Chin. Ann. Math.
34B(2), 2013, 193–212
DOI: 10.1007/s11401-013-0766-z

Chinese Annals of Mathematics, Series B

The Rain on Underground Porous Media Part I: Analysis of a Richards Model

Christine BERNARDI[1] Adel BLOUZA[2] Linda EL ALAOUI[3]

(*In honor of the scientific heritage of Jacques-Louis Lions*)

Abstract The Richards equation models the water flow in a partially saturated underground porous medium under the surface. When it rains on the surface, boundary conditions of Signorini type must be considered on this part of the boundary. The authors first study this problem which results into a variational inequality and then propose a discretization by an implicit Euler's scheme in time and finite elements in space. The convergence of this discretization leads to the well-posedness of the problem.

Keywords Richards equation, Porous media, Euler's implicit scheme, Finite element discretization, Parabolic variational inequality

2000 MR Subject Classification 76S05, 76M10, 65M12

1 Introduction

The following equation:

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

models the flow of a wetting fluid, mainly water, in the underground surface, hence in an unsaturated medium (see [15] for the introduction of this type of models). In opposite to Darcy's or Brinkman's systems (see [14] for all these models), this equation, which is derived by combining Darcy's generalized equation with the mass conservation law, 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. The unknown ψ is the difference between the pressure of water and the atmospherical pressure.

This equation is usually provided with Dirichlet or Neumann type boundary conditions. Indeed, Neumann boundary conditions on the underground part of the boundary are linked to the draining of water outside of the domain, and Dirichlet boundary conditions on the surface are introduced to take into account the rain. However, when the porous media can no longer absorb

Manuscript received August 25, 2012.
[1]Laboratoire Jacques-Louis Lions, CNRS & Université Pierre et Marie Curie, BC 187, 4 Place Jussieu, 75252 Paris Cedex 05, France. E-mail: bernardi@ann.jussieu.fr
[2]Laboratoire de Mathématiques Raphaël Salem (UMR 6085 CNRS), Université de Rouen, Avenue de l'Université, BP 12, 76801 Saint-Étienne-du-Rouvray, France. E-mail: Adel.Blouza@univ-rouen.fr
[3]Universit Paris 13, Sorbonne Paris Cit, LAGA, CNRS (UMR 7539), F-93430, Villetaneuse, France. E-mail: elalaoui@math.univ-paris13.fr

the rainwater that falls, the upper surface of the domain allows to exfiltration and infiltration. In other words, the upper surface is divided into a saturated zone and an unsaturated zone. We assume that the re-infiltration process is negligible. This leads to variational inequalities of the following type:

$$-\psi \geq 0, \quad \boldsymbol{v}(\psi) \cdot \boldsymbol{n} \geq \boldsymbol{v}_r \cdot \boldsymbol{n}, \quad \psi(\boldsymbol{v}(\psi) \cdot \boldsymbol{n} - \boldsymbol{v}_r \cdot \boldsymbol{n}) = 0, \tag{1.2}$$

where $\boldsymbol{v}(\psi)$ is the flux

$$\boldsymbol{v}(\psi) = -K_w(\Theta(\psi))\nabla(\psi + z), \tag{1.3}$$

and $\boldsymbol{n}$ stands for the unit outward normal vector to the surface, and $\boldsymbol{v}_r$ stands for a given rain fall rate. We refer to the thesis of Berninger [4] for the full derivation of this model from hydrology laws and more specifically to [4, Section 1.5] for the derivation of the boundary inequalities (1.2).

It is not so easy to give a mathematical sense to the system (1.1)–(1.2). As a standard, the key argument for the analysis of the problem (1.1) is to use Kirchhoff's change of unknowns. Indeed, after this transformation, the new equation fits the general framework proposed in [1] (see also [6] for the analysis of a different model). Thus, the existence and uniqueness of a solution to this equation with appropriate linear initial and boundary conditions can be derived from standard arguments. In order to handle the inequality in (1.2), we again use a variational formulation. We refer to [2] for the first analysis of very similar systems (see also [5]). We prove that the problem (1.1)–(1.2) is well-posed when the data are smooth enough but in the first step with a rather restrictive assumption on the coefficients.

The discretization of the problem (1.1) was proposed and/or studied in many papers with standard boundary conditions (see [3, 7, 13, 16, 18, 19] and [17] for a more general equation). However, it does not seem to be treated for the case of the boundary inequality (1.2). We propose here a discretization of system (1.1)–(1.2), in two steps as follows:

(i) We first use the Euler's implicit scheme to build a time semi-discrete problem, where one of the nonlinear terms is treated in an explicit way for simplicity.

(ii) We then construct a fully discrete problem that relies on the Galerkin method and finite elements in the spatial domain.

In both cases, we prove that the corresponding variational problem is well-posed.

To conclude, we prove that the solution to this discrete problem converges to a solution to the continuous one when the discretization parameters tend to zero. This ends the proof of our existence result, since no restrictive condition is needed here.

The outline of the paper is as follows.

In Section 2, we present the variational formulation of the full system, and investigate its well-posedness in appropriate Sobolev spaces.

Section 3 is devoted to the descriptions of the time semi-discrete problem and of the fully discrete problem. We check their well-posedness.

In Section 4, we investigate the convergence of the solution of the discrete problem to a solution of the continuous one.

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 three parts Γ_B, Γ_F and Γ_G (these indices mean "bottom", "flux" and "ground", respectively), and that Γ_B has a positive measure. Let also T be a positive real number.

In order to perform the Kirchhoff's change of unknowns in the problem (1.1), we observe that, since the conductivity coefficient K_w is positive, the mapping

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

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

$$u = \mathcal{K}(\psi), \quad b(u) = \Theta \circ \mathcal{K}^{-1}(u), \quad k(\cdot) = K_w(\cdot),$$

and thanks to an appropriate choice of the function $\widetilde{\Theta}$, we derive the equation (more details are given in [3, Remark 2.1] for instance)

$$\alpha\, \partial_t u + \partial_t b(u) - \nabla \cdot (\nabla u + k \circ b(u)\boldsymbol{e}_z) = 0 \quad \text{in } \Omega \times [0, T],$$

where $-\boldsymbol{e}_z$ stands for the unit vector in the direction of gravity. Moreover, the Kirchhoff's change of unknowns has the further property of preserving the positivity: u is positive if and only if ψ is positive; u is negative if and only if ψ is negative. So, writing the inequality (1.2) in terms of the unknown u is easy.

As a consequence, from now on, we work with the following system:

$$\begin{cases} \alpha\, \partial_t u + \partial_t b(u) - \nabla \cdot (\nabla u + k \circ b(u)\boldsymbol{e}_z) = 0 & \text{in } \Omega \times [0, T], \\ u = u_B & \text{on } \Gamma_B \times [0, T], \\ -(\nabla u + k \circ b(u)\boldsymbol{e}_z) \cdot \boldsymbol{n} = f_F & \text{on } \Gamma_F \times [0, T], \\ u \leq 0, \quad -(\nabla u + k \circ b(u)\boldsymbol{e}_z) \cdot \boldsymbol{n} \geq \boldsymbol{q}_r \cdot \boldsymbol{n}, & \\ u\, (\nabla u + k \circ b(u)\boldsymbol{e}_z + \boldsymbol{q}_r) \cdot \boldsymbol{n} = 0 & \text{on } \Gamma_G \times [0, T], \\ u|_{t=0} = u_0 & \text{in } \Omega. \end{cases} \tag{2.1}$$

The unknown is now the quantity u. The data are the Dirichlet boundary condition u_B on $\Gamma_B \times [0, T]$ and the initial condition u_0 on Ω, together with the boundary conditions f_F and $\boldsymbol{q}_r$ on the normal component of the flux, where f_F corresponds to the draining of water, and $\boldsymbol{q}_r$ corresponds to the rain. Finally, b and k are supposed to be known, while α is a positive constant. From now on, we assume that

(i) the function b is of class $\mathscr{C}^2$ on $\mathbb{R}$, with bounded and Lipschitz-continuous derivatives, and is nondecreasing,

(ii) the function $k \circ b$ is continuous, bounded, and uniformly Lipschitz-continuous on $\mathbb{R}$.

Remark 2.1 It must be noted that the parameter α has a physical meaning. Indeed, the function $\widetilde{\Theta}$ in (1.1) is usually the sum of Θ and a term linked to the saturation state. But it can also be considered as a regularization parameter, since it avoids the degeneracy of the equation, where the derivative of b vanishes. So, adding the term $\alpha\, \partial_t u$ is a standard technique in the analysis of such problems, which has been used with success for constructing effective numerical algorithms (see e.g., [12–13]).

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 the seminorm $|\cdot|_{W^{m,p}(\Omega)}$, with the usual notation $H^m(\Omega)$ when $p = 2$. As a standard, the range of $H^1(\Omega)$ by the trace operator on any part Γ of $\partial\Omega$ is denoted by $H^{\frac{1}{2}}(\Gamma)$. 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 on $[0,T]$ 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: $v \mapsto \|\partial_t^\ell v\|_E$, $0 \leq \ell \leq m$, are square-integrable on $]0,T[$.

To write a variational formulation for the problem, we introduce the time-dependent subset

$$\mathbb{V}(t) = \{v \in H^1(\Omega);\ v|_{\Gamma_B} = u_B(\cdot,t) \text{ and } v|_{\Gamma_G} \leq 0\}. \tag{2.2}$$

It is readily checked that each $\mathbb{V}(t)$ is closed and convex (see [4, Proposition 1.5.5]), when u_B belongs to $\mathscr{C}^0(0,T;H^{\frac{1}{2}}(\Gamma_B))$. Thus, we are led to consider the following variational problem (with obvious notation for $L^2(0,T;\mathbb{V})$).

Find u in $L^2(0,T;\mathbb{V})$ with $\partial_t u$ in $L^2(0,T;L^2(\Omega))$, such that

$$u|_{t=0} = u_0, \tag{2.3}$$

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

$$\begin{aligned}\forall v \in \mathbb{V}(t), \quad & \alpha \int_\Omega (\partial_t u)(\boldsymbol{x},t)(v-u)(\boldsymbol{x},t)\mathrm{d}\boldsymbol{x} + \int_\Omega (\partial_t b(u))(\boldsymbol{x},t)(v-u)(\boldsymbol{x},t)\mathrm{d}\boldsymbol{x} \\ & + \int_\Omega (\nabla u + k \circ b(u)\boldsymbol{e}_z)(\boldsymbol{x},t)\cdot(\nabla(v-u))(\boldsymbol{x},t)\mathrm{d}\boldsymbol{x} \\ & \geq -\int_{\Gamma_F} f_F(\tau,t)(v-u)(\tau,t)\mathrm{d}\tau - \int_{\Gamma_G} (\boldsymbol{q}_r\cdot\boldsymbol{n})(\tau,t)(v-u)(\tau,t)\mathrm{d}\tau,\end{aligned} \tag{2.4}$$

where τ denotes the tangential coordinates on $\partial\Omega$. The reason for this follows.

Proposition 2.1 *The problems* (2.1) *and* (2.3)–(2.4) *are equivalent, and more precisely:*

(i) *Any solution to the problem* (2.1) *in* $L^2(0,T;H^1(\Omega)) \cap H^1(0,T;L^2(\Omega))$ *is a solution to* (2.3)–(2.4).

(ii) *Any solution to the problem* (2.3)–(2.4) *is a solution to the problem* (2.1) *in the distribution sense.*

Proof We check successively the two assertions of the proposition.

(1) Let u be any solution to (2.1) in $L^2(0,T;H^1(\Omega))\cap H^1(0,T;L^2(\Omega))$. Obviously, it belongs to $L^2(0,T;\mathbb{V})$ and satisfies (2.3). Next, we observe that, for any v in $\mathbb{V}(t)$, the function $v-u$ vanishes on Γ_B. Multiplying the first line in (2.1) by this function and integrating it by parts on Ω, we have

$$\begin{aligned}& \alpha \int_\Omega (\partial_t u)(\boldsymbol{x},t)(v-u)(\boldsymbol{x},t)\mathrm{d}\boldsymbol{x} + \int_\Omega (\partial_t b(u))(\boldsymbol{x},t)(v-u)(\boldsymbol{x},t)\mathrm{d}\boldsymbol{x} \\ & + \int_\Omega (\nabla u + k\circ b(u)\boldsymbol{e}_z)(\boldsymbol{x},t)\cdot(\nabla(v-u))(\boldsymbol{x},t)\mathrm{d}\boldsymbol{x} \\ = & \int_{\Gamma_F\cup\Gamma_G} (\nabla u + k\circ b(u)\boldsymbol{e}_z)\cdot\boldsymbol{n}(\tau)(v-u)(\tau,t)\mathrm{d}\tau.\end{aligned}$$

To conclude, we observe on Γ_G, either u is zero and $\nabla u + k \circ b(u)\boldsymbol{e}_z$ is smaller than $-q_\tau \cdot \boldsymbol{n}$, or u is not zero and $\nabla u + k \circ b(u)\boldsymbol{e}_z$ is equal to $-\boldsymbol{q}_r \cdot \boldsymbol{n}$. All these yield (2.4).

(2) Conversely, let u be any solution to (2.3)–(2.4).

(i) By noting that for any function w in $\mathscr{D}(\Omega)$, $(u+w)(\cdot,t)$ belongs to $\mathbb{V}(t)$. Taking v equal to $u \pm w$ in (2.4), we obtain the first line of (2.1) in the distribution sense.

(ii) The second line in (2.1) follows from the definition of $\mathbb{V}(t)$.

(iii) By taking v equal to $u \pm w$ for any w in $\mathscr{D}(\Omega \cup \Gamma_F)$, we also derive the third line in (2.1).

(iv) The fact that u is nonpositive on Γ_G, comes from the definition of $\mathbb{V}(t)$. On the other hand, the previous equations imply that for any v in $\mathbb{V}(t)$,

$$\int_{\Gamma_G} (\nabla u + k \circ b(u)\boldsymbol{e}_z) \cdot \boldsymbol{n}(\tau)(v-u)(\tau,t)\mathrm{d}\tau \geq -\int_{\Gamma_G} (\boldsymbol{q}_r \cdot \boldsymbol{n})(\tau,t)(v-u)(\tau,t)\mathrm{d}\tau.$$

Taking v equal to $u+w$, where w vanishes on Γ_B and is nonpositive on Γ_G, yields that $-(\nabla u + k \circ b(u)\boldsymbol{e}_z) \cdot \boldsymbol{n}$ is larger than $\boldsymbol{q}_r \cdot \boldsymbol{n}$. Finally, taking v equal to zero on Γ_G, leads to

$$\int_{\Gamma_G} (\nabla u + k \circ b(u)\boldsymbol{e}_z + \boldsymbol{q}_r) \cdot \boldsymbol{n}(\tau)u(\tau,t)\mathrm{d}\tau \leq 0.$$

Since the two quantities u and $(\nabla u + k \circ b(u)\boldsymbol{e}_z + \boldsymbol{q}_r)$ are nonpositive on Γ_G, their product is zero.

(v) Finally the last line of (2.1) is written in (2.3).

Proving that the problem (2.3)–(2.4) is well-posed and is not at all obvious. We begin with the simpler result, i.e., the uniqueness of the solution. For brevity, we set

$$\mathbb{X} = L^2(0,T;\mathbb{V}) \cap H^1(0,T;L^2(\Omega)). \tag{2.5}$$

We also refer to [11, Chapitre 1, Théorèrne 11.7] for the definition of the space $H_{00}^{\frac{1}{2}}(\Gamma_B)$.

Proposition 2.2 *For any data u_B, f_F, $\boldsymbol{q}_r$ and u_0 satisfying*

$$\begin{aligned} &u_B \in H^1(0,T;H_{00}^{\frac{1}{2}}(\Gamma_B)), \quad f_F \in L^2(0,T;L^2(\Gamma_F)), \\ &\boldsymbol{q}_r \in L^2(0,T;L^2(\Gamma_G)^d), \qquad u_0 \in H^1(\Omega), \end{aligned} \tag{2.6}$$

the problem (2.3)–(2.4) *has at most a solution in* $\mathbb{X}$.

Proof Let u_1 and u_2 be two solutions to the problem (2.3)–(2.4). Thus, the function $u = u_1 - u_2$ vanishes on Γ_B and at $t=0$. Taking v equal to u_2 in the problem satisfied by u_1 and equal to u_1 in the problem satisfied by u_2, and subtracting the second problem from the first one, we obtain

$$\begin{aligned} &\alpha \int_\Omega (\partial_t u)(\boldsymbol{x},t)u(\boldsymbol{x},t)\mathrm{d}\boldsymbol{x} + \int_\Omega (\partial_t b(u_1) - \partial_t b(u_2))(\boldsymbol{x},t)u(\boldsymbol{x},t)\mathrm{d}\boldsymbol{x} \\ &+ \int_\Omega (\nabla u)^2(\boldsymbol{x},t)\mathrm{d}\boldsymbol{x} + \int_\Omega (k \circ b(u_1) - k \circ b(u_2))(\boldsymbol{x},t)\boldsymbol{e}_z \cdot (\nabla u))(\boldsymbol{x},t)\mathrm{d}\boldsymbol{x} \ \leq 0. \end{aligned} \tag{2.7}$$

We integrate this inequality with respect to t and evaluate successively the four integrals.

(1) The first and third ones are obvious

$$\alpha \int_0^t \int_\Omega (\partial_t u)(\boldsymbol{x},s) u(\boldsymbol{x},s) \mathrm{d}\boldsymbol{x}\mathrm{d}s + \int_0^t \int_\Omega (\nabla u)^2(\boldsymbol{x},s) \mathrm{d}\boldsymbol{x}\mathrm{d}s$$
$$= \frac{\alpha}{2} \|u(\cdot,t)\|^2_{L^2(\Omega)} + \int_0^t |u(\cdot,s)|^2_{H^1(\Omega)} \mathrm{d}s.$$

(2) To evaluate the second one, we use the decomposition

$$\int_\Omega (\partial_t b(u_1) - \partial_t b(u_2))(\boldsymbol{x},t) u(\boldsymbol{x},t) \mathrm{d}\boldsymbol{x}$$
$$= \int_\Omega b'(u_1)(\boldsymbol{x},t)(\partial_t u)(\boldsymbol{x},t) u(\boldsymbol{x},t) \mathrm{d}\boldsymbol{x}$$
$$+ \int_\Omega (b'(u_1) - b'(u_2))(\boldsymbol{x},t)(\partial_t u_2)(\boldsymbol{x},t) u(\boldsymbol{x},t) \mathrm{d}\boldsymbol{x},$$

and integrate the first term by parts with respect to t, which gives

$$\int_0^t \int_\Omega (\partial_t b(u_1) - \partial_t b(u_2))(\boldsymbol{x},s) u(\boldsymbol{x},s) \mathrm{d}\boldsymbol{x}\mathrm{d}s$$
$$= \int_\Omega \frac{b'(u_1)(\boldsymbol{x},t)}{2} u^2(\boldsymbol{x},t) \mathrm{d}\boldsymbol{x}$$
$$- \frac{1}{2} \int_0^t \int_\Omega b''(u_1)(\boldsymbol{x},s)(\partial_t u_1)(\boldsymbol{x},s) u^2(\boldsymbol{x},s) \mathrm{d}\boldsymbol{x}\mathrm{d}s$$
$$+ \int_0^t \int_\Omega (b'(u_1) - b'(u_2))(\boldsymbol{x},s)(\partial_t u_2)(\boldsymbol{x},s) u(\boldsymbol{x},s) \mathrm{d}\boldsymbol{x}\mathrm{d}s.$$

Next, the nonnegativity of b', the boundedness of b'' and the Lipschitz-continuity of b' yield

$$\int_0^t \int_\Omega (\partial_t b(u_1) - \partial_t b(u_2))(\boldsymbol{x},s) u(\boldsymbol{x},s) \mathrm{d}\boldsymbol{x}\mathrm{d}s \geq -c(u_1,u_2) \int_0^t \|u(\cdot,s)\|^2_{L^4(\Omega)} \mathrm{d}s,$$

where $c(u_1,u_2) > 0$ depends on $\|\partial_t u_i\|_{L^2(0,T;L^2(\Omega))}$. Next, we use an interpolation inequality (see [11, Chapitre 1, Proposition 2.3]) and the Poincaré-Friedrichs inequality

$$\|u\|_{L^4(\Omega)} \leq \|u\|^{1-\frac{d}{4}}_{L^2(\Omega)} \left(c|u|_{H^1(\Omega)}\right)^{\frac{d}{4}} \leq c'\Big(1 - \frac{d}{4}\Big)\|u\|_{L^2(\Omega)} + \frac{d}{4}|u|_{H^1(\Omega)},$$

and conclude with a Young's inequality.

(3) Finally, to bound the last one, we combine the Lipschitz-continuity of $k \circ b$ together with a Young's inequality

$$\int_0^t \int_\Omega (k \circ b(u_1) - k \circ b(u_2))(\boldsymbol{x},s) \boldsymbol{e}_z \cdot (\nabla u)(\boldsymbol{x},s) \mathrm{d}\boldsymbol{x}\mathrm{d}s$$
$$\leq \frac{1}{4} \Big(\int_0^t |u(\cdot,s)|^2_{H^1(\Omega)} \mathrm{d}s \Big) + c\Big(\int_0^t \|u(\cdot,s)\|^2_{L^2(\Omega)} \mathrm{d}s \Big).$$

All these give

$$\frac{\alpha}{2} \|u(\cdot,t)\|^2_{L^2(\Omega)} + \frac{1}{2} \int_0^t |u(\cdot,s)|^2_{H^1(\Omega)} \mathrm{d}s \leq c(u_1,u_2) \int_0^t \|u(\cdot,s)\|^2_{L^2(\Omega)} \mathrm{d}s.$$

Thus, applying Grönwall's lemma yields that u is zero, whence the uniqueness result follows.

Proving the existence is much more complex. We begin with a basic result.

Lemma 2.1 *If the function u_B belongs to $\mathscr{C}^0(0,T;H_{00}^{\frac{1}{2}}(\Gamma_B))$, then for all t in $[0,T]$, the convex set $\mathbb{V}(t)$ is not empty.*

Proof Denoting by $\overline{u}_B(\cdot,t)$ the extension by zero of $u_B(\cdot,t)$ to $\partial\Omega$, we observe that any lifting of $\overline{u}_B(\cdot,t)$ in $H^1(\Omega)$ belongs to $\mathbb{V}(t)$, whence the desired result follows.

In the first step, we consider the linear problem, for any datum F in $L^2(0,T;L^2(\Omega))$.

Find u in $L^2(0,T;\mathbb{V})$ with $\partial_t u$ in $L^2(0,T;L^2(\Omega))$ satisfying (2.3) and such that, for a.e. t in $[0,T]$,

$$\begin{aligned}\forall v\in\mathbb{V}(t),\quad \alpha\int_\Omega(\partial_t u)(\boldsymbol{x},t)(v-u)(\boldsymbol{x},t)\mathrm{d}\boldsymbol{x}+\int_\Omega(\nabla u)(\boldsymbol{x},t)\cdot(\nabla(v-u))(\boldsymbol{x},t)\mathrm{d}\boldsymbol{x}\\ \geq-\int_\Omega F(\boldsymbol{x},t)(v-u)(\boldsymbol{x},t)\mathrm{d}\boldsymbol{x}-\int_{\Gamma_F}f_F(\tau,t)(v-u)(\tau,t)\mathrm{d}\tau\\ -\int_{\Gamma_G}(\boldsymbol{q}_r\cdot\boldsymbol{n})(\tau,t)(v-u)(\tau,t)\mathrm{d}\tau.\end{aligned}\tag{2.8}$$

However a weaker formulation of this problem can be derived by integrating with respect to t. It reads as follows.

Find u in $L^2(0,T;\mathbb{V})$ satisfying (2.3), such that

$$\begin{aligned}\forall v\in\mathbb{X},\quad \alpha\int_0^T\int_\Omega(\partial_t u)(\boldsymbol{x},t)(v-u)(\boldsymbol{x},t)\mathrm{d}\boldsymbol{x}\mathrm{d}t+\int_0^T\int_\Omega(\nabla u)(\boldsymbol{x},t)\cdot(\nabla(v-u))(\boldsymbol{x},t)\mathrm{d}\boldsymbol{x}\mathrm{d}t\\ \geq-\int_0^T\int_\Omega F(\boldsymbol{x},t)(v-u)(\boldsymbol{x},t)\mathrm{d}\boldsymbol{x}\mathrm{d}t-\int_0^T\int_{\Gamma_F}f_F(\tau,t)(v-u)(\tau,t)\mathrm{d}\tau\mathrm{d}t\\ -\int_0^T\int_{\Gamma_G}(\boldsymbol{q}_r\cdot\boldsymbol{n})(\tau,t)(v-u)(\tau,t)\mathrm{d}\tau\mathrm{d}t.\end{aligned}\tag{2.9}$$

We recall in the next lemma the properties of this problem which are standard.

Lemma 2.2 *Assume that the data u_B, f_F, $\boldsymbol{q}_r$ and u_0 satisfy* (2.6)*. Then, for any F in $L^2(0,T;L^2(\Omega))$, the problem* (2.3)–(2.9) *has a unique solution u in $L^2(0,T;\mathbb{V})$.*

Proof It follows from Lemma 2.1 and the further assumption on u_B that $\mathbb{X}$ is a non-empty closed convex set. We also consider a lifting $\overline{u}_B$ of the extension by zero of u_B to $\partial\Omega$ in $H^1(0,T;H^1(\Omega))$. Then, it is readily checked that $u-\overline{u}_B$ is the solution to a problem, which satisfies all the assumptions in [10, Chapitre 6, Theorem 2.2], whence the existence and uniqueness result follows.

Any solution to (2.3)–(2.8) is a solution to (2.3)–(2.9), but the converse property is not obvious in the general case (see [10, Chapitre 6]). However, in our specific case, it is readily checked by a density argument that (2.9) is satisfied for any v in $L^2(0,T;\mathbb{V})$, so that problems (2.3)–(2.8) and (2.3)–(2.9) are fully equivalent.

To go further, we assume that the following compatibility condition holds:

$$u_0(\boldsymbol{x})=u_B(\boldsymbol{x},0)\quad\text{for }\boldsymbol{x}\in\Gamma_B\text{ a.e.}\quad\text{and}\quad u_0(\boldsymbol{x})\leq 0\quad\text{for }\boldsymbol{x}\in\Gamma_G\text{ a.e.}\tag{2.10}$$

Moreover, we introduce a lifting u_B^* of an extension of u_B to $\partial\Omega$, which belongs to $H^1(0,T;\mathbb{V})$ and satisfies

$$u_B^*(\boldsymbol{x},0) = u_0(\boldsymbol{x}) \quad \text{for } x \in \Omega \text{ a.e.,} \tag{2.11}$$

together with the stability property

$$\|u_B^*\|_{H^1(0,T;H^1(\Omega))} \le c\,\|u_B\|_{H^1(0,T;H_{00}^{\frac{1}{2}}(\Gamma_B))}. \tag{2.12}$$

Then, it is readily checked that u is a solution to the problem (2.3)–(2.4) if and only if the function $u^* = u - u_B^*$ is a solution to the following problem.

Find u^* in $L^2(0,T;\mathbb{V}_0)$ with $\partial_t u^*$ in $L^2(0,T;L^2(\Omega))$, such that

$$u^*|_{t=0} = 0, \tag{2.13}$$

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

$$\begin{aligned}\forall v \in \mathbb{V}_0, \quad & \alpha \int_\Omega (\partial_t u^*)(\boldsymbol{x},t)(v-u^*)(\boldsymbol{x},t)\mathrm{d}\boldsymbol{x} + \int_\Omega (\partial_t b_*(u^*))(\boldsymbol{x},t)(v-u^*)(\boldsymbol{x},t)\mathrm{d}\boldsymbol{x} \\ & + \int_\Omega (\nabla u^* + k \circ b_*(u^*)\boldsymbol{e}_z)(\boldsymbol{x},t) \cdot (\nabla(v-u^*))(\boldsymbol{x},t)\mathrm{d}\boldsymbol{x} \\ & \ge - \int_\Omega F_B(\boldsymbol{x},t)(v-u^*)(\boldsymbol{x},t)\mathrm{d}\boldsymbol{x} - \int_{\Gamma_F} f_F(\tau,t)(v-u^*)(\tau,t)\mathrm{d}\tau \\ & - \int_{\Gamma_G} (\boldsymbol{q}_r\cdot\boldsymbol{n})(\tau,t)(v-u^*)(\tau,t)\mathrm{d}\tau \end{aligned} \tag{2.14}$$

with the definition of the subset $\mathbb{V}_0$,

$$\mathbb{V}_0 = \{v \in H^1(\Omega);\ v|_{\Gamma_B} = 0 \text{ and } v|_{\Gamma_G} \le 0\}, \tag{2.15}$$

where the new application b_* is defined by $b_*(u^*) = b(u^* + u_B^*)$. The datum F_B is defined by, for a.e. t in $]0,T[$,

$$\begin{aligned} & \int_\Omega F_B(\boldsymbol{x},t)v(\boldsymbol{x})\mathrm{d}\boldsymbol{x} \\ & = \alpha \int_\Omega (\partial_t u_B^*)(\boldsymbol{x},t)v(\boldsymbol{x})\mathrm{d}\boldsymbol{x} + \int_\Omega (\nabla u_B^*)(\boldsymbol{x},t) \cdot (\nabla v)(\boldsymbol{x})\mathrm{d}\boldsymbol{x}, \end{aligned} \tag{2.16}$$

and clearly belongs to $L^2(0,T;\mathbb{W}')$, where $\mathbb{W}$ is the smallest linear space containing $\mathbb{V}_0$, namely

$$\mathbb{W} = \{v \in H^1(\Omega);\ v|_{\Gamma_B} = 0\}. \tag{2.17}$$

It can be noted that the existence result stated in Lemma 2.2 is still valid for any F in $L^2(0,T;\mathbb{W}')$.

We denote by $\mathcal{T}$ the operator, which associates with any pair (F,D), with F in $L^2(0,T;\mathbb{W}')$ and the datum $D = (0, f_F, \boldsymbol{q}_r, 0)$ satisfying (2.6), the solution u to the problem (2.3)–(2.8). It follows from (2.13)–(2.14) that u^* satisfies

$$u^* - \mathcal{T}(F_B + F(u^*), D) = 0, \tag{2.18}$$

where the quantity $F(u)$ is defined by duality, for a.e. t in $]0,T[$,

$$\langle F(u), v\rangle = \int_\Omega (\partial_t b_*(u))(\boldsymbol{x},t)v(\boldsymbol{x})\mathrm{d}\boldsymbol{x} + \int_\Omega k\circ b_*(u)(\boldsymbol{x},t)\boldsymbol{e}_z\,\cdot\,(\nabla v)(\boldsymbol{x})\mathrm{d}\boldsymbol{x}. \tag{2.19}$$

We first prove some further properties of the operator $\mathcal{T}$.

Lemma 2.3 *The operator $\mathcal{T}$ is continuous from $L^2(0,T;\mathbb{W}')\times L^2(0,T;L^2(\Gamma_F))\times L^2(0,T;$ $L^2(\Gamma_G)^d)$ into the space $L^2(0,T;\mathbb{V}_0)$. Moreover, the following estimate holds:*

$$\begin{aligned}&\Big(\int_0^T |\mathcal{T}(F,f_F,\boldsymbol{q}_r)(\cdot,t)|^2_{H^1(\Omega)}\mathrm{d}t\Big)^{\frac12}\\ &\le \|F\|_{L^2(0,T;\mathbb{W}')} + c\,\|f_F\|_{L^2(0,T:L^2(\Gamma_F))} + c\,\|\boldsymbol{q}_r\|_{L^2(0,T;L^2(\Gamma_G)^d)}.\end{aligned} \tag{2.20}$$

Proof We set $u=\mathcal{T}(F,f_F,\boldsymbol{q}_r)$ and only prove the estimate (indeed, it is readily checked that it implies the continuity property). We take v equal to $\frac{u}{2}$ in the problem (2.8). This obviously gives

$$\begin{aligned}&\frac{\alpha}{2}\int_\Omega(\partial_t u^2)(\boldsymbol{x},t)\mathrm{d}\boldsymbol{x} + |u(\cdot,t)|^2_{H^1(\Omega)}\\ &\le (\|F(\cdot,t)\|_{\mathbb{W}'} + c\,\|f_F(\cdot,t)\|_{L^2(\Gamma_F)} + c\,\|\boldsymbol{q}_r(\cdot,t)\|_{L^2(\Gamma_G)^d})|u(\cdot,t)|_{H^1(\Omega)},\end{aligned}$$

where c is the norm of the trace operator. Thus, integrating with respect to t gives the estimate (2.20).

Lemma 2.4 *The operator $\mathcal{T}$ is continuous from $L^2(0,T;L^2(\Omega))\times H^1(0,T;\ L^2(\Gamma_F))\times$ $H^1(0,T;L^2(\Gamma_G)^d)$ into the space $H^1(0,T;L^2(\Omega))$. Moreover, the following estimate holds: for any positive ε,*

$$\begin{aligned}&\alpha\,\|\partial_t\mathcal{T}(F,f_F,\boldsymbol{q}_r)\|_{L^2(0,T;L^2(\Omega))}\\ &\le (1+\varepsilon)\|F\|_{L^2(0,T;L^2(\Omega))} + c\,\|f_F\|_{H^1(0,T:L^2(\Gamma_F))} + c\,\|\boldsymbol{q}_r\|_{H^1(0,T;L^2(\Gamma_G)^d)}.\end{aligned} \tag{2.21}$$

Proof The continuity property of $\mathcal{T}$ is proved in [10, Chapitre 6, Théorème 2.1]. Next, setting $u=\mathcal{T}(F,f_F,\boldsymbol{q}_r)$, we take v equal to $u-\eta\,\partial_t u$ in (2.8) for a positive η. Indeed, we have that:

(1) Since u vanishes on Γ_B, so does $\partial_t u$.

(2) Since u is nonpositive on Γ_G and $u(\boldsymbol{x},t-\eta)$, which is close to $u(\boldsymbol{x},t)-\eta\,\partial_t u(\boldsymbol{x},t)$, is also nonpositive, there exists an $\eta>0$, such that $u-\eta\,\partial_t u$ belongs to $\mathbb{V}_0$.

This yields

$$\begin{aligned}&\alpha\,\|\partial_t u\|^2_{L^2(\Omega)} + \frac12\partial_t|u|^2_{H^1(\Omega)}\\ &\le \|F\|_{L^2(\Omega)}\|\partial_t u\|_{L^2(\Omega)} - \int_{\Gamma_F} f_F(\tau,t)\partial_t u(\tau,t)\mathrm{d}\tau - \int_{\Gamma_G}(\boldsymbol{q}_r\cdot\boldsymbol{n})(\tau,t)\partial_t u(\tau,t)\mathrm{d}\tau.\end{aligned}$$

To bound the first term, we use Young's inequality

$$\|F\|_{L^2(\Omega)}\|\partial_t u\|_{L^2(\Omega)} \le \frac{\alpha}{2}\|\partial_t u\|^2_{L^2(\Omega)} + \frac{1}{2\alpha}\|F\|^2_{L^2(\Omega)}.$$

To handle the last two integrals, we integrate them by parts with respect to t. For instance, we have, for any $\varepsilon > 0$,

$$\int_0^t \int_{\Gamma_F} f_F(\tau,s)\partial_t u(\tau,s)\mathrm{d}\tau\mathrm{d}s$$
$$= \int_{\Gamma_F} f_F(\tau,t)u(\tau,t)\mathrm{d}\tau\mathrm{d}s - \int_0^t \int_{\Gamma_F} \partial_t f_F(\tau,s)\, u(\tau,s)\mathrm{d}\tau\mathrm{d}s$$
$$\leq \frac{1}{4}|u(\cdot,t)|^2_{H^1(\Omega)} + c\|f_F(\cdot,t)\|^2_{L^2(\Gamma_F)} + c\|\partial_t f_F\|^2_{L^2(0,t;L^2(\Gamma_F))} + \varepsilon\,\|u\|^2_{L^2(0,t;H^1(\Omega))}.$$

Thus, the desired estimate follows by combining all of those and using (2.20).

We are thus in a position to prove the first existence result.

Theorem 2.1 *Assume that the coefficient α satisfies*

$$\frac{1}{\alpha}\,\|b'\|_{L^\infty(\mathbb{R})} < 1. \tag{2.22}$$

For any data u_B, f_F, $\boldsymbol{q}_r$ and u_0 satisfying

$$\begin{aligned} &u_B \in H^1(0,T;H^{\frac{1}{2}}_{00}(\Gamma_B)), \quad f_F \in H^1(0,T;L^2(\Gamma_F)),\\ &\boldsymbol{q}_r \in H^1(0,T;L^2(\Gamma_G)^d), \quad u_0 \in H^1(\Omega) \end{aligned} \tag{2.23}$$

and (2.10)*, the problem* (2.3)–(2.4) *has at least a solution in* $\mathbb{X}$.

Proof We proceed in several steps.

(1) Let $\mathbb{X}_0$ be the space of functions of $\mathbb{X}$ vanishing at $t = 0$. We provide it with the norm

$$\|v\|_{\mathbb{X}_0} = \|\partial_t v\|_{L^2(0,T;L^2(\Omega))}.$$

It follows from the Lemma 2.4 that

$$\|\mathcal{T}(F_B + F(u^*), D)\|_{\mathbb{X}_0} \leq \frac{1+\varepsilon}{\alpha}\,\|F(u^*)\|_{L^2(0,T;L^2(\Omega))} + c(D),$$

where the constant $c(D)$ only depends on the data u_B, f_F and $\boldsymbol{q}_r$. Due to the boundedness of b' and $k \circ b$ (see (2.19) for the definition of $F(u^*)$), we have

$$\|\mathcal{T}(F_B + F(u^*), D)\|_{\mathbb{X}_0} \leq \frac{1+\varepsilon}{\alpha}\,\|b'\|_{L^\infty(\mathbb{R})}\,\|u^*\|_{\mathbb{X}_0} + c'(D).$$

Thus, due to (2.22), the application: $u^* \mapsto \mathcal{T}(F_B + F(u^*), D)$ maps the ball in $\mathbb{X}_0$ with radius R into itself for all R, such that, for an appropriate ε,

$$\Big(1 - \frac{1+\varepsilon}{\alpha}\,\|b'\|_{L^\infty(\mathbb{R})}\Big)R > c'(D). \tag{2.24}$$

(2) Since $\mathbb{X}_0$ is separable, there exists an increasing sequence of finite-dimensional spaces $\mathbb{X}_n$, which is dense in $\mathbb{X}_0$. If Π_n denotes the orthogonal projection operator (for the scalar product associated with the norm of $\mathbb{X}_0$) onto $\mathbb{X}_n$, the mapping: $u \mapsto \Pi_n\mathcal{T}(F_B + F(u), D)$ is continuous from $\mathbb{X}_n$ into itself. The same arguments as previously yield that it maps the ball of $\mathbb{X}_n$ with radius R into itself for all R satisfying (2.24). Thus, applying the Brouwer's

fixed point theorem (see [9, Chapter IV, Theorem 1.1] for instance), implies that this mapping admits a fixed point in this same ball, namely, there exists a u_n in $\mathbb{X}_n$ satisfying the equation $u_n = \Pi_n \mathcal{T}(F_B + F(u_n), D)$. Moreover, it follows from Lemma 2.3 that this sequence is also bounded in $L^2(0,T;H^1(\Omega))$.

(3) The function u_n thus satisfies,

$$\begin{aligned}
\forall v \in \mathbb{X}_n, \quad & \alpha \int_\Omega (\partial_t u_n)(\boldsymbol{x},t)(v-u_n)(\boldsymbol{x},t)\mathrm{d}\boldsymbol{x} + \int_\Omega (\partial_t b_*(u_n))(\boldsymbol{x},t)(v-u_n)(\boldsymbol{x},t)\mathrm{d}\boldsymbol{x} \\
& + \int_\Omega (\nabla u_n + k \circ b_*(u_n)\boldsymbol{e}_z)(\boldsymbol{x},t) \cdot (\nabla(v-u_n))(\boldsymbol{x},t)\mathrm{d}\boldsymbol{x} \\
& \geq -\int_\Omega F_B(\boldsymbol{x},t)(v-u_n)(\boldsymbol{x},t)\mathrm{d}\boldsymbol{x} - \int_{\Gamma_F} f_F(\tau,t)(v-u_n)(\tau,t)\mathrm{d}\tau \\
& - \int_{\Gamma_G} (\boldsymbol{q}_r \cdot \boldsymbol{n})(\tau,t)(v-u_n)(\tau,t)\mathrm{d}\tau.
\end{aligned} \tag{2.25}$$

Moreover, due to the boundedness properties of the sequence $(u_n)_n$, there exists a subsequence still denoted by $(u_n)_n$ for simplicity, which converges to a function u^* of $\mathbb{X}_0$ weakly in $\mathbb{X}$ and strongly in $L^2(0,T;L^2(\Omega))$. Next, we observe that, for a fixed v in $\mathbb{X}_n$:

(i) The convergence of all terms in the right-hand side follows from the weak convergence in $L^2(0,T;\mathbb{W})$.

(ii) The convergence of the first term is derived by writing the expansion

$$\begin{aligned}
& \int_\Omega (\partial_t u_n)(\boldsymbol{x},t)(v-u_n)(\boldsymbol{x},t)\mathrm{d}\boldsymbol{x} \\
&= \int_\Omega (\partial_t u^*)(\boldsymbol{x},t)(v-u^*)(\boldsymbol{x},t)\mathrm{d}\boldsymbol{x} \\
&\quad + \int_\Omega \partial_t(u_n - u^*)(\boldsymbol{x},t)(v-u^*)(\boldsymbol{x},t)\mathrm{d}\boldsymbol{x} \\
&\quad + \int_\Omega (\partial_t u_n)(\boldsymbol{x},t)(u^*-u_n)(\boldsymbol{x},t)\mathrm{d}\boldsymbol{x}
\end{aligned}$$

and by checking that the last two terms converge.

(iii) The convergence of the term $\int_\Omega (\nabla u_n)(\boldsymbol{x},t) \cdot (\nabla(v-u_n))(\boldsymbol{x},t)\mathrm{d}\boldsymbol{x}$ is obtained by using the weak lower semi-continuity of the norm $|u_n|_{H^1(\Omega)}$.

Moreover, the convergence of the nonlinear terms follows from the expansions

$$\begin{aligned}
& \int_\Omega (\partial_t b_*(u_n))(\boldsymbol{x},t)(v-u_n)(\boldsymbol{x},t)\mathrm{d}\boldsymbol{x} \\
&= \int_\Omega (\partial_t b_*(u^*))(\boldsymbol{x},t)(v-u^*)(\boldsymbol{x},t)\mathrm{d}\boldsymbol{x} \\
&\quad + \int_\Omega (\partial_t b_*(u_n) - \partial_t b_*(u^*))(\boldsymbol{x},t)(v-u^*)(\boldsymbol{x},t)\mathrm{d}\boldsymbol{x} \\
&\quad + \int_\Omega (\partial_t b_*(u_n))(\boldsymbol{x},t)(u^*-u_n)(\boldsymbol{x},t)\mathrm{d}\boldsymbol{x}
\end{aligned}$$

and

$$\begin{aligned}&\int_\Omega k\circ b_*(u_n)(\boldsymbol{x},t)\boldsymbol{e}_z\cdot(\nabla(v-u_n))(\boldsymbol{x},t)\mathrm{d}\boldsymbol{x}\\ =&\int_\Omega k\circ b_*(u^*)(\boldsymbol{x},t)\boldsymbol{e}_z\cdot(\nabla(v-u^*))(\boldsymbol{x},t)\mathrm{d}\boldsymbol{x}\\ &+\int_\Omega k\circ b_*(u^*)(\boldsymbol{x},t)\boldsymbol{e}_z\cdot(\nabla(u^*-u_n))(\boldsymbol{x},t)\mathrm{d}\boldsymbol{x}\\ &+\int_\Omega \big(k\circ b_*(u_n)-k\circ b_*(u^*)\big)(\boldsymbol{x},t)\boldsymbol{e}_z\cdot(\nabla(v-u_n))(\boldsymbol{x},t)\mathrm{d}\boldsymbol{x},\end{aligned}$$

combined with the Lipschitz-continuity of b' and $k\circ b$. Finally, using the density of the sequence $(\mathbb{X}_n)_n$ in $\mathbb{X}_0$, u^* is a solution to the problem (2.13)–(2.14). Thus, u is a solution to the problem (2.3)–(2.4).

Condition (2.22) is rather restrictive, since, in practical situations, α is small. However, this condition can be relaxed when b satisfies, for a positive constant b_0,

$$b'(\xi)\ge b_0\quad\forall\xi\in\mathbb{R}. \tag{2.26}$$

Indeed, all the previous arguments are still valid when we replace α by $\alpha+b_0$ and replace the coefficient $b(\xi)$ by $b(\xi)-b_0\,\xi$.

Corollary 2.1 *Assume that b satisfies* (2.26), *and that the coefficient α satisfies*

$$\frac{1}{\alpha+b_0}\,\|b'-b_0\|_{L^\infty(\mathbb{R})}<1. \tag{2.27}$$

For any data u_B, f_F, $\boldsymbol{q}_r$ and u_0 satisfying (2.10) *and* (2.23), *the problem* (2.3)–(2.4) *has at least a solution in $\mathbb{X}$.*

Assume that b satisfies

$$\min_{\xi\in\mathbb{R}} b'(\xi)>0,\quad \max_{\xi\in\mathbb{R}} b'(\xi)<2\,\min_{\xi\in\mathbb{R}} b'(\xi). \tag{2.28}$$

Under this condition, the problem (2.3)–(2.4) has a solution even for $\alpha=0$. We refer to [2] for another proof of this result of a similar problem.

3 The Discrete Problems

We present first the time semi-discrete problem constructed from the backward Euler's scheme. Next, we consider a finite element discretization of this problem relying on standard, conforming, finite element spaces.

3.1 A Time semi-discrete problem

Since we intend to work with nonuniform 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 $(\tau_1,\cdots,\tau_N)$ 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 as follows.

Find $(u^n)_{0 \le n \le N}$ in $\prod_{n=0}^{N} \mathbb{V}(t_n)$, such that

$$u^0 = u_0 \quad \text{in } \Omega, \tag{3.1}$$

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

$$\begin{aligned}\forall v \in \mathbb{V}(t_n), \quad & \alpha \int_\Omega \Big(\frac{u^n - u^{n-1}}{\tau_n}\Big)(\boldsymbol{x})(v-u^n)(\boldsymbol{x})\mathrm{d}\boldsymbol{x} + \int_\Omega \Big(\frac{b(u^n) - b(u^{n-1})}{\tau_n}\Big)(\boldsymbol{x})(v-u^n)(\boldsymbol{x})\mathrm{d}\boldsymbol{x} \\ & + \int_\Omega (\nabla u^n + k \circ b(u^{n-1}))(\boldsymbol{x})\boldsymbol{e}_z \cdot \nabla(v-u^n)\boldsymbol{x}\mathrm{d}\boldsymbol{x} \\ & \ge -\int_{\Gamma_F} f_F(\tau, t_n)(v-u^n)(\tau)\mathrm{d}\tau - \int_{\Gamma_G} (\boldsymbol{q}_r \cdot \boldsymbol{n})(\tau, t_n)(v-u^n)(\tau)\mathrm{d}\tau. \end{aligned} \tag{3.2}$$

It can be noted that this problem makes sense when both f_F and $\boldsymbol{q}_r$ are continuous in time. Proving its well-posedness relies on rather different arguments as previously.

Theorem 3.1 *For any data u_B, f_F, $\boldsymbol{q}_r$ and u_0 satisfying*

$$\begin{aligned} & u_B \in H^1(0,T; H^{\frac{1}{2}}_{00}(\Gamma_B)), \quad f_F \in \mathscr{C}^0(0,T; L^2(\Gamma_F)), \\ & \boldsymbol{q}_r \in \mathscr{C}^0(0,T; L^2(\Gamma_G)^d), \quad u_0 \in H^1(\Omega), \end{aligned} \tag{3.3}$$

and (2.10)*, for any nonnegative coefficient α, the problem* (3.1)–(3.2) *has a unique solution in* $\prod_{n=0}^{N} \mathbb{V}(t_n)$.

Proof We proceed by induction on n. Since u^0 is given by (3.1), we assume that u^{n-1} is known. We consider problem (3.2) for a fixed n, called $(3.2)_n$, that can equivalently be written as

$$\begin{aligned}\forall v \in \mathbb{V}(t_n), \quad & \int_\Omega (\alpha\, u^n + b(u^n))(\boldsymbol{x})(v-u^n)(\boldsymbol{x})\mathrm{d}\boldsymbol{x} + \tau_n \int_\Omega \nabla u^n(\boldsymbol{x}) \cdot \nabla(v-u^n)(\boldsymbol{x})\mathrm{d}\boldsymbol{x} \\ & \ge \int_\Omega (\alpha\, u^{n-1} + b(u^{n-1}))(\boldsymbol{x})(v-u^n)(\boldsymbol{x})\mathrm{d}\boldsymbol{x} \\ & - \tau_n \int_\Omega k \circ b(u^{n-1})(\boldsymbol{x})\boldsymbol{e}_z \cdot \nabla(v-u^n)(\boldsymbol{x})\mathrm{d}\boldsymbol{x} - \tau_n \int_{\Gamma_F} f_F(\tau, t_n)(v-u^n)(\tau)\mathrm{d}\tau \\ & - \tau_n \int_{\Gamma_G} (\boldsymbol{q}_r \cdot \boldsymbol{n})(\tau, t_n)(v-u^n)(\tau)\mathrm{d}\tau. \end{aligned}$$

Let us now set

$$\varphi(z) = \int_0^z (\alpha\, \zeta + b(\zeta))\mathrm{d}\zeta, \quad \Phi(v) = \int_\Omega \varphi(v(\boldsymbol{x}))\mathrm{d}\boldsymbol{x}.$$

It is readily checked that, since b' is nonnegative, both φ and Φ are convex, and moreover, that

$$D\Phi(u) \cdot (v-u^n) = \int_\Omega (\alpha\, u + b(u))(\boldsymbol{x})(v-u^n)(\boldsymbol{x})\mathrm{d}\boldsymbol{x}.$$

Thus, taking

$$a(u,v) = \int_\Omega \nabla u(\boldsymbol{x}) \cdot \nabla v(\boldsymbol{x}) \mathrm{d}\boldsymbol{x},$$

$$\ell(v) = \int_\Omega (\alpha\, u^{n-1} + b(u^{n-1}))(\boldsymbol{x}) v(\boldsymbol{x}) \mathrm{d}\boldsymbol{x} - \tau_n \int_\Omega k \circ b(u^{n-1})(\boldsymbol{x}) \boldsymbol{e}_z \cdot \nabla v(\boldsymbol{x}) \mathrm{d}\boldsymbol{x}$$

$$- \tau_n \int_{\Gamma_F} f_F(\tau, t_n) v(\tau) \mathrm{d}\tau - \tau_n \int_{\Gamma_G} (\boldsymbol{q}_r \cdot \boldsymbol{n})(\tau, t_n) v(\tau) \mathrm{d}\tau,$$

the problem $(3.2)_n$ can also be written as

$$D\Phi(u^n) \cdot (v - u^n) + a(u^n, v - u^n) - \ell(v - u^n) \geq 0, \quad \forall v \in \mathbb{V}(t_n).$$

We now set $\Psi(v) = \Phi(v) + J(v)$ with $J(v) = \frac{1}{2} a(v,v) - \ell(v)$. The problem $(3.2)_n$ can finally be written as

$$\forall v \in \mathbb{V}(t_n), \quad D\Psi(u^n) \cdot (v - u^n) \geq 0,$$

or

$$\forall v \in \mathbb{V}(t_n), \quad \Psi(u^n) \leq \Psi(v).$$

So it is equivalent to the minimization of a convex functional on the convex set $\mathbb{V}(t_n)$. Hence it admits a unique solution. This completes the proof.

It can be noted that, in contrast with the continuous problem, the existence of a solution to the semi-discrete problem (3.1)–(3.2) does not require any limitation on α.

3.2 A Fully discrete problem

From now on, we assume that Ω is a polygon ($d = 2$) or a polyhedron ($d = 3$). Let $(\mathcal{T}_h)_h$ be a regular family of triangulations of Ω (by triangles or tetrahedra), in the sense that, for each h,

(i) $\overline{\Omega}$ is the union of all elements of $\mathcal{T}_h$.

(ii) The intersection of two different elements of $\mathcal{T}_h$, if not empty, is a vertex or a whole edge or a whole face of both of them.

(iii) The ratio of the diameter h_K of any element K of $\mathcal{T}_h$ to the diameter of its inscribed circle or sphere is smaller than a constant σ independent of h.

As usual, h stands for the maximum of the diameters h_K ($K \in \mathcal{T}_h$). We make the further and nonrestrictive assumption that $\overline{\Gamma}_B$, $\overline{\Gamma}_F$ and $\overline{\Gamma}_G$ are the union of whole edges ($d = 2$) or whole faces ($d = 3$) of elements of $\mathcal{T}_h$. From now on, c, c', $\cdots$ stand for generic constants that may vary from line to line and are always independent of τ and h.

We now introduce the finite element space

$$\overline{\mathbb{V}}_h = \{v_h \in H^1(\Omega);\ \forall K \in \mathcal{T}_h,\ v_h|_K \in \mathcal{P}_1(K)\}, \tag{3.4}$$

where $\mathcal{P}_1(K)$ is the space of restrictions to K of affine functions on $\mathbb{R}^d$. Let $\mathcal{I}_h$ denote the Lagrange interpolation operator at all the vertices of elements of $\mathcal{T}_h$ with values in $\overline{\mathbb{V}}_h$, and i_h^B denote the corresponding interpolation operator on Γ_B. Assuming that u_B is continuous where needed, we then define for each n ($0 \leq n \leq N$), the subset of $\overline{\mathbb{V}}_h$,

$$\mathbb{V}_h(t_n) = \{v_h \in \overline{\mathbb{V}}_h;\ v_h|_{\Gamma_B} = i_h^B u_B(\cdot, t_n) \text{ and } v_h|_{\Gamma_G} \leq 0\}. \tag{3.5}$$

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

Find $(u_h^n)_{0\le n\le N}$ in $\prod_{n=0}^{N} \mathbb{V}_h(t_n)$, such that

$$u_h^0 = \mathcal{I}_h u_0 \quad \text{in } \Omega, \tag{3.6}$$

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

$$\begin{aligned}\forall v_h \in \mathbb{V}_h(t_n), \quad & \alpha \int_\Omega \Big(\frac{u_h^n - u_h^{n-1}}{\tau_n}\Big)(\boldsymbol{x})(v_h - u_h^n)(\boldsymbol{x})\mathrm{d}\boldsymbol{x} \\ & + \int_\Omega \Big(\frac{b(u_h^n) - b(u_h^{n-1})}{\tau_n}\Big)(\boldsymbol{x})(v_h - u_h^n)(\boldsymbol{x})\mathrm{d}\boldsymbol{x} \\ & + \int_\Omega (\nabla u_h^n + k \circ b(u_h^{n-1}))(\boldsymbol{x})\boldsymbol{e}_z \cdot \nabla(v_h - u_h^n)(\boldsymbol{x})\mathrm{d}\boldsymbol{x} \\ \ge & - \int_{\Gamma_F} f_F(\tau, t_n)(v_h - u_h^n)(\tau)\mathrm{d}\tau - \int_{\Gamma_G} (\boldsymbol{q}_r \cdot \boldsymbol{n})(\tau, t_n)(v_h - u_h^n)(\tau)\mathrm{d}\tau.\end{aligned} \tag{3.7}$$

The proof of the next theorem is exactly the same as the proof of Theorem 3.1, so we omit it.

Theorem 3.2 *For any data u_B, f_F, $\boldsymbol{q}_r$ and u_0 satisfying* (2.10), (3.3) *and*

$$u_B \in \mathscr{C}^0(\overline{\Gamma}_B \times [0,T]), \quad u_0 \in \mathscr{C}^0(\overline{\Omega}) \tag{3.8}$$

for any nonnegative coefficient α, the problem (3.6)–(3.7) *has a unique solution.*

Here also the existence result is unconditional.

4 A Convergence Result

The aim of this section is to prove a convergence result for the solutions $(u_h^n)_{0\le n\le N}$ to the problem (3.6)–(3.7), when $|\tau|$ and h tend to zero. In order to do that, as in Section 2, we use the lifting u_B^* of u_B which satisfies (2.11)–(2.12), and assume moreover that it is continuous on $\overline{\Omega} \times [0,T]$. Indeed, if $(u_h^n)_{0\le n\le N}$ is a solution to (3.6)–(3.7), and the family $(u_h^{*n})_{0\le n\le N}$ with $u_h^{*n} = u_h^n - \mathcal{I}_h u_B^*(t_n)$ is a solution to the following problem:

Find $(u_h^{*n})_{0\le n\le N}$ in $\mathbb{V}_{h0}^{N+1}$, such that

$$u_h^{*0} = 0 \quad \text{in } \Omega, \tag{4.1}$$

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

$$\begin{aligned}\forall v_h \in \mathbb{V}_{h0}, \quad & \alpha \int_\Omega \Big(\frac{u_h^{*n} - u_h^{*n-1}}{\tau_n}\Big)(\boldsymbol{x})(v_h - u_h^{*n})(\boldsymbol{x})\mathrm{d}\boldsymbol{x} \\ & + \int_\Omega \Big(\frac{b_{*n}(u_h^{*n}) - b_{*n-1}(u_h^{*n-1})}{\tau_n}\Big)(\boldsymbol{x})(v_h - u_h^{*n})(\boldsymbol{x})\mathrm{d}\boldsymbol{x} \\ & + \int_\Omega (\nabla u_h^{*n} + k \circ b_{*n-1}(u_h^{*n-1}))(\boldsymbol{x})\boldsymbol{e}_z \cdot \nabla(v_h - u_h^{*n})(\boldsymbol{x})\mathrm{d}\boldsymbol{x} \\ \ge & - \int_\Omega F_{Bh}(\boldsymbol{x}, t_n)(v_h - u_h^{*n})\mathrm{d}\boldsymbol{x} - \int_{\Gamma_F} f_F(\tau, t_n)(v_h - u_h^{*n})(\tau)\mathrm{d}\tau \\ & - \int_{\Gamma_G} (\boldsymbol{q}_r \cdot \boldsymbol{n})(\tau, t_n)(v_h - u_h^{*n})(\tau)\mathrm{d}\tau,\end{aligned} \tag{4.2}$$

where the convex set $\mathbb{V}_{h0}$ and the function F_{Bh} are defined, in analogy with (2.15)–(2.16), by

$$\mathbb{V}_{h0} = \overline{\mathbb{V}}_h \cap \mathbb{V}_0 \tag{4.3}$$

and

$$\begin{aligned}&\int_\Omega F_{Bh}(\boldsymbol{x},t)v(\boldsymbol{x})\mathrm{d}\boldsymbol{x}\\ &= \alpha \int_\Omega (\partial_t \mathcal{I}_h u_B^*)(\boldsymbol{x},t)v(\boldsymbol{x})\mathrm{d}\boldsymbol{x} + \int_\Omega (\nabla \mathcal{I}_h u_B^*)(\boldsymbol{x},t) \cdot (\nabla v)(\boldsymbol{x})\mathrm{d}\boldsymbol{x},\end{aligned} \tag{4.4}$$

while each function b_{*n} is given by $b_{*n}(\xi) = b(\xi + \mathcal{I}_h u_B^*(\cdot, t_n))$. We now investigate the boundedness of the sequence $(u_h^{*n})_{0\le n\le N}$ in appropriate norms. We need a preliminary lemma for that.

Lemma 4.1 *For each part Γ of $\partial\Omega$, which is the union of whole edges ($d = 2$) or whole faces ($d = 3$) of elements of $\mathcal{T}_h$, the following inequality holds for all functions w_h in $\overline{\mathbb{V}}_h$:*

$$\|w_h\|_{H^{-\frac{1}{2}}(\Gamma)} \le c\, \|w_h\|_{L^2(\Omega)}. \tag{4.5}$$

Proof It relies on standard arguments. We have

$$\|w_h\|_{H^{-\frac{1}{2}}(\Gamma)} = \sup_{z\in H^{\frac{1}{2}}(\Gamma)} \frac{\int_\Gamma z(\tau)w_h(\tau)\mathrm{d}\tau}{\|z\|_{H^{\frac{1}{2}}(\Gamma)}}.$$

Let e be any edge or face of an element K of $\mathcal{T}_h$ which is contained in Γ. Denoting by $\widehat{K}$ the reference triangle or tetrahedron, we have, with obvious notation for $\widehat{e}$, $\widehat{w}$, $\widehat{z}$,

$$\int_e z(\tau)w_h(\tau)\mathrm{d}\tau \le c\, h_e^{d-1} \int_{\widehat{e}} \widehat{z}(\widehat{\tau})\widehat{w}_h(\widehat{\tau})\mathrm{d}\widehat{\tau} \le c'\, h_K^{d-1}\, \|\widehat{z}\|_{L^2(\widehat{e})}\|\widehat{w}_h\|_{L^2(\widehat{e})}.$$

By using the equivalence of norms on $\mathcal{P}_1(\widehat{K})$ and an appropriate stable lifting operator $\widehat{\pi}$ which maps traces on $\widehat{e}$ into functions of K vanishing at the vertex of K which does not belong to $\overline{\Gamma}$, we derive

$$\int_e z(\tau)w_h(\tau)\mathrm{d}\tau \le c'\, h_K^{d-1}\, |\widehat{\pi}\widehat{z}|_{H^1(\widehat{K})}\|\widehat{w}_h\|_{L^2(\widehat{K})} \le c'\, h_K^{d-1}\, h_K^{1-\frac{d}{2}}|\pi z|_{H^1(K)} h_K^{-\frac{d}{2}}\|w_h\|_{L^2(K)},$$

there also with an obvious definition of π. We conclude by summing this last inequality on e and by using a Cauchy-Schwarz inequality and the stability of $\widehat{\pi}$,

$$\int_\Gamma z(\tau)w_h(\tau)\mathrm{d}\tau \le c\|z\|_{H^{\frac{1}{2}}(\Gamma)}\|w_h\|_{L^2(\Omega)},$$

whence the desired result follows.

Lemma 4.2 *For any data u_B, f_F, $\boldsymbol{q}_r$ and u_0 satisfying*

$$\begin{aligned}&u_B \in H^1(0,T; H_{00}^{\frac{1}{2}}(\Gamma_B)), \quad f_F \in \mathscr{C}^0(0,T; H^{\frac{1}{2}}(\Gamma_F)),\\ &\boldsymbol{q}_r \in \mathscr{C}^0(0,T; H^{\frac{1}{2}}(\Gamma_G)^d), \quad u_0 \in H^1(\Omega)\end{aligned} \tag{4.6}$$

and (2.10)*, the sequence* $(u_h^{*n})_{0\le n\le N}$ *satisfies the following inequality, for* $1 \le n \le N$*,*

$$\begin{aligned} &\alpha \sum_{m=1}^{n} \tau_m \left\| \frac{u_h^{*m} - u_h^{*m-1}}{\tau_m} \right\|_{L^2(\Omega)}^2 + |u_h^{*n}|_{H^1(\Omega)}^2 \\ &\le c\,(1 + \|\mathcal{I}_h u_B^*\|_{H^1(0,T;H^1(\Omega))}^2 + \|f_F\|_{\mathscr{C}^0(0,T;H^{\frac12}(\Gamma_F))}^2 + \|\boldsymbol{q}_r\|_{\mathscr{C}^0(0,T;H^{\frac12}(\Gamma_G)^d)}^2). \end{aligned} \tag{4.7}$$

Proof Taking v equal to u_h^{*n-1} in (4.2), leads to

$$\begin{aligned} &\alpha\,\tau_n \left\| \frac{u_h^{*n} - u_h^{*n-1}}{\tau_n} \right\|_{L^2(\Omega)}^2 + \int_\Omega \nabla u_h^{*n}(\boldsymbol{x}) \cdot \nabla (u_h^{*n} - u_h^{*n-1})(\boldsymbol{x})\mathrm{d}\boldsymbol{x} \\ &\le - \int_\Omega \Big(\frac{b_{*n}(u_h^{*n}) - b_{*n-1}(u_h^{*n-1})}{\tau_n} \Big)(\boldsymbol{x})(u_h^{*n} - u_h^{*n-1})(\boldsymbol{x})\mathrm{d}\boldsymbol{x} \\ &\quad - \int_\Omega k \circ b_{*n-1}(u_h^{*n-1})(\boldsymbol{x})\boldsymbol{e}_z \cdot \nabla (u_h^{*n} - u_h^{*n-1})(\boldsymbol{x})\mathrm{d}\boldsymbol{x} + \langle \mathcal{G}, u_h^{*n} - u_h^{*n-1} \rangle, \end{aligned}$$

where the data depending quantity $\mathcal{G}$ is defined by

$$\langle \mathcal{G}, v \rangle = - \int_\Omega F_{Bh}(\boldsymbol{x}, t_n) v(\boldsymbol{x})\mathrm{d}\boldsymbol{x} - \int_{\Gamma_F} f_F(\tau, t_n) v(\tau)\mathrm{d}\tau - \int_{\Gamma_G} (\boldsymbol{q}_r \cdot \boldsymbol{n})(\tau, t_n) v(\tau)\mathrm{d}\tau.$$

To handle the second term, we use the identity

$$\int_\Omega \nabla u_h^{*n} \cdot \nabla (u_h^{*n} - u_h^{*n-1})(\boldsymbol{x})\mathrm{d}\boldsymbol{x} = \frac{1}{2}(|u_h^{*n}|_{H^1(\Omega)}^2 + |u_h^{*n} - u_h^{*n-1}|_{H^1(\Omega)}^2 - |u_h^{*n-1}|_{H^1(\Omega)}^2).$$

To handle the third term, we write the expansion

$$\begin{aligned} &\int_\Omega \Big(\frac{b_{*n}(u_h^{*n}) - b_{*n-1}(u_h^{*n-1})}{\tau_n} \Big)(\boldsymbol{x})(u_h^{*n} - u_h^{*n-1})(\boldsymbol{x})\mathrm{d}\boldsymbol{x} \\ &= \int_\Omega \Big(\frac{b(u_h^{*n} + \mathcal{I}_h u_B^*(t_n)) - b(u_h^{*n-1} + \mathcal{I}_h u_B^*(t_n))}{\tau_n} \Big)(\boldsymbol{x})(u_h^{*n} - u_h^{*n-1})(\boldsymbol{x})\mathrm{d}\boldsymbol{x} \\ &\quad + \int_\Omega \Big(\frac{b(u_h^{*n-1} + \mathcal{I}_h u_B^*(t_n)) - b(u_h^{*n-1} + \mathcal{I}_h u_B^*(t_{n-1}))}{\tau_n} \Big)(\boldsymbol{x})(u_h^{*n} - u_h^{*n-1})(\boldsymbol{x})\mathrm{d}\boldsymbol{x}. \end{aligned}$$

By using the nonnegativity of b', together with the Lipschitz-continuity of b, we derive

$$\begin{aligned} &\int_\Omega \Big(\frac{b_{*n}(u_h^{*n}) - b_{*n-1}(u_h^{*n-1})}{\tau_n} \Big)(\boldsymbol{x})(u_h^{*n} - u_h^{*n-1})(\boldsymbol{x})\mathrm{d}\boldsymbol{x} \\ &\le \frac{\alpha}{4}\,\tau_n \Big\| \frac{u_h^{*n} - u_h^{*n-1}}{\tau_n} \Big\|_{L^2(\Omega)}^2 + \frac{1}{\alpha}\,\tau_n \Big\| \frac{\mathcal{I}_h u_B^*(t_n) - \mathcal{I}_h u_B^*(t_{n-1})}{\tau_n} \Big\|_{L^2(\Omega)}^2. \end{aligned}$$

Finally, evaluating the last term is an easy consequence of Lemma 4.1,

$$\begin{aligned} \langle \mathcal{G}, u_h^{*n} - u_h^{*n-1} \rangle &\le \frac{\alpha}{4}\,\tau_n \Big\| \frac{u_h^{*n} - u_h^{*n-1}}{\tau_n} \Big\|_{L^2(\Omega)}^2 \\ &\quad + c\tau_n\,(\|F_{Bh}(\cdot, t_n)\|_{L^2(\Omega)}^2 + \|f_F(\cdot, t_n)\|_{H^{\frac12}(\Gamma_F)}^2 + \|\boldsymbol{q}_r(\cdot, t_n)\|_{H^{\frac12}(\Gamma_G)^d}^2). \end{aligned}$$

By combining, we obtain

$$\begin{aligned}
&\frac{\alpha}{2}\,\tau_n \Big\|\frac{u_h^{*n}-u_h^{*n-1}}{\tau_n}\Big\|^2_{L^2(\Omega)} + \frac{1}{2}\,|u_h^{*n}|^2_{H^1(\Omega)} \\
&\le \frac{1}{2}\,|u_h^{*n-1}|^2_{H^1(\Omega)} \\
&\quad + c'\tau_n\,(\|F_{Bh}(\cdot,t_n)\|^2_{L^2(\Omega)} + \|f_F(\cdot,t_n)\|^2_{H^{\frac{1}{2}}(\Gamma_F)} + \|\boldsymbol{q}_r(\cdot,t_n)\|^2_{H^{\frac{1}{2}}(\Gamma_G)^d}) \\
&\quad - \int_\Omega k\circ b_{*n-1}(u_h^{*n-1})(\boldsymbol{x})\boldsymbol{e}_z\cdot\nabla(u_h^{*n}-u_h^{*n-1})(\boldsymbol{x})\mathrm{d}\boldsymbol{x}.
\end{aligned}$$

We sum up this inequality on n. To handle the last term, we observe that

$$\begin{aligned}
&-\sum_{m=1}^{n}\int_\Omega k\circ b_{*m-1}(u_h^{*m-1})(\boldsymbol{x})\boldsymbol{e}_z\cdot\nabla(u_h^{*m}-u_h^{*m-1})(\boldsymbol{x})\mathrm{d}\boldsymbol{x} \\
&= -\int_\Omega k\circ b_{*n-1}(u_h^{*n-1})(\boldsymbol{x})\boldsymbol{e}_z\cdot\nabla u_h^{*n}(\boldsymbol{x})\mathrm{d}\boldsymbol{x} \\
&\quad + \sum_{m=1}^{n-1}\int_\Omega (k\circ b_{*m}(u_h^{*m}) - k\circ b_{*m-1}(u_h^{*m-1}))(\boldsymbol{x})\boldsymbol{e}_z\cdot\nabla u_h^{*m}(\boldsymbol{x})\mathrm{d}\boldsymbol{x}.
\end{aligned}$$

Hence, thanks to the boundedness of k and the Lipschitz continuity of $k\circ b$, we derive

$$\begin{aligned}
&-\sum_{m=1}^{n}\int_\Omega k\circ b_{*m-1}(u_h^{*m-1})(\boldsymbol{x})\boldsymbol{e}_z\cdot\nabla(u_h^{*m}-u_h^{*m-1})(\boldsymbol{x})\mathrm{d}\boldsymbol{x} \\
&\le c + \frac{1}{4}\,|u_h^{*n}|^2_{H^1(\Omega)} + \frac{\alpha}{4}\sum_{m=1}^{n-1}\tau_m\Big\|\frac{u_h^{*m}-u_h^{*m-1}}{\tau_m}\Big\|^2_{L^2(\Omega)} \\
&\quad + c'\sum_{m=1}^{n-1}\tau_m\Big\|\frac{\mathcal{I}_h u_B^*(t_m)-\mathcal{I}_h u_B^*(t_m-1)}{\tau_m}\Big\|^2_{L^2(\Omega)} + c''\sum_{m=1}^{n-1}\tau_m\,|u^{*m}|^2_{H^1(\Omega)}.
\end{aligned}$$

We conclude by using the discrete Grönwall's lemma (see [8, Chap. V, Lemma 2.4]).

Let us now introduce the function $u^*_{h\tau}$, which is affine on each interval $[t_{n-1},t_n]$ $(1\le n\le N)$, and equal to u_h^{*n} at time t_n $(0\le n\le N)$. When the data u_B, f_F, $\boldsymbol{q}_r$ and u_0 satisfy

$$\begin{aligned}
&u_B\in H^1(0,T;H^s(\Gamma_B)), \qquad f_F\in\mathscr{C}^0(0,T;H^{\frac{1}{2}}(\Gamma_F)), \\
&\boldsymbol{q}_r\in\mathscr{C}^0(0,T;H^{\frac{1}{2}}(\Gamma_G)^d), \quad u_0\in H^{s+\frac{1}{2}}(\Omega),
\end{aligned}\tag{4.8}$$

for some $s>\frac{d-1}{2}$ (in order to ensure the stability of the operator $\mathcal{I}_h$), it follows from Lemma 4.2 that this function belongs to the set $\mathbb{X}_0 = L^2(0,T;\mathbb{V}_0)\cap H^1(0,T;L^2(\Omega))$ (see (2.5) and (2.14)). More precisely, it satisfies

$$\|u^*_{h\tau}\|_{L^2(0,T;H^1(\Omega))\cap H^1(0,T;L^2(\Omega))} \le c(u_B,f_F,\boldsymbol{q}_r), \tag{4.9}$$

where the constant $c(u_B,f_F,\boldsymbol{q}_r)$ only depends on the data. Thus, we are in a position to derive the next result.

Theorem 4.1 *For any data u_B, f_F, $\boldsymbol{q}_r$ and u_0 satisfying* (4.8) *and* (2.10)*, and for any positive coefficient α, the problem* (2.3)–(2.4) *has at least a solution in $\mathbb{X}$.*

Proof Thanks to (4.9), the family of functions $u^*_{h\tau}$ is bounded in $\mathbb{X}_0$ independently of h and τ. Thus, there exist a sequence $(\mathcal{T}_{hk})_k$ of triangulations $\mathcal{T}_h$ and a sequence $(\tau_k)_k$ of parameters τ, such that the sequence $(u^*_k)_k$ converges to a function u^* of $\mathbb{X}_0$ weakly in $L^2(0,T;H^1(\Omega)) \cap H^1(0,T;L^2(\Omega))$ and strongly in $L^2(0,T;L^2(\Omega))$. We now intend to prove that u^* is a solution to the problem (2.13)–(2.14). Since it obviously satisfies (2.13), we now investigate the convergence of all terms in (4.2). For clarity, we keep the notation u_h^{*n} for $u^*_k(t_n)$.

(1) The convergence of the first term follows from the expansion

$$\begin{aligned}
&\alpha \int_\Omega \Big(\frac{u_h^{*n} - u_h^{*n-1}}{\tau_n}\Big)(\boldsymbol{x})(v_h - u_h^{*n})(\boldsymbol{x})\mathrm{d}\boldsymbol{x} \\
&= \alpha \int_\Omega (\partial_t u^*)(\boldsymbol{x}, t_n)(v_h - u^*)(\boldsymbol{x}, t_n)\mathrm{d}\boldsymbol{x} \\
&\quad + \alpha \int_\Omega (\partial_t(u^*_k - u^*))(\boldsymbol{x}, t_n)(v_h - u^*)(\boldsymbol{x}, t_n)\mathrm{d}\boldsymbol{x} \\
&\quad + \alpha \int_\Omega (\partial_t u^*_k)(\boldsymbol{x}, t_n)(u^* - u_h^{*n})(\boldsymbol{x}, t_n)\mathrm{d}\boldsymbol{x}.
\end{aligned}$$

(2) To prove the convergence of the term

$$\int_\Omega \Big(\frac{b_{*n}(u_h^{*n}) - b_{*n-1}(u_h^{*n-1})}{\tau_n}\Big)(\boldsymbol{x})(v_h - u_h^{*n})(\boldsymbol{x})\mathrm{d}\boldsymbol{x},$$

we use a rather complex expansion that we skip for brevity, combined with the dominated convergence theorem of Lebesgue. Indeed, since $(u^*_k)_k$ converges to a function u^* in $L^2(0,T;L^2(\Omega))$, it converges almost everywhere in $\Omega \times [0,T]$, so that $(b'(u^*_k))_k$ also converges a.e. to $b'(u^*)$. Thus, since b' is bounded, $(b'(u^*_k))_k$ also converges to $b'(u^*)$ in $L^2(0,T;L^2(\Omega))$.

(3) The convergence of the term $\int_\Omega \nabla u_h^{*n}(\boldsymbol{x}, t_n)\boldsymbol{e}_z \cdot \nabla(v_h - u_h^{*n})(\boldsymbol{x}, t_n))\mathrm{d}\boldsymbol{x}$ is a consequence of the weak lower semi-continuity of the norm.

(4) The convergence of the term $\int_\Omega k \circ b_{*n-1}(u_h^{*n-1})(\boldsymbol{x})\boldsymbol{e}_z \cdot \nabla(v_h - u_h^{*n})(\boldsymbol{x})\mathrm{d}\boldsymbol{x}$ is easily derived from the expansion

$$\begin{aligned}
&\int_\Omega k \circ b_{*n-1}(u_h^{*n-1})(\boldsymbol{x})\boldsymbol{e}_z \cdot \nabla(v_h - u_h^{*n})(\boldsymbol{x})\mathrm{d}\boldsymbol{x} \\
&= \int_\Omega k \circ b_*(u^*)(\boldsymbol{x}, t_n)\boldsymbol{e}_z \cdot \nabla(v_h - u^*)(\boldsymbol{x}, t_n)\mathrm{d}\boldsymbol{x} \\
&\quad + \int_\Omega (k \circ b_{*n-1} - k \circ b_*)(u^*)(\boldsymbol{x}, t_n)\boldsymbol{e}_z \cdot \nabla(v_h - u^*)(\boldsymbol{x}, t_n)\mathrm{d}\boldsymbol{x} \\
&\quad + \int_\Omega k \circ b_{*n-1}(u^*)(\boldsymbol{x}, t_n)\boldsymbol{e}_z \cdot \nabla(u^* - u_h^{*n})(\boldsymbol{x})\mathrm{d}\boldsymbol{x} \\
&\quad + \int_\Omega (k \circ b_{*n-1}(u_h^{*n-1}) - k \circ b_{*n-1}(u^*))(\boldsymbol{x})\boldsymbol{e}_z \cdot \nabla(v_h - u_h^{*n})(\boldsymbol{x}, t_n)\mathrm{d}\boldsymbol{x},
\end{aligned}$$

and from the dominated convergence theorem of Lebesgue.

(5) The convergence of all terms in the right-hand side of (4.2) is obviously derived from the weak convergence of the sequence $(u^*_k)_k$.

Finally, using the density of the union of the $\mathbb{V}_{h0}$ in $\mathbb{V}_0$, we derive that u^* is a solution to the problem (2.13)–(2.14). Thus, the function $u = u^* + u^*_B$ is a solution to the problem (2.3)–(2.4).

Even if this requires a slightly different regularity of the data, Theorem 4.1 combined with Proposition 2.2 yields that, for any positive coefficient α, the problem (2.3)–(2.4) is well-posed in $\mathbb{X}$. Of course, this is a great improvement of the results in Section 2 and leads to considering that the discretization proposed in Section 3 is rather efficient. We shall check this in the second part of this work.

References

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

[2] Alt, H. W., Luckhaus, S. and Visintin, A., On nonstationary flow through porous media, *Ann. Mat. Pura Appl.*, **136**, 1984, 303–316.

[3] Bernardi, C., El Alaoui, L. and Mghazli, Z., A posteriori analysis of a space and time discretization of a nonlinear model for the flow in variably saturated porous media, submitted.

[4] Berninger, H., Domain decomposition methods for elliptic problems with jumping nonlinearities and application to the Richards equation, Ph. D. Thesis, Freie Universität, Berlin, Germany, 2007.

[5] Brezzi, F., Hager, W. W. and Raviart, P. A., Error estimates for the finite element solution to variational inequalities. II. Mixed methods, *Numer. Math.*, **31**, 1978/1979, 1–16.

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

[7] Gabbouhy, M., 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, University Ibn Tofail, Kénitra, Morocco, 2000.

[8] Girault, V. and Raviart, P. A., Finite Element Approximation of the Navier-Stokes Equations, Lecture Notes in Mathematics, **749**, Springer-Verlag, Berlin, New York, 1979.

[9] Girault, V. and Raviart, P. A., Finite Element Methods for Navier-Stokes Equations, Theory and Algorithms, Springer-Verlag, Berlin, 1986.

[10] Glowinski, R., Lions, J. L. and Trémolières, R., Analyse numérique des inéquations variationnelles. 2. Applications aux phénomènes stationnaires et d'évolution, Collection "Méthodes Mathématiques de l'Informatique" 5, Dunod, Paris, 1976.

[11] Lions, J. L. and Magenes, E., Problèmes aux limites non homogènes et applications, Vol. I, Dunod, Paris, 1968.

[12] Nochetto, R. H. and Verdi, C., Approximation of degenerate parabolic problems using numerical integration, *SIAM J. Numer. Anal.*, **25**, 1988, 784–814.

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

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

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

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

[17] Sochala, P. and Ern, A., Numerical methods for subsurface flows and coupling with runoff, to appear.

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

[19] Woodward, C. S. and Dawson, C. N., 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.