Discretization schemes for heterogeneous and anisotropic diffusion problems on general nonconforming meshes

Robert Eymard, Thierry Gallouët, Raphaele Herbin

Discretization schemes for heterogeneous and anisotropic diffusion problems on general nonconforming meshes[1]

R. Eymard[2], T. Gallouët[3] and R. Herbin[4]

January 22, 2008

Abstract: A discretization of heterogeneous and anisotropic diffusion problems on general nonconforming meshes is developed and studied. The unknowns of this scheme are the values at the center of the control volumes and at some internal interfaces, chosen because of some irregularity of the diffusion tensor. If the tensor is regular enough, the values on the interfaces may be deduced from the values at the center, at the expense of loosing the local conservativity of the fluxes. This scheme is shown to be accurate on several numerical examples. Mathematical convergence to the continuous solution is obtained for homogeneous and heterogeneous tensors. An error estimate may be drawn under sufficient regularity assumptions on the solution.

Keywords : Heterogeneous anisotropic diffusion, nonconforming grids, finite volume schemes
AMS Classificaton : 65N30

1 Introduction

Anisotropic heterogenous diffusion problems arise in a wide range of scientific fields such as hydrogeology, oil reservoir simulation, plasma physics, semiconductor modelling, biology.... When implementing numerical methods for this kind of problem, one needs to find an approximation of u, weak solution to the following equation:

$$-\mathrm{div}(\Lambda(\boldsymbol{x})\nabla u) = f \text{ in } \Omega, \tag{1}$$

with boundary condition

$$u = 0 \text{ on } \partial\Omega, \tag{2}$$

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

$$\Omega \text{ is an open bounded connected polyhedral subset of } \mathbb{R}^d,\ d \in \mathbb{N} \setminus \{0\}, \tag{3}$$

$$\Lambda \text{ is a measurable function from } \Omega \text{ to } \mathcal{M}_d(\mathbb{R}), \tag{4}$$

where we denote by $\mathcal{M}_d(\mathbb{R})$ the set of $d \times d$ matrices, such that for a.e. $\boldsymbol{x} \in \Omega$, $\Lambda(\boldsymbol{x})$ is symmetric, and such that the set of its eigenvalues is included in $[\underline{\lambda}, \overline{\lambda}]$, where $\underline{\lambda}, \overline{\lambda} \in L^\infty(\Omega)$ are such that $0 < \alpha_0 \leq \underline{\lambda}(\boldsymbol{x}) \leq \overline{\lambda}(\boldsymbol{x})$ for a.e. $\boldsymbol{x} \in \Omega$, and

$$f \in L^2(\Omega). \tag{5}$$

Under these hypotheses, the weak solution of (1)–(2) is the unique function u satisfying:

$$\left\{ \begin{array}{l} u \in H_0^1(\Omega), \\ \displaystyle\int_\Omega \Lambda(\boldsymbol{x})\nabla u(\boldsymbol{x}) \cdot \nabla v(\boldsymbol{x}) \mathrm{d}\boldsymbol{x} = \int_\Omega f(\boldsymbol{x}) v(\boldsymbol{x}) \mathrm{d}\boldsymbol{x}, \quad \forall v \in H_0^1(\Omega). \end{array} \right. \tag{6}$$

Usual discretization schemes for Problem (6) include finite difference, finite element or finite volume methods. Finite volume methods are actually very popular in oil engineering, the main reason problably being that

[1]This work was supported by GDR MOMAS, CNRS/PACEN
[2]Université de Paris-Est, France, Robert.Eymard@univ-mlv.fr
[3]Université de Provence, France, Thierry.Gallouet@cmi.univ-mrs.fr
[4]Université de Provence, France, Raphaele.Herbin@cmi.univ-mrs.fr

complex coupled physical phenomena may be discretized on the same grids. The well known five point and four point schemes on rectangles [23] or triangles [20] are not easily adapted to heterogeneous anisotropic diffusion operators [5]. An enlarged stencil scheme which handles anisotropy on meshes satisfying an orthogonality property was proposed and analysed in [14, 15]. Another problem which has to be faced in several fields of applications (such as hydrogeology and oil engineering) is the fact that the discretization meshes are imposed by engineering and computing considerations; therefore, we have to deal with distorted and possibly non conforming meshes.
A huge litterature exists in the engineering setting, so that we shall not try to be exhaustive. Let us nevertheless mention the finite volume schemes using the well known multipoint flux approximation [1, 2, 3]. These schemes involve the reconstruction of the gradient in order to evaluate the fluxes, which is also the case in [10, 22]. Among other approaches let us cite [19], which uses a parametrization technique. However, even though these schemes perform well in a number of cases, their convergence analysis often seems to remain out of reach, except under geometrical conditions [10].
In the two-dimensional case, we also mention [6], which is based on vertex reconstructions, and the family of double mesh schemes [21, 11, 7]. The generalization of this type of scheme to 3D is ongoing work.
In [16] we presented a "hybrid finite volume" scheme for any space dimension, which involves edges unkowns in addition to the usual cell unknowns. This is also the case for the mimetic finite difference schemes [8]. Along the same line of thought, a "mixed finite volume" scheme was proposed in [12]. These schemes perform quite well but seem rather expensive at first glance, because of the edge unknowns and equations. In [18] a cell centred scheme is used for the approximation of the Laplace operator on non conforming grids. It is cheaper than the two above mentioned scheme because it is based on cell unknowns only. In the present work, we construct discretization schemes for any kind of polyhedral mesh which take the best of these two latter schemes: unknowns on the edges are only introduced when there is strong heterogeneity of the medium at these edges.

The outline of this paper is as follows. In Section 2, we present the guidelines which have led to convergent schemes on general nonconforming meshes. The practical properties of the schemes thus obtained are shown on numerical examples in Section 3. Then the mathematical analysis of convergence and error estimate is performed in Section 4. This analysis needs some discrete functional analysis, such as discrete Sobolev inequalities, provided in section 5. Conclusions and perspectives are drawn in section 6.

2 Fundamentals for a class of non conforming schemes

Let us first present the desired properties which have led us to the design of the schemes under study:

1. The schemes must apply on any type of grid: conforming or non conforming, 2D and 3D (or more, see for instance the frameworks of kinetic formulations or financial mathematics), made with control volumes which are only assumed to be polyhedral (the boundary of each control volume is a finite union of subsets of hyperplanes).

2. The matrices of the generated linear systems are expected to be sparse, symmetric, positive and definite.

3. We wish to be able to prove the convergence of the discrete solution and an associate gradient to the solution of the continuous problem and its gradient, and to show error estimates.

In order to describe the schemes we now introduce some notations for the space discretization.

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

1. *$\mathcal{M}$ is a finite family of non empty connex open disjoint subsets of Ω (the "control volumes") such that $\overline{\Omega} = \cup_{K\in\mathcal{M}} \overline{K}$. For any $K \in \mathcal{M}$, let $\partial K = \overline{K} \setminus K$ be the boundary of K; let $\mathrm{m}(K) > 0$ denote the measure of K and h_K denote the diameter of K.*

2. *$\mathcal{E}$ is a finite family of disjoint subsets of $\overline{\Omega}$ (the "edges" of the mesh), such that, for all $\sigma \in \mathcal{E}$, σ is a non empty open subset of a hyperplane of $\mathbb{R}^d$, whose (d-1)-dimensional measure $\mathrm{m}(\sigma)$ is strictly positive. We also assume that, for all $K \in \mathcal{M}$, there exists a subset $\mathcal{E}_K$ of $\mathcal{E}$ such that $\partial K = \cup_{\sigma\in\mathcal{E}_K} \overline{\sigma}$. For any $\sigma \in \mathcal{E}$, we denote by $\mathcal{M}_\sigma = \{K \in \mathcal{M}, \sigma \in \mathcal{E}_K\}$. We then assume that, for all $\sigma \in \mathcal{E}$, either $\mathcal{M}_\sigma$ has exactly one element and then $\sigma \subset \partial\Omega$ (the set of these interfaces, called boundary interfaces, is denoted by $\mathcal{E}_{\mathrm{ext}}$) or $\mathcal{M}_\sigma$ has exactly two elements (the set of these interfaces, called interior interfaces, is denoted by $\mathcal{E}_{\mathrm{int}}$). For all $\sigma \in \mathcal{E}$, we denote by $\boldsymbol{x}_\sigma$ the barycenter of σ. For all $K \in \mathcal{M}$ and $\sigma \in \mathcal{E}_K$, we denote by $\boldsymbol{n}_{K,\sigma}$ the unit vector normal to σ outward to K.*

3. *$\mathcal{P}$ is a family of points of Ω indexed by $\mathcal{M}$, denoted by $\mathcal{P} = (\boldsymbol{x}_K)_{K\in\mathcal{M}}$, such that for all $K \in \mathcal{M}$, $\boldsymbol{x}_K \in K$ and K is assumed to be $\boldsymbol{x}_K$-star-shaped, which means that for all $\boldsymbol{x} \in K$, the property $[\boldsymbol{x}_K, \boldsymbol{x}] \subset K$ holds. Denoting by $d_{K,\sigma}$ the Euclidean distance between $\boldsymbol{x}_K$ and the hyperplane including σ, one assumes that $d_{K,\sigma} > 0$. We then denote by $D_{K,\sigma}$ the cone with vertex $\boldsymbol{x}_K$ and basis σ.*

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

2.1 From a "hybrid" finite volume scheme...

The idea of the "hybrid" schemes (among them one may include the mixed finite elements, the mixed finite volume or the mimetic finite difference schemes) is to find an approximation of the solution of (1)–(2) by setting up a system of discrete equations for a family of values $((u_K)_{K\in\mathcal{M}}, (u_\sigma)_{\sigma\in\mathcal{E}})$ in the control volumes and on the interfaces. The number of unknowns is therefore $\mathrm{card}(\mathcal{M}) + \mathrm{card}(\mathcal{E})$. Following the idea of the finite volume framework, equation (1) is integrated over each control volume $K \in \mathcal{M}$, which formally gives (assuming sufficient regularity on u and Λ) the following balance equation on the control volume K:

$$\sum_{\sigma\in\mathcal{E}_K} \left(-\int_\sigma \Lambda(\boldsymbol{x})\nabla u(\boldsymbol{x}) \cdot \boldsymbol{n}_{K,\sigma} \mathrm{d}\gamma(\boldsymbol{x}) \right) = \int_K f(\boldsymbol{x})\mathrm{d}\boldsymbol{x}.$$

The flux $-\int_\sigma \Lambda(\boldsymbol{x})\nabla u(\boldsymbol{x}) \cdot \boldsymbol{n}_{K,\sigma} \mathrm{d}\gamma(\boldsymbol{x})$ is approximated by a function $F_{K,\sigma}(u)$ of the values $((u_K)_{K\in\mathcal{M}}, (u_\sigma)_{\sigma\in\mathcal{E}})$ at the "centers" and at the interfaces of the control volumes (in all practical cases, $F_{K,\sigma}(u)$ only depends on u_K and all $(u_{\sigma'})_{\sigma'\in\mathcal{E}_K}$). A discrete equation corresponding to (1) is then:

$$\sum_{\sigma\in\mathcal{E}_K} F_{K,\sigma}(u) = \int_K f(\boldsymbol{x})\mathrm{d}\boldsymbol{x}, \ \forall K \in \mathcal{M}. \tag{7}$$

The values u_σ on the interfaces are then introduced so as to allow for a consistent approximation of the normal fluxes in the case of an anisotropic operator and a general, possibly nonconforming mesh. We thus have $\mathrm{card}(\mathcal{E})$ supplementary unknowns, and need $\mathrm{card}(\mathcal{E})$ equations to ensure that the problem is well posed. For the boundary faces or edges, these equations are obtained by writing the discrete counterpart of the boundary condition (2):

$$u_\sigma = 0, \forall \sigma \in \mathcal{E}_{\mathrm{ext}}. \tag{8}$$

Following the finite volume ideas, we may write the continuity of the discrete flux for all interior edges, that is to say:

$$F_{K,\sigma}(u) + F_{L,\sigma}(u) = 0, \text{ for } \sigma \in \mathcal{E}_{\rm int} \text{ such that } \mathcal{M}_\sigma = \{K, L\}. \tag{9}$$

We now have $\mathrm{card}(\mathcal{M}) + \mathrm{card}(\mathcal{E}_{\rm int})$ unknowns and equations.

Remark 2.2 *In the case $\Lambda = \lambda(\boldsymbol{x})\mathrm{Id}$, on meshes satisfying an orthogonality condition, a consistent numerical flux is obtained using two point formula $F_{K,\sigma}(u) = \lambda_K \mathrm{m}(\sigma)(u_K - u_\sigma)/d_{K,\sigma}$, where λ_K is an average value for λ in K. Then, writing (9) for all $\sigma \in \mathcal{E}_{\rm int}$ such that $\mathcal{M}_\sigma = \{K, L\}$, we obtain u_σ as a linear combination of u_K and u_L. Plugging this expression in (7), we get a scheme with card($\mathcal{M}$) equations and card($\mathcal{M}$) unknowns (see [13] for more details). In the case of a rectangular (resp. triangular) mesh, this is the well known five points (resp four points) scheme with harmonic averages of the diffusion.*

With a proper choice of the expression $F_{K,\sigma}(u)$ which we shall introduce below, this scheme, first introduced in [16], is quite efficient, and may be shown to converge. It is in fact quite well suited for finite volume discretisation in heterogeneous media, where harmonic averages for Λ are preferred to arithmetic ones (see [4]). This scheme does have one drawback: since the number of unknowns is the sum of the number of control volumes and of interior interfaces, the resulting scheme is quite expensive (although it is sometimes possible to algebraically eliminate the values at the control volumes, in the same way it is done in the framework of the mixed hybrid finite elements, see [24]).

Remark 2.3 *Note that in the case of regular simplicial conforming meshes (triangles in 2D, tetrahedra in 3D), there is an algebraic possibility to express the unknowns $(u_\sigma)_{\sigma\in\mathcal{E}}$ as local affine combinations of the values $(u_K)_{K\in\mathcal{M}}$ [25] and therefore to eliminate them. The idea is to remark that the linear system constituted by the equations (7) for all $K \in \mathcal{M}_S$, where $\mathcal{M}_S$ is the set of all simplices sharing the same interior vertex S, and (9) for all the interior edges such that $\mathcal{M}_\sigma \subset \mathcal{M}_S$, presents as many equations as unknowns u_σ, for $\sigma \in \cup_{K\in\mathcal{M}_S}\mathcal{E}_K$. Indeed, the number of edges in $\cup_{K\in\mathcal{M}_S}\mathcal{E}_K$ such that $\mathcal{M}_\sigma \not\subset \mathcal{M}_S$ is equal to the number of control volumes in $\mathcal{M}_S$. Unfortunately, there is at this time no general result on the invertibility nor the symmetry of the matrix of this system, and this method does not apply to other types of meshes than the simplicial ones.*

In order to reduce the computational cost of the scheme, we developped in [18] an idea which is in fact close to the finite element philosophy since we express the finite volume scheme in a weak form; to this aim, let us first define the sets $X_\mathcal{D}$ and $X_{\mathcal{D},0}$ where the discrete unknowns lie, that is to say:

$$X_\mathcal{D} = \{v = ((v_K)_{K\in\mathcal{M}}, (v_\sigma)_{\sigma\in\mathcal{E}}), v_K \in \mathbb{R}, v_\sigma \in \mathbb{R}\}, \tag{10}$$

$$X_{\mathcal{D},0} = \{v \in X_\mathcal{D} \text{ such that } v_\sigma = 0 \,\forall \sigma \in \mathcal{E}_{\rm ext}\}. \tag{11}$$

Multiplying, for any $v \in X_{\mathcal{D},0}$, equation (7) by the value v_K of v on the control volume K and summing on $K \in \mathcal{M}$ leads to:

$$\sum_{K\in\mathcal{M}} v_K \sum_{\sigma\in\mathcal{E}_K} F_{K,\sigma}(u) = \sum_{K\in\mathcal{M}} v_K \int_K f(\boldsymbol{x})\mathrm{d}\boldsymbol{x}.$$

Using (9), we get the following discrete weak formulation:

$$\left\{ \begin{array}{l} \text{Find } u \in X_{\mathcal{D},0} \text{ such that:} \\ \langle u, v\rangle_F = \displaystyle\sum_{K\in\mathcal{M}} v_K \int_K f(\boldsymbol{x})\mathrm{d}\boldsymbol{x}, \quad \text{for all } v \in X_{\mathcal{D},0}, \end{array} \right. \tag{12}$$

with

$$\langle u, v\rangle_F = \sum_{K\in\mathcal{M}} \sum_{\sigma\in\mathcal{E}_K} F_{K,\sigma}(u)(v_K - v_\sigma). \tag{13}$$

Note that choosing $v \in X_{\mathcal{D},0}$ such that $v_K = 1$, $v_L = 0$ for any $L \in \mathcal{M}, L \neq K$ and $v_\sigma = 0$ for any $\sigma \in \mathcal{E}$ yields (7). Similarly, choosing $v \in X_{\mathcal{D},0}$ such that $v_K = 0$ for any $K \in \mathcal{M}$, and $v_\sigma = 1$ and $v_\tau = 0$ for any $\tau \in \mathcal{E}, \tau \neq \sigma$ leads to (9). Therefore the hybrid finite volume scheme (7)–(9) is equivalent to the discrete weak formulation (12).

2.2 . . . to a nonconforming finite element scheme. . .

We may then choose to use the weak discrete form (13) as an approximation of the bilinear form $a(\cdot,\cdot)$, but with a space of dimension smaller than that of $X_{\mathcal{D},0}$. This can be achieved by expressing the value of u on any interior interface $\sigma \in \mathcal{E}_{\rm int}$ as a consistent barycentric combination of the values u_K:

$$u_\sigma = \sum_{K\in\mathcal{M}} \beta_\sigma^K u_K, \tag{14}$$

where $(\beta_\sigma^K)_{\substack{K\in\mathcal{M}\\ \sigma\in\mathcal{E}_{\rm int}}}$ is a family of real numbers, with $\beta_\sigma^K \neq 0$ only for some control volumes K close to σ, and such that

$$\sum_{K\in\mathcal{M}} \beta_\sigma^K = 1 \text{ and } \boldsymbol{x}_\sigma = \sum_{K\in\mathcal{M}} \beta_\sigma^K \boldsymbol{x}_K, \forall \sigma \in \mathcal{E}_{\rm int}. \tag{15}$$

We recall that the values $u_\sigma, \sigma \in \mathcal{E}_{\rm ext}$ are set to 0 in order to respect the boundary conditions. This ensures that if φ is a regular function, then $\varphi_\sigma = \sum_{K\in\mathcal{M}} \beta_\sigma^K \varphi(\boldsymbol{x}_K)$ is a consistent approximation of $\varphi(\boldsymbol{x}_\sigma)$ for $\sigma \in \mathcal{E}_{\rm int}$. Hence the new scheme reads:

$$\left\{\begin{array}{l} \text{Find } u \in X_{\mathcal{D},0} \text{ such that } u_\sigma = \displaystyle\sum_{K\in\mathcal{M}} \beta_\sigma^K u_K,\ \forall \sigma \in \mathcal{E}_{\rm int}, \text{ and} \\ \langle u, v\rangle_F = \displaystyle\sum_{K\in\mathcal{M}} v_K \int_K f(\boldsymbol{x})\mathrm{d}\boldsymbol{x}, \quad \text{for all } v \in X_{\mathcal{D},0} \text{ with } v_\sigma = \displaystyle\sum_{K\in\mathcal{M}} \beta_\sigma^K v_K,\ \forall \sigma \in \mathcal{E}_{\rm int}. \end{array}\right. \tag{16}$$

This method has been shown in [18] to be efficient in the case of a problem where $\Lambda = \mathrm{Id}$ (for the approximation of the viscous terms in the Navier-Stokes problem). With an appropriate choice for the expression of the numerical flux, it also yields some kind of conservativity (more on this below), but no longer to the classical (in the finite volume framework) equation (9): indeed, since the degrees of freedom on the edges are no longer present, one may not use $v_\sigma = 1$ to recover (9). Note also that taking $v_K = 1$ does not yield (7). This scheme has been implemented for the discretization of the diffusive term in the incompressible Navier Stokes equations on general 2 or 3D grids, and gives excellent results [9]. Unfortunately, because of a poor approximation of the local flux at strongly hetereogenous interfaces, this approach is not sufficient to provide accurate results for some types of flows in heterogeneous media, as we shall show in section 3. This is especially true when using coarse meshes, as is often the case in industrial problems.

2.3 . . . to an optimal compromise ?

Therefore we now propose a scheme which has the advantage of both techniques: we shall use equation (13) and keep the unknowns u_σ on the edges which require them, for instance those where the matrix Λ is discontinuous: hence (9) will hold for all edges associated to these unknowns; for all other interfaces, we shall impose the values of u using (14), and therefore eliminate these unknowns. Let us decompose the set $\mathcal{E}_{\rm int}$ of interfaces into two non intersecting subsets, that is: $\mathcal{E}_{\rm int} = \mathcal{B} \cup \mathcal{H}, \mathcal{H} = \mathcal{E}_{\rm int} \setminus \mathcal{B}$. The interface unknowns associated with $\mathcal{B}$ will be computed by using the barycentric formula (14).

Remark 2.4 *Note that, although the accuracy of the scheme is increased in practice when the points where the matrix Λ is discontinuous are located within the set $\bigcup_{\sigma\in\mathcal{H}} \sigma$, such a property is not needed in the mathematical study of the scheme.*

Let us introduce the space $X_{\mathcal{D},\mathcal{B}} \subset X_{\mathcal{D},0}$ defined by:

$$X_{\mathcal{D},\mathcal{B}} = \{v \in X_{\mathcal{D}} \text{ such that } v_\sigma = 0 \text{ for all } \sigma \in \mathcal{E}_{\rm ext} \text{ and } v_\sigma \text{ satisfying (14) for all } \sigma \in \mathcal{B}\}. \tag{17}$$

The composite scheme which we consider in this work reads:

$$\left\{ \begin{array}{l} \text{Find } u \in X_{\mathcal{D},\mathcal{B}} \text{ such that:} \\ \langle u, v\rangle_F = \sum_{K\in\mathcal{M}} v_K \int_K f(\boldsymbol{x})\mathrm{d}\boldsymbol{x}, \quad \text{for all } v \in X_{\mathcal{D},\mathcal{B}}. \end{array} \right. \tag{18}$$

We therefore obtain a scheme with $\mathrm{card}(\mathcal{M}) + \mathrm{card}(\mathcal{H})$ equations and unknowns. It is thus less expensive while it remains precise (for the choice of numerical flux given below) even in the case of strong heterogeneities (see section 3).
Note that with the present scheme, (9) holds for all $\sigma \in \mathcal{H}$, but not generally for any $\sigma \in \mathcal{B}$. However, fluxes between pairs of control volumes can nevertheless be identified. These pairs are no longer necessarily connected by a common boundary, but determined by the stencil used in relation (14).

Remark 2.5 (Other boundary conditions) *In the case of Neumann or Robin boundary conditions, the discrete space $X_{\mathcal{D},\mathcal{B}}$ is modified to include the unknowns associated to the corresponding edges, and the resulting discrete weak formulation is then straightforward.*

Remark 2.6 (Extension of the scheme) *There is no additional difficulty to replace* (14) *in the definition of* (17) *by*

$$u_\sigma = \sum_{K\in\mathcal{M}} \beta_\sigma^K u_K + \sum_{\sigma'\in\mathcal{H}} \beta_\sigma^{\sigma'} u_{\sigma'}, \ \forall \sigma \in \mathcal{B},$$

with

$$\sum_{K\in\mathcal{M}} \beta_\sigma^K + \sum_{\sigma'\in\mathcal{H}} \beta_\sigma^{\sigma'} = 1 \text{ and } \boldsymbol{x}_\sigma = \sum_{K\in\mathcal{M}} \beta_\sigma^K \boldsymbol{x}_K + \sum_{\sigma'\in\mathcal{H}} \beta_\sigma^{\sigma'} \boldsymbol{x}_{\sigma'}, \forall \sigma \in \mathcal{B}.$$

Then all the mathematical properties shown below still hold.

2.4 Construction of the fluxes using a discrete gradient

For the definition of the schemes to be complete, there now remains to explain how we find a convenient expression for $F_{K,\sigma}(u)$ with respect to the discrete unknowns. An idea which has been used in several of the schemes presented in the introduction, is to look for a consistent expression of the flux by using adequate linear combinations of the unknowns; however, refering to the beginning of Section 2, such a recontruction does not in general lead to the desired properties 2 (symmetric definite positive matrices) and 3 (convergence). Our idea here is different: it is based on the identification of the numerical fluxes $F_{K,\sigma}(u)$ through the mesh dependent bilinear form $\langle\cdot,\cdot\rangle_F$ defined in (13), using the expression of a discrete gradient. Indeed let us assume that, for all $u \in X_{\mathcal{D}}$, we have constructed a discrete gradient $\nabla_{\mathcal{D}} u$, we then seek a family $(F_{K,\sigma}(u))_{\substack{K\in\mathcal{M}\\ \sigma\in\mathcal{E}_K}}$ such that

$$\sum_{K\in\mathcal{M}} \sum_{\sigma\in\mathcal{E}_K} F_{K,\sigma}(u)(v_K - v_\sigma) = \int_\Omega \nabla_{\mathcal{D}} u(\boldsymbol{x}) \cdot \Lambda(\boldsymbol{x}) \nabla_{\mathcal{D}} v(\boldsymbol{x})\mathrm{d}\boldsymbol{x}, \ \forall u, v \in X_{\mathcal{D}}. \tag{19}$$

Remark 2.7 *It is then always possible to deduce an expression for $F_{K,\sigma}(u)$ satisfying* (19)*, under the sufficient condition that, for all $K \in \mathcal{M}$ and a.e. $\boldsymbol{x} \in K$, $\nabla_{\mathcal{D}} u(\boldsymbol{x})$ may be expressed as a linear combination of $(u_\sigma - u_K)_{\sigma\in\mathcal{E}_K}$, the coefficients of which are measurable bounded functions of $\boldsymbol{x}$. This property is ensured in the construction of $\nabla_{\mathcal{D}} u(\boldsymbol{x})$ given below.*

Then, in order to ensure the desired properties 2 and 3, we shall see in Section 4 that it suffices that the discrete gradient satisfies the following properties.

1. For a sequence of space discretisations of Ω with mesh size tending to 0, if the sequence of associated grid functions is bounded in some sense, then their discrete gradient is expected to converge, at least weakly in $L^2(\Omega)^d$, to the gradient of an element of $H_0^1(\Omega)$;

2. If φ is a regular function from $\overline{\Omega}$ to $\mathbb{R}$, the discrete gradient of the piecewise function defined by taking the value $\varphi(\boldsymbol{x}_K)$ on each control volume K and $\varphi(\boldsymbol{x}_\sigma)$ on each edge σ is a consistent approximation of the gradient of φ.

Le us first define:

$$\nabla_K u = \frac{1}{\mathrm{m}(K)} \sum_{\sigma \in \mathcal{E}_K} \mathrm{m}(\sigma)(u_\sigma - u_K)\boldsymbol{n}_{K,\sigma},\ \forall K \in \mathcal{M},\ \forall u \in X_{\mathcal{D}}, \tag{20}$$

where $\boldsymbol{n}_{K,\sigma}$ is the outward to K normal unit vector, $\mathrm{m}(K)$ and $\mathrm{m}(\sigma)$ are the usual measures (volumes, areas, or lengths) of K and σ. The consistency of formula (20) stems from the following geometrical relation:

$$\sum_{\sigma \in \mathcal{E}_K} \mathrm{m}(\sigma)\boldsymbol{n}_{K,\sigma}(\boldsymbol{x}_\sigma - \boldsymbol{x}_K)^t = \mathrm{m}(K)\mathrm{Id},\ \forall K \in \mathcal{M}, \tag{21}$$

where $(\boldsymbol{x}_\sigma - \boldsymbol{x}_K)^t$ is the transpose of $\boldsymbol{x}_\sigma - \boldsymbol{x}_K \in \mathbb{R}^d$, and Id is the $d \times d$ identity matrix. Indeed, for any linear function defined on Ω by $\psi(\boldsymbol{x}) = \boldsymbol{G} \cdot \boldsymbol{x}$ with $\boldsymbol{G} \in \mathbb{R}^d$, assuming that $u_\sigma = \psi(\boldsymbol{x}_\sigma)$ and $u_K = \psi(\boldsymbol{x}_K)$, we get $u_\sigma - u_K = (\boldsymbol{x}_\sigma - \boldsymbol{x}_K)^t\boldsymbol{G} = (\boldsymbol{x}_\sigma - \boldsymbol{x}_K)^t\nabla\psi$, hence (20) leads to $\nabla_K u = \nabla\psi$.
Since the coefficient of u_K in (20) is in fact equal to zero, a reconstruction of the discrete gradient $\nabla_{\mathcal{D}} u$ solely based on (20) cannot lead to a definite discrete bilinear form in the general case. Hence, we now introduce:

$$\nabla_{K,\sigma} u = \nabla_K u + R_{K,\sigma} u\ \boldsymbol{n}_{K,\sigma}, \tag{22}$$

with

$$R_{K,\sigma} u = \frac{\sqrt{d}}{d_{K,\sigma}} \left(u_\sigma - u_K - \nabla_K u \cdot (\boldsymbol{x}_\sigma - \boldsymbol{x}_K)\right), \tag{23}$$

(recall that d is the space dimension and $d_{K,\sigma}$ is the Euclidean distance between $\boldsymbol{x}_K$ and σ). We may then define $\nabla_{\mathcal{D}} u$ as the piecewise constant function equal to $\nabla_{K,\sigma} u$ a.e. in the cone $D_{K,\sigma}$ with vertex $\boldsymbol{x}_K$ and basis σ:

$$\nabla_{\mathcal{D}} u(\boldsymbol{x}) = \nabla_{K,\sigma} u \text{ for a.e. } \boldsymbol{x} \in D_{K,\sigma}. \tag{24}$$

We can then prove that the discrete gradient defined by (20)-(24) meets the required properties (see Lemmas 4.2 and 4.3). In order to identify the numerical fluxes $F_{K,\sigma}(u)$ using the relation (19), we put the discrete gradient under the form

$$\nabla_{K,\sigma} u = \sum_{\sigma' \in \mathcal{E}_K} (u_{\sigma'} - u_K)\boldsymbol{y}^{\sigma\sigma'},$$

with

$$\boldsymbol{y}^{\sigma\sigma'} = \begin{cases} \dfrac{\mathrm{m}(\sigma)}{\mathrm{m}(K)}\boldsymbol{n}_{K,\sigma} + \dfrac{\sqrt{d}}{d_{K,\sigma}}\left(1 - \dfrac{\mathrm{m}(\sigma)}{\mathrm{m}(K)}\boldsymbol{n}_{K,\sigma} \cdot (\boldsymbol{x}_\sigma - \boldsymbol{x}_K)\right)\boldsymbol{n}_{K,\sigma} & \text{if } \sigma = \sigma' \\ \dfrac{\mathrm{m}(\sigma')}{\mathrm{m}(K)}\boldsymbol{n}_{K,\sigma'} - \dfrac{\sqrt{d}}{d_{K,\sigma}\mathrm{m}(K)}\mathrm{m}(\sigma')\boldsymbol{n}_{K,\sigma'} \cdot (\boldsymbol{x}_\sigma - \boldsymbol{x}_K)\boldsymbol{n}_{K,\sigma} & \text{otherwise .} \end{cases} \tag{25}$$

Thus:

$$\int_\Omega \nabla_{\mathcal{D}} u(\boldsymbol{x}) \cdot \Lambda(\boldsymbol{x}) \nabla_{\mathcal{D}} v(\boldsymbol{x}) \mathrm{d}\boldsymbol{x} = \sum_{K \in \mathcal{M}} \sum_{\sigma \in \mathcal{E}_K} \sum_{\sigma' \in \mathcal{E}_K} A_K^{\sigma\sigma'}(u_\sigma - u_K)(v_{\sigma'} - v_K),\ \forall u, v \in X_{\mathcal{D}}, \tag{26}$$

with:

$$A_K^{\sigma\sigma'} = \sum_{\sigma'' \in \mathcal{E}_K} \boldsymbol{y}^{\sigma''\sigma} \cdot \Lambda_{K,\sigma''} \boldsymbol{y}^{\sigma''\sigma'} \text{ and } \Lambda_{K,\sigma''} = \int_{D_{K,\sigma''}} \Lambda(\boldsymbol{x}) \mathrm{d}\boldsymbol{x}. \tag{27}$$

Then we get that the local matrices $(A_K^{\sigma\sigma'})_{\sigma\sigma'\in\mathcal{E}_K}$ are symmetric and positive, and the identification of the numerical fluxes using (19) leads to the expression:

$$F_{K,\sigma}(u) = \sum_{\sigma'\in\mathcal{E}_K} A_K^{\sigma\sigma'}(u_K - u_{\sigma'}). \tag{28}$$

The properties provided by this definition (which could not have been obtained using natural expansions of regular functions) are shown in Lemma 4.4. Then Theorem 4.1 shows that these properties are sufficient to provide the convergence of the scheme. Note that the proof of this property holds for general heterogeneous, anisotropic and possibly discontinuous fields Λ, for which the solution u of (6) is not in general more regular than $u \in H_0^1(\Omega)$. The local consistency property provided by definition (28) is only detailed in the error estimate theorem 4.2, in the case where Λ and u are regular enough.

Remark 2.8 *The choice of the coefficient $\sqrt{d}$ in (23) is not compulsory, and any fixed positive value could be substituted; it is motivated by the fact that it provides a diagonal matrix A_K in the case of meshes which satisfy $\boldsymbol{n}_{K,\sigma} = \frac{\boldsymbol{x}_\sigma - \boldsymbol{x}_K}{d_{K,\sigma}}$ (triangular, rectangular, orthogonal parallelepipedic meshes but unfortunately not general tetrahedric meshes), and yields the usual two point scheme. In this case, the formula (20) leads to the discrete gradient which was introduced in [15].*

3 Numerical results

We present some numerical results obtained with various choices of $\mathcal{B}$ in the scheme (18),(13) with the flux (28), which we synthetize here for the sake of clarity:

$$\left\{\begin{array}{l} \text{Find } u \in X_{\mathcal{D},\mathcal{B}} \text{ (that is } (u_K)_{K\in\mathcal{M}},\ (u_\sigma)_{\sigma\in\mathcal{H}}), \text{ such that:} \\ \displaystyle\sum_{K\in\mathcal{M}}\sum_{\sigma\in\mathcal{E}_K} F_{K,\sigma}(u)(v_K - v_\sigma) = \sum_{K\in\mathcal{M}} v_K \int_K f(\boldsymbol{x})\mathrm{d}\boldsymbol{x}, \quad \text{for all } v \in X_{\mathcal{D},\mathcal{B}}, \\ \text{with } F_{K,\sigma}(u) = \displaystyle\sum_{\sigma'\in\mathcal{E}_K} A_K^{\sigma\sigma'}(u_{\sigma'} - u_K), \forall K \in \mathcal{M}, \forall \sigma \in \mathcal{E}_K. \end{array}\right. \tag{29}$$

3.1 Order of convergence

We consider here the numerical resolution of Equation (1) supplemented by a homogeneous Dirichlet boundary condition; the right hand side is chosen so as to obtain an exact solution to the problem, so as to easily compute the error between the exact and approximate solutions. We consider Problem (1)-(2) with a constant matrix Λ:

$$\Lambda = \begin{pmatrix} 1.5 & .5 \\ .5 & 1.5 \end{pmatrix}, \tag{30}$$

and choose $f : \Omega \to \mathbb{R}$ and f such that the exact solution to Problem (1)–(30) is $\bar{u} : \Omega \to \mathbb{R}$ defined by $\bar{u}(x, y) = 16x(1 - x)y(1 - y)$ for any $(x, y) \in \Omega$. Note that in this case, the composite scheme is in fact the cell centred scheme, there are no edge unknowns.
Let us first consider conforming meshes, such as the triangular meshes which are depicted on Figure 1 and uniform square meshes.
For both $\mathcal{B} = \emptyset$ (hybrid scheme) and $\mathcal{B} = \mathcal{E}_{\text{int}}$ (cell centred scheme), the order of convergence is close to 2 for the unknown u and 1 for its gradient. Of course, the hybrid scheme is almost three times more costly in terms of number of unknowns than the cell centred scheme for a given precision. However, the number of nonzero terms in the matrix is, again for a given precision on the approximate solution, larger for the cell centred scheme than for the hybrid scheme. Hence the number of unknowns is probably not a sufficient criterion for assessing the cost of the scheme.

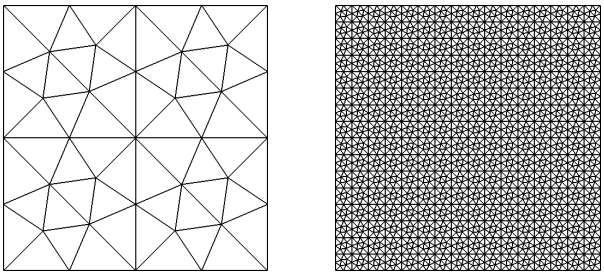

Figure 1: Regular conforming coarse and fine triangular grids

	NU		NM		$\epsilon(u)$		$\epsilon(\nabla u)$	
n	Hyb	Cent	Hyb	Cent	Hyb	Cent	Hyb	Cent
C1	130	48	874	488	1.28E-01	1.20E-01	1.64E-02	3.57E-02
NC	182	64	1334	724	1.03E-01	9.43E-02	1.66E-02	3.69E-02
C2	222	80	1542	864	7.61E-02	7.09E-02	9.18E-03	2.44E-02

Table 1: Error for the non conforming rectangular mesh, hybrid scheme (Hyb) and centred (Cent) schemes. For both schemes: NU is the number of unknowns in the resulting linear system, NM the number of non zero terms in the matrix, $\epsilon(u)$ the discrete L^2 norm of the error on the solution and $\epsilon(\nabla u)$ the discrete L^2 norm of the error on the gradient. C1 and C2 are the two conforming meshes represented on the left and the right in Figure 2, and NC the non conforming one represented in the middle.

Results were also obtained in the case of uniform square or rectangular meshes. They show a better rate of convergence of the gradient (order 2 in the case of the hybrid scheme and 1.5 in the case of the centred scheme), even though the rate of convergence of the approximate solution remains unchanged and close to 2.
We then use a rectangular nonconforming mesh, obtained by cutting the domain in two vertical sides and using a rectangular grid of $3n \times 2n$ (resp. $5n \times 2n$) on the first (resp. second side), where n is the number of the mesh, $n = 1, \dots 7$. . Again, the order of convergence which we obtain is 2 for u and around 1.8 for the gradient. We give in Table 1 below the errors obtained in the discrete L^2 norm for u and ∇u for a nonconforming mesh and (in terms of number of unknowns) and for the rectangular 4×6 and 4×10 conforming rectangular meshes, for both the hybrid and cell centred schemes. We show on Figure 2 the solutions for the corresponding grids (which looks very much the same for both schemes).

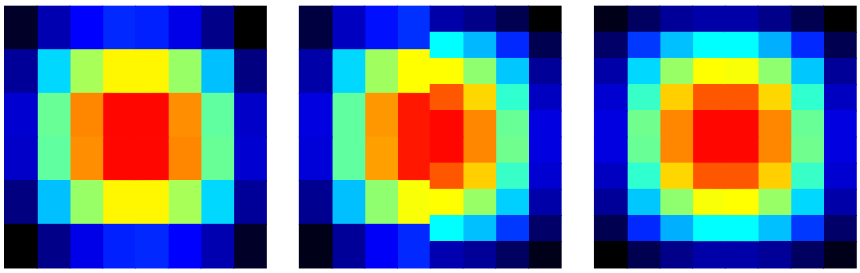

Figure 2: The approximate solution for conforming and nonconforming meshes. Left: conforming 8×6 mesh, center: non conforming $4 \times 6, 4 \times 10$ mesh, right: conforming 10×10.

Further detailed results on several problems and conforming, non conforming and distorted meshes may be found in [17].

3.2 The case of a highly heterogeneous tilted barrier

We now turn to a heterogeneous case. The domain $\Omega =]0,1[\times]0,1[$ is composed of 3 subdomains, which are depicted in Figure 3: $\Omega_1 = \{(x,y) \in \Omega; \varphi_1(x,y) < 0\}$, with $\varphi_1(x,y) = y - \delta(x - .5) - .475$, $\Omega_2 = \{(x,y) \in \Omega; \varphi_1(x,y) > 0, \varphi_2(x,y) < 0\}$, with $\varphi_2(x,y) = \varphi_1(x,y) - 0.05$, $\Omega_3 = \{(x,y) \in \Omega; \varphi_2(x,y) > 0\}$, and $\delta = 0.2$ is the slope of the drain (see Figure 3). Dirichlet boundary conditions are imposed by setting the boundary values to those of the analytical solution given by $u(x,y) = -\varphi_1(x,y)$ on $\Omega_1 \cup \Omega_3$ and $u(x,y) = -\varphi_1(x,y)/10^{-2}$ on Ω_2. The permeability tensor Λ is heterogeneous and isotropic, given by $\Lambda(\boldsymbol{x}) = \lambda(\boldsymbol{x})\mathrm{Id}$, with $\lambda(\boldsymbol{x}) = 1$ for a.e. $x \in \Omega_1 \cup \Omega_3$ and $\lambda(\boldsymbol{x}) = 10^{-2}$ for a.e. $x \in \Omega_2$. Note that the isolines of the exact solution are parallel to the boundaries of the subdomain, and that the tangential component of the gradient is 0. We use the meshes depicted in figure 3. Mesh 3 (containing 10×25 control volumes) is obtained from Mesh 1 by the addition of two layers of very thin control volumes around each of the two lines of discontinuity of Λ: because of the very low thickness of these layers, equal to 1/10000, the picture representing Mesh 3 is not different from that of Mesh 1.

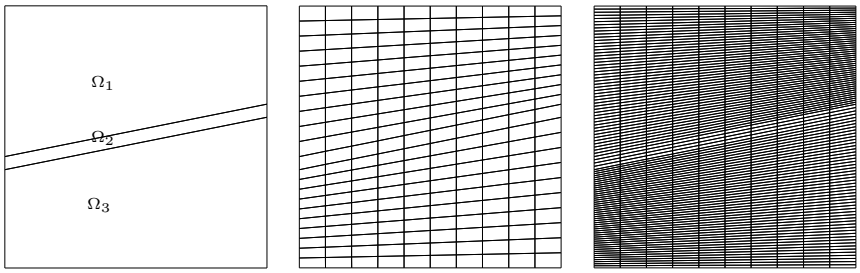

Figure 3: Domain and meshes used for the tilted barrier test: mesh 1 (10×21 center), mesh 2 (10×100 right)

We get the following results for the approximations of the four fluxes at the boundary.

		nb. unknowns	matrix size	$x = 0$	$x = 1$	$y = 0$	$y = 1$
analytical				-0.2	0.2	1.	-1.
centred	mesh 1	210	2424	-1.17	1.17	3.51	-3.51
	mesh 2	1000	11904	-0.237	0.237	1.104	-1.104
	mesh 3	250	2904	-0.208	0.208	1.02	-1.02
composite	mesh 1	239	2583	-0.2	0.2	1.	-1.
	mesh 2	1020	12036	-0.2	0.2	1.	-1.
hybrid	mesh 1	599	4311	-0.2	0.2	1.	-1.
	mesh 2	2890	21138	-0.2	0.2	1.	-1.

Note that the values of the numerical solution given by the hybrid and composite schemes are equal to those of the analytical one (this holds under the only condition that the interfaces located on the lines $\varphi_i(x,y) = 0$, $i = 1,2$ are not included in $\mathcal{B}$, and that, for all $\sigma \in \mathcal{B}$, all $K \in \mathcal{M}$ with $\beta_\sigma^K \neq 0$ are included in the same subdomain Ω_i). Note that Mesh 3, which leads to acceptable results for the computation of the fluxes, is not well designed for such a coupled problem, because of too small measures of control volumes. Hence the composite method on Mesh 1 appears to be the most suitable method for this problem.

4 Convergence study

Let us first introduce some notations related to the mesh. Let $\mathcal{D} = (\mathcal{M}, \mathcal{E}, \mathcal{P})$ be a discretization of Ω in the sense of definition 2.1. The size of the discretization $\mathcal{D}$ is defined by:

$$h_{\mathcal{D}} = \sup\{\mathrm{diam}(K), K \in \mathcal{M}\},$$

and the regularity of the mesh by:

$$\theta_{\mathcal{D}} = \max\left(\max_{\sigma\in\mathcal{E}_{\rm int},K,L\in\mathcal{M}_\sigma} \frac{d_{K,\sigma}}{d_{L,\sigma}}, \max_{K\in\mathcal{M},\sigma\in\mathcal{E}_K} \frac{h_K}{d_{K,\sigma}}\right). \tag{31}$$

For a given set $\mathcal{B} \subset \mathcal{E}_{\rm int}$ and for a given family $(\beta_\sigma^K)_{\substack{K\in\mathcal{M}\\ \sigma\in\mathcal{E}_{\rm int}}}$ satisfying property (15), we introduce some measure of the resulting regularity with

$$\theta_{\mathcal{D},\mathcal{B}} = \max\left(\theta_{\mathcal{D}}, \max_{K\in\mathcal{M},\sigma\in\mathcal{E}_K\cap\mathcal{B}} \frac{\sum_{L\in\mathcal{M}} |\beta_\sigma^L||\boldsymbol{x}_L - \boldsymbol{x}_\sigma|^2}{h_K^2}\right). \tag{32}$$

Remark 4.1 *Note that, for any mesh, it is easy to choose the family $(\beta_\sigma^K)_{\substack{K\in\mathcal{M}\\ \sigma\in\mathcal{E}_{\rm int}}}$ such that $\theta_{\mathcal{D},\mathcal{B}}$ remains small. It suffices to express $\boldsymbol{x}_\sigma$ as the barycenter of $d+1$ points $\boldsymbol{x}_L$ (which is always possible), for L sufficiently close to K, so that $\boldsymbol{x}_L - \boldsymbol{x}_\sigma$ is close to h_K when $\beta_\sigma^K \neq 0$. Note also that in fact, it would be sufficient to have h_K^η with $\eta > 1$ instead of h_K^2 in (32) thus allowing farther points.*

Remark that, thanks to the assumption that K is $\boldsymbol{x}_K$-star-shaped, the property

$$\sum_{\sigma\in\mathcal{E}_K} \mathrm{m}(\sigma) d_{K,\sigma} = d\ \mathrm{m}(K),\ \forall K \in \mathcal{M} \tag{33}$$

holds.
The space $X_{\mathcal{D}}$ defined in (10) is equipped with the following semi-norm:

$$\forall v \in X_{\mathcal{D}},\ |v|_X^2 = \sum_{K\in\mathcal{M}} \sum_{\sigma\in\mathcal{E}_K} \frac{\mathrm{m}(\sigma)}{d_{K,\sigma}} (v_\sigma - v_K)^2, \tag{34}$$

which is a norm on the spaces $X_{\mathcal{D},0}$ and $X_{\mathcal{D},\mathcal{B}}$ respectively defined by (11) and (17).
Let $H_{\mathcal{M}}(\Omega) \subset L^2(\Omega)$ be the set of piecewise constant functions on the control volumes of the mesh $\mathcal{M}$. We then denote, for all $v \in H_{\mathcal{M}}(\Omega)$ and for all $\sigma \in \mathcal{E}_{\rm int}$ with $\mathcal{M}_\sigma = \{K, L\}$, $D_\sigma v = |v_K - v_L|$ and $d_\sigma = d_{K,\sigma} + d_{L,\sigma}$, and for all $\sigma \in \mathcal{E}_{\rm ext}$ with $\mathcal{M}_\sigma = \{K\}$, we denote $D_\sigma v = |v_K|$ and $d_\sigma = d_{K,\sigma}$. We then define the following norm:

$$\forall v \in H_{\mathcal{M}}(\Omega),\ \|v\|_{1,2,\mathcal{M}} = \sum_{K\in\mathcal{M}} \sum_{\sigma\in\mathcal{E}_K} \mathrm{m}(\sigma) d_{K,\sigma} (\frac{D_\sigma v}{d_\sigma})^2 = \sum_{\sigma\in\mathcal{E}} \mathrm{m}(\sigma) \frac{(D_\sigma v)^2}{d_\sigma}. \tag{35}$$

(Note that this norm is also defined by (72) in Lemma 5.2, setting $p = 2$).
For all $v \in X_{\mathcal{D}}$, we denote by $\Pi_{\mathcal{M}} v \in H_{\mathcal{M}}(\Omega)$ the piecewise function from Ω to $\mathbb{R}$ defined by $\Pi_{\mathcal{M}} v(\boldsymbol{x}) = v_K$ for a.e. $\boldsymbol{x} \in K$, for all $K \in \mathcal{M}$. Using the Cauchy-Schwarz inequality, we have for all $\sigma \in \mathcal{E}_{\rm int}$ with $\mathcal{M}_\sigma = \{K, L\}$,

$$\frac{(v_K - v_L)^2}{d_\sigma} \leq \frac{(v_K - v_\sigma)^2}{d_{K,\sigma}} + \frac{(v_\sigma - v_L)^2}{d_{L,\sigma}},\ \forall v \in X_{\mathcal{D}},$$

which leads to the relation

$$\|\Pi_{\mathcal{M}} v\|_{1,2,\mathcal{M}}^2 \leq |v|_X^2,\ \forall v \in X_{\mathcal{D},0}. \tag{36}$$

For all $\varphi \in C(\Omega, \mathbb{R})$, we denote by $P_{\mathcal{D}}\varphi$ the element of $X_{\mathcal{D}}$ defined by $((\varphi(\boldsymbol{x}_K))_{K\in\mathcal{M}}, (\varphi(\boldsymbol{x}_\sigma))_{\sigma\in\mathcal{E}})$, by $P_{\mathcal{D},\mathcal{B}}\varphi$ the element $v \in X_{\mathcal{D},\mathcal{B}}$ such that $v_K = \varphi(\boldsymbol{x}_K)$ for all $K \in \mathcal{M}$, $v_\sigma = 0$ for all $\sigma \in \mathcal{E}_{\rm ext}$, $v_\sigma = \sum_{K\in\mathcal{M}} \beta_\sigma^K \varphi(\boldsymbol{x}_K)$ for all $\sigma \in \mathcal{B}$ and $v_\sigma = \varphi(\boldsymbol{x}_\sigma)$ for all $\sigma \in \mathcal{E}_{\rm int} \setminus \mathcal{B}$.
We denote by $P_{\mathcal{M}}\varphi \in H_{\mathcal{M}}(\Omega)$ the function such that $P_{\mathcal{M}}\varphi(\boldsymbol{x}) = \varphi(\boldsymbol{x}_K)$ for a.e. $\boldsymbol{x} \in K$, for all $K \in \mathcal{M}$ (we then have $P_{\mathcal{M}}\varphi = \Pi_{\mathcal{M}} P_{\mathcal{D}}\varphi = \Pi_{\mathcal{M}} P_{\mathcal{D},\mathcal{B}}\varphi$).
The following lemma provides an equivalence property between the L^2-norm of the discrete gradient, defined by (20)-(24) and the norm $|\cdot|_X$.

Lemma 4.1 *Let $\mathcal{D}$ be a discretization of Ω in the sense of Definition 2.1, and let $\theta \geq \theta_{\mathcal{D}}$ be given (where $\theta_{\mathcal{D}}$ is defined by* (31)*). Then there exists $C_1 > 0$ and $C_2 > 0$ only depending on θ and d such that:*

$$C_1 |u|_X \leq \|\nabla_{\mathcal{D}} u\|_{L^2(\Omega)} \leq C_2 |u|_X, \ \forall u \in X_{\mathcal{D}}, \tag{37}$$

where $\nabla_{\mathcal{D}}$ is defined by (20)-(24)*.*

Proof. By definition,

$$\|\nabla_{\mathcal{D}} u\|_{L^2(\Omega)^d}^2 = \sum_{K\in\mathcal{M}} \sum_{\sigma\in\mathcal{E}_K} \frac{\mathrm{m}(\sigma) d_{K,\sigma}}{d} |\nabla_{K,\sigma} u|^2.$$

From the definition (23), thanks to (21) and to the definition (20), we get that

$$\sum_{\sigma\in\mathcal{E}_K} \frac{\mathrm{m}(\sigma) d_{K,\sigma}}{d} R_{K,\sigma} u \ \boldsymbol{n}_{K,\sigma} = 0, \forall K \in \mathcal{M}. \tag{38}$$

Therefore,

$$\|\nabla_{\mathcal{D}} u\|_{L^2(\Omega)^d}^2 = \sum_{K\in\mathcal{M}} \left(\mathrm{m}(K) |\nabla_K u|^2 + \sum_{\sigma\in\mathcal{E}_K} \frac{\mathrm{m}(\sigma) d_{K,\sigma}}{d} (R_{K,\sigma} u)^2 \right). \tag{39}$$

Let us now notice that the following inequality holds:

$$(a-b)^2 \geq \frac{\lambda}{1+\lambda} a^2 - \lambda b^2, \ \forall a, b \in \mathbb{R}, \ \forall \lambda > -1. \tag{40}$$

We apply this inequality to $(R_{K,\sigma} u)^2$ for some $\lambda > 0$ and obtain:

$$(R_{K,\sigma} u)^2 \geq \frac{\lambda d}{1+\lambda} \left(\frac{u_\sigma - u_K}{d_{K,\sigma}} \right)^2 - \lambda d |\nabla_K u|^2 \left(\frac{|\boldsymbol{x}_\sigma - \boldsymbol{x}_K|}{d_{K,\sigma}} \right)^2. \tag{41}$$

This leads to

$$\sum_{\sigma\in\mathcal{E}_K} \frac{\mathrm{m}(\sigma) d_{K,\sigma}}{d} (R_{K,\sigma} u)^2 \geq \frac{\lambda}{1+\lambda} \sum_{\sigma\in\mathcal{E}_K} \mathrm{m}(\sigma) d_{K,\sigma} \left(\frac{u_\sigma - u_K}{d_{K,\sigma}} \right)^2 - \lambda \mathrm{m}(K) d |\nabla_K u|^2 \theta^2.$$

Choosing λ as

$$\lambda = \frac{1}{d\theta^2}, \tag{42}$$

we get that

$$\|\nabla_{\mathcal{D}} u\|_{(L^2(\Omega))^d}^2 \geq \frac{\lambda}{1+\lambda} |u|_X^2,$$

which shows the left inequality of (37).
Let us now prove the right inequality. On one hand, using the definition (20) of $\nabla_K u$ and (33), the Cauchy–Schwarz inequality leads to:

$$|\nabla_K u|^2 \leq (\frac{1}{\mathrm{m}(K)})^2 \sum_{\sigma\in\mathcal{E}_K} \frac{\mathrm{m}(\sigma)}{d_{K,\sigma}}(u_\sigma - u_K)^2 \sum_{\sigma\in\mathcal{E}_K} \mathrm{m}(\sigma) d_{K,\sigma} = \frac{d}{\mathrm{m}(K)} \sum_{\sigma\in\mathcal{E}_K} \frac{\mathrm{m}(\sigma)}{d_{K,\sigma}}(u_\sigma - u_K)^2. \tag{43}$$

On the other hand, by definition (23), and thanks to the regularity of the mesh (50), we have:

$$(R_{K,\sigma}u)^2 \leq 2d\left((\frac{u_\sigma - u_K}{d_{K,\sigma}})^2 + |\nabla_K u|^2 |\frac{\boldsymbol{x}_\sigma - \boldsymbol{x}_K}{d_{K,\sigma}}|^2\right) \leq 2d\left((\frac{u_\sigma - u_K}{d_{K,\sigma}})^2 + \theta^2 |\nabla_K u|^2\right). \tag{44}$$

From (39), (43) et (44), we conclude that the right inequality of (37) holds. □
We can now state a result of weak convergence for the discrete gradient of a sequence of bounded discrete functions.

Lemma 4.2 *Let $\mathcal{F}$ be a family of discretizations in the sense of definition 2.1 such that there exists $\theta > 0$ with $\theta \geq \theta_\mathcal{D}$ for all $\mathcal{D} \in \mathcal{F}$. Let $(u_\mathcal{D})_{\mathcal{D}\in\mathcal{F}}$ be a family of functions, such that:*

- *$u_\mathcal{D} \in X_{\mathcal{D},0}$ for all $\mathcal{D} \in \mathcal{F}$,*
- *there exists $C > 0$ with $|u_\mathcal{D}|_X \leq C$ for all $\mathcal{D} \in \mathcal{F}$,*
- *there exists $u \in L^2(\Omega)$ with $\lim_{h_\mathcal{D}\to 0} \|\Pi_\mathcal{M} u_\mathcal{D} - u\|_{L^2(\Omega)} = 0$,*

Then $u \in H_0^1(\Omega)$ and $\nabla_\mathcal{D} u_\mathcal{D}$ weakly converge in $L^2(\Omega)^d$ to ∇u as $h_\mathcal{D} \to 0$, where the operator $\nabla_\mathcal{D}$ is defined by (20)-(24).

Remark 4.2 *Note that the proof that $u \in H_0^1(\Omega)$ also results from (36), which allows to apply Lemma 5.7 of the appendix in the particular case $p = 2$.*

Proof. Let us prolong $\Pi_\mathcal{M} u_\mathcal{D}$ and $\nabla_\mathcal{D} u_\mathcal{D}$ by 0 outside of Ω. Thanks to Lemma 4.1, up to a subsequence, there exists some function $\boldsymbol{G} \in L^2(\mathbb{R}^d)^d$ such that $\nabla_\mathcal{D} u_\mathcal{D}$ weakly converges in $L^2(\mathbb{R}^d)^d$ to $\boldsymbol{G}$ as $h_\mathcal{D} \to 0$. Let us show that $\boldsymbol{G} = \nabla u$. Let $\boldsymbol{\psi} \in C_c^\infty(\mathbb{R}^d)^d$ be given. Let us consider the term $T_1^\mathcal{D}$ defined by

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

We get that $T_1^\mathcal{D} = T_2^\mathcal{D} + T_3^\mathcal{D}$, with

$$T_2^\mathcal{D} = \sum_{K\in\mathcal{M}} \sum_{\sigma\in\mathcal{E}_K} \mathrm{m}(\sigma)(u_\sigma - u_K)\boldsymbol{n}_{K,\sigma} \cdot \boldsymbol{\psi}_K, \text{ with } \boldsymbol{\psi}_K = \frac{1}{\mathrm{m}(K)} \int_K \boldsymbol{\psi}(\boldsymbol{x}) \mathrm{d}\boldsymbol{x},$$

and

$$T_3^\mathcal{D} = \sum_{K\in\mathcal{M}} \sum_{\sigma\in\mathcal{E}_K} R_{K,\sigma} u \; \boldsymbol{n}_{K,\sigma} \cdot \int_{D_{K,\sigma}} \boldsymbol{\psi}(\boldsymbol{x}) \mathrm{d}\boldsymbol{x}.$$

We compare $T_2^\mathcal{D}$ with $T_4^\mathcal{D}$ defined by

$$T_4^\mathcal{D} = \sum_{K\in\mathcal{M}} \sum_{\sigma\in\mathcal{E}_K} \mathrm{m}(\sigma)(u_\sigma - u_K)\boldsymbol{n}_{K,\sigma} \cdot \boldsymbol{\psi}_\sigma,$$

with

$$\boldsymbol{\psi}_\sigma = \frac{1}{\mathrm{m}(\sigma)} \int_\sigma \boldsymbol{\psi}(\boldsymbol{x}) \mathrm{d}\gamma(\boldsymbol{x}).$$

We get that

$$(T_2^{\mathcal{D}} - T_4^{\mathcal{D}})^2 \leq \sum_{K\in\mathcal{M}}\sum_{\sigma\in\mathcal{E}_K} \frac{\mathrm{m}(\sigma)}{d_{K,\sigma}}(u_\sigma - u_K)^2 \sum_{K\in\mathcal{M}}\sum_{\sigma\in\mathcal{E}_K} \mathrm{m}(\sigma) d_{K,\sigma}|\boldsymbol{\psi}_K - \boldsymbol{\psi}_\sigma|^2,$$

which leads to $\lim_{h_{\mathcal{D}}\to 0}(T_2^{\mathcal{D}} - T_4^{\mathcal{D}}) = 0$.
Since

$$T_4^{\mathcal{D}} = -\sum_{K\in\mathcal{M}}\sum_{\sigma\in\mathcal{E}_K} \mathrm{m}(\sigma) u_K \boldsymbol{n}_{K,\sigma}\cdot\boldsymbol{\psi}_\sigma = -\int_{\mathbb{R}^d} \Pi_{\mathcal{M}} u_{\mathcal{D}}(\boldsymbol{x})\mathrm{div}\boldsymbol{\psi}(\boldsymbol{x})\mathrm{d}\boldsymbol{x},$$

we get that $\lim_{h_{\mathcal{D}}\to 0} T_4^{\mathcal{D}} = -\int_{\mathbb{R}^d} u(\boldsymbol{x})\mathrm{div}\boldsymbol{\psi}(\boldsymbol{x})\mathrm{d}\boldsymbol{x}$. Let us now turn to the study of $T_3^{\mathcal{D}}$. Noting again that (38) holds, we have:

$$T_3^{\mathcal{D}} = \sum_{K\in\mathcal{M}}\sum_{\sigma\in\mathcal{E}_K} R_{K,\sigma} u\ \boldsymbol{n}_{K,\sigma}\cdot\int_{D_{K,\sigma}} (\boldsymbol{\psi}(\boldsymbol{x}) - \boldsymbol{\psi}_K)\mathrm{d}\boldsymbol{x}.$$

Since $\boldsymbol{\psi}$ is a regular function, there exists $C_{\boldsymbol{\psi}}$ only depending on $\boldsymbol{\psi}$ such that $|\int_{D_{K,\sigma}}(\boldsymbol{\psi}(\boldsymbol{x}) - \boldsymbol{\psi}_K)\mathrm{d}\boldsymbol{x}| \leq C_{\boldsymbol{\psi}} h_{\mathcal{D}} \dfrac{\mathrm{m}(\sigma) d_{K,\sigma}}{d}$. From (44) and the Cauchy-Schwarz inequality, we thus get:

$$\lim_{h_{\mathcal{D}}\to 0} T_3^{\mathcal{D}} = 0.$$

This proves that the function $\boldsymbol{G} \in L^2(\mathbb{R}^d)^d$ is a.e. equal to ∇u in $\mathbb{R}^d$. Since $u = 0$ outside of Ω, we get that $u \in H_0^1(\Omega)$, and the uniqueness of the limit implies that the whole family $\nabla_{\mathcal{D}} u_{\mathcal{D}}$ weakly converges in $L^2(\mathbb{R}^d)^d$ to ∇u as $h_{\mathcal{D}} \to 0$.
□
Let us now state some strong consistency property of the discrete gradient applied to the interpolation of a regular function.

Lemma 4.3 *Let $\mathcal{D}$ be a discretization of Ω in the sense of Definition 2.1, and let $\theta \geq \theta_{\mathcal{D}}$ be given. Then, for any function $\varphi \in C^2(\overline{\Omega})$, there exists C_3 only depending on d, θ and φ such that:*

$$\|\nabla_{\mathcal{D}} P_{\mathcal{D}}\varphi - \nabla\varphi\|_{(L^\infty(\Omega))^d} \leq C_3\, h_{\mathcal{D}}, \tag{45}$$

where $\nabla_{\mathcal{D}}$ is defined by (20)-(24).

Proof. Taking into account definition (24), and using definition (22), we write:

$$|\nabla_{K,\sigma} P_{\mathcal{D}}\varphi - \nabla\varphi(\boldsymbol{x}_K)| \leq |\nabla_K P_{\mathcal{D}}\varphi - \nabla\varphi(\boldsymbol{x}_K)| + |R_{K,\sigma} P_{\mathcal{D}}\varphi|$$

From (20), we have, for any $K \in \mathcal{M}$,

$$\begin{aligned}\nabla_K P_{\mathcal{D}}\varphi &= \frac{1}{\mathrm{m}(K)}\sum_{\sigma\in\mathcal{E}_K} \mathrm{m}(\sigma)(\varphi(\boldsymbol{x}_\sigma) - \varphi(\boldsymbol{x}_K)) \\ &= \frac{1}{\mathrm{m}(K)}\sum_{\sigma\in\mathcal{E}_K} \mathrm{m}(\sigma)\big(\nabla\varphi(\boldsymbol{x}_K)\cdot(\boldsymbol{x}_\sigma - \boldsymbol{x}_K) + h_K^2\rho_{K,\sigma}\big)\boldsymbol{n}_{K,\sigma},\end{aligned}$$

where $|\rho_{K,\sigma}| \leq C_\varphi$ with C_φ only depending on φ. Thanks to (21) and to the regularity of the mesh, we get:

$$|\nabla_K P_{\mathcal{D}}\varphi - \nabla\varphi(\boldsymbol{x}_K)| \leq \frac{1}{\mathrm{m}(K)}\sum_{\sigma\in\mathcal{E}_K} \mathrm{m}(\sigma) h_K^2 |\rho_{K,\sigma}| \leq h_K\ d\ C_\varphi\theta.$$

From this last inequality, using Definition 23, we get:

$$\begin{aligned} |R_{K,\sigma}P_{\mathcal{D}}\varphi| &= \frac{\sqrt{d}}{d_{K,\sigma}}|\varphi(\boldsymbol{x}_\sigma)-\varphi(\boldsymbol{x}_K)-\nabla_K P_{\mathcal{D}}\varphi\cdot(\boldsymbol{x}_\sigma-\boldsymbol{x}_K)| \\ &\le \frac{\sqrt{d}}{d_{K,\sigma}}\left(h_K^2\rho_{K,\sigma}+h_K^2\ d\ C_\varphi\theta\right) \\ &\le \sqrt{d}\theta(h_K C_\varphi + h_K d C_\varphi\theta), \end{aligned}$$

which concludes the proof.□
We now give the abstract properties of the discrete fluxes which are necessary to prove the convergence of the general scheme (18),(13), and then prove that the fluxes which we constructed in Section 2.4 indeed satisfy these properties.

Definition 4.1 (Continuous, coercive, consistent and symmetric families of fluxes) *Let $\mathcal{F}$ be a family of discretizations in the sense of definition 2.1. For $\mathcal{D}=(\mathcal{M},\mathcal{E},\mathcal{P})\in\mathcal{F}$, $K\in\mathcal{M}$ and $\sigma\in\mathcal{E}$, we denote by $F_{K,\sigma}^{\mathcal{D}}$ a linear mapping from $X_{\mathcal{D}}$ to $\mathbb{R}$, and we denote by $\Phi=((F_{K,\sigma}^{\mathcal{D}})_{\substack{K\in\mathcal{M}\\ \sigma\in\mathcal{E}}})_{\mathcal{D}\in\mathcal{F}}$. We consider the bilinear form defined by*

$$\langle u,v\rangle_F=\sum_{K\in\mathcal{M}}\sum_{\sigma\in\mathcal{E}_K}F_{K,\sigma}^{\mathcal{D}}(u)(v_K-v_\sigma),\ \forall(u,v)\in X_{\mathcal{D}}^2. \tag{46}$$

The family of numerical fluxes Φ is said to be continuous if there exists $M>0$ such that

$$\langle u,v\rangle_F\le M|u|_X|v|_X,\ \forall(u,v)\in X_{\mathcal{D}}^2,\ \forall\mathcal{D}=(\mathcal{M},\mathcal{E},\mathcal{P})\in\mathcal{F}. \tag{47}$$

The family of numerical fluxes Φ is said to be coercive if there exists $\alpha>0$ such that

$$\alpha|u|_X^2\le\langle u,u\rangle_F,\ \forall u\in X_{\mathcal{D}},\ \forall\mathcal{D}=(\mathcal{M},\mathcal{E},\mathcal{P})\in\mathcal{F}. \tag{48}$$

The family of numerical fluxes Φ is said to be consistent (with Problem (1)–(2)) if for any family $(u_{\mathcal{D}})_{\mathcal{D}\in\mathcal{F}}$ sastisfying:

- *$u_{\mathcal{D}}\in X_{\mathcal{D},0}$ for all $\mathcal{D}\in\mathcal{F}$,*
- *there exists $C>0$ with $|u_{\mathcal{D}}|_X\le C$ for all $\mathcal{D}\in\mathcal{F}$,*
- *there exists $u\in L^2(\Omega)$ with $\lim_{h_{\mathcal{D}}\to0}\|\Pi_{\mathcal{M}}u_{\mathcal{D}}-u\|_{L^2(\Omega)}=0$ (recall that, from Lemma 5.7, we get that $u\in H_0^1(\Omega)$),*

then:

$$\lim_{h_{\mathcal{D}}\to0}\langle u_{\mathcal{D}},P_{\mathcal{D}}\varphi\rangle_F=\int_\Omega\Lambda(\boldsymbol{x})\nabla\varphi(\boldsymbol{x})\cdot\nabla u(\boldsymbol{x})\mathrm{d}\boldsymbol{x},\ \forall\varphi\in C_c^\infty(\Omega). \tag{49}$$

Finally the family of numerical fluxes Φ is said to be symmetric if

$$\langle u,v\rangle_F=\langle v,u\rangle_F,\ \forall(u,v)\in X_{\mathcal{D}}^2,\ \forall\mathcal{D}=(\mathcal{M},\mathcal{E},\mathcal{P})\in\mathcal{F}.$$

We now show that the family of fluxes defined by (25)-(28) satisfies the definition of a consistent, coercive and symmetric family of fluxes. Recall that the composite scheme (29) is studied numerically in section 3 under this choice for the family of fluxes.

Lemma 4.4 *Let $\mathcal{F}$ be a family of discretizations in the sense of definition 2.1. We assume that there exists $\theta > 0$ with*

$$\theta_{\mathcal{D}} \leq \theta, \ \forall \mathcal{D} \in \mathcal{F}, \tag{50}$$

where $\theta_{\mathcal{D}}$ is defined by (31)*. Let $\Phi = ((F^{\mathcal{D}}_{K,\sigma})_{\substack{K\in\mathcal{M}^{\mathcal{D}}\\ \sigma\in\mathcal{E}_K}})_{\mathcal{D}\in\mathcal{F}}$ be the family of fluxes defined by* (25)*-*(28)*. Then the family Φ is a continuous, coercive, consistent and symmetric family of numerical fluxes in the sense of definition 4.1.*

Proof. Since the family of fluxes is defined by (25)-(28), it satisfies (19), and therefore we have:

$$\langle u, v\rangle_F = \int_\Omega \nabla_{\mathcal{D}} u(\boldsymbol{x}) \cdot \Lambda(\boldsymbol{x}) \nabla_{\mathcal{D}} v(\boldsymbol{x}) \mathrm{d}\boldsymbol{x}, \ \forall u, v \in X_{\mathcal{D}}.$$

Hence the property $\langle u, v\rangle_F = \langle v, u\rangle_F$ holds. The continuity and coercivity of the family Φ result from Lemma 4.1 and the properties of Λ, which give: $\langle u, v\rangle_F \leq \overline{\lambda} \|\nabla_{\mathcal{D}} u\|_{L^2(\Omega)} \|\nabla_{\mathcal{D}} v\|_{L^2(\Omega)}$ and $\langle u, v\rangle_F \geq \underline{\lambda} \|\nabla_{\mathcal{D}} u\|^2_{L^2(\Omega)}$ for any $u, v \in X_{\mathcal{D}}$. The consistency results from the weak and strong convergence properties of lemmas 4.2 and 4.3, which give $\nabla_{\mathcal{D}} u_{\mathcal{D}} \to \nabla u$ weakly in $L^2(\Omega)$ and $\nabla_{\mathcal{D}} P_{\mathcal{D}} \varphi \to \nabla \varphi$ in $L^2(\Omega)$ as the mesh size tends to 0. □

We may now state the general convergence theorem.

Theorem 4.1 *Let $\mathcal{F}$ be a family of discretizations in the sense of definition 2.1; for any $\mathcal{D} \in \mathcal{F}$, let $\mathcal{B} \subset \mathcal{E}_{\rm int}$ and $(\beta^K_\sigma)_{\substack{K\in\mathcal{M}\\ \sigma\in\mathcal{E}_{\rm int}}}$ satisfying property* (15)*. We assume that there exists $\theta > 0$ such that $\theta_{\mathcal{D},\mathcal{B}} \leq \theta$, for all $\mathcal{D} \in \mathcal{F}$, where $\theta_{\mathcal{D},\mathcal{B}}$ is defined by* (32)*. Let $\Phi = ((F^{\mathcal{D}}_{K,\sigma})_{\substack{K\in\mathcal{M}\\ \sigma\in\mathcal{E}}})_{\mathcal{D}\in\mathcal{F}}$ be a continuous, coercive and symmetric and consistent family of numerical fluxes in the sense of definition 4.1. Let $(u_{\mathcal{D}})_{\mathcal{D}\in\mathcal{F}}$ be the family of functions solution to* (18) *for all $\mathcal{D} \in \mathcal{F}$. Then $\Pi_{\mathcal{M}} u_{\mathcal{D}}$ converges in $L^2(\Omega)$ to the unique solution u of* (6) *as $h_{\mathcal{D}} \to 0$. Moreover $\nabla_{\mathcal{D}} u_{\mathcal{D}}$ converges to ∇u in $L^2(\Omega)^d$ as $h_{\mathcal{D}} \to 0$.*

Proof. We let $v = u_{\mathcal{D}}$ in (18), we apply the Cauchy-Schwarz inequality to the right hand side. We get

$$\langle u_{\mathcal{D}}, u_{\mathcal{D}}\rangle_F = \int_\Omega f(\boldsymbol{x}) \Pi_{\mathcal{M}} u_{\mathcal{D}}(\boldsymbol{x}) \mathrm{d}\boldsymbol{x} \leq \|f\|_{L^2(\Omega)} \|\Pi_{\mathcal{M}} u_{\mathcal{D}}\|_{L^2(\Omega)}.$$

We apply the Sobolev inequality (75) with $p = 2$, which gives in this case $\|\Pi_{\mathcal{M}} u_{\mathcal{D}}\|_{L^2(\Omega)} \leq C_4 \|\Pi_{\mathcal{M}} u_{\mathcal{D}}\|_{1,2,\mathcal{M}}$. Using (36) and the coercivity (48) of the family Φ of fluxes, we then have

$$\alpha |u_{\mathcal{D}}|^2_X \leq C_4 \, \|f\|_{L^2(\Omega)} |u_{\mathcal{D}}|_X.$$

This leads to the inequality

$$\|\Pi_{\mathcal{M}} u_{\mathcal{D}}\|_{1,2,\mathcal{M}} \leq |u_{\mathcal{D}}|_X \leq \frac{C_4}{\alpha} \|f\|_{L^2(\Omega)}. \tag{51}$$

Thanks to lemma 5.7, we get the existence of $u \in H^1_0(\Omega)$, and of a subfamily extracted from $\mathcal{F}$, such that $\|\Pi_{\mathcal{M}} u_{\mathcal{D}} - u\|_{L^2(\Omega)}$ tends to 0 as $h_{\mathcal{D}} \to 0$. For a given $\varphi \in C^\infty_c(\Omega)$, let us take $v = P_{\mathcal{D},\mathcal{B}} \varphi$ in (18) (recall that $P_{\mathcal{D},\mathcal{B}} \varphi \in X_{\mathcal{D},\mathcal{B}}$). We get

$$\langle u_{\mathcal{D}}, P_{\mathcal{D},\mathcal{B}} \varphi\rangle_F = \int_\Omega f(\boldsymbol{x}) P_{\mathcal{M}} \varphi(\boldsymbol{x}) \mathrm{d}\boldsymbol{x}.$$

Let us remark that, thanks to the continuity of the family Φ of fluxes, we have

$$\langle u_{\mathcal{D}}, P_{\mathcal{D},\mathcal{B}} \varphi - P_{\mathcal{D}} \varphi\rangle_F \leq M \frac{C_4}{\alpha} \|f\|_{L^2(\Omega)} \ |P_{\mathcal{D},\mathcal{B}} \varphi - P_{\mathcal{D}} \varphi|_X.$$

Thanks to (15) and (32), we get the existence of C_φ, only depending on φ such that, for all $K \in \mathcal{M}$ and all $\sigma \in \mathcal{B} \cap \mathcal{E}_K$,

$$|\sum_{L\in\mathcal{M}} \beta^L_\sigma \varphi(\boldsymbol{x}_L) - \varphi(\boldsymbol{x}_\sigma)| \leq \sum_{L\in\mathcal{M}} |\beta^L_\sigma| |\boldsymbol{x}_L - \boldsymbol{x}_\sigma|^2 C_\varphi \leq \theta_{\mathcal{D},\mathcal{B}} C_\varphi h^2_K. \tag{52}$$

We can then deduce

$$\lim_{h_{\mathcal{D}}\to 0} |P_{\mathcal{D},\mathcal{B}}\varphi - P_{\mathcal{D}}\varphi|_X = 0. \tag{53}$$

Thanks to the properties of subfamily extracted from $\mathcal{F}$, we can apply the consistency hypothesis on the family Φ of fluxes, which gives

$$\lim_{h_{\mathcal{D}}\to 0} \langle u_{\mathcal{D}}, P_{\mathcal{D}}\varphi\rangle_F = \int_\Omega \Lambda(\boldsymbol{x})\nabla\varphi(\boldsymbol{x})\cdot\nabla u(\boldsymbol{x})\mathrm{d}\boldsymbol{x}.$$

Gathering the two above results leads to

$$\lim_{h_{\mathcal{D}}\to 0} \langle u_{\mathcal{D}}, P_{\mathcal{D},\mathcal{B}}\varphi\rangle_F = \int_\Omega \Lambda(\boldsymbol{x})\nabla\varphi(\boldsymbol{x})\cdot\nabla u(\boldsymbol{x})\mathrm{d}\boldsymbol{x},$$

which concludes the proof that

$$\int_\Omega \Lambda(\boldsymbol{x})\nabla\varphi(\boldsymbol{x})\cdot\nabla u(\boldsymbol{x})\mathrm{d}\boldsymbol{x} = \int_\Omega f(\boldsymbol{x})\varphi(\boldsymbol{x})\mathrm{d}\boldsymbol{x}.$$

Therefore, u is the unique solution of (6), and we get that the whole family $(u_{\mathcal{D}})_{\mathcal{D}\in\mathcal{F}}$ converges to u as $h_{\mathcal{D}}\to 0$.

Let us now prove the second part of the theorem, which we begin by proving in the particular (simple) case where the family of fluxes is defined by (25)-(28). In such a case, we get from Lemma 4.2 that $\nabla_{\mathcal{D}}u_{\mathcal{D}}$ weakly converges to ∇u in $L^2(\Omega)^d$. Therefore we have

$$\int_\Omega \Lambda(\boldsymbol{x})\nabla u(\boldsymbol{x})\cdot\nabla u(\boldsymbol{x})\mathrm{d}\boldsymbol{x} \le \liminf_{h_{\mathcal{D}}\to 0}\int_\Omega \Lambda(\boldsymbol{x})\nabla_{\mathcal{D}}u_{\mathcal{D}}(\boldsymbol{x})\cdot\nabla_{\mathcal{D}}u_{\mathcal{D}}(\boldsymbol{x})\mathrm{d}\boldsymbol{x}. \tag{54}$$

Therefore, noting that, passing to the limit $h_{\mathcal{D}}\to 0$ in the scheme (18), we get:

$$\begin{aligned}\limsup_{h_{\mathcal{D}}\to 0}\int_\Omega \Lambda(\boldsymbol{x})\nabla_{\mathcal{D}}u_{\mathcal{D}}(\boldsymbol{x})\cdot\nabla_{\mathcal{D}}u_{\mathcal{D}}(\boldsymbol{x})\mathrm{d}\boldsymbol{x} &= \limsup_{h_{\mathcal{D}}\to 0}\int_\Omega f(\boldsymbol{x})\Pi_{\mathcal{M}}u_{\mathcal{D}}(\boldsymbol{x})d\boldsymbol{x}\\ &= \int_\Omega f(\boldsymbol{x})u(\boldsymbol{x})d\boldsymbol{x} = \int_\Omega \Lambda(\boldsymbol{x})\nabla u(\boldsymbol{x})\cdot\nabla u(\boldsymbol{x})\mathrm{d}\boldsymbol{x}.\end{aligned}$$

Together with (54), this yields that $\int_\Omega \Lambda(\boldsymbol{x})\nabla_{\mathcal{D}}u_{\mathcal{D}}(\boldsymbol{x})\cdot\nabla_{\mathcal{D}}u_{\mathcal{D}}(\boldsymbol{x})\mathrm{d}\boldsymbol{x}$ tends to $\int_\Omega \Lambda(\boldsymbol{x})\nabla u(\boldsymbol{x})\cdot\nabla u(\boldsymbol{x})\mathrm{d}\boldsymbol{x}$ as the mesh size tends to 0. Thanks to the weak convergence of the discrete gradient in $L^2(\Omega)^d$, we obtain that

$$\int_\Omega \Lambda(\boldsymbol{x})(\nabla_{\mathcal{D}}u_{\mathcal{D}}(\boldsymbol{x}) - \nabla u(\boldsymbol{x}))\cdot(\nabla_{\mathcal{D}}u_{\mathcal{D}}(\boldsymbol{x}) - \nabla u(\boldsymbol{x}))\mathrm{d}\boldsymbol{x} \to 0,$$

which allows to conclude that $\nabla_{\mathcal{D}}u_{\mathcal{D}}(\boldsymbol{x})$ tends to $\nabla u(\boldsymbol{x})$ in $L^2(\Omega)^d$.

We now turn to the proof in the case of a general family of continuous, coercive, symmetric and consistent family of numerical fluxes in the sense of definition 4.1. Let $\varphi\in \mathrm{C}_c^\infty(\Omega)$ be given (this function is devoted to approximate u in $H_0^1(\Omega)$). Thanks to the Cauchy-Schwarz inequality, we have

$$\int_\Omega |\nabla_{\mathcal{D}}u_{\mathcal{D}}(\boldsymbol{x}) - \nabla u(\boldsymbol{x})|^2\mathrm{d}\boldsymbol{x} \le 3\ (T_5^{\mathcal{D}} + T_6^{\mathcal{D}} + T_7)$$

where

$$T_5^{\mathcal{D}} = \int_\Omega |\nabla_{\mathcal{D}}u_{\mathcal{D}}(\boldsymbol{x}) - \nabla_{\mathcal{D}}P_{\mathcal{D}}\varphi(\boldsymbol{x})|^2\mathrm{d}\boldsymbol{x},$$

$$T_6^{\mathcal{D}} = \int_\Omega |\nabla_{\mathcal{D}}P_{\mathcal{D}}\varphi(\boldsymbol{x}) - \nabla\varphi(\boldsymbol{x})|^2\mathrm{d}\boldsymbol{x},$$

and

$$T_7 = \int_\Omega |\nabla\varphi(\boldsymbol{x}) - \nabla u(\boldsymbol{x})|^2 \mathrm{d}\boldsymbol{x}.$$

We have, thanks to Lemma 4.3,

$$\lim_{h_\mathcal{D}\to 0} T_6^\mathcal{D} = 0. \tag{55}$$

We have, thanks to Lemma 4.1 and to the coercivity of the family of fluxes, that there exists C_5 such that

$$\|\nabla_\mathcal{D} v\|^2_{L^2(\Omega)^d} \leq {C_2}^2 |v|^2_X \leq C_5 \langle v, v\rangle_F, \forall v \in X_\mathcal{D},$$

with $C_5 = \frac{{C_2}^2}{\alpha}$. Taking $v = u_\mathcal{D} - P_\mathcal{D}\varphi$, we have

$$T_5^\mathcal{D} \leq C_5 \left(\langle u_\mathcal{D}, u_\mathcal{D}\rangle_F - 2\langle u_\mathcal{D}, P_\mathcal{D}\varphi\rangle_F + \langle P_\mathcal{D}\varphi, P_\mathcal{D}\varphi\rangle_F\right).$$

Using the result of convergence proved for $u_\mathcal{D}$ and the consistency of the family of fluxes, we get

$$\lim_{h_\mathcal{D}\to 0} \langle u_\mathcal{D}, P_\mathcal{D}\varphi\rangle_F = \int_\Omega \nabla u(\boldsymbol{x}) \cdot \Lambda(\boldsymbol{x}) \nabla\varphi(\boldsymbol{x}) \mathrm{d}\boldsymbol{x}. \tag{56}$$

Since it is " that $|P_\mathcal{D}\varphi|_X$ remains bounded, using the regularity of φ and the regularity hypotheses of the family of discretizations, we can use the consistency of the family of fluxes, which writes in this case

$$\lim_{h_\mathcal{D}\to 0} \langle P_\mathcal{D}\varphi, P_\mathcal{D}\varphi\rangle_F = \int_\Omega \nabla\varphi(\boldsymbol{x}) \cdot \Lambda(\boldsymbol{x}) \nabla\varphi(\boldsymbol{x}) \mathrm{d}\boldsymbol{x}. \tag{57}$$

Remarking that passing to the limit $h_\mathcal{D} \to 0$ in (18) with $v = u_\mathcal{D}$ provides that $\langle u_\mathcal{D}, u_\mathcal{D}\rangle_F$ converges to $\int_\Omega \nabla u \cdot \Lambda \nabla u \mathrm{d}\boldsymbol{x}$, we get that

$$\lim_{h_\mathcal{D}\to 0} \langle u_\mathcal{D} - P_\mathcal{D}\varphi, u_\mathcal{D} - P_\mathcal{D}\varphi\rangle_F = \int_\Omega \nabla(u-\varphi) \cdot \Lambda \nabla(u-\varphi) \mathrm{d}\boldsymbol{x} \leq \overline{\lambda} \int_\Omega |\nabla u - \nabla\varphi|^2 \mathrm{d}\boldsymbol{x},$$

which yields

$$\limsup_{h_\mathcal{D}\to 0} T_5^\mathcal{D} \leq C_5\, \overline{\lambda} \int_\Omega |\nabla u - \nabla\varphi|^2 \mathrm{d}\boldsymbol{x}.$$

From the above results, we obtain that there exists C_6, independent of $\mathcal{D}$, such that

$$\int_\Omega |\nabla_\mathcal{D} u_\mathcal{D}(\boldsymbol{x}) - \nabla u(\boldsymbol{x})|^2 \mathrm{d}\boldsymbol{x} \leq C_6 \int_\Omega |\nabla\varphi(\boldsymbol{x}) - \nabla u(\boldsymbol{x})|^2 \mathrm{d}\boldsymbol{x} + T_8^\mathcal{D},$$

with (noting that φ is fixed)

$$\lim_{h_\mathcal{D}\to 0} T_8^\mathcal{D} = 0. \tag{58}$$

Let $\varepsilon > 0$. We can choose φ such that $\int_\Omega |\nabla\varphi(\boldsymbol{x}) - \nabla u(\boldsymbol{x})|^2 \mathrm{d}\boldsymbol{x} \leq \varepsilon$, and we can then choose $h_\mathcal{D}$ small enough such that $T_8^\mathcal{D} \leq \varepsilon$. This completes the proof that

$$\lim_{h_\mathcal{D}\to 0} \int_\Omega |\nabla_\mathcal{D} u_\mathcal{D}(\boldsymbol{x}) - \nabla u(\boldsymbol{x})|^2 \mathrm{d}\boldsymbol{x} = 0 \tag{59}$$

in the case of a general continuous, coercive, consistent and symmetric family of fluxes. □
Let us write an error estimate, in the particular case that $\Lambda = \mathrm{Id}$ and that the solution of (6) is regular enough.

Theorem 4.2 (Error estimate, isotropic case) *We consider the particular case $\Lambda = \mathrm{Id}$, and we assume that the solution $u \in H_0^1(\Omega)$ of (6) is in $C^2(\overline{\Omega})$. Let $\mathcal{D} = (\mathcal{M}, \mathcal{E}, \mathcal{P})$ be a discretization in the sense of definition 2.1, let $\mathcal{B} \subset \mathcal{E}_{\rm int}$ be given, let $\boldsymbol{\mathcal{B}} = (\beta_\sigma^K)_{\sigma\in\mathcal{B},K\in\mathcal{M}}$ be a family of real numbers such that (15) holds, and let $\theta \geq \theta_{\mathcal{D},\mathcal{B}}$ be given (see (32)). Let $(F_{K,\sigma})_{K\in\mathcal{M},\sigma\in\mathcal{E}}$ be a family of linear mappings from $X_\mathcal{D}$ to $\mathbb{R}$, such that there exists $\alpha > 0$ with*

$$\alpha |u|_X^2 \leq \langle u, u\rangle_F, \ \forall u \in X_\mathcal{D}, \tag{60}$$

defining $\langle u, v\rangle_F$ by (46). We denote by

$$Ru = \left(\sum_{K\in\mathcal{M}} \sum_{\sigma\in\mathcal{E}_K} \frac{d_{K,\sigma}}{\mathrm{m}(\sigma)} \left(F_{K,\sigma}(P_{\mathcal{D},\mathcal{B}} u) + \int_\sigma \nabla u(\boldsymbol{x}) \cdot \boldsymbol{n}_{K,\sigma} \mathrm{d}\gamma(\boldsymbol{x}) \right)^2 \right)^{1/2}. \tag{61}$$

Then the solution $u_\mathcal{D}$ of (18) verifies that there exists C_7, only depends on α and on θ, such that

$$\|\Pi_\mathcal{M} u_\mathcal{D} - P_\mathcal{M} u\|_{L^2(\Omega)} \leq C_7 \, Ru, \tag{62}$$

and verifies that there exists C_8, only depending on α, θ and u such that

$$\|\nabla_\mathcal{D} u_\mathcal{D} - \nabla u\|_{L^2(\Omega)^d} \leq C_8 \, (Ru + h_\mathcal{D}). \tag{63}$$

Moreover, in the particular case where $(F_{K,\sigma})_{K\in\mathcal{M},\sigma\in\mathcal{E}}$ is defined by (25)-(28), there exists C_9, only depending on α, θ and u, such that

$$Ru \leq C_9 \, h_\mathcal{D}. \tag{64}$$

Remark 4.3 *The extension of Theorem 4.2 to the case $u \in H^2(\Omega)$ could be studied in the case $d = 2$ or $d = 3$. Such a study, which demands a rather longer and more technical proof, is not expected to provide more information on the link between accuracy and the regularity of the mesh than the result presented here.*

Proof. Let $v \in X_\mathcal{D}$, since $-\Delta u = f$, we get:

$$-\sum_{K\in\mathcal{M}} v_K \int_K \Delta u(\boldsymbol{x}) \mathrm{d}\boldsymbol{x} = \int_\Omega f(\boldsymbol{x}) \Pi_\mathcal{M} v(\boldsymbol{x}) \mathrm{d}\boldsymbol{x}.$$

Thanks to the following equality (recall that $u \in C^2(\overline{\Omega})$ and therefore $\nabla u \cdot \boldsymbol{n}_{K,\sigma}$ is defined on each edge σ)

$$-\sum_{K\in\mathcal{M}} v_K \int_K \Delta u(\boldsymbol{x}) \mathrm{d}\boldsymbol{x} = -\sum_{K\in\mathcal{M}} \sum_{\sigma\in\mathcal{E}_K} (v_K - v_\sigma) \int_\sigma \nabla u(\boldsymbol{x}) \cdot \boldsymbol{n}_{K,\sigma} \mathrm{d}\gamma(\boldsymbol{x}),$$

we get that

$$\langle P_{\mathcal{D},\mathcal{B}} u, v\rangle_F = \int_\Omega f(\boldsymbol{x}) \Pi_\mathcal{M} v(\boldsymbol{x}) \mathrm{d}\boldsymbol{x} + \sum_{K\in\mathcal{M}} \sum_{\sigma\in\mathcal{E}_K} \left(F_{K,\sigma}^\mathcal{D}(P_{\mathcal{D},\mathcal{B}} u) + \int_\sigma \nabla u(\boldsymbol{x}) \cdot \boldsymbol{n}_{K,\sigma} \mathrm{d}\gamma(\boldsymbol{x}) \right) (v_K - v_\sigma).$$

Taking $v = P_{\mathcal{D},\mathcal{B}} u - u_\mathcal{D} \in X_{\mathcal{D},\mathcal{B}}$ in this latter equality and using (18) we get

$$\langle v, v\rangle_F = \sum_{K\in\mathcal{M}} \sum_{\sigma\in\mathcal{E}_K} \left(F_{K,\sigma}^\mathcal{D}(P_{\mathcal{D},\mathcal{B}} u) + \int_\sigma \nabla u(\boldsymbol{x}) \cdot \boldsymbol{n}_{K,\sigma} \mathrm{d}\gamma(\boldsymbol{x}) \right) (v_K - v_\sigma),$$

which leads, using (60) and the Cauchy-Schwarz inequality, to

$$\alpha |v|_X \leq Ru. \tag{65}$$

Using (36) and the Sobolev inequality (75) with $p = 2$ provides the conclusion of (62). Let us now prove (63). We have

$$\|\nabla_{\mathcal{D}} u_{\mathcal{D}} - \nabla u\|_{L^2(\Omega)^d} \leq \|\nabla_{\mathcal{D}} u_{\mathcal{D}} - \nabla_{\mathcal{D}} P_{\mathcal{D},\mathcal{B}} u\|_{L^2(\Omega)^d} + \|\nabla_{\mathcal{D}} P_{\mathcal{D},\mathcal{B}} u - \nabla u\|_{L^2(\Omega)^d}.$$

The bound of the first term in the above right hand side is bounded thanks to Lemma 4.1 and (65). The inequality $\|\nabla_{\mathcal{D}} P_{\mathcal{D},\mathcal{B}} u - \nabla u\|_{L^2(\Omega)^d} \leq C_{10} h_{\mathcal{D}}$ is obtained thanks to Lemma 4.3 and using a similar inequality to (52), replacing φ by u.

Let us now turn to the proof of (64) in the particular case where the family of fluxes is defined by (25)-(28). Indeed, we get in this case that, for all $v \in X_{\mathcal{D}}$,

$$F_{K,\sigma}(v) = -\sum_{\sigma' \in \mathcal{E}_K} (\nabla_K v + R_{K,\sigma'} v \; \boldsymbol{n}_{K,\sigma'}) \cdot \frac{\mathrm{m}(\sigma') d_{K,\sigma'}}{d} \boldsymbol{y}^{\sigma'\sigma},$$

with

$$\boldsymbol{y}^{\sigma'\sigma} = \begin{cases} \dfrac{\mathrm{m}(\sigma)}{\mathrm{m}(K)} \boldsymbol{n}_{K,\sigma} + \dfrac{\sqrt{d}}{d_{K,\sigma}} \left(1 - \dfrac{\mathrm{m}(\sigma)}{\mathrm{m}(K)} \boldsymbol{n}_{K,\sigma} \cdot (\boldsymbol{x}_\sigma - \boldsymbol{x}_K)\right) \boldsymbol{n}_{K,\sigma} & \text{if } \sigma = \sigma' \\ \dfrac{\mathrm{m}(\sigma)}{\mathrm{m}(K)} \boldsymbol{n}_{K,\sigma} - \dfrac{\sqrt{d}}{d_{K,\sigma'} \mathrm{m}(K)} \mathrm{m}(\sigma) \boldsymbol{n}_{K,\sigma} \cdot (\boldsymbol{x}_{\sigma'} - \boldsymbol{x}_K) \boldsymbol{n}_{K,\sigma'} & \text{otherwise .} \end{cases}$$

Using (21), we get that

$$\sum_{\sigma' \in \mathcal{E}_K} \frac{\mathrm{m}(\sigma') d_{K,\sigma'}}{d} \boldsymbol{y}^{\sigma'\sigma} = \mathrm{m}(\sigma) \boldsymbol{n}_{K,\sigma}.$$

Since it is easy to see that there exists $C_{11} \in \mathbb{R}_+$ such that $|R_{K,\sigma'} P_{\mathcal{D},\mathcal{B}} u| \leq C_{11} \, h_K$, we then obtain that there exists some $C_{12} \in \mathbb{R}_+$ with

$$\left| F_{K,\sigma}(P_{\mathcal{D},\mathcal{B}} u) + \int_\sigma \nabla u(\boldsymbol{x}) \cdot \boldsymbol{n}_{K,\sigma} \mathrm{d}\gamma(\boldsymbol{x}) \right| \leq C_{12} \, \mathrm{m}(\sigma) h_K.$$

This easily leads to the conclusion of (64). □

5 Discrete functional analysis

This section is devoted to some functional analysis results which are useful for the proof of convergence of numerical schemes when the approximate solution is piecewise constant on each cell of the mesh. Although some of the results presented here were already introduced in previous works ot the authors, they were mostly presented (even when not needed, see [13, Remark 9.13 p. 793]) in the framework of "admissible" meshes, that is meshes with an orthogonality condition.
We recall that in the proof of the main convergence theorem 4.1, we first obtain from the scheme some estimates on the approximate solutions in the discrete H^1 norm. We now show how, from a general discrete $W^{1,p}$ estimate (this generalization to $p \neq 2$ is useful in the case of non linear problems) we obtain a discrete L^q estimate for some $q > p$ (Lemma 5.3). We then obtain some compactness in L^1 (Lemma 5.5 and therefore in L^p (Lemma 5.6), which in turn allows to show that the limit of the approximate solution is in $W_0^{1,p}(\Omega)$ (Lemma 5.7).

5.1 Discrete Sobolev embedding

5.1.1 Discrete embedding of $W^{1,1}$ in $L^{1^\star}$

Definition 5.1 (Polyhedral partition of Ω) *Let $d \geq 1$ and Ω be an open bounded set of $\mathbb{R}^d$, whose boundary is a finite union of part of hyperplanes. A polyhedral partition of Ω is a finite partition of Ω such that each element of this partition is measurable and has a boundary which is composed of a finite union of part of*

hyperplanes. Let $\mathcal{M}$ be such a partition and $\mathcal{E}$ be the set of interfaces of this partition. If $\sigma \in \mathcal{E}$ is an interface (or an edge) of this partition, one denotes by $\mathrm{m}(\sigma)$ the $d-1$–Lebesgue measure of σ. Let $H_{\mathcal{M}}(\Omega)$ be the set of functions from Ω to $\mathbb{R}$, constant on each element of $\mathcal{M}$. Let $u \in H_{\mathcal{M}}(\Omega)$. If $\sigma \in \mathcal{E}$ is a common interface to K, $L \in \mathcal{M}_\sigma$, (that is $\sigma = \overline{K} \cap \overline{L}$), one sets $D_\sigma u = |u_K - u_L|$. If $\sigma \in \mathcal{E}$ is on the boundary of Ω and $K \in \mathcal{M}$ (that is $\sigma = \partial\Omega \cap \overline{K}$), one sets $D_\sigma u = |u_K|$. For $u \in H_{\mathcal{M}}(\Omega)$, one sets :

$$\|u\|_{1,1,\mathcal{M}} = \sum_{\sigma \in \mathcal{M}} \mathrm{m}(\sigma) D_\sigma u. \tag{66}$$

Lemma 5.1 *Let $d \geq 1$ and Ω be an open bounded set of $\mathbb{R}^d$, whose boundary is a finite union of part of hyperplanes. Let $\mathcal{M}$ be a polyhedral partition of Ω (see Definition 5.1). Then, with the notation of Definition 5.1 :*

$$\|u\|_{L^{1^\star}(\Omega)} \leq \frac{1}{2\sqrt{d}} \|u\|_{1,1,\mathcal{M}}, \ \ \forall u \in H_{\mathcal{M}}(\Omega), \tag{67}$$

where $1^\star = \frac{d}{d-1}$.

Proof.
Different proofs of this lemma are possible. A first proof consists in adapting to this discrete setting the classical proof of the Sobolev embedding due to L. Nirenberg (actually, it gives $1/2$ instead of $1/(2\sqrt{d})$ in (67)): it is based on an induction on d. This proof is essentially given in [13, Lemma 9.5 page 790], with slightly less general hypotheses; however, an easy adaptation of the proof leads to the present lemma (with $1/2$ instead of $1/(2\sqrt{d})$ in (67)). We give here another proof by directly using L. Nirenberg's result, namely:

$$\|u\|_{L^{1^\star}(\mathbb{R}^d)} \leq \frac{1}{2d} \|u\|_{W^{1,1}(\mathbb{R}^d)}, \ \ \forall u \in W^{1,1}(\mathbb{R}^d), \tag{68}$$

where $\|u\|_{W^{1,1}(\mathbb{R}^d)} = \sum_{i=1}^d \|D_i u\|_{L^1(\mathbb{R}^d)}$ and $D_i u$ is the weak derivative (or derivative in the sense of distributions) of u in the direction x_i (with $\boldsymbol{x} = (x_1, \ldots, x_d) \in \mathbb{R}^d$).

For $u \in L^1(\mathbb{R}^d)$, one sets $\|u\|_{BV} = \sum_{i=1}^d \|D_i u\|_M$ with, for $i = 1, \ldots, d$, $\|D_i u\|_M = \sup\{\int u \frac{\partial \varphi}{\partial x_i} \mathrm{d}\boldsymbol{x}, \varphi \in C_c^\infty(\mathbb{R}^d), \|\varphi\|_{L^\infty(\mathbb{R}^d)} \leq 1\}$. One says that the function u is in the space BV if $u \in L^1(\mathbb{R}^d)$ and $\|u\|_{BV} < \infty$. We first remark that (68) is true with $\|u\|_{BV}$ instead of $\|u\|_{W^{1,1}(\mathbb{R}^d)}$, and if $u \in BV$ instead of $W^{1,1}(\mathbb{R}^d)$. Indeed, to prove this result (which is classical), let $\rho \in C_c^\infty(\mathbb{R}^d, \mathbb{R}_+)$ with $\int \rho \mathrm{d}\boldsymbol{x} = 1$. For $n \in \mathbb{N}^\star$, define $\rho_n = n^d \rho(n\cdot)$. Let $u \in BV$ and $u_n = u \star \rho_n$ so that, with (68):

$$\|u_n\|_{L^{1^\star}(\mathbb{R}^d)} \leq \frac{1}{2d} \sum_{i=1}^d \|D_i u_n\|_{L^1(\mathbb{R}^d)}. \tag{69}$$

But, $\|D_i u_n\|_{L^1(\mathbb{R}^d)} = \|D_i u_n\|_M$, and, for $\varphi \in C_c^\infty(\mathbb{R}^d)$, using Fubini's theorem:

$$\int_{\mathbb{R}^d} u_n \frac{\partial \varphi}{\partial x_i} \mathrm{d}\boldsymbol{x} = \int_{\mathbb{R}^d} u \frac{\partial}{\partial x_i} (\varphi \star \rho_n) \mathrm{d}\boldsymbol{x} \leq \|D_i u\|_M \|\varphi\|_{L^\infty(\mathbb{R}^d)}.$$

This leads to $\|D_i u_n\|_{L^1(\mathbb{R}^d)} \leq \|D_i u\|_M$. Since $u_n \to u$ a.e., as $n \to \infty$, at least for a subsequence, Fatou's lemma gives, from (69):

$$\|u\|_{L^{1^\star}(\mathbb{R}^d)} \leq \frac{1}{2d} \|u\|_{BV} \ \forall u \in BV. \tag{70}$$

Let $u \in H_{\mathcal{M}}(\Omega)$. One sets $u = 0$ outside Ω so that $u \in L^1(\mathbb{R}^d)$. One has $\|u\|_{BV} = \sup\{\int_{\mathbb{R}^d} u \ \mathrm{div}\varphi \ \mathrm{d}\boldsymbol{x}$, $\varphi \in C_c^\infty(\mathbb{R}^d, \mathbb{R}^d)$, $\|\varphi\|_{L^\infty(\mathbb{R}^d)} \leq 1\}$, with $\|\varphi\|_{L^\infty(\mathbb{R}^d)} = \sup_{i=1,\ldots,d} \|\varphi_i\|_{L^\infty(\mathbb{R}^d)}$ and $\varphi = (\varphi_1, \ldots, \varphi_d)$. But, for

$\varphi \in C_c^\infty(\mathbb{R}^d, \mathbb{R}^d)$ such that $\|\varphi\|_{L^\infty(\mathbb{R}^d)} \leq 1$, an integration by parts on each element of $\mathcal{M}$ gives (where n_σ is a normal vector to σ and γ is the $(d-1)-$Lebesgue measure on σ):

$$\int_{\mathbb{R}^d} u \ \mathrm{div}\varphi \ \mathrm{d}\boldsymbol{x} = \sum_{\sigma\in\mathcal{E}} D_\sigma u \int_\sigma |\varphi \cdot n_\sigma| d\gamma(\boldsymbol{x}) \leq \sqrt{d}\|u\|_{1,1,\mathcal{M}}.$$

Then, one has $\|u\|_{BV} \leq \sqrt{d}\|u\|_{1,1,\mathcal{M}}$ and (70) leads to (67).
□

5.1.2 Discrete embedding of $W^{1,p}$ in $L^{p^\star}$, $1 < p < d$

We now prove a discrete Sobolev embedding for $1 < p < d$ and for meshes in the sense of Definition 2.1.

Lemma 5.2 *Let $d > 1$, $1 < p < d$ and Ω be a polyhedral open bounded connected subset of $\mathbb{R}^d$. Let $\mathcal{D}$ be a mesh of Ω in the sense of Definition 2.1. Let $\eta > 0$ such that $\eta \leq d_{K,\sigma}/d_{L,\sigma} \leq 1/\eta$ for all $\sigma \in \mathcal{E}$, where $\mathcal{M}_\sigma = \{K, L\}$ (see the definitions in Section 4 and Definition 5.1). Then, there exists C_{13}, only depending on d, p and η such that:*

$$\|u\|_{L^{p^\star}(\Omega)} \leq C_{13} \|u\|_{1,p,\mathcal{M}} \ \forall u \in H_\mathcal{D}(\Omega), \tag{71}$$

where $p^\star = \frac{pd}{d-p}$ and

$$\|u\|_{1,p,\mathcal{M}}^p = \sum_{K\in\mathcal{M}} \sum_{\sigma\in\mathcal{E}_K} \mathrm{m}(\sigma) d_{K,\sigma} (\frac{D_\sigma u}{d_\sigma})^p, \tag{72}$$

with $d_\sigma = d_{K,\sigma} + d_{L,\sigma}$, if $\mathcal{M}_\sigma = \{K, L\}$, and $d_\sigma = d_{K,\sigma}$, if $\mathcal{M}_\sigma = \{K\}$.

Proof. We again follow here L. Nirenberg's proof of the Sobolev embeddings. Let α be such that $\alpha 1^\star = p^\star$ (that is $\alpha = p(d-1)/(d-p) > 1$). Let $u \in H_\mathcal{D}(\Omega)$. Inequality (67) applied with $|u|^\alpha$ instead of u leads to:

$$(\int_\Omega |u|^{p^\star} \mathrm{d}\boldsymbol{x})^{\frac{d-1}{d}} \leq \sum_{\sigma\in\mathcal{E}} \mathrm{m}(\sigma) D_\sigma |u|^\alpha.$$

For $\sigma \in \mathcal{E}_{\mathrm{int}}$, $\mathcal{M}_\sigma = \{K, L\}$, one has $D_\sigma|u|^\alpha \leq \alpha(|u_K|^{\alpha-1} + |u_L|^{\alpha-1}) D_\sigma u$. For $\sigma \in \mathcal{E}_{\mathrm{ext}}$, $\mathcal{M}_\sigma = \{K\}$, one has $D_\sigma|u|^\alpha \leq \alpha|u_K|^{\alpha-1} D_\sigma u$. This yields:

$$(\int_\Omega |u|^{p^\star} \mathrm{d}\boldsymbol{x})^{\frac{d-1}{d}} \leq \sum_{K\in\mathcal{M}} \sum_{\sigma\in\mathcal{E}_K} \mathrm{m}(\sigma)\alpha|u_K|^{\alpha-1} D_\sigma u, \tag{73}$$

For all $\sigma \in \mathcal{E}$, one has $1 \leq (1+\eta)(d_{K,\sigma}/d_\sigma)$, if $\sigma \in \mathcal{E}_{\mathrm{int}}$, $\mathcal{M}_\sigma = \{K, L\}$, or if $\sigma \in \mathcal{E}_{\mathrm{ext}}$, $\mathcal{M}_\sigma = \{K\}$. Then, Hölder Inequality applied to (73) yields, with $q = p/(p-1)$:

$$(\int_\Omega |u|^{p^\star} \mathrm{d}\boldsymbol{x})^{\frac{d-1}{d}} \leq \alpha(1+\eta)(\sum_{K\in\mathcal{M}} \sum_{\sigma\in\mathcal{E}_K} \mathrm{m}(\sigma) d_{K,\sigma} |u_K|^{(\alpha-1)q})^{\frac{1}{q}} \|u\|_{1,p,\mathcal{M}}. \tag{74}$$

Since $(\alpha-1)q = p^\star$, one has:

$$\sum_{K\in\mathcal{M}} \sum_{\sigma\in\mathcal{E}_K} \mathrm{m}(\sigma) d_{K,\sigma} |u_K|^{(\alpha-1)q} = d \int_\Omega |u|^{p^\star} \mathrm{d}\boldsymbol{x}.$$

Then, noticing that $d/(d-1) - 1/q = 1/p^\star$, we deduce (71) follows from (74) with $C_{13} = \alpha(1+\eta)d^{1/q}$ only depending on d, p and η. □

5.1.3 Discrete embedding of $W^{1,p}$ in L^q, for some $q > p$

Let $1 \leq p < \infty$, we now deduce from lemma 5.3 the following Lemma, which gives the discrete embedding of $W^{1,p}$ in L^q, for some $q > p$.

Lemma 5.3 *Let $d \geq 1$, $1 \leq p < \infty$ and Ω be a polyhedral open bounded connected subset of $\mathbb{R}^d$. Let $\mathcal{D}$ be a mesh of Ω in the sense of Definition 2.1. Let $\eta > 0$ such that $\eta \leq d_{K,\sigma}/d_{L,\sigma} \leq 1/\eta$ for all $\sigma \in \mathcal{E}$, where $\mathcal{M}_\sigma = \{K, L\}$. Then, there exists $q > p$ only depending on p and there exists C_{14}, only depending on d, Ω, p and η such that (the definitions in Section 4 and Definition 5.1):*

$$\|u\|_{L^q(\Omega)} \leq C_{14} \|u\|_{1,p,\mathcal{M}} \ \forall u \in H_{\mathcal{D}}(\Omega), \tag{75}$$

where $\|u\|^p_{1,p,\mathcal{M}}$ is defined in (72).

Proof. If $p = 1$, one takes $q = 1^\star$ and the result follows from lemma 5.1 (in this case C_{14} does not depend on η). If $1 < p < d$, one takes $q = p^\star$ and the result is in lemma 5.2.

If $p \geq d$, one chooses any $q \in]p, \infty[$ and $p_1 < d$ such that $p_1^\star = q$ (this is possible since $p_1^\star$ tends to ∞ as p_1 tends to d). lemma 5.2 gives, for some C_{13} only depending on p, d and η, $\|u\|_{L^q(\Omega)} \leq C_{13} \|u\|_{1,p_1,\mathcal{M}}$. But, using Hölder inequality, there existe C_{15}, only depending on d, p, Ω, such that $\|u\|_{1,p_1,\mathcal{M}} \leq C_{15} \|u\|_{1,p,\mathcal{M}}$. Inequality (5.3) follows with $C_{14} = C_{13} C_{15}$. □

5.2 Compactness results for bounded families in discrete $W^{1,p}$ norm

5.2.1 Compactness in L^p

We prove in this section that bounded families in the discrete $W^{1,p}$ norms are relatively compact in L^p. We begin here also with the case $p = 1$, giving in this case a crucial inequality which holds for general polyhedral partitions of Ω.

Lemma 5.4 *Let $d \geq 1$ and Ω be an open bounded set of $\mathbb{R}^d$, whose boundary is a finite union of part of hyperplanes. Let $\mathcal{M}$ be a polyhedral partition of Ω (see Definition 5.1). Then, with the notation of Definition 5.1 :*

$$\|u(\cdot + \boldsymbol{y}) - u\|_{L^1(\mathbb{R}^d)} \leq |\boldsymbol{y}|\sqrt{d}\|u\|_{1,1,\mathcal{M}}, \ \ \forall u \in H_{\mathcal{M}}(\Omega), \ \forall \boldsymbol{y} \in \mathbb{R}^d, \tag{76}$$

where u is defined on the whole $\mathbb{R}^d$, taking $u = 0$ outside Ω, and $|h|$ is the euclidean norm of $h \in \mathbb{R}^d$.

Proof. As in lemma 5.1, a proof of this result is possible with a method similar to the method for proving compactness results for bounded families in the discrete $W^{1,p}$ norms, given in [13] (and we obtain (76) without $\sqrt{d}$). Indeed, this proof of [13] holds here in this case of a general partition, thanks to the fact that $p = 1$. More restrictive assumptions are needed for the case $p > 1$. We give here another proof, using the $BV-$space, as in lemma 5.1.

Let $u \in C_c^\infty(\mathbb{R}^d)$. For $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^d$, one has:

$$|u(\boldsymbol{x} + \boldsymbol{y}) - u(\boldsymbol{x})| = |\int_0^1 \nabla u(\boldsymbol{x} + t\boldsymbol{y}) \cdot \boldsymbol{y} \mathrm{d}t| \leq |h| \int_0^1 |\nabla u(\boldsymbol{x} + t\boldsymbol{y})| \mathrm{d}t.$$

Integrating with respect to $\boldsymbol{x}$ and using Fubini's Theorem gives the well known result

$$\|u(\cdot + \boldsymbol{y}) - u\|_{L^1(\mathbb{R}^d)} \leq |\boldsymbol{y}| \int_{\mathbb{R}^d} |\nabla u| \mathrm{d}\boldsymbol{x} \leq |\boldsymbol{y}| \sum_{i=1}^d \|D_i u\|_{L^1(\mathbb{R}^d)}, \tag{77}$$

where $\nabla u = (D_1 u, \ldots, D_d u)$. By density of $C_c^\infty(\mathbb{R}^d)$ in $W^{1,1}(\mathbb{R}^d)$, Inequality (77) is also true for $u \in W^{1,1}(\mathbb{R}^d)$.

We proceed now as in lemma 5.1, using the same notations. Let $u \in BV$ and $u_n = u \star \rho_n$. Since $u_n \in W^{1,1}(\mathbb{R}^n)$, Inequality (77) gives, for all $\boldsymbol{y} \in \mathbb{R}^d$, $\|u_n(\cdot + \boldsymbol{y}) - u_n\|_{L^1(\mathbb{R}^d)} \leq |\boldsymbol{y}| \sum_{i=1}^d \|D_i u_n\|_{L^1(\mathbb{R}^d)}$. But, for $i = 1, \ldots, d$, as in lemma 5.1, $\|D_i u_n\|_{L^1(\mathbb{R}^d)} \leq \|D_i u\|_M$. Then, since $u_n \to u$ in $L^1(\mathbb{R}^d)$, as $n \to \infty$, we obtain:

$$\|u(\cdot + \boldsymbol{y}) - u\|_{L^1(\mathbb{R}^d)} \leq |\boldsymbol{y}| \sum_{i=1}^d \|D_i u\|_M = |\boldsymbol{y}| \|u\|_{BV}, \ \forall u \in BV, \ \forall \boldsymbol{y} \in \mathbb{R}^d. \tag{78}$$

Let now $u \in H_{\mathcal{M}}(\Omega)$. One sets $u = 0$ outside Ω so that $u \in L^1(\mathbb{R}^d)$. lemma 5.1 gives $\|u\|_{BV} \leq \sqrt{d}\|u\|_{1,1,\mathcal{M}}$, then:

$$\|u(\cdot + \boldsymbol{y}) - u\|_{L^1(\mathbb{R}^d)} \leq |\boldsymbol{y}|\sqrt{d}\|u\|_{1,1,\mathcal{M}}, \ \forall \boldsymbol{y} \in \mathbb{R}^d.$$

□

An easy consequence of lemmas 5.1 and 5.4 is a compactness result in L^1 given in the following lemma.

Lemma 5.5 *Let $d \geq 1$ and Ω be an open bounded set of $\mathbb{R}^d$, whose boundary is a finite union of part of hyperplanes. Let F be a family of polyhedral partition of Ω (see Definition 5.1). For $\mathcal{M} \in F$, let $u_{\mathcal{M}} \in H_{\mathcal{M}}(\Omega)$ and assume that there exists $C \in \mathbb{R}$ such, for all $\mathcal{M} \in F$, $\|u_{\mathcal{M}}\|_{1,1,\mathcal{M}} \leq C$. Then the family $(u_{\mathcal{M}})_{\mathcal{M} \in F}$ is relatively compact in $L^1(\Omega)$ and also in $L^1(\mathbb{R}^d)$ taking $u_{\mathcal{M}} = 0$ outside Ω.*

Proof. The proof is quite easy. Lemma 5.1 gives that the family $(u_{\mathcal{M}})_{\mathcal{M} \in F}$ is bounded in $L^{1^\star}(\Omega)$. Then, since Ω is bounded, the family $(u_{\mathcal{M}})_{\mathcal{M} \in F}$ is bounded in $L^1(\Omega)$ and also in $L^1(\mathbb{R}^d)$ taking $u_{\mathcal{M}} = 0$ outside Ω. Then, thanks to Kolmogorov Compactness Theorem, lemma 5.4 gives that the family $(u_{\mathcal{M}})_{\mathcal{M} \in F}$ is relatively compact in $L^1(\Omega)$ and also in $L^1(\mathbb{R}^d)$ taking $u_{\mathcal{M}} = 0$ outside Ω. □

For $p > 1$, we need some an additional hypothesis on the meshes, actually given by Definition 2.1 with a "uniform η".

Lemma 5.6 *Let $d \geq 1$, $1 \leq p < \infty$ and Ω be a polyhedral open bounded connected subset of $\mathbb{R}^d$. Let F be a family of meshes of Ω in the sense of Definition 2.1. Let $\eta > 0$ such that, for all $\mathcal{D} \in F$, one has $\eta \leq d_{K,\sigma}/d_{L,\sigma} \leq 1/\eta$ for all $\sigma \in \mathcal{E}$, where $\mathcal{M}_\sigma = \{K, L\}$. For $\mathcal{D} \in F$, let $u_{\mathcal{D}} \in H_{\mathcal{D}}(\Omega)$ and assume that there exists $C \in \mathbb{R}$ such, for all $\mathcal{D} \in F$, $\|u_{\mathcal{D}}\|_{1,p,\mathcal{M}} \leq C$. Then the family $(u_{\mathcal{D}})_{\mathcal{D} \in F}$ is relatively compact in $L^p(\Omega)$ and also in $L^p(\mathbb{R}^d)$ taking $u_{\mathcal{D}} = 0$ outside Ω.*

Proof. Here also the proof is quite simple. Thanks to lemma 5.3 and the fact that Ω is bounded, the family $(u_{\mathcal{D}})_{\mathcal{D} \in F}$ is bounded in $L^1(\Omega)$ and also in $L^1(\mathbb{R}^d)$ taking $u_{\mathcal{D}} = 0$ outside Ω. Thanks, once again, to the fact that Ω is bounded the family $(\|u_{\mathcal{D}}\|_{1,1,\mathcal{M}})_{\mathcal{D} \in F}$ is bounded in $\mathbb{R}$. Then, as in the preceding lemma, Kolmogorov Copmpactness Theorem gives that the family $(\|u_{\mathcal{D}}\|_{1,1,\mathcal{M}})_{\mathcal{D} \in F}$ is relatively compact in $L^1(\Omega)$ and also in $L^1(\mathbb{R}^d)$ taking $u_{\mathcal{D}} = 0$ outside Ω.

In order to conclude we use, once again, lemma 5.3. It gives that the family $(u_{\mathcal{D}})_{\mathcal{D} \in F}$ is bounded in $L^q(\Omega)$ for some $q > p$. With the relative compactness in $L^1(\Omega)$, this leads to the fact that the family $(u_{\mathcal{D}})_{\mathcal{D} \in F}$ is relatively compact in $L^p(\Omega)$ (and then also in $L^p(\mathbb{R}^d)$ taking $u_{\mathcal{D}} = 0$ outside Ω). □

5.2.2 Regularity of the limit

With the hypotheses of lemma 5.6, assume that $u_{\mathcal{D}} \to u$ in L^p as $\text{size}(\mathcal{D}) \to 0$ (lemma 5.6 gives that this is possible, at least for subsequences of sequences of meshes with vanishing size). We prove below that $u \in W_0^{1,p}(\Omega)$.

Lemma 5.7 *Let $d \geq 1$, $1 \leq p < \infty$ and Ω be a polyhedral open bounded connected subset of $\mathbb{R}^d$. Let $(\mathcal{D}_n)_{n \in \mathbb{N}}$ be a family of discretizations of Ω in the sense of Definition 2.1. Let $\eta > 0$ such that, for any discretization $\mathcal{D}_n = (\mathcal{M}_n, \mathcal{E}_n, \mathcal{P}_n)$, one has $\eta \leq d_{K,\sigma}/d_{L,\sigma} \leq 1/\eta$ for all $\sigma \in \mathcal{E}$, where $\mathcal{M}_\sigma = \{K, L\}$. For $n \in \mathbb{N}$, let $u^{(n)} \in H_{\mathcal{D}_n}(\Omega)$ and assume that there exists $C \in \mathbb{R}$ such, for all $n \in \mathbb{N}$, $\|u^{(n)}\|_{1,p,\mathcal{M}_n} \leq C$. Assume also that* $\text{size}(\mathcal{D}_n) \to 0$ *as $n \to \infty$. Then:*

1. *There exists a subsequence of* $(u^{(n)})_{n\in\mathbb{N}}$*, still denoted by* $(u^{(n)})_{n\in\mathbb{N}}$*, and* $u \in L^p(\Omega)$ *such that* $u^{(n)} \to u$ *in* $L^p(\Omega)$ *as* $n \to \infty$*.*

2. $u \in W_0^{1,p}(\Omega)$ *and*

$$\|\nabla u\|_{L^p(\Omega)^d} = \||\nabla u|\|_{L^p(\Omega)} \leq \frac{(1+\eta)d^{\frac{p-1}{p}}}{\eta} C. \tag{79}$$

Proof. The fact that there exists a subsequence of $(u^{(n)})_{n\in\mathbb{N}}$, still denoted by $(u^{(n)})_{n\in\mathbb{N}}$, and $u \in L^p(\Omega)$ such that $u^{(n)} \to u$ in $L^p(\Omega)$ as $n \to \infty$ is a consequence of the relative compactness of $(u^{(n)})_{n\in\mathbb{N}}$ in L^p given in lemma 5.6. Assuming that $u^{(n)} \to u$ in $L^p(\Omega)$ as $n \to \infty$, we have now to prove that $u \in W_0^{1,p}(\Omega)$.

Letting $u^{(n)} = 0$ and $u = 0$ outside Ω, one also has $u^{(n)} \to u$ in $L^p(\mathbb{R}^d)$. The method consists to construct some approximate gradient, namely $\tilde{\nabla}_{\mathcal{D}_n} u^{(n)}$, bounded in $L^p(\Omega)$, equal to 0 outside Ω and converging, at least in the distribution sense, to ∇u.

Step 1 Construction of $\tilde{\nabla}_{\mathcal{D}} u$, for $u \in H_{\mathcal{D}}(\Omega)$, and properties
Let $n \in \mathbb{N}$ and $\mathcal{D} = \mathcal{D}_n$. For this step, one sets $u = u^{(n)}$ (not to be confused with the limit of the sequence $(u^{(n)})_{n\in\mathbb{N}}$). For $\sigma \in \mathcal{E}$, one sets $u_\sigma = 0$ if σ is on the boundary of Ω. Otherwise, one has $\mathcal{M}_\sigma = \{K, L\}$ and we choose a value u_σ between u_K and u_L. (it is possible to choose, for instance; $u_\sigma = \frac{1}{2}(u_K + u_L)$ but any other choice between u_K and u_L is possible). Then, one defines $\tilde{\nabla}_{\mathcal{D}} u$ on $K \in \mathcal{D}$ on the following way:

$$\tilde{\nabla}_{\mathcal{D}} u = \frac{1}{m_K} \sum_{\sigma\in\mathcal{E}_K} \mathrm{m}(\sigma) \boldsymbol{n}_{K,\sigma} (u_\sigma - u_K).$$

The function $\tilde{\nabla}_{\mathcal{D}} u$ is constant on each $K \in \mathcal{M}$ and, on K, using Hölder Inequality:

$$|\tilde{\nabla}_{\mathcal{D}} u|^p \leq \frac{1}{(m_K)^p} (\sum_{\sigma\in\mathcal{E}_K} \mathrm{m}(\sigma) \boldsymbol{n}_{K,\sigma} (u_\sigma - u_K))^p \leq \frac{1}{(m_K)^p} (\sum_{\sigma\in\mathcal{E}_K} \mathrm{m}(\sigma) d_{K,\sigma})^{p-1} \sum_{\sigma\in\mathcal{E}_K} \mathrm{m}(\sigma) d_{K,\sigma} (\frac{D_\sigma u}{d_{K,\sigma}})^p.$$

Since $\sum_{\sigma\in\mathcal{E}_K} \mathrm{m}(\sigma) d_{K,\sigma} = d m_K$, one deduces

$$|\tilde{\nabla}_{\mathcal{D}} u|^p \leq \frac{d^{p-1}}{m_K} \sum_{\sigma\in\mathcal{E}_K} \mathrm{m}(\sigma) d_{K,\sigma} (\frac{D_\sigma u}{d_{K,\sigma}})^p.$$

This gives a L^p- estimate on $\tilde{\nabla}_{\mathcal{D}} u$ in $(L^p(\Omega))^d$ (or in $(L^p(\mathbb{R}^d))^d$, setting $\tilde{\nabla}_{\mathcal{D}} u = 0$ outside Ω), in terms of $\|u\|_{1,p,\mathcal{M}}$, namely:

$$\||\tilde{\nabla}_{\mathcal{D}} u|\|_{L^p} \leq \frac{(1+\eta)d^{\frac{p-1}{p}}}{\eta} \|u\|_{1,p,\mathcal{M}}. \tag{80}$$

In order to prove, in the next step, the convergence of this approximate gradient, we compute now the integral of this gradient against a test function. Let $\varphi \in C_c^\infty(\mathbb{R}^d; \mathbb{R}^d)$, φ_K the mean value of φ on $K \in \mathcal{D}$ and φ_σ the mean value of φ on σ. Then:

$$\int_{\mathbb{R}^d} \tilde{\nabla}_{\mathcal{D}} u \cdot \varphi \mathrm{d}\boldsymbol{x} = \sum_{K\in\mathcal{D}} \sum_{\sigma\in\mathcal{E}_K} \mathrm{m}(\sigma) \boldsymbol{n}_{K,\sigma} (u_\sigma - u_K) \varphi_K = \sum_{K\in\mathcal{D}} \sum_{\sigma\in\mathcal{E}_K} \mathrm{m}(\sigma) \boldsymbol{n}_{K,\sigma} (-u_K) \varphi_\sigma + R(u,\varphi), \tag{81}$$

with

$$R(u,\varphi) = \sum_{K\in\mathcal{D}} \sum_{\sigma\in\mathcal{E}_K} \mathrm{m}(\sigma) \boldsymbol{n}_{K,\sigma} (u_\sigma - u_K)(\varphi_K - \varphi_\sigma).$$

Then there exists C_φ only depending on φ, d, p and Ω such that $|R(u,\varphi)| \leq C_\varphi \mathrm{size}(\mathcal{D}) \|u\|_{1,p,\mathcal{M}}$. Equation (81) can also be written as:

$$\int_{\mathbb{R}^d} \tilde{\nabla}_{\mathcal{D}} u \cdot \varphi \mathrm{d}\boldsymbol{x} = \sum_{K\in\mathcal{D}} \int_K (-u_K)\ \mathrm{div}(\varphi)\ \mathrm{d}\boldsymbol{x} + R(u,\varphi) = -\int_{\mathbb{R}^d} u\ \mathrm{div}(\varphi)\ \mathrm{d}\boldsymbol{x} + R(u,\varphi). \tag{82}$$

Step 2 Convergence of $\tilde{\nabla}_{\mathcal{D}_n} u^{(n)}$ to ∇u and proof of $u \in W_0^{1,p}(\Omega)$.
We consider now the sequence $(u^{(n)})_{n\in\mathbb{N}}$. Inequality (80) gives:

$$|||\tilde{\nabla}_{\mathcal{D}} u^{(n)}|||_{L^p} \leq \frac{(1+\eta)d^{\frac{p-1}{p}}}{\eta} \|u^{(n)}\|_{1,p,\mathcal{M}}.$$

Then, the sequence $(\tilde{\nabla}_{\mathcal{D}} u^{(n)})_{n\in\mathbb{N}}$ is bounded in $L^p(\mathbb{R}^d)^d$ and we can assume, up to a subsequence, that $\tilde{\nabla}_{\mathcal{D}} u^{(n)}$ converges to some w weakly in $L^p(\mathbb{R}^d)^d$, as $n \to \infty$ and $|||w|||_{L^p} \leq \frac{(1+\eta)d^{\frac{p-1}{p}}}{\eta} C$.

Let $\varphi \in C_c^\infty(\mathbb{R}^d;\mathbb{R}^d)$, Equation (82) gives

$$\int_{\mathbb{R}^d} \tilde{\nabla}_{\mathcal{D}} u^{(n)} \cdot \varphi \mathrm{d}\boldsymbol{x} = -\int_{\mathbb{R}^d} u^{(n)}\ \mathrm{div}(\varphi)\ \mathrm{d}\boldsymbol{x} + R(u^{(n)},\varphi). \tag{83}$$

Thanks to $|R(u^{(n)},\varphi)| \leq C_\varphi \mathrm{size}(\mathcal{D}_n)\|u^{(n)}\|_{1,p,\mathcal{M}_n}$, one has $R(u^{(n)},\varphi) \to 0$, as $n \to \infty$. Since $u^{(n)} \to u$ in $L^p(\mathbb{R}^d)$ as $n \to \infty$, passing to the limit in (83) gives:

$$\int_{\mathbb{R}^d} w \cdot \varphi \mathrm{d}\boldsymbol{x} = -\int_{\mathbb{R}^d} u\ \mathrm{div}(\varphi)\ \mathrm{d}\boldsymbol{x}.$$

Since φ is arbitrary in $C_c^\infty(\mathbb{R}^d,\mathbb{R}^d)$, one deduces that $\nabla u = w$. Then $u \in W^{1,p}(\mathbb{R}^d)$ and $|||\nabla u|||_{L^p} \leq \frac{(1+\eta)d^{\frac{p-1}{p}}}{\eta} C$. Finally, since $u = 0$ outside Ω, one has $u \in W_0^{1,p}(\Omega)$. □

6 Conclusion and perspectives

A discretization scheme was introduced for anisotropic heterogeneous problems on distorted nonconforming meshes. Although this scheme stems from the finite volume analysis which was developed these past years, its formulation is actually derived from a discrete weak formulation; in this respect it may be seen a non conforming finite element method. Discrete analysis tools were obtained which allow a mathematical analysis of the scheme; the convergence of the discrete solution to the exact solution of the continuous problem is shown with no regularity assumption on the solution (other than the natural assumption that it is in $H_0^1(\Omega)$). Even though this convergence result yields no rate of convergence, it is problably more interesting than error estimates which require some assumptions on the diffusion tensor. Nevertheless, we show an order 1 estimate in the case of the Laplace operator, which is readily extendable to regular (say piecewise C^1) isotropic diffusion operators. The numerical results presented here show the high performance of the scheme (in particular order 2 is obtained for the convergence in the L^2 norm of the solution), and so do three dimensional experiments which were performed in [9] for the incompressible Navier–Stokes equations on general grids. Note that the convergence analysis which is performed here readily extends to the nonlinear setting of Leray-Schauder operators. This will be the object of a future paper.

References

[1] I. Aavatsmark, T. Barkve, O. Boe, and T. Mannseth. Discretization on non-orthogonal, quadrilateral grids for inhomogeneous, anisotropic media. *J. Comput. Phys.*, 127(1):2–14, 1996.

[2] I. Aavatsmark, T. Barkve, O. Boe, and T. Mannseth. Discretization on unstructured grids for inhomogeneous, anisotropic media. part i: Derivation of the methods. *SIAM Journal on Sc. Comp.*, 19:1700–1716, 1998.

[3] I. Aavatsmark, T. Barkve, O. Boe, and T. Mannseth. Discretization on unstructured grids for inhomogeneous, anisotropic media. part ii: Discussion and numerical results. *SIAM Journal on Sc. Comp.*, 19:1717–1736, 1998.

[4] K. Aziz and A. Settari. *Petroleum reservoir simulation*. Applied Science, London, 1979.

[5] F. Benkhaldoun and R. Vilsmeier, editors. *Finite volume methods for diffusion convection equations on general meshes*. Hermès, 1996.

[6] E. Bertolazzi and G. Manzini. On vertex reconstructions for cell-centered finite volume approximations of 2D anisotropic diffusion problems. *Math. Models Methods Appl. Sci.*, 17(1):1–32, 2007.

[7] F. Boyer and Hubert F. Finite volume method for 2d linear and nonlinear elliptic problems wiht discontinuities. submitted, 2006.

[8] F. Brezzi, K. Lipnikov, and M. Shashkov. Convergence of the mimetic finite difference method for diffusion problems on polyhedral meshes. *SIAM J. Numer. Anal.*, 43(5):1872–1896, 2005.

[9] E. Chénier, R. Eymard, and R. Herbin. A collocated finite volume schemes for navier–stokes and energy equations under the boussinesq assumption for general grids. submitted, 2007.

[10] Y. Coudière, J.-P. Vila, and Ph. Villedieu. Convergence rate of a finite volume scheme for a two-dimensional convection-diffusion problem. *M2AN Math. Model. Numer. Anal.*, 33(3):493–516, 1999.

[11] K. Domelevo and P. Omnes. A finite volume method for the laplace equation on almost arbitrary two-dimensional grids. *M2AN Math. Model. Numer. Anal.*, 39(6):1203–1249, 2005.

[12] J. Droniou and R. Eymard. A mixed finite volume scheme for anisotropic diffusion problems on any grid. *Numer. Math.*, 105(1):35–71, 2006.

[13] R. Eymard, T. Gallouët, and R. Herbin. Finite volume methods. In P. G. Ciarlet and J.-L. Lions, editors, *Techniques of Scientific Computing, Part III*, Handbook of Numerical Analysis, VII, pages 713–1020. North-Holland, Amsterdam, 2000.

[14] R. Eymard, T. Gallouët, and R. Herbin. A finite volume scheme for anisotropic diffusion problems. *C. R. Math. Acad. Sci. Paris*, 339(4):299–302, 2004.

[15] R. Eymard, T. Gallouët, and R. Herbin. A cell-centered finite-volume approximation for anisotropic diffusion operators on unstructured meshes in any space dimension. *IMA J. Numer. Anal.*, 26(2):326–353, 2006.

[16] R. Eymard, T. Gallouët, and R. Herbin. A new finite volume scheme for anisotropic diffusion problems on general grids: convergence analysis. *C. R., Math., Acad. Sci. Paris*, 344(6):403–406, 2007.

[17] R. Eymard, T. Gallouët, and R. Herbin. A discretization scheme for anisotropic heterogeneous diffusion problems. Benchmark on discretization schemes for anisotropic diffusion problems on general grids, FVCA5, http://www.latp.univ-mrs.fr/fvca5, 2008.

[18] R. Eymard and R. Herbin. A new colocated finite volume scheme for the incompressible navier-stokes equations on general non matching grids. *C. R. Math. Acad. Sci. Paris*, 344(10):659–662, 2007.

[19] Ph. Guillaume and V. Latocha. Numerical convergence of a parameterisation method for the solution of a highly anisotropic two-dimensional elliptic problem. *J. Sci. Comput.*, 25(3):423–444, 2005.

[20] R. Herbin. An error estimate for a finite volume scheme for a diffusion-convection problem on a triangular mesh. *Numer. Methods Partial Differential Equations*, 11(2):165–173, 1995.

[21] F. Hermeline. Approximation of diffusion operators with discontinuous tensor coefficients on distorted meshes. *Comput. Methods Appl. Mech. Engrg.*, 192(16-18):1939–1959, 2003.

[22] C. Le Potier. Schéma volumes finis monotone pour des opérateurs de diffusion fortement anisotropes sur des maillages de triangles non structurés. *C. R. Math. Acad. Sci. Paris*, 341(12):787–792, 2005.

[23] S.V. Patankar. *Numerical heat transfer and fluid flow.* Series in Computational Methods in Mechanics and Thermal Sciences. Washington - New York - London: Hemisphere Publishing Corporation; New York etc.: McGraw-Hill Book Company. XIII, 197 p., 1980.

[24] J. E. Roberts and J.-M. Thomas. Mixed and hybrid methods. In *Handbook of numerical analysis, Vol. II*, Handb. Numer. Anal., II, pages 523–639. North-Holland, Amsterdam, 1991.

[25] M. Vohralík. Equivalence between mixed finite element and multi-point finite volume methods. *C. R. Acad. Sci. Paris., Ser. I*, 339:525–528, 2004.