INCOMPRESSIBLE IMMISCIBLE MULTIPHASE FLOWS IN POROUS MEDIA: A VARIATIONAL APPROACH

CLÉMENT CANCÈS, THOMAS O. GALLOUËT, AND LÉONARD MONSAINGEON

Abstract. We describe the competitive motion of $(N+1)$ incompressible immiscible phases within a porous medium as the gradient flow of a singular energy in the space of non-negative measures with prescribed mass endowed with some tensorial Wasserstein distance. We show the convergence of the approximation obtained by a minimization scheme *à la* [R. Jordan, D. Kinderlehrer & F. Otto, SIAM J. Math. Anal, 29(1):1–17, 1998]. This allow to obtain a new existence result for a physically well-established system of PDEs consisting in the Darcy-Muskat law for each phase, N capillary pressure relations, and a constraint on the volume occupied by the fluid. Our study does not require the introduction of any global or complementary pressure.

Keywords. Multiphase porous media flows, Wasserstein gradient flows, constrained parabolic system, minimizing movement scheme

AMS subjects classification. 35K65, 35A15, 49K20, 76S05

1. Introduction

1.1. **Equations for multiphase flows in porous media.** We consider a connected open bounded set $\Omega \subset \mathbb{R}^d$ representing a porous medium. $N+1$ incompressible and immiscible phases, labeled by subscripts $i \in \{0, \dots, N\}$ are supposed to flow within the pores. Let us present now some classical equations that describe the motion of such a mixture. The physical justification of these equations can be found for instance in [10, Chapter 5]. We denote by $s_i : \Omega \times (0,T) =: Q \to [0,1]$ the content of the phase i, i.e., the volume ratio of the phase i compared to all the phases and the solid matrix, and by $\boldsymbol{v}_i$ the filtration speed of the phase i, then the conservation of the volume of each phase writes

$$\partial_t s_i + \boldsymbol{\nabla} \cdot (s_i \boldsymbol{v}_i) = 0 \quad \text{in } Q, \quad \forall i \in \{0, \dots, N\}, \tag{1}$$

where $T > 0$ is an arbitrary finite time horizon. The filtration speed of each phase is supposed to be given by the Darcy law

$$\boldsymbol{v}_i = -\frac{1}{\mu_i} \mathbb{K} \left(\boldsymbol{\nabla} p_i - \rho_i \mathbf{g} \right) \quad \text{in } Q, \quad \forall i \in \{0, \dots, N\}. \tag{2}$$

In the above relation, $\mathbf{g}$ is the gravity vector, μ_i denotes the viscosity of the phase i, p_i its pressure, and ρ_i its density. The intrinsic permeability tensor $\mathbb{K} : \overline{\Omega} \to \mathbb{R}^{d \times d}$ is supposed to be smooth, symmetric $\mathbb{K} = \mathbb{K}^T$, and uniformly positive definite: there exist $\kappa_\star, \kappa^\star > 0$ such that:

$$\kappa_\star |\boldsymbol{\xi}|^2 \le \mathbb{K}(\boldsymbol{x}) \boldsymbol{\xi} \cdot \boldsymbol{\xi} \le \kappa^\star |\boldsymbol{\xi}|^2, \qquad \forall \boldsymbol{\xi} \in \mathbb{R}^d, \ \forall \boldsymbol{x} \in \overline{\Omega}. \tag{3}$$

The pore volume is supposed to be saturated by the fluid mixture, i.e.,

$$\sigma := \sum_{i=0}^{N} s_i = \omega(\boldsymbol{x}) \quad \text{a.e. in } Q, \tag{4}$$

where the porosity $\omega : \overline{\Omega} \to (0,1)$ of the porous medium is assumed to be smooth. In particular, there exists $0 < \omega_\star \le \omega^\star$ such that $\omega_\star \le \omega(\boldsymbol{x}) \le \omega^\star$ for all $\boldsymbol{x} \in \overline{\Omega}$. In what follows, we denote by $\boldsymbol{s} = (s_0, \dots, s_N)$, by

$$\Delta(\boldsymbol{x}) = \left\{ \boldsymbol{s} \in (\mathbb{R}_+)^{N+1} \,\middle|\, \sum_{i=0}^{N} s_i = \omega(\boldsymbol{x}) \right\},$$

and by

$$\boldsymbol{\mathcal{X}} = \left\{ \boldsymbol{s} \in L^1(\Omega; \mathbb{R}_+^{N+1}) \;\middle|\; \boldsymbol{s}(\boldsymbol{x}) \in \Delta(\boldsymbol{x}) \text{ a.e. in } \Omega \right\}.$$

There is an obvious one-to-one mapping between the sets $\Delta(\boldsymbol{x})$ and

$$\Delta^*(\boldsymbol{x}) = \left\{ \boldsymbol{s}^* = (s_1, \dots, s_N) \in (\mathbb{R}_+)^N \,\middle|\, \sum_{i=1}^{N} s_i \le \omega(\boldsymbol{x}) \right\},$$

and consequently also between $\boldsymbol{\mathcal{X}}$ and

$$\boldsymbol{\mathcal{X}}^* = \left\{ \boldsymbol{s}^* \in L^1(\Omega; \mathbb{R}_+^{N}) \;\middle|\; \boldsymbol{s}^*(\boldsymbol{x}) \in \Delta^*(\boldsymbol{x}) \text{ a.e. in } \Omega \right\}.$$

In what follows, we denote by $\boldsymbol{\Upsilon} = \bigcup_{\boldsymbol{x} \in \overline{\Omega}} \Delta^*(\boldsymbol{x}) \times \{\boldsymbol{x}\}$.

In order to close the system, we impose N capillary pressure relations

$$p_i - p_0 = \pi_i(\boldsymbol{s}^*, \boldsymbol{x}) \quad \text{a.e in } Q, \quad \forall i \in \{1, \dots, N\}, \tag{5}$$

where the capillary pressure functions $\pi_i : \boldsymbol{\Upsilon} \to \mathbb{R}$ are assumed to be continuously differentiable and to derive from a strictly convex potential $\Pi : \boldsymbol{\Upsilon} \to \mathbb{R}_+$:

$$\pi_i(\boldsymbol{s}^*, \boldsymbol{x}) = \frac{\partial \Pi}{\partial s_i}(\boldsymbol{s}^*, \boldsymbol{x}) \qquad \forall i \in \{1, \dots, N\}. \tag{6}$$

We assume that Π is uniformly convex w.r.t. its first variable. More precisely, we assume that there exist two positive constants $\varpi_\star$ and $\varpi^\star$ such that, for all $\boldsymbol{x} \in \overline{\Omega}$ and all $\boldsymbol{s}^*, \widehat{\boldsymbol{s}}^* \in \Delta^*(\boldsymbol{x})$, one has

$$\frac{\varpi^\star}{2} \left| \widehat{\boldsymbol{s}}^* - \boldsymbol{s}^* \right|^2 \ge \Pi(\widehat{\boldsymbol{s}}^*, \boldsymbol{x}) - \Pi(\boldsymbol{s}^*, \boldsymbol{x}) - \boldsymbol{\pi}(\boldsymbol{s}^*, \boldsymbol{x}) \cdot (\widehat{\boldsymbol{s}}^* - \boldsymbol{s}^*) \ge \frac{\varpi_\star}{2} \left| \widehat{\boldsymbol{s}}^* - \boldsymbol{s}^* \right|^2, \tag{7}$$

where we introduced the notation

$$\boldsymbol{\pi} : \begin{cases} \boldsymbol{\Upsilon} \to \mathbb{R}^N \\ (\boldsymbol{s}^*, \boldsymbol{x}) \mapsto \boldsymbol{\pi}(\boldsymbol{s}^*, \boldsymbol{x}) = (\pi_1(\boldsymbol{s}^*, \boldsymbol{x}), \dots, \pi_N(\boldsymbol{s}^*, \boldsymbol{x})). \end{cases}$$

The relation (7) implies that $\boldsymbol{\pi}$ is monotone and injective w.r.t. its first variable. Denoting by

$$\boldsymbol{z} \mapsto \boldsymbol{\phi}(\boldsymbol{z}, \boldsymbol{x}) = (\phi_1(\boldsymbol{z}, \boldsymbol{x}), \dots, \phi_N(\boldsymbol{z}, \boldsymbol{x})) \in \Delta^*(\boldsymbol{x})$$

the inverse of $\boldsymbol{\pi}(\cdot, \boldsymbol{x})$, it follows from (7) that

$$0 < \frac{1}{\varpi^\star} \le \mathbb{J}_{\boldsymbol{z}} \boldsymbol{\phi}(\boldsymbol{z}, \boldsymbol{x}) \le \frac{1}{\varpi_\star} \qquad \text{for all } \boldsymbol{x} \in \overline{\Omega} \text{ and all } \boldsymbol{z} \in \boldsymbol{\pi}(\Delta^*(\boldsymbol{x}), \boldsymbol{x}), \tag{8}$$

where $\mathbb{J}_{\boldsymbol{z}}$ stands for the Jacobian with respect to $\boldsymbol{z}$, and where the above inequality has to be understood in the sense of positive definite matrices. Moreover, due to the regularity of $\boldsymbol{\pi}$ w.r.t. the space variable, there exists $M_{\boldsymbol{\phi}} > 0$ such that

$$|\boldsymbol{\nabla}_{\boldsymbol{x}} \boldsymbol{\phi}(\boldsymbol{z}, \boldsymbol{x})| \le M_{\boldsymbol{\phi}} \qquad \text{for all } \boldsymbol{x} \in \overline{\Omega} \text{ and all } \boldsymbol{z} \in \boldsymbol{\pi}(\Delta^*(\boldsymbol{x}), \boldsymbol{x}), \tag{9}$$

where $\boldsymbol{\nabla}_{\boldsymbol{x}}$ denote the gradient w.r.t. to the second variable only.

The problem is complemented with no-flux boundary conditions

$$\boldsymbol{v}_i \cdot \boldsymbol{n} = 0 \quad \text{on } \partial\Omega \times (0,T), \quad \forall i \in \{0,\dots,N\}, \tag{10}$$

and by the initial content profile $\boldsymbol{s}^0 = \left(s_0^0, \dots, s_N^0\right) \in \boldsymbol{\mathcal{X}}$:

$$s_i(\cdot, 0) = s_i^0 \quad \forall i \in \{0,\dots,N\}, \quad \text{with} \quad \sum_{i=0}^N s_i^0 = \omega \text{ a.e. in } \Omega. \tag{11}$$

Since we did not consider sources, and since we imposed no-flux boundary conditions, the volume of each phase is conserved along time

$$\int_\Omega s_i(\boldsymbol{x},t)\mathrm{d}\boldsymbol{x} = \int_\Omega s_i^0(\boldsymbol{x})\mathrm{d}\boldsymbol{x} =: m_i > 0, \qquad \forall i \in \{0,\dots,N\}. \tag{12}$$

We can now give a proper definition of what we call a weak solution to the problem (1)–(2), (4)–(5), and (10)–(11).

Definition 1.1 (Weak solution)**.** *A measurable function $\boldsymbol{s} : Q \to (\mathbb{R}_+)^{N+1}$ is said to be a weak solution if $\boldsymbol{s} \in \Delta$ a.e. in Q, if there exists $\boldsymbol{p} = (p_0, \dots, p_N) \in L^2((0,T);H^1(\Omega))^{N+1}$ such that the relations (5) hold, and such that, for all $\phi \in C_c^\infty(\overline{\Omega} \times [0,T))$ and all $i \in \{0,\dots,N\}$, one has*

$$\iint_Q s_i \partial_t \phi \mathrm{d}\boldsymbol{x}\mathrm{d}t + \int_\Omega s_i^0 \phi(\cdot,0)\mathrm{d}\boldsymbol{x} - \iint_Q \frac{s_i}{\mu_i}\mathbb{K}\left(\boldsymbol{\nabla} p_i - \rho_i \mathbf{g}\right)\cdot \boldsymbol{\nabla}\phi \mathrm{d}\boldsymbol{x}\mathrm{d}t = 0. \tag{13}$$

1.2. **Wasserstein gradient flow of the energy.**

1.2.1. *Energy of a configuration.*

First, we extend the convex function $\Pi : \boldsymbol{\Upsilon} \to [0,+\infty]$, called *capillary energy density*, into a convex function (still denoted by) $\Pi : \mathbb{R}^{N+1} \times \overline{\Omega} \to [0,+\infty]$ by setting

$$\Pi(\boldsymbol{s},\boldsymbol{x}) = \begin{cases} \Pi\left(\omega \frac{\boldsymbol{s}^*}{\sigma}, \boldsymbol{x}\right) = \Pi\left(\omega\frac{s_1}{\sigma}, \dots, \omega\frac{s_N}{\sigma}, \boldsymbol{x}\right) & \text{if } \boldsymbol{s} \in \mathbb{R}_+^{N+1} \text{ and } \sigma \le \omega(\boldsymbol{x}), \\ +\infty & \text{otherwise,} \end{cases}$$

σ being defined by (4). The extension of Π by $+\infty$ where $\sigma > \omega$ is natural because of the incompressibility of the fluid mixture. The extension on $\{\sigma < \omega\} \cup \mathbb{R}_+^{N+1}$ has been designed so that the energy density only depends on the composition of the fluid mixture. However, this extension is somehow arbitrary, and, as it will appear in what follows, it has no influence on the flow since the solution $\boldsymbol{s}$ remains in $\boldsymbol{\mathcal{X}}$. In our previous note [17], the appearance of void $\sigma < \omega$ was directly prohibited by a penalization in the energy.

The second part in the energy comes from the gravity. In order to lighten the notations, we introduce the functions

$$\Psi_i : \left\{ \begin{array}{ccc} \overline{\Omega} & \to & \mathbb{R}_+, \\ \boldsymbol{x} & \mapsto & -\rho_i \mathbf{g}\cdot\boldsymbol{x}, \end{array}\right. \qquad \forall i \in \{0,\dots,N\}, \tag{14}$$

and

$$\boldsymbol{\Psi} : \left\{ \begin{array}{ccc} \overline{\Omega} & \to & \mathbb{R}_+^{N+1}, \\ \boldsymbol{x} & \mapsto & \left(\Psi_0(\boldsymbol{x}), \dots, \Psi_N(\boldsymbol{x})\right). \end{array}\right.$$

The fact that Ψ_i can be supposed to be positive come from the fact that Ω is bounded. Even though the physically relevant potentials are indeed the gravitational $\Psi_i(\boldsymbol{x}) = -\rho_i \mathbf{g} \cdot \boldsymbol{x}$, the subsequent analysis allows for a broader class of external potentials and for the sake of generality we shall therefore consider arbitrary $\Psi_i \in \mathcal{C}^1(\overline{\Omega})$ in the sequel.

We can now define the convex energy functional $\mathcal{E} : L^1(\Omega, \mathbb{R}^{N+1}) \to \mathbb{R} \cup \{+\infty\}$ by adding the capillary energy to the gravitational one:

$$\mathcal{E}(\boldsymbol{s}) = \int_\Omega (\Pi(\boldsymbol{s}, \boldsymbol{x}) + \boldsymbol{s} \cdot \boldsymbol{\Psi}) \, \mathrm{d}\boldsymbol{x} \geq 0, \qquad \forall \boldsymbol{s} \in L^1(\Omega; \mathbb{R}^{N+1}). \tag{15}$$

Note moreover that $\mathcal{E}(\boldsymbol{s}) < \infty$ iff $\boldsymbol{s} \geq 0$ and $\sigma \leq \omega$ a.e. in Ω. It follows from the conservation of the total mass (12) that

$$\int_\Omega \sigma(\boldsymbol{x}) \mathrm{d}\boldsymbol{x} = \sum_{i=0}^N m_i = \int_\Omega \omega(\boldsymbol{x}) \mathrm{d}\boldsymbol{x}.$$

Assume that there exists a non-negligible subset A of Ω such that $\sigma < \omega$ on A, then necessarily, there must be a non-negligible subset B of Ω such that $\sigma > \omega$ so that the above equation holds, hence $\mathcal{E}(\boldsymbol{s}) = +\infty$. Therefore,

$$\mathcal{E}(\boldsymbol{s}) < \infty \quad \Leftrightarrow \quad \boldsymbol{s} \in \boldsymbol{\mathcal{X}}. \tag{16}$$

Let $\boldsymbol{p} = (p_0, \dots, p_N) : \Omega \to \mathbb{R}^{N+1}$ be such that $\boldsymbol{p} \in \partial_{\boldsymbol{s}} \Pi(\boldsymbol{s}, \boldsymbol{x})$ for a.e. $\boldsymbol{x}$ in Ω, then, defining $h_i = p_i + \Psi_i(\boldsymbol{x})$ for all $i \in \{0, \dots, N\}$ and $\boldsymbol{h} = (h_i)_{0 \leq i \leq N}$, $\boldsymbol{h}$ belongs to the subdifferential $\partial_{\boldsymbol{s}} \mathcal{E}(\boldsymbol{s})$ of $\mathcal{E}$ at $\boldsymbol{s}$, i.e.,

$$\mathcal{E}(\widehat{\boldsymbol{s}}) \geq \mathcal{E}(\boldsymbol{s}) + \sum_{i=0}^N \int_\Omega h_i (\widehat{s}_i - s_i) \mathrm{d}\boldsymbol{x}, \qquad \forall \widehat{\boldsymbol{s}} \in L^1(\Omega; \mathbb{R}^{N+1}).$$

The reverse inclusion also holds, hence

$$\partial_{\boldsymbol{s}} \mathcal{E}(\boldsymbol{s}) = \left\{ \boldsymbol{h} : \Omega \to \mathbb{R}^{N+1} \;\middle|\; h_i - \Psi_i(\boldsymbol{x}) \in \partial_{\boldsymbol{s}} \Pi(\boldsymbol{s}, \boldsymbol{x}) \text{ for a.e. } \boldsymbol{x} \in \Omega \right\}. \tag{17}$$

Thanks to (16), we know that a configuration $\boldsymbol{s}$ has a finite energy iff $\boldsymbol{s} \in \boldsymbol{\mathcal{X}}$. Since we are interested in finite energy configurations, it is relevant to consider the restriction of $\mathcal{E}$ to $\boldsymbol{\mathcal{X}}$. Then using the one-to-one mapping between $\boldsymbol{\mathcal{X}}$ and $\boldsymbol{\mathcal{X}}^*$, we define the energy of a configuration $\boldsymbol{s}^* \in \boldsymbol{\mathcal{X}}^*$, that we denote by $\mathcal{E}(\boldsymbol{s}^*)$ by setting $\mathcal{E}(\boldsymbol{s}^*) = \mathcal{E}(\boldsymbol{s})$ where $\boldsymbol{s}$ is the unique element of $\boldsymbol{\mathcal{X}}$ corresponding to $\boldsymbol{s}^* \in \boldsymbol{\mathcal{X}}^*$.

1.2.2. *Geometry of Ω and Wasserstein distance.* Inspired by the paper of Lisini [34], where heterogeneous anisotropic degenerate parabolic equations are studied from a variational point of view, we introduce $(N+1)$ distances on Ω that takes the permeability of the porous medium and the phase viscosities into account. Given two points $\boldsymbol{x}, \boldsymbol{y}$ in $\overline{\Omega}$, we denote by

$$P(\boldsymbol{x}, \boldsymbol{y}) = \left\{ \boldsymbol{\gamma} \in C^1([0,1]; \overline{\Omega}) \,\middle|\, \boldsymbol{\gamma}(0) = \boldsymbol{x} \text{ and } \boldsymbol{\gamma}(1) = \boldsymbol{y} \right\}$$

the set of the smooth paths joining $\boldsymbol{x}$ to $\boldsymbol{y}$, and we introduce distances d_i, $i \in \{0, \dots, N\}$ between elements on $\overline{\Omega}$ by setting

$$d_i(\boldsymbol{x}, \boldsymbol{y}) = \inf_{\boldsymbol{\gamma} \in P(\boldsymbol{x}, \boldsymbol{y})} \left(\int_0^1 \mu_i \mathbb{K}^{-1}(\boldsymbol{\gamma}(\tau)) \boldsymbol{\gamma}'(\tau) \cdot \boldsymbol{\gamma}'(\tau) \mathrm{d}\tau \right)^{1/2}, \qquad \forall (\boldsymbol{x}, \boldsymbol{y}) \in \overline{\Omega}^2. \tag{18}$$

It follows from (3) that

$$\sqrt{\frac{\mu_i}{\kappa^\star}}|\boldsymbol{x}-\boldsymbol{y}| \le d_i(\boldsymbol{x},\boldsymbol{y}) \le \sqrt{\frac{\mu_i}{\kappa_\star}}|\boldsymbol{x}-\boldsymbol{y}|, \qquad \forall(\boldsymbol{x},\boldsymbol{y})\in\overline{\Omega}^2. \tag{19}$$

Since $\mathbb{K}^{-1}$ is smooth, at least $C^2(\overline{\Omega})$, the Ricci curvature of (Ω, d_i) is uniformly bounded, i.e., there exists C depending only on $(\mu_i)_{0\le i\le N}$ and $\mathbb{K}$ such that

$$|\mathrm{Ric}_{\boldsymbol{x}}(\boldsymbol{v})| \le C\mu_i\mathbb{K}^{-1}\boldsymbol{v}\cdot\boldsymbol{v}, \qquad \forall\boldsymbol{x}\in\Omega,\ \forall\boldsymbol{v}\in\mathbb{R}^d. \tag{20}$$

We refer to [44, Chap. 14] for further details on the Ricci curvature and its links with optimal transportation.

Let $i\in\{0,\dots,N\}$, then define

$$\mathcal{A}_i = \left\{ s_i \in L^1(\Omega;\mathbb{R}_+) \;\middle|\; \int_\Omega s_i \mathrm{d}\boldsymbol{x} = m_i \right\}.$$

Let $s_i, \widehat{s}_i$ be two elements of $\mathcal{A}_i$, the set of the transport plans between s_i and $\widehat{s}_i$ is given by

$$\Gamma_i(s_i,\widehat{s}_i) = \left\{ \theta_i \in \mathcal{M}_+(\Omega\times\Omega) \;\middle|\; \theta_i(\Omega\times\Omega) = m_i,\ \theta_i^{(1)} = s_i \text{ and } \theta_i^{(2)} = \widehat{s}_i \right\},$$

where $\mathcal{M}_+(\Omega\times\Omega)$ stands for the set of Borel measures on $\Omega\times\Omega$, and where $\theta_i^{(k)}$ is the k^{th} marginal of the measure θ_i. We define the quadratic Wasserstein distance W_i of $\mathcal{A}_i$ by setting

$$W_i(s_i,\widehat{s}_i) = \left(\inf_{\theta_i\in\Gamma(s_i,\widehat{s}_i)} \iint_{\Omega\times\Omega} d_i(\boldsymbol{x},\boldsymbol{y})^2 \mathrm{d}\theta_i(\boldsymbol{x},\boldsymbol{y}) \right)^{1/2}. \tag{21}$$

Due to the permeability tensor $\mathbb{K}(\boldsymbol{x})$, the porous medium Ω might be heterogeneous and anisotropic. Therefore, some directions and areas might me privileged by the fluid motions. This is encoded in the distances d_i we put on Ω. Moreover, the more the phase is viscous, the more costly are its displacements, hence the μ_i in the definition (18) of d_i. But it follows from (19) that

$$\sqrt{\frac{\mu_i}{\kappa^\star}}W_{\text{ref}}(s_i,\widehat{s}_i) \le W_i(s_i,\widehat{s}_i) \le \sqrt{\frac{\mu_i}{\kappa_\star}}W_{\text{ref}}(s_i,\widehat{s}_i)., \qquad \forall s_i,\widehat{s}_i\in\mathcal{A}_i, \tag{22}$$

where W_{ref} denotes the classical quadratic Wasserstein distance defined by

$$W_{\text{ref}}(s_i,\widehat{s}_i) = \left(\inf_{\theta_i\in\Gamma(s_i,\widehat{s}_i)} \iint_{\Omega\times\Omega} |\boldsymbol{x}-\boldsymbol{y}|^2 \mathrm{d}\theta_i(\boldsymbol{x},\boldsymbol{y}) \right)^{1/2}. \tag{23}$$

Once the phase Wasserstein distances $(W_i)_{0\le i\le N}$ at hand, we can define the global Wasserstein distance $\boldsymbol{W}$ on $\boldsymbol{\mathcal{A}} := \mathcal{A}_0\times\cdots\times\mathcal{A}_N$ by setting

$$\boldsymbol{W}(\boldsymbol{s},\widehat{\boldsymbol{s}}) = \left(\sum_{i=0}^N W_i(s_i,\widehat{s}_i)^2 \right)^{1/2}, \quad \forall\boldsymbol{s},\widehat{\boldsymbol{s}}\in\boldsymbol{\mathcal{A}}.$$

1.2.3. *Gradient flow of the energy.* The content of this section is formal. Our aim is to write the problem as a gradient flow, i.e.

$$\frac{\mathrm{d}\boldsymbol{s}}{\mathrm{d}t} \in -\mathbf{grad}_{\boldsymbol{W}}\mathcal{E}(\boldsymbol{s}) = -\left(\mathrm{grad}_{W_0}\mathcal{E}(\boldsymbol{s}), \dots, \mathrm{grad}_{W_N}\mathcal{E}(\boldsymbol{s})\right) \tag{24}$$

where $\mathbf{grad}_{\boldsymbol{W}}\mathcal{E}(\boldsymbol{s})$ denotes the Wasserstein gradient of $\mathcal{E}(\boldsymbol{s})$, and where grad_{W_i} denote the gradient of $s_i \mapsto \mathcal{E}(\boldsymbol{s})$ for the Wasserstein distance W_i. The Wasserstein distance W_i was built so that $\dot{\boldsymbol{s}} = (\dot{s}_i)_i \in \mathbf{grad}_{\boldsymbol{W}}\mathcal{E}(\boldsymbol{s})$ iff there exists $\boldsymbol{h} \in \partial_{\boldsymbol{s}}\mathcal{E}(\boldsymbol{s})$ such that

$$\partial_t s_i = -\boldsymbol{\nabla} \cdot \left(s_i \frac{\mathbb{K}}{\mu_i} \boldsymbol{\nabla} h_i \right), \qquad \forall i \in \{0, \dots, N\}.$$

Such a construction was already performed by Lisini in the case of a single equation. Owing to the definitions (15) and (17) of the energy $\mathcal{E}(\boldsymbol{s})$ and its subdifferential $\partial_{\boldsymbol{s}}\mathcal{E}(\boldsymbol{s})$, the partial differential equations can be (at least formally) recovered. This was roughly speaking to purpose of our note [17].

In order to define rigorously the gradient $\mathbf{grad}_{\boldsymbol{W}}\mathcal{E}$ in (24), $\boldsymbol{\mathcal{A}}$ has to be a Riemannian manifold. The so-called Otto's calculus (see [40] and [44, Chapter 15]) allows to put a formal Riemannian structure on $\boldsymbol{\mathcal{A}}$. But as far as we know, this structure cannot be made rigorous and $\boldsymbol{\mathcal{A}}$ is a mere metric space. This leads us to consider generalized gradient flows in metric spaces (cf. [5]). We won't go deep into details in this direction, but we will prove that weak solutions can be obtained as limits of a minimizing movement scheme presented in the next section. This characterizes the gradient flow structure of the problem.

1.3. **Minimizing movement scheme and main result.**

1.3.1. *The scheme and existence of a solution.* For a fixed time-step $\tau > 0$, the so-called minimizing movement scheme [24, 5] or JKO scheme [30] consists in computing recursively $(\boldsymbol{s}^n)_{n\geq 1}$ as the solution to the minimization problem

$$\boldsymbol{s}^n = \underset{\boldsymbol{s}\in\boldsymbol{\mathcal{A}}}{\mathrm{argmin}} \left(\frac{\boldsymbol{W}(\boldsymbol{s}, \boldsymbol{s}^{n-1})^2}{2\tau} + \mathcal{E}(\boldsymbol{s}) \right), \tag{25}$$

the initial data $\boldsymbol{s}^0$ being given (11).

1.3.2. *Approximate solution and main result.* Anticipating that the JKO scheme (25) is well posed (this is the purpose of Proposition 2.1 below), we can now define the piecewise constant w.r.t. time approximate solution $\boldsymbol{s}^\tau \in L^\infty((0,T); \boldsymbol{\mathcal{X}} \cap \boldsymbol{\mathcal{A}})$ by

$$\boldsymbol{s}^\tau(0,\cdot) = \boldsymbol{s}^0, \quad \text{and} \quad \boldsymbol{s}^\tau(t,\cdot) = \boldsymbol{s}^n \quad \forall t \in ((n-1)\tau, n\tau],\ \forall n \geq 1. \tag{26}$$

The main theorem of our paper is the following.

Theorem 1.2. *Let $(\tau_k)_{k\geq 1}$ be a sequence of time steps tending to 0, then there exists one weak solution $\boldsymbol{s}$ in the sense of Definition 1.1 such that, up to an unlabeled subsequence, $(\boldsymbol{s}^{\tau_k})_{k\geq 1}$ converges a.e. in Q towards s as k tends to ∞.*

As a direct by-product of Theorem 1.2, the continuous problem admits (at least) one solution in the sense of Definition 1.1. As far as we know, this existence result is new.

Remark 1.3. *It is worth stressing that our final solution will satisfy* a posteriori $\partial_t s_i \in L^2((0,T); H^1(\Omega)')$, $s_i \in L^2((0,T); H^1(\Omega))$, *and thus* $s_i \in \mathcal{C}([0,T]; L^2(\Omega))$. *This regularity is enough to retrieve the so-called* Energy-Dissipation-Equality

$$\frac{d}{dt}\mathcal{E}(\boldsymbol{s}(t)) = -\sum_{i=0}^{N}\int_\Omega \mathbb{K}\frac{s_i(t)}{\mu_i}\nabla(p_i(t)+\Psi_i)\cdot\nabla(p_i(t)+\Psi_i)\mathrm{d}\boldsymbol{x} \leq 0 \quad \textit{for a.e. } t \in (0,T),$$

which is another admissible formulation of gradient flows in metric spaces [5].

1.4. **Goal and positioning of the paper.** The aims of the paper are twofolds. First, we aim to provide rigorous foundations to the formal variational approach exposed in the authors' recent note [17]. It gives new insights for the modeling of complex porous media flows and their numerical approximation. It appears to be very natural since only physically motivated quantities appear in the study. Indeed, we manage to avoid the introduction of the so-called Kirchhoff transform and global pressure that classically appear in the mathematical study of multiphase flows in porous media (see for instance [19, 9, 21, 26, 27, 23, 20, 2, 3]).

Second, the existence result that we deduce from the convergence of the variational scheme is new as soon as there are at least three phases ($N \geq 2$). Indeed, since our study does not require the introduction of a global pressure, we get rid of many structural assumptions on the data among which the so-called *total differentiability condition*, see for instance Assumption **(H3)** in the paper by Fabrie and Saad [26]. This structural condition is not naturally satisfied by the models, and suitable algorithms have to be employed in order to adapt the data to this constraint [22]. However, our approach suffers from another technical difficulty: we are stuck to the case of linear relative permeabilities. The extension to the case of nonlinear concave relative permeabilities, i.e., where (1) is replaced by

$$\partial_t s_i + \boldsymbol{\nabla}\cdot(k_i(s_i)\boldsymbol{v}_i) = 0,$$

may be reachable thanks to the contribution of Dolbeault, Nazaret, and Savaré [25], but we did not push into this direction since the relative permeabilities k_i are in general supposed to be convex in models coming from engineering.

Since the seminal paper of Jordan, Kinderlehrer, and Otto [30], gradient flows in metric spaces (and particularly in the space of probability measures endowed with the quadratic Wasserstein distance) were the object of many studies. Let us for instance refer to the monograph of Ambrosio, Gigli, and Savaré [5] and to Villani's book [44, Part II] for a complete overview. Applications are numerous. We refer for instance to [39] for an application to magnetic fluids, to [41, 7, 6] for applications to supra-conductivity, to [12, 11, 45] for applications to chemotaxis, to [35] for phase field models, to [37] for a macroscopic model of crowd motion, to [13] for an application to granular media, to [18] for aggregation equations, or to [31] for a model of ionic transport that applies in semi-conductors. In the context of porous media flows, this framework has been used by Otto [40] to study the asymptotic behavior of the porous medium equation, that is a simplified model for the filtration of a gaz in a porous medium. The gradient flow approach in Wasserstein metric spaces was used more recently by Laurençot and Matioc [33] on a thin film approximation model for two-phase flows in porous media. The variational structure of the system governing incompressible immiscible two-phase flows in porous media was recently depicted by the authors in their note [17]. Whereas the purpose of [17] is formal, our goal is here to give a rigorous foundation to the variational approach for complex flows in porous media. Finally, let us mention the work of Gigli and Otto [28] where it was noticed that multiphase linear transportation with saturation constraint (as we have here thanks to (1) and (4)) yields nonlinear transport with mobilities that appear naturally in the two-phase flow context.

The paper is organized as follows. In Section 2, we derive estimates on the solution $\boldsymbol{s}^\tau$ for a fixed τ. Beyond the classical energy and distance estimates explicited in §2.1, we obtain enhanced estimates thanks to an adaptation of the so-called *flow interchange* technique of Matthes, McCann, and Savaré [36] to our inhomogeneous context in §2.2. Because of the constraint on the pore volume (4), the auxiliary flow we use is no longer the heat flow, and a drift term has to be added. An important effort is then done in §3 to derive the Euler-Lagrange equations that follow from the optimality of $\boldsymbol{s}^n$. Our proof is inspired from the work of Maury, Roudneff-Chupin, and Santambrogio [37]. It relies on an intensive use of the dual characterization of the optimal transportation problem and the corresponding Kantorovitch potentials. However, additional difficulties come from the multiphase aspect of our problem, in particular when there are at least three phases (i.e., $N \geq 2$). These are overpassed using a minimax theorem and computing the associated Lagrange multipliers in §3.1. This key step then allows to define the notion of discrete phase and capillary pressures in §3.2. Then Section 4 is devoted to the convergence of the approximate solutions $(\boldsymbol{s}^{\tau_k})_k$ towards a weak solution $\boldsymbol{s}$ as τ_k tends to 0. The estimates we obtained in Section 2 are integrated w.r.t. time in §4.1. In §4.2, we show that these estimates are sufficient to enforce the relative compactness of $(\boldsymbol{s}^{\tau_k})_k$ in the strong $L^1(Q)^{N+1}$ topology. Finally, it is shown in §4.3 that any limit $\boldsymbol{s}$ of $(\boldsymbol{s}^{\tau_k})_k$ is a weak solution in the sense of Definition 1.1.

2. One-step regularity estimates

The first thing to do is to show that the JKO scheme (25) is well-posed. This is the purpose of the following Proposition.

Proposition 2.1. *Let $n \geq 1$ and $\boldsymbol{s}^{n-1} \in \boldsymbol{\mathcal{X}} \cap \boldsymbol{\mathcal{A}}$, then there exists a unique solution $\boldsymbol{s}^n$ to the scheme (25). Moreover, one has $\boldsymbol{s}^n \in \boldsymbol{\mathcal{X}} \cap \boldsymbol{\mathcal{A}}$.*

Proof. Any $\boldsymbol{s}^{n-1} \in \boldsymbol{\mathcal{X}} \cap \boldsymbol{\mathcal{A}}$ has finite energy thanks to (16). Let $(\boldsymbol{s}^{n,k})_k \subset \boldsymbol{\mathcal{A}}$ be a minimizing sequence in (25). Testing $\boldsymbol{s}^{n-1}$ in (25) it is easy to see that $\mathcal{E}(\boldsymbol{s}^{n,k}) \leq \mathcal{E}(\boldsymbol{s}^{n-1}) < \infty$ for large k, thus $(\boldsymbol{s}^{n,k})_k \subset \boldsymbol{\mathcal{X}} \cap \boldsymbol{\mathcal{A}}$ thanks to (16). Hence, one has $0 \leq s_i^{n,k}(\boldsymbol{x}) \leq \omega(\boldsymbol{x})$ for all k. By Dunford-Pettis' theorem, we can therefore assume that $s_i^{n,k} \rightharpoonup s_i^n$ weakly in $L^1(\Omega)$. It is then easy to check that the limit $\boldsymbol{s}^n$ of $\boldsymbol{s}^{n,k}$ belongs to $\boldsymbol{\mathcal{X}} \cap \boldsymbol{\mathcal{A}}$. The lower semi-continuity of the Wasserstein distance with respect to weak L^1 convergence is well known (see, e.g., [42, Prop. 7.4]), and since the energy functional is convex thus l.s.c., we conclude that $\boldsymbol{s}^n$ is indeed a minimizer. Uniqueness follows from the strict convexity of the energy as well as from the convexity of the Wasserstein distances (w.r.t. linear interpolation $\boldsymbol{s}_\theta = (1-\theta)\boldsymbol{s}_0 + \theta \boldsymbol{s}_1$). ☐

The rest of this section is devoted to improving the regularity of the successive minimizers.

2.1. **Energy and distance estimates.** Testing $\boldsymbol{s} = \boldsymbol{s}^{n-1}$ in (25) we obtain

$$\frac{\boldsymbol{W}(\boldsymbol{s}^n, \boldsymbol{s}^{n-1})^2}{2\tau} + \mathcal{E}(\boldsymbol{s}^n) \le \mathcal{E}(\boldsymbol{s}^{n-1}), \tag{27}$$

As a consequence we have the monotonicity

$$\ldots \le \mathcal{E}(\boldsymbol{s}^n) \le \mathcal{E}(\boldsymbol{s}^{n-1}) \le \ldots \le \mathcal{E}(\boldsymbol{s}^0) < \infty$$

at the discrete level, thus $\boldsymbol{s}^n \in \boldsymbol{\mathcal{X}}$ for all $n \ge 0$ thanks to (16). Summing (27) over n we also obtain the classical *total square distance* estimate

$$\frac{1}{\tau} \sum_{n \ge 0} \boldsymbol{W}^2(\boldsymbol{s}^{n+1}, \boldsymbol{s}^n) \le 2\mathcal{E}(\boldsymbol{s}^0) \le C\left(\Omega, \Pi, \boldsymbol{\Psi}\right), \tag{28}$$

the last inequality coming from the fact that $\boldsymbol{s}^0$ is uniformly bounded since it belongs to $\boldsymbol{\mathcal{X}}$, thus so is $\mathcal{E}(\boldsymbol{s}^0)$. By the Cauchy-Schwartz inequality, this readily gives the approximate 1/2-Hölder estimate

$$\boldsymbol{W}(\boldsymbol{s}^{n_1}, \boldsymbol{s}^{n_2}) \le C\sqrt{(|n_2 - n_1|\tau + 1)\tau}. \tag{29}$$

2.2. **Flow interchange, entropy estimate and enhanced regularity.** The goal of this section is to obtain some additional Sobolev regularity on the capillary pressure field $\boldsymbol{\pi}(\boldsymbol{s}^{n*}, \boldsymbol{x})$, where $\boldsymbol{s}^{n*}$ is the unique element of $\boldsymbol{\mathcal{X}}^*$ corresponding to the minimizer $\boldsymbol{s}^n$ of (25). In what follows, we denote by

$$\pi_i^n : \left\{ \begin{array}{ccl} \Omega & \to & \mathbb{R}, \\ \boldsymbol{x} & \mapsto & \pi_i(\boldsymbol{s}^{n*}(\boldsymbol{x}), \boldsymbol{x}), \end{array} \right. \qquad \forall i \in \{1, \ldots, N\}$$

and $\boldsymbol{\pi}^n = (\pi_1^n, \ldots, \pi_N^n)$. Bearing in mind that $\omega(\boldsymbol{x}) \ge \omega_\star > 0$ in $\overline{\Omega}$, we can define the relative Boltzmann entropy $\mathcal{H}_\omega$ with respect to ω by

$$\mathcal{H}_\omega(s) := \int_\Omega s(\boldsymbol{x}) \log\left(\frac{s(\boldsymbol{x})}{\omega(\boldsymbol{x})}\right) d\boldsymbol{x}, \quad \text{for all measurable } s : \Omega \to \mathbb{R}_+.$$

Lemma 2.2. *There exists C depending only on $\Omega, \Pi, \omega, \mathbb{K}, (\mu_i)_i$, and $\boldsymbol{\Psi}$ such that, for all $n \ge 1$ and all $\tau > 0$, one has*

$$\sum_{i=0}^N \|\nabla \pi_i^n\|_{L^2(\Omega)}^2 \le C\left(1 + \frac{\boldsymbol{W}^2(\boldsymbol{s}^n, \boldsymbol{s}^{n-1})}{\tau} + \sum_{i=0}^N \frac{\mathcal{H}_\omega(s_i^{n-1}) - \mathcal{H}_\omega(s_i^n)}{\tau}\right). \tag{30}$$

Proof. The argument relies on the *flow interchange* technique introduced by Matthes, McCann, and Savaré in [36]. Throughout the proof, C denotes a fluctuating quantity that depends on the prescribed data $\Omega, \Pi, \omega, \mathbb{K}, (\mu_i)_i$, and $\boldsymbol{\Psi}$, but neither on t, τ, nor on n. For $i = 0 \ldots N$ consider the auxiliary flows

$$\left\{ \begin{array}{ll} \partial_t \check{s}_i = \operatorname{div}(\mathbb{K}\nabla \check{s}_i - \check{s}_i \mathbb{K}\nabla \log \omega), & t > 0,\ \boldsymbol{x} \in \Omega, \\ \mathbb{K}(\nabla \check{s}_i - \check{s}_i \nabla \log \omega) \cdot \nu = 0, & t > 0,\ \boldsymbol{x} \in \partial\Omega, \\ \check{s}_i\big|_{t=0} = s_i^n, & \boldsymbol{x} \in \Omega. \end{array} \right. \tag{31}$$

By standard parabolic theory (see for instance [32, Chapter III, Theorem 12.2]), these Initial-Boundary value problems are well-posed, and their solutions $\check{s}_i(\boldsymbol{x})$ belong to $\mathcal{C}^{1,2}((0,1] \times \overline{\Omega}) \cap \mathcal{C}([0,1]; L^p(\Omega))$ for all $p \in (1, \infty)$ if $\omega \in \mathcal{C}^{2,\alpha}(\overline{\Omega})$ and $\mathbb{K} \in \mathcal{C}^{1,\alpha}(\overline{\Omega})$ for some $\alpha > 0$. Therefore, $t \mapsto \check{s}_i(\cdot, t)$ is absolutely continuous in $L^1(\Omega)$, thus in $\mathcal{A}_i$ endowed with the usual quadratic distance W_{ref} (23) thanks to [42, Prop. 7.4]. Because of (22), the curve $t \mapsto \check{s}_i(\cdot, t)$ is also absolutely continuous in $\mathcal{A}_i$ endowed with W_i.

From Lisini's results [34], we know that the evolution $t \mapsto \check{s}_i(\cdot, t)$ can be interpreted as the gradient flow of the relative Boltzmann functional $\frac{1}{\mu_i}\mathcal{H}_\omega$ with respect to the metric W_i, the scaling factor $\frac{1}{\mu_i}$ appearing due to the definition (21) of the distance W_i. As a consequence of (20), The Ricci curvature of (Ω, d_i) is bounded, hence bounded from below. Since $\omega \in \mathcal{C}^2(\overline{\Omega})$ we also have that $\frac{1}{\mu_i}\mathcal{H}_\omega$ is λ_i-displacement convex with respect to W_i for some $\lambda_i \in \mathbb{R}$ depending on ω and the geometry of (Ω, d_i), see [44, Chapter 14]. Therefore, we can use the so-called *Evolution Variational Inequality* characterization of gradient flows (see for instance [4, Definition 4.5]) centered at s_i^{n-1}, namely

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}W_i^2(\check{s}_i(t), s_i^{n-1}) + \frac{\lambda_i}{2}W_i^2(\check{s}_i(t), s_i^{n-1}) \le \frac{1}{\mu_i}\mathcal{H}_\omega(s_i^{n-1}) - \frac{1}{\mu_i}\mathcal{H}_\omega(\check{s}_i(t)). \tag{32}$$

Denote by $\check{\boldsymbol{s}} = (\check{s}_0, \dots, \check{s}_N)$, and by $\check{\boldsymbol{s}}^* = (\check{s}_1, \dots, \check{s}_N)$. Summing (32) over $i \in \{0, \dots, N\}$ leads, for all $t > 0$, to

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{1}{2\tau}\boldsymbol{W}^2(\check{\boldsymbol{s}}(t), \boldsymbol{s}^{n-1})\right) \le C\left(\frac{\boldsymbol{W}^2(\check{\boldsymbol{s}}(t), \boldsymbol{s}^{n-1})}{\tau} + \sum_{i=0}^N \frac{\mathcal{H}_\omega(s_i^{n-1}) - \mathcal{H}_\omega(\check{s}_i(t))}{\tau}\right). \tag{33}$$

In order to estimate the internal energy contribution in (25), we first note that $\sum_{i=0}^N s_i^n(\boldsymbol{x}) = \omega(\boldsymbol{x})$ for all $\boldsymbol{x} \in \overline{\Omega}$, thus by the linearity of the system (31), we have $\sum_{i=0}^N \check{s}_i(\boldsymbol{x}, t) = \omega(\boldsymbol{x})$ as well for all $\boldsymbol{x} \in \overline{\Omega}$ and all $t > 0$. Moreover, the problem (31) is monotone, thus order preserving, and admits 0 as a subsolution. Hence $\check{s}_i(\boldsymbol{x}, t) \ge 0$, so that $\check{\boldsymbol{s}}(t) \in \boldsymbol{\mathcal{A}} \cap \boldsymbol{\mathcal{X}}$ is an admissible competitor in (25) for all $t > 0$. The smoothness of $\check{\boldsymbol{s}}$ for $t > 0$ allows to write

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\int_\Omega \Pi(\check{\boldsymbol{s}}^*(\boldsymbol{x}, t), \boldsymbol{x})\mathrm{d}\boldsymbol{x}\right) = \sum_{i=1}^N \int_\Omega \check{\pi}_i(\boldsymbol{x}, t)\partial_t \check{s}_i(\boldsymbol{x}, t)\mathrm{d}\boldsymbol{x} = I_1(t) + I_2(t), \tag{34}$$

where $\check{\pi}_i := \pi_i(\check{\boldsymbol{s}}^*, \cdot)$, and where, for all $t > 0$, we have set

$$I_1(t) = -\sum_{i=1}^N \int_\Omega \boldsymbol{\nabla}\check{\pi}_i(t) \cdot \mathbb{K}\boldsymbol{\nabla}\check{s}_i(t)\mathrm{d}\boldsymbol{x}, \quad I_2(t) = -\sum_{i=1}^N \int_\Omega \frac{\check{s}_i(t)}{\omega}\boldsymbol{\nabla}\check{\pi}_i(t) \cdot \mathbb{K}\boldsymbol{\nabla}\omega\mathrm{d}\boldsymbol{x}.$$

To estimate I_1, we first use the inversibility of $\boldsymbol{\pi}$ to write

$$\check{\boldsymbol{s}}(\boldsymbol{x}, t) = \boldsymbol{\phi}(\check{\boldsymbol{\pi}}(\boldsymbol{x}, t), \boldsymbol{x}) =: \check{\boldsymbol{\phi}}(\boldsymbol{x}, t),$$

yielding

$$\boldsymbol{\nabla}\check{\boldsymbol{s}}(\boldsymbol{x}, t) = \mathbb{J}_{\boldsymbol{z}}\boldsymbol{\phi}(\check{\boldsymbol{\pi}}(\boldsymbol{x}, t), \boldsymbol{x})\boldsymbol{\nabla}\check{\boldsymbol{\pi}}(\boldsymbol{x}, t) + \boldsymbol{\nabla}_{\boldsymbol{x}}\boldsymbol{\phi}(\check{\boldsymbol{\pi}}(\boldsymbol{x}, t), \boldsymbol{x}). \tag{35}$$

Combining (3), (8), (9) and the elementary inequality

$$ab \le \delta\frac{a^2}{2} + \frac{b^2}{2\delta} \quad \text{with } \delta > 0 \text{ arbitrary,} \tag{36}$$

we get that for all $t > 0$, there holds

$$I_1(t) \le -\frac{\kappa_\star}{\varpi^\star}\int_\Omega |\boldsymbol{\nabla}\check{\boldsymbol{\pi}}(t)|^2\mathrm{d}\boldsymbol{x} + \kappa^\star\left(\delta\int_\Omega |\boldsymbol{\nabla}\check{\boldsymbol{\pi}}(t)|^2\mathrm{d}\boldsymbol{x} + \frac{1}{\delta}\int_\Omega |\boldsymbol{\nabla}_{\boldsymbol{x}}\boldsymbol{\phi}(\check{\boldsymbol{\pi}}(t))|^2\mathrm{d}\boldsymbol{x}\right).$$

Choosing $\delta = \frac{\kappa_\star}{4\kappa^\star \varpi^\star}$, we get that

$$I_1(t) \le -\frac{3\kappa_\star}{4\varpi^\star} \int_\Omega |\boldsymbol{\nabla}\check{\boldsymbol{\pi}}(t)|^2 \mathrm{d}\boldsymbol{x} + C, \qquad \forall t > 0. \tag{37}$$

In order to estimate I_2, we use that $\check{\boldsymbol{s}}(t) \in \boldsymbol{\mathcal{X}}$ for all $t > 0$, so that $0 \le \check{s}_i(\boldsymbol{x}, t) \le \omega(\boldsymbol{x})$, hence we deduce that $\sum_{i=1}^N \left(\frac{\check{s}_i}{\omega}\right)^2 \le 1$. Therefore, using (36) again, we get

$$I_2(t) \le \delta\kappa^\star \int_\Omega |\boldsymbol{\nabla}\check{\boldsymbol{\pi}}(t)|^2 \mathrm{d}\boldsymbol{x} + \frac{\kappa^\star}{\delta} \int_\Omega |\boldsymbol{\nabla}\omega|^2 \mathrm{d}\boldsymbol{x}.$$

Choosing again $\delta = \frac{\kappa_\star}{4\kappa^\star \varpi^\star}$ yields

$$I_2(t) \le \frac{\kappa_\star}{4\varpi^\star} \int_\Omega |\boldsymbol{\nabla}\check{\boldsymbol{\pi}}(t)|^2 \mathrm{d}\boldsymbol{x} + C. \tag{38}$$

Taking (37)–(38) into account in (34) provides

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\int_\Omega \Pi(\check{\boldsymbol{s}}^*(\boldsymbol{x}, t), \boldsymbol{x}) \mathrm{d}\boldsymbol{x}\right) \le -\frac{\kappa_\star}{2\varpi^\star} \int_\Omega |\boldsymbol{\nabla}\check{\boldsymbol{\pi}}(t)|^2 \mathrm{d}\boldsymbol{x} + C, \qquad \forall t > 0. \tag{39}$$

Let us now focus on the potential (gravitational) energy. Since $\check{\boldsymbol{s}}(t)$ belongs to $\boldsymbol{\mathcal{X}} \cap \boldsymbol{\mathcal{A}}$ for all $t > 0$, we can make use of the relation

$$\check{s}_0(\boldsymbol{x}, t) = \omega(\boldsymbol{x}) - \sum_{i=1}^N \check{s}_i(\boldsymbol{x}, t), \qquad \text{for all } (\boldsymbol{x}, t) \in \Omega \times \mathbb{R}_+,$$

to write: for all $t > 0$,

$$\sum_{i=0}^N \int_\Omega \check{s}_i(\boldsymbol{x}, t)\Psi_i(\boldsymbol{x}) \mathrm{d}\boldsymbol{x} = \sum_{i=1}^N \int_\Omega \check{s}_i(\boldsymbol{x}, t)(\Psi_i - \Psi_0)(\boldsymbol{x}) \mathrm{d}\boldsymbol{x} + \int_\Omega \omega(\boldsymbol{x})\Psi_0(\boldsymbol{x}) \mathrm{d}\boldsymbol{x}.$$

This leads to

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\sum_{i=0}^N \int_\Omega \check{s}_i(t)\Psi_i \mathrm{d}\boldsymbol{x}\right) = \sum_{i=1}^N \int_\Omega (\Psi_i(\boldsymbol{x}) - \Psi_0(\boldsymbol{x}))\partial_t s_i(\boldsymbol{x}, t) \mathrm{d}\boldsymbol{x} = J_1(t) + J_2(t), \tag{40}$$

where, using the equations (31), we have set

$$J_1(t) = -\sum_{i=1}^N \int_\Omega \boldsymbol{\nabla}(\Psi_i - \Psi_0) \cdot \mathbb{K}\boldsymbol{\nabla}\check{s}_i(t) \mathrm{d}\boldsymbol{x},$$
$$J_2(t) = \sum_{i=1}^N \int_\Omega \frac{\check{s}_i(t)}{\omega} \boldsymbol{\nabla}(\Psi_i - \Psi_0) \cdot \mathbb{K}\boldsymbol{\nabla}\omega \mathrm{d}\boldsymbol{x}.$$

The term J_1 can be overestimated using (36). More precisely, for all $\delta > 0$, we have

$$J_1(t) \le \kappa^\star \left(\delta \|\boldsymbol{\nabla}\check{\boldsymbol{s}}^*(t)\|_{L^2}^2 + \frac{1}{\delta}\sum_{i=1}^N \|\nabla(\Psi_i - \Psi_0)\|_{L^2}^2\right). \tag{41}$$

Using (35) together with (8)–(9), we get that

$$\|\boldsymbol{\nabla}\check{\boldsymbol{s}}^*\|_{L^2}^2 \le \left(\frac{1}{\varpi_\star}\|\boldsymbol{\nabla}\check{\boldsymbol{\pi}}\|_{L^2} + |\Omega| M_{\boldsymbol{\phi}}\right)^2 \le \frac{2}{(\varpi_\star)^2}\|\boldsymbol{\nabla}\check{\boldsymbol{\pi}}\|_{L^2}^2 + 2\left(|\Omega| M_{\boldsymbol{\phi}}\right)^2.$$

Therefore, choosing $\delta = \frac{(\varpi_\star)^2 \kappa_\star}{8\kappa^\star \varpi^\star}$ in (41), we infer from the regularity of $\boldsymbol{\Psi}$ that

$$J_1(t) \leq \frac{\kappa_\star}{4\varpi^\star} \int_\Omega |\boldsymbol{\nabla} \check{\boldsymbol{\pi}}(t)|^2 \mathrm{d}\boldsymbol{x} + C, \qquad \forall t > 0. \tag{42}$$

Finally, it follows from the fact that $\sum_{i=1}^N \check{s}_i \leq \omega$, from Cauchy-Schwarz inequality, and from the definition (14) of Ψ_i that

$$J_2(t) \geq -\kappa^\star \sum_{i=1}^N \|\boldsymbol{\nabla} \Psi_i - \boldsymbol{\nabla} \Psi_0\|_{L^2} \|\boldsymbol{\nabla} \omega\|_{L^2} = C. \tag{43}$$

Combining (40), (42), and (43) with (39), we get that

$$\frac{\mathrm{d}}{\mathrm{d}t} \mathcal{E}(\check{\boldsymbol{s}}(t)) \leq -\frac{\kappa_\star}{4\varpi^\star} \int_\Omega |\boldsymbol{\nabla} \check{\boldsymbol{\pi}}(t)|^2 \mathrm{d}\boldsymbol{x} + C, \qquad \forall t > 0. \tag{44}$$

Denote by

$$\mathcal{F}_\tau^n(\boldsymbol{s}) = \frac{1}{2\tau} \boldsymbol{W}^2(\boldsymbol{s}, \boldsymbol{s}^{n-1}) + \mathcal{E}(\boldsymbol{s}) \tag{45}$$

the functional to be minimized in (25), then gathering (33) and (44) provides

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t} \mathcal{F}_\tau^n(\check{\boldsymbol{s}}(t)) &+ \frac{\kappa_\star}{4\varpi^\star} \|\boldsymbol{\nabla} \check{\boldsymbol{\pi}}\|_{L^2}^2 \\ &\leq C \left(1 + \frac{\boldsymbol{W}^2(\check{\boldsymbol{s}}(t), \boldsymbol{s}^{n-1})}{\tau} + \sum_{i=0}^N \frac{\mathcal{H}_\omega(s_i^{n-1}) - \mathcal{H}_\omega(\check{s}_i(t))}{\tau} \right) \qquad \forall t > 0. \end{aligned}$$

Since $\check{\boldsymbol{s}}(0) = \boldsymbol{s}^n$ is a minimizer of (25) we must have

$$0 \leq \limsup_{t \to 0^+} \left(\frac{\mathrm{d}}{\mathrm{d}t} \mathcal{F}_\tau^n(\check{\boldsymbol{s}}(t)) \right),$$

otherwise $\check{\boldsymbol{s}}(t)$ would be a strictly better competitor than $\boldsymbol{s}^n$ for small strictly positive t. As a consequence, we get

$$\liminf_{t \to 0^+} \|\boldsymbol{\nabla} \check{\boldsymbol{\pi}}(t)\|_{L^2}^2 \leq C \limsup_{t \to 0^+} \left(1 + \frac{\boldsymbol{W}^2(\check{\boldsymbol{s}}(t), \boldsymbol{s}^{n-1})}{\tau} + \sum_{i=0}^N \frac{\mathcal{H}_\omega(s_i^{n-1}) - \mathcal{H}_\omega(\check{s}_i(t))}{\tau} \right).$$

Since $\check{s}_i$ belongs to $C([0,1]; L^p(\Omega))$ for all $p \in [1, \infty)$ (see for instance [16]) for all $i \in \{0, \dots, N\}$, then thanks to the continuity of the Wasserstein distance and of the Boltzmann entropy with respect to the L^p-convergence, we get that

$$\boldsymbol{W}^2(\check{\boldsymbol{s}}(t), \boldsymbol{s}^{n-1}) \underset{t \to 0^+}{\longrightarrow} \boldsymbol{W}^2(\boldsymbol{s}^n, \boldsymbol{s}^{n-1}) \quad \text{and} \quad \mathcal{H}_\omega(\check{s}_i(t)) \underset{t \to 0^+}{\longrightarrow} \mathcal{H}_\omega(s_i^n).$$

Therefore, we obtain that

$$\liminf_{t \to 0^+} \|\boldsymbol{\nabla} \check{\boldsymbol{\pi}}(t)\|_{L^2}^2 \leq C \left(1 + \frac{\boldsymbol{W}^2(\boldsymbol{s}^n, \boldsymbol{s}^{n-1})}{\tau} + \sum_{i=0}^N \frac{\mathcal{H}_\omega(s_i^{n-1}) - \mathcal{H}_\omega(s_i^n)}{\tau} \right). \tag{46}$$

It follows from the regularity of $\boldsymbol{\pi}$ that

$$\check{\boldsymbol{\pi}}(t) \underset{t \to 0^+}{\longrightarrow} \boldsymbol{\pi}^n = \boldsymbol{\pi}(\boldsymbol{s}^{n*}, \boldsymbol{x}) \quad \text{in } L^p(\Omega).$$

Finally, let $(t_\ell)_{\ell\geq 1}$ be a decreasing sequence tending to 0 realizing the lim inf in (46), then the sequence $(\boldsymbol{\nabla}\check{\boldsymbol{\pi}}(t_\ell))_{\ell\geq 1}$ converges weakly in $L^2(\Omega)^{N\times d}$ towards $\boldsymbol{\nabla}\boldsymbol{\pi}^n$. The lower semi-continuity of the norm w.r.t. the weak convergence leads to

$$\sum_{i=1}^N \|\boldsymbol{\nabla}\pi_i^n\|_{L^2}^2 \leq \lim_{\ell\to\infty} \|\boldsymbol{\nabla}\check{\boldsymbol{\pi}}(t_\ell)\|_{L^2}^2 = \liminf_{t\to 0^+} \|\boldsymbol{\nabla}\check{\boldsymbol{\pi}}(t)\|_{L^2}^2$$
$$\leq C\left(1 + \frac{\boldsymbol{W}^2(\boldsymbol{s}^n, \boldsymbol{s}^{n-1})}{\tau} + \sum_{i=0}^N \frac{\mathcal{H}_\omega(s_i^{n-1}) - \mathcal{H}_\omega(s_i^n)}{\tau}\right)$$

and the proof is complete. □

3. The Euler-Lagrange equations and pressure bounds

The goal of this section is to extract informations coming from the optimality of $\boldsymbol{s}^n$ in the minimization (25). The main difficulty consists in constructing the phase and capillary pressures from this optimality condition. Our proof is inspired from [37] and makes an extensive use of the Kantorovich potentials. Therefore, we first recall their definition and some useful properties. We refer to [42, §1.2] or [44, Chapter 5] for details.

Let $(\nu_1, \nu_2) \in \mathcal{M}_+(\Omega)^2$ be two nonnegative measures with same total mass. A pair of Kantorovich potentials $(\varphi_i, \psi_i) \in L^1(\nu_1)\times L^1(\nu_2)$ associated to the measures ν_1 and ν_2 and to the cost function $\frac{1}{2}d_i^2$ defined by (18), $i \in \{0,\dots,N\}$, is a solution of the Kantorovich *dual problem*

$$DP_i(\nu_1,\nu_2) = \max_{\substack{(\varphi_i,\psi_i)\in L^1(\nu_1)\times L^1(\nu_2)\\ \varphi_i(\boldsymbol{x})+\psi_i(\boldsymbol{y})\leq \frac{1}{2}d_i^2(\boldsymbol{x},\boldsymbol{y})}} \int_\Omega \varphi_i(\boldsymbol{x})\nu_1(\boldsymbol{x})\mathrm{d}\boldsymbol{x} + \int_\Omega \psi_i(\boldsymbol{y})\nu_2(\boldsymbol{y})\mathrm{d}\boldsymbol{y}.$$

We will use the three following important properties of the Kantorovich potentials:

(a) There is always duality

$$DP_i(\nu_1,\nu_2) = \frac{1}{2}W_i^2(\nu_1,\nu_2), \qquad \forall i \in \{0,\dots,N\}.$$

(b) A pair of Kantorovich potentials (φ_i,ψ_i) is $\mathrm{d}\nu_1\otimes\mathrm{d}\nu_2$ unique, up to additive constants.

(c) The Kantorovich potentials φ_i and ψ_i are $\frac{1}{2}d_i^2$-conjugate, that is

$$\varphi_i(\boldsymbol{x}) = \inf_{\boldsymbol{y}\in\Omega} \frac{1}{2}d_i^2(\boldsymbol{x},\boldsymbol{y}) - \psi_i(\boldsymbol{y}), \quad \forall\, \boldsymbol{x}\in\Omega,$$
$$\psi_i(\boldsymbol{y}) = \inf_{\boldsymbol{x}\in\Omega} \frac{1}{2}d_i^2(\boldsymbol{x},\boldsymbol{y}) - \varphi_i(\boldsymbol{x}), \quad \forall\, \boldsymbol{y}\in\Omega.$$

Remark 3.1. *Since Ω is bounded, the cost functions $(\boldsymbol{x},\boldsymbol{y}) \mapsto \frac{1}{2}d_i^2(\boldsymbol{x},\boldsymbol{y})$, $i \in \{1,\dots,N\}$, are globally Lipschitz continuous, see* (19)*. Thus item* (c) *shows that φ_i and ψ_i are also Lipschitz continuous.*

3.1. **A decomposition result.**

Our next result essentially states that, since $\boldsymbol{s}^n$ is a minimizer of (25), it is also a minimizer of the linearized problem. This linearization will be useful later on to deduce the aforementioned structural decomposition, and leverages the Kantorovich potentials in order to linearize the squared Wasserstein distances.

Lemma 3.2. *For $n \geq 1$ and $i = 0, \dots, N$ there exist some (backwards, optimal) Kantorovich potentials φ_i^n from s_i^n to s_i^{n-1} such that, setting*

$$
\left\{ \begin{array}{l} F_0^n = 0, \\ F_i^n = \frac{\varphi_i^n}{\tau} - \frac{\varphi_0^n}{\tau} + \pi_i^n + \Psi_i - \Psi_0, \quad \forall i \in \{1, \dots, N\}, \end{array} \right. \tag{47}
$$

and $\mathbf{F}^n = (F_i^n)_{0 \leq i \leq N}$, there holds

$$
\boldsymbol{s}^n = \underset{\boldsymbol{s} \in \boldsymbol{\mathcal{X}} \cap \boldsymbol{\mathcal{A}}}{\operatorname{Argmin}} \int_\Omega \mathbf{F}^n(\boldsymbol{x}) \cdot \boldsymbol{s}(\boldsymbol{x}) \mathrm{d}\boldsymbol{x}. \tag{48}
$$

Moreover, $F_i^n \in L^\infty \cap H^1(\Omega)$ for all $i \in \{0, \dots, N\}$.

Proof. The proof is inspired from that of [37, Lemma 3.1]. We assume first that $s_i^{n-1}(\boldsymbol{x}) > 0$ everywhere in $\overline{\Omega}$ for all $i \in \{1, \dots, N\}$, so that the Kantorovich potentials (φ_i^n, ψ_i^n) from s_i^n to s_i^{n-1} are uniquely determined after normalizing $\varphi_i^n(\boldsymbol{x}_{\mathrm{ref}}) = 0$ for some arbitrary point $\boldsymbol{x}_{\mathrm{ref}} \in \Omega$ (cf. [42, Proposition 7.18]). Given any $\boldsymbol{s}^* = (s_i)_{1 \leq i \leq N} \in \boldsymbol{\mathcal{X}}^* \cap \boldsymbol{\mathcal{A}}^*$ we define

$$
s_i^\varepsilon := (1 - \varepsilon) s_i^n + \varepsilon s_i \quad \text{for } i \in \{1, \dots, N\}, \qquad \text{and} \qquad s_0^\varepsilon := \omega - \sum_{i=1}^N s_i^\varepsilon.
$$

Note that for $\varepsilon \in (0,1)$ this ε-perturbation $\boldsymbol{s}^\varepsilon = (s_0^\varepsilon, s_1^\varepsilon, \dots, s_N^\varepsilon)$ sums to $\omega(\boldsymbol{x})$ and is admissible in the sense that $\boldsymbol{s}^\varepsilon \in \boldsymbol{\mathcal{X}} \cap \boldsymbol{\mathcal{A}}$. Let $(\varphi_i^\varepsilon, \psi_i^\varepsilon)$ be the unique Kantorovich potentials from s_i^ε to s_i^{n-1}, similarly normalized as $\varphi_i^\varepsilon(\boldsymbol{x}_0) = 0$. Then by characterization of the squared Wasserstein distance in terms of the dual Kantorovich problem we have

$$
\begin{cases} \dfrac{1}{2} W_i^2(s_i^\varepsilon, s_i^{n-1}) = \displaystyle\int_\Omega \varphi_i^\varepsilon(\boldsymbol{x}) s_i^\varepsilon(\boldsymbol{x}) \mathrm{d}\boldsymbol{x} + \int_\Omega \psi_i^\varepsilon(\boldsymbol{y}) s_i^{n-1}(\boldsymbol{y}) \mathrm{d}\boldsymbol{y}, \\ \dfrac{1}{2} W_i^2(s_i^n, s_i^{n-1}) \geq \displaystyle\int_\Omega \varphi_i^\varepsilon(\boldsymbol{x}) s_i^n(\boldsymbol{x}) \mathrm{d}\boldsymbol{x} + \int_\Omega \psi_i^\varepsilon(\boldsymbol{y}) s_i^{n-1}(\boldsymbol{y}) \mathrm{d}\boldsymbol{y}. \end{cases}
$$

By definition of the perturbation $\boldsymbol{s}^\varepsilon$ it is easy to check that $s_i^\varepsilon - s_i^n = \varepsilon(s_i - s_i^n)$ for $i \in \{1, \dots, N\}$, while $s_0^\varepsilon - s_0^n = -\varepsilon \sum\limits_{i=1}^N (s_i - s_i^n)$. Subtracting the previous inequalities we get, respectively,

$$
\frac{1}{2\tau} \left(W_i^2(s_i^\varepsilon, s_i^{n-1}) - W_i^2(s_i^n, s_i^{n-1}) \right) \leq \frac{\varepsilon}{\tau} \int_\Omega \varphi_i^\varepsilon (s_i - s_i^n) \mathrm{d}\boldsymbol{x} \tag{49}
$$

for all $i \in \{1, \dots, N\}$, and

$$
\frac{1}{2\tau} \left(W_0^2(s_0^\varepsilon, s_0^{n-1}) - W_0^2(s_0^n, s_0^{n-1}) \right) \leq -\frac{\varepsilon}{\tau} \sum_{i=1}^N \int_\Omega \varphi_0^\varepsilon (s_i - s_i^n) \mathrm{d}\boldsymbol{x}. \tag{50}
$$

Denote by $\boldsymbol{s}^{\varepsilon *} = (s_1^\varepsilon, \dots, s_N^\varepsilon)$ and by $\boldsymbol{\pi}^\varepsilon = \boldsymbol{\pi}(\boldsymbol{s}^{\varepsilon *}, \cdot)$. The convexity of Π implies that

$$
\int_\Omega \left(\Pi(\boldsymbol{s}^{n*}, \boldsymbol{x}) - \Pi(\boldsymbol{s}^{\varepsilon *}, \boldsymbol{x}) \right) \mathrm{d}\boldsymbol{x} \geq \int_\Omega \boldsymbol{\pi}^\varepsilon \cdot (\boldsymbol{s}^{n*} - \boldsymbol{s}^{\varepsilon *}) \, \mathrm{d}\boldsymbol{x} = -\varepsilon \int_\Omega \boldsymbol{\pi}^\varepsilon \cdot (\boldsymbol{s}^* - \boldsymbol{s}^{n*}) \, \mathrm{d}\boldsymbol{x}. \tag{51}
$$

As for the potential energy, we substitute $s_0^\varepsilon = \omega - \sum_{i=1}^N s_i^\varepsilon$ and immediately obtain by linearity

$$\int_\Omega (\boldsymbol{s}^\varepsilon - \boldsymbol{s}^n) \cdot \boldsymbol{\Psi} \, \mathrm{d}\boldsymbol{x} = \varepsilon \sum_{i=1}^N \int_\Omega (\Psi_i - \Psi_0)(s_i - s_i^n) \mathrm{d}\boldsymbol{x}. \tag{52}$$

Summing (49)–(52), dividing by ε, and recalling that $\boldsymbol{s}^n$ minimizes the functional $\mathcal{F}_\tau^n$ defined by (45), we obtain

$$0 \le \frac{\mathcal{F}_\tau^n(\boldsymbol{s}^\varepsilon) - \mathcal{F}_\tau^n(\boldsymbol{s}^n)}{\varepsilon} \le \sum_{i=1}^N \int_\Omega \left(\frac{\varphi_i^\varepsilon}{\tau} - \frac{\varphi_0^\varepsilon}{\tau} + \pi_i^\varepsilon + \Psi_i - \Psi_0 \right) (s_i - s_i^n) \mathrm{d}\boldsymbol{x} \tag{53}$$

for all $\boldsymbol{s} \in \boldsymbol{\mathcal{X}} \cap \boldsymbol{\mathcal{A}}$ and all $\varepsilon \in (0,1)$. Because Ω is bounded, any Kantorovich potential is globally Lipschitz with bounds uniform in ε (see for instance the proof of [42, Theorem 1.17]). Since $\boldsymbol{s}^\varepsilon$ converges uniformly towards $\boldsymbol{s}^n$ when ε tends to 0, we infer from [42, Theorem 1.52] that φ_i^ε converges uniformly towards φ_i^n as ε tends to 0, where φ_i^n is a Kantorovich potential form s_i^n to s_i^{n-1}. Moreover, since $\boldsymbol{\pi}$ is uniformly continuous, we also know that $\boldsymbol{\pi}^\varepsilon$ converges uniformly towards $\boldsymbol{\pi}^n$. Then we can pass to the limit in (53) and infer that

$$0 \le \int_\Omega \mathbf{F}^{n*} \cdot (\boldsymbol{s}^* - \boldsymbol{s}^{n*}) \mathrm{d}\boldsymbol{x} = \int_\Omega \mathbf{F}^n \cdot (\boldsymbol{s} - \boldsymbol{s}^n) \mathrm{d}\boldsymbol{x}, \qquad \forall \boldsymbol{s} \in \boldsymbol{\mathcal{X}} \cap \boldsymbol{\mathcal{A}} \tag{54}$$

thanks to the choice $F_0^n = 0$. As a consequence, (48) holds.

If $s_i^{n-1} > 0$ does not hold everywhere we argue by approximation. Running the flow (31) for short times $\delta > 0$ starting from $\boldsymbol{s}^{n-1}$, we construct an approximation $\boldsymbol{s}^{n-1,\delta} = (s_0^{n-1,\delta}, \dots, s_N^{n-1,\delta})$ converging to $\boldsymbol{s}^{n-1} = (s_0^{n-1}, \dots, s_N^{n-1})$ at least in $L^1(\Omega)$ when δ tends to 0. By construction $\boldsymbol{s}^{n,\delta} \in \boldsymbol{\mathcal{X}} \cap \boldsymbol{\mathcal{A}}$, and it follows from the strong maximum principle that $s_i^{n-1,\delta} > 0$ in $\overline{\Omega}$ for all $\delta > 0$. We denote by $\boldsymbol{s}^{n,\delta *} = \left(s_1^{n,\delta}, \dots, s_N^{n,\delta}\right) \in \boldsymbol{\mathcal{X}}^* \cap \boldsymbol{\mathcal{A}}^*$. There exists a unique minimizer $\boldsymbol{s}^{n,\delta}$ to the functional

$$\mathcal{F}_\tau^{n,\delta} : \begin{cases} \boldsymbol{\mathcal{X}} \cap \boldsymbol{\mathcal{A}} \to \mathbb{R}_+ \\ \boldsymbol{s} \mapsto \frac{1}{2\tau} \boldsymbol{W}^2(\boldsymbol{s}, \boldsymbol{s}^{n-1,\delta}) + \mathcal{E}(\boldsymbol{s}) \end{cases}$$

Thanks to the positivity of $\boldsymbol{s}^{n,\delta}$, there exist unique Kantorovich potentials $(\varphi_i^{n,\delta}, \psi_i^{n,\delta})$ from $s_i^{n,\delta}$ to $s_i^{n-1,\delta}$. This allows to construct $\mathbf{F}^{n,\delta}$ using (47) where φ_i^n (resp. π_i^n) has been replaced by $\varphi_i^{n,\delta}$ (resp. $\pi_i^{n,\delta} := \pi_i(\boldsymbol{s}^{n,\delta *}, \cdots)$). Thanks to the above discussion,

$$0 \le \int_\Omega \mathbf{F}^{n,\delta *} \cdot (\boldsymbol{s}^* - \boldsymbol{s}^{n,\delta *}) \mathrm{d}\boldsymbol{x}, \qquad \forall \boldsymbol{s}^* \in \boldsymbol{\mathcal{X}}^* \cap \boldsymbol{\mathcal{A}}^*. \tag{55}$$

We can now let δ tend to 0. Because of the time continuity of the solutions to (31), we know that $\boldsymbol{s}^{n-1,\delta}$ converges towards $\boldsymbol{s}^{n-1}$ in $L^1(\Omega)$. On the other hand, from the definition of $s^{n,\delta}$ and Lemma 2.2 (in particular (30) with $s^{n-1,\delta}, s^{n,\delta}, \boldsymbol{\pi}^{n,\delta}$ instead of $s^{n-1}, s^n, \boldsymbol{\pi}^n$) it is not difficult to check that $\boldsymbol{\pi}^{n,\delta}$ is bounded in $H^1(\Omega)^{N+1}$ uniformly in $\delta > 0$. Using next $\boldsymbol{s}^{n,\delta}(\boldsymbol{x}) = \boldsymbol{\phi}(\boldsymbol{\pi}^{n,\delta}(\boldsymbol{x}), \boldsymbol{x})$ and (8)–(9), one gets that $\boldsymbol{s}^{n,\delta}$ is uniformly bounded in $H^1(\Omega)^{N+1}$. Then, thanks to Rellich's compactness theorem, we can assume that $\boldsymbol{s}^{n,\delta}$ converges strongly in $L^2(\Omega)^{N+1}$ as δ tends to 0. By the strong convergence $\boldsymbol{s}^{n-1,\delta} \to \boldsymbol{s}^{n-1}$ and standard properties of the squared Wasserstein distance, one readily checks that $\mathcal{F}_\tau^{n,\delta}$ Γ-converges towards $\mathcal{F}_\tau^n$, and we can therefore identify the limit of $\boldsymbol{s}^{n,\delta}$ as the unique minimizer $\boldsymbol{s}^n$ of $\mathcal{F}_\tau^n$. Thanks to Lebesgue's dominated convergence theorem, we also infer that $\pi_i^{n,\delta}$ converges in $L^2(\Omega)$ towards π_i^n. Finally, using once again the stability of the Kantorovich potentials [42, Theorem 1.52], we know that $\varphi_i^{n,\delta}$ converges uniformly towards some Kantorovich potential φ_i^n. Then we can pass to the limit in (55) and claim that (54) is satisfied even when some coordinates of $\boldsymbol{s}^{n-1}$ vanish on some parts of Ω.

Finally, note that since the Kantorovich potentials φ_i^n are Lipschitz continuous and because $\pi_i^n \in H^1$ (cf. Lemma 2.2) and Ψ_i, Ψ_0 are smooth, we have $F_i^n \in H^1$. Since the phases are bounded $0 \le s_i^n(\boldsymbol{x}) \le \omega(\boldsymbol{x}) \le 1$ and $\boldsymbol{\pi}$ is continuous we have $\boldsymbol{\pi}^n \in L^\infty$, thus $F_i^n \in L^\infty$ as well and the proof is complete. □

The purpose of the following Lemma is to suitably decompose the vector field $\mathbf{F}^n = (F_i^n)_{0 \le i \le N}$ defined by (47).

Lemma 3.3. *There exists $\boldsymbol{\alpha}^n \in \mathbb{R}^{N+1}$ and $\boldsymbol{\lambda}^n \in L^\infty \cap H^1(\Omega; \mathbb{R}^{N+1})$ such that*

$$F_i^n = \lambda_i^n - \lambda_0^n - \alpha_i^n \ \ \text{a.e. in } \Omega, \qquad \forall i \in \{0, \dots, N\}, \tag{56}$$

$$\boldsymbol{\nabla} \lambda_i^n = 0 \ \text{ in } \{s_i > 0\}, \qquad \forall i \in \{0, \dots, N\}. \tag{57}$$

Proof. The proof is based on a duality formula for the minimization problem (48). The quantities $\boldsymbol{\alpha}^n \in \mathbb{R}^{N+1}$, $\boldsymbol{\lambda}^n \in L^\infty(\Omega, \mathbb{R}^{N+1})$ will be related to some Lagrange multipliers for the constraints of total saturation ($\sum s_i = \omega$), individual masses ($\int s_i = m_i$), and non-negativity ($s_i \ge 0$). We start with some notations.

Notations. Let $E = L^\infty(\Omega)$ and $E^{N+1} = L^\infty(\Omega)^{N+1}$. The norms are

$$\|\mu\|_E = \|\mu\|_{L^\infty(\Omega)}, \qquad \|\boldsymbol{\lambda}\|_{E^{N+1}} = \left(\sum_{i=0}^N \|\lambda_i\|^2_{L^\infty(\Omega)} \right)^{1/2}.$$

We denote by E' and $(E')^{N+1}$ their respective topological dual spaces equipped with usual dual norms. We also define the positive dual cone

$$E'_+ = \left\{ s \in E' \ \middle| \ \langle \mu, s \rangle_{E,E'} \ge 0 \text{ for all } \mu \in L^\infty(\Omega; \mathbb{R}^+) \right\}$$

and the set

$$\overline{\boldsymbol{\mathcal{X}}} = \left\{ \boldsymbol{s} \in (E'_+)^{N+1} \ \middle| \ \sum_{i=0}^N s_i = \omega \right\}.$$

It is easy to verify that $\overline{\boldsymbol{\mathcal{X}}} = \boldsymbol{\mathcal{X}}$ since, for all measurable subset B of Ω, and all $\boldsymbol{s} \in \overline{\boldsymbol{\mathcal{X}}}$, one has

$$0 \le s_i[B] \le \sum_{i=0}^N s_i[B] = \omega[B] \le |B|, \qquad \forall i \in \{0, \dots, N\},$$

where we used the notation $s[B] = \langle \chi_B, s \rangle_{E,E'}$. We also define the set

$$\overline{\boldsymbol{\mathcal{A}}} = \left\{ \boldsymbol{s} \in (E'_+)^{N+1} \ \middle| \ s_i[\Omega] = m_i, \ i \in \{0, \dots, N\} \right\}$$

as well as the functional

$$\mathcal{I}(\boldsymbol{s}) = \int_\Omega \mathbf{F}^n \cdot \boldsymbol{s} \mathrm{d}\boldsymbol{x}, \qquad \forall \boldsymbol{s} \in \boldsymbol{\mathcal{X}} \cap \boldsymbol{\mathcal{A}},$$

where $\mathbf{F}^n$ was defined in (47). The characterization (48) of $\boldsymbol{s}^n$ and the choice $F_0^n = 0$ implies that

$$\underline{I} = \min_{\boldsymbol{s} \in \boldsymbol{\mathcal{X}} \cap \boldsymbol{\mathcal{A}}} \mathcal{I}(\boldsymbol{s}) = \min_{\boldsymbol{s}^* \in \boldsymbol{\mathcal{X}}^* \cap \boldsymbol{\mathcal{A}}^*} \int_\Omega \mathbf{F}^{n*} \cdot \boldsymbol{s}^* \mathrm{d}\boldsymbol{x} = \mathcal{I}(\boldsymbol{s}^n). \tag{58}$$

Since $\mathbf{F}^n \in L^\infty(\Omega)^{N+1}$ (cf. Lemma 3.2), the functional $\mathcal{I}(\boldsymbol{s})$ can be naturally extended to $\boldsymbol{s} \in (E')^{N+1}$ by setting

$$\mathcal{I}(\boldsymbol{s}) = \sum_{i=0}^{N} \langle F_i^n, s_i \rangle_{E,E'}, \qquad \forall \boldsymbol{s} \in (E')^{N+1}.$$

However, since $\overline{\boldsymbol{\mathcal{X}}} = \boldsymbol{\mathcal{X}}$, we know that

$$\underline{I} = \inf_{\boldsymbol{s} \in \overline{\boldsymbol{\mathcal{X}}} \cap \overline{\boldsymbol{\mathcal{A}}}} \mathcal{I}(\boldsymbol{s}) = \min_{\boldsymbol{s} \in \boldsymbol{\mathcal{X}} \cap \boldsymbol{\mathcal{A}}} \mathcal{I}(\boldsymbol{s}). \tag{59}$$

To construct the dual problem, we consider $\boldsymbol{\alpha} \in \mathbb{R}^{N+1}$ and define $\mu_{\boldsymbol{\alpha}} \in E$ by

$$\mu_{\boldsymbol{\alpha}} = \sup_{i \in \{0,\dots,N\}} \left(-F_i^n - \alpha_i \right). \tag{60}$$

Finally let us define the functional $\mathcal{J} : \mathbb{R}^{N+1} \to \mathbb{R}$ by

$$\mathcal{J}(\boldsymbol{\alpha}) = -\int_\Omega \mu_{\boldsymbol{\alpha}} \omega \mathrm{d}\boldsymbol{x} - \sum_{i=0}^{N} \alpha_i m_i, \qquad \forall \boldsymbol{\alpha} \in \mathbb{R}^{N+1},$$

and its supremum

$$\overline{J} = \sup_{\boldsymbol{\alpha} \in \mathbb{R}^{N+1}} \mathcal{J}(\boldsymbol{\alpha}).$$

Remark 3.4. *The condition $\sum\limits_{i=0}^{N} m_i = \int_\Omega \omega$ implies that $\mathcal{J}$ is invariant under translation, that is $\mathcal{J}(\boldsymbol{\alpha}) = \mathcal{J}(\boldsymbol{\alpha} + \delta.\mathbf{1})$ for any $\delta \in \mathbb{R}$. Therefore one can either set $\underline{\alpha} := \min_{i \in \{0,\dots,N\}} \alpha_i = 0$ or $\alpha_0 = 0$. Both normalizations will be used in the following.*

In order to make the link between $\underline{I}$ and $\overline{J}$ we compute (58), rewriting the constraint with Lagrange multipliers. The Lagrange multiplier corresponding to the total saturation constraint ($\sum s_i = \omega$) will be $\mu \in L^\infty(\Omega)$, while those for the mass of each phase ($s_i[\Omega] = m_i$) will be $\boldsymbol{\alpha} \in \mathbb{R}^{N+1}$. More precisely,

$$\begin{aligned}
\underline{I} &= \min_{\boldsymbol{s} \in \overline{\boldsymbol{\mathcal{X}}} \cap \overline{\boldsymbol{\mathcal{A}}}} \sum_{i=0}^{N} \langle F_i^n \cdot s_i \rangle_{E,E'} \\
&= \min_{\boldsymbol{s} \in (E'_+)^{N+1}} \sup_{(\mu, \boldsymbol{\alpha}) \in E \times \mathbb{R}^{N+1}} \left[\sum_{i=0}^{N} \langle F_i^n, s_i \rangle_{E,E'} \right. \\
&\qquad\qquad \left. - \left\langle \mu, \omega - \sum_{i=0}^{N} s_i \right\rangle_{E,E'} - \sum_{i=0}^{N} \alpha_i \left(m_i - s_i[\Omega] \right) \right] \\
&= \min_{\boldsymbol{s} \in (E'_+)^{N+1}} \sup_{(\mu, \boldsymbol{\alpha}) \in E \times \mathbb{R}^{N+1}} \sum_{i=0}^{N} \langle F_i^n + \mu + \alpha_i, s_i \rangle_{E,E'} - \langle \mu, \omega \rangle_{E,E'} - \sum_{i=0}^{N} \alpha_i m_i.
\end{aligned}$$

As both ω and μ belong to $L^\infty(\Omega)$, we have $\langle \mu, \omega\rangle_{E,E'} = \int_\Omega \mu\omega \mathrm{d}\boldsymbol{x}$. Hence we proved

$$\underline{I} = \min_{\boldsymbol{s}\in(E'_+)^{N+1}} \sup_{(\mu,\boldsymbol{\alpha})\in E\times\mathbb{R}^{N+1}} \sum_{i=0}^{N} \langle F_i^n + \mu + \alpha_i, s_i\rangle_{E,E'} - \int_\Omega \mu\omega \mathrm{d}\boldsymbol{x} - \sum_{i=0}^{N} \alpha_i m_i. \tag{61}$$

Let us now focus on $\overline{J}$. Since $\omega \geq 0$, one can relax the problem on $E \times R^{N+1}$:

$$\overline{J} = \sup_{\substack{(\mu,\boldsymbol{\alpha})\in E\times\mathbb{R}^{N+1}\\ \mu\geq\mu_{\boldsymbol{\alpha}}}} - \int_\Omega \mu\omega \,\mathrm{d}\boldsymbol{x} - \sum_{i=0}^{N} \alpha_i m_i.$$

The constraint $\mu \geq \mu_{\boldsymbol{\alpha}}$ a.e. in Ω rewrites $F_i^n(\boldsymbol{x}) + \mu(\boldsymbol{x}) + \alpha_i \geq 0$ for a.e. $\boldsymbol{x} \in \Omega$ and all $i \in \{0, \dots, N\}$. Thus by definition of $(E'_+)^{N+1}$, one has

$$\overline{J} = \sup_{(\mu,\boldsymbol{\alpha})\in E\times\mathbb{R}^{N+1}} \inf_{\boldsymbol{s}\in(E'_+)^{N+1}} \sum_{i=0}^{N} \langle F_i^n + \mu + \alpha_i, s_i\rangle_{E,E'} - \int_\Omega \mu_{\boldsymbol{\alpha}}\omega \mathrm{d}\boldsymbol{x} - \sum_{i=0}^{N} \alpha_i m_i. \tag{62}$$

Remark that $\underline{I}$ and $\overline{J}$ are equal, provided we can swap the inf and the sup. In order to establish our statement, we shall prove three things:

- We can swap the inf and the sup, that is $\underline{I} = \overline{J}$.
- The sup in $\mathcal{J}$ is achieved.
- The optimality conditions in $\underline{I} = \overline{J}$ leads to (56)-(57).

We consider each item one by one.

Step 1: Inf-Sup duality. We will use the following *min-max* theorem, whose proof can be found in [43, Thm. 3.1] (see also [14, Chapter 1, Prop. 1.1]):

Theorem 3.5. *Let A, B be nonempty convex sets of some vector spaces and let us suppose that A is endowed with an Hausdorff topology. Let $\mathcal{L} : A \times B \to \mathbb{R}$ be a function such that*

$$a \mapsto \mathcal{L}(a, b) \text{ is convex and lower semicontinuous in } A \text{ for every } b \in B,$$
$$b \mapsto \mathcal{L}(a, b) \text{ is concave in } B \text{ for every } a \in A.$$

If there exists $b_\star \in B$ and $M > \sup_{b\in B} \inf_{a\in A} \mathcal{L}(a, b)$ such that

$$S = \{a \in A : \quad \mathcal{L}(a, b_\star) \leq M\} \text{ is compact in } A \tag{63}$$

then

$$\sup_{b\in B} \inf_{a\in A} \mathcal{L}(a, b) = \inf_{b\in B} \sup_{a\in A} \mathcal{L}(a, b).$$

In order to apply this duality result to our specific problem, we define the primal space as

$$B = \left\{(\mu, \boldsymbol{\alpha}) \in E \times \mathbb{R}^{N+1}\right\}$$

and the dual space by

$$A = (E'_+)^{N+1}.$$

Notice that A is a convex subset of $(E')^{N+1}$ and B is also convex. We endow $(E')^{N+1}$ with the weak-$*$ topology, denoted $\sigma((E')^{N+1}, E^{N+1})$. Because E is a Banach space the topology $\sigma((E')^{N+1}, E^{N+1})$ is Hausdorff, [15, Prop III. 11]. The weak-$*$ topology on $(E')^{N+1}$ induces a topology on A, denoted $\sigma_+((E')^{N+1}, E^{N+1})$, which is also Hausdorff. We define the functional $\mathcal{L} : A \times B \to \mathbb{R}$ by

$$\mathcal{L}\left(\boldsymbol{s}, (\mu, \boldsymbol{\alpha})\right) = \sum_{i=0}^{N} \langle F_i^n + \mu + \alpha_i, s_i \rangle_{E,E'} - \int_\Omega \mu\omega \, \mathrm{d}\boldsymbol{x} - \sum_{i=0}^{N} \alpha_i m_i.$$

Given $(\mu, \boldsymbol{\alpha}) \in B$, the partial map $\boldsymbol{s} \mapsto \mathcal{L}(\boldsymbol{s}, (\mu, \boldsymbol{\alpha}))$ is affine thus convex. Because $\mathbf{F}^n \in E^{N+1}$ and given $(\mu, \boldsymbol{\alpha}) \in B$, this map is moreover continuous on $(E')^{N+1}$ for the weak-$*$ topology, thus in particular lower semi-continuous on A for the induced weak-$*$ topology $\sigma_+((E')^{N+1}, E^{N+1})$. Given $\boldsymbol{s} \in A$, the partial map $(\mu, \boldsymbol{\alpha}) \mapsto \mathcal{L}\left(\boldsymbol{s}, (\mu, \boldsymbol{\alpha})\right)$ is affine hence concave. It is always true that sup inf $\le$ inf sup, which together with (59) and (61) gives

$$\sup_{(\mu,\boldsymbol{\alpha})\in B} \inf_{\boldsymbol{s}\in A} \mathcal{L}\left(\boldsymbol{s}, (\mu, \boldsymbol{\alpha})\right) \le \inf_{\boldsymbol{s}\in A} \sup_{(\mu,\boldsymbol{\alpha})\in B} \mathcal{L}\left(\boldsymbol{s}, (\mu, \boldsymbol{\alpha})\right) = \underline{I} < +\infty.$$

To check the compactness assumption (63) in Theorem 3.5, fix

$$S = \left\{ \boldsymbol{s} \in A \;\middle|\; \mathcal{L}\big(\boldsymbol{s}, (\|\mathbf{F}^n\|_{E^{N+1}}, \mathbf{1})\big) \le M = 1 + \underline{I} \right\}.$$

For any $\boldsymbol{s} \in S$ we have

$$\sum_{i=0}^{N} \langle F_i^n + \mu + \alpha_i, s_i \rangle_{E,E'} \le 1 + \underline{I} + \|\mathbf{F}^n\|_{E^{N+1}} \|\omega\|_{L^1} + \sum_{i=0}^{N} m_i.$$

Therefore, using $\sum_{i=0}^{N} m_i = \|\omega\|_{L^1}$, we obtain that

$$\sum_{i=0}^{N} \langle F_i^n + \|\mathbf{F}^n\|_\infty + 1, s_i \rangle_{E,E'} \le 1 + \underline{I} + \left(1 + \|\mathbf{F}^n\|_{E^{N+1}}\right) \|\omega\|_{L^1}.$$

Since $F_i^n + \|\mathbf{F}^n\|_\infty \ge 0$ and $s_i \ge 0$ we get

$$0 \le \langle 1, s_i \rangle_{E,E'} \le \sum_{i=0}^{N} \langle 1, s_i \rangle_{E,E'} \le 1 + \underline{I} + \left(1 + \|\mathbf{F}^n\|_\infty\right) \|\omega\|_{L^1} = C. \tag{64}$$

Since $s_i \ge 0$ one readily checks that $\langle 1, s_i \rangle_{E,E'} = \|s_i\|_{E'}$, hence (64) gives the bounds $\|s_i\|_{E'} \le C$. By the Banach-Alaoglu-Bourbaki Theorem, the ball $\mathcal{B}_C = \left\{ \boldsymbol{s} \in (E')^{N+1} \middle| \quad \|s\|_{(E')^{N+1}} \le C \right\}$ is compact in $(E')^{N+1}$ for the weak-$*$ topology $\sigma((E')^{N+1}, E^{N+1})$. Observe now that $(E'_+)^{N+1} = \left\{ \boldsymbol{s} \in (E')^{N+1} \;\middle|\; \boldsymbol{s} \ge 0 \right\}$ is closed for the weak-$*$ topology, as an intersection of pre-images of closed sets by continuous applications. Thus $\mathcal{B}_C \cap (E'_+)^{N+1}$ is compact in A for the induced weak-$*$ topology $\sigma_+((E')^{N+1}, E^{N+1})$. Since $\mathcal{L}(., (\mu_\star, \boldsymbol{\alpha}_\star))$ is weakly-$*$ continuous, the sublevelset S is moreover closed in A as the preimage of $(-\infty, C]$. Thus we conclude that $S = S \cap (E'_+)^{N+1} \cap \mathcal{B}_C$ is compact in A as the intersection of the closed set S with the compact set $\mathcal{B}_C \cap (E'_+)^{N+1}$. As a conclusion, the min-max Theorem 3.5 gives

$$\sup_{(\mu,\boldsymbol{\alpha})\in B} \inf_{\boldsymbol{s}\in A} \mathcal{L}\left(\boldsymbol{s}, (\mu, \boldsymbol{\alpha})\right) = \inf_{\boldsymbol{s}\in A} \sup_{(\mu,\boldsymbol{\alpha})\in B} \mathcal{L}\left(\boldsymbol{s}, (\mu, \boldsymbol{\alpha})\right).$$

Combining with (61) and (62) yields $\underline{I} = \overline{J}$.

Step 2: $\overline{J}$ is achieved. Let $(\boldsymbol{\alpha}^k)_{k\in\mathbb{N}}$ be a maximizing sequence for $\overline{J}$, there exists $k_\star$ such that for any $k \geq k_\star$, one has

$$-\int_\Omega \mu_{\boldsymbol{\alpha}^k}(\boldsymbol{x})\omega(\boldsymbol{x})\mathrm{d}\boldsymbol{x} - \sum_{i=0}^N \alpha_i^k m_i \geq \overline{J} - \frac{1}{2},$$

where $-\mu_{\boldsymbol{\alpha}^k}(\boldsymbol{x}) = \min\limits_{i\in\{0,\dots,N\}}\{F_i^n(\boldsymbol{x}) + \alpha_i^k\}$. Choosing $j_k \in \{0,\dots,N\}$ such that $\alpha_{j_k}^k = \min\limits_i \alpha_i^k = \underline{\alpha}^k$, there holds

$$-\mu_{\boldsymbol{\alpha}^k} \leq F_{j_k}^n + \underline{\alpha}^k \leq \|\mathbf{F}^n\|_{E^{N+1}} + \underline{\alpha}^k.$$

Therefore

$$\sum_{i=0}^N \alpha_i^k m_i \leq \frac{1}{2} - \overline{J} + \|\mathbf{F}^n\|_{E^{N+1}} \int_\Omega \omega(\boldsymbol{x})\mathrm{d}\boldsymbol{x} + \int_\Omega \underline{\alpha}^k \omega(\boldsymbol{x})\mathrm{d}\boldsymbol{x}. \tag{65}$$

According to Remark 3.4, we can set $\underline{\alpha}^k = 0$ without lost of generality, which implies in particular $\alpha_i^k \geq 0$ for any $i \in \{0,\dots,N\}$. Combining with (65) and since $m_i > 0$ for all $i \in \{0,\dots,N\}$, it ensures that $(\boldsymbol{\alpha}^k)_{k\in\mathbb{N}}$ is bounded in $\mathbb{R}^n$. We can thus extract a convergent subsequence and denote its limit by $\boldsymbol{\alpha}^\infty$. The continuity of $\mathcal{J}$ yields $\mathcal{J}(\boldsymbol{\alpha}^\infty) = \overline{J}$, hence $\mathcal{J}$ reaches its supremum.

Step 3: optimality conditions. Thanks to Remark 3.4 we shift $(\mu^\infty, \boldsymbol{\alpha}^\infty) \rightsquigarrow (\mu^\infty + \delta, \boldsymbol{\alpha}^\infty + \delta.\mathbf{1})$ into $(\mu^n, \boldsymbol{\alpha}^n)$ such that $\alpha_0^n = 0$ and $\mathcal{J}(\boldsymbol{\alpha}^n) = \overline{J}$. Then we use $\sum_{i=0}^N s_i^n = \omega$, $\int_\Omega s_i^n = m_i$, and $\underline{I} = \overline{J}$ to infer

$$\begin{aligned}
\underline{I} = \mathcal{I}(\boldsymbol{s}^n) &= \int_\Omega \mathbf{F}^n \cdot \boldsymbol{s}^n \mathrm{d}\boldsymbol{x} = \overline{J} = \mathcal{J}(\boldsymbol{\alpha}^n) \\
&= -\int_\Omega \mu^n \omega \mathrm{d}\boldsymbol{x} - \sum_{i=0}^N \alpha_i^n m_i = -\sum_{i=0}^N \int_\Omega (\mu^n + \alpha_i^n)\, s_i \mathrm{d}\boldsymbol{x}
\end{aligned}$$

which leads to

$$\sum_{i=0}^N \int_\Omega (F_i^n(\boldsymbol{x}) + \mu^n(\boldsymbol{x}) + \alpha_i^n)\, s_i(\boldsymbol{x})\mathrm{d}\boldsymbol{x} = 0. \tag{66}$$

Setting

$$\lambda_i^n(\boldsymbol{x}) := F_i^n(\boldsymbol{x}) + \mu^n(\boldsymbol{x}) + \alpha_i^n, \qquad i \in \{0,\dots,N\},$$

equation (66) rewrites

$$\sum_{i=0}^N \int_\Omega \lambda_i^n s_i(\boldsymbol{x})\mathrm{d}\boldsymbol{x} = 0. \tag{67}$$

The definition of $\boldsymbol{\lambda}^n = (\lambda_i^n)_{0\leq i\leq N}$ proves (56). Remark that (60) implies

$$\lambda_i^n \geq 0, \ a.e. \text{ in } \Omega.$$

In particular for $i = 0$, this provides that $\lambda_0^n = F_0^n + \mu^n + \alpha_0^n = \mu^n \geq 0$, since we normalized $\alpha_0^n = 0$ and by definition $F_0^n \equiv 0$. Notice that, since $F_i^n \in H^1(\Omega)$ (cf. Lemma 3.2), we have $\mu^n = -\min\limits_i\{F_i^n + \alpha_i^n\} \in H^1(\Omega)$ with

$$\boldsymbol{\nabla}\mu^n = -\chi_{\{\mu^n = -F_i^n - \alpha_i^n\}} \boldsymbol{\nabla}(F_i^n + \alpha_i^n). \tag{68}$$

Therefore, $\lambda_i = F_i^n + \mu^n + \alpha_i^n$ belongs to $H^1(\Omega)$ as well. We deduce from (67) and the non-negativity of λ_i^n and s_i^n that for all $i \in \{0, \dots, N\}$ (and up to negligible sets), there holds

$$s_i^n(\boldsymbol{x}) > 0 \quad \Rightarrow \quad \lambda_i^n(\boldsymbol{x}) = 0 \quad \Rightarrow \quad \mu^n(\boldsymbol{x}) = -F_i^n(\boldsymbol{x}) - \alpha_i^n.$$

Together with (68) and the definition of $\boldsymbol{\lambda}^n$, this yields

$$\begin{aligned} \text{in } \{s_i^n > 0\}: \quad \boldsymbol{\nabla}\lambda_i^n &= \boldsymbol{\nabla}\left(F_i^n + \mu^n + \alpha_i^n\right) = \boldsymbol{\nabla}F_i^n + \boldsymbol{\nabla}\mu^n \\ &= \boldsymbol{\nabla}F_i^n - \chi_{\{\mu^n = -F_i^n - \alpha_i^n\}}\boldsymbol{\nabla}(F_i^n + \alpha_i^n) = \boldsymbol{\nabla}F_i^n - \boldsymbol{\nabla}F_i^n = 0, \end{aligned}$$

for all $i \in \{0, \dots, N\}$ and the proof is complete. $\square$

Remark 3.6. *Remark that $\boldsymbol{\alpha}^n$ is the Lagrange multiplier for the mass of each phase ($s_i[\Omega] = m_i$), μ^n is the Lagrange multiplier corresponding to the total saturation constraint ($\sum s_i = \omega$) and $\boldsymbol{\lambda}^n$ is the Lagrange multiplier associated to the non-negativity of each phase ($s_i \geq 0$). The choice $\alpha_0^n = 0$ leads $\lambda_0^n = \mu^n$.*

3.2. **The discrete capillary pressure law and pressure estimates.** In this section, some calculations in the Riemannian settings (Ω, d_i) will be carried out. In order to make them as readable as possible, we have to introduce a few basics. We refer to [44, Chapter 14] for a more detail presentation.

Let $i \in \{0, \dots, N\}$, then consider the Riemannian geometry (Ω, d_i), and let $\boldsymbol{x} \in \Omega$, then we denote by $g_{i,\boldsymbol{x}} : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ the local metric tensor defined by

$$g_{i,\boldsymbol{x}}(\boldsymbol{v}, \boldsymbol{v}) = \mu_i \mathbb{K}^{-1}(\boldsymbol{x})\boldsymbol{v} \cdot \boldsymbol{v} = \mathbb{G}_i(\boldsymbol{x})\boldsymbol{v} \cdot \boldsymbol{v}, \qquad \forall \boldsymbol{v} \in \mathbb{R}^d. \tag{69}$$

In this framework, the gradient $\boldsymbol{\nabla}_{g_i}\varphi$ of a function $\varphi \in \mathcal{C}^1(\Omega)$ is defined by

$$\varphi(\boldsymbol{x} + h\boldsymbol{v}) = \varphi(\boldsymbol{x}) + h g_{i,\boldsymbol{x}}\left(\boldsymbol{\nabla}_{g_{i,\boldsymbol{x}}}\varphi(\boldsymbol{x}), \boldsymbol{v}\right) + o(h), \qquad \forall \boldsymbol{v} \in \mathbb{S}^{d-1},\ \forall \boldsymbol{x} \in \Omega.$$

It is easy to check that this leads to the formula

$$\boldsymbol{\nabla}_{g_i}\varphi = \frac{1}{\mu_i}\mathbb{K}\boldsymbol{\nabla}\varphi, \tag{70}$$

where $\boldsymbol{\nabla}\varphi$ stands for the usual (euclidean) gradient. The formula (70) can be extended to Lipschitz continuous functions φ thanks to Rademacher theorem.

For φ belonging to $\mathcal{C}^2$, we can also define the Hessian $D^2_{g_i}\varphi$ of φ in the Riemannian setting by

$$g_{i,\boldsymbol{x}}\left(D^2_{g_i}\varphi(\boldsymbol{x}) \cdot \boldsymbol{v}, \boldsymbol{v}\right) = \left.\frac{d^2}{dt^2}\varphi(\boldsymbol{\gamma}_t)\right|_{t=0}$$

for any geodesic $\boldsymbol{\gamma}_t = \exp_{\boldsymbol{x}}(t\boldsymbol{v})$ starting from $\boldsymbol{x}$ with initial speed $\boldsymbol{v}$.

Denote by φ_i^n the backward Kantorovich potential sending s_i^n to s_i^{n-1} associated to the cost $\frac{1}{2}d_i^2$, $i \in \{0, \dots, N\}$. By the usual definition of the Wasserstein distance through the Monge problem, one has

$$W_i^2(s_i^n, s_i^{n-1}) = \int_\Omega d_i^2(\boldsymbol{x}, \mathbf{t}_i^n(\boldsymbol{x}))s_i^n(\boldsymbol{x})\mathrm{d}\boldsymbol{x},$$

where $\mathbf{t}_i^n$ denotes the optimal map sending s_i^n on s_i^{n-1}. It follows from [44, Theorem 10.41] that

$$\mathbf{t}_i^n(\boldsymbol{x}) = \exp_{i,\boldsymbol{x}}\left(-\boldsymbol{\nabla}_{g_i}\varphi_i^n(\boldsymbol{x})\right), \qquad \forall \boldsymbol{x} \in \Omega. \tag{71}$$

Moreover, using the definition of the exponential and the relation (70), one gets that

$$d_i^2(\boldsymbol{x}, \exp_{i,\boldsymbol{x}}(-\boldsymbol{\nabla}_{g_i}\varphi_i^n(\boldsymbol{x})) = g_{i,\boldsymbol{x}}\left(\boldsymbol{\nabla}_{g_i}\varphi_i^n(\boldsymbol{x}), \boldsymbol{\nabla}_{g_i}\varphi_i^n(\boldsymbol{x})\right) = \frac{1}{\mu_i}\mathbb{K}(\boldsymbol{x})\boldsymbol{\nabla}\varphi_i^n(\boldsymbol{x})\cdot\boldsymbol{\nabla}\varphi_i^n(\boldsymbol{x}),$$

we obtain the formula

$$W_i^2(s_i^n, s_i^{n-1}) = \int_\Omega \frac{s_i^n}{\mu_i}\mathbb{K}\boldsymbol{\nabla}\varphi_i^n\cdot\boldsymbol{\nabla}\varphi_i^n \mathrm{d}\boldsymbol{x}, \qquad \forall i\in\{0,\dots,N\}. \tag{72}$$

We have now introduced the necessary material in order to reconstruct the phase and capillary pressures. This is the purpose of the following Proposition 3.7 and of Corollary 3.8

Proposition 3.7. *For $n\geq 1$ let $\varphi_i^n : s_i^n \to s_i^{n-1}$ be the (backward) Kantorovich potentials from the previous Lemma 3.2, then there exists $\boldsymbol{h} = (h_0^n,\dots,h_N^n)\in H^1(\Omega)^{N+1}$ such that*

(i) $\boldsymbol{\nabla}h_i^n = -\frac{\boldsymbol{\nabla}\varphi_i^n}{\tau}$ for $\mathrm{d}s_i^n$-a.e. $\boldsymbol{x}\in\Omega$
(ii) $h_i^n(\boldsymbol{x}) - h_0^n(\boldsymbol{x}) = \pi_i^n(\boldsymbol{x}) + \Psi_i(\boldsymbol{x}) - \Psi_0(\boldsymbol{x})$ for $\mathrm{d}\boldsymbol{x}$-a.e. $\boldsymbol{x}\in\Omega$, $i\in\{0,\dots,N\}$
(iii) there exists C depending only on $\Omega,\Pi,\omega,\mathbb{K},(\mu_i)_i$, and $\boldsymbol{\Psi}$ such that, for all $n\geq 1$ and all $\tau>0$, one has

$$\|\boldsymbol{h}^n\|^2_{H^1(\Omega)^{N+1}} \leq C\left(1 + \frac{\boldsymbol{W}^2(\boldsymbol{s}^n, \boldsymbol{s}^{n-1})}{\tau^2} + \sum_{i=0}^N \frac{\mathcal{H}_\omega(s_i^{n-1}) - \mathcal{H}_\omega(s_i^n)}{\tau}\right).$$

Proof. Let φ_i^n be the Kantorovich potentials from Lemma 3.2, and $\boldsymbol{\alpha}^n\in\mathbb{R}^{N+1}$, $\boldsymbol{\lambda}^n\in\left(L^\infty\cap H^1(\Omega)\right)^{N+1}$ as in Lemma 3.3. Note from $F_0^n\equiv 0$ and (56) that we implicitly normalized $\alpha_0^n = 0$. Setting

$$h_i^n := -\frac{\varphi_i^n}{\tau} + \lambda_i^n - \alpha_i^n, \qquad \forall i\in\{0,\dots,N\}, \tag{73}$$

we have $h_i^n\in H^1(\Omega)$ as the sum of Lipschitz functions (the Kantorovich potentials φ_i^n) and H^1 functions (the corrector λ_i^n, see Lemma 3.3). We deduce next from (47) and (56) that

$$\begin{aligned} h_i^n - h_0^n &= (\lambda_i^n - \lambda_0^n - \alpha_i^n + 0) - \left(\frac{\varphi_i^n}{\tau} - \frac{\varphi_0^n}{\tau}\right) \\ &= F_i^n - \left(\frac{\varphi_i^n}{\tau} - \frac{\varphi_0^n}{\tau}\right) = \pi_i^n + \Psi_i - \Psi_0 \end{aligned} \tag{74}$$

for all $i\in\{1,\dots,N\}$, which is exactly our statement (ii).

For (i), we simply use (57) and the definition (73) of h_i^n to compute

$$\boldsymbol{\nabla}h_i^n = -\frac{\boldsymbol{\nabla}\varphi_i^n}{\tau} + \boldsymbol{\nabla}\lambda_i^n = -\frac{\boldsymbol{\nabla}\varphi_i^n}{\tau} \quad \text{for } \mathrm{d}s_i^n \text{ a.e. } \boldsymbol{x}\in\Omega, \qquad \forall i\in\{0,\dots,N\}. \tag{75}$$

Denote by

$$\mathcal{U}_i = \left\{\boldsymbol{x}\in\Omega \;\middle|\; s_i^n \geq \frac{\omega_\star}{N+1}\right\},$$

then since $\sum_{i=0}^N s_i^n = \omega \geq \omega_\star$, one gets that, up to a negligible set,

$$\bigcup_{i=0}^N \mathcal{U}_i = \Omega, \quad \text{hence} \quad (\mathcal{U}_i)^c \subset \bigcup_{j\neq i}\mathcal{U}_j. \tag{76}$$

In order to establish (iii), we first estimate $\boldsymbol{\nabla} h_0^n$. To this end, we write

$$\|\boldsymbol{\nabla} h_0^n\|_{L^2}^2 \leq \frac{1}{\kappa_\star}\int_\Omega \mathbb{K}\boldsymbol{\nabla} h_0^n \cdot \boldsymbol{\nabla} h_0^n \mathrm{d}\boldsymbol{x} \leq A + B, \tag{77}$$

where we have set

$$A = \frac{1}{\kappa_\star}\int_{\mathcal{U}_0} \mathbb{K}\boldsymbol{\nabla} h_0^n \cdot \boldsymbol{\nabla} h_0^n \mathrm{d}\boldsymbol{x}, \qquad B = \frac{1}{\kappa_\star}\int_{(\mathcal{U}_0)^c} \mathbb{K}\boldsymbol{\nabla} h_0^n \cdot \boldsymbol{\nabla} h_0^n \mathrm{d}\boldsymbol{x}.$$

Thanks to (75), one has $\boldsymbol{\nabla} h_0^n = -\frac{\boldsymbol{\nabla}\varphi_0}{\tau}$ on $\mathcal{U}_0 \subset \Omega$. Therefore,

$$A \leq \frac{(N+1)\mu_0}{\omega_\star\kappa_\star}\int_{\mathcal{U}_0} \frac{s_0^n}{\mu_0}\mathbb{K}\boldsymbol{\nabla} h_0^n \cdot \boldsymbol{\nabla} h_0^n \mathrm{d}\boldsymbol{x} \leq \frac{(N+1)\mu_0}{\tau^2\omega_\star\kappa_\star}\int_\Omega \frac{s_0^n}{\mu_0}\mathbb{K}\boldsymbol{\nabla}\varphi_0^n \cdot \boldsymbol{\nabla}\varphi_0^n \mathrm{d}\boldsymbol{x}.$$

Then it results from formula (72) that

$$A \leq \frac{C}{\tau^2} W_0^2(s_0^n, s_0^{n-1}) \tag{78}$$

where C depends neither on n nor on τ. Combining (76) and (74), we infer

$$B \leq \frac{1}{\kappa_\star}\sum_{i=1}^N \int_{\mathcal{U}_i} \mathbb{K}\boldsymbol{\nabla}[h_i^n - (\pi_i^n + \Psi_i - \Psi_0)] \cdot \boldsymbol{\nabla}[h_i^n - (\pi_i^n + \Psi_i - \Psi_0)]\mathrm{d}\boldsymbol{x}.$$

Using $(a+b+c)^2 \leq 3(a^2+b^2+c^2)$ and (3), we get that

$$B \leq \frac{3}{\kappa_\star}\sum_{i=1}^N \int_{\mathcal{U}_i} \mathbb{K}\boldsymbol{\nabla} h_i \cdot \boldsymbol{\nabla} h_i \mathrm{d}\boldsymbol{x} + \frac{3\kappa^\star}{\kappa_\star}\sum_{i=1}^N \left(\|\boldsymbol{\nabla}\pi_i^n\|_{L^2}^2 + \|\boldsymbol{\nabla}(\Psi_i - \Psi_0)\|_{L^2}^2\right). \tag{79}$$

Similar calculations to those carried out to estimate A yield

$$\int_{\mathcal{U}_i} \mathbb{K}\boldsymbol{\nabla} h_i \cdot \boldsymbol{\nabla} h_i \mathrm{d}\boldsymbol{x} \leq \frac{C}{\tau^2} W_i^2(s_i^n, s_i^{n-1})$$

for some C depending neither on n, i nor on τ. Combining this inequality with Lemma 2.2 and the regularity of $\boldsymbol{\Psi}$, we get from (79) that

$$B \leq C\left(1 + \frac{\boldsymbol{W}^2(\boldsymbol{s}^n, \boldsymbol{s}^{n-1})}{\tau^2} + \sum_{i=0}^N \frac{\mathcal{H}_\omega(s_i^{n-1}) - \mathcal{H}_\omega(s_i^n)}{\tau}\right) \tag{80}$$

for some C not depending on n and τ (here we also used $1/\tau \leq 1/\tau^2$ for small τ in the W^2 terms). Gathering (78) and (80) in (77) provides

$$\|\boldsymbol{\nabla} h_0^n\|_{L^2}^2 \leq C\left(1 + \frac{\boldsymbol{W}^2(\boldsymbol{s}^n, \boldsymbol{s}^{n-1})}{\tau^2} + \sum_{i=0}^N \frac{\mathcal{H}_\omega(s_i^{n-1}) - \mathcal{H}_\omega(s_i^n)}{\tau}\right).$$

Note that (i)(ii) remain invariant under subtraction of the same constant $h_0^n, h_i^n \rightsquigarrow h_0^n - C, h_i^n - C$, as the gradients remain unchanged in (i) and only the differences $h_i^n - h_0^n$ appear in (ii) for $i \in \{1 \dots N\}$. We can therefore assume without loss of generality that $\int_\Omega h_0^n \mathrm{d}\boldsymbol{x} = 0$. Hence by Poincaré-Wirtinger inequality, we get that

$$\|h_0^n\|_{H^1}^2 \leq C\|\boldsymbol{\nabla} h_0^n\|_{L^2}^2 \leq C\left(1 + \frac{\boldsymbol{W}^2(\boldsymbol{s}^n, \boldsymbol{s}^{n-1})}{\tau^2} + \sum_{i=0}^N \frac{\mathcal{H}_\omega(s_i^{n-1}) - \mathcal{H}_\omega(s_i^n)}{\tau}\right).$$

Finally, from (ii) $h_i^n = h_0^n + \pi_i^n + \Psi_i - \Psi_0$, the smoothness of Ψ_i, Ψ_0, and using again the estimate (30) for $\|\boldsymbol{\pi}^n\|^2$ we finally get that for all $i \in \{1, \dots, N\}$, one has

$$\|h_i^n\|_{H^1}^2 \leq C(\|h_0^n\|_{H^1}^2 + \|\pi_i^n\|_{H^1}^2 + \|\Psi_i - \Psi_0\|_{H^1}^2)$$
$$\leq C\left(1 + \frac{\boldsymbol{W}^2(\boldsymbol{s}^n, \boldsymbol{s}^{n-1})}{\tau^2} + \sum_{i=0}^{N} \frac{\mathcal{H}_\omega(s_i^{n-1}) - \mathcal{H}_\omega(s_i^n)}{\tau}\right),$$

and the proof of Proposition 3.7 is complete. □

We can now define the phase pressures $(p_i^n)_{i=0^N}$ by setting

$$p_i^n := h_i^n - \Psi_i, \qquad \forall i \in \{0, \dots, N\}. \tag{81}$$

The following corollary is a straightforward consequence of Proposition 3.7 and of the regularity of Ψ_i.

Corollary 3.8. *The phase pressures $\boldsymbol{p}^n = (p_i^n)_{0 \leq i \leq N} \in H^1(\Omega)^{N+1}$ satisfy*

$$\|\boldsymbol{p}^n\|_{H^1(\Omega)}^2 \leq C\left(1 + \frac{\boldsymbol{W}^2(\boldsymbol{s}^n, \boldsymbol{s}^{n-1})}{\tau^2} + \sum_{i=0}^{N} \frac{\mathcal{H}_\omega(s_i^{n-1}) - \mathcal{H}_\omega(s_i^n)}{\tau}\right) \tag{82}$$

for some C depending only on $\Omega, \Pi, \omega, \mathbb{K}, (\mu_i)_i$, and $\boldsymbol{\Psi}$ (but neither on n nor on τ), and the capillary pressure relations are fulfilled:

$$p_i^n - p_0^n = \pi_i^n, \qquad \forall i \in \{1, \dots, N\}. \tag{83}$$

The following Lemma is a first step towards the recovery of the PDEs.

Lemma 3.9. *There exists C depending depending only on $\Omega, \Pi, \omega, \mathbb{K}, (\mu_i)_i$, and $\boldsymbol{\Psi}$ (but neither on n nor on τ) such that, for all $i \in \{0, \dots, N\}$ and all $\xi \in \mathcal{C}^2(\overline{\Omega})$, one has*

$$\left|\int_\Omega \left(s_i^n - s_i^{n-1}\right)\xi \mathrm{d}\boldsymbol{x} + \tau \int_\Omega \mathbb{K}\frac{s_i^n}{\mu_i}\boldsymbol{\nabla}\left(p_i^n + \Psi_i\right) \cdot \boldsymbol{\nabla}\xi \mathrm{d}\boldsymbol{x}\right|$$
$$\leq C W_i^2(s_i^n, s_i^{n-1}) \|D_{g_i}^2 \xi\|_\infty. \tag{84}$$

Proof. Let φ_i^n denote the (backward) optimal Kantorovich potential from Lemma 3.2 sending s_i^n to s_i^{n-1}, and let $\mathbf{t}_i^n$ be the optimal transportation map as in (71). For fixed $\xi \in \mathcal{C}^2(\overline{\Omega})$ let us first Taylor expand (in the g_i Riemaniann framework)

$$\left|\xi(\mathbf{t}_i^n(\boldsymbol{x})) - \xi(\boldsymbol{x}) + \frac{1}{\mu_i}\mathbb{K}(\boldsymbol{x})\boldsymbol{\nabla}\xi(\boldsymbol{x}) \cdot \boldsymbol{\nabla}\varphi_i^n(\boldsymbol{x})\right| \leq \frac{1}{2}\|D_{g_i}^2 \xi\|_\infty d_i^2(\boldsymbol{x}, \mathbf{t}_i^n(\boldsymbol{x})).$$

Using the definition of the pushforward $s_i^{n-1} = \mathbf{t}_i^n \# s_i^n$, we then compute

$$\left|\int_\Omega (s_i^n(\boldsymbol{x}) - s_i^{n-1}(\boldsymbol{x}))\xi(\boldsymbol{x})\mathrm{d}\boldsymbol{x} - \int_\Omega \frac{\mathbb{K}(\boldsymbol{x})}{\mu_i}\boldsymbol{\nabla}\xi(x) \cdot \boldsymbol{\nabla}\varphi_i^n(\boldsymbol{x}) s_i^n(\boldsymbol{x})\mathrm{d}\boldsymbol{x}\right|$$
$$= \left|\int_\Omega (\xi(\boldsymbol{x}) - \xi(\mathbf{t}_i^n(\boldsymbol{x})) s_i^n(\boldsymbol{x})\mathrm{d}\boldsymbol{x} - \int_\Omega \frac{\mathbb{K}(\boldsymbol{x})}{\mu_i}\boldsymbol{\nabla}\xi(x) \cdot \boldsymbol{\nabla}\varphi_i^n(\boldsymbol{x}) s_i^n(\boldsymbol{x})\mathrm{d}\boldsymbol{x}\right|$$
$$\leq \int_\Omega \frac{1}{2}\|D_{g_i}^2 \xi\|_\infty d_i^2(\boldsymbol{x}, \mathbf{t}_i^n(\boldsymbol{x})) s_i^n(\boldsymbol{x})\mathrm{d}\boldsymbol{x} = \frac{1}{2}\|D_{g_i}^2 \xi\|_\infty W_i^2(s_i^n, s_i^{n-1}).$$

From Proposition 3.7(i) we have $\boldsymbol{\nabla}\varphi_i^n = -\tau\boldsymbol{\nabla}h_i^n$ for $\mathrm{d}s_i^n$ a.e. $\boldsymbol{x} \in \Omega$, thus by the definition (81) of p_i^n, we get $\boldsymbol{\nabla}\varphi^n = -\tau\boldsymbol{\nabla}(p_i^n + \Psi_i)$. Substituting in the previous inequality gives exactly (84) and the proof is complete. □

4. Convergence towards a weak solution

The goal is now to prove the convergence of the piecewise constant w.r.t. time approximate solutions $\boldsymbol{s}^\tau$ defined by (26) towards a weak solution $\boldsymbol{s}$ as τ tends to 0. Similarly, the τ superscript denotes the piecewise constant interpolation of any previous discrete quantity (e.g. $p_i^\tau(t)$ stands for the piecewise constant time interpolation of the discrete pressures p_i^n). In what follows, we will also use the notations $\boldsymbol{s}^{\tau *} = (s_1^\tau, \dots, s_N^\tau) \in L^\infty((0,T); \boldsymbol{\mathcal{X}}^*)$ and $\boldsymbol{\pi}^\tau = \boldsymbol{\pi}(\boldsymbol{s}^{\tau *}, \boldsymbol{x})$.

4.1. **Time integrated estimates.** We immediately deduce from (29) that

$$\boldsymbol{W}(\boldsymbol{s}^\tau(t_2), \boldsymbol{s}^\tau(t_1)) \le C|t_2 - t_1 + \tau|^{\frac{1}{2}}, \qquad \forall t_2 \ge 0, \forall t_1 \in [0, t_2]. \tag{85}$$

From the total saturation $\sum_{i=0}^{N} s_i^n(\boldsymbol{x}) = \omega(\boldsymbol{x}) < 1$ and $s_i^\tau \ge 0$, we have the L^∞ estimates

$$0 \le s_i^\tau \le \omega \le 1 \quad \text{a.e. in } Q \text{ for all } i \in \{0, \dots, N\}. \tag{86}$$

Lemma 4.1. *There exists C depending only on $\Omega, T, \Pi, \omega, \mathbb{K}, (\mu_i)_i$, and $\boldsymbol{\Psi}$ such that*

$$\|\boldsymbol{p}^\tau\|^2_{L^2((0,T);H^1(\Omega)^{N+1})} + \|\boldsymbol{\pi}^\tau\|^2_{L^2((0,T);H^1(\Omega)^N)} \le C. \tag{87}$$

Proof. Summing (82) from $n = 1$ to $n = N_\tau := \lceil T/\tau \rceil$, we get

$$\begin{aligned}
\|\boldsymbol{p}^\tau\|^2_{L^2(H^1)} &= \sum_{n=1}^{N_\tau} \tau \|\boldsymbol{p}^n\|^2_{H^1} \\
&\le C \sum_{n=1}^{N_\tau} \tau \left(1 + \frac{\boldsymbol{W}^2(\boldsymbol{s}^n, \boldsymbol{s}^{n-1})}{\tau^2} + \sum_{i=0}^{N_\tau} \frac{\mathcal{H}_\omega(s_i^{n-1}) - \mathcal{H}_\omega(s_i^n)}{\tau} \right) \\
&\le C \left((T+1) + \sum_{n=1}^{N_\tau} \frac{\boldsymbol{W}^2(\boldsymbol{s}^n, \boldsymbol{s}^{n-1})}{\tau} + \sum_{i=0}^{N} \Big(\mathcal{H}_\omega(s_i^0) - \mathcal{H}_\omega(s_i^{N_\tau}) \Big) \right).
\end{aligned}$$

We use that

$$0 \ge \mathcal{H}_\omega(s) \ge -\frac{1}{e}\|\omega\|_{L^1} \ge -\frac{|\Omega|}{e}, \qquad \forall s \in L^\infty(\Omega) \text{ with } 0 \le s \le \omega$$

together with the total square distance estimate (28) to infer that $\|\boldsymbol{p}\|^2_{L^2(H^1)} \le C$. The proof is identical for the capillary pressure $\boldsymbol{\pi}^\tau$ (simply summing the one-step estimate from Lemma 2.2). □

4.2. **Compactness of approximate solutions.** We denote by $H' = H^1(\Omega)'$.

Lemma 4.2. *For each $i \in \{0, \dots, N\}$, there exists C depending only on Ω, Π, $\boldsymbol{\Psi}$, $\mathbb{K}$, and μ_i (but not on τ) such that*

$$\|s_i^\tau(t_2) - s_i^\tau(t_1)\|_{H'} \le C|t_2 - t_1 + \tau|^{\frac{1}{2}}, \qquad \forall t_2 \in [0,T], \forall t_1 \in [0, t_2].$$

Proof. Thanks to (86), we can apply [37, Lemma 3.4] to get

$$\left| \int_\Omega f(\boldsymbol{x}) \mathrm{d}(s_i^\tau(t) - s_i^\tau(s))(\boldsymbol{x}) \right| \le \|\boldsymbol{\nabla} f\|_{L^2(\Omega)} W_{\mathrm{ref}}(s_i^\tau(t), s_i^\tau(s)), \qquad \forall f \in H^1(\Omega).$$

Thus by duality and thanks to the distance estimate (85) and to the lower bound in (22), we obtain that

$$\|s_i^\tau(t) - s_i^\tau(s)\|_{'} \le W_{\mathrm{ref}}(s_i^\tau(t), s_i^\tau(s)) \le C W_i(s_i^\tau(t), s_i^\tau(s)) \le C|t-s+\tau|^{\frac{1}{2}}$$

for some fluctuating quantity C depending only on Ω, Π, $(\rho_i)_i$, $\mathbf{g}$, $(\mu_i)_i$, $\mathbb{K}$. ☐

From the previous equi-continuity in time, we deduce full compactness of the capillary pressure:

Lemma 4.3. *The family $(\boldsymbol{\pi}^\tau)_{\tau>0}$ is sequentially relatively compact in $L^2(Q)^N$.*

Proof. We use Alt & Luckhaus' trick [1] (an alternate solution would consist in slightly adapting the nonlinear time compactness results [38, 8] to our context). Let $h > 0$ be a small time shift, then the Lipschitz continuity of the capillary pressure function $\boldsymbol{\pi}$ yields

$$\begin{aligned}
&\|\boldsymbol{\pi}^\tau(.+h) - \boldsymbol{\pi}^\tau(.)\|^2_{L^2((0,T-h);L^2(\Omega)^N)} \\
&\quad \le \frac{1}{\kappa_\star}\int_0^{T-h}\int_\Omega (\boldsymbol{\pi}^\tau(t+h,\boldsymbol{x}) - \boldsymbol{\pi}^\tau(t,\boldsymbol{x}))\cdot(\boldsymbol{s}^{\tau*}(t+h,\boldsymbol{x}) - \boldsymbol{s}^{\tau*}(t,\boldsymbol{x}))\mathrm{d}\boldsymbol{x}\mathrm{d}t \\
&\qquad\qquad \le \frac{2\sqrt{T}}{\kappa_\star}\|\boldsymbol{\pi}^\tau\|_{L^2((0,T);H^1(\Omega)^N)}\|\boldsymbol{s}^{\tau*}(\cdot+h,\cdot) - \boldsymbol{s}^{\tau*}\|_{L^\infty((0,T-h);H')^N}.
\end{aligned}$$

Then it follows from Lemmas 4.1 and 4.2 that there exists $C > 0$, depending neither on h nor on τ, such that

$$\|\boldsymbol{\pi}^\tau(\cdot+h,\cdot) - \boldsymbol{\pi}^\tau\|_{L^2((0,T-h);L^2(\Omega)^N)} \le C|h+\tau|^{1/2}.$$

On the other hand, the (uniform w.r.t. τ) $L^2((0,T);H^1(\Omega)^N)$- and $L^\infty(Q)^N$-estimates on $\boldsymbol{\pi}^\tau$ ensure that

$$\|\boldsymbol{\pi}^\tau(\cdot,\cdot+y)) - \boldsymbol{\pi}^\tau\|_{L^2(0,T;L^2)} \le C\sqrt{|\boldsymbol{y}|}(1+\sqrt{|\boldsymbol{y}|}), \qquad \forall\, \boldsymbol{y}\in\mathbb{R}^d,$$

where $\boldsymbol{\pi}^\tau$ is extended by 0 outside Ω. This allows to apply Kolmogorov's compactness theorem (see, for instance, [29]) and entails the desired relative compactness. ☐

4.3. **Identification of the limit.** In this section we prove our main Theorem 1.2, and the proof goes in two steps: we first retrieve strong convergence of the phase contents $\boldsymbol{s}^\tau \to \boldsymbol{s}$ and weak convergence of the pressures $\boldsymbol{p}^\tau \rightharpoonup \boldsymbol{p}$, and then use the strong-weak limit of products to show that the limit is a weak solution. All along this section, $(\tau_k)_{k\ge1}$ denotes a sequence of times steps tending to 0 as k tends to ∞.

Lemma 4.4. *There exist $\boldsymbol{s} = (s_i)_{0\le i\le N}$ belonging to $L^\infty(Q)^{N+1}$ with $\boldsymbol{s}(\cdot,t)\in \boldsymbol{\mathcal{X}}\cap\boldsymbol{\mathcal{A}}$ for a.e. $t\in(0,T)$, and $\boldsymbol{p}$ belonging to $L^2((0,T);H^1(\Omega)^{N+1})$ such that, up to an unlabeled subsequence, the following convergence properties hold:*

$$\boldsymbol{s}^{\tau_k} \underset{k\to\infty}{\longrightarrow} \boldsymbol{s} \quad \text{a.e. in } Q, \tag{88}$$

$$\boldsymbol{\pi}^{\tau_k} \underset{k\to\infty}{\rightharpoonup} \boldsymbol{\pi}(\boldsymbol{s}^*,\cdot) \quad \text{weakly in } L^2((0,T);H^1(\Omega)^N), \tag{89}$$

$$\boldsymbol{p}^{\tau_k} \underset{k\to\infty}{\rightharpoonup} \boldsymbol{p} \quad \text{weakly in } L^2((0,T);H^1(\Omega)^{N+1}). \tag{90}$$

Moreover, the capillary pressure relations (5) hold.

Proof. From Lemma 4.3, we can assume that $\boldsymbol{\pi}^{\tau_k} \to \boldsymbol{z}$ strongly in $L^2(Q)^N$ for some limit $\boldsymbol{z}$, thus a.e. up to the extraction of an additional subsequence. Since $\boldsymbol{z} \mapsto \boldsymbol{\phi}(\boldsymbol{z}, \boldsymbol{x}) = \boldsymbol{\pi}^{-1}(\boldsymbol{z}, \boldsymbol{x})$ is continuous, we have that

$$\boldsymbol{s}^{\tau_k *} = \boldsymbol{\phi}(\boldsymbol{\pi}^{\tau_k}, \boldsymbol{x}) \underset{k\to\infty}{\longrightarrow} \boldsymbol{\phi}(\boldsymbol{\pi}, \boldsymbol{x}) =: \boldsymbol{s}^* \quad \text{a.e. in } Q.$$

In particular, this yields $\boldsymbol{\pi}^{\tau_k} \underset{k\to\infty}{\longrightarrow} \boldsymbol{\pi}(\boldsymbol{s}^*, \cdot)$ a.e. in Q. Since we had the total saturation $\sum_{i=0}^{N} s_i^{\tau_k}(t, \boldsymbol{x}) = \omega(\boldsymbol{x})$, we conclude that the first component $i = 0$ converges pointwise as well. Therefore, (88) holds. Thanks to Lebesgue's dominated convergence theorem, it is easy to check that $\boldsymbol{s}(\cdot, t) \in \boldsymbol{\mathcal{X}} \cap \boldsymbol{\mathcal{A}}$ for a.e. $t \in (0, T)$. The convergences (89) and (90) are straightforward consequences of Lemma 4.1. Lastly, it follows from (83) that

$$p_i^{\tau_k} - p_0^{\tau_k} = \boldsymbol{\pi}_i(\boldsymbol{s}^{\tau_k *}, \cdot), \qquad \forall i \in \{1, \dots, N\},\ \forall k \geq 1.$$

We can pass to the limit $k \to \infty$ in the above relation thanks to (89)–(90) and infer

$$p_i - p_0 = \boldsymbol{\pi}_i(\boldsymbol{s}^*, \boldsymbol{x}) \quad \text{in } L^2((0,T); H^1(\Omega)), \qquad \forall i \in \{1, \dots, N\}.$$

This immediately implies (5) as claimed. □

Lemma 4.5. *Up to the extraction of an additional subsequence, the limit* $\boldsymbol{s}$ *of* $(\boldsymbol{s}^{\tau_k})_{k\geq 1}$ *can be supposed to belong to* $\mathcal{C}([0,T]; \boldsymbol{\mathcal{A}})$ *where* $\boldsymbol{\mathcal{A}}$ *is equipped with the metric* $\boldsymbol{W}$*. Moreover,* $\boldsymbol{W}(\boldsymbol{s}^{\tau_k}(t), \boldsymbol{s}(t)) \underset{k\to\infty}{\longrightarrow} 0$ *for all* $t \in [0,T]$*.*

Proof. Let $i \in \{0, \dots, N\}$, then it follows from the uniform boundedness of s_i (cf. (86)) that for all $t \in [0,T]$, the sequence $(s_i^{\tau_k})_k$ is weakly compact in $L^1(\Omega)$. It is also compact in $\mathcal{A}_i$ equipped with the metric W_i due to the continuity of W_i with respect to the weak convergence in $L^1(\Omega)$ (this is for instance a consequence of [42, Theorem 5.10] together with the equivalence of W_i with W_{ref} stated at (22)). Thanks to (85), one has

$$\limsup_{k\to\infty} W_i\left(s_i^{\tau_k}(t_2), s_i^{\tau_k}(t_1)\right) \leq |t_2 - t_1|^{1/2}, \qquad \forall t_1, t_2 \in [0,T].$$

Applying a refined version of the Arzelà-Ascoli theorem [5, Prop. 3.3.1] then provides the desired result. □

In order to conclude the proof of Theorem 1.2, it only remains to show that $\boldsymbol{s} = \lim \boldsymbol{s}^{\tau_k}$ and $\boldsymbol{p} = \lim \boldsymbol{p}^{\tau_k}$ satisfy the weak formulation (13). This is the purpose of the last statement of the paper.

Proposition 4.6. *Let* $(\tau_k)_{k\geq 1}$ *be a sequence such that the convergences in Lemmas 4.4 and 4.5 hold, then the limit* $\boldsymbol{s}$ *of* $(\boldsymbol{s}^{\tau_k})_{k\geq 1}$ *is a weak solution in the sense of Definition 1.1 (with* $-\rho_i \mathbf{g}$ *replaced by* $+\nabla \Psi_i$ *in the general case).*

Proof. Let $(t_1, t_2) \in [0,T]^2$ with $t_1 \leq t_2$, and denote $n_{j,k} = \left\lceil \frac{t_j}{\tau_k} \right\rceil$ and $\tilde{t}_j = n_{j,k}\tau_k$ for $j \in \{1, 2\}$. Let $\xi \in \mathcal{C}^2(\overline{\Omega})$ be arbitrary, then summing (84) from $n = n_{1,k} + 1$ to

$n = n_{2,k}$ yields, for all $i \in \{0, \dots, N\}$,

$$\begin{aligned}(91)\qquad \int_\Omega (s_i^{\tau_k}(t_2) - s_i^{\tau_k}(t_1))\xi \mathrm{d}\boldsymbol{x} &= \sum_{n=n_{1,k}+1}^{n_{2,k}} \int_\Omega (s_i^n - s_i^{n-1})\xi \mathrm{d}\boldsymbol{x} \\ &= -\int_{\tilde{t}_1}^{\tilde{t}_2} \int_\Omega \frac{s_i^{\tau_k}}{\mu_i} \mathbb{K}\boldsymbol{\nabla}\left(p_i^{\tau_k} + \Psi_i\right) \cdot \boldsymbol{\nabla}\xi \mathrm{d}\boldsymbol{x}\mathrm{d}t + \mathcal{O}\left(\sum_{n=n_{1,k}+1}^{n_{2,k}} W_i^2(s_i^n, s_i^{n-1})\right).\end{aligned}$$

Since $0 \le \tilde{t}_j - t_j \le \tau_k$, and since $\frac{s_i^{\tau_k}}{\mu_i}\mathbb{K}\boldsymbol{\nabla}\left(p_i^{\tau_k} + \Psi_i\right) \cdot \boldsymbol{\nabla}\xi$ is uniformly bounded in $L^2(Q)$, one has

$$\begin{aligned}\int_{\tilde{t}_1}^{\tilde{t}_2} \int_\Omega \frac{s_i^{\tau_k}}{\mu_i} \mathbb{K}\boldsymbol{\nabla}\left(p_i^{\tau_k} + \Psi_i\right) \cdot \boldsymbol{\nabla}\xi \mathrm{d}\boldsymbol{x}\mathrm{d}t \\ = \int_{t_1}^{t_2} \int_\Omega \frac{s_i^{\tau_k}}{\mu_i} \mathbb{K}\boldsymbol{\nabla}\left(p_i^{\tau_k} + \Psi_i\right) \cdot \boldsymbol{\nabla}\xi \mathrm{d}\boldsymbol{x}\mathrm{d}t + \mathcal{O}(\tau_k).\end{aligned}$$

Combining the above estimate with the total square distance estimate (28) in (91), we obtain that for all $t_1, t_2 \in [0, T]$ with $t_1 \le t_2$,

$$(92)\qquad \int_\Omega (s_i^{\tau_k}(t_2) - s_i^{\tau_k}(t_1))\xi \mathrm{d}\boldsymbol{x} + \int_{t_1}^{t_2} \int_\Omega \frac{s_i^{\tau_k}}{\mu_i} \mathbb{K}\boldsymbol{\nabla}\left(p_i^{\tau_k} + \Psi_i\right) \cdot \boldsymbol{\nabla}\xi \mathrm{d}\boldsymbol{x}\mathrm{d}t = \mathcal{O}\left(\tau_k\right).$$

Thanks to Lemma 4.5, and since the convergence in $(\mathcal{A}_i, W_i)$ is equivalent to the narrow convergence of measures (i.e., the convergence in $\mathcal{C}(\overline{\Omega})'$, see for instance [42, Theorem 5.10]), we get that

$$(93)\qquad \int_\Omega (s_i^{\tau_k}(t_2) - s_i^{\tau_k}(t_1))\xi \mathrm{d}\boldsymbol{x} \underset{k\to\infty}{\longrightarrow} \int_\Omega (s_i(t_2) - s_i(t_1))\xi \mathrm{d}\boldsymbol{x}.$$

Moreover, thanks to Lemma 4.4, one has

$$(94)\qquad \int_{t_1}^{t_2} \int_\Omega \frac{s_i^{\tau_k}}{\mu_i} \mathbb{K}\boldsymbol{\nabla}\left(p_i^{\tau_k} + \Psi_i\right) \cdot \boldsymbol{\nabla}\xi \mathrm{d}\boldsymbol{x}\mathrm{d}t \underset{k\to\infty}{\longrightarrow} \int_{t_1}^{t_2} \int_\Omega \frac{s_i}{\mu_i} \mathbb{K}\boldsymbol{\nabla}\left(p_i + \Psi_i\right) \cdot \boldsymbol{\nabla}\xi \mathrm{d}\boldsymbol{x}\mathrm{d}t.$$

Gathering (92)–(94) yields, for all $\xi \in \mathcal{C}^2(\overline{\Omega})$ and all $t_1, t_2 \in [0, T]$ with $t_1 \le t_2$,

$$(95)\qquad \int_\Omega (s_i(t_2) - s_i(t_1))\xi \mathrm{d}\boldsymbol{x} + \int_{t_1}^{t_2} \int_\Omega \frac{s_i}{\mu_i} \mathbb{K}\boldsymbol{\nabla}\left(p_i + \Psi_i\right) \cdot \boldsymbol{\nabla}\xi \mathrm{d}\boldsymbol{x}\mathrm{d}t = 0.$$

In order to conclude the proof, it remains to check that the formulation (95) is stronger the formulation (13). Let $\varepsilon > 0$ be a time step (not related to the one appearing in the minimization scheme (25)), and set $L_\varepsilon = \left\lfloor \frac{T}{\varepsilon} \right\rfloor$. Let $\phi \in \mathcal{C}_c^\infty(\overline{\Omega} \times [0, T))$, one sets $\phi_\ell = \phi(\cdot, \ell\varepsilon)$ for $\ell \in \{0, \dots, L_\varepsilon\}$. Since $t \mapsto \phi(\cdot, t)$ is compactly supported in $[0, T)$, then there exists $\varepsilon^\star > 0$ such that $\phi_{L_\varepsilon} \equiv 0$ for all $\varepsilon \in (0, \varepsilon^\star]$. Then define by

$$\phi^\varepsilon : \left\{ \begin{array}{ccll} \overline{\Omega} \times [0, T] & \to & \mathbb{R} & \\ (\boldsymbol{x}, t) & \mapsto & \phi_\ell(\boldsymbol{x}) & \text{if } t \in [\ell\varepsilon, (\ell+1)\varepsilon). \end{array} \right.$$

Choose $t_1 = \ell\varepsilon$, $t_2 = (\ell+1)\varepsilon$, $\xi = \phi_\ell$ in (95) and sum over $\ell \in \{0, \dots, L_\varepsilon - 1\}$. This provides

$$(96)\qquad A(\varepsilon) + B(\varepsilon) = 0, \qquad \forall \varepsilon > 0.$$

where

$$A(\varepsilon) = \sum_{\ell=0}^{L_\varepsilon - 1} \int_\Omega \left(s_i((\ell+1)\varepsilon) - s_i(\ell\varepsilon) \right) \phi^\ell \mathrm{d}\boldsymbol{x},$$

$$B(\varepsilon) = \iint_Q \frac{s_i}{\mu_i} \mathbb{K} \boldsymbol{\nabla} \left(p_i + \Psi_i \right) \cdot \boldsymbol{\nabla} \phi^\varepsilon \mathrm{d}\boldsymbol{x} \mathrm{d}t.$$

Due to the regularity of ϕ, $\boldsymbol{\nabla}\phi^\varepsilon$ converges uniformly towards ϕ as ε tends to 0, so that

$$B(\varepsilon) \underset{\varepsilon \to 0}{\longrightarrow} \iint_Q \frac{s_i}{\mu_i} \mathbb{K} \boldsymbol{\nabla} \left(p_i + \Psi_i \right) \cdot \boldsymbol{\nabla} \phi \mathrm{d}\boldsymbol{x} \mathrm{d}t. \tag{97}$$

Reorganizing the first term and using that $\phi_{L_\varepsilon} \equiv 0$, we get that

$$A(\varepsilon) = - \sum_{\ell=1}^{L_\varepsilon} \varepsilon \int_\Omega s_i(\ell\varepsilon) \frac{\phi_\ell - \phi_{\ell-1}}{\varepsilon} \mathrm{d}\boldsymbol{x} - \int_\Omega s_i^0 \phi(\cdot, 0) \mathrm{d}\boldsymbol{x}.$$

It follows from the continuity of $t \mapsto s_i(\cdot, t)$ in $\mathcal{A}_i$ equipped with W_i and from the uniform convergence of

$$(\boldsymbol{x}, t) \mapsto \frac{\phi_\ell(\boldsymbol{x}) - \phi_{\ell-1}(\boldsymbol{x})}{\varepsilon} \quad \text{if } t \in [(\ell-1)\varepsilon, \ell\varepsilon)$$

towards $\partial_t \phi$ that

$$A(\varepsilon) \underset{\varepsilon \to 0}{\longrightarrow} - \iint_Q s_i \partial_t \phi \mathrm{d}\boldsymbol{x} \mathrm{d}t - \int_\Omega s_i^0 \phi(\cdot, 0) \mathrm{d}\boldsymbol{x}. \tag{98}$$

Combining (96)–(98) shows that the weak formulation (13) is fulfilled. □

Acknowledgements. This project was supported by the ANR GEOPOR project (ANR-13-JS01-0007-01). TOG was supported by the ANR ISOTACE project (ANR-12-MONU-0013). LM was supported by the Portuguese Science Fundation through FCT fellowship SFRH/BPD/88207/2012 and the UT Austin — Portugal CoLab project. Part of this work was carried out during the stay of CC and TOG at CAMGSD, Instituto Superior Técnico, Universidade de Lisboa. The authors thank Quentin Mérigot for the numerous fruitful discussion they shared.

References

[1] H. W. Alt and S. Luckhaus. Quasilinear elliptic-parabolic differential equations. *Math. Z.*, 183(3):311–341, 1983.

[2] B. Amaziane, M. Jurak, and A. Vrbaški. Existence for a global pressure formulation of water-gas flow in porous media. *Electron. J. Differential Equations*, pages No. 102, 22, 2012.

[3] B. Amaziane, M. Jurak, and A. Žgaljić Keko. Modeling compositional compressible two-phase flow in porous media by the concept of the global pressure. *Comput. Geosci.*, 18(3-4):297–309, 2014.

[4] L. Ambrosio and N. Gigli. A user's guide to optimal transport. In *Modelling and optimisation of flows on networks*, volume 2062 of *Lecture Notes in Math.*, pages 1–155. Springer, Heidelberg, 2013.

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

[6] L. Ambrosio, E. Mainini, and S. Serfaty. Gradient flow of the Chapman-Rubinstein-Schatzman model for signed vortices. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 28(2):217–246, 2011.

[7] L. Ambrosio and S. Serfaty. A gradient flow approach to an evolution problem arising in superconductivity. *Comm. Pure Appl. Math.*, 61(11):1495–1539, 2008.
[8] B. Andreianov, C. Cancès, and A. Moussa. A nonlinear time compactness result and applications to discretization of degenerate parabolic-elliptic PDEs. HAL: hal-01142499, 2015.
[9] S. N. Antoncev and V. N. Monahov. Three-dimensional problems of transient two-phase filtration in inhomogeneous anisotropic porous media. *Dokl. Akad. Nauk SSSR*, 243(3):553–556, 1978.
[10] J. Bear and Y. Bachmat. *Introduction to modeling of transport phenomena in porous media.* Kluwer Academic Publishers, Dordrecht, The Netherlands, 1990.
[11] A. Blanchet. A gradient flow approach to the Keller-Segel systems. RIMS Kokyuroku's lecture notes, 2014.
[12] A. Blanchet, V. Calvez, and J. A. Carrillo. Convergence of the mass-transport steepest descent scheme for the subcritical Patlak-Keller-Segel model. *SIAM J. Numer. Anal.*, 46(2):691–721, 2008.
[13] F. Bolley, I. Gentil, and A. Guillin. Uniform convergence to equilibrium for granular media. *Arch. Ration. Mech. Anal.*, 208(2):429–445, 2013.
[14] H. Brézis. *Opérateurs maximaux monotones et semi-groupes de contractions dans les espaces de Hilbert.* North-Holland Publishing Co., Amsterdam-London; American Elsevier Publishing Co., Inc., New York, 1973. North-Holland Mathematics Studies, No. 5. Notas de Matemática (50).
[15] H. Brezis. *Functional analysis, Sobolev spaces and partial differential equations.* Springer Science & Business Media, 2010.
[16] C. Cancès and T. Gallouët. On the time continuity of entropy solutions. *J. Evol. Equ.*, 11(1):43–55, 2011.
[17] C. Cancès, T. O. Gallouët, and L. Monsaingeon. The gradient flow structure of immiscible incompressible two-phase flows in porous media. *C. R. Acad. Sci. Paris Sér. I Math.*, 353:985–989, 2015.
[18] J. A. Carrillo, M. DiFrancesco, A. Figalli, T. Laurent, and D. Slepčev. Global-in-time weak measure solutions and finite-time aggregation for nonlocal interaction equations. *Duke Math. J.*, 156(2):229–271, 2011.
[19] G. Chavent. A new formulation of diphasic incompressible flows in porous media. In *Applications of methods of functional analysis to problems in mechanics (Joint Sympos., IUTAM/IMU, Marseille, 1975)*, pages 258–270. Lecture Notes in Math., 503. Springer, Berlin, 1976.
[20] G. Chavent. A fully equivalent global pressure formulation for three-phases compressible flows. *Appl. Anal.*, 88(10-11):1527–1541, 2009.
[21] G. Chavent and J. Jaffré. *Mathematical Models and Finite Elements for Reservoir Simulation*, volume 17. North-Holland, Amsterdam, stud. math. appl. edition, 1986.
[22] G. Chavent and G. Salzano. Un algorithme pour la détermination de perméabilités relatives triphasiques satisfaisant une condition de différentielle totale. Technical Report 355, INRIA, 1985.
[23] Z. Chen. Degenerate two-phase incompressible flow. I. Existence, uniqueness and regularity of a weak solution. *J. Differential Equations*, 171(2):203–232, 2001.
[24] E. De Giorgi. New problems on minimizing movements. In *Boundary value problems for partial differential equations and applications*, volume 29 of *RMA Res. Notes Appl. Math.*, pages 81–98. Masson, Paris, 1993.
[25] J. Dolbeault, B. Nazaret, and G. Savaré. A new class of transport distances between measures. *Calc. Var. Partial Differential Equations*, 34(2):193–231, 2009.
[26] P. Fabrie and M. Saad. Existence de solutions faibles pour un modèle d'écoulement triphasique en milieu poreux. *Ann. Fac. Sci. Toulouse Math. (6)*, 2(3):337–373, 1993.
[27] G. Gagneux and M. Madaune-Tort. *Analyse mathématique de modèles non linéaires de l'ingénierie pétrolière*, volume 22 of *Mathématiques & Applications (Berlin) [Mathematics & Applications]*. Springer-Verlag, Berlin, 1996. With a preface by Charles-Michel Marle.
[28] N. Gigli and F. Otto. Entropic Burgersd5 equation via a minimizing movement scheme based on the Wasserstein metric. *Calc. Var. Partial Differential Equations*, 47(1-2):181–206, 2013.
[29] H. Hance-Olsen and H. Holden. The Kolmogorov-Riesz compactness theorem. *Expo. Math.*, 28:385–394, 2010.

[30] R. Jordan, D. Kinderlehrer, and F. Otto. The variational formulation of the Fokker-Planck equation. *SIAM J. Math. Anal.*, 29(1):1–17, 1998.

[31] D. Kinderlehrer, L. Monsaingeon, and X. Xu. A Wasserstein gradient flow approach to Poisson-Nernst-Planck equations. arXiv:1501.04437, to appear in ESAIM: COCV.

[32] O. A. Ladyženskaja, V. A. Solonnikov, and N. N. Ural′ceva. *Linear and quasilinear equations of parabolic type*. Translated from the Russian by S. Smith. Translations of Mathematical Monographs, Vol. 23. American Mathematical Society, Providence, R.I., 1968.

[33] P. Laurençot and B.-V. Matioc. A gradient flow approach to a thin film approximation of the muskat problem. *Calc. Var. Partial Differential Equations*, 47((1-2)):319–341, 2013.

[34] S. Lisini. Nonlinear diffusion equations with variable coefficients as gradient flows in Wasserstein spaces. *ESAIM Control Optim. Calc. Var.*, 15(3):712–740, 2009.

[35] S. Lisini, D. Matthes, and G. Savaré. Cahn-Hilliard and thin film equations with nonlinear mobility as gradient flows in weighted-Wasserstein metrics. *J. Differential Equations*, 253(2):814–850, 2012.

[36] D. Matthes, R. J. McCann, and G. Savaré. A family of nonlinear fourth order equations of gradient flow type. *Communications in Partial Differential Equations*, 34(11):1352–1397, 2009.

[37] B. Maury, A. Roudneff-Chupin, and F. Santambrogio. A macroscopic crowd motion model of gradient flow type. *Math. Models Methods Appl. Sci.*, 20(10):1787–1821, 2010.

[38] A. Moussa. Some variants of the classical Aubin-Lions Lemma. *J. Evol. Equ.*, 16(1):65–93, 2016.

[39] F. Otto. Dynamics of labyrinthine pattern formation in magnetic fluids: a mean-field theory. *Arch. Rational Mech. Anal.*, 141(1):63–103, 1998.

[40] F. Otto. The geometry of dissipative evolution equations: the porous medium equation. *Comm. Partial Differential Equations*, 26(1-2):101–174, 2001.

[41] E. Sandier and S. Serfaty. Gamma-convergence of gradient flows with applications to Ginzburg-Landau. *Comm. Pure Appl. Math.*, 57(12):1627–1672, 2004.

[42] F. Santambrogio. *Optimal Transport for Applied Mathematicians: Calculus of Variations, PDEs, and Modeling*. Progress in Nonlinear Differential Equations and Their Applications 87. Birkhäuser Verlag, Basel, 1 edition, 2015.

[43] S. Simons. *Minimax and monotonicity*, volume 1693 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1998.

[44] C. Villani. *Optimal transport*, volume 338 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 2009. Old and new.

[45] J. Zinsl and D. Matthes. Exponential convergence to equilibrium in a coupled gradient flow system modeling chemotaxis. *Anal. PDE*, 8(2):425–466, 2015.

Clément Cancès (clement.cances@inria.fr). Team RAPSODI, Inria Lille – Nord Europe, 40 av. Halley, F-59650 Villeneuve d'Ascq, France.

Thomas O. Gallouët (thomas.gallouet@polytechnique.edu). CMLS, UMR 7640, Ecole Polytechnique, FR-91128 Palaiseau Cedex.

Léonard Monsaingeon (leonard.monsaingeon@tecnico.ulisboa.pt). CAMGSD, Instituto Superior Técnico, Universidade de Lisboa, Av. Rovisco Pais, 1049-001 Lisboa, Portugal