Finite volume approximation of degenerate two-phase flow model with unlimited air mobility

Boris Andreianov, Robert Eymard, Mustapha Ghilani, Nouzha Marhraoui

FINITE VOLUME APPROXIMATION OF DEGENERATE TWO-PHASE FLOW MODEL WITH UNLIMITED AIR MOBILITY. *, **

Boris Andreianov[1], Robert Eymard[2], Mustapha Ghilani[3] and Nouzha Marhraoui[4]

Abstract. Models of two-phase flows in porous media, used in petroleum engineering, lead to a coupled system of two equations, one elliptic, the other degenerate parabolic, with two unknowns: the saturation and the pressure. In view of applications in hydrogeology, we are interested at the singular limit of this model, as the ratio μ of air/liquid mobility goes to infinity, and in a comparison with the one-phase Richards model. We construct a robust finite volume scheme that can apply for large values of the parameter μ. This scheme is shown to satisfy a priori estimates (the saturation is shown to remain in a fixed interval, and a discrete $L^2(0,T;H^1(\Omega))$ estimate is proved for both the pressure and a function of the saturation) which are sufficient to derive the convergence of a subsequence to a weak solution of the continuous equations, as the size of the discretization tends to zero. At the limit as the mobility of the air phase tends to infinity, we obtain the two-phase flow model introduced in the work Henry, Hilhorst and Eymard [14] (see also [13]) which we call the quasi-Richards equation.

2000 Mathematics Subject Classification. 65M12, 65J15.

July 7, 2011.

Contents

Keywords and phrases: Flow in porous media, two-phase flow model, infinite mobility limit, Richards model, Finite Volume method, discrete a priori estimates, convergence of approximate solutions.

* *This work was supported by the project PARS MI06 CNRST and the Project CNRS-CNRST SPM08/10 No.24506.*

** *Equipe EMMACS, ENSAM, UMI, Meknès, Maroc; Equipe AMN-TA, FSM, UMI, Meknès, Maroc.*

[1] UMR CNRS 6623, Université de Franche-Comté, 16 Route de Gray, 25030 Besançon Cedex, France.

[2] Université Marne-la-Vallèe, 5, boulevard Descartes Champs-sur-Marne F-77454 Marne La Vallée CEDEX 2 , France.

[3] ENSAM, BP 4024 Bni M'hamed 50 000, Meknès, Morocco.

[4] Faculté des Sciences, BP 11201 Zitoune 50 000, Meknès, Morocco.

1. Introduction

Hydrologists have been studying the unsaturated flow in soils, mainly using the so-called Richards model. This one doesn't take into account of the air-phase balance equation, replacing it by the assumption that the air phase remains essentially at atmospheric pressure. This hypothesis is not always verified. Published laboratory and field results show that air effects can reduce infiltration rates considerably. Separate or joint effects of the air compressibility, its viscous resistance, or the capillary entry pressure will produce this reduction, see [21,22]. So, the more general two-phase flow model, well known in the context of petroleum engeneering, must be considered. However, the Richards model remains reasonable in most cases because the mobility of air is much larger than that of water, due to the viscosity difference between the two fluids [8].

The main propose of this work is to explore the limit of the two-phase flow model as the mobility of the air phase tends to infinity (cf. Henry, Hilhorst and Eymard [14]), and to propose a finite volume scheme for the two-phase model which can apply to the case of elevated ratio μ of phase mobilities. The main results and ideas of the present paper were previously published in the note [13].

In this paper, we assume that the air and water phases are incompressible and immiscible. For this first study of the limit of the two-phase flow model, the geometric domain is supposed to be horizontal, homogeneous and isotropic; in particular, gravity effects are avoided. The equations of the two-phase flow model in this particular case, using Darcy's law, can be written as :

$$\left\{ \begin{array}{lcl} u_t - \mathrm{div}(k_w(u)\nabla p) & = & s_w, \\ (1-u)_t - \mathrm{div}(\mu k_a(u)\nabla(p + p_c(u))) & = & s_a, \end{array} \right. \tag{1}$$

where u and p, respectively, are the saturation and the pressure of the water phase, k_w and k_a, respectively, represent the relative permeabilities of the water and the air phase, μ is the ratio between the mobility of the air phase and that of the water phase, p_c is the capillary pressure; internal source term s_w for the water phase and s_a for the air phase are present (these source terms are used to represent the exchange terms with the outside of the domain). We suppose in particular that the physical functions k_w, k_a and p_c only depend on the saturation u of the water phase (see Figure 1). The mobility of the air phase is greater than that of the water phase, and, assuming that the functions k_w and k_a are normalized by $k_w(1) = k_a(0) = 1$ and $k_w(0) = k_a(1) = 0$, the aim of this paper is the study of the limit of this physical model as $\mu \to +\infty$. We consider the following particular form of the source terms s_a, s_w:

$$s_w = f_\mu(c)\,\overline{s} - f_\mu(u)\,\underline{s}, \quad s_a = (1 - f_\mu(c))\,\overline{s} - (1 - f_\mu(u))\,\underline{s} \tag{2}$$

(the function f_μ is defined as the fractional flow of the water phase, see the definitions (13) below). This is physically meaningful; indeed, c represents the saturation of the **w**etting injected fluid, and $\overline{s}$, $\underline{s}$ represent the intensity (**t**he injection and extraction velocities) of the sources and sinks that act on the mixture of the phases. The representation (2) is crucial in our study of the two-phase flow.

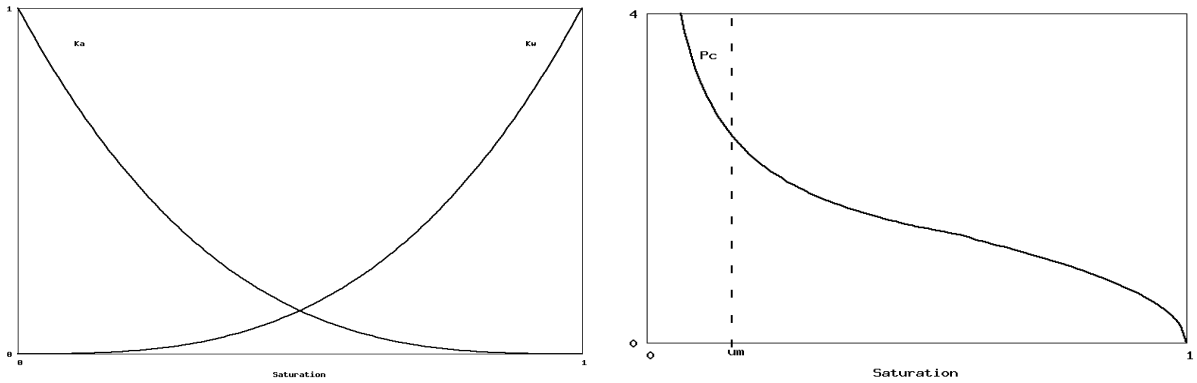

FIGURE 1. Curves shapes of k_w, k_a and p_c

The mathematical analysis of the resulting equations (under varying assumptions) and finite volume schemes for their approximation have been developed for some time now; we refer in particular to the monograph Gagneux and Madaune-Tort [17] and the papers [1, 2, 6, 7, 9, 16, 18]. In what concerns finite volume analysis of the two-phase problem, we refer to the introduction of the work [19] of Michel, and to the papers [15, 20].

Unfortunately, the results given in the above references do not give enough mathematical background to perform the limit analysis as $\mu \to \infty$. Analysis leading to estimates uniform in μ was initiated by the three last authors in [13] and then by Eymard, Henry and Hilhorst in [14]. The present paper extends the work [13]; and it is the numerical counterpart of the analysis carried out in [14], where the authors used parabolic regularization in order to construct solutions to the degenerate problem (5),(6). The regularity assumptions we take on k_a, k_w and p_c are slightly weaker then those of [13] and [14].

Using a specially designed finite volume numerical scheme, we obtain the existence of a solution to the two-phase flow model, which satisfies estimates sufficient to get compactness of the *ad hoc* quantities. Moreover, the dependence of the estimates on the air/liquid mobility ration μ is controlled. Thus one can rigorously pass to the limit in the model (5),(6) as the mobility of the air phase tends to infinity, and obtain a singular limit formulation corresponding to $\mu = \infty$. This limit formulation should be compared to the Richards model.

Indeed, the usual assumption made by engineers is that this limit is the Richards model [12,26], which writes

$$\left\{ \begin{array}{lcl} u_t - \mathrm{div}(k_w(u)\nabla p) & = & s_w, \\ u - {p_c}^{-1}(p_{atm} - p) & = & 0, \end{array} \right. \tag{3}$$

where the function ${p_c}^{-1} : \mathbb{R} \to [0,1]$ is defined by ${p_c}^{-1}(p) = 1$ for all $p \leq 0$, ${p_c}^{-1}(p) = u$ such that $p_c(u) = p$, for all $p \in [0, p_c(0)]$, ${p_c}^{-1}(p) = 0$ for all $p \geq p_c(0)$. A consequence of such model is that $u = 1$ if $p \geq p_{atm}$. The existence and uniqueness of the solution of Richards model have been obtained by different authors [6, 23].

In this paper, we prove that the singular limit of the two-phase flow model as $\mu \to \infty$, denoted (u, p), with $u > 0$ a.e., is a solution to a different formulation, namely,

$$\left\{ \begin{array}{lcl} u_t - \mathrm{div}(k_w(u)\nabla p) & = & s_w, \\ u = 1 \quad \text{or} \quad \nabla(p + p_c(u)) = 0 \text{ a.e. in } \Omega \times (0,T). \end{array} \right. \tag{4}$$

Note that a solution of (3) with $u > 0$ a.e. does satisfy (4). The relation between problems (3) and (4) is further discussed in [13, 14].

This paper is organized as follows. In Section 2 we make precise the assumptions on the data and on the nonlinearities, and recall the notion of solution to (1). We then state the existence result of a solution (u_μ, p_μ) which satisfies additional estimates (15), (16), (17) and (18) (these estimates ensure compactness properties as $\mu \to \infty$). We can then directly deduce a theorem of existence of a singular limit as $\mu \to +\infty$, which satisfies in a weak sense the equations (4). In Section 3, we focus on the limit problem (4), of which a more detailed analysis is provided in our work [4]. We justify well-posedness for the particular but important situation where information concerning $\overline{s}$ in the saturation zone $[u = 1]$ is available; this includes e.g. the case of equation without the source term.

Further, the existence results of Section 2 follow from the stability and convergence analysis of the finite volume scheme presented in Section 4. In Section 5 we get *a priori* estimates and deduce the existence of a discrete solution. In Section 6 we convert the estimates into compactness results for sequences of discrete solutions, as the discretization parameters go to zero. Finally, the passage to the limit on the scheme, performed in Section 7, concludes the proof of the convergence of the numerical scheme to a solution of the two-phase flow problem. This solution inherits the desired estimates (15), (16), (17) and (18).

Numerical results in 1D and 2D are given in Section 8. We perform qualitative tests, comparing different values of μ. We exhibit rates of convergence close to h^1 for the saturation with respect the discretization parameter h (which are similar for different values of μ), and the rate of convergence of the saturation in $1/\mu$ to the discretized Richards model.

2. A mathematical formulation of the two-phase flow model

The problem can be formulated mathematically as follows: let Ω be an open bounded subset of $\mathbb{R}^d$ $(d = 2, 3)$, $T \in \mathbb{R}^+$, find $u : \Omega \times (0, T) \to \mathbb{R}$ and $p : \Omega \times (0, T) \to \mathbb{R}$ solution to the following coupled system:

$$u_t - \operatorname{div}(k_w(u)\nabla p) = f_\mu(c)\,\overline{s} - f_\mu(u)\,\underline{s} \text{ on } \Omega \times (0, T), \tag{5}$$

$$(1-u)_t - \operatorname{div}(\mu k_a(u)\nabla(p + p_c(u))) = (1 - f_\mu(c))\,\overline{s} - (1 - f_\mu(u))\,\underline{s} \text{ on } \Omega \times (0, T), \tag{6}$$

with the following Neumann boundary conditions:

$$\nabla p \cdot \mathbf{n} = 0 \text{ on } \partial\Omega \times (0, T), \tag{7}$$

$$\nabla(p + p_c(u)) \cdot \mathbf{n} = 0 \text{ on } \partial\Omega \times (0, T), \tag{8}$$

the following initial condition:

$$u(\cdot, 0) = u_0 \text{ on } \Omega, \tag{9}$$

and such that p satisfies the arbitrary condition (in order to fix the degree of freedom due to the incompressibility of the fluids)

$$\int_\Omega p(x, t)\, dx = 0 \text{ on } (0, T). \tag{10}$$

In this model, u and p are respectively the saturation and the pressure of the water phase, k_w and μk_a are respectively the mobilities of the water phase and the mobility of the non-water phase and p_c is the capillary pressure. We suppose in particular that the physical functions k_w, k_a and p_c only depend on the saturation u of the water phase. Without any exterior action, the state of the system would be stationary because the Neumann boundary conditions (7)-(8) are homogeneous. Here, we suppose that the flow of the water phase in the reservoir Ω is driven by an internal source with rate $f_\mu(c)\,\overline{s}$ and an internal sink with rate $f_\mu(u)\,\underline{s}$ where $\overline{s}$ and $\underline{s}$ represent respectively a source term in injection and at the sinks, c is the saturation of the injected fluid,

f_μ is the fractional flow of the water phase i.e.:

$$f_\mu(x) = \frac{k_w(x)}{M_\mu(x)}, \tag{11}$$

where the function M_μ denotes the total mobility of the two phases, defined by

$$M_\mu(x) = k_w(x) + \mu k_a(x); \tag{12}$$

thus f_μ is a non-decreasing function given by

$$f_\mu(x) = \begin{cases} \left(1 + \mu \frac{k_a(x)}{k_w(x)}\right)^{-1}, & 0 \le x < 1 \\ 1, & x = 1 \end{cases} \tag{13}$$

In the sequel, the following assumptions on the data are referred to as hypothesis (H):

hypothesis (H)

(1) Ω is a polygonal subset of $\mathbb{R}^d$, $d = 2$ or 3,
(2) $T > 0$ is given,
(3) $u_m \in (0, 1)$,
(4) $u_0 \in L^\infty(\Omega)$ and $u_m \le u_0(x) \le 1$ a.e $x \in \Omega$,
(5) $c \in L^\infty(\Omega \times (0,T))$, $u_m \le c(x) \le 1$ a.e. $x \in \Omega$,
(6) $\overline{s} \in L^2(\Omega \times (0,T))$, $\overline{s} \ge 0$, $\underline{s} \in L^2(\Omega \times (0,T))$, $\underline{s} \ge 0$ and $\int_\Omega (\overline{s}(x) - \underline{s}(x))dx = 0$,
(7) $k_w \in \mathcal{C}^0([0,1], \mathbb{R})$, k_w is non-decreasing with $k_w(0) = 0$, $k_w(1) = 1$ and $k_w(u_m) > 0$,
(8) $k_a \in \mathcal{C}^0([0,1], \mathbb{R})$, k_a is non-increasing with $k_a(1) = 0$, $k_a(0) = 1$ and $k_a(s) > 0$ for all $s \in [0, 1)$,
(9) $p_c \in \mathcal{C}^0([u_m, 1], \mathbb{R})$, $p_c \in \mathrm{Lip}_{loc}([u_m, 1), \mathbb{R})$, p_c is strictly decreasing
(10) $\mu \in [1, +\infty)$.

Actually, the absolute continuity of p_c on $[u_m, 1]$ is enough (one has to recast (14) below as a Stiltjes integral), but we stick to Lipschitz continuity assumptions in order to use freely the writing ${p_c}'$.

Let g (known as the Kirchoff transform) and ζ (that we will call "1/2-Kirchhoff" transform) be the functions defined by

$$\forall\, s \in [0,1] \quad g(s) = \int_{u_m}^{s} k_a(\sigma)\, {p_c}'(\sigma) d\sigma, \quad \zeta(s) = \int_{u_m}^{s} \sqrt{k_a(\sigma)}\, {p_c}'(\sigma) d\sigma. \tag{14}$$

Then we have $g, \zeta \in \mathcal{C}^0([u_m, 1], \mathbb{R})$. Our assumptions on k_a, k_w and p_c are slightly more general than those of Henry, Hilhorst and Eymard [14]; in particular, we avoid the Lipshitz continuity assumptions on g and ζ (this forces us, in particular, to use L^1 time translation estimates in the spirit of Andreianov, Bendahmane and Karlsen [3]). Further, unlike in [14], in our technique the Kirchhoff transform $g(u)$ only plays an auxiliary role, whereas the estimates of the space and time translates of u are obtained via controlling the translates of the 1/2-Kirchhoff transform $\zeta(u)$. Notice that, formally, $\nabla g(u) = \sqrt{k_a(u)} \nabla \zeta(u)$, hence $\zeta(u) \in L^2(0,T; H^1(\Omega))$ implies $g(u) \in L^2(0,T; H^1(\Omega))$.

Definition 2.1 (Weak solution)**.** *Under hypothesis (H), a pair (u, p) is a weak solution of Problem (5)-(10) if*

$$\begin{array}{l} u \in L^\infty(\Omega \times (0,T)), \text{ with } 0 \le u(x,t) \le 1 \text{ for a.e. } (x,t) \in \Omega \times (0,T), \\ p \in L^2(0,T; H^1(\Omega)), \\ \zeta(u) \in L^2(0,T; H^1(\Omega)) \quad (\text{thus also } g(u) \in L^2(0,T; H^1(\Omega)), \end{array}$$

if $\int_\Omega p(x,t)dx = 0$ *for a.e.* $t \in (0,T)$*, and for all* $\varphi \in \mathcal{C}^\infty(\mathbb{R}^d \times (0,T])$ *there holds*

$$\int_0^T\int_\Omega [u(x,t)\varphi_t(x,t) - k_w(u(x,t))\nabla p(x,t)\cdot\nabla\varphi(x,t)]dxdt + \int_0^T\int_\Omega [f_\mu(c)\,\overline{s} - f_\mu(u)\,\underline{s}](x,t)\varphi(x,t)dxdt + \int_\Omega u_0(x)\varphi(x,0)dx = 0,$$

$$\int_0^T\int_\Omega [(1-u(x,t))\varphi_t(x,t) - \mu(k_a(u)\nabla p + \nabla g(u))(x,t)\cdot\nabla\varphi(x,t)]dxdt + \int_0^T\int_\Omega [(1-f_\mu(c))\,\overline{s} - (1-f_\mu(u))\,\underline{s}](x,t)\varphi(x,t)dxdt + \int_\Omega (1-u_0(x))\varphi(x,0)dx = 0,$$

In this paper, by constructing a convergent finite volume scheme for Problem (5)-(10) we prove the following existence result for a weak solution with additional estimates related to the μ-dependence of the solutions:

Theorem 2.1 (see also Henry, Hilhorst and Eymard [14])**.** *Let us assume that Hypothesis (H) are satisfied. Then there exists a weak solution* (u_μ, p_μ) *of Problem (5)-(10) in the sense of Definition 2.1 which satisfies the following property:there exists a real value* $C_1 > 0$ *which only depends on* Ω,k_w, k_a, p_c, T, u_m, $\overline{s}$, $\underline{s}$, *and not on* μ*, such that the following inequalities hold:*

$$\int_0^T\int_\Omega k_a(u_\mu)(\nabla p_\mu + \nabla p_c(u_\mu))^2 dxdt \le \frac{C_1}{\mu}, \tag{15}$$

$$\int_0^T\int_\Omega |\nabla p_\mu|^2 dxdt \le C_1\,, \tag{16}$$

$$\int_0^T\int_\Omega |\nabla\zeta(u_\mu)|^2 dxdt \le C_1\,, \tag{17}$$

and for all $\tau \in (0,T)$*, the following estimate holds:*

$$\int_0^{T-\tau}\int_\Omega |\zeta(u_\mu(x,t+\tau)) - \zeta(u_\mu(x,t))|\,dx\,dt \le C_1\,\omega(\tau) \tag{18}$$

where $\omega \in \mathcal{C}^0(\mathbb{R}^+,\mathbb{R}^+)$ *is a modulus of continuity, in particular,* $\omega(0) = 0$.

Remark 2.1. *In absence of a uniqueness result for weak solutions of (5)-(10), the estimates (15),(16),(17) can be viewed as additional constraints on the weak solutions constructed in this paper. As to the translation estimate (18), it can be deduced from the aforementioned ones (see the proof of Theorem 4.1 below).*

As in [14], we readily get the existence of a sequence $(\mu_n)_{n\in\mathbb{N}}$ which tends to infinity such that the sequence $(\zeta(u_{\mu_n}))_{n\in\mathbb{N}}$ converges to a function $\zeta(u) \in L^2(0,T;H^1(\Omega))$ weakly in $L^2(\Omega\times(0,T))$ and a.e. on $\Omega\times(0,T)$; and the sequence $(p_{\mu_n})_{n\in\mathbb{N}}$ weakly converges to $p \in L^2(0,T;H^1(\Omega))$.

Since the sequence $(f_{\mu_n}(u_{\mu_n}))_{n\in\mathbb{N}}$ is bounded in $L^\infty(\Omega\times(0,T))$, a subsequence can be extracted which converges for the weak-* topology of $L^\infty(\Omega\times(0,T))$ to a function θ taking values in $[0,1]$. Moreover, it is not difficult to identify θ to zero a.e. on the set $[u<1] := \{(x,t)\,|\,u(x,t)<1\}$. Indeed, $f_{\mu_n} \to 0$ uniformly on every interval $[0,1-\alpha]$, $\alpha > 0$. The function $\zeta(\cdot)$ being strictly increasing and $(\zeta(u_{\mu_n}))_{n\in\mathbb{N}}$ being strongly convergent to $\zeta(u)$, we deduce that $f_{\mu_n}(u_{\mu_n})$ converge to zero in measure on every set $[u\le 1-\delta] := \{(x,t)\,|\,u(x,t)\le 1-\delta\}$, $\delta > 0$. Therefore we actually have $\theta \equiv \theta \mathbb{1}_{\{1\}}(u)$, where the function $\mathbb{1}_{\{1\}} : [0,1] \to \{0,1\}$ is defined by $\mathbb{1}_{\{1\}}(1) = 1$ and $\mathbb{1}_{\{1\}}(s) = 0$ for all $s \in [0,1)$.

In conclusion, the following theorem, already proved by Henry, Hilhorst and Eymard in [14] under slightly more restrictive assumptions on $p_c(\cdot)$, holds true.

Theorem 2.2. *Assume that Hypothesis (H) are satisfied. Then there exists a triple (u,p,θ) (which is an accumulation point for the sequence $(u_\mu,p_\mu,f_\mu(u_\mu)$ constructed in Theorem 2.1) such that*

$$\begin{array}{l} u\in L^\infty(\Omega\times(0,T)),\ \text{with } 0\le u(x,t)\le 1 \text{ for a.e. } (x,t)\in\Omega\times(0,T),\\ \theta\in L^\infty(\Omega\times(0,T)),\ \text{with } 0\le \theta(x,t)\le 1 \text{ for a.e. } (x,t)\in\Omega\times(0,T),\\ p\in L^2(0,T;H^1(\Omega)),\\ \zeta(u)\in L^2(0,T;H^1(\Omega)),\end{array}$$

and for all $\varphi\in\mathcal{C}^\infty(\mathbb{R}^d\times\mathbb{R})\quad$ with $\quad\varphi(\cdot,T)=0$,

$$\begin{array}{l}\displaystyle\int_0^T\int_\Omega [u(x,t)\varphi_t(x,t)-k_w(u(x,t))\nabla p(x,t)\cdot\nabla\varphi(x,t)]dxdt+\\ \displaystyle\int_0^T\int_\Omega [\mathbb{1}_{\{1\}}(c)\,\overline{s}-\theta\mathbb{1}_{\{1\}}(u)\,\underline{s}](x,t)\varphi(x,t)dxdt+\int_\Omega u_0(x)\varphi(x,0)dx=0;\end{array}$$

moreover, $k_a(u)(\nabla p+\nabla p_c(u))=0$ a.e. on $\Omega\times(0,T)\quad$ and $\displaystyle\int_\Omega p(x,t)dx=0$ for a.e. $t\in(0,T)$.

3. The singular limit of the two-phase flow equation

The result of Theorem 2.2 shows the convergence, as $\mu\to\infty$, of the couple (u_μ,p_μ) solving the two-phase problem (5)-(10) to a (weak) solution (u,p) of problem (4) with the source

$$s_w=\mathbb{1}_{\{1\}}(c)\,\overline{s}-\theta\mathbb{1}_{\{1\}}(u)\,\underline{s}\ \text{ where } \theta \text{ is some } [0,1] \text{ valued function.} \tag{19}$$

Remark 3.1. *It should be stressed that, either we are not able to identify θ in general (except by saying that θ can be calculated from (4),(19) as soon as u and p are known), in some practical situations the relation $\theta\equiv 1$ holds true (recall that θ should only be defined on the subset $[u=1]:=\{(x,t)\,|\,u(t,x)=1\}$ of $\Omega\times(0,T)$). Indeed, it is sometimes observed that the family $(u_\mu)_\mu$ is monotone non-increasing in the regions where $u=\lim_{\mu\to\infty}u_\mu$ is close to the saturation value 1; in particular, for μ large one observes $u_\mu=1$ on $[u=1]$. It is clear from the definition of θ that one must have $\theta=1$ a.e. on $[u=1]$ in these cases.*

We refer to Henry, Hilhorst and Eymard [14] for a comparative discussion of (4),(19) and the Richards model (3). In this section, we establish a uniqueness and continuous dependence result for the limit problem (4),(19) under the additional restrictive assumption that the source term $\overline{s}\mathbb{1}_{\{1\}}(c)$ is zero in the saturated zone $[u=1]$. Indeed, it turns out that, although p and θ may not be unique, the saturation u in a triple (u,p,θ) solving (4),(19) is uniquely defined. More precisely, we prove the order-preserving L^1 contraction principle for (4),(19):

Theorem 3.1. *Let the assumptions (**H**) be verified. Let (u,p,θ) and $(\hat u,\hat p,\hat\theta)$ be two solutions of the quasi-Richards model (4),(19) corresponding to data u_0, $\hat u_0$ and the same source and sink terms $\overline{s}$, $\underline{s}$. Assume in addition that we consider solutions such that $\overline{s}\mathbb{1}_{\{1\}}(c)\mathbb{1}_{\{1\}}(u)=0$ a.e. on Ω (a sufficient, u-independent condition is that there is no water injection, i.e., $\overline{s}\mathbb{1}_{\{1\}}(c)=0$ a.e. on Ω).*

Then for a.e. $t\in(0,T)$,

$$\int_\Omega (u(\cdot,t)-\hat u(\cdot,t))^+\le\int_\Omega (u(\cdot,t)-\hat u(\cdot,t))^+. \tag{20}$$

In particular, for all $u_m>0$, for all $[u_m,1]$-valued measurable initial datum u_0 there exists a unique function u such that (u,p,θ) solves problem (4),(19); the function u depends monotonically and continuously in $L^\infty(0,T;L^1(\Omega))$ on the initial datum u_0.

Before turning to the proof, we make precise the definition of a solution to the quasi-Richards equation and give some notation and a key preparatory lemma.

As in Theorem 2.2, by a weak solution of (4) we will understand a couple (u,p) formed by a $[u_m,1]$-valued measurable function u with $\zeta(u)\in L^2(0,T;H^1(\Omega))$ and by $p\in L^2(0,T;H^1(\Omega))$ normalized by (10); we require

that the first equation in (4) (supplemented with the homogeneous Neumann boundary condition) hold in the sense of distributions (or, equivalently, in $L^2(0,T;(H^1(\Omega))^*)$), and that the second equation hold in the sense that the function $k_a(u)\nabla p_c(u) := \sqrt{k_a(u)}\nabla\zeta(u)$ coincides with the function $-k_a(u)\nabla p$ a.e. on $\Omega\times(0,T)$ (recall that $k_a(u)=0$ if and only if $u=1$; in particular, we can state that $\nabla p = \nabla p_c(u)$ a.e. on $[u<1]$).

Remark 3.2. *The function k_a, which is formally absent from the quasi-Richards formulation, appears in the above definition as the multiplicative term of the constraint $k_a(u)\nabla(p+p_c(u))=0$; yet $k_a(u)=0$ if and only if $u=1$. More importantly, dependence of the notion of solution on the profile of k_a appears implicitly in the integrability constraint $\nabla\zeta(u)\in L^2(\Omega\times(0,T))$ (recall that $\zeta' := \sqrt{k_a}p_c'$). In [4] we give an intrinsic weak formulation from which k_a is discarded. In particular, the fact of being a solution to the quasi-Richards equation does not depend on the profile of k_a as soon as assumptions (**H**) are fulfilled.*

In order to prove Theorem 2.2, we use the technique of renormalized solutions, following the ideas of [24,25]. For the sake of simplicity, let us make the assumption that $-p_c'\geq\delta>0$ (the assumption is verified in several models one finds in the literature); see [4] for the general case. Then, according to (**H**)(8), for all $\alpha>0$, ζ^{-1} is a Lipschitz function on $[0,1-\alpha]$.

Consider a sequence $(T_n)_{n\in\mathbb{N}}$ of functions on $[0,1]$ with the following properties:

$$T_n\in\mathcal{C}^1([0,1]),\ \ T_n'\leq 0,\ \ T_n|_{[0,1-\frac{1}{n}]}\equiv 1,\ \ T_n|_{[1-\frac{1}{2n},1]}=0.$$

Then the following properties are obvious:

$$b_n(z) := \int_0^z T_n(s)\,ds \ \text{ tends to the identity function;} \tag{21}$$

$$c_n(z) = \int_0^z (1-T_n(s))\,ds \ \text{ tends to the zero function.} \tag{22}$$

We define in addition the functions

$$\varphi_n(z) = \int_0^z k_w(s)T_n(s)(-p_c'(s))\,ds \quad\text{and}\quad \psi_n(z)=\int_0^z\sqrt{k_w(s)(-T_n)'(s)(-p_c'(s))}\,ds.$$

The function T_n is supported in the interval $[0,1-\frac{1}{2n}]$, moreover, k_a is separated from zero and p_c' is bounded on this interval (according to assumptions (**H**)(8),(9)); thus we see that both functions $\varphi_n\circ\zeta^{-1}$ and $\psi_n\circ\zeta^{-1}$ are Lipschitz continuous. Therefore we have $\varphi_n(u),\psi_n(u)\in L^2(0,T;H^1(\Omega))$ as soon as $\zeta(u)\in L^2(0,T;H^1(\Omega))$, and the following statement makes sense.

Lemma 3.1. *Assume that u is a solution of (4) in the above sense. Then, with the above notation, there holds in $\mathcal{D}'(\overline{\Omega}\times[0,T))$ the renormalized formulation*

$$b_n(u)_t - \Delta\varphi_n(u) - |\nabla\psi_n(u)|^2 = s_w T_n(u) \tag{23}$$

with $\nabla\varphi_n(u)\cdot\mathbf{n}=0$ on $\partial\Omega\times(0,T)$ and $b_n(u)|_{t=0}=b_n(u_0)$.

Moreover, for all $t\in(0,T)$ there holds

$$\lim_{n\to\infty}\int_0^t\int_\Omega|\nabla\psi_n(u)|^2 = \int_0^t\int_\Omega s_w\mathbb{1}_{\{1\}}(u)\equiv\int_0^t\int_{[u=1]}s_w. \tag{24}$$

Proof. By the definition of a solution to (4) and because we have assumed that ζ^{-1} is a Lipschitz function, the composition of $u\in L^2(0,T;H^1(\Omega))$ by the Lipschitz function T_n is an admissible test function in the weak formulation of the first equation of (4). More precisely, having $u_t\in L^2(0,T;(H^1(\Omega))^*)$ and $T_n(u)\in$

$L^2(0,T;H^1(\Omega))$, we can apply the integration-by-parts formula in time (see [2, 23]) and get for all $\xi \in \mathcal{D}'(\overline{\Omega}\times[0,T))$,

$$\int_0^T\int_\Omega \Big(-b_n(u)\xi_t + k_w(u)\nabla p\cdot\nabla(T_n(u)\xi)\Big) = \int_0^T\int_\Omega s_w T_n(u)\xi + \int_\Omega b_n(u_0)\xi(\cdot,0). \tag{25}$$

In the second term in (25), we split the integral over the two sets $[u<1]$ and $[u=1]$. A.e. on $[u=1]$, we have $\nabla\zeta(u)=0$ ($\zeta(u)$ being a Sobolev function); then, according to our assumptions on T_n, also $\nabla(T_n(u)\xi)=0$ a.e. on $[u=1]$. Further, the second equation in (4) allows to replace, a.e. on $[u<1]$, the function ∇p by the function $-\nabla p_c(u)$ (interpreted e.g. as $-\frac{1}{\sqrt{k_a(u)}}\nabla\zeta(u)$; recall that we work on $[u<1]$, and even on $[0,1-\frac{1}{2n}]$ according to the properties of T_n). Developing $\nabla(T_n(u)\xi)$, we finally write

$$\int_0^T\int_\Omega k_w(u)\nabla p\cdot\nabla(T_n(u)\xi) = \int_0^T\int_{[u<1]}\Big(\frac{k_w(u)}{\sqrt{k_a(u)}}T_n(u)(-\nabla\zeta(u))\cdot\nabla\xi + \frac{k_w(u)}{\sqrt{k_a(u)}}(-\nabla\zeta(u))\cdot\nabla T_n(u)\,\xi\Big). \tag{26}$$

Developing the definitions of ζ, φ_n and ψ_n, we see that in the right-hand side above can be rewritten as

$$\int_0^t\int_{[u<1]} \nabla\varphi_n(u)\cdot\nabla\xi - |\nabla\psi_n(u)|^2\,\xi$$

(both $\varphi_n(u)$ and $\psi_n(u)$ being $L^2(0,T;H^1(\Omega))$ functions). Then, the integrand above being zero a.e. on $[u=1]$, we end up with the claim (23).

Similarly, replacing $T_n(u)$ by $(1-T_n(u))$ in the above argument, we find

$$\int_0^T\int_\Omega\Big(-c_n(u)\xi_t + |\nabla\psi_n(u)|^2\xi + k_w(u)(1-T_n(u))\nabla p\cdot\nabla\xi\Big) = \int_0^T\int_\Omega s_w(1-T_n(u))\xi + \int_\Omega c_n(u_0)\xi(\cdot,0). \tag{27}$$

Here we take for ξ a constant in x function converging to $\mathbb{1}_{[0,t]}$; we let $n\to\infty$, and find $(1-T_n(u))\xi \to \mathbb{1}_{\{1\}}(u)$, $(1-T_n(u))\nabla\xi = 0$ and $c_n(u)\to 0$. Hence the claim (24) follows. □

We are now in a position to prove Theorem 3.1.

Proof. Consider u, $\hat{u}$ two solutions of (4),(19). For each of those, write the renormalized formulation (23) and consider it as degenerate elliptic-parabolic equation with a source term:

$$b_n(u)_t - \Delta\varphi_n(u) = f,\ \ f = |\nabla\psi_n(u)|^2 + s_w T_n(u)$$
$$b_n(\hat{u})_t - \Delta\varphi_n(\hat{u}) = \hat{f},\ \ \hat{f} = |\nabla\psi_n(\hat{u})|^2 + \hat{s}_w T_n(\hat{u})$$

and the same initial condition $b_n(u)|_{t=0} = b_n(u_0) = b_n(\hat{u})|_{t=0}$. Then the technique of doubling of the time variable due to Otto [23] yields the L^1 order-preserving contraction inequality:

$$\text{for a.e. } t\in(0,T)\quad \int_\Omega (b_n(u)-b_n(\hat{u}))^+(t,\cdot) - \int_\Omega (b_n(u_0)-b_n(\hat{u}_0))^+ \le \int_0^t\int_\Omega \operatorname{sgn}^+(u-\hat{u})(f-\hat{f}). \tag{28}$$

At the limit $n\to\infty$, thanks to (21) and to the fact that $T_n\to\mathbb{1}_{[0,1)}$, we find for a.e. $t\in(0,T)$

$$\int_\Omega (u-\hat{u})^+(t,\cdot) - \int_\Omega (u_0-\hat{u}_0)^+ \le \liminf_{n\to\infty}\int_0^t\int_{[u>\hat{u}]}(|\nabla\psi_n(u)|^2-|\nabla\psi_n(\hat{u})|^2) + \int_0^t\int_{[u>\hat{u}]}(s_w\mathbb{1}_{[u<1]} - \hat{s}_w\mathbb{1}_{[\hat{u}<1]}). \tag{29}$$

Gathering (29), (19) and (24), dropping the non-positive term containing θ, we get

$$\int_\Omega (u-\hat{u})^+(t,\cdot) - \int_\Omega (u_0-\hat{u}_0)^+ \le \int_0^t\int_{[c=1,u=\hat{u}=1]} \overline{s},$$

of which the right-hand side is zero due to the assumption of the theorem. □

Remark 3.3. *Based on this result, one easily concludes that in the no-injection case* $\overline{s}\mathbb{1}_{\{1\}}(c) = 0$ *Richards and quasi-Richards models coincide, see [4]. Thus convergence of the two-phase model to the Richards equation holds true, in this case. In general, we expect that the two models can be different in what concerns the "air trapping" phenomenon. In general, the question of uniqueness of the quasi-Richards model is open.*

4. The finite volume scheme

4.1. Finite volume definitions and notations

We mainly follow here the notations of [11].

Definition 4.1 (Admissible mesh of Ω)**.** *An admissible mesh* $\mathcal{T}_h$ *of* Ω *is given by a set of open bounded polygonal convex subsets of* Ω *called control volumes and a family of points (the "centers" of control volumes) satisfying the following properties:*

1. *The closure of the union of all control volumes is* $\bar{\Omega}$*. We denote by* m_K *the measure of* K*.*
2. *For any* $(K,L) \in {\mathcal{T}_h}^2$ *with* $K \neq L$*, then* $K \cap L = \emptyset$*. One denotes by* $\mathcal{E}_h \subset {\mathcal{T}_h}^2$ *the set of* (K,L) *such that the* $(d-1)$*-dimensional Lebesgue measure of* $\bar{K} \cap \bar{L}$ *is positive. For* $(K,L) \in \mathcal{E}_h$*, one denotes* $K|L = \bar{K} \cap \bar{L}$ *and* $m(K|L)$ *the* $(d-1)$*-dimensional Lebesgue measure of* $K|L$*.*
3. *For any* $K \in \mathcal{T}_h$*, one defines* $\mathcal{N}_K = \{L \in \mathcal{T}_h, (K,L) \in \mathcal{E}_h\}$ *and one assumes that* $\partial K = \bar{K} \backslash K = (\bar{K} \cap \partial\Omega) \cup \cup_{L \in \mathcal{N}_K} K|L$*.*
4. *The family of points* $(x_K)_{K \in \mathcal{T}_h}$ *is such that* $x_K \in K$ *(for all* $K \in \mathcal{T}_h$*)*[1] *and, if* $L \in \mathcal{N}_K$*, it is assumed that the straight line* (x_K, x_L) *is orthogonal to* $K|L$*.*
5. *We set* $d_{K|L} = \|\overrightarrow{x_K x_L}\|$ *and* $\tau_{K|L} = \frac{m(K|L)}{d_{K|L}}$*, that is sometimes called the "transmissibility" through* $K|L$*. We set* $\vec{n}_{K,L}$ *the unit normal vector to* $K|L$ *pointing from* K *to* L*; i.e.,* $\vec{n}_{K,L} = \frac{\overrightarrow{x_K x_L}}{d_{K|L}}$*.*
6. *Finally, given an interface* $K|L$ *and the associated neighbour centers* x_K, x_L *we define the diamond* $D_{K|L}$ *as the convex hull of* x_K*,* x_L *and* $K|L$*, so that the diamonds are disjoint and cover* Ω *up to a neighbourhood of* $\partial\Omega$*.*

The problem under consideration is time-dependent, hence we also need to discretize the time interval $(0,T)$.

Definition 4.2 (Time discretization)**.** *A time discretization of* $(0,T)$ *is given by an integer value* N *and by a strictly increasing sequence of real values* $(t^n)_{n \in [\![0,N+1]\!]}$ *with* $t^0 = 0$ *and* $t^{N+1} = T$*. The time steps are then defined by* $\delta t^n = t^{n+1} - t^n$*, for* $n \in [\![0,N]\!]$*.*

We may then define a discretization of the whole domain $\Omega \times (0,T)$ in the following way:

Definition 4.3 (Discretization of $\Omega \times (0,T)$)**.** *A finite volume discretization* $\mathcal{D}$ *of* $\Omega \times (0,T)$ *is a family*

$$\mathcal{D} = (\mathcal{T}_h, \mathcal{E}_h, (x_K)_{K \in \mathcal{T}_h}, N, (t^n)_{n \in [\![0,N]\!]}),$$

where $\mathcal{T}_h$*,* $\mathcal{E}_h$*,* $(x_K)_{K \in \mathcal{T}_h}$ *are described in Definition 4.1 of admissible mesh of* Ω*, and* N*,* $(t^n)_{n \in [\![0,N+1]\!]}$ *is a time discretization of* $(0,T)$ *in the sense of Definition 4.2. One then sets*

$$\text{size}(\mathcal{D}) = \max(\text{size}(\mathcal{T}_h), (\delta t^n)_{n \in [\![0,N+1]\!]}), \quad \textit{where } \text{size}(\mathcal{T}_h) = \sup\{diam(K), K \in \mathcal{T}_h\}.$$

Definition 4.4 (Discrete functions and finite differencing)**.** *Let* $\mathcal{D}$ *be a discretization of* $\Omega \times (0,T)$*, we denote by* $X_\mathcal{D}$ *the set of the discrete functions associated to* $\mathcal{D}$ *i.e.* $X_\mathcal{D} = \mathbb{R}^{\mathcal{T}_h \times [\![0,N]\!]}$*. An element of* $X_\mathcal{D}$ *is denoted with capital letters and the index* $\mathcal{D}$ *(*$U_\mathcal{D}$ *or* $P_\mathcal{D}$ *for instance) and the value at point* (K,n) *with the index* K *and the*

[1] this constraint can be relaxed, i.e. for triangular meshes in 2D satisfying the so-called Delaunay condition, we can pick for x_K the center of the circonscribed circle of K (see [11]).

upper index n (U_K^n or P_K^{n+1} for instance). To a discrete function $U_\mathcal{D}$ corresponds an approximate function $u_\mathcal{D}$ defined almost everywhere on $\Omega \times (0,T)$ by:

$$u_\mathcal{D}(x,t) = U_K^{n+1} \text{ for all } (x,t) \in K \times (t^n, t^{n+1}].$$

The subscript h denotes a quantity defined per diamond $D_{K|L}$, e.g. $\nabla_h U_\mathcal{D}$ (the discrete gradient of $U_\mathcal{D}$) is the piecewise constant function taking the value $\nabla_{K|L} U_\mathcal{D}$ of the discrete gradient on the interface $K|L$ of the discrete function $U_\mathcal{D}$.

For any function $f_\mu : \mathbb{R} \mapsto \mathbb{R}$, $f_\mu(U_\mathcal{D})$ denotes the discrete function $(K,n) \mapsto f_\mu(U_K^{n+1})$. If $L \in \mathcal{N}_K$, and $U_\mathcal{D}$ is a discrete function, we denote by $\delta_{K,L}^{n+1}(U) = U_L^{n+1} - U_K^{n+1}$. For example, $\delta_{K,L}^{n+1}(f_\mu(U)) = f_\mu(U_L^{n+1}) - f_\mu(U_K^{n+1})$. At the same time, the notation $\mathfrak{d}_{K,L}^{n+1}[g(U)]$ will denote a discretization different from $g(U_L^{n+1}) - g(U_K^{n+1})$, see (37) below.

Let us now give the regularity property of the discretization mesh we need in order to use the discrete Poincaré inequality of [11]. [2]

Definition 4.5 (Regularity of the mesh)**.** *Let $\xi > 0$. A discretization $\mathcal{D}$ of $\Omega \times (0,T)$ is ξ-regular if*

$$\forall K \in \mathcal{T}_h, \sum_{L \in \mathcal{N}_K} m(K|L) d_{K|L} \leq \xi m_K \tag{30}$$

We will therefore require that the family of discretizations considered be ξ-regular with $\mathcal{D}$-independent ξ.

4.2. The coupled finite volume scheme

The finite volume scheme is obtained by writing the balance equations of the fluxes on each control volume. Let $\mathcal{D}$ be a discretization of $\Omega \times (0,T)$. Let us integrate equations (5)-(6) over each control volume K. By using the Green-Riemann formula, if Φ is a vector field, the integral of $\mathrm{div}(\Phi)$ on a control volume K is equal to the sum of the normal fluxes of Φ on the edges. Here we apply this formula to $\Phi_1 = k_w(u)\nabla p$ and $\Phi_2 = \mu(k_a(u)\nabla p + \nabla g(u))$. The resulting equation is discretized with a time implicit finite difference scheme; the normal gradients are discretized with a centered finite difference scheme. If we denote by $U_\mathcal{D} = \{U_K^n\}_{n \in [\![0,N+1]\!], K \in \mathcal{T}_h}$ and $P_\mathcal{D} = \{P_K^n\}_{n \in [\![1,N+1]\!], K \in \mathcal{T}_h}$ the discrete unknowns corresponding to u and p, the finite volume scheme that we obtain is the following set of equations:

$$U_K^0 = \frac{1}{m_K} \int_K u_0(x)dx, \text{ for all } K \in \mathcal{T}_h, \tag{31}$$

for all $(K,n) \in \mathcal{T}_h \times [\![0,N]\!]$,

$$\begin{aligned} &\frac{U_K^{n+1} - U_K^n}{\delta t^n} m_K - \sum_{L \in \mathcal{N}_K} \tau_{K|L} f_\mu(U_{K|L}^{n+1}) M_\mu(\bar{U}_{K|L}^{n+1}) \delta_{K,L}^{n+1}(P_\mathcal{D}) \\ &- m_K \left(f_\mu(c_K^{n+1})\, \overline{s}_K^{n+1} - f_\mu(U_K^{n+1})\, \underline{s}_K^{n+1} \right) = 0, \end{aligned} \tag{32}$$

$$\begin{aligned} &\frac{(1 - U_K^{n+1}) - (1 - U_K^n)}{\delta t^n} m_K \\ &- \sum_{L \in \mathcal{N}_K} \tau_{K|L} (1 - f_\mu(U_{K|L}^{n+1})) M_\mu(\bar{U}_{K|L}^{n+1}) \delta_{K,L}^{n+1}(P_\mathcal{D}) - \mu \sum_{L \in \mathcal{N}_K} \tau_{K|L} \mathfrak{d}_{K,L}^{n+1}[g(U_\mathcal{D})] \\ &- m_K \left((1 - f_\mu(c_K^{n+1}))\, \overline{s}_K^{n+1} - (1 - f_\mu(U_K^{n+1}))\, \underline{s}_K^{n+1} \right) = 0, \end{aligned} \tag{33}$$

and

$$\sum_{K \in \mathcal{T}_h} m_K P_K^{n+1} = 0, \text{ for all } n \in [\![0,N]\!], \tag{34}$$

[2] We guess that restriction (30) can be avoided. Indeed, at least for the case of Dirichlet boundary conditions, this regularity assumption is unnecessary for the Poincaré inequality to hold; see [5].

where

- c_K^{n+1} is the mean value of c over the time-space cell $K \times (t^n, t^{n+1})$,
- $\overline{s}_K^{n+1}$ and $\underline{s}_K^{n+1}$ denote the mean values of $\overline{s}$ and $\underline{s}$ over the time-space cell $K \times (t^n, t^{n+1})$,
- $U_{K|L}^{n+1}$ denotes the upwind value of U on the interface $K|L$, which is defined by:

$$U_{K|L}^{n+1} = \begin{cases} U_L^{n+1} \text{ if } \delta_{K,L}^{n+1}(P_{\mathcal{D}}) \geq 0, \\ U_K^{n+1} \text{ otherwise,} \end{cases} \tag{35}$$

- $\bar{U}_{K|L}^{n+1} \in [\min(U_K^{n+1}, U_L^{n+1}), \max(U_K^{n+1}, U_L^{n+1})]$ denotes a value on the interface $K|L$, which is defined by:

$$\sqrt{k_a(\bar{U}_{K|L}^{n+1})}\, \delta_{K,L}^{n+1}(p_c(U_{\mathcal{D}})) = \delta_{K,L}^{n+1}(\zeta(U_{\mathcal{D}})), \text{ i.e.,}$$

$$k_a(\bar{U}_{K|L}^{n+1}) := \left(\frac{\zeta(U_L^{n+1}) - \zeta(U_K^{n+1})}{p_c(U_L^{n+1}) - p_c(U_K^{n+1})} \right)^2 \text{ unless } U_L^{n+1} = U_K^{n+1}; \tag{36}$$

then we make the following choice:

$$\mathfrak{d}_{K,L}^{n+1}[g(U_{\mathcal{D}})] := \sqrt{k_a(\bar{U}_{K|L}^{n+1})}\, \delta_{K,L}^{n+1}(\zeta(U_{\mathcal{D}})), \text{ i.e.,}$$

$$\mathfrak{d}_{K,L}^{n+1}[g(U_{\mathcal{D}})] := \frac{(\zeta(U_L^{n+1}) - \zeta(U_K^{n+1}))^2}{p_c(U_L^{n+1}) - p_c(U_K^{n+1})} \text{ unless } U_L^{n+1} = U_K^{n+1}. \tag{37}$$

In the case $U_K^{n+1} = U_L^{n+1}$ we put $\bar{U}_{K|L}^{n+1} = U_K^{n+1} = U_L^{n+1}$.

Remark 4.1. *While it is clear that nonlinear chain rules, like e.g. $g'(u) = \sqrt{k_a(u)}\zeta'(u) = k_a(u)p_c{}'(u)$, cannot be exactly preserved at the discretization level, the tricky choice (37) preserves some of the chain rule structure. In fact, this specific choice allows us to get estimates on the discrete gradient of $\zeta(U_{\mathcal{D}})$.*

We show below (see Proposition 5.3)) that there exists at least a solution to this scheme.

Remark 4.2. *The discretization scheme yields a nonlinear system of equations which is solved in practice by the Newton method. Numerical experiments show that if the time step is adequately chosen, the Newton procedure converges with a small number of iterations. Hence, although is implicit, this scheme is cheaper than the analogous explicit one.*

We may now state the main convergence result wich readly implies the existence result of Theorem 2.1.

Theorem 4.1. *Assume that hypothesis (H) are satisfied. Let $\{\mathcal{D}_m\}_{m\in\mathbb{N}}$ be a sequence of discretizations of $\Omega \times (0,T)$ in the sense of Definition 4.3, such that there exists $\xi > 0$ with $\mathcal{D}_m$ ξ-regular in the sense of Definition 4.5, and such that $\lim_{m\to\infty} \text{size}(\mathcal{D}_m) = 0$. Let $(u_{\mathcal{D}_m}, p_{\mathcal{D}_m})$ be the approximate solutions corresponding to $\mathcal{D}_m$.*

Then exists a subsequence again denoted by $(u_{\mathcal{D}_m}, p_{\mathcal{D}_m})$ which converges to a weak solution (u,p) of Problem (5)-(10) in the sense of Definition 2.1; moreover, this limit solution satisfies the properties (15)-(18).

For the exact list of the convergence properties, we refer to Section 6.

5. Discrete a priori estimates and existence

In this section, we develop the first part of the proof of Theorem 4.1. Since these estimates mimic the continuous ones, we give a sketch of the continuous case before the proof of the discrete counterpart.

5.1. **The maximum principle**

Let us show here that the scheme implies the satisfaction of the maximum principle.

Proposition 5.1 (Maximum principle)**.** *Assume that hypothesis (H) are fulfilled and $(U_\mathcal{D}, P_\mathcal{D})$ is a solution of the finite volume scheme (31)-(34) then we have the following maximum principle :*

$$u_m \le U_K^n \le 1, \qquad \forall K \in \mathcal{T}_h, \forall\, n \in [\![0, N+1]\!]. \tag{38}$$

Proof. Let us prove the property by induction on n. It is true for $n = 0$, using hypothesis (H). We assume that it is true at the level n. We get, from the sum of (32) and (33),

$$-\sum_{L\in\mathcal{N}_K} \tau_{K|L} M_\mu(\bar U_{K|L}^{n+1}) \delta_{K,L}^{n+1}(P) - \mu \sum_{L\in\mathcal{N}_K} \tau_{K|L} \mathfrak{d}_{K,L}^{n+1}[g(U)] \;=\; m_K(\,\overline{s}_K^{n+1} - \underline{s}_K^{n+1}). \tag{39}$$

Here and in the sequel, we drop the subscript $\mathcal{D}$ for $U_\mathcal{D}, P_\mathcal{D}$, the discretization $\mathcal{D}$ being fixed. Multiplying (39) by $f_\mu(U_K^{n+1})$ and subtracting from (32) gives

$$\begin{aligned}&\frac{U_K^{n+1}-U_K^n}{\delta t^n} m_K - \sum_{L\in\mathcal{N}_K} \Big(f_\mu(U_{K|L}^{n+1}) - f_\mu(U_K^{n+1})\Big)\, \tau_{K|L} M_\mu(\bar U_{K|L}^{n+1}) \delta_{K,L}^{n+1}(P)\\ &+\mu f_\mu(U_K^{n+1}) \sum_{L\in\mathcal{N}_K} \tau_{K|L} \mathfrak{d}_{K,L}^{n+1}[g(U)] \;=\; m_K \left(f_\mu(c_K^{n+1}) - f_\mu(U_K^{n+1})\right) \overline{s}_K^{n+1},\end{aligned}$$

which gives, thanks to (35),

$$\begin{aligned}&\frac{U_K^{n+1}-U_K^n}{\delta t^n} m_K + \sum_{L\in\mathcal{N}_K} \big(f_\mu(U_K^{n+1}) - f_\mu(U_L^{n+1})\big)\, \tau_{K|L} M_\mu(\bar U_{K|L}^{n+1}) (\delta_{K,L}^{n+1}(P))^+\\ &+\mu f_\mu(U_K^{n+1}) \sum_{L\in\mathcal{N}_K} \tau_{K|L} \mathfrak{d}_{K,L}^{n+1}[g(U)] \;=\; m_K \left(f_\mu(c_K^{n+1}) - f_\mu(U_K^{n+1})\right) \overline{s}_K^{n+1},\end{aligned} \tag{40}$$

This yields

$$\begin{aligned}&\frac{U_K^{n+1}-U_K^n}{\delta t^n} m_K + \sum_{L\in\mathcal{N}_K} \big(f_\mu(U_K^{n+1}) - f_\mu(U_L^{n+1})\big)\, \tau_{K|L} M_\mu(\bar U_{K|L}^{n+1}) (\delta_{K,L}^{n+1}(P))^+\\ &-\mu f_\mu(U_K^{n+1}) \sum_{L\in\mathcal{N}_K} \tau_{K|L} \sqrt{k_a(\bar U_{K|L}^{n+1})} (\zeta(U_K^{n+1}) - \zeta(U_L^{n+1})) + m_K \left(f_\mu(U_K^{n+1}) - f_\mu(c_K^{n+1})\right) \overline{s}_K^{n+1} \;=\; 0.\end{aligned} \tag{41}$$

Since all the values U_K^n and c_K^{n+1} belong to $[u_m, 1]$, and since f and $-g$ are non decreasing, we then get from (41) that the maximum value of U_K^{n+1} over $K \in \mathcal{T}_h$ is lower than 1 and the minimum value is greater than u_m. □

5.2. **Estimates on the discrete gradients**

We can state the following property.

Proposition 5.2 (Estimates on the discrete gradients)**.** *Under hypothesis (H), let $\mathcal{D}$ be a finite volume discretization of $\Omega \times (0,T)$ in the sense and with the notations of Definition 4.3; let $(U_\mathcal{D}, P_\mathcal{D})$ be a solution of the finite volume scheme (31)-(34) . Then there exists $C_2 > 0$, which only depends on k_w, k_a, p_c, Ω, T, u_m, $\|\,\overline{s}\|_{L^2(\Omega\times(0,T))}$, $\|\,\underline{s}\|_{L^2(\Omega\times(0,T))}$, and nor on $\mathcal{D}$ neither on μ, such that the following discrete $L^2(0,T;H^1(\Omega))$ estimates hold:*

$$\sum_{n=0}^{N} \delta t^n \sum_{K\in\mathcal{T}_h} \sum_{L\in\mathcal{N}_K} \tau_{K|L} (\delta_{K,L}^{n+1}(P_\mathcal{D}))^2 \le C_2\,, \tag{42}$$

$$\sum_{n=0}^{N}\delta t^n\sum_{K\in\mathcal{T}_h}\sum_{L\in\mathcal{N}_K}\tau_{K|L}k_a(\bar U_{K|L}^{n+1})\left(\delta_{K,L}^{n+1}(P_{\mathcal{D}}+p_c(U_{\mathcal{D}}))\right)^2\le\frac{C_2}{\mu},\tag{43}$$

$$\sum_{n=0}^{N}\delta t^n\sum_{K\in\mathcal{T}_h}\sum_{L\in\mathcal{N}_K}\tau_{K|L}(\delta_{K,L}^{n+1}(\zeta(U_{\mathcal{D}})))^2\le C_2\,.\tag{44}$$

Proof. In the sequel, we drop the subscript $\mathcal{D}$ for $U_{\mathcal{D}}, P_{\mathcal{D}}$, the discretization $\mathcal{D}$ being fixed.

Let us introduce the following quantities:

$$\begin{array}{rclcrcl} A_{KL}^{n+1} &=& \tau_{K|L}M_\mu(\bar U_{K|L}^{n+1}), && B_{KL}^{n+1} &=& A_{KL}^{n+1}f_\mu(U_{K|L}^{n+1}),\\ s_K^{n+1} &=& \overline{s}_K^{n+1}-\underline{s}_K^{n+1}, && \sigma_K^{n+1} &=& f_\mu(c_K^{n+1})\,\overline{s}_K^{n+1}-f_\mu(U_K^{n+1})\,\underline{s}_K^{n+1}; \end{array}\tag{45}$$

we will write $\delta_K^{n+1,n}(\beta(U))=\beta(U_K^{n+1})-\beta(U_K^n)$ where $\beta:\mathbb{R}\longrightarrow\mathbb{R}$ is a continuous function.

With this notation, the scheme can be rewritten as

$$\frac{m_K}{\delta t^n}\delta_K^{n+1,n}(U)-\sum_{L\in\mathcal{N}_K}B_{KL}^{n+1}\delta_{K,L}^{n+1}(P)=m_K\sigma_K^{n+1},\tag{46}$$

$$-\frac{m_K}{\delta t^n}\delta_K^{n+1,n}(U)-\sum_{L\in\mathcal{N}_K}(A_{KL}^{n+1}-B_{KL}^{n+1})\delta_{K,L}^{n+1}(P)-\mu\sum_{L\in\mathcal{N}_K}\tau_{K|L}\mathfrak{d}_{K,L}^{n+1}[g(U)]=m_K\left(s_K^{n+1}-\sigma_K^{n+1}\right).\tag{47}$$

Note the two following facts. First, the summation-by-parts procedure yields:

$$\text{If }\left\{\begin{array}{l}\forall K|L\in{\mathcal{E}_h}^{int}\ \ \theta_{KL}=-\theta_{LK}\\ \forall K|L\in{\mathcal{E}_h}^{ext}\ \ \theta_{KL}=0\end{array}\right.\quad\text{then }\forall\,W\in\mathbb{R}^{\mathcal{T}_h},\ \sum_{K\in\mathcal{T}_h}W_K\sum_{L\in\mathcal{N}_K}\theta_{KL}=-\sum_{K|L\in\mathcal{E}_h}\theta_{KL}\delta_{KL}(W).\tag{48}$$

Next, the monotonicity of $-p_c$ yields

$$\forall a,b\in[0,1]\qquad -p_c(b)\,(b-a)\ge\psi(b)-\psi(a),\text{where }\psi:z\mapsto-\int_0^z p_c(s)\,ds.\tag{49}$$

Now we derive the L^2 estimates on the discrete gradients of P and of $(P+p_c(U))$. Multiplying (46) by P_K^{n+1}, multiplying (47) by $P_K^{n+1}+p_c(U_K^{n+1})$ and summing in $K\in\mathcal{T}_h$, we get after some cancellations the equality

$$\begin{array}{rcl} &&-\displaystyle\sum_{K\in\mathcal{T}_h}\frac{m_K}{\delta t^n}p_c(U_K^{n+1})\,\delta_K^{n+1,n}(U)\\ &&-\displaystyle\sum_{K\in\mathcal{T}_h}\left(P_K^{n+1}\sum_{L\in\mathcal{N}_K}A_{KL}^{n+1}\delta_{K,L}^{n+1}(P)-p_c(U_K^{n+1})\sum_{L\in\mathcal{N}_K}(A_{KL}^{n+1}-B_{KL}^{n+1})\delta_{K,L}^{n+1}(P)\right)\\ &&-\displaystyle\sum_{K\in\mathcal{T}_h}\mu(P_K^{n+1}+p_c(U_K^{n+1}))\sum_{L\in\mathcal{N}_K}\tau_{K|L}\mathfrak{d}_{K,L}^{n+1}[g(U)]\\ &=&-\displaystyle\sum_{K\in\mathcal{T}_h}m_K\sigma_K^{n+1}p_c(U_K^{n+1})+\sum_{K\in\mathcal{T}_h}m_Ks_K^{n+1}(P_K^{n+1}+p_c(U_K^{n+1})). \end{array}$$

On the first term, we use (49). On the other terms in the left-hand side, we use the summation-by-parts (48) (here the discrete zero-flux boundary condition on P and $p_c(U)$ is used, and the fact that $\delta_{K,L}^{n+1}\equiv-\delta_{LK}^{n+1}$). In the right-hand side, we use the boundedness of p_c and the weighted Young inequality with parameter $\alpha>0$ (α

will be chosen later). This results in the inequality

$$\begin{aligned}
&\sum_{K\in\mathcal{T}_h}\frac{m_K}{\delta t^n}\delta_K^{n+1,n}(\psi(U))\\
&+\sum_{K|L\in\mathcal{E}_h}\Big(A_{KL}^{n+1}|\delta_{K,L}^{n+1}(P)|^2-(A_{KL}^{n+1}-B_{KL}^{n+1})\delta_{K,L}^{n+1}(P)\delta_{K,L}^{n+1}(p_c(U))+\mu\tau_{K|L}\mathfrak{d}_{K,L}^{n+1}[g(U)]\delta_{K,L}^{n+1}\Big(P+p_c(U)\Big)\Big)\\
&\quad\leq\quad\alpha\sum_{K\in\mathcal{T}_h}m_K|P_K^{n+1}|^2+C_\alpha\sum_{K\in\mathcal{T}_h}m_K(|\,\overline{s}_K^{n+1}|^2+|\,\underline{s}_K^{n+1}|^2)+C.
\end{aligned}$$

Here and in the sequel, C denotes a generic constant that only depends on the data of the problem (in particular, C depend neither on μ nor on h); notation C_α indicates that C depends on α. Next, we multiply the obtained inequalities by δt^n and sum up in n. Using for each n the constraint $\sum_{K\in\mathcal{T}_h}m_KP_K^{n+1}=0$, the discrete Poincaré inequality (see [11]) with the proportionality constraint (30), from the L^2 bound on $\overline{s}$, $\underline{s}$ we get

$$\sum_{K\in\mathcal{T}_h}m_K\psi(U_K^{N+1})+\sum_{n=0}^N\delta t^n\sum_{K|L\in\mathcal{E}_h}E_{KL}^{n+1}\;\leq\;\alpha\sum_{n=0}^N\delta t^n\sum_{K|L\in\mathcal{E}_h}\tau_{K|L}|\delta_{K,L}^{n+1}(P)|^2+C_\alpha,\tag{50}$$

where $\alpha>0$ still denotes a generic number, and E_{KL}^{n+1} denotes the exchange term given by

$$\begin{aligned}
E_{KL}^{n+1}\;&=\;A_{KL}^{n+1}|\delta_{K,L}^{n+1}(P)|^2\;+\;A_{KL}^{n+1}(1-f_\mu(U_{K|L}^{n+1}))\delta_{K,L}^{n+1}(P)\delta_{K,L}^{n+1}(p_c(U))\\
&\quad+\mu\tau_{K|L}\mathfrak{d}_{K,L}^{n+1}[g(U)]\Big[\delta_{K,L}^{n+1}(P)+\delta_{K,L}^{n+1}(p_c(U))\Big]\;=:\;{}^{(1)}E_{KL}^{n+1}+{}^{(2)}E_{KL}^{n+1}+{}^{(3)}E_{KL}^{n+1}.
\end{aligned}$$

The first term in the above right-hand side can be rewritten as

$${}^{(1)}E_{KL}^{n+1}=\tau_{K|L}k_w(\bar{U}_{K|L}^{n+1})|\delta_{K,L}^{n+1}(P)|^2+\mu\tau_{K|L}k_a(\bar{U}_{K|L}^{n+1})|\delta_{K,L}^{n+1}(P)|^2.\tag{51}$$

In the second term, we insert $f_\mu(\bar{U}_{K|L}^{n+1})$ in the place of $f_\mu(U_{K|L}^{n+1})$ and then use the fact that

$$A_{KL}^{n+1}(1-f_\mu(\bar{U}_{K|L}^{n+1}))=\tau_{K|L}M_\mu(\bar{U}_{K|L}^{n+1})\frac{\mu k_a(\bar{U}_{K|L}^{n+1})}{M_\mu(\bar{U}_{K|L}^{n+1})}=\mu\tau_{K|L}k_a(\bar{U}_{K|L}^{n+1});$$

this yields

$${}^{(2)}E_{KL}^{n+1}=\mu\tau_{K|L}k_a(\bar{U}_{K|L}^{n+1})\delta_{K,L}^{n+1}(P)\delta_{K,L}^{n+1}(p_c(U))+D_{KL}^{n+1},\tag{52}$$

$$D_{KL}^{n+1}:=A_{KL}^{n+1}\Big(f_\mu(\bar{U}_{K|L}^{n+1})-f_\mu(U_{K|L}^{n+1})\Big)\delta_{K,L}^{n+1}(p_c(U))\delta_{K,L}^{n+1}(P).\tag{53}$$

We claim that the last term is non-negative. Indeed, recall that $\bar{U}_{K|L}^{n+1}$ is a value located between U_K^{n+1} and U_L^{n+1}, so that $\mathrm{sgn}(\bar{U}_{K|L}^{n+1}-U_L^{n+1})=\mathrm{sgn}(U_K^{n+1}-U_L^{n+1})$. Consider e.g. the case $\delta_{K,L}^{n+1}(P)\geq0$. Then by the upwind choice of $U_{K|L}^{n+1}$, we have $U_{K|L}^{n+1}=U_L^{n+1}$; by the monotonicity of f_μ, $-p_c$, we deduce

$$\begin{aligned}
\mathrm{sgn}D_{KL}^{n+1}\;&=\;\mathrm{sgn}\Big(f_\mu(\bar{U}_{K|L}^{n+1})-f_\mu(U_L^{n+1})\Big)\,\mathrm{sgn}\Big(p_c(U_L^{n+1})-p_c(U_K^{n+1})\Big)\\
&=\;\mathrm{sgn}\Big(\bar{U}_{K|L}^{n+1}-U_L^{n+1}\Big)\,\mathrm{sgn}\Big(U_K^{n+1}-U_L^{n+1}\Big)=\mathrm{sgn}\Big(U_K^{n+1}-U_L^{n+1}\Big)\,\mathrm{sgn}\Big(U_K^{n+1}-U_L^{n+1}\Big)\geq0.
\end{aligned}$$

The case $\delta_{K,L}^{n+1}(P)<0$ is analogous; thus $D_{KL}^{n+1}\geq0$.

Finally, we use the definition of $\bar{U}_{K|L}^{n+1}$ to derive

$$^{(3)}E_{KL}^{n+1} = \mu\tau_{K|L}k_a(\bar{U}_{K|L}^{n+1})\Big(|\delta_{K,L}^{n+1}(p_c(U))|^2 + \delta_{K,L}^{n+1}(p_c(U))\delta_{K,L}^{n+1}(P)\Big). \tag{54}$$

Estimating $k_w(\bar{U}_{K|L}^{n+1})$ from below by $k_w(u_m)$, gathering the terms from (51)-(54) we get

$$E_{KL}^{n+1} \;\geq\; k_w(u_m)\tau_{K|L}\,|\delta_{K,L}^{n+1}(P)|^2 \;+\; \mu\tau_{K|L}\,k_a(\bar{U}_{K|L}^{n+1})\,|\delta_{K,L}^{n+1}(P) + \delta_{K,L}^{n+1}(p_c(U))|^2.$$

Therefore, choosing $\alpha = k_w(u_m)/2$ in (50), we get the claims (42) and (43) of the proposition.

Before turning to the proof of the remaining estimate, let us point out that the system (5),(6) can be rewritten in the following "global flux" formulation:

$$-\mathrm{div}\, q = \overline{s} - \underline{s}, \quad q\cdot n|_{\partial\Omega\times(0,T)} = 0, \tag{55}$$

$$u_t - \mathrm{div}(f_\mu(u)q - k_w(u)\nabla Q(u)) \;=\; f_\mu(c)\,\overline{s} - f_\mu(u)\,\underline{s}, \quad \nabla Q(u)\cdot n|_{\partial\Omega\times(0,T)} = 0, \tag{56}$$

where

$$q = M_\mu(u)\,\nabla(P+Q(u)) \quad \text{and} \quad Q : z \mapsto \int_0^z (1-f_\mu(s))p_c'(s)\,ds.$$

Indeed, (55) is the sum of the two equations (5),(6); and (56) is just the equation (5) rewritten in terms of the global flux q. The idea is then to eliminate q from the system (55),(56). To this end, one can e.g. take the test function $-p_c(u)$ in (56), the test function $F_\mu(u)$ in (55), with

$$F_\mu : z \mapsto \int_0^z f_\mu(s)p_c'(s)\,ds,$$

and sum up. In our discrete setting, we follow the same procedure but without introducing q and Q. Indeed, the formulation (55),(56) of system (5),(6) relies on the use of chain rules that cannot extend directly if one replaces ∇ with the discrete differencing $\delta_{K,L}^{n+1}$.

Now we turn to the proof of (44). Sum up the discrete equations (46),(47) to produce an analogue of (55):

$$-\sum_{L\in\mathcal{N}_K} A_{KL}^{n+1}\delta_{K,L}^{n+1}(P) - \mu\sum_{L\in\mathcal{N}_K}\tau_{K|L}k_a(\bar{U}_{K|L}^{n+1})\delta_{K,L}^{n+1}(P_c(U)) = m_K s_K^{n+1}. \tag{57}$$

Here, we have used the definition on $\bar{U}_{K|L}^{n+1}$ and the identity $M_\mu(1-f_\mu) = \mu k_a$. Further, write (47) under the form

$$\frac{m_K}{\delta t^n}\delta_K^{n+1,n}(U) - \sum_{L\in\mathcal{N}_K} A_{KL}^{n+1}f_\mu(U_{K|L}^{n+1})\delta_{K,L}^{n+1}(P) = m_K\sigma_K^{n+1}. \tag{58}$$

Now we multiply (58) by $-p_c(U_K^{n+1})$, add (57) multiplied by $F(U_K^{n+1})$, and sum up in $K\in\mathcal{T}_h$. Using inequality (49) and estimating the right-hand side (notice that F_μ is bounded uniformly in μ, because $0\leq f_\mu\leq 1$), with the help of the summation-by-parts (48) we get

$$\begin{aligned}&\sum_{K\in\mathcal{T}_h}\frac{m_K}{\delta t^n}\delta_K^{n+1,n}(\psi(U)) + \sum_{K|L\in\mathcal{E}_h} A_{KL}^{n+1}\delta_{K,L}^{n+1}(P)\Big(\delta_{K,L}^{n+1}(F_\mu(U)) - f_\mu(U_{K|L}^{n+1})\delta_{K,L}^{n+1}(p_c(U))\Big)\\ &+\mu\sum_{K|L\in\mathcal{E}_h}\tau_{K|L}k_a(\bar{U}_{K|L}^{n+1})\delta_{K,L}^{n+1}(p_c(U))\delta_{K,L}^{n+1}(F_\mu(U)) \;\leq\; C\sum_{K\in\mathcal{T}_h} m_K(\,\overline{s}_K^{n+1} + \underline{s}_K^{n+1}).\end{aligned}$$

Now, observe that

$$\delta_{K,L}^{n+1}(F_\mu(U)) = \int_{U_K^{n+1}}^{U_L^{n+1}} f_\mu(s)p_c'(s)\,ds = f_\mu(\Theta_{KL}^{n+1})\delta_{K,L}^{n+1}(p_c(U)),$$

where Θ_{KL}^{n+1} is an intermediate point between U_K^{n+1} and U_L^{n+1}. Therefore

$$A_{KL}^{n+1}\delta_{K,L}^{n+1}(P)\Big(\delta_{K,L}^{n+1}(F_\mu(U)) - f_\mu(U_{K|L}^{n+1})\delta_{K,L}^{n+1}(p_c(U))\Big) = A_{KL}^{n+1}\delta_{K,L}^{n+1}(P)\Big(f_\mu(\Theta_{KL}^{n+1}) - f_\mu(U_{K|L}^{n+1})\Big)\delta_{K,L}^{n+1}(p_c(U)).$$

We claim that this term is non-negative; the arguments of the proof are the same as used to treat the term (53).

Using the definition (36) of $U_{K|L}^{n+1}$ and the fact that

$$\mu f_\mu(\Theta_{KL}^{n+1}) \geq \mu f_\mu(u_m) = \frac{\mu k_w(u_m)}{k_w(u_m)+\mu k_a(u_m)} \geq \frac{k_w(u_m)}{k_w(u_m)+k_a(u_m)} =: \frac{1}{C}$$

(here we assume that $\mu \geq 1$), we deduce the estimate

$$\sum_{K\in\mathcal{T}_h} \frac{m_K}{\delta t^n}\delta_K^{n+1,n}(\psi(U)) + \frac{1}{C}\sum_{K|L\in\mathcal{E}_h} \tau_{K|L}\mathfrak{d}_{K,L}^{n+1}[g(U)]\delta_{K,L}^{n+1}(p_c(U)) \leq \sum_{K\in\mathcal{T}_h} m_K(\overline{s}_K^{n+1} + \underline{s}_K^{n+1}).$$

Multiplication by δt^n and summation in n yield the estimate

$$\sum_{n=0}^{N}\delta t^n \sum_{K|L\in\mathcal{E}_h} \tau_{K|L}\mathfrak{d}_{K,L}^{n+1}[g(U)]\delta_{K,L}^{n+1}(p_c(U)) \leq C \tag{59}$$

Using the definition of $\mathfrak{d}_{K,L}^{n+1}[g(U)]$ we get

$$\sum_{n=0}^{N}\delta t^n \sum_{K|L\in\mathcal{E}_h} \tau_{K|L}\delta_{K,L}^{n+1}(\zeta(U))^2 \leq C. \tag{60}$$

and estimate (44) follows. The proof is concluded. □

5.3. **Existence of a discrete solution**

We prove here the existence of a solution to the scheme, which is a consequence of Leray-Schauder fixed point theorem. The idea of the proof is the following: if we can modify continuously the scheme to obtain a system which has a solution and if the modification preserves in the same time the estimates, then the scheme also has a solution.

Proposition 5.3. *Under Hypothesis (H), there exists a solution $(U_\mathcal{D}, P_\mathcal{D})$ to the scheme (31)-(34) .*

Proof. We define the vector space of discrete solutions $E_{h,N}$ by

$$E_{h,N} = \Big\{(U,P)\in\mathbb{R}^{\mathcal{T}_h\times[\![1,N+1]\!]}\times\mathbb{R}^{\mathcal{T}_h\times[\![1,N+1]\!]}\;\Big|\;\forall\, n\in[\![0,N]\!]\;\sum_{K\in\mathcal{T}_h} m_K P_K^{n+1} = 0\Big\}. \tag{61}$$

One easily checks that

$$\|(U,P)\|_{h,N} := \max\{|U_K^{n+1}|\,|\,(K,n)\in\mathcal{T}_h\times[\![0,N]\!]\} + \Big(\sum_{n=0}^{N}\delta t^n\sum_{K|L\in\mathcal{E}_h}\tau_{K|L}|\delta_{K,L}^{n+1}(P)|^2\Big)^{1/2}$$

is a norm on $E_{h,N}$.

For $t\in[0,1]$, set

$$u_0^t = tu_0 + (1-t)u_m, \quad \overline{s}^t = t\,\overline{s}, \quad \underline{s}^t = t\,\underline{s}. \tag{62}$$

Now we can define a continuous application $\mathcal{F}_\mu : [0,1] \times E_{h,N} \to E_{h,N}$ by $\mathcal{F}_\mu(t,(U,P)) = (W^t, Q^t)$. To define W^t and Q^t, for a given value $t \in [0,1]$, assign

$$U_K^0 := \frac{1}{m_K}\int_K u_0^t, \quad \overline{s}_K^{n+1} := \frac{1}{\delta t^n m_K}\int_{t^n}^{t^{n+1}}\int_K \overline{s}^t, \quad \underline{s}_K^{n+1} := \frac{1}{\delta t^n m_K}\int_{t^n}^{t^{n+1}}\int_K \underline{s}^t$$

(in the notation, we have dropped the dependency of t). Then for all $(K,n) \in \mathcal{T}_h \times [\![0,N]\!]$, W_K^{n+1} is the expression in the left-hand side of (32) and Q_K^{n+1} is the sum of W_K^{n+1} and of the expression in the left-hand side of (33), the t-dependent values U_K^0, $\overline{s}_K^{n+1}$, $\underline{s}_K^{n+1}$ being defined above. Here, we mean that the quantities $U_{K|L}^{n+1}, \bar{U}_{K|L}^{n+1}$ are defined by (35) and (36), respectively, starting from the values U_K^{n+1}, P_K^{n+1} contained in (U,P). Notice that we actually have $\sum_{K\in\mathcal{T}_h} m_K Q_K^{n+1} = 0$; this is because the scheme is conservative, and for all $n \in [\![0,N]\!]$, $\sum_{K\in\mathcal{T}_h} m_K(\overline{s}_K^{n+1} - \underline{s}_K^{n+1}) = 0$. By construction, $(U,P) \in E_{h,N}$ is a solution for the discrete scheme with the datum u_0^t and the source functions $\overline{s}^t$, $\underline{s}^t$ if and only if $\mathcal{F}_\mu(t,(U,P)) = 0$.

Let us now complete the proof. First of all, equation $\mathcal{F}_\mu(0,(U,P)) = 0$ admits a trivial solution given by $U_K^{n+1} = u_m$, $P_K^{n+1} = 0$ for all $(K,n) \in \mathcal{T}_h \times [\![0,N]\!]$. The function $\mathcal{F}_\mu$ is continuous. If X is a ball with a sufficiently large radius in $E_{h,N}$, the equation $\mathcal{F}_\mu(t,(U,P)) = 0$ has no solution on the boundary of X. Indeed, Proposition 5.2 applies for all value of t, with the bound C independent of t. From the estimate (42) and the definition of the norm on $E_{h,N}$, we get a bound, independent of t, on all possible solutions of equation $\mathcal{F}_\mu(t,(U,P)) = 0$. Therefore we can apply the Leray-Schauder topological degree theorem (see [10]). We get

$$degree(\mathcal{F}_\mu(1,\cdot),X) = degree(\mathcal{F}_\mu(0,\cdot),X) \neq 0, \tag{63}$$

and thus there exists at least a solution to equation $\mathcal{F}_\mu(1,(U_\mathcal{D},P_\mathcal{D})) = 0$. This solution is a solution to our scheme (31)-(34) . □

6. Compactness properties

From now on, we consider the the couple of piecewise constant functions

$$u_\mathcal{D} = \sum_{n=0}^{N}\sum_{K\in\mathcal{T}_h} U_K^{n+1}\mathbb{1}_{K\times(t^n,t^{n+1}]}, \qquad p_\mathcal{D} = \sum_{n=0}^{N}\sum_{K\in\mathcal{T}_h} P_K^{n+1}\mathbb{1}_{K\times(t^n,t^{n+1}]}. \tag{64}$$

defined starting from a discrete solution $(U_\mathcal{D}, P_\mathcal{D})$.

In this section, we pass to the limit in the sequence $(u_\mathcal{D}, p_\mathcal{D})$ of discrete solutions, as the discretization parameters tend to zero. Because the system considered is linear in p and nonlinear in u, weak "$L^2(0,T;H^1(\Omega))$ convergence" (see Lemma 6.2 for the exact statement) for $p_\mathcal{D}$ is sufficient; whereas strong (a.e. on $\Omega\times(0,T)$) convergence of $u_\mathcal{D}$ is required. The weak convergence of $p_\mathcal{D}$ is a consequence of the L^2 estimate of the discrete gradient, of the discrete Poincaré inequality and of the consistency of the discrete divergence operator used in our finite volume scheme. The strong convergence of $u_\mathcal{D}$ is obtained from the uniform $L^1(\Omega\times(0,T))$-translation estimates on $\zeta(u_\mathcal{D})$, the Kolmogorov theorem and the invertibility of ζ.

6.1. Weak compactness of $p_\mathcal{D}$ and $\zeta(u_\mathcal{D})$

Take $\varphi \in (\mathcal{D}(\Omega\times[0,T)))^d$, $d = 2,3$. Given a discretization $\mathcal{D}$ for all $K \in \mathcal{T}_h$, for all $n \in [\![0,N+1]\!]$ set

$$\varphi_K^n = \varphi(x_K,t^n), \quad \varphi_{K|L}^n = \frac{1}{m(K|L)}\int_{K|L}\varphi(\sigma,t^n)\,d\sigma, \quad \mathrm{div}_K\,\varphi_\mathcal{D}^n = \frac{1}{m_K}\sum_{L\in\mathcal{N}_K} m(K|L)\varphi_{K|L}^n\cdot\vec{n}_{K,L}; \tag{65}$$

and define the following functions on $\Omega \times (0,T)$:

$$\begin{aligned} &\varphi_{\mathcal{D}} = \sum_{n=0}^{N} \sum_{K\in\mathcal{T}_h} \varphi_K^{n+1} \mathbb{1}_{K\times(t^n,t^{n+1}]}, \quad \varphi_h = \sum_{n=0}^{N} \sum_{K|L\in\mathcal{E}_h} \varphi_{K|L}^{n+1} \mathbb{1}_{D_{K|L}\times(t^n,t^{n+1}]}, \\ &\text{and } \operatorname{div}_{\mathcal{D}} \varphi_{\mathcal{D}} = \sum_{n=0}^{N} \sum_{K\in\mathcal{T}_h} \Big(\operatorname{div}_K \varphi_{\mathcal{D}}^{n+1}\Big) \mathbb{1}_{K\times(t^n,t^{n+1}]}. \end{aligned} \tag{66}$$

Note the following consistency lemma.

Lemma 6.1. *With the notation (65),(66), we have $\varphi_{\mathcal{D}} \longrightarrow \varphi$, $\varphi_h \longrightarrow \varphi$ in $(L^\infty(\Omega\times(0,T)))^d$ and $\operatorname{div}_{\mathcal{D}} \varphi_{\mathcal{D}} \longrightarrow \operatorname{div}\varphi$ in $L^\infty(\Omega\times(0,T))$ as $\operatorname{size}(\mathcal{D}) \to 0$.*

Proof. The convergence of $\varphi_{\mathcal{D}}$ and φ_h is straightforward from the continuity of φ. For the second claim, notice that $\operatorname{div}_K \varphi_{\mathcal{D}}^n = \frac{1}{m_K}\int_K \operatorname{div}\varphi(x,t^n)\,dx$; thus the convergence follows from the continuity of $\operatorname{div}\varphi$. □

Now we can prove the "discrete $L^2(0,T;H^1(\Omega))$" compactness result.

Lemma 6.2. *Given a family of discretizations $\mathcal{D}$ such that $\operatorname{size}(\mathcal{D}) \to 0$, consider a family of corresponding discrete functions $V_{\mathcal{D}}$ satisfying the uniform bounds*

$$\sum_{n=0}^{N} \delta t^n \sum_{K\in\mathcal{T}_h} m_K (V_K^{n+1})^2 \le C, \qquad \sum_{n=0}^{N} \delta t^n \sum_{K|L\in\mathcal{E}_h} \tau_{K|L} \big(\delta_{K,L}^{n+1}(V_{\mathcal{D}})\big)^2 \le C. \tag{67}$$

Let denote $v_{\mathcal{D}} = \sum_{n=0}^{N} \sum_{K\in\mathcal{T}_h} V_K^{n+1} \mathbb{1}_{K\times(t^n,t^{n+1})}$ and $\nabla_h V_{\mathcal{D}} = \sum_{n=0}^{N} \sum_{K|L\in\mathcal{E}_h} \nabla_{K|L} V_{\mathcal{D}}^{n+1}\ \mathbb{1}_{D_{K|L}\times(t^n,t^{n+1})}$, where we use the following definition of the discrete gradient:

$$\nabla_{K|L} V_{\mathcal{D}}^{n+1} := d\frac{\delta_{K,L}^{n+1}(V_{\mathcal{D}})}{d_{KL}} n_{\vec{K}L} \tag{68}$$

Then there exists $v \in L^2(0,T;H^1(\Omega))$ such that, up to extraction of a subsequence, $v_{\mathcal{D}} \longrightarrow v$ in $L^2(\Omega\times(0,T))$ weakly and $\nabla_h V_{\mathcal{D}} \longrightarrow \nabla v$ in $(L^2(\Omega\times(0,T)))^d$ weakly.

Proof. The inequality (67) exactly means that $\nabla_h V_{\mathcal{D}}$ are bounded in $(L^2(\Omega\times(0,T)))^d$ uniformly in $\mathcal{D}$ and $v_{\mathcal{D}}$ are uniformly bounded in $L^2(\Omega\times(0,T))$. Extracting weakly convergent subsequences, we get $(v_{\mathcal{D}}, \nabla_h V_{\mathcal{D}}) \longrightarrow (v,q)$ in L^2 weakly. It remains to identify q to the gradient of v in the sense of distributions.

Taking $\varphi \in (\mathcal{D}(\Omega\times(0,T)))^d$, by Lemma 6.1 we have $\displaystyle\int_0^T\!\!\int_\Omega v_{\mathcal{D}} \operatorname{div}_{\mathcal{D}}\varphi_{\mathcal{D}} \longrightarrow \int_0^T\!\!\int_\Omega v \operatorname{div}\varphi$ as $\operatorname{size}(\mathcal{D}) \to 0$. On the other hand, summing by parts and using the definition (68) of the components of the discrete gradient, using again Lemma 6.1 we have

$$\begin{aligned} \int_0^T\!\!\int_\Omega v_{\mathcal{D}} \operatorname{div}_{\mathcal{D}}\varphi_{\mathcal{D}} &= \sum_{n=0}^{N} \delta t^n \sum_{K\in\mathcal{T}_h} m_K\, V_K^{n+1}\, (\operatorname{div}_K \varphi_{\mathcal{D}}^{n+1}) = \sum_{n=0}^{N} \delta t^n \sum_{K\in\mathcal{T}_h} V_K^{n+1} \sum_{L\in\mathcal{N}_K} m(K|L)\, \varphi_{K|L}^{n+1} \cdot \vec{n}_{K,L} \\ &= -\sum_n \delta t^n \sum_{K|L\in\mathcal{E}_h} m(K|L)(V_L^{n+1} - V_K^{n+1}) \vec{n}_{K,L} \cdot \varphi_{K|L}^{n+1} \\ &= -\sum_n \delta t^n \sum_{K|L\in\mathcal{E}_h} \Big(\frac{m(K|L)d_{KL}}{d}\Big)\Big(d\frac{\delta_{KL}^{n+1}(V_{\mathcal{D}})}{d_{KL}} \vec{n}_{K,L}\Big) \cdot \varphi_{K|L}^{n+1} \\ &= -\sum_n \delta t^n \sum_{K|L\in\mathcal{E}_h} m_{D_{K|L}} \nabla_h V_{\mathcal{D}} \cdot \varphi_{K|L}^{n+1} = -\int_0^T\!\!\int_\Omega \nabla_h V_{\mathcal{D}} \cdot \varphi_h \longrightarrow -\int_0^T\!\!\int_\Omega q\cdot\varphi. \end{aligned}$$

This leads to the conclusion that $\nabla v = q$ in the sense of distributions, and in particular, $v \in L^2(0,T;H^1(\Omega))$. $\square$

According to Proposition 5.2, the families of discrete functions $P_{\mathcal{D}}$ and $\zeta(U_{\mathcal{D}})$ verify the assumptions of this lemma (notice that the L^2 estimate of $P_{\mathcal{D}}$ comes from the L^2 estimate of the discrete gradient, the normalization of $P_{\mathcal{D}}$ and the corresponding discrete Poincaré inequality; and $\zeta(U_{\mathcal{D}})$ is L^∞ bounded). By Lemma 6.2 we deduce they weakly converge, respectively, to p and $\tilde{\zeta}$ in the sense of the lemma.

6.2. **Estimates of space and time translates of $\zeta(u_{\mathcal{D}})$**

It is classical (see [11]) that the estimate (44) on the discrete space gradient of $\zeta(U_{\mathcal{D}})$, given in Proposition 5.2, implies the estimate of the space translates. For the sake of completeness, we state the L^1 result and give the proof.

Corollary 6.1. *[L^1 translation estimate] Assume that (H) are fulfilled, and for all discretization $\mathcal{D}$, $V_{\mathcal{D}}$ is a corresponding discrete function. Assume that there exists a constant C, independent of $\mathcal{D}$, such that*

$$\sum_{n=0}^{N} \delta t^n \sum_{K|L\in\mathcal{E}_h} m(K|L)|\delta_{K,L}^{n+1}(V_{\mathcal{D}})| \le C. \tag{69}$$

Then for any $\xi \in \mathbb{R}^d$ such that $|\xi| \le \mathrm{diam}(\Omega)$, there holds

$$\int_0^T \int_{\Omega_\xi} \Big| v_{\mathcal{D}}(x+\xi,t) - v_{\mathcal{D}}(x,t) \Big| dxdt \le C|\xi|, \tag{70}$$

where $\Omega_\xi = \{x \in \Omega, [x, x+\xi] \subset \Omega\}$ and $v_{\mathcal{D}} = \displaystyle\sum_{n=0}^{N} \sum_{K\in\mathcal{T}_h} V_K^{n+1} \mathbb{1}_{K\times(t^n,t^{n+1}]}$.

Proof. For a given $\xi \in \mathbb{R}^d$, for $x \in \mathbb{R}^d$ set $\overline{\psi}_{K|L}(x) = 1$, in case the segment $[x, x+\xi]$ crosses $K|L$, and $\overline{\psi}_{K|L}(x) = 0$ otherwise. By the triangle inequality,

$$|v_{\mathcal{D}}(x,t) - v_{\mathcal{D}}(x+\xi,t)| \le \sum_{K|L\in\mathcal{E}_h} \overline{\psi}_{K|L}(x)\, |V_L^{n(t)+1} - V_K^{n(t)+1}|,$$

where $n(t)$ is defined by the fact that $t \in (t^{n(t)}, t^{n(t)+1}]$. In addition, we have $\displaystyle\int_{\mathbb{R}^d} \overline{\psi}_{K|L}(x)\, dx \le m(K|L)\, |\xi|$. Hence (70) follows, since

$$\begin{aligned}\int_0^T \int_{\Omega_\xi} |v_{\mathcal{D}}(x,t) - v_{\mathcal{D}}(x+\xi,t)|\, dxdt &\le \sum_{n=0}^{N} \delta t^n \int_{\Omega_\xi} \sum_{K|L\in\mathcal{E}_h} \overline{\psi}_{K|L}(x)\, |V_L^{n+1} - V_K^{n+1}|\, dx \\ &\le\ |\xi| \sum_{n=0}^{N} \delta t^n \sum_{K|L\in\mathcal{E}_h} m(K|L)\, |V_L^{n+1} - V_K^{n+1}| \le C|\xi|.\end{aligned}$$

$\square$

Obtaining the time translates is technical, although the approach is classical (see [11]). By making appeal to L^1 rather than to L^2 translates, we avoid the Lipschitz regularity assumption on ζ which the L^2 approach would have required.

Notice that, for technical reasons, we give a μ-dependent estimate; yet we point out at the end of Section 7 that the continuous limit of the scheme admits a uniform in μ time translation estimate.

Proposition 6.1 (Time translates of $\zeta(u)$)**.** *Under hypothesis (H), let $\mathcal{D}$ be a finite volume discretization of $\Omega\times(0,T)$ in the sense of Definition 4.3 and let $(U_\mathcal{D},P_\mathcal{D})$ be a solution of the finite volume scheme (31)-(34) .*

Then there exists a function $\tilde\omega\in\mathcal{C}^0(\mathbb{R}^+;\mathbb{R}^+)$, with $\tilde\omega(0)=0$, which only depends on k_w, k_a, p_c, Ω, T, u_m, $\|\overline{s}\|_{L^2(\Omega\times(0,T))}$, $\|\underline{s}\|_{L^2(\Omega\times(0,T))}$ and on μ but not on $\mathcal{D}$ such that, for all $\tau\in(0,T)$, the following estimate holds:

$$\int_0^{T-\tau}\int_\Omega\Big|\,\zeta(u_\mathcal{D}(x,t+\tau))-\zeta(u_\mathcal{D}(x,t))\,\Big|\,dx\,dt\le\tilde\omega(\tau) \tag{71}$$

Proof. In order to obtain the L^1 translates of $\zeta(u_\mathcal{D})$ and avoid restrictions on the modulus of continuity of ζ, let us first make the following observation. Let π be a concave, strictly increasing modulus of continuity of ζ on $[0,1]$, Π be the inverse of π, and $\overline{\Pi}(r)=r\,\Pi(r)$. Let $\overline{\pi}$ be the inverse of $\overline{\Pi}$. Note that $\overline{\pi}$ is concave strictly increasing, continuous, $\overline{\pi}(0)=0$. Denote $v=u_\mathcal{D}(x,t+\tau)$ and $y=u_\mathcal{D}(x,t)$; we will now omit "$dx\,dt$" in the integrals over $\Omega\times(0,T-\tau)$. Using the Jensen inequality, we have

$$\begin{aligned}\int_0^{T-\tau}\int_\Omega|\zeta(v)-\zeta(y)|&=\int_0^{T-\tau}\int_\Omega\overline{\pi}\Big(\overline{\Pi}(|\zeta(v)-\zeta(y)|)\Big)\\&\le(T-\tau)|\Omega|\,\overline{\pi}\bigg(\frac{1}{(T-\tau)|\Omega|}\int_0^{T-\tau}\int_\Omega\overline{\Pi}(|\zeta(v)-\zeta(y)|)\bigg).\end{aligned}$$

Since $|\zeta(v)-\zeta(y)|\le\pi(|v-y|)$, we have $\Pi(|\zeta(v)-\zeta(y)|)\le|v-y|$ and

$$\overline{\Pi}(|\zeta(v)-\zeta(y)|)=\Pi(|\zeta(v)-\zeta(y)|)|\zeta(v)-\zeta(y)|\le|v-y|\,|\zeta(v)-\zeta(y)|.$$

Therefore, in order to prove the proposition it is enough to estimate the quantity

$$\int_0^{T-\tau}\int_\Omega-\Big(u_\mathcal{D}(x,t+\tau)-u_\mathcal{D}(x,t)\Big)\Big(\zeta(u_\mathcal{D}(x,t+\tau))-\zeta(u_\mathcal{D}(x,t))\Big)dxdt=:\int_0^{T-\tau}A(t)dt \tag{72}$$

(recall that $-\zeta(\cdot)$ is a non-decreasing function). The estimate is similar to the one given in [11]. The idea is that the quantity $A(t)$ naturally appears when we "integrate" the discrete equations (32) between t and $t+\tau$ and then multiply by the discrete test function $-\Big(\zeta(u_\mathcal{D}(\cdot,t+\tau))-\zeta(u_\mathcal{D}(\cdot,t))\Big)$ and sum up in $K\in\mathcal{T}_h$. After summation in n, we use the discrete version of the Fubini theorem (see Lemma 6.3 below) to make appear τ, and the estimate of the discrete gradient of $\zeta(u_\mathcal{D})$ yields a control of the integral (72) by $C\tau$, with a constant C independent of $\mathcal{D}$ and μ.

For $t\in[0,t)$, let us denote by $n(t)$ the integer $n\in[\![0,N+1]\!]$ such that $t\in(t^n,t^{n+1}]$. With this notation,

$$\begin{aligned}A(t)&=-\sum_{K\in\mathcal{T}_h}m_K(\zeta(U_K^{n(t)+1})-\zeta(U_K^{n(t+\tau)+1}))(U_K^{n(t+\tau)+1}-U_K^{n(t)+1})\\&=-\sum_{K\in\mathcal{T}_h}(\zeta(U_K^{n(t)+1})-\zeta(U_K^{n(t+\tau)+1}))\sum_{n=n(t)+1}^{n(t+\tau)}m_K(U_K^{n+1}-U_K^n)\\&=-\sum_{K\in\mathcal{T}_h}(\zeta(U_K^{n(t)+1})-\zeta(U_K^{n(t+\tau)+1}))\sum_{n=n(t)+1}^{n(t+\tau)}\delta t^n\left(\begin{array}{c}\displaystyle\sum_{L\in\mathcal{N}_K}\tau_{K|L}f_\mu(U_{K|L}^{n+1})M_\mu(\bar U_{K|L}^{n+1})\delta_{K,L}^{n+1}(P)+\\ m_K(f_\mu(c_K^{n+1})\,\overline{s}_K^{n+1}-f_\mu(U_K^{n+1})\,\underline{s}_K^{n+1})\end{array}\right).\end{aligned}$$

Gathering by edges, we get

$$\begin{aligned} A(t) \;\leq\; & -\sum_{n=n(t)+1}^{n(t+\tau)} \delta t^n \sum_{K\in\mathcal{T}_h} \left(\sum_{L\in\mathcal{N}_K} \tau_{K|L}\, f_\mu(U_{K|L}^{n+1}) M_\mu(\bar U_{K|L}^{n+1})\, \delta_{K,L}^{n+1}(P)\, \delta_{KL}^{n(t)+1}(\zeta(U)) \right) \\ & + \sum_{n=n(t)+1}^{n(t+\tau)} \delta t^n \sum_{K\in\mathcal{T}_h} \left(\sum_{L\in\mathcal{N}_K} \tau_{K|L}\, f_\mu(U_{K|L}^{n+1}) M_\mu(\bar U_{K|L}^{n+1})\, \delta_{K,L}^{n+1}(P)\, \delta_{KL}^{n(t+\tau)+1}(\zeta(U)) \right) \\ & - \sum_{n=n(t)+1}^{n(t+\tau)} \delta t^n \sum_{K\in\mathcal{T}_h} m_K \left(\left(f_\mu(c_K^{n+1})\, \overline{s}_K^{n+1} - f_\mu(U_K^{n+1})\, \underline{s}_K^{n+1} \right) \left(\zeta(U_K^{n(t)+1}) - \zeta(U_K^{n(t+\tau)+1}) \right) \right). \end{aligned}$$

Thanks to the Young inequality and the uniform bound on $\zeta(U)$, we get

$$A(t) \;\leq\; \tfrac{1}{2}\Big(2A_{(1)}(t) + A_{(2)}(t) + A_{(3)}(t) + 2A_{(4)}(t) \Big)$$

with

$$\begin{aligned} A_{(1)}(t) &:= \sum_{n=n(t)+1}^{n(t+\tau)} \delta t^n \sum_{K\in\mathcal{T}_h} \left(\sum_{L\in\mathcal{N}_K} \tau_{K|L} f_\mu(U_{K|L}^{n+1}) M_\mu(\bar U_{K|L}^{n+1}) \big| \delta_{K,L}^{n+1}(P) \big|^2 \right) =: \sum_{n=n(t)+1}^{n(t+\tau)} a_{(1)}^{n+1} \\ A_{(2)}(t) &:= \sum_{n=n(t)+1}^{n(t+\tau)} \delta t^n \sum_{K\in\mathcal{T}_h} \left(\sum_{L\in\mathcal{N}_K} \tau_{K|L} f_\mu(U_{K|L}^{n+1}) M_\mu(\bar U_{K|L}^{n+1}) \big| \delta_{KL}^{n(t)+1}(\zeta(U)) \big|^2 \right) =: \sum_{n=n(t)+1}^{n(t+\tau)} a_{(2)}^{n+1} \\ A_{(3)}(t) &:= \sum_{n=n(t)+1}^{n(t+\tau)} \delta t^n \sum_{K\in\mathcal{T}_h} \left(\sum_{L\in\mathcal{N}_K} \tau_{K|L} f_\mu(U_{K|L}^{n+1}) M_\mu(\bar U_{K|L}^{n+1}) \big| \delta_{KL}^{n(t+\tau)+1}(\zeta(U)) \big|^2 \right) =: \sum_{n=n(t)+1}^{n(t+\tau)} a_{(3)}^{n+1} \\ A_{(4)}(t) &:= C \sum_{n=n(t)+1}^{n(t+\tau)} \sum_{K\in\mathcal{T}_h} (\overline{s}_K^{n+1} + \underline{s}_K^{n+1}) =: \sum_{n=n(t)+1}^{n(t+\tau)} a_{(4)}^{n+1}. \end{aligned}$$

In order to conclude the proof, we use the following lemma shown in [12].

Lemma 6.3. *Let $T>0$, $\tau\in(0,T)$ and $(a^n)_{n\in\mathbb{N}}$ be a family of non negative real values. Then*

$$\int_0^{T-\tau} \sum_{n=n(t)+1}^{n(t+\tau)} a^{n+1} dt \leq \tau \sum_{n=0}^{N} a^{n+1}, \tag{73}$$

and for any $\sigma\in[0,\tau]$

$$\int_0^{T-\tau} \sum_{n=n(t)+1}^{n(t+\tau)} a^{n(t+\sigma)+1} dt \leq \tau \sum_{n=0}^{N} a^{n+1}. \tag{74}$$

Notice that all the quantities $\displaystyle\sum_{n=0}^{N} a_{(i)}^{n+1}$, $i=1,2,3$, are controlled by the uniform in $\mathcal{D}$ estimates of Proposition 5.2; the bound depends on μ, because we roughly estimate $f_\mu(U_{K|L}^{n+1})M_\mu(\bar U_{K|L}^{n+1})$ by $k_w(1)+\mu k_a(u_m)$. Hence we easily derive the desired estimate $\displaystyle\int_0^{T-\tau} A(t)dt \leq C\tau$; the estimate (71) follows. □

6.3. **Strong compactness of $u_{\mathcal{D}}$**

Firstly, the L^2-kind estimate (44), uniform in $\mathcal{D}$ and μ, implies the L^1-kind estimate (69) on $V_{\mathcal{D}} = \zeta(U_{\mathcal{D}})$:

$$\begin{aligned}\sum_{n=0}^{N}\delta t^n \sum_{K|L\in\mathcal{E}_h} m(K|L)|\delta_{K,L}^{n+1}(V_{\mathcal{D}})| &= \sum_{n=0}^{N}\delta t^n \sum_{K|L\in\mathcal{E}_h} m(K|L)d_{K|L}\Big|\frac{\delta_{K,L}^{n+1}(V_{\mathcal{D}})}{d_{K|L}}\Big| \\ &\le \Big(\sum_{n=0}^{N}\delta t^n \sum_{K|L\in\mathcal{E}_h} m(K|L)d_{K|L}\Big|\frac{\delta_{K,L}^{n+1}(V_{\mathcal{D}})}{d_{K|L}}\Big|^2\Big)^{1/2}\Big(\sum_{n=0}^{N}\delta t^n \sum_{K|L\in\mathcal{E}_h} m(K|L)d_{K|L}\Big)^{1/2} \\ &= (T|\Omega|)^{1/2}\Big(\sum_{n=0}^{N}\delta t^n \sum_{K|L\in\mathcal{E}_h} \tau_{K|L}|\delta_{K,L}^{n+1}(V_{\mathcal{D}})|^2\Big)^{1/2} \le C.\end{aligned}$$

Secondly, the uniform bound on $U_{\mathcal{D}}$ of Proposition 5.1 and the translation estimates of Propositions 6.1 and 6.1 permit to extend $u_{\mathcal{D}}$ by zero outside $\Omega\times(0,T)$ and get a uniform estimate of

$$\int_{\mathbb{R}}\int_{\mathbb{R}^d}\Big|\,\zeta\Big(u_{\mathcal{D}}(x+\xi,t+\tau)\Big)-\zeta\Big(u_{\mathcal{D}}(x,t)\Big)\,\Big|\,dx\,dt$$

by a quantity vanishing as $\xi\to 0$, $\tau\to 0$. Therefore $\zeta(u_{\mathcal{D}})$ satisfies the assumptions of the Kolmogorov's compactness theorem; in particular, if we take a sequence of discretizations D_m with $\lim_{n\to+\infty}\mathrm{size}(D_m)=0$, there exists a function $\tilde{\zeta}$ such that up to a subsequence, $\zeta(u_{\mathcal{D}_m})\to\tilde{\zeta}$ in $L^1(\Omega\times(0,T))$ and a.e. on $\Omega\times(0,T)$. Since ζ is coninuous and strictly decreasing, its inverse is also continuous; therefore $u_{\mathcal{D}_m}\to u:=\zeta^{-1}(\tilde{\zeta})$ a.e on $\Omega\times(0,T)$.

Recall that we have already shown in Section 6.1 that the a.e. limit $\tilde{\zeta}=\zeta(u)$ of $\zeta(u_{\mathcal{D}})$ (which is also the weak L^2 limit) actually belongs to $L^2(0,T;H^1(\Omega))$.

The last step in the proof of the convergence theorem 4.1 it to justify that (u,p) is a weak solution. This is the aim of the next section.

7. Proof of convergence

We now finish the proof of Theorem 4.1 started in Section 6. The proof follows classical guidelines: using discrete integration-by-parts arguments, we write down the "weak discrete formulation" in which we pass to the limit, with the help of the consistency properties (the discrete differential operators applied to a smooth test function φ converge to the corresponding derivatives of φ) and the compactness results shown in Section 6.

Take $\varphi\in\mathcal{C}^\infty(\mathbb{R}^d\times[0,T))$. We set $\varphi_K^{n+1}:=\varphi(x_K,t^{n+1})$ and multiply Equations (32) and (33) by $\varphi_K^{n+1}\delta t^n$; then we sum over $K\in\mathcal{T}_h$ and $n\in[\![0,N]\!]$. Thus we get

$$\left\{\begin{array}{l} T1_h+F1_h=\Sigma 1_h+\sigma 1_h \\ T2_h+F2_h+G_h=\Sigma 2_h+\sigma 2_h\end{array}\right. \tag{75}$$

with the following notation:

$$T1_h=\sum_{n=0}^{N}\sum_{K\in\mathcal{T}_h} m_K\big(U_K^{n+1}-U_K^n\big)\varphi_K^{n+1},\qquad T2_h=\sum_{n=0}^{N}\sum_{K\in\mathcal{T}_h}\big((1-U_K^{n+1})-(1-U_K^n)\big)\,m_K\varphi_K^{n+1};$$

$$F1_h = -\sum_{n=0}^{N} \delta t^n \sum_{K\in\mathcal{T}_h} \sum_{L\in\mathcal{N}_K} \tau_{K|L} f_\mu(U_{K|L}^{n+1}) M_\mu(\bar{U}_{K|L}^{n+1}) \delta_{K,L}^{n+1}(P) \varphi_K^{n+1},$$
$$F2_h = -\sum_{n=0}^{N} \delta t^n \sum_{K\in\mathcal{T}_h} \sum_{L\in\mathcal{N}_K} \tau_{K|L} (1 - f_\mu(U_{K|L}^{n+1})) M_\mu(\bar{U}_{K|L}^{n+1}) \delta_{K,L}^{n+1}(P) \varphi_K^{n+1},$$
$$G_h = -\sum_{n=0}^{N} \delta t^n \sum_{K\in\mathcal{T}_h} \mu \sum_{L\in\mathcal{N}_K} \tau_{K|L} \mathfrak{d}_{K,L}^{n+1}[g(U)] \varphi_K^{n+1};$$

$$\Sigma 1_h = \sum_{n=0}^{N} \delta t^n \sum_{K\in\mathcal{T}_h} m_K f_\mu(c_K^{n+1})\, \overline{s}_K^{n+1} \varphi_K^{n+1}, \qquad \sigma 1_h = -\sum_{n=0}^{N} \delta t^n \sum_{K\in\mathcal{T}_h} m_K f_\mu(U_K^{n+1})\, \underline{s}_K^{n+1} \varphi_K^{n+1},$$
$$\Sigma 2_h = \sum_{n=0}^{N} \delta t^n \sum_{K\in\mathcal{T}_h} m_K (1 - f_\mu(c_K^{n+1}))\, \overline{s}_K^{n+1} \varphi_K^{n+1}, \qquad \sigma 2_h = -\sum_{n=0}^{N} \delta t^n \sum_{K\in\mathcal{T}_h} m_K (1 - f_\mu(U_K^{n+1}))\, \underline{s}_K^{n+1} \varphi_K^{n+1}.$$

Now we show that the discrete terms converge to the corresponding integral terms of the weak formulation of Problem (5)-(10). Let us set

$$\begin{cases} \varphi_{\mathcal{D}}(.,0) = \displaystyle\sum_{n=0}^{N} \varphi_K^0 \mathbb{1}_K(.) \quad \text{where } \varphi_K^0 = \varphi(x_K, 0) \\ \varphi_{\mathcal{D}}(x,t) = \displaystyle\sum_{n=0}^{N} \sum_{K\in\mathcal{T}_h} \varphi(x_k, t^{n+1}) \mathbb{1}_{K\times(t^n,t^{n+1}]}, \text{ for } t \in]0,T] \\ \dfrac{\partial_{\mathcal{D}}}{\partial t} \varphi_{\mathcal{D}}(x,t) = \displaystyle\sum_{n=0}^{N} \sum_{K\in\mathcal{T}_h} \frac{\varphi_K^{n+1} - \varphi_K^n}{\delta t^n} \mathbb{1}_{K\times(t^n,t^{n+1}]} \\ \bar{\nabla}_h \varphi_{\mathcal{D}} = \displaystyle\sum_{n=0}^{N} \sum_{K|L\in\mathcal{E}_h} \left(\bar{\nabla}_{K|L} \varphi_{\mathcal{D}}^{n+1}\right) \mathbb{1}_{D_{K|L}\times(t^n,t^{n+1}]} \\ \qquad \text{where we set} \quad \bar{\nabla}_{K|L} \varphi_{\mathcal{D}}^{n+1} = \displaystyle\int_0^1 \nabla\varphi(\theta x_K + (1-\theta) x_L, t^{n+1})\, d\theta. \end{cases} \tag{76}$$

Notice the following fact:

$$\bar{\nabla}_{K|L} \varphi_{\mathcal{D}}^{n+1} \cdot \vec{n}_{K,L} = \frac{\delta_{K,L}^{n+1}(\varphi_{\mathcal{D}})}{d_{K|L}}. \tag{77}$$

Because φ is smooth, we have the following consistency properties as size$(\mathcal{D}) \to 0$ (cf. Lemma 6.1):

$$\begin{array}{l} \varphi_{\mathcal{D}}(.,0) \longrightarrow \varphi(.,0) \text{ a.e. on } \Omega \text{ with a uniform } L^\infty \text{ bound;} \\ \varphi_{\mathcal{D}} \longrightarrow \varphi, \quad \dfrac{\partial_{\mathcal{D}}}{\partial t} \varphi_{\mathcal{D}} \longrightarrow \dfrac{\partial \varphi}{\partial t}, \quad \bar{\nabla}_h \varphi_{\mathcal{D}} \longrightarrow \nabla\varphi \quad \text{a.e. on } \Omega\times(0,T) \text{ with a uniform } L^\infty \text{ bound.} \end{array} \tag{78}$$

We have shown in Section 6 that (up to extraction of a subsequence) $u_{\mathcal{D}}$ converge to u a.e. on $\Omega\times(0,T)$ with a uniform L^∞ bound. In addition, define

$$U_h(x,t) = \sum_{n=0}^{N} \sum_{K|L\in\mathcal{E}_h} U_{K|L}^{n+1} \mathbb{1}_{D_{K|L}\times(t^n,t^{n+1}]} \quad \text{and} \quad \overline{U}_h(x,t) = \sum_{n=0}^{N} \sum_{K|L\in\mathcal{E}_h} \bar{U}_{K|L}^{n+1} \mathbb{1}_{D_{K|L}\times(t^n,t^{n+1}]}.$$

Lemma 7.1. *Both U_h and $\bar{U}_h$ converge (up to extraction of a subsequence) to u a.e. on $\Omega\times(0,T)$ with a uniform L^∞ bound.*

Proof. For all couple K, L of neighbours, we have either $U_{KL} = U_K$, or $U_{KL} = U_L$; hence

$$\max\{|\zeta(U_{KL}) - \zeta(U_K)|, |\zeta(U_{KL}) - \zeta(U_L)|\} \leq \text{size}(\mathcal{D})|\delta^{n+1}_{K,L}(\zeta(U))|. \tag{79}$$

Then directly from the definitions of $u_\mathcal{D}$, U_h and $\nabla_h\zeta(U_\mathcal{D})$ we have

$$\begin{aligned}\|\zeta(U_h) - \zeta(u_\mathcal{D})\|^2_{L^2(\Omega\times(0,T))} &= \\ &= \sum_{n=0}^{N}\delta t^n \sum_{K|L\in\mathcal{E}_h}\left(m_{D_{K|L}\cap K}|\zeta(U^{n+1}_{K|L}) - \zeta(U^{n+1}_K)|^2 + m_{D_{K|L}\cap L}|\zeta(U^{n+1}_{K|L}) - \zeta(U^{n+1}_L)|^2\right) \\ &\leq \text{size}(\mathcal{D})\sum_{n=0}^{N}\delta t^n \sum_{K|L\in\mathcal{E}_h} m_{D_{K|L}}|\delta^{n+1}_{K,L}(\zeta(U))|^2 \;=\; \text{size}(\mathcal{D})\,\|\nabla_h\zeta(U_\mathcal{D})\|^2_{L^2(\Omega\times(0,T))}.\end{aligned}$$

Therefore the uniform estimate (44) yields the convergence of $\zeta(U_h) - \zeta(u_\mathcal{D})$ to zero in $L^2(\Omega\times(0,T))$ as $\text{size}(\mathcal{D}) \to 0$. Due to the strict monotonicity of $\zeta(\cdot)$, up to extraction of a subsequence $U_h - u_\mathcal{D} \to 0$ a.e.; thus $U_h - u \to 0$ a.e. on $\Omega\times(0,T)$. Because U_h takes values in $[0,1]$, the first claim of the lemma follows.

Because ζ is monotone and $\bar{U}_{KL}$ is a value between U_K and U_L, the inequality (79) holds with U_{KL} replaced with $\bar{U}_{KL}$; thus the second claim of the lemma is justified in the same way. □

As a consequence of (78) and the above lemma, by the dominated convergence theorem we have in particular

$$f_\mu(U_h)M_\mu(\bar{U}_h)\bar{\nabla}_h\varphi_\mathcal{D} \longrightarrow f_\mu(u)M_\mu(u)\nabla\varphi = k_w(u)\nabla\varphi \;\text{ in } L^2(\Omega\times(0,T)). \tag{80}$$

Now we pass to the limit in (75), term by term. Using the summation-by-parts procedure and the fact that $\varphi(\cdot, t^{N+1}) = 0$, we can write

$$\begin{aligned}T1_h &= -\sum_{n=0}^{N}\delta t^n\sum_{K\in\mathcal{T}_h} m_K U^{n+1}_K\frac{\varphi^{n+1}_k - \varphi^n_k}{\delta t^n} + \sum_{K\in\mathcal{T}_h} m_K U^N_K\varphi^{N+1}_K - \sum_{K\in\mathcal{T}_h} m_K U^0_K\varphi^1_k \\ &= -\int_0^T\int_\Omega u_\mathcal{D}\frac{\partial_\mathcal{D}\varphi_\mathcal{D}}{\partial t} - \int_\Omega {u_\mathcal{D}}^0(x)\varphi_\mathcal{D}(0,x);\end{aligned}$$

the term $T2_h$ is treated analogously. Thanks of the fact that $u_\mathcal{D} \to u$ (resp., ${u_\mathcal{D}}^0 \to u_0$) a.e. on $\Omega\times(0,T)$) (resp., a.e. on Ω) with a uniform L^∞ bound, from (78) we infer

$$T1_h \longrightarrow -\int_0^T\int_\Omega u\frac{\partial\varphi}{\partial t} - \int_\Omega u_0(x)\varphi(x,0), \qquad T2_h \longrightarrow -\int_0^T\int_\Omega(1-u)\frac{\partial}{\partial t}\varphi - \int_\Omega(1-u_0)\varphi(x,0).$$

Next, consider the term $F1_h$. Using the summation-by-parts and taking (68),(77) into account, we re-write the term as

$$\begin{aligned}F1_h &= \sum_{n=0}^{N}\delta t^n\sum_{K|L\in\mathcal{E}_h}\tau_{K|L}f_\mu(U^{n+1}_{K|L})M_\mu(\bar{U}^{n+1}_{K|L})\delta^{n+1}_{K,L}(P_\mathcal{D})\delta^{n+1}_{K,L}(\varphi_\mathcal{D}) \\ &= \sum_{n=0}^{N}\delta t^n\sum_{K|L\in\mathcal{E}_h}\frac{m_{KL}d_{K|L}}{d}f_\mu(U^{n+1}_{K|L})M_\mu(\bar{U}^{n+1}_{K|L})\left(d\frac{\delta^{n+1}_{K,L}(P_\mathcal{D})}{d_{K|L}}\vec{n}_{KL}\right)\cdot\left(\frac{\delta^{n+1}_{K,L}(\varphi_\mathcal{D})}{d_{K|L}}\vec{n}_{KL}\right) \\ &= \sum_{n=0}^{N}\delta t^n\sum_{K|L\in\mathcal{E}_h} m_{D_{K|L}}f_\mu(U^{n+1}_{K|L})M_\mu(\bar{U}^{n+1}_{K|L})\;\nabla_{K|L}P^{n+1}_\mathcal{D}\cdot\bar{\nabla}_{K|L}\varphi^{n+1}_\mathcal{D} \\ &= \int_0^T\int_\Omega f_\mu(U_h)M_\mu(\overline{U}_h)\;\nabla_h p_\mathcal{D}\cdot\bar{\nabla}_h\varphi_\mathcal{D}\end{aligned}$$

By Lemma 6.2 and property (80), we conclude that $F1_h$ converges to $\displaystyle\int_0^T\int_\Omega k_w(u)\nabla p\cdot\nabla\varphi$. In the same way, we get

$$\begin{aligned} F2_h &= \sum_{n=0}^{N}\delta t^n\sum_{K|L\in\mathcal{E}_h}\tau_{K|L}(1-f_\mu(U_{K|L}^{n+1}))M_\mu(\bar U_{K|L}^{n+1})\delta_{K,L}^{n+1}(P_\mathcal{D})\delta_{K,L}^{n+1}(\varphi_\mathcal{D})\\ &= \int_0^T\int_\Omega (1-f_\mu(U_h))\,M_\mu(\overline{U}_h)\nabla_h P_\mathcal{D}\cdot\bar\nabla_h\varphi_\mathcal{D}\longrightarrow\int_0^T\int_\Omega k_a(u)\nabla p\cdot\nabla\varphi;\end{aligned}$$

$$\begin{aligned} G_h &= \mu\sum_{n=0}^{N}\delta t^n\sum_{K|L\in\mathcal{E}_h}\tau_{K|L}\mathfrak{d}_{K,L}^{n+1}[g(U_\mathcal{D})]\delta_{K,L}^{n+1}(\varphi_\mathcal{D}) = \mu\sum_{n=0}^{N}\delta t^n\sum_{K|L\in\mathcal{E}_h}\tau_{K|L}\sqrt{k_a(U_{K|L}^{n+1})}\delta_{K,L}^{n+1}(\zeta(U_\mathcal{D}))\delta_{K,L}^{n+1}(\varphi_\mathcal{D})\\ &= \mu\int_0^T\int_\Omega\sqrt{k_a(\overline{U}_h)}\,\nabla_h\zeta(u_\mathcal{D})\cdot\bar\nabla_h\varphi_\mathcal{D}\longrightarrow\mu\int_0^T\int_\Omega\sqrt{k_a(u)}\,\nabla\zeta(u)\cdot\nabla\varphi=\mu\int_0^T\int_\Omega\nabla g(u)\cdot\nabla\varphi.\end{aligned}$$

Finally, setting $c_\mathcal{D}=\displaystyle\sum_{n=0}^{N}\sum_{K\in\mathcal{T}_h}C_K^{n+1}\mathbb{1}_{K\times(t^n,t^{n+1}]}$ and with the analogous meaning of the notation $\overline{s}_\mathcal{D}$, $\underline{s}_\mathcal{D}$, from (78) and the L^2 consistency of the approximation of c, $\overline{s}$, $\underline{s}$ by $c_\mathcal{D}$, $\overline{s}_\mathcal{D}$, $\underline{s}_\mathcal{D}$, respectively, we readily get

$$\begin{aligned} \Sigma1_h &= \int_0^T\int_\Omega f_\mu(c_\mathcal{D})\ \overline{s}_\mathcal{D}\ \varphi_\mathcal{D}\ \longrightarrow\ \int_0^T\int_\Omega f_\mu(c)\ \overline{s}\ \varphi,\\ \Sigma2_h &= \int_0^T\int_\Omega (1-f_\mu(c_\mathcal{D}))\ \overline{s}_\mathcal{D}\ \varphi_\mathcal{D}\ \longrightarrow\ \int_0^T\int_\Omega (1-f_\mu(c))\ \overline{s}\ \varphi;\\ \sigma1_h &= \int_0^T\int_\Omega f_\mu(u_\mathcal{D})\ \underline{s}_h\ \varphi_\mathcal{D}\ \longrightarrow\ \int_0^T\int_\Omega f_\mu(u)\ \underline{s}\ \varphi,\\ \sigma2_h &= \int_0^T\int_\Omega (1-f_\mu(u_\mathcal{D}))\ \underline{s}_\mathcal{D}\ \varphi_\mathcal{D}\ \longrightarrow\ \int_0^T\int_\Omega (1-f_\mu(u))\ \underline{s}\ \varphi.\end{aligned}$$

Thus passing to the limit in (75), we justify the convergence of the scheme (up to extraction of a subsequence) to a weak solution of Problem (5)-(10).

Finally, recall the estimates (42),(43),(44) and the weak convergences proved in Section 6. Thanks to the strong convergence of $k_a(u_\mathcal{D})$ and the lower semi-continuity of L^2 norms under the weak convergence, we get the bounds (15),(16),(17). As to the translation estimate (18), we get it directly from the weak formulation of Problem (5)-(10) following the scheme of the proof of (71) (cf. [3]). The difference with the proof of Proposition 6.1 is that the quantity $f_\mu(U_{K|L}^{n+1})M_\mu(\bar U_{K|L}^{n+1})$, that was upper bounded, quite roughly, by $k_w(1)+\mu k_a(u_m)$, is now replaced by $f_\mu(u)M_\mu(u)=k_w(u)$ which is, clearly, bounded uniformly in μ. Thus (15)–(18) are established; this ends the proof of Theorem 4.1 and also establishes Theorem 2.1.

8. Numerical results

Here we present a series of numerical experiments with the implicit finite volume scheme (31)-(34) for the two-phase equation with different values of μ; we also looked at the associated explicit scheme (see [14]). In Section 8.1, for fixed values on μ we illustrate the phenomena of diffusion and injection in the two-phase model, and study the speed of convergence of the implicit scheme as the discretization step goes to zero.

In Section 8.2, the numerical results obtained for a sequence of large values of μ are compared to the numerical solution obtained by a time-implicit discretization of the Richards equation on the same mesh. For discretizing the Richards equation, we replace the second equation of the two-phase system (5)-(6) with the equation $p_c(u)-p_{atm}+p=0$, where p_{atm} is normalized by taking the value zero.

The profiles of nonlinearities are taken from the work [14] of Eymard, Henry and Hilhorst. Namely, the capillar pressure takes the form $\mathbf{p_c(s)=0.1\sqrt{1-s}}$; we set to zero the athmospheric pressure $\mathbf{p_{atm}}$; the relative permeability of the air phase is $\mathbf{k_a(s)=(1-s)^2}$ and the one of the water phase is $\mathbf{k_w(s)=\sqrt{s}}$. The source

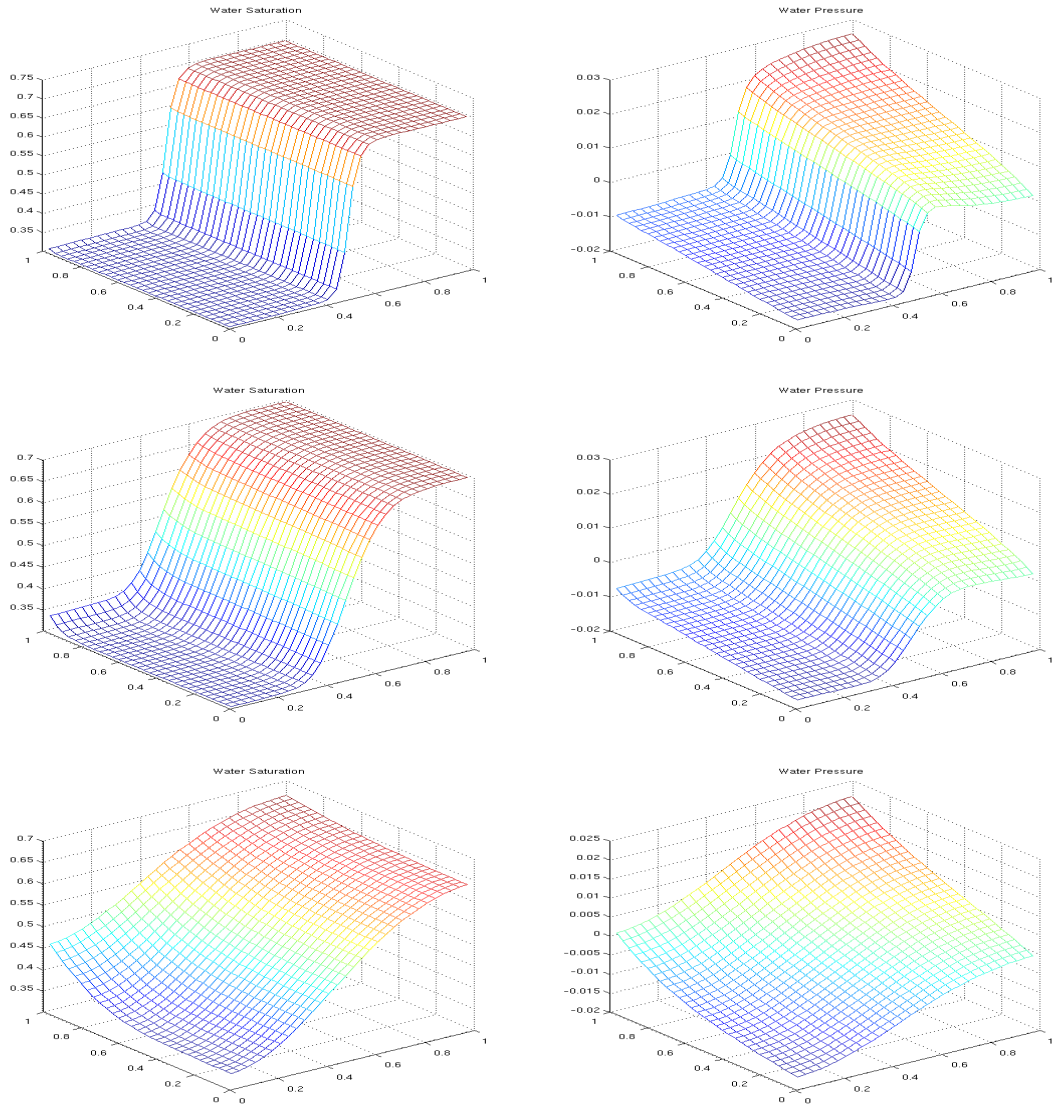

FIGURE 2. Test 1: 2D water saturation and pressure for different times: T= 0.01, 0.1 and 1

terms are different in different tests (we mainly use sources and sinks under the form of Dirac delta functions sitting at the opposite parts of the domain of calculation).

For all implicit nonlinear schemes, a Newton algorithm with time step adaptation is used (time step Δt is refined when the number of iterations exceeds 6, and it is coarsened when less than 3 iterations are needed). The initial pressure values needed for initialization of the Newton method are found by solving numerically the linear elliptic system obtained by summing the two equations with u equal to u_0. In most of the tests performed, we observe convergence of Newton algorithm in 3 to 6 iterations; the initialization step may take more iterations in case strong source and sink terms are present. Close to the saturation values, much smaller time steps were necessary in order to keep the saturation within the interval $[u_m, 1]$ and avoid divergence of the Newton method.

8.1. **Behaviour of the scheme for fixed values of mobility μ**

Test 1. Qualitative behaviour of solutions in 2D.

In this test, the domain of calculation is $[0,1]\times[0,1]$, the initial datum is $u_0(x,y) = 0.3\mathbb{1}_{[x<0.5]} + 0.7\mathbb{1}_{[x>0.5]}$, the source term is distributed uniformly at the line $\{y=1\}$ and the sink term, at the line $\{y=0\}$. The saturation of the injected fluid is $c = 0.7$. In this way, we can observe the relative importance of the two phenomena, a simple diffusion in the direction x and the displacement by injection/draining in the direction y. The strength of the source and sink terms is chosen in such a way that the two phenomena be perceptible.

In Figure 2 we give snapshots of the evolution of the system at times $T = 0.01$, $T = 0.1$ and $T = 1$; 30×30 volumes are in use.

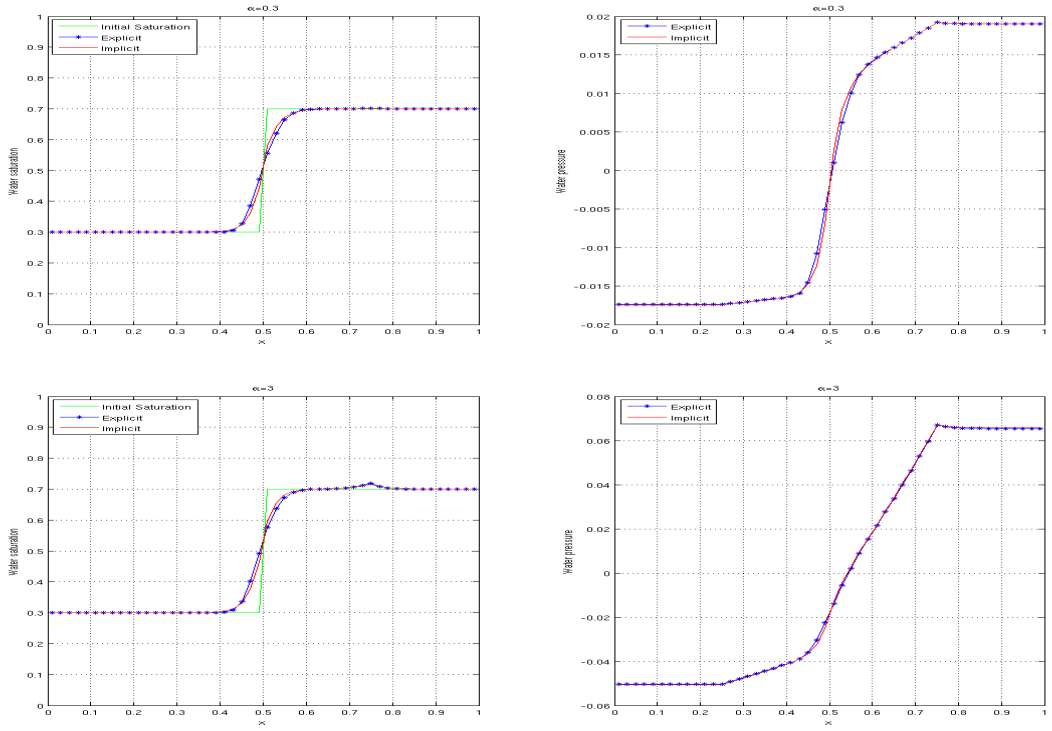

FIGURE 3. Test 2: Implicit/explicit schemes comparison with $C = 0.75$ and $\alpha = 0.3$, then $\alpha = 3$.

Test 2. Comparison of implicit and explicit schemes in 1D.

In [14], the computations were performed by an explicit scheme; in our simulations, we compared implicit and explicit scheme behaviour. In most of the experiments, we have obtained very good accordance of the numerical results obtained by the two schemes.

Yet, stability of the explicit scheme requires a CFL condition of order two ($\Delta t \leq Const\,\Delta x^2$). While the implicit *nonlinear* scheme is unconditionnally stable, in practice the linearization by the Newton algorithm works well under the order one CFL condition ($\Delta t \leq Const\,\Delta x$). In both cases, the constant $Const$ should be taken smaller when stronger source/sink terms are present. For tests on fine meshes and with strong source/sink terms, the advantage in speed of computation and robustness of the implicit scheme is very clear. At the same time, the implicit scheme appears to be slightly more diffusive. The following test uses the initial datum $u_0(x) = 0.3\mathbb{1}_{[x<0.5]} + 0.7\mathbb{1}_{[x>0.5]}$ with source term $\overline{s}(x) = \alpha\delta_0(x-0.75)$ and sink term $\underline{s}(x) = \alpha\delta_0(x-0.25)$,

where δ_0 is the Dirac delta; the source saturation $c = 0.75$ is chosen slightly larger then the initial saturation 0.7 in the injection zone; see Figures 3–5.

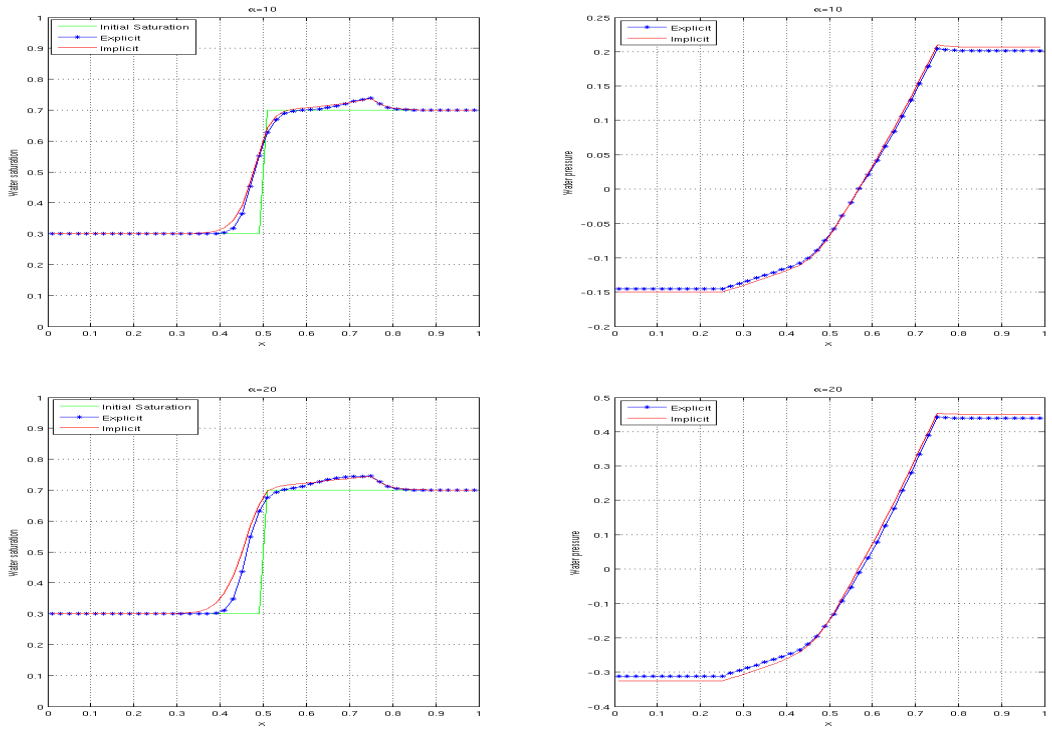

FIGURE 4. Test 2: Implicit/explicit schemes comparison with $C = 0.75$ and $\alpha = 10$, then $\alpha = 20$.

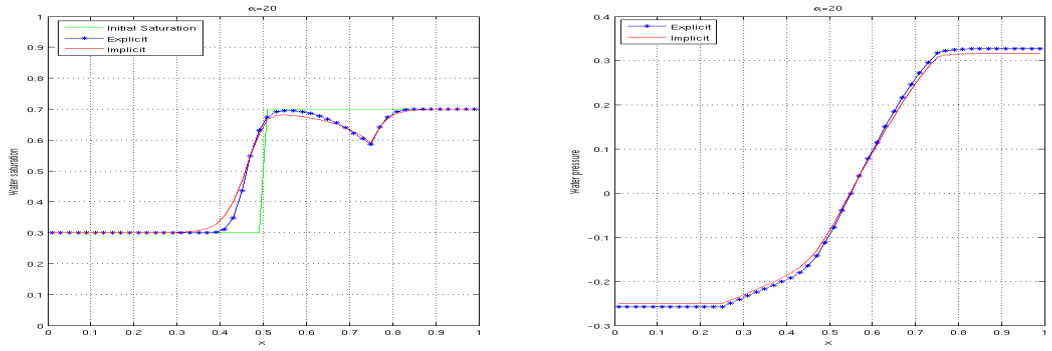

FIGURE 5. Test 2: Implicit and explicit schemes comparison with $C = 0.5$ and $\alpha = 20$.

Test 3. Orders of convergence of the implicit scheme in 1D.

The orders of convergence in Δx (Δt for the implicit scheme solved by Newton algorithm is adapted to a linear CFL condition) that we obtain for different μ (including very large ones) and for different stopping times T are

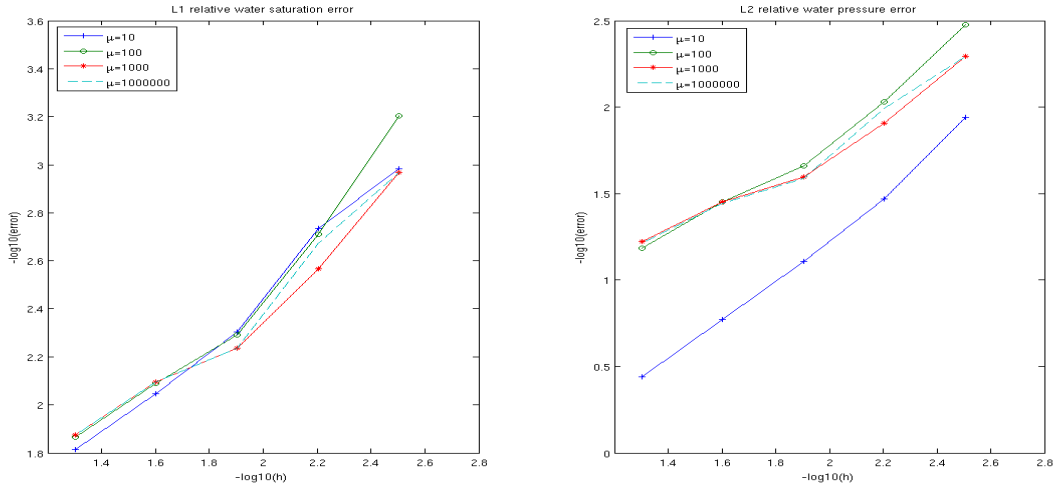

FIGURE 6. Test 3: Orders of convergence of the implicit scheme in 1D, for different μ.

very comparable. In the next fugure, we plot the saturation and pressure convergence curves (in L^1 norm for saturation, in L^2 norm for pressure); the orders observed are close to 0.9 in the two cases (the reference solution is the one computed on fine mesh, 640 points in $[0,1]$); see Figure 6. This experiment confirms the robustness of the scheme: *convergence properties are quite similar for small and large values of* μ.

8.2. **Behaviour of the scheme as $\mu \to \infty$, comparison with the Richards equation**

The main goal of this work is to justify that the Richards equation is the singular limit of the two-phase flow, at least is the *gradually saturated* regime. Indeed, if $u < 1$ in Ω the coincidence of (4) and (3) is clear. In this section, we also investigate the numerical convergence:

> *do the discrete solutions of our scheme for the two-phase flow tend, as* $\mu \to \infty$*, to the discrete solution to the Richards equation ?*

The experiments we perform exhibit the behaviour

$$\|u_{\mathcal{D}}^{\mu} - u_{\mathcal{D}}^{Rich}\|_{L^1} \approx \omega(\Delta x) + \frac{Const}{\mu}, \quad \|p_{\mathcal{D}}^{\mu} - p_{\mathcal{D}}^{Rich}\|_{L^2} \approx \omega(\Delta x) + \frac{Const}{\mu^q},$$

with q between 1 and 2 and $\omega(\Delta x) \to 0$ as $\Delta x \to 0$, in the gradually saturated regime. Indeed, the way one discretizes the Richards equation is quite different in spirit from the way the two-phase flow is discretized (as a matter of fact, the two problems have different nature). Therefore, at fixed Δx and $\mu \to \infty$ we do not observe convergence of the two-phase finite volume scheme to the scheme used for the Richards equation. But as Δx diminishes, we find better and better accordance of $\lim_{\mu\to\infty}(u_{\mathcal{D}}^{\mu}, p_{\mathcal{D}}^{\mu})$ with $(u_{\mathcal{D}}^{Rich}, p_{\mathcal{D}}^{Rich})$.

Test 4. Orders of convergence to the Richards equation.

We take the pure jump function u_0 and moderate source and sink terms located at $x = 0.75$ and at $x = 0.25$, respectively. In Figure 8.2 we plot the curves of saturation and pressure with 120 points in the domain of computation and with different values $\mu = 10^r$, $r = 1, 2, \ldots$; the last curve is computed with an implicit scheme for the Richards equation on the same mesh.

Finally, Figure 8.2 exhibits a quite accurate affine dependence of the saturation errors $\|u_{\mathcal{D}}^{\mu} - u_{\mathcal{D}}^{Rich}\|_{L^1}$ and $\|u_{\mathcal{D}}^{\mu} - u_{\mathcal{D}}^{Rich}\|_{L^\infty}$ on $1/\mu$; a residual error is observed, which is diminished if we refine the mesh.

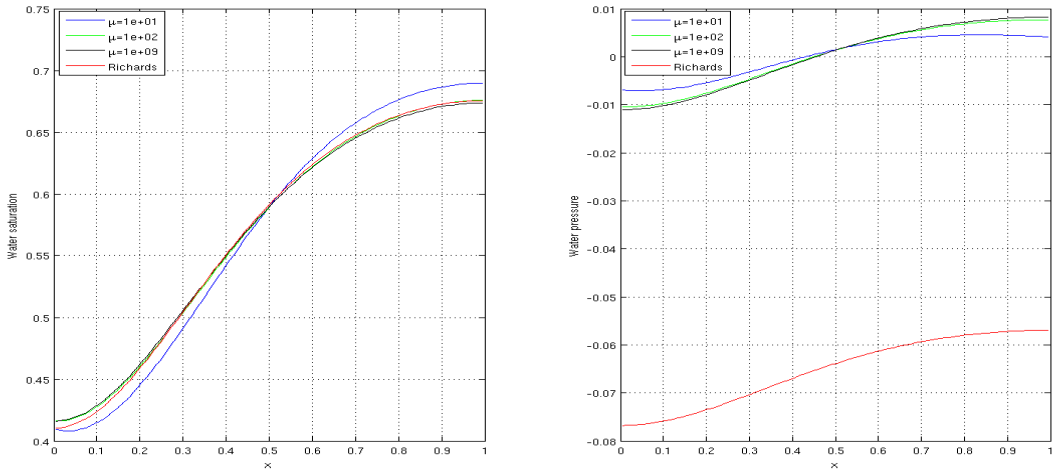

FIGURE 7. Test 4: Two-phase flow versus Richards, $T = 1$, $C = 0.7$, $\alpha = 1$.

8.3. **Conclusions from the numerical evidence**

In conclusion, for a fixed μ the finite volume scheme (31)-(34) for the two-phase flow system (5)-(6) that we have studied in this paper converges to a weak solution of the system as $\Delta x \to 0$; the rates observed are close to 1 (both for the saturation and the pressure), without apparent dependence on μ. The scheme is robust with respect to possibly very large values of μ.

While we have shown in the previous sections that $\lim_{\mu\to\infty} \lim_{\text{size}(\mathcal{D})\to 0}(u^{\mu}_{\mathcal{D}}, p^{\mu}_{\mathcal{D}}) = (u^{Rich}, p^{Rich})$ (at least in the gradually saturated regime), the scheme (31)-(34) agrees with discrete scheme for the Richards equation in gradually saturated regime as the couple $(\Delta x, \mu)$ tends to $(0, \infty)$. For a fixed Δx, a residual $(u^{Rich}_{\mathcal{D}}, p^{Rich}_{\mathcal{D}}) - \lim_{\mu\to\infty}(u^{\mu}_{\mathcal{D}}, p^{\mu}_{\mathcal{D}})$ (residual which vanishes with Δx) is observed: this is due to different mathematical nature of two-phase flow system and Richards equation and, consequently, to different strategies employed for their respective discretization.

For reasonably small Δx, and values of μ between 10^2 and 10^3 which are those observed in the practical applications in hydrogeology, we have obtained very similar numerical results while using a finite volume scheme for the Richards equation and the scheme (31)-(34) for the two-phase flow system.

REFERENCES

[1] H.W. Alt and E. Di Benedetto. Nonsteady flow of water and oil through inhomogeneous porous media. *Annali della seno la Normale Superiore di Pisa*, 12(3):335–392, 1985.
[2] H.W. Alt and S. Luckhaus. Quasilinear elliptic-parabolic differential equations. *Math-Z*, 183(3):311–341, 1983.
[3] B. Andreianov, M. Bendahmane, and K.H. Karlsen. Discrete duality finite volume schemes for doubly nonlinear degenerate hyperbolic-parabolic equations. *J. Hyp. Diff. Eq.*, 7:1–67, 2010.
[4] B. Andreianov, R. Eymard, M. Ghilani, and N. Marhraoui. On intrinsic formulation and well-posedness of a singular limit of two-phase flow equations in porous media. 2011. submitted.
[5] B. Andreianov, M. Gutnic, and P. Wittbold. Convergence of finite volume approximations for a nonlinear elliptic-parabolic problem: a "continuous" approach. *SIAM J. Numer. Anal.*, 42(1):228–251, 2004.
[6] J. Carrillo. Unicité des solutions du type kruskhov pour des problèmes elliptiques avec des termes de transport non linéaires. *C. R. Acad. Sci. Paris Sér. I Math*, 303(5):189–192, 1986.
[7] Z. Chen. Degenerate two phase flow incompressible flow 1 : existence, uniqueness and regularity of a weak solution. Smu math report 97-10, Departement of Mathematics, Southern Methodist University, 1997.
[8] Z. Chen, M. Espedal, and R. Ewing. Continuous time finite element analysis of multiphase flow in groundwater hydrology. *Appl. Math.*, (40):203–226, 1995.

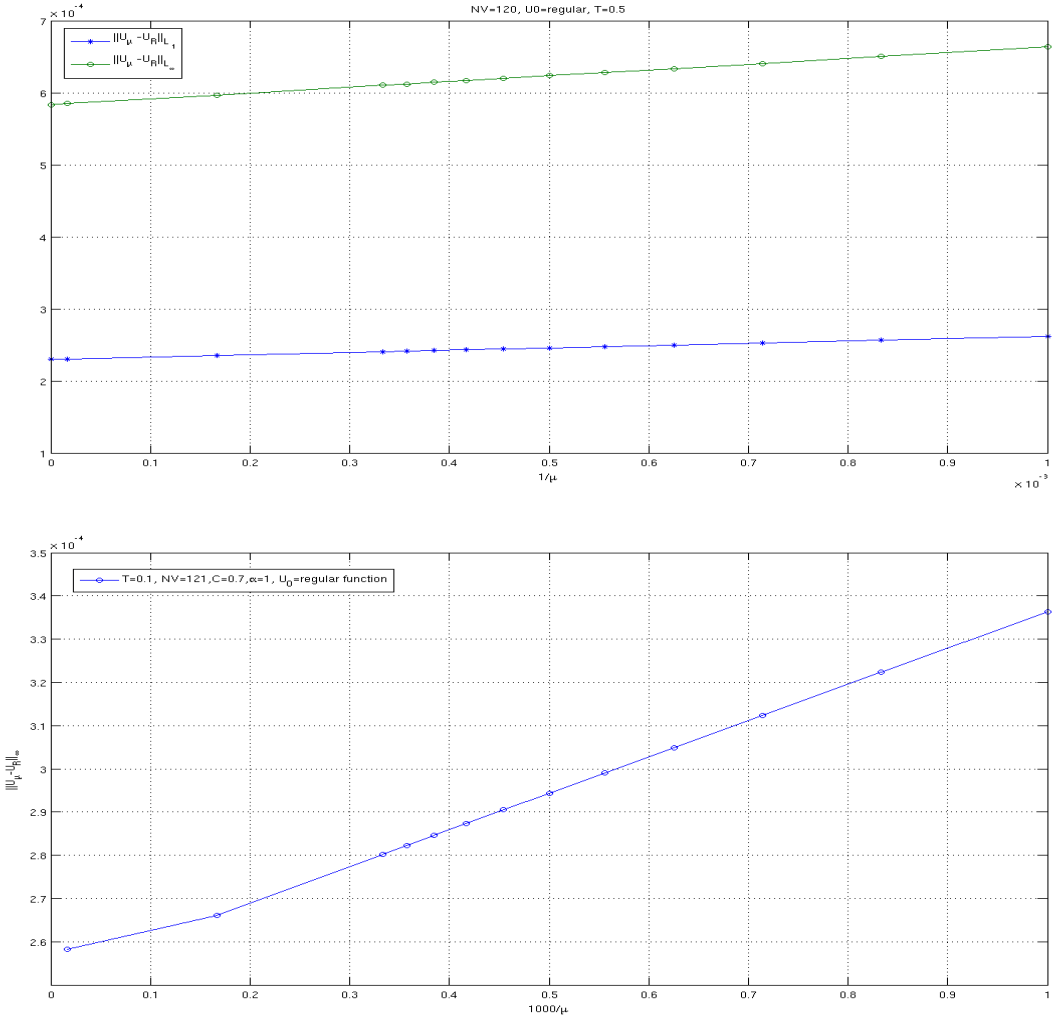

FIGURE 8. Test 4: Dependence of the errors $\|u_{\mathcal{D}}^{\mu} - u_{\mathcal{D}}^{Rich}\|_{L^1}$ and $\|u_{\mathcal{D}}^{\mu} - u_{\mathcal{D}}^{Rich}\|_{L^\infty}$ with respect to μ

[9] Z. Chen and R. Ewing. Mathematical analysis for reservoir models. *SIAM J. Math. Anal.*, 30(2):431–453, 1999.
[10] K. Deimling. *Nonlinear Functional Analysis*. Springer-Verlag, 1985.
[11] R. Eymard, T. Gallouët, and R. Herbin. *The finite volume methods*. The Handbook of Numerical Analysis, 2000.
[12] R. Eymard, T. Gallouët, R. Herbin, and A. Michel. Convergence of a finite volume scheme for nonlinear degenerate parabolic equations. *Num. Math.*, 92(1):41–82, 2002.
[13] R. Eymard, M. Ghilani, and N. Marhraoui. Convergence of two phase flow to richards model. In F. Benkhaldoun, editor, *Finite Volumes for Complex Applications IV*. ISTE, London, 2005.
[14] R. Eymard, M. Henry, and D. Hilhorst. Singular limit of a two-phase flow problem in porous medium as the air viscosity tends to zero. *Discrete Cont. Dynamical Syst. S*, 2011. to appear.
[15] R. Eymard, R. Herbin, and A. Michel. Mathematical study of a petroleum-engineering scheme. *M2AN Mathematical Modelling and Numerical Analysis*, 37(6):937–972, 2003.
[16] P. Fabrie and T. Gallouët. Modelling wells in porous media. *M3AS Math. Models Meth. Qppl. Sci.*, 10(5):673–709, 2000.
[17] G. Gagneux and M. Madaune-Tort. *Analyse Mathématique de Modèles non linéaires de l'ingénierie pétrolière, Mathématiques & Applications*, volume 22 of *SMAI*. Springer-Verlag, Berlin,1996.
[18] D. Kroener and S. Luckhaus. Flow of oil and water in a porous medium. *J. Differ. Equ.*, (55):276–288, 1984.

[19] A. Michel. *Convergence de schŕnas volumes finis pour des problèmes de convection diffusion non linéaires*. PhD thesis, Université de Provence, Marseille, France, 2001.
[20] A. Michel. A finite volume scheme for two-phase immiscible flow in porous media. *SIAM J. Numer. Anal.*, 41(4):1301–1317, 2004.
[21] H. J. Morel-Seytoux. Pour une théorie modifiée de l'infiltration. In *Cahiers O.R.S.T.O.M.*, volume X of *Hydrologie*, pages 185–194. 1973.
[22] H.J. Morel-Seytoux. Pour une théorie modifiée de l'infiltration. In *Cahiers O.R.S.T.O.M.*, volume X of *Hydrologie*, pages 199–209. 1973.
[23] F. Otto. L1-contraction and uniqueness for quasilinear elliptic-parabolic equations. *J. Differ. Equ.*, 131(1):2038, 1996.
[24] A. Plouvier-Debaight. Solutions renormalisées pour des équations autonomes des milieux poreux. *Ann. Fac. Sci. Toulouse Math.*, 6(4):727–743, 1997.
[25] A. Plouvier-Debaight, B. Donné, G. Gagneux, and P. Urruty. Solutions renormalisées pour des modèles des milieux poreux. *C. R. Acad. Sci. Paris Sér. I Math.*, 325(10):1091–1095, 1997.
[26] L. A. Richards. Capillary conduction of liquids through porous media. *Phys.*, 1:318–333, 1931.