ANALYSIS & PDE

Volume 10 No. 8 2017

CLÉMENT CANCÈS, THOMAS O. GALLOUËT
AND LÉONARD MONSAINGEON

INCOMPRESSIBLE IMMISCIBLE MULTIPHASE FLOWS IN POROUS MEDIA
A VARIATIONAL APPROACH

msp

ANALYSIS AND PDE
Vol. 10, No. 8, 2017
dx.doi.org/10.2140/apde.2017.10.1845

INCOMPRESSIBLE IMMISCIBLE MULTIPHASE FLOWS IN POROUS MEDIA A VARIATIONAL APPROACH

Clément Cancès, Thomas O. Gallouët and Léonard Monsaingeon

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 nonnegative measures with prescribed masses, endowed with some tensorial Wasserstein distance. We show the convergence of the approximation obtained by a minimization scheme á la R. Jordan, D. Kinderlehrer and F. Otto (*SIAM J. Math. Anal.* **29**:1 (1998) 1–17). This allows us to obtain a new existence result for a physically well-established system of PDEs consisting of 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.

1. Introduction

Equations for multiphase flows in porous media. We consider a convex 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 [Bear and Bachmat 1990, Chapter 5]. Let $T > 0$ be an arbitrary finite time horizon. 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 can be written as

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

The filtration speed of each phase is assumed to be given by Darcy's law

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

In the above relation, $\boldsymbol{g}$ is the gravity vector, μ_i denotes the constant 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, that is, $\mathbb{K} = \mathbb{K}^T$, and uniformly positive definite: there exist $\kappa_\star, \kappa^\star > 0$ such that

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

MSC2010: 35K65, 35A15, 49K20, 76S05.
Keywords: multiphase porous media flows, Wasserstein gradient flows, constrained parabolic system, minimizing movement scheme.

The pore volume is supposed to be saturated by the fluid mixture

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

where the porosity $\omega : \overline{\Omega} \to (0, 1)$ of the surrounding porous matrix 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 set $\boldsymbol{s} = (s_0, \dots, s_N)$,

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

and

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

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

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

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

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

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

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

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

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 $\Pi : \boldsymbol{\Upsilon} \to \mathbb{R}_+$; that is,

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

We assume that Π is uniformly convex with respect to 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 $\boldsymbol{s}^*, \hat{\boldsymbol{s}}^* \in \Delta^*(\boldsymbol{x})$, one has

$$\tfrac{1}{2}\varpi^\star |\hat{\boldsymbol{s}}^* - \boldsymbol{s}^*|^2 \ge \Pi(\hat{\boldsymbol{s}}^*, \boldsymbol{x}) - \Pi(\boldsymbol{s}^*, \boldsymbol{x}) - \boldsymbol{\pi}(\boldsymbol{s}^*, \boldsymbol{x}) \cdot (\hat{\boldsymbol{s}}^* - \boldsymbol{s}^*) \ge \tfrac{1}{2}\varpi_\star |\hat{\boldsymbol{s}}^* - \boldsymbol{s}^*|^2, \tag{6}$$

where we introduced the notation

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

The relation (6) implies that $\boldsymbol{\pi}$ is monotone and injective with respect to its first variable. Denoting by

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

the inverse of $\boldsymbol{\pi}(\cdot, \boldsymbol{x})$, it follows from (6) that

$$0 < \frac{1}{\varpi^\star} \le \mathbb{J}_z \boldsymbol{\phi}(z, \boldsymbol{x}) \le \frac{1}{\varpi_\star} \quad \forall \boldsymbol{x} \in \overline{\Omega}, \; \forall z \in \boldsymbol{\pi}(\Delta^*(\boldsymbol{x}), \boldsymbol{x}), \tag{7}$$

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

$$|\nabla_{\boldsymbol{x}}\boldsymbol{\phi}(z, \boldsymbol{x})| \le M_{\boldsymbol{\phi}} \quad \forall \boldsymbol{x} \in \overline{\Omega},\ \forall z \in \boldsymbol{\pi}(\Delta^*(\boldsymbol{x}), \boldsymbol{x}), \tag{8}$$

where $\nabla_{\boldsymbol{x}}$ denotes the gradient with respect to the second variable only.

The problem is complemented with no-flux boundary conditions

$$\boldsymbol{v}_i \cdot \boldsymbol{n} = 0 \quad \text{on } \partial\Omega \times (0, T),\ \forall i \in \{0, \dots, N\}, \tag{9}$$

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

$$s_i(\cdot, 0) = s_i^0 \quad \forall i \in \{0, \dots, N\},\ \text{with } \sum_{i=0}^{N} s_i^0 = \omega \ \text{ a.e. in } \Omega. \tag{10}$$

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

$$\int_\Omega s_i(\boldsymbol{x}, t)\, \mathrm{d}\boldsymbol{x} = \int_\Omega s_i^0(\boldsymbol{x})\, \mathrm{d}\boldsymbol{x} =: m_i > 0 \quad \forall i \in \{0, \dots, N\}. \tag{11}$$

We can now give a proper definition of what we call a weak solution to the problem (1)–(2), (4)–(5), and (9)–(10).

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

$$\iint_Q s_i\, \partial_t \phi\, \mathrm{d}\boldsymbol{x}\, \mathrm{d}t + \int_\Omega s_i^0 \phi(\cdot, 0)\, \mathrm{d}\boldsymbol{x} - \iint_Q \frac{s_i}{\mu_i} \mathbb{K}(\nabla p_i - \rho_i \boldsymbol{g}) \cdot \nabla\phi\, \mathrm{d}\boldsymbol{x}\, \mathrm{d}t = 0. \tag{12}$$

Wasserstein gradient flow of the energy.

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

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

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 to $\{\sigma < \omega\} \cup \mathbb{R}_+^{N+1}$ is designed so that the energy density only depends on the relative composition of the fluid mixture. However, this extension is somehow arbitrary, and, as it will appear in the sequel, it has no influence on the flow since the solution $\boldsymbol{s}$ remains in $\boldsymbol{\mathcal{X}}$; i.e., $\sum_{i=0}^N s_i = \omega$. In our previous note [Cancès et al. 2015] 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 notation, we introduce the functions

$$\Psi_i : \overline{\Omega} \to \mathbb{R}_+, \quad \boldsymbol{x} \mapsto -\rho_i \boldsymbol{g} \cdot \boldsymbol{x}, \quad \forall i \in \{0, \dots, N\},$$

and

$$\boldsymbol{\Psi} : \overline{\Omega} \to \mathbb{R}_+^{N+1}, \quad \boldsymbol{x} \mapsto (\Psi_0(\boldsymbol{x}), \dots, \Psi_N(\boldsymbol{x})).$$

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

$$\mathcal{E}(\boldsymbol{s}) = \int_\Omega (\Pi(\boldsymbol{s}, \boldsymbol{x}) + \boldsymbol{s} \cdot \boldsymbol{\Psi})\, \mathrm{d}\boldsymbol{x} \geq 0 \quad \forall \boldsymbol{s} \in L^1(\Omega; \mathbb{R}^{N+1}). \tag{13}$$

Note moreover that $\mathcal{E}(\boldsymbol{s}) < \infty$ if and only if $\boldsymbol{s} \geq 0$ and $\sigma \leq \omega$ a.e. in Ω. It follows from the mass conservation (11) 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 nonnegligible subset A of Ω such that $\sigma < \omega$ on A; then necessarily, there must be a nonnegligible subset B of Ω such that $\sigma > \omega$ so that the above equation holds, and hence $\mathcal{E}(\boldsymbol{s}) = +\infty$. Therefore,

$$\mathcal{E}(\boldsymbol{s}) < \infty \iff \boldsymbol{s} \in \boldsymbol{\mathcal{X}}. \tag{14}$$

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 $\boldsymbol{h} = (h_i)_{0 \leq i \leq N}$, we have $\boldsymbol{h}$ belongs to the subdifferential $\partial_{\boldsymbol{s}} \mathcal{E}(\boldsymbol{s})$ of $\mathcal{E}$ at $\boldsymbol{s}$; i.e.,

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

The reverse inclusion also holds; hence

$$\partial_{\boldsymbol{s}} \mathcal{E}(\boldsymbol{s}) = \{\boldsymbol{h} : \Omega \to \mathbb{R}^{N+1} \mid h_i - \Psi_i(\boldsymbol{x}) \in \partial_{\boldsymbol{s}} \Pi(\boldsymbol{s}, \boldsymbol{x}) \text{ for a.e. } \boldsymbol{x} \in \Omega\}. \tag{15}$$

Thanks to (14), we know that a configuration $\boldsymbol{s}$ has finite energy if and only if $\boldsymbol{s} \in \boldsymbol{\mathcal{X}}$. 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 $\boldsymbol{s}^* \in \boldsymbol{\mathcal{X}}^*$, which 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 $\boldsymbol{s}^* \in \boldsymbol{\mathcal{X}}^*$.

Geometry of Ω and Wasserstein distance. Inspired by [Lisini 2009], where heterogeneous anisotropic degenerate parabolic equations are studied from a variational point of view, we introduce $N+1$ distances on Ω that take into account the permeability of the porous medium and the phase viscosities. Given two points $\boldsymbol{x}$, $\boldsymbol{y}$ in Ω, we denote by

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

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 Ω by setting

$$d_i(\boldsymbol{x}, \boldsymbol{y}) = \inf_{\boldsymbol{\gamma} \in P(\boldsymbol{x},\boldsymbol{y})} \left(\int_0^1 \mu_i \mathbb{K}^{-1}(\boldsymbol{\gamma}(\tau)) \boldsymbol{\gamma}'(\tau) \cdot \boldsymbol{\gamma}'(\tau) \, \mathrm{d}\tau \right)^{1/2} \quad \forall (\boldsymbol{x}, \boldsymbol{y}) \in \overline{\Omega}. \tag{16}$$

It follows from (3) that

$$\sqrt{\frac{\mu_i}{\kappa^\star}} |\boldsymbol{x} - \boldsymbol{y}| \le d_i(\boldsymbol{x}, \boldsymbol{y}) \le \sqrt{\frac{\mu_i}{\kappa_\star}} |\boldsymbol{x} - \boldsymbol{y}| \quad \forall (\boldsymbol{x}, \boldsymbol{y}) \in \overline{\Omega}^2. \tag{17}$$

For $i \in \{0, \dots, N\}$ we define

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

Given $s_i, \hat{s}_i \in \mathcal{A}_i$, the set of admissible transport plans between s_i and $\hat{s}_i$ is given by

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

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

$$W_i(s_i, \hat{s}_i) = \left(\inf_{\theta_i \in \Gamma(s_i, \hat{s}_i)} \iint_{\Omega \times \Omega} d_i(\boldsymbol{x}, \boldsymbol{y})^2 \, \mathrm{d}\theta_i(\boldsymbol{x}, \boldsymbol{y}) \right)^{1/2}. \tag{18}$$

Due to the permeability tensor $\mathbb{K}(\boldsymbol{x})$, the porous medium Ω might be heterogeneous and anisotropic. Therefore, some directions and areas might be privileged by the fluid motions. This is encoded in the distances d_i we put on Ω. Moreover, the more viscous the phase is, the more costly are its displacements, hence the μ_i in the definition (16) of d_i. But it follows from (17) that

$$\sqrt{\frac{\mu_i}{\kappa^\star}} W_{\mathrm{ref}}(s_i, \hat{s}_i) \le W_i(s_i, \hat{s}_i) \le \sqrt{\frac{\mu_i}{\kappa_\star}} W_{\mathrm{ref}}(s_i, \hat{s}_i) \quad \forall s_i, \hat{s}_i \in \mathcal{A}_i, \tag{19}$$

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

$$W_{\mathrm{ref}}(s_i, \hat{s}_i) = \left(\inf_{\theta_i \in \Gamma(s_i, \hat{s}_i)} \iint_{\Omega \times \Omega} |\boldsymbol{x} - \boldsymbol{y}|^2 \, \mathrm{d}\theta_i(\boldsymbol{x}, \boldsymbol{y}) \right)^{1/2}. \tag{20}$$

With 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

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

Finally for technical reasons we also assume that there exists a smooth extension $\widetilde{\mathbb{K}}$ to $\mathbb{R}^d$ of the permeability tensor such that (3) holds on $\mathbb{R}^d$. This allows us to define distances $\tilde{d}_i$ on the whole $\mathbb{R}^d$ by

$$\tilde{d}_i(\boldsymbol{x},\boldsymbol{y}) = \inf_{\boldsymbol{\gamma}\in\widetilde{P}(\boldsymbol{x},\boldsymbol{y})} \left(\int_0^1 \mu_i \widetilde{\mathbb{K}}^{-1}(\boldsymbol{\gamma}(\tau))\boldsymbol{\gamma}'(\tau)\cdot\boldsymbol{\gamma}'(\tau)\,\mathrm{d}\tau \right)^{1/2} \quad \forall \boldsymbol{x},\boldsymbol{y}\in\mathbb{R}^d, \tag{21}$$

where $\widetilde{P}(\boldsymbol{x},\boldsymbol{y}) = \{\boldsymbol{\gamma}\in C^1([0,1];\mathbb{R}^d) \mid \boldsymbol{\gamma}(0)=\boldsymbol{x} \text{ and } \boldsymbol{\gamma}(1)=\boldsymbol{y}\}$. In the sequel, we assume that the extension $\widetilde{\mathbb{K}}$ of $\mathbb{K}$ is such that

$$\Omega \text{ is geodesically convex in } \mathcal{M}_i = (\mathbb{R}^d, \tilde{d}_i) \text{ for all } i. \tag{22}$$

In particular $\tilde{d}_i = d_i$ on $\Omega\times\Omega$. Since $\widetilde{\mathbb{K}}^{-1}$ is smooth, at least $C_b^2(\mathbb{R}^d)$, the Ricci curvature of the smooth complete Riemannian manifold $\mathcal{M}_i$ is uniformly bounded; i.e., there exists C depending only on $(\mu_i)_{0\le i\le N}$ and $\widetilde{\mathbb{K}}$ such that

$$|\mathrm{Ric}_{\mathcal{M}_i,\boldsymbol{x}}(\boldsymbol{v})| \le C\mu_i \mathbb{K}^{-1}\boldsymbol{v}\cdot\boldsymbol{v} \quad \forall \boldsymbol{x}\in\mathbb{R}^d,\ \forall \boldsymbol{v}\in\mathbb{R}^d. \tag{23}$$

We deduce from the lower bound on the Ricci curvature and on the geodesic convexity of Ω that the Boltzmann relative entropy $\mathcal{H}_\omega$ with respect to ω_i, defined by

$$\mathcal{H}_\omega(s) = \int_{\mathbb{R}^d} s \log\left(\frac{s}{\omega}\right) \mathrm{d}\boldsymbol{x} \quad \text{for all measurable } s:\Omega\to\mathbb{R}_+, \tag{24}$$

is λ_i-displacement convex on $\mathcal{P}^{\mathrm{ac}}(\Omega)$ for some $\lambda_i\in\mathbb{R}$. Here, $\mathcal{P}^{\mathrm{ac}}(\Omega)$ denotes the set of probability measures on Ω that are absolutely continuous with respect to the Lebesgue measure. Then mass scaling implies that $\mathcal{H}_\omega$ is also λ_i-displacement convex on $(\mathcal{A}_i, W_i)$. We refer to [Villani 2009, Chapters 14 and 17] for further details on the Ricci curvature and its links with optimal transportation.

In the homogeneous and isotropic case $\mathbb{K}(\boldsymbol{x}) = \mathrm{Id}$, condition (22) simply amounts to assuming that Ω is convex. A simple sufficient condition implying (22) is given in Appendix A in the isotropic but heterogeneous case $\mathbb{K}(\boldsymbol{x}) = \kappa(\boldsymbol{x})\mathbb{I}_d$.

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.,

$$\frac{\mathrm{d}\boldsymbol{s}}{\mathrm{d}t} \in -\,\mathbf{grad}_{\boldsymbol{W}}\,\mathcal{E}(\boldsymbol{s}) = -\big(\mathrm{grad}_{W_0}\,\mathcal{E}(\boldsymbol{s}),\dots,\mathrm{grad}_{W_N}\,\mathcal{E}(\boldsymbol{s})\big), \tag{25}$$

where $\mathbf{grad}_{\boldsymbol{W}}\,\mathcal{E}(\boldsymbol{s})$ denotes the full Wasserstein gradient of $\mathcal{E}(\boldsymbol{s})$, and $\mathrm{grad}_{W_i}\,\mathcal{E}(\boldsymbol{s})$ stands for the partial gradient of $s_i\mapsto\mathcal{E}(\boldsymbol{s})$ with respect to the Wasserstein distance W_i. The Wasserstein distance W_i was built so that $\dot{\boldsymbol{s}} = (\dot{s}_i)_i \in \mathbf{grad}_{\boldsymbol{W}}\,\mathcal{E}(\boldsymbol{s})$ if and only if there exists $\boldsymbol{h}\in\partial_{\boldsymbol{s}}\mathcal{E}(\boldsymbol{s})$ such that

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

Such a construction was already performed by Lisini in the case of a single equation. Owing to the definitions (13) and (15) 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, the purpose of our note [Cancès et al. 2015].

In order to define rigorously the gradient $\mathbf{grad}_W\, \mathcal{E}$ in (25), $\boldsymbol{\mathcal{A}}$ has to be a Riemannian manifold. The so-called Otto's calculus [2001], see also [Villani 2009, 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; see [Ambrosio et al. 2008]. 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.

Minimizing movement scheme and main result.

The scheme and existence of a solution. For a fixed time-step $\tau > 0$, the so-called minimizing movement scheme [De Giorgi 1993; Ambrosio et al. 2008] or JKO scheme [Jordan et al. 1998] consists in computing recursively $(\boldsymbol{s}^n)_{n\geq 1}$ as the solution to the minimization problem

$$\boldsymbol{s}^n = \underset{\boldsymbol{s}\in\boldsymbol{\mathcal{A}}}{\operatorname{Argmin}}\left(\frac{\boldsymbol{W}(\boldsymbol{s},\boldsymbol{s}^{n-1})^2}{2\tau} + \mathcal{E}(\boldsymbol{s})\right), \tag{26}$$

the initial data $\boldsymbol{s}^0$ being given in (10).

Approximate solution and main result. Anticipating that the JKO scheme (26) is well-posed (this is the purpose of Proposition 2.1 below), we can now define the piecewise constant interpolation $\boldsymbol{s}^\tau \in L^\infty((0,T); \boldsymbol{\mathcal{X}} \cap \boldsymbol{\mathcal{A}})$ by

$$\boldsymbol{s}^\tau(0,\cdot) = \boldsymbol{s}^0 \qquad \text{and} \qquad \boldsymbol{s}^\tau(t,\cdot) = \boldsymbol{s}^n \quad \forall t \in ((n-1)\tau, n\tau],\ \forall n \geq 1. \tag{27}$$

The main result 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* $\boldsymbol{s}$ *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*

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

which is another admissible formulation of gradient flows in metric spaces [Ambrosio et al. 2008].

Goal and positioning of the paper. The aims of the paper are twofold. First, we aim to provide a rigorous foundation to the formal variational approach introduced in the authors' recent note [Cancès et al. 2015]. This gives new insights into the modeling of complex porous media flows and their numerical approximation. Our approach 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, which classically appear in the mathematical study of multiphase flows in porous media; see, for instance, [Chavent 1976; 2009; Antoncev and Monahov 1978; Chavent and Jaffré 1986; Fabrie and Saad 1993; Gagneux and Madaune-Tort 1996; Chen 2001; Amaziane et al. 2012; 2014].

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 any global pressure, we get rid of many structural assumptions on the data, among which is the so-called *total differentiability condition*; see, for instance, Assumption (H3) in [Fabrie and Saad 1993]. 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 [Chavent and Salzano 1985]. However, our approach suffers from another technical difficulty: we are limited to the case of linear relative permeabilities. The extension to the case of nonlinear concave relative permeabilities, i.e., where (1) is replaced by

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

may be reachable thanks to the contributions of Dolbeault, Nazaret, and Savaré [Dolbeault et al. 2009], see also [Zinsl and Matthes 2015b], but we did not push in 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 [Jordan et al. 1998], 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 [Ambrosio et al. 2008] and to Villani's book [2009, Part II] for a complete overview. Applications are numerous. We refer for instance to [Otto 1998] for an application to magnetic fluids, to [Sandier and Serfaty 2004; Ambrosio and Serfaty 2008; Ambrosio et al. 2011] for applications to superconductivity to [Blanchet et al. 2008; Blanchet 2013; Zinsl and Matthes 2015a] for applications to chemotaxis, to [Lisini et al. 2012] for phase field models, to [Maury et al. 2010] for a macroscopic model of crowd motion, to [Bolley et al. 2013] for an application to granular media, to [Carrillo et al. 2011] for aggregation equations, and to [Kinderlehrer et al. 2017] for a model of ionic transport that applies in semiconductors. In the context of porous media flows, this framework has been used by Otto [2001] to study the asymptotic behavior of the porous medium equation, which is a simplified model for the filtration of a gas in a porous medium. The gradient flow approach in Wasserstein metric spaces was used more recently by Laurençot and Matioc [2013] on a thin film approximation model for two-phase flows in porous media. Finally, let us mention that similar ideas were successfully applied for multicomponent systems; see, e.g., [Carlier and Laborde 2015; Laborde 2016; Zinsl and Matthes 2015b; Zinsl 2014].

The variational structure of the system governing incompressible immiscible two-phase flows in porous media was recently depicted by the authors in their note [Cancès et al. 2015]. Whereas the purpose of that paper 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 [2013], 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 detailed in the first subsection, in the second subsection we obtain enhanced regularity estimates thanks to an adaptation of the so-called *flow interchange* technique of Matthes, McCann, and Savaré [Matthes et al. 2009] to our inhomogeneous context. 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 Section 3 to derive the Euler–Lagrange equations that follow from the optimality of $\boldsymbol{s}^n$. Our proof is inspired by the work of Maury, Roudneff-Chupin, and Santambrogio [Maury et al. 2010]. It relies on an intensive use of the dual characterization of the optimal transportation problem and the corresponding Kantorovich potentials. However, additional difficulties arise from the multiphase aspect of our problem, in particular when there are at least three phases (i.e., $N \geq 2$). These are bypassed using a generalized multicomponent bathtub principle (Theorem B.1 in Appendix B) and computing the associated Lagrange multipliers in the first subsection. This key step then allows to define the notion of discrete phase and capillary pressures in the second subsection. 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 with respect to time in the first subsection. In the second subsection, 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 the third subsection 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 (26) 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 (26). 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 (14). Let $(\boldsymbol{s}^{n,k})_k \subset \boldsymbol{\mathcal{A}}$ be a minimizing sequence in (26). Plugging $\boldsymbol{s}^{n-1}$ into (26), 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 (14). Hence, one has $0 \leq s_i^{n,k}(\boldsymbol{x}) \leq \omega(\boldsymbol{x})$ for all k. By the 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 semicontinuity of the Wasserstein distance with respect to weak L^1 convergence is well known, see, e.g., [Santambrogio 2015, Proposition 7.4], and since the energy functional is convex and thus lower semicontinuous, 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 (with respect to 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.

Energy and distance estimates. Plugging $\boldsymbol{s} = \boldsymbol{s}^{n-1}$ into (26) we obtain

$$\frac{\boldsymbol{W}(\boldsymbol{s}^n, \boldsymbol{s}^{n-1})^2}{2\tau} + \mathcal{E}(\boldsymbol{s}^n) \leq \mathcal{E}(\boldsymbol{s}^{n-1}). \tag{28}$$

As a consequence we have the monotonicity

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

at the discrete level; thus $\boldsymbol{s}^n \in \boldsymbol{\mathcal{X}}$ for all $n \ge 0$ thanks to (14). Summing (28) over n we also obtain the classical *total square distance* estimate

$$\frac{1}{\tau}\sum_{n\ge 0} \boldsymbol{W}^2(\boldsymbol{s}^{n+1}, \boldsymbol{s}^n) \le 2\mathcal{E}(\boldsymbol{s}^0) \le C(\Omega, \Pi, \boldsymbol{\Psi}), \tag{29}$$

where the last inequality comes from the fact that $\boldsymbol{s}^0$ is uniformly bounded since it belongs to $\boldsymbol{\mathcal{X}}$, and thus so is $\mathcal{E}(\boldsymbol{s}^0)$. This readily gives the approximate $\frac{1}{2}$-Hölder estimate

$$\boldsymbol{W}(\boldsymbol{s}^{n_1}, \boldsymbol{s}^{n_2}) \le C\sqrt{|n_2 - n_1|\tau}. \tag{30}$$

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 $\boldsymbol{s}^{n*} = (s_1^n, \dots, s_N^n)$ is the unique element of $\boldsymbol{\mathcal{X}}^*$ corresponding to the minimizer $\boldsymbol{s}^n$ of (26). In what follows, we set

$$\pi_i^n : \Omega \to \mathbb{R}, \quad \boldsymbol{x} \mapsto \pi_i(\boldsymbol{s}^{n*}(\boldsymbol{x}), \boldsymbol{x}), \quad \forall i \in \{1, \dots, N\}$$

and $\boldsymbol{\pi}^n = (\pi_1^n, \dots, \pi_N^n)$. 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 (24).

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

$$\sum_{i=0}^{N} \|\nabla \pi_i^n\|_{L^2(\Omega)}^2 \le C\left(1 + \frac{\boldsymbol{W}^2(\boldsymbol{s}^n, \boldsymbol{s}^{n-1})}{\tau} + \sum_{i=0}^{N} \frac{\mathcal{H}_\omega(s_i^{n-1}) - \mathcal{H}_\omega(s_i^n)}{\tau}\right). \tag{31}$$

Proof. The argument relies on the *flow interchange* technique introduced by Matthes, McCann, and Savaré [Matthes et al. 2009]. Throughout the proof, C denotes a fluctuating constant that depends on the prescribed data Ω, Π, ω, $\mathbb{K}$, $(\mu_i)_i$, and $\boldsymbol{\Psi}$, but neither on t, τ, nor on n. For $i = 0, \dots, N$ consider the auxiliary flows

$$\begin{cases} \partial_t \check{s}_i = \operatorname{div}(\mathbb{K}\nabla \check{s}_i - \check{s}_i \mathbb{K}\nabla \log \omega), & t > 0,\ \boldsymbol{x} \in \Omega, \\ \mathbb{K}(\nabla \check{s}_i - \check{s}_i \nabla \log \omega) \cdot \nu = 0, & t > 0,\ \boldsymbol{x} \in \partial\Omega, \\ \check{s}_i|_{t=0} = s_i^n, & \boldsymbol{x} \in \Omega \end{cases} \tag{32}$$

for each $i \in \{0, \dots, N\}$. By standard parabolic theory, see for instance [Ladyženskaja et al. 1968, 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 $L^1(\Omega)$, and thus in $\mathcal{A}_i$ endowed with the usual quadratic distance W_{ref} (20) thanks to [Santambrogio 2015, Proposition 7.4]. Because of (19), 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 [2009], we know that the evolution $t \mapsto \check{s}_i(\cdot, t)$ can be interpreted as the gradient flow of the relative Boltzmann functional $(1/\mu_i)\mathcal{H}_\omega$ with respect to the metric W_i, the scaling factor $1/\mu_i$ appearing due to the definition (18) of the distance W_i. As a consequence of (23), The Ricci curvature of (Ω, d_i) is bounded, and hence bounded from below. Since $\omega \in \mathcal{C}^2(\overline{\Omega})$, and with our assumption (22), we also have that $(1/\mu_i)\mathcal{H}_\omega$ 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 [Villani 2009, Chapter 14]. Therefore, we can use the so-called *evolution variational inequality* characterization of gradient flows, see for instance [Ambrosio and Gigli 2013, Definition 4.5], centered at s_i^{n-1}, namely

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

Define $\check{\boldsymbol{s}} = (\check{s}_0, \dots, \check{s}_N)$ and $\check{\boldsymbol{s}}^* = (\check{s}_1, \dots, \check{s}_N)$. Summing the previous inequality over $i \in \{0, \dots, N\}$ leads 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 (26), we first note that $\sum s_i^n(\boldsymbol{x}) = \omega(\boldsymbol{x})$ for all $\boldsymbol{x} \in \overline{\Omega}$; thus by the linearity of (32) and since ω is a stationary solution we have $\sum \check{s}_i(\boldsymbol{x}, t) = \omega(\boldsymbol{x})$ as well. Moreover, the problem (32) 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 (26) for all $t > 0$. The smoothness of $\check{\boldsymbol{s}}$ for $t > 0$ allows us 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

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

To estimate I_1, we first use the invertibility 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

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

Combining (3), (7), (8) 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 $t > 0$,

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

Choosing $\delta = \kappa_\star/(4\kappa^\star\varpi^\star)$, we get that

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

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 $\sum_{i=1}^N(\check{s}_i/\omega)^2\le 1$. Therefore, using (36) again, we get

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

Choosing again $\delta = \kappa_\star/(4\kappa^\star\varpi^\star)$ yields

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

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

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

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

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

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 (32), we have set

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

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

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

Using (35) together with (7)–(8), we get that

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

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

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

Finally, it follows from the fact that $\sum_{i=1}^N \check{s}_i \le \omega$, from the Cauchy–Schwarz inequality, and from the regularity of $\boldsymbol{\Psi}$, ω that

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

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

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

Denote by

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

the functional to be minimized in (26); then combining (33) and (44) provides

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

Since $\check{\boldsymbol{s}}(0) = \boldsymbol{s}^n$ is a minimizer of (26), we must have

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

otherwise $\check{\boldsymbol{s}}(t)$ would be a strictly better competitor than $\boldsymbol{s}^n$ for small $t > 0$. As a consequence, we get

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

Since $\check{s}_i$ belongs to $C([0, 1]; L^p(\Omega))$ for all $p \in [1, \infty)$, see for instance [Cancès and Gallouët 2011], the continuity of the Wasserstein distance and of the Boltzmann entropy with respect to strong L^p-convergence imply that

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

Therefore, we obtain that

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

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

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

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

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

3. The Euler–Lagrange equations and pressure bounds

The goal of this section is to extract information coming from the optimality of $\boldsymbol{s}^n$ in the JKO minimization (26). The main difficulty consists in constructing the phase and capillary pressures from this optimality condition. Our proof is inspired by [Maury et al. 2010] and makes extensive use of the Kantorovich potentials. Therefore, we first recall their definition and some useful properties. We refer to [Santambrogio 2015, §1.2; Villani 2009, 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 (16), $i \in \{0, \dots, N\}$, is a solution of the Kantorovich *dual problem*

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

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

(a) There is always duality; that is,

$$\mathrm{DP}_i(\nu_1, \nu_2) = \tfrac{1}{2}W_i^2(\nu_1, \nu_2) \quad \forall i \in \{0, \dots, N\}.$$

(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,

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

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 (17). Thus item (c) shows that φ_i and ψ_i are also Lipschitz continuous.

A decomposition result. The next lemma is an adaptation of [Maury et al. 2010, Lemma 3.1] to our framework. It essentially states that, since $\boldsymbol{s}^n$ is a minimizer of (26), it is also a minimizer of the linearized problem.

Lemma 3.2. *For $n \geq 1$ and $i = 0, \dots, N$ there exist some (backward, optimal) Kantorovich potentials φ_i^n from s_i^n to s_i^{n-1} such that, using the convention $\pi_0^n = (\partial \Pi / \partial s_0)(s_1^n, \dots, s_N^n, \boldsymbol{x}) = 0$, setting*

$$F_i^n := \frac{\varphi_i^n}{\tau} + \pi_i^n + \Psi_i, \quad \forall i \in \{0, \dots, N\}, \tag{47}$$

and defining $\boldsymbol{F}^n = (F_i^n)_{0 \leq i \leq N}$, we have

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

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

Proof. We assume first that $s_i^{n-1}(\boldsymbol{x}) > 0$ everywhere in Ω 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 $\boldsymbol{x}_{\mathrm{ref}} \in \Omega$; see [Santambrogio 2015, Proposition 7.18]. Given any $\boldsymbol{s} = (s_i)_{1 \leq 0 \leq N} \in \boldsymbol{\mathcal{X}} \cap \boldsymbol{\mathcal{A}}$ and $\varepsilon \in (0, 1)$ we define the perturbation

$$\boldsymbol{s}^\varepsilon := (1 - \varepsilon)\boldsymbol{s}^n + \varepsilon \boldsymbol{s}.$$

Note that $\boldsymbol{\mathcal{X}} \cap \boldsymbol{\mathcal{A}}$ is convex; thus $\boldsymbol{s}^\varepsilon$ is an admissible competitor for all $\varepsilon \in (0, 1)$. Let $(\varphi_i^\varepsilon, \psi_i^\varepsilon)$ be the unique Kantorovich potentials from s_i^ε to s_i^{n-1}, similarly normalized as $\varphi_i^\varepsilon(\boldsymbol{x}_{\mathrm{ref}}) = 0$. Then by characterization of the squared Wasserstein distance in terms of the dual Kantorovich problem we have

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

By definition of the perturbation $\boldsymbol{s}^\varepsilon$ it is easy to check that $s_i^\varepsilon - s_i^n = \varepsilon(s_i - s_i^n)$. Subtracting the previous inequalities we get

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

Define $\boldsymbol{s}^{\varepsilon *} = (s_1^\varepsilon, \dots, s_N^\varepsilon)$, $\boldsymbol{\pi}^\varepsilon = \boldsymbol{\pi}(\boldsymbol{s}^{\varepsilon *}, \cdot)$, and extend to the zeroth component $\bar{\boldsymbol{\pi}}^\varepsilon = (0, \boldsymbol{\pi}^\varepsilon)$. The convexity of Π as a function of $s_1, \dots, s_N$ implies

$$\int_\Omega \big(\Pi(\boldsymbol{s}^{n*}, \boldsymbol{x}) - \Pi(\boldsymbol{s}^{\varepsilon *}, \boldsymbol{x})\big) \, \mathrm{d}\boldsymbol{x} \geq \int_\Omega \boldsymbol{\pi}^\varepsilon \cdot (\boldsymbol{s}^{n*} - \boldsymbol{s}^{\varepsilon *}) \, \mathrm{d}\boldsymbol{x} = \int_\Omega \bar{\boldsymbol{\pi}}^\varepsilon \cdot (\boldsymbol{s}^n - \boldsymbol{s}^\varepsilon) \, \mathrm{d}\boldsymbol{x} = -\varepsilon \int_\Omega \bar{\boldsymbol{\pi}}^\varepsilon \cdot (\boldsymbol{s} - \boldsymbol{s}^n) \, \mathrm{d}\boldsymbol{x}. \tag{50}$$

For the potential energy, we obtain by linearity that

$$\int_\Omega (\boldsymbol{s}^\varepsilon - \boldsymbol{s}^n) \cdot \boldsymbol{\Psi} \, \mathrm{d}\boldsymbol{x} = \varepsilon \int_\Omega (\boldsymbol{s} - \boldsymbol{s}^n) \cdot \boldsymbol{\Psi} \, \mathrm{d}\boldsymbol{x}. \tag{51}$$

Summing (49)–(51), dividing by ε, and recalling that $\boldsymbol{s}^n$ minimizes the functional $\mathcal{F}_\tau^n$ defined by (45), we obtain

$$0 \leq \frac{\mathcal{F}_\tau^n(\boldsymbol{s}^\varepsilon) - \mathcal{F}_\tau^n(\boldsymbol{s}^n)}{\varepsilon} \leq \sum_{i=0}^N \int_\Omega \left(\frac{\varphi_i^\varepsilon}{\tau} + \bar{\pi}_i^\varepsilon + \Psi_i \right) (s_i - s_i^n) \, \mathrm{d}\boldsymbol{x} \tag{52}$$

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 [Santambrogio 2015, Theorem 1.17]. Since $\boldsymbol{s}^\varepsilon$ converges uniformly towards $\boldsymbol{s}^n$ when ε tends to 0, we infer from Theorem 1.52 of the same paper that φ_i^ε converges uniformly towards φ_i^n as ε tends to 0, where φ_i^n is a Kantorovich potential from s_i^n to s_i^{n-1}. Moreover, since $\boldsymbol{\pi}$ is uniformly continuous in $\boldsymbol{s}$, we also know that $\boldsymbol{\pi}^\varepsilon$ converges uniformly towards $\boldsymbol{\pi}^n$ and thus $\bar{\boldsymbol{\pi}}^\varepsilon \to \bar{\boldsymbol{\pi}}^n = (0, \boldsymbol{\pi}^n)$ as well. Then we can pass to the limit in (52) and infer that

$$0 \le \int_\Omega \boldsymbol{F}^n \cdot (\boldsymbol{s} - \boldsymbol{s}^n)\,\mathrm{d}\boldsymbol{x} \quad \forall \boldsymbol{s} \in \boldsymbol{\mathcal{X}} \cap \boldsymbol{\mathcal{A}} \tag{53}$$

and (48) holds.

If $s_i^{n-1} > 0$ does not hold everywhere, we argue by approximation. Running the flow (32) for a short time $\delta > 0$ 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 $\boldsymbol{s}^{n-1} = (s_0^{n-1}, \dots, s_N^{n-1})$ in $L^1(\Omega)$ as δ tends to 0. By construction $\boldsymbol{s}^{n-1,\delta} \in \boldsymbol{\mathcal{X}} \cap \boldsymbol{\mathcal{A}}$, and it follows from the strong maximum principle that $s_i^{n-1,\delta} > 0$ in $\overline{\Omega}$ for all $\delta > 0$. By Proposition 2.1 there exists a unique minimizer $\boldsymbol{s}^{n,\delta}$ to the functional

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

Since $\boldsymbol{s}^{n-1,\delta} > 0$, 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 us to construct $\boldsymbol{F}^{n,\delta}$ using (47), where φ_i^n and π_i^n have been replaced by $\varphi_i^{n,\delta}$ and $\pi_i^{n,\delta}$. Thanks to the above discussion,

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

We can now let δ tend to 0. Because of the time continuity of the solutions to (32), 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 $\boldsymbol{s}^{n,\delta}$ and Lemma 2.2 (in particular (31) with $s^{n-1,\delta}$, $s^{n,\delta}$, $\boldsymbol{\pi}^{n,\delta}$ instead of s^{n-1}, s^n, $\boldsymbol{\pi}^n$) we see that $\boldsymbol{\pi}^{n,\delta}$ is bounded in $H^1(\Omega)^{N+1}$ uniformly in $\delta > 0$. Using next the Lipschitz continuity (8) of $\boldsymbol{\phi}$, one deduces 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. Using once again the stability of the Kantorovich potentials [Santambrogio 2015, 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 (54) and claim that (53) 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 $\pi_i^n \in H^1$ (see Lemma 2.2) and $\boldsymbol{\Psi}$ is smooth, we have $F_i^n \in H^1$. Since the phases are bounded $0 \le s_i^n(\boldsymbol{x}) \le \omega(\boldsymbol{x})$ 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. □

We can now suitably decompose the vector field $\boldsymbol{F}^n = (F_i^n)_{0 \le i \le N}$ defined by (47).

Corollary 3.3. *Let $\boldsymbol{F}^n = (F_0^n, \dots, F_N^n)$ be as in Lemma 3.2. There exists $\boldsymbol{\alpha}^n \in \mathbb{R}^{N+1}$ such that, setting $\lambda^n(\boldsymbol{x}) := \min_j (F_j^n(\boldsymbol{x}) + \alpha_j^n)$, we have $\lambda^n \in H^1(\Omega)$ and*

$$F_i^n + \alpha_i^n = \lambda^n \qquad \mathrm{d}s_i^n\text{-a.e. in } \Omega,\ \forall i \in \{0, \dots, N\}, \tag{55}$$

$$\nabla F_i^n = \nabla \lambda^n \quad \mathrm{d}s_i^n\text{-a.e. in } \Omega,\ \forall i \in \{0, \dots, N\}. \tag{56}$$

Proof. By Lemma 3.2 we know that $\boldsymbol{s}^n$ minimizes $\boldsymbol{s} \mapsto \int \boldsymbol{F}^n \cdot \boldsymbol{s}$ among all admissible $\boldsymbol{s} \in \boldsymbol{\mathcal{X}} \cap \boldsymbol{\mathcal{A}}$. Applying the multicomponent bathtub principle, Theorem B.1 in Appendix B, we infer that there exists $\boldsymbol{\alpha}^n = (\alpha_0^n, \dots, \alpha_N^n) \in \mathbb{R}^{N+1}$ such that $F_i^n + \alpha_i^n = \lambda^n$ for $\mathrm{d}s_i^n$-a.e. $\boldsymbol{x} \in \Omega$ and $\lambda^n = \min_j (F_j^n + \alpha_j^n)$ as in our statement. Note first that $\lambda^n \in H^1(\Omega)$ as the minimum of finitely many H^1 functions $F_0, \dots, F_N \in H^1(\Omega)$. From the usual Serrin's chain rule we have moreover that

$$\nabla \lambda^n = \nabla \min_j (F_j^n + \alpha_j^n) = \nabla F_i . \chi_{[F_i^n + \alpha_i^n = \lambda^n]},$$

and since $s_i^n = 0$ inside $[F_i^n + \alpha_i^n \neq \lambda^n]$, the proof is complete. □

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 [Villani 2009, Chapter 14] for a more detailed presentation.

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

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

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

$$\varphi(\boldsymbol{x} + h\boldsymbol{v}) = \varphi(\boldsymbol{x}) + h g_{i,\boldsymbol{x}}(\nabla_{g_{i,x}} \varphi(\boldsymbol{x}), \boldsymbol{v}) + o(h) \quad \forall \boldsymbol{v} \in \mathbb{S}^{d-1},\ \forall \boldsymbol{x} \in \Omega.$$

It is easy to check that this leads to the formula

$$\nabla_{g_i} \varphi = \frac{1}{\mu_i} \mathbb{K} \nabla \varphi, \tag{57}$$

where $\nabla \varphi$ stands for the usual (euclidean) gradient. The formula (57) can be extended to Lipschitz continuous functions φ thanks to Rademacher's theorem.

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

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

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

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

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

where $\boldsymbol{t}_i^n$ denotes the optimal map sending s_i^n to s_i^{n-1}. It follows from [Villani 2009, Theorem 10.41] that

$$\boldsymbol{t}_i^n(\boldsymbol{x}) = \exp_{i,\boldsymbol{x}}(-\nabla_{g_i}\varphi_i^n(\boldsymbol{x})) \quad \forall \boldsymbol{x} \in \Omega. \tag{58}$$

Moreover, using the definition of the exponential and the relation (57), one gets that

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

This yields the formula

$$W_i^2(s_i^n, s_i^{n-1}) = \int_\Omega \frac{s_i^n}{\mu_i}\mathbb{K}\nabla\varphi_i^n\cdot\nabla\varphi_i^n \,\mathrm{d}\boldsymbol{x} \quad \forall i \in \{0, \dots, N\}. \tag{59}$$

We have now introduced the necessary material in order to reconstruct the phase and capillary pressures. This is the purpose of the following Proposition 3.4 and then of Corollary 3.5.

Proposition 3.4. *For $n \geq 1$ let $\varphi_i^n : s_i^n \to s_i^{n-1}$ be the (backward) Kantorovich potentials from Lemma 3.2. There exists $\boldsymbol{h} = (h_0^n, \dots, h_N^n) \in H^1(\Omega)^{N+1}$ such that*

(i) $\nabla h_i^n = -\nabla\varphi_i^n/\tau$ *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.* $\boldsymbol{x} \in \Omega$, $i \in \{1, \dots, N\}$,

(iii) *there exists C depending only on Ω, Π, ω, $\mathbb{K}$, $(\mu_i)_i$, 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 $F_i^n \in L^\infty \cap H^1(\Omega)$ as in (47), as well as $\boldsymbol{\alpha}^n \in \mathbb{R}^{N+1}$ and $\lambda^n = \min_j(F_j^n + \alpha_j^n) \in L^\infty \cap H^1(\Omega)$ as in Corollary 3.3. Setting

$$h_i^n := -\frac{\varphi_i^n}{\tau} + F_i^n - \lambda^n \quad \forall i \in \{0, \dots, N\},$$

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 F_i^n, λ^n. Recalling that we use the notation $\pi_0 = \partial\Pi/\partial s_0 = 0$, we see from the definition (47) of F_i^n that

$$h_i^n - h_0^n = \left(F_i^n - \frac{\varphi_i^n}{\tau}\right) - \left(F_0^n - \frac{\varphi_0^n}{\tau}\right) = (\pi_i^n + \Psi_i) - (\pi_0^n + \Psi_0) = \pi_i^n + \Psi_i - \Psi_0 \tag{60}$$

for all $i \in \{1, \dots, N\}$ and $\mathrm{d}\boldsymbol{x}$-a.e. x, which is exactly our statement (ii).

For (i), we simply use (56) to compute

$$\nabla h_i^n = -\frac{\nabla\varphi_i^n}{\tau} + \nabla(F_i^n - \lambda_i^n) = -\frac{\nabla\varphi_i^n}{\tau} \quad \text{for } \mathrm{d}s_i^n\text{-a.e. } \boldsymbol{x} \in \Omega, \ \forall i \in \{0, \dots, N\}. \tag{61}$$

In order to establish now the H^1 estimate (iii), let us define

$$\mathcal{U}_i = \{\boldsymbol{x} \in \Omega \mid s_i^n(\boldsymbol{x}) \geq \omega_\star/(N+1)\}.$$

Then since $\sum s_i^n(\boldsymbol{x}) = \omega(\boldsymbol{x}) \ge \omega_\star > 0$, one gets that, up to a negligible set,

$$\bigcup_{i=0}^N \mathcal{U}_i = \Omega, \quad \text{hence} \quad (\mathcal{U}_i)^c \subset \bigcup_{j\neq i} \mathcal{U}_j. \tag{62}$$

We first estimate ∇h_0^n. To this end, we write

$$\|\nabla h_0^n\|_{L^2}^2 \le \frac{1}{\kappa_\star}\int_\Omega \mathbb{K}\nabla h_0^n \cdot \nabla h_0^n \,\mathrm{d}\boldsymbol{x} \le A + B, \tag{63}$$

where we have set

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

Owing to (61) one has $\nabla h_0^n = -\nabla\varphi_0/\tau$ on $\mathcal{U}_0 \subset \Omega$, where $s_0^n \ge \omega_\star/(N+1)$. Therefore,

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

Then it results from formula (59) that

$$A \le \frac{C}{\tau^2} W_0^2(s_0^n, s_0^{n-1}), \tag{64}$$

where C depends neither on n nor on τ. Combining (62) and (60), we infer

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

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

$$B \le \frac{3}{\kappa_\star}\sum_{i=1}^N \int_{\mathcal{U}_i} \mathbb{K}\nabla h_i \cdot \nabla h_i \,\mathrm{d}\boldsymbol{x} + \frac{3\kappa^\star}{\kappa_\star}\sum_{i=1}^N \big(\|\nabla\pi_i^n\|_{L^2}^2 + \|\nabla(\Psi_i - \Psi_0)\|_{L^2}^2\big). \tag{65}$$

Similar calculations to those carried out to estimate A yield

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

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 (65) that

$$B \le C\left(1 + \frac{\boldsymbol{W}^2(\boldsymbol{s}^n, \boldsymbol{s}^{n-1})}{\tau^2} + \sum_{i=0}^N \frac{\mathcal{H}_\omega(s_i^{n-1}) - \mathcal{H}_\omega(s_i^n)}{\tau}\right) \tag{66}$$

for some C not depending on n and τ (here we also used $1/\tau \le 1/\tau^2$ for small τ in the W^2 terms). Gathering (64) and (66) in (63) provides

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

Note that (i) and (ii) remain invariant under subtraction of the same constant, that is, $h_0^n, h_i^n \rightsquigarrow h_0^n - C, h_i^n - C$, 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 the Poincaré–Wirtinger inequality, we get that

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

Finally, from (ii) $h_i^n = h_0^n + \pi_i^n + \Psi_i - \Psi_0$, the smoothness of $\boldsymbol{\Psi}$, and using again the estimate (31) for $\|\nabla \boldsymbol{\pi}^n\|_{L^2}^2$ we finally get that for all $i \in \{1, \dots, N\}$, one has

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

and the proof of Proposition 3.4 is complete. □

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

$$p_i^n := h_i^n - \Psi_i \quad \forall i \in \{0, \dots, N\}. \tag{67}$$

The following corollary is a straightforward consequence of Proposition 3.4 and of the regularity of Ψ_i.

Corollary 3.5. *The phase pressures $\boldsymbol{p}^n = (p_i^n)_{0 \le i \le N} \in H^1(\Omega)^{N+1}$ satisfy*

$$\|\boldsymbol{p}^n\|_{H^1(\Omega)}^2 \le C\left(1 + \frac{\boldsymbol{W}^2(\boldsymbol{s}^n, \boldsymbol{s}^{n-1})}{\tau^2} + \sum_{i=0}^N \frac{\mathcal{H}_\omega(s_i^{n-1}) - \mathcal{H}_\omega(s_i^n)}{\tau}\right) \tag{68}$$

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

$$p_i^n - p_0^n = \pi_i^n \quad \forall i \in \{1, \dots, N\}. \tag{69}$$

Our next result is a first step towards the recovery of the PDEs.

Lemma 3.6. *There exists C depending only on Ω, Π, ω, $\mathbb{K}$, $(\mu_i)_i$, 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*

$$\left|\int_\Omega (s_i^n - s_i^{n-1})\xi \, \mathrm{d}\boldsymbol{x} + \tau \int_\Omega s_i^n \frac{\mathbb{K}}{\mu_i} \nabla(p_i^n + \Psi_i) \cdot \nabla \xi \, \mathrm{d}\boldsymbol{x}\right| \le C W_i^2(s_i^n, s_i^{n-1}) \|D_{g_i}^2 \xi\|_\infty. \tag{70}$$

This is of course a discrete approximation to the continuity equation $\partial_t s_i = \nabla \cdot (s_i(\mathbb{K}/\mu_i)\nabla(p_i + \Psi_i))$.

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 $\boldsymbol{t}_i^n$ be the corresponding optimal map as in (58). For fixed $\xi \in \mathcal{C}^2(\overline{\Omega})$ let us first Taylor expand (in the g_i Riemannian framework)

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

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

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

From Proposition 3.4(i) we have $\nabla\varphi_i^n = -\tau\nabla h_i^n$ for $\mathrm{d}s_i^n$-a.e. $\boldsymbol{x}\in\Omega$; thus by the definition (67) of p_i^n, we get $\nabla\varphi^n = -\tau\nabla(p_i^n + \Psi_i)$. Substituting in the second integral of the left-hand side gives exactly (70). □

4. Convergence towards a weak solution

The goal is now to prove the convergence of the piecewise constant interpolated solutions $\boldsymbol{s}^\tau$, defined by (27), towards a weak solution $\boldsymbol{s}$ as $\tau \to 0$. 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 notation $\boldsymbol{s}^{\tau *} = (s_1^\tau, \dots, s_N^\tau) \in L^\infty((0,T); \boldsymbol{\mathcal{X}}^*)$ and $\boldsymbol{\pi}^\tau = \boldsymbol{\pi}(\boldsymbol{s}^{\tau *}, \boldsymbol{x})$.

Time integrated estimates. We immediately deduce from (30) that

$$\boldsymbol{W}(\boldsymbol{s}^\tau(t_2), \boldsymbol{s}^\tau(t_1)) \leq C|t_2 - t_1 + \tau|^{1/2} \quad \forall\, 0 \leq t_1 \leq t_2 \leq T. \tag{71}$$

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

$$0 \leq s_i^\tau(\boldsymbol{x}, t) \leq \omega^\star \quad \text{a.e. in } Q \text{ for all } i \in \{0, \dots, N\}. \tag{72}$$

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

$$\|\boldsymbol{p}^\tau\|^2_{L^2((0,T);H^1(\Omega)^{N+1})} + \|\boldsymbol{\pi}^\tau\|^2_{L^2((0,T);H^1(\Omega)^N)} \leq C. \tag{73}$$

Proof. Summing (68) from $n = 1$ to $n = N_\tau := \lceil T/\tau \rceil$, we get

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

We use that

$$0 \geq \mathcal{H}_\omega(s) \geq -\frac{1}{e}\|\omega\|_{L^1} \geq -\frac{|\Omega|}{e} \quad \forall s \in L^\infty(\Omega) \text{ with } 0 \leq s \leq \omega$$

together with the total square distance estimate (29) to infer that $\|\boldsymbol{p}\|^2_{L^2(H^1)} \leq C$. The proof is identical for the capillary pressure $\boldsymbol{\pi}^\tau$ (simply summing the one-step estimate from Lemma 2.2). □

Compactness of approximate solutions. We define $H' = H^1(\Omega)'$.

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*

$$\|s_i^\tau(t_2) - s_i^\tau(t_1)\|_{H'} \le C|t_2 - t_1 + \tau|^{1/2} \quad \forall\, 0 \le t_1 \le t_2 \le T.$$

Proof. Thanks to (72), we can apply [Maury et al. 2010, Lemma 3.4] to get

$$\left| \int_\Omega f\{s_i^\tau(t_2) - s_i^\tau(t_1)\}\, \mathrm{d}\boldsymbol{x} \right| \le \|\nabla f\|_{L^2(\Omega)} W_{\mathrm{ref}}(s_i^\tau(t_1), s_i^\tau(t_2)) \quad \forall f \in H^1(\Omega).$$

Thus by duality and thanks to the distance estimate (71) and to the lower bound in (19), we obtain that

$$\|s_i^\tau(t_2) - s_i^\tau(t_1)\|_{H'} \le W_{\mathrm{ref}}(s_i^\tau(t_1), s_i^\tau(t_2)) \le C\, W_i(s_i^\tau(t_1), s_i^\tau(t_2)) \le C|t_2 - t_1 + \tau|^{1/2}$$

for some C depending only on Ω, Π, $(\rho_i)_i$, $\boldsymbol{g}$, $(\mu_i)_i$, $\mathbb{K}$. □

From the previous equicontinuity 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 and Luckhaus' trick [1983] (an alternate solution would consist in slightly adapting the nonlinear time compactness results [Moussa 2016; Andreianov et al. 2015] to our context). Let $h > 0$ be a small time shift; then by monotonicity and Lipschitz continuity of the capillary pressure function $\boldsymbol{\pi}(\,\cdot\,, \boldsymbol{x})$,

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

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 with respect to τ) $L^2((0, T); H^1(\Omega)^N)$- and $L^\infty(Q)^N$-estimates on $\boldsymbol{\pi}^\tau$ ensure that

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

where $\boldsymbol{\pi}^\tau$ is extended by 0 outside Ω. This allows us to apply Kolmogorov's compactness theorem, see, for instance, [Hanche-Olsen and Holden 2010], and gives the desired relative compactness. □

Identification of the limit. In this section we prove our main result, 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. Throughout this section, $(\tau_k)_{k\ge1}$ denotes a sequence of times steps tending to 0 as $k \to \infty$.

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

$$\boldsymbol{s}^{\tau_k} \xrightarrow{k\to\infty} \boldsymbol{s} \qquad \text{a.e. in } Q, \tag{74}$$

$$\boldsymbol{\pi}^{\tau_k} \xrightharpoonup{k\to\infty} \boldsymbol{\pi}(\boldsymbol{s}^*,\cdot) \qquad \text{weakly in } L^2((0,T); H^1(\Omega)^N), \tag{75}$$

$$\boldsymbol{p}^{\tau_k} \xrightharpoonup{k\to\infty} \boldsymbol{p} \qquad \text{weakly in } L^2((0,T); H^1(\Omega)^{N+1}). \tag{76}$$

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

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

In particular, this yields $\boldsymbol{\pi}^{\tau_k} \xrightarrow{k\to\infty} \boldsymbol{\pi}(\boldsymbol{s}^*,\cdot)$ a.e. in Q. Since we have 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, (74) 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 (75) and (76) are straightforward consequences of Lemma 4.1. Lastly, it follows from (69) that

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

We can finally pass to the limit $k\to\infty$ in the above relation thanks to (75)–(76) and infer

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

which immediately implies (5) as claimed. □

Lemma 4.5. *Up to the extraction of an additional subsequence, the limit $\boldsymbol{s}$ of $(\boldsymbol{s}^{\tau_k})_{k\ge1}$ belongs to $C([0,T]; \boldsymbol{\mathcal{A}})$, where $\boldsymbol{\mathcal{A}}$ is equipped with the metric $\boldsymbol{W}$. Moreover, $\boldsymbol{W}(\boldsymbol{s}^{\tau_k}(t),\boldsymbol{s}(t)) \xrightarrow{k\to\infty} 0$ for all $t \in [0,T]$.*

Proof. It follows from the bounds (72) on s_i 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 [Santambrogio 2015, Theorem 5.10] together with the equivalence of W_i with W_{ref} stated in (19). Thanks to (71), one has

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

Applying a refined version of the Arzelà–Ascoli theorem [Ambrosio et al. 2008, Proposition 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 (12):

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 $-\rho_i \boldsymbol{g}$ replaced by $+\nabla\Psi_i$ in the general case).*

Proof. Let $0 \leq t_1 \leq t_2 \leq T$, and define $n_{j,k} = \lceil t_j/\tau_k \rceil$ and $\tilde{t}_j = n_{j,k}\tau_k$ for $j \in \{1, 2\}$. Fixing an arbitrary $\xi \in \mathcal{C}^2(\overline{\Omega})$ and summing (70) from $n = n_{1,k} + 1$ to $n = n_{2,k}$ yields

$$\begin{aligned}\int_\Omega (s_i^{\tau_k}(t_2) - s_i^{\tau_k}(t_1))\xi \,\mathrm{d}\boldsymbol{x} &= \sum_{n=n_{1,k}+1}^{n_{2,k}} \int_\Omega (s_i^n - s_i^{n-1})\xi \,\mathrm{d}\boldsymbol{x} \\ &= -\int_{\tilde{t}_1}^{\tilde{t}_2}\int_\Omega \frac{s_i^{\tau_k}}{\mu_i}\mathbb{K}\nabla(p_i^{\tau_k} + \Psi_i)\cdot\nabla\xi \,\mathrm{d}\boldsymbol{x}\,\mathrm{d}t + \mathcal{O}\Bigg(\sum_{n=n_{1,k}+1}^{n_{2,k}} W_i^2(s_i^n, s_i^{n-1})\Bigg). \end{aligned} \tag{77}$$

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

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

Combining the above estimate with the total square distance estimate (29) in (77), we obtain

$$\int_\Omega (s_i^{\tau_k}(t_2) - s_i^{\tau_k}(t_1))\xi \,\mathrm{d}\boldsymbol{x} + \int_{t_1}^{t_2}\int_\Omega \frac{s_i^{\tau_k}}{\mu_i}\mathbb{K}\nabla(p_i^{\tau_k} + \Psi_i)\cdot\nabla\xi \,\mathrm{d}\boldsymbol{x}\,\mathrm{d}t = \mathcal{O}(\sqrt{\tau_k}). \tag{78}$$

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 [Santambrogio 2015, Theorem 5.10]), we get

$$\int_\Omega (s_i^{\tau_k}(t_2) - s_i^{\tau_k}(t_1))\xi \,\mathrm{d}\boldsymbol{x} \xrightarrow{k\to\infty} \int_\Omega (s_i(t_2) - s_i(t_1))\xi \,\mathrm{d}\boldsymbol{x}. \tag{79}$$

Moreover, thanks to Lemma 4.4, one has

$$\int_{t_1}^{t_2}\int_\Omega \frac{s_i^{\tau_k}}{\mu_i}\mathbb{K}\nabla(p_i^{\tau_k} + \Psi_i)\cdot\nabla\xi \,\mathrm{d}\boldsymbol{x}\,\mathrm{d}t \xrightarrow{k\to\infty} \int_{t_1}^{t_2}\int_\Omega \frac{s_i}{\mu_i}\mathbb{K}\nabla(p_i + \Psi_i)\cdot\nabla\xi \,\mathrm{d}\boldsymbol{x}\,\mathrm{d}t. \tag{80}$$

Combining (78)–(80) yields, for all $\xi \in \mathcal{C}^2(\overline{\Omega})$ and all $0 \leq t_1 \leq t_2 \leq T$,

$$\int_\Omega (s_i(t_2) - s_i(t_1))\xi \,\mathrm{d}\boldsymbol{x} + \int_{t_1}^{t_2}\int_\Omega \frac{s_i}{\mu_i}\mathbb{K}\nabla(p_i + \Psi_i)\cdot\nabla\xi \,\mathrm{d}\boldsymbol{x}\,\mathrm{d}t = 0. \tag{81}$$

In order to conclude the proof, it remains to check that the formulation (81) is stronger the formulation (12). Let $\varepsilon > 0$ be a time step, unrelated to that appearing in the minimization scheme (26), and set $L_\varepsilon = \lfloor T/\varepsilon \rfloor$. Let $\phi \in \mathcal{C}_c^\infty(\overline{\Omega} \times [0, T))$, and set $\phi_\ell = \phi(\,\cdot\,, \ell\varepsilon)$ for $\ell \in \{0, \ldots, L_\varepsilon\}$. Since $t \mapsto \phi(\,\cdot\,, t)$ is compactly supported in $[0, T)$, there exists $\varepsilon^\star > 0$ such that $\phi_{L_\varepsilon} \equiv 0$ for all $\varepsilon \in (0, \varepsilon^\star]$. Then define

$$\phi^\varepsilon : \overline{\Omega} \times [0, T] \to \mathbb{R}, \qquad (\boldsymbol{x}, t) \mapsto \phi_\ell(\boldsymbol{x}) \quad \text{if } t \in [\ell\varepsilon, (\ell + 1)\varepsilon).$$

Choose $t_1 = \ell\varepsilon,\ t_2 = (\ell + 1)\varepsilon,\ \xi = \phi_\ell$ in (81) and sum over $\ell \in \{0, \ldots, L_\varepsilon - 1\}$. This provides

$$A(\varepsilon) + B(\varepsilon) = 0 \quad \forall \varepsilon > 0, \tag{82}$$

where

$$A(\varepsilon) = \sum_{\ell=0}^{L_\varepsilon - 1} \int_\Omega \big(s_i((\ell+1)\varepsilon) - s_i(\ell\varepsilon)\big)\phi^\ell \,\mathrm{d}\boldsymbol{x}, \quad B(\varepsilon) = \iint_Q \frac{s_i}{\mu_i} \mathbb{K}\nabla(p_i + \Psi_i)\cdot\nabla\phi^\varepsilon \,\mathrm{d}\boldsymbol{x}\,\mathrm{d}t.$$

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

$$B(\varepsilon) \xrightarrow{\varepsilon\to 0} \iint_Q \frac{s_i}{\mu_i}\mathbb{K}\nabla(p_i+\Psi_i)\cdot\nabla\phi \,\mathrm{d}\boldsymbol{x}\,\mathrm{d}t. \tag{83}$$

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

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

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

$$A(\varepsilon) \xrightarrow{\varepsilon\to 0} -\iint_Q s_i\,\partial_t\phi\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}t - \int_\Omega s_i^0\phi(\,\cdot\,,0)\,\mathrm{d}\boldsymbol{x}. \tag{84}$$

Combining (82)–(84) shows that the weak formulation (12) is fulfilled. □

Appendix A: A simple condition for the geodesic convexity of (Ω, d_i)

The goal of this appendix is to provide a simple condition on the permeability tensor in order to ensure that condition (22) is fulfilled. For the sake of simplicity, we only consider here the case of isotropic permeability tensors

$$\mathbb{K}(\boldsymbol{x}) = \kappa(\boldsymbol{x})\mathbb{I}_d \quad \forall \boldsymbol{x} \in \overline{\Omega} \tag{85}$$

with $\kappa_\star \le \kappa(\boldsymbol{x}) \le \kappa^\star$ for all $\boldsymbol{x} \in \overline{\Omega}$. Let us stress that the condition we provide is not optimal.

As in the core of the paper, Ω denotes a convex open subset of $\mathbb{R}^d$ with C^2 boundary $\partial\Omega$. For $\bar{\boldsymbol{x}} \in \partial\Omega$, we denote by $\boldsymbol{n}(\bar{\boldsymbol{x}})$ the outward-pointing normal. Since $\partial\Omega$ is smooth, there exists $\ell_0 > 0$ such that, for all $\boldsymbol{x} \in \Omega$ such that $\mathrm{dist}(\boldsymbol{x}, \partial\Omega) < \ell_0$, there exists a unique $\bar{\boldsymbol{x}} \in \partial\Omega$ such that $\mathrm{dist}(\boldsymbol{x},\partial\Omega) = |\boldsymbol{x} - \bar{\boldsymbol{x}}|$ (here dist denotes the usual euclidean distance between sets in $\mathbb{R}^d$). As a consequence, one can rewrite $\boldsymbol{x} = \bar{\boldsymbol{x}} - \ell\boldsymbol{n}(\bar{\boldsymbol{x}})$ for some $\ell \in (0, \ell_0)$.

In what follows, a function $f : \overline{\Omega} \to \mathbb{R}$ is said to be normally nondecreasing (resp. nonincreasing) on a neighborhood of $\partial\Omega$ if there exists $\ell_1 \in (0, \ell_0]$ such that $\ell \mapsto f(\bar{\boldsymbol{x}} - \ell\boldsymbol{n}(\bar{\boldsymbol{x}}))$ is nonincreasing (resp. nondecreasing) on $[0, \ell_1]$.

Proposition A.1. *Assume that*

(i) *the permeability field $\boldsymbol{x} \mapsto \kappa(\boldsymbol{x})$ is normally nonincreasing in a neighborhood of $\partial\Omega$;*

(ii) *for all $\bar{\boldsymbol{x}} \in \partial\Omega$, either $\nabla\kappa(\bar{\boldsymbol{x}})\cdot\boldsymbol{n}(\bar{\boldsymbol{x}}) < 0$, or $\nabla\kappa(\bar{\boldsymbol{x}})\cdot\boldsymbol{n}(\bar{\boldsymbol{x}}) = 0$ and $D^2\kappa(\bar{\boldsymbol{x}})\boldsymbol{n}(\boldsymbol{x})\cdot\boldsymbol{n}(\boldsymbol{x}) = 0$.*

Then there exists a C^2 extension $\tilde\kappa : \mathbb{R}^d \to \left[\frac{1}{2}\kappa_\star, \kappa^\star\right]$ of κ and a Riemannian metric

$$\tilde\delta(\boldsymbol{x}, \boldsymbol{y}) = \inf_{\boldsymbol{\gamma}\in\widetilde{P}(\boldsymbol{x},\boldsymbol{y})} \left(\int_0^1 \frac{1}{\tilde\kappa(\boldsymbol{\gamma}(\tau))} |\boldsymbol{\gamma}'(\tau)|^2 \,\mathrm{d}\tau\right)^{1/2} \quad \forall \boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^d \tag{86}$$

with $\widetilde{P}(\boldsymbol{x}, \boldsymbol{y}) = \{\boldsymbol{\gamma} \in C^1([0, 1]; \mathbb{R}^d) \mid \boldsymbol{\gamma}(0)=\boldsymbol{x}$ and $\boldsymbol{\gamma}(1)=\boldsymbol{y}\}$, such that $(\Omega, \tilde\delta)$ is geodesically convex.

Proof. Since Ω is convex, for all $\boldsymbol{x} \in \mathbb{R}^d \setminus \Omega$ there exists a unique $\bar{\boldsymbol{x}} \in \partial\Omega$ such that $\mathrm{dist}(\boldsymbol{x}, \Omega) = |\boldsymbol{x} - \bar{\boldsymbol{x}}|$. Then one can extend κ in a C^2 way into the whole $\mathbb{R}^d$ by defining

$$\kappa(\boldsymbol{x}) = \kappa(\bar{\boldsymbol{x}}) + |\boldsymbol{x} - \bar{\boldsymbol{x}}|\nabla\kappa(\bar{\boldsymbol{x}}) \cdot \boldsymbol{n}(\bar{\boldsymbol{x}}) + \tfrac{1}{2}|\boldsymbol{x} - \bar{\boldsymbol{x}}|^2 D^2\kappa(\bar{\boldsymbol{x}})\boldsymbol{n}(\bar{\boldsymbol{x}}) \cdot \boldsymbol{n}(\bar{\boldsymbol{x}}), \quad \forall \boldsymbol{x} \in \mathbb{R}^d \setminus \Omega.$$

Thanks to assumptions (i) and (ii), the function $\ell \mapsto \kappa(\bar{\boldsymbol{x}} - \ell\boldsymbol{n}(\bar{\boldsymbol{x}}))$ is nondecreasing on $(-\infty, \ell_1]$ for all $\bar{\boldsymbol{x}} \in \partial\Omega$. Since $\partial\Omega$ is compact, there exists $\ell_2 > 0$ such that

$$\kappa(\bar{\boldsymbol{x}} - \ell\boldsymbol{n}(\bar{\boldsymbol{x}})) \geq \tfrac{1}{2}\kappa_\star \quad \forall \ell \in (-\ell_2, 0].$$

Let $\rho : \mathbb{R}_+ \to \mathbb{R}$ be a nondecreasing C^2 function such that $\rho(0) = 1$, $\rho'(0) = \rho''(0) = 0$ and $\rho(\ell) = 0$ for all $\ell \geq \ell_2$. Then define

$$\tilde\kappa(\boldsymbol{x}) = \rho(\mathrm{dist}(\boldsymbol{x}, \Omega))\kappa(\boldsymbol{x}) + (1 - \rho(\mathrm{dist}(\boldsymbol{x}, \Omega)))\tfrac{1}{2}\kappa_\star \quad \forall \boldsymbol{x} \in \mathbb{R}^d,$$

so that the function $\ell \mapsto \tilde\kappa(\bar{\boldsymbol{x}} - \ell\boldsymbol{n}(\bar{\boldsymbol{x}}))$ is nonincreasing on $(-\infty, \ell_1)$ and bounded from below by $\frac{1}{2}\kappa_\star$.

Let $\boldsymbol{x}, \boldsymbol{y} \in \Omega$; then there exists $\varepsilon > 0$ such that $\mathrm{dist}(\boldsymbol{x}, \partial\Omega) \geq \varepsilon$, $\mathrm{dist}(\boldsymbol{y}, \partial\Omega) \geq \varepsilon$, and κ is normally nonincreasing on $\partial\Omega_\varepsilon := \{\boldsymbol{x} \in \overline{\Omega} \mid \mathrm{dist}(\boldsymbol{x}, \partial\Omega) < \varepsilon\}$. A sufficient condition for $(\Omega, \tilde\delta)$ to be geodesic is that the geodesic $\boldsymbol{\gamma}^{\mathrm{opt}}_{\boldsymbol{x},\boldsymbol{y}}$ from $\boldsymbol{x}$ to $\boldsymbol{y}$ is such that

$$\mathrm{dist}(\boldsymbol{\gamma}^{\mathrm{opt}}_{\boldsymbol{x},\boldsymbol{y}}(t), \partial\Omega) \geq \varepsilon, \quad \forall t \in [0, 1]. \tag{87}$$

In order to ease the reading, we denote by $\boldsymbol{\gamma} = \boldsymbol{\gamma}^{\mathrm{opt}}_{\boldsymbol{x},\boldsymbol{y}}$ any geodesic such that

$$\tilde\delta^2(\boldsymbol{x}, \boldsymbol{y}) = \int_0^1 \frac{1}{\tilde\kappa(\boldsymbol{\gamma}(\tau))} |\boldsymbol{\gamma}'(\tau)|^2 \,\mathrm{d}\tau. \tag{88}$$

We define the continuous and piecewise C^1 path $\boldsymbol{\gamma}_\varepsilon$ from $\boldsymbol{x}$ to $\boldsymbol{y}$ by setting

$$\boldsymbol{\gamma}_\varepsilon(t) = \mathrm{proj}_{\overline{\Omega}_\varepsilon}(\boldsymbol{\gamma}(t)) \quad \forall t \in [0, 1], \tag{89}$$

where $\overline{\Omega}_\varepsilon := \{\boldsymbol{x} \in \Omega \mid \mathrm{dist}(\boldsymbol{x}, \partial\Omega) \geq \varepsilon\}$ is convex, and the orthogonal (with respect to the euclidean distance dist) projection $\mathrm{proj}_{\overline{\Omega}_\varepsilon}$ onto $\overline{\Omega}_\varepsilon$ is therefore uniquely defined.

Assume that condition (87) is violated. Then by continuity there exists a nonempty interval $[a, b] \subset [0, 1]$ such that

$$\mathrm{dist}(\boldsymbol{\gamma}(t), \partial\Omega) < \varepsilon \quad \forall t \in (a, b);$$

that is, the geodesic between $\gamma(a)$ and $\gamma(b)$ coincides with the part of the geodesic between $\boldsymbol{x}$ and $\boldsymbol{y}$. Then, changing $\boldsymbol{x}$ into $\boldsymbol{\gamma}(a)$ and $\boldsymbol{y}$ into $\boldsymbol{\gamma}(b)$, we can assume without loss of generality that

$$\mathrm{dist}(\boldsymbol{\gamma}(t), \partial\Omega) < \varepsilon \quad \forall t \in (0, 1).$$

It is easy to verify that

$$|\boldsymbol{\gamma}_\varepsilon'(t)| \le |\boldsymbol{\gamma}'(t)| \quad \forall t \in [0,1] \qquad \text{and} \quad |\boldsymbol{\gamma}_\varepsilon'(t)| < |\boldsymbol{\gamma}'(t)| \quad \text{on } (a,b) \tag{90}$$

for some nonempty interval $(a,b) \subset [0,1]$. It follows from (86) that

$$\tilde{\delta}^2(\boldsymbol{x},\boldsymbol{y}) \le \int_0^1 \frac{1}{\tilde{\kappa}(\boldsymbol{\gamma}_\varepsilon(\tau))} |\boldsymbol{\gamma}_\varepsilon'(\tau)|^2 \, \mathrm{d}\tau.$$

Since κ is normally nonincreasing, one has

$$\tilde{\delta}^2(\boldsymbol{x},\boldsymbol{y}) \le \int_0^1 \frac{1}{\tilde{\kappa}(\boldsymbol{\gamma}(\tau))} |\boldsymbol{\gamma}_\varepsilon'(\tau)|^2 \, \mathrm{d}\tau.$$

Thanks to (90), one obtains that

$$\tilde{\delta}^2(\boldsymbol{x},\boldsymbol{y}) < \int_0^1 \frac{1}{\tilde{\kappa}(\boldsymbol{\gamma}(\tau))} |\boldsymbol{\gamma}'(\tau)|^2 \, \mathrm{d}\tau,$$

providing a contradiction with the optimality (88) of $\boldsymbol{\gamma}$. Thus condition (87) holds; hence (Ω,δ) is a geodesic space. □

Appendix B: A multicomponent bathtub principle

The following theorem can be seen as a generalization of the classical scalar bathtub principle; see, for instance, [Lieb and Loss 2001, Theorem 1.14]. In what follows, N is a positive integer and Ω denotes an arbitrary measurable subset of $\mathbb{R}^d$.

Theorem B.1. *Let $\omega \in L^1_+(\Omega)$, and let $\boldsymbol{m} = (m_0,\dots,m_N) \in (\mathbb{R}^*_+)^{N+1}$ be such that $\sum_{i=0}^N m_i = \int_\Omega \omega \, \mathrm{d}\boldsymbol{x}$. We define*

$$\boldsymbol{\mathcal{X}} \cap \boldsymbol{\mathcal{A}} = \left\{ \boldsymbol{s} = (s_0,\dots,s_N) \in L^1_+(\Omega)^{N+1} \;\middle|\; \int_\Omega s_i \, \mathrm{d}\boldsymbol{x} = m_i \text{ and } \sum_{i=0}^N s_i = \omega \text{ a.e. in } \Omega \right\}.$$

Then for any $\boldsymbol{F} = (F_0,\dots,F_N) \in (L^\infty(\Omega))^{N+1}$, the functional

$$\mathcal{F} : \boldsymbol{s} \mapsto \int_\Omega \boldsymbol{F} \cdot \boldsymbol{s} \, \mathrm{d}\boldsymbol{x}$$

has a minimizer in $\boldsymbol{\mathcal{X}} \cap \boldsymbol{\mathcal{A}}$. Moreover, there exists $\boldsymbol{\alpha} = (\alpha_0,\dots,\alpha_N) \in \mathbb{R}^{N+1}$ such that, defining

$$\lambda(\boldsymbol{x}) := \min_{0 \le j \le N} \{F_j(\boldsymbol{x}) + \alpha_j\}, \qquad \boldsymbol{x} \in \Omega,$$

any minimizer $\underline{\boldsymbol{s}} = (\underline{s}_0,\dots,\underline{s}_N)$ satisfies

$$F_i + \alpha_i = \lambda \quad \mathrm{d}\underline{s}_i\text{-a.e. in } \Omega, \ \forall i \in \{0,\dots,N\}.$$

One can think of this as: $\underline{s}_i = 0$ in $\{F_i + \alpha_i > \lambda\}$ and $F_i + \alpha_i \ge \lambda$ everywhere; i.e., $\underline{s}_i > 0$ can only occur in the "contact set" $\{\boldsymbol{x} \mid F_i(\boldsymbol{x}) + \alpha_i = \min_j (F_j(\boldsymbol{x}) + \alpha_j)\}$.

Proof. For the existence part, note that $\mathcal{F}$ is continuous for the weak L^1 convergence, and that $\boldsymbol{\mathcal{X}} \cap \boldsymbol{\mathcal{A}}$ is weakly closed. Since $\sum s_i = \omega$ and $s_i \geq 0$, we have in particular $0 \leq s_i \leq \omega \in L^1$ for all i and $\boldsymbol{s} \in \boldsymbol{\mathcal{X}} \cap \boldsymbol{\mathcal{A}}$. This implies that $\boldsymbol{\mathcal{X}} \cap \boldsymbol{\mathcal{A}}$ is uniformly integrable, and since the mass $\|s_i\|_{L^1} = \int s_i = m_i$ is prescribed, the Dunford–Pettis theorem shows that $\boldsymbol{\mathcal{X}} \cap \boldsymbol{\mathcal{A}}$ is L^1-weakly relatively compact. Hence from any minimizing sequence we can extract a weakly-L^1 converging subsequence, and by weak L^1 continuity the weak limit is a minimizer.

Let us now introduce a dual problem: for fixed $\boldsymbol{\alpha} = (\alpha_0, \dots, \alpha_N) \in \mathbb{R}^{N+1}$ we set

$$\lambda_{\boldsymbol{\alpha}}(\boldsymbol{x}) := \min_i \{F_i(\boldsymbol{x}) + \alpha_i\} \tag{91}$$

and define

$$J(\boldsymbol{\alpha}) := \int_\Omega \lambda_{\boldsymbol{\alpha}}(\boldsymbol{x})\omega(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} - \sum_{i=0}^N \alpha_i m_i.$$

We shall prove below that

(i) $\sup_{\boldsymbol{\alpha} \in \mathbb{R}^{N+1}} J(\boldsymbol{\alpha}) = \max_{\boldsymbol{\alpha} \in \mathbb{R}^{N+1}} J(\boldsymbol{\alpha})$ is achieved,

(ii) $\min_{\boldsymbol{s} \in \boldsymbol{\mathcal{X}} \cap \boldsymbol{\mathcal{A}}} \mathcal{F}(\boldsymbol{s}) = \max_{\boldsymbol{\alpha} \in \mathbb{R}^{N+1}} J(\boldsymbol{\alpha})$.

The desired decomposition will then follow from equality conditions in (ii), and $\lambda(\boldsymbol{x}) = \lambda_{\bar{\boldsymbol{\alpha}}}(\boldsymbol{x})$ will be retrieved from any maximizer $\bar{\boldsymbol{\alpha}} \in \operatorname{Argmax} J$.

Remark B.2. The above dual problem can be guessed by introducing suitable Lagrange multipliers $\lambda(\boldsymbol{x}), \boldsymbol{\alpha}$ for the total saturation and mass constraints, respectively, and writing the convex indicator of the constraints as a supremum over these multipliers. Formally exchanging inf sup and sup inf and computing the optimality conditions in the rightmost infimum relates λ to $\boldsymbol{\alpha}$ as in (91), which in turn yields exactly the duality $\inf_{\boldsymbol{s}} \mathcal{F} = \max_{\boldsymbol{\alpha}} J$.

Let us first establish property (i). For all $\boldsymbol{\alpha} \in \mathbb{R}^{N+1}$ and all $\boldsymbol{s} \in \boldsymbol{\mathcal{X}} \cap \boldsymbol{\mathcal{A}}$, we first observe that

$$\begin{aligned} J(\boldsymbol{\alpha}) &= \int_\Omega \min_j \{F_j(\boldsymbol{x}) + \alpha_j\}\omega(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} - \sum_{i=0}^N \alpha_i m_i \\ &= \int_\Omega \min_j \{F_j(\boldsymbol{x}) + \alpha_j\} \sum_{i=0}^N s_i(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} - \sum_{i=0}^N \alpha_i \int_\Omega s_i(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} \\ &= \sum_{i=0}^N \int_\Omega \Big(\min_j \{F_j(\boldsymbol{x}) + \alpha_j\} - \alpha_i\Big) s_i(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} \leq \int_\Omega \boldsymbol{F} \cdot \boldsymbol{s}\,\mathrm{d}\boldsymbol{x} = \mathcal{F}(\boldsymbol{s}). \end{aligned}$$

In particular J is bounded from above and

$$\sup_{\boldsymbol{\alpha} \in \mathbb{R}^{N+1}} J(\boldsymbol{\alpha}) \leq \min_{\boldsymbol{s} \in \boldsymbol{\mathcal{X}} \cap \boldsymbol{\mathcal{A}}} \mathcal{F}(\boldsymbol{s}). \tag{92}$$

Since $\int \omega\,\mathrm{d}\boldsymbol{x} = \sum m_i$, the function J is invariant under diagonal shifts, i.e., $J(\boldsymbol{\alpha} + c\mathbf{1}) = J(\boldsymbol{\alpha})$ for any constant $c \in \mathbb{R}$. As a consequence we can choose a maximizing sequence $\{\boldsymbol{\alpha}^k\}_{k \geq 1}$ such that $\min_j \alpha_j^k = 0$

for all $k \geq 0$. Let $j(k)$ be an index such that $\alpha^k_{j(k)} = \min_j \alpha^k_j = 0$. Then, since $\boldsymbol{\alpha}^k$ is maximizing and $\omega(\boldsymbol{x}) \geq 0$, we get, for k large enough,

$$\sup J - 1 \leq J(\boldsymbol{\alpha}^k) = \int_\Omega \min_j \{F_j(\boldsymbol{x}) + \alpha^k_j\}\omega(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} - \sum \alpha^k_i m_i$$
$$\leq \int_\Omega \big(F_{j(k)}(\boldsymbol{x}) + \underbrace{\alpha^k_{j(k)}}_{=0}\big)\omega(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} - \sum \alpha^k_i m_i \leq \|\boldsymbol{F}\|_{L^\infty}\|\omega\|_{L^1} - \sum \alpha^k_i m_i.$$

Thus $\sum \alpha^k_i m_i \leq C$, and since $\alpha^k_i \geq 0$ and $m_i > 0$ we deduce that $(\boldsymbol{\alpha}^k)_k$ is bounded. Hence, up to extraction of a nonrelabeled subsequence, we can assume that $\boldsymbol{\alpha}^k$ converges towards some $\bar{\boldsymbol{\alpha}} \in \mathbb{R}^{N+1}_+$. The map J is continuous; hence $\bar{\boldsymbol{\alpha}}$ is a maximizer.

Let us now focus on property (ii). Note from (92) and (i) it suffices to prove the reverse inequality

$$\max_{\boldsymbol{\alpha} \in \mathbb{R}^{N+1}} J(\boldsymbol{\alpha}) \geq \min_{\boldsymbol{s} \in \boldsymbol{\mathcal{X}} \cap \boldsymbol{\mathcal{A}}} \mathcal{F}(\boldsymbol{s}).$$

We show below that, for any maximizer $\bar{\boldsymbol{\alpha}}$ of J, we can always construct a suitable $\boldsymbol{s} \in \boldsymbol{\mathcal{X}} \cap \boldsymbol{\mathcal{A}}$ such that $\mathcal{F}(\boldsymbol{s}) = J(\bar{\boldsymbol{\alpha}})$. This will immediately imply the reverse inequality and thus our claim (ii). In order to do so, we first observe that J is concave; thus the optimality condition at $\bar{\boldsymbol{\alpha}}$ can be written in terms of superdifferentials as $\mathbf{0}_{\mathbb{R}^{N+1}} \in \partial J(\bar{\boldsymbol{\alpha}})$. Denoting by

$$\Lambda(\boldsymbol{\alpha}) = \int_\Omega \lambda_{\boldsymbol{\alpha}}\,\omega\,\mathrm{d}\boldsymbol{x} = \int_\Omega \min_j \{F_j(x) + \alpha_j\}\omega(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}$$

the first contribution in J, this optimality can be recast as

$$\boldsymbol{m} \in \partial\Lambda(\bar{\boldsymbol{\alpha}}). \tag{93}$$

For fixed $\boldsymbol{x} \in \Omega$ and by usual properties of the min function, the superdifferential $\partial\lambda_{\boldsymbol{\alpha}}(\boldsymbol{x})$ of the concave map $\boldsymbol{\alpha} \mapsto \lambda_{\boldsymbol{\alpha}}(\boldsymbol{x})$ at $\boldsymbol{\alpha} \in \mathbb{R}^{N+1}$ is characterized by

$$\partial\lambda_{\boldsymbol{\alpha}}(\boldsymbol{x}) = \Big\{\boldsymbol{\theta} \in \mathbb{R}^{N+1}_+ \;\Big|\; \textstyle\sum_{i=0}^N \theta_i = 1 \text{ and } \theta_i = 0 \text{ if } F_i(\boldsymbol{x}) + \alpha_i > \lambda_{\boldsymbol{\alpha}}(\boldsymbol{x})\Big\}.$$

Therefore, it follows from the extension of the formula of differentiation under the integral to the nonsmooth case, see [Clarke 1990, Theorem 2.7.2], that

$$\partial\Lambda(\boldsymbol{\alpha}) = \Big\{\boldsymbol{w} \in \mathbb{R}^{N+1}_+ \;\Big|\; \boldsymbol{w} = \textstyle\int_\Omega \boldsymbol{\theta}(\boldsymbol{x})\omega(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} \text{ for some } \boldsymbol{\theta}(\boldsymbol{x}) \in \partial\lambda_{\boldsymbol{\alpha}}(\boldsymbol{x}) \text{ a.e. in } \Omega\Big\}. \tag{94}$$

The optimality criterion (93) at any maximizer $\bar{\boldsymbol{\alpha}}$ gives the existence of some function $\boldsymbol{\theta}$ as in (94) such that

$$m_i = \int_\Omega \theta_i(\boldsymbol{x})\omega(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} \quad \forall i \in \{0, \dots, N\}.$$

Defining

$$s_i(\boldsymbol{x}) := \theta_i(\boldsymbol{x})\omega(\boldsymbol{x}) \quad \forall i \in \{0, \dots, N\}, \tag{95}$$

we have by construction that $s_i \geq 0$, $\int s_i = m_i$, and $\sum s_i = \big(\sum_i \theta_i\big)\omega = \omega$ a.e.; thus $\boldsymbol{s} \in \boldsymbol{\mathcal{X}} \cap \boldsymbol{\mathcal{A}}$. Exploiting again $\sum s_i = \omega$ as well as the crucial property that $\theta_i = 0$ a.e. in $\{\boldsymbol{x} \mid F_i + \bar{\alpha}_i > \lambda_{\bar{\boldsymbol{\alpha}}}\}$, or in other words that $F_i + \bar{\alpha}_i = \lambda_{\bar{\boldsymbol{\alpha}}}$ for $\mathrm{d}s_i$-a.e $\boldsymbol{x} \in \Omega$, we get

$$J(\bar{\boldsymbol{\alpha}}) = \int_\Omega \lambda_{\bar{\boldsymbol{\alpha}}}\,\omega\,\mathrm{d}\boldsymbol{x} - \sum_{i=0}^N \bar{\alpha}_i m_i = \sum_{i=0}^N \int_\Omega \lambda_{\bar{\boldsymbol{\alpha}}} s_i\,\mathrm{d}\boldsymbol{x} - \sum_{i=0}^N \bar{\alpha}_i m_i = \sum_{i=0}^N \int_\Omega (F_i + \bar{\alpha}_i) s_i\,\mathrm{d}\boldsymbol{x} - \sum_{i=0}^N \bar{\alpha}_i m_i = \mathcal{F}(\boldsymbol{s})$$

as claimed. Therefore $\boldsymbol{s}$ constructed by (95) is a minimizer of $\mathcal{F}$ and

$$J(\bar{\boldsymbol{\alpha}}) = \mathcal{F}(\underline{\boldsymbol{s}}). \tag{96}$$

In order to finally retrieve the desired decomposition, choose any minimizer $\underline{\boldsymbol{s}} \in \boldsymbol{\mathcal{X}} \cap \boldsymbol{\mathcal{A}}$ of $\mathcal{F}$ and any maximizer $\bar{\boldsymbol{\alpha}} \in \mathbb{R}^{N+1}$ of J. Then it follows from (96) that

$$0 = \mathcal{F}(\underline{\boldsymbol{s}}) - J(\bar{\boldsymbol{\alpha}}) = \sum_{i=0}^N \int_\Omega F_i \underline{s}_i\,\mathrm{d}\boldsymbol{x} - \int_\Omega \lambda_{\bar{\boldsymbol{\alpha}}}\,\omega\,\mathrm{d}\boldsymbol{x} + \sum_{i=0}^N \bar{\alpha}_i m_i.$$

Using once again that $\int \underline{s}_i = m_i$ and $\sum_i \underline{s}_i = \omega$, we get that

$$\sum_{i=0}^N \int_\Omega (F_i + \bar{\alpha}_i - \lambda_{\bar{\boldsymbol{\alpha}}}) \underline{s}_i\,\mathrm{d}\boldsymbol{x} = 0.$$

By the definition of $\lambda_{\bar{\boldsymbol{\alpha}}}$, the above integrand is nonnegative; hence $F_i + \bar{\alpha}_i = \lambda_{\bar{\boldsymbol{\alpha}}}$ a.e. in $\{\underline{s}_i > 0\}$. □

Acknowledgements

This project was supported by the ANR GEOPOR project (ANR-13-JS01-0007-01). Cancès also acknowledges the support of Labex CEMPI (ANR-11-LABX-0007-01). Gallouët was supported by the ANR ISOTACE project (ANR-12-MONU-0013) and by the Fonds de la Recherche Scientifique - FNRS under Grant MIS F.4539.16. Monsaingeon was supported by the Portuguese Science Foundation 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 Cancès and Gallouët at CAMGSD, Instituto Superior Técnico, Universidade de Lisboa. The authors wish to thank Quentin Mérigot for fruitful discussion.

References

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

[Amaziane et al. 2012] B. Amaziane, M. Jurak, and A. Vrbaški, "Existence for a global pressure formulation of water-gas flow in porous media", *Electron. J. Differential Equations* **2012**:102 (2012), 1–22. MR Zbl

[Amaziane et al. 2014] 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 (2014), 297–309. MR

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

[Ambrosio and Serfaty 2008] L. Ambrosio and S. Serfaty, "A gradient flow approach to an evolution problem arising in superconductivity", *Comm. Pure Appl. Math.* **61**:11 (2008), 1495–1539. MR Zbl

[Ambrosio et al. 2008] L. Ambrosio, N. Gigli, and G. Savaré, *Gradient flows in metric spaces and in the space of probability measures*, 2nd ed., Birkhäuser, Basel, 2008. MR Zbl

[Ambrosio et al. 2011] 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 (2011), 217–246. MR Zbl

[Andreianov et al. 2015] B. Andreianov, C. Cancès, and A. Moussa, "A nonlinear time compactness result and applications to discretization of degenerate parabolic-elliptic PDEs", preprint, 2015, available at https://hal.archives-ouvertes.fr/hal-01142499/document.

[Antoncev and Monahov 1978] 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 (1978), 553–556. In Russian; translated in *Soviet Math., Dokl.* **19** (1978), 1354–1358. MR Zbl

[Bear and Bachmat 1990] J. Bear and Y. Bachmat, *Introduction to modeling of transport phenomena in porous media*, Springer, 1990. Zbl

[Blanchet 2013] A. Blanchet, "A gradient flow approach to the Keller–Segel systems", *RIMS Kôkyûroku* **1837** (2013), 52–73.

[Blanchet et al. 2008] 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 (2008), 691–721. MR Zbl

[Bolley et al. 2013] F. Bolley, I. Gentil, and A. Guillin, "Uniform convergence to equilibrium for granular media", *Arch. Ration. Mech. Anal.* **208**:2 (2013), 429–445. MR Zbl

[Cancès and Gallouët 2011] C. Cancès and T. Gallouët, "On the time continuity of entropy solutions", *J. Evol. Equ.* **11**:1 (2011), 43–55. MR Zbl

[Cancès et al. 2015] C. Cancès, T. O. Gallouët, and L. Monsaingeon, "The gradient flow structure for incompressible immiscible two-phase flows in porous media", *C. R. Math. Acad. Sci. Paris* **353**:11 (2015), 985–989. MR Zbl

[Carlier and Laborde 2015] G. Carlier and M. Laborde, "On systems of continuity equations with nonlinear diffusion and nonlocal drifts", preprint, 2015. arXiv

[Carrillo et al. 2011] 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 (2011), 229–271. MR Zbl

[Chavent 1976] G. Chavent, "A new formulation of diphasic incompressible flows in porous media", pp. 258–270 in *Applications of methods of functional analysis to problems in mechanics* (Marseille, 1975), Lecture Notes in Math. **503**, Springer, 1976. MR Zbl

[Chavent 2009] G. Chavent, "A fully equivalent global pressure formulation for three-phases compressible flows", *Appl. Anal.* **88**:10-11 (2009), 1527–1541. MR Zbl

[Chavent and Jaffré 1986] G. Chavent and J. Jaffré, *Mathematical models and finite elements for reservoir simulation*, Studies in Mathematics and its Applications **17**, North Holland, Amsterdam, 1986. Zbl

[Chavent and Salzano 1985] G. Chavent and G. Salzano, "Un algorithme pour la détermination de perméabilités relatives triphasiques satisfaisant une condition de différentielle totale", INRIA Technical Report 335, 1985, available at https://hal.inria.fr/inria-00076202v1.

[Chen 2001] Z. Chen, "Degenerate two-phase incompressible flow, I: Existence, uniqueness and regularity of a weak solution", *J. Differential Equations* **171**:2 (2001), 203–232. MR Zbl

[Clarke 1990] F. H. Clarke, *Optimization and nonsmooth analysis*, 2nd ed., Classics in Applied Mathematics **5**, SIAM, Philadelphia, PA, 1990. MR Zbl

[De Giorgi 1993] E. De Giorgi, "New problems on minimizing movements", pp. 81–98 in *Boundary value problems for partial differential equations and applications*, edited by J.-L. Lions and C. Baiocchi, RMA Res. Notes Appl. Math. **29**, Masson, Paris, 1993. MR Zbl

[Dolbeault et al. 2009] J. Dolbeault, B. Nazaret, and G. Savaré, "A new class of transport distances between measures", *Calc. Var. Partial Differential Equations* **34**:2 (2009), 193–231. MR Zbl

[Fabrie and Saad 1993] 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 (1993), 337–373. MR Zbl

[Gagneux and Madaune-Tort 1996] G. Gagneux and M. Madaune-Tort, *Analyse mathématique de modèles non linéaires de l'ingénierie pétrolière*, Mathématiques & Applications (Berlin) **22**, Springer, 1996. MR Zbl

[Gigli and Otto 2013] N. Gigli and F. Otto, "Entropic Burgers' equation via a minimizing movement scheme based on the Wasserstein metric", *Calc. Var. Partial Differential Equations* **47**:1-2 (2013), 181–206. MR Zbl

[Hanche-Olsen and Holden 2010] H. Hanche-Olsen and H. Holden, "The Kolmogorov–Riesz compactness theorem", *Expo. Math.* **28**:4 (2010), 385–394. MR Zbl

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

[Kinderlehrer et al. 2017] D. Kinderlehrer, L. Monsaingeon, and X. Xu, "A Wasserstein gradient flow approach to Poisson–Nernst–Planck equations", *ESAIM Control Optim. Calc. Var.* **23**:1 (2017), 137–164. MR Zbl

[Laborde 2016] M. Laborde, *Systèmes de particules en interaction, approche par flot de gradient dans l'espace de Wasserstein*, Ph.D. thesis, Université Paris-Dauphine, 2016, available at https://basepub.dauphine.fr/handle/123456789/16518.

[Ladyženskaja et al. 1968] O. A. Ladyženskaja, V. A. Solonnikov, and N. N. Ural'ceva, *Linear and quasilinear equations of parabolic type*, Translations of Mathematical Monographs **23**, Amer. Math. Soc., Providence, RI, 1968. MR Zbl

[Laurençot and Matioc 2013] 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 (2013), 319–341. MR Zbl

[Lieb and Loss 2001] E. H. Lieb and M. Loss, *Analysis*, 2nd ed., Graduate Studies in Mathematics **14**, Amer. Math. Soc., Providence, RI, 2001. MR Zbl

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

[Lisini et al. 2012] 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 (2012), 814–850. MR Zbl

[Matthes et al. 2009] D. Matthes, R. J. McCann, and G. Savaré, "A family of nonlinear fourth order equations of gradient flow type", *Comm. Partial Differential Equations* **34**:10-12 (2009), 1352–1397. MR Zbl

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

[Moussa 2016] A. Moussa, "Some variants of the classical Aubin–Lions lemma", *J. Evol. Equ.* **16**:1 (2016), 65–93. MR Zbl

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

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

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

[Santambrogio 2015] F. Santambrogio, *Optimal transport for applied mathematicians; calculus of variations, PDEs, and modeling*, Progress in Nonlinear Differential Equations and their Applications **87**, Springer, 2015. MR Zbl

[Villani 2009] C. Villani, *Optimal transport: old and new*, Grundlehren der Mathematischen Wissenschaften **338**, Springer, 2009. MR Zbl

[Zinsl 2014] J. Zinsl, "Existence of solutions for a nonlinear system of parabolic equations with gradient flow structure", *Monatsh. Math.* **174**:4 (2014), 653–679. MR Zbl

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

[Zinsl and Matthes 2015b] J. Zinsl and D. Matthes, "Transport distances and geodesic convexity for systems of degenerate diffusion equations", *Calc. Var. Partial Differential Equations* **54**:4 (2015), 3397–3438. MR Zbl

Received 13 Jul 2016. Revised 23 May 2017. Accepted 29 Jun 2017.

CLÉMENT CANCÈS: clement.cances@inria.fr
Team RAPSODI, Inria Lille – Nord Europe, Lille, France

THOMAS O. GALLOUËT: thomas.gallouet@ulg.ac.be
Département de Mathématiques, Université de Liège, Liège, Belgium

LÉONARD MONSAINGEON: leonard.monsaingeon@univ-lorraine.fr
Institut Élie Cartan de Lorraine, Université de Lorraine, Nancy, France

mathematical sciences publishers

Analysis & PDE

msp.org/apde

See inside back cover or msp.org/apde for submission instructions.

The subscription price for 2017 is US $265/year for the electronic version, and $470/year (+$55, if shipping outside the US) for print and electronic. Subscriptions, requests for back issues from the last three years and changes of subscriber address should be sent to MSP.

Analysis & PDE (ISSN 1948-206X electronic, 2157-5045 printed) at Mathematical Sciences Publishers, 798 Evans Hall #3840, c/o University of California, Berkeley, CA 94720-3840, is published continuously online. Periodical rate postage paid at Berkeley, CA 94704, and additional mailing offices.

APDE peer review and production are managed by EditFlow® from MSP.

PUBLISHED BY

mathematical sciences publishers

nonprofit scientific publishing

http://msp.org/

ANALYSIS & PDE

Volume 10 No. 8 2017