Weierstraß-Institut
für Angewandte Analysis und Stochastik

im Forschungsverbund Berlin e.V.

Preprint ISSN 0946 – 8633

A Finite Volume Scheme for Nonlinear Parabolic Equations Derived from One-Dimensional Local Dirichlet Problems

Robert Eymard
Université Marne La Valleé
eymard@uni-mlv.fr

Jürgen Fuhrmann, Klaus Gärtner
Weierstrass Institute for Applied Analysis and Stochastics
fuhrmann@wias-berlin.de, gaertner@wias-berlin.de

submitted: September 15, 2004

No. 966
Berlin 2004

2000 *Mathematics Subject Classification.* 65M12.

Key words and phrases. Finite Volume Methods, Convergence, Nonlinear parabolic PDEs.

Edited by
Weierstraß-Institut für Angewandte Analysis und Stochastik (WIAS)
Mohrenstraße 39
10117 Berlin
Germany

Fax: + 49 30 2044975
E-Mail: preprint@wias-berlin.de
World Wide Web: http://www.wias-berlin.de/

A Finite Volume Scheme for Nonlinear Parabolic Equations Derived from One-Dimensional Local Dirichlet Problems

ROBERT EYMARD, JÜRGEN FUHRMANN, AND KLAUS GÄRTNER

Abstract. In this paper, we propose a new method to compute the numerical flux of a finite volume scheme, used for the approximation of the solution of the nonlinear partial differential equation $u_t + \mathrm{div}(\vec{q}f(u)) - \Delta\phi(u) = 0$ in a 1D, 2D or 3D domain. The function ϕ is supposed to be strictly increasing, but some values s such that $\phi'(s) = 0$ can exist. The method is based on the solution, at each interface between two control volumes, of the nonlinear elliptic two point boundary value problem $(qf(v) + (\phi(v))')' = 0$ with Dirichlet boundary conditions given by the values of the discrete approximation in both control volumes. We prove the existence of a solution to this two point boundary value problem. We show that the expression for the numerical flux can be yielded without referring to this solution. Furthermore, we prove that the so designed finite volume scheme has the expected stability properties and that its solution converges to the weak solution of the continuous problem. Numerical results show the increase of accuracy due to the use of this scheme, compared to some other schemes.

1. Introduction

The simulation of two-phase flow in porous media must be performed in different industrial settings: for example, in the petroleum engineering setting, one must approximate the solution of oil/water flow (see for example [3]); in the hydrological setting, one has to simulate air/water flow. In both cases, the equations to be considered include many mathematical features. In particular, an equation of nonlinear convective/diffusive type arises, the solution of which is the saturation of one of the two phases. This equation is obtained after eliminating the pressure of the other of the two phases. We thus study in this paper the following problem: find a numerical approximate solution to the equation

$$u_t + \mathrm{div}(\vec{q}f(u)) - \Delta\phi(u) = 0 \tag{1.1}$$

with an initial condition

$$u(x, 0) = u_0(x) \tag{1.2}$$

and a boundary condition

$$u(x, t) = \bar{u}(x, t) \tag{1.3}$$

on $\partial\Omega \times (0, T)$. We suppose that the following hypotheses are fulfilled:

1.1. **Assumption.**

(i) Ω is an open bounded connected subset of $\mathbb{R}^d$, with $d = 1, 2$ or 3,
(ii) $\bar{u}$ is the trace on $\partial\Omega \times (0, T)$ of a function, again denoted $\bar{u}$, which is assumed to satisfy $\bar{u} \in H^1(\Omega \times (0, T))$,
(iii) $u_0 \in L^2(\Omega)$,
(iv) $\vec{q} \in C^1(\bar{\Omega}, \mathbb{R}^d)$ is such that $\mathrm{div}\vec{q} = 0$,
(v) let $\phi \in C^1(\mathbb{R}, \mathbb{R})$ be Lipschitz continuous and strictly monotonically increasing
(vi) $f \in C^0(\mathbb{R}, \mathbb{R})$

Denote $U^\flat = \mathrm{ess\,inf}\,\bar{u} \bot \mathrm{ess\,inf}\, u_0$, $U^\sharp = \mathrm{ess\,sup}\,\bar{u} \top \mathrm{ess\,sup}\, u_0$.

The derivative ϕ' is understood in the distribution sense. As ϕ is Lipschitz continuous, we have that $\phi' \in L^\infty(\mathbb{R})$.

Moreover, we will write $a \top b = \max(a, b), a \bot b = \min(a, b)$.

The existence and uniqueness of a weak solution to problem (1.1)-(1.3) has been extensively studied (see [2] and [4]). Many numerical schemes have been proposed to approximate the solution of this problem. In particular, approximate solutions using Finite Volume methods have been shown to be convenient (see [10], [13], [11]) for two reasons: first, these methods are suitable for purely nonlinear hyperbolic equations, secondly they have been proved to be efficient in the case of degenerate parabolic equations. The use of these methods implies to formulate a numerical approximation G of $(\vec{q}f(u) - \nabla\phi(u)) \cdot \vec{n}$ at the interface between two grid blocks, where $\vec{n}$ is the unit vector normal to the interface between the two grid blocks, oriented from the first to the second one. We denote by a and b the respective approximate values of u in the first and in the second grid block at a given time step, by h the distance between the respective "centers" of these grid blocks. We suppose that q is an average value of $\vec{q}\cdot\vec{n}$ at the interface. A classical method consists in defining the numerical flux G by the relation

$$G = \psi(a,b,q) + \frac{\phi(a)-\phi(b)}{h},$$

where $\psi(a,b,q)$ is a numerical flux for the hyperbolic term $\mathrm{div}(\vec{q}f(u))$, which is not necessarily the same for all interfaces. Many choices can be taken for the function ψ, provided they fulfill properties of *regularity* (ψ must be continuous in order to prove the existence of a discrete solution), of *consistency* ($\psi(a,a,q) = qf(a)$ for all reals a) and of *monotonicity* ($\psi(a,b,q)$ is increasing with respect to a and decreasing with respect to b). Among others, the classical Godunov scheme [14]

$$\psi_{godunov}(a,b,q) = \begin{cases} \min\limits_{s\in[a,b]}(qf(s)) \text{ if } a\le b, \\ -\psi(b,a,-q) = \max\limits_{s\in[b,a]}(qf(s)) \text{ otherwise} \end{cases} \tag{1.4}$$

fulfill these conditions. A drawback of this method is that it does not make use of the nonlinear diffusion generated by the function ϕ in order to get, as far as possible, a less diffusive scheme. Indeed, if – in the case of Lipschitz continuous f – the following Péclet-type condition holds:

$$h\ |q|\ \operatorname{ess\,sup}\{|f'(a)|,\ a\in\mathbb{R}\} \le 2\ \operatorname{ess\,inf}\{\phi'(a),\ a\in\mathbb{R}\}, \tag{1.5}$$

the choice

$$\psi(a,b,q) = \frac{q}{2}(f(a)+f(b)) \tag{1.6}$$

can be done, leading to an accurate and stable scheme. Unfortunately, functions ϕ such as the ones resulting from actual examples of two-phase flow problems do not satisfy the existence of some $\alpha > 0$ such that $\phi'(a) > \alpha$ for a.e. a (the diffusion degenerates when one of the phases vanishes) and a local Péclet condition cannot be satisfied.

A scheme which in the linear case ($\phi = Id, f = Id$) minimizes artificial diffusion is the Il'in scheme [15] (also known as Allen-Southwell [1], Scharfetter-Gummel [18]). In this case, we express G by the relation

$$G = q\ \frac{ae^{qh}-b}{e^{qh}-1} = \frac{aB(-qh)-bB(qh)}{h} \tag{1.7}$$

where $B(x) = \frac{x}{e^x-1}$ is the Bernoulli function. It is well known that it is possible to obtain this scheme by solving a two point boundary value problem on $(0,h)$

$$\begin{cases} [-v'+qv]' &= 0 \text{ on } (0,h), \\ v(0) &= a, \\ v(h) &= b. \end{cases} \tag{1.8}$$

and taking G as the constant value between 0 and h of $-v'+qv$.

In this paper, we generalize this method to nonlinear problems like (1.1). We set $G = g(a,b,q,h)$, where the function $g(a,b,q,h)$ is defined from the solution of the following local problem:

$$\begin{cases} [-\phi(v)'+qf(v)]' &= 0 \text{ on } (0,h), \\ v(0) &= a, \\ v(h) &= b. \end{cases} \tag{1.9}$$

Then $g(a,b,q,h)$ is set to be equal to the constant value $-[\phi(v(x))]'+qf(v(x))$ for all $x\in(0,h)$. The above problem appears as a nonlinear 1D elliptic problem, with non homogeneous Dirichlet boundary conditions. We prove in this paper that the value of $g(a,b,q,h)$ is given by the following relation:

if $a \le b$ then, setting $G^\flat_{a,b} = \min\limits_{s\in[a,b]}(qf(s))$,

if $\forall\varepsilon>0$, $\displaystyle\int_a^b \frac{\phi'(s)ds}{qf(s)-G^\flat_{a,b}+\varepsilon} \le h$, then

$g(a,b,q,h) = G^\flat_{a,b}$

else

$g(a,b,q,h) = G \in (-\infty, G^\flat_{a,b})$ such that $\displaystyle\int_a^b \frac{\phi'(s)ds}{qf(s)-G} = h$

else if $a > b$ then

$g(a,b,q,h) = -g(b,a,-q,h)$. (1.10)

Note that Definition (1.10) is meaningful since it is obvious that $\displaystyle\int_a^b \frac{\phi'(s)ds}{qf(s)-G}$ is a strictly increasing expression with respect to $G \in (-\infty, G^\flat_{a,b})$ and tends to 0 as G tends to $-\infty$. Furthermore, formula (1.10) is then easy to numerically implement. We also see, in the case where $\phi' \equiv 0$, that formula (1.10) gives back the Godunov flux (1.4). However, such a situation will not be studied in this paper.

The paper is organized as follows: In section 2, we deduce properties of the function $g(a, b, q, h)$ as given by (1.10). Section 3 is devoted to the study of the boundary value problem (1.9), and its relation to the function $g(a, b, q, h)$. In both sections 2 and 3, we distinguish between the cases where $F(u) = qf \circ \phi^{-1}$ is continuously differentiable resp. continuous. In Section 4, we study the finite volume scheme based on the flux functions $g(a, b, q, h)$ and we prove its convergence. Finally, in section 5, we present some numerical examples which show the increasing of precision due to the use of this new numerical scheme.

1.2. *Remark.* It would be sufficient to assume that ϕ and f are defined on $[U^\flat, U^\sharp]$. To proceed in this case, we would need to extend the domains of definition in an appropriate way.

1.3. *Remark.* The convergence study of the case where ϕ is increasing, but not strictly increasing, could be done as well. However, in such a case, we would have to use some more complex tools [11].

2. Properties of the numerical flux

Before obtaining the results on the function $g(a, b, q, h)$ defined by (1.10), we regard the function $\mathcal{G}(F, A, B, h)$ which is given, for all $(A, B, h) \in \mathbb{R} \times \mathbb{R} \times (0, +\infty)$ and $F \in C[A, B]$, by

$$\boxed{\begin{array}{l}
\text{if } A \le B \text{ then, setting } F^\flat_{A,B} = \min\limits_{s\in[A,B]} F(s),\\
\boxed{\begin{array}{l}
\text{if } \forall\varepsilon > 0,\ \displaystyle\int_A^B \frac{ds}{F(s) - F^\flat_{A,B} + \varepsilon} \le h \text{ then}\\
\boxed{\mathcal{G}(F, A, B, h) = F^\flat_{A,B}}\\
\text{else}\\
\boxed{\mathcal{G}(F, A, B, h) = G \in (-\infty, F^\flat_{A,B}) \ \text{ such that } \displaystyle\int_A^B \frac{ds}{F(s) - G} = h}
\end{array}}\\
\text{else if } A > B \text{ then}\\
\boxed{\mathcal{G}(F, A, B, h) = -\mathcal{G}(-F, B, A, h).}
\end{array}} \tag{2.1}$$

Note again that definition (2.1) is meaningful since $\displaystyle\int_A^B \frac{ds}{F(s)-G}$ is a strictly increasing expression with respect to $G \in (-\infty, F^\flat_{A,B})$ and tends to 0 as G tends to $-\infty$.

2.1. **Lemma** (Properties of $\mathcal{G}(F, A, B, h)$ in the case where $F \in C^1[A, B]$)**.** *Let $h > 0$, $A, B \in \mathbb{R}$ with $A < B$ and let $F \in C^1[A, B]$ be given. Let $\mathcal{G}(F, A, B, h)$ be defined as in* (2.1)*.*

(i) *For fixed A, B, the function $h \mapsto \mathcal{G}(F, A, B, h)$ defined for all $h > 0$ is continuously differentiable and strictly monotonically increasing, and satisfies $\lim\limits_{h\to 0} \mathcal{G}(F, A, B, h) = -\infty$ and $\lim\limits_{h\to\infty} \mathcal{G}(F, A, B, h) = \min\limits_{s\in[A,B]} F(s)$.*

(ii) *For fixed h, the function $(\alpha, \beta) \mapsto \mathcal{G}(F, \alpha, \beta, h)$, defined for all $(\alpha, \beta) \in [A, B] \times [A, B]$, is strictly increasing with respect to α, and strictly decreasing with respect to β.*

(iii) *For fixed h, the function $(\alpha, \beta) \mapsto \mathcal{G}(F, \alpha, \beta, h)$ is Lipschitz continuous with a Lipschitz constant lower or equal to $2 \max\limits_{s\in[A,B]} |F'(s)| + \frac{1}{h}$.*

Proof. (i). Let us denote by $F^\flat_{A,B} = \min\limits_{s\in[A,B]} F(s)$. The property $\lim\limits_{G\to-\infty} \int_A^B \frac{ds}{F(s)-G} = 0$ is straightforward. Let us now prove that $\lim\limits_{G\to F^\flat_{A,B}} \int_A^B \frac{ds}{F(s)-G} = +\infty$. Let us first assume that $M = \max\limits_{s\in[A,B]} |F'(s)|$ is such that $M > 0$. Let $C \in [A, B]$ be such that $F(C) = F^\flat_{A,B} = \min\limits_{s\in[A,B]} F(s)$. Then we get for all

$s \in [A, B]$, $0 \le F(s) - F(C) \le M\ |s - C|$, and therefore

$$\begin{aligned}\int_A^B \tfrac{ds}{F(s)-G} &\ge \int_A^B \frac{ds}{M|s-C| + F(C) - G} \\ &= \int_A^C \frac{ds}{M(C-s) + F(C) - G} + \int_C^B \frac{ds}{M(s-C) + F(C) - G} \\ &= \frac{1}{M}\left(\log\left(\frac{M(C-A)}{F(C)-G} + 1\right) + \log\left(\frac{M(B-C)}{F(C)-G} + 1\right)\right)\end{aligned}$$

Since one of the two values $M(C-A)$ and $M(B-C)$ is necessarily strictly positive, it is then clear that

$$\lim_{G \to F(C)} \frac{1}{M}\left(\log\left(\frac{M(C-A)}{F(C)-G} + 1\right) + \log\left(\frac{M(B-C)}{F(C)-G} + 1\right)\right) = +\infty.$$

In the case $M = 0$, then $F(s) = F^\flat_{A,B}$ for all $s \in [A, B]$. Then

$$\int_A^B \frac{ds}{F(s) - G} = \int_A^B \frac{ds}{F^\flat_{A,B} - G} = \frac{B - A}{F^\flat_{A,B} - G},$$

and the conclusion $\lim\limits_{G \to F^\flat_{A,B}} \int_A^B \frac{ds}{F(s)-G} = +\infty$ follows.

(ii). Let us first show the monotonicity of the function $\alpha \mapsto \mathcal{G}(F, \alpha, \beta, h)$, for $\beta \in [A, B]$ given. Let $A \le \bar\alpha < \alpha < \beta \le B$ be given. Let us denote by $G = \mathcal{G}(F, \alpha, \beta, h)$ and $\bar G = \mathcal{G}(F, \bar\alpha, \beta, h)$, which means in this case

$$\int_{\bar\alpha}^\beta \frac{ds}{F(s) - \bar G} = h \tag{2.2}$$

and

$$\int_\alpha^\beta \frac{ds}{F(s) - G} = h. \tag{2.3}$$

Subtracting (2.3) to (2.2) gives

$$\int_{\bar\alpha}^\alpha \frac{ds}{F(s) - \bar G} + \int_\alpha^\beta \frac{\bar G - G}{(F(s) - G)(F(s) - \bar G)} ds = 0,$$

which proves that

$$G - \bar G = \left(\int_{\bar\alpha}^\alpha \frac{ds}{F(s) - \bar G}\right) / \left(\int_\alpha^\beta \frac{ds}{(F(s) - G)(F(s) - \bar G)}\right),$$

and therefore $G - \bar G > 0$. We thus get that $\mathcal{G}(F, \alpha, \beta, h)$ is strictly increasing with respect to α for $\alpha < \beta$. A similar argument shows that $\mathcal{G}(F, \alpha, \beta, h)$ is strictly decreasing with respect to β.

(iii). Let now $A \le \bar\alpha < \alpha < \beta < \bar\beta \le B$ be given. Let us denote by $\bar G = \mathcal{G}(F, \bar\alpha, \bar\beta, h)$. The later means in this case

$$\int_{\bar\alpha}^{\bar\beta} \frac{ds}{F(s) - \bar G} = h. \tag{2.4}$$

We note that, by monotonicity, $\bar G < G$. Let us define $\bar F : [\alpha, \beta] \to \mathbb{R}$, $\bar s \mapsto \bar F(\bar s) = F\left(\bar\alpha + \frac{\bar\beta - \bar\alpha}{\beta - \alpha}(\bar s - \alpha)\right)$. Since $\bar G$ satisfies (2.4), we get, thanks to the change of variable $s = \bar\alpha + \frac{\bar\beta-\bar\alpha}{\beta-\alpha}(\bar s - \alpha)$,

$$\frac{\bar\beta - \bar\alpha}{\beta - \alpha} \int_\alpha^\beta \frac{d\bar s}{\bar F(\bar s) - \bar G} = h. \tag{2.5}$$

After replacing the symbol $\bar s$ by s in (2.5), we subtract the above relation from (2.3). We get

$$\int_\alpha^\beta \frac{(\beta - \alpha)(\bar F(s) - \bar G) - (\bar\beta - \bar\alpha)(F(s) - G)}{(F(s) - G)(\bar F(s) - \bar G)} ds = 0.$$

This gives

$$\int_\alpha^\beta \frac{(\bar\beta - \bar\alpha)(\bar F(s) - \bar G - F(s) + G) + (\bar\alpha - \alpha - \bar\beta + \beta)(\bar F(s) - \bar G)}{(F(s) - G)(\bar F(s) - \bar G)} ds = 0,$$

leading to

$$G - \bar G = \frac{\displaystyle\int_\alpha^\beta \frac{F(s) - \bar F(s) + \frac{\alpha - \bar\alpha - \beta + \bar\beta}{\bar\beta - \bar\alpha}(\bar F(s) - \bar G)}{(F(s) - G)(\bar F(s) - \bar G)} ds}{\displaystyle\int_\alpha^\beta \frac{ds}{(F(s) - G)(\bar F(s) - \bar G)}}.$$

Since $F(s) - G > 0$ and $\bar{F}(s) - \bar{G} > 0$ for all $s \in [\alpha, \beta]$, we get

$$G - \bar{G} \leq \max_{s \in [\alpha,\beta]} |F(s) - \bar{F}(s)| + \frac{|\alpha - \bar{\alpha}| + |\beta - \bar{\beta}|}{\bar{\beta} - \bar{\alpha}} \left(\max_{s \in [\alpha,\beta]} \bar{F}(s) - \bar{G} \right).$$

We first remark that $\max_{s \in [\alpha,\beta]} \bar{F}(s) = \max_{s \in [\bar{\alpha},\bar{\beta}]} F(s)$ and that, using the mean value theorem in (2.2), there exists $\bar{\gamma} \in [\bar{\alpha}, \bar{\beta}]$ such that

$$\int_{\bar{\alpha}}^{\bar{\beta}} \frac{ds}{F(s) - \bar{G}} = \frac{\bar{\beta} - \bar{\alpha}}{F(\bar{\gamma}) - \bar{G}} = h,$$

which produces

$$\bar{G} = F(\bar{\gamma}) - \frac{\bar{\beta} - \bar{\alpha}}{h}.$$

We thus get

$$G - \bar{G} \leq \max_{s \in [\alpha,\beta]} |F(s) - \bar{F}(s)| + \frac{\alpha - \bar{\alpha} - \beta + \bar{\beta}}{\bar{\beta} - \bar{\alpha}} \left(\max_{s \in [\bar{\alpha},\bar{\beta}]} F(s) - F(\bar{\gamma}) + \frac{\bar{\beta} - \bar{\alpha}}{h} \right).$$

Since $M = \max_{s \in [A,B]} |F'(s)|$, we have, for all $s \in [\alpha, \beta]$,

$$\begin{aligned} |F(s) - \bar{F}(s)| &= \left| F(s) - F\left(\bar{\alpha} + \frac{\bar{\beta} - \bar{\alpha}}{\beta - \alpha}(s - \alpha) \right) \right| \\ &\leq \frac{M}{\beta - \alpha} \big((\beta - \alpha)(s - \bar{\alpha}) - (\bar{\beta} - \bar{\alpha})(s - \alpha) \big) \\ &\leq M \max(|\alpha - \bar{\alpha}|, |\beta - \bar{\beta}|) \\ &\leq M(|\alpha - \bar{\alpha}| + |\beta - \bar{\beta}|) \end{aligned}$$

and

$$\max_{s \in [\bar{\alpha},\bar{\beta}]} F(s) - F(\bar{\gamma}) \leq M(\bar{\beta} - \bar{\alpha}).$$

Gathering the three above inequalities, we get

$$G - \bar{G} \leq M(|\alpha - \bar{\alpha}| + |\beta - \bar{\beta}|) + \frac{|\alpha - \bar{\alpha}| + |\beta - \bar{\beta}|}{\bar{\beta} - \bar{\alpha}} (\bar{\beta} - \bar{\alpha})(M + \frac{1}{h}),$$

which gives

$$G - \bar{G} \leq (2\,M + \frac{1}{h})(|\alpha - \bar{\alpha}| + |\beta - \bar{\beta}|).$$

Let us now consider the case where $A \leq \bar{\alpha} < \alpha = \beta < \bar{\beta} \leq B$. In this case, the relations $G = F(\alpha) = F(\beta)$ and $\bar{G} = F(\bar{\gamma}) - \frac{\bar{\beta} - \bar{\alpha}}{h}$, with $\bar{\gamma} \in [\bar{\alpha}, \bar{\beta}]$, lead to

$$\begin{aligned} G - \bar{G} &= F(\alpha) - F(\bar{\gamma}) + \frac{\bar{\beta} - \bar{\alpha}}{h} \\ &\leq M(\bar{\beta} - \bar{\alpha}) + \frac{\bar{\beta} - \bar{\alpha}}{h} \\ &= M(\bar{\beta} - \beta + \alpha - \bar{\alpha}) + \frac{\bar{\beta} - \beta + \alpha - \bar{\alpha}}{h} \\ &\leq (M + \frac{1}{h})(|\alpha - \bar{\alpha}| + |\beta - \bar{\beta}|). \end{aligned}$$

This suffices to prove that $(\alpha, \beta) \mapsto \mathcal{G}(F, \alpha, \beta, h)$ is Lipschitz continuous with the constant $2\,M + \frac{1}{h}$ for all $\alpha, \beta \in [A, B]$ with $\alpha \leq \beta$. Since, for all $\alpha \geq \beta$, $\mathcal{G}(F, \alpha, \beta, h) = -\mathcal{G}(-F, \beta, \alpha, h)$, the same proofs hold with $-F$ instead of F, leading to the same Lipschitz constant $2\,M + \frac{1}{h}$ for all $\alpha, \beta \in [A, B]$. □

Now, we examine the case of a weaker hypothesis on F, namely $F \in C[0, h]$ instead of $F \in C^1[0, h]$.

2.2. **Lemma** (Regularity of $\mathcal{G}(F, A, B, h)$ with respect to F)**.** *Let $h > 0$, $A, B \in \mathbb{R}$ with $A \leq B$ and let F_1, $F_2 \in C[A, B]$. Then*

$$|\mathcal{G}(F_1, A, B, h) - \mathcal{G}(F_2, A, B, h)| \leq \max_{s \in [A,B]} |F_1(s) - F_2(s)|. \tag{2.6}$$

Proof. Let us denote, for $i = 1, 2$, by $F_i^\flat = \min_{s\in[A,B]} F_i(s)$ and $G_i = \mathcal{G}(F_i, A, B, h)$. We first assume that for both values $i = 1, 2$, $G_i < F_i^\flat$ holds, which leads to $\int_A^B \frac{ds}{F_i(s)-G_i} = h$. Since we have

$$\int_A^B \frac{ds}{F_1(s) - G_1} - \int_A^B \frac{ds}{F_2(s) - G_2} = \int_A^B \frac{F_2(s) - F_1(s) + G_1 - G_2}{(F_1(s) - G_1)(F_2(s) - G_2)} ds = 0,$$

we get

$$G_1 - G_2 = \frac{\displaystyle\int_A^B \frac{F_1(s) - F_2(s)}{(F_1(s) - G_1)(F_2(s) - G_2)} ds}{\displaystyle\int_A^B \frac{1}{(F_1(s) - G_1)(F_2(s) - G_2)} ds},$$

which gives (2.6), thanks to $(F_1(s) - G_1)(F_2(s) - G_2) > 0$ for all $s \in [A, B]$. Let us now assume that $G_1 < F_1^\flat$ and that $G_2 = F_2^\flat$, which means, thanks to Definition (2.1), that for any $\varepsilon > 0$,

$$\int_A^B \frac{ds}{F_2(s) - F_2^\flat + \varepsilon} \le h.$$

The above inequality gives

$$\int_A^B \frac{ds}{F_1(s) - G_1} - \int_A^B \frac{ds}{F_2(s) - F_2^\flat + \varepsilon} = \int_A^B \frac{F_2(s) - F_1(s) + G_1 - F_2^\flat + \varepsilon}{(F_1(s) - G_1)(F_2(s) - F_2^\flat + \varepsilon)} ds \ge 0,$$

which produces

$$G_1 - F_2^\flat + \varepsilon \ge \frac{\displaystyle\int_A^B \frac{F_1(s) - F_2(s)}{(F_1(s) - G_1)(F_2(s) - F_2^\flat + \varepsilon)} ds}{\displaystyle\int_A^B \frac{1}{(F_1(s) - G_1)(F_2(s) - F_2^\flat + \varepsilon)} ds}.$$

This proves that

$$G_1 \ge F_2^\flat - \varepsilon - \max_{s\in[A,B]} |F_1(s) - F_2(s)|.$$

Since the above inequality holds for all $\varepsilon > 0$ and since $G_1 < F_1^\flat \le F_2^\flat + \max_{s\in[A,B]} |F_1(s) - F_2(s)|$, we get

$$F_2^\flat + \max_{s\in[A,B]} |F_1(s) - F_2(s)| > G_1 \ge F_2^\flat - \max_{s\in[A,B]} |F_1(s) - F_2(s)|,$$

which also leads to (2.6). The case $G_i = F_i^\flat$ for $i = 1, 2$ results from $F_1^\flat \le F_2^\flat + \max_{s\in[A,B]} |F_1(s) - F_2(s)|$ and $F_2^\flat \le F_1^\flat + \max_{s\in[A,B]} |F_1(s) - F_2(s)|$. This completes the proof of (2.6) in the general case. □

2.3. **Lemma** (Properties of $\mathcal{G}(F, A, B, h)$ in the case where $F \in C[A, B]$)**.** *Let $F \in C[A, B]$. We regard the function $(\alpha, \beta) \mapsto \mathcal{G}(F, \alpha, \beta, h)$, defined for all $(\alpha, \beta) \in [A, B] \times [A, B]$.*

(i) *It is increasing with respect to α, and decreasing with respect to β.*
(ii) *It is continuous.*
(iii) *There exists $\gamma \in [\alpha\perp\beta, \alpha\top\beta]$ such that*

$$\mathcal{G}(F, \alpha, \beta, h) = F(\gamma) - \frac{\beta - \alpha}{h}. \tag{2.7}$$

Proof. (i). Consider, for $F \in C[A, B]$, a sequence of functions $F_n \in C^1[A, B]$ which uniformly converge to F. Applying Lemma 2.1 and (2.6), we get that for all $(\alpha, \beta) \in [A, B] \times [A, B]$, $|\mathcal{G}(F, \alpha, \beta, h) - \mathcal{G}(F_n, \alpha, \beta, h)| \le \max_{s\in[A,B]} |F(s) - F_n(s)|$, which proves that the sequence of functions $(\alpha, \beta) \mapsto \mathcal{G}(F_n, \alpha, \beta, h)$ uniformly converges to the function $(\alpha, \beta) \mapsto \mathcal{G}(F, \alpha, \beta, h)$. Hence the function $(\alpha, \beta) \mapsto \mathcal{G}(F, \alpha, \beta, h)$, is increasing with respect to α, and decreasing with respect to β.

(ii). The function is continuous as the limit of a sequence of uniformly converging continuous functions.

(iii). Let us denote $G = \mathcal{G}(F, \alpha, \beta, h)$. Let us assume that $\alpha \le \beta$ and that $\int_\alpha^\beta \frac{ds}{F(s)-G} = h$. Then the mean value theorem gives that there exists $\gamma \in [\alpha, \beta]$ with

$$\int_\alpha^\beta \frac{ds}{F(s) - G} = (\beta - \alpha)\frac{1}{F(\gamma) - G} = h,$$

which gives (2.7). Otherwise, let us suppose that for all $\varepsilon > 0$,

$$\int_\alpha^\beta \frac{ds}{F(s) - G + \varepsilon} \le h.$$

We then have that there exists $\gamma_\varepsilon \in [\alpha, \beta]$ such that

$$\int_\alpha^\beta \frac{ds}{F(s) - G} = (\beta - \alpha)\frac{1}{F(\gamma_\varepsilon) - G + \varepsilon},$$

and therefore

$$G \leq -\frac{\beta - \alpha}{h} + F(\gamma_\varepsilon) + \varepsilon.$$

Extracting a sequence from γ_ε which converges as $\varepsilon \to 0$, we get that there exists $\gamma_0 \in [\alpha, \beta]$ with

$$G \leq -\frac{\beta - \alpha}{h} + F(\gamma_0).$$

Since $G = \min_{s\in[\alpha,\beta]} F(s)$, we get

$$\min_{s\in[\alpha,\beta]} F(s) \leq \min_{s\in[\alpha,\beta]} F(s) + \frac{\beta - \alpha}{h} \leq F(\gamma_0).$$

Therefore there exists $\gamma \in [\alpha, \beta]$ such that

$$F(\gamma) = \min_{s\in[\alpha,\beta]} F(s) + \frac{\beta - \alpha}{h},$$

which again gives (2.7) in this case. Finally, the case $\beta \leq \alpha$ is similar. □

We can now obtain the results on the function g defined by (1.10), which will be sufficient to derive the convergence of the numerical scheme in section 4.

Note that Definitions (1.10) and (2.1) meet the property $g(a, b, q, h) = \mathcal{G}(qf \circ \phi^{-1}, \phi(a), \phi(b), h)$ since for all $G < F^\flat_{A,B} = G^\flat$,

$$\int_a^b \frac{\phi'(s)ds}{qf(s) - G} = \int_A^B \frac{dt}{F(t) - G}.$$

Indeed, thanks to the following Lemma 2.4, we can apply the change of variable $t = \phi(s)$.

2.4. **Lemma.** *(Change of variable) Let $a, b \in \mathbb{R}$ be given with $a \leq b$, and let $w \in C[a, b]$ be a Lipschitz continuous monotonous function. Let A, B be defined by $A = w(a)$ and $B = w(b)$. Let $f \in C[A, B]$ be given. Then the identity*

$$\int_a^b f(w(x))w'(x)dx = \int_A^B f(s)ds \tag{2.8}$$

holds.

Proof. Let us suppose that $w' \geq 0$ a.e. in (a, b) and that $B > A$ (otherwise, the function w is a constant, and then (2.8) holds). Let $\xi_\varepsilon \in C[a, b]$, for all $\varepsilon > 0$, be a strictly positive function such that $\|\xi_\varepsilon - w'\|_{L^1(a,b)} \to 0$ as $\varepsilon \to 0$ and $\int_a^b \xi_\varepsilon(s)ds = B - A$ (it suffices to prolong w' by periodicity on $\mathbb{R}$, to regularize with a mollifier, to add $\varepsilon(B - A)/(b - a)$ and multiply by $1/(1 + \varepsilon)$). Let the function w_ε be defined by $w_\varepsilon(x) = A + \int_a^x \xi_\varepsilon(s)ds$. Then the function w_ε is a diffeomorphism from $[a, b]$ to $[A, B]$ such that w_ε uniformly converges to w in $C(a, b)$ and w'_ε converges to w' in $L^1(a, b)$ as $\varepsilon \to 0$. Then the change of variable $y = w_\varepsilon(x)$ is possible, and the equality

$$\int_a^b f(w_\varepsilon(x))w'_\varepsilon(x)dx = \int_A^B f(s)ds$$

holds. Thanks to the dominated convergence theorem of Lebesgue and to the continuity of f, we get

$$\int_a^b f(w(x))w'(x)dx = \lim_{\varepsilon\to0} \int_a^b f(w_\varepsilon(x))w'_\varepsilon(x)dx = \lim_{\varepsilon\to0} \int_A^B f(s)ds = \int_A^B f(s)ds$$

resulting in (2.8). □

2.5. **Lemma** (Properties of $g(a, b, q, h)$)**.** *Let $a, b, q \in \mathbb{R}$ and let $h \in (0, +\infty)$ be given. Let $f \in C(\mathbb{R})$ be given and let $\phi \in C(\mathbb{R})$ be a strictly increasing Lipschitz continuous function. Let $g(a, b, q, h)$ be defined by* (1.10)*. The following properties are then available.*

(i) *The function g is continuous with respect to (a, b), and it is (not strictly) increasing with respect to a and (not strictly) decreasing with respect to b.*

(ii) *There exists $c \in [a\bot b, a\top b]$ such that*

$$g(a, b, q, h) = -\frac{\phi(b) - \phi(a)}{h} + qf(c), \tag{2.9}$$

(iii) *The following inequality holds.*

$$(a-b)g(a,b,q,h) \geq \frac{(\zeta(a)-\zeta(b))^2}{h} + q\int_a^b f(s)ds, \tag{2.10}$$

where $\zeta \in C(\mathbb{R})$ *is a Lipschitz continuous function such that, for a.e.* $s \in \mathbb{R}$, $\zeta'(s) = \sqrt{\phi'(s)}$.

Proof. The first item is a consequence of $g(a,b,q,h) = \mathcal{G}(qf\circ\phi^{-1}, \phi(a), \phi(b), h)$, and of Lemma 2.3. We then set $F = qf\circ\phi^{-1}$, $A = \phi(a)$, $B = \phi(b)$ and $G = g(a,b,q,h) = \mathcal{G}(F,A,B,h)$. The second item is a consequence of (2.7) and of $F = qf\circ\phi^{-1}$. The third item can be proved by the following method. For all $\varepsilon > 0$, we have,

$$\int_a^b \frac{\phi'(s)ds}{q\ f(s) - G + \varepsilon} \leq h.$$

From the Cauchy-Schwarz inequality, we get, for all function $\alpha(s) > 0$

$$\left(\int_a^b \zeta'(s)ds\right)^2 \leq \left(\int_a^b \frac{\zeta'(s)^2 ds}{\alpha(s)}\right)\left(\int_a^b \alpha(s)ds\right).$$

We choose $\alpha(s) = q\ f(s) - G + \varepsilon$. It leads to

$$\left(\int_a^b \zeta'(s)ds\right)^2 \leq \left(\int_a^b \frac{\phi'(s)ds}{q\ f(s) - G + \varepsilon}\right)\left(\int_a^b (q\ f(s) - G + \varepsilon)ds\right),$$

and therefore

$$(\zeta(b)-\zeta(a))^2 \leq h\left(\int_a^b q\ f(s)ds - G(b-a) + \varepsilon(b-a)\right).$$

Letting ε tend to 0 in the above inequality, we get

$$(\zeta(b)-\zeta(a))^2 \leq h\left(\int_a^b q\ f(s)ds - G(b-a)\right),$$

which is (2.10). □

2.6. *Remark.* As an immediate consequence of (ii), we get, using Definition (1.4), that, for $a \leq b$,

$$g(a,b,q,h) \geq \psi_{godunov}(a,b,q) - \frac{\phi(b)-\phi(a)}{h},$$

and for $a \geq b$,

$$g(a,b,q,h) \leq \psi_{godunov}(a,b,q) - \frac{\phi(b)-\phi(a)}{h}.$$

3. Study of a 1D steady nonlinear elliptic problem

First, we transform Problem (1.9), setting $A = \phi(a)$, $B = \phi(B)$ and $F = qf\circ\phi^{-1}$ and changing the unknown v in $\phi(v)$. Our objective is then the following: for $A < B$, $h > 0$ and for a given function $F \in C[A,B]$, prove that any function $v \in C^1[0,h]$ solution to the problem

$$\begin{cases} (-v(x)' + F(v(x)))' &= 0,\ \forall x \in (0,h) \\ v(0) = A \\ v(h) = B, \end{cases} \tag{3.1}$$

in the sense that it satisfies that there exists $G \in \mathbb{R}$ such that

$$\begin{cases} -v(x)' + F(v(x)) &= G,\ \forall x \in [0,h] \\ v(0) = A \\ v(h) = B, \end{cases} \tag{3.2}$$

is such that $G = \mathcal{G}(F,A,B,h)$ as defined in (2.1).

First, we will consider the more obvious case of continuously differentiable F (lemmas 3.1, 3.2). A limit process then leads to results for the case of continuous F (lemma 3.4).

Hence we consider the auxiliary problem

$$\begin{cases} -v''(x) + p(x)v'(x) &= 0 \quad \text{on} \quad (0,h) \\ v(0) &= A \\ v(h) &= B. \end{cases} \tag{3.3}$$

3.1. **Lemma** (Existence of solution of (3.3)). *Let $h > 0$, $A, B \in \mathbb{R}$ with $A \le B$ and let $p \in C[0,h]$. Then Problem* (3.3) *has a unique solution $v \in C^2[0,h]$ which is strictly monotonically increasing and verifies*

$$\frac{B-A}{h}\exp(-2Mh) \le v'(x) \le \frac{B-A}{h}\exp(2Mh), \tag{3.4}$$

denoting $M = \max_{x\in[0,h]} |p(x)|$.

Proof. Let us define the functions $e, E\ :\ [0,h] \to \mathbb{R}$, by

$$e(x) = \exp\left(\int_0^x p(\eta)d\eta\right) \quad \text{and} \quad E(x) = \int_0^x e(\xi)d\xi,\ \forall x \in [0,h].$$

A standard integration of the above problem gives

$$v(x) = A + (B-A)\frac{E(x)}{E(h)} \text{ and } v'(x) = (B-A)\frac{e(x)}{E(h)},\ \forall x \in [0,h]. \tag{3.5}$$

In order to obtain (3.4), we note that $\exp(-Mh) \le e(x) \le \exp(Mh)$, and $h\exp(-Mh) \le E(x) \le h\exp(Mh)$. □

3.2. **Lemma** (Existence of a solution of (3.2) for $F \in C^1[A,B]$). *Let $h > 0$, $A, B \in \mathbb{R}$ with $A \le B$ and let $F \in C^1[A,B]$. Let us denote by $M = \max_{s\in[A,B]} |F'(s)|$. Then Problem* (3.1) *has a unique, strictly monotonically increasing solution $v \in C^2[0,h]$ such that* (3.4) *is satisfied. Furthermore, the constant $G \in \mathbb{R}$ such that $G = -v(x)' + F(v(x))$ for all $x \in [0,h]$ is such that $G = \mathcal{G}(F,A,B,h)$.*

Proof. Let K be the closed convex subset of $C[0,h]$ defined by $K = \{u \in C[0,h],\ u(x) \in [A,B],\ \forall x \in [0,h]\}$. In order to prove the existence of a solution, we consider the mapping $L\ :\ K \to C^2[0,h]$, such that for all $w \in K$, $v = L(w)$ is the solution of problem (3.3) with $p \in C[0,h]$ defined by $p\ :\ x \mapsto F'(w(x))$ (the existence and uniqueness of $L(w)$ is given by Lemma 3.1). Let us remark that, since F' is uniformly continuous on $[0,h]$, for a given $\varepsilon > 0$, there exists $\eta > 0$ such that, for all $\tilde{w} \in K$ with $\max_{x\in[0,h]} |w(x) - \tilde{w}(x)| \le \eta$, $\max_{x\in[0,h]} |F'(w(x)) - F'(\tilde{w}(x))| \le \varepsilon$. We then get, thanks to (3.5), that the corresponding functions e, E, $\tilde{e}$, $\tilde{E}$ satisfy

$$\max_{x\in[0,h]} |e(x) - \tilde{e}(x)| \le \exp(Mh)h\varepsilon$$

and thus

$$\max_{x\in[0,h]} |E(x) - \tilde{E}(x)| \le \exp(Mh)h^2\varepsilon,$$

which gives

$$\max_{x\in[0,h]} |L(w)(x) - L(\tilde{w})(x)| \le (B-A)2h\exp(4Mh)\varepsilon,$$

which proves that $L : C[0,h] \to C[0,h]$ is continuous. Since $L(K)$ is bounded in $C^1[0,h]$, it is thus compact in $C[0,h]$ by the Ascoli theorem [16]. Thus by Schauder's fix point theorem, L has a fix point, denoted v, which therefore satisfies $v = L(v)$ and thus $v \in C^2[0,h]$ and v is a solution of (3.1). Moreover, the conclusion of Lemma 3.1 applies to v solution of (3.3) with $p = F'(v)$, and thus we get that this function v satisfies (3.4) with $M = \max_{s\in[A,B]} |F'(s)|$. Integrating (3.1) gives the existence of the constant G such that, for all $x \in [0,h]$,

$$G = -v(x)' + F(v(x)),\ \forall x \in [0,h]. \tag{3.6}$$

Thanks to (3.4), we get that $F(v(x)) - G \ge \frac{B-A}{h}\exp(-2Mh) > 0$ for all $x \in [0,h]$. Integrating (3.6) yields

$$\int_0^h \frac{v(x)'}{F(v(x)) - G}dx = h. \tag{3.7}$$

Thanks to (3.4), v is a diffeomorphism, which allows for the change of variable $s = v(x)$ in (3.7). This leads to the relation

$$\int_A^B \frac{ds}{F(s) - G} = h, \tag{3.8}$$

which proves that $G = \mathcal{G}(F,A,B,h)$.

Finally, we have to prove uniqueness of the solution. Assume, that v and w both are solutions of (3.1). As $F \in C^1[A,B]$, the mean value theorem allows us to obtain a continuous function $\alpha(x)$

such that $F(v) - F(w) = \alpha(x)(v - w)$. Further, by (3.8), the flux G for both solutions is equal. Therefore, we obtain that for $u = v - w$, the following holds:

$$\begin{cases} -u' + \alpha(x)u &= 0 \\ u(0) &= 0 \\ u(1) &= 0 \end{cases}$$

Standard integration of this problem yields $u = 0$, and thus, $v = w$. □

In the case where F is only continuous, let us consider the following example:

3.3. *Example.*

$$\begin{cases} -u' + \sqrt{1 - u^2} &= G \quad \text{in} \quad (0, h) \\ u(0) &= -1 \\ u(h) &= 1. \end{cases}$$

The differential equation has solutions -1, 1 and $\sin(x - x_0)$ with $G = 0 = F^\flat$. These can be glued together in order to fulfill the boundary conditions. Namely, for $h > \pi$, we get the solutions

$$u(x) = \begin{cases} -1, & x \in (0, x_0 - \frac{\pi}{2}] \\ \sin(x - x_0), & x \in (x_0 - \frac{\pi}{2}, x_0 + \frac{\pi}{2}) \\ 1, & x \in [x_0 + \frac{\pi}{2}, h) \end{cases}$$

for any $x_0 \in (\frac{\pi}{2}, h - \frac{\pi}{2})$, showing that this problem has an infinite number of solutions.

Example 3.3 indicates the way to extend the results from the case of continuously differentiable F to the case of continuous F. Problem (3.1) must be taken in the sense of Problem (3.2). The following lemma states that Problem (3.2) has at least one solution which is necessarily monotone (not strictly monotone) and which belongs to $C^1[0, h]$ (not to $C^2[0, h]$), and that a uniqueness property is available for G, but not for u.

3.4. **Lemma** (Existence of a solution of (3.2) for $F \in C[A, B]$)**.** *Let $h > 0$, $A, B \in \mathbb{R}$ with $A \le B$ and let $F \in C[A, B]$. Let us denote by $F^\flat_{A,B} = \min_{s \in [A,B]} F(s)$. Then the following statements hold:*

(i) *Problem (3.2) has at least one solution $(v, G) \in C^1[0, h] \times (-\infty, F^\flat_{A,B}]$ with $v'(x) \ge 0$ for all $x \in [0, h]$.*
(ii) *For any solution (v, G) of (3.2), G is the unique value given by $G = \mathcal{G}(F, A, B, h)$ defined by (2.1).*

Proof. Let $F_n \in C^1[A, B]$, for all $n \in \mathbb{N}$, be given such that $\max_{s \in [A,B]} |F(s) - F_n(s)|$ tends to 0 as $n \to \infty$. Thanks to Lemma 3.2, let us define $v_n \in C^2[0, h]$ as the solution of Problem (3.1) for $F = F_n$. It thus satisfies

$$-v_n(x)' + F_n(v_n(x)) = G_n, \ \forall x \in [0, h], \tag{3.9}$$

where $G_n = \mathcal{G}(F_n, A, B, h)$. Thanks to Lemma 2.2, the sequence $(G_n)_{n \in \mathbb{N}}$ converges to $G = \mathcal{G}(F, A, B, h)$. We thus get from (3.9) that the sequence $(v_n)_{n \in \mathbb{N}}$ is bounded in $C^1[0, h]$. From Ascoli's theorem, we get that there exists a subsequence, again denoted $(v_n)_{n \in \mathbb{N}}$, and a function $v \in C[0, h]$, such that $(v_n)_{n \in \mathbb{N}}$ converges to v in $C[0, h]$. Moreover, using (3.9), we get

$$v_n(x)' = F(v_n(x)) + (F_n(v_n(x)) - F(v_n(x)) - G_n, \ \forall x \in [0, h],$$

which proves that the sequence $(v_n')_{n \in \mathbb{N}}$ converges to $F(v) - G$ in $C[0, h]$. We thus get that the function v satisfies $v \in C^1[0, h]$ and that (v, G) is solution to Problem (3.2). We have thus proved the first part of the lemma. Let us now prove that any solution $(v, G) \in C^1[0, h] \times \mathbb{R}$ to Problem (3.2) is monotone and that $G = \mathcal{G}(F, A, B, h)$. Let us assume that v is not monotone. Then there exist two values x_1, x_2 in $[0, h]$ such that $v(x_1) = v(x_2)$ and $v'(x_1)v'(x_2) < 0$. But as $-v(x_1)' + F(v(x_1)) = -v(x_2)' + F(v(x_2))$, we conclude that $v'(x_1) = v'(x_2)$, leading to a contradiction. Consequently, $\forall x \in [0, h], v(x) \in [A, B]$. From the mean value theorem, we get that there exists $\bar{x} \in [0, h]$ such that $v'(\bar{x}) = \frac{B-A}{h} > 0$ which proves that for all $x \in [0, h]$, $v'(x) \ge 0$. Therefore, we necessarily get $G \le F^\flat_{A,B}$. Assume that there exists $x_0 \in [0, h]$ such that $v'(x_0) = 0$. In such a case, we get $G = F(v(x_0))$, which implies $G = F^\flat_{A,B}$. We thus get, for all $\varepsilon > 0$,

$$v(x)' + \varepsilon = F(v(x)) - F^\flat_{A,B} + \varepsilon, \ \forall x \in [0, h],$$

which gives

$$\int_0^h \frac{(v'(x) + \varepsilon)dx}{F(v(x)) - F^\flat_{A,B} + \varepsilon} = h,$$

yielding

$$\int_0^h \frac{v'(x)dx}{F(v(x)) - F^\flat_{A,B} + \varepsilon} \leq h.$$

Thanks to Lemma 2.4, we can apply the change of variable $s = v(x)$ which gives

$$\int_A^B \frac{ds}{F(s) - F^\flat_{A,B} + \varepsilon} \leq h, \ \forall \varepsilon > 0.$$

We thus proved that in this case $G = \mathcal{G}(F, A, B, h)$. Otherwise, if for all $x \in [0, h]$, $v'(x) > 0$, we get that $G < F^\flat_{A,B}$, and since v is a diffeomorphism from $[0, h]$ to $[A, B]$, we can make the change of variable $s = v(x)$ in

$$\int_0^h \frac{v'(x)dx}{F(v(x)) - G} = h,$$

thus obtaining

$$\int_A^B \frac{ds}{F(s) - G} = h.$$

We thus have proved in this case as well that $G = \mathcal{G}(F, A, B, h)$. □

4. Study of the resulting finite volume scheme

We give in this section the main steps which permit to define a finite volume scheme, we give the results of convergence. We first introduce the notion of admissible discretization [10] which is useful to define a finite volume scheme.

4.1. **Definition** (Admissible discretization)**.** Under assumption 1.1, an admissible finite volume discretization of $\Omega \times (0, T)$, denoted by $\mathcal{D}$, is given by $\mathcal{D} = (\mathcal{T}, \mathcal{E}, \mathcal{P}, \tau)$, where:

- $\mathcal{T}$ is a finite family of non empty open polygonal convex disjoint subsets of Ω (the "control volumes") such that $\overline{\Omega} = \cup_{K \in \mathcal{T}} \overline{K}$. We then denote, for all $K \in \mathcal{T}$, by $\partial K = \overline{K} \setminus K$ the boundary of K and $m_K > 0$ the N-dimensional Lebesgue measure of K (it is the area of K in the two-dimensional case and the volume in the three-dimensional case).
- $\mathcal{E}$ is a finite family of disjoint subsets of $\overline{\Omega}$ (the "edges" of the mesh), such that, for all $\sigma \in \mathcal{E}$, there exists a hyperplane E of $\mathbb{R}^N$ and $K \in \mathcal{T}$ with $\overline{\sigma} = \partial K \cap E$ and σ is a non empty open subset of E. We then denote $m_\sigma > 0$ the $(N-1)$-dimensional measure of σ. We assume that, for all $K \in \mathcal{T}$, there exists a subset $\mathcal{E}_K$ of $\mathcal{E}$ such that $\partial K = \cup_{\sigma \in \mathcal{E}_K} \overline{\sigma}$. It then results from the previous hypotheses that, for all $\sigma \in \mathcal{E}$, either $\sigma \subset \partial\Omega$ or there exists $(K, L) \in \mathcal{T}^2$ with $K \neq L$ such that $\overline{K} \cap \overline{L} = \overline{\sigma}$; we denote in the latter case $\sigma = K|L$. We denote by $\mathcal{N}_K$ the set of control volumes $L \neq K$ such that there exists $\sigma \in \mathcal{E}_K$ with $\sigma = K|L$. The subset of $\mathcal{E}$ of the edges σ such that there exist two control volumes K and L with $\sigma = K|L$ is denoted by $\mathcal{E}_{\text{int}}$.
- $\mathcal{P}$ is a family of points of Ω indexed by $\mathcal{T}$, denoted by $\mathcal{P} = (x_K)_{K \in \mathcal{T}}$. This family is such that, for all $K \in \mathcal{T}$, $x_K \in \bar{K}$. For all $\sigma \in \mathcal{E}$ such that there exists $(K, L) \in \mathcal{T}^2$ with $\sigma = K|L$, it is assumed that $x_K \neq x_L$ and that the straight line (x_K, x_L) going through x_K and x_L is orthogonal to $K|L$. We denote $d_{KL} = \text{dist}(x_k, x_l)$. For all $K \in \mathcal{T}$ such that $\text{meas}(\partial K \cap \partial\Omega) \neq 0$ we assume that $x_K \in \partial\Omega$ (the set of such control volumes is denoted $\mathcal{T}_{\text{ext}}$, the set of control volumes such that $\text{meas}(\partial K \cap \partial\Omega) = 0$ is denoted $\mathcal{T}_{\text{int}}$). We set, for all $K \in \mathcal{T}$ and $\sigma \in \mathcal{E}_K$,
$$D_{K,\sigma} = \{t x_K + (1-t)y, t \in (0,1), \ y \in \sigma\}$$
and we assume that for all $K \in \mathcal{T}$, $\bar{K} = \bigcup_{L \in \mathcal{N}_K} \bar{D}_{K,K|L}$.
- $\tau > 0$ is the time step. Let N_T be the largest integer such that $N_T \tau \leq T$.

The size respectively regularity of the discretization are defined by

$$\text{size}(\mathcal{D}) = \sup\{\text{diam}(K), K \in \mathcal{T}\} \cup \{\tau\}$$

$$\text{reg}(\mathcal{D}) = \inf\left\{\frac{\text{dist}(x_K, K|L)}{\text{diam}(K)}, K \in \mathcal{T}, L \in \mathcal{N}_K\right\}.$$

For all $K \in \mathcal{T}$ and $L \in \mathcal{N}_K$, we denote by $\vec{n}_{K,L}$ the unit vector normal to $K|L$ outward to K.

For any family of values $\{v_K^n\}_{K \in \mathcal{D}, n \in \mathbb{N}}$, we define the piecewise constant function $v_\mathcal{D} : \Omega \times \mathbb{R} \to \mathbb{R}$ as

$$v_\mathcal{D}(x, t) = v_K^{n+1}, \ \forall (x, t) \in K \times (n\tau, (n+1)\tau), \ \forall (K, n) \in \mathcal{T} \times \mathbb{N}.$$

We now give a scheme which permits to obtain an approximation of the solution of the continuous problem given in the introduction of this paper.

4.2. **Definition** (The implicit finite volume scheme). Under assumption 1.1, let $\mathcal{D}$ be an admissible discretization of $\Omega \times (0,T)$ in the sense of Definition 4.1. We set, as given in the introduction of this paper,

$$q_{KL} = \frac{1}{m_{KL}} \int_{K|L} \vec{q}(x) \cdot \vec{n}_{KL} ds(x), \ \forall K \in \mathcal{T}, \ \forall L \in \mathcal{N}_K. \tag{4.1}$$

Setting

$$\bar{u}_K^{n+1} = \frac{1}{\tau \ m_K} \int_{n\tau}^{(n+1)\tau} \int_K \bar{u}(x,s) dx ds, \ \forall K \in \mathcal{T}, \ \forall n \in \mathbb{N}, \tag{4.2}$$

we take into account the boundary condition by

$$u_K^{n+1} = \bar{u}_K^{n+1}, \ \forall K \in \mathcal{T}_{\text{ext}}, \ \forall n \in \mathbb{N}, \tag{4.3}$$

and we take into account the initial condition by

$$u_K^0 = \frac{1}{m_K} \int_K u_0(x) dx, \ \forall K \in \mathcal{T}. \tag{4.4}$$

We then define the scheme by

$$m_K(u_K^{n+1} - u_K^n) + \tau \sum_{L \in \mathcal{N}_K} m_{KL} \ g(u_K^{n+1}, u_L^{n+1}, q_{KL}, d_{KL}) = 0, \forall K \in \mathcal{T}_{\text{int}}, \ \forall n \in \mathbb{N}. \tag{4.5}$$

Then an approximate solution to problem (1.1)-(1.3) is given by $u_{\mathcal{D}}(x,t)$.

4.3. *Remark.* The following results cannot be extended to the explicit scheme since we do not exhibit in this paper the Lipschitz continuity of g in the general case.

4.4. **Lemma** (Stability, existence and uniqueness of a discrete solution). *Under assumptions 1.1, let $\mathcal{D}$ be an admissible discretization of $\Omega \times (0,T)$ in the sense of Definition 4.1.*

(i) *For any solution $u_{\mathcal{D}}$ of the scheme (4.1)-(4.5), we have the estimates*

$$\begin{aligned} u_K^{n+1} &\le \max\{u_K^n\} \cup \{u_L^{n+1}\}_{L \in \mathcal{N}_K} \ \forall K \in \mathcal{T}_{\text{int}}, n > 0 \\ u_K^{n+1} &\ge \min\{u_K^n\} \cup \{u_L^{n+1}\}_{L \in \mathcal{N}_K} \ \forall K \in \mathcal{T}_{\text{int}}, n > 0 \end{aligned} \tag{4.6}$$

(ii)

$$u_K^{n+1} \in [U^\flat, U^\sharp] \ \forall K \in \mathcal{T}_{\text{int}}, n > 0 \tag{4.7}$$

(iii) *There exists one and only one solution $\{u_K^{n+1}\}_{K \in \mathcal{T}_{\text{int}}, \, n \in \mathbb{N}}$ of (4.1)-(4.5).*

Proof. First, for $\lambda \in [0,1]$, define

$$g_{KL}^\lambda(u,v) = m_{KL} \left((1-\lambda) \frac{u-v}{d_{KL}} + \lambda g(u,v,q_{KL},d_{KL}) \right).$$

It is straightforward that g_{KL}^λ is continuous and for each fixed $\lambda \in [0,1]$, it has the same monotonicity properties as g. Let us regard the scheme

$$m_K \frac{u_K^{n+1} - u_K^n}{\tau} + \sum_{L \in \mathcal{N}_K} g_{KL}^\lambda(u_K^{n+1}, u_L^{n+1}) = 0, \forall K \in \mathcal{T}_{\text{int}}, \ \forall n \in \mathbb{N}. \tag{4.8}$$

Positivity of m_K and monotonicity of g^λ in the first argument allow to solve each individual equation of (4.8) for u_K^{n+1}, thus yielding a function U_K^λ such that

$$u_K^{n+1} = U_K^\lambda(u_K^n, \{u_L^{n+1}\}_{L \in \mathcal{N}_K}). \tag{4.9}$$

U_K^λ is non decreasing with respect to all of its arguments: Assume $L_0 \in \mathcal{N}_K$, and let $v_{L_0}^{n+1} \ge u_{L_0}^{n+1}$ and define v_K^{n+1} by

$$m_K \frac{v_K^{n+1} - u_K^n}{\tau} + \sum_{L \in \mathcal{N}_K \setminus \{L_0\}} g_{KL}^\lambda(v_K^{n+1}, u_L^{n+1}) + g_{KL_0}^\lambda(v_K^{n+1}, v_{L_0}^{n+1}) = 0$$

Subtracting this equation from (4.8) and adding $0 = -g_{KL_0}^\lambda(v_K^{n+1}, u_{L_0}^{n+1}) + g_{KL_0}^\lambda(v_K^{n+1}, u_{L_0}^{n+1})$ yields

$$\begin{aligned} m_K \frac{v_K^{n+1} - u_K^{n+1}}{\tau} &+ \sum_{L \in \mathcal{N}_K \setminus \{L_0\}} \left(g_{KL}^\lambda(v_K^{n+1}, u_L^{n+1}) - g_{KL}^\lambda(u_K^{n+1}, u_L^{n+1}) \right) \\ &+ g_{KL_0}^\lambda(v_K^{n+1}, u_{L_0}^{n+1}) - g_{KL_0}^\lambda(u_K^{n+1}, u_{L_0}^{n+1}) + g_{KL_0}^\lambda(v_K^{n+1}, v_{L_0}^{n+1}) - g_{KL_0}^\lambda(v_K^{n+1}, u_{L_0}^{n+1}) = 0 \end{aligned}$$

Define

$$\tilde{U}_K^\lambda(\xi) = \frac{m_K}{\tau} \xi + \sum_{L \in \mathcal{N}_K} g_{KL}^\lambda(\xi, u_L^{n+1})$$

By monotonicity of g in the second argument we get

$$\tilde{U}_K^\lambda(v_K^{n+1}) - \tilde{U}_K^\lambda(u_K^{n+1}) = g_{KL_0}^\lambda(v_K^{n+1}, u_{L_0}^{n+1}) - g_{KL_0}^\lambda(v_K^{n+1}, v_{L_0}^{n+1}) \geq 0$$

But by construction, $\tilde{U}_K^\lambda$ is strictly monotone, so necessarily, $v_K^{n+1} \geq u_K^{n+1}$.

Now, let $v_K^n \geq u_K^n$ and define v_K^{n+1} by

$$m_K \frac{v_K^{n+1} - v_K^n}{\tau} + \sum_{L \in \mathcal{N}_K} g_{KL}^\lambda(v_K^{n+1}, u_L^{n+1}) = 0$$

Subtracting from (4.8) yields

$$m_K \frac{v_K^{n+1} - u_K^n}{\tau} + \sum_{L \in \mathcal{N}_K} \left(g_{KL}^\lambda(v_K^{n+1}, u_L^{n+1}) - g_{KL}^\lambda(u_K^{n+1}, u_L^{n+1}) \right) = m_K \frac{v_K^n - u_K^n}{\tau} \geq 0,$$

therefore, as above $\tilde{U}_K^\lambda(v_K^{n+1}) - \tilde{U}_K^\lambda(u_K^{n+1}) \geq 0$, again yielding $v_K^{n+1} \geq u_K^{n+1}$.

As the next step, we verify that for all $a \in \mathbb{R}$,

$$a = U_K^\lambda(a, \{a\}_{L \in \mathcal{N}_K}). \tag{4.10}$$

Looking at (4.8), it remains to show that

$$\sum_{L \in \mathcal{N}_K} g_{KL}^\lambda(a, a) = 0, \forall K \in \mathcal{T}_{\text{int}}, \ \forall n \in \mathbb{N}.$$

Equation (2.9) yields $g_{KL}^\lambda(a, a) = \lambda m_{KL} q_{KL} f(a)$. Equation (4.1), the Gaussian integral theorem and the assumption $\text{div}\vec{q} = 0$ from 1.1 ensure that $\sum_{L \in \mathcal{N}_K} m_{KL} q_{KL} = 0$.

Let $u_{\max} = \max\{u_K^n\} \cup \{u_L^{n+1}\}_{L \in \mathcal{N}_K}$. Then by (4.10) and monotonicity of U_K^λ,

$$u_{\max} = U_K^\lambda(u_{\max} \ldots u_{\max}) \geq U_K^\lambda(u_K^n, \{u_L^{n+1}\}_{L \in \mathcal{N}_K}) = u_K^{n+1},$$

thus verifying the first inequality of (4.6). A similar discussion yields the second inequality. Let K be such that $u_K^{n+1} = u_{\max} = \max\{u_L^{n+1}\}_{L \in \mathcal{T}}$. Assume that $K \in \mathcal{T}_{\text{int}}$. we then remark that

$$u_K^{n+1} = u_{\max} = U_K^\lambda(u_K^n, \{u_L^{n+1}\}_{L \in \mathcal{N}_K}) \leq U_K^\lambda(u_K^n, \{u_{\max}\}_{L \in \mathcal{N}_K}),$$

while at the other hand, $u_{\max} = U_K^\lambda(u_{\max}, \{u_{\max}\}_{L \in \mathcal{N}_K})$. This implies $u_{\max} \leq u_K^n$, and the proof of the L^∞ estimate (4.7) follows by induction.

All the results shown in this proof so far hold for any $\lambda \in [0, 1]$ under the assumption that a solution exists. Let $N_{int} = \text{card}\,\mathcal{T}_{\text{int}}$. Then the scheme (4.8) defines a discrete nonlinear equation in $\mathbb{R}^{N_{int}}$

$$U^\lambda(u^{n+1}) - Mu^n = 0 \tag{4.11}$$

where M is a diagonal matrix consisting of the values m_K. For $\lambda = 0$, the system is linear. Its matrix is diagonally dominant with non-positive off diagonal entries, and its graph is connected, so it has the M-property, implying that $U^0(u^{n+1}) - Mu^n = 0$ has a unique solution. Thus the topological degree of the affine mapping $U^0(\cdot) - Mu^n$ is nonzero. Furthermore, $U^\lambda(\cdot) - Mu^n$ is continuous in its argument and in λ, and for any λ, any solutions are bounded by the L^∞ estimate. Being a bounded set of solutions of a continuous operator, the set of possible solutions is compact. Therefore the homotopy invariance of the topological degree yields that the degree of $U^1(\cdot) - Mu^n$ is nonzero proving the existence of at least one solution of $U^1(u^{n+1}) - Mu^n = 0$ which is equivalent to our original scheme (4.5) [17, 6].

To prove uniqueness, we follow the reasoning of [10]. Let $\lambda = 1$ and $U_K = U_K^1$ and $g_{KL} = g_{KL}^1$. Assume that u^{n+1} and v^{n+1} are two solutions of (4.8). From equation (4.9) and the monotonicity of U we get

$$u_K^{n+1} - U_K(u_K^n \top v_K^n, \{u_L^{n+1} \top v_L^{n+1}\}_{L \in \mathcal{N}_K}) \leq 0$$
$$v_K^{n+1} - U_K(u_K^n \top v_K^n, \{u_L^{n+1} \top v_L^{n+1}\}_{L \in \mathcal{N}_K}) \leq 0,$$

and, therefore

$$u_K^{n+1} \top v_K^{n+1} - U_K(u_K^n \top v_K^n, \{u_L^{n+1} \top v_L^{n+1}\}_{L \in \mathcal{N}_K}) \leq 0$$

Similarly, we obtain

$$u_K^{n+1} \bot v_K^{n+1} - U_K(u_K^n \bot v_K^n, \{u_L^{n+1} \bot v_L^{n+1}\}_{L \in \mathcal{N}_K}) \geq 0$$

Subtracting and taking into account $a \top b - a \bot b = |a - b|$ and the monotonicity of g_{KL} in the first argument allows to obtain

$$m_K \frac{|u_K^{n+1} - v_K^{n+1}| - |u_K^n - v_K^n|}{\tau} + \sum_{L \in \mathcal{N}_K} (g_{KL}(u_K^{n+1} \top v_K^{n+1}, u_L^{n+1} \top v_L^{n+1}) - g_{KL}(u_K^{n+1} \bot v_K^{n+1}, u_L^{n+1} \bot v_L^{n+1})) \leq 0$$

Taking into account that $g_{KL}(u, v) = -g_{LK}(v, u)$ leads to canceling of all fluxes along interior edges and to the fact that for any $\{v_K\}_{K \in \mathcal{T}}$,

$$\sum_{K \in \mathcal{T}_{\mathrm{int}}} \sum_{L \in \mathcal{N}_K} g_{KL}(v_K, v_L) = - \sum_{K \in \mathcal{T}_{\mathrm{ext}}} \sum_{L \in \mathcal{N}_K} g_{KL}(v_K, v_L).$$

Therefore,

$$\sum_{K \in \mathcal{T}_{\mathrm{int}}} m_K \frac{|u_K^{n+1} - v_K^{n+1}| - |u_K^n - v_K^n|}{\tau} \leq \sum_{K \in \mathcal{T}_{\mathrm{ext}}} \sum_{L \in \mathcal{N}_K} \left(g_{KL}(u_K^{n+1} \top v_K^{n+1}, u_L^{n+1} \top v_L^{n+1}) - g_{KL}(u_K^{n+1} \bot v_K^{n+1}, u_L^{n+1} \bot v_L^{n+1}) \right).$$

But for any $K \in \mathcal{T}_{\mathrm{ext}}$,

$$\begin{aligned} g_{KL}(u_K^{n+1} \top v_K^{n+1}, u_L^{n+1} \top v_L^{n+1}) - g_{KL}(u_K^{n+1} \bot v_K^{n+1}, u_L^{n+1} \bot v_L^{n+1}) \\ = g_{KL}(\bar{u}_K^{n+1}, u_L^{n+1} \top v_L^{n+1}) - g_{KL}(\bar{u}_K^{n+1}, u_L^{n+1} \bot v_L^{n+1}) \leq 0 \end{aligned}$$

Adding trivial terms at the boundary yields the L^1 contraction estimate

$$\sum_{K \in \mathcal{T}} m_K |u_K^{n+1} - v_K^{n+1}| \leq \sum_{K \in \mathcal{T}} m_K |u_K^n - v_K^n|$$

leading to the uniqueness of the discrete solution. □

For a discrete function $v_{\mathcal{D}} = \{v_K^n\}_{K \in \mathcal{T}, n \in \mathbb{N}}$ define

$$\mathcal{N}_{\mathcal{D}}(v_{\mathcal{D}})^2 = \sum_{n=0}^{N_T} \tau \sum_{K \in \mathcal{T}} \sum_{L \in \mathcal{N}_K} \frac{m_{KL}}{d_{KL}} (v_K^{n+1} - v_L^{n+1})^2$$

4.5. **Lemma** (Discrete $L^2(0, T; H^1(\Omega))$ estimate). *Under assumptions 1.1, let $\mathcal{D}$ be an admissible discretization of $\Omega \times (0, T)$ in the sense of Definition 4.1. Let $\{u_K^{n+1}\}_{K \in \mathcal{T}, n \in \mathbb{N}}$ and $\{\bar{u}_K^{n+1}\}_{K \in \mathcal{T}, n \in \mathbb{N}}$ be given by (4.1)-(4.5). Let $\rho > 0$ such that* $\mathrm{reg}(\mathcal{D}) \geq \rho$. *Then, there exists a constant C_1, only depending on Ω, f, ϕ, u_0, $\bar{u}$, ρ, and not on* $\mathrm{size}(\mathcal{D})$ *such that*

$$\mathcal{N}_{\mathcal{D}}(\zeta(u_{\mathcal{D}}) - \zeta(\bar{u}_{\mathcal{D}}))^2 \leq C_1$$

Proof. We mainly follow the reasoning in [11]. We multiply (4.5) by $w_K^{n+1} = u_K^{n+1} - \bar{u}_K^{n+1}$ (which vanishes for all $K \in \mathcal{T}_{\mathrm{ext}}$), and sum on n and K. We obtain: $A + \bar{A} + B = 0$, where

$$A = \sum_{n=0}^{N_T} \sum_{K \in \mathcal{T}_{\mathrm{int}}} m_K (w_K^{n+1} - w_K^n) w_K^{n+1}$$

$$\bar{A} = \sum_{n=0}^{N_T} \sum_{K \in \mathcal{T}_{\mathrm{int}}} m_K (\bar{u}_K^{n+1} - \bar{u}_K^n) w_K^{n+1}$$

$$B = \tau \sum_{n=0}^{N_T} \sum_{K \in \mathcal{T}_{\mathrm{int}}} \sum_{L \in \mathcal{N}_K} m_{KL} \; g(u_K^{n+1}, u_L^{n+1}, q_{KL}, d_{KL}) w_K^{n+1}$$

Thanks to (1.10) and to $w_K^{n+1} = 0$ if $K \in \mathcal{T}_{\mathrm{ext}}$ we get

$$B = \tau \sum_{n=0}^{N_T} \frac{1}{2} \sum_{K \in \mathcal{T}} \sum_{L \in \mathcal{N}_K} m_{KL} \; g(u_K^{n+1}, u_L^{n+1}, q_{KL}, d_{KL})(w_K^{n+1} - w_L^{n+1})$$

which delivers $B = B' - \bar{B}$ with

$$B' = \tau \sum_{n=0}^{N_T} \frac{1}{2} \sum_{K \in \mathcal{T}} \sum_{L \in \mathcal{N}_K} m_{KL} \; g(u_K^{n+1}, u_L^{n+1}, q_{KL}, d_{KL})(u_K^{n+1} - u_L^{n+1})$$

$$\bar{B} = \quad \tau \sum_{n=0}^{N_T} \frac{1}{2} \sum_{K\in\mathcal{T}} \sum_{L\in\mathcal{N}_K} m_{KL}\; g(u_K^{n+1}, u_L^{n+1}, q_{KL}, d_{KL})(\bar{u}_K^{n+1} - \bar{u}_L^{n+1}),$$

leading to $B' = \bar{B} - A - \bar{A}$.

Estimate of A: We get, thanks to $b(b-a) = \frac{1}{2}b^2 + \frac{1}{2}(b-a)^2 - \frac{1}{2}a^2$ that

$$A \geq \frac{1}{2} \sum_{K\in\mathcal{T}_{\text{int}}} m_K \left((w_K^{N_T+1})^2 - (w_K^0)^2 \right) \geq -\frac{1}{2} \sum_{K\in\mathcal{T}_{\text{int}}} m_K (w_K^0)^2 \geq -\frac{1}{2} ||u_0 - \bar{u}(\cdot, 0)||_{L^2(\Omega)} = C_A \tag{4.12}$$

Estimate of $\bar{A}$: For $K \in \mathcal{T}$, $n \leq N_T$ define

$$\tilde{u}_K^n = \frac{1}{m_K} \int_K u(x, t^n)\, dx$$

Then

$$\bar{A} = \sum_{n=0}^{N_T} \sum_{K\in\mathcal{T}_{\text{int}}} m_K (\bar{u}_K^{n+1} - \tilde{u}_K^n) w_K^{n+1} + \sum_{n=0}^{N_T} \sum_{K\in\mathcal{T}_{\text{int}}} m_K (\tilde{u}_K^n - \bar{u}_K^n) w_K^{n+1}.$$

A density argument allows to estimate

$$|\bar{u}_K^{n+1} - \tilde{u}_K^n| \leq ||\bar{u}_t||_{L^1(K\times(t^n, t^{n+1}))}$$
$$|\bar{u}_K^n - \tilde{u}_K^n| \leq ||\bar{u}_t||_{L^1(K\times(t^{n-1}, t^n))}$$

Using the L^∞ bound we obtain

$$|\bar{A}| \leq 2||\bar{u}_t||_{L^1(\Omega\times[0,T])} |u^\sharp - u^\flat| = C_{\bar{A}}.$$

Estimate of $\bar{B}$: Equation (2.9) allows to obtain

$$\bar{B} = \tau \sum_{n=0}^{N_T} \frac{1}{2} \sum_{K\in\mathcal{T}} \sum_{L\in\mathcal{N}_K} \frac{m_{KL}}{d_{KL}} (\phi(u_K^{n+1}) - \phi(u_L^{n+1}))(\bar{u}_K^{n+1} - \bar{u}_L^{n+1})$$
$$+ \tau \sum_{n=0}^{N_T} \frac{1}{2} \sum_{K\in\mathcal{T}} \sum_{L\in\mathcal{N}_K} m_{KL} q_{KL} f(u_{KL}^{n+1})(\bar{u}_K^{n+1} - \bar{u}_L^{n+1}) \tag{4.13}$$

where $u_{KL}^{n+1} \in [u_K \perp u_L, u_K \top u_L]$. We apply Young's inequality for the first sum. In the second sum we use continuity of f and boundedness of u_{KL}^{n+1}. This allows to obtain for any $\alpha > 0$

$$|\bar{B}| \;\leq\; \frac{\alpha}{2} \mathcal{N}_{\mathcal{D}}(\phi(u_{\mathcal{D}}))^2 \;+\; \frac{1}{2\alpha} \mathcal{N}_{\mathcal{D}}(\bar{u}_{\mathcal{D}})^2 \;+\; C_{q,f} \sum_{n=0}^{N_T} \frac{1}{2} \sum_{K\in\mathcal{T}} \sum_{L\in\mathcal{N}_K} \frac{m_{KL}}{\sqrt{d_{KL}}} \sqrt{d_{KL}} |\bar{u}_K^{n+1} \;-\; \bar{u}_L^{n+1}|$$

For the first sum, we remark that there is a constant C_ϕ such that $|\phi(u) - \phi(v)| \leq C_\phi |\zeta(u) - \zeta(v)|$. This allows to estimate

$$|\bar{B}| \leq C_\phi^2 \frac{\alpha}{2} \mathcal{N}_{\mathcal{D}}(\zeta(u_{\mathcal{D}}))^2 + \frac{1}{2\alpha} \mathcal{N}_{\mathcal{D}}(\bar{u}_{\mathcal{D}})^2$$
$$+ C_{q,f} \mathcal{N}_{\mathcal{D}}(\bar{u}_{\mathcal{D}}) \left(\tau \sum_{n=0}^{N_T} \sum_{K\in\mathcal{T}} \sum_{L\in\mathcal{N}_K} m_{KL} d_{KL} \right)^{\frac{1}{2}}$$
$$\leq C_\phi^2 \mathcal{N}_{\mathcal{D}}(\zeta(u_{\mathcal{D}}))^2 + \frac{1}{2\alpha} \mathcal{N}_{\mathcal{D}}(\bar{u}_{\mathcal{D}})^2 + C_{q,f} \mathcal{N}_{\mathcal{D}}(\bar{u}_{\mathcal{D}}) d m_\Omega$$

As we can estimate $\mathcal{N}_{\mathcal{D}}(\bar{u}_{\mathcal{D}}) \leq F(\rho) ||\bar{u}||_{L^2((0,T), H^1(\Omega)}$ [9] we obtain the estimate

$$|\bar{B}| \leq C_\phi \frac{\alpha}{2} \mathcal{N}_{\mathcal{D}}(\zeta(u_{\mathcal{D}}))^2 + C_{\bar{B}}(\alpha, \rho)$$

Final estimate: Using (2.10) we get $B' \geq \frac{1}{2} \mathcal{N}_{\mathcal{D}}(\zeta(u_{\mathcal{D}}))^2$ which for any $\alpha > 0$ leads to

$$\frac{1}{2} \mathcal{N}_{\mathcal{D}}(\zeta(u_{\mathcal{D}}))^2 \leq C_\phi \frac{\alpha}{2} \mathcal{N}_{\mathcal{D}}(\zeta(u_{\mathcal{D}}))^2 + C_{\bar{B}}(\alpha, \rho) + C_A + C_{\bar{A}}$$

allowing to deduce for $\alpha = \frac{1}{2C_\phi}$ $\mathcal{N}_{\mathcal{D}}(\zeta(u_{\mathcal{D}}))^2 \leq C^* = 4(C_{\bar{B}}(\alpha, \rho) + C_A + C_{\bar{A}})$. But then, [9]

$$\mathcal{N}_{\mathcal{D}}(\zeta(u_{\mathcal{D}}) - \zeta(\bar{u}_{\mathcal{D}}))^2 \leq 2\mathcal{N}_{\mathcal{D}}(\zeta(u_{\mathcal{D}}))^2 + 2\mathcal{N}_{\mathcal{D}}(\zeta(\bar{u}_{\mathcal{D}}))^2$$
$$\leq 2C^* + 2C_\zeta^2 F(\rho) ||\bar{u}_{\mathcal{D}}||_{L^2((0,T), H^1(\Omega)}.$$

□

We can then obtain similar results to those of [10] and [11].

4.6. **Lemma** (Space and time translate estimates). *Under assumptions 1.1, let $\mathcal{D}$ be an admissible discretization of $\Omega \times (0,T)$ in the sense of Definition 4.1 such that* $\mathrm{reg}(\mathcal{D}) \geq \rho > 0$ *Let u_K^{n+1} and $\bar{u}_K^{n+1}$ for all $K \in \mathcal{T}$ and $n \in \mathbb{N}$ such that $n\tau \leq T$, be given by (4.1)-(4.5). Then we get the existence of C_2, only depending on Ω, f, ϕ, u_0, $\bar{u}$, ρ and not on $\mathcal{D}$ such that the function $z_{\mathcal{D}} = \zeta(u_{\mathcal{D}}) - \zeta(\bar{u}_{\mathcal{D}})$ (prolonged by 0 outside of $\Omega \times (0,T)$) satisfies*

$$\int_0^T \int_\Omega (z_{\mathcal{D}}(x+\xi,t) - z_{\mathcal{D}}(x,t))^2 dxdt \leq C_2|\xi| \ (|\xi| + 4 \ \mathrm{size}(\mathcal{D})), \ \forall \xi \in \mathbb{R}^d, \tag{4.14}$$

and

$$\int_0^T \int_\Omega (z_{\mathcal{D}}(x,t+\theta) - z_{\mathcal{D}}(x,t))^2 dxdt \leq C_2|\theta|, \ \forall \theta \in \mathbb{R}. \tag{4.15}$$

Proof. The proof of (4.14) is a classical consequence of Lemma 4.5 (the results of [10] apply, since $z_K^n = 0$ holds for any $n \in \mathbb{N}$ and $K \in \mathcal{T}_{\mathrm{ext}}$). Let us now turn to the proof of (4.15). Denoting, for all $t \in \mathbb{R}_+$, we notice that

$$\int_0^T \int_\Omega (z_{\mathcal{D}}(x,t+\theta) - z_{\mathcal{D}}(x,t))^2 dxdt = \sum_{K\in\mathcal{T}_{\mathrm{int}}} \int_0^T m_K (z_K^{N_{t+\theta}+1} - z_K^{N_t+1})^2 dt = A - 2B + \bar{A},$$

with

$$A = \sum_{K\in\mathcal{T}_{\mathrm{int}}} \int_0^T m_K (\zeta(u_K^{N_{t+\theta}+1}) - \zeta(u_K^{N_t+1}))^2 dt,$$

$$B = \sum_{K\in\mathcal{T}_{\mathrm{int}}} \int_0^T m_K (\zeta(u_K^{N_{t+\theta}+1}) - \zeta(u_K^{N_t+1}))(\zeta(\bar{u}_K^{N_{t+\theta}+1}) - \zeta(\bar{u}_K^{N_t+1}))dt,$$

and

$$\bar{A} = \sum_{K\in\mathcal{T}_{\mathrm{int}}} \int_0^T m_K (\zeta(\bar{u}_K^{N_{t+\theta}+1}) - \zeta(\bar{u}_K^{N_t+1}))^2 dt.$$

Using the hypothesis $\bar{u} \in H^1(\Omega \times (0,T))$, and the L^∞ estimate on the discrete solution, we easily get that B and $\bar{A}$ are respectively bounded by expressions under the form $C|\theta|$. Let us turn to the study of A. Introducing a Lipschitz constant C_ζ for the function ζ, we get

$$A \leq C_\zeta \sum_{K\in\mathcal{T}_{\mathrm{int}}} \int_0^T m_K (\zeta(u_K^{N_{t+\theta}+1}) - \zeta(u_K^{N_t+1}))(u_K^{N_{t+\theta}+1} - u_K^{N_t+1})dt,$$

which gives $A \leq C_\zeta A'$ with

$$A' = \sum_{K\in\mathcal{T}_{\mathrm{int}}} \int_0^T m_K (\zeta(u_K^{N_{t+\theta}+1}) - \zeta(u_K^{N_t+1})) \sum_{n=N_t+1}^{N_{t+\theta}} (u_K^{n+1} - u_K^n)dt.$$

Using the scheme (4.1)-(4.5) and after gathering by edges, we get that

$$A' = \frac{1}{2} \sum_{K\in\mathcal{T}} \tau \sum_{L\in\mathcal{N}_K} \int_0^T \sum_{n=N_t+1}^{N_{t+\theta}} g(u_K^{n+1}, u_L^{n+1}, q_{KL}, d_{KL}) \times$$
$$\times \Big(\ \zeta(u_L^{N_{t+\theta}+1}) - \zeta(u_K^{N_{t+\theta}+1}) - \zeta(u_L^{N_t+1}) + \zeta(u_K^{N_t+1}) \ \Big) dt.$$

Using expression (2.9), the discrete L^∞ estimate and the Young inequality, we then get the existence of some $C_3 > 0$, only depending on Ω, f, ϕ, u_0, $\bar{u}$ and not on $\mathcal{D}$ such that

$$A' \leq C_3 \int_0^T \sum_{n=N_t+1}^{N_{t+\theta}} (a^n + b^{N_t+1} + b^{N_{t+\theta}+1} + \tau)dt,$$

with

$$a^n = \sum_{K\in\mathcal{T}} \tau \sum_{L\in\mathcal{N}_K} \frac{m_{KL}}{d_{KL}} (\phi(u_K^{n+1}) - \phi(u_L^{n+1}))^2,$$

and

$$b^n = \sum_{K\in\mathcal{T}} \tau \sum_{L\in\mathcal{N}_K} \frac{m_{KL}}{d_{KL}} (\zeta(\bar{u}_K^{n+1}) - \zeta(\bar{u}_L^{n+1}))^2.$$

Since $\sum_{n=0}^{N_T} a^n$ and $\sum_{n=0}^{N_T} b^n$ are bounded depending only on ρ (thanks to Lemma 4.5), we can then apply Lemma 4.6 of [12] (also proved in [10]). This gives a result under the form $A' \leq C|\theta|$. Thanks to the discrete L^∞ estimate, and to the prolongation by 0 outside $\Omega \times (0,T)$, we then conclude that (4.15) holds. □

We are now able to prove the following result.

4.7. **Theorem** (Convergence of the scheme). *Under assumptions 1.1, let $\mathcal{D}$ be an admissible discretization of $\Omega \times (0,T)$ in the sense of Definition 4.1. Let $u_{\mathcal{D}}$ be given by (4.1)-(4.5). Then the function $u_{\mathcal{D}}$ converges in $L^2(\Omega \times (0,T))$ to u as* $\mathrm{size}(\mathcal{D})$ *tends to 0, and* $\mathrm{reg}(\mathcal{D}) \geq \rho$, *where u is the unique weak solution of Problem 1 in the following sense:*

(i) *$u \in L^\infty(\Omega \times (0,T))$ is such that $\zeta(u) - \zeta(\bar{u}) \in L^2(0,T;H_0^1(\Omega))$,*

(ii) *for all test function $\varphi \in C_c^\infty(\Omega \times [0,T))$, we have*

$$\int_0^T \int_\Omega u(x,t)\varphi_t(x,t)dxdt + \int_0^T \int_\Omega (f(u(x,t))\ \vec{q}(x) - \nabla\phi(u)(x,t)) \cdot \nabla\varphi(x,t)dxdt + \int_\Omega u_0(x)\varphi(x,0)dx = 0.$$

Proof. This proof is similar to that of [10], [11] and [12]. We consider a sequence of discretizations $(\mathcal{D}_m)_{m\in\mathbb{N}}$ with size tending to 0. Applying Lemma 4.6 and Kolmogorov's theorem, we can extract a subsequence such that the sequence $(z_{\mathcal{D}_m})_{m\in\mathbb{N}}$ (we recall that $z_{\mathcal{D}_m} = \zeta(u_{\mathcal{D}_m}) - \zeta(\bar{u}_{\mathcal{D}_m})$) converges in $L^2(\Omega \times (0,T))$ to some function $\tilde{z}$ such that $\tilde{z} \in L^2(0,T;H_0^1(\Omega))$. Since the sequence $(\bar{u}_{\mathcal{D}_m})_{m\in\mathbb{N}}$ strongly converges and since ζ is strictly increasing and continuous, we thus get that the sequence $(u_{\mathcal{D}_m})_{m\in\mathbb{N}}$ converges in $L^2(\Omega \times (0,T))$ as well to some function $\tilde{u} \in L^2(\Omega \times (0,T))$ such that $\tilde{z} = \zeta(\tilde{u}) - \zeta(\bar{u}) \in L^2(0,T;H_0^1(\Omega))$, which shows that the boundary condition is satisfied. It now suffices to obtain the weak relation satisfied by $\tilde{u}$, in order to show that $\tilde{u} = u$. Equation (2.9) can be rewritten as

$$g(u_K^{n+1}, u_L^{n+1}, q_{KL}, d_{KL}) = q_{KL} f(u_{KL}^{n+1}) + \frac{\phi(u_K^{n+1}) - \phi(u_L^{n+1})}{d_{KL}},$$

where $u_{KL}^{n+1} = u_{LK}^{n+1}$ is a value belonging to the interval with ends u_K^{n+1} and u_L^{n+1}. Then, the multiplication of the scheme (4.5) by $\varphi(x_K,(n+1)\tau)$, where φ is a regular test function with a compact support, leads to $A_m + B_m + C_m = 0$, in which we define:

$$A_m = \sum_{n=0}^{N_T} \sum_{K\in\mathcal{T}_{\mathrm{int}}} m_K (u_K^{n+1} - u_K^n)\varphi(x_K,(n+1)\tau),$$

$$B_m = \sum_{n=0}^{N_T} \tau \sum_{K\in\mathcal{T}_{\mathrm{int}}} \sum_{L\in\mathcal{N}_K} m_{KL} q_{KL} f(u_{KL}^{n+1})\varphi(x_K,(n+1)\tau)$$

and

$$C_m = \sum_{n=0}^{N_T} \tau \sum_{K\in\mathcal{T}_{\mathrm{int}}} \sum_{L\in\mathcal{N}_K} m_{KL} \frac{\phi(u_K^{n+1}) - \phi(u_L^{n+1})}{d_{KL}} \varphi(x_K,(n+1)\tau).$$

These terms then classically verify (see [10], [11] and [12]) the following convergence properties:

$$\lim_{m\to\infty} A_m = -\int_0^T \int_\Omega \tilde{u}(x,t)\varphi_t(x,t)dxdt - \int_\Omega u_0(x)\varphi(x,0)dx,$$

and

$$\lim_{m\to\infty} C_m = \int_0^T \int_\Omega \nabla\phi(\tilde{u})(x,t) \cdot \nabla\varphi(x,t)dxdt.$$

The only relation which is slightly different from the above references is

$$\lim_{m\to\infty} B_m = -\int_0^T \int_\Omega f(\tilde{u}(x,t))\ \vec{q}(x) \cdot \nabla\varphi(x,t)dxdt. \tag{4.16}$$

Let

$$\tilde{B}_m = \sum_{n=0}^{N_T} \tau \sum_{K\in\mathcal{T}_{\mathrm{int}}} \sum_{L\in\mathcal{N}_K} m_{KL} \tilde{q}_{KL} f(u_{KL}^{n+1})\varphi(x_K,(n+1)\tau)$$

with

$$\tilde{q}_{KL} = \frac{1}{\mathrm{meas}(D_{KL})} \int_{D_{KL}} \vec{q}(x) \cdot \vec{n}_{KL} dx,$$

where $D_{KL} = D_{K,K|L} \cup D_{L,K|L}$ is the so called "diamond". The regularity of $\vec{q}$ allows to assume that q_{KL} can be estimated by $\tilde{q}_{KL}$ assuring that B_m and $\tilde{B}_m$ have the same limit.

Let us define the function $\nabla_{\mathcal{D}_m}\varphi$, by $\nabla_{\mathcal{D}_m}\varphi(x,t) = \frac{d}{d_{KL}}(\varphi(x_L,(n+1)\tau) - \varphi(x_K,(n+1)\tau))\vec{n}_{KL}$ for a.e. $(x,t) \in D_{KL} \times (n\tau,(n+1)\tau)$ and for all $\sigma = K|L$, and by $\nabla_{\mathcal{D}_m}\varphi(x,t) = 0$ for a.e. $(x,t) \in D_{K,\sigma} \times (n\tau,(n+1)\tau)$, for all $\sigma \in \mathcal{E}_{\mathrm{ext}}$. We get from [8] that $\nabla_{\mathcal{D}_m}\varphi$ weakly converges to $\nabla\varphi$ in $L^2(\Omega \times (0,T))$. We then define the function $\hat{u}_{\mathcal{D}_m}$ by the value u_{KL}^{n+1} in $D_{KL} \times (n\tau,(n+1)\tau)$ and, for all $\sigma \in \mathcal{E}_{\rm ext}$, by the value u_K^{n+1} in $D_{K,\sigma} \times (n\tau,(n+1)\tau)$ where K is such that $\sigma \in \mathcal{E}_K$. It is easy to see, thanks to Lemma 4.5 that $\zeta(\hat{u}_{\mathcal{D}_m}) - \zeta(u_{\mathcal{D}_m})$ tends to 0 in $L^2(\Omega \times (0,T))$ as $m \to \infty$. This proves that $\hat{u}_{\mathcal{D}_m}$ also converges in $L^2(\Omega \times (0,T))$ to $\tilde{u}$. Since, gathering by edges, we observe that

$$\tilde{B}_m = -\int_0^T \int_\Omega f(\hat{u}_{\mathcal{D}_m}(x,t))\ \vec{q}(x) \cdot \nabla_{\mathcal{D}_m} \varphi(x,t) dx dt,$$

passing to the strong-weak limit in the above expression gives (4.16), which concludes the proof that $\tilde{u}$ is a weak solution of Problem 1. The uniqueness of this weak solution yields the convergence of all the sequence, and achieves the proof of the theorem. □

5. Numerical experiment

This section illustrates the method and its typical properties on an example such that a simple analytical solution is known. It does not go into implementational details but touches on this example the two main issues: an initial guess for the local Newton's method (needed to solve the nonlinear problem (1.10)), and the evaluation of derivatives close to singularities by asymptotic expressions. The techniques used should ease the application of the suggested method in other cases, too.

We consider the case where, in (1.1), we take $\Omega = (0,1)$, $\phi\ :\ s \mapsto s^2$, $f\ :\ s \mapsto s$, $q \in [0,+\infty)$, in (1.2) we take $u_0 = 0$ and in (1.3), we take, for a given $v \in (q,+\infty)$, $\bar{u}(0,t) = (v-q)vt/2$ and we set $\bar{u}(1,t) = 0$ for $t < 1/v$ and $\bar{u}(1,t) = (v-q)(vt-1)/2$ otherwise. The unique weak solution of this problem is then given by

$$u(x,t) = \begin{cases} (v-q)(vt-x)/2 & \text{if } x < vt, \\ 0 & \text{if } x \geq vt. \end{cases}$$

Since this situation corresponds to a decreasing solution, we now examine the computation of the numerical flux $G = g(a,b,q,h) = -g(b,a,-q,h)$, defined by (1.10) in the case $0 \leq b < a$. In this case, we get that G is such that $G > qa$ and that $\int_b^a \frac{2sds}{G-qs} = h$. The case $q = 0$ leads to $G = \frac{1}{h}(a^2-b^2)$. Let us suppose that $q > 0$. This gives, by integration, the equation

$$\frac{G}{q} \ln \frac{\frac{G}{q} - b}{\frac{G}{q} - a} = \frac{qh}{2} + a - b.$$

We set $\gamma = \frac{G}{q} - a$ and $d = \frac{qh}{2} + a - b$. This leads to the equations

$$(\gamma + a) \ln \left(1 + \frac{a-b}{\gamma}\right) - d = 0 \tag{5.1}$$

and

$$\gamma(e^{d/(\gamma+a)} - 1) - (a-b) = 0. \tag{5.2}$$

Both equations (5.1) and (5.2) are useful to derive close bounds of γ. Note that, from Lemma 2.5, we know that $\gamma \leq \gamma_0$, where $\gamma_0 = (G - \psi_{godunov}(a,b,q))/q$ is the Godunov approximation, defined by

$$\gamma_0 := \frac{a^2 - b^2}{qh}.$$

Using (5.2), a super-solution for all h with the correct asymptotics for small and large h (see Figure 1) can be derived by

$$\gamma \leq \gamma_2 := \frac{a-b}{e^{d/(\gamma_0+a)} - 1}.$$

The same argument can be used to construct a sub-solution for large h by choosing the sub-solution $\gamma = 0$ in the exponent of (5.2)

$$\gamma \geq \gamma_3 := \frac{a-b}{e^{d/a} - 1}.$$

Using in (5.1), the inequality $\ln(1+x) \geq x - x^2/2$ for $x \geq 0$ gives, in the case $(4+\epsilon)qha \leq (a+b)^2$,

$$\gamma_1 \leq \gamma,$$

with

$$\gamma_1 = \frac{a^2-b^2}{2qh} \left(1 + \sqrt{1 - \frac{4hqa}{(a+b)^2}}\right).$$

The convex combinations of super- and sub-solutions

$$\gamma_{init} = \begin{cases} \theta\gamma_2 + (1-\theta)\gamma_1, \\ \theta\gamma_2 + (1-\theta)\gamma_3. \end{cases}$$

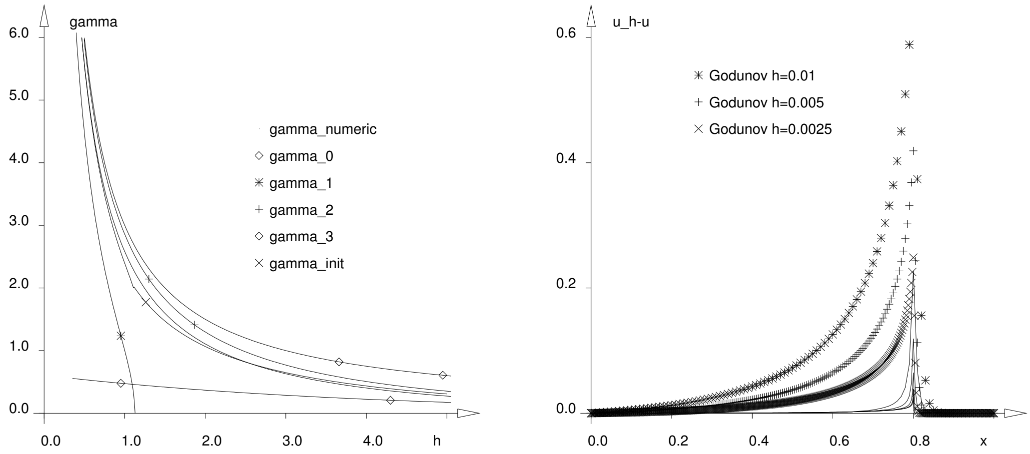

FIGURE 1. Asymptotic expansions (γ_1, γ_2, γ_3, γ_{init}, the numeric solution $\gamma_{numeric}$, and the Godunov flux function γ_0 for $a = 2$, $b = 1$ $q = 1$ (left); the local error of the implicit scheme (lines, $t = 0.004$, $\tau = 3.125 \cdot 10^{-7}$, $h = 0.01,\ 0.01/2, \ldots, 0.01/16$) and the Godunov scheme (markers) with respect to the analytic solution (right).

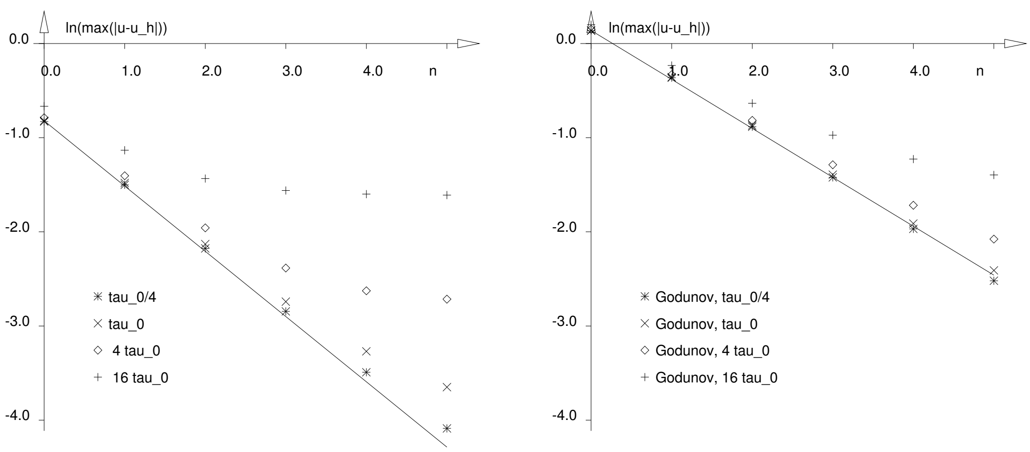

FIGURE 2. Experimental order of convergence for spatial step sizes $h_n = h_0(1/2)^n$ and different time step sizes $\tau_0 = 3.125 \cdot 10^{-7}$, $(||u_h - u||_\infty)$ of the implicit scheme (left, line: $c\, h^1$) and Godunov scheme (right, line: $c\, h^{3/4}$).

($\theta = 0.75$) and using the above inequality with $\epsilon = 0.05$ provide initial values to solve either (5.1) or (5.2) by a local Newton's method in typically 4 steps to rounding precision 10^{-15}. The asymptotic expression γ_{init} is used to compute the derivatives $\partial g/\partial a$, $\partial g/\partial b$ directly when the parameters are to close to the singularities to proceed via the chain rule and the original equations.

The following pictures, obtained with $q = 100$, $v = 200$, $t = 0.004$ (in this case, the analytical solution vanishes at $x_0 = 0.8$) illustrate these generic properties of the flux functions, show the error and the order of convergence of the implicit Euler scheme for different time step sizes (τ). Results for the Godunov case are added, too.

The experiments suggest differences in the order of convergence with respect to h and the Godunov scheme, the interference with the time discretization error can not be neglected. The qualitative properties of the scheme (maximum principle, stability for all h, direct relation to the equation) may be the points of interest for some applications.

6. Conclusions

Dealing in this paper with a nonlinear hyperbolic - degenerate parabolic problem, a new finite volume scheme can be obtained, taking for numerical flux at each interface between two grid blocks the solution of a 1D steady flow problem. This last problem, which is formulated as a nonlinear elliptic equation, is shown to possess an analytical solution which can easily be numerically approached. Thus, taking advantage from the nonlinear diffusion to stabilize the convective discrete

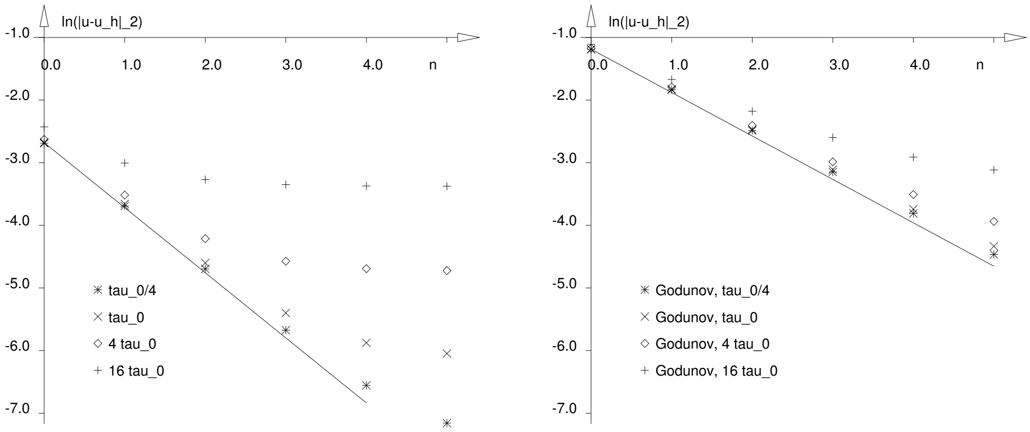

Figure 3. As Figure 2, but $||u_h - u||_2$ (left, line: $c\,h^{3/2}$, right, line: $c\,h^1$).

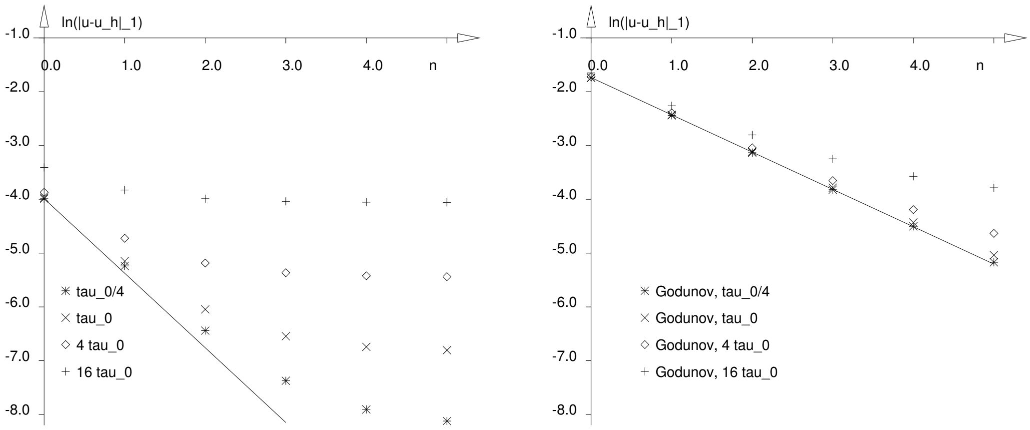

Figure 4. As Figure 2, but $||u_h - u||_1$ (left, line: $c\,h^2$, right, line: $c\,h^1$).

term, this finite volume scheme produces an increased accurateness compared to schemes where the convective part is approximated independently with the nonlinear diffusive part. Some further works remain to be done, in order to ameliorate the algorithm for solving the nonlinear equation providing the numerical flux.

References

[1] D. N. Allen and R. V. Southwell, *Relaxation methods applied to determine the motion, in two dimensions, of a viscous fluid past a fixed cylinder*, Quart. J. Mech. and Appl. Math., 8 (1955), pp. 129–145.

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

[3] Aziz, K, Settari, A., Petroleum Reservoir Simulation, *Applied Science* (1979), London.

[4] J. Carrillo, Entropy solutions for nonlinear degenerate problems. *Arch. Rat. Mech. Anal.*, **147** (1999), 269-361.

[5] Chen, Z. Degenerate Two-Phase Incompressible Flow, *J. of Differential Equations*, **171** (2001), 203–232.

[6] K. Deimling *Nonlinear Functional Analysis.* Springer 1985.

[7] Enchéry G. Ph.D thesis, *Université de Marne-la-Vallée and Institut Français du Pétrole*, 2004.

[8] R. Eymard, T. Gallouët H-convergence and numerical schemes for elliptic problems, *SIAM J. of Numer. Anal.*, **41** (2003), 2, 539–562.

[9] R. Eymard, T. Gallouët, R. Herbin Convergence of finitie volume schemes for semilinear convection diffusion equations, *Numer. Math* **82**(1999),1, 91–116.

[10] R. Eymard, T. Gallouët, R. Herbin The finite volume method, *Handbook for Numerical Analysis, Ph. Ciarlet J.L. Lions eds, North Holland* **VII** (2000), 715–1022.

[11] R. Eymard, T. Gallouët, R. Herbin and A. Michel Convergence of a finite volume scheme for nonlinear degenerate parabolic equations, *Numer.Math* **92**(2002) 41–82.

[12] R. Eymard, T. Gallouët, M. Gutnic, R. Herbin, D. Hilhorst Approximation by the finite volume method of an elliptic-parabolic equation arising in environmental studies, *Mathematical Models and Methods in Applied Sciences (M3AS)***11** (2001) 1505-1528,

[13] J. Fuhrmann and H. Langmach Stability and existence of solutions of time-implicit finite volume schemes for viscous nonlinear conservation laws. *Applied Numerical Mathematics*, 37(1-2):201–230, 2001.

[14] S. K. Godunov, *A finite difference method for the numerical computation of discontinuous solutions of the equations of fluid dynamics*, Mat. Sb., 47 (1959), pp. 271–290.

[15] A. M. Il'in, *A difference scheme for a differential equation with a small parameter multiplying the second derivative*, Matematičeskije zametki, 6 (1969), pp. 237–248.

[16] A. Kufner, O.John, and S.Fučik *Function Spaces.* Academia, Prague, 1977.

[17] J. M. Ortega and W. C. Rheinboldt. *Iterative solution of nonlinear equations in several variables.* Academic Press, New York, 1970.

[18] D. L. Scharfetter and H. K. Gummel, *Large signal analysis of a silicon Read diode*, IEEE Transactions on Electron Devices, 16 (1969), pp. 64–77.

R.Eymard
Université de Marne-la-Vallée
5, boulevard Descartes
Champs-sur-Marne
F-77454 Marne La Vallée Cedex 2

J.Fuhrmann, K.Gärtner
Weierstraß-Institut für Angewandte Analysis und Stochastik
Mohrenstraße 39
D-10117 Berlin