Finite volume schemes for the p-laplacian on cartesian meshes

Boris Andreianov,* Franck Boyer†and Florence Hubert†

Abstract. This paper is concerned with the finite volume approximation of the p-laplacian equation with homogeneous Dirichlet boundary conditions on rectangular meshes.

A reconstruction of the norm of the gradient on the mesh's interfaces is needed in order to discretize the p-laplacian operator. We give a detailed description of the possible nine points schemes ensuring that the solution of the resulting finite dimensional nonlinear system exists and is unique. These schemes, called admissible, are locally conservative and in addition derive from the minimization of a strictly convexe and coercive discrete functional.

The convergence rate is analyzed when the solution lies in $W^{2,p}$. Numerical results are given in order to compare different admissible and non-admissible schemes.

Keywords. finite volume methods, p-laplacian, error estimates.

AMS Subject Classification. 35J65 - 65N15 - 74S10

1 Introduction

1.1 Mathematical background

In this paper, we are interested in the study of the numerical approximation of solutions to the 2D p-laplacian with homogeneous Dirichlet boundary conditions ($1 < p < +\infty$):

$$\begin{cases} -\text{div}\left(|\nabla \overline{u}|^{p-2}\nabla\overline{u}\right) = f, & \text{on } \Omega, \\ \overline{u} = 0, & \text{on } \partial\Omega. \end{cases} \tag{1.1}$$

This equation arises as a model problem in various physical situations such as non-Newtonian flows (for example in glaciology models [13, 16]), turbulent flows, or in the study of flows in porous media. Some of these models are described and studied for instance in [8].

Throughout this paper, we will denote by $p' = \frac{p}{p-1}$ the conjugate exponent of p. Let us recall that for any data $f \in W^{-1,p'}(\Omega)$, problem (1.1) is well-posed in $W_0^{1,p}(\Omega)$. This can be proved easily using the monotony of the p-laplacian operator or equivalently using the fact that (1.1) is the Euler-Lagrange equation for the problem of minimizing the coercive and strictly convex functional

$$\begin{aligned} J : W_0^{1,p}(\Omega) &\rightarrow \mathbb{R} \\ u &\mapsto \frac{1}{p}\int_\Omega |\nabla u(z)|^p\, dz - \int_\Omega f(z)u(z)\, dz \end{aligned} \tag{1.2}$$

on the whole space $W_0^{1,p}(\Omega)$. Formula (1.2) is only valid for $f \in L^{p'}(\Omega)$ but we may replace the last integral by a duality bracket if f is in $W^{-1,p'}(\Omega)$. Furthermore, we have the *a priori* estimate

$$\|\overline{u}\|_{W^{1,p}} \leq \|f\|_{W^{-1,p'}}^{\frac{1}{p-1}} \leq C\|f\|_{L^{p'}}^{\frac{1}{p-1}}. \tag{1.3}$$

*Département de Mathématiques, Université de Franche-Comté, 16 route de Gray, 25030 Besançon Cedex, France
†Laboratoire d'Analyse, Topologie et Probabilités, 39 rue F. Joliot Curie, 13453 Marseille Cedex 13, France

Many authors (see for instance [4, 5, 6, 12]) have studied the finite elements approximation of (1.1). In this work, we propose a new approach in the framework of cell-centered finite volume schemes. Indeed, the choice of finite volume methods is natural as soon as the equation we consider is in a divergence form. It consists, for this kind of problems, in approaching the equation expressing the balance of fluxes in each control volume in a conservative way.

It appears that the approximation of the fluxes for (1.1) is not straightforward. Indeed, the difficulty arises from the need to approximate the complete gradient of the solution (and not only its normal derivatives on the edges as for the Laplace equation on admissible meshes). This problem of the complete gradient approximation for finite volume methods has interested many authors in the last few years in different contexts. We refer for instance to [7, 10, 11] for anisotropic diffusion problems and to [14, 15] for the Laplace equation on almost arbitrary meshes. In the context of non-linear diffusion problems, the finite volume approach with gradient reconstruction was only considered, to our knowledge, in the work by B. Andreianov, M. Gutnic and P. Wittbold in [3], for more general elliptic-parabolic equations and systems involving the p-laplacian. Their notion of admissible scheme concern particular meshes which are dual to triangular ones. They obtain a convergence result but no error analysis is made.

In this paper, we consider the approximation of (1.1) on rectangular meshes. In this particular case, we first give a complete description of the possible nine points finite volume schemes for which existence and uniqueness of the approximate solution is ensured in spite of the non-linearity of the problem. Moreover, this approximate solution can be computed by solving a finite dimensional minimization problem. We also analyze the convergence properties of these schemes and obtain error estimates.

1.2 Outline

This article is organized as follows. In section 2, we introduce the notations used in this work then we give the construction of the finite volume schemes we consider. We give the general form of conservative finite volume schemes that lead to a symmetric non-linear system (and thus derive from the minimization of a discrete functional approaching J), and that are consistent with piecewise affine functions. The symmetry property is crucial in order to ensure that for any given data, the scheme has a unique solution (Theorem 2.1).

The construction of the schemes is summed up in Definition 2.2 or equivalently in (2.26) which gives the explicit formula for the discrete energy associated to the scheme.

Section 3 is concerned with the proof of error estimates for the approximate solutions of problem (1.1) in the case where $\overline{u} \in W^{2,p}(\Omega)$. This assumption is a natural extension of the usual $H^2(\Omega)$ regularity for the Laplace equation (case $p = 2$), which is studied for example in [9]. Furthermore, it is proved in [4] that, when $1 < p \leq 2$ and $f \in L^q(\Omega)$ $(q > 2)$ then $\overline{u}$ actually belongs to $W^{2,p}(\Omega)$.

In this case we use similar methods than in [9] to prove error estimates, and we obtain the same kind of results than in the work by R. Glowinski and A. Marrocco [12] in the finite elements framework, namely a convergence order in $h^{\frac{1}{p-1}}$ when $p \geq 2$ (Theorem 3.1).

Notice that in the finite elements case, the convergence orders given in [12] have been improved by J.W. Barrett and W.B. Liu in [4, 5] and by S. Chow in [6]. The method used by S. Chow cannot be used in the finite volume case to improve the theorem 3.1 but it is useful to obtain error estimates for less regular solution. Indeed, in [2], we use it in the framework of Besov spaces, which are the natural regularity spaces for the p-laplacian (see [17]).

We conclude this paper by giving some numerical results in Section 4.

1.3 Remarks

Of course, many physical applications occur in domains with more complex geometry and the use of rectangular meshes seems to be too restrictive. Nevertheless, it is well known (especially in fluid mechanics) that one can cope with complex geometries with cartesian meshes using a penalization method. More precisely, in order to solve (1.1) in any bounded domain Ω, we take $\mathcal{U}$ a rectangle containing Ω and we solve instead of (1.1) the

penalized problem

$$\begin{cases} -\text{div}(|\nabla u_\varepsilon|^{p-2}\nabla u_\varepsilon) + \dfrac{1}{\varepsilon^p}\mathbf{1}_{\mathcal{U}\setminus\Omega}|u_\varepsilon|^{p-2}u_\varepsilon = f, & \text{in } \mathcal{U}, \\[2mm] \quad u_\varepsilon = 0, & \text{on } \partial\mathcal{U}. \end{cases} \tag{1.4}$$

Note that the penalization term is monotone so that this problem has a unique solution. Furthermore one can easily show that

$$\|\nabla u_\varepsilon\|_{L^p(\mathcal{U}\setminus\Omega)} + \frac{1}{\varepsilon}\|u_\varepsilon\|_{L^p(\mathcal{U}\setminus\Omega)} + \|\nabla\overline{u} - \nabla u_\varepsilon\|_{L^p(\Omega)} \le C\varepsilon^{\frac{1}{p}}. \tag{1.5}$$

Hence, one can use our schemes on a rectangular mesh on $\mathcal{U}$ to solve (1.4) which gives a good approximation of the solution to (1.1). Notice that the schemes we propose in this paper can be easily extended to the problem (1.4) without changing their basic properties. In particular, all the error estimates remain valid for the penalized problem. Notice that the estimate (1.5) may not be optimal. Indeed, for the Laplace equation as for the Stokes problem, we can obtain sharper results by studying the corresponding boundary layer problem (see for instance [1]).

This method can be useful in many situations. Indeed, the practical computation of the approximate solution on cartesian meshes can be easily optimized using multigrids methods and/or parallel computing. Furthermore, some glaciology models consist in free boundary problems in which a stationnary non-linear diffusion equation like (1.1) has to be solved at each time step in a time dependent domain (see for instance [16]). Such problems can also be handled with our methods by computing the solution to (1.4) on a fixed cartesian mesh with a time dependent penalization term. Finally, the extension of our schemes to the 3D case would be much more simple than for unstructured grids.

We also want to emphasize the fact that our study can be applied to more general non-linear diffusion equation of the form

$$-\text{div}(k(|\nabla\overline{u}|)\nabla\overline{u}) = f,$$

provided that k is smooth enough and satisfies some monotony properties (see [13]).

2 Construction of the finite volume schemes

2.1 Notations

Let Ω be a rectangular bounded domain of $\mathbb{R}^2$. We consider $\mathcal{T}$ a set of disjoint rectangular control volumes $\kappa \subset \Omega$ such that $\cup\bar{\kappa} = \overline{\Omega}$. We denote by h the maximum of the diameters of the control volumes in $\mathcal{T}$.

Let x_κ be the center of the control volume κ. For all adjacent control volumes κ and $\mathcal{L}$, denote by $\kappa|\mathcal{L}$ the edge between κ and $\mathcal{L}$.

In order to take the boundary conditions into account, we introduce artificial points constructed by symmetry with respect to the boundaries of Ω, as shown in Figure 2.1. This procedure is useful to state the scheme in a uniform way in the whole domain. We will give the details below.

The dual mesh $\mathcal{T}^*$ of $\mathcal{T}$ is defined to be the set of rectangle dual control volumes whose vertices are the points x_κ as well as the artificial points introduced above, see Figure 2.1.

Furthermore, for any control volume $\kappa \in \mathcal{T}$, we define (see Figure 2.2)

- $m(\kappa)$ the measure of κ;

- ν_κ the outward unit normal vector on $\partial\kappa$;

- $V_\kappa = \{\kappa^* \in \mathcal{T}^*/m(\kappa \cap \kappa^*) \neq 0\}$, the set of the four dual control volumes around the control volume κ;

- For any $\kappa^* \in V_\kappa$, $\sigma^h_{\kappa,\kappa^*}$ (resp. $\sigma^v_{\kappa,\kappa^*}$) is the horizontal (resp. vertical) half-edge of $\partial\kappa$ included in κ^* and $\mathcal{E}_\kappa = \{\sigma^h_{\kappa,\kappa^*}, \sigma^v_{\kappa,\kappa^*} / \kappa^* \in V_\kappa\}$ is the set of all the half-edges included in $\partial\kappa$.
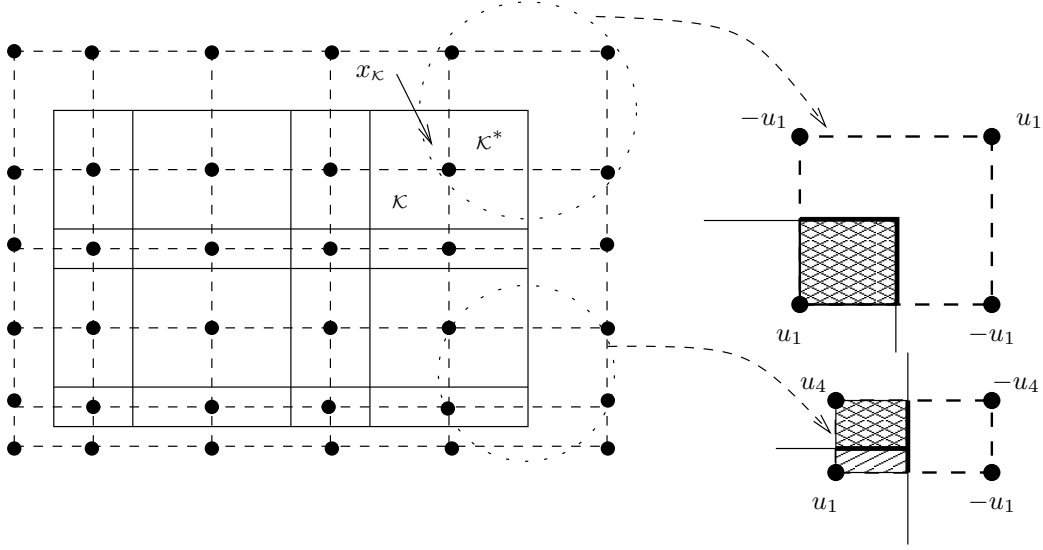
3

Figure 2.1: Non uniform rectangular mesh and treatment of the boundary conditions

We will see in the sequel that the schemes we propose are naturally written in each dual control volume. That is the reason why, given a dual control volume κ^*, we introduce some notations to describe the situation in κ^* (see Figure 2.3):

- $(x_i^{\kappa^*})_{i=1,2,3,4}$ are the vertices of the dual control volume κ^* numbered counter clockwise starting from the lower left hand corner.

- $(\kappa_i^{\kappa^*})_{i=1,2,3,4}$ are the corresponding control volumes with centers $(x_i^{\kappa^*})_{i=1,2,3,4}$.

- $l_i^{\kappa^*}$ is the distance between $x_i^{\kappa^*}$ and $x_{i+1}^{\kappa^*}$.

- $\sigma_i^{\kappa^*}$ is the half-edge between $\kappa_i^{\kappa^*}$ and $\kappa_{i+1}^{\kappa^*}$ located in κ^*,

- $m_i^{\kappa^*}$ is the measure of $\sigma_i^{\kappa^*}$.

Conventionnally, in a given dual control volume, the indices $i \in \mathbb{Z}$ are understood modulo 4.

The finite volume method associates to all control volumes κ an unknown value u_κ. We denote by $u^{\mathcal{T}} = (u_\kappa)_{\kappa \in \mathcal{T}} \in \mathbb{R}^{\mathcal{T}}$ the approximate solution on the mesh $\mathcal{T}$.

For any continuous function v on Ω, we will also denote by $v^{\mathcal{T}} = (v_\kappa)_{\kappa \in \mathcal{T}}$, with $v_\kappa = v(x_\kappa)$, the projection of v on the space $\mathbb{R}^{\mathcal{T}}$ of discrete functions.

• **Boundary conditions:**

We only consider in this paper the homogeneous Dirichlet condition. For a given discrete function $u^{\mathcal{T}} \in \mathbb{R}^{\mathcal{T}}$, which is a set of values on each control volume $\kappa \in \mathcal{T}$, the boundary conditions are taken into account by using the so-called ghost cells method. More precisely, we extend the values of $u^{\mathcal{T}}$ on artificial points outside of Ω, as shown in Figure 2.1, each time it is needed.

Given a dual control volume κ^*, we define the projection operator T_{κ^*} which associates to each $u^{\mathcal{T}} \in \mathbb{R}^{\mathcal{T}}$ its values

$$T_{\kappa^*}(u^{\mathcal{T}}) \overset{\text{def}}{=} (u_{1,\kappa^*}^{\mathcal{T}}, u_{2,\kappa^*}^{\mathcal{T}}, u_{3,\kappa^*}^{\mathcal{T}}, u_{4,\kappa^*}^{\mathcal{T}}) \tag{2.1}$$

in the four control volumes $(\kappa_i^{\kappa^*})_i$ around κ^*. Notice that for boundary dual control volumes, we need the ghost cells described above in order to give sense to the definition of T_{κ^*}. For instance, if κ^* is located at the
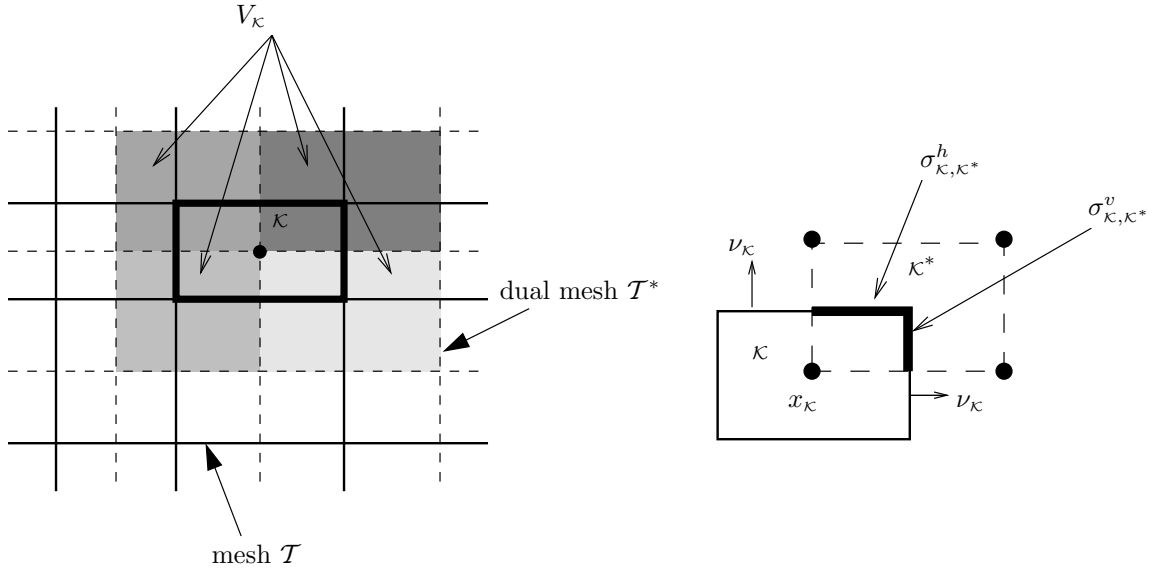
4

Figure 2.2: Notations

right boundary of Ω, then

$$u^{\mathcal{T}}_{2,\kappa^*} = -u^{\mathcal{T}}_{1,\kappa^*}, u^{\mathcal{T}}_{3,\kappa^*} = -u^{\mathcal{T}}_{4,\kappa^*},$$

as shown in Figure 2.1.

   • **Regularity of mesh families:**
   In the sequel, we will prove error estimates for families of meshes which satisfy the following regularity assumption:

   There exists a constant c_1 (independent of the mesh $\mathcal{T}$) such that

$$m(\kappa) \geq c_1 h^2, \ \forall \kappa \in \mathcal{T}. \tag{2.2}$$

Remark 2.1 *Under the previous regularity assumption, there exist constants c_2, c_3 independent of $\mathcal{T}$ such that, for all dual control volume $\kappa^* \in \mathcal{T}^*$,*

$$\mathrm{diam}(\kappa^*) \leq c_2 h \ \text{and} \ m(\kappa^*) \geq c_3 h^2.$$

2.2   Discrete norms and classical inequalities

Each discrete function $u^{\mathcal{T}} \in \mathbb{R}^{\mathcal{T}}$ can be considered as a (piecewise constant) function in $L^r(\Omega)$ (for any $r \in [1, +\infty])$ by noting abusively

$$u^{\mathcal{T}}(z) = \sum_{\kappa \in \mathcal{T}} u_\kappa \mathbf{1}_\kappa(z), \ \ \forall z \in \Omega.$$

Then, we naturally define

$$\|u^{\mathcal{T}}\|_{L^r} = \left(\sum_{\kappa \in \mathcal{T}} m(\kappa)|u_\kappa|^r \right)^{\frac{1}{r}}, \ \ \forall r < +\infty,$$

and

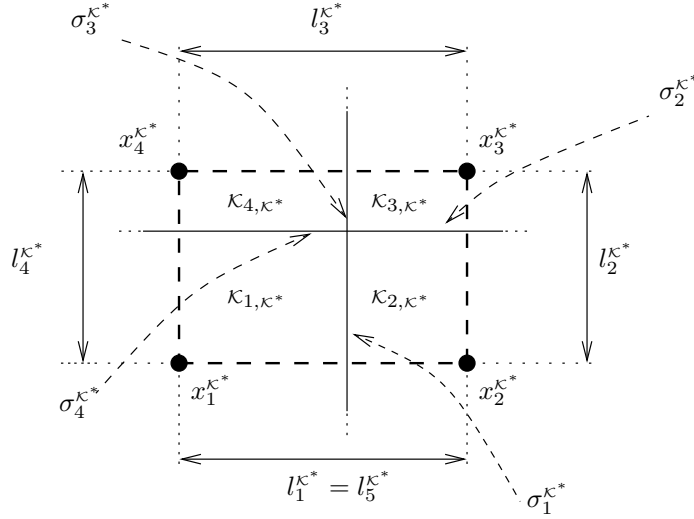$$\|u^{\mathcal{T}}\|_{L^\infty} = \sup_{\kappa \in \mathcal{T}} |u_\kappa|.$$

5

Figure 2.3: Notations in a dual control volume κ^*

Now let us define a discrete Sobolev norm for the elements of $\mathbb{R}^{\mathcal{T}}$. For any $u^{\mathcal{T}} \in \mathbb{R}^{\mathcal{T}}$, and any $\kappa^* \in \mathcal{T}^*$, denote by $\delta_i^{\kappa^*}(u^{\mathcal{T}})$ the differential quotient

$$\delta_i^{\kappa^*}(u^{\mathcal{T}}) = \frac{u_{i+1,\kappa^*}^{\mathcal{T}} - u_{i,\kappa^*}^{\mathcal{T}}}{l_i^{\kappa^*}}, \quad \forall i \in \{1,\dots,4\}.$$

One can remark that $\delta_1^{\kappa^*}$ and $-\delta_3^{\kappa^*}$ are approximations of the horizontal derivative u_x while $\delta_2^{\kappa^*}$ and $-\delta_4^{\kappa^*}$ are approximations of the vertical derivative u_y.

Definition 2.1 *For any $u^{\mathcal{T}} \in \mathbb{R}^{\mathcal{T}}$ and any $\kappa^* \in \mathcal{T}^*$, we introduce an approximation of $|\nabla u|$ on κ^* defined by*

$$|u^{\mathcal{T}}|_{1,\kappa^*} = \left(\frac{1}{2}\sum_{i=1}^4 \left|\delta_i^{\kappa^*}(u^{\mathcal{T}})\right|^2\right)^{\frac{1}{2}}.$$

The discrete $W_0^{1,p}$ norm of $u^{\mathcal{T}}$ is then defined by

$$\|u^{\mathcal{T}}\|_{1,p,\mathcal{T}} = \left(\sum_{\kappa^* \in \mathcal{T}^*} m(\kappa^* \cap \Omega)|u^{\mathcal{T}}|_{1,\kappa^*}^p\right)^{\frac{1}{p}}.$$

Remark 2.2

- *The Dirichlet boundary conditions are taken into account in the definition of the discrete $W_0^{1,p}$ norm through the introduction of the ghost cells. Actually, this implies that $\|\cdot\|_{1,p,\mathcal{T}}$ is a norm on $\mathbb{R}^{\mathcal{T}}$.*

- *For any function $v \in W_0^{1,p}(\Omega)$, with $p > 2$, we have*

$$\|v^{\mathcal{T}}\|_{1,p,\mathcal{T}} \leq C\|v\|_{W^{1,p}}, \tag{2.3}$$

where $v^{\mathcal{T}} = (v(x_\kappa))_{\kappa \in \mathcal{T}}$ is the projection of v on the mesh $\mathcal{T}$, and C depends only on the mesh regularity constant c_1 in (2.2).

Let us state the discrete version of the Poincaré inequality. This result is classical in the case $p = 2$ (see for example [9]). When $p \neq 2$, it is proved in a slightly different context in [3]. Note that one does not assume any regularity assumptions on the mesh.

Lemma 2.1 (Discrete Poincaré inequality) *Let $\mathcal{T}$ be a mesh of the rectangle Ω. There exists a constant C which only depends on p such that for any $u^{\mathcal{T}} \in \mathbb{R}^{\mathcal{T}}$, we have*

$$\|u^{\mathcal{T}}\|_{L^p} \leq C \operatorname{diam}(\Omega) \|u^{\mathcal{T}}\|_{1,p,\mathcal{T}}.$$

Proof :
It is a straightforward adaptation from the proof in [3], taking into account the ghost cells method used for the boundary conditions. ∎

2.3 Construction of the schemes

A family of nine-points finite volume schemes is proposed to compute approximate solutions of problem (1.1). Such schemes are obtained by integrating equation (1.1) on each control volume $\kappa \in \mathcal{T}$:

$$
\begin{aligned}
\int_{\kappa} f(z)\, dz &= \int_{\kappa} -\operatorname{div}\left(|\nabla \overline{u}|^{p-2} \nabla \overline{u}\right)\, dz = -\int_{\partial \kappa} |\nabla \overline{u}|^{p-2} \nabla \overline{u} \cdot \nu_{\kappa}\, ds \\
&= \sum_{\kappa^* \in V_{\kappa}} \left(-\int_{\sigma_{\kappa,\kappa^*}^h} |\nabla \overline{u}|^{p-2} \nabla \overline{u} \cdot \nu_{\kappa}\, ds - \int_{\sigma_{\kappa,\kappa^*}^v} |\nabla \overline{u}|^{p-2} \nabla \overline{u} \cdot \nu_{\kappa}\, ds \right)
\end{aligned}
\tag{2.4}
$$

Let f_{κ} denote the mean value of the function f on the control volume κ

$$f_{\kappa} = \frac{1}{m(\kappa)} \int_{\kappa} f(z)\, dz. \tag{2.5}$$

The finite volume method consists in approaching the exact relation between the fluxes (2.4) by a system of discrete equations

$$a_{\kappa}(u^{\mathcal{T}}) \stackrel{\text{def}}{=} \sum_{\kappa^* \in V_{\kappa}} a_{\kappa,\kappa^*}(u^{\mathcal{T}}) = m(\kappa) f_{\kappa}, \quad \forall \kappa \in \mathcal{T}, \tag{2.6}$$

where $a_{\kappa,\kappa^*}(u^{\mathcal{T}})$ is a numerical flux to be determined which is supposed to approach the exact flux, that is:

$$a_{\kappa,\kappa^*}(u^{\mathcal{T}}) \approx -\int_{\sigma_{\kappa,\kappa^*}^h} |\nabla \overline{u}|^{p-2} \nabla \overline{u} \cdot \nu_{\kappa}\, ds - \int_{\sigma_{\kappa,\kappa^*}^v} |\nabla \overline{u}|^{p-2} \nabla \overline{u} \cdot \nu_{\kappa}\, ds. \tag{2.7}$$

2.3.1 General structure

The classical finite volume theory for the Laplace equation (see [10] for instance) provides very natural ways to approach the normal derivative $\nabla \overline{u} \cdot \nu$ on the edges. On the other hand, in the literature there is no systematic method to discretize the norm of the gradient of the solution.

In this paragraph, we describe a wide class of possible choices for the numerical fluxes $a_{\kappa,\kappa^*}(u^{\mathcal{T}})$ in which the approximations of the terms $\nabla \overline{u} \cdot \nu$ and $|\nabla \overline{u}|^{p-2}$ are chosen independently. Actually, in the next paragraph we show that these two discretizations have to be compatible (in a sense to be precised) to ensure good properties of the schemes, in particular the well-posedness of the discrete non-linear equations.

The general framework we propose is the following.

1. On each dual control volume κ^*, we choose to approximate $|\nabla \overline{u}|^2$ by a quadratic form (to be precised later) in the variables $(u_{i,\kappa^*}^{\mathcal{T}})_{\{i=1,2,3,4\}}$, namely

$$|\nabla \overline{u}|^2 \approx q_{\kappa^*}(T_{\kappa^*}(u^{\mathcal{T}})) \quad \text{on } \kappa^*, \tag{2.8}$$

7

where $T_{\kappa^*}(u^{\mathcal{T}})$ is defined at the end of section 2.1. This leads to a first approximation of the fluxes on any half edge $\sigma \subset \kappa^*$ by

$$-\int_\sigma |\nabla \overline{u}|^{p-2} \nabla \overline{u} \cdot \nu_\kappa \, ds \approx -q_{\kappa^*}(T_{\kappa^*}(u^{\mathcal{T}}))^{\frac{p-2}{2}} \int_\sigma \nabla \overline{u} \cdot \nu_\kappa \, ds.$$

2. The problem is now reduced to the approximation of the same linear fluxes as for the Laplace equation. We suppose that we are given linear forms A^h_{κ,κ^*}, A^v_{κ,κ^*}, in the variables $(u^{\mathcal{T}}_{i,\kappa^*})_{\{i=1,2,3,4\}}$, such that

$$-\int_{\sigma^h_{\kappa,\kappa^*}} \nabla \overline{u} \cdot \nu_\kappa \, ds \approx A^h_{\kappa,\kappa^*}(T_{\kappa^*}(u^{\mathcal{T}})) \quad \text{and} \quad -\int_{\sigma^v_{\kappa,\kappa^*}} \nabla \overline{u} \cdot \nu_\kappa \, ds \approx A^v_{\kappa,\kappa^*}(T_{\kappa^*}(u^{\mathcal{T}})). \tag{2.9}$$

3. The numerical flux satisfying (2.7) is then given by

$$a_{\kappa,\kappa^*}(u^{\mathcal{T}}) \stackrel{\text{def}}{=} q_{\kappa^*}(T_{\kappa^*}(u^{\mathcal{T}}))^{\frac{p-2}{2}} \left(A^h_{\kappa,\kappa^*}(T_{\kappa^*}(u^{\mathcal{T}})) + A^v_{\kappa,\kappa^*}(T_{\kappa^*}(u^{\mathcal{T}})) \right).$$

2.3.2 Choice of q_{κ^*} and A^h_{κ,κ^*}, A^v_{κ,κ^*}

First of all, it is convenient to represent q_{κ^*} through a symmetric non negative matrix on $\mathbb{R}^4$ denoted by B_{κ^*} so that we have

$$q_{\kappa^*}(v) = (B_{\kappa^*}v, v), \ \forall v \in \mathbb{R}^4.$$

Notice that the schemes proposed above can also be written as a sum of contributions on each dual control volume. Indeed, each numerical flux $a_{\kappa,\kappa^*}(u^{\mathcal{T}})$ depends only on the values $T_{\kappa^*}(u^{\mathcal{T}}) = (u^{\mathcal{T}}_{i,\kappa^*})_{\{i=1,2,3,4\}}$ of $u^{\mathcal{T}}$ at the four vertices of the dual control volume κ^*. More precisely, one can easily see that the map $a : u^{\mathcal{T}} \in \mathbb{R}^{\mathcal{T}} \mapsto (a_\kappa(u^{\mathcal{T}}))_\kappa \in \mathbb{R}^{\mathcal{T}}$ which represents the scheme and defined in (2.6), reads

$$a(u^{\mathcal{T}}) = \sum_{\kappa^* \in \mathcal{T}^*} m(\kappa^* \cap \Omega) T^t_{\kappa^*} \circ a_{\kappa^*} \circ T_{\kappa^*}(u^{\mathcal{T}}) \tag{2.10}$$

with

$$a_{\kappa^*}(v) \stackrel{\text{def}}{=} \frac{1}{m(\kappa^*)} (B_{\kappa^*}v, v)^{\frac{p-2}{2}} \begin{pmatrix} A^h_{\kappa_1,\kappa^*}v + A^v_{\kappa_1,\kappa^*}v \\ A^h_{\kappa_2,\kappa^*}v + A^v_{\kappa_2,\kappa^*}v \\ A^h_{\kappa_3,\kappa^*}v + A^v_{\kappa_3,\kappa^*}v \\ A^h_{\kappa_4,\kappa^*}v + A^v_{\kappa_4,\kappa^*}v \end{pmatrix} = \frac{1}{m(\kappa^*)} (B_{\kappa^*}v, v)^{\frac{p-2}{2}} A_{\kappa^*}v, \ \forall v \in \mathbb{R}^4, \tag{2.11}$$

where A_{κ^*} is the corresponding 4×4 matrix.

We first derive some properties on B_{κ^*} and A_{κ^*}, which ensure the consistency of the scheme but also the existence and uniqueness of a solution of the scheme.

● **Consistency of q_{κ^*} :**
By definition, B_{κ^*} is a symmetric non-negative matrix. Moreover, q_{κ^*} is supposed to approach $|\nabla \overline{u}|^2$ so that is natural to impose the following consistency property:

$$\begin{aligned} &\text{Approximation (2.8) is exact for affine functions on } \kappa^*, \\ &\text{and } q_{\kappa^*}(v) = 0 \text{ if and only if } v_1 = v_2 = v_3 = v_4. \end{aligned} \tag{2.12}$$

In particular, this imposes that the kernel of B_{κ^*} is spanned by the vector $e_0 = (1,1,1,1)^t$, so that the rank of B_{κ^*} is 3.

● **Conservativity of the linear fluxes:**

8

Any reasonable finite volume scheme has to be conservative, in our case it means that, for any dual control volume κ^*, the linear fluxes have to be in the following form

$$
\begin{cases}
A^h_{\kappa_1,\kappa^*} = \;\;\; A_{4,\kappa^*}, & A^v_{\kappa_1,\kappa^*} = -A_{1,\kappa^*}, \\
A^h_{\kappa_2,\kappa^*} = -A_{2,\kappa^*}, & A^v_{\kappa_2,\kappa^*} = \;\;\; A_{1,\kappa^*}, \\
A^h_{\kappa_3,\kappa^*} = \;\;\; A_{2,\kappa^*}, & A^v_{\kappa_3,\kappa^*} = -A_{3,\kappa^*}, \\
A^h_{\kappa_4,\kappa^*} = -A_{4,\kappa^*}, & A^v_{\kappa_4,\kappa^*} = \;\;\; A_{3,\kappa^*},
\end{cases}
\tag{2.13}
$$

where A_{i,κ^*} is an approximation of the linear flux $\displaystyle\int_{\sigma_i^{\kappa^*}} \nabla \overline{u} \cdot \nu_i^{\kappa^*}\, ds$, $\nu_i^{\kappa^*}$ being the unit normal to $\sigma_i^{\kappa^*}$ oriented from $\kappa_i^{\kappa^*}$ to $\kappa_{i+1}^{\kappa^*}$ (see Figure 2.3).

• **Construction of the linear fluxes:**

To ensure the consistency of the numerical fluxes A_{i,κ^*}, the approximation has to be exact on affine functions. One easily checks that this requirement amounts to

$$
A_{i,\kappa^*} = m_i^{\kappa^*} \left(\mu_i^{\kappa^*} \frac{u^{\mathcal{T}}_{i+1,\kappa^*} - u^{\mathcal{T}}_{i,\kappa^*}}{l_i^{\kappa^*}} + (1 - \mu_i^{\kappa^*}) \frac{u^{\mathcal{T}}_{i+2,\kappa^*} - u^{\mathcal{T}}_{i+3,\kappa^*}}{l_{i+2}^{\kappa^*}} \right),
\tag{2.14}
$$

where $\mu_i^{\kappa^*}$ is a real parameter. Notice that the classical finite volume approximation of the Laplace equation (i.e. when $p = 2$) corresponds to the case $\mu_i^{\kappa^*} = 1, \forall i \in \{1,\ldots,4\}, \forall \kappa^* \in \mathcal{T}^*$.

For any $\kappa^* \in \mathcal{T}^*$ and any i, let us define

$$
\tau_i^{\kappa^*} = \frac{m_i^{\kappa^*}}{l_i^{\kappa^*}},
\tag{2.15}
$$

which depends only on the geometrical structure of the mesh. Using (2.13) the matrix A_{κ^*} defined in (2.11) reads for each dual control volume κ^*

$$
\begin{pmatrix}
\tau_1\mu_1 + \tau_4\mu_4 & -\tau_1\mu_1 + \tau_4(1-\mu_4) & -\tau_1(1-\mu_1) - \tau_4(1-\mu_4) & -\tau_4\mu_4 + \tau_1(1-\mu_1) \\
\tau_2(1-\mu_2) - \tau_1\mu_1 & \tau_2\mu_2 + \tau_1\mu_1 & -\tau_2\mu_2 + \tau_1(1-\mu_1) & -\tau_2(1-\mu_2) - \tau_1(1-\mu_1) \\
-\tau_3(1-\mu_3) - \tau_2(1-\mu_2) & \tau_3(1-\mu_3) - \tau_2\mu_2 & \tau_3\mu_3 + \tau_2\mu_2 & -\tau_3\mu_3 + \tau_2(1-\mu_2) \\
\tau_3(1-\mu_3) - \tau_4\mu_4 & -\tau_4(1-\mu_4) - \tau_3(1-\mu_3) & \tau_4(1-\mu_4) - \tau_3\mu_3 & \tau_4\mu_4 + \tau_3\mu_3
\end{pmatrix}
\tag{2.16}
$$

where we dropped the subscript κ^* for $\tau_i^{\kappa^*}$ and $\mu_i^{\kappa^*}$.

• **Symmetry property of the scheme:**

One of the fundamental properties of the continuous equation (1.1) is that it derives from a minimization problem. It is natural to look for numerical schemes which preserve this property.

Indeed, the scheme (2.23) is the Euler-Lagrange equation for critical points of a functional $J_{\mathcal{T}} : \mathbb{R}^{\mathcal{T}} \mapsto \mathbb{R}$ if and only if the differential of the map

$$
a : u^{\mathcal{T}} = (u_\kappa)_{\kappa \in \mathcal{T}} \mapsto (a_\kappa(u^{\mathcal{T}}))_{\kappa \in \mathcal{T}},
$$

is symmetric at each point. In that case, we shall see later that we can give an explicit formula for $J_{\mathcal{T}}$, and that $J_{\mathcal{T}}$ is strictly convex and coercive.

For non-symmetric schemes, the existence and uniqueness of a solution to the non-linear discrete equations is not guaranteed. Even if the solution exists, its practical computation is not obvious. That is the reason, from now on, we focus our attention on symmetric schemes.

We now have the following proposition which characterizes such schemes in our framework. The proof is given in Section 2.6.

Proposition 2.1 *The differential of the map* $a : u^{\mathcal{T}} = (u_\kappa)_{\kappa \in \mathcal{T}} \mapsto (a_\kappa(u^{\mathcal{T}}))_{\kappa \in \mathcal{T}}$ *is symmetric at each point if and only if one of the two following conditions is fulfilled.*

9

- $p = 2$ and A_{κ^*} is a symmetric matrix.

- $p \neq 2$ and for any $\kappa^* \in \mathcal{T}^*$ there exists a $\lambda_{\kappa^*} \in \mathbb{R}$ such that

$$A_{\kappa^*} = \lambda_{\kappa^*} B_{\kappa^*}.$$

In order for the scheme to be consistent with the balance equation (2.4) we will see later that, necessarily, $\lambda_{\kappa^*} = m(\kappa^*)$.

Hence, at this point, the choice of the discretization of the linear part of the fluxes across half-edges in each dual control volume (i.e. the choice of the parameters $\mu_i^{\kappa^*}$ in (2.14)) determines the scheme completely. In other words, if we want the scheme to be symmetric we cannot choose independently the discretization of the norm of the gradient $|\nabla \overline{u}|$ (represented by B_{κ^*}) and the discretization of the normal derivative $\nabla \overline{u} \cdot \nu_\kappa$ (represented by A_{κ^*}).

Furthermore, Proposition 2.1 also implies that A_{κ^*} is necessarily symmetric. The expression (2.16) shows that the $(\mu_i^{\kappa^*})_i$ have to fulfill the following constraints

$$\begin{cases} \tau_4^{\kappa^*}(\mu_4^{\kappa^*} - 1) = \tau_2^{\kappa^*}(\mu_2^{\kappa^*} - 1), \\ \tau_3^{\kappa^*}(\mu_3^{\kappa^*} - 1) = \tau_1^{\kappa^*}(\mu_1^{\kappa^*} - 1). \end{cases} \tag{2.17}$$

The scheme on the dual control volume κ^* now seems to be determined by the choice of $\mu_1^{\kappa^*}$ and $\mu_4^{\kappa^*}$ for instance. Nevertheless, using (2.17), it is easy to see that the matrix A_{κ^*} given by (2.16) depends only on the quantity $\tau_1^{\kappa^*}\mu_1^{\kappa^*} + \tau_4^{\kappa^*}\mu_4^{\kappa^*}$. Consequently, without loss of generality the coefficients $(\mu_i^{\kappa^*})_i$ can be chosen so that

$$\tau_1^{\kappa^*}(\mu_1^{\kappa^*} - 1) = \tau_2^{\kappa^*}(\mu_2^{\kappa^*} - 1) = \tau_3^{\kappa^*}(\mu_3^{\kappa^*} - 1) = \tau_4^{\kappa^*}(\mu_4^{\kappa^*} - 1) \stackrel{\text{def}}{=} \xi^{\kappa^*}. \tag{2.18}$$

Therefore, for any κ^* the matrix A_{κ^*} only depends on a real parameter ξ^{κ^*} in the following way:

$$A_{\kappa^*} = \begin{pmatrix} 2\xi^{\kappa^*} + \tau_1^{\kappa^*} + \tau_4^{\kappa^*} & -2\xi^{\kappa^*} - \tau_1^{\kappa^*} & 2\xi^{\kappa^*} & -2\xi^{\kappa^*} - \tau_4^{\kappa^*} \\ -2\xi^{\kappa^*} - \tau_1^{\kappa^*} & 2\xi^{\kappa^*} + \tau_1^{\kappa^*} + \tau_2^{\kappa^*} & -2\xi^{\kappa^*} - \tau_2^{\kappa^*} & 2\xi^{\kappa^*} \\ 2\xi^{\kappa^*} & -2\xi^{\kappa^*} - \tau_2^{\kappa^*} & 2\xi^{\kappa^*} + \tau_2^{\kappa^*} + \tau_3^{\kappa^*} & -2\xi^{\kappa^*} - \tau_3^{\kappa^*} \\ -2\xi^{\kappa^*} - \tau_4^{\kappa^*} & 2\xi^{\kappa^*} & -2\xi^{\kappa^*} - \tau_3^{\kappa^*} & 2\xi^{\kappa^*} + \tau_4^{\kappa^*} + \tau_3^{\kappa^*} \end{pmatrix}, \tag{2.19}$$

and B_{κ^*} is given by Proposition 2.1.

- **Consistency of the scheme:**

As we have claimed above, the consistency of the whole scheme determines the value of the coefficient λ_{κ^*} introduced in Proposition 2.1.

Indeed, let w be an affine function in a dual control volume κ^*, whose gradient is $(\alpha, \beta)^t$, and let $v = T_{\kappa^*}w^\mathcal{T}$. We easily check that

$$v_2 = v_1 + \alpha l_1^{\kappa^*}, v_3 = v_1 + \alpha l_1^{\kappa^*} + \beta l_2^{\kappa^*}, v_4 = v_1 + \beta l_2^{\kappa^*},$$

so that, using (2.19), we have

$$A_{\kappa^*}v = \begin{pmatrix} -m_4^{\kappa^*}\beta - m_1^{\kappa^*}\alpha \\ -m_2^{\kappa^*}\beta + m_1^{\kappa^*}\alpha \\ m_2^{\kappa^*}\beta + m_3^{\kappa^*}\alpha \\ m_4^{\kappa^*}\beta - m_3^{\kappa^*}\alpha \end{pmatrix}.$$

As $l_1^{\kappa^*} = m_2^{\kappa^*} + m_4^{\kappa^*}$ and $l_2^{\kappa^*} = m_1^{\kappa^*} + m_3^{\kappa^*}$ (see Figure 2.3), we get

$$\begin{aligned} (A_{\kappa^*}v, v) &= -m_1^{\kappa^*}\alpha(v_1 - v_2) - m_3^{\kappa^*}\alpha(v_4 - v_3) - m_2^{\kappa^*}\beta(v_2 - v_3) - m_4^{\kappa^*}\beta(v_1 - v_4) \\ &= l_1^{\kappa^*}m_1^{\kappa^*}\alpha^2 + l_1^{\kappa^*}m_3^{\kappa^*}\alpha^2 + m_2^{\kappa^*}l_2^{\kappa^*}\beta^2 + m_4^{\kappa^*}l_2^{\kappa^*}\beta^2 \\ &= l_1^{\kappa^*}l_2^{\kappa^*}(\alpha^2 + \beta^2) = m(\kappa^*)|\nabla w|^2. \end{aligned}$$

By Proposition 2.1, we also have $A_{\kappa^*} = \lambda_{\kappa^*} B_{\kappa^*}$ so that

$$\lambda_{\kappa^*} (B_{\kappa^*} v, v) = m(\kappa^*)|\nabla w|^2.$$

Hence, the consistency requirement (2.12) holds if and only if $\lambda_{\kappa^*} = m(\kappa^*)$.

• Admissible schemes:

Given a dual control volume κ^* and a value of ξ^{κ^*}, the matrices A_{κ^*} and B_{κ^*} are defined by (2.19) and by $B_{\kappa^*} = \frac{1}{m(\kappa^*)} A_{\kappa^*}$. It remains to check that B_{κ^*} is a non negative matrix of rank 3, as required previously.

It is easy to check that the kernel of the matrix B_{κ^*} defined above contains e_0. Consider now the quadratic form $v \mapsto (B_{\kappa^*} v, v)$ restricted to the hyperplane $\{e_0\}^\perp$ orthogonal to the vector e_0. In the basis $(-1, 1, 1, -1)^t$, $(1, 1, -1, -1)^t$, $(1, -1, 1, -1)^t$, its matrix reads

$$\frac{4}{m(\kappa^*)} \begin{pmatrix} \tau_1^{\kappa^*} + \tau_3^{\kappa^*} & 0 & \tau_3^{\kappa^*} - \tau_1^{\kappa^*} \\ 0 & \tau_2^{\kappa^*} + \tau_4^{\kappa^*} & \tau_4^{\kappa^*} - \tau_2^{\kappa^*} \\ \tau_3^{\kappa^*} - \tau_1^{\kappa^*} & \tau_4^{\kappa^*} - \tau_2^{\kappa^*} & 8\xi^{\kappa^*} + \tau_1^{\kappa^*} + \tau_2^{\kappa^*} + \tau_3^{\kappa^*} + \tau_4^{\kappa^*} \end{pmatrix}. \tag{2.20}$$

The computation of main minor determinants show that this matrix is positive definite if and only if

$$2\xi^{\kappa^*} + \frac{m_1^{\kappa^*} m_3^{\kappa^*} + m_2^{\kappa^*} m_4^{\kappa^*}}{l_1^{\kappa^*} l_2^{\kappa^*}} > 0.$$

This gives a condition on ξ^{κ^*} which ensures the admissibility of the scheme. It will be shown that this condition guarantees existence and uniqueness for the approximate solution $u^\mathcal{T}$ as well as the discrete $W^{1,p}$ estimate (see Lemmas 2.3 and 3.1).

2.3.3 Admissible finite volume schemes

Let us sum up our construction of the schemes in the following definition. We recall that the projector T_{κ^*} is defined in (2.1). In particular, the boundary conditions are taken into account in the scheme through the definition of this projector for boundary dual control volumes.

Definition 2.2 *For each dual control volume κ^*, we choose a real number ξ^{κ^*} such that*

$$2\xi^{\kappa^*} + \frac{m_1^{\kappa^*} m_3^{\kappa^*} + m_2^{\kappa^*} m_4^{\kappa^*}}{l_1^{\kappa^*} l_2^{\kappa^*}} > 0. \tag{2.21}$$

Recall that $\tau_i^{\kappa^*} = \frac{m_i^{\kappa^*}}{l_i^{\kappa^*}}$. Then, we define A_{κ^*} by (2.19) and B_{κ^*} by

$$B_{\kappa^*} = \frac{1}{m(\kappa^*)} A_{\kappa^*}. \tag{2.22}$$

The finite volume scheme associated to the parameters $(\xi^{\kappa^*})_{\kappa^* \in \mathcal{T}^*}$ is then defined by

$$a(u^\mathcal{T}) = (m(\kappa) f_\kappa)_{\kappa \in \mathcal{T}}, \tag{2.23}$$

where $a = \sum_{\kappa^* \in \mathcal{T}^*} m(\kappa^* \cap \Omega) T_{\kappa^*}^t \circ a_{\kappa^*} \circ T_{\kappa^*}$, and

$$a_{\kappa^*}(v) = (B_{\kappa^*} v, v)^{\frac{p-2}{2}} B_{\kappa^*} v, \quad \text{for all } v \in \mathbb{R}^4.$$

Note that the admissibility condition (2.21) is fundamental to give sense to the definition of a_{κ^*} since it ensures that $(B_{\kappa^*} v, v) \geq 0$ for any $v \in \mathbb{R}^4$.

2.4   Discrete energy

2.4.1   Construction of the discrete energy

The schemes defined in Definition 2.2 are built in such a way that the symmetry of the differential of the map a on $\mathbb{R}^{\mathcal{T}}$ is ensured. Hence, there exists a discrete functional $J_{\mathcal{T}} : u^{\mathcal{T}} \in \mathbb{R}^{\mathcal{T}} \mapsto \mathbb{R}$, which we call the discrete energy of the system, such that we have

$$a_{\kappa}(u^{\mathcal{T}}) - m(\kappa)f_{\kappa} = \frac{\partial J_{\mathcal{T}}}{\partial u_{\kappa}}, \quad \forall \kappa \in \mathcal{T}. \tag{2.24}$$

Using the homogeneity of the map a, it is possible to obtain an explicit formula for $J_{\mathcal{T}}$. Indeed, let us check that the functional defined by

$$J_{\mathcal{T}}(u^{\mathcal{T}}) = \frac{1}{p}\sum_{\kappa \in \mathcal{T}} a_{\kappa}(u^{\mathcal{T}})u_{\kappa} - \sum_{\kappa \in \mathcal{T}} m(\kappa)f_{\kappa}u_{\kappa}$$

is the discrete energy of the scheme (which is unique up to a constant). An easy computation shows that

$$\nabla J_{\mathcal{T}}(u^{\mathcal{T}}) = \frac{1}{p}(da(u^{\mathcal{T}}))^{t}.u^{\mathcal{T}} + \frac{1}{p}a(u^{\mathcal{T}}) - (m(\kappa)f_{\kappa})_{\kappa}. \tag{2.25}$$

Since $a : (u_{\kappa})_{\kappa} \mapsto (a_{\kappa}(u^{\mathcal{T}}))_{\kappa}$ is a positively homogeneous function of degree $p-1$, we have

$$da(u^{\mathcal{T}}).u^{\mathcal{T}} = (p-1)\,a(u^{\mathcal{T}}).$$

Since, by construction (see Proposition 2.1), we know that $da(u^{\mathcal{T}})$ is symmetric, we obtain

$$(da(u^{\mathcal{T}}))^{t}.u^{\mathcal{T}} = da(u^{\mathcal{T}}).u^{\mathcal{T}} = (p-1)\,a(u^{\mathcal{T}}).$$

Hence, (2.25) yields (2.24).

Let us now obtain a more precise expression for the energy $J_{\mathcal{T}}$.

Lemma 2.2 *Let $u^{\mathcal{T}}, v^{\mathcal{T}} \in \mathbb{R}^{\mathcal{T}}$, we have*

$$a(u^{\mathcal{T}}) \cdot v^{\mathcal{T}} \stackrel{\text{def}}{=} \sum_{\kappa \in \mathcal{T}^{*}} a_{\kappa}(u^{\mathcal{T}})v_{\kappa} = \sum_{\kappa^{*} \in \mathcal{T}^{*}} m(\kappa^{*} \cap \Omega)(B_{\kappa^{*}}T_{\kappa^{*}}(u^{\mathcal{T}}), T_{\kappa^{*}}(u^{\mathcal{T}}))^{\frac{p-2}{2}}(B_{\kappa^{*}}T_{\kappa^{*}}(u^{\mathcal{T}}), T_{\kappa^{*}}(v^{\mathcal{T}})).$$

Proof :
It is a straightforward application of formulas (2.10) and (2.22). ∎

Denote by $B_{\kappa^{*}}^{\frac{1}{2}}$ the square root of the non-negative symmetric matrix $B_{\kappa^{*}}$. From Lemma 2.2, we deduce that

$$J_{\mathcal{T}}(u^{\mathcal{T}}) = \frac{1}{p}\sum_{\kappa^{*} \in \mathcal{T}^{*}} m(\kappa^{*} \cap \Omega)|B_{\kappa^{*}}^{\frac{1}{2}}T_{\kappa^{*}}(u^{\mathcal{T}})|^{p} - \sum_{\kappa \in \mathcal{T}} m(\kappa)f_{\kappa}u_{\kappa}. \tag{2.26}$$

It is now clear that this discrete functional can be seen as an approximation of the continuous functional J associated to problem (1.1).

2.4.2   Properties of $J_{\mathcal{T}}$

Let us begin with the following technical lemma:

Lemma 2.3 *Consider a mesh $\mathcal{T}$ satisfying the regularity assumption (2.2). Consider the matrices A_{κ^*} and B_{κ^*} defined by (2.19) and B_{κ^*} (2.22).*

Let $\gamma > 0$ be such that

$$\xi^{\kappa^*} \leq \frac{1}{\gamma}, \quad \text{and} \quad 2\xi^{\kappa^*} + \frac{m_1^{\kappa^*} m_3^{\kappa^*} + m_2^{\kappa^*} m_4^{\kappa^*}}{l_1^{\kappa^*} l_2^{\kappa^*}} \geq \gamma, \quad \forall \kappa^* \in \mathcal{T}^*. \tag{2.27}$$

There exist $\beta_1, \beta_2 > 0$, depending only on γ and on c_1 in (2.2), such that

$$\beta_1 |u^{\mathcal{T}}|_{1,\kappa^*} \leq |B_{\kappa^*}^{\frac{1}{2}} T_{\kappa^*}(u^{\mathcal{T}})| \leq \beta_2 |u^{\mathcal{T}}|_{1,\kappa^*}, \quad \forall u^{\mathcal{T}} \in \mathbb{R}^{\mathcal{T}}, \forall \kappa^* \in \mathcal{T}^*.$$

Proof :

Consider the quadratic form $q_{\kappa^*}(v) = (B_{\kappa^*}v, v) = |B_{\kappa^*}^{\frac{1}{2}}v|^2$. For any $v \in \mathbb{R}^4$, we can write $v = v_0 + v_0^\perp$ with $v_0 \in \mathbb{R}e_0$ and $v_0^\perp \in (\mathbb{R}e_0)^\perp$, and we have $q_{\kappa^*}(v) = q_{\kappa^*}(v_0^\perp)$. Consider the linear map $\Psi : \mathbb{R}^4 \to \mathbb{R}^4$, defined by:

$$\Psi : v \mapsto \bar{\eta} = (\bar{\eta}_i)_{i=1,2,3,4}, \quad \text{with} \quad \bar{\eta}_i = v_{i+1} - v_i,$$

where we recall that v_5 stands for v_1. The map Ψ is one-to-one from $(\mathbb{R}e_0)^\perp$ onto itself. Since $\Psi(v) = \Psi(v_0^\perp)$, we have $q_{\kappa^*}(v) = q_{\kappa^*}(\Psi^{-1}(\bar{\eta}))$. We immediately deduce that for any $v \in \mathbb{R}^4$,

$$\min Sp\left(B_{\kappa^* | (\mathbb{R}e_0)^\perp}\right) \frac{\|\bar{\eta}\|^2}{\|\Psi\|^2} \leq q_{\kappa^*}(v) \leq \max Sp\left(B_{\kappa^* | (\mathbb{R}e_0)^\perp}\right) \|\bar{\eta}\|^2 \|\Psi^{-1}\|^2,$$

where $\|.\|$ denotes the Euclidean norm on $\mathbb{R}^4$ but also the corresponding norm on the space of linear maps on $\mathbb{R}^4$. The matrix $B_{\kappa^* | \mathbb{R}e_0^\perp}$ is given by (2.20). Therefore, using assumption (2.27), we get

$$\frac{1}{\|\Psi\|^2} \min Sp\left(B_{\kappa^* | (\mathbb{R}e_0)^\perp}\right) \geq C \left\|\left(B_{\kappa^* | (\mathbb{R}e_0)^\perp}\right)^{-1}\right\|_\infty^{-1} \geq C \frac{2\xi^{\kappa^*} + \frac{m_1^{\kappa^*} m_3^{\kappa^*} + m_2^{\kappa^*} m_4^{\kappa^*}}{l_1^{\kappa^*} l_2^{\kappa^*}}}{(1 + \xi^{\kappa^*})} \frac{1}{m(\kappa^*)} \geq C_1 \gamma^2 \frac{1}{m(\kappa^*)},$$

and

$$\|\Psi^{-1}\|^2 \max Sp\left(B_{\kappa^* | (\mathbb{R}e_0)^\perp}\right) \leq C \left\|B_{\kappa^* | (\mathbb{R}e_0)^\perp}\right\|_\infty \leq C(1 + \xi^{\kappa^*}) \frac{1}{m(\kappa^*)} \leq C_2 \frac{1}{\gamma} \frac{1}{m(\kappa^*)},$$

where C_1 and C_2 only depend on the constant c_1 in (2.2). Thus for all $v \in \mathbb{R}^4$, we have

$$C_1(\gamma) \frac{1}{m(\kappa^*)} \|\bar{\eta}\|^2 \leq \left|B_{\kappa^*}^{\frac{1}{2}} v\right| \leq C_2(\gamma) \frac{1}{m(\kappa^*)} \|\bar{\eta}\|^2.$$

If we take $v = T_{\kappa^*}(u^{\mathcal{T}})$, we have $\bar{\eta}_i = l_i^{\kappa^*} \delta_i^{\kappa^*}(u^{\mathcal{T}})$. From (2.2) we have $c_5 m(\kappa^*) \leq |l_i^{\kappa^*}|^2 \leq c_6 m(\kappa^*)$, therefore the claim of the lemma follows. $\blacksquare$

We can now prove the monotony of the scheme, which is equivalent to the convexity of $J_{\mathcal{T}}$. We first recall the following inequalities whose proofs can be found for example in [5, 12]

Lemma 2.4 *For any $p > 1$ and $\delta \geq 0$, there exist C_1 and C_2 such that for any $n \geq 1$ and for any $(\eta, \xi) \in (\mathbb{R}^n)^2$, we have*

$$||\xi|^{p-2}\xi - |\eta|^{p-2}\eta| \leq C_1 |\xi - \eta|^{1-\delta} (|\xi| + |\eta|)^{p-2+\delta},$$

$$\left(|\xi|^{p-2}\xi - |\eta|^{p-2}\eta, \xi - \eta\right) \geq C_2 |\xi - \eta|^{2+\delta} (|\xi| + |\eta|)^{p-2-\delta}.$$

Using Lemma 2.3 and the previous result, one can obtain easily the coercivity of $J_{\mathcal{T}}$. The proof is formally similar to the one for the continuous problem that one can find for example in [6, 12].

13

Lemma 2.5 *Let $\mathcal{T}$ be a mesh satisfying the regularity assumption (2.2). Consider an admissible scheme as defined in Definition 2.2 and satisfying the uniform admissibility condition (2.27).*

There exists a constant $C > 0$ (depending only on the regularity constants of the mesh and on the parameter γ) such that for any elements $u^{\mathcal{T}}$ and $v^{\mathcal{T}}$ of $\mathbb{R}^{\mathcal{T}}$ we have the following inequality :

$$\left(\nabla J_{\mathcal{T}}(u^{\mathcal{T}}) - \nabla J_{\mathcal{T}}(v^{\mathcal{T}}), u^{\mathcal{T}} - v^{\mathcal{T}}\right) \geq C\|u^{\mathcal{T}} - v^{\mathcal{T}}\|_{1,p,\mathcal{T}}^{p}, \quad \text{if } p \geq 2, \tag{2.28}$$

$$\left(\nabla J_{\mathcal{T}}(u^{\mathcal{T}}) - \nabla J_{\mathcal{T}}(v^{\mathcal{T}}), u^{\mathcal{T}} - v^{\mathcal{T}}\right) \geq C\|u^{\mathcal{T}} - v^{\mathcal{T}}\|_{1,p,\mathcal{T}}^{2}\left(\|u^{\mathcal{T}}\|_{1,p,\mathcal{T}}^{p} + \|v^{\mathcal{T}}\|_{1,p,\mathcal{T}}^{p}\right)^{\frac{p-2}{p}}, \quad \text{if } 1 < p \leq 2. \tag{2.29}$$

Corollary 2.1 *Under the assumptions of the previous Lemma, there exists a constant $C > 0$ such that for any $u^{\mathcal{T}}, v^{\mathcal{T}} \in \mathbb{R}^{\mathcal{T}}$ we have*

$$J_{\mathcal{T}}(v^{\mathcal{T}}) - J_{\mathcal{T}}(u^{\mathcal{T}}) - (\nabla J_{\mathcal{T}}(u^{\mathcal{T}}), v^{\mathcal{T}} - u^{\mathcal{T}}) \geq C\|u^{\mathcal{T}} - v^{\mathcal{T}}\|_{1,p,\mathcal{T}}^{p}, \quad \text{if } p \geq 2,$$

and

$$J_{\mathcal{T}}(v^{\mathcal{T}}) - J_{\mathcal{T}}(u^{\mathcal{T}}) - (\nabla J_{\mathcal{T}}(u^{\mathcal{T}}), v^{\mathcal{T}} - u^{\mathcal{T}}) \geq C\|u^{\mathcal{T}} - v^{\mathcal{T}}\|_{1,p,\mathcal{T}}^{2}(\|u^{\mathcal{T}}\|_{1,p,\mathcal{T}} + \|v^{\mathcal{T}}\|_{1,p,\mathcal{T}})^{p-2}, \quad \text{if } 1 < p < 2.$$

We can now prove the fundamental theorem of this section.

Theorem 2.1 *Consider an admissible finite volume scheme for (1.1) on a mesh $\mathcal{T}$ of Ω, in the sense of Definition 2.2. Then the system (2.23) admits a unique solution $u^{\mathcal{T}}$, which is the unique minimizing point of the discrete energy $J_{\mathcal{T}}$.*

Proof :
Lemma 2.1, Lemma 2.5 and Corollary 2.1 ensure that the functional $J_{\mathcal{T}}$ is strictly convex and that $J_{\mathcal{T}}(u^{\mathcal{T}})$ tends to infinity as the norm of $u^{\mathcal{T}}$ tends to infinity. The claim follows immediately. ∎

This theorem shows that the approximate solution of the equation exists and is unique under the admissibility conditions. Furthermore, it shows that this unique solution can be obtained numerically, for instance, thanks to minimization algorithms such as nonlinear conjugate gradient methods (Polak-Ribière for instance). Hence, the schemes we are studying can be implemented in a straightforward way.

Remark 2.3

- *Assume that the family of meshes satisfies the uniform regularity assumption (2.2). Then, the scheme obtained with $\xi^{\kappa^*} = 0$, $\forall \kappa^* \in \mathcal{T}^*$, is always admissible and satisfies the uniform admissibility condition (2.27).*

- *For a uniform mesh, i.e. a mesh consisting of control volumes isometric to a reference control volume $]0, h_x[\times]0, h_y[$, assumption (2.27) becomes*

$$\frac{2}{\gamma} \geq 2\xi^{\kappa^*} \geq -\frac{1}{4}\left(\frac{h_y}{h_x} + \frac{h_x}{h_y}\right) + \gamma, \quad \forall \kappa^* \in \mathcal{T}^*.$$

2.5 Examples of schemes

As we have seen before, an admissible scheme is completely determined if we know, for each dual control volume κ^*, the quadratic form

$$q_{\kappa^*}(v) = \left|B_{\kappa^*}^{\frac{1}{2}}v\right|^2 = \frac{1}{m(\kappa^*)}\sum_{i=1}^{4}\tau_i^{\kappa^*}(v_{i+1} - v_i)^2 + \frac{2\xi^{\kappa^*}}{m(\kappa^*)}(v_1 + v_3 - v_2 - v_4)^2. \tag{2.30}$$

We have already seen that it is convenient to express this quadratic form in v as a quadratic form Q_{κ^*} of the quantities $\delta_i = \frac{v_{i+1} - v_i}{l_i^{\kappa^*}}$. Note that this representation is not unique since $\sum_i l_i^{\kappa^*}\delta_i = 0$. The representations we choose below are, in a sense, the canonical ones.

- As we have seen before, the scheme given by $\xi^{\kappa^*} = 0$ is always admissible. It corresponds to the case

$$Q_{\kappa^*}(\delta_1, \delta_2, \delta_3, \delta_4) = \frac{1}{m(\kappa^*)} \sum_{i=1}^{4} m_i^{\kappa^*} l_i^{\kappa^*} \delta_i^2 = \sum_{i=1}^{4} \frac{m_i^{\kappa^*}}{l_{i+1}^{\kappa^*}} \delta_i^2. \qquad (2.31)$$

Note that for $p = 2$, this scheme corresponds to the classical finite volume scheme for the laplacian operator (see for example [9]) on a cartesian rectangular mesh. Proposition (2.1) shows that (2.31) is the unique way to extend this scheme to the case $p \neq 2$ while preserving the symmetric structure of the equations. Let us emphasize that, if the mesh is not uniform, we cannot approach $|\nabla \overline{u}|^2$ using on each

$\kappa^* \in \mathcal{T}^*$, the quadratic form

$$Q_{\kappa^*}(\delta_1, \delta_2, \delta_3, \delta_4) = Q^0(\delta_1, \delta_2, \delta_3, \delta_4) = \frac{1}{2} \sum_{i=1}^{4} \delta_i^2, \qquad (2.32)$$

which seems to be a natural choice. Indeed, it is possible to use the form Q^0 in order to define a discrete energy $J_{\mathcal{T}}^0$ by the analogue of (2.26). This new functional is convex and coercive, so that there is a unique minimum satisfying $\nabla J_{\mathcal{T}}^0(u^{\mathcal{T}}) = 0$. Nevertheless, if we write down the Euler-Lagrange equation for $u^{\mathcal{T}}$, it is easy to see that the scheme is not a consistent approximation of the flux balance equation (2.4).

- If $p = 2$ and if $\xi^{\kappa^*} \neq 0$ satisfies (2.27), we obtain non classical nine points finite volume schemes for the Laplace equation. As we will see in the sequel of this paper, these schemes enjoy the same convergence properties that the classical five points scheme (see [9]).

- If the mesh is uniform, then we have $\tau_1^{\kappa^*} = \tau_3^{\kappa^*} = \frac{h_y}{2h_x}$ and $\tau_2^{\kappa^*} = \tau_4^{\kappa^*} = \frac{h_x}{2h_y}$, for any dual control volume. Hence, the general form of the schemes is given by

$$Q_{\kappa^*}(\delta) = \frac{1}{2} \sum_{i=1}^{4} \delta_i^2 + \xi^{\kappa^*} \left(\frac{h_x}{h_y}(\delta_1 + \delta_3)^2 + \frac{h_y}{h_x}(\delta_2 + \delta_4)^2 \right).$$

- In the case of a uniform mesh, consider the scheme obtained for $\mu_1^{\kappa^*} = \mu_2^{\kappa^*} = \mu_3^{\kappa^*} = \mu_4^{\kappa^*} = \frac{1}{2}$ in (2.14), corresponding to a centered approximation with respect to the dual control volumes of the normal derivative of $\overline{u}$ through the edges of the mesh. An easy computation shows that this scheme is the one obtained for $\xi^{\kappa^*} = -\frac{1}{8} \left(\frac{h_y}{h_x} + \frac{h_x}{h_y} \right)$, which is exactly the limiting case in the admissibility condition of Definition 2.2. Therefore, in this case the rank of the matrix B_{κ^*} is only equal to 2. Our analysis does not apply for this particular scheme. Nevertheless, some numerical computations seem to show that the approximate solutions obtained through this scheme converge towards the exact solution $\overline{u}$.

2.6 Proof of Proposition 2.1

In this section, we prove that the scheme defined by (2.10) and (2.11) is symmetric if and only if for each dual control volume κ^*, the matrix A_{κ^*} is proportional to the matrix B_{κ^*}.

It is quite clear that if the matrix A_{κ^*} is linked to the symmetric matrix B_{κ^*} by the relation $A_{\kappa^*} = \lambda_{\kappa^*} B_{\kappa^*}$, the differential of each dual contribution a_{κ^*} is symmetric and so far is the differential of a.

Conversely, suppose that the differential of a is symmetric: for any control volume κ and $\mathcal{L}$,

$$\frac{\partial a_{\kappa}}{\partial u_{\mathcal{L}}} = \frac{\partial a_{\mathcal{L}}}{\partial u_{\kappa}}. \qquad (2.33)$$

Let us fix some dual control volume κ^* and take two control volumes κ and $\mathcal{L}$ whose centers are vertices of κ^*. Let us notice that the following proofs are only given if $\kappa^* \subset \Omega$. The arguments can be adapted without
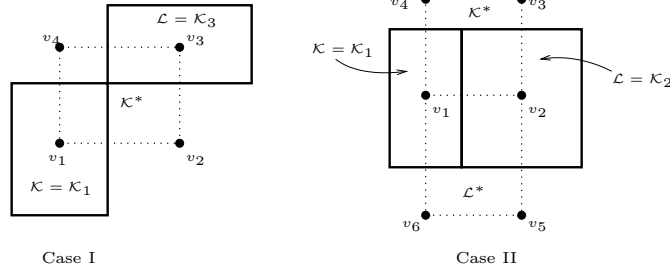
15

Figure 2.4: Notations

difficulty to the case of boundary dual control volumes by taking into account the values assigned to the ghost cells (see Figure 2.1).

Figure 2.4 exhibits the two different situations we have to deal with. In case I, the symmetry property (2.33) readily yields that

$$\frac{\partial a_{\mathcal{K},\mathcal{K}^*}}{\partial u_{\mathcal{L}}} - \frac{\partial a_{\mathcal{L},\mathcal{K}^*}}{\partial u_{\mathcal{K}}} = \frac{\partial a_{\mathcal{K}}}{\partial u_{\mathcal{L}}} - \frac{\partial a_{\mathcal{L}}}{\partial u_{\mathcal{K}}} = 0. \tag{2.34}$$

In case II, $x_{\mathcal{K}}, x_{\mathcal{L}}$ belong simultaneously to two dual control volumes $\mathcal{K}^*$ and $\mathcal{L}^*$, so that (2.33) gives

$$\frac{\partial a_{\mathcal{K},\mathcal{K}^*}}{\partial u_{\mathcal{L}}} + \frac{\partial a_{\mathcal{K},\mathcal{L}^*}}{\partial u_{\mathcal{L}}} = \frac{\partial a_{\mathcal{L},\mathcal{K}^*}}{\partial u_{\mathcal{K}}} + \frac{\partial a_{\mathcal{L},\mathcal{L}^*}}{\partial u_{\mathcal{K}}}$$

or, equivalently,

$$\frac{\partial a_{\mathcal{K},\mathcal{K}^*}}{\partial u_{\mathcal{L}}} - \frac{\partial a_{\mathcal{L},\mathcal{K}^*}}{\partial u_{\mathcal{K}}} = \frac{\partial a_{\mathcal{L},\mathcal{L}^*}}{\partial u_{\mathcal{K}}} - \frac{\partial a_{\mathcal{K},\mathcal{L}^*}}{\partial u_{\mathcal{L}}}. \tag{2.35}$$

We will show that both sides of (2.35) are in fact zero. In the sequel, we fix some dual control volume $\mathcal{K}^*$, so that we have

$$a_{\mathcal{K}^*}(v) = \frac{1}{m(\mathcal{K}^*)}(Bv,v)^{\frac{p-2}{2}}Av \in \mathbb{R}^4,$$

where we drop the $\mathcal{K}^*$ dependency of the matrices A and B, in order to simplify the notations. Let us denote by $(a_{i,j})_{1\leq i,j\leq 4}$ and $(b_{i,j})_{1\leq i,j\leq 4}$ the entries of the matrices A and B. We want to prove that, there exists $\lambda \in \mathbb{R}$ such that $A = \lambda B$.

Remark 2.4 *Following* (2.12), *the value of* (Bv,v) *is zero if and only if* v *is parallel to* $e_0 = (1,1,1,1)^t$. *Hence,* $a_{\mathcal{K}^*}$ *is a differentiable function except on the line* $\mathbb{R}e_0$ *if* $p < 4$. *If* $v \notin \mathbb{R}e_0$, *we have*

$$\begin{aligned} da_{\mathcal{K}^*}(v).w &= \frac{1}{m(\mathcal{K}^*)}(Bv,v)^{\frac{p-2}{2}}Aw + \frac{1}{m(\mathcal{K}^*)}(p-2)(Bv,v)^{\frac{p-4}{2}}(Bv,w)Av \\ &= \frac{1}{m(\mathcal{K}^*)}(Bv,v)^{\frac{p-4}{2}}\left((Bv,v)Aw + (p-2)(Av)\otimes(Bv)w\right). \end{aligned} \tag{2.36}$$

We refer to Figures 2.3 and 2.4 for the notations $(\kappa_i)_i$ corresponding to the dual control volume $\mathcal{K}^*$. For $i \neq j$, set

$$\rho_{i,j}(v) = \frac{\partial a_{\mathcal{K}_i,\mathcal{K}^*}(v)}{\partial v_j} - \frac{\partial a_{\mathcal{K}_j,\mathcal{K}^*}(v)}{\partial v_i}.$$

We recall that indices i, j are understood modulo 4.

Lemma 2.6 *We have:*

   1. For any $i \in \{1,\ldots,4\}$, *the function* $\rho_{i,i+2}$ *is identically zero.*

16

2. For any $i \in \{1, \ldots, 4\}$, the function $\rho_{i,i+1}$ only depends on the variables v_i, v_{i+1}.

Proof :
Without loss of generality, we can assume that $i = 1$. The equalities (2.34) taken for $\kappa = \kappa_1$, and $\mathcal{L} = \kappa_3$ immediately yield $\rho_{1,3} \equiv 0$. The equality (2.35) taken for $\kappa = \kappa_1$ and $\mathcal{L} = \kappa_2$ shows that $\rho_{1,2}$ is a function of the variables v_1, v_2, v_3, v_4 on the one hand, and of the variables v_1, v_2, v_5, v_6 on the other hand. Therefore, $\rho_{1,2}$ can only depend on v_1, v_2. ∎

We can now state the following Lemma which is essentially an homogeneity argument.

Lemma 2.7 *Suppose that $p \notin \{2, 4\}$. Let q be a quadratic form defined on $\mathbb{R}^4$, B a symmetric non-negative matrix, whose kernel is spanned by e_0, and ρ the function defined on $\mathbb{R}^4$ by*

$$\rho(v) = (Bv, v)^{\frac{p-4}{2}} q(v).$$

Assume that ρ only depends on the variables (v_1, v_2). Then, ρ is identically zero.

Proof :
Let us fix v_1 and v_2. As B is a definite positive matrix on $\{e_0\}^\perp$, its entry $b_{3,3}$ is positive. Moreover we can consider $q(v)$ as a quadratic polynomial in the variable v_3:

$$q(v) = \alpha_2 v_3^2 + \alpha_1 v_3 + \alpha_0,$$

where $\alpha_2, \alpha_1, \alpha_0$ may depend on v_1, v_2, v_4.

- Suppose that $\alpha_2 \neq 0$, we deduce that

$$\lim_{|v_3| \to +\infty} \rho(v) \left(b_{3,3}^{\frac{p-4}{2}} \alpha_2 |v_3|^{p-2} \right)^{-1} = 1,$$

which is impossible since $p \neq 2$ and ρ does not depend on v_3. Thus, $\alpha_2 = 0$.

- Now if $\alpha_1 \neq 0$ and $p \neq 3$, we get

$$\lim_{|v_3| \to +\infty} \rho(v) \left(b_{3,3}^{\frac{p-4}{2}} \alpha_1 |v_3|^{p-3} \right)^{-1} = 1.$$

This is also a contradiction. Thus, $\alpha_1 = 0$ unless $p = 3$.

- Now if $\alpha_0 \neq 0$ and $p \neq 3$, we have $\alpha_1 = 0$ and then

$$\lim_{|v_3| \to +\infty} \rho(v) \left(b_{3,3}^{\frac{p-4}{2}} \alpha_0 |v_3|^{p-4} \right)^{-1} = 1.$$

Since $p \neq 4$, we get another contradiction.

Hence, if $p \neq 3$, we have proved that q is identically zero, which gives the claim.

It remains to treat the case where $p = 3$ and $\alpha_1 \neq 0$. In this case, we have

$$\rho(v) = (Bv, v)^{-\frac{1}{2}} q(v) = \psi(v_1, v_2)$$

or, equivalently,

$$q(v) = \psi(v_1, v_2)(Bv, v)^{\frac{1}{2}}.$$

As $\alpha_2 = 0$ and $\alpha_1 \neq 0$, the quadratic form q is linear with respect to the variable v_3.

Let us consider any $v_1 \neq v_2$, then for any choice of v_3, v_4, we have $(Bv, v) \neq 0$ (because the kernel of B is $\mathbb{R}e_0$). Moreover, we can choose $v_4 = 0$ and then, as $\alpha_2 = 0$ and $\alpha_1 \neq 0$, there exists a value of v_3 (depending on v_1, v_2, v_4) such that $q(v) = 0$. We immediately deduce that $\psi(v_1, v_2) = 0$. As ψ is a continuous function, the result still holds for $v_1 = v_2$. Hence, for $p = 3$, ρ vanishes identically. ∎

Let us now prove the following result:

Lemma 2.8 *Let $i, j \in \{1, 2, 3, 4\}$, $i \neq j$. Suppose that $\rho_{i,j}$ is identically zero.*

- *If $p = 2$, we have $a_{i,j} = a_{j,i}$.*

- *If $p \neq 2$, there exists a real number $\lambda_{i,j}$ such that*

$$a_{i,i} = \lambda_{i,j} b_{i,i}, \quad a_{j,j} = \lambda_{i,j} b_{j,j}, \quad a_{i,j} = a_{j,i} = \lambda_{i,j} b_{i,j} \tag{2.37}$$

Proof :

The assertion $\rho_{i,j} \equiv 0$ is equivalent to

$$\frac{\partial a_{\kappa_i, \kappa^*}}{\partial v_j}(v) = \frac{\partial a_{\kappa_j, \kappa^*}}{\partial v_i}(v), \quad \forall v \in \mathbb{R}^4.$$

Hence, from (2.36) we deduce that the quadratic form defined on $\mathbb{R}^4$ by

$$q_{ij}(v) = (Bv, v)(a_{i,j} - a_{j,i}) + (p - 2)((Av)_i (Bv)_j - (Bv)_i (Av)_j), \quad \forall v \in \mathbb{R}^4$$

is identically zero. This implies that

$$(Bv, w)(a_{i,j} - a_{j,i}) + \frac{p-2}{2}((Av)_i (Bw)_j + (Aw)_i (Bv)_j - (Bv)_i (Aw)_j - (Bw)_i (Av)_j) = 0, \quad \forall v, w \in \mathbb{R}^4.$$

In particular, using the symmetry of B, for (v, w) equal respectively to (e_i, e_i), (e_j, e_j) and (e_i, e_j), we get the following relations:

$$b_{i,i}(a_{i,j} - a_{j,i}) + (p - 2)(a_{i,i} b_{i,j} - b_{i,i} a_{j,i}) = 0,$$

$$b_{j,j}(a_{i,j} - a_{j,i}) + (p - 2)(a_{i,j} b_{j,j} - b_{i,j} a_{j,j}) = 0,$$

$$\frac{p}{2} b_{i,j}(a_{i,j} - a_{j,i}) + \frac{p-2}{2}(a_{i,i} b_{j,j} - b_{i,i} a_{j,j}) = 0.$$

These equations can be written as an underdetermined linear system:

$$\begin{pmatrix} (p-2)b_{i,j} & 0 & b_{i,i} & -(p-1)b_{i,i} \\ 0 & -(p-2)b_{i,j} & (p-1)b_{j,j} & -b_{j,j} \\ \frac{p-2}{2}b_{j,j} & -\frac{p-2}{2}b_{i,i} & \frac{p}{2}b_{i,j} & -\frac{p}{2}b_{i,j} \end{pmatrix} \begin{pmatrix} a_{i,i} \\ a_{j,j} \\ a_{i,j} \\ a_{j,i} \end{pmatrix} = 0. \tag{2.38}$$

Let us extract a 3×3 matrix, whose determinant is given by

$$\begin{vmatrix} 0 & b_{i,i} & -(p-1)b_{i,i} \\ (2-p)b_{i,j} & (p-1)b_{j,j} & -b_{j,j} \\ -\frac{p-2}{2}b_{i,i} & \frac{p}{2}b_{i,j} & -\frac{p}{2}b_{i,j} \end{vmatrix} = (p-2)^2 \frac{p}{2} b_{i,i}(b_{i,j}^2 - b_{i,i} b_{j,j}). \tag{2.39}$$

As the kernel of B is reduced to $\mathbb{R}e_0$, the restriction of the quadratic form $v \mapsto (Bv, v)$ to the plane defined by e_i et e_j is positive. Then we obtain

$$b_{i,i} > 0, \ b_{j,j} > 0, \quad \text{and} \quad (b_{i,i} b_{j,j} - b_{i,j}^2) > 0.$$

Hence, the 3×3 determinant (2.39) is non zero and it follows that the system (2.38) admits a one dimensional vector space of solutions. The vector $(b_{i,i}, b_{j,j}, b_{i,j}, b_{i,j})^t$ is an obvious non zero solution of the system. Therefore, there exists a real $\lambda_{i,j}$ such that

$$a_{i,i} = \lambda_{i,j} b_{i,i}, \quad a_{j,j} = \lambda_{i,j} b_{j,j}, \quad a_{i,j} = a_{j,i} = \lambda_{i,j} b_{i,j}.$$

∎

Collecting all the previous facts, we get the claim of Proposition 2.1 for the case $p \notin \{2,4\}$. Indeed, Lemmas 2.6 and 2.7 show that for any $i \neq j$, we have $\rho_{i,j} = 0$. Hence by Lemma 2.8, we deduce that equations (2.37) must be satisfied for any (i,j), $i \neq j$. Consequently, the real numbers $\lambda_{i,j}$ are independent of i and j. Therefore, equations (2.37) imply that there exists $\lambda \in \mathbb{R}$ such that

$$A = \lambda B.$$

In order to conclude the proof of Proposition 2.1, we have to consider separately the cases $p = 2$ and $p = 4$.

Lemma 2.9 *For any $p > 1$, and any $\kappa^* \in \mathcal{T}^*$, the matrix A_{κ^*} is symmetric.*

Proof :
For $p \notin \{2,4\}$, we just have proved that $A_{\kappa^*} = \lambda_{\kappa^*} B_{\kappa^*}$. The claim follows since B_{κ^*} is symmetric.

For $p = 2$ and $p = 4$, Lemmas 2.6 and 2.8 show, in particular, that $a_{1,3} = a_{3,1}$ and $a_{2,4} = a_{4,2}$. These two equalities together with the general expression for the matrix A_{κ^*} given by (2.16), imply that $\tau_1(1 - \mu_1) = \tau_3(1 - \mu_3)$ and $\tau_2(1 - \mu_2) = \tau_4(1 - \mu_4)$. It is easily seen that these relations imply the symmetry of the matrix A_{κ^*}. ∎

Note that in the case $p \notin \{2,4\}$ we do not use the specific expression (2.16) for the matrix A_{κ^*}, whereas this expression is useful when $p \in \{2,4\}$. In fact, we believe that this particular form of the matrix is not necessary to conclude even when $p = 4$, but the proof seems to be heavier.

Lemma 2.9 shows Proposition 2.1 in the case $p = 2$. Now suppose $p = 4$. In this case, Lemma 2.9 shows that A_{κ^*} is symmetric and Lemmas 2.6 and 2.8 show that relations (2.37) are valid for $j = i + 2$. Now using the fact that for $i = 1, 2, 3, 4$, $\rho_{i,i+1}$ does not depend on v_{i+2}, v_{i+3}, we get in particular

$$\frac{\partial^2 \rho_{i,i+1}}{\partial^2 v_{i+2}} = 0, \frac{\partial^2 \rho_{i,i+1}}{\partial v_{i+1} \partial v_{i+3}} = 0 \text{ and } \frac{\partial^2 \rho_{i,i+1}}{\partial^2 v_{i+3}} = 0, \frac{\partial^2 \rho_{i,i+1}}{\partial v_i \partial v_{i+2}} = 0.$$

Consider the case $i = 1$. The previous relations are equivalent to

$$\begin{cases} (a_{4,1} - \lambda_{2,4} b_{4,1}) b_{4,2} = 0, \\ (a_{2,1} - \lambda_{2,4} b_{2,1}) b_{4,2} + (a_{4,1} - \lambda_{2,4} b_{4,1}) b_{2,2} = 0, \end{cases}$$

and

$$\begin{cases} (a_{3,2} - \lambda_{1,3} b_{3,2}) b_{3,1} = 0, \\ (a_{2,1} - \lambda_{1,3} b_{2,1}) b_{3,1} + (a_{3,2} - \lambda_{1,3} b_{3,2}) b_{1,1} = 0. \end{cases}$$

Combining these equations, and using the fact that $b_{1,1} > 0$ and $b_{2,2} > 0$, we deduce that

$$a_{4,1} - \lambda_{2,4} b_{4,1} = 0 \text{ and } a_{3,2} - \lambda_{1,3} b_{3,2} = 0. \tag{2.40}$$

For $i = 2, 3, 4$, we get in a similar way

$$\begin{array}{ll} a_{2,1} - \lambda_{1,3} b_{2,1} = 0, & a_{4,3} - \lambda_{2,4} b_{4,3} = 0, \\ a_{4,1} - \lambda_{1,3} b_{4,1} = 0, & a_{3,2} - \lambda_{2,4} b_{3,2} = 0, \\ a_{4,3} - \lambda_{1,3} b_{4,3} = 0, & a_{2,1} - \lambda_{2,4} b_{2,1} = 0. \end{array} \tag{2.41}$$

First of all, it is impossible that $b_{1,2} = b_{2,3} = b_{1,4} = b_{3,4} = 0$, because in that case the vector $\tilde{e}_0 = (0,1,0,1)^t$ would have been in the kernel of B which is spanned by $e_0 = (1,1,1,1)^t$ by assumption (2.12).

Hence using relations (2.40) and (2.41), we see that we have $\lambda_{1,3} = \lambda_{2,4} \overset{\text{def}}{=} \lambda$, so that we can conclude that $A_{\kappa^*} = \lambda_{1,3} B_{\kappa^*}$.

3 Error estimates

Let f be a given data in $L^{p'}(\Omega)$, so that there exists a unique solution $\overline{u}$ of (1.1) in $W_0^{1,p}(\Omega)$. Throughout this section, we consider an admissible scheme, in the sense of Definition 2.2, on a finite volume mesh $\mathcal{T}$ on Ω. We recall that h is the maximum of the diameters of the control volumes in $\mathcal{T}$ and we assume that (2.2) and (2.27) hold. Let $u^{\mathcal{T}}$ be the unique solution to (2.23). In the sequel, we denote by C a generic constant that depends only on Ω, p, c_1 in (2.2) and γ in (2.27).

Our main result is the following.

Theorem 3.1 *Assume that the solution $\overline{u}$ of problem* (1.1) *belongs to $W_0^{1,p}(\Omega) \cap W^{2,p}(\Omega)$, then we have*

$$\begin{cases} \|\overline{u}^{\mathcal{T}} - u^{\mathcal{T}}\|_{1,p,\mathcal{T}} \le Ch^{p-1}\|\overline{u}\|_{W^{2,p}}^{p-1}\|f\|_{L^{p'}}^{\frac{2-p}{p-1}}, & \text{if } p \le 2, \\[2mm] \|\overline{u}^{\mathcal{T}} - u^{\mathcal{T}}\|_{1,p,\mathcal{T}} \le Ch\|\overline{u}\|_{W^{2,p}} + Ch^{\frac{1}{p-1}}\|\overline{u}\|_{W^{2,p}}^{\frac{3p-4}{p(p-1)}}\|f\|_{L^{p'}}^{\frac{(p-2)^2}{p(p-1)^2}}, & \text{if } p > 2. \end{cases}$$

Recall that J.W. Barrett and W.B. Liu proved in [4], that if $f \in L^q(\Omega)$ and if $1 < p \le 2$ then $\overline{u}$ belongs to $H^2(\Omega)$ and then to $W^{2,p}(\Omega)$, so that the assumption in the previous theorem is fulfilled.

On the other hand when $p > 2$, there exists solutions of (1.1) with $f \in L^{p'}(\Omega)$ which are not in $W^{2,p}(\Omega)$. We will address the problem of proving error estimates for these solutions in [2].

Remark 3.1 *Using the Poincaré inequality (Lemma 2.1), we also deduce error estimates for the $L^p(\Omega)$ norm of order h^{p-1} when $p < 2$ and $h^{\frac{1}{p-1}}$ when $p \ge 2$.*

3.1 *A priori* estimates on the approximate solution

Lemma 3.1 *Under the previous assumptions, the approximate solution $u^{\mathcal{T}}$ satisfies the estimate*

$$\|u^{\mathcal{T}}\|_{1,p,\mathcal{T}} \le C\|f\|_{L^{p'}}^{\frac{1}{p-1}}. \tag{3.1}$$

Furthermore, if $u_1^{\mathcal{T}}$ and $u_2^{\mathcal{T}}$ are the solutions corresponding respectively to the data $f_1, f_2 \in L^{p'}(\Omega)$ then we have

$$\begin{cases} \|u_1^{\mathcal{T}} - u_2^{\mathcal{T}}\|_{1,p,\mathcal{T}} \le C\|f_1 - f_2\|_{L^{p'}}^{\frac{1}{p-1}}, & \text{if } p \ge 2, \\[2mm] \|u_1^{\mathcal{T}} - u_2^{\mathcal{T}}\|_{1,p,\mathcal{T}} \le C\|f_1 - f_2\|_{L^{p'}}(\|u_1^{\mathcal{T}}\|_{1,p,\mathcal{T}}^{2-p} + \|u_2^{\mathcal{T}}\|_{1,p,\mathcal{T}}^{2-p}), & \text{if } 1 < p < 2. \end{cases} \tag{3.2}$$

Proof :

1. By Lemmas 2.2 and 2.3, we have

$$0 = \left(\nabla J_{\mathcal{T}}(u^{\mathcal{T}}), u^{\mathcal{T}}\right) = \sum_{\kappa \in \mathcal{T}} a_\kappa(u^{\mathcal{T}})u_\kappa - m(\kappa)f_\kappa u_\kappa \ge C\|u^{\mathcal{T}}\|_{1,p,\mathcal{T}}^p - \sum_{\kappa \in \mathcal{T}} m(\kappa)f_\kappa u_\kappa.$$

It follows by the Hölder inequality that

$$\|u^{\mathcal{T}}\|_{1,p,\mathcal{T}}^p \le C\sum_{\kappa \in \mathcal{T}} m(\kappa)|f_\kappa||u_\kappa| \le C\left(\sum_{\kappa \in \mathcal{T}} m(\kappa)|f_\kappa|^{\frac{p}{p-1}}\right)^{\frac{p-1}{p}}\left(\sum_{\kappa \in \mathcal{T}} m(\kappa)|u_\kappa|^p\right)^{\frac{1}{p}}.$$

Now, the discrete Poincaré inequality yields

$$\|u^{\mathcal{T}}\|_{1,p,\mathcal{T}}^{p-1} \le C\left(\sum_{\kappa \in \mathcal{T}} m(\kappa)|f_\kappa|^{\frac{p}{p-1}}\right)^{\frac{p-1}{p}}. \tag{3.3}$$

20

Using the definition of f_κ given by (2.5), by the Jensen inequality we have

$$|f_\kappa|^{\frac{p}{p-1}} \le \left(\frac{1}{m(\kappa)} \int_\kappa |f(z)| \, dz \right)^{\frac{p}{p-1}} \le \frac{1}{m(\kappa)} \int_\kappa |f(z)|^{\frac{p}{p-1}} \, dz,$$

so that (3.3) gives (3.1).

2. In order to prove (3.2), let us substract the discrete equations for $u_1^{\mathcal{T}}$ and $u_2^{\mathcal{T}}$:

$$a_\kappa(u_1^{\mathcal{T}}) - a_\kappa(u_2^{\mathcal{T}}) = m(\kappa)((f_1)_\kappa - (f_2)_\kappa), \quad \forall \kappa \in \mathcal{T}.$$

Multiplying by $(u_1^{\mathcal{T}})_\kappa - (u_2^{\mathcal{T}})_\kappa$ and summing over κ, we deduce, using (2.28) and (2.29), that

$$\|u_1^{\mathcal{T}} - u_2^{\mathcal{T}}\|_{1,p,\mathcal{T}}^p \le C \|u_1^{\mathcal{T}} - u_2^{\mathcal{T}}\|_{L^p} \|f_1 - f_2\|_{L^{p'}}, \quad \text{if } p \ge 2,$$

$$\|u_1^{\mathcal{T}} - u_2^{\mathcal{T}}\|_{1,p,\mathcal{T}}^2 (\|u_1^{\mathcal{T}}\|_{1,p,\mathcal{T}}^{p-2} + \|u_2^{\mathcal{T}}\|_{1,p,\mathcal{T}}^{p-2}) \le C \|u_1^{\mathcal{T}} - u_2^{\mathcal{T}}\|_{L^p} \|f_1 - f_2\|_{L^{p'}}, \quad \text{if } 1 < p < 2.$$

We conclude using Lemma 2.1.

∎

Remark 3.2 *This result gives the Hölder continuity of the inverse discrete p-laplacian operator. As we have already seen, the approximate solution $u^{\mathcal{T}}$ of problem (1.1) is obtained as the minimum of the discrete energy $J_{\mathcal{T}}$. Hence, the practical computation of the solution can be performed using a classical iterative method. The appropriate stopping criterion is given by the previous Lemma.*

Indeed, consider for instance, the case where $p \ge 2$. If we denote by $u^{\mathcal{T}}$ the exact approximate solution and $u_k^{\mathcal{T}}$ an iteration of any numerical algorithm used to solve the scheme, then the error between $u^{\mathcal{T}}$ and $u_k^{\mathcal{T}}$ is estimated by

$$\|u^{\mathcal{T}} - u_k^{\mathcal{T}}\|_{1,p,\mathcal{T}} \le C \|f - a(u_k^{\mathcal{T}})\|_{L^{p'}}^{\frac{1}{p-1}}.$$

Consequently, if we know that the numerical scheme has an order h^s, we can choose the stopping criterion

$$\|f - a(u_k^{\mathcal{T}})\|_{L^{p'}} \le C h^{s(p-1)},$$

so that finally

$$\|\overline{u}^{\mathcal{T}} - u_k^{\mathcal{T}}\|_{1,p,\mathcal{T}} \le \|\overline{u}^{\mathcal{T}} - u^{\mathcal{T}}\|_{1,p,\mathcal{T}} + \|u^{\mathcal{T}} - u_k^{\mathcal{T}}\|_{1,p,\mathcal{T}} \le C_1 h^s + (C_2 h^{s(p-1)})^{\frac{1}{p-1}} \le C h^s.$$

3.2 Consistency error of the scheme

Definition 3.1 (Local consistency error) *For any control volume κ and for any half-edge $\sigma \subset \partial\kappa$, we define the local consistency error by*

$$R_{\sigma,\kappa} = a_{\kappa,\sigma}(\overline{u}^{\mathcal{T}}) - \left(-\int_\sigma |\nabla \overline{u}|^{p-2} \nabla \overline{u} \cdot \nu_\kappa \, ds \right),$$

where

$$\begin{cases} a_{\kappa,\sigma}(\overline{u}^{\mathcal{T}}) = q_{\kappa^*}(T_{\kappa^*}(\overline{u}^{\mathcal{T}}))^{\frac{p-2}{2}} A_{\kappa,\kappa^*}^h(T_{\kappa^*}(\overline{u}^{\mathcal{T}})) & \text{if } \sigma = \sigma_{\kappa,\kappa^*}^h \\ a_{\kappa,\sigma}(\overline{u}^{\mathcal{T}}) = q_{\kappa^*}(T_{\kappa^*}(\overline{u}^{\mathcal{T}}))^{\frac{p-2}{2}} A_{\kappa,\kappa^*}^v(T_{\kappa^*}(\overline{u}^{\mathcal{T}})) & \text{if } \sigma = \sigma_{\kappa,\kappa^*}^v. \end{cases}$$

Remark that $a_{\kappa,\sigma}(\overline{u}^{\mathcal{T}})$ can be written (in a non-unique way) as a function of $\delta^{\kappa^*}(\overline{u}^{\mathcal{T}}) = (\delta_i^{\kappa^*}(\overline{u}^{\mathcal{T}}))_{i=1,2,3,4}$. Let us define g_σ to be a function such that

$$m(\sigma) g_\sigma(\delta^{\kappa^*}(v^{\mathcal{T}})) \overset{\text{def}}{=} a_{\kappa,\sigma}(v^{\mathcal{T}}), \quad \forall v^{\mathcal{T}} \in \mathbb{R}^{\mathcal{T}}.$$

For example, suppose that $\sigma = \sigma^v_{\kappa_1, \kappa^*}$ so that with (2.13) we have $a_{\kappa, \sigma} = q_{\kappa^*}^{\frac{p-2}{2}} A_{1,\kappa^*}$. In this case, following (2.14), (2.18) and (2.30), we can take

$$g_\sigma(\delta_1, \delta_2, \delta_3, \delta_4) = \left(\frac{1}{m(\kappa^*)} \sum_{i=1}^{4} m_i^{\kappa^*} l_i^{\kappa^*} \delta_i^2 + \frac{2\xi^{\kappa^*}}{m(\kappa^*)} (\delta_1 + \delta_3)^2 \right)^{\frac{p-2}{2}} \left(-\delta_1 - \xi^{\kappa^*} \frac{l_1^{\kappa^*}}{m_1^{\kappa^*}} (\delta_1 + \delta_3) \right). \tag{3.4}$$

Hence the local consistency error can be written as follows:

$$R_{\sigma, \kappa} = m(\sigma) g_\sigma(\delta^{\kappa^*}(\overline{u}^{\mathcal{T}})) + \int_\sigma |\nabla \overline{u}|^{p-2} \nabla \overline{u} \cdot \nu_\kappa \, ds. \tag{3.5}$$

Lemma 3.2 *Let κ^* be a dual control volume and $\sigma \subset \kappa^*$. For any point $s \in \sigma$, let us write $\delta^{\kappa^*}(\overline{u}^{\mathcal{T}}) = \delta^0(s) + \varepsilon(s)$, where $\delta^0(s) = (\delta_i^0)_{i=1,2,3,4}$ is defined by $\delta_i^0(s) = \nabla \overline{u}(s) \cdot \nu_i$.*

- *If $\kappa^* \subset \Omega$, we have*

$$\varepsilon_i(s) = \frac{1}{l_i} \int_0^1 (1-t) D^2 \overline{u}(tx_{i+1} + (1-t)s) \cdot (x_{i+1} - s)^2 \, dt - \frac{1}{l_i} \int_0^1 (1-t) D^2 \overline{u}(tx_i + (1-t)s) \cdot (x_i - s)^2 \, dt.$$

- *If κ^* is a boundary dual control volume, we have*

$$\varepsilon_i(s) = s_i^1 \frac{1}{l_i} \int_0^1 (1-t) D^2 \overline{u}(ty_{i+1} + (1-t)s) \cdot (y_{i+1} - s)^2 \, dt + s_i^2 \frac{1}{l_i} \int_0^1 (1-t) D^2 \overline{u}(ty_i + (1-t)s) \cdot (y_i - s)^2 \, dt,$$

where y_i denotes either the point x_i if $x_i \in \Omega$, either the point symmetric of x_i with respect to $\partial\Omega$, if $x_i \notin \Omega$, and $s_i^1, s_i^2 \in \{-1, 1\}$.

Proof :
The result is a straightforward consequence of the Taylor expansion :

$$\overline{u}_i - \overline{u}(s) = \overline{u}(x_i) - \overline{u}(s) = \nabla \overline{u}(y) \cdot (x_i - s) + \int_0^1 (1-t) D^2 \overline{u}(tx_i + (1-t)s) \cdot (x_i - s)^2 \, dt.$$

When κ^* is a boundary dual control volume, it suffices to take into account the values imposed in the ghost cells (see Figure 2.1). ∎

Lemma 3.3 *There exists a constant $C > 0$ such that for any $\sigma \subset \kappa^* \cap \overline{\Omega}$, we have*

$$\int_\sigma |\varepsilon(s)|^p \, ds \le C h^{p-1} \int_{\kappa^* \cap \Omega} |D^2 \overline{u}(z)|^p \, dz.$$

Proof :
Using the previous Lemma, we see that the estimate of $\int_\sigma |\varepsilon(s)|^p \, ds$ requires the control of the terms:

$$A_i \overset{\text{def}}{=} \int_\sigma \left| \frac{1}{l_i} \int_0^1 (1-t) D^2 \overline{u}(tx_i + (1-t)s) \cdot (x_i - s)^2 \, dt \right|^p \, ds, \quad i = 1, \dots, 4.$$

Thanks to the regularity of the mesh, using the change of variables $(t, s) \mapsto z = tx_i + (1-t)s$, we get:

$$A_i \le \int_\sigma \frac{1}{l_i^p} \int_0^1 \left| (1-t) D^2 \overline{u}(tx_i + (1-t)s) \cdot (x_i - s)^2 \right|^p \, dt \, ds \le C h^{p-1} \int_{\Sigma_{\sigma, i}} |D^2 \overline{u}(z)|^p dz,$$

where the domains $\Sigma_{\sigma,1}, \Sigma_{\sigma,2}, \Sigma_{\sigma,3}, \Sigma_{\sigma,4}$ are introduced in Figure 3.1. Note that the Jacobian determinant of the change of variables is $C(1-t)h$.

The lemma follows immediately.

∎

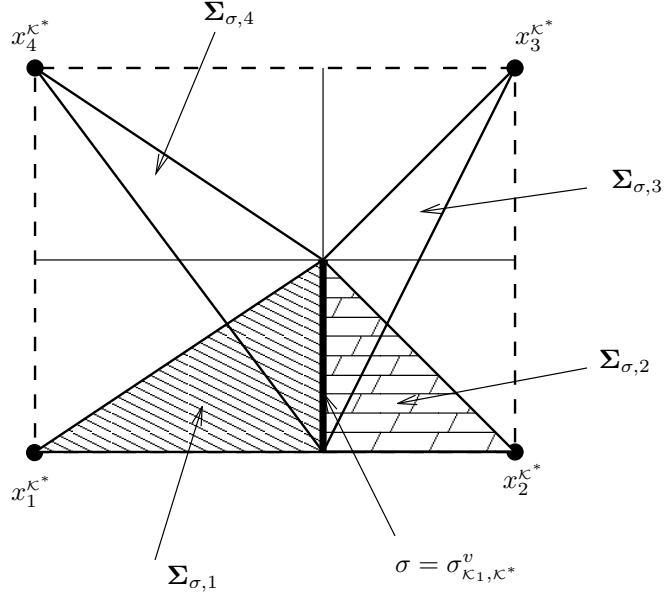We can now estimate the local consistency error of the scheme.

Figure 3.1: Definition of the domains $\boldsymbol{\Sigma}_{\sigma,1}, \boldsymbol{\Sigma}_{\sigma,2}, \boldsymbol{\Sigma}_{\sigma,3}$ and $\boldsymbol{\Sigma}_{\sigma,4}$

Lemma 3.4 *For any control volume κ, and for any half-edge $\sigma \subset \partial\kappa$, we have*

$$|R_{\sigma,\kappa}| \leq Ch^{\frac{(p-1)^2+1}{p}} \left(\int_{\kappa^* \cap \Omega} |D^2\overline{u}(z)|^p \, dz \right)^{\frac{p-1}{p}}, \quad \text{if } p \leq 2,$$

$$|R_{\sigma,\kappa}| \leq C \left(h^{\frac{(p-1)^2+1}{p}} \left(\int_{\kappa^* \cap \Omega} |D^2\overline{u}(z)|^p \, dz \right)^{\frac{p-1}{p}} + h^{\frac{2(p-1)}{p}} \|\nabla\overline{u}\|_{L^\infty(\Omega)}^{p-2} \left(\int_{\kappa^* \cap \Omega} |D^2\overline{u}(z)|^p \, dz \right)^{\frac{1}{p}} \right), \quad \text{if } p > 2,$$

where κ^* is the dual control volume which contains σ.

Proof :

Without loss of generality, we still assume that $\sigma = \sigma^v_{\kappa_1,\kappa^*}$ is the vertical half-edge in κ^* between $\kappa_1^{\kappa^*}$ and $\kappa_2^{\kappa^*}$. As in Lemma 3.2, we write $\delta^{\kappa^*}(\overline{u}^{\mathcal{T}}) = \delta^0(s) + \varepsilon(s)$. Let us first prove that

$$R_{\sigma,\kappa} = \int_\sigma \int_0^1 \nabla g_\sigma(\delta^0(s) + t\varepsilon(s)) \cdot \varepsilon(s) \, dt \, ds. \tag{3.6}$$

Indeed, for all $s \in \sigma$,

$$g_\sigma(\delta^{\kappa^*}(\overline{u}^{\mathcal{T}})) = g_\sigma(\delta^0(s) + \varepsilon(s)) = g_\sigma(\delta^0(s)) + \int_0^1 \nabla g_\sigma(\delta^0(s) + t\varepsilon(s)) \cdot \varepsilon(s) \, dt.$$

Integrating this relation over the edge σ and using the fact that $\delta^{\kappa^*}(\overline{u}^{\mathcal{T}})$ does not depend on s, we have

$$m(\sigma)g_\sigma(\delta^{\kappa^*}(\overline{u}^{\mathcal{T}})) = \int_\sigma g_\sigma(\delta^0(s)) \, ds + \int_\sigma \int_0^1 \nabla g(\delta^0(s) + t\varepsilon(s)) \cdot \varepsilon(s) \, dt \, ds.$$

Hence, (3.6) follows from (3.5).

23

We remark that, thanks to (3.4), for any $\alpha, \beta \in \mathbb{R}$ we have $g_\sigma(\alpha, \beta, -\alpha, -\beta) = -\left(\alpha^2 + \beta^2\right)^{\frac{p-2}{2}} \alpha$ (this amounts to the consistency of the scheme on piecewise affine functions). Hence, since we have $\delta^0(s) = (\overline{u}_x(s), \overline{u}_y(s), -\overline{u}_x(s), -\overline{u}_y(s))$, we deduce that

$$\int_\sigma g_\sigma(\delta^0(s))\, ds = -\int_\sigma |\nabla \overline{u}|^{p-2} \nabla \overline{u} \cdot \nu_\kappa\, ds,$$

which proves (3.6).

According that for any $\eta_1, \eta_2 \in \mathbb{R}^4$, we have $|\nabla g_\sigma(\eta_1) \cdot \eta_2| \leq C|\eta_1|^{p-2}|\eta_2|$, we obtain

$$|R_{\sigma,\kappa}| \leq C \int_\sigma |\varepsilon(s)| \left(\int_0^1 |\delta^0(s) + t\varepsilon(s)|^{p-2}\, dt\right) ds. \tag{3.7}$$

Remark that, for any $1 < p \leq 2$, there exists a constant C such that, for any $n \geq 0$ and for any $(\eta, \xi) \in (\mathbb{R}^n)^2 \backslash \{(0,0)\}$, we have

$$\int_0^1 |\eta||\xi + t\eta|^{p-2}\, dt \leq C|\eta|^{p-1}.$$

Therefore, we deduce from (3.7) that

$$|R_{\sigma,\kappa}| \leq C \int_\sigma |\varepsilon(s)| \left(|\delta^0(s)|^{p-2} + |\varepsilon(s)|^{p-2}\right) ds, \quad \text{if } p > 2,$$

$$|R_{\sigma,\kappa}| \leq C \int_\sigma |\varepsilon(s)|^{p-1}\, ds, \quad \text{if } p \leq 2.$$

Hence we readily get

$$|R_{\sigma,\kappa}| \leq Cm(\sigma)^{\frac{1}{p}} \left(\int_\sigma |\varepsilon(s)|^p\, ds\right)^{\frac{p-1}{p}}, \quad \text{if } 1 < p \leq 2.$$

For $p > 2$, note that $W^{1,p}(\Omega) \subset L^\infty(\Omega)$, so that we have $|\delta^0(s)| \leq \|\nabla \overline{u}\|_{L^\infty}$, and finally

$$|R_{\sigma,\kappa}| \leq C \left(m(\sigma)^{\frac{p-1}{p}} \|\nabla \overline{u}\|_{L^\infty(\Omega)}^{p-2} \left(\int_\sigma |\varepsilon(s)|^p\, ds\right)^{\frac{1}{p}} + m(\sigma)^{\frac{1}{p}} \left(\int_\sigma |\varepsilon(s)|^p\, ds\right)^{\frac{p-1}{p}}\right), \quad \text{if } p > 2.$$

We conclude using Lemma 3.3. ∎

3.3 Proof of Theorem 3.1

Using the estimate of the local consistency error obtained in Lemma 3.4, we can now prove the error estimates stated in Theorem 3.1.

Thanks to Lemma 2.5, we have for $p \geq 2$

$$\|\overline{u}^{\mathcal{T}} - u^{\mathcal{T}}\|_{1,p,\mathcal{T}}^p \leq C \left(\nabla J_{\mathcal{T}}(\overline{u}^{\mathcal{T}}) - \nabla J_{\mathcal{T}}(u^{\mathcal{T}}), \overline{u}^{\mathcal{T}} - u^{\mathcal{T}}\right),$$

and for $1 < p \leq 2$, we derive

$$\|\overline{u}^{\mathcal{T}} - u^{\mathcal{T}}\|_{1,p,\mathcal{T}}^2 \left(\|u^{\mathcal{T}}\|_{1,p,\mathcal{T}}^p + \|\overline{u}^{\mathcal{T}}\|_{1,p,\mathcal{T}}^p\right)^{\frac{p-2}{p}} \leq C \left(\nabla J_{\mathcal{T}}(\overline{u}^{\mathcal{T}}) - \nabla J_{\mathcal{T}}(u^{\mathcal{T}}), \overline{u}^{\mathcal{T}} - u^{\mathcal{T}}\right).$$

By (2.24), we have

$$\left(\nabla J_{\mathcal{T}}(\overline{u}^{\mathcal{T}}) - \nabla J_{\mathcal{T}}(u^{\mathcal{T}}), \overline{u}^{\mathcal{T}} - u^{\mathcal{T}}\right) = \sum_{\kappa \in \mathcal{T}} \sum_{\sigma \in \mathcal{E}_\kappa} \left(a_{\kappa,\sigma}(\overline{u}^{\mathcal{T}}) - a_{\kappa,\sigma}(u^{\mathcal{T}}), \overline{u}_\kappa - u_\kappa\right).$$

24

Integrating equation (1.1) over a control volume κ and using the definition of the scheme (2.23), we have

$$-\sum_{\sigma \in \mathcal{E}_\kappa} \int_\sigma |\nabla \overline{u}|^{p-2} \nabla \overline{u} \cdot \nu_\kappa \, ds = \int_\kappa f(z) \, dz = m(\kappa) f_\kappa = \sum_{\sigma \in \mathcal{E}_\kappa} a_{\kappa,\sigma}(u^\mathcal{T}).$$

Therefore, the definition of $R_{\sigma,\kappa}$ implies

$$\left(\nabla J_\mathcal{T}(\overline{u}^\mathcal{T}) - \nabla J_\mathcal{T}(u^\mathcal{T}), \overline{u}^\mathcal{T} - u^\mathcal{T}\right) = \sum_{\kappa \in \mathcal{T}} \sum_{\sigma \in \mathcal{E}_\kappa} \left(R_{\sigma,\kappa}, \overline{u}_\kappa - u_\kappa\right). \tag{3.8}$$

The conservativity property ensures that for any half-edge $\sigma \subset \kappa | \mathcal{L}$, we have $R_{\sigma,\mathcal{L}} = -R_{\sigma,\kappa}$. Bringing together the terms corresponding to neighbors κ and $\mathcal{L}$, we get a sum over all the half-edges in the mesh which reads

$$\left(\nabla J_\mathcal{T}(\overline{u}^\mathcal{T}) - \nabla J_\mathcal{T}(u^\mathcal{T}), \overline{u}^\mathcal{T} - u^\mathcal{T}\right) \leq \sum_{\sigma = \kappa | \mathcal{L}} \left|\left(R_{\sigma,\kappa}, (\overline{u}_\kappa - u_\kappa - \overline{u}_\mathcal{L} + u_\mathcal{L})\right)\right|,$$

where the ghost cells values are taken into account near the boundary. Let us note $R_\sigma = |R_{\sigma,\kappa}| = |R_{\sigma,\mathcal{L}}|$ so that we get

$$\left(\nabla J_\mathcal{T}(\overline{u}^\mathcal{T}) - \nabla J_\mathcal{T}(u^\mathcal{T}), \overline{u}^\mathcal{T} - u^\mathcal{T}\right) \leq \|\overline{u}^\mathcal{T} - u^\mathcal{T}\|_{1,p,\mathcal{T}} \left(\sum_{\kappa^* \in \mathcal{T}^*} \frac{1}{m(\kappa^*)^{\frac{1}{p-1}}} \sum_{\sigma \subset \kappa^* \cap \overline{\Omega}} \mathrm{diam}(\kappa^*)^{\frac{p}{p-1}} |R_\sigma|^{\frac{p}{p-1}}\right)^{\frac{p-1}{p}}.$$

The mesh regularity assumption (2.2) ensures that

$$\left(\nabla J_\mathcal{T}(\overline{u}^\mathcal{T}) - \nabla J_\mathcal{T}(u^\mathcal{T}), \overline{u}^\mathcal{T} - u^\mathcal{T}\right) \leq C \|\overline{u}^\mathcal{T} - u^\mathcal{T}\|_{1,p,\mathcal{T}} h^{\frac{p-2}{p}} \left(\sum_{\kappa^* \in \mathcal{T}^*} \sum_{\sigma \subset \kappa^* \cap \overline{\Omega}} |R_\sigma|^{\frac{p}{p-1}}\right)^{\frac{p-1}{p}}.$$

According to Lemma 3.4, the term $\displaystyle\sum_{\kappa^* \in \mathcal{T}^*} \sum_{\sigma \subset \kappa^* \cap \overline{\Omega}} |R_\sigma|^{\frac{p}{p-1}}$ is controled using the following inequalities

- For any $p > 1$, we have

$$h^{\frac{(p-1)^2+1}{p-1}} \sum_{\kappa^* \in \mathcal{T}^*} \sum_{\sigma \subset \kappa^* \cap \overline{\Omega}} \int_{\kappa^* \cap \Omega} |D^2 \overline{u}(z)|^p \, dz \leq C h^{\frac{(p-1)^2+1}{p-1}} \int_\Omega |D^2 \overline{u}(z)|^p \, dz.$$

- For any $p \geq 2$, using (2.2) we also have

$$h^2 |\nabla \overline{u}|_\infty^{\frac{p-2}{p-1}} \sum_{\kappa^* \in \mathcal{T}^*} \sum_{\sigma \subset \kappa^* \cap \overline{\Omega}} \left(\int_{\kappa^* \cap \Omega} |D^2 \overline{u}(z)|^p \, dz\right)^{\frac{1}{p-1}} \leq C h^2 |\nabla \overline{u}|_\infty^{\frac{p-2}{p-1}} \left(\sum_{\kappa^* \in \mathcal{T}^*} \int_{\kappa^* \cap \Omega} |D^2 \overline{u}(z)|^p \, dz\right)^{\frac{1}{p-1}} \left(\sum_{\kappa^* \in \mathcal{T}^*} 1\right)^{\frac{p-2}{p-1}}$$

$$\leq C h^{\frac{2}{p-1}} |\nabla \overline{u}|_\infty^{\frac{p-2}{p-1}} \left(\int_\Omega |D^2 \overline{u}(z)|^p \, dz\right)^{\frac{1}{p-1}}.$$

Hence, for $p > 2$, using the interpolation inequality $\|\nabla \overline{u}\|_\infty \leq C \|\overline{u}\|_{W^{2,p}}^{1-\frac{2}{p}} \|\overline{u}\|_{W^{1,p}}^{\frac{2}{p}}$, we obtain

$$\|\overline{u}^\mathcal{T} - u^\mathcal{T}\|_{1,p,\mathcal{T}} \leq C \left(h \|\overline{u}\|_{W^{2,p}} + h^{\frac{1}{p-1}} \|\overline{u}\|_{W^{2,p}}^{\frac{3p-4}{p(p-1)}} \|\overline{u}\|_{W^{1,p}}^{\frac{(p-2)^2}{p(p-1)}}\right),$$

and we can conclude using (1.3).

In the case where $1 < p \leq 2$, we first obtain

$$\|\overline{u}^{\mathcal{T}} - u^{\mathcal{T}}\|_{1,p,\mathcal{T}} \leq Ch^{p-1}\|\overline{u}\|_{W^{2,p}}^{p-1}\left(\|\overline{u}^{\mathcal{T}}\|_{1,p,\mathcal{T}}^{2-p} + \|u^{\mathcal{T}}\|_{1,p,\mathcal{T}}^{2-p}\right),$$

then with (1.3), (2.3) and (3.1) we deduce

$$\|\overline{u}^{\mathcal{T}} - u^{\mathcal{T}}\|_{1,p,\mathcal{T}} \leq Ch^{p-1}\|\overline{u}\|_{W^{2,p}}^{p-1}\|f\|_{L^{p'}}^{\frac{2-p}{p-1}}.$$

3.4 Improved error estimates

The convergence estimate of theorem 3.1 can be improved for $p = 2$ and $p > 3$, in the case where :

(H1) The mesh is uniform, i.e. all the control volumes are isometric to a given reference rectangular volume $[0, h_x] \times [0, h_y]$. In this case, the regularity assumptions on the mesh (2.2) writes

$$\exists c_1 > 0, \quad \text{such that} \quad c_1 \leq \frac{h_y}{h_x} \leq \frac{1}{c_1}. \tag{3.9}$$

(H2) The exact solution $\overline{u}$ of (1.1) is smooth enough.

(H3) The same approximation scheme is chosen on all dual control volumes.

The meaning of the last assumption has to be precised. As we have seen before, our schemes are determined by the value of the parameter ξ^{κ^*} in each dual control volume. As the mesh is supposed to be uniform in this section, it makes sense to impose $\xi^{\kappa^*} \overset{\text{def}}{=} \xi$ independently of the dual control volume κ^*. As usual (see (2.27)) we suppose that for a given family of meshes and associated schemes, there exists $\gamma > 0$ such that

$$\xi \leq \frac{1}{\gamma}, \text{ and } 2\xi + \frac{1}{4}\left(\frac{h_y}{h_x} + \frac{h_x}{h_y}\right) \geq \gamma. \tag{3.10}$$

Theorem 3.2 *Assume that $p = 2$ or $p > 3$. Consider an admissible scheme, in the sense of Definition 2.2, on a uniform mesh which satisfies (2.2) and (3.10). There exists a constant $C > 0$ such that, for any $\overline{u} \in W^{4,1}(\Omega) \cap W_0^{1,p}(\Omega)$ solution of (1.1), we have*

$$\|u^{\mathcal{T}} - \overline{u}^{\mathcal{T}}\|_{1,p,\mathcal{T}} \leq C\|\overline{u}\|_{W^{4,1}}h^{m/(p-1)} + C\left(\sup_{\partial\Omega}|f|\right)^{\frac{1}{p-1}} h^{\frac{1}{p}+\frac{1}{p-1}},$$

where $m = 2$ for $p = 2$ and for $p \geq 4$, and $m = p - 2$ for $3 < p < 4$.

Note that in each case we have $m > 1$, so that the convergence order obtained here is improved compared to the order $h^{\frac{1}{p-1}}$ obtained in the theorem 3.1, for less regular solutions.

The proof of this result is quite technical and can be found in [2]. We only give here a sketch of this proof.

The idea is to treat the right-hand side of (3.8) in a different manner, bringing together the terms originated from all the eight half-edges which delimit a control volume $\kappa \in \mathcal{T}$ in order to take advantage of the symmetries of the schemes. More precisely, we rewrite (3.8) under the form

$$\left(\nabla J_{\mathcal{T}}(\overline{u}^{\mathcal{T}}) - \nabla J_{\mathcal{T}}(u^{\mathcal{T}}), \overline{u}^{\mathcal{T}} - u^{\mathcal{T}}\right) = \sum_{\kappa \in \mathcal{T}}\left(\sum_{\sigma \in \mathcal{E}_\kappa} R_{\sigma,\kappa}\right)(u_\kappa - \overline{u}_\kappa).$$

As it has been shown in Lemma 3.4, for any σ and κ, the local consistency error $R_{\sigma,\kappa}$ is in general of order h^2 if u is smooth enough (say in $\mathcal{C}^2(\overline{\Omega})$). We can show that compensations due to symmetries of the scheme imply that the term $\sum_{\sigma \in \mathcal{E}_\kappa} R_{\sigma,\kappa}$ is of order h^4 (if κ is an interior control volume) provided that $\overline{u}$ is regular enough (roughly speaking, in $\mathcal{C}^4(\overline{\Omega})$) and g_σ lies at least in $\mathcal{C}^2(\mathbb{R}^4)$. For this last reason, this result is restricted to $p \geq 3$ or $p = 2$.
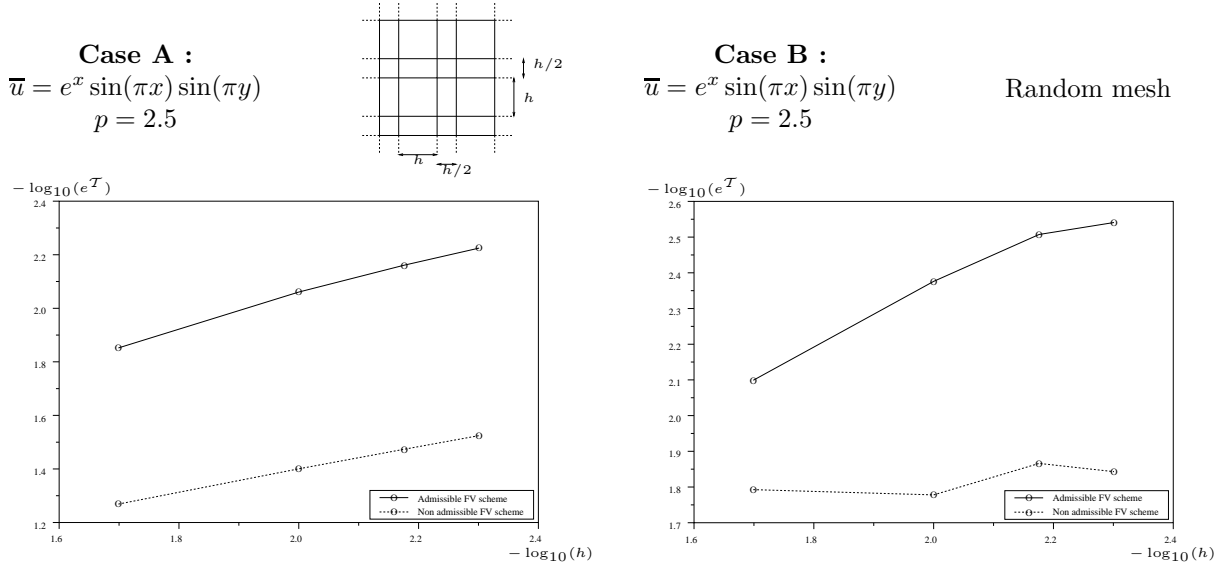
Figure 4.1: Comparison between Admissible and Non-admissible schemes

4 Numerical results

The admissible finite volume schemes we have constructed derive from a discrete energy $J_\mathcal{T}$. Hence, the practical computation of the approximate solution has been performed using the Polak-Ribire nonlinear conjugate gradient method. Note that for large values of p, there exist more efficient methods for this minimization problem (see for instance [12]).

As it is classical in such problems (see for instance [5]), we observe numerically better convergence orders than the results given in Theorems 3.1 and 3.2. Hence, we concentrate in this section on the qualitative properties of the schemes.

All the figures show the relative discrete $W^{1,p}$ error $e^\mathcal{T} = \frac{\|\overline{u}^\mathcal{T} - u^\mathcal{T}\|_{1,p,\mathcal{T}}}{\|\overline{u}^\mathcal{T}\|_{1,p,\mathcal{T}}}$ as a function of the size of the mesh h, in a logarithmic scale. The domain Ω is the unit square $]0, 1[\times]0, 1[$.

First, we compare admissible schemes as defined by Definition 2.2 (we choose here $\xi_{\mathcal{K}^*} = 0$) and the non-admissible scheme introduced in Section 2.5, that is the scheme defined by the functional $J_\mathcal{T}^0$ issued from the quadratic form (2.32). We recall that the Euler-Lagrange equation for $J_\mathcal{T}^0$ is not a consistent finite volume discretization of (1.1) on the mesh $\mathcal{T}$ if the mesh is not uniform.

Two kinds of behavior are observed and presented in Figure 4.1. In case A, the mesh is non-uniform but quite regular. We see that the convergence order is roughly the same for the two schemes but the error is smaller for the admissible scheme than for the non-admissible scheme. Note that the dual mesh $\mathcal{T}^*$ is uniform in this case.

In case B, the dimensions of the control volumes are randomly chosen between $h/2$ and $3h/2$ so that the regularity condition (2.2) is satisfied. We see that the non-admissible scheme has very bad convergence properties whereas our admissible scheme is convergent.

In Figure 4.2, we give another example of the bad convergence properties of non-admissible schemes in the case of random meshes.

Secondly, we compare in Figure 4.3 the various admissible schemes obtained by choosing different values of the parameter $\xi_{\mathcal{K}^*}$. In all the following results, we have chosen the same $\xi^{\mathcal{K}^*}$ on each dual control volume.

We recall that the natural scheme (extended from the classical finite volume scheme for the Laplace equation) is obtained for $\xi = 0.0$ (see Section 2.5). We see in Figure 4.3 that the convergence order is the same for each value of ξ. For $p = 2.5$, the smallest error is obtained for $\xi = 0.0$, i.e. with the "classical"

$$\overline{u} = \sin(3\pi x)\sin(3\pi y)$$
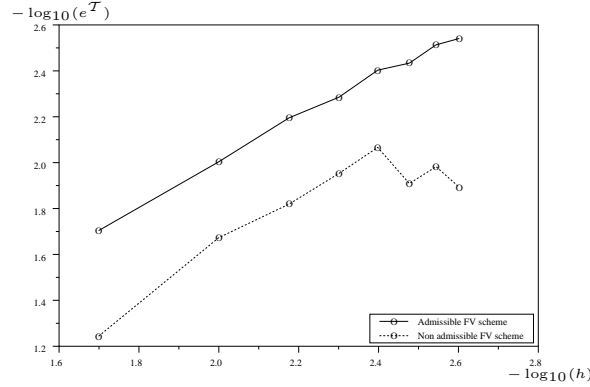$$p = 2.5$$

Random mesh

Figure 4.2: Comparison between Admissible and Non-admissible schemes (2)

$$\overline{u} = e^x \sin(3\pi x)\sin(3\pi y)$$
$$p = 2.5$$

Uniform mesh
$N \times N$ control volumes

$$\overline{u} = e^x \sin(3\pi x)\sin(3\pi y)$$
$$p = 4$$

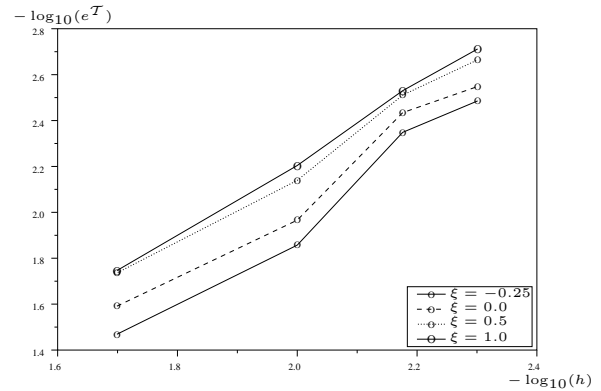Uniform mesh
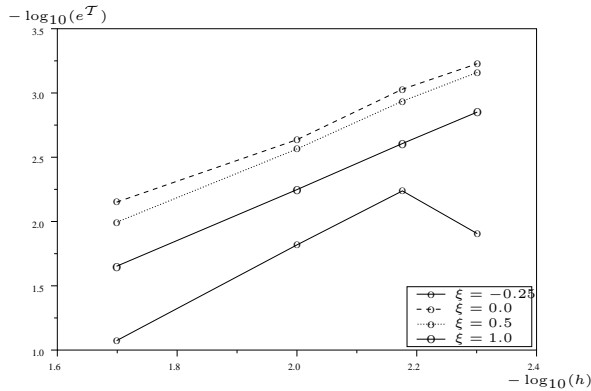$N \times N$ control volumes

Figure 4.3: Comparison of Admissible schemes for different ξ

finite volume scheme. On the other hand, for $p = 4$ we see that the smallest error is obtained for $\xi = 1.0$, so that the "classical" scheme ($\xi = 0.0$) is not the best scheme in this case.

Note that the scheme for $\xi = -0.25$ is not admissible in the sense of Definition 2.2 since this is the limiting case in condition (2.21). Even though this particular scheme does not fulfill the assumptions of our error estimates theorem, we can see that the scheme seems to be convergent, for instance for $p = 4$.

References

[1] P. ANGOT, C.-H. BRUNEAU AND P. FABRIE, A penalization method to take into account obstacles in incompressible viscous flows, *Numer. Math.*, **81**, pp. 497-520, 1999.

[2] B. ANDREIANOV, F. BOYER AND F. HUBERT, Finite volume schemes for the p-laplacian. Further error estimates, *Preprint no 03-29, LATP Université de Provence*, submitted, 2003.

[3] B. ANDREIANOV, M. GUTNIC AND P. WITTBOLD, Convergence of finite volume approximations for a nonlinear elliptic-parabolic problem : a "continuous" approach, *SIAM J. on Numer. Anal.,* **42** no 1, pp. 228-251, 2004.

[4] J.W. BARRETT AND W.B. LIU, A remark on the regularity of the solutions of the *p*-laplacian and its application to the finite element approximation, *Journal of Math. anal. Appl.* **178**, pp. 470-487, 1993.

[5] J.W. BARRETT AND W.B. LIU, Finite element approximation of the *p*-laplacian, *Math. Comp.* **61**, pp. 523-537, 1993.

[6] S. CHOW, Finite element error estimates for non-linear elliptic equations of monotone type, *Numer. Math.* **54** , pp. 373-393, 1989.

[7] Y. COUDIÈRE, J.-P. VILA AND P. VILLEDIEU Convergence rate of a finite volume scheme for a two dimensional convection-diffusion problem, M^2AN, **33**, no 3, pp. 493-516, 1999.

[8] J.I. DIAZ AND F. DE THELIN, On a nonlinear parabolic problem arising in some models related to turbulent flows, *SIAM J. Math. Anal.* **25** , no 4, pp. 1085-1111, 1994.

[9] R. EYMARD, T. GALLOUËT AND R. HERBIN Finite Volume Methods, *Handbook of Numerical Analysis, Vol. VII,* edited by P.G. Ciarlet and J.L. Lions, North-Holland, 2000.

[10] R. EYMARD, T. GALLOUËT AND R. HERBIN Finite volume approximation of elliptic problems and convergence of an approximate gradient, *Appl. Num. Math.*, **37**, no 1-2, pp. 31-53, 2001.

[11] R. EYMARD, T. GALLOUËT AND R. HERBIN A finite volume scheme for anisotropic diffusion problems, *C.R. Acad. Sci. Paris Srie 1*, to appear, 2004.

[12] R. GLOWINSKI AND A. MARROCCO, Sur l'approximation par éléments finis d'ordre un, et la résolution, par pénalisation-dualité, d'une classe de problèmes de Dirichlet non linéaires, *Journal of Math. anal. Appl. RAIRO Anal. Num.* **2**, pp. 41-64, 1975.

[13] R. GLOWINSKI AND J. RAPPAZ, Approximation of a nonlinear elliptic problem arising in a non-Newtonian fluid flow model in glaciology, *Math. Model. Numer. Anal.* **37**, no 1, pp. 175-186, 2003.

[14] P. OMNES AND K. DOMELEVO, A finite volume method for the Laplace equation on almost arbitrary two-dimensional grids. Part I : construction, *submitted*, 2004.

[15] P. OMNES, A finite volume method for the Laplace equation on almost arbitrary two-dimensional grids. Part II : error estimates, *submitted*, 2004.

[16] M. PICASSO, J. RAPPAZ, A. REIST, M. FUNK AND H. BLATTER, Numerical simulation of the motion of a two dimensional glacier, *Int. J. Numer. Methods Eng.*, **60**, pp. 995-1009, 2004.

[17] J. SIMON, Régularité de la solution d'un problème aux limites non linéaires, *Ann. Fac. Sciences Toulouse* **3** , pp. 247-274, 1981.