

# WEAK SOLUTIONS TO THE MUSKAT PROBLEM WITH SURFACE TENSION VIA OPTIMAL TRANSPORT

MATT JACOBS, INWON KIM, AND ALPÁR R. MÉSZÁROS

ABSTRACT. Inspired by recent works on the threshold dynamics scheme for multi-phase mean curvature flow (by Esedoğlu-Otto and Laux-Otto), we introduce a novel framework to approximate solutions of the Muskat problem with surface tension. Our approach is based on interpreting the Muskat problem as a gradient flow in a product Wasserstein space. This perspective allows us to construct weak solutions via a minimizing movements scheme. Rather than working directly with the singular surface tension force, we instead relax the perimeter functional with the heat content energy approximation of Esedoğlu-Otto. The heat content energy allows us to show the convergence of the associated minimizing movement scheme in the Wasserstein space, and makes the scheme far more tractable for numerical simulations. Under a typical energy convergence assumption, we show that our scheme converges to weak solutions of the Muskat problem with surface tension. We then conclude the paper with a discussion of equilibrium configurations.

## 1. INTRODUCTION

The Muskat problem was first introduced by Morris Muskat [27] as a model for the flow of two immiscible fluids through a porous medium. Since its introduction, the Muskat problem has received sustained attention in a variety of fields. The Muskat problem is used to model flows in oil reservoirs (water is injected into the oil well to drive oil extraction), and in hydrology to model flows of groundwater through aquifers.

In this paper we are interested in obtaining the global existence of weak solutions for the Muskat problem with surface tension, based on its gradient flow structure. We begin by introducing a variational formulation of the problem, which will motivate our subsequent analysis. The fluid evolution can be written as Darcy's law

$$(1) \quad v_i + b_i^{-1} \nabla \delta_{\rho_i} \mathcal{E}(\boldsymbol{\rho}) = 0,$$

coupled with the continuity equation

$$(2) \quad \partial_t \rho_i + \nabla \cdot (\rho_i v_i) = 0,$$

where  $v_i$  is the velocity of phase  $i$ ,  $\boldsymbol{\rho} = (\rho_1, \rho_2)$  is the collection of relative concentrations for each phase,  $\nabla \delta_{\rho_i} E$  denotes the gradient of the first variation of the internal energy with respect to  $\rho_i$ , and  $b_i > 0$  ( $i = 1, 2$ ) denotes constant mobilities. For convenience, throughout the rest of the paper, we will refer to  $\boldsymbol{\rho}$  as a collection of density functions; however, one should note that  $\boldsymbol{\rho}$  only encodes information about the volume occupied by the fluids and nothing about their mass.

The physical setting for our problem is a bounded, convex open domain  $\Omega \subset \mathbb{R}^d$  with smooth boundary. We shall suppose that the two fluids fill the entire domain, and that they are confined to  $\Omega$  for all time. We then take the internal energy to be a sum of three distinct terms:

$$(3) \quad \mathcal{E}(\boldsymbol{\rho}) = \mathcal{E}_p(\boldsymbol{\rho}) + \mathcal{E}_s(\boldsymbol{\rho}) + \Phi(\boldsymbol{\rho}).$$

The first term in the energy, describing incompressibility and containment of the fluids, is given by

$$(4) \quad \mathcal{E}_p(\boldsymbol{\rho}) = \begin{cases} 0, & \text{if } \rho_1(x) + \rho_2(x) = 1 \text{ for a.e. } x \in \Omega \\ +\infty, & \text{otherwise.} \end{cases}$$

The immiscibility of the fluids and the surface tension force arise from the highly non-convex interaction energy

$$(5) \quad \mathcal{E}_s(\boldsymbol{\rho}) = \begin{cases} \frac{\sigma}{2} |D\rho_1|(\Omega) + \frac{\sigma}{2} |D\rho_2|(\Omega), & \text{if } \rho_1, \rho_2 \in BV(\Omega; \{0, 1\}) \text{ and } \rho_1(x)\rho_2(x) = 0 \text{ for a.e. } x \in \Omega \\ +\infty, & \text{otherwise,} \end{cases}$$

where  $|D\rho_i|(\Omega)$  denotes the total variation of  $\rho_i$  in  $\Omega$  and  $\sigma > 0$  is a surface tension constant. Finally,  $\Phi(\boldsymbol{\rho})$  denotes the potential energy of the fluid configuration, i.e.

$$\Phi(\boldsymbol{\rho}) = \int_{\Omega} \Phi_1 d\rho_1 + \int_{\Omega} \Phi_2 d\rho_2,$$

where  $\Phi_1, \Phi_2 : \Omega \rightarrow \mathbb{R}$  are given Lipschitz continuous potentials. A typical example is when one assumes these to be gravitational potentials, i.e.

$$(6) \quad \Phi_i = g_i x \cdot e_d, \quad g_i > 0, \quad i = 1, 2$$

where  $g_i$ 's are proportional to the specific gravity of each fluid.

Although the internal energy is singular, when  $\rho_1$  and  $\rho_2$  are separated by a smooth interface  $\Gamma := \partial\{\rho_1 > 0\} \cap \partial\{\rho_2 > 0\}$ , one can formulate a classical solution to the Muskat problem equations (1-2). In the classical solution, the flow is driven by the pressure variables  $p_i$  for each phase, which are Lagrange multipliers generated by  $\mathcal{E}_p$  above. The continuity equation becomes

$$(MP_1) \quad \partial_t \rho_i - b_i^{-1} \nabla \cdot ((\nabla p_i + \nabla \Phi_i) \rho_i) = 0$$

and the pressure is determined by solving the free boundary problem

$$(MP_2) \quad \begin{cases} -\Delta p_i = \Delta \Phi_i & \text{in } \text{spt}(\rho_i); \\ \partial_n(p_i + \Phi_i) = 0 & \text{on } \partial\Omega; \\ V = b_1^{-1} \partial_n(p_1 + \Phi_1) = b_2^{-1} \partial_n(p_2 + \Phi_2) & \text{on } \Gamma; \\ [p] := (p_1 - p_2) = \frac{\sigma}{2} \kappa & \text{on } \Gamma; \end{cases}$$

where  $\kappa$  denotes the mean curvature of  $\Gamma$ , oriented to be positive when  $\{\rho_2 > 0\}$  is convex at the point, and  $n$  denotes the outer normal along  $\partial\Omega$  and along  $\Gamma$ . Problems like (MP<sub>2</sub>) received a lot of attention in the past decades. Most of the works focus on the zero surface tension model ( $\sigma = 0$ ) and well-posedness of regular solutions with graph property [1, 5, 7, 8, 9]. In the presence of surface tension, the problem has stronger regularity properties in stable settings [30], but still, topological singularities can occur in finite time, for instance when heavier fluid is placed on top of the lighter one [13]. Thus, our aim is to construct global-in-time weak solutions to the Muskat problem (MP<sub>2</sub>), which exist past the formation of singularities.

To construct global-in-time solutions, we exploit the gradient flow structure of the Muskat problem. As noted by Otto in [28, 29], Darcy's law can be approximated by the Euler-Lagrange equation for the minimizing movements scheme (or JKO [17] scheme) with time step size  $\tau > 0$ ,

$$(7) \quad \boldsymbol{\rho}^{n+1} = \arg \min_{\boldsymbol{\rho}} \left\{ \mathcal{E}(\boldsymbol{\rho}) + \sum_{i=1}^2 \frac{b_i}{2\tau} W_2^2(\rho_i, \rho_i^n) \right\}$$

where  $W_2(\rho_i, \rho_i^n)$  denotes the 2-Wasserstein or 2-Monge-Kantorovich distance. In this context, the squared  $W_2$  distance has a physical interpretation as the energy dissipated by friction as the fluids flow through the porous media.

Let us note that Wasserstein gradient flows of energies involving total variation terms have been considered before in the literature, though only in the case of one phase models (see e.g. [24, 4]), and hence with no incompressibility or interaction constraints. As a result, the techniques developed in those papers do not appear to be applicable here — the constrained two phase setting adds many additional difficulties.

Indeed, it is not easy to obtain a complete characterization of the solutions to the minimizing movements problem (7). The interaction energy (5) is sufficiently non-convex that problem (7) is non-convex for *any*  $\tau > 0$ . As a result, one must be careful in using duality to introduce the pressure as a Lagrange multiplier. Furthermore, we are interested in developing a scheme which could be used for numerical implementations. The formulation (7) is poorly suited for numerical methods. Optimizing over the non-convex constraint set  $\{\rho_1, \rho_2 \in BV(\Omega; \{0, 1\}) : \rho_1(x)\rho_2(x) = 0 \text{ a.e.}\}$  is extremely difficult. For these reasons, we instead consider a relaxed version of minimizing movements scheme inspired by [12] and [20].

### Approximation of the perimeter by the Heat Content

In our analysis, we replace the interaction energy in  $\mathcal{E}_s$  in (7) by the *heat content energy*

$$(8) \quad \text{HC}_\varepsilon(\boldsymbol{\rho}) := \sigma \sqrt{\frac{2\pi}{\varepsilon}} \int_{\Omega} (G_\varepsilon * \rho_1)(x) \, d\rho_2(x) = \sigma \sqrt{\frac{2\pi}{\varepsilon}} \int_{\Omega} \int_{\mathbb{R}^d} G(z) \, d\rho_1(x + \sqrt{\varepsilon}z) \, d\rho_2(x).$$

Here  $G_\varepsilon : \mathbb{R}^d \rightarrow \mathbb{R}$  stands for the standard heat kernel (with mean 0 and variance  $\varepsilon > 0$ ) and the densities  $\rho_i$  are assumed to be defined on all of  $\mathbb{R}^d$  by extending them to zero off of  $\Omega$ .

The heat content energy was first introduced by Esedoğlu and Otto in [12]. It was observed in [12] that the threshold dynamics, a well known numerical scheme for mean curvature motion introduced by Merriman, Bence, and Osher [26], is precisely a minimizing movements scheme for the heat content energy. Esedoğlu and Otto also showed that  $\text{HC}_\varepsilon$   $\Gamma$ -converges (with respect to the  $L^1$  topology) to  $\mathcal{E}_s$  as  $\varepsilon \rightarrow 0$ . Building off of these results, Laux and Otto showed that under an energy convergence assumption, the threshold dynamics scheme produces weak solutions to the multi-phase motion by mean curvature in the limit  $\varepsilon \rightarrow 0$  [20].

Our goal is to consider such a framework in the context of the Muskat problem by studying the minimizing movements scheme

$$(9) \quad \boldsymbol{\rho}^{n+1} = \arg \min_{\boldsymbol{\rho}} \left\{ \mathcal{E}_\varepsilon(\boldsymbol{\rho}) + \sum_{i=1}^2 \frac{b_i}{2\tau} W_2^2(\rho_i, \rho_i^n) \right\},$$

where we used the notation

$$\mathcal{E}_\varepsilon(\boldsymbol{\rho}) := \mathcal{E}_p(\boldsymbol{\rho}) + \text{HC}_\varepsilon(\boldsymbol{\rho}) + \Phi(\boldsymbol{\rho}).$$

As we alluded above, the scheme (9) has a number of numerical advantages over (7). Unlike  $\mathcal{E}_s$  which is neither convex nor concave, the heat content is a strictly concave functional of the densities. This concavity can be exploited to simplify numerical implementations, along the same lines as the linearization trick noted in [12]. Furthermore, the concavity of the heat content is controlled by  $\varepsilon$ , thus, if one chooses  $\tau$  sufficiently small the overall problem becomes displacement convex (c.f. Lemma 2.1). Indeed, displacement convexity is crucial to guarantee that the problem can be solved efficiently with standard numerical techniques.

Although the heat content in principle allows mixing of the phases, we shall show that the discrete in time solutions constructed by the JKO scheme always stay unmixed with a sharp interface between the phases for all time (see Proposition 2.3 below). This phenomenon is due to the fact that  $\text{HC}_\varepsilon$  behaves like a strictly concave functional (see Lemma 2.2). Thus, one retains the essential properties of the Muskat problem evolution.

In the context of the Muskat problem, the heat content also has a natural physical interpretation. In a discrete statistical mechanics model with  $N$  particles, surface tension can be seen to arise from short range interactions between particles in different phases (cf. [16]). Typically, the discrete surface tension takes the form

$$\frac{1}{N^2} \sum_{i \in P_1, j \in P_2} V(|x_i - x_j|) = \frac{1}{N^2} \sum_{i \in P_1, j \in P_2} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} V(|x - x'|) \delta(x - x_i) \delta(x' - x_j)$$

where  $P_r$  is the particle index in each phase  $r$ ,  $V$  is some decreasing function, and  $x_i$  is the location of particle  $i$  ([16]). By taking the limit  $N \rightarrow \infty$  in above formula, one obtains an analogue of the heat content energy where the kernel  $G_\varepsilon$  is replaced with  $V$ . Here it is worth noting that we choose to work with the heat kernel for computational convenience, indeed a different choice of kernel may be more physically relevant.

Finally, let us also emphasize that the heat content approach can be naturally extended to the multi-phase Muskat Problem evolution (with any number of phases), without incurring any additional difficulties. This includes scenarios where the surface tension force depends on the phases that are interacting [12]. Although we do not pursue the multiphase case further in this paper, this presents an interesting direction for future work.

### Statement of our main results

From the relaxed minimizing movements scheme (9), we obtain a sequence of discrete in time approximations to the Muskat flow. When we take the time step  $\tau$  and the heat content approximation parameter  $\varepsilon$  to zero together, we hope to recover weak solutions to the Muskat problem. Our main results show that this is indeed the case under the assumption that there is no loss of perimeter when passing from discrete to continuous solutions. For the precise statements of the convergence results we refer to Theorem 3.1 and Theorem 3.2.

**Theorem 1.1.** *Let  $\varepsilon, \tau > 0$  be fixed and let  $(\rho_1^\varepsilon, \rho_2^\varepsilon)$  be the discrete in time interpolations of the densities obtained from the minimizing movements scheme (9). Let moreover  $p^\varepsilon$  stand for the discrete in time interpolations between the scalar pressure fields, obtained as Lagrange multipliers associated to the incompressibility constraint in (9). Then*

- $(\rho_1^\varepsilon, \rho_2^\varepsilon)$  are such that  $\rho_1^\varepsilon \rho_2^\varepsilon = 0$  and  $\rho_1^\varepsilon + \rho_2^\varepsilon = 1$  a.e. in  $[0, T] \times \Omega$ .
- There exists  $\rho_i \in L^1([0, T]; BV(\Omega; \{0, 1\})) \cap AC^2([0, T]; \mathcal{P}(\Omega))$  ( $i = 1, 2$ ) such that as  $\max\{\varepsilon, \tau\} \downarrow 0$  and along a subsequence  $(\rho_1^\varepsilon, \rho_2^\varepsilon) \rightarrow (\rho_1, \rho_2)$  strongly in  $L^1([0, T] \times \Omega) \times L^1([0, T] \times \Omega)$ . Moreover,  $\rho_1, \rho_2 \in L^1([0, T]; BV(\Omega))$  and they are characteristic functions that sum up to one, i.e.

$$(10) \quad \rho_1(t, \cdot) = \chi_{A_t \cap \Omega} \text{ and } \rho_2(t, \cdot) = \chi_{\Omega \setminus A_t},$$

for a measurable family of sets  $(A_t)_{t \in [0, T]}$  which are of finite perimeter.

- There exists a scalar pressure field  $p \in L^2([0, T]; (C^{0, \alpha}(\Omega))^*)$  such that along a subsequence  $\nabla p^\varepsilon \xrightarrow{*} \nabla p$  weakly- $*$  in  $L^2([0, T]; (C^1(\Omega))^*)$  as  $\max\{\varepsilon, \tau\} \downarrow 0$ .
- Finally, under the assumption of energy convergence

$$(EC) \quad \int_0^T \text{HC}_\varepsilon(\rho_1^\varepsilon, \rho_2^\varepsilon) dt \rightarrow \int_0^T \left( \sum_i \frac{\sigma}{2} \int_\Omega |D\rho_i| \right) dx dt,$$

$(\rho_i, v_i, p)$  with  $i = 1, 2$  solves, in the weak sense (see Definition 3.1), the problem (MP<sub>1</sub>)-(MP<sub>2</sub>). Here formally  $v_i$  corresponds to  $-b_i^{-1}(\nabla p + \nabla \Phi_i)$ .

It is challenging to study qualitative or geometric properties of our solutions. We will illustrate heuristically in Section 4 that, even for global minimizers of the energy, there are diverse possibilities depending on the values of the specific gravity and volumes of the two phases.

### Remarks on our results

- (1) Let us underline the fact that our discrete-time scheme (9) produces minimizers that are characteristic functions of a partition of  $\Omega$ . To prove this fact, we exploit the strict concavity of the heat content along the admissible set. This seems to be an interesting property in its own right, and ensures that numerical implementations of the scheme maintain a sharp interface at every time step.
- (2) From the point of view of Wasserstein gradient flows, an interesting remark on our results is that while one cannot expect strong compactness for the interpolated densities  $(\rho_1^\varepsilon, \rho_2^\varepsilon)$  in the case when  $\varepsilon > 0$  is fixed and  $\tau \downarrow 0$ , when sending both parameters to 0 in the same time, we regain the strong compactness. This phenomenon is mainly due to the fact that in the limit as  $\max\{\tau, \varepsilon\} \downarrow 0$  we recover total variation estimates on the densities. This compactness is obtained via a careful Aubin-Lions type argument.
- (3) The energy convergence assumption (EC) is rather natural and the same as the ones given in [20] and [21]. This assumption ensures that there is no sudden loss of boundary between phases in the limit  $\varepsilon \rightarrow 0$ . If there is a loss of interface then one cannot obtain weak solutions. Indeed, if two components of the support of  $\rho_1$  merge into each other and remove a sizable part of its boundary, one can expect a discontinuous change of the pressure term  $p$  in the entire domain  $\Omega$ , creating an inconsistency between the discrete and limiting evolutions.

Replacing the assumption with a direct argument has been discussed for mean curvature flows [36, 10]. Unfortunately, these results rely strongly on certain properties of the mean curvature flow (especially the comparison principle), which do not hold for fourth order equations like the Muskat problem.

There are also results in the literature [22, 31] which eliminate the assumption for fourth order curvature driven flows (specifically the Stefan problem and the Mullins-Sekerka flow respectively).

However, the solutions constructed in [22] are discontinuous in time (the interface may experience sudden jumps in time) and the formulation in [31] only keeps track of the regular part of the interface. Let us also note that the Stefan problem and the Mullins-Sekerka flow are not similar to the Muskat problem. In particular, the jump condition for the pressure across the interface is different for the Muskat problem, which leads to qualitatively different behavior.

### Paper summary

The rest of the paper is organized as follows. In Section 2, we derive the basic properties of the minimizing movements scheme (9) and construct our discrete-time quantities. We begin by showing that solutions to the minimization problem are characteristic functions at every time step of the discrete scheme. We then derive the existence of pressure as a Lagrange multiplier for the incompressibility constraint and obtain the Euler-Lagrange equation for the minimization problem. Our derivation of these equations were inspired by previous results from [11, 25, 19, 18]. In particular, the definition of the discrete in time pressure variable is very much inspired from [25].

In Section 3, we take  $\tau$  and  $\varepsilon$  to zero together to obtain weak solutions to the Muskat problem, under the assumption that the internal energy of the discrete solutions converges to the internal energy of the limiting solutions. The main task in this section amounts to showing that that one can pass to the limit in the Euler-Lagrange equation obtained in Section 2. This can be done using the standard theory for Wasserstein gradient flows if  $\varepsilon$  is held fixed. However, the joint limit,  $\tau, \varepsilon_\tau \rightarrow 0$ , requires an adaptation of the arguments of [20] to the case of Wasserstein gradient flows. In particular, we develop a careful Aubin-Lions type argument, by applying a version of flow-exchange technique introduced in [24]. Thus, while we largely follow ideas from [20] in our analysis, in some cases we give simplified proofs of the results we need, and in others we develop completely new arguments to conclude our results.

Finally, in Section 4, we conclude the paper with a discussion of the global minimizers of the approximated internal energy associated to the Muskat problem. While this discussion remains at the heuristic level, we conjecture the validity of the geometric description of the equilibrium configurations.

## 2. THE WASSERSTEIN MINIMIZING MOVEMENTS SCHEME FOR THE HEAT CONTENT

**2.1. Some preliminary results.** Recall that the setting for our problem is a smooth convex domain  $\Omega \subset \mathbb{R}^d$  with  $\mathcal{L}^d(\Omega) = 2$ . By  $\mathcal{P}(\Omega)$  we denote the space of Borel probability measures on  $\mathbb{R}^d$  supported on  $\bar{\Omega}$ .  $\mathcal{P}^{\text{ac}}(\Omega)$  stands for the elements of  $\mathcal{P}(\Omega)$  that are absolutely continuous with respect to  $\mathcal{L}^d \llcorner \Omega$ .

Let  $\Phi_1, \Phi_2 : \Omega \rightarrow \mathbb{R}$  be given Lipschitz potentials and let us recall the definition of the potential energy  $\Phi : \mathcal{P}(\Omega) \times \mathcal{P}(\Omega) \rightarrow \mathbb{R}$ , given as

$$\Phi(\rho) := \int_{\Omega} \Phi_1 d\rho_1 + \int_{\Omega} \Phi_2 d\rho_2.$$

Let  $\varepsilon > 0$ . We consider the heat content  $\text{HC}_\varepsilon : \mathcal{P}(\Omega) \times \mathcal{P}(\Omega) \rightarrow \mathbb{R}$  defined in (8), using the standard heat kernel  $G_\varepsilon : \mathbb{R}^d \rightarrow \mathbb{R}$ , i.e.

$$G_\varepsilon(x) = \frac{1}{(4\pi\varepsilon)^{d/2}} e^{-\frac{|x|^2}{4\varepsilon}}.$$

We also use the notations  $K_\varepsilon, G : \mathbb{R}^d \rightarrow \mathbb{R}$  to denote  $K_\varepsilon(x) = \sigma \sqrt{\frac{2\pi}{\varepsilon}} G_\varepsilon(x)$  and  $G(x) = G_1(x)$ . We have the following preliminary results.

**Lemma 2.1.** *Let  $\text{HC}_\varepsilon$  be defined as in (8). We have the following properties.*

- (1)  $\text{HC}_\varepsilon$  is bounded from below and continuous w.r.t. the weak- $\star$  convergence on  $\mathcal{P}(\Omega) \times \mathcal{P}(\Omega)$ .
- (2)  $\text{HC}_\varepsilon$  displacement  $\lambda$ -convex on  $\mathcal{P}(\Omega) \times \mathcal{P}(\Omega)$ , with  $\lambda = -\frac{\sigma\sqrt{2\pi}}{(4\pi)^{d/2}} \frac{1}{\varepsilon^{(d+3)/2}}$ .

*Proof.* (1) Is immediate by the definition of  $\text{HC}_\varepsilon$ .

To show (2), it is enough to show that the function  $z \mapsto K_\varepsilon(z)$  is  $\lambda$ -convex in the classical sense for some  $\lambda \in \mathbb{R}$ . We have

$$D^2 K_\varepsilon(x) = \frac{\sigma\sqrt{2\pi}}{2(4\pi)^{d/2}} \frac{1}{\varepsilon^{(d+3)/2}} e^{-\frac{|x|^2}{4\varepsilon}} \left( \frac{1}{2\varepsilon} x \otimes x - I_d \right).$$

Since the matrix  $x \otimes x$  is nonnegative definite for any  $x \in \mathbb{R}^d$ , setting  $\lambda = -\frac{1}{2(4\pi)^{d/2}} \frac{\sigma\sqrt{2\pi}}{\varepsilon^{(d+3)/2}}$ , we have that the matrix  $D^2K_\varepsilon(x) - \lambda I_d$  is nonnegative definite for any  $x \in \mathbb{R}^d$ , which implies in particular that  $K_\varepsilon$  is  $\lambda$ -convex.

We conclude similarly as in [11, Lemma 2.1] the displacement  $2\lambda$ -convexity of  $\text{HC}_\varepsilon$  on  $\mathcal{P}(\Omega) \times \mathcal{P}(\Omega)$ .  $\square$

**Lemma 2.2.** *If  $P \subset \mathcal{P}(\Omega) \times \mathcal{P}(\Omega)$  denotes the set of pairs  $(\rho_1, \rho_2)$  such that  $\rho_1 + \rho_2 = 1$  a.e. in  $\Omega$ , then the heat content is strictly concave along line segments in  $P$ .*

*Proof.* For any pair  $(\rho_1, \rho_2) \in P$  we may write  $\rho_2(x) = 1 - \rho_1(x)$ . Thus, extending the densities by 0 outside of  $\Omega$ , we have

$$\text{HC}_\varepsilon(\rho_1, \rho_2) = \sigma \sqrt{\frac{2\pi}{\varepsilon}} \int_{\mathbb{R}^d} (G_\varepsilon * \rho_1)(x)(1 - \rho_1(x)) \, dx = \sigma \sqrt{\frac{2\pi}{\varepsilon}} \int_{\mathbb{R}^d} (G_\varepsilon * \rho_1)(x) \, dx - \sigma \sqrt{\frac{2\pi}{\varepsilon}} \int_{\mathbb{R}^d} (G_\varepsilon * \rho_1)(x)\rho_1(x) \, dx$$

Ignoring the constant multiples, both terms can be expressed conveniently in the Fourier domain:

$$\hat{\rho}_1(0) - \int_{\mathbb{R}^d} \hat{G}(\xi\sqrt{\varepsilon}) |\hat{\rho}_1(\xi)|^2 \, d\xi.$$

Now the strict concavity follows immediately as  $\hat{G}(\xi\sqrt{\varepsilon}) > 0$  for all  $\xi \in \mathbb{R}^d$ .  $\square$

**2.2. The minimizing movements scheme.** Now we are ready to discuss the minimizing movements scheme. Our first result confirms the existence of minimizers, and shows that any minimizing configuration  $\rho = (\rho_1, \rho_2)$  is a completely unmixed partition of the domain. As we will see, the phases stay unmixed thanks to the concavity of the heat content.

**Proposition 2.3.** *Suppose that  $\rho_1^n, \rho_2^n \in \mathcal{P}(\Omega)$  and let  $\tau > 0$ . Then the set of minimizers of the problem*

$$(11) \quad \inf \left\{ \text{HC}_\varepsilon(\rho) + \Phi(\rho) + \frac{b_1}{2\tau} W_2^2(\rho_1, \rho_1^n) + \frac{b_2}{2\tau} W_2^2(\rho_2, \rho_2^n) : \rho_1, \rho_2 \in \mathcal{P}^{\text{ac}}(\Omega), \rho_1 + \rho_2 = 1 \text{ a.e.} \right\}$$

*is non-empty and any solution  $(\rho_1^*, \rho_2^*) \in \mathcal{P}(\Omega) \times \mathcal{P}(\Omega)$  is the characteristic function of a partition of  $\Omega$ .*

*Proof.* The existence of a solution of the optimization problem is an easy consequence of the weak lower semicontinuity of the objective functional and the (weak- $\star$ ) compactness of  $\mathcal{P}(\Omega) \times \mathcal{P}(\Omega)$ . Let us remark that the constraint  $\rho_1 + \rho_2 = 1$  a.e. is closed under weak convergence, since  $\int_\Omega \rho_1 \, dx + \int_\Omega \rho_2 \, dx = \mathcal{L}^d(\Omega)$ .

To show that an arbitrary solution  $(\rho_1^*, \rho_2^*)$  is the characteristic functions of a partition of  $\Omega$ , let us rewrite equivalently the minimization problem in terms of transport plans  $\pi_i \in \mathcal{P}(\Omega \times \Omega)$ . Recall that  $\pi_i$  is a plan between  $\rho_i^n$  and  $\rho_i$ , whenever

$$\int_{\Omega \times \Omega} \varphi(x) \, d\pi_i(x, y) = \int_\Omega \varphi(x) \, d\rho_i^n(x) \quad \text{and} \quad \int_{\Omega \times \Omega} \psi(y) \, d\pi_i(x, y) = \int_\Omega \psi(y) \, d\rho_i(y),$$

for any  $\varphi, \psi \in C(\Omega)$ . Moreover, since we are always working with measures  $\rho_i^n, \rho_i$  that are absolutely continuous w.r.t.  $\mathcal{L}^d \llcorner \Omega$ , we can restrict our search to plans that are absolutely continuous w.r.t.  $\mathcal{L}^d \llcorner \Omega \otimes \mathcal{L}^d \llcorner \Omega$ . For such measures we will have

$$(12) \quad \int_{\Omega \times \Omega} \varphi(x) \pi_i(x, y) \, d(x, y) = \int_\Omega \varphi(x) \rho_i^n(x) \, dx \quad \text{and} \quad \int_{\Omega \times \Omega} \psi(y) \pi_i(x, y) \, d(x, y) = \int_\Omega \psi(y) \rho_i(y) \, dy,$$

for any  $\varphi, \psi \in C(\Omega)$ . Also, in this case as a consequence of Fubini's and the disintegration theorem, we have that both

$$\Omega \ni y \mapsto \int_\Omega \pi_i(x, y) \, dx \quad \text{and} \quad \Omega \ni x \mapsto \int_\Omega \pi_i(x, y) \, dy$$

belong to  $L^1(\Omega)$  and in particular, as a corollary of (12)

$$\rho_i^n(x) = \int_\Omega \pi_i(x, y) \, dy, \quad \text{for a.e. } x \in \Omega \quad \text{and} \quad \rho_i(y) = \int_\Omega \pi_i(x, y) \, dx, \quad \text{for a.e. } y \in \Omega.$$

Thus, we aim to solve

$$(13) \quad \inf \left\{ \mathcal{F}(\pi_1, \pi_2) + \mathcal{G}(\pi_1, \pi_2) + \sum_{i=1}^2 \int_{\Omega} \int_{\Omega} \frac{b_i}{2\tau} |x-y|^2 \pi_i(x, y) \, dx \, dy \right\} =: \inf \mathcal{S}(\pi_1, \pi_2)$$

subject to  $\int_{\Omega} \pi_i(x, y) \, dy = \rho_i^n(x)$  and  $\sum_{i=1}^2 \int_{\Omega} \pi_i(x, y) \, dx = 1$  a.e.,

where we denote

$$\mathcal{G}(\pi_1, \pi_2) := \int_{\Omega} \int_{\Omega} \sum_{i=1}^2 \Phi_i(y) \pi_i(x, y) \, dx \, dy$$

and we define

$$\begin{aligned} \mathcal{F}(\pi_1, \pi_2) &:= \sigma \sqrt{\frac{2\pi}{\varepsilon}} \int_{\Omega} \int_{\mathbb{R}^d} \int_{\Omega} G(z) \pi_1(x_1, x + \sqrt{\varepsilon}z) \, dx_1 \, dz \, dx \\ &\quad - \sigma \sqrt{\frac{2\pi}{\varepsilon}} \int_{\Omega} \int_{\mathbb{R}^d} \int_{\Omega} \int_{\Omega} G(z) \pi_1(x_1, x + \sqrt{\varepsilon}z) \pi_1(x_2, x) \, dx_1 \, dx_2 \, dz \, dx \end{aligned}$$

and

$$\mathcal{S}(\pi_1, \pi_2) := \mathcal{F}(\pi_1, \pi_2) + \mathcal{G}(\pi_1, \pi_2) + \sum_{i=1}^2 \int_{\Omega} \int_{\Omega} \frac{b_i}{2\tau} |x-y|^2 \pi_i(x, y) \, dx \, dy.$$

The minimization is carried out over a weakly compact set and  $\mathcal{S}$  is weakly lower semicontinuous and bounded below, thus minimizers exist. If  $\boldsymbol{\pi}^* = (\pi_1, \pi_2)$  is a minimizer of (13) then we can construct a minimizer  $\boldsymbol{\rho}^* = (\rho_1^*, \rho_2^*)$  of the original problem by taking  $\rho_i^*(y) = \int_{\Omega} \pi_i^*(x, y) \, dx$ .

Now we consider the properties of minimizers. Clearly,  $\mathcal{S}$  is Gâteaux differentiable on  $L^2(\Omega \times \Omega) \times L^2(\Omega \times \Omega)$ . From Lemma 2.2 it follows that  $\mathcal{S}$  is concave along line segments  $\boldsymbol{\pi} + t\boldsymbol{d}$  ( $t \in (-1, 1)$ ), where  $\boldsymbol{\pi}$  is a feasible point and  $\boldsymbol{d} = (d_1, d_2)$  is a feasible direction at  $\boldsymbol{\pi}$  i.e.

$$(14) \quad \int_{\Omega} d_i(x, y) \, dy = 0 \text{ a.e. } x \in \Omega, \quad \int_{\Omega} \int_{\Omega} d_i(x, y) \, dx \, dy = 0 \quad \text{and} \quad \sum_{i=1}^2 \int_{\Omega} d_i(x, y) \, dx = 0 \text{ a.e. } y \in \Omega,$$

and for some  $\delta > 0$

$$(15) \quad \pi_i(x, y) + td_i(x, y) \geq 0$$

for all  $t \in [0, \delta)$ . Furthermore, it follows from Lemma 2.2 that  $\mathcal{S}$  is strictly concave on line segments  $\boldsymbol{\pi} + t\boldsymbol{d}$ , if for some  $i$  the marginal  $\int_{\Omega} d_i(x, y) \, dx$  is not 0 for almost every  $y$ . Therefore, if  $\boldsymbol{\pi}^* = (\pi_1^*, \pi_2^*)$  is a minimizer and  $\boldsymbol{d}$  is a feasible direction at  $\boldsymbol{\pi}^*$  with at least one non-trivial marginal then

$$\langle \delta_{L^2} \mathcal{S}(\boldsymbol{\pi}^*), \boldsymbol{d} \rangle_{L^2} > 0,$$

where  $\delta_{L^2} \mathcal{S}(\boldsymbol{\pi}^*)$  stands for the standard  $L^2$  variation of  $\mathcal{S}$  at  $\boldsymbol{\pi}^*$  and  $\langle \cdot, \cdot \rangle_{L^2}$  is used for the  $L^2$  inner product.

Let  $\boldsymbol{\pi}$  be a feasible solution and let  $\rho_i(y) = \int_{\Omega} \pi_i(x, y) \, dx$ . Now suppose that there exists  $0 < \alpha < 1$  such that the set  $\Omega_{\alpha} = \{y \in \Omega : \rho_1(y), \rho_2(y) \in (\alpha, 1 - \alpha)\}$  has positive measure. Partition  $\Omega_{\alpha}$  into two sets  $E_1, E_2$  of equal measure. Then there exist measure preserving maps  $T : E_1 \rightarrow E_2$  and  $S : E_2 \rightarrow E_1$  such that  $S \circ T$  and  $T \circ S$  are the identity almost everywhere on their respective domains (for example one may choose  $T = \nabla \psi$  to be the optimal transport map between the densities  $\mathbf{1}_{E_1}$  and  $\mathbf{1}_{E_2}$  and  $S = \nabla \psi^*$ ). Now we construct a feasible direction  $\boldsymbol{d}$  at  $\boldsymbol{\pi}$  as follows. Let  $r(y) = \frac{\rho_1(T(y))}{\rho_2(y)}$  for  $y \in E_1$  and for a.e.  $(x, y) \in \Omega \times \Omega$  define  $d_i$  so that

$$(16) \quad d_1(x, y) = \begin{cases} \pi_1(x, T(y)), & \text{if } y \in E_1, \\ -\pi_1(x, y), & \text{if } y \in E_2, \\ 0, & \text{otherwise,} \end{cases}$$

and

$$(17) \quad d_2(x, y) = \begin{cases} -r(y)\pi_2(x, y), & \text{if } y \in E_1, \\ r(S(y))\pi_2(x, S(y)), & \text{if } y \in E_2, \\ 0, & \text{otherwise.} \end{cases}$$

Since  $\int_{\Omega} d_1(x, y) dx = \rho_1(T(y)) > \alpha$  a.e. on  $E_1$  we see that  $\mathbf{d}$  has a nontrivial marginal.

Let us now check that  $\mathbf{d}$  is feasible, i.e. it satisfies (14) and (15). If  $\beta < \min\{1 - \alpha, \alpha\}$  then  $\pi \pm \beta \mathbf{d}$  is everywhere non-negative. Next, we check that  $\mathbf{d}$  satisfies  $\int_{\Omega} d_i(x, y) dy = 0$  for a.e.  $x \in \Omega$  and  $i = 1, 2$  and  $\sum_{i=1}^2 \int_{\Omega} d_i(x, y) dx = 0$  for a.e.  $y \in \Omega$ . We have

$$\begin{aligned} \int_{\Omega} d_1(x, y) dy &= \int_{E_1} \pi_1(x, T(y)) dy - \int_{E_2} \pi_1(x, y) dy = 0 \\ \int_{\Omega} d_2(x, y) dy &= - \int_{E_1} r(y) \pi_2(x, y) dy + \int_{E_2} r(S(y)) \pi_2(x, S(y)) dy = 0 \end{aligned}$$

where the final equalities hold because both  $T$  and  $S$  are measure preserving between  $E_1$  and  $E_2$  and vice-versa, respectively. Let us notice that these arguments also imply in particular that  $\int_{\Omega} \int_{\Omega} d_i(x, y) dx dy = 0$ ,  $i = 1, 2$ . Now, for a.e.  $y \in E_1$  we have

$$\sum_{i=1}^2 \int_{\Omega} d_i(x, y) dx = \int_{\Omega} \pi_1(x, T(y)) - r(y) \pi_2(x, y) dx = \rho_1(T(y)) - r(y) \rho_2(y) = 0,$$

otherwise if  $y \in E_2$  we have

$$\sum_{i=1}^2 \int_{\Omega} d_i(x, y) dx = \int_{\Omega} -\pi_1(x, y) + r(S(y)) \pi_2(x, S(y)) dx = -\rho_1(y) + r(S(y)) \rho_2(S(y)) = 0.$$

Note that our arguments in fact show that  $-\mathbf{d}$  is also a feasible direction at  $\pi$ .  $\pm \mathbf{d}$  have nontrivial marginals, and it is not possible to have both  $\langle \delta_{L^2} \mathcal{S}(\pi^*), \mathbf{d} \rangle_{L^2} > 0$  and  $-\langle \delta_{L^2} \mathcal{S}(\pi^*), \mathbf{d} \rangle_{L^2} > 0$ . Therefore,  $\pi$  cannot be a minimizer. This allows us to conclude that for any minimizer  $\rho^*$  of the original problem each density  $\rho_i^*$  takes values  $\{0, 1\}$  almost everywhere.  $\square$

**2.3. Optimality conditions and construction of the pressure variables.** In the next Lemma, we give a more complete characterization of the minimizers in terms of certain necessary inequalities. In particular, this is the first place where we see the appearance of the pressure variable, which plays an essential role in all of the subsequent analysis. Note that for convenience we express this result using the notation  $K_{\varepsilon} := \sigma \sqrt{\frac{2\pi}{\varepsilon}} G_{\varepsilon}$ .

**Lemma 2.4.** *Let  $(\rho_1^*, \rho_2^*)$  be an optimizer in (11) and let  $(\rho_1, \rho_2)$  a pair of probability measures such that  $\rho_1 + \rho_2 = 1$  a.e. Then,*

(i) *we have the following optimality condition*

$$(18) \quad \int_{\Omega} [K_{\varepsilon} \star \rho_2^* + \Phi_1 + b_1 \varphi_1 / \tau] (\rho_1 - \rho_1^*) dx + \int_{\Omega} [K_{\varepsilon} \star \rho_1^* + \Phi_2 + b_2 \varphi_2 / \tau] (\rho_2 - \rho_2^*) dx \geq 0,$$

*for a suitable pair of Kantorovich potentials  $(\varphi_1, \varphi_2)$  in the optimal transport of  $\rho_1^*$  onto  $\rho_1^n$  and  $\rho_2^*$  onto  $\rho_2^n$ , respectively.*

(ii) *there exists a Lipschitz continuous function  $p : \Omega \rightarrow \mathbb{R}$  such that for any probability densities  $\rho_1, \rho_2$  with  $\rho_1 + \rho_2 = 1$  a.e. we have*

$$(19) \quad \int_{\Omega} [K_{\varepsilon} \star \rho_2^* + \Phi_1 + b_1 \varphi_1 / \tau + p] (\rho_1 - \rho_1^*) dx + \int_{\Omega} [K_{\varepsilon} \star \rho_1^* + \Phi_2 + b_2 \varphi_2 / \tau + p] (\rho_2 - \rho_2^*) dx \geq 0,$$

*moreover, we have*

$$(20) \quad \nabla p = -\nabla K_{\varepsilon} \star \rho_2^* - \nabla \Phi_1 - b_1 \nabla \varphi_1 / \tau \quad \text{a.e. in } \text{spt}(\rho_1^*)$$

and

$$\nabla p = -\nabla K_{\varepsilon} \star \rho_1^* - \nabla \Phi_2 - b_2 \nabla \varphi_2 / \tau \quad \text{a.e. in } \text{spt}(\rho_2^*).$$

*Proof.* Let  $(\rho_1, \rho_2)$  be a pair of probability measures such that  $\rho_1 + \rho_2 = 1$  a.e. For  $\delta \in [0, 1]$  let us consider the competitors  $(\rho_1^{\delta}, \rho_2^{\delta}) := (\rho_1^* + \delta(\rho_1 - \rho_1^*), \rho_2^* + \delta(\rho_2 - \rho_2^*))$ , which by construction satisfy the constraint.



By optimality, we have

$$\begin{aligned} & \text{HC}_\varepsilon(\rho_1^\delta, \rho_2^\delta) - \text{HC}_\varepsilon(\rho_1^*, \rho_2^*) + \Phi(\rho_1^\delta, \rho_2^\delta) - \Phi(\rho_1^*, \rho_2^*) \\ & + \frac{1}{2\tau} (W_2^2(\rho_1^\delta, \rho_1^n) - W_2^2(\rho_1^*, \rho_1^n) + W_2^2(\rho_2^\delta, \rho_2^n) - W_2^2(\rho_2^*, \rho_2^n)) \geq 0. \end{aligned}$$

Using the exact same argument as in [25, Lemma 3.1] to develop the Wasserstein part on the one hand and the first variations of  $\text{HC}_\varepsilon$  and  $\Phi$  on the other hand, we find (18).

For (ii) (similarly as in [19, Proposition 4.7]), let us notice first that (18) can be written in the form

$$\int_{\Omega} [K_\varepsilon \star \rho_2^* + \Phi_1 + b_1 \varphi_1 / \tau] h_1 \, dx + \int_{\Omega} [K_\varepsilon \star \rho_1^* + \Phi_2 + b_2 \varphi_2 / \tau] h_2 \, dx \geq 0,$$

where  $(h_1, h_2) \in L^\infty(\Omega) \times L^\infty(\Omega)$  is such that  $\rho_1^* + \delta h_1 + \rho_2^* + \delta h_2 = 1$  and  $\rho_i^* + \delta h_i \in [0, 1]$  for  $i = 1, 2$ . We know from Proposition 2.3 that  $\rho_1^*, \rho_2^*$  forms a partition of  $\Omega$ . Therefore, we must take  $h_1 \leq 0$  on  $\text{spt}(\rho_1^*)$ , and  $h_1 \geq 0$  on  $\text{spt}(\rho_2^*)$  (and vice-versa for  $h_2$ ). To preserve the constraint  $\rho_1^* + \delta h_1 + \rho_2^* + \delta h_2 = 1$  and the mass of each density, we must also take  $h_1 + h_2 = 0$  a.e. and  $\int_{\Omega} h_1 \, dx = \int_{\Omega} h_2 \, dx = 0$ .

Now if we set  $h_2 = -h_1$ , we find that

$$\int_{\Omega} \left[ K_\varepsilon \star \rho_2^* + \Phi_1 + b_1 \varphi_1 / \tau - \left( K_\varepsilon \star \rho_1^* + \Phi_2 + b_2 \varphi_2 / \tau \right) \right] h_1 \, dx \geq 0,$$

for any  $h_1 \in L^\infty(\Omega)$  with  $0$  mean such that  $h_1 \leq 0$  a.e. on  $\text{spt}(\rho_1)$ . This implies that there exist constants  $C_1, C_2 \in \mathbb{R}$  such that

$$\begin{aligned} K_\varepsilon \star \rho_2^* + \Phi_1 + b_1 \varphi_1 / \tau - C_1 &\leq K_\varepsilon \star \rho_1^* + \Phi_2 + b_2 \varphi_2 / \tau - C_2 \quad \text{a.e. } \text{spt}(\rho_1) \\ K_\varepsilon \star \rho_1^* + \Phi_2 + b_2 \varphi_2 / \tau - C_2 &\leq K_\varepsilon \star \rho_2^* + \Phi_1 + b_1 \varphi_1 / \tau - C_1 \quad \text{a.e. } \text{spt}(\rho_2) \end{aligned}$$

From here, we can define the pressure variable as

(21)

$$p := C_1 - K_\varepsilon \star \rho_2^* - \Phi_1 - b_1 \varphi_1 / \tau \quad \text{a.e. on } \text{spt}(\rho_1^*) \quad \text{and} \quad p := C_2 - K_\varepsilon \star \rho_1^* - \Phi_2 - b_2 \varphi_2 / \tau \quad \text{a.e. on } \text{spt}(\rho_2^*).$$

Since the Kantorovich potentials are Lipschitz continuous and the other terms are smooth, we find that  $p$  is Lipschitz continuous (note this regularity may degenerate as  $\tau \downarrow 0$ ). By construction,  $p$  clearly satisfies the inequality in (19), and in particular the value of the l.h.s. is equal to zero. Since the functions under consideration are all Lipschitz continuous, we obtain (20).  $\square$

**2.4. Continuous in time solutions for  $\varepsilon > 0$  fixed.** Now we are ready to begin constructing time interpolations from the discrete scheme. In this section we restrict ourselves to the case where  $\varepsilon$  is held fixed. In this special case, we can use standard arguments from the theory of Wasserstein gradient flows to obtain continuous in time equations in the limit  $\tau \downarrow 0$ . When  $\varepsilon$  is held fixed, we have strong compactness on the pressure term. However, similarly to the models in [18, 19], we will lack strong compactness on the density variables, which would mean that in the limit when  $\tau \downarrow 0$ , only a weaker version of the system will be available (see (27)). Let us also note that later on in Section 3, we will need some of the following pressure estimates when we take  $\varepsilon$  to zero along with  $\tau$ .

Let  $T > 0$  be a given time horizon and  $N \in \mathbb{N}$  and  $\tau > 0$  such that  $N\tau = T$ . Let  $\varepsilon > 0$  be fixed. By now, it is standard how to construct weak solutions to PDEs that have gradient flow structure, using minimizing movement schemes as

(22)

$$\left( \rho_1^{n+1, \tau}, \rho_2^{n+1, \tau} \right) \in \arg \min \left\{ \text{HC}_\varepsilon(\rho_1, \rho_2) + \Phi(\rho_1, \rho_2) + \frac{b_1}{2\tau} W_2^2(\rho_1, \rho_1^n) + \frac{b_2}{2\tau} W_2^2(\rho_2, \rho_2^n) : \rho_1, \rho_2 \in \mathcal{D}^{\text{ac}}(\Omega), \rho_1 + \rho_2 = 1 \text{ a.e.} \right\}.$$

Let  $\rho_i^\tau : [0, T] \rightarrow \mathcal{P}(\Omega)$  be defined as

$$(23) \quad \rho_i^\tau(t) := \rho_i^{n+1}, \quad \text{if } t \in [n\tau, (n+1)\tau).$$

We define the corresponding velocities, pressures and momentum variables as

$$(24) \quad \begin{aligned} v_i^\tau(t, \cdot) &:= \frac{\nabla \varphi_i^{n+1}}{\tau}, \quad \text{if } t \in [n\tau, (n+1)\tau), \\ p_i^\tau(t, \cdot) &:= p_i^{n+1}, \quad \text{if } t \in [n\tau, (n+1)\tau), \\ E_i^\tau(t, \cdot) &:= v_i^\tau(t, \cdot) \rho_i^\tau, \quad \text{if } t \in [n\tau, (n+1)\tau), \end{aligned}$$

where  $\varphi_i^{n+1}$  and  $\rho_i^{n+1}$  are defined in Lemma 2.4. Following the same steps as in [18, 25], the analysis boils down to obtain a sufficient amount of uniform (in  $\tau$ ) estimates and compactness for the previously obtained functions, then pass to the limit  $\tau \downarrow 0$ . Also, as additional tools we construct the corresponding continuous in time (geodesic) interpolations between the densities and the corresponding velocities and momentum variables,  $(\tilde{\rho}_i^\tau, \tilde{v}_i^\tau, \tilde{E}_i^\tau)$ . We refer to [18, Section 3.1] (see also [25, 33]) for the precise construction. In particular, as a consequence of this construction, we have

$$\partial_t \tilde{\rho}_i^\tau + \nabla \cdot \tilde{E}_i^\tau = 0,$$

in the sense of distribution on  $(0, T) \times \Omega$ .

Based on the same techniques as in [25, 18], it is easy to obtain the following estimates.

**Lemma 2.5.** *Let  $(\rho_i^\tau, v_i^\tau, E_i^\tau)$  and  $(\tilde{\rho}_i^\tau, \tilde{v}_i^\tau, \tilde{E}_i^\tau)$  be the previously constructed piecewise constant and continuous in time interpolations, respectively. Then there exists  $C > 0$  independent of  $\tau > 0$  and depending only on  $\text{HC}_\varepsilon(\rho_{1,0}, \rho_{2,0})$  such that*

- (i)  $W_2(\rho_i^\tau(t), \rho_i^\tau(s)) \leq C\sqrt{t-s+\tau}$  and  $W_2(\tilde{\rho}_i^\tau(t), \tilde{\rho}_i^\tau(s)) \leq C\sqrt{t-s}$  for any  $0 \leq s \leq t \leq T$ . Moreover, up to passing to subsequences  $(\rho_i^\tau)_{\tau>0}$  and  $(\tilde{\rho}_i^\tau)_{\tau>0}$  converge (uniformly with respect to  $W_2$ ) as  $\tau \downarrow 0$  to the same limit.
- (ii)  $(v_i^\tau)_{\tau>0}$  is uniformly bounded in  $L^2([0, T]; L^2_{\rho_i^\tau})$ .
- (iii)  $(E_i^\tau)_{\tau>0}$  and  $(\tilde{E}_i^\tau)_{\tau>0}$  are uniformly bounded in  $\mathcal{M}^d([0, T] \times \Omega)$  and up to passing to subsequences they have the same distributional limits as  $\tau \downarrow 0$ .
- (iv)  $\int_0^T \int_{\mathbb{R}^d} |\nabla G_\varepsilon \star \rho_i^\tau| dx dt \leq CT \text{HC}_\varepsilon(\rho_{1,0}, \rho_{2,0})$ .
- (v)  $\int_0^T \int_\Omega |\rho_i^\tau(t, x + \delta d) - \rho_i^\tau(t, x)| dx dt \leq CT(\delta + \sqrt{\varepsilon}) \text{HC}_\varepsilon(\rho_{1,0}, \rho_{2,0})$  for any  $d \in \mathbb{R}^d$  with  $|d| = 1$  and any  $\delta > 0$ .
- (vi)  $\int_0^T \int_\Omega |\nabla p^\tau|^2 dx dt \leq C \int_0^T \int_\Omega \frac{1}{\varepsilon^2} \left( (G_{2\varepsilon} \star \rho_2^\tau)^2 \rho_1^\tau + (G_{2\varepsilon} \star \rho_1^\tau)^2 \rho_2^\tau \right) dx dt + C$ , and as a consequence,  $(\nabla p^\tau)_{\tau>0}$  is uniformly bounded in  $L^2([0, T] \times \Omega; \mathbb{R}^d)$  and  $(p^\tau - \int_\Omega p^\tau(\cdot, x) dx)_{\tau>0}$  is uniformly bounded in  $L^2([0, T] \times \Omega)$  by a constant of the form  $TC(1 + 1/\varepsilon^2)$ .
- (vii)  $(\nabla p^\tau)_{\tau>0}$  is uniformly bounded in  $L^2([0, T]; (C^1(\Omega))^*)$ , independently of  $\tau$  and  $\varepsilon$ . In particular,  $\nabla p^\tau$  is uniformly bounded in  $\mathcal{D}'((0, T) \times \Omega; \mathbb{R}^d)$  and  $\nabla p^\tau(t, \cdot)$  defines a uniformly bounded distribution of order one, for a.e.  $t \in [0, T]$ .

*Proof.* Let us notice that the proofs of point (i), (ii) and (iii) follow the exact same lines and the proofs of [25, Lemma 3.3] and [18, Lemma 3.3], so we omit them.

From Proposition 2.3 we know that that  $\rho_i^n$ 's are characteristic functions, therefore the proofs of (iv) and (v) follow the exact same lines as the proof of [20, Lemma 2.4.].

Let us give the details on (vi). From the identities (20) from Lemma 2.4, we have that there exists  $C > 0$  (independent of  $\tau > 0$ , which might increase from one inequality to the next) such that

$$\begin{aligned} \int_0^T \int_\Omega |\nabla p^\tau|^2 (\rho_1^\tau + \rho_2^\tau) dx dt &\leq C \sum_{n=1}^N \tau \int_\Omega (|\nabla K_\varepsilon \star \rho_2^n|^2 + |\nabla \Phi_1|^2) \rho_1^n + (|\nabla K_\varepsilon \star \rho_1^n|^2 + |\nabla \Phi_2|^2) \rho_2^n dx \\ &\quad + \frac{C\tau}{\tau^2} \sum_{n=1}^N \sum_{i=1}^2 b_i W_2^2(\rho_i^n, \rho_i^{n-1}) \\ &\leq C \int_0^T \int_\Omega \frac{1}{\varepsilon^2} \left( (G_{2\varepsilon} \star \rho_2^\tau)^2 \rho_1^\tau + (G_{2\varepsilon} \star \rho_1^\tau)^2 \rho_2^\tau \right) dx dt + C. \end{aligned}$$

where we have used the uniform bounds on  $(\rho_i^\tau)_{\tau>0}$  from the previous points and

$$|\nabla K_\varepsilon \star \rho_i^n|^2 = \frac{2\pi\sigma}{\varepsilon} (|\nabla G_\varepsilon| \star \rho_i^n)^2 \leq \frac{2\pi\sigma}{\varepsilon} \left( \frac{1}{\sqrt{\varepsilon}} G_{2\varepsilon} \star \rho_i^n \right)^2 = \frac{2\pi\sigma}{\varepsilon^2} (G_{2\varepsilon} \star \rho_i^n)^2$$

Since  $\rho_1^n + \rho_2^n = 1$  a.e., this previous bound implies that  $(\nabla p^\tau)_{\tau>0}$  is uniformly bounded in  $L^2([0, T] \times \Omega; \mathbb{R}^d)$ . As a consequence of Poincaré's inequality, we have that  $(p^\tau - \int_\Omega p^\tau(\cdot, x) dx)_{\tau>0}$  is uniformly bounded in  $L^2([0, T] \times \Omega)$ . Both uniform bounds have the form  $TC(1 + 1/\varepsilon^2)$ .

To show (vii), let  $\xi \in C^1([0, T] \times \Omega; \mathbb{R}^d)$ . Fix  $t \in [n\tau, (n+1)\tau)$ . Taking the inner product of both sides in (20) with  $\xi(t, \cdot)$  and integrating on  $\Omega$  w.r.t.  $\rho_i^{n+1}$  (we drop the dependence on  $t$  in the notation of  $\xi$ ), we obtain

$$(25) \quad \int_{\Omega} \nabla p^{n+1} \cdot \xi \rho_1^{n+1} dx = -\sigma\sqrt{2\pi} \int_{\Omega} \frac{1}{\sqrt{\varepsilon}} \nabla G_{\varepsilon} \star \rho_2^{n+1} \cdot \xi \rho_1^{n+1} dx - \int_{\Omega} \nabla \Phi_1 \cdot \xi \rho_1^{n+1} dx - \int_{\Omega} \frac{b_1}{\tau} \nabla \varphi_1^{n+1} \cdot \xi \rho_1^{n+1} dx$$

and interchanging the roles of  $\rho_1^{n+1}$  and  $\rho_2^{n+1}$ , we get

$$(26) \quad \int_{\Omega} \nabla p^{n+1} \cdot \xi \rho_2^{n+1} dx = -\sigma\sqrt{2\pi} \int_{\Omega} \frac{1}{\sqrt{\varepsilon}} \nabla G_{\varepsilon} \star \rho_1^{n+1} \cdot \xi \rho_2^{n+1} dx - \int_{\Omega} \nabla \Phi_2 \cdot \xi \rho_2^{n+1} dx - \int_{\Omega} \frac{b_2}{\tau} \nabla \varphi_2^{n+1} \cdot \xi \rho_2^{n+1} dx$$

First, we have

$$\begin{aligned} & \int_{\Omega} \frac{1}{\sqrt{\varepsilon}} \nabla G_{\varepsilon} \star \rho_2^{n+1} \cdot \xi \rho_1^{n+1} dx + \int_{\Omega} \frac{1}{\sqrt{\varepsilon}} \nabla G_{\varepsilon} \star \rho_1^{n+1} \cdot \xi \rho_2^{n+1} dx \\ &= \frac{1}{2\varepsilon} \int_{\mathbb{R}^d} \int_{\Omega} zG(z) [\rho_2^{n+1}(x + \sqrt{\varepsilon}z) \cdot \xi(x) \rho_1^{n+1}(x) + \rho_1^{n+1}(x + \sqrt{\varepsilon}z) \cdot \xi(x) \rho_2^{n+1}(x)] dx dz \\ &= \frac{1}{2\varepsilon} \int_{\mathbb{R}^d} \int_{\Omega} zG(z) [\rho_2^{n+1}(x + \sqrt{\varepsilon}z) \cdot \xi(x) \rho_1^{n+1}(x) + \rho_1^{n+1}(x) \cdot \xi(x - \sqrt{\varepsilon}z) \rho_2^{n+1}(x - \sqrt{\varepsilon}z)] dx dz \\ &= \frac{1}{2\varepsilon} \int_{\Omega} \int_{\mathbb{R}^d} zG(z) [\rho_2^{n+1}(x + \sqrt{\varepsilon}z) \cdot \xi(x) \rho_1^{n+1}(x) - \rho_1^{n+1}(x) \cdot \xi(x + \sqrt{\varepsilon}z) \rho_2^{n+1}(x + \sqrt{\varepsilon}z)] dz dx \\ &= \frac{1}{2\varepsilon} \int_{\Omega} \int_{\mathbb{R}^d} zG(z) \rho_2^{n+1}(x + \sqrt{\varepsilon}z) \rho_1^{n+1}(x) \cdot [\xi(x) - \xi(x + \sqrt{\varepsilon}z)] dz dx \\ &= \frac{1}{2\sqrt{\varepsilon}} \int_{\Omega} \int_{\mathbb{R}^d} \int_1^0 zG(z) \rho_2^{n+1}(x + \sqrt{\varepsilon}z) \rho_1^{n+1}(x) \cdot (D\xi(x + s\sqrt{\varepsilon}z)z) ds dz dx \end{aligned}$$

where in the second and third equalities we have used the change of variables  $x \mapsto x - \sqrt{\varepsilon}z$  and  $z \mapsto -z$ , respectively and the fundamental theorem of calculus in the last equality. Now, since

$$|z|^2 G(z) \leq \alpha G(z/\beta), \quad \forall z \in \mathbb{R}^d,$$

for some universal constants  $\alpha, \beta > 0$ , by the previous chain of equalities, we obtain

$$\begin{aligned} \left| \int_{\Omega} \frac{1}{\sqrt{\varepsilon}} \nabla G_{\varepsilon} \star \rho_2^{n+1} \cdot \xi \rho_1^{n+1} dx + \int_{\Omega} \frac{1}{\sqrt{\varepsilon}} \nabla G_{\varepsilon} \star \rho_1^{n+1} \cdot \xi \rho_2^{n+1} dx \right| &\leq C \text{HC}_{2\varepsilon}(\rho_1^{n+1}, \rho_2^{n+1}) \|D\xi(t, \cdot)\|_{L^\infty} \\ &\leq C \text{HC}_{\varepsilon}(\rho_1^0, \rho_2^0) \|D\xi(t, \cdot)\|_{L^\infty}, \end{aligned}$$

where, in the last inequality we used the monotonicity of  $\text{HC}_{\varepsilon}$  along the sequence  $(\rho_1^n, \rho_2^n)_n$ .

Furthermore, we have that

$$\int_{\Omega} \frac{1}{\tau} |\nabla \varphi_i^{n+1}| |\xi| \rho_i^{n+1} dx = \int_{\Omega} |v_i^{n+1}| \rho_i^{n+1} |\xi| dx$$

and

$$\left| \int_{\Omega} \nabla \Phi_1 \cdot \xi \rho_1^{n+1} dx + \int_{\Omega} \nabla \Phi_1 \cdot \xi \rho_1^{n+1} dx \right| \leq C \|\xi\|_{L^1}$$

Using the fact that we have piecewise constant interpolations, integrating the last equality in time on  $[0, T]$ , we get

$$\int_0^T \int_{\Omega} |v_i^\tau| \rho_i^\tau |\xi| dx dt \leq \left( \int_0^T \int_{\Omega} |v_i^\tau|^2 \rho_i^\tau dx dt \right)^{\frac{1}{2}} \left( \int_0^T \int_{\Omega} |\xi|^2 dx dt \right)^{\frac{1}{2}}$$

Now, adding together (25) and (26), then integrating in time on  $[0, T]$  and using (ii), we obtain

$$\left| \int_0^T \int_{\Omega} \nabla p^\tau \cdot \xi dx dt \right| \leq C \|D\xi\|_{L^1(C^0)} + C \|\xi\|_{L^2(L^2)} \leq C \|\xi\|_{L^2(C^1)},$$

where the constant  $C > 0$  is independent of  $\tau > 0$  and  $\varepsilon > 0$ . The thesis of follows.  $\square$

Now we can use the above pressure estimates to derive a system of continuous in time equations in the limit  $\tau \rightarrow 0$  when  $\varepsilon$  is held fixed.

**Theorem 2.1.** *Let  $\varepsilon > 0$ ,  $\rho_{1,0}, \rho_{2,0} \in \mathcal{P}(\Omega)$  such that  $\rho_{1,0} + \rho_{2,0} = 1$  a.e. There exists  $\rho_i \in \text{AC}^2([0, T]; \mathcal{P}(\Omega))$ ,  $i = 1, 2$ ,  $\bar{p} \in L^2([0, T]; H^1(\Omega))$  and  $\zeta_1, \zeta_2 \in L^2([0, T] \times \Omega; \mathbb{R}^d)$  such that  $\rho_1 + \rho_2 = 1$  a.e. in  $[0, T] \times \Omega$ ,  $b_1 \zeta_1 + b_2 \zeta_2 = \nabla \bar{p}$  a.e. in  $[0, T] \times \Omega$  and the system*

$$(27) \quad \begin{cases} \partial_t \rho_1 - \nabla \cdot [b_1^{-1} \rho_1 (\nabla K_\varepsilon \star \rho_2 + \nabla \Phi_1) + \zeta_1] = 0, & (0, T) \times \Omega, \\ \partial_t \rho_2 - \nabla \cdot [b_2^{-1} \rho_2 (\nabla K_\varepsilon \star \rho_1 + \nabla \Phi_2) + \zeta_2] = 0, & (0, T) \times \Omega, \\ \rho_1(0, \cdot) = \rho_{1,0}, \quad \rho_2(0, \cdot) = \rho_{2,0}, & \Omega. \end{cases}$$

is satisfied in the sense of distributions on  $(0, T) \times \Omega$ .

*Proof.* Using the uniform (in  $\tau$ ) estimates in Lemma 2.5, we have the existence of  $\rho_i \in \text{AC}^2([0, T]; \mathcal{P}(\Omega))$ ,  $i = 1, 2$  such that up to passing to a subsequence, both  $(\rho_1^\tau, \rho_2^\tau)$  and  $(\tilde{\rho}_1^\tau, \tilde{\rho}_2^\tau)$  converge to them uniformly in time, w.r.t.  $W_2$ .

Since  $(p^\tau - \int_\Omega p^\tau(\cdot, x) dx)_{\tau > 0}$  is uniformly bounded in  $L^2([0, T]; H^1(\Omega))$ , there exists  $\bar{p} \in L^2([0, T]; H^1(\Omega))$  such that up to passing to a subsequence,  $p^\tau - \int_\Omega p^\tau(\cdot, x) dx \rightharpoonup \bar{p}$ , weakly in  $L^2([0, T] \times \Omega)$  and  $\nabla p^\tau \rightharpoonup \nabla \bar{p}$ , weakly in  $L^2([0, T] \times \Omega; \mathbb{R}^d)$  as  $\tau \downarrow 0$ .

We only need to identify the limits of the momentum variables  $(E_i^\tau)_{\tau > 0}$ . From the weak convergence of the density variables, we have that up to passing to a subsequence,  $\rho_1^\tau \nabla K_\varepsilon \star \rho_2^\tau \xrightarrow{*} \rho_1 \nabla K_\varepsilon \star \rho_2$ , as  $\tau \downarrow 0$  and similarly  $\rho_2^\tau \nabla K_\varepsilon \star \rho_1^\tau \xrightarrow{*} \rho_2 \nabla K_\varepsilon \star \rho_1$ , and  $\rho_i^\tau \nabla \Phi_i \xrightarrow{*} \rho_i \nabla \Phi_i$ , as  $\tau \downarrow 0$  as vector measures.

Furthermore, since  $(\rho_i^\tau \nabla p^\tau)_{\tau > 0}$  is uniformly bounded in  $L^2([0, T] \times \Omega; \mathbb{R}^d)$ , there exists  $\zeta_i \in L^2([0, T] \times \Omega; \mathbb{R}^d)$ ,  $i = 1, 2$  such that up to passing to a subsequence  $b_i^{-1} \rho_i^\tau \nabla p^\tau \rightharpoonup \zeta_i$  weakly in  $L^2([0, T] \times \Omega; \mathbb{R}^d)$ , as  $\tau \downarrow 0$ .

Last, by the previous arguments, we have

$$b_1 b_1^{-1} \rho_1^\tau \nabla p^\tau + b_2 b_2^{-1} \rho_2^\tau \nabla p^\tau = \nabla p^\tau \rightharpoonup \nabla \bar{p} = b_1 \zeta_1 + b_2 \zeta_2,$$

as  $\tau \downarrow 0$  in  $L^2([0, T] \times \Omega; \mathbb{R}^d)$ . Therefore, the thesis of the theorem follows.  $\square$

It remains open whether in the previous theorem we have strong convergence  $\rho_i^\tau \rightarrow \rho_i$  in  $L^2([0, T] \times \Omega)$  as  $\tau \downarrow 0$ , and in particular whether one can claim that  $\zeta_i = b_i^{-1} \rho_i \nabla \bar{p}$ . It is also open whether the continuum in time densities in Theorem 2.1 are characteristic functions. Indeed, since we can only guarantee that the discrete densities converge weakly to the continuum densities, the characteristic function property may be lost in the limit. We point out that due to the energy bounds, the densities are ‘‘almost’’ characteristic functions, i.e. we have

**Lemma 2.6.** *For  $(\rho_1^\tau, \rho_2^\tau)$  given as above with  $\varepsilon > 0$  fixed, for any  $\alpha \in (0, 1)$  we have*

$$\mathcal{L}^d(\alpha \leq \rho_i^\tau \leq 1 - \alpha) \leq \frac{3\sqrt{\varepsilon}}{\sigma\sqrt{2\pi\alpha(1-\alpha)}} \text{HC}_\varepsilon(\rho_1^\tau, \rho_2^\tau) \text{ for all } \tau.$$

*Proof.* Let us set  $i = 1$ , the other case will be parallel. We have

$$\mathcal{L}^d(\alpha \leq \rho_1^\tau \leq 1 - \alpha) \leq \frac{1}{\alpha(1-\alpha)} \int_\Omega \rho_1^\tau(x) (1 - \rho_1^\tau(x)) dx = \frac{1}{\alpha(1-\alpha)} \int_\Omega [\rho_1^\tau(x) - \rho_1^\tau(x)^2] dx.$$

By Jensen’s inequality, the above is smaller than

$$\begin{aligned} & \frac{1}{\alpha(1-\alpha)} \int_\Omega [\rho_1^\tau(x) - (G_\varepsilon * \rho_1^\tau)(x)^2] dx \\ & \leq \frac{\sqrt{\varepsilon}}{\sigma\sqrt{2\pi\alpha(1-\alpha)}} \text{HC}_\varepsilon(\rho_1^\tau, \rho_2^\tau) + \frac{1}{\alpha(1-\alpha)} \int_\Omega (G_\varepsilon * \rho_1^\tau)(x) (\rho_1^\tau(x) - (G_\varepsilon * \rho_1^\tau)(x)) dx. \end{aligned}$$

We then estimate

$$\frac{1}{\alpha(1-\alpha)} \int_\Omega (G_\varepsilon * \rho_1^\tau)(x) (\rho_1^\tau(x) - (G_\varepsilon * \rho_1^\tau)(x)) dx \leq \frac{1}{\alpha(1-\alpha)} \int_\Omega \int_{\mathbb{R}^d} G(z) |\rho_1^\tau(x) - \rho_1^\tau(x - \sqrt{\varepsilon}z)| dz dx$$

Since  $\rho_1^\tau$  and  $\rho_2^\tau$  take values in  $\{0, 1\}$ , we have  $|\rho_1^\tau(x) - \rho_1^\tau(x - \sqrt{\varepsilon}z)| \leq \rho_1^\tau(x) \rho_2^\tau(x - \sqrt{\varepsilon}z) + \rho_1^\tau(x - \sqrt{\varepsilon}z) \rho_2^\tau(x)$ . Applying these inequalities, the result follows.  $\square$

## 3. MUSKAT FLOW WITH SURFACE TENSION

In this section, we complete the proof of Theorem 1.1 and show that when  $\varepsilon = \varepsilon_\tau$  goes to zero along with  $\tau$ , the time interpolated minimizing movements scheme constructed in (23)-(24) converges to a weak formulation of the Muskat Problem with surface tension under the energy convergence assumption (EC). This amounts to showing that each quantity in the system of Euler-Lagrange equations given in (20) converges (in an appropriate sense) to the correct limiting object. In particular, we will need strong  $L^1$  convergence of the density functions in  $[0, T] \times \Omega$ . To obtain the necessary compactness for strong  $L^1$  convergence, we develop an adaptation of an Aubin-Lions type lemma in Proposition 3.2. We then conclude our result by verifying the convergence of the Euler-Lagrange equations to the weak formulation of the Muskat problem in a similar manner to the approach in [20].

Before we introduce the weak formulation of the Muskat Problem, let us recall the classical formulation of the problem. When the two phases are separated by a smooth interface  $\Gamma$ , the Muskat Problem is given by the continuity equation

$$(MP_1) \quad \partial_t \rho_i - b_i^{-1} \nabla \cdot ((\nabla p_i + \nabla \Phi_i) \rho_i) = 0$$

along with the free boundary problem for the pressure

$$(MP_2) \quad \begin{cases} -\Delta p_i = \Delta \Phi_i & \text{in } \text{spt}(\rho_i); \\ \partial_n(p_i + \Phi_i) = 0 & \text{on } \partial\Omega; \\ V = b_1^{-1} \partial_n(p_1 + \Phi_1) = b_2^{-1} \partial_n(p_2 + \Phi_2) & \text{on } \Gamma; \\ [p] := (p_1 - p_2) = \frac{\sigma}{2} \kappa & \text{on } \Gamma, \end{cases}$$

where  $\kappa$  is the mean curvature of  $\Gamma$ , oriented to be positive when  $\text{spt}(\rho_2)$  is convex at the point. The weak formulation of the Muskat Problem with surface tension is provided in Definition 3.1 below.

**Definition 3.1.** We say that  $(\rho_i, v_i, p)$  is a weak solution to the Muskat problem with surface tension, if  $\rho_i \in L^1([0, T]; BV(\Omega; \{0, 1\})) \cap AC^2([0, T]; \mathcal{P}(\Omega))$ ,  $\rho_1 \rho_2 = 0$  and  $\rho_1 + \rho_2 = 1$  a.e.,  $v_i \in L^2([0, T]; L^2_{\rho_i, t}(\Omega; \mathbb{R}^d))$ ,  $p \in L^2([0, T]; (C^{0, \alpha}(\Omega))^*)$  and for any  $\psi \in C^\infty([0, T] \times \Omega)$  and for any  $\xi \in C^\infty([0, T] \times \Omega; \mathbb{R}^d)$  vector field we have

$$(28) \quad \int_0^T \int_\Omega (\rho_i v_i \cdot \nabla \psi + \rho_i \partial_t \psi) \, dx \, dt = \left[ \int_\Omega \rho_i \psi \, dx \right]_0^T$$

with

$$(29) \quad \int_0^T \int_\Omega \sum_i \rho_i (b_i v_i + \nabla \Phi_i) \cdot \xi - p \nabla \cdot \xi + \frac{\sigma}{2} \left( \frac{D\rho_1}{|D\rho_1|} \otimes \frac{D\rho_1}{|D\rho_1|} : \nabla \xi + \nabla \cdot \xi \right) (|D\rho_1| + |D\rho_2|) \, dx \, dt = 0$$

*Remark 3.1.* Let us remark here that  $\frac{D\rho_1}{|D\rho_1|}$  stands for the  $L^\infty$  density of  $D\rho_1$  with respect to the total variation measure  $|D\rho_1|$  (after using the Radon-Nikodym differentiation). Furthermore, let us notice that the term  $\left( \frac{D\rho_1}{|D\rho_1|} \otimes \frac{D\rho_1}{|D\rho_1|} : \nabla \xi \right) (|D\rho_1| + |D\rho_2|)$  is meaningful in the sense that  $f(\nu/|\nu|) \, d|\nu|$  defines a matrix valued element of  $\mathcal{M}(\Omega)$  for any  $f : \mathbb{R}^d \rightarrow \mathbb{R}^{d \times d}$  which is continuous and 1-homogeneous (see for instance [2, Proposition 3.15] in the case when  $\nu = D\rho$  for some  $\rho \in BV(\Omega)$ ). In particular, if  $\rho \in BV(\Omega; \{0, 1\})$ , then  $D\rho/|D\rho|$  stands for the measure theoretic normal to the boundary of the set  $\text{spt}(\rho)$  and by  $\nu/|\nu|$  we mean the density of  $\nu$  with respect to its total variation measure  $|\nu|$ .

Now we are ready to state the main result of our paper.

**Theorem 3.1.** *Given initial data  $\rho_{1,0}, \rho_{2,0} \in BV(\Omega)$  such that  $\rho_{1,0} \rho_{2,0} = 0$  and  $\rho_{1,0} + \rho_{2,0} = 1$  a.e. in  $\Omega$  and Lipschitz continuous potential functions  $\Phi_1, \Phi_2 : \Omega \rightarrow \mathbb{R}$ , there exists  $(\rho_i, v_i, p)$  with  $i = 1, 2$  such that under the energy convergence assumption (EC), it solves (MP<sub>1</sub>)-(MP<sub>2</sub>) in the sense of Definition 3.1.*

We postpone the proof of the previous theorem to the end of this section. First, let us show a consistency result, i.e. that classical solutions of the Muskat problem satisfy the weak formulation (28) and (29).

**Lemma 3.1.** *If smooth solutions of (MP<sub>1</sub>) and (MP<sub>2</sub>) exist with  $\rho_i = \chi_{A_i}$  and with a  $C^2$  hypersurface  $\Gamma = \partial A_1 = \partial A_2$ , then  $\rho_i$  satisfies (28) and (29) with the choice of*

$$(30) \quad v_i := -b_i^{-1}(\nabla p_i + \nabla \Phi_i) \text{ and } p = p_1 \rho_1 + p_2 \rho_2 + \frac{\sigma}{2} d\mu,$$

where  $\mu \in (C([0, T] \times \Omega))^*$  is the surface measure of  $\Gamma = \cup_{t>0}(\Gamma_t \times \{t\})$ , i.e. after disintegration

$$\mu = \mathcal{H}^{d-1} \llcorner \Gamma_t \otimes \mathcal{L}^1 \llcorner [0, T]$$

or using test functions

$$\int_0^T \int_{\Omega} f d\mu = \int_0^T \int_{\Gamma_t} f d\mathcal{H}^{d-1} dt \text{ for any } f \in C([0, T] \times \Omega).$$

*Remark 3.2.* Note that our notion of solution requires adding the surface measure  $\frac{\sigma}{2} d\mu$  to the classical pressure variable. This singular term appears from the minimizing movement scheme, where it ensures that a vacuum does not form at the interface. In general, we expect that the singular part in the weak pressure variable corresponds to the surface measure in (30).

*Proof.* (28) is a standard weak expression of (MP<sub>1</sub>) with  $b_i v_i = -\nabla p_i - \nabla \Phi_i$ . Let us write again  $\Gamma = \cup_{t>0}(\Gamma_t \times \{t\})$ . Then we have

$$\begin{aligned} \int_0^T \int_{\Omega} \left[ \sum_i \rho_i (b_i v_i + \nabla \Phi_i) \cdot \xi - p \nabla \cdot \xi \right] dx dt &= \int_0^T \int_{\Omega} \left[ \sum_i \rho_i (b_i v_i + \nabla \Phi_i) \cdot \xi - (p_1 \rho_1 + p_2 \rho_2) \nabla \cdot \xi \right] dx dt \\ &\quad - \int_0^T \int_{\Omega} \frac{\sigma}{2} (\nabla \cdot \xi) d\mu \\ &= \int_0^T \int_{\Gamma_t} (p_1 - p_2) \xi \cdot \nu d\mathcal{H}^{d-1} dt - \frac{\sigma}{2} \int_0^T \int_{\Gamma_t} \nabla \cdot \xi d\mathcal{H}^{d-1} dt \\ &= \frac{\sigma}{2} \int_0^T \int_{\Gamma_t} \kappa \xi \cdot \nu d\mathcal{H}^{d-1} dt - \frac{\sigma}{2} \int_0^T \int_{\Gamma_t} \nabla \cdot \xi d\mathcal{H}^{d-1} dt. \end{aligned}$$

Here for the second equality we used integration by parts for the first integral, using the fact that

$$\nabla \left( \sum_i p_i \rho_i \right) = \sum_i \nabla p_i \rho_i + (p_1 - p_2) d\mathcal{H}^{d-1} \llcorner \Gamma_t,$$

and for the third equality we used the curvature jump condition at the interface. To conclude, let us recall that  $\kappa$  is oriented to be positive when  $\text{spt}(\rho_2)$  is convex at the point. With this orientation, observe that for any  $C^1$  vector field  $\xi$  we have

$$\int_{\Gamma_t} \kappa \xi \cdot \nu d\mathcal{H}^{d-1} = \int_{\Gamma_t} [\nabla \cdot \xi - \nu \cdot \nabla \xi \nu] d\mathcal{H}^{d-1}$$

where  $\nu = D\rho_1/|D\rho_1|$  is normal to  $\Gamma$  toward the support of  $\rho_1$ . Hence we conclude that (29) holds.  $\square$

**3.1. Preliminary estimates.** We present below a compactness result on the piecewise constant interpolations of the density variables  $(\rho_1^\tau, \rho_2^\tau)_{\tau>0}$ , when the parameter  $\varepsilon$  is vanishing together with  $\tau$ .

**Proposition 3.2.** *Let  $(\rho_1^\tau, \rho_2^\tau)_{\tau>0}$  be the piecewise constant interpolations constructed in Section 2 using the minimizing movements scheme (22). Let moreover  $\rho_{1,0}, \rho_{2,0} \in \mathcal{P}(\Omega)$  be such that  $\rho_{i,0} \in BV(\Omega; \{0, 1\})$  with  $\rho_{1,0} + \rho_{2,0} = 1$  a.e. in  $\Omega$  and in particular  $\mathcal{E}(\rho_{1,0}, \rho_{2,0}) < +\infty$ .*

*Then, there exists  $\rho_1, \rho_2 \in L^1([0, T]; BV(\Omega; \{0, 1\}))$  such that  $\rho_1 + \rho_2 = 1$  a.e. in  $[0, T] \times \Omega$  and up to passing to a subsequence,  $\rho_i^\tau \rightarrow \rho_i$  strongly in  $L^1([0, T] \times \Omega)$  as  $\varepsilon_\tau := \max\{\tau, \varepsilon\} \downarrow 0$ .*

*Proof.* First, let us notice that by the uniform quasi-Hölder estimate from Lemma 2.5(i), we have that there exists a subsequence of  $(\rho_i^\tau)_{\tau>0}$  (that we do not relabel) and  $\rho_i \in AC^2([0, T]; \mathcal{P}(\Omega))$  such that  $W_2(\rho_{i,t}^\tau, \rho_{i,t}) \rightarrow 0$  as  $\varepsilon_\tau \downarrow 0$ , uniformly in  $t$ . We shall work with this subsequence from now on.

The rest of the proof relies on a careful adaptation of the Aubin-Lions lemma, developed in [32] and used in similar context for instance in [18, 19].

Let us notice that in order to use the Aubin-Lions argument, we need to show a tightness condition (cf. [32, Definition 1.3, Remark 1.5]) for the time-dependent family  $(\rho_i^\tau)_{\tau>0}$ . This is typically done by providing a compact set (of the space of probability measures), where ‘most’ of the sequences  $(\rho_i^\tau(t))_{\tau>0}$  lie. Note that the estimate in Lemma 2.5(v) does not provide such a compact set. Indeed, for  $\tau > 0$ , the densities  $(\rho_i^\tau(t))_{\tau>0}$  are not actually BV in space. Inspired by arguments in [12], we will work first with an auxiliary sequence  $\bar{\rho}_i^\tau : [0, T] \times \Omega \rightarrow [0, 1]$ , defined as

$$\bar{\rho}_i^\tau := G_\varepsilon \star \rho_i^\tau,$$

where we have performed a convolution with the heat kernel  $G_\varepsilon$  only in the spacial variable. It worth noticing that this ‘perturbation’ of  $\rho_i^\tau$  is reminiscent to the one obtained via the so-called flow interchange technique introduced in [24]. We have the following properties for this new sequence.

*Claim 1.*  $(\bar{\rho}_i^\tau)_{\tau>0}$  is uniformly (w.r.t.  $\varepsilon$  and  $\tau$ ) in  $L^1([0, T]; BV(\Omega))$ .

*Proof of Claim 1.* The uniform  $L^\infty$  bounds on  $(\rho_i^\tau)_{\tau>0}$  are also preserved for  $(\bar{\rho}_i^\tau)_{\tau>0}$ . Now, as in the proof of [20, Lemma 2.4] we conclude that

$$(31) \quad \int_0^T \int_\Omega |\nabla \bar{\rho}_{i,t}^\tau| dx dt \leq C \text{THC}_\varepsilon(\rho_{1,0}, \rho_{2,0}),$$

for a uniform constant  $C$ , which proves our claim.

*Claim 2.* There exists a subsequence of  $(\bar{\rho}_i^\tau)_{\tau>0}$  (that we do not relabel) and  $\bar{\rho}_i \in L^1([0, T] \times \Omega)$  such that  $\bar{\rho}_i^\tau \rightarrow \bar{\rho}_i$  strongly in  $L^1([0, T] \times \Omega)$  as  $\varepsilon_\tau \downarrow 0$ .

*Proof of Claim 2.* For  $t \in (0, T)$  fixed, let us notice that  $\bar{\rho}_{i,t}^\tau$  actually can be seen as the evolution of  $\rho_{i,t}^\tau$  via the heat flow for time  $\varepsilon > 0$ . It is well-known that  $W_2$  is contractive along the heat flow, so we have

$$(32) \quad W_2(\bar{\rho}_{i,t}^\tau, \bar{\rho}_{i,s}^\tau) \leq W_2(\rho_{i,t}^\tau, \rho_{i,s}^\tau), \quad \forall 0 \leq s \leq t \leq T.$$

Now let us set  $g : L^1(\Omega) \times L^1(\Omega) \rightarrow [0, +\infty]$  and  $\mathcal{F} : L^1(\Omega) \rightarrow [0, +\infty]$  defined as

$$g(\mu, \nu) := W_2(\mu, \nu)$$

and

$$\mathcal{F}(\mu) := \begin{cases} |D\mu|(\Omega), & \text{if } \mu \in BV(\Omega), \\ +\infty, & \text{otherwise.} \end{cases}$$

By construction,  $\mathcal{F}$  is convex, l.s.c. in  $L^1(\Omega)$  and its sublevel sets are compact in  $L^1(\Omega)$ , therefore it defines a normal coercive integrand.

From (31) and from the definition of  $\mathcal{F}$ , we have

$$\sup_{\varepsilon_\tau} \int_0^T \mathcal{F}(\bar{\rho}_{i,t}^\tau) dt < +\infty.$$

And similarly from Lemma 2.5(i) and from (32), we have

$$\limsup_{h \downarrow 0} \sup_{\varepsilon_\tau} \int_0^{T-h} g(\bar{\rho}_{i,t+h}^\tau, \bar{\rho}_{i,t}^\tau) dt < +\infty,$$

therefore the assumptions of [32, Theorem 2] are fulfilled and one can conclude that there exists  $\bar{\rho}_i \in L^1([0, T] \times \Omega)$  and a subsequence of  $(\bar{\rho}_i^\tau)_{\tau>0}$  (that we do not relabel) which is converging in measure, and in particular pointwise a.e. to  $\bar{\rho}_i$  as  $\varepsilon_\tau \downarrow 0$ . The strong convergence in  $L^1([0, T] \times \Omega)$  follows from Lebesgue’s dominated convergence theorem, since  $(\bar{\rho}_i^\tau)_{\tau>0}$  is uniformly bounded. The claim follows.

*Claim 3.* There exists a subsequence of the original sequence  $(\rho_i^\tau)_{\tau>0}$  which is converging to  $\bar{\rho}_i$  strongly in  $L^1([0, T] \times \Omega)$  as  $\varepsilon_\tau \downarrow 0$ .

*Proof of Claim 3.* Passing to the same subsequence in the original sequence  $(\rho_i^\tau)_{\tau>0}$  (that we do not relabel) as in the last step of the proof of Claim 2, we have

$$\|\rho_i^\tau - \bar{\rho}_i\|_{L^1([0, T] \times \Omega)} \leq \|\rho_i^\tau - \bar{\rho}_i^\tau\|_{L^1([0, T] \times \Omega)} + \|\bar{\rho}_i^\tau - \bar{\rho}_i\|_{L^1([0, T] \times \Omega)} \leq C\sqrt{\varepsilon} + \|\bar{\rho}_i^\tau - \bar{\rho}_i\|_{L^1([0, T] \times \Omega)},$$

for a constant  $C > 0$  (independent of  $\tau$  and  $\varepsilon$ ) and the claim follows by taking  $\varepsilon_\tau \downarrow 0$ . In the last inequality we have used

$$\begin{aligned} \|\rho_1^\tau - \bar{\rho}_1^\tau\|_{L^1([0,T] \times \Omega)} &= \int_0^T \int_\Omega |\rho_1^\tau - G_\varepsilon \star \rho_1^\tau| dx dt \leq \int_0^T \int_\Omega \int_{\mathbb{R}^d} G(z) |\rho_1^\tau(x) - \rho_1^\tau(x - \sqrt{\varepsilon}z)| dz dx dt \\ &\leq CT\sqrt{\varepsilon} \text{HC}_\varepsilon(\rho_{1,0}, \rho_{2,0}), \end{aligned}$$

where we relied on the fact that since  $\rho_1^\tau$  and  $\rho_2^\tau$  take values in  $\{0, 1\}$ , we have  $|\rho_1^\tau(x) - \rho_1^\tau(x - \sqrt{\varepsilon}z)| \leq \rho_1^\tau(x)\rho_2^\tau(x - \sqrt{\varepsilon}z) + \rho_1^\tau(x - \sqrt{\varepsilon}z)\rho_2^\tau(x)$ .

To conclude, let us notice that since  $(\rho_i^\tau)_{\tau>0}$  converges uniformly w.r.t.  $W_2$  to  $\rho_i$ ,  $\bar{\rho}_i$  and  $\rho_i$  have to coincide. Also, for this last subsequence, we can pass to the limit as  $\varepsilon_\tau \downarrow 0$  in the estimation of Lemma 2.5(v) to obtain that actually  $\rho_i \in L^1([0, T]; BV(\Omega; \{0, 1\}))$ .  $\square$

Let  $\boldsymbol{\rho} = (\rho_1, \rho_2)$  the limit point obtained in Proposition 3.2. Later in this section, we will profit from the assumption

$$(EC) \quad \int_0^T \text{HC}_{\varepsilon_\tau}(\boldsymbol{\rho}^\tau) dt \rightarrow \int_0^T \mathcal{E}_s(\boldsymbol{\rho}) dt, \text{ as } \varepsilon \downarrow 0.$$

**3.2. Derivation of the weak curvature equation for  $\varepsilon$  going to zero together with  $\tau$ .** Let  $\xi \in C^2([0, T] \times \Omega; \mathbb{R}^d)$ . Fix  $t \in [n\tau, (n+1)\tau)$ . We consider the piecewise constant interpolations  $(\rho_i^\tau, v_i^\tau, p^\tau)$ . Let us take the inner product of the equations (20) with  $\xi(t, \cdot)$ , multiply with the corresponding  $\rho_i^\tau$  and integrate over  $[0, T] \times \Omega$ . Adding up the two equations we get

$$\int_0^T \int_\Omega [\nabla K_\varepsilon \star \rho_2^\tau + \nabla \Phi_1 + v_1^\tau + \nabla p^\tau] \cdot \xi \rho_1^\tau dx dt + \int_0^T \int_\Omega [\nabla K_\varepsilon \star \rho_1^\tau + \nabla \Phi_2 + v_2^\tau + \nabla p^\tau] \cdot \xi \rho_2^\tau dx dt = 0,$$

rearranging the terms yields

$$(33) \quad \int_0^T \int_\Omega \xi \cdot [(\nabla K_\varepsilon \star \rho_2^\tau) \rho_1^\tau + (\nabla K_\varepsilon \star \rho_1^\tau) \rho_2^\tau] + (\rho_1^\tau(v_1^\tau + \nabla \Phi_1) + \rho_2^\tau(v_2^\tau + \nabla \Phi_1)) \cdot \xi + \nabla p^\tau \cdot \xi dx dt = 0,$$

Our aim is now to pass to the limit in this last expression (33) as  $\varepsilon_\tau \downarrow 0$  to recover (29).

**Theorem 3.2.** *Let  $(\rho_i^\tau, v_i^\tau, p^\tau)_{\tau>0}$  be the piecewise constant interpolations constructed in (23)-(24). Then there exists  $\rho_i \in L^1([0, T]; BV(\Omega; \{0, 1\})) \cap AC^2([0, T]; \mathcal{P}(\Omega))$ ,  $v_i \in L^2([0, T]; L^2_{\rho_i}(\Omega; \mathbb{R}^d))$  and  $p \in L^2([0, T]; (C^{0,\alpha}(\Omega))^*)$  such that, up to passing to a subsequence that we do not relabel, we have the following*

- (i)  $\rho_i^\tau \rightarrow \rho_i$ , as  $\varepsilon_\tau \downarrow 0$ , strongly in  $L^1([0, T] \times \Omega)$ ;
- (ii)  $v_i^\tau \rho_i^\tau \xrightarrow{*} v_i \rho_i$ , as  $\varepsilon_\tau \downarrow 0$ , weakly- $\star$  in  $\mathcal{M}^d([0, T] \times \Omega)$ ;
- (iii)  $\nabla p^\tau \xrightarrow{*} \nabla p$ , as  $\varepsilon_\tau \downarrow 0$ , weakly- $\star$  in  $L^2([0, T]; (C^1(\Omega))^*)$ ;

*Proof.* (i) is a consequence of Proposition 3.2 via the Aubin-Lions type argument.

(ii) Let us notice that the estimates in Lemma 2.5(i-iii) are independent of  $\varepsilon > 0$ , therefore there exists  $E_i \in \mathcal{M}^d([0, T] \times \Omega)$ ,  $i = 1, 2$  such that  $E_i^\tau \xrightarrow{*} E_i$  and  $\tilde{E}_i^\tau \xrightarrow{*} E_i$  as  $\varepsilon_\tau \downarrow 0$ . Moreover, we have that  $(\rho_i, E_i)$  solves  $\partial_t + \nabla \cdot (E_i) = 0$  in the sense of distributions and  $\rho_i \in AC^2([0, T]; (\mathcal{P}(\Omega), W_2))$ . Therefore, there exists  $v_i \in L^2([0, T]; L^2_{\rho_i}(\Omega; \mathbb{R}^d))$  such that  $\partial_t \rho_i + \nabla \cdot (\rho_i v_i) = 0$  in the sense of distributions.

(iii) Let us notice first that from Lemma 2.5(vii) we have that the sequence  $(\nabla p^\tau)_\tau$  is uniformly bounded in  $L^2([0, T]; (C^1(\Omega))^*)$ , independently of  $\varepsilon$ . Then, the Banach-Bourbaki-Alaouglu theorem yields that it is sequentially pre-compact in that space, so we have that there exists  $\zeta \in L^2([0, T]; (C^1(\Omega))^*)$  such that up to passing to a subsequence  $\nabla p^\tau \xrightarrow{*} \zeta$ , as  $\varepsilon_\tau \downarrow 0$ .

Thus, it only remains to show that  $\zeta$  is a gradient. Let us notice that by construction of  $p^\tau$ ,

$$\int_0^T \int_\Omega \nabla p^\tau \cdot \xi dx dt = 0,$$

for any incompressible smooth field  $\xi$ . Therefore, we also have that  $\langle \zeta, \xi \rangle = 0$  for any incompressible smooth field  $\xi$ . So we would be done, if one would have a Helmholtz decomposition in this corresponding



space. This result is well known (see for instance [35, Lemma 2.2.1]), if  $\zeta(t, \cdot) \in W^{-1,q}(\Omega)^d$ , for some  $q \in (1, +\infty)$ . However, our limit object has slightly worse regularity.

To overcome this issue, let us argue using the following claim. This is a consequence of classical result in the theory of elliptic equations and Schauder estimates (see for instance [14, Theorem 5.23-Theorem 5.24])

*Claim.* Let  $\varphi \in C^{0,\alpha}(\Omega)$ . Then the problem  $-\Delta u = \varphi$  (with periodic or Neumann boundary conditions) has a unique (modulo constants) solution  $u \in C^{2,\alpha}(\Omega)$  such that  $\|u\|_{C^{2,\alpha}} \leq \|\varphi\|_{C^{0,\alpha}}$ .

Now, let  $\varphi \in C^{0,\alpha}(\Omega)$  and  $u \in C^{2,\alpha}(\Omega)$  as in the claim. Let  $t \in [0, T]$ . Then we have

$$\left| \int_{\Omega} p^\tau(t, \cdot) \varphi \, dx \right| = \left| \int_{\Omega} p^\tau(t, \cdot) \Delta u \, dx \right| = \left| \int_{\Omega} \nabla p^\tau(t, \cdot) \cdot \nabla u \, dx \right| \leq C \|\nabla u\|_{C^1} \leq C \|u\|_{C^{2,\alpha}} \leq C \|\varphi\|_{C^{0,\alpha}},$$

where in the first inequality we have used the last uniform estimate from the proof of Lemma 2.5(vii). This implies that  $(p^\tau(t, \cdot))_\tau$  is uniformly bounded in  $(C^{0,\alpha}(\Omega))^*$ .

To conclude the thesis of the lemma, we observe that (by possibly subtracting the mean)  $(p^\tau)_\tau$  is also converging weakly- $\star$  to some  $p$ , therefore we will have that  $\zeta = \nabla p$  in  $\mathcal{D}'((0, T) \times \Omega)$ .  $\square$

To complete the proof of of Theorem 1.1, it remains to show that limit of equation (33) converges to the weak formulation of the Muskat problem. This result is proven in Proposition 3.3, which in turn uses the results of Lemmas 3.4 and 3.5. Note that the energy convergence assumption is finally used in Lemma 3.4. The assumption plays a crucial role in the following arguments as it allows us to convert a limit requiring uniform convergence to one which only requires pointwise convergence. In what follows, it will be useful for notational simplicity to introduce the following definition:

**Definition 3.2.** For a smooth rapidly decaying kernel  $J : \mathbb{R}^d \rightarrow \mathbb{R}$  and a vector valued Radon measure  $\nu \in \mathcal{M}^d(\Omega)$  we define the Radon measure  $\sigma_J(\nu) \in \mathcal{M}(\Omega)$  as

$$\int_{\Omega} \psi(x) \, d\sigma_J(\nu)(x) := \int_{\mathbb{R}^d} J(z) \int_{\Omega} \psi(x) \left| z \cdot \frac{\nu(x)}{|\nu(x)|} \right| \, d|\nu|(x) \, dz, \quad \forall \psi \in C(\Omega).$$

**Proposition 3.3.** Let  $(\rho_i^\tau)_{\tau>0}$  be the piecewise constant interpolations constructed in (23)-(24) and let  $\rho_i \in L^1([0, T]; BV(\Omega; \{0, 1\}))$ ,  $i = 1, 2$  be their strong  $L^1$  limits. If the assumption (EC) is fulfilled, then up to passing to a subsequence that we do not relabel, we have

$$(34) \quad \int_0^T \int_{\Omega} \xi \cdot [(\nabla K_{\varepsilon_\tau} \star \rho_2^\tau) \rho_1^\tau + (\nabla K_{\varepsilon_\tau} \star \rho_1^\tau) \rho_2^\tau] \, dx \, dt \rightarrow \sum_i \int_0^T \int_{\Omega} \frac{1}{4} \sum_k \zeta_k(x) \, d\sigma_J(D\rho_i)(x) \, dt \\ = \int_0^T \int_{\Omega} \frac{\sigma}{2} \left( \frac{D\rho_1}{|D\rho_1|} \otimes \frac{D\rho_1}{|D\rho_1|} : \nabla \xi + \nabla \cdot \xi \right) (|D\rho_1| + |D\rho_2|) \, dx \, dt$$

as  $\varepsilon_\tau \downarrow 0$ , for any  $\xi \in C^3([0, T] \times \Omega; \mathbb{R}^d)$ . Here, we denoted  $J(z) := \sigma \sqrt{2\pi} |z \cdot e_1|^2 G(z)$  and we used the decomposition  $\nabla \xi^{\text{sym}}(x) = \sum_k \zeta_k(x) n_k \otimes n_k$  with  $\zeta_k \in C^\infty(\Omega)$  and  $n_k \in S^{d-1}$  such that  $\{n_k \otimes n_k\}_{k=1}^{d(d+1)/2}$  is an appropriate basis of the space of symmetric matrices.

*Proof.* Let  $\xi \in C^3([0, T] \times \Omega; \mathbb{R}^d)$ . Let us show that along a subsequence

$$\int_0^T \int_{\Omega} \xi \cdot [(\nabla K_{\varepsilon_\tau} \star \rho_2^\tau) \rho_1^\tau + (\nabla K_{\varepsilon_\tau} \star \rho_1^\tau) \rho_2^\tau] \, dx \, dt \rightarrow \sum_i \int_0^T \int_{\Omega} \frac{1}{4} \sum_k \zeta_k(x) \, d\sigma_J(D\rho_i)(x) \, dt,$$

as  $\varepsilon_\tau \downarrow 0$ . This result is not straight forward, since both the functional and the densities depend on  $\tau$ . This difficulty is the key reason that we need the energy convergence assumption (EC).

Arguing exactly as in the proof of [20, Lemma 2.8], in order to show the convergence result (34), it is enough to show its time-independent version, i.e.

$$\int_{\Omega} \xi \cdot [(\nabla K_{\varepsilon_\tau} \star \rho_2^\tau) \rho_1^\tau + (\nabla K_{\varepsilon_\tau} \star \rho_1^\tau) \rho_2^\tau] \, dx \rightarrow \sum_i \int_{\Omega} \frac{1}{4} \sum_k \zeta_k(x) \, d\sigma_J(D\rho_i)(x),$$

as  $\varepsilon_\tau \downarrow 0$ , for  $\mathcal{L}^1$ -a.e.  $t \in [0, T]$ .

A computation similar to the one in the proof of Lemma 2.5(viii) reveals

$$\begin{aligned}
(35) \quad & \int_{\Omega} \xi \cdot [(\nabla K_{\varepsilon_\tau} \star \rho_2^\tau) \rho_1^\tau + (\nabla K_{\varepsilon_\tau} \star \rho_1^\tau) \rho_2^\tau] dx \\
&= \frac{\sigma\sqrt{2\pi}}{2\varepsilon_\tau} \int_{\Omega} \int_{\mathbb{R}^d} zG(z) \rho_2^\tau(x + \sqrt{\varepsilon_\tau}z) \rho_1^\tau(x) \cdot [\xi(x) - \xi(x + \sqrt{\varepsilon_\tau}z)] dz dx \\
&= -\frac{\sigma\sqrt{2\pi}}{\sqrt{\varepsilon_\tau}} \int_{\mathbb{R}^d} \int_{\Omega} \rho_2^\tau(x + \sqrt{\varepsilon_\tau}z) \rho_1^\tau(x) [z \otimes \nabla G(z) : \nabla \xi(x)] dx dt \\
&+ O\left(\|D^2\xi\|_{L^\infty} \int_{\mathbb{R}^d} |z|^2 |\nabla G(z)| \int_{\Omega} \rho_1^\tau(x) \rho_2^\tau(x + \sqrt{\varepsilon_\tau}z) dx dz\right)
\end{aligned}$$

where, we were using a second order Taylor expansion (and the smoothness of  $\xi$ ) in the last equality. Let us observe that the very last term in (35) is converging to 0 as  $\varepsilon_\tau \downarrow 0$  (due to the strong convergence  $\rho_i^\tau \rightarrow \rho_i$  in  $L^1(\Omega)$  and the fact that  $\rho_1\rho_2 \equiv 0$  a.e., see Proposition 3.2).

Therefore, it remains to prove

$$-\frac{\sigma\sqrt{2\pi}}{\sqrt{\varepsilon_\tau}} \int_{\mathbb{R}^d} \int_{\Omega} \rho_2^\tau(x + \sqrt{\varepsilon_\tau}z) \rho_1^\tau(x) [z \otimes \nabla G(z) : \nabla \xi(x)] dx dz \rightarrow \sum_i \int_{\Omega} \frac{1}{4} \sum_k \zeta_k(x) d\sigma_J(D\rho_i)(x),$$

as  $\varepsilon_\tau \downarrow 0$ . We note that

$$-z \otimes \nabla G(z) : \nabla \xi = (z \otimes z)G(z) : \nabla \xi = (z \otimes z)G(z) : \nabla \xi^{\text{sym}}$$

where  $\nabla \xi^{\text{sym}}$  denotes the symmetric part of  $\nabla \xi$ . A basis for the space of  $d \times d$  symmetric matrices is given by the  $d + \frac{d(d-1)}{2}$  matrices

$$\{e_i \otimes e_i\}_{i=1}^d \cup \left\{ \frac{1}{2}(e_i + e_j) \otimes (e_i + e_j) \right\}_{1 \leq i < j \leq d}$$

where  $e_i \in \mathbb{R}^d$  is the  $i^{\text{th}}$  standard basis vector. All of these basis matrices have the form  $n \otimes n$  for some  $n \in S^{d-1}$ . Thus, we may write

$$\nabla \xi^{\text{sym}}(x) = \sum_k \zeta_k(x) n_k \otimes n_k$$

where  $\{n_k\}$  is any indexing of the above matrices and  $\zeta_k(x) n_k \otimes n_k = P_k \nabla \xi^{\text{sym}}(x)$  where  $P_k$  is the appropriate change of basis matrix.

Therefore we have

$$\begin{aligned}
& \frac{\sigma\sqrt{2\pi}}{\sqrt{\varepsilon_\tau}} \int_{\mathbb{R}^d} \int_{\Omega} \rho_2^\tau(x + \sqrt{\varepsilon_\tau}z) \rho_1^\tau(x) [z \otimes z G(z) : \nabla \xi(x)] dx dz \\
&= \frac{\sigma\sqrt{2\pi}}{\sqrt{\varepsilon_\tau}} \sum_k \int_{\mathbb{R}^d} \int_{\Omega} \rho_2^\tau(x + \sqrt{\varepsilon_\tau}z) \rho_1^\tau(x) \zeta_k(x) |z \cdot n_k|^2 G(z) dx dz
\end{aligned}$$

Since the functions  $z \mapsto |z \cdot n_k|^2 G(z)$  are rotation invariant, when integrating on  $\mathbb{R}^d$ , it is enough to consider only  $n_k = e_1$ , where  $e_1$  is the first element of the standard basis on  $\mathbb{R}^d$ . Now, defining  $J(z) := \sigma\sqrt{2\pi}|z \cdot e_1|^2 G(z)$ , we deduce the claim in (iii) from Lemma 3.4 and Lemma 3.5.

Now, let us remark that a computation completely parallel to [20, Proof of Lemma 3.6.] reveals furthermore that

$$(36) \quad \frac{\sigma\sqrt{2\pi}}{\sqrt{\varepsilon_\tau}} \int_{\mathbb{R}^d} \int_{\Omega} \rho_2^\tau(x + \sqrt{\varepsilon_\tau}z) \rho_1^\tau(x) |z \cdot n|^2 G(z) \zeta(x) dx dz \rightarrow \int_{\Omega} \frac{\sigma}{2} \left( \left| \frac{D\rho_1}{|D\rho_1|} \cdot n \right|^2 + 1 \right) \zeta(x) (|D\rho_1| + |D\rho_2|) dx$$

as  $\varepsilon_\tau \rightarrow 0$ , and so, in this case we would have

$$\begin{aligned}
& \frac{\sigma\sqrt{2\pi}}{\sqrt{\varepsilon_\tau}} \sum_k \int_{\mathbb{R}^d} \int_{\Omega} \rho_2^\tau(x + \sqrt{\varepsilon_\tau}z) \rho_1^\tau(x) |z \cdot n_k|^2 G(z) \zeta_k(x) dx dz \\
&\rightarrow \int_{\Omega} \frac{\sigma}{2} \sum_k \left( \frac{D\rho_1}{|D\rho_1|} \otimes \frac{D\rho_1}{|D\rho_1|} : n_k \otimes n_k + 1 \right) \zeta_k(x) (|D\rho_1| + |D\rho_2|) dx.
\end{aligned}$$

This would in fact complete the claim, as we can compute

$$\begin{aligned} \sum_k \left( \frac{D\rho_1}{|D\rho_1|} \otimes \frac{D\rho_1}{|D\rho_1|} : n_k \otimes n_k + 1 \right) \zeta_k(x) &= \frac{D\rho_1}{|D\rho_1|} \otimes \frac{D\rho_1}{|D\rho_1|} : \sum_k \zeta_k(x) n_k \otimes n_k + \sum_k \zeta_k(x) \operatorname{Tr}(n_k \otimes n_k) \\ &= \frac{D\rho_1}{|D\rho_1|} \otimes \frac{D\rho_1}{|D\rho_1|} : \nabla \xi^{\operatorname{sym}}(x) + \operatorname{Tr}(\nabla \xi^{\operatorname{sym}}(x)) \\ &= \frac{D\rho_1}{|D\rho_1|} \otimes \frac{D\rho_1}{|D\rho_1|} : \nabla \xi(x) + \nabla \cdot \xi(x) \end{aligned}$$

where we have used  $\operatorname{Tr}(n_k \otimes n_k) = 1$  and the fact that the antisymmetric parts of  $\nabla \xi$  are annihilated by the above operations.  $\square$

The conclusion in the previous proposition is made by the following two lemmas.

**Lemma 3.4.** *If  $\rho^\tau \rightarrow \rho$  in  $L^1(\Omega) \times L^1(\Omega)$  and  $\operatorname{HC}_{\varepsilon_\tau}(\rho^\tau) \rightarrow \mathcal{E}(\rho)$  as  $\varepsilon_\tau \downarrow 0$ , and  $J$  is a nonnegative kernel with rapid decay (i.e.  $J(z) \leq |P(z)|G(z)$  for some polynomial  $P$ ) then*

$$(37) \quad \lim_{\tau \rightarrow 0} \frac{1}{\sqrt{\varepsilon_\tau}} \int_{\mathbb{R}^d} J(z) \left| \int_{\Omega} \left( \rho_1^\tau(x + \sqrt{\varepsilon_\tau} z) \rho_2^\tau(x) - \rho_1(x + \sqrt{\varepsilon_\tau} z) \rho_2(x) \right) dx \right| dz = 0$$

*Proof.* The proof of this result is the same as the one of [20, Lemma 3.7].  $\square$

Thanks to Lemma 3.4 we just need the following *pointwise* convergence result. This applies to the multiphase case and is stronger than what is needed here. Similarly to [20, Lemma 2.8, Lemma 3.6] we can formulate the following result.

**Lemma 3.5.** *Suppose that  $\rho_1, \rho_2 \in \operatorname{BV}(\Omega; \{0, 1\})$  such that  $\rho_1(x)\rho_2(x) = 0$  for a.e.  $x \in \Omega$ . Then for any smooth function  $\zeta : \Omega \rightarrow \mathbb{R}$  and any even nonnegative kernel  $J : \mathbb{R}^d \rightarrow \mathbb{R}$  with rapid decay we have*

$$\begin{aligned} \lim_{\varepsilon \rightarrow 0} \int_{\Omega} \zeta(x) [\rho_1(x)(J_\varepsilon * \rho_2(x))] dx &= \frac{1}{4} \left[ \int_{\Omega} \zeta(x) d\sigma_J(D\rho_1)(x) + \int_{\Omega} \zeta(x) d\sigma_J(D\rho_2)(x) \right] \\ &\quad - \frac{1}{4} \int_{\Omega} \zeta(x) d\sigma_J(D(\rho_1 + \rho_2))(x) \end{aligned}$$

where we define the Radon measure  $\sigma_J(D\rho_i) \in \mathcal{M}(\Omega)$  as in Definition 3.2 and we have used the notation  $J_\varepsilon(z) := J(z/\varepsilon)$ .

The proof supplied below is different from the one by Laux-Otto in [20, Lemma 2.8, Lemma 3.6]. Instead of disintegrating on  $\mathbb{R}^d$  and using one-dimensional arguments we obtain upper and lower bounds using mollifiers.

*Proof.* We begin by showing

$$\lim_{\varepsilon \rightarrow 0} \int_{\Omega} \zeta(x) [\rho_1(x)J_\varepsilon * \rho_2(x)] dx = \lim_{\varepsilon \rightarrow 0} \int_{\Omega} \frac{1}{2} \zeta(x) [\rho_1(x)J_\varepsilon * \rho_2(x) + \rho_2(x)J_\varepsilon * \rho_1(x)] dx$$

which amounts to showing that

$$\lim_{\varepsilon \rightarrow 0} \int_{\Omega} \zeta(x) [\rho_1(x)J_\varepsilon * \rho_2(x) - \rho_2(x)J_\varepsilon * \rho_1(x)] dx = 0.$$

Expanding out the convolution we have

$$\frac{1}{\varepsilon} \int_{\Omega} \int_{\mathbb{R}^d} J(z) [\rho_1(x)\rho_2(x + \varepsilon z)\zeta(x) - \rho_1(x + \varepsilon z)\rho_2(x)\zeta(x)] dz dx$$

Changing variables  $x \mapsto x - \varepsilon z$  and then  $z \mapsto -z$  in the second term of the integral we get

$$\begin{aligned} \left| \frac{1}{\varepsilon} \int_{\Omega} \int_{\mathbb{R}^d} J(z) \rho_1(x)\rho_2(x + \varepsilon z) [\zeta(x) - \zeta(x + \varepsilon z)] dz dx \right| \\ \leq \|\nabla \zeta\|_\infty \left| \int_{\Omega} \int_{\mathbb{R}^d} |z| J(z) \rho_1(x)\rho_2(x + \varepsilon z) dz dx \right| \end{aligned}$$

Thus, the dominated convergence theorem with the fact that  $\rho_1\rho_2 = 0$  a.e. yield that the quantity vanishes as  $\varepsilon \rightarrow 0$ .

Now we can restrict our attention to the limit

$$\lim_{\varepsilon \rightarrow 0} \int_{\Omega} \frac{1}{2} \zeta(x) [\rho_1(x) J_{\varepsilon} * \rho_2(x) + \rho_2(x) J_{\varepsilon} * \rho_1(x)] dx$$

Since  $\rho_i \in \text{BV}(\Omega; \{0, 1\})$  and  $\rho_1(x)\rho_2(x) = 0$  a.e.,

$$\begin{aligned} & \rho_1(x + \varepsilon z)\rho_2(x) + \rho_1(x)\rho_2(x + \varepsilon z) \\ &= \frac{1}{2} |\rho_1(x + \varepsilon z) - \rho_1(x)| + \frac{1}{2} |\rho_2(x + \varepsilon z) - \rho_2(x)| - \frac{1}{2} |(\rho_1 + \rho_2)(x + \varepsilon z) - (\rho_1 + \rho_2)(x)|, \end{aligned}$$

which follows from directly by evaluating both sides. Thus, it suffices to prove

$$(38) \quad \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} \int_{\mathbb{R}^d} J(z) \int_{\Omega} \zeta(x) |\chi(x + \varepsilon z) - \chi(x)| dx dz = \int_{\Omega} \zeta(x) d\sigma_J(\nabla\chi)(x)$$

for any  $\chi \in \text{BV}(\Omega; \{0, 1\})$ .

For  $\delta > 0$  let  $\eta_{\delta}$  be a smooth approximation to the identity, and set  $\chi_{\delta} = \eta_{\delta} * \chi$ , such that  $\chi_{\delta} \rightarrow \chi$ , as  $\delta \rightarrow 0$  in the sense of strict convergence of BV functions (see [2, Definition 3.14]).

Then, we also have

$$(39) \quad \frac{1}{\varepsilon} \int_{\mathbb{R}^d} J(z) \int_{\Omega} \zeta(x) |\chi(x + \varepsilon z) - \chi(x)| dx dz = \lim_{\delta \rightarrow 0} \int_{\mathbb{R}^d} J(z) \int_{\Omega} \zeta(x) \frac{1}{\varepsilon} |\chi_{\delta}(x + \varepsilon z) - \chi_{\delta}(x)| dx dz.$$

Let us define  $f : (0, 1) \times (0, 1) \rightarrow \mathbb{R}$  and  $f : (0, 1) \rightarrow \mathbb{R}$  as

$$\begin{aligned} f(\varepsilon, \delta) &:= \int_{\mathbb{R}^d} J(z) \int_{\Omega} \zeta(x) \frac{1}{\varepsilon} |\chi_{\delta}(x + \varepsilon z) - \chi_{\delta}(x)| dx dz = \int_{\mathbb{R}^d} J(z) \int_{\Omega} \zeta(x) \left| \int_0^1 \nabla \chi_{\delta}(x + t\varepsilon z) \cdot z dt \right| dx dz \\ &= \int_{\mathbb{R}^d} J(z) \int_{\Omega} \zeta(x) |\nabla \chi_{\delta}(x + t_{\delta}\varepsilon z) \cdot z| dx dz = \int_{\mathbb{R}^d} J(z) \int_{\Omega} \zeta(x - t_{\delta}\varepsilon z) |\nabla \chi_{\delta}(x) \cdot z| dx dz \end{aligned}$$

and

$$f(\delta) := \int_{\mathbb{R}^d} J(z) \int_{\Omega} \zeta(x) |\nabla \chi_{\delta}(x) \cdot z| dx dz,$$

where we have used the mean value theorem to find  $t_{\delta} \in [0, 1]$  in the definition of  $f(\varepsilon, \delta)$ .

*Claim.*  $f(\varepsilon, \delta) \rightarrow f(\delta)$ , uniformly w.r.t.  $\delta \in (0, 1)$  as  $\varepsilon \rightarrow 0$ .

*Proof of the Claim.* We easily check the pointwise convergence  $f(\varepsilon, \delta) \rightarrow f(\delta)$  as  $\varepsilon \rightarrow 0$ , for any  $\delta \in (0, 1)$ . Moreover,

$$\begin{aligned} |f(\varepsilon, \delta) - f(\delta)| &\leq \int_{\mathbb{R}^d} \int_{\Omega} |\zeta(x - t_{\delta}\varepsilon z) - \zeta(x)| J(z) |\nabla \chi_{\delta}(x) \cdot z| dx dz \\ &\leq \|\nabla \zeta\|_{\infty} \varepsilon \int_{\mathbb{R}^d} \int_{\Omega} J(z) |z|^2 |\nabla \chi_{\delta}(x)| dx dz \end{aligned}$$

Taking now  $\sup_{\delta}$  in the previous inequality and using the fact that  $\int_{\Omega} |\nabla \chi_{\delta}(x)| dx \rightarrow |D\chi|(\Omega)$ , as  $\delta \rightarrow 0$  by our assumption of strict convergence, yields the desired uniform convergence.

Certainly, as we have seen in (39), we also have that  $\lim_{\delta \rightarrow 0} f(\varepsilon, \delta)$  exists for any  $\varepsilon \in (0, 1)$ . Therefore, a classical Moore-Osgood type theorem yields that both

$$\lim_{\varepsilon \rightarrow 0} \lim_{\delta \rightarrow 0} f(\varepsilon, \delta) \quad \text{and} \quad \lim_{\delta \rightarrow 0} \lim_{\varepsilon \rightarrow 0} f(\varepsilon, \delta)$$

exist and they are equal. Therefore, by [2, Proposition 3.15] we conclude the limit (38) by computing  $\lim_{\delta \rightarrow 0} f(\delta)$ .  $\square$

We finish this section with the proof of Theorem 3.1.

*Proof of Theorem 3.1.* This proof is a direct consequence of Theorem 3.2 and Proposition 3.3.  $\square$

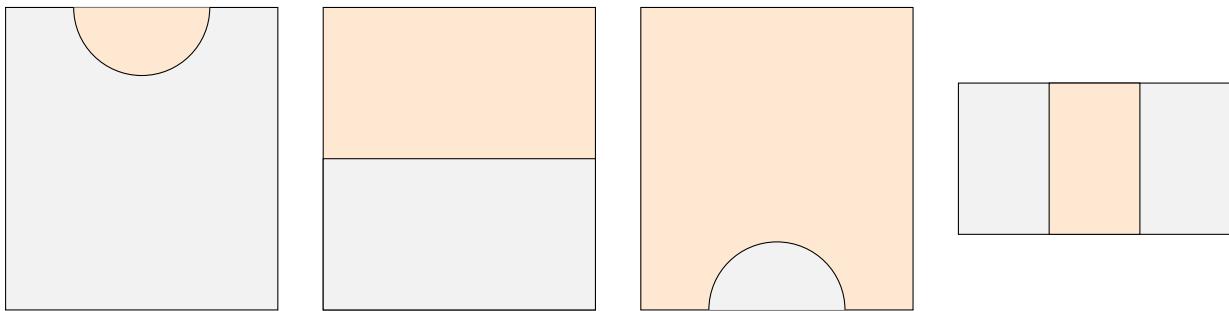


FIGURE 1. Equilibrium shapes in any dimensions, depending on the proportion of the lighter vs. the heavier fluid: orange stands for the lighter fluid, while grey stands for the heavier fluid

#### 4. A HEURISTIC DISCUSSION ON THE EQUILIBRIUM SHAPES

In this section we discuss global equilibrium of the energy with gravity potentials

$$\tilde{\mathcal{E}}_\varepsilon(\rho_1, \rho_2) := \text{HC}_\varepsilon(\rho_1, \rho_2) + \Phi(\rho_1, \rho_2).$$

where  $\Phi_i(x) = c_i x_d$ , with  $0 < c_1 < c_2$ . The order of the constants denote that  $\rho_1$  is the lighter fluid, where the vector  $-e_d$  denotes the direction of gravity. The coordinate here is  $x = (x', x_d)$  and for simplicity we consider a cylindrical domain,

$$\Sigma := \{|x'| \leq 1\} \times [0, h] = B_1^{d-1}(0) \times [0, h].$$

Here the convolution is taken with the extension of the density function as zero outside of the domain, with the density constraint  $\rho_1 + \rho_2 = 1$ ,  $\int_{\mathbb{R}^d} d\rho_i = M_i$ .

Let us mention that away from the global equilibrium, there are diverse possibilities of stationary states for  $\tilde{\mathcal{E}}_\varepsilon$  even with zero potentials. For instance any choice of characteristic functions  $\rho_1$  and  $\rho_2$  generating the interface as a disjoint union of spheres,  $\{|x - a_i| = r_i\}$ , is a stationary solution of the limit energy.

In the limit  $\varepsilon \rightarrow 0$ , the  $\Gamma$ -convergence properties indicate that the global equilibrium of the  $\varepsilon$ -energy converges to the limiting density pair  $(\rho_1, \rho_2) = (\chi_A, \chi_{A^c})$ , which is the global minimizer of the limit energy

$$E_\infty(A) := \text{Per}(A) + \int_A \Psi(x) dx, \text{ where } \Psi(x) := (\Phi_1 - \Phi_2)(x).$$

under the volume constraint  $\mathcal{L}^d(A) = M$ . Away from the domain boundary, the classical minimal surface theory yields the  $C^{2,\alpha}$  regularity of  $\partial A$  with the Euler-Lagrange equation

$$-\kappa - \Psi(x) = \lambda \text{ on } \partial A,$$

where  $\lambda$  is the Lagrange multiplier associated to the volume constraint and  $\kappa$  stands for the curvature.

When  $\partial A$  is away from the lateral and bottom portion of the cylinder, this corresponds to the classical pendant liquid droplet problem where the minimizer is known to be rotationally symmetric and convex (see [15]). When the droplet boundary touches the cylinder boundary, various shapes of drops are possible (see Figure 1) and the complete description of possibly non-smooth global minimizers appear to be open. In general, when  $\Sigma$  has  $C^{1,\alpha}$  boundary, it is shown in [37] that  $\partial A$  is  $C^{1,\alpha}$  up to the boundary and meets  $\partial\Sigma$  orthogonally. For further discussion of available results we refer to [23]. An ongoing work on numerical simulations for our flow suggest that the equilibrium states even for the  $\varepsilon$ -energy can be categorized as the ones on Figure 1. In dimension two, we could observe an additional equilibrium shape, as in Figure 2.

ACKNOWLEDGEMENT. We thank Tim Laux for the helpful discussions and references. I.K. is partially supported by NSF DMS-1566578. M.J. is partially supported by Simons Math + X Investigator - 510776, DARPA FA8750-18-2-0066, and NSF ATD-1737770.

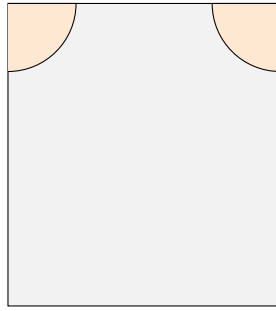


FIGURE 2. An additional equilibrium shape in dimension two

## REFERENCES

- [1] D.M. AMBROSE, Well-posedness of two-phase Hele-Shaw flow without surface tension, *European J. Appl. Math.*, 15 (2004), no. 5, 597–607.
- [2] L. AMBROSIO, N. FUSCO, D. PALLARA, *Functions of bounded variation and free discontinuity problems*, Oxford Mathematical Monographs, The Clarendon Press, Oxford University Press, New York, (2000).
- [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] G. CARLIER, C. POON, On the total variation Wasserstein gradient flow and the TV-JKO scheme, *ESAIM: COCV*, to appear.
- [5] Á. CASTRO, D. CÓRDOBA, C. FEFFERMAN, F. GANCEDO, Breakdown of smoothness for the Muskat problem, *Arch. Ration. Mech. Anal.*, 208, (2013), 3, 805–909.
- [6] Á. CASTRO, D. CÓRDOBA, C. FEFFERMAN, F. GANCEDO, M. LÓPEZ-FERNÁNDEZ, Rayleigh-Taylor breakdown for the Muskat problem with applications to water waves, *Ann. of Math. (2)*, 175 (2012), no. 2, 909–948.
- [7] P. CONSTANTIN, D. CÓRDOBA, F. GANCEDO, R.M. STRAIN, On the global existence for the Muskat problem, *J. Eur. Math. Soc. (JEMS)*, 15 (2013), no. 1, 201–227.
- [8] P. CONSTANTIN, D. CÓRDOBA, F. GANCEDO, L. RODRÍGUEZ-PIAZZA, R. M. STRAIN, On the Muskat problem: global in time results in 2D and 3D, *Amer. J. Math.*, 138 (2016), no. 6, 1455–1494.
- [9] A. CÓRDOBA, D. CÓRDOBA, F. GANCEDO, Interface evolution: the Hele-Shaw and Muskat problems, *Ann. of Math. (2)*, 173 (2011), no. 1, 477–542.
- [10] G. DE PHILIPPIS, T. LAUX, Implicit time discretization for the mean curvature flow of outward minimizing sets, *Ann. Sc. Norm. Super. Pisa Cl. Sci.*, (2019), to appear
- [11] M. DI FRANCESCO, S. FAGIOLI, Measure solutions for non-local interaction PDEs with two species, *Nonlinearity* 26 (2013), no. 10, 2777–2808.
- [12] S. ESEDOĞLU, F. OTTO, Threshold dynamics for networks with arbitrary surface tensions, *Commun. Pure Appl. Math.*, 68 (2015), 5, 808–864.
- [13] J. ESCHER, B. MATIOC, On the Parabolicity of the Muskat Problem: Well-Posedness, Fingering, and Stability Results. *Zeitschrift für Analysis und ihre Anwendungen*. 30. (2010).
- [14] M. GIAQUINTA, L. MARTINAZZI, *An introduction to the regularity theory for elliptic systems, harmonic maps and minimal graphs*, Second edition, Appunti, Scuola Normale Superiore di Pisa (Nuova Serie) [Lecture Notes. Scuola Normale Superiore di Pisa (New Series)], 11. Edizioni della Normale, Pisa, (2012).
- [15] E. GONZALEZ, Regolarità per il problema della goccia appoggiata, *Rend. Sem. Mat. Univ. Padova*, (1977), 58, 25–33.
- [16] J. H. IRVING, J. G. KIRKWOOD, The statistical mechanical theory of transport processes. IV. The equations of hydrodynamics, *J. Chem. Phys.*, (1950) 18, 817–829.
- [17] R. JORDAN, D. KINDERLEHRER, F. OTTO, The variational formulation of the Fokker-Planck equation, *SIAM J. Math. Anal.*, 29 (1998), no. 1, 1–17.
- [18] I. KIM, A.R. MÉSZÁROS, On nonlinear cross-diffusion systems: an optimal transport approach, *Calc. Var. Partial Differential Equations*, 57 (2018), no. 3, Art. 79, 40 pp.
- [19] M. LABORDE, On cross-diffusion systems for two populations subject to a common congestion effect, *Appl. Math. Optim.*, (2018), to appear.
- [20] T. LAUX, F. OTTO, Convergence of the thresholding scheme for multi-phase mean-curvature flow, *Calc. Var. Partial Differential Equations*, 55 (2016), no. 5, Art. 129, 74 pp.
- [21] S. LUCKHAUS, T. STURZENHECKER, Implicit time discretization for the mean curvature flow equation, *Calc. Var. Partial Differential Equations*, 3 (1995), no. 2, 253–271.
- [22] S. LUCKHAUS, The Stefan problem with surface tension, in *Variational and free boundary problems*, 153–157, *IMA Vol. Math. Appl.*, 53, Springer, New York, (1993).
- [23] F. MAGGI, C. MIHAILA, On the shape of capillarity droplets in a container, *Calc. Var. Partial Differential Equations* 55 (2016), no. 5, Art. 122, 42 pp.
- [24] D. MATTHES, R.J. MCCANN, G. SAVARÉ, A family of nonlinear fourth order equations of gradient flow type, *Comm. Partial Differential Equations*, 34 (2009), no. 10-12, 1352–1397.

- [25] 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.
- [26] B. MERRIMAN, J. K. BENICE, S. OSHER, Motion of multiple functions: a level set approach, *J. Comput. Phys.*, 112 (1994), no. 2, 334–363.
- [27] M. MUSKAT, Two fluid systems in porous media. The encroachment of water into an oil sand, *Physics* 5 (1934), 250–264.
- [28] F. OTTO, Dynamics of labyrinthine pattern formation in magnetic fluids: a mean-field theory, *Arch. Rational Mech. Anal.*, 141 (1998), no. 1, 63–103.
- [29] F. OTTO, The geometry of dissipative evolution equations: the porous medium equation, *Comm. Partial Differential Equations*, 26 (2001), no. 1-2, 101–174.
- [30] J. PRÜSS, G. SIMONETT, On the Muskat problem, *Evol. Equ. Control Theory*, 5 (2016), no. 4, 631–645.
- [31] M. RÖGER, Existence of weak solutions for the Mullins-Sekerka flow, *SIAM J. Math. Anal.*, 37 (2005), no. 1, 291–301.
- [32] R. ROSSI, G. SAVARÉ, Tightness, integral equicontinuity and compactness for evolution problems in Banach spaces, *Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)*, (2003) 2(2), 395–431.
- [33] F. SANTAMBROGIO, *Optimal Transport for Applied Mathematicians*, Birkhäuser, (2015).
- [34] M. SIEGEL, R. E. CAFLISCH, S. HOWISON, Global existence, singular solutions, and ill-posedness for the Muskat problem, *Comm. Pure Appl. Math.*, 57:10 (2004), 1374–1411.
- [35] H. SOHR, *The Navier-Stokes equations. An elementary functional analytic approach*, Birkhäuser Advanced Texts, Basel, (2001).
- [36] K. TAKASAO, Existence of Weak Solution for Volume-Preserving Mean Curvature Flow via Phase Field Method, *Indiana University Mathematics Journal*, 66 (2017), no. 6, 2015–2035.
- [37] J. TAYLOR, Crystalline variational problems, *Bulletin of the American Mathematical Society*, 84 (1978), no.4, 568–588.

DEPARTMENT OF MATHEMATICS, UCLA, 520 PORTOLA PLAZA, LOS ANGELES, CA 90095, USA  
*E-mail address:* `majaco@math.ucla.edu`

DEPARTMENT OF MATHEMATICS, UCLA, 520 PORTOLA PLAZA, LOS ANGELES, CA 90095, USA  
*E-mail address:* `ikim@math.ucla.edu`

DEPARTMENT OF MATHEMATICS, UCLA, 520 PORTOLA PLAZA, LOS ANGELES, CA 90095, USA  
*E-mail address:* `alpar@math.ucla.edu`