POSITIVE NONLINEAR CONTROL VOLUME FINITE ELEMENT SCHEME FOR AN ANISOTROPIC SEAWATER INTRUSION MODEL WITH CROSS-DIFFUSION IN AN UNCONFINED AQUIFER

AHMED AIT HAMMOU OULHAJ AND DAVID MALTESE

Abstract. We consider a degenerate parabolic system modelling the flow of fresh and saltwater in an anisotropic porous medium in the context of seawater intrusion. We propose and analyze a nonlinear Control Volume Finite Element scheme. This scheme ensures the nonnegativity of the discrete solution without any restriction on the mesh and on the anisotropy tensor. Moreover It also provides a control on the entropy. Based on these nonlinear stability results, we show that the scheme converges towards a weak solution to the problem. Numerical results are provided to illustrate the behavior of the model and of the scheme.

Keywords. Unsaturated porous media flow, seawater intrusion, nonlinear discretization, entropy stability, convergence analysis, cross-diffusion, unconfined aquifer

AMS subjects classification. 65M12, 65M08, 76S05

1. Introduction

1.1. **Presentation of the continuous problem.** In this paper, we are interested in degenerate nonlinear parabolic cross-diffusion systems modelling the so-called seawater intrusion over general mesh, with anisotropic and heterogeneous diffusion tensors. The model we are interested in is derived by Jazar and Monneau in [33], with isotropic tensor. The authors consider the Dupuit approximation of an unsaturated immiscible two-phase (freshwater and saltwater) within an unconfined aquifer assuming that the interface between both fluids is sharp (the fluids occupy disjoint regions), see also [25, 44] for alternative derivations of the same model. This yields a 2D reduced model obtained from a full 3D model where the unknowns are the heights of the fluid layers. More precisely the impermeable interface between the saltwater and the bedrock is set at $\{z = b(x)\}$, whereas the height of the freshwater (resp. saltwater) layer is denoted by $\{z = f(t,x)\}$ (resp. $\{z = g(t,x)\}$), see Figure 1. We refer to [9, 10, 11, 43, 42, 5] for more details about this approach. In the literature, there exists other modelling approaches (see [19, 40, 41, 20, 21]). The model proposed in [33] with anisotropic tensor is

$$\begin{cases} \partial_t f - \nabla.\big(\nu\Lambda f\nabla(f+g+b)\big) &= 0 \quad \text{in} \quad \Omega\times(0,t_f), \\ \partial_t g - \nabla.\big(\Lambda g\nabla(\nu f+g+b)\big) &= 0 \quad \text{in} \quad \Omega\times(0,t_f), \end{cases} \tag{1}$$

This work was supported by the GeoPor project funded by the French National Research Agency (ANR) with the grant ANR-13-JS01-0007-01 (project GEOPOR) .

where $\Omega \subset \mathbb{R}^2$ is a polygonal open bounded subset and $t_f > 0$ is a finite time horizon. The parameter ν is given by

$$\nu = \frac{\rho_f}{\rho_s} \in (0,1), \tag{2}$$

where ρ_s (resp. ρ_f) is the mass density of the fluid saltwater (resp. freshwater) (assumed to be constant with $0 < \rho_f < \rho_s$).

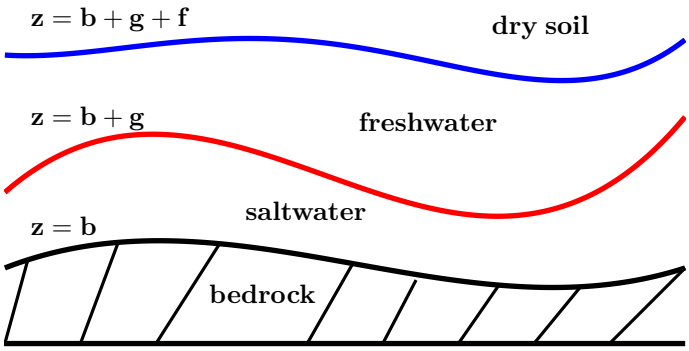


FIGURE 1. Description of an unconfined aquifer

It is supplemented with no-flux boundary conditions on $\partial\Omega \times (0, t_f)$

$$f\Lambda\nabla(f+g+b)\cdot\mathbf{n} = 0, \quad g\Lambda\nabla(\nu f+g+b)\cdot\mathbf{n} = 0, \tag{3}$$

where $\mathbf{n}$ is the unit normal to the boundary $\partial\Omega$. Initial data are

$$f_{|t=0} = f_0, \quad g_{|t=0} = g_0, \tag{4}$$

where f_0 and g_0 are two mesurable functions such that

$$f_0, g_0 \geq 0, \quad \text{for a.e } x \in \Omega. \tag{5}$$

The permeability tensor Λ belongs $(L^\infty(\Omega))^{2\times 2}$, and it is supposed to be symmetric and uniformly elliptic on Ω, i.e, there exists $(\overline{\Lambda}, \underline{\Lambda}) \in \mathbb{R}_+^* \times \mathbb{R}_+^*$ such that

$$\underline{\Lambda}|\mathbf{v}|^2 \leq \Lambda(\mathbf{x})\mathbf{v}\cdot\mathbf{v} \leq \overline{\Lambda}|\mathbf{v}|^2, \quad \forall \mathbf{v} \in \mathbb{R}^2, \quad \text{for a.e.} \quad \mathbf{x} \in \Omega.$$

In the isotropic case, let us mention that the problems of kind (1), have been the object of several studies. The authors in [25, 26] studied the classical solutions of the system (1) (with $b = 0$). Moreover, nonnegative weak solutions are established under different assumptions in [24, 34, 35]. We recall (see [24, 34, 35]) the definition of entropy functional :

$$\mathfrak{E}(f,g) = \int_\Omega \Big[\Gamma(g) + \frac{1}{\nu}\Gamma(f)\Big]\,\mathrm{d}\boldsymbol{x}, \quad \text{where } \Gamma(s) = s\log s - s + 1.$$

Multiplying (formally) the first equation of (1) by $\dfrac{1}{\nu}\log f$ and the second equation by $\log g$, integrating over Ω and summing both equations

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathfrak{E}(f,g) + \int_\Omega \Lambda(\boldsymbol{x})\Big[|\nabla f|^2 + |\nabla g|^2\Big]\,\mathrm{d}\boldsymbol{x} + (\nu+1)\int_\Omega \Lambda(\boldsymbol{x})\nabla f\cdot\nabla g\,\mathrm{d}\boldsymbol{x}$$
$$= -\int_\Omega \Lambda(\boldsymbol{x})\nabla b\cdot(\nabla f + \nabla g)\,\mathrm{d}\boldsymbol{x}.$$

Since Λ is symmetric and positive, using the Hölder inequality and the Young inequality one has

$$
\begin{aligned}
-\int_\Omega \Lambda(\boldsymbol{x})\nabla b\cdot(\nabla f+\nabla g)\,\mathrm{d}\boldsymbol{x} &= -\int_\Omega \sqrt{\Lambda(\boldsymbol{x})}\nabla b\cdot\sqrt{\Lambda(\boldsymbol{x})}(\nabla f+\nabla g)\,\mathrm{d}\boldsymbol{x}\\
&\le \|\sqrt{\Lambda(\boldsymbol{x})}\nabla b\|_{L^2(\Omega)}\|\sqrt{\Lambda(\boldsymbol{x})}(\nabla f+\nabla g)\|_{L^2(\Omega)}\\
\le \frac{\varepsilon}{2}\int_\Omega \Lambda(\boldsymbol{x})|\nabla b|^2\,\mathrm{d}\boldsymbol{x} &+ \frac{1}{2\varepsilon}\int_\Omega \Lambda(\boldsymbol{x})\left(|\nabla f|^2+|\nabla g|^2\right)\mathrm{d}\boldsymbol{x}+\frac{1}{\varepsilon}\int_\Omega \Lambda(\boldsymbol{x})\nabla f\cdot\nabla g\,\mathrm{d}\boldsymbol{x},
\end{aligned}
$$

for $\varepsilon>0$. We choose $\varepsilon=\dfrac{1}{\nu+1}$, yields the classical entropy/dissipation property :

$$
\frac{\mathrm{d}}{\mathrm{d}t}\mathfrak{E}(f,g)+\frac{1-\nu}{2}\int_\Omega \Lambda(\boldsymbol{x})\Big[|\nabla f|^2+|\nabla g|^2\Big]\,\mathrm{d}\boldsymbol{x}\le\frac{\overline{\Lambda}}{2(\nu+1)}\int_\Omega|\nabla b|^2\,\mathrm{d}\boldsymbol{x}. \tag{6}
$$

Let us mention that the cross-diffusion systems are extensively presented in different domain as ecology, biology, chemistry and others. In [30] the author propose and analyze a finite volume scheme for the Patlak-Keller-Segel (PKS) chemotaxis model. Moreover, a finite-volume scheme for a PKS system with additional cross diffusion and discrete entropy dissipation property was investigated in [12]. We refer to [7] for the analysis of a finite volume method for a cross diffusion model in population dynamics. See [39, 37] for the numerical analysis for a seawater intrusion problem in a unconfined aquifer with finite element method approximation. In [1] the authors propose a finite element method and a finite volume method and compare the results given by these two methods. In [22] the authors address the question of global existence for the sharp interface approach. For an analysis of a finite volume scheme for two-phase immiscible flow in porous media, used in petroleum engineering, we can refer to these papers [38, 29]. An upwind two-point flux approximation was used in [2, 3] for a seawater intrusion cross-diffusion model. An implicit Euler finite-volume scheme for a degenerate cross-diffusion system describing the ion transport through biological membranes is analyzed in [13]. We refer to [17] for a positive nonlinear CVFE scheme for degenerate anisotropic PKS system.

The method we study here was designed on the following specifications:

(a) to handle anisotropic and heterogeneous anisotropy tensors;
(b) to preserve the nonnegativity of the discrete solution;
(c) to preserve the control of the entropy;
(d) to conserve locally the mass of fluid;
(e) to converge towards the solution to the continuous problem.

The scheme we propose belongs to the family of the so-called Control Volume Finite Element schemes introduced in the context of porous media flows by Forsyth [31, 32]. Roughly speaking, it consists in an interpretation of Finite Elements with mass lumping as a locally conservative method on dual cells. This is an extension of the one studied in [14, 15, 4]. Especially, the diffusion terms are discretized by means of a conforming piecewise linear finite element method on a primal triangular mesh and on a suitable upwinding of the mobility that allows to preserve the physical bounds (but not the monotonicity as in [32]). Moreover, we show that our method provides a control on the entropy and that this control is sufficient to perform a convergence proof based on compactness arguments.

2. The numerical scheme

2.1. Discretization of Q_{t_f}.

2.1.1. *Discretizations of* Ω. The CVFE method requires the introduction of two different space discretizations of Ω: a *primal triangular mesh* and a *dual barycentric mesh*.

The *primal triangular mesh* is denoted by $\mathcal{T}$. It is a conformal triangular discretization of the polygonal domain Ω, consisting in open bounded separated triangles satisfying $\bigcup_{T\in\mathcal{T}} \overline{T} = \overline{\Omega}$. For $T \in \mathcal{T}$, we denote by $\boldsymbol{x}_T$ the center of gravity of T, by h_T the diameter of the triangle T, and by ρ_T the diameter of the largest ball inscribed in the triangle T. Then, we define the mesh diameter h and the mesh regularity $\theta_{\mathcal{T}}$ by

$$h = \max_{T\in\mathcal{T}} h_T, \quad \theta_{\mathcal{T}} = \max_{T\in\mathcal{T}} \frac{h_T}{\rho_T}.$$

We denote by $\mathcal{V}$ the set of the vertices of the discretization $\mathcal{T}$, located at positions $(\boldsymbol{x}_K)_{K\in\mathcal{V}}$. The set $\mathcal{E}$ of the edges of $\mathcal{T}$ is made of straight segments σ joining two vertices of $\mathcal{V}$. Given $T, T' \in \mathcal{T}$, we assume that $\overline{T} \cap \overline{T'}$ is either empty, or it is reduced to $\boldsymbol{x}_K$ for some $K \in \mathcal{V}$, or it consists in an edge σ belonging $\mathcal{E}$. For $T \in \mathcal{T}$, we denote by $\mathcal{E}_T$ the set of the edges of T: $\bigcup_{\sigma\in\mathcal{E}_T} \overline{\sigma} = \partial T$. We assume that $\mathcal{E} = \bigcup_{T\in\mathcal{T}} \mathcal{E}_{\mathcal{T}}$. Given two vertices $K, L \in \mathcal{V}$ of a triangle T, then the edge joining $\boldsymbol{x}_K$ and $\boldsymbol{x}_L$ is denoted by σ_{KL}. For $K \in \mathcal{V}$, one denotes by $\mathcal{T}_K$ the subset of $\mathcal{T}$ made the triangles admitting K as a vertex, by $\mathcal{E}_K$ the set of edges having the vertex K as an extremity, and by $\mathcal{V}_K$ the subset of $\mathcal{V}$ such that, if $L \in \mathcal{V}_K$, then $[\boldsymbol{x}_K, \boldsymbol{x}_L]$ is an edge of $\mathcal{E}_K$.

Once the *primal triangular mesh* has been built, we can define its *dual barycentric mesh* $\mathcal{M}$ as follows. To each $K \in \mathcal{V}$, we associate a cell ω_K whose vertices are the isobarycenters $\boldsymbol{x}_T$ of the triangles $T \in \mathcal{T}_K$ and the isobarycenters $\boldsymbol{x}_\sigma$ of the edges $\sigma \in \mathcal{E}_K$. Note that $\overline{\Omega} = \bigcup_{K\in\mathcal{V}} \overline{\omega}_K$. We refer to Figure 2 for an illustration of the primary and dual barycentric meshes. The 2-dimensional Lebesgue measure of ω_K is denoted by m_K.

Let us now introduce some useful functional spaces. The space $V_{\mathcal{T}} \subset \mathcal{C}(\overline{\Omega})$ is made of piecewise affine functions on the primal mesh, i.e.,

$$V_{\mathcal{T}} = \{f \in H^1(\Omega) \mid f_{|_T} \text{ is affine}, \ \forall T \in \mathcal{T}\}.$$

For all $K \in \mathcal{V}$, we denote by e_K the unique element of $V_{\mathcal{T}}$ such that $e_K(\boldsymbol{x}_K) = 1$ and $e_K(\boldsymbol{x}_L) = 0$ if $L \in \mathcal{V} \setminus \{K\}$. The geometrical construction of ω_K ensures that

$$\int_\Omega e_K(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = \int_{\omega_K} \mathrm{d}\boldsymbol{x} =: m_K, \qquad \forall K \in \mathcal{V}.$$

We can also define the set of the piecewise constant functions on $\mathcal{M}$, $X_{\mathcal{M}}$, by

$$X_{\mathcal{M}} = \{f : \Omega \longrightarrow \mathbb{R} \text{ measurable} \mid f_{|_{\omega_K}} \in \mathbb{R} \text{ is constant}, \quad \forall K \in \mathcal{V}\}.$$

Given a vector $(u_K)_{K\in\mathcal{V}} \in \mathbb{R}^{\#\mathcal{V}}$, there exists a unique $u_{\mathcal{T}} \in V_{\mathcal{T}}$ and a unique $u_{\mathcal{M}} \in X_{\mathcal{M}}$ such that $u_{\mathcal{T}}(\boldsymbol{x}_K) = u_{\mathcal{M}}(\boldsymbol{x}_K) = u_K$ for all $K, L \in \mathcal{V}$. Let us note that $u_{\mathcal{T}} = \sum_{K\in\mathcal{V}} u_K e_K$. Moreover, for all $q \in [1, \infty)$, there exist C_1 and C_2 depending only on q and on $\theta_{\mathcal{T}}$ such that

$$C_1 \|u_{\mathcal{T}}\|_{L^q(\Omega)} \le \|u_{\mathcal{M}}\|_{L^q(\Omega)} \le C_2 \|u_{\mathcal{T}}\|_{L^q(\Omega)}, \qquad \forall\ (u_K)_{K\in\mathcal{V}} \in \mathbb{R}^{\#\mathcal{V}}. \tag{7}$$

A proof of the above inequalities can be found for instance in [16, Lemma A.6].

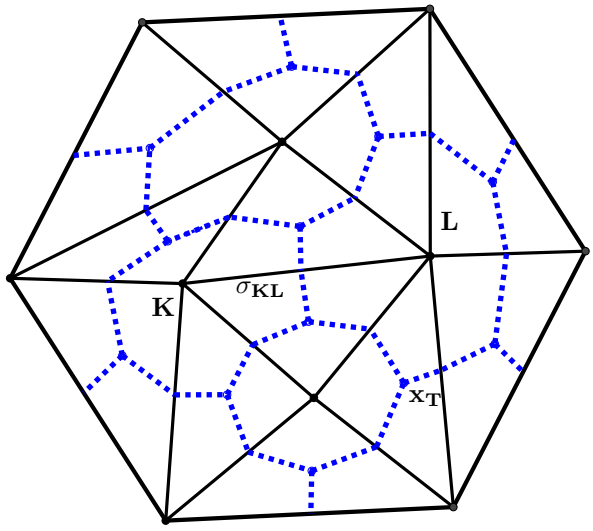


FIGURE 2. The triangular mesh $\mathcal{T}$ (solid line) and its corresponding dual barycentric dual mesh $\mathcal{M}$ (dashed line).

2.1.2. *Space-time discretizations.* In order to avoid heavier notations, we restrict our study to the case of a uniform time discretization of $(0, t_f)$. However, all the results presented in this paper can be extended to general time discretizations without any technical difficulty. In what follows, we assume that the spatial mesh is fixed and does not change with the time step.

Let N be a nonnegative integer, then we define $\Delta t = \dfrac{t_f}{N+1}$, and $t_n = n\Delta t$ for all $n \in \{0, ..., N+1\}$, so that $t_0 = 0$, and $t_{N+1} = t_f$.

We define the space and time discrete spaces $V_{\mathcal{T},\Delta t}$ and $X_{\mathcal{M},\Delta t}$ as the set of piecewise constant functions in time with values in $V_{\mathcal{T}}$ and $X_{\mathcal{M}}$ respectively:

$$V_{\mathcal{T},\Delta t} = \{f : Q_{t_f} \to \overline{\mathbb{R}} \mid f(x,t) = f(x,t^{n+1}) \in V_{\mathcal{T}}, \quad \forall t \in (t_n, t_{n+1}]\},$$
$$X_{\mathcal{M},\Delta t} = \{f : Q_{t_f} \to \overline{\mathbb{R}} \mid f(x,t) = f(x,t^{n+1}) \in X_{\mathcal{M}}, \quad \forall t \in (t_n, t_{n+1}]\}.$$

For a given $(u_K^{n+1})_{n\in\{0,...,N\},K\in\mathcal{V}} \in \mathbb{R}^{(N+1)\#\mathcal{V}}$, we denote by $u_{\mathcal{T},\Delta t}$ and $u_{\mathcal{M},\Delta t}$ the unique elements of $V_{\mathcal{T},\Delta t}$ and $X_{\mathcal{M},\Delta t}$ respectively such that

$$u_{\mathcal{T},\Delta t}(x_K, t) = u_{\mathcal{M},\Delta t}(x_K, t) = u_K^{n+1}, \quad \forall K \in \mathcal{V},\ \forall t \in (t_n, t_{n+1}]. \tag{8}$$

2.2. **Finite elements.** The method we propose, and more generally the CVFE method, is based on P_1-finite elements. We introduce in this section the technical material that is needed in order to define the scheme and to perform its analysis. We define the transmissibility coefficients

$$a_{KL}^T = -\int_T \Lambda \nabla e_K \cdot \nabla e_L \, \mathrm{d}\boldsymbol{x} = a_{LK}^T, \quad \forall T \in \mathcal{T},\ \forall (K,L) \in \mathcal{V}^2, \tag{9}$$

and

$$a_{KL} = a_{LK} = -\int_\Omega \Lambda \nabla e_K \cdot \nabla e_L \, \mathrm{d}\boldsymbol{x} = \sum_{T\in\mathcal{T}} a_{KL}^T, \quad \forall (K,L) \in \mathcal{V}^2. \tag{10}$$

Note that $a_{KL} = 0$ unless $\sigma_{KL} \in \mathcal{E}$. Moreover, since $\sum_{K\in\mathcal{V}} \nabla e_K = 0$, we have that :

$$-a_{KK} = \sum_{L\neq K} a_{KL} > 0. \tag{11}$$

As a consequence of (10)-(11), given $u_\mathcal{T}$ and $v_\mathcal{T}$ two elements of $V_\mathcal{T}$, one has

$$\begin{aligned}\int_\Omega \Lambda \nabla u_\mathcal{T} \cdot \nabla v_\mathcal{T} \, \mathrm{d}\boldsymbol{x} &= \sum_{\sigma_{KL}\in\mathcal{E}} a_{KL}(u_K - u_L)(v_K - v_L) \\ &= \sum_{T\in\mathcal{T}} \sum_{\sigma_{KL}\in\mathcal{E}_T} a_{KL}^T (u_K - u_L)(v_K - v_L).\end{aligned} \tag{12}$$

The following lemma plays a crucial role in the numerical analysis carried out in this paper. We refer to [15, Lemma 3.2] for its proof.

Lemma 2.1. *There exists C_3 depending only on $\theta_\mathcal{T}$, $\overline{\Lambda}$ and $\underline{\Lambda}$ such that, for all $u_\mathcal{T} \in V_\mathcal{T}$, one has*

$$\sum_{\sigma_{KL}\in\mathcal{E}} |a_{KL}|(u_K - u_L)^2 \le \sum_{T\in\mathcal{T}} \sum_{\sigma_{KL}\in\mathcal{E}_T} |a_{KL}^T|(u_K - u_L)^2 \le C_3 \int_\Omega \Lambda \nabla u_\mathcal{T} \cdot \nabla u_\mathcal{T} \, \mathrm{d}\boldsymbol{x}.$$

We will need the following control of a_{KL}.

Lemma 2.2. *There exists C_4 depending only on $\theta_\mathcal{T}$ and $\overline{\Lambda}$ such that, for any $\sigma_{KL} \in \mathcal{E}$ where $(K,L) \in \mathcal{V}^2$ one has*

$$|a_{KL}| \le C_4.$$

Proof. Using classical properties of the affine change of variable between simplexes, one has for any $a \in \mathcal{V}$ and for any $T \in \mathcal{T}$

$$\|\nabla e_a\|_{L^2(T)} \le 2\frac{\sqrt{m_T}}{\rho_T}. \tag{13}$$

We have also for any $a \in \mathcal{V}$

$$\mathrm{supp} e_a \subset \cup_{T\in\mathcal{T}_a} \overline{T}.$$

Let $(T,T') \in \mathcal{T}^2$ such that $\sigma_{KL} = \overline{T} \cap \overline{T'}$. In the case where T and T' are distinct we have

$$a_{KL} = -\int_T \Lambda \nabla e_K \cdot \nabla e_L \, \mathrm{d}\boldsymbol{x} - \int_{T'} \Lambda \nabla e_K \cdot \nabla e_L \, \mathrm{d}\boldsymbol{x}.$$

Consequently using (13) and Hölder's inequality we obtain that

$$|a_{KL}| \le 4\overline{\Lambda}\left(\frac{m_T}{\rho_T^2} + \frac{m_{T'}}{\rho_{T'}^2}\right).$$

In the case where T and T' are equal we have

$$a_{KL} = -\int_T \Lambda \nabla e_K \cdot \nabla e_L \, \mathrm{d}\boldsymbol{x},$$

which gives using also (13) and Hölder's inequality

$$|a_{KL}| \le 4\overline{\Lambda}\frac{m_T}{\rho_T^2}.$$

By virtue of the definition of $\theta_{\mathcal{T}}$ one has for any $T \in \mathcal{T}$

$$\frac{\pi}{4}\rho_T^2 \leq m_T \leq \pi\theta_{\mathcal{T}}^2\rho_T^2,$$

which gives the expected results where $C_4 = 8\pi\|\Lambda\|_{L^\infty(\Omega)^{2\times 2}}\theta_{\mathcal{T}}^2$. □

The following Lemma and its proof can be found in [15].

Lemma 2.3. *Let $u_{\mathcal{T}} \in V_{\mathcal{T}}$, then one defines the piecewise constant functions $\overline{u}_{\mathcal{T}}$ and $\underline{u}_{\mathcal{T}}$ mapping Ω to $\mathbb{R}$ by*

$$\overline{u}_{\mathcal{T}}(\boldsymbol{x}) = \sup_{\boldsymbol{x}\in T} u(\boldsymbol{x}) \text{ for any } \boldsymbol{x} \in T,$$

$$\underline{u}_{\mathcal{T}}(\boldsymbol{x}) = \inf_{\boldsymbol{x}\in T} u(\boldsymbol{x}) \text{ for any } \boldsymbol{x} \in T.$$

Then there exists a constant C_5 only depending on $1 \leq p < \infty$ such that

$$\|\overline{u}_{\mathcal{T}} - \underline{u}_{\mathcal{T}}\|_{L^p(\Omega)} \leq C_5 h\|\nabla u_{\mathcal{T}}\|_{L^p(\Omega)^2}.$$

The following Lemma is a consequence of the previous Lemma.

Lemma 2.4. *Let $(u_K)_{K\in\mathcal{V}} \in \mathbb{R}^{\#\mathcal{V}}$. Let $u_{\mathcal{T}} \in V_{\mathcal{T}}$ and $u_{\mathcal{M}} \in X_{\mathcal{M}}$ be the corresponding piecewise linear and piecewise constant reconstructions defined in §2.1.1. Then there exists a constant C_6 only depending on $1 \leq p < \infty$ such that*

$$\|u_{\mathcal{T}} - u_{\mathcal{M}}\|_{L^p(\Omega)} \leq C_6 h\|\nabla u_{\mathcal{T}}\|_{L^p(\Omega)^2}.$$

2.3. **The nonlinear CVFE scheme.** In this section, we explicit the discretization of the problem (1) we will study in this paper. The time discretization relies on backward Euler scheme, while the space discretization relies on finite elements with mass lumping and a suitable upwinding of the mobility.

The discretization $s_{\mathcal{M}}^0 \in X_{\mathcal{M}}$ of the initial data is defined by

$$s_K^0 = \frac{1}{m_K}\int_{\omega_K} s_0(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}, \qquad \forall K \in \mathcal{V}, \quad \text{with} \quad s = f \text{ or } g. \tag{14}$$

and we denote $(b_K)_{K\in\mathcal{V}}$ the family given by

$$b_K = b(\boldsymbol{x}_K) \text{ for any } K \in \mathcal{V}.$$

In the sequel, we set

$$u = f + g + b, \qquad v = \nu f + g + b.$$

Let us now introduce the scheme. For all $n \in \{0, ..., N\}$, a solution $\left(f_K^{n+1}\right)_{K\in\mathcal{V}}$ and $\left(g_K^{n+1}\right)_{K\in\mathcal{V}}$ to the scheme at the time step $n+1$ have to satisfy the following

equations : for all $K \in \mathcal{V}$,

$$\frac{f_K^{n+1} - f_K^n}{\Delta t} m_K + \sum_{\sigma_{KL} \in \mathcal{E}_K} \nu f_{KL}^{n+1} a_{KL} (u_K^{n+1} - u_L^{n+1}) = 0, \tag{15a}$$

$$\frac{g_K^{n+1} - g_K^n}{\Delta t} m_K + \sum_{\sigma_{KL} \in \mathcal{E}_K} g_{KL}^{n+1} a_{KL} (v_K^{n+1} - v_K^{n+1}) = 0, \tag{15b}$$

$$u_K^{n+1} = f_K^{n+1} + g_K^{n+1} + b_K, \tag{15c}$$

$$v_K^{n+1} = \nu f_K^{n+1} + g_K^{n+1} + b_K, \tag{15d}$$

$$f_{KL}^{n+1} = \begin{cases} \left(f_K^{n+1}\right)^+ & \text{if} \quad a_{KL}(u_K^{n+1} - u_L^{n+1}) \geq 0, \\ \left(f_L^{n+1}\right)^+ & \text{if} \quad a_{KL}(u_K^{n+1} - u_L^{n+1}) < 0, \end{cases} \tag{15e}$$

$$g_{KL}^{n+1} = \begin{cases} \left(g_K^{n+1}\right)^+ & \text{if} \quad a_{KL}(v_K^{n+1} - v_L^{n+1}) \geq 0, \\ \left(g_L^{n+1}\right)^+ & \text{if} \quad a_{KL}(v_K^{n+1} - v_L^{n+1}) < 0, \end{cases} \tag{15f}$$

with $x^+ = \max(x, 0)$.

This scheme, whose construction is based on finite elements *via* (10), can be interpreted as a finite volume scheme. Indeed denoting by

$$F_{KL}^{n+1} = \nu a_{KL} f_{KL}^{n+1} (u_K^{n+1} - u_L^{n+1}), \quad G_{KL}^{n+1} = a_{KL} g_{KL}^{n+1} (v_K^{n+1} - v_L^{n+1}),$$

the scheme (15) can be rewritten under the locally conservative form on the dual cells ω_K :

$$\begin{cases} F_{KL}^{n+1} + F_{LK}^{n+1} = 0, & \text{for all } \sigma_{KL} \in \mathcal{E}_K \\ \dfrac{f_K^{n+1} - f_K^n}{\Delta t} m_K + \displaystyle\sum_{\sigma_{KL} \in \mathcal{E}_K} F_{KL}^{n+1} = 0, & \text{for all } K \in \mathcal{V}, \\ G_{KL}^{n+1} + G_{LK}^{n+1} = 0, & \text{for all } \sigma_{KL} \in \mathcal{E}_K \\ \dfrac{g_K^{n+1} - g_K^n}{\Delta t} m_K + \displaystyle\sum_{\sigma_{KL} \in \mathcal{E}_K} G_{KL}^{n+1} = 0, & \text{for all } K \in \mathcal{V}. \end{cases} \tag{16}$$

As a straightforward consequence, we can claim that the scheme (15) is globally conservative, i.e.,

$$\sum_{K \in \mathcal{V}} m_K s_K^{n+1} = \sum_{K \in \mathcal{V}} m_K s_K^n = \int_\Omega s_0(\boldsymbol{x}) \mathrm{d}\boldsymbol{x}, \qquad \forall n \geq 0, \quad \text{with } s = f \text{ or } g. \tag{17}$$

2.4. **Main results.** The scheme (15) amounts to a nonlinear system to be solved at each time step. The existence of a solution to this system is therefore non trivial. The first result we highlight is thus the existence of a solution to the scheme (15) and the stability in terms of the discrete entropy defined as follows :

$$\mathfrak{E}^n := \mathfrak{E}(f_K^n, g_K^n) = \sum_{K \in \mathcal{V}} m_K \Big(\frac{1}{\nu} \Gamma(f_K^n) + \Gamma(g_K^n) \Big).$$

Theorem 2.5. *There exists (at least) one solution* $\left(f_K^{n+1}, g_K^{n+1}\right)_{K \in \mathcal{V}, n \in \{0, \dots, N\}}$ *to the scheme* (15). *Moreover,* $f_K^n \geq 0$, $g_K^n \geq 0$ *for all* $K \in \mathcal{V}$ *and for all* $n \in$

$\{0,\dots,N\}$ and there exists C depending only on $\Omega, f_0, g_0, \nu, t_f, \overline{\Lambda}$ and b such that

$$\sup_{n\in\{0,\dots,N\}} \mathfrak{E}^{n+1} + \sum_{n=0}^{N} \Delta t \sum_{\sigma_{KL}\in\mathcal{E}} a_{KL}\Big[(f_K^{n+1} - f_L^{n+1})^2 + (g_K^{n+1} - g_L^{n+1})^2\Big] \leq C.$$

Once we have the discrete solution $\left(f_K^{n+1}, g_K^{n+1}\right)_{K\in\mathcal{V}, n\in\{0,\dots,N\}}$ at hand for all meshes and all time discretizations, then we can study the convergence of the scheme when the discretization parameters tend to 0. More precisely, consider a sequence $(\mathcal{T}_m)_{m\geq 1}$ of triangulations of Ω such that

$$h_m = \max_{T\in\mathcal{T}_m} \operatorname{diam}(T) \underset{m\to\infty}{\longrightarrow} 0, \tag{18}$$

and such that there exists $\theta^\star > 0$ such that

$$\theta_{\mathcal{T}_m} \leq \theta^\star, \quad \forall m \geq 1. \tag{19}$$

A sequence of dual meshes $(\mathcal{M}_m)_{m\geq 1}$ corresponding to the triangular meshes $(\mathcal{T}_m)_{m\geq 1}$ is built as in §2.1.1. Let $(N_m)_{m\geq 1}$ be an increasing sequence of integers, then we define the corresponding sequence of time steps $\Delta t_m = \frac{t_f}{N_m+1}$ tending to 0 as m tends to ∞.

Theorem 2.6. *Let $(\mathcal{T}_m)_{m\geq 1}$ be a sequence conformal triangular discretization of Ω such that (18) and (19) hold. Let $(f_{\mathcal{M}_m,\Delta t_m}, g_{\mathcal{M}_m,\Delta t_m})_{m\geq 1}$ be a sequence of solutions to the scheme (15). Then there exists $f : \Omega_{t_f} \to \mathbb{R}_+$ and $g : \Omega_{t_f} \to \mathbb{R}_+$ two mesurable functions such that $(f, g) \in L^\infty(0, t_f; L^1(\Omega)) \cap L^2(0, t_f; H^1(\Omega))$ and such that, up a subsequence*

$$\begin{aligned}
f_{\mathcal{M}_m,\Delta t_m} &\longrightarrow f \quad \text{weakly in } L^2(0,t_f;H^1(\Omega)),\\
g_{\mathcal{M}_m,\Delta t_m} &\longrightarrow g \quad \text{weakly } L^2(0,t_f;H^1(\Omega)),\\
f_{\mathcal{M}_m,\Delta t_m} &\longrightarrow f \quad \text{a.e in } (0,t_f)\times\Omega,\\
g_{\mathcal{M}_m,\Delta t_m} &\longrightarrow g \quad \text{a.e in } (0,t_f)\times\Omega,
\end{aligned}$$

and $(f, g) \in L^2(0, t_f; H^1(\Omega))^2$ is a weak solution to (1)-(3) in the following sense

$$\int_0^{t_f}\int_\Omega (f\partial_t\psi - \nu f\Lambda\nabla(f+g+b)\cdot\nabla\psi)\,\mathrm{d}x\,\mathrm{d}t + \int_\Omega f_0\psi(.,0)\,\mathrm{d}x = 0, \tag{20}$$

$$\int_0^{t_f}\int_\Omega (g\partial_t\psi - g\Lambda\nabla(\nu f+g+b)\cdot\nabla\psi)\,\mathrm{d}x\,\mathrm{d}t + \int_\Omega g_0\psi(.,0)\,\mathrm{d}x = 0, \tag{21}$$

for all test functions $\psi \in C_0^\infty(\overline{\Omega}\times[0,t_f))$.

The paper is organized as follows. The existence of nonnegative solution is shown in Section 3. Discrete counterpart of the entropy/entropy-dissipation (6) is established in Section 4. Section 5 is devoted to the convergence proof of the scheme. This proof is based first on the compactness of the sequence of approximate solutions and then on the identification of the limit. We finally present numerical experiments in Section 6, to illustrate the behaviour of the model and of the scheme.

3. Existence of a nonnegative discrete solutions

First all, we prove the positivity of the discrete solutions. This estimate allows to prove the existence of a solution to the nonlinear system (15).

Proposition 3.1. *For all $K \in \mathcal{V}, n \geq 0$,*

$$f_K^n \geq 0, \quad g_K^n \geq 0, \tag{22}$$

hence

$$\sum_{K\in\mathcal{V}} m_K f_K^n = \sum_{K\in\mathcal{V}} m_K f_K^0 = \|f_0\|_{L^1(\Omega)}, \tag{23}$$

and

$$\sum_{K\in\mathcal{V}} m_K g_K^n = \sum_{K\in\mathcal{V}} m_K g_K^0 = \|g_0\|_{L^1(\Omega)}. \tag{24}$$

Proof. The property (22) clearly holds for $n = 0$ thanks to (5). Assume now that (22) holds at time step n. Let us assume that

$$f_K^{n+1} < 0, \quad \text{for some } K \in \mathcal{V}.$$

In view of the definition (15e) of f_{KL}^{n+1} one has that

$$f_K^{n+1} = -\frac{\nu\Delta t}{m_K} \sum_{\sigma_{KL}\in\mathcal{E}_K} \Bigg(\underbrace{(f_K^{n+1})^+}_{=0} \Big[a_{KL}(u_K^{n+1} - u_L^{n+1})\Big]^+ - (f_L^{n+1})^+ \Big[a_{KL}(u_K^{n+1} - u_L^{n+1})\Big]^- \Bigg) + f_K^n \geq 0,$$

yielding a contradiction, ensuring that

$$f_K^{n+1} \geq 0, \quad \forall K \in \mathcal{V}, \forall n \geq 0.$$

Proving that $g_K^n \geq 0$ for all $K \in \mathcal{V}, \forall n \geq 0$, is similar. □

Now, one can apply the same strategy as in [3, Proposition 2.2] for proving the existence of a solution to the scheme (15).

Proposition 3.2. *Given $(f_K^n, g_K^n)_{K\in\mathcal{V}}$ such that $f_{\mathcal{M},\Delta t}(.,n\Delta t)$ and $g_{\mathcal{M},\Delta t}(.,n\Delta t)$ are nonnegative. There exists (at least) one solution $(f_K^{n+1}, g_K^{n+1})_{K\in\mathcal{V}}$ to the scheme (15). Moreover $f_{\mathcal{M},\Delta t}(.,(n+1)\Delta t)$ and $g_{\mathcal{M},\Delta t}(.,(n+1)\Delta t)$ are nonnegative.*

The proof of Proposition 3.2 is not detailed here since it mimics the one of [3, Proposition 2.2]. Let us just mention that we follow the methodology proposed in [27], using a topological degree argument [23, 36].

4. Entropy estimate

The goal of this section is to establish discrete counterpart to the entropy/entropy-dissipation estimate (6) that appears to be sufficient to establish Theorem 2.6. In what follows, $(f_K^n, g_K^n)_{K\in\mathcal{V},n\in\{0,\dots,N+1\}}$ denotes a solution to the scheme (15). We recall the discrete version of entropy functional :

$$\mathfrak{E}^n := \mathfrak{E}(f_K^n, g_K^n) = \sum_{K\in\mathcal{V}} m_K \Big(\frac{1}{\nu}\Gamma(f_K^n) + \Gamma(g_K^n)\Big).$$

Proposition 4.1. *(Entropy stability) For all $n \in \{0, ..., N\}$, one has*

$$\mathfrak{E}^{n+1} - \mathfrak{E}^n + \frac{1-\nu}{2}\Delta t \sum_{\sigma_{KL}\in\mathcal{E}} a_{KL}\Big[(f_K^{n+1} - f_L^{n+1})^2 + (g_K^{n+1} - g_L^{n+1})^2\Big]$$
$$\leq \frac{\Delta t}{2(\nu+1)} \sum_{\sigma_{KL}\in\mathcal{E}} |a_{KL}|(b_K - b_L)^2. \tag{25}$$

Proof. Let us assume momentarly that $f_K^{n+1} > 0$ for any $K \in \mathcal{V}$. We multiply (15a) by $\Delta t \dfrac{\log f_K^{n+1}}{\nu}$ and summing over $K \in \mathcal{V}$ and (15b) by $\Delta t \log g_K^{n+1}$ and summing over $K \in \mathcal{V}$, provides that :

$$A = B + C,$$

where

$$A = \sum_{K\in\mathcal{V}} m_K\Big[\frac{1}{\nu}(f_K^{n+1} - f_K^n)\log f_K^{n+1} + (g_K^{n+1} - g_K^n)\log g_K^{n+1}\Big],$$
$$B = -\Delta t \sum_{K\in\mathcal{V}} \sum_{\sigma_{KL}\in\mathcal{E}_K} a_{KL} f_{KL}^{n+1}\Big(u_K^{n+1} - u_L^{n+1}\Big)\log f_K^{n+1},$$
$$C = -\Delta t \sum_{K\in\mathcal{V}} \sum_{\sigma_{KL}\in\mathcal{E}_K} a_{KL} g_{KL}^{n+1}\Big(v_K^{n+1} - v_L^{n+1}\Big)\log g_K^{n+1}.$$

By the convexity of Γ, we find that

$$A \geq \sum_{K\in\mathcal{V}} m_K\Big[\frac{1}{\nu}(\Gamma(f_K^{n+1}) - \Gamma(f_K^n)) + \Gamma(g_K^{n+1}) - \Gamma(g_K^n)\Big] = \mathfrak{E}^{n+1} - \mathfrak{E}^n.$$

We perform a summation by parts and using the symmetry of a_{KL} :

$$B = -\Delta t \sum_{\sigma_{KL}\in\mathcal{E}} a_{KL} f_{KL}^{n+1}(u_K^{n+1} - u_L^{n+1})(\log f_K^{n+1} - \log f_L^{n+1})$$
$$= -\Delta t \sum_{\sigma_{KL}\in\mathcal{E}} \big(f_K^{n+1}[a_{KL}(u_K^{n+1} - u_L^{n+1})]^+ - f_L^{n+1}[a_{KL}(u_K^{n+1} - u_L^{n+1})]^-\big)$$
$$\times (\log f_K^{n+1} - \log f_L^{n+1}).$$

The Taylor expansion around f_K^{n+1} shows that

$$\overline{f}_{KL}^{n+1}(\log f_K^{n+1} - \log f_L^{n+1}) = f_K^{n+1} - f_L^{n+1}, \tag{26}$$

where $\overline{f}_{KL}^{n+1} = t_{KL} f_K^{n+1} + (1 - t_{KL}) f_L^{n+1}$ for some $t_{KL} \in (0, 1)$. It is shown in [30, p. 468] that

$$B \leq -\Delta t \sum_{\sigma_{KL}\in\mathcal{E}} a_{KL}(u_K^{n+1} - u_L^{n+1})\overline{f}_{KL}^{n+1}(\log f_K^{n+1} - \log f_L^{n+1}).$$

Using (26), one has

$$B \leq -\Delta t \sum_{\sigma_{KL}\in\mathcal{E}} a_{KL}(u_K^{n+1} - u_L^{n+1})(f_K^{n+1} - f_L^{n+1}).$$

In the same way, we have

$$C \leq -\Delta t \sum_{\sigma_{KL}\in\mathcal{E}} a_{KL}(v_K^{n+1} - v_L^{n+1})(g_K^{n+1} - g_L^{n+1}).$$

$$B+C \le -\Delta t \sum_{\sigma_{KL}\in\mathcal{E}} a_{KL}\Big[(f_K^{n+1}-f_L^{n+1})^2+(g_K^{n+1}-g_L^{n+1})^2\Big]$$
$$-\Delta t \sum_{\sigma_{KL}\in\mathcal{E}} a_{KL}(b_K-b_L)\Big[(f_K^{n+1}-f_L^{n+1})+(g_K^{n+1}-g_L^{n+1})\Big]$$
$$-(\nu+1)\Delta t \sum_{\sigma_{KL}\in\mathcal{E}} a_{KL}(f_K^{n+1}-f_L^{n+1})(g_K^{n+1}-g_L^{n+1}).$$

Using the Young inequality, we have

$$D := -\Delta t \sum_{\sigma_{KL}\in\mathcal{E}} a_{KL}(b_K-b_L)\Big[(f_K^{n+1}-f_L^{n+1})+(g_K^{n+1}-g_L^{n+1})\Big]$$
$$\le \frac{\Delta t}{2\varepsilon}\sum_{\sigma_{KL}\in\mathcal{E}}|a_{KL}|(b_K-b_L)^2+\frac{\varepsilon}{2}\Delta t\sum_{\sigma_{KL}\in\mathcal{E}}|a_{KL}|\Big[(f_K^{n+1}-f_L^{n+1})^2+(g_K^{n+1}-g_L^{n+1})^2\Big]$$
$$+\varepsilon\Delta t\sum_{\sigma_{KL}\in\mathcal{E}}\left|a_{KL}(f_K^{n+1}-f_L^{n+1})(g_K^{n+1}-g_L^{n+1})\right|.$$

We choose $\varepsilon=\nu+1$, then since $|x|=x+2x^-$, with $x^-=\max(-x,0)$, one has

$$D \le \frac{\nu+1}{2}\Delta t\sum_{\sigma_{KL}\in\mathcal{E}} a_{KL}\Big[(f_K^{n+1}-f_L^{n+1})^2+(g_K^{n+1}-g_L^{n+1})^2\Big]$$
$$+(\nu+1)\Delta t\sum_{\sigma_{KL}\in\mathcal{E}} a_{KL}^-\Big[(f_K^{n+1}-f_L^{n+1})^2+(g_K^{n+1}-g_L^{n+1})^2\Big]$$
$$+(\nu+1)\Delta t\sum_{\sigma_{KL}\in\mathcal{E}} a_{KL}(f_K^{n+1}-f_L^{n+1})(g_K^{n+1}-g_L^{n+1})$$
$$+2(\nu+1)\Delta t\sum_{\sigma_{KL}\in\mathcal{E}}|a_{KL}(f_K^{n+1}-f_L^{n+1})(g_K^{n+1}-g_L^{n+1})|^-$$
$$+\frac{\Delta t}{2(\nu+1)}\sum_{\sigma_{KL}\in\mathcal{E}}|a_{KL}|(b_K-b_L)^2,$$

concluding the proof of Proposition 4.1. In the general case that is $f_K^{n+1}\ge 0$ for any $K\in\mathcal{V}$ and for any $n\in\{0,...,N+1\}$ we introduce $f_K^{n,\epsilon}=f_K^n+\epsilon$. Following the previous proof on obtain that for any $n\in\{0,...,N\}$,

$$\mathfrak{E}^{n+1,\epsilon}-\mathfrak{E}^{n,\epsilon}+\frac{1-\nu}{2}\Delta t\sum_{\sigma_{KL}\in\mathcal{E}} a_{KL}\Big[(f_K^{n+1}-f_L^{n+1})^2+(g_K^{n+1}-g_L^{n+1})^2\Big]$$
$$-\epsilon\Delta t\sum_{\sigma_{KL}\in\mathcal{E}} a_{KL}(u_K^{n+1}-u_L^{n+1})(\log(f_K^{n+1,\epsilon})-\log(f_L^{n+1,\epsilon}))$$
$$-\epsilon\Delta t\sum_{\sigma_{KL}\in\mathcal{E}} a_{KL}(v_K^{n+1}-v_L^{n+1})(\log(g_K^{n+1,\epsilon})-\log(g_L^{n+1,\epsilon}))$$
$$\le \frac{\Delta t}{2(\nu+1)}\sum_{\sigma_{KL}\in\mathcal{E}}|a_{KL}|(b_K-b_L)^2. \tag{27}$$

where the entropy $\mathfrak{E}^{n,\epsilon}$ is given by

$$\mathfrak{E}^{n,\epsilon}=\sum_{K\in\mathcal{V}} m_K\Big(\frac{1}{\nu}\Gamma(f_K^{n,\epsilon})+\Gamma(g_K^{n,\epsilon})\Big).$$

Passing to the limit when $\epsilon\to 0$ in (27) gives the expected result. □

Remark 4.2. *We define the* $L^2(0,t_f;H^1(\Omega))$ *semi-norm on* $V_{\mathcal{T}}$ *by :*

$$|u|_{1,\mathcal{T}}^2=\left(\sum_{n=0}^{N}\Delta t\sum_{\sigma_{KL}\in\mathcal{E}}a_{KL}(u_K^{n+1}-u_L^{n+1})^2\right)=\int_0^{t_f}\int_\Omega \Lambda\nabla u_{\mathcal{T}}\cdot\nabla u_{\mathcal{T}}\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}t.$$

Corollary 4.3. *There exists* C_7 *depending only on* $t_f,\Omega,f_0,g_0,\nu,b,\theta_{\mathcal{T}},\overline{\Lambda}$ *and* $\underline{\Lambda}$ *such that*

$$\sup_{n\in\{0,\dots,N\}}\mathfrak{E}^{n+1}+\frac{1-\nu}{2}\Big(|f_{\mathcal{T}}|_{1,\mathcal{T}}^2+|g_{\mathcal{T}}|_{1,\mathcal{T}}^2\Big)\le C_7. \tag{28}$$

Proof. Summing (25) over $n=0,\dots,m$ where $m\in\{0,\dots,N\}$ provides

$$\mathfrak{E}^{m+1}\le\mathfrak{E}^0.$$

Summing (25) over $n=0,\dots,N$ provides

$$\mathfrak{E}^{N+1}+\frac{1-\nu}{2}\sum_{n=0}^{N}\Delta t\sum_{\sigma_{KL}\in\mathcal{E}}a_{KL}\Big[(f_K^{n+1}-f_L^{n+1})^2+(g_K^{n+1}-g_L^{n+1})^2\Big]$$
$$\le\mathfrak{E}^0+\frac{1}{2(\nu+1)}\sum_{n=0}^{N}\Delta t\sum_{\sigma_{KL}\in\mathcal{E}}|a_{KL}|(b_K-b_L)^2.$$

As a consequence of Jensen's inequality, one has

$$\mathfrak{E}^0=\sum_{K\in\mathcal{V}}m_K\Big[\frac{1}{\nu}\Gamma(f_K^0)+\Gamma(g_K^0)\Big]\le\int_\Omega\Big[\frac{1}{\nu}\Gamma(f_0)+\Gamma(g_0)\Big]\,\mathrm{d}\boldsymbol{x}<+\infty.$$

Moreover

$$\sum_{n=0}^{N}\Delta t\sum_{\sigma_{KL}\in\mathcal{E}}|a_{KL}|(b_K-b_L)^2\le C_4t_f\|\nabla b\|_\infty^2\sum_{\sigma_{KL}\in\mathcal{E}}|\boldsymbol{x}_K-\boldsymbol{x}_L|^2.$$

Using the fact $|\boldsymbol{x}_K-\boldsymbol{x}_L|^2\le\frac{4}{\pi}\theta_{\mathcal{T}}^2m_T$ for $\sigma_{KL}\in\mathcal{E}_T$ we have

$$\sum_{\sigma_{KL}\in\mathcal{E}}|\boldsymbol{x}_K-\boldsymbol{x}_L|^2\le\frac{12}{\pi}\theta_{\mathcal{T}}^2m_\Omega.$$

Consequently

$$\sum_{n=0}^{N}\Delta t\sum_{\sigma_{KL}\in\mathcal{E}}|a_{KL}|(b_K-b_L)^2\le\frac{12}{\pi}C_4t_fm_\Omega\theta_{\mathcal{T}}^2\|\nabla b\|_\infty^2,$$

concluding the proof of Corollary 4.3 where $C_7=\mathfrak{E}(f_0,g_0)+\frac{8C_4}{\pi(\nu+1)}t_fm_\Omega\theta_{\mathcal{T}}^2\|\nabla b\|_\infty^2$. □

Corollary 4.4. *There exist* C_8 *and* C_9 *depending only on* $t_f,\Omega,f_0,g_0,\nu,b,\theta_{\mathcal{T}}$, $\overline{\Lambda}$ *and* $\underline{\Lambda}$ *such that*

$$\|f_{\mathcal{T},\Delta t}\|_{L^\infty(0,t_f;L^1(\Omega))}+\|g_{\mathcal{T},\Delta t}\|_{L^\infty(0,t_f;L^1(\Omega))}\le C_8, \tag{29}$$

and

$$\|f_{\mathcal{T},\Delta t}\|_{L^2(0,t_f;H^1(\Omega))}+\|g_{\mathcal{T},\Delta t}\|_{L^2(0,t_f;H^1(\Omega))}\le C_9. \tag{30}$$

Moreover for any $1 \leq r < 3$ *there exists* C_{10} *only depending on* $t_f, \Omega, f_0, g_0, \nu, b, \theta_{\mathcal{T}}$, $\overline{\Lambda}, \underline{\Lambda}$ *and* r *such that*

$$\|f_{\mathcal{T},\Delta t}\|_{L^r(\Omega_{t_f})} + \|g_{\mathcal{T},\Delta t}\|_{L^r(\Omega_{t_f})} \leq C_{10}. \tag{31}$$

Proof. By virtue of Corollary 4.3 and the following inequality

$$t - 1 \leq \Gamma(t) \text{ for any } t \geq 1,$$

one has

$$\|f_{\mathcal{M},\Delta t}\|_{L^\infty(0,t_f;L^1(\Omega))} + \|g_{\mathcal{M},\Delta t}\|_{L^\infty(0,t_f;L^1(\Omega))} \leq 2(2m_\Omega + C_7).$$

Using (7) we obtain (29). Using the fact that Λ is uniformly elliptic we have

$$\frac{1-\nu}{2}\underline{\Lambda}(\|\nabla f_{\mathcal{T},\Delta t}\|^2_{L^2(0,t_f;L^2(\Omega)^2)} + \|\nabla g_{\mathcal{T},\Delta t}\|^2_{L^2(0,t_f;L^2(\Omega)^2)}) \leq C_7, \tag{32}$$

which gives (30). □

5. Convergence towards a weak solution

The proof of the convergence properties stated in Theorem 2.6 is based on compactness arguments. As a first step, we show in §5.1 the appropriate compactness properties on the reconstructed discrete solutions. Then we identify in §5.2 the limit value (whose existence is ensured thanks to the compactness properties) as the weak solution to the problem (1).

5.1. Compactness properties of discrete solutions.

As it is classical for unsteady problems, we need to prove some time-compactness for the approximate solutions. We make use of the time-compactness result for degenerate parabolic equations proposed in [8], as an alternative to the classical technique that consists in estimating the time-translates (see [6] in the continuous setting and [28] in the discrete setting). In particular we use [18, Theorem A.1] which is an adaptation to the discretization considered in this article. To this end, we need the following lemma.

Lemma 5.1. *There exist* C_{11} *and* C_{12} *depending only on* $t_f, f_0, g_0, \nu, b, ...$ *such that*

$$\sum_{n=0}^{N}\sum_{K\in\mathcal{V}} m_K(f_K^{n+1} - f_K^n)\varphi(\boldsymbol{x}_K, t_{n+1}) \leq C_{11}\|\nabla\varphi\|_{L^\infty(\Omega_{t_f})}, \quad \varphi \in \mathcal{C}_c^\infty(\Omega_{t_f}). \tag{33}$$

$$\sum_{n=0}^{N}\sum_{K\in\mathcal{V}} m_K(g_K^{n+1} - g_K^n)\varphi(\boldsymbol{x}_K, t_{n+1}) \leq C_{12}\|\nabla\varphi\|_{L^\infty(\Omega_{t_f})}, \quad \varphi \in \mathcal{C}_c^\infty(\Omega_{t_f}). \tag{34}$$

Proof. We only establish (33) since the proof of (34) is similar. For the sake of readability, we denote by $\varphi_K^{n+1} = \varphi(\boldsymbol{x}_K, t_{n+1})$ for all $K \in \mathcal{T}$ and all $n \in \{0, ..., N\}$. We multiply the scheme (15a) by $\Delta t \varphi_K^{n+1}$ and sum for $K \in \mathcal{V}$, for $n \in \{0, ..., N\}$. This yields :

$$A = B,$$

where

$$A = \sum_{n=0}^{N}\sum_{K\in\mathcal{V}} m_K(f_K^{n+1} - f_K^n)\varphi_K^{n+1},$$

and

$$B = -\nu \sum_{n=0}^{N} \Delta t \sum_{\sigma_{KL}\in\mathcal{E}} a_{KL} f_{KL}^{n+1} (u_K^{n+1} - u_L^{n+1})(\varphi_K^{n+1} - \varphi_L^{n+1}).$$

We have

$$|B|^2 \leq \nu^2 \sum_{n=0}^{N} \Delta t \sum_{\sigma_{KL}\in\mathcal{E}} |\boldsymbol{x}_K - \boldsymbol{x}_L|^2 |a_{KL}| (f_{KL}^{n+1})^2 \times \sum_{n=0}^{N} \Delta t \sum_{\sigma_{KL}\in\mathcal{E}} |a_{KL}| (u_K^{n+1} - u_L^{n+1})^2 \|\nabla\varphi\|^2_{L^\infty(\Omega_{t_f})}.$$

Using Lemma 2.2 we have

$$|B|^2 \leq C_4 \nu^2 \sum_{n=0}^{N} \Delta t \sum_{\sigma_{KL}\in\mathcal{E}} |\boldsymbol{x}_K - \boldsymbol{x}_L|^2 (f_{KL}^{n+1})^2 \times \sum_{n=0}^{N} \Delta t \sum_{\sigma_{KL}\in\mathcal{E}} |a_{KL}| (u_K^{n+1} - u_L^{n+1})^2 \|\nabla\varphi\|^2_{L^\infty(\Omega_{t_f})}.$$

Using $(f_{KL}^{n+1})^2 \leq 2((f_K^{n+1})^2 + (g_K^{n+1})^2)$ and $|\boldsymbol{x}_K - \boldsymbol{x}_L|^2 \leq \frac{12}{\pi}\theta_{\mathcal{T}}^2 m_a$ for $a \in \{K, L\}$ we have

$$\sum_{n=0}^{N} \Delta t \sum_{\sigma_{KL}\in\mathcal{E}} |\boldsymbol{x}_K - \boldsymbol{x}_L|^2 (f_{KL}^{n+1})^2 \leq \frac{24}{\pi}\theta_{\mathcal{T}}^2 \sum_{n=0}^{N} \Delta t \sum_{\sigma_{KL}\in\mathcal{E}} (m_K (f_K^{n+1})^2 + m_L (f_L^{n+1})^2).$$

Moreover we have

$$\sum_{\sigma_{KL}\in\mathcal{E}} (m_K (f_K^{n+1})^2 + m_L (f_L^{n+1})^2) = \sum_{K\in\mathcal{V}_{int}} \mathrm{card}(\mathcal{T}_K) m_K (f_K^{n+1})^2 + \sum_{K\in\mathcal{V}_{ext}} (\mathrm{card}(\mathcal{T}_k) + 1)(\mathcal{T}_K) m_K (f_K^{n+1})^2.$$

Let us remark that there exists C_{13} depending only on $\theta_{\mathcal{T}}$ and such that $\mathrm{card}(\mathcal{T}_a) \leq C_{13}$, where $\mathcal{T}_a$ denotes the subset of $\mathcal{T}$ made of the simplices admitting $a \in \mathcal{V}$ as a vertex which gives

$$\frac{24}{\pi}\theta_{\mathcal{T}}^2 \sum_{n=0}^{N} \Delta t \sum_{\sigma_{KL}\in\mathcal{E}} (m_K (f_K^{n+1})^2 + m_L (f_L^{n+1})^2) \leq \frac{24}{\pi}\theta_{\mathcal{T}}^2 (C_{13} + 1) \|f_{\mathcal{M},\Delta t}\|^2_{L^2(\Omega_{t_f})}.$$

Consequently using (7) and (31) one has

$$\nu^2 \sum_{n=0}^{N} \Delta t \sum_{\sigma_{KL}\in\mathcal{E}} |\boldsymbol{x}_K - \boldsymbol{x}_L|^2 |a_{KL}| (f_{KL}^{n+1})^2 \leq \frac{24}{\pi}\nu^2 C_2^2 C_4 C_{10}^2 (C_{13} + 1)\theta_{\mathcal{T}}^2. \tag{35}$$

Using Lemma 2.1 we have

$$\Delta t \sum_{\sigma_{KL}\in\mathcal{E}} |a_{KL}| (u_K^{n+1} - u_L^{n+1})^2 \leq C_3 \int_{\Omega_{t_f}} \Lambda \nabla (f_{\mathcal{T},\Delta t} + g_{\mathcal{T},\Delta t}) \cdot \nabla (f_{\mathcal{T},\Delta t} + g_{\mathcal{T},\Delta t}) \, \mathrm{d}\boldsymbol{x} \, \mathrm{d}t.$$

Consequently using the fact that Λ is uniformly elliptic we have

$$\Delta t \sum_{\sigma_{KL}\in\mathcal{E}} |a_{KL}| (u_K^{n+1} - u_L^{n+1})^2 \leq 2C_3 \overline{\Lambda} (\|\nabla f_{\mathcal{T},\Delta t}\|^2_{L^2(\Omega_{t_f})} + \|\nabla g_{\mathcal{T},\Delta t}\|^2_{L^2(\Omega_{t_f})}).$$

Using (30) we obtain that

$$\Delta t \sum_{\sigma_{KL}\in\mathcal{E}} |a_{KL}|(u_K^{n+1}-u_L^{n+1})^2 \leq 2C_3\overline{\Lambda}C_9^2. \tag{36}$$

Therefore (33) holds with $C_{11}^2 = \frac{48}{\pi}\nu^2\overline{\Lambda}\theta_{\mathcal{T}}^2 C_2^2 C_4 C_{10}^2(C_{13}+1)C_3C_9^2$.

□

We can now state the expected compactness properties.

Proposition 5.2. *There exist $f:\Omega_{t_f}\to\mathbb{R}_+$ and $g:\Omega_{t_f}\to\mathbb{R}_+$ two measurable functions such that, up to an unlabeled subsequence, one has*

$$f_{\mathcal{T}_m,\Delta_m}\to f \text{ a.e in } (0,t_f)\times\Omega \text{ and } g_{\mathcal{T}_m,\Delta_m}\to g \text{ a.e in } (0,t_f)\times\Omega,$$

$$f_{\mathcal{T}_m,\Delta_m}\to f \text{ weakly in } L^2(0,t_f;H^1(\Omega)) \text{ and } g_{\mathcal{T}_m,\Delta_m}\to g \text{ weakly in } L^2(0,t_f;H^1(\Omega)).$$

Proof. We only establish the convergence results for the sequence $(f_{\mathcal{T}_m,\Delta_m})_{m\geq 1}$ since the proof for the sequence $(g_{\mathcal{T}_m,\Delta_m})_{m\geq 1}$ is similar. From (29) we obtain that the sequence $(f_{\mathcal{T}_m,\Delta_m})_{m\geq 1}$ is uniformly bounded in $L^2(0,t_f;H^1(\Omega))$. Therefore, there $f\in L^2(0,t_f;H^1(\Omega))$ such that up to an unlabeled subsequence

$$f_{\mathcal{T}_m,\Delta_m}\to f \text{ weakly in } L^2(0,t_f;H^1(\Omega)). \tag{37}$$

Thanks to (31) and Lemma 5.1, we can apply [8, Thorem 4.1] to obtain up to an unlabeled subsequence

$$f_{\mathcal{T}_m,\Delta_m}\to f \text{ a.e in } (0,t_f)\times\Omega.$$

By virtue of (22) one has

$$f_{\mathcal{M}_m,\Delta t_m}\geq 0 \text{ a.e in } (0,t_f)\times\Omega,$$

which gives

$$f\geq 0 \text{ a.e in } (0,t_f)\times\Omega,$$

which conclude the proof of Proposition 5.2. □

$(f_{\mathcal{T}_m,\Delta_m})_{m>0}$ and $(g_{\mathcal{T}_m,\Delta_m})_{m>0}$ are bounded in $L^r(\Omega_{t_f})$ with $1\leq r<3$ thanks to (31). Hence $(f_{\mathcal{T}_m,\Delta_m})_{m>0}$ and $(g_{\mathcal{T}_m,\Delta_m})_{m>0}$ are equi-integrable in $L^r(\Omega_{t_f})$. Applying the Vitali's convergence theorem we deduce that

Lemma 5.3. *Keeping the assumption and notations of Proposition 5.2 , one has*

$$f_{\mathcal{T}_m,\Delta_m}\underset{m\to+\infty}{\longrightarrow} f, \quad \text{strongly in } L^r(\Omega_{t_f}), \quad \text{for all } r<3,$$

and

$$g_{\mathcal{T}_m,\Delta_m}\underset{m\to+\infty}{\longrightarrow} g, \quad \text{strongly in } L^r(\Omega_{t_f}), \quad \text{for all } r<3.$$

5.2. **Identification as a weak solution.**

Proposition 5.4. *Let (f,g) be as in Proposition 5.2, then f and g are a weak solution to (1)-(3) in the sense of (20) and (21).*

Proof. Let $\psi\in C_0^\infty(\overline{\Omega}\times[0,t_f))$ be a test function and $\psi_K^n=\psi(x_K,t_n)$ for all $K\in\mathcal{V}_m$ and $n\in\{0,...,N_m+1\}$. We first establish (20) from (15a) and to obtain (21) from (15b) is similar. In order to prove that f is a weak solution, we multiply (15a) by $\Delta t_m\psi_K^n$ and sum over $n\in\{0,...,N_m\}$ and $K\in\mathcal{V}_m$, we obtain

$$A_m+B_m=0,$$

where

$$A_m = \sum_{n=0}^{N_m} \sum_{K\in\mathcal{V}_m} m_K (f_K^{n+1} - f_L^{n+1})\psi_K^n,$$

$$B_m = \nu \sum_{n=0}^{N_m} \Delta t_m \sum_{K\in\mathcal{V}_m} \psi_K^n \sum_{\sigma_{KL}\in\mathcal{E}_K} a_{KL} f_{KL}^{n+1}(u_K^{n+1} - u_L^{n+1}).$$

Note that $\psi_K^{N_m+1} = 0$ for all $K \in \mathcal{V}_m$, then a discrete integration part yields

$$\begin{aligned} A_m &= -\sum_{n=0}^{N_m} \Delta t_m \sum_{K\in\mathcal{V}_m} m_K \frac{\psi_K^{n+1} - \psi_K^n}{\Delta t_m} f_K^{n+1} - \sum_{K\in\mathcal{V}_m} m_K f_K^0 \psi_K^0 \\ &= -\int_0^{t_f}\int_\Omega f_{\mathcal{M}_m,\Delta t_m} \delta\psi_{\mathcal{M}_m,\Delta t_m} \,\mathrm{d}\boldsymbol{x}\,\mathrm{d}t - \int_\Omega f_m^0 \psi_m(.,0)\,\mathrm{d}\boldsymbol{x}, \end{aligned}$$

where the function $\delta\psi_{\mathcal{M}_m,\Delta t_m}(\boldsymbol{x}_K, t) = \dfrac{\psi_K^{n+1} - \psi_K^n}{\Delta t_m}$, if $(\boldsymbol{x}_K, t) \in \omega_K \times (t^n, t^{n+1})$. Thanks to the regularity of ψ, the function $\delta\psi_{\mathcal{M}_m,\Delta t_m}$ converges uniformly towards $\partial_t\psi$ on Ω_{t_f}. Moreover by virtue of Lemma 5.3 we have

$$f_{\mathcal{M}_m,\Delta t_m} \longrightarrow f \quad \text{in } L^1(\Omega_{t_f}) \text{ as } m \to \infty.$$

Therefore

(38)
$$\int_0^{t_f}\int_\Omega f_{\mathcal{M}_m,\Delta t_m}(\delta\psi)_{\mathcal{M}_m,\Delta t_m} \,\mathrm{d}\boldsymbol{x}\,\mathrm{d}t \longrightarrow \int_0^{t_f}\int_\Omega f(t,\boldsymbol{x})\partial_t\psi \,\mathrm{d}\boldsymbol{x}\,\mathrm{d}t \quad \text{as } m \to \infty.$$

Moreover, $(f_m^0)_{m\geq 1}$ converges strongly in $L^1(\Omega)$ towards the initial data f_0 and $(\psi_{\mathcal{M}_m,\Delta t_m}(.,0))_{m\geq 1}$ converges uniformly towards $\psi(.,0)$. Therefore, we get that

(39)
$$\int_\Omega f^0_{\mathcal{M}_m,\Delta t_m} \psi_{\mathcal{M}_m,\Delta t_m}(.,0)\,\mathrm{d}\boldsymbol{x} \longrightarrow \int_\Omega f_0(\boldsymbol{x})\psi(.,0)\,\mathrm{d}\boldsymbol{x} \quad m \to \infty.$$

We deduce from (38) and (39) that

(40)
$$A_m \longrightarrow -\int_0^{t_f}\int_\Omega f(t,\boldsymbol{x})\partial_t\psi \,\mathrm{d}\boldsymbol{x}\,\mathrm{d}t - \int_\Omega f_0(\boldsymbol{x})\psi(.,0)\,\mathrm{d}\boldsymbol{x} \quad m \to \infty.$$

Let us now focus on the term B_m and prove that

$$B_m \to \int_{\Omega_{t_f}} \nu f \Lambda \nabla(f + g + b)\cdot\nabla\psi \,\mathrm{d}\boldsymbol{x}\,\mathrm{d}t.$$

To do this, let us introduce the term B'_m defined by

$$B'_m = \int_{\Omega_{t_f}} \nu \tilde{f}_{\mathcal{T}_m,\Delta t_m} \Lambda \nabla(f_{\mathcal{T}_m,\Delta t_m} + g_{\mathcal{T}_m,\Delta t_m} + b_{\mathcal{T}_m})\cdot\nabla\psi_{\mathcal{T}_m,\Delta t_m}(\cdot, t - \Delta t_m)\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}t.$$

where $\tilde{f}_{\mathcal{T}_m,\Delta t_m}$ is a piecewise constant function (on the triangular mesh) such that for any $\boldsymbol{x} \in T$ and for any $t \in (t^n, t^{n+1})$

$$\tilde{f}_{\mathcal{T}_m,\Delta t_m}(\boldsymbol{x}, t) = f_{\mathcal{T}_m}^{n+1}(\boldsymbol{x}_T) = \frac{1}{3}\sum_{K\in\mathcal{V}_T} f_K^{n+1}.$$

Clearly one has

(41)
$$\|f_{\mathcal{T}_m,\Delta_m} - \tilde{f}_{\mathcal{T}_m,\Delta_m}\|_{L^2((0,t_f)\times\Omega)} \leq h_{\mathcal{T}_m}\|\nabla f_{\mathcal{T}_m,\Delta_m}\|_{L^2((0,t_f)\times\Omega)}.$$

By virtue of the strong convergence in $L^2(\Omega_{t_f})$ of $(f_{\mathcal{T}_m,\Delta t_m})_{m\geq 1}$ towards f and the weak convergence in $L^2(\Omega_{t_f})^2$ of $\nabla(f_{\mathcal{T}_m,\Delta t_m}+g_{\mathcal{T}_m,\Delta t_m}+b_{\mathcal{T}_m})$ towards $\nabla(f+g+b)$ which are consequences of (31) and Lemma 5.2 we have

$$B'_m \to \int_{\Omega_{t_f}} \nu\Lambda f\nabla(f+g+b)\cdot\nabla\psi \,\mathrm{d}\boldsymbol{x}\,\mathrm{d}t. \tag{42}$$

It remains to prove that $R_m = B_m - B'_m \to 0$ as $m\to\infty$. To do this we remark that

$$B_m = \nu\sum_{n=0}^{N_m}\Delta t_m \sum_{\sigma_{KL}\in\mathcal{E}} a_{KL} f_{KL}^{n+1}(u_K^{n+1}-u_L^{n+1})(\psi_K^n-\psi_L^n).$$

and using the properties of the family $(a_{KL})_{(K,L)\in\mathcal{V}^2}$ one has

$$B_m = \nu\sum_{n=0}^{N_m}\Delta t_m \sum_{T\in\mathcal{T}_m}\sum_{\sigma_{KL}\in\mathcal{E}_T} a_{KL}^T f_{KL}^{n+1}(u_K^{n+1}-u_L^{n+1})(\psi_K^n-\psi_L^n).$$

The term B'_m can be written as follows

$$B'_m = \nu\sum_{n=0}^{N_m}\Delta t_m \sum_{T\in\mathcal{T}_m} f_T^{n+1}\sum_{\sigma_{KL}\in\mathcal{E}_T} a_{KL}^T(u_K^{n+1}-u_L^{n+1})(\psi_K^n-\psi_L^n).$$

Indeed we can write

$$B'_m = \nu\sum_{n=0}^{N_m}\Delta t_m \sum_{T\in\mathcal{T}_m}\int_{\mathcal{T}} f_{\mathcal{T}_m}^{n+1}\Lambda\nabla(f_{\mathcal{T}_m}^{n+1}+g_{\mathcal{T}_m}^{n+1}+b_{\mathcal{T}_m})\cdot\nabla\psi_{\mathcal{T}_m}^n \,\mathrm{d}\boldsymbol{x}\,\mathrm{d}t.$$

One has for any $T\in\mathcal{T}_m$

$$\int_{\mathcal{T}} \nu f_{\mathcal{T}_m}^{n+1}\Lambda\nabla u_{\mathcal{T}_m}^{n+1}\cdot\nabla\psi_{\mathcal{T}_m}^n \,\mathrm{d}\boldsymbol{x}\,\mathrm{d}t = -\nu f^{n+1}(\boldsymbol{x}_T)\sum_{K\in\mathcal{V}} u_K^{n+1}\sum_{L\in\mathcal{V}}\psi_L^n a_{KL}^T.$$

As a consequence of (10)-(11) one has for any $T\in\mathcal{T}_m$

$$\sum_{K\in\mathcal{V}_m} u_K^{n+1}\sum_{L\in\mathcal{V}_m}\psi_L^n a_{KL}^T = -\sum_{\sigma_{KL}\in\mathcal{E}_T} a_{KL}^T(u_K^{n+1}-u_L^{n+1})(\psi_K^n-\psi_L^n),$$

which gives the expected result. Consequently

$$B_m - B'_m = \nu\sum_{n=0}^{N_m}\Delta t_m\sum_{T\in\mathcal{T}_m}\sum_{\sigma_{KL}\in\mathcal{E}_T} a_{KL}^T(f_{KL}^{n+1}-f_T^{n+1})(u_K^{n+1}-u_L^{n+1})(\psi_K^n-\psi_L^n).$$

We remark that for any $T\in\mathcal{T}_m$ and for any $(K,L)\in\mathcal{V}_m^2$ such that $\sigma_{KL}\in\mathcal{E}_{\mathcal{T}}$ we have

$$|f_{KL}^{n+1}-f_T^{n+1}| \leq \frac{2}{3}(\max_{\boldsymbol{x}\in T} f_{\mathcal{T}_m}^{n+1}(\boldsymbol{x}) - \min_{\boldsymbol{x}\in T} f_{\mathcal{T}_m}^{n+1}(\boldsymbol{x})) = A_T^{n+1},$$

and following the proof of Lemma 2.1 (cf. [9, Lemma 3.2]), we can prove that

$$\sum_{\sigma_{KL}\in\mathcal{E}_T}|a_{KL}^T|(\psi_K^n-\psi_L^n)^2 \leq C_3\overline{\Lambda}\|\nabla\psi\|^2_{L^\infty(\Omega_{t_f})^2} m_T.$$

Consequently by virtue of Holder's inequality and Lemma 2.1 we have

$$|B_m - B'_m|^2 \leq \nu^2 \sum_{n=0}^{N_m} \Delta t_m \sum_{T\in\mathcal{T}_m} (A_T^{n+1})^2 \sum_{\sigma_{KL}\in\mathcal{E}_T} |a_{KL}^T||\psi_K^n - \psi_L^n|^2$$
$$\times \sum_{n=0}^{N_m} \Delta t_m \sum_{T\in\mathcal{T}_m} \sum_{\sigma_{KL}\in\mathcal{E}_T} |a_{KL}^T||u_K^{n+1} - u_L^{n+1}|^2$$
$$\leq C_3\nu^2\overline{\Lambda}\|\nabla\psi\|_{L^\infty(\Omega_{t_f})^2}\Big(\sum_{n=0}^{N_m} \Delta t_m \sum_{T\in\mathcal{T}_m} m_T(A_T^{n+1})^2\Big) \int_{\Omega_{t_f}} \Lambda\nabla u_{\mathcal{T}_m,\Delta_m}\cdot\nabla u_{\mathcal{T}_m,\Delta_m}\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}t.$$

Moreover using [15, Lemma A.1] there exists C_{14} such that

$$\sum_{n=0}^{N_m} \Delta t_m \sum_{T\in\mathcal{T}_m} m_T(A_T^{n+1})^2 \leq C_{14} h_{\mathcal{T}_m}^2 \|\nabla f_{\mathcal{T}_m,\Delta t_m}\|_{L^2(\Omega_{t_f})}^2,$$

which gives

$$|B_m - B'_m|^2 \leq C_3^2\nu^2\overline{\Lambda}\|\nabla\psi\|_{L^\infty(\Omega_{t_f})^2} C_{14} h_{\mathcal{T}_m}^2 \|\nabla f_{\mathcal{T}_m,\Delta t_m}\|_{L^2(\Omega_{t_f})}^2 \times$$
$$\int_{\Omega_{t_f}} \Lambda\nabla u_{\mathcal{T}_m,\Delta_m}\cdot\nabla u_{\mathcal{T}_m,\Delta_m}\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}t.$$

Consequently using (30) we have

$$B_m - B'_m \to 0 \tag{43}$$

From (40), (42) and (43) we obtain that f and g satisfy (20). □

6. Numerical results

Let us provide some illustrations of the behavior of the numerical scheme (15). The nonlinear system (15) is solved thanks to Newton Raphson method. In our test case, the domain is the unit square, i.e., $\Omega = (0,1)^2$. We consider an admissible triangular mesh made of 7297 vertices. An illustration of a mesh type used here is given in Figure 3. These triangle meshes show no symmetry which could artificially increase the convergence rate. All angles are acute, so that, in the case of an isotropic tensor Λ, the coefficients a_{KL} are all non-negative. This is no longer the case when Λ is chosen to be anisotropic. To be more precise concerning the diffusion tensor, we have considered constant diagonal tensors

$$\Lambda = \begin{pmatrix} \Lambda_{xx} & 0 \\ 0 & \Lambda_{yy} \end{pmatrix}$$

where Λ_{xx} and Λ_{yy} are chosen constant in Ω.

The numerical analysis of the scheme was carried out for a uniform time discretization of $(0, t_f)$ only in order to avoid heavy notations. In order to increase the robustness of the algorithm and to ensure the convergence of the Newton-Raphson iterative procedure, we used an adaptive time step procedure in the practical implementation. More precisely, we associate a maximal time step $\Delta t_{\max} = 0.00001$ for the mesh . If the Newton-Raphson method fails to converge after 30 iterations —we choose that the ℓ^∞ norm of the residual has to be smaller than 10^{-7} as stopping criterion—, the time step is divided by two. If the Newton-Raphson method

converges, the time step is multiplied by two and projected on $[0, \Delta t_{\max}]$. The first time step Δt is equal to $\Delta t_{\max}$ in the test case presented below. We perform the

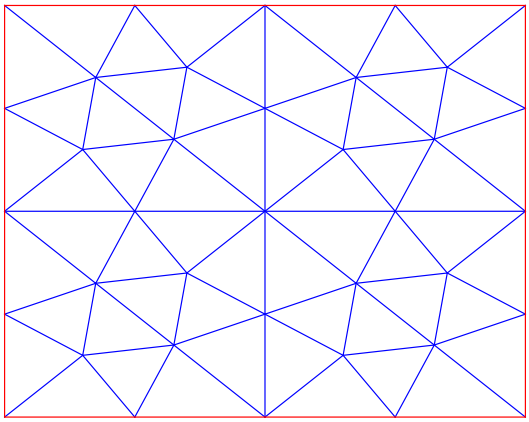


FIGURE 3. Mesh type used in the numerical test

numerical experiments with the following data taken from [2, 3]

$$b(x,y) = \max\left(0, \frac{1}{2}\Big(1 - 16(x-1/2)^2\Big)\Big((\cos(\pi y) + 2\Big)\right), \quad \nu = 0.9.$$

As an initial condition we take

$$f_0(x,y) = \begin{cases} \frac{1}{2} & \text{if } x \le \frac{1}{4}, \\ 0 & \text{elsewhere,} \end{cases} \qquad g_0(x,y) = \begin{cases} b\Big(\frac{1}{2}, 0\Big) - b(x,y) - \Big(x - \frac{1}{2}\Big) & \text{if } x \le \frac{1}{2}, \\ 0 & \text{elsewhere.} \end{cases}$$

Let us provide some illustrations of the behavior of the numerical scheme (15).

Figure 4 and 5 show the evolution of $b(x)+g(x,t)$ (*red*) and $b(x)+g(x,t)+f(x,t)$ (*blue*) in the case isotropic and anisotropic. There is convergence towards a steady state, with horizontal interfaces as expected (see [24]).

Figure 6 shows the control of the entropy in the case isotropic and anisotropic.

ACKNOWLEDGEMENTS

The authors thanks the team Inria/Rapsodi, the project GEOPOR (ANR-13-JS01-0007-01) for their support. The author also thank Clément Cancès for fruitful discussions.

REFERENCES

[1] A. Abudawia and C. Rosier. Numerical analysis for a seawater intrusion problem in a confined aquifer. *Math. Comput. Simulation*, 118:2–16, 2015.

[2] A. Ait Hammou Oulhaj. A finite volume scheme for a seawater intrusion model with cross-diffusion. In *Finite volumes for complex applications VIII—methods and theoretical aspects*, volume 199 of *Springer Proc. Math. Stat.*, pages 421–429. Springer, Cham, 2017.

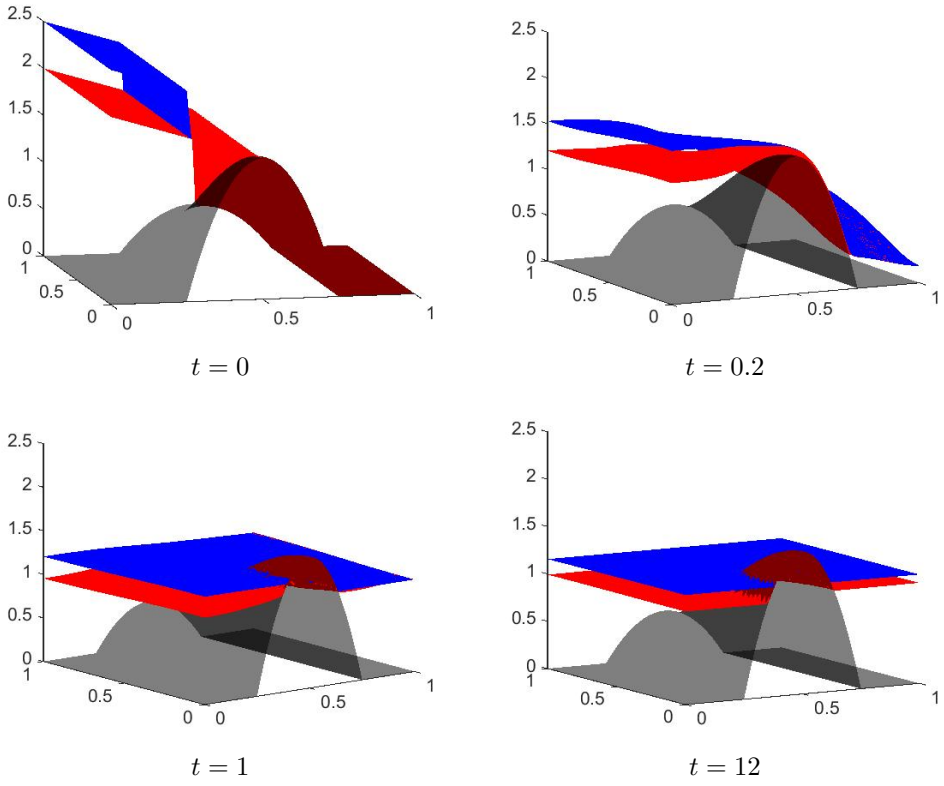


FIGURE 4. Behaviour of the model with an isotropic tensor $\Lambda = \mathbf{I}$.

[3] 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(3):857–880, 2018.

[4] A. Ait Hammou Oulhaj, C. Cancès, and C. Chainais-Hillairet. Numerical analysis of a nonlinearly stable and positive Control Volume Finite Element scheme for Richards equation with anisotropy. *ESAIM: Mathematical Modelling and Numerical Analysis*, 2017.

[5] A. Al Bitar. Modélisation des écoulements en milieu poreux hétérogenes 2d/3d, avec couplages surface/souterrain et densitaires. 2007. Thèse Institut National Polytechnique de Toulouse.

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

[7] B. Andreianov, M. Bendahmane, and R. Ruiz-Baier. Analysis of a finite volume method for a cross-diffusion model in population dynamics. *Math. Models Methods Appl. Sci.*, 21(02):307–344, 2011.

[8] 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(12):3633–3670, 2017.

[9] J. Bear. *Dynamic of fuilds in Porous media*. American Elsevier, New York, 1972.

[10] J. Bear. Hydraulics of groundwater, McGraw-Hill series in water resources and environmental engineering. 1979.

[11] J. Bear, A. H.-D. Cheng, S. Sorek, D. Ouazar, and I. Herrera. *Seawater intrusion in coastal aquifers: concepts, methods and practices*, volume 14. Springer Science & Business Media, 1999.

[12] 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.

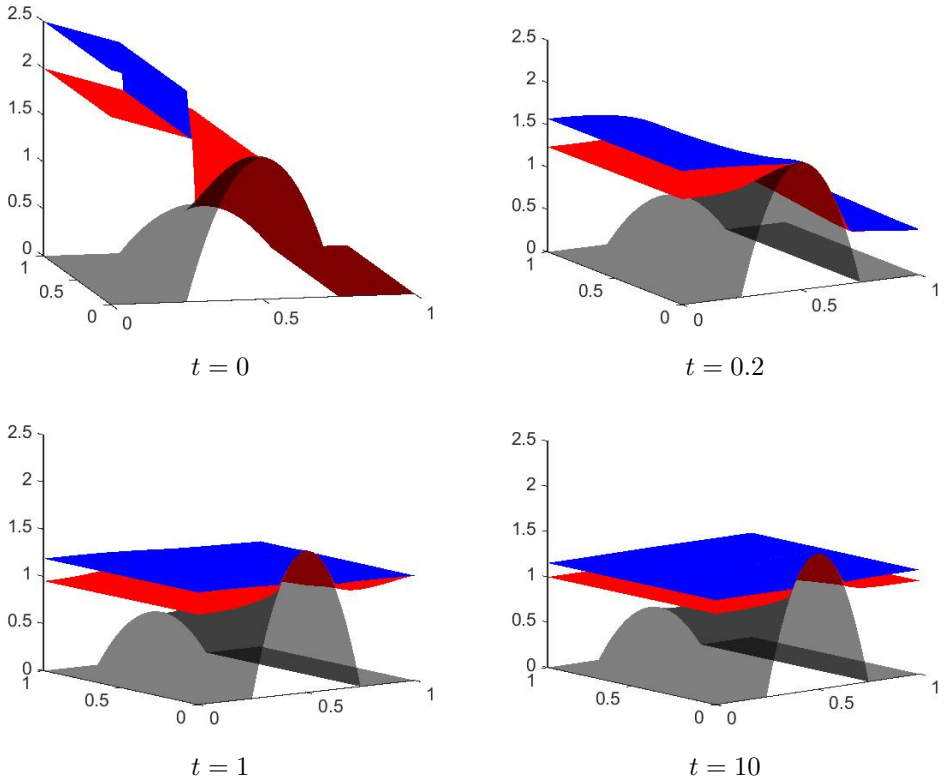


FIGURE 5. Behaviour of the model with an anisotropic tensor : $\Lambda_{xx} = 1, \Lambda_{yy} = 20$.

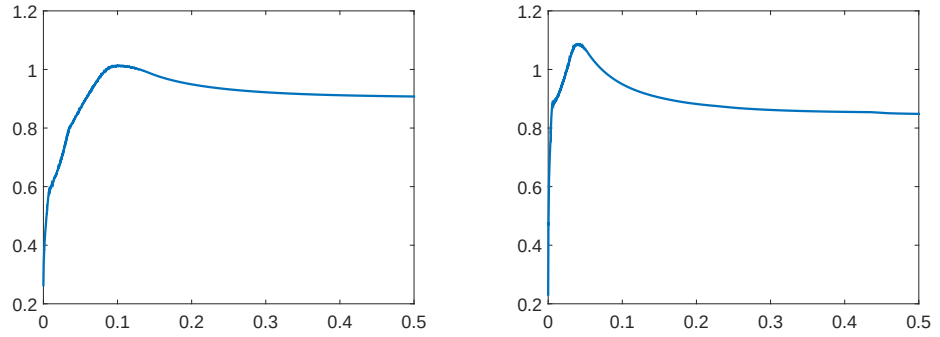


FIGURE 6. Control of the entropy.

[13] C. Cancès, C. Chainais-Hillairet, A. Gerstenmayer, and A. Jüngel. Convergence of a Finite-Volume Scheme for a Degenerate Cross-Diffusion Model for Ion Transport. working paper or preprint, Jan. 2018.
[14] C. Cancès and C. Guichard. Entropy-diminishing CVFE scheme for solving anisotropic degenerate diffusion equations. In *Finite volumes for complex applications. VII. Methods and theoretical aspects*, volume 77 of *Springer Proc. Math. Stat.*, pages 187–196. Springer, Cham, 2014.
[15] 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.
[16] C. Cancès and C. Guichard. Numerical analysis of a robust free energy-diminishing Finite Volume scheme for degenerate parabolic equations with gradient structure. *Found. Comput. Math.*, 2016. online first, DOI: 10.1007/s10208-016-9328-6.
[17] C. Cancès, M. Ibrahim, and M. Saad. Positive nonlinear CVFE scheme for degenerate anisotropic Keller-Segel system. *SMAI J. Comput. Math.*, 3:1–28, 2017.
[18] C. Cancès, F. Nabet, and M. Vohralík. Convergence and a posteriori error analysis for energy-stable finite element approximations of degenerate parabolic equations. working paper or preprint, Oct. 2018.
[19] Z. Chen and R. Ewing. Mathematical analysis for reservoir models. *SIAM J. Math. Anal.*, 30(2):431–453, 1999.
[20] C. Choquet. Parabolic and degenerate parabolic models for pressure-driven transport problems. *Math. Models Methods Appl. Sci.*, 20(4):543–566, 2010.
[21] C. Choquet, M. M. Diédhiou, and C. Rosier. Derivation of a sharp-diffuse interfaces model for seawater intrusion in a free aquifer. numerical simulations. *SIAM J. Appl. Math.*, 76(1):138–158, 2016.
[22] C. Choquet, J. Li, and C. Rosier. Global existence for seawater intrusion models: comparison between sharp interface and sharp-diffuse interface approaches. *Electron. J. Differential Equations*, pages No. 126, 27, 2015.
[23] K. Deimling. *Nonlinear functional analysis*. Springer-Verlag, Berlin, 1985.
[24] J. Escher, P. Laurençot, and B.-V. Matioc. Existence and stability of weak solutions for a degenerate parabolic system modelling two-phase flows in porous media. *Ann. Inst. H. Poincaré Anal. Non Linéaire.*, 28(4):583–598, 2011.
[25] J. Escher, A.-V. Matioc, and B.-V. Matioc. Modelling and analysis of the Muskat problem for thin fluid layers. *J. Math. Fluid Mech.*, 14(2):267–277, 2012.
[26] J. Escher and B.-V. Matioc. Existence and stability of solutions for a strongly coupled system modelling thin fluid films. *NoDEA Nonlinear Differential Equations Appl.*, 20(3):539–555, 2013.
[27] R. Eymard, T. Gallouët, M. Ghilani, and R. Herbin. Error estimates for the approximate solutions of a nonlinear hyperbolic equation given by finite volume schemes. *IMA J. Numer. Anal.*, 18(4):563–594, 1998.
[28] R. Eymard, T. Gallouët, and R. Herbin. Finite volume methods. In *Handbook of numerical analysis, Vol. VII*, Handb. Numer. Anal., VII, pages 713–1020. North-Holland, Amsterdam, 2000.
[29] R. Eymard, R. Herbin, and A. Michel. Mathematical study of a petroleum-engineering scheme. *M2AN Math. Model. Numer. Anal.*, 37(6):937–972, 2003.
[30] F. Filbet. A finite volume scheme for the Patlak-Keller-Segel chemotaxis model. *Numer. Math.*, 104(4):457–488, 2006.
[31] P. A. Forsyth. A control volume finite element approach to NAPL groundwater contamination. *SIAM J. Sci. Statist. Comput.*, 12(5):1029–1057, 1991.
[32] P. A. Forsyth and M. C. Kropinski. Monotonicity considerations for saturated–unsaturated subsurface flow. *SIAM J. Sci. Comput.*, 18(5):1328–1354, 1997.
[33] M. Jazar and R. Monneau. Derivation of seawater intrusion models by formal asymptotics. *SIAM J. Appl. Math.*, 74(4):1152–1173, 2014.
[34] P. Laurençot and B.-V. Matioc. A gradient flow approach to a thin film approximation of the Muskat problem. *Calc. Var. Partial Differential Equations*, 47(1-2):319–341, 2013.
[35] P. Laurençot and B.-V. Matioc. A thin film approximation of the Muskat problem with gravity and capillary forces. *J. Math. Soc. Japan*, 66(4):1043–1071, 2014.

[36] J. Leray and J. Schauder. Topologie et équations fonctionnelles. *Ann. Sci. École Norm. Sup. (3)*, 51:45–78, 1934.
[37] P. Marion, K. Najib, and C. Rosier. Numerical simulations for a seawater intrusion problem in a free aquifer. *Appl. Numer. Math.*, 75:48–60, 2014.
[38] A. Michel. A mathematical comparison of two finite volume methods for two phase flow in porous media. In *Finite volumes for complex applications, III (Porquerolles, 2002)*, pages 213–220. Hermes Sci. Publ., Paris, 2002.
[39] K. Najib and C. Rosier. On the global existence for a degenerate elliptic–parabolic seawater intrusion problem. *Math. Comput. Simulation*, 81(10):2282–2295, 2011.
[40] O. Strack, R. Barnes, and A. Verruijt. Vertically integrated flows, discharge potential, and the Dupuit-Forchheimer approximation. *Ground water*, 44(1):72–75, 2006.
[41] O. D. Strack. A Dupuit-Forchheimer model for three-dimensional flow with variable density. *Water Resources Research*, 31(12):3007–3017, 1995.
[42] M. E. A. Talibi and M. H. Tber. Existence of solutions for a degenerate seawater intrusion problem. *Electron. J. Differential Equations*, 2005(72):1–14, 2005.
[43] M. H. Tber, M. E. A. Talibi, and D. Ouazar. Parameters identification in a seawater intrusion model using adjoint sensitive method. *Math. Comput. Simulation*, 77(2):301–312, 2008.
[44] A. W. Woods and R. Mason. The dynamics of two-layer gravity-driven flows in permeable rock. *J. Fluid Mech.*, 421:83–114, 2000.

Ahmed Ait Hammou Oulhaj
Laboratoire de Mathématiques Appliquées de Compiègne
Centre de Recherches de Royallieu
Département Génie Informatique
B.P. 20529, 60205 Compiègne Cedex
ahmed.ait-hammou-oulhaj@utc.fr

David Maltese
Institut Polytechnique des Sciences Avancées, 63 Boulevard de Brandebourg, 94200 Ivry-sur-Seine, France.
Inria, Univ. Lille, CNRS, UMR 8524 - Laboratoire Paul Painlevé, F-59000 Lille
david.maltese@ipsa.fr
david.maltese@inria.fr