



HAL
open science

Cell centred discretisation of non linear elliptic problems on general multidimensional polyhedral grids

Robert Eymard, Thierry Gallouët, Raphaelae Herbin

► **To cite this version:**

Robert Eymard, Thierry Gallouët, Raphaelae Herbin. Cell centred discretisation of non linear elliptic problems on general multidimensional polyhedral grids. *Journal of Numerical Mathematics*, 2009, 17 (3), pp.173–193. 10.1515/JNUM.2009.010 . hal-00351867

HAL Id: hal-00351867

<https://hal.science/hal-00351867>

Submitted on 12 Jan 2009

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

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

Cell centred discretisation of non linear elliptic problems on general multidimensional polyhedral grids*

R. Eymard[†] T. Gallouët[‡] and R. Herbin[§]

January 10, 2009

Abstract

This work is devoted to the discretisation of non linear elliptic problems on general polyhedral meshes in several space dimensions. The SUSHI scheme which was recently studied for anisotropic heterogeneous problems is applied in its full barycentric version, thus resulting into a cell centred scheme written under variational form, also known as 'SUCCES'. We prove the existence of the approximate solution and its convergence to the weak solution of the continuous solution as the mesh size tends to 0. Numerical examples are shown for the p -Laplacian.

1 Introduction

We study the following problem: find an approximation of u , weak solution to the following equation:

$$\begin{aligned} -\operatorname{div} \mathbf{a}(\mathbf{x}, \nabla u) &= f \text{ in } \Omega, \\ u &= 0 \text{ on } \partial\Omega, \end{aligned} \tag{1}$$

where we denote by $\partial\Omega = \overline{\Omega} \setminus \Omega$ the boundary of Ω , under the following assumptions:

$$\Omega \text{ is an open bounded connected polygonal subset of } \mathbb{R}^d, \quad d \in \mathbb{N}^*, \tag{2a}$$

$$\mathbf{a} : \Omega \times \mathbb{R}^d \rightarrow \mathbb{R}^d \text{ is a Caratheodory function,} \tag{2b}$$

(e.g. a function such that for a.e. $\mathbf{x} \in \Omega$, $\xi \mapsto a(\mathbf{x}, \xi)$ is continuous, and for any $\xi \in \mathbb{R}^d$ the function $\mathbf{x} \mapsto a(\mathbf{x}, \xi)$ is measurable)

$$\exists \underline{a} \in (0, +\infty), p \in (1, +\infty) ; \mathbf{a}(\mathbf{x}, \xi) \cdot \xi \geq \underline{a} |\xi|^p, \text{ for a.e. } \mathbf{x} \in \Omega, \forall \xi \in \mathbb{R}^d, \tag{2c}$$

$$(\mathbf{a}(\mathbf{x}, \xi) - \mathbf{a}(\mathbf{x}, \chi)) \cdot (\xi - \chi) > 0, \text{ for a.e. } \mathbf{x} \in \Omega, \forall \xi, \chi \in \mathbb{R}^d \text{ with } \xi \neq \chi, \tag{2d}$$

$$\exists \bar{a} \in L^{p'}(\Omega), \Lambda \in (0, +\infty) ; |\mathbf{a}(\mathbf{x}, \xi)| \leq \bar{a}(\mathbf{x}) + \Lambda |\xi|^{p-1}, \text{ for a.e. } \mathbf{x} \in \Omega, \forall \xi \in \mathbb{R}^d, \tag{2e}$$

and

$$f \in L^{p'}(\Omega) \text{ where } p' = \frac{p}{p-1}. \tag{2f}$$

*This work was supported by Groupement MOMAS, CNRS/PACEN

[†]Université Paris Est, France, Robert.Eymard@univ-paris-est.fr

[‡]Université de Provence, France, Thierry.Gallouet@cmi.univ-mrs.fr

[§]Université de Provence, France, Raphael.Herbin@cmi.univ-mrs.fr

If the function \mathbf{a} satisfies (2b)-(2e), then the mapping $u \mapsto -\operatorname{div}\mathbf{a}(\cdot, \nabla u)$ is a Leray-Lions operator, a classical example of which is the p -Laplacian operator, obtained by setting

$$\begin{aligned}\mathbf{a}(\mathbf{x}, \xi) &= |\xi|^{p-2}\xi, \quad \forall \xi \in \mathbb{R}^d \setminus \{0\}, \quad \text{for a.e. } \mathbf{x} \in \Omega, \\ \mathbf{a}(\mathbf{x}, 0) &= 0, \quad \text{for a.e. } \mathbf{x} \in \Omega.\end{aligned}\tag{3}$$

It is well-known [27] that under hypotheses (2), there exists a unique weak solution (1), that is a function u satisfying:

$$\begin{cases} u \in W_0^{1,p}(\Omega), \\ \int_{\Omega} \mathbf{a}(\mathbf{x}, \nabla u(\mathbf{x})) \cdot \nabla v(\mathbf{x}) d\mathbf{x} = \int_{\Omega} f(\mathbf{x})v(\mathbf{x})d\mathbf{x}, \quad \forall v \in W_0^{1,p}(\Omega).\end{cases}\tag{4}$$

The numerical approximation of (4) has been the subject of numerous works: finite element methods have been extensively studied [23, 24, 22, 11, 28, 9, 25, 16]. Finite volume schemes have also been addressed, first on Cartesian grids [5, 6]; more recent works [10, 8, 7] are concerned with the DDFV (Discrete Duality Finite Volume) scheme for general two-dimensional grids, introduced in [26] and first analysed in [12] for linear problems; these schemes require the use of two dual grids and sets of unknowns. In [13] the mixed finite volume scheme involving unknowns at the faces of the mesh, first introduced and analysed in [14] is shown to converge for general multidimensional grids.

In this paper, we study a numerical scheme which satisfies the following properties:

- It is a cell centred scheme, that is there is only one unknown associated to each grid cell. The advantage of such method is that it can be more easily extended to nonlinear coupled problems.
- It is defined for general, possibly non-conforming, polygonal meshes in any space dimension.
- The approximate problem has one and only one solution, so that the resulting nonlinear system of equations may be solved by an adequate (iterative) method.
- The convergence of the approximate solution and gradient to the exact solution and gradient when the mesh size goes to 0 are proven.

The scheme "SUCCES" (Scheme Using Conservativity and Consistency error Stabilisation) that we use here was first introduced in [21] for the discretisation of the Laplace operator in the incompressible Navier-Stokes equations on general multidimensional grids. Its numerical performance for the numerical simulation of flow in heterogeneous porous media is demonstrated in [3]. Its convergence analysis for linear problems is carried out in [19] in the more general framework of the SUSHI (Scheme Using Stabilisation and Hybrid Interfaces) scheme, for which edge unknowns may be present or not, according to the user's choice.

The contents of this paper is the following. In Section 2, we present the discretisation scheme and a few basic properties. Some estimates (using discrete Sobolev embeddings) are presented in Section 3, along with the proof of the existence and uniqueness of the approximate solution. In Section 4, we prove the convergence results, which are based on some "discrete functional analysis" tools, mimicking classical functional analysis tools: the Kolmogorov theorem is used to get some compactness results (strong convergence for the approximate solution, weak convergence for the approximate gradient). Then the well-known 'Minty trick' [27] is used to overcome the problem of passing to the limit in a nonlinear monotone problem under only weak convergence properties for the approximate gradient of the solution. Finally, the Leray-Lions trick of [27] allows us to recover the strong convergence of this approximate gradient. The numerical behaviour of the scheme is shown on an example in Section 5.

2 The finite volume scheme

Let us first begin with the detailed description of the mesh under consideration in this paper.

Definition 2.1 (Space discretisation) *Let Ω be a polyhedral open bounded connected subset of \mathbb{R}^d , with $d \in \mathbb{N} \setminus \{0\}$, and $\partial\Omega = \bar{\Omega} \setminus \Omega$ its boundary. A discretisation of Ω , denoted by \mathcal{D} , is defined as the triplet $\mathcal{D} = (\mathcal{M}, \mathcal{E}, \mathcal{P})$, where:*

1. \mathcal{M} is a finite family of non empty connected open disjoint subsets of Ω (the “control volumes”) such that $\bar{\Omega} = \cup_{K \in \mathcal{M}} \bar{K}$. For any $K \in \mathcal{M}$, let $\partial K = \bar{K} \setminus K$ be the boundary of K ; let $|K| > 0$ denote the measure of K and h_K denote the diameter of K .
2. \mathcal{E} is a finite family of disjoint subsets of $\bar{\Omega}$ (the “edges” of the mesh), such that, for all $\sigma \in \mathcal{E}$, σ is a non empty open subset of a hyperplane of \mathbb{R}^d , whose $(d-1)$ -dimensional measure $|\sigma|$ is strictly positive. We also assume that, for all $K \in \mathcal{M}$, there exists a subset \mathcal{E}_K of \mathcal{E} such that $\partial K = \cup_{\sigma \in \mathcal{E}_K} \bar{\sigma}$. For any $\sigma \in \mathcal{E}$, we denote by $\mathcal{M}_\sigma = \{K \in \mathcal{M}, \sigma \in \mathcal{E}_K\}$. We then assume that, for all $\sigma \in \mathcal{E}$, either \mathcal{M}_σ has exactly one element and then $\sigma \subset \partial\Omega$ (the set of these interfaces, called boundary interfaces, is denoted by \mathcal{E}_{ext}) or \mathcal{M}_σ has exactly two elements (the set of these interfaces, called interior interfaces, is denoted by \mathcal{E}_{int}). For all $\sigma \in \mathcal{E}$, we denote by \mathbf{x}_σ the barycentre of σ . For all $K \in \mathcal{M}$ and $\sigma \in \mathcal{E}_K$, we denote by $\mathbf{n}_{K,\sigma}$ the unit vector normal to σ outward to K .
3. \mathcal{P} is a family of points of Ω indexed by \mathcal{M} , denoted by $\mathcal{P} = (\mathbf{x}_K)_{K \in \mathcal{M}}$, such that for all $K \in \mathcal{M}$, $\mathbf{x}_K \in K$ and K is assumed to be \mathbf{x}_K -star-shaped, which means that for all $\mathbf{x} \in K$, the property $[\mathbf{x}_K, \mathbf{x}] \subset K$ holds. Denoting by $d_{K,\sigma}$ the Euclidean distance between \mathbf{x}_K and the hyperplane including σ , one assumes that $d_{K,\sigma} > 0$. We then denote by $D_{K,\sigma}$ the cone with vertex \mathbf{x}_K and basis σ .

Remark 2.1 (Non convex generalized hexahedra) *The above definition applies to a large variety of meshes. Note that no hypothesis is made on the convexity of the control volumes; in fact, generalised hexahedra, i.e. with faces which may be composed of several planar sub-faces may be used. Often encountered in underground flow simulations, such hexahedra may have up to 12 faces (resp. 24 faces) if each non planar face is composed of two triangles (resp. four triangles), but only 6 neighbouring control volumes.*

Let $\mathcal{D} = (\mathcal{M}, \mathcal{E}, \mathcal{P})$ be a discretisation of Ω in the sense of Definition 2.1. The size of the discretisation \mathcal{D} is defined by:

$$h_{\mathcal{D}} = \sup\{h_K, K \in \mathcal{M}\}.$$

Let $H_{\mathcal{D}} \subset L^2(\Omega)$ be the set of piece-wise constant functions on the control volumes of the mesh and let u_K denote the (constant) value of u on K .

A first simple idea to find a scheme that will approximate (1) is to use the usual non conforming Galerkin method: assuming that for $v \in H_{\mathcal{D}}$ we know how to construct an adequate discrete gradient $\nabla_{\mathcal{D}} v$ (expected to be an approximation of the gradient of the exact solution), we then seek $u_{\mathcal{D}} \in H_{\mathcal{D}}$ such that

$$\int_{\Omega} \mathbf{a}(\mathbf{x}, \nabla_{\mathcal{D}} u_{\mathcal{D}}(\mathbf{x})) \cdot \nabla_{\mathcal{D}} v(\mathbf{x}) \, d\mathbf{x} = \int_{\Omega} f(\mathbf{x}) v(\mathbf{x}) \, d\mathbf{x}, \quad \forall v \in H_{\mathcal{D}}. \quad (5)$$

The construction of the discrete gradient is obtained from the values of the discrete function on the cell and from its reconstructed values on the edges. These are constructed in the following way. For $K \in \mathcal{M}$, one chooses $\mathbf{x}_K \in \mathcal{M}$ (this choice is possible for any set K which is “ \mathbf{x}_K star shaped”). For any interior edge (interface) σ of \mathcal{M} , choose some points \mathbf{x}_M of the mesh close to σ and write \mathbf{x}_σ (recall that \mathbf{x}_σ is the barycentre of σ) as a combination of these points: $\mathbf{x}_\sigma = \sum_{M \in \mathcal{M}} \beta_\sigma^M \mathbf{x}_M$. Note

that there is no need for this combination to be convex. The coefficients of the combination, however, should be bounded.

Then, for any $v \in H_{\mathcal{D}}$, one sets:

$$\Pi_{\sigma}u = \begin{cases} \sum_{M \in \mathcal{M}} \beta_{\sigma}^M u_M & \text{if } \sigma \in \mathcal{E}_{\text{int}}, \\ 0 & \text{if } \sigma \in \mathcal{E}_{\text{ext}}. \end{cases} \quad (6)$$

Next, denoting by Id the $d \times d$ identity matrix, we use the following geometrical relationship

$$\frac{1}{|K|} \sum_{\sigma \in \mathcal{E}_K} |\sigma| \mathbf{n}_{K,\sigma} (\mathbf{x}_{\sigma} - \mathbf{x}_K)^t = \text{Id}, \quad \forall K \in \mathcal{M}, \quad (7)$$

(which is an easy consequence of the Stokes formula), to define a consistent discrete gradient $\nabla_{\mathcal{D}}v \in H_{\mathcal{D}}$ of a function $v \in H_{\mathcal{D}}$ as the piece-wise constant function equal to its constant value $\nabla_K u$ a.e. in K :

$$\nabla_{\mathcal{D}}u(\mathbf{x}) = \nabla_K u \text{ for a.e. } \mathbf{x} \in K, \quad \forall K \in \mathcal{M}. \quad (8a)$$

$$\nabla_K u = \frac{1}{|K|} \sum_{\sigma \in \mathcal{E}_K} |\sigma| (\Pi_{\sigma}u - u_K) \mathbf{n}_{K,\sigma}. \quad (8b)$$

We may then try to find $u_{\mathcal{D}} \in H_{\mathcal{D}}$ solution of (5); however, it is easily seen that the problem lacks coercivity because of the definition of the discrete gradient. Take for instance $d = 1$, $\Omega = (0, 1)$ and a uniform mesh with step size $h = 1/N$. For each cell $K_i = ((i-1)h, ih)$ ($i = 1, \dots, N$), take $x_{K_i} = x_i = (i-1)h + \frac{h}{2}$. A natural combination to compute the interface values is $x_{i+1/2} = \frac{1}{2}(x_i + x_{i+1})$, $i = 1, \dots, N-1$ (and the boundary conditions give $x_{\frac{1}{2}} = x_{N+\frac{1}{2}} = 0$). Then, the discrete gradient of $v = (v_i)_{i=1,N}$ has the value $\frac{1}{2h}(v_{i+1} - v_{i-1})$ on cell K_i for $i = 2, \dots, N-1$, the value $\frac{1}{2h}(v_1 + v_2)$ on cell K_1 and the value $\frac{-1}{2h}(v_{N-1} + v_N)$ on cell K_N . Taking $v_i = (-1)^i$ for all i leads to a discrete gradient equal 0 in all cells. This problem must be cured to obtain the coercivity of the discrete problem and the uniqueness of its solution.

A remedy to this problem was found by using a stabilisation which uses a consistency estimate [18, 21, 19]. Let us introduce, for any $v \in H_{\mathcal{D}}$, for any $K \in \mathcal{M}$ and $\sigma \in \mathcal{E}_K$, the value $R_{K,\sigma}v$ defined by:

$$R_{K,\sigma}v = \frac{1}{d_{K,\sigma}} (\Pi_{\sigma}v - v_K - \nabla_K u \cdot (\mathbf{x}_{\sigma} - \mathbf{x}_K)). \quad (9)$$

Using this value, we define the function $R_{\mathcal{D}}v$ by the constant value $R_{K,\sigma}v$ in the cone $D_{K,\sigma}$:

$$R_{\mathcal{D}}v(\mathbf{x}) = R_{K,\sigma}v \text{ for a.e. } \mathbf{x} \in D_{K,\sigma}, \quad \forall K \in \mathcal{M}, \quad \forall \sigma \in \mathcal{E}_K. \quad (10)$$

We now consider the following approximate problem:

find $u \in H_{\mathcal{D}}$ such that

$$\langle u, v \rangle_{\mathcal{D}} = \int_{\Omega} \mathbf{a}(\mathbf{x}, \nabla_{\mathcal{D}}u(\mathbf{x})) \cdot \nabla_{\mathcal{D}}v(\mathbf{x}) \, d\mathbf{x} + b(u, v) = \int_{\Omega} f(\mathbf{x}) v(\mathbf{x}) \, d\mathbf{x}, \quad \forall v \in H_{\mathcal{D}}. \quad (11)$$

with

$$b(u, v) = \int_{\Omega} |R_{\mathcal{D}}u(\mathbf{x})|^{p-1} \text{sgn}(R_{\mathcal{D}}u(\mathbf{x})) R_{\mathcal{D}}v(\mathbf{x}) \, d\mathbf{x}, \quad (12)$$

introducing the function $\text{sgn}(x) = 1$ for all $x \in]0, +\infty[$, $\text{sgn}(x) = -1$ for all $x \in]-\infty, 0[$ and $\text{sgn}(0) = 0$. We now define, for all $K \in \mathcal{M}$, the function

$$\begin{aligned} \mathbf{a}_K &: \mathbb{R}^d \rightarrow \mathbb{R}^d \\ \forall \xi \in \mathbb{R}^d, \mathbf{a}_K(\xi) &= \frac{1}{|K|} \int_K \mathbf{a}(\mathbf{x}, \xi) \, d\mathbf{x}, \end{aligned} \quad (13)$$

we can then write that

$$\langle u, v \rangle_{\mathcal{D}} = \sum_{K \in \mathcal{M}} \left(|K| \mathbf{a}_K (\nabla_K u) \cdot \nabla_K v + \sum_{\sigma \in \mathcal{E}_K} \frac{|\sigma| d_{K,\sigma}}{d} |R_{K,\sigma} u|^{p-1} \operatorname{sgn}(R_{K,\sigma} u) R_{K,\sigma} v \right).$$

Remark 2.2 (An alternate scheme) We could also proceed as in [19]: define $\nabla_{K,\sigma} v$ in the cone $D_{K,\sigma}$:

$$\nabla_{K,\sigma} v = \nabla_K v + R_{K,\sigma} v \mathbf{n}_{K,\sigma},$$

then define $\nabla_{\mathcal{D}} v$ by the constant value $\nabla_{K,\sigma} v$ in the cone $D_{K,\sigma}$, and then define the scheme by (5). This scheme presents similar properties of convergence to the one which is studied here, under conditions of regularity on the mesh which ensure that there exist two positive reals α, β independent of the mesh such that

$$\alpha \|v\|_{p,\Pi} \leq \|\nabla_{\mathcal{D}} v\|_{L^p(\Omega)} \leq \beta \|v\|_{p,\Pi},$$

where $\|\cdot\|_{p,\Pi}$ is defined by (18). This can be shown under more restrictive hypotheses than those used here.

3 Estimates, existence and uniqueness of the approximate solution

We define the following discrete $W_0^{1,p}$ norm on $H_{\mathcal{D}}$:

$$\|u\|_{1,p,\mathcal{D}}^p = \sum_{K \in \mathcal{M}} \sum_{\sigma \in \mathcal{E}_K} |\sigma| d_{K,\sigma} \left(\frac{D_{\sigma} u}{d_{\sigma}} \right)^p, \quad (14)$$

where $d_{\sigma} = d_{K,\sigma} + d_{L,\sigma}$, if $\mathcal{M}_{\sigma} = \{K, L\}$, and $d_{\sigma} = d_{K,\sigma}$, if $\mathcal{M}_{\sigma} = \{K\}$ and $D_{\sigma} u = |u_K - u_L|$ if $\mathcal{M}_{\sigma} = \{K, L\}$ and $D_{\sigma} u = |u_K - 0|$ if $\mathcal{M}_{\sigma} = \{K\}$.

Let us first recall the following results, which are proven in [19].

Lemma 3.1 (Discrete Sobolev inequality) Let $d \geq 1$, $1 \leq p < \infty$ and let Ω be a polyhedral open bounded connected subset of \mathbb{R}^d . Let \mathcal{D} be a mesh of Ω in the sense of Definition 2.1. Let $\eta > 0$ be such that $\eta \leq d_{K,\sigma}/d_{L,\sigma} \leq 1/\eta$ for all $\sigma \in \mathcal{E}$, where $\mathcal{M}_{\sigma} = \{K, L\}$. Then, there exists $q > p$ only depending on p ($q = \frac{pd}{d-p}$, for instance, in the case $1 < p < d$, and q is any value in $]p, \infty[$ in the case $d \leq p$) and there exists C_1 , only depending on d , Ω , p and η such that

$$\|u\|_{L^q(\Omega)} \leq C_1 \|u\|_{1,p,\mathcal{D}} \quad \forall u \in H_{\mathcal{D}}, \quad (15)$$

where $\|u\|_{1,p,\mathcal{D}}^p$ is defined by (14).

Lemma 3.2 (Compactness in L^p) Let $d \geq 1$, $1 \leq p < \infty$ and Ω be a polyhedral open bounded connected subset of \mathbb{R}^d . Let \mathcal{F} be a family of meshes of Ω in the sense of Definition 2.1. Let $\eta > 0$ be such that, for all $\mathcal{D} \in \mathcal{F}$, one has $\eta \leq d_{K,\sigma}/d_{L,\sigma} \leq 1/\eta$ for all $\sigma \in \mathcal{E}$, where $\mathcal{M}_{\sigma} = \{K, L\}$. For $\mathcal{D} \in \mathcal{F}$, let $u_{\mathcal{D}} \in H_{\mathcal{D}}$ and assume that there exists $C \in \mathbb{R}$ such, for all $\mathcal{D} \in \mathcal{F}$, $\|u_{\mathcal{D}}\|_{1,p,\mathcal{D}} \leq C$. Then, the family $(u_{\mathcal{D}})_{\mathcal{D} \in \mathcal{F}}$ is relatively compact in $L^p(\Omega)$ and also in $L^p(\mathbb{R}^d)$ taking $u_{\mathcal{D}} = 0$ outside Ω .

We introduce a measure of the regularity of the discretisation by

$$\theta_{\mathcal{D}} = \max \left\{ \frac{h_K}{d_{K,\sigma}}, \frac{d_{K,\sigma}}{d_{L,\sigma}}, K \in \mathcal{M}, \sigma \in \mathcal{E}_K \cap \mathcal{E}_{\text{int}}, \mathcal{M}_{\sigma} = \{K, L\} \right\}. \quad (16)$$

Note that if $\theta_{\mathcal{D}}$ is bounded, the hypotheses of Lemmas 3.1 and 3.2 hold. We next define:

$$\theta_{\mathcal{D},\Pi} = \max \left(\theta_{\mathcal{D}}, \left\{ \frac{\sum_{M \in \mathcal{M}} |\beta_{\sigma}^M| |\mathbf{x}_M - \mathbf{x}_{\sigma}|^2}{h_K^2}, K \in \mathcal{M}, \sigma \in \mathcal{E}_K \cap \mathcal{E}_{\text{int}} \right\} \right) \quad (17)$$

We then define the following discrete $W_0^{1,p}$ norms on $H_{\mathcal{D}}$, depending on the mappings Π_{σ} :

$$\|u\|_{p,\Pi}^p = \sum_{K \in \mathcal{M}} \sum_{\sigma \in \mathcal{E}_K} |\sigma| d_{K,\sigma} \left(\frac{|\Pi_{\sigma} u - u_K|}{d_{K,\sigma}} \right)^p \quad (18)$$

Let us remark that the following inequality holds:

$$\|u\|_{1,p,\mathcal{D}}^p \leq (1 + \theta_{\mathcal{D}}) \|u\|_{p,\Pi}^p, \quad \forall u \in H_{\mathcal{D}}, \quad (19)$$

using the following inequalities

$$\begin{aligned} \left(\frac{|u_K - u_L|}{d_{\sigma}} \right)^p &\leq \left(\left(\frac{d_{K,\sigma}}{d_{\sigma}} \right)^{p'} + \left(\frac{d_{L,\sigma}}{d_{\sigma}} \right)^{p'} \right)^{p/p'} \left(\frac{|\Pi_{\sigma} u - u_K|^p}{d_{K,\sigma}^p} + \frac{|\Pi_{\sigma} u - u_L|^p}{d_{L,\sigma}^p} \right), \\ \left(\frac{d_{K,\sigma}}{d_{\sigma}} \right)^{p'} + \left(\frac{d_{L,\sigma}}{d_{\sigma}} \right)^{p'} &\leq \frac{d_{K,\sigma}}{d_{\sigma}} + \frac{d_{L,\sigma}}{d_{\sigma}} = 1, \end{aligned}$$

and $d_{L,\sigma} \leq \theta_{\mathcal{D}} d_{K,\sigma}$. We also have

$$\|\nabla_{\mathcal{D}} u\|_{L^p(\Omega)}^p \leq d^{p-1} \|u\|_{p,\Pi}^p, \quad \forall u \in H_{\mathcal{D}}, \quad (20)$$

using the inequality

$$\left| \sum_{\sigma \in \mathcal{E}_K} \frac{|\sigma| d_{K,\sigma}}{|K|} \frac{\Pi_{\sigma} u - u_K}{d_{K,\sigma}} \mathbf{n}_{K,\sigma} \right|^p \leq \left(\sum_{\sigma \in \mathcal{E}_K} \frac{|\sigma| d_{K,\sigma}}{|K|} 1^{p'} \right)^{p/p'} \left(\sum_{\sigma \in \mathcal{E}_K} \frac{|\sigma| d_{K,\sigma}}{|K|} \frac{|\Pi_{\sigma} u - u_K|^p}{|d_{K,\sigma}|^p} \right).$$

Lemma 3.3 (Coerciveness) *Under hypotheses (2), let \mathcal{D} be a discretisation in the sense of Definition 2.1. Let $\theta \geq \theta_{\mathcal{D}}$ (defined by (16)). Then there exists $C_2 > 0$, only depending on d , Ω , p , \underline{a} and θ such that*

$$\forall u \in H_{\mathcal{D}}, C_2 \|u\|_{p,\Pi}^p \leq \langle u, u \rangle_{\mathcal{D}}, \quad (21)$$

where $\langle u, u \rangle_{\mathcal{D}}$ is defined in (11).

PROOF. Thanks to (2c), we have

$$\int_{\Omega} \mathbf{a}(\mathbf{x}, \nabla_{\mathcal{D}} u(\mathbf{x})) \cdot \nabla_{\mathcal{D}} u(\mathbf{x}) \, d\mathbf{x} \geq \underline{a} \sum_{K \in \mathcal{M}} |K| |\nabla_K u|^p.$$

We get, from (12),

$$b(u, u) = \sum_{K \in \mathcal{M}} \sum_{\sigma \in \mathcal{E}_K} \frac{|\sigma| d_{K,\sigma}}{d} |R_{K,\sigma} u|^p.$$

Using Hölder's inequality, we have

$$\forall x, y \in \mathbb{R}, \forall \mu > 0, |x + y| \leq |x| + |y| \leq (1 + \mu^{p'})^{1/p'} (|x|^p + \frac{|y|^p}{\mu^p})^{1/p}.$$

We apply the above inequality to $x = R_{K,\sigma} u$ and $y = \frac{1}{d_{K,\sigma}} \nabla_K u \cdot (\mathbf{x}_{\sigma} - \mathbf{x}_K)$, for a value μ which will be chosen later. We get

$$\left(\frac{|\Pi_{\sigma} u - u_K|}{d_{K,\sigma}} \right)^p \leq (1 + \mu^{p'})^{p/p'} \left(|R_{K,\sigma} u|^p + \left(\frac{|\nabla_K u \cdot (\mathbf{x}_{\sigma} - \mathbf{x}_K)|}{\mu d_{K,\sigma}} \right)^p \right).$$

We then choose $\mu = \frac{\theta}{\underline{a}^{1/p}}$, which ensures

$$\left(\frac{|\mathbf{x}_\sigma - \mathbf{x}_K|}{\mu d_{K,\sigma}} \right)^p \leq \underline{a}, \forall K \in \mathcal{M}, \forall \sigma \in \mathcal{E}_K.$$

We then get, for this value of μ ,

$$(1 + \mu^{p'})^{-p/p'} \sum_{K \in \mathcal{M}} \sum_{\sigma \in \mathcal{E}_K} \frac{|\sigma| d_{K,\sigma}}{d} \left(\frac{|\Pi_\sigma u - u_K|}{d_{K,\sigma}} \right)^p \leq \langle u, u \rangle_{\mathcal{D}},$$

which concludes the proof of (21). \square

Lemma 3.4 (Estimate, existence and uniqueness of the solution) *Under hypotheses (2), let \mathcal{D} be a discretisation in the sense of Definition 2.1. Let $\theta \geq \theta_{\mathcal{D}}$ (defined by (16)). Let u be a solution of (11).*

Then there exists $C_3 > 0$, only depending on $d, \Omega, p, \underline{a}$ and θ such that

$$\|u\|_{p,\Pi}^p \leq C_3 \|f\|_{L^{p'}(\Omega)}^{p'}. \quad (22)$$

Moreover, there exists one and only one solution $u \in H_{\mathcal{D}}$ to (11).

PROOF. Let $v = u$ in (11); applying (21) in Lemma 3.3, we get

$$C_2 \|u\|_{p,\Pi}^p \leq \int f(\mathbf{x}) u(\mathbf{x}) \, d\mathbf{x}.$$

We now apply Hölder's inequality: we get

$$\int f(\mathbf{x}) u(\mathbf{x}) \, d\mathbf{x} \leq \|f\|_{L^{p'}(\Omega)} \|u\|_{L^p(\Omega)}.$$

We then use (3.1) in Lemma 3.1 and (19) to obtain that there exists C_4 , only depending on θ, p and Ω , such that

$$\|u\|_{L^p(\Omega)} \leq C_4 \|u\|_{p,\Pi}.$$

The last three inequalities yield (22).

The existence of the discrete solution follows, for instance by a topological degree argument.

Let us now assume that u and \hat{u} are two solutions of (11). Then, $\langle u, u - \hat{u} \rangle_{\mathcal{D}} + \langle \hat{u}, \hat{u} - u \rangle_{\mathcal{D}} = 0$, which yields

$$\begin{aligned} & \sum_{K \in \mathcal{M}} |K| (\mathbf{a}_K(\nabla_K u) - \mathbf{a}_K(\nabla_K \hat{u})) \cdot \nabla_K (u - \hat{u}) \\ & + \sum_{K \in \mathcal{M}} \sum_{\sigma \in \mathcal{E}_K} \alpha_{K,\sigma} |\sigma| d_{K,\sigma} (\text{sgn}(R_{K,\sigma} u) |R_{K,\sigma} u|^{p-1} - \text{sgn}(R_{K,\sigma} \hat{u}) |R_{K,\sigma} \hat{u}|^{p-1}) R_{K,\sigma} (u - \hat{u}) = 0. \end{aligned}$$

Since

$$(\text{sgn}(R_{K,\sigma} u) |R_{K,\sigma} u|^{p-1} - \text{sgn}(R_{K,\sigma} \hat{u}) |R_{K,\sigma} \hat{u}|^{p-1}) (R_{K,\sigma} u - R_{K,\sigma} \hat{u}) \geq 0,$$

and

$$(\mathbf{a}_K(\nabla_K u) - \mathbf{a}_K(\nabla_K \hat{u})) \cdot \nabla_K (u - \hat{u}) \geq 0,$$

we get that both terms vanish. Thus, $\nabla_K (u - \hat{u}) = 0$ and $R_{K,\sigma} (u - \hat{u}) = 0$. We then get from the inequality (21) that $u - \hat{u} = 0$.

\square

4 Convergence of the scheme

The convergence of the scheme requires a compactness property which extends that of [19] proven in the case where $p = 2$:

Lemma 4.1 (Weak discrete $W^{1,p}$ compactness) *Under hypotheses (2), let \mathcal{F} be a family of discretisations in the sense of Definition 2.1. For any $\mathcal{D} \in \mathcal{F}$, let $(\Pi_\sigma)_{\sigma \in \mathcal{E}}$ be a family of linear mappings from $H_{\mathcal{D}}$ to \mathbb{R} defined by (6). Assume that there exists $\theta > 0$ such that for all $\mathcal{D} \in \mathcal{F}$, $\theta_{\mathcal{D}} \leq \theta$, with $\theta_{\mathcal{D}}$ defined by (16). Let $(u_{\mathcal{D}})_{\mathcal{D} \in \mathcal{F}}$ be a family of piecewise constant functions such that:*

- $u_{\mathcal{D}} \in H_{\mathcal{D}}$ for all $\mathcal{D} \in \mathcal{F}$,
- there exists $C > 0$ with $\|u_{\mathcal{D}}\|_{p,\Pi} \leq C$ for all $\mathcal{D} \in \mathcal{F}$,
- there exists $u \in L^p(\Omega)$ with $\lim_{h_{\mathcal{D}} \rightarrow 0} \|u_{\mathcal{D}} - u\|_{L^p(\Omega)} = 0$.

Then, the limit u belongs to $W_0^{1,p}(\Omega)$; moreover, defining $\nabla_{\mathcal{D}} u_{\mathcal{D}}$ by (8), the sequence $(\nabla_{\mathcal{D}} u_{\mathcal{D}})_{\mathcal{D} \in \mathcal{F}}$ weakly converges in $L^p(\Omega)^d$ to ∇u as $h_{\mathcal{D}} \rightarrow 0$. Prolonging all functions by 0 outside of Ω , the convergence also holds in $L^p(\mathbb{R}^d)^d$.

PROOF. Thanks to (20), we get that, up to a sub-sequence, there exists some function $\mathbf{G} \in L^p(\mathbb{R}^d)^d$ such that $\nabla_{\mathcal{D}} u_{\mathcal{D}}$ weakly converges in $L^p(\mathbb{R}^d)^d$ to \mathbf{G} as $h_{\mathcal{D}} \rightarrow 0$. Let us show that $\mathbf{G} = \nabla u$. Let $\psi \in C_c^\infty(\mathbb{R}^d)^d$ be given. Let us consider the term $T_1^{\mathcal{D}}$ defined by

$$T_1^{\mathcal{D}} = \int_{\mathbb{R}^d} \nabla_{\mathcal{D}} u_{\mathcal{D}}(\mathbf{x}) \cdot \psi(\mathbf{x}) d\mathbf{x};$$

this term may also be written:

$$T_1^{\mathcal{D}} = \sum_{K \in \mathcal{M}} \sum_{\sigma \in \mathcal{E}_K} |\sigma| (\Pi_\sigma u - u_K) \mathbf{n}_{K,\sigma} \cdot \psi_K, \text{ with } \psi_K = \frac{1}{|K|} \int_K \psi(\mathbf{x}) d\mathbf{x}.$$

Let us then compare $T_1^{\mathcal{D}}$ with

$$T_2^{\mathcal{D}} = \sum_{K \in \mathcal{M}} \sum_{\sigma \in \mathcal{E}_K} |\sigma| (\Pi_\sigma u - u_K) \mathbf{n}_{K,\sigma} \cdot \psi_\sigma, \text{ where } \psi_\sigma = \frac{1}{|\sigma|} \int_\sigma \psi(\mathbf{x}) d\gamma(\mathbf{x}).$$

We get that

$$|T_1^{\mathcal{D}} - T_2^{\mathcal{D}}| \leq \left(\sum_{K \in \mathcal{M}} \sum_{\sigma \in \mathcal{E}_K} |\sigma| d_{K,\sigma} \frac{|\Pi_\sigma u - u_K|^p}{d_{K,\sigma}^p} \right)^{1/p} \left(\sum_{K \in \mathcal{M}} \sum_{\sigma \in \mathcal{E}_K} |\sigma| d_{K,\sigma} |\psi_K - \psi_\sigma|^{p'} \right)^{1/p'},$$

which leads to $\lim_{h_{\mathcal{D}} \rightarrow 0} (T_1^{\mathcal{D}} - T_2^{\mathcal{D}}) = 0$.

Since

$$T_2^{\mathcal{D}} = - \sum_{K \in \mathcal{M}} \sum_{\sigma \in \mathcal{E}_K} |\sigma| u_K \mathbf{n}_{K,\sigma} \cdot \psi_\sigma = - \int_{\mathbb{R}^d} \Pi_{\mathcal{M}} u_{\mathcal{D}}(\mathbf{x}) \operatorname{div} \psi(\mathbf{x}) d\mathbf{x},$$

we deduce that $\lim_{h_{\mathcal{D}} \rightarrow 0} T_4^{\mathcal{D}} = - \int_{\mathbb{R}^d} u(\mathbf{x}) \operatorname{div} \psi(\mathbf{x}) d\mathbf{x}$. This proves that the function $\mathbf{G} \in L^p(\mathbb{R}^d)^d$ is a.e. equal to ∇u in \mathbb{R}^d . Since $u = 0$ outside of Ω , we get that $u \in W_0^{1,p}(\Omega)$, and the uniqueness of the limit implies that the whole family $\nabla_{\mathcal{D}} u_{\mathcal{D}}$ weakly converges in $L^p(\mathbb{R}^d)^d$ to ∇u as $h_{\mathcal{D}} \rightarrow 0$.

□

For any $\varphi \in C(\bar{\Omega}, \mathbb{R})$, we denote by $P_{\mathcal{D}} \varphi$ the element of $H_{\mathcal{D}}$ defined by:

$$P_{\mathcal{D}} \varphi = (\varphi(\mathbf{x}_K))_{K \in \mathcal{M}}, \tag{23}$$

and we prove a consistency result of the discrete gradient, which was already used in [19] in the more general setting of the SUSHI scheme.

Lemma 4.2 (Discrete gradient consistency) *Let \mathcal{D} be a discretisation of Ω in the sense of Definition 2.1. and let $(\Pi_\sigma)_{\sigma \in \mathcal{E}}$ be a family of linear mappings from $H_{\mathcal{D}}$ to \mathbb{R} defined by (6). Assume that $\theta \geq \theta_{\mathcal{D}, \Pi}$ (given by (17)). Then, for any function $\varphi \in C^2(\overline{\Omega})$, there exists C_5 only depending on d , θ and φ such that:*

$$\|\nabla_{\mathcal{D}} P_{\mathcal{D}} \varphi - \nabla \varphi\|_{(L^\infty(\Omega))^d} \leq C_5 h_{\mathcal{D}}, \quad (24)$$

where $\nabla_{\mathcal{D}}$ is defined by (8), and

$$\|R_{\mathcal{D}} P_{\mathcal{D}} \varphi\|_{L^\infty(\Omega)} \leq C_5 h_{\mathcal{D}}, \quad (25)$$

where $R_{\mathcal{D}}$ is defined by (9)-(10) and $P_{\mathcal{D}} \varphi$ is defined by (23).

PROOF. From (8b), we have, for any $K \in \mathcal{M}$,

$$\nabla_K P_{\mathcal{D}} \varphi = \frac{1}{|K|} \sum_{\sigma \in \mathcal{E}_K} |\sigma| (\Pi_\sigma P_{\mathcal{D}} \varphi - \varphi(\mathbf{x}_K)) \mathbf{n}_{K, \sigma}.$$

Using definition (17), we can write, for $\sigma \in \mathcal{E}_K$,

$$|\Pi_\sigma P_{\mathcal{D}} \varphi - \varphi(\mathbf{x}_\sigma)| \leq C_\varphi \theta h_K^2,$$

and on the other hand, we have

$$\frac{1}{|K|} \sum_{\sigma \in \mathcal{E}_K} |\sigma| (\varphi(\mathbf{x}_\sigma) - \varphi(\mathbf{x}_K)) \mathbf{n}_{K, \sigma} = \frac{1}{|K|} \sum_{\sigma \in \mathcal{E}_K} |\sigma| (\nabla \varphi(\mathbf{x}_K) \cdot (\mathbf{x}_\sigma - \mathbf{x}_K) + h_K^2 \rho_{K, \sigma}) \mathbf{n}_{K, \sigma},$$

where $|\rho_{K, \sigma}| \leq C_\varphi$ with C_φ only depending on φ . Thanks to (7) and to the regularity of the mesh, we get the existence of C_6 , only depending on θ , such that

$$|\nabla_K P_{\mathcal{D}} \varphi - \nabla \varphi(\mathbf{x}_K)| \leq h_K C_6 C_\varphi.$$

From this last inequality, using Definition 9, we get the existence of C_7 , only depending on θ , such that

$$\begin{aligned} |R_{K, \sigma} P_{\mathcal{D}} \varphi| &= \frac{1}{d_{K, \sigma}} |\Pi_\sigma P_{\mathcal{D}} \varphi - \varphi(\mathbf{x}_K) - \nabla_K P_{\mathcal{D}} \varphi \cdot (\mathbf{x}_\sigma - \mathbf{x}_K)| \\ &\leq h_K C_\varphi C_7 \end{aligned}$$

which concludes the proof. \square

Theorem 4.1 (Convergence of the scheme) *Under hypotheses (2), let \mathcal{F} be a family of discretisations in the sense of Definition 2.1. For any $\mathcal{D} \in \mathcal{F}$, let $(\Pi_\sigma)_{\sigma \in \mathcal{E}}$ be a family defined by (6). We assume that, there exists $\theta > 0$ with, for all $\mathcal{D} \in \mathcal{F}$, $\theta \geq \theta_{\mathcal{D}, \Pi}$ (see (17)).*

For all $\mathcal{D} \in \mathcal{F}$, let $u_{\mathcal{D}} \in H_{\mathcal{D}}$ be the unique solution of (11). Then, as $h_{\mathcal{D}} \rightarrow 0$:

$$u_{\mathcal{D}} \longrightarrow u \in W_0^{1,p}(\Omega) \text{ for the strong topology of } L^p(\Omega), \quad (26a)$$

where $u \in W_0^{1,p}(\Omega)$ is the unique solution to (1); furthermore,

$$\nabla_{\mathcal{D}} u_{\mathcal{D}} \longrightarrow \nabla u \text{ for the strong topology of } L^p(\Omega)^d \text{ and} \quad (26b)$$

$$\mathbf{a}(\cdot, \nabla_{\mathcal{D}} u_{\mathcal{D}}) \longrightarrow \mathbf{a}(\cdot, \nabla u) \text{ for the strong topology of } (L^{p'}(\Omega))^d. \quad (26c)$$

PROOF.

Step 1. Convergence of the approximate solution (proof of (26a)).

Thanks to (22) in Lemma 3.4, the family $(\|u_{\mathcal{D}}\|_{p,\Pi})_{\mathcal{D} \in \mathcal{F}}$ is bounded independently of \mathcal{D} . Therefore, thanks to inequality (19), we may apply Lemma 3.2. Thus, for any sequence $(\mathcal{D}_n)_{n \in \mathbb{N}}$ of discretisations in the family \mathcal{F} such that $h_{\mathcal{D}_n}$ tends to 0 as $n \rightarrow \infty$, there exist $u \in L^p(\Omega)$ such that, up to a sub-sequence, $u_{\mathcal{D}_n} \rightarrow u$ in $L^p(\Omega)$ as $n \rightarrow \infty$. For short, we replace the index \mathcal{D}_n by n in the remainder of Step 1.

We then apply Lemma 4.1 and deduce that $u \in W_0^{1,p}(\Omega)$ and that $\nabla_n u_n \rightarrow \nabla u$ weakly in $(L^p(\Omega))^d$ as $n \rightarrow \infty$. Using Hypothesis (2e), we then get that the function $\mathbf{x} \rightarrow \mathbf{a}(\mathbf{x}, \nabla_n u_n(\mathbf{x}))$ is bounded in $(L^{p'}(\Omega))^d$; hence, there exists a sub-sequence of $(\mathcal{D}_n)_{n \in \mathbb{N}}$, again denoted $(\mathcal{D}_n)_{n \in \mathbb{N}}$, and $\mathbf{A} \in (L^{p'}(\Omega))^d$ such that $\mathbf{a}(\cdot, \nabla_n u_n) \rightarrow \mathbf{A}$ weakly in $(L^{p'}(\Omega))^d$.

Let us now prove that $\operatorname{div}(\mathbf{A} - \mathbf{a}(\nabla u)) = 0$, and that u is the weak solution of (1).

Let $\psi \in C_c^\infty(\Omega)$. We introduce $P_n \psi$ as a test function in (11) (with $P_n = P_{\mathcal{D}_n}$ defined by (23)). We get

$$\int_{\Omega} \mathbf{a}(\mathbf{x}, \nabla_n u_n(\mathbf{x})) \cdot \nabla_n P_n \psi(\mathbf{x}) \, d\mathbf{x} + b(u_n, P_n \psi) = \int_{\Omega} f(\mathbf{x}) P_n \psi(\mathbf{x}) \, d\mathbf{x}.$$

Thanks to the estimate (22) on u_n (in the $\|\cdot\|_{1,p,\mathcal{D}}$ norm) and thanks to (20), the sequence $(\nabla_n u_n)_{n \in \mathbb{N}}$ is bounded in $L^p(\Omega)^d$. Hence, the sequence $(R_n u)_{n \in \mathbb{N}}$ is bounded in $L^p(\Omega)$, which implies, using (25) in Lemma 4.2, that

$$\lim_{n \rightarrow \infty} b(u_n, P_n \psi) = 0. \quad (27)$$

Hence, using (24) in Lemma 4.2 and passing to the limit as $n \rightarrow +\infty$ in (11), we get that

$$\int_{\Omega} \mathbf{A}(\mathbf{x}) \cdot \nabla \psi(\mathbf{x}) \, d\mathbf{x} = \int_{\Omega} f(\mathbf{x}) \psi(\mathbf{x}) \, d\mathbf{x}. \quad (28)$$

By density, this also holds for all $\psi \in W_0^{1,p}(\Omega)$. It remains to prove that

$$\int_{\Omega} \mathbf{A}(\mathbf{x}) \cdot \nabla \psi(\mathbf{x}) \, d\mathbf{x} = \int_{\Omega} \mathbf{a}(\mathbf{x}, \nabla u(\mathbf{x})) \cdot \nabla \psi(\mathbf{x}) \, d\mathbf{x}.$$

This is the object of the famous ‘‘Minty trick’’ [27]. Indeed, for any $\psi \in C_c^\infty(\Omega)$, since \mathbf{a} satisfies (2d), the following inequality holds for any $n \in \mathbb{N}$:

$$\int_{\Omega} (\mathbf{a}(\mathbf{x}, \nabla_n u_n(\mathbf{x})) - \mathbf{a}(\mathbf{x}, \nabla_n P_n \psi(\mathbf{x}))) \cdot (\nabla_n u_n(\mathbf{x}) - \nabla_n P_n \psi(\mathbf{x})) \, d\mathbf{x} \geq 0, \quad (29)$$

On the other hand, thanks to the positivity of $b(u_n, u_n)$ and because u_n verifies (11), we get

$$\int_{\Omega} \mathbf{a}(\mathbf{x}, \nabla_n u_n(\mathbf{x})) \cdot \nabla_n u_n(\mathbf{x}) \, d\mathbf{x} \leq \int_{\Omega} \mathbf{a}(\mathbf{x}, \nabla_n u_n(\mathbf{x})) \cdot \nabla_n u_n(\mathbf{x}) \, d\mathbf{x} + b(u_n, u_n) = \int_{\Omega} f(\mathbf{x}) u_n(\mathbf{x}) \, d\mathbf{x}.$$

Passing to the limit $n \rightarrow +\infty$ (up to the considered sub-sequence), we thus get

$$\limsup_{n \rightarrow \infty} \int_{\Omega} \mathbf{a}(\mathbf{x}, \nabla_n u_n(\mathbf{x})) \cdot \nabla_n u_n(\mathbf{x}) \, d\mathbf{x} \leq \int_{\Omega} f(\mathbf{x}) u(\mathbf{x}) \, d\mathbf{x} = \int_{\Omega} \mathbf{A}(\mathbf{x}) \cdot \nabla u(\mathbf{x}) \, d\mathbf{x},$$

thanks to (28) replacing ψ by u .

Now, thanks to the continuity on \mathbf{a} with respect to its second argument (assumption (2b)) and to assumption (2e), we obtain by the Lebesgue dominated convergence theorem that

$$\mathbf{a}(\cdot, \nabla_n P_n \psi(\mathbf{x})) \rightarrow \mathbf{a}(\cdot, \nabla \psi) \text{ in } L^{p'}(\Omega) \text{ as } n \rightarrow +\infty. \quad (30)$$

Hence, passing to the limit as $n \rightarrow \infty$ in (29) leads to

$$\int_{\Omega} (\mathbf{A}(\mathbf{x}) - \mathbf{a}(\mathbf{x}, \nabla\psi(\mathbf{x}))) \cdot (\nabla u(\mathbf{x}) - \nabla\psi(\mathbf{x})) \, d\mathbf{x} \geq 0.$$

By density, this last inequality remains true for any $\psi \in W_0^{1,p}(\Omega)$. Taking $\psi = u + t\varphi$ with $\varphi \in C_c^\infty(\Omega)$ and $t > 0$, we get

$$\int_{\Omega} (\mathbf{A}(\mathbf{x}) - \mathbf{a}(\mathbf{x}, \nabla u(\mathbf{x}) + t\nabla\varphi(\mathbf{x}))) \cdot \nabla\varphi(\mathbf{x}) \, d\mathbf{x} \geq 0,$$

This gives, letting $t \rightarrow 0$,

$$\int_{\Omega} (\mathbf{A}(\mathbf{x}) - \mathbf{a}(\mathbf{x}, \nabla u(\mathbf{x}))) \cdot \nabla\varphi(\mathbf{x}) \, d\mathbf{x} \geq 0.$$

Changing φ in $-\varphi$, we get

$$- \int_{\Omega} (\mathbf{A}(\mathbf{x}) - \mathbf{a}(\mathbf{x}, \nabla u(\mathbf{x}))) \cdot \nabla\varphi(\mathbf{x}) \, d\mathbf{x} \geq 0.$$

This proves that $\int_{\Omega} (\mathbf{A}(\mathbf{x}) - \mathbf{a}(\mathbf{x}, \nabla u(\mathbf{x}))) \cdot \nabla\varphi(\mathbf{x}) \, d\mathbf{x} = 0$ and therefore, with (28), that u is the weak solution of (1). Then, using the uniqueness of the solution of (1), a classical argument gives that the whole sequence converges and therefore, this convergence holds for the whole family \mathcal{F} as $h_{\mathcal{D}} \rightarrow 0$ which proves (26a).

Step 2. Strong convergence of the gradient (proof of (26b)).

We now prove the (strong) convergence of $\nabla_{\mathcal{D}} u_{\mathcal{D}}$ to ∇u in $(L^p(\Omega))^d$ as $h_{\mathcal{D}} \rightarrow 0$. We already remarked (thanks to the positivity of $b(u_{\mathcal{D}}, u_{\mathcal{D}})$ and using that $u_{\mathcal{D}}$ verifies (11)) that:

$$\limsup_{h_{\mathcal{D}} \rightarrow 0} \int_{\Omega} \mathbf{a}(\mathbf{x}, \nabla_{\mathcal{D}} u_{\mathcal{D}}(\mathbf{x})) \cdot \nabla_{\mathcal{D}} u_{\mathcal{D}}(\mathbf{x}) \, d\mathbf{x} \leq \int_{\Omega} f(\mathbf{x}) u(\mathbf{x}) \, d\mathbf{x} = \int_{\Omega} \mathbf{a}(\mathbf{x}, \nabla u(\mathbf{x})) \cdot \nabla u(\mathbf{x}) \, d\mathbf{x}.$$

Hence we get

$$\limsup_{h_{\mathcal{D}} \rightarrow 0} \int_{\Omega} (\mathbf{a}(\mathbf{x}, \nabla_{\mathcal{D}} u_{\mathcal{D}}(\mathbf{x})) - \mathbf{a}(\mathbf{x}, \nabla u(\mathbf{x}))) \cdot (\nabla_{\mathcal{D}} u_{\mathcal{D}}(\mathbf{x}) - \nabla u(\mathbf{x})) \, d\mathbf{x} \leq 0.$$

Since $(\mathbf{a}(\mathbf{x}, \nabla_{\mathcal{D}} u_{\mathcal{D}}) - \mathbf{a}(\mathbf{x}, \nabla u)) \cdot (\nabla_{\mathcal{D}} u_{\mathcal{D}} - \nabla u) \geq 0$ for a.e. $\mathbf{x} \in \Omega$, we then have

$$(\mathbf{a}(\cdot, \nabla_{\mathcal{D}} u_{\mathcal{D}}) - \mathbf{a}(\cdot, \nabla u)) \cdot (\nabla_{\mathcal{D}} u_{\mathcal{D}} - \nabla u) \rightarrow 0 \text{ in } L^1(\Omega), \quad (31)$$

and therefore a.e. for a sub-sequence. Then, Lemma 4.3, (which states the ‘‘Leray-Lions trick’’ [27] and which we give below for the sake of completeness) shows that $\nabla_{\mathcal{D}} u_{\mathcal{D}} \rightarrow \nabla u$ a.e. as $h_{\mathcal{D}} \rightarrow 0$, at least for the same sub-sequence. Since the family $(\nabla_{\mathcal{D}} u_{\mathcal{D}})_{\mathcal{D} \in \mathcal{F}}$ is bounded in $L^p(\Omega)^d$, this a.e. convergence implies the convergence in $L^q(\Omega)^d$ for any $q \in [1, p)$. The convergence of the whole family $(\nabla_{\mathcal{D}} u_{\mathcal{D}})_{\mathcal{D} \in \mathcal{F}}$ to ∇u in $L^q(\Omega)^d$ classically follows, for any $q \in [1, p)$, as $h_{\mathcal{D}} \rightarrow 0$. The boundedness of the family $(\nabla_{\mathcal{D}} u_{\mathcal{D}})_{\mathcal{D} \in \mathcal{F}}$ in $L^p(\Omega)^d$ also entails that:

$$\nabla_{\mathcal{D}} u_{\mathcal{D}} \text{ converges to } \nabla u \text{ weakly in } L^p(\Omega)^d \text{ as } h_{\mathcal{D}} \rightarrow 0.$$

In order to obtain the strong convergence of $\nabla_{\mathcal{D}} u_{\mathcal{D}}$ in $L^p(\Omega)^d$ (and not only in $L^q(\Omega)^d$ for $q < p$), we then remark that (31) gives:

$$\lim_{h_{\mathcal{D}} \rightarrow 0} \int_{\Omega} \mathbf{a}(\mathbf{x}, \nabla_{\mathcal{D}} u_{\mathcal{D}}(\mathbf{x})) \cdot \nabla_{\mathcal{D}} u_{\mathcal{D}}(\mathbf{x}) \, d\mathbf{x} = \int_{\Omega} \mathbf{a}(\mathbf{x}, \nabla u(\mathbf{x})) \cdot \nabla u(\mathbf{x}) \, d\mathbf{x}. \quad (32)$$

Next, we notice that, for any sequence of discretisations, we can assume, up to a sub-sequence, the a.e. convergence of $\mathbf{a}(\cdot, \nabla_{\mathcal{D}} u_{\mathcal{D}}) \cdot \nabla_{\mathcal{D}} u_{\mathcal{D}}$ to $\mathbf{a}(\cdot, \nabla u) \cdot \nabla u$; thus, since $\mathbf{a}(\cdot, \nabla_{\mathcal{D}} u_{\mathcal{D}}) \cdot \nabla_{\mathcal{D}} u_{\mathcal{D}} \geq 0$ a.e., we also have

by Lemma 4.4 (which is again classical [27] and which we give below for the sake of completeness) $\mathbf{a}(\cdot, \nabla_{\mathcal{D}} u_{\mathcal{D}}) \cdot \nabla_{\mathcal{D}} u_{\mathcal{D}} \rightarrow \mathbf{a}(\cdot, \nabla u) \cdot \nabla u$ in $L^1(\Omega)$ as $h_{\mathcal{D}} \rightarrow 0$ (see Lemma 4.4). This L^1 -convergence gives the equi-integrability of the family of functions $\mathbf{a}(\cdot, \nabla_{\mathcal{D}} u_{\mathcal{D}}) \cdot \nabla_{\mathcal{D}} u_{\mathcal{D}}$, which, in turn, gives, thanks to (2c), that the family of functions $|\nabla_{\mathcal{D}} u_{\mathcal{D}}|^p$ is equi-integrable. Finally, we obtain (using Vitali's theorem) the $L^p(\Omega)^d$ convergence of $\nabla_{\mathcal{D}} u_{\mathcal{D}}$ to ∇u , as $h_{\mathcal{D}} \rightarrow 0$.

Note that in the case $\mathbf{a}(\mathbf{x}, \xi) = |\xi|^{p-2}\xi$, a simple proof is possible since (32) gives the convergence of the L^p -norm of the approximate gradient to the L^p -norm of ∇u (which is sufficient, thanks to the a.e. convergence of "sub-sequences", to obtain the $L^p(\Omega)^d$ convergence).

Step 3. Convergence of fluxes (proof of (26c)).

Since $\nabla_{\mathcal{D}} u_{\mathcal{D}}$ converges to ∇u in $L^p(\Omega)^d$ as $h_{\mathcal{D}} \rightarrow 0$, the convergence of $\mathbf{a}(\cdot, \nabla_{\mathcal{D}} u_{\mathcal{D}})$ to $\mathbf{a}(\cdot, \nabla u)$ in $L^{p'}(\Omega)^d$ (namely assertion (26c)) follows classically from hypotheses (2b) and (2e) on \mathbf{a} .

□

Lemma 4.3 (The "Leray-Lions trick") *Let \mathbf{b} be a continuous function from \mathbb{R}^d to \mathbb{R}^d such that $(\mathbf{b}(\delta) - \mathbf{b}(\gamma)) \cdot (\delta - \gamma) > 0$ if $\delta, \gamma \in \mathbb{R}^d$, $\delta \neq \gamma$. Let $(\beta_n)_{n \in \mathbb{N}}$ be a sequence in \mathbb{R}^d and $\beta \in \mathbb{R}^d$ such that $(\mathbf{b}(\beta_n) - \mathbf{b}(\beta)) \cdot (\beta_n - \beta) \rightarrow 0$ as $n \rightarrow \infty$. Then, $\beta_n \rightarrow \beta$ as $n \rightarrow \infty$.*

PROOF. We begin the proof with a preliminary remark. Let $\delta \in \mathbb{R}^d$, $\delta \neq 0$. We define the function h_{δ} from \mathbb{R} to \mathbb{R} by $h_{\delta}(s) = (\mathbf{b}(\beta + s\delta) - \mathbf{b}(\beta)) \cdot \delta$. The hypothesis on \mathbf{b} gives that h_{δ} is an increasing function since, for $s > s'$, one has :

$$h_{\delta}(s) - h_{\delta}(s') = (\mathbf{b}(\beta + s\delta) - \mathbf{b}(\beta + s'\delta)) \cdot \delta > 0.$$

We prove now, by contradiction, that $\lim_{n \rightarrow \infty} \beta_n = \beta$. If the sequence $(\beta_n)_{n \in \mathbb{N}}$ does not converge to β , there exists $\varepsilon > 0$ and a subsequence, still denoted $(\beta_n)_{n \in \mathbb{N}}$, such that $|\beta_n - \beta| \geq \varepsilon$, for all $n \in \mathbb{N}$. Then, we set $\delta_n = \frac{\beta_n - \beta}{|\beta_n - \beta|}$ and we can assume, up to a subsequence, that $\delta_n \rightarrow \delta$ as $n \rightarrow \infty$, for some $\delta \in \mathbb{R}^d$ with $|\delta| = 1$. Taking $s_n = |\beta_n - \beta|$, we then have, since $s_n \geq \varepsilon$:

$$(\mathbf{b}(\beta_n) - \mathbf{b}(\beta)) \cdot \frac{\beta_n - \beta}{s_n} = h_{\delta_n}(s_n) \geq h_{\delta_n}(\varepsilon) = (\mathbf{b}(\beta + \varepsilon\delta_n) - \mathbf{b}(\beta)) \cdot \delta_n.$$

Then, passing to the limit as $n \rightarrow \infty$,

$$0 = \lim_{n \rightarrow \infty} \frac{1}{s_n} (\mathbf{b}(\beta_n) - \mathbf{b}(\beta)) \cdot (\beta_n - \beta) \geq (\mathbf{b}(\beta + \varepsilon\delta) - \mathbf{b}(\beta)) \cdot \delta > 0.$$

which is impossible.

□

Lemma 4.4 *Let $(F_n)_{n \in \mathbb{N}}$ be a sequence non-negative functions in $L^1(\Omega)$. Let $F \in L^1(\Omega)$ be such that $F_n \rightarrow F$ a.e. in Ω and $\int_{\Omega} F_n(\mathbf{x}) d\mathbf{x} \rightarrow \int_{\Omega} F(\mathbf{x}) d\mathbf{x}$, as $n \rightarrow \infty$. Then, $F_n \rightarrow F$ in $L^1(\Omega)$ as $n \rightarrow \infty$.*

PROOF. The proof of this lemma is very classical. Applying the Dominated Convergence Theorem to the sequence $(F - F_n)^+$ leads to $\int_{\Omega} (F(\mathbf{x}) - F_n(\mathbf{x}))^+ d\mathbf{x} \rightarrow 0$ as $n \rightarrow \infty$. Then, since $|F - F_n| = 2(F - F_n)^+ - (F - F_n)$, we conclude that $F_n \rightarrow F$ in $L^1(\Omega)$ as $n \rightarrow \infty$.

□

5 Numerical results

We consider the particular case $\Omega =]0, 1]^2$, and f given by the following: $f(\mathbf{x}) = 1$ for a.e. \mathbf{x} in a sub-domain of Ω and 0 elsewhere, as shown in the right part of Figure 5, and \mathbf{a} is given by (3) for $p \in]1, +\infty[$.

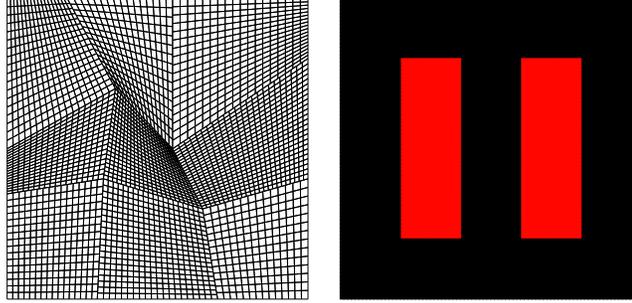


Figure 1: The mesh used (left), the function f ($f = 1$ in the red part)

We then compute, for $p = 1.3, 1.6, 2, 3, 6$, an approximate solution of the solution of (11), using an under-relaxed fixed point method (consisting in computing $|\nabla_{\mathcal{D}} u|^{p-2}$ one iteration late). We then get the following results, on two meshes. On one hand, we use the mesh shown in the left part of Figure 5, on the other hand, we use a regular 80×80 square mesh. The results are provided in Figures 2, 3, 4, 5 and 6. We obtain the approximate solutions with a precision of 10^{-8} with a quite small number (less than 60) thanks to the under-relaxed fixed point method (we observed that without relaxation the method does not converge for $p \geq 3$).

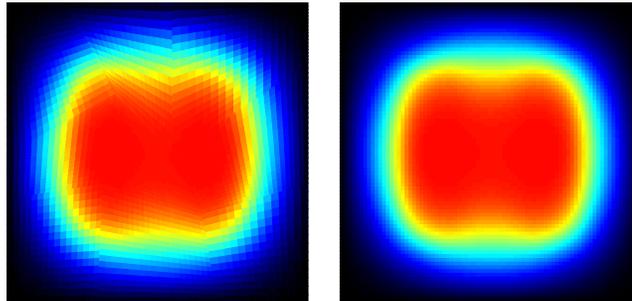


Figure 2: Results for $p = 1.3$ on the irregular mesh (left) and on the regular mesh (right)

6 Conclusion and generalisations

We showed here the convergence of a cell centred scheme for the discretisation of nonlinear elliptic equations of the Leray-Lions type and showed its numerical efficiency on an example. The convergence analysis is performed by mimicking the functional analysis tools of [27] (such as the Minty-Browder trick and the Leray-Lions trick) in the discrete setting. It may be generalised to the case of pseudo-monotone operators, that is \mathbf{a} of the form $u \mapsto \mathbf{a}(\cdot, u, \nabla u)$ with adequate assumptions on \mathbf{a} (see [27]). The convergence result can also be extended, using the discrete functional tools which are introduced here and the results of [15], to handle right hand sides $f \in W^{-1,p'}$. Both of these extensions are presented in [13] in the framework of the mixed finite volume scheme.

Let us also remark that the SUCCES scheme which we applied here to the discretisation of the Leray-Lions operator is closer, in its formulation, to a low order non conforming finite element scheme than

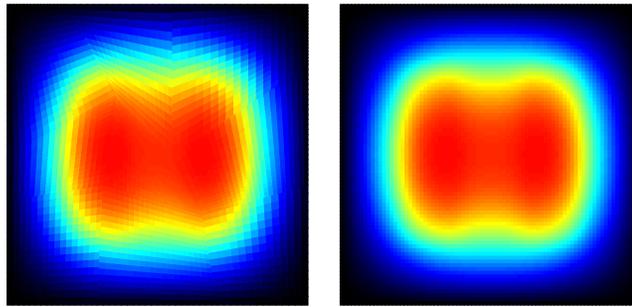


Figure 3: Results for $p = 1.6$ on the irregular mesh (left) and on the regular mesh (right)

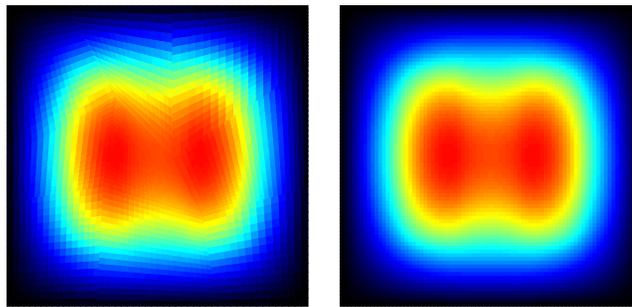


Figure 4: Results for $p = 2$ on the irregular mesh (left) and on the regular mesh (right)

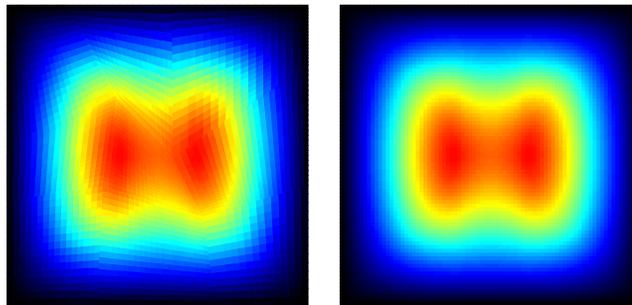


Figure 5: Results for $p = 3$ on the irregular mesh (left) and on the regular mesh (right)

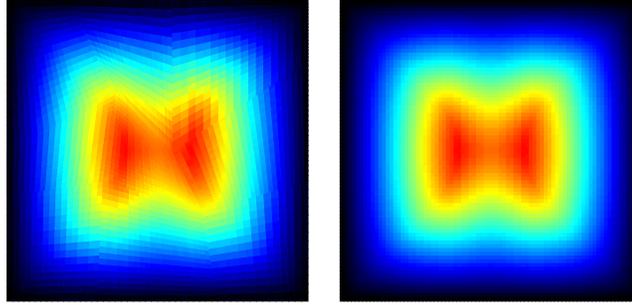


Figure 6: Results for $p = 6$ on the irregular mesh (left) and on the regular mesh (right)

to finite volume scheme: indeed, even though global conservativity of the fluxes is ensured, the scheme may not in general be written as a system of discrete balance laws over the discretisation cells, such as for the “classical” finite volume schemes, such as the two point flux scheme on admissible meshes, [17, chapter 3], the previously mentioned DDFV scheme [12] and mixed finite volume scheme [14], and the SUSHI scheme in its hybrid form [18, 19]. A scheme adapted from SUCCES, localising the fluxes on half edges as in the O scheme [1, 2] was recently introduced [4]: this scheme is a classical finite volume scheme in the sense that it may be written as a system of discrete mass balances and that local numerical fluxes are conservative. Nevertheless, it preserves the main properties of the SUCCES scheme (cell centred unknowns, consistent gradient, compactness properties) and therefore, the convergence theory which was presented here should also extend to this latter scheme (which, at the present time, is only written in the two dimensional setting).

Finally, we again stress that the tools used here essentially mimic the functional analysis tools of [27]. Since the obtained results include the strong convergence of the approximate solution, gradient, and fluxes, they may also be used for nonlinear evolution problems such as the level set equation [20].

References

- [1] I. Aavatsmark, T. Barkve, O. Boe, and T. Mannseth. Discretization on unstructured grids for inhomogeneous, anisotropic media. part i: Derivation of the methods. *SIAM Journal on Sc. Comp.*, 19:1700–1716, 1998.
- [2] I. Aavatsmark, T. Barkve, O. Boe, and T. Mannseth. Discretization on unstructured grids for inhomogeneous, anisotropic media. part ii: Discussion and numerical results. *SIAM Journal on Sc. Comp.*, 19:1717–1736, 1998.
- [3] L. Agelas, D.A. Di Pietro, and R. Masson. A symmetric and coercive finite volume scheme for multiphase porous media flow problems with applications in the oil industry. In R. Eymard and J.-M. Hérard, editors, *Finite Volumes for Complex Applications V*, pages 35–51. Wiley, 2008.
- [4] L. Agelas, R. Eymard, and R. Herbin. A nine-point finite volume scheme for the simulation of diffusion in heterogeneous media. submitted, see also <http://hal.archives-ouvertes.fr/hal-00350139/fr/>.
- [5] B. Andreianov, F. Boyer, and F. Hubert. Finite volume schemes for the p -Laplacian on Cartesian meshes. *M2AN Math. Model. Numer. Anal.*, 38(6):931–959, 2004.

- [6] B. Andreianov, F. Boyer, and F. Hubert. Besov regularity and new error estimates for finite volume approximations of the p -Laplacian. *Numer. Math.*, 100(4):565–592, 2005.
- [7] B. Andreianov, F. Boyer, and F. Hubert. On the finite-volume approximation of regular solutions of the p -Laplacian. *IMA J. Numer. Anal.*, 26(3):472–502, 2006.
- [8] B. Andreianov, F. Boyer, and F. Hubert. Discrete duality finite volume schemes for Leray-Lions-type elliptic problems on general 2D meshes. *Numer. Methods Partial Differential Equations*, 23(1):145–195, 2007.
- [9] J. W. Barrett and W. B. Liu. Finite element approximation of the p -Laplacian. *Math. Comp.*, 61(204):523–537, 1993.
- [10] F. Boyer and F. Hubert. Finite volume method for 2d linear and nonlinear elliptic problems with discontinuities. *SIAM J. Numer. Anal.*, 46(6):3032–3070, 2008.
- [11] S.-S. Chow. Finite element error estimates for nonlinear elliptic equations of monotone type. *Numer. Math.*, 54(4):373–393, 1989.
- [12] K. Domelevo and P. Omnes. A finite volume method for the laplace equation on almost arbitrary two-dimensional grids. *M2AN Math. Model. Numer. Anal.*, 39(6):1203–1249, 2005.
- [13] J. Droniou. Finite volume schemes for fully non-linear elliptic equations in divergence form. *M2AN Math. Model. Numer. Anal.*, 40(6):1069–1100, 2006.
- [14] J. Droniou and R. Eymard. A mixed finite volume scheme for anisotropic diffusion problems on any grid. *Numer. Math.*, 105(1):35–71, 2006.
- [15] J. Droniou and T. Gallouët. Finite volume methods for convection-diffusion equations with right-hand side in H^{-1} . *M2AN Math. Model. Numer. Anal.*, 36(4):705–724, 2002.
- [16] C. Ebmeyer and W. B. Liu. Finite element approximation of the fast diffusion and the porous medium equations. *SIAM J. Numer. Anal.*, 46(5):2393–2410, 2008.
- [17] R. Eymard, T. Gallouët, and R. Herbin. Finite volume methods. In P. G. Ciarlet and J.-L. Lions, editors, *Techniques of Scientific Computing, Part III*, Handbook of Numerical Analysis, VII, pages 713–1020. North-Holland, Amsterdam, 2000.
- [18] R. Eymard, T. Gallouët, and R. Herbin. A new finite volume scheme for anisotropic diffusion problems on general grids: convergence analysis. *C. R., Math., Acad. Sci. Paris*, 344(6):403–406, 2007.
- [19] R. Eymard, T. Gallouët, and R. Herbin. Discretisation of heterogeneous and anisotropic diffusion problems on general non-conforming meshes, sushi: a scheme using stabilisation and hybrid interfaces. to appear in IMAJNA, see also <http://hal.archives-ouvertes.fr/docs/00/21/18/28/PDF/suchi.pdf> , 2008.
- [20] R. Eymard, A. Handlovičová, and K. Mikula. Convergence of a finite volume method to a solution of the level set equation. work in progress, 2009.
- [21] R. Eymard and R. Herbin. A new collocated finite volume scheme for the incompressible Navier-Stokes equations on general non matching grids. *C. R. Math. Acad. Sci. Paris*, 344(10):659–662, 2007.
- [22] M. Feistauer and V. Sobotíková. Finite element approximation of nonlinear elliptic problems with discontinuous coefficients. *RAIRO Modél. Math. Anal. Numér.*, 24(4):457–500, 1990.

- [23] M. Feistauer and A. Ženíšek. Finite element solution of nonlinear elliptic problems. *Numer. Math.*, 50(4):451–475, 1987.
- [24] M. Feistauer and A. Ženíšek. Compactness method in the finite element theory of nonlinear elliptic problems. *Numer. Math.*, 52(2):147–163, 1988.
- [25] R. Glowinski and J. Rappaz. Approximation of a nonlinear elliptic problem arising in a non-Newtonian fluid flow model in glaciology. *M2AN Math. Model. Numer. Anal.*, 37(1):175–186, 2003.
- [26] F. Hermeline. Approximation of diffusion operators with discontinuous tensor coefficients on distorted meshes. *Comput. Methods Appl. Mech. Engrg.*, 192(16-18):1939–1959, 2003.
- [27] J. Leray and J.-L. Lions. Quelques résultats de Višik sur les problèmes elliptiques nonlinéaires par les méthodes de Minty-Browder. *Bull. Soc. Math. France*, 93:97–107, 1965.
- [28] A. Ženíšek. The finite element method for nonlinear elliptic equations with discontinuous coefficients. *Numer. Math.*, 58(1):51–77, 1990.