



**HAL**  
open science

# Convergence of approximate solutions to an elliptic-parabolic equation without the structure condition

Boris Andreianov, Petra Wittbold

► **To cite this version:**

Boris Andreianov, Petra Wittbold. Convergence of approximate solutions to an elliptic-parabolic equation without the structure condition. *Nonlinear Differential Equations and Applications*, 2012, 19 (6), pp. 695-717. 10.1007/s00030-011-0148-8 . hal-00608521

**HAL Id: hal-00608521**

**<https://hal.science/hal-00608521>**

Submitted on 13 Jul 2011

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

# Convergence of approximate solutions to an elliptic-parabolic equation without the structure condition

B. Andreianov\* and P. Wittbold†

July 13, 2011

## Abstract

We study the Cauchy-Dirichlet problem for the elliptic-parabolic equation

$$b(u)_t + \operatorname{div} F(u) - \Delta u = f$$

in a bounded domain. We do not assume the structure condition “ $b(z) = b(\hat{z}) \Rightarrow F(z) = F(\hat{z})$ ”. Our main goal is to investigate the problem of continuous dependence of the solutions on the data of the problem and the question of convergence of discretization methods. As in the work of Ammar and Wittbold [3] where existence was established, monotonicity and penalization are the main tools of our study.

In the case of a Lipschitz continuous flux  $F$ , we justify the uniqueness of  $u$  (the uniqueness of  $b(u)$  is well-known) and prove the continuous dependence in  $L^1$  for the case of strongly convergent finite energy data. We also prove convergence of the  $\varepsilon$ -discretized solutions used in the semigroup approach to the problem; and we prove convergence of a monotone time-implicit finite volume scheme. In the case of a merely continuous flux  $F$ , we show that the problem admits a maximal and a minimal solution.

**Keywords:** Elliptic-Parabolic Equation, Structure Condition, Convergence of Approximate Solutions, Finite Volume Method.

## 1 Introduction

Let  $\Omega$  be a bounded domain of  $\mathbb{R}^N$  with Lipschitz boundary. We study the problem

$$\begin{cases} b(u)_t + \operatorname{div} F(u) - \Delta u = f & \text{in } Q = (0, T) \times \Omega \\ u = 0 & \text{on } \Sigma = (0, T) \times \partial\Omega \\ b(u) = b_0 & \text{on } \{0\} \times \Omega, \end{cases} \quad (1)$$

where  $b : \mathbb{R} \rightarrow \mathbb{R}$  is a continuous non-decreasing function normalized by  $b(0) = 0$ . We assume that  $f \in L^2(Q)$  and  $(B \circ b^{-1})(b_0) \in L^1(\Omega)$ , where

$$B(z) = \int_0^1 z (b(z) - b(\sigma z)) d\sigma = z b(z) - \int_0^z b(s) ds \geq 0.$$

We assume that  $F : \mathbb{R} \rightarrow \mathbb{R}^N$  is at least continuous. For the existence of weak solutions with unbounded data, one needs the following growth assumption on  $F$ :

$$\forall z \in \mathbb{R} \quad |F(z)|^2 \leq C(1 + B(z) + |z|^2), \quad (2)$$

---

\*Laboratoire de Mathématiques CNRS UMR 6623, Université de Franche-Comté  
16 route de Gray, 25030 Besançon, France. E-mail: boris.andreianov@univ-fcomte.fr

†Universität Duisburg-Essen, Campus Essen, Fakultät für Mathematik,  
Universitätsstr. 2-17, 45117 Essen, Germany. E-mail: petra.wittbold@uni-due.de

where  $C > 0$ .

We do not assume that  $F(z) = F(\hat{z})$  whenever  $b(z) = b(\hat{z})$ . The latter assumption, known as the “structure condition”, permits to develop a complete well-posedness theory for (1) (see Alt and Luckhaus [2], Otto [23], Bénéilan and Wittbold [12], Carrillo and Wittbold [17]); see also [4] for the explicit continuous dependence results). Our goal is to give some partial uniqueness and continuous dependence results for problem (1) without the structure condition; the closely related issue is convergence of different discretization methods for (1), including those that could be used for a practical computation of solutions.

Existence and uniqueness in dimension one for (1) *without the structure condition*, with a Lipschitz flux  $F$  and under some additional restrictions on the nonlinearity  $b$  and the data, was shown by Bénéilan and the second author in [13]. The proof used the semigroup techniques (see e.g. [11]) in an indirect way. Indeed, the semigroup techniques for (1) are naturally concerned with the convergence of  $b(u^\varepsilon)$ , not with the one of  $u^\varepsilon$ ; and, as a matter of fact, no convergence of  $u^\varepsilon$  was shown. Therefore, beyond the natural generalizations (multi-dimensional case, etc.), the paper [13] left open the following question:

*can one show existence by a passage to the limit  
in a sequence of (well-chosen) approximations of (1)?*

The difficulty here comes from the fact that the standard a priori estimates permit to get the “compactness in  $x$ ” for  $u^\varepsilon$  and the “compactness in  $t$ ” for  $b(u^\varepsilon)$ ; in case  $b$  is constant in some interval of values of  $u$ , oscillations of  $u^\varepsilon$  within this interval cannot be precluded by these a priori estimates.

A first answer to this question was given by Ammar and the second author in [3]. The approximation procedure of [3] is very special: “bi-monotone” approximations of the data and the introduction of a penalization by absorption term of the kind  $\frac{1}{m}\psi(u)$  ( $\psi$  being a bounded strictly increasing continuous function) in the left-hand side of (1) are used. Then the comparison result for the solutions  $u^m$  of the penalized problem (1) permits to use the monotone convergence theorem in order to pass to the limit in the sequence of approximate solutions (i.e., the compactness of the sequence  $(u^m)_m$  of approximate solutions is ensured by its monotonicity). In this way, the existence of solutions for (1) is shown without restrictive assumptions (in fact, [3] treats a much more general equation in the more general framework of renormalized solutions). The method of [3] found many subsequent applications; in particular, it simplifies considerably the study of renormalized solutions to various problems (see [1] for one example).

Still the work [3] left open the question of convergence of “more natural” approximations of the solutions to problem (1). One may think in particular of the sensitivity of the solutions of (1) with respect to small perturbations of the data and of convergence of numerical approximations of (1). In the present paper, we give a (still partial) answer to both questions. The penalization and comparison techniques of [3] remain the essential tools in our study; we combine them with  $\liminf / \limsup$  constructions, “weak” time translation arguments and order-preserving approximation methods. We also prove convergence of sequences  $u^\varepsilon$  generated by the time-implicit discretization scheme used in the nonlinear semigroup approach (recall that the issue of convergence of  $u^\varepsilon$  is a delicate question here, the natural question being the convergence of  $b(u^\varepsilon)$ ); this result was essentially contained in the work of Zimmermann [24]. Continuous dependence is justified in the case of strongly convergent sequences of data. The case of weakly convergent data remains open (this includes the issue of convergence of finite element methods that do not enjoy the order-preservation property).

It is clear that the uniqueness of solutions to (1) (justified under the Lipschitz continuity assumption on  $F$ ) is necessary in order to address the question of stability and convergence of approximation methods. Therefore our results are, in a sense, optimal: indeed, under structure restrictions on approximations (those that allow to use comparison techniques) uniqueness implies continuous dependence on the data and convergence of discretization methods (see the proofs of Theorem 2.3, Theorem 4.1, and Theorem 5.1(ii) below). It is worth mentioning that for a merely continuous convection flux  $F$ , uniqueness of a solution remains a “generic” property with respect to the choice of the data (see Section 2 for the precise statements).

The paper is organized as follows. Uniqueness, stability and existence of maximal/minimal solutions for the continuous problem is studied in the two next sections: definitions and results are given in Section 2, with the proofs collected in Section 3. Section 4 is devoted to the proof of convergence of time-implicit discretizations; its essential ingredient is the time-translation estimate for  $u$  (formula (12), proved in [24], and Lemma 4.2) obtained in presence of a strictly monotone penalization by absorption term  $\psi(u)$ . Section 5 contains the description of a monotone finite volume scheme for (1) together with a translation estimate and a convergence proof inspired by those of Section 4.

## 2 Definitions and results for the continuous problem

Let us introduce  $\Phi : z \mapsto \int_0^z b(s) ds$ , and denote by  $\Phi^*$  the Legendre (Fenchel) transform of  $\Phi$ . We have

$$B(z) = z b(z) - \Phi(z) \quad \text{and} \quad B(z) = \Phi^*(b(z)). \quad (3)$$

Note that for all  $\delta > 0$ , the convexity inequality

$$B(z) - B(r) \geq r(b(z) - b(r))$$

with the choice  $r = \pm \frac{1}{\delta}$  yields the upper bound

$$|b(z)| \leq \delta \Phi^*(b(z)) + \max\{b(\frac{1}{\delta}), -b(-\frac{1}{\delta})\} = \delta B(z) + \max\{b(\frac{1}{\delta}), -b(-\frac{1}{\delta})\}. \quad (4)$$

**Definition 2.1.** Let  $b_0 \in L^1(\Omega)$ ,  $f \in L^1(Q)$ . A weak solution of (1) is a function  $u : Q \mapsto \mathbb{R}$  satisfying

- (i)  $u \in L^2(0, T; H_0^1(\Omega))$ ,  $B(u) \in L^\infty(0, T; L^1(\Omega))$  (whence  $b(u) \in L^1(Q)$  by (3),(4));
- (ii) for all  $\zeta \in L^2(0, T; H_0^1(\Omega))$  with  $\zeta_t \in L^\infty(Q)$  and  $\zeta(T) = 0$ ,

$$\int \int_Q b(u) \zeta_t + \int_\Omega b(u_0)(\cdot) \zeta(0, \cdot) = \int \int_Q (\nabla u + F(u)) \cdot \nabla \zeta - \int \int_Q f \zeta. \quad (5)$$

Note that weak solutions only exist under additional integrability assumptions on  $b_0$  and  $f$ . More exactly, we consider the class

$$D := \{(b_0, f) \in L^1(\Omega) \times L^1(Q) \mid \Phi^*(b_0) \in L^1(\Omega), f \in L^2(Q)\},$$

which is the subclass of so-called ‘‘finite energy data’’. An appropriate framework for (1) with pure  $L^1$  data is the one of renormalized solutions (see [3]); in this framework, also the growth assumption (2) can be dropped, see e.g. [17, 3, 24]. Although the results of Sections 2, 4 extend to this framework (the issue of numerical approximation of renormalized solutions can be more delicate, thus extension of results of Section 5 is delicate), we focus on weak solutions in order to make the ideas clear.

**Remark 2.2.** It is well known that relation (5) implies that the distributional derivative  $b(u)_t$  of  $b(u)$  can be identified with an element  $\chi$  of the space  $L^2(0, T; H^{-1}(\Omega)) + L^1(Q)$ , which is included into the dual space  $E'$  of the space  $E := L^2(0, T; H_0^1(\Omega)) \cap L^\infty(Q)$ . More exactly, we have

$$- \int \int_Q b(u) \zeta_t - \int_\Omega b_0(\cdot) \zeta(0, \cdot) = \int_0^T \langle \chi, \zeta \rangle$$

for all  $\zeta \in L^2(0, T; H_0^1(\Omega))$  with  $\zeta_t \in L^\infty(Q)$  and  $\zeta(T) = 0$  (which implies that  $\zeta \in L^\infty(Q)$ ). Here and in the sequel,  $\langle \cdot, \cdot \rangle$  denotes the duality pairing between  $W^{-1, p'} + L^1$  and  $W_0^{1, p} \cap L^\infty$ .

This interpretation allows for the crucial chain rule formula (see e.g. [2, 23]) :

$$\begin{aligned} - \int_\Omega (\Psi_h \circ b_0) \mu(0) - \int_0^T \int_\Omega (\Psi_h \circ b(u))(t) \sigma_t(t) \\ = \int_0^T \langle \chi(t), (h \circ u)(t) \rangle > \sigma(t) \quad \forall \sigma \in \mathcal{D}([0, T]), \end{aligned} \quad (6)$$

for all Carathéodory function  $h : \Omega \times \mathbb{R} \mapsto \mathbb{R}$ , with  $\Psi_h : (x, z) \in \Omega \times \mathbb{R} \mapsto \int_0^{b^{-1}(z)} h(x, s) db(s)$ . Notice that  $\Psi_h$  is well defined. For the sake of being definite, we can replace the maximal monotone graph  $b^{-1}$  by its minimal section  $[b^{-1}]^0$ ; we also have  $\Psi_h(z) = \int_0^z (h \circ [b^{-1}]^0)(r) dr$ . In particular, if  $h(x, z) = z$ , then  $\Psi_h(x, \cdot) \equiv \Phi^*(\cdot)$  for all  $x \in \Omega$ .

We show the following results.

**Theorem 2.3.** *Assume that  $F$  satisfies (2).*

(i) *Assume in addition that  $F$  is Lipschitz continuous. Then for all data  $(b_0, f) \in D$ , there exists one and only one weak solution  $u_{b_0, f}$  of (1). Furthermore, consider a sequence  $(b_0^n, f^n)_{n \in \mathbb{N}} \subset D$  of data such that  $f^n \rightarrow f$  in  $L^2(Q)$ ,  $b_0^n \rightarrow b_0$  a.e. on  $\Omega$  and  $\Phi^*(b_0^n) \rightarrow \Phi^*(b_0)$  in  $L^1(\Omega)$ . Then the sequence  $u^n = u_{b_0^n, f^n}$  of the associated weak solutions of (1) converges to  $u_{b_0, f}$  weakly in  $L^2(0, T; H_0^1(\Omega))$  and strongly in  $L^2(Q)$ .*

(ii) *For a general merely continuous  $F$ , for all data  $(b_0, f) \in D$ , there exist weak solutions  $\underline{u}_{b_0, f}, \bar{u}_{b_0, f}$  of (1) such that for all weak solution  $u$  with the same data one has  $\underline{u}_{b_0, f} \leq u \leq \bar{u}_{b_0, f}$  a.e. on  $Q$ . In addition, if  $f \leq \hat{f}$  a.e. on  $Q$  and  $b_0 \leq \hat{b}_0$  a.e. on  $\Omega$ , then the associated maximal and minimal solutions satisfy  $\underline{u}_{b_0, f} \leq \underline{u}_{\hat{b}_0, \hat{f}}$ ,  $\bar{u}_{b_0, f} \leq \bar{u}_{\hat{b}_0, \hat{f}}$  a.e. on  $Q$ . If  $f < \hat{f}$  a.e. on  $Q$  and  $b_0 \leq \hat{b}_0$  a.e. on  $\Omega$ , then  $\bar{u}_{b_0, f} \leq \underline{u}_{\hat{b}_0, \hat{f}}$  a.e. on  $Q$ .*

(iii) *With the notation of (ii), consider a sequence of data  $(b_0^n, f^n)_{n \in \mathbb{N}} \subset D$  satisfying the assumptions of the point (i). Then each sequence  $(u^n)_{n \in \mathbb{N}}$  of solutions associated with these data admits a subsequence weakly convergent in  $L^2(0, T; H_0^1(\Omega))$ , and such limits take their values in  $[\underline{u}_{b_0, f}(t, x), \bar{u}_{b_0, f}(t, x)]$ , for a.e.  $(t, x) \in Q$ . Furthermore,  $b(u)$  does not depend on the choice of the subsequence, so that the whole sequence  $b(u^n)$  converges to  $b(\underline{u}_{b_0, f}) \equiv b(\bar{u}_{b_0, f})$  in  $L^1(Q)$ .*

*In case there exists a unique solution to (1) with the data  $(b_0, f)$ , the claim of (i) still holds.*

(iv) *Let  $(b_0, f) \in D$  and  $g \in L^2(Q)$ ,  $g > 0$  a.e. on  $Q$ . Then there is uniqueness of a weak solution of (1) with the datum  $(b_0, f + \lambda g)$  for all  $\lambda \in \mathbb{R}$  except for, maybe, an at most countable set.*

Notice that, although we are unable to deduce well-posedness for weak solutions of (1) with a general flux  $F$  satisfying (2), the result of (iv) indicates that the non-uniqueness of a weak solution is a rather exceptional situation with respect to the choice of the source term.

**Remark 2.4.**

(i) *In case  $b(\pm\infty) = \pm\infty$ , for  $b_0 \in L^\infty(\Omega)$  and  $f$  such that  $\int_0^T \|f(t, \cdot)\|_{L^\infty(\Omega)} dt < +\infty$ , the comparison principle for  $b(u)$  yields an  $L^\infty$  bound on  $u$ . For such data, the local Lipschitz continuity of  $F$  is sufficient for the uniqueness claim of Theorem 2.3(i) (cf. [13] for a similar uniqueness result).*

(ii) *Notice that the same results can be obtained for the same equation as in (1) with non-homogeneous Dirichlet boundary condition on  $u$  in the space  $L^2(0, T; H^{1/2}(\partial\Omega))$  (see [8] for the comparison principle including the variation of boundary conditions).*

(iii) *For a general Dirichlet boundary condition, the point (i) of Theorem 2.3 can be shown with the Laplacian  $-\Delta u$  replaced by the  $p$ -Laplacian operator with  $p \leq 2$ , or more generally, with a Leray-Lions operator  $-\operatorname{div} \mathbf{a}(\nabla u)$  satisfying the coercivity estimate*

$$\forall \xi, \hat{\xi} \in \mathbb{R}^N, \quad |\xi - \hat{\xi}|^2 \leq \operatorname{const}(1 + |\xi|^p + |\hat{\xi}|^p)^{\frac{2-p}{p}} (\mathbf{a}(\xi) - \mathbf{a}(\hat{\xi})) \cdot (\xi - \hat{\xi}). \quad (7)$$

(iv) *In case of a general Leray-Lions operator (e.g., in case of the  $p$ -Laplacian with  $p > 2$ ), the points (ii)-(iii) of Theorem 2.3 are still true. Recall that uniqueness in this case in general does not hold, see Boccardo et al. [14].*

### 3 Uniqueness and continuous dependence: the proofs

Let us start with two lemmas.

**Lemma 3.1.** *Assume  $F$  is a merely continuous function. Assume that  $u_i$ ,  $i = 1, 2$ , are weak solutions of (1) associated with the data  $(b_0^i, f^i) \in D$  such that  $b_0^1 \leq b_0^2$  and  $f_1 < f_2$  a.e. on  $Q$ . Then  $u_1 \leq u_2$  a.e. on  $Q$ .*

The proof of the lemma given below generalizes directly to a general Leray-Lions operator with the homogeneous Dirichlet boundary condition, as well as to the Laplacian with inhomogeneous Dirichlet conditions (see the references [3, 7, 8, 9]).

PROOF : It was shown by Carrillo in [16] (see also [17, 3, 7, 8] for generalizations) that under the assumptions of the lemma,  $b(u_1) \leq b(u_2)$  a.e. on  $Q$ . Let us recall the arguments of the proof; our claim will be a simple consequence of this proof.

One uses the Kruzhkov doubling of variables in time, as in Otto [23], and (in case  $F$  is not Hölder continuous of order at least 1/2) in space, as in Carrillo [16]. We set  $T_\varepsilon(r) = \min\{\varepsilon, r^+\}$ . Taking the test functions  $\frac{1}{\varepsilon} T_\varepsilon(u_1(t, x) - u_2(s, y)) \xi(t, s, x, y)$  in the weak formulations for  $u_1(t, x)$  and  $u_2(s, y)$ , using the chain rule (6), with  $\varepsilon \downarrow 0$  we find the inequality

$$\begin{aligned} & - \iint_Q \iint_Q (b(u_1) - b(u_2))^+ (\xi_t + \xi_s) - \kappa (f_1 - f_2) \xi \\ & + \iint_Q \iint_Q \text{sign}^+(u_1 - u_2) (F(u_1) - \nabla u_1 - F(u_2) + \nabla u_2) \cdot (\nabla_x \xi + \nabla_y \xi) \leq 0, \end{aligned}$$

for all  $\xi \in \mathcal{D}((0, T)^2 \times \Omega^2)^+$ ; here  $\kappa \in L^\infty(Q)$  is such that  $((u_1 - u_2), \kappa)$  belongs to the maximal monotone graph  $\text{sign}^+$  almost everywhere. Here the functions  $u_1, f_1, \nabla u_1$  depend on  $(t, x) \in Q$ , and  $u_2, f_2, \nabla u_2$  depend on  $(s, y) \in Q$ . Choosing the test function  $\xi$  under the form  $\zeta(t, x) \rho_n(t - s, x - y)$ , where  $(\rho_n)_n$  approximates the Dirac  $\delta$ -function at 0, we deduce the so-called Kato inequality

$$\begin{aligned} & - \iint_Q (b(u_1) - b(u_2))^+ \zeta_t + \text{sign}^+(u_1 - u_2) (F(u_1) - \nabla u_1 - F(u_2) + \nabla u_2) \cdot \nabla \zeta \\ & \leq \iint_Q \text{sign}^+(u_1 - u_2) (f_1 - f_2) \zeta + \iint_Q \mathbb{1}_{\{(t, x) \mid u_1(t, x) = u_2(t, x)\}} (f_1 - f_2)^+ \zeta \leq 0, \end{aligned} \quad (8)$$

where both  $u_1$  and  $u_2$  now depend on the same variables  $(t, x) \in Q$ , with  $\zeta \in \mathcal{D}((0, T) \times \Omega)^+$ . Using the assumption  $(b_0^1 - b_0^2)^+ = 0$  and exploiting the entropy inequalities deduced from the weak formulation, we extend (8) to the case  $\zeta \in \mathcal{D}([0, T] \times \Omega)^+$  (see e.g. [16, 5, 9]). As in [8, 9], letting a well-chosen sequence  $(\zeta_n)_{n \in \mathbb{N}} \subset \mathcal{D}([0, T] \times \Omega)^+$  tend to the characteristic function of  $(0, t) \times \Omega$ , we deduce that

$$\text{for a.e. } t \in (0, T), \quad \int_\Omega (b(u_1) - b(u_2))^+(t) \leq 0.$$

Different techniques for the treatment of the homogeneous Dirichlet boundary condition for problem (1) can be found in [16, 7].

Now notice that in the above formula (8), one can keep the term

$$\iint_Q \kappa (f_1 - f_2) \zeta.$$

Then the same reasoning yields in addition the inequality

$$\iint_{\{(t, x) \mid u_1(t, x) > u_2(t, x)\}} (f_1 - f_2) \geq 0.$$

As we have assumed that  $f_1 < f_2$  a.e. on  $Q$ , it follows that  $u_1 \leq u_2$  a.e. on  $Q$ .  $\diamond$

**Remark 3.2.** Note that in the setting of Lemma 3.1, one can show the strict inequality  $u_1 < u_2$  at least in the case where  $F$  is Hölder continuous of order 1/2 (in this case, the doubling of variables in space can be avoided, cf. [2, 23], and inequality (8) can be strengthened).

**Lemma 3.3.** Assume that  $F$  is Lipschitz continuous on  $\mathbb{R}$ . Assume that  $u_i, i = 1, 2$ , are weak solutions of (1) associated with the data  $(b_0^i, f^i) \in D$  such that  $b_0^1 \leq b_0^2$  and  $f_1 \leq f_2$  a.e. on  $Q$ . Then  $u_1 \leq u_2$  a.e. on  $Q$ .

The proof of this lemma can be generalized to Leray-Lions operators with Dirichlet boundary conditions, but it requires the coercivity estimate (7).

For the proof, we use capacity estimates; an alternative proof is obtained using the technique of test functions introduced by Brézis, Kinderlehrer and Stampaccia in [15].

PROOF : We start with inequality (8) deduced e.g. in the paper [23] by Otto (see also the proof of the above Lemma 3.1). As we already know that  $(b(u_1) - b(u_2))^+ = 0$  a.e. on  $Q$ , we get

$$\iint_Q \text{sign}^+(u_1 - u_2) \left( \nabla(u_1 - u_2) + (F(u_1) - F(u_2)) \right) \cdot \nabla \zeta \leq 0 \quad (9)$$

for all  $\zeta \in L^2(0, T; H_0^1(\Omega))^+$ .

Recall that one can define the capacity (relative to the set  $\Omega$ )  $\text{cap}(A)$  associated with a family of subsets  $A$  of  $\Omega$  in the following way (see e.g. [19]):

$$\text{cap}(A) = \inf \left\{ \int_{\Omega} |\nabla w|^2 \mid w \in H_0^1(\Omega), w = 1 \text{ a.e. in a neighbourhood of } A, 0 \leq w \leq 1 \text{ a.e. in } \Omega \right\}.$$

Note that if  $\text{cap}(A) = 0$ , then  $A$  is Lebesgue measurable and the measure of  $A$  is zero. Moreover, any function  $u \in H_0^1(\Omega)$  has a precise (and even quasi-continuous) representative which is defined quasi-everywhere, and one can prove that, for any set  $A \subset \Omega$ ,

$$\text{cap}(A) = \inf \left\{ \int_{\Omega} |\nabla w|^2 \mid w \in H_0^1(\Omega), w \geq \mathbb{1}_A \text{ quasi-everywhere on } \Omega \right\}.$$

Now take in (9) the test function  $\zeta = \frac{1}{\varepsilon^2} T_{\varepsilon}(u_1 - u_2)^+ \in L^2(0, T; H_0^1(\Omega))^+$  Using the Lipschitz continuity of  $F$ , we deduce that

$$\iint_Q \left| \nabla \frac{1}{\varepsilon} T_{\varepsilon}(u_1 - u_2)^+ \right|^2 \leq L \iint_Q \frac{1}{\varepsilon} (u_1 - u_2)^+ \left| \nabla \frac{1}{\varepsilon} T_{\varepsilon}(u_1 - u_2)^+ \right|$$

with some constant  $L$ . Setting  $v := (u_1 - u_2)^+$  (we pick here a representative such that  $v(t, \cdot)$  is quasicontinuous in  $\Omega$ , for a.e.  $t \in (0, T)$ ) and using the Cauchy-Schwarz inequality, we get

$$\iint_Q \left| \nabla \frac{1}{\varepsilon} T_{\varepsilon}(v) \right|^2 \leq L^2 \iint_{\{(t,x) \mid 0 < v(t,x) < \varepsilon\}} \left| \frac{1}{\varepsilon} T_{\varepsilon}(v) \right|^2. \quad (10)$$

Now, the right-hand side of (10) is upper bounded by the measure of the set  $\{(t, x) \mid 0 < v(t, x) < \varepsilon\}$ , which tends to zero as  $\varepsilon \rightarrow 0$ , because  $\frac{1}{\varepsilon} T_{\varepsilon}$  is upper bounded by 1. The left-hand side of (10) is lower bounded by the integral in  $t$  of the capacity of the set  $S_{\varepsilon}^t := \{x \in \Omega \mid v(t, x) > \varepsilon\}$ . Indeed, for a.e.  $t \in (0, T)$ , one has  $w(\cdot) := \frac{1}{\varepsilon} T_{\varepsilon}(v(t, \cdot)) \in H_0^1(\Omega)$  and  $w \geq 1$  quasi-everywhere on  $S_{\varepsilon}^t$  by the choice of the representative of  $v$ . Because  $\text{cap}(S_{\varepsilon}^t)$  does not decrease as  $\varepsilon \rightarrow 0^+$ , we deduce that for all  $\varepsilon > 0$ ,  $\int_0^T \text{cap}(S_{\varepsilon}^t) dt = 0$ . Therefore we have for a.e.  $t \in (0, T)$ ,  $u_1(t, \cdot) \geq u_2(t, \cdot)$  a.e. on  $\Omega$ .  $\diamond$

Now we deduce the claims of Theorem 2.3 using the standard  $\liminf / \limsup$  hint.

PROOF OF THEOREM 2.3:

(i) The existence of a so-called renormalized solution  $u_{b_0, f}$  to (1) was shown by Ammar and Wittbold in [3]. The assumption  $(b_0, f) \in D$  and the growth assumption (2) lead to a uniform  $L^2(0, T; H_0^1(\Omega))$  estimate on the sequence of the approximate solutions used in [3]. It follows that the so constructed  $u_{b_0, f}$  is also a weak solution of (1).

The uniqueness of  $u_{b_0, f}$  follows from Lemma 3.3. Let us prove the continuous dependence property. Denote  $u^n := u_{b_0^n, f^n}$ . Because of the uniqueness of  $u_{b_0, f}$ , it is sufficient to show that each subsequence of  $(u^n)_n$  possesses itself a subsequence converging to  $u_{b_0, f}$  strongly in  $L^2(Q)$  and weakly in  $L^2(0, T; H_0^1(\Omega))$ .

In the sequel, we work with an extracted (not relabelled) subsequence such that  $(f^n)_n$  is dominated by a function  $h \in L^2(Q)$ , and  $(b_0^n)_n$  satisfies  $g^- \leq b_0^n \leq g^+$ ,  $\pm g^{\pm} \geq 0$  a.e. on  $\Omega$ , and  $\Phi^*(g^{\pm}) \in L^1(\Omega)$ . We can define pointwise (a.e. on  $Q$ ) the functions  $m^n := \inf_{k \geq n} f^k$  and  $M^n := \sup_{k \geq n} f^k$ ; these functions are a.e. finite and belong to  $L^2(Q)$  because of the inequality  $-h \leq m^n \leq f^n \leq M^n \leq h$  a.e. on  $Q$ . Similarly, we define  $\mu^n := \inf_{k \geq n} b_0^k$  and  $\mathcal{M}^n := \sup_{k \geq n} b_0^k$ ; we have  $\Phi^*(\mu^n), \Phi^*(\mathcal{M}^n) \in L^1(\Omega)$  because  $g^- \leq \mu^n \leq b_0^n \leq \mathcal{M}^n \leq g^+$  a.e. on  $\Omega$ .

As the data  $(\mu^n, m^n)$  and  $(\mathcal{M}^n, M^n)$  belong to  $D$ , there exist weak solutions of (1) associated with these data, which we denote by  $\underline{u}^n$  and  $\bar{u}^n$ , respectively. Similarly, there exist weak solutions  $\underline{U}$  and  $\bar{U}$  of (1) associated with the data  $(g^-, -h)$  and  $(g^+, h)$ , respectively. Also the uniform estimates on  $\|\nabla \underline{u}^n\|_{L^2(Q)}$  and on  $\|\underline{u}^n\|_{L^2(Q)}$  in terms of  $\int_{\Omega} \Phi^*(g^{\pm})$  and  $\iint_Q |h|^2$  hold true; the same estimates hold with  $\underline{u}^n$  replaced with  $\bar{u}^n$ .

By construction,  $(\mu^n)_n, (m^n)_n$ , resp.  $(\mathcal{M}^n)_n, (M^n)_n$ , are non-decreasing, resp. non-increasing sequences. By the comparison principle of Lemma 3.3,  $(\underline{u}^n)_n$ , resp.  $(\bar{u}^n)_n$ , is also non-decreasing, resp. non-increasing. Similarly, both sequences are lower bounded by  $\underline{U} \in L^2(Q)$ , and upper bounded by  $\bar{U} \in L^2(Q)$ . Therefore  $(\underline{u}^n)_n$ , resp.  $(\bar{u}^n)_n$ , converge a.e. on  $Q$  and in  $L^2(Q)$  to some limits  $\underline{u}$ , resp.  $\bar{u}$ , as  $n \rightarrow \infty$ . Using the fact that  $|F(\underline{u}^n)|^2, |F(\bar{u}^n)|^2$  are dominated by the  $L^1(Q)$  function  $(B(\underline{U}) + B(\bar{U}) + \underline{U}^2 + \bar{U}^2)$ , we deduce that  $F(\underline{u}^n)$ , resp.  $F(\bar{u}^n)$  converge in  $L^2(Q)$  to  $F(\underline{u})$ , resp. to  $F(\bar{u})$ . In addition, one can extract from  $(\underline{u}^n)_n$  and  $(\bar{u}^n)_n$  subsequences weakly convergent in the space  $L^2(0, T; H_0^1(\Omega))$ ; by uniqueness of a weak limit in  $L^2(Q)$ , these limits coincide with  $\underline{u}$  and  $\bar{u}$ , respectively. Finally, by construction we have  $m^n, M^n \rightarrow f$  in  $L^2(Q)$  as  $n \rightarrow \infty$ ; also  $\mu^n, \mathcal{M}^n \rightarrow b_0$  in  $L^1(\Omega)$ , because of (4).

The convergences obtained above permit to pass to the limit and deduce that  $\underline{u}$  and  $\bar{u}$  are weak solutions of (1); by the uniqueness, both functions coincide with  $u_{b_0, f}$ . Because  $\underline{u}^n \leq u^n \leq \bar{u}^n$  and thanks to the uniform bound on  $\|\nabla u^n\|_{L^2(Q)}$ ,  $(u^n)_n$  admits a subsequence that converges to  $u_{b_0, f}$  in  $L^2(Q)$  strongly and in  $L^2(0, T; H_0^1(\Omega))$  weakly. Hence the claim follows.

(ii) The proof is the same as the continuous dependence proof in (i), except that we use the comparison principle of Lemma 3.1. We start with the sequence of data  $\mu^n = b_0 = \mathcal{M}^n$ ,  $m^n = f - \frac{1}{n}$ ,  $M^n = f + \frac{1}{n}$ ,  $n \in \mathbb{N}$ . These data have all the properties used in the proof of (i), moreover, for all  $n \in \mathbb{N}$ ,  $m^n < m^{n+1}$  and  $M^n > M^{n+1}$  a.e. on  $Q$ . By  $\underline{u}^n$ , resp. by  $\bar{u}^n$  we denote any weak solution of (1) associated with the data  $(\mu^n, m^n)$ , resp. with  $(\mathcal{M}^n, M^n)$ . The existence is justified as in the point (i). By Lemma 3.1, we deduce that  $(\underline{u}^n)_n$  is non-decreasing,  $(\bar{u}^n)_n$  is non-increasing, and for all weak solution  $u$  of (1),  $\underline{u}^n \leq u \leq \bar{u}^n$ . As in (i), we deduce that the limits  $\underline{u}_{b_0, f}$ , resp.  $\bar{u}_{b_0, f}$ , of  $(\underline{u}^n)_n$ , resp. of  $(\bar{u}^n)_n$ , are weak solutions of (1), and  $\underline{u}_{b_0, f} \leq u \leq \bar{u}_{b_0, f}$ .

In the sequel, we slightly modify the construction of  $m^n, M^n$  in order to deduce the subsequent claims. Notice that, because the smallest solution  $\underline{u}_{b_0, f}$  and the greatest solution  $\bar{u}_{b_0, f}$  of (1) are obviously unique, any other approximations of  $f$  by  $m^n, M^n$  having the same strict monotonicity properties converge to the same limits.

The monotonicity of the maps  $(b_0, f) \mapsto \underline{u}_{b_0, f}$ ,  $(b_0, f) \mapsto \bar{u}_{b_0, f}$  is obtained by taking  $m^n = f - \frac{1}{n}$ ,  $M^n = f + \frac{1}{n}$  and  $\hat{m}^n = \hat{f} - \frac{1}{2n}$ ,  $\hat{M}^n = \hat{f} + \frac{1}{2n}$ . Then for all  $n \in \mathbb{N}$ ,  $m^n < \hat{m}^n$ ,  $M^n < \hat{M}^n$  a.e. on  $Q$ , and Lemma 3.1 ensures the comparison principle for the associated solutions.

In case we have in addition  $f < \hat{f}$  a.e. on  $Q$ , we define the a.e. positive function  $d := \hat{f} - f$  and use the data  $M^n := f + \frac{d}{3n}$ ,  $\hat{m}^n := \hat{f} - \frac{d}{3n}$ . We have  $M^n < \hat{m}^n$  a.e. on  $Q$ , which implies that  $\bar{u}_{b_0, f} \leq \underline{u}_{\hat{b}_0, \hat{f}}$  a.e. on  $Q$ .

(iii) The first claim is similar to the point (ii). Starting with the appropriately chosen subsequences  $(f^n)_n, (b_0^n)_n$  as in (i), we denote by  $\underline{u}^n$ , resp. by  $\bar{u}^n$  the solutions of (1) associated to the data  $(\mu^n, m^n - \frac{1}{n})$ , resp. with  $(\mathcal{M}^n, M^n + \frac{1}{n})$ . We have  $\underline{u}^n \uparrow \underline{u}_{b_0, f}$ ,  $\bar{u}^n \downarrow \bar{u}_{b_0, f}$ , and  $\underline{u}^n \leq u^n \leq \bar{u}^n$  by Lemma 3.1. The latter inequality being preserved by the passage to the weak limit in  $L^2(0, T; H_0^1(\Omega))$ , we infer that  $\underline{u}_{b_0, f} \leq u \leq \bar{u}_{b_0, f}$ .

The convergence claim for  $b(u^n)$  follows directly from the uniqueness and continuous dependence result for  $b(u)$  shown in [16, 7, 8]. Finally, uniqueness of a weak solution for the data  $(b_0, f)$  means that  $\underline{u}_{b_0, f} \equiv \bar{u}_{b_0, f}$ , in which case the limit  $u$  does not depend on the choice of the weakly convergent in  $L^2(0, T; H_0^1(\Omega))$  subsequence of  $u^n$ , and  $u^n$  converges strongly in  $L^2(Q)$ .

(iv) Consider the function  $\pi : \lambda \in \mathbb{R} \mapsto \iint_Q \arctan(\bar{u}_{b_0, f + \lambda g})$ . Clearly, it is well defined. By (ii),  $\pi$  is non-decreasing; therefore it has an at most countable number of discontinuities. If  $\lambda_0$  is a continuity point of  $\pi$ , it follows from the monotonicity of the map  $\lambda \mapsto \bar{u}_{b_0, f + \lambda g}$  that  $\bar{u}_{b_0, f + \lambda g} \uparrow \bar{u}_{b_0, f + \lambda_0 g}$  as  $\lambda \uparrow \lambda_0$ . But for  $\lambda < \lambda_0$ , the point (ii) and the assumption  $g > 0$  also yield the inequality  $\bar{u}_{b_0, f + \lambda g} \leq \underline{u}_{b_0, f + \lambda_0 g}$ . Thus  $\bar{u}_{b_0, f + \lambda_0 g} \leq \underline{u}_{b_0, f + \lambda_0 g}$ , which implies uniqueness of a weak solution of (1) with the data  $(b_0, f + \lambda_0 g)$ .  $\diamond$



## 4 Convergence of $\varepsilon$ -discretizations

The paper [13] left open the question of convergence of solutions to  $\varepsilon$ -discretizations for problem (1). Recall that in the context of the nonlinear semigroup approach (see e.g. [11]), one constructs mild solutions of the abstract evolution problem

$$v_t + Av = f, \quad v(0) = b_0$$

(here  $A$  is the multi-valued nonlinear operator on  $L^1$  corresponding to the formal expression  $\operatorname{div} F \circ b^{-1}(v) - \Delta b^{-1}(v)$  with the homogeneous Dirichlet boundary condition for  $u \in b^{-1}(v)$ ).

More precisely, mild solutions are  $C([0, T]; L^1(\Omega))$ -limits of  $v^\varepsilon := b(u^\varepsilon)$ , where  $u^\varepsilon$  is the solution of the time-implicit  $\varepsilon$ -discretized problem:  $u^\varepsilon(t, x) = \sum_{n=1}^{N_\varepsilon} u^n(x) \mathbb{1}_{(t_{n-1}, t_n)}(t)$  with

$$\begin{cases} \frac{b(u^n) - b(u^{n-1})}{\Delta t_n} + \operatorname{div} F(u^n) - \Delta u^n + \frac{1}{m} \psi(u^n) = f_\varepsilon^n, \\ u^n \in H_0^1(\Omega), \\ b(u^0) = b_0 \end{cases} \quad n = 1 \dots N_\varepsilon, \quad (11)$$

with  $0 < \Delta t_n \leq \varepsilon$  for all  $n$ ,  $t_0 = 0$ ,  $t_i := \sum_{n=1}^i \Delta t_n$  with  $t_{N_\varepsilon} = T$ , and with  $(f_\varepsilon^n)_n \subset L^1(\Omega)$  satisfying  $\sum_{n=1}^{N_\varepsilon} f_\varepsilon^n(x) \mathbb{1}_{(t_{n-1}, t_n)}(t) \rightarrow f(t, x)$  in  $L^1(Q)$ . In the  $\varepsilon$ -discretization (11) of problem (1), we formally set  $m = \infty$  (values  $m \in \mathbb{N}$  and strictly increasing functions  $\psi \in C(\mathbb{R})$  will be needed later for our arguments).

The function  $b$  not being invertible, convergence of  $v^\varepsilon = b(u^\varepsilon)$  does not imply the convergence of  $u^\varepsilon$ . Now, assume a solution  $u$  to (1) is unique (according to Lemma 3.3, this is always the case if  $F$  is Lipschitz continuous; according to Theorem 2.3, uniqueness remains a “generic” property if  $F$  is only continuous). Then a proof of convergence of  $u^\varepsilon$  to  $u$  is essentially contained in the work of Zimmermann [24].

Indeed, solutions  $u_{\psi, m}^\varepsilon$  of (11) for  $m \in \mathbb{N}$  and a non-decreasing continuous function  $\psi : \mathbb{R} \mapsto \mathbb{R}$  exist (for each  $n = 1 \dots N_\varepsilon$ , the proof of existence of  $u_{\psi, m}^\varepsilon$  is analogous to the proof of existence of  $u^n$ , see e.g. [24]); moreover, [24, Lemma 4.3.10] states that whenever  $\psi$  in (11) is non-decreasing and  $m < \infty$  is fixed, the associated family of solutions  $((u_{\psi, m}^\varepsilon)_\varepsilon)$  admits a uniform estimate of the time translates of  $\psi(u_{\psi, m}^\varepsilon)$  under the following form:

$$\forall s, \Delta > 0 \text{ with } s + \Delta < T, \quad \int_s^{T-\Delta} \int_\Omega |\psi(u_{\psi, m}^\varepsilon)(t+\Delta, x) - \psi(u_{\psi, m}^\varepsilon)(t, x)| \leq C(m, s) \omega(\Delta) \quad (12)$$

where  $\omega : \mathbb{R}^+ \mapsto \mathbb{R}^+$  is a modulus of continuity, in particular,  $\omega(\Delta) \rightarrow 0$  as  $\Delta \rightarrow 0^+$ . Taking  $\psi$  continuous and strictly increasing, using the continuity of  $\psi^{-1}$  and the uniform  $L^2(0, T; H_0^1(\Omega))$  estimate on  $u_{\psi, m}^\varepsilon$  (recall that we take data with finite energy), by the Fréchet-Kolmogorov compactness criterion we deduce the relative compactness of  $(u_{\psi, m}^\varepsilon)_\varepsilon$  in  $L^1((s, T) \times \Omega)$ . Taking  $s_k := \frac{1}{k}$  with  $k \in \mathbb{N}$ , by the diagonal extraction argument we obtain a sequence  $(\varepsilon_l)_{l \in \mathbb{N}}$  convergent to zero (in the sequel, the subscript  $l$  is omitted) such that  $u_{\psi, m}^{\varepsilon_l}$  converge to some limit  $u_{\psi, m}$  in  $L^1(Q)$  as  $\varepsilon_l \rightarrow 0$ . Then it is a standard issue (see, e.g., [24]) to show that  $u_{\psi, m}$  is a weak solution of the problem

$$\begin{cases} b(u)_t + \operatorname{div} F(u) - \Delta u + \frac{1}{m} \psi(u) = f & \text{in } Q = (0, T) \times \Omega \\ u = 0 & \text{on } \Sigma = (0, T) \times \partial\Omega \\ b(u) = b_0 & \text{on } \{0\} \times \Omega. \end{cases} \quad (1_{\psi, m})$$

Now observe that the following comparison principle is easy to justify using essentially the same tools that ensure convergence of  $(b(u_{\psi, m}^\varepsilon))_\varepsilon$  (see [24] again):

$$\begin{cases} \text{if } \psi > 0 \text{ on } \mathbb{R}, & \text{then } (u_{\psi, m}^\varepsilon)_m \text{ is non-decreasing;} \\ \text{if } \psi < 0 \text{ on } \mathbb{R}, & \text{then } (u_{\psi, m}^\varepsilon)_m \text{ is non-increasing;} \\ \text{if } \psi_- < 0 < \psi_+ \text{ on } \mathbb{R}, & \text{then } u_{\psi_-, m}^\varepsilon \geq u_{\psi_+, m}^\varepsilon \text{ a.e. on } Q. \end{cases} \quad (13)$$

In particular, using

$$\psi_\pm : z \mapsto \pm \frac{\pi}{2} + \arctan z, \quad (14)$$

we get the inequalities

$$\forall m \in \mathbb{N} \quad \forall \varepsilon > 0 \quad u_{\psi_+, m}^\varepsilon \leq u^\varepsilon \leq u_{\psi_-, m}^\varepsilon \quad (15)$$

and for the limits  $u_{\psi_{\pm},m}$  of  $(u_{\psi_{\pm},m}^{\varepsilon})_{\varepsilon}$ , we get

$$\forall m \in \mathbb{N} \quad u_{\psi_+,m} \leq u_{\psi_+,m+1} \leq u_{\psi_-,m+1} \leq u_{\psi_-,m}.$$

This provides the following information. Firstly,  $(u_{\psi_{\pm},m})_m$  are monotone sequences. As in the work of Ammar and Wittbold [3], the monotonicity allows for an easy passage to the limit in  $(\mathbf{1}_{\psi,m})$  as  $m \rightarrow \infty$ ; since both the terms  $\frac{1}{m}\psi_{\pm}(\cdot)$  vanish in  $L^{\infty}$  as  $m \rightarrow \infty$ , the functions

$$\underline{u} := \lim_{m \rightarrow \infty} u_{\psi_+,m} \quad \text{and} \quad \bar{u} := \lim_{m \rightarrow \infty} u_{\psi_-,m}$$

are weak solutions of (1). By our uniqueness assumption, we have  $\underline{u} = \bar{u}$ . Finally, inequalities (15) and the definition of  $u_{\psi_{\pm},m}$  yield

$$\begin{aligned} \underline{u} &= \lim_{m \rightarrow \infty} u_{\psi_+,m} = \lim_{m \rightarrow \infty} \lim_{\varepsilon \rightarrow 0} u_{\psi_+,m}^{\varepsilon} \leq \liminf_{\varepsilon \rightarrow 0} u^{\varepsilon} \\ &\leq \limsup_{\varepsilon \rightarrow 0} u^{\varepsilon} \leq \lim_{m \rightarrow \infty} \lim_{\varepsilon \rightarrow 0} u_{\psi_-,m}^{\varepsilon} = \lim_{m \rightarrow \infty} u_{\psi_-,m} = \bar{u}. \end{aligned}$$

Consequently, there exists  $\lim_{\varepsilon \rightarrow 0} u^{\varepsilon} = \underline{u} = \bar{u}$ , and this limit is the solution of (1).

We have justified

**Theorem 4.1.** *Assume the data  $(b_0, f) \in D$  are such that the corresponding weak solution  $u$  of (1) is unique. Then  $u$  is the a.e. limit of  $\varepsilon$ -discretizations of problem (1).*

Let us end this section with an argument that gives insight into the proof of the estimate (12). For the sake of simplicity, we prove the analogous result for weak solutions of problem  $(\mathbf{1}_{\psi,m})$ . It is enough to consider  $m = 1$ .

**Lemma 4.2.** *Let  $u_{\psi}$  be a weak solution of problem  $(\mathbf{1}_{\psi,m})$  with  $m = 1$ . Then for all  $\delta, \Delta > 0$  with  $\delta + \Delta < T$ ,*

$$\int_{\delta}^{T-\Delta} \int_{\Omega} |\psi(u_{\psi})(t+\Delta, x) - \psi(u_{\psi})(t, x)| \leq \frac{1}{\delta} \omega(\Delta) \quad (16)$$

for some modulus of continuity  $\omega(\cdot)$ .

PROOF : Firstly, the following time translation estimate is rather well known:

$$\int_0^{T-\Delta} \int_{\Omega} |b(u_{\psi})(s+\Delta, x) - b(u_{\psi})(s, x)| \leq \omega_0(\Delta), \quad (17)$$

where  $\omega_0(\cdot)$  is some modulus of continuity that depends only on the energy  $\|\Psi^*(b_0)\|_{L^1(\Omega)}$  of the initial datum, on  $\|f\|_{L^2(Q)}$ , and on the modulus of continuity of  $b(\cdot)$ . Indeed, the time translation estimate can be obtained, e.g. in the way of Alt and Luckhaus [2]. One starts with the *a priori* estimates on solutions (at this point, let us denote  $u_{\psi}$  by  $u$ ). Namely,  $B(u)$  is bounded in  $L^{\infty}(0, T; L^1(\Omega))$  and  $u$  is bounded in  $L^2(0, T; H_0^1(\Omega))$ . These bounds are obtained by testing the approximate equations by  $u$ ; let us stress that the integral containing the convective term vanishes. Then, by the equation and the growth bound on  $F$  it follows that  $\partial_t b(u)$  is bounded in  $L^2(0, T; H^{-1}(\Omega))$  and therefore, for any  $\delta > 0$

$$\begin{aligned} &\int_0^{T-\Delta} \int_{\Omega} (b(u(\tau+\Delta)) - b(u(\tau)))(u(\tau+\Delta) - u(\tau)) \, dx \, d\tau \\ &= \int_0^{T-\Delta} \int_{\tau}^{\tau+\Delta} \langle \partial_t b(u)(t), u(\tau+\Delta) - u(\tau) \rangle_{H^{-1}, H_0^1} \, dt \, d\tau \\ &\leq \int_0^{T-\Delta} \|u(\tau+\Delta) - u(\tau)\|_{H_0^1} \int_{\tau}^{\tau+\Delta} \|\partial_t b(u)(t)\|_{H^{-1}} \, dt \, d\tau \\ &\leq \int_0^{T-\Delta} \|u(\tau+\Delta) - u(\tau)\|_{H_0^1} \left( \int_{\tau}^{\tau+\Delta} 1^2 \, dt \right)^{1/2} \left( \int_{\tau}^{\tau+\Delta} \|\partial_t b(u)(t)\|_{H^{-1}}^2 \, dt \right)^{1/2} \, d\tau \\ &\leq 2h^{1/2} \|\partial_t b(u)\|_{L^2(0, T; H^{-1}(\Omega))} \|u\|_{L^2(0, T; H_0^1(\Omega))} \leq Ch^{1/2} \end{aligned}$$

where the constant  $C$  does only depend on  $\|f\|_{L^2(Q)}$  and on  $\|\Psi^*(b_0)\|_{L^1(\Omega)}$ . Then one can estimate

$$\begin{aligned} & \int_0^{T-\Delta} \int_{\Omega} |b(u(t+\Delta)) - b(u(t))| \\ & \leq \int_0^{T-\Delta} \int_{\{ |u(t+\Delta) - u(t)| > h^{1/4} \}} |b(u(t+\Delta)) - b(u(t))| + \omega_b(h^{1/4})T|\Omega| \\ & \leq \frac{1}{h^{1/4}} \int_0^{T-\Delta} \int_Q (b(u(t+\Delta)) - b(u(t)))(u(t+\Delta) - u(t)) + \omega_b(h^{1/4})T|\Omega| \\ & \leq Ch^{1/4} + \omega(h^{1/4})T|\Omega| =: \omega(h). \end{aligned}$$

Notice that another idea for the above estimation is due to Kruzhkov [22] (cf. [4, 10]). Now, restore the notation  $u_\psi$  and denote  $u_\psi^\Delta : (t, x) \mapsto u(t+\Delta, x)$  and  $f^\Delta : (t, x) \mapsto f(t+\Delta, x)$  for  $(t, x) \in (0, T-\Delta) \times \Omega$ . Assume  $s, s+\Delta$  and  $t-\Delta, t$  are Lebesgue points of the  $L^1(0, T; L^1(\Omega))$  map  $b(u_\psi) : t \mapsto b(u_\psi)(t, \cdot)$  (we prefer not to use the  $C([0, T]; L^1(\Omega))$  regularity of  $b(u_\psi)$ , because its justification is relatively involved). Then  $u_\psi, u_\psi^\Delta$  are weak solutions of  $(1_{\psi, m})$  on  $[s, T-\Delta] \times \Omega$  with data  $(b(u_\psi)(s, \cdot), f)$  and  $(b(u_\psi)(s+\Delta, \cdot), f^\Delta)$ , respectively. Using the technique of doubling the time variable due to Otto [23], we deduce the following  $L^1$  contraction property:

$$\begin{aligned} & \int_{\Omega} |b(u_\psi^\Delta) - b(u_\psi)|(t-\Delta) + \int_s^{t-\Delta} \int_{\Omega} |\psi(u_\psi^\Delta) - \psi(u_\psi)| \\ & \leq \int_{\Omega} |b(u_\psi^\Delta)(s+\Delta, \cdot) - b(u_\psi)(s, \cdot)| + \int_0^{t-\Delta} \int_{\Omega} |f^\Delta - f|. \quad (18) \end{aligned}$$

We drop the first term on the left-hand side, then we can let  $t$  converge to  $T$ . Further, we estimate the last term in the right-hand side by  $\omega_f(\Delta)$  where  $\omega_f(\cdot)$  is the  $L^1(Q)$  modulus of continuity of  $f$ , and integrate in  $s$  for  $s \in (\delta, T-\Delta)$  (inequality (18) makes sense for  $s$  in a set of full measure). Using the Fubini theorem to exchange the integrals in  $s$  and in  $t$ , we obtain the inequality

$$\int_{\delta}^{T-\Delta} \int_{\Omega} t |\psi(u_\psi^\Delta) - \psi(u_\psi)|(t, \cdot) dt \leq \int_{\delta}^{T-\Delta} \int_{\Omega} |b(u_\psi)(s+\Delta, x) - b(u_\psi)(s, x)|.$$

Now the claim of the lemma follows by (17).  $\diamond$

In the next section, we study convergence of finite volume approximations of problem (1) with the help of a discrete version of Lemma 4.2.

## 5 Convergence of finite volume approximations

The strategy of the previous proof applies to any numerical scheme that enjoys comparison and  $L^1$  contraction properties. Moreover, it is well known that these properties come hand in hand, thanks to the Crandall-Tartar lemma. To give an example, we analyze in this section a fully implicit finite volume scheme that enjoys the above properties. Similar approach applies to the fully explicit scheme, which is monotone under the Lipschitz continuity assumption on  $F$  and the standard parabolic CFL condition.

### 5.1 A monotone time-implicit finite volume scheme

In the previous section, implicit semi-discretization in time was studied. For a numerical study of convection-diffusion problems, of which (1) is a particular case, time-implicit schemes for the diffusion term are classical. In our case, also the convection term should be discretized as time-implicit. Indeed, the possible degeneracy of  $b$  makes it impossible to guarantee a Courant-Friedrichs-Levy condition on the scheme.

We follow essentially the book [20] of Eymard, Gallouët and Herbin (see also the work [21] of the same authors and Michel); some of the arguments are borrowed from the work [5] of Bendahmane, Karlsen and the first author.

Consider an admissible finite volume mesh in the sense of [20]. In the case  $\Omega$  is a polygonal domain, such a mesh can be constructed if a partition of  $\Omega$  in simplexes satisfying the so-called

Delaunay condition is given (here and below, we refer to [20] for the terminology and notation). For a non-polygonal domain  $\Omega$ , one can use for instance a structured mesh of  $\mathbb{R}^N$  consisting of hypercubes (called volumes) with a sufficiently small edge. Then the volumes  $\kappa$  contained in  $\Omega$  are considered as the control volumes of the scheme (an unknown is associated with each of these volumes), and the volumes intersecting  $\partial\Omega$  are considered as boundary volumes (the homogeneous Dirichlet boundary condition is imposed in these boundary volumes).

In both cases, we have a family  $\mathfrak{T}_h = \{\kappa\}$  of interior volumes, and a time step  $\Delta$  (here  $h$  refers to the discretization size, in particular,  $\Delta \leq h$  and for all  $\kappa \in \mathfrak{T}_h$ ,  $\text{diam}(\kappa) \leq h$ ). The set of boundary volumes will be denoted  $\partial\mathfrak{T}_h$ . The (interior or boundary) volumes  $L$  that have a common face with  $\kappa$  form the set  $\mathcal{N}(\kappa) \subset \mathfrak{T}_h \cup \partial\mathfrak{T}_h$  of the neighbours of  $\kappa$ . For  $\kappa \in \mathfrak{T}_h$  and  $L \in \mathcal{N}(\kappa)$ , the interface between  $\kappa$  and  $L$  is denoted  $\kappa|L$ . The admissibility assumption on the meshes means in particular that each (interior or boundary) volume  $\kappa$  is supplied with a “center”  $x_\kappa$  in such a way that the vector  $x_\kappa x_L$  is aligned with the normal vector  $\nu_{\kappa,L}$  to  $\kappa|L$  pointing from  $\kappa$  to  $L$ . The orthogonality condition allows for approximation of the normal fluxes for the laplacian operator by simple two-point divided differences; for more general operators, more involved schemes for the diffusion part (co-volume, DDFV, etc) can be used; see e.g. [10, 5] and references therein.

Proceeding further with notation, we denote by  $m_\kappa$  the  $N$ -dimensional Lebesgue measure of the volume  $\kappa$ ; the  $(N-1)$ -dimensional measure of  $\kappa|L$  is denoted by  $m_{\kappa|L}$ . The distance between  $x_\kappa$  and  $x_L$  is denoted  $d_{\kappa,L}$ . For every (ordered) couple of neighbours  $\kappa$  and  $L$ , we consider a *numerical convection flux*  $g_{\kappa,L}$  satisfying the classical assumptions:

$$\left\{ \begin{array}{l} \text{(a) } g_{\kappa,L}(\cdot, b) \text{ is nondecreasing for all } b \in \mathbb{R}, \\ \quad \text{and } g_{\kappa,L}(a, \cdot) \text{ is nonincreasing for all } a \in \mathbb{R}; \\ \text{(b) } g_{\kappa,L}(a, a) = F(a) \cdot \nu_{\kappa,L} \text{ for all } a \in \mathbb{R}; \\ \text{(c) } g_{\kappa,L}(a, b) = -g_{L,\kappa}(b, a) \quad \forall a, b \in \mathbb{R}, \text{ for all } \kappa \in \mathfrak{T}_h, L \in \mathcal{N}(\kappa); \\ \text{(d) the family of fluxes } g_{\kappa,L} \\ \quad \text{is locally uniformly (in } \kappa, L \in \mathfrak{T}_h \text{ and in } h) \text{ continuous.} \end{array} \right. \quad (19)$$

Recall that typically, we require the Lipschitz continuity of  $F$  to get uniqueness of  $u$  solution to (1) needed for the conclusion of the convergence argument. In this case, it is not a restriction to assume that the numerical fluxes  $g_{\kappa,L}$  are locally uniformly Lipschitz continuous.

Many examples of numerical fluxes satisfying (19) are given e.g. in [20].

For the sake of simplicity, consider the case without the source term (i.e.,  $f = 0$ ), and with  $L^\infty$  initial condition  $b_0$  (this permits to avoid the issue of  $L^1$  estimates on the discrete energy  $B(u^h)$  which would involve a control of the growth of  $F$ ; yet such estimates can be obtained with the same techniques as in the continuous case, e.g. under the assumption (2)). We discretize the initial condition by approximating  $b_0$  by a piecewise constant function

$$b_0^h(\cdot) = \sum_{\kappa \in \mathfrak{T}_h} b_\kappa^0 \mathbb{1}_\kappa(\cdot), \quad (20)$$

where  $(b_\kappa^0)_\kappa \subset \mathbb{R}$  is chosen in such a way that  $b_0^h, \Phi^* \circ b_0^h$  converge in  $L^1(\Omega)$  to  $b_0$  and  $\Phi^* \circ b_0$ , respectively. For the purpose of theoretical approximation, one can take e.g.  $b_\kappa^0 := \frac{1}{m_\kappa} \int_\kappa b_0$ ; if  $b_0$  is continuous, the standard choice is  $b_\kappa^0 := b_0(x_\kappa)$ .

With the above notation, the finite volume scheme for problem (1) writes

$$\left\{ \begin{array}{l} \frac{m_\kappa}{\Delta t} (b(u_\kappa^n) - b(u_\kappa^{n-1})) - \sum_{L \in \mathcal{N}(\kappa)} m_{\kappa|L} \left( \frac{u_L^n - u_\kappa^n}{d_{\kappa,L}} - g_{\kappa,L}(u_\kappa^n, u_L^n) \right) + m_\kappa \psi(u_\kappa^n) = 0 \\ \quad \forall \kappa \in \mathfrak{T}_h \quad n = 1, \dots, N_h, \\ \text{with } u_\kappa^n = 0 \text{ for all } \kappa \in \partial\mathfrak{T}_h \text{ and } n = 1, \dots, N_h, \\ \text{and with } b(u_\kappa^0) := b_\kappa^0, \kappa \in \mathfrak{T}_h, \text{ according to (20),} \end{array} \right. \quad (21)$$

where  $N_h := T/\Delta t$  (for the sake of simplicity, assume  $T$  be a multiple of  $\Delta t$ ).

Given a solution  $(u_\kappa^n)_{\kappa \in \mathfrak{T}_h, n=1..N_h}$ , we consider the function  $u^h$  on  $Q$  given by

$$u^h := \sum_{n=1}^{N_h} u_\kappa^n \mathbb{1}_{((n-1)\Delta t, n\Delta t]}(t) \mathbb{1}_\kappa(x)$$

as the discrete solution.

Let us state the main result of this section; to make the arguments clear, we restrict the choice of the data (yet the case of general data  $(b_0, f) \in D$  is analogous).

**Theorem 5.1.**

- (i) *There exists a solution  $u_h$  to the scheme (21) with  $b_0 \in L^\infty(\Omega)$  and  $f = 0$ .*
- (ii) *Take  $b_0 \in L^\infty(\Omega)$  and  $f = 0$ . Under the assumption that the corresponding weak solution  $u$  of (1) is unique,  $u^h$  converge to  $u$  a.e. on  $Q$ .*

The point (ii) will be proved in Section 5.4.

Further, as soon as *a priori* estimates are established, the proof of (i) uses standard arguments (topological degree or Brouwer fixed-point theorem), see e.g. [20]; the existence proofs in [21] and in [5] are very close to our situation. The estimate we use here is the  $L^\infty$  bound

$$\text{for all } \kappa \in \mathfrak{T}_h, \text{ for all } n \leq N_h, |u_\kappa^n| \leq \max\{[b^{-1}]^0(b_\kappa^0) \mid \kappa \in \mathfrak{T}_h\}$$

( $[b^{-1}]^0$  is the minimal section of the monotone multivalued graph  $b^{-1}$ ). The above  $L^\infty$  bound (see (24) below) is an easy byproduct of the monotonicity of the scheme that we establish now.

## 5.2 Monotonicity of the scheme

Recall that  $\psi$  in (21) is a continuous non-decreasing function. We have

**Proposition 5.2.** *Consider two solutions  $u_h, \hat{u}_h$  of the scheme corresponding to two initial data  $b_0^h, \hat{b}_0^h$ , respectively.*

- (i) *Assume that  $b_\kappa^0 \leq \hat{b}_\kappa^0$  for all  $\kappa \in \mathfrak{T}_h$ . Then  $u^h \leq \hat{u}^h$  on  $Q$ .*
- (ii) *In general, we have for all  $\nu = 0, \dots, N_h$  for all  $N = (\nu+1), \dots, N_h$ ,*

$$\sum_{\kappa} m_{\kappa} \left( |b(u_{\kappa}^N) - b(\hat{u}_{\kappa}^N)| + \sum_{n=\nu+1}^N \Delta t |\psi(u_{\kappa}^n) - \psi(\hat{u}_{\kappa}^n)| \right) \leq \sum_{\kappa} m_{\kappa} |b(u_{\kappa}^{\nu}) - b(\hat{u}_{\kappa}^{\nu})|. \quad (22)$$

PROOF :

- (i) The proof is by induction in  $n$ ; we thus assume that for some  $n \geq 1$ ,  $u_{\kappa}^{n-1} \leq \hat{u}_{\kappa}^{n-1}$  for all  $\kappa \in \mathfrak{T}_h$ . Arguing by contradiction, define

$$\kappa_+ := \{\kappa \in \mathfrak{T}_h \mid u_{\kappa}^n > \hat{u}_{\kappa}^n\} \text{ and } \mathcal{N}(\kappa_+) := \{L \in \mathfrak{T}_h \cup \partial\mathfrak{T}_h \mid L \in \mathcal{N}(\kappa) \text{ for some } \kappa \in \kappa_+\} \setminus \kappa_+.$$

Expressing  $b(u_{\kappa}^n) + \Delta t \psi(u_{\kappa}^n)$  from the scheme equations, summing up in  $\kappa \in \kappa_+$ , using the conservativity of the fluxes, we end up with

$$S := \sum_{\kappa \in \kappa_+} m_{\kappa} (b(u_{\kappa}^n) + \Delta t \psi(u_{\kappa}^n)) = \sum_{\kappa \in \kappa_+} m_{\kappa} b(u_{\kappa}^{n-1}) + \sum_{(K,L) \in \mathcal{E}_+} \Delta t m_{K|L} \left( \frac{u_L^n - u_K^n}{d_{KL}} - g_{K,L}(u_K^n, u_L^n) \right),$$

where  $\mathcal{E}_+ := \{(K, L) \mid K \in \kappa_+, L \in \mathcal{N}(K) \setminus \kappa_+\} \subset \kappa_+ \times \mathcal{N}(\kappa_+)$ . It is then clear (using, in particular, the monotonicity (19)(a) of the convection fluxes  $g_{K,L}$ ) that

$$S = \sum_{\kappa \in \kappa_+} m_{\kappa} b(u_{\kappa}^{n-1}) + D\left((u_{\kappa}^n)_{\kappa \in \kappa_+}, (u_L^n)_{L \in \mathcal{N}(\kappa_+)}\right)$$

where the function  $D$  is strictly decreasing in all the arguments  $u_{\kappa}^n$ ,  $\kappa \in \kappa_+$ , and it is strictly increasing in all the arguments  $u_L^n$ ,  $L \in \mathcal{N}(\kappa_+)$ . By the definition of  $\kappa_+$  and since we have  $\mathcal{N}(\kappa_+) \subset \{L \in \mathfrak{T}_h \mid u_L^n \geq \hat{u}_L^n\}$  (in particular for  $L \in \partial\mathfrak{T}_h$ , we have  $u_L^n = 0 = \hat{u}_L^n$ ), we infer

$$S < \sum_{\kappa \in \kappa_+} m_{\kappa} b(\hat{u}_{\kappa}^{n-1}) + D\left((\hat{u}_{\kappa}^n)_{\kappa \in \kappa_+}, (\hat{u}_L^n)_{L \in \mathcal{N}(\kappa_+)}\right) =: \hat{S}.$$

Yet with the same calculation as for  $u$ , the scheme for  $\hat{u}$  yields

$$\hat{S} = \sum_{\kappa \in \kappa_+} m_{\kappa} (b(\hat{u}_{\kappa}^n) + \Delta t \psi(\hat{u}_{\kappa}^n)) \leq \sum_{\kappa \in \kappa_+} m_{\kappa} (b(u_{\kappa}^n) + \Delta t \psi(u_{\kappa}^n)) = S,$$

which is contradictory. This contradiction proves the claim (i).

(ii) Consider the map  $T : b_0^h \in L^1(\Omega) \mapsto \sum_{\kappa} \left( b(u_{\kappa}^1) + \Delta t \psi(u_{\kappa}^1) \right) \mathbb{1}_{\kappa}$ . By (i),  $T$  is an order-preserving map from the subset  $C$  of piecewise constant functions on  $L^1(\Omega)$  to  $L^1(\Omega)$ . It is therefore contractive, thanks to the Crandall-Tartar lemma [18]. Indeed, although the map  $T$  does not satisfy the  $L^1$  conservativity condition  $\int_{\Omega} T(b_0) = \int_{\Omega} b_0$ , it is easily seen (from the conservativity of the fluxes, the definition of the scheme (21) for volumes  $\kappa$  adjacent to the boundary, from the monotonicity (19)(a) of the convection fluxes, and from the above claim (i)) that the map  $T$  verifies

$$\hat{b}_0 \geq b_0 \implies \int (T(\hat{b}_0) - T(b_0)) \leq \int (\hat{b}_0 - b_0).$$

Then the proof of the Crandall-Tartar lemma still applies. Alternatively, we can re-define

$$\tilde{T} : \begin{array}{l} C \times \{0\} \\ (b_0, 0) \end{array} \mapsto \begin{array}{l} L^1(\Omega) \times L^1(\partial\Omega) \\ \left( T(b_0), \frac{1}{\text{meas}(\partial\Omega)} \int_{\Omega} (b_0 - T(b_0)) \right), \end{array}$$

in which case both the conservativity and the monotonicity hold, so that we can apply the original result of [18]. Both functions  $b$  and  $\psi$  being non-decreasing, we have

$$\begin{aligned} \sum_{\kappa} m_{\kappa} |b_{\kappa}^0 - \hat{b}_{\kappa}^0| &= \int_{\Omega} |b_0^h - \hat{b}_0^h| \geq \int_{\Omega} |T(b_0^h) - T(\hat{b}_0^h)| \\ &= \sum_{\kappa} m_{\kappa} \left| b(u_{\kappa}^1) - b(\hat{u}_{\kappa}^1) + \Delta t (\psi(u_{\kappa}^1) - \psi(\hat{u}_{\kappa}^1)) \right| \\ &= \sum_{\kappa} m_{\kappa} \left( |b(u_{\kappa}^1) - b(\hat{u}_{\kappa}^1)| + \Delta t |\psi(u_{\kappa}^1) - \psi(\hat{u}_{\kappa}^1)| \right). \end{aligned}$$

Now, applying the contraction result recursively at time steps  $n = (\nu+1)..N$ , we get (22).  $\diamond$

**Remark 5.3.** If we replace  $\psi$  in scheme (21) by  $\frac{1}{m}\psi$  and if we denote  $u_{\psi, m}^h$  the corresponding discrete solutions, as a consequence of the above proof we have the following comparison principles analogous to (13):

$$\begin{cases} \text{if } \psi > 0 \text{ on } \mathbb{R}, & \text{then } (u_{\psi, m}^h)_m \text{ is non-decreasing;} \\ \text{if } \psi < 0 \text{ on } \mathbb{R}, & \text{then } (u_{\psi, m}^h)_m \text{ is non-increasing;} \\ \text{if } \psi_- < 0 < \psi_+ \text{ on } \mathbb{R}, & \text{then } u_{\psi_-, m}^h \geq u_{\psi_+, m}^h \text{ a.e. on } Q. \end{cases} \quad (23)$$

### 5.3 Translation estimates

Before we continue, we need *a priori* estimates on the scheme, which are analogous to the estimates of the continuous problem (1) (recall that we have made the simplifying assumption  $b_0 \in L^{\infty}(\Omega)$ ).

**Proposition 5.4.** If  $u^h$  is a discrete solution of scheme (21), then

$$\|u^h\|_{L^{\infty}(Q)} \leq \| [b^{-1}]^0(b_0) \|_{L^{\infty}(Q)}, \quad (24)$$

$$\sum_{n=1}^{N_h} \Delta t \sum_{\kappa} \sum_{L \in \mathcal{N}(\kappa)} m_{\kappa|L} d_{KL} \left| \frac{u_L - u_{\kappa}}{d_{KL}} \right|^2 \leq C \quad (25)$$

with some  $C$  independent of  $h$ ; in particular, extending  $u^h$  by zero on  $(0, T) \times (\mathbb{R}^N \setminus \Omega)$ , we have

$$\forall \Delta > 0 \quad \int_0^T \int_{\Omega} |b(u^h)(t, x + \Delta) - b(u^h)(t, x)| \leq \omega(\Delta) \quad (26)$$

for some modulus of continuity  $\omega$ , uniform in  $h$ .

PROOF (SKETCHED): The claims of Proposition 5.4 are standard.

First, (24) comes from a slight modification of Proposition 5.2(i). Consider the constant function  $\hat{u}^h \equiv + \| [b^{-1}]^0(b_0) \|_{\infty}$ ; this discrete function is a *super-solution* of the scheme (with

any of the customary discretizations  $b_0^h$  of  $b_0$ ), and the comparison principle still holds giving  $u^h \leq \|[b^{-1}]^0(b_0)\|_\infty$ . Similarly, we get  $-\|[b^{-1}]^0(b_0)\|_\infty \leq u^h$ .

Next, one multiplies the  $n, \kappa$ 'th equation of the scheme (21) by  $\Delta t u_\kappa^n$ , sums up in  $n$  and  $\kappa$ , and proceeds "gathering by edges" (this is the discrete analogue of the integration-by-parts procedure). Using the above  $L^\infty$  bound and the local uniform continuity assumption on the convection fluxes, we easily control the contribution of the discrete convection term and end up with estimate (25).

Finally, from the Cauchy-Schwarz inequality and the standard technique for estimating  $L^1$  translates of discrete functions via the  $L^1$  norm of the discrete gradient (see [20, 21, 10, 5]) we get the estimate

$$\int_0^T \int_\Omega |u^h(t, x + \Delta) - u^h(t, x)| \leq C\Delta,$$

for some  $C$  independent of  $h$ . Then using the Jensen inequality with a concave modulus of continuity of  $b$ , we end up with (26).  $\diamond$

Now, we prove the discrete analogue of Lemma 4.2 (to deduce  $L^1_{loc}$  compactness with the diagonal procedure, it is enough to treat the case where the time step  $\Delta t$  is small enough):

**Lemma 5.5.** *Let  $\delta > 0$ ; let  $u^h$  be a solution to the scheme (21) with size  $h \leq \delta$ . Then*

$$\forall \Delta < \delta \quad \int_\delta^{T-\Delta} \int_\Omega |\psi(u^h)(t+\Delta, x) - \psi(u^h)(t, x)| \leq \frac{1}{\delta} \omega(\Delta) \quad (27)$$

for some modulus of continuity  $\omega(\cdot)$ .

PROOF : First, having assumed that  $\delta \geq \Delta t$ , we can assert that  $\delta \in [l\Delta t, (l+1)\Delta t)$  for some  $l \geq 1$ , and  $\frac{1}{l\Delta t} = \frac{l+1}{l} \frac{1}{(l+1)\Delta t} \leq \frac{2}{\delta}$ . Thus changing  $\omega(\cdot)$  into  $2\omega(\cdot)$ , it is enough to prove the estimate

$$J(\Delta) := \int_{l\Delta t}^{T-\Delta} \int_\Omega |\psi(u^h)(t+\Delta, x) - \psi(u^h)(t, x)| \leq \frac{1}{l\Delta t} \omega(\Delta). \quad (28)$$

Further, it is enough to prove (28) for  $\Delta = r\Delta t$ ,  $r = 0, \dots, N_h/2$ . Indeed, let  $0 < \Delta < \delta < T/2$ ; we have  $\Delta/\Delta t = (k-1) + \alpha$  for some  $k \in \{1, \dots, N_h/2\}$  and  $\alpha \in [0, 1)$ . Since  $u^h$  is piecewise constant in  $t$  with step  $\Delta t$ , we have

$$J(\Delta) = J((k-1)\Delta t + \alpha\Delta t) \leq \alpha J(k\Delta t) + (1-\alpha)J((k-1)\Delta t).$$

Assuming (28) for  $\Delta = k\Delta t$  and  $\Delta = (k-1)\Delta t$ , we have from the concavity of  $\omega$  (recall that a modulus of continuity can always be assumed to be concave):

$$J(\Delta) \leq \frac{1}{l\Delta t} \left( \alpha \omega(k\Delta t) + (1-\alpha) \omega((k-1)\Delta t) \right) \leq \frac{1}{l\Delta t} \omega((k-1) + \alpha)\Delta t = \frac{1}{l\Delta t} \omega(\Delta).$$

Now we prove (28) for  $\Delta = r\Delta t$ ,  $r = 0, \dots, N_h/2$ . Set  $\hat{u}^h(t, x) := u^h(t + r\Delta t, x)$ , which amounts to  $\hat{u}_K^n = u_K^{n+r}$ , for all  $K \in \mathfrak{T}_h$  and  $n = 0, \dots, (N_h - r)$ . Applying the contraction inequality (22) to the solutions  $u^h, \hat{u}^h$  between time steps  $\nu$  and  $N = N_h - r$ , we get in particular

$$\sum_{n=\nu+1}^{N_h-r} \Delta t \sum_K m_K |\psi(u_K^n) - \psi(\hat{u}_K^n)| \leq \sum_K m_K |b(u_K^\nu) - b(\hat{u}_K^\nu)|. \quad (29)$$

Then we multiply (29) by  $\Delta t$  and sum up in  $\nu = 0, \dots, (N_h - r - 1)$ :

$$\sum_{\nu=0}^{N_h-r-1} \Delta t \sum_{n=\nu+1}^{N_h-r} \Delta t \sum_K m_K |\psi(u_K^n) - \psi(u_K^{n+r})| \leq \sum_{\nu=0}^{N_h-r-1} \Delta t \sum_K m_K |b(u_K^\nu) - b(u_K^{\nu+r})|. \quad (30)$$

The right-hand side  $R$  of (30) can be estimated, because the discrete time derivative of  $b(u^h)$  is controlled by the equations of the scheme (21). In practice, this can be done via the discrete Kruzhkov lemma (see e.g. [6]). Another way to proceed is the standard multiplication technique which goes back to [2] (see e.g. [20, 21, 5]); the function  $b$  being merely continuous, we have to work with the modulus of continuity of  $b$ , e.g. in the way of [5, p.10 and pp.37-39]. In either way,

we find  $R \leq \omega(r\Delta t)$  for some modulus of continuity  $\omega$ . As to the left-hand side  $L$  of inequality (30), exchanging the order of summation we find

$$L = \sum_{n=1}^{N_h-r} \Delta t (n\Delta t) \sum_{\mathcal{K}} m_{\mathcal{K}} |\psi(u_{\mathcal{K}}^n) - \psi(u_{\mathcal{K}}^{n+r})| \leq \omega(r\Delta t).$$

In particular, for all  $l = 1..(N_h - r)$ , dropping the terms corresponding to  $n < l$  we have  $n\Delta t \geq l\Delta t$  and thus

$$\int_{l\Delta t}^{T-r\Delta t} \int_{\Omega} |\psi(t, x) - \psi(t + r\Delta t, x)| = \sum_{n=l}^{N_h-r} \Delta t \sum_{\mathcal{K}} m_{\mathcal{K}} |\psi(u_{\mathcal{K}}^n) - \psi(u_{\mathcal{K}}^{n+r})| \leq \frac{1}{l\Delta t} \omega(r\Delta t).$$

This establishes the desired inequality and concludes the proof.  $\diamond$

## 5.4 Convergence of the scheme

We are now in a position to prove Theorem 5.1(ii).

As in Section 4, we fix  $\psi_{\pm}$  given by (14). For every given  $m > 0$  and  $\psi := \frac{1}{m}\psi_-$  (resp.,  $\psi := \frac{1}{m}\psi_+$ ) we denote the associated discrete solution of (21) by  $u_{\psi_-,m}^h$  (resp.,  $u_{\psi_+,m}^h$ ). Then the translation estimates (26) and (27) (combined with the fact that  $\psi_{\pm}^{-1}$  are continuous) and the diagonal extraction argument permit to get the convergences  $u_{\psi_{\pm},m}^h \rightarrow u_{\psi_{\pm},m}$  (as usual, all convergences are for a subsequence; at the final step, the uniqueness of  $u$  solution to (1) is used to suppress the subsequence extractions).

Standard arguments allow to pass to the limit in the weak formulation of the scheme (for instance, we have to combine the consistency (19)(b) and the discrete  $H_0^1$  estimate (25) to ensure convergence of the convection fluxes  $g(u_{\mathcal{K}}, u_L)$  in the “diamond” containing  $\kappa L$  to  $F(u)$ . We refer to [20] for details.

Thus we infer that  $u_{\psi_{\pm},m}$  are weak solutions of the associated problems  $(1_{\psi,m})$ . Starting from this point, the proof follows the lines of the proof of Theorem 4.1, with (13) replaced by (23), and with the discretization parameter  $\varepsilon$  replaced by  $h$ .

**Acknowledgement** The work on this paper was supported by the Procope project No.21806TL. The first author thanks Technische Universität Berlin, where a part of this work was carried on, for the hospitality.

## References

- [1] N. Alibaud, B. Andreianov and M. Bendahmane. Renormalized solutions of the fractional Laplace equation. *C. R. Acad. Sci., Paris, Sér. A* 348(2010), no.13-14, pp.759–762.
- [2] H.W. Alt and S. Luckhaus. Quasilinear elliptic-parabolic differential equations. *Mat. Z.* V.183(1983), pp. 311–341.
- [3] K. Ammar and P. Wittbold. Existence of renormalized solutions of degenerate elliptic-parabolic problems. *Proc. Royal Soc. Edinburgh*, 133A(2003), no.3, pp.477–496.
- [4] B. Andreianov. *Quelques problèmes de la théorie des systèmes paraboliques dégénérés non-linéaires et des lois de conservation*. PhD thesis, Univ. de Franche-Comté, 2000.
- [5] B. Andreianov, M. Bendahmane and K. H. Karlsen. Finite volume schemes for doubly nonlinear degenerate parabolic equations, *J. Hyperbolic Differ. Equ.* 7(2010), no.1, pp.1–67.
- [6] B. Andreianov, M. Bendahmane and R. Ruiz Baier. Analysis of a finite volume method to solve a cross-diffusion population system. *M3AS Math. Models and Meth. in Applied Sci.* (2011), to appear.
- [7] B. Andreianov and N. Igbida. Revising Uniqueness for a Nonlinear Diffusion-Convection Equation. *J. Differ. Equ.*, 227(2006), no.1, pp.69–79.
- [8] B. Andreianov and N. Igbida. Uniqueness for the Dirichlet elliptic-parabolic problem. *Proc. Royal Soc. Edinburgh*, 137A(2007), pp.1119–1133.



- [9] B. Andreianov and N. Igbida. On uniqueness techniques for degenerate convection-diffusion problems. *Int. J. Dynamical Syst. Differ. Equ.*, accepted; available at <http://hal.archives-ouvertes.fr/hal-00553819>.
- [10] 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(2004), no.1, pp.228–251.
- [11] Ph. Bénilan, M.G. Crandall and A. Pazy. *Nonlinear evolution equations in Banach spaces*. Preprint book.
- [12] Ph. Benilan and P. Wittbold. On mild and weak solutions of elliptic-parabolic problems. *Adv. Differential Equations*, 1(1996), no.6, pp.1053–1073.
- [13] Ph. Bénilan and P. Wittbold. Sur un problème parabolique-elliptique. *M2AN Math. Modelling and Num. Anal.* 33(1999), no.1, pp. 121-127.
- [14] L. Boccardo, Th. Gallouët and F. Murat. Unicité de la solution de certaines équations elliptiques non linéaires. *C. R. Acad. Sci. Paris Sér. I Math.* 315 (1992), no. 11, 1159–1164.
- [15] H. Brézis, D. Kinderlehrer, and G. Stampacchia. Sur une nouvelle formulation du problème de l'écoulement à travers une digue. (French) *C. R. Acad. Sci., Paris, Sér. A* 287(1978), pp.711–714.
- [16] J. Carrillo. Entropy solutions for nonlinear degenerate problems. *Arch. Rational Mech. Anal.*, 147(1999), pp. 269–361.
- [17] J. Carrillo and P. Wittbold. Uniqueness of renormalized solutions of degenerate elliptic-parabolic problems. *J. Differ. Equ.* 156(1):93–121, 1999.
- [18] M.G. Crandall and L. Tartar. Some relations between nonexpansive and order preserving mappings. *Proc. AMS* 78(1980), no.3, pp. 385–390.
- [19] L.C. Evans and R.F. Gariepy. *Measure theory and fine properties of functions*. Studies in Advanced Mathematics. CRC Press, 1992.
- [20] R. Eymard, T. Gallouët and R. Herbin. Finite Volume Methods. *Handbook of Numerical Analysis*, Vol. VII, P. Ciarlet, J.-L. Lions, eds., North-Holland, 2000.
- [21] R. Eymard, T. Gallouët, R. Herbin and A. Michel. Convergence of a finite volume scheme for nonlinear degenerate parabolic equations. *Numer. Math.* 92(1):41–82, 2002.
- [22] S.N. Kruzhkov. Results on the nature of the continuity of solutions of parabolic equations and some of their applications. *Mat. Zametki* 6(1):97–108, 1969; english tr. in *Math. Notes* 6(1):517-523, 1969.
- [23] F. Otto.  $L^1$  contraction and uniqueness for quasilinear elliptic-parabolic equations. *J. Differ. Equ.*, 131(1996), pp. 20–38.
- [24] A. Zimmermann. *Renormalized solutions for nonlinear partial differential equations with variable exponents and  $L^1$ -data*. PhD thesis, Technische Univ. Berlin, 2010.