Finite Volume Scheme and Renormalized Solutions for an Elliptic Operator with Discontinuous Matrix Diffusion Coefficients

Arouna Ouédraogo and Wendlassida Basile Yaméogo,

***Abstract*—We are interested in this paper to show that the approximate solution, by the finite volumes method, converges to the renormalized solution of an elliptic operator with discontinuous matrix diffusion coefficients and L^1-data. By adapting the strategy developed in the finite volume method, we show that the approximate solution converges to the unique renormalized solution.**

***Index Terms*—Discontinuous matrix diffusion, L^1-data, Renormalized Solutions, Finite Volume schemes.**

I. Introduction

We are interested here in the discretization of an elliptic operator with discontinuous matrix diffusion coefficients, which may appear in real case problems such as electrical or thermal transfer problems or, more generally, diffusion problems in heterogeneous media. Let's Ω be an open bounded polygonal subset of $\mathbb{R}^d$, $d \geq 2$, with boundary $\partial\Omega$, and consider the problem

$$\begin{cases} -\text{div}(A\nabla u) + \ \text{div}(vu) + bu & = f \qquad \text{in } \Omega, \\ \qquad\qquad u & = 0 \quad \text{on } \partial\Omega, \end{cases} \tag{1}$$

with the following assumptions on the data (one denotes by $\mathbb{R}^{d\times d}$ the set of $d \times d$ matrices with real coefficients):

(H_1) A is a bounded measurable function from Ω to $\mathbb{R}^{d\times d}$ such that for any $x \in \Omega$, $A(x)$ is symmetric, and that there exists α and $\beta \in \mathbb{R}_+^*$ such that

$$\alpha\xi\cdot\xi \leq A(x)\xi\cdot\xi \leq \beta\xi\cdot\xi \text{ for any } x \in \Omega \text{ and any } \xi \in \mathbb{R}^d.$$

(H_2) $v \in (L^p(\Omega))^d$, $2 < p < +\infty$ if $d = 2$, $p = d$, if $d \geq 3$.

(H_3) $b \in L^2(\Omega)$ is a positive function.

(H_4) $f \in L^1(\Omega)$.

In the sequel, we use the notation Avw for the scalar product of the vector Av by the vector w (which is often denoted by ${}^t w \cdot Av$).

The main difficulties in dealing with the existence and the uniqueness of a solution to problem (1) are due to the discontinuous character of the matrix and the L^1-data.

The theory of renormalized solutions has been introduced in [6] for Boltzmann equations and has been adapted in [13], [14] for elliptic problems with L^1-data. It is well 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 discretization methods, several techniques are developed in [9], [16] and [17]. The convergence of the cell-centered finite volume scheme for equation (1) has been studied in [10] when $v = 0$ and with measure data. In [9] the authors consider a bounded piecewise continuous function f with Non-homogeneous Dirichlet boundary conditions: They prove that the solution of this scheme for equation (1) converges to the unique variational solution $u \in H^1(\Omega)$ of (1).

In [12] the author studied problem (1) with Δu instead of -div$(A\nabla u)$.

Recall that a renormalized solution of (1) is a measurable function u defined from Ω to $\overline{\mathbb{R}}$, such that u is finite a.e. in Ω and

$$\forall k > 0, T_k(u) \in H_0^1(\Omega), \tag{2}$$

$$\lim_{k\to+\infty} \frac{1}{k}\int_\Omega |\nabla T_k(u)|^2 dx = 0, \tag{3}$$

$$\begin{cases} \forall h \in C_c^1(\mathbb{R}), \forall\psi \in H_0^1(\Omega)\cap L^\infty(\Omega), \\ \displaystyle\int_\Omega A\nabla T_k(u)\,\nabla\psi\, h(u)\,dx \\ \quad \displaystyle+ \int_\Omega h'(u)\, A\nabla T_k(u)\,\nabla T_k(u)\,\psi\,dx \\ \quad\displaystyle - \int_\Omega u\,h(u)\,v\nabla\psi\,dx - \int_\Omega u\,h'(u)\,\psi\,v\,\nabla u\,dx \\ \qquad \displaystyle+ \int_\Omega b\,u\,h(u)\,\psi\,dx = \int_\Omega \psi\,h(u)\,f\,dx, \end{cases} \tag{4}$$

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

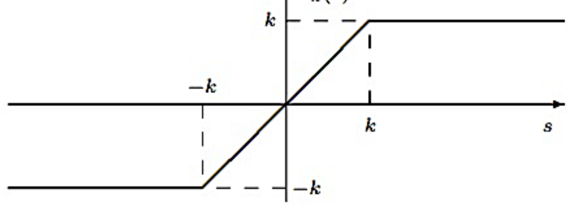

Fig. 1. **The function T_k**

Since h has a compact support, each term of (4) is well defined.

Manuscript received June 1, 2022; revised December 28, 2022.

A. Ouédraogo is a professor of the Department of Mathematics, Université Norbert Zongo, Koudougou, BP 316, Burkina Faso (e-mail: arouna.ouedraogo@unz.bf).

W. B. Yaméogo is a PhD student in the Mathematics, Computer Science and Applications Laboratory, Université Norbert Zongo, Koudougou, BP 316, Burkina Faso (e-mail: wendlassbas@gmail.com).

The existence and the uniqueness of a renormalized solution to (1) for L^1- data is proved in [14] by F. Murat.
In the present paper, 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 [4], [5], [13], [14], [15]).
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 $\varphi h(u)$ as a test function in the finite volume scheme. The first difficulty is to establish a discrete version of the estimate on the energy (3). Moreover it is worth noting that in (4) 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. The method developed in [12] allows us to deal with nonlinear version of (1) in the sense that the solution of the discrete scheme converges to the unique renormalized solution (see Section 4).
The rest of the paper is organized as follows. In Section 2, we present the finite volume scheme and the properties of the discrete gradient. In Section 3, we prove existence and uniqueness of the solution to the schemes. Section 4 is devoted to prove several estimates, especially the discrete equivalent to (4) which is crucial to pass to the limit in the finite volume scheme. In Section 5, we prove the convergence of the cell-centered finite volume scheme via a density argument.

II. Finite volume scheme

As in [9], let us define the admissibility mesh in the present work.

Definition 2.1: (**Admissible meshes**)
Let Ω be an open bounded polygonal subset of $\mathbb{R}^d$, $d = 2$, or 3. An admissible finite volume mesh of Ω, denoted by $\mathcal{T}$, is given by a family of "control volumes", which are open polygonal convex subsets of Ω, a family of subsets of $\overline{\Omega}$ contained in hyperplanes of $\mathbb{R}^d$, denoted by $\mathcal{E}$ (these are the edges (two-dimensional) or sides (three-dimensional) of the control volumes), with strictly positive $(d-1)$-dimensional measure, and a family of points of Ω denoted by $\mathcal{P}$ satisfying the following properties (in fact, we shall denote by $\mathcal{T}$ the family of control volumes):

(i) The closure of the union of all the control volumes is $\overline{\Omega}$.

(ii) For any $K \in \mathcal{T}$, there exists a subset $\mathcal{T}(K)$ of $\mathcal{T}$ denoted $\mathcal{E}(K) = \{\sigma \in \mathcal{E}; \sigma \in \partial K\}$, such that $\partial K = \overline{K} \setminus K = \cup_{\sigma \in \mathcal{E}(K)} \overline{\sigma}$ for all $K \in \mathcal{T}$. Furthermore, $\mathcal{E} = \cup_{K \in \mathcal{T}} \mathcal{E}(K)$.

(iii) For any $(K, L) \in \mathcal{T}^2$ with $K \neq L$, either the $(d-1)$-dimensional Lebesgue measure of $\overline{K} \cap \overline{L}$ is 0 or $\overline{K} \cap \overline{L} = \overline{\sigma}$ for some $\sigma \in \mathcal{E}$, which will then be denoted by $\sigma = K/L$.

(iv) For all $K \in \mathcal{T}$, x_K is in the interior of K.

(v) For any $K \in \mathcal{T}$, let A_K denote the mean value of A on K, that is

$$A_K = \frac{1}{|K|} \int_K A(x)\, dx.$$

There exists a family of points

$$\mathcal{P} = (x_K)_{K \in \mathcal{T}} \text{ such that } x_K = \cup_{\sigma \in \mathcal{E}(K)} \mathcal{D}_{K,\sigma},$$

where $\mathcal{D}_{K,\sigma}$ is a straight line perpendicular to σ with respect to the scalar product induced by A_K^{-1} such that $\mathcal{D}_{K,\sigma} \cap \sigma = \mathcal{D}_{L,\sigma} \cap \sigma \neq \emptyset$ if $\sigma = K/L$. Furthermore, if $\sigma = K/L$, let $y_\sigma = \mathcal{D}_{K,\sigma} \cap \sigma (= \mathcal{D}_{L,\sigma} \cap \sigma)$ and assume that $x_K \neq x_L$.

(vi) For any $\sigma \in \mathcal{E}_{ext}$, let K be the control volume such that $\sigma \in \mathcal{E}(K)$ and let $\mathcal{D}_{K,\sigma}$ be the straight line going through x_K and orthogonal to σ with respect to the scalar product induced by A_K^{-1}; then, there exists $y_\sigma \in \mathcal{D}_{K,\sigma} \cap \sigma$.

In the sequel, the following notations are used.
The mesh size is defined by: $h_{\mathcal{T}} = \sup_{K \in \mathcal{T}} diam(K)$.
For any $K \in \mathcal{T}$ and $\sigma \in \mathcal{E}$, we denote by $|K|$ the d-dimensional Lebesgue measure of K (it is the area of K in the two-dimensional case and the volume in the three-dimensional case) and $|\sigma|$ the $(d-1)$-dimensional measure of σ.
The unit normal to $\sigma \in \mathcal{E}(K)$ outward to K is denoted by $\eta_{K,\sigma}$.
The set of interior (resp. boundary) edges is denoted by $\mathcal{E}_{int}$ (resp. $\mathcal{E}_{ext}$), that is $\mathcal{E}_{int} = \{\sigma \in \mathcal{E};\ \sigma \not\subset \partial\Omega\}$ (resp. $\mathcal{E}_{ext} = \{\sigma \in \mathcal{E};\ \sigma \subset \partial\Omega\}$).
The set of neighbours of K is denoted by $\mathcal{N}(K)$, that is $\mathcal{N}(K) = \{L \in \mathcal{T}; \exists \sigma \in \mathcal{E}(K), \overline{\sigma} = \overline{K} \cap \overline{L}\}$.
For any $K \in \mathcal{T}$ and $\sigma \in \mathcal{E}$, we denote by $d_{K,\sigma}$ the Euclidean distance between x_K and σ.
For any $\sigma \in \mathcal{E}$, d_σ is defined by $d_\sigma = d_{K,\sigma} + d_{L,\sigma}$, if $\sigma = K/L \in \mathcal{E}_{int}$ (in which case d_σ is the Euclidean distance between x_K and x_L) and $d_\sigma = d_{K.\sigma}$, if $\sigma \in \mathcal{E}_{ext} \cap \mathcal{E}(K)$.
For any $\sigma \in \mathcal{E}$, the "transmissibility" through σ is defined by $\tau_\sigma = \dfrac{|\sigma|}{d_\sigma}$ if $d_\sigma \neq 0$.

We will need discrete Sobolev inequalities (see Lemma 2.9), which depend on the constant ζ appearing in the following assumption.

$$\exists \zeta \text{ such that } \forall K \in \mathcal{T},\ \forall \sigma \in \mathcal{E}_K,\ d_{K.\sigma} \geq \zeta d_\sigma. \tag{5}$$

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 [4]).
Let us define the discrete $W_0^{1,q}$ norm.
Definition 2.2: (discrete $W_0^{1,q}$ norm) Let Ω 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 Ω 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

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

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

Definition 2.3: The set of measurable functions of $L^p(\Omega)$ which admit some small derivative up to $p \in \mathbb{N}$ is denoted by $H^p(\Omega)$.
The sub-set vectorial space $H^p(\Omega)$ of the functions at compact support in Ω is denoted by $H_0^p(\Omega) \subset H^p(\Omega)$.

Definition 2.4: The set of measurable functions of $L^p(\Omega)$ whose all the derived are also in $L^P(\Omega)$ up to the total derivation order of $q \in \mathbb{N}$ is denoted by $W^{q,p}(\Omega)$.
For $1 \leq p \leq \infty$, a norm in $W^{q,p}(\Omega)$ is

$$\| u \|_{W^{q,p}(\Omega)} = \sum_{k_1+\cdots+k_q \leq q} \| u^{(k_1+\cdots+k_q)} \|_{L^p(\Omega)} .$$

The sub-set vectorial space of $W^{q,p}(\Omega)$ consisting of the functions at compact support in Ω is denoted $W_0^{q,p}(\Omega) \subset W^{q,p}(\Omega)$.

Theorem 2.5: (of Rellich) [14] Let Ω be an open set of $\mathbb{R}$ that we suppose boundary and with border regular enough. Then, the injection $H_0^p(\Omega) \hookrightarrow L^p(\Omega)$ is compact.

Before writing the finite volume scheme, let us define a discrete finite volume gradient (see [12])).

Definition 2.6: (Discrete finite volume gradient) For all $K \in \mathcal{T}$ and for all $\sigma \in \mathcal{E}(K)$, we define the volume $D_{K,\sigma}$ as the cone of basis σ and of opposite vertex x_K. Then, we define the "diamond-cell" D_σ by:

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

For any $v_\mathcal{T} \in X(\mathcal{T}_m)$ and notice that $|D_\sigma| = \dfrac{|\sigma|\, d_\sigma}{d}$,
we define the discrete gradient $\nabla_\mathcal{T} v_\mathcal{T}$ by: for all x in D_σ,

$$\forall \sigma \in \mathcal{E}_{int}, \sigma = K/L, \nabla_\mathcal{T} v(x) = d\frac{v_L - v_K}{d_\sigma}\eta_{K,\sigma},$$

$$\forall \sigma \in \mathcal{E}_{ext} \cap \mathcal{E}(K), \nabla_\mathcal{T} v(x) = d\frac{0 - v_K}{d_\sigma}\eta_{K,\sigma}.$$

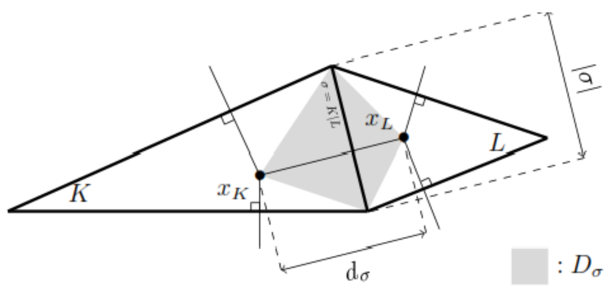

Fig. 2. **Example of control volume for the method of finite volume in two dimensions of space**

Proposition 2.7: (Discrete Poincaré inequality) Let $\mathcal{T}$ be an admissible mesh and $v_\mathcal{T} \in X(\mathcal{T})$. Then, if $1 \leq q \leq 2$, $\| v_\mathcal{T} \|_{L^q(\Omega)} \leq diam(\Omega) \| \nabla v_\mathcal{T} \|_{1,q,\mathcal{T}}$.

Proposition 2.8: (Discrete Sobolev inequality) Let $1 \leq q \leq 2$, $\mathcal{T}$ be an admissible mesh and $\zeta > 0$ satisfying for all $K \in \mathcal{T}$ and all $\sigma \in \mathcal{E}(K)$, $d_{K,\sigma} \geq \zeta d_\sigma$, Then, with $q* = \dfrac{dq}{d-q}$ if $q < d$ and $q* < \infty$ if $q = d = 2$, there exists $C > 0$ only depending on $(\Omega, q, q*,)$ such that, for all $v_\mathcal{T} \in X(\mathcal{T})$,we have $\| v_\mathcal{T} \|_{L^{q*}(\Omega)} \leq C \| v_\mathcal{T} \|_{1,q,\mathcal{T}}$.

Lemma 2.9: (Weak convergence of the finite volume gradient) Let $(\mathcal{T}_m)_{m\geq 1}$ be a sequence of admissible meshes such that there exists $\zeta > 0$ satisfying for all $m > 1$, for all $K \in \mathcal{T}$ and for all $\sigma \in \mathcal{E}(K)$, $d_{K,\sigma} \geq \zeta d_\sigma$, and such that $h_{\mathcal{T}_m} \longrightarrow 0$. Let $v_{\mathcal{T}_m} \in X(\mathcal{T}_m)$ and let us assume that there exists $C > 0$ such that $\| v_{\mathcal{T}_m} \|_{1,\mathcal{T}_m} \leq C$, and that $v_{\mathcal{T}_m}$ converges in $L^1(\Omega)$ to $v \in H^1(\Omega)$. then $\nabla \mathcal{T}_m v_{\mathcal{T}_m}$ converges to ∇v weakly in $L^2(\Omega)^d$.

Let $\mathcal{T}$ be an admissible mesh, we can define the finite volume discretization of (1). For $K \in \mathcal{T}$, we define

$$b_K = \frac{1}{|K|}\int_K b\, dx, \tag{6}$$

$$v_{K,\sigma} = \frac{1}{|D_\sigma|}\int_{D_\sigma} v.\eta_{K,\sigma} dx \tag{7}$$

and

$$f_K = \frac{1}{|K|}\int_K f\, dx. \tag{8}$$

Let $(u_K)_{K\in\mathcal{T}}$ denote the discrete unknowns, which aim to be approximations of the values $u(x_K)$, for all $K \in \mathcal{T}$. In order to describe the scheme in the most general way, one introduces some auxiliary unknowns, namely the fluxes $F_{K,\sigma}$, for all $K \in \mathcal{T}$ and $\sigma \in \mathcal{E}(K)$, and some (expected) approximation of u in σ, denoted by u_σ, for all $\sigma \in \mathcal{E}$. These auxiliary unknowns are helpful to write the scheme, but they can be eliminated locally so that the discrete equations will only be written with respect to the primary unknowns $(u_K)_{K\in\mathcal{T}}$. The finite volume scheme for the numerical approximation of the solution to Problem (1) is obtained by integrating it over each control volume K, and approximating the fluxes over each edge σ of K. This yields for all $K \in \mathcal{T}$

$$\begin{aligned} \sum_{\sigma\in\mathcal{E}(K)} F_{K,\sigma} + \sum_{\sigma\in\mathcal{E}(K)} |\sigma| v_{K,\sigma}\, u_{\sigma,+} + |K| b_K\, u_K \\ = |K| f_K, \end{aligned} \tag{9}$$

with

$$\forall \sigma = K/L \in \mathcal{E}_{int}, \begin{cases} u_{\sigma,+} = u_K \text{ if } v_{K,\sigma} \geq 0, \\ u_{\sigma,+} = u_L, \text{ otherwise} \end{cases} \tag{10}$$

and

$$\forall \sigma \in \mathcal{E}_{ext} \cap \mathcal{E}(K), \begin{cases} u_{\sigma,+} = u_K \text{ if } v_{K,\sigma} \geq 0, \\ u_{\sigma,+} = 0, \text{ otherwise.} \end{cases} \tag{11}$$

We denote by $u_{\sigma,-}$ the downstream choice of u which is such that $\{u_{\sigma,+}, u_{\sigma,-}\} = \{u_K, u_L\}$ (with $u_L = 0$ if $\sigma \in \mathcal{E}_{ext} \cap \mathcal{E}(K)$).
$F_{K,\sigma}$ is an approximation of $\displaystyle\int_\sigma -A_K \nabla u(x) \cdot \eta_{K,\sigma}\, d\gamma(x)$; the approximation $F_{K,\sigma}$ is written with respect to the discrete unknowns $(u_K)_{K\in\mathcal{T}}$ and $(u_\sigma)_{\sigma\in\mathcal{E}}$. For $K \in \mathcal{T}$ and $\sigma \in \mathcal{E}(K)$, let $\lambda_{K,\sigma} = |A_K\, \eta_{K,\sigma}|$. A natural expression for $F_{K,\sigma}$ is then

$$F_{K,\sigma} = -|\sigma|\lambda_{K,\sigma}\frac{u_\sigma - u_K}{d_{K,\sigma}}.$$

Writing the conservativity of the scheme, i.e. $F_{L,\sigma} = -F_{K,\sigma}$ if $\sigma = K/L \subset \Omega$, yields the value of u_σ, with respect to $(u_K)_{K\in\mathcal{T}}$;

$$u_\sigma = \frac{1}{\frac{\lambda_{K,\sigma}}{d_{K,\sigma}} + \frac{\lambda_{L,\sigma}}{d_{L,\sigma}}}\Big(\frac{\lambda_{K,\sigma}}{d_{K,\sigma}}u_K + \frac{\lambda_{L,\sigma}}{d_{L,\sigma}}u_L\Big).$$

Hence the value of $F_{K,\sigma}$;

$$F_{K,\sigma} = -\tau_\sigma(u_L - u_K), \text{ if } \sigma \in \mathcal{E}(K), \tag{12}$$

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

$$\tau_\sigma = \begin{cases} |\sigma| \dfrac{\lambda_{K,\sigma}\,\lambda_{L,\sigma}}{\lambda_{K,\sigma}\,d_{L,\sigma} + \lambda_{L,\sigma}\,d_{K,\sigma}} & \text{if } \sigma = K/L \in \mathcal{E}_{int}, \\ |\sigma| \dfrac{\lambda_{K,\sigma}}{d_{K,\sigma}} & \text{if } \sigma \in \mathcal{E}_{ext} \cap \mathcal{E}(K). \end{cases}$$

Now we are led to give our main results.

Theorem 2.10: (Existence of a solution for the scheme on (1)) Let $\mathcal{T}$ be an admissible mesh of Ω. Then, there exists a unique solution $u_\mathcal{T} = (u_K)_{K\in\mathcal{T}}$ to (9)-(12).

Theorem 2.11: (Convergence of the solution for the scheme on (1)) If $(\mathcal{T}_m)_{m\geq 1}$ is a sequence of admissible meshes such that there exists $\zeta > 0$ satisfying for all $m \geq 1$, for all $K \in \mathcal{T}$ and all $\sigma \in \mathcal{E}(K)$, $d_{K,\sigma} \geq \xi d_\sigma$, and such that $h_{\mathcal{T}_m} \to 0$, then if $u_{\mathcal{T}_m} = (u_K)_{K\in\mathcal{T}_m}$ is the solution to (9)-(12), 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 $H_0^1(\Omega)$, when u is the unique renormalized solution of (1).

III. Existence and uniqueness of the solution to the schemes

Proof: (**Proof of Theorem 2.10**)

We intend here to prove Theorem 2.10 by means of linear algebra tools. We set

$$\mathcal{F}_{K,\sigma} = \tau_\sigma(u_K - u_L) + |\sigma|(v^+_{K,\sigma}\,u_K - v^-_{K,\sigma}\,u_L)\forall\sigma \in \mathcal{E}(K),$$

where $s^+ = \max(s, 0)$ and $s^- = \max(-s, 0)$ are the positive and negative parts of a real number s. The quantity $|\sigma|(v^+_{K,\sigma}\,u_K - v^-_{K,\sigma}\,u_L)$ is a upwind discretization approximing the convective flux $\displaystyle\int_\sigma u\,v\cdot\eta_{K,\sigma}$, which stabilizes the scheme (at the cost of the introduction of an additional numerical diffusion).

Defining $\mathcal{B}(s) = 1 + (-s)^+ = 1 + s^-$, then, using the fact that $\tau_\sigma = \dfrac{|\sigma|}{d_\sigma}$, $\mathcal{F}_{K,\sigma}$ can be written as

$$\begin{aligned} \mathcal{F}_{K,\sigma} &= \tau_\sigma\Big(\mathcal{B}\Big(-v_{K,\sigma}\frac{|\sigma|}{\tau_\sigma}\Big)\,u_K - \mathcal{B}\Big(v_{K,\sigma}\frac{|\sigma|}{\tau_\sigma}\Big)\,u_L\Big) \\ &= \frac{|\sigma|}{d_\sigma}\big(\mathcal{B}(-v_{K,\sigma}\,d_\sigma)\,u_K - \mathcal{B}(v_{K,\sigma}\,d_\sigma)\,u_L\big) \end{aligned} \tag{13}$$

As in [4], we note that the function $\mathcal{B}$ satisfies the following:

$$\mathcal{B} \text{ is Lipschitz-continuous on } \mathbb{R}, \tag{14}$$

$$\mathcal{B}(0) = 1 \text{ and } \mathcal{B}(s) > 0 \quad \forall\, s \in \mathbb{R}, \tag{15}$$

$$\mathcal{B}(s) - \mathcal{B}(-s) = -s \quad \forall\, s \in \mathbb{R}. \tag{16}$$

The scheme (9)-(12) leads to a linear system of equations that can be written as

$$(\mathbb{A} + B\mathbb{D})U = F, \tag{17}$$

where $U = (u_K)_{K\in\mathcal{T}}$, $B = (|K|b_K)_{K\in\mathcal{T}}$, $F = (|K|f_K)_{K\in\mathcal{T}}$, $\mathbb{D}$ is the diagonal matrix whose diagonal entries are $\mathbb{D}_{K,K} = |K|$ and $\mathbb{A}$ is the square matrix of size $Card(\mathcal{T}) \times Card(\mathcal{T})$ with entries

$$\mathbb{A}_{K,K} = \sum_{\sigma\in\mathcal{E}(K)} \frac{|\sigma|}{d_\sigma}\mathcal{B}\big(-v_{K,\sigma}d_\sigma\big) \quad \forall K \in \mathcal{T}, \tag{18}$$

$$\mathbb{A}_{K,L} = -\frac{|\sigma|}{d_\sigma}\mathcal{B}\big(v_{K,\sigma}\,d_\sigma\big)\forall\, K \in \mathcal{T}, L \in \mathcal{N}(K), \tag{19}$$

$$\mathbb{A}_{K,L} = 0 \quad \forall\, K \in \mathcal{T}, \forall\, L \notin \mathcal{N}(K). \tag{20}$$

Theorem 2.10 is thus a logical consequence of the following proposition.

Proposition 3.1: For all $b \in L^2(\Omega)$ a nonnegative function, the diagonal coefficients of the matrix $\mathbb{A}_b = \mathbb{A} + B\mathbb{D}$ are positive, the extra-diagonal coefficients of $\mathbb{A}_b$ are nonpositive and the sum of the coefficients in each column of $\mathbb{A}_b$ is positive. Therefore $\mathbb{A}_b$ is an M-matrix and is invertible.

Proof: (**Proof of Proposition 3.1**) Due to (18)-(20) and (15) we note that all of the diagonal entries of $\mathbb{A}$ are strictly positive, whereas the extra-diagonal coefficients are nonpositive. This is therefore also the case for $\mathbb{A}_b$.

Moreover, since $v_{L,\sigma} = -v_{K,\sigma}$ whenever $\sigma = K/L \in \mathcal{E}_{int}$, we have

$$\mathbb{A}_{K,K} = -\sum_{L\in\mathcal{N}(K)} \mathbb{A}_{L,K} \quad \forall\, K \in \mathcal{T}. \tag{21}$$

In other words, in each column the diagonal term is the opposite of the sum of the extra-diagonal terms. This has the following consequence:

the sum of the coefficients in the column K of $\mathbb{A}_b$ is equal to $B|K|$, and Proposition 3.1 is proved. ■

The proof of Theorem 2.10 is then complete. ■

IV. Estimations

In this section, we first establish in Proposition 4.1 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 4.3 an estimate on $T_n(u_\mathcal{T})$ and the convergence of $T_n(u_\mathcal{T})$ to $T_n(u)$. Finally, we prove in Proposition 4.4 a discrete version of the decay of the energy.

Proposition 4.1: (see [12]) Let $\mathcal{T}$ be an admissible mesh. If $u_\mathcal{T} = (u_K)_{K\in\mathcal{T}}$ is a solution to (9)-(12), then

$$\|\ln(1 + |u_\mathcal{T}|)\|^2_{1,\mathcal{T}} \leq 2\|f\|_{L^1(\Omega)} + d|\Omega|^{\frac{p-2}{p}}\|v\|^2_{L^p(\Omega)^d}, \tag{22}$$

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

Let us state a corollary, which is used in the proof of the estimate of Proposition 4.3.

Corollary 4.2: Let $\mathcal{T}$ be an admissible mesh.

If $u_\mathcal{T} = (u_K)_{K\in\mathcal{T}}$ is a solution to (9)-(12) and, for $n > 0$, $E_n = \{|u_\mathcal{T}| > n\}$, then there exists $C > 0$ only depending on (Ω, v, f, d, p) such that

$$|E_n| \leq \frac{C(1 + \|f\|_{L^1(\Omega)})}{(\ln(1 + n))^2}. \tag{23}$$

The following proposition checks that our schemes satisfy properties that are well known for the continuous equations, namely, if b is positive then it is easy to obtain a priori estimates for the solution to (1) and the L^2-norm of the gradient of the solution to (1) is always controlled by the L^2-norm of the solution. Of course, the main difference with respect to the continuous case is that, at the discrete level, we have to make sure that these estimates do not depend on the size of the mesh.

Proposition 4.3: (Estimation 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 (9)-(12), then for all $n \in \mathbb{N}^*$,

$$\|T_n(u_{\mathcal{T}})\|_{1,\mathcal{T}} \leq 2n\|f\|_{L^1(\Omega)} + n^2 d\|v\|^2_{L^2(\Omega)^d}. \tag{24}$$

Moreover, if $(\mathcal{T}_m)_{m\geq 1}$ is a sequence of admissible meshes such that there exists $\zeta > 0$ satisfying for all $m \leq 1$, for all $K \in \mathcal{T}$ and for all $\sigma \in \mathcal{E}(K)$, $d_{K\sigma} \geqslant \zeta d_\sigma$, there exists a measurable function u finite a.e. in Ω such that, up to a subsequence $T_n(u_{\mathcal{T}_m})$ converges to $T_n(u)$ weakly in $H^1(\Omega)$, strongly in $L^2(\Omega)$ and a.e. in Ω.

Proof: The proof is divided into two steps. In Step 1 we derive the estimate (24) on the truncate on $u_{\mathcal{T}}$. The step 2 is devoted to extract a Cauchy sub-sequences in measure.

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

We multiply equation (9) by $T_n(u_K)$ and sum over $K \in \mathcal{T}$. Due to the conservativity of the fluxes and to (13), gathering by edges, we find that

$$\sum_{\sigma\in\mathcal{E}} \frac{|\sigma|}{d_\sigma}\big(\mathcal{B}(-v_{K,\sigma}\, d_\sigma)\, u_K - \mathcal{B}(v_{K,\sigma}\, d_\sigma)\, u_L\big) \times$$
$$\times\big(T_n(u_K) - T_n(u_L)\big) + \sum_{K\in\mathcal{T}} |K| b_K u_K T_n(u_K)$$
$$= \sum_{K\in\mathcal{T}} \int_K f T_n(u_K) dx. \tag{25}$$

Since b is nonnegative and since $rT_n(r) \geqslant 0 \quad \forall r$, we notice that the second term in the left hand side of (25) is nonnegative. Moreover, since T_n is bounded by n, we deduce that

$$\left|\sum_{K\in\mathcal{T}} \int_K f T_n(u_K) dx\right| \leqslant n\|f\|_{L^1(\Omega)}.$$

Using (16), the first term of (25) can be rewritten as

$$\sum_{\sigma\in\mathcal{E}} \frac{|\sigma|}{d_\sigma}\big(\mathcal{B}(-v_{K,\sigma}\, d_\sigma)\, u_K - \mathcal{B}(v_{K,\sigma}\, d_\sigma)\, u_L\big)\times$$
$$\times\big(T_n(u_K) - T_n(u_L)\big)$$
$$= \sum_{\sigma\in\mathcal{E}} \frac{|\sigma|}{d_\sigma}\, B\big(v_{K,\sigma}\, d_\sigma\big)(u_K - u_L)(T_n(u_K) - T_n(u_L))$$
$$+ \sum_{\sigma\in\mathcal{E}} |\sigma|\, v_{K,\sigma}\, u_K\big(T_n(u_K) - T_n(u_L)\big)$$
$$= I_1 + I_2,$$

with

$$I_1 = \sum_{\sigma\in\mathcal{E}} \frac{|\sigma|}{d_\sigma}\, B\big(v_{K,\sigma}\, d_\sigma\big)(u_K - u_L)(T_n(u_K) - T_n(u_L)),$$
$$I_2 = \sum_{\sigma\in\mathcal{E}} |\sigma|\, v_{K,\sigma}\, u_K\big(T_n(u_K) - T_n(u_L)\big).$$

Therefore, we deduce from (25)

$$I_1 \leqslant n\|f\|_{L^1(\Omega)} - I_2. \tag{26}$$

As in [7], I_2 can be rewritten as

$$-I_2 = \sum_{\sigma\in\mathcal{E}} |\sigma||v_{K,\sigma}|u_{\sigma,+}(T_n(u_{\sigma,-}) - T_n(u_{\sigma,+})).$$

As in [7], we define the subset $\mathcal{A}$ of edges by

$$\begin{aligned}\mathcal{A} \;&=\; \{\sigma \in \mathcal{E}; u_{\sigma,+} \geqslant u_{\sigma,-}, u_{\sigma,+} < 0\} \\ &\quad\cup\{\sigma \in \mathcal{E}; u_{\sigma,+} < u_{\sigma,-}, u_{\sigma,+} \geq 0\}\end{aligned} \tag{27}$$

and since T_n is non decreasing we have

$$-I_2 = \sum_{\sigma\in\mathcal{E}} |\sigma||v_{K,\sigma}|u_{\sigma,+}(T_n(u_{\sigma,-}) - T_n(u_{\sigma,+})).$$

Notice that $\forall\sigma \in \mathcal{A}, |u_{\sigma,+}| \geq n$ implies $|u_{\sigma,-}| \geq n$. So, we deduce that for all $\sigma \in \mathcal{A}$,

$$u_{\sigma,+}(T_n(u_{\sigma,-}) - T_n(u_{\sigma,+})) = T_n(u_{\sigma,+})(T_n(u_{\sigma,-}) - T_n(u_{\sigma,+})).$$

It follows that

$$\begin{aligned}-I_2 \;&\leq\; \sum_{\sigma\in\mathcal{A}} |\sigma||v_{K,\sigma}|u_{\sigma,+}T_n(u_{\sigma,+})(T_n(u_{\sigma,-}) - T_n(u_{\sigma,+}))\\
&\leq\; nd^{\frac{1}{2}}\|v\|_{L^2(\Omega)^d}\Big(\sum_{\sigma\in\mathcal{A}} \frac{|\sigma|}{d_\sigma}(T_n(u_{\sigma,-}) - T_n(u_{\sigma,+}))^2\Big)^{\frac{1}{2}}\\
&\leq\; \frac{1}{2}n^2 d^{\frac{1}{2}}\|v\|^2_{L^2(\Omega)^d} + \frac{1}{2}\sum_{\sigma\in\mathcal{A}} \frac{|\sigma|}{d_\sigma}(T_n(u_{\sigma,-}) - T_n(u_{\sigma,+}))^2\\
&\leq\; \frac{1}{2}n^2 d^{\frac{1}{2}}\|v\|^2_{L^2(\Omega)^d} + \frac{1}{2}\sum_{\sigma\in\mathcal{E}} \frac{|\sigma|}{d_\sigma}(T_n(u_K) - T_n(u_L))^2.\end{aligned}$$

Since $|T_n(u_K) - T_n(u_L)| \leq |u_K - u_L|$ (because T_n is 1-Lipschitz function), we have

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

We deduce from (26) and the fact that $\mathcal{B}(s) \geq 1,\ \forall s \in \mathbb{R}$,

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

Therefore, using again the fact that T_n is 1-Lipschitz, we can write :

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

which yields (24).

Applying Lemma 2.9 and the diagonal process, up to a subsequence still denoted by $\mathcal{T}_m$, for any $n \geq 1$, there exists v_n in $H^1(\Omega)$ such that $T_n(u_{\mathcal{T}_m}) \longrightarrow v_n$ and $T_n(u_{\mathcal{T}_m}) \rightharpoonup v_n$ in the finite volume gradient sense.

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

In this step, we follow a proof of [5] to show that $u_{\mathcal{T}_m}$ converges a.e. to u . For all $n > 0$ and all sequences $(\mathcal{T}_m)_{m\geq 1}$ and $(\mathcal{T}_p)_{p\geq 1}$ of admissible meshes, we have

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

Let $\varepsilon > 0$ fixed. By (14), 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

$$\forall h_{\mathcal{T}_m}, h_{\mathcal{T}_p} < h_0,\ \text{meas}(\{|T_n(u_{\mathcal{T}_m}) - T_n(u_{\mathcal{T}_p})| > n\}) < \frac{\varepsilon}{2}.$$

Therefore, we deduce that

$$\forall h_{\mathcal{T}_m}, h_{\mathcal{T}_p} < h_0,\ \text{meas}(\{|u_{\mathcal{T}_m} - u_{\mathcal{T}_p}| > n\}) < \varepsilon.$$

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 $u_{\mathcal{T}_m} \longrightarrow u$ a.e. in Ω. Due to Corollary 4.2, u is finite a.e. in Ω. Moreover from convergences obtained in Step 1 we get that

$$T_n(u) \in H^1(\Omega), \nabla_{\mathcal{T}_m} T_n(u_{\mathcal{T}}) \to \nabla T_n(u) \text{ in } (L^2(\Omega))^d. \tag{28}$$

■

In the following proposition we prove a uniform estimate on the truncated energy of $u_{\mathcal{T}}$ (see (29) below) which is crucial to pass to the limit in the approximate problem. We explicitly observe that (29) is the discrete version of (3) which is imposed in the definition of the renormalized solution for elliptic equation with L^1-data. As in the continuous case (29) is related to the regularity of f : $f \in L^1(\Omega)$ and does not charge any zero-Lebesgue set. If we replace $div(vu)$, we also have to uniformly control the discrete version of $\frac{1}{n}\int_\Omega vu\nabla T_n(u)dx$ which is stated in (30).

Proposition 4.4: **(Discrete estimate on the energy)** Let $(\mathcal{T}_m)m \geq 1$ be a sequence of admissible meshes such that there exists $\zeta > 0$ satisfying $\forall m \geqslant 1, \forall K \in \mathcal{T}$ and $\forall \sigma \in \mathcal{E}(K), d_{K,\sigma} \geq \zeta d_\sigma$.
If $u_{\mathcal{T}_m} = (u_K)_{K\in\mathcal{T}_m}$ is a solution to (9)-(12), then

$$\begin{aligned}\lim_{n\to+\infty} \overline{\lim_{h_{\mathcal{T}_m}\to 0}} \frac{1}{n} \sum_{\sigma\in\mathcal{E}} \frac{|\sigma|}{d_\sigma}(u_K - u_L) \times \\ \times(T_n(u_K) - T_n(u_L)) = 0\end{aligned} \tag{29}$$

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

$$\begin{aligned}\lim_{n\to+\infty} \overline{\lim_{h_{\mathcal{T}_m}\to 0}} \frac{1}{n} \sum_{\sigma\in\mathcal{E}} |\sigma||v_{K,\sigma}||u_{\sigma,+}| \times \\ \times|(T_n(u_{\sigma,+}) - T_n(u_{\sigma,-})| = 0.\end{aligned} \tag{30}$$

Proof: We first establish (29). Let $\mathcal{T}$ be an admissible mesh and let $u_{\mathcal{T}}$ be a solution of (9)-(12). 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

$$\begin{aligned}&\frac{1}{n}\sum_{\sigma\in\mathcal{E}} \frac{|\sigma|}{d_\sigma}\big(\mathcal{B}(-v_{K,\sigma}\, d_\sigma)\, u_K - \mathcal{B}(v_{K,\sigma}\, d_\sigma)\, u_L\big) \times \\ &\times\big(T_n(u_K) - T_n(u_L)\big) + \frac{1}{n}\sum_{K\in\mathcal{T}} |K| b_K u_K T_n(u_K) \\ &= \frac{1}{n}\sum_{K\in\mathcal{T}} \int_K f T_n(u_K) dx.\end{aligned} \tag{31}$$

Since b is non-negative and since $rT_n(r) \geq 0\ \forall r$, we get

$$\frac{1}{n}\sum_{K\in\mathcal{T}} |K| b_K u_K T_n(u_K) \geq 0. \tag{32}$$

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

$$\frac{1}{n}\sum_{K\in\mathcal{T}} \int_K f T_n(u_K) dx = \int_\Omega f \frac{T_n(u_{\mathcal{T}})}{n} dx.$$

In view of the point-wise convergence of $u_{\mathcal{T}}$ to u, we obtain that $T_n(u_{\mathcal{T}})$ converges to $T_n(u)$ a.e. and weak * as $h_{\mathcal{T}} \to 0$. It follows that

$$\lim_{h_{\mathcal{T}}\to 0} \int_\Omega f \frac{T_n(u_{\mathcal{T}})}{n} dx = \int_\Omega f \frac{T_n(u)}{n} dx.$$

Since u is finite a.e. in Ω , $\frac{T_n(u)}{n}$ converges to 0 a.e. and in L^∞ weak, and since f belongs to $L^1(\Omega)$, the Lebesgue dominated convergence theorem implies that

$$\lim_{n\to+\infty} \lim_{h_{\mathcal{T}}\to 0} \frac{1}{n}\sum_{K\in\mathcal{T}} \int_K f T_n(u_K) dx = 0. \tag{33}$$

Using (16), the first term of (31) can be rewritten as

$$\begin{aligned}&\frac{1}{n}\sum_{\sigma\in\mathcal{E}} \frac{|\sigma|}{d_\sigma}\big(\mathcal{B}(-v_{K,\sigma}\, d_\sigma)\, u_K - \mathcal{B}(v_{K,\sigma}\, d_\sigma)\, u_L\big) \times \\ &\qquad \times\big(T_n(u_K) - T_n(u_L)\big) \\ &= \frac{1}{n}\sum_{\sigma\in\mathcal{E}} \frac{|\sigma|}{d_\sigma} B\big(v_{K,\sigma}\, d_\sigma\big)(u_K - u_L)(T_n(u_K) - T_n(u_L)) \\ &\quad + \frac{1}{n}\sum_{\sigma\in\mathcal{E}} |\sigma|\, v_{K,\sigma}\, u_K\big(T_n(u_K) - T_n(u_L)\big) \\ &= T_1 + T_2,\end{aligned}$$

with

$$T_1 = \frac{1}{n}\sum_{\sigma\in\mathcal{E}} \frac{|\sigma|}{d_\sigma} B\big(v_{K,\sigma}\, d_\sigma\big)(u_K - u_L)(T_n(u_K) - T_n(u_L)),$$

$$T_2 = \frac{1}{n}\sum_{\sigma\in\mathcal{E}} |\sigma|\, v_{K,\sigma}\, u_K\big(T_n(u_K) - T_n(u_L)\big).$$

Using the same arguments as in [12], we prove that

$$\begin{aligned}-T_2 &\leq \frac{1}{n}\frac{r^2 d\|v\|^2_{L^2(\Omega)^d}}{2} \\ &\quad + \frac{1}{2n}\sum_{\sigma\in\mathcal{E}(K)} \frac{|\sigma|}{d_\sigma}(u_K - u_L)(T_n(u_K) - T_n(u_L)).\end{aligned} \tag{34}$$

Using the fact that $\mathcal{B}(s) \geq 1,\ \forall s \in \mathbb{R}$, we get from the second term in the right hand side of (34)

$$\begin{aligned}&\frac{1}{2n}\sum_{\sigma\in\mathcal{E}(K)} \frac{|\sigma|}{d_\sigma}(u_K - u_L)(T_n(u_K) - T_n(u_L)) \\ &\leq \frac{1}{2n}\sum_{\sigma\in\mathcal{E}(K)} \frac{|\sigma|}{d_\sigma} B\big(v_{K,\sigma}\, d_\sigma\big)(u_K - u_L) \times \\ &\qquad \times(T_n(u_K) - T_n(u_L)).\end{aligned} \tag{35}$$

From (34) and (35), we deduce that

$$-T_2 \leq \frac{1}{n}\frac{r^2 d\|v\|^2_{L^2(\Omega)^d}}{2} + \frac{1}{2}T_1. \tag{36}$$

Combining (32), (33) and (36) we deduce (29).
By the same manage as in [12], we prove (30). ■

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

Corollary 4.5: (see [12]) Let $(\mathcal{T}_m)_{m\geq 1}$ be a sequence of admissible meshes such that there exists $\xi > 0$ satisfying for all $m \leq 1$ for all $K \in \mathcal{T}$ 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 (9)-(12), then

$$\lim_{n\to+\infty} \overline{\lim_{h_\mathcal{T}\to 0}} \sum_{\substack{\sigma\in\mathcal{E},\\ |u_K|\leq 2n,\\ |u_L>4n}} \frac{|\sigma|}{d_\sigma}|u_L| = 0. \tag{37}$$

V. Convergence results

We intend here to prove Theorem 2.11 that is the main result of this paper.
Before proving Theorem 2.11, we recall the following convergence result (see [12]) concerning the function (h_n) defined, for any $n \geq 1$, by

$$h_n(s) = \begin{cases} 0, & \text{if } s \leq -2n; \\ \dfrac{s}{n} + 2, & \text{if } -2n \leq s \leq -n, \\ 1, & \text{if } -n \leq s \leq n, \\ \dfrac{-s}{n} + 2, & \text{if } n \leq s \leq 2n, \\ 0, & \text{if } s \geq 2n. \end{cases} \tag{38}$$

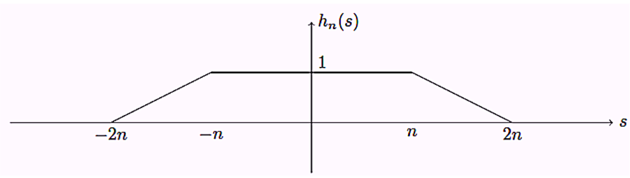

Fig. 3. **The function h_n**

Lemma 5.1: Let $(\mathcal{T}_m)_{m\geq 1}$ be a sequence of admissible meshes such that there exists $\xi > 0$ satisfying for all $m \geq 1$, for all $K \in T$ and for 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 (9)-(12). We define the function $\widetilde{h}_n$ by

$$\forall \sigma = K/L \in \mathcal{E}_{int}, \forall x \in D_\sigma, \widetilde{h}_n(x) = \frac{h_n(x_K) + h_n(x_L)}{2},$$

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

Proof: **Proof of Theorem 2.11**. Let $\varphi \in C_c^\infty(\Omega)$ and h_n the function defined by (22). 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 (9)-(12) 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

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

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

$$T_4 = \int_\Omega f\varphi_\mathcal{T} h_n(u_\mathcal{T})dx \underset{h_\mathcal{T}\to 0}{\longrightarrow} \int_\Omega f\varphi h_n(u)dx. \tag{39}$$

In view of the definition of $b_\mathcal{T}$, and since b belongs to $L^1(\Omega)$, $b_\mathcal{T} = (b_K)_{K\in\mathcal{T}}$ converges to b in $L^1(\Omega)$ as $h_\mathcal{T} \to 0$. With already used arguments we can assert that

$$\begin{aligned} T_3 &= \int_\Omega b_\mathcal{T} T_{2n}(u_\mathcal{T})\varphi_\mathcal{T} h_n(u_\mathcal{T})dx \\ &\underset{h_\mathcal{T}\to 0}{\longrightarrow} \int_\Omega bT_{2n}(u)\,\varphi\, h_n(u)dx. \end{aligned} \tag{40}$$

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

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

with

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

In view of the definitions of h_n and τ_σ 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 (29), we deduce that

$$\lim_{n\to+\infty} \overline{\lim_{h_\mathcal{T}\to 0}} T_{1,2} = 0. \tag{41}$$

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

$$\begin{aligned} T_{1,1} &= \sum_{\sigma\in\mathcal{E}} \tau_\sigma\, h_n(u_K)(T_{2n}(u_K) - u_L)(\varphi(x_K) - \varphi(x_L)) \\ &= T_{1,1}^1 + T_{1,1}^2 + T_{1,1}^3, \end{aligned}$$

with

$$T_{1,1}^1 = \sum_{\sigma\in\mathcal{E}} \tau_\sigma \frac{h_n(u_K)+h_n(u_L)}{2}(T_{4n}(u_K)-T_{4n}(u_L)) \times \times(\varphi(x_K)-\varphi(x_L)),$$

$$T_{1,1}^2 = \sum_{\sigma\in\mathcal{E}} \tau_\sigma h_n(u_K)(T_{4n}(u_L)-u_L)(\varphi(x_K)-\varphi(x_L)),$$

$$T_{1,1}^3 = \sum_{\sigma\in\mathcal{E}} \tau_\sigma \frac{h_n(u_K)-h_n(u_L)}{2}(T_{4n}(u_K)-T_{4n}(u_L)) \times \times(\varphi(x_K)-\varphi(x_L)).$$

Since in $T_{1,1}^1$, u_K and u_L are both truncated we can pass to the limit in I as $h_T \to 0$ by writing

$$\begin{aligned} T_{1,1}^1 &= \sum_{\sigma\in\mathcal{E}} \frac{d_\sigma}{d}\tau_\sigma\big(T_{4n}(u_K)-T_{4n}(u_L)\big) \times \\ &\quad \times \frac{h_n(u_K)+h_n(u_L)}{2} d\frac{(\varphi(x_K)-\varphi(x_L))}{d_\sigma} \\ &= I_1 + I_2, \end{aligned}$$

with

$$I_1 = -\sum_{\sigma\in\mathcal{E}} \int_\sigma A_\mathcal{T}\nabla_\mathcal{T} T_{4n}(u_\mathcal{T})\cdot\nabla_\mathcal{T}\varphi_\mathcal{T}\,\widetilde{h_n}(x)\,d\gamma(x),$$

$$\begin{aligned} I_2 &= \sum_{\sigma\in\mathcal{E}} |\sigma| \frac{h_n(u_K)+h_n(u_L)}{2}\,\frac{(\varphi(x_K)-\varphi(x_L))}{d_\sigma} \times \\ &\quad \times\big(T_{4n}(u_K)-T_{4n}(u_L)\big) \\ &\quad -\sum_{\sigma\in\mathcal{E}} \frac{h_n(u_K)+h_n(u_L)}{2}\,\frac{(\varphi(x_K)-\varphi(x_L))}{d_\sigma} \times \\ &\quad \times\frac{1}{|D_\sigma|}\int_\sigma A_\mathcal{T}\nabla_\mathcal{T} T_{4n}(u_\mathcal{T})\cdot\eta_{K,\sigma} d\gamma(x). \end{aligned}$$

By Lemma 5.1, $\widetilde{h_n}(u_\mathcal{T}) \to h_n(u)$ in $L^q(\Omega)$ $\forall q\in[2,+\infty[$, while $A_\mathcal{T}\nabla_\mathcal{T}T_{4n}(u_\mathcal{T})$ tends to $A\nabla T_{4n}(u_\mathcal{T})$ weakly in $(L^2(\Omega))^2$. Since φ belongs to $C_c^\infty(\Omega)$ we conclude that

$$\lim_{h_\mathcal{T}\to 0} I_1 = \int_\Omega h_n(u)\,A\nabla T_{4n}(u).\nabla\varphi dx. \tag{42}$$

By the regularity of φ and due to hypothesis (H_1), we see that

$$|I_2| \le \beta\ \|\varphi\|_{W^{1,\infty}(\Omega)}\,h_\mathcal{T}\,\|T_{4n}(u_\mathcal{T})\|^2_{1,2,\mathcal{T}}\,.$$

Therefore,

$$\lim_{h_\mathcal{T}\to 0} I_2 = 0. \tag{43}$$

We now turn to $T_{1,1}^2$.

$$\begin{aligned} |T_{1,1}^2| &\le \sum_{\sigma\in\mathcal{E}} \frac{|\sigma|}{d_\sigma}|h_n(u_K)||T_{4n}(u_L-u_L||(\varphi(x_K)-\varphi(x_L))| \\ &\le 2\|\varphi\|_{L^\infty(\Omega)} \sum_{\substack{\sigma\in\mathcal{E}\\ |u_K\le 2n|,\ |u_L>4n|}} \frac{|\sigma|}{d_\sigma}|u_L|. \end{aligned}$$

We can deduce by Corollary 4.5 that

$$\lim_{n\to+\infty}\overline{\lim_{h_\mathcal{T}\to 0}} T_{1,1}^2 = 0. \tag{44}$$

Noticing that

$$|T_{1,1}^3| \le \frac{||\varphi||_{L^\infty(\Omega)}}{n}\sum_{\sigma\in\mathcal{E}}\frac{|\sigma|}{d_\sigma}|u_K-u_L||T_n(u_K)-T_n(u_L)|,$$

and using estimate (29), we deduce that

$$\lim_{n\to+\infty}\overline{\lim_{h_\mathcal{T}\to 0}} T_{1,1}^3 = 0. \tag{45}$$

For the convection term we have

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

with

$$T_{2,1} = \sum_{\sigma\in\mathcal{E}} |\sigma| v_{K,\sigma}u_{\sigma,+}h_n(u_{\sigma,+})(\varphi(x_K)-\varphi(x_L)),$$

$$T_{2,2} = \sum_{\substack{\sigma\in\mathcal{E}\\ v_{K,\sigma}\ge 0}} |\sigma| v_{K,\sigma}u_{\sigma,+}\varphi(x_L)(h_n(u_{\sigma,+})-h_n(u_{\sigma,-})),$$

$$T_{2,3} = \sum_{\substack{\sigma\in\mathcal{E}\\ v_{K,\sigma}< 0}} |\sigma| v_{K,\sigma}u_{\sigma,+}\varphi(x_K)(h_n(u_{\sigma,+})-h_n(u_{\sigma,-})).$$

We prove that according to [12]

$$\lim_{n\to+\infty}\overline{\lim_{h_\mathcal{T}\to 0}}(T_{2,2}-T_{2,3}) = 0 \tag{46}$$

and

$$\lim_{h_\mathcal{T}\to 0} T_{2,1} = -\int_\Omega T_{2n}(u)h_n(u)v.\nabla\varphi dx. \tag{47}$$

We are now in position to pass to the limit as $h_\mathcal{T}\to 0$ in the scheme (9)-(12). Gathering equations from (39) to (47), we can assert that

$$\begin{aligned} &\int_\Omega h_n(u)\,A\nabla u\cdot\nabla\varphi dx - \int_\Omega u\,h_n(u)\,v\cdot\nabla\varphi dx \\ &\quad + \int_\Omega b\,u\,h_n(u)\,\varphi dx - \int_\Omega f\,\varphi\,h_n(u) dx \\ &\quad = \lim_{h_\mathcal{T}\to 0} T(n,\varphi), \end{aligned} \tag{48}$$

where $\lim_{h_\mathcal{T}\to 0}|T(n,\varphi)| \le ||\varphi||_{L^\infty(\Omega)}\,\omega(n)$ with $\omega(n)\to 0$ as $n\to+\infty$.

Let $h\in C_c^1(\mathbb{R})$ and $\psi\in C_c^1(\Omega)\cap H_0^1(\Omega)$. In view of the regularity of $T_n(u)$(see (2)) the function $h(u)\psi$ belongs to $L^\infty(\Omega)\cap H_0^1(\Omega)$. By the density of $C_c^\infty(\Omega)$ 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 $C_c^\infty(\Omega)$ which is bounded in $L^\infty(\Omega)$). So, taking $\varphi = h(u)\,\psi$ in (48) and passing to the limit as $n\to+\infty$, we deduce that

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

which is Equality (4) in the definition of a renormalized solution. It remains to prove that u satisfies the decay (3) of the truncate energy.

Thanks to the discrete estimate on the energy (27) we get,

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

and

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

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

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

which leads to

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

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 $H_0^1(\Omega)$. ■

REFERENCES

[1] M. Ben Cheick and O. Guibé, *Nonlinear and non-coercive elliptic problems with integrable data*. Adv. Math. Sci. Appl. 16 $n^\circ 1$, $275-297, 2006$.

[2] L. Boccardo and T. Gallouët, *On some nonlinear elliptic and parabolic equations involving measure data*. J. Funct. Anal., 87, 149-169, 1989.

[3] J. Carrillo and M. Chipot, *On some nonlinear elliptic equations involving derivatives of the nonlinearity*. Proc. Roy. Soc. Edinburgh Sect $A100$, $n^\circ 3-4$, 281-294, 1985.

[4] C. Chainais-Hillairet and J. Droniou, *Finite volume schemes for non-coercive elliptic problems with Neumann boundary conditions*. IMA J. Numer. Anal. 31, $n^\circ 1$, 61-85, 2011.

[5] 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.

[6] R.-J. DiPerna and P.-L. Lions, *On the Cauchy problem for Boltzmann equations: Global existence and weak stability*. Ann. Math. 130, 321-366, 1989.

[7] J. Droniou, T. Gallouët and R. Herbin, *A finite volume scheme for a noncoercive elliptic equation with measure data*. SIAM J. Numer. Anal., 6, 1997-2031, 2003.

[8] R. Eymard, T. Gallouët and R. Herbin, *Convergence of finite volume approximations to the solutions of semilinear convection diffusion reaction equations*. Numer. Math., 82, 91-116 (1999).

[9] R. Eymard, T. Gallouët and R. Herbin, Finite volume methods. In P. Ciarlet and J.L. Lions, editors, Handbook of Numerical Analysis, Vol VII, p. 713-1020. North Holland, 2000.

[10] T. Gallouët and R. Herbin, Finite volume approximation of elliptic problems with irregular data. Finite volumes for complex applications II, 155-162, Hermes Sci. Publ.,Paris, 1999.

[11] O. Guibé and A. Mercaldo, *Uniqueness results for noncoercive nonlinear elliptic equations with two lower order terms*. Commun. Pure Appl. Anal. 7, $n^\circ 1$, 163-192, 2008.

[12] S. Leclavier, *Finite volume scheme and renormalized solutions for a noncoercive elliptic problem with L^1- data*, Comput. Methods Appl. Math. 17, No. 1, 85-104 (2017).

[13] P.-L. Lions and F. Murat. Solutions renormalisées d'équations elliptiques non linéaires. In preparation.

[14] F. Murat, Soluciones renormalizadas de EDP elipticas no lineales. Technical Report R93023, Laboratoire d'Analyse Numérique, Paris VI, 1993.

[15] A. Ouédraogo, T. Valea and W. B. Yaméogo, *Renormalized solutions for a non-coercive elliptic problem with measure data*. Submitted.

[16] K. Pananu, S. Sungnul, S. Sirisubtawee and S. Phongthanapanich, *Convergence and Applications of the Implicit Finite Difference Method for Advection-Diffusion-Reaction Equations*, IAENG International Journal of Computer Science, vol. 47, no.4, pp645-663, 2020.

[17] K. Para, S. Sungnul, S. Sirisubtawee and S. Phongthanapanich, *A Comparison of Numerical Solutions for Advection-Diffusion-Reaction Equations between Finite Volume and Finite Difference Methods*, Engineering Letters, vol. 30, no.2, pp566-581, 2022.