A posteriori analysis of a space and time discretization of a nonlinear model for the flow in variably saturated porous media

by **Christine Bernardi**[1], **Linda El Alaoui**[2], **and Zoubida Mghazli**[3]

Abstract: We consider the equation due to Richards which models the water flow in a partially saturated underground porous medium under the surface. We propose a discretization of this equation by an implicit Euler's scheme in time and finite elements in space. We perform the a posteriori analysis of this discretization, in order to improve its efficiency via time step and mesh adaptivity. Some numerical experiments confirm the interest of this approach.

Résumé: Nous considérons l'équation dite de Richards qui modélise l'écoulement d'eau dans un milieu poreux partiellement saturé souterrain, situé juste sous la surface. Nous écrivons une discrétisation de cette équation par schéma d'Euler implicite en temps et éléments finis en espace. Nous en effectuons l'analyse a posteriori, le but étant d'améliorer son efficacité par adaptation du pas de temps et du maillage. Quelques expériences numériques confirment l'intérêt de cette approche.

[1] Laboratoire Jacques-Louis Lions, C.N.R.S. & Université Pierre et Marie Curie,
B.C. 187, 4 place Jussieu, 75252 Paris Cedex 05, France.
e-mail address: bernardi@ann.jussieu.fr

[2] Université Paris 13, C.N.R.S., U.M.R. 7539,
L.A.G.A., 99 avenue Jean-Baptiste Clément, 93430 Villetaneuse, France.
e-mail address: elalaoui@math.univ-paris13.fr

[3] Équipe d'Ingénierie Mathématique (EIMA- LIRNE),
Faculté des sciences, Université Ibn Tofail, B.P. 133, Kénitra, Maroc.
e-mail address: mghazli_zoubida@yahoo.com

1. Introduction.

The following equation

$$\partial_t \widetilde{\Theta}(h_w) - \nabla \cdot K_w\big(\Theta(h_w)\big)\nabla(h_w + z) = 0, \tag{1.1}$$

models the flow of a wetting fluid, mainly water, in the underground surface, hence in an unsaturated medium, see L.A. Richards [24] for the introduction of this type of models. In opposite to Darcy's or Brinkman's systems (see [22] for all these models), this equation is highly nonlinear: This follows from the fact that, due to the presence of air above the surface, the porous medium is only partially saturated with water. Indeed, this model is derived by combining Darcy's generalized equation with the mass conservation law: When denoting by $\boldsymbol{q}_w$ the flux of water (also called Darcy's velocity), these equations read

$$\boldsymbol{q}_w = -K_w\big(\Theta(h_w)\big)\nabla(h_w + z), \qquad \partial_t \widetilde{\Theta}(h_w) + \nabla \cdot \boldsymbol{q}_w = 0.$$

The unknown is the pressure head h_w, where the index w means "water". The coefficients are the water content Θ and a perturbation of it denoted by $\widetilde{\Theta}$, the permeability term K_w here supposed to be scalar, and the height against the gravitational direction, denoted by z. We refer to [2] for physical values of these coefficients that we use in the numerical experiments.

The key argument for the analysis of problem (1.1) is to use Kirchoff's change of unknowns. Indeed, after this transformation, the new equation fits the general framework proposed in [1] but is simpler (see also [9] for the analysis of a different model). Thus, the existence and uniqueness of a solution to this equation when provided with appropriate initial and boundary conditions are easily derived from standard arguments.

We refer *e.g.* to [13] and [14] for pioneering papers on the finite element discretizations of similar problems, and to [8] for the first study of their finite volume discretization. More recently, several discretizations of Richards equation have been proposed in [11], [21], [26], [27] and [31], see also [28] for a more general equation. All of them rely on a mixed formulation of the previous equation, where the flux $\boldsymbol{q}_w$ is introduced as a second unknown, and fully optimal a priori error estimates are derived. We recall this mixed formulation and its well-posedness. We then propose a discretization that combines the Euler implicit scheme in time and Raviart–Thomas finite elements in space. We prove the well-posedness of the discrete problem.

The goal of the present work is to perform the a posteriori error analysis of this discretization, more precisely to exhibit indicators that uncouple as much as possible the space and time errors, as first proposed in [3] for time-dependent problems. We prove that all these indicators satisfy optimal or quasioptimal error estimates. They allow us to adapt both the time step and the space triangulation in order to optimize the discretization. We thus derive an efficient strategy for adaptivity, following the approach in [4]. Numerical

experiments confirm both the efficiency of this strategy and the interest of the discretization that we propose.

Acknowledgement: This work was partially supported by the GNR MoMaS (PACEN/ CNRS, ANDRA, BRGM, CEA, EdF, IRSN, France).

An outline of the paper is as follows.

- In Section 2, we present the variational formulation of problem (1.1) and investigate its wellposedness in appropriate Sobolev spaces. We also write its mixed formulation.
- Section 3 is devoted to the description of the time semi-discrete problem and of the fully discrete problem. We check their well-posedness.
- In Section 4, we propose error indicators. Next, we prove upper and lower bounds of the error as a function of these indicators.
- Section 5 is devoted to the description of our adaptivity strategy relying on these indicators and to the presentation of some numerical experiments.

2. The continuous problem and its well-posedness.

Let Ω be a bounded connected open set in $\mathbb{R}^d$, $d = 2$ or 3, with a Lipschitz-continuous boundary $\partial\Omega$, and let $\boldsymbol{n}$ denote the unit outward normal vector to Ω on $\partial\Omega$. We assume that $\partial\Omega$ admits a partition without overlap into two parts Γ_D and Γ_F, and that Γ_D has a positive measure. Let also T be a positive real number. From now on, we are interested in the following system

$$\begin{cases} \alpha\,\partial_t u + \partial_t b(u) - \nabla\cdot\Big(\nabla u + k\circ b(u)\boldsymbol{e}_z\Big) = 0 & \text{in } \Omega\times]0,T[,\\ u = u_D & \text{on } \Gamma_D\times]0,T[,\\ \Big(\nabla u + k\circ b(u)\boldsymbol{e}_z\Big)\cdot\boldsymbol{n} = f & \text{on } \Gamma_F\times]0,T[,\\ u|_{t=0} = u_0 & \text{in } \Omega, \end{cases} \tag{2.1}$$

where $-\boldsymbol{e}_z$ stands for the unit vector in the direction of gravity. The unknown is now the quantity u. The coefficients b and k are supposed to be known, and their properties are made precise later on, while α is a positive constant. The data are the Dirichlet boundary condition u_D on Γ_D and the initial condition u_0 on Ω, together with the boundary condition f on the normal component of the flux.

Remark 2.1. The links between equation (1.1) and the first line of system (2.1) follow from Kirchoff's change of unknowns. Indeed, since the conductivity coefficient K_w is positive, the mapping:

$$x \mapsto \mathcal{K}(x) = \int_0^x K_w\big(\Theta(\xi)\big)\,d\xi,$$

is one-to-one from $\mathbb{R}$ into itself. Thus, by setting

$$u = \mathcal{K}(h_w), \qquad b(u) = \Theta\circ\mathcal{K}^{-1}(u), \qquad k\circ b(u) = K_w\circ\Theta\circ\mathcal{K}^{-1}(u),$$

we easily derive the equivalence of (1.1) and the first line of (2.1), for a specific choice of the difference $\widetilde{\Theta} - \Theta$ which is made for mathematical simplicity (see e.g. [31] for a more realistic case where the quantity $\alpha\,\partial_t u$ is replaced by $\alpha\,\partial_t \max\{u, 0\}$). We refer to [25] and [26, §1] for more details. It can be observed that the quantity u has no physical meaning, so that returning to the unknown h_w is needed at the end of each computation. However the importance of using Kirchoff's change of unknowns for degenerate problems has been brought to light in [31].

In what follows, we use the whole scale of Sobolev spaces $W^{m,p}(\Omega)$, with $m \geq 0$ and $1 \leq p \leq +\infty$, equipped with the norm $\|\cdot\|_{W^{m,p}(\Omega)}$ and seminorm $|\cdot|_{W^{m,p}(\Omega)}$, with the usual notation $H^m(\Omega)$ when $p = 2$. For any separable Banach space E equipped with the norm $\|\cdot\|_E$, we denote by $\mathscr{C}^0(0,T;E)$ the space of continuous functions from $[0,T]$ with values in E. For each integer $m \geq 0$, we also introduce the space $H^m(0,T;E)$ as the space

of measurable functions on $]0,T[$ with values in E such that the mappings: $v \mapsto \|\partial_t^\ell v\|_E$, $0 \le \ell \le m$, are square-integrable on $]0,T[$. Finally, we need the spaces $L^\infty(\Omega)$ and $L^\infty(\Omega\times]0,T[)$ of essentially bounded functions on Ω and $\Omega\times]0,T[$, respectively. We are led to make the following assumption concerning the coefficients and the data.

Assumption 2.2.
(i) The mapping b is of class $\mathscr{C}^1$, non-decreasing and globally Lipschitz–continuous on $\mathbb{R}$, with Lipschitz constant c_b;
(ii) The mapping: $x \mapsto k\circ b(x)$ is continuous, bounded on $\mathbb{R}$ and satisfies for a constant c_k

$$\forall x_1 \in \mathbb{R}, \forall x_2 \in \mathbb{R}, \quad \big|k\circ b(x_1) - k\circ b(x_2)\big|^2 \le c_k \big(b(x_1) - b(x_2)\big)(x_1 - x_2); \tag{2.2}$$

(iii) The function u_0 belongs to $H^1(\Omega)$;
(iv) The function u_D admits a lifting, still denoted by u_D for simplicity, which belongs to $L^2(0,T;H^1(\Omega)) \cap H^1(0,T;L^2(\Omega))$ and satisfies $u_D(\cdot,0) = u_0$;
(v) The function f belongs to $H^1(0,T;L^2(\Gamma_F))$.

In order to take into account the boundary condition on Γ_D, we now introduce the space

$$H^1_D(\Omega) = \big\{v \in H^1(\Omega);\ v = 0 \text{ on } \Gamma_D\big\}. \tag{2.3}$$

We denote by $H_D^{-1}(\Omega)$ its dual space and by $\langle\cdot,\cdot\rangle$ the duality pairing between $H_D^{-1}(\Omega)$ and $H^1_D(\Omega)$. Next, we consider the following variational problem

Find u in $L^2(0,T;H^1(\Omega))$ with $\partial_t u$ in $L^2(0,T;H_D^{-1}(\Omega))$ such that

$$u = u_D \quad \text{on } \Gamma_D\times]0,T[\qquad \text{and} \qquad u|_{t=0} = u_0 \quad \text{in } \Omega, \tag{2.4}$$

and, for a.e. t in $]0,T[$,

$$\begin{aligned}\forall v \in H^1_D(\Omega), \quad & \alpha\,\langle \partial_t u(\cdot,t), v\rangle + \langle \partial_t b(u)(\cdot,t), v\rangle \\ & + \int_\Omega \Big(\nabla u + k\circ b(u)\boldsymbol{e}_z\Big)(\boldsymbol{x},t)\ \cdot\ (\nabla v)(\boldsymbol{x})\,d\boldsymbol{x} = \int_{\Gamma_F} f(\boldsymbol{\tau},t)v(\boldsymbol{\tau})\,d\boldsymbol{\tau}.\end{aligned} \tag{2.5}$$

Indeed, the equivalence of such a problem with system (2.1) (in the distribution sense) only requires that the partition of $\partial\Omega$ into Γ_D and Γ_F is sufficiently smooth (in order that $\mathscr{D}(\Omega\cup\Gamma_F)$ is dense into $H^1_D(\Omega)$).

If Assumption 2.2 is satisfied, the mapping b_α defined by

$$b_\alpha(x) = b(x) + \alpha\, x, \tag{2.6}$$

is one-to-one from $\mathbb{R}$ into itself. So, by using the further change of unknown $v_\alpha = b_\alpha(u)$, we can prove the existence result in a simple way by applying the Cauchy–Lipschitz theorem and using the separability of the space $L^2(0,T;H^1(\Omega))$. The uniqueness is then a consequence of Gronwall's lemma. We refer to [18, §2.1] for a detailed proof of these results

relying on the monotonicity of the function b. Note that part of Assumption 2.2 can be weakened for this. However, we have no applications for these weaker properties.

Theorem 2.3. *If Assumption* 2.2 *is satisfied, problem* $(2.4)-(2.5)$ *has a unique solution* u*. Moreover, the quantities* $\partial_t u$ *and* $\partial_t b(u)$ *belong to* $L^2(0,T;L^2(\Omega))$*.*

Remark 2.4. In the more complex case where $\alpha = 0$, the existence and uniqueness of a less regular solution can be derived thanks to the arguments in [1, Thms 2.3 & 2.4] (see also [12] for a similar proof in the case of a biphasic flow water–air). However, this requires slightly different assumptions on the coefficients and the data.

To go further, we prove an a priori estimate for the solution u exhibited in Theorem 2.3.

Proposition 2.5. *If Assumption* 2.2 *is satisfied, the following estimate holds for the solution* u *of problem* $(2.4)-(2.5)$*, for all* t *in* $]0,T[$*,*

$$\begin{aligned}\alpha\,\|u(\cdot,t)\|^2_{L^2(\Omega)} + \int_0^t |u(\cdot,s)|^2_{H^1(\Omega)}\,ds \\ \leq c\,\big(t + \|u_D\|^2_{L^2(0,T;H^1(\Omega))\cap H^1(0,T;L^2(\Omega))} + \|f\|^2_{L^2(0,T;L^2(\Gamma_F))}\big).\end{aligned} \tag{2.7}$$

Proof: We set:

$$u(\boldsymbol{x},t) = u_D(\boldsymbol{x},t) + u_*(\boldsymbol{x},t), \qquad b_*(w) = b(u_D + w).$$

Thus, it is readily checked that u_* belongs to $L^2(0,T;H^1_D(\Omega))$ and satisfies

$$\begin{aligned}\forall v \in H^1_D(\Omega), \quad \alpha\,\langle \partial_t u_*(\cdot,t), v\rangle + \langle \partial_t b_*(u_*)(\cdot,t), v\rangle \\ + \int_\Omega \Big(\nabla u_* + k\circ b_*(u_*)\boldsymbol{e}_z\Big)(\boldsymbol{x},t)\,\cdot\,(\nabla v)(\boldsymbol{x})\,d\boldsymbol{x} = \mathcal{L}_t(v),\end{aligned}$$

where the linear form $\mathcal{L}_t$, defined by

$$\mathcal{L}_t(v) = -\alpha\,\langle \partial_t u_D(\cdot,t), v\rangle - \int_\Omega \big(\nabla u_D\big)(\boldsymbol{x},t)\,\cdot\,(\nabla v)(\boldsymbol{x})\,d\boldsymbol{x} + \int_{\Gamma_F} f(\boldsymbol{\tau},t)v(\boldsymbol{\tau})\,d\boldsymbol{\tau},$$

is obviously continuous on $H^1_D(\Omega)$, with norm $c(t)$ satisfying for a.e. t in $]0,T[$,

$$c(t) \leq \alpha\,\|\partial_t u_D(\cdot,t)\|_{H^{-1}_D(\Omega)} + |u_D(\cdot,t)|_{H^1(\Omega)} + c\,\|f(\cdot,t)\|_{L^2(\Gamma_F)}.$$

Next, we take v equal to $u_*(\cdot,t)$ and integrate the equation with respect to t. Since b_*' is nonnegative, this leads to (note that u_* vanishes at $t=0$)

$$\alpha\,\|u_*(\cdot,t)\|^2_{L^2(\Omega)} + \int_0^t |u_*(\cdot,s)|^2_{H^1(\Omega)}\,ds \leq c\,\big(t + \int_0^t c(s)^2\,ds\big).$$

We conclude by using the properties of u_D.

In view of the discretization, we finally introduce a mixed formulation of problem $(2.4)-(2.5)$. To this aim, we consider the domain $H(\mathrm{div},\Omega)$ of the divergence operator, namely

$$H(\mathrm{div},\Omega)=\big\{\boldsymbol{\varphi}\in L^2(\Omega)^d;\ \nabla\cdot\boldsymbol{\varphi}\in L^2(\Omega)\big\},\tag{2.8}$$

equipped with the graph norm. Since the normal trace operator: $\boldsymbol{\varphi}\mapsto\boldsymbol{\varphi}\cdot\boldsymbol{n}$ can be defined from $H(\mathrm{div},\Omega)$ onto $H^{-\frac{1}{2}}(\partial\Omega)$, see *e.g.* [15, Chap. I, Thm 2.5], and its restriction to Γ_F maps $H(\mathrm{div},\Omega)$ into the dual space of $H^{\frac{1}{2}}_{00}(\Gamma_F)$ (see [19, Chap. 1, Th. 11.7] for the definition of this last space), we also introduce the space

$$H_F(\mathrm{div},\Omega)=\big\{\boldsymbol{\varphi}\in H(\mathrm{div},\Omega);\ \boldsymbol{\varphi}\cdot\boldsymbol{n}=0\text{ on }\Gamma_F\big\}.\tag{2.9}$$

The mixed variational problem then reads

Find $(u,\boldsymbol{q})$ in $L^2(0,T;L^2(\Omega))\times L^2(0,T;H(\mathrm{div},\Omega))$ with $\partial_t u$ in $L^2(0,T;L^2(\Omega))$ such that

$$\boldsymbol{q}\cdot\boldsymbol{n}=-f\quad\text{on }\Gamma_F\times]0,T[\qquad\text{and}\qquad u|_{t=0}=u_0\quad\text{in }\Omega,\tag{2.10}$$

and, for a.e. t in $]0,T[$,

$$\begin{aligned}
&\forall w\in L^2(\Omega),\quad \alpha\int_\Omega\big(\partial_t u\big)(\boldsymbol{x},t)w(\boldsymbol{x})\,d\boldsymbol{x}+\int_\Omega\big(\partial_t b(u)\big)(\boldsymbol{x},t)w(\boldsymbol{x})\,d\boldsymbol{x}\\
&\qquad\qquad+\int_\Omega\big(\nabla\cdot\boldsymbol{q}\big)(\boldsymbol{x},t)w(\boldsymbol{x})\,d\boldsymbol{x}=0,\\
&\forall\boldsymbol{\varphi}\in H_F(\mathrm{div},\Omega),\quad \int_\Omega\boldsymbol{q}(\boldsymbol{x},t)\cdot\boldsymbol{\varphi}(\boldsymbol{x})\,d\boldsymbol{x}-\int_\Omega u(\boldsymbol{x},t)(\nabla\cdot\boldsymbol{\varphi})(\boldsymbol{x})\,d\boldsymbol{x}\\
&\qquad\qquad+\int_\Omega\big(k\circ b(u)\big)(\boldsymbol{x},t)\boldsymbol{e}_z\cdot\boldsymbol{\varphi}(\boldsymbol{x})\,d\boldsymbol{x}=-\langle u_D(\cdot,t),\boldsymbol{\varphi}\cdot\boldsymbol{n}\rangle_{\Gamma_D},
\end{aligned}\tag{2.11}$$

where $\langle\cdot,\cdot\rangle_{\Gamma_D}$ now denotes the duality pairing between $H^{\frac{1}{2}}(\Gamma_D)$ and its dual space. We now check its equivalence with problem $(2.4)-(2.5)$.

Proposition 2.6. *If Assumption 2.2 is satisfied, problems $(2.4)-(2.5)$ and $(2.10)-(2.11)$ are equivalent, in the following sense:*
(i) For any solution u of $(2.4)-(2.5)$, there exists a function $\boldsymbol{q}$ in $L^2(0,T;H(\mathrm{div},\Omega))$ such that the pair $(u,\boldsymbol{q})$ is a solution of problem $(2.10)-(2.11)$;
(ii) For any solution $(u,\boldsymbol{q})$ of $(2.10)-(2.11)$, the function u belongs to $L^2(0,T;H^1(\Omega))$ and is a solution of problem $(2.4)-(2.5)$.

Proof: We check successively the two assertions of the proposition.
1) Let u be a solution of problem $(2.4)-(2.5)$. Letting v run through $\mathscr{D}(\Omega)$ yields the first line of system (2.1) and letting v run through $\mathscr{D}(\Omega\cup\Gamma_F)$ yields the third line of this system. Thus, the function $\boldsymbol{q}=-\nabla u-k\circ b(u)\boldsymbol{e}_z$ belongs to $L^2(0,T;L^2(\Omega)^d)$ and,

since $\partial_t u$ and $\partial_t b(u)$ belong to $L^2(0,T;L^2(\Omega))$ thanks to Theorem 2.3, the same property holds for $\nabla \cdot \boldsymbol{q}$. All this yields that $\boldsymbol{q}$ belongs to $L^2(0,T;H(\mathrm{div},\Omega))$ and also satisfies the first part of (2.10). Moreover, multiplying the first line of (2.1) by any function w in $\mathscr{D}(\Omega)$ and using the density of $\mathscr{D}(\Omega)$ in $L^2(\Omega)$, we derive the first equation in (2.11). The second equation follows by multiplying the equation $\boldsymbol{q} = -\nabla u - k \circ b(u)\boldsymbol{e}_z$ by any $\boldsymbol{\varphi}$ in $H_F(\mathrm{div},\Omega)$ and integrating by parts thanks to the Stokes formula. Thus, $(u,\boldsymbol{q})$ is a solution of (2.10) – (2.11).
2) Conversely, let $(u,\boldsymbol{q})$ be a solution of problem (2.10) – (2.11). Letting $\boldsymbol{\varphi}$ run through $\mathscr{D}(\Omega)^d$ in the second line of (2.11) yields the equation

$$\boldsymbol{q} = -\nabla u - k \circ b(u)\boldsymbol{e}_z, \tag{2.12}$$

and letting it run through $\mathscr{D}(\Omega \cup \Gamma_D)^d$ leads to the boundary condition $u = u_D$ on Γ_D. Thus, since the mapping k is bounded, it follows from the previous equation that u belongs to $L^2(0,T;H^1(\Omega))$ and satisfies (2.4). On the other hand, equation (2.5) follows from the first equation in (2.11), by taking w in $H^1_D(\Omega)$ and using the first part of (2.10) and equation (2.12). Thus, u is a solution of (2.4) – (2.5).

The following corollary is now a direct consequence of Theorem 2.3 and Proposition 2.6.

Corollary 2.7. *If Assumption* 2.2 *is satisfied, problem* (2.10)–(2.11) *has a unique solution* $(u,\boldsymbol{q})$.

Note that, in contrast with u, the flux $\boldsymbol{q}$ has a physical meaning. Indeed, it follows from Remark 2.1 that $\boldsymbol{q}$ is equal to $-K_w\big(\Theta(h_w)\big)\nabla(h_w + z)$, which is the flux in problem (1.1).

3. The discrete problem and its well-posedness.

As already explained in Section 1, we propose a discretization of the problem in two steps: time semi-dicretization, full discretization. The next analysis requires hypotheses which are slightly stronger than Assumption 2.2 but still not restrictive.

Assumption 3.1.
(i) The mappings b and k and the data u_0, u_D, and f satisfy Assumption 2.2;
(ii) The function k is Lipschitz–continuous on $\mathbb{R}$, with Lipschitz constant c_k^*;
(iii) The function u_D belongs to $\mathscr{C}^0(0,T;H^1(\Omega))$.

3.1. The time semi-discrete problem.

Since we intend to work with non uniform time steps, we introduce a partition of the interval $[0,T]$ into subintervals $[t_{n-1},t_n]$, $1\le n\le N$, such that $0=t_0<t_1<\cdots<t_N=T$. We denote by τ_n the time step t_n-t_{n-1}, by τ the N-tuple $(\tau_1,\ldots,\tau_N)$ and by $|\tau|$ the maximum of the τ_n, $1\le n\le N$.

As already hinted in Section 1, the time discretization mainly relies on a backward Euler's scheme, where the nonlinear term $k\circ b(u)$ is treated in an explicit way for simplicity. Thus, the semi-discrete problem reads

Find $(u^n)_{0\le n\le N}$ *in* $L^2(\Omega)^{N+1}$ *and* $(\boldsymbol{q}^n)_{1\le n\le N}$ *in* $H(\mathrm{div},\Omega)^N$ *such that*

$$\boldsymbol{q}^n\cdot\boldsymbol{n}=-f(\cdot,t_n)\quad\text{on }\Gamma_F,\ 1\le n\le N,\qquad\text{and}\qquad u^0=u_0\quad\text{in }\Omega,\tag{3.1}$$

and, for $1\le n\le N$,

$$\begin{aligned}&\forall w\in L^2(\Omega),\\ &\quad\alpha\int_\Omega\big(\frac{u^n-u^{n-1}}{\tau_n}\big)(\boldsymbol{x})w(\boldsymbol{x})\,d\boldsymbol{x}+\int_\Omega\big(\frac{b(u^n)-b(u^{n-1})}{\tau_n}\big)(\boldsymbol{x})w(\boldsymbol{x})\,d\boldsymbol{x}\\ &\qquad\qquad+\int_\Omega\big(\nabla\cdot\boldsymbol{q}^n\big)(\boldsymbol{x})w(\boldsymbol{x})\,d\boldsymbol{x}=0,\\ &\forall\boldsymbol{\varphi}\in H_F(\mathrm{div},\Omega),\quad\int_\Omega\boldsymbol{q}^n(\boldsymbol{x})\cdot\boldsymbol{\varphi}(\boldsymbol{x})\,d\boldsymbol{x}-\int_\Omega u^n(\boldsymbol{x})(\nabla\cdot\boldsymbol{\varphi})(\boldsymbol{x})\,d\boldsymbol{x}\\ &\qquad\qquad+\int_\Omega\big(k\circ b(u^{n-1})\big)(\boldsymbol{x})\boldsymbol{e}_z\cdot\boldsymbol{\varphi}(\boldsymbol{x})\,d\boldsymbol{x}=-\langle u_D(\cdot,t_n),\boldsymbol{\varphi}\cdot\boldsymbol{n}\rangle_{\Gamma_D}.\end{aligned}\tag{3.2}$$

It can be noted that this problem makes sense since both u_D and f are continuous in time. Proving its well-posedness relies on rather different arguments as previously.

Proposition 3.2. *Assume the partition* $\{\Gamma_D,\Gamma_F\}$ *of* $\partial\Omega$ *sufficiently smooth for* $\mathscr{D}(\Omega\cup\Gamma_F)$ *to be dense in* $H^1_D(\Omega)$. *If Assumption* 3.1 *is satisfied, problem* (3.1) – (3.2) *has a unique solution* $(u^n,\boldsymbol{q}^n)_n$.

Proof: We proceed by induction on n and deduce from the same arguments as for Proposition 2.6 that, at each step n, $1 \leq n \leq N$, problem (3.1) – (3.2) admits the equivalent formulation: Find u^n in $H^1(\Omega)$, with u^n equal to $u_D(\cdot, t_n)$ on Γ_D, such that

$$\forall v \in H^1_D(\Omega), \quad \alpha \langle \frac{u^n - u^{n-1}}{\tau_n}, v \rangle + \langle \frac{b(u^n) - b(u^{n-1})}{\tau_n}, v \rangle + \int_\Omega \big(\nabla u^n + k \circ b(u^{n-1}) \boldsymbol{e}_z\big)(\boldsymbol{x}) \cdot (\nabla v)(\boldsymbol{x})\, d\boldsymbol{x} = \int_{\Gamma_F} f(\boldsymbol{\tau}, t_n) v(\boldsymbol{\tau})\, d\boldsymbol{\tau}. \tag{3.3}$$

We check successively the uniqueness and the existence of the solution.
1) Let $(u^n, \boldsymbol{q}^n)_n$ and $(\tilde{u}^n, \tilde{\boldsymbol{q}}^n)_n$ be two solutions of problem (3.1) – (3.2). Due to the induction hypothesis, the function $u^n - \tilde{u}^n$ satisfies, for all v in $H^1_D(\Omega)$,

$$\alpha \langle \frac{u^n - \tilde{u}^n}{\tau_n}, v \rangle + \langle \frac{b(u^n) - b(\tilde{u}^n)}{\tau_n}, v \rangle + \int_\Omega \big(\nabla (u^n - \tilde{u}^n)\big)(\boldsymbol{x}) \cdot (\nabla v)(\boldsymbol{x})\, d\boldsymbol{x} = 0.$$

Thus, taking v equal to $u^n - \tilde{u}^n$ (which belongs to $H^1_D(\Omega)$) and recalling that the product $\big(b(u^n) - b(\tilde{u}^n)\big)(u^n - \tilde{u}^n)$ is nonnegative, we obtain that u^n and $\tilde{u}^n$ are equal. Then, by using the second equation in problem (3.2), we deduce that $\boldsymbol{q}^n$ and $\tilde{\boldsymbol{q}}^n$ coincide, whence the uniqueness result.
2) Due to the induction hypothesis and by setting: $u^n = u^n_* + u_D(\cdot, t_n)$, we must prove the existence of a solution for the problem: Find u^n_* in $H^1_D(\Omega)$ such that

$$\forall v \in H^1_D(\Omega), \quad \alpha \int_\Omega u^n_*(\boldsymbol{x}) v(\boldsymbol{x})\, d\boldsymbol{x} + \int_\Omega b(u^n_*(\boldsymbol{x}) + u_D(\boldsymbol{x}, t_n)) v(\boldsymbol{x})\, d\boldsymbol{x} + \tau_n \int_\Omega (\nabla u^n_*)(\boldsymbol{x}) \cdot (\nabla v)(\boldsymbol{x})\, d\boldsymbol{x} = \mathcal{L}^n(v),$$

where $\mathcal{L}^n$ is a linear form continuous on $H^1_D(\Omega)$. We perform the proof in several steps.
• We define the mapping Φ from $H^1_D(\Omega)$ into its dual space by duality:

$$\forall v \in H^1_D(\Omega), \quad \langle \Phi(w), v \rangle = \alpha \int_\Omega w(\boldsymbol{x}) v(\boldsymbol{x})\, d\boldsymbol{x} + \int_\Omega b(w(\boldsymbol{x}) + u_D(\boldsymbol{x}, t_n)) v(\boldsymbol{x})\, d\boldsymbol{x} + \tau_n \int_\Omega (\nabla w)(\boldsymbol{x}) \cdot (\nabla v)(\boldsymbol{x})\, d\boldsymbol{x} - \mathcal{L}^n(v).$$

The mapping Φ is clearly continuous on $H^1_D(\Omega)$. Moreover, by noting that the quantity $b(w + u_D(\cdot, t_n))w$ is greater than $b(u_D(\cdot, t_n))w$ and that b is Lipschitz–continuous, we observe that

$$\langle \Phi(w), w \rangle \geq \alpha \|w\|^2_{L^2(\Omega)} + \tau_n |w|^2_{H^1(\Omega)} - c_n \|w\|_{L^2(\Omega)},$$

where the constant c_n only depends on the norm of $\mathcal{L}^n$, the Lipschitz constant c_b and the norm of $u_D(\cdot, t_n)$. Thus, the quantity $\langle \Phi(w), w \rangle$ is nonnegative on the "ellipse"

$$\frac{\alpha}{2} \|w\|^2_{L^2(\Omega)} + \tau_n |w|^2_{H^1(\Omega)} = \frac{c_n^2}{2\alpha}.$$

• It follows from the density assumption that there exists an increasing sequence $(\mathbb{H}_m)_m$ of finite-dimensional subspaces of $H^1_D(\Omega)$ such that $\cup_{m\in\mathbb{N}}\mathbb{H}_m$ is dense in $H^1_D(\Omega)$. The restriction of the mapping Φ to each $\mathbb{H}_m$ obviously satisfies the same properties as previously, so that applying Brouwer's fixed point theorem (see [15, Chap. IV, Corollary 1.1] for instance) yields the existence of a function u_m in $\mathbb{H}_m$ such that

$$\forall v_m \in \mathbb{H}_m, \quad \langle \Phi(u_m), v_m\rangle = 0 \qquad \text{and} \qquad \frac{\alpha}{2}\,\|u_m\|^2_{L^2(\Omega)} + \tau_n\,|u_m|^2_{H^1(\Omega)} \le \frac{c_n^2}{2\alpha}.$$

• Since the sequence $(u_m)_m$ is bounded in $H^1(\Omega)$, there exists a subsequence, still denoted by $(u_m)_m$ for simplicity, which converges to a limit weakly in $H^1(\Omega)$ and strongly in $L^2(\Omega)$. We denote this limit by u^n_*. On the other hand, the subsequence satisfies for $m \ge k$

$$\begin{aligned}\forall v_k \in \mathbb{H}_k, \quad \alpha \int_\Omega u_m(\boldsymbol{x})v_k(\boldsymbol{x})\,d\boldsymbol{x} + \int_\Omega b(u_m(\boldsymbol{x}) + u_D(\boldsymbol{x},t_n))v_k(\boldsymbol{x})\,d\boldsymbol{x} \\ + \tau_n \int_\Omega (\nabla u_m)(\boldsymbol{x})\,\cdot\,(\nabla v_k)(\boldsymbol{x})\,d\boldsymbol{x} = \mathcal{L}^n(v_k).\end{aligned} \tag{3.4}$$

Passing to the limit in the linear terms is easy. We also derive from the Lipschitz property of b that

$$\|b(u_m + u_D(t_n)) - b(u^n_* + u_D(t_n))\|_{L^2(\Omega)} \le c\,\|u_m - u^n_*\|_{L^2(\Omega)},$$

whence the convergence of the nonlinear term. Thus, the function u^n_* still satisfies equation (3.4) for all v_k in $\mathbb{H}_k$, whence, owing to the density of $\cup_{k\in\mathbb{N}}\mathbb{H}_k$ in $H^1_D(\Omega)$, for all v in $H^1_D(\Omega)$. Thus, the pair $(u^n, \boldsymbol{q}^n)$, with $u^n = u^n_* + u_D(\cdot,t_n)$ and $\boldsymbol{q}^n = -\nabla u^n - k\circ b(u^{n-1})\boldsymbol{e}_z$ is a solution of problem (3.1) – (3.2), which concludes the proof.

In analogy with Proposition 2.5, we prove a stability property of the solution $(u^n, \boldsymbol{q}^n)$ which is needed later on.

Lemma 3.3. *If Assumption* 3.1 *is satisfied, the following estimate holds for the solutions* $(u^n, \boldsymbol{q}^n)$ *of problems* (3.1) – (3.2), $1 \le n \le N$,

$$\begin{aligned}&\Big(\alpha \sum_{m=1}^n \tau_m \,\|\frac{u^m - u^{m-1}}{\tau_m}\|^2_{L^2(\Omega)}\Big)^{\frac12} + |u^n|_{H^1(\Omega)} \\ &\quad\le c\,\Big(\sqrt{n} + |u_0|_{H^1(\Omega)} \\ &\qquad + \Big(\sum_{m=1}^n \tau_m\,\|\frac{u_D(\cdot,t_m) - u_D(\cdot,t_{m-1})}{\tau_m}\|^2_{L^2(\Omega)}\Big)^{\frac12} + \Big(\sum_{m=1}^n |u_D(\cdot,t_m)|^2_{H^1(\Omega)}\Big)^{\frac12} \\ &\qquad\qquad + \Big(\sum_{m=1}^n \|f(\cdot,t_m)\|^2_{L^2(\Gamma_F)}\Big)^{\frac12}\Big).\end{aligned} \tag{3.5}$$

Proof: Setting as previously $u^n = u_*^n + u_D(\cdot, t_n)$, we observe that problem (3.2) can equivalent be written as

$$\forall v \in H_D^1(\Omega), \quad \alpha \langle \frac{u_*^n - u_*^{n-1}}{\tau_n}, v\rangle + \langle \frac{b(u_*^n + u_D(\cdot,t_n)) - b(u_*^{n-1} + u_D(\cdot,t_n))}{\tau_n}, v\rangle$$
$$+ \int_\Omega \big(\nabla u_*^n + k \circ b(u^{n-1})\boldsymbol{e}_z\big)(\boldsymbol{x}) \cdot (\nabla v)(\boldsymbol{x})\, d\boldsymbol{x} = \langle g^n, v\rangle,$$

where the quantity g^n is defined by

$$\langle g^n, v\rangle = -\alpha \langle \frac{u_D(\cdot,t_n) - u_D(\cdot,t_{n-1})}{\tau_n}, v\rangle$$
$$- \langle \frac{b\big(u_*^{n-1} + u_D(\cdot,t_n)\big) - b\big(u_*^{n-1} + u_D(\cdot,t_{n-1})\big)}{\tau_n}, v\rangle$$
$$- \int_\Omega (\nabla u_D)(\boldsymbol{x},t_n) \cdot (\nabla v)(\boldsymbol{x})\, d\boldsymbol{x} + \int_{\Gamma_F} f(\boldsymbol{\tau},t_n) v(\boldsymbol{\tau})\, d\boldsymbol{\tau}.$$

Thus, taking v equal to $u_*^n - u_*^{n-1}$, noting that the quantity

$$\langle b(u_*^n + u_D(\cdot,t_n)) - b(u_*^{n-1} + u_D(\cdot,t_n)), u_*^n - u_*^{n-1}\rangle$$

is nonnegative and using the formula

$$\nabla u_*^n \cdot \nabla(u_*^n - u_*^{n-1}) = \frac{1}{2}\Big(|\nabla(u_*^n - u_*^{n-1})|^2 + |\nabla u_*^n|^2 - |\nabla u_*^{n-1}|^2\Big),$$

together with the boundedness of the mapping k lead to, with $g^n = g_1^n + g_2^n$,

$$\alpha\,\tau_n \,\|\frac{u_*^n - u_*^{n-1}}{\tau_n}\|_{L^2(\Omega)}^2 + \frac{1}{2}|u_*^n|_{H^1(\Omega)}^2 + \frac{1}{2}|u_*^n - u_*^{n-1}|_{H^1(\Omega)}^2$$
$$\leq \frac{1}{2}|u_*^{n-1}|_{H^1(\Omega)}^2 + (c + \|g_1^n\|_{H_D^{-1}(\Omega)})|u_*^n - u_*^{n-1}|_{H^1(\Omega)} + \tau_n \,\|g_2^n\|_{L_D^2(\Omega)}\|\frac{u_*^n - u_*^{n-1}}{\tau_n}\|_{L^2(\Omega)}.$$

By using the inequality $ab \leq \frac{1}{2}(a^2 + b^2)$ and summing on the n, we obtain

$$\frac{\alpha}{2}\sum_{m=1}^n \tau_m \,\|\frac{u_*^m - u_*^{m-1}}{\tau_m}\|_{L^2(\Omega)}^2 + \frac{1}{2}|u_*^n|_{H^1(\Omega)}^2$$
$$\leq cn + \sum_{m=1}^n \|g_1^m\|_{H_D^{-1}(\Omega)}^2 + \frac{1}{2\alpha}\sum_{m=1}^n \tau_m \,\|g_2^m\|_{L_D^2(\Omega)}^2.$$

We conclude thanks to an appropriate choice of g_1^n and g_2^n and by using a triangle inequality.

Remark 3.4. From now on, we denote by $c_0(\tau)$ the maximum of the quantities that appear in the right-hand side of estimate (3.5), namely

$$c_0(\tau) = c\,\Big(\sqrt{N} + |u_0|_{H^1(\Omega)} + \Big(\sum_{m=1}^{N} \tau_m \,\|\frac{u_D(\cdot,t_m) - u_D(\cdot,t_{m-1})}{\tau_m}\|^2_{L^2(\Omega)}\Big)^{\frac{1}{2}} + \Big(\sum_{m=1}^{N} |u_D(\cdot,t_m)|^2_{H^1(\Omega)}\Big)^{\frac{1}{2}} + \Big(\sum_{m=1}^{N} \|f(\cdot,t_m)\|^2_{L^2(\Gamma_F)}\Big)^{\frac{1}{2}}\Big). \tag{3.6}$$

There is no reason for the last terms in this quantity to be bounded independently of τ. Assumption 3.1 only implies that

$$c_0(\tau) \le c\,\sqrt{N}. \tag{3.7}$$

3.2. The fully discrete problem.

From now on, we assume that Ω is a polygon ($d = 2$) or a polyhedron ($d = 3$). For each n, $0 \le n \le N$, let $(\mathcal{T}_h^n)_{h_n}$ be a regular family of triangulations of Ω (by triangles or tetrahedra), in the sense that, for each h_n:
• $\overline{\Omega}$ is the union of all elements of $\mathcal{T}_h^n$;
• The intersection of two different elements of $\mathcal{T}_h^n$, if not empty, is a vertex or a whole edge or a whole face of both of them;
• The ratio of the diameter h_K of any element K of $\mathcal{T}_h^n$ to the diameter of its inscribed circle or sphere is smaller than a constant σ independent of h and n.
As usual, h_n stands for the maximum of the diameters h_K, $K \in \mathcal{T}_h^n$. We make the further and non restrictive assumption that both $\overline{\Gamma}_D$ and $\overline{\Gamma}_F$ are the union of whole edges ($d = 2$) or whole faces ($d = 3$) of elements of $\mathcal{T}_h^n$.

We now introduce two finite element spaces, first the space $\mathbb{X}_h^n$:

$$\mathbb{X}_h^n = \big\{v_h \in L^2(\Omega);\ \forall K \in \mathcal{T}_h^n,\ v_h|_K \in \mathcal{P}_0(K)\big\}, \tag{3.8}$$

where $\mathcal{P}_0(K)$ is the space of constant functions on K, next the space $\mathbb{Y}_h^n$ associated with Raviart–Thomas finite elements [23]:

$$\mathbb{Y}_h^n = \big\{\boldsymbol{\varphi}_h \in H(\mathrm{div},\Omega);\ \forall K \in \mathcal{T}_h^n,\ \boldsymbol{\varphi}_h|_K \in \mathcal{RT}(K)\big\}, \tag{3.9}$$

where $\mathcal{RT}(K)$ stands for the space of restrictions to K of polynomials of the form $\boldsymbol{\alpha} + \beta\,\boldsymbol{x}$, $\boldsymbol{\alpha} \in \mathbb{R}^d$, $\beta \in \mathbb{R}$. In order to take into account the boundary conditions, we also need its subspace

$$\mathbb{Y}_{hF}^n = \mathbb{Y}_h^n \cap H_F(\mathrm{div},\Omega). \tag{3.10}$$

Recalling that normal traces of elements of $\mathbb{Y}_h^n$ on $\partial\Omega$ are piecewise constant, we define for each n, $1 \le n \le N$, an approximation f_h^n of $f(\cdot, t_n)$ by

$$\forall K \in \mathcal{T}_h^n, \quad f_h^n|_{K\cap\Gamma_F} = \frac{1}{\mathrm{meas}(K \cap \Gamma_F)} \int_{K\cap\Gamma_F} f(\boldsymbol{\tau}, t_n)\, d\boldsymbol{\tau}. \tag{3.11}$$

Similarly, we define u_{0h} as the image of u_0 by the orthogonal projection operator from $L^2(\Omega)$ onto $\mathbb{X}_h^0$. As standard for the discretization of parabolic equations with different triangulations, we also introduce the orthogonal projection operator Π_h^n from $L^2(\Omega)$ onto the space $\mathbb{X}_h^n$.

We are thus in a position to write the discrete problem, constructed from problem (3.1) – (3.2) by the Galerkin method,

Find $(u_h^n)_{0\le n\le N}$ *in* $\prod_{n=0}^N \mathbb{X}_h^n$ *and* $(\boldsymbol{q}_h^n)_{1\le n\le N}$ *in* $\prod_{n=1}^N \mathbb{Y}_h^n$ *such that*

$$\boldsymbol{q}_h^n \cdot \boldsymbol{n} = -f_h^n \quad \text{on } \Gamma_F,\ 1 \le n \le N, \qquad \text{and} \qquad u_h^0 = u_{0h} \quad \text{in } \Omega, \tag{3.12}$$

and, for $1 \le n \le N$,

$$\begin{aligned}
&\forall w_h \in \mathbb{X}_h^n,\\
&\quad \alpha \int_\Omega \Big(\frac{u_h^n - u_h^{n-1}}{\tau_n}\Big)(\boldsymbol{x}) w_h(\boldsymbol{x})\, d\boldsymbol{x} + \int_\Omega \Big(\frac{b(u_h^n) - b(u_h^{n-1})}{\tau_n}\Big)(\boldsymbol{x}) w_h(\boldsymbol{x})\, d\boldsymbol{x}\\
&\qquad\qquad + \int_\Omega \big(\nabla \cdot \boldsymbol{q}_h^n\big)(\boldsymbol{x}) w_h(\boldsymbol{x})\, d\boldsymbol{x} = 0,\\
&\forall \boldsymbol{\varphi}_h \in \mathbb{Y}_{hF}^n,\\
&\quad \int_\Omega \boldsymbol{q}_h^n(\boldsymbol{x}) \cdot \boldsymbol{\varphi}_h(\boldsymbol{x})\, d\boldsymbol{x} - \int_\Omega u_h^n(\boldsymbol{x})(\nabla \cdot \boldsymbol{\varphi}_h)(\boldsymbol{x})\, d\boldsymbol{x}\\
&\qquad + \int_\Omega \Pi_h^n\big(k \circ b(u_h^{n-1})\big)(\boldsymbol{x}) \boldsymbol{e}_z \cdot \boldsymbol{\varphi}_h(\boldsymbol{x})\, d\boldsymbol{x} = -\int_{\Gamma_D} \boldsymbol{u}_D(\boldsymbol{\tau}, t_n)(\boldsymbol{\varphi}_h \cdot \boldsymbol{n})(\boldsymbol{\tau})\, d\boldsymbol{\tau}.
\end{aligned} \tag{3.13}$$

Remark 3.5. Owing to the definition of Π_h^n, the quantities u_h^{n-1} and $b(u_n^{n-1})$ can be replaced by $\Pi_h^n u_h^{n-1}$ and $\Pi_h^n b(u_n^{n-1})$ in (3.13) without modifying the discrete problem. Moreover, since u_h^{n-1} is piecewise constant, computing $\Pi_h^n g(u_h^{n-1})$ for any continuous function g is not expensive at all. Indeed, we have the formula, for all K in $\mathcal{T}_h^n$,

$$\big(\Pi_h^n g(u_h^{n-1})\big)|_K = \frac{1}{\mathrm{meas}(K)} \sum_{K' \in \mathcal{T}_h^{n-1}} \mathrm{meas}(K' \cap K)\, g(u_h^{n-1})|_{K'}. \tag{3.14}$$

Thus, where the mesh is only refined, i.e., for the K in $\mathcal{T}_h^n$ which are contained in one element of $\mathcal{T}_h^{n-1}$, $\big(\Pi_h^n g(u_h^{n-1})\big)|_K$ coincides with $g(u_h^{n-1})|_K$. It must also be noted that, in practice, the operator Π_h^n is often replaced by the Lagrange interpolation operator. We avoid to present this modification for simplicity.

Fortunately, proving the well-posedness of this problem is easier than for the previous ones.

Proposition 3.6. *If Assumption* 3.1 *is satisfied, each problem* (3.12) – (3.13), $1 \le n \le N$, *has a unique solution* $(u_h^n, \boldsymbol{q}_h^n)$.

Proof: There also, we proceed by induction on n and check successively the uniqueness and the existence of the solution.
1) Let $(u_h^n, \boldsymbol{q}_h^n)$ and $(\tilde{u}_h^n, \tilde{\boldsymbol{q}}_h^n)$ be two solutions of problem (3.12) – (3.13). Their difference satisfies

$$\begin{aligned}
\forall w_h \in \mathbb{X}_h^n, \quad & \alpha \int_\Omega \big(\frac{u_h^n - \tilde{u}_h^n}{\tau_n}\big)(\boldsymbol{x},) w_h(\boldsymbol{x})\, d\boldsymbol{x} + \int_\Omega \big(\frac{b(u_h^n) - b(\tilde{u}_h^n)}{\tau_n}\big)(\boldsymbol{x}) w_h(\boldsymbol{x})\, d\boldsymbol{x} \\
& \qquad + \int_\Omega \big(\nabla \cdot (\boldsymbol{q}_h^n - \tilde{\boldsymbol{q}}_h^n)\big)(\boldsymbol{x}) w_h(\boldsymbol{x})\, d\boldsymbol{x} = 0, \qquad (3.15) \\
\forall \boldsymbol{\varphi}_h \in \mathbb{Y}_{hF}^n, \quad & \int_\Omega (\boldsymbol{q}_h^n - \tilde{\boldsymbol{q}}_h^n)(\boldsymbol{x}) \cdot \boldsymbol{\varphi}_h(\boldsymbol{x})\, d\boldsymbol{x} - \int_\Omega (u_h^n - \tilde{u}_h^n)(\boldsymbol{x})(\nabla \cdot \boldsymbol{\varphi}_h)(\boldsymbol{x})\, d\boldsymbol{x} = 0.
\end{aligned}$$

When taking w_h equal to $u_h^n - \tilde{u}_h^n$ and $\boldsymbol{\varphi}_h$ equal to $\boldsymbol{q}_h^n - \tilde{\boldsymbol{q}}_h^n$, summing the two equations and noting that $\big(b(u_h^n) - b(\tilde{u}_h^n)\big)(u_n^n - \tilde{u}_h^n)$ is nonnegative, we obtain

$$\frac{\alpha}{\tau_n} \|u_h^n - \tilde{u}_h^n\|^2_{L^2(\Omega)} + \|\boldsymbol{q}_h^n - \tilde{\boldsymbol{q}}_h^n\|^2_{L^2(\Omega)^d} \le 0.$$

So, $(u_h^n, \boldsymbol{q}_h^n)$ and $(\tilde{u}_h^n, \tilde{\boldsymbol{q}}_h^n)$ are equal.
2) We introduce a lifting $\boldsymbol{\chi}_h^n$ of f_h^n in the following way: If $\mathcal{E}_h^n$ denotes the set of all edges ($d = 2$) or faces (d = 3) of elements of $\mathcal{T}_h^n$ and $\mathcal{E}_h^{nF}$ its subset made of all edges or faces contained in Γ_F, there exists a unique function $\boldsymbol{\chi}_h^n$ in $\mathbb{Y}_h^n$ such that

$$\boldsymbol{\chi}_h^n \cdot \boldsymbol{n} = \begin{cases} f_h^n & \text{on } e \in \mathcal{E}_h^{nF}, \\ 0 & \text{on } e \in \mathcal{E}_h^n \setminus \mathcal{E}_h^{nF}, \end{cases}$$

(indeed, these degrees of freedom are $\mathcal{RT}(K)$-unisolvent on each K, see [23]). Next, we define the mapping Ψ on $\mathbb{X}_h^n \times \mathbb{Y}_{hF}^n$ by

$$\begin{aligned}
\forall (w_h, \boldsymbol{\varphi}_h) \in \mathbb{X}_h^n \times \mathbb{Y}_{hF}^n, & \\
\langle \Psi(u_h, \boldsymbol{q}_h), (w_h, \boldsymbol{\varphi}_h) \rangle = \alpha & \int_\Omega \big(\frac{u_h}{\tau_n}\big)(\boldsymbol{x}) w_h(\boldsymbol{x})\, d\boldsymbol{x} + \int_\Omega \big(\frac{b(u_h)}{\tau_n}\big)(\boldsymbol{x}) w_h(\boldsymbol{x})\, d\boldsymbol{x} \\
& + \int_\Omega \big(\nabla \cdot \boldsymbol{q}_h\big)(\boldsymbol{x}, t) w_h(\boldsymbol{x})\, d\boldsymbol{x} + \int_\Omega \boldsymbol{q}_h(\boldsymbol{x}) \cdot \boldsymbol{\varphi}_h(\boldsymbol{x})\, d\boldsymbol{x} \\
& - \int_\Omega u_h(\boldsymbol{x})(\nabla \cdot \boldsymbol{\varphi}_h)(\boldsymbol{x})\, d\boldsymbol{x} - \mathcal{M}^n(w_h, \boldsymbol{\varphi}_h),
\end{aligned}$$

where the linear form $\mathcal{M}^n(\cdot,\cdot)$ is defined by

$$\begin{aligned}\mathcal{M}^n(w_h,\boldsymbol{\varphi}_h) = \alpha \int_\Omega \big(\frac{\Pi_h^n u_h^{n-1}}{\tau_n}\big)(\boldsymbol{x}) w_h(\boldsymbol{x})\,d\boldsymbol{x} + \int_\Omega \big(\frac{\Pi_h^n b(u_h^{n-1})}{\tau_n}\big)(\boldsymbol{x}) w_h(\boldsymbol{x})\,d\boldsymbol{x}\\ - \int_\Omega \big(\nabla\cdot\boldsymbol{\chi}_h^n\big)(\boldsymbol{x},t) w_h(\boldsymbol{x})\,d\boldsymbol{x} - \int_\Omega \boldsymbol{\chi}_h^n(\boldsymbol{x})\cdot\boldsymbol{\varphi}_h(\boldsymbol{x})\,d\boldsymbol{x}\\ - \int_\Omega \big(\Pi_h^n k\circ b(u_h^{n-1})\big)(\boldsymbol{x})\boldsymbol{e}_z\cdot\boldsymbol{\varphi}_h(\boldsymbol{x})\,d\boldsymbol{x} - \int_{\Gamma_D} \boldsymbol{u}_D(\boldsymbol{\tau},t_n)(\boldsymbol{\varphi}_h\cdot\boldsymbol{n})(\boldsymbol{\tau})\,d\boldsymbol{\tau}.\end{aligned}$$

When the space $\mathbb{X}_h^n \times \mathbb{Y}_{hF}^n$ is equipped with the norm of $L^2(\Omega)\times H(\mathrm{div},\Omega)$, the mapping Ψ is continuous and satisfies

$$\langle\Psi(u_h,\boldsymbol{q}_h),(u_h,\boldsymbol{q}_h)\rangle \geq \frac{\alpha}{\tau_n}\|u_h\|^2_{L^2(\Omega)} + \|\boldsymbol{q}_h\|^2_{L^2(\Omega)^d} - C_n\big(\frac{\alpha}{\tau_n}\|u_h\|^2_{L^2(\Omega)} + \|\boldsymbol{q}_h\|^2_{H(\mathrm{div},\Omega)}\big)^{\frac{1}{2}},$$

where C_n only depends on the norm of $\mathcal{M}^n$ in an appropriate dual space. So, using the fact that on the finite-dimensional space $\mathbb{X}_h^n \times \mathbb{Y}_{hF}^n$ all norms are equivalent and applying once more Brouwer's fixed point theorem yield the existence of a pair $(u_h^n, \boldsymbol{q}_{h*}^n)$ such that

$$\forall (w_h,\boldsymbol{\varphi}_h)\in \mathbb{X}_h^n\times\mathbb{Y}_{hF}^n,\quad \langle\Psi(u_h^n,\boldsymbol{q}_{h*}^n),(w_h,\boldsymbol{\varphi}_h)\rangle = 0.$$

Then, the pair $(u_h^n, \boldsymbol{q}_h^n = \boldsymbol{q}_{h*}^n + \boldsymbol{\chi}_h^n)$ is a solution of problem (3.12) – (3.13).

Since our aim is adaptivity of the time steps and the triangulations, we do not prove a priori error estimates for the semi-discrete and discrete problems. We refer to [13], [14], [21], and [26] for these estimates in the case of very similar problems or discretizations.

4. A posteriori analysis of the discretization.

We first recall some notation which is standard in the a posteriori analysis of finite element discretization. Next, we give the definition of two families of indicators linked respectively to the time semi-discretization and to the finite element discretization. We then prove successively upper and lower bounds for the error. All these estimates are summed up in a conclusion.

We also make precise the new assumptions that are needed in this section.

Assumption 4.1.
(i) The mappings b and k and the data u_0, u_D, and f satisfy Assumption 3.1;
(ii) The mapping b is of class $\mathscr{C}^2$ on $\mathbb{R}$ with bounded and Lipschitz-continuous derivatives;
(iii) The mapping k is of class $\mathscr{C}^1$ on $\mathbb{R}$ with bounded and Lipschitz-continuous derivative;
(iv) The function u_0 belongs to $L^\infty(\Omega)$;
(v) The function u_D is continuous on $[0,T]\times\Gamma_D$.

4.1. Some notation.

Foe each n, $1\le n\le N$, and with each K in $\mathcal{T}_h^n$, we associate
(i) the set $\mathcal{E}_K^0$ of all edges ($d=2$) or faces ($d=3$) of K which are not contained in $\partial\Omega$;
(ii) the sets $\mathcal{E}_K^D$ and $\mathcal{E}_K^F$ of all edges ($d=2$) or faces ($d=3$) of K which are contained in $\overline{\Gamma}_D$ and $\overline{\Gamma}_F$, respectively;
(iii) the domain ω_K equal to the union of all elements of $\mathcal{T}_h^n$ that share at least an edge ($d=2$) or a face ($d=3$) with K.
For each e in $\mathcal{E}_K^0$, we denote by $[\cdot]_e$ the jump through e (the introduction of a vector normal to e is necessary to make precise the sign of this jump, however we do not need it in what follows).

With each family of values (v^n), $0\le n\le N$, we associate the function v_τ which is affine on each interval $[t_{n-1},t_n]$, $1\le n\le N$, and equal to v^n in t_n, $0\le n\le N$. For each function v continuous on $[0,T]$, we also introduce the functions $\pi_\tau^+ v$ and $\pi_\tau^- v$ which are constant, equal to $v(t_n)$ and $v(t_{n-1})$, respectively, on each interval $]t_{n-1},t_n]$, $1\le n\le N$. For brevity, we use the notation $\pi_\tau^\pm v=\pi_\tau^\pm v_\tau$.

Finally, we introduce an approximation u_{Dh}^n of $u_D(\cdot,t_n)$ on Γ_D: For each K in $\mathcal{T}_h^n$ and each edge ($d=2$) or face ($d=3$) e in $\mathcal{E}_K^D$, the restriction of u_{Dh}^n to e is equal to the Lagrange interpolate in $\mathcal{P}_1(e)$ of $u_D(\cdot,t_n)$.

4.2. The error indicators.

For the reasons explained above, we consider two families of error indicators. All of them are defined for each n, $1\le n\le N$, and for each K in $\mathcal{T}_h^n$.

(i) Error indicators linked to the time discretization

$$\eta_K^{n(\tau)} = \tau_n^{\frac{1}{2}} \, \|u_h^n - u_h^{n-1}\|_{L^2(K)} + \tau_n^{\frac{1}{2}} \, \|\frac{b(u_h^n) - b(u_h^{n-1})}{\tau_n} - B_h^n \frac{u_h^n - u_h^{n-1}}{\tau_n}\|_{L^2(K)}, \tag{4.1}$$

where the function B_h^n is defined on each K as $\frac{b'(\Pi_h^n u_h^{n-1}) + b'(u_h^n)}{2}$.

(ii) Error indicators linked to the space discretization

$$\begin{aligned}\eta_K^{n(h)} = \tau_n^{\frac{1}{2}} \, \|\alpha \, \frac{u_h^n - \Pi_h^n u_h^{n-1}}{\tau_n} + \frac{b(u_h^n) - \Pi_h^n b(u_h^{n-1})}{\tau_n} + \nabla \cdot \boldsymbol{q}_h^n\|_{L^2(K)} \\ + \tau_n^{\frac{1}{2}} \, \|\boldsymbol{q}_h^n + \Pi_h^n k \circ b(u_h^{n-1}) \boldsymbol{e}_z\|_{L^2(K)^d} \\ + \sum_{e \in \mathcal{E}_K^0} \tau_n^{\frac{1}{2}} h_e^{-\frac{1}{2}} \, \|[u_h^n]_e\|_{L^2(e)} + \sum_{e \in \mathcal{E}_K^D} \tau_n^{\frac{1}{2}} h_e^{-\frac{1}{2}} \, \|u_h^n - u_{Dh}^n\|_{L^2(e)}.\end{aligned} \tag{4.2}$$

It can be noted that all these indicators only depend on the discrete solution $(u_h^n, \boldsymbol{q}_h^n)_n$ and are very easy to compute (only polynomials of degree at most 1 appear in the norms). However the terms involving u_h^{n-1} are reinterpolated on the new mesh $\mathcal{T}_h^n$ thanks to the operator Π_h^n.

4.3. Upper bounds for the error.

As now standard for multi-step discretizations (see [3] for the introduction of this approach), we proceed in two steps, in order to uncouple the two sources of error. Due to the nonlinearity of all problems, we use the theorem due to Pousin and Rappaz [20] at each step.

With the notation presented in Section 4.1, we first estimate the norm of the term $(u - u_\tau, \boldsymbol{q} - \pi_\tau^+ \boldsymbol{q})$. We introduce the subspace

$$\mathbb{X} = H^1(0, T; L^2(\Omega)). \tag{4.3}$$

Indeed, it is clear that the mapping $\mathcal{F} = (\mathcal{F}_1, \mathcal{F}_2)$ defined with the notation $U = (u, \boldsymbol{q})$ and for a.e. t in $[0, T]$ by

$$\begin{aligned}\forall w \in L^2(\Omega), \quad \langle \mathcal{F}_1(U)(t), w \rangle = \alpha \int_\Omega (\partial_t u)(\boldsymbol{x}, t) w(\boldsymbol{x}) \, d\boldsymbol{x} + \int_\Omega \big(\partial_t b(u)\big)(\boldsymbol{x}, t) w(\boldsymbol{x}) \, d\boldsymbol{x} \\ + \int_\Omega \big(\nabla \cdot \boldsymbol{q}\big)(\boldsymbol{x}, t) w(\boldsymbol{x}) \, d\boldsymbol{x}, \\ \forall \boldsymbol{\varphi} \in H_F(\mathrm{div}, \Omega), \quad \langle \mathcal{F}_2(U)(t), \boldsymbol{\varphi} \rangle = \int_\Omega \boldsymbol{q}(\boldsymbol{x}, t) \cdot \boldsymbol{\varphi}(\boldsymbol{x}) \, d\boldsymbol{x} - \int_\Omega u(\boldsymbol{x}, t) (\nabla \cdot \boldsymbol{\varphi})(\boldsymbol{x}) \, d\boldsymbol{x} \\ + \int_\Omega \big(k \circ b(u)\big)(\boldsymbol{x}, t) \boldsymbol{e}_z \cdot \boldsymbol{\varphi}(\boldsymbol{x}) \, d\boldsymbol{x} + \langle u_D, \boldsymbol{\varphi} \cdot \boldsymbol{n} \rangle_{\Gamma_D},\end{aligned} \tag{4.4}$$

is continuous from $\mathbb{X} \times L^2(0,T;H_F(\mathrm{div},\Omega))$ into $L^2(0,T;L^2(\Omega)) \times L^2(0,T;H_F(\mathrm{div},\Omega)')$. Unfortunately, it is only differentiable on a smaller domain, as stated in the next lemma. We now introduce the space $\mathscr{L}$ of linear mappings from $\mathbb{X} \times L^2(0,T;H_F(\mathrm{div},\Omega))$ into $L^2(0,T;L^2(\Omega)) \times L^2(0,T;H_F(\mathrm{div},\Omega)')$.

Lemma 4.2. *If Assumption* 4.1 *is satisfied, the mapping $\mathcal{F}$ is continuously differentiable on the space $\mathbb{Y} \times L^2(0,T;H(\mathrm{div},\Omega))$ with values in $\mathscr{L}$, where the space $\mathbb{Y}$ is given by*

$$\mathbb{Y} = H^1(0,T;L^\infty(\Omega)). \tag{4.5}$$

Moreover, the mapping: $V \mapsto D\mathcal{F}(V)$ is locally Lipschitz-continuous on this same space.

Proof: We have, for any $Z = (z,\boldsymbol{\psi})$ in $\mathbb{X} \times L^2(0,T;H(\mathrm{div},\Omega))$,

$$\begin{aligned}
\langle D\mathcal{F}_1(U)(t)\cdot Z, w\rangle &= \int_\Omega \big(\alpha + b'(u)\big)(\boldsymbol{x},t)\big(\partial_t z\big)(\boldsymbol{x},t)w(\boldsymbol{x})\,d\boldsymbol{x} \\
&\quad + \int_\Omega b''(u)(\boldsymbol{x},t)z(\boldsymbol{x},t)\big(\partial_t u\big)(\boldsymbol{x},t)w(\boldsymbol{x})\,d\boldsymbol{x} + \int_\Omega \big(\nabla\cdot\boldsymbol{\psi}\big)(\boldsymbol{x},t)w(\boldsymbol{x})\,d\boldsymbol{x}, \\
\langle D\mathcal{F}_2(U)(t)\cdot Z, \boldsymbol{\varphi}\rangle &= \int_\Omega \boldsymbol{\psi}(\boldsymbol{x},t)\cdot\boldsymbol{\varphi}(\boldsymbol{x})\,d\boldsymbol{x} - \int_\Omega z(\boldsymbol{x},t)(\nabla\cdot\boldsymbol{\varphi})(\boldsymbol{x})\,d\boldsymbol{x} \\
&\quad + \int_\Omega \big(k\circ b\big)'(u)(\boldsymbol{x},t)z(\boldsymbol{x},t)\boldsymbol{e}_z\cdot\boldsymbol{\varphi}(\boldsymbol{x})\,d\boldsymbol{x}.
\end{aligned}$$

The continuity and Lipschitz property of $D\mathcal{F}$ is then easily derived from these formulas and the choice of $\mathbb{Y}$, see Assumption 4.1.

It can be checked that equation (2.11) can be written in the abridged form, with obvious notation,

$$\mathcal{F}(U) = 0. \tag{4.6}$$

On the other hand, the residual equation satisfied by $U_\tau = (u_\tau, \pi_\tau^+\boldsymbol{q})$, where the $(u^n,\boldsymbol{q}^n)$ are the solutions of problem (3.1) – (3.2), reads, for all t in $[t_{n-1},t_n]$,

$$\begin{aligned}
&\forall w\in L^2(\Omega),\quad \langle \mathcal{F}_1(U_\tau)(t),w\rangle = \int_\Omega \Big(\partial_t\big(b(u_\tau) - b_\tau(u_\tau)\big)\Big)(\boldsymbol{x},t)w(\boldsymbol{x})\,d\boldsymbol{x}, \\
&\forall \boldsymbol{\varphi}\in H_F(\mathrm{div},\Omega),\quad \langle \mathcal{F}_2(U_\tau)(t),\boldsymbol{\varphi}\rangle = -\int_\Omega (u_\tau - u^n)(\boldsymbol{x},t)(\nabla\cdot\boldsymbol{\varphi})(\boldsymbol{x})\,d\boldsymbol{x} \\
&\qquad\qquad + \int_\Omega \big(k\circ b(u_\tau) - k\circ b(u^{n-1})\big)(\boldsymbol{x},t)\boldsymbol{e}_z\cdot\boldsymbol{\varphi}(\boldsymbol{x})\,d\boldsymbol{x},
\end{aligned} \tag{4.7}$$

where $b_\tau(u_\tau)$ stands for the function which is affine on each interval $[t_{n-1},t_n]$, $1\le n\le N$, and equal to $b(u^n)$ in t_n, $0\le n\le N$. We are thus in a position to apply the theorem of Pousin and Rappaz [20] in its more precise form given in [29, Prop. 2.1].

We need the lifting operator $\mathcal{L}$ which associates with any function g in the dual space of $H^{\frac{1}{2}}_{00}(\Gamma_F)$ the function $\mathbf{grad}\,\chi$, where χ is the solution of the problem

$$\begin{cases} -\Delta\chi = 0 & \text{in } \Omega, \\ \chi = 0 & \text{on } \Gamma_D, \\ \partial_n\chi = g & \text{on } \Gamma_F. \end{cases} \tag{4.8}$$

It is readily checked that $\mathcal{L}$ is continuous from $H^{\frac{1}{2}}_{00}(\Gamma_F)'$ into $H(\mathrm{div},\Omega)$. In what follows, we denote by $[\![\cdot]\!]$ the norm of a continuous linear mapping from a Banach space into another one (without making precise the spaces for simplicity). We finally define the quantities

$$\varepsilon_K^{n(\tau)} = \|\big(b'(u_{h\tau}) - B_h^n\big)\frac{u_h^n - u_h^{n-1}}{\tau_n}\|_{L^2(t_{n-1},t_n;L^2(K))}. \tag{4.9}$$

Proposition 4.3. *If Assumption* 4.1 *is satisfied, for any solution $U = (u, \boldsymbol{q})$ of problem* (2.10) – (2.11) *in $\mathbb{Y} \times L^2(0,T;H(\mathrm{div},\Omega))$ such that $D\mathcal{F}(U)$ is an isomorphism in $\mathscr{L}$, there exists a bounded neighbourhood of U in $\mathbb{Y} \times L^2(0,T;H(\mathrm{div},\Omega))$ such that the following a posteriori estimate holds for any solution $U_\tau = (u_\tau, \pi_\tau^+\boldsymbol{q})$ associated with problems* (3.1) – (3.2) *in this neighbourhood*

$$\begin{aligned} &\sup_{0\le t\le T}\|(u-u_\tau)(\cdot,t)\|_{L^2(\Omega)} + \alpha^{\frac{1}{2}}\|\partial_t(u-u_\tau)\|_{L^2(0,T;L^2(\Omega))} + \|\boldsymbol{q}-\pi_\tau^+\boldsymbol{q}\|_{L^2(0,T;H(\mathrm{div},\Omega))} \\ &\quad\le c\,\Bigg(\Big(\sum_{n=1}^{N}\sum_{K\in\mathcal{T}_h^n}\big((\eta_K^{n(\tau)})^2 + (\varepsilon_K^{n(\tau)})^2\big)\Big)^{\frac{1}{2}} + \|u_\tau - u_{h\tau}\|_{L^2(0,T;L^2(\Omega))} \\ &\qquad\qquad + \max_{0\le n\le N}\|u^n - u_h^n\|_{L^2(\Omega)} + \alpha^{\frac{1}{2}}\|\partial_t(u_\tau - u_{h\tau})\|_{L^2(0,T;L^2(\Omega))} \\ &\qquad\qquad\qquad + \|f - \pi_\tau^+ f\|_{L^2(0,T;H^{\frac{1}{2}}_{00}(\Gamma_F)')}\Bigg). \end{aligned} \tag{4.10}$$

where the constant c only depends on α and the norm $[\![D\mathcal{F}(U)^{-1}]\!]$.

Proof: As previously, we set:

$$\tilde{U} = \big(u - u_0, \boldsymbol{q} - \mathcal{L}f(\cdot,t)\big), \qquad \tilde{U}_\tau = \big(u_\tau - u_0, \pi_\tau^+(\boldsymbol{q} - \mathcal{L}f(\cdot,t))\big).$$

Due to the continuity of $\mathcal{L}$ and with obvious notation, we have

$$\|\boldsymbol{q}-\pi_\tau^+\boldsymbol{q}\|_{L^2(0,T;H(\mathrm{div},\Omega))} \le \|\tilde{\boldsymbol{q}} - \widetilde{\pi_\tau^+\boldsymbol{q}}\|_{L^2(0,T;H(\mathrm{div},\Omega))} + c\,\|f-\pi_\tau^+ f\|_{L^2(0,T;H^{\frac{1}{2}}_{00}(\Gamma_F)')}.$$

Owing to Lemma 4.2, applying a slight extension of [29, Prop. 2.1] to the mapping $\widetilde{\mathcal{F}}$ defined by $\widetilde{\mathcal{F}}(\tilde{U}) = \mathcal{F}(U)$ yields

$$\begin{aligned} &\sup_{0\le t\le T}\|(u-u_\tau)(\cdot,t)\|_{L^2(\Omega)} + \alpha^{\frac{1}{2}}\|\partial_t(u-u_\tau)\|_{L^2(0,T;L^2(\Omega))} + \|\tilde{\boldsymbol{q}} - \widetilde{\pi_\tau^+\boldsymbol{q}}\|_{L^2(0,T;H(\mathrm{div},\Omega))} \\ &\quad\le c\,\Bigg(\|\partial_t\big(b(u_\tau) - b_\tau(u_\tau)\big)\|_{L^2(0,T;L^2(\Omega))} \\ &\qquad + \|u_\tau - \pi_\tau^+ u\|_{L^2(0,T;L^2(\Omega))} + \|k\circ b(u_\tau) - k\circ b(\pi_\tau^- u)\|_{L^2(0,T;L^2(\Omega))}\Bigg). \end{aligned}$$

We now evaluate successively the three terms in the right-hand side of this inequality, beginning by the last two ones which are easier to handle.
1) We start from the triangle inequality

$$\|u_\tau - \pi_\tau^+ u\|_{L^2(0,T;L^2(\Omega))} \le \|u_\tau - u_{h\tau}\|_{L^2(0,T;L^2(\Omega))} + \|\pi_\tau^+(u-u_h)\|_{L^2(0,T;L^2(\Omega))} + \|u_{h\tau} - \pi_\tau^+ u_h\|_{L^2(0,T;L^2(\Omega))},$$

and observe by the same arguments as in [3, Lemma 2.1] that

$$\|\pi_\tau^+(u-u_h)\|_{L^2(0,T;L^2(\Omega))} = \Big(\sum_{n=1}^N \tau_n \|u^n - u_h^n\|_{L^2(\Omega)}^2\Big)^{\frac12} \le c\,\|u_\tau - u_{h\tau}\|_{L^2(0,T;L^2(\Omega))}.$$

On the other hand, it follows from the formula, valid on each interval $]t_{n-1}, t_n]$,

$$u_\tau = u^n - \frac{t_n - t}{\tau_n}(u^n - u^{n-1}).$$

that

$$\|u_{h\tau} - \pi_\tau^+ u_h\|_{L^2(0,T;L^2(\Omega))}^2 = \sum_{n=1}^N \frac{\tau_n}{3}\|u_h^n - u_h^{n-1}\|_{L^2(\Omega)}^2.$$

All this yields

$$\|u_\tau - \pi_\tau^+ u\|_{L^2(0,T;L^2(\Omega))} \le c\,\|u_\tau - u_{h\tau}\|_{L^2(0,T;L^2(\Omega))} + \frac{1}{\sqrt{3}}\Big(\sum_{n=1}^N \tau_n\,\|u_h^n - u_h^{n-1}\|_{L^2(\Omega)}^2\Big)^{\frac12}. \tag{4.11}$$

2) Since both functions k and b are Lipschitz-continuous, we have

$$\|k\circ b(u_\tau) - k\circ b(\pi_\tau^- u)\|_{L^2(0,T;L^2(\Omega))} \le c\,\|u_\tau - \pi_\tau^- u\|_{L^2(0,T;L^2(\Omega))}.$$

By using the same arguments as previously (in particular a modified version of [3, Lemma 2.1]) and the formula

$$u_\tau = u^{n-1} + \frac{t - t_{n-1}}{\tau_n}(u^n - u^{n-1}),$$

we derive

$$\|k\circ b(u_\tau) - k\circ b(\pi_\tau^- u)\|_{L^2(0,T;L^2(\Omega))} \le c\left(\|u_\tau - u_{h\tau}\|_{L^2(0,T;L^2(\Omega))} + \Big(\sum_{n=1}^N \tau_n\,\|u_h^n - u_h^{n-1}\|_{L^2(\Omega)}^2\Big)^{\frac12}\right). \tag{4.12}$$

3) Here, we use the triangle inequality

$$\|\partial_t\big(b(u_\tau) - b_\tau(u_\tau)\big)\|_{L^2(0,T;L^2(\Omega))} \le \|\partial_t\big(b(u_\tau) - b(u_{h\tau})\big)\|_{L^2(0,T;L^2(\Omega))} + \|\partial_t\big(b_\tau(u_\tau) - b_\tau(u_{h\tau})\big)\|_{L^2(0,T;L^2(\Omega))} + \|\partial_t\big(b(u_{h\tau}) - b_\tau(u_{h\tau})\big)\|_{L^2(0,T;L^2(\Omega))}.$$

To bound the first term, we observe that

$$\partial_t\big(b(u_\tau)-b(u_{h\tau})\big)=b'(u_\tau)\,\partial_t u_\tau-b'(u_{h\tau})\,\partial_t u_{h\tau}\\ =\big(b'(u_\tau)-b'(u_{h\tau})\big)\,\partial_t u_\tau+b'(u_{h\tau})\,\partial_t(u_\tau-u_{h\tau}).$$

We thus combine the boundedness of b' and its Lipschitz property, see Assumption 4.1, with the boundedness of $\partial_t u_\tau$ in $L^\infty(\Omega)$ (indeed, u_τ belongs to a bounded neighbourhoood of u in $\mathbb{Y}$). This leads to

$$\|\partial_t\big(b(u_\tau)-b(u_{h\tau})\big)\|_{L^2(0,T;L^2(\Omega))}\\ \le c\Big(\max_{0\le n\le N}\|u^n-u_h^n\|_{L^2(\Omega)}+\|\partial_t(u_\tau-u_{h\tau})\|_{L^2(0,T;L^2(\Omega))}\Big).$$

To bound the second term, we note that $b_\tau(u_\tau)$ is the Lagrange interpolate of $b(u_\tau)$ in piecewise affine functions. Thus, using the stability of the corresponding interpolation operator in $H^1(0,T)$, we obtain

$$\|\partial_t\big(b_\tau(u_\tau)-b_\tau(u_{h\tau})\big)\|_{L^2(0,T;L^2(\Omega))}\le c\,\|\partial_t\big(b(u_\tau)-b(u_{h\tau})\big)\|_{L^2(0,T;L^2(\Omega))},$$

and we use the previous estimate. Finally, to bound the third term, we use the expansion on the interval $[t_{n-1},t_n]$

$$\partial_t\big(b(u_{h\tau})-b_\tau(u_{h\tau})\big)=b'(u_{h\tau})\frac{u_h^n-u_h^{n-1}}{\tau_n}-\frac{b(u_h^n)-b(u_h^{n-1})}{\tau_n}.$$

Thus, it follows from the definition (4.9) and a triangle inequality that

$$\|\partial_t\big(b(u_{h\tau})-b_\tau(u_{h\tau})\big)\|_{L^2(t_{n-1},t_n;L^2(K))}\\ \le\tau_n^{\frac12}\,\|\frac{b(u_h^n)-b(u_h^{n-1})}{\tau_n}-B_h^n\frac{u_h^n-u_h^{n-1}}{\tau_n}\|_{L^2(K)}+\varepsilon_K^{n(\tau)}.$$

Combining all this gives

$$\begin{aligned}&\|\partial_t\big(b(u_\tau)-b_\tau(u_\tau)\big)\|_{L^2(0,T;L^2(\Omega))}\\ &\quad\le c\Bigg(\max_{0\le n\le N}\|u^n-u_h^n\|_{L^2(\Omega)}+\|\partial_t(u_\tau-u_{h\tau})\|_{L^2(0,T;L^2(\Omega))}\\ &\qquad+\Big(\sum_{n=1}^N\sum_{K\in\mathcal{T}_h^n}\big(\tau_n\,\|B_h^n\frac{u_h^n-u_h^{n-1}}{\tau_n}-\frac{b(u_h^n)-b(u_h^{n-1})}{\tau_n}\|_{L^2(K)}^2+(\varepsilon_K^{n(\tau)})^2\big)\Big)^{\frac12}\Bigg).\end{aligned}\tag{4.13}$$

Owing to the definition (4.1) of the $\eta_K^{n(\tau)}$, the desired estimate is now a direct consequence of (4.11), (4.12), and (4.13).

Remark 4.4. Due to Assumption 4.1, we have

$$\varepsilon_K^{n(\tau)}\le\tau_n^{\frac12}\Big(\sup_{x\in I_h^n}|b''(x)|\Big)\,\|u_h^n-u_h^{n-1}\|_{L^\infty(K)}\|\frac{u_h^n-u_h^{n-1}}{\tau_n}\|_{L^2(K)},\tag{4.14}$$

where I_h^n stands for a small interval only depending on the minimal and maximal values of u_h^{n-1} and u_h^n. So, at least when the quantity $\|u_h^n - u_h^{n-1}\|_{L^\infty(\Omega)}$ tends to zero, which is rather likely, the quantity $\varepsilon_K^{n(\tau)}$ can be considered as negligible. Moreover, it is zero when b' is constant on I_h^n.

We now bound the error betweeen $U_\tau = (u_\tau, \pi_\tau^+ \boldsymbol{q})$ and $U_{h\tau} = (u_{h\tau}, \pi_\tau^+ \boldsymbol{q}_h)$ by very similar but simpler arguments. Indeed, let us introduce the mapping $\mathcal{F}_\tau = (\mathcal{F}_{1\tau}, \mathcal{F}_{2\tau})$ defined with the notation $U = (u, \boldsymbol{q})$ and for a.e. t in $[0,T]$ by

$$\begin{aligned}
\forall w \in L^2(\Omega), \quad \langle \mathcal{F}_{1\tau}(U)(t), w\rangle &= \alpha \int_\Omega (\partial_t u)(\boldsymbol{x},t) w(\boldsymbol{x})\, d\boldsymbol{x} + \int_\Omega \big(\partial_t b_\tau(u)\big)(\boldsymbol{x},t) w(\boldsymbol{x})\, d\boldsymbol{x} \\
&\qquad + \int_\Omega \big(\nabla \cdot \boldsymbol{q}\big)(\boldsymbol{x},t) w(\boldsymbol{x})\, d\boldsymbol{x}, \\
\forall \boldsymbol{\varphi} \in H_F(\mathrm{div},\Omega), \quad \langle \mathcal{F}_{2\tau}(U)(t), \boldsymbol{\varphi}\rangle &= \int_\Omega \boldsymbol{q}(\boldsymbol{x},t) \cdot \boldsymbol{\varphi}(\boldsymbol{x})\, d\boldsymbol{x} - \int_\Omega \pi_\tau^+ u(\boldsymbol{x},t) (\nabla \cdot \boldsymbol{\varphi})(\boldsymbol{x})\, d\boldsymbol{x} \\
&\qquad + \int_\Omega \big(k \circ b(\pi_\tau^- u)\big)(\boldsymbol{x},t) \boldsymbol{e}_z \cdot \boldsymbol{\varphi}(\boldsymbol{x})\, d\boldsymbol{x} + \langle u_D, \boldsymbol{\varphi} \cdot \boldsymbol{n}\rangle_{\Gamma_D}.
\end{aligned} \tag{4.15}$$

The solution $U_\tau = (u_\tau, \pi_\tau^+ \boldsymbol{q})$ associated with problems $(3.1) - (3.2)$ satisfies $\mathcal{F}_\tau(U_\tau) = 0$, while the solution $U_{h\tau} = (u_{h\tau}, \pi_\tau^+ \boldsymbol{q}_h)$ associated with problems $(3.12) - (3.13)$ satisfies, for all t in $[t_{n-1}, t_n]$

$$\begin{aligned}
\langle \mathcal{F}_{1\tau}(U_{h\tau})(t), w\rangle &= \alpha \int_\Omega \big(\frac{u_h^n - u_h^{n-1}}{\tau_n}\big)(\boldsymbol{x}) w(\boldsymbol{x})\, d\boldsymbol{x} \\
&\qquad + \int_\Omega \big(\frac{b(u_h^n) - b(u_h^{n-1})}{\tau_n}\big)(\boldsymbol{x}) w(\boldsymbol{x})\, d\boldsymbol{x} + \int_\Omega \big(\nabla \cdot \boldsymbol{q}_h^n\big)(\boldsymbol{x}) w(\boldsymbol{x})\, d\boldsymbol{x}, \\
\langle \mathcal{F}_{2\tau}(U_{h\tau})(t), \boldsymbol{\varphi}\rangle &= \int_\Omega \boldsymbol{q}_h^n(\boldsymbol{x}) \cdot \boldsymbol{\varphi}(\boldsymbol{x})\, d\boldsymbol{x} - \int_\Omega u_n^n(\boldsymbol{x}) (\nabla \cdot \boldsymbol{\varphi})(\boldsymbol{x})\, d\boldsymbol{x} \\
&\qquad + \int_\Omega \big(k \circ b(u_h^{n-1})\big)(\boldsymbol{x}) \boldsymbol{e}_z \cdot \boldsymbol{\varphi}(\boldsymbol{x})\, d\boldsymbol{x} + \langle u_D, \boldsymbol{\varphi} \cdot \boldsymbol{n}\rangle_{\Gamma_D}.
\end{aligned} \tag{4.16}$$

We omit the proof of the next lemma, since it results from the definition of $D\mathcal{F}_\tau$.

Lemma 4.5. *If Assumption* 4.1 *is satisfied, the mapping $\mathcal{F}_\tau$ is continuously differentiable on the space $\mathbb{Y} \times L^2(0,T; H(\mathrm{div},\Omega))$ with values in $\mathscr{L}$ and, moreover, the operator $D\mathcal{F}_\tau$ is locally Lipschitz-continuous on this same space.*

We can now prove the second error estimate. This requires the quantities

$$\begin{aligned}
\varepsilon_K^{n(h)} = \|(\mathrm{Id} - \Pi_h^n)\big(\frac{u_h^n - u_h^{n-1}}{\tau_n}\big)\|_{L^2(K)} &+ \|(\mathrm{Id} - \Pi_h^n)\big(\frac{b(u_h^n) - b(u_h^{n-1})}{\tau_n}\big)\|_{L^2(K)} \\
&+ \|(\mathrm{Id} - \Pi_h^n) k \circ b(u_h^{n-1})\|_{L^2(K)}.
\end{aligned} \tag{4.17}$$

As already hinted in Remark 3.5, they are not 0 only for the K where the triangulation $\mathcal{T}_h^n$ is coarsened with respect to $\mathcal{T}_h^{n-1}$.

Proposition 4.6. *If Assumption* 4.1 *is satisfied, for any solution* $U_\tau = (u_\tau, \pi_\tau^+ \boldsymbol{q})$ *associated with problems* (3.1) – (3.2) *in* $\mathbb{Y} \times L^2(0,T; H(\mathrm{div},\Omega))$ *such that* $D\mathcal{F}_\tau(U_\tau)$ *is an isomorphism in* $\mathscr{L}$*, there exists a neighbourhood of* U_τ *in* $\mathbb{Y} \times L^2(0,T; H(\mathrm{div},\Omega))$ *such that the following a posteriori estimate holds for any solution* $U_{h\tau} = (u_{h\tau}, \pi_\tau^+ \boldsymbol{q}_h)$ *associated with problems* (3.12) – (3.13) *in this neighbourhood*

$$\begin{aligned}
&\sup_{0\le t\le T} \|(u_\tau - u_{h\tau})(\cdot,t)\|_{L^2(\Omega)} + \alpha^{\frac12} \|\partial_t(u_\tau - u_{h\tau})\|_{L^2(0,T;L^2(\Omega))} \\
&\qquad\qquad + \|\pi_\tau^+(\boldsymbol{q} - \boldsymbol{q}_h)\|_{L^2(0,T;H(\mathrm{div},\Omega))} \\
&\le c(\tau) \Bigg(\Big(\sum_{n=1}^{N} \sum_{K\in\mathcal{T}_h^n} \big((\eta_K^{n(h)})^2 + (\varepsilon_K^{n(h)})^2 \big) \Big)^{\frac12} + \|u_0 - u_{0h}\|_{L^2(\Omega)} \\
&\qquad + \Big(\sum_{n=1}^{N} \tau_n \left(\|f(\cdot,t_n) - f_h^n\|^2_{H_{00}^{\frac12}(\Gamma_F)'} + \|u_D(\cdot,t_n) - u_{Dh}^n\|^2_{H^{\frac12}(\Gamma_D)} \right) \Big)^{\frac12} \Bigg),
\end{aligned} \tag{4.18}$$

where $c(\tau)$ *only depends on the norm* $[\![D\mathcal{F}_\tau(U_\tau)^{-1}]\!]$.

Proof: Here, we set:

$$\tilde{U}_\tau = U_\tau - (u_0, \pi_\tau^+ \ell), \qquad \text{with} \quad \ell^n = \mathcal{L} f(\cdot,t_n),$$
$$\tilde{U}_{h\tau} = U_{h\tau} - (u_{0h}, \pi_\tau^+ \ell_h), \qquad \text{with} \quad \ell_h^n = \mathcal{L} f_h^n,$$

where the operator $\mathcal{L}$ is defined from (4.8). To bound the error, we observe that

$$\begin{aligned}
\|U_\tau - U_{h\tau}\|_{\mathbb{X}\times L^2(0,T;H(\mathrm{div},\Omega))} &\le \|\tilde{U}_\tau - \tilde{U}_{h\tau}\|_{\mathbb{X}\times L^2(0,T;H(\mathrm{div},\Omega))} \\
&\quad + \|u_0 - u_{0h}\|_{L^2(\Omega)} + \Big(\sum_{n=1}^{N} \tau_n \, \|\mathcal{L}\big(f(\cdot,t_n) - f_h^n\big)\|^2_{H(\mathrm{div},\Omega)} \Big)^{\frac12},
\end{aligned}$$

and bounding the last term follows from the continuity of $\mathcal{L}$. To estimate the first one, we apply [29, Prop. 2.1] to the function $\widetilde{\mathcal{F}}_\tau$ defined by $\widetilde{\mathcal{F}}_\tau(\tilde{U}_\tau) = \mathcal{F}_\tau(U_\tau)$. This gives

$$\begin{aligned}
&\|\tilde{U}_\tau - \tilde{U}_{h\tau}\|_{\mathbb{X}\times L^2(0,T;H(\mathrm{div},\Omega))} \\
&\qquad \le c(\tau) \Big(\|\mathcal{F}_{1\tau}(U_{h\tau})\|_{L^2(0,T;L^2(\Omega))} + \|\mathcal{F}_{2\tau}(U_{h\tau})\|_{L^2(0,T;H_F(\mathrm{div},\Omega)')} \Big).
\end{aligned}$$

Owing to (4.16), evaluating the right-hand side relies on an integration by parts on each K for the second component $\mathcal{F}_{2\tau}(U_{h\tau})$ and also on the inverse inequality, valid for each φ_h in $\mathcal{P}_1(e)$, $e \in \mathcal{E}_K^0 \cup \mathcal{E}_K^F$, (note that this requires the introduction of u_{Dh}^n)

$$\|\varphi_h\|_{H^{\frac12}(e)} \le c\, h_K^{-\frac12} \, \|\varphi_h\|_{L^2(e)}.$$

We conclude with triangle inequalities which involve the $\varepsilon_K^{n(h)}$.

Remark 4.7. The following equality holds for each piecewise affine function v_τ,

$$\begin{aligned}\sup_{0\le t\le T}\|v_\tau(\cdot,t)\|_{L^2(\Omega)}+\alpha^{\frac12}\|\partial_t v_\tau\|_{L^2(0,T;L^2(\Omega))}&\\ =\max_{0\le n\le N}\|v^n\|_{L^2(\Omega)}+\big(\alpha\sum_{n=1}^{N}\tau_n\,\|\frac{v^n-v^{n-1}}{\tau_n}\|^2_{L^2(\Omega)}\big)^{\frac12}.&\end{aligned}\tag{4.19}$$

So the norm which appears in the left-hand side of (4.18) is in fact semi-discrete in time.

To make the statement of Proposition 4.6 more complete, we now prove that, in most cases, the quantity $[\![D\mathcal{F}_\tau(U_\tau)^{-1}]\!]$ is bounded independently of τ.

Lemma 4.8. *Assume that the solution U_τ associated with problems (3.1) – (3.2) satisfies*

$$\lim_{|\tau|\to 0}\max_{1\le n\le N}\|u^n-u^{n-1}\|_{L^\infty(\Omega)}=0.\tag{4.20}$$

Let $U=(u,\boldsymbol{q})$ be any pair in $\mathbb{Y}\times L^2(0,T;H(\mathrm{div},\Omega))$ such that $D\mathcal{F}(U)$ is an isomorphism in $\mathscr{L}$. Thus, there exists a constant $\tau_0>0$ and a neighbourhood of U in $\mathbb{Y}\times L^2(0,T;H(\mathrm{div},\Omega))$ such that, for any τ, $|\tau|\le\tau_0$, and for any pair U_τ in this neighbourhood,
(i) $D\mathcal{F}_\tau(U_\tau)$ is an isomorphism in this same space,
(ii) the norm of its inverse is bounded independently of τ.

Proof: We use the expansion

$$D\mathcal{F}_\tau(U_\tau)=D\mathcal{F}(U)-\big(D\mathcal{F}(U)-D\mathcal{F}(U_\tau)\big)-\big(D\mathcal{F}(U_\tau)-D\mathcal{F}_\tau(U_\tau)\big).$$

So, there exists a constant c only depending on $[\![D\mathcal{F}(U)^{-1}]\!]$ such that the desired result holds when

$$[\![D\mathcal{F}(U)-D\mathcal{F}(U_\tau)]\!]\le c,\qquad [\![D\mathcal{F}(U_\tau)-D\mathcal{F}_\tau(U_\tau)]\!]\le c.$$

For an appropriate choice of the neighbourhood of U, the first inequality is a direct consequence of the Lipschitz property of $D\mathcal{F}$, see Lemma 4.2. On the other hand, we have, for each t in $[t_{n-1},t_n]$,

$$\begin{aligned}\langle\big(D\mathcal{F}_1-D\mathcal{F}_{1\tau}\big)(U_\tau)(t)\cdot Z,w\rangle&=\int_\Omega\big(b'(u_\tau)-\frac{b(u^n)-b(u^{n-1})}{\tau_n}\big)(\boldsymbol{x},t)\big(\partial_t z\big)(\boldsymbol{x},t)w(\boldsymbol{x})\,d\boldsymbol{x}\\&\quad+\int_\Omega\big(b''(u_\tau)-\frac{b'(u^n)-b'(u^{n-1})}{\tau_n}\big)(\boldsymbol{x},t)z(\boldsymbol{x},t)\big(\partial_t u\big)(\boldsymbol{x},t)w(\boldsymbol{x})\,d\boldsymbol{x},\\ \langle\big(D\mathcal{F}_2-D\mathcal{F}_{2\tau}\big)(U_\tau)(t)\cdot Z,\boldsymbol{\varphi}\rangle&=-\int_\Omega(z-\pi_\tau^+ z)(\boldsymbol{x},t)(\nabla\cdot\boldsymbol{\varphi})(\boldsymbol{x})\,d\boldsymbol{x}\\&\quad+\int_\Omega\big((k\circ b)'(u_\tau)-(k\circ b)'(\pi_\tau^- u)\big)(\boldsymbol{x},t)z(\boldsymbol{x},t)\boldsymbol{e}_z\cdot\boldsymbol{\varphi}(\boldsymbol{x})\,d\boldsymbol{x},\end{aligned}$$

so that the second inequality follows from assumption (4.20) combined with the properties of b and k.

Note that assumption (4.20) is not restrictive and should follow from the convergence of the Euler scheme.

4.4. Upper bounds for the indicators.

We bound successively all the indicators introduced in Section 4.2. Note that evaluating the indicators $\eta_K^{n(\tau)}$ requires a preliminary lemma that we now state and prove. We need the following notation: For each z_0 in $\mathbb{R}$, Ω^{z_0} and $\Gamma_D^{z_0}$ denotes the intersection of Ω and Γ_D, respectively, with the line or plane $z = z_0$.

Lemma 4.9. *Assume that, for each z in $\mathbb{R}$, and for each connected component Ω_k^z of Ω^z, the measure of $\Gamma_D^z \cap \partial\Omega_k^z$ is positive. For any function g in $L^2(\Omega)$, there exists a function $\boldsymbol{\varphi} = (\varphi_x, \varphi_y, 0)$ in $H_F(\mathrm{div}, \Omega)$ such that*

$$\nabla \cdot \boldsymbol{\varphi} = g \quad \text{in } \Omega \qquad \text{and} \qquad \|\boldsymbol{\varphi}\|_{H(\mathrm{div},\Omega)} \leq c\, \|g\|_{L^2(\Omega)}. \tag{4.21}$$

Proof: Assuming that the domain Ω lies between the lines or planes $z = z_1$ and $z = z_2$. For a.e. z in $]z_1, z_2[$, and for each connected component Ω_k^z, we solve the $(d-1)$-dimensional problem, in dimension $d = 3$ for instance: Find ψ in $H_D^1(\Omega_k^z)$ (with obvious notation for this space) such that

$$\forall \chi \in H_D^1(\Omega_k^z), \quad \int_{\Omega_k^z} \big((\partial_x\psi)(x,y)(\partial_x\chi)(x,y) + (\partial_y\psi)(x,y)(\partial_y\chi)(x,y)\big)\, dxdy \\ = -\int_{\Omega_k^z} g(x,y,z)\chi(x,y)\, dxdy.$$

Thus, the vector field $\boldsymbol{\varphi} = (\partial_x\psi, \partial_y\psi, 0)$ satisfies the first part of (4.21). On the other hand, the second part of (4.21) follows from the Poincaré-Friedrichs inequality

$$\forall \chi \in H_D^1(\Omega_k^z), \quad \|\chi\|_{L^2(\Omega_k^z)} \leq c_z \left(\|\partial_x\chi\|_{L^2(\Omega_k^z)} + \|\partial_y\chi\|_{L^2(\Omega_k^z)}\right).$$

To conclude, we note that that each Ω_k^z is, up to an homothety, an interval or a polygon in a finite family of polygons (where "finite" means only depending on Ω) with possible small diameter. Thus, it is readily checked that c_z is bounded independently of z.

Note that assumptions of Lemma 4.9 are not restrictive and can be avoided by using more sophisticated arguments.

Proposition 4.10. *If the assumptions of Lemma* 4.9 *hold, the following estimate holds for the indicators* $\eta_K^{n(\tau)}$ *defined in* (4.1), $1 \le n \le N$*:*

$$\begin{aligned}
&\Big(\sum_{K\in\mathcal{T}_h^n} (\eta_K^{n(\tau)})^2\Big)^{\frac12} \\
&\quad\le c\,\Big(\alpha\,\|\partial_t(u-u_\tau)\|_{L^2(t_{n-1},t_n;L^2(\Omega))} + \|\partial_t\big(b(u)-b(u_\tau)\big)\|_{L^2(t_{n-1},t_n;L^2(\Omega))} \\
&\qquad + \|\boldsymbol{q}-\pi_\tau^+\boldsymbol{q}\|_{L^2(t_{n-1},t_n;H(\mathrm{div},\Omega))} + \|\partial_t\big(b(u_\tau)-b(u_{h\tau})\big)\|_{L^2(t_{n-1},t_n;L^2(\Omega))} \\
&\qquad\qquad + \Big(\sum_{K\in\mathcal{T}_h^n} (\varepsilon_K^{n(\tau)})^2\Big)^{\frac12}\Big).
\end{aligned} \tag{4.22}$$

Proof: We bound successively the two terms in the $\eta_K^{n(\tau)}$.
1) Using the first lines in (4.4) and in the residual equation (4.7), we obtain, for all t in $[t_{n-1},t_n]$,

$$\begin{aligned}
\|\partial_t\big(b(u_\tau)-b_\tau(u_\tau)\big)\|_{L^2(\Omega)} &\le \alpha\,\|\partial_t(u-u_\tau)\|_{L^2(\Omega))} + \|\partial_t\big(b(u)-b(u_\tau)\big)\|_{L^2(\Omega)} \\
&\quad + \|\boldsymbol{q}-\boldsymbol{q}_\tau\|_{H(\mathrm{div},\Omega)}.
\end{aligned}$$

We also have the triangle inequality

$$\begin{aligned}
\|\partial_t\big(b(u_{h\tau})-b_\tau(u_{h\tau})\big)\|_{L^2(\Omega)} &\le \|\partial_t\big(b(u_\tau)-b_\tau(u_\tau)\big)\|_{L^2(\Omega)} \\
&\quad + \|\partial_t\big(b(u_\tau)-b(u_{h\tau})\big)\|_{L^2(\Omega)} + \|\partial_t\big(b_\tau(u_\tau)-b_\tau(u_{h\tau})\big)\|_{L^2(\Omega)}.
\end{aligned}$$

Integrating the square of these inequalities on $[t_{n-1},t_n]$, noting that b_τ is the Lagrange intepolate of b in the space of piecewise affine functions and using the stability of the corresponding interpolation operator in $H^1(0,T)$ finally give

$$\begin{aligned}
&\|\partial_t\big(b(u_{h\tau})-b_\tau(u_{h\tau})\big)\|_{L^2(t_{n-1},t_n;L^2(\Omega))} \\
&\quad\le c\,\Big(\alpha\,\|\partial_t(u-u_\tau)\|_{L^2(t_{n-1},t_n;L^2(\Omega))} + \|\partial_t\big(b(u)-b(u_\tau)\big)\|_{L^2(t_{n-1},t_n;L^2(\Omega))} \\
&\qquad + \|\boldsymbol{q}-\boldsymbol{q}_\tau\|_{L^2(t_{n-1},t_n;H(\mathrm{div},\Omega))} + \|\partial_t\big(b(u_\tau)-b(u_{h\tau})\big)\|_{L^2(t_{n-1},t_n;L^2(\Omega))}\Big).
\end{aligned}$$

Then, the desired bound for the last term in (4.1) follows from the definition (4.9) of the $\varepsilon_K^{n(\tau)}$ and a triangle inequality.
2) With the function $g = u_\tau - u^n$, we associate the function $\boldsymbol{\varphi}$ exhibited in Lemma 4.9. By inserting this $\boldsymbol{\varphi}$ in (4.11) and noting that $\mathcal{F}_2(U)$ is zero, we obtain the bound for the first term in (4.1).

Bounding the indicators $\eta_K^{n(h)}$ defined in (4.2) relies on more standard arguments, see [29, Chap. 3].

Proposition 4.11. *The following estimate holds for the indicators $\eta_K^{n(h)}$ defined in* (4.2), $1 \le n \le N$, $K \in \mathcal{T}_h^n$*:*

$$\eta_K^{n(h)} \le c \Big(\alpha \, \|\partial_t(u_\tau - u_{h\tau})\|_{L^2(t_{n-1},t_n;L^2(\omega_K))} + \|\partial_t\big(b(u_\tau) - b(u_{h\tau})\big)\|_{L^2(t_{n-1},t_n;L^2(\omega_K))} + \|\pi_\tau^+(u_\tau - u_{h\tau})\|_{L^2(t_{n-1},t_n;L^2(\omega_K))} + \|\pi_\tau^+(\boldsymbol{q} - \boldsymbol{q}_h)\|_{L^2(t_{n-1},t_n;H(\mathrm{div},\omega_K))} + \varepsilon_K^{n(h)} \Big). \tag{4.23}$$

Proof: There also, we bound successively the different terms in the $\eta_K^{n(h)}$.
1) For each K in $\mathcal{T}_h^n$, we set

$$w_K = \begin{cases} \big(\alpha \, \frac{u_h^n - \Pi_h^n u_h^{n-1}}{\tau_n} + \frac{b(u_h^n) - \Pi_h^n b(u_h^{n-1})}{\tau_n} + \nabla \cdot \boldsymbol{q}_h^n\big) \, \psi_K & \text{on } K, \\ 0 & \text{on } \Omega \setminus K, \end{cases}$$

where ψ_K stands for the bubble function on K (equal to the product of the barycentric coordinates associated with the vertices of K). Taking w equal to w_K in the first lines of (4.15) and (4.16) and using triangle inequalities thus yield

$$\|\big(\alpha \, \frac{u_h^n - \Pi_h^n u_h^{n-1}}{\tau_n} + \frac{b(u_h^n) - \Pi_h^n b(u_h^{n-1})}{\tau_n} + \nabla \cdot \boldsymbol{q}_h^n\big)\psi_K^{\frac{1}{2}}\|_{L^2(K)}^2 \le \Big(\alpha \, \|\partial_t(u_\tau - u_{h\tau})\|_{L^2(K)} + \|\partial_t\big(b(u_\tau) - b(u_{h\tau})\big)\|_{L^2(K)} + \|\pi_\tau^+(\boldsymbol{q} - \boldsymbol{q}_h)\|_{H(\mathrm{div},K)} + \varepsilon_K^{n(h)} \Big) \|w_K\|_{L^2(K)}.$$

By noting that w_K is a constant times ψ_K on K, using the inverse inequalities (see [29, Lemma 1.3] for instance)

$$\forall v \in \mathcal{P}_0(K), \quad \|v\|_{L^2(K)} \le c \, \|v \, \psi_K^{\frac{1}{2}}\|_{L^2(K)} \quad \text{and} \quad \|v \, \psi_K\|_{L^2(K)} \le \|v \, \psi_K^{\frac{1}{2}}\|_{L^2(K)},$$

and integrating the square of the previous estimate between t_{n-1} and t_n, we obtain the desired bound for $\tau_n^{\frac{1}{2}} \, \|\alpha \, \frac{u_h^n - \Pi_h^n u_h^{n-1}}{\tau_n} + \frac{b(u_h^n) - \Pi_h^n b(u_h^{n-1})}{\tau_n} + \nabla \cdot \boldsymbol{q}_h^n\|_{L^2(K)}$.
2) Similarly, we set

$$\boldsymbol{\varphi}_K = \begin{cases} \big(\boldsymbol{q}_h^n + \Pi_h^n k \circ b(u_h^{n-1}) \boldsymbol{e}_z\big) \, \psi_K & \text{on } K, \\ \mathbf{0} & \text{on } \Omega \setminus K. \end{cases}$$

By taking $\boldsymbol{\varphi}$ equal to $\boldsymbol{\varphi}_K$ in the second lines of (4.15) and (4.16) and using a further integration by parts together wih the same inverse inequality as above, we derive the desired bound for $\tau_n^{\frac{1}{2}} \, \|\boldsymbol{q}_h^n + \Pi_h^n k \circ b(u_h^{n-1}) \boldsymbol{e}_z\|_{L^2(K)^d}$.
3) For each e in $\mathcal{E}_K^0$, denoting by K and K' the two elements of $\mathcal{T}_h$ that share e, we set

$$\boldsymbol{\varphi}_e = \begin{cases} \mathcal{L}_{e,\kappa}\big([u_h^n]_e \, \psi_e\big) & \text{on } \kappa \in \big\{K, K'\big\}, \\ 0 & \text{on } \Omega \setminus (K \cup K'). \end{cases}$$

Here ψ_e is now the bubble function on e and $\mathcal{L}_{e,\kappa}$ is a lifting operator of the normal trace; it is defined from $L^2(e)$ into the space of functions in $H(\mathrm{div},\kappa)$ with zero normal traces on $\partial K \setminus e$ and is constructed from the harmonic lifting operator on a reference element $\hat{K}$. The following inequality is thus readily checked, for any q_e in $\mathcal{P}_0(e)$

$$\|\mathcal{L}_{e,\kappa}(q_e)\|_{H(\mathrm{div},\kappa)} \leq c\, h_e^{\frac{1}{2}}\, \|q_e\|_{L^2(e)}.$$

Taking $\boldsymbol{\varphi}$ equal to $\boldsymbol{\varphi}_e$ in the second lines of (4.15) and (4.16), using this inequality and the previous results lead to the bound for $\tau_n^{\frac{1}{2}} h_e^{-\frac{1}{2}}\, \|[u_h^n]_e\|_{L^2(e)}$.
4) Finally, for each e in $\mathcal{E}_K^D$, we set

$$\boldsymbol{\varphi}_e = \begin{cases} \mathcal{L}_{e,K}\big(u_h^n - u_{Dh}^n\big) & \text{on } K, \\ 0 & \text{on } \Omega \setminus K. \end{cases}$$

Exactly the same arguments as previously, combined with a separate treatment of the terms $u_h^n - u_{Dh}^n$ and $u_D^n - u_{Dh}^n$ yields the bound for $\tau_n^{\frac{1}{2}} h_e^{-\frac{1}{2}}\, \|u_h^n - u_{Dh}^n\|_{L^2(e)}$.
This concludes the proof.

4.5. Conclusions.

Let us introduce the full error

$$\begin{aligned} E = \sup_{0\leq t\leq T} \|(u-u_\tau)(\cdot,t)\|_{L^2(\Omega)} + \alpha^{\frac{1}{2}} \|\partial_t(u-u_\tau)\|_{L^2(0,T;L^2(\Omega))} \\ + \|\boldsymbol{q} - \pi_\tau^+ \boldsymbol{q}\|_{L^2(0,T;H(\mathrm{div},\Omega))} \\ + \sup_{0\leq t\leq T} \|(u_\tau - u_{h\tau})(\cdot,t)\|_{L^2(\Omega)} + \alpha^{\frac{1}{2}} \|\partial_t(u_\tau - u_{h\tau})\|_{L^2(0,T;L^2(\Omega))} \\ + \|\pi_\tau^+(\boldsymbol{q} - \boldsymbol{q}_h)\|_{L^2(0,T;H(\mathrm{div},\Omega))}, \end{aligned} \tag{4.24}$$

and also the terms depending on the data

$$\begin{aligned} \varepsilon^{(d)} = \|u_0 - u_{0h}\|_{L^2(\Omega)} + \|f - \pi_\tau^+ f\|_{L^2(0,T;H_{00}^{\frac{1}{2}}(\Gamma_F)')} \\ + \Big(\sum_{n=1}^N \tau_n \big(\|f(\cdot,t_n) - f_h^n\|^2_{H_{00}^{\frac{1}{2}}(\Gamma_F)'} + \|u_D(\cdot,t_n) - u_{Dh}^n\|^2_{H^{\frac{1}{2}}(\Gamma_D)}\big)\Big)^{\frac{1}{2}}, \end{aligned} \tag{4.25}$$

and finally the terms due to the time or space interpolation of the coefficients

$$\varepsilon^{(c)} = \Big(\sum_{n=1}^N \sum_{K\in\mathcal{T}_h^n} \big((\varepsilon_K^{n(\tau)})^2 + (\varepsilon_K^{n(h)})^2\big)\Big)^{\frac{1}{2}}. \tag{4.26}$$

Next, we make the following hypotheses:
(i) Assumption 4.1 holds;
(ii) The solution $(u, \boldsymbol{q})$ of problem (2.10) − (2.11) belongs to $\mathbb{Y} \times L^2(0,T;H(\mathrm{div},\Omega))$ and $D\mathcal{F}(U)$ is an isomorphism in $\mathscr{L}$;
(iii) The solution $(u_\tau, \pi_\tau^+ \boldsymbol{q})$ associated with (3.1) − (3.2) belongs to $\mathbb{Y} \times L^2(0,T;H(\mathrm{div},\Omega))$;
(iv) The quantity $|\tau|$ is smaller than τ_0 (see Lemma 4.8);
(v) The assumptions of Lemmas 4.8 and 4.9 hold.
Thus, the following equivalence property is satisfied

$$c\left(\Big(\sum_{n=1}^{N}\sum_{K\in\mathcal{T}_h^n}(\eta_K^{n(\tau)}+\eta_K^{n(h)})^2\Big)^{\frac{1}{2}}-\varepsilon^{(d)}-\varepsilon^{(c)}\right)\le E$$
$$\le c'\left(\Big(\sum_{n=1}^{N}\sum_{K\in\mathcal{T}_h^n}(\eta_K^{n(\tau)}+\eta_K^{n(h)})^2\Big)^{\frac{1}{2}}+\varepsilon^{(d)}+\varepsilon^{(c)}\right).$$

This result is fully optimal. Note also that a simple appropriate choice of the time steps and the meshes makes the quantity $\varepsilon^{(d)}$ negligible in comparison with the Hilbertian sum of the indicators. Finally, for reasons explained above, $\varepsilon^{(c)}$ is most often negligible.

Remark 4.12. The assumption that u and u_τ belong to $\mathbb{Y}$ is rather strong and could be weakened by using more technical arguments that we prefer to avoid here. Moreover it is not unlikely in all cases: For instance, if the domain Ω is two-dimensional and convex, for smooth enough data, the solution u belongs to $L^2(0,T;H^2(\Omega))$ and to $H^1(0,T;L^2(\Omega))$, hence to $\mathscr{C}^0(0,T;L^p(\Omega))$ for any p, $1 \le p < +\infty$; similar results can be obtained for $\partial_t u$ by differentiating equation (2.1) with respect to t and also for u_τ by simpler arguments.

Finally, estimate (4.22) is local in time and estimate (4.23) is local in space and time. Thus, the indicators $\eta_K^{n(\tau)}$ and $\eta_K^{n(h)}$ seem to be the right tools for an efficient adaptation strategy.

5. An adaptivity strategy and numerical experiments.

We first describe an adaptivity strategy which relies on the error indicators introduced in Section 4.2. Next, we check the efficiency of this strategy by presenting some numerical experiments. All of them have been performed on the finite element code FreeFem++, see [17].

5.1. An adaptivity strategy.

The strategy that we now propose is very similar to that in [4, §6], even if the problem that we consider is rather different. Let η^* be a fixed tolerance.

Initialization: We first choose an initial time step τ_0 such that

$$\|f - \pi_\tau^+ f\|_{L^2(t_0,t_1;H_{00}^{\frac{1}{2}}(\Gamma_F)')} \leq \eta^*. \tag{5.1}$$

Indeed this term appears in the definition (4.25) of $\varepsilon^{(d)}$. Next, we choose the triangulation $\mathcal{T}_h^0 = \mathcal{T}_h^1$ such that all other terms which appear in $\varepsilon^{(d)}$ and only depend on the data on the interval $[t_0, t_1]$ are small enough, which means:

$$\|u_0 - u_{0h}\|_{L^\infty(\Omega)} + \tau_1^{\frac{1}{2}} \left(\|f(\cdot,t_1) - f_h^1\|_{H_{00}^{\frac{1}{2}}(\Gamma_F)'} + \|u_D(\cdot,t_1) - u_{Dh}^1\|_{H^{\frac{1}{2}}(\Gamma_D)} \right) \leq \eta^*. \tag{5.2}$$

We then compute the solution $(u_h^1, \boldsymbol{q}_h^1)$ of problem (3.12) – (3.13).

Time adaptivity: Assuming that the time step τ_n, the mesh $\mathcal{T}_h^n$ and the discrete solution u_h^n are known, we first choose τ_{n+1} equal to τ_n and $\mathcal{T}_h^{n+1}$ equal to $\mathcal{T}_h^n$. We compute a first solution $(u_h^{n+1}, \boldsymbol{q}_h^{n+1})$ of problem (3.12) – (3.13), the corresponding error indicators $\eta_K^{n+1(\tau)}$ defined in (4.1) and their Hilbertian sum

$$\eta_h^{n+1(\tau)} = \Big(\sum_{K \in \mathcal{T}_h^{n+1}} (\eta_K^{n+1(\tau)})^2 \Big)^{\frac{1}{2}}. \tag{5.3}$$

Next,

- if $\eta_h^{n+1(\tau)}$ is smaller than η^*, we proceed to the spatial adaptivity step;
- if not, we divide τ_{n+1} by two (or by a constant times $\eta_h^{n+1(\tau)}/\eta^*$) and perform a new computation.

Of course, this step can be iterated a number of times. This leads to the final value of τ_{n+1}.

Space adaptivity: Assuming that the time step τ_{n+1} is known and that a first solution $(u_h^{n+1}, \boldsymbol{q}_h^{n+1})$ has been computed on the mesh $\mathcal{T}_h^{n+1}$, we compute the indicators $\eta_K^{n+1(h)}$,

$K \in \mathcal{T}_h^{n+1}$, and their mean value $\overline{\eta}^{n+1}$. Then, we perform mesh adaptivity in the usual way: The diameter of any element in the new triangulation which contains or is contained in an element K of $\mathcal{T}_h^{n+1}$ is proportional to the diameter of K times the ratio $\overline{\eta}^{n+1}/\eta_K^{n+1(h)}$. We refer to [10, Chap. 21] for the way of constructing such a mesh. This step can be iterated three or four times, and the final mesh is called $\mathcal{T}_h^{n+1}$.

Remark 5.1. At each step of time or space adaptivity, we must verify that the new terms which appear in $\varepsilon^{(d)}$, namely

$$\|f - \pi_\tau^+ f\|_{L^2(t_n,t_{n+1};H_{00}^{\frac12}(\Gamma_F)')} + \tau_{n+1}^{\frac12}\left(\|f(\cdot,t_{n+1}) - f_h^{n+1}\|_{H_{00}^{\frac12}(\Gamma_F)'} + \|u_D(\cdot,t_{n+1}) - u_{Dh}^{n+1}\|_{H^{\frac12}(\Gamma_D)}\right),$$

remain smaller than η^*. If it is not the case, a further adaptation is needed to handle these terms.

Remark 5.2. The a priori estimates (see [21, Section 4.2] for instance) indicate that, for a smooth solution $(u, \boldsymbol{q})$, the global error behaves like $c\,(\delta t + h)$. So no convergence can be hoped when performing only time adaptivity or only space adaptivity. On the other hand, when the τ_{n+1} resulting from time adaptivity is much smaller than τ_n, it could be reasonable to also replace the initial triangulation $\mathcal{T}_h^{n+1}$ by a new one which is uniformly refined from $\mathcal{T}_h^n$.

5.2. Validation of the discretization.

We work on the model domain

$$\Omega =]0,1[^2, \qquad \Gamma_F = \{1\}\times]0,1[, \qquad \Gamma_D = \partial\Omega \setminus \overline{\Gamma}_F, \qquad T = 5, \tag{5.4}$$

with the simple coefficients

$$\alpha = 0.01, \qquad b(x) = \frac{x^3}{3}, \qquad k = 0. \tag{5.5}$$

The data are given by

$$u_0(x,y) = 0, \qquad u_D(0,y) = \begin{cases} \exp(t)\, y^2(\frac12 - y)^2 & \text{if } 0 \le y \le \frac12, \\ 0 & \text{if } \frac12 \le y \le 1, \end{cases} \qquad u_D(x,0) = u_D(x,1) = 0, \qquad f(1,y) = 1. \tag{5.6}$$

We first compute a reference solution, denoted by $(u_{\rm ref}, \boldsymbol{q}_{\rm ref})$ for a very small time step $\tau = 10^{-3}$ and on a uniform very fine mesh made of 5932 triangles. Figure 1 presents the isovalues of the part $u_{\rm ref}$ of this solution at the final time T.

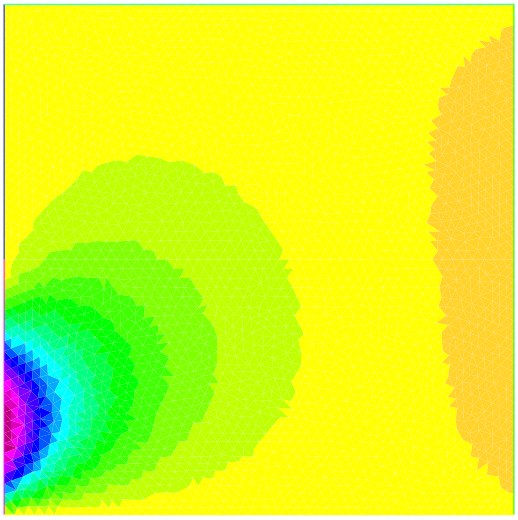

Figure 1. The isovalues of the reference solution u_{ref} at the final time

In a first step we do not perform adaptivity, i.e., we work with a fixed time step τ and all triangulations $\mathcal{T}_h^n$ equal to $\mathcal{T}_h$ which is uniform. In the following tables, we present the errors at the final time $T = t_N$ (indeed, it is too expensive to keep in memory the values of $(u_{\text{ref}}, \boldsymbol{q}_{\text{ref}})$ on the whole time interval), namely

$$\begin{aligned} E_u &= \|u_{\text{ref}}(\cdot, t_N) - u_h^N\|_{L^2(\Omega)}, \\ E_{\partial_t u} &= \tau^{-\frac{1}{2}} \|(u_{\text{ref}}(\cdot, t_N) - u_h^N) - (u_{\text{ref}}(\cdot, t_{N-1}) - u_h^{N-1})\|_{L^2(\Omega)}, \\ E_{\boldsymbol{q}} &= \tau^{\frac{1}{2}} \|\boldsymbol{q}_{\text{ref}}(\cdot, t_N) - \boldsymbol{q}_h^N\|_{H(\text{div},\Omega))}, \end{aligned} \tag{5.7}$$

and the Hilbertian sum of the indicators

$$\begin{aligned} \eta_1^{N(\tau)} &= \tau^{\frac{1}{2}} \|u_h^N - u_h^{N-1}\|_{L^2(\Omega)}, \\ \eta_2^{N(\tau)} &= \tau^{\frac{1}{2}} \|\frac{b(u_h^N) - b(u_h^{N-1})}{\tau} - B_h^N \frac{u_h^N - u_h^{N-1}}{\tau}\|_{L^2(\Omega)}, \\ \eta_1^{N(h)} &= \tau^{\frac{1}{2}} \|\alpha \frac{u_h^N - u_h^{N-1}}{\tau} + \frac{b(u_h^N) - b(u_h^{N-1})}{\tau} + \nabla \cdot \boldsymbol{q}_h^N\|_{L^2(\Omega)}, \\ \eta_2^{N(h)} &= \tau^{\frac{1}{2}} \|\boldsymbol{q}_h^N\|_{L^2(\Omega)^d}, \qquad \eta_3^{N(h)} = \tau^{\frac{1}{2}} \Big(\sum_{e \in \mathcal{E}_h^0} h_e^{-1} \|[u_h^N]_e\|_{L^2(e)}^2\Big)^{\frac{1}{2}}, \end{aligned} \tag{5.8}$$

first for τ fixed and h decreasing (Table 1), second for h fixed and τ decreasing (Table 2), third for τ et h decreasing simultaneously (Table 3).

h	τ	E_u	$E_{\partial_t u}$	E_q	$\eta_1^{N(\tau)}$	$\eta_2^{N(\tau)}$	$\eta_1^{N(h)}$	$\eta_2^{N(h)}$	$\eta_3^{N(h)}$
0.108	0.03125	0.02187	0.00427	0.28761	0.00197	0.08108	0.04686	0.55912	0.27086
0.05050	0.03125	0.01966	0.00450	0.15664	0.00206	0.05462	0.04801	0.59864	0.44028
0.03285	0.03125	0.01919	0.00456	0.11553	0.00207	0.04386	0.04039	0.60589	0.57340
0.02533	0.03125	0.01893	0.00458	0.09180	0.00208	0.03282	0.02777	0.61066	0.65367

Table 1. The errors and indicators for uniformly refined meshes

h	τ	E_u	$E_{\partial_t u}$	E_q	$\eta_1^{N(\tau)}$	$\eta_2^{N(\tau)}$	$\eta_1^{N(h)}$	$\eta_2^{N(h)}$	$\eta_3^{N(h)}$
0.03713	0.125	0.05053	0.01286	0.30305	0.01738	0.11002	0.09546	1.32405	2.18646
0.03713	0.0625	0.02700	0.00804	0.18353	0.00595	0.07048	0.06119	0.87980	1.02748
0.03713	0.03125	0.01790	0.00454	0.12406	0.00207	0.04750	0.04124	0.60308	0.49806
0.03713	0.015625	0.01516	0.00170	0.08666	0.00072	0.03280	0.02848	0.41989	0.24520

Table 2. The errors and indicators for decreasing time steps

h	τ	E_u	$E_{\partial_t u}$	E_q	$\eta_1^{N(\tau)}$	$\eta_2^{N(\tau)}$	$\eta_1^{N(h)}$	$\eta_2^{N(h)}$	$\eta_3^{N(h)}$
0.108	0.125	0.04196	0.01223	0.58873	0.01656	0.18834	0.10855	1.2274	1.18867
0.05050	0.0625	0.02723	0.00800	0.22919	0.00592	0.08103	0.07125	0.87333	0.90828
0.03285	0.03125	0.01919	0.00456	0.11553	0.00207	0.04386	0.04039	0.60589	0.57340
0.02533	0.015625	0.01540	0.00173	0.06347	0.00073	0.02266	0.01917	0.42514	0.32181

Table 3. The errors and indicators for uniformly refined meshes and decreasing time steps

From these tables, the convergence of the discretization in this situation is undeniable. It is rather slow, which seems correct for a low order discretization of a nonlinear problem.

Remark 5.3. Since k is equal to zero, the only nonlinear term involves the time derivative. It also follows from the previous calculation that the time error is now the leading term in the error. As a consequence, refining the mesh does not improve the convergence. On the other hand, since k is equal to zero and the discrete solution u_h^n is piecewise constant,

the indicator $\eta_2^{N(h)}$ involves the norm of $\boldsymbol{q}_h^N$, so that it decreases only when τ diminishes. A remedy to this consists in reinterpolating u_h^n in an enriched space, as first proposed for the Laplace equation in [30]. But, for simplicity, we prefer here to follow the approach in [5]: In this paper, it is proved that, still in the simplest case of the Laplace equation, this term can be omitted without destroying the optimality of the a posteriori estimates. So, from now on, we do not compute the indicator $\eta_2^{N(h)}$ when the function k is zero.

5.3. Validation of the adaptivity strategy.

To check the efficiency of our adaptivity strategy, we work with the domain and final time given by

$$\Omega =]0,1[^2, \qquad \Gamma_F = \{1\}\times]0,1[, \qquad \Gamma_D = \partial\Omega \setminus \overline{\Gamma}_F, \qquad T = 1, \tag{5.9}$$

and the coefficients given in (5.5), but now for the exact solution

$$u_{\text{ex}}(x,y) = \sin(\pi x)\sin(\pi y)\sin(10\pi x t). \tag{5.10}$$

Indeed, u_{ex} now satisfies problem (2.1) with a non-zero datum g in the right-hand side of the first equation. Of course, this induces a slight modification in the definition of the $\eta_K^{n(h)}$, but the estimates established in Section 4 remain valid in this case (with some further terms involving the function g and its approximation).

In the following Table 4, we present for some iterations n the time t_n, the final adapted time step τ_n, and the number of vertices N_h^n of the final adapted mesh $\mathcal{T}_h^n$. It can be observed that, even if in our adaptivity strategy we have decided not to increase τ_n, the final result is reasonable.

n	1	8	13	26	42
t_n	0.1	0.7	0.825	0.9	1
τ_n	0.1	0.0125	0.00625	0.00625	0.00625
N_h^n	49	236	554	1177	1923

Table 4. The parameters issued from adaptivity

Figures 2 and 3 present the adapted meshes at time $T/2 = 0.5$ and at time $T = 1$, respectively. They fit very well the increasing oscillations of the solution.

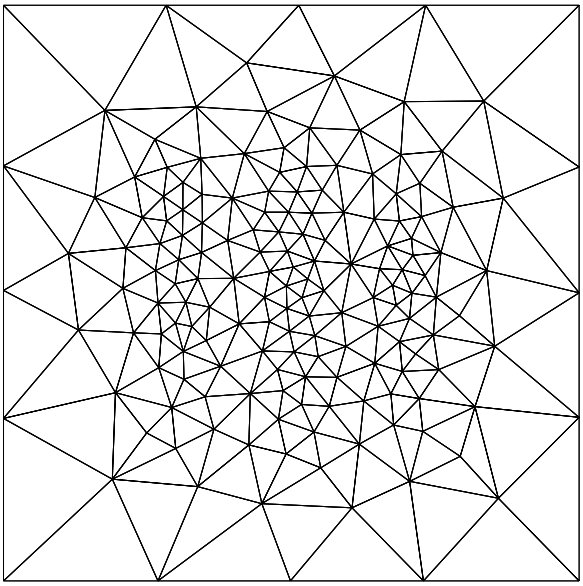

Figure 2. The final adapted mesh at time $T/2$

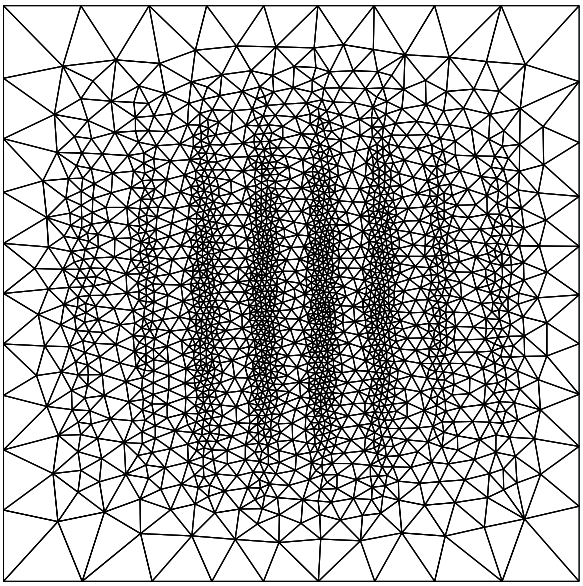

Figure 3. The final adapted mesh at time T

5.4. A more realistic experiment.

We now work in the framework proposed in [6] (see also [16] for a very similar model), where a sand ground is modelized. Numerical simulations of the flow have been performed in [7] and [8], relying on finite element and finite volume discretizations, respectively.

The domain Ω here is a rectangle:

$$\Omega =]0,100[\times]0,40[, \qquad \Gamma_D =]0,100[\times\{0,40\}, \qquad \Gamma_F = \{0,100\}\times]0,40[, \qquad T = 1. \tag{5.11}$$

The coefficients of system (1.1) are defined by

$$\Theta(h) = \begin{cases} \frac{\beta(\beta_s-\beta_r)}{\beta+|100\,h|^\delta} + \beta_r & \text{if } h < 0, \\ \beta_s & \text{if } h \geq 0, \end{cases}, \qquad K_w\big(\Theta(h)\big) = \begin{cases} K_s \frac{A}{A+|100\,h|^\xi} & \text{if } h < 0, \\ K_s & \text{if } h \geq 0, \end{cases} \tag{5.12}$$

with constants given by

$$\beta = 0.075, \qquad \beta_s = 0.287, \qquad \beta_r = 0.075, \qquad \delta = 3.96, \qquad K_s = 0.00944, \qquad A = 1.175 \times 10^6, \qquad \xi = 4.74. \tag{5.13}$$

From these equations, the coefficients b and $k \circ b$ of system (2.1) are easily recovered from the Kirchoff's change of unknowns, see Remark 2.1. The parameter α is arbitrarily chosen equal to 0.01.

The boundary and initial data are specified on the unknown h_w. They read

$$h_w(x,0;t) = -61.5, \qquad h_w(x,40;t) = -20.7, \qquad \big(K_w\big(\Theta(h_w)\big)\frac{\partial h_w}{\partial n}\big)(0,z;t) = \big(K_w\big(\Theta(h_w)\big)\frac{\partial h_w}{\partial n}\big)(100,z;t) = 0, \qquad h_w(x,z;0) = -61.5. \tag{5.14}$$

There also, the data u_D, f and u_0 can easily be recovered from that. Moreover, it can be noted that the solution u, hence h_w, are independent of x, so that the problem is in fact one-dimensional.

Figure 4 (left part) presents the curves of the values of the solution u as a function of z, $z \in]0,40[$ at different times $t = 0.05$, $t = 0.1$, $t = 0.3$, $t = 0.67$. The values of the physical unknown h_w at the same times are given in Figure 4 (right part).

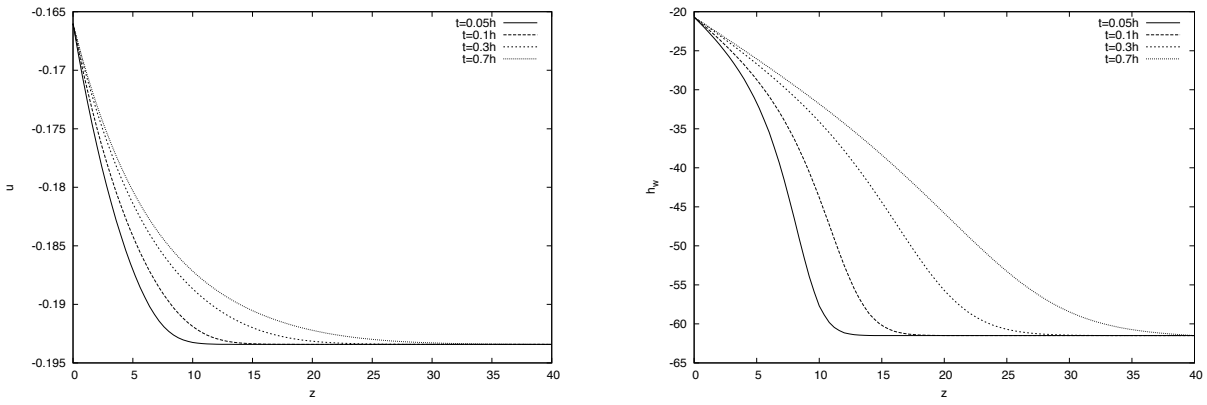

Figure 4. The solutions u and h_w

These figures are very coherent with the curves in [6]. Moreover, we do think that our adaptation process improves the efficiency of the computation.

References

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

[2] J. Bear — *Dynamics of Fluids in Porous Media*, Elsevier (1972).

[3] A. Bergam, C. Bernardi, Z. Mghazli — A posteriori analysis of the finite element discretization of some parabolic equations, *Math. Comput.* **74** (2005), 1117–1138.

[4] C. Bernardi, E. Süli — Time and space adaptivity for the second–order wave equation, *Math. Models Methods Appl. Sci.* **15** (2005), 199–225.

[5] C. Carstensen, R. Verfürth — Edge residuals dominate a posteriori error estimates for low order finite element methods, *SIAM J. Numer. Anal.* **36** (1999), 1571–1587.

[6] M.A. Celia, E.T. Bouloutas, R.L. Zarba — A general mass-conservative numerical solution for the unsaturated flow in porous media, *Water Resour. Res.* **26** (1990), 1483–1496.

[7] L.M. Chounet, D. Hilhorst, C. Jouron, Y. Kelanemer, P. Nicolas — Saturated–unsaturated simulation for coupled heat and mass transfer in the ground by means of a mixed finite element method, *Adv. Water Resources* **22** (1999), 445–460.

[8] R. Eymard, M. Gutnic, D. Hilhorst — The finite volume method for Richards equation, *Computational Geosciences* **3** (1999), 259–294.

[9] P. Fabrié, T. Gallouët — Modelling wells in porous media flows, *Math. Models Methods Appl. Sci.* **10** (2000), 673–709.

[10] P.J. Frey, P.-L. George — *Maillages, applications aux éléments finis*, Hermès (1999).

[11] M. Gabbouhy — Analyse mathématique et simulation numérique des phénomènes d'écoulement et de transport en milieux poreux non saturés. Application à la région du Gharb, Ph.D. Thesis, Université Ibn Tofail, Kénitra, Maroc (2000).

[12] M. Gabbouhy, Z. Mghazli — Un résultat d'existence de solutions faibles du problème d'écoulement non saturé modélisé par un système parabolique-elliptique non linéaire doublement dégénéré. *C. R. Acad. Sci. Paris Sér. I Math.* **330** (2000), 403–408.

[13] S.M.F. Garcia — Improved error estimates for mixed finite-element approximations for nonlinear parabolic equations: the continuous-time case, *Numer. Methods Partial Differential Equations* **10** (1994), 129–147.

[14] S.M.F. Garcia — Improved error estimates for mixed finite-element approximations for nonlinear parabolic equations: the discrete-time case, *Numer. Methods Partial Differential Equations* **10** (1994), 149–169.

[15] V. Girault, P.-A. Raviart — *Finite Element Methods for Navier–Stokes Equations, Theory and Algorithms*, Springer–Verlag (1986).

[16] R. Haverkampf, M. Vauclin, M. Touma, P.J. Wierenga, G. Vachaud — A comparison of numerical simulation models for one-dimensional infiltration, *Soil Sci. Soc. Am. J.* **41** (1977), 285–294.

[17] F. Hecht — Freefem++, Third Edition, Version 3.9-1, Université Pierre et Marie Curie, Paris

(2010), on the web at http://www.freefem.org/ff++/ftp/freefem++doc.pdf.

[18] J.-L. Lions — *Quelques méthodes de résolution des problèmes aux limites non linéaires*, Dunod & Gauthier-Villars (1969).

[19] J.-L. Lions, E. Magenes — *Problèmes aux limites non homogènes et applications*, Vol. I, Dunod, Paris (1968).

[20] J. Pousin, J. Rappaz — Consistency, stability, a priori and a posteriori errors for Petrov-Galerkin methods applied to nonlinear problems, *Numer. Math.* **69** (1994), 213–231.

[21] F. Radu, I.S. Pop, P. Knabner — Order of convergence estimates for an Euler implicit, mixed finite element discretization of Richards' equation, *SIAM J. Numer. Anal.* **42** (2004), 1452–1478.

[22] K.R. Rajagopal — On a hierarchy of approximate models for flows of incompressible fluids through porous solid, *Math. Models Methods Appl. Sci.* **17** (2007), 215–252.

[23] P.-A. Raviart, J.-M. Thomas — A mixed finite element method for second order elliptic problems, *Mathematical Aspects of Finite Element Methods*, Lecture Notes in Mathematics **606**, Springer (1977), 292–315.

[24] L.A. Richards — Capillary conduction of liquids through porous mediums. *Physics* **1** (1931), 318–333.

[25] C.A. San Soucie — Mixed finite element methods for variably saturated subsurface flow, Ph.D. Thesis, Rice University, United States (1996).

[26] E. Schneid, P. Knabner, F. Radu — A priori error estimates for a mixed finite element discretization of the Richards' equation, *Numer. Math.* **98** (2004), 353–370.

[27] P. Sochala, A. Ern, S. Piperno — Mass conservative BDF-discontinuous Galerkin/explicit finite volume schemes for coupling subsurface and overland flows, *Comput. Methods Appl. Mech. Engrg.* **198** (2009), 2122–2136.

[28] P. Sochala, A. Ern, S. Piperno — Numerical methods for subsurface flows and coupling with surface runoff, in preparation.

[29] R. Verfürth — *A Review of A Posteriori Error Estimation and Adaptive Mesh-Refinement Techniques*, Wiley & Teubner (1996).

[30] M. Vohralík — Guaranteed and fully robust a posteriori error estimates for conforming discretizations of diffusion problems with discontinuous coefficients, *J. Sci. Comput.* **46** (2011), 397–438.

[31] C.S. Woodward, C.N. Dawson — Analysis of expanded mixed finite element methods for a nonlinear parabolic equation modeling flow into variably saturated porous media, *SIAM J. Numer. Anal.* **37** (2000), 701–724.