A finite volume scheme for a noncoercive elliptic equation with measure data

Jérôme Droniou[1], Thierry Gallouët[2], Raphaèle Herbin[3]
May 14, 2003.

Abstract
We show here the convergence of the finite volume approximate solutions of a convection-diffusion equation to a weak solution, without the usual coercitivity assumption on the elliptic operator and with weak regularity assumptions on the data. Numerical experiments are performed to obtain some rates of convergence in two and three space dimensions.

1 Introduction

The scope of this work is the discretization by the cell-centered finite volume method of convection-diffusion problems on general structured or non structured grids. Let Ω be a polygonal (or polyhedral) open subset of $\mathbb{R}^d$ ($d = 2$ or 3); the problem under study writes:

$$\left\{ \begin{array}{ll} -\Delta u + \mathrm{div}(\mathbf{v}u) + bu = \mu & \text{in } \Omega, \\ u = 0 & \text{on } \partial\Omega \end{array} \right. \tag{1}$$

with the following hypotheses on the data:

$$\begin{array}{c} \mathbf{v} \in (C(\overline{\Omega}))^d\,, \\ b \in L^2(\Omega)\,,\ b \geq 0 \text{ a.e. on } \Omega\,, \\ \mu \in M(\overline{\Omega})\,, \end{array} \tag{2}$$

where $M(\overline{\Omega}) = (C(\overline{\Omega}))'$ is the dual space of $C(\overline{\Omega})$, which may also be identified to the set of bounded measures on $\overline{\Omega}$. In the sequel, we shall consider the usual infinity norm on $C(\overline{\Omega})$), and we shall denote by $||\cdot||_{M(\overline{\Omega})}$ its dual norm on $M(\overline{\Omega})$.
Our purpose is to prove the convergence of the cell-centered finite volume scheme for the discretization of Problem (1). Cell-centered schemes for convection-diffusion equations using rectangular, triangular or Voronoï grids were analysed in a number of papers including [27],[18], [23], [26], [29], [8]. The analysis which we develop here uses some of the tools which were developped in [14], [20], [15] and [19]. In [15], a convergence result without any assumption of regularity of the solution is proved. An approximate gradient was constructed in [7]. Noncoercive elliptic equations with a regular H^{-1} right-hand-side were also recently studied [12]. Finally, a thorough study of finite volume schemes for linear or nonlinear elliptic, parabolic and hyperbolic equations may be found in [14], which we refer to for further details. The discretization grids which are considered here and in these latter works consist of polygonal (or polyhedral) control volumes satisfying adequate geometrical conditions (which are stated in the sequel) and not necessarily ordered in a cartesian grid.

Let us remark that the analysis which is developed here still holds for equations of the type

$$-\mathrm{div}\Big(k(x)\nabla u(x)\Big) + \mathrm{div}\Big(\mathbf{v}(x)\,u(x)\Big) + b(x)u(x) = f(x), \quad x \in \Omega, \tag{3}$$

with the following hypotheses on k:

$$\begin{array}{l} k \text{ is a piecewise } C^1 \text{ function from } \overline{\Omega} \text{ to } \mathbb{R}; \\ \text{there exists } k_0 \in \mathbb{R}_+^* \text{ such that } k(x) \geq k_0 \text{ for a.e. } x \in \Omega. \end{array} \tag{4}$$

[1]Département de Mathématiques, Université de Montpellier, droniou@math.univ-montp2.fr
[2]L.A.T.P., UMR 6632, Université de Provence, gallouet@cmi.univ-mrs.fr
[3]L.A.T.P., UMR 6632, Université de Provence, herbin@cmi.univ-mrs.fr

For the sake of the simplicity of notations we prefer to deal with the Laplace operator here but we shall point out the modifications which take place if the operator $\mathrm{div}(k\nabla .)$ is considered instead: see remarks 2.2, 2.4 and 2.6. If now k is a tensor satisfying the following hypotheses:

$$\begin{array}{l} k \text{ is a piecewise } C^1 \text{ function from } \overline{\Omega} \text{ to } \mathbb{R}^{d\times d}, \\ \text{for all } x \in \overline{\Omega}, k(x) \text{ is a symmetric matrix}, \\ \text{there exists } k_0 \in \mathbb{R}_+^* \text{ such that } k(x)\xi\cdot\xi \geq k_0 \text{ for a.e. } x \in \Omega \text{ and for all } \xi \in \mathbb{R}^d, \end{array} \tag{5}$$

then one may still write the finite volume scheme and obtain some error estimates in the regular case, but the assumptions on the mesh have to be modified see [20], [24] and [8]. However if the mesh is Cartesian and if for all $x \in \overline{\Omega}$ the matrix $k(x)$ is diagonal then it is "aligned" with the grid and the analysis is similar to the (non constant) scalar case of Equation (3).
The originality of the present work with respect to the above cited works is threefold: first, the elliptic operator associated to the convection-diffusion equation is not assumed to be coercive; second, the convection velocity $\mathbf{v}$ is only assumed to be continuous (it was assumed C^1 in previous works); third, the right hand side μ is only supposed to be a Radon measure.

In the next section, the finite volume scheme for the discretization of (1) is presented, along with the admissible meshes. We then state the main convergence theorem of this paper (Theorem 2.1), along with some preliminary technical results similar to those used in [14], [20], [15], and the proof of which is given in an appendix. Section 3 is devoted to *a priori* estimates on the approximate solutions (existence is not proven at this stage), which will be needed in order to obtain compactness results, and which also yield the existence and uniqueness of the approximate solution. The proof of Theorem 2.1, that is the proof of the convergence of the approximate solutions to the weak solution of (1), is then given in Section 4. Section 5 presents a modified finite volume scheme where the measure data whose support is on the edges of the mesh are taken into account through a jump of the flux between two neighboring cells; comparing this scheme to the scheme of Section 2, the convergence result is easy to obtain. Finally, we present in Section 6 some numerical results in two and three space dimensions, using Cartesian or unstructured triangular meshes (in 2D), as well as for a spherical geometry. These results allow to derive some rates of convergence of the method, even though no error estimate is known theoretically.

2 Conservative finite volume discretization and convergence result

Definition 2.1 *An admissible mesh of Ω, denoted by $\mathcal{M}$, is given by a finite partition $\mathcal{T}$ of Ω in polygonal (or polyhedral) convex sets (the "control volumes"), by a finite family $\mathcal{E}$ of disjoint subsets of $\overline{\Omega}$ contained in affine hyperplanes (the "edges") and by a family $\mathcal{P} = (x_K)_{K\in\mathcal{T}}$ of points in Ω such that*

i) each $\sigma \in \mathcal{E}$ is a non-empty open subset of ∂K for some $K \in \mathcal{T}$,

ii) by denoting $\mathcal{E}_K = \{\sigma \in \mathcal{E} \mid \sigma \subset \partial K\}$, one has $\partial K = \cup_{\sigma\in\mathcal{E}_K} \overline{\sigma}$ for all $K \in \mathcal{T}$,

iii) for all $K \neq L$ in $\mathcal{T}$, either the $(d-1)$-dimensional measure of $\overline{K}\cap\overline{L}$ is null, or $\overline{K}\cap\overline{L} = \overline{\sigma}$ for some $\sigma \in \mathcal{E}$, that we denote then $\sigma = K|L$,

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

v) for all $\sigma = K|L \in \mathcal{E}$, the line (x_K, x_L) intersects and is orthogonal to σ,

vi) for all $\sigma \in \mathcal{E}$, $\sigma \subset \partial\Omega \cap \partial K$, the line which is orthogonal to σ and going through x_K intersects σ.

Remark 2.1 (Other admissible meshes) *Note that Property v) in the above definition is required so as to obtain a consistent discretization of the normal fluxes over the boundary of the control domains when*

using the two points finite difference scheme to discretize the normal flux. In fact, the above definition of an admissible mesh may be extended to other geometries of Ω than a polygone or a polyhedron. For instance, if $\Omega = \{x \in \mathbb{R}^d; |x| \leq r\}$ is a spherical ball of radius r, then a natural mesh is defined by the control volumes $K_0 = \{x \in \mathbb{R}^d; |x| \leq r_{1/2}\}$ and, for $i = 1, N$, $K_i = \{x \in \mathbb{R}^d; r_{i-1/2} \leq |x| \leq r_{i+1/2}\}$ where $(r_{i+1/2})_{i=1,N} \subset (0, r]$ is a given increasing sequence such that $r_{N+1/2} = r$. Let $x_0 = 0$ and, for $i = 1, \ldots, N$, $r_i \in (r_{i-1/2}, r_{i+1/2})$, then a discretization of the normal diffusive flux $\nabla u \cdot \mathbf{n}$ (where $\mathbf{n}$ is the outward normal unit vector) over the sphere $\{x \in \mathbb{R}^d; |x| = r_{i+1/2}\}$ by the two points scheme $\frac{u_{i+1}-u_i}{r_{i+1}-r_i}$ is clearly consistent if the solution u to (1) only depends on r. Moreover, if $r_{i+1/2} = \frac{1}{2}(r_{i+1} - r_i)$, it is consistent of order 2. Hence this class of spherical discretizations is clearly admissible for the analysis which will be derived in the sequel.

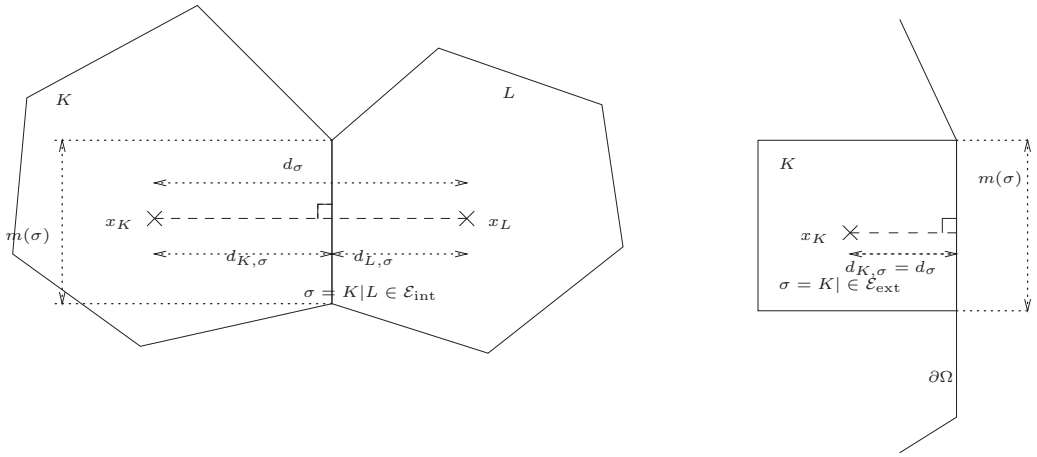


Figure 1: Notations for an admissible mesh

The size of the mesh is then defined by $\text{size}(\mathcal{M}) = \sup_{K\in\mathcal{T}} \text{diam}(K)$. We denote by $\text{meas}(K)$ the Lebesgue measure of $K \in \mathcal{T}$. The unit normal to $\sigma \in \mathcal{E}_K$ outward to K is denoted by $\mathbf{n}_{K,\sigma}$.
We define $\mathcal{E}_{\rm int} = \{\sigma \in \mathcal{E} \mid \sigma \not\subset \partial\Omega\}$ and $\mathcal{E}_{\rm ext} = \mathcal{E}\backslash\mathcal{E}_{\rm int}$. If $\sigma \in \mathcal{E}$, $\text{meas}(\sigma)$ is the $(d-1)$-dimensional measure of σ; if $\sigma = K|L \in \mathcal{E}_{\rm int}$, d_σ is the distance between the points (x_K, x_L) and $d_{K,\sigma}$ denotes the distance between x_K and σ; if $\sigma \in \mathcal{E}_{\rm ext} \cap \mathcal{E}_K$, $d_\sigma = d_{K,\sigma}$ is the distance between x_K and σ. The transmissivity through an edge σ is

$$\tau_\sigma = \frac{\text{meas}(\sigma)}{d_\sigma}.$$

Within the integrals, the letter λ (resp. γ) stands for the d (resp.$(d-1)$)-dimensional measure on the domain Ω (resp. on the edges of the mesh). Note that moth measures are denoted by "meas" when applied to a control volume or an edge.
We shall naturally identify the set $\mathbb{R}^{\text{Card}(\mathcal{T})}$ to the set $X(\mathcal{T})$ of functions defined a.e. on Ω and constant on each control volume $K \in \mathcal{T}$.

Remark 2.2 *In the case of the operator* $\text{div}(k\nabla.)$ *which is considered in Equation (3) where k is a function from $\overline{\Omega}$ to $\mathbb{R}$ or $\mathbb{R}^{d\times d}$ which satisfies (4) or (5), admissible meshes must satisfy the following additional condition:*

(vi) For any $K \in \mathcal{T}$, the restriction $k|_K$ of the function k to any given control volume K belongs to $C^1(\overline{K})$.

Furthermore if k is a piecewise C^1 function from $\overline{\Omega}$ to $\mathbb{R}^{d\times d}$, the orthogonality conditions (iv) and (v) have to be modified into:

(iv)' For any $K \in \mathcal{T}$, let k_K denote the mean value of k on K, that is

$$k_K = \frac{1}{meas(K)} \int_K k d\lambda. \tag{6}$$

The set $\mathcal{T}$ is such that there exists a family of points

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

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

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

If $\mathcal{M}$ is an admissible mesh, and under Hypothesis (2), we can define the finite volume discretization of (1).
By denoting, for $K \in \mathcal{T}$ and $\sigma \in \mathcal{E}_K$,

$$b_K = \frac{1}{\text{meas}(K)} \int_K b d\lambda \quad \text{and} \quad v_{K,\sigma} = \int_\sigma \mathbf{v} \cdot \mathbf{n}_{K,\sigma} \, d\gamma \tag{7}$$

the scheme is defined by

$$\forall K \in \mathcal{T}\,, \ \sum_{\sigma\in\mathcal{E}_K} F_{K,\sigma} + \sum_{\sigma\in\mathcal{E}_K} v_{K,\sigma} u_{\sigma,+} + \text{meas}(K) b_K u_K = \mu(K)\,, \tag{8}$$

$$\begin{array}{ll} \forall \sigma = K|L \in \mathcal{E}_{\text{int}}\,, & F_{K,\sigma} = -\tau_\sigma(u_L - u_K)\,, \\ \forall \sigma \in \mathcal{E}_{\text{ext}} \cap \mathcal{E}_K\,, & F_{K,\sigma} = \tau_\sigma u_K\,, \end{array} \tag{9}$$

$$\begin{array}{l} \forall \sigma = K|L \in \mathcal{E}_{\text{int}}\,, \quad u_{\sigma,+} = u_K \ \text{ if } \ v_{K,\sigma} \geq 0\,, \quad u_{\sigma,+} = u_L \ \text{ otherwise,} \\ \forall \sigma \in \mathcal{E}_{\text{ext}} \cap \mathcal{E}_K\,, \quad u_{\sigma,+} = u_K \ \text{ if } \ v_{K,\sigma} \geq 0\,, \quad u_{\sigma,+} = 0 \ \text{ otherwise.} \end{array} \tag{10}$$

Equations (8)—(10) form a linear system in $(u_K)_{K\in\mathcal{T}}$ of size Card($\mathcal{T}$). Notice that this scheme is conservative in the sense that if $\sigma = K|L$, then $F_{K,\sigma} = -F_{L,\sigma}$ and $v_{K,\sigma} = -v_{L,\sigma}$.

Remark 2.3 *The approximation (10) of the convective flux is the classical upwind scheme, which we choose here because it ensures both the existence of a solution to the scheme (and the maximum principle) without any condition on the size of the mesh. If instead of the upwind scheme, we used the central difference scheme, then we would need a condition on the size of the mesh in order to have existence of a solution to the scheme, and in order for the maximum principle to hold. However, when the size of the mesh tends to 0, the centered scheme may also be shown to converge. The upwind scheme is often preferred in applications because of its robustness on coarse meshes.*
Also note that if $v_{K,\sigma} = 0$, for some $\sigma = K|L$ for example, then (10) does not determine $u_{\sigma,+}$ uniquely since one may take either $u_{\sigma,+} = u_K$ (since $v_{K,\sigma} \geq 0$) or $u_{\sigma,+} = u_L$ (since $v_{L,\sigma} = -v_{K,\sigma} = 0 \geq 0$). However, this is no real problem since $u_{\sigma,+}$ always appears multiplied by $v_{K,\sigma}$ or $v_{L,\sigma}$ and thus, if $v_{K,\sigma} = 0$, the value of $u_{\sigma,+}$ does not matter (one can, for example, reduce the second sum of (8) to the $\sigma \in \mathcal{E}_K$ such that $v_{K,\sigma} \neq 0$).

Remark 2.4 *In the case of a non constant diffusion coefficient as in Equation (3) where k is a function from Ω to $\mathbb{R}$ satisfying (4) or from Ω to $\mathbb{R}^{d\times d}$ satisfying (5), one considers admissible meshes satisfying*

(vi) of Remark 2.2 and in the tensor case also (iv)′ and (v)′ instead of (iv) and (v). For $K \in \mathcal{T}$ and $\sigma \in \mathcal{E}_K$, let

$$k_{K,\sigma} = \left| \frac{1}{\mathrm{meas}(K)} \int_K k \, d\lambda \, \mathbf{n}_{K,\sigma}, \right| \tag{11}$$

(where $|\cdot|$ denotes the Euclidean norm). Note that in the scalar case, this yields in fact $k_{K,\sigma} = \frac{1}{\mathrm{meas}(K)} \int_K k d\lambda$. The exact diffusion fluxes $k(x)\nabla u \cdot \mathbf{n}_{K,\sigma}$ on an edge σ of the mesh may then be approximated in a consistent way (see [14] and [24]) by replacing the formulae in (9) by:

- *internal edges:*

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

  where

$$\tau_\sigma = \mathrm{meas}(\sigma) \frac{k_{K,\sigma} k_{L,\sigma}}{k_{K,\sigma} d_{L,\sigma} + k_{L,\sigma} d_{K,\sigma}};$$

- *boundary edges:*

$$F_{K,\sigma} = -\tau_\sigma(u_\sigma - u_K), \ \text{if } \sigma \in \mathcal{E}_{\mathrm{ext}} \text{ and } x_K \notin \sigma, \tag{13}$$

  where

$$\tau_\sigma = \mathrm{meas}(\sigma) \frac{k_{K,\sigma}}{d_{K,\sigma}}.$$

Let us now state our main result, which we shall prove in the following sections.

Theorem 2.1 *If $\mathcal{M}$ is an admissible mesh, then there exists a unique solution to (8)—(10). Moreover, if $(\mathcal{M}_n)_{n\geq 1}$ is a sequence of admissible meshes such that there exists $\zeta > 0$ satisfying*

$$\text{for all } n \geq 1, \text{ for all } K \in \mathcal{T}_n, \text{ for all } \sigma \in \mathcal{E}_K, \ d_{K,\sigma} \geq \zeta d_\sigma,$$

and such that $\mathrm{size}(\mathcal{M}_n) \to 0$, then, by denoting $u_n \in X(\mathcal{T}_n)$ the solution of (8)—(10) with $\mathcal{M} = \mathcal{M}_n$, $(u_n)_{n\geq 1}$ converges to u in $L^p(\Omega)$ for all $p \in [1, \frac{d}{d-2})$, where u is the unique solution to (1) in the sense

$$\begin{cases} u \in \bigcap\limits_{q < \frac{d}{d-1}} W_0^{1,q}(\Omega), \\ \displaystyle\int_\Omega \nabla u \cdot \nabla \varphi d\lambda - \int_\Omega u \mathbf{v} \cdot \nabla \varphi d\lambda + \int_\Omega b u \varphi d\lambda = \int_\Omega \varphi \, d\mu, \ \forall \varphi \in \bigcup\limits_{s>d} W_0^{1,s}(\Omega), \end{cases} \tag{14}$$

where $\int_\Omega \varphi \, d\mu =< \mu, \varphi >_{(C(\bar\Omega))', C(\bar\Omega)}$. (We recall that $W^{1,q}(\Omega)$ is the set of functions which belong to $L^q(\Omega)$ and such that their derivatives are also in $L^q(\Omega)$, and $W_0^{1,q}(\Omega) = \overline{C_c^\infty(\Omega)}^{W^{1,q}(\Omega)}$. We also recall that $W_0^{1,s}(\Omega) \subset C_0(\bar\Omega)$ for $s > d$.)

Remark 2.5 *Notice that we do not suppose the existence and uniqueness of a solution to (14); we will prove both.*

Remark 2.6 *A convergence result still holds if a non constant piecewise C^1 diffusion scalar coefficient is considered i.e. if k satisfies (4) and if Equation (3) is discretized by the scheme (7),(8),(11)—(13). In fact, in the two-dimensional case, the proof follows the one given below in the case $k = Id$. In the three-dimensional case however, the regularity of the solution to the dual problem (47), which is used in the proof of the uniqueness of a solution to (14) (see section 4) is not so clear. Hence in the 3D case, uniqueness of a solution to (14) is not known, and the convergence result of Theorem (2.1) still holds, but only up to a subsequence.*
If one now considers the general tensor case, then some more restrictive assumptions are needed on the mesh in order to obtain consistency of the fluxes, see [14] and [24].

The proof of existence and uniqueness of a solution to (8)—(10) is based on *a priori* estimates on the solutions to this problem, which are obtained with the following discrete $W_0^{1,q}$ norm, defined as follows.

Definition 2.2 (Discrete $W^{1,q}$ norm) *If $\mathcal{M}$ is an admissible mesh, $v_\mathcal{T} = (v_K)_{K\in\mathcal{T}} \in \mathbb{R}^{\mathrm{Card}(\mathcal{T})}$ and $1 \le q < \infty$, we define*

$$||v_\mathcal{T}||_{1,q,\mathcal{M}} = \left(\sum_{\sigma\in\mathcal{E}} meas(\sigma) d_\sigma \left(\frac{D_\sigma v_\mathcal{T}}{d_\sigma}\right)^q\right)^{1/q},$$

where $D_\sigma v_\mathcal{T} = |v_K - v_L|$ if $\sigma = K|L \in \mathcal{E}_{\rm int}$ and $D_\sigma v_\mathcal{T} = |v_K|$ if $\sigma \in \mathcal{E}_{\rm ext} \cap \mathcal{E}_K$.

Let us now state the main *a priori* estimate, which will be proven in Section 3. This estimate is crucial to prove the existence of a solution to (8)—(10), and also to obtain the compactness properties on approximate solutions which will eventually yield the convergence result.

Theorem 2.2 *Let $\mathcal{M}$ be an admissible mesh and $\zeta > 0$ satisfying*

$$\text{for all } K \in \mathcal{T} \text{ and all } \sigma \in \mathcal{E}_K,\ d_{K,\sigma} \ge \zeta d_\sigma. \tag{15}$$

Then, for all $q \in [1, \frac{d}{d-1})$, there exists $C > 0$ only depending on $(\Omega, \mathbf{v}, q, \zeta)$ such that, if $u_\mathcal{T} \in X(\mathcal{T})$ is a solution to (8)—(10), then $||u_\mathcal{T}||_{1,q,\mathcal{M}} \le C||\mu||_{M(\overline{\Omega})}$.

In the sequel, we shall use the following properties of the discrete $W_0^{1,q}$ norm:

Proposition 2.1 (Discrete Poincaré inequality) *If $1 \le q \le 2$, $\mathcal{M}$ is an admissible mesh and $v_\mathcal{T} \in X(\mathcal{T})$, then*

$$\|v_\mathcal{T}\|_{L^q(\Omega)} \le \mathrm{diam}(\Omega)||v_\mathcal{T}||_{1,q,\mathcal{M}}. \tag{16}$$

Proposition 2.2 (Discrete Sobolev Inequality) *Let $1 \le q \le 2$, $\mathcal{M}$ be an admissible mesh and $\zeta > 0$ satisfying (15). Then, with $q^* = \frac{dq}{d-q}$ if $q < d$ and $q^* < \infty$ if $q = d = 2$, there exists $C > 0$ only depending on (Ω, q, q^*, ζ) such that, for all $v_\mathcal{T} \in X(\mathcal{T})$,*

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

In fact, it is easily seen that the above inequality also holds for any $r \le q^*$, that is:

$$||v_\mathcal{T}||_{L^r(\Omega)} \le C||v_\mathcal{T}||_{1,q,\mathcal{M}}, \text{ for any } r \le q^*.$$

Proposition 2.3 (Discrete Rellich Theorem) *Let $1 \le q \le 2$ and $\mathcal{M}$ be an admissible mesh. Then there exists $C > 0$ only depending on (Ω, q) such that, for all $h \in \mathbb{R}^d$ and all $v_\mathcal{T} \in X(\mathcal{T})$, denoting $w_\mathcal{T}$ the extension of $v_\mathcal{T}$ to $\mathbb{R}^d$ by 0 outside Ω, we have*

$$\int_{\mathbb{R}^d} |w_\mathcal{T}(x+h) - w_\mathcal{T}(x)|^q \, d\lambda(x) \le |h|(|h| + C\mathrm{size}(\mathcal{M}))^{q-1}||v_\mathcal{T}||^q_{1,q,\mathcal{M}}. \tag{17}$$

In particular, if $(\mathcal{M}_n)_{n\ge1}$ is a sequence of admissible meshes and $v_n \in X(\mathcal{T}_n)$ is such that $(||v_n||_{1,q,\mathcal{M}_n})_{n\ge1}$ is bounded, then $(v_n)_{n\ge1}$ is relatively compact in $L^q(\Omega)$.

Proposition 2.4 (Regularity of the limit) *Let $q \in (1, 2]$ and $(\mathcal{M}_n)_{n\ge1}$ be a sequence of admissible meshes such that $\mathrm{size}(\mathcal{M}_n) \to 0$. If $v_n \in X(\mathcal{T}_n)$, $(||v_n||_{1,q,\mathcal{M}_n})_{n\ge1}$ is bounded and $v_n \to v$ in $L^q(\Omega)$, then $v \in W_0^{1,q}(\Omega)$.*

These propositions are easy adaptations of similar results in [14] for the case $q = 2$ (see also [6] for Proposition 2.2 and [19] for Proposition 2.3). We sketch the proofs of these propositions in the appendix for the sake of completeness.

3 A Priori Estimates

The aim of this section is to prove the discrete $W^{1,q}$ *a priori* estimate of Theorem 2.2, which is crucial in the proof of existence of the scheme, and also in the obtention of a compactness result which will allow to prove the convergence of a sequence of approximate solutions (Theorem 2.1 and its proof in Section 4).

Such *a priori* estimates were already used for the study of the finite volume approximation of nonlinear elliptic or parabolic equations, see e.g. [15], [16]. But in these previous works, the estimates were obtained in a discrete H^1 norm, accordingly with the regularity of the solution of the continuous problem.

We prove here some *a priori* estimates on the solution to (8)—(10) in a discrete $W^{1,q}$ norm, since the solution to the continuous problem is in $W^{1,q}$. As in the continuous case, it is difficult to obtain an estimate on $u_{\mathcal{T}}$ itself (note that in the continuous case, u is not allowed as a test function in (14)). Hence, as in [19], we shall obtain estimates on truncations of the approximate solutions, that is the functions $T_k(u_{\mathcal{T}})$, where T_k is defined in Figure 2. However, in [19], we only dealt with the Laplace operator, whereas here we allow non-coercive convection-diffusion operators. Because of this non-coercivity, we shall need to start with some weaker estimates, namely an estimate on $\ln(1+|u_{\mathcal{T}}|)$, as was done in [9] in the continuous case. In order to obtain this estimate, we shall obtain some estimate on $S_k(u_{\mathcal{T}})$, where $S_k = Id - T_k$ is also defined in Figure 2 and section 3.2. Note that in the diffusion dominated case, the operator is becomes coercive and the discrete $W^{1,q}$ estimate may be directly obtained from the estimates on $T_k(u_{\mathcal{T}})$ as in [19].

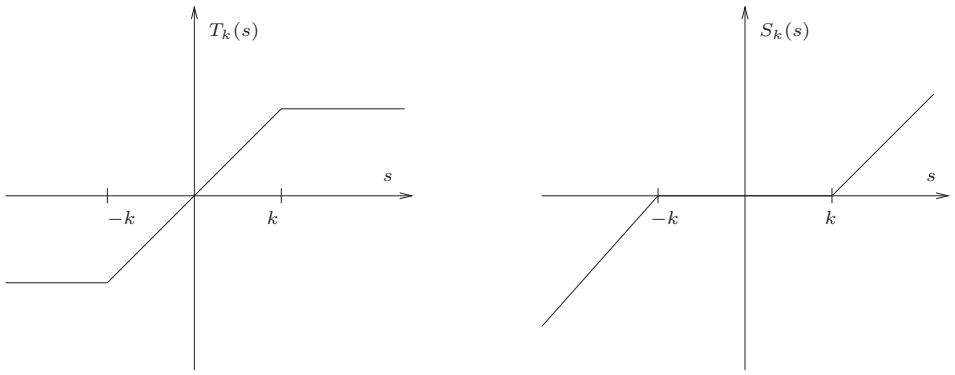


Figure 2: The functions T_k and S_k

Since the function T_k is bounded, the estimate on $T_k(u_{\mathcal{T}})$ is easy to obtain. The estimate on $S_k(u_{\mathcal{T}})$ is more tricky. The convective term is controlled through a bound of $\mathrm{meas}(E_k)$ where $E_k = \{|u_{\mathcal{T}}| > k\}$ (see Corollary 3.1), which is a consequence of an estimate on $\ln(1+|u_{\mathcal{T}}|)$ (see Proposition 3.1).

Each of the estimates we present here has a continuous counterpart; see for example [2], [3] for estimates on nonlinear elliptic equations with measure data and [9], [10] for estimates on linear and nonlinear noncoercive variational elliptic problems. Mixing the techniques of [3] and [9] (or [10]), we can prove estimates (and an existence result) on solutions to linear or nonlinear noncoercive elliptic equations with measure data.

To obtain the estimates on the solutions to (8)—(10), we adapt to the discrete setting this mix of techniques of [3] and [9]. Thus, to make the following proofs easier to understand, we sketch, for each of the discrete estimate, the proof of the corresponding continuous estimate.

3.1 Estimate on $\ln(1+|u_{\mathcal{T}}|)$

Proposition 3.1 *Let $\mathcal{M}$ be an admissible mesh. If $u_{\mathcal{T}} = (u_K)_{K\in\mathcal{T}}$ is a solution to (8)—(10), then*

$$|| \ln(1+|u_{\mathcal{T}}|)||^2_{1,2,\mathcal{T}} \leq 2||\mu||_{M(\overline{\Omega})} + d\mathrm{meas}(\Omega)\, ||\, |\mathbf{v}|\, ||^2_{L^\infty(\Omega)} \tag{18}$$

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

Before we prove Proposition 3.1, let us state an easy corollary, which is used in the proof of the estimate of Proposition 3.2.

Corollary 3.1 *Let $\mathcal{M}$ be an admissible mesh. If $u_{\mathcal{T}} = (u_K)_{K\in\mathcal{T}}$ is a solution to (8)—(10) and, for $k>0$, $E_k = \{|u_{\mathcal{T}}| > k\}$, then there exists $C \in \mathbb{R}^*_+$ only depending on $(\Omega, \mathbf{v})$ such that*

$$\mathrm{meas}(E_k) \leq \frac{C(1+||\mu||_{M(\overline{\Omega})})}{(\ln(1+k))^2}.$$

Proof of Corollary 3.1
By Proposition 3.1, we get that

$$|| \ln(1+|u_{\mathcal{T}}|)||^2_{1,2,\mathcal{T}} \leq (2 + d\mathrm{meas}(\Omega)\, ||\, |\mathbf{v}|\, ||^2_{L^\infty(\Omega)})(1+||\mu||_{M(\overline{\Omega})}).$$

Therefore, using the discrete Poincaré inequality (Proposition 2.1), we get that there exists $C \in \mathbb{R}^*_+$ only depending on $(\Omega, \mathbf{v})$ such that:

$$|| \ln(1+|u_{\mathcal{T}}|)||^2_{L^2(\Omega)} \leq C(1+||\mu||_{M(\overline{\Omega})}).$$

Finally, since $\mathrm{meas}(E_k) = \mathrm{meas}(\{\ln(1+|u_{\mathcal{T}}|) \geq \ln(1+k)\})$, the Chebyshev inequality yields that $\mathrm{meas}(E_k) \leq \frac{C(1+||\mu||_{M(\overline{\Omega})})}{(\ln(1+k))^2}$. ■

Proof of Proposition 3.1
Step 0: sketch of the proof in the continuous case.
Let $\varphi(s) = \int_0^s \frac{dt}{(1+|t|)^2}$. Suppose that $\mu \in H^{-1}(\Omega)\cap L^1(\Omega)$ and let $u \in H^1_0(\Omega)$ be a variational solution of (1). Using $\varphi(u)$ as a test function in the equation satisfied by u, and since φ is bounded by 1, we find:

$$\int_\Omega \nabla u \cdot \frac{\nabla u}{(1+|u|)^2}\, d\lambda + \int_\Omega bu\varphi(u)\, d\lambda \leq ||\mu||_{L^1(\Omega)} + ||\, |\mathbf{v}|\, ||_{L^\infty(\Omega)} \int_\Omega |u| \frac{|\nabla u|}{(1+|u|)^2}\, d\lambda \leq C + C\int_\Omega \frac{|\nabla u|}{(1+|u|)}\, d\lambda,$$

where C only depends on $||\mu||_{L^1(\Omega)}$ and $\mathbf{v}$. Since $\nabla(\ln(1+|u|)) = \mathrm{sgn}(u)\frac{\nabla u}{(1+|u|)}$ and $bu\varphi(u) \geq 0$ (b is nonnegative and $\varphi(s)$ has the same sign as s), we deduce that

$$||\, |\nabla(\ln(1+|u|))|\, ||^2_{L^2(\Omega)} \leq C + C\mathrm{meas}(\Omega)^{1/2} ||\, |\nabla(\ln(1+|u|))|\, ||_{L^2(\Omega)},$$

which gives an estimate on $||\, |\nabla(\ln(1+|u|))|\, ||_{L^2(\Omega)}$ (and thus, by the Poincaré inequality, also on $||\ln(1+|u|)||_{L^2(\Omega)}$).

Step 1: proof of a first discrete estimate.
Let $\varphi(s) = \int_0^s \frac{dt}{(1+|t|)^2}$. Multiplying each equality of (8) by $\varphi(u_K)$ and summing on $K\in\mathcal{T}$, we have

$$\sum_{K\in\mathcal{T}}\sum_{\sigma\in\mathcal{E}_K} F_{K,\sigma}\varphi(u_K) + \sum_{K\in\mathcal{T}}\sum_{\sigma\in\mathcal{E}_K} v_{K,\sigma}u_{\sigma,+}\varphi(u_K) + \sum_{K\in\mathcal{T}} \mathrm{meas}(K) b_K u_K \varphi(u_K) = \sum_{K\in\mathcal{T}} \mu(K)\varphi(u_K). \tag{19}$$

Gathering by edges and using (9), we can write

$$\sum_{K\in\mathcal{T}}\sum_{\sigma\in\mathcal{E}_K} F_{K,\sigma}\varphi(u_K) = \sum_{\sigma\in\mathcal{E}} \tau_\sigma(u_K - u_L)(\varphi(u_K) - \varphi(u_L)) \tag{20}$$

where we let $\sigma = K|L$ if $\sigma \in \mathcal{E}_{\rm int}$ and $u_L = 0$ if $\sigma \in \mathcal{E}_{\rm ext} \cap \mathcal{E}_K$.
By the conservativity of the fluxes, still gathering by edges, we find

$$\sum_{K\in\mathcal{T}}\sum_{\sigma\in\mathcal{E}_K} v_{K,\sigma}u_{\sigma,+}\varphi(u_K) = \sum_{\sigma\in\mathcal{E}} u_{\sigma,+}v_{K,\sigma}(\varphi(u_K) - \varphi(u_L))$$

(recall that $u_L = 0$ — so that $\varphi(u_L) = 0$ — if $\sigma \in \mathcal{E}_{\rm ext} \cap \mathcal{E}_K$). If $\sigma \in \mathcal{E}$, we denote $v_\sigma = |v_{K,\sigma}|$ for a $K \in \mathcal{T}$ such that $\sigma \in \mathcal{E}_K$ (the definition of v_σ does not depend on the choice of such a K) and $u_{\sigma,-}$ the downstream choice of u, i.e. $u_{\sigma,-}$ is such that $\{u_{\sigma,+}, u_{\sigma,-}\} = \{u_K, u_L\}$ (where $\sigma = K|L$ if $\sigma \in \mathcal{E}_{\rm int}$ and $u_L = 0$ if $\sigma \in \mathcal{E}_{\rm ext} \cap \mathcal{E}_K$).
Let $\sigma \in \mathcal{E}$; if $v_{K,\sigma} \geq 0$, then $u_{\sigma,+} = u_K$ and $u_{\sigma,-} = u_L$ so that $v_{K,\sigma}(\varphi(u_K) - \varphi(u_L)) = v_\sigma(\varphi(u_{\sigma,+}) - \varphi(u_{\sigma,-}))$; if $v_{K,\sigma} < 0$, then $u_{\sigma,+} = u_L$ and $u_{\sigma,-} = u_K$, which gives $v_{K,\sigma}(\varphi(u_K) - \varphi(u_L)) = -v_\sigma(\varphi(u_{\sigma,-}) - \varphi(u_{\sigma,+})) = v_\sigma(\varphi(u_{\sigma,+}) - \varphi(u_{\sigma,-}))$. Thus,

$$\sum_{K\in\mathcal{T}}\sum_{\sigma\in\mathcal{E}_K} v_{K,\sigma}u_{\sigma,+}\varphi(u_K) = \sum_{\sigma\in\mathcal{E}} v_\sigma u_{\sigma,+}(\varphi(u_{\sigma,+}) - \varphi(u_{\sigma,-})). \tag{21}$$

b being nonnegative and $\varphi(s)$ having the same sign as s,

$$\sum_{K\in\mathcal{T}} \mathrm{meas}(K) b_K u_K \varphi(u_K) \geq 0. \tag{22}$$

Since φ is bounded by 1 and $\mathcal{T}$ is a partition of Ω,

$$\left|\sum_{K\in\mathcal{T}} \mu(K)\varphi(u_K)\right| \leq \sum_{K\in\mathcal{T}} |\mu(K)| \leq |\mu|(\Omega) = ||\mu||_{M(\overline{\Omega})}. \tag{23}$$

Using (20), (21), (22) and (23) in (19), we get

$$\sum_{\sigma\in\mathcal{E}} \tau_\sigma(u_K - u_L)(\varphi(u_K) - \varphi(u_L)) \leq ||\mu||_{M(\overline{\Omega})} + \sum_{\sigma\in\mathcal{E}} v_\sigma u_{\sigma,+}(\varphi(u_{\sigma,-}) - \varphi(u_{\sigma,+})). \tag{24}$$

We now study each term of the last sum a little more precisely. We use the fact that φ is nondecreasing.

- If $u_{\sigma,+} \geq u_{\sigma,-}$ and $u_{\sigma,+} \geq 0$, then $\varphi(u_{\sigma,-}) - \varphi(u_{\sigma,+}) \leq 0$ and $u_{\sigma,+}(\varphi(u_{\sigma,-}) - \varphi(u_{\sigma,+})) \leq 0$.
- If $u_{\sigma,+} \geq u_{\sigma,-}$ and $u_{\sigma,+} < 0$, then $0 > u_{\sigma,+} \geq u_{\sigma,-}$, so that $(u_{\sigma,+}, u_{\sigma,-})$ have the same sign and $|u_{\sigma,+}| \leq |u_{\sigma,-}|$.
- If $u_{\sigma,+} < u_{\sigma,-}$ and $u_{\sigma,+} \geq 0$, then $0 \leq u_{\sigma,+} < u_{\sigma,-}$, so that $(u_{\sigma,+}, u_{\sigma,-})$ have the same sign and $|u_{\sigma,+}| \leq |u_{\sigma,-}|$.
- If $u_{\sigma,+} < u_{\sigma,-}$ and $u_{\sigma,+} < 0$, then $\varphi(u_{\sigma,-}) - \varphi(u_{\sigma,+}) \geq 0$ and $u_{\sigma,+}(\varphi(u_{\sigma,-}) - \varphi(u_{\sigma,+})) \leq 0$.

By denoting $\mathcal{A} = \{\sigma \in \mathcal{E} \mid u_{\sigma,+} \geq u_{\sigma,-}\,,\ u_{\sigma,+} < 0\} \cup \{\sigma \in \mathcal{E} \mid u_{\sigma,+} < u_{\sigma,-}\,,\ u_{\sigma,+} \geq 0\}$, we notice thus that, for all $\sigma \in \mathcal{E}\backslash\mathcal{A}$, $v_\sigma u_{\sigma,+}(\varphi(u_{\sigma,-}) - \varphi(u_{\sigma,+})) \leq 0$. This gives

$$\sum_{\sigma\in\mathcal{E}} v_\sigma u_{\sigma,+}(\varphi(u_{\sigma,-}) - \varphi(u_{\sigma,+})) \leq \sum_{\sigma\in\mathcal{A}} v_\sigma u_{\sigma,+}(\varphi(u_{\sigma,-}) - \varphi(u_{\sigma,+})).$$

As $v_\sigma \le \mathrm{meas}(\sigma)|| \, |\mathbf{v}| \, ||_{L^\infty(\Omega)}$, we deduce, using the Cauchy-Schwarz inequality, that

$$\begin{aligned}
&\sum_{\sigma\in\mathcal{E}} v_\sigma u_{\sigma,+}(\varphi(u_{\sigma,-})-\varphi(u_{\sigma,+})) \\
&\quad\le \quad || \, |\mathbf{v}| \, ||_{L^\infty(\Omega)} \sum_{\sigma\in\mathcal{A}} \mathrm{meas}(\sigma)|u_{\sigma,+}||\varphi(u_{\sigma,-})-\varphi(u_{\sigma,+})| \\
&\quad\le \quad || \, |\mathbf{v}| \, ||_{L^\infty(\Omega)} \left(\sum_{\sigma\in\mathcal{A}} \mathrm{meas}(\sigma) d_\sigma\right)^{\frac12} \left(\sum_{\sigma\in\mathcal{A}} \tau_\sigma u_{\sigma,+}^2(\varphi(u_{\sigma,-})-\varphi(u_{\sigma,+}))^2\right)^{\frac12}.
\end{aligned}$$

But $\sum_{\sigma\in\mathcal{A}} \mathrm{meas}(\sigma) d_\sigma \le \sum_{\sigma\in\mathcal{E}} \mathrm{meas}(\sigma) d_\sigma = d\mathrm{meas}(\Omega)$ and, if $\sigma\in\mathcal{A}$, $(u_{\sigma,+}, u_{\sigma,-})$ have the same sign and $|u_{\sigma,+}| \le |u_{\sigma,-}|$, thus, by Lemma 3.1 below and Young's inequality,

$$\begin{aligned}
&\sum_{\sigma\in\mathcal{E}} v_\sigma u_{\sigma,+}(\varphi(u_{\sigma,-})-\varphi(u_{\sigma,+})) \\
&\quad\le \quad (d\mathrm{meas}(\Omega))^{1/2} || \, |\mathbf{v}| \, ||_{L^\infty(\Omega)} \left(\sum_{\sigma\in\mathcal{A}} \tau_\sigma (u_{\sigma,-}-u_{\sigma,+})(\varphi(u_{\sigma,-})-\varphi(u_{\sigma,+}))\right)^{\frac12} \\
&\quad\le \quad \frac12 d\mathrm{meas}(\Omega) \, || \, |\mathbf{v}| \, ||^2_{L^\infty(\Omega)} + \frac12 \sum_{\sigma\in\mathcal{E}} \tau_\sigma (u_{\sigma,-}-u_{\sigma,+})(\varphi(u_{\sigma,-})-\varphi(u_{\sigma,+})).
\end{aligned}$$

For all $\sigma\in\mathcal{E}$, we have $\{u_{\sigma,+}, u_{\sigma,-}\} = \{u_K, u_L\}$, so that $(u_{\sigma,-}-u_{\sigma,+})(\varphi(u_{\sigma,-})-\varphi(u_{\sigma,+})) = (u_K-u_L)(\varphi(u_K)-\varphi(u_L))$. Coming back to (24), we obtain

$$\sum_{\sigma\in\mathcal{E}} \tau_\sigma (u_K-u_L)(\varphi(u_K)-\varphi(u_L)) \le 2||\mu||_{M(\overline{\Omega})} + d\mathrm{meas}(\Omega) \, || \, |\mathbf{v}| \, ||^2_{L^\infty(\Omega)}, \tag{25}$$

which concludes this step.

Step 2: Estimate on $\ln(1+|u_\mathcal{T}|)$.
We notice that, for all $s\in\mathbb{R}$, $\ln(1+|s|) = \int_0^s \frac{\mathrm{sgn}(t)\,dt}{1+|t|}$. Thus, for all $(x,y)\in\mathbb{R}^2$, by the Cauchy-Schwarz inequality and since φ is nondecreasing,

$$\begin{aligned}
(\ln(1+|x|)-\ln(1+|y|))^2 &= \left(\int_y^x \frac{\mathrm{sgn}(t)\,dt}{1+|t|}\right)^2 \\
&\le |x-y| \left|\int_y^x \frac{dt}{(1+|t|)^2}\right| = |x-y||\varphi(x)-\varphi(y)| = (x-y)(\varphi(x)-\varphi(y)).
\end{aligned}$$

Using this upper bound and (25), we deduce the result of the proposition. ■

Let us now state and prove the technical result which was used in Step 1 of the above proof.

Lemma 3.1 *Let $\varphi(s) = \int_0^s \frac{dt}{(1+|t|)^2}$. If $(x,y)\in\mathbb{R}^2$ have the same sign and $|x|\le|y|$, then*

$$x^2(\varphi(y)-\varphi(x))^2 \le (y-x)(\varphi(y)-\varphi(x)). \tag{26}$$

Proof of Lemma 3.1
Since φ is C^1-continuous on $\mathbb{R}$, there exists $\theta\in[x,y]$ such that $\varphi(y)-\varphi(x) = \varphi'(\theta)(y-x)$, so that, since φ is nondecreasing,

$$\begin{aligned}
x^2(\varphi(y)-\varphi(x))^2 &\le \frac{x^2}{(1+|\theta|)^2}|y-x| \, |\varphi(y)-\varphi(x)| \\
&\le \frac{x^2}{(1+|\theta|)^2}(y-x)(\varphi(y)-\varphi(x)).
\end{aligned}$$

But $|x| \leq |y|$ and x and y have the same sign, so that, since $\theta \in [x, y]$, we have $|\theta| \geq |x|$, and (26) is thus a consequence of the previous inequality. ■

3.2 Estimate on $||u_{\mathcal{T}}||_{1,q,\mathcal{M}}$

We denote, for $k > 0$, $T_k(s) = \max(-k, \min(s, k))$ and $S_k(s) = s - T_k(s)$ (see Figure 2).

Proposition 3.2 *Let $\mathcal{M}$ be an admissible mesh and $\zeta > 0$ satisfying (15). We suppose that μ satisfies $||\mu||_{M(\overline{\Omega})} \leq 1$. Then there exists $k_0 > 0$ only depending on $(\Omega, \mathbf{v}, \zeta)$ and, for all $m \in (1, 2)$, $C > 0$ only depending on $(\Omega, \mathbf{v}, m, \zeta)$ such that, if $u_{\mathcal{T}}$ is a solution to (8)—(10) and $\varphi_m(s) = \int_0^s \frac{dt}{(1+|t|)^m}$, we have*

$$\sum_{\sigma \in \mathcal{E}} \tau_\sigma (S_{k_0}(u_K) - S_{k_0}(u_L))(\varphi_m(S_{k_0}(u_K)) - \varphi_m(S_{k_0}(u_L))) \leq C \tag{27}$$

and

$$\sum_{\sigma \in \mathcal{E}} \tau_\sigma (T_{k_0}(u_K) - T_{k_0}(u_L))(\varphi_m(T_{k_0}(u_K)) - \varphi_m(T_{k_0}(u_L))) \leq C, \tag{28}$$

where we let $\sigma = K|L$ if $\sigma \in \mathcal{E}_{\rm int}$ and $u_L = 0$ if $\sigma \in \mathcal{E}_{\rm ext} \cap \mathcal{E}_K$.

Remark 3.1 *Problem (8)—(10) being linear, there is no loss of generality in the estimate if we consider measures of norm less than* 1*, as we will see in Theorem 2.2*

Proof of Proposition 3.2
Step 0: sketch of the estimate in the continuous case.
Suppose that $u \in H^1_0(\Omega)$ is a variational solution of (1) with $\mu \in H^{-1}(\Omega) \cap L^1(\Omega)$ satisfying $||\mu||_{L^1(\Omega)} \leq 1$, and take $\varphi_m(S_k(u))$ as a test function in (14). Using the fact that $bu\varphi_m(S_k(u)) \geq 0$ (b is nonnegative and $\varphi_m(s)$ and $S_k(s)$ have the same sign as s), that $\nabla(S_k(u)) = \nabla u$ where $\nabla(S_k(u)) \neq 0$ and that φ_m is bounded by $1/(m-1)$, we have

$$\int_\Omega \frac{|\nabla(S_k(u))|^2}{(1+|S_k(u)|)^m}\, d\lambda \leq \frac{1}{m-1} + ||\,|\mathbf{v}|\,||_{L^\infty(\Omega)} \int_\Omega |u| \frac{|\nabla(S_k(u))|}{(1+|S_k(u)|)^m}\, d\lambda.$$

But $|u| \leq k + |S_k(u)|$ and $(1+|S_k(u)|)^{2m} \geq (1+|S_k(u)|)^m$, so that, by the Cauchy-Schwarz inequality,

$$\begin{aligned}
&\int_\Omega \frac{|\nabla(S_k(u))|^2}{(1+|S_k(u)|)^m}\, d\lambda \\
&\quad \leq \frac{1}{m-1} + C_1 k \left(\int_\Omega \frac{|\nabla(S_k(u))|^2}{(1+|S_k(u)|)^{2m}}\, d\lambda \right)^{\frac{1}{2}} + C_1 \int_\Omega \frac{|S_k(u)|}{(1+|S_k(u)|)^{\frac{m}{2}}} \frac{|\nabla(S_k(u))|}{(1+|S_k(u)|)^{\frac{m}{2}}}\, d\lambda \\
&\quad \leq \frac{1}{m-1} + C_1 k \left(\int_\Omega \frac{|\nabla(S_k(u))|^2}{(1+|S_k(u)|)^m}\, d\lambda \right)^{\frac{1}{2}} + C_1 ||\psi(S_k(u))||_{L^2(\Omega)} \left(\int_\Omega \frac{|\nabla(S_k(u))|^2}{(1+|S_k(u)|)^m}\, d\lambda \right)^{\frac{1}{2}},
\end{aligned} \tag{29}$$

where C_1 only depends on $(\Omega, \mathbf{v})$ and $\psi(s) = \frac{|s|}{(1+|s|)^{\frac{m}{2}}}$.
Now, by the Hölder inequality and the Sobolev injection, and since $\psi(S_k(u)) = 0$ outside $E_k = \{|u| > k\}$, there exists $r > 2$ only depending on d, and C_2 only depending on (Ω, r) (notice that a dependence on Ω takes into account a dependence on d), such that

$$||\psi(S_k(u))||_{L^2(\Omega)} \leq \mathrm{meas}(E_k)^{\frac{1}{2}-\frac{1}{r}} ||\psi(S_k(u))||_{L^r(\Omega)} \leq C_2 \mathrm{meas}(E_k)^{\frac{1}{2}-\frac{1}{r}} ||\,|\nabla(\psi(S_k(u)))|\,||_{L^2(\Omega)}. \tag{30}$$

Since $|\psi'(s)| \leq \frac{1+\frac{m}{2}}{(1+|s|)^{\frac{m}{2}}} \leq \frac{2}{(1+|s|)^{\frac{m}{2}}}$, one has

$$||\,|\nabla(\psi(S_k(u)))|\,||_{L^2(\Omega)} \leq 2 \left(\int_\Omega \frac{|\nabla(S_k(u))|^2}{(1+|S_k(u)|)^m}\, d\lambda \right)^{\frac{1}{2}}. \tag{31}$$

Gathering (29), (30) and (31), we find C_3 only depending on $(\Omega, \mathbf{v})$ such that

$$\begin{aligned}&\int_\Omega \frac{|\nabla(S_k(u))|^2}{(1+|S_k(u)|)^m}\,d\lambda \\ &\quad\le \frac{C_1}{m-1} + C_1 k\left(\int_\Omega \frac{|\nabla(S_k(u))|^2}{(1+|S_k(u)|)^m}\,d\lambda\right)^{\frac12} + C_3\mathrm{meas}(E_k)^{\frac12-\frac1r}\int_\Omega \frac{|\nabla(S_k(u))|^2}{(1+|S_k(u)|)^m}\,d\lambda.\end{aligned} \tag{32}$$

Thanks to a continuous equivalent of Corollary 3.1, there exists C_4 only depending on $(\Omega, \mathbf{v})$ such that $\mathrm{meas}(E_k) \le \frac{C_4}{(\ln(1+k))^2}$. Thus, there exists $k_0 > 0$ only depending on (C_4, C_3, r) (i.e. on $(\Omega, \mathbf{v})$) such that $C_3\mathrm{meas}(E_{k_0})^{\frac12-\frac1r} \le \frac12$. Applying (32) to this k_0 gives

$$\int_\Omega \frac{|\nabla(S_{k_0}(u))|^2}{(1+|S_{k_0}(u)|)^m}\,d\lambda \le C_5$$

where C_5 only depends on $(\Omega, \mathbf{v}, m)$, which is the continuous equivalent of (27).

The estimate on $T_{k_0}(u)$ is quite simple and well known (see [2]). Take $\varphi_m(T_{k_0}(u))$ as a test function in the equation satisfied by u; since $\nabla(T_{k_0}(u)) = 0$ outside $\{|u| \le k_0\}$ and $(1+|T_{k_0}(u)|)^{2m} \ge (1+|T_{k_0}(u)|)^m$, we find

$$\begin{aligned}\int_\Omega \frac{|\nabla(T_{k_0}(u))|^2}{(1+|T_{k_0}(u)|)^m}\,d\lambda &\le \frac{1}{m-1} + ||\,|\mathbf{v}|\,||_{L^\infty(\Omega)}\int_{\{|u|\le k_0\}} |u|\frac{|\nabla(T_{k_0}(u))|}{(1+|T_{k_0}(u)|)^m}\,d\lambda \\ &\le \frac{1}{m-1} + ||\,|\mathbf{v}|\,||_{L^\infty(\Omega)} k_0 \mathrm{meas}(\Omega)^{1/2}\left(\int_\Omega \frac{|\nabla(T_{k_0}(u))|^2}{(1+|T_{k_0}(u)|)^m}\,d\lambda\right)^{\frac12}.\end{aligned}$$

This gives an estimate on $T_{k_0}(u)$ which is the continuous equivalent of (28).

Step 1: estimate on $S_k(u_\mathcal{T})$.
Let $\mathcal{M}$ be an admissible mesh and take $u_\mathcal{T}$ a solution of (8)—(10). Multiplying each equation of (8) by $\varphi_m(S_k(u_K))$, summing on $K \in \mathcal{T}$ and gathering by edges, we find

$$\begin{aligned}&\sum_{\sigma\in\mathcal{E}} \tau_\sigma(u_K - u_L)(\varphi_m(S_k(u_K)) - \varphi_m(S_k(u_L))) + \sum_{K\in\mathcal{T}} \mathrm{meas}(K) b_K u_K \varphi_m(S_k(u_K)) \\ &\quad= \sum_{K\in\mathcal{T}} \mu(K)\varphi_m(S_k(u_K)) - \sum_{\sigma\in\mathcal{E}} v_{K,\sigma} u_{\sigma,+}(\varphi_m(S_k(u_K)) - \varphi_m(S_k(u_L)))\end{aligned} \tag{33}$$

(recall that, if $\sigma \in \mathcal{E}_{\mathrm{int}}$, we use the notation $\sigma = K|L$ and, if $\sigma \in \mathcal{E}_{\mathrm{ext}} \cap \mathcal{E}_K$, we set $u_L = 0$).
The function φ_m is bounded by $\frac{1}{m-1}$ and $\mathcal{T}$ is a partition of Ω, so that

$$\left|\sum_{K\in\mathcal{T}} \mu(K)\varphi_m(S_k(u_K))\right| \le \frac{1}{m-1}\sum_{K\in\mathcal{T}} |\mu(K)| \le \frac{||\mu||_{M(\overline{\Omega})}}{m-1} \le \frac{1}{m-1} \tag{34}$$

We again denote $u_{\sigma,-}$ the downstream choice of u_σ (i.e. $u_{\sigma,-} = u_L$ if $v_{K,\sigma} \ge 0$ and $u_{\sigma,-} = u_K$ otherwise) and $v_\sigma = |v_{K,\sigma}|$ (for a $K \in \mathcal{T}$ such that $\sigma \in \mathcal{E}_K$); we have then:

$$-\sum_{\sigma\in\mathcal{E}} v_{K,\sigma} u_{\sigma,+}(\varphi_m(S_k(u_K)) - \varphi_m(S_k(u_L))) = \sum_{\sigma\in\mathcal{E}} v_\sigma u_{\sigma,+}(\varphi_m(S_k(u_{\sigma,-})) - \varphi_m(S_k(u_{\sigma,+}))).$$

But, as in the proof of Proposition 3.1 (because $\varphi_m \circ S_k$ is nondecreasing), we have $u_{\sigma,+}(\varphi_m(S_k(u_{\sigma,-})) - \varphi_m(S_k(u_{\sigma,+}))) \le 0$ if $\sigma \notin \mathcal{A}$, where $\mathcal{A} = \{\sigma \in \mathcal{E} \mid u_{\sigma,+} \ge u_{\sigma,-}\ ,\ u_{\sigma,+} < 0\} \cup \{\sigma \in \mathcal{E} \mid u_{\sigma,+} < u_{\sigma,-}\ ,\ u_{\sigma,+} \ge$

0}. Thus,

$$\begin{aligned}
&-\sum_{\sigma\in\mathcal{E}} v_{K,\sigma}u_{\sigma,+}(\varphi_m(S_k(u_K))-\varphi_m(S_k(u_L)))\\
&\quad\le \sum_{\sigma\in\mathcal{A}} v_\sigma u_{\sigma,+}(\varphi_m(S_k(u_{\sigma,-}))-\varphi_m(S_k(u_{\sigma,+})))\\
&\quad\le ||\,|\mathbf{v}|\,||_{L^\infty(\Omega)}\sum_{\sigma\in\mathcal{A}}\mathrm{meas}(\sigma)|u_{\sigma,+}|\,|\varphi_m(S_k(u_{\sigma,-}))-\varphi_m(S_k(u_{\sigma,+}))|. \qquad (35)
\end{aligned}$$

Let $a_{k,\sigma}=\int_0^1\varphi_m'(S_k(u_{\sigma,+})+t(S_k(u_{\sigma,-})-S_k(u_{\sigma,+})))\,dt\ge 0$, so that

$$\varphi_m(S_k(u_{\sigma,-}))-\varphi_m(S_k(u_{\sigma,+}))=a_{k,\sigma}(S_k(u_{\sigma,-})-S_k(u_{\sigma,+})). \qquad (36)$$

We can write

$$\begin{aligned}
&\sum_{\sigma\in\mathcal{A}}\mathrm{meas}(\sigma)|u_{\sigma,+}|\,|\varphi_m(S_k(u_{\sigma,-}))-\varphi_m(S_k(u_{\sigma,+}))|\\
&\quad=\sum_{\sigma\in\mathcal{A}}\mathrm{meas}(\sigma)a_{k,\sigma}^{1/2}|u_{\sigma,+}|a_{k,\sigma}^{1/2}|S_k(u_{\sigma,-})-S_k(u_{\sigma,+})|\\
&\quad\le\left(\sum_{\sigma\in\mathcal{A}}\mathrm{meas}(\sigma)d_\sigma a_{k,\sigma}u_{\sigma,+}^2\right)^{\frac12}\left(\sum_{\sigma\in\mathcal{A}}\tau_\sigma a_{k,\sigma}(S_k(u_{\sigma,-})-S_k(u_{\sigma,+}))^2\right)^{\frac12}.
\end{aligned}$$

But, by (36), $a_{k,\sigma}(S_k(u_{\sigma,-})-S_k(u_{\sigma,+}))^2=(S_k(u_{\sigma,-})-S_k(u_{\sigma,+}))(\varphi_m(S_k(u_{\sigma,-}))-\varphi_m(S_k(u_{\sigma,+})))$, so that

$$\begin{aligned}
&\sum_{\sigma\in\mathcal{A}}\mathrm{meas}(\sigma)|u_{\sigma,+}|\,|\varphi_m(S_k(u_{\sigma,-}))-\varphi_m(S_k(u_{\sigma,+}))|\\
&\quad\le\left(\sum_{\sigma\in\mathcal{A}}\mathrm{meas}(\sigma)d_\sigma a_{k,\sigma}u_{\sigma,+}^2\right)^{\frac12}\\
&\qquad\times\left(\sum_{\sigma\in\mathcal{A}}\tau_\sigma(S_k(u_{\sigma,-})-S_k(u_{\sigma,+}))(\varphi_m(S_k(u_{\sigma,-}))-\varphi_m(S_k(u_{\sigma,+})))\right)^{\frac12}. \qquad (37)
\end{aligned}$$

Moreover, for all $\sigma\in\mathcal{A}$, $u_{\sigma,+}$ and $u_{\sigma,-}$ have the same sign and $|u_{\sigma,+}|\le|u_{\sigma,-}|$. Thus, for such σ, $(S_k(u_{\sigma,+}),S_k(u_{\sigma,-}))$ have the same sign and $|S_k(u_{\sigma,+})|\le|S_k(u_{\sigma,-})|$ and, by Lemma 3.2 stated after this proof, we deduce that

$$a_{k,\sigma}\le\frac{1}{(1+|S_k(u_{\sigma,+})|)^m}\le 1.$$

Since $|u_{\sigma,+}|\le k+|S_k(u_{\sigma,+})|$, we deduce that

$$a_{k,\sigma}u_{\sigma,+}^2\le 2k^2+2\frac{|S_k(u_{\sigma,+})|^2}{(1+|S_k(u_{\sigma,+})|)^m},$$

which gives, in (37), using $\sum_{\sigma\in\mathcal{A}}\mathrm{meas}(\sigma)d_\sigma\le\sum_{\sigma\in\mathcal{E}}\mathrm{meas}(\sigma)d_\sigma=d\mathrm{meas}(\Omega)$ and $(\alpha+\beta)^{1/2}\le\alpha^{1/2}+\beta^{1/2}$ for all nonnegative (α,β),

$$\begin{aligned}
&\sum_{\sigma\in\mathcal{A}}\mathrm{meas}(\sigma)|u_{\sigma,+}|\,|\varphi_m(S_k(u_{\sigma,-}))-\varphi_m(S_k(u_{\sigma,+}))|\\
&\quad\le\sqrt{2d\mathrm{meas}(\Omega)}kA_k+\sqrt{2}A_k\left(\sum_{\sigma\in\mathcal{A}}\mathrm{meas}(\sigma)d_\sigma\psi(S_k(u_{\sigma,+}))^2\right)^{\frac12}, \qquad (38)
\end{aligned}$$

where $\psi(s) = \frac{|s|}{(1+|s|)^{\frac{m}{2}}}$ and

$$\begin{aligned} A_k &= \left(\sum_{\sigma\in\mathcal{E}} \tau_\sigma (S_k(u_{\sigma,-}) - S_k(u_{\sigma,+}))(\varphi_m(S_k(u_{\sigma,-})) - \varphi_m(S_k(u_{\sigma,+})))\right)^{\frac{1}{2}} \\ &= \left(\sum_{\sigma\in\mathcal{E}} \tau_\sigma (S_k(u_K) - S_k(u_L))(\varphi_m(S_k(u_K)) - \varphi_m(S_k(u_L)))\right)^{\frac{1}{2}} \end{aligned}$$

(recall that $\sigma = K|L$ if $\sigma \in \mathcal{E}_{\rm int}$, that $u_L = 0$ if $\sigma \in \mathcal{E}_{\rm ext} \cap \mathcal{E}_K$ and that $\{u_{\sigma,+}, u_{\sigma,-}\} = \{u_K, u_L\}$ for all $\sigma \in \mathcal{E}$).
We have, since $d_{K,\sigma} \geq \zeta d_\sigma$ for all $K \in \mathcal{T}$ and all $\sigma \in \mathcal{E}_K$,

$$\begin{aligned} \sum_{\sigma\in\mathcal{A}} {\rm meas}(\sigma) d_\sigma \psi(S_k(u_{\sigma,+}))^2 &\leq \sum_{K\in\mathcal{T}} \psi(S_k(u_K))^2 \left(\sum_{\sigma\in\mathcal{A}\cap\mathcal{E}_K \mid v_{K,\sigma}\geq 0} {\rm meas}(\sigma) d_\sigma\right) \\ &\leq \frac{1}{\zeta} \sum_{K\in\mathcal{T}} \psi(S_k(u_K))^2 \left(\sum_{\sigma\in\mathcal{E}_K} {\rm meas}(\sigma) d_{K,\sigma}\right) \\ &= \frac{1}{\zeta} \sum_{K\in\mathcal{T}} \psi(S_k(u_K))^2 \times d{\rm meas}(K) = \frac{d}{\zeta} ||\psi(S_k(u_\mathcal{T}))||^2_{L^2(\Omega)}. \end{aligned}$$

By Proposition 2.2, and since $\psi(S_k(u_\mathcal{T})) = 0$ outside $E_k = \{|u_\mathcal{T}| > k\}$, we can thus find $r > 2$ and $C_1 > 0$ only depending on (Ω, ζ) such that

$$\left(\sum_{\sigma\in\mathcal{A}} {\rm meas}(\sigma) d_\sigma \psi(S_k(u_{\sigma,+}))^2\right)^{\frac{1}{2}} \leq C_1 {\rm meas}(E_k)^{\frac{1}{2}-\frac{1}{r}} ||\psi(S_k(u_\mathcal{T}))||_{1,2,\mathcal{M}}.$$

But, by Lemma 3.3 below and the definition of A_k,

$$||\psi(S_k(u_\mathcal{T}))||^2_{1,2,\mathcal{M}} = \sum_{\sigma\in\mathcal{E}} \tau_\sigma (\psi(S_k(u_K)) - \psi(S_k(u_L))^2 \leq 4A_k^2,$$

so that

$$\left(\sum_{\sigma\in\mathcal{A}} {\rm meas}(\sigma) d_\sigma \psi(S_k(u_{\sigma,+}))^2\right)^{\frac{1}{2}} \leq 2C_1 A_k {\rm meas}(E_k)^{\frac{1}{2}-\frac{1}{r}}.$$

Returning to (38), we thus find

$$\sum_{\sigma\in\mathcal{A}} {\rm meas}(\sigma) |u_{\sigma,+}| \, |\varphi_m(S_k(u_{\sigma,-})) - \varphi_m(S_k(u_{\sigma,+}))| \leq \sqrt{2d{\rm meas}(\Omega)} k A_k + 2\sqrt{2} C_1 {\rm meas}(E_k)^{\frac{1}{2}-\frac{1}{r}} A_k^2. \quad (39)$$

(33), (34), (35), (39) and the fact that $b_K u_K \varphi_m(S_k(u_K)) \geq 0$ then give

$$\begin{aligned} &\sum_{\sigma\in\mathcal{E}} \tau_\sigma (u_K - u_L)(\varphi_m(S_k(u_K)) - \varphi_m(S_k(u_L))) \\ &\quad\leq \frac{1}{m-1} + ||\,|\mathbf{v}|\,||_{L^\infty(\Omega)} \sqrt{2d{\rm meas}(\Omega)} k A_k + 2\sqrt{2} ||\,|\mathbf{v}|\,||_{L^\infty(\Omega)} C_1 {\rm meas}(E_k)^{\frac{1}{2}-\frac{1}{r}} A_k^2 \\ &\quad\leq \frac{1}{m-1} + C_2 k^2 + \frac{1}{2} A_k^2 + C_2 {\rm meas}(E_k)^{\frac{1}{2}-\frac{1}{r}} A_k^2, \end{aligned} \quad (40)$$

where C_2 only depends on $(\Omega, \mathbf{v}, \zeta)$. But φ_m and S_k are nondecreasing and S_k is Lipschitz-continuous with Lipschitz constant 1 so that

$$(S_k(u_K) - S_k(u_L))(\varphi_m(S_k(u_K)) - \varphi_m(S_k(u_L))) \leq (u_K - u_L)(\varphi_m(S_k(u_K)) - \varphi_m(S_k(u_L)))$$

and (40) gives

$$\sum_{\sigma\in\mathcal{E}}\tau_\sigma(S_k(u_K)-S_k(u_L))(\varphi_m(S_k(u_K))-\varphi_m(S_k(u_L)))\le\frac{2}{m-1}+2C_2k^2$$
$$+2C_2\mathrm{meas}(E_k)^{\frac12-\frac1r}\sum_{\sigma\in\mathcal{E}}\tau_\sigma(S_k(u_K)-S_k(u_L))(\varphi_m(S_k(u_K))-\varphi_m(S_k(u_L))). \quad (41)$$

By Corollary 3.1, there exists $k_0>0$ only depending on $(\Omega,\mathbf{v},C_2,r)$ (i.e. only depending on $(\Omega,\mathbf{v},\zeta)$) such that $2C_2\mathrm{meas}(E_k)^{\frac12-\frac1r}\le\frac12$. We deduce from (41) that

$$\sum_{\sigma\in\mathcal{E}}\tau_\sigma(S_{k_0}(u_K)-S_{k_0}(u_L))(\varphi_m(S_{k_0}(u_K))-\varphi_m(S_{k_0}(u_L)))\le\frac{4}{m-1}+4C_2k_0^2,$$

which gives (27).

Step 2: Estimate on $T_{k_0}(u_\mathcal{T})$.
Multiplying each equation of (8) by $\varphi_m(T_{k_0}(u_K))$, summing on $K\in\mathcal{T}$ and re-ordering the sums on the edges, we find

$$\begin{aligned}&\sum_{\sigma\in\mathcal{E}}\tau_\sigma(u_K-u_L)(\varphi_m(T_{k_0}(u_K))-\varphi_m(T_{k_0}(u_L)))+\sum_{K\in\mathcal{T}}\mathrm{meas}(K)b_Ku_K\varphi_m(T_{k_0}(u_K))\\ &=\sum_{K\in\mathcal{T}}\mu(K)\varphi_m(T_{k_0}(u_K))-\sum_{\sigma\in\mathcal{E}}v_{K,\sigma}u_{\sigma,+}(\varphi_m(T_{k_0}(u_K))-\varphi_m(T_{k_0}(u_L))).\end{aligned}\quad(42)$$

As before, we have

$$\left|\sum_{K\in\mathcal{T}}\mu(K)\varphi_m(T_{k_0}(u_K))\right|\le\frac{1}{m-1}\quad(43)$$

and, with the previous notations,

$$\begin{aligned}-\sum_{\sigma\in\mathcal{E}}v_{K,\sigma}u_{\sigma,+}(\varphi_m(T_{k_0}(u_K))-\varphi_m(T_{k_0}(u_L)))&=\sum_{\sigma\in\mathcal{E}}v_\sigma u_{\sigma,+}(\varphi_m(T_{k_0}(u_{\sigma,-}))-\varphi_m(T_{k_0}(u_{\sigma,+})))\\&\le\sum_{\sigma\in\mathcal{A}}v_\sigma u_{\sigma,+}(\varphi_m(T_{k_0}(u_{\sigma,-}))-\varphi_m(T_{k_0}(u_{\sigma,+}))).\end{aligned}$$

If $\sigma\in\mathcal{A}$, then $0\le u_{\sigma,+}\le u_{\sigma,-}$ or $u_{\sigma,-}\le u_{\sigma,+}\le0$. In either case, if $|u_{\sigma,+}|\ge k_0$, then $T_{k_0}(u_{\sigma,+})=T_{k_0}(u_{\sigma,-})$, so that $\varphi_m(T_{k_0}(u_{\sigma,-}))-\varphi_m(T_{k_0}(u_{\sigma,+}))=0$. Thus, in the previous sum, we can suppress the terms $\sigma\in\mathcal{A}$ such that $|u_{\sigma,+}|\ge k_0$ and we have

$$\begin{aligned}&-\sum_{\sigma\in\mathcal{E}}v_{K,\sigma}u_{\sigma,+}(\varphi_m(T_{k_0}(u_K))-\varphi_m(T_{k_0}(u_L)))\\&\le k_0\sum_{\sigma\in\mathcal{A}}v_\sigma|\varphi_m(T_{k_0}(u_{\sigma,-}))-\varphi_m(T_{k_0}(u_{\sigma,+}))|\\&\le k_0||\,|\mathbf{v}|\,||_{L^\infty(\Omega)}\left(\sum_{\sigma\in\mathcal{E}}\mathrm{meas}(\sigma)d_\sigma\right)^{\frac12}\left(\sum_{\sigma\in\mathcal{E}}\tau_\sigma(\varphi_m(T_{k_0}(u_{\sigma,-}))-\varphi_m(T_{k_0}(u_{\sigma,+})))^2\right)^{\frac12}.\end{aligned}$$

φ_m and T_{k_0} are nondecreasing and φ_m is Lipschitz-continuous with Lipschitz constant 1, thus, for all $\sigma\in\mathcal{E}$,

$$\begin{aligned}(\varphi_m(T_{k_0}(u_{\sigma,-}))-\varphi_m(T_{k_0}(u_{\sigma,+})))^2&\le(T_{k_0}(u_{\sigma,-})-T_{k_0}(u_{\sigma,+}))(\varphi_m(T_{k_0}(u_{\sigma,-}))-\varphi_m(T_{k_0}(u_{\sigma,+})))\\&=(T_{k_0}(u_K)-T_{k_0}(u_L))(\varphi_m(T_{k_0}(u_K))-\varphi_m(T_{k_0}(u_L))).\end{aligned}$$

Using this inequality and the fact that $\sum_{\sigma\in\mathcal{E}} \mathrm{meas}(\sigma)d_\sigma = d\mathrm{meas}(\Omega)$, we find

$$
\begin{aligned}
&-\sum_{\sigma\in\mathcal{E}} v_{K,\sigma}u_{\sigma,+}(\varphi_m(T_{k_0}(u_K)) - \varphi_m(T_{k_0}(u_L))) \\
&\quad\leq\quad k_0||\,|\mathbf{v}|\,||_{L^\infty(\Omega)}\sqrt{d\mathrm{meas}(\Omega)}\left(\sum_{\sigma\in\mathcal{E}}\tau_\sigma(T_{k_0}(u_K)-T_{k_0}(u_L))(\varphi_m(T_{k_0}(u_K))-\varphi_m(T_{k_0}(u_L)))\right)^{\frac{1}{2}}. \quad (44)
\end{aligned}
$$

Since φ_m and T_{k_0} are nondecreasing and T_{k_0} is Lipschitz-continuous with Lipschitz constant 1, we have

$$(T_{k_0}(u_K)-T_{k_0}(u_L))(\varphi_m(T_{k_0}(u_K))-\varphi_m(T_{k_0}(u_L))) \leq (u_K-u_L)(\varphi_m(T_{k_0}(u_K))-\varphi_m(T_{k_0}(u_L))).$$

Combined with (42), (43), (44) and the fact that $b_K u_K \varphi_m(T_{k_0}(u_K)) \geq 0$, this inequality gives

$$
\begin{aligned}
&\sum_{\sigma\in\mathcal{E}}\tau_\sigma(T_{k_0}(u_K)-T_{k_0}(u_L))(\varphi_m(T_{k_0}(u_K))-\varphi_m(T_{k_0}(u_L))) \\
&\quad\leq\quad \frac{1}{m-1}+k_0||\,|\mathbf{v}|\,||_{L^\infty(\Omega)}\sqrt{d\mathrm{meas}(\Omega)}\left(\sum_{\sigma\in\mathcal{E}}\tau_\sigma(T_{k_0}(u_K)-T_{k_0}(u_L))(\varphi_m(T_{k_0}(u_K))-\varphi_m(T_{k_0}(u_L)))\right)^{\frac{1}{2}}
\end{aligned}
$$

from which we deduce (28). ■

There remains to state and prove the two technical lemmas which were used in Step 1 of the above proof.

Lemma 3.2 *Let $m \in (1,2)$ and $\varphi_m(s) = \int_0^s \frac{dt}{(1+|t|)^m}$. If (x,y) have the same sign and $|x| \leq |y|$, then*

$$\int_0^1 \varphi_m'(x+t(y-x))\,dt \leq \frac{1}{(1+|x|)^m}.$$

Proof of Lemma 3.2
Suppose that $0 \leq x \leq y$. Then, for all $t \in [0,1]$, $0 \leq x \leq x+t(y-x)$, so that $\varphi_m'(x+t(y-x)) = \frac{1}{(1+(x+t(y-x)))^m} \leq \frac{1}{(1+|x|)^m}$. Integrating this relation on $[0,1]$ gives the desired inequality. If $y \leq x \leq 0$, we use the fact that φ_m' is even and apply the previous result to $(-x,-y)$. ■

Lemma 3.3 *Let $m \in (1,2)$, $\varphi_m(s) = \int_0^s \frac{dt}{(1+|t|)^m}$ and $\psi(s) = \frac{|s|}{(1+|s|)^{\frac{m}{2}}}$. Then for all $(x,y) \in \mathbb{R}^2$, one has*

$$(\psi(x)-\psi(y))^2 \leq 4(x-y)(\varphi_m(x)-\varphi_m(y)).$$

Proof of Lemma 3.3
The function ψ is Lipschitz-continuous and, for all $s \in \mathbb{R}$,

$$|\psi'(s)| = \left|\frac{\mathrm{sgn}(s)}{(1+|s|)^{\frac{m}{2}}} - \frac{\frac{m}{2}\mathrm{sgn}(s)|s|}{(1+|s|)^{1+\frac{m}{2}}}\right| \leq \frac{1+\frac{m}{2}}{(1+|s|)^{\frac{m}{2}}} \leq \frac{2}{(1+|s|)^{\frac{m}{2}}},$$

so that, for all $(x,y) \in \mathbb{R}^2$, by the Cauchy-Schwarz inequality,

$$|\psi(x)-\psi(y)| = \left|\int_y^x \psi'(s)\,ds\right| \leq \left|\int_y^x \frac{4\,ds}{(1+|s|)^m}\right|^{1/2} |x-y|^{1/2} \leq 2|\varphi_m(x)-\varphi_m(y)|^{1/2}|x-y|^{1/2}.$$

Taking the power 2 of this inequality and using the fact that φ_m is nondecreasing, we deduce the desired inequality. ■

We shall now deduce the key estimate on $u_\mathcal{T}$ (Theorem 2.2) from Proposition 3.2 and the following lemma.

Lemma 3.4 *Let $\mathcal{M}$ be an admissible mesh and $\zeta > 0$ satisfying (15). Let $F : (1,2) \to \mathbb{R}^+$ be a function. For $m \in (1,2)$, we denote $\varphi_m(s) = \int_0^s \frac{dt}{(1+|t|)^m}$. If $v_{\mathcal{T}} = (v_K)_{K\in\mathcal{T}} \in X(\mathcal{T})$ satisfies, for all $m \in (1,2)$,*

$$\sum_{\sigma\in\mathcal{E}} \tau_\sigma (v_K - v_L)(\varphi_m(v_K) - \varphi_m(v_L)) \le F(m)$$

(where we have denoted, as usual, $\sigma = K|L$ if $\sigma \in \mathcal{E}_{\rm int}$ and $u_L = 0$ if $\sigma \in \mathcal{E}_{\rm ext} \cap \mathcal{E}_K$), then, for all $q \in [1, \frac{d}{d-1})$, there exists $C > 0$ only depending on (Ω, ζ, F, q) such that $||v_{\mathcal{T}}||_{1,q,\mathcal{M}} \le C$.

Proof of Lemma 3.4
Let $q \in [1, \frac{d}{d-1})$.
Take $m \in (1,2)$ (fixed later on as a function of d and q) and denote $a_{m,\sigma} = \int_0^1 \varphi_m'(v_K + t(v_L - v_K))\, dt$. We have $\varphi_m(v_K) - \varphi_m(v_L) = (v_K - v_L)a_{m,\sigma}$, so that

$$\sum_{\sigma\in\mathcal{E}} \tau_\sigma a_{m,\sigma} (D_\sigma v_{\mathcal{T}})^2 \le F(m).$$

By Hölder's inequality, we have, since $1 \le q < 2$,

$$\begin{aligned}\sum_{\sigma\in\mathcal{E}} \mathrm{meas}(\sigma) d_\sigma \left(\frac{D_\sigma v_{\mathcal{T}}}{d_\sigma}\right)^q &\le \left(\sum_{\sigma\in\mathcal{E}} \mathrm{meas}(\sigma) d_\sigma a_{m,\sigma} \left(\frac{D_\sigma v_{\mathcal{T}}}{d_\sigma}\right)^2\right)^{\frac{q}{2}} \left(\sum_{\sigma\in\mathcal{E}} \mathrm{meas}(\sigma) d_\sigma a_{m,\sigma}^{-\frac{q}{2-q}}\right)^{\frac{2-q}{2}} \\ &\le F(m)^{\frac{q}{2}} \left(\sum_{\sigma\in\mathcal{E}} \mathrm{meas}(\sigma) d_\sigma a_{m,\sigma}^{-\frac{q}{2-q}}\right)^{\frac{2-q}{2}}. \qquad (45)\end{aligned}$$

For all $(x,y) \in \mathbb{R}^2$ and all $t \in [0,1]$, one has $|x + t(y-x)| \le \sup(|x|, |y|)$, so that

$$\varphi_m'(x + t(y-x)) = \frac{1}{(1 + |x + t(y-x)|)^m} \ge \frac{1}{(1 + \sup(|x|,|y|))^m} \ge \inf\left(\frac{1}{(1+|x|)^m}, \frac{1}{(1+|y|)^m}\right).$$

Taking $x = v_K$, $y = v_L$ and integrating the previous inequality on $t \in [0,1]$, we find

$$a_{m,\sigma} \ge \inf\left(\frac{1}{(1+|v_K|)^m}, \frac{1}{(1+|v_L|)^m}\right),$$

which implies

$$a_{m,\sigma}^{-\frac{q}{2-q}} \le \sup\left((1+|v_K|)^{\frac{mq}{2-q}}, (1+|v_L|)^{\frac{mq}{2-q}}\right) \le 2^{\frac{mq}{2-q}} (1 + |v_K|^{\frac{mq}{2-q}} + |v_L|^{\frac{mq}{2-q}}).$$

We deduce from (45), using the fact that $\sum_{\sigma\in\mathcal{E}} \mathrm{meas}(\sigma) d_\sigma = d\mathrm{meas}(\Omega)$ and re-ordering the sum on the control volumes,

$$||v_{\mathcal{T}}||_{1,q,\mathcal{M}}^q \le C_1 \left(1 + \sum_{K\in\mathcal{T}} |v_K|^{\frac{mq}{2-q}} \left(\sum_{\sigma\in\mathcal{E}_K} \mathrm{meas}(\sigma) d_\sigma\right)\right)^{\frac{2-q}{2}},$$

where C_1 only depends on (F, m, q, Ω). But since $d_{K,\sigma} \ge \zeta d_\sigma$ for all $K \in \mathcal{T}$ and all $\sigma \in \mathcal{E}_K$, we have $\sum_{\sigma\in\mathcal{E}_K} \mathrm{meas}(\sigma) d_\sigma \le \frac{1}{\zeta} \sum_{\sigma\in\mathcal{E}_K} \mathrm{meas}(\sigma) d_{K,\sigma} = \frac{d}{\zeta}\mathrm{meas}(K)$ and we obtain thus

$$||v_{\mathcal{T}}||_{1,q,\mathcal{M}}^q \le C_2 \left(1 + ||v_{\mathcal{T}}||_{L^{\frac{mq}{2-q}}(\Omega)}^{\frac{mq}{2}}\right), \qquad (46)$$

where C_2 only depends on (F, m, q, Ω) (notice that, since $m > 1$, we always have $\frac{mq}{2-q} \ge 1$).
By Proposition 2.2, there exists C_3 only depending on (Ω, q, ζ) such that, if $q^* = \frac{dq}{d-q}$ (note that $q < \frac{d}{d-1} \le d$),

$$||v_{\mathcal{T}}||_{L^{q^*}(\Omega)} \le C_3 ||v_{\mathcal{T}}||_{1,q,\mathcal{M}}.$$

Using this in (46), we obtain

$$||v_{\mathcal{T}}||^q_{L^{q^*}(\Omega)} \leq C_3^q C_2 \left(1 + ||v_{\mathcal{T}}||^{\frac{mq}{2}}_{L^{\frac{mq}{2-q}}(\Omega)}\right).$$

If $q < \frac{d}{d-1}$, one has $\frac{q}{2-q} < q^*$, so that we can choose $m \in (1,2)$ (only depending on (q,d)) such that $\frac{mq}{2-q} \leq q^*$. We obtain thus, with such a choice of m and Hölder's inequality,

$$||v_{\mathcal{T}}||^q_{L^{q^*}(\Omega)} \leq C_4 \left(1 + ||v_{\mathcal{T}}||^{\frac{mq}{2}}_{L^{q^*}(\Omega)}\right),$$

where C_4 only depends on (Ω, ζ, q, F). Since $\frac{mq}{2} < q$ (recall that $m < 2$), this inequality gives us C_5 only depending on (Ω, ζ, q, F) such that $||v_{\mathcal{T}}||_{L^{q^*}(\Omega)} \leq C_5$ and, returning to (46), we deduce the desired estimate on $||v_{\mathcal{T}}||_{1,q,\mathcal{M}}$. ■

3.3 Proof of Theorem 2.2

We give here the proof of the key estimate on $||u_{\mathcal{T}}||_{1,q,\mathcal{M}}$ which was stated in Theorem 2.2, and which is crucial to show existence and convergence of the solution to the finite volume scheme.

Proof of Theorem 2.2

Let $\Lambda > ||\mu||_{M(\overline{\Omega})}$ (to avoid dividing by 0). (8)—(10) being a linear problem, we see that $u_{\mathcal{T}}/\Lambda$ is a solution to (8)—(10) with μ/Λ instead of μ.

Since $||\mu/\Lambda||_{M(\overline{\Omega})} \leq 1$, we can apply Proposition 3.2 to $u_{\mathcal{T}}/\Lambda$; let $k_0 > 0$ only depending on $(\Omega, \mathbf{v}, \zeta)$ given by this proposition. $S_{k_0}(u_{\mathcal{T}}/\Lambda)$ and $T_{k_0}(u_{\mathcal{T}}/\Lambda)$ satisfy then the hypotheses of Lemma 3.4 with a function F only depending on $(\Omega, \mathbf{v}, \zeta)$. We deduce from this lemma that, for all $q \in [1, \frac{d}{d-1})$, there exists $C > 0$ only depending on $(\Omega, \mathbf{v}, \zeta, q)$ such that

$$||S_{k_0}(u_{\mathcal{T}}/\Lambda)||_{1,q,\mathcal{M}} \leq C \quad \text{and} \quad ||T_{k_0}(u_{\mathcal{T}}/\Lambda)||_{1,q,\mathcal{M}} \leq C.$$

Since $u_{\mathcal{T}}/\Lambda = S_{k_0}(u_{\mathcal{T}}/\Lambda) + T_{k_0}(u_{\mathcal{T}}/\Lambda)$ and $||\cdot||_{1,q,\mathcal{M}}$ is a norm, this gives $||u_{\mathcal{T}}/\Lambda||_{1,q,\mathcal{M}} \leq C$, that is to say $||u_{\mathcal{T}}||_{1,q,\mathcal{M}} \leq C\Lambda$. Letting then Λ tend to $||\mu||_{M(\overline{\Omega})}$, we obtain the desired estimate on $u_{\mathcal{T}}$. ■

4 Proof of Theorem 2.1

We first prove the uniqueness of the solution to (14), which does not involve numerical analysis methods, and then the existence and convergence of the approximate solutions (which yields the existence of a solution to (14)).

Proof of the uniqueness of the solution to (14)

This proof uses the regularity results of [22] on the variational solution to $-\Delta v = f \in L^2(\Omega)$, $v_{|\partial\Omega} = 0$, for Ω polygonal (or polyhedral) open set in $\mathbb{R}^d$, $d = 2$ or 3.

Problem (14) being linear, it is sufficient to prove that, if u is a solution to (14) with $\mu = 0$, then $u = 0$. Let $\theta \in L^\infty(\Omega)$ and take $\varphi \in H^1_0(\Omega) \cap L^\infty(\Omega)$ the solution to

$$\int_\Omega \nabla\varphi \cdot \nabla\psi \, d\lambda - \int_\Omega \psi \mathbf{v} \cdot \nabla\varphi \, d\lambda + \int_\Omega b\varphi\psi \, d\lambda = \int_\Omega \theta\psi \, d\lambda \,, \ \forall \psi \in H^1_0(\Omega). \tag{47}$$

The existence of such a φ is ensured by the results of [9]. Letting $\Theta = \theta + \mathbf{v} \cdot \nabla\varphi - b\varphi \in L^2(\Omega)$, we see that $\varphi \in H^1_0(\Omega)$ satisfies $-\Delta\varphi = \Theta$ on Ω.

Since Ω is a polygonal (or polyhedral) open set in $\mathbb{R}^2$ or $\mathbb{R}^3$, the results of [22] give us $\eta > 0$ such that $\varphi \in H^{\frac{3}{2}+\eta}(\Omega)$. Thus, by the Sobolev injections (see [1]), there exists $s > d$ such that $\varphi \in W^{1,s}_0(\Omega)$ (in the case $d = 2$, to obtain such a $s > 2$, we could also have used the result of [28] — which is stated for regular open sets but is also true for open sets with Lipschitz-continuous boundary, see [21]).

Thanks to this additional regularity, a density argument allows to see that (47) is also true for $\psi \in W_0^{1,s'}(\Omega)$, where s' is the conjugate exponent to s, that is, such that $\frac{1}{s} + \frac{1}{s'} = 1$.
We can thus use φ in the equation satisfied by u and u in the equation satisfied by φ to obtain

$$0 = \int_\Omega \nabla u \cdot \nabla \varphi \, d\lambda - \int_\Omega u \mathbf{v} \cdot \nabla \varphi \, d\lambda + \int_\Omega bu\varphi \, d\lambda = \int_\Omega \theta u \, d\lambda.$$

We deduce from this equality, satisfied for all $\theta \in L^\infty(\Omega)$, that $u = 0$, i.e. the uniqueness of the solution to (14). ■

Proof of the existence and convergence results
The existence of a unique solution to (8)—(10) is an immediate consequence of the estimate of Theorem 2.2: indeed, if $\mu = 0$, then this theorem shows that any solution to (8)—(10) is null, that is to say that the square matrix defining this linear system is invertible.

Let us now prove the convergence result. The techniques used here are easy adaptations of the convergence proof of [14].
Let $(u_n)_{n\in\mathbb{N}}$ be a sequence of functions of $L^2(\Omega)$ such that u_n is solution to (8)—(10) for $\mathcal{M} = \mathcal{M}_n$, where $(\mathcal{M}_n)_{n\in\mathbb{N}}$ is a family of admissible meshes $\mathcal{M}$ satisfying (15) (for some fixed $\zeta > 0$), and such that size$(\mathcal{M}_n)$ tends to 0 as n tends to $+\infty$.
We first prove (steps 0 to 5), that if $(u_n)_{n\in\mathbb{N}}$ tends to u in $L^p(\Omega)$ for all $p < \frac{d}{d-2}$, as n tends to $+\infty$ (and size$(\mathcal{M}_n) \to 0$), with $u \in \cap_{q<\frac{d}{d-1}} W_0^{1,q}(\Omega)$, then u is a solution to (14).
We then prove (step 6), thanks to the a priori estimates of Section 3, the compactness of the sequence $(u_n)_{n\in\mathbb{N}}$ and conclude, thanks to the uniqueness result which was proved above, to the convergence of $(u_n)_{n\in\mathbb{N}}$ to the solution u to (14).

Step 0: Density argument
By density of $C_c^\infty(\Omega)$ in $W_0^{1,s}(\Omega)$ for all $s \in (d, \infty)$ and by the regularity results on u, it is clearly sufficient to prove that u satisfies the equation of (14) for all $\varphi \in C_c^\infty(\Omega)$. Take such a φ. Multiplying (8) by $\varphi(x_K)$ and summing over $K \in \mathcal{T}$ we have, by conservativity of the fluxes and by dropping the index n:

$$\sum_{\sigma\in\mathcal{E}} \tau_\sigma (u_K - u_L)(\varphi(x_K) - \varphi(x_L)) + \sum_{\sigma\in\mathcal{E}} v_{K,\sigma} u_{\sigma,+} (\varphi(x_K) - \varphi(x_L)) + \sum_{K\in\mathcal{T}} \mathrm{meas}(K) b_K u_K \varphi(x_K) = \sum_{K\in\mathcal{T}} \varphi(x_K)\mu(K). \tag{48}$$

We shall now pass to the limit as size$(\mathcal{M})$ tends to 0 in (48), and prove the convergence of each of the terms to the corresponding term in (14). In fact, the proof of convergence of the first and third terms of the left hand side can be found in [14] or [15] and so can the proof of the second term under a stronger regularity condition. The proof of convergence of the right hand side may be found in [19], so that the only new part in this proof is Step 4 which shows the convergence of the convective term with a continuous convection velocity (rather than C^1 in previous works). However, we give a quick proof for all terms for the sake of completeness.

Step 1: convergence of the lower order terms.
Denote $\varphi_\mathcal{T} \in X(\mathcal{T})$ the function defined by $\varphi_K = \varphi(x_K)$ for all $K \in \mathcal{T}$. By regularity of φ, we have $\varphi_\mathcal{T} \to \varphi$ uniformly on Ω as size$(\mathcal{M}) \to 0$, thus

$$\sum_{K\in\mathcal{T}} \varphi(x_K)\mu(K) = \int_\Omega \varphi_\mathcal{T} \, d\mu \to \int_\Omega \varphi \, d\mu \tag{49}$$

as size$(\mathcal{M}) \to 0$ (notice that $\varphi_\mathcal{T} = 0$ near $\partial\Omega$ for size$(\mathcal{M})$ small enough).

By regularity of b, $b_{\mathcal{T}} = (b_K)_{K\in\mathcal{T}}$ tends to b in $L^2(\Omega)$ as size$(\mathcal{M}) \to 0$; thus, since $\varphi_{\mathcal{T}} \to \varphi$ in $L^\infty(\Omega)$ and $u_{\mathcal{T}} \to u$ in $L^2(\Omega)$ (because $2 < d/(d-2)$) as size$(\mathcal{M}) \to 0$, we have

$$\sum_{K\in\mathcal{T}} \text{meas}(K) b_K u_K \varphi(x_K) = \int_\Omega b_{\mathcal{T}} u_{\mathcal{T}} \varphi_{\mathcal{T}}\, d\lambda \to \int_\Omega bu\varphi\, d\lambda \tag{50}$$

as size$(\mathcal{M}) \to 0$.

Step 2: convergence of the diffusion term.
Gathering by control volumes, we have

$$\sum_{\sigma\in\mathcal{E}} \tau_\sigma (u_K - u_L)(\varphi(x_K) - \varphi(x_L)) = \sum_{K\in\mathcal{T}} u_K \sum_{\sigma\in\mathcal{E}_K} \tau_\sigma(\varphi(x_K) - \varphi(x_L)).$$

But, by regularity of φ,

$$\tau_\sigma(\varphi(x_K) - \varphi(x_L)) = -\int_\sigma \nabla\varphi\cdot\mathbf{n}_{K,\sigma}\, d\gamma + \text{meas}(\sigma) R_{K,\sigma},$$

where $|R_{K,\sigma}| \le C_1 \text{size}(\mathcal{M})$ (C_1 does not depend on the mesh) and $R_{K,\sigma} = -R_{L,\sigma}$ whenever $\sigma = K|L \in \mathcal{E}_{\text{int}}$. Thus, gathering by edges,

$$\begin{aligned}\sum_{\sigma\in\mathcal{E}} \tau_\sigma (u_K - u_L)(\varphi(x_K) - \varphi(x_L)) + \sum_{K\in\mathcal{T}} u_K \int_{\partial K} \nabla\varphi\cdot\mathbf{n}_K\, d\gamma &= \sum_{K\in\mathcal{T}} u_K \sum_{\sigma\in\mathcal{E}_K} \text{meas}(\sigma) R_{K,\sigma} \\ &= \sum_{\sigma\in\mathcal{E}} \text{meas}(\sigma) R_{K,\sigma}(u_K - u_L).\end{aligned}$$

But

$$\left|\sum_{\sigma\in\mathcal{E}} \text{meas}(\sigma) R_{K,\sigma}(u_K - u_L)\right| \le C_1 \text{size}(\mathcal{M}) \sum_{\sigma\in\mathcal{E}} \text{meas}(\sigma) d_\sigma \frac{D_\sigma u_{\mathcal{T}}}{d_\sigma} = C_1 \text{size}(\mathcal{M}) ||u_{\mathcal{T}}||_{1,1,\mathcal{M}},$$

and this last quantity tends to 0 as size$(\mathcal{M}) \to 0$ (because, by Theorem 2.2, $||u_{\mathcal{T}}||_{1,1,\mathcal{M}}$ stays bounded). By noticing that

$$\sum_{K\in\mathcal{T}} u_K \int_{\partial K} \nabla\varphi\cdot\mathbf{n}_K\, d\gamma = \sum_{K\in\mathcal{T}} u_K \int_K \Delta\varphi\, d\lambda = \int_\Omega u_{\mathcal{T}} \Delta\varphi\, d\lambda,$$

and since $u_{\mathcal{T}} \to u$ in $L^1(\Omega)$ as size$(\mathcal{M}) \to 0$, we deduce that

$$\sum_{\sigma\in\mathcal{E}} \tau_\sigma (u_K - u_L)(\varphi(x_K) - \varphi(x_L)) \to -\int_\Omega u\Delta\varphi\, d\lambda = \int_\Omega \nabla u\cdot\nabla\varphi\, d\lambda \tag{51}$$

as size$(\mathcal{M}) \to 0$.

Step 3: Preliminary to the convergence of the convection term (in fact, we prove here the convergence of the convection term if $\mathbf{v}$ is regular).
Let $\mathbf{w} \in (C^1(\overline{\Omega}))^d$ and define, for $K \in \mathcal{T}$ and $\sigma \in \mathcal{E}_K$, $w_{K,\sigma} = \int_\sigma \mathbf{w}\cdot\mathbf{n}_{K,\sigma}\, d\gamma$ (notice that, if $\sigma = K|L \in \mathcal{E}_{\text{int}}$, then $w_{K,\sigma} = -w_{L,\sigma}$). We want to study the limit, as size$(\mathcal{M}) \to 0$, of $\sum_{\sigma\in\mathcal{E}} w_{K,\sigma} u_{\sigma,+}(\varphi(x_K) - \varphi(x_L))$ (that is to say the convection term of (48) with $\mathbf{w}$ instead of $\mathbf{v}$).
We have

$$\begin{aligned}&\sum_{\sigma\in\mathcal{E}} w_{K,\sigma} u_{\sigma,+}(\varphi(x_K) - \varphi(x_L)) \\ &\quad= \sum_{K\in\mathcal{T}} \sum_{\sigma\in\mathcal{E}_K} w_{K,\sigma} u_{\sigma,+} \varphi(x_K) \\ &\quad= \sum_{K\in\mathcal{T}} \sum_{\sigma\in\mathcal{E}_K} w_{K,\sigma}(u_{\sigma,+} - u_K)\varphi(x_K) + \sum_{K\in\mathcal{T}} \varphi(x_K) u_K \sum_{\sigma\in\mathcal{E}_K} w_{K,\sigma}.\end{aligned} \tag{52}$$

Since $\sum_{\sigma\in\mathcal{E}_K} w_{K,\sigma} = \int_{\partial K} \mathbf{w}\cdot\mathbf{n}_K\,d\gamma = \int_K \mathrm{div}(\mathbf{w})\,d\lambda$, we have

$$\sum_{K\in\mathcal{T}} \varphi(x_K)u_K \sum_{\sigma\in\mathcal{E}_K} w_{K,\sigma} = \int_\Omega u_\mathcal{T}\varphi_\mathcal{T}\mathrm{div}(\mathbf{w})\,d\lambda \to \int_\Omega u\varphi\mathrm{div}(\mathbf{w})\,d\lambda \tag{53}$$

as size$(\mathcal{M})\to 0$.
Moreover,

$$\begin{aligned}\sum_{K\in\mathcal{T}}\sum_{\sigma\in\mathcal{E}_K} w_{K,\sigma}(u_{\sigma,+}-u_K)\varphi(x_K) = &\sum_{K\in\mathcal{T}}\sum_{\sigma\in\mathcal{E}_K}(u_{\sigma,+}-u_K)\int_\sigma \varphi\mathbf{w}\cdot\mathbf{n}_{K,\sigma}\,d\gamma\\ &+\sum_{K\in\mathcal{T}}\sum_{\sigma\in\mathcal{E}_K}(u_{\sigma,+}-u_K)\int_\sigma(\varphi(x_K)-\varphi)\mathbf{w}\cdot\mathbf{n}_{K,\sigma}\,d\gamma.\end{aligned}$$

Since, for size$(\mathcal{M})$ small enough, the support of φ does not intersect the cells K such that $\partial K\cap\partial\Omega\neq\emptyset$, we have

$$\sum_{K\in\mathcal{T}}\sum_{\sigma=K|L\in\mathcal{E}_K} u_{\sigma,+}\int_\sigma \varphi\mathbf{w}\cdot\mathbf{n}_{K,\sigma}\,d\gamma = \sum_{\sigma\in\mathcal{E}_{\rm int}} u_{\sigma,+}\left(\int_\sigma \varphi\mathbf{w}\cdot\mathbf{n}_{K,\sigma}\,d\gamma+\int_\sigma\varphi\mathbf{w}\cdot\mathbf{n}_{L,\sigma}\,d\gamma\right)=0,$$

because $\mathbf{n}_{K,\sigma}=-\mathbf{n}_{L,\sigma}$ if $\sigma=K|L\in\mathcal{E}_{\rm int}$. On the other hand,

$$-\sum_{K\in\mathcal{T}}\sum_{\sigma\in\mathcal{E}_K} u_K\int_\sigma\varphi\mathbf{w}\cdot\mathbf{n}_{K,\sigma}\,d\gamma = -\sum_{K\in\mathcal{T}} u_K\int_K\mathrm{div}(\varphi\mathbf{w})\,d\lambda = -\int_\Omega u_\mathcal{T}\mathrm{div}(\varphi\mathbf{w})\,d\lambda\to-\int_\Omega u\mathrm{div}(\varphi\mathbf{w})\,d\lambda$$

as size$(\mathcal{M})\to 0$. By regularity of φ, we have C_5 only depending on φ such that

$$\begin{aligned}&\left|\sum_{K\in\mathcal{T}}\sum_{\sigma\in\mathcal{E}_K}(u_{\sigma,+}-u_K)\int_\sigma(\varphi(x_K)-\varphi)\mathbf{w}\cdot\mathbf{n}_{K,\sigma}\,d\gamma\right|\\ &\quad\le\ C_5||\,|\mathbf{w}|\,||_{C(\overline{\Omega})}\mathrm{size}(\mathcal{M})\sum_{K\in\mathcal{T}}\sum_{\sigma\in\mathcal{E}_K}\mathrm{meas}(\sigma)|u_{\sigma,+}-u_K|\\ &\quad\le\ C_5||\,|\mathbf{w}|\,||_{C(\overline{\Omega})}\mathrm{size}(\mathcal{M})\sum_{\sigma\in\mathcal{E}}\mathrm{meas}(\sigma)D_\sigma u_\mathcal{T}\\ &\quad=\ C_5||\,|\mathbf{w}|\,||_{C(\overline{\Omega})}\mathrm{size}(\mathcal{M})||u_\mathcal{T}||_{1,1,\mathcal{M}}.\end{aligned}$$

The last quantity tending to 0 as size$(\mathcal{M})\to 0$, we deduce from what preceeds that

$$\sum_{K\in\mathcal{T}}\sum_{\sigma\in\mathcal{E}_K} w_{K,\sigma}(u_{\sigma,+}-u_K)\varphi(x_K)\to-\int_\Omega u\mathrm{div}(\varphi\mathbf{w})\,d\lambda \tag{54}$$

as size$(\mathcal{M})\to 0$.
Using (53) and (54) in (52), we obtain

$$\sum_{\sigma\in\mathcal{E}} w_{K,\sigma}u_{\sigma,+}(\varphi(x_K)-\varphi(x_L))\to\int_\Omega u\varphi\mathrm{div}(\mathbf{w})\,d\lambda-\int_\Omega u\mathrm{div}(\varphi\mathbf{w})\,d\lambda=-\int_\Omega u\mathbf{w}\cdot\nabla\varphi\,d\lambda \tag{55}$$

as size$(\mathcal{M})\to 0$.

Step 4: convergence of the convection term.
Let $\varepsilon>0$ and take $\mathbf{w}\in(C^1(\overline{\Omega}))^d$ such that $||\,|\mathbf{v}-\mathbf{w}|\,||_{C(\overline{\Omega})}\le\varepsilon$. By regularity of φ,

$$\left|\sum_{\sigma\in\mathcal{E}} v_{K,\sigma}u_{\sigma,+}(\varphi(x_K)-\varphi(x_L))-\sum_{\sigma\in\mathcal{E}} w_{K,\sigma}u_{\sigma,+}(\varphi(x_K)-\varphi(x_L))\right|\le C_2\varepsilon\sum_{\sigma\in\mathcal{E}}\mathrm{meas}(\sigma)d_\sigma|u_{\sigma,+}|$$

where C_2 only depends on φ. Gathering by control volumes, we deduce that

$$\left|\sum_{\sigma\in\mathcal{E}} v_{K,\sigma}u_{\sigma,+}(\varphi(x_K)-\varphi(x_L)) - \sum_{\sigma\in\mathcal{E}} w_{K,\sigma}u_{\sigma,+}(\varphi(x_K)-\varphi(x_L))\right| \leq C_2\varepsilon \sum_{K\in\mathcal{T}} |u_K| \sum_{\sigma\in\mathcal{E}_K \,|\, v_{K,\sigma}\geq 0} \mathrm{meas}(\sigma)d_\sigma.$$

But, by hypothesis on the mesh, $\sum_{\sigma\in\mathcal{E}_K \,|\, v_{K,\sigma}\geq 0} \mathrm{meas}(\sigma)d_\sigma \leq \zeta^{-1}\sum_{\sigma\in\mathcal{E}_K} \mathrm{meas}(\sigma)d_{K,\sigma} = \zeta^{-1}d\mathrm{meas}(K)$, so that

$$\begin{aligned}\left|\sum_{\sigma\in\mathcal{E}} v_{K,\sigma}u_{\sigma,+}(\varphi(x_K)-\varphi(x_L)) - \sum_{\sigma\in\mathcal{E}} w_{K,\sigma}u_{\sigma,+}(\varphi(x_K)-\varphi(x_L))\right| &\leq C_3\varepsilon \sum_{K\in\mathcal{T}} \mathrm{meas}(K)|u_K| \\ &\leq C_4\varepsilon, \end{aligned} \tag{56}$$

where C_3 and C_4 do not depend on the mesh $\mathcal{M}$ nor on ε ($\sum_{K\in\mathcal{T}} \mathrm{meas}(K)|u_K| = ||u_\mathcal{T}||_{L^1(\Omega)}$ is bounded). We also notice that

$$\left|\int_\Omega u\mathbf{v}\cdot\nabla\varphi\, d\lambda - \int_\Omega u\mathbf{w}\cdot\nabla\varphi\, d\lambda\right| \leq C_6\varepsilon \tag{57}$$

where C_6 does not depend on ε.
Using then (55) and (57) in (56), we obtain

$$\limsup_{\mathrm{size}(\mathcal{M})\to 0} \left|\sum_{\sigma\in\mathcal{E}} v_{K,\sigma}u_{\sigma,+}(\varphi(x_K)-\varphi(x_L)) + \int_\Omega u\mathbf{v}\cdot\nabla\varphi\, d\lambda\right| \leq C_7\varepsilon,$$

where C_7 does not depend on ε. This being true for any $\varepsilon > 0$, we deduce that

$$\sum_{\sigma\in\mathcal{E}} v_{K,\sigma}u_{\sigma,+}(\varphi(x_K)-\varphi(x_L)) \to -\int_\Omega u\mathbf{v}\cdot\nabla\varphi\, d\lambda \tag{58}$$

as $\mathrm{size}(\mathcal{M}) \to 0$.

Step 5: Passage to the limit in the scheme.
Using (49), (50), (51) and (58), we may pass to the limit in (48) to obtain:

$$\int_\Omega \nabla u\cdot\nabla\varphi\, d\lambda - \int_\Omega u\mathbf{v}\cdot\nabla\varphi\, d\lambda + \int_\Omega bu\varphi\, d\lambda = \int_\Omega \varphi\, d\mu,$$

which proves that u is a solution to (14).

Step 6: proof of the convergence of $(u_n)_{n\in\mathbb{N}}$.
Thanks to Theorem 2.2 and to Propositions 2.3 and 2.4, we see that $(u_n)_{n\geq 1}$ is relatively compact in $L^q(\Omega)$ for all $q \in [1, \frac{d}{d-1})$ and that the adherence values (in $L^q(\Omega)$) of this sequence are in $W_0^{1,q}(\Omega)$ (for $q \in (1, \frac{d}{d-1})$). Up to a subsequence, we can thus suppose that $u_n \to u$ in $L^q(\Omega)$ for all $q \in [1, \frac{d}{d-1})$, with $u \in \cap_{q<\frac{d}{d-1}} W_0^{1,q}(\Omega)$; by Proposition 2.2 and Theorem 2.2, $(u_n)_{n\geq 1}$ is also bounded in $L^p(\Omega)$ for all $p < \frac{d}{d-2}$ so that, by an easy consequence of the Vitali convergence theorem, $u_n \to u$ in $L^p(\Omega)$ for all $p < \frac{d}{d-2}$.
By what we have just proved, we see that u is then a solution to (14); since this solution is unique, this proves that the whole sequence $(u_n)_{n\geq 1}$ converges to u.
As a by-product, this convergence entails the existence of a solution to (14) (which can be deduced from previous works, [2] and [9] for instance). ■

5 A scheme with jump of the fluxes

Until now, we considered, in the definition of "admissible mesh", a partition of Ω into convex polygonal (or polyhedral) sets. We then defined a finite volume scheme where the conservativity of the numerical fluxes writes : $F_{K,\sigma} = -F_{L,\sigma}$ for all $\sigma = K|L \in \mathcal{E}_{\rm int}$.
There is, however, another manner to deal with the discretization of a right-hand side measure, which was implemented, for instance, in [17] for the numerical simulation of fuel cells. In this formulation, we write that if the support of the measure intersects a given edge, then there is a jump of the flux on this edge. This leads to the following scheme.

The mesh $\mathcal{M}$ we consider now is defined by a finite family $\mathcal{T}$ of polygonal (or polyhedral) open disjoint subsets of Ω, by a finite family $\mathcal{E}$ of subsets of $\overline{\Omega}$ contained in affine hyperplanes and by a finite family $\mathcal{P} = (x_K)_{K\in\mathcal{T}}$ of points of Ω such that

a) $\mathcal{T} \cup \mathcal{E}$ is a partition of $\overline{\Omega}$,

b) for each $\sigma \in \mathcal{E}$, there exists $K \in \mathcal{T}$ and a non-empty open subset O of ∂K such that $O \subset \sigma \subset \overline{O}$,

c) items iii)—vi) of Definition 2.1 hold.

The notations concerning the mesh are the same as before, and the reader can easily verify that Propositions 2.1 — 2.4 are still true for such meshes.
Still defining $(b_K)_{K\in\mathcal{T}}$ and $(v_{K,\sigma})_{K\in\mathcal{T},\,\sigma\in\mathcal{E}_K}$ by (7), the new scheme is

$$\forall K \in \mathcal{T}\,,\ \sum_{\sigma\in\mathcal{E}_K} F_{K,\sigma} + \sum_{\sigma\in\mathcal{E}_K} v_{K,\sigma}u_{\sigma,+} + \mathrm{meas}(K)b_K u_K = \mu(K)\,, \tag{59}$$

$$\begin{array}{ll} \forall \sigma = K|L \in \mathcal{E}_{\rm int}\,, & F_{K,\sigma} = -\frac{\mathrm{meas}(\sigma)}{d_{K,\sigma}}(u_\sigma - u_K)\,, \\ \forall \sigma \in \mathcal{E}_{\rm ext}\cap\mathcal{E}_K\,, & F_{K,\sigma} = \tau_\sigma u_K\,, \end{array} \tag{60}$$

$$\forall \sigma = K|L \in \mathcal{E}_{\rm int}\,,\quad F_{K,\sigma} + F_{L,\sigma} = -\mu(\sigma)\,, \tag{61}$$

$$\begin{array}{lll} \forall \sigma = K|L \in \mathcal{E}_{\rm int}\,, & u_{\sigma,+} = u_K \ \text{ if } v_{K,\sigma} \geq 0\,, & u_{\sigma,+} = u_L \ \text{ otherwise,} \\ \forall \sigma \in \mathcal{E}_{\rm ext}\cap\mathcal{E}_K\,, & u_{\sigma,+} = u_K \ \text{ if } v_{K,\sigma} \geq 0\,, & u_{\sigma,+} = 0 \ \text{ otherwise.} \end{array} \tag{62}$$

Notice that the unknowns of this scheme are $(u_K)_{K\in\mathcal{T}}$ and $(u_\sigma)_{\sigma\in\mathcal{E}}$ (which represent approximate values on the edges), but that Relation (61) allows to eliminate the $(u_\sigma)_{\sigma\in\mathcal{E}}$; this scheme can thus be considered as a linear system on $(u_K)_{K\in\mathcal{T}}$.

In fact, the elimination of u_σ thanks to (61) gives, for $\sigma = K|L \in \mathcal{E}_{\rm int}$,

$$F_{K,\sigma} = \frac{\mathrm{meas}(\sigma)}{d_\sigma}(u_K - u_L) - \frac{d_{L,\sigma}}{d_\sigma}\mu(\sigma).$$

Thus, this new scheme is in fact the scheme (8)—(10) where we have changed, for all $K \in \mathcal{T}$, $\mu(K)$ by $\widetilde{\mu}_K = \mu(K) + \sum_{\sigma\in\mathcal{E}_K} \frac{d_{L,\sigma}}{d_\sigma}\mu(\sigma)$ (with $\sigma = K|L$ if $\sigma \in \mathcal{E}_{\rm int}$ and $d_{L,\sigma} = 0$ if $\sigma \in \mathcal{E}_{\rm ext}\cap\mathcal{E}_K$), which is just another way to discretize the measure μ (forgetting the values of μ on the boundary of the domain, which does not modify the problem since we consider Dirichlet boundary conditions).
The matrix of (59)—(62) is thus the same as the matrix of (8)—(10) and, since $(\widetilde{\mu}_K)_{K\in\mathcal{T}}$ satisfies

$$\sum_{K\in\mathcal{T}} |\widetilde{\mu}_K| \leq \sum_{K\in\mathcal{T}} |\mu(K)| + \sum_{\sigma\in\mathcal{E}} \left(\frac{d_{K,\sigma}}{d_\sigma} + \frac{d_{L,\sigma}}{d_\sigma}\right)|\mu(\sigma)| = \sum_{K\in\mathcal{T}} |\mu(K)| + \sum_{\sigma\in\mathcal{E}} |\mu(\sigma)| \leq ||\mu||_{M(\overline{\Omega})}$$

(because $\mathcal{T} \cup \mathcal{E}$ is a partition of $\overline{\Omega}$), the *a priori* estimates on the solutions to (59)—(62) are obtained exactly the same way as the estimates on the solutions to (8)—(10).

We also have, for $\varphi \in C_c(\Omega)$, for $\sigma = K|L \in \mathcal{E}_{\rm int}$,

$$\left| \frac{d_{L,\sigma}}{d_\sigma} \varphi(x_K)\mu(\sigma) + \frac{d_{K,\sigma}}{d_\sigma} \varphi(x_L)\mu(\sigma) - \int_\sigma \varphi \, d\mu \right| \leq \omega(\varphi, \text{size}(\mathcal{M}))|\mu(\sigma)|,$$

where $\omega(\varphi, h)$ is the modulus of continuity of φ; thus,

$$\sum_{K\in\mathcal{T}} \varphi(x_K)\widetilde{\mu}_K \to \int_\Omega \varphi \, d\mu$$

as size$(\mathcal{M}) \to 0$ and the convergence of the solution of (59)—(62) as size$(\mathcal{M}) \to 0$ is obtained by the same technique as in the proof of Theorem 2.1.

6 Numerical results

We performed a few simple numerical experiments on problems to which the exact solution is known, in order to try and obtain some rates of convergence of the finite volume scheme in presence of a non regular right hand side. Numerical results were also shown in [11] in the non coercive case with right hand side in H^{-1}, so we shall concentrate here on tests in the irregular data case.

6.1 Comparison of the two finite volume schemes

The first numerical experiment is concerned with the comparison of the treatment of the singularity in the one-dimensional case. In this case, the Dirac is not a very "mean" measure, in the sense that the solution of the problem is continous, the jump is only on the derivative. In the first version of the FV scheme (scheme (8)-(10), which we shall call Scheme 1 in the sequel), the Dirac measure is taken in its integral form in the right hand side while in the second version (scheme (59)-(62), which we shall call Scheme 2), the mesh is adapted so as to be able to write the numerical jump of the flux on a cell interface. We solve $-u'' = \delta_{1/2}$, $u(0) = 0$, $u(1) = 0$, on the interval $(0, 1)$; the exact solution is $u(x) = \frac{x}{2}$ for $x < .5$, $u(x) = \frac{(1-x)}{2}$ for $x \geq .5$. We use a uniform mesh, and ensure that the number of cells is even, so that in the second scheme, the flux jump is located on a cell interface. The error function e is defined by $e(x) = u(x_K) - u_K$ for any $x \in K$, where $u(x_K)$ denotes the value of the exact solution of the continuous problem at point x_K and $(u_K)_{K\in\mathcal{T}}$ the solution to the finite volume scheme.
We analyse the rate of convergence by showing the L^1, L^2 and L^∞ norms of the error e versus the number of cells with a log-log scale in Figure 3.

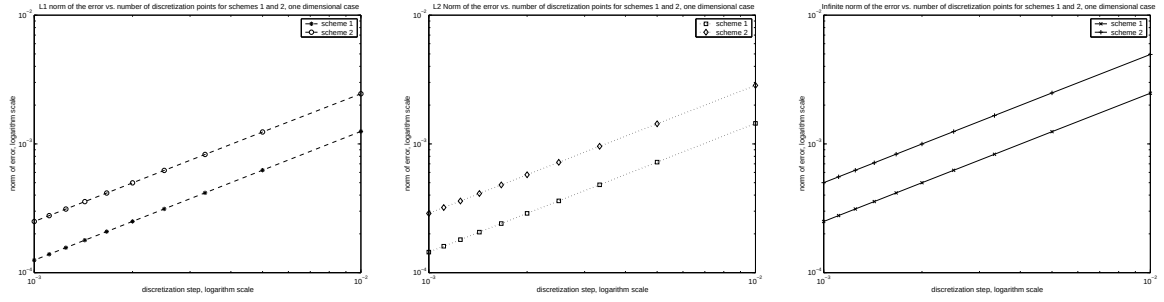


Figure 3: Convergence rate in the one-dimensional case.

The results show straight lines for all three norms, so that it is natural to try and evaluate the norms of the error as $\|e\| \equiv Ch^\alpha$. The computation of the coefficients C and α from the numerical results are given in Table 1. These coefficients are computed using the two finest meshes.

α	L^1 norm	L^2 norm	L^∞ norm
Scheme 1	1.0000	1.0000	0.9961
Scheme 2	0.9923	0.9941	0.9961

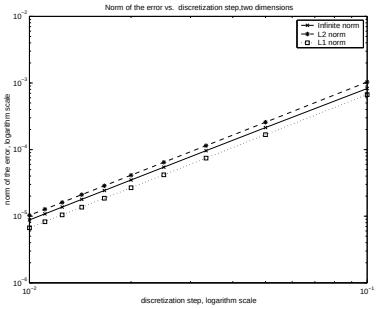
C	L^1 norm	L^2 norm	L^∞ norm
Scheme 1	0.1250	0.1443	0.2431
Scheme 2	0.2365	0.2768	0.4861

Table 1: Values of (C, α) for schemes 1 and 2, one dimensional case.

These results show that the two schemes have a rate of convergence which is roughly the same (close to one) and that the constant C is about twice as large for Scheme 2 (jump of flux) than for Scheme 1 (Dirac in one cell). This is quite in accordance with what can be seen from the implementation the scheme, because Scheme 2 amounts to spreading the Dirac measure over two cells, instead of one in Scheme 1.

6.2 Two and three-dimensional tests on a Cartesian mesh

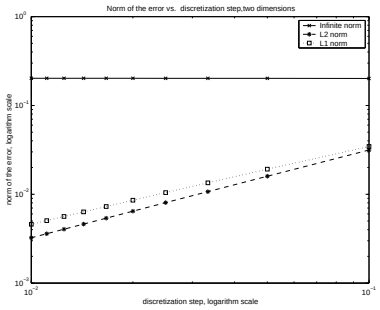
We also implemented the finite volume scheme on the square (resp. cubic) domain $\Omega = (-1, 1)^2$ (resp. $\Omega = (-1, 1)^3$). The domain is discretized with a uniform mesh, and the L^p norm of the error is computed for an increasing number of cells, so as to evaluate the rate of convergence.
We first tested the two-dimensional code for a regular data, yielding the exact solution $u(x, y) = \sin x \sin y$. In this case, since the mesh is rectangular and the exact solution regular, the consistency error on the flux is of order 2 and the rate of convergence in the L^2 norm can be theoretically shown to be of order 2 ([14], [20], see also [5] for a related co-volume scheme). The rate of convergence was computed for the piecewise constant error function defined by $e_K = u(x_K) - u_K$ for $K \in \mathcal{T}$, where u is the exact solution and $(u_K)_{K\in\mathcal{T}}$ is the solution to the finite volume scheme.



	α	C
L^1 norm	2.0000	.1031
L^2 norm	2.0000	.0428
L^∞ norm	1.7931	.0690

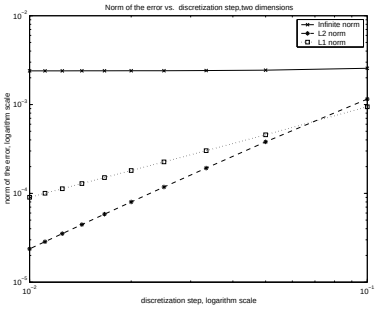
Figure 4: Convergence rate, two dimensional case, regular right hand side.

We then performed some tests with a right hand side given by a Dirac measure at 0. The boundary conditions were taken such that the exact solution be the restriction of the solution of $-\Delta u = \delta_0$ in the whole set $\mathbb{R}^2$ (resp. $\mathbb{R}^3$). It is well-known that this function lies in $L^p(\mathbb{R}^2)$ for $p \in [1, +\infty)$ (resp. $L^p(\mathbb{R}^3)$ for $p \in [1, 3)$).
We obtain the results (in log-log scale) given in Figure 5. The coefficients C and α such that $\|e\| = Ch^\alpha$ are again evaluated for the norms $L^1(\Omega)$ and $L^2(\Omega)$, and are also given in Figure 5.
In these tests, the mesh is such that the point $(0, 0)$ is located at the corner of the cell $[0, h] \times [0, h]$, where h is the discretization step of the mesh. Hence the radial symmetry of the solution is broken by the mesh. If we restore it by allocating one fourth of the Dirac measure to each of the four cells $[0, h] \times [0, h]$, $[0, h] \times [0, -h]$, $[-h, 0] \times [0, h]$ and $[-h, 0] \times [-h, 0]$, we gain in the order of convergence, as can be seen in Figure 6. Hence the order of convergence depends on the singularity of the data, but also on the preservation of the symmetry of the solution.



	α	C
L^1 norm	.9047	.2421
L^2 norm	.9965	.3181

Figure 5: Convergence rate, two dimensional case, right hand side Dirac at zero, non symmetric discrete problem.



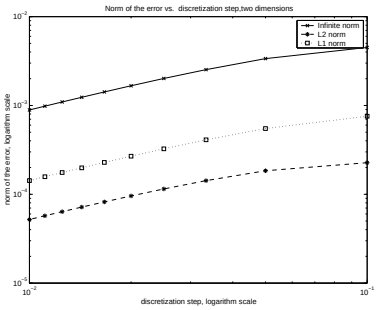
	α	C
L^1 norm	1.7740	.0073
L^2 norm	1.0010	.0837

Figure 6: Convergence rate, two dimensional case, right hand side Dirac at zero, symmetric discrete problem.

A question of interest is to know whether the singular data influences the rate of convergence outside of the region of singularity. In order to check this point, we compute the norm of the error between the exact and approximate solutions on the region $\{x \leq -.5\} \times \{y \leq -.5\}$. We find that in this case, we recover an order of convergence close to one in all norms if the Dirac is located at the corner of a cell, in which case the symmetry of the solution is not preserved by the discretization (see Figure 7). In this case, the rate of convergence in the regular zone is perturbed by the singularity outside this zone (recall that the theoretical rate of convergence for regular solutions on rectangular meshes is 2 [20], [14]). However, if we restore the symmetry of the problem as described above, then the rate of convergence is close to two (see Figure 8).

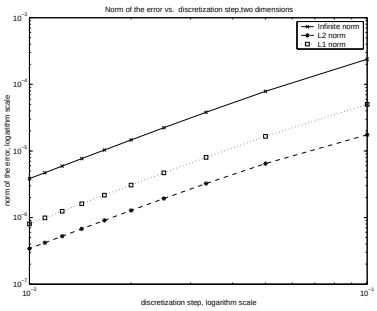
We then implemented a three dimensional cartesian mesh and found, for the non-symmetric discrete problem (Dirac located at a corner of the cell $[0, h]^3$) a rate of convergence close to 1 in norm L^1 and .5 in norm L^2, as shown in Figure 9. Recall that in this case the exact solution is in L^p for $1 \leq p < 3$.

If the Dirac is distributed on the eight cells neighbouring the origin, in order to symmtrize the discrete problem, as was done in the two-dimensional case, then one obtains a rate of convergence of 1.631 in the L^1 norm and .504 in the L^2 norm. This seems to indicate a super-convergence in the L^1 norm, although not to the second order (see also Remark 6.1).



	α	C
L^1 norm	0.9131	.0035
L^2 norm	0.9350	.0086
L^∞ norm	0.9360	.0586

Figure 7: Convergence rate, two dimensional case, right hand side Dirac at zero, non symmetric discrete problem, norm computed on a "regular zone".



	α	C
L^1 norm	1.9486	.0247
L^2 norm	1.9571	.0042
L^∞ norm	1.9305	.0025

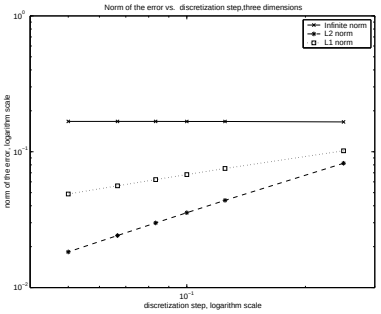
Figure 8: Convergence rate, two dimensional case, right hand side Dirac at zero, located at the center of the center cell, norm computed on a "regular zone".

6.3 Two-dimensional tests on an unstructured mesh

We also tested our algorithm on an unstructured triangular mesh. Numerical experiments for the cell centered scheme on triangular meshes were performed in [4] and in [7] in the case of coercive convection diffusion equations and regular data. These experiments show a convergence rate of order two, as in the finite element case, although this superconvergence is still, to our knowledge, an open problem in the finite volume case. We show in figure 10 the rate of convergence which we obtain for the Poisson equation where the right hand side is a Dirac at 0 and the boundary conditions are such that the exact solution is $u(x_1, x_2) = \ln(x_1^2 + x_2^2)$. The refined meshes are not imbedded, so that the convergence lines are not straight, but one can figure out that the L^1 and L^2 norms of the error between the exact and approximate solutions are bounded by 0.1 size $(\mathcal{M})^{0.7}$.

6.4 Spherical domain and mesh

We also made some experiments for a three dimensional spherical problem : we search for the solution of $-\Delta u = \delta_0$ on the Euclidean unit ball $B(0,1)$ of $\mathbb{R}^3$, with boundary conditions such that the exact solution be the restriction of the solution of $-\Delta u = \delta_0$ in the whole set $\mathbb{R}^3$. The control volumes are defined by $K_i = \{x \in B(0,1); ih \leq |x| \leq (i+1)h\}$, for $i = 0, \ldots, N$, where $h = \frac{1}{N+1/2}$. As we noted in Remark 2.1, such domain and mesh are not strictly contained in Definition 2.1 of an admissible mesh, since a sphere



	α	C
L^1 norm	0.9670	.3314
L^2 norm	0.4809	.2062

Figure 9: Convergence rate, three dimensional case, right hand side Dirac at zero, nonsymmetric discrete problem.

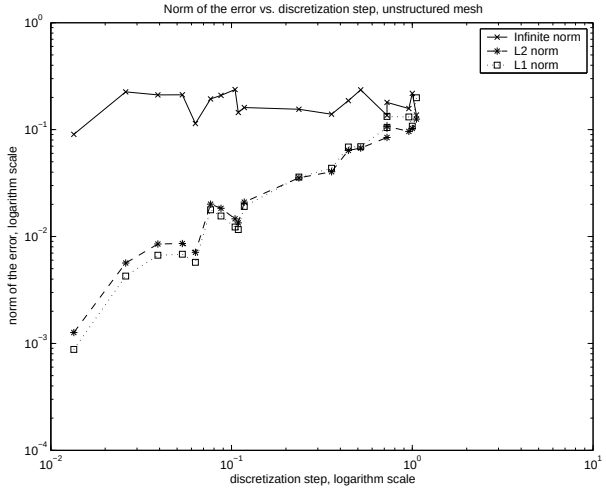
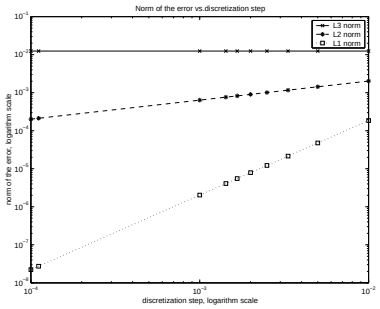


Figure 10: Convergence rate, two dimensional case, right hand side Dirac at zero, triangular mesh.

is hardly a polyhedral domain, but in fact, the discretization of the normal flux on the boundaries of such a spherical mesh is clearly consistent when looking at spherical solutions of Problem 1. Indeed, the numerical flux at interface $i+1/2$ is taken as $F_{i+1/2} = \frac{4\pi i^2 h^2}{h}(u_{i+1} - u_i)$, where the $(u_i)_{i=0,\dots,N}$ denote the discrete unknowns. In this case, the rate of convergence of the method was found to be 2 in norm L^1 and .5 in norm L^2 : see Figure 11.
Hence the symmetry of the problem seems to improve the performance of the method, at least on the L^1 norm.

Remark 6.1 *We recall that in the three-dimensional case, the exact solution $-\Delta u = \delta_0$ is in $L^{3-\varepsilon}$ for any $\varepsilon > 0$, hence we can expect a convergence in L^p for $1 \leq p < 3$. From a convergence in $L^{3-\varepsilon}$ for any $\varepsilon > 0$, and a convergence with a rate h^α in L^1, one may deduce (from Hölder's inequality) a convergence in the L^2 norm with a rate of at least $h^{\frac{\alpha}{4}-\varepsilon}$ for any $\varepsilon > 0$. The above numerical results are in accordance with this estimate, both in the spherical case and in the Cartesian case of section 6.2.*

We also give in Table 2 below the rate of convergence obtained when computing the norm of the error on a zone where the solution is regular, i.e. on the set $\{x \in \mathbb{R}^3, |x| > 1/2\}$. Again, we find in this case a rate of convergence of 2 (even a little more than 2) for all norms.



	α	C
L^1 norm	1.9288	.4993
L^2 norm	0.5000	.1879

Figure 11: Convergence rate, three dimensional case, right hand side Dirac at zero, spherical case.

	α	C
L^1 norm	2.0506	.3411
L^2 norm	2.1164	.1720
L^∞ norm	2.1295	.2331

	α	C
L^1 norm	1.0506	.1874
L^2 norm	0.9993	.1787
L^∞ norm	0.9983	.2006

Table 2: Convergence rate, three dimensional case. left: the right hand side is a Dirac measure at zero, spherical case, norm computed on a "regular zone", right: the right hand side is a two dimensional Lebesgue measure supported on the sphere of radius 1/2. The norm is computed on the whole set Ω.

If we now search for the solution of $-\Delta u = \mu$ on the three-dimensional unit ball, with μ the two dimensional Lebesgue measure supported on the sphere of radius .5, then the obtained convergence rate is again 1, even though the exact solution is more regular than the solution to the Dirac problem, see Figure 8. Note that in this case, the exact solution is in L^∞ (and even in H^1).

7 Appendix

Throughout this section, for any $q \in (1, +\infty)$, we denote by q' its conjugate exponent, that is, $q' \in (1, +\infty)$ such that $\frac{1}{q} + \frac{1}{q'} = 1$.

Proof of Proposition 2.1

The case $q = 2$ is done in [14]. We use the same method for $q \in [1, 2)$.

Define, for $\sigma \in \mathcal{E}$ and $(x, y) \in \mathbb{R}^d$, $\chi_\sigma(x, y) = 1$ if $\sigma \cap [x, y] \neq \emptyset$ and $\chi_\sigma(x, y) = 0$ otherwise. Let $\mathbf{d}$ be a unit vector and define, for $x \in \Omega$, $y(x)$ as the point on the semi-line, with origin x and direction $\mathbf{d}$, such that $y(x) \in \partial\Omega$ and $[x, y(x)] \subset \overline{\Omega}$. If $\sigma \in \mathcal{E}$, we let $c_\sigma = |\mathbf{n}_\sigma \cdot \mathbf{d}|$, where $\mathbf{n}_\sigma$ is a unit normal to σ.

For all $x \in \Omega$ such that x does not belong to an affine hyperplane generated by some $\sigma \in \mathcal{E}$, i.e. for a.e. $x \in \Omega$, we have

$$|v_\mathcal{T}(x)| \leq \sum_{\sigma \in \mathcal{E}} \chi_\sigma(x, y(x)) D_\sigma v_\mathcal{T}$$

(recall that $v_\mathcal{T}(x) = v_K$ for the $K \in \mathcal{T}$ such that $x \in K$). Take such an x and suppose that, for some $\sigma \in \mathcal{E}$, $c_\sigma = 0$; we have then $\chi_\sigma(x, y(x)) = 0$ (indeed, otherwise x would belong to the affine hyperplane generated by σ). Thus, the preceding sum can be reduced to the $\sigma \in \mathcal{E}$ such that $c_\sigma \neq 0$ and we can write, thanks to Hölder's inequality, for a.e. $x \in \Omega$,

$$|v_\mathcal{T}(x)|^q \leq \left(\sum_{\sigma \in \mathcal{E} \mid c_\sigma \neq 0} \chi_\sigma(x, y(x)) d_\sigma c_\sigma^{-\frac{q}{q'}} \left(\frac{D_\sigma v_\mathcal{T}}{d_\sigma} \right)^q \right) \left(\sum_{\sigma \in \mathcal{E} \mid c_\sigma \neq 0} \chi_\sigma(x, y(x)) d_\sigma c_\sigma \right)^{\frac{q}{q'}}. \tag{63}$$

Since $\sum_{\sigma\in\mathcal{E}} \chi_\sigma(x,y(x))d_\sigma c_\sigma \le \mathrm{diam}(\Omega)$ for all $x\in\Omega$ (see [14]) and $\int_\Omega \chi_\sigma(x,y(x))\,d\lambda(x) \le \mathrm{diam}(\Omega)\mathrm{meas}(\sigma)c_\sigma$, we obtain, integrating (63) on Ω,

$$\int_\Omega |v_\mathcal{T}|^q\,d\lambda \le \mathrm{diam}(\Omega)^{\frac{q}{q'}} \sum_{\sigma\in\mathcal{E}\,|\,c_\sigma\neq 0} \mathrm{diam}(\Omega)\mathrm{meas}(\sigma)d_\sigma c_\sigma^{1-\frac{q}{q'}} \left(\frac{D_\sigma v_\mathcal{T}}{d_\sigma}\right)^q.$$

But $q\le 2$, so that $1-\frac{q}{q'} = 2-q \ge 0$ and $c_\sigma^{2-q}\le 1$, which concludes this proof. ■

Proof of Proposition 2.2
The case $d=2$ has already been done in the course of the proof of the discrete Sobolev inequalities in [14] (inequality (9.73), page 791).
For $d=3$, the case $q=2$ may be found in [8]. The case of a general q is similar; we use the following inequality (inequality (9.75) page 793 of [14]) : for any $w_\mathcal{T}\in X(\mathcal{T})$,

$$\int_\Omega |w_\mathcal{T}|^{\frac{3}{2}}\,d\lambda \le \left(\sum_{\sigma\in\mathcal{E}} \mathrm{meas}(\sigma) D_\sigma w_\mathcal{T}\right)^{3/2}.$$

Applying this to $w_K = |v_K|^{\frac{2q}{3-q}}\mathrm{sgn}(v_K)$, and since $D_\sigma w_\mathcal{T} \le \frac{2q}{3-q}(|v_K|^{\frac{3(q-1)}{3-q}} + |v_L|^{\frac{3(q-1)}{3-q}})D_\sigma v_\mathcal{T}$ (with $\sigma = K|L\in\mathcal{E}_{\mathrm{int}}$ or $v_L=0$ if $\sigma\in\mathcal{E}_{\mathrm{ext}}\cap\mathcal{E}_K$), we deduce, by the Hölder inequality,

$$\begin{aligned}
&\left(\int_\Omega |v_\mathcal{T}|^{\frac{3q}{3-q}}\,d\lambda\right)^{2/3} \\
&\quad\le \frac{2q}{3-q}\sum_{\sigma\in\mathcal{E}} \mathrm{meas}(\sigma)d_\sigma(|v_K|^{\frac{3(q-1)}{3-q}} + |v_L|^{\frac{3(q-1)}{3-q}})\frac{D_\sigma v_\mathcal{T}}{d_\sigma} \\
&\quad\le \frac{2q}{3-q}\left(\sum_{\sigma\in\mathcal{E}} \mathrm{meas}(\sigma)d_\sigma\left(\frac{D_\sigma v_\mathcal{T}}{d_\sigma}\right)^q\right)^{1/q}\left(\sum_{\sigma\in\mathcal{E}} \mathrm{meas}(\sigma)d_\sigma(2^{q'-1}|v_K|^{\frac{3q}{3-q}} + 2^{q'-1}|v_L|^{\frac{3q}{3-q}})\right)^{1/q'}.
\end{aligned}$$

But, by hypothesis on ζ,

$$\begin{aligned}
\sum_{\sigma\in\mathcal{E}} \mathrm{meas}(\sigma)d_\sigma|v_K|^{\frac{3q}{3-q}} &= \sum_{K\in\mathcal{T}} |v_K|^{\frac{3q}{3-q}} \sum_{\sigma\in\mathcal{E}_K} \mathrm{meas}(\sigma)d_\sigma \\
&\le \frac{1}{\zeta}\sum_{K\in\mathcal{T}} |v_K|^{\frac{3q}{3-q}} \sum_{\sigma\in\mathcal{E}_K} \mathrm{meas}(\sigma)d_{K,\sigma} \\
&= \frac{3}{\zeta}\sum_{K\in\mathcal{T}} \mathrm{meas}(K)|v_K|^{\frac{3q}{3-q}} \\
&= \frac{3}{\zeta}||v_\mathcal{T}||^{\frac{3q}{3-q}}_{L^{\frac{3q}{3-q}}(\Omega)}.
\end{aligned}$$

Thus, we finally have

$$\left(\int_\Omega |v_\mathcal{T}|^{\frac{3q}{3-q}}\,d\lambda\right)^{2/3} \le C||v_\mathcal{T}||_{1,q,\mathcal{M}}||v_\mathcal{T}||^{\frac{3(q-1)}{3-q}}_{L^{\frac{3q}{3-q}}(\Omega)}$$

where C only depends on (q,ζ), and this gives the desired estimate. ■

Proof of Proposition 2.3
Define $\chi_\sigma(x,y)$ as at the beginning of the proof of Proposition 2.1.
Suppose first that $q>1$ and take $h\in\mathbb{R}^d\backslash\{0\}$. Denote, for $\sigma\in\mathcal{E}$, $c_\sigma = |\mathbf{n}_\sigma\cdot\frac{h}{|h|}|$ (where $\mathbf{n}_\sigma$ is a unit normal to σ).

We have, for a.e. $x \in \Omega$ (in fact for all x which does not belong to an affine hyperplane generated by some $\sigma \in \mathcal{E}$),

$$|w_{\mathcal{T}}(x+h) - w_{\mathcal{T}}(x)| \leq \sum_{\sigma \in \mathcal{E}} \chi_\sigma(x+h,x) D_\sigma v_{\mathcal{T}}. \tag{64}$$

As in the proof of Proposition 2.1, this sum can be limited to those $\sigma \in \mathcal{E}$ such that $c_\sigma \neq 0$, and we have then, by Hölder, for a.e. $x \in \Omega$,

$$|w_{\mathcal{T}}(x+h) - w_{\mathcal{T}}(x)| \leq \left(\sum_{\sigma \in \mathcal{E} \mid c_\sigma \neq 0} \frac{\chi_\sigma(x+h,x) d_\sigma}{c_\sigma} \left(\frac{D_\sigma v_{\mathcal{T}}}{d_\sigma} \right)^q \right)^{1/q} \left(\sum_{\sigma \in \mathcal{E}} \chi_\sigma(x+h,x) d_\sigma c_\sigma^{q'/q} \right)^{1/q'}.$$

Since $q \leq 2$ (and hence $q'/q \geq 1$) and $c_\sigma \in [0,1]$, we have $c_\sigma^{q'/q} \leq c_\sigma$; but (see [14]) $\sum_{\sigma \in \mathcal{E}} \chi_\sigma(x+h,x) d_\sigma c_\sigma \leq |h| + C\text{size}(\mathcal{M})$, where C only depends on Ω. Thus,

$$|w_{\mathcal{T}}(x+h) - w_{\mathcal{T}}(x)|^q \leq (|h| + C\text{size}(\mathcal{M}))^{q-1} \sum_{\sigma \in \mathcal{E} \mid c_\sigma \neq 0} \frac{\chi_\sigma(x+h,x) d_\sigma}{c_\sigma} \left(\frac{D_\sigma v_{\mathcal{T}}}{d_\sigma} \right)^q.$$

Since $\int_{\mathbb{R}^d} \chi_\sigma(x+h,x)\, d\lambda(x) \leq \text{meas}(\sigma) c_\sigma |h|$, we deduce, after integrating, the desired estimate (17).
If $q = 1$, we simply integrate (64) and this directly gives (bounding $\int_{\mathbb{R}^d} \chi_\sigma(x+h,x)\, d\lambda(x)$ by $\text{meas}(\sigma)|h|$) the estimate.
The compactness result is then an immediate application of Kolmogorov's Theorem, with the use of Proposition 2.1 to obtain a bound in $L^q(\Omega)$. ■

Proof of Proposition 2.4
Applying (17) to v_n and passing to the limit $n \to \infty$, we get, for $h \in \mathbb{R}^d \backslash \{0\}$,

$$\int_{\mathbb{R}^d} \frac{|w(x+h) - w(x)|^q}{h^q}\, d\lambda(x) \leq C,$$

where C does not depend on h and w is the extension of v to $\mathbb{R}^d$ by 0 outside Ω.
Since $q > 1$, this estimate classically gives $w \in W^{1,q}(\mathbb{R}^d)$ and, by the regularity of Ω, since w is the extension of v by 0 outside Ω, $v \in W_0^{1,q}(\Omega)$. ■

References

[1] Adams R.A., Sobolev Spaces. Academic Press (1975).

[2] Boccardo L., Gallouët T., *Nonlinear elliptic and parabolic equations involving measure data.* J. Funct. Anal. **87** (1989), 241-273.

[3] Boccardo L., Gallouët T., Vazquez J-L., *Nonlinear elliptic equations in $\mathbb{R}^N$ without grow restrictions on the data.* Jour. Diff. Eqns. **105** (1993), 334-363.

[4] Boivin S., Cayré F., Hérard J.M., *A finite volume method to solve the Navier-Stokes equations for incompressible flows on unstructured meshes.* Int. J. Therm. Sci. **39** (2000), 806-825.

[5] Chou, S. H., Vassilevski P.S. , A general mixed covolume framework for constructing conservative schemes for elliptic problems, Math. Comp. **68** (1999), 991-1011.

[6] Coudière Y., Gallouët T., Herbin R., *Discrete Sobolev Inequalities and L^p error estimates for approximate finite volume solutions of convection diffusion equations*, M2AN, 35,4, 767-778, 2001.

[7] Eymard R., Gallouët T., Herbin R., *Finite volume approximation of elliptic problems and convergence of an approximate gradient*, Appl. Num. Math., 37/1-2, pp. 31 - 53 (2001).

[8] Coudière Y., Vila J.P., Villedieu p., *Convergence Rate of a Finite Volume Scheme for a Two Dimensionnal Convection Diffusion Problem*, M2AN, 33(3): 493-516, 1999.

[9] Droniou J., *Noncoercive Linear Elliptic Problems*, accepted for publication in Potential Analysis.

[10] Droniou J., *PhD Thesis*, CMI, Université de Provence, 2001.

[11] Droniou J., Gallouët T, *A finite volume scheme for noncoercive Dirichlet problems with right-hand side in H^{-1}*, in Finite volume for complex applications III, R. Herbin and D. Kröner eds, Hermes Penton Science (2002), 195—202.

[12] Droniou J., Gallouët T, *Finite volume methods for convection-diffusion equations with right-hand side in H^{-1}*, M2AN, Vol. 36 No. 4 (2002).

[13] Droniou J., Gallouët T., *A uniqueness result for quasilinear elliptic equations with measures as data*, Rendiconti di Matematica, Serie VII, Volume 21, Roma (2001), 57-86.

[14] Eymard R., Gallouët T., Herbin R., Finite Volume Methods, Handbook of Numerical Analysis, Vol. VII, pp. 713-1020. Edited by P.G. Ciarlet and J.L. Lions (North Holland).

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

[16] Eymard R., Gallouët T., Herbin R., Michel A., Convergence of a finite volume scheme for nonlinear degenerate parabolic equations, Num. Math, 2002, 92: 41-82. R. Eymard, T. Gallou?t, R. Herbin, A. Michel.

[17] Fiard J.M., Herbin R., *Comparison between finite volume finite element methods for the numerical simulation of an elliptic problem arising in electrochemical engineering*, Comput. Meth. Appl. Mech. Engin., 115, 315-338 (1994).

[18] Forsyth P.A., Sammon P.H., *Quadratic Convergence for Cell-Centered Grids*, Appl. Num. Math. 4 (1988), 377-394.

[19] Gallouët T., Herbin R., *Finite volume methods for diffusion problems and irregular data*, in Finite volumes for complex applications, Problems and Perspectives, II, F. Benkhaldoun, M. Hänel and R. Vilsmeier eds, Hermes, 155–162 (1999).

[20] Gallouët T., Herbin R., Vignal M.H., *Error estimates for the approximate finite volume solution of convection diffusion equations with general boundary conditions*, SIAM J. Numer. Anal, 37, 6, 1935 - 1972, 2000.

[21] Gallouët T., Monier A., *On the regularity of solutions to elliptic equations*, Rend. Mat., VII **19** (1999), 471-488.

[22] Grisvard P., Elliptic problems in nonsmooth domains, Pitman 1985.

[23] Herbin R., *An error estimate for a finite volume scheme for a diffusion-convection problem on a triangular mesh*, Num. Meth. P.D.E. **11**, 165-173 (1995).

[24] Herbin R., *Finite volume approximation of elliptic problems with irregular data*, Finite volumes for complex applications, Problems and Perspectives, F. Benkhaldoun, D. Hanel and R. Vilsmeier eds, Hermes, 1999,153-160.

[25] Gallouët T., Herbin R., *Finite volume methods for diffusion convection equations on general meshes*, Finite volumes for complex applications, Problems and Perspectives, F. Benkhaldoun and R. Vilsmeier eds, Hermes, 1996, 153-160.

[26] Lazarov R.D., Mishev I.D., Vassilevski P.S., *Finite volume methods for convection-diffusion problems*, SIAM J. Numer. Anal., 33, 1996, 31-55.

[27] Manteufel T.A., White A.B. (1986), *The numerical solution of second order boundary value problem on non uniform meshes*, Math. Comput. **47**, 511-536.

[28] Meyers N.G., *An L^p estimate for the gradient of solutions of second order divergence equations*, Ann. Sc. Norm. Sup. Pisa, **17** (1963), 189-206.

[29] Mishev I.D., *Finite volume methods on Voronoï meshes*, Num. Meth. P.D.E. 14, 2, 193-212 (1998).