



**HAL**  
open science

# Volumes finis et solutions renormalisées, applications à des systèmes couplés.

Sarah Leclavier

► **To cite this version:**

Sarah Leclavier. Volumes finis et solutions renormalisées, applications à des systèmes couplés.. Equations aux dérivées partielles [math.AP]. Normandie Université, 2017. Français. NNT : 2017NORMR029 . tel-01665136

**HAL Id: tel-01665136**

**<https://theses.hal.science/tel-01665136>**

Submitted on 15 Dec 2017

**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.



Normandie Université

## THESE

**Pour obtenir le diplôme de doctorat**

**Spécialité Mathématiques (4200001)**

**Préparée au sein de l'université de Rouen-Normandie**

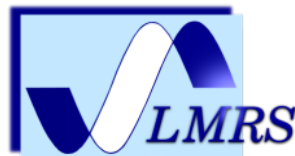
**Volumes finis et solutions renormalisées,  
application à des systèmes couplés.**

**Présentée et soutenue par  
Sarah LECLAVIER**

**Thèse soutenue publiquement le 12/12/2017  
devant le jury composé de**

M. Boris ANDREIANOV	Professeur, Université de Tours	Rapporteur
M. Thierry GALLOUET	Professeur, Université de Marseille	Rapporteur
Mme Patrizia DONATO	Professeur, Université de Rouen-Normandie	Examineur
M. Nicolas FORCADEL	Professeur, INSA de Rouen	Examineur
Mme Claire CHAINAIS-HILLAIRET	Professeur, Université de Lille	Examineur
M. François MURAT	Directeur de recherche émérite, CNRS	Examineur
M. Olivier GUIBE	Maître de conférences, Université de Rouen-Normandie	Directeur de thèse

**Thèse dirigée par Olivier GUIBE, laboratoire LMRS, université de Rouen-Normandie**



---

# Remerciements

Mes premiers remerciements vont à Olivier Guibé qui a dirigé ma thèse et avec qui ce fut un réel plaisir de travailler. Ses connaissances scientifiques, ses qualités pédagogiques et sa capacité d'écoute en font un directeur remarquable. Sans oublier le doux bruit de son moulin à café sans lequel le laboratoire serait resté bien silencieux certains matins.

Je voudrais également exprimer toute ma reconnaissance à Boris Andreianov et Thierry Gallouët pour avoir accepté d'être les rapporteurs de ma thèse. Leurs travaux ont grandement participé à l'élaboration de cette thèse.

Je souhaiterais ensuite remercier Claire Chainais-Hillairet, Patrizia Donato, Nicolas Forcadel et François Murat qui ont accepté de faire partie du jury en cette fin d'année chargée en soutenances.

A côté du travail de thèse, il a aussi fallu gérer quelques tâches administratives. Pour cela je tiens à remercier Edwige, Hamed, Marc, Pierre et Sandrine sur qui nous pouvons toujours compter. Bien des personnes seraient perdues sans eux au laboratoire.

Mes remerciements vont aussi à Hélène sans qui ma première année au LMRS n'aurait pas été la même et à toute la bande des Rouennais pour tous ces moments qui m'ont permis de garder un esprit sain. Je remercie également Déborah et Sahar pour leur présence et leur bonne humeur quotidienne.

Enfin, je souhaiterais remercier tous mes proches, spécialement mes parents et ma soeur qui m'ont toujours encouragée et soutenue tout au long de ces huit années d'études. Sans eux je ne me serais sûrement pas lancée dans ce doctorat.

Ma dernière pensée va à Paul qui, lui-même doctorant, a su mettre de côté ses propres doutes pour me remonter le moral et me motiver lors de mes passages à vide. Je lui dois beaucoup.



# Résumé

On s'intéresse dans cette thèse à montrer que la solution approchée, par la méthode des volumes finis, converge vers la solution renormalisée de problèmes elliptiques ou paraboliques à donnée  $L^1$ .

Dans la première partie nous étudions une équation de convection-diffusion elliptique à donnée  $L^1$ . En adaptant la stratégie développée pour les solutions renormalisées à la méthode des volumes finis, nous montrons que la solution approchée converge vers l'unique solution renormalisée.

Dans la deuxième partie nous nous intéressons à un problème parabolique non-linéaire à donnée  $L^1$ . En utilisant une version discrète de résultats de compacité classiques, nous montrons que les résultats obtenus dans le cas elliptique restent vrais dans le cas parabolique.

Dans la troisième partie nous montrons des résultats similaires pour une équation parabolique doublement non-linéaire à donnée  $L^1$ . Le caractère doublement non-linéaire de l'équation crée des difficultés supplémentaires par rapport à la partie précédente, notamment car la règle de dérivation en chaîne ne s'applique pas dans le cas discret.

Enfin, dans la quatrième partie, nous utilisons les résultats établis précédemment pour étudier un système de type thermoviscoélasticité. Nous montrons que la solution approchée, obtenue par un schéma éléments finis-volumes finis, converge vers une solution faible-renormalisée du système.

**Mots-clés** : Problème Elliptique, Problème Parabolique, Donnée  $L^1$ , Solutions Renormalisées, Volumes Finis, Systèmes Couplés

# Abstract

In this thesis we are interested in proving that the approximate solution, obtained by the finite volume method, converges to the unique renormalized solution of elliptic and parabolic equations with  $L^1$  data.

In the first part we study an elliptic convection-diffusion equation with  $L^1$  data. Mixing the strategy developed for renormalized solution and the finite volume method, we prove that the approximate solution converges to the unique renormalized solution.

In the second part we investigate a nonlinear parabolic equation with  $L^1$  data. Using a discrete version of classical compactness results, we show that the results obtained previously in the elliptic case hold true in the parabolic case.

In the third part we prove similar results for a doubly nonlinear parabolic equation with  $L^1$  data. The doubly nonlinear character of the equation makes new difficulties with respect to the previous part, especially since the chain rule formula does not apply in the discrete case.

Finally, in the fourth part we use the results established previously to investigate a system of thermoviscoelasticity kind. We show that the approximate solution, obtained by finite element-finite volume scheme, converges to a weak-renormalized solution of the system.

**Keywords** : Elliptic Problem, Parabolic Problem,  $L^1$  Data, Renormalized Solutions, Finite Volume, Coupled Systems

# Table des matières

Remerciements . . . . .	iii
Résumé . . . . .	v
Abstract . . . . .	vi
Table des matières . . . . .	vii
<b>Introduction</b>	<b>1</b>
<b>1 Finite volume scheme and renormalized solutions for a noncoercive elliptic problem with <math>L^1</math> data *</b>	<b>9</b>
1.1 Introduction . . . . .	9
1.2 The finite volume scheme . . . . .	11
1.3 Estimates . . . . .	14
1.4 Convergence analysis . . . . .	22
1.5 Concluding remark . . . . .	30
<b>2 Finite volume scheme and renormalized solutions for a noncoercive and nonlinear parabolic problem with <math>L^1</math> data</b>	<b>33</b>
2.1 Introduction . . . . .	33
2.2 The finite volume scheme and the time discretization . . . . .	35
2.3 Estimates . . . . .	40
2.4 Convergence analysis . . . . .	55
<b>3 Finite volume scheme and renormalized solutions for a doubly nonlinear parabolic equation with <math>L^1</math> data</b>	<b>69</b>
3.1 Introduction . . . . .	69
3.2 The finite volume scheme and the time discretization . . . . .	71
3.3 Estimates . . . . .	74
3.4 Convergence analysis . . . . .	84
<b>4 Convergence of a scheme for a coupled PDE system of thermoviscoelasticity type</b>	<b>97</b>
4.1 Introduction . . . . .	97
4.2 Discretization . . . . .	99
4.3 Estimates . . . . .	104
4.4 Convergence analysis . . . . .	115

---

\*. Publié dans le journal "Computational Methods for Applied Mathematics" (voir [53])



<b>5 Quelques perspectives</b>	<b>121</b>
5.1 Perspectives sur les solutions renormalisées et les volumes finis . . . .	121
5.2 Perspectives sur les systèmes . . . . .	122
<b>Bibliographie</b>	<b>125</b>

# Introduction

Dans cette thèse nous étudions des schémas aux volumes finis de type "cell-centered" pour des problèmes elliptiques ou paraboliques à donnée  $L^1$ . Notre but est de montrer que la solution approchée, obtenue par la méthode des volumes finis, converge vers l'unique solution renormalisée du problème. Une application, la discrétisation d'un système de deux EDP couplées de type thermoviscoélasticité, est aussi abordée.

Rappelons tout d'abord quelques éléments sur les notions de solution de l'équation

$$\begin{aligned} -\Delta_p u &= f \text{ dans } \Omega, \\ u &= 0 \text{ sur } \partial\Omega, \end{aligned} \tag{0.1}$$

où  $p > 1$  et  $\Omega$  est un ouvert borné de  $\mathbb{R}^d$ .

Si  $f \in L^{p'}(\Omega)$  ou plus généralement  $f \in W^{-1,p'}(\Omega)$  la notion de solution variationnelle

$$\begin{cases} u \in W_0^{1,p}(\Omega), \\ \int_{\Omega} |Du|^{p-1} Du \cdot Dv \, dx = \int_{\Omega} f v \, dx \quad \forall v \in W_0^{1,p}(\Omega), \end{cases} \tag{0.2}$$

est adaptée dans le sens où il y a existence et unicité. Pour des opérateurs non-linéaires plus généraux, dits de Leray-Lions (voir [54]), nous avons aussi des résultats d'existence. c'est le cadre variationnel usuel.

Si  $f \in L^1(\Omega)$  ou  $f$  est une mesure de Radon à variations bornées, nous ne pouvons pas en général obtenir l'existence d'une solution variationnelle,  $\int_{\Omega} f v \, dx$ , ou  $\langle f, v \rangle$ , n'a aucune raison d'être bien défini et il n'est pas possible d'espérer avoir  $u \in W_0^{1,p}(\Omega)$  pour une donnée peu régulière quand  $p < d$ .

Quand  $f \in L^1(\Omega)$  (ou  $f$  mesure de Radon à variations bornées) et  $p > 2 - 1/d$ , Boccardo et Gallouët ont démontré dans [23] (voir aussi [24]) l'existence d'une solution au sens des distributions de (0.1), à savoir

$$\begin{cases} u \in W_0^{1,q}(\Omega) \quad \forall q < d(p-1)/(d-1), \\ \int_{\Omega} |Du|^{p-1} Du \cdot D\varphi \, dx = \int_{\Omega} f \varphi \, dx \quad \forall \varphi \in \cup_{q>d} W_0^{1,q}(\Omega). \end{cases} \tag{0.3}$$

Ces résultats s'étendent pour des opérateurs plus généraux ([12, 23, 24]) et au cas parabolique ([22, 23]). Cependant même dans le cas linéaire la solution au sens des distributions n'est pas unique (voir le contre-exemple de Serrin [68] et Prignet [65]) essentiellement par manque de régularité de la solution et donc à cause d'un espace de fonctions test plus réduit.

Pour palier ce manque d'unicité, plusieurs notions de solutions ont été développées, la notion de solution par transposition dans le cas linéaire (voir [70], pour des

équations de convection-diffusion voir [37] et pour une extension partielle au cas non-linéaire voir [25, 60]). Dans les années 90, pour  $f \in L^1(\Omega)$  et des opérateurs non-linéaires à croissance  $p > 1$  les notions de SOLA (solutions obtenues comme limites d'approximation), solutions entropiques et solutions renormalisées ont été développées. La notion de solution renormalisée a été généralisée pour une donnée mesure dans [33]. Rappelons que dans le cas d'une donnée  $L^1$  ces notions sont équivalentes (voir [33] pour le cas elliptique). Nous nous intéressons dans cette thèse uniquement à la notion de solution renormalisée.

La théorie des solutions renormalisées a été introduite en 1989 par DiPerna et Lions dans le cadre des équations de Boltzmann (voir [36]). Cette théorie a ensuite été adaptée aux problèmes elliptiques à donnée  $L^1$  par Lions et Murat dans [58, 59]. Une généralisation pour des problèmes elliptiques à donnée mesure a été faite dans [33]. Concernant les problèmes d'évolution à donnée  $L^1$  et les solutions renormalisées un des premiers travaux est celui de [17] qui s'appuie sur [14]. Par la suite de nombreux articles traitent d'équations paraboliques à donnée  $L^1$  ou mesure et de solutions renormalisées, notamment [16, 18, 20, 21, 35] pour une donnée  $L^1$  et [19, 63, 64] pour une donnée mesure.

Dans le cadre de problème à donnée peu régulière, la notion de solution renormalisée est plus avantageuse que celle de solution faible puisqu'elle nous donne en général existence, stabilité et parfois unicité de la solution. Le manque de régularité de la solution, dû à la donnée peu régulière, crée une difficulté pour montrer l'unicité de la solution faible (voir [66, 68]) et cette difficulté est contournée par la technique des solutions renormalisées. Prenons l'exemple du problème

$$\begin{aligned} -\Delta u &= f \text{ dans } \Omega, \\ u &= 0 \text{ sur } \partial\Omega, \end{aligned} \tag{0.4}$$

où  $\Omega$  est un ouvert borné de  $\mathbb{R}^d$  et  $f \in L^1(\Omega)$ . La définition de solution renormalisée pour ce problème est la suivante,  $u : \Omega \rightarrow \mathbb{R}$  mesurable finie presque partout est une solution renormalisée de (0.4) si,

$$\forall k > 0, T_k(u) \in H_0^1(\Omega), \tag{0.5}$$

$$\lim_{k \rightarrow +\infty} \frac{1}{k} \int_{\Omega} |\nabla T_k(u)|^2 dx = 0, \tag{0.6}$$

$$\begin{cases} \forall h \in \mathcal{C}_c^1(\mathbb{R}), \forall \psi \in H_0^1(\Omega) \cap L^\infty(\Omega), \\ \int_{\Omega} \nabla u h(u) \nabla \psi dx + \int_{\Omega} \nabla u \psi \nabla u h'(u) dx = \int_{\Omega} \psi h(u) f dx, \end{cases} \tag{0.7}$$

avec  $T_k$  la fonction troncature de hauteur  $k$  :

$$T_k(r) = \min(k, \max(r, -k)).$$

L'existence d'une solution est établie par un résultat de stabilité : si  $f_\varepsilon \rightarrow f$  dans  $L^1$  fort et si  $u_\varepsilon$  est solution renormalisée de  $-\Delta u_\varepsilon = f_\varepsilon$  dans  $\Omega$ ,  $u_\varepsilon = 0$  sur  $\partial\Omega$  alors  $T_k(u_\varepsilon) \rightarrow T_k(u)$  dans  $H_0^1(\Omega)$  pour tout  $k > 0$  où  $u$  est solution renormalisée de (0.4). Si  $f_\varepsilon \in L^2(\Omega)$  il est aisé de vérifier qu'une solution variationnelle est aussi une

solution renormalisée en prenant  $f_\varepsilon \in L^2(\Omega)$  convergeant fort vers  $f$  dans  $L^1(\Omega)$ , on obtient donc l'existence.

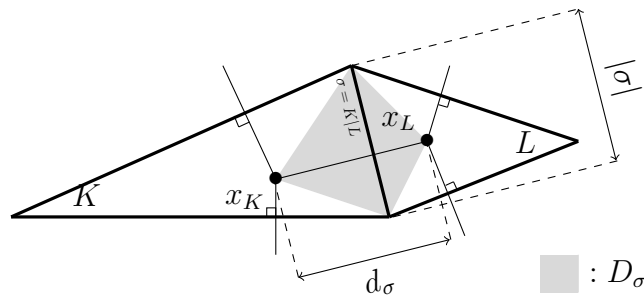
La régularité est imposée non pas sur  $u$  globalement mais sur les tronquées. Si  $u$  est finie presque partout et  $T_k(u) \in H_0^1(\Omega)$  on peut (voir [12]) définir  $\nabla u$  dans le sens où  $\nabla T_k(u) = \mathbb{1}_{\{|u|<k\}} \nabla u$ .

La formulation (0.7) est non linéaire même si le problème (0.4) est linéaire. Comme (0.7) ne "voit" que les zones où  $u$  est bornée tous les termes sont bien définis et la condition (0.6) de décroissance de l'énergie donne une information supplémentaire cruciale pour la stabilité et l'unicité (quand il y a unicité). Quand on peut prouver l'unicité la démonstration se fait, comme la plupart du temps, en prenant deux solutions renormalisées  $u$  et  $v$  mais avec un choix judicieux de troncature de  $u$ , de  $v$  ou de  $u - v$  pour les fonctions test.

Le second outil utilisé tout au long de cette thèse est celui des volumes finis. Il s'agit d'une méthode de discrétisation qui a été conçue pour résoudre numériquement des problèmes faisant intervenir des lois de conservation (mécanique des fluides, transfert de chaleur, ingénierie pétrolière...). Cette méthode est très proche de celle des éléments finis puisque les deux se basent sur des approximations d'intégrales. Tandis que la méthode des éléments finis repose sur une formulation variationnelle de l'équation, la méthode des volumes finis se base quant à elle sur la forme forte de l'équation. Le principe est de mailler le domaine en petits volumes de contrôle  $K$  et d'intégrer l'équation sur chacun de ces volumes. Les volumes de contrôle pouvant être de forme quelconque, un des avantages de cette méthode est de pouvoir traiter des problèmes posés sur des géométries complexes. De plus les termes de divergence présents dans les lois de conservation sont plus simple à traiter puisque les intégrales de volume de ces termes sont transformées en intégrales de surface grâce au théorème de la divergence. Par exemple, regardons l'équation

$$\begin{aligned} -\Delta u + \operatorname{div}(\mathbf{v}u) &= f \text{ dans } \Omega, \\ u &= 0 \text{ sur } \partial\Omega. \end{aligned} \quad (0.8)$$

Le maillage serait de cette forme



avec  $K$  et  $L$  des volumes de contrôle voisins et  $\sigma$  les arrêtes de ces volumes. Le schéma aux volumes finis s'écrit alors

$$\forall K \in \mathcal{T}, \quad \sum_{\sigma \in \mathcal{E}(K)} \frac{|\sigma|}{d_\sigma} (u_K - u_L) + \sum_{\sigma \in \mathcal{E}(K)} |\sigma| \mathbf{v}_{K,\sigma} u_{\sigma,+} = \int_K f \, dx, \quad (0.9)$$

avec  $u_L = 0$  if  $\sigma \in \mathcal{E}_{ext} \cap \mathcal{E}(K)$  et

$$\forall \sigma = K|L \in \mathcal{E}_{int}, u_{\sigma,+} = u_K \text{ si } \mathbf{v}_{K,\sigma} \geq 0, u_{\sigma,+} = u_L \text{ sinon,} \quad (0.10)$$

$$\forall \sigma \in \mathcal{E}_{ext} \cap \mathcal{E}(K), u_{\sigma,+} = u_K \text{ si } \mathbf{v}_{K,\sigma} \geq 0, u_{\sigma,+} = 0 \text{ sinon.} \quad (0.11)$$

Les avantages des volumes finis ont poussé les ingénieurs, notamment pétroliers, à développer cette méthode dans les années 60, celle-ci permettant, contrairement aux différences finies et aux éléments finis, de calculer le déplacement des hydrocarbures en réservoir géologiquement complexe. La recherche des ses fondements mathématiques (formalisation, convergence...) n'a débuté que plus tard.

Le livre "Finite volume methods" [44] écrit par Eymard, Gallouët et Herbin recense les principales techniques de résolution avec les volumes finis pour des problèmes elliptiques, paraboliques, hyperboliques et même pour des systèmes d'équations. Comme dans le cas continu il est nécessaire d'avoir des estimations a priori et de la compacité pour pouvoir passer à la limite. On utilise des versions discrètes des inégalités de Poincaré et des injections de Sobolev. Une des difficultés vient de la reconstruction du gradient. L'approximation par les volumes finis étant constante sur chaque cellule, cette approximation n'appartient pas à un espace de Sobolev. Il faut définir un "gradient au sens des volumes finis" qui ne pourra converger que faiblement.

Dans le cadre de problèmes elliptiques à donnée  $L^1$  ou mesure on peut citer les articles [40, 46] dans lesquels les auteurs prouvent la convergence de la solution approchée par un schéma volumes finis de type "cell-centered" vers la solution au sens des distributions de l'équation. Dans [51] les auteurs étudient un système d'équations dont une est à donnée  $L^1$  et est traitée avec la méthode des volumes finis. Ils montrent que la solution approchée converge vers une solution faible qui dans ce cas n'est pas forcément unique.

En ce qui concerne les problèmes paraboliques à donnée peu régulière, la difficulté dans le cas discret, comme dans le cas continu, est d'avoir des résultats de compacité, notamment la convergence presque partout. Pour faire face à cette difficulté dans le cas continu l'idée est de démontrer des estimations sur la solution ou ses tronquées et d'utiliser un lemme de type Aubin-Simon (voir [69]). On peut adapter cette stratégie dans le cas discret comme cela a été fait dans [47, 48] où les auteurs ont prouvé un équivalent discret du résultat de compacité de [69]. Une autre technique est de trouver des estimations sur les différences des translatées en temps et en espace dans le but de montrer un résultat de compacité (voir [1, 4, 6, 8, 45]).

Dans la plupart des articles dans lesquels sont étudiés des problèmes à donnée  $L^1$  (elliptiques ou paraboliques) avec la méthode des volumes finis, les auteurs montrent que la solution du schéma converge vers une solution faible du problème. Cependant, comme le montre le contre-exemple de Serrin [68], cette solution n'est pas toujours unique. La contribution principale de cette thèse est de combiner la méthode des volumes finis et la notion de solution renormalisée. Nous utiliserons des schémas volumes finis de type "cell-centered" et l'idée est d'obtenir des estimations sur la tronquée en norme discrète, de montrer une version discrète de la décroissance de l'énergie à l'infini (1.4) et de prendre une version discrète de la fonction test  $\psi h(u)$  de (1.5) dans le schéma pour pouvoir passer à la limite.

Adapter au cas discret la stratégie développée pour les solutions renormalisées n'est pas chose aisée puisque la fonction test ne tronque pas toute l'équation contrai-

rement au cas continu. Il faut donc faire apparaître "manuellement" le gradient au sens des tronquées et contrôler des termes résiduels. Dans le cas parabolique d'autres difficultés s'ajoutent car la formule de dérivation en chaîne ne s'applique pas dans le cas discret. Il y a donc de nouveaux termes résiduels à contrôler. Enfin la formulation renormalisée est non linéaire et fait apparaître un gradient au carré; alors que dans le cas continu  $\nabla T_k(u_\varepsilon) \rightarrow \nabla T_k(u)$  fort dans  $L^2$  on ne peut avoir un résultat similaire pour les volumes finis et le "gradient au sens des volumes finis".

Dans le premier chapitre nous étudions l'équation de convection-diffusion

$$\begin{aligned} -\Delta u + \operatorname{div}(\mathbf{v}u) + bu &= f \text{ dans } \Omega, \\ u &= 0 \text{ sur } \partial\Omega, \end{aligned} \tag{0.12}$$

où  $\Omega$  est un ouvert borné de  $\mathbb{R}^d$ ,  $d \geq 2$ ,  $\mathbf{v} \in L^p(\Omega)^d$ ,  $2 < p < +\infty$  si  $d = 2$ ,  $p = d$  si  $d \geq 3$ ,  $b \in L^2(\Omega)$ ,  $b \geq 0$  et  $f \in L^1(\Omega)$ . Dans le cas continu ce type d'équations a été abordé par Droniou à l'aide d'une méthode de dualité ([37]) et par Ben Cheikh et Guibé (dans le cas non-linéaire [10]) avec des résultats d'existence et d'unicité. Pour une donnée mesure et quand  $\mathbf{v} \in (\mathcal{C}(\bar{\Omega}))^d$  la convergence du schéma volumes finis pour (0.12) est abordée dans [40], les auteurs montrent que la solution converge vers la solution faible de (0.12) (qui sera ici unique, la linéarité de l'équation permettant d'utiliser une technique de dualité).

La principale nouveauté de ce chapitre est que nous prouvons la convergence de la solution approchée, obtenue par la méthode des volumes finis, vers l'(unique) solution renormalisée. Comme dans le cas continu les principales difficultés sont de gérer le caractère non coercif de l'équation et la donnée  $L^1$ . En combinant à la fois les techniques développées pour les solutions renormalisées ([33, 59]) et celles de la méthode des volumes finis ([40]) nous établissons des estimations sur les solutions discrètes. Nous montrons en particulier une version discrète de la décroissance à l'infini de l'énergie tronquée

$$\lim_{n \rightarrow \infty} \frac{1}{n} \int_{\Omega} |\nabla T_n(u)|^2 dx = 0,$$

cette estimation étant, comme dans le cas continu, cruciale pour montrer que la limite de la solution approchée est la solution renormalisée. Lors du passage à la limite une autre difficulté apparaît puisqu'en version discrète la fonction test, de type  $h(u)\varphi$  avec  $h$  à support compact, ne permet pas de tronquer toute l'équation. De plus en adaptant les preuves la méthode se généralise aux équations non-linéaires de type

$$\begin{aligned} -\operatorname{div}(\lambda(u)\nabla u) + \operatorname{div}(\mathbf{v}u) + bu &= f \text{ dans } \Omega, \\ u &= 0 \text{ sur } \partial\Omega, \end{aligned} \tag{0.13}$$

où  $\lambda$  est une fonction continue telle que  $\lambda_\infty \geq \lambda(u) \geq \mu > 0$  avec  $\lambda_\infty$  et  $\mu$  deux nombres réels.

Ce travail a été publié dans le journal "Computational Methods in Applied Mathematics" (voir [53]).

Dans le deuxième chapitre, nous nous intéressons au problème

$$\begin{aligned} \frac{\partial u}{\partial t} - \operatorname{div}(\lambda(u)\nabla u) + \operatorname{div}(\mathbf{v}u) &= f \text{ dans } Q, \\ u &= 0 \text{ sur } \partial\Omega \times (0, T), \\ u(x, 0) &= u^0(x) \quad \forall x \in \Omega, \end{aligned} \tag{0.14}$$

où  $\Omega$  est un ouvert borné de  $\mathbb{R}^d$ ,  $d \geq 2$ ,  $T$  est un nombre positif et  $Q = \Omega \times (0, T)$ . De plus  $\lambda$  est une fonction continue telle que  $\lambda(u) \geq \mu > 0$ , avec  $\mu$  un réel positif,  $\mathbf{v} \in (L^{d+2}(Q))^d$ ,  $u^0 \in L^1(\Omega)$  et  $f \in L^1(Q)$ .

Dans le cas continu des résultats d'existence et d'unicité pour l'équation (0.14) avec  $\lambda$  non nécessairement bornée ont été montrés dans [16, 20, 21]. Si  $\lambda$  n'est pas bornée il n'y a aucune raison d'avoir pour  $f$  dans  $L^1$  une solution au sens des distributions. Dans le cadre des volumes finis des équations similaires ont été étudiées dans [47] pour le cas linéaire  $\lambda = 1$  et dans [48] avec une fonction  $\lambda$  continue bornée. Les auteurs y montrent la convergence de la solution approchée vers une solution au sens des distributions.

Dans ce chapitre, en utilisant les outils développés pour les schémas de type volumes finis ([44, 47]), nous adaptons la stratégie utilisée pour montrer l'existence d'une solution renormalisée de problèmes paraboliques à donnée  $L^1$  (voir [16, 20, 21]). Il faut souligner l'avantage de la méthode de [20] par rapport à celle précédemment utilisée dans [18] qui est plus courte et plus simple. Dans [18] une approximation de type Landes [50], de nombreuses étapes (un lemme technique concernant la dualité et l'approximation de Landes) sont nécessaires tandis que [20] utilise les moyennes de Steklov et des inégalités de convexité. Adapter la méthode de [18] au cas des volumes finis paraît hors de portée. Dans ce chapitre nous mélangeons la stratégie de [20] avec la méthode des volumes finis.

Comme dans le cas continu l'obstacle principal pour les équations dépendant du temps est d'obtenir un résultat de compacité pour la solution approchée et de contrôler les dérivées en temps afin de passer à la limite. Pour la compacité nous utilisons la version discrète du lemme d'Aubin-Simon montrée dans [48]. Les autres difficultés sont de gérer les non-coercivité et non-linéarité du problème, la donnée  $L^1$  ainsi que le fait que  $\lambda$  n'est pas supposée bornée. De plus la discrétisation de la dérivée par rapport au temps fait apparaître des termes supplémentaires par rapport au cas elliptique dans le chapitre 1. Il faut donc trouver un moyen de contrôler ces termes résiduels à l'aide de fonctions tests bien choisies.

Dans le troisième chapitre nous étudions l'équation doublement non linéaire

$$\begin{aligned} \frac{\partial b(u)}{\partial t} - \operatorname{div}(\lambda(u)\nabla u) &= f \text{ dans } Q, \\ u &= 0 \text{ sur } \partial\Omega \times (0, T), \\ b(u(x, 0)) &= b(u^0(x)) \quad \forall x \in \Omega, \end{aligned} \tag{0.15}$$

où  $\Omega$  est un ouvert borné de  $\mathbb{R}^d$ ,  $d \geq 2$ ,  $T$  est un nombre positif et  $Q = \Omega \times (0, T)$ ,  $\lambda$  est une fonction continue telle que  $\lambda(u) \geq \mu > 0$ , avec  $\mu$  un réel positif, et  $f \in L^1(Q)$ . De plus  $b$  est une fonction  $\mathcal{C}^1$  strictement croissante telle que  $b(0) = 0$ ,  $b' > 0$  sur  $\mathbb{R}$ . Comme  $\lambda$  n'est pas supposée bornée et de même pour  $b$ , le changement de variable

$v = b(u)$  peut mener à un problème dégénéré et n'est donc pas traité dans le chapitre précédent.

Les équations de ce type ont été étudiées avec les solutions renormalisées entre autres dans [16,20,21] et dans [6,8] avec les volumes finis. Nous prouvons ici la convergence de la solution approchée, obtenue par la méthode des volumes finis, vers l'unique solution renormalisée. A cause du caractère doublement non linéaire de l'équation nous faisons face à une nouvelle difficulté par rapport au chapitre 2. En effet la règle de dérivation en chaîne ne s'applique pas dans le cas discret au terme  $\frac{\partial b(u)}{\partial t}$ ; ceci nous force à contrôler de nouveaux termes additionnels.

La principale nouveauté de ce chapitre est de traiter avec les volumes finis et les solutions renormalisées une équation parabolique à la fois doublement non-linéaire et à donnée  $L^1$ .

Dans le quatrième chapitre nous étudions à l'aide des éléments finis et des volumes finis le système de type thermoviscoélasticité  $\mathbf{S}$  :

$$\frac{\partial^2 u}{\partial t^2} - \operatorname{div} \left( \varepsilon(u) + \varepsilon \left( \frac{\partial u}{\partial t} \right) \right) = g - Df(\theta) \text{ dans } Q, \quad (0.16)$$

$$\frac{\partial \theta}{\partial t} - \operatorname{div} (\lambda(\theta) \nabla \theta) = \varepsilon \left( \frac{\partial u}{\partial t} \right) \cdot \varepsilon \left( \frac{\partial u}{\partial t} \right) - f(\theta) \operatorname{tr} \left( \varepsilon \left( \frac{\partial u}{\partial t} \right) \right) \text{ dans } Q, \quad (0.17)$$

$$u(t=0) = u_0, \quad \frac{\partial u}{\partial t}(t=0) = v_0, \quad \theta(t=0) = \theta_0 \text{ dans } \Omega, \quad (0.18)$$

$$u = 0, \quad \theta = 0, \quad \text{sur } \partial\Omega \times (0, T), \quad (0.19)$$

où  $\Omega$  est un ouvert borné de  $\mathbb{R}^d$ ,  $d = 2$  ou  $d = 3$ ,  $T$  est un nombre positif et  $Q = \Omega \times (0, T)$ . De plus  $f$  et  $\lambda$  sont deux fonctions continues bornées et  $g \in L^2(Q)^d$ . L'équation (0.16) est l'équation de conservation du mouvement tandis que (0.17) est l'équation de conservation de l'énergie dans laquelle la dissipation mécanique n'a pas été linéarisée et est donc du type  $\varepsilon \left( \frac{\partial u}{\partial t} \right) \cdot \varepsilon \left( \frac{\partial u}{\partial t} \right) - f(\theta) \operatorname{tr} \left( \varepsilon \left( \frac{\partial u}{\partial t} \right) \right)$  et appartient donc à  $L^1(Q)$  en général.

Dans [15] une version plus générale ( $\frac{\partial b(u)}{\partial t}$  à la place de  $\frac{u}{\partial t}$ ,  $f$  vérifiant des hypothèses de croissance) est étudiée. Les auteurs démontrent l'existence d'une solution faible-renormalisée dans le sens où  $u$  est solution faible de (0.16) et  $\theta$  est solution renormalisée de (0.17).

Un système de turbulence (modèle réduit) qui couple l'équation de Navier-Stokes avec l'équation de conservation de l'énergie est traité par les volumes finis et les éléments finis dans [47]. Les auteurs montrent que la solution discrète converge vers un couple  $(u, \theta)$  où  $u$  est solution faible de Navier-Stokes et  $\theta$  est solution au sens des distributions de l'équation de conservation de l'énergie.

En nous basant sur [15], [47], sur les résultats établis dans le chapitre 2 et sur les travaux [42, 43] pour la discrétisation de l'équation d'ordre 2 en temps avec les éléments finis, nous montrons que la solution approchée par le schéma converge vers une solution faible-renormalisée du système.





# Chapitre 1

## Finite volume scheme and renormalized solutions for a noncoercive elliptic problem with $L^1$ data \*

### 1.1 Introduction

In this work, we consider the discretization by the cell-centered finite volume method of the following convection-diffusion problem :

$$\begin{aligned} -\Delta u + \operatorname{div}(\mathbf{v}u) + bu &= f \text{ in } \Omega, \\ u &= 0 \text{ on } \partial\Omega, \end{aligned} \tag{1.1}$$

where  $\Omega$  is an open bounded polygonal subset of  $\mathbb{R}^d$ ,  $d \geq 2$ ,  $\mathbf{v} \in L^p(\Omega)^d$ ,  $2 < p < +\infty$  if  $d = 2$ ,  $p = d$  if  $d \geq 3$ ,  $b \in L^2(\Omega)$ ,  $b \geq 0$  and  $f \in L^1(\Omega)$ .

The main difficulties in dealing with the existence and the uniqueness of a solution to problem (1.1) are due to the noncoercive character of the operator  $u \mapsto -\Delta u + \operatorname{div}(\mathbf{v}u) + bu$  and to the  $L^1$  data  $f$ .

The difficulty of the irregular data is overcome by Boccardo and Gallouët in [23] where they had shown the existence of a solution in the sense of distribution in the coercive case for general nonlinear operators. However it is well known that this solution is not unique in general (see the counterexample in [68]). In the linear case it is possible to use the duality method and appropriate notion of weak solution to obtain uniqueness. Particularly in [37] the author had proved that equation (1.1) admits a unique solution using the duality method (the author considers also more general boundary conditions and a measure right-hand side). Equation (1.1) is also studied in [40] where  $\mathbf{v} \in \mathcal{C}(\bar{\Omega})^d$  and  $f$  is a measure. Nonlinear versions of (1.1) are considered in [38] proving local estimates outside the support of the singular part of the right-hand side and in [10, 11] using the framework of renormalized solution.

The theory of renormalized solutions has been introduced in [36] for Boltzmann equations and has been adapted in [58, 59] for elliptic problems with  $L^1$  data. A generalization to elliptic problem with measure data is addressed in [33]. It is well

---

\*. Publié dans le journal "Computational Methods for Applied Mathematics" (voir [53])

known that the renormalized solutions are a convenient framework for parabolic and elliptic equations with  $L^1$  data which provides in general existence, stability and uniqueness results.

Concerning the finite volume discretization, several techniques are developed in [44]. The convergence of the cell-centered finite volume scheme for equation (1.1) has been studied in [46] when  $\mathbf{v} = 0$  and with measure data. In [40] the authors consider the case  $\mathbf{v} \in \mathcal{C}(\bar{\Omega})^d$ ,  $f \in \mathcal{M}(\bar{\Omega})$  (bounded measure in  $\bar{\Omega}$ ) : They prove that the solution of this scheme for equation (1.1) converges to the unique solution of (1.1) in the sense

$$\left\{ \begin{array}{l} u \in \bigcap_{q < \frac{d}{d-1}} W_0^{1,q}(\Omega), \\ \int_{\Omega} \nabla u \cdot \nabla \psi \, d\lambda - \int_{\Omega} u \mathbf{v} \cdot \nabla \psi \, d\lambda + \int_{\Omega} b u \psi \, d\lambda = \int_{\Omega} \psi \, d\mu, \quad \forall \psi \in \bigcup_{s > d} W_0^{1,s}(\Omega). \end{array} \right. \quad (1.2)$$

Discretizations of equation (1.1) with Neumann boundary conditions and right hand side belonging to  $L^2(\Omega)$  are studied in [28] (see also [41] for the continuous problem). In [51] the authors studied problem (1.1) with  $-\operatorname{div}(\lambda(u)\nabla u)$  instead of  $-\Delta u$  and with  $\operatorname{div}(\mathbf{v}) = 0$ . It is also worth mentioning that the convergence of the finite elements approximation of the equation  $-\operatorname{div}(A\nabla u) = f$ , with  $A$  a coercive matrix with coefficients in  $L^\infty(\Omega)$  and  $f \in L^1(\Omega)$ , is proved in [27].

In the present chapter by using the tools developed for finite volume schemes we adapt the strategy used to deal with the existence of a renormalized solution for elliptic equations with  $L^1$  data (see [33, 58, 59]). Recall that a renormalized solution of (1.1) is a measurable function  $u$  defined from  $\Omega$  to  $\bar{\mathbb{R}}$ , such that  $u$  is finite a.e. in  $\Omega$  and

$$\forall k > 0, \quad T_k(u) \in H_0^1(\Omega), \quad (1.3)$$

$$\lim_{k \rightarrow +\infty} \frac{1}{k} \int_{\Omega} |\nabla T_k(u)|^2 \, dx = 0, \quad (1.4)$$

$$\left\{ \begin{array}{l} \forall h \in \mathcal{C}_c^1(\mathbb{R}), \quad \forall \psi \in H_0^1(\Omega) \cap L^\infty(\Omega), \\ \int_{\Omega} \nabla u h(u) \nabla \psi \, dx + \int_{\Omega} \nabla u \psi \nabla u h'(u) \, dx - \int_{\Omega} u h(u) \mathbf{v} \cdot \nabla \psi \, dx \\ - \int_{\Omega} u h'(u) \psi \mathbf{v} \cdot \nabla u \, dx + \int_{\Omega} b u h(u) \psi \, dx = \int_{\Omega} \psi h(u) f \, dx, \end{array} \right. \quad (1.5)$$

with  $T_k$  the truncate function at height  $k$  (see Figure 1 below).

Since  $h$  has a compact support, each term of (1.5) is well defined. The existence and the uniqueness of a renormalized solution to (1.1) is proved in [11]. The main originality in the present work is that we pass to the limit in a "renormalized discrete version", this is to say that we take a discrete version of  $\psi h(u)$  as test function in the finite volume scheme. A first difficulty is to establish a discrete version of the estimate on the energy (1.4). Moreover it is worth noting that in (1.5) all the terms are "truncated" while a discrete version of  $\varphi h(u)$  in the finite volume scheme leads to some residual terms which are not "truncated". The second difficulty is then to handle these residual terms. With respect to [40, 46] in the context of finite volume method or to [27] in the context of finite element method it is worth noting that we

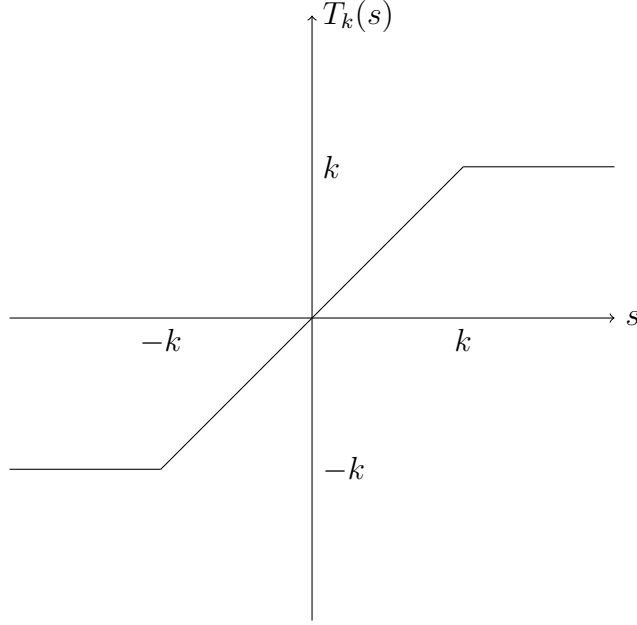


FIGURE 1.1 – The function  $T_k$

do not use the discrete or continuous Boccardo-Gallouët estimates. Moreover the method developed in the present paper allows one to deal with nonlinear version of (1.1) in the sense that the solution of the discrete scheme converges to the unique renormalized solution (see Section 5).

The paper is organized as follows. In Section 2, we present the finite volume scheme and the properties of the discrete gradient. Section 3 is devoted to prove several estimates, especially the discrete equivalent to (1.4) which is crucial to pass to the limit in the scheme. In Section 4, we prove the convergence of the cell-centered finite volume scheme, passing to the limit and using a density argument. At last we explain in Section 5 that we are able to show similar results for a quasilinear problem with  $-\operatorname{div}(\lambda(u)\nabla u)$  in place of  $-\Delta u$ .

## 1.2 The finite volume scheme

Let us define the admissibility of the mesh in the present work.

**Definition 1.2.1 (Admissible mesh)** *Let  $\Omega$  be an open bounded polygonal subset of  $\mathbb{R}^d$ . An admissible finite volume mesh of  $\Omega$  is given by a finite partition  $\mathcal{T}$  of  $\Omega$  in polygonal convex sets, called the "control volumes", by a finite family  $\mathcal{E}$  of disjoint subsets of  $\bar{\Omega}$  contained in affine hyperplanes, called the "edges", and by a family  $\mathcal{P} = (x_K)_{K \in \mathcal{T}}$  of points in  $\Omega$  such that*

- *each  $\sigma \in \mathcal{E}$  is a non-empty open subset of  $\partial K$  for some  $K \in \mathcal{T}$ ,*
- *by denoting  $\mathcal{E}(K) = \{\sigma \in \mathcal{E}; \sigma \in \partial K\}$ , one has  $\partial K = \cup_{\sigma \in \mathcal{E}(K)} \bar{\sigma}$  for all  $K \in \mathcal{T}$ ,*
- *for all  $K \neq L$  in  $\mathcal{T}$ , either the  $(d-1)$ -dimensional measure of  $\bar{K} \cap \bar{L}$  is null, or  $\bar{K} \cap \bar{L} = \bar{\sigma}$  for some  $\sigma \in \mathcal{E}$ , that we denote then  $\sigma = K|L$ ,*
- *for all  $K \in \mathcal{T}$ ,  $x_K$  is in the interior of  $K$ ,*

- for all  $\sigma = K|L \in \mathcal{E}$ , the line  $(x_K, x_L)$  intersects and is orthogonal to  $\sigma$ ,
- for all  $\sigma \in \mathcal{E}$ ,  $\sigma \subset \partial\Omega \cap \partial K$ , the line which is orthogonal to  $\sigma$  and going through  $x_K$  intersects  $\sigma$ .

We denote by  $|K|$  (resp.  $|\sigma|$ ) the Lebesgue measure of  $K \in \mathcal{T}$  (resp.  $\sigma \in \mathcal{E}$ ). The unit normal to  $\sigma \in \mathcal{E}(K)$  outward to  $K$  is denoted by  $\eta_{K,\sigma}$ .  $\mathcal{E}_{int}$  (resp.  $\mathcal{E}_{ext}$ ) is defined as the set of interior (resp. boundary) edges.

For any  $K \in \mathcal{T}$  and  $\sigma \in \mathcal{E}(K)$  we denote by  $d_{K,\sigma}$  the Euclidean distance between  $x_K$  and  $\sigma$ . For any  $\sigma \in \mathcal{E}$ , we define  $d_\sigma = d_{K,\sigma} + d_{L,\sigma}$  if  $\sigma = K|L \in \mathcal{E}_{int}$  (in which case  $d_\sigma$  is the Euclidean distance between  $x_K$  and  $x_L$ ) and  $d_\sigma = d_{K,\sigma}$  if  $\sigma \in \mathcal{E}_{ext} \cap \mathcal{E}(K)$ . The size of the mesh, denoted by  $h_{\mathcal{T}}$ , is defined by  $h_{\mathcal{T}} = \sup_{K \in \mathcal{T}} \text{diam}(K)$ .

In the continuous case the usual tools to solve the problem are Poincaré and Sobolev inequalities. In the discrete case we will need such estimates, so we have to establish their discrete versions (see [32]). Let us define the discrete  $W_0^{1,q}$  norm.

**Definition 1.2.2 (Discrete  $W_0^{1,q}$  norm)** *Let  $\Omega$  be an open bounded polygonal subset of  $\mathbb{R}^d$ ,  $d \geq 2$ , and let  $\mathcal{T}$  be an admissible mesh. Define  $X(\mathcal{T})$  as the set of functions from  $\Omega$  to  $\mathbb{R}$  which are constant over each control volume of the mesh. For  $v_{\mathcal{T}} \in X(\mathcal{T})$  and  $q \in [1, +\infty[$ , we define the discrete  $W_0^{1,q}$  norm by*

$$\|v_{\mathcal{T}}\|_{1,q,\mathcal{T}}^q = \sum_{\substack{\sigma \in \mathcal{E}_{int} \\ \sigma = K|L}} |\sigma| d_\sigma \left| \frac{v_K - v_L}{d_\sigma} \right|^q + \sum_{\substack{\sigma \in \mathcal{E}_{ext} \\ \sigma \in \mathcal{E}(K)}} |\sigma| d_\sigma \left| \frac{v_K}{d_\sigma} \right|^q$$

where  $v_K$  denotes the value taken by  $v$  on the control volume  $K$ .

**Proposition 1.2.3 (Discrete Poincaré inequality)** *Let  $\mathcal{T}$  be an admissible mesh and  $v_{\mathcal{T}} \in X(\mathcal{T})$ . Then, if  $1 \leq q \leq 2$ ,*

$$\|v_{\mathcal{T}}\|_{L^q(\Omega)} \leq \text{diam}(\Omega) \|v_{\mathcal{T}}\|_{1,q,\mathcal{T}}.$$

**Proposition 1.2.4 (Discrete Sobolev inequality)** *Let  $1 \leq q \leq 2$ ,  $\mathcal{T}$  be an admissible mesh and  $\xi > 0$  satisfying*

$$\text{for all } K \in \mathcal{T} \text{ and all } \sigma \in \mathcal{E}(K), \quad d_{K,\sigma} \geq \xi d_\sigma.$$

*Then, with  $q^* = \frac{dq}{d-q}$  if  $q < d$  and  $q^* < \infty$  if  $q = d = 2$ , there exists  $C > 0$  only depending on  $(\Omega, q, q^*, \xi)$  such that, for all  $v_{\mathcal{T}} \in X(\mathcal{T})$ ,*

$$\|v_{\mathcal{T}}\|_{L^{q^*}(\Omega)} \leq C \|v_{\mathcal{T}}\|_{1,q,\mathcal{T}}.$$

Before writing the finite volume scheme, let us define a discrete finite volume gradient (see [51]).

**Definition 1.2.5 (Discrete finite volume gradient)** *For  $K \in \mathcal{T}$  and  $\sigma \in \mathcal{E}(K)$ , we define the volume  $D_{K,\sigma}$  as the cone of basis  $\sigma$  and of opposite vertex  $x_K$ . Then, we define the "diamond-cell"  $D_\sigma$  (see Figure 1.2) by*

$$\begin{aligned} D_\sigma &= D_{K,\sigma} \cup D_{L,\sigma} && \text{if } \sigma = K|L \in \mathcal{E}_{int}, \\ D_\sigma &= D_{K,\sigma} && \text{if } \sigma \in \mathcal{E}_{ext} \cap \mathcal{E}(K). \end{aligned}$$

For any  $v_{\mathcal{T}} \in X(\mathcal{T})$ , we define the discrete gradient  $\nabla_{\mathcal{T}} v_{\mathcal{T}}$  as the piecewise constant function over each diamond cell and given by

$$\begin{aligned} \forall \sigma \in \mathcal{E}_{int}, \sigma = K|L, & \quad \nabla_{\mathcal{T}} v(\mathbf{x}) = d \frac{v_L - v_K}{d_{\sigma}} \eta_{K,\sigma}, \quad \forall \mathbf{x} \in D_{\sigma}, \\ \forall \sigma \in \mathcal{E}_{ext} \cap \mathcal{E}(K), & \quad \nabla_{\mathcal{T}} v(\mathbf{x}) = d \frac{0 - v_K}{d_{\sigma}} \eta_{K,\sigma}, \quad \forall \mathbf{x} \in D_{\sigma}. \end{aligned}$$

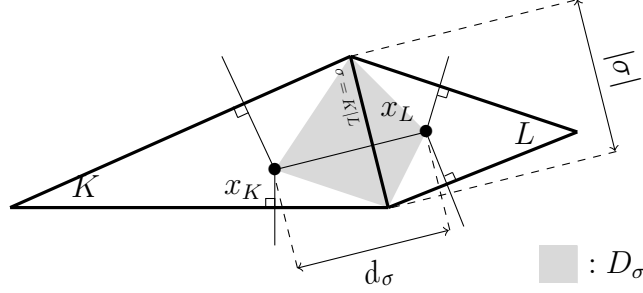


FIGURE 1.2 – The diamond  $D_{\sigma}$

**Lemma 1.2.6 (Weak convergence of the finite volume gradient)** Let  $(\mathcal{T}_m)_{m \geq 1}$  be a sequence of admissible meshes such that there exists  $\xi > 0$  satisfying

$$\text{for all } m \geq 1 \text{ for all } K \in \mathcal{T} \text{ and all } \sigma \in \mathcal{E}(K), \quad d_{K,\sigma} \geq \xi d_{\sigma},$$

and such that  $h_{\mathcal{T}_m} \rightarrow 0$ . Let  $v_{\mathcal{T}_m} \in X(\mathcal{T}_m)$  and let us assume that there exists  $\alpha \in [1, +\infty[$  and  $C > 0$  such that  $\|v_{\mathcal{T}_m}\|_{1,\alpha,\mathcal{T}_m} \leq C$ , and that  $v_{\mathcal{T}_m}$  converges in  $L^1(\Omega)$  to  $v \in W_0^{1,\alpha}(\Omega)$ . Then  $\nabla_{\mathcal{T}_m} v_{\mathcal{T}_m}$  converges to  $\nabla v$  weakly in  $L^{\alpha}(\Omega)^d$ .

The proof of a more general result can be found in [51].

Let  $\mathcal{T}$  be an admissible mesh, we can define the finite volume discretization of (1.1). For  $K \in \mathcal{T}$  and  $\sigma \in \mathcal{E}(K)$ , we define

$$b_K = \frac{1}{|K|} \int_K b \, dx \quad \text{and} \quad \mathbf{v}_{K,\sigma} = \frac{1}{|D_{\sigma}|} \int_{D_{\sigma}} \mathbf{v} \cdot \eta_{K,\sigma} \, dx, \quad (1.6)$$

and we can write the scheme as the following set of equations

$$\forall K \in \mathcal{T}, \quad \sum_{\sigma \in \mathcal{E}(K)} \frac{|\sigma|}{d_{\sigma}} (u_K - u_L) + \sum_{\sigma \in \mathcal{E}(K)} |\sigma| \mathbf{v}_{K,\sigma} u_{\sigma,+} + |K| b_K u_K = \int_K f \, dx, \quad (1.7)$$

with  $u_L = 0$  if  $\sigma \in \mathcal{E}_{ext} \cap \mathcal{E}(K)$  and

$$\forall \sigma = K|L \in \mathcal{E}_{int}, \quad u_{\sigma,+} = u_K \text{ if } \mathbf{v}_{K,\sigma} \geq 0, \quad u_{\sigma,+} = u_L \text{ otherwise}, \quad (1.8)$$

$$\forall \sigma \in \mathcal{E}_{ext} \cap \mathcal{E}(K), \quad u_{\sigma,+} = u_K \text{ if } \mathbf{v}_{K,\sigma} \geq 0, \quad u_{\sigma,+} = 0 \text{ otherwise}. \quad (1.9)$$

We denote  $u_{\sigma,-}$  the downstream choice of  $u$ , i.e.  $u_{\sigma,-}$  is such that  $\{u_{\sigma,+}, u_{\sigma,-}\} = \{u_K, u_L\}$  (with  $u_L = 0$  if  $\sigma \in \mathcal{E}_{ext} \cap \mathcal{E}(K)$ ).

### 1.3 Estimates

In this section, we first establish in Proposition 1.3.1 an estimate on  $\ln(1 + |u_{\mathcal{T}}|)$  which is crucial to control the measure of the set  $\{|u_{\mathcal{T}}| > n\}$ . Then, we show in Proposition 1.3.3 an estimate on  $T_n(u_{\mathcal{T}})$  and the convergence of  $T_n(u_{\mathcal{T}})$  to  $T_n(u)$ . The estimates in Propositions 1.3.1 and 1.3.3 are similar to the ones derived in [40]; the fact that we consider here  $\mathbf{v}$  in  $L^p(\Omega)^d$  and  $f$  in  $L^1(\Omega)$  in place of  $\mathbf{v}$  in  $\mathcal{C}^1(\Omega)^d$  and  $f$  as a measure in [40] does not play any role. Finally, we prove in Proposition 1.3.4 a discrete version of the decay of the energy of the truncate (see (1.4)).

**Proposition 1.3.1** *Let  $\mathcal{T}$  be an admissible mesh. If  $u_{\mathcal{T}} = (u_K)_{K \in \mathcal{T}}$  is a solution to (1.7), then*

$$\|\ln(1 + |u_{\mathcal{T}}|)\|_{1,2,\mathcal{T}}^2 \leq 2\|f\|_{L^1(\Omega)} + d|\Omega|^{\frac{p-2}{p}} \|\mathbf{v}\|_{L^p(\Omega)^d}^2 \quad (1.10)$$

where  $|\mathbf{v}|$  denotes the Euclidean norm of  $\mathbf{v}$  in  $\mathbb{R}^d$ .

**Proof** We can easily adapt the proof of the *log*-estimate derived in [40] where  $\mathbf{v} \in \mathcal{C}(\bar{\Omega})^d$  and  $f \in L^1(\Omega)$ . Using a Sobolev inequality the following term appears

$$\left( \sum_{\sigma \in \mathcal{E}} |\sigma| d_{\sigma} |\mathbf{v}_{K,\sigma}|^p \right)$$

whereas it is  $\|\mathbf{v}\|_{L^\infty(\Omega)}$  in [40]. In order to bound this term, we can use Jensen's inequality such as in [28] (where  $\mathbf{v} \in L^p(\Omega)^d$  and  $f \in L^2(\Omega)$ , with Neumann boundary conditions). Thus

$$\left( \sum_{\sigma \in \mathcal{E}} |\sigma| d_{\sigma} |\mathbf{v}_{K,\sigma}|^p \right)^{\frac{1}{p}} \leq \left( \sum_{\sigma \in \mathcal{E}} \frac{|\sigma| d_{\sigma}}{|D_{\sigma}|} \int_{D_{\sigma}} |\mathbf{v} \cdot \eta_{K,\sigma}|^p \right)^{\frac{1}{p}} \leq d^{\frac{1}{p}} \|\mathbf{v}\|_{L^p(\Omega)^d},$$

and the proof is complete.  $\square$

**Corollary 1.3.2** *Let  $\mathcal{T}$  be an admissible mesh. If  $u_{\mathcal{T}} = (u_K)_{K \in \mathcal{T}}$  is a solution to (1.7) and, for  $n > 0$ ,  $E_n = \{|u_{\mathcal{T}}| > n\}$ , then there exists  $C > 0$  only depending on  $(\Omega, \mathbf{v}, f, d, p)$  such that*

$$|E_n| \leq \frac{C}{(\ln(1+n))^2}. \quad (1.11)$$

**Proof** Using Proposition 1.3.1 and the discrete Poincaré inequality, we get that

$$\|\ln(1 + |u_{\mathcal{T}}|)\|_{L^2(\Omega)}^2 \leq C(2\|f\|_{L^1(\Omega)} + d|\Omega|^{\frac{p-2}{p}} \|\mathbf{v}\|_{L^p(\Omega)^d}^2),$$

where  $C$  only depends on  $\text{diam}(\Omega)$ . Hence,

$$(\ln(1+n))^2 |E_n| \leq C(2\|f\|_{L^1(\Omega)} + d|\Omega|^{\frac{p-2}{p}} \|\mathbf{v}\|_{L^p(\Omega)^d}^2).$$

$\square$

**Proposition 1.3.3 (Estimate on  $T_n(u_{\mathcal{T}})$ )** *Let  $\mathcal{T}$  be an admissible mesh. If  $u_{\mathcal{T}} = (u_K)_{K \in \mathcal{T}}$  is a solution to (1.7), then there exists  $C > 0$  only depending on  $(\Omega, \mathbf{v}, f, n, d)$  such that*

$$\|T_n(u_{\mathcal{T}})\|_{1,2,\mathcal{T}} \leq C \quad \forall n > 0. \quad (1.12)$$

*Moreover, if  $(\mathcal{T}_m)_{m \geq 1}$  is a sequence of admissible meshes such that there exists  $\xi > 0$  satisfying*

$$\text{for all } m \geq 1 \text{ for all } K \in \mathcal{T} \text{ and all } \sigma \in \mathcal{E}(K), \quad d_{K,\sigma} \geq \xi d_{\sigma},$$

*there exists a measurable function  $u$  finite a.e. in  $\Omega$  such that, up to a subsequence,  $T_n(u_{\mathcal{T}_m})$  converges to  $T_n(u)$  weakly in  $H_0^1(\Omega)$ , strongly in  $L^2(\Omega)$  and a.e. in  $\Omega$ .*

**Proof** The proof is divided into two steps. In Step 1 we derive the estimate (1.12) on the truncate on  $u_{\mathcal{T}}$ . Step 2 is devoted to extract subsequences.

**Step 1 : Estimate on  $T_n(u_{\mathcal{T}})$**

By multiplying each equation of the scheme by  $T_n(u_K)$ , by summing over each control volume and by reordering the sums, we obtain  $T_1 + T_2 + T_3 = T_4$  with

$$\begin{aligned} T_1 &= \sum_{\sigma \in \mathcal{E}} \frac{|\sigma|}{d_{\sigma}} (u_K - u_L) (T_n(u_K) - T_n(u_L)), \\ T_2 &= \sum_{\sigma \in \mathcal{E}} |\sigma| \mathbf{v}_{K,\sigma} u_{\sigma,+} (T_n(u_K) - T_n(u_L)), \\ T_3 &= \sum_{K \in \mathcal{T}} |K| b_K u_K T_n(u_K), \\ T_4 &= \sum_{K \in \mathcal{T}} \int_K f T_n(u_K) dx. \end{aligned}$$

Since  $b$  is nonnegative and since  $r T_n(r) \geq 0 \quad \forall r$ , we notice that  $T_3 \geq 0$ . Moreover, since  $T_n$  is bounded by  $n$ , we deduce that  $|T_4| \leq n \|f\|_{L^1(\Omega)}$ . Thus  $T_1 \leq n \|f\|_{L^1(\Omega)} - T_2$ . As in [40],  $T_2$  can be rewritten as

$$-T_2 = \sum_{\sigma \in \mathcal{E}} |\sigma| |\mathbf{v}_{K,\sigma}| u_{\sigma,+} (T_n(u_{\sigma,-}) - T_n(u_{\sigma,+})).$$

As in [40], we define the subset  $\mathcal{A}$  of edges by

$$\mathcal{A} = \{\sigma \in \mathcal{E}; u_{\sigma,+} \geq u_{\sigma,-}, u_{\sigma,+} < 0\} \cup \{\sigma \in \mathcal{E}; u_{\sigma,+} < u_{\sigma,-}, u_{\sigma,+} \geq 0\}, \quad (1.13)$$

and since  $T_n$  is non decreasing we have

$$-T_2 \leq \sum_{\sigma \in \mathcal{A}} |\sigma| |\mathbf{v}_{K,\sigma}| u_{\sigma,+} (T_n(u_{\sigma,-}) - T_n(u_{\sigma,+})).$$

By observing that  $\forall \sigma \in \mathcal{A} \quad |u_{\sigma,+}| \geq n$  implies that  $|u_{\sigma,-}| \geq n$  we deduce that

$$\forall \sigma \in \mathcal{A}, \quad u_{\sigma,+} (T_n(u_{\sigma,-}) - T_n(u_{\sigma,+})) = T_n(u_{\sigma,+}) (T_n(u_{\sigma,-}) - T_n(u_{\sigma,+})).$$



It follows that

$$\begin{aligned}
-T_2 &\leq \sum_{\sigma \in \mathcal{A}} |\sigma| |\mathbf{v}_{K,\sigma}| T_n(u_{\sigma,+}) (T_n(u_{\sigma,-}) - T_n(u_{\sigma,+})) \\
&\leq \left( \sum_{\sigma \in \mathcal{E}} |\sigma| d_\sigma |\mathbf{v}_{K,\sigma}|^2 \right)^{\frac{1}{2}} \left( \sum_{\sigma \in \mathcal{A}} \frac{|\sigma|}{d_\sigma} T_n(u_{\sigma,+})^2 (T_n(u_{\sigma,-}) - T_n(u_{\sigma,+}))^2 \right)^{\frac{1}{2}} \\
&\leq n d^{\frac{1}{2}} \|\mathbf{v}\|_{L^2(\Omega)^d} \left( \sum_{\sigma \in \mathcal{A}} \frac{|\sigma|}{d_\sigma} (T_n(u_{\sigma,-}) - T_n(u_{\sigma,+}))^2 \right)^{\frac{1}{2}} \\
&\leq \frac{1}{2} n^2 d \|\mathbf{v}\|_{L^2(\Omega)^d}^2 + \frac{1}{2} \sum_{\sigma \in \mathcal{A}} \frac{|\sigma|}{d_\sigma} (T_n(u_{\sigma,-}) - T_n(u_{\sigma,+}))^2 \\
&\leq \frac{1}{2} n^2 d \|\mathbf{v}\|_{L^2(\Omega)^d}^2 + \frac{1}{2} \sum_{\sigma \in \mathcal{E}} \frac{|\sigma|}{d_\sigma} (T_n(u_K) - T_n(u_L))^2.
\end{aligned}$$

Since  $|T_n(u_K) - T_n(u_L)| \leq |u_K - u_L|$ , we have

$$-T_2 \leq \frac{1}{2} n^2 d \|\mathbf{v}\|_{L^2(\Omega)^d}^2 + \frac{1}{2} \sum_{\sigma \in \mathcal{E}} \frac{|\sigma|}{d_\sigma} (u_K - u_L) (T_n(u_K) - T_n(u_L)),$$

and we can deduce that

$$\frac{1}{2} \sum_{\sigma \in \mathcal{E}} \frac{|\sigma|}{d_\sigma} (u_K - u_L) (T_n(u_K) - T_n(u_L)) \leq n \|f\|_{L^1(\Omega)} + \frac{1}{2} n^2 d \|\mathbf{v}\|_{L^2(\Omega)^d}^2.$$

Therefore, using again the fact that  $|T_n(u_K) - T_n(u_L)| \leq |u_K - u_L|$  yields

$$\frac{1}{2} \sum_{\sigma \in \mathcal{E}} \frac{|\sigma|}{d_\sigma} (T_n(u_K) - T_n(u_L))^2 \leq n \|f\|_{L^1(\Omega)} + \frac{1}{2} n^2 d \|\mathbf{v}\|_{L^2(\Omega)^d}^2.$$

Applying Lemma 1.2.6 and the diagonal process, up to a subsequence still denoted by  $\mathcal{T}_m$ , for any  $n \geq 1$  there exists  $v_n$  belonging to  $H_0^1(\Omega)$  such that  $T_n(u_{\mathcal{T}}) \rightarrow v_n$  and  $\nabla_{\mathcal{T}} T_n(u_{\mathcal{T}}) \rightarrow \nabla v_n$  in  $L^2(\Omega)^d$ .

## Step 2 : Up to a subsequence, $u_{\mathcal{T}}$ is a Cauchy sequence in measure

In this step, we follow a proof of [33] to show that  $u_{\mathcal{T}_m}$  converges a.e. to  $u$ .

Let  $\omega > 0$ . For all  $n > 0$  and all sequences of admissible meshes  $(\mathcal{T}_m)_{m \geq 1}$  and  $(\mathcal{T}_p)_{p \geq 1}$ , we have

$$\{|u_{\mathcal{T}_m} - u_{\mathcal{T}_p}| > \omega\} \subset \{|u_{\mathcal{T}_m}| > n\} \cup \{|u_{\mathcal{T}_p}| > n\} \cup \{|T_n(u_{\mathcal{T}_m}) - T_n(u_{\mathcal{T}_p})| > \omega\}.$$

Let  $\varepsilon > 0$  fixed. By Corollary 1.3.2, let  $n > 0$  such that, for all admissible meshes  $\mathcal{T}_m$  and  $\mathcal{T}_p$

$$\text{meas}(\{|u_{\mathcal{T}_m}| > n\}) + \text{meas}(\{|u_{\mathcal{T}_p}| > n\}) < \frac{\varepsilon}{2}.$$

Once  $n$  is chosen, we deduce from Step 1 that  $T_n(u_{\mathcal{T}_m})$  is a Cauchy sequence in measure, thus

$$\exists h_0 > 0; \forall h_{\mathcal{T}_m}, h_{\mathcal{T}_p} < h_0 \quad \text{meas}(\{|T_n(u_{\mathcal{T}_m}) - T_n(u_{\mathcal{T}_p})| > \omega\}) \leq \frac{\varepsilon}{2}.$$

Therefore, we deduce that  $\forall h_{\mathcal{T}_m}, h_{\mathcal{T}_p} < h_0$

$$\text{meas}\{|u_{\mathcal{T}_m} - u_{\mathcal{T}_p}| > \omega\} < \varepsilon.$$

Hence  $u_{\mathcal{T}_m}$  is a Cauchy sequence in measure. Consequently, up to a subsequence still indexed by  $\mathcal{T}_m$ , there exists a measurable function  $u$  such that  $u_{\mathcal{T}_m} \rightarrow u$  a.e. in  $\Omega$ . Due to Corollary 1.3.2  $u$  is finite a.e. in  $\Omega$ . Moreover from convergences obtained in Step 1 we get that  $T_n(u) \in H_0^1(\Omega)$  and

$$\nabla_{\mathcal{T}} T_n(u_{\mathcal{T}}) \rightharpoonup \nabla T_n(u) \text{ in } (L^2(\Omega))^d. \quad (1.14)$$

□

In the following proposition we prove a uniform estimate on the truncated energy of  $u_{\mathcal{T}}$  (see (1.15)) which is crucial to pass to the limit in the approximate problem. We explicitly observe that (1.15) is the discrete version of (1.4) which is imposed in the definition of the renormalized solution for elliptic equation with  $L^1$  data. As in the continuous case (1.15) is related to the regularity of  $f : f \in L^1(\Omega)$  and does not charge any zero-Lebesgue set. If we replace  $f$  by a bounded Radon measure we cannot expect to have (1.15) (see [33] for the definition of the renormalized solution in this case). Since the equation contains the term  $\text{div}(\mathbf{v} u)$  we also have to uniformly control the discrete version of  $\frac{1}{n} \int_{\Omega} \mathbf{v} u \nabla T_n(u) dx$  which is stated in (1.16).

**Proposition 1.3.4 (Discrete estimate on the energy)**

Let  $(\mathcal{T}_m)_{m \geq 1}$  be a sequence of admissible meshes such that there exists  $\xi > 0$  satisfying

$$\text{for all } m \geq 1 \text{ for all } K \in \mathcal{T} \text{ and all } \sigma \in \mathcal{E}(K), d_{K,\sigma} \geq \xi d_{\sigma}.$$

If  $u_{\mathcal{T}_m} = (u_K)_{K \in \mathcal{T}_m}$  is a solution to (1.7), then

$$\lim_{n \rightarrow +\infty} \overline{\lim}_{h_{\mathcal{T}_m} \rightarrow 0} \frac{1}{n} \sum_{\sigma \in \mathcal{E}} \frac{|\sigma|}{d_{\sigma}} (u_K - u_L)(T_n(u_K) - T_n(u_L)) = 0, \quad (1.15)$$

where  $u_L = 0$  if  $\sigma \in \mathcal{E}_{ext}$ , and

$$\lim_{n \rightarrow +\infty} \overline{\lim}_{h_{\mathcal{T}} \rightarrow 0} \frac{1}{n} \sum_{\sigma \in \mathcal{E}} |\sigma| |\mathbf{v}_{K,\sigma}| |u_{\sigma,+}| |T_n(u_{\sigma,+}) - T_n(u_{\sigma,-})| = 0. \quad (1.16)$$

**Proof** We first establish (1.15). Let  $\mathcal{T}$  be an admissible mesh and let  $u_{\mathcal{T}}$  be a solution of (1.7). Multiplying each equation of the scheme by  $\frac{T_n(u_K)}{n}$ , summing on  $K \in \mathcal{T}$  and gathering by edges lead to

$$T_1 + T_2 + T_3 = T_4$$

with

$$T_1 = \frac{1}{n} \sum_{\sigma \in \mathcal{E}} \frac{|\sigma|}{d_{\sigma}} (u_K - u_L)(T_n(u_K) - T_n(u_L)), \quad (1.17)$$

$$T_2 = \frac{1}{n} \sum_{\sigma \in \mathcal{E}} |\sigma| \mathbf{v}_{K,\sigma} u_{\sigma,+} (T_n(u_K) - T_n(u_L)), \quad (1.18)$$

$$T_3 = \frac{1}{n} \sum_{K \in \mathcal{T}} |K| b_K u_K T_n(u_K), \quad (1.19)$$

$$T_4 = \frac{1}{n} \sum_{K \in \mathcal{T}} \int_K f T_n(u_K) dx. \quad (1.20)$$

Since  $b$  is nonnegative and since  $r T_n(r) \geq 0 \forall r$ , we get  $T_3 \geq 0$ .  
 Due to the definition of  $u_{\mathcal{T}}$  we have

$$T_4 = \int_{\Omega} f \frac{T_n(u_{\mathcal{T}})}{n} dx.$$

In view of the pointwise convergence of  $u_{\mathcal{T}}$  to  $u$  we obtain that  $T_n(u_{\mathcal{T}})$  converges to  $T_n(u)$  a.e. and in  $L^\infty$  weak  $\star$  as  $h_{\mathcal{T}} \rightarrow 0$ . It follows that  $\lim_{h_{\mathcal{T}} \rightarrow 0} T_4 = \int_{\Omega} f \frac{T_n(u)}{n} dx$ .  
 Since  $u$  is finite a.e. in  $\Omega$ ,  $\frac{T_n(u)}{n}$  converges to 0 a.e. and in  $L^\infty$  weak  $\star$ , and since  $f$  belongs to  $L^1(\Omega)$ , the Lebesgue theorem implies that

$$\lim_{n \rightarrow +\infty} \lim_{h_{\mathcal{T}} \rightarrow 0} T_4 = 0. \quad (1.21)$$

We now prove that

$$-T_2 \leq \frac{1}{2} T_1 + R(n, h_{\mathcal{T}})$$

with  $R$  verifying  $\lim_{n \rightarrow +\infty} \overline{\lim}_{h_{\mathcal{T}} \rightarrow 0} R(n, h_{\mathcal{T}}) = 0$ .

Let  $\sigma \in \mathcal{E}$ . By the definition (1.8) of  $u_{\sigma,+}$  and recalling that  $u_{\sigma,-}$  is the downstream choice of  $u$ , if  $\mathbf{v}_{K,\sigma} \geq 0$  it gives

$$\mathbf{v}_{K,\sigma}(T_n(u_K) - T_n(u_L)) = \mathbf{v}_{K,\sigma}(T_n(u_{\sigma,+}) - T_n(u_{\sigma,-})),$$

and if  $\mathbf{v}_{K,\sigma} < 0$  it gives

$$\mathbf{v}_{K,\sigma}(T_n(u_K) - T_n(u_L)) = -\mathbf{v}_{K,\sigma}(T_n(u_{\sigma,+}) - T_n(u_{\sigma,-})).$$

In consequence,  $T_2$  can be rewritten as

$$-T_2 = \frac{1}{n} \sum_{\sigma \in \mathcal{E}} |\sigma| |\mathbf{v}_{K,\sigma}| u_{\sigma,+} (T_n(u_{\sigma,-}) - T_n(u_{\sigma,+})).$$

In view of the definition (1.13) of  $\mathcal{A}$  and since  $T_n$  is non decreasing we have

$$-T_2 \leq \frac{1}{n} \sum_{\sigma \in \mathcal{A}} |\sigma| |\mathbf{v}_{K,\sigma}| u_{\sigma,+} (T_n(u_{\sigma,-}) - T_n(u_{\sigma,+})). \quad (1.22)$$

Because for any  $\sigma \in \mathcal{A}$ ,  $|u_{\sigma,+}| \geq n$  implies  $|u_{\sigma,-}| \geq n$ , we can split the previous sum on  $\{|u_{\sigma,+}| \leq r\} \cup \{r \leq |u_{\sigma,+}| \leq n\}$  where  $r > 0$  will be chosen later. It leads to  $-T_2 \leq I_1 + I_2$  with

$$I_1 = \frac{1}{n} \sum_{\substack{\sigma \in \mathcal{A} \\ |u_{\sigma,+}| \leq r}} |\sigma| |\mathbf{v}_{K,\sigma}| u_{\sigma,+} (T_n(u_{\sigma,-}) - T_n(u_{\sigma,+})), \quad (1.23)$$

$$I_2 = \frac{1}{n} \sum_{\substack{\sigma \in \mathcal{A} \\ r \leq |u_{\sigma,+}| \leq n}} |\sigma| |\mathbf{v}_{K,\sigma}| u_{\sigma,+} (T_n(u_{\sigma,-}) - T_n(u_{\sigma,+})). \quad (1.24)$$

For any  $\sigma \in \mathcal{A}$   $|u_{\sigma,+}| \geq r$  implies  $|u_{\sigma,-}| \geq r$ , thus for  $d \geq 3$  :

$$\begin{aligned}
|I_2| &\leq \frac{1}{n} \sum_{\substack{\sigma \in \mathcal{A} \\ r \leq |u_{\sigma,+}| \leq n \\ r \leq |u_{\sigma,-}|}} |\sigma| |\mathbf{v}_{K,\sigma}| T_n(u_{\sigma,+}) [T_n(u_{\sigma,-}) - T_n(u_{\sigma,+})] \\
&\leq \frac{1}{n} \left( \sum_{\substack{\sigma \in \mathcal{A} \\ r \leq |u_{\sigma,+}| \leq n \\ r \leq |u_{\sigma,-}|}} |\sigma| d_\sigma |\mathbf{v}_{K,\sigma}|^d \right)^{\frac{1}{d}} \left( \sum_{\substack{\sigma \in \mathcal{A} \\ r \leq |u_{\sigma,+}| \leq n \\ r \leq |u_{\sigma,-}|}} |\sigma| d_\sigma |T_n(u_{\sigma,+})|^{\frac{2d}{d-2}} \right)^{\frac{d-2}{2d}} \\
&\quad \left( \sum_{\substack{\sigma \in \mathcal{A} \\ r \leq |u_{\sigma,+}| \leq n \\ r \leq |u_{\sigma,-}|}} \frac{|\sigma|}{d_\sigma} [T_n(u_{\sigma,-}) - T_n(u_{\sigma,+})]^2 \right)^{\frac{1}{2}}.
\end{aligned}$$

Using the discrete Sobolev inequality (Proposition 1.2.4)

$$|I_2| \leq \frac{d^{\frac{1}{2}} \|\mathbf{v}\|_{L^d(E_r)^d}}{n} \sum_{\substack{\sigma \in \mathcal{A} \\ r \leq |u_{\sigma,+}| \leq n \\ r \leq |u_{\sigma,-}|}} \frac{|\sigma|}{d_\sigma} (T_n(u_K) - T_n(u_L))^2. \quad (1.25)$$

For  $d = 2$  we see that

$$\begin{aligned}
|I_2| &\leq \frac{1}{n} \sum_{\substack{\sigma \in \mathcal{A} \\ r \leq |u_{\sigma,+}| \leq n \\ r \leq |u_{\sigma,-}|}} |\sigma| |\mathbf{v}_{K,\sigma}| T_n(u_{\sigma,+}) [T_n(u_{\sigma,-}) - T_n(u_{\sigma,+})] \\
&\leq \frac{1}{n} \left( \sum_{\substack{\sigma \in \mathcal{A} \\ r \leq |u_{\sigma,+}| \leq n \\ r \leq |u_{\sigma,-}|}} |\sigma| d_\sigma |\mathbf{v}_{K,\sigma}|^4 \right)^{\frac{1}{4}} \left( \sum_{\substack{\sigma \in \mathcal{A} \\ r \leq |u_{\sigma,+}| \leq n \\ r \leq |u_{\sigma,-}|}} |\sigma| d_\sigma |T_n(u_{\sigma,+})|^4 \right)^{\frac{1}{4}} \\
&\quad \left( \sum_{\substack{\sigma \in \mathcal{A} \\ r \leq |u_{\sigma,+}| \leq n \\ r \leq |u_{\sigma,-}|}} \frac{|\sigma|}{d_\sigma} [T_n(u_{\sigma,-}) - T_n(u_{\sigma,+})]^2 \right)^{\frac{1}{2}},
\end{aligned}$$

and using again the Discrete Sobolev inequality

$$|I_2| \leq \frac{d^{\frac{1}{2}} \|\mathbf{v}\|_{L^4(E_r)^d}}{n} \sum_{\substack{\sigma \in \mathcal{A} \\ r \leq |u_{\sigma,+}| \leq n \\ r \leq |u_{\sigma,-}|}} \frac{|\sigma|}{d_\sigma} (T_n(u_K) - T_n(u_L))^2. \quad (1.26)$$

By Corollary 1.3.2 and since  $\mathbf{v} \in L^p(\Omega)^d$  ( $2 < p < +\infty$  if  $d = 2$ ,  $p = d$  if  $d \geq 3$ ), the absolute continuity of the integral implies that there exists  $r > 0$  (independent of  $\mathcal{T}$ ) such that for all admissible mesh  $\mathcal{T}$

$$d^{\frac{1}{2}} \|\mathbf{v}\|_{L^p(E_r)^d} \leq \frac{1}{2}. \quad (1.27)$$

Then from (1.25) and (1.27) we obtain

$$|I_2| \leq \frac{1}{2n} \sum_{\substack{\sigma \in \mathcal{A} \\ r \leq |u_{\sigma,+}| \leq n \\ r \leq |u_{\sigma,-}|}} \frac{|\sigma|}{d_\sigma} (u_K - u_L)(T_n(u_K) - T_n(u_L)). \quad (1.28)$$

We now turn to  $I_1$ . Using the Cauchy-Schwarz inequality and the Young inequality we deduce that

$$\begin{aligned} |I_1| &= \frac{1}{n} \sum_{\substack{\sigma \in \mathcal{A} \\ |u_{\sigma,+}| \leq r}} |\sigma| |\mathbf{v}_{K,\sigma}| u_{\sigma,+} (T_n(u_{\sigma,-}) - T_n(u_{\sigma,+})) \\ &\leq \frac{1}{n} \left( \sum_{\substack{\sigma \in \mathcal{A} \\ |u_{\sigma,+}| \leq r}} |\sigma| d_\sigma |\mathbf{v}_{K,\sigma}|^2 \right)^{\frac{1}{2}} \left( \sum_{\substack{\sigma \in \mathcal{A} \\ |u_{\sigma,+}| \leq r}} \frac{|\sigma|}{d_\sigma} u_{\sigma,+}^2 (T_n(u_{\sigma,-}) - T_n(u_{\sigma,+}))^2 \right)^{\frac{1}{2}} \\ &\leq \frac{r}{n} \left( \sum_{\substack{\sigma \in \mathcal{A} \\ |u_{\sigma,+}| \leq r}} |\sigma| d_\sigma |\mathbf{v}_{K,\sigma}|^2 \right)^{\frac{1}{2}} \left( \sum_{\substack{\sigma \in \mathcal{A} \\ |u_{\sigma,+}| \leq r}} \frac{|\sigma|}{d_\sigma} (u_K - u_L)(T_n(u_K) - T_n(u_L)) \right)^{\frac{1}{2}} \\ &\leq \frac{1}{n} \left[ \frac{r^2 d \|\mathbf{v}\|_{L^2(\Omega)^d}^2}{2} + \frac{1}{2} \sum_{\substack{\sigma \in \mathcal{A} \\ |u_{\sigma,+}| \leq r}} \frac{|\sigma|}{d_\sigma} (u_K - u_L)(T_n(u_K) - T_n(u_L)) \right]. \quad (1.29) \end{aligned}$$

Gathering (1.28) and (1.29) we get

$$\begin{aligned} -T_2 \leq |I_1 + I_2| &\leq \frac{1}{n} \frac{r^2 d \|\mathbf{v}\|_{L^2(\Omega)^d}^2}{2} + \frac{1}{2n} \sum_{\substack{\sigma \in \mathcal{A} \\ |u_{\sigma,+}| \leq n}} \frac{|\sigma|}{d_\sigma} (u_K - u_L)(T_n(u_K) - T_n(u_L)) \\ &\leq \frac{1}{n} \frac{r^2 d \|\mathbf{v}\|_{L^2(\Omega)^d}^2}{2} + \frac{1}{2} T_1, \quad (1.30) \end{aligned}$$

and we deduce that

$$\lim_{n \rightarrow +\infty} \overline{\lim}_{h_{\mathcal{T}} \rightarrow 0} \frac{1}{n} \sum_{\sigma \in \mathcal{E}} \frac{|\sigma|}{d_\sigma} (u_K - u_L)(T_n(u_K) - T_n(u_L)) = 0.$$

As far as (1.16) is concerned, using some computations of Step 1, we get

$$T_1 \leq -T_2 + T_4 \leq T_2' + T_4 \leq \frac{1}{2} T_1 + R + T_4 \text{ with}$$

$$T_2' = \frac{1}{n} \sum_{\sigma \in \mathcal{A}} |\sigma| |\mathbf{v}_{K,\sigma}| u_{\sigma,+} (T_n(u_{\sigma,-}) - T_n(u_{\sigma,+})),$$

$$\lim_{n \rightarrow +\infty} \overline{\lim}_{h_{\mathcal{T}} \rightarrow 0} R = 0,$$

$$\lim_{n \rightarrow +\infty} \overline{\lim}_{h_{\mathcal{T}} \rightarrow 0} T_4 = 0.$$

By defining the set  $\mathcal{B} = \{\sigma \in \mathcal{E}; u_{\sigma,+}(T_n(u_{\sigma,+}) - T_n(u_{\sigma,-})) > 0\}$  we have

$$\begin{aligned} T_2 &= \frac{1}{n} \sum_{\sigma \in \mathcal{A}} |\sigma| |\mathbf{v}_{K,\sigma}| u_{\sigma,+}(T_n(u_{\sigma,+}) - T_n(u_{\sigma,-})) \\ &\quad + \frac{1}{n} \sum_{\sigma \in \mathcal{B}} |\sigma| |\mathbf{v}_{K,\sigma}| u_{\sigma,+}(T_n(u_{\sigma,+}) - T_n(u_{\sigma,-})) \\ &= T_{2,1} + T_{2,2}. \end{aligned}$$

Since  $T_n$  is a non decreasing function and since  $T_n(0) = 0$

$$\begin{aligned} |T_{2,1}| &= -T_{2,1} = -\frac{1}{n} \sum_{\sigma \in \mathcal{A}} |\sigma| |\mathbf{v}_{K,\sigma}| u_{\sigma,+}(T_n(u_{\sigma,+}) - T_n(u_{\sigma,-})), \\ |T_{2,2}| &= T_{2,2} = \frac{1}{n} \sum_{\sigma \in \mathcal{B}} |\sigma| |\mathbf{v}_{K,\sigma}| u_{\sigma,+}(T_n(u_{\sigma,+}) - T_n(u_{\sigma,-})), \end{aligned}$$

and  $|T_2| = -T_{2,1} + T_{2,2}$ . From (1.30) and (1.15) we know that

$$\lim_{n \rightarrow +\infty} \overline{\lim}_{h_{\mathcal{T}} \rightarrow 0} |T_{2,1}| = 0. \quad (1.31)$$

Recalling that  $T_1 + T_2 + T_3 = T_4$  we obtain

$$T_1 + T_{2,2} \leq |T_{2,1}| + T_4$$

and (1.15), (1.21) and (1.31) allow one to conclude that

$$\lim_{n \rightarrow +\infty} \overline{\lim}_{h_{\mathcal{T}} \rightarrow 0} T_{2,2} = 0,$$

which gives (1.16). □

The following corollary is necessary to pass to the limit in the diffusion term.

**Corollary 1.3.5** *Let  $(\mathcal{T}_m)_{m \geq 1}$  be a sequence of admissible meshes such that there exists  $\xi > 0$  satisfying*

$$\text{for all } m \geq 1 \text{ for all } K \in \mathcal{T} \text{ and all } \sigma \in \mathcal{E}(K), d_{K,\sigma} \geq \xi d_\sigma.$$

*If  $u_{\mathcal{T}_m} = (u_K)_{K \in \mathcal{T}_m}$  is a solution to (1.7), then*

$$\lim_{n \rightarrow +\infty} \overline{\lim}_{h_{\mathcal{T}} \rightarrow 0} \sum_{\substack{\sigma \in \mathcal{E} \\ |u_K| \leq 2n \\ |u_L| > 4n}} \frac{|\sigma|}{d_\sigma} |u_L| = 0. \quad (1.32)$$

**Proof** Thanks to the discrete estimate on the energy, we see that

$$\lim_{n \rightarrow +\infty} \lim_{h_{\mathcal{T}} \rightarrow 0} \frac{1}{4n} \sum_{\sigma \in \mathcal{E}} \frac{|\sigma|}{d_\sigma} (u_K - u_L)(T_{4n}(u_K) - T_{4n}(u_L)) = 0.$$

If  $u_L > 4n$  and  $|u_K| \leq 2n$

$$(u_K - u_L)(T_{4n}(u_K) - T_{4n}(u_L)) \geq \frac{u_L}{2} 2n \geq 0,$$

and if  $u_L < -4n$  and  $|u_K| \leq 2n$

$$(u_K - u_L)(T_{4n}(u_K) - T_{4n}(u_L)) \geq \frac{-u_L}{2} 2n \geq 0.$$

We conclude that

$$\sum_{\substack{\sigma \in \mathcal{E} \\ |u_K| \leq 2n \\ |u_L| > 4n}} \frac{|\sigma|}{d_\sigma} |u_L| \leq \frac{1}{n} \sum_{\sigma \in \mathcal{E}} \frac{|\sigma|}{d_\sigma} (u_K - u_L)(T_{4n}(u_K) - T_{4n}(u_L))$$

and then from (1.15) we get (1.32).  $\square$

## 1.4 Convergence analysis

Let us now state the main result of this paper.

**Theorem 1.4.1** *If  $\mathcal{T}$  is an admissible mesh, then there exists a unique solution to (1.7).*

*If  $(\mathcal{T}_m)_{m \geq 1}$  is a sequence of admissible meshes such that there exists  $\xi > 0$  satisfying*

$$\text{for all } m \geq 1 \text{ for all } K \in \mathcal{T} \text{ and all } \sigma \in \mathcal{E}(K), d_{K,\sigma} \geq \xi d_\sigma,$$

*and such that  $h_{\mathcal{T}_m} \rightarrow 0$ , then if  $u_{\mathcal{T}_m} = (u_K)_{K \in \mathcal{T}_m}$  is the solution to (1.7) with  $\mathcal{T} = \mathcal{T}_m$ ,  $u_{\mathcal{T}_m}$  converges to  $u$  in the sense that for all  $n > 0$ ,  $T_n(u_{\mathcal{T}_m})$  converges weakly to  $T_n(u)$  in  $H_0^1(\Omega)$ , where  $u$  is the unique renormalized solution of (1.1).*

**Remark 1.4.2** *When  $v \in C(\bar{\Omega})^d$  the authors prove in [40] that  $u_{\mathcal{T}_m}$  converges to  $u$  where  $u$  is the unique weak solution of (1.1) in the sense of (1.2). It is well known, for this linear problem, that this weak solution and the renormalized solution coincide. A continuity argument then can give a similar result as Theorem 1.4.1. In the context of finite element method, for the Dirichlet problem  $-\operatorname{div}(A\nabla u) = f$  with  $f \in L^1$ , in [27] the authors prove that the discrete solution converges to the unique renormalized solution. As in [40, 46] the authors use the discrete and continuous Boccardo-Gallouët estimates and a duality argument to show that the limit of the discrete solution which is a weak solution coincides with the unique renormalized solution. We present here a different method which can be generalized to some nonlinear problem (see Section 5) for which the nonlinearity can be an obstacle to derive the uniqueness of the weak solution (in the sense of (1.2)).*

Before proving Theorem 1.4.1, we have to prove the following convergence result concerning the function  $h_n$  defined, for any  $n \geq 1$ , by

$$h_n(s) = \begin{cases} 0 & \text{if } s \leq -2n, \\ \frac{s}{n} + 2 & \text{if } -2n \leq s \leq -n, \\ 1 & \text{if } -n \leq s \leq n, \\ \frac{-s}{n} + 2 & \text{if } n \leq s \leq 2n, \\ 0 & \text{if } s \geq 2n. \end{cases} \quad (1.33)$$

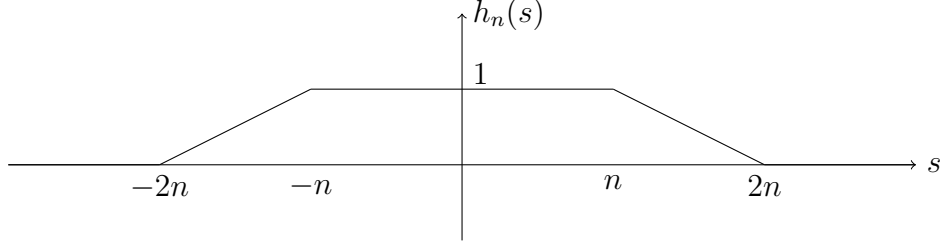


FIGURE 1.3 – The function  $h_n$

**Lemma 1.4.3** *Let  $(\mathcal{T}_m)_{m \geq 1}$  be a sequence of admissible meshes such that there exists  $\xi > 0$  satisfying*

$$\text{for all } m \geq 1 \text{ for all } K \in \mathcal{T} \text{ and all } \sigma \in \mathcal{E}(K), d_{K,\sigma} \geq \xi d_\sigma.$$

*Let  $u_{\mathcal{T}_m} \in X(\mathcal{T}_m)$  be a sequence of solution of (1.7). We define the function  $\tilde{h}_n$  by*

$$\forall \sigma \in \mathcal{E}, \forall x \in D_\sigma, \tilde{h}_n(x) = \frac{h_n(u_K) + h_n(u_L)}{2},$$

*then  $\tilde{h}_n \rightarrow h_n(u)$  in  $L^q(\Omega) \forall q \in [2, +\infty[$  as  $h_{\mathcal{T}_m} \rightarrow 0$ , where  $u$  is the limit of  $u_{\mathcal{T}_m}$ .*

**Proof** Since  $u_{\mathcal{T}_m}$  converges, up to a subsequence, to  $u$  a.e. in  $\Omega$ ,  $h_n(u_{\mathcal{T}_m}) \rightarrow h_n(u)$  a.e. in  $\Omega$ , thus for  $q \geq 2$

$$\begin{aligned} \|\tilde{h}_n - h_n(u_{\mathcal{T}_m})\|_{L^q(\Omega)}^q &= \int_{\Omega} |\tilde{h}_n(x) - h_n(u_{\mathcal{T}_m}(x))|^q dx \\ &= \sum_{\sigma \in \mathcal{E}} \int_{D_\sigma} |\tilde{h}_n(x) - h_n(u_{\mathcal{T}_m}(x))|^q dx \\ &= \sum_{\sigma \in \mathcal{E}} \int_{D_\sigma} \left| \frac{h_n(u_K) + h_n(u_L)}{2} - h_n(u_{\mathcal{T}_m}(x)) \right|^q dx \\ &= \frac{1}{2^q} \sum_{\sigma \in \mathcal{E}} |D_\sigma| |h_n(u_K) - h_n(u_L)|^q \\ &\leq \frac{1}{4} \sum_{\sigma \in \mathcal{E}} |D_\sigma| |h_n(u_K) - h_n(u_L)|^2 \\ &\leq \frac{1}{4} \sum_{\sigma \in \mathcal{E}} |D_\sigma| |T_{2n}(u_K) - T_{2n}(u_L)|^2 \\ &= \frac{1}{4d} \sum_{\sigma \in \mathcal{E}} |\sigma| d_\sigma \left| \frac{T_{2n}(u_K) - T_{2n}(u_L)}{d_\sigma} \right|^2 (d_\sigma)^2 \\ &\leq \frac{1}{4d} \|T_{2n}(u_{\mathcal{T}_m})\|_{1,2,\mathcal{T}}^2 (h_{\mathcal{T}_m})^2, \end{aligned}$$

and the proof is complete. □

#### Proof of Theorem 1.4.1

The proof consists into two steps. In Step 1 we prove the existence and the uniqueness



of the solution of (1.7). Concerning the uniqueness of the renormalized solution, the proof is done in [10]. To prove the second point, we adapt some uniqueness techniques developed in the continuous case in [26, 49]. It is worth noting that we use here a different method to the one developed in [40]. Using the results of Section 3, Step 2 is devoted to pass to the limit in the scheme. It is worth noting that we take in the scheme a discrete version of what is a test function in the renormalized formulation.

### Existence and uniqueness of the solution of the scheme

Since the scheme is a linear system of  $n$  equations with  $n$  unknowns, it is sufficient to show that the solution of the scheme is null with  $f = 0$ .

For  $\varepsilon > 0$ , let  $\varphi_\varepsilon(s) = \int_0^s \frac{dt}{(\varepsilon+|t|)^2}$ . Taking  $\varphi_\varepsilon(u_{\mathcal{T}})$  as a test function in the scheme with  $f = 0$  and reordering the sums yield

$$\begin{aligned} \sum_{\sigma \in \mathcal{E}} \frac{|\sigma|}{d_\sigma} (u_K - u_L) (\varphi_\varepsilon(u_K) - \varphi_\varepsilon(u_L)) + \sum_{\sigma \in \mathcal{E}} |\sigma| |\mathbf{v}_{K,\sigma}| u_{\sigma,+} (\varphi_\varepsilon(u_{\sigma,+}) - \varphi_\varepsilon(u_{\sigma,-})) \\ + \sum_{K \in \mathcal{T}} |K| b_K u_K \varphi_\varepsilon(u_K) = 0. \end{aligned}$$

Since  $b$  is nonnegative and  $\varphi_\varepsilon(s)$  has the same sign as  $s$ ,

$$\sum_{K \in \mathcal{T}} |K| b_K u_K \varphi_\varepsilon(u_K) \geq 0,$$

hence

$$\sum_{\sigma \in \mathcal{E}} \frac{|\sigma|}{d_\sigma} (u_K - u_L) (\varphi_\varepsilon(u_K) - \varphi_\varepsilon(u_L)) \leq \sum_{\sigma \in \mathcal{E}} |\sigma| |\mathbf{v}_{K,\sigma}| u_{\sigma,+} (\varphi_\varepsilon(u_{\sigma,-}) - \varphi_\varepsilon(u_{\sigma,+})).$$

As in the proof of estimate (1.15) we use the set  $\mathcal{A} = \{\sigma \in \mathcal{E}; u_{\sigma,+} \geq u_{\sigma,-}, u_{\sigma,+} < 0\} \cup \{\sigma \in \mathcal{E}; u_{\sigma,+} < u_{\sigma,-}, u_{\sigma,+} \geq 0\}$ . Since  $\varphi_\varepsilon$  is an increasing function

$$\sum_{\sigma \in \mathcal{E}} |\sigma| |\mathbf{v}_{K,\sigma}| u_{\sigma,+} (\varphi_\varepsilon(u_{\sigma,-}) - \varphi_\varepsilon(u_{\sigma,+})) \leq \sum_{\sigma \in \mathcal{A}} |\sigma| |\mathbf{v}_{K,\sigma}| u_{\sigma,+} (\varphi_\varepsilon(u_{\sigma,-}) - \varphi_\varepsilon(u_{\sigma,+})).$$

Consequently

$$\begin{aligned} & \sum_{\sigma \in \mathcal{E}} \frac{|\sigma|}{d_\sigma} (u_K - u_L) (\varphi_\varepsilon(u_K) - \varphi_\varepsilon(u_L)) \\ & \leq \sum_{\sigma \in \mathcal{A}} |\sigma| |\mathbf{v}_{K,\sigma}| u_{\sigma,+} (\varphi_\varepsilon(u_{\sigma,-}) - \varphi_\varepsilon(u_{\sigma,+})) \\ & \leq \left( \sum_{\sigma \in \mathcal{A}} |\sigma| d_\sigma |\mathbf{v}_{K,\sigma}|^2 \right)^{\frac{1}{2}} \left( \sum_{\sigma \in \mathcal{A}} \frac{|\sigma|}{d_\sigma} u_{\sigma,+}^2 (\varphi_\varepsilon(u_{\sigma,-}) - \varphi_\varepsilon(u_{\sigma,+}))^2 \right)^{\frac{1}{2}} \\ & \leq \left( \sum_{\sigma \in \mathcal{A}} d |D_\sigma| |\mathbf{v}_{K,\sigma}|^2 \right)^{\frac{1}{2}} \left( \sum_{\sigma \in \mathcal{A}} \frac{|\sigma|}{d_\sigma} u_{\sigma,+}^2 (\varphi_\varepsilon(u_{\sigma,-}) - \varphi_\varepsilon(u_{\sigma,+}))^2 \right)^{\frac{1}{2}} \\ & \leq d^{\frac{1}{2}} \|\mathbf{v}\|_{L^2(\Omega)^d} \left( \sum_{\sigma \in \mathcal{A}} \frac{|\sigma|}{d_\sigma} u_{\sigma,+}^2 (\varphi_\varepsilon(u_{\sigma,-}) - \varphi_\varepsilon(u_{\sigma,+}))^2 \right)^{\frac{1}{2}}. \end{aligned}$$

In order to deal with this last term, we use a similar argument as in [40]. Noticing that  $\varphi_\varepsilon$  is  $\mathcal{C}^1$ -continuous on  $\mathbb{R}$ , there exists  $w_\sigma \in [u_{\sigma,+}, u_{\sigma,-}]$  such that

$$\begin{aligned} u_{\sigma,+}^2 (\varphi_\varepsilon(u_{\sigma,-}) - \varphi_\varepsilon(u_{\sigma,+}))^2 &\leq u_{\sigma,+}^2 \varphi'_\varepsilon(w_\sigma) (u_{\sigma,-} - u_{\sigma,+}) (\varphi_\varepsilon(u_{\sigma,-}) - \varphi_\varepsilon(u_{\sigma,+})) \\ &\leq \frac{u_{\sigma,+}^2}{(\varepsilon + |w_\sigma|)^2} (u_{\sigma,-} - u_{\sigma,+}) (\varphi_\varepsilon(u_{\sigma,-}) - \varphi_\varepsilon(u_{\sigma,+})). \end{aligned}$$

Recalling that  $\forall \sigma \in \mathcal{A}$ ,  $|u_{\sigma,+}| \leq |u_{\sigma,-}|$  and that  $u_{\sigma,+}$  and  $u_{\sigma,-}$  have the same sign, since  $w_\sigma \in [u_{\sigma,+}, u_{\sigma,-}]$ , one deduces that  $\frac{u_{\sigma,+}^2}{(\varepsilon + |w_\sigma|)^2} \leq 1$ . Therefore

$$\begin{aligned} &\sum_{\sigma \in \mathcal{E}} \frac{|\sigma|}{d_\sigma} (u_K - u_L) (\varphi_\varepsilon(u_K) - \varphi_\varepsilon(u_L)) \\ &\leq d^{\frac{1}{2}} \|\mathbf{v}\|_{L^2(\Omega)^d} \left( \sum_{\sigma \in \mathcal{A}} \frac{|\sigma|}{d_\sigma} (u_{\sigma,-} - u_{\sigma,+}) (\varphi_\varepsilon(u_{\sigma,-}) - \varphi_\varepsilon(u_{\sigma,+})) \right)^{\frac{1}{2}} \\ &\leq \frac{d \|\mathbf{v}\|_{L^2(\Omega)^d}^2}{2} + \frac{1}{2} \sum_{\sigma \in \mathcal{E}} \frac{|\sigma|}{d_\sigma} (u_K - u_L) (\varphi_\varepsilon(u_K) - \varphi_\varepsilon(u_L)). \end{aligned}$$

We conclude that  $\forall \varepsilon > 0$

$$\sum_{\sigma \in \mathcal{E}} \frac{|\sigma|}{d_\sigma} (u_K - u_L) (\varphi_\varepsilon(u_K) - \varphi_\varepsilon(u_L)) \leq d \|\mathbf{v}\|_{L^2(\Omega)^d}^2.$$

Moreover, for all  $(x, y) \in \mathbb{R}^2$

$$\begin{aligned} \left( \ln\left(1 + \frac{|x|}{\varepsilon}\right) - \ln\left(1 + \frac{|y|}{\varepsilon}\right) \right)^2 &= \left( \int_{|y|}^{|x|} \frac{dt}{\varepsilon + |t|} \right)^2 \\ &\leq |x - y| \left| \int_{|y|}^{|x|} \frac{dt}{(\varepsilon + |t|)^2} \right| \\ &\leq |x - y| \int_y^x \frac{dt}{(\varepsilon + |t|)^2} \\ &\leq (x - y) (\varphi_\varepsilon(x) - \varphi_\varepsilon(y)). \end{aligned}$$

It follows that

$$\sum_{\sigma \in \mathcal{E}} \frac{|\sigma|}{d_\sigma} \left( \ln\left(1 + \frac{|u_K|}{\varepsilon}\right) - \ln\left(1 + \frac{|u_L|}{\varepsilon}\right) \right)^2 \leq d \|\mathbf{v}\|_{L^2(\Omega)^d}^2,$$

so that  $\|\ln(1 + \frac{|u_\mathcal{T}|}{\varepsilon})\|_{1,2,\mathcal{T}}^2 \leq d \|\mathbf{v}\|_{L^2(\Omega)^d}^2$  for any  $\varepsilon > 0$ .

Letting  $\varepsilon \rightarrow 0$ , we deduce that  $u_\mathcal{T} = 0$  if  $f = 0$ . (1.7) being a linear system, we have proved that there exists a unique solution of the scheme.

### Convergence

Let  $\varphi \in \mathcal{C}_c^\infty(\Omega)$  et  $h_n$  the function defined by (1.33). We denote by  $\varphi_\mathcal{T}$  the function defined by  $\varphi_K = \varphi(x_K)$  for all  $K \in \mathcal{T}$ .

Multiplying each equation of the scheme by  $\varphi(x_K)h_n(u_K)$  (which is a discrete version of the test function used in the renormalized formulation), summing over the control volumes and gathering by edges, we get  $T_1 + T_2 + T_3 = T_4$  with

$$\begin{aligned} T_1 &= \sum_{\sigma \in \mathcal{E}} \frac{|\sigma|}{d_\sigma} (u_K - u_L) (\varphi(x_K)h_n(u_K) - \varphi(x_L)h_n(u_L)), \\ T_2 &= \sum_{\sigma \in \mathcal{E}} |\sigma| \mathbf{v}_{K,\sigma} u_{\sigma,+} (\varphi(x_K)h_n(u_K) - \varphi(x_L)h_n(u_L)), \\ T_3 &= \sum_{K \in \mathcal{T}} |K| b_K u_K \varphi(x_K) h_n(u_K), \\ T_4 &= \sum_{K \in \mathcal{T}} \int_K f \varphi(x_K) h_n(u_K). \end{aligned}$$

As far as the lower order term  $T_4$  is concerned, by the regularity of  $\varphi$ , we have  $\varphi_{\mathcal{T}} \rightarrow \varphi$  uniformly on  $\Omega$  as  $h_{\mathcal{T}} \rightarrow 0$ . We now pass to the limit as  $h_{\mathcal{T}} \rightarrow 0$ . Since  $h_n(u_{\mathcal{T}}) \rightarrow h_n(u)$  a.e. and  $L^\infty$  weak  $\star$ ,  $\varphi_{\mathcal{T}} \rightarrow \varphi$  uniformly and  $|f \varphi_{\mathcal{T}} h_n(u_{\mathcal{T}})| \leq C_\varphi |f| \in L^1(\Omega)$ , the Lebesgue theorem ensures that

$$T_4 = \int_{\Omega} f \varphi_{\mathcal{T}} h_n(u_{\mathcal{T}}) dx \xrightarrow{h_{\mathcal{T}} \rightarrow 0} \int_{\Omega} f \varphi h_n(u) dx. \quad (1.34)$$

In view of the definition of  $b_{\mathcal{T}}$ , since  $b$  belongs to  $L^2(\Omega)$ ,  $b_{\mathcal{T}} = (b_K)_{K \in \mathcal{T}}$  converges to  $b$  in  $L^2(\Omega)$  as  $h_{\mathcal{T}} \rightarrow 0$ . With already used arguments we can assert that

$$T_3 = \int_{\Omega} b_{\mathcal{T}} T_{2n}(u_{\mathcal{T}}) \varphi_{\mathcal{T}} h_n(u_{\mathcal{T}}) dx \xrightarrow{h_{\mathcal{T}} \rightarrow 0} \int_{\Omega} b T_{2n}(u) \varphi h_n(u) dx. \quad (1.35)$$

We now study the convergence of the diffusion term. We write

$$\begin{aligned} T_1 &= \sum_{\sigma \in \mathcal{E}} \frac{|\sigma|}{d_\sigma} (u_K - u_L) (\varphi(x_K)h_n(u_K) - \varphi(x_L)h_n(u_L)) \\ &= T_{1,1} + T_{1,2} \end{aligned}$$

with

$$\begin{aligned} T_{1,1} &= \sum_{\sigma \in \mathcal{E}} \frac{|\sigma|}{d_\sigma} h_n(u_K) (u_K - u_L) (\varphi(x_K) - \varphi(x_L)), \\ T_{1,2} &= \sum_{\sigma \in \mathcal{E}} \frac{|\sigma|}{d_\sigma} \varphi(x_L) (u_K - u_L) (h_n(u_K) - h_n(u_L)). \end{aligned}$$

In view of the definition of  $h_n$  we get

$$|T_{1,2}| \leq \frac{1}{n} \sum_{\sigma \in \mathcal{E}} \frac{|\sigma|}{d_\sigma} \varphi(x_L) (u_K - u_L) (T_{2n}(u_K) - T_{2n}(u_L)).$$

From the discrete estimate on the energy (see (1.15) in Proposition 1.3.4), it may be concluded that

$$\lim_{n \rightarrow +\infty} \overline{\lim}_{h_{\mathcal{T}} \rightarrow 0} T_{1,2} = 0. \quad (1.36)$$

As far as  $T_{1,1}$  is concerned we observe that  $u_K$  is truncated while  $u_L$  is not truncated. To deal with  $T_{1,1}$  we write

$$\begin{aligned}
T_{1,1} &= \sum_{\sigma \in \mathcal{E}} \frac{|\sigma|}{d_\sigma} h_n(u_K) (T_{2n}(u_K) - u_L) (\varphi(x_K) - \varphi(x_L)) \\
&= \underbrace{\sum_{\sigma \in \mathcal{E}} \frac{|\sigma|}{d_\sigma} \frac{h_n(u_K) + h_n(u_L)}{2} (T_{4n}(u_K) - T_{4n}(u_L)) (\varphi(x_K) - \varphi(x_L))}_I \\
&\quad + \underbrace{\sum_{\sigma \in \mathcal{E}} \frac{|\sigma|}{d_\sigma} h_n(u_K) (T_{4n}(u_L) - u_L) (\varphi(x_K) - \varphi(x_L))}_{II} \\
&\quad + \underbrace{\sum_{\sigma \in \mathcal{E}} \frac{|\sigma|}{d_\sigma} \frac{h_n(u_K) - h_n(u_L)}{2} (T_{4n}(u_K) - T_{4n}(u_L)) (\varphi(x_K) - \varphi(x_L))}_{III}.
\end{aligned}$$

Since in  $I$   $u_K$  and  $u_L$  are both truncated we can pass to the limit in  $I$  by writing

$$\begin{aligned}
I &= \sum_{\sigma \in \mathcal{E}} |\sigma| d_\sigma \frac{h_n(u_K) + h_n(u_L)}{2} \frac{T_{2n}(u_K) - T_{4n}(u_L)}{d_\sigma} \frac{\varphi(x_K) - \varphi(x_L)}{d_\sigma} \\
&= \underbrace{\sum_{\sigma \in \mathcal{E}} \int_{D_\sigma} \tilde{h}_n(x) \nabla_{\mathcal{T}} T_{4n}(u_{\mathcal{T}}) \cdot \nabla \varphi \, dx}_{I_1} \\
&\quad + \underbrace{\sum_{\sigma \in \mathcal{E}} d |D_\sigma| \frac{h_n(u_K) + h_n(u_L)}{2} \frac{T_{4n}(u_K) - T_{4n}(u_L)}{d_\sigma} \left[ \frac{\varphi(x_K) - \varphi(x_L)}{d_\sigma} - \frac{1}{|D_\sigma|} \int_{D_\sigma} \nabla \varphi \cdot \eta_{K,\sigma} \, dx \right]}_{I_2}.
\end{aligned}$$

By Lemma 1.4.3  $\tilde{h}_n \rightarrow h_n(u)$  in  $L^q(\Omega)$  for all  $q \in [2, +\infty[$ , while  $\nabla_{\mathcal{T}} T_{4n}(u_{\mathcal{T}})$  tends to  $\nabla T_{4n}(u)$  weakly in  $(L^2(\Omega))^d$ . Since  $\varphi$  belongs to  $\mathcal{C}_c^\infty(\Omega)$  we conclude that

$$\lim_{h_{\mathcal{T}} \rightarrow 0} I_1 = \int_{\Omega} h_n(u) \nabla T_{4n}(u) \cdot \nabla \varphi \, dx. \quad (1.37)$$

By the regularity of  $\varphi$  we see that

$$\begin{aligned}
|I_2| &\leq \sum_{\sigma \in \mathcal{E}} d |D_\sigma| \left| \frac{h_n(u_K) + h_n(u_L)}{2} \right| \left| \frac{T_{4n}(u_K) - T_{4n}(u_L)}{d_\sigma} \right| \left| \frac{\varphi(x_K) - \varphi(x_L)}{d_\sigma} - \frac{1}{|D_\sigma|} \int_{D_\sigma} \nabla \varphi \cdot \eta_{K,\sigma} \, dx \right| \\
&\leq \|\varphi\|_{W^{2,\infty}(\Omega)} h_{\mathcal{T}} \|T_{4n}(u_{\mathcal{T}})\|_{1,1,\mathcal{T}},
\end{aligned}$$

thus

$$\lim_{h_{\mathcal{T}} \rightarrow 0} I_2 = 0. \quad (1.38)$$

We now turn to  $II$ .

$$\begin{aligned}
|II| &\leq \sum_{\sigma \in \mathcal{E}} \frac{|\sigma|}{d_\sigma} |h_n(u_K)| |T_{4n}(u_L) - u_L| |\varphi(x_K) - \varphi(x_L)| \\
&\leq 2 \|\varphi\|_{L^\infty(\Omega)} \sum_{\substack{\sigma \in \mathcal{E} \\ |u_K| \leq 2n \\ |u_L| > 4n}} \frac{|\sigma|}{d_\sigma} |u_L|.
\end{aligned}$$

We can deduce by Corollary 1.3.5 that

$$\lim_{n \rightarrow +\infty} \overline{\lim}_{h_T \rightarrow 0} II = 0. \quad (1.39)$$

Then, noticing that

$$|III| \leq \frac{\|\varphi\|_{L^\infty(\Omega)}}{n} \sum_{\sigma \in \mathcal{E}} \frac{|\sigma|}{d_\sigma} |u_K - u_L| |T_n(u_K) - T_n(u_L)|,$$

and using estimate (1.15), we deduce that

$$\lim_{n \rightarrow +\infty} \overline{\lim}_{h_T \rightarrow 0} III = 0. \quad (1.40)$$

For the convection term we have

$$\begin{aligned} T_2 &= \sum_{\sigma \in \mathcal{E}} |\sigma| \mathbf{v}_{K,\sigma} u_{\sigma,+} (\varphi(x_K) h_n(u_K) - \varphi(x_L) h_n(u_L)) \\ &= \sum_{\substack{\sigma \in \mathcal{E} \\ \mathbf{v}_{K,\sigma} \geq 0}} |\sigma| \mathbf{v}_{K,\sigma} u_{\sigma,+} (\varphi(x_K) h_n(u_{\sigma,+}) - \varphi(x_L) h_n(u_{\sigma,-})) \\ &\quad + \sum_{\substack{\sigma \in \mathcal{E} \\ \mathbf{v}_{K,\sigma} < 0}} |\sigma| \mathbf{v}_{K,\sigma} u_{\sigma,+} (\varphi(x_K) h_n(u_{\sigma,-}) - \varphi(x_L) h_n(u_{\sigma,+})) \\ &= \sum_{\substack{\sigma \in \mathcal{E} \\ \mathbf{v}_{K,\sigma} \geq 0}} |\sigma| \mathbf{v}_{K,\sigma} u_{\sigma,+} h_n(u_{\sigma,+}) (\varphi(x_K) - \varphi(x_L)) \\ &\quad + \sum_{\substack{\sigma \in \mathcal{E} \\ \mathbf{v}_{K,\sigma} \geq 0}} |\sigma| \mathbf{v}_{K,\sigma} u_{\sigma,+} \varphi(x_L) (h_n(u_{\sigma,+}) - h_n(u_{\sigma,-})) \\ &\quad - \sum_{\substack{\sigma \in \mathcal{E} \\ \mathbf{v}_{K,\sigma} < 0}} |\sigma| \mathbf{v}_{K,\sigma} u_{\sigma,+} h_n(u_{\sigma,+}) (\varphi(x_L) - \varphi(x_K)) \\ &\quad - \sum_{\substack{\sigma \in \mathcal{E} \\ \mathbf{v}_{K,\sigma} < 0}} |\sigma| \mathbf{v}_{K,\sigma} u_{\sigma,+} \varphi(x_K) (h_n(u_{\sigma,+}) - h_n(u_{\sigma,-})) \\ &= T_{2,1} + T_{2,2} + T_{2,3} \end{aligned}$$

with

$$\begin{aligned} T_{2,1} &= \sum_{\sigma \in \mathcal{E}} |\sigma| \mathbf{v}_{K,\sigma} u_{\sigma,+} h_n(u_{\sigma,+}) (\varphi(x_K) - \varphi(x_L)) \\ T_{2,2} &= \sum_{\substack{\sigma \in \mathcal{E} \\ \mathbf{v}_{K,\sigma} \geq 0}} |\sigma| \mathbf{v}_{K,\sigma} u_{\sigma,+} \varphi(x_L) (h_n(u_{\sigma,+}) - h_n(u_{\sigma,-})) \\ T_{2,3} &= - \sum_{\substack{\sigma \in \mathcal{E} \\ \mathbf{v}_{K,\sigma} < 0}} |\sigma| \mathbf{v}_{K,\sigma} u_{\sigma,+} \varphi(x_K) (h_n(u_{\sigma,+}) - h_n(u_{\sigma,-})) \end{aligned}$$

Since

$$|T_{2,2} + T_{2,3}| \leq \frac{\|\varphi\|_{L^\infty(\Omega)}}{n} \sum_{\sigma \in \mathcal{E}} |\sigma| |\mathbf{v}_{K,\sigma}| |u_{\sigma,+}| |T_{2n}(u_{\sigma,+}) - T_{2n}(u_{\sigma,-})|,$$

using (1.16) we show that

$$\lim_{n \rightarrow +\infty} \overline{\lim}_{h_{\mathcal{T}} \rightarrow 0} T_{2,2} + T_{2,3} = 0. \quad (1.41)$$

Concerning  $T_{2,1}$  we obtain

$$\begin{aligned} T_{2,1} &= \sum_{\sigma \in \mathcal{E}} |\sigma| \mathbf{v}_{K,\sigma} u_{\sigma,+} h_n(u_{\sigma,+}) (\varphi(x_K) - \varphi(x_L)) \\ &= \sum_{\sigma \in \mathcal{E}} \frac{|\sigma| d_\sigma}{d|D_\sigma|} u_{\sigma,+} h_n(u_{\sigma,+}) d \frac{\varphi(x_K) - \varphi(x_L)}{d_\sigma} \int_{D_\sigma} \mathbf{v} \cdot \eta_{K,\sigma} d\gamma \\ &= - \sum_{\sigma \in \mathcal{E}} \int_{D_\sigma} T_{2n}(u_{\sigma,+}) h_n(u_{\sigma,+}) \mathbf{v} \cdot \nabla_{\mathcal{T}} \varphi_{\mathcal{T}} dx. \end{aligned}$$

Let  $\forall \sigma \in \mathcal{E}, \forall x \in D_\sigma, \overline{T_{2n}(u)h_n(u)}^{\mathcal{T}_m}(x) = T_{2n}(u_{\sigma,+}) h_n(u_{\sigma,+})$ , we can rewrite

$$T_{2,1} = - \int_{\Omega} \overline{T_{2n}(u)h_n(u)}^{\mathcal{T}_m} \mathbf{v} \cdot \nabla_{\mathcal{T}} \varphi_{\mathcal{T}} dx.$$

By Lemma 1.2.6,  $\nabla_{\mathcal{T}} \varphi_{\mathcal{T}}$  converges weakly to  $\nabla \varphi$  in  $L^\alpha(\Omega)^d \forall \alpha \in [1, +\infty[$ , and then in  $L^{p'}(\Omega)$ . Moreover, with a proof similar to the one of Lemma 1.4.3, we can show that  $\overline{T_{2n}(u)h_n(u)}^{\mathcal{T}_m}$  tends to  $T_{2n}(u)h_n(u)$  in  $L^\infty$  weak  $\star$  as  $h_{\mathcal{T}} \rightarrow 0$ , therefore

$$\lim_{h_{\mathcal{T}} \rightarrow 0} T_{2,1} = - \int_{\Omega} T_{2n}(u) h_n(u) \mathbf{v} \cdot \nabla \varphi dx. \quad (1.42)$$

We are now in position to pass to the limit as  $h_{\mathcal{T}} \rightarrow 0$  in the scheme. Gathering equations (1.34) to (1.42), we can assert that

$$\begin{aligned} \int_{\Omega} h_n(u) \nabla u \cdot \nabla \varphi dx - \int_{\Omega} u h_n(u) \mathbf{v} \cdot \nabla \varphi dx + \int_{\Omega} b u h_n(u) \varphi dx \\ - \int_{\Omega} f \varphi h_n(u) dx = \overline{\lim}_{h_{\mathcal{T}} \rightarrow 0} T(n, \varphi) \end{aligned} \quad (1.43)$$

where  $\overline{\lim}_{h_{\mathcal{T}} \rightarrow 0} |T(n, \varphi)| \leq \|\varphi\|_{L^\infty(\Omega)} \omega(n)$  with  $\omega(n) \rightarrow 0$  as  $n \rightarrow +\infty$ .

Let  $h \in \mathcal{C}_c^1(\mathbb{R})$  and  $\psi \in \mathcal{C}_c^\infty(\Omega) \cap H_0^1(\Omega)$ . In view of the regularity of  $T_n(u)$  (see (1.3)) the function  $h(u)\psi$  belongs to  $L^\infty(\Omega) \cap H_0^1(\Omega)$ . By the density of  $\mathcal{C}_c^\infty(\Omega)$  in  $L^\infty(\Omega) \cap H_0^1(\Omega)$  (here any element of  $L^\infty(\Omega) \cap H_0^1(\Omega)$  can be approached by a sequence of  $\mathcal{C}_c^\infty(\Omega)$  which is bounded in  $L^\infty(\Omega)$ ) from (1.43) we deduce that

$$\begin{aligned} \left| \int_{\Omega} \nabla u h_n(u) h(u) \nabla \psi dx + \int_{\Omega} \nabla u h_n(u) \psi \nabla u h'(u) dx \right. \\ \left. - \int_{\Omega} u h_n(u) h(u) \mathbf{v} \cdot \nabla \psi dx - \int_{\Omega} u h_n(u) h'(u) \psi \mathbf{v} \cdot \nabla u dx \right. \\ \left. + \int_{\Omega} b u h_n(u) h(u) \psi dx - \int_{\Omega} \psi h(u) h_n(u) f dx \right| \\ \leq \|\varphi\|_{L^\infty(\Omega)} \omega(n). \end{aligned}$$

Passing to the limit as  $n \rightarrow +\infty$  in the previous inequality yields that

$$\begin{aligned} \int_{\Omega} \nabla u h(u) \nabla \psi \, dx + \int_{\Omega} \nabla u \psi \nabla u h'(u) \, dx \\ - \int_{\Omega} u h(u) \mathbf{v} \cdot \nabla \psi \, dx - \int_{\Omega} u h'(u) \psi \mathbf{v} \cdot \nabla u \, dx \\ + \int_{\Omega} b u h(u) \psi \, dx = \int_{\Omega} \psi h(u) f \, dx, \end{aligned}$$

which is Equality (1.5) in the definition of a renormalized solution.

It remains to prove that  $u$  satisfies the decay (1.4) of the truncate energy. Thanks to the discrete estimate on the energy (1.15) we get

$$\lim_{n \rightarrow \infty} \overline{\lim}_{h_{\mathcal{T}} \rightarrow 0} \frac{1}{n} \sum_{\sigma \in \mathcal{E}} \frac{|\sigma|}{d_{\sigma}} (T_{2n}(u_K) - T_{2n}(u_L))^2 = 0,$$

and

$$\begin{aligned} \sum_{\sigma \in \mathcal{E}} \frac{|\sigma|}{d_{\sigma}} (T_{2n}(u_K) - T_{2n}(u_L))^2 &= \sum_{\sigma \in \mathcal{E}} |\sigma| d_{\sigma} \left( \frac{T_{2n}(u_K) - T_{2n}(u_L)}{d_{\sigma}} \right)^2 \\ &= \sum_{\sigma \in \mathcal{E}} d |D_{\sigma}| \left( \frac{T_{2n}(u_K) - T_{2n}(u_L)}{d_{\sigma}} \right)^2 \\ &= \frac{1}{d} \sum_{\sigma \in \mathcal{E}} |D_{\sigma}| \left( d \frac{T_{2n}(u_K) - T_{2n}(u_L)}{d_{\sigma}} \right)^2 \\ &= \frac{1}{d} \int_{\Omega} |\nabla_{\mathcal{T}} T_{2n}(u_{\mathcal{T}})|^2 \, dx, \end{aligned}$$

hence,  $\lim_{n \rightarrow \infty} \overline{\lim}_{h_{\mathcal{T}} \rightarrow 0} \frac{1}{n} \int_{\Omega} |\nabla_{\mathcal{T}} T_{2n}(u_{\mathcal{T}})|^2 \, dx = 0$ . Since  $\nabla T_{2n}(u_{\mathcal{T}})$  converges weakly in  $L^2(\Omega)^d$ , we have also

$$\frac{1}{n} \int_{\Omega} |\nabla T_{2n}(u)|^2 \, dx \leq \liminf_{h_{\mathcal{T}} \rightarrow 0} \frac{1}{n} \int_{\Omega} |\nabla_{\mathcal{T}} T_{2n}(u_{\mathcal{T}})|^2 \, dx,$$

which leads to

$$\lim_{n \rightarrow \infty} \frac{1}{n} \int_{\Omega} |\nabla T_{2n}(u)|^2 \, dx = 0.$$

Since the renormalized solution  $u$  is unique we conclude that the whole sequence  $u_{\mathcal{T}_m}$  converges to  $u$  in the sense that for all  $n > 0$ ,  $T_n(u_{\mathcal{T}_m})$  converges weakly to  $T_n(u)$  in  $H_0^1(\Omega)$ .  $\square$

## 1.5 Concluding remark

We are able to show similar results for the nonlinear problem

$$\begin{aligned} -\operatorname{div}(\lambda(u) \nabla u) + \operatorname{div}(\mathbf{v}u) + bu &= f \text{ in } \Omega, \\ u &= 0 \text{ on } \partial\Omega, \end{aligned} \tag{1.44}$$

where  $\lambda$  is a continuous function such that  $\lambda_\infty \geq \lambda(u) \geq \mu > 0$  with  $\lambda_\infty$  and  $\mu$  two real numbers.

For  $\mathbf{v} \in (L^p(\Omega))^d$ ,  $b \in L^2(\Omega)$  ( $b \geq 0$ ) and  $f \in L^1(\Omega)$ , a solution in a weak sense similar to the one of (1.1) exists but due to the nonlinear character the methods developed in [37, 40] may fail to obtain the uniqueness of such a weak solution. On the contrary the renormalized solution exists and is unique (see [11]).

Our study is based on the same finite volume scheme as in [51]. This scheme writes

$$\forall K \in \mathcal{T}, \quad \sum_{\sigma \in \mathcal{E}(K)} \frac{|\sigma|}{d_\sigma} \lambda(u)_\sigma (u_K - u_L) + \sum_{\sigma \in \mathcal{E}(K)} |\sigma| \mathbf{v}_{K,\sigma} u_{\sigma,+} = \int_K f \, dx, \quad (1.45)$$

with  $u_L = 0$  if  $\sigma \in \mathcal{E}_{ext} \cap \mathcal{E}(K)$  and

$$\forall \sigma = K|L \in \mathcal{E}_{int}, \quad \min[\lambda(u_K), \lambda(u_L)] \leq \lambda(u)_\sigma \leq \max[\lambda(u_K), \lambda(u_L)], \quad (1.46)$$

$$\forall \sigma \in \mathcal{E}_{ext} \cap \mathcal{E}(K), \quad \lambda(u)_\sigma = \lambda(u_K), \quad (1.47)$$

where  $\lambda(u)_\sigma$  is for example the mean value of  $\lambda(u_K)$  and  $\lambda(u_L)$  if  $\sigma \in \mathcal{E}_{int}$ .

When  $\mathbf{v}$  has a null divergence the existence of a solution to (1.45) is proved in the context of a coupled system in [51] and the authors prove that the solution converges to a solution of (1.44) in a weak sense.

If  $\mathcal{T}$  is an admissible mesh and without assuming that  $\operatorname{div}(u) = 0$ , the Brouwer fixed point theorem and a log-type estimate similar to Proposition 1.3.1 allow one to prove that the scheme (1.45) admits at least one solution.

Indeed by replacing  $\lambda(u)_\sigma$  by  $\lambda(\bar{u})_\sigma$  where  $\bar{u}_\mathcal{T} \in X(\mathcal{T})$ , the assumptions on  $\lambda$  ( $0 < \mu \leq \lambda \leq \lambda_\infty$ ) and the arguments developed in the proof of Theorem 1.4.1 show that there exists a unique  $u_\mathcal{T} \in X(\mathcal{T})$  solution of

$$\forall K \in \mathcal{T}, \quad \sum_{\sigma \in \mathcal{E}(K)} \frac{|\sigma|}{d_\sigma} \lambda(\bar{u})_\sigma (u_K - u_L) + \sum_{\sigma \in \mathcal{E}(K)} |\sigma| \mathbf{v}_{K,\sigma} u_{\sigma,+} = \int_K f \, dx. \quad (1.48)$$

Moreover since  $0 < \mu \leq \lambda \leq \lambda_\infty$  we can derive a log type estimate similar to (1.10)

$$\|\ln(1 + |u_\mathcal{T}|)\|_{1,2,\mathcal{T}}^2 \leq c(\lambda) \left( 2\|f\|_{L^1(\Omega)} + d|\Omega|^{\frac{p-2}{p}} \|\mathbf{v}\|_{L^p(\Omega)^d}^2 \right),$$

where  $c(\lambda)$  is a positive constant depending on  $\mu$  and  $\lambda_\infty$ .

The Brouwer theorem allows one to conclude that (1.45) admits at least one solution.

We now turn to the convergence of the scheme (1.45) as  $h_\mathcal{T} \rightarrow 0$ . Since  $0 < \mu \leq \lambda \leq \lambda_\infty$ , if  $\mathcal{T}$  is an admissible mesh such that there exists  $\xi > 0$  satisfying

$$\text{for all } m \geq 1 \text{ for all } K \in \mathcal{T} \text{ and all } \sigma \in \mathcal{E}(K), \quad d_{K,\sigma} \geq \xi d_\sigma,$$

we can adapt all the results of Section 3 (the estimates) and in particular the decay of the discrete truncate energy which is crucial in the process.

As far as the convergence in the discrete renormalized version is concerned we have



to modify the study of  $I$  which reads as

$$I = \underbrace{\sum_{\sigma \in \mathcal{E}} \int_{D_\sigma} \tilde{\lambda}(x) \tilde{h}_n(x) \nabla_{\mathcal{T}} T_{4n}(u_{\mathcal{T}}) \cdot \nabla \varphi \, dx}_{I_1} + \underbrace{\sum_{\sigma \in \mathcal{E}} d|D_\sigma| \lambda(u)_\sigma h_n(u_K) \frac{T_{4n}(u_K) - T_{4n}(u_L)}{d_\sigma} \left[ \frac{\varphi(x_K) - \varphi(x_L)}{d_\sigma} - \frac{1}{|D_\sigma|} \int_{D_\sigma} \nabla \varphi \cdot \eta_{K,\sigma} \, dx \right]}_{I_2}.$$

where  $I_2 \rightarrow 0$  as  $h_{\mathcal{T}} \rightarrow 0$  and  $\tilde{\lambda}$  is defined by

$$\forall \sigma \in \mathcal{E}, \forall x \in D_\sigma, \tilde{\lambda}(x) = \lambda(u)_\sigma.$$

Recalling that  $\tilde{T}_n(u)$  converges strongly to  $T_n(u)$  (where  $\tilde{T}_n(u)$  is the mean value of  $T_n(u_K)$  and  $T_n(u_L)$  on  $D_\sigma$ ), since  $\lambda$  is continuous and bounded, an equi-integrability argument allows one to prove that  $\tilde{\lambda}$  converges to  $\lambda(u)$  in  $L^\infty$  weak  $\star$ .

Then we can repeat the end of the proof of Theorem 1.4.1. It follows that  $u$  is a renormalized solution of (1.44).

# Chapitre 2

## Finite volume scheme and renormalized solutions for a noncoercive and nonlinear parabolic problem with $L^1$ data

### 2.1 Introduction

In this work we consider a finite volume discretization of the following parabolic problem :

$$\begin{aligned} \frac{\partial u}{\partial t} - \operatorname{div}(\lambda(u)\nabla u) + \operatorname{div}(\mathbf{v}u) &= f \text{ in } Q, \\ u &= 0 \text{ on } \partial\Omega \times (0, T), \\ u(x, 0) &= u^0(x) \quad \forall x \in \Omega, \end{aligned} \tag{2.1}$$

where  $\Omega$  is an open bounded polygonal subset of  $\mathbb{R}^d$ ,  $d \geq 2$ ,  $T$  is a positive number and  $Q = \Omega \times (0, T)$ . Moreover  $\lambda$  is a continuous function such that  $\lambda(u) \geq \mu > 0$ , with  $\mu$  a positive real number,  $\mathbf{v} \in (L^{d+2}(Q))^d$  (see [34]),  $u^0 \in L^1(\Omega)$  and  $f \in L^1(Q)$ .

The main difficulties in dealing with the existence and the uniqueness of a solution to problem (2.1) are due to the noncoercive and nonlinear characters of the operator and to the  $L^1$  data  $f$ . These difficulties naturally have an impact on any finite volume or finite element scheme to handle approximation of solution of (2.1).

As far as parabolic equations with  $L^1$  data are concerned Boccardo and Gallouët have proved in [23] the existence of a distributional solution for nonlinear and coercive equations. Since the solution in the sense of distribution is not unique in general (see [69] in the elliptic case) different notions of solution have been developed for parabolic (or elliptic) equations with  $L^1$  data (and measure data) : the notion of renormalized solution which have been introduced by DiPerna and Lions in [36] for first order equations. This notion has been adapted to elliptic equations ([58]) and to parabolic equations ([18, 21]) with  $L^1$  or measure data. Renormalized solutions allow one to obtain in general existence, stability and uniqueness results when the data belongs to  $L^1$ . Nonlinear version of (2.1) with  $p$ -growth operator is addressed in [35] : existence and uniqueness are proved (under more restrictive assumptions). A large number of papers is devoted to renormalized or entropy solutions of parabolic

problems with  $L^1$  data : see, among others [16, 20, 63, 66].

As far as finite volume schemes for parabolic equations are concerned the linear and variational case is dealt in [44]. As in the continuous case one of the main obstacle for time dependent equations is to derive some compactness (and then the pointwise convergence) for the approximate solution which is necessary (but not sufficient) to pass to the limit. One of the methods is to derive with the equation an estimate on differences of time and space translates (see [44] and [1, 4, 6, 8, 45] in the context of degenerate parabolic equations). Another method is to use the compactness result proved in [48] which is a discrete version of Aubin-Lions Lemma (see [69]).

The convergence of a finite volume scheme for the linear version of (2.1) (i.e.  $\lambda = 1$ ) and when  $\mathbf{v}$  is more regular ( $\mathbf{v} \in \mathcal{C}^1(\bar{\Omega} \times [0, T])$  and  $\operatorname{div}(\mathbf{v}) = 0$ ) has been studied in [47] (see also [40] for the stationary case). They proved that the solution of the scheme for equation  $\partial_t u + \operatorname{div}(u\mathbf{v}) - \Delta u = f$  converges to a weak solution of this equation in the sense

$$\left\{ \begin{array}{l} u \in \bigcap_{1 \leq q \leq (d+2)/(d+1)} L^q(0, T; W_0^{1,q}(\Omega)), \\ \int_Q u \partial_t \varphi \, dx \, dt + \int_Q \nabla u \cdot \nabla \varphi \, dx \, dt - \int_Q u \mathbf{v} \cdot \nabla \varphi \, dx \, dt \\ \qquad \qquad \qquad = \int_Q f \varphi \, dx \, dt + \int_{\Omega} u_0 \varphi(0) \, dx \quad \forall \varphi \in \mathcal{C}_c^\infty(\Omega \times [0, T]). \end{array} \right. \quad (2.2)$$

The case where  $\lambda$  is a continuous and bounded function is dealt in [48, 51], where the authors prove the convergence of the solution of the finite volume scheme to a weak solution (in the sense of (2.2)) of (2.1). However the uniqueness of such a solution in the sense of (2.2) remains an open question even in the linear case.

The main novelty of the present paper is to mix the tools developed for finite volume schemes ([44, 47, 48]) with the strategy to deal with the existence of a renormalized solution for parabolic equations with  $L^1$  data ([20, 21]). Let us mention that such an analysis is performed in [53] for the stationary case, namely the equation  $-\operatorname{div}(\lambda(u)\nabla u) + \operatorname{div}(\mathbf{v}u) = f$  with  $f \in L^1(\Omega)$ ,  $\mathbf{v} \in L^p(\Omega)^d$ ,  $2 < p < +\infty$  if  $d = 2$ ,  $p = d$  if  $d \geq 3$  and with Dirichlet boundary conditions. It is worth noting that time dependent equations give additional difficulties due to the lack of regularity of  $\frac{\partial u}{\partial t}$ . We prove here that the approximate solution obtained by a finite volume scheme converges, up to a subsequence, to a renormalized solution of (2.1). In the case where  $\lambda$  is a local Lipschitz continuous function or in the case where  $\mathbf{v}$  is independent of the time, the uniqueness of the renormalized solution implies that the whole sequence converges.

Recall (see [20, 21]) that a renormalized solution of (2.1) is a measurable function  $u$  defined from  $Q$  to  $\mathbb{R}$  such that

$$u \in L^\infty(0, T; L^1(\Omega)), \quad (2.3)$$

$$\forall k > 0, T_k(u) \in L^2(0, T; H_0^1(\Omega)), \quad (2.4)$$

$$\lim_{k \rightarrow +\infty} \frac{1}{k} \int_0^T \int_{\Omega} \lambda(u) |\nabla T_k(u)|^2 \, dx \, dt = 0, \quad (2.5)$$

and for every functions  $\xi \in \mathcal{C}_c^\infty(\Omega \times [0, T])$  and  $S \in W^{2,\infty}(\mathbb{R})$  such that  $S'$  has a

compact support and  $S'(u)\xi \in L^2(0, T; H_0^1(\Omega))$ .

$$\begin{aligned}
& - \int_0^T \int_{\Omega} \xi_t \int_0^u S'(r) dr dx dt - \int_{\Omega} \xi(x, 0) \int_0^{u^0} S'(r) dr dx \\
& + \int_0^T \int_{\Omega} \lambda(u) \nabla u \cdot \nabla (\xi S'(u)) dx dt - \int_0^T \int_{\Omega} u \mathbf{v} \cdot \nabla (\xi S'(u)) dx dt \\
& = \int_0^T \int_{\Omega} f \xi S'(u) dx dt.
\end{aligned} \tag{2.6}$$

Here  $T_k$  is the truncate function at height  $k$  :

$$T_k(s) = \begin{cases} s & \text{if } |s| \leq k, \\ \text{sgn}(s)k & \text{if } |s| > k. \end{cases}$$

The existence of a renormalized solution of (2.1) is established in [34] for a non linear version with  $p$ -growth condition while the uniqueness is dealt in [35].

The chapter is organized as follows. In Section 2, we present the finite volume scheme, the time discretization and the properties of the discrete gradient. Section 3 is devoted to prove several estimates, especially an estimate on the time derivative and the discrete equivalent to (2.5) which is crucial to pass to the limit in the scheme. In Section 4, we prove the convergence of the finite volume scheme, passing to the limit and using both a density argument and Steklov averages.

## 2.2 The finite volume scheme and the time discretization

Let us define the admissibility of the mesh in the present work.

**Definition 2.2.1 (Admissible mesh)** *Let  $\Omega$  be an open bounded polygonal subset of  $\mathbb{R}^d$ . An admissible finite volume mesh of  $\Omega$  is given by a finite partition  $\mathcal{T}$  of  $\Omega$  in polygonal convex sets, called the "control volumes", by a finite family  $\mathcal{E}$  of disjoint subsets of  $\bar{\Omega}$  contained in affine hyperplanes, called the "edges", and by a family  $\mathcal{P} = (x_K)_{K \in \mathcal{T}}$  of points in  $\Omega$  such that*

- *each  $\sigma \in \mathcal{E}$  is a non-empty open subset of  $\partial K$  for some  $K \in \mathcal{T}$ ,*
- *by denoting  $\mathcal{E}(K) = \{\sigma \in \mathcal{E}; \sigma \in \partial K\}$ , one has  $\partial K = \cup_{\sigma \in \mathcal{E}(K)} \bar{\sigma}$  for all  $K \in \mathcal{T}$ ,*
- *for all  $K \neq L$  in  $\mathcal{T}$ , either the  $(d-1)$ -dimensional measure of  $\bar{K} \cap \bar{L}$  is null, or  $\bar{K} \cap \bar{L} = \bar{\sigma}$  for some  $\sigma \in \mathcal{E}$ , that we denote then  $\sigma = K|L$ ,*
- *for all  $K \in \mathcal{T}$ ,  $x_K$  is in the interior of  $K$ ,*
- *for all  $\sigma = K|L \in \mathcal{E}$ , the line  $(x_K, x_L)$  intersects and is orthogonal to  $\sigma$ ,*
- *for all  $\sigma \in \mathcal{E}$ ,  $\sigma \subset \partial\Omega \cap \partial K$ , the line which is orthogonal to  $\sigma$  and going through  $x_K$  intersects  $\sigma$ .*

We denote by  $|K|$  (resp.  $|\sigma|$ ) the Lebesgue measure of  $K \in \mathcal{T}$  (resp.  $\sigma \in \mathcal{E}$ ). The unit normal to  $\sigma \in \mathcal{E}(K)$  outward to  $K$  is denoted by  $\eta_{K,\sigma}$ .  $\mathcal{E}_{int}$  (resp.  $\mathcal{E}_{ext}$ ) is defined as the set of interior (resp. boundary) edges.

For any  $K \in \mathcal{T}$  and  $\sigma \in \mathcal{E}(K)$  we denote by  $d_{K,\sigma}$  the Euclidean distance between  $x_K$

and  $\sigma$ . For any  $\sigma \in \mathcal{E}$ , we define  $d_\sigma = d_{K,\sigma} + d_{L,\sigma}$  if  $\sigma = K|L \in \mathcal{E}_{int}$  (in which case  $d_\sigma$  is the Euclidean distance between  $x_K$  and  $x_L$ ) and  $d_\sigma = d_{K,\sigma}$  if  $\sigma \in \mathcal{E}_{ext} \cap \mathcal{E}(K)$ . The size of the mesh, denoted by  $h_{\mathcal{T}}$ , is defined by  $h_{\mathcal{T}} = \sup_{K \in \mathcal{T}} \text{diam}(K)$ .

In the continuous case, Poincaré and Sobolev inequalities are necessary to find a priori estimates, and then to solve the problem. Since, in the discrete case we also need to find such estimates, we will need discrete versions of these two inequalities (see [32]). Before that, we have to define the discrete  $W_0^{1,q}$  norm.

**Definition 2.2.2 (Discrete  $W_0^{1,q}$  norm)** *Let  $\Omega$  be an open bounded polygonal subset of  $\mathbb{R}^d$ ,  $d \geq 2$ , and let  $\mathcal{T}$  be an admissible mesh. Define  $X(\mathcal{T})$  as the set of functions from  $\Omega$  to  $\mathbb{R}$  which are constant over each control volume of the mesh. For  $v_{\mathcal{T}} \in X(\mathcal{T})$  and  $q \in [1, +\infty[$ , we define the discrete  $W_0^{1,q}$  norm by*

$$\|v_{\mathcal{T}}\|_{1,q,\mathcal{T}}^q = \sum_{\substack{\sigma \in \mathcal{E}_{int} \\ \sigma = K|L}} |\sigma| d_\sigma \left| \frac{v_K - v_L}{d_\sigma} \right|^q + \sum_{\substack{\sigma \in \mathcal{E}_{ext} \\ \sigma \in \mathcal{E}(K)}} |\sigma| d_\sigma \left| \frac{v_K}{d_\sigma} \right|^q$$

where  $v_K$  denotes the value taken by  $v$  on the control volume  $K$ .

Let us also define

$$\begin{aligned} \|v_{\mathcal{T}}\|_{1,\infty,\mathcal{T}} = \max \{ & \left\{ \frac{|v_K - v_L|}{d_\sigma}, \sigma \in \mathcal{E}_{int}, \sigma = K|L \right\} \\ & \cup \left\{ \frac{|v_K|}{d_\sigma}, \sigma \in \mathcal{E}_{ext}, \sigma \in \mathcal{E}(K) \right\} \cup \{|v_K|, K \in \mathcal{T}\} \}. \end{aligned}$$

**Definition 2.2.3 (Discrete dual norm)** *Let  $q > 1$  and let  $q'$  be given by  $\frac{1}{q} + \frac{1}{q'} = 1$  if  $q$  is finite and  $q' = 1$  if  $q = +\infty$ . We define the dual norm by*

$$\|v\|_{-1,q',\mathcal{T}} = \sup_{\substack{w \in X(\mathcal{T}) \\ w \neq 0}} \frac{1}{\|w\|_{1,q,\mathcal{T}}} \int_{\Omega} v w \, dx.$$

**Proposition 2.2.4 (Discrete Poincaré inequality)** *Let  $\mathcal{T}$  be an admissible mesh and  $v_{\mathcal{T}} \in X(\mathcal{T})$ . Then, if  $1 \leq q \leq 2$ ,*

$$\|v_{\mathcal{T}}\|_{L^q(\Omega)} \leq \text{diam}(\Omega) \|v_{\mathcal{T}}\|_{1,q,\mathcal{T}}.$$

**Proposition 2.2.5 (Discrete Sobolev inequality)** *Let  $1 \leq q \leq 2$ ,  $\mathcal{T}$  be an admissible mesh and  $\xi > 0$  satisfying*

$$\text{for all } K \in \mathcal{T} \text{ and all } \sigma \in \mathcal{E}(K), \quad d_{K,\sigma} \geq \xi d_\sigma.$$

*Then, with  $q^* = \frac{dq}{d-q}$  if  $q < d$  and  $q^* < \infty$  if  $q = d = 2$ , there exists  $C > 0$  only depending on  $(\Omega, q, q^*, \xi)$  such that, for all  $v_{\mathcal{T}} \in X(\mathcal{T})$ ,*

$$\|v_{\mathcal{T}}\|_{L^{q^*}(\Omega)} \leq C \|v_{\mathcal{T}}\|_{1,q,\mathcal{T}}.$$

As far as the time discretization is concerned, we suppose given a uniform partition of the interval  $[0, T)$  such that  $[0, T] = \cup_{1 \leq i \leq N} [t^{i-1}, t^i]$  where  $t^i = i\delta t$  and  $\delta t = \frac{T}{N}$ . We define  $X_{\mathcal{D}}$  the space of piecewise constant functions over each  $K \times (t^{i-1}, t^i)$  for  $K \in \mathcal{T}$  and  $1 \leq i \leq N$ . To each sequence  $(v_{\mathcal{T}}^i)_{1 \leq i \leq N}$  of functions of  $X(\mathcal{T})$ , we associate the function  $v_{\mathcal{D}} \in X_{\mathcal{D}}$  defined by :

$$v_{\mathcal{D}}(x, t) = v_{\mathcal{T}}^i(x) \quad \forall x \in \Omega, \quad \forall t \in (t^{i-1}, t^i), \quad 1 \leq i \leq N,$$

$$\text{i.e. } v_{\mathcal{D}}(x, t) = v_K^i \quad \forall x \in K, \quad K \in \mathcal{T}, \quad \forall t \in (t^{i-1}, t^i), \quad 1 \leq i \leq N.$$

We now define the discrete time derivative of  $v_{\mathcal{D}}$  by

$$\partial_{t, \mathcal{D}} v^i(x) = \frac{v^i(x) - v^{i-1}(x)}{\delta t} \quad \forall 1 \leq i \leq N,$$

$$\text{i.e. } \partial_t v_K^i = \frac{v_K^i - v_K^{i-1}}{\delta t} \quad \forall K \in \mathcal{T}, \quad \forall 1 \leq i \leq N.$$

Before writing the finite volume scheme, let us recall the definition of the discrete finite volume gradient (see [51]).

**Definition 2.2.6 (Discrete finite volume gradient)** For  $K \in \mathcal{T}$  and  $\sigma \in \mathcal{E}(K)$ , we define the volume  $D_{K, \sigma}$  as the cone of basis  $\sigma$  and of opposite vertex  $x_K$ . Then, we define the "diamond-cell"  $D_{\sigma}$  (see Figure 2.1) by

$$\begin{aligned} D_{\sigma} &= D_{K, \sigma} \cup D_{L, \sigma} && \text{if } \sigma = K|L \in \mathcal{E}_{int}, \\ D_{\sigma} &= D_{K, \sigma} && \text{if } \sigma \in \mathcal{E}_{ext} \cap \mathcal{E}(K). \end{aligned}$$

For any  $v_{\mathcal{D}} \in X_{\mathcal{D}}$ , we define the discrete gradient  $\nabla_{\mathcal{D}} v_{\mathcal{D}}$  as the piecewise constant function over each  $D_{\sigma} \times (t^{i-1}, t^i)$  and given by

$$\forall \sigma \in \mathcal{E}_{int}, \quad \sigma = K|L, \quad \forall 1 \leq i \leq N,$$

$$\nabla_{\mathcal{D}} v_{\mathcal{D}}(x, t) = d \frac{v_L^i - v_K^i}{d_{\sigma}} \eta_{K, \sigma}, \quad \forall x \in D_{\sigma}, \quad \forall t \in (t^{i-1}, t^i],$$

$$\forall \sigma \in \mathcal{E}_{ext} \cap \mathcal{E}(K), \quad \forall 1 \leq i \leq N,$$

$$\nabla_{\mathcal{T}} v_{\mathcal{D}}(x, t) = d \frac{0 - v_K^i}{d_{\sigma}} \eta_{K, \sigma}, \quad \forall x \in D_{\sigma}, \quad \forall t \in (t^{i-1}, t^i].$$

**Lemma 2.2.7 (Weak convergence of the finite volume gradient)** Let  $(\mathcal{T}_m)_{m \geq 1}$  be a sequence of admissible meshes such that there exists  $\xi > 0$  satisfying

$$\text{for all } m \geq 1 \text{ for all } K \in \mathcal{T} \text{ and all } \sigma \in \mathcal{E}(K), \quad d_{K, \sigma} \geq \xi d_{\sigma},$$

and such that  $h_{\mathcal{T}_m} \rightarrow 0$  and let  $\delta t_m$  be a sequence of time steps such that  $\delta t_m \rightarrow 0$ . Let  $v_{\mathcal{D}_m} \in X_{\mathcal{D}_m}$  and let us assume that there exists  $C > 0$  such that  $\sum_{i=1}^N \delta t_m \|v_{\mathcal{T}_m}^i\|_{1, 2, \mathcal{T}_m}^2 \leq C$ , and that  $v_{\mathcal{D}_m}$  converges weakly in  $L^1(Q)$  to  $v \in L^2(0, T; H_0^1(\Omega))$ . Then  $\nabla_{\mathcal{D}_m} v_{\mathcal{D}_m}$  converges to  $\nabla v$  weakly in  $L^2(Q)^d$ .

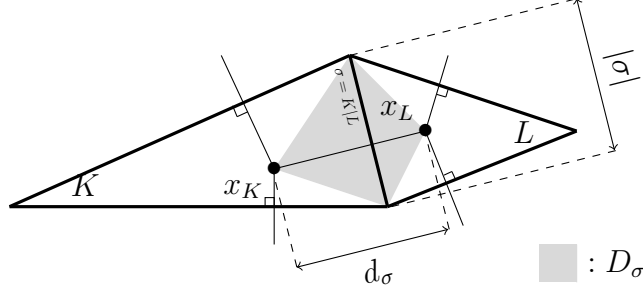


FIGURE 2.1 – The diamond  $D_\sigma$

**Proof** This proof is done in Lemma 5.2 of [5] where the idea is to adapt the proof done in the elliptic case in [51]. Let  $\varphi \in C_c^\infty(Q)^d$ . Since the size of the mesh tends to zero we can suppose that the intersection of the support of  $\varphi$  and the diamond cells associated to the external faces of the meshes is the empty set. We have

$$\begin{aligned} \int_0^T \int_\Omega \nabla_{\mathcal{D}_m} v \cdot \varphi dx dt &= \sum_{i=1}^N \sum_{\sigma=K|L \in \mathcal{E}_{int}} d \frac{v_L^i - v_K^i}{d_\sigma} \int_{t_{i-1}}^{t_i} \int_{D_\sigma} \varphi dx \cdot \eta_{K,\sigma} \\ &= T_1 + T_2, \end{aligned}$$

where

$$\begin{aligned} T_1 &= \sum_{i=1}^N \sum_{\sigma=K|L \in \mathcal{E}_{int}} (v_L^i - v_K^i) \int_{t_{i-1}}^{t_i} \int_\sigma \varphi \cdot \eta d\gamma, \\ T_2 &= \sum_{i=1}^N \sum_{\sigma=K|L \in \mathcal{E}_{int}} |\sigma| (v_L^i - v_K^i) \left[ \frac{1}{|D_\sigma|} \int_{t_{i-1}}^{t_i} \int_{D_\sigma} \varphi dx - \frac{1}{|\sigma|} \int_{t_{i-1}}^{t_i} \int_\sigma \varphi d\gamma \right] \cdot \eta_{K,\sigma}. \end{aligned}$$

Reordering the sums, we obtain

$$\begin{aligned} T_1 &= - \sum_{i=1}^N \sum_{K \in \mathcal{T}_m} v_K^i \sum_{\sigma=K|L} \int_{t_{i-1}}^{t_i} \int_\sigma \varphi \cdot \eta_{K,\sigma} d\gamma \\ &= - \sum_{i=1}^N \sum_{K \in \mathcal{T}_m} \int_{t_{i-1}}^{t_i} \int_K v_{\mathcal{D}_m} \operatorname{div} \varphi dx \\ &= - \int_0^T \int_\Omega v_{\mathcal{D}_m} \operatorname{div} \varphi dx dt. \end{aligned}$$

Since  $v_{\mathcal{D}_m}$  converges weakly in  $L^1(Q)$  we get

$$\lim_{\substack{h_{\mathcal{T}_m} \rightarrow 0 \\ \delta t_m \rightarrow 0}} T_1 = - \int_0^T \int_\Omega v \operatorname{div} \varphi dx dt.$$

As far as  $T_2$  is concerned, thanks to the regularity of  $\varphi$ , we have

$$|T_2| \leq C_\varphi h_{\mathcal{T}_m} \sum_{i=1}^N \delta t_m \|v_{\mathcal{T}_m}^i\|_{1,1,\mathcal{T}},$$

and by Hölder's inequality and the fact that  $\sum_{i=1}^N \delta t_m \|v_{\mathcal{T}_m}^i\|_{1,2,\mathcal{T}_m}^2 \leq C$  we obtain that

$$\lim_{\substack{h_{\mathcal{T}_m} \rightarrow 0 \\ \delta t_m \rightarrow 0}} T_2 = 0.$$

Since  $v \in L^2(0, T; H_0^1(\Omega))$  it follows that

$$\forall \varphi \in (C_c^\infty(Q))^d, \quad \lim_{\substack{h_{\mathcal{T}_m} \rightarrow 0 \\ \delta t_m \rightarrow 0}} \int_0^T \int_{\Omega} \nabla_{\mathcal{D}_m} v_{\mathcal{D}_m} \cdot \varphi \, dx \, dt = \int_0^T \int_{\Omega} \nabla v \cdot \varphi \, dx \, dt.$$

Recalling that  $\nabla_{\mathcal{D}_m} v_{\mathcal{D}_m}$  is bounded in  $L^2(Q)^d$ , by density of  $C_c^\infty(Q)^d$  in  $L^2(Q)^d$  the proof is complete.  $\square$

Let  $\mathcal{T}$  be an admissible mesh, we are now in position to give the finite volume discretization of (2.1). For  $1 \leq i \leq N$ ,  $K \in \mathcal{T}$  and  $\sigma \in \mathcal{E}(K)$ , we define

$$\mathbf{v}_{K,\sigma}^i = \frac{1}{\delta t |D_\sigma|} \int_{t_{i-1}}^{t_i} \int_{D_\sigma} \mathbf{v} \cdot \eta_{K,\sigma} \, dx \, dt, \quad (2.7)$$

$$f_K^i = \frac{1}{\delta t |K|} \int_{t_{i-1}}^{t_i} \int_K f \, dx \, dt, \quad (2.8)$$

$$u_K^0 = \frac{1}{|K|} \int_K u^0(x) \, dx, \quad (2.9)$$

and we can write the scheme as the following set of equations :

For all  $1 \leq i \leq N$  and for all  $K \in \mathcal{T}$

$$|K| \frac{u_K^i - u_K^{i-1}}{\delta t} + \sum_{\sigma \in \mathcal{E}(K)} \frac{|\sigma|}{d_\sigma} \lambda(u)_\sigma^i (u_K^i - u_L^i) + \sum_{\sigma \in \mathcal{E}(K)} |\sigma| \mathbf{v}_{K,\sigma}^i u_{\sigma,+}^i = |K| f_K^i, \quad (2.10)$$

with  $u_L^i = 0$  if  $\sigma \in \mathcal{E}_{ext} \cap \mathcal{E}(K)$  and

$$\forall \sigma = K|L \in \mathcal{E}_{int}, \quad u_{\sigma,+}^i = u_K^i \text{ if } \mathbf{v}_{K,\sigma}^i \geq 0, \quad u_{\sigma,+}^i = u_L^i \text{ otherwise}, \quad (2.11)$$

$$\forall \sigma \in \mathcal{E}_{ext} \cap \mathcal{E}(K), \quad u_{\sigma,+}^i = u_K^i \text{ if } \mathbf{v}_{K,\sigma}^i \geq 0, \quad u_{\sigma,+}^i = 0 \text{ otherwise}. \quad (2.12)$$

We denote  $u_{\sigma,-}^i$  the downstream choice of  $u$ , i.e.  $u_{\sigma,-}^i$  is such that  $\{u_{\sigma,+}^i, u_{\sigma,-}^i\} = \{u_K^i, u_L^i\}$  (with  $u_L^i = 0$  if  $\sigma \in \mathcal{E}_{ext} \cap \mathcal{E}(K)$ ).

For all  $1 \leq i \leq N$  we define  $\lambda(u)_\sigma^i$  in the following way

$$\forall \sigma = K|L \in \mathcal{E}_{int}, \quad \min[\lambda(u_K^i), \lambda(u_L^i)] \leq \lambda(u)_\sigma^i \leq \max[\lambda(u_K^i), \lambda(u_L^i)], \quad (2.13)$$

$$\forall \sigma \in \mathcal{E}_{ext} \cap \mathcal{E}(K), \quad \lambda(u)_\sigma^i = \lambda(u_K^i), \quad (2.14)$$

where  $\lambda(u)_\sigma^i$  is for example the mean value of  $\lambda(u_K^i)$  and  $\lambda(u_L^i)$ .



## 2.3 Estimates

This section is devoted to derive estimates on the sequence  $u_{\mathcal{D}}$ . The very first estimate,  $u_{\mathcal{D}}$  bounded in  $L^\infty(0, T; L^1(\Omega))$ , is proved in Proposition 2.3.1. Since we do not assume that  $\operatorname{div} \mathbf{v} = 0$  we use a different method than the one developed in [47]. A log-type estimate is first obtained in the same spirit that in the elliptic case (see [40]) which in turn gives the  $L^\infty(L^1)$  bound. We then prove in Proposition 2.3.3 uniform estimates on the truncated discrete energy which are discrete versions of the continuous renormalized case. Using Proposition 2.3.3, the discrete equation and the discrete compactness result proved in [48] we are then able to obtain in Proposition 2.3.8 the pointwise convergence of  $u_{\mathcal{D}}$  which is crucial to perform the convergence analysis.

**Proposition 2.3.1** *Let  $\mathcal{T}$  be an admissible mesh and let  $\delta t$  be a time step. If  $u_{\mathcal{D}} = (u_{\mathcal{T}}^i)_{1 \leq i \leq N}$  is a solution to (2.10), then*

$$\sum_{i=1}^N \delta t \|\ln(1 + |u_{\mathcal{T}}^i|)\|_{1,2,\mathcal{T}}^2 \leq C. \quad (2.15)$$

and

$$\|u_{\mathcal{D}}\|_{L^\infty(0,T;L^1(\Omega))} \leq C. \quad (2.16)$$

**Proof** The proof is divided in two steps. In Step 1 we adapt the proof of the log-estimate derived in [40] to the parabolic case. Step 2 is devoted to derive (2.16). Let  $\varphi(s) = \int_0^s \frac{dr}{(1+|r|)^2}$  and  $\tilde{\varphi}(s) = \int_0^s \varphi(r) dr$ . By multiplying each equation of the scheme (2.10) by  $\varphi(u_K^i)$  and by summing over each control volume and each time step, we obtain  $T_1 + T_2 + T_3 = T_4$  with

$$\begin{aligned} T_1 &= \sum_{i=1}^N \sum_{K \in \mathcal{T}} |K| (u_K^i - u_K^{i-1}) \varphi(u_K^i), \\ T_2 &= \sum_{i=1}^N \delta t \sum_{\sigma \in \mathcal{E}} \frac{|\sigma|}{d_\sigma} \lambda(u)_\sigma^i (u_K^i - u_L^i) (\varphi(u_K^i) - \varphi(u_L^i)), \\ T_3 &= \sum_{i=1}^N \delta t \sum_{\sigma \in \mathcal{E}} |\sigma| \mathbf{v}_{K,\sigma}^i u_{\sigma,+}^i (\varphi(u_K^i) - \varphi(u_L^i)), \\ T_4 &= \sum_{i=1}^N \delta t \sum_{K \in \mathcal{T}} |K| f_K^i \varphi(u_K^i). \end{aligned}$$

Since  $\varphi$  is bounded by 1, we have that  $|T_4| \leq \|f\|_{L^1(Q)}$ . By the convexity of  $\tilde{\varphi}$  we have

$$\tilde{\varphi}(u_K^{i-1}) \geq \tilde{\varphi}(u_K^i) + \varphi(u_K^i) (u_K^{i-1} - u_K^i),$$

so we find that

$$T_1 \geq - \sum_{K \in \mathcal{T}} |K| \tilde{\varphi}(u_K^0) + \sum_{K \in \mathcal{T}} |K| \tilde{\varphi}(u_K^N).$$

By the definition of the function  $\tilde{\varphi}$

$$\sum_{K \in \mathcal{T}} |K| |\tilde{\varphi}(u_K^0)| \leq \sum_{K \in \mathcal{T}} |K| |u_K^0| = \|u^0\|_{L^1(\Omega)},$$

and since  $\tilde{\varphi}$  is a convex function, there exists  $C_{\tilde{\varphi}} > 0$  such that

$$C_{\tilde{\varphi}}(1 + \tilde{\varphi}(u_K^N)) \geq \|u_{\mathcal{T}}^N\|_{L^1(\Omega)}.$$

Then there exists a constant  $C > 0$  only depending on  $\|f\|_{L^1(Q)}$ ,  $\|u^0\|_{L^1(\Omega)}$  and  $C_{\tilde{\varphi}}$  such that

$$T_2 + T_3 + \|u_{\mathcal{T}}^N\|_{L^1(\Omega)} \leq C. \quad (2.17)$$

Moreover thanks to the definition of  $\lambda$  we have

$$T_2 \geq \mu \sum_{i=1}^N \delta t \sum_{\sigma \in \mathcal{E}} \frac{|\sigma|}{d_{\sigma}} (u_K^i - u_L^i) (\varphi(u_K^i) - \varphi(u_L^i)) \geq 0.$$

As in [40],  $T_3$  can be rewritten as

$$-T_3 = \sum_{i=1}^N \delta t \sum_{\sigma \in \mathcal{E}} |\sigma| |\mathbf{v}_{K,\sigma}^i| u_{\sigma,+}^i (\varphi(u_{\sigma,-}^i) - \varphi(u_{\sigma,+}^i)).$$

As in [40], we define for all  $1 \leq i \leq N$  the subsets  $\mathcal{A}^i$  of edges by

$$\mathcal{A}^i = \{\sigma \in \mathcal{E}; u_{\sigma,+}^i \geq u_{\sigma,-}^i, u_{\sigma,+}^i < 0\} \cup \{\sigma \in \mathcal{E}; u_{\sigma,+}^i < u_{\sigma,-}^i, u_{\sigma,+}^i \geq 0\}, \quad (2.18)$$

and since  $\varphi$  is non decreasing we have

$$\begin{aligned} -T_3 &\leq \sum_{i=1}^N \delta t \sum_{\sigma \in \mathcal{A}^i} |\sigma| |\mathbf{v}_{K,\sigma}^i| u_{\sigma,+}^i (\varphi(u_{\sigma,-}^i) - \varphi(u_{\sigma,+}^i)) \\ &\leq \underbrace{\left( \sum_{i=1}^N \delta t \sum_{\sigma \in \mathcal{A}^i} |\sigma| |\mathbf{v}_{K,\sigma}^i|^2 \right)^{\frac{1}{2}}}_I \underbrace{\left( \sum_{i=1}^N \delta t \sum_{\sigma \in \mathcal{A}^i} \frac{|\sigma|}{d_{\sigma}} (u_{\sigma,+}^i)^2 (\varphi(u_{\sigma,-}^i) - \varphi(u_{\sigma,+}^i))^2 \right)^{\frac{1}{2}}}_{II}. \end{aligned}$$

Using Jensen's inequality we find

$$\begin{aligned} I &\leq \left( \sum_{i=1}^N \delta t \sum_{\sigma \in \mathcal{A}^i} \frac{|\sigma| d_{\sigma}}{\delta t |D_{\sigma}|} \int_{t_{i-1}}^{t_i} \int_{D_{\sigma}} |\mathbf{v}|^2 dx dt \right)^{\frac{1}{2}} \\ &= d^{\frac{1}{2}} \|\mathbf{v}\|_{L^2(Q)^d} \\ &\leq C_1. \end{aligned}$$

As far as  $II$  is concerned, in view of the definition of  $\varphi$ , we have (see Lemma 3.6 in [40]) that  $\forall |x| < |y|$  such that  $xy > 0$

$$x^2(\varphi(y) - \varphi(x))^2 \leq (y - x)(\varphi(y) - \varphi(x)).$$

We then obtain

$$\begin{aligned} II &\leq \left( \sum_{i=1}^N \delta t \sum_{\sigma \in \mathcal{A}^i} \frac{|\sigma|}{d_{\sigma}} (u_{\sigma,-}^i - u_{\sigma,+}^i) (\varphi(u_{\sigma,-}^i) - \varphi(u_{\sigma,+}^i)) \right)^{\frac{1}{2}} \\ &\leq \left( \sum_{i=1}^N \delta t \sum_{\sigma \in \mathcal{E}} \frac{|\sigma|}{d_{\sigma}} (u_K^i - u_L^i) (\varphi(u_K^i) - \varphi(u_L^i)) \right)^{\frac{1}{2}}. \end{aligned}$$

Finally we get with Young's inequality that

$$-T_3 \leq \frac{C_1^2}{2\mu} + \frac{\mu}{2} \sum_{i=1}^N \delta t \sum_{\sigma \in \mathcal{E}} \frac{|\sigma|}{d_\sigma} (u_K^i - u_L^i) (\varphi(u_K^i) - \varphi(u_L^i)).$$

Thanks to (2.17) we deduce that

$$\sum_{i=1}^N \delta t \sum_{\sigma \in \mathcal{E}} \frac{|\sigma|}{d_\sigma} (u_K^i - u_L^i) (\varphi(u_K^i) - \varphi(u_L^i)) \leq C_2,$$

where  $C_2 > 0$  is a constant only depending on  $f, u^0, \varphi, \mu, \mathbf{v}$ .

In view of the definition of  $\varphi$  and using the Cauchy-Schwarz inequality, we notice that

$$(\ln(1 + |u_K^i|) - \ln(1 + |u_L^i|))^2 \leq (u_K^i - u_L^i)(\varphi(u_K^i) - \varphi(u_L^i)).$$

We then obtain the desired estimate, i.e.

$$\sum_{i=1}^N \delta t \|\ln(1 + |u_{\mathcal{T}}^i|)\|_{1,2,\mathcal{T}}^2 \leq C_2.$$

We now turn to Step 2. Summing the computations of the previous step only from 1 to  $M$  for all  $1 \leq M \leq N$ , we find

$$\|u_{\mathcal{T}}^M\|_{L^1(\Omega)} \leq C, \quad \forall 1 \leq M \leq N.$$

We thus get

$$\|u_{\mathcal{D}}\|_{L^\infty(0,T;L^1(\Omega))} \leq C.$$

□

**Corollary 2.3.2** *Let  $\mathcal{T}$  be an admissible mesh and let  $\delta t$  be a time step. If  $u_{\mathcal{D}} = (u_{\mathcal{T}}^i)_{1 \leq i \leq N}$  is a solution to (2.10) and, for  $n > 0$ ,  $E_n = \{(x, t) \in \Omega \times (0, T); |u_{\mathcal{D}}(x, t)| > n\}$ , then there exists  $C > 0$  only depending on  $(\Omega, \mathbf{v}, f, d, p)$  such that*

$$|E_n| \leq \frac{C}{(\ln(1 + n))^2}. \quad (2.19)$$

**Proof** Using Proposition 2.3.1 and the discrete Poincaré inequality, we get that

$$\sum_{i=1}^N \delta t \|\ln(1 + |u_{\mathcal{T}}^i|)\|_{L^2(\Omega)}^2 \leq C \quad \text{i.e.} \quad \|\ln(1 + |u_{\mathcal{D}}|)\|_{L^2(0,T;L^2(\Omega))}^2 \leq C,$$

where  $C$  only depends on  $\text{diam}(\Omega)$ ,  $d$ ,  $\|f\|_{L^1(Q)}$ ,  $\|u^0\|_{L^1(\Omega)}$  and  $\|\mathbf{v}\|_{L^2(Q)^d}$ . Thus,

$$(\ln(1 + n))^2 |E_n| \leq C.$$

□

In the following proposition we prove a uniform estimate on the truncate energy of  $u_{\mathcal{D}}$  (see (2.20)) which is crucial to pass to the limit in the approximate problem. We explicitly observe that (2.20) is the discrete version of (2.5) which is imposed in the definition of the renormalized solution. In the elliptic case, in order to prove such estimate, we use the convergence of  $u_{\mathcal{D}}$  (see [53]) which we do not have yet. We will see that Corollary 2.3.2 and an equi-integrability argument are sufficient to obtain the desired estimate. Since the equation contains the term  $\operatorname{div}(\mathbf{v} u)$  we also have to uniformly control the discrete version of  $\frac{1}{n} \int_0^T \int_{\Omega} \mathbf{v} u \nabla T_n(u) dx dt$  which is stated in (2.21). Estimate (2.22) is a control of a rest which will be needed to pass to the limit.

**Proposition 2.3.3 (Discrete estimate on the energy)**

Let  $(\mathcal{T}_m)_{m \geq 1}$  be a sequence of admissible meshes such that there exists  $\xi > 0$  satisfying

$$\text{for all } m \geq 1 \text{ for all } K \in \mathcal{T} \text{ and all } \sigma \in \mathcal{E}(K), d_{K,\sigma} \geq \xi d_{\sigma},$$

and  $(\delta t_m)_{m \geq 1}$  be a sequence of time steps.

If  $u_{\mathcal{D}_m} = (u_{\mathcal{T}_m}^i)_{1 \leq i \leq N}$  is a solution to (2.10), then

$$\lim_{n \rightarrow +\infty} \overline{\lim}_{\substack{h_{\mathcal{T}_m} \rightarrow 0 \\ \delta t_m \rightarrow 0}} \frac{1}{n} \sum_{i=1}^N \delta t \sum_{\sigma \in \mathcal{E}} \frac{|\sigma|}{d_{\sigma}} \lambda(u)_{\sigma}^i (u_K^i - u_L^i) (T_n(u_K^i) - T_n(u_L^i)) = 0, \quad (2.20)$$

$$\lim_{n \rightarrow +\infty} \overline{\lim}_{\substack{h_{\mathcal{T}_m} \rightarrow 0 \\ \delta t_m \rightarrow 0}} \frac{1}{n} \sum_{i=1}^N \delta t \sum_{\sigma \in \mathcal{E}} |\sigma| |\mathbf{v}_{K,\sigma}^i| |u_{\sigma,+}^i| |T_n(u_{\sigma,+}^i) - T_n(u_{\sigma,-}^i)| = 0, \quad (2.21)$$

$$\lim_{n \rightarrow +\infty} \overline{\lim}_{\substack{h_{\mathcal{T}_m} \rightarrow 0 \\ \delta t_m \rightarrow 0}} \frac{1}{n} \sum_{i=1}^N \sum_{K \in \mathcal{T}} |K| \int_{u_K^{i-1}}^{u_K^i} (r - u_K^{i-1}) T_n'(r) dr = 0, \quad (2.22)$$

where  $u_L^i = 0$  if  $\sigma \in \mathcal{E}_{ext}$

**Proof** We first establish (2.20). Let  $\mathcal{T}$  be an admissible mesh and  $\delta t$  a time step and let  $u_{\mathcal{D}}$  be a solution of (2.10). Multiplying each equation of the scheme by  $\frac{T_n(u_K^i)}{n}$ , summing on  $K \in \mathcal{T}$ , gathering by edges and from 1 to  $M$  where  $1 \leq M \leq N$  lead to

$$T_1 + T_2 + T_3 = T_4 \quad (2.23)$$

with

$$T_1 = \frac{1}{n} \sum_{i=1}^M \delta t \sum_{K \in \mathcal{T}} |K| \frac{u_K^i - u_K^{i-1}}{\delta t} T_n(u_K^i), \quad (2.24)$$

$$T_2 = \frac{1}{n} \sum_{i=1}^M \delta t \sum_{\sigma \in \mathcal{E}} \frac{|\sigma|}{d_{\sigma}} \lambda(u)_{\sigma}^i (u_K^i - u_L^i) (T_n(u_K^i) - T_n(u_L^i)), \quad (2.25)$$

$$T_3 = \frac{1}{n} \sum_{i=1}^M \delta t \sum_{\sigma \in \mathcal{E}} |\sigma| \mathbf{v}_{K,\sigma}^i u_{\sigma,+}^i (T_n(u_K^i) - T_n(u_L^i)), \quad (2.26)$$

$$T_4 = \frac{1}{n} \sum_{i=1}^M \delta t \sum_{K \in \mathcal{T}} |K| f_K^i T_n(u_K^i). \quad (2.27)$$

As far as  $T_1$  is concerned, we find that

$$T_1 = \frac{1}{n} \sum_{K \in \mathcal{T}} |K| \tilde{T}_n(u_K^M) - \frac{1}{n} \sum_{K \in \mathcal{T}} |K| \tilde{T}_n(u_K^0) + \frac{1}{n} \sum_{i=1}^N \sum_{K \in \mathcal{T}} |K| \int_{u_K^{i-1}}^{u_K^i} (r - u_K^{i-1}) T_n'(r) dr = 0,$$

where  $\tilde{T}_n(v) = \int_0^v T_n(r) dr$ . Since

$$\tilde{T}_n(u_K^M) \geq \frac{1}{2} u_K^M T_n(u_K^M) \geq \frac{1}{2} |T_n(u_K^M)|^2$$

we have

$$\frac{1}{n} \sum_{K \in \mathcal{T}} |K| \tilde{T}_n(u_K^M) \geq \frac{1}{2n} \|T_n(u_{\mathcal{T}}^M)\|_{L^2(\Omega)}^2. \quad (2.28)$$

Moreover the convexity of  $\tilde{T}_n$  yields that

$$\frac{1}{n} \sum_{i=1}^N \sum_{K \in \mathcal{T}} |K| \int_{u_K^{i-1}}^{u_K^i} (r - u_K^{i-1}) T_n'(r) dr \geq 0. \quad (2.29)$$

We now study  $T_3$ . If  $\sigma \in \mathcal{E}$ , by the definition (2.11) of  $u_{\sigma,+}^i$  and recalling that  $u_{\sigma,-}^i$  is the downstream choice of  $u_{\mathcal{T}}^i$ , if  $\mathbf{v}_{K,\sigma}^i \geq 0$  we have

$$\mathbf{v}_{K,\sigma}^i (T_n(u_K^i) - T_n(u_L^i)) = \mathbf{v}_{K,\sigma}^i (T_n(u_{\sigma,+}^i) - T_n(u_{\sigma,-}^i)),$$

and if  $\mathbf{v}_{K,\sigma}^i < 0$  we have

$$\mathbf{v}_{K,\sigma}^i (T_n(u_K^i) - T_n(u_L^i)) = -\mathbf{v}_{K,\sigma}^i (T_n(u_{\sigma,+}^i) - T_n(u_{\sigma,-}^i)).$$

It follows that  $T_3$  can be rewritten as

$$-T_3 = \frac{1}{n} \sum_{i=1}^M \delta t \sum_{\sigma \in \mathcal{E}} |\sigma| |\mathbf{v}_{K,\sigma}^i| u_{\sigma,+}^i (T_n(u_{\sigma,-}^i) - T_n(u_{\sigma,+}^i)).$$

In view of the definition (2.18) of the subsets  $\mathcal{A}^i$  and since  $T_n$  is non decreasing we have,  $\forall 1 \leq M \leq N$

$$-T_3 \leq \frac{1}{n} \sum_{i=1}^N \delta t \sum_{\sigma \in \mathcal{A}^i} |\sigma| |\mathbf{v}_{K,\sigma}^i| u_{\sigma,+}^i (T_n(u_{\sigma,-}^i) - T_n(u_{\sigma,+}^i)). \quad (2.30)$$

By taking the supremum in (2.23) in  $M \in \{1, \dots, N\}$ , (2.28), (2.29) and (2.30) allow one to obtain

$$0 \leq T'_{1,1} + T'_{1,3} + T'_2 \leq 2T'_4 + 2T'_{1,2} - 2T'_3, \quad (2.31)$$

where

$$T'_{1,1} = \frac{1}{2n} \sup_{1 \leq i \leq N} \|T_n(u^i)\|_{L^2(\Omega)}^2, \quad (2.32)$$

$$T'_{1,2} = \frac{1}{n} \sum_{K \in \mathcal{T}} |K| \tilde{T}_n(u_K^0), \quad (2.33)$$

$$T'_{1,3} = \frac{1}{n} \sum_{i=1}^N \sum_{K \in \mathcal{T}} |K| \int_{u_K^{i-1}}^{u_K^i} (r - u_K^{i-1}) T'_n(r) dr, \quad (2.34)$$

$$T'_2 = \frac{1}{n} \sum_{i=1}^N \delta t \sum_{\sigma \in \mathcal{E}} \frac{|\sigma|}{d_\sigma} \lambda(u)_\sigma^i (u_K^i - u_L^i) (T_n(u_K^i) - T_n(u_L^i)), \quad (2.35)$$

$$-T'_3 = \frac{1}{n} \sum_{i=1}^N \delta t \sum_{\sigma \in A^i} |\sigma| |\mathbf{v}_{K,\sigma}^i| u_{\sigma,+}^i (T_n(u_{\sigma,-}^i) - T_n(u_{\sigma,+}^i)), \quad (2.36)$$

$$T'_4 = \frac{1}{n} \sum_{i=1}^N \delta t \sum_{K \in \mathcal{T}} |K| |f_K^i| |T_n(u_K^i)|. \quad (2.37)$$

Due to the definition of  $u_{\mathcal{D}}$  we have

$$T'_4 = \frac{1}{n} \int_0^T \int_{\Omega} |f_{\mathcal{D}}| |T_n(u_{\mathcal{D}})| dx dt,$$

where  $f_{\mathcal{D}} = (f_{\mathcal{T}}^i)_{i=1,\dots,N}$  and  $f_{\mathcal{D}}$  converges strongly to  $f$  in  $L^1(Q)$ . Since  $\forall k > 0$  we have

$$\begin{aligned} T'_4 &= \frac{1}{n} \int_{E_k} |f_{\mathcal{D}}| |T_n(u_{\mathcal{D}})| + \frac{1}{n} \int_{Q \setminus E_k} |f_{\mathcal{D}}| |T_n(u_{\mathcal{D}})| \\ &\leq \int_{E_k} |f_{\mathcal{D}}| + \frac{k}{n} \int_{Q \setminus E_k} |f_{\mathcal{D}}|, \end{aligned}$$

and Corollary 2.3.2 and an equi-integrability argument lead to

$$\lim_{n \rightarrow +\infty} \lim_{\substack{h_{\mathcal{T}} \rightarrow 0 \\ \delta t \rightarrow 0}} T'_4 = 0. \quad (2.38)$$

As far as  $T'_{1,2}$  is concerned since  $(u_K^0)_{K \in \mathcal{T}}$  converges almost everywhere in  $\Omega$  to  $u^0$  we have

$$\lim_{\substack{h_{\mathcal{T}} \rightarrow 0 \\ \delta t \rightarrow 0}} T'_{1,2} = \frac{1}{n} \int_{\Omega} \tilde{T}_n(u^0) dx.$$

Recalling that  $\frac{1}{n} \tilde{T}_n(u^0)$  goes to 0 almost everywhere in  $\Omega$  while  $|\frac{1}{n} \tilde{T}_n(u^0)| \leq |u^0| \in L^1(\Omega)$  the Lebesgue theorem implies that

$$\lim_{n \rightarrow +\infty} \lim_{\substack{h_{\mathcal{T}} \rightarrow 0 \\ \delta t \rightarrow 0}} T'_{1,2} = 0, \quad (2.39)$$

We now prove that

$$-T'_3 \leq \frac{1}{4} T'_2 + \frac{1}{2} T'_{1,1} + R(n, h_{\mathcal{T}}, \delta t)$$

with  $R$  verifying  $\lim_{n \rightarrow +\infty} \overline{\lim}_{\substack{h_{\mathcal{T}} \rightarrow 0 \\ \delta t \rightarrow 0}} R(n, h_{\mathcal{T}}, \delta t) = 0$ .

Because for any  $\sigma \in \mathcal{A}^i$ ,  $|u_{\sigma,+}^i| \geq n$  implies  $|u_{\sigma,-}^i| \geq n$ , we can split  $T'_3$  into a sum on  $\{|u_{\sigma,+}^i| \leq r\}$  and a sum on  $\{r \leq |u_{\sigma,+}^i| \leq n\}$  where  $r > 0$  will be chosen later. It leads to  $-T'_3 \leq I_1 + I_2$  with

$$I_1 = \frac{1}{n} \sum_{i=1}^N \delta t \sum_{\substack{\sigma \in \mathcal{A}^i \\ |u_{\sigma,+}^i| \leq r}} |\sigma| |\mathbf{v}_{K,\sigma}^i| u_{\sigma,+}^i (T_n(u_{\sigma,-}^i) - T_n(u_{\sigma,+}^i)), \quad (2.40)$$

$$I_2 = \frac{1}{n} \sum_{i=1}^N \delta t \sum_{\substack{\sigma \in \mathcal{A}^i \\ r \leq |u_{\sigma,+}^i| \leq n}} |\sigma| |\mathbf{v}_{K,\sigma}^i| u_{\sigma,+}^i (T_n(u_{\sigma,-}^i) - T_n(u_{\sigma,+}^i)). \quad (2.41)$$

For any  $\sigma \in \mathcal{A}$   $|u_{\sigma,+}^i| \geq r$  implies  $|u_{\sigma,-}^i| \geq r$ , thus using Hölder inequality :

$$I_2 \leq \frac{1}{n} I_{2,1} \cdot I_{2,2} \cdot I_{2,3},$$

where

$$\begin{aligned} I_{2,1} &= \left( \sum_{i=1}^N \delta t \sum_{\substack{\sigma \in \mathcal{A}^i \\ r \leq |u_{\sigma,+}^i| \leq n \\ r \leq |u_{\sigma,-}^i|}} |\sigma| d_{\sigma} |\mathbf{v}_{K,\sigma}^i|^{d+2} \right)^{\frac{1}{d+2}}, \\ I_{2,2} &= \left( \sum_{i=1}^N \delta t \sum_{\substack{\sigma \in \mathcal{A}^i \\ r \leq |u_{\sigma,+}^i| \leq n \\ r \leq |u_{\sigma,-}^i|}} |\sigma| d_{\sigma} |T_n(u_{\sigma,+}^i)|^{\frac{2(d+2)}{d}} \right)^{\frac{d}{2(d+2)}}, \\ I_{2,3} &= \left( \sum_{i=1}^N \delta t \sum_{\substack{\sigma \in \mathcal{A}^i \\ r \leq |u_{\sigma,+}^i| \leq n \\ r \leq |u_{\sigma,-}^i|}} \frac{|\sigma|}{d_{\sigma}} [T_n(u_{\sigma,-}^i) - T_n(u_{\sigma,+}^i)]^2 \right)^{\frac{1}{2}}. \end{aligned}$$

Using Jensen's inequality we have

$$I_{2,1} \leq d^{\frac{1}{d+2}} \|\mathbf{v}\|_{L^{d+2}(E_r)}^d. \quad (2.42)$$

We now control  $I_{2,2}$ . If  $d \geq 3$ , using the Sobolev inequality (see Proposition 2.2.5) we get

$$\begin{aligned} & \sum_{\substack{\sigma \in \mathcal{A}^i \\ r \leq |u_{\sigma,+}^i| \leq n \\ r \leq |u_{\sigma,-}^i|}} |\sigma| d_{\sigma} |T_n(u_{\sigma,+}^i)|^{\frac{4}{d}+2} \\ & \leq \left( \sum_{\substack{\sigma \in \mathcal{A}^i \\ r \leq |u_{\sigma,+}^i| \leq n \\ r \leq |u_{\sigma,-}^i|}} |\sigma| d_{\sigma} |T_n(u_{\sigma,+}^i)|^2 \right)^{\frac{2}{d}} \left( \sum_{\substack{\sigma \in \mathcal{A}^i \\ r \leq |u_{\sigma,+}^i| \leq n \\ r \leq |u_{\sigma,-}^i|}} |\sigma| d_{\sigma} |T_n(u_{\sigma,+}^i)|^{2^*} \right)^{\frac{2}{2^*}} \end{aligned}$$

$$\leq C \left( \sup_{1 \leq i \leq N} \sum_{\substack{\sigma \in \mathcal{A}^i \\ r \leq |u_{\sigma,+}^i| \leq n \\ r \leq |u_{\sigma,-}^i|}} |\sigma| d_\sigma |T_n(u_{\sigma,+}^i)|^2 \right)^{\frac{2}{d}} \left( \sum_{\substack{\sigma \in \mathcal{A}^i \\ r \leq |u_{\sigma,+}^i| \leq n \\ r \leq |u_{\sigma,-}^i|}} \frac{|\sigma|}{d_\sigma} |T_n(u_{\sigma,+}^i) - T_n(u_{\sigma,-}^i)|^2 \right)$$

Summing over the time steps and raising to the power  $\frac{d}{2(d+2)}$  we obtain

$$I_{2,2} \leq C \left( \sup_{1 \leq i \leq N} \sum_{\substack{\sigma \in \mathcal{A}^i \\ r \leq |u_{\sigma,+}^i| \leq n \\ r \leq |u_{\sigma,-}^i|}} |\sigma| d_\sigma |T_n(u_{\sigma,+}^i)|^2 \right)^{\frac{1}{d+2}} \left( \sum_{i=1}^N \delta t \sum_{\substack{\sigma \in \mathcal{A}^i \\ r \leq |u_{\sigma,+}^i| \leq n \\ r \leq |u_{\sigma,-}^i|}} \frac{|\sigma|}{d_\sigma} |T_n(u_{\sigma,+}^i) - T_n(u_{\sigma,-}^i)|^2 \right)^{\frac{d}{2(d+2)}}. \quad (2.43)$$

In the case of dimension  $d = 2$ , using a discrete version of Gagliardo-Nirenberg inequality (see [13]) and Hölder inequality allows one to obtain (2.43). Then from (2.42) and (2.43) and using Young's inequality it follows that

$$|I_2| \leq \frac{C \|\mathbf{v}\|_{L^{d+2}(E_r)^d}}{n} \left( \frac{1}{\mu^{d+1}(d+2)} \sup_{1 \leq i \leq N} \sum_{\substack{\sigma \in \mathcal{A}^i \\ r \leq |u_{\sigma,+}^i| \leq n \\ r \leq |u_{\sigma,-}^i|}} |D_\sigma| |T_n(u_{\sigma,+}^i)|^2 + \frac{d+1}{d+2} \mu \sum_{i=1}^N \delta t \sum_{\substack{\sigma \in \mathcal{A}^i \\ r \leq |u_{\sigma,+}^i| \leq n \\ r \leq |u_{\sigma,-}^i|}} \frac{|\sigma|}{d_\sigma} |T_n(u_{\sigma,+}^i) - T_n(u_{\sigma,-}^i)|^2 \right).$$

By Corollary 2.3.2 and since  $\mathbf{v} \in L^{d+2}(Q)^d$ , the absolute continuity of the integral implies that there exists  $r > 0$  (independent of  $\mathcal{T}$  and  $\delta t$ ) such that for all admissible mesh  $\mathcal{T}$  and all time step  $\delta t$

$$\max \left\{ C \|\mathbf{v}\|_{L^{d+2}(E_r)^d} \frac{1}{\mu^{d+1}(d+2)}, C \|\mathbf{v}\|_{L^{d+2}(E_r)^d} \frac{d+1}{d+2} \right\} \leq \frac{1}{4}. \quad (2.44)$$

Therefore

$$|I_2| \leq \frac{1}{n} \left( \frac{1}{4} \sup_{1 \leq i \leq N} \sum_{\substack{\sigma \in \mathcal{A}^i \\ r \leq |u_{\sigma,+}^i| \leq n \\ r \leq |u_{\sigma,-}^i|}} |D_\sigma| |T_n(u_{\sigma,+}^i)|^2 + \frac{\mu}{4} \sum_{i=1}^N \delta t \sum_{\substack{\sigma \in \mathcal{A}^i \\ r \leq |u_{\sigma,+}^i| \leq n \\ r \leq |u_{\sigma,-}^i|}} \frac{|\sigma|}{d_\sigma} |T_n(u_{\sigma,+}^i) - T_n(u_{\sigma,-}^i)|^2 \right). \quad (2.45)$$



We now turn to  $I_1$ . Using the Cauchy-Schwarz inequality and the Young inequality we have

$$\begin{aligned}
|I_1| &= \frac{1}{n} \sum_{i=1}^N \delta t \sum_{\substack{\sigma \in \mathcal{A}^i \\ |u_{\sigma,+}^i| \leq r}} |\sigma| |\mathbf{v}_{K,\sigma}^i| u_{\sigma,+}^i (T_n(u_{\sigma,-}^i) - T_n(u_{\sigma,+}^i)) \\
&\leq \frac{1}{n} \left( \sum_{i=1}^N \delta t \sum_{\substack{\sigma \in \mathcal{A}^i \\ |u_{\sigma,+}^i| \leq r}} |\sigma| d_\sigma |\mathbf{v}_{K,\sigma}^i|^2 \right)^{\frac{1}{2}} \left( \sum_{i=1}^N \delta t \sum_{\substack{\sigma \in \mathcal{A}^i \\ |u_{\sigma,+}^i| \leq r}} \frac{|\sigma|}{d_\sigma} (u_{\sigma,+}^i)^2 (T_n(u_{\sigma,-}^i) - T_n(u_{\sigma,+}^i))^2 \right)^{\frac{1}{2}} \\
&\leq \frac{r}{n} \left( \sum_{i=1}^N \delta t \sum_{\substack{\sigma \in \mathcal{A}^i \\ |u_{\sigma,+}^i| \leq r}} |\sigma| d_\sigma |\mathbf{v}_{K,\sigma}^i|^2 \right)^{\frac{1}{2}} \left( \sum_{i=1}^N \delta t \sum_{\substack{\sigma \in \mathcal{A}^i \\ |u_{\sigma,+}^i| \leq r}} \frac{|\sigma|}{d_\sigma} (T_n(u_{\sigma,-}^i) - T_n(u_{\sigma,+}^i))^2 \right)^{\frac{1}{2}} \\
&\leq \frac{1}{n} \left[ \frac{r^2 d \|\mathbf{v}\|_{L^2(Q)^d}^2}{\mu} + \frac{\mu}{4} \sum_{i=1}^N \delta t \sum_{\substack{\sigma \in \mathcal{A}^i \\ |u_{\sigma,+}^i| \leq r}} \frac{|\sigma|}{d_\sigma} (T_n(u_{\sigma,-}^i) - T_n(u_{\sigma,+}^i))^2 \right]. \tag{2.46}
\end{aligned}$$

Gathering (2.45) and (2.46) we get

$$\begin{aligned}
-T_3' \leq |I_1 + I_2| &\leq \frac{1}{n} \frac{r^2 d \|\mathbf{v}\|_{L^2(Q)^d}^2}{\mu} + \frac{1}{4n} \sup_{1 \leq i \leq N} \sum_{\substack{\sigma \in \mathcal{A}^i \\ r \leq |u_{\sigma,+}^i| \leq n \\ r \leq |u_{\sigma,-}^i|}} |D_\sigma| |T_n(u_{\sigma,+}^i)|^2 \\
&\quad + \frac{\mu}{4n} \sum_{i=1}^N \delta t \sum_{\substack{\sigma \in \mathcal{A}^i \\ |u_{\sigma,+}^i| \leq n}} \frac{|\sigma|}{d_\sigma} (u_K^i - u_L^i) (T_n(u_K^i) - T_n(u_L^i)) \\
&\leq \frac{1}{n} \frac{r^2 d \|\mathbf{v}\|_{L^2(Q)^d}^2}{\mu} + \frac{1}{4n} \sup_{1 \leq i \leq N} \sum_{\substack{\sigma \in \mathcal{A}^i \\ r \leq |u_{\sigma,+}^i| \leq n \\ r \leq |u_{\sigma,-}^i|}} |D_\sigma| |T_n(u_{\sigma,+}^i)|^2 \\
&\quad + \frac{1}{4n} \sum_{i=1}^N \delta t \sum_{\substack{\sigma \in \mathcal{A}^i \\ |u_{\sigma,+}^i| \leq n}} \frac{|\sigma|}{d_\sigma} \lambda(u)_\sigma^i (u_K^i - u_L^i) (T_n(u_K^i) - T_n(u_L^i)) \\
&\leq \frac{1}{n} \frac{r^2 d \|\mathbf{v}\|_{L^2(Q)^d}^2}{\mu} + \frac{1}{4} T_2' + \frac{1}{4n} \sup_{1 \leq i \leq N} \|T_n(u_{\mathcal{T}}^i)\|_{L^2(\Omega)}, \tag{2.47}
\end{aligned}$$

and we deduce from (2.31), (2.39) and (2.47) that

$$\lim_{n \rightarrow +\infty} \overline{\lim}_{\substack{h_{\mathcal{T}} \rightarrow 0 \\ \delta t \rightarrow 0}} \frac{1}{n} \sum_{i=1}^N \sum_{\sigma \in \mathcal{E}} \frac{|\sigma|}{d_\sigma} (u_K^i - u_L^i) (T_n(u_K^i) - T_n(u_L^i)) = 0.$$

and

$$\lim_{n \rightarrow +\infty} \overline{\lim}_{\substack{h_{\mathcal{T}} \rightarrow 0 \\ \delta t \rightarrow 0}} \frac{1}{n} \sum_{i=1}^N \sum_{K \in \mathcal{T}} |K| \int_{u_K^{i-1}}^{u_K^i} (r - u_K^{i-1}) T'_n(r) dr = 0.$$

We now turn to (2.21). By (2.23), we have  $T_1 + T_2 + T_3 = T_4$  with  $M = N$ . Corollary 2.3.2 together with (2.32), (2.39) and (2.22) ensures us that

$$\lim_{n \rightarrow +\infty} \overline{\lim}_{\substack{h_{\mathcal{T}} \rightarrow 0 \\ \delta t \rightarrow 0}} T_1 = 0.$$

By defining the set

$$\mathcal{B}^i = \{\sigma \in \mathcal{E}; u_{\sigma,+}^i (T_n(u_{\sigma,+}^i) - T_n(u_{\sigma,-}^i)) > 0\}$$

we have  $T_3 = T'_3 + T''_3$  with

$$\begin{aligned} \lim_{n \rightarrow +\infty} \overline{\lim}_{\substack{h_{\mathcal{T}} \rightarrow 0 \\ \delta t \rightarrow 0}} T'_3 &= 0, \\ T''_3 &= \frac{1}{n} \sum_{i=1}^N \delta t \sum_{\sigma \in \mathcal{B}^i} |\sigma| |\mathbf{v}_{K,\sigma}^i| u_{\sigma,+}^i (T_n(u_{\sigma,+}^i) - T_n(u_{\sigma,-}^i)). \end{aligned}$$

Still with (2.23), we obtain that

$$\lim_{n \rightarrow +\infty} \overline{\lim}_{\substack{h_{\mathcal{T}} \rightarrow 0 \\ \delta t \rightarrow 0}} T''_3 = 0.$$

Finally, remarking that  $|T_3| = -T'_3 + T''_3$  yields that

$$\lim_{n \rightarrow +\infty} \overline{\lim}_{\substack{h_{\mathcal{T}} \rightarrow 0 \\ \delta t \rightarrow 0}} \frac{1}{n} \sum_{i=1}^N \delta t \sum_{\sigma \in \mathcal{E}} |\sigma| |\mathbf{v}_{K,\sigma}^i| u_{\sigma,+}^i |T_n(u_{\sigma,+}^i) - T_n(u_{\sigma,-}^i)| = 0.$$

□

**Remark 2.3.4** *In the present paper the data  $f$  is given and the "discrete" right-hand side  $f_K^i$  is such that  $f_{\mathcal{D}}$  converges strongly in  $L^1(Q)$  to  $f$ . If we replace  $f_{\mathcal{D}}$  by a  $L^1$  bounded sequence  $f_{\mathcal{D}}$ , that is if we assume that  $f_{\mathcal{D}}$  is only bounded in  $L^1(Q)$ , the attentive reader may notice that Proposition 2.3.1 and then Corollary 2.3.2 hold true. However as in the continuous case, Proposition 2.3.3 should be replaced by a weaker one. Indeed except the precise behavior of  $T'_4$  all the computations and the arguments of the proof hold. Since  $T'_4$  is (only) bounded we obtain  $\forall n \geq 1$*

$$\sum_{i=1}^N \delta t \sum_{\sigma \in \mathcal{E}} \frac{|\sigma|}{d_{\sigma}} \lambda(u)_{\sigma}^i (u_K^i - u_L^i) (T_n(u_K^i) - T_n(u_L^i)) \leq C, \quad (2.48)$$

$$\sum_{i=1}^N \delta t \sum_{\sigma \in \mathcal{E}} |\sigma| |\mathbf{v}_{K,\sigma}^i| u_{\sigma,+}^i |T_n(u_{\sigma,+}^i) - T_n(u_{\sigma,-}^i)| \leq C, \quad (2.49)$$

$$\sum_{i=1}^N \sum_{K \in \mathcal{T}} |K| \int_{u_K^i}^{u_K^{i-1}} (r - u_K^i) T'_n(r) dr \leq C. \quad (2.50)$$

**Corollary 2.3.5 (Estimate on  $T_n(u_{\mathcal{D}})$ )** Let  $\mathcal{T}$  be an admissible mesh and let  $\delta t$  be a time step. If  $u_{\mathcal{D}} = (u_{\mathcal{T}}^i)_{1 \leq i \leq N}$  is a solution to (2.10), then for all  $n > 0$  there exists  $C > 0$  only depending on  $(\Omega, \mathbf{v}, f, n, d)$  such that

$$\sum_{i=1}^N \delta t \|T_n(u_{\mathcal{T}}^i)\|_{1,2,\mathcal{T}} \leq C. \quad (2.51)$$

**Proof** Thanks to (2.20) and the fact that  $\lambda(u) \geq \mu$ , we have

$$\frac{\mu}{2} \sum_{i=1}^N \delta t \sum_{\sigma \in \mathcal{E}} \frac{|\sigma|}{d_{\sigma}} (u_K^i - u_L^i) (T_n(u_K^i) - T_n(u_L^i)) \leq C.$$

Since  $|T_n(u_K^i) - T_n(u_L^i)| \leq |u_K^i - u_L^i|$ , we obtain

$$\frac{\mu}{2} \sum_{i=1}^N \delta t \sum_{\sigma \in \mathcal{E}} \frac{|\sigma|}{d_{\sigma}} (T_n(u_K^i) - T_n(u_L^i))^2 \leq C,$$

which is (2.51). □

Let us introduce the function  $h_n$  defined by

$$h_n(r) = \begin{cases} 0 & \text{if } r \leq -2n, \\ \frac{r}{n} + 2 & \text{if } -2n \leq r \leq -n, \\ 1 & \text{if } -n \leq r \leq n, \\ \frac{-r}{n} + 2 & \text{if } n \leq r \leq 2n, \\ 0 & \text{if } r \geq 2n. \end{cases} \quad (2.52)$$

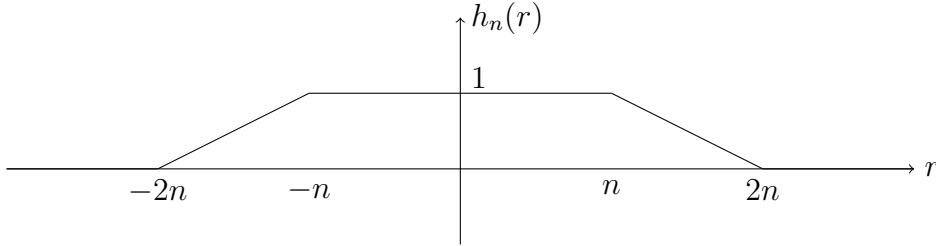


FIGURE 2.2 – The function  $h_n$

**Proposition 2.3.6 (Estimate on  $\tilde{h}_n(u_{\mathcal{D}})$ )** Let  $\mathcal{T}$  be an admissible mesh and let  $\delta t$  be a time step. If  $u_{\mathcal{D}} = (u_{\mathcal{T}}^i)_{1 \leq i \leq N}$  is a solution to (2.10), then for all  $n > 0$  there exists  $C > 0$  only depending on  $(\Omega, \mathbf{v}, f, n, d)$  such that

$$\sum_{i=1}^N \delta t \|\tilde{h}_n(u_{\mathcal{T}}^i)\|_{1,2,\mathcal{T}}^2 \leq C, \quad (2.53)$$

where  $\tilde{h}_n(s) = \int_0^s h_n(r) dr$ .

**Proof** We can perform the proof thanks to (2.51) and the fact that

$$\begin{aligned}
\sum_{i=1}^N \delta t \|\tilde{h}_n(u_{\mathcal{T}}^i)\|_{1,2,\mathcal{T}}^2 &= \sum_{i=1}^N \delta t \sum_{\sigma \in \mathcal{E}} \frac{|\sigma|}{d_\sigma} |\tilde{h}_n(u_K^i) - \tilde{h}_n(u_L^i)|^2 \\
&= \sum_{i=1}^N \delta t \sum_{\sigma \in \mathcal{E}} \frac{|\sigma|}{d_\sigma} |\tilde{h}_n(T_{2n}(u_K^i)) - \tilde{h}_n(T_{2n}(u_L^i))|^2 \\
&\leq \sum_{i=1}^N \delta t \sum_{\sigma \in \mathcal{E}} \frac{|\sigma|}{d_\sigma} |T_{2n}(u_K^i) - T_{2n}(u_L^i)|^2 \\
&= \sum_{i=1}^N \delta t \|T_{2n}(u_{\mathcal{T}}^i)\|_{1,2,\mathcal{T}}^2.
\end{aligned}$$

□

**Proposition 2.3.7 (Estimate on  $\partial_t \tilde{h}_n(u_{\mathcal{D}})$ )** *Let  $\mathcal{T}$  be an admissible mesh and let  $\delta t$  be a time step. If  $u_{\mathcal{D}} = (u_{\mathcal{T}}^i)_{1 \leq i \leq N}$  is a solution to (2.10), then for all  $n > 0$  there exists  $C > 0$  such that*

$$\sum_{i=1}^N \delta t \|\partial_t \tilde{h}_n(u_{\mathcal{T}}^i)\|_{-1,1,\mathcal{T}} \leq C. \quad (2.54)$$

**Proof** Let  $w \in X(\mathcal{T})$ . Multiplying each equation of the scheme by  $h_n(u_K^i)w_K$ , summing over the control volumes and gathering by edges, we get  $T_1^i + T_2^i + T_3^i = T_4^i$  with

$$T_1^i = \sum_{K \in \mathcal{T}} |K| \frac{u_K^i - u_K^{i-1}}{\delta t} h_n(u_K^i) w_K, \quad (2.55)$$

$$T_2^i = \sum_{\sigma \in \mathcal{E}} \frac{|\sigma|}{d_\sigma} \lambda(u)_\sigma^i (u_K^i - u_L^i) (h_n(u_K^i) w_K - h_n(u_L^i) w_L), \quad (2.56)$$

$$T_3^i = \sum_{\sigma \in \mathcal{E}} |\sigma| \mathbf{v}_{K,\sigma}^i u_{\sigma,+}^i (h_n(u_K^i) w_K - h_n(u_L^i) w_L), \quad (2.57)$$

$$T_4^i = \sum_{K \in \mathcal{T}} |K| f_K^i h_n(u_K^i) w_K. \quad (2.58)$$

As far as  $T_1^i$  is concerned we have  $T_1^i = T_{1,1}^i + T_{1,2}^i$  with

$$\begin{aligned}
T_{1,1}^i &= \sum_{K \in \mathcal{T}} |K| \frac{\tilde{h}_n(u_K^i) - \tilde{h}_n(u_K^{i-1})}{\delta t} w_K, \\
T_{1,2}^i &= \sum_{K \in \mathcal{T}} \frac{|K|}{\delta t} w_K \int_{u_K^{i-1}}^{u_K^i} (r - u_K^{i-1}) h'_n(r) dr,
\end{aligned}$$

where  $\tilde{h}_n(s) = \int_0^s h_n(r) dr$ .

By the definition of the dual norm  $\|\cdot\|_{-1,1,\mathcal{T}}$  we have

$$\sup_{\substack{w \in X(\mathcal{T}) \\ w \neq 0}} \frac{1}{\|w\|_{1,\infty,\mathcal{T}}} T_{1,1}^i = \|\partial_t \tilde{h}_n(u_{\mathcal{T}}^i)\|_{-1,1,\mathcal{T}}. \quad (2.59)$$

Since  $h_n$  is bounded by 1 we have  $|T_4^i| \leq \|w\|_{L^\infty(\Omega)} \|f^i\|_{L^1(\Omega)}$  where  $f^i = \frac{1}{\delta t} \int_{t_{i-1}}^{t_i} f dt$ . We thus get

$$\sum_{i=1}^N \delta t \sup_{\substack{w \in X(\mathcal{T}) \\ w \neq 0}} \frac{1}{\|w\|_{1,\infty,\mathcal{T}}} |T_4^i| \leq \|f\|_{L^1(Q)}. \quad (2.60)$$

As far as  $T_2^i$  is concerned we have  $T_2^i = T_{2,1}^i + T_{2,2}^i$  with

$$\begin{aligned} T_{2,1}^i &= \sum_{\sigma \in \mathcal{E}} \frac{|\sigma|}{d_\sigma} \lambda(u)_\sigma^i h_n(u_K^i)(u_K^i - u_L^i)(w_K - w_L), \\ T_{2,2}^i &= \sum_{\sigma \in \mathcal{E}} \frac{|\sigma|}{d_\sigma} \lambda(u)_\sigma^i w_L(u_K^i - u_L^i)(h_n(u_K^i) - h_n(u_L^i)). \end{aligned}$$

In view of the definition of  $h_n$  we get

$$|T_{2,2}^i| \leq \frac{\|w\|_{L^\infty(\Omega)}}{n} \sum_{\sigma \in \mathcal{E}} \frac{|\sigma|}{d_\sigma} \lambda(u)_\sigma^i (u_K^i - u_L^i) (T_{2n}(u_K^i) - T_{2n}(u_L^i)),$$

so that taking the sup on  $X(\mathcal{T})$  and summing over the time steps give

$$\begin{aligned} &\sum_{i=1}^N \delta t \sup_{\substack{w \in X(\mathcal{T}) \\ w \neq 0}} \frac{1}{\|w\|_{1,\infty,\mathcal{T}}} |T_{2,2}^i| \\ &\leq \frac{1}{n} \sum_{i=1}^N \delta t \sum_{\sigma \in \mathcal{E}} \frac{|\sigma|}{d_\sigma} \lambda(u)_\sigma^i (u_K^i - u_L^i) (T_{2n}(u_K^i) - T_{2n}(u_L^i)), \end{aligned} \quad (2.61)$$

which is bounded thanks to the estimate (2.20). We now turn to  $T_{2,1}^i$ , which reads as

$$T_{2,1}^i = T_{2,1,1}^i + T_{2,1,2}^i \text{ where}$$

$$\begin{aligned} T_{2,1,1}^i &= \sum_{\sigma \in \mathcal{E}} \frac{|\sigma|}{d_\sigma} \lambda(u)_\sigma^i h_n(u_K^i) h_{2n}(u_L^i) (u_K^i - u_L^i) (w_K - w_L), \\ T_{2,1,2}^i &= \sum_{\sigma \in \mathcal{E}} \frac{|\sigma|}{d_\sigma} \lambda(u)_\sigma^i h_n(u_K^i) (1 - h_{2n}(u_L^i)) (u_K^i - u_L^i) (w_K - w_L). \end{aligned}$$

Since  $h_n(u_K^i)(1 - h_{2n}(u_L^i)) = h_n(u_K^i)(h_{2n}(u_K^i) - h_{2n}(u_L^i))$  we have

$$|T_{2,1,2}^i| \leq \frac{\|w\|_{L^\infty(\Omega)}}{n} \sum_{\sigma \in \mathcal{E}} \frac{|\sigma|}{d_\sigma} \lambda(u)_\sigma^i (T_{4n}(u_K^i) - T_{4n}(u_L^i)) (u_K^i - u_L^i),$$

which gives thanks to (2.20)

$$\sum_{i=1}^N \delta t \sup_{\substack{w \in X(\mathcal{T}) \\ w \neq 0}} \frac{1}{\|w\|_{1,\infty,\mathcal{T}}} |T_{2,1,2}^i| \leq C. \quad (2.62)$$

As far as  $T_{2,1,1}^i$  is concerned, we have

$$T_{2,1,1}^i = \sum_{\sigma \in \mathcal{E}} |\sigma| \lambda(u)_\sigma^i h_n(u_K^i) h_{2n}(u_L^i) (T_{4n}(u_K^i) - T_{4n}(u_L^i)) \frac{w_K - w_L}{d_\sigma},$$

and by the definition of  $\lambda(u)_\sigma^i$ , we can write

$$|T_{2,1,1}^i| \leq \|w\|_{1,\infty,\mathcal{T}} \left( \max_{r \in [-4n, 4n]} |\lambda(r)| \right) \|T_{4n}(u_{\mathcal{T}}^i)\|_{1,1,\mathcal{T}}.$$

Thus, we deduce that

$$\sum_{i=1}^N \delta t \sup_{\substack{w \in X(\mathcal{T}) \\ w \neq 0}} \frac{1}{\|w\|_{1,\infty,\mathcal{T}}} |T_{2,1,1}^i| \leq C. \quad (2.63)$$

Gathering (2.62) and (2.63) we get

$$\sum_{i=1}^N \delta t \sup_{\substack{w \in X(\mathcal{T}) \\ w \neq 0}} \frac{1}{\|w\|_{1,\infty,\mathcal{T}}} |T_2^i| \leq C. \quad (2.64)$$

The term  $T_3^i$  can be rewritten as

$$T_3^i = \sum_{\sigma \in \mathcal{E}} |\sigma| |\mathbf{v}_{K,\sigma}^i| u_{\sigma,+}^i (h_n(u_{\sigma,+}^i) w_{\sigma,+} - h_n(u_{\sigma,-}^i) w_{\sigma,-}),$$

where  $w_{\sigma,+} = w_K$  if  $\mathbf{v}_{K,\sigma}^i \geq 0$ ,  $w_L$  otherwise and  $(w_{\sigma,+}, w_{\sigma,-}) = (w_K, w_L)$ . We have  $T_3^i = T_{3,1}^i + T_{3,2}^i$  with

$$\begin{aligned} T_{3,1}^i &= \sum_{\sigma \in \mathcal{E}} |\sigma| |\mathbf{v}_{K,\sigma}^i| u_{\sigma,+}^i h_n(u_{\sigma,+}^i) (w_{\sigma,+} - w_{\sigma,-}), \\ T_{3,2}^i &= \sum_{\sigma \in \mathcal{E}} |\sigma| |\mathbf{v}_{K,\sigma}^i| u_{\sigma,+}^i w_{\sigma,-} (h_n(u_{\sigma,+}^i) - h_n(u_{\sigma,-}^i)). \end{aligned}$$

By the definition of  $h_n$  we have

$$\begin{aligned} |T_{3,1}^i| &\leq \sum_{\sigma \in \mathcal{E}} |\sigma| d_\sigma |\mathbf{v}_{K,\sigma}^i| |T_{2n}(u_{\sigma,+}^i)| \frac{|w_{\sigma,+} - w_{\sigma,-}|}{d_\sigma} \\ &\leq \|w\|_{1,\infty,\mathcal{T}} \sum_{\sigma \in \mathcal{E}} |\sigma| d_\sigma |\mathbf{v}_{K,\sigma}^i| |T_{2n}(u_{\sigma,+}^i)|. \end{aligned}$$

Using Hölder and Poincaré inequalities we get

$$\begin{aligned} \sum_{i=1}^N \delta t \sup_{\substack{w \in X(\mathcal{T}) \\ w \neq 0}} \frac{1}{\|w\|_{1,\infty,\mathcal{T}}} |T_{3,1}^i| &\leq \sum_{i=1}^N \delta t \sum_{\sigma \in \mathcal{E}} |\sigma| d_\sigma |\mathbf{v}_{K,\sigma}^i| |T_{2n}(u_{\sigma,+}^i)| \\ &\leq \left( \sum_{i=1}^N \delta t \sum_{\sigma \in \mathcal{E}} |\sigma| d_\sigma |\mathbf{v}_{K,\sigma}^i|^2 \right)^{\frac{1}{2}} \left( \sum_{i=1}^N \delta t \sum_{\sigma \in \mathcal{E}} |\sigma| d_\sigma |T_{2n}(u_{\sigma,+}^i)|^2 \right)^{\frac{1}{2}} \\ &\leq d^{\frac{1}{2}} \|\mathbf{v}\|_{L^2(Q)^d} C \left( \sum_{i=1}^N \delta t \|T_{2n}(u_{\mathcal{T}}^i)\|_{1,2,\mathcal{T}}^2 \right)^{\frac{1}{2}}. \quad (2.65) \end{aligned}$$

For  $T_{3,2}^i$  it is sufficient to observe that

$$|T_{3,2}^i| \leq \frac{\|w\|_{L^\infty(\Omega)}}{n} \sum_{\sigma \in \mathcal{E}} |\sigma| |\mathbf{v}_{K,\sigma}^i| |u_{\sigma,+}^i| |T_{2n}(u_{\sigma,+}^i) - T_{2n}(u_{\sigma,-}^i)|,$$

from which it follows

$$\sum_{i=1}^N \delta t \sup_{\substack{w \in X(\mathcal{T}) \\ w \neq 0}} \frac{1}{\|w\|_{1,\infty,\mathcal{T}}} |T_{3,2}^i| \leq \frac{1}{n} \sum_{i=1}^N \delta t \sum_{\sigma \in \mathcal{E}} |\sigma| |\mathbf{v}_{K,\sigma}^i| |u_{\sigma,+}^i| |T_{2n}(u_{\sigma,+}^i) - T_{2n}(u_{\sigma,-}^i)|, \quad (2.66)$$

which is bounded thanks to (2.21).

It remains to control  $T_{1,2}^i$ . Noticing that

$$\left| \int_{u_K^{i-1}}^{u_K^i} (r - u_K^{i-1}) h'_n(r) dr \right| \leq \int_{u_K^{i-1}}^{u_K^i} (r - u_K^{i-1}) T'_{2n}(r) dr, \quad (2.67)$$

(2.67) and (2.22) yield that

$$\sum_{i=1}^N \delta t \sup_{\substack{w \in X(\mathcal{T}) \\ w \neq 0}} \frac{1}{\|w\|_{1,\infty,\mathcal{T}}} |T_{1,2}^i| \leq C. \quad (2.68)$$

Gathering (2.59) to (2.68), we obtain the estimate

$$\sum_{i=1}^N \delta t \|\partial_t \tilde{h}_n(u_{\mathcal{T}}^i)\|_{-1,1,\mathcal{T}} \leq C. \quad (2.69)$$

□

We are now in position to extract a convergent subsequence to  $T_n(u_{\mathcal{D}})$ .

**Proposition 2.3.8** *Let  $(\mathcal{T}_m)_{m \geq 1}$  be a sequence of admissible meshes such that there exists  $\xi > 0$  satisfying*

$$\text{for all } m \geq 1 \text{ for all } K \in \mathcal{T} \text{ and all } \sigma \in \mathcal{E}(K), \quad d_{K,\sigma} \geq \xi d_\sigma,$$

*and let  $\delta t_m$  be a sequence of time steps. If  $u_{\mathcal{D}_m} = (u_{\mathcal{T}_m}^i)_{1 \leq i \leq N}$  is a solution to (2.10) with  $\mathcal{T} = \mathcal{T}_m$  and  $\delta t = \delta t_m$ , then there exists a measurable function  $u$  finite a.e. in  $Q$  such that, up to a subsequence,  $T_n(u_{\mathcal{D}_m})$  converges to  $T_n(u)$  weakly in  $L^2(0, T; H_0^1(\Omega))$  and  $u_{\mathcal{D}_m}$  converges almost everywhere to  $u$ .*

In order to prove this proposition we use a discrete version of the Aubin-Simon Lemma proved in [47] and in [48].

**Proof** Thanks to estimate (2.51) we can extract a subsequence, still indexed by  $m$ , such that

$$T_n(u_{\mathcal{D}_m}) \rightharpoonup v_n \text{ in } L^2(0, T; H_0^1(\Omega)) \quad (2.70)$$

in the finite volume sense.

Moreover, since we have estimates (2.53) and (2.54), we also apply the discrete version of the Aubin-Simon Lemma to the function  $\tilde{h}_n(u_{\mathcal{D}_m})$ . It implies that for any  $n > 0$ ,  $\tilde{h}_n(u_{\mathcal{D}_m})$  lies in a compact set of  $L^1(Q)$ . Using the diagonal process, we can deduce that there exists a subsequence of  $u_{\mathcal{D}_m}$ , still indexed by  $m$ , such that

$\tilde{h}_n(u_{\mathcal{D}_m})$  converges almost everywhere to a measurable function  $w_n$  defined on  $Q$  for any  $n > 0$ . Due to the definition of  $\tilde{h}_n$ ,  $T_n(u_{\mathcal{D}_m}) = T_n(\tilde{h}_n(u_{\mathcal{D}_m}))$ , thus the sequence  $T_n(u_{\mathcal{D}_m})$  converges almost everywhere in  $Q$  to the measurable function  $v_n = T_n(w_n)$ .

Using Corollary 2.3.2 we now prove that  $u_{\mathcal{D}_m}$  is a Cauchy sequence in measure. Let  $\omega > 0$ . For all  $n > 0$  and all sequences of admissible meshes and time steps  $(\mathcal{T}_m, \delta t_m)_{m \geq 1}$  and  $(\mathcal{T}_p, \delta t_p)_{p \geq 1}$ , we have

$$\{|u_{\mathcal{D}_m} - u_{\mathcal{D}_p}| > \omega\} \subset \{|u_{\mathcal{D}_m}| > n\} \cup \{|u_{\mathcal{D}_p}| > n\} \cup \{|T_n(u_{\mathcal{D}_m}) - T_n(u_{\mathcal{D}_p})| > \omega\}.$$

Let  $\varepsilon > 0$  fixed. By (2.3.2), let  $n > 0$  such that, for all admissible meshes and time steps  $\mathcal{T}_m, \delta t_m$  and  $\mathcal{T}_p, \delta t_p$

$$\text{meas}(\{|u_{\mathcal{D}_m}| > n\}) + \text{meas}(\{|u_{\mathcal{D}_p}| > n\}) < \frac{\varepsilon}{2}.$$

Once  $n$  is chosen, we use the fact that  $T_n(u_{\mathcal{D}_m})$  is a Cauchy sequence in measure, thus

$$\exists h_0 > 0; \forall h_{\mathcal{D}_m}, h_{\mathcal{D}_p} < h_0 \quad \text{meas}(\{|T_n(u_{\mathcal{D}_m}) - T_n(u_{\mathcal{D}_p})| > \omega\}) \leq \frac{\varepsilon}{2}.$$

Therefore, we deduce that  $\forall h_{\mathcal{D}_m}, h_{\mathcal{D}_p} < h_0$

$$\text{meas}\{|u_{\mathcal{D}_m} - u_{\mathcal{D}_p}| > \omega\} < \varepsilon.$$

Hence  $u_{\mathcal{D}_m}$  is a Cauchy sequence in measure. It follows that, up to a subsequence still indexed by  $\mathcal{D}_m$ , there exists a measurable function  $u$  such that  $u_{\mathcal{D}_m} \rightarrow u$  a.e. in  $Q$ . Thus  $v_n = T_n(u)$  and  $w_n = \tilde{h}_n(u)$ . Due to Corollary 2.3.2  $u$  is finite almost everywhere. Moreover from convergences obtained before we get that  $T_n(u) \in L^2(0, T; H_0^1(\Omega))$  and

$$\begin{aligned} \nabla_{\mathcal{D}} T_n(u_{\mathcal{D}}) &\rightharpoonup \nabla T_n(u) \text{ in } (L^2(Q))^d, \\ T_n(u_{\mathcal{D}}) &\rightarrow T_n(u) \text{ almost everywhere in } Q. \end{aligned}$$

□

**Remark 2.3.9** *Following Remark 2.3.4 it is worth noting that Corollaries 2.3.5, 2.3.6 and Proposition 2.3.7 hold true if instead of (2.20), (2.21) and (2.22) (see Proposition 2.3.3) we have (2.48), (2.49) and (2.50) (which are weaker). It follows that if we replace in the scheme  $f_{\mathcal{D}}$  (an approximation of a fixed data  $f$ ) by a sequence  $f_{\mathcal{D}}$  which is only bounded in  $L^1(Q)$  Proposition 2.3.8 holds true.*

## 2.4 Convergence analysis

Let us now state the main result of this paper.

**Theorem 2.4.1** *If  $\mathcal{T}$  is an admissible mesh and  $\delta t$  a time step, then there exists at least one solution to (2.10).*

*If  $(\mathcal{T}_m)_{m \geq 1}$  is a sequence of admissible meshes such that there exists  $\xi > 0$  satisfying*

$$\text{for all } m \geq 1 \text{ for all } K \in \mathcal{T} \text{ and all } \sigma \in \mathcal{E}(K), \quad d_{K,\sigma} \geq \xi d_\sigma,$$



and such that  $h_{\mathcal{T}_m} \rightarrow 0$  and if  $(\delta t_m)_{m \geq 1}$  is a sequence of time steps such that  $\delta t_m \rightarrow 0$ , then if  $u_{\mathcal{D}_m} = (u_{\mathcal{T}_m}^i)_{1 \leq i \leq N^m}$  is the solution to (2.10) with  $\mathcal{T} = \mathcal{T}_m$  and  $\delta t_m = \delta t$ , up to a subsequence,  $u_{\mathcal{D}_m}$  converges to  $u$  in the sense that for all  $n > 0$ ,  $T_n(u_{\mathcal{D}_m})$  converges weakly to  $T_n(u)$  in  $L^2(0, T; H_0^1(\Omega))$  and  $u_{\mathcal{D}_m}$  converges to  $u$  almost everywhere in  $Q$ , where  $u$  is a renormalized solution of (2.1).

Moreover if

1.  $\lambda$  is a local Lipschitz continuous function, or
2.  $\mathbf{v}(x, t) = \mathbf{v}(x)$  is independent of  $t$

then the whole sequence  $u_{\mathcal{D}_m}$  converges almost everywhere to  $u$  where  $u$  is the unique renormalized solution of (2.1).

### Proof of Theorem 2.4.1

Concerning the existence of a solution to the scheme, using the Brouwer theorem, the proof is quite easy when  $\lambda$  is a bounded function (see [44]). In the case where  $\lambda$  is not bounded, we replace  $\lambda$  by  $\lambda(T_{1/\varepsilon})$  and we show with the Brouwer theorem that the modified scheme admits at least one solution  $u_{\mathcal{D}}^\varepsilon$ . Since we have a bound on  $u_{\mathcal{D}}^\varepsilon$  (uniform in  $\varepsilon$ ), for  $\varepsilon$  small enough,  $\lambda(T_{1/\varepsilon}(u_{\mathcal{D}}^\varepsilon)) = \lambda(u_{\mathcal{D}}^\varepsilon)$ . Then the scheme (2.10) admits at least a solution. We then split the rest of the proof in four steps. In Step 1 we pass to the limit in the scheme as the size of the mesh and the time step tend to zero. It is worth noting that we take in the scheme a discrete version of what is a test function in the renormalized formulation. This is to say that we multiply each equation of the scheme by a discrete version of  $h_n(u)\varphi$  where  $\varphi \in \mathcal{C}_c^\infty(\Omega \times [0, T])$  and  $h_n$  is defined by (2.52). Step 2 is devoted to pass to the limit as  $n$  tends to infinity and to show that the renormalized formulation (2.6) holds. As in the continuous case the main difficulty is the lack of regularity of the time derivative  $\frac{\partial u}{\partial t}$ . We use here Steklov averages and the method developed in [20] (see also [63]). In Step 3 we conclude the convergence analysis, showing that the limit  $u$  satisfies the estimate on the energy (2.5). In Step 4 we recall that under assumption 1 or 2 the uniqueness of the renormalized solution provides that the whole sequence converges to the renormalized solution.

**Step 1** We prove that the limit  $u$  satisfies (2.84). Let  $\varphi \in \mathcal{C}_c^\infty(\Omega \times [0, T])$  and  $h_n$  the function defined by (2.52). We denote by  $\varphi_{\mathcal{D}}$  the function defined by  $\varphi_K^i = \varphi(x_K, t_i)$  for all  $K \in \mathcal{T}$  and all  $1 \leq i \leq N$ .

Multiplying each equation of the scheme by  $\varphi_K^{i-1} h_n(u_K^i)$  (which is a discrete version of the test function used in the renormalized formulation), summing over the control volumes, gathering by edges and summing over the time steps, we get  $T_1 + T_2 + T_3 = T_4$  with

$$\begin{aligned} T_1 &= \sum_{i=1}^N \sum_{K \in \mathcal{T}} |K| (u_K^i - u_K^{i-1}) h_n(u_K^i) \varphi_K^{i-1}, \\ T_2 &= \sum_{i=1}^N \delta t \sum_{\sigma \in \mathcal{E}} \frac{|\sigma|}{d_\sigma} \lambda(u)_\sigma (u_K^i - u_L^i) (h_n(u_K^i) \varphi_K^{i-1} - h_n(u_L^i) \varphi_L^{i-1}), \\ T_3 &= \sum_{i=1}^N \delta t \sum_{\sigma \in \mathcal{E}} |\sigma| \mathbf{v}_{K,\sigma}^i u_{\sigma,+}^i (h_n(u_K^i) \varphi_K^{i-1} - h_n(u_L^i) \varphi_L^{i-1}), \end{aligned}$$

$$T_4 = \sum_{i=1}^N \delta t \sum_{K \in \mathcal{T}} |K| f_K^i h_n(u_K^i) \varphi_K^{i-1}.$$

We now pass to the limit as  $h_{\mathcal{T}} \rightarrow 0$  and  $\delta t \rightarrow 0$ .

As far as the lower order term  $T_4$  is concerned, we have  $T_4 = T_{4,1} + T_{4,2}$  with

$$\begin{aligned} T_{4,1} &= \sum_{i=1}^N \sum_{K \in \mathcal{T}} \int_{t_{i-1}}^{t_i} \int_K f h_n(u_K^i) \varphi_K^i dx dt, \\ T_{4,2} &= \sum_{i=1}^N \sum_{K \in \mathcal{T}} \int_{t_{i-1}}^{t_i} \int_K f h_n(u_K^i) (\varphi_K^{i-1} - \varphi_K^i) dx dt. \end{aligned}$$

By the regularity of  $\varphi$ , we have  $\varphi_{\mathcal{D}} \rightarrow \varphi$  uniformly on  $Q$  as  $h_{\mathcal{T}} \rightarrow 0$  and  $\delta t \rightarrow 0$ . Since  $h_n(u_{\mathcal{D}}) \rightarrow h_n(u)$  a.e. and  $L^\infty$  weak- $\star$ ,  $\varphi_{\mathcal{D}} \rightarrow \varphi$  uniformly and  $|f \varphi_{\mathcal{T}} h_n(u_{\mathcal{D}})| \leq C_\varphi |f| \in L^1(Q)$ , the Lebesgue theorem ensures that

$$\begin{aligned} \lim_{\substack{h_{\mathcal{T}} \rightarrow 0 \\ \delta t \rightarrow 0}} T_{4,1} &= \lim_{\substack{h_{\mathcal{T}} \rightarrow 0 \\ \delta t \rightarrow 0}} \int_{\Omega} f \varphi_{\mathcal{D}} h_n(u_{\mathcal{D}}) dx dt \\ &= \int_0^T \int_{\Omega} f \varphi h_n(u) dx dt. \end{aligned} \quad (2.71)$$

In view of the definition of  $h_n$  and the regularity of  $\varphi$ ,

$$|T_{4,2}| \leq c \|\varphi\|_{\infty} \delta t \|f\|_{L^1(Q)},$$

thus

$$\lim_{\substack{h_{\mathcal{T}} \rightarrow 0 \\ \delta t \rightarrow 0}} T_{4,2} = 0. \quad (2.72)$$

We now study the convergence of the diffusion term. We have

$$\begin{aligned} T_2 &= \sum_{i=1}^N \delta t \sum_{\sigma \in \mathcal{E}} \frac{|\sigma|}{d_\sigma} \lambda(u)_\sigma^i (u_K^i - u_L^i) (h_n(u_K^i) \varphi_K^{i-1} - h_n(u_L^i) \varphi_L^{i-1}) \\ &= T_{2,1} + T_{2,2} \end{aligned}$$

with

$$\begin{aligned} T_{2,1} &= \sum_{i=1}^N \delta t \sum_{\sigma \in \mathcal{E}} \frac{|\sigma|}{d_\sigma} \lambda(u)_\sigma^i h_n(u_K^i) (u_K^i - u_L^i) (\varphi_K^{i-1} - \varphi_L^{i-1}), \\ T_{2,2} &= \sum_{i=1}^N \delta t \sum_{\sigma \in \mathcal{E}} \frac{|\sigma|}{d_\sigma} \lambda(u)_\sigma^i \varphi_L^{i-1} (u_K^i - u_L^i) (h_n(u_K^i) - h_n(u_L^i)). \end{aligned}$$

Not assuming any bound on the function  $\lambda$  gives additional technical difficulties in order to control  $T_{2,1}$  and  $T_{2,2}$ . In view of the definition of  $h_n$  and  $\lambda$  we get

$$|T_{2,2}| \leq \frac{c \|\varphi\|_{\infty}}{n} \sum_{i=1}^N \delta t \sum_{\sigma \in \mathcal{E}} \frac{|\sigma|}{d_\sigma} \lambda(u)_\sigma^i (u_K^i - u_L^i) (T_{2n}(u_K^i) - T_{2n}(u_L^i)).$$

From the discrete estimate on the energy (see (2.20) in Proposition 2.3.3), it follows that

$$\lim_{n \rightarrow +\infty} \overline{\lim}_{\substack{h_{\mathcal{T}} \rightarrow 0 \\ \delta t \rightarrow 0}} T_{2,2} = 0. \quad (2.73)$$

As far as  $T_{2,1}$  is concerned we have  $T_{2,1} = T_{2,1,1} + T_{2,1,2}$  with

$$\begin{aligned} T_{2,1,1} &= \sum_{i=1}^N \delta t \sum_{\sigma \in \mathcal{E}} \frac{|\sigma|}{d_{\sigma}} \lambda(u)_{\sigma}^i h_n(u_K^i) h_{2n}(u_L^i) (u_K^i - u_L^i) (\varphi_K^{i-1} - \varphi_L^{i-1}), \\ T_{2,1,2} &= \sum_{i=1}^N \delta t \sum_{\sigma \in \mathcal{E}} \frac{|\sigma|}{d_{\sigma}} \lambda(u)_{\sigma}^i h_n(u_K^i) (1 - h_{2n}(u_L^i)) (u_K^i - u_L^i) (\varphi_K^{i-1} - \varphi_L^{i-1}). \end{aligned}$$

Observing that  $h_n(u_K^i) (1 - h_{2n}(u_L^i)) = h_n(u_K^i) (h_{2n}(u_K^i) - h_{2n}(u_L^i))$  gives

$$|T_{2,1,2}| \leq \frac{c \|\varphi\|_{\infty}}{n} \sum_{i=1}^N \delta t \sum_{\sigma \in \mathcal{E}} \frac{|\sigma|}{d_{\sigma}} \lambda(u)_{\sigma}^i (T_{4n}(u_K^i) - T_{4n}(u_L^i)) (u_K^i - u_L^i),$$

and from (2.20) it follows that

$$\lim_{n \rightarrow +\infty} \overline{\lim}_{\substack{h_{\mathcal{T}} \rightarrow 0 \\ \delta t \rightarrow 0}} T_{2,1,2} = 0. \quad (2.74)$$

For  $T_{2,1,1}$  we write  $T_{2,1,1} = I + II$  with

$$\begin{aligned} I &= \sum_{i=1}^N \delta t \sum_{\sigma \in \mathcal{E}} \frac{|\sigma|}{d_{\sigma}} \lambda(u)_{\sigma}^i \frac{h_n(u_K^i) h_{2n}(u_L^i) + h_n(u_L^i) h_{2n}(u_K^i)}{2} (u_K^i - u_L^i) (\varphi_K^{i-1} - \varphi_L^{i-1}), \\ II &= \sum_{i=1}^N \delta t \sum_{\sigma \in \mathcal{E}} \frac{|\sigma|}{d_{\sigma}} \lambda(u)_{\sigma}^i \frac{h_n(u_K^i) h_{2n}(u_L^i) - h_n(u_L^i) h_{2n}(u_K^i)}{2} (u_K^i - u_L^i) (\varphi_K^{i-1} - \varphi_L^{i-1}). \end{aligned}$$

Since

$$\begin{aligned} &h_n(u_K^i) h_{2n}(u_L^i) - h_n(u_L^i) h_{2n}(u_K^i) \\ &= h_n(u_K^i) (h_{2n}(u_L^i) - h_{2n}(u_K^i)) + (h_n(u_K^i) - h_n(u_L^i)) h_{2n}(u_K^i), \end{aligned}$$

and since

$$\begin{aligned} |h_{2n}(u_L^i) - h_{2n}(u_K^i)| &\leq \frac{1}{2n} |T_{4n}(u_L^i) - T_{4n}(u_K^i)|, \\ |h_n(u_L^i) - h_n(u_K^i)| &\leq \frac{1}{n} |T_{2n}(u_L^i) - T_{2n}(u_K^i)|, \end{aligned}$$

estimate (2.20) leads to

$$|II| \leq \|\varphi\|_{\infty} \omega(n, h_{\mathcal{T}}, \delta t), \quad (2.75)$$

with

$$\lim_{n \rightarrow +\infty} \overline{\lim}_{\substack{h_{\mathcal{T}} \rightarrow 0 \\ \delta t \rightarrow 0}} \omega(n, h_{\mathcal{T}}, \delta t) = 0.$$

As far as  $I$  is concerned, we have  $I = I_1 + I_2$  with

$$I_1 = \sum_{i=1}^N \delta t \sum_{\sigma \in \mathcal{E}} \frac{|\sigma|}{d_\sigma} \lambda(u)_\sigma^i \frac{h_n(u_K^i)h_{2n}(u_L^i) + h_n(u_L^i)h_{2n}(u_K^i)}{2} (u_K^i - u_L^i)(\varphi_K^i - \varphi_L^i),$$

$$I_2 = \sum_{i=1}^N \delta t \sum_{\sigma \in \mathcal{E}} \frac{|\sigma|}{d_\sigma} \lambda(u)_\sigma^i \frac{h_n(u_K^i)h_{2n}(u_L^i) + h_n(u_L^i)h_{2n}(u_K^i)}{2} \\ \times (u_K^i - u_L^i)(\varphi_K^{i-1} - \varphi_K^i + \varphi_L^i - \varphi_L^{i-1}),$$

Both  $u_K^i$  and  $u_L^i$  being truncated in  $I_2$ , we get

$$|I_2| \leq C_\varphi(h_{\mathcal{T}} + \delta t) \left( \max_{r \in [-4n, 4n]} |\lambda(r)| \right) \sum_{i=1}^N \delta t \|T_{4n}(u_{\mathcal{T}}^i)\|_{1,1,\mathcal{T}},$$

thus

$$\lim_{\substack{h_{\mathcal{T}} \rightarrow 0 \\ \delta t \rightarrow 0}} I_2 = 0. \quad (2.76)$$

Let us define  $\bar{h}_n(x, t)$  and  $\bar{\lambda}(x, t) \forall t \in [t_{i-1}, t_i]$ ,  $1 \leq i \leq N$ ,  $\forall x \in D_\sigma$ ,  $\sigma \in \mathcal{E}$ , by

$$\bar{h}_n(x, t) = \frac{h_n(u_K^i)h_{2n}(u_L^i) + h_n(u_L^i)h_{2n}(u_K^i)}{2},$$

$$\bar{\lambda}(x, t) = \lambda(T_{4n}(u))_\sigma^i.$$

We then write  $I_1 = I_{1,1} + I_{1,2}$  with

$$I_{1,1} = \int_0^T \int_\Omega \bar{\lambda} \bar{h}_n \nabla_{\mathcal{D}} T_{4n}(u_{\mathcal{D}}) \cdot \nabla \varphi \, dx \, dt,$$

$$I_{1,2} = \sum_{i=1}^N \delta t \sum_{\sigma \in \mathcal{E}} d|D_\sigma| \lambda(u)_\sigma^i \frac{h_n(u_K^i)h_{2n}(u_L^i) + h_n(u_L^i)h_{2n}(u_K^i)}{2} \\ \frac{T_{4n}(u_K^i) - T_{4n}(u_L^i)}{d_\sigma} C(\varphi, \mathcal{T}, \delta t),$$

where  $C(\varphi, \mathcal{T}, \delta t) = \frac{\varphi_K^i - \varphi_L^i}{d_\sigma} - \frac{1}{\delta t |D_\sigma|} \int_{t_{i-1}}^{t_i} \int_{D_\sigma} \nabla \varphi \cdot \eta_{K,\sigma} \, dx \, dt$ . Thanks to the regularity of  $\varphi$ , we have

$$|I_{1,2}| \leq C_\varphi(h_{\mathcal{T}} + \delta t) \left( \max_{r \in [-4n, 4n]} |\lambda(r)| \right) \sum_{i=1}^N \delta t \|T_{4n}(u_{\mathcal{T}}^i)\|_{1,1,\mathcal{T}},$$

which implies

$$\lim_{\substack{h_{\mathcal{T}} \rightarrow 0 \\ \delta t \rightarrow 0}} I_{1,2} = 0. \quad (2.77)$$

Since  $\lambda(T_{4n}(\cdot))$  is continuous and bounded, an equi-integrability argument allows one to prove that  $\bar{\lambda}$  converges to  $\lambda(T_{4n}(u))$  in  $L^\infty$  weak- $\star$ . In view of the definition of  $\bar{h}_n$  and since  $\bar{h}_n$  is constant on each diamond cell  $D_\sigma$ , we deduce that  $\bar{h}_n \rightarrow h_n(u)$

in  $L^q(Q)$  for all  $q \in [2, +\infty)$ . Moreover  $\nabla_{\mathcal{D}} T_{4n}(u_{\mathcal{D}})$  tends to  $\nabla T_{4n}(u)$  weakly in  $L^2(0, T; L^2(\Omega)^d)$ . Since  $\varphi$  belongs to  $\mathcal{C}_c^\infty(\Omega \times [0, T])$  we obtain

$$\lim_{\substack{h_T \rightarrow 0 \\ \delta t \rightarrow 0}} I_1 = \int_0^T \int_{\Omega} \lambda(u) h_n(u) \nabla T_{4n}(u) \cdot \nabla \varphi \, dx \, dt. \quad (2.78)$$

For the convection term we have

$$\begin{aligned} T_3 &= \sum_{i=1}^N \delta t \sum_{\sigma \in \mathcal{E}} |\sigma| \mathbf{v}_{K,\sigma}^i u_{\sigma,+}^i (\varphi_K^{i-1} h_n(u_K^i) - \varphi_L^{i-1} h_n(u_L^i)) \\ &= \sum_{i=1}^N \delta t \sum_{\substack{\sigma \in \mathcal{E} \\ \mathbf{v}_{K,\sigma}^i \geq 0}} |\sigma| \mathbf{v}_{K,\sigma}^i u_{\sigma,+}^i h_n(u_{\sigma,+}^i) (\varphi_K^{i-1} - \varphi_L^{i-1}) \\ &\quad + \sum_{i=1}^N \delta t \sum_{\substack{\sigma \in \mathcal{E} \\ \mathbf{v}_{K,\sigma}^i \geq 0}} |\sigma| \mathbf{v}_{K,\sigma}^i u_{\sigma,+}^i \varphi_L^{i-1} (h_n(u_{\sigma,+}^i) - h_n(u_{\sigma,-}^i)) \\ &\quad - \sum_{i=1}^N \delta t \sum_{\substack{\sigma \in \mathcal{E} \\ \mathbf{v}_{K,\sigma}^i < 0}} |\sigma| \mathbf{v}_{K,\sigma}^i u_{\sigma,+}^i h_n(u_{\sigma,+}^i) (\varphi_L^{i-1} - \varphi_K^{i-1}) \\ &\quad - \sum_{i=1}^N \delta t \sum_{\substack{\sigma \in \mathcal{E} \\ \mathbf{v}_{K,\sigma}^i < 0}} |\sigma| \mathbf{v}_{K,\sigma}^i u_{\sigma,+}^i \varphi_K^{i-1} (h_n(u_{\sigma,+}^i) - h_n(u_{\sigma,-}^i)) \\ &= T_{3,1} + T_{3,2} + T_{3,3} \end{aligned}$$

with

$$\begin{aligned} T_{3,1} &= \sum_{i=1}^N \delta t \sum_{\sigma \in \mathcal{E}} |\sigma| \mathbf{v}_{K,\sigma}^i u_{\sigma,+}^i h_n(u_{\sigma,+}^i) (\varphi_K^{i-1} - \varphi_L^{i-1}) \\ T_{3,2} &= \sum_{i=1}^N \delta t \sum_{\substack{\sigma \in \mathcal{E} \\ \mathbf{v}_{K,\sigma}^i \geq 0}} |\sigma| \mathbf{v}_{K,\sigma}^i u_{\sigma,+}^i \varphi_L^{i-1} (h_n(u_{\sigma,+}^i) - h_n(u_{\sigma,-}^i)) \\ T_{3,3} &= - \sum_{i=1}^N \delta t \sum_{\substack{\sigma \in \mathcal{E} \\ \mathbf{v}_{K,\sigma}^i < 0}} |\sigma| \mathbf{v}_{K,\sigma}^i u_{\sigma,+}^i \varphi_K^{i-1} (h_n(u_{\sigma,+}^i) - h_n(u_{\sigma,-}^i)). \end{aligned}$$

Since

$$|T_{3,2} + T_{3,3}| \leq \frac{c \|\varphi\|_\infty}{n} \sum_{i=1}^N \delta t \sum_{\sigma \in \mathcal{E}} |\sigma| |\mathbf{v}_{K,\sigma}^i| |u_{\sigma,+}^i| |T_{2n}(u_{\sigma,+}^i) - T_{2n}(u_{\sigma,-}^i)|,$$

using (2.21) we show that

$$\lim_{n \rightarrow +\infty} \overline{\lim}_{\substack{h_T \rightarrow 0 \\ \delta t \rightarrow 0}} T_{3,2} + T_{3,3} = 0. \quad (2.79)$$

Concerning  $T_{3,1}$  we have  $T_{3,1} = T_{3,1,1} + T_{3,1,2}$  with

$$\begin{aligned} T_{3,1,1} &= \sum_{i=1}^N \delta t \sum_{\sigma \in \mathcal{E}} |\sigma| \mathbf{v}_{K,\sigma}^i u_{\sigma,+}^i h_n(u_{\sigma,+}^i) (\varphi_K^i - \varphi_L^i), \\ T_{3,1,2} &= \sum_{i=1}^N \delta t \sum_{\sigma \in \mathcal{E}} |\sigma| \mathbf{v}_{K,\sigma}^i u_{\sigma,+}^i h_n(u_{\sigma,+}^i) (\varphi_K^{i-1} - \varphi_K^i + \varphi_L^i - \varphi_L^{i-1}). \end{aligned}$$

Still by the regularity of  $\varphi$  and using Hölder's inequality we obtain

$$|T_{3,1,2}| \leq C_\varphi (h_{\mathcal{T}} + \delta t) \|\mathbf{v}\|_{L^2(Q)^d} \|T_{2n}(u_{\mathcal{D}})\|_{L^2(Q)},$$

which implies that

$$\lim_{\substack{h_{\mathcal{T}} \rightarrow 0 \\ \delta t \rightarrow 0}} T_{3,1,2} = 0. \quad (2.80)$$

As far as  $T_{3,1,1}$  is concerned we have

$$\begin{aligned} T_{3,1,1} &= \sum_{i=1}^N \delta t \sum_{\sigma \in \mathcal{E}} |\sigma| \mathbf{v}_{K,\sigma}^i u_{\sigma,+}^i h_n(u_{\sigma,+}^i) (\varphi_K^i - \varphi_L^i) \\ &= \sum_{i=1}^N \sum_{\sigma \in \mathcal{E}} \frac{|\sigma| d_\sigma}{d|D_\sigma|} u_{\sigma,+}^i h_n(u_{\sigma,+}^i) d \frac{\varphi_K^i - \varphi_L^i}{d_\sigma} \int_{t_{i-1}}^{t_i} \int_{D_\sigma} \mathbf{v} \cdot \eta_{K,\sigma} dx dt \\ &= - \sum_{i=1}^N \sum_{\sigma \in \mathcal{E}} \int_{t_{i-1}}^{t_i} \int_{D_\sigma} T_{2n}(u_{\sigma,+}^i) h_n(u_{\sigma,+}^i) \mathbf{v} \cdot \nabla_{\mathcal{D}} \varphi_{\mathcal{D}} dx dt. \end{aligned}$$

Let us define  $\overline{T_{2n}(u)h_n(u)}^{\mathcal{D}^m}$ ,  $\forall t \in [t_{i-1}, t_i)$ ,  $\forall \sigma \in \mathcal{E}$ ,  $\forall x \in D_\sigma$ , by

$$\overline{T_{2n}(u)h_n(u)}^{\mathcal{D}^m}(x, t) = T_{2n}(u_{\sigma,+}^i) h_n(u_{\sigma,+}^i).$$

We write

$$T_{3,1,1} = - \int_0^T \int_{\Omega} \overline{T_{2n}(u)h_n(u)}^{\mathcal{D}^m} \mathbf{v} \cdot \nabla_{\mathcal{D}} \varphi_{\mathcal{D}} dx.$$

By Lemma 2.2.7,  $\nabla_{\mathcal{D}} \varphi_{\mathcal{D}}$  converges weakly to  $\nabla \varphi$  in  $(L^2(Q))^d$ , and then in  $(L^{(d+2)'(Q)})^d$ . Moreover, in view of the definition of  $\overline{T_{2n}(u)h_n(u)}^{\mathcal{D}^m}$ , we can show that it tends to  $T_{2n}(u)h_n(u)$  in  $L^\infty$  weak- $\star$  as  $h_{\mathcal{T}} \rightarrow 0$  and  $\delta t \rightarrow 0$ . Then we obtain

$$\lim_{\substack{h_{\mathcal{T}} \rightarrow 0 \\ \delta t \rightarrow 0}} T_{3,1,1} = - \int_0^T \int_{\Omega} T_{2n}(u) h_n(u) \mathbf{v} \cdot \nabla \varphi dx dt. \quad (2.81)$$

It remains to study the time derivative term. Noting  $\tilde{h}_n(v) = \int_0^v h_n(r) dr$ , and observing that

$$\tilde{h}_n(u_K^i) - \tilde{h}_n(u_K^{i-1}) = (u_K^i - u_K^{i-1}) h_n(u_K^i) - \int_{u_K^{i-1}}^{u_K^i} (r - u_K^{i-1}) h'_n(r) dr,$$

we write  $T_1 = T_{1,1} + T_{1,2}$  with

$$T_{1,1} = \sum_{i=1}^N \delta t \sum_{K \in \mathcal{T}} |K| \frac{\tilde{h}_n(u_K^i) - \tilde{h}_n(u_K^{i-1})}{\delta t} \varphi_K^{i-1},$$

$$T_{1,2} = \sum_{i=1}^N \sum_{K \in \mathcal{T}} |K| \varphi_K^{i-1} \int_{u_K^{i-1}}^{u_K^i} (r - u_K^{i-1}) h'_n(r) dr.$$

Reordering the sums leads to

$$T_{1,1} = - \sum_{i=1}^N \delta t \sum_{K \in \mathcal{T}} |K| \tilde{h}_n(u_K^i) \frac{\varphi_K^i - \varphi_K^{i-1}}{\delta t} - \sum_{K \in \mathcal{T}} |K| \tilde{h}_n(u_K^0) \varphi_K^0$$

$$= - \int_0^T \int_{\Omega} \tilde{h}_n(u_{\mathcal{T}}) \partial_t \varphi \, dx \, dt + R_1 - \int_{\Omega} \tilde{h}_n(u_{\mathcal{T}}^0) \varphi(x, 0) \, dx + R_2,$$

where

$$R_1 = - \sum_{i=1}^N \delta t \sum_{K \in \mathcal{T}} |K| \tilde{h}_n(u_K^i) \left[ \frac{\varphi_K^i - \varphi_K^{i-1}}{\delta t} - \frac{1}{|K| \delta t} \int_{t_{i-1}}^{t_i} \int_K \partial_t \varphi \, dx \, dt \right],$$

$$R_2 = - \sum_{K \in \mathcal{T}} \int_K \tilde{h}_n(u_{\mathcal{T}}^0) [\varphi_K^0 - \varphi(x, 0)] \, dx.$$

By the regularity of  $\varphi$  we have

$$|R_1| \leq C_{\varphi} (h_{\mathcal{T}} + \delta t) \sum_{i=1}^N \delta t \sum_{K \in \mathcal{T}} |K| |\tilde{h}_n(u_K^i)|$$

$$\leq C_{\varphi} (h_{\mathcal{T}} + \delta t) \sum_{i=1}^N \delta t \sum_{K \in \mathcal{T}} |K| |T_{2n}(u_K^i)|$$

$$\leq C_{\varphi} (h_{\mathcal{T}} + \delta t) \|T_{2n}(u_{\mathcal{D}})\|_{L^1(Q)}$$

and  $|R_2| \leq C_{\varphi} (h_{\mathcal{T}}) \|u^0\|_{L^1(\Omega)}$ . Since  $u_{\mathcal{D}}$  converges a.e. in  $Q$  to  $u$  and since  $u_{\mathcal{T}}^0 = (u_K^0)_{K \in \mathcal{T}}$  converges a.e. in  $\Omega$  to  $u^0$  Lebesgue theorem leads to

$$\lim_{\substack{h_{\mathcal{T}} \rightarrow 0 \\ \delta t \rightarrow 0}} T_{1,1} = - \int_0^T \int_{\Omega} \tilde{h}_n(u) \partial_t \varphi \, dx \, dt - \int_{\Omega} \tilde{h}_n(u^0) \varphi(x, 0) \, dx. \quad (2.82)$$

Since

$$|T_{1,2}| \leq \frac{\|\varphi\|_{\infty}}{n} \sum_{i=1}^N \sum_{K \in \mathcal{T}} |K| \int_{u_K^{i-1}}^{u_K^i} (r - u_K^{i-1}) T'_{2n}(r) \, dr,$$

(2.22) implies that

$$\lim_{n \rightarrow +\infty} \overline{\lim}_{\substack{h_{\mathcal{T}} \rightarrow 0 \\ \delta t \rightarrow 0}} T_{1,2} = 0. \quad (2.83)$$

We are now in position to pass to the limit as  $h_{\mathcal{T}} \rightarrow 0$  and  $\delta t \rightarrow 0$  in the scheme. Gathering equations (2.71) to (2.83), we obtain that

$$\begin{aligned} & - \int_0^T \int_{\Omega} \tilde{h}_n(u) \partial_t \varphi \, dx \, dt - \int_{\Omega} \tilde{h}_n(u^0) \varphi(x, 0) \, dx \\ & \quad + \int_0^T \int_{\Omega} \lambda(u) h_n(u) \nabla u \cdot \nabla \varphi \, dx \, dt - \int_0^T \int_{\Omega} u h_n(u) \mathbf{v} \cdot \nabla \varphi \, dx \, dt \\ & \quad - \int_0^T \int_{\Omega} f \varphi h_n(u) \, dx \, dt = \overline{\lim}_{\substack{h_{\mathcal{T}} \rightarrow 0 \\ \delta t \rightarrow 0}} T(n, \varphi) \end{aligned} \quad (2.84)$$

where  $\overline{\lim}_{\substack{h_{\mathcal{T}} \rightarrow 0 \\ \delta t \rightarrow 0}} |T(n, \varphi)| \leq C_{\varphi} \omega(n)$  with  $\omega(n) \rightarrow 0$  as  $n \rightarrow +\infty$ , where  $C_{\varphi}$  depends on  $\|\varphi\|_{L^{\infty}(Q)}$ .

**Step 2** In this step we prove that the renormalized equation (2.6) holds, namely

$$\begin{aligned} & - \int_0^T \int_{\Omega} \xi_t \int_0^u S'(r) \, dr \, dx \, dt - \int_{\Omega} \xi(x, 0) \int_0^{u^0} S'(r) \, dr \, dx \\ & \quad + \int_0^T \int_{\Omega} \lambda(u) \nabla u \cdot \nabla (\xi S'(u)) \, dx \, dt - \int_0^T \int_{\Omega} u \mathbf{v} \cdot \nabla (\xi S'(u)) \, dx \, dt \\ & \quad = \int_0^T \int_{\Omega} f \xi S'(u) \, dx \, dt, \end{aligned}$$

where  $\xi \in C_c^{\infty}(\Omega \times [0, T])$  and  $S \in W^{2, \infty}(\mathbb{R})$  is such that  $S'$  has a compact support and  $S'(u)\xi \in L^2(0, T; H_0^1(\Omega))$ . We pass to the limit in (2.84) as  $n \rightarrow +\infty$  using the techniques developed in [20] which are based on mean time regularizations. In the elliptic case, using a density argument, it is sufficient to take  $\varphi = h(u)\psi$  as a test function and to pass to the limit as  $n \rightarrow +\infty$  (see [53]). Because of the presence of the time derivative in (2.84), we can not use the same technique. This is why we use the tool of Steklov averages and Lemmas 2.1 and 2.3 of [20].

Let  $\xi \in C_c^{\infty}([0, T] \times \Omega)$  and  $S \in W^{2, \infty}(\mathbb{R})$  such that  $S'$  has a compact support and  $S'(u)\xi \in L^2(0, T; H_0^1(\Omega))$ . By density we can take  $\varphi = \xi(S'(u))_h$  as test function in (2.84) where

$$(S'(u))_h(x, t) = \frac{1}{h} \int_t^{t+h} S'(u(x, \tau)) \, d\tau,$$

which gives

$$\begin{aligned} & - \int_0^T \int_{\Omega} \partial_t (\xi(S'(u))_h) \tilde{h}_n(u) \, dx \, dt - \int_{\Omega} \xi(x, 0) (S'(u))_h(x, 0) \tilde{h}_n(u^0) \, dx \\ & \quad + \int_0^T \int_{\Omega} \lambda(u) h_n(u) \nabla u \cdot \nabla (\xi(S'(u))_h) \, dx \, dt - \int_0^T \int_{\Omega} u h_n(u) \mathbf{v} \cdot \nabla (\xi(S'(u))_h) \, dx \, dt \\ & \quad = \int_0^T \int_{\Omega} f h_n(u) \xi(S'(u))_h \, dx \, dt + T(C(\xi, S), n), \end{aligned}$$



which reads as

$$\begin{aligned}
& - \int_0^T \int_{\Omega} \partial_t (\xi(S'(u))_h) (\tilde{h}_n(u) - \tilde{h}_n(u^0)) dx dt \\
& + \int_0^T \int_{\Omega} \lambda(u) h_n(u) \nabla u \cdot \nabla (\xi(S'(u))_h) dx dt - \int_0^T \int_{\Omega} u h_n(u) \mathbf{v} \cdot \nabla (\xi(S'(u))_h) dx dt \\
& = \int_0^T \int_{\Omega} f h_n(u) \xi(S'(u))_h dx dt + T(C(\xi, S), n). \tag{2.85}
\end{aligned}$$

Using Lemma 2.1 of [20], we obtain

$$\begin{aligned}
& - \int_0^T \int_{\Omega} \partial_t (\xi(S'(u))_h) (\tilde{h}_n(u) - \tilde{h}_n(u^0)) dx dt \\
& \geq - \int_0^T \int_{\Omega} \xi_t \left( \tilde{h}_n(u) \frac{1}{h} \int_t^{t+h} S'(u(\tau)) d\tau - \frac{1}{h} \int_t^{t+h} \int_0^{u(\tau)} S''(r) \tilde{h}_n(r) dr d\tau \right) dx dt \\
& \quad - \int_{\Omega} \xi(x, 0) \left( \tilde{h}_n(u^0) S'(u^0) - \int_0^{u^0} S''(r) \tilde{h}_n(r) dr \right) dx,
\end{aligned}$$

and letting  $h$  tend to zero :

$$\begin{aligned}
& \liminf_{h \rightarrow 0} \left\{ - \int_0^T \int_{\Omega} \partial_t (\xi(S'(u))_h) (\tilde{h}_n(u) - \tilde{h}_n(u^0)) dx dt \right\} \\
& \geq - \int_0^T \int_{\Omega} \xi_t \left( \tilde{h}_n(u) S'(u) - \int_0^u S''(r) \tilde{h}_n(r) dr \right) dx dt \\
& \quad - \int_{\Omega} \xi(x, 0) \left( \tilde{h}_n(u^0) S'(u^0) - \int_0^{u^0} S''(r) \tilde{h}_n(r) dr \right) dx.
\end{aligned}$$

Injecting this last inequality in (2.85) and letting  $h$  tend to zero we get

$$\begin{aligned}
& - \int_0^T \int_{\Omega} \xi_t \left( \tilde{h}_n(u) S'(u) - \int_0^u S''(r) \tilde{h}_n(r) dr \right) dx dt \\
& - \int_{\Omega} \xi(x, 0) \left( \tilde{h}_n(u^0) S'(u^0) - \int_0^{u^0} S''(r) \tilde{h}_n(r) dr \right) dx \\
& + \int_0^T \int_{\Omega} \lambda(u) h_n(u) \nabla u \cdot \nabla (\xi S'(u)) dx dt - \int_0^T \int_{\Omega} u h_n(u) \mathbf{v} \cdot \nabla (\xi S'(u)) dx dt \\
& \leq \int_0^T \int_{\Omega} f h_n(u) \xi S'(u) dx dt + T(C(\xi, S), n).
\end{aligned}$$

Passing to the limit as  $n \rightarrow +\infty$  we find

$$\begin{aligned}
& - \int_0^T \int_{\Omega} \xi_t \left( u S'(u) - \int_0^u S''(r) r \right) dx dt \\
& - \int_{\Omega} \xi(x, 0) \left( u^0 S'(u^0) - \int_0^{u^0} S''(r) r dr \right) dx
\end{aligned}$$

$$\begin{aligned}
& + \int_0^T \int_{\Omega} \lambda(u) \nabla u \cdot \nabla (\xi S'(u)) dx dt - \int_0^T \int_{\Omega} u \mathbf{v} \cdot \nabla (\xi S'(u)) dx dt \\
& \leq \int_0^T \int_{\Omega} f \xi S'(u) dx dt.
\end{aligned}$$

Noticing that  $zS'(z) - \int_0^z S''(r)dr = \int_0^z S'(r)dr$  we finally get

$$\begin{aligned}
& - \int_0^T \int_{\Omega} \xi_t \int_0^u S'(r)dr dx dt - \int_{\Omega} \xi(x, 0) \int_0^{u^0} S'(r)dr dx \\
& + \int_0^T \int_{\Omega} \lambda(u) \nabla u \cdot \nabla (\xi S'(u)) dx dt - \int_0^T \int_{\Omega} u \mathbf{v} \cdot \nabla (\xi S'(u)) dx dt \\
& \leq \int_0^T \int_{\Omega} f \xi S'(u) dx dt. \tag{2.86}
\end{aligned}$$

We prove now the reverse inequality taking  $\varphi = \xi(S'(u))_{-h}$  where

$$(S'(u))_{-h}(x, t) = \frac{1}{h} \int_{t-h}^t S'(\hat{u}(x, \tau)) d\tau,$$

with

$$\hat{u}(x, t) = \begin{cases} u(x, t) & \text{if } t \geq 0, \\ U_j & \text{if } t < 0, \end{cases}$$

where  $U_j$  is a sequence in  $\mathcal{C}_c^\infty(\Omega)$  such that  $U_j \rightarrow u^0$  strongly in  $L^1(\Omega)$ . We thus get

$$\begin{aligned}
& - \int_{\varepsilon}^T \int_{\Omega} \partial_t (\xi(S'(u))_{-h}) (\tilde{h}_n(u) - \tilde{h}_n(u^0)) dx dt \\
& + \int_{\varepsilon}^T \int_{\Omega} \lambda(u) h_n(u) \nabla u \cdot \nabla (\xi(S'(u))_{-h}) dx dt - \int_{\varepsilon}^T \int_{\Omega} u h_n(u) \mathbf{v} \cdot \nabla (\xi(S'(u))_{-h}) dx dt \\
& = \int_{\varepsilon}^T \int_{\Omega} f h_n(u) \xi(S'(u))_{-h} dx dt + T(C(\xi, S), n). \tag{2.87}
\end{aligned}$$

Using Lemma 2.3 of [20], we obtain

$$\begin{aligned}
& - \int_0^T \int_{\Omega} \partial_t (\xi(S'(u))_{-h}) (\tilde{h}_n(u) - \tilde{h}_n(u^0)) dx dt \\
& \leq - \int_0^T \int_{\Omega} \xi_t \left( \tilde{h}_n(u) \frac{1}{h} \int_{t-h}^t S'(\hat{u}(\tau)) d\tau - \frac{1}{h} \int_{t-h}^t \int_0^{\hat{u}(\tau)} S''(r) \tilde{h}_n(r) dr d\tau \right) dx dt \\
& \quad - \int_{\Omega} \xi(x, 0) \left( \tilde{h}_n(u^0) S'(U_j) - \int_0^{U_j} S''(r) \tilde{h}_n(r) dr \right) dx.
\end{aligned}$$

We now let  $h$  tend to zero :

$$\begin{aligned}
& \limsup_{h \rightarrow 0} \left\{ - \int_0^T \int_{\Omega} \partial_t (\xi(S'(u))_{-h}) (\tilde{h}_n(u) - \tilde{h}_n(u^0)) dx dt \right\} \\
& \leq - \int_0^T \int_{\Omega} \xi_t \left( \tilde{h}_n(u) S'(u) - \int_0^u S''(r) \tilde{h}_n(r) dr \right) dx dt \\
& \quad - \int_{\Omega} \xi(x, 0) \left( \tilde{h}_n(u^0) S'(U_j) - \int_0^{U_j} S''(r) \tilde{h}_n(r) dr \right) dx,
\end{aligned}$$

and letting  $j$  tends to infinity :

$$\begin{aligned} & \limsup_{h \rightarrow 0} \left\{ - \int_0^T \int_{\Omega} \partial_t (\xi(S'(u))_{-h}) (\tilde{h}_n(u) - \tilde{h}_n(u^0)) dx dt \right\} \\ & \leq - \int_0^T \int_{\Omega} \xi_t \left( \tilde{h}_n(u) S'(u) - \int_0^u S''(r) \tilde{h}_n(r) dr \right) dx dt \\ & \quad - \int_{\Omega} \xi(x, 0) \left( \tilde{h}_n(u^0) S'(u^0) - \int_0^{u^0} S''(r) \tilde{h}_n(r) dr \right) dx, \end{aligned}$$

Injecting this last inequality in (2.87) and letting  $h$  tends to zero we get

$$\begin{aligned} & \int_0^T \int_{\Omega} f h_n(u) \xi S'(u) dx dt + T(C(\xi, S), n) \\ & \leq - \int_0^T \int_{\Omega} \xi_t \left( \tilde{h}_n(u) S'(u) - \int_0^u S''(r) \tilde{h}_n(r) dr \right) dx dt \\ & \quad - \int_{\Omega} \xi(x, 0) \left( \tilde{h}_n(u^0) S'(u^0) - \int_0^{u^0} S''(r) \tilde{h}_n(r) dr \right) dx \\ & \quad + \int_0^T \int_{\Omega} \lambda(u) h_n(u) \nabla u \cdot \nabla (\xi S'(u)) dx dt - \int_0^T \int_{\Omega} u h_n(u) \mathbf{v} \cdot \nabla (\xi S'(u)) dx dt, \end{aligned}$$

and letting  $n \rightarrow +\infty$

$$\begin{aligned} & \int_0^T \int_{\Omega} f \xi S'(u) dx dt \\ & \leq - \int_0^T \int_{\Omega} \xi_t \left( u S'(u) - \int_0^u S''(r) r \right) dx dt \\ & \quad - \int_{\Omega} \xi(x, 0) \left( u^0 S'(u^0) - \int_0^{u^0} S''(r) r \right) dx \\ & \quad + \int_0^T \int_{\Omega} \lambda(u) \nabla u \cdot \nabla (\xi S'(u)) dx dt - \int_0^T \int_{\Omega} u \mathbf{v} \cdot \nabla (\xi S'(u)) dx dt. \end{aligned}$$

We finally obtain the inequality

$$\begin{aligned} & \int_0^T \int_{\Omega} f \xi S'(u) dx dt \leq \\ & \quad - \int_0^T \int_{\Omega} \xi_t \int_0^u S''(r) dr dx dt - \int_{\Omega} \xi(x, 0) \int_0^{u^0} S''(r) dr dx \\ & \quad + \int_0^T \int_{\Omega} \lambda(u) \nabla u \cdot \nabla (\xi S'(u)) dx dt - \int_0^T \int_{\Omega} u \mathbf{v} \cdot \nabla (\xi S'(u)) dx dt. \end{aligned} \quad (2.88)$$

Gathering (2.86) and (2.88) we get the desired result.

**Step 3** We now prove that  $u$  satisfies the decay (2.5) of the truncate energy. Thanks to the discrete estimate on the energy (2.20) we have

$$\lim_{n \rightarrow \infty} \overline{\lim}_{\substack{h_T \rightarrow 0 \\ \delta t \rightarrow 0}} \frac{1}{n} \sum_{i=1}^N \delta t \sum_{\sigma \in \mathcal{E}} \frac{|\sigma|}{d_{\sigma}} \lambda(u)_{\sigma}^i (T_{2n}(u_K^i) - T_{2n}(u_L^i))^2 = 0,$$

and

$$\begin{aligned}
& \sum_{i=1}^N \delta t \sum_{\sigma \in \mathcal{E}} \frac{|\sigma|}{d_\sigma} \lambda(u)_\sigma^i (T_{2n}(u_K^i) - T_{2n}(u_L^i))^2 \\
&= \sum_{i=1}^N \delta t \sum_{\sigma \in \mathcal{E}} |\sigma| d_\sigma \lambda(u)_\sigma^i \left( \frac{T_{2n}(u_K^i) - T_{2n}(u_L^i)}{d_\sigma} \right)^2 \\
&= \sum_{i=1}^N \delta t \sum_{\sigma \in \mathcal{E}} d |D_\sigma| \lambda(u)_\sigma^i \left( \frac{T_{2n}(u_K^i) - T_{2n}(u_L^i)}{d_\sigma} \right)^2 \\
&= \sum_{i=1}^N \delta t \frac{1}{d} \sum_{\sigma \in \mathcal{E}} |D_\sigma| \lambda(u)_\sigma^i \left( d \frac{T_{2n}(u_K^i) - T_{2n}(u_L^i)}{d_\sigma} \right)^2 \\
&\geq \frac{1}{d} \sum_{i=1}^N \delta t \sum_{\sigma \in \mathcal{E}} |D_{K,\sigma}| h_n(u_K^i) \frac{\lambda(u_K^i)}{2} \left( d \frac{T_{2n}(u_K^i) - T_{2n}(u_L^i)}{d_\sigma} \right)^2 \\
&\quad + \frac{1}{d} \sum_{i=1}^N \delta t \sum_{\sigma \in \mathcal{E}} |D_{L,\sigma}| h_n(u_L^i) \frac{\lambda(u_L^i)}{2} \left( d \frac{T_{2n}(u_K^i) - T_{2n}(u_L^i)}{d_\sigma} \right)^2 \\
&= \frac{1}{d} \int_0^T \int_\Omega h_n(u_{\mathcal{D}}) \frac{\lambda(T_{2n}(u_{\mathcal{D}}))}{2} |\nabla_{\mathcal{D}} T_{2n}(u_{\mathcal{D}})|^2 dx dt,
\end{aligned}$$

thus,  $\lim_{n \rightarrow \infty} \overline{\lim}_{\substack{h_T \rightarrow 0 \\ \delta t \rightarrow 0}} \frac{1}{n} \int_0^T \int_\Omega h_n(u_{\mathcal{D}}) \frac{\lambda(u_{\mathcal{D}})}{2} |\nabla_{\mathcal{D}} T_{2n}(u_{\mathcal{D}})|^2 dx dt = 0$ . Since  $\nabla T_{2n}(u_{\mathcal{D}})$  converges weakly in  $L^2(Q)^d$ ,  $\lambda(T_{2n}(u_{\mathcal{D}}))$  converges to  $\lambda(u)$   $L^\infty$  weak- $\star$  and  $h_n(u_{\mathcal{D}})$  converge to  $h_n(u)$  almost everywhere in  $Q$  we also have

$$\frac{1}{n} \int_0^T \int_\Omega h_n(u) \lambda(u) |\nabla T_{2n}(u)|^2 dx \leq \liminf_{\substack{h_T \rightarrow 0 \\ \delta t \rightarrow 0}} \frac{1}{n} \int_0^T \int_\Omega h_n(u_{\mathcal{D}}) \lambda(u_{\mathcal{D}}) |\nabla_{\mathcal{D}} T_{2n}(u_{\mathcal{D}})|^2 dx dt,$$

which, since  $h_n(\cdot) = 1$  on  $[-n, n]$ , leads to

$$\lim_{n \rightarrow \infty} \frac{1}{n} \int_0^T \int_\Omega \lambda(u) |\nabla T_n(u)|^2 dx dt = 0.$$

**Step 4** We prove here the uniqueness of the renormalized solution. In view of the convergence analysis, if the renormalized solution of (2.1)  $u$  is unique usual arguments imply that the whole sequence  $u_{\mathcal{D}_m}$  converges to  $u$ . To our knowledge due to the nonlinearity of the operator and the presence of the term  $-\operatorname{div}(\mathbf{v}(x, t)u)$  the uniqueness is proved under additional assumptions on  $\lambda$  or  $\mathbf{v}$ .

If  $\lambda$  is local Lipschitz continuous, namely

$$\forall n \in \mathbb{N} \exists C_n > 0 \text{ such that } \forall |r|, |t| \leq n, |\lambda(r) - \lambda(t)| \leq C_n |r - t|,$$

then we are in the scope of [35] (see also [18]) : the renormalized solution is unique.

If  $\lambda$  is a continuous function the price to pay to have uniqueness is to assume more regularity on  $\mathbf{v}$ . If  $\mathbf{v}$  depends only on  $x$  then the time doubling variable method, see e.g. [61] (which is not the method developed in [18]), allows one to prove the

uniqueness of the renormalized solution. Indeed the change of variable  $w = \tilde{\lambda}(u)$  where  $\tilde{\lambda}(r) = \int_0^r \lambda(s) ds$  implies that  $w$  is a renormalized solution of

$$\begin{aligned} \frac{\partial \tilde{\lambda}^{-1}(w)}{\partial t} - \Delta w + \operatorname{div}(\mathbf{v} \tilde{\lambda}^{-1}(w)) &= f \text{ in } Q, \\ w &= 0 \text{ on } \partial\Omega \times (0, T), \\ w(x, 0) &= \tilde{\lambda}(u^0(x)) \quad \forall x \in \Omega. \end{aligned}$$

Since the function  $\tilde{\lambda}(r)$  is local Lipschitz continuous and since  $\mathbf{v}$  does not depend on  $t$  the results of [20] (see Section 5 comments and remarks) give the uniqueness of the renormalized solution  $w$  which in turn gives the uniqueness of  $u$ .  $\square$

# Chapitre 3

## Finite volume scheme and renormalized solutions for a doubly nonlinear parabolic equation with $L^1$ data

### 3.1 Introduction

In this chapter we consider a finite volume discretization of the following parabolic problem :

$$\begin{aligned} \frac{\partial b(u)}{\partial t} - \operatorname{div}(\lambda(u)\nabla u) &= f \text{ in } Q, \\ u &= 0 \text{ on } \partial\Omega \times (0, T), \\ b(u(x, 0)) &= b(u^0(x)) \quad \forall x \in \Omega, \end{aligned} \tag{3.1}$$

where  $\Omega$  is an open bounded polygonal subset of  $\mathbb{R}^d$ ,  $d \geq 2$ ,  $T$  is a positive number and  $Q = \Omega \times (0, T)$ ,  $\lambda$  is a continuous function such that  $\lambda(u) \geq \mu > 0$ , with  $\mu$  a positive real number,  $f \in L^1(Q)$  and  $b(u^0) \in L^1(\Omega)$ . Moreover  $b$  is assumed to be a strictly increasing  $\mathcal{C}^1$ -function such that  $b(0) = 0$  and  $b'(r) > 0 \quad \forall r \in \mathbb{R}$ .

In [23] the authors proved the existence of a distributional solution of (3.1) in the case where  $b(u) = u$ ,  $\lambda$  is bounded and the right-hand side is a bounded measure. But since the solution in the sense of distribution is not unique in general (see [68] in the elliptic case) we have to choose an appropriate framework. Moreover even in the case  $f \in L^2(Q)$ , without any growth assumption on  $\lambda$ , if  $u \in L^2(0, T; H_0^1(\Omega))$  we cannot expect to have  $\lambda(u)\nabla u \in L^2(Q)$  nor  $\lambda(u)\nabla u \in L_{loc}^1(Q)$  and then we cannot give in general an existence result for weak solution or solution in the sense of distribution of equation (3.1). Here we use the framework of renormalized solutions which provides in general existence, stability and uniqueness for a large class of parabolic equations with  $L^1$  data. Recall that the notion of renormalized solution have been introduced by DiPerna and Lions in [36] for the study of Boltzmann equation. This notion has been adapted to elliptic equations ([58]) and to parabolic equations ([18, 21]) with  $L^1$  or measure data.

When  $b(r) = r$  the existence of a renormalized solution of (3.1) has been proved

in [16] where the authors considered a more general Leray-Lions operator  $-\operatorname{div}(a(x, t, u, \nabla u))$ , which is coercive and which grows like  $|\nabla u|^{p-1}$  with respect to  $\nabla u$ , and a data belonging to  $L^1(Q) + L^{p'}(0, T; W^{-1, p'}(\Omega))$ . Under additional assumptions on  $a$  uniqueness results are given in [16]. In [21] existence and uniqueness of a renormalized solution have been established in the case where  $b(r)$  is strictly increasing and can possibly blow up for some finite  $r_0$  and  $a(x, t, s, \xi)$  is independent of  $s$  and linear with respect to  $\xi$ . When  $b$  is a maximal monotone graph the existence of a renormalized solution and the uniqueness (when  $a$  does not depend on  $t$ ) are proved in [20]. It is worth noting that the proof of the existence in [20] relies on a different strategy to the one used in [21]. Indeed in [21] the authors use Landes regularization and technical lemmas while the technique of [20], using Steklov averages and convex inequalities, is simpler and shorter.

As far as finite volume schemes for parabolic equations are concerned the linear and variational case is dealt in [44]. As in the continuous case one of the main obstacle for time dependent equations is to derive some compactness (and then the pointwise convergence) for the approximate solution which is necessary (but not sufficient) to pass to the limit. One of the methods is to derive with the equation an estimate on differences of time and space translates (see [44] and [4, 8] in the context of degenerate parabolic equations). Another method is to use the compactness result proved in [48] which is a discrete version of the Aubin-Lions Lemma (see [69]).

In the context parabolic equations with a time derivative term  $\frac{\partial b(u)}{\partial t}$ , finite volume schemes have been studied in [2, 6–8] where the data are more regular.

In the present paper we mix the tools developed for finite volume schemes ([44, 47, 48]) with the strategy to deal with the existence of a renormalized solution for parabolic equations with  $L^1$  data ([16, 20, 21]). Let us mention that such an analysis is performed for a convection-diffusion equation in [53] in the elliptic case and in [52] in the parabolic case. It is worth noting that the presence of the time derivative term  $\frac{\partial b(u)}{\partial t}$  gives additional difficulties due to the lack of regularity and the nonlinearity of this term.

We prove here that the approximate solution obtained by a finite volume scheme converges to the unique renormalized solution of (3.1). Recall that (see [16]) a renormalized solution of (3.1) is a measurable function  $u$  defined from  $Q$  to  $\mathbb{R}$  such that

$$b(u) \in L^\infty(0, T; L^1(\Omega)), \quad (3.2)$$

$$\forall k > 0, T_k(u) \in L^2(0, T; H_0^1(\Omega)), \quad (3.3)$$

$$\lim_{k \rightarrow +\infty} \frac{1}{k} \int_0^T \int_\Omega \lambda(u) |\nabla T_k(u)|^2 dx dt = 0, \quad (3.4)$$

and for every functions  $\xi \in C_c^\infty([0, T[ \times \Omega)$  and  $S \in W^{2, \infty}(\mathbb{R})$  such that  $S'$  has a compact support and  $S'(u)\xi \in L^2(0, T; H_0^1(\Omega))$ ,

$$\begin{aligned} & - \int_0^T \int_\Omega \xi_t \int_0^u b'(r) S'(r) dr dx dt - \int_\Omega \xi(x, 0) \int_0^{u^0} b'(r) S'(r) dr dx \\ & + \int_0^T \int_\Omega \lambda(u) \nabla u \cdot \nabla (\xi S'(u)) dx dt = \int_0^T \int_\Omega f \xi S'(u) dx dt. \end{aligned} \quad (3.5)$$

where  $T_k$  is the truncate function at height  $k$  :

$$T_k(s) = \begin{cases} s & \text{if } |s| \leq k, \\ \text{sgn}(s)k & \text{if } |s| > k. \end{cases}$$

The chapter is organized as follows. In Section 2, we present the finite volume scheme, the time discretization and the properties of the discrete gradient. Section 3 is devoted to prove several estimates, especially an estimate on the time derivative and the discrete equivalent to (3.4) which is crucial to pass to the limit in the scheme. In Section 4, we prove the convergence of the finite volume scheme, passing to the limit and using both a density argument and Steklov averages.

## 3.2 The finite volume scheme and the time discretization

Let us define the admissibility of the mesh in the present work.

**Definition 3.2.1 (Admissible mesh)** *Let  $\Omega$  be an open bounded polygonal subset of  $\mathbb{R}^d$ . An admissible finite volume mesh of  $\Omega$  is given by a finite partition  $\mathcal{T}$  of  $\Omega$  in polygonal convex sets, called the "control volumes", by a finite family  $\mathcal{E}$  of disjoint subsets of  $\bar{\Omega}$  contained in affine hyperplanes, called the "edges", and by a family  $\mathcal{P} = (x_K)_{K \in \mathcal{T}}$  of points in  $\Omega$  such that*

- *each  $\sigma \in \mathcal{E}$  is a non-empty open subset of  $\partial K$  for some  $K \in \mathcal{T}$ ,*
- *by denoting  $\mathcal{E}(K) = \{\sigma \in \mathcal{E}; \sigma \in \partial K\}$ , one has  $\partial K = \cup_{\sigma \in \mathcal{E}(K)} \bar{\sigma}$  for all  $K \in \mathcal{T}$ ,*
- *for all  $K \neq L$  in  $\mathcal{T}$ , either the  $(d-1)$ -dimensional measure of  $\bar{K} \cap \bar{L}$  is null, or  $\bar{K} \cap \bar{L} = \bar{\sigma}$  for some  $\sigma \in \mathcal{E}$ , that we denote then  $\sigma = K|L$ ,*
- *for all  $K \in \mathcal{T}$ ,  $x_K$  is in the interior of  $K$ ,*
- *for all  $\sigma = K|L \in \mathcal{E}$ , the line  $(x_K, x_L)$  intersects and is orthogonal to  $\sigma$ ,*
- *for all  $\sigma \in \mathcal{E}$ ,  $\sigma \subset \partial\Omega \cap \partial K$ , the line which is orthogonal to  $\sigma$  and going through  $x_K$  intersects  $\sigma$ .*

We denote by  $|K|$  (resp.  $|\sigma|$ ) the Lebesgue measure of  $K \in \mathcal{T}$  (resp.  $\sigma \in \mathcal{E}$ ). The unit normal to  $\sigma \in \mathcal{E}(K)$  outward to  $K$  is denoted by  $\eta_{K,\sigma}$ .  $\mathcal{E}_{int}$  (resp.  $\mathcal{E}_{ext}$ ) is defined as the set of interior (resp. boundary) edges.

For any  $K \in \mathcal{T}$  and  $\sigma \in \mathcal{E}(K)$  we denote by  $d_{K,\sigma}$  the Euclidean distance between  $x_K$  and  $\sigma$ . For any  $\sigma \in \mathcal{E}$ , we define  $d_\sigma = d_{K,\sigma} + d_{L,\sigma}$  if  $\sigma = K|L \in \mathcal{E}_{int}$  (in which case  $d_\sigma$  is the Euclidean distance between  $x_K$  and  $x_L$ ) and  $d_\sigma = d_{K,\sigma}$  if  $\sigma \in \mathcal{E}_{ext} \cap \mathcal{E}(K)$ . The size of the mesh, denoted by  $h_{\mathcal{T}}$ , is defined by  $h_{\mathcal{T}} = \sup_{K \in \mathcal{T}} \text{diam}(K)$ .

In the continuous case the usual tools to solve the problem are Poincaré and Sobolev inequalities. In the discrete case we will need such estimates, so we have to establish their discrete versions (see [32]). Let us define the discrete  $W_0^{1,q}$  norm.

**Definition 3.2.2 (Discrete  $W_0^{1,q}$  norm)** *Let  $\Omega$  be an open bounded polygonal subset of  $\mathbb{R}^d$ ,  $d \geq 2$ , and let  $\mathcal{T}$  be an admissible mesh. Define  $X(\mathcal{T})$  as the set of*



functions from  $\Omega$  to  $\mathbb{R}$  which are constant over each control volume of the mesh. For  $v_{\mathcal{T}} \in X(\mathcal{T})$  and  $q \in [1, +\infty[$ , we define the discrete  $W_0^{1,q}$  norm by

$$\|v_{\mathcal{T}}\|_{1,q,\mathcal{T}}^q = \sum_{\substack{\sigma \in \mathcal{E}_{int} \\ \sigma = K|L}} |\sigma| d_{\sigma} \left| \frac{v_K - v_L}{d_{\sigma}} \right|^q + \sum_{\substack{\sigma \in \mathcal{E}_{ext} \\ \sigma \in \mathcal{E}(K)}} |\sigma| d_{\sigma} \left| \frac{v_K}{d_{\sigma}} \right|^q$$

where  $v_K$  denotes the value taken by  $v$  on the control volume  $K$ .

Let us also define

$$\begin{aligned} \|v_{\mathcal{T}}\|_{1,\infty,\mathcal{T}} = \max \{ & \left\{ \frac{|v_K - v_L|}{d_{\sigma}}, \sigma \in \mathcal{E}_{int}, \sigma = K|L \right\} \\ & \cup \left\{ \frac{|v_K|}{d_{\sigma}}, \sigma \in \mathcal{E}_{ext}, \sigma \in \mathcal{E}(K) \right\} \cup \{|v_K|, K \in \mathcal{T}\} \}. \end{aligned}$$

**Definition 3.2.3 (Discrete dual norm)** Let  $q > 1$  and let  $q'$  be given by  $\frac{1}{q} + \frac{1}{q'} = 1$  if  $q$  is finite and  $q' = 1$  if  $q = +\infty$ . We define the dual norme by

$$\|v\|_{-1,q',\mathcal{T}} = \sup_{\substack{w \in X(\mathcal{T}) \\ w \neq 0}} \frac{1}{\|w\|_{1,q,\mathcal{T}}} \int_{\Omega} v w \, dx.$$

**Proposition 3.2.4 (Discrete Poincaré inequality)** Let  $\mathcal{T}$  be an admissible mesh and  $v_{\mathcal{T}} \in X(\mathcal{T})$ . Then, if  $1 \leq q \leq 2$ ,

$$\|v_{\mathcal{T}}\|_{L^q(\Omega)} \leq \text{diam}(\Omega) \|v_{\mathcal{T}}\|_{1,q,\mathcal{T}}.$$

**Proposition 3.2.5 (Discrete Sobolev inequality)** Let  $1 \leq q \leq 2$ ,  $\mathcal{T}$  be an admissible mesh and  $\xi > 0$  satisfying

$$\text{for all } K \in \mathcal{T} \text{ and all } \sigma \in \mathcal{E}(K), \, d_{K,\sigma} \geq \xi d_{\sigma}.$$

Then, with  $q^* = \frac{dq}{d-q}$  if  $q < d$  and  $q^* < \infty$  if  $q = d = 2$ , there exists  $C > 0$  only depending on  $(\Omega, q, q^*, \xi)$  such that, for all  $v_{\mathcal{T}} \in X(\mathcal{T})$ ,

$$\|v_{\mathcal{T}}\|_{L^{q^*}(\Omega)} \leq C \|v_{\mathcal{T}}\|_{1,q,\mathcal{T}}.$$

As far as the time discretization is concerned, we suppose given a uniform partition of the interval  $[0, T]$  such that  $[0, T] = \cup_{1 \leq i \leq N} [t^{i-1}, t^i]$  where  $t^i = i\delta t$  and  $\delta t = \frac{T}{N}$ . We define  $X_{\mathcal{D}}$  the space of piecewise constant functions over each  $K \times (t^{i-1}, t^i)$  for  $K \in \mathcal{T}$  and  $1 \leq i \leq N$ . To each sequence  $(v_{\mathcal{T}}^i)_{1 \leq i \leq N}$  of functions of  $X(\mathcal{T})$ , we associate the function  $v_{\mathcal{D}} \in X_{\mathcal{D}}$  defined by :

$$v_{\mathcal{D}}(x, t) = v_{\mathcal{T}}^i(x) \, \forall x \in \Omega, \, \forall t \in (t^{i-1}, t^i), \, 1 \leq i \leq N,$$

$$\text{i.e. } v_{\mathcal{D}}(x, t) = v_K^i \, \forall x \in K, \, K \in \mathcal{T}, \, \forall t \in (t^{i-1}, t^i), \, 1 \leq i \leq N.$$

We now define the discrete time derivative of  $v_{\mathcal{D}}$  by

$$\partial_{i,\mathcal{D}} v^i(x) = \frac{v^i(x) - v^{i-1}(x)}{\delta t} \, \forall 1 \leq i \leq N,$$

$$\text{i.e. } \partial_t v_K^i = \frac{v_K^i - v_K^{i-1}}{\delta t} \, \forall K \in \mathcal{T}, \, \forall 1 \leq i \leq N.$$

Before writing the finite volume scheme, let us define a discrete finite volume gradient (see [51]).

**Definition 3.2.6 (Discrete finite volume gradient)** For  $K \in \mathcal{T}$  and  $\sigma \in \mathcal{E}(K)$ , we define the volume  $D_{K,\sigma}$  as the cone of basis  $\sigma$  and of opposite vertex  $x_K$ . Then, we define the "diamond-cell"  $D_\sigma$  (see Figure 3.1) by

$$\begin{aligned} D_\sigma &= D_{K,\sigma} \cup D_{L,\sigma} && \text{if } \sigma = K|L \in \mathcal{E}_{int}, \\ D_\sigma &= D_{K,\sigma} && \text{if } \sigma \in \mathcal{E}_{ext} \cap \mathcal{E}(K). \end{aligned}$$

For any  $v_{\mathcal{D}} \in X_{\mathcal{D}}$ , we define the discrete gradient  $\nabla_{\mathcal{D}} v_{\mathcal{D}}$  as the piecewise constant function over each diamond cell and given by

$$\begin{aligned} \forall \sigma \in \mathcal{E}_{int}, \sigma = K|L, \quad \nabla_{\mathcal{D}} v_{\mathcal{D}}(x, t) &= d \frac{v_L^i - v_K^i}{d_\sigma} \eta_{K,\sigma}, \quad \forall x \in D_\sigma, \quad \forall t \in (t^{i-1}, t^i], \quad 1 \leq i \leq N \\ \forall \sigma \in \mathcal{E}_{ext} \cap \mathcal{E}(K), \quad \nabla_{\mathcal{T}} v_{\mathcal{D}}(x, t) &= d \frac{0 - v_K^i}{d_\sigma} \eta_{K,\sigma}, \quad \forall x \in D_\sigma, \quad \forall t \in (t^{i-1}, t^i], \quad 1 \leq i \leq N. \end{aligned}$$

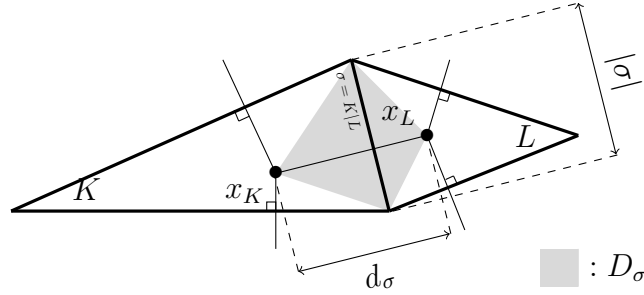


FIGURE 3.1 – The diamond  $D_\sigma$

**Lemma 3.2.7 (Weak convergence of the finite volume gradient)** Let  $(\mathcal{T}_m)_{m \geq 1}$  be a sequence of admissible meshes such that there exists  $\xi > 0$  satisfying

$$\text{for all } m \geq 1 \text{ for all } K \in \mathcal{T} \text{ and all } \sigma \in \mathcal{E}(K), \quad d_{K,\sigma} \geq \xi d_\sigma,$$

and such that  $h_{\mathcal{T}_m} \rightarrow 0$  and let  $\delta t_m$  be a sequence of time steps such that  $\delta t_m \rightarrow 0$ . Let  $v_{\mathcal{D}_m} \in X_{\mathcal{D}_m}$  and let us assume that there exists  $C > 0$  such that  $\sum_{i=1}^N \delta t_m \|v_{\mathcal{T}_m}^i\|_{1,2,\mathcal{T}_m}^2 \leq C$ , and that  $v_{\mathcal{D}_m}$  converges in  $L^1(Q)$  to  $v \in L^2(0, T; H_0^1(\Omega))$ . Then  $\nabla_{\mathcal{D}_m} v_{\mathcal{D}_m}$  converges to  $\nabla v$  weakly in  $L^2(Q)^d$ .

**Proof** The proof is done in the elliptic case in [51] and adapted to the evolution case in Lemma 5.2 of [5, 52].  $\square$

Let  $\mathcal{T}$  be an admissible mesh, we can define the finite volume discretization of (3.1). For  $1 \leq i \leq N$ ,  $K \in \mathcal{T}$  and  $\sigma \in \mathcal{E}(K)$ , we define

$$f_K^i = \frac{1}{\delta t |K|} \int_{t_{i-1}}^{t_i} \int_K f \, dx \, dt, \quad (3.6)$$

$$b(u_K^0) = \frac{1}{|K|} \int_K b(u^0(x)) \, dx, \quad (3.7)$$

and we can write the scheme as the following set of equations

For all  $1 \leq i \leq N$  and for all  $K \in \mathcal{T}$

$$|K| \frac{b(u_K^i) - b(u_K^{i-1})}{\delta t} + \sum_{\sigma \in \mathcal{E}(K)} \frac{|\sigma|}{d_\sigma} \lambda(u)_\sigma^i (u_K^i - u_L^i) = |K| f_K^i, \quad (3.8)$$

with  $u_L^i = 0$  if  $\sigma \in \mathcal{E}_{ext} \cap \mathcal{E}(K)$  and for all  $1 \leq i \leq N$  we define  $\lambda(u)_\sigma^i$  in the following way

$$\forall \sigma = K|L \in \mathcal{E}_{int}, \min[\lambda(u_K^i), \lambda(u_L^i)] \leq \lambda(u)_\sigma^i \leq \max[\lambda(u_K^i), \lambda(u_L^i)], \quad (3.9)$$

$$\forall \sigma \in \mathcal{E}_{ext} \cap \mathcal{E}(K), \lambda(u)_\sigma^i = \lambda(u_K^i), \quad (3.10)$$

where  $\lambda(u)_\sigma^i$  is for example the mean value of  $\lambda(u_K^i)$  and  $\lambda(u_L^i)$ .

### 3.3 Estimates

This section is devoted to derive estimates on the sequence  $u_{\mathcal{D}}$  (see Propositions 3.3.1, 3.3.3 and 3.3.4) and then to show the pointwise convergence of  $u_{\mathcal{D}}$ , up to a subsequence (see Proposition 3.3.6). At last we prove in Proposition 3.3.7 a uniform estimate on the truncated discrete energy which is crucial to pass to the limit in the convergence analysis. The main tool in Proposition 3.3.6 is the discrete version proved in [48] of Aubin-Simon type Lemma (see [69]), which requires estimates on the functions and on the discrete time derivative of the functions in appropriate discrete spaces. In view of the  $L^1$  character of (3.1) we cannot expect to derive estimates on  $u_{\mathcal{D}}$  and  $\partial_{t,\mathcal{D}} u_{\mathcal{D}}$  nor on  $b(u_{\mathcal{D}})$  and  $\partial_{t,\mathcal{D}} b(u_{\mathcal{D}})$ . Here we adapt the strategy developed in [21] for the continuous case, namely the equation

$$\frac{\partial b(u_\varepsilon)}{\partial t} - \operatorname{div}(A \nabla u_\varepsilon) + \operatorname{div} \phi(u_\varepsilon) = f_\varepsilon, \quad (3.11)$$

where  $f_\varepsilon \in L^2(Q)$  and  $f_\varepsilon \rightarrow f \in L^1(Q)$ , which consists in proving estimates on  $\tilde{h}_n(b(u_\varepsilon))$  and on  $\frac{\partial \tilde{h}_n(b(u_\varepsilon))}{\partial t}$ , where  $\tilde{h}_n$  is a regularized truncation function, in applying Aubin-Simon Lemma (see [69]) and in using a diagonal process which gives the pointwise convergence of  $b(u_\varepsilon)$  which in turn implies the pointwise convergence of  $u_\varepsilon$ . In the continuous case the chain rule formula

$$h_n(b(u_\varepsilon)) \frac{\partial b(u_\varepsilon)}{\partial t} = \frac{\partial \tilde{h}_n(b(u_\varepsilon))}{\partial t} \quad (3.12)$$

and multiplying formally equation (3.11) by  $h_n(b(u_\varepsilon))$  allows one to obtain estimate on  $\frac{\partial \tilde{h}_n(b(u_\varepsilon))}{\partial t}$  (see [16, 21]). However in the discrete case instead of chain rule formula we have the equality

$$\partial_{t,\mathcal{D}} \tilde{h}_n(b(u_{\mathcal{D}}^i)) = h_n(b(u_{\mathcal{D}}^i)) \partial_{t,\mathcal{D}} b(u_{\mathcal{D}}^i) - \frac{1}{\delta t} \int_{b(u_{\mathcal{D}}^{i-1})}^{b(u_{\mathcal{D}}^i)} (r - b(u_{\mathcal{D}}^{i-1})) h'_n(r) dr. \quad (3.13)$$

As a consequence the control of  $\partial_{t,\mathcal{D}} \tilde{h}_n(b(u_{\mathcal{D}}^i))$  requires estimates on  $\partial_{t,\mathcal{D}} b(u_{\mathcal{D}}^i)$  (which is performed roughly speaking through pointwise multiplication of the scheme by  $h_n(b(u_{\mathcal{D}}))$ ) and on the additional term  $\int_{b(u_{\mathcal{D}}^{i-1})}^{b(u_{\mathcal{D}}^i)} (r - b(u_{\mathcal{D}}^{i-1})) h'_n(r) dr$  (see Lemma 3.3.5).

**Proposition 3.3.1 (Estimate on  $T_n(u_{\mathcal{D}})$ )** Let  $\mathcal{T}$  be an admissible mesh and let  $\delta t$  be a time step. If  $u_{\mathcal{D}} = (u_{\mathcal{T}}^i)_{1 \leq i \leq N}$  is a solution to (3.8), then for all  $n > 0$  there exists  $C > 0$  only depending on  $(\Omega, T, f, b(u^0))$  such that

$$\sum_{i=1}^N \delta t \sum_{\sigma \in \mathcal{E}} \frac{|\sigma|}{d_{\sigma}} \lambda(u)_{\sigma}^i (u_K^i - u_L^i) (T_n(u_K^i) - T_n(u_L^i)) \leq Cn. \quad (3.14)$$

**Proof** By multiplying each equation of the scheme by  $T_n(u_K^i)$ , by summing over each control volume and over each time step and by reordering the sums, we obtain  $T_1 + T_2 + T_3 = T_4$  with

$$\begin{aligned} T_1 &= \sum_{i=1}^N \delta t \sum_{K \in \mathcal{T}} |K| \frac{b(u_K^i) - b(u_K^{i-1})}{\delta t} T_n(u_K^i), \\ T_2 &= \sum_{i=1}^N \delta t \sum_{\sigma \in \mathcal{E}} \frac{|\sigma|}{d_{\sigma}} \lambda(u)_{\sigma}^i (u_K^i - u_L^i) (T_n(u_K^i) - T_n(u_L^i)), \\ T_3 &= \sum_{i=1}^N \delta t \sum_{K \in \mathcal{T}} |K| f_K^i T_n(u_K^i). \end{aligned}$$

Since  $T_n$  is bounded by  $n$ , we deduce that  $|T_3| \leq n \|f\|_{L^1(Q)}$ . Let us define  $B_n(s) = \int_0^s b'(r) T_n(r) dr \geq 0$ . We have

$$\begin{aligned} B_n(u_K^i) - B_n(u_K^{i-1}) &= \int_{u_K^{i-1}}^{u_K^i} b'(r) T_n(r) dr \\ &= b(u_K^i) T_n(u_K^i) - b(u_K^{i-1}) T_n(u_K^{i-1}) - \int_{u_K^{i-1}}^{u_K^i} b(r) T_n'(r) dr \\ &= (b(u_K^i) - b(u_K^{i-1})) T_n(u_K^i) + b(u_K^{i-1}) (T_n(u_K^i) - T_n(u_K^{i-1})) - \int_{u_K^{i-1}}^{u_K^i} b(r) T_n'(r) dr \\ &= (b(u_K^i) - b(u_K^{i-1})) T_n(u_K^i) + b(u_K^{i-1}) \int_{u_K^{i-1}}^{u_K^i} T_n'(r) dr - \int_{u_K^{i-1}}^{u_K^i} b(r) T_n'(r) dr \\ &= (b(u_K^i) - b(u_K^{i-1})) T_n(u_K^i) + \int_{u_K^{i-1}}^{u_K^i} (b(u_K^{i-1}) - b(r)) T_n'(r) dr. \end{aligned}$$

Thus, we can rewrite  $T_1$  as

$$T_1 = \sum_{i=1}^N \sum_{K \in \mathcal{T}} |K| \left( B_n(u_K^i) - B_n(u_K^{i-1}) + \int_{u_K^{i-1}}^{u_K^i} (b(r) - b(u_K^{i-1})) T_n'(r) dr \right).$$

Since  $\int_{u_K^{i-1}}^{u_K^i} (b(r) - b(u_K^{i-1})) T_n'(r) dr$  is positive, we deduce

$$\begin{aligned} T_1 &\geq \sum_{i=1}^N \sum_{K \in \mathcal{T}} |K| (B_n(u_K^i) - B_n(u_K^{i-1})) \\ &= \sum_{K \in \mathcal{T}} |K| B_n(u_K^N) - \sum_{K \in \mathcal{T}} |K| B_n(u_K^0). \end{aligned}$$

Then we obtain

$$\begin{aligned} \sum_{K \in \mathcal{T}} |K| B_n(u_K^N) + \sum_{i=1}^N \delta t \sum_{\sigma \in \mathcal{E}} \frac{|\sigma|}{d_\sigma} \lambda(u)_\sigma^i (u_K^i - u_L^i) (T_n(u_K^i) - T_n(u_L^i)) \\ \leq n \|f\|_{L^1(Q)} + \sum_{K \in \mathcal{T}} |K| B_n(u_K^0). \end{aligned}$$

Noticing that  $\sum_{K \in \mathcal{T}} |K| B_n(u_K^N) \geq 0$  and  $\sum_{K \in \mathcal{T}} |K| B_n(u_K^0) \leq n \|b(u^0)\|_{L^1(\Omega)}$ , we conclude the proof.  $\square$

**Remark 3.3.2** In view of (3.14) since  $|T_n(u_K^i) - T_n(u_L^i)| \leq |u_K^i - u_L^i|$  and  $\mu > 0$  we also get

$$\sum_{i=1}^N \delta t \|T_n(u_{\mathcal{T}}^i)\|_{1,2,\mathcal{T}} \leq Cn,$$

and the discrete Poincaré inequality leads to  $\|T_n(u_{\mathcal{D}})\|_{L^2(Q)}^2 \leq Cn$ , thus we deduce that

$$\text{meas}\{|u_{\mathcal{D}}| > n\} \leq \frac{C}{n}, \quad (3.15)$$

for any  $n > 0$ .

Let us introduce the function  $h_n$  defined by

$$h_n(r) = \begin{cases} 0 & \text{if } r \leq -2n, \\ \frac{r}{n} + 2 & \text{if } -2n \leq r \leq -n, \\ 1 & \text{if } -n \leq r \leq n, \\ \frac{-r}{n} + 2 & \text{if } n \leq r \leq 2n, \\ 0 & \text{if } r \geq 2n. \end{cases} \quad (3.16)$$

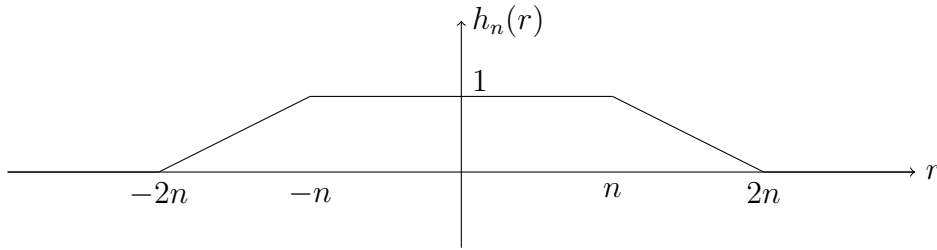


FIGURE 3.2 – The function  $h_n$

and let us also introduce the function  $s_n$  defined by

$$s_n(r) = \begin{cases} -1 & \text{if } r \leq -2n, \\ \frac{r}{n} + 1 & \text{if } -2n \leq r \leq -n, \\ 0 & \text{if } -n \leq r \leq n, \\ \frac{r}{n} - 1 & \text{if } n \leq r \leq 2n, \\ 1 & \text{if } r \geq 2n. \end{cases} \quad (3.17)$$

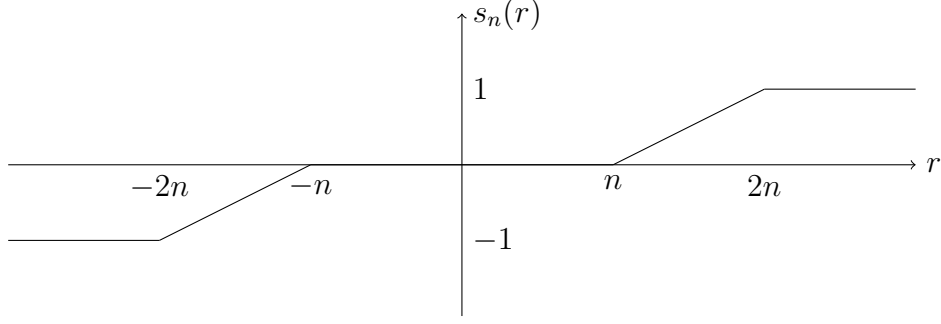


FIGURE 3.3 – The function  $s_n$

In the two following propositions we adapt the strategy developed in [21] in the continuous case. We prove estimates on  $\tilde{h}_n(b(u_{\mathcal{D}}))$  and on  $\partial_t \tilde{h}_n(u_{\mathcal{D}})$  where  $\tilde{h}_n$  is a smooth truncation function.

**Proposition 3.3.3 (Estimate on  $\tilde{h}_n(b(u_{\mathcal{D}}))$ )** *Let  $\mathcal{T}$  be an admissible mesh and let  $\delta t$  be a time step. If  $u_{\mathcal{D}} = (u_{\mathcal{T}}^i)_{1 \leq i \leq N}$  is a solution to (3.8), then for all  $n > 0$  there exists  $C > 0$  only depending on  $(\Omega, T, f, b(u^0), n)$  such that*

$$\sum_{i=1}^N \delta t \|\tilde{h}_n(b(u_{\mathcal{T}}^i))\|_{1,2,\mathcal{T}}^2 \leq C, \quad (3.18)$$

where  $\tilde{h}_n(s) = \int_0^s h_n(r) dr$ .

**Proof** We can perform the proof thanks to (3.14). As in [21], due to the properties of the function  $b$  if  $|b(s)| \leq n$ , then  $|s| \leq M_n$  where  $M_n$  is defined by  $M_n = \max\{-b^{-1}(-n); b^{-1}(n)\}$ . Thus

$$\begin{aligned} \sum_{i=1}^N \delta t \|\tilde{h}_n(b(u_{\mathcal{T}}^i))\|_{1,2,\mathcal{T}}^2 &= \sum_{i=1}^N \delta t \sum_{\sigma \in \mathcal{E}} \frac{|\sigma|}{d_{\sigma}} |\tilde{h}_n(b(u_K^i)) - \tilde{h}_n(b(u_L^i))|^2 \\ &= \sum_{i=1}^N \delta t \sum_{\sigma \in \mathcal{E}} \frac{|\sigma|}{d_{\sigma}} |\tilde{h}_n(T_{2n}(b(u_K^i))) - \tilde{h}_n(T_{2n}(b(u_L^i)))|^2 \\ &= \sum_{i=1}^N \delta t \sum_{\sigma \in \mathcal{E}} \frac{|\sigma|}{d_{\sigma}} |\tilde{h}_n(b(T_{M_{2n}}(u_K^i))) - \tilde{h}_n(b(T_{M_{2n}}(u_L^i)))|^2 \\ &\leq \left( \max_{r \in [b^{-1}(-2n), b^{-1}(2n)]} |b'(r)| \right)^2 \sum_{i=1}^N \delta t \sum_{\sigma \in \mathcal{E}} \frac{|\sigma|}{d_{\sigma}} |T_{M_{2n}}(u_K^i) - T_{M_{2n}}(u_L^i)|^2 \\ &= \left( \max_{r \in [b^{-1}(-2n), b^{-1}(2n)]} |b'(r)| \right)^2 \sum_{i=1}^N \delta t \|T_{M_{2n}}(u_{\mathcal{T}}^i)\|_{1,2,\mathcal{T}}^2, \end{aligned}$$

where  $M_{2n} = \max\{-b^{-1}(-2n); b^{-1}(2n)\}$ . □

**Proposition 3.3.4 (Estimate on  $\partial_t \tilde{h}_n(u_{\mathcal{D}})$ )** Let  $\mathcal{T}$  be an admissible mesh and let  $\delta t$  be a time step. If  $u_{\mathcal{D}} = (u_{\mathcal{T}}^i)_{1 \leq i \leq N}$  is a solution to (3.8), then for all  $n > 0$  there exists  $C > 0$  only depending on  $(\Omega, T, \lambda, f, b(u^0), n)$  such that

$$\sum_{i=1}^N \delta t \|\partial_t \tilde{h}_n(b(u_{\mathcal{T}}^i))\|_{-1,1,\mathcal{T}} \leq C. \quad (3.19)$$

Before proving Proposition 3.3.4 we establish Lemma 3.3.5 which allows one to control additional terms with respect to the continuous case, see (3.12) in the continuous case and (3.13) in the discrete one.

**Lemma 3.3.5** Let  $(\mathcal{T}_m)_{m \geq 1}$  be a sequence of admissible meshes and  $(\delta t_m)_{m \geq 1}$  be a sequence of time steps. If  $u_{\mathcal{D}} = (u_{\mathcal{T}}^i)_{1 \leq i \leq N}$  is a solution to (3.8), then for all  $n > 0$  there exists  $C > 0$  only depending on  $(\Omega, T, f, b(u^0), n)$  such that

$$\sum_{i=1}^N \sum_{K \in \mathcal{T}} |K| \int_{b(u_K^{i-1})}^{b(u_K^i)} (r - b(u_K^{i-1})) s'_n(r) dr \leq C. \quad (3.20)$$

**Proof** By multiplying each equation of the scheme by  $s_n(b(u_K^i))$ , by summing over each control volume and each time step and reordering the sums, we get  $S_1 + S_2 = S_3$  where

$$\begin{aligned} S_1 &= \sum_{i=1}^N \delta t \sum_{K \in \mathcal{T}} |K| \frac{b(u_K^i) - b(u_K^{i-1})}{\delta t} s_n(b(u_K^i)), \\ S_2 &= \sum_{i=1}^N \delta t \sum_{\sigma \in \mathcal{E}} \frac{|\sigma|}{d_\sigma} (u_K^i - u_L^i) (s_n(b(u_K^i)) - s_n(b(u_L^i))), \\ S_3 &= \sum_{i=1}^N \delta t \sum_{K \in \mathcal{T}} |K| f_K^i s_n(b(u_K^i)). \end{aligned}$$

By the definition of  $s_n$  we have that  $S_3 \leq \|f\|_{L^1(Q)}$  and  $S_2 \geq 0$ , thus  $S_1 \leq \|f\|_{L^1(Q)}$ . We now turn to  $S_1$  and we note  $\tilde{s}_n(v) = \int_0^v s_n(r) dr$ .

$$\begin{aligned} S_1 &= \sum_{i=1}^N \sum_{K \in \mathcal{T}} |K| \left[ \tilde{s}_n(b(u_K^i)) - \tilde{s}_n(b(u_K^{i-1})) + \int_{b(u_K^{i-1})}^{b(u_K^i)} (r - b(u_K^{i-1})) s'_n(r) dr \right] \\ &= \sum_{K \in \mathcal{T}} |K| \tilde{s}_n(b(u_K^N)) - \sum_{K \in \mathcal{T}} |K| \tilde{s}_n(b(u_K^0)) + \sum_{i=1}^N \sum_{K \in \mathcal{T}} |K| \int_{b(u_K^{i-1})}^{b(u_K^i)} (r - b(u_K^{i-1})) s'_n(r) dr. \end{aligned}$$

Since  $\sum_{K \in \mathcal{T}} |K| \tilde{s}_n(b(u_K^0)) \leq \|b(u^0)\|_{L^1(\Omega)}$  and  $\sum_{K \in \mathcal{T}} |K| \tilde{s}_n(b(u_K^N)) \geq 0$  we obtain that  $\sum_{i=1}^N \sum_{K \in \mathcal{T}} |K| \int_{b(u_K^{i-1})}^{b(u_K^i)} (r - b(u_K^{i-1})) s'_n(r) dr$  is bounded independently of  $n$ .  $\square$

We are now in position to show Proposition 3.3.4.

**Proof of Proposition 3.3.4** Let  $w \in X(\mathcal{T})$ . Multiplying each equation of the scheme by  $h_n(b(u_K^i))w_K$ , summing over the control volumes and gathering by edges, we get  $T_1^i + T_2^i + T_3^i = T_4^i$  with

$$\begin{aligned} T_1^i &= \sum_{K \in \mathcal{T}} |K| \frac{b(u_K^i) - b(u_K^{i-1})}{\delta t} h_n(b(u_K^i)) w_K, \\ T_2^i &= \sum_{\sigma \in \mathcal{E}} \frac{|\sigma|}{d_\sigma} \lambda(u)_\sigma^i (u_K^i - u_L^i) (h_n(b(u_K^i)) w_K - h_n(b(u_L^i)) w_L), \\ T_3^i &= \sum_{K \in \mathcal{T}} |K| f_K^i h_n(b(u_K^i)) w_K. \end{aligned}$$

Since  $h_n$  is bounded by 1 we have  $|T_3^i| \leq \|w\|_{L^\infty(\Omega)} \|f^i\|_{L^1(\Omega)}$  where  $f^i = \frac{1}{\delta t} \int_{t_{i-1}}^{t_i} f \, dt$ . We thus get

$$\sum_{i=1}^N \delta t \sup_{\substack{w \in X(\mathcal{T}) \\ w \neq 0}} \frac{1}{\|w\|_{1,\infty,\mathcal{T}}} |T_3^i| \leq \|f\|_{L^1(Q)}. \quad (3.21)$$

As far as  $T_2^i$  is concerned we have  $T_2^i = T_{2,1}^i + T_{2,2}^i$  with

$$\begin{aligned} T_{2,1}^i &= \sum_{\sigma \in \mathcal{E}} \frac{|\sigma|}{d_\sigma} \lambda(u)_\sigma^i h_n(b(u_K^i)) (u_K^i - u_L^i) (w_K - w_L), \\ T_{2,2}^i &= \sum_{\sigma \in \mathcal{E}} \frac{|\sigma|}{d_\sigma} \lambda(u)_\sigma^i w_L (u_K^i - u_L^i) (h_n(b(u_K^i)) - h_n(b(u_L^i))). \end{aligned}$$

In view of the definitions of  $h_n$  and  $b$  we get

$$|T_{2,2}^i| \leq \frac{\|w\|_{L^\infty(\Omega)}}{n} \max_{r \in [b^{-1}(-2n), b^{-1}(2n)]} |b'(r)| \sum_{\sigma \in \mathcal{E}} \frac{|\sigma|}{d_\sigma} \lambda(u)_\sigma^i (u_K^i - u_L^i) (T_{M_{2n}}(u_K^i) - T_{M_{2n}}(u_L^i)),$$

where  $M_{2n} = \max\{-b^{-1}(-2n); b^{-1}(2n)\}$ .

Taking the sup on  $X(\mathcal{T})$  and summing over the time steps we obtain that

$$\begin{aligned} & \sum_{i=1}^N \delta t \sup_{\substack{w \in X(\mathcal{T}) \\ w \neq 0}} \frac{1}{\|w\|_{1,\infty,\mathcal{T}}} |T_{2,2}^i| \\ & \leq \frac{\left( \max_{r \in [b^{-1}(-2n), b^{-1}(2n)]} |b'(r)| \right)}{n} \sum_{i=1}^N \delta t \sum_{\sigma \in \mathcal{E}} \frac{|\sigma|}{d_\sigma} \lambda(u)_\sigma^i (u_K^i - u_L^i) (T_{M_{2n}}(u_K^i) - T_{M_{2n}}(u_L^i)), \end{aligned} \quad (3.22)$$

which is bounded thanks to the estimate (3.14).



We now turn to  $T_{2,1}^i$ . Since only  $u_K^i$  is truncated in  $T_{2,1}^i$  we have to split it into

$$\begin{aligned} T_{2,1}^i &= \sum_{\sigma \in \mathcal{E}} \frac{|\sigma|}{d_\sigma} \lambda(u)_\sigma^i h_n(b(u_K^i))(u_K^i - u_L^i)(w_K - w_L) \\ &= \underbrace{\sum_{\sigma \in \mathcal{E}} \frac{|\sigma|}{d_\sigma} \lambda(u)_\sigma^i h_n(b(u_K^i)) h_{2n}(b(u_L^i))(u_K^i - u_L^i)(w_K - w_L)}_I \\ &\quad + \underbrace{\sum_{\sigma \in \mathcal{E}} \frac{|\sigma|}{d_\sigma} \lambda(u)_\sigma^i h_n(b(u_K^i))(1 - h_{2n}(b(u_L^i)))(u_K^i - u_L^i)(w_K - w_L)}_{II}. \end{aligned}$$

We have

$$\begin{aligned} h_n(b(u_K^i))(1 - h_{2n}(b(u_L^i))) &= h_n(b(u_K^i))(h_{2n}(b(u_K^i)) - h_{2n}(b(u_L^i))) \\ &= h_n(b(u_K^i))(h_{2n}(T_{4n}(b(u_K^i))) - h_{2n}(T_{4n}(b(u_L^i)))) \\ &= h_n(b(u_K^i))(h_{2n}(b(T_{M_{4n}}(u_K^i))) - h_{2n}(b(T_{M_{4n}}(u_L^i))))), \end{aligned}$$

where  $M_{4n} = \max\{-b^{-1}(-4n); b^{-1}(4n)\}$ . Thus

$$|II| \leq \frac{\|w\|_{L^\infty(\Omega)}}{n} \left( \max_{r \in [b^{-1}(-4n), b^{-1}(4n)]} |b'(r)| \right) \sum_{\sigma \in \mathcal{E}} \frac{|\sigma|}{d_\sigma} \lambda(u)_\sigma^i (T_{M_{4n}}(u_K^i) - T_{M_{4n}}(u_L^i))(u_K^i - u_L^i).$$

We deduce that

$$\sum_{i=1}^N \delta t \sup_{\substack{w \in X(\mathcal{T}) \\ w \neq 0}} \frac{1}{\|w\|_{1,\infty,\mathcal{T}}} |II| \leq C. \quad (3.23)$$

We now turn to  $I$  :

$$\begin{aligned} I &= \sum_{\sigma \in \mathcal{E}} \frac{|\sigma|}{d_\sigma} \lambda(u)_\sigma^i h_n(b(u_K^i)) h_{2n}(b(u_L^i))(u_K^i - u_L^i)(w_K - w_L) \\ &= \sum_{\sigma \in \mathcal{E}} \frac{|\sigma|}{d_\sigma} \lambda(u)_\sigma^i h_n(T_{2n}(b(u_K^i))) h_{2n}(T_{4n}(b(u_L^i)))(u_K^i - u_L^i)(w_K - w_L) \\ &= \sum_{\sigma \in \mathcal{E}} \frac{|\sigma|}{d_\sigma} \lambda(u)_\sigma^i h_n(b(T_{M_{2n}}(u_K^i))) h_{2n}(b(T_{M_{4n}}(u_L^i)))(u_K^i - u_L^i)(w_K - w_L) \\ &= \sum_{\sigma \in \mathcal{E}} |\sigma| \lambda(u)_\sigma^i h_n(b(T_{M_{2n}}(u_K^i))) h_{2n}(b(T_{M_{4n}}(u_L^i)))(T_{M_{4n}}(u_K^i) - T_{M_{4n}}(u_L^i)) \frac{w_K - w_L}{d_\sigma}, \end{aligned}$$

Thanks to the definition of  $\lambda(u)_\sigma^i$  we get

$$|I| \leq \|w\|_{1,\infty,\mathcal{T}} \max_{r \in [-M_{4n}, M_{4n}]} |\lambda(r)| \|T_{M_{4n}}(u_{\mathcal{T}}^i)\|_{1,1,\mathcal{T}},$$

thus

$$\sum_{i=1}^N \delta t \sup_{\substack{w \in X(\mathcal{T}) \\ w \neq 0}} \frac{1}{\|w\|_{1,\infty,\mathcal{T}}} |I| \leq C. \quad (3.24)$$

Gathering (3.22), (3.23) and (3.24) we obtain that

$$\sum_{i=1}^N \delta t \sup_{\substack{w \in X(\mathcal{T}) \\ w \neq 0}} \frac{1}{\|w\|_{1,\infty,\mathcal{T}}} |T_2^i| \leq C. \quad (3.25)$$

As far as  $T_1^i$  is concerned we have  $T_1^i = T_{1,1}^i + T_{1,2}^i$  with

$$\begin{aligned} T_{1,1}^i &= \sum_{K \in \mathcal{T}} |K| \frac{\tilde{h}_n(b(u_K^i)) - \tilde{h}_n(b(u_K^{i-1}))}{\delta t} w_K, \\ T_{1,2}^i &= \sum_{K \in \mathcal{T}} \frac{|K|}{\delta t} w_K \int_{b(u_K^{i-1})}^{b(u_K^i)} (r - b(u_K^{i-1})) h'_n(r) dr, \end{aligned}$$

where  $\tilde{h}_n(v) = \int_0^v h_n(r) dr$ .

By the definition of the dual norm  $\|\cdot\|_{-1,1,\mathcal{T}}$  we have

$$\sum_{i=1}^N \delta t \sup_{\substack{w \in X(\mathcal{T}) \\ w \neq 0}} \frac{1}{\|w\|_{1,\infty,\mathcal{T}}} T_{1,1}^i = \sum_{i=1}^N \delta t \|\partial_t \tilde{h}_n(b(u_{\mathcal{T}}^i))\|_{-1,1,\mathcal{T}}. \quad (3.26)$$

In order to control  $T_{1,2}^i$ , we notice that

$$\left| \int_{b(u_K^{i-1})}^{b(u_K^i)} (r - b(u_K^{i-1})) h'_n(r) dr \right| \leq \int_{b(u_K^{i-1})}^{b(u_K^i)} (r - b(u_K^{i-1})) s'_n(r) dr. \quad (3.27)$$

Then by (3.20) of Lemma 3.3.5 we get

$$\sum_{i=1}^N \delta t \sup_{\substack{w \in X(\mathcal{T}) \\ w \neq 0}} \frac{1}{\|w\|_{1,\infty,\mathcal{T}}} |T_{1,2}^i| \leq C. \quad (3.28)$$

Gathering (3.21) to (3.28), we get the estimate

$$\sum_{i=1}^N \delta t \|\partial_t \tilde{h}_n(b(u_{\mathcal{T}}^i))\|_{-1,1,\mathcal{T}} \leq C. \quad (3.29)$$

□

We are now in position to extract a pointwise convergent subsequence to  $T_n(u_{\mathcal{D}})$ .

**Proposition 3.3.6** *Let  $(\mathcal{T}_m)_{m \geq 1}$  be a sequence of admissible meshes such that there exists  $\xi > 0$  satisfying*

$$\text{for all } m \geq 1 \text{ for all } K \in \mathcal{T} \text{ and all } \sigma \in \mathcal{E}(K), \quad d_{K,\sigma} \geq \xi d_\sigma,$$

*and let  $\delta t_m$  be a sequence of time steps. If  $u_{\mathcal{D}_m} = (u_{\mathcal{T}_m}^i)_{1 \leq i \leq N}$  is a solution to (3.8) with  $\mathcal{T} = \mathcal{T}_m$  and  $\delta t = \delta t_m$ , then there exists a measurable function  $u$  finite a.e. in  $Q$  such that  $T_n(u) \in L^2(0, T; H_0^1(\Omega))$  and, up to a subsequence,  $\nabla_{\mathcal{D}} T_n(u_{\mathcal{D}_m})$  converges to  $\nabla T_n(u)$  weakly in  $L^2(Q)^d$  and  $u_{\mathcal{D}_m}$  converges almost everywhere to  $u$  in  $Q$ .*

In order to prove this proposition we use a discrete version of the Aubin-Simon Lemma (see [47, 48]).

**Proof** Thanks to estimate (3.14) we can extract a subsequence, still indexed by  $m$ , such that  $v_n \in L^2(0, T; H_0^1(\Omega))$  and

$$\nabla_{\mathcal{D}} T_n(u_{\mathcal{D}_m}) \rightharpoonup \nabla v_n \text{ in } L^2(Q)^d$$

in the finite volume sense.

Moreover, since we have estimates (3.18) and (3.19) we also apply the discrete version of the Aubin-Simon Lemma to the function  $\tilde{h}_n(b(u_{\mathcal{D}_m}))$ . It implies that for any  $n > 0$ ,  $\tilde{h}_n(b(u_{\mathcal{D}_m}))$  lies in a compact set of  $L^1(Q)$ . Using the diagonal process, we can deduce that there exists a subsequence of  $b(u_{\mathcal{D}_m})$ , still indexed by  $m$ , such that  $\tilde{h}_n(b(u_{\mathcal{D}_m}))$  converges almost everywhere to a measurable function  $w_n$  defined on  $Q$  for any  $n > 0$ . Due to the definition of  $\tilde{h}_n$  and the hypothesis on  $b$ , the sequence  $T_n(b(u_{\mathcal{D}_m}))$  converges almost everywhere in  $Q$ , which in turn implies that  $T_{M_n}(u_{\mathcal{D}_m})$  converges almost everywhere in  $Q$  to the measurable function  $v_n$  for any  $n > 0$ .

Using Inequality (3.15) we now prove that  $u_{\mathcal{D}_m}$  is a Cauchy sequence in measure. Let  $\omega > 0$ . For all  $n > 0$  and all sequences of admissible meshes and time steps  $(\mathcal{T}_m, \delta t_m)_{m \geq 1}$  and  $(\mathcal{T}_p, \delta t_p)_{p \geq 1}$ , we have

$$\{|u_{\mathcal{D}_m} - u_{\mathcal{D}_p}| > \omega\} \subset \{|u_{\mathcal{D}_m}| > n\} \cup \{|u_{\mathcal{D}_p}| > n\} \cup \{|T_n(u_{\mathcal{D}_m}) - T_n(u_{\mathcal{D}_p})| > \omega\}.$$

Let  $\varepsilon > 0$  fixed. By (3.15), let  $n > 0$  such that, for all admissible meshes and time steps  $\mathcal{T}_m, \delta t_m$  and  $\mathcal{T}_p, \delta t_p$

$$\text{meas}(\{|u_{\mathcal{D}_m}| > n\}) + \text{meas}(\{|u_{\mathcal{D}_p}| > n\}) < \frac{\varepsilon}{2}.$$

Once  $n$  is chosen, we use the fact that  $T_n(u_{\mathcal{D}_m})$  is a Cauchy sequence in measure, thus

$$\exists h_0 > 0; \forall h_{\mathcal{D}_m}, h_{\mathcal{D}_p} < h_0 \quad \text{meas}(\{|T_n(u_{\mathcal{D}_m}) - T_n(u_{\mathcal{D}_p})| > \omega\}) \leq \frac{\varepsilon}{2}.$$

Therefore, we deduce that  $\forall h_{\mathcal{D}_m}, h_{\mathcal{D}_p} < h_0$

$$\text{meas}\{|u_{\mathcal{D}_m} - u_{\mathcal{D}_p}| > \omega\} < \varepsilon.$$

Hence  $u_{\mathcal{D}_m}$  is a Cauchy sequence in measure. It follows that, up to a subsequence still indexed by  $\mathcal{D}_m$ , there exists a measurable function  $u$  such that  $u_{\mathcal{D}_m} \rightarrow u$  a.e. in  $Q$ . Due to Inequality (3.15)  $u$  is finite almost everywhere. Moreover from convergences obtained before we get that  $T_n(u) \in L^2(0, T; H_0^1(\Omega))$  and

$$\begin{aligned} \nabla_{\mathcal{D}} T_n(u_{\mathcal{D}}) &\rightharpoonup \nabla T_n(u) \text{ in } (L^2(Q))^d, \\ T_n(u_{\mathcal{D}}) &\rightarrow T_n(u) \text{ almost everywhere in } Q. \end{aligned}$$

□

In the following proposition we prove a uniform estimate on the truncate energy of  $u_{\mathcal{D}}$  (see (3.30)) which is crucial to pass to the limit in the approximate problem. We explicitly observe that (3.30) is the discrete version of (3.4) which is imposed in the definition of the renormalized solution. Moreover we have to control the additional term (3.31) in the discrete case. The pointwise convergence of  $u_{\mathcal{D}_m}$  allows one to improve the behavior of the additional term coming from (3.13) with respect to (3.20). Instead of a bound we show a decay to zero, see (3.31), which is also crucial in the convergence analysis.

**Proposition 3.3.7** *Let  $(\mathcal{T}_m)_{m \geq 1}$  be a sequence of admissible meshes such that there exists  $\xi > 0$  satisfying*

$$\text{for all } m \geq 1 \text{ for all } K \in \mathcal{T} \text{ and all } \sigma \in \mathcal{E}(K), d_{K,\sigma} \geq \xi d_\sigma,$$

and  $(\delta t_m)_{m \geq 1}$  a sequence of times steps.

If  $u_{\mathcal{D}_m} = (u_{\mathcal{T}_m}^i)_{1 \leq i \leq N}$  is a solution to (3.8), then

$$\lim_{n \rightarrow +\infty} \overline{\lim}_{\substack{h_{\mathcal{T}_m} \rightarrow 0 \\ \delta t_m \rightarrow 0}} \frac{1}{n} \sum_{i=1}^N \delta t \sum_{\sigma \in \mathcal{E}} \frac{|\sigma|}{d_\sigma} \lambda(u)_\sigma^i (u_K^i - u_L^i) (T_n(u_K^i) - T_n(u_L^i)) = 0, \quad (3.30)$$

where  $u_L^i = 0$  if  $\sigma \in \mathcal{E}_{ext}$  and

$$\lim_{n \rightarrow +\infty} \overline{\lim}_{\substack{h_{\mathcal{T}_m} \rightarrow 0 \\ \delta t_m \rightarrow 0}} \sum_{i=1}^N \sum_{K \in \mathcal{T}} |K| \int_{b(u_K^{i-1})}^{b(u_K^i)} (r - b(u_K^{i-1})) s'_n(r) dr = 0. \quad (3.31)$$

**Proof** We first prove (3.30). Let  $\mathcal{T}$  be an admissible mesh and  $\delta t$  a time step and let  $u_{\mathcal{D}}$  be a solution of (3.8). Multiplying each equation of the scheme by  $\frac{T_n(u_K^i)}{n}$ , summing on  $K \in \mathcal{T}$  and gathering by edges lead to  $T_1 + T_2 = T_3$  with

$$T_1 = \frac{1}{n} \sum_{i=1}^M \delta t \sum_{K \in \mathcal{T}} |K| \frac{b(u_K^i) - b(u_K^{i-1})}{\delta t} T_n(u_K^i), \quad (3.32)$$

$$T_2 = \frac{1}{n} \sum_{i=1}^M \delta t \sum_{\sigma \in \mathcal{E}} \frac{|\sigma|}{d_\sigma} \lambda(u)_\sigma^i (u_K^i - u_L^i) (T_n(u_K^i) - T_n(u_L^i)), \quad (3.33)$$

$$T_3 = \frac{1}{n} \sum_{i=1}^M \delta t \sum_{K \in \mathcal{T}} |K| f_K^i T_n(u_K^i). \quad (3.34)$$

As far as  $T_1$  is concerned, reusing the computations of the proof of (3.14) we find that

$$T_1 \geq \frac{1}{n} \sum_{K \in \mathcal{T}} |K| B_n(u_K^N) - \frac{1}{n} \sum_{K \in \mathcal{T}} |K| B_n(u_K^0),$$

where  $B_n(s) = \int_0^s b'(r) T_n(r) dr$ .

We have

$$\frac{1}{n} \sum_{K \in \mathcal{T}} |K| B_n(u_K^N) \geq 0, \quad (3.35)$$

and

$$\lim_{n \rightarrow +\infty} \lim_{\substack{h_{\mathcal{T}} \rightarrow 0 \\ \delta t \rightarrow 0}} \frac{1}{n} \sum_{K \in \mathcal{T}} |K| B_n(u_K^0) = 0. \quad (3.36)$$

Moreover, since  $u_{\mathcal{D}}$  is finite almost everywhere, we deduce that

$$\lim_{n \rightarrow +\infty} \lim_{\substack{h_{\mathcal{T}} \rightarrow 0 \\ \delta t \rightarrow 0}} T_3 = 0. \quad (3.37)$$

Combining (3.35), (3.36) and (3.37), we obtain the desired estimate.

We now turn to (3.31). Since the limit  $u$  is also finite almost everywhere, we have in fact more that (3.20). Indeed, reusing the computations of the proof of Lemma 3.3.5 we get

$$\begin{aligned} \sum_{i=1}^N \sum_{K \in \mathcal{T}} |K| \int_{b(u_K^{i-1})}^{b(u_K^i)} (r - b(u_K^{i-1})) s'_n(r) dr &\leq \int_0^T \int_{\Omega} f s_n(b(u_{\mathcal{D}})) dx dt \\ &+ \int_{\Omega} \tilde{s}_n(b(u_{\mathcal{T}}^0)) dx. \end{aligned} \quad (3.38)$$

Due to the definition of  $s_n$  and since  $b(u_{\mathcal{T}}^0)$  converges almost everywhere in  $\Omega$  to  $b(u^0)$  (which belongs to  $L^1(\Omega)$ ) the Lebesgue theorem ensures us that

$$\lim_{n \rightarrow +\infty} \lim_{\substack{h_{\mathcal{T}} \rightarrow 0 \\ \delta t \rightarrow 0}} \int_{\Omega} \tilde{s}_n(b(u_{\mathcal{T}}^0)) dx = 0.$$

Moreover  $b(u_{\mathcal{D}})$  converges almost everywhere in  $Q$  to  $b(u)$  which is finite almost everywhere and the function  $s_n$  converges to zero as  $n \rightarrow +\infty$ . Thanks to the Lebesgue theorem,  $s_n(b(u_{\mathcal{D}})) \rightarrow 0$   $L^\infty$  weak-\*. We thus obtain (3.31).  $\square$

## 3.4 Convergence analysis

Let us now state the main result of this paper.

**Theorem 3.4.1** *If  $\mathcal{T}$  is an admissible mesh and  $\delta t$  a time step, then there exists at least one solution to (3.8).*

*If  $(\mathcal{T}_m)_{m \geq 1}$  is a sequence of admissible meshes such that there exists  $\xi > 0$  satisfying*

$$\text{for all } m \geq 1 \text{ for all } K \in \mathcal{T} \text{ and all } \sigma \in \mathcal{E}(K), d_{K,\sigma} \geq \xi d_\sigma,$$

*and such that  $h_{\mathcal{T}_m} \rightarrow 0$  and if  $(\delta t_m)_{m \geq 1}$  is a sequence of time steps such that  $\delta t_m \rightarrow 0$ , then if  $u_{\mathcal{D}_m} = (u_{\mathcal{T}_m}^i)_{1 \leq i \leq N^m}$  is the solution to (3.8) with  $\mathcal{T} = \mathcal{T}_m$  and  $\delta t_m = \delta t$ ,  $u_{\mathcal{D}_m}$  converges to  $u$  in the sense that for all  $n > 0$*

$$\begin{aligned} \nabla_{\mathcal{D}} T_n(u_{\mathcal{D}}) &\rightharpoonup \nabla T_n(u) \text{ in } (L^2(Q))^d, \\ T_n(u_{\mathcal{D}}) &\rightarrow T_n(u) \text{ almost everywhere in } Q, \end{aligned}$$

*where  $T_n(u) \in L^2(0, T; H_0^1(\Omega))$  and  $u$  is the unique renormalized solution of (3.1).*

### Proof of Theorem 3.4.1

We split the proof in four steps. Step 1 is devoted to prove the existence of a solution to (3.8). In Step 2 we pass to the limit in the scheme as the size of the mesh and the time step tend to zero. It is worth noting that we take in the scheme a discrete version of what is a test function in the renormalized formulation. This is to say that we multiply each equation of the scheme by a discrete version of  $h_n(u)\varphi$  where  $\varphi \in \mathcal{C}_c^\infty(\Omega \times [0, T])$  and  $h_n$  is defined by (3.16). In Step 3 we use the method developed in [20] to show that the renormalized equation holds true. It consists in passing to the limit as  $n \rightarrow +\infty$  and using Steklov averages. In Step 4 we conclude the proof, showing that the limit  $u$  satisfies the estimate on the energy (3.4).

#### Step 1

We proceed as in as in [6] to show the existence of a solution to (3.8). Let  $1 \leq i \leq N$  be fixed. We assume that we already know the values  $(u_K^{i-1})_{K \in \mathcal{T}}$ . Recalling that  $\text{card}(\mathcal{T})$  is the number of control volumes  $K$  in the mesh, let  $\mathcal{U}$  be a vector of  $\mathbb{R}^{\text{card}(\mathcal{T})}$  with entries  $(u_K^i)_{K \in \mathcal{T}}$ . Let us consider the operator  $\Phi$  defined with the scheme by

$$\Phi(\mathcal{U}) = \left[ |K| \frac{b(u_K^i) - b(u_K^{i-1})}{\delta t} + \sum_{\sigma \in \mathcal{E}(K)} \frac{|\sigma|}{d_\sigma} \lambda(u)_\sigma^i (u_K^i - u_L^i) - |K| f_K^i \right]_{K \in \mathcal{T}}.$$

To find a solution of the scheme (3.8) is to find a solution of equation  $\Phi(\mathcal{U}) = 0$ . Let us consider the scalar product  $\Phi(\mathcal{U}) \cdot \mathcal{U}$  in  $\mathbb{R}^{\text{card}(\mathcal{T})}$ .

$$\begin{aligned} \Phi(\mathcal{U}) \cdot \mathcal{U} &= \frac{1}{\delta t} \sum_{K \in \mathcal{T}} |K| b(u_K^i) u_K^i - \frac{1}{\delta t} \sum_{K \in \mathcal{T}} |K| b(u_K^{i-1}) u_K^i + \sum_{\sigma \in \mathcal{E}} \lambda(u)_\sigma^i (u_K^i - u_L^i)^2 \\ &\quad - \sum_{K \in \mathcal{T}} |K| f_K^i u_K^i. \end{aligned}$$

Thanks to the properties of the function  $b$  we have

$$\frac{1}{\delta t} \sum_{K \in \mathcal{T}} |K| b(u_K^i) u_K^i \geq 0.$$

Since all the norms are equivalent on  $\mathbb{R}^{\text{card}(\mathcal{T})}$ , we get

$$\frac{1}{\delta t} \sum_{K \in \mathcal{T}} |K| b(u_K^{i-1}) u_K^i \leq |\mathcal{U}| \frac{1}{\delta t} \sum_{K \in \mathcal{T}} |K| b(u_K^{i-1}) \leq C_1 |\mathcal{U}|,$$

where  $|\mathcal{U}|$  denotes a norm of  $\mathcal{U}$  and  $C_1$  is a positive constant.

We then have using Poincaré inequality

$$\begin{aligned} \sum_{\sigma \in \mathcal{E}} \lambda(u)_\sigma^i (u_K^i - u_L^i)^2 &\geq \mu \sum_{\sigma \in \mathcal{E}} |\sigma| d_\sigma \left| \frac{u_K^i - u_L^i}{d_\sigma} \right|^2 \\ &\geq C \sum_{K \in \mathcal{T}} |K| |u_K^i|^2 \\ &\geq C_2 |\mathcal{U}|^2, \end{aligned}$$

where  $C_2$  is a positive constant. Thus we obtain that

$$\Phi(\mathcal{U}) \cdot \mathcal{U} \geq C_2 |\mathcal{U}|^2 - C_1 |\mathcal{U}| - \|f_{\mathcal{T}}^i\|_{L^1(\Omega)} |\mathcal{U}|,$$

which is positive for  $|\mathcal{U}|$  large enough. Applying the Brouwer fixed point theorem (more precisely Proposition 3.4 of [56]) we get the existence of a solution to  $\Phi(\mathcal{U}) = 0$ .

## Step 2

Let  $\varphi \in C_c^\infty(\Omega \times [0, T])$  et  $h_n$  the function defined by (3.16). We denote by  $\varphi_{\mathcal{D}}$  the function defined by  $\varphi_K^i = \varphi(x_K, t_i)$  for all  $K \in \mathcal{T}$  and all  $1 \leq i \leq N$ .

Multiplying each equation of the scheme by  $\varphi_K^{i-1} h_n(u_K^i)$  (which is a discrete version of the test function used in the renormalized formulation), summing over the control volumes, gathering by edges and summing over the time steps, we get  $T_1 + T_2 + T_3 = T_4$  with

$$\begin{aligned} T_1 &= \sum_{i=1}^N \sum_{K \in \mathcal{T}} |K| (b(u_K^i) - b(u_K^{i-1})) h_n(u_K^i) \varphi_K^{i-1}, \\ T_2 &= \sum_{i=1}^N \delta t \sum_{\sigma \in \mathcal{E}} \frac{|\sigma|}{d_\sigma} (u_K^i - u_L^i) (h_n(u_K^i) \varphi_K^{i-1} - h_n(u_L^i) \varphi_L^{i-1}), \\ T_3 &= \sum_{i=1}^N \delta t \sum_{K \in \mathcal{T}} |K| f_K^i h_n(u_K^i) \varphi_K^{i-1}. \end{aligned}$$

As far as the lower order term  $T_3$  is concerned, we have  $T_3 = T_{3,1} + T_{3,2}$  with

$$\begin{aligned} T_{3,1} &= \sum_{i=1}^N \sum_{K \in \mathcal{T}} \int_{t_{i-1}}^{t_i} \int_K f h_n(u_K^i) \varphi_K^i dx dt, \\ T_{3,2} &= \sum_{i=1}^N \sum_{K \in \mathcal{T}} \int_{t_{i-1}}^{t_i} \int_K f h_n(u_K^i) (\varphi_K^{i-1} - \varphi_K^i) dx dt. \end{aligned}$$

By the regularity of  $\varphi$ , we have  $\varphi_{\mathcal{D}} \rightarrow \varphi$  uniformly on  $Q$  as  $h_{\mathcal{T}} \rightarrow 0$  and  $\delta t \rightarrow 0$ . We now pass to the limit as  $h_{\mathcal{T}} \rightarrow 0$  and  $\delta t \rightarrow 0$ . Since  $h_n(u_{\mathcal{D}}) \rightarrow h_n(u)$  a.e. and  $L^\infty$  weak  $\star$ ,  $\varphi_{\mathcal{D}} \rightarrow \varphi$  uniformly and  $|f \varphi_{\mathcal{T}} h_n(u_{\mathcal{D}})| \leq C_\varphi |f| \in L^1(Q)$ , the Lebesgue theorem ensures that

$$\begin{aligned} \lim_{\substack{h_{\mathcal{T}} \rightarrow 0 \\ \delta t \rightarrow 0}} T_{3,1} &= \lim_{\substack{h_{\mathcal{T}} \rightarrow 0 \\ \delta t \rightarrow 0}} \int_{\Omega} f \varphi_{\mathcal{D}} h_n(u_{\mathcal{D}}) dx dt \\ &= \int_0^T \int_{\Omega} f \varphi h_n(u) dx dt. \end{aligned} \tag{3.39}$$

In view of the definition of  $h_n$  and the regularity of  $\varphi$ ,

$$|T_{3,2}| \leq C_\varphi \delta t \|f\|_{L^1(Q)},$$

thus

$$\lim_{\substack{h_{\mathcal{T}} \rightarrow 0 \\ \delta t \rightarrow 0}} T_{3,2} = 0. \tag{3.40}$$

We now study the convergence of the diffusion term. We have

$$\begin{aligned} T_2 &= \sum_{i=1}^N \delta t \sum_{\sigma \in \mathcal{E}} \frac{|\sigma|}{d_\sigma} \lambda(u)_\sigma^i (u_K^i - u_L^i) (h_n(u_K^i) \varphi_K^{i-1} - h_n(u_L^i) \varphi_L^{i-1}) \\ &= T_{2,1} + T_{2,2} \end{aligned}$$

with

$$\begin{aligned} T_{2,1} &= \sum_{i=1}^N \delta t \sum_{\sigma \in \mathcal{E}} \frac{|\sigma|}{d_\sigma} \lambda(u)_\sigma^i h_n(u_K^i) (u_K^i - u_L^i) (\varphi_K^{i-1} - \varphi_L^{i-1}), \\ T_{2,2} &= \sum_{i=1}^N \delta t \sum_{\sigma \in \mathcal{E}} \frac{|\sigma|}{d_\sigma} \lambda(u)_\sigma^i \varphi_L^{i-1} (u_K^i - u_L^i) (h_n(u_K^i) - h_n(u_L^i)). \end{aligned}$$

Not assuming any bound on the function  $\lambda$  gives additional technical difficulties in order to control  $T_{2,1}$  and  $T_{2,2}$ . In view of the definition of  $h_n$  and  $\lambda$  we get

$$|T_{2,2}| \leq \frac{c \|\varphi\|_\infty}{n} \sum_{i=1}^N \delta t \sum_{\sigma \in \mathcal{E}} \frac{|\sigma|}{d_\sigma} \lambda(u)_\sigma^i (u_K^i - u_L^i) (T_{2n}(u_K^i) - T_{2n}(u_L^i)).$$

From the discrete estimate on the energy (see (3.30) in Proposition 3.3.7), it follows that

$$\lim_{n \rightarrow +\infty} \overline{\lim}_{\substack{h_{\mathcal{T}} \rightarrow 0 \\ \delta t \rightarrow 0}} T_{2,2} = 0. \quad (3.41)$$

As far as  $T_{2,1}$  is concerned we have  $T_{2,1} = T_{2,1,1} + T_{2,1,2}$  with

$$\begin{aligned} T_{2,1,1} &= \sum_{i=1}^N \delta t \sum_{\sigma \in \mathcal{E}} \frac{|\sigma|}{d_\sigma} \lambda(u)_\sigma^i h_n(u_K^i) h_{2n}(u_L^i) (u_K^i - u_L^i) (\varphi_K^{i-1} - \varphi_L^{i-1}), \\ T_{2,1,2} &= \sum_{i=1}^N \delta t \sum_{\sigma \in \mathcal{E}} \frac{|\sigma|}{d_\sigma} \lambda(u)_\sigma^i h_n(u_K^i) (1 - h_{2n}(u_L^i)) (u_K^i - u_L^i) (\varphi_K^{i-1} - \varphi_L^{i-1}). \end{aligned}$$

Observing that  $h_n(u_K^i) (1 - h_{2n}(u_L^i)) = h_n(u_K^i) (h_{2n}(u_K^i) - h_{2n}(u_L^i))$  gives

$$|T_{2,1,2}| \leq \frac{c \|\varphi\|_\infty}{n} \sum_{i=1}^N \delta t \sum_{\sigma \in \mathcal{E}} \frac{|\sigma|}{d_\sigma} \lambda(u)_\sigma^i (T_{4n}(u_K^i) - T_{4n}(u_L^i)) (u_K^i - u_L^i),$$

and from (3.30) it follows that

$$\lim_{n \rightarrow +\infty} \overline{\lim}_{\substack{h_{\mathcal{T}} \rightarrow 0 \\ \delta t \rightarrow 0}} T_{2,1,2} = 0. \quad (3.42)$$

For  $T_{2,1,1}$  we write  $T_{2,1,1} = I + II$  with

$$\begin{aligned} I &= \sum_{i=1}^N \delta t \sum_{\sigma \in \mathcal{E}} \frac{|\sigma|}{d_\sigma} \lambda(u)_\sigma^i \frac{h_n(u_K^i) h_{2n}(u_L^i) + h_n(u_L^i) h_{2n}(u_K^i)}{2} (u_K^i - u_L^i) (\varphi_K^{i-1} - \varphi_L^{i-1}), \\ II &= \sum_{i=1}^N \delta t \sum_{\sigma \in \mathcal{E}} \frac{|\sigma|}{d_\sigma} \lambda(u)_\sigma^i \frac{h_n(u_K^i) h_{2n}(u_L^i) - h_n(u_L^i) h_{2n}(u_K^i)}{2} (u_K^i - u_L^i) (\varphi_K^{i-1} - \varphi_L^{i-1}). \end{aligned}$$



Since

$$\begin{aligned} & h_n(u_K^i)h_{2n}(u_L^i) - h_n(u_L^i)h_{2n}(u_K^i) \\ &= h_n(u_K^i)(h_{2n}(u_L^i) - h_{2n}(u_K^i)) + (h_n(u_K^i) - h_n(u_L^i))h_{2n}(u_K^i), \end{aligned}$$

and since

$$\begin{aligned} |h_{2n}(u_L^i) - h_{2n}(u_K^i)| &\leq \frac{1}{2n} |T_{4n}(u_L^i) - T_{4n}(u_K^i)|, \\ |h_n(u_L^i) - h_n(u_K^i)| &\leq \frac{1}{n} |T_{2n}(u_L^i) - T_{2n}(u_K^i)|, \end{aligned}$$

estimate (3.30) leads to

$$|II| \leq \|\varphi\|_\infty \omega(n, h_{\mathcal{T}}, \delta t), \quad (3.43)$$

with

$$\lim_{n \rightarrow +\infty} \overline{\lim}_{\substack{h_{\mathcal{T}} \rightarrow 0 \\ \delta t \rightarrow 0}} \omega(n, h_{\mathcal{T}}, \delta t) = 0.$$

As far as  $I$  is concerned, we have  $I = I_1 + I_2$  with

$$\begin{aligned} I_1 &= \sum_{i=1}^N \delta t \sum_{\sigma \in \mathcal{E}} \frac{|\sigma|}{d_\sigma} \lambda(u)_\sigma^i \frac{h_n(u_K^i)h_{2n}(u_L^i) + h_n(u_L^i)h_{2n}(u_K^i)}{2} (u_K^i - u_L^i)(\varphi_K^i - \varphi_L^i), \\ I_2 &= \sum_{i=1}^N \delta t \sum_{\sigma \in \mathcal{E}} \frac{|\sigma|}{d_\sigma} \lambda(u)_\sigma^i \frac{h_n(u_K^i)h_{2n}(u_L^i) + h_n(u_L^i)h_{2n}(u_K^i)}{2} \\ &\quad \times (u_K^i - u_L^i)(\varphi_K^{i-1} - \varphi_K^i + \varphi_L^i - \varphi_L^{i-1}), \end{aligned}$$

Both  $u_K^i$  and  $u_L^i$  being truncated in  $I_2$ , we get

$$|I_2| \leq C_\varphi(h_{\mathcal{T}} + \delta t) \left( \max_{r \in [-4n, 4n]} |\lambda(r)| \right) \sum_{i=1}^N \delta t \|T_{4n}(u_{\mathcal{T}}^i)\|_{1,1,\mathcal{T}},$$

thus

$$\lim_{\substack{h_{\mathcal{T}} \rightarrow 0 \\ \delta t \rightarrow 0}} I_2 = 0. \quad (3.44)$$

Let us define  $\bar{h}_n(x, t)$  and  $\bar{\lambda}(x, t) \forall t \in [t_{i-1}, t_i]$ ,  $1 \leq i \leq N$ ,  $\forall x \in D_\sigma$ ,  $\sigma \in \mathcal{E}$ , by

$$\begin{aligned} \bar{h}_n(x, t) &= \frac{h_n(u_K^i)h_{2n}(u_L^i) + h_n(u_L^i)h_{2n}(u_K^i)}{2}, \\ \bar{\lambda}(x, t) &= \lambda(T_{4n}(u))_\sigma^i. \end{aligned}$$

We then write  $I_1 = I_{1,1} + I_{1,2}$  with

$$\begin{aligned} I_{1,1} &= \int_0^T \int_\Omega \bar{\lambda} \bar{h}_n \nabla_{\mathcal{D}} T_{4n}(u_{\mathcal{D}}) \cdot \nabla \varphi \, dx \, dt, \\ I_{1,2} &= \sum_{i=1}^N \delta t \sum_{\sigma \in \mathcal{E}} d|D_\sigma| \lambda(u)_\sigma^i \frac{h_n(u_K^i)h_{2n}(u_L^i) + h_n(u_L^i)h_{2n}(u_K^i)}{2} \\ &\quad \frac{T_{4n}(u_K^i) - T_{4n}(u_L^i)}{d_\sigma} C(\varphi, \mathcal{T}, \delta t), \end{aligned}$$

where  $C(\varphi, \mathcal{T}, \delta t) = \frac{\varphi_K^i - \varphi_L^i}{d_\sigma} - \frac{1}{\delta t |D_\sigma|} \int_{t_{i-1}}^{t_i} \int_{D_\sigma} \nabla \varphi \cdot \eta_{K,\sigma} dx dt$ . Thanks to the regularity of  $\varphi$ , we have

$$|I_{1,2}| \leq C_\varphi (h_{\mathcal{T}} + \delta t) \left( \max_{r \in [-4n, 4n]} |\lambda(r)| \right) \sum_{i=1}^N \delta t \|T_{4n}(u_{\mathcal{T}}^i)\|_{1,1,\mathcal{T}},$$

which implies

$$\lim_{\substack{h_{\mathcal{T}} \rightarrow 0 \\ \delta t \rightarrow 0}} I_{1,2} = 0. \quad (3.45)$$

Since  $\lambda(T_{4n}(\cdot))$  is continuous and bounded, an equi-integrability argument allows one to prove that  $\bar{\lambda}$  converges to  $\lambda(T_{4n}(u))$  in  $L^\infty$  weak- $\star$ . In view of the definition of  $\bar{h}_n$  and since  $\bar{h}_n$  is constant on each diamond cell  $D_\sigma$ , we deduce that  $\bar{h}_n \rightarrow h_n(u)$  in  $L^q(Q)$  for all  $q \in [2, +\infty)$ . Moreover  $\nabla_{\mathcal{D}} T_{4n}(u_{\mathcal{D}})$  tends to  $\nabla T_{4n}(u)$  weakly in  $L^2(0, T; L^2(\Omega)^d)$ . Since  $\varphi$  belongs to  $\mathcal{C}^\infty(\Omega \times [0, T])$  we obtain

$$\lim_{\substack{h_{\mathcal{T}} \rightarrow 0 \\ \delta t \rightarrow 0}} I_1 = \int_0^T \int_{\Omega} \lambda(u) h_n(u) \nabla T_{4n}(u) \cdot \nabla \varphi dx dt. \quad (3.46)$$

It remains to study the time derivative term. Noting  $B_n(s) = \int_0^s b'(r) h_n(r) dr$  and observing that

$$B_n(u_K^i) - B_n(u_K^{i-1}) = (b(u_K^i) - b(u_K^{i-1})) h_n(u_K^i) + \int_{u_K^{i-1}}^{u_K^i} (b(u_K^{i-1}) - b(r)) h'_n(r) dr,$$

we can write  $T_1 = T_{1,1} + T_{1,2}$  with

$$\begin{aligned} T_{1,1} &= \sum_{i=1}^N \delta t \sum_{K \in \mathcal{T}} |K| \frac{B_n(u_K^i) - B_n(u_K^{i-1})}{\delta t} \varphi_K^{i-1}, \\ T_{1,2} &= \sum_{i=1}^N \sum_{K \in \mathcal{T}} |K| \varphi_K^{i-1} \int_{u_K^{i-1}}^{u_K^i} (b(r) - b(u_K^{i-1})) h'_n(r) dr. \end{aligned}$$

Reordering the sums we find that

$$\begin{aligned} T_{1,1} &= - \sum_{i=1}^N \delta t \sum_{K \in \mathcal{T}} |K| B_n(u_K^i) \frac{\varphi_K^i - \varphi_K^{i-1}}{\delta t} - \sum_{K \in \mathcal{T}} |K| B_n(u_K^0) \varphi_K^0 \\ &= - \int_0^T \int_{\Omega} B_n(u_{\mathcal{D}}) \partial_t \varphi dx dt + R_1 - \int_{\Omega} B_n(u_{\mathcal{T}}^0) \varphi(x, 0) dx + R_2, \end{aligned}$$

where

$$\begin{aligned} R_1 &= - \sum_{i=1}^N \delta t \sum_{K \in \mathcal{T}} |K| B_n(u_K^i) \left[ \frac{\varphi_K^i - \varphi_K^{i-1}}{\delta t} - \frac{1}{|K| \delta t} \int_{t_{i-1}}^{t_i} \int_K \partial_t \varphi dx dt \right], \\ R_2 &= - \sum_{K \in \mathcal{T}} \int_K B_n(u_{\mathcal{T}}^0) [\varphi_K^0 - \varphi(x, 0)] dx. \end{aligned}$$

By the regularity of  $\varphi$  we have

$$\begin{aligned}
|R_1| &\leq C_\varphi(h_{\mathcal{T}} + \delta t) \sum_{i=1}^N \delta t \sum_{K \in \mathcal{T}} |K| |B_n(u_K^i)| \\
&\leq C_\varphi(h_{\mathcal{T}} + \delta t) \max_{r \in [-b^{-1}(-2n), b^{-1}(2n)]} |b'(r)| \sum_{i=1}^N \delta t \sum_{K \in \mathcal{T}} |K| |T_{2n}(u_K^i)| \\
&\leq C_\varphi(h_{\mathcal{T}} + \delta t) \max_{r \in [-b^{-1}(-2n), b^{-1}(2n)]} |b'(r)| \|T_{2n}(u_{\mathcal{D}})\|_{L^1(Q)}
\end{aligned}$$

and  $|R_2| \leq C_\varphi(h_{\mathcal{T}}) \|b(u^0)\|_{L^1(\Omega)}$ .

Since  $u_{\mathcal{D}}$  converges a.e. to  $u$  and in  $L^\infty$  weak  $\star$  and since  $u_{\mathcal{T}}^0 = (u_K^0)_{K \in \mathcal{T}}$  converges a.e. to  $u^0$  and in  $L^\infty$  weak  $\star$  we get

$$\lim_{\substack{h_{\mathcal{T}} \rightarrow 0 \\ \delta t \rightarrow 0}} T_{1,1} = - \int_0^T \int_{\Omega} B_n(u) \partial_t \varphi \, dx \, dt - \int_{\Omega} B_n(u^0) \varphi(x, 0) \, dx. \quad (3.47)$$

Using results (3.31) of Proposition 3.3.7 we can prove that

$$\lim_{n \rightarrow +\infty} \overline{\lim}_{\substack{h_{\mathcal{T}} \rightarrow 0 \\ \delta t \rightarrow 0}} T_{1,2} = 0. \quad (3.48)$$

We are now in position to pass to the limit as  $h_{\mathcal{T}} \rightarrow 0$  and  $\delta t \rightarrow 0$  in the scheme. Gathering equations (3.39) to (3.48), we obtain that

$$\begin{aligned}
& - \int_0^T \int_{\Omega} B_n(u) \partial_t \varphi \, dx \, dt - \int_{\Omega} B_n(u^0) \varphi(x, 0) \, dx + \int_0^T \int_{\Omega} \lambda(u) h_n(u) \nabla u \cdot \nabla \varphi \, dx \, dt \\
& \quad - \int_0^T \int_{\Omega} f \varphi h_n(u) \, dx \, dt = \overline{\lim}_{\substack{h_{\mathcal{T}} \rightarrow 0 \\ \delta t \rightarrow 0}} T(n, \varphi) \quad (3.49)
\end{aligned}$$

where  $\overline{\lim}_{\substack{h_{\mathcal{T}} \rightarrow 0 \\ \delta t \rightarrow 0}} |T(n, \varphi)| \leq C_\varphi \omega(n)$  with  $\omega(n) \rightarrow 0$  as  $n \rightarrow +\infty$ .

**Step 3** In this step we prove that the renormalized equation (3.5) holds, namely

$$\begin{aligned}
& - \int_0^T \int_{\Omega} \xi_t \int_0^u S'(r) \, dr \, dx \, dt - \int_{\Omega} \xi(x, 0) \int_0^{u^0} S'(r) \, dr \, dx \\
& + \int_0^T \int_{\Omega} \lambda(u) \nabla u \cdot \nabla (\xi S'(u)) \, dx \, dt = \int_0^T \int_{\Omega} f \xi S'(u) \, dx \, dt,
\end{aligned}$$

where  $\xi \in \mathcal{C}_c^\infty(\Omega \times [0, T])$  and  $S \in W^{2,\infty}(\mathbb{R})$  is such that  $S'$  has a compact support and  $S'(u)\xi \in L^2(0, T; H_0^1(\Omega))$ . We pass to the limit in (3.49) as  $n \rightarrow +\infty$  using the techniques developed in [20] which are based on mean time regularizations. In the elliptic case, using a density argument, it is sufficient to take  $\varphi = h(u)\psi$  as a test function and to pass to the limit as  $n \rightarrow +\infty$  (see [53]). Because of the presence of the time derivative in (3.49), we can not use the same technique. This is why we use the tool of Steklov averages and Lemmas 2.1 and 2.3 of [20].

Let  $\xi \in \mathcal{C}_c^\infty([0, T] \times \Omega)$  and  $S \in W^{2, \infty}(\mathbb{R})$  such that  $S'$  has a compact support and  $S'(u)\xi \in L^2(0, T; H_0^1(\Omega))$ . By density we take  $\varphi = \xi(S'(u))_h$  as test function in (3.49) where

$$(S'(u))_h(x, t) = \frac{1}{h} \int_t^{t+h} S'(u(x, \tau)) d\tau,$$

thus

$$\begin{aligned} & - \int_0^T \int_\Omega \partial_t (\xi(S'(u))_h) B_n(u) dx dt - \int_\Omega \xi(x, 0) (S'(u))_h(x, 0) B_n(u^0) dx \\ & + \int_0^T \int_\Omega \lambda(u) h_n(u) \nabla u \cdot \nabla (\xi(S'(u))_h) \xi dx dt \\ & = \int_0^T \int_\Omega f h_n(u) \xi(S'(u))_h dx dt + T(C(\xi, S), n), \end{aligned}$$

which gives us

$$\begin{aligned} & - \int_0^T \int_\Omega \partial_t (\xi(S'(u))_h) (B_n(u) - B_n(u^0)) dx dt \\ & + \int_0^T \int_\Omega \lambda(u) h_n(u) \nabla u \cdot \nabla (\xi(S'(u))_h) dx dt \\ & = \int_0^T \int_\Omega f h_n(u) \xi(S'(u))_h dx dt + T(C(\xi, S), n). \end{aligned} \quad (3.50)$$

Using Lemma 2.1 of [20], we obtain

$$\begin{aligned} & - \int_0^T \int_\Omega \partial_t (\xi(S'(u))_h) (B_n(u) - B_n(u^0)) dx dt \\ & \geq - \int_0^T \int_\Omega \xi_t \left( B_n(u) \frac{1}{h} \int_t^{t+h} S'(u(\tau)) d\tau - \frac{1}{h} \int_t^{t+h} \int_0^{u(\tau)} S''(r) B_n(r) dr d\tau \right) dx dt \\ & \quad - \int_\Omega \xi(x, 0) \left( B_n(u^0) S'(u^0) - \int_0^{u^0} S''(r) B_n(r) dr \right) dx, \end{aligned}$$

and letting  $h$  tends to zero :

$$\begin{aligned} & \liminf_{h \rightarrow 0} \left\{ - \int_0^T \int_\Omega \partial_t (\xi(S'(u))_h) (B_n(u) - B_n(u^0)) dx dt \right\} \\ & \geq - \int_0^T \int_\Omega \xi_t \left( B_n(u) S'(u) - \int_0^u S''(r) B_n(r) dr \right) dx dt \\ & \quad - \int_\Omega \xi(x, 0) \left( B_n(u^0) S'(u^0) - \int_0^{u^0} S''(r) B_n(r) dr \right) dx. \end{aligned}$$

Injecting this last inequality in (3.50) and letting  $h$  tends to zero we get

$$\begin{aligned}
& - \int_0^T \int_{\Omega} \xi_t \left( B_n(u) S'(u) - \int_0^u S''(r) B_n(r) dr \right) dx dt \\
& - \int_{\Omega} \xi(x, 0) \left( B_n(u^0) S'(u^0) - \int_0^{u^0} S''(r) B_n(r) dr \right) dx \\
& + \int_0^T \int_{\Omega} \lambda(u) h_n(u) \nabla u \cdot \nabla (\xi S'(u)) dx dt \\
& \leq \int_0^T \int_{\Omega} f h_n(u) \xi S'(u) dx dt + T(C(\xi, S), n).
\end{aligned}$$

Passing to the limite as  $n \rightarrow +\infty$  we find

$$\begin{aligned}
& - \int_0^T \int_{\Omega} \xi_t \left( b(u) S'(u) - \int_0^u S''(r) b(r) dr \right) dx dt \\
& - \int_{\Omega} \xi(x, 0) \left( b(u^0) S'(u^0) - \int_0^{u^0} S''(r) b(r) dr \right) dx \\
& + \int_0^T \int_{\Omega} \lambda(u) \nabla u \cdot \nabla (\xi S'(u)) dx dt \\
& \leq \int_0^T \int_{\Omega} f \xi S'(u) dx dt.
\end{aligned}$$

Noticing that  $b(z) S'(z) - \int_0^z S''(r) b(r) dr = \int_0^z S'(r) b'(r) dr$  we finally get

$$\begin{aligned}
& - \int_0^T \int_{\Omega} \xi_t \int_0^u S'(r) b'(r) dr dx dt - \int_{\Omega} \xi(x, 0) \int_0^{u^0} S'(r) b'(r) dr dx \\
& + \int_0^T \int_{\Omega} \lambda(u) \nabla u \cdot \nabla (\xi S'(u)) dx dt \leq \int_0^T \int_{\Omega} f \xi S'(u) dx dt. \tag{3.51}
\end{aligned}$$

We prove now the inverse inequality taking  $\varphi = \xi(S'(u))_{-h}$  where

$$(S'(u))_{-h}(x, t) = \frac{1}{h} \int_{t-h}^t S'(\hat{u}(x, \tau)) d\tau,$$

with

$$\hat{u}(x, t) = \begin{cases} u(x, t) & \text{if } t \geq 0, \\ U_j & \text{if } t < 0, \end{cases}$$

where  $U_j$  is a sequence in  $\mathcal{C}_c^\infty$  such that  $U_j \rightarrow u^0$  a.e. in  $\Omega$ . We thus get

$$\begin{aligned}
& - \int_0^T \int_{\Omega} \partial_t (\xi(S'(u))_{-h}) (B_n(u) - B_n(u^0)) dx dt \\
& + \int_0^T \int_{\Omega} \lambda(u) h_n(u) \nabla u \cdot \nabla (\xi(S'(u))_{-h}) dx dt \\
& = \int_0^T \int_{\Omega} f h_n(u) \xi(S'(u))_{-h} dx dt + T(C(\xi, S), n). \tag{3.52}
\end{aligned}$$

Using Lemma 2.3 of [20], we obtain

$$\begin{aligned}
& - \int_0^T \int_{\Omega} \partial_t (\xi(S'(u))_{-h}) (B_n(u) - B_n(u^0)) dx dt \\
& \leq - \int_0^T \int_{\Omega} \xi_t \left( B_n(u) \frac{1}{h} \int_{t-h}^t S'(\hat{u}(\tau)) d\tau - \frac{1}{h} \int_{t-h}^t \int_0^{\hat{u}(\tau)} S''(r) B_n(r) dr d\tau \right) dx dt \\
& \quad - \int_{\Omega} \xi(x, 0) \left( B_n(u^0) S'(U_j) - \int_0^{U_j} S''(r) B_n(r) dr \right) dx.
\end{aligned}$$

We now let  $h$  tend to zero :

$$\begin{aligned}
& \limsup_{h \rightarrow 0} \left\{ - \int_0^T \int_{\Omega} \partial_t (\xi(S'(u))_{-h}) (B_n(u) - B_n(u^0)) dx dt \right\} \\
& \leq - \int_0^T \int_{\Omega} \xi_t \left( B_n(u) S'(u) - \int_0^u S''(r) B_n(r) dr \right) dx dt \\
& \quad - \int_{\Omega} \xi(x, 0) \left( B_n(u^0) S'(U_j) - \int_0^{U_j} S''(r) B_n(r) dr \right) dx,
\end{aligned}$$

and letting  $j$  tends to infinity :

$$\begin{aligned}
& \limsup_{h \rightarrow 0} \left\{ - \int_0^T \int_{\Omega} \partial_t (\xi(S'(u))_{-h}) (B_n(u) - B_n(u^0)) dx dt \right\} \\
& \leq - \int_0^T \int_{\Omega} \xi_t \left( B_n(u) S'(u) - \int_0^u S''(r) B_n(r) dr \right) dx dt \\
& \quad - \int_{\Omega} \xi(x, 0) \left( B_n(u^0) S'(u^0) - \int_0^{u^0} S''(r) B_n(r) dr \right) dx,
\end{aligned}$$

Injecting this last inequality in (3.52) and letting  $h$  tends to zero we get

$$\begin{aligned}
& \int_0^T \int_{\Omega} f h_n(u) \xi S'(u) dx dt + T(C(\xi, S), n) \\
& \leq - \int_0^T \int_{\Omega} \xi_t \left( B_n(u) S'(u) - \int_0^u S''(r) B_n(r) dr \right) dx dt \\
& \quad - \int_{\Omega} \xi(x, 0) \left( B_n(u^0) S'(u^0) - \int_0^{u^0} S''(r) B_n(r) dr \right) dx \\
& \quad + \int_0^T \int_{\Omega} \lambda(u) h_n(u) \nabla u \cdot \nabla (\xi S'(u)) dx dt,
\end{aligned}$$

and letting  $n \rightarrow +\infty$

$$\begin{aligned}
& \int_0^T \int_{\Omega} f \xi S'(u) dx dt \\
& \leq - \int_0^T \int_{\Omega} \xi_t \left( b(u) S'(u) - \int_0^u S''(r) b(r) dr \right) dx dt \\
& \quad - \int_{\Omega} \xi(x, 0) \left( b(u^0) S'(u^0) - \int_0^{u^0} S''(r) b(r) dr \right) dx \\
& \quad + \int_0^T \int_{\Omega} \lambda(u) \nabla u \cdot \nabla (\xi S'(u)) dx dt.
\end{aligned}$$

We finally obtain the inequality

$$\begin{aligned}
& \int_0^T \int_{\Omega} f \xi S'(u) dx dr \leq \\
& - \int_0^T \int_{\Omega} \xi_t \int_0^u S'(r) b'(r) dr dx dt - \int_{\Omega} \xi(x, 0) \int_0^{u^0} S'(r) b'(r) dr dx \\
& + \int_0^T \int_{\Omega} \lambda(u) \nabla u \cdot \nabla (\xi S'(u)) dx dt. \tag{3.53}
\end{aligned}$$

Gathering (3.51) and (3.53) we get the desired renormalized formulation

$$\begin{aligned}
& - \int_0^T \int_{\Omega} \xi_t \int_0^u S'(r) b'(r) dr dx dt - \int_{\Omega} \xi(x, 0) \int_0^{u^0} S'(r) b'(r) dr dx \\
& + \int_0^T \int_{\Omega} \lambda(u) \nabla u \cdot \nabla (\xi S'(u)) dx dt = \int_0^T \int_{\Omega} f \xi S'(u) dx dt.
\end{aligned}$$

#### Step 4

It remains to prove that  $u$  satisfies the decay (3.4) of the truncate energy. Thanks to the discrete estimate on the energy (3.30) we have

$$\lim_{n \rightarrow \infty} \overline{\lim}_{\substack{h_T \rightarrow 0 \\ \delta t \rightarrow 0}} \frac{1}{n} \sum_{i=1}^N \delta t \sum_{\sigma \in \mathcal{E}} \frac{|\sigma|}{d_{\sigma}} \lambda(u)_{\sigma}^i (T_{2n}(u_K^i) - T_{2n}(u_L^i))^2 = 0,$$

and

$$\begin{aligned}
& \sum_{i=1}^N \delta t \sum_{\sigma \in \mathcal{E}} \frac{|\sigma|}{d_{\sigma}} \lambda(u)_{\sigma}^i (T_{2n}(u_K^i) - T_{2n}(u_L^i))^2 \\
& = \sum_{i=1}^N \delta t \sum_{\sigma \in \mathcal{E}} |\sigma| d_{\sigma} \lambda(u)_{\sigma}^i \left( \frac{T_{2n}(u_K^i) - T_{2n}(u_L^i)}{d_{\sigma}} \right)^2 \\
& = \sum_{i=1}^N \delta t \sum_{\sigma \in \mathcal{E}} d |D_{\sigma}| \lambda(u)_{\sigma}^i \left( \frac{T_{2n}(u_K^i) - T_{2n}(u_L^i)}{d_{\sigma}} \right)^2 \\
& = \sum_{i=1}^N \delta t \frac{1}{d} \sum_{\sigma \in \mathcal{E}} |D_{\sigma}| \lambda(u)_{\sigma}^i \left( d \frac{T_{2n}(u_K^i) - T_{2n}(u_L^i)}{d_{\sigma}} \right)^2 \\
& \geq \frac{1}{d} \sum_{i=1}^N \delta t \sum_{\sigma \in \mathcal{E}} |D_{K,\sigma}| h_n(u_K^i) \frac{\lambda(u_K^i)}{2} \left( d \frac{T_{2n}(u_K^i) - T_{2n}(u_L^i)}{d_{\sigma}} \right)^2 \\
& + \frac{1}{d} \sum_{i=1}^N \delta t \sum_{\sigma \in \mathcal{E}} |D_{L,\sigma}| h_n(u_L^i) \frac{\lambda(u_L^i)}{2} \left( d \frac{T_{2n}(u_K^i) - T_{2n}(u_L^i)}{d_{\sigma}} \right)^2 \\
& = \frac{1}{d} \int_0^T \int_{\Omega} h_n(u_{\mathcal{D}}) \frac{\lambda(T_{2n}(u_{\mathcal{D}}))}{2} |\nabla_{\mathcal{D}} T_{2n}(u_{\mathcal{D}})|^2 dx dt,
\end{aligned}$$

thus,  $\lim_{n \rightarrow \infty} \overline{\lim}_{\substack{h_T \rightarrow 0 \\ \delta t \rightarrow 0}} \frac{1}{n} \int_0^T \int_{\Omega} h_n(u_{\mathcal{D}}) \frac{\lambda(u_{\mathcal{D}})}{2} |\nabla_{\mathcal{D}} T_{2n}(u_{\mathcal{D}})|^2 dx dt = 0$ . Since  $\nabla T_{2n}(u_{\mathcal{D}})$  converges weakly in  $L^2(Q)^d$ ,  $\lambda(T_{2n}(u_{\mathcal{D}}))$  converges to  $\lambda(u)$   $L^\infty$  weak- $\star$  and  $h_n(u_{\mathcal{D}})$  converge to  $h_n(u)$  almost everywhere in  $Q$  we also have

$$\frac{1}{n} \int_0^T \int_{\Omega} h_n(u) \lambda(u) |\nabla T_{2n}(u)|^2 dx \leq \liminf_{\substack{h_T \rightarrow 0 \\ \delta t \rightarrow 0}} \frac{1}{n} \int_0^T \int_{\Omega} h_n(u_{\mathcal{D}}) \lambda(u_{\mathcal{D}}) |\nabla_{\mathcal{D}} T_{2n}(u_{\mathcal{D}})|^2 dx dt,$$

which leads to

$$\lim_{n \rightarrow \infty} \frac{1}{n} \int_0^T \int_{\Omega} \lambda(u) |\nabla T_n(u)|^2 dx dt = 0.$$

Since the renormalized solution  $u$  is unique (see [20] for the proof of the uniqueness) we conclude that the whole sequence  $u_{\mathcal{D}_m}$  converges to  $u$  in the sense that for all  $n > 0$ ,  $T_n(u_{\mathcal{D}_m})$  converges weakly to  $T_n(u)$  in  $L^2(0, T; H_0^1(\Omega))$ .  $\square$





# Chapitre 4

## Convergence of a scheme for a coupled PDE system of thermoviscoelasticity type

### 4.1 Introduction

We consider in this chapter a finite element-finite volume discretization of the following nonlinear thermoviscoelasticity at small strains system  $\mathcal{S}$  :

$$\frac{\partial^2 u}{\partial t^2} - \operatorname{div} \left( \varepsilon(u) + \varepsilon \left( \frac{\partial u}{\partial t} \right) \right) = g - Df(\theta) \text{ in } Q, \quad (4.1)$$

$$\frac{\partial \theta}{\partial t} - \operatorname{div} (\lambda(\theta) \nabla \theta) = \varepsilon \left( \frac{\partial u}{\partial t} \right) \cdot \varepsilon \left( \frac{\partial u}{\partial t} \right) - f(\theta) \operatorname{tr} \left( \varepsilon \left( \frac{\partial u}{\partial t} \right) \right) \text{ in } Q, \quad (4.2)$$

$$u(t=0) = u_0, \quad \frac{\partial u}{\partial t}(t=0) = v_0, \quad \theta(t=0) = \theta_0 \text{ in } \Omega, \quad (4.3)$$

$$u = 0, \quad \theta = 0, \quad \text{on } \partial\Omega \times (0, T), \quad (4.4)$$

where  $\Omega$  is an open bounded polygonal subset of  $\mathbb{R}^d$ ,  $d = 2$  or  $d = 3$ ,  $T$  is a positive number and  $Q = \Omega \times (0, T)$ . Moreover  $f$  and  $\lambda$  are two bounded continuous functions ( $\lambda_\infty \geq \lambda \geq \mu > 0$  with  $\lambda_\infty$  and  $\mu$  positive real numbers) and  $g \in L^2(Q)^d$ . The assumptions on the initial data are the following :  $\theta_0 \in L^1(\Omega)$ ,  $u_0 \in (H_0^1(\Omega))^d$  and  $v_0 \in (L^2(\Omega))^d$ .

In the above equations  $u : \Omega \times (0, T) \rightarrow \mathbb{R}^d$  is the displacement field,  $\theta : \Omega \times (0, T) \rightarrow \mathbb{R}$  is the temperature and  $\varepsilon(\cdot)$  is the small-strain tensor, i.e.  $\varepsilon(u) = \frac{1}{2}(\nabla u + (\nabla u)^t)$  and  $\varepsilon \left( \frac{\partial u}{\partial t} \right) = \frac{1}{2} \left( \nabla \frac{\partial u}{\partial t} + (\nabla \frac{\partial u}{\partial t})^t \right)$ . Equation (4.1) is the equation of motion while equation (4.2) is the energy balance equation. The mechanical dissipation  $\varepsilon \left( \frac{\partial u}{\partial t} \right) \cdot \varepsilon \left( \frac{\partial u}{\partial t} \right) - f(\theta) \operatorname{tr} \left( \varepsilon \left( \frac{\partial u}{\partial t} \right) \right)$  is not linearized, thus the right-hand side of (4.2) belongs to  $L^1(Q)$ . We then choose to use the notion of renormalized solution to investigate the existence of a solution to equation (4.2).

A more general version of system  $\mathcal{S}$  is studied in [15, 62, 67] in the framework of a Kelvin-Voigt type model and in [29–31] in the framework of a Norton-Hoff type model. In [15, 29–31] the authors use the notion of renormalized solution to deal with equation (4.2) while in [67] the author use the notion of solution in the sense of distribution for a nonlinear version of (4.2) with Neumann boundary condition. In

fluid mechanics other coupled systems (reduced turbulence model, Boussinesq type) with an energy balance equation with  $L^1$  data have been investigated in [55, 57, 71]. The finite volume version of [55] is addressed in [48] (and in [51] for the stationary case). The authors use nonconforming finite elements to discretize the Navier-Stokes equation and the finite volume method to discretize the balance equation for the turbulent energy. They prove that the approximate solution of the Navier-Stokes equation converges to a weak solution and that the approximate solution of the energy balance equation converges to a solution in the sense of distribution.

In the present work our purpose is to show, using the results of [52], that the solution obtained by a finite element-finite volume discretization converges to a weak-renormalized solution of  $\mathbf{S}$ . As in [48] we use finite element to discretize (4.1) and finite volume to discretize (4.2). It is worth noting that the presence of a second order time derivative gives additional difficulties in the study of the scheme. To handle these difficulties we use the techniques developed in [42, 43] for the approximation of second order nonlinear evolution equations. The authors use a conforming finite element method in space and an explicit-implicit Euler method in time and show that, up to a subsequence, the approximate solution converges to the weak solution of the equation. We prove here that the approximate solution obtained by a finite element-finite volume scheme converges, up to a subsequence, to a weak-renormalized solution of  $\mathbf{S}$ . Recall (see [15]) that a weak-renormalized solution of  $\mathbf{S}$  is a couple of functions  $(u, \theta)$  defined on  $\Omega \times (0, T)$  such that :

- $u \in \mathcal{C}^0(0, T; (H_0^1(\Omega))^d)$ , (4.5)

- $v := \frac{\partial u}{\partial t} \in L^2(0, T; (H_0^1(\Omega))^d) \cap \mathcal{C}^0(0, T; (L^2(\Omega))^d)$ ; (4.6)

- $\theta \in \mathcal{C}^0([0, T], L^1(\Omega))$ ,  $T_n(\theta) \in L^2(0, T; H_0^1(\Omega)) \forall n \geq 0$ , (4.7)

- $f(\theta) \in L^2(Q)$ ; (4.8)

and if  $(u, \theta)$  satisfies :

- $\frac{\partial v}{\partial t} - \operatorname{div}[\varepsilon(u) + \varepsilon(v)] = g - Df(\theta)$  in  $(\mathcal{D}'(Q))^d$ ; (4.9)

for all function  $S \in \mathcal{C}^\infty(\mathbb{R})$  such that  $S'$  has a compact support in  $\mathbb{R}$ ,

- $\frac{\partial S(\theta)}{\partial t} - \operatorname{div}(S'(\theta)\lambda(\theta)\nabla\theta) + S''(\theta)\lambda(\theta)\nabla\theta \cdot \nabla\theta =$  (4.10)

$$S'(\theta) \times \{\varepsilon(v)\varepsilon(v) - f(\theta)\operatorname{tr}(\varepsilon(v))\} \text{ in } \mathcal{D}'(Q);$$

- $\theta(t=0) = \theta_0$  in  $\Omega$ ; (4.11)

- $\lim_{n \rightarrow +\infty} \frac{1}{n} \int_0^T \int_\Omega \lambda(\theta) |\nabla T_n(\theta)|^2 dx dt = 0$ ; (4.12)

- $u - u_0 = \mathcal{K}v$  with  $v(t) = \int_0^t v(s) ds$  and  $v(t=0) = v_0$  in  $\Omega$ . (4.13)

Here  $T_n$  is the truncate function at height  $n$  :

$$T_n(s) = \begin{cases} s & \text{if } |s| \leq n, \\ \operatorname{sgn}(s)n & \text{if } |s| > n. \end{cases}$$

**Remark 4.1.1** Thanks to (4.6), (4.7) and (4.8) the term  $\varepsilon(v)\varepsilon(v) - f(\theta)\operatorname{tr}(\varepsilon(v))$

belongs to  $L^1(Q)$ . Thus, equations (4.10), (4.11) and (4.12) represent the definition of a renormalized solution of (4.2) (see [20, 21]).

The chapter is organized as follows. In Section 2, we present the finite element-finite volume scheme. Section 3 is devoted to prove several estimates adapting the techniques developed in [42, 43] and in [52]. In Section 4, we prove the convergence of the scheme.

## 4.2 Discretization

Let us define the admissibility of the mesh in the present work.

**Definition 4.2.1 (Admissible mesh)** *Let  $\Omega$  be an open bounded polygonal subset of  $\mathbb{R}^d$ . An admissible finite volume mesh of  $\Omega$  is given by a finite partition  $\mathcal{T}$  of  $\Omega$  in polygonal convex sets, called the "control volumes", by a finite family  $\mathcal{E}$  of disjoint subsets of  $\bar{\Omega}$  contained in affine hyperplanes, called the "edges", and by a family  $\mathcal{P} = (x_K)_{K \in \mathcal{T}}$  of points in  $\Omega$  such that*

- each  $\sigma \in \mathcal{E}$  is a non-empty open subset of  $\partial K$  for some  $K \in \mathcal{T}$ ,
- by denoting  $\mathcal{E}(K) = \{\sigma \in \mathcal{E}; \sigma \in \partial K\}$ , one has  $\partial K = \cup_{\sigma \in \mathcal{E}(K)} \bar{\sigma}$  for all  $K \in \mathcal{T}$ ,
- for all  $K \neq L$  in  $\mathcal{T}$ , either the  $(d-1)$ -dimensional measure of  $\bar{K} \cap \bar{L}$  is null, or  $\bar{K} \cap \bar{L} = \bar{\sigma}$  for some  $\sigma \in \mathcal{E}$ , that we denote then  $\sigma = K|L$ ,
- for all  $K \in \mathcal{T}$ ,  $x_K$  is in the interior of  $K$ ,
- for all  $\sigma = K|L \in \mathcal{E}$ , the line  $(x_K, x_L)$  intersects and is orthogonal to  $\sigma$ ,
- for all  $\sigma \in \mathcal{E}$ ,  $\sigma \subset \partial\Omega \cap \partial K$ , the line which is orthogonal to  $\sigma$  and going through  $x_K$  intersects  $\sigma$ .

We denote by  $|K|$  (resp.  $|\sigma|$ ) the Lebesgue measure of  $K \in \mathcal{T}$  (resp.  $\sigma \in \mathcal{E}$ ). The unit normal to  $\sigma \in \mathcal{E}(K)$  outward to  $K$  is denoted by  $\eta_{K,\sigma}$ .  $\mathcal{E}_{int}$  (resp.  $\mathcal{E}_{ext}$ ) is defined as the set of interior (resp. boundary) edges.

For any  $K \in \mathcal{T}$  and  $\sigma \in \mathcal{E}(K)$  we denote by  $d_{K,\sigma}$  the Euclidean distance between  $x_K$  and  $\sigma$ . For any  $\sigma \in \mathcal{E}$ , we define  $d_\sigma = d_{K,\sigma} + d_{L,\sigma}$  if  $\sigma = K|L \in \mathcal{E}_{int}$  (in which case  $d_\sigma$  is the Euclidean distance between  $x_K$  and  $x_L$ ) and  $d_\sigma = d_{K,\sigma}$  if  $\sigma \in \mathcal{E}_{ext} \cap \mathcal{E}(K)$ . The size of the mesh, denoted by  $h_{\mathcal{T}}$ , is defined by  $h_{\mathcal{T}} = \sup_{K \in \mathcal{T}} \text{diam}(K)$ .

In the continuous case the usual tools to solve the problem are Poincaré and Sobolev inequalities. In the discrete case we will need such estimates, so we have to establish their discrete versions (see [32]). Let us define the discrete  $W_0^{1,q}$  norm.

**Definition 4.2.2 (Discrete  $W_0^{1,q}$  norm)** *Let  $\Omega$  be an open bounded polygonal subset of  $\mathbb{R}^d$ ,  $d \geq 2$ , and let  $\mathcal{T}$  be an admissible mesh. Define  $X(\mathcal{T})$  as the set of functions from  $\Omega$  to  $\mathbb{R}$  which are constant over each control volume of the mesh. For  $v_{\mathcal{T}} \in X(\mathcal{T})$  and  $q \in [1, +\infty[$ , we define the discrete  $W_0^{1,q}$  norm by*

$$\|v_{\mathcal{T}}\|_{1,q,\mathcal{T}}^q = \sum_{\substack{\sigma \in \mathcal{E}_{int} \\ \sigma = K|L}} |\sigma| d_\sigma \left| \frac{v_K - v_L}{d_\sigma} \right|^q + \sum_{\substack{\sigma \in \mathcal{E}_{ext} \\ \sigma \in \mathcal{E}(K)}} |\sigma| d_\sigma \left| \frac{v_K}{d_\sigma} \right|^q$$

where  $v_K$  denotes the value taken by  $v$  on the control volume  $K$ .

Let us also define

$$\begin{aligned} \|v_{\mathcal{T}}\|_{1,\infty,\mathcal{T}} &= \max \left\{ \left\{ \frac{|v_K - v_L|}{d_\sigma}, \sigma \in \mathcal{E}_{int}, \sigma = K|L \right\} \right. \\ &\quad \left. \cup \left\{ \frac{|v_K|}{d_\sigma}, \sigma \in \mathcal{E}_{ext}, \sigma \in \mathcal{E}(K) \right\} \cup \{|v_K|, K \in \mathcal{T}\} \right\}. \end{aligned}$$

**Definition 4.2.3 (Discrete dual norm)** Let  $q > 1$  and let  $q'$  be given by  $\frac{1}{q} + \frac{1}{q'} = 1$  if  $q$  is finite and  $q' = 1$  if  $q = +\infty$ . We define the dual norme by

$$\|v\|_{-1,q',\mathcal{T}} = \sup_{\substack{w \in X(\mathcal{T}) \\ w \neq 0}} \frac{1}{\|w\|_{1,q,\mathcal{T}}} \int_{\Omega} v w \, dx.$$

**Proposition 4.2.4 (Discrete Poincaré inequality)** Let  $\mathcal{T}$  be an admissible mesh and  $v_{\mathcal{T}} \in X(\mathcal{T})$ . Then, if  $1 \leq q \leq 2$ ,

$$\|v_{\mathcal{T}}\|_{L^q(\Omega)} \leq \text{diam}(\Omega) \|v_{\mathcal{T}}\|_{1,q,\mathcal{T}}.$$

**Proposition 4.2.5 (Discrete Sobolev inequality)** Let  $1 \leq q \leq 2$ ,  $\mathcal{T}$  be an admissible mesh and  $\xi > 0$  satisfying

$$\text{for all } K \in \mathcal{T} \text{ and all } \sigma \in \mathcal{E}(K), \quad d_{K,\sigma} \geq \xi d_\sigma.$$

Then, with  $q^* = \frac{dq}{d-q}$  if  $q < d$  and  $q^* < \infty$  if  $q = d = 2$ , there exists  $C > 0$  only depending on  $(\Omega, q, q^*, \xi)$  such that, for all  $v_{\mathcal{T}} \in X(\mathcal{T})$ ,

$$\|v_{\mathcal{T}}\|_{L^{q^*}(\Omega)} \leq C \|v_{\mathcal{T}}\|_{1,q,\mathcal{T}}.$$

As far as the finite element discretization is concerned we define the discrete associated finite element space for each component of the displacement field by

$$V_h = \{w_h \in L^2(\Omega) \cap \mathcal{C}(\bar{\Omega}); w_h|_K \in P_1\},$$

and  $\mathbf{V}_h = (V_h)^d$  is the approximation space of the displacement field.

We now turn to the time discretization. We suppose given a uniform partition of the interval  $[0, T)$  such that  $[0, T] = \cup_{1 \leq i \leq N} [t_{i-1}, t_i]$  where  $t_i = i\delta t$  and  $\delta t = \frac{T}{N}$ . We also define  $t_{i-1/2} = t_i - \delta t/2$  and  $t_{i+1/2} = t_i + \delta t/2$  for all  $2 \leq i \leq N-1$ . We define  $X_{\mathcal{D}}$  the space of piecewise constant functions over each  $K \times (t_{i-1}, t_i)$  for  $K \in \mathcal{T}$  and  $1 \leq i \leq N$ . To each sequence  $(v_{\mathcal{T}}^i)_{1 \leq i \leq N}$  of functions of  $X(\mathcal{T})$ , we associate the function  $v_{\mathcal{D}} \in X_{\mathcal{D}}$  defined by :

$$v_{\mathcal{D}}(x, t) = v_{\mathcal{T}}^i(x) \quad \forall x \in \Omega, \quad \forall t \in (t_{i-1}, t_i), \quad 1 \leq i \leq N,$$

$$\text{i.e. } v_{\mathcal{D}}(x, t) = v_K^i \quad \forall x \in K, \quad K \in \mathcal{T}, \quad \forall t \in (t_{i-1}, t_i), \quad 1 \leq i \leq N.$$

We now define the discrete time derivative of  $v_{\mathcal{D}}$  by

$$\partial_{t,\mathcal{D}} v^i(x) = \frac{v^i(x) - v^{i-1}(x)}{\delta t} \quad \forall 1 \leq i \leq N,$$

$$\text{i.e. } \partial_t v_K^i = \frac{v_K^i - v_K^{i-1}}{\delta t} \quad \forall K \in \mathcal{T}, \quad \forall 1 \leq i \leq N.$$

Before writing the scheme, let us define a discrete finite volume gradient (see [51]).

**Definition 4.2.6 (Discrete finite volume gradient)** For  $K \in \mathcal{T}$  and  $\sigma \in \mathcal{E}(K)$ , we define the volume  $D_{K,\sigma}$  as the cone of basis  $\sigma$  and of opposite vertex  $x_K$ . Then, we define the "diamond-cell"  $D_\sigma$  (see Figure 4.1) by

$$\begin{aligned} D_\sigma &= D_{K,\sigma} \cup D_{L,\sigma} && \text{if } \sigma = K|L \in \mathcal{E}_{int}, \\ D_\sigma &= D_{K,\sigma} && \text{if } \sigma \in \mathcal{E}_{ext} \cap \mathcal{E}(K). \end{aligned}$$

For any  $v_{\mathcal{D}} \in X_{\mathcal{D}}$ , we define the discrete gradient  $\nabla_{\mathcal{D}} v_{\mathcal{D}}$  as the piecewise constant function over each  $D_\sigma \times (t^{i-1}, t^i)$  and given by

$$\begin{aligned} \forall \sigma \in \mathcal{E}_{int}, \sigma = K|L, \forall 1 \leq i \leq N, \\ \nabla_{\mathcal{D}} v_{\mathcal{D}}(x, t) &= d \frac{v_L^i - v_K^i}{d_\sigma} \eta_{K,\sigma}, \quad \forall x \in D_\sigma, \quad \forall t \in (t^{i-1}, t^i], \\ \forall \sigma \in \mathcal{E}_{ext} \cap \mathcal{E}(K), \forall 1 \leq i \leq N, \\ \nabla_{\mathcal{T}} v_{\mathcal{D}}(x, t) &= d \frac{0 - v_K^i}{d_\sigma} \eta_{K,\sigma}, \quad \forall x \in D_\sigma, \quad \forall t \in (t^{i-1}, t^i]. \end{aligned}$$

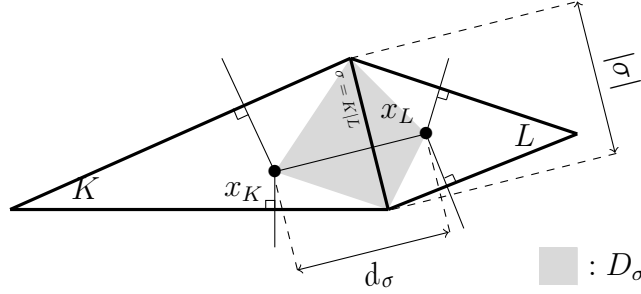


FIGURE 4.1 – The diamond  $D_\sigma$

**Lemma 4.2.7 (Weak convergence of the finite volume gradient)** Let  $(\mathcal{T}_m)_{m \geq 1}$  be a sequence of admissible meshes such that there exists  $\xi > 0$  satisfying

$$\text{for all } m \geq 1 \text{ for all } K \in \mathcal{T} \text{ and all } \sigma \in \mathcal{E}(K), \quad d_{K,\sigma} \geq \xi d_\sigma,$$

and such that  $h_{\mathcal{T}_m} \rightarrow 0$  and let  $\delta t_m$  be a sequence of time steps such that  $\delta t_m \rightarrow 0$ . Let  $v_{\mathcal{D}_m} \in X_{\mathcal{D}_m}$  and let us assume that there exists  $C > 0$  such that  $\sum_{i=1}^N \delta t_m \|v_{\mathcal{T}_m}^i\|_{1,2,\mathcal{T}_m}^2 \leq C$ , and that  $v_{\mathcal{D}_m}$  converges weakly in  $L^1(Q)$  to  $v \in L^1(0, T; H_0^1(\Omega))$ . Then  $\nabla_{\mathcal{D}_m} v_{\mathcal{D}_m}$  converges to  $\nabla v$  weakly in  $L^2(Q)^d$ .

**Proof** This proof is done in Lemma 5.2 of [5], the idea is to adapt the proof done in the elliptic case in [51]. Let  $\varphi \in \mathcal{C}_c^\infty(Q)^d$ . Since the size of the mesh tends to zero we can suppose that the intersection of the support of  $\varphi$  and the diamond cells associated to the external faces of the meshes is the empty set. We have

$$\begin{aligned} \int_0^T \int_\Omega \nabla_{\mathcal{D}_m} v \cdot \varphi dx dt &= \sum_{i=1}^N \sum_{\sigma=K|L \in \mathcal{E}_{int}} d \frac{v_L^i - v_K^i}{d_\sigma} \int_{t_{i-1}}^{t_i} \int_{D_\sigma} \varphi dx \cdot \eta_{K,\sigma} \\ &= T_1 + T_2, \end{aligned}$$

where

$$T_1 = \sum_{i=1}^N \sum_{\sigma=K|L \in \mathcal{E}_{int}} (v_L^i - v_L^i) \int_{t_{i-1}}^{t_i} \int_{\sigma} \varphi \cdot \eta d\gamma,$$

$$T_2 = \sum_{i=1}^N \sum_{\sigma=K|L \in \mathcal{E}_{int}} |\sigma| (v_L^i - v_L^i) \left[ \frac{1}{|D_{\sigma}|} \int_{t_{i-1}}^{t_i} \int_{D_{\sigma}} \varphi dx - \frac{1}{|\sigma|} \int_{t_{i-1}}^{t_i} \int_{\sigma} \varphi d\gamma \right] \cdot \eta_{K,\sigma}.$$

Reordering the sum, we obtain

$$\begin{aligned} T_1 &= - \sum_{i=1}^N \sum_{K \in \mathcal{T}_m} v_K^i \sum_{\sigma=K|L} \int_{t_{i-1}}^{t_i} \int_{\sigma} \varphi \cdot \eta_{K,\sigma} d\gamma \\ &= - \sum_{i=1}^N \sum_{K \in \mathcal{T}_m} \int_{t_{i-1}}^{t_i} \int_K v_{\mathcal{D}_m} \operatorname{div} \varphi dx \\ &= - \int_0^T \int_{\Omega} v_{\mathcal{D}_m} \operatorname{div} \varphi dx dt. \end{aligned}$$

Since  $v_{\mathcal{D}_m}$  converges weakly in  $L^1(Q)$  we get

$$\lim_{\substack{h_{\mathcal{T}_m} \rightarrow 0 \\ \delta t_m \rightarrow 0}} T_1 = - \int_0^T \int_{\Omega} v \operatorname{div} \varphi dx dt.$$

As far as  $T_2$  is concerned, thanks to the regularity of  $\varphi$ , we have

$$|T_2| \leq C_{\varphi} h_{\mathcal{T}_m} \sum_{i=1}^N \delta t_m \|v_{\mathcal{T}_m}^i\|_{1,1,\mathcal{T}},$$

and by Hölder's inequality and the fact that  $\sum_{i=1}^N \delta t_m \|v_{\mathcal{T}_m}^i\|_{1,2,\mathcal{T}_m}^2 \leq C$  we obtain that

$$\lim_{\substack{h_{\mathcal{T}_m} \rightarrow 0 \\ \delta t_m \rightarrow 0}} T_2 = 0.$$

Since  $v \in L^1(0, T; H_0^1(\Omega))$  it follows that

$$\forall \varphi \in (C_c^{\infty}(Q))^d, \quad \lim_{\substack{h_{\mathcal{T}_m} \rightarrow 0 \\ \delta t_m \rightarrow 0}} \int_0^T \int_{\Omega} \nabla_{\mathcal{D}_m} v_{\mathcal{D}_m} \cdot \varphi dx dt = \int_0^T \int_{\Omega} \nabla v \cdot \varphi dx dt.$$

Recalling that  $\nabla_{\mathcal{D}_m} v_{\mathcal{D}_m}$  is bounded in  $L^2(Q)^d$ , by density of  $C_c^{\infty}(Q)^d$  in  $L^2(Q)^d$  the proof is complete.  $\square$

We are now in position to write a scheme for the thermoviscoelasticity system  $\mathbf{S}$ . We combine both finite element and finite volume methods. This technique is used for a turbulence model in [51] in the elliptic case and in [48] in the parabolic case. In these two papers the authors discretized Stokes equations with a nonconforming finite element method and the convection-diffusion equation with a finite volume scheme. As in [52], a discrete version of the Aubin-Lions Lemma is used in [48] in order to extract convergent subsequences. In our work equation (4.1) is approximated

by a  $P_1$  finite element technique while equation (4.2) is discretized by a finite volume method.

Let  $\mathcal{T}$  be an admissible mesh and  $\delta t$  a time step. We will write the discretization of (4.1)-(4.2). We first write the finite element scheme of (4.1). We look for  $\{u_h^i\}_{i=0}^N \subset \mathbf{V}_h$  such that  $\forall \varphi \in \mathbf{V}_h$  and  $\forall 1 \leq i \leq N$

$$\begin{aligned} \int_{\Omega} \frac{u_h^{i+1} - 2u_h^i + u_h^{i-1}}{\delta t^2} \cdot \varphi \, dx + \int_{\Omega} \left[ \varepsilon(u_h^i) + \varepsilon\left(\frac{u_h^{i+1} - u_h^i}{\delta t}\right) \right] \cdot \nabla \varphi \, dx \\ = \int_{\Omega} g_h^i \cdot \varphi \, dx + \int_{\Omega} f(\theta_{\mathcal{D}_h}^i) \cdot \nabla \varphi \, dx, \end{aligned}$$

where  $u_h^0 \approx u_0$ ,  $v_h^0 = \frac{u_h^1 - u_h^0}{\delta t} \approx v_0$  and  $g_h^i = \frac{1}{\delta t} \int_{t_{i-1/2}}^{t_{i+1/2}} g \, dt$ . In order to avoid any problem when rebuilding the integral we made the following choice for  $\theta_{\mathcal{D}_h}^i$  :

$$\theta_{\mathcal{D}_h}^i = \frac{1}{\delta t} \int_{t_{i-1/2}}^{t_{i+1/2}} \theta_{\mathcal{D}} \, dt. \quad (4.14)$$

Using the fact that equation (4.2.5) can be rewritten with  $v = \frac{\partial u}{\partial t}$ , we get the following scheme  $\forall \varphi \in \mathbf{V}_h$

$$\begin{aligned} \int_{\Omega} \frac{u_h^{i+1} - u_h^i}{\delta t} \cdot \varphi \, dx - \int_{\Omega} v^i \cdot \varphi \, dx = 0 \quad i = 0, \dots, N-1 \\ \int_{\Omega} \frac{v_h^i - v_h^{i-1}}{\delta t} \cdot \varphi \, dx + \int_{\Omega} [\varepsilon(u_h^i) + \varepsilon(v_h^i)] \cdot \nabla \varphi \, dx \\ = \int_{\Omega} g_h^i \cdot \varphi \, dx + \int_{\Omega} f(\theta_{\mathcal{D}_h}^i) \cdot \nabla \varphi \, dx \quad i = 1, \dots, N-1 \end{aligned}$$

Then, with  $u(t) = u_0 + \mathcal{K}v(t)$  where  $\mathcal{K}v(t) = \int_0^t v(s) \, ds$ , we obtain  $\forall \varphi \in \mathbf{V}_h$

$$\begin{aligned} \int_{\Omega} \frac{v_h^i - v_h^{i-1}}{\delta t} \cdot \varphi \, dx + \int_{\Omega} [\varepsilon(u_h^0 + \mathcal{K}_I v_h^i) + \varepsilon(v_h^i)] \cdot \nabla \varphi \, dx \\ = \int_{\Omega} g_h^i \cdot \varphi \, dx + \int_{\Omega} f(\theta_{\mathcal{D}_h}^i) \cdot \nabla \varphi \, dx \quad i = 1, \dots, N-1, \end{aligned}$$

where  $\mathcal{K}_I v_h^i = \sum_{j=0}^{i-1} \delta t v_h^j$ .

As far as equation (4.2) is concerned, we first define for  $1 \leq i \leq N$ ,  $K \in \mathcal{T}$  and  $\sigma \in \mathcal{E}(K)$

$$\theta_K^0 = \frac{1}{|K|} \int_K \theta_0(x) \, dx, \quad (4.15)$$

and we can write the scheme as the following set of equations

For all  $1 \leq i \leq N$  and for all  $K \in \mathcal{T}$

$$|K| \frac{\theta_K^i - \theta_K^{i-1}}{\delta t} + \sum_{\sigma \in \mathcal{E}(K)} \frac{|\sigma|}{d_{\sigma}} \lambda(\theta)_{\sigma}^i (\theta_K^i - \theta_L^i) = \frac{1}{|K|} \int_K (\varepsilon(v_h^i) \cdot \varepsilon(v_h^i) - f(\theta_{\mathcal{D}_h}^i) \text{tr}(\varepsilon(v_h^i))) \, dx, \quad (4.16)$$



with  $\theta_L^i = 0$  if  $\sigma \in \mathcal{E}_{ext} \cap \mathcal{E}(K)$ ,  $v_h^N = 0$  and  $\theta_{\mathcal{D}_h}^N = 0$ .

For all  $1 \leq i \leq N$  we define  $\lambda(\theta)_\sigma^i$  in the following way

$$\forall \sigma = K|L \in \mathcal{E}_{int}, \min[\lambda(\theta_K^i), \lambda(\theta_L^i)] \leq \lambda(\theta)_\sigma^i \leq \max[\lambda(\theta_K^i), \lambda(\theta_L^i)], \quad (4.17)$$

$$\forall \sigma \in \mathcal{E}_{ext} \cap \mathcal{E}(K), \lambda(\theta)_\sigma^i = \lambda(\theta_K^i), \quad (4.18)$$

where  $\lambda(\theta)_\sigma^i$  is for example the mean value of  $\lambda(\theta_K^i)$  and  $\lambda(\theta_L^i)$ .

Finally the complete scheme writes  $\forall \varphi \in \mathbf{V}_h, \forall K \in \mathcal{T}$

$$\begin{aligned} \int_{\Omega} \frac{v_h^i - v_h^{i-1}}{\delta t} \cdot \varphi \, dx + \int_{\Omega} [\varepsilon(u_h^0 + K_I v_h^i) + \varepsilon(v_h^i)] \cdot \nabla \varphi \, dx \\ = \int_{\Omega} g_h^i \cdot \varphi \, dx + \int_{\Omega} f(\theta_{\mathcal{D}_h}^i) \cdot \nabla \varphi \, dx \quad i = 1, \dots, N-1, \end{aligned} \quad (4.19)$$

$$\begin{aligned} |K| \frac{\theta_K^i - \theta_K^{i-1}}{\delta t} + \sum_{\sigma \in \mathcal{E}(K)} \frac{|\sigma|}{d_\sigma} \lambda(\theta)_\sigma^i (\theta_K^i - \theta_L^i) \\ = \frac{1}{|K|} \int_K (\varepsilon(v_h^i) \cdot \varepsilon(v_h^i) - f(\theta_{\mathcal{D}_h}^i) \operatorname{tr}(\varepsilon(v_h^i))) \, dx \quad i = 1, \dots, N. \end{aligned} \quad (4.20)$$

The study of equation (4.19) is based on the work of Emmrich and al. in [42, 43] while equation (4.20) is studied with the techniques developed in [52].

### 4.3 Estimates

This section is devoted to derive estimates on the sequences  $u_h$  and  $\theta_{\mathcal{D}}$ . In Proposition 4.3.1, we prove several estimates on  $u_h$  and  $v_h$  which allow one to extract convergent subsequences (see [42, 43]). We then prove classical finite volume estimates using the techniques developed in [53] in the elliptic case and in [52] in the parabolic case.

**Proposition 4.3.1** *Let  $\mathcal{T}$  be an admissible mesh and let  $\delta t$  be a time step. If  $u_h = (u_h^i)_{1 \leq i \leq N}$  is a solution to (4.19), then for all  $1 \leq n \leq N-1$*

$$\begin{aligned} \|v_h^n\|_{L^2(\Omega)}^2 + \sum_{i=1}^n \|v_h^i - v_h^{i-1}\|_{L^2(\Omega)}^2 + C \sum_{i=1}^n \delta t \|v_h^i\|_{H_0^1(\Omega)}^2 + \|u_h^{n+1}\|_{H_0^1(\Omega)}^2 \\ \leq \|v_h^0\|_{L^2(\Omega)}^2 + \|u_h^0\|_{H_0^1(\Omega)}^2 + C(\|g\|_{L^2(Q)}, M, T, |\Omega|). \end{aligned} \quad (4.21)$$

**Proof** The proof is a direct application of the tools used in [43] to show Theorem 3. Moreover we use Korn's inequality to handle the term in  $\varepsilon(\cdot)$ .

Taking  $v_h^i$  as test function in (4.19) we get  $T_1^i + T_2^i = T_3^i + T_4^i$  with

$$\begin{aligned} T_1^i &= \int_{\Omega} \frac{v_h^i - v_h^{i-1}}{\delta t} v_h^i dx, \\ T_2^i &= \int_{\Omega} [\varepsilon(u_h^0 + \mathcal{K}_I v_h^i) + \varepsilon(v_h^i)] \cdot \nabla v_h^i dx, \\ T_3^i &= \int_{\Omega} g_h^i v_h^i dx \\ T_4^i &= \int_{\Omega} f(\theta_{\mathcal{D}_h}^i) \nabla v_h^i dx. \end{aligned}$$

Rewriting  $T_1^i$  we find

$$\begin{aligned} T_1^i &= \frac{1}{2\delta t} \int_{\Omega} (v_h^i - v_h^{i-1})(v_h^i + v_h^{i-1} + v_h^i - v_h^{i-1}) dx \\ &= \frac{1}{2\delta t} \int_{\Omega} (v_h^i - v_h^{i-1})(v_h^i + v_h^{i-1}) + (v_h^i - v_h^{i-1})^2 dx \\ &= \frac{1}{2\delta t} \int_{\Omega} (v_h^i)^2 - (v_h^{i-1})^2 + (v_h^i - v_h^{i-1})^2 dx. \end{aligned}$$

Thus, for all  $1 \leq n \leq N-1$

$$\sum_{i=1}^n \delta t T_1^i = \frac{1}{2} \int_{\Omega} (v_h^n)^2 dx - \frac{1}{2} \int_{\Omega} (v_h^0)^2 dx + \frac{1}{2} \sum_{i=1}^n \int_{\Omega} |v_h^i - v_h^{i-1}|^2 dx. \quad (4.22)$$

As far as  $T_2^i$  is concerned, we write  $T_2^i = T_{2,1}^i + T_{2,2}^i$  with

$$\begin{aligned} T_{2,1}^i &= \int_{\Omega} \varepsilon(u_h^0 + \mathcal{K}_I v_h^i) \cdot \nabla v_h^i dx, \\ T_{2,2}^i &= \int_{\Omega} \varepsilon(v_h^i) \cdot \nabla v_h^i dx. \end{aligned}$$

We easily have  $T_{2,2}^i = \int_{\Omega} \varepsilon(v_h^i) \cdot \varepsilon(v_h^i) dx$  and using the fact that

$$(a-b)b = \frac{1}{2}(a^2 - b^2 - (a-b)^2),$$

we get

$$\begin{aligned} T_{1,2}^i &= \frac{1}{\delta t} \int_{\Omega} \varepsilon(u_h^i) \cdot \nabla (u_h^{i+1} - u_h^i) dx \\ &= \frac{1}{\delta t} \int_{\Omega} \varepsilon(u_h^i) \cdot (\varepsilon(u_h^{i+1}) - \varepsilon(u_h^i)) dx \\ &= \frac{1}{2\delta t} \int_{\Omega} \varepsilon(u_h^{i+1})^2 - \varepsilon(u_h^i)^2 - (\varepsilon(u_h^{i+1}) - \varepsilon(u_h^i))^2 dx \\ &= \frac{1}{2\delta t} \int_{\Omega} \varepsilon(u_h^{i+1})^2 - \varepsilon(u_h^i)^2 - \delta t \varepsilon(v_h^i)^2 dx, \end{aligned}$$

and summing on the time steps we have

$$\sum_{i=1}^n \delta t T_2^i = \frac{1}{2} \sum_{i=1}^n \delta t \int_{\Omega} \varepsilon(v_h^i)^2 dx + \frac{1}{2} \int_{\Omega} \varepsilon(u_h^{n+1})^2 dx - \frac{1}{2} \int_{\Omega} \varepsilon(u_h^0)^2 dx. \quad (4.23)$$

Due to the regularity of  $g$  and  $f$  we get for the right-hand side terms

$$\begin{aligned} \sum_{i=1}^n \delta t T_3^i &= \sum_{i=1}^n \delta t \int_{\Omega} g_h^i v_h^i dx \\ &\leq \|g\|_{L^2(Q)} \sum_{i=1}^n \delta t \|v_h^i\|_{H_0^1(\Omega)}, \end{aligned} \quad (4.24)$$

and

$$\begin{aligned} \sum_{i=1}^n \delta t T_4^i &= \sum_{i=1}^n \delta t \int_{\Omega} f(\theta_{\mathcal{D}_h}^i) \nabla v_h^i dx \\ &\leq \sum_{i=1}^n \delta t \int_{\Omega} M |\nabla v_h^i| dx \\ &\leq M^2 T |\Omega| \sum_{i=1}^n \delta t \|v_h^i\|_{H_0^1(\Omega)}. \end{aligned} \quad (4.25)$$

Gathering (4.22)-(4.25) and using both Korn and Young inequalities yields the desired estimate.  $\square$

The following propositions are a set of estimates on  $\theta_{\mathcal{D}}$ , which are classical when using the finite volume method (see [40, 52]). If the right-hand side of (4.2) is bounded in  $L^1(Q)$ , all these results are already proved in [52]. This is why we only give a sketch of proof for all the propositions.

**Proposition 4.3.2** *Let  $\mathcal{T}$  be an admissible mesh and let  $\delta t$  be a time step. If  $\theta_{\mathcal{D}} = (\theta_{\mathcal{T}}^i)_{1 \leq i \leq N}$  is a solution to (4.20), then*

$$\sum_{i=1}^N \delta t \|\ln(1 + |\theta_{\mathcal{T}}^i|)\|_{1,2,\mathcal{T}}^2 \leq C, \quad (4.26)$$

and

$$\|\theta_{\mathcal{D}}\|_{L^\infty(0,T;L^1(\Omega))} \leq C. \quad (4.27)$$

**Proof** We start by checking that the right-hand side of (4.20) belongs to  $L^1(Q)$ . We have

$$\begin{aligned} &\sum_{i=1}^N \delta t \sum_{K \in \mathcal{T}} |K| \left[ \frac{1}{|K|} \int_K (\varepsilon(v_h^i) \varepsilon(v_h^i) - f(\theta_{\mathcal{D}_h}^i) \text{tr}(\varepsilon(v_h^i))) dx \right] \\ &= \sum_{i=1}^{N-1} \delta t \int_{\Omega} (\varepsilon(v_h^i) \varepsilon(v_h^i) - f(\theta_{\mathcal{D}_h}^i) \text{tr}(\varepsilon(v_h^i))) dx \\ &\leq \sum_{i=1}^{N-1} \delta t \|v_h^i\|_{H_0^1(\Omega)}^2 + C(M, |\Omega|, T) \sum_{i=1}^{N-1} \delta t \|v_h^i\|_{H_0^1(\Omega)}^2 \\ &\leq C \sum_{i=1}^{N-1} \delta t \|v_h^i\|_{H_0^1(\Omega)}^2, \end{aligned} \quad (4.28)$$

which is bounded thanks to (4.3.1). Thus we can use the same techniques as in [52], which are an easy adaptation of the proof of the log-estimate in [40] in the elliptic case.

Let  $\varphi(s) = \int_0^s \frac{dr}{(1+|r|)^2}$  and  $\tilde{\varphi}(s) = \int_0^s \varphi(r) dr$ . We multiply each equation of the scheme (4.20) by  $\varphi(\theta_K^i)$  and we sum over each control volume and each time step. Using the fact that the right-hand side of (4.20) is bounded in  $L^1(Q)$  and  $\varphi$  is bounded by 1, the convexity of  $\tilde{\varphi}$  and Young and Cauchy-Schwarz inequalities, we obtain the desired estimates.  $\square$

**Corollary 4.3.3** *Let  $\mathcal{T}$  be an admissible mesh and let  $\delta t$  be a time step. If  $\theta_{\mathcal{D}} = (\theta_{\mathcal{T}}^i)_{1 \leq i \leq N}$  is a solution to (4.20) and, for  $n > 0$ ,  $E_n = \{(x, t) \in \Omega \times (0, T); |\theta_{\mathcal{D}}(x, t)| > n\}$ , then there exists  $C > 0$  only depending on  $(\Omega, f, u, d)$  such that*

$$|E_n| \leq \frac{C}{(\ln(1+n))^2}. \quad (4.29)$$

Unlike what is performed in [52] we cannot directly show the discrete estimate on the energy. Indeed we have not yet the convergence of the right-hand side of (4.20). Thus we only have for now the following estimates.

**Proposition 4.3.4** *Let  $\mathcal{T}$  be an admissible mesh and let  $\delta t$  be a time step. If  $\theta_{\mathcal{D}} = (\theta_{\mathcal{T}}^i)_{1 \leq i \leq N}$  is a solution to (4.20), then for all  $n > 0$  there exists  $C > 0$  such that*

$$\sum_{i=1}^N \delta t \sum_{\sigma \in \mathcal{E}} \frac{|\sigma|}{d_{\sigma}} \lambda(\theta)_{\sigma}^i (\theta_K^i - \theta_L^i) (T_n(\theta_K^i) - T_n(\theta_L^i)) \leq C, \quad (4.30)$$

$$\sum_{i=1}^N \delta t \sum_{K \in \mathcal{T}} |K| \int_{\theta_K^{i-1}}^{\theta_K^i} (r - \theta_K^{i-1}) T_n'(r) dr \leq C. \quad (4.31)$$

**Proof** By multiplying each equation of the scheme by  $T_n(\theta_K^i)$ , by summing over each control volume and over each time step and by reordering the sums, we obtain  $T_1 + T_2 = T_3$  with

$$\begin{aligned} T_1 &= \sum_{i=1}^N \delta t \sum_{K \in \mathcal{T}} |K| \frac{\theta_K^i - \theta_K^{i-1}}{\delta t} T_n(\theta_K^i), \\ T_2 &= \sum_{i=1}^N \delta t \sum_{\sigma \in \mathcal{E}} \frac{|\sigma|}{d_{\sigma}} \lambda(\theta)_{\sigma}^i (\theta_K^i - \theta_L^i) (T_n(\theta_K^i) - T_n(\theta_L^i)), \\ T_3 &= \sum_{i=1}^N \delta t \sum_{K \in \mathcal{T}} \int_K (\varepsilon(v_h^i) \cdot \varepsilon(v_h^i) - f(\theta_{\mathcal{D}_n}^i) \text{tr}(\varepsilon(v_h^i))) T_n(\theta_K^i) dx. \end{aligned}$$

Since  $T_n$  is bounded by  $n$ , we deduce with (4.28) that  $|T_3| \leq nC \sum_{i=1}^{N-1} \delta t \|v_h^i\|_{H_0^1(\Omega)}^2$ . Let us define  $\tilde{T}_n(s) = \int_0^s T_n(r) dr$ . Since  $\tilde{T}_n$  is a convex function, we have  $\forall x, y \in \mathbb{R}$ ,

$$\tilde{T}_n(x) - \tilde{T}_n(y) = T_n(x)(x - y) - \int_y^x (r - y) T_n'(r) dr;$$

we thus obtain  $T_1 = T_{1,1} + T_{1,2} + T_{1,3}$  with

$$\begin{aligned} T_{1,1} &= - \sum_{K \in \mathcal{T}} |K| \tilde{T}_n(\theta_K^0), \\ T_{1,2} &= \sum_{K \in \mathcal{T}} |K| \tilde{T}_n(\theta_K^N), \\ T_{1,3} &= \sum_{K \in \mathcal{T}} |K| \int_{\theta_K^{i-1}}^{\theta_K^i} (r - \theta_K^{i-1}) T_n'(r) dr \geq 0. \end{aligned}$$

Since  $-T_{1,1} \leq n \|\theta^0\|_{L^1(\Omega)}$  and  $T_{1,2} \geq 0$ , we get

$$\begin{aligned} \sum_{i=1}^N \delta t \sum_{\sigma \in \mathcal{E}} \frac{|\sigma|}{d_\sigma} \lambda(\theta)_\sigma^i (\theta_K^i - \theta_L^i) (T_n(\theta_K^i) - T_n(\theta_L^i)) + \sum_{i=1}^N \delta t \sum_{K \in \mathcal{T}} |K| \int_{\theta_K^{i-1}}^{\theta_K^i} (r - \theta_K^{i-1}) T_n'(r) dr \\ \leq nC \sum_{i=1}^{N-1} \delta t \|v_h^i\|_{H_0^1(\Omega)}^2 + n \|\theta^0\|_{L^1(\Omega)}. \end{aligned}$$

Thanks to (4.21) we obtain (4.30) and (4.31).  $\square$

**Corollary 4.3.5** *Let  $\mathcal{T}$  be an admissible mesh and let  $\delta t$  be a time step. If  $\theta_{\mathcal{D}} = (\theta_{\mathcal{T}}^i)_{1 \leq i \leq N}$  is a solution to (4.20), then for all  $n > 0$  there exists  $C > 0$  such that*

$$\sum_{i=1}^N \delta t \|T_n(\theta_{\mathcal{T}}^i)\|_{1,2,\mathcal{T}} \leq C. \quad (4.32)$$

Let us introduce the function  $h_n$  defined by

$$h_n(r) = \begin{cases} 0 & \text{if } r \leq -2n, \\ \frac{r}{n} + 2 & \text{if } -2n \leq r \leq -n, \\ 1 & \text{if } -n \leq r \leq n, \\ \frac{-r}{n} + 2 & \text{if } n \leq r \leq 2n, \\ 0 & \text{if } r \geq 2n. \end{cases} \quad (4.33)$$

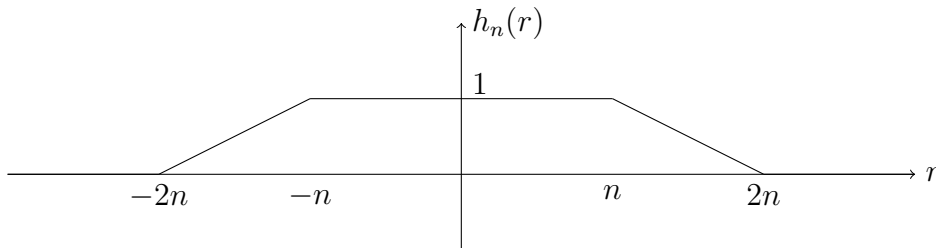


FIGURE 4.2 – The function  $h_n$

The two estimates of the following proposition are necessary to apply the discrete version of the Aubin-Lions lemma (see [47,69]) and extract a convergent subsequence from  $\theta_{\mathcal{D}}$ .

**Proposition 4.3.6** *Let  $\mathcal{T}$  be an admissible mesh and let  $\delta t$  be a time step. If  $\theta_{\mathcal{D}} = (\theta_{\mathcal{T}}^i)_{1 \leq i \leq N}$  is a solution to (4.20), then for all  $n > 0$  there exists  $C > 0$  such that*

$$\sum_{i=1}^N \delta t \|\tilde{h}_n(\theta_{\mathcal{T}}^i)\|_{1,2,\mathcal{T}}^2 \leq C, \quad (4.34)$$

$$\sum_{i=1}^N \delta t \|\partial_t \tilde{h}_n(\theta_{\mathcal{T}}^i)\|_{-1,1,\mathcal{T}} \leq C. \quad (4.35)$$

**Proof** Estimate (4.34) is a direct corollary of estimate (4.32). As in [52], let  $w \in X(\mathcal{T})$ , to prove (4.35) we multiply each equation of the scheme (4.20) by  $h_n(\theta_K^i)w_K$ , we sum over the control volumes and we gather by edges. We achieve the proof with estimates (4.30) and (4.31), the definition of the dual norm and the fact that the right-hand side of (4.20) is bounded in  $L^1(Q)$ .  $\square$

Before extracting convergent subsequences let us define the following functions.

$$u_h(t) := \begin{cases} 0 & \text{if } t \in [0, t_{1/2}], \\ u_h^i & \text{if } t \in (t_{i-1/2}, t_{i+1/2}], \quad i = 1, \dots, N-1, \\ 0 & \text{if } t \in (t_{N-1/2}, T], \end{cases}$$

$$v_h(t) := \begin{cases} 0 & \text{if } t \in [0, t_{1/2}], \\ v_h^i & \text{if } t \in (t_{i-1/2}, t_{i+1/2}], \quad i = 1, \dots, N-1, \\ 0 & \text{if } t \in (t_{N-1/2}, T], \end{cases}$$

$$\hat{v}_h(t) := \begin{cases} v_h^0 & \text{if } t \in [0, t_{1/2}], \\ v_h^i + \frac{t-t_{i-1/2}}{\delta t}(v_h^i - v_h^{i-1}) & \text{if } t \in (t_{i-1/2}, t_{i+1/2}], \quad i = 1, \dots, N-1, \\ v_h^{N-1} & \text{if } t \in (t_{N-1/2}, T], \end{cases}$$

$$g_h(t) := \begin{cases} 0 & \text{if } t \in [0, t_{1/2}], \\ g_h^i & \text{if } t \in (t_{i-1/2}, t_{i+1/2}], \quad i = 1, \dots, N-1, \\ 0 & \text{if } t \in (t_{N-1/2}, T], \end{cases}$$

$$\theta_{\mathcal{D}_h}(t) := \begin{cases} 0 & \text{if } t \in [0, t_{1/2}], \\ \frac{\theta_{\mathcal{T}}^i + \theta_{\mathcal{T}}^{i+1}}{2} & \text{if } t \in (t_{i-1/2}, t_{i+1/2}], \quad i = 1, \dots, N-1, \\ 0 & \text{if } t \in (t_{N-1/2}, T], \end{cases}$$

These functions are necessary to rebuild the integrals and pass to the limit.

**Assumption 4.3.7** *Let  $(\mathcal{T}_m)_{m \geq 1}$  be a sequence of admissible meshes such that  $h_{\mathcal{T}_m} \rightarrow 0$  as  $m \rightarrow +\infty$ , let  $\delta t_m$  be a sequence of time steps such that  $\delta t_m \rightarrow 0$  as  $m \rightarrow +\infty$  and let  $\mathbf{V}_{h_m}$  be a sequence of Galerkin schemes. The initial values for (4.19) and (4.20) satisfy*

$$u_{h_m}^0 \rightarrow u_0 \text{ in } H_0^1(\Omega), \quad v_{h_m}^0 \rightarrow v_0 \text{ in } L^2(\Omega) \text{ as } m \rightarrow +\infty, \quad (4.36)$$

$$\theta_{\mathcal{D}_m}^0 = \sum_{K \in \mathcal{T}_m} u_K^0 \mathbf{1}_K \rightarrow \theta_0 \text{ a.e. in } \Omega \text{ as } m \rightarrow +\infty. \quad (4.37)$$

We are now in position to extract convergent subsequences from  $u_{h_m}$  and  $T_n(\theta_{\mathcal{D}_m})$ .

**Proposition 4.3.8** *Let  $(\mathcal{T}_m)_{m \geq 1}$  be a sequence of admissible meshes, let  $(\delta t_m)_{m \geq 1}$  be a sequence of time steps and let  $\mathbf{V}_{h_m}$  be a sequence of Galerkin schemes. If  $u_{h_m}$  is a solution to (4.19) with  $\mathcal{T} = \mathcal{T}_m$  and  $\delta t = \delta t_m$ , then there exist  $u \in L^\infty(0, T; H_0^1(\Omega)) \cap C^0(0, T; H_0^1(\Omega))$ ,  $v \in L^\infty(0, T; L^2(\Omega)) \cap L^2(0, T; H_0^1(\Omega)) \cap C^0(0, T; L^2(\Omega))$  with  $u - u_0 = \mathcal{K}v \in C^0(0, T; H_0^1(\Omega)) \cap L^\infty(0, T; H_0^1(\Omega))$  and  $u' = v \in L^2(0, T; H_0^1(\Omega))$ , such that up to a subsequence*

$$\begin{aligned} u_{h_m} &\overset{*}{\rightharpoonup} u \text{ in } L^\infty(0, T; H_0^1(\Omega)), \\ v_{h_m} &\overset{*}{\rightharpoonup} v \text{ in } L^\infty(0, T; L^2(\Omega)) \text{ and } v_{h_m} \rightharpoonup v \text{ in } L^2(0, T; H_0^1(\Omega)), \\ \hat{v}_{h_m} &\overset{*}{\rightharpoonup} \hat{v} \text{ in } L^\infty(0, T; L^2(\Omega)) \text{ and } \hat{v}_{h_m} \rightharpoonup \hat{v} \text{ in } L^2(0, T; H_0^1(\Omega)), \\ \mathcal{K}v_{h_m} &\overset{*}{\rightharpoonup} \mathcal{K}v \text{ in } L^\infty(0, T; H_0^1(\Omega)) \text{ and } u_{h_m} - u_0 - \mathcal{K}v_{h_m} \rightarrow 0 \text{ in } L^2(0, T; H_0^1(\Omega)). \end{aligned}$$

**Proof** The proof is divided in two steps. In Step 1 we use estimate (4.21) to extract convergent subsequences, still indexed by  $m$  from  $u_{h_m}$ ,  $v_{h_m}$  and  $\hat{v}_{h_m}$ . We also show in this step that the limits of  $v_{h_m}$  and  $\hat{v}_{h_m}$  are the same. Step 2 is devoted to prove that  $\mathcal{K}v_{h_m} \overset{*}{\rightharpoonup} \mathcal{K}v$  in  $L^\infty(0, T; H_0^1(\Omega))$  and  $u_{h_m} - u_0 - \mathcal{K}v_{h_m} \rightarrow 0$  in  $L^2(0, T; H_0^1(\Omega))$ . This allows one to identify  $v$ , the limit of  $v_{h_m}$ , and  $u'$ , where  $u$  is the limit of  $u_{h_m}$ . Thanks to estimate (4.21) we have

$$\forall n = 2, \dots, N - 1, \quad \|u_h^{n+1}\|_{H_0^1(\Omega)} \leq C.$$

Moreover  $u_h^1 = u_h^0 + \delta t v_h^0$ , thus  $\|u_h^1\|_{H_0^1(\Omega)}$  is also bounded. We then obtain that

$$\|u_h\|_{L^\infty(0, T; H_0^1(\Omega))} \leq C. \quad (4.38)$$

Estimate (4.21) also gives

$$\forall n = 1, \dots, N - 1, \quad \|v_h^n\|_{L^2(\Omega)} \leq C,$$

thus

$$\|v_h\|_{L^\infty(0, T; L^2(\Omega))} \leq C. \quad (4.39)$$

Since  $\sum_{i=1}^n \delta t \|v_h^i\|_{H_0^1(\Omega)}^2 \leq C$  pour tout  $n = 1, \dots, N - 1$ , we get

$$\|v_h\|_{L^2(0, T; H_0^1(\Omega))} \leq C. \quad (4.40)$$

In view of the definition of  $\hat{v}_h$  and thanks to (4.39) we easily have

$$\|\hat{v}_h\|_{L^\infty(0, T; L^2(\Omega))} \leq C. \quad (4.41)$$

We then consider the  $L^2(H_0^1)$  norm of  $\hat{v}_h$  :

$$\begin{aligned}
\|\hat{v}_h\|_{L^2(0,T;H_0^1(\Omega))}^2 &= \int_0^T \|\hat{v}_h\|_{H_0^1(\Omega)}^2 dt \\
&= \int_0^{t_{1/2}} \|v_h^0\|_{H_0^1(\Omega)}^2 dt + \sum_{i=1}^{N-1} \int_{t_{i-1/2}}^{t_{i+1/2}} \|v_h^i + \frac{t-t_{i-1/2}}{\delta t}(v_h^i - v_h^{i-1})\|_{H_0^1(\Omega)}^2 dt \\
&\quad + \int_{t_{N-1/2}}^T \|v_h^{N-1}\|_{H_0^1(\Omega)}^2 dt \\
&= \frac{\delta t}{2} \|v_h^0\|_{H_0^1(\Omega)}^2 + \frac{\delta t}{2} \|v_h^{N-1}\|_{H_0^1(\Omega)}^2 \\
&\quad + \sum_{i=1}^{N-1} \int_{t_{i-1/2}}^{t_{i+1/2}} \|v_h^i + \frac{t-t_{i-1/2}}{\delta t}(v_h^i - v_h^{i-1})\|_{H_0^1(\Omega)}^2 dt \\
&\leq \frac{\delta t}{2} \|v_h^0\|_{H_0^1(\Omega)}^2 + \frac{\delta t}{2} \|v_h^{N-1}\|_{H_0^1(\Omega)}^2 \\
&\quad + C \left[ \sum_{i=1}^{N-1} \delta t \|v_h^i\|_{H_0^1(\Omega)}^2 + \sum_{i=1}^{N-1} \int_{t_{i-1/2}}^{t_{i+1/2}} \left\| \frac{t-t_{i-1/2}}{\delta t}(v_h^i - v_h^{i-1}) \right\|_{H_0^1(\Omega)}^2 dt \right] \\
&\leq \frac{\delta t}{2} \|v_h^0\|_{H_0^1(\Omega)}^2 + \frac{\delta t}{2} \|v_h^{N-1}\|_{H_0^1(\Omega)}^2 + C \sum_{i=1}^{N-1} \delta t \|v_h^i\|_{H_0^1(\Omega)}^2 \\
&\quad + \tilde{C} \sum_{i=1}^{N-1} \delta t \|v_h^i - v_h^{i-1}\|_{H_0^1(\Omega)}^2 \\
&\leq \frac{\delta t}{2} \|v_h^0\|_{H_0^1(\Omega)}^2 + \frac{\delta t}{2} \|v_h^{N-1}\|_{H_0^1(\Omega)}^2 + \bar{C} \sum_{i=1}^{N-1} \delta t \|v_h^i\|_{H_0^1(\Omega)}^2 \\
&\quad + \hat{C} \sum_{i=1}^{N-1} \delta t \|v_h^{i-1}\|_{H_0^1(\Omega)}^2 \\
&= \leq \frac{\delta t}{2} \|v_h^0\|_{H_0^1(\Omega)}^2 + \frac{\delta t}{2} \|v_h^{N-1}\|_{H_0^1(\Omega)}^2 + \bar{C} \sum_{i=1}^{N-1} \delta t \|v_h^i\|_{H_0^1(\Omega)}^2 \\
&\quad + \hat{C} \sum_{j=0}^{N-2} \delta t \|v_h^j\|_{H_0^1(\Omega)}^2 \\
&\leq C_1 \|v_h^0\|_{H_0^1(\Omega)}^2 + C_2 \|v_h^{N-1}\|_{H_0^1(\Omega)}^2 + C_3 \sum_{i=0}^{N-2} \delta t \|v_h^i\|_{H_0^1(\Omega)}^2.
\end{aligned}$$

Thanks to estimate (4.21) we conclude that

$$\|\hat{v}_h\|_{L^2(0,T;H_0^1(\Omega))} \leq C. \quad (4.42)$$

Gathering (4.38) to (4.42), we are now in position to extract convergent subsequences (still denoted by  $h_m$ ). There exist  $u \in L^\infty(0, T; H_0^1(\Omega))$ ,  $v \in L^\infty(0, T; L^2(\Omega)) \cap$



$L^2(0, T; H_0^1(\Omega))$  and  $\hat{v} \in L^\infty(0, T; L^2(\Omega)) \cap L^2(0, T; H_0^1(\Omega))$  such that

$$\begin{aligned} u_{h_m} &\overset{*}{\rightharpoonup} u \text{ in } L^\infty(0, T; H_0^1(\Omega)), \\ v_{h_m} &\overset{*}{\rightharpoonup} v \text{ in } L^\infty(0, T; L^2(\Omega)) \text{ and } v_{h_m} \rightharpoonup v \text{ in } L^2(0, T; H_0^1(\Omega)), \\ \hat{v}_{h_m} &\overset{*}{\rightharpoonup} \hat{v} \text{ in } L^\infty(0, T; L^2(\Omega)) \text{ and } \hat{v}_{h_m} \rightharpoonup \hat{v} \text{ in } L^2(0, T; H_0^1(\Omega)). \end{aligned}$$

We now prove that  $\hat{v} = v$  :

$$\begin{aligned} \|\hat{v}_h - v_h\|_{L^2(0, T; L^2(\Omega))}^2 &= \int_0^T \|\hat{v}_h(t) - v_h(t)\|_{L^2(\Omega)}^2 dt \\ &= \int_0^{t_{1/2}} \|v_h^0\|_{L^2(\Omega)}^2 dt + \sum_{i=1}^{N-1} \int_{t_{i-1/2}}^{t_{i+1/2}} \|v_h^i + \frac{t - t_{i-1/2}}{\delta t} (v_h^i - v_h^{i-1}) - v_h^i\|_{L^2(\Omega)}^2 dt \\ &\quad + \int_{t_{N-1/2}}^T \|v_h^{N-1}\|_{L^2(\Omega)}^2 dt \\ &= \int_0^{t_{1/2}} \|v_h^0\|_{L^2(\Omega)}^2 dt + \sum_{i=1}^{N-1} \int_{t_{i-1/2}}^{t_{i+1/2}} \left\| \frac{t - t_{i-1/2}}{\delta t} (v_h^i - v_h^{i-1}) \right\|_{L^2(\Omega)}^2 dt \\ &\quad + \int_{t_{N-1/2}}^T \|v_h^{N-1}\|_{L^2(\Omega)}^2 dt \\ &\leq \frac{\delta t}{2} \|v_h^0\|_{L^2(\Omega)}^2 + \frac{\delta t}{2} \|v_h^{N-1}\|_{L^2(\Omega)}^2 + \delta t \sum_{i=1}^{N-1} \|v_h^i - v_h^{i-1}\|_{L^2(\Omega)}^2 \\ &\leq C(\delta t) \left[ \|v_h^0\|_{L^2(\Omega)}^2 + \|v_h^{N-1}\|_{L^2(\Omega)}^2 + \sum_{i=1}^{N-1} \|v_h^i - v_h^{i-1}\|_{L^2(\Omega)}^2 \right]. \end{aligned}$$

Thus, with estimate (4.21), we obtain that  $\hat{v} = v$ .

We then show that  $\mathcal{K}v_{h_m} \overset{*}{\rightharpoonup} \mathcal{K}v$  in  $L^\infty(0, T; H_0^1(\Omega))$ . For all  $\varphi \in L^1(0, T; H^{-1}(\Omega))$  we have

$$\begin{aligned} \langle \mathcal{K}v_{h_m} - \mathcal{K}v, \varphi \rangle &= \int_0^T \left\langle \varphi(t), \int_0^t (v_{h_m}(s) - v(s)) ds \right\rangle dt \\ &= \int_0^T \int_0^t \langle \varphi(t), v_{h_m}(s) - v(s) \rangle ds dt \\ &= \int_0^T \int_s^T \langle \varphi(t), v_{h_m}(s) - v(s) \rangle dt ds \\ &= \int_0^T \left\langle \int_s^T \varphi(t) dt, v_{h_m}(s) - v(s) \right\rangle ds. \end{aligned}$$

Since  $\int_s^T \varphi(t) dt \in L^\infty(0, T; H^{-1}(\Omega))$ , we get the desired convergence.

The last step of the proof is to show that  $u_{h_m} - u_0 - \mathcal{K}v_{h_m} \rightarrow 0$  in  $L^2(0, T; H_0^1(\Omega))$ .

$$\begin{aligned} \|u_{h_m} - u_0 - \mathcal{K}v_{h_m}\|_{L^2(0, T; H_0^1(\Omega))}^2 &= \int_0^T \|u_{h_m}(t) - u_0 - \mathcal{K}v_{h_m}(t)\|_{H_0^1(\Omega)}^2 dt \\ &= T_1 + T_2 + T_3, \end{aligned}$$

with

$$\begin{aligned}
T_1 &= \int_0^{t_{1/2}} \|u_0\|_{H_0^1(\Omega)}^2 dt, \\
T_2 &= \sum_{i=1}^{N-1} \int_{t_{i-1/2}}^{t_{i+1/2}} \|u_{h_m}^i - u_0 - \int_0^t v_{h_m}(s) ds\|_{H_0^1(\Omega)}^2 dt, \\
T_3 &= \int_{t_{N-1/2}}^T \|-u_0 - \int_0^t v_{h_m}(s) ds\|_{H_0^1(\Omega)}^2 dt.
\end{aligned}$$

We get for the first term

$$T_1 \leq \frac{\delta t}{2} \|u_0\|_{H_0^1(\Omega)}^2 \leq C(\delta t) \|u_0\|_{H_0^1(\Omega)}^2. \quad (4.43)$$

As far as  $T_3$  is concerned we have

$$\begin{aligned}
T_3 &= \int_{t_{N-1/2}}^T \|u_0 + \int_0^{t_{1/2}} v_{h_m}(d) ds + \dots + \int_{t_{N-1/2}}^t v_{h_m}(s) ds\|_{H_0^1(\Omega)}^2 dt \\
&= \int_{t_{N-1/2}}^T \|u_0 + \int_{t_{1/2}}^{t_{3/2}} v_{h_m}^1 ds + \dots + \int_{t_{N-3/2}}^{t_{N-1/2}} v_{h_m}^{N-1} ds\|_{H_0^1(\Omega)}^2 dt \\
&\leq \int_{t_{N-1/2}}^T \|u_0 + \sum_{j=1}^{N-1} \delta t v_{h_m}^j\|_{H_0^1(\Omega)}^2 dt \\
&\leq c \left[ \int_{t_{N-1/2}}^T \|u_0\|_{H_0^1(\Omega)}^2 dt + \int_{t_{N-1/2}}^T \left\| \sum_{j=1}^{N-1} \delta t v_{h_m}^j \right\|_{H_0^1(\Omega)}^2 dt \right] \\
&\leq C \frac{\delta t}{2} \|u_0\|_{H_0^1(\Omega)}^2 + C \frac{\delta t}{2} \left\| \sum_{j=1}^{N-1} \delta t v_{h_m}^j \right\|_{H_0^1(\Omega)}^2 \\
&\leq C(\delta t) \|u_0\|_{H_0^1(\Omega)}^2 + \tilde{C}(\delta t) \sum_{j=1}^{N-1} \delta t \|v_{h_m}^j\|_{H_0^1(\Omega)}^2.
\end{aligned} \quad (4.44)$$

We now turn to  $T_2$ .

$$\begin{aligned}
T_2 &= \sum_{i=1}^{N-1} \int_{t_{i-1/2}}^{t_{i+1/2}} \|u_{h_m}^i - u_0 - \int_0^t v_{h_m}(s) ds\|_{H_0^1(\Omega)}^2 dt \\
&= \sum_{i=1}^{N-1} \int_{t_{i-1/2}}^{t_{i+1/2}} \|u_{h_m}^i - u_0 - \int_0^{t_{1/2}} v_{h_m}(s) ds - \dots - \int_{t_{i-1/2}}^t v_{h_m}(s) ds\|_{H_0^1(\Omega)}^2 dt \\
&= \sum_{i=1}^{N-1} \int_{t_{i-1/2}}^{t_{i+1/2}} \|u_{h_m}^i - u_0 - \sum_{j=1}^{i-1} \delta t v_{h_m}^j - (t - t_{i-1/2}) v_{h_m}^i\|_{H_0^1(\Omega)}^2 dt \\
&= \sum_{i=1}^{N-1} \int_{t_{i-1/2}}^{t_{i+1/2}} \|u_{h_m}^i + u_{h_m}^0 - u_0 - u_{h_m}^0 - \sum_{j=1}^{i-1} \delta t v_{h_m}^j - (t - t_{i-1/2}) v_{h_m}^i\|_{H_0^1(\Omega)}^2 dt \\
&\leq C \sum_{i=1}^{N-1} \int_{t_{i-1/2}}^{t_{i+1/2}} \|u_{h_m}^0 - u_0\|_{H_0^1(\Omega)}^2 dt
\end{aligned}$$

$$\begin{aligned}
& + C \sum_{i=1}^{N-1} \int_{t_{i-1/2}}^{t_{i+1/2}} \|u_{h_m}^i - u_{h_m}^0 - \sum_{j=1}^{i-1} \delta t v_{h_m}^j - (t - t_{i-1/2})v_{h_m}^i\|_{H_0^1(\Omega)}^2 dt \\
& \leq \bar{C} \|u_{h_m}^0 - u_0\|_{H_0^1(\Omega)}^2 \\
& + C \sum_{i=1}^{N-1} \int_{t_{i-1/2}}^{t_{i+1/2}} \|u_{h_m}^i - u_{h_m}^{i-1} + u_{h_m}^{i-1} - \dots - u_{h_m}^0 - \sum_{j=1}^{i-1} \delta t v_{h_m}^j - (t - t_{i-1/2})v_{h_m}^i\|_{H_0^1(\Omega)}^2 dt \\
& \leq \bar{C} \|u_{h_m}^0 - u_0\|_{H_0^1(\Omega)}^2 \\
& + C \sum_{i=1}^{N-1} \int_{t_{i-1/2}}^{t_{i+1/2}} \left\| \sum_{j=1}^i \delta t v_{h_m}^{j-1} - \sum_{j=1}^{i-1} \delta t v_{h_m}^j - (t - t_{i-1/2})v_{h_m}^i \right\|_{H_0^1(\Omega)}^2 dt \\
& \leq \bar{C} \|u_{h_m}^0 - u_0\|_{H_0^1(\Omega)}^2 + C \sum_{i=1}^{N-1} \int_{t_{i-1/2}}^{t_{i+1/2}} \|v_{h_m}^0 - (t - t_{i-1/2})v_{h_m}^i\|_{H_0^1(\Omega)}^2 dt \\
& \leq \bar{C} \|u_{h_m}^0 - u_0\|_{H_0^1(\Omega)}^2 + \tilde{C} \delta t \|v_{h_m}^0\|_{H_0^1(\Omega)}^2 + \hat{C} \sum_{i=1}^{N-1} \int_{t_{i-1/2}}^{t_{i+1/2}} \|(t - t_{i-1/2})v_{h_m}^i\|_{H_0^1(\Omega)}^2 dt \\
& \leq \bar{C} \|u_{h_m}^0 - u_0\|_{H_0^1(\Omega)}^2 + \tilde{C} \delta t \|v_{h_m}^0\|_{H_0^1(\Omega)}^2 + \hat{C} \delta t^2 \sum_{i=1}^{N-1} \delta t \|v_{h_m}^i\|_{H_0^1(\Omega)}^2 \\
& \leq \bar{C} \|u_{h_m}^0 - u_0\|_{H_0^1(\Omega)}^2 + C(\delta t) \left[ \|v_{h_m}^0\|_{H_0^1(\Omega)}^2 + \sum_{i=1}^{N-1} \delta t \|v_{h_m}^i\|_{H_0^1(\Omega)}^2 \right]. \tag{4.45}
\end{aligned}$$

Gathering (4.43), (4.44), (4.45) and letting  $m$  tends to  $\infty$  we obtain the desired convergence. Moreover we have

$$u - u_0 - \mathcal{K}v = u_{h_m} - u_0 - \mathcal{K}v_{h_m} + u - u_{h_m} + \mathcal{K}v_{h_m} - \mathcal{K}v,$$

thus  $u - u_0 = \mathcal{K}v \in \mathcal{C}^0(0, T; H_0^1(\Omega)) \cap L^\infty(0, T; H_0^1(\Omega))$  and  $u' = v \in L^2(0, T; H_0^1(\Omega))$ . Moreover since  $\frac{\partial v}{\partial t} \in L^2(0, T; H^{-1}(\Omega))$ , by Lemma 1.2 of [56] we have that  $v \in \mathcal{C}^0(0, T; L^2(\Omega))$ .  $\square$

The following proposition is devoted to extract a convergent subsequence from  $\theta_{\mathcal{D}_m}$ .

**Proposition 4.3.9** *Let  $(\mathcal{T}_m)_{m \geq 1}$  be a sequence of admissible meshes such that there exists  $\xi > 0$  satisfying*

$$\text{for all } m \geq 1 \text{ for all } K \in \mathcal{T} \text{ for all } \sigma \in \mathcal{E}(K), d_{K,\sigma} \geq \xi d_\sigma,$$

*and let  $(\delta t_m)_{m \geq 1}$  be a sequence of time steps. If  $\theta_{\mathcal{D}_m} = (\theta_{\mathcal{T}_m}^i)_{1 \leq i \leq N}$  is a solution to (4.20) with  $\mathcal{T} = \mathcal{T}_m$  and  $\delta t = \delta t_m$ , then there exists a measurable function  $\theta$  finite a.e. in  $Q$  such that, up to a subsequence,  $T_n(\theta_{\mathcal{D}_m})$  converges to  $T_n(\theta)$  weakly in  $L^2(0, T; H_0^1(\Omega))$  and  $\theta_{\mathcal{D}_m}$  converges almost everywhere to  $\theta$ .*

In order to prove this proposition we use a discrete version of the Aubin-Simon Lemma proved in [47] and in [48].

**Proof** Thanks to estimates (4.32) we can extract a subsequence, still indexed by  $m$ , such that

$$T_n(\theta_{\mathcal{D}_m}) \rightharpoonup v_n \text{ in } L^2(0, T; H_0^1(\Omega)) \quad (4.46)$$

in the finite volume sense.

Moreover, since we have estimates (4.34) and (4.35), we also apply the discrete version of the Aubin-Simon Lemma to the function  $\tilde{h}_n(\theta_{\mathcal{D}_m})$ . It implies that for any  $n > 0$ ,  $\tilde{h}_n(\theta_{\mathcal{D}_m})$  lies in a compact set of  $L^1(Q)$ . Using the diagonal process, we can deduce that there exists a subsequence of  $\theta_{\mathcal{D}_m}$ , still indexed by  $m$ , such that  $\tilde{h}_n(\theta_{\mathcal{D}_m})$  converges almost everywhere to a measurable function  $w_n$  defined on  $Q$  for any  $n > 0$ . Due to the definition of  $\tilde{h}_n$ , the sequence  $T_n(\theta_{\mathcal{D}_m})$  converges almost everywhere in  $Q$  to the measurable function  $v_n$ .

Using Corollary 4.3.3 we now prove that  $\theta_{\mathcal{D}_m}$  is a Cauchy sequence in measure. Let  $\omega > 0$ . For all  $n > 0$  and all sequences of admissible meshes and time steps  $(\mathcal{T}_m, \delta t_m)_{m \geq 1}$  and  $(\mathcal{T}_p, \delta t_p)_{p \geq 1}$ , we have

$$\{|\theta_{\mathcal{D}_m} - \theta_{\mathcal{D}_p}| > \omega\} \subset \{|\theta_{\mathcal{D}_m}| > n\} \cup \{|\theta_{\mathcal{D}_p}| > n\} \cup \{|T_n(\theta_{\mathcal{D}_m}) - T_n(\theta_{\mathcal{D}_p})| > \omega\}.$$

Let  $\varepsilon > 0$  fixed. By (4.29), let  $n > 0$  such that, for all admissible meshes and time steps  $\mathcal{T}_m, \delta t_m$  and  $\mathcal{T}_p, \delta t_p$

$$\text{meas}(\{|\theta_{\mathcal{D}_m}| > n\}) + \text{meas}(\{|\theta_{\mathcal{D}_p}| > n\}) < \frac{\varepsilon}{2}.$$

Once  $n$  is chosen, we deduce that  $T_n(\theta_{\mathcal{D}_m})$  is a Cauchy sequence in measure, thus

$$\exists h_0 > 0; \forall h_{\mathcal{D}_m}, h_{\mathcal{D}_p} < h_0 \quad \text{meas}(\{|T_n(\theta_{\mathcal{D}_m}) - T_n(\theta_{\mathcal{D}_p})| > \omega\}) \leq \frac{\varepsilon}{2}.$$

Therefore, we deduce that  $\forall h_{\mathcal{D}_m}, h_{\mathcal{D}_p} < h_0$

$$\text{meas}\{|\theta_{\mathcal{D}_m} - \theta_{\mathcal{D}_p}| > \omega\} < \varepsilon.$$

Hence  $\theta_{\mathcal{D}_m}$  is a Cauchy sequence in measure. It follows that, up to a subsequence still indexed by  $\mathcal{D}_m$ , there exists a measurable function  $\theta$  such that  $\theta_{\mathcal{D}_m} \rightarrow \theta$  a.e. in  $Q$ . Due to Corollary 4.3.3  $\theta$  is finite almost everywhere. Moreover from convergences obtained before we get that  $T_n(\theta) \in L^2(0, T; H_0^1(\Omega))$  and

$$\begin{aligned} \nabla_{\mathcal{D}} T_n(\theta_{\mathcal{D}}) &\rightharpoonup \nabla T_n(\theta) \text{ in } (L^2(Q))^d, \\ T_n(\theta_{\mathcal{D}}) &\rightarrow T_n(\theta) \text{ almost everywhere in } Q. \end{aligned}$$

□

## 4.4 Convergence analysis

Let us now state the main result of this paper.

**Theorem 4.4.1** *If  $\mathcal{T}$  is an admissible mesh and  $\delta t$  a time step, then there exists at least one solution to (4.19)-(4.20).*

*If  $(\mathcal{T}_m)_{m \geq 1}$  is a sequence of admissible meshes such that there exists  $\xi > 0$  satisfying*

$$\text{for all } m \geq 1 \text{ for all } K \in \mathcal{T} \text{ and all } \sigma \in \mathcal{E}(K), d_{K,\sigma} \geq \xi d_\sigma,$$

and such that  $h_{\mathcal{T}_m} \rightarrow 0$ , if  $(\delta t_m)_{m \geq 1}$  is a sequence of time steps such that  $\delta t_m \rightarrow 0$  and if  $\mathbf{V}_{h_m}$  is a sequence of the associated finite element spaces, then if  $(u_{h_m}, \theta_{\mathcal{D}_m})$  is the solution to (4.16) with  $\mathcal{T} = \mathcal{T}_m$  and  $\delta t_m = \delta t$ ,

—  $u_{h_m}$  converges to  $u$  in  $L^\infty(0, T; H_0^1(\Omega))$  weak-\*, with  $v_{h_m}$  which converges to  $\frac{\partial u}{\partial t}$  strongly in  $L^2(0, T, H_0^1(\Omega))$ ,

—  $\theta_{\mathcal{D}_m}$  converges to  $\theta$  almost everywhere in  $Q$ ,

where  $(u, \theta)$  is the unique renormalized solution of (4.1) to (4.4).

**Proof** The proof is divided into four steps. In Step 1 using the Brouwer fixed point theorem we show that there exists at least one solution to (4.19)-(4.20). Step 2 is devoted to pass to the limit in (4.19) and show that  $u$  is a weak solution of (4.1). In Step 3 we prove that  $\nabla v_h$  converges to  $\nabla v$  strongly in  $L^2(Q)^d$  and show that  $\theta$  is a renormalized solution of (4.2). In Step 4 we conclude the convergence analysis, showing that the limit  $\theta$  satisfies the estimate on the energy (4.12).

**Step 1** Let  $\hat{\theta} \in L^1(Q)$  be fixed. We define for all  $1 \leq i \leq N - 1$

$$\hat{\theta}^i = \frac{1}{\delta t} \int_{t_{i-1/2}}^{t_{i+1/2}} \hat{\theta}(t) dt.$$

Let us consider the following scheme :  $\forall \varphi \in \mathbf{V}_h, \forall K \in \mathcal{T}$

$$\begin{aligned} \int_{\Omega} \frac{v_h^i - v_h^{i-1}}{\delta t} \cdot \varphi dx + \int_{\Omega} [\varepsilon(u_h^0 + \mathcal{K}_T v_h^i) + \varepsilon(v_h^i)] \cdot \nabla \varphi dx \\ = \int_{\Omega} g_h^i \cdot \varphi dx + \int_{\Omega} f(\hat{\theta}^i) \cdot \nabla \varphi dx \quad i = 1, \dots, N - 1, \end{aligned} \quad (4.47)$$

$$\begin{aligned} |K| \frac{\theta_K^i - \theta_K^{i-1}}{\delta t} + \sum_{\sigma \in \mathcal{E}(K)} \frac{|\sigma|}{d_\sigma} \lambda(\hat{\theta}^i)_\sigma (\theta_K^i - \theta_L^i) \\ = \frac{1}{|K|} \int_K \left( \varepsilon(v_h^i) \cdot \varepsilon(v_h^i) - f(\hat{\theta}^i) \text{tr}(\varepsilon(v_h^i)) \right) dx \quad i = 1, \dots, N, \end{aligned} \quad (4.48)$$

where  $\lambda(\hat{\theta}^i)_\sigma$  is the mean value of  $\frac{1}{|K|} \int_K \hat{\theta}^i dx$  and  $\frac{1}{|L|} \int_L \hat{\theta}^i dx$ . This scheme is initialized by  $u_h^0 = 0$  and  $\theta_{\mathcal{D}}^0 = 0$ .

Each equation is linear and each solutions  $u_h$  and  $\theta_{\mathcal{D}}$  satisfy respectively estimates (4.3.1) and (4.27). Taking null right-hand sides, thanks to the estimates, we find that the solutions of this system are equal to zero. Thus each equation admits a unique solution.

We now consider the two following applications

$$\Gamma_1 : \begin{cases} W \rightarrow \widehat{V}_h \\ \hat{\theta} \mapsto u_h \text{ unique solution of (4.47),} \end{cases}$$

and

$$\Gamma_2 : \begin{cases} \widehat{V}_h \times W \rightarrow W \\ (\hat{v}_h, \hat{\theta}) \mapsto \theta_{\mathcal{D}} \text{ unique solution of (4.48),} \end{cases}$$

where

$$\begin{aligned} W &= \{ w \in L^1(Q); w_{|[t_{i-1}, t_i] \times K} \text{ constant} \}, \\ \widehat{V}_h &= \{ \hat{v}_h \in L^2(Q); \hat{v}_h^i \in \mathbf{V}_h \quad \forall 1 \leq i \leq N-1 \}. \end{aligned}$$

The function  $\Gamma_1$  is continuous. Indeed, let  $(\hat{\theta}_n)_{n \in \mathbb{N}} \in W$  such that  $\hat{\theta}_n$  converges strongly in  $L^1(Q)$  to  $\hat{\theta} \in W$ . We define  $u_{h_n} = \Gamma_1(\hat{\theta}_n)$  and  $u_h = \Gamma_1(\hat{\theta})$ . We now show that  $u_{h_n}$  converges strongly in  $L^2(0, T; H_0^1(\Omega))$  to  $u_h \in \widehat{V}_h$ . Taking the difference between the equations in  $(u_{h_n}, \hat{\theta}_n)$  and in  $(u_h, \hat{\theta})$  and using the techniques developed to prove estimate (4.3.1) we find for all  $1 \leq i \leq N-1$

$$\|u_{h_n}^i - u_h^i\|_{H_0^1(\Omega)}^2 \leq \|v_{h_n}^0 - v_h^0\|_{L^2(\Omega)}^2 + \|u_{h_n}^0 - u_h^0\|_{H_0^1(\Omega)}^2 + C \|f(\hat{\theta}_n) - f(\hat{\theta})\|_{L^2(Q)}^2.$$

Since  $f$  is a continuous function we obtain the desired convergence, thus  $\Gamma_1$  is continuous.

We then prove that  $\Gamma_2$  is also continuous. Let  $(\hat{v}_{h_n}, \hat{\theta}_n)_{n \in \mathbb{N}} \in \widehat{V}_h \times W$  such that  $(\hat{v}_{h_n}, \hat{\theta}_n)$  converges strongly in  $L^2(0, T; H_0^1(\Omega)) \times L^1(Q)$  to  $(\hat{v}_h, \hat{\theta}) \in \widehat{V}_h \times W$  and let us define  $\theta_{\mathcal{D}_n} = \Gamma_2(\hat{v}_{h_n}, \hat{\theta}_n)$ ,  $\theta_{\mathcal{D}} = \Gamma_2(\hat{v}_h, \hat{\theta})$ . Applying the finite volumes techniques to equation (4.48) with  $\hat{v}_{h_n}$  and  $\hat{\theta}_n$ , we find that

$$\|\theta_{\mathcal{D}_n}\|_{L^1(Q)} \leq C.$$

Then since the sets  $W$  and  $\widehat{V}_h$  are of finite dimension, up to subsequence,  $\theta_{\mathcal{D}_n}$  converges strongly in  $L^1(Q)$  to a limit  $\overline{\theta_D}$  which is a solution to (4.48). Since the solution is unique  $\overline{\theta_D} = \theta_{\mathcal{D}}$ . Thus  $\Gamma_2$  is continuous.

Applying the Brouwer fixed point theorem to the function  $\Gamma_2(\Gamma_1(\cdot), \cdot)$ , we deduce that there exists at least one solution to the scheme (4.19)-(4.20).

**Step 2** For all  $1 \leq i \leq N-1$  we multiply (4.19) by  $\psi(t)$  where  $\psi \in (C_c^1((0, T)))^d$  and we integrate on the time.

$$\begin{aligned} & - \int_0^T \left( \int_{\Omega} \hat{v}_h(t) \varphi \, dx \right) \psi'(t) \, dt + \int_0^T \left( \int_{\Omega} \varepsilon(u_h(t)) \cdot \nabla \varphi \, dx \right) \psi(t) \, dt \\ & + \int_0^T \left( \int_{\Omega} \varepsilon(v_h(t)) \cdot \nabla \varphi \, dx \right) \psi(t) \, dt = \int_0^T \left( \int_{\Omega} g_h(t) \varphi \, dx \right) \psi(t) \, dt \\ & \quad + \int_0^T \left( \int_{\Omega} f(\theta_{\mathcal{D}_h})(t) \cdot \nabla \varphi \, dx \right) \psi(t) \, dt. \end{aligned} \quad (4.49)$$

In view of the definition of  $\varepsilon$  and in view of the convergence results we get

$$\begin{aligned} \varepsilon(v_h) &\rightharpoonup \varepsilon(v) \text{ in } (L^2(Q))^{d \times d}, \\ \varepsilon(u_h) &\rightharpoonup \varepsilon(u) \text{ in } (L^2(Q))^{d \times d}. \end{aligned}$$

By definition of  $g_h$  we have

$$g_h \rightarrow g \text{ in } (L^2(Q))^d.$$

Since the right-hand side of (4.20) belongs to  $L^1(Q)$ , we have Boccardo-Gallouët estimates on  $\theta_{\mathcal{D}}$  (see [47] for instance). Thus by the discrete version of Aubin-Simon

Lemma (proved in [47]),  $\theta_{\mathcal{D}}$  converges to  $\theta$  in  $L^1(Q)$  and almost everywhere in  $Q$ . Due to the fact that  $f$  is a continuous function,  $f(\theta_{\mathcal{D}})$  converges to  $f(\theta)$  in  $L^2(Q)$  and almost everywhere in  $Q$ . We now turn to the convergence of  $f(\theta_{\mathcal{D}_h})$ . We notice that

$$\theta_{\mathcal{D}_h}(t) = \begin{cases} 0 & \text{on } [0, t_{1/2}], \\ \frac{\theta_{\mathcal{D}}(t) + \theta_{\mathcal{D}}(t + \delta t)}{2} & \text{on } (t_{i-1/2}, t_i], \\ \frac{\theta_{\mathcal{D}}(t - \delta t) + \theta_{\mathcal{D}}(t)}{2} & \text{on } (t_i, t_{i+1/2}], \\ 0 & \text{on } (t_{N-1/2}, T]. \end{cases}$$

Since we have Boccardo-Gallouët estimates on  $\theta_{\mathcal{D}}$  and since the time step  $\delta t$  is fixed, we can apply Lemma 3.4 of [47] to  $\theta_{\mathcal{D}}$ . Thus we have the estimate of the time translates :

$$\forall \varepsilon > 0, \|\theta_{\mathcal{D}}(\cdot, t + \delta t) - \theta_{\mathcal{D}}(\cdot, t)\|_{L^1(Q)} \leq \varepsilon.$$

We then deduce that

$$\forall \varepsilon > 0, \|\theta_{\mathcal{D}_h} - \theta_{\mathcal{D}}\|_{L^1(Q)} \leq \varepsilon.$$

Thus  $f(\theta_{\mathcal{D}_h})$  converges to  $f(\theta)$  in  $L^2(Q)$  and almost everywhere in  $Q$ .

Passing to the limit in (4.49) we get for all  $\varphi \in \mathbf{V}_h$  and for all  $\psi \in \mathcal{C}_c^1((0, T))$

$$\begin{aligned} & - \int_0^T \left( \int_{\Omega} v(t) \varphi \, dx \right) \psi'(t) \, dt + \int_0^T \left( \int_{\Omega} \varepsilon(u(t)) \cdot \nabla \varphi \, dx \right) \psi(t) \, dt \\ & + \int_0^T \left( \int_{\Omega} \varepsilon(v(t)) \cdot \nabla \varphi \, dx \right) \psi(t) \, dt = \int_0^T \left( \int_{\Omega} g(t) \varphi \, dx \right) \psi(t) \, dt \\ & \quad + \int_0^T \left( \int_{\Omega} f(\theta)(t) \cdot \nabla \varphi \, dx \right) \psi(t) \, dt, \end{aligned} \quad (4.50)$$

which holds true for all  $\varphi \in H_0^1(\Omega)$ . Thus  $u$  is a weak solution of equation (4.1)

**Step 3** In order to pass to the limit in the renormalized sense in (4.20), we have to obtain a discrete estimate on the energy as in [52] which is the discrete equivalent of (4.12). This is possible only if we can pass to the limit in the right-hand side of (4.20), that is if  $\nabla v_{h_m}$  converges strongly to  $\nabla v$  in  $L^2(Q)^d$ .

Let  $u$  be the limit of  $u_h$  and  $\theta$  the limit of  $\theta_{\mathcal{D}}$  and let us define

$$\begin{aligned} \tilde{u}^i &= \frac{1}{\delta t} \int_{t_{i-1/2}}^{t_{i+1/2}} u(t) \, dt \text{ and } \tilde{v}^i = \frac{\tilde{u}^{i+1} - \tilde{u}^i}{\delta t} \quad \forall 1 \leq i \leq N-1, \\ \tilde{\theta}^i &= \frac{1}{\delta t} \int_{t_{i-1/2}}^{t_{i+1/2}} \theta(t) \, dt \quad \forall 1 \leq i \leq N-1. \end{aligned}$$

Then  $u$  satisfies the following equation :

$$\begin{aligned} & \int_{\Omega} \frac{\tilde{v}^i - \tilde{v}^{i-1}}{\delta t} \varphi \, dx + \int_{\Omega} [\varepsilon(u_0 + \mathcal{K}_I \tilde{v}^i) + \varepsilon(\tilde{v}^i)] \cdot \nabla \varphi \, dx \\ & = \int_{\Omega} g^i \varphi \, dx + \int_{\Omega} f(\tilde{\theta}^i) \cdot \nabla \varphi \, dx \quad \forall \varphi \in \mathbf{V}_h. \end{aligned} \quad (4.51)$$

Thanks to the linearity of equation (4.19) it is sufficient to take the difference between (4.19) and (4.51). In [42], because of the nonlinearity of the operators the proof

is not as simple. We get

$$\begin{aligned} & \int_{\Omega} \frac{(v_h^i - \tilde{v}^i) - (v_h^{i-1} - \tilde{v}^{i-1})}{\delta t} \varphi \, dx \\ & + \int_{\Omega} [\varepsilon((u_h^0 - u_0) + \mathcal{K}_I(v_h^i - \tilde{v}^i)) + \varepsilon(v_h^i - \tilde{v}^i)] \cdot \nabla \varphi \, dx \\ & = \int_{\Omega} [f(\theta_{\mathcal{D}}^i) - f(\tilde{\theta}^i)] \cdot \nabla \varphi \, dx. \end{aligned}$$

Using the techniques developed in the proof of estimate (4.21), we find  $\forall 1 \leq n \leq N-1$

$$\sum_{i=1}^n \delta t \|v_h^i - \tilde{v}^i\|_{H_0^1(\Omega)}^2 \leq \|v_h^0 - v_0\|_{L^2(\Omega)}^2 + \|u_h^0 - u_0\|_{L^2(\Omega)}^2 + \|f(\theta_{\mathcal{D}}) - f(\theta)\|_{L^2(\Omega)}^2.$$

Thus  $v_h \rightarrow \tilde{v}$  in  $L^2(0, T; H_0^1(\Omega))$ . Remarking that for all  $1 \leq i \leq N-1$

$$\tilde{v}^i = \frac{1}{\delta t} \int_{t_{i-1/2}}^{t_{i+1/2}} \frac{u(t + \delta t) - u(t)}{\delta t} \, dt,$$

we can prove with the help of Steklov averages and the regularity of  $u$  and  $v$  that  $\tilde{v} \rightarrow v$  in  $L^2(0, T; H_0^1(\Omega))$ . Then  $\nabla v_h \rightarrow \nabla v$  in  $L^2(0, T; L^2(\Omega))$  and the right-hand side of (4.20) converges to

$$\int_0^T \int_{\Omega} (\varepsilon(v)\varepsilon(v) - f(\theta)\text{tr}(\varepsilon(v))) \, dx \, dt.$$

We also have the discrete estimate on the energy

$$\lim_{n \rightarrow +\infty} \lim_{\substack{h_T \rightarrow 0 \\ \delta t \rightarrow 0}} \frac{1}{n} \sum_{i=1}^N \delta t \sum_{\sigma \in \mathcal{E}} \frac{|\sigma|}{d_{\sigma}} \lambda(\theta)_{\sigma}^i (\theta_K^i - \theta_L^i) (T_n(\theta_K^i) - T_n(\theta_L^i)) = 0. \quad (4.52)$$

Thanks to remarks 3.4 and 3.9 of Chapter 2, the results established in this paper hold true. Thus  $\theta$  is a renormalized solution of (4.2).

**Step 4** We now prove that  $\theta$  satisfies the decay (4.12) of the truncate energy. Thanks to the discrete estimate on the energy (4.52) we have

$$\lim_{n \rightarrow \infty} \overline{\lim}_{\substack{h_T \rightarrow 0 \\ \delta t \rightarrow 0}} \frac{1}{n} \sum_{i=1}^N \delta t \sum_{\sigma \in \mathcal{E}} \frac{|\sigma|}{d_{\sigma}} \lambda(\theta)_{\sigma}^i (T_{2n}(\theta_K^i) - T_{2n}(\theta_L^i))^2 = 0,$$

and

$$\begin{aligned} \sum_{i=1}^N \delta t \sum_{\sigma \in \mathcal{E}} \frac{|\sigma|}{d_{\sigma}} \lambda(\theta)_{\sigma}^i (T_{2n}(\theta_K^i) - T_{2n}(\theta_L^i))^2 &= \sum_{i=1}^N \delta t \sum_{\sigma \in \mathcal{E}} |\sigma| d_{\sigma} \lambda(\theta)_{\sigma}^i \left( \frac{T_{2n}(\theta_K^i) - T_{2n}(\theta_L^i)}{d_{\sigma}} \right)^2 \\ &= \sum_{i=1}^N \delta t \sum_{\sigma \in \mathcal{E}} d |D_{\sigma}| \lambda(\theta)_{\sigma}^i \left( \frac{T_{2n}(\theta_K^i) - T_{2n}(\theta_L^i)}{d_{\sigma}} \right)^2 \\ &= \sum_{i=1}^N \delta t \frac{1}{d} \sum_{\sigma \in \mathcal{E}} |D_{\sigma}| \lambda(\theta)_{\sigma}^i \left( \frac{T_{2n}(\theta_K^i) - T_{2n}(\theta_L^i)}{d_{\sigma}} \right)^2 \\ &\geq \frac{1}{d} \int_0^T \int_{\Omega} h_n(\theta_{\mathcal{D}}) \frac{\lambda(T_{2n}(\theta_{\mathcal{D}}))}{2} |\nabla_{\mathcal{D}} T_{2n}(\theta_{\mathcal{D}})|^2 \, dx \, dt, \end{aligned}$$



thus,  $\lim_{n \rightarrow \infty} \overline{\lim}_{\substack{h_T \rightarrow 0 \\ \delta t \rightarrow 0}} \frac{1}{n} \int_0^T \int_{\Omega} h_n(\theta_{\mathcal{D}}) \frac{\lambda(T_{2n}(\theta_{\mathcal{D}}))}{2} |\nabla_{\mathcal{D}} T_{2n}(\theta_{\mathcal{D}})|^2 dx dt = 0$ . Since  $\nabla T_{2n}(\theta_{\mathcal{D}})$  converges weakly in  $(L^2(Q))^d$ ,  $\lambda(T_{2n}(\theta_{\mathcal{D}}))$  converges to  $\lambda(\theta)$   $L^\infty$  weak- $\star$  and  $h_n(\theta_{\mathcal{D}})$  converges to  $h_n(\theta)$  almost everywhere in  $Q$  we also have

$$\frac{1}{n} \int_0^T \int_{\Omega} h_n(\theta) \lambda(\theta) |\nabla T_{2n}(\theta)|^2 dx \leq \liminf_{\substack{h_T \rightarrow 0 \\ \delta t \rightarrow 0}} \frac{1}{n} \int_0^T \int_{\Omega} h_n(\theta_{\mathcal{D}}) \frac{\lambda(\theta_{\mathcal{D}})}{2} |\nabla_{\mathcal{D}} T_{2n}(\theta_{\mathcal{D}})|^2 dx dt,$$

which leads to

$$\lim_{n \rightarrow \infty} \frac{1}{n} \int_0^T \int_{\Omega} \lambda(\theta) |\nabla T_n(\theta)|^2 dx dt = 0.$$

□

# Chapitre 5

## Quelques perspectives

### 5.1 Perspectives sur les solutions renormalisées et les volumes finis

Concernant les problèmes stationnaires un prochain travail pourrait être d'étudier une équation elliptique sous forme de divergence complètement non linéaire. Une difficulté supplémentaire viendrait alors de la discrétisation avec les volumes finis du terme  $-\operatorname{div}(a(\nabla(u)))$ . Pour gérer cette difficulté venant du terme en gradient une possibilité serait d'utiliser en dimension 2 les maillages duaux (voir [3] dans le cas variationnel) ou la méthode des gradients (voir [39] dans un cas très général variationnel).

Par la suite on pourrait aussi se pencher sur l'étude de l'équation de convection-diffusion (1.1) avec d'autres conditions au bord (Neumann, Fourier...). Par exemple, un schéma aux volumes finis est considéré dans [28] pour l'équation

$$\begin{aligned} -\Delta u + \operatorname{div}(Vu) &= g \text{ dans } \Omega, \\ \nabla u \cdot \eta - (V \cdot \eta)u &= 0 \text{ sur } \partial\Omega, \end{aligned}$$

où  $g \in L^2(\Omega)$ . Les auteurs prouvent que, si  $\int_{\Omega} g = 0$ , la solution du schéma converge vers l'unique solution faible  $u$  de moyenne nulle. L'idée serait de mixer les techniques développées dans [28] avec celles utilisées dans ce chapitre pour étudier une équation plus générale  $(-\operatorname{div}(\lambda(u)\nabla u) + \dots)$  et prouver que la solution du schéma converge vers l'unique solution renormalisée.

Concernant les problèmes d'évolution, si l'étude d'une équation elliptique sous forme de divergence complètement non linéaire se révèle concluante, la prochaine étape serait de voir si les résultats restent vrais dans le cas parabolique.

Une autre idée pour un prochain travail serait de rajouter une non-linéarité dans l'équation. On pourrait par exemple étudier l'équation

$$\frac{\partial u}{\partial t} - \operatorname{div}(A\nabla u) + \operatorname{div}(\psi(u)) = f \text{ dans } Q,$$

où  $\psi$  est une fonction continue. Dans [21] les auteurs ont montré l'existence et l'unicité d'une solution renormalisée pour cette équation. En utilisant leur travail et les

techniques développées dans le chapitre 2, nous pourrions peut-être montrer que la solution approchée par la méthode des volumes finis converge vers l'unique solution renormalisée.

Une extension naturelle du chapitre 3 serait l'ajout d'une non-linéarité à l'équation. On se retrouverait alors à étudier une équation parabolique doublement non-linéaire de la forme

$$\frac{\partial b(u)}{\partial t} - \operatorname{div}(A\nabla u) + \operatorname{div}(\psi(u)) = f,$$

ou encore

$$\frac{\partial b(u)}{\partial t} - \operatorname{div}(\lambda(u)\nabla u) + \operatorname{div}(c(x, t)u) = f.$$

L'idée serait de montrer que les résultats du chapitre restent encore valides dans ce cas, c'est à dire que la solution du schéma de type volumes finis converge vers l'unique solution renormalisée du problème.

Enfin dans le chapitre 3 nous nous sommes restreints au cas  $b \in \mathcal{C}^1$ ,  $b' > 0$ . Dans le cas continu et une donnée  $L^1$  il n'y a pas cette restriction et des résultats d'existence et d'unicité de la solution renormalisée sont démontrés dans [20] pour une fonction  $b$  assez générale (graphe maximal monotone). Pour une donnée plus régulière ( $L^2$  par exemple) et à l'aide des translatés en temps et en espace des résultats de convergence ont été obtenus (voir [7, 8]). Affaiblir la condition sur  $b$  du chapitre 3 paraît très intéressant et délicat.

## 5.2 Perspectives sur les systèmes

La première perspective au travail du chapitre 4 serait d'ajouter les tenseurs et les données dans les équations (4.1)-(4.2) et de considérer le modèle de Kelvin-Voigt. Comme dans [15] on aurait alors à traiter le système

$$\begin{aligned} \frac{\partial^2 u}{\partial t^2} - \operatorname{div} \left( B_1 \varepsilon \left( \frac{\partial u}{\partial t} \right) + B_2 \varepsilon(u) \right) + \kappa D(\theta + T_0) &= g, \\ (T_0 + \theta) \frac{\partial \theta}{\partial t} - \Delta \theta = B_1 \varepsilon \left( \frac{\partial u}{\partial t} \right) \cdot \varepsilon \left( \frac{\partial u}{\partial t} \right) - \kappa (T_0 + \theta) \operatorname{tr} \left( \varepsilon \left( \frac{\partial u}{\partial t} \right) \right). \end{aligned}$$

Par rapport au système traité dans le chapitre 4, il y a plusieurs difficultés supplémentaires. D'une part la fonction  $\theta + T_0$  n'est pas bornée par rapport à  $\theta$ . D'autre part  $(T_0 + \theta) \frac{\partial \theta}{\partial t}$  correspond à  $\frac{1}{2} \left( \frac{\partial(T_0 + \theta)}{\partial t} \right)^2$  : il faut montrer dans la version discrète que  $T_0 + \theta \geq 0$  et comme  $T_0 + \theta \geq 0$  nous ne sommes pas dans le cas du chapitre 3 avec  $b' > 0$ , mais dans celui d'un problème dégénéré.

Une autre perspective serait d'appliquer les techniques décrites dans cette thèse

à l'étude d'un système de type Boussinesq (voir [9] pour le cas continu) :

$$\begin{aligned}\frac{\partial u}{\partial t} + (u \cdot \nabla)u - 2\operatorname{div}(\mu(\theta)Du) + \nabla p &= F(\theta), \\ \frac{\partial b(\theta)}{\partial t} + u \cdot \nabla b(\theta) - \Delta \theta &= 2\mu(\theta)|Du|^2, \\ \operatorname{div} u &= 0.\end{aligned}$$

En utilisant les techniques développées dans les chapitres 3 et 4, nous devrions être en mesure de prouver que la solution approchée par un schéma combinant éléments finis et volumes finis converge vers une solution faible-renormalisée du système.



# Bibliographie

- [1] B. Andreianov. Time compactness tools for discretized evolution equations and applications to degenerate parabolic PDEs. In *Finite volumes for complex applications VI. Problems & perspectives. Volume 1, 2*, volume 4 of *Springer Proc. Math.*, pages 21–29. Springer, Heidelberg, 2011.
- [2] B. Andreianov, M. Bendahmane, and K. H. Karlsen. Discrete duality finite volume schemes for doubly nonlinear degenerate hyperbolic-parabolic equations. *J. Hyperbolic Differ. Equ.*, 7(1) :1–67, 2010.
- [3] B. Andreianov, F. Boyer, and F. Hubert. Discrete duality finite volume schemes for Leray-Lions-type elliptic problems on general 2D meshes. *Numer. Methods Partial Differential Equations*, 23(1) :145–195, 2007.
- [4] B. Andreianov, C. Cancès, and A. Moussa. A nonlinear time compactness result and applications to discretization of degenerate parabolic-elliptic pdes. *HAL : hal-01142499*, 2015.
- [5] B. Andreianov, R. Eymard, M. Ghilani, and N. Marhraoui. Finite volume approximation of degenerate two-phase flow model with unlimited air mobility. *Numer. Methods Partial Differential Equations*, 29(2) :441–474, 2013.
- [6] B. Andreianov, M. Gutnic, and P. Wittbold. Convergence of finite volume approximations for a nonlinear elliptic-parabolic problem : a “continuous” approach. *SIAM J. Numer. Anal.*, 42(1) :228–251, 2004.
- [7] B. Andreianov and P. Wittbold. Convergence of approximate solutions to an elliptic-parabolic equation without the structure condition. *NoDEA Nonlinear Differential Equations Appl.*, 19(6) :695–717, 2012.
- [8] O. Angelini, K. Brenner, and D. Hilhorst. A finite volume method on general meshes for a degenerate parabolic convection-reaction-diffusion equation. *Numer. Math.*, 123(2) :219–257, 2013.
- [9] A. Attaoui, D. Blanchard, and O. Guibé. Weak-renormalized solution for a nonlinear Boussinesq system. *Differential Integral Equations*, 22(5-6) :465–494, 2009.
- [10] A. Ben Cheikh and O. Guibé. Résultats d’existence et d’unicité pour une classe de problèmes non linéaires et non coercifs. *C. R. Acad. Sci. Paris Sér. I Math.*, 329(11) :967–972, 1999.
- [11] A. Ben Cheikh and O. Guibé. Nonlinear and non-coercive elliptic problems with integrable data. *Adv. Math. Sci. Appl.*, 16(1) :275–297, 2006.
- [12] P. Bénilan, L. Boccardo, T. Gallouët, R. Gariepy, M. Pierre, and J.L. Vázquez. An  $L^1$ -theory of existence and uniqueness of solutions of nonlinear elliptic equations. *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4)*, 22(2) :241–273, 1995.

- [13] M. Bessemoulin-Chatard, C. Chainais-Hillairet, and F. Filbet. On discrete functional inequalities for some finite volume schemes. *IMA J. Numer. Anal.*, 35(3) :1125–1149, 2015.
- [14] D. Blanchard. Renormalized solutions for parabolic problems with  $L^1$  data. In *Free boundary problems, theory and applications (Zakopane, 1995)*, volume 363 of *Pitman Res. Notes Math. Ser.*, pages 177–185. Longman, Harlow, 1996.
- [15] D. Blanchard and O. Guibé. Existence of a solution for a nonlinear system in thermoviscoelasticity. *Adv. Differential Equations*, 5(10-12) :1221–1252, 2000.
- [16] D. Blanchard, O. Guibé, and H. Redwane. Existence and uniqueness of a solution for a class of parabolic equations with two unbounded nonlinearities. *Commun. Pure Appl. Anal.*, 15(1) :197–217, 2016.
- [17] D. Blanchard and F. Murat. Renormalised solutions of nonlinear parabolic problems with  $L^1$  data : existence and uniqueness. *Proc. Roy. Soc. Edinburgh Sect. A*, 127(6) :1137–1152, 1997.
- [18] D. Blanchard, F. Murat, and H. Redwane. Existence and uniqueness of a renormalized solution for a fairly general class of nonlinear parabolic problems. *J. Differential Equations*, 177(2) :331–374, 2001.
- [19] D. Blanchard and A. Porretta. Nonlinear parabolic equations with natural growth terms and measure initial data. *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4)*, 30(3-4) :583–622 (2002), 2001.
- [20] D. Blanchard and A. Porretta. Stefan problems with nonlinear diffusion and convection. *J. Differential Equations*, 210(2) :383–428, 2005.
- [21] D. Blanchard and H. Redwane. Renormalized solutions for a class of nonlinear evolution problems. *J. Math. Pures Appl. (9)*, 77(2) :117–151, 1998.
- [22] L. Boccardo, A. Dall’Aglia, T. Gallouët, and L. Orsina. Nonlinear parabolic equations with measure data. *J. Funct. Anal.*, 147(1) :237–258, 1997.
- [23] L. Boccardo and T. Gallouët. Nonlinear elliptic and parabolic equations involving measure data. *J. Funct. Anal.*, 87(1) :149–169, 1989.
- [24] L. Boccardo and T. Gallouët. Nonlinear elliptic equations with right-hand side measures. *Comm. Partial Differential Equations*, 17(3-4) :641–655, 1992.
- [25] L. Boccardo, T. Gallouët, and F. Murat. Unicité de la solution de certaines équations elliptiques non linéaires. *C. R. Acad. Sci. Paris Sér. I Math.*, 315(11) :1159–1164, 1992.
- [26] J. Carrillo and M. Chipot. On some nonlinear elliptic equations involving derivatives of the nonlinearity. *Proc. Roy. Soc. Edinburgh Sect. A*, 100(3-4) :281–294, 1985.
- [27] J. Casado-Dí az, T. Chacón Rebollo, V. Girault, M. Gómez Mármol, and F. Murat. Finite elements approximation of second order linear elliptic equations in divergence form with right-hand side in  $L^1$ . *Numer. Math.*, 105(3) :337–374, 2007.
- [28] C. Chainais-Hillairet and J. Droniou. Finite-volume schemes for noncoercive elliptic problems with Neumann boundary conditions. *IMA J. Numer. Anal.*, 31(1) :61–85, 2011.

- [29] K. Chełmiński and S. Owczarek. Renormalised solutions in thermo-viscoplasticity for a Norton-Hoff type model. Part II : the limit case. *Nonlinear Anal. Real World Appl.*, 31 :643–660, 2016.
- [30] K. Chełmiński and S. Owczarek. Renormalized solutions in thermo-viscoplasticity for a Norton-Hoff type model. Part I : the truncated case. *Nonlinear Anal. Real World Appl.*, 28 :140–152, 2016.
- [31] K. Chełmiński and S. Owczarek. Corrigendum to “Renormalised solutions in thermo-viscoplasticity for a Norton-Hoff type model. Part I : the truncated case” [Nonlinear Anal. RWA 28 (2016) 140–152][ MR3422816]. *Nonlinear Anal. Real World Appl.*, 37 :489–492, 2017.
- [32] Y. Coudière, T. Gallouët, and R. Herbin. Discrete Sobolev inequalities and  $L^p$  error estimates for finite volume solutions of convection diffusion equations. *M2AN Math. Model. Numer. Anal.*, 35(4) :767–778, 2001.
- [33] G. Dal Maso, F. Murat, L. Orsina, and A. Prignet. Renormalized solutions of elliptic equations with general measure data. *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4)*, 28(4) :741–808, 1999.
- [34] R. Di Nardo, F. Feo, and O. Guibé. Existence result for nonlinear parabolic equations with lower order terms. *Anal. Appl. (Singap.)*, 9(2) :161–186, 2011.
- [35] R. Di Nardo, F. Feo, and O. Guibé. Uniqueness of renormalized solutions to nonlinear parabolic problems with lower-order terms. *Proc. Roy. Soc. Edinburgh Sect. A*, 143(6) :1185–1208, 2013.
- [36] R.J. DiPerna and P.L. Lions. On the Cauchy problem for Boltzmann equations : global existence and weak stability. *Ann. of Math. (2)*, 130(2) :321–366, 1989.
- [37] J. Droniou. Solving convection-diffusion equations with mixed, Neumann and Fourier boundary conditions and measures as data, by a duality method. *Adv. Differential Equations*, 5(10-12) :1341–1396, 2000.
- [38] J. Droniou. Global and local estimates for nonlinear noncoercive elliptic equations with measure data. *Comm. Partial Differential Equations*, 28(1-2) :129–153, 2003.
- [39] J. Droniou. Finite volume schemes for fully non-linear elliptic equations in divergence form. *M2AN Math. Model. Numer. Anal.*, 40(6) :1069–1100 (2007), 2006.
- [40] J. Droniou, T. Gallouët, and R. Herbin. A finite volume scheme for a noncoercive elliptic equation with measure data. *SIAM J. Numer. Anal.*, 41(6) :1997–2031, 2003.
- [41] J. Droniou and J.L. Vázquez. Noncoercive convection-diffusion elliptic problems with Neumann boundary conditions. *Calc. Var. Partial Differential Equations*, 34(4) :413–434, 2009.
- [42] E. Emmrich and D. Siska. Full discretisation of second-order nonlinear evolution equations : strong convergence and applications. *Comput. Methods Appl. Math.*, 11(4) :441–459, 2011.
- [43] E. Emmrich and M. Thalhammer. Doubly nonlinear evolution equations of second order : existence and fully discrete approximation. *J. Differential Equations*, 251(1) :82–118, 2011.



- [44] R. Eymard, T. Gallouët, and R. Herbin. Finite volume methods. In *Handbook of numerical analysis, Vol. VII*, Handb. Numer. Anal., VII, pages 713–1020. North-Holland, Amsterdam, 2000.
- [45] 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.
- [46] T. Gallouët and R. Herbin. Finite volume approximation of elliptic problems with irregular data. In *Finite volumes for complex applications II*, pages 155–162. Hermes Sci. Publ., Paris, 1999.
- [47] T. Gallouët, A. Larcher, and J.C. Latché. Convergence of a finite volume scheme for the convection-diffusion equation with  $L^1$  data. *Math. Comp.*, 81(279) :1429–1454, 2012.
- [48] T. Gallouët and J.C. Latché. Compactness of discrete approximate solutions to parabolic PDEs—application to a turbulence model. *Commun. Pure Appl. Anal.*, 11(6) :2371–2391, 2012.
- [49] O. Guibé and A. Mercaldo. Uniqueness results for noncoercive nonlinear elliptic equations with two lower order terms. *Commun. Pure Appl. Anal.*, 7(1) :163–192, 2008.
- [50] R. Landes. On the existence of weak solutions for quasilinear parabolic initial-boundary value problems. *Proc. Roy. Soc. Edinburgh Sect. A*, 89(3-4) :217–237, 1981.
- [51] A. Larcher and J.C. Latché. Convergence analysis of a finite element-finite volume scheme for a rans turbulence model. *Institut de radioprotection et de sureté nucléaire (IRSN)*, 2012.
- [52] S. Leclavier. Finite volume scheme and renormalized solutions for a noncoercive and nonlinear parabolic problem with  $L^1$  data. *In preparation*.
- [53] S. Leclavier. Finite volume scheme and renormalized solutions for a noncoercive elliptic problem with  $L^1$  data. *Comput. Methods Appl. Math.*, 17(1) :85–104, 2017.
- [54] J. Leray and J.L. Lions. Quelques résultats de Višik sur les problèmes elliptiques nonlinéaires par les méthodes de Minty-Browder. *Bull. Soc. Math. France*, 93 :97–107, 1965.
- [55] R. Lewandowski. *Analyse mathématique et océanographie*, volume 39 of *Recherches en Mathématiques Appliquées [Research in Applied Mathematics]*. Masson, Paris, 1997.
- [56] J.L. Lions. *Quelques méthodes de résolution des problèmes aux limites non linéaires*. Dunod ; Gauthier-Villars, Paris, 1969.
- [57] P.L. Lions. *Mathematical topics in fluid mechanics. Vol. 1*, volume 3 of *Oxford Lecture Series in Mathematics and its Applications*. The Clarendon Press, Oxford University Press, New York, 1996. Incompressible models, Oxford Science Publications.
- [58] P.L. Lions and F. Murat. Solutions renormalisées d'équations elliptiques non linéaires. *In preparation*.

- [59] F. Murat. Soluciones renormalizadas de edp elípticas no lineales. *Technical Report R93023*, 1993.
- [60] F. Murat. Equations elliptiques non linéaires avec second membre  $L^1$  ou mesure. *Compte Rendu du 26ème Congrès d'Analyse Numérique*, 1994.
- [61] F. Otto.  $L^1$ -contraction and uniqueness for quasilinear elliptic-parabolic equations. *J. Differential Equations*, 131(1) :20–38, 1996.
- [62] I. Pawłow and W.M. Zajączkowski. Global regular solutions to a Kelvin-Voigt type thermoviscoelastic system. *SIAM J. Math. Anal.*, 45(4) :1997–2045, 2013.
- [63] F. Petitta, A. Ponce, and A. Porretta. Diffuse measures and nonlinear parabolic equations. *J. Evol. Equ.*, 11(4) :861–905, 2011.
- [64] F. Petitta and A. Porretta. On the notion of renormalized solution to nonlinear parabolic equations with general measure data. *J. Elliptic Parabol. Equ.*, 1 :201–214, 2015.
- [65] A. Prignet. Remarks on existence and uniqueness of solutions of elliptic problems with right-hand side measures. *Rend. Mat. Appl. (7)*, 15(3) :321–337, 1995.
- [66] A. Prignet. Existence and uniqueness of “entropy” solutions of parabolic problems with  $L^1$  data. *Nonlinear Anal.*, 28(12) :1943–1954, 1997.
- [67] T. Roubíček. Thermo-visco-elasticity at small strains with  $L^1$ -data. *Quart. Appl. Math.*, 67(1) :47–71, 2009.
- [68] J. Serrin. Pathological solutions of elliptic differential equations. *Ann. Scuola Norm. Sup. Pisa (3)*, 18 :385–387, 1964.
- [69] J. Simon. Compact sets in the space  $L^p(0, T; B)$ . *Ann. Mat. Pura Appl. (4)*, 146 :65–96, 1987.
- [70] G. Stampacchia. Le problème de Dirichlet pour les équations elliptiques du second ordre à coefficients discontinus. *Ann. Inst. Fourier (Grenoble)*, 15(fasc. 1) :189–258, 1965.
- [71] C. Vaz and E. Fernández-Cara. Renormalized solutions to a system of type Navier-Stokes. *J. Math. Anal. Appl.*, 378(2) :442–449, 2011.

