Minimizing movement schemes and properties of the minimizers

2.1. **Setting and notations.** Let us introduce the setting of the problem and some notations. Let $\Omega \subseteq \mathbb{R}^d$ be a bounded domain with smooth boundary. We denote by $\mathscr{P}(\Omega)$ the space of probability measures on Ω. For $M > 0$ we denote by $\mathscr{P}^M(\Omega)$ the space of finite nonnegative Radon measures on Ω ($\mathscr{M}_+(\Omega)$) with mass M.

For a Borel measurable map $T : \Omega \to \Omega$ and $\mu, \nu \in \mathscr{P}^M(\Omega)$, we say that T *pushes forward μ onto ν*, and write $\nu = T_\# \mu$ if $\nu(B) = \mu(T^{-1}(B))$ for every $B \subseteq \Omega$ Borel measurable set. Using test functions, the definition of pushforward translates to

$$\int_\Omega \phi(y)\,\mathrm{d}\nu(y) = \int_\Omega \phi(T(x))\,\mathrm{d}\mu(x), \ \ \forall \phi : \Omega \to \mathbb{R}, \ \ \text{bounded and measurable.}$$

We equip the space $\mathscr{P}^M(\Omega)$ with the well-known 2-Wasserstein distance $W_{2,M}$, i.e. For $\mu, \nu \in \mathscr{P}^M(\Omega)$,

$$W_{2,M}^2(\mu,\nu) := \min\left\{ \int_{\Omega \times \Omega} |x - y|^2\,\mathrm{d}\gamma \ : \ \gamma \in \Pi^M(\mu,\nu) \right\},$$

where $\Pi^M(\mu,\nu)$ is the set of the so-called *transport plans*, i.e. $\Pi^M(\mu,\nu) := \{\gamma \in \mathscr{P}^{M^2}(\Omega \times \Omega) : (\pi^x)_\# \gamma = \mu, \ (\pi^y)_\# \gamma = \nu\}$. In particular if $\mu \ll \mathscr{L}^d \llcorner \Omega$ then the previous problem has a unique solution, which is of the form $\gamma_T := (\mathrm{id}, T)_\# \mu$. Here, in particular we adjusted the usual distance defined on probability measures to measures having mass $M > 0$. Since it shall be clear from the context, from now on we write W_2 instead of $W_{2,M}$. On the forthcoming pages we shall use classical results from the optimal transport theory. All of these can be found for instance in [43, 3, 44].

We denote by $\mathscr{M}^d(\Omega)$ the space of finite vector-valued Radon measures on Ω. If $\mathbf{E} \in \mathscr{M}^d(\Omega)$, we denote by $|\mathbf{E}|$ its variation. We denote the subspaces of absolutely continuous measures (w.r.t. $\mathscr{L}^d \llcorner \Omega$) by $\mathscr{P}^{\mathrm{ac}}(\Omega)$, $\mathscr{P}^{\mathrm{ac},M}(\Omega)$, etc.; we always identify these absolutely continuous measures with their densities and write ρ instead of $\rho \cdot \mathscr{L}^d$ or $\rho \mathrm{d}x$. If $\rho \in \mathscr{M}_+^{\mathrm{ac}}(\Omega)$ and $c \geq 0$ by $\{\rho > c\}$ we mean the set (up to $\mathscr{L}^d$-negligible sets) where $\rho(x) > c$ a.e. In particular a property holds a.e. in $\{\rho > 0\}$ if and only if it holds $\rho-$a.e. Notice also that $\{\rho > 0\} \subseteq_{\mathrm{a.e.}} \mathrm{spt}(\rho)$. For a measurable set $B \subset \mathbb{R}^d$, we denote the set of its Lebesgue point by $\mathrm{Leb}(B)$.

2.2. **Minimizing movements.** The heart of our analysis is the well-known *minimizing movement* or *JKO* scheme (see for instance [2, 3, 42, 21]) on a product Wasserstein space.

Let us introduce the functionals. We consider $\mathcal{F}_m, \mathcal{F}_\infty : \mathscr{P}^{M_1}(\Omega) \times \mathscr{P}^{M_2}(\Omega) \to \mathbb{R} \cup \{+\infty\}$ and $\mathcal{G} : \mathscr{P}^{M_1}(\Omega) \times \mathscr{P}^{M_2}(\Omega) \to \mathbb{R}$ to be defined as

$$(2.1) \qquad \mathcal{F}_m(\boldsymbol{\rho}) = \begin{cases} \displaystyle \int_\Omega \frac{1}{m-1}(\rho^1(x) + \rho^2(x))^m \mathrm{d}x, & \text{if } (\rho^1 + \rho^2)^m \in L^1(\Omega), \\ +\infty, & \text{otherwise,} \end{cases}$$

$$(2.2) \qquad \mathcal{F}_\infty(\boldsymbol{\rho}) = \begin{cases} 0, & \text{if } \|\rho^1 + \rho^2\|_{L^\infty} \leq 1, \\ +\infty, & \text{otherwise,} \end{cases}$$

and $\mathcal{G} : \mathscr{P}^{M_1}(\Omega) \times \mathscr{P}^{M_2}(\Omega) \to \mathbb{R}$

$$(2.3) \qquad \mathcal{G}(\boldsymbol{\rho}) = \int_\Omega \Phi_1(x)\,\mathrm{d}\rho^1(x) + \int_\Omega \Phi_2(x)\,\mathrm{d}\rho^2(x),$$

where $\boldsymbol{\rho} := (\rho^1, \rho^2)$, $m > 1$ is fixed and $\Phi_1, \Phi_2 : \Omega \to \mathbb{R}$ are given continuous potentials. Notice that $\mathcal{F}_\infty$ is the indicator function (in the sense of convex analysis) of the set

$$\mathcal{K}_1 := \left\{ (\rho^1, \rho^2) \in \mathscr{P}^{M_1, \mathrm{ac}}(\Omega) \times \mathscr{P}^{M_2, \mathrm{ac}}(\Omega) : \rho^1 + \rho^2 \le 1 \text{ a.e.} \right\}.$$

It is classical that $\mathcal{F}_m, \mathcal{F}_\infty$ and $\mathcal{G}$ are l.s.c. w.r.t. the weak convergence of measures on $\mathscr{P}^{M_1}(\Omega) \times \mathscr{P}^{M_2}(\Omega)$. It is immediate to see that they are convex, moreover $\mathcal{F}_m$ is also strictly convex (in the usual sense) on $\mathscr{P}^{M_1}(\Omega) \times \mathscr{P}^{M_2}(\Omega)$. We remark also that in general $\mathcal{F}_m$ is not *displacement convex* (in the sense of [34]) on the product space. To see this, let us consider for simplicity $m = 2$. In this case, for $\rho^1, \rho^2 \in L^2(\Omega)$ we can write $\mathcal{F}_2(\rho^1, \rho^2) = \int_\Omega (\rho^1)^2 \, \mathrm{d}x + \int_\Omega (\rho^2)^2 \, \mathrm{d}x + 2 \int_\Omega \rho^1 \rho^2 \, \mathrm{d}x$. If $\mathcal{F}_2$ would be λ-displacement convex (for some $\lambda \in \mathbb{R}$), then the map $\rho^1 \mapsto \mathcal{F}_2(\rho^1, \rho^2)$ would share at least the same modulus of convexity for any $\rho^2 \in \mathscr{P}^{M_2}(\Omega) \cap L^2(\Omega)$ fixed. While the first term in the development of $\mathcal{F}_2$ is 0-displacement convex and the second term is a constant for fixed ρ^2, the last term would be λ-displacement convex if and only if ρ^2 would be λ-convex, i.e. $D^2 \rho^2 \ge \lambda I_d$ in the sense of distributions. However, ρ^2 can be chosen in a way that the lower bound on its Hessian is arbitrarily negative. Therefore, this term fails to be λ-displacement convex for any $\lambda \in \mathbb{R}$ and so does the functional $\mathcal{F}_2$.

We proceed as in the classical setting (see for instance [3, 21]): we define a recursive sequence of densities associated to a fixed time step τ, then we introduce suitable interpolations between these densities and take the limit as $\tau \downarrow 0$.

Let us introduce now the scheme. For this, we consider $\tau > 0$ a fixed time step and $N \in \mathbb{N}$ such that $N\tau = T$. Let (ρ_0^1, ρ_0^2) be two given initial densities. For all $k \in \{0, \dots, N\}$ we define $\boldsymbol{\rho}_k^\tau := (\rho_k^{1,\tau}, \rho_k^{2,\tau})$ as

$$\boldsymbol{\rho}_0^\tau = (\rho_0^{1,\tau}, \rho_0^{2,\tau}) := (\rho_0^1, \rho_0^2)$$

and for $k \ge 0$

$$(\mathrm{MM}_m) \qquad \boldsymbol{\rho}_{k+1}^\tau = (\rho_{k+1}^{1,\tau}, \rho_{k+1}^{2,\tau}) = \operatorname{argmin}_{\boldsymbol{\rho} \in \mathscr{P}^{M_1}(\Omega) \times \mathscr{P}^{M_2}(\Omega)} \left\{ \mathcal{F}_m(\boldsymbol{\rho}) + \mathcal{G}(\boldsymbol{\rho}) + \frac{1}{2\tau} \mathbf{W}_2^2(\boldsymbol{\rho}, \boldsymbol{\rho}_k^\tau) \right\}.$$

In this scheme either $m > 1$ but finite, or $m = \infty$. Here $\mathbf{W}_2$ denotes the Wasserstein distance on the product space $\mathscr{P}^{M_1}(\Omega) \times \mathscr{P}^{M_2}(\Omega)$, i.e. $\mathbf{W}_2^2(\boldsymbol{\mu}, \boldsymbol{\nu}) := W_2^2(\mu^1, \nu^1) + W_2^2(\mu^2, \nu^2)$, where $\boldsymbol{\mu} := (\mu^1, \mu^2)$, and $\boldsymbol{\nu} := (\nu^1, \nu^2)$.

We state the following well-known lemma.

Lemma 2.1. *The objective functional in the minimization problem* (MM_m) *is l.s.c. and bounded from below and* $\mathscr{P}^{M_1}(\Omega) \times \mathscr{P}^{M_2}(\Omega)$ *is compact, thus the optimizer exists. Moreover,* $\mathcal{F}_m$ *($m \in [1, +\infty]$) and* $\mathcal{G}$ *are convex functionals and the functional* $\boldsymbol{\rho} \mapsto \mathbf{W}_2^2(\boldsymbol{\rho}, \boldsymbol{\mu})$ *is strictly convex whenever* $\boldsymbol{\mu} = (\mu^1, \mu^2)$ *has absolutely continuous density coordinates (see for instance [43]). Therefore, if the densities* (ρ_0^1, ρ_0^2) *are absolutely continuous w.r.t.* $\mathscr{L}^d \llcorner \Omega$, *the optimizer* $\boldsymbol{\rho}_k^\tau$ *is also unique at each step.*

2.3. **Different diffusion coefficients for the two densities.** In many cross-diffusion models (coming mainly from mathematical biology or fluid mechanics, see for instance in [6, 28]) considered in the literature, it is important to have different diffusion coefficients for the two densities. In our setting, this could be formulated as follows. Given κ_1, κ_2 positive constants, consider a system similar to (PME_m) or (PME_∞), i.e.

$$(2.4) \qquad \begin{cases} \partial_t \rho^1 - \nabla \cdot \left(\kappa_1 \nabla p \rho^1 + \nabla \Phi_1 \rho^1 \right) = 0 \\ \partial_t \rho^2 - \nabla \cdot \left(\kappa_2 \nabla p \rho^2 + \nabla \Phi_2 \rho^2 \right) = 0 \end{cases}$$

on $[0, T] \times \Omega$, where $p := \dfrac{m}{m-1} (\rho^1 + \rho^2)^{m-1}$, with $m > 1$ and where $\Phi_1, \Phi_2 : \Omega \to \mathbb{R}$ are given potentials. Observe that (PME_m) corresponds to $\kappa_1 = \kappa_2 = 1$. Actually, even for $\kappa_1 \ne \kappa_2$, this system enters naturally into the framework of gradient flows considered in this paper. Indeed, we can define the minimizing movement scheme as

$$(\rho_{k+1}^{1,\tau}, \rho_{k+1}^{2,\tau}) = \operatorname{argmin}_{(\rho^1, \rho^2)} \left\{ \mathcal{F}(\rho^1, \rho^2) + \int_\Omega \frac{\Phi_1}{\kappa_1} \rho^1 \, \mathrm{d}x + \int_\Omega \frac{\Phi_2}{\kappa_2} \rho^2 \, \mathrm{d}x + \frac{1}{2\tau \kappa_1} W_2^2(\rho^1, \rho_k^{1,\tau}) + \frac{1}{2\tau \kappa_2} W_2^2(\rho^2, \rho_k^{2,\tau}) \right\}.$$

Actually a part of the analysis that we perform in the forthcoming sections will be valid in this case as well. In particular the results from Section 3 can be easily adapted to the system (2.4).

2.4. **Properties of the minimizers.** We discuss now some properties of the minimizers in (MM_m). For this, let us consider the following hypotheses

$$(\mathrm{H}_\rho^m) \qquad \begin{cases} \rho_0^1, \rho_0^2 \in L^m(\Omega), & \text{if } m \in (1, +\infty), \\ \|\rho_0^1 + \rho_0^2\|_{L^\infty(\Omega)} \leq 1 \text{ and } \mathscr{L}^d(\Omega) > M_1 + M_2, & \text{if } m = +\infty; \end{cases}$$

Notice that the structural condition $\mathscr{L}^d(\Omega) > M_1 + M_2$ in the case of $m = +\infty$ is needed in order to have nontrivial competitors that satisfy the upper bound constraint.

$$(\mathrm{H}_\Phi) \qquad\qquad\qquad\qquad\qquad \Phi_1, \Phi_2 \in W^{1,\infty}(\Omega).$$

First, let us derive the first order necessary optimality conditions for the minimizers in (MM_m).

Lemma 2.2 (Optimality conditions: m finite)**.** *Let $m \in (1, +\infty)$ and let Φ_1 and Φ_2 satisfy (H_Φ) and (ρ_0^1, ρ_0^2) satisfy (H_ρ^m). Let (ρ^1, ρ^2) be the unique minimizer in (MM_m) with $k = 0$. Then*

(1) *there exist Kantorovich potentials φ^i, $i = 1, 2$, in the transport of ρ^i onto ρ_0^i and $C_i \in \mathbb{R}$ ($i = 1, 2$) such that*

$$(2.5) \qquad\qquad \frac{m}{m-1}(\rho^1 + \rho^2)^{m-1} = \max\left(C_1 - \Phi_1 - \varphi^1/\tau; C_2 - \Phi_2 - \varphi^2/\tau; 0\right),$$

In particular $(\rho^1 + \rho^2)^{m-1}$ is Lipschitz continuous, $\rho^1 + \rho^2 \in C^{0,1/(m-1)}(\Omega)$, and these regularities degenerate as $\tau \downarrow 0$.

(2) *One can differentiate the above equality a.e. and the optimal transport maps T^i ($i = 1, 2$) in the transport of ρ^i onto ρ_0^i have the form*

$$T^i = \mathrm{id} + \tau\left(\frac{m}{m-1}\nabla(\rho^1 + \rho^2)^{m-1} + \nabla\Phi_i\right)$$

Proof. The proof of these results are just easy adaptations of the ones from Lemma A.1, thus we omit it. $\qquad\qquad\square$

Lemma 2.3 (Optimality conditions: $m = \infty$)**.** *Let $m = \infty$ and let Φ_1 and Φ_2 satisfy (H_Φ) and (ρ_0^1, ρ_0^2) satisfy (H_ρ^m). Let (ρ^1, ρ^2) be the unique minimizer in (MM_m) with $k = 0$. Then*

(1) *there exist Kantorovich potentials φ^i in the transport of ρ^i onto ρ_0^i ($i = 1, 2$) such that*

$$(2.6) \qquad\qquad \int_\Omega (\Phi_1 + \varphi^1/\tau)(\mu^1 - \rho^1)\mathrm{d}x + \int_\Omega (\Phi_2 + \varphi^2/\tau)(\mu^2 - \rho^2)\mathrm{d}x \geq 0,$$

for any $(\mu^1, \mu^2) \in \mathscr{P}^{M_1}(\Omega) \times \mathscr{P}^{M_2}(\Omega)$ such that $\mu^1 + \mu^2 \leq 1$ a.e. in Ω.

(2) *There exists a Lipschitz continuous pressure function p that can be defined via the Kantorovich potentials φ^1, φ^2 from (1) as*

$$(2.7) \qquad\qquad \nabla p = -\nabla\varphi^i/\tau - \nabla\Phi_i, \ \rho^i - \text{a.e.}, \ i = 1, 2,$$

and $p \geq 0$ and $p(1 - (\rho^1 + \rho^2)) = 0$ a.e. in Ω. In particular, the optimal transport map T^i in the transportation of ρ^i onto ρ_0^i ($i = 1, 2$) has the form

$$T^i = \mathrm{id} + \tau\left(\nabla p + \nabla\Phi_i\right).$$

Proof. The proof of the above results are adaptations of the ones from [30, Lemma 3.1-3.2] and [25, Lemma 6.11-Proposition 6.12], so we omit it.

$\qquad\qquad\square$

2.5. **Equilibrium solutions when $m < +\infty$.** Let us study the equilibrium solutions $(\bar{\rho}^1, \bar{\rho}^2)$ of the scheme (MM_m), meaning that $(\bar{\rho}^1, \bar{\rho}^2)$ is a minimizer of the *free energy* $\mathcal{F} + \mathcal{G}$. This exists by the l.s.c. and boundedness from below of the functional and the compactness of $\mathscr{P}^{M_1}(\Omega) \times \mathscr{P}^{M_2}(\Omega)$. Then writing down the first order optimality conditions as in Lemma 2.2, one obtains that

$$(2.8) \qquad\qquad \begin{cases} \frac{m}{m-1}(\bar{\rho}^1 + \bar{\rho}^2)^{m-1} = C_i - \Phi_i, & \text{in } \{\bar{\rho}^i > 0\}, \\ \frac{m}{m-1}(\bar{\rho}^1 + \bar{\rho}^2)^{m-1} \geq C_i - \Phi_i, & \text{in } \{\bar{\rho}^i = 0\}, \end{cases}$$

or in short

$$\frac{m}{m-1}(\bar{\rho}^1 + \bar{\rho}^2)^{m-1} = \max[C_1 - \Phi_1; C_2 - \Phi_2; 0],$$

for $i = 1, 2$ and for some constants $C_1, C_2 \in \mathbb{R}$. For simplicity in this informal discussion one may suppose that both Φ_1 and Φ_2 are strictly convex with a unique minimizer in Ω. Otherwise the constants C_i may vary on each connected component of $\{\bar{\rho}^i > 0\}$. Observe that the above conditions imply in particular that whenever the potentials Φ_1 and Φ_2 are different and their difference is not only a constant, then the phases $\bar{\rho}^1$ and $\bar{\rho}^2$ are separated, i.e. $\mathscr{L}^d\left(\{\bar{\rho}^1 > 0\} \cap \{\bar{\rho}^2 > 0\}\right) = 0$. Moreover, in general the interface $\{\bar{\rho}^1 > 0\} \cap \{\bar{\rho}^2 > 0\}$ is present and on the interface the densities ρ^i $(i = 1, 2)$ are positive. For instance this is the case when we take potentials $\Phi_1(x) = |x|^2$ and $\Phi_2 = 2|x|^2$ and C_1, C_2 are such that $0 < C_1 < C_2$ and both densities are present.

In fact, with the above choice of potentials Φ_i, $i = 1, 2$, suppose that we start our minimizing movements with initial configuration of densities $\rho_0^1 = \chi_{\{|x| \le 1\}}$ and $\rho_0^2 = \chi_{\{1 < |x| < 2\}}$. In the equilibrium limit we have $\{\bar{\rho}^2 > 0\} = \{|x| \le r_1\}$ and $\{\bar{\rho}^2 > 0\} = \{r_1 \le |x| \le r_2\}$ for some $0 < r_1 < r_2$. Thus, if solutions (ρ^1, ρ^2) of the system (PME_m) exist with these initial data and potentials, heuristically it is inevitable that the supports of ρ_t^1 and ρ_t^2 get mixed for some finite time $t > 0$, while ρ^1 "filtrates" through ρ^2 to change the ordering of their supports from the initial configuration. Such situation indicates low regularity for each density, and illustrates the difficulty in obtaining a strong notion of limit solutions for (PME_m) in the continuum limit. Indeed in general we are only able to obtain a very weak notion of solutions in the continuum limit, as we will see in Theorem 3.8. Deriving this weak notion of solutions in general settings is our first main result in the paper. To the best of the authors' knowledge, there does not seem to be a PDE approach to yield well-posedness on the continuum PDE (PME_m), especially when $\nabla \Phi_1 \ne \nabla \Phi_2$.

On the other hand, if the initial configuration of above example is in line with the potentials, i.e. if we switch the roles of ρ_0^1 and ρ_0^2, we expect the solutions to be well-behaved and to stay separated throughout the evolution, with stable interface in between them. It turns out that we can indeed show such separation in one spacial dimension. In this case stronger results are available, and one can derive stronger notion of solutions as well as the properties of the solutions and their interfaces in the incompressible limit $m \to \infty$, which in some cases leads to a type of two-phase Hele-Shaw flow with drifts (see Section 4.4).

2.6. Regularity of the minimizers in the (MM_m) scheme.

Theorem 2.4. *Let* $m \in (1, +\infty)$. *Let* $(\rho_0^1, \rho_0^2) \in \mathscr{P}^{M_1}(\Omega) \times \mathscr{P}^{M_2}(\Omega)$ *satisfying* (H_ρ^m) *and let* (H_Φ) *be fulfilled. Let* (ρ^1, ρ^2) *be the minimizer in* (MM_m) *constructed with the help of* (ρ_0^1, ρ_0^2). *Then*

$$(2.9) \qquad \rho^1, \rho^2 \in L^m(\Omega)$$

and

$$(2.10) \qquad (\rho^1 + \rho^2)^{m-1/2} \in H^1(\Omega).$$

If $m = +\infty$, $\rho^1 + \rho^2 \le 1$ *a.e. in* Ω.

Proof. First, setting $\boldsymbol{\rho} = (\rho^1, \rho^2)$ and $\boldsymbol{\rho}_0 = (\rho_0^1, \rho_0^2)$, by the optimality of $\boldsymbol{\rho}$ in (MM_m) w.r.t. $\boldsymbol{\rho}_0$, one obtains

$$\mathcal{F}_m(\boldsymbol{\rho}) = \frac{1}{m-1} \int_\Omega (\rho^1 + \rho^2)^m \, \mathrm{d}x \le \frac{1}{2\tau} \mathbf{W}_2^2(\boldsymbol{\rho}, \boldsymbol{\rho}_0) + \mathcal{F}_m(\boldsymbol{\rho}_0) + \mathcal{G}(\boldsymbol{\rho}_0) - \mathcal{G}(\boldsymbol{\rho})$$

$$\le \frac{1}{2\tau} \mathbf{W}_2^2(\boldsymbol{\rho}, \boldsymbol{\rho}_0) + \mathcal{F}_m(\boldsymbol{\rho}_0) + 2(M_1 \|\Phi_1\|_{L^\infty} + M_2 \|\Phi_2\|_{L^\infty}),$$

which by the assumptions (H_ρ^m) and (H_Φ) implies (2.9) for m finite. If $m = \infty$, then clearly $\rho^1 + \rho^2 \le 1$ a.e. in Ω.

Second, writing down the first order optimality conditions (see Lemma 2.2) for the above problem, one obtains

$$\frac{m}{m-1}(\rho^1 + \rho^2)^{m-1} + \Phi_i + \frac{\varphi^i}{\tau} = C_i, \ \text{ in } \{\rho^i > 0\}, \ i = 1, 2,$$

where φ^i is a Kantorovich potential in the optimal transport of ρ^i onto ρ_0^i. This potential is linked to the optimal transport map between these densities as $T^i(x) = x - \nabla\varphi^i(x)$. So, by Lemma 2.2(2) one can write

$$(2.11) \qquad -\frac{m}{m-1}\nabla(\rho^1 + \rho^2)^{m-1} - \nabla\Phi_i = \frac{\nabla\varphi^i}{\tau}, \ \rho^i - \text{a.e.}, \ i = 1, 2.$$

Since the r.h.s. of (2.11) is in $L^2_{\rho^i}(\Omega)$ with $\int_\Omega \frac{1}{\tau^2}|\nabla\varphi^i|^2\rho^i\,\mathrm{d}x = \frac{1}{\tau^2}W_2^2(\rho^i,\rho_0^i)$ and $\nabla\Phi_i \in L^2_{\rho^i}(\Omega;\mathbb{R}^d)$ we have the estimation

$$\int_\Omega \left|\nabla(\rho^1+\rho^2)^{m-1}\right|^2 \rho^i\,\mathrm{d}x \leq \frac{2(m-1)^2}{m^2}\left(\frac{1}{\tau^2}W_2^2(\rho^i,\rho_0^i) + M_i\|\nabla\Phi_i\|_{L^\infty}^2\right).$$

Adding up the two inequalities for $i = 1, 2$, one obtains after rearranging

$$(2.12) \qquad \int_\Omega \left|\nabla(\rho^1+\rho^2)^{m-1/2}\right|^2 \mathrm{d}x \leq \frac{2(m-1/2)^2}{m^2}\left(\frac{1}{\tau^2}\mathbf{W}_2^2(\boldsymbol{\rho},\boldsymbol{\rho}_0) + M_1\|\nabla\Phi_1\|_{L^\infty}^2 + M_2\|\nabla\Phi_2\|_{L^\infty}^2\right).$$

By the estimation (2.9) $\rho^1+\rho^2$ is bounded in $L^m(\Omega)$, so by the fact that Ω is compact, $\rho^1+\rho^2$ is summable in $L^q(\Omega)$ for any $1 \leq q \leq m$. This means in particular that the average can be bounded as

$$\fint_\Omega (\rho^1+\rho^2)^{m-1/2}\,\mathrm{d}x \leq \|\rho^1+\rho^2\|_{L^m}^{m-1/2}\mathscr{L}^d(\Omega)^{1/(2m)-1}$$

hence Poincaré's inequality yields that $\|(\rho^1+\rho^2)^{m-1/2}\|_{L^2(\Omega)}$ is bounded, more precisely

$$\|(\rho^1+\rho^2)^{m-1/2}\|_{L^2(\Omega)} \leq C_\Omega\|\nabla(\rho^1+\rho^2)^{m-1/2}\|_{L^2(\Omega)} + \mathscr{L}^d(\Omega)^{\frac{1}{2}}\fint_\Omega(\rho^1+\rho^2)^{m-1/2}\,\mathrm{d}x$$

$$= C_\Omega\|\nabla(\rho^1+\rho^2)^{m-1/2}\|_{L^2(\Omega)} + \|\rho^1+\rho^2\|_{L^m}^{m-1/2}\mathscr{L}^d(\Omega)^{1/(2m)-1/2}$$

where $C_\Omega > 0$ is the Poincaré constant associated to Ω. Thus, (2.10) follows. $\square$

3. The continuum limit solutions in general dimension

In this section we study the convergence of the time-discrete solutions in the continuum limit. The limit solutions can be interpreted as a *very weak* solution for both systems (PME$_m$) and (PME$_\infty$) in the following sense:

Definition 3.1 (Notion of weak solution). *By a* weak solution *of system* (PME$_m$) *we mean a pair* (ρ^1,ρ^2) *such that* $\rho^i \in AC^2([0,T];\mathscr{P}^{M_i}(\Omega)) \cap L^{2m-1}([0,T]\times\Omega)$, *and setting* $p := \frac{m}{m-1}(\rho^1+\rho^2)^{m-1}$, $\nabla p\rho^i \in L^r([0,T]\times\Omega;\mathbb{R}^d)$, *for some* $1 < r < 2$ $(i = 1,2)$. *Moreover* $\rho^i|_{t=0} = \rho_0^i$ $(i = 1,2)$ *and the equation*

$$(\text{Weak}) \qquad -\int_s^t\int_\Omega \rho^i\partial_t\phi\,\mathrm{d}x\mathrm{d}\tau + \int_s^t\int_\Omega \boldsymbol{v}^i\cdot\nabla\phi\rho^i\,\mathrm{d}x\mathrm{d}\tau = \int_\Omega \rho_s^i(x)\phi(s,x)\,\mathrm{d}x - \int_\Omega \rho_t^i(x)\phi(t,x)\,\mathrm{d}x,$$

holds true for all $\phi \in C^1([0,T]\times\Omega)$ *and for all* $0 \leq s < t \leq T$, *where*

$$\boldsymbol{v}^i := \nabla p + \nabla\Phi_i.$$

Similarly, by a weak solution *of* (PME$_\infty$) *we mean a triple* $(\rho^{1,\infty},\rho^{2,\infty},p^\infty)$ *such that* $\rho^{i,\infty} \in AC^2([0,T]; \mathscr{P}^{M_i}(\Omega)) \cap L^\infty([0,T]\times\Omega)$, $i = 1,2$ *with* $\|\rho^{1,\infty}+\rho^{2,\infty}\|_{L^\infty} \leq 1$, $p^\infty \in L^2([0,T];H^1(\Omega))$, $p^\infty \geq 0$ *and* $p^\infty(1-\rho^{1,\infty}-\rho^{2,\infty}) = 0$ *a.e. in* $[0,T]\times\Omega$. *Moreover* $\rho^i|_{t=0} = \rho_0^i$ $(i = 1,2)$ *and the equation* (Weak) *holds true with* p *replaced by* p^∞ *for all* $\phi \in C^1([0,T]\times\Omega)$ *and for all* $0 \leq s < t \leq T$.

We underline that the above weak formulations encode in particular no-flux boundary conditions on $[0,T]\times\partial\Omega$.

Remark 3.2. (a) *Notice that by density arguments, in the definition of the weak solution of* (PME$_m$) *one can consider* $\phi \in W^{1,1}([0,T];L^q(\Omega))\cap L^q([0,T];W^{1,q}(\Omega))$ *where* $q = \max\{r', (2m-1)'\}$ *and in the case of* (PME$_\infty$) *one can consider test functions in* $W^{1,1}([0,T];L^1(\Omega))\cap L^2([0,T];H^1(\Omega))$.

(b) *Also, by the fact that we impose that the densities are absolutely continuous curves in the Wasserstein space[1], imposing the initial conditions is meaningful.*

(c) *The uniqueness question of weak solutions seems to be very delicate and challenging.*

---

[1]See the Appendix on optimal transportation

3.1. **Interpolations between the densities.** Let $m \in (1, +\infty]$ and let us consider (ρ_0^1, ρ_0^2) and Φ_1 and Φ_2 satisfying the hypotheses (H_ρ^m) and (H_Φ) respectively. We consider also $T > 0$ a fixed time horizon, a time step $\tau > 0$ and $N \in \mathbb{N}$ such that $N\tau = T$ and the densities $(\rho_k^{1,\tau}, \rho_k^{2,\tau})_{k=0}^N$ obtained via the (MM_m) scheme starting from (ρ_0^1, ρ_0^2). We denote the optimal transport maps and the corresponding Kantorovich potentials between two consecutive densities $\rho_{k+1}^{i,\tau}$ and $\rho_k^{i,\tau}$ by T_k^i and φ_k^i respectively ($k \in \{0, \dots, N-1\}$, $i = 1, 2$). If $m = \infty$, we consider also the pressure variables p_k^τ ($k \in \{1, \dots, N\}$) constructed as in Lemma 2.3.

Since $\frac{\nabla \varphi_k^i}{\tau} = \frac{\mathrm{id} - T_k^i}{\tau}$ can be seen as a discrete velocity (displacement divided by time), it is reasonable to define the discrete velocity of the particles of the ith fluid located at $x \in \Omega$ (for a.e. $x \in \Omega$) as

$$(3.1) \qquad \mathbf{v}_k^{i,\tau}(x) := \begin{cases} -\frac{m}{m-1} \nabla(\rho_{k+1}^{1,\tau}(x) + \rho_{k+1}^{2,\tau}(x))^{m-1} - \nabla \Phi_i(x), & \text{if } m \in (1, +\infty), \\ -\nabla p_{k+1}^\tau(x) - \nabla \Phi_i(x), & \text{if } m = \infty. \end{cases}$$

As technical tools, we shall consider continuous and piecewise constant interpolations between the discrete densities. We will also work with the associated velocities and momenta. These constructions and the estimates on them are standard for experts and are very similar to the ones from [43, Chapter 8.3] and from [30]. We refer to [42] as well, as an overview of these techniques.

Continuous interpolations. Using McCann's interpolation – as it is done for instance in [43, Chapter 8.3] – we can consider families of continuous interpolations $[0, T] \ni t \mapsto (\rho_t^{1,\tau}, \rho_t^{2,\tau}) \in \mathscr{P}^{M_1}(\Omega) \times \mathscr{P}^{M_2}(\Omega)$ between the discrete in time densities parametrized with $\tau > 0$. We denote the corresponding time dependent families of velocities and momenta by $\mathbf{v}^{i,\tau}, \mathbf{E}^{i,\tau}$.

It is worth to notice that the above construction implies in particular that $(\rho^{i,\tau}, \mathbf{E}^{i,\tau})$ ($i = 1, 2$) solves the continuity equation

$$(3.2) \qquad \partial_t \rho^{i,\tau} + \nabla \cdot \mathbf{E}^{i,\tau} = 0.$$

on $[0, T] \times \Omega$ in the weak sense, i.e.

$$(3.3) \qquad \int_s^t \int_\Omega \rho^i \partial_t \phi \, \mathrm{d}x \mathrm{d}\tau + \int_s^t \int_\Omega \mathbf{E}^i \cdot \nabla \phi \, \mathrm{d}x \mathrm{d}\tau = -\int_\Omega \rho_s^i(x) \phi(s, x) \, \mathrm{d}x + \int_\Omega \rho_t^i(x) \phi(t, x) \, \mathrm{d}x$$

for all $\phi \in C^1([0, T] \times \Omega)$ and $0 \leq s < t \leq T$.

Piecewise constant interpolations. We consider a second family of interpolations, simply taking

$$(3.4) \qquad \tilde{\rho}_t^{i,\tau} := \rho_{k+1}^{i,\tau}, \quad \tilde{\mathbf{v}}_t^{i,\tau} := \mathbf{v}_k^{i,\tau}, \text{ and } \widetilde{\mathbf{E}}_t^{i,\tau} := \tilde{\rho}_t^{i,\tau} \tilde{\mathbf{v}}_t^{i,\tau} \text{ for } t \in [k\tau, (k+1)\tau).$$

We consider the piecewise constant interpolation for the pressure variable (see Lemma 2.3) as well, i.e.
$$(3.5)$$
$$\tilde{p}_t^\tau := \begin{cases} (C_1 - \Phi_1 - \varphi_k^1/\tau)_+, & \text{in } \{\rho_{k+1}^{1,\tau} > 0\}, \\ (C_2 - \Phi_2 - \varphi_k^2/\tau)_+, & \text{in } \{\rho_{k+1}^{2,\tau} > 0\}, \qquad \text{for } t \in [k\tau, (k+1)\tau), k \in \{0, \dots, N-1\}. \\ 0, & \text{in } \Omega \setminus \left(\{\rho_{k+1}^{1,\tau} > 0\} \cup \{\rho_{k+1}^{2,\tau} > 0\} \right), \end{cases}$$

In addition we set $\tilde{\boldsymbol{\rho}}_t^\tau := (\tilde{\rho}_t^{1,\tau}, \tilde{\rho}_t^{2,\tau})$ for all $t \in [0, T]$. We remark that by construction one has $\tilde{\rho}_t^{i,\tau} = \rho_t^{i,\tau}$ for $t = k\tau$, $k \in \{1, \dots N\}$.

3.2. **A priori estimates for the interpolations.** We discuss now some estimates on the interpolations that will be useful to pass to the limit as $\tau \downarrow 0$. In general, all the constants in the estimates depend on the data $\rho_0^1, \rho_0^2, \Phi_1, \Phi_2, T$ and m, however it will be especially important to keep track the precise dependence of them on m (in particular we use these estimates also in the limiting procedure when $m \to +\infty$). To highlight this dependence, we denote the constants as $C(m)$.

Lemma 3.1. *For any $m \in (1, +\infty]$, $\tau > 0$ and any $k \in \{0, \dots, N-1\}$ one has*

$$(3.6) \qquad \frac{1}{2\tau} \sum_{k=0}^{N-1} \boldsymbol{W}_2^2(\boldsymbol{\rho}_{k+1}^\tau, \boldsymbol{\rho}_k^\tau) \leq \mathcal{F}_m(\boldsymbol{\rho}_0^\tau) + \mathcal{G}(\boldsymbol{\rho}_0^\tau) - \mathcal{F}_m(\boldsymbol{\rho}_N^\tau) - \mathcal{G}(\boldsymbol{\rho}_N^\tau),$$

and

$$(3.7) \qquad \|\rho_{k+1}^{1,\tau} + \rho_{k+1}^{2,\tau}\|_{L^m(\Omega)} \leq C_1(m),$$

where

$$(3.8) \quad C_1(m) := \begin{cases} \left((2m-2) \left(M_1 \|\Phi_1\|_{L^\infty(\Omega)} + M_2 \|\Phi_2\|_{L^\infty(\Omega)} \right) + \|\rho_0^1 + \rho_0^2\|_{L^m(\Omega)}^m \right)^{1/m}, & \text{if } m < \infty, \\ 1, & \text{if } m = \infty. \end{cases}$$

Proof. The proofs of both inequalities are immediate by the optimality of $\boldsymbol{\rho}_{k+1}^\tau$ w.r.t. $\boldsymbol{\rho}_k^\tau$ in (MM_m). So we omit them. $\qquad\square$

Corollary 3.2. *Hypotheses* (H_Φ) *and* (H_ρ^m) *imply that*

$$\frac{1}{\tau} \sum_{k=0}^{N-1} \left(W_2^2(\rho_k^{1,\tau}, \rho_{k-1}^{1,\tau}) + W_2^2(\rho_k^{2,\tau}, \rho_{k-1}^{2,\tau}) \right) \leq C_2(m),$$

where

$$(3.9) \qquad C_2(m) := \begin{cases} \frac{2}{m-1} \|\rho_0^1 + \rho_0^2\|_{L^m}^m + 4M_1 \|\Phi_1\|_{L^\infty} + 4M_2 \|\Phi_2\|_{L^\infty}, & \text{if } m \in (1, +\infty), \\ 4M_1 \|\Phi_1\|_{L^\infty} + 4M_2 \|\Phi_2\|_{L^\infty}, & \text{if } m = +\infty. \end{cases}$$

is independent of τ.

Lemma 3.3 (Bounds for $\rho^{i,\tau}$, $\mathbf{v}^{i,\tau}$ and $\mathbf{E}^{i,\tau}$)**.** *Assume that we constructed the discrete densities $\rho_k^{i,\tau}$ for $\tau > 0$, $k \in \{0, \dots, N\}$ and $i = 1, 2$. Let $\rho^{i,\tau}$ be the continuous interpolations and let $\mathbf{v}^{i,\tau}$ and $\mathbf{E}^{i,\tau}$ be the associated velocity field and momentum variables respectively. Then*

(1) *$\rho^{i,\tau}$ is bounded in $AC^2([0,T]; (\mathscr{P}^{M_i}(\Omega), W_2))$ uniformly in $\tau > 0$;*
(2) *$\mathbf{v}^{i,\tau}$ is bounded in $L^2([0,T]; L^2_{\rho^{i,\tau}}(\Omega; \mathbb{R}^d))$ uniformly in $\tau > 0$;*
(3) *$\mathbf{E}^{i,\tau}$ and $\widetilde{\mathbf{E}}^{i,\tau}$ are bounded in $\mathscr{M}^d([0,T] \times \Omega)$ uniformly in $\tau > 0$.*

Proof. For $\tau > 0$, by construction $\rho^{i,\tau}$ is a constant speed geodesic interpolation with the corresponding velocity field $\mathbf{v}^{i,\tau}$. This implies that

$$\int_0^T \|\mathbf{v}_t^{i,\tau}\|_{L^2_{\rho_t^{i,\tau}}}^2 \, \mathrm{d}t = \int_0^T |(\rho^{i,\tau})'|_{W_2}^2(t) \mathrm{d}t = \sum_{k=1}^{N-1} \frac{1}{\tau} W_2^2(\rho_{k-1}^{i,\tau}, \rho_k^{i,\tau}) \leq C_2(m).$$

Now, by Corollary 3.2 we obtain that (1)-(2) hold true.

To estimate the total variation of $\mathbf{E}^{i,\tau}$ we write

$$|\mathbf{E}^{i,\tau}|([0,T] \times \Omega) = \int_0^T \int_\Omega |\mathbf{v}_t^{i,\tau}| \rho_t^{i,\tau} \, \mathrm{d}x \mathrm{d}t \leq \int_0^T \left(\int_\Omega |\mathbf{v}_t^{i,\tau}|^2 \rho_t^{i,\tau} \, \mathrm{d}x \right)^{\frac{1}{2}} \left(\int_\Omega \rho_t^{i,\tau} \, \mathrm{d}x \right)^{\frac{1}{2}} \mathrm{d}t$$

$$\leq \sqrt{M_i} \sqrt{T} \left(\int_0^T \int_\Omega |\mathbf{v}_t^{i,\tau}|^2 \rho_t^{i,\tau} \, \mathrm{d}x \mathrm{d}t \right)^{\frac{1}{2}} \leq \sqrt{M_i T C_2(m)}.$$

In the last inequality we used the previously obtained bound on $\mathbf{v}^{i,\tau}$. The bound on $\widetilde{\mathbf{E}}^{i,\tau}$ rely on the same argument. $\qquad\square$

Lemma 3.4 (Bounds on $\tilde{p}^\tau$)**.** *Let us consider the piecewise constant interpolation $[0,T] \ni t \mapsto \tilde{p}_t^\tau$ of the pressure variables defined in* (3.5). *Then $\tilde{p}^\tau$ is bounded $L^2([0,T]; H^1(\Omega))$ independently of $\tau > 0$.*

Proof. The proof is similar to the ones in [25, Proposition 6.13] and [36, Lemma 3.6]. We sketch it below. Let us use the fact that $\nabla p_k^\tau = -\nabla \varphi_k^i / \tau - \nabla \Phi_i$, $\rho_{k+1}^{i,\tau}$ — a.e., for all $k \in \{0, \dots, N-1\}$ where φ_k^i is an optimal Kantorovich potential in the transport of $\rho_{k+1}^{i,\tau}$ onto $\rho_k^{i,\tau}$. First, let us compute

$$\int_\Omega |\nabla p_k^\tau|^2 \rho_{k+1}^i \mathrm{d}x \leq \frac{2}{\tau^2} \int_\Omega |\nabla \varphi_k^i|^2 \rho_{k+1}^i \mathrm{d}x + 2 \int_\Omega |\nabla \Phi_i|^2 \rho^i \mathrm{d}x = \frac{2}{\tau^2} W_2^2(\rho_{k+1}^{i,\tau}, \rho_k^{i,\tau}) + 2\|\nabla \Phi_i\|_{L^\infty(\Omega)}^2,$$

then adding up the two inequalities for $i = 1, 2$ (using the fact that p_k^τ is supported on $\{\rho_{k+1}^1 + \rho_{k+1}^2 = 1\}$), one finds

$$\int_\Omega |\nabla p_k^\tau|^2 \mathrm{d}x = \int_\Omega |\nabla p_k^\tau|^2 (\rho_{k+1}^1 + \rho_{k+1}^2) \mathrm{d}x \leq \sum_{i=1}^2 \left(\frac{2}{\tau^2} W_2^2(\rho_{k+1}^{i,\tau}, \rho_k^{i,\tau}) + 2\|\nabla \Phi_i\|_{L^\infty(\Omega)}^2 \right).$$

Integrating in time $\|\nabla \tilde{p}_t^\tau\|_{L^2(\Omega)}$ using Corollary 3.2 one has that $\nabla \tilde{p}^\tau \in L^2([0,T] \times \Omega)$. Using the fact that $\mathscr{L}^d(\{\tilde{p}_t^\tau = 0\}) \geq \mathscr{L}^d(\{\tilde{\rho}_t^{1,\tau} + \tilde{\rho}_t^{1,\tau} < 1\}) \geq \mathscr{L}^d(\Omega) - M_1 - M_2 > 0$ for a.e. $t \in [0,T]$, one concludes by a suitable version of Poincaré's inequality that $\tilde{p}^\tau$ uniformly bounded in $L^2([0,T]; H^1(\Omega))$ as desired. $\qquad\square$

We show now gradient estimates, derived from the optimality conditions (2.11).

Theorem 3.5. *Let $m \in (1, +\infty)$. Then for the piecewise constant interpolation $\tilde{\rho}^{i,\tau}$ $(i = 1, 2)$ introduced in (3.4) one has*

$$(3.10) \qquad \|\nabla(\tilde{\rho}^{1,\tau} + \tilde{\rho}^{2,\tau})^{m-1/2}\|_{L^2([0,T]\times\Omega)} \leq C_3(m), \quad \|(\tilde{\rho}^{1,\tau} + \tilde{\rho}^{2,\tau})^{m-1/2}\|_{L^2([0,T]\times\Omega)} \leq C_4(m),$$

and

$$(3.11) \qquad \|\tilde{\rho}^{1,\tau} + \tilde{\rho}^{2,\tau}\|_{L^q([0,T];L^m(\Omega))} \leq C_5(q, m), \ \forall q \geq 1,$$

where $C_3(m), C_4(m), C_5(q, m) > 0$ are constants independent of $\tau > 0$. The first two bounds imply in particular that

$$(3.12) \qquad \|(\tilde{\rho}^{1,\tau} + \tilde{\rho}^{2,\tau})^{m-1/2}\|_{L^2([0,T];H^1(\Omega))} \leq \left(C_3(m)^2 + C_4(m)^2\right)^{1/2}.$$

Proof. We use the inequality (2.12), writing for $(\rho_{k+1}^{1,\tau}, \rho_{k+1}^{2,\tau})$, i.e.

$$\int_\Omega \left|\nabla(\rho_{k+1}^{1,\tau} + \rho_{k+1}^{2,\tau})^{m-1/2}\right|^2 \mathrm{d}x \leq \frac{2(m-1/2)^2}{m^2}\left(\sum_{i=1}^2 \frac{1}{\tau^2}W_2^2(\rho_{k+1}^{i,\tau}, \rho_k^{i,\tau}) + \sum_{i=1}^2 M_i\|\nabla\Phi_i\|_{L^\infty}^2\right).$$

Since the curves $\tilde{\rho}^{i,\tau}$ $(i = 1, 2)$ are piecewise constant interpolations, i.e. $\tilde{\rho}_t^{i,\tau} = \rho_{k+1}^{i,\tau}$ for $t \in (k\tau, (k+1)\tau]$, one has

$$\int_0^T \int_\Omega \left|\nabla(\tilde{\rho}_t^{1,\tau} + \tilde{\rho}_t^{2,\tau})^{m-1/2}\right|^2 \mathrm{d}x\mathrm{d}t = \tau \sum_{k=0}^{N-1} \int_\Omega \left|\nabla(\rho_{k+1}^{1,\tau} + \rho_{k+1}^{2,\tau})^{m-1/2}\right|^2 \mathrm{d}x$$

$$\leq \frac{2(m-1/2)^2}{m^2} \sum_{i=1}^2 \sum_{k=0}^{N-1} \frac{1}{\tau}W_2^2(\rho_{k+1}^{i,\tau}, \rho_k^{i,\tau})$$

$$+ \frac{2(m-1/2)^2}{m^2}\tau \sum_{i=1}^2 \sum_{k=0}^{N-1} M_i\|\nabla\Phi_i\|_{L^\infty}^2$$

$$\leq \frac{2(m-1/2)^2}{m^2}\left(C_2(m) + T\sum_{i=1}^2 M_i\|\nabla\Phi_i\|_{L^\infty}^2\right) =: C_3(m)^2,$$

which implies (3.10), with

$$(3.13) \qquad C_3(m) := \frac{\sqrt{2}(m-1/2)}{m}\left(C_2(m) + T\sum_{i=1}^2 M_i\|\nabla\Phi_i\|_{L^\infty}^2\right)^{1/2}.$$

Similarly, using the estimations from Theorem 2.4 and (3.7), we can write

$$\int_0^T \|(\tilde{\rho}_t^{1,\tau} + \tilde{\rho}_t^{2,\tau})^{m-1/2}\|_{L^2(\Omega)}^2 \mathrm{d}t = \tau \sum_{k=0}^{N-1} \|(\rho_{k+1}^{1,\tau} + \rho_{k+1}^{2,\tau})^{m-1/2}\|_{L^2(\Omega)}^2$$

$$\leq \tau \sum_{k=0}^{N-1}\left(C_\Omega\|\nabla(\rho_{k+1}^{1,\tau} + \rho_{k+1}^{2,\tau})^{m-1/2}\|_{L^2(\Omega)} + \|\rho_{k+1}^{1,\tau} + \rho_{k+1}^{2,\tau}\|_{L^m(\Omega)}^{m-1/2}\mathscr{L}^d(\Omega)^{1/(2m)-1/2}\right)^2$$

$$\leq 2\tau \sum_{k=0}^{N-1}\left(C_\Omega^2\|\nabla(\rho_{k+1}^{1,\tau} + \rho_{k+1}^{2,\tau})^{m-1/2}\|_{L^2(\Omega)}^2 + \|\rho_{k+1}^{1,\tau} + \rho_{k+1}^{2,\tau}\|_{L^m(\Omega)}^{2m-1}\mathscr{L}^d(\Omega)^{1/m-1}\right)$$

$$\leq 2\left(C_\Omega^2 C_3(m)^2 + TC_1(m)^{2m-1}\mathscr{L}^d(\Omega)^{1/m-1}\right)$$

Thus, the second estimation in (3.10) holds true with

$$(3.14) \qquad C_4(m) := \sqrt{2}\left(C_\Omega^2 C_3(m)^2 + TC_1(m)^{2m-1}\mathscr{L}^d(\Omega)^{1/m-1}\right)^{1/2}.$$

Using (3.7), for any $q \geq 1$ we can write similarly as before

$$\int_0^T \|\tilde{\rho}_t^{1,\tau} + \tilde{\rho}_t^{2,\tau}\|_{L^m(\Omega)}^q \, dt = \tau \sum_{k=0}^{N-1} \|\rho_{k+1}^{1,\tau} + \rho_{k+1}^{2,\tau}\|_{L^m(\Omega)}^q \leq TC_1(m)^q.$$

So defining $C_5(q,m) := T^{1/q} C_1(m)$, one obtains the last estimation (3.11)   □

In what follows – using a refined version of the Aubin-Lions lemma – we prove a strong compactness result for $\tilde{\rho}^{1,\tau} + \tilde{\rho}^{2,\tau}$ where $\tilde{\rho}^{1,\tau}$ and $\tilde{\rho}^{2,\tau}$ are the piecewise constant interpolations.

Proposition 3.6. *Let $m \in (1, +\infty)$. Then the sequence of curves defined as $\tilde{\rho}^{1,\tau_n} + \tilde{\rho}^{2,\tau_n}$ (for any sequence $(\tau_n)_{n \geq 0}$ of positive reals that converges to 0) is strongly pre-compact in $L^{2m-1}([0,T] \times \Omega)$.*

Proof. We will use a refined version of the classical Aubin-Lions lemma to prove this result (see [40] and Theorem B.1). Then we will argue as in in [19].

Let us set $B := L^{2m-1}(\Omega)$, $\mathfrak{F} : L^{2m-1}(\Omega) \to [0, +\infty]$ defined as

$$\mathfrak{F}(\rho) := \begin{cases} \|\rho^{m-1/2}\|_{H^1(\Omega)}, & \text{if } \rho \in H^1(\Omega) \cap \mathscr{P}^{M_1+M_2}(\Omega), \\ +\infty, & \text{otherwise} \end{cases}$$

and $g : L^{2m-1}(\Omega) \times L^{2m-1}(\Omega) \to [0, +\infty]$ defined as

$$g(\mu, \nu) := \begin{cases} W_2(\mu, \nu), & \text{if } \mu, \nu \in \mathscr{P}^{M_1+M_2}(\Omega), \\ +\infty, & \text{otherwise.} \end{cases}$$

In this setting, $\left(\tilde{\rho}^{1,\tau_n} + \tilde{\rho}^{2,\tau_n}\right)_{n \geq 0}$ and $\mathfrak{F}$ satisfy the assumptions of Theorem B.1. Indeed, from Theorem 3.5 one has in particular that $\int_0^T \|(\tilde{\rho}_t^{1,\tau} + \tilde{\rho}_t^{2,\tau})^{m-1/2}\|_{H^1(\Omega)}^2 \, dt \leq C_3(m)^2 + C_4(m)^2$. The injection $H^1(\Omega) \hookrightarrow L^2(\Omega)$ is compact, the injection $i : \eta \mapsto \eta^{\frac{2}{2m-1}}$ is continuous from $L^2(\Omega)$ to $L^{2m-1}(\Omega)$ and the sub-level sets of $\rho \mapsto \|\rho^{m-1/2}\|_{H^1(\Omega)}$ are compact in $L^{2m-1}(\Omega)$.

Moreover, by Corollary 3.2, Lemma A.3 and by the fact that g defines a distance on $D(\mathfrak{F})$, one has that g also satisfies the assumptions from Theorem B.1, hence the implication of the theorem holds and one has that $\left(\tilde{\rho}^{1,\tau_n} + \tilde{\rho}^{2,\tau_n}\right)_{n \geq 0}$ is pre-compact in $\mathscr{M}(0,T; L^{2m-1})$. Finally, the uniform bound (3.7) implies the strong pre-compactness of $\left(\tilde{\rho}^{1,\tau_n} + \tilde{\rho}^{2,\tau_n}\right)_{n \geq 0}$ in $L^{2m-1}([0,T] \times \Omega)$.   □

3.3. **The limit systems as $\tau \downarrow 0$.** We proceed with the final step of our scheme, i.e. as the time step size goes to zero, we show that along a subsequence the discrete solutions converge to yield a *very weak solution* of the PDE systems, in the sense of Definition 3.1. We use the convention of $L^{2m-1}([0,T] \times \Omega) = L^\infty([0,T] \times \Omega)$ whenever $m = +\infty$.

Proposition 3.7. *Let $m \in (1, +\infty]$ and let us consider any sequence $(\tau_n)_{n \geq 0}$ which converges to zero. Then, along a subsequence the following holds:*

(1) *There exists $\rho^i \in AC^2([0,T]; (\mathscr{P}^{M_i}(\Omega), W_2)) \cap L^{2m-1}([0,T] \times \Omega)$ $(i = 1, 2)$ s.t. $\rho^{i,\tau_n} \to \rho^i$ and $\tilde{\rho}^{i,\tau_n} \to \rho^i$ as $n \to +\infty$ uniformly on $[0,T]$ w.r.t. W_2, in particular weakly-$\star$ in $\mathscr{P}^{M_i}(\Omega)$ for all $t \in [0,T]$.*

(2) *There exists $\mathbf{E}^i \in \mathscr{M}^d([0,T] \times \Omega)$ $(i = 1, 2)$ s.t. $\mathbf{E}^{i,\tau_n} \overset{\star}{\rightharpoonup} \mathbf{E}^i$ and $\widetilde{\mathbf{E}}^{i,\tau_n} \overset{\star}{\rightharpoonup} \mathbf{E}^i$ as $n \to +\infty$.*

Proof. In Lemma 3.3 we obtained uniform bounds on the metric derivative of the continuous interpolations ρ^{i,τ_n} $(i = 1, 2)$, which is enough to get compactness. More precisely there exist $[0,T] \ni t \mapsto \rho_t^i \in \mathscr{P}^{M_i}(\Omega)$, $i = 1, 2$ continuous curves such that (up to taking subsequences for τ_n) $\rho_t^{i,\tau_n} \to \rho_t^i$ uniformly on $[0,T]$ w.r.t. W_2 as $n \to +\infty$, in particular weakly-$\star$ in $\mathscr{P}^{M_i}(\Omega)$ for all $t \in [0,T]$.

The other interpolation $\tilde{\rho}^{i,\tau_n}$ coincides with ρ^{i,τ_n} at every node point $k\tau$, hence it is straightforward that (up to a subsequence taken for τ_n) it converges to the same curve ρ^i uniformly on $[0,T]$ w.r.t. W_2.

Lemma 3.3 states also that $\mathbf{E}^{i,\tau_n}$ and $\widetilde{\mathbf{E}}^{i,\tau_n}$ are uniformly bounded sequences in $\mathscr{M}^d([0,T] \times \Omega)$, hence there exist $\mathbf{E}^i \in \mathscr{M}^d([0,T] \times \Omega)$ such that (up to a subsequence taken for τ_n) $\mathbf{E}^{i,\tau_n} \overset{\star}{\rightharpoonup} \mathbf{E}^i$ and $\widetilde{\mathbf{E}}^{i,\tau_n} \overset{\star}{\rightharpoonup} \mathbf{E}^i$ $(i = 1, 2)$ in $\mathscr{M}^d([0,T] \times \Omega)$ as $n \to +\infty$. The convergence of $\mathbf{E}^{i,\tau_n}$ and $\widetilde{\mathbf{E}}^{i,\tau_n}$ to the same limit $\mathbf{E}^i$ follows from the same argument as in the proof of [36, Theorem 3.1].   □

These convergences imply that one can pass to the limit in the weak formulation (3.3) as $\tau \downarrow 0$ and obtain that $(\rho^i, \mathbf{E}^i)$ solves as well the continuity equation

$$(3.15) \qquad \partial_t \rho^i + \nabla \cdot \mathbf{E}^i = 0$$

on $[0, T] \times \Omega$ (with initial condition $\rho^i(0, \cdot) = \rho_0^i$) in the same weak sense.

In particular, by Lemma 3.3 one has that the sequence $\left(\mathcal{B}_2(\rho^{i,\tau_n}, \mathbf{E}^{i,\tau_n})\right)_{n \in \mathbb{N}}$ is uniformly bounded for any positive vanishing sequence $(\tau_n)_{n \in \mathbb{N}}$, where $\mathcal{B}_2$ denotes the Benamou-Brenier action functional (see its precise definition and properties in Appendix A). In particular, by the lower semicontinuity of this functional, there exists $\mathbf{v}^i$ such that $\mathbf{v}_t^i \in L_{\rho^i}^2(\Omega; \mathbb{R}^d)$ for a.e. $t \in [0, T]$ and at the limit (as $\tau \downarrow 0$) $\mathbf{E}^i = \mathbf{v}^i \cdot \rho^i$. This implies further that the equation (3.15) has the form

$$(3.16) \qquad \partial_t \rho^i + \nabla \cdot (\mathbf{v}^i \rho^i) = 0.$$

3.3.1. *Precise form of the limit systems.* Now we shall work with the piecewise constant interpolations $\tilde{\rho}^{i,\tau}, i = 1, 2$ and with the corresponding momenta $\widetilde{\mathbf{E}}^{i,\tau}, i = 1, 2$ to determine more properties of the limit systems. The more precise convergence results are summarized in Theorem 3.8 and 3.9 below.

Theorem 3.8. *Let $m \in (1, +\infty)$ and let $\tilde{\rho}^{i,\tau}, i = 1, 2$ be the piecewise constant interpolations between the densities $(\rho_k^{i,\tau})_{k=0}^N$ and $\widetilde{\mathbf{E}}^{i,\tau}, i = 1, 2$ the corresponding momentum variables. Taking any sequence $(\tau_n)_{n \geq 1}$ that goes to zero, the following holds along a subsequence:*

(1) $\left(\tilde{\rho}^{1,\tau_n} + \tilde{\rho}^{2,\tau_n}\right)_{n \geq 0}$ *converges strongly in $L^{2m-1}([0, T] \times \Omega)$ to $\rho^1 + \rho^2$;*

(2) $\left(\widetilde{\boldsymbol{E}}^{1,\tau_n} + \widetilde{\boldsymbol{E}}^{2,\tau_n}\right)_{n \geq 0}, i = 1, 2$ *converges in the sense of distributions to $-\nabla(\rho^1 + \rho^2)^m - \nabla\Phi_1\rho^1 - \nabla\Phi_2\rho^2$, i.e.,*

$$(3.17) \qquad \boldsymbol{v}^1\rho^1 + \boldsymbol{v}^2\rho^2 = -\nabla(\rho^1 + \rho^2)^m - \nabla\Phi_1\rho^1 - \nabla\Phi_2\rho^2$$

in the sense of distributions on $[0, T] \times \Omega$.

Proof. Let us show (1). Proposition 3.6 implies already that $\tilde{\rho}^{1,\tau_n} + \tilde{\rho}^{2,\tau_n}$ (up to some subsequence that we do not relabel) converges strongly in $L^{2m-1}([0, T] \times \Omega)$. Also, by Proposition 3.7 we have that $\tilde{\rho}_t^{i,\tau_n} \xrightarrow{\star} \rho_t^i$ as $n \to \infty$ for all $t \in [0, T]$. Hence the limit of $(\tilde{\rho}^{1,\tau_n} + \tilde{\rho}^{2,\tau_n})_{n \geq 0}$ is precisely $\rho^1 + \rho^2$ and $\rho^i \in L^{2m-1}([0, T] \times \Omega) \cap AC^2([0, T]; (\mathscr{P}^{M_i}(\Omega), W_2)), i = 1, 2$.

We show now (2). By definition of $\widetilde{\mathbf{E}}^{i,\tau}$ on has that

$$\widetilde{\mathbf{E}}^{1,\tau} + \widetilde{\mathbf{E}}^{2,\tau} = -\nabla(\tilde{\rho}^{1,\tau} + \tilde{\rho}^{2,\tau})^m - \nabla\Phi_1\tilde{\rho}^{1,\tau} - \nabla\Phi_2\tilde{\rho}^{2,\tau}.$$

Since by (1) $\tilde{\rho}^{1,\tau_n} + \tilde{\rho}^{2,\tau_n} \to \rho^1 + \rho^2$ strongly in $L^{2m-1}([0, T] \times \Omega)$ as $n \to +\infty$, one has that $(\tilde{\rho}^{1,\tau_n} + \tilde{\rho}^{2,\tau_n})^m \to (\rho^1 + \rho^2)^m$ strongly in $L^{2-\frac{1}{m}}([0, T] \times \Omega)$ as $n \to +\infty$. This, together with the weak$-\star$ convergence of $(\tilde{\rho}^{i,\tau_n})_{n \geq 0}$ to ρ^i implies the first part of the statement. On the other hand one has obtained already that $\widetilde{\mathbf{E}}^{i,\tau_n} \rightharpoonup \mathbf{E}^i = \mathbf{v}^i\rho^i$ as $n \to +\infty$, thus (3.17) follows as well. $\square$

Remark 3.3. (1) *Let us underline the fact that it is unclear whether we could show a stronger version of Theorem 3.8(2), i.e. the convergence (up to passing to a subsequence) of $\left(\widetilde{\boldsymbol{E}}^{i,\tau_n}\right)_{n \geq 0}$ to $-\frac{m}{m-1}\rho^i\nabla(\rho^1 + \rho^2)^{m-1} - \nabla\Phi_i\rho^i \ i = 1, 2$, which is necessary in order to obtain the weak formulation of the PDE system at the limit .*

(2) *In Theorem 3.8 if in addition $\left(\tilde{\rho}^{i,\tau_n}\right)_{n \geq 0}$ either for $i = 1$ or $i = 2$ converges a.e. in $[0, T] \times \Omega$ then both sequences $(i = 1, 2)$ converge strongly in $L^{2m-1}([0, T] \times \Omega)$ to ρ^i and the corresponding momentum $\left(\widetilde{\boldsymbol{E}}^{i,\tau_n}\right)_{n \geq 0}$ $(i = 1, 2)$ converge in the sense of distributions to $-\frac{m}{m-1}\rho^i\nabla(\rho^1 + \rho^2)^{m-1} - \nabla\Phi_i\rho^i$. The study of a stable scenario when this holds true is the subject of Section 4.*

Proof of Remark 3.3(2). First, clearly the pointwise convergence of $\left(\tilde{\rho}^{i,\tau_n}\right)_{n \geq 0}$ and the strong convergence of $\left(\tilde{\rho}^{1,\tau_n} + \tilde{\rho}^{2,\tau_n}\right)_{n \geq 0}$ in $L^{2m-1}([0, T] \times \Omega)$ by a suitable version of Vitali's convergence theorem imply the strong convergence of $\left(\tilde{\rho}^{i,\tau_n}\right)_{n \geq 0}$ in $L^{2m-1}([0, T] \times \Omega)$ to ρ^i. Since one of the terms in the sum of these sequences and the sum itself converges strongly, so does the other term as well. Let us recall the formula

$$\widetilde{\mathbf{E}}^{i,\tau_n} = -\frac{m}{m-1}\tilde{\rho}^{i,\tau_n}\nabla(\tilde{\rho}^{1,\tau_n} + \tilde{\rho}^{2,\tau_n})^{m-1} - \nabla\Phi_1\tilde{\rho}^{1,\tau_n}.$$

The strong convergence of $\left(\tilde\rho^{i,\tau_n}\right)_{n\geq 0}$ implies that the second term of $\widetilde{\mathbf{E}}^{i,\tau_n}$, i.e. $-\nabla\Phi_i\tilde\rho^{i,\tau_n}$ (since $\nabla\Phi_i$ is in $L^\infty(\Omega;\mathbb{R}^d)$) converges strongly in $L^{2m-1}([0,T]\times\Omega;\mathbb{R}^d)$ to $-\nabla\Phi_i\rho^i$. Thus in particular weakly-$\star$ in $\mathscr{M}^d([0,T]\times\Omega)$. The first term of $\widetilde{\mathbf{E}}^{i,\tau_n}$ can be written as

$$-\frac{m}{m-1}\tilde\rho^{i,\tau_n}\nabla(\tilde\rho^{1,\tau_n}+\tilde\rho^{2,\tau_n})^{m-1} = -\frac{m}{m-1}\left[\nabla(\tilde\rho^{1,\tau_n}+\tilde\rho^{2,\tau_n})^{m-1}\right](\tilde\rho^{1,\tau_n}+\tilde\rho^{2,\tau_n})^{1/2}\frac{\tilde\rho^{i,\tau_n}}{(\tilde\rho^{1,\tau_n}+\tilde\rho^{2,\tau_n})^{1/2}},$$

and notice furthermore that

$$-\frac{m}{m-1}\left[\nabla(\tilde\rho^{1,\tau_n}+\tilde\rho^{2,\tau_n})^{m-1}\right](\tilde\rho^{1,\tau_n}+\tilde\rho^{2,\tau_n})^{1/2} = -\frac{m}{m-1/2}\nabla(\tilde\rho^{1,\tau_n}+\tilde\rho^{2,\tau_n})^{m-1/2}.$$

Theorem 3.5 implies that the sequence $\left(-\frac{m}{m-1/2}(\tilde\rho^{1,\tau_n}+\tilde\rho^{2,\tau_n})^{m-1/2}\right)_{n\geq 0}$ is uniformly bounded in the space $L^2([0,T];H^1(\Omega))$, hence there exists a subsequence (not relabeled) and some $\xi\in L^2([0,T];H^1(\Omega))$ such that $\left(-\frac{m}{m-1/2}(\tilde\rho^{1,\tau_n}+\tilde\rho^{2,\tau_n})^{m-1/2}\right)_{n\geq 0}$ is converging weakly to ξ as $n\to+\infty$. In particular, $-\frac{m}{m-1/2}(\tilde\rho^{1,\tau_n}+\tilde\rho^{2,\tau_n})^{m-1/2}\rightharpoonup\xi$ weakly in $L^2([0,T]\times\Omega)$ and $-\frac{m}{m-1/2}\nabla(\tilde\rho^{1,\tau_n}+\tilde\rho^{2,\tau_n})^{m-1/2}\rightharpoonup\nabla\xi$ weakly in $L^2([0,T]\times\Omega;\mathbb{R}^d)$ as $n\to+\infty$.

By Proposition 3.6 one has that $\left(\tilde\rho^{1,\tau_n}+\tilde\rho^{2,\tau_n}\right)_{n\geq 0}$ converges strongly to $\rho^1+\rho^2$ in $L^{2m-1}([0,T]\times\Omega)$, which implies in particular that $\left((\tilde\rho^{1,\tau_n}+\tilde\rho^{2,\tau_n})^{m-1/2}\right)_{n\geq 0}$ converges strongly in $L^2([0,T]\times\Omega)$ to $(\rho^1+\rho^2)^{m-1/2}$. This together with the above weak convergences implies that $\nabla\xi = -\frac{m}{m-1/2}\nabla(\rho^1+\rho^2)^{m-1/2}\in L^2([0,T]\times\Omega;\mathbb{R}^d)$.

Now, by the strong convergence of $\tilde\rho^{i,\tau_n}$ to ρ^i in $L^{2m-1}([0,T]\times\Omega)$ one has that

$$\frac{\tilde\rho^{i,\tau_n}}{(\tilde\rho^{1,\tau_n}+\tilde\rho^{2,\tau_n})^{1/2}}\to\frac{\rho^i}{(\rho^1+\rho^2)^{1/2}}\quad\text{as}\ \ n\to+\infty$$

pointwisely a.e. in $[0,T]\times\Omega$. Moreover, since the densities are non-negative one has $\frac{\tilde\rho^{i,\tau_n}}{(\tilde\rho^{1,\tau_n}+\tilde\rho^{2,\tau_n})^{1/2}}\leq(\tilde\rho^{i,\tau_n})^{\frac12}$ and $(\tilde\rho^{i,\tau_n})^{\frac12}$ converges to $(\rho^i)^{\frac12}$ strongly in $L^{2(2m-1)}([0,T]\times\Omega)$. Hence Lebesgue's dominated convergence theorem implies that

$$\frac{\tilde\rho^{i,\tau_n}}{(\tilde\rho^{1,\tau_n}+\tilde\rho^{2,\tau_n})^{1/2}}\to\frac{\rho^i}{(\rho^1+\rho^2)^{1/2}}\quad\text{as}\ n\to+\infty,$$

strongly in $L^{2(2m-1)}([0,T]\times\Omega)$.

Gluing together the two previous results, one obtains that $-\frac{m}{m-1}\nabla(\tilde\rho^{1,\tau_n}+\tilde\rho^{2,\tau_n})^{m-1}\tilde\rho^{i,\tau_n}$ converges weakly to $-\frac{m}{m-1}\nabla(\rho^1+\rho^2)^{m-1}\rho^i$ in $L^r([0,T]\times\Omega;\mathbb{R}^d)$ as $n\to+\infty$, where $\frac12+\frac{1}{2(2m-1)}+\frac1r=1$, i.e. $r=\frac{2(2m-1)}{2m-2}>1$. So in particular the convergence is weakly-$\star$ in $\mathscr{M}^d([0,T]\times\Omega)$, which together with the strong convergence of the term $\nabla\Phi_i\tilde\rho^{i,\tau_n}$ implies the thesis. $\qquad\square$

Theorem 3.9. *Let $m=+\infty$ and let and let us consider $\tilde\rho^{i,\tau}$, $i=1,2$ the piecewise constant interpolations between the densities $(\rho_k^{i,\tau})_{k=0}^N$ and $\widetilde{\mathbf{E}}^{i,\tau}$, $i=1,2$ the corresponding momentum variables. Let us consider moreover p^τ the piecewise constant interpolations between the pressure variables $(p_k^\tau)_{k=0}^{N-1}$. Let us take any positive sequence $(\tau_n)_{n\geq 0}$ such that $\tau_n\downarrow 0$ as $n\to+\infty$, and let us consider the weak limit p of $(p^{\tau_n})_{n\geq 0}$ in $L^2([0,T];H^1(\Omega))$ and ρ^i the limit of $(\rho^{i,\tau_n})_{n\geq 0}$ in $L^\infty([0,T];(\mathscr{P}^{M_i}(\Omega),W_2))$ (up to passing to a subsequence that we do not relabel). Then we have the following:*

$$p(1-(\rho^1+\rho^2))=0\quad\text{a.e. in }[0,T]\times\Omega.$$

Proof. First notice that by Lemma 3.4 and Proposition 3.7 the weak limits p and ρ^i ($i=1,2$) exist. Furthermore, (1) follows from the previously mentioned results, Lemma A.3 and [36, Lemma 3.5]. $\qquad\square$

Remark 3.4. *In Theorem 3.9 if $\tilde\rho^{1,\tau_n}$ and $\tilde\rho^{2,\tau_n}$ are such that $\mathscr{L}^d(\{\tilde\rho_t^{1,\tau_n}>0\}\cap\{\tilde\rho_t^{2,\tau_n}>0\})=0$ for all $t\in[0,T]$ and for all $n\in\mathbb{N}$, then*

$$\widetilde{\boldsymbol{E}}^{i,\tau_n}\overset{\star}{\rightharpoonup}-\nabla p-\nabla\Phi_i\rho^i,\quad\text{as }n\to+\infty,\text{ in }\mathscr{M}^d([0,T]\times\Omega),\ i=1,2.$$

If moreover $\mathscr{L}^d(\{\rho_t^1 > 0\} \cap \{\rho_t^2 > 0\}) = 0$ *for a.e.* $t \in [0, T]$, *then*

$$\nabla p = \rho^i \nabla p, \text{ a.e. in } [0, T] \times \Omega, \quad i = 1, 2.$$

Proof. To show the remark let us recall the form of the momentum variables, i.e.

$$\widetilde{\mathbf{E}}^{i,\tau_n} = \check{\mathbf{v}}^{i,\tau_n} \tilde{\rho}^{i,\tau_n} = -\nabla p^{\tau_n} \tilde{\rho}^{i,\tau_n} - \nabla \Phi_i \tilde{\rho}^{i,\tau_n} = -\nabla p^{\tau_n} - \nabla \Phi_i \tilde{\rho}^{i,\tau_n},$$

where the last equality holds true since $p^{\tau_n}(1 - (\tilde{\rho}^{1,\tau_n} + \tilde{\rho}^{2,\tau_n})) = 0$ for all $t \in [0, T]$ and a.e. in Ω and by the assumption $\mathscr{L}^d(\{\tilde{\rho}_t^{1,\tau_n} > 0\} \cap \{\tilde{\rho}_t^{2,\tau_n} > 0\}) = 0$ for all $t \in [0, T]$, one has $\{\tilde{\rho}_t^{1,\tau_n} + \tilde{\rho}_t^{2,\tau_n} = 1\} =_{\text{a.e.}}$ $\{\tilde{\rho}_t^{1,\tau_n} = 1\} \cup \{\tilde{\rho}_t^{2,\tau_n} = 1\}$ for all $t \in [0, T]$ and the two sets are disjoint a.e. in Ω. By the weak convergences of $(p^{\tau_n})_{n \geq 0}$ to p and $(\tilde{\rho}^{i,\tau_n})_{n \geq 0}$ to ρ^i $(i = 1, 2)$ one can easily conclude that

$$\widetilde{\mathbf{E}}^{i,\tau_n} \overset{\star}{\rightharpoonup} -\nabla p - \nabla \Phi_i \rho^i, \text{ as } n \to +\infty, \text{ in } \mathscr{M}^d([0, T] \times \Omega), \ i = 1, 2.$$

Also, $\nabla p = \rho^i \nabla p$, $i = 1, 2$ a.e. in $[0, T] \times \Omega$ follows easily from (1) and the assumption $\mathscr{L}^d(\{\tilde{\rho}_t^1 > 0\} \cap \{\tilde{\rho}_t^2 > 0\}) = 0$ for all $t \in [0, T]$, as desired. $\square$

3.4. **Segregation of the densities.** As mentioned in the introduction, it seems natural to look for initial configurations of the system (PME$_m$) where there is no mixing of the densities, to strengthen our convergence results in the continuum limit. We shall describe such initial configurations in one space dimension in the next section. Here we describe some properties of the time-discrete solutions (obtained by the JKO scheme) which hold for all dimensions. In particular, we show that when $\Phi_2 = \Phi_1 + C$ (for some $C \in \mathbb{R}$), then the densities stay segregated if initially they were so. We derive also some properties of the mixed region, when the two initial densities are mixed in a special way. Still, these statements are only true for time-discrete solutions, and we cannot rule out the possibility that the limiting densities end up mixed, for instance, due to "fingering" phenomena (see also the numerical observations in [28] which displays fingering phenomena when a system, similar to (2.4), has unstable combination of diffusion constants and source terms). It seems that additional geometric property is required to preserve the segregation property in the continuum limit.

Proposition 3.10. *Let* $m \in (1, +\infty]$. *Let us assume moreover that* Φ_1 *and* Φ_2 *are such that* $\Phi_2 = \Phi_1 + C$ *on* Ω *for some* $C \in \mathbb{R}$, *with the hypothesis* (H$_\Phi$) *fulfilled. Let* $(\rho_0^1, \rho_0^2) \in \mathscr{P}^{M_1}(\Omega) \times \mathscr{P}^{M_2}(\Omega)$ *satisfy* (H$_\rho^m$) *and let* (ρ^1, ρ^2) *be the minimizers in* (MM$_m$) *constructed with the help of* (ρ_0^1, ρ_0^2). *Then the following statements hold true.*

(1) *If* $\mathscr{L}^d(\{\rho_0^1 > 0\} \cap \{\rho_0^2 > 0\}) = 0$, *then* $\mathscr{L}^d(\{\rho^1 > 0\} \cap \{\rho^2 > 0\}) = 0$.

(2) *Let us define the Borel measurable sets* $A := \{\rho_0^1 > 0\} \cap \{\rho_0^2 > 0\}$ *and* $B := \{\rho^1 > 0\} \cap \{\rho^2 > 0\}$. *Let us suppose that* $\mathscr{L}^d(A) > 0$ *and* $\mathscr{L}^d(B) > 0$. *If there exists* $r > 0$ *such that* $r\rho_0^1 \leq \rho_0^2$ *a.e. in* A, *then* $r\rho^1 \leq \rho^2$ *a.e. in* B.

Proof. Let us use the notation $\nabla \Phi := \nabla \Phi_1 = \nabla \Phi_2$. Using (2.11), the optimal transport maps T^i $(i = 1, 2)$ in the transport of ρ^i onto ρ_0^i (see Lemma 2.2-2.3) can be written (ρ^i − a.e.) as

$$T^i = \begin{cases} \text{id} + \tau\left(\frac{m}{m-1}\nabla(\rho^1 + \rho^2)^{m-1} + \nabla\Phi\right), & \text{if } m \in (1, +\infty), \\ \text{id} + \tau(\nabla p + \nabla\Phi), & \text{if } m = \infty. \end{cases}$$

In particular, observe that $T^1 = T^2$ a.e. in $\{\rho^1 > 0\} \cap \{\rho^2 > 0\}$.

We show (1). Suppose that the Borel measurable set $B := \{\rho^1 > 0\} \cap \{\rho^2 > 0\}$ has positive Lebesgue measure. For any $x_0 \in B$ such that x_0 is a Lebesgue point of ρ^1, ρ^2 and T^1, T^2 and $T^1(x_0) = T^2(x_0)$ is a Lebesgue point for both ρ_0^1 and ρ_0^2, one has (since $T^1(x_0) = T^2(x_0)$) that $T^1(x_0) = T^2(x_0) \in \{\rho_0^1 > 0\} \cap \{\rho_0^2 > 0\}$. In particular the positive mass of each ρ^i $(i = 1, 2)$ on B is transported onto $\{\rho_0^1 > 0\} \cap \{\rho_0^2 > 0\}$. On the other hand, since both ρ_0^1 and ρ_0^2 are absolutely continuous w.r.t. $\mathscr{L}^d$, this mass cannot be supported on an $\mathscr{L}^d$-null set, which is a contradiction to the assumption $\mathscr{L}^d(\{\rho_0^1 > 0\} \cap \{\rho_0^2 > 0\}) = 0$.

We show (2). First observe that one can write two Jacobian equation in a weak sense, i.e.

$$(3.18) \qquad \qquad \det(DT^i) = \frac{\rho^i}{\rho_0^i \circ T^i}, \quad \rho^i - \text{a.e.}.$$

Since the measures ρ^i, $i = 1, 2$ are absolutely continuous w.r.t. the Lebesgue measure, the maps T^i are differentiable ρ^i − a.e. and the previous equation holds true pointwisely ρ^i − a.e. (see for instance

[17, Theorem 3.1]). Let us choose $x_0 \in B$ such that it is a Lebesgue point of both ρ^1 and ρ^2 and it is a point of differentiability of both T^1 and T^2 (in particular $\rho^1(x_0) > 0$ and $\rho^2(x_0) > 0$). Since the optimal transport maps coincide on the common support of ρ^1 and ρ^2, one may assume that $T(x_0) := T^1(x_0) = T^2(x_0)$ is a Lebesgue point of both ρ_0^1 and ρ_0^2. The Jacobian equation (3.18) yields that $\rho^2(x_0)/\rho^1(x_0) = \rho_0^2(T(x_0))/\rho_0^1(T(x_0)) \geq r$, which concludes the proof.   □

4. Segregated weak solutions in 1D

In this section we study the local segregation property of the supports for the time-discrete solutions. As a consequence we show the existence of *segregated weak solutions* of the systems (PME$_m$) and (PME$_\infty$) in one spacial dimension.

4.1. Separation of the supports and ordering property in one space dimension.

Framework Hyp-1D. *We set the following geometric framework (see also Figure 1 below for illustration).*
(1) *$d = 1$, Ω a bounded open interval, the potentials Φ_i, $i = 1, 2$ are semi-convex and $C^1(\Omega)$;*
(2) *The drifts are 'ordered', in the sense that $\partial_x\Phi_2(x) \geq \partial_x\Phi_1(x)$ for all $x \in \Omega$. This means in particular that $\Phi_2 - \Phi_1$ is increasing;*
(3) *ρ_0^1 and ρ_0^2 are two densities such that for a.e. $x \in \{\rho_0^1 > 0\}$ and $y \in \{\rho_0^2 > 0\}$ one has that $y < x$. We refer to this last property as "ordering of the supports" of the initial densities. This implies in particular that $\mathscr{L}^1(\{\rho_0^1 > 0\} \cap \{\rho_0^2 > 0\}) = 0$.*

Let us point out that the assumptions from Hyp-1D immediately imply with reasoning parallel to Proposition 3.10 that $\mathscr{L}^1(\{\rho^1 > 0\} \cap \{\rho^2 > 0\}) = 0$ for one-step minimizers (ρ^1, ρ^2) given by (MM$_m$). To see this, suppose $\{\rho^1 > 0\} \cap \{\rho^2 > 0\} =_{a.e} B$ for some Borel measurable set B such that $\mathscr{L}^1(B) > 0$. As in the proof of Proposition 3.10, The optimal transport map T^i ($i = 1, 2$) in the transport of ρ^i onto ρ_0^i is given by

$$T^i = \begin{cases} \mathrm{id} + \tau\left(\frac{m}{m-1}\partial_x(\rho^1 + \rho^2)^{m-1} + \partial_x\Phi_i\right), & \text{if } m \in (1, +\infty), \\ \mathrm{id} + \tau(\partial_x p + \partial_x\Phi_i), & \text{if } m = \infty. \end{cases}$$

The above formula and the assumption $\partial_x\Phi_2 - \partial_x\Phi_1 \geq 0$ in Ω yield that $T^2(x) \geq T^1(x)$ a.e. in B, which contradicts the ordering property of the initial data.

Still, this separation property is not enough to iterate over time steps unless the ordering property of the initial configuration is preserved for (ρ^1, ρ^2). This is what we prove next.

Proposition 4.1. *Let $m \in (1, +\infty]$ and suppose the assumptions in (Hyp-1D) and the hypotheses (H$_\Phi$)-(H$_\rho^m$) are in place. Let us denote by (ρ^1, ρ^2) the one-step time discrete solutions given by (MM$_m$) for $k = 0$. Then the ordering property from (Hyp-1D) holds true for $\{\rho^1 > 0\}$ and $\{\rho^2 > 0\}$.*
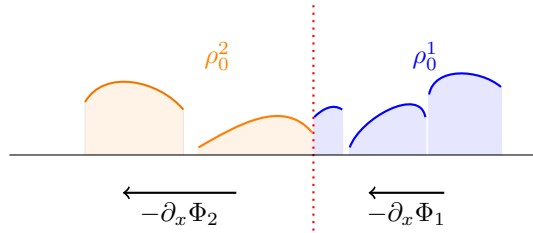

FIGURE 1. Ordering of the supports of the initial data

Proof. Suppose the contrary, i.e. there exist $B_1 \subseteq \{\rho^1 > 0\}$ and $B^2 \subseteq \{\rho^2 > 0\}$ with $\mathscr{L}^1(B^1) > 0$ and $\mathscr{L}^1(B_2) > 0$ such that for a.e. $x \in B^1$ and $y \in B^2$ $x < y$ (see Figure 2 for illustration).

   Claim: there exist $E^i \subseteq B^i$, $i = 1, 2$ Borel measurable sets, $\theta > 0$ and $\delta > 0$ such that $\mathscr{L}^1(E^1) = \mathscr{L}^1(E^2) > 0$, $E^2 = E^1 + \theta$ and $\rho^i \geq \delta$ a.e. on E^i, $i = 1, 2$ (see Figure 2 for illustration).

Proof of the claim. Let us take $x_0 \in B^1, y_0 \in B^2$ Lebesgue points. This means in particular that $\rho^1(x_0) > 0$, $\rho^2(y_0) > 0$ and

$$(4.1) \qquad \lim_{r \downarrow 0} \fint_{B_r(x_0)} \left| \rho^1(x) - \rho^1(x_0) \right| \, \mathrm{d}x = 0, \quad \lim_{r \downarrow 0} \fint_{B_r(y_0)} \left| \rho^2(x) - \rho^2(y_0) \right| \, \mathrm{d}x = 0.$$

Now let us take $r > 0$ small (we fix it later) and let $\delta := \min \left\{ \rho^1(x_0)/2, \rho^2(y_0)/2 \right\}$. Let us consider moreover the measurable sets $\tilde{E}^1 \subseteq B^1 \cap B_r(x_0)$ and $\tilde{E}^2 \subseteq B^2 \cap B_r(y_0)$ defined as $\tilde{E}^i := \left\{ \rho^i \geq \delta \right\}$, $i = 1, 2$. By construction, for $r > 0$ small enough one has that $\mathscr{L}^1(B_r(x_0) \setminus \tilde{E}^1)/\mathscr{L}^1(B_r(x_0)) \leq 1/3$ and $\mathscr{L}^1(B_r(y_0) \setminus \tilde{E}^2)/\mathscr{L}^1(B_r(y_0)) \leq 1/3$. Indeed, one has

$$\fint_{B_r(x_0)} \left| \rho^1(x) - \rho^1(x_0) \right| \, \mathrm{d}x \geq \frac{1}{\mathscr{L}^1(B_r(x_0))} \int_{B_r(x_0) \setminus \tilde{E}^1} \left| \rho^1(x) - \rho^1(x_0) \right| \, \mathrm{d}x$$
$$\geq \frac{\rho^1(x_0)}{2} \frac{\mathscr{L}^1(B_r(x_0) \setminus \tilde{E}^1)}{\mathscr{L}^1(B_r(x_0))},$$

and by (4.1) the l.h.s. tends to 0 as $r \downarrow 0$, so for $r > 0$ small enough $\mathscr{L}^1(B_r(x_0) \setminus \tilde{E}^1)/\mathscr{L}^1(B_r(x_0)) \leq 1/3$. Similarly for ρ^2 and $\tilde{E}^2$. Fix such an $r > 0$.

Furthermore, set $\theta := y_0 - x_0$ and define $E^1 := \tilde{E}^1 \cap (\tilde{E}^2 - \theta)$ and $E^2 := E^1 + \theta$. Thus,

$$\frac{\mathscr{L}^1(E^2)}{\mathscr{L}^1(B_r(y_0))} = \frac{\mathscr{L}^1(E^1)}{\mathscr{L}^1(B_r(x_0))} \geq 1 - \frac{\mathscr{L}^1(B_r(x_0) \setminus \tilde{E}^1)}{\mathscr{L}^1(B_r(x_0))} - \frac{\mathscr{L}^1(B_r(y_0) \setminus \tilde{E}^2)}{\mathscr{L}^1(B_r(y_0))} = \frac{1}{3}.$$

This finishes the proof of the claim, since $r > 0$ is a fixed small number.

Now we construct a new competitor $(\tilde{\rho}^1, \tilde{\rho}^2)$ in (MM_m) which has less energy (we refer to Figure 2 for the illustration) than (ρ^1, ρ^2), yielding the contradiction. Define $\tilde{\rho}^1$ and $\tilde{\rho}^2$ as

$$\tilde{\rho}^1 = \begin{cases} \rho^1, & \text{in } \Omega \setminus (E^1 \cup E^2), \\ \rho^1 - \delta, & \text{in } E^1, \\ \delta, & \text{in } E^2, \end{cases} \qquad \text{and} \quad \tilde{\rho}^2 = \begin{cases} \rho^2, & \text{in } \Omega \setminus (E^1 \cup E^2), \\ \delta, & \text{in } E^1, \\ \rho^2 - \delta, & \text{in } E^2. \end{cases}$$

We construct corresponding transport maps (not necessarily optimal ones), $\tilde{T}^1$ between $\tilde{\rho}^1$ and ρ_0^1 and $\tilde{T}^2$ between $\tilde{\rho}^2$ and ρ_0^2 as

$$\tilde{T}^1 = \begin{cases} T^1, & \text{in } \Omega \setminus E^2, \\ T^1(\cdot - \theta), & \text{in } E^2, \end{cases} \qquad \text{and} \quad \tilde{T}^2 = \begin{cases} T^2, & \text{in } \Omega \setminus E^1, \\ T^2(\cdot + \theta), & \text{in } E^1. \end{cases}$$

By construction $\tilde{T}^i_\# \tilde{\rho}^i = \rho_0^i$, $i = 1, 2$. Let us use the notation $E_0^1 := T^1(E^1)$ and $E_0^2 := T^2(E^2)$, these are Borel measurable sets and subsets of $\{\rho_0^1 > 0\}$ and $\{\rho_0^2 > 0\}$ respectively.
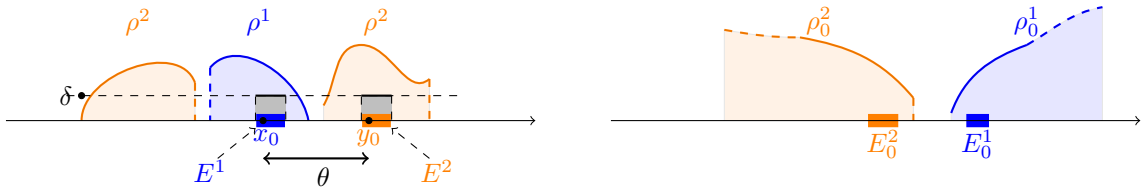

FIGURE 2. Ordering property for $\{\rho_0^1 > 0\}$ and $\{\rho_0^2 > 0\}$ (on the right). This is is violated by $\{\rho^1 > 0\}$ and $\{\rho^2 > 0\}$ (on the left)

Notice that by construction $\tilde{\rho}^1 + \tilde{\rho}^2 = \rho^1 + \rho^2$ in Ω, hence

$$(4.2) \qquad \mathcal{F}_m(\tilde{\rho}^1, \tilde{\rho}^2) = \mathcal{F}_m(\rho^1, \rho^2).$$

Now let us see how the other two energy terms in (MM_m) change by considering $(\tilde\rho^1, \tilde\rho^2)$ as competitors. Let us use the notation $h(x) := \partial_x \Phi_2(x) - \partial_x \Phi_1(x)$. First,

$$E_{\mathcal{G}} := \mathcal{G}(\tilde\rho^1, \tilde\rho^2) - \mathcal{G}(\rho^1, \rho^2) = \int_\Omega \Phi_1 \tilde\rho^1 \, \mathrm{d}x + \int_\Omega \Phi_2 \tilde\rho^2 \, \mathrm{d}x - \int_\Omega \Phi_1 \rho^1 \, \mathrm{d}x - \int_\Omega \Phi_2 \rho^2 \, \mathrm{d}x$$

$$= \delta \left(\int_{E^2} \Phi_1(x) \, \mathrm{d}x - \int_{E^1} \Phi_1(x) \, \mathrm{d}x \right) + \delta \left(\int_{E^1} \Phi_2(x) \, \mathrm{d}x - \int_{E^2} \Phi_2(x) \, \mathrm{d}x \right)$$

$$= \int_{E^1} \delta \left[(\Phi_1(x+\theta) - \Phi_2(x+\theta)) - (\Phi_1(x) - \Phi_2(x)) \right] \mathrm{d}x$$

$$= \int_{E^1} \delta\theta \left[\partial_x \Phi_1(\xi_{x,\theta}) - \partial_x \Phi_2(\xi_{x,\theta}) \right] \mathrm{d}x$$

where in the last equality we used the mean value theorem and $\xi_{x,\theta}$ is some point in $(x, x+\theta)$. We compute now the change in the W_2 terms. Recall the structure of the transport maps $\tilde{T}^i$, $i = 1, 2$ and mind that they might be not optimal. Thus one has

$$E_{W_2} := \frac{1}{2\tau} W_2^2(\tilde\rho^1, \rho_0^1) + \frac{1}{2\tau} W_2^2(\tilde\rho^2, \rho_0^2) - \frac{1}{2\tau} W_2^2(\rho^1, \rho_0^1) - \frac{1}{2\tau} W_2^2(\rho^2, \rho_0^2)$$

$$\leq \frac{1}{2\tau} \int_{E^2} |x - T^1(x-\theta)|^2 \delta \, \mathrm{d}x + \frac{1}{2\tau} \int_{E^1} |x - T^2(x+\theta)|^2 \delta \, \mathrm{d}x$$

$$- \frac{1}{2\tau} \int_{E^1} |x - T^1(x)|^2 \delta \, \mathrm{d}x - \frac{1}{2\tau} \int_{E^2} |x - T^2(x)|^2 \delta \, \mathrm{d}x$$

$$= \frac{1}{2\tau} \int_{E^1} \left(|x + \theta - T^1(x)|^2 - |x - T^1(x)|^2 \right) \delta \, \mathrm{d}x$$

$$+ \frac{1}{2\tau} \int_{E^2} \left(|x - \theta - T^2(x)|^2 - |x - T^2(x)|^2 \right) \delta \, \mathrm{d}x$$

$$= \frac{\delta\theta}{\tau} \int_{E^1} \left(T^2(x+\theta) - T^1(x) \right) \mathrm{d}x$$

where $\eta^i = \delta \cdot \mathscr{L}^1 \llcorner E^i$ and $\eta_0^i = T_\#^i \eta^i$, $i = 1, 2$.

Now, it is easy to see that $E_{\mathcal{G}} + E_{W_2} < 0$. Indeed, by the assumptions (2) from (Hyp-1D) one has that $\partial_x \Phi_1 - \partial_x \Phi_2$ nonpositive, thus

$$(4.3) \qquad E_{\mathcal{G}} + E_{W_2} \leq \delta\theta \int_{E^1} \left\{ [\partial_x \Phi_1(\xi_{x,\theta}) - \partial_x \Phi_2(\xi_{x,\theta})] + \frac{1}{\tau}[T^2(x+\theta) - T^1(x)] \right\} \mathrm{d}x$$

is negative since by the assumption (3) from (Hyp-1D) $T^2(x+\theta) - T^1(x) < 0$.

Thus one concludes that $E_{\mathcal{G}} + E_{W_2} < 0$, which together with (4.2) imply that $(\tilde\rho^1, \tilde\rho^2)$ is a better competitor than (ρ^1, ρ^2). This is clearly a contradiction to the uniqueness of the minimizer in (MM_m). Thus the ordering property for $\{\rho^1 > 0\}$ and $\{\rho^2 > 0\}$ follows. $\qquad\square$

4.2. Discussion on possibly mixed initial data.
Extending the above proposition to more general cases seems to be challenging, due to possible presence of the *mixing zone* $\{\rho_1 > 0\} \cap \{\rho_2 > 0\}$. The main issue, for instance to localize our argument, would be to ensure the finite propagation of mixing zone. The only available result in this direction arises in the case of the stiff pressure limit, $m = \infty$, $\Phi_i = c_i x$, and with full saturation, that is when we have the constraint $\rho_1 + \rho_2 = 1$. In this case Otto ([38]) showed in one dimensional setting that there is a unique description of the mixing zone that propagates with finite speed generated by the entropy solution of a conservation law. While we are not sure whether the same uniqueness results hold for our undersaturated case, we believe that the mixing zone should travel with finite speed at least in one dimension.

4.3. Existence of a solution for (PME_m) supposing (Hyp-1D).

Theorem 4.2. *Let us suppose that $m \in (1, +\infty]$ and the setting of (Hyp-1D) takes place. Let us consider (ρ^1, ρ^2) to be any subsequential limit (uniformly in time w.r.t. W_2) of the piecewise constant interpolation curves $(\tilde\rho^{1,\tau_n}, \tilde\rho^{2,\tau_n})$ when $\tau_n \downarrow 0$, with the initial densities (ρ_0^1, ρ_0^2). Then (ρ^1, ρ^2) satisfies*

$$\mathscr{L}^1 \left(\{\rho_t^1 > 0\} \cap \{\rho_t^2 > 0\} \right) = 0, \ \forall t \in [0, T]$$

and the sets $\{\rho_t^1 > 0\}$ and $\{\rho_t^2 > 0\}$ are ordered in the sense of (Hyp-1D) for all $t \in [0, T]$.

Proof. First, let us recall that the *ordering* of $\{\rho_0^1 > 0\}$ and $\{\rho_0^2 > 0\}$ in (Hyp-1D) is such that $\{\rho_0^2 > 0\}$ is to the left of $\{\rho_0^1 > 0\}$.

Second, let us underline that by Proposition 3.7 (1) ρ^i is obtained as the uniform limit in time w.r.t. W_2 (as $\tau \downarrow 0$) of the piecewise constant interpolation curves $\tilde{\rho}^{i,\tau}$ ($i = 1, 2$). For a fixed time step $\tau > 0$, considering the above mentioned interpolations, we introduce the following functions $I^{1,\tau}, I^{2,\tau} : [0, T] \to \overline{\Omega}$ defined as

$$I^{1,\tau}(t) := \inf \left\{ x : x \in \text{Leb}\left(\{\tilde{\rho}_t^{1,\tau} > 0\}\right) \right\} \quad \text{and} \quad I^{2,\tau}(t) := \sup \left\{ x : x \in \text{Leb}\left(\{\tilde{\rho}_t^{2,\tau} > 0\}\right) \right\}.$$

These functions are well-defined, since Ω is bounded and in particular Proposition 4.1 implies that $I^{2,\tau}(t) \leq I^{1,\tau}(t)$ for all $t \in [0, T]$ and for any $\tau > 0$. Also, by the boundedness of Ω, these functions are uniformly bounded in t and τ.

Let us take a sequence $(\tau_n)_{n \geq 0}$, s.t. $\tau_n \downarrow 0$ as $n \to +\infty$ and $\sup_{t \in [0,T]} W_2(\tilde{\rho}_t^{i,\tau}, \rho_t^i) \to 0$ as $n \to +\infty$, ($i = 1, 2$). $\left(I^{i,\tau_n}(t)\right)_{n \geq 0}$ is a bounded sequence for each $t \in [0, T]$, so up to passing to a subsequence (that we do not relabel), it has a poitwise limit as $n \to +\infty$ that we denote by $I^i(t)$ for $t \in [0, T]$ and $i = 1, 2$. Now we show the following.

Claim:

(1) $\rho_t^2(y) = 0$ for a.e. $y > I^2(t)$ and (2) $\rho_t^1(x) = 0$ for a.e. $x < I^1(t)$.

Proof of the claim. Let us suppose that the claim is false, i.e. the first statement fails to be true (the proof of (2) is parallel). Then there exits $r > 0$ and $\delta > 0$ small such that

$$\int_{I^2(t)+r}^{I^2(t)+2r} \rho_t^2(x)\,\mathrm{d}x > \delta > 0.$$

But, for $n \in \mathbb{N}$ large enough such that $\left|I^{2,\tau_n}(t) - I^2(t)\right| < r/2$ one has that

$$W_2^2(\tilde{\rho}_t^{2,\tau_n}, \rho_t^2) \geq (r/2)^2 \int_{I^2(t)+r}^{I^2(t)+2r} \rho_t^2(x)\,\mathrm{d}x = (r/2)^2 \delta,$$

which yields a contradiction to the fact that $W_2(\tilde{\rho}_t^{2,\tau}, \rho_t^2) \to 0$ as $n \to +\infty$. A similar argument can be performed to show (2), thus the claim follows.

Now, since $I^{2,\tau_n}(t) \leq I^{1,\tau_n}(t)$ for all $n \in \mathbb{N}$ and $t \in [0, T]$, after passing to subsequences if necessary, one has that $I^2(t) \leq I^1(t)$ for any limit points $I^1(t), I^2(t)$ and for all $t \in [0, T]$. This together with the Claim imply that $\mathscr{L}^1\left(\{\rho_t^1 > 0\} \cap \{\rho_t^2 > 0\}\right) = 0$, $\forall t \in [0, T]$ and that the sets $\{\rho_t^1 > 0\}$ and $\{\rho_t^2 > 0\}$ are ordered in the sense of (Hyp-1D) for all $t \in [0, T]$. The result follows. $\qquad \square$

Remark 4.1. *When $m \in (1, +\infty)$, the above result allows to obtain the strong convergence result of the density sequences $(\tilde{\rho}^{i,\tau_n})_{n \geq 0}$, $i = 1, 2$ separately. When $m = +\infty$, together with Proposition 4.1 this result is crucial to fulfill the hypotheses in Remark 3.4, which will lead to the precise weak form of the (PME_∞) system.*

Theorem 4.3. *Let us suppose that $m \in (1, +\infty)$ and the setting of (Hyp-1D) takes place. Consider the piecewise constant interpolations $\tilde{\rho}^{i,\tau_n}$ ($i = 1, 2$) for some $(\tau_n)_{n \geq 0}$ such that $\tau_n \downarrow 0$ as $n \to +\infty$. Then up to passing to a subsequence with $(\tau_n)_{n \geq 0}$, $\left(\tilde{\rho}^{i,\tau_n}\right)_{n \geq 0}$ ($i = 1, 2$) converges strongly in $L^{2m-1}([0, T] \times \Omega)$, in particular pointwise a.e. in $[0, T] \times \Omega$.*

Proof. Let us show first that $(\tilde{\rho}^{i,\tau_n})_{n \geq 0}$ (up to passing to a subsequence) converges strongly to ρ^i ($i = 1, 2$) in $L^1([0, T] \times \Omega)$. We pass to subsequences if necessary (that we do not relabel) to ensure that $(\tilde{\rho}^{1,\tau_n} + \tilde{\rho}^{2,\tau_n})_{n \geq 0}$ converges strongly to $\rho^1 + \rho^2$ in $L^{2m-1}([0, T] \times \Omega)$ and $(\tilde{\rho}^{i,\tau_n})_{n \geq 0}$ converges to ρ^i

$(i = 1, 2)$ weakly in $L^{2m-1}([0,T] \times \Omega)$ as $n \to +\infty$. We compute

$$
\begin{aligned}
\|\tilde{\rho}^{i,\tau_n} - \rho^i\|_{L^1([0,T]\times\Omega)} &= \int_{[0,T]\times\Omega} |\tilde{\rho}^{i,\tau_n} - \rho^i| \, \mathrm{d}t \otimes \mathrm{d}x \\
&= \int_{\{\rho^i>0\}} |\tilde{\rho}^{i,\tau_n} - \rho^i| \, \mathrm{d}t \otimes \mathrm{d}x + \int_{([0,T]\times\Omega)\setminus\{\rho^i>0\}} \tilde{\rho}^{i,\tau_n} \, \mathrm{d}t \otimes \mathrm{d}x \\
&= \int_{\{\rho^i>0\}} |\tilde{\rho}^{1,\tau_n} + \tilde{\rho}^{2,\tau_n} - (\rho^1 + \rho^2)| \, \mathrm{d}t \otimes \mathrm{d}x + \int_{([0,T]\times\Omega)\setminus\{\rho^i>0\}} \tilde{\rho}^{i,\tau_n} \, \mathrm{d}t \otimes \mathrm{d}x \\
&\leq \int_{[0,T]\times\Omega} |\tilde{\rho}^{1,\tau_n} + \tilde{\rho}^{2,\tau_n} - (\rho^1 + \rho^2)| \, \mathrm{d}t \otimes \mathrm{d}x \\
&\quad + \int_0^T \int_{([0,T]\times\Omega)\setminus\{\rho^i>0\}} \tilde{\rho}^{i,\tau_n} \, \mathrm{d}t \otimes \mathrm{d}x \\
&\to 0, \quad \text{as } n \to +\infty,
\end{aligned}
$$

where in the third equality we used the facts (see Theorem 4.2) that $\rho^i = 0$ a.e. in $\{\rho^{i+1} > 0\}$ and $\tilde{\rho}^{i,\tau_n} = 0$ a.e. in $\{\rho^{i+1} > 0\}$, with the convention $i + 1 = 1$, when $i = 2$. Moreover, both terms in the last sum converge to 0. Indeed, the convergence of the first term is a consequence of the strong convergence of $(\tilde{\rho}^{1,\tau_n} + \tilde{\rho}^{2,\tau_n})_{n\geq 0}$ to $\rho^1 + \rho^2$ in $L^{2m-1}([0,T] \times \Omega)$ as $n \to +\infty$. The convergence to 0 of the last term is a consequence of the weak convergence of $\tilde{\rho}^{i,\tau_n}$ to ρ^i in $L^{2m-1}([0,T] \times \Omega)$.

This together with Theorem 3.5 and Proposition 3.6 imply that (up to passing to a subsequence) $(\rho^{i,\tau_n})_{n\geq 0}$ converges strongly in $L^{2m-1}([0,T] \times \Omega)$. $\qquad\square$

We state now the results on the existence of weak solutions of the PDE systems (PME_m) and (PME_∞).

Theorem 4.4. *Let us assume that $m \in (1, +\infty)$, the hypotheses (H_ρ^m) and (H_Φ) are fulfilled and the setting in $(\mathrm{Hyp\text{-}1D})$ takes place. Then the system (PME_m) has a weak solution (ρ^1, ρ^2) in the sense of (Weak) such that $\rho^i \in L^{2m-1}([0,T] \times \Omega) \cap AC^2([0,T]; (\mathscr{P}^{M_i}(\Omega), W_2)), i = 1, 2$ and $(\rho^1 + \rho^2)^{m-1/2} \in L^2([0,T]; H^1(\Omega))$. In addition, $\rho^i \in L^q([0,T] \times \Omega)$ for all $1 \leq q \leq m$ and $\boldsymbol{E}^i := -\frac{m}{m-1}\partial_x(\rho^1 + \rho^2)^{m-1}\rho^i - \partial_x\Phi_i\rho^i$ belongs to $L^r([0,T] \times \Omega; \mathbb{R}^d)$ for some $1 \leq r < 2$ with uniform bounds in m. Lastly, if $m \to +\infty$, q can be arbitrary large and r can be chosen arbitrary close to 2.*

Proof. By Theorem 3.8 one has that the limit densities ρ^1 and ρ^2 belong to

$$L^{2m-1}([0,T] \times \Omega) \cap AC^2([0,T]; (\mathscr{P}^{M_i}(\Omega), W_2)).$$

The same theorem establishes the convergence of $(\tilde{\rho}^{i,\tau}, \widetilde{\mathbf{E}}^{i,\tau})$ and the precise form of the limit. By the fact that $(\tilde{\rho}^{i,\tau}, \widetilde{\mathbf{E}}^{i,\tau})$ and $(\rho^{i,\tau}, \mathbf{E}^{i,\tau})$ converge weakly as measures to the same limit $(\rho^i, \mathbf{E}^i)$ and by the fact that this latter pair solves the continuity equation (3.15) in the weak sense (3.3), so does the precise limit of $(\tilde{\rho}^{i,\tau}, \widetilde{\mathbf{E}}^{i,\tau})$ developed in Remark 3.3(2) (notice that by Theorem 4.3 the assumptions in Remark 3.3(2) are fulfilled). This means in particular that the limit equation reads (for $i = 1, 2$) as

$$\partial_t\rho^i - \partial_x\left(\frac{m}{m-1}\partial_x(\rho^1 + \rho^2)^{m-1}\rho^i + \partial_x\Phi_i\rho^i\right) = 0,$$

that has to be understood in the weak sense (Weak) with no-flux boundary condition.

Finally, let us obtain the uniform (w.r.t m) bounds on ρ^i and $\mathbf{E}^i$. First, by Theorem 3.5 (3.11) one has that the limit curves are bounded in $L^p([0,T]; L^m(\Omega))$ for all $p \geq 1$ with uniform bounds. Take $p = m$ and any $1 \leq q \leq m$. Then Hölder's inequality yields

$$(4.4) \qquad \|\rho^i\|_{L^q([0,T]\times\Omega)} \leq (T\mathscr{L}^1(\Omega))^{\frac{m-q}{qm}}\|\rho^i\|_{L^m([0,T]\times\Omega)} \leq T^{\frac{1}{q}}\mathscr{L}^1(\Omega)^{\frac{m-q}{qm}}C_1(m).$$

Second, let us write

$$\mathbf{E}^i = \frac{m}{m-1/2}\partial_x(\rho^1 + \rho^2)^{m-1/2}\frac{\rho^i}{(\rho^1 + \rho^2)^{1/2}} + \partial_x\Phi_i\rho^i.$$

Notice that by (3.10) (Theorem 3.5) the L^2 bound for $\partial_x(\rho^1 + \rho^2)^{m-1/2}$ remains the same after passing to the limit with the time step τ. Also, by the previous bound on ρ^i, $\frac{\rho^i}{(\rho^1+\rho^2)^{1/2}}$ is bounded uniformly in

$L^{2q}([0,T] \times \Omega)$ (since $\frac{\rho^i}{(\rho^1+\rho^2)^{1/2}} \le (\rho^i)^{1/2}$ a.e.). These observations, together with the fact that $\partial_x \Phi_i$ is uniformly bounded let us conclude by Hölder's inequality that

$$(4.5) \qquad \|\mathbf{E}^i\|_{L^r} \le \frac{m}{m-1/2} \|\partial_x(\rho^1+\rho^2)^{m-1/2}\|_{L^2} \|\rho^i\|_{L^{r/(2-r)}}^{1/2} + \|\partial_x \Phi_i\|_{L^\infty} \|\rho^i\|_{L^r},$$

provided $1 \le r < 2$ and $\max\left\{\frac{r}{2-r}, r\right\} \le q$.

Thus the thesis of the theorem follows. $\qquad\qquad\square$

Lemma 4.5. *Let* $m \in (1,+\infty)$ *and let us consider* (ρ^1, ρ^2) *the solution of* (PME$_m$) *supposing* (Hyp-1D) *with given initial data* (ρ_0^1, ρ_0^2). *We assume – similarly to the hypotheses* (H$_\rho^m$) *in the* $m = +\infty$ *case – that the measure of* Ω *is large enough, i.e.*

$$(4.6) \qquad \mathscr{L}^1(\Omega) > (M_1 + M_2),$$

where M_i *denotes the total mass of* ρ_0^i. *Then – uniformly in* m *– we have the following regularity estimates*

(1) $(\rho^1+\rho^2)^m \in L^1([0,T]; C^{0,\alpha}(\Omega))$ *for some* $0 < \alpha < 1/2$ *and in particular it is uniformly bounded in* $L^r([0,T] \times \Omega)$ *for some* $1 < r < 2$;
(2) $(\rho^1+\rho^2)^{m-1/2} \in L^1([0,T]; C^{0,1/2}(\Omega))$ *and in particular it is uniformly bounded in* $L^2([0,T] \times \Omega)$.

Proof. We show (1). Using the notations from Theorem 4.4, one has that

$$\mathbf{E}^1 + \mathbf{E}^2 = -\partial_x(\rho^1+\rho^2)^m - \partial_x\Phi_1\rho^1 - \partial_x\Phi_2\rho^2.$$

By the estimations from Theorem 4.4 we know that the quantities $\mathbf{E}^1 + \mathbf{E}^2$ and $\partial_x\Phi_1\rho^1 + \partial_x\Phi_2\rho^2$ are bounded uniformly in $L^r([0,T] \times \Omega; \mathbb{R}^d)$ for some $1 < r < 2$.

This implies first that $\partial_x(\rho^1+\rho^2)^m$ is uniformly bounded in $L^r([0,T] \times \Omega; \mathbb{R}^d)$. Furthermore, the Poincaré-Wirtinger inequality yields that

$$(4.7) \qquad \left\| (\rho^1+\rho^2)^m - \frac{1}{T\mathscr{L}^1(\Omega)} \int_0^T \int_\Omega (\rho^1+\rho^2)^m \, \mathrm{d}x\mathrm{d}t \right\|_{L^r([0,T]\times\Omega)} \le \left\| \partial_x(\rho^1+\rho^2)^m \right\|_{L^r([0,T]\times\Omega)}.$$

So the l.h.s. is uniformly bounded. Let us show that the average of $(\rho^1+\rho^2)^m$ is uniformly bounded.

Let us fix $0 < \varepsilon < 1 - (M_1 + M_2)/\mathscr{L}^1(\Omega)$.

Claim 1. For every $t \in [0,T]$ *there exists* $r > 0$ *and* $x_0 \in \Omega$ *such that* $\rho_t^1 + \rho_t^2 \le 1 - \varepsilon$ *a.e. in* $B_r(x_0)$.

Fix $t \in [0,T]$. Let us suppose that the claim is not true. Then in every ball $B_r(x_0)$, $\rho_t^1 + \rho_t^2 > 1 - \varepsilon$ a.e. Since Ω is a bounded interval, this in particular means that $\rho_t^1 + \rho_t^2 > 1 - \varepsilon$ a.e. in Ω. Furthermore,

$$\int_\Omega (\rho_t^1 + \rho_t^2) \, \mathrm{d}x > (1-\varepsilon)\mathscr{L}^1(\Omega) > M_1 + M_2,$$

and this is clearly a contradiction (by the choice of ε) to fact that $\int_\Omega (\rho_t^1 + \rho_t^2) \, \mathrm{d}x = M_1 + M_2$, thus the claim follows.

Claim 2. $(\rho^1+\rho^2)^m \in L^1([0,T]; C^{0,\alpha}(\Omega))$ *for some* $0 < \alpha < 1/2$. *In particular, for a.e.* $t \in [0,T]$, $(\rho_t^1 + \rho_t^2)^m$ *has bounded oscillation uniformly in* m.

Notice that for $f : [0,T] \times \Omega \to \mathbb{R}$ measurable such that $\partial_x f \in L^r([0,T] \times \Omega)$ for some $r > 1$ and $a, b \in \Omega$, $a < b$ defining

$$\mathrm{osc}_{[a,b]} f_t := \sup_{x \in [a,b]} f_t(x) - \inf_{x \in [a,b]} f_t(x),$$

one has the estimate

$$\int_0^T (\mathrm{osc}_{[a,b]} f_t) \, \mathrm{d}t \le \int_0^T \int_a^b |\partial_x f_t| \, \mathrm{d}x\mathrm{d}t \le \left(\int_0^T \int_a^b |\partial_x f_t|^r \, \mathrm{d}x\mathrm{d}t \right)^{\frac{1}{r}} (T|a-b|)^{\frac{1}{r'}}$$

$$\le \left(\int_0^T \int_\Omega |\partial_x f_t|^r \, \mathrm{d}x\mathrm{d}t \right)^{\frac{1}{r}} (T|a-b|)^{\frac{1}{r'}},$$

where $1/r + 1/r' = 1$. Since the integrand on the l.h.s. of the previous inequality is non-negative, this implies that

$$\mathrm{osc}_{[a,b]} f_t \le C|a-b|^{\frac{1}{r'}}$$

for a.e. $t \in [0,T]$, hence in particular f_t has bounded oscillation, with a constant that depends only on $\|\partial_x f\|_{L^r}$ and T. Applying this reasoning to $(\rho^1 + \rho^2)^m$, one obtains the statement of the claim.

Now Claim 1 and Claim 2 imply that $(\rho_t^1 + \rho_t^2)^m$ is uniformly bounded for a.e. $t \in T$. This means furthermore that the average $\frac{1}{T\mathscr{L}^1(\Omega)} \int_0^T \int_\Omega (\rho^1 + \rho^2)^m \, \mathrm{d}x \mathrm{d}t$ is uniformly bounded, which together with (4.7) implies (1).

The proof of (2) follows the same lines. The bound $\|\partial_x (\rho^{1,m} + \rho^{2,m})^{m-1/2}\|_{L^2([0,T]\times\Omega)} \le C_3(m)$ in (3.10) from Theorem 3.5 remains uniform, since $C_3(m)$ remains bounded uniformly when $m \to +\infty$. This bound is enough to perform the same analysis as in (1), thus we can conclude the same way.

$\square$

Theorem 4.6. *Let us assume that $m = +\infty$, the hypotheses (H_ρ^m) and (H_Φ) are fulfilled and the setting in (Hyp-1D) takes place. Then the system (PME_∞) has a weak solution (ρ^1, ρ^2, p) in the sense of (Weak) such that $\rho^i \in L^\infty([0,T] \times \Omega) \cap AC^2([0,T]; (\mathscr{P}^{M_i}(\Omega), W_2)), i = 1, 2$ and $p \in L^2([0,T]; H^1(\Omega))$. One has moreover $\rho^1 + \rho^2 \le 1$ a.e. in $[0,T] \times \Omega$, $p \ge 0$, $p(1 - \rho^i) = 0$ a.e. in $\{\rho^i > 0\}$, $i = 1, 2$.*

Proof. Let us take a positive vanishing sequence of time steps $(\tau_n)_{n \ge 0}$ and consider the piecewise constant and continuous interpolations of density curves $(\tilde{\rho}^{i,\tau_n})_{n \ge 0}$, $(\rho^{i,\tau_n})_{n \ge 0}$ and momenta $(\widetilde{\mathbf{E}}^{i,\tau_n})_{n \ge 0}$, $(\mathbf{E}^{i,\tau_n})_{n \ge 0}$. By Proposition 3.7 (up to passing to subsequences) these objects converge (to ρ^i and $\mathbf{E}^i$ respectively) in the appropriate weak senses and one has a limit system as in (3.15)-(3.16). To identify a precise form of the system, we use the fact that the momentum sequences $(\widetilde{\mathbf{E}}^{i,\tau_n})_{n \ge 0}$ and $(\mathbf{E}^{i,\tau_n})_{n \ge 0}$ and the curve sequences $(\tilde{\rho}^{i,\tau_n})_{n \ge 0}$ and $(\rho^{i,\tau_n})_{n \ge 0}$ converge to the same limit.

Now observe that the setting in (Hyp-1D) implies that Theorem 4.2 can be applied, so the assumptions of Remark 3.4 are fulfilled. This implies that the limit momenta have the form

$$\mathbf{E}^i = -\partial_x p - \partial_x \Phi_i \rho^i = -\partial_x p \rho^i - \partial_x \Phi_i \rho^i, \quad i = 1, 2.$$

Here $p \in L^2([0,T]; H^1(\Omega))$ is the weak limit of $(p^{\tau_n})_{n \ge 0}$ obtained in Theorem 3.9, so in particular $p \ge 0$ and $p(1 - (\rho^1 + \rho^2)) = 0$ a.e. in $[0,T] \times \Omega$. These imply that the limit system has the form

$$\partial_t \rho^i - \partial_x \left(\partial_x p \rho^i + \partial_x \Phi_i \rho^i \right) = 0, \quad i = 1, 2,$$

which has to be understood in the weak sense with no-flux boundary conditions.

At last, since Theorem 4.2 implies in particular that $\mathscr{L}^1(\{\rho_t^1 > 0\} \cap \{\rho_t^2 > 0\}) = 0$ for all $t \in [0,T]$, the relation $p(1 - (\rho^1 + \rho^2)) = 0$ a.e. in $[0,T] \times \Omega$ reads as $p(1 - \rho^i) = 0$ a.e. in $\{\rho^i > 0\}$, $i = 1, 2$.

$\square$

It is not hard to verify that, for a fixed $\tau > 0$, the functionals in (MM_m) Γ-convergence as $m \to \infty$ to the functional where $\mathcal{F}_m$ is replaced by $\mathcal{F}_\infty$. Thus, it is natural to pose the question about the convergence of the corresponding gradient flow solutions in the spirit of Sandier and Serfaty (see [41]). Unfortunately, one cannot use these kinds of results directly, and obtain the convergence of the continuum solutions of (PME_m) to the solutions of (PME_∞), mainly due to the lack of uniqueness. Hence, it is necessary to proceed by studying the convergence of the continuum solutions at the PDE level. This will be addressed in the next section.

4.4. **Passing to the limit as $m \to +\infty$.** We will show that solutions of (PME_m) converge, along a subsequence as $m \to +\infty$, to a solution of (PME_∞).

We suppose that the initial data satisfy

$$\|\rho_0^1 + \rho_0^2\|_{L^\infty} \le 1.$$

Let us recall (see Definition 3.1) that a triple of nonnegative functions $(\rho^{1,\infty}, \rho^{2,\infty}, p^\infty)$, such that $\rho^{i,\infty} \in AC^2([0,T]; \mathscr{P}^{M_i}(\Omega))$, $\|\rho^{1,\infty} + \rho^{2,\infty}\|_{L^\infty([0,T]\times\Omega)} \le 1$, and $p^\infty \in L^2([0,T]; H^1(\Omega))$, is a weak solution of (PME_∞) if for any $\phi \in C^1([0,T] \times \Omega)$ and $0 < s < t \le T$ we have

$$-\int_s^t \int_\Omega \rho^{i,\infty} \partial_t \phi \, \mathrm{d}x \mathrm{d}\tau - \int_s^t \int_\Omega (\partial_x p + \partial_x \Phi_i) \rho^{i,\infty} \cdot \partial_x \phi \, \mathrm{d}x \mathrm{d}\tau = \int_\Omega \rho^{i,\infty}(s,x) \phi(s,x) \, \mathrm{d}x - \int_\Omega \rho^{i,\infty}(t,x) \phi(t,x) \, \mathrm{d}x,$$

and $p^\infty (1 - \rho^{1,\infty} - \rho^{2,\infty}) = 0$ a.e. in $[0,T] \times \Omega$.

Theorem 4.7. *Let $(\rho^{1,m}, \rho^{2,m})$ be a weak solution to* (PME$_m$) *in the setting of* (Hyp-1D) *with initial data satisfying $\|\rho_0^1 + \rho_0^2\|_{L^\infty} \leq 1$, where now we have noted m as a parameter. We assume moreover that the geometric condition* (4.6) *holds true for the domain Ω.*

Then, there exist $\rho^{i,\infty} \in L^\infty([0,T] \times \Omega) \cap AC^2([0,T]; (\mathscr{P}^{M_i}(\Omega), W_2))$ and $p^\infty \in L^r([0,T]; W^{1,r}(\Omega))$ for all $r \in (1,2)$, such that along a subsequence when $m \to +\infty$, $\rho^{i,m} \rightharpoonup \rho^{i,\infty}$ weakly in $L^q([0,T] \times \Omega)$ for any $q \geq 1$, $(\rho^{1,m} + \rho^{2,m})^m \rightharpoonup p^\infty$ in $L^r([0,T]; W^{1,r}(\Omega))$ and $\mathbf{E}^{i,m} \rightharpoonup \partial_x p^\infty \rho^{i,\infty} + \partial_x \Phi_i \rho^{i,\infty}$ weakly in $L^r([0,T] \times \Omega)$.

Moreover $(\rho^{1,\infty}, \rho^{2,\infty}, p^\infty)$ is a weak solution of (PME$_\infty$).

Proof. Let us recall the weak formulation of the system (PME$_m$).

$$(4.8) \qquad -\int_s^t \int_\Omega \rho^{i,m} \partial_t \phi \, dx \, d\tau - \int_s^t \int_\Omega \mathbf{E}^{i,m} \cdot \partial_x \phi \, dx \, d\tau = \int_\Omega \rho_s^{i,m}(x) \phi(s,x) \, dx - \int_\Omega \rho_t^{i,m}(x) \phi(t,x) \, dx,$$

for all $0 \leq s < t \leq T$ and $\phi \in C^1([0,T] \times \Omega)$, where $\mathbf{E}^{i,m} := -\left(\frac{m}{m-1} \partial_x (\rho^{1,m} + \rho^{2,m})^{m-1} + \partial_x \Phi_i \right) \rho^{i,m}$.

First, by the assumption $\|\rho_0^1 + \rho_0^2\|_{L^\infty} \leq 1$, the bounds for $\rho^{i,m}$ in $AC^2([0,T]; (\mathscr{P}^{M_i}(\Omega), W_2))$ (see Lemma 3.3 and (3.9)) are uniform in m, so clearly up to passing to a subsequence with m, $(\rho^{i,m})_{m>1}$ converges weakly-$\star$ to some $\rho^{i,\infty} \in AC^2([0,T]; (\mathscr{P}^{M_i}(\Omega), W_2))$. In particular this convergence is uniform in time w.r.t. W_2. By the uniform estimation (4.4), it follows that along a subsequence $(\rho^{i,m})_{m>1}$ converges weakly to $\rho^{i,\infty}$ in $L^q([0,T] \times \Omega)$ for all $q \geq 1$. In particular, these weak convergences allow us to obtain that in (4.8) the first term on the l.h.s., and both terms on the r.h.s. pass to the limit.

Second, Lemma 4.5 ensures the uniform boundedness of $(\rho^{1,m} + \rho^{2,m})^m$ in $L^r([0,T]; W^{1,r}(\Omega))$ for some $r \in (1,2)$ (where r can be chosen arbitrarily close to 2 for m large enough) hence there exists $p^\infty \in L^r([0,T]; W^{1,r}(\Omega))$ such that up to passing to a subsequence in m, $(\rho^{1,m} + \rho^{2,m})^m \rightharpoonup p^\infty$ weakly in $L^r([0,T]; W^{1,r}(\Omega))$. In particular, one has also that $\partial_x (\rho^{1,m} + \rho^{2,m})^m \rightharpoonup \partial_x p^\infty$ weakly in $L^r([0,T] \times \Omega)$.

Notice that the convergence $\mathbf{E}^{i,m} \rightharpoonup \partial_x p^\infty \rho^{i,\infty} + \partial_x \Phi_i \rho^{i,\infty}$ weakly in $L^r([0,T] \times \Omega)$ is much more delicate, since both terms in the product $-\frac{m}{m-1} \partial_x (\rho^{1,m} + \rho^{2,m})^{m-1} \rho^{i,m}$ (in the definition of $\mathbf{E}^{i,m}$) converge only weakly.

We shall provide the convergence $\mathbf{E}^{i,m} \rightharpoonup \partial_x p^\infty \rho^{2,\infty} + \partial_x \Phi_i \rho^{i,\infty}$ only for $i = 2$, the other case is analogous. Observe that by the uniform estimation (in m) on $\mathbf{E}^{2,m}$ in $L^r([0,T] \times \Omega)$ (for some $1 < r < 2$) (see (4.5)), there exists $\mathbf{E}^{2,\infty} \in L^r([0,T] \times \Omega)$ such that up to passing to a subsequence, $\mathbf{E}^{2,m} \rightharpoonup \mathbf{E}^{2,\infty}$ weakly in $L^r([0,T] \times \Omega)$ as $m \to +\infty$. Now let us identify the limit $\mathbf{E}^{2,\infty}$. Let us fix a subsequence (that for simplicity of notation we denote by m), such that $(\rho^{2,m})_{m>1}$ converges weakly to $\rho^{2,\infty}$ and $(\mathbf{E}^{2,m})_{m>1}$ converges weakly to $\mathbf{E}^{2,\infty}$ as $m \to +\infty$ in the previously described spaces.

Let us fix $0 \leq s < t \leq T$. For all $\tau \in [s,t]$, we define $I^{2,m}(\tau)$ be the "right-most point" of the support of $\rho_\tau^{2,m}$, i.e.

$$I^{2,m}(\tau) := \sup \left\{ x : x \in \text{Leb}\left(\{\rho_\tau^{2,m} > 0\} \right) \right\}$$

and let us define the set

$$\mathcal{I} = \left\{ (\tau, x) \in [s,t] \times \Omega : \exists \, (\tau_n, x_n)_{n \geq 0}, \text{ s.t. } x_n = I^{2,m_n}(\tau_n), \text{ and } (\tau_n, x_n) \to (\tau, x) \text{ as } n \to \infty \right\},$$

where $(m_n)_{n \geq 0}$ is a subsequence of the previously chosen subsequence. Then $\mathcal{I}$ is a closed subset of $[s,t] \times \Omega$. Let $\mathcal{I}(\tau) := \mathcal{I} \cap (\{\tau\} \times \Omega)$. Note that in particular $\mathcal{I}(\tau)$ is the collection of all subsequential limits of $I^{2,m}(\tau)$.

Observe that if for some $\tau \in (s,t)$, $y \in \Omega$ lies to the left of $\mathcal{I}(\tau)$, i.e. if $y < x$ for any $x \in \mathcal{I}(\tau)$, then (τ, y) lies in the complement of $\{\rho^{1,m} > 0\}$ for sufficiently large m, and thus defining

$$\mathcal{J}^- := \{(\tau, y) \in (s,t) \times \Omega : y < x \text{ for any } x \in \mathcal{I}(\tau)\},$$

one has that when restricted to $\mathcal{J}^-$,

$$\mathbf{E}^{2,m} = -\partial_x (\rho^{2,m})^m - \partial_x \Phi \rho^{2,m},$$

in the sense of distributions for sufficiently large m. Similarly, defining

$$\mathcal{J}^+ := \{(\tau, y) \in (s,t) \times \Omega : y > x \text{ for any } x \in \mathcal{I}(\tau)\},$$

one has that $\rho^{2,m} = 0$ a.e. on $\mathcal{J}^+$, hence when restricted to $\mathcal{J}^+$, $\mathbf{E}^{2,m} = 0$ a.e. for sufficiently large m. Clearly, $\mathcal{J}^-, \mathcal{I}, \mathcal{J}^+$ are Lebesgue measurable and one can write $(s,t) \times \Omega = \mathcal{J}^- \cup \mathcal{I} \cup \mathcal{J}^+$, $\mathscr{L}^1 \llcorner [0,T] \otimes \mathscr{L}^1 \llcorner \Omega$−a.e. Thus, we write furthermore

$$\int_s^t \int_\Omega \mathbf{E}^{2,m} \cdot \partial_x \phi \, \mathrm{d}x \mathrm{d}\tau = \int_{\mathcal{J}^-} \mathbf{E}^{2,m} \cdot \partial_x \phi \mathrm{d}\tau \otimes \mathrm{d}x + \int_{\mathcal{I}} \mathbf{E}^{2,m} \cdot \partial_x \phi \mathrm{d}\tau \otimes \mathrm{d}x + \int_{\mathcal{J}^+} \mathbf{E}^{2,m} \cdot \partial_x \phi \mathrm{d}\tau \otimes \mathrm{d}x$$
$$= \int_{\mathcal{J}^-} \mathbf{E}^{2,m} \cdot \partial_x \phi \mathrm{d}\tau \otimes \mathrm{d}x + \int_{\mathcal{I}} \mathbf{E}^{2,m} \cdot \partial_x \phi \mathrm{d}\tau \otimes \mathrm{d}x.$$

Moreover, the very same decomposition remains valid for the weak limit $\mathbf{E}^{2,\infty}$ as well.

Claim 1. $\int_{\mathcal{I}} \mathbf{E}^{2,\infty} \cdot \partial_x \phi \mathrm{d}\tau \otimes \mathrm{d}x$ can be made arbitrarily small for any smooth test function ϕ.

If $(\mathscr{L}^1 \otimes \mathscr{L}^1)(\mathcal{I}) = 0$, then this is obvious. So one can suppose that this set has positive measure. To show the claim, let us define the width of $\mathcal{I}(\tau)$, i.e.

$$W(\tau) := \max\{|x - y| : x, y \in \mathcal{I}(\tau)\} = x_\tau^2 - x_\tau^1,$$

where $x_\tau^2 := \max\{x : x \in \mathcal{I}(\tau)\}$ and $x_\tau^1 := \min\{x : x \in \mathcal{I}(\tau)\}$ and these are well-defined since $\mathcal{I}(\tau)$ is compact. Let $T_n := \left\{\tau \in (s,t) : W(\tau) \geq \frac{2}{n}\right\}$. Then $\mathcal{I} \subset A_n \cup B_n$, where

$$A_n := \bigcup_{\tau \in T_n} \left(\{\tau\} \times (x_\tau^1 + 1/n, x_\tau^2)\right) \quad \text{and} \quad B_n = \mathcal{I} \setminus A_n.$$

Note that A_n and B_n are Lebesgue measurable and the measure of B_n in $[0,T] \times \Omega$ is at most $2T/n$, which goes to zero as $n \to \infty$.
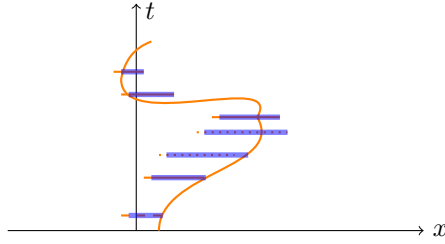

FIGURE 3. The sets $\mathcal{I}$ (with orange) and A_n
(with blue)

We know that $W_2(\rho_\tau^{i,m}, \rho_\tau^{2,\infty}) \to 0$ as $m \to +\infty$, uniformly in τ. This implies in particular that the sequence $(\rho_\tau^{2,m})_m$ are Cauchy w.r.t. W_2 uniformly in τ, which means that for any $n \in \mathbb{N}$, there exists $N(n) > 0$ such that for all $\tau \in [0,T]$

$$(4.9) \qquad\qquad W_2^2(\rho_\tau^{2,m_k}, \rho_\tau^{2,m_l}) \leq \frac{1}{n^3}, \ \forall k, l > N(n),$$

where m_k and m_l denote elements of the sequence denoted by m. Now, let us define $A_n(\tau) := A_n \cap \{\tau\} \times \Omega$.

Let us show that for all $\tau \in (s,t)$ on $A_n(\tau)$ all elements of the sequence $(\rho_\tau^{2,m})_m$ have small mass. Indeed, on the one hand, for any point $x \in A_n(\tau) \cap \mathcal{I}(\tau)$ (if the intersection is $\emptyset$, then there is nothing to show), there is a subsequence $\tilde{m}$ of m such that $I^{2,\tilde{m}}(\tau) \to x$ as $\tilde{m} \to +\infty$. On the other hand, since $x_\tau^1 \in \mathcal{I}(\tau)$, there exists another subsequence $\overline{m}$ of m such that $I^{2,\overline{m}}(\tau) \to x_\tau^1$ as $\overline{m} \to +\infty$. The inequality (4.9) implies that

$$\frac{1}{n^2} \int_{A_n(\tau)} \rho_t^{2,\tilde{m}} \mathrm{d}x \leq W_2^2(\rho_\tau^{2,\tilde{m}}, \rho_\tau^{2,\overline{m}}) \leq \frac{1}{n^3}$$

for n large enough and $\tilde{m}$ and $\overline{m}$ larger than $N(n)$. The first inequality holds because of the fact that $x - x_\tau^1 \geq 1/n$. Thus for all $\tau \in (s,t)$

$$\int_{A_n(\tau)} \rho_\tau^{2,m} \mathrm{d}x \leq 1/n, \quad \text{for any } m \text{ large enough.}$$

Considering any smooth test function ϕ supported in $A_n \cup \mathcal{J}^+$, the weak formulation (4.8) together with the fact that $\mathbf{E}^{2,m}$ and $\rho^{2,m}$ vanish on $\mathcal{J}^+$ yield

$$\left| \int_{A_n} \mathbf{E}^{2,m} \cdot \partial_x \phi \mathrm{d}t \otimes \mathrm{d}x \right| \leq \sup_{A_n} |\partial_t \phi| \frac{T}{n} + \sup_{A_n} |\phi| \frac{2}{n},$$

for sufficiently large m. This together with the fact that the measure of B_n is at most $2T/n$ and $\mathcal{I} \subset A_n \cup B_n$ implies that $\int_{\mathcal{I}} E^{2,m} \cdot \partial_x \phi \mathrm{d}t \otimes \mathrm{d}x$ is arbitrary small provided m is large enough, which implies in particular that $\int_{\mathcal{I}} E^{2,\infty} \cdot \partial_x \phi \mathrm{d}t \otimes \mathrm{d}x$ is as small as we would like.

The above claim shows that one needs to describe the weak limit $\mathbf{E}^{2,\infty}$ only on the set $\mathcal{J}^-$. When restricted to $\mathcal{J}^-$, one can write

$$\mathbf{E}^{2,m} = -\partial_x (\rho^{2,m})^m - \partial_x \Phi_2 \rho^{2,m} =: -\partial_x p_m - \partial_x \Phi_2 \rho^{2,m},$$

and the second term on the r.h.s. of the previous formula passes to the limit due to the weak convergence of $(\rho^{2,m})_{m>1}$. Let us consider a smooth test function ϕ compactly supported in $\mathcal{J}^-$. Then one has

$$\lim_{m \to +\infty} - \int_{\mathcal{J}^-} \partial_x p_m \cdot \partial_x \phi \mathrm{d}\tau \otimes \mathrm{d}x = \lim_{m \to +\infty} \int_{\mathcal{J}^-} p_m \partial^2_{xx} \phi \mathrm{d}\tau \otimes \mathrm{d}x = \lim_{m \to +\infty} \int_{\mathcal{J}^-} (\rho^{2,m})^m \partial^2_{xx} \phi \mathrm{d}\tau \otimes \mathrm{d}x$$

$$= \lim_{m \to +\infty} \int_s^t \int_\Omega (\rho^{1,m} + \rho^{2,m})^m \partial^2_{xx} \phi \mathrm{d}x \mathrm{d}\tau = \int_s^t \int_\Omega p^\infty \partial^2_{xx} \phi \mathrm{d}x \mathrm{d}\tau$$

$$= - \int_s^t \int_\Omega \partial_x p^\infty \partial_x \phi \mathrm{d}x \mathrm{d}\tau = - \int_{\mathcal{J}^-} \partial_x p^\infty \cdot \partial_x \phi \mathrm{d}\tau \otimes \mathrm{d}x$$

Hence when restricted to $\mathcal{J}^-$, $\mathbf{E}^{2,\infty} = -\partial_x p^\infty - \partial_x \Phi_2 \rho^{2,\infty}$. Note that $\rho^{2,\infty} = 0$ as well in $\mathcal{I} \cup \mathcal{J}^+$, so we can write

$$\mathbf{E}^{2,\infty} = -\partial_x p^\infty \chi_{\{\rho^{2,\infty} > 0\}} - \partial_x \Phi_2 \rho^{2,\infty}.$$

Below we will show that p^∞ vanishes in $\rho^{2,\infty} < 1$. This allows us to write

$$\mathbf{E}^{2,\infty} = -\partial_x p^\infty \rho^{2,\infty} - \partial_x \Phi_2 \rho^{2,\infty}.$$

Similar reasoning yields the concrete form of $\mathbf{E}^{1,\infty}$ as well.

Let us show that $\|\rho^{i,\infty}\|_{L^\infty} \leq 1$. By Lemma 4.5 we know that for a.e. $t \in [0,T]$, $\int_\Omega (\rho_t^{i,m})^m \mathrm{d}x \leq C$, where the constant C is independent of m. Thus for any $\delta > 0$, on the set where $\rho_t^{i,m} \geq 1 + \delta$ a.e. we have by Chebyshev's inequality that

(4.10) $$(1+\delta)^m \mathscr{L}^1(\{\rho_t^{i,m} \geq 1 + \delta\}) \leq \int_{\{\rho_t^{i,m} \geq 1+\delta\}} (\rho_t^{i,m})^m \mathrm{d}x \leq \int_\Omega (\rho_t^{i,m})^m \mathrm{d}x \leq C.$$

This implies

$$\mathscr{L}^1(\{\rho_t^{i,m} \geq 1 + \delta\}) \leq C/(1+\delta)^m \to 0 \text{ as } m \to \infty,$$

and thus by the arbitrariness of $\delta > 0$ we conclude $\rho^{i,\infty} \leq 1$ a.e in $[0,T] \times \Omega$.

At last, it remains to show that p^∞ is supported in the region $\{\rho^{1,\infty} + \rho^{2,\infty} = 1\}$. Notice that since Theorem 4.2 yields that $\mathscr{L}^1(\{\rho_t^{1,\infty} > 0\} \cap \{\rho_t^{2,\infty} > 0\}) = 0$ for all $t \in [0,T]$, it is enough to show that $p^\infty(1 - \rho^{i,\infty}) = 0$ a.e. in $\{\rho^{i,\infty} > 0\}$ $i = 1, 2$. We show this property only in the case of $i = 2$, the other case is analogous. Let us use the notations

$$p_m := (\rho^{1,m} + \rho^{2,m})^m =_{\text{a.e.}} (\rho^{2,m})^m, \quad \text{in } \{\rho^{2,\infty} > 0\} \subseteq_{\text{a.e.}} \mathcal{J}^-$$

$$\tilde{p}_m := (\rho^{1,m} + \rho^{2,m})^{m-1/2} =_{\text{a.e.}} (\rho^{2,m})^{m-1/2} = (p_m)^{1-\frac{1}{2m}}, \quad \text{in } \{\rho^{2,\infty} > 0\} \subseteq_{\text{a.e.}} \mathcal{J}^-$$

Claim 2. When $m \to +\infty$, $\int_0^T \int_\Omega p_m(1 - \rho^{2,m}) \mathrm{d}x \mathrm{d}t$ and $\int_0^T \int_\Omega \tilde{p}_m(1 - \rho^{2,m}) \mathrm{d}x \mathrm{d}t$ are arbitrary small.

Let $\delta > 0$ be small. To show the claim, first observe that $[0,T] \ni t \mapsto \int_\Omega \chi_{\{\rho_t^{2,m} \geq 1+\delta\}} \mathrm{d}x$ is a measurable function and by (4.10), $[0,T] \ni t \mapsto \int_\Omega \chi_{\{\rho_t^{2,m} \geq 1+\delta\}} \mathrm{d}x$ is integrable. Similar properties are valid for other

characteristic functions of the (sub)level sets. Thus,

$$
\left| \int_0^T \int_\Omega p_m (1 - \rho^{2,m}) \,\mathrm{d}x\,\mathrm{d}t \right| \le \int_0^T \int_{\{\rho_t^{2,m} < 1 - \delta\}} p_m (1 - \rho^{2,m}) \,\mathrm{d}x\,\mathrm{d}t + \int_0^T \int_{\{1 - \delta \le \rho_t^{2,m} \le 1 + \delta\}} p_m |1 - \rho^{2,m}| \,\mathrm{d}x\,\mathrm{d}t
$$
$$
+ \int_0^T \int_{\{\rho_t^{2,m} > 1 + \delta\}} p_m (\rho^{2,m} - 1) \,\mathrm{d}x\,\mathrm{d}t
$$
$$
\le (1 - \delta)^m T \mathscr{L}^1(\Omega) + \delta C
$$
$$
+ \|p_m\|_{L^{q_1}([0,T]\times\Omega)} \|\rho^{2,m}\|_{L^{q_2}([0,T]\times\Omega)} \left(\int_0^T \int_{\{\rho_t^{2,m} > 1 + \delta\}} 1 \,\mathrm{d}x\,\mathrm{d}t \right)^{1/q_3}
$$
$$
\le (1 - \delta)^m T \mathscr{L}^1(\Omega) + \delta C + \left(\frac{TC}{(1 + \delta)^m} \right)^{1/q_3},
$$

where $\frac{1}{q_1} + \frac{1}{q_2} + \frac{1}{q_3} = 1$, $q_1 < r$, $q_2 \le m$ and C is independent of m. Here we used also the uniform estimations from Lemma 4.5(1) and Theorem 4.4. Now the last sum is as small as desired by choosing $\delta > 0$ small and m large enough, which shows the first part of the claim. A similar reasoning and Lemma 4.5(2) yield the second part of the claim.

Last, one can use the same arguments as in [30, Lemma 3.4, Step 3. in the proof of Theorem 2.1.] to conclude that $p^\infty(1 - \rho^{2,\infty}) = 0$ a.e. in $\mathcal{J}^-$. Indeed, we know that $\tilde{p}_m$ is uniformly bounded in $L^2([0,T]; H^1(\Omega))$ and by Lemma 4.5(2) $\rho_t^{2,m}$ is uniformly bounded for a.e. $t \in [0,T]$. These together with *Claim 2* imply in particular that mentioned results from [30] can be applied to obtain that

$$
\tilde{p}^\infty(1 - \rho^{2,\infty}) = 0, \quad \text{a.e. in } \mathcal{J}^-,
$$

where $\tilde{p}^\infty$ is the weak limit of $(\tilde{p}_m)_{m>1}$ in $L^2([0,T]; H^1(\Omega))$ as $m \to +\infty$. It remains to show only that $p^\infty = \tilde{p}^\infty$ a.e. in $\mathcal{J}^-$. This is straight forward. Indeed, notice that $\tilde{p}_m = (p_m)^{1 - \frac{1}{2m}}$ a.e. in $\mathcal{J}^-$. Lemma 4.5 implies that for a.e. $t \in [0,T]$ both $p_m(t, \cdot)$ and $\tilde{p}_m(t, \cdot)$ are (uniformly) Hölder continuous and converge uniformly (up to passing to a subsequence). This means that p_t^∞ and $\tilde{p}_t^\infty$ are the uniform limits for a.e. $t \in [0,T]$ and $p_t^\infty = \tilde{p}_t^\infty$ for a.e. $t \in [0,T]$. The result follows. $\qquad \square$

4.5. Characterization of the pressure in the case of $m = +\infty$.

Here we establish the optimal regularity of the pressure, which is Lipschitz continuity for a.e. time. The pressure can be discontinuous in time even in the single density case, when two components of the congested density zone merge into one (see the discussion in [23] for instance.)

Proposition 4.8. *Let $(\rho^{1,\infty}, \rho^{2,\infty}, p^\infty)$ be a solution of the system* (PME$_\infty$) *in the setting of* (Hyp-1D)*. Then $p^\infty(t, \cdot)$ is uniformly Lipschitz continuous for a.e. $t \in [0,T]$. Moreover, $p^\infty(t, \cdot)$ is as smooth as Φ_i in the interior of the sets $\{\rho^{i,\infty} = 1\}$, $i = 1,2$ for a.e. $t \in [0,T]$.*

Let us remark first that p^∞ may have positive boundary data on $\partial\{\rho^{1,\infty} = 1\} \cap \partial\{\rho^{2,\infty} = 1\}$. Also, notice that the set $\{\rho^{i,\infty} = 1\}$ may have empty interior.

Proof. Considering the sum of the two weak equations tested against smooth test functions supported in the interior of $\{\rho_t^{1,\infty} + \rho_t^{2,\infty} = 1\}$ one obtains that

$$
(4.11) \qquad -\partial_{xx}^2 p^\infty(t, \cdot) = \partial_x(\partial_x \Phi_1 \rho^{1,\infty} + \partial_x \Phi_2 \rho^{2,\infty}), \quad \text{in the interior of } \{\rho_t^{1,\infty} + \rho_t^{2,\infty} = 1\},
$$

for a.e. $t \in [0,T]$ with homogeneous Dirichlet boundary data (this is because of the fact that $p^\infty(t, \cdot)$ is Hölder continuous a.e. $t \in [0,T]$). Since $\mathscr{L}^1\left(\{\rho_t^{1,\infty} > 0\} \cap \{\rho_t^{2,\infty} > 0\}\right) = 0$ and the sets $\{\rho_t^{1,\infty} > 0\}$ and $\{\rho_t^{2,\infty} > 0\}$ are ordered in the sense of (Hyp-1D), one gets

$$
(4.12) \qquad -\partial_{xx}^2 p^\infty(t, \cdot) = \partial_{xx}^2 \Phi_i, \quad \text{in the interior of } \{\rho_t^{i,\infty} = 1\}.
$$

Therefore $p^\infty(t, \cdot)$ is as smooth as Φ_i in the interior of the sets $\{\rho_t^{i,\infty} = 1\}$, $i = 1, 2$, for a.e. $t \in [0, T]$. Let us remark also that since $p^\infty(t, \cdot)$ is Hölder continuous, the set $\{p^\infty(t, \cdot) > 0\} \subseteq \{\rho_t^{1,\infty} + \rho_t^{2,\infty} = 1\}$ is open. Thus $p^\infty(t, \cdot)$ is as smooth as Φ_i in $\{p^\infty(t, \cdot) > 0\} \cap \text{int}\{\rho_t^{i,\infty} = 1\}$.

Let us show that $p^\infty(t, \cdot)$ is Lipschitz continuous for a.e. $t \in [0, T]$. Notice that by the previous arguments, $p^\infty(t, \cdot)$ fails to be differentiable in at most countably many points of Ω, and except these points it is smooth. By the fact that $p^\infty(t, \cdot)$ is a Sobolev function, we know that it is absolutely continuous for a.e $t \in [0, T]$. Moreover

$$|\partial_x p^\infty(t, \cdot)| \le C, \quad \text{a.e. in } \Omega,$$

where $C > 0$ is a positive constant that depends only on $\max\{\|\partial_x \Phi_i\|_{L^\infty} : i = 1, 2\}$. This together with the absolute continuity imply that $p^\infty(t, \cdot)$ Lipschitz continuous for a.e. $t \in [0, T]$.

$\square$

4.6. Patch solutions.

Proposition 4.9. *Let $(\rho^{1,\infty}, \rho^{2,\infty}, p^\infty)$ be a solution of the system (PME$_\infty$) in the setting of (Hyp-1D). Let us suppose that $\rho_0^{1,\infty}$ and $\rho_0^{2,\infty}$ are patches, i.e. $\rho_0^{i,\infty} = \chi_{A_0^i}$ for open intervals A_0^i in Ω, $i = 1, 2$. We suppose moreover that the drifts $-\partial_x \Phi_i$, $i = 1, 2$ are compressive, meaning that $\partial_{xx}^2 \Phi_i \ge 0$ on Ω. Then $\rho_t^{i,\infty}$, $i = 1, 2$ is a patch for all $t \in [0, T]$, i.e. there exists $\{A_0^i(t)\}_{t \in [0,T]}$: a family of open intervals such that $\rho_t^{i,\infty} = \chi_{A_0^i(t)}$.*

Remark 4.2. *While we believe our argument can be extended to measurable sets instead of open intervals, we do not pursue this generalization for simplicity.*

Proof. Let us recall that the sets $\{\rho_t^{2,\infty} > 0\}$ and $\{\rho_t^{1,\infty} > 0\}$ are ordered for all $t \in [0, T]$ in the sense of (Hyp-1D), with $\rho^{2,\infty}$ supported to the left of $\rho^{1,\infty}$. We show the proposition only for $\rho^{2,\infty}$, the case of $\rho^{1,\infty}$ is analogous.

We define an extension $\tilde{p}$ of p^∞ to the right of the support of $\rho^{2,\infty}$. For a.e. t where p^∞ satisfies (4.12), let $x_0(t) := \sup \text{Leb}\left(\{\rho_t^{2,\infty} > 0\}\right)$. If $p^\infty(t, x_0(t)) = 0$ then we let $\tilde{p}(t, \cdot) = p^\infty(t, \cdot)$. If $p(t, x_0(t)) > 0$, this means $x_0(t)$ lies in the interior of $\{\rho_t^{1,\infty} + \rho_t^{2,\infty} = 1\}$. In this case let us define $\tilde{p}$ to be a C^2 extension of p^∞ to the right of $x_0(t)$ such that $-\partial_{xx}^2 \tilde{p} = \partial_{xx}^2 \Phi_2$.

Let us consider the test function ϕ, as the solution of the transport equation

$$\partial_t \phi_t - v^\varepsilon \partial_x \phi = 0$$

with initial condition $\phi(0, \cdot) = \chi_{A_0^2}$, where we define $v^\varepsilon := v \star \eta_\varepsilon$ with $v = (\partial_x \tilde{p} + \partial_x \Phi_i)$ and η_ε is a standard mollifier (the mollification being performed only w.r.t. the space variable). Then $\|v^\varepsilon(t, \cdot) - v(t, \cdot)\|_{L^q} \le C\varepsilon$ in the set $\{\rho_t^{2,\infty} > 0\}$ for any $q > 0$ and a.e. $t \in [0, T]$, where C depends on the Lipschitz constant of p^∞ and Φ_2. Then, from the weak expression we have

$$\left| \int_\Omega (\rho^{2,\infty} \phi)(t, \cdot) \mathrm{d}x - \int_\Omega (\rho^{2,\infty} \phi)(0, \cdot) \mathrm{d}x \right| = \left| \int_0^t \int_\Omega (v_\varepsilon - v) \rho^{2,\infty} \partial_x \phi \, \mathrm{d}x \, \mathrm{d}\tau \right| \le C\varepsilon.$$

Since ϕ solves a transport equation with spatially smooth velocity which is integrable in time, this equation is well-posed and ϕ can be represented using the method of characteristics. In particular $\phi(t, \cdot) = \chi_{A_t^2}$ for some measurable set A_t^2 for each time $t > 0$. Hence the l.h.s. of the above equation is

$$(4.13) \qquad \int_\Omega (\rho^{2,\infty} \phi)(t, \cdot) \mathrm{d}x - \int_\Omega (\rho^{2,\infty} \phi)(0, \cdot) \mathrm{d}x = \int_{A_t^2} \rho_t^{2,\infty} \mathrm{d}x - \int_{A_0^2} 1 \mathrm{d}x.$$

Now we claim that

$$(4.14) \qquad\qquad\qquad\qquad\qquad \partial_x v^\varepsilon \ge 0.$$

This is true because $\tilde{p}$ satisfies $-\partial_{xx}^2 \tilde{p} = \partial_{xx}^2 \Phi_2$ (or $\tilde{p} = p^\infty$) to the right of $x_0(t)$ as well as in any interior point of $\{\rho_t^{2,\infty} = 1\}$, and to the left of $x_0(t)$ we have $\tilde{p} = p^\infty = \max\{p^\infty, 0\}$, which makes p^∞ convex at any boundary point of the set $\{\rho^{2,\infty} = 1\}$ to the left of $x_0(t)$. Thus we conclude that

$$\partial_x v = \partial_{xx}^2 p^\infty + \partial_{xx}^2 \Phi_2 \ge 0 \quad \text{on } (-\infty, x_0(t)) \cap \Omega,$$

in the distributional sense, which yields $\partial_x v^\varepsilon \ge 0$.

Due to (4.14) we have $\mathscr{L}^1(A_t^2) \leq \mathscr{L}^1(A_0^2)$. This, the fact that $\rho_t^{2,\infty} \leq 1$ a.e., and (4.13) yield that

$$\int_{A_t^2} (1 - \rho_t^{2,\infty}) \mathrm{d}x \leq \int_{A_t^2} (1 - \rho_t^{2,\infty}) \mathrm{d}x + \mathscr{L}^1(A_0^2) - \mathscr{L}^2(A_t^2) \leq C\varepsilon.$$

We still need to let $\varepsilon \downarrow 0$ to achieve the desired result. If A_0^2 is an interval then A_t^2 is an interval (that depends on $\varepsilon > 0$) with uniformly bounded velocity with respect to ε, hence along a subsequence the endpoints (as a function of t they are equicontinuous) uniformly converge to limiting endpoints $(a(t), b(t))_{\{t>0\}}$ as $\varepsilon \downarrow 0$. Let $A_0^2(t) := (a(t), b(t))$. Here $\rho_t^{2,\infty}$ should be identically one (because of the previous inequality). But this means that along this subsequence, v was incompressible except in a small set in A_t^2, so that makes $\mathscr{L}^1(A_t^2)$ very close to $\mathscr{L}^1(A_0^2)$ and $|b(t) - a(t)| = \mathscr{L}^1(A_0^2)$. Since $\rho^{2,\infty}$ preserves mass over time, this means that

$$\rho_t^{2,\infty} = \chi_{A_0^2(t)}.$$

$\square$

It remains to describe the evolution of the patches $\{\rho^{i,\infty} = 1\}$. As we see in [24] in [35], the evolution laws are different depending on whether there are regions of the densities with values between zero and one. In the above Proposition, we have patch solutions supported on an interval, and the continuity of the densities over time in W_2-distance yields that each patch $\{\rho^{i,\infty} = 1\}$ evolves continuously in time. Therefore it follows that the space-time interior of those sets taken at time t equals spatial interior at time t. Thus from (4.12) we have $p^\infty(t, \cdot) \in C^2$ at every time in the interior of $\{\rho_t^{1,\infty} + \rho_t^{2,\infty} = 1\}$.

Remark 4.3. *With the aforementioned regularity of p^∞ and $\{\rho^{i,\infty} = 1\}$ at hand, one can verify with test functions in the weak formulation that the following holds: the velocity law on one-phase boundary points is given by*

$$(4.15) \qquad V = \nu_x^i(-\partial_x^i p^\infty - \partial_x \Phi_i) \quad \text{on} \quad \partial\{\rho^{i,\infty} = 1\},$$

where ν_x is the outward normal of the set $\{\rho^{i,\infty} = 1\}$, V is the normal velocity of the interface and ∂_x^i denotes the x-derivative taken from the interior of the set $\{\rho^{i,\infty} = 1\}$. This yields the flux matching across different densities,

$$(4.16) \qquad \partial_x^1 p^\infty + \partial_x \Phi_1 = \partial_x^2 p^\infty + \partial_x \Phi_2 \quad \text{on} \quad \partial\{\rho^{1,\infty} = 1\} \cap \partial\{\rho^{2,\infty} = 1\}.$$

The equations (4.12), (4.15) and (4.16) corresponds to a generalized two-phase Hele-Shaw flow evolving by the pressure variable p^∞ where different drift potentials are present for each phase $\rho^{i,\infty}$.

APPENDIX A. OPTIMAL TRANSPORT TOOLBOX

Lemma A.1. *Let $f : [0, +\infty) \to \mathbb{R}$ be a C^1 convex function that is superlinear at $+\infty$. Let $M > 0$. We consider $\mathcal{F} : \mathscr{P}^M(\Omega) \to \mathbb{R} \cup +\infty$ defined as*

$$\mathcal{F}(\rho) = \begin{cases} \displaystyle\int_\Omega f(\rho(x)) \mathrm{d}x, & \text{if } \rho \ll \mathscr{L}^d, \\ +\infty, & \text{otherwise.} \end{cases}$$

Let $\nu \in \mathscr{P}^M(\Omega)$ be given. Then there exists a solution $\varrho \in \mathscr{P}^{\mathrm{ac},M}(\Omega)$ of the minimization problem

$$\min_{\rho \in \mathscr{P}^M(\Omega)} \left\{ \mathcal{F}(\rho) + \frac{1}{2} W_2^2(\rho, \nu) \right\}.$$

If in addition $\nu \ll \mathscr{L}^d$ or if f is strictly convex, then ϱ is unique.

Moreover, $\exists C \in \mathbb{R}$ such that for a suitable Kantorovich potential φ in the optimal transport of ϱ onto ν one has the following first order necessary optimality condition fulfilled

$$(A.1) \qquad \begin{cases} f'(\varrho) + \varphi = C, & \varrho - \text{a.e.,} \\ f'(\varrho) + \varphi \geq C, & \text{on } \{\varrho = 0\}. \end{cases}$$

If $f'(0)$ is finite, then one can express the above condition as $f'(\varrho) = \max\{C - \varphi, f'(0)\}$.

Proof. The proof of the previous results can be found in [11] or [43, Chapter 7]. $\square$

It turns out that $(\mathscr{P}^M(\Omega), W_2)$ is a geodesic space and constant speed geodesics (and absolutely continuous curves in general) can be characterized by special solutions of continuity equations. Since this characterization is true for any $M > 0$, we simply set $M = 1$ in the theorem below.

Theorem A.2 (see [3, 43]). (1) *Let $\Omega \subset \mathbb{R}^d$ compact and $(\mu_t)_{t\in[0,T]}$ be an absolutely continuous curve in $(\mathscr{P}(\Omega), W_2)$. Then for a.e. $t \in [0, T]$ there exists a vector field $\boldsymbol{v}_t \in L^2_{\mu_t}(\Omega; \mathbb{R}^d)$ s.t.*
- *the continuity equation $\partial_t \mu_t + \nabla \cdot (\boldsymbol{v}_t \mu_t) = 0$ is satisfied in the weak sense;*
- *for a.e. $t \in [0, T]$, one has $\|v_t\|_{L^2_{\mu_t}} \leq |\mu'|_{W_2}(t)$, where*

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

denotes the metric derivative of the curve $[0, T] \ni t \mapsto \mu_t$ w.r.t. W_2, provided the limit exists.

(2) *Conversely, if $(\mu_t)_{t\in[0,T]}$ is a family of measures in $\mathscr{P}(\Omega)$ and for each t one has a vector field $\boldsymbol{v}_t \in L^2_{\mu_t}(\Omega; \mathbb{R}^d)$ s.t. $\int_0^T \|\boldsymbol{v}_t\|_{L^2_{\mu_t}} \mathrm{d}t < +\infty$ and $\partial_t \mu_t + \nabla \cdot (\boldsymbol{v}_t \mu_t) = 0$ in the weak sense, then $[0, T] \ni t \mapsto \mu_t$ is an absolutely continuous curve in $(\mathscr{P}(\Omega), W_2)$, with $|\mu'|_{W_2}(t) \leq \|v_t\|_{L^2_{\mu_t}}$ for a.e. $t \in [0, T]$ and $W_2(\mu_{t_1}, \mu_{t_2}) \leq \int_{t_1}^{t_2} |\mu'|_{W_2}(t)\mathrm{d}t$. If moreover $|\mu'| \in L^2(0, T)$, then we say that μ belongs to the space $AC^2([0, T]; (\mathscr{P}(\Omega), W_2))$.*

(3) *For curves $(\mu_t)_{t\in[0,1]}$ that are geodesics in $(\mathscr{P}(\Omega), W_2)$ one has the equality*

$$W_2(\mu_0, \mu_1) = \int_0^1 |\mu'|_{W_2}(t)\mathrm{d}t = \int_0^1 \|v_t\|_{L^2_{\mu_t}} \mathrm{d}t.$$

(4) *For $\mu_0, \mu_1 \in \mathscr{P}^{\mathrm{ac}}(\Omega)$, a constant speed geodesic connecting them is a curve $(\mu_t)_{t\in[0,1]}$ such that $W_2(\mu_s, \mu_t) = |t - s| W_2(\mu_0, \mu_1)$ for any $t, s \in [0, 1]$. One can compute this constant speed geodesic using McCann's interpolation, i.e. $\mu_t := (T_t)_\# \mu_0$, for all $t \in [0, 1]$, where $T_t := (1 - t)\mathrm{id} + tT$ with $T_\# \mu_0 = \mu_1$ the optimal transport map between μ_0 and μ_1. Moreover, the velocity field in the continuity equation is given by $\boldsymbol{v}_t := (T - \mathrm{id}) \circ (T_t)^{-1}$.*

Let us introduce the *Benamou-Brenier functional* $\mathcal{B}_2 : \mathscr{M}([0, T] \times \Omega) \times \mathscr{M}^d([0, T] \times \Omega) \to \mathbb{R} \cup \{+\infty\}$ defined as

$$\mathcal{B}_2(\mu, \mathbf{E}) := \begin{cases} \int_0^T \int_\Omega |\mathbf{v}_t|^2 \, \mathrm{d}\mu_t(x)\mathrm{d}t, & \text{if } \mathbf{E} = \mathbf{E}_t \otimes \mathrm{d}t, \mu = \mu_t \otimes \mathrm{d}t \text{ and } \mathbf{E}_t = \mathbf{v}_t \cdot \mu_t, \\ +\infty, & \text{otherwise.} \end{cases}$$

It is well-known (see for instance [43, Proposition 5.18]) that $\mathcal{B}_2$ is jointly convex and lower semicontinuous w.r.t. the weak$-\star$ convergence. In particular if $(\mu, \mathbf{E})$ solves $\partial_t \mu + \nabla \cdot \mathbf{E} = 0$ in the weak sense with $\mathcal{B}_2(\mu, \mathbf{E}) < +\infty$, implies that $t \mapsto \mu_t$ is a curve in $AC^2([0, T]; (\mathscr{P}(\Omega), W_2))$.

The following comparison result appears to be well-known but we write it here for completeness.

Lemma A.3. *Let $\mu^1, \nu^1 \in \mathscr{P}^{M_1}(\Omega)$ and $\mu^2, \nu^2 \in \mathscr{P}^{M_2}(\Omega)$. Then the following inequality holds true*

$$(A.2) \qquad W_2^2(\mu^1 + \mu^2, \nu^1 + \nu^2) \leq W_2^2(\mu^1, \nu^1) + W_2^2(\mu^2, \nu^2).$$

Remark A.1. *Note that with the abuse of notation, W_2 on the l.h.s. of (A.2) denotes the $2-$Wasserstein distance on $\mathscr{P}^{M_1+M_2}(\Omega)$, while on the r.h.s. W_2 denotes the corresponding distances on $\mathscr{P}^{M_1}(\Omega)$ and $\mathscr{P}^{M_2}(\Omega)$ respectively.*

Proof of Lemma A.3. The quantity on the l.h.s. of (A.2) is realized by an optimal plan $\gamma \in \Pi^{M_1+M_2}(\mu^1 + \mu^2, \nu^1 + \nu^2)$ i.e.

$$W_2^2(\mu^1 + \mu^2, \nu^1 + \nu^2) = \int_{\Omega \times \Omega} |x - y|^2 \mathrm{d}\gamma.$$

Similarly the quantities on the r.h.s. can be written with the help of some optimal plans $\gamma^i \in \Pi^{M_i}(\mu^i, \nu^i)$, $i = 1, 2$, i.e.

$$W_2^2(\mu^i, \nu^i) = \int_{\Omega \times \Omega} |x - y|^2 \mathrm{d}\gamma^i.$$

Now set $\tilde{\gamma} := \gamma^1 + \gamma^2$. Clearly since $(\pi^x)_\# \tilde{\gamma} = \mu^1 + \nu^1$ and $(\pi^y)_\# \tilde{\gamma} = \mu^2 + \nu^2$ one has $\tilde{\gamma} \in \Pi^{M_1 + M_2}(\mu^1 + \mu^2, \nu^1 + \nu^2)$. Hence

$$W_2^2(\mu^1 + \mu^2, \nu^1 + \nu^2) = \int_{\Omega \times \Omega} |x - y|^2 \mathrm{d}\gamma \leq \int_{\Omega \times \Omega} |x - y|^2 \mathrm{d}\tilde{\gamma}$$
$$= \int_{\Omega \times \Omega} |x - y|^2 \mathrm{d}\gamma^1 + \int_{\Omega \times \Omega} |x - y|^2 \mathrm{d}\gamma^2$$
$$\leq W_2^2(\mu^1, \nu^1) + W_2^2(\mu^2, \nu^2).$$

Therefore, inequality (A.2) follows. $\square$

APPENDIX B. A REFINED AUBIN-LIONS LEMMA

In [40] the authors present the following version of the classical Aubin-Lions lemma (see [4]):

Theorem B.1. [40, Theorem 2] *Let B be a Banach space and $\mathcal{U}$ be a family of measurable B-valued function. Let us suppose that there exist a normal coercive integrand $\mathfrak{F} : (0, T) \times B \to [0, +\infty]$, meaning that*

(1) *$\mathfrak{F}$ is $\mathscr{B}(0, T) \otimes \mathscr{B}(B)$-measurable, where $\mathscr{B}(0, T)$ and $\mathscr{B}(B)$ denote the σ-algebgras of the Lebesgue measurable subsets of $(0, T)$ and of the Borel subsets of B respectively;*
(2) *the maps $v \mapsto \mathfrak{F}_t(v) := \mathfrak{F}(t, v)$ are l.s.c. for a.e. $t \in (0, T)$;*
(3) *$\{v \in B : \mathfrak{F}_t(v) \leq c\}$ are compact for any $c \geq 0$ and for a.e. $t \in (0, T)$,*

and a l.s.c. map $g : B \times B \to [0, +\infty]$ with the property

$$[u, v \in D(\mathfrak{F}_t), \ g(u, v) = 0] \Rightarrow u = w, \ \text{for a.e. } t \in (0, T).$$

If

$$\sup_{u \in \mathcal{U}} \int_0^T \mathfrak{F}(t, u(t)) \mathrm{d}t < +\infty \ \ and \ \ \lim_{h \downarrow 0} \sup_{u \in \mathcal{U}} \int_0^{T-h} g(u(t+h), u(t)) \mathrm{d}t = 0,$$

then $\mathcal{U}$ is relatively compact in $\mathscr{M}(0, T; B)$.

Conflict of Interest – The authors declare that they have no conflict of interest.

REFERENCES

[1] D. ALEXANDER, I. KIM, Y. YAO, Quasi-static evolution and congested crowd transport, *Nonlinearity*, 27 (2014), No. 4, 823–858.
[2] L. AMBROSIO, Movimenti minimizzanti, *Rend. Accad. Naz. Sci. XL Mem. Mat. Sci. Fis. Nat.* 113 (1995), 191–246.
[3] L. AMBROSIO, N. GIGLI, G. SAVARÉ, *Gradient flows in metric spaces and in the space of probability measures,* Lectures in Mathematics ETH Zürich, Birkhäuser Verlag, Basel, (2008).
[4] J.-P. AUBIN, Un théorème de compacité, (French), *C. R. Acad. Sci. Paris*, 256 (1963), 5042–5044.
[5] J. BERENDSEN, M. BURGER, J.-F. PIETSCHMANN, On a cross-diffusion model for multiple species with nonlocal inter-action and size exclusion, *Nonlinear Anal.* (2017), to appear.
[6] M. BERTSCH, M.E. GURTIN, D. HILHORST, L.A. PELETIER, On interacting populations that disperse to avoid crowding: preservation of segregation, *J. Math. Biol.*, 23 (1985), No. 1, 1–13.
[7] M. BERTSCH, M.E. GURTIN, D. HILHORST, On interacting populations that disperse to avoid crowding: the case of equal dispersal velocities, *Nonlinear Anal.* 11 (1987), no. 4, 493–499.
[8] M. BURGER, M. DI FRANCESCO, J.-F. PIETSCHMANN, B. SCHLAKE, Nonlinear cross-diffusion with size exclusion, *SIAM J. Math. Anal.* 42 (2010), no. 6, 2842–2871.
[9] M. BURGER, M. DI FRANCESCO, S. FAGIOLI, A. STEVENS, Sorting Phenomena in a Mathematical Model For Two Mutually Attracting/Repelling Species, (2017) *preprint* https://arxiv.org/abs/1704.04179.
[10] M. BRUNA, M. BURGER, H. RANETBAUER, M.-T. WOLFRAM, Cross-Diffusion Systems with Excluded-Volume Effects and Asymptotic Gradient Flow Structures, *J. Nonlinear Sci.* (2017) 27, 687–719.
[11] G. BUTTAZZO, F. SANTAMBROGIO, A model for the optimal planning of an urban area, *SIAM J. Math. Anal.*, 37 (2005), no. 2, 514–530.
[12] C. CANCÈS, T. GALLOUËT, L. MONSAINGEON, The gradient flow structure for incompressible immiscible two-phase flows in porous media, *C. R. Math. Acad. Sci. Paris* 353 (2015), no. 11, 985–989.
[13] C. CANCÈS, T. GALLOUËT, L. MONSAINGEON, Incompressible immiscible multiphase flows in porous media: a variational approach, (2016), *preprint* https://arxiv.org/abs/1607.04009.
[14] G. CARLIER, M. LABORDE, Remarks on continuity equations with nonlinear diffusion and nonlocal drifts, *J. Math. Anal. Appl.*, 444 (2016), no. 2, 1690–1702.

[15] G. Carlier, M. Laborde, A splitting method for nonlinear diffusions with nonlocal, nonpotential drifts, *Nonlinear Anal.* 150 (2017), 1–18.

[16] J. Dambrine, N. Meunier, B. Maury, A. Roudneff-Chupin, A congestion model for cell migration, *Commun. Pure Appl. Anal.* 11 (2012), no. 1, 243–260.

[17] G. De Philippis, A. Figalli, The Monge-Ampre equation and its link to optimal transportation, *Bull. Amer. Math. Soc. (N.S.)*, 51 (2014), no. 4, 527–580.

[18] M. Di Francesco, S. Fagioli, Measure solutions for non-local interaction PDEs with two species, *Nonlinearity* 26 (2013), no. 10, 2777–2808.

[19] M. Di Francesco, D. Matthes, Curves of steepest descent are entropy solutions for a class of degenerate convection-diffusion equations, *Calc. Var. Partial Differential Equations*, 50 (2014), No. 1-2, 199–230.

[20] J. Escher, Ph. Laurençot, B.-V. Matioc, Existence and stability of weak solutions for a degenerate parabolic system modelling two-phase flows in porous media, *Ann. Inst. H. Poincar Anal. Non Linéaire* 28 (2011), no. 4, 583–598.

[21] R. Jordan, D. Kinderlehrer, F. Otto, The variational formulation of the Fokker-Planck equation, *SIAM J. Math. Anal.*, 29 (1998), no. 1, 1–17.

[22] A. Jüngel, N. Zamponi, A cross-diffusion system derived from a Fokker-Planck equation with partial averaging, *Z. Angew. Math. Phys.* 68 (2017), no. 1, Art. 28, 15 pp.

[23] I. C. Kim, Uniqueness and existence results on the Hele-Shaw and the Stefan problems, *Arch. Rational Mech. Anal.*, 168 (2003), no. 4, 299–328.

[24] I. Kim, N. Pozar, Porous medium equation to Hele-Shaw flow with general initial density, to appear in *Transactions of AMS*, https://arxiv.org/abs/1509.06287.

[25] M. Laborde, *Systèmes de particules en interaction, approche par flot de gradient dans l'espace de Wasserstein*, PhD thesis, Université Paris-Dauphine, (2016).

[26] M. Laborde, On some non linear evolution systems which are perturbations of Wasserstein gradient flows, in: *Topological Optimization and Optimal Transport*, in: *Radon Ser. Comput. Appl. Math.*, (2017), in press.

[27] Ph. Laurençot, B.-V. Matioc, A gradient flow approach to a thin film approximation of the Muskat problem, *Calc. Var. Partial Differential Equations* 47 (2013), no. 1-2, 319–341.

[28] T. Lorenzi, A. Lorz, B. Perthame, On interfaces between cell populations with different mobilities, *Kinet. Relat. Models*, 10 (2017), No. 1, 299–311.

[29] D. Matthes, R. McCann, G. Savaré, A family of nonlinear fourth order equations of gradient flow type, *Comm. PDE*, 34 (2009), No. 10-12, 1352–1397.

[30] B. Maury, A. Roudneff-Chupin, F. Santambrogio, A macroscopic crowd motion model of gradient flow type, *Math. Models and Meth. in Appl. Sci.*, 20 (2010), No. 10, 1787–1821.

[31] B. Maury, A. Roudneff-Chupin, F. Santambrogio, Congestion-driven dendritic growth, *Discrete Contin. Dyn. Syst.*, 34 (2014), no. 4, 1575–1604.

[32] B. Maury, A. Roudneff-Chupin, F. Santambrogio, J. Venel, Handling congestion in crowd motion modeling, *Netw. Heterog. Media*, 6 (2011), No. 3, 485–519.

[33] B. Maury, J. Venel, Handling of contacts in crowd motion simulations, *Traffic and Granular Flow,* Springer (2007).

[34] R.J. McCann, A convexity principle for interacting gases, *Adv. Math.*, 128 (1997), No. 1, 153–179.

[35] A. Mellet, B. Perthame, F. Quirós A Hele-Shaw problem for tumor growth, *preprint* https://arxiv.org/abs/1512.06995.

[36] A.R. Mészáros, F. Santambrogio, Advection-diffusion equations with density constraints, *Analysis & PDE*, Vol. 9 (2016), No. 3, 615–644.

[37] F. Otto, The geometry of dissipative evolution equations: the porous medium equation, *Comm. PDE*, 26 (2001), No. 1-2, 101–174.

[38] F. Otto, Evolution of Microstructure in Unstable Porous Media Flow: A Relaxational Approach, *Comm. Pure Appl. Math*, 52 (1999), 873–915.

[39] B. Perthame, F. Quirós, J.L. Vázquez, The Hele-Shaw asymptotics for mechanical models of tumor growth, *Arch. Rational Mech. Anal.*, 212 (2014), 93–127.

[40] R. Rossi, G. Savaré, Tightness, integral equicontinuity and compactness for evolution problems in Banach spaces, *Ann. Sc. Norm. Super. Pisa Cl. Sci. (5),* 2 (2003), no. 2, 395–431.

[41] E. Sandier, S. Serfaty, Gamma-convergence of gradient flows with applications to Ginzburg-Landau, *Comm. Pure Appl. Math.*, 57 (2004), no. 12, 1627–1672.

[42] F. Santambrogio, {Euclidean, metric, and Wasserstein} gradient flows: an overview, *Bull. Math. Sci.*, (2017) 7, 87–154.

[43] F. Santambrogio, *Optimal Transport for Applied Mathematicians,* Birkäuser, (2015).

[44] C. Villani *Topics in Optimal Transportation.* Graduate Studies in Mathematics, AMS, (2003).

[45] N. Zamponi, A. Jüngel, Analysis of degenerate cross-diffusion population models with volume filling, *Ann. Inst. H. Poincar Anal. Non Linéaire* 34 (2017), no. 1, 1–29.

Department of Mathematics, UCLA, California, USA
E-mail address: ikim@math.ucla.edu

Department of Mathematics, UCLA, California, USA
E-mail address: alpar@math.ucla.edu