



HAL
open science

A convergent finite volume scheme for dissipation driven models with volume filling constraint

Clément Cancès, Antoine Zurek

► **To cite this version:**

Clément Cancès, Antoine Zurek. A convergent finite volume scheme for dissipation driven models with volume filling constraint. *Numerische Mathematik*, 2022, 151, pp.279–328. 10.1007/s00211-022-01270-7 . hal-03166069v2

HAL Id: hal-03166069

<https://hal.science/hal-03166069v2>

Submitted on 19 Jan 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

A CONVERGENT FINITE VOLUME SCHEME FOR DISSIPATION DRIVEN MODELS WITH VOLUME FILLING CONSTRAINT

CLÉMENT CANCÈS AND ANTOINE ZUREK

ABSTRACT. In this paper we propose and study an implicit finite volume scheme for a general model which describes the evolution of the composition of a multi-component mixture in a bounded domain. We assume that the whole domain is occupied by the different phases of the mixture which leads to a volume filling constraint. In the continuous model this constraint yields the introduction of a pressure, which should be thought as a Lagrange multiplier for the volume filling constraint. The pressure solves an elliptic equation, to be coupled with parabolic equations, possibly including cross-diffusion terms, which govern the evolution of the mixture composition. Besides the system admits an entropy structure which is at the cornerstone of our analysis. More precisely, the main objective of this work is to design a two-point flux approximation finite volume scheme which preserves the key properties of the continuous model, namely the volume filling constraint and the control of the entropy production. Thanks to these properties, and in particular the discrete entropy-entropy dissipation relation, we are able to prove the existence of solutions to the scheme and its convergence. Finally, we illustrate the behavior of our scheme through different applications.

1. INTRODUCTION

In this work we consider in Ω a connected bounded domain of \mathbb{R}^N and for any arbitrary finite time horizon $T > 0$ the evolution of a mixture made of $n + 1$ components. The evolution of the mixture composition $\mathbf{u} = (u_i)_{0 \leq i \leq n}$ is given by the following partial differential equations

$$(1) \quad \partial_t u_i + \operatorname{div}(-d_i u_i \nabla \mu_i) = 0 \quad \text{in } Q_T = \Omega \times (0, T), \quad i = 0, \dots, n,$$

where $d_i > 0$ is the diffusivity of the i^{th} component, μ_i is its potential and u_i denotes its volume fraction. The whole domain Ω is occupied by the $n + 1$ phases, leading to the following volume filling constraint

$$(2) \quad \langle \mathbf{u}, \mathbf{1} \rangle = \sum_{i=0}^n u_i = 1 \quad \text{in } Q_T,$$

where $\langle \cdot, \cdot \rangle$ stands for the canonical scalar product of \mathbb{R}^{n+1} . In what follows, we use the same notation for the canonical scalar product of $\mathbb{R}^{(n+1) \times N}$. The difference of the potentials between the phase i and the phase 0 is given by

$$(3a) \quad \mu_i - \mu_0 = \frac{\delta E}{\delta u_i}(\mathbf{u}) - \frac{\delta E}{\delta u_0}(\mathbf{u}) \quad \text{in } Q_T, \quad i = 1, \dots, n.$$

Here, $\frac{\delta E}{\delta \mathbf{u}}(\mathbf{u}) = \left(\frac{\delta E}{\delta u_i}(\mathbf{u}) \right)_{0 \leq i \leq n} : Q_T \rightarrow \mathbb{R}^{n+1}$ denotes the first variation (or Fréchet derivative) of some energy functional $E : [0, 1]^{n+1} \rightarrow \mathbb{R}$. Motivated by applications to be discussed later on in Section 8, the energy functional E is assumed to have the generic form

$$E(\mathbf{u}) = \sum_{i=0}^n \int_{\Omega} [\alpha_i (u_i (\log(u_i) - 1) + 1) + u_i \Psi_i] \, dx + \int_{\Omega} B(\mathbf{u}) \, dx.$$

In the above definition of E , the coefficients $\alpha_i > 0$ can be thought as related to thermal agitation, Ψ_i a given external potential acting on phase i (and assumed to be independent on time for

Date: January 19, 2022.

Key words and phrases. Cross-diffusion system, volume filling, finite volume method, convergence analysis.

simplicity) and $B : \mathbb{R}^{n+1} \rightarrow \mathbb{R}$ is a $C^{1,1}$ convex function. Using the above definition of E we can rewrite the equation on the difference of the potentials as

$$(3b) \quad \mu_i - \mu_0 = \alpha_i \log(u_i) - \alpha_0 \log(u_0) + \Psi_i - \Psi_0 + b_i(\mathbf{u}) - b_0(\mathbf{u}), \quad i = 1, \dots, n,$$

where we denoted by $b_i(\mathbf{u}) = \frac{\partial B}{\partial u_i}(\mathbf{u})$ for $i = 0, \dots, n$. Finally, we prescribe no-flux boundary and initial conditions:

$$(4) \quad d_i u_i \nabla \mu_i \cdot \nu = 0 \quad \text{on } \partial\Omega \times (0, T), \quad u_i(\cdot, 0) = u_i^0 \quad \text{in } \Omega \text{ for } i = 0, \dots, n,$$

where we assume that u_i^0 is nonnegative for $i = 0, \dots, n$ with $\sum_{i=0}^n u_i^0(x) = 1$ for a.e. $x \in \Omega$ and ν denotes the exterior unit normal vector to $\partial\Omega$.

The system (1)–(4) is not yet closed. Indeed, the potentials $\boldsymbol{\mu} = (\mu_i)_{0 \leq i \leq n}$ are defined up to a common additive constant due to the no-flux boundary conditions. The approach proposed in [33] for fixing this degree of freedom consists in enforcing a zero-mean condition on some averaged potential, i.e.,

$$\sum_{i=0}^n \int_{\Omega} u_i(x, t) \mu_i(x, t) dx = 0 \quad \text{for } 0 \leq t \leq T.$$

Here, we rather impose the condition

$$(5) \quad \sum_{i=0}^n \int_{\Omega} u_i(x, t) \mu_i(x, t) dx = \sum_{i=0}^n \int_{\Omega} u_i(x, t) \frac{\delta E}{\delta u_i}(\mathbf{u}(x, t)) dx \quad \text{for } 0 \leq t \leq T,$$

the meaning of which will become clear after a reformulation of the problem to be presented in Section 2.2.

2. PRELIMINARY CONSIDERATIONS ON THE CONTINUOUS MODEL

Prior to discretizing the model under consideration, we discuss several aspects of the continuous model. First, we highlight in Section 2.1 that the continuous model can be interpreted as the generalized gradient flow of the energy E in a geometry related to optimal transportation constrained by the volume filling condition (2). A reformulation of the problem is then proposed in Section 2.2 in order to make explicit the corresponding Lagrange multiplier, while we introduce in Section 2.3 an entropy production estimate on which the definition for weak solutions and our numerical analysis will rely. Finally, in Section 2.4 we present the main objectives of this work and give a state of the art.

2.1. Variational interpretation of the model. The system (1)–(5) can be interpreted as a generalized gradient flow in some constrained Wasserstein space. More precisely, the space

$$\mathcal{M} = \left\{ \mathbf{u} \in L^1(\Omega; \mathbb{R}_+^{n+1}) \mid \langle \mathbf{u}, \mathbf{1} \rangle = 1 \text{ a.e. in } \Omega \text{ and } \int_{\Omega} \mathbf{u} dx = \bar{\mathbf{u}} \right\},$$

where the mean values $\bar{\mathbf{u}} = (\bar{u}_i)_{0 \leq i \leq n} \in (0, 1)^{n+1}$ are such that $\mathbf{u}^0 = (u_i^0)_{0 \leq i \leq n}$ belongs to \mathcal{M} . In particular, $\langle \bar{\mathbf{u}}, \mathbf{1} \rangle = 1$. The space \mathcal{M} can be equipped by the constrained Wasserstein metric defined as follows: let $\mathbf{u}_0, \mathbf{u}_1$ be two elements of \mathcal{M} , then define the squared distance

$$(6) \quad \mathcal{W}^2(\mathbf{u}_0, \mathbf{u}_1) = \inf_{\mathbf{F} \in \mathcal{X}} \sum_{i=0}^n \frac{1}{d_i} \int_0^1 \int_{\Omega} \mathcal{B}(u_i, F_i) dx dt,$$

where $\mathcal{B} : \mathbb{R} \times \mathbb{R}^N \rightarrow \mathbb{R}_+ \cup \{+\infty\}$ is the convex lower semi-continuous function introduced by Benamou and Brenier [10] defined by

$$\mathcal{B}(u, F) = \begin{cases} \frac{|F|^2}{u} & \text{if } u > 0, \\ 0 & \text{if } F = 0 \text{ and } u = 0, \\ +\infty & \text{otherwise.} \end{cases}$$

In (6), the fluxes $\mathbf{F} = (F_i)_{0 \leq i \leq n}$ are required to belong to the set

$$\mathcal{X} = \left\{ \mathbf{F} : \Omega \times (0, T) \rightarrow \mathbb{R}^{n \times N} \mid \right.$$

$$\partial_t u_i + \operatorname{div} F_i = 0, \quad F_i \cdot \nu = 0, \quad \mathbf{u}(\cdot, 0) = \mathbf{u}_0, \quad \mathbf{u}(\cdot, 1) = \mathbf{u}_1, \quad \langle \operatorname{div} \mathbf{F}, \mathbf{1} \rangle = 0 \}.$$

While the first four constraints in the definition of \mathcal{X} classically appear in the dynamical formulation of the optimal transportation problem, the last constraint $\langle \operatorname{div} \mathbf{F}, \mathbf{1} \rangle = 0$ is there to enforce that $\mathbf{u}(\cdot, t) \in \mathcal{M}$ for all $t \in [0, 1]$, i.e., all along the geodesics joining \mathbf{u}_0 to \mathbf{u}_1 . We refer to [11] for the computation of such geodesics.

Remark 1. *The definition of the space \mathcal{M} depends on the initial data \mathbf{u}^0 via the mass allocated to each species, so that $\mathcal{M}_{\bar{\mathbf{u}}}$ would have been a more rigorous notation. For the ease of reading, this dependency of \mathcal{M} with respect to the mass repartition will be highlighted in what follows.*

Saying that the dynamical system (1)–(5) can be seen as the gradient flow of the energy E in the metric space \mathcal{M} equipped with the metric \mathcal{W} means that this system evolves along the steepest descent direction of the energy E for the geometry of optimal transportation with constraint (2). While the gradient flow structure in this complex geometry was already pointed out in the pioneering works [68], the mathematical analysis for such problems using a variational approach has been proposed much more recently. The model for incompressible multiphase porous media flows considered in [27] is very close to our problem, whereas [33, 32] deal with the extension to fourth-order dissipative system.

As a consequence of the gradient flow structure of the problem (1), the energy $E(\mathbf{u})$ decays along time. More precisely, multiplying (1) by μ_i , summing over $i = 0, \dots, n$ and integrating over Ω yields after integration by parts

$$\int_{\Omega} \langle \partial_t \mathbf{u}, \boldsymbol{\mu} \rangle dx + \sum_{i=0}^n \int_{\Omega} d_i u_i |\nabla \mu_i|^2 dx = 0.$$

Using (2) and (3), the first term rewrites

$$\int_{\Omega} \langle \partial_t \mathbf{u}, \boldsymbol{\mu} \rangle dx = \int_{\Omega} \langle \partial_t \mathbf{u}, \frac{\delta E}{\delta \mathbf{u}} \rangle dx = \frac{d}{dt} E(\mathbf{u}).$$

Therefore,

$$(7) \quad \frac{d}{dt} E(\mathbf{u}) = - \sum_{i=0}^n \int_{\Omega} d_i u_i |\nabla \mu_i|^2 dx \leq 0.$$

This energy-dissipation relation is one of the core properties of the system. However, the analysis carried out in the paper does not rely on it, but rather on the control of the production rate of an entropy functional to be introduced in Section 2.3.

2.2. Reformulation of the model (1)–(4). Let $\mathbf{u}, \boldsymbol{\mu}$ be C^1 functions satisfying (1)–(4). It follows from the condition (3) that the potentials $\boldsymbol{\mu}$ can be rewritten

$$(8) \quad \mu_i = \frac{\delta E}{\delta u_i}(\mathbf{u}) + \pi = \alpha_i \log(u_i) + \Psi_i + b_i(\mathbf{u}) + \pi, \quad i = 0, \dots, n.$$

Following [68, 27], the pressure $\pi : Q_T \rightarrow \mathbb{R}$ defined by the algebraic relation (8) shall be thought as the Lagrange multiplier for the constraint (2). Now, our goal is to rewrite the system (1)–(4) with the $2n + 2$ unknown functions μ_0, \dots, μ_n and u_0, \dots, u_n as a system of $n + 1$ unknown functions u_1, \dots, u_n, π . In this purpose, we first rewrite, using the expression (8) of the potential μ_i , the equation (1) as

$$(9) \quad \partial_t u_i + \operatorname{div} (-\alpha_i d_i \nabla u_i - d_i u_i \nabla \Psi_i - d_i u_i \nabla b_i(\mathbf{u}) - d_i u_i \nabla \pi) = 0, \quad i = 0, \dots, n.$$

Then, summing over $i = 0, \dots, n$ these equations and using the constraint (2) we deduce that

$$\sum_{i=0}^n \operatorname{div} (-\alpha_i d_i \nabla u_i - d_i u_i \nabla \Psi_i - d_i u_i \nabla b_i(\mathbf{u}) - d_i u_i \nabla \pi) = 0.$$

Thus, the function π solves the following elliptic equation

$$\operatorname{div} \left(- \left(\sum_{i=0}^n d_i u_i \right) \nabla \pi \right) = \sum_{i=0}^n \operatorname{div} (\alpha_i d_i \nabla u_i + d_i u_i \nabla \Psi_i + d_i u_i \nabla b_i(\mathbf{u})).$$

Now, we rewrite the system (1)–(4) as: find a solution u_1, \dots, u_n, π of the system

$$(10a) \quad \partial_t u_i + \operatorname{div} \mathcal{F}_i = 0 \quad \text{in } Q_T, \quad i = 1, \dots, n,$$

$$(10b) \quad \operatorname{div} \left(- \left(\sum_{i=0}^n d_i u_i \right) \nabla \pi \right) + \operatorname{div} \mathcal{G} = 0 \quad \text{in } Q_T,$$

where the fluxes are given by

$$(11a) \quad \mathcal{F}_i = -\alpha_i d_i \nabla u_i - d_i u_i \nabla \Psi_i - d_i u_i \nabla b_i(\mathbf{u}) - d_i u_i \nabla \pi, \quad \text{for } i = 1, \dots, n,$$

$$(11b) \quad \mathcal{G} = - \sum_{i=0}^n (\alpha_i d_i \nabla u_i + d_i u_i \nabla \Psi_i + d_i u_i \nabla b_i(\mathbf{u})),$$

and u_0 is defined as

$$(12) \quad u_0 = 1 - \sum_{i=1}^n u_i \quad \text{in } Q_T.$$

So far, π is defined up to an additive constant. We eliminate this degree of freedom by imposing (5), which in the current context simply rewrites

$$(13) \quad \int_{\Omega} \pi(x, t) dx = 0 \quad \text{for } 0 \leq t \leq T.$$

Eventually we prescribe no-flux boundary conditions

$$(14) \quad \mathcal{F}_i \cdot \nu = 0, \quad \text{on } \partial\Omega \times (0, T), \quad \text{for } i = 1, \dots, n,$$

and initial conditions

$$(15) \quad \mathbf{u}(\cdot, 0) = \mathbf{u}^0 \in \mathcal{M},$$

the vector $\bar{\mathbf{u}}$ of the average values appearing in the definition of \mathcal{M} being adapted to the choice of \mathbf{u}^0 .

From the previous computations we deduce that any classical solution of (1)–(4) such that (5) holds, is also a solution to (10)–(15). In the following statement we prove the other way around.

Lemma 2. *Any classical solution solution of (10)–(15) is also a solution to (1)–(4) satisfying (5).*

Proof. First, using equation (10b) and the definition of \mathcal{G} we obtain that

$$(16) \quad \operatorname{div} (\alpha_0 d_0 \nabla u_0 + d_0 u_0 \nabla \Psi_0 + d_0 u_0 \nabla b_0(\mathbf{u}) + d_0 u_0 \nabla \pi) = \sum_{i=1}^n \operatorname{div} \mathcal{F}_i.$$

Then, we simply derive in time the constraint (12) and we have

$$\partial_t u_0 = - \sum_{i=1}^n \partial_t u_i = \sum_{i=1}^n \operatorname{div} \mathcal{F}_i.$$

Inserting (16) in the above equation implies that

$$\partial_i u_i + \operatorname{div} (-d_i u_i \nabla \mu_i) = 0 \quad \text{in } Q_T, \quad i = 0, \dots, n.$$

Moreover, the equality (5) is a direct consequence of (13) and the constraint (12) since

$$\begin{aligned} 0 &= \int_{\Omega} \pi(x, t) dx = \sum_{i=0}^n \int_{\Omega} u_i(x, t) \pi(x, t) dx \\ &= \sum_{i=0}^n \int_{\Omega} u_i(x, t) \left(\mu_i(x, t) - \frac{\delta E}{\delta u_i}(\mathbf{u}(x, t)) \right) dx. \end{aligned}$$

This finishes the proof of Lemma 2. □

2.3. Control on the entropy production and weak formulation. As already mentioned in Section 2.1, the numerical analysis to be carried out in the paper does not rely on the energy dissipation estimate (7) but rather on the control of the dissipation of another functional, that we call entropy functional in what follows. The purpose of this section is to present this entropy production estimate at the continuous level, as well as an estimation on the pressure field π . First, assuming the existence of a classical solution to (10)–(15) with positive volume fractions u_i we establish some a priori estimates required to deduce the sufficient regularity needed to define a proper notion of weak solution for the system. Second, the Finite Volume scheme presented in Section 3 has been designed in order to ensure that discrete counterparts of these estimates are satisfied. The rigorous proof for the convergence of the scheme will then rely on these discrete versions of the estimates below.

Let us introduce the entropy functional $H : \mathcal{M} \rightarrow \mathbb{R}_+$ defined as

$$(17) \quad H(\mathbf{u}) = \sum_{i=0}^n \frac{1}{d_i} \int_{\Omega} (u_i(\log(u_i) - 1) + 1) \, dx.$$

Note that H is bounded on \mathcal{M} , hence $H(\mathbf{u}^0) < +\infty$ whatever $\mathbf{u}^0 \in \mathcal{M}$.

Lemma 3. *Let (\mathbf{u}, π) be a classical solution to (10)–(15) where $\mathbf{u} : [0, T] \rightarrow \mathcal{M}$ is assumed positive, then*

$$(18) \quad \frac{d}{dt} H(\mathbf{u}) + \sum_{i=0}^n 2\alpha_i \int_{\Omega} |\nabla \sqrt{u_i}|^2 \, dx + \int_{\Omega} \langle \nabla \mathbf{u}, \nabla \mathbf{b}(\mathbf{u}) \rangle \, dx \leq \sum_{i=0}^n \frac{1}{2\alpha_i} \int_{\Omega} |\nabla \Psi_i|^2 \, dx.$$

In particular, if $\Psi = (\Psi_i)_{0 \leq i \leq n} = \mathbf{0}$, then H is a Lyapunov functional for the system (10)–(15) since the convexity of B implies that $\langle \nabla \mathbf{u}, \nabla \mathbf{b}(\mathbf{u}) \rangle \geq 0$.

Proof. In this purpose we use the equivalence established in Lemma 2 and we multiply (9) by $\log(u_i)/d_i$ for $i = 0, \dots, n$ (which is well defined since we assume that u_i is positive, i.e., $u_i > 0$ everywhere on Q_T), we integrate in space, we use some integration by parts and we sum over i to obtain the following relation

$$\sum_{i=0}^n \frac{1}{d_i} \int_{\Omega} \partial_t u_i \log(u_i) \, dx + \sum_{i=0}^n \int_{\Omega} \left(\alpha_i \nabla u_i + u_i \nabla \Psi_i + u_i \nabla b_i(\mathbf{u}) + u_i \nabla \pi \right) \cdot \nabla \log(u_i) \, dx = 0.$$

Since $\frac{\delta H}{\delta u_i}(\mathbf{u}) = \frac{1}{d_i} \log(u_i)$, one has

$$\sum_{i=0}^n \frac{1}{d_i} \int_{\Omega} \partial_t u_i \log(u_i) \, dx = \frac{d}{dt} H(\mathbf{u}).$$

On the other hand, since $u_i \nabla \log(u_i) = \nabla u_i$ and since $\langle \nabla \mathbf{u}, \mathbf{1} \rangle = 0$ in view of (2), one gets that

$$\sum_{i=0}^n \int_{\Omega} \left(u_i \nabla \Psi_i + u_i \nabla b_i(\mathbf{u}) + u_i \nabla \pi \right) \cdot \nabla \log(u_i) \, dx = \int_{\Omega} \langle \nabla \mathbf{u}, \nabla \mathbf{b}(\mathbf{u}) + \nabla \Psi \rangle \, dx.$$

Besides, a simple chain rule together with the uniform bound $0 < u_i \leq 1$ leads to

$$\nabla u_i \cdot \nabla \log(u_i) = 4|\nabla \sqrt{u_i}|^2 \geq |\nabla u_i|^2.$$

Altogether, one obtains that

$$\frac{d}{dt} H(\mathbf{u}) + \sum_{i=0}^n \int_{\Omega} \left(2\alpha_i |\nabla \sqrt{u_i}|^2 + \frac{\alpha_i}{2} |\nabla u_i|^2 \right) dx + \int_{\Omega} \langle \nabla \mathbf{u}, \nabla \mathbf{b}(\mathbf{u}) \rangle \, dx = - \int_{\Omega} \langle \nabla \mathbf{u}, \nabla \Psi \rangle \, dx.$$

Then it only remains to use the weighted Young inequality

$$-\langle \nabla \mathbf{u}, \nabla \Psi \rangle \leq \sum_{i=0}^n \left(\frac{\alpha_i}{2} |\nabla u_i|^2 + \frac{1}{2\alpha_i} |\nabla \Psi_i|^2 \right),$$

to recover (18). Finally, if $\Psi = \mathbf{0}$, then the right-hand side in (18) vanishes. As B is a convex function, one obtains that $\frac{d}{dt} H(\mathbf{u}) \leq 0$, hence H is a Lyapunov functional. \square

Thanks to Lemma 3 together with the $L^\infty(Q_T)$ estimate stemming from $\mathbf{u}(t) \in \mathcal{M}$ for $t \in [0, T]$, we conclude that \mathbf{u} is bounded in $L^2(0, T; H^1(\Omega))^{n+1}$. It remains now to establish an a priori estimate for the function π .

Lemma 4. *There exists a constant $C > 0$ which depends on $d_i, \alpha_i, T, H(u^0), \|Db_i\|_\infty, \|\nabla\Psi_i\|_{L^2(\Omega)^N}$ and $\|\nabla\sqrt{u_i}\|_{L^2(\Omega)^N}$, which is uniformly bounded thanks to (18), for $i = 0, \dots, n$ such that*

$$(19) \quad \|\nabla\pi\|_{L^2(Q_T)^N} \leq C.$$

Proof. We multiply (10b) by π , we integrate over Q_T , we use an integration by parts, the Cauchy-Schwarz inequality, the L^∞ estimate $0 \leq u_i \leq 1$ and the Lipschitz regularity of the functions b_i for $i = 0, \dots, n$ to get

$$\begin{aligned} \iint_{Q_T} \left(\sum_{i=0}^n d_i u_i \right) |\nabla\pi|^2 dx &\leq \sum_{i=0}^n d_i \left(\alpha_i \|\nabla u_i\|_{L^2(Q_T)^N} + T^{1/2} \|\nabla\Psi_i\|_{L^2(\Omega)^N} \right. \\ &\quad \left. + \|Db_i\|_\infty \sum_{j=0}^n \|\nabla u_j\|_{L^2(Q_T)^N} \right) \|\nabla\pi\|_{L^2(Q_T)^N}. \end{aligned}$$

To conclude the proof, it remains to apply the L^∞ estimate $0 \leq u_i \leq 1$ and to notice that thanks to (2), one has $(\sum_{i=0}^n d_i u_i) \geq \min_{i=0, \dots, n} d_i =: d_*$, so that the estimate

$$\|\nabla\pi\|_{L^2(Q_T)^N} \leq \frac{1}{d_*} \sum_{i=0}^n d_i \left(\alpha_i \|\nabla u_i\|_{L^2(Q_T)^N} + T^{1/2} \|\nabla\Psi_i\|_{L^2(\Omega)^N} + 2\|Db_i\|_\infty \sum_{j=0}^n \|\nabla\sqrt{u_j}\|_{L^2(Q_T)^N} \right),$$

holds. \square

Now applying the Poincaré-Wirtinger inequality and using the zero-mean constraint (13) we deduce that there exists a constant, still denoted by C and depending only on Ω such that

$$\|\pi(t)\|_{L^2(\Omega)} \leq C \|\nabla\pi(t)\|_{L^2(\Omega)^N}, \quad \text{for a.e. } t \in (0, T).$$

In particular, as a direct consequence of Lemma 4 we conclude that $\pi \in L^2(0, T; H^1(\Omega))$.

Thanks to the results obtained in this section we can define a notion of weak solution for the system (10)–(15).

Definition 1. *(\mathbf{u}, π) is said to be a weak solution of (10)–(15) if*

- $\pi \in L^2(0, T; H^1(\Omega))$, $\mathbf{u} \in L^\infty(Q_T)^{n+1} \cap L^2(0, T; H^1(\Omega))^{n+1}$ with $\mathbf{u}(\cdot, t) \in \mathcal{M}$ for a.e. $t \in (0, T)$ and π satisfying the zero-mean condition (13) for a.e. $t \in (0, T)$;
- for all $\varphi \in C_c^\infty(\bar{\Omega} \times [0, T])$ and all $i \in \{0, \dots, n\}$, there holds

$$(20) \quad \begin{aligned} \iint_{Q_T} u_i \partial_t \varphi dx dt + \int_{\Omega} u_i^0 \varphi(\cdot, 0) dx \\ = d_i \iint_{Q_T} (\alpha_i \nabla u_i + u_i \nabla\Psi_i + u_i \nabla b_i(\mathbf{u}) + u_i \nabla\pi) \cdot \nabla\varphi dx dt. \end{aligned}$$

Note that summing (20), one recovers the weak formulation

$$(21) \quad \sum_{i=0}^n d_i \iint_{Q_T} u_i \nabla\pi \cdot \nabla\varphi dx dt = - \sum_{i=0}^n d_i \iint_{Q_T} (\alpha_i \nabla u_i + u_i \nabla\Psi_i + u_i \nabla b_i(\mathbf{u})) \cdot \nabla\varphi dx dt,$$

with $\varphi \in C_c^\infty(\bar{\Omega} \times [0, T])$ arbitrary, which implies that (10b) holds in the distributional sense.

2.4. Main objectives and state of the art. Before to enter in the core of this work let us explain the main goals of the paper and describe its position within the literature of entropy preserving numerical methods. Our main objective is to design and analyze a fully implicit two-point flux approximation finite volume scheme for the system (10)–(15) preserving at the discrete level the following properties:

- (i) The conservation of the mass (see Lemma 10);
- (ii) The nonnegativity of the mixture composition \mathbf{u} (see Lemma 11);

- (iii) The volume filling constraint;
- (iv) the entropy-entropy dissipation relation (see Proposition 9).

The property (i) will be a consequence of the conservativity of the finite volume scheme introduced in Section 3.2, the properties (ii) and (iii) will be established thanks to the monotonicity of the scheme while the proof of (iv) will rely on the entropy stability of our numerical method. In particular, this last property is at the cornerstone of the proofs of our main results, namely the existence of nonnegative solutions to the scheme at each iteration (Theorem 7) and its convergence (Theorem 8).

Let us notice that the idea to build finite-volume schemes which preserved at the discrete level the entropy production inequality, in order to study these numerical methods, for dissipation driven models is far from being new. Indeed, there exists a wide literature concerning the design and the analysis of entropy stable finite volume schemes. For instance, let us mention the following (non-exhaustive) list of contributions [7, 13, 15, 16, 17, 18, 22, 24, 30, 34, 35, 36, 51, 66, 54, 70].

Now, in the particular framework of cross-diffusion systems which admit an entropy structure as described in [60, 61], we would like to mention the papers [1, 3] where the authors proposed some convergent two-point flux approximation with upwind mobilities schemes for a seawater intrusion model in an unconfined and confined aquifer. In [5] a first order in space convergent finite volume scheme is considered for the approximation of the cross-diffusion Shigesada-Kawazaki-Teramoto (SKT) model [71]. The results established in this last work have been recently generalized in [64] by employing a two-point flux approximation finite volume scheme with mobilities given by logarithmic mean ensuring the preservation of the entropy structure of the models exhibited in [39]. We also refer to [37] where an entropy preserving finite volume scheme is analyzed for a spinorial matrix drift-diffusion model for semiconductors, and to [49] where two energy stable two-point flux approximation finite volume schemes are studied for a generalized Poisson-Nernst-Planck system accounting for the excess chemical potential of the solvent in multicomponent ionic liquids.

Let us point out that the previous contributions did not deal with the development and the analysis of finite volume schemes for cross-diffusion systems with volume filling constraint. In fact there exists only few papers on this subject. Let us cite [23, 25, 28, 40] where the authors studied, using similar techniques than in this paper, some convergent finite volume schemes for an ion transport, the Maxwell-Stefan, a thin-film solar cells and a biofilm cross-diffusion system with a volume filling constraint. However, in these former works the numerical schemes were designed for particular models. This differs from the present study where we consider, thanks to (10)–(15), a rather general model (including or not cross-diffusion terms). Up to our knowledge, there does not exist, so far, “black-box” results which state the existence of solutions to a finite volume scheme and its convergence for a general class of (cross-diffusion) systems with volume filling constraint. This fact is one of the main source of originality of this paper.

Let us also notice that in the literature other numerical methods have been proposed for the discretization of models (with or without volume filling constraint) admitting an entropy structure. Indeed, some authors designed finite element/finite difference schemes. Let us mention the following list [8, 20, 48, 50, 56, 62, 63] (again far from being exhaustive) of such contributions. Closer to our study we mention [26], where a model admitting a gradient flow structure for a constraint Wasserstein metric, similar to the one presented in Section 2.1, is discretized thanks to a minimizing JKO scheme [12, 59]. An other approach dealing with discontinuous Galerkin method has been used in [21, 72, 73]. Finally, we also refer to [2, 29, 57] for numerical schemes based on the control volume finite element method which preserved the entropy structure of some dissipative models.

3. NUMERICAL SCHEME AND MAIN RESULTS

3.1. Notation and definitions. We present the discretization of the domain $Q_T = \Omega \times (0, T)$. Let $\Omega \subset \mathbb{R}^N$ be a bounded, polygonal (or polyhedral if $N \geq 3$) domain. An admissible mesh of Ω is given by (i) a family \mathcal{T} of open polygonal (or polyhedral) control volumes (also referred as cells in what follows), (ii) a family \mathcal{E} of edges (or faces if $N \geq 3$), and (iii) a family \mathcal{P} of points $(x_K)_{K \in \mathcal{T}}$ associated to the control volumes and satisfying Definition 9.1 in [45]. This definition

implies that the straight line $\overline{x_K x_L}$ between two centers of neighboring cells is orthogonal to the edge $\sigma = K|L$ between two cells. For instance, Voronoï meshes satisfy this condition [45, Example 9.2]. The size of the mesh is denoted by $\Delta x = \max_{K \in \mathcal{T}} \text{diam}(K)$. The family of edges \mathcal{E} can be split into interior edges \mathcal{E}_{int} satisfying $\sigma \in \Omega$ and boundary edges $\sigma \in \mathcal{E}_{\text{ext}}$ satisfying $\sigma \subset \partial\Omega$. For given $K \in \mathcal{T}$, \mathcal{E}_K is the set of edges of K , i.e. $\bigcup_{\sigma \in \mathcal{E}_K} \bar{\sigma} = \partial K$. It splits into $\mathcal{E}_K = \mathcal{E}_{\text{int},K} \cup \mathcal{E}_{\text{ext},K}$, with $\mathcal{E}_{\text{int},K} = \mathcal{E}_K \cap \mathcal{E}_{\text{int}}$ and $\mathcal{E}_{\text{ext},K} = \mathcal{E}_K \cap \mathcal{E}_{\text{ext}}$. For any $\sigma \in \mathcal{E}$, there exists at least one cell $K \in \mathcal{T}$ such that $\sigma \in \mathcal{E}_K$.

We need the following definitions. For $\sigma \in \mathcal{E}$, we introduce the distance

$$d_\sigma = \begin{cases} d(x_K, x_L) & \text{if } \sigma = K|L \in \mathcal{E}_{\text{int},K}, \\ d(x_K, \sigma) & \text{if } \sigma \in \mathcal{E}_{\text{ext},K}, \end{cases}$$

where d is the Euclidean distance in \mathbb{R}^N , and the transmissibility coefficient

$$(22) \quad \tau_\sigma = \frac{m(\sigma)}{d_\sigma},$$

where $m(\sigma)$ denotes the $(N-1)$ -dimensional Lebesgue measure of σ , which is set to 1 if $N=1$. The mesh is assumed to satisfy the following regularity assumption: There exists $\zeta > 0$ such that for all $K \in \mathcal{T}$ and $\sigma \in \mathcal{E}_K$,

$$(23) \quad d(x_K, \sigma) \geq \zeta d_\sigma \quad \text{and} \quad \text{Card}(\mathcal{E}_K) \leq \frac{1}{\zeta}.$$

Let $T > 0$, let $N_T \in \mathbb{N}$ be the number of time steps, and set $\Delta t = T/N_T$ the time step as well as $t_k = k\Delta t$ for $k = 0, \dots, N_T$. We denote by \mathcal{D} the admissible space-time discretization of Q_T composed of an admissible space discretization $(\mathcal{T}, \mathcal{E}, (x_K)_{K \in \mathcal{T}})$ and the values $(\Delta t, N_T)$. Note that the choice of a uniform time discretization is made only to avoid extra notations, and that our study can be extended in a straightforward way to the case of variable time steps, as used in practice for the simulations presented in Section 8.

We also introduce suitable function spaces for the analysis of the numerical scheme. The space of piecewise constant functions is defined by

$$\mathcal{H}_{\mathcal{T}} = \left\{ v : \Omega \rightarrow \mathbb{R} : \exists (v_K)_{K \in \mathcal{T}} \subset \mathbb{R}, v(x) = \sum_{K \in \mathcal{T}} v_K \mathbf{1}_K(x) \right\},$$

where $\mathbf{1}_K$ is the characteristic function on K . Since there is a one to one correspondence between a function $v \in \mathcal{H}_{\mathcal{T}}$ and a vector $(v_K)_{K \in \mathcal{T}}$ in \mathbb{R}^θ (with $\theta = \#\mathcal{T}$) we will make a slight abuse of notation by writing $\mathbf{v} = (v_K)_{K \in \mathcal{T}} \in \mathcal{H}_{\mathcal{T}}$. In order to define a norm on this space, we also introduce the notation

$$v_{K,\sigma} = \begin{cases} v_L & \text{if } \sigma = K|L \in \mathcal{E}_{\text{int},K}, \\ v_K & \text{if } \sigma \in \mathcal{E}_{\text{ext},K}, \end{cases}$$

for $K \in \mathcal{T}$, $\sigma \in \mathcal{E}_K$ and the discrete operators

$$D_{K,\sigma} \mathbf{v} = v_{K,\sigma} - v_K, \quad D_\sigma \mathbf{v} = |D_{K,\sigma} \mathbf{v}| \quad \forall \mathbf{v} \in \mathcal{H}_{\mathcal{T}}.$$

The discrete H^1 seminorm and discrete H^1 norm on $\mathcal{H}_{\mathcal{T}}$ are given by

$$|\mathbf{v}|_{1,2,\mathcal{T}}^2 = \sum_{\sigma \in \mathcal{E}} \tau_\sigma (D_\sigma \mathbf{v})^2, \quad \|\mathbf{v}\|_{1,2,\mathcal{T}}^2 = |\mathbf{v}|_{1,2,\mathcal{T}}^2 + \|\mathbf{v}\|_{0,2,\mathcal{T}}^2,$$

where $\|\mathbf{v}\|_{0,2,\mathcal{T}}$ denotes the $L^2(\Omega)$ norm for $\mathbf{v} \in \mathcal{H}_{\mathcal{T}}$. More precisely, for every $1 \leq p < \infty$ and $\mathbf{v} \in \mathcal{H}_{\mathcal{T}}$, the $L^p(\Omega)$ norm is given by

$$\|\mathbf{v}\|_{0,p,\mathcal{T}} = \left(\sum_{K \in \mathcal{T}} m(K) |v_K|^p \right)^{1/p}.$$

Finally, we introduce the space $\mathcal{H}_{\mathcal{D}}$ of piecewise constant in time functions with values in $\mathcal{H}_{\mathcal{T}}$,

$$\mathcal{H}_{\mathcal{D}} = \left\{ v : \bar{\Omega} \times [0, T] \rightarrow \mathbb{R} : \exists (\mathbf{v}^k)_{k=1, \dots, N_T} \subset \mathcal{H}_{\mathcal{T}}, v(x, t) = \sum_{k=1}^{N_T} \mathbf{v}^k(x) \mathbf{1}_{(t_{k-1}, t_k]}(t) \right\},$$

equipped with the discrete $L^2(0, T; H^1(\Omega))$ norm

$$\left(\sum_{k=1}^{N_T} \Delta t \|\mathbf{v}^k\|_{1,2,\mathcal{T}}^2 \right)^{1/2}.$$

3.2. Numerical scheme. We define now the finite volume scheme for the cross-diffusion model (10)–(15). We first approximate the functions u_i^0 and the external potential fields Ψ_i by

$$(24) \quad u_{i,K}^0 = \frac{1}{m(K)} \int_K u_i^0(x) dx, \quad \Psi_{i,K} = \frac{1}{m(K)} \int_K \Psi_i dx, \quad \text{for } K \in \mathcal{T}, i = 0, \dots, n.$$

Remark 5. With the above definition (24), $(\mathbf{u}_K^0)_{K \in \mathcal{T}} \in (\mathcal{H}_{\mathcal{T}})^{n+1}$ belongs to \mathcal{M} if \mathbf{u}^0 does. From a practical perspective, the initial data is rather approximated by

$$u_{i,K}^0 = u_i^0(x_K), \quad K \in \mathcal{T}, i = 0, \dots, n,$$

provided \mathbf{u}^0 is continuous so that the previous formula has a sense. This definition possibly introduces some shift on the mean values, hence \mathbf{u}^0 and $(\mathbf{u}_K^0)_{K \in \mathcal{T}}$ do no longer belong to the same \mathcal{M} (cf. Remark 1). This shift tends to 0 for continuous \mathbf{u}^0 as the mesh size Δx tends to 0, and it can be quantified under further regularity assumptions on \mathbf{u}^0 , typically Lipschitz continuity. To ease the presentation, we assume that the discretization of the initial data and of the external potentials does not introduce mass shifts and that formulas (24) are satisfied.

Similarly, a classical approach for discretizing Ψ is to set

$$\Psi_{i,K}^0 = \Psi_i^0(x_K), \quad K \in \mathcal{T}, i = 0, \dots, n.$$

Provided Ψ is Lipschitz continuous, the estimate $|\Psi_i|_{1,2,\mathcal{T}} \leq C \|\nabla \Psi_i\|_{\infty}$ can be readily established and allows to carry out the convergence proof. However, we stick to the theoretical definition (24) of $\Psi_{i,K}$ since it allows for merely $H^1(\Omega)$ and possibly not continuous functions Ψ_i .

For $k \geq 1$, let $\mathbf{u}_K^{k-1} = (u_{1,K}^{k-1}, \dots, u_{n,K}^{k-1})$ be given for each $K \in \mathcal{T}$. Then, the vector $(\mathbf{u}^k, \boldsymbol{\pi}^k)$ is solution to the following nonlinear system

$$(25) \quad m(K) \frac{u_{i,K}^k - u_{i,K}^{k-1}}{\Delta t} + \sum_{\sigma \in \mathcal{E}_K} \mathcal{F}_{i,K,\sigma}^k = 0, \quad \forall K \in \mathcal{T}, i = 1, \dots, n,$$

where for all $K \in \mathcal{T}$ and $\sigma \in \mathcal{E}_K$ the fluxes $\mathcal{F}_{i,K,\sigma}^k$ are given by

$$(26) \quad \mathcal{F}_{i,K,\sigma}^k = -\tau_{\sigma} d_i \left[\alpha_i D_{K,\sigma} \mathbf{u}_i^k + u_{i,\sigma}^k (D_{K,\sigma} b_i(\mathbf{u}^k) + D_{K,\sigma} \Psi_i) + u_{i,\sigma}^{k,\pi} D_{K,\sigma} \boldsymbol{\pi}^k \right], \quad i = 1, \dots, n,$$

and

$$(27) \quad - \sum_{\sigma \in \mathcal{E}_K} \tau_{\sigma} \left(\sum_{i=0}^n d_i u_{i,\sigma}^{k,\pi} \right) D_{K,\sigma} \boldsymbol{\pi}^k + \sum_{\sigma \in \mathcal{E}_K} \mathcal{G}_{K,\sigma}^k = 0, \quad \forall K \in \mathcal{T},$$

with

$$(28) \quad \mathcal{G}_{K,\sigma}^k = - \sum_{i=0}^n \tau_{\sigma} d_i (\alpha_i D_{K,\sigma} \mathbf{u}_i^k + u_{i,\sigma}^k (D_{K,\sigma} b_i(\mathbf{u}^k) + D_{K,\sigma} \Psi_i)), \quad \forall K \in \mathcal{T}, \sigma \in \mathcal{E}_K.$$

where $u_{0,K}^k$ is discretized thanks to the constraint (12) by

$$(29) \quad u_{0,K}^k = 1 - \sum_{i=1}^n u_{i,K}^k, \quad \forall K \in \mathcal{T}.$$

Furthermore, in the relations (26)–(28) the components of the vector $\mathbf{u}_{\sigma}^k = (u_{i,\sigma}^k)_{0 \leq i \leq n}$ are defined as

$$(30) \quad u_{i,\sigma}^k = \begin{cases} u_{i,K}^k & \text{if } u_{i,K}^k = u_{i,K,\sigma}^k > 0, \\ 0 & \text{if } \min(u_{i,K}^k, u_{i,K,\sigma}^k) \leq 0, \\ \frac{D_{K,\sigma} \mathbf{u}_i^k}{D_{K,\sigma} \log(\mathbf{u}_i^k)} & \text{otherwise,} \end{cases}$$

while $u_{i,\sigma}^{k,\pi}$ is an upwind mobility given by

$$(31) \quad u_{i,\sigma}^{k,\pi} = \begin{cases} (u_{i,K}^k)^+ = \max(0, u_{i,K}^k) & \text{if } \pi_K^k \geq \pi_{K,\sigma}^k, \\ (u_{i,K,\sigma}^k)^+ = \max(0, u_{i,K,\sigma}^k) & \text{if } \pi_{K,\sigma}^k > \pi_K^k, \end{cases}$$

for every $\sigma \in \mathcal{E}_{\text{int},K}$ and $i = 0, \dots, n$. As it will appear later on in Lemma 11, the introduction of the positive part in the definition of $u_{i,\sigma}^{k,\pi}$ is purely technical and can be removed in practice. Finally, we approximate (13) in a straightforward way by imposing

$$(32) \quad \sum_{K \in \mathcal{T}} m(K) \pi_K^k = 0.$$

Let us notice that we design our scheme in such a way that $u_{0,K}^k$ also satisfies an equation of the form (25) for every $K \in \mathcal{T}$ and $k \geq 1$. Indeed, summing over $i = 1, \dots, n$ equation (25), using (27) and the constraint (29) we obtain that $u_{0,K}^k$ is solution to

$$(33) \quad m(K) \frac{u_{0,K}^k - u_{0,K}^{k-1}}{\Delta t} + \sum_{\sigma \in \mathcal{E}_K} \mathcal{F}_{0,K,\sigma}^k = 0, \quad \forall K \in \mathcal{T},$$

where $\mathcal{F}_{0,K,\sigma}^k$ is defined by (26) with $i = 0$. Thus the scheme (24)–(32) is equivalent to (24)–(26) and (29)–(33). In the sequel depending on our needs we will consider either of them.

Remark 6. *A more classical approach would consist in solving the problem in its form (1)–(5), i.e. in terms of discrete unknowns approximating $(\mathbf{u}, \boldsymbol{\mu})$ rather than $(\mathbf{u}, \boldsymbol{\pi})$. A typical example of such approaches is the widely used upstream mobility scheme (see for instance [46, 26]) or smoothed counterparts taking advantage of the self-diffusion (see for instance [19, 49]). Such schemes would naturally be energy diminishing in accordance with the gradient flow structure highlighted in Section 2.1. However, their convergence analysis can only be carried out in particular situations, like when $\mu_i - \mu_0$ is a function of u_i but not of u_j , $j \neq i$ in (3a) for the upstream mobility scheme (as it is always the case in the two-phase case $n = 1$), cf. [26]. Putting the pressure π apart in the fluxes (26) allows to get enough control on it to state the convergence of the scheme, cf. Theorem 8 below, but at the price of (slightly) jeopardizing the energy dissipative character of the problem.*

3.3. Main results. Let us first gather our hypotheses.

- (H1) Domain: $\Omega \subset \mathbb{R}^N$ is a connected bounded, polygonal (or polyhedral if $N \geq 3$) domain.
- (H2) Discretization: \mathcal{D} is an admissible discretization of Q_T satisfying (23).
- (H3) Initial data: $\mathbf{u}^0 \in \mathcal{M}$.
- (H4) External potentials: $\Psi_i \in H^1(\Omega)$ for $i = 0, \dots, n$.
- (H5) Function B : B is a $C^{1,1}$ convex function.
- (H6) Given constants: $d_i > 0$ and $\alpha_i > 0$ for $i = 0, \dots, n$.

For any nonnegative $\mathbf{v} = (\mathbf{v}_0, \dots, \mathbf{v}_n) \in \mathcal{H}_{\mathcal{T}}^{n+1}$, we introduce the discrete counterpart of the entropy functional H as follows

$$H(\mathbf{v}) = \sum_{i=0}^n \frac{1}{d_i} \sum_{K \in \mathcal{T}} m(K) (v_{i,K} (\log(v_{i,K}) - 1) + 1),$$

where the Boltzmann entropy density $v \log v$ is extended by continuity by 0 whenever $v = 0$ for $K \in \mathcal{T}$ and $i = 0, \dots, n$.

Our first main result deals with the existence of nonnegative solutions to scheme (24)–(32) at each time step.

Theorem 7 (Existence of discrete solutions). *Let Hypotheses (H1)–(H6) hold. Then there exists (at least) one solution $(\mathbf{u}^k, \boldsymbol{\pi}^k)_{k \geq 1}$ to scheme (24)–(32). Moreover, this solution satisfies the following properties:*

(i) *Mass conservation:*

$$\sum_{K \in \mathcal{T}} m(K) u_{i,K}^k = \int_{\Omega} u_i^0 dx, \quad \forall k \geq 0, i = 1, \dots, n.$$

(ii) *L^∞ bounds:* $0 < u_{i,K}^k < 1$ for all $K \in \mathcal{T}$, $k \geq 1$ and $i = 0, \dots, n$.

(iii) *Volume-filling constraint:* $\sum_{i=0}^n u_{i,K}^k = 1$ for all $K \in \mathcal{T}$ and $k \geq 1$.

(iv) *Entropy production estimate:* for all $k \geq 1$ there exists a constant $C_\zeta > 0$ only depending on ζ such that

$$\begin{aligned} \frac{H(\mathbf{u}^k) - H(\mathbf{u}^{k-1})}{\Delta t} + \sum_{i=0}^n 2\alpha_i \sum_{\sigma \in \mathcal{E}} \tau_\sigma \left(D_\sigma \sqrt{u_i^k} \right)^2 \\ + \sum_{\sigma \in \mathcal{E}} \tau_\sigma \langle D_{K,\sigma} \mathbf{u}^k, D_{K,\sigma} \mathbf{b}(\mathbf{u}^k) \rangle \leq \sum_{i=0}^n \frac{C_\zeta}{2\alpha_i} \|\Psi_i\|_{H^1(\Omega)}^2. \end{aligned}$$

Note that, due to the convexity of B , each term $\langle D_{K,\sigma} \mathbf{u}^k, D_{K,\sigma} \mathbf{b}(\mathbf{u}^k) \rangle$ is nonnegative. The proof of Theorem 7 is based on a topological degree argument. In this purpose, we establish in Section 4 some a priori estimates and we prove Theorem 7 in Section 5.

Now, thanks to Theorem 7 we deduce the existence of a solution $(\mathbf{u}^k, \boldsymbol{\pi}^k)_{k \geq 1}$ for each Δx and Δt . In the sequel we will prove that when $\Delta x \rightarrow 0$ and $\Delta t \rightarrow 0$ then there exists a subsequence of solutions to the scheme (24)–(32) which converges towards a weak solution to (10)–(15) in the sense of Definition 1. In order to state precisely our convergence result, we need some notation. For $K \in \mathcal{T}$ and $\sigma \in \mathcal{E}_K$, we define the cell $T_{K,\sigma}$ of the so-called dual mesh:

- If $\sigma = K|L \in \mathcal{E}_{\text{int},K}$, then T_σ is that cell (“diamond”), the vertices of which are given by x_K, x_L , and the end points of the edge σ .
- If $\sigma \in \mathcal{E}_{\text{ext},K}$, then T_σ is that cell (“half-diamond”), the vertices of which are given by x_K and the end points of the edge σ .

The cells $(T_\sigma)_{\sigma \in \mathcal{E}}$ define a partition of Ω up to a negligible set. It follows from the property that the straight line $\overline{x_K x_L}$ between two neighboring centers of cells is orthogonal to the edge $\sigma = K|L$ that

$$m(\sigma) d_\sigma = N m(T_\sigma) \quad \text{for } \sigma \in \mathcal{E}.$$

Following [38, 44], the approximate gradient of $v \in \mathcal{H}_{\mathcal{D}}$ is defined by

$$\nabla^{\mathcal{D}} v(x, t) = \frac{m(\sigma)}{m(T_\sigma)} (D_{K,\sigma} v^k) \nu_{K,\sigma} \quad \text{for } x \in T_{K,\sigma}, t \in (t_{k-1}, t_k],$$

where $\nu_{K,\sigma}$ is the unit vector that is normal to σ and points outwards of K .

We introduce a family $(\mathcal{D}_m)_{m \in \mathbb{N}}$ of admissible space-time discretizations of Q_T indexed by the size $\eta_m = \max\{\Delta x_m, \Delta t_m\}$ of the mesh, satisfying $\eta_m \rightarrow 0$ as $m \rightarrow \infty$. We denote by \mathcal{T}_m the corresponding meshes of Ω and by Δt_m the corresponding time steps. Finally, for every $m \in \mathbb{N}$ we set $(\mathbf{u}_m, \boldsymbol{\pi}_m) = (u_{1,m}, \dots, u_{n,m}, \pi_m)_{m \in \mathbb{N}} \in \mathcal{H}_{\mathcal{D}_m}$ and $\nabla^m = \nabla^{\mathcal{D}_m}$.

Theorem 8 (Convergence of the scheme). *Let the assumptions of Theorem 7 hold, let $(\mathcal{D}_m)_{m \in \mathbb{N}}$ be a family of admissible space-time discretizations of Q_T satisfying (23) uniformly in $m \in \mathbb{N}$. Let $(\mathbf{u}_m, \boldsymbol{\pi}_m)_{m \in \mathbb{N}}$ be a family of finite volume solutions to (24)–(32) constructed in Theorem 7. Then there exists a weak solution $(\mathbf{u}, \boldsymbol{\pi})$ to (10)–(15) in the sense of Definition 1 such that, for $i = 0, \dots, n$, the following convergences hold up to a subsequence:*

$$\begin{aligned} u_{i,m} &\rightarrow u_i && \text{strongly in } L^2(Q_T) \quad \text{as } m \rightarrow \infty, \\ \nabla^m u_{i,m} &\rightharpoonup \nabla u_i && \text{weakly in } L^2(Q_T)^N \quad \text{as } m \rightarrow \infty, \\ \pi_m &\rightharpoonup \pi && \text{weakly in } L^2(Q_T) \quad \text{as } m \rightarrow \infty, \\ \nabla^m \pi_m &\rightharpoonup \nabla \pi && \text{weakly in } L^2(Q_T)^N \quad \text{as } m \rightarrow \infty. \end{aligned}$$

In order to prove Theorem 8 we establish in Section 6 some uniform w.r.t. Δx and Δt estimates. These estimates will allow us to apply a compactness result obtained in [6] and prove the existence

of (\mathbf{u}, π) . Finally, following [38], we will identify the functions (\mathbf{u}, π) as a weak solution to (10)–(15).

The remaining of the paper is organized as follows. In Section 4 we establish some a priori estimates needed for the proof of our existence result. Then, we prove Theorem 7 in Section 5. In Section 6 we show some uniform estimates w.r.t. Δx and Δt , while we prove Theorem 8 in Section 7. Section 8 is devoted to numerical experiments in one and two space dimensions for some models which enter in the framework described by (10)–(15). Eventually, Section 9 gathers some concluding remarks on possible extensions of the present work.

4. A PRIORI ESTIMATES

One of our main objectives in this section is to show the following entropy dissipation inequality:

Proposition 9. *Let the assumptions (H1)–(H6) hold and let $(\mathbf{u}^k, \boldsymbol{\pi}^k)_{k \geq 1}$ be a solution to (24)–(26) and (29)–(33). Then for all $k \geq 1$ there exists a constant $C_\zeta > 0$ only depending on ζ such that*

$$(34) \quad \frac{H(\mathbf{u}^k) - H(\mathbf{u}^{k-1})}{\Delta t} + \sum_{i=0}^n 2\alpha_i \sum_{\sigma \in \mathcal{E}} \tau_\sigma \left(D_\sigma \sqrt{u_i^k} \right)^2 \\ + \sum_{\sigma \in \mathcal{E}} \tau_\sigma \langle D_{K,\sigma} \mathbf{u}^k, D_{K,\sigma} \mathbf{b}(\mathbf{u}^k) \rangle \leq \sum_{i=0}^n \frac{C_\zeta}{2\alpha_i} \|\Psi_i\|_{H^1(\Omega)}^2.$$

The proof of Proposition 9 is a transcription to the discrete setting of the proof of Lemma 3. To prove that (34) holds true we will use $\log(u_{i,K}^k)$ as “test function” in the equations (25) and (33). Then, we first need to show that $u_{i,K}^k > 0$ for all $K \in \mathcal{T}$ and $i = 0, \dots, n$. In this purpose, let us first prove that

$$\sum_{K \in \mathcal{T}} m(K) u_{i,K}^k > 0, \quad i = 0, \dots, n.$$

Lemma 10. *Let the assumptions of Proposition 9 hold. Then, the solutions to (24)–(26) and (29)–(33) satisfy*

$$(35) \quad \sum_{K \in \mathcal{T}} m(K) u_{i,K}^k = \int_{\Omega} u_i^0 dx > 0, \quad \forall k \geq 0, i = 0, \dots, n.$$

Proof. We simply sum equations (25) and (33) over K and thanks to the conservativity of the scheme we deduce that it holds

$$\sum_{K \in \mathcal{T}} m(K) u_{i,K}^k = \sum_{K \in \mathcal{T}} m(K) u_{i,K}^{k-1}, \quad \forall k \geq 1, i = 0, \dots, n.$$

We conclude the proof by induction and thanks to definition (24) of $u_{i,K}^0$ for all $K \in \mathcal{T}$ and assumption (H3) on the initial data. \square

Let us now show that the discrete concentrations are positive.

Lemma 11. *Let the assumptions of Proposition 9 hold. Then the solutions to (24)–(26) and (29)–(33) satisfy*

$$(36) \quad 0 < u_{i,K}^k < 1, \quad \forall K \in \mathcal{T}, k \geq 1, i = 0, \dots, n.$$

Proof. Let $i \in \{0, \dots, n\}$ be fixed. Assume by induction that for $k \geq 1$ we have $0 \leq u_{i,K}^{k-1}$ and $\sum_K m(K) u_{i,K}^{k-1} > 0$ (we recall that this property has been already established for every $k \geq 1$ in Lemma 10). Let us now define the subset $\mathcal{K}_i \subseteq \mathcal{T}$ by

$$(37) \quad \mathcal{K}_i^k = \{K \in \mathcal{T} \mid u_{i,K}^k \leq 0\}.$$

Our goal is to show that $\mathcal{K}_i^k = \emptyset$. In this purpose, we proceed by contradiction by assuming that $\mathcal{K}_i^k \neq \emptyset$.

The case $\mathcal{K}_i^k = \mathcal{T}$ easily leads to a contradiction with $\sum_{K \in \mathcal{T}} m(K) u_{i,K}^k > 0$ obtained in Lemma 10. We restrict then our attention to the situation where $\mathcal{K}_i^k \subsetneq \mathcal{T}$.

Let us introduce the subset

$$\mathcal{E}_i^k = \{\sigma \in \mathcal{E}_K \mid K \in \mathcal{K}_i^k\} \subset \mathcal{E}.$$

Such edges $\sigma \in \mathcal{E}_i^k$ are of three different types:

- (i) $\sigma \in \mathcal{E}_{i,\text{ext}}^k = \mathcal{E}_i^k \cap \mathcal{E}_{\text{ext}}$;
- (ii) $\sigma = K|L \in \mathcal{E}_{i,-}^k$ if both K and L belong to \mathcal{K}_i^k ;
- (iii) $\sigma = K|L \in \mathcal{E}_{i,+}^k$ if $K \in \mathcal{K}_i^k$ whereas $L \notin \mathcal{K}_i^k$.

Since Ω is connected, and since $\mathcal{K}_i^k \subsetneq \mathcal{T}$, then the subset $\mathcal{E}_{i,+}^k$ of \mathcal{E}_i^k is not empty. It follows from the definition (30) that $u_{i,\sigma}^k = 0$ for any $\sigma \in \mathcal{E}_i^k$, and that $u_{i,\sigma}^{k,\pi} = 0$ if $\sigma \in \mathcal{E}_{i,-}^k$ or if $\sigma \in \mathcal{E}_{i,+}^k$ and $\pi_K^k \geq \pi_{K,\sigma}^k$ due to the positive part in (31). We set now $\tilde{\mathcal{K}}_i^k$ a connected component of \mathcal{K}_i^k . Hence, assuming that $(\pi_K^k)_{K \in \mathcal{T}}$ and $(u_{i,K}^k)_{K \in \mathcal{T} \setminus \mathcal{K}_i^k}$ are given, the vector $\tilde{\mathbf{u}}_i^k = (u_{i,K}^k)_{K \in \tilde{\mathcal{K}}_i^k}$ solves the linear system: $\forall K \in \tilde{\mathcal{K}}_i^k$,

$$(38) \quad u_{i,K}^k m(K) + \Delta t d_i \alpha_i \left[\sum_{\sigma \in \mathcal{E}_{i,-}^k} \tau_\sigma (u_{i,K}^k - u_{i,K,\sigma}^k) + \sum_{\sigma \in \mathcal{E}_{i,+}^k} \tau_\sigma u_{i,K}^k \right] \\ = u_{i,K}^{k-1} m(K) + \Delta t \sum_{\sigma \in \mathcal{E}_{i,+}^k} \tau_\sigma d_i \left[\alpha_i u_{i,K,\sigma}^k + u_{i,K,\sigma}^k (\pi_{K,\sigma}^k - \pi_K^k)^+ \right].$$

The above right-hand side is non-negative, and even positive if K is such that $\mathcal{E}_{i,+}^k \neq \emptyset$. Let us first consider the case where $\tilde{\mathcal{K}}_i^k = \{K\}$ for some $K \in \mathcal{K}_i^k$. Then, the equality (38) becomes

$$u_{i,K}^k m(K) + \Delta t d_i \alpha_i \sum_{\sigma \in \mathcal{E}_{i,+}^k} \tau_\sigma u_{i,K}^k \\ = u_{i,K}^{k-1} m(K) + \Delta t \sum_{\sigma \in \mathcal{E}_{i,+}^k} \tau_\sigma d_i \left[\alpha_i u_{i,K,\sigma}^k + u_{i,K,\sigma}^k (\pi_{K,\sigma}^k - \pi_K^k)^+ \right].$$

This yields a contradiction since the left-hand side is non-positive while the right-hand side is positive. Now, if the subset $\tilde{\mathcal{K}}_i^k$ is not reduced to a single cell, we set

$$u_{i,K}^k := \min_{L \in \tilde{\mathcal{K}}_i^k} u_{i,L}^k.$$

If $u_{i,K}^k = 0$ then $u_{i,L}^k = 0$ for every $L \in \tilde{\mathcal{K}}_i^k \setminus \{K\}$ and for at least one cell $L_0 \in \tilde{\mathcal{K}}_i^k$ there exists $\sigma \in \mathcal{E}_{L_0,\text{int}} \cap \mathcal{E}_{i,+}^k$ (otherwise $\tilde{\mathcal{K}}_i^k = \mathcal{T}$). In particular, for this cell L_0 , the equation (38) provides a contradiction since the left-hand side is equal to 0 while the right-hand side is positive. Finally, the case $u_{i,K}^k < 0$ also leads to a contradiction since the left-hand side is negative while the right-hand side is non-negative. Arguing similarly for each connected component of \mathcal{K}_i^k we conclude that $\mathcal{K}_i^k = \emptyset$.

Finally, thanks to the constraint (29) we conclude that $u_{i,K}^k < 1$ for every $K \in \mathcal{T}$ and $i = 0, \dots, n$. \square

Thanks to the previous result, we can remove the positive parts in (31), and the values $\log(u_{i,K}^k)$ have a meaning for all $K \in \mathcal{T}$ and $i = 0, \dots, n$. Therefore we are in position to prove Proposition 9.

Proof of Proposition 9. We multiply (25) by $\log(u_{i,K}^k)/d_i$, equation (33) by $\log(u_{0,K}^k)/d_0$, we sum over $K \in \mathcal{T}$ and $i = 0, \dots, n$, and after some discrete integration by parts we have

$$J_1 + J_2 = 0,$$

with

$$J_1 = \sum_{i=0}^n \sum_{K \in \mathcal{T}} m(K) \frac{(u_{i,K}^k - u_{i,K}^{k-1})}{\Delta t} \frac{1}{d_i} \log(u_{i,K}^k),$$

$$J_2 = \sum_{i=0}^n \sum_{\substack{\sigma \in \mathcal{E}_{\text{int}} \\ \sigma = K|L}} \tau_\sigma \left(\alpha_i D_{K,\sigma} \mathbf{u}_i^k + u_{i,\sigma}^k \left(D_{K,\sigma} b_i(\mathbf{u}^k) + D_{K,\sigma} \Psi_i \right) + u_{i,\sigma}^{k,\pi} D_{K,\sigma} \boldsymbol{\pi}^k \right) D_{K,\sigma} \log(\mathbf{u}_i^k).$$

For J_1 , using the convexity of $x \mapsto x(\log(x) - 1) + 1$ allows us to conclude that

$$(39) \quad J_1 \geq \frac{H(\mathbf{u}^k) - H(\mathbf{u}^{k-1})}{\Delta t}.$$

Now, applying the Cauchy-Schwarz inequality, we first notice that it holds

$$(\sqrt{a} - \sqrt{b})^2 = \left(\int_b^a \frac{dx}{2\sqrt{x}} \right)^2 \leq \int_b^a \frac{dx}{4} \int_b^a \frac{dx}{x} = \frac{1}{4} (a - b)(\log(a) - \log(b)) \quad \forall a, b \in (0, \infty).$$

Then, thanks to this inequality and the definition (30) of $u_{i,\sigma}^k$ we deduce that

$$\begin{aligned} J_2 \geq \sum_{i=0}^n \sum_{\substack{\sigma \in \mathcal{E}_{\text{int}} \\ \sigma = K|L}} \tau_\sigma \left(4\alpha_i \left(D_\sigma \sqrt{\mathbf{u}_i^k} \right)^2 + D_{K,\sigma} \mathbf{u}_i^k \left(D_{K,\sigma} b_i(\mathbf{u}^k) + D_{K,\sigma} \Psi_i \right) \right) \\ + \sum_{i=0}^n \sum_{\substack{\sigma \in \mathcal{E}_{\text{int}} \\ \sigma = K|L}} \tau_\sigma D_{K,\sigma} \boldsymbol{\pi}^k u_{i,\sigma}^{k,\pi} D_{K,\sigma} \log(\mathbf{u}_i^k). \end{aligned}$$

Besides using the definition (31) of $u_{i,\sigma}^{k,\pi}$ and the inequality

$$a(\log a - \log b) \geq a - b \geq b(\log b - \log a) \quad \forall a, b \in [0, \infty),$$

which follows from the convexity of the exponential function we get

$$J_2 \geq \sum_{i=0}^n \sum_{\substack{\sigma \in \mathcal{E}_{\text{int}} \\ \sigma = K|L}} \tau_\sigma \left(4\alpha_i \left(D_\sigma \sqrt{\mathbf{u}_i^k} \right)^2 + D_{K,\sigma} \mathbf{u}_i^k \left(D_{K,\sigma} b_i(\mathbf{u}^k) + D_{K,\sigma} \Psi_i \right) + D_{K,\sigma} \boldsymbol{\pi}^k D_{K,\sigma} \mathbf{u}_i^k \right).$$

For the last term, thanks to the constraint (29), we notice that the following relation

$$\sum_{\substack{\sigma \in \mathcal{E}_{\text{int}} \\ \sigma = K|L}} \tau_\sigma \left(\sum_{i=1}^n D_{K,\sigma} \mathbf{u}_i^k + D_{K,\sigma} \mathbf{u}_0^k \right) D_{K,\sigma} \boldsymbol{\pi}^k = 0,$$

holds. Thus

$$(40) \quad J_2 \geq \sum_{i=0}^n \sum_{\substack{\sigma \in \mathcal{E}_{\text{int}} \\ \sigma = K|L}} \tau_\sigma \left(4\alpha_i \left(D_\sigma \sqrt{\mathbf{u}_i^k} \right)^2 + D_{K,\sigma} \mathbf{u}_i^k \left(D_{K,\sigma} b_i(\mathbf{u}^k) + D_{K,\sigma} \Psi_i \right) \right).$$

Gathering the inequalities (39) and (40), we conclude that

$$\begin{aligned} \frac{H(\mathbf{u}^k) - H(\mathbf{u}^{k-1})}{\Delta t} + \sum_{i=0}^n 4\alpha_i \sum_{\sigma \in \mathcal{E}} \tau_\sigma \left(D_\sigma \sqrt{\mathbf{u}_i^k} \right)^2 \\ + \sum_{\sigma \in \mathcal{E}} \tau_\sigma \langle D_{K,\sigma} \mathbf{u}^k, D_{K,\sigma} \mathbf{b}(\mathbf{u}^k) \rangle \leq \sum_{\sigma \in \mathcal{E}} \tau_\sigma \langle D_\sigma \Psi, D_\sigma \mathbf{u}^k \rangle. \end{aligned}$$

Finally, we use that $2D_\sigma \sqrt{\mathbf{u}_i^k} \geq D_\sigma \mathbf{u}_i^k$ since $0 \leq u_{i,K}^k \leq 1$ to control the right-hand side of the above inequality, leading to

$$\begin{aligned} \frac{H(\mathbf{u}^k) - H(\mathbf{u}^{k-1})}{\Delta t} + \sum_{i=0}^n 2\alpha_i \sum_{\sigma \in \mathcal{E}} \tau_\sigma \left(D_\sigma \sqrt{\mathbf{u}_i^k} \right)^2 \\ + \sum_{\sigma \in \mathcal{E}} \tau_\sigma \langle D_{K,\sigma} \mathbf{u}^k, D_{K,\sigma} \mathbf{b}(\mathbf{u}^k) \rangle \leq \sum_{i=0}^n \frac{1}{2\alpha_i} \sum_{\sigma \in \mathcal{E}} \tau_\sigma (D_\sigma \Psi_i)^2. \end{aligned}$$

It remains to apply [45, Lemma 9.4] in order to complete the proof of Proposition 9. \square

Let us now establish the discrete counterpart of Lemma 4.

Proposition 12. *Let the assumptions (H1)–(H6) hold and let $(\mathbf{u}^k, \boldsymbol{\pi}^k)_{k \geq 1}$ be a solution to (24)–(26) and (29)–(33). Then, for all $k \geq 1$ we have*

$$(41) \quad |\boldsymbol{\pi}^k|_{1,2,\mathcal{T}} \leq \frac{d^*}{d_*} \sum_{i=0}^n \left(\alpha_i |\mathbf{u}_i^k|_{1,2,\mathcal{T}} + \|Db_i\|_\infty \sum_{j=0}^n |\mathbf{u}_j^k|_{1,2,\mathcal{T}} + |\boldsymbol{\Psi}_i|_{1,2,\mathcal{T}} \right),$$

with $d_* = \min_{i=0,\dots,n} d_i$ and $d^* = \max_{i=0,\dots,n} d_i$.

Proof. In this purpose we multiply (27) by π_K^k , we sum over K and after some integration by parts we have

$$\sum_{\substack{\sigma \in \mathcal{E}_{\text{int}} \\ \sigma = K|L}} \left(\sum_{i=0}^n d_i u_{i,\sigma}^{k,\pi} \right) (\mathbb{D}_\sigma \boldsymbol{\pi}^k)^2 = J_3,$$

where

$$J_3 = - \sum_{i=0}^n \sum_{\substack{\sigma \in \mathcal{E}_{\text{int}} \\ \sigma = K|L}} \tau_\sigma d_i (\alpha_i \mathbb{D}_{K,\sigma} \mathbf{u}_i^k + u_{i,\sigma}^{k,\pi} (\mathbb{D}_{K,\sigma} b_i(\mathbf{u}^k) + \mathbb{D}_{K,\sigma} \boldsymbol{\Psi}_{i,K})) \mathbb{D}_{K,\sigma} \boldsymbol{\pi}^k.$$

Now thanks to Lemma 11 we notice that for every $K \in \mathcal{T}$ we have $0 < u_{i,K}^k < 1$ for $i = 0, \dots, n$. Then for $i = 0, \dots, n$ and $\sigma \in \mathcal{E}_{\text{int},K}$ we obtain by definition (30)

$$0 < \min(u_{i,K}^k, u_{i,K,\sigma}^k) \leq u_{i,\sigma}^{k,\pi} \leq \max(u_{i,K}^k, u_{i,K,\sigma}^k) < 1.$$

Hence, using these L^∞ bounds, the Lipschitz regularity of the functions b_i for $i = 0, \dots, n$ and the Cauchy-Schwarz inequality we find that

$$(42) \quad |J_3| \leq \sum_{i=0}^n d_i \left(\alpha_i |\mathbf{u}_i^k|_{1,2,\mathcal{T}} + \|Db_i\|_\infty \sum_{j=0}^n |\mathbf{u}_j^k|_{1,2,\mathcal{T}} + |\boldsymbol{\Psi}_i|_{1,2,\mathcal{T}} \right) |\boldsymbol{\pi}^k|_{1,2,\mathcal{T}}.$$

Let us notice that depending on the sign of $\mathbb{D}_{K,\sigma} \boldsymbol{\pi}^k$ either $\sum_{i=0}^n u_{i,\sigma}^{k,\pi} = \sum_{i=0}^n u_{i,K}^k = 1$ or $\sum_{i=0}^n u_{i,\sigma}^{k,\pi} = \sum_{i=0}^n u_{i,K,\sigma}^k = 1$ which implies

$$(43) \quad \sum_{\substack{\sigma \in \mathcal{E}_{\text{int}} \\ \sigma = K|L}} \left(\sum_{i=0}^n d_i u_{i,\sigma}^{k,\pi} \right) (\mathbb{D}_\sigma \boldsymbol{\pi}^k)^2 \geq d_* |\boldsymbol{\pi}^k|_{1,2,\mathcal{T}}^2,$$

where $d_* = \min_{i=0,\dots,n} d_i$. Thus, collecting (42) and (43) we obtain

$$|\boldsymbol{\pi}^k|_{1,2,\mathcal{T}} \leq \sum_{i=0}^n \frac{d_i}{d_*} \left(\alpha_i |\mathbf{u}_i^k|_{1,2,\mathcal{T}} + \|Db_i\|_\infty \sum_{j=0}^n |\mathbf{u}_j^k|_{1,2,\mathcal{T}} + |\boldsymbol{\Psi}_i|_{1,2,\mathcal{T}} \right).$$

This concludes the proof of Proposition 12. \square

5. PROOF OF THEOREM 7

First, we notice that the scheme (24)–(26) and (29)–(33), denoted (S) in the sequel, is an overdetermined system of equations, $(n+2)\theta + 1$ equations for $(n+2)\theta$ unknowns with $\theta = \#\mathcal{T}$. However, for $k \geq 1$ and $u^{k-1} \in \mathcal{M}$ given, consider the set of equations

$$(44) \quad m(K) \frac{u_{i,K}^k - u_{i,K}^{k-1}}{\Delta t} + \sum_{\sigma \in \mathcal{E}_K} \mathcal{F}_{i,K,\sigma}^k = 0, \quad \forall K \in \mathcal{T}, i = 0, \dots, n,$$

where for all $K \in \mathcal{T}$ and $\sigma \in \mathcal{E}_K$ the fluxes $\mathcal{F}_{i,K,\sigma}^k$ are given by (26). Now we choose arbitrarily $K_0 \in \mathcal{T}$ and we impose the constraint

$$(45) \quad \sum_{i=0}^n u_{i,K}^k = 1, \quad \forall K \in \mathcal{T} \setminus \{K_0\},$$

and the zero-mean constraint

$$(46) \quad \sum_{K \in \mathcal{T}} m(K) \pi_K^k = 0.$$

Hence the system (26), (44)–(46), simply denoted by (44)–(46) in the sequel, has $(n+2)\theta$ equations for $(n+2)\theta$ unknowns. Moreover, we claim that (S) and (44)–(46) are equivalent. Indeed, it is clear that a solution to (S) is also a solution to (44)–(46). Besides, we have the following result:

Lemma 13. *For $k \geq 1$, let $\mathbf{u}^{k-1} \in \mathcal{M}$ be given. Then every solutions to (44)–(46) is also a solution to (S).*

Proof. In this purpose we need to show that $\sum_{i=0}^n u_{i,K_0}^k = 1$. We first notice, thanks to equation (44) that it holds

$$\sum_{K \in \mathcal{T}} m(K) u_{i,K}^k = \sum_{K \in \mathcal{T}} m(K) u_{i,K}^{k-1} \quad i = 0, \dots, n.$$

Then

$$\sum_{i=0}^n \sum_{K \in \mathcal{T}} m(K) u_{i,K}^k = \sum_{i=0}^n \sum_{K \in \mathcal{T}} m(K) u_{i,K}^{k-1} = m(\Omega),$$

since $\mathbf{u}^{k-1} \in \mathcal{M}$. Moreover, the constraint (45) yields

$$\sum_{K \in \mathcal{T}} \sum_{i=0}^n m(K) u_{i,K}^k = \sum_{i=0}^n m(K_0) u_{i,K_0}^k + m(\Omega \setminus K_0).$$

Therefore we deduce that

$$m(K_0) \left(\sum_{i=0}^n u_{i,K_0}^k - 1 \right) = 0,$$

which implies the result. \square

Hence the systems are equivalent and in the remainder of this section we will focus on (44)–(46). Now for $k \geq 1$, we assume that $\mathbf{u}^{k-1} \in \mathcal{M}$ is given. In order to prove the existence of $(\mathbf{u}^k, \boldsymbol{\pi}^k)$ we will use a topological degree argument, see [42, Chap. 1]. In particular, we will build $(\mathbf{u}^k, \boldsymbol{\pi}^k)$ as a solution to a nonlinear system $F^1(\mathbf{u}^k, \boldsymbol{\pi}^k) = 0$ where the map F^1 corresponds to the scheme (44)–(46). Roughly speaking in the sequel we will define an homotopy F^λ , where λ denotes a parameter in $[0, 1]$, such that if $\lambda = 0$ the system $F^0(\mathbf{u}, \boldsymbol{\pi}) = 0$ is linear and admits a unique solution. Then, we will conclude thanks to the invariance by homotopy of the Brouwer topological degree that the nonlinear set of equation $F^1(\mathbf{u}, \boldsymbol{\pi}) = 0$ admits at least one solution, i.e., the scheme (44)–(46) (or equivalently (S)) possesses at least one solution at time t_k denoted by $(\mathbf{u}^k, \boldsymbol{\pi}^k)$.

In this purpose we introduce, for $R > 0$, the set

$$Z_R = \left\{ (\mathbf{u}, \boldsymbol{\pi}) = (\mathbf{u}_0, \dots, \mathbf{u}_n, \boldsymbol{\pi}) \in \mathcal{H}_{\mathcal{T}}^{n+2} \mid \|\boldsymbol{\pi}\|_{0,2,\mathcal{T}} + \sum_{i=0}^n \|\mathbf{u}_i\|_{0,1,\mathcal{T}} < R \right\},$$

and for each $\lambda \in [0, 1]$ a map $F^\lambda : Z_R \rightarrow \mathbb{R}^{(n+2)\theta}$ with $\theta = \#\mathcal{T}$. Then $(\mathbf{u}^{k,\lambda}, \boldsymbol{\pi}^{k,\lambda})$ stands for a zero of F^λ , i.e. $(\mathbf{u}^{k,\lambda}, \boldsymbol{\pi}^{k,\lambda})$ satisfies $F^\lambda(\mathbf{u}^{k,\lambda}, \boldsymbol{\pi}^{k,\lambda}) = 0$ for every $\lambda \in [0, 1]$. Since the construction of the homotopy F^λ is non-trivial we split those definitions in two steps.

Step 1. If $\lambda \in [0, 1/2]$, then the problem $F^\lambda(\mathbf{u}^{k,\lambda}, \boldsymbol{\pi}^{k,\lambda}) = 0$ is given for every $K \in \mathcal{T}$ and $i = 0, \dots, n$ by

$$\mathfrak{m}(K) \frac{u_{i,K}^{k,\lambda} - u_{i,K}^{k-1}}{\Delta t} - \frac{d_*}{2} \sum_{\sigma \in \mathcal{E}_K} \tau_\sigma \left(\frac{\alpha_*}{2} \mathbb{D}_{K,\sigma} \mathbf{u}_i^{k,\lambda} + \left(\bar{u}_i + 2\lambda \left(u_{i,\sigma}^{k,\lambda,\pi} - \bar{u}_i \right) \right) \mathbb{D}_{K,\sigma} \boldsymbol{\pi}^{k,\lambda} \right) = 0,$$

where $d_* = \min_{i=0,\dots,n} d_i$, $\alpha_* = \min_{i=0,\dots,n} \alpha_i$ and $\bar{u}_i = \mathfrak{m}(\Omega)^{-1} \sum_{K \in \mathcal{T}} \mathfrak{m}(K) u_{i,K}^{k,\lambda} = \mathfrak{m}(\Omega)^{-1} \int_\Omega u_i^0(x) dx$. Moreover we add the constraints

$$\sum_{i=0}^n u_{i,K}^{k,\lambda} = 1 \quad \forall K \in \mathcal{T} \setminus \{K_0\}, \quad \sum_{K \in \mathcal{T}} \mathfrak{m}(K) \boldsymbol{\pi}_K^{k,\lambda} = 0.$$

We notice that for $\lambda = 0$ the map F^0 reduces to

(47)

$$\mathfrak{m}(K) u_{i,K}^{k,0} - \Delta t \frac{d_*}{2} \sum_{\sigma \in \mathcal{E}_K} \tau_\sigma \left(\frac{\alpha_*}{2} \mathbb{D}_{K,\sigma} \mathbf{u}_i^{k,0} + \bar{u}_i \mathbb{D}_{K,\sigma} \boldsymbol{\pi}^{k,0} \right) = \mathfrak{m}(K) u_{i,K}^{k-1}, \quad \forall K \in \mathcal{T}, i = 0, \dots, n,$$

and

$$(48) \quad \sum_{i=0}^n u_{i,K}^{k,0} = 1 \quad \forall K \in \mathcal{T} \setminus \{K_0\}, \quad \sum_{K \in \mathcal{T}} \mathfrak{m}(K) \boldsymbol{\pi}_K^{k,0} = 0.$$

This system is linear and corresponds to a matrix $\mathbb{L} \in \mathcal{M}_{(n+2)\theta, (n+2)\theta}$. Now our objective is to show that the system (47)–(48) is well-posed. In this purpose we study the kernel of \mathbb{L} , i.e., we set the right hand side of the previous system to 0. Using the constraint $\sum_{i=0}^n u_{i,K}^{k,0} = 0$ for all $K \in \mathcal{T}$ (we simply adapt the proof of Lemma 13) and the fact that by construction we have $\sum_{i=0}^n \bar{u}_i = 1$, then summing (47) over $i = 0, \dots, n$ we readily deduce that $\sum_{\sigma \in \mathcal{E}_K} \tau_\sigma \mathbb{D}_{K,\sigma} \boldsymbol{\pi}^{k,0} = 0$ for all $K \in \mathcal{T}$. Then, multiplying this equality by $\pi_K^{k,0}$, summing over $K \in \mathcal{T}$ and applying a discrete integration by parts we obtain $|\boldsymbol{\pi}^{k,0}|_{1,2,\mathcal{T}} = 0$. Thus, the connectedness of Ω , see assumption (H1), implies that $\boldsymbol{\pi}^{k,0}$ is a constant vector. Moreover, thanks to the constraint $\sum_{K \in \mathcal{T}} \mathfrak{m}(K) \boldsymbol{\pi}_K^{k,0} = 0$ we conclude that $\boldsymbol{\pi}_K^{k,0} = 0$ for all $K \in \mathcal{T}$. Thus, $\mathbf{u}_i^{k,0}$ is the solution to the classical backward Euler TPFA scheme for the heat equation for $i = 0, \dots, n$ which is well-posed. In particular this implies that $\ker(\mathbb{L}) = \{\mathbf{0}_{(n+2)\theta}\}$, i.e., the linear system (47)–(48) admits a unique solution. Using the constraint (48) (which is also valid for K_0 thanks to Lemma 13), one can reproduce the calculation discussed above to show that $\boldsymbol{\pi}_K^{k,0} = 0$ for every $K \in \mathcal{T}$ also in the non-homogeneous case, then that $\mathbf{u}_i^{k,0}$ is the “solution” to the discrete heat equation with the same diffusion constant equal to $\alpha_* d_*/4$ for $i = 0, \dots, n$.

We conclude that $\deg(F^0, Z_R, 0) = 1$, where \deg denotes the Brouwer topological degree, for any $R > 0$ such that

$$(49) \quad R > \mathfrak{m}(\Omega)^{1/2} \sum_{i=0}^n \|\mathbf{u}_i^{k-1}\|_{0,2,\mathcal{T}} = C_1.$$

Indeed, to obtain this condition on R we first multiply (47) by $u_{i,K}^{k,0}$, we sum over $K \in \mathcal{T}$ and we use an integration by parts to get

$$\|\mathbf{u}_i^{k,0}\|_{0,2,\mathcal{T}}^2 + \frac{\Delta t d_* \alpha_*}{2} |\mathbf{u}_i^{k,0}|_{1,2,\mathcal{T}}^2 = \sum_{K \in \mathcal{T}} \mathfrak{m}(K) u_{i,K}^{k-1} u_{i,K}^{k,0}, \quad \text{for } i = 0, \dots, n.$$

Then, the Cauchy-Schwarz inequality yields

$$\|\mathbf{u}_i^{k,0}\|_{0,2,\mathcal{T}} \leq \|\mathbf{u}_i^{k-1}\|_{0,2,\mathcal{T}}, \quad \text{for } i = 0, \dots, n,$$

which implies

$$\|\mathbf{u}_i^{k,0}\|_{0,1,\mathcal{T}} \leq \mathfrak{m}(\Omega)^{1/2} \|\mathbf{u}_i^{k-1}\|_{0,2,\mathcal{T}}, \quad \text{for } i = 0, \dots, n.$$

Now, passing from $\lambda = 0$ to $\lambda = 1/2$ allows to introduce some nonlinearity at the level of the Lagrange multiplier, even though it remains inactive. Indeed, in the case $\lambda \in (0, 1/2]$, adapting

the proof of Lemma 13 we notice that it holds $\sum_{i=0}^n u_{i,\sigma}^{k,\lambda,\pi} = 1$ for every $\sigma \in \mathcal{E}_{\text{int},K}$ which implies $\pi_K^{k,\lambda} = 0$ for all $K \in \mathcal{T}$. Then, arguing as before one can prove that $\deg(F^\lambda, Z_R, 0) = 1$ for all $\lambda \in (0, 1/2]$ where the solution $(\mathbf{u}^{k,\lambda}, \boldsymbol{\pi}^{k,\lambda})$ to $F^\lambda(\mathbf{u}^{k,\lambda}, \boldsymbol{\pi}^{k,\lambda}) = 0$ is given by $(\mathbf{u}^{k,0}, \boldsymbol{\pi}^{k,0})$.

Step 2. If $\lambda \in [1/2, 1]$, then the problem $F^\lambda(\mathbf{u}^{k,\lambda}, \boldsymbol{\pi}^{k,\lambda}) = 0$ is given for every $K \in \mathcal{T}$ and $i = 0, \dots, n$ by

$$(50) \quad \text{m}(K) \frac{u_{i,K}^{k,\lambda} - u_{i,K}^{k-1}}{\Delta t} - d_i^\lambda \sum_{\sigma \in \mathcal{E}_K} \tau_\sigma \left(\alpha_i^\lambda \text{D}_{K,\sigma} \mathbf{u}_i^{k,\lambda} \right. \\ \left. + 2 \left(\lambda - \frac{1}{2} \right) u_{i,\sigma}^{k,\lambda} (\text{D}_{K,\sigma} b_i(\mathbf{u}^{k,\lambda}) + \text{D}_{K,\sigma} \boldsymbol{\Psi}_i) + u_{i,\sigma}^{k,\lambda,\pi} \text{D}_{K,\sigma} \boldsymbol{\pi}^{k,\lambda} \right) = 0,$$

where $d_i^\lambda = (\lambda - 1/2)(2d_i - d_*) + d_*/2$, $\alpha_i^\lambda = (\lambda - 1/2)(2\alpha_i - \alpha_*) + \alpha_*/2$ and we impose the constraints

$$(51) \quad \sum_{i=0}^n u_{i,K}^{k,\lambda} = 1 \quad \forall K \in \mathcal{T} \setminus \{K_0\}, \quad \sum_{K \in \mathcal{T}} \text{m}(K) \pi_K^{k,\lambda} = 0.$$

Let us notice that for $\lambda = 1$ and thanks to Lemma 13 we recover the scheme (S).

Now our main objective is to prove that $\deg(F^1, Z_R, 0) = 1$ for $R > 0$ large enough. In this purpose, thanks to the invariance by homotopy of \deg , and since we already know that $\deg(F^\lambda, Z_R, 0) = 1$ for every $\lambda \in [0, 1/2]$, it is sufficient to prove that any solution $(\mathbf{u}^{k,\lambda}, \boldsymbol{\pi}^{k,\lambda}) \in \overline{Z}_R$ to $F^\lambda(\mathbf{u}^{k,\lambda}, \boldsymbol{\pi}^{k,\lambda}) = 0$ satisfies $(\mathbf{u}^{k,\lambda}, \boldsymbol{\pi}^{k,\lambda}) \notin \partial Z_R$ for every $\lambda \in [1/2, 1]$ and for sufficiently large values of $R > 0$.

First, adapting the proof of Lemma 11 we have

$$0 < u_{i,K}^{k,\lambda} < 1 \quad \forall K \in \mathcal{T}, i = 0, \dots, n.$$

Now our main objective is to adapt the proof of Proposition 9. In this purpose we multiply (50) by $\log(u_{i,K}^{k,\lambda})/d_i^\lambda$, we sum over $K \in \mathcal{T}$ and $i = 0, \dots, n$ and after a discrete integration by parts we obtain

$$J_4 + J_5 = 0,$$

with

$$J_4 = \sum_{i=0}^n \sum_{K \in \mathcal{T}} \text{m}(K) \frac{(u_{i,K}^{k,\lambda} - u_{i,K}^{k-1})}{\Delta t} \frac{1}{d_i^\lambda} \log(u_{i,K}^{k,\lambda}),$$

and

$$J_5 = \sum_{i=0}^n \sum_{\substack{\sigma \in \mathcal{E}_{\text{int}} \\ \sigma = K|L}} \tau_\sigma \left(\alpha_i^\lambda \text{D}_{K,\sigma} \mathbf{u}_i^{k,\lambda} + 2 \left(\lambda - \frac{1}{2} \right) u_{i,\sigma}^{k,\lambda} (\text{D}_{K,\sigma} b_i(\mathbf{u}^{k,\lambda}) + \text{D}_{K,\sigma} \boldsymbol{\Psi}_i) \right. \\ \left. + u_{i,\sigma}^{k,\lambda,\pi} \text{D}_{K,\sigma} \boldsymbol{\pi}^{k,\lambda} \right) \text{D}_{K,\sigma} \log(u_i^{k,\lambda}).$$

Then, using the same arguments as in the proof of Proposition 9 we get

$$J_4 \geq \frac{H_\lambda(\mathbf{u}^{k,\lambda}) - H_\lambda(\mathbf{u}^{k-1})}{\Delta t},$$

where for any nonnegative $\mathbf{v} = (v_0, \dots, v_n) \in \mathcal{H}_T^{n+1}$ we define H_λ as

$$H_\lambda(\mathbf{v}) = \sum_{i=0}^n \frac{1}{d_i^\lambda} \sum_{K \in \mathcal{T}} \text{m}(K) (v_{i,K} (\log(v_{i,K}) - 1) + 1),$$

and

$$J_5 \geq \sum_{i=0}^n \sum_{\substack{\sigma \in \mathcal{E}_{\text{int}} \\ \sigma = K|L}} \tau_\sigma \left(4\alpha_i^\lambda \left(D_\sigma \sqrt{\mathbf{u}_i^{k,\lambda}} \right)^2 + 2 \left(\lambda - \frac{1}{2} \right) D_{K,\sigma} \mathbf{u}_i^{k,\lambda} \left(D_{K,\sigma} b_i(\mathbf{u}^{k,\lambda}) + D_{K,\sigma} \Psi_i \right) \right).$$

Hence gathering the previous inequalities, following the computations done in the proof of Proposition 9 and noticing that $d_i^\lambda \leq d_i$ for $i = 0, \dots, n$ we deduce that $(\mathbf{u}^{k,\lambda}, \boldsymbol{\pi}^{k,\lambda})$ satisfies the following inequality

$$(52) \quad \frac{H_\lambda(\mathbf{u}^{k,\lambda}) - H_\lambda(\mathbf{u}^{k-1})}{\Delta t} + \sum_{i=0}^n 2\alpha_i^\lambda \sum_{\sigma \in \mathcal{E}} \tau_\sigma \left(D_\sigma \sqrt{\mathbf{u}_i^{k,\lambda}} \right)^2 + 2 \left(\lambda - \frac{1}{2} \right) \sum_{\sigma \in \mathcal{E}} \tau_\sigma \langle D_{K,\sigma} \mathbf{u}^{k,\lambda}, D_{K,\sigma} \mathbf{b}(\mathbf{u}^{k,\lambda}) \rangle \leq \sum_{i=0}^n \frac{C_\zeta}{2\alpha_i^\lambda} \|\Psi_i\|_{H^1(\Omega)}^2.$$

Now, using classical arguments, we notice that for every $C > 0$ it holds

$$\log(1+C)x - C \leq x(\log(x) - 1) + 1, \quad \forall x \geq 0,$$

where the equality is reached at $x = 1 + C$. Then, applying the convexity of the function B and the fact that $d_i^\lambda \geq d_i^{1/2} = d_*/2$ for all $\lambda \in [1/2, 1]$, $i = 0, \dots, n$, we deduce thanks to (52) that

$$(53) \quad \sum_{i=0}^n \|\mathbf{u}_i^{k,\lambda}\|_{0,1,\mathcal{T}} \leq \frac{d^*}{\log(1+C)} \left(H_{1/2}(\mathbf{u}^{k-1}) + \frac{(n+1)C \text{m}(\Omega)}{d_*} + \Delta t \sum_{i=0}^n \frac{C_\zeta}{2\alpha_i} \|\Psi_i\|_{H^1(\Omega)}^2 \right) = C_2,$$

with $d^* = \max_{i=0,\dots,n} d_i$ and

$$H_{1/2}(\mathbf{u}^{k-1}) = \sum_{i=0}^n \frac{2}{d_*} \sum_{K \in \mathcal{T}} \text{m}(K) \left(u_{i,K}^{k-1} (\log(u_{i,K}^{k-1}) - 1) + 1 \right).$$

Moreover, adapting the proof of Proposition 12 and noticing that $d_i^\lambda \leq d_i$ for all $\lambda \in [1/2, 1]$ for $i = 0, \dots, n$, we have

$$|\boldsymbol{\pi}^{k,\lambda}|_{1,2,\mathcal{T}} \leq \sum_{i=0}^n \frac{d_i}{d_*} \left(\alpha_i^\lambda |\mathbf{u}_i^{k,\lambda}|_{1,2,\mathcal{T}} + \|Db_i\|_\infty \sum_{j=0}^n |\mathbf{u}_j^{k,\lambda}|_{1,2,\mathcal{T}} + |\Psi_i|_{1,2,\mathcal{T}} \right).$$

Now applying the inequality

$$(a-b)^2 \leq 4(\sqrt{a} - \sqrt{b})^2 \quad \forall a, b \in [0, 1],$$

we obtain

$$|\boldsymbol{\pi}^{k,\lambda}|_{1,2,\mathcal{T}} \leq \sum_{i=0}^n \frac{d_i}{d_*} \left(2\alpha_i^\lambda \left| \sqrt{\mathbf{u}_i^{k,\lambda}} \right|_{1,2,\mathcal{T}} + 2\|Db_i\|_\infty \sum_{j=0}^n \left| \sqrt{\mathbf{u}_j^{k,\lambda}} \right|_{1,2,\mathcal{T}} + |\Psi_i|_{1,2,\mathcal{T}} \right).$$

Thanks to (52) and the fact that $\alpha_i^\lambda \geq \alpha_*/2$ for all $\lambda \in [1/2, 1]$, we remark that it holds

$$2 \left| \sqrt{\mathbf{u}_i^{k,\lambda}} \right|_{1,2,\mathcal{T}}^2 \leq \frac{2H_{1/2}(\mathbf{u}^{k-1})}{\alpha_* \Delta t} + \frac{C_\zeta}{\alpha_*^2} \sum_{i=0}^n \|\Psi_i\|_{H^1(\Omega)}^2 = C_3, \quad \text{for } i = 0, \dots, n.$$

Thus

$$(54) \quad |\boldsymbol{\pi}^{k,\lambda}|_{1,2,\mathcal{T}} \leq \sum_{i=0}^n \frac{d_i}{d_*} \left(\alpha_* \sqrt{2C_3} + (n+1)\sqrt{2C_3} \|Db_i\|_\infty + |\Psi_i|_{1,2,\mathcal{T}} \right) = C_4,$$

where $\alpha_* = \max_{i=0,\dots,n} \alpha_i$. Furthermore, applying the zero-mean condition $\sum_{K \in \mathcal{T}} \text{m}(K) \pi_K^{k,\lambda} = 0$ and the discrete Poincaré-Wirtinger inequality obtained in [14, Theorem 5] we deduce the existence of a constant $C_5 > 0$ only depending on Ω such that

$$\|\boldsymbol{\pi}^{k,\lambda}\|_{0,2,\mathcal{T}} \leq \frac{C_5}{\zeta^{1/2}} |\boldsymbol{\pi}^{k,\lambda}|_{1,2,\mathcal{T}}.$$

Hence, thanks to (54) we conclude that

$$(55) \quad \|\boldsymbol{\pi}^{k,\lambda}\|_{0,2,\mathcal{T}} \leq \frac{C_4 C_5}{\zeta^{1/2}} = C_6.$$

Finally, defining $R = \max\{C_1 + 1, C_2 + C_6 + 1\}$ we deduce from (49), (53) and (55) that

$$\|\boldsymbol{\pi}^{k,\lambda}\|_{0,2,\mathcal{T}} + \sum_{i=0}^n \|\mathbf{u}_i^{k,\lambda}\|_{0,1,\mathcal{T}} < R \quad \forall \lambda \in [0, 1],$$

which implies that $(\mathbf{u}^{k,\lambda}, \boldsymbol{\pi}^{k,\lambda}) \notin \partial Z_R$ and $\deg(F^\lambda, Z_R, 0) = 1$ for all $\lambda \in [0, 1]$. Thus, the scheme (S) (or equivalently (24)–(32)) admits at least one solution denoted $(\mathbf{u}^k, \boldsymbol{\pi}^k)$. Moreover, thanks to the a priori estimates established in Section 4 we conclude that $(\mathbf{u}^k, \boldsymbol{\pi}^k)$ satisfies the properties (i)–(iii) and we notice that taking $\lambda = 1$ in (52) yields the property (iv) stated in Theorem 7.

6. UNIFORM ESTIMATES

Let us first collect the a priori estimates uniform w.r.t. Δx and Δt already established in Section 4 and Section 5.

Proposition 14. *Let the assumptions of Theorem 7 hold. Then, there exists a constant $C_7 > 0$ only depending on $\Omega, T, \alpha_* = \min_{i=0,\dots,n} \alpha_i, \zeta, H(\mathbf{u}^0)$ and the sum over $i = 0, \dots, n$ of the H^1 norm of Ψ_i such that*

$$(56) \quad \sum_{i=0}^n \sum_{k=1}^{N_T} \Delta t \|\mathbf{u}_i^k\|_{1,2,\mathcal{T}}^2 \leq C_7 \quad i = 0, \dots, n.$$

Furthermore, there exists a constant $C_8 > 0$ only depending on $\Omega, T, \alpha^* = \max_{i=0,\dots,n} \alpha_i, \zeta, H(\mathbf{u}^0), d_*, d^*$ and the sum over $i = 0, \dots, n$ of the L^∞ norm of Db_i and the H^1 norm of Ψ_i such that

$$(57) \quad \sum_{k=1}^{N_T} \Delta t \|\boldsymbol{\pi}^k\|_{1,2,\mathcal{T}}^2 \leq C_8.$$

Proof. First let us notice that (56) is a direct consequence of the L^∞ estimate $u_{i,K}^k \leq 1$ for every $K \in \mathcal{T}, k \geq 1$ and $i = 1, \dots, n$ obtained in Theorem 7 and the entropy dissipation inequality (34). Indeed applying the inequality $D_\sigma u_i^k \leq 2D_\sigma \sqrt{u_i^k}$ for all $\sigma \in \mathcal{E}$ and summing over k (34) yield

$$\sum_{i=0}^n \sum_{k=1}^{N_T} \Delta t \|\mathbf{u}_i^k\|_{1,2,\mathcal{T}}^2 \leq \frac{1}{\alpha_*} \left(2H(\mathbf{u}^0) + \frac{T}{\alpha_*} \sum_{i=0}^n C_\zeta \|\Psi_i\|_{H^1(\Omega)}^2 \right).$$

Moreover, estimate (57) is a straightforward consequence of Proposition 12 (see also the end of the proof of Theorem 7). \square

Let us now establish another estimate uniform w.r.t. Δx and Δt needed for the convergence proof.

Proposition 15. *Let the assumptions of Theorem 7 hold. Then there exists a constant $C_9 > 0$ only depending on $\Omega, N, T, \zeta, \alpha_i, d_i, C_7, C_8$ and the sum over $i = 0, \dots, n$ of the L^∞ norm of Db_i and the H^1 norm of Ψ_i such that for all $\varphi \in C_0^\infty(Q_T)$ we have*

$$(58) \quad \sum_{k=1}^{N_T} \sum_{K \in \mathcal{T}} m(K) \left(u_{i,K}^k - u_{i,K}^{k-1} \right) \varphi(x_K, t_k) \leq C_9 \|\nabla \varphi\|_{L^\infty(Q_T)}, \quad i = 1, \dots, n.$$

Proof. Let $i \in \{1, \dots, n\}$ be fixed. In the sequel for $\varphi \in C_0^\infty(Q_T)$ we denote for all $K \in \mathcal{T}$ and $k \geq 1$ by $\varphi_K^k = m(K)^{-1} \int_K \varphi(x, t_k) dx$ and by $\boldsymbol{\varphi}^k$ the sequence $(\varphi_K^k)_{K \in \mathcal{T}}$. Then, we multiply (25) by $\Delta t \varphi_K^k$, we sum over $K \in \mathcal{T}$ and $1 \leq k \leq N_T$ and we obtain after some discrete integration by parts

$$\sum_{k=1}^{N_T} \sum_{K \in \mathcal{T}} m(K) \left(u_{i,K}^k - u_{i,K}^{k-1} \right) \varphi_K^k = J_6 + J_7 + J_8 + J_9,$$

with

$$\begin{aligned}
J_6 &= - \sum_{k=1}^{N_T} \Delta t \sum_{\substack{\sigma \in \mathcal{E}_{\text{int}} \\ \sigma = K|L}} \tau_\sigma d_i \alpha_i D_{K,\sigma} \mathbf{u}_i^k D_{K,\sigma} \boldsymbol{\varphi}^k, \\
J_7 &= - \sum_{k=1}^{N_T} \Delta t \sum_{\substack{\sigma \in \mathcal{E}_{\text{int}} \\ \sigma = K|L}} \tau_\sigma d_i u_{i,\sigma}^k D_{K,\sigma} b_i(\mathbf{u}^k) D_{K,\sigma} \boldsymbol{\varphi}^k, \\
J_8 &= - \sum_{k=1}^{N_T} \Delta t \sum_{\substack{\sigma \in \mathcal{E}_{\text{int}} \\ \sigma = K|L}} \tau_\sigma d_i u_{i,\sigma}^k D_{K,\sigma} \boldsymbol{\Psi}_i D_{K,\sigma} \boldsymbol{\varphi}^k, \\
J_9 &= - \sum_{k=1}^{N_T} \Delta t \sum_{\substack{\sigma \in \mathcal{E}_{\text{int}} \\ \sigma = K|L}} \tau_\sigma d_i u_{i,\sigma}^{k,\pi} D_{K,\sigma} \boldsymbol{\pi}^k D_{K,\sigma} \boldsymbol{\varphi}^k.
\end{aligned}$$

For $|J_6|$, using the regularity of φ , the Cauchy-Schwarz inequality and the regularity assumption (23) we have

$$\begin{aligned}
|J_6| &\leq d_i \alpha_i \|\nabla \varphi\|_{L^\infty(Q_T)} \sum_{k=1}^{N_T} \Delta t |\mathbf{u}_{i,K}^k|_{1,2,\mathcal{T}} \left(\sum_{K \in \mathcal{T}} \sum_{\sigma \in \mathcal{E}_{\text{int},K}} m(\sigma) d_\sigma \right)^{1/2} \\
&\leq \frac{d_i \alpha_i}{\zeta^{1/2}} \|\nabla \varphi\|_{L^\infty(Q_T)} \sum_{k=1}^{N_T} \Delta t |\mathbf{u}_{i,K}^k|_{1,2,\mathcal{T}} \left(\sum_{K \in \mathcal{T}} \sum_{\sigma \in \mathcal{E}_{\text{int},K}} m(\sigma) d(x_K, \sigma) \right)^{1/2}.
\end{aligned}$$

Now, we notice that it holds

$$\sum_{K \in \mathcal{E}_K} \sum_{\sigma \in \mathcal{E}_{\text{int},K}} m(\sigma) d(x_K, \sigma) \leq N m(\Omega),$$

which implies together with the Cauchy-Schwarz inequality and estimate (56) that

$$(59) \quad |J_6| \leq d_i \alpha_i \sqrt{\frac{N C_7 m(\Omega) T}{\zeta}} \|\nabla \varphi\|_{L^\infty(Q_T)}.$$

For $|J_7|$ since

$$0 \leq \min(u_{i,K}^k, u_{i,K,\sigma}^k) \leq u_{i,\sigma}^k \leq \max(u_{i,K}^k, u_{i,K,\sigma}^k) \leq 1, \quad \forall K \in \mathcal{T}, \sigma \in \mathcal{E}_{\text{int},K}, k \geq 1,$$

we deduce, thanks to the Lipschitz regularity of the functions b_i , that

$$|J_7| \leq d_i \|Db_i\|_\infty \sum_{j=0}^n \sum_{k=1}^{N_T} \Delta t \sum_{\substack{\sigma \in \mathcal{E}_{\text{int}} \\ \sigma = K|L}} \tau_\sigma D_\sigma \mathbf{u}_j^k D_\sigma \boldsymbol{\varphi}^k.$$

Then, using the same techniques as for the term J_6 we directly deduce that it holds

$$(60) \quad |J_7| \leq d_i \|Db_i\|_\infty \sqrt{\frac{N C_7 m(\Omega) T}{\zeta}} \|\nabla \varphi\|_{L^\infty(Q_T)}.$$

For $|J_8|$ arguing as for the term J_7 we get

$$|J_8| \leq d_i T \sqrt{\frac{N m(\Omega)}{\zeta}} |\boldsymbol{\Psi}_i|_{1,2,\mathcal{T}} \|\nabla \varphi\|_{L^\infty(Q_T)}.$$

Thanks to [45, Lemma 9.4] we deduce that there exists a constant $C_\zeta > 0$ only depending on ζ such that

$$(61) \quad |J_8| \leq d_i T C_\zeta \sqrt{\frac{N m(\Omega)}{\zeta}} \|\boldsymbol{\Psi}_i\|_{H^1(\Omega)} \|\nabla \varphi\|_{L^\infty(Q_T)}.$$

Finally, bearing in mind the L^∞ estimate $u_{i,\sigma}^{k,\pi} \leq 1$ for all $K \in \mathcal{T}$, $\sigma \in \mathcal{E}_{\text{int},K}$ and $k \geq 1$ and using the same techniques as for the term J_6 (for instance) we conclude that

$$(62) \quad |J_9| \leq d_i \sqrt{\frac{NC_8 m(\Omega)T}{\zeta}} \|\nabla \varphi\|_{L^\infty(Q_T)}.$$

Collecting estimates (59)–(62) yields the existence of $C_9 > 0$ such that (58) holds true. This finishes the proof of Proposition 15. \square

7. PROOF OF THEOREM 8

Throughout this section we will assume that the assumptions of Theorem 8 hold and we will consider the notations introduced in Section 3.3. Before to prove Theorem 8, we show some compactness properties.

7.1. Compactness properties. Let $(\mathbf{u}_m, \pi_m)_{m \in \mathbb{N}}$ be a family of finite volume solutions to (24)–(32) constructed in Theorem 7. Our first aim is to prove some compactness properties on $(\mathbf{u}_m)_{m \in \mathbb{N}}$. In this purpose we will use the black-box result [6, Theorem 3.9] (see also [34, Lemma B.1]).

Proposition 16. *Let the assumptions of Theorem 8 hold and let $(\mathbf{u}_m, \pi_m)_{m \in \mathbb{N}}$ be a sequence of discrete solutions to (24)–(32) constructed in Theorem 7. Then there exists a subsequence of (\mathbf{u}_m) , which is not relabeled, and $\mathbf{u} = (u_0, u_1, \dots, u_n)$ with $\langle \mathbf{u}, \mathbf{1} \rangle = 1$ a.e. in Q_T such that*

$$u_{i,m} \rightarrow u_i \quad \text{a.e. in } Q_T \text{ as } m \rightarrow \infty, \quad i = 0, \dots, n.$$

Proof. Let $i \in \{1, \dots, n\}$. Assumptions (A_x1) and (A_x3) in [6, Theorem 3.9] are satisfied due to the choice of finite volumes method. Assumption (A_t) is always fulfilled for one-step methods like the implicit Euler discretization. Assumptions (a) and (b) are a consequence of the L^∞ bound, while Proposition 15 ensures assumption (c). Thus, the result, for $i = 1, \dots, n$, is a direct consequence of the black-box result obtained in [6, Theorem 3.9]. Now, since

$$u_{0,m} = 1 - \sum_{i=1}^n u_{i,m},$$

we also deduce the existence of a u_0 such that $u_0 = 1 - \sum_{i=1}^n u_i$ with $u_{0,m} \rightarrow u_0$ as $m \rightarrow \infty$. \square

Proposition 17. *Under the assumptions of Proposition 16, there exists $\mathbf{u} = (u_1, \dots, u_n) \in L^2(0, T; H^1(\Omega))^n \cap L^\infty(Q_T)^n$ and $u_0 = 1 - \sum_{i=1}^n u_i$ and a subsequence of $(\mathbf{u}_m)_{m \in \mathbb{N}}$ such that for $i = 0, \dots, n$, as $m \rightarrow \infty$,*

$$(63) \quad u_{i,m} \rightarrow u_i \quad \text{strongly in } L^p(Q_T), \quad 1 \leq p < \infty,$$

$$(64) \quad \nabla^m u_{i,m} \rightharpoonup \nabla u_i \quad \text{weakly in } L^2(Q_T),$$

where ∇^m is defined in Section 3.3.

Proof. The strong convergence (63) is a consequence of Proposition 16 and the L^∞ estimate

$$0 < u_{i,m} \leq 1 \quad \text{a.e. in } Q_T,$$

obtained in Theorem 7 which allow to apply the dominated convergence theorem. Moreover, the estimate (56) implies that $(\nabla^m u_{i,m})_{m \in \mathbb{N}}$ is bounded in $L^2(Q_T)$. Thus, for a subsequence, $\nabla^m u_{i,m} \rightharpoonup v_i$ weakly in $L^2(Q_T)$ as $m \rightarrow \infty$. Following the proof of [38, Lemma 4.4] we show that $v_i = \nabla u_i$. This concludes the proof of Proposition 17. \square

Proposition 18. *Let the assumptions of Theorem 8 hold and let $(\mathbf{u}_m, \pi_m)_{m \in \mathbb{N}}$ be a sequence of discrete solutions to (24)–(32) constructed in Theorem 7. Then there exists a subsequence of (π_m) , which is not relabeled, and $\pi \in L^2(0, T; H^1(\Omega))$ such that*

$$(65) \quad \pi_m \rightharpoonup \pi \quad \text{weakly in } L^2(Q_T),$$

$$(66) \quad \nabla^m \pi_m \rightharpoonup \nabla \pi \quad \text{weakly in } L^2(Q_T).$$

Proof. This result is a direct consequence of estimate (57). Indeed from this estimate we deduce that the sequence $(\pi_m)_{m \in \mathbb{N}}$ is uniformly bounded in $L^2(Q_T)$ which implies the existence of $\pi \in L^2(Q_T)$ such that (65) holds. The same argument shows that $(\nabla^m \pi_m)_{m \in \mathbb{N}}$ is uniformly bounded in $L^2(Q_T)$ and applying the proof of [38, Lemma 4.4] we directly deduce that the convergence result (66) also holds. \square

7.2. Convergence of the scheme. To finish the proof of Theorem 8, we need to show that the function (\mathbf{u}, π) obtained in Proposition 16–Proposition 18 is a weak solution to (10)–(15) in the sense of Definition 1. To this end, we first notice that we can rewrite (32) for every $1 \leq k \leq N_T$ as

$$\int_{\Omega} \pi_m(x, t) dx = 0, \quad \forall t \in ((k-1)\Delta t_m, k\Delta t_m].$$

Now for an arbitrary function $\phi \in C_0^\infty(0, T)$, we multiply the previous equality by

$$\frac{1}{\Delta t_m} \int_{(k-1)\Delta t_m}^{k\Delta t_m} \phi(t) dt,$$

and we sum over $k \in \{1, \dots, N_T\}$ and we have

$$\int \int_{Q_T} \pi_m(x, t) \phi(t) dx dt = 0.$$

Now using the weak convergence in $L^2(Q_T)$ of π_m towards π established in Proposition 18, we deduce that after passing to the limit $m \rightarrow +\infty$ it holds

$$\int \int_{Q_T} \pi(x, t) \phi(t) dx dt = 0.$$

Thus π satisfies (13) for a.e. $t \in (0, T)$.

Let us now prove that (\mathbf{u}, π) satisfies (20) (the proof that (\mathbf{u}, π) fulfills (21) being similar). In this purpose we follow the strategy of [38]. Let $\varphi \in C_0^\infty(\Omega \times [0, T])$ be given, and let $\eta_m = \max\{\Delta x_m, \Delta t_m\}$ be sufficiently small such that $\text{supp}(\varphi) \subset \{x \in \Omega : d(x, \partial\Omega) > \eta_m\} \times [0, T]$. For the limit, we introduce the following notation:

$$\begin{aligned} F_{10}^m &= - \int_0^T \int_{\Omega} u_{i,m} \partial_t \varphi dx dt - \int_{\Omega} u_{i,m}(x, 0) \varphi(x, 0) dx, \\ F_{20}^m &= d_i \alpha_i \int_0^T \int_{\Omega} \nabla^m u_{i,m} \cdot \nabla \varphi dx dt, \\ F_{30}^m &= d_i \int_0^T \int_{\Omega} u_{i,m} \nabla^m \Psi_{i,m} \cdot \nabla \varphi dx dt, \\ F_{40}^m &= d_i \int_0^T \int_{\Omega} u_{i,m} \nabla^m b_i(\mathbf{u}_m) \cdot \nabla \varphi dx dt, \\ F_{50}^m &= d_i \int_0^T \int_{\Omega} u_{i,m} \nabla^m \pi_m \cdot \nabla \varphi dx dt. \end{aligned}$$

The convergence results of Proposition 16, Proposition 17 and Proposition 18, the regularity of the functions b_i and Ψ_i , and the assumption on the initial data show that

$$\begin{aligned} \sum_{j=1}^5 F_{j0}^m &\xrightarrow{m \rightarrow +\infty} - \int_0^T \int_{\Omega} u_i \partial_t \varphi dx dt - \int_{\Omega} u_i^0(x) \varphi(x, 0) dx \\ &+ d_i \alpha_i \int_0^T \int_{\Omega} \nabla u_i \cdot \nabla \varphi dx dt + d_i \int_0^T \int_{\Omega} u_i \nabla \Psi_i \cdot \varphi dx dt \\ &+ d_i \int_0^T \int_{\Omega} u_i \nabla b_i(\mathbf{u}) \cdot \nabla \varphi dx dt + d_i \int_0^T \int_{\Omega} u_i \nabla \pi \cdot \nabla \varphi dx dt. \end{aligned}$$

We proceed with the limit $m \rightarrow \infty$ in (25). For this, we set $\varphi_K^k = \varphi(x_K, t_k)$, multiply (25) by $\Delta t_m \varphi_K^{k-1}$ and sum over $K \in \mathcal{T}_m$, leading to

$$(67) \quad \begin{aligned} F_1^m + F_2^m + F_3^m + F_4^m + F_5^m &= 0, \quad \text{where} \\ F_1^m &= \sum_{k=1}^{N_T} \sum_{K \in \mathcal{T}} m(K) (u_{i,K}^k - u_{i,K}^{k-1}) \varphi_K^{k-1}, \\ F_2^m &= -d_i \alpha_i \sum_{k=1}^{N_T} \Delta t_m \sum_{K \in \mathcal{T}} \sum_{\sigma \in \mathcal{E}_{\text{int},K}} \tau_\sigma D_{K,\sigma} \mathbf{u}_i^k \varphi_K^{k-1}, \\ F_3^m &= -d_i \sum_{k=1}^{N_T} \Delta t_m \sum_{K \in \mathcal{T}} \sum_{\sigma \in \mathcal{E}_{\text{int},K}} \tau_\sigma u_{i,\sigma}^k D_{K,\sigma} \Psi_i \varphi_K^{k-1}, \\ F_4^m &= -d_i \sum_{k=1}^{N_T} \Delta t_m \sum_{K \in \mathcal{T}} \sum_{\sigma \in \mathcal{E}_{\text{int},K}} \tau_\sigma u_{i,\sigma}^k D_{K,\sigma} b_i(\mathbf{u}^k) \varphi_K^{k-1}, \\ F_5^m &= -d_i \sum_{k=1}^{N_T} \Delta t_m \sum_{K \in \mathcal{T}} \sum_{\sigma \in \mathcal{E}_{\text{int},K}} \tau_\sigma u_{i,\sigma}^{k,\pi} D_{K,\sigma} \boldsymbol{\pi}^k \varphi_K^{k-1}. \end{aligned}$$

The aim is to show that $F_{j0}^m - F_j^m \rightarrow 0$ as $m \rightarrow \infty$ for $j = 1, \dots, 5$. Then (67) shows that $F_{10}^m + F_{20}^m + F_{30}^m + F_{40}^m + F_{50}^m \rightarrow 0$, which finishes the proof.

It is proved in [38, Theorem 5.2], using the $L^1(Q_T)$ bound for u_m and the regularity of φ , that $F_{10}^m - F_1^m \rightarrow 0$. Using the bound (56) and arguing as in the proof of [38, Theorem 5.1] we easily show that $F_{20}^m - F_2^m \rightarrow 0$. Moreover, applying similar arguments as in the proof of [64, Theorem 2] we deduce that $F_{40}^m - F_4^m \rightarrow 0$.

Let us now prove that $F_{30}^m - F_3^m \rightarrow 0$ as $m \rightarrow \infty$. In this purpose we first apply a discrete integration by parts and write $F_3^m = F_{31}^m + F_{32}^m$ where

$$\begin{aligned} F_{31}^m &= d_i \sum_{k=1}^{N_T} \Delta t_m \sum_{K \in \mathcal{T}} \sum_{\sigma \in \mathcal{E}_{\text{int},K}} \tau_\sigma u_{i,K}^k D_{K,\sigma} \Psi_i D_{K,\sigma} \varphi^{k-1}, \\ F_{32}^m &= d_i \sum_{k=1}^{N_T} \Delta t_m \sum_{K \in \mathcal{T}} \sum_{\sigma \in \mathcal{E}_{\text{int},K}} \tau_\sigma (u_{i,\sigma}^k - u_{i,K}^k) D_{K,\sigma} \Psi_i D_{K,\sigma} \varphi^{k-1}. \end{aligned}$$

The definition of the discrete gradient ∇^m in Section 3.3 gives

$$\begin{aligned} |F_{30}^m - F_{31}^m| &\leq d_i \sum_{k=1}^{N_T} \sum_{K \in \mathcal{T}} \sum_{\sigma \in \mathcal{E}_{\text{int},K}} m(\sigma) |u_{i,K}^k| |D_{K,\sigma} \Psi_i| \\ &\quad \times \left| \int_{t_{k-1}}^{t_k} \left(\frac{D_{K,\sigma} \varphi^{k-1}}{d_\sigma} - \frac{1}{m(T_{K,\sigma})} \int_{T_{K,\sigma}} \nabla \varphi \cdot \nu_{K,\sigma} dx \right) dt \right|. \end{aligned}$$

It is shown in the proof of [38, Theorem 5.1] that there exists a constant $C > 0$ only depending on φ such that

$$\left| \int_{t_{k-1}}^{t_k} \left(\frac{D_{K,\sigma} \varphi^{k-1}}{d_\sigma} - \frac{1}{m(T_{K,\sigma})} \int_{T_{K,\sigma}} \nabla \varphi \cdot \nu_{K,\sigma} dx \right) dt \right| \leq C \Delta t_m \eta_m.$$

Hence, thanks to the bound $u_{i,K}^k \leq 1$ for all $K \in \mathcal{T}$, $k \geq 1$ and by the Cauchy-Schwarz inequality,

$$\begin{aligned} |F_{30}^m - F_{31}^m| &\leq d_i C \eta_m T \sum_{K \in \mathcal{T}} \sum_{\sigma \in \mathcal{E}_{\text{int},K}} m(\sigma) |D_{K,\sigma} \Psi_i| \\ &\leq d_i C \eta_m T |\Psi_i|_{1,2,\mathcal{T}_m} \left(\sum_{K \in \mathcal{T}} \sum_{\sigma \in \mathcal{E}_{\text{int},K}} m(\sigma) d_\sigma \right)^{1/2}. \end{aligned}$$

It follows from the mesh regularity (23) that

$$\sum_{\sigma \in \mathcal{E}_{\text{int}}} m(\sigma) d_\sigma \leq \frac{1}{\zeta} \sum_{K \in \mathcal{T}} \sum_{\sigma \in \mathcal{E}_K} m(\sigma) d(x_K, \sigma) = \frac{N}{\zeta} \sum_{K \in \mathcal{T}} m(K) = \frac{N m(\Omega)}{\zeta}.$$

Moreover, applying [45, Lemma 9.4] there exists a constant, still denoted C_ζ , only depending on ζ such that

$$|\Psi_i|_{1,2,\mathcal{T}_m} \leq C_\zeta \|\Psi_i\|_{H^1(\Omega)}$$

Therefore, we obtain

$$|F_{30}^m - F_{31}^m| \leq \sqrt{\frac{N m(\Omega)}{\zeta}} C C_\zeta d_i T \|\Psi_i\|_{H^1(\Omega)} \eta_m \rightarrow 0, \quad \text{as } m \rightarrow \infty.$$

Finally, we estimate F_{32}^m according to

$$\begin{aligned} |F_{32}^m| &\leq C \eta_m \|\varphi\|_{C^1(\bar{\Omega}_T)} G^m, \quad \text{where} \\ G^m &= \sum_{k=1}^{N_T} \Delta t_m \sum_{K \in \mathcal{T}} \sum_{\sigma \in \mathcal{E}_{\text{int},K}} \tau_\sigma |u_{i,\sigma}^k - u_{i,K}^k| |\mathbf{D}_{K,\sigma} \Psi_i|. \end{aligned}$$

Now since for $k = 1, \dots, N_T$ it holds that $\min(u_{i,K}^k, u_{i,L}^k) \leq u_{i,\sigma}^k \leq \max(u_{i,K}^k, u_{i,L}^k)$ for every $\sigma \in \mathcal{E}_{\text{int},K}$ with $\sigma = K/L$, we deduce that $|u_{i,\sigma}^k - u_{i,K}^k| \leq \mathbf{D}_\sigma \mathbf{u}_i^k$. Thus, from the Cauchy-Schwarz inequality we conclude that

$$G^m \leq T^{1/2} \left(\sum_{k=1}^{N_T} \Delta t_m \sum_{\sigma \in \mathcal{E}} \tau_\sigma (\mathbf{D}_\sigma \mathbf{u}_i^k)^2 \right)^{1/2} \left(\sum_{\sigma \in \mathcal{E}} \tau_\sigma (\mathbf{D}_\sigma \Psi_i)^2 \right)^{1/2}.$$

By Proposition 14 and [45, Lemma 9.4], the right-hand side is bounded uniformly in m . Thus, we infer that $|F_{32}^m| \leq C \eta_m \rightarrow 0$ and $F_{30}^m - F_3^m \rightarrow 0$ as $m \rightarrow \infty$. Finally, in order to prove that $F_{50}^m - F_5^m \rightarrow 0$ as $m \rightarrow \infty$ we follow the previous computations done for the term $F_{30}^m - F_3^m$. Indeed, if we look closely at the proof, we remark that we only need to have a uniform bound on the discrete H^1 seminorm of π_m and $u_{i,m}$. In particular these uniform bounds are established in Proposition 14. Therefore, we can claim that $F_{50}^m - F_5^m \rightarrow 0$ as $m \rightarrow \infty$ which finishes the proof.

8. NUMERICAL EXPERIMENTS

In this section we present some numerical experiments obtained in one and two space dimensions.

8.1. About the practical implementation. The finite-volume scheme (24)–(32) is implemented in MATLAB. More precisely, we define and implement a system equivalent to the scheme (24)–(32). In this purpose, we eliminate the unknowns $(u_{0,K}^k)_{K \in \mathcal{T}}$ thanks to (29) and, for some cell $K_0 \in \mathcal{T}$, we add the zero-mean condition (32) in equation (27). In this way, we equivalently rewrite (24)–(32) as a nonlinear system to be solved at each time step composed of $(n+1)\theta$ equations with $(n+1)\theta$ unknowns $u_{1,K}^k, \dots, u_{n,K}^k, \pi_K^k$ for all $K \in \mathcal{T}$ (where $\theta = \#\mathcal{T}$). Then, we solve this system thanks to a Newton-Raphson based iterative method. This iterative method stops as soon as the ℓ_∞ norm of the residual vector is smaller than 10^{-8} after at most 40 iterations. Finally, the vector of the obtained concentrations, denoted $\mathbf{u}^{k-2/3}$, is then projected onto \mathcal{M} by setting:

$$\mathbf{u}^{k-1/3} = \max(\mathbf{u}^{k-2/3}, 10^{-10}) \quad \text{and} \quad u_{i,K}^k = \frac{u_{i,K}^{k-1/3}}{\sum_{i=0}^n u_{i,K}^{k-2/3}} \quad \forall K \in \mathcal{T}, \quad i = 0, \dots, n.$$

8.2. A vertically-integrated model for three-phase porous media flows. Multiphase porous media flows have an important role in many applications, including oil engineering, water resource management, carbone dioxide sequestration, or subsurface hydrogen storage. Dupuit's approximation [43] leads to vertically integrated model with reduced dimension. The computational gain when compared to the direct resolution of full Darcy problems makes this approach popular. We refer for instance to [9, 55, 52, 53, 67, 58, 31] for examples of such reduced models, the list being far from being exhaustive. Here, we are interested in the situation where three phases $i \in \{0, 1, 2\}$ (say two liquid phases $i \in \{1, 2\}$ and one gas phase $i = 0$) are flowing within a $(N + 1)$ -dimensional porous medium with impervious boundaries. This situation typically occurs in the context of oil engineering where water, oil and gas share the available porous space. We assume that the $(N + 1)$ -dimensional physical domains is of the form

$$\Omega \times (0, h), \quad \Omega \subset \mathbb{R}^N,$$

where $h > 0$ denotes the thickness of the porous layer. The phases flowing within the porous medium are supposed to be separated because of gravity and to fill the whole porous space, see figure 1.

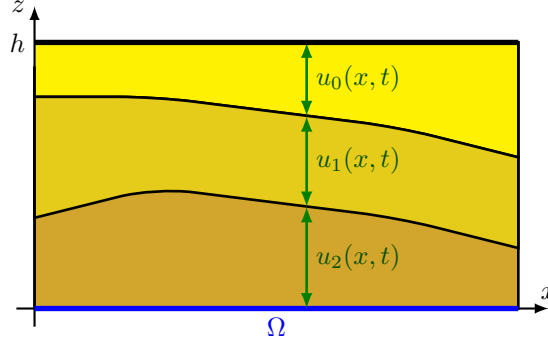


FIGURE 1. Schematic representation of the $(N + 1)$ dimensional porous medium with the stratified repartition of the 3 phases.

Taking inspiration in [65] for the Wasserstein gradient flow, in [58] for the treatment of the congestion, and approach in [41] for the so-called sharp-diffused interface approximation, consistent models for the evolution of the phase heights enter our framework. Concerning the contributions appearing in the definition of the energy $E(\mathbf{u})$, the external potentials $\Psi = (\Psi_i)_{0 \leq i \leq n}$ are all taken identically equal to 0, whereas the convex function B appearing in the energy is defined by

$$B(\mathbf{u}) = \frac{1}{2} \langle \mathbf{u}, \mathbb{B} \mathbf{u} \rangle \quad \text{with} \quad \mathbb{B} = g \begin{pmatrix} \rho_0 & \rho_0 & \rho_0 \\ \rho_0 & \rho_0 + \rho_1 & \rho_0 + \rho_1 \\ \rho_0 & \rho_0 + \rho_1 & \rho_0 + \rho_1 + \rho_2 \end{pmatrix},$$

where $\rho_0 < \rho_1 < \rho_2$ denote the densities of the phases, while g is the gravitational constant. Concerning the quantities appearing in the definition of \mathcal{M} and in the constrained Wasserstein metric (6), we set $d_i = \frac{\kappa}{\phi \eta_i}$ where $\kappa > 0$ is the permeability of the porous medium, where $\phi \in (0, 1)$ is its porosity, and where $\eta_i > 0$ denotes the viscosity of the i^{th} phase.

We use this example to validate experimentally our convergence result. We set $N = 1$, $\Omega = (0, 1)$ and $h = 1$. The characteristics of the porous medium are $\phi = 0.35$, $\kappa = 10^{-11} \text{ m}^2$, while the gravity constant is set to $g = 9.81 \text{ kg} \cdot \text{s}^{-2}$. The properties of the fluids are described in Tabular 1.

The initial profile \mathbf{u}^0 is defined by

$$u_1^0(x) = \max(0, \min(1, -2 + 3x)), \quad u_2^0(x) = \max(0, \min(1, 2 - 3x)), \quad u_0^0(x) = 1 - u_1^0(x) - u_2^0(x)$$

for $x \in (0, 1)$, while the final time is set to $T = 1$.

While it is clear that the time discretization is first order accurate since it relies on the backward Euler scheme, let us show that the space discretization is also of first order. To this end, we

	ρ_i	d_i	α_i
$i = 0$	0.2	3.4	5×10^{-2}
$i = 1$	0.8	2.5×10^{-2}	5×10^{-2}
$i = 2$	1	5.5×10^{-2}	5×10^{-2}

TABLE 1. Fluid properties

perform a convergence study by refining the discretization w.r.t. space only. Since we do not have an explicit exact solution at hand, we use a solution $(\mathbf{u}_{\text{ref}}, \pi_{\text{ref}})$ computed on a reference (uniform) grid made of 25600 cells to compare the results obtained on coarser meshes. The same non-uniform time discretization is used for all the meshes, including the reference one. The k^{th} time step is given by

$$\Delta t^k = \min(\Delta t_{\text{max}}, (1.1)^{k-1} \times \Delta t^1), \quad \text{where } \Delta t^1 = 2.10^{-8} \text{ and } \Delta t_{\text{max}} = 10^{-2}.$$

In Figure 2, we plot the evolution of the errors

$$(68) \quad \text{err}_u = \int_0^T \int_{\Omega} (|u_{1,m} - u_{1,\text{ref}}| + |u_{2,m} - u_{2,\text{ref}}|) dx dt \quad \text{and} \quad \text{err}_{\pi} = \int_0^T \int_{\Omega} |\pi_m - \pi_{\text{ref}}| dx dt$$

as a function of the number of cells in the m^{th} mesh. Only the error related to the space discretization is recorded here since the time discretization is the same for all grids. One observes a first order error decay, which was expected since we used upwinding in the definition of the scheme.

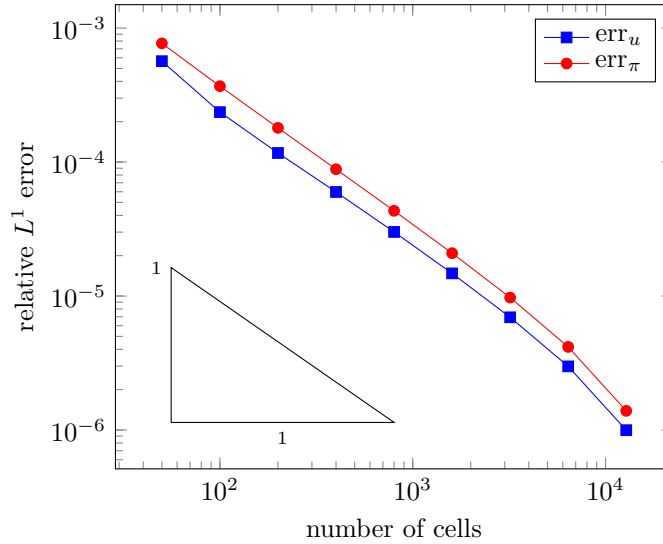


FIGURE 2. Evolution of the space-time L^1 errors (68) as a function of the number of cells in the spatial discretization.

8.3. Diffusion in multicomponent solutions. In our second example, we consider an incompressible mixture made of three inert species, inert in the sense that there is no reaction. The vector $\mathbf{u} = (u_i)_{0 \leq i \leq 2}$ denotes the volume fractions. Each specie has its own diffusion coefficient as well as its own density ρ_i , the external potential $\Psi_i(x) = -\rho_i g \cdot x$ accounting for gravity, where $|g| = 9.81$, $\rho_0 = 0.4$, $\rho_1 = 1.6$, and $\rho_2 = 6.4$. The convex function B is identically equal to 0, and all the self-diffusion coefficients α_i are taken equal to 0.1. The model then corresponds to the one discussed for instance in [69, Sec. 6.2].

For the test case, we consider a two-dimensional domain $\Omega = (0, 1)^2$, which is discretized with a Delaunay triangulation made of 97908 triangles. We used an adaptive time stepping strategy. Starting from an initial time step $\Delta t^0 = 10^{-10}$, the time step Δt^k at the k^{th} iteration in the march

in time is initially set to $\Delta t^k = \min(1.1 \times \Delta t^{k-1}, 5 \times 10^{-4})$. If the Newton-like method fails to converge after 40 iterations, we restart the simulation with a reduced time step divided by 5.

All the phases aim at going to the bottom of the domain, but the Lagrange multiplier π introduces buoyancy, so that the lighter phase progressively moves to the top of the domain, cf. Figure 3. The most dense phase will concentrate at the bottom of the computational domain, as observed on Figure 5, while the one with mid-range density locates at intermediate height, cf. Figure 4.

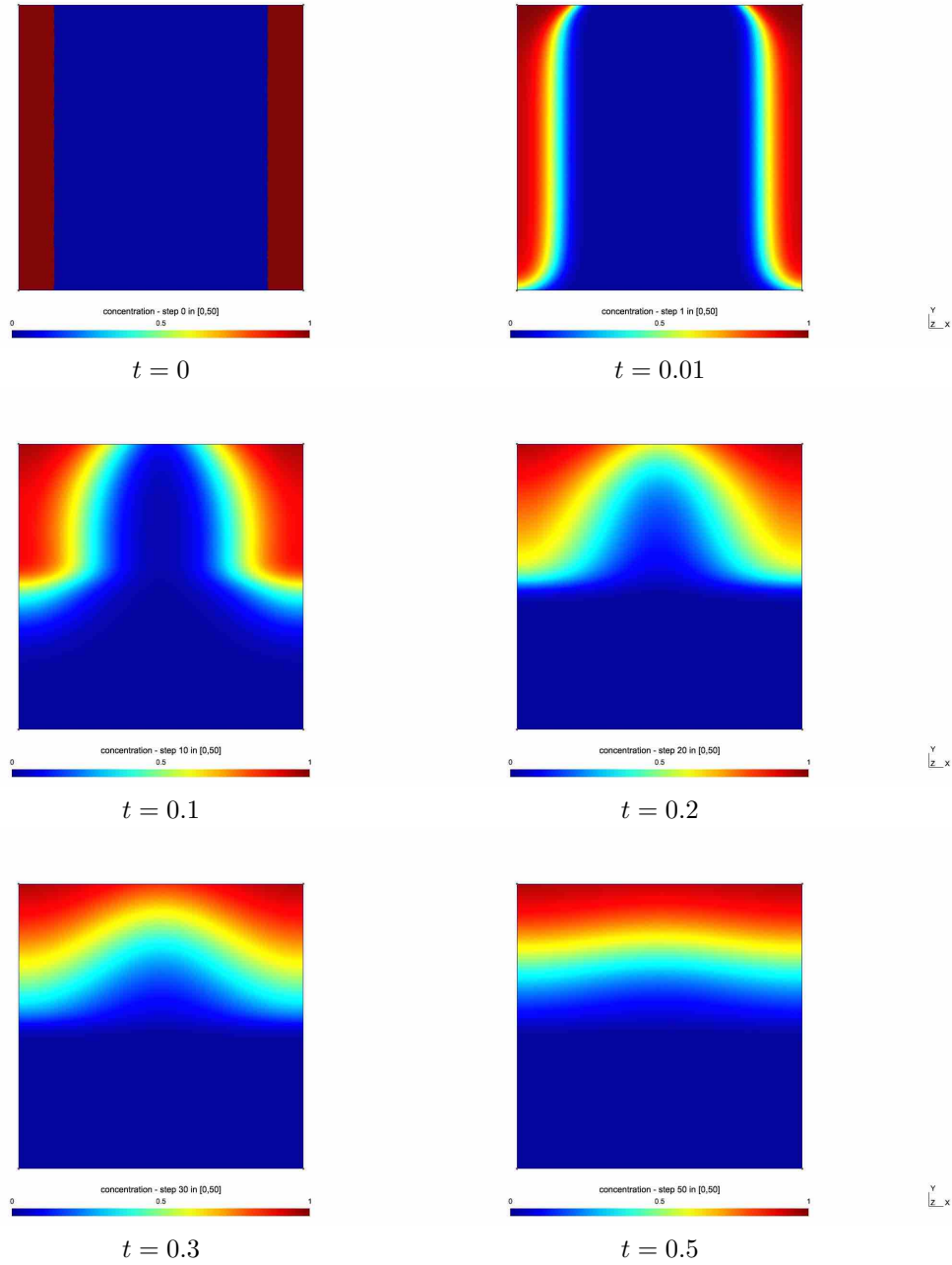
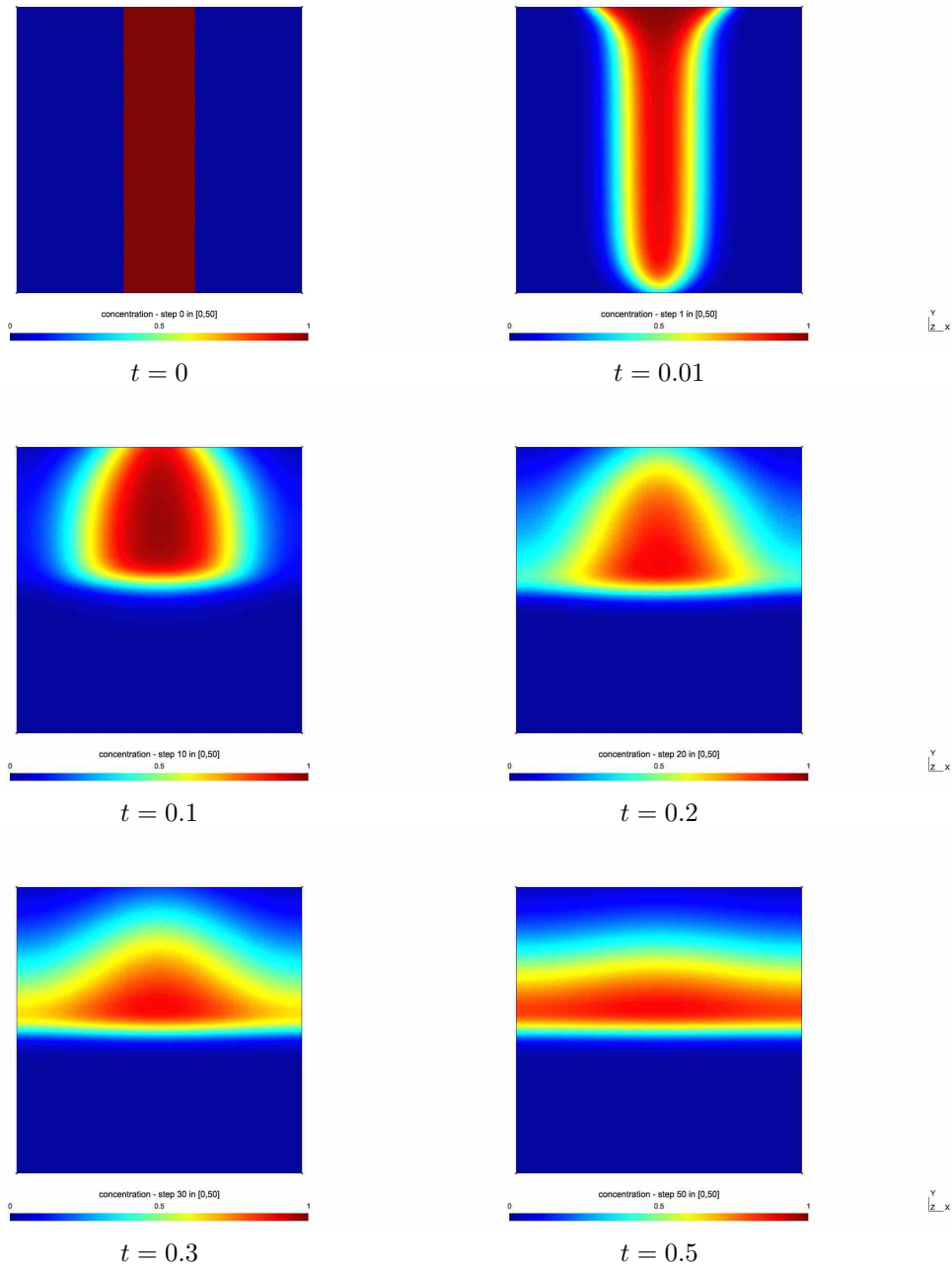


FIGURE 3. Evolution of the volume fraction u_0 along time.

FIGURE 4. Evolution of the volume fraction u_1 along time.

9. CONCLUDING REMARKS AND POSSIBLE GENERALIZATIONS

In this paper we have shown the convergence of a finite volume scheme for a class of parabolic systems (possibly including cross-diffusion terms) coupled with an elliptic equation ensuring the volume filling constraint. In particular, our scheme preserves the conservation of the mass, the nonnegativity of \mathbf{u} , the volume filling constraint $\langle \mathbf{u}, \mathbf{1} \rangle = 1$ and the decay of the entropy. However, even though the problem we consider in this paper is rather general, several generalizations are natural and worth being discussed:

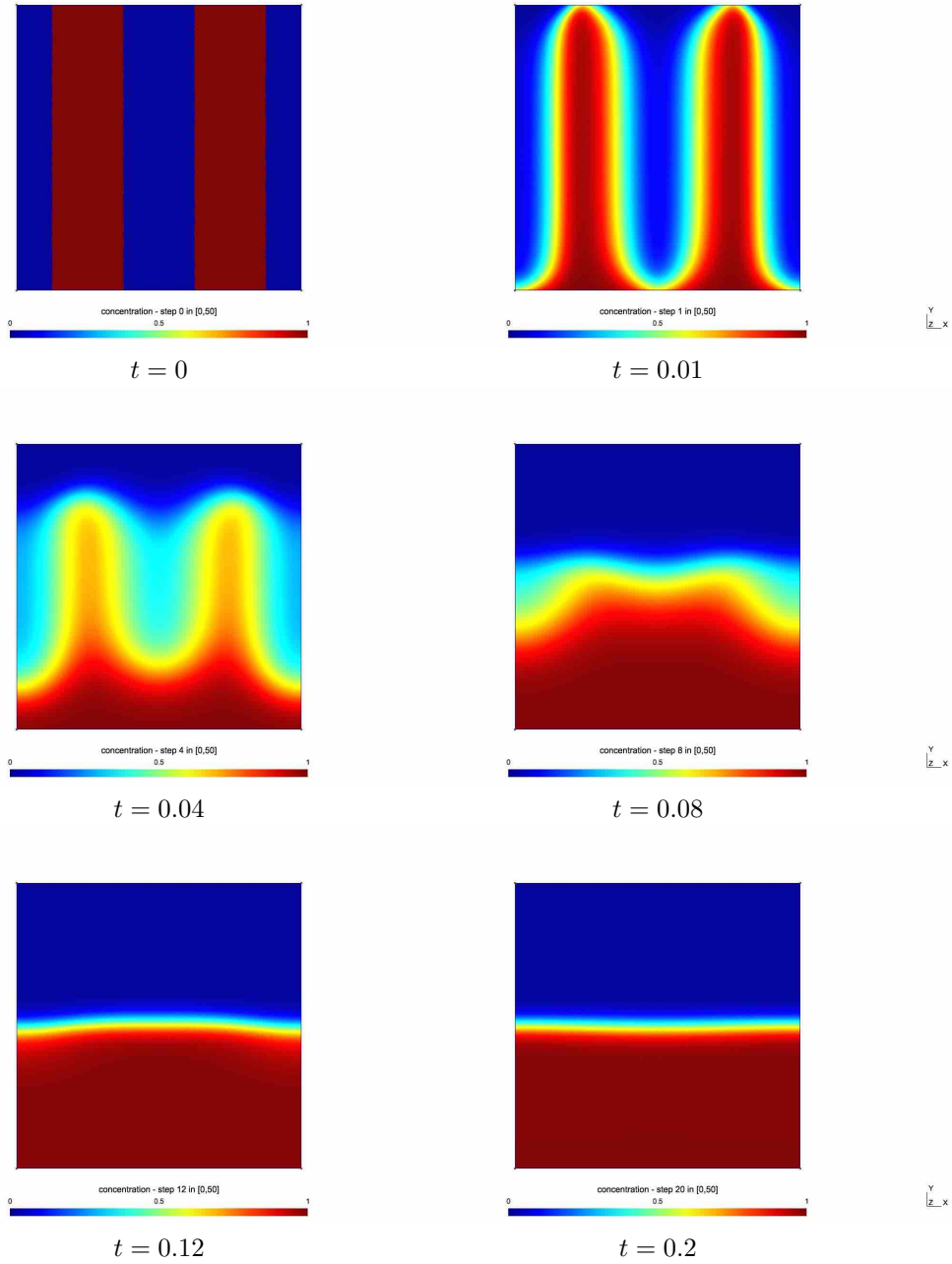
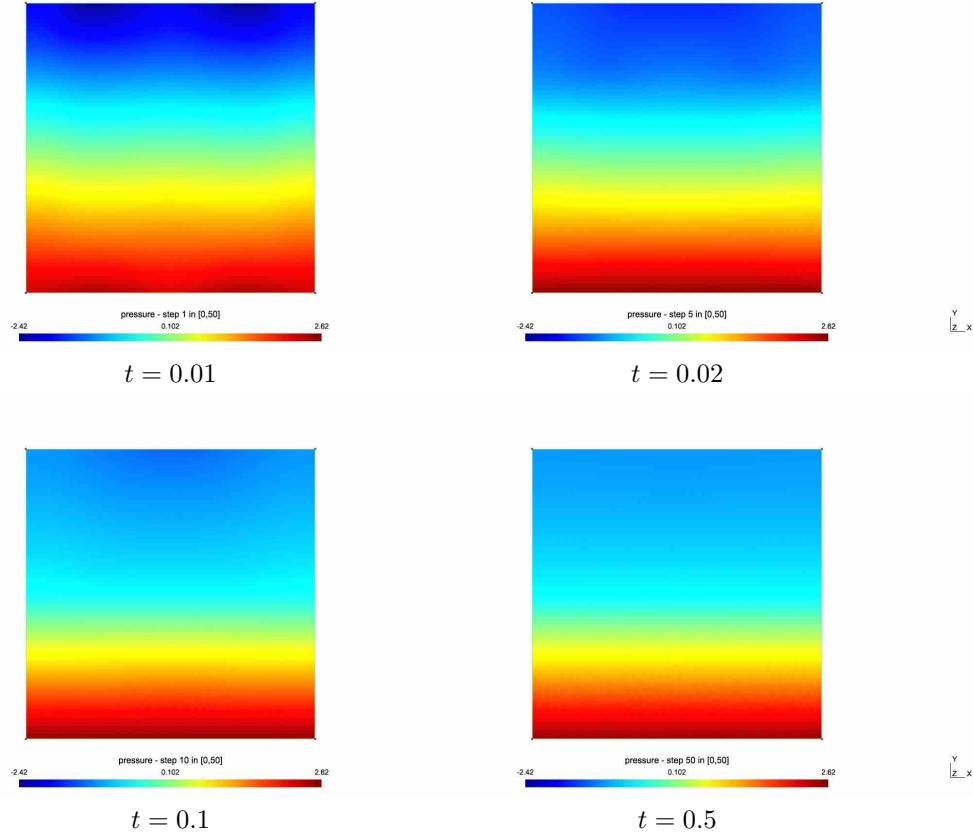


FIGURE 5. Evolution of the volume fraction u_2 along time.

- (i) Generalizing the volume filling constraint. In the context of multicomponent mixtures as discussed in Section 8.3, rather than choosing the volume fractions as main unknowns, it is more usual to consider concentrations. Doing so, the volume filling constraint (2) should turn to

$$\langle \mathbf{u}, \mathbf{v} \rangle = \sum_{i=0}^n u_i v_i = 1,$$

FIGURE 6. Evolution of the Lagrange multiplier π along time.

where v_i is the molar volume of the i^{th} chemical component. Such non-constant molar volumes leads to a change in (8), which then turns to

$$\mu_i = \frac{\delta E}{\delta u_i}(\mathbf{u}) + v_i \pi, \quad i = 0, \dots, n.$$

Another natural generalization of the volume filling constraint (2) consists in incorporating a space dependent constraint

$$\langle \mathbf{u}, \mathbf{1} \rangle = \sum_{i=0}^n u_i(x, t) = \omega(x),$$

for some smoothly varying ω bounded away from 0 and from above. This situation naturally occurs for multiphase Darcy flows where ω is the porosity, or in their Dupuit approximation, cf. Section 8.2, where the thickness h could vary. To generalize our approach to this case, the entropy $H(\mathbf{u})$ has to be modified into

$$H_\omega(\mathbf{u}) = \sum_{i=0}^n \frac{1}{d_i} \int_{\Omega} \left(u_i \log \left(\frac{u_i}{\omega} \right) + \omega - u_i \right) dx,$$

as it was for instance done in [27].

- (ii) Dirichlet boundary conditions. Following [40], we could consider mixed boundary conditions and include (at least) constant Dirichlet boundary conditions on Γ^D a subset of $\Gamma = \partial\Omega$ and no-flux boundary condition on $\Gamma^N = \Gamma \setminus \Gamma^D$ such that $\Gamma = \Gamma^D \cup \Gamma^N$. However, in this case, the mean fractions $\bar{\mathbf{u}}$ are no longer constant with respect to time. Given $\mathbf{u}^D =$

$(u_0^D, \dots, u_n^D) \in \mathbb{R}_+^{n+1}$ with $\langle \mathbf{u}^D, \mathbf{1} \rangle = 1$ be the Dirichlet boundary data, i.e., $u_i(x, t) = u_i^D$ on $\Gamma^D \times (0, T)$, we need to consider the relative entropy functional given by

$$H(\mathbf{u}|\mathbf{u}^D) = \sum_{i=0}^n \frac{1}{d_i} \int_{\Omega} \left(u_i \log \left(\frac{u_i}{u_i^D} \right) + u_i^D - u_i \right) dx,$$

rather than the mere Boltzmann entropy (17). Then, since $H(\mathbf{u}, \mathbf{u}^D)$ distinguishes from $H(\mathbf{u})$ by linear terms we can easily adapt the proof of Lemma 3 to obtain that any regular enough solution $\mathbf{u} : Q_T \rightarrow \mathbb{R}_+^{n+1}$ satisfies $\langle \mathbf{u}(x, t), \mathbf{1} \rangle = 1$ in Q_T and

$$\frac{d}{dt} H(\mathbf{u}|\mathbf{u}^D) + \sum_{i=0}^n 2\alpha_i \int_{\Omega} |\nabla \sqrt{u_i}|^2 dx + \int_{\Omega} \langle \nabla \mathbf{u}, \nabla \mathbf{b}(\mathbf{u}) \rangle dx \leq \sum_{i=0}^n \frac{1}{2\alpha_i} \int_{\Omega} |\nabla \Psi_i|^2 dx.$$

Moreover, we can establish a discrete counterpart of this modified entropy production inequality.

- (iii) Reaction terms. If we want to incorporate reaction terms $\mathbf{f} = (f_0, \dots, f_n)$ in the system (10)–(15), we first need to assume that $\sum_{i=0}^n f_i(\mathbf{u}) = 0$ for all $\mathbf{u} \in \mathbb{R}^{n+1}$ such that the total volume is conserved and the volume filling constraint $\langle \mathbf{u}, \mathbf{1} \rangle = 1$ still holds. We also need to preserve the positivity of the solutions to our scheme. In this purpose we impose the condition $f_i(\mathbf{u}) \geq 0$ for all $\mathbf{u} \in \mathbb{R}^{n+1}$ with $u_i \leq 0$ for $i = 0, \dots, n$. Furthermore, in order to include a suitable reaction term \mathbf{f} which fits with the entropy structure of the model (10)–(15), we need to assume, following [60], that there exists a constant $C_{\mathbf{f}} > 0$ such that

$$\sum_{i=0}^n \frac{1}{d_i} f_i(\mathbf{u}) \log(u_i) \leq C_{\mathbf{f}}(1 + h(\mathbf{u})),$$

where

$$h(\mathbf{u}) = \sum_{i=0}^n \frac{1}{d_i} (u_i(\log(u_i) - 1) + 1).$$

Then, assuming that $\Delta t < 1/C_{\mathbf{f}}$ and adapting the proof of [64, Lemma 5] we can show that the solutions to our scheme satisfy

$$\begin{aligned} (1 - C_{\mathbf{f}}\Delta t)H(\mathbf{u}^k) + \Delta t \sum_{i=0}^n 2\alpha_i \sum_{\sigma \in \mathcal{E}} \tau_{\sigma} \left(D_{\sigma} \sqrt{u_i^k} \right)^2 \\ + \Delta t \sum_{\sigma \in \mathcal{E}} \tau_{\sigma} \langle D_{K,\sigma} \mathbf{u}^k, D_{K,\sigma} \mathbf{b}(\mathbf{u}^k) \rangle \leq H(\mathbf{u}^{k-1}) + C_{\mathbf{f}}\Delta t m(\Omega) + \Delta t \sum_{i=0}^n \frac{C_{\zeta}}{2\alpha_i} \|\Psi_i\|_{H^1(\Omega)}^2. \end{aligned}$$

Let us remark that in this case, even if $\Psi_i = 0$ for $i = 0, \dots, n$, then H is no longer a Lyapunov functional for the system (10)–(15) but the above inequality yields enough estimates in order to generalize the proofs of our main results.

- (iv) Deep quench limit $\alpha_i = 0$. In this case we have to assume that B is a C^2 strictly convex function such that its Hessian matrix \mathbb{B} is a positive definite matrix with smallest eigenvalue $\mu > 0$. Then, adapting the proof of Lemma 3, for every regular enough function $\mathbf{u} : [0, T] \rightarrow \mathcal{M}$ we have

$$\frac{d}{dt} H(\mathbf{u}) + \int_{\Omega} \langle \nabla \mathbf{u}, \mathbb{B}(\mathbf{u}) \nabla \mathbf{u} \rangle dx = - \int_{\Omega} \langle \nabla \mathbf{u}, \nabla \Psi \rangle dx.$$

Moreover, since it holds

$$\int_{\Omega} \langle \nabla \mathbf{u}, \mathbb{B}(\mathbf{u}) \nabla \mathbf{u} \rangle dx \geq \mu \sum_{i=0}^n \int_{\Omega} |\nabla u_i|^2 dx,$$

we obtain

$$\frac{d}{dt} H(\mathbf{u}) + \frac{\mu}{2} \sum_{i=0}^n \int_{\Omega} |\nabla u_i|^2 dx \leq \frac{1}{2\mu} \sum_{i=0}^n \int_{\Omega} |\nabla \Psi_i|^2 dx.$$

Finally, replacing in the scheme (24)–(32) the logarithmic mean mobilities by upwind mobilities, i.e. for all $\sigma \in \mathcal{E}$ and $i = 0, \dots, n$ we consider

$$u_{i,\sigma}^k = \begin{cases} u_{i,K}^k & \text{if } D_{K,\sigma} b_i(\mathbf{u}^k) + D_{K,\sigma} \Psi_i \leq 0, \\ u_{i,K,\sigma}^k & \text{if } D_{K,\sigma} b_i(\mathbf{u}^k) + D_{K,\sigma} \Psi_i > 0, \end{cases}$$

then we can extend the proofs of our main results in this setting, which is the one of [26].

ACKNOWLEDGMENTS

We thank the two referees for their very careful reading of the manuscript and their constructive comments. The authors have been supported by the multilateral project of the Austrian Agency for International Co-operation in Education and Research (OeAD), grant FR 01/2021 and by the Campus France PHC AMADEUS program (grant 46397PA). C. Cancès acknowledges support from the Labex CEMPI (ANR-11-LABX-0007-01) and from the COMODO project (ANR-19-CE46-0002). A. Zurek has been partially supported by the Austrian Science Fund (FWF), grants P30000, P33010, F65, and W1245.

REFERENCES

- [1] A. Ait Hammou Oulhaj. Numerical analysis of a finite volume scheme for a seawater intrusion model with cross-diffusion in an unconfined aquifer. *Numer. Methods Partial Differential Equations*, 34:857–880, 2018.
- [2] A. Ait Hammou Oulhaj and D. Maltese. Convergence of a positive nonlinear control volume finite element scheme for an anisotropic seawater intrusion model with sharp interfaces. *Numer. Methods Partial Differential Equations*, 36(1):133–153, 2019.
- [3] A. Ait Hammou Oulhaj, D. Maltese, and N. Staïli. Convergence and long time behavior of a Finite Volume scheme for an isotropic seawater intrusion model with a sharp-diffuse interface in a confined aquifer. *Hal-02865698* (2020).
- [4] L. Ambrosio, N. Gigli, and G. Savaré. *Gradient flows in metric spaces and in the space of probability measures*, Lectures in Mathematics ETH Zürich, Birkhäuser Verlag, 2008.
- [5] B. Andreianov, M. Bendahmane, and R. Ruiz Baier. Analysis of a finite volume method for a cross-diffusion model in population dynamics. *Math. Models Meth. Appl. Sci.*, 21:307–344, 2011.
- [6] B. Andreianov, C. Cancès, and A. Moussa. A nonlinear time compactness result and applications to discretization of degenerate parabolic-elliptic PDEs. *J. Funct. Anal.*, 273:3633–3670, 2017.
- [7] R. Bailo, J. A. Carrillo, H. Murakawa, and M. Schmidtchen. Convergence of a fully discrete and energy-dissipating finite-volume scheme for aggregation-diffusion equations. *Math. Models Methods Appl. Sci.*, 30(13):2487–2522, 2020.
- [8] J.W. Barrett and J.F. Blowey. Finite element approximation of a nonlinear cross-diffusion population model. *Numer. Math.*, 98(2):195–221, 2004.
- [9] J. Bear. *J. Bear. Hydraulics of groundwater*. McGraw-Hill, Inc., New York, 1979.
- [10] J.-D. Benamou and Y. Brenier. A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem. *Numer. Math.*, 84:375–393, 2000.
- [11] J.-D. Benamou, Y. Brenier, and K. Guittet. Numerical analysis of a multi-phasic mass transport problem. *Contemp. Math.*, 353:1–17, 2004.
- [12] J.-D. Benamou, G. Carlier, and M. Laborde. An augmented Lagrangian approach to Wasserstein gradient flows and applications. In *Gradient flows: from theory to application*, volume 54 of *ESAIM Proc. Surveys*, pages 1–17. EDP Sci., Les Ulis, 2016.
- [13] M. Bessemoulin-Chatard. A finite volume scheme for convection-diffusion equations with nonlinear diffusion derived from the Scharfetter-Gummel scheme. *Numer. Math.*, 121(4):637–670, 2012.
- [14] M. Bessemoulin-Chatard, C. Chainais-Hillairet, and F. Filbet. On discrete functional inequalities for some finite volume schemes. *IMA J. Numer. Anal.*, 35:1125–1149, 2015.
- [15] M. Bessemoulin-Chatard, C. Chainais-Hillairet, and H. Mathis. Analysis of numerical schemes for semiconductors energy-transport models. *Hal-02940224*, 2020.
- [16] M. Bessemoulin-Chatard, C. Chainais-Hillairet, and M.-H. Vignal. Study of a finite volume scheme for the drift-diffusion system. Asymptotic behavior in the quasi-neutral limit. *SIAM J. Numer. Anal.*, 52(4):1666–1691, 2014.
- [17] M. Bessemoulin-Chatard and A. Jüngel. A finite volume scheme for a Keller-Segel model with additional cross-diffusion. *IMA J. Numer. Anal.*, 34(1):96–122, 2014.
- [18] M. Bessemoulin-Chatard, G. Lissoni, and H. Mathis. Numerical analysis of DDFV schemes for semiconductors energy-transport models. *Hal-03080236*, 2020.
- [19] S. Bosma, F. P. Hamon, B. T. Mallison, and H. A. Tchelepi. Smooth implicit hybrid upwinding for compositional multiphase flow in porous media. *arXiv:2106.03317*, 2021.

- [20] M. Braukhoff and A. Jüngel. Entropy-dissipating finite-difference schemes for nonlinear fourth-order parabolic equations. *Discrete Cont. Dyn. Sys. B*, 26:3335–3355, 2021.
- [21] M. Braukhoff, I. Perugia, and P. Stocker. An entropy structure preserving space-time Galerkin method for cross-diffusion systems. *arXiv:2006.13069*, 2020.
- [22] C. Cancès, C. Chainais-Hillairet, J. Fuhrmann, and B. Gaudeul. A numerical analysis focused comparison of several finite volume schemes for a Unipolar Degenerated Drift-Diffusion Model. *IMA J. Numer. Anal.*, 41:271–314, 2021.
- [23] C. Cancès, C. Chainais-Hillairet, A. Gerstenmayer, and A. Jüngel. Finite-volume scheme for a degenerate cross-diffusion model motivated from ion transport. *Numer. Meth. Partial Diff. Eqs.*, 35:545–575, 2019.
- [24] C. Cancès, C. Chainais-Hillairet, M. Herda, S. Krell. Large time behavior of nonlinear finite volume schemes for convection-diffusion equations. *SIAM J. Numer. Anal.*, 58(5):2544–2571, 2020.
- [25] C. Cancès, V. Ehrlacher, and L. Monasse. Finite volumes for the Maxwell–Stefan cross-diffusion system. Submitted for publication, *arXiv:2007.09951*, 2020.
- [26] C. Cancès, T. O. Gallouët, M. Laborde, and L. Monsaingeon. Simulation of multiphase porous media flows with minimizing movement and finite volume schemes. *European J. Appl. Math*, 30(6):1123–1152, 2019
- [27] C. Cancès, T. O. Gallouët, and L. Monsaingeon. Incompressible immiscible multiphase flows in porous media: a variational approach. *Analysis & PDE*, 10:1845–1876, 2017.
- [28] C. Cancès and B. Gaudeul. A convergent entropy diminishing finite volume scheme for a cross-diffusion system. *SIAM J. Numer. Anal.*, 58(5):2684–2710, 2020.
- [29] C. Cancès and C. Guichard. Convergence of a nonlinear entropy diminishing Control Volume Finite Element scheme for solving anisotropic degenerate parabolic equations. *Math. Comp.*, 85(298):549–580, 2016.
- [30] C. Cancès and C. Guichard. Numerical analysis of a robust free energy diminishing finite volume scheme for parabolic equations with gradient structure. *Found. Comput. Math.*, 17(6):1525–1584, 2017.
- [31] C. Cancès and D. Maltese. A gravity current model with capillary trapping for oil migration in multilayer geological basins. *SIAM J. Appl. Math*, 81:454–484, 2021.
- [32] C. Cancès and D. Matthes. Construction of a two-phase flow with singular energy by gradient flow methods. *arXiv:2003.07632*, 2020.
- [33] C. Cancès, D. Matthes, and F. Nabet. A two-phase two-fluxes degenerate Cahn-Hilliard model as constrained Wasserstein gradient flow. *Arch. Rational Mech. Anal.*, 232:837–866, 2019.
- [34] C. Cancès and F. Nabet. Finite Volume approximation of a two-phase two fluxes degenerate Cahn-Hilliard model. *ESAIM Math. Model. Numer. Anal.*, 55:969–1003, 2021.
- [35] J. A. Carrillo, F. Filbet, and M. Schmidtchen. Convergence of a finite volume scheme for a system of interacting species with cross-diffusion. *Numer. Math.*, 145:473–511, 2020.
- [36] C. Chainais-Hillairet. Entropy method and asymptotic behaviours of finite volume schemes. In *Finite volumes for complex applications VII. Methods and theoretical aspects*, volume 77 of *Springer Proc. Math. Stat.*, pages 17–35. Springer, Cham, 2014.
- [37] C. Chainais-Hillairet, A. Jüngel, and P. Shpartko. A finite-volume scheme for a spinorial matrix drift-diffusion model for semiconductors. *Numer. Meth. Part. Diff. Eqs.*, 32:819–846, 2016.
- [38] C. Chainais-Hillairet, J.-G. Liu, and Y.-J. Peng. Finite volume scheme for multi-dimensional drift-diffusion equations and convergence analysis. *ESAIM Math. Model. Numer. Anal.*, 37:319–338, 2003.
- [39] X. Chen, E. Daus, and A. Jüngel. Global existence analysis of cross-diffusion population systems for multiple species. *Arch. Ration. Mech. Anal.* 227:715–747, 2018.
- [40] E. S. Daus, A. Jüngel, and A. Zurek. Convergence of a finite-volume scheme for a degenerate-singular cross-diffusion system for biofilms. *IMA J. Numer. Anal.*, doi:10.1093/imanum/draa040, 2020.
- [41] C. Choquet, M. M. Diédhiou, and C. Rosier. Derivation of a sharp-diffuse interfaces model for seawater intrusion in a free aquifer. *SIAM J. of Appl. Math.*, 76:138–158, 2016.
- [42] K. Deimling. *Nonlinear Functional Analysis*. Springer, Berlin, 1985.
- [43] J. Dupuit. *Études théoriques et pratiques sur le mouvement des eaux dans les canaux découverts et à travers les terrains perméables*. Dunod, Paris, 1863.
- [44] R. Eymard and T. Gallouët. H -convergence and numerical schemes for elliptic problems. *SIAM J. Numer. Anal.*, 41:539–562, 2003.
- [45] R. Eymard, T. Gallouët, and R. Herbin. Finite volume methods. Ciarlet, P. G. (ed.) et al., in *Handbook of numerical analysis*. North-Holland, Amsterdam, pages 713–1020, 2000.
- [46] R. Eymard, R. Herbin and A. Michel. Mathematical study of a petroleum-engineering scheme. *ESAIM Math. Model. Numer. Anal.*, 37:937–972, 2003.
- [47] J. Fuhrmann. Existence and uniqueness of solutions of certain systems of algebraic equations with off-diagonal nonlinearity. *Appl. Numer. Math.*, 37:359–370, 2001.
- [48] G. Galiano, M. Garzón, and A. Jüngel. Analysis and numerical solution of a nonlinear cross-diffusion system arising in population dynamics. *RACSAM Rev. R. Acad. Cien. Ser. A*, 95:281–295, 2001.
- [49] B. Gaudeul and J. Fuhrmann. Entropy and convergence analysis for two finite volume schemes for a Nernst-Planck-Poisson system with ion volume constraints. *HAL: hal-03129529*, 2021.
- [50] A. Gerstenmayer and A. Jüngel. Comparison of a finite-element and finite-volume scheme for a degenerate cross-diffusion system for ion transport. *Comput. Appl. Math.*, 38(3):Art. 108, 23, 2019.

- [51] A. Glitzky. Exponential decay of the free energy for discretized electro-reaction–diffusion systems. *Nonlinearity*, 21: 1989–2009, 2009.
- [52] M. J. Golding, J. A. Neufeld, M. A. Hesse, and Huppert. Two-phase gravity currents in porous media. *J. Fluid Mech.*, 678:248–270, 2011.
- [53] I. Gunn and A. W. Woods. On the flow of buoyant fluid injected into a confined, inclined aquifer. *J. Fluid Mech.*, 672:109–129, 2011.
- [54] M. Heida. Convergences of the squareroot approximation scheme to the Fokker-Planck operator *Math. Models Methods Appl. Sci.*, 28(13):2599–2635, 2018.
- [55] H. E. Huppert and A. W. Woods. Gravity-driven flows in porous layers. *J. Fluid Mech.*, 292:55–69, 1995.
- [56] X. Huo, H. Liu, A. E. Tzavaras, and S. Wang. An energy stable and positivity-preserving scheme for the Maxwell-Stefan diffusion system. *arXiv:2005.08062*, 2020.
- [57] M. Ibrahim and M. Saad. On the efficacy of a control volume finite element method for the capture of patterns for a volume-filling chemotaxis model. *Comput. Math. Appl.*, 68(9):1032–1051, 2014.
- [58] M. Jazar and R. Monneau. Derivation of seawater intrusion models by formal asymptotics. *SIAM J. Appl. Math.*, 74:1152–1173, 2014.
- [59] R. Jordan, D. Kinderlehrer, and F. Otto. The variational formulation of the Fokker-Planck equation. *SIAM J. Math. Anal.*, 29(1):1–17, 1998.
- [60] A. Jüngel. The boundedness-by-entropy method for cross-diffusion systems. *Nonlinearity*, 28:1963–2001, 2015.
- [61] A. Jüngel. *Entropy Methods for Diffusive Partial Differential Equations*. BCAM Springer Briefs, Springer, 2016.
- [62] A. Jüngel and O. Leingang. Convergence of an implicit Euler Galerkin scheme for Poisson-Maxwell-Stefan systems. *Adv. Comput. Math.*, 45(3):1469–1498, 2019.
- [63] A. Jüngel and I. V. Stelzer. Existence analysis of Maxwell–Stefan systems for multicomponent mixtures. *SIAM J. Math. Anal.*, 45(4):2421–2440, 2013.
- [64] A. Jüngel and A. Zurek. A convergent structure-preserving finite-volume scheme for the Shigesada-Kawasaki-Teramoto population system. *arXiv:2011.08731*, 2020.
- [65] P. Laurençot and B.-V. Matioc. A gradient flow approach to a thin film approximation of the Muskat problem *Calc. Var. Partial Differential Equations*, 47:319–341, 2013.
- [66] J. Maas and D. Matthes. Long-time behavior of a finite volume discretization for a fourth order diffusion equation. *Nonlinearity*, 29:1992–2023, 2016.
- [67] J. M. Nordbotten and M. Celia. *Geological Storage of CO₂: Modeling Approaches for Large scale Simulation*. John Wiley & Sons, New Jersey, 2012.
- [68] F. Otto and W. E. Thermodynamically driven incompressible fluid mixtures. *J. Chem. Phys.*, 107:10177–10184, 1997.
- [69] M. A. Peletier. *Variational Modelling: Energies, gradient flows, and large deviations*. Lecture notes, *arXiv:1402.1990*, 2014.
- [70] C. Seis and A. Schlichting. The Scharfetter-Gummel scheme for aggregation-diffusion equations *arXiv:2004.13981*, 2020.
- [71] N. Shigesada, K. Kawasaki, and E. Teramoto. Spatial segregation of interacting species. *J. Theor. Biol.*, 79:83–99, 1979.
- [72] Z. Sun, J. A. Carrillo, and C.-W. Shu. A discontinuous Galerkin method for nonlinear parabolic equations and gradient flow problems with interaction potentials. *J. Comput. Phys.*, 352:76–104, 2018.
- [73] Z. Sun, J. A. Carrillo, and C.-W. Shu. An entropy stable high-order discontinuous Galerkin method for cross-diffusion gradient flow systems. *Kinet. Relat. Models*, 12(4):885–908, 2019.

INRIA, UNIV. LILLE, CNRS, UMR 8524–LABORATOIRE PAUL PAINLEVÉ, F-59000 LILLE
 Email address: clement.cances@inria.fr

UNIVERSITÉ DE TECHNOLOGIE DE COMPIÈGNE, LMAC, 60200 COMPIÈGNE, FRANCE
 Email address: antoine.zurek@utc.fr