IMA Journal of Numerical Analysis (2006) **26**, 326−353
doi:10.1093/imanum/dri036
Advance Access publication on October 7, 2005

A cell-centred finite-volume approximation for anisotropic diffusion operators on unstructured meshes in any space dimension

R. EYMARD†
Université de Marne-la-Vallée, France

AND

T. GALLOUËT‡ AND R. HERBIN§
Université de Provence, France

[Received on 22 April 2005; revised on 15 August 2005]

Finite-volume methods for problems involving second-order operators with full diffusion matrix can be used thanks to the definition of a discrete gradient for piecewise constant functions on unstructured meshes satisfying an orthogonality condition. This discrete gradient is shown to satisfy a strong convergence property for the interpolation of regular functions, and a weak one for functions bounded in a discrete H^1-norm. To highlight the importance of both properties, the convergence of the finite-volume scheme for a homogeneous Dirichlet problem with full diffusion matrix is proven, and an error estimate is provided. Numerical tests show the actual accuracy of the method.

Keywords: anisotropic diffusion; finite-volume methods; discrete gradient; convergence analysis.

1. Introduction

The approximation of convection–diffusion problems in anisotropic media is an important issue in several engineering fields. Let us briefly review four particular situations where the discretization of a non-diagonal second-order operator is required.

1. In the case of a contaminant transported by a single-phase flow, one must account for the diffusion–dispersion operator $\mathrm{div}(\Lambda\nabla u)$, where the matrix $\Lambda(x) = \lambda(x)I_d + \mu(x)\mathbf{q}(x)\cdot\mathbf{q}(x)^t$ depends on the space variable x and $\mathbf{q}(x)$ is the velocity of the fluid flow in the porous medium. The real parameter $\lambda(x)$ corresponds to a resulting isotropic diffusion term, including dispersion in the directions orthogonal to the flow, and the real parameter $\mu(x)$ to an additional diffusion in the direction of the flow (Chénier *et al.*, 2004). The term $\mathbf{q}(x)$ is then given by $\mathbf{q}(x) = -K(x)\nabla p(x)$, where $p(x)$ is a pressure and $K(x)$ another non-diagonal matrix (the absolute permeability matrix, depending on the geological layers), and satisfies the incompressibility equation $\mathrm{div}\mathbf{q}(x) = 0$. In this coupled problem, one must simultaneously compute this pressure and the contaminant concentration $u(x)$.
2. In the study of undersaturated flows in porous media (e.g. air–water flows), two equations of conservation have to be solved, associated with two unknowns, pressure and saturation. These equations include non-linear hyperbolic and degenerate parabolic terms with respect to the saturation

†Email: eymard@math.univ-mlv.fr
‡Email: gallouet@cmi.univ-mrs.fr
§Corresponding author. Email: herbin@cmi.univ-mrs.fr

unknown. As in the preceding case, one must discretize such terms as $\mathrm{div}\mathbf{q}(x) = \mathrm{div}(K(x)\nabla p(x))$, where again $K(x)$ is a non-diagonal matrix depending on the geological layers.

3. In the case of the compressible Navier–Stokes equations, one has to discretize the operator representing viscous forces, which can be written in the form $a\Delta\mathbf{u} + b\nabla\mathrm{div}\mathbf{u}$ (a and b are deduced from the dynamic viscosity coefficients and $\mathbf{u}$ is the fluid velocity). In this problem, the term $\nabla\mathrm{div}\mathbf{u}$ involves all the cross derivatives $\partial^2_{ij}\mathbf{u}$.
4. Some problems arising in financial mathematics lead to anisotropic diffusion equations in high-dimensional domains (e.g. dimension equal to five or more). Under some assumptions on financial markets (Lamberton & Lapeyre, 1995), the price of a European or an American option is obtained by solving a linear or non-linear partial differential equation, involving the second-order anisotropic diffusion matrix $\Lambda = \Sigma\Sigma^t$, where Σ is a real matrix.

All these cases involve a term of the form $\mathrm{div}(\Lambda\nabla u)$, where Λ is a (generally) non-diagonal matrix depending on the space variable and u is a function of the space variable in steady problems and of the space and time variables in transient problems. Finite-element schemes are known to allow for an easy discretization of such a term on triangular or tetrahedral meshes (Putti & Cordes, 1998). However, in engineering situations such as the ones described above, one also has to discretize convection and reaction terms, and avoid numerical instabilities. Unfortunately, classical finite-element methods (and more generally centred schemes) are known to generate instabilities on coarse grids, although some cures have been proposed, see Angermann (2000) and Forsyth (1991); therefore, many numerical codes (Aavatsmark *et al.*, 1998a,b; Forsyth, 1991; Jayantha & Turner, 2003, 2005) use finite-volume or finite-volume–finite-element type schemes, which admit the implementation of discretization techniques (such as the classical upwind schemes) that prevent the appearance of instabilities. Let us also note that finite-volume schemes are known for their simplicity of implementation, particularly so when discretizing coupled systems of equations of various nature.

Besides, a thorough mathematical analysis has now been developed, showing that finite-volume methods are well suited and convergent for a simple convection–diffusion equation in the case where $\Lambda(x) = \lambda(x)I_d$. Indeed, this analysis has been completed (Herbin, 1995; Gallouët *et al.*, 2000; Mishev, 1998; Eymard *et al.*, 1999), see also Eymard *et al.* (2000) for a review, in the case of grids (called admissible in the sense of Eymard *et al.* (2000), see also Definition 2.1) satisfying an orthogonality condition: the line joining two cell centres is orthogonal to the interface between the two cells, thus ensuring a consistency property when approximating the normal flux at the cell interface by centred finite differences. Some examples of such admissible grids are the Delaunay triangular meshes or tetrahedral meshes, rectangular or parallelepiped meshes in 2D or 3D and the Voronoï meshes in any dimension.

But the situation is quite different in the case where the condition $\Lambda(x) = \lambda(x)I_d$ no longer holds: only few of the actual discretization methods used for handling non-diagonal second-order terms on finite-volume grids have a complete mathematical analysis of stability or convergence. Let us briefly review some of them. The first one, in the case where $\Lambda(x) = \lambda(x)M$, where M is a symmetric positive definite matrix, consists of adapting the above orthogonality condition by stating that the line joining two cell centres is orthogonal to the interface between the two cells with respect to the dot product induced by the matrix Λ^{-1}. Indeed, it is also possible to consider the case where M depends on the discretization cell, by using, in each cell, the orthogonal bisectors for the metric induced by M^{-1} (see Herbin, 1996, and Eymard *et al.*, 2000, section 11, p. 815). In the case of triangular grids, this yields a well-defined scheme under some restriction on the admissible anisotropy for a given geometry, since the cell centre is chosen as the intersection of the orthogonal bisectors of the triangle for the metric defined by M^{-1}. Another method consists of defining the finite-volume method as a dual method to a finite-element one (e.g. a P1 finite element (Chénier *et al.*, 2004) or a Crouzeix–Raviart one, see e.g. Eymard *et al.* (2004b)) or, for example, to interpret node-centred finite-volume discretizations in terms of mixed finite-element methods (Angermann, 2003).

Another possibility to derive a finite-volume scheme for problems including anisotropic diffusion is to construct a local discrete gradient, allowing to get, at each edge σ of the mesh, a consistent approximate value for the flux $-\int_\sigma (\Lambda(x)\nabla u(x))\cdot\mathbf{n}_\sigma\, \mathrm{d}\gamma(x)$ involved in the finite-volume scheme ($\mathbf{n}_\sigma$ is a unit vector normal to the edge σ and $\mathrm{d}\gamma(x)$ is the $(d-1)$-dimensional Lebesgue measure on the edge σ). In two space dimensions, such a scheme was introduced in Coudière *et al.* (1999) for parallelograms and generalized to triangles in Bertolazzi & Manzini (2004). However, no proof of convergence is available in the general case. Still in 2D, a technique using dual meshes is introduced in Domelevo & Omnes (2005) and Hermeline (2000), which generalizes the idea of Hu & Nicolaides (1992) and Nicolaides (1992) for div–curl problems to meshes with no orthogonality conditions; however, the use of a dual mesh renders the scheme computationally expensive; moreover, it does not seem to be easily extendable to 3D. Another scheme for 2D and 3D problems, using reconstructed gradients at the vertices, has been introduced in Le Potier (2005); although it leads to symmetric definite positive matrices, the proof of its convergence does not seem to have been actually completed. For problems in any space dimension and for any irregular grid, a mixed finite-volume method is proposed in Droniou & Eymard (2005), involving the simultaneous computation of an approximate gradient and of the unknown; although this method is shown to converge, its drawback is that the linear systems involve unknowns at the edges, in a similar way to mixed hybrid finite-element methods. In Eymard *et al.* (2001), we used Raviart–Thomas shape functions, generalized to the case of any admissible mesh (again in the sense specified in Eymard *et al.* (2000); see also Definition 2.1), in order to define a discrete gradient for piecewise constant functions. The strong convergence of this discrete gradient was then shown in the case of the elliptic equation $-\Delta u = f$. A drawback of this definition was the difficulty to find an approximation of these generalized shape functions in cases other than triangles or rectangles.

We therefore propose in this paper a new cheap and simple method of constructing a discrete gradient for a piecewise constant function, on arbitrary admissible meshes in any space dimension (this method has been first introduced in Eymard *et al.* (2004a)). We prove that the discrete gradients of any sequence of piecewise constant functions converging to some $u \in H_0^1(\Omega)$ weakly converges to ∇u in $L^2(\Omega)$. Moreover, the discrete gradient is shown to be consistent, in the sense that it satisfies a strong convergence property for the interpolation of regular function. In order to show the efficiency of this approximation method, we use this discrete gradient to design a scheme for the approximation of the weak solution $\bar{u}$ of the following diffusion problem with full anisotropic tensor:

$$\begin{aligned} -\mathrm{div}(\Lambda\nabla\bar{u}) &= f, \quad \text{in } \Omega,\\ \bar{u} &= 0, \quad \text{on } \partial\Omega, \end{aligned} \tag{1.1}$$

under the following assumptions:

$$\Omega \text{ is an open bounded connected polygonal subset of } \mathbb{R}^d,\ d \in \mathbb{N}^\star = \mathbb{N}\setminus\{0\}, \tag{1.2}$$

$$\begin{aligned} &\Lambda \text{ is a measurable function from } \Omega \text{ to } \mathcal{M}_d(\mathbb{R}), \text{ where } \mathcal{M}_d(\mathbb{R}) \text{ denotes the set of}\\ &d \times d \text{ matrices, such that for a.e. } x \in \Omega,\ \Lambda(x) \text{ is symmetric,}\\ &\text{the lowest and the largest eigenvalues of } \Lambda(x), \text{ denoted by } \underline{\lambda}(x) \text{ and } \overline{\lambda}(x),\\ &\text{are such that } \underline{\lambda}, \overline{\lambda} \in L^\infty(\Omega) \text{ and there exists } \lambda_0 \in \mathbb{R} \text{ with}\\ &0 < \lambda_0 \leqslant \underline{\lambda}(x) \leqslant \overline{\lambda}(x) \text{ for a.e. } x \in \Omega, \end{aligned} \tag{1.3}$$

and

$$f \in L^2(\Omega). \tag{1.4}$$

We give the classical weak formulation in the following definition.

DEFINITION 1.1 (Weak solution) Under Hypotheses (1.2)–(1.4), we say that $\bar{u}$ is a weak solution of (1.1) if

$$\begin{cases} \bar{u} \in H_0^1(\Omega), \\ \displaystyle\int_\Omega \Lambda(x)\nabla\bar{u}(x)\cdot\nabla v(x)\,\mathrm{d}x = \int_\Omega f(x)v(x)\,\mathrm{d}x \quad \forall v \in H_0^1(\Omega). \end{cases} \tag{1.5}$$

REMARK 1.1 For the sake of clarity, we restrict ourselves here to the numerical analysis of Problem (1.1); however, the present analysis readily extends to convection–diffusion–reaction problems and coupled problems. Indeed, we emphasize that proofs of convergence or error estimates can easily be adapted to such situations, since the discretization methods of all these terms are independent of one another, and the treatment of the convection and reaction terms is well-known to be exact (see Gallouët *et al.*, 2000, or Eymard *et al.*, 2000).

The outline of this paper is as follows. In Section 2, we present the method for approximating the gradient of a piecewise constant function, and we show some functional properties which help to understand why the present definition of a gradient is well suited for second-order diffusion problems. In Section 3, we present the finite-volume scheme for Problem (1.1), and we show the strong convergence of the discrete solution and of its discrete gradient. In Section 3.4, we give an error estimate for Problem (1.1), and we illustrate this study by some numerical examples in Section 4. Some short conclusions are drawn in Section 5.

2. A discrete gradient for piecewise constant functions

We present in this section a method for the approximation of the gradient of piecewise constant functions, in the case of grids satisfying some orthogonality condition as defined below.

2.1 *Admissible discretization of Ω*

We first present the following notion of admissible discretization, which is taken from Eymard *et al.* (2000). The notations are summarized in Fig. 1 for the particular case $d = 2$ (we recall that the case $d \geqslant 3$ is considered as well).

In the following definition, we shall say that a bounded subset of $\mathbb{R}^d$ is polygonal if its boundary is included in the union of a finite number of hyperplanes.

DEFINITION 2.1 [Admissible discretization] Let Ω be an open bounded polygonal subset of $\mathbb{R}^d$ and $\partial\Omega = \overline{\Omega} \setminus \Omega$ its boundary. An admissible finite-volume discretization of Ω, denoted by $\mathcal{D}$, is given by $\mathcal{D} = (\mathcal{M}, \mathcal{E}, \mathcal{P})$, where

- $\mathcal{M}$ is a finite family of non-empty open disjoint convex polygonal 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 and $\mathrm{m}(K) > 0$ denote the measure of K.
- $\mathcal{E}$ is a finite family of disjoint subsets of $\overline{\Omega}$ (the 'edges' of the mesh), such that, for all $\sigma \in \mathcal{E}$, there exists a hyperplane E of $\mathbb{R}^d$ and $K \in \mathcal{M}$ with $\overline{\sigma} = \partial K \cap E$ and σ is a non-empty open subset of E. We then denote by $m_\sigma > 0$ the $(d-1)$-dimensional measure of σ. We assume that, for all $K \in \mathcal{M}$, there exists a subset $\mathcal{E}_K$ of $\mathcal{E}$ such that $\partial K = \cup_{\sigma\in\mathcal{E}_K}\overline{\sigma}$. It then follows from the previous hypotheses

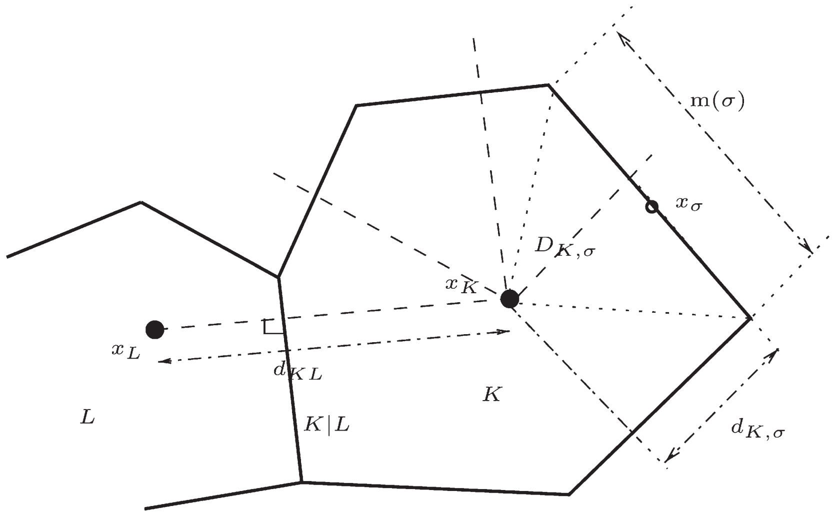

FIG. 1. Notations for a control volume K in the case $d = 2$.

that, for all $\sigma \in \mathcal{E}$, either $\sigma \subset \partial\Omega$ or there exists $(K, L) \in \mathcal{M}^2$ with $K \neq L$ such that $\overline{K} \cap \overline{L} = \overline{\sigma}$; in the latter case we write $\sigma = K|L$.

- $\mathcal{P}$ is a family of points of Ω indexed by $\mathcal{M}$, denoted by $\mathcal{P} = (x_K)_{K\in\mathcal{M}}$. The coordinates of x_K are denoted by $x_K^{(i)}$, $i = 1, \ldots, d$. The family $\mathcal{P}$ is such that, for all $K \in \mathcal{M}$, $x_K \in K$. Furthermore, for all $\sigma \in \mathcal{E}$ such that there exists $(K, L) \in \mathcal{M}^2$ with $\sigma = K|L$, it is assumed that the straight line (x_K, x_L) going through x_K and x_L is orthogonal to $K|L$. For all $K \in \mathcal{M}$ and all $\sigma \in \mathcal{E}_K$, let z_σ be the orthogonal projection of x_K on σ. We suppose that $z_\sigma \in \sigma$ if $\sigma \subset \partial\Omega$.

The following notations are used. The size of the discretization is defined by:

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

For all $K \in \mathcal{M}$ and $\sigma \in \mathcal{E}_K$, we denote by $\mathbf{n}_{K,\sigma}$ the unit vector normal to σ outward to K. We denote by $d_{K,\sigma}$ the Euclidean distance between x_K and σ. We then define

$$\tau_{K,\sigma} = \frac{m_\sigma}{d_{K,\sigma}}.$$

The set of interior (resp. boundary) edges is denoted by $\mathcal{E}_{\mathrm{int}}$ (resp. $\mathcal{E}_{\mathrm{ext}}$), i.e. $\mathcal{E}_{\mathrm{int}} = \{\sigma \in \mathcal{E};\ \sigma \not\subset \partial\Omega\}$ (resp. $\mathcal{E}_{\mathrm{ext}} = \{\sigma \in \mathcal{E};\ \sigma \subset \partial\Omega\}$). For all $K \in \mathcal{M}$, we denote by $\mathcal{N}_K$ the subset of $\mathcal{M}$ of the neighbouring control volumes, and we denote by $\mathcal{E}_{K,\mathrm{ext}} = \mathcal{E}_K \cap \mathcal{E}_{\mathrm{ext}}$. For all $\sigma \in \mathcal{E}_{\mathrm{int}}$, let $K, L \in \mathcal{M}$ be such that $\sigma = K|L$; we define by $d_{K|L}$ the Euclidean distance between x_K and x_L, by $\mathbf{n}_{KL}$ the unit normal vector to $K|L$ from K to L, and we set

$$\tau_\sigma = \frac{m_\sigma}{d_{K|L}}. \tag{2.1}$$

For all $\sigma \in \mathcal{E}_{\mathrm{ext}}$, let $K \in \mathcal{M}$ be such that $\sigma \in \mathcal{E}_K$; we define

$$\tau_\sigma = \tau_{K,\sigma}. \tag{2.2}$$

For all $K \in \mathcal{M}$ and $\sigma \in \mathcal{E}_K$, we define

$$D_{K,\sigma} = \{t x_K + (1-t)y,\ t \in (0,1),\ y \in \sigma\}.$$

For all $\sigma \in \mathcal{E}_{\text{int}}$, let $K, L \in \mathcal{M}$ be such that $\sigma = K|L$; we set $D_\sigma = D_{K,\sigma} \cup D_{L,\sigma}$. For all $\sigma \in \mathcal{E}_{\text{ext}}$, let $K \in \mathcal{M}$ be such that $\sigma \in \mathcal{E}_K$; we define $D_\sigma = D_{K,\sigma}$.

For all $\sigma \in \mathcal{E}$, we define

$$x_\sigma = \frac{1}{\mathrm{m}(\sigma)} \int_\sigma x \, \mathrm{d}\gamma(x). \tag{2.3}$$

We shall measure the regularity of the mesh through the function $\theta_\mathcal{D}$ defined by

$$\theta_\mathcal{D} = \inf \left\{ \frac{d_{K,\sigma}}{\operatorname{diam}(K)}, K \in \mathcal{M}, \ \sigma \in \mathcal{E}_K \right\}. \tag{2.4}$$

DEFINITION 2.2 Let Ω be an open bounded polygonal subset of $\mathbb{R}^d$ and $\mathcal{D}$ an admissible discretization of Ω in the sense of Definition 2.1. We define $H_\mathcal{D}$ as the set of functions $u \in L^2(\Omega)$ which are constant in each control volume. For $u \in H_\mathcal{D}$, we denote by u_K the constant value of u in K. We define the interpolation operator $P_\mathcal{D}$: $C(\overline{\Omega}) \to H_\mathcal{D}$, by $\bar{u} \mapsto P_\mathcal{D}\bar{u}$ such that

$$P_\mathcal{D}\bar{u}(x) = \bar{u}(x_K) \text{ for a.e. } x \in K \quad \forall K \in \mathcal{M}. \tag{2.5}$$

For $(u, v) \in (H_\mathcal{D})^2$ and for any function $\alpha \in L^\infty(\Omega)$, we introduce the following symmetric bilinear form:

$$[u, v]_{\mathcal{D},\alpha} = \sum_{K|L \in \mathcal{E}_{\text{int}}} \tau_{K|L} \alpha_{K|L} (u_L - u_K)(v_L - v_K) + \sum_{K \in \mathcal{M}} \sum_{\sigma \in \mathcal{E}_{K,\text{ext}}} \tau_{K,\sigma} \alpha_\sigma u_K v_K, \tag{2.6}$$

where we set

$$\alpha_\sigma = \frac{1}{\mathrm{m}(D_\sigma)} \int_{D_\sigma} \alpha(x) \, \mathrm{d}x \quad \forall \sigma \in \mathcal{E}. \tag{2.7}$$

Note that (2.7) is not the only possible choice (see Remark 3.1). We then define a norm in $H_\mathcal{D}$ (thanks to the discrete Poincaré inequality (2.8) given below) by

$$\|u\|_\mathcal{D} = ([u, u]_{\mathcal{D},1})^{1/2}$$

(where 1 denotes the constant function equal to 1). Indeed, the discrete Poincaré inequality states that (see Eymard *et al.*, 2000):

$$\|w\|_{L^2(\Omega)} \leqslant \operatorname{diam}(\Omega) \|w\|_\mathcal{D} \quad \forall w \in H_\mathcal{D}. \tag{2.8}$$

Let us now give a relative compactness result, which is also partly stated in some other papers concerning finite-volume methods (Herbin & Marchand, 2001; Eymard *et al.*, 2002).

LEMMA 2.1 (Relative compactness in $L^2(\Omega)$) Let Ω be an bounded open connected polygonal subset of $\mathbb{R}^d$, $d \in \mathbb{N}^\star$ and let $(\mathcal{D}_n, u_n)_{n \in \mathbb{N}}$ be a sequence such that, for all $n \in \mathbb{N}$, $\mathcal{D}_n$ is an admissible finite-volume discretization of Ω in the sense of Definition 2.1 and $u_n \in H_{\mathcal{D}_n}(\Omega)$ (cf. Definition 2.2). Let us assume that $\lim_{n \to \infty} h_{\mathcal{D}_n} = 0$ and that there exists $C_1 > 0$ such that $\|u_n\|_{\mathcal{D}_n} \leqslant C_1$, for all $n \in \mathbb{N}$.

Then there exists a subsequence of $(\mathcal{D}_n, u_n)_{n \in \mathbb{N}}$, again denoted $(\mathcal{D}_n, u_n)_{n \in \mathbb{N}}$, and $\bar{u} \in H_0^1(\Omega)$ such that u_n tends to $\bar{u}$ in $L^2(\Omega)$ as $n \to +\infty$, and the inequality

$$\int_\Omega |\nabla \bar{u}(x)|^2 \, \mathrm{d}x \leqslant \liminf_{n \to \infty} \|u_n\|^2_{\mathcal{D}_n} \tag{2.9}$$

holds. Moreover, for all function $\alpha \in L^\infty(\Omega)$, we have

$$\lim_{n\to\infty} \big[u_n, P_{\mathcal{D}_n}\varphi\big]_{\mathcal{D}_n,\alpha} = \int_\Omega \alpha(x)\nabla\bar{u}(x)\cdot\nabla\varphi(x)\,\mathrm{d}x \quad \forall\varphi\in \mathrm{C}_c^\infty(\Omega). \tag{2.10}$$

Proof. The proof of the existence of the subsequence, again denoted $(\mathcal{D}_n, u_n)_{n\in\mathbb{N}}$, and of $\bar{u}\in H_0^1(\Omega)$ such that u_n tends to $\bar{u}$ in $L^2(\Omega)$ as $n\to\infty$, is given in Eymard *et al.* (2000). Assertion (2.9) was proven in Eymard *et al.* (2002, Lemma 5.2). Let us first show (2.10) in the case $\alpha\in\mathrm{C}^1(\overline{\Omega})$. Let $\varphi\in\mathrm{C}_c^\infty(\Omega)$. Defining, for all $n\in\mathbb{N}$, $T_1^{(n)} = -\int_\Omega u_n(x)\mathrm{div}(\alpha(x)\nabla\varphi(x))\,\mathrm{d}x$, we get that

$$\lim_{n\to\infty} T_1^{(n)} = -\int_\Omega \bar{u}(x)\mathrm{div}(\alpha(x)\nabla\varphi(x))\,\mathrm{d}x = \int_\Omega \alpha(x)\nabla\bar{u}(x)\cdot\nabla\varphi(x)\,\mathrm{d}x.$$

We consider a value n sufficiently large such that for all $K\in\mathcal{M}_n$ and $x\in K$, if $\varphi(x)\neq 0$, then $\partial K\cap\partial\Omega=\emptyset$. Defining $T_2^{(n)} = \big[u_n, P_{\mathcal{D}_n}\varphi\big]_{\mathcal{D}_n,\alpha} - T_1^{(n)}$, we obtain

$$T_2^{(n)} = \sum_{\sigma\in\mathcal{E}_{\mathrm{int}},\sigma=K|L} \mathrm{m}(K|L)(u_L-u_K)R_{KL},$$

with

$$R_{KL} = \alpha_{K|L}\frac{\varphi(x_L)-\varphi(x_K)}{d_{K|L}} - \int_{K|L}\alpha(x)\nabla\varphi(x)\cdot\mathbf{n}_{KL}\,\mathrm{d}\gamma(x) \quad \forall K\in\mathcal{M},\ \forall L\in\mathcal{N}_K.$$

Since there exists some real value C_2, which does not depend on $\mathcal{D}_n$, such that $|R_{KL}|\leqslant C_2 h_{\mathcal{D}_n}$, we conclude in a similar way as in Eymard *et al.* (2000) that $\lim_{n\to\infty} T_2^{(n)} = 0$, which gives (2.10) in this case. Let us now consider the general case $\alpha\in L^\infty(\Omega)$. Let $\varepsilon>0$ be given. We first choose a function $\tilde{\alpha}\in\mathrm{C}^1(\overline{\Omega})$ such that $\|\alpha-\tilde{\alpha}\|_{L^2(\Omega)}\leqslant\varepsilon$. Then we have, for all $n\in\mathbb{N}$, using the Cauchy–Schwarz inequality,

$$\Big(\big[u_n, P_{\mathcal{D}_n}\varphi\big]_{\mathcal{D}_n,\tilde{\alpha}} - \big[u_n, P_{\mathcal{D}_n}\varphi\big]_{\mathcal{D}_n,\alpha}\Big)^2 \leqslant \left(\sum_{K|L\in\mathcal{E}_{\mathrm{int}}} \tau_{K|L}(\tilde{\alpha}_{KL}-\alpha_{KL})^2|\varphi(x_L)-\varphi(x_K)|^2\right) \times \left(\sum_{K|L\in\mathcal{E}_{\mathrm{int}}} \tau_{K|L}|u_L-u_K|^2\right).$$

Therefore, setting $C_3 = \|\nabla\varphi\|_{L^\infty(\Omega)}$, the properties $|\varphi(x_L)-\varphi(x_K)|\leqslant C_3 d_{K|L}$ and $\mathrm{m}(K|L)d_{K|L} = d\mathrm{m}(D_{K|L})$ lead to

$$\big(\big[u_n, P_{\mathcal{D}_n}\varphi\big]_{\mathcal{D}_n,\tilde{\alpha}} - \big[u_n, P_{\mathcal{D}_n}\varphi\big]_{\mathcal{D}_n,\alpha}\big)^2 \leqslant dC_3^2\|\alpha-\tilde{\alpha}\|^2_{L^2(\Omega)}C_1 \leqslant dC_3^2\varepsilon^2 C_1.$$

In the same manner, we get

$$\left(\int_\Omega \tilde{\alpha}(x)\nabla\bar{u}(x)\cdot\nabla\varphi(x)\,\mathrm{d}x - \int_\Omega \alpha(x)\nabla\bar{u}(x)\cdot\nabla\varphi(x)\,\mathrm{d}x\right)^2 \leqslant C_3^2\varepsilon^2\|\nabla\bar{u}\|^2_{L^2(\Omega)^d}.$$

Since $\tilde{\alpha}\in\mathrm{C}^1(\Omega)$, we can apply (2.10), proven above for such a function. It then suffices to choose n large enough such that

$$\left|\big[u_n, P_{\mathcal{D}_n}\varphi\big]_{\mathcal{D}_n,\tilde{\alpha}} - \int_\Omega \tilde{\alpha}(x)\nabla\bar{u}(x)\cdot\nabla\varphi(x)\,\mathrm{d}x\right| \leqslant \varepsilon,$$

to prove that

$$\left|\left[u_n, P_{\mathcal{D}_n}\varphi\right]_{\mathcal{D}_n,\alpha} - \int_\Omega \alpha(x)\nabla\bar{u}(x)\cdot\nabla\varphi(x)\,\mathrm{d}x\right| \leqslant C_4\varepsilon,$$

where the real $C_4 > 0$ does not depend on n. This concludes the proof of (2.10) in the general case. □

2.2 *Definition of a discrete gradient*

We now define a discrete gradient for piecewise constant functions on an admissible discretization.

DEFINITION 2.3 (Discrete gradient) Let Ω be an bounded open connected polygonal subset of $\mathbb{R}^d$, $d \in \mathbb{N}^\star$. Let $\mathcal{D} = (\mathcal{M}, \mathcal{E}, \mathcal{P})$ be an admissible finite-volume discretization of Ω in the sense of Definition 2.1. Let us define, for all $K \in \mathcal{M}$, for all $L \in \mathcal{N}_K$,

$$A_{K,L} = \tau_{K|L}(x_{K|L} - x_K), \tag{2.11}$$

and for all $\sigma \in \mathcal{E}_{K,\mathrm{ext}}$, we define

$$A_{K,\sigma} = \tau_\sigma(x_\sigma - x_K). \tag{2.12}$$

(Recall that x_σ is defined by (2.3) and that $x_{K|L} = x_\sigma$ if $\sigma = K|L$.)

We define the discrete gradient $\nabla_{\mathcal{D}}\colon H_{\mathcal{D}} \to H_{\mathcal{D}}^d$, for any $u \in H_{\mathcal{D}}$, by:

$$\begin{aligned}\nabla_{\mathcal{D}}u(x) &= (\nabla_{\mathcal{D}}u)_K\\ &= \frac{1}{\mathrm{m}(K)}\left(\sum_{L\in\mathcal{N}_K} A_{K,L}\,(u_L - u_K) - \sum_{\sigma\in\mathcal{E}_{K,\mathrm{ext}}} A_{K,\sigma}\,u_K\right),\\ &\quad\text{for a.e. } x \in K \quad \forall K \in \mathcal{M}.\end{aligned}$$

Let us first state a bound on the $L^2(\Omega)^d$-norm of the discrete gradient of any element of $H_{\mathcal{D}}$.

LEMMA 2.2 (Bound on $\nabla_{\mathcal{D}}u$) Let Ω be an bounded open connected polygonal subset of $\mathbb{R}^d$, $d \in \mathbb{N}^\star$, let $\mathcal{D}$ be an admissible finite-volume discretization of Ω in the sense of Definition 2.1 and let $\theta \in (0, \theta_{\mathcal{D}}]$. Then, there exists C_5, only depending on d and θ, such that, for all $u \in H_{\mathcal{D}}$:

$$\|\nabla_{\mathcal{D}}u\|_{L^2(\Omega)^d} \leqslant C_5\|u\|_{\mathcal{D}}. \tag{2.13}$$

Proof. Let $u \in H_{\mathcal{D}}$. Let us denote, for all $K \in \mathcal{M}$, $L \in \mathcal{N}_K$ and $\sigma = K|L$, $\delta_{K,\sigma}u = u_L - u_K$, and for $\sigma \in \mathcal{E}_{K,\mathrm{ext}}$, $\delta_{K,\sigma}u = -u_K$. Then Definition (2.6) leads to

$$\|u\|_{\mathcal{D}}^2 = \sum_{K\in\mathcal{M}}\left(\frac{1}{2}\sum_{L\in\mathcal{N}_K}\tau_{K|L}(\delta_{K,K|L}u)^2 + \sum_{\sigma\in\mathcal{E}_{K,\mathrm{ext}}}\tau_\sigma(\delta_{K,\sigma}u)^2\right),$$

and Definition (2.3) leads, for a given $K \in \mathcal{M}$, to

$$\mathrm{m}(K)(\nabla_{\mathcal{D}}u)_K = \sum_{\sigma\in\mathcal{E}_K}\tau_\sigma(x_\sigma - x_K)\delta_{K,\sigma}u.$$

Using the Cauchy–Scharwz inequality, we obtain

$$\mathrm{m}(K)^2|(\nabla_{\mathcal{D}}u)_K|^2 \leqslant \sum_{\sigma\in\mathcal{E}_K}\tau_\sigma|x_\sigma - x_K|^2\sum_{\sigma\in\mathcal{E}_K}\tau_\sigma(\delta_{K,\sigma}u)^2,$$

and since, for $\sigma \in \mathcal{E}_K$, one has $|x_\sigma - x_K| = d(x_\sigma, x_K) \leqslant \frac{d_{K,\sigma}}{\theta}$,

$$\mathrm{m}(K)^2|(\nabla_{\mathcal{D}}u)_K|^2 \leqslant \sum_{\sigma\in\mathcal{E}_K} \frac{1}{\theta^2}\mathrm{m}(\sigma)d_{K,\sigma} \sum_{\sigma\in\mathcal{E}_K} \tau_\sigma(\delta_{K,\sigma}u)^2. \tag{2.14}$$

Since $\sum_{\sigma\in\mathcal{E}_K} \mathrm{m}(\sigma)d_{K,\sigma} = d\mathrm{m}(K)$, (2.14) gives

$$\mathrm{m}(K)|(\nabla_{\mathcal{D}}u)_K|^2 \leqslant \frac{d}{\theta^2} \sum_{\sigma\in\mathcal{E}_K} \tau_\sigma(\delta_{K,\sigma}u)^2.$$

Summing over $K \in \mathcal{M}$, we get

$$\|\nabla_{\mathcal{D}}u\|^2_{L^2(\Omega)^d} \leqslant 2\frac{d}{\theta^2}\|u\|^2_{\mathcal{D}},$$

which gives (2.13) with $C_5 = (\frac{2d}{\theta^2})^{\frac{1}{2}}$. □

We now state a weak convergence property for the discrete gradient.

LEMMA 2.3 (Weak convergence of the discrete gradient) Let Ω be an bounded open connected polygonal subset of $\mathbb{R}^d$, $d \in \mathbb{N}^\star$, let $\mathcal{D}$ be an admissible finite-volume discretization of Ω in the sense of Definition 2.1. We assume that there exist $u_{\mathcal{D}} \in H_{\mathcal{D}}$ and a function $\bar{u} \in H^1_0(\Omega)$ such that $u_{\mathcal{D}}$ tends to $\bar{u}$ in $L^2(\Omega)$ as $h_{\mathcal{D}}$ tends to 0 while $\|u_{\mathcal{D}}\|_{\mathcal{D}}$ remains bounded. Then $\nabla_{\mathcal{D}}u_{\mathcal{D}}$ weakly tends to $\nabla\bar{u}$ in $L^2(\Omega)^d$ as $h_{\mathcal{D}} \to 0$.

Proof. Let $\varphi \in C_c^\infty(\Omega)$. We assume that $h_{\mathcal{D}}$ is small enough to ensure that for all $K \in \mathcal{M}$ and $x \in K$, if $\varphi(x) \neq 0$, then $\mathcal{E}_{K,\mathrm{ext}} = \emptyset$. The expression $T_3^{\mathcal{D}}$, defined by

$$T_3^{\mathcal{D}} = \int_\Omega P_{\mathcal{D}}\varphi(x)\nabla_{\mathcal{D}}u_{\mathcal{D}}(x)\,\mathrm{d}x,$$

satisfies, using (2.11),

$$T_3^{\mathcal{D}} = \sum_{K|L\in\mathcal{E}_{\mathrm{int}}} \tau_{K|L}(u_L - u_K)((x_{K|L} - x_K)\varphi(x_K) + (x_L - x_{K|L})\varphi(x_L)),$$

where we denote, for the sake of simplicity, $u_K = (u_{\mathcal{D}})_K$ for all $K \in \mathcal{M}$. We thus get $T_3^{\mathcal{D}} = T_4^{\mathcal{D}} + T_5^{\mathcal{D}}$ with

$$T_4^{\mathcal{D}} = \sum_{K|L\in\mathcal{E}_{\mathrm{int}}} \tau_{K|L}(u_L - u_K)(x_L - x_K)\frac{\varphi(x_K) + \varphi(x_L)}{2}$$

and

$$T_5^{\mathcal{D}} = \sum_{K|L\in\mathcal{E}_{\mathrm{int}}} \tau_{K|L}(u_L - u_K)\left(x_{K|L} - \frac{x_L + x_K}{2}\right)(\varphi(x_L) - \varphi(x_K)).$$

Thanks to the Cauchy–Schwarz inequality, we get

$$(T_5^{\mathcal{D}})^2 \leqslant \sum_{K|L\in\mathcal{E}_{\mathrm{int}}} \tau_{K|L}(u_L - u_K)^2 \sum_{K|L\in\mathcal{E}_{\mathrm{int}}} \tau_{K|L}(\varphi(x_L) - \varphi(x_K))^2 \left|x_{K|L} - \frac{x_L + x_K}{2}\right|^2.$$

Since $|x_{K|L} - \frac{x_L + x_K}{2}| \leqslant \frac{1}{2}|x_{K|L} - x_L| + \frac{1}{2}|x_{K|L} - x_K| \leqslant h_{\mathcal{D}}$, there exists $C_6 > 0$, depending on d, Ω and φ such that

$$(T_5^{\mathcal{D}})^2 \leqslant \|u_{\mathcal{D}}\|_{\mathcal{D}}^2 C_6 h_{\mathcal{D}}^2 \mathrm{m}(\Omega),$$

and therefore we get

$$\lim_{h_{\mathcal{D}} \to 0} T_5^{\mathcal{D}} = 0.$$

We then compare $T_4^{\mathcal{D}}$ with

$$T_6^{\mathcal{D}} = -\int_{\Omega} u_{\mathcal{D}}(x) \nabla \varphi(x)\, \mathrm{d}x = \sum_{K|L \in \mathcal{E}_{\mathrm{int}}} (u_L - u_K) \int_{K|L} \varphi(x) \mathbf{n}_{K,L}\, \mathrm{d}\gamma(x).$$

Since

$$\mathbf{n}_{K,L} = \frac{x_L - x_K}{d_{K|L}}$$

and since

$$\left| \frac{1}{\mathrm{m}(K|L)} \int_{K|L} \varphi(x)\, \mathrm{d}\gamma(x) - \frac{\varphi(x_K) + \varphi(x_L)}{2} \right| \leqslant \|\nabla \varphi\|_{L^\infty(\Omega)}\, h_{\mathcal{D}},$$

we get, thanks to the Cauchy–Schwarz inequality,

$$\lim_{h_{\mathcal{D}} \to 0} (T_4^{\mathcal{D}} - T_6^{\mathcal{D}})^2 = 0.$$

Since

$$\lim_{h_{\mathcal{D}} \to 0} T_6^{\mathcal{D}} = -\int_{\Omega} \bar{u}(x) \nabla \varphi(x)\, \mathrm{d}x = \int_{\Omega} \varphi(x) \nabla \bar{u}(x)\, \mathrm{d}x,$$

we have thus proven, thanks to the density of $\mathrm{C}_c^\infty(\Omega)$ in $L^2(\Omega)$, the weak convergence of $\nabla_{\mathcal{D}} u_{\mathcal{D}}$ to $\nabla \bar{u}(x)$ as $h_{\mathcal{D}} \to 0$. This completes the proof of the lemma. □

We now study, for a regular function φ, the strong convergence of the discrete gradient $\nabla_{\mathcal{D}} P_{\mathcal{D}} \varphi$ to $\nabla \varphi$. This study uses the following lemma.

LEMMA 2.4 Let Ω be an bounded open connected polygonal subset of $\mathbb{R}^d$, $d \in \mathbb{N}^\star$, let $\mathcal{D}$ be an admissible finite-volume discretization of Ω in the sense of Definition 2.1. Then we have

$$v = \frac{1}{\mathrm{m}(K)} \sum_{\sigma \in \mathcal{E}_K} \mathrm{m}(\sigma)(x_\sigma - x_0)\, (\mathbf{n}_{K,\sigma} \cdot v) \quad \forall K \in \mathcal{M}, \;\; \forall x_0 \in \mathbb{R}^d, \; \forall v \in \mathbb{R}^d. \tag{2.15}$$

Proof. For any $K \in \mathcal{M}$, we denote, for a.e. $x \in \partial K$, by $\mathbf{n}_{\partial K}(x)$ the normal vector to ∂K at the point x outward K. Let v and $w \in \mathbb{R}^d$ be given. We have, considering vectors as $d \times 1$ matrices and denoting by w^t the transposed $1 \times d$ matrix of w,

$$\begin{aligned} w^t \left(\int_{\partial K} (x - x_0) \mathbf{n}_K^t(x)\, \mathrm{d}\gamma(x) \right) v &= \int_{\partial K} w^t (x - x_0) \mathbf{n}_K^t(x) v\, \mathrm{d}\gamma(x) \\ &= \int_{\partial K} w^t (x - x_0) v^t \mathbf{n}_K(x)\, \mathrm{d}\gamma(x) \\ &= \int_{\partial K} (v(x - x_0)^t w) \cdot \mathbf{n}_K(x)\, \mathrm{d}\gamma(x) \\ &= \int_K \mathrm{div}(v(x - x_0)^t w)\, \mathrm{d}x = \mathrm{m}(K)\, v^t w. \end{aligned}$$

This gives (2.15). □

LEMMA 2.5 (Consistency property of the discrete gradient) Let Ω be an bounded open connected polygonal subset of $\mathbb{R}^d$, $d \in \mathbb{N}^\star$, let $\mathcal{D}$ be an admissible finite-volume discretization in the sense of Definition 2.1 and let $\theta \in (0, \theta_{\mathcal{D}}]$. Let $\bar{u} \in \mathrm{C}^2(\overline{\Omega})$ be such that $\bar{u} = 0$ on the boundary of Ω. Then, there exists C_7, only depending on Ω, θ and $\bar{u}$, such that:

$$\|\nabla_{\mathcal{D}} P_{\mathcal{D}} \bar{u} - \nabla \bar{u}\|_{L^2(\Omega)^d} \leqslant C_7 h_{\mathcal{D}}. \tag{2.16}$$

(Recall that $P_{\mathcal{D}}$ is defined by (2.5) and $\nabla_{\mathcal{D}}$ in Definition 2.3.)

Proof. From Definition 2.3 and (2.5), we can write for any $K \in \mathcal{M}$

$$\mathrm{m}(K)(\nabla_{\mathcal{D}} P_{\mathcal{D}} \bar{u})_K = \sum_{L \in \mathcal{N}_K} \tau_{K|L}(x_{K|L} - x_K)(\bar{u}(x_L) - \bar{u}(x_K)) - \sum_{\sigma \in \mathcal{E}_{K,\mathrm{ext}}} \tau_\sigma (x_\sigma - x_K)\bar{u}(x_K). \tag{2.17}$$

Let $(\nabla \bar{u})_K$ be the mean value of $\nabla \bar{u}$ on K:

$$(\nabla \bar{u})_K = \frac{1}{\mathrm{m}(K)} \int_K \nabla \bar{u}(x)\,\mathrm{d}x.$$

Thanks to the regularity of $\bar{u}$ (and the fact that $\bar{u} = 0$ on the boundary of Ω), there exists C_8, only depending on $\bar{u}$ (indeed, C_8 only depends on the L^∞-norm of the second derivatives of $\bar{u}$), such that, for all $\sigma = K|L \in \mathcal{E}_{\mathrm{int}}$,

$$|e_\sigma| \leqslant C_8 h_{\mathcal{D}}, \quad \text{with } e_\sigma = (\nabla \bar{u})_K \cdot \mathbf{n}_{K,\sigma} - \frac{\bar{u}(x_L) - \bar{u}(x_K)}{d_\sigma}, \tag{2.18}$$

and, for all $\sigma \in \mathcal{E}_{K,\mathrm{ext}}$,

$$|e_\sigma| \leqslant C_8 h_{\mathcal{D}}, \quad \text{with } e_\sigma = (\nabla \bar{u})_K \cdot \mathbf{n}_{K,\sigma} - \frac{-\bar{u}(x_K)}{d_{K,\sigma}}. \tag{2.19}$$

Thanks to (2.17), (2.18) and (2.19), we get, for all $K \in \mathcal{M}$:

$$\mathrm{m}(K)(\nabla_{\mathcal{D}} P_{\mathcal{D}} \bar{u})_K = \sum_{\sigma \in \mathcal{E}_K} \mathrm{m}(\sigma)(x_\sigma - x_K)(\nabla \bar{u})_K \cdot \mathbf{n}_{K,\sigma} + R_K,$$

with $R_K = -\sum_{\sigma \in \mathcal{E}_K} e_\sigma \mathrm{m}(\sigma) d(x_\sigma, x_K)$. Applying (2.15) gives

$$\mathrm{m}(K)(\nabla_{\mathcal{D}} P_{\mathcal{D}} \bar{u})_K = \mathrm{m}(K)(\nabla \bar{u})_K + R_K. \tag{2.20}$$

Using the inequalities (2.18) and (2.19), we have

$$|R_K| \leqslant \frac{C_8}{\theta} h_{\mathcal{D}} \sum_{\sigma \in \mathcal{E}_K} \mathrm{m}(\sigma) d_{K,\sigma} = \frac{d\, C_8}{\theta} h_{\mathcal{D}} \mathrm{m}(K). \tag{2.21}$$

Then, from (2.20) and (2.21), we obtain

$$\begin{aligned} \sum_{K \in \mathcal{M}} |(\nabla_{\mathcal{D}} P_{\mathcal{D}} \bar{u})_K - (\nabla \bar{u})_K|^2 \mathrm{m}(K) &\leqslant \sum_{K \in \mathcal{M}} \left(\frac{d\, C_8}{\theta}\right)^2 h_{\mathcal{D}}^2 \mathrm{m}(K) \\ &= \mathrm{m}(\Omega) \left(\frac{d\, C_8}{\theta}\right)^2 h_{\mathcal{D}}^2. \end{aligned} \tag{2.22}$$

In order to conclude, we remark that, thanks to the regularity of $\bar{u}$, there exists C_9, only depending on $\bar{u}$ (here also, C_9 only depends on the L^∞-norm of the second derivatives of $\bar{u}$), such that:

$$\sum_{K\in\mathcal{M}} \int_K |\nabla \bar{u}(x) - (\nabla \bar{u})_K|^2 \, \mathrm{d}x \leqslant C_9 h_{\mathcal{D}}^2. \tag{2.23}$$

Then, using (2.22) and (2.23), we get the existence of C_7, only depending on Ω, θ and $\bar{u}$, such that (2.16) holds. □

REMARK 2.1 (Choice of the points x_K and x_σ) Note that in the proof of Lemma 2.3, one is free to choose any point lying on $K|L$ instead of $x_{K|L}$ in the definition of the coefficients $A_{K,L}$. However, we need this choice in the proof of the strong consistency of the discrete gradient (Lemma 2.5). Conversely, in the proof of Lemma 2.5, we could take any point of K instead of x_K in the definition of $A_{K,L}$. However, the choice of x_K is crucial in the proof of Lemma 2.3: when comparing the terms T_5 and T_6, one needs the property of consistency of the normal flux, which follows from the fact that $\mathbf{n}_{K,L} = \frac{x_L - x_K}{d_{K|L}}$.

LEMMA 2.6 (A sufficient condition for the strong convergence of the discrete gradient) Let Ω be an bounded open connected polygonal subset of $\mathbb{R}^d$, $d \in \mathbb{N}^\star$, let $\theta > 0$ and let $\mathcal{D}$ be an admissible finite-volume discretization in the sense of Definition 2.1, such that $\theta_{\mathcal{D}} \geqslant \theta$. Assume that there exists a function $u_{\mathcal{D}} \in H_{\mathcal{D}}$ and a function $\bar{u} \in H_0^1(\Omega)$ such that $u_{\mathcal{D}}$ tends to $\bar{u}$ in $L^2(\Omega)$ as $h_{\mathcal{D}}$ tends to 0. Assume also that there exists a function $\alpha \in L^\infty(\Omega)$ and $\alpha_0 > 0$ such that $\alpha(x) \geqslant \alpha_0$ for a.e. $x \in \Omega$ and $[u_{\mathcal{D}}, u_{\mathcal{D}}]_{\mathcal{D},\alpha}$ tends to $\int_\Omega \alpha(x)\nabla\bar{u}(x)^2 \, \mathrm{d}x$ as $h_{\mathcal{D}}$ tends to 0. Then $\nabla_{\mathcal{D}} u_{\mathcal{D}}$ tends to $\nabla\bar{u}$ in $L^2(\Omega)^d$ as $h_{\mathcal{D}}$ tends to 0.

Proof. Let $\varphi \in C_c^\infty(\Omega)$ be given (this function is meant to approximate $\bar{u}$ in $H_0^1(\Omega)$). Thanks to the Cauchy–Schwarz inequality, we have

$$\int_\Omega (\nabla_{\mathcal{D}} u_{\mathcal{D}}(x) - \nabla\bar{u}(x))^2 \, \mathrm{d}x \leqslant 3\,(T_7^{\mathcal{D}} + T_8^{\mathcal{D}} + T_9)$$

with

$$T_7^{\mathcal{D}} = \int_\Omega (\nabla_{\mathcal{D}} u_{\mathcal{D}}(x) - \nabla_{\mathcal{D}} P_{\mathcal{D}}\varphi(x))^2 \, \mathrm{d}x,$$
$$T_8^{\mathcal{D}} = \int_\Omega (\nabla_{\mathcal{D}} P_{\mathcal{D}}\varphi(x) - \nabla\varphi(x))^2 \, \mathrm{d}x$$

and

$$T_9 = \int_\Omega (\nabla\varphi(x) - \nabla\bar{u}(x))^2 \, \mathrm{d}x.$$

We have, thanks to Lemma 2.5,

$$\lim_{h_{\mathcal{D}}\to 0} T_8^{\mathcal{D}} = 0. \tag{2.24}$$

Thanks to Lemma 2.2, we have

$$\int_\Omega (\nabla_{\mathcal{D}} v(x))^2 \, \mathrm{d}x \leqslant C_5{}^2 [v, v]_{\mathcal{D},1} \leqslant \frac{C_5{}^2}{\alpha_0} [v, v]_{\mathcal{D},\alpha}, \quad \forall v \in H_{\mathcal{D}}.$$

We thus get, setting $v = u_{\mathcal{D}} - P_{\mathcal{D}}\varphi$ in the above inequality, that

$$T_7^{\mathcal{D}} \leqslant \frac{C_5{}^2}{\alpha_0} ([u_{\mathcal{D}}, u_{\mathcal{D}}]_{\mathcal{D},\alpha} - 2[u_{\mathcal{D}}, P_{\mathcal{D}}\varphi]_{\mathcal{D},\alpha} + [P_{\mathcal{D}}\varphi, P_{\mathcal{D}}\varphi]_{\mathcal{D},\alpha}).$$

We have, applying twice Lemma 2.1, that

$$\lim_{h_{\mathcal{D}}\to 0}[u_{\mathcal{D}}, P_{\mathcal{D}}\varphi]_{\mathcal{D},\alpha} = \int_{\Omega} \alpha(x)\nabla\bar{u}(x)\cdot\nabla\varphi(x)\,\mathrm{d}x \tag{2.25}$$

and

$$\lim_{h_{\mathcal{D}}\to 0}[P_{\mathcal{D}}\varphi, P_{\mathcal{D}}\varphi]_{\mathcal{D},\alpha} = \int_{\Omega} \alpha(x)\nabla\varphi(x)^2\,\mathrm{d}x. \tag{2.26}$$

Under the hypotheses of the lemma, we then get that

$$\limsup_{h_{\mathcal{D}}\to 0} T_7^{\mathcal{D}} \leqslant \frac{C_5^{\,2}}{\alpha_0}\int_{\Omega} \alpha(x)(\nabla\bar{u}(x)-\nabla\varphi(x))^2\,\mathrm{d}x.$$

We then get, gathering the above results, setting $C_{10} = \frac{C_5^{\,2}}{\alpha_0}\operatorname{ess\,sup}_{x\in\Omega}\alpha(x)+1$, that

$$\int_{\Omega}(\nabla_{\mathcal{D}}u_{\mathcal{D}}(x)-\nabla\bar{u}(x))^2\,\mathrm{d}x \leqslant C_{10}\int_{\Omega}(\nabla\varphi(x)-\nabla\bar{u}(x))^2\,\mathrm{d}x + T_{10}^{\mathcal{D}},$$

with

$$\lim_{h_{\mathcal{D}}\to 0} T_{10}^{\mathcal{D}} = 0. \tag{2.27}$$

Let $\varepsilon > 0$. We can choose φ such that $\int_{\Omega}(\nabla\varphi(x)-\nabla\bar{u}(x))^2\,\mathrm{d}x \leqslant \varepsilon$, and we can then choose $h_{\mathcal{D}}$ such that $T_{10}^{\mathcal{D}} \leqslant \varepsilon$. Hence we deduce that

$$\lim_{h_{\mathcal{D}}\to 0}\int_{\Omega}(\nabla_{\mathcal{D}}u_{\mathcal{D}}(x)-\nabla\bar{u}(x))^2\,\mathrm{d}x = 0. \tag{2.28}$$

□

REMARK 2.2 Thanks to Lemma 2.6, we get the strong convergence of the discrete gradient in the case of the classical finite-volume scheme for an isotropic problem. Note that in the above proof, we did not use the weak convergence of the discrete gradient, and therefore, any point of K can be taken instead of x_K in the definition of the coefficients $A_{K,L}$. We thus find that the average value in K of the gradient defined in Eymard *et al.* (2001) is also strongly convergent (the average of this gradient, defined by the generalized Raviart–Thomas basis functions, is obtained by replacing x_K by the barycentre of K in the definition of $A_{K,L}$). Note that the drawback of the generalization of the Raviart–Thomas basis was the difficulty to compute approximate values of the gradients. This drawback no longer exists for an averaged gradient. Nevertheless, the properties of convergence of the finite-volume method shown here for non-isotropic problems are only proven for the choice (2.11) in the definition of $A_{K,L}$, and not for the Raviart–Thomas basis.

3. Discretization of the anisotropic diffusion problem (1.1)

3.1 *The finite-volume scheme*

Under Hypotheses (1.2)–(1.4), let $\mathcal{D}$ be an admissible discretization of Ω in the sense of Definition 2.1. The finite-volume approximation to Problem (1.1) is given, for a suitable choice of the function $\alpha \in L^{\infty}(\Omega)$, as the solution of the following equation:

$$\begin{cases} u_{\mathcal{D}} \in H_{\mathcal{D}}, \\ \int_{\Omega}(\Lambda(x)-\alpha(x)I_d)\nabla_{\mathcal{D}}u_{\mathcal{D}}(x)\cdot\nabla_{\mathcal{D}}v(x)\,\mathrm{d}x + [u_{\mathcal{D}}, v]_{\mathcal{D},\alpha} = \int_{\Omega} f(x)v(x)\,\mathrm{d}x, \ \forall v \in H_{\mathcal{D}}, \end{cases} \tag{3.1}$$

denoting by I_d the identity mapping in $\mathbb{R}^d$. In the sequel, we shall refer to Scheme (3.1) as the 'gradient scheme', since it is obtained from a discretization formula for the gradient. Note that in this formulation, we use the discrete gradient on part of the the operator only, while on the homogeneous part, we write the usual cell-centred scheme. This needs to be done in order to obtain the stability of the scheme through a certain *a priori* estimate on the discrete solution. If we take $\alpha = 0$ in (3.1), we are no longer able to prove the discrete H^1 estimate (3.7) below. Taking for v the characteristic function of a control volume K in (3.1), we may note that (3.1) is equivalent to finding the values $(u_K)_{K\in\mathcal{M}}$ (we again denote u_K instead of $(u_{\mathcal{D}})_K$) that comprise the solution of the following system of equations:

$$\sum_{L\in\mathcal{N}_K} F_{KL} + \sum_{\sigma\in\mathcal{E}_{K,\mathrm{ext}}} F_{K\sigma} = \int_K f(x)\,\mathrm{d}x \quad \forall K \in \mathcal{M}, \tag{3.2}$$

where

$$F_{KL} = \tau_{K|L}\alpha_{K|L}(u_K - u_L) + (\Lambda_L A_{LK}\cdot\nabla_{\mathcal{D}}u_L - \Lambda_K A_{KL}\cdot\nabla_{\mathcal{D}}u_K) \quad \forall K|L \in \mathcal{E}_{\mathrm{int}}, \tag{3.3}$$

and

$$F_{K\sigma} = \tau_{K\sigma}\alpha_\sigma u_K + \Lambda_K A_{K\sigma}\cdot\nabla_{\mathcal{D}}u_K \quad \forall\sigma\in\mathcal{E}_{K,\mathrm{ext}}. \tag{3.4}$$

In (3.3) and (3.4), the matrices $(\Lambda_K)_{K\in\mathcal{M}}$ are defined by:

$$\Lambda_K = \frac{1}{\mathrm{m}(K)}\int_K (\Lambda(x) - \alpha(x)I_d)\,\mathrm{d}x. \tag{3.5}$$

One can then complete the discrete expressions of F_{KL} and $F_{K\sigma}$ using Definition 2.3 for A_{KL}, $A_{K\sigma}$ and $\nabla_{\mathcal{D}}u_K$ for all $K\in\mathcal{M}$, $L\in\mathcal{N}_K$ and $\sigma\in\mathcal{E}_K$.

This is indeed a finite-volume scheme, since

$$F_{KL} = -F_{LK} \quad \forall K|L\in\mathcal{E}_{\mathrm{int}}.$$

The mathematical study of this scheme (existence of a solution to (3.1) and proof of the convergence of this solution to that of Problem (1.1)) is carried out in this paper under the following assumption for α, denoted by Hypothesis (3.6): the function $\alpha\in L^\infty(\Omega)$ is such that there exists $\alpha_0 > 0$ with

$$0 < \alpha_0 \leqslant \alpha(x) \leqslant \underline{\lambda}(x) \quad \text{for a.e. } x\in\Omega, \tag{3.6}$$

where $\underline{\lambda}$ is defined in (1.3).

REMARK 3.1 Note that, in some cases of actual engineering studies involving multiphase flow in soils, because of computing costs, only coarse meshes can be used and accurate results can only be obtained if fluxes are computed by harmonic-averaging of absolute permeabilities (this is done in most of the commercial codes). Hence, in practical implementations of the gradient scheme, one can define the function α by $\alpha(x) = |\Lambda(x)\mathbf{n}_{K,\sigma}|$ for all $K\in\mathcal{M}$, $\sigma\in\mathcal{E}_K$ and a.e. $x\in D_{K,\sigma}$, and then replace (2.7) by

$$\frac{1}{\alpha_\sigma} = \frac{1}{\mathrm{m}(D_\sigma)}\int_{D_\sigma}\frac{1}{\alpha(x)}\,\mathrm{d}x.$$

In the case where Λ is heterogeneous and isotropic, the above formula reduces to the classical harmonic average method. However, the mathematical analysis of the gradient scheme in such a case remains an open problem since, on one hand, this function α depends on the mesh, and on the other hand Hypothesis (3.6) no longer holds.

3.2 *Discrete $H^1(\Omega)$ estimate*

We now prove the following estimate.

LEMMA 3.1 (Discrete H^1 estimate) Under Hypotheses (1.2)–(1.4) and (3.6), let $\mathcal{D}$ be an admissible discretization of Ω in the sense of Definition 2.1. Let $u \in H_{\mathcal{D}}$ be a solution to (3.1). Then the following inequality holds:

$$\alpha_0 \|u\|_{\mathcal{D}} \leqslant \operatorname{diam}(\Omega) \|f\|_{(L^2(\Omega))^2}. \tag{3.7}$$

Proof. We apply (3.1) setting $v = u$. We get

$$\int_\Omega (\Lambda(x) - \alpha(x) I_d) \nabla_{\mathcal{D}} u(x) \cdot \nabla_{\mathcal{D}} u(x) \,\mathrm{d}x + [u, u]_{\mathcal{D},\alpha} = \int_\Omega f(x) u(x) \,\mathrm{d}x.$$

Thanks to (3.6), the above relation implies

$$\alpha_0 [u, u]_{\mathcal{D}} \leqslant \int_\Omega f(x) u(x) \,\mathrm{d}x.$$

Then the conclusion follows from the discrete Poincaré inequality (2.8). □

We can now state the existence and the uniqueness of a discrete solution to (3.1).

COROLLARY 3.1 (Existence and uniqueness of a solution to the finite-volume scheme). Under Hypotheses (1.2)–(1.4) and (3.6), let $\mathcal{D}$ be an admissible discretization of Ω in the sense of Definition 2.1. Then there exists a unique solution $u_{\mathcal{D}}$ to (3.1).

Proof. Note that (3.1) is a finite-dimensional linear problem. Assume that $f = 0$. From the discrete Poincaré inequality (2.8), we get that $u = 0$. This proves that the linear problem (3.1) is uniquely solvable. □

3.3 *Convergence*

We have the following result, which states the convergence of the gradient scheme (3.1).

THEOREM 3.1 (Convergence of the finite-volume scheme) Under Hypotheses (1.2)–(1.4) and (3.6), let $\theta > 0$. Let $\mathcal{D}$ be an admissible discretization of Ω in the sense of Definition 2.1, such that $\theta_{\mathcal{D}} \geqslant \theta$. Let $u_{\mathcal{D}} \in H_{\mathcal{D}}(\Omega)$ be the solution to (3.1). Then

- $u_{\mathcal{D}}$ converges in $L^2(\Omega)$ to $\bar{u}$, the weak solution of Problem (1.1) in the sense of Definition 1.1,
- the discrete gradient $\nabla_{\mathcal{D}} u_{\mathcal{D}}$ converges in $L^2(\Omega)^d$ to $\nabla \bar{u}$,

as $h_{\mathcal{D}}$ tends to 0.

Proof. We consider a sequence of admissible discretizations $(\mathcal{D}_n)_{n\in\mathbb{N}}$ such that $h_{\mathcal{D}_n}$ tends to 0 as $n \to \infty$ and $\theta_{\mathcal{D}_n} \geqslant \theta$ for all $n \in \mathbb{N}$. Thanks to Lemma 3.1, we can apply the compactness result (2.1), which gives the existence of a subsequence (again denoted $(\mathcal{D}_n)_{n\in\mathbb{N}}$) and of $\bar{u} \in H_0^1(\Omega)$ such that $u_{\mathcal{D}_n}$ (given by (3.1) with $\mathcal{D} = \mathcal{D}_n$) tends to $\bar{u}$ in $L^2(\Omega)$ as $n \to \infty$. Let $\varphi \in C_c^\infty(\Omega)$ be given; we choose $v = P_{\mathcal{D}_n}\varphi$ as test function in (3.1). We obtain

$$\int_\Omega (\Lambda(x) - \alpha(x) I_d) \nabla_{\mathcal{D}_n} u_{\mathcal{D}_n}(x) \cdot \nabla_{\mathcal{D}_n} P_{\mathcal{D}_n}\varphi(x) \,\mathrm{d}x + \left[u_{\mathcal{D}_n}, P_{\mathcal{D}_n}\varphi\right]_{\mathcal{D}_n,\alpha} = \int_\Omega f(x) P_{\mathcal{D}_n}\varphi(x) \,\mathrm{d}x. \tag{3.8}$$

We let $n \to \infty$ in (3.8). Thanks to Lemma 2.3 and Lemma 2.5 (which provide a weak/strong convergence result), we get that

$$\lim_{n\to\infty} \int_\Omega (\Lambda(x) - \alpha(x) I_d) \nabla_{\mathcal{D}_n} u_{\mathcal{D}_n}(x) \cdot \nabla_{\mathcal{D}_n} P_{\mathcal{D}_n} \varphi(x)\, \mathrm{d}x = \int_\Omega (\Lambda(x) - \alpha(x) I_d) \nabla \bar{u}(x) \cdot \nabla \varphi(x)\, \mathrm{d}x.$$

Using Lemma 2.1, we deduce that

$$\lim_{n\to\infty} \left[u_{\mathcal{D}_n}, P_{\mathcal{D}_n}\varphi\right]_{\mathcal{D}_n,\alpha} = \int_\Omega \alpha(x) \nabla \bar{u}(x) \cdot \nabla \varphi(x)\, \mathrm{d}x.$$

Since it is easy to see that

$$\lim_{n\to\infty} \int_\Omega f(x) P_{\mathcal{D}_n}\varphi(x)\, \mathrm{d}x = \int_\Omega f(x)\varphi(x)\, \mathrm{d}x,$$

we thus get that any limit $\bar{u}$ of a subsequence of solutions satisfies (1.5) with $v = \varphi$. A classical density argument and the uniqueness of the solution to (1.5) permit to conclude the convergence in $L^2(\Omega)$ of $u_{\mathcal{D}}$ to $\bar{u}$, the weak solution of the problem in the sense of Definition 1.1, as $h_{\mathcal{D}}$ tends to 0, thanks to the fact that $\theta_{\mathcal{D}} \geqslant \theta$. Let us now prove the strong convergence of $\nabla_{\mathcal{D}} u_{\mathcal{D}}$ to $\nabla \bar{u}$. We have, using (3.1) with $v = u_{\mathcal{D}}$,

$$\int_\Omega (\Lambda(x) - \alpha(x) I_d) \nabla_{\mathcal{D}} u_{\mathcal{D}}(x) \cdot \nabla_{\mathcal{D}} u_{\mathcal{D}}(x)\, \mathrm{d}x = \int_\Omega f(x) u_{\mathcal{D}}(x)\, \mathrm{d}x - [u_{\mathcal{D}}, u_{\mathcal{D}}]_{\mathcal{D},\alpha}. \tag{3.9}$$

Thanks to Lemma 2.1, we have

$$\int_\Omega \alpha(x) \nabla \bar{u}(x)^2\, \mathrm{d}x \leqslant \liminf_{h_{\mathcal{D}}\to 0} [u_{\mathcal{D}}, u_{\mathcal{D}}]_{\mathcal{D},\alpha},$$

and therefore, passing to the limit in (3.9), we get that

$$\limsup_{h_{\mathcal{D}}\to 0} \int_\Omega (\Lambda(x) - \alpha(x) I_d) \nabla_{\mathcal{D}} u_{\mathcal{D}}(x) \cdot \nabla_{\mathcal{D}} u_{\mathcal{D}}(x)\, \mathrm{d}x \leqslant \int_\Omega f(x) u_{\mathcal{D}}(x)\, \mathrm{d}x - \int_\Omega \alpha(x) \nabla \bar{u}(x)^2\, \mathrm{d}x.$$

We then have, letting $v = \bar{u}$ in (1.5),

$$\int_\Omega (\Lambda(x) - \alpha(x) I_d) \nabla \bar{u}(x) \cdot \nabla \bar{u}(x)\, \mathrm{d}x = \int_\Omega f(x) \bar{u}(x)\, \mathrm{d}x - \int_\Omega \alpha(x) \nabla \bar{u}(x)^2\, \mathrm{d}x. \tag{3.10}$$

This leads to

$$\limsup_{h_{\mathcal{D}}\to 0} \int_\Omega (\Lambda(x) - \alpha(x) I_d) \nabla_{\mathcal{D}} u_{\mathcal{D}}(x) \cdot \nabla_{\mathcal{D}} u_{\mathcal{D}}(x)\, \mathrm{d}x \leqslant \int_\Omega (\Lambda(x) - \alpha(x) I_d) \nabla \bar{u}(x) \cdot \nabla \bar{u}(x)\, \mathrm{d}x.$$

Using Lemma 2.3, which states the weak convergence of the gradient $\nabla_{\mathcal{D}} u_{\mathcal{D}}$ to $\nabla \bar{u}$, we get that

$$\int_\Omega (\Lambda(x) - \alpha(x) I_d) \nabla \bar{u}(x) \cdot \nabla \bar{u}(x)\, \mathrm{d}x \leqslant \liminf_{h_{\mathcal{D}}\to 0} \int_\Omega (\Lambda(x) - \alpha(x) I_d) \nabla_{\mathcal{D}} u_{\mathcal{D}}(x) \cdot \nabla_{\mathcal{D}} u_{\mathcal{D}}(x)\, \mathrm{d}x.$$

The above inequalities yield

$$\lim_{h_{\mathcal{D}}\to 0} \int_\Omega (\Lambda(x) - \alpha(x) I_d) \nabla_{\mathcal{D}} u_{\mathcal{D}}(x) \cdot \nabla_{\mathcal{D}} u_{\mathcal{D}}(x)\, \mathrm{d}x = \int_\Omega (\Lambda(x) - \alpha(x) I_d) \nabla \bar{u}(x) \cdot \nabla \bar{u}(x)\, \mathrm{d}x. \tag{3.11}$$

From (3.9), (3.10) and (3.11), we thus obtain that

$$\lim_{h_{\mathcal{D}}\to 0}[u_{\mathcal{D}}, u_{\mathcal{D}}]_{\mathcal{D},\alpha} = \int_{\Omega} \alpha(x)\nabla\bar{u}(x)^2\,\mathrm{d}x.$$

Therefore, we can apply Lemma 2.6. This completes the proof of the strong convergence of the discrete gradient. □

3.4 *Error estimate*

We now give an error estimate, assuming first that the solution of (1.5) is in $\mathrm{C}^2(\overline{\Omega})$. In Theorem 3.3, we will consider the weaker hypothesis that the solution of (1.5) is only in $H^2(\Omega)$ under the assumption $d \leqslant 3$.

THEOREM 3.2 (C^2 error estimate) Assume Hypotheses (1.2)–(1.4) and (3.6) and that Λ and α are of class C^1 on $\overline{\Omega}$. Let $\mathcal{D}$ be an admissible finite-volume discretization (in the sense of Definition 2.1). Let $\theta \in (0, \theta_{\mathcal{D}}]$, where $\theta_{\mathcal{D}}$ is defined by (2.4). Let $u_{\mathcal{D}} \in H_{\mathcal{D}}$ be the solution of (3.1) and $\bar{u} \in H_0^1(\Omega)$ be the solution of (1.5). We assume that $\bar{u} \in \mathrm{C}^2(\overline{\Omega})$.

Let us first assume that

$$\int_{\sigma} \Lambda(x)\mathbf{n}_{\partial\Omega}(x)\cdot(x_\sigma - z_\sigma)\,\mathrm{d}\gamma(x) = 0 \quad \forall\sigma\in\mathcal{E}_{\mathrm{ext}}, \tag{3.12}$$

where $\mathbf{n}_{\partial\Omega}(x)$ is the unit normal vector to $\partial\Omega$ at point x, outward to Ω.

Then, there exists C_{11} only depending on Ω, θ, α_0, α, β, Λ and $\|\bar{u}\|_{\mathrm{C}^2(\Omega)}$, such that:

$$\|u_{\mathcal{D}} - P_{\mathcal{D}}\bar{u}\|_{\mathcal{D}} \leqslant C_{11}h_{\mathcal{D}}, \tag{3.13}$$

$$\|u_{\mathcal{D}} - \bar{u}\|_{L^2(\Omega)} \leqslant C_{11}h_{\mathcal{D}} \tag{3.14}$$

and

$$\|\nabla_{\mathcal{D}}u_{\mathcal{D}} - \nabla\bar{u}\|_{L^2(\Omega)^d} \leqslant C_{11}h_{\mathcal{D}}. \tag{3.15}$$

If (3.12) does not hold, then there exists C_{12}, only depending on Ω, θ, α, β, Λ and $\|\bar{u}\|_{H^2(\Omega)}$, such that (3.38), (3.39) and (3.40) hold with $C_{12}\sqrt{h_{\mathcal{D}}}$ instead of $C_{11}h_{\mathcal{D}}$.

REMARK 3.2 Let us give some sufficient (and practical) conditions for (3.12) to hold.

- If the normal vector to $\partial\Omega$ is an eigenvector of $\Lambda(x)$ for a.e. $x \in \partial\Omega$, then (3.12) holds. Since this property is always satisfied in the isotropic case, the error estimate on the gradient (3.15) holds for the classical cell-centred scheme, for any admissible mesh.
- If for all $\sigma \in \mathcal{E}_{\mathrm{ext}}$ with $\sigma \in \mathcal{E}_K$, the barycentre x_σ of σ is equal to the orthogonal projection z_σ of x_K on σ, then (3.12) holds. This hypothesis is easy to ensure on rectangular and triangular meshes.

Note also that one could replace (3.12) by $|z_\sigma - x_\sigma| \leqslant \frac{1}{\theta}\mathrm{diam}(K)(h_{\mathcal{D}})^{\frac{1}{2}}$ for all $\sigma \in \mathcal{E}_{\mathrm{ext}}$.

Proof. In the proof, we denote by C_i ($i \in \mathbb{N}$), various quantities only depending on Ω, θ, α_0, α, β, Λ and $\|\bar{u}\|_{\mathrm{C}^2(\Omega)}$.

Step 1 Let $v \in H_{\mathcal{D}}$. We first perform an estimation of a consistency error, namely, a bound for $|T_{11}(v)|$, where $T_{11}(v)$ is defined by:

$$\int_{\Omega}(\Lambda(x) - \alpha(x)I_d)\nabla_{\mathcal{D}}P_{\mathcal{D}}\bar{u}(x)\cdot\nabla_{\mathcal{D}}v(x)\,\mathrm{d}x + [P_{\mathcal{D}}\bar{u}, v]_{\mathcal{D},\alpha} = \int_{\Omega} f(x)v(x)\,\mathrm{d}x + T_{11}(v). \tag{3.16}$$

We first consider the second term of the left-hand side of (3.16). Using classical consistency error estimation (also used in the proof of Lemma 2.1), one has:

$$[P_{\mathcal{D}}\bar{u}, v]_{\mathcal{D},\alpha} = -\int_{\Omega} \operatorname{div}(\alpha\nabla\bar{u})(x)v(x)\,\mathrm{d}x + T_{12}(v), \tag{3.17}$$

with

$$|T_{12}(v)| \leqslant \sum_{\sigma\in\mathcal{E}} \mathrm{m}(\sigma)|R_\sigma|\delta_\sigma v,$$

where $\delta_\sigma v = |v_K - v_L|$ if $\sigma = K|L$ is an interior edge, $\delta_\sigma v = |v_K|$ is $\sigma \in \mathcal{E}_{\mathrm{ext}}$ and $|R_\sigma| \leqslant C_{13}h_{\mathcal{D}}$. Using the Cauchy–Schwarz inequality, this leads to:

$$|T_{12}(v)| \leqslant C_{14}h_{\mathcal{D}}\|v\|_{\mathcal{D}}. \tag{3.18}$$

We now consider the first term of the left-hand side of (3.16). We have

$$\int_{\Omega} (\Lambda(x) - \alpha(x)I_d)\nabla_{\mathcal{D}}P_{\mathcal{D}}\bar{u}(x)\cdot\nabla_{\mathcal{D}}v(x)\,\mathrm{d}x = T_{13}(v) + T_{14}(v), \tag{3.19}$$

with

$$T_{13}(v) = \int_{\Omega} (\Lambda(x) - \alpha(x)I_d)\nabla\bar{u}(x)\cdot\nabla_{\mathcal{D}}v(x)\,\mathrm{d}x$$

and

$$|T_{14}(v)| \leqslant C_{15}\|\nabla_{\mathcal{D}}P_{\mathcal{D}}\bar{u} - \nabla\bar{u}\|_{L^2(\Omega)^d}\|\nabla_{\mathcal{D}}v\|_{L^2(\Omega)^d}.$$

Using Lemma 2.5 and Lemma 2.2, we obtain

$$|T_{14}(v)| \leqslant C_{16}h_{\mathcal{D}}\|v\|_{\mathcal{D}}. \tag{3.20}$$

We now compute $T_{13}(v)$. For $K \in \mathcal{M}$ and $\sigma \in \mathcal{E}$, let μ_K and μ_σ, respectively, be the mean values of $(\Lambda(x) - \alpha(x)I_d)\nabla\bar{u}$ on K and σ:

$$\mu_K = \frac{1}{\mathrm{m}(K)}\int_K (\Lambda(x) - \alpha(x)I_d)\nabla\bar{u}(x)\,\mathrm{d}x, \quad \mu_\sigma = \frac{1}{\mathrm{m}(\sigma)}\int_\sigma (\Lambda(x) - \alpha(x)I_d)\nabla\bar{u}(x)\,\mathrm{d}\gamma(x).$$

The regularity of $\bar{u}$, Λ and α gives, for all $K \in \mathcal{M}$ and all $\sigma \in \mathcal{E}_K$ (recall that $|\cdot|$ denotes the Euclidean norm in $\mathbb{R}^d$):

$$|\mu_K - \mu_\sigma| \leqslant C_{17}h_{\mathcal{D}}. \tag{3.21}$$

Indeed, C_{17} only depends on the L^∞-norms of Λ, α and $\nabla\bar{u}$ and on the L^∞-norms of the derivatives of Λ, α and $\nabla\bar{u}$.

We now use (3.21) in order to give a bound on $T_{13}(v)$ as a function of $h_{\mathcal{D}}$. Indeed, the definition of $\nabla_{\mathcal{D}}v$ leads to:

$$\begin{aligned}
T_{13}(v) &= \sum_{K\in\mathcal{M}} \mu_K\cdot\mathrm{m}(K)(\nabla_{\mathcal{D}}v)_K \\
&= \sum_{K\in\mathcal{M}}\left(\sum_{L\in\mathcal{N}_K}\mu_K\cdot A_{K,L}\,(v_L - v_K) - \sum_{\sigma\in\mathcal{E}_{K,\mathrm{ext}}}\mu_K\cdot A_{K,\sigma}v_K\right) \\
&= \sum_{K\in\mathcal{M}}\left(\sum_{L\in\mathcal{N}_K}\mu_{K|L}\cdot A_{K,L}\,(v_L - v_K) - \sum_{\sigma\in\mathcal{E}_{K,\mathrm{ext}}}\mu_\sigma\cdot A_{K,\sigma}v_K\right) + T_{15}(v),
\end{aligned}$$

with

$$|T_{15}(v)| \leqslant C_{17} h_{\mathcal{D}} \sum_{K\in\mathcal{M}} \left(\sum_{L\in\mathcal{N}_K} |A_{K,L}||v_L - v_K| + \sum_{\sigma\in\mathcal{E}_{K,\mathrm{ext}}} |A_{K,\sigma}||v_K| \right)$$
$$\leqslant C_{17} h_{\mathcal{D}} \left(\sum_{\sigma=K|L\in\mathcal{E}_{\mathrm{int}}} (|A_{K,L}| + |A_{L,K}|)|v_L - v_K| + \sum_{K\in\mathcal{M}} \sum_{\sigma\in\mathcal{E}_{K,\mathrm{ext}}} |A_{K,\sigma}||v_K| \right).$$

Since $A_{K,L} = \tau_{K|L}(x_{K|L} - x_K)$ and $A_{K,\sigma} = \tau_\sigma(x_\sigma - x_K)$, one deduces from the preceding inequality, thanks to the definition of $\theta_{\mathcal{D}}$ (which gives $\mathrm{d}(x_\sigma, x_K) \leqslant (d_{K,\sigma}/\theta)$ if $\sigma \in \mathcal{E}_K$) and using the Cauchy–Schwarz inequality:

$$|T_{15}(v)| \leqslant C_{18} h_{\mathcal{D}} \|v\|_{\mathcal{D}}. \tag{3.22}$$

We now remark that:

$$T_{13}(v) - T_{15}(v) = \sum_{K\in\mathcal{M}} \left(\sum_{L\in\mathcal{N}_K} \mu_{K|L} \cdot A_{K,L}\, (v_L - v_K) - \sum_{\sigma\in\mathcal{E}_{K,\mathrm{ext}}} \mu_\sigma \cdot A_{K,\sigma} v_K \right)$$
$$= \sum_{\sigma=K|L\in\mathcal{E}_{\mathrm{int}}} \mu_\sigma \cdot (x_L - x_K)\tau_\sigma(v_L - v_K) - \sum_{K\in\mathcal{M}} \sum_{\sigma\in\mathcal{E}_{K,\mathrm{ext}}} \mu_\sigma \cdot (x_\sigma - x_K)\tau_\sigma v_K. \tag{3.23}$$

For $\sigma \in \mathcal{E}_{\mathrm{int}}$, one has $\sigma = K|L$ and $(x_L - x_K) = d_\sigma \mathbf{n}_{K,\sigma}$, where $\mathbf{n}_{K,\sigma}$ is the normal vector to σ exterior to K.
For $\sigma \in \mathcal{E}_{\mathrm{ext}}$, one has $\sigma \in \mathcal{E}_K$. Thanks to the fact that under homogeneous Dirichlet boundary conditions, the gradient of $\bar{u}$ is normal to the boundary, using Assumption (3.12), we get that

$$\mu_\sigma \cdot (x_\sigma - x_K)\tau_\sigma = \int_\sigma (\Lambda(x) - \alpha(x) I_d)\nabla\bar{u}(x) \cdot \mathbf{n}_{\partial\Omega}(x)\, \mathrm{d}\gamma(x).$$

Then, one deduces from (3.23):

$$T_{13}(v) - T_{15}(v) = -\int_\Omega \mathrm{div}((\Lambda - \alpha I_d)\nabla\bar{u})(x) v(x)\, \mathrm{d}x. \tag{3.24}$$

Therefore, since $-\mathrm{div}(\Lambda\nabla\bar{u}) = f$, one has (3.16) with $T_{11}(v) = T_{12}(v) + T_{14}(v) + T_{15}(v)$. This gives, with (3.18), (3.20) and (3.22):

$$|T_{11}(v)| \leqslant C_{19} h_{\mathcal{D}} \|v\|_{\mathcal{D}}. \tag{3.25}$$

This concludes Step 1.

Step 2 Let $e_{\mathcal{D}} = P_{\mathcal{D}}\bar{u} - u_{\mathcal{D}}$ be the discrete discretization error. Using (3.16) and (3.1) gives, for all $v \in H_{\mathcal{D}}$:

$$\int_{\Omega} (\Lambda(x) - \alpha(x) I_d) \nabla_{\mathcal{D}} e_{\mathcal{D}}(x) \cdot \nabla_{\mathcal{D}} v(x)\, \mathrm{d}x + [e_{\mathcal{D}}, v]_{\mathcal{D},\alpha} = T_{11}(v).$$

Taking $v = e_{\mathcal{D}}$ in this formula gives, with (3.25), $[e_{\mathcal{D}}, e_{\mathcal{D}}]_{\mathcal{D},\alpha} \leqslant C_{19} h_{\mathcal{D}} \|e_{\mathcal{D}}\|_{\mathcal{D}}$ and then, with $C_{20} = C_{19}/\alpha_0$ (since $\alpha_0 \|e_{\mathcal{D}}\|_{\mathcal{D}}^2 \leqslant [e_{\mathcal{D}}, e_{\mathcal{D}}]_{\mathcal{D},\alpha}$):

$$\|e_{\mathcal{D}}\|_{\mathcal{D}} \leqslant C_{20} h_{\mathcal{D}}, \tag{3.26}$$

which is exactly (3.13).

Using the discrete Poincaré estimate (2.8) and the fact that $\bar{u} \in C(\overline{\Omega})$, one deduces (3.14) from (3.13).

The last estimate, inequality (3.15), is a direct consequence of (3.26), (2.16) and (2.13). This concludes the first part of the theorem, i.e. assuming (3.12).

If $\mathcal{D}$ no longer satisfies the hypothesis (3.12), one has to replace (3.24) by:

$$T_{13}(v) - T_{15}(v) = -\int_{\Omega} \mathrm{div}((\Lambda - \alpha I_d)\nabla \bar{u})(x) v(x)\, \mathrm{d}x + T_{15}(v),$$

where, recalling that by z_σ the orthogonal projection of x_K on σ (see Definition 2.1):

$$T_{16}(v) = \sum_{K \in \mathcal{M}} \sum_{\sigma \in \mathcal{E}_{K,\mathrm{ext}}} \mu_\sigma \cdot (z_\sigma - x_\sigma) \tau_\sigma v_K.$$

Thanks to the Cauchy–Schwarz inequality, we get

$$T_{16}(v)^2 \leqslant \sum_{K \in \mathcal{M}} \sum_{\sigma \in \mathcal{E}_{K,\mathrm{ext}}} \tau_\sigma \mu_\sigma^2 (\mathrm{diam}(K))^2 \sum_{K \in \mathcal{M}} \sum_{\sigma \in \mathcal{E}_{K,\mathrm{ext}}} \tau_\sigma v_K^2,$$

which leads to

$$T_{16}(v)^2 \leqslant \frac{h_{\mathcal{D}}}{\theta} \mathrm{m}(\partial\Omega) \|\nabla \bar{u}\|_\infty^2 \|v\|_{\mathcal{D}}^2,$$

where $\mathrm{m}(\partial\Omega)$ is the $(d-1)$-dimensional Lebesgue measure of $\partial\Omega$. This gives (3.25) with $h_{\mathcal{D}}^{\frac{1}{2}}$ instead of $h_{\mathcal{D}}$. Following Step 2, this allows to conclude the proof. □

We now want to derive an error estimate when the solution of (1.5) is in $H^2(\Omega)$ instead of $\mathrm{C}^2(\overline{\Omega})$, in the case where the space dimension is less than or equal to 3. Indeed, the C^2-regularity of the solution of (1.5) was used, in the preceding proofs, only four times, namely, to prove (2.18), (2.19) and (2.23) in Lemma 2.5 and to prove (3.21) in Theorem 3.2 (in fact, it is also used for the classical consistency error (3.17), but, for this term, the generalization to the case where the solution of (1.5) is in $H^2(\Omega)$ instead of $\mathrm{C}^2(\overline{\Omega})$, in the case $d \leqslant 3$, has already been done in Eymard *et al.* (2000)). We will now prove similar inequalities for $\bar{u} \in H^2(\Omega) \cap H_0^1(\Omega)$ (instead of $\bar{u} \in \mathrm{C}^2(\Omega)$ with $\bar{u} = 0$ on the boundary of Ω) which will allow us to obtain the desired error estimate.

LEMMA 3.2 (Consistency of the gradient, $\bar{u} \in H^2(\Omega)$) Under hypothesis (1.2), with $d \leqslant 3$, let $\mathcal{D}$ be an admissible finite-volume discretization in the sense of Definition 2.1, and let $\theta \in (0, \theta_{\mathcal{D}}]$. Let $\bar{u} \in H^2(\Omega) \cap H_0^1(\Omega)$. Then, there exists C_{21}, only depending on Ω, θ and $\bar{u}$, such that:

$$\|\nabla_{\mathcal{D}}(P_{\mathcal{D}}\bar{u}) - \nabla \bar{u}\|_{L^2(\Omega)^d} \leqslant C_{21} h_{\mathcal{D}} \|\bar{u}\|_{H^2(\Omega)}. \tag{3.27}$$

(Recall that $P_{\mathcal{D}}$ is defined in (2.5) and $\nabla_{\mathcal{D}}$ in Definition 2.3.)

Proof. The proof follows along the same lines as the proof of Lemma 2.5 (in particular, recall that $H^2(\Omega) \subset C(\overline{\Omega})$ since $d \leqslant 3$). The C^2-regularity was only used to prove (2.18), (2.19) and (2.23). We now prove similar inequalities in the case $\bar{u} \in H^2(\Omega)$.

We begin with providing inequalities similar to (2.18) and (2.19). We denote by $(\nabla\bar{u})_\sigma$ the mean value of $\nabla\bar{u}$ on σ (recall that $(\nabla\bar{u})_K$ is the mean value of $\nabla\bar{u}$ on K). We use inequality (9.63) of Eymard *et al.* (2000) (in the proof of Theorem 9.4, using the H^2-regularity). This inequality states the existence of C_{22}, only depending on d and θ, such that, for all $\sigma = K|L \in \mathcal{E}_{\mathrm{int}}$:

$$|E_\sigma|^2 \leqslant C_{22}\frac{h_{\mathcal{D}}^2}{\mathrm{m}(\sigma)d_\sigma}\int_{D_\sigma}|H(\bar{u})(\mathrm{z})|^2\,\mathrm{dz},\quad \text{with } E_\sigma = (\nabla\bar{u})_\sigma\cdot\mathbf{n}_{K,\sigma} - \frac{\bar{u}(x_L)-\bar{u}(x_K)}{d_\sigma}, \tag{3.28}$$

and, for all $\sigma \in \mathcal{E}_{\mathrm{ext}}$, if $\sigma \in \mathcal{E}_K$:

$$|E_\sigma|^2 \leqslant C_{22}\frac{h_{\mathcal{D}}^2}{\mathrm{m}(\sigma)d_\sigma}\int_{D_\sigma}|H(\bar{u})(\mathrm{z})|^2\,\mathrm{dz},\ \text{with } E_\sigma = (\nabla\bar{u})_\sigma\cdot\mathbf{n}_{K,\sigma} - \frac{-\bar{u}(x_K)}{d_{\mathrm{K},\sigma}}, \tag{3.29}$$

where:

$$|H(\bar{u})(\mathrm{z})|^2 = \sum_{i,j=1}^{d}|D_iD_j\bar{u}(\mathrm{z})|^2.$$

We have now to compare $(\nabla\bar{u})_\sigma$ and $(\nabla\bar{u})_K$. This is possible thanks to inequality (9.38) in Lemma 9.4 of Eymard *et al.* (2000). Following this result, there exists C_{23}, only depending on d and θ, such that, for all $K \in \mathcal{M}$, all $\sigma \in \mathcal{E}_K$ and all $v \in H^1(K)$:

$$\begin{aligned}\left|\frac{1}{\mathrm{m}(K)}\int_K v(x)\,\mathrm{d}x - \frac{1}{\mathrm{m}(\sigma)}\int_\sigma v(x)\,\mathrm{d}\gamma(x)\right|^2 &\leqslant C_{23}\frac{\mathrm{diam}(K)}{\mathrm{m}(\sigma)}\int_K|\nabla v(x)|^2\,\mathrm{d}x\\ &\leqslant 2C_{23}\frac{h_{\mathcal{D}}^2}{\mathrm{m}(\sigma)d_\sigma}\int_K|\nabla v(x)|^2\,\mathrm{d}x.\end{aligned} \tag{3.30}$$

Using (3.30) with the derivatives of u, one deduces from (3.28) and (3.29), that there exists some real value C_{24} only depending on d and θ, such that

$$|e_\sigma|^2 \leqslant C_{24}\frac{h_{\mathcal{D}}^2}{\mathrm{m}(\sigma)d_\sigma}\int_{D_\sigma}|H(\bar{u})(\mathrm{z})|^2\,\mathrm{dz},\quad \text{with } e_\sigma = (\nabla\bar{u})_K\cdot\mathbf{n}_{K,\sigma} - \frac{\bar{u}(x_L)-\bar{u}(x_K)}{d_\sigma}, \tag{3.31}$$

and, for all $\sigma \in \mathcal{E}_{\mathrm{ext}}$, if $\sigma \in \mathcal{E}_K$:

$$|e_\sigma|^2 \leqslant C_{24}\frac{h_{\mathcal{D}}^2}{\mathrm{m}(\sigma)d_\sigma}\int_{D_\sigma}|H(\bar{u})(\mathrm{z})|^2\,\mathrm{dz},\quad \text{with } e_\sigma = (\nabla\bar{u})_K\cdot\mathbf{n}_{K,\sigma} - \frac{-\bar{u}(x_K)}{d_{\mathrm{K},\sigma}}. \tag{3.32}$$

Since $|R_K| \leqslant \sum_{\sigma\in\mathcal{E}_K}\frac{\mathrm{m}(\sigma)d_{K,\sigma}}{\theta}|e_\sigma|$ (where R_K is defined in (2.20)), using the Cauchy–Schwarz inequality, (3.31) and (3.32) lead to the following bound:

$$\begin{aligned}R_K^2 &\leqslant \frac{1}{\theta^2}\sum_{\sigma\in\mathcal{E}_K}\mathrm{m}(\sigma)d_{K,\sigma}\sum_{\sigma\in\mathcal{E}_K}\mathrm{m}(\sigma)d_{K,\sigma}e_\sigma^2\\ &\leqslant \frac{d\mathrm{m}(K)}{\theta^2}\sum_{\sigma\in\mathcal{E}_K}\mathrm{m}(\sigma)d_{K,\sigma}C_{24}\frac{h_{\mathcal{D}}^2}{\mathrm{m}(\sigma)d_\sigma}\int_{D_\sigma}|H(\bar{u})(\mathrm{z})|^2\,\mathrm{dz}\end{aligned}$$

and since $d_{K,\sigma} \leqslant d_\sigma$ and $\theta_{\mathcal{D}} \geqslant \theta$:

$$\left(\frac{R_K}{\mathrm{m}(K)}\right)^2 \mathrm{m}(K) \leqslant \frac{d\,C_{24}}{\theta^2} h_{\mathcal{D}}^2 \sum_{\sigma\in\mathcal{E}_K} \int_{D_\sigma} |H(\bar{u})(\mathrm{z})|^2\,\mathrm{dz}.$$

Then, (2.22) becomes:

$$\sum_{K\in\mathcal{M}} |(\nabla_{\mathcal{D}} P_{\mathcal{D}}\bar{u})_K - (\nabla\bar{u})_K|^2 \mathrm{m}(K) \leqslant \sum_K \frac{d\,C_{24}}{\theta^2} h_{\mathcal{D}}^2 \sum_{\sigma\in\mathcal{E}_K} \int_{D_\sigma} |H(\bar{u})(\mathrm{z})|^2\,\mathrm{dz},$$

which gives the existence of C_{25}, only depending on d and θ such that:

$$\sum_{K\in\mathcal{M}} |(\nabla_{\mathcal{D}} P_{\mathcal{D}}\bar{u})_K - (\nabla\bar{u})_K|^2 \mathrm{m}(K) \leqslant C_{25} h_{\mathcal{D}}^2 \|\bar{u}\|_{H^2(\Omega)}^2. \tag{3.33}$$

We have now to obtain an inequality similar to (2.23) (but without using $\bar{u} \in \mathrm{C}^2(\overline{\Omega})$). We will use here the fact that $d_{K,\sigma} \geqslant \theta \mathrm{diam}(K)$ if $\sigma \in \mathcal{E}_K$.

If ω is a convex, bounded, open subset of $\mathbb{R}^d$, the well-known 'Mean Poincaré Inequality' gives, for all $v \in H^1(\omega)$:

$$\int_\omega |v(x) - m_\omega v|^2\,\mathrm{d}x \leqslant \frac{1}{\mathrm{m}(\omega)} d_\omega^2 \mathrm{m}(B(0, d_\omega)) \int_\omega |\nabla v(x)|^2\,\mathrm{d}x, \tag{3.34}$$

where $m_\omega(v)$ is the mean value of v on ω, d_ω is the diameter of ω, $B(a, \delta)$ is the ball in $\mathbb{R}^d$ of centre a and radius δ and $\mathrm{m}(\omega)$ (resp. $\mathrm{m}(B(a, \delta))$ is the d-dimensional Lebesgue measure of ω (resp. $B(a, \delta)$). (A discrete counterpart of (3.34) is given, for instance, in Eymard *et al.* (2000, Lemma 10.2.))
Let $K \in \mathcal{M}$. We will use (3.34) for $\omega = K$. Since $d_{K,\sigma}$ is the distance between x_K to σ (for $\sigma \in \mathcal{E}_K$), there exists $\sigma \in \mathcal{E}_K$ such that $B(x_K, d_{K,\sigma}) \subset K$. Then, one has $\mathrm{m}(B(0, 1))d_{K,\sigma}^d = \mathrm{m}(B(x_K, d_{K,\sigma})) \leqslant \mathrm{m}(K)$ and, using $d_{K,\sigma} \geqslant \theta \mathrm{diam}(K)$, one obtains:

$$\mathrm{m}(K) \geqslant \mathrm{m}(B(0, 1))(\theta)^d (\mathrm{diam}(K))^d. \tag{3.35}$$

Taking $\omega = K$ in (3.34) gives, for all $K \in \mathcal{M}$ and all $v \in H^1(K)$:

$$\int_K |v(x) - m_\omega v|^2\,\mathrm{d}x \leqslant \frac{1}{\theta^d} \mathrm{diam}(K)^2 \int_K |\nabla v(x)|^2\,\mathrm{d}x. \tag{3.36}$$

Taking v equal to the derivatives of $\bar{u}$ (which are in $H^1(K)$ for all $K \in \mathcal{M}$) in (3.36) gives the existence of C_{26}, only depending on d and θ, such that:

$$\sum_{K\in\mathcal{M}} \int_K |\nabla\bar{u}(x) - (\nabla\bar{u})_K|^2\,\mathrm{d}x \leqslant C_{26} h_{\mathcal{D}}^2 \|\bar{u}\|_{H^2(\Omega)}^2. \tag{3.37}$$

Then, we conclude as in Lemma 2.5, using (3.33) and (3.37), that there exists C_{21} only depending on Ω, θ and $\bar{u}$ such that (3.27) holds. □

THEOREM 3.3 (H^2 error estimate) Assume hypotheses (1.2)–(1.4) and (3.6), with $d \leqslant 3$ and that Λ and α are of class C^1 on $\overline{\Omega}$. Let $\mathcal{D}$ be an admissible finite-volume discretization in the sense of Definition 2.1, and let $\theta \in (0, \theta_{\mathcal{D}}]$. We assume that the $\mathrm{card}(\mathcal{E}_K) \leqslant \frac{1}{\theta}$ for all $K \in \mathcal{M}$. Let $u_{\mathcal{D}} \in H_{\mathcal{D}}$ be the solution of

(3.1) and $\bar{u} \in H_0^1(\Omega)$ be the solution of (1.5). We assume that $\bar{u} \in H^2(\Omega)$ (which is necessarily true if Ω is convex).

Let us first assume that hypothesis (3.12) holds. Then, there exists C_{27}, only depending on Ω, θ, α, β, Λ and $\|\bar{u}\|_{H^2(\Omega)}$, such that:

$$\|u_{\mathcal{D}} - P_{\mathcal{D}}\bar{u}\|_{\mathcal{D}} \leqslant C_{27} h_{\mathcal{D}}, \tag{3.38}$$

$$\|u_{\mathcal{D}} - \bar{u}\|_{L^2(\Omega)} \leqslant C_{27} h_{\mathcal{D}} \tag{3.39}$$

and

$$\|\nabla_{\mathcal{D}} u_{\mathcal{D}} - \nabla\bar{u}\|_{L^2(\Omega)^d} \leqslant C_{27} h_{\mathcal{D}}. \tag{3.40}$$

(Recall that $H_{\mathcal{D}}$, $\nabla_{\mathcal{D}}$ and $\|\cdot\|_{\mathcal{D}}$ are defined in Definition 2.3, $P_{\mathcal{D}}$ is defined in (2.5).)

Let us then assume that (3.12) no longer holds, then there exists C_{28}, only depending on $\Omega, \theta, \alpha, \beta$, Λ and $\|\bar{u}\|_{H^2(\Omega)}$, such that (3.38), (3.39) and (3.40) hold with $C_{28}\sqrt{h_{\mathcal{D}}}$ instead of $C_{27}h_{\mathcal{D}}$.

Proof. The proof of Theorem 3.3 follows the same lines as the proof of Theorem 3.2. The quantities C_{25} and C_{26}, depending on θ, are now used to get a bound for $T_{12}(v)$ (as in Aavatsmark *et al.*, 1998a), and the quantity C_{16}, also depending on θ since it is obtained with (3.27) (Lemma 3.2) instead of (2.16) (Lemma 3.2), is used to obtain a bound for $T_{14}(v)$.

In order to obtain a bound for $T_{15}(v)$ (and then to conclude the proof of Theorem 3.3), we need to obtain an inequality similar to (3.21) (where the C^2-regularity of $\bar{u}$ was used), which gives a bound for the difference between the mean values of $(\Lambda(x) - \alpha(x) I_d)\nabla\bar{u}$ on K and σ if $\sigma \in \mathcal{E}_K$. Here, we will obtain a bound for the difference between these mean values using once again the consequence (3.30) of inequality (9.38) in Lemma 9.4 of Eymard *et al.* (2000). Applying (3.30) to the derivatives of $(\Lambda - \alpha I_d)\nabla\bar{u}$, there exists C_{29} only depending on Ω, θ, Λ and α (indeed, the C^1-norms of Λ and α), such that, for all $K \in \mathcal{M}$, all $\sigma \in \mathcal{E}_K$ and all $v \in H^1(K)$:

$$|\mu_K - \mu_\sigma|^2 \leqslant C_{29} \frac{\mathrm{diam}(K)}{\mathrm{m}(\sigma)} \|\bar{u}\|^2_{H^2(K)}. \tag{3.41}$$

Following the proof of Theorem 3.2, (3.41) is used to obtain a bound for $T_{15}(v)$:

$$\begin{aligned}
|T_{15}(v)| &\leqslant \sum_{K\in\mathcal{M}} \left(\sum_{L\in\mathcal{N}_K} |\mu_{K|L} - \mu_K||A_{K,L}\ (v_L - v_K)| + \sum_{\sigma\in\mathcal{E}_{K,\mathrm{ext}}} |\mu_\sigma - \mu_K||A_{K,\sigma} v_K| \right) \\
&\leqslant \sum_{\sigma=K|L\in\mathcal{E}_{\mathrm{int}}} \frac{|\mu_\sigma - \mu_K| + |\mu_\sigma - \mu_L|}{\theta} \mathrm{m}(\sigma) d_\sigma \frac{\delta_\sigma v}{d_\sigma} + \sum_{\sigma\in\mathcal{E}_{\mathrm{ext}}} \frac{|\mu_\sigma - \mu_K|}{\theta} \mathrm{m}(\sigma) d_\sigma \frac{\delta_\sigma v}{d_\sigma},
\end{aligned}$$

where, in the last term, K is such that $\sigma \in \mathcal{E}_K$ and where $\delta_\sigma v = |v_K - v_L|$ if $\sigma = K|L \in \mathcal{E}_{\mathrm{int}}$ and $\delta_\sigma v = |v_K|$ if $\sigma \in \mathcal{E}_{\mathrm{ext}} \cap \mathcal{E}_K$. (We also used the fact that $|A_{K,L}| \leqslant \frac{\mathrm{m}(\sigma)}{\theta}$ and $|A_{K,\sigma}| \leqslant \frac{\mathrm{m}(\sigma)}{\theta}$, thanks to $\theta_{\mathcal{D}} \geqslant \theta$.)

Then, using Cauchy–Schwarz inequality and (3.41), one obtains:

$$\begin{aligned}
|T_{15}(v)| \leqslant \|v\|_{\mathcal{D}} \frac{\sqrt{2C_5}}{\theta} \Bigg(&\sum_{\sigma=K|L\in\mathcal{E}_{\mathrm{int}}} d_\sigma (\mathrm{diam}(K)\|\bar{u}\|^2_{H^2(K)} + \mathrm{diam}(L)\|\bar{u}\|^2_{H^2(L)}) \\
&+ \sum_{\sigma\in\mathcal{E}_{\mathrm{ext}}} d_\sigma \mathrm{diam}(K)\|\bar{u}\|^2_{H^2(K)} \Bigg)^{\frac{1}{2}}.
\end{aligned}$$

Using $d_\sigma \leqslant 2h_{\mathcal{D}}$, $\operatorname{diam}(K) \leqslant h_{\mathcal{D}}$ and the fact that $\operatorname{card}(\mathcal{E}_K) \leqslant \frac{1}{\theta}$ for all $K \in \mathcal{M}$, one deduces the existence of C_6, only depending on Ω, θ, Λ and α, such that:

$$|T_{15}(v)| \leqslant C_6 h_{\mathcal{D}} \|\bar{u}\|_{H^2(\Omega)} \|v\|_{\mathcal{D}}. \tag{3.42}$$

Then, we conclude the proof of Theorem 3.3 exactly as in the proof of Theorem 3.2 ((3.42) replaces (3.22)). □

4. Numerical results

The gradient scheme was tested for various academic problems, for which the analytical solution is known. We first recall that, in the case where $\Lambda(x) = \alpha_0 I_d$ and $\alpha(x) = \alpha_0$ for some $\alpha_0 > 0$ and for a.e. $x \in \Omega$, the gradient scheme reduces to the classical cell-centred finite-volume scheme for the computation of the approximate solution, while providing in addition an approximate value for the gradient of the continuous solution. First note that in the classical cell-centred scheme, the equation relative to a given cell involves the neighbours of this cell, while in the gradient scheme, it involves the neighbours of this cell and the neighbours of the neighbours. Hence, in the case of a rectangular (resp. parallelepiped) mesh, the classical cell-centred scheme is a 5-point (resp. 7 point) scheme, while the gradient scheme is a 13-point (resp. 24 point) scheme. Similarly, if one uses a triangular (resp. tetrahedral) mesh, the classical scheme is a 4-point (resp. 5 point) scheme, while the gradient scheme is at most a 10-point (resp. at most 17 point) scheme. Hence, on isotropic problems, the gradient scheme is more expensive in terms of time and memory than the classical finite-volume scheme; however, in anisotropic cases where the classical finite-volume scheme no longer applies, its cost is comparable to that of the piecewise bilinear finite-element scheme in the case of a parallelepiped mesh (or a finite-volume scheme with vertex-centred reconstructed gradient), which leads to a 9-point scheme (2D) and a 27-point scheme (3D).

Experiments were carried out for $\Omega = (0,1) \times (0,1)$, using triangular and rectangular meshes. In order to obtain convergence ratios, the number of cells was varied from 1600 to 102400 for rectangular meshes (note that, in fact, rectangles are squares) and from 1400 to 89600 for triangular meshes.

Let us first consider the case $\Lambda(x) = I_d$ for a.e. $x \in \Omega$, with the solution of Problem (1.1) given by $\bar{u}(x_1, x_2) = x_1(1-x_1)x_2(1-x_2)$. In this case, we found an order 2 for the convergence of the approximate solution for both types of meshes and an order 1.5 for the gradient computed on rectangular meshes and 1 in the case of triangular meshes (the convergence rates being fitted, using a least-square regression on the logarithmic values of the errors and of the characteristic size of the mesh).

We then tested the gradient scheme for a heterogeneous anisotropic problem inspired by Le Potier (2005). Let us define $\bar{x} = (-0.1, -0.1)$, and for any $x = (x_1, x_2)$, denoting $\tilde{x}_i = x_i - \bar{x}_i$, $i = 1, 2$, set

$$\Lambda(x) = \begin{pmatrix} \tilde{x}_2^2 + \varepsilon \tilde{x}_1^2 & -(1-\varepsilon)\tilde{x}_1\tilde{x}_2 \\ -(1-\varepsilon)\tilde{x}_1\tilde{x}_2 & \tilde{x}_1^2 + \varepsilon \tilde{x}_2^2 \end{pmatrix},$$

for the three values $\varepsilon = 1$, $\varepsilon = 10^{-2}$ and $\varepsilon = 10^{-4}$. Then the eigenvalues of $\Lambda(x)$ are equal to $\underline{\lambda}(x) = \varepsilon|x - \bar{x}|^2$ and $\overline{\lambda}(x) = |x - \bar{x}|^2$: the anisotropy ratio is therefore $1/\varepsilon$ in the whole domain. Note that, thanks to the choice $\bar{x} = (-0.1, -0.1)$, we have $\inf_{x\in\Omega} \underline{\lambda}(x) = 0.02\varepsilon$ and $\sup_{x\in\Omega} \underline{\lambda}(x) = 2.42\varepsilon$ (similar relations hold on $\overline{\lambda}$) which corresponds to a highly heterogeneous case from both the point of view of the magnitude of the ratio of the eigenvalues and that of the directions of anisotropy: indeed, since these directions are not constant in space, one may not solve the problem by a classical finite-volume method on a tilted rectangular mesh.

REMARK 4.1 We have also successfully performed the computations with $\bar{x} = (0,0)$, as in Le Potier (2005), and obtained results similar to that of Le Potier (2005). However, in this case, one has $\inf_{x\in\Omega} \underline{\lambda}(x) = 0$, so that inequality (1.3) does not hold, and no theoretical estimate may be obtained in this case.

Assume that the solution of problem (1.1) is given by $\bar{u}(x) = \sin(\pi x_1)\sin(\pi x_2)$; in this case, $\|\bar{u}\|_{L^2(\Omega)} = 1/2$ and $\|\nabla\bar{u}\|_{L^2(\Omega)} = \pi/\sqrt{2}$, and the function f satisfies:

$$f(x) = \pi^2(1+\varepsilon)\sin(\pi x_1)\sin(\pi x_2)|x-\bar{x}|^2 + \pi(1-3\varepsilon)\cos(\pi x_1)\sin(\pi x_2)\tilde{x}_1 + \pi(1-3\varepsilon)\sin(\pi x_1)\cos(\pi x_2)\tilde{x}_2 + 2\pi^2(1-\varepsilon)\cos(\pi x_1)\cos(\pi x_2)\tilde{x}_1\tilde{x}_2.$$

The rates of convergence obtained in this case, with the choice $\alpha(x) = \underline{\lambda}(x)$ for a.e. $x\in\Omega$, are given in the following table.

	$\varepsilon = 1$		$\varepsilon = 10^{-2}$	$\varepsilon = 10^{-4}$
	Triangles	Rectangles	Rectangles	Rectangles
u	2.0	2.0	2.1	2.4
∇u	1.0	2.0	1.7	1.9

These results indicate that the use of the gradient scheme leads to a correct numerical behaviour, indeed, comparable with low-degree finite-element schemes on similar problems.

Next, we tested different choices for the function α, taken to be of the form $\alpha(x) = \widehat{\alpha}\,\overline{\lambda}(x)$ for a given $\widehat{\alpha} > 0$ and for a.e. $x\in\Omega$ (note that, for $\widehat{\alpha} > \varepsilon$, assumption (3.6) no longer holds). The results obtained are shown in Fig. 2. It seems that there is some dependence of the optimal value of $\widehat{\alpha}$ on the type of mesh and the eigenvalues of Λ and that this optimal value seems to differ for the solution and its gradient. In our numerical experiments, we found the choice $\alpha = \underline{\lambda}$ (namely, the lowest eigenvalue which was indeed chosen in the theoretical study) to be robust and convenient. We also performed some tests with large and small values of $\widehat{\alpha}$ (not shown in Fig. 2, for ease of readability) and found the error to tend to $1/2$ as $\widehat{\alpha}$ tends to infinity irrespective of the value of ε, whereas the error stays bounded and small as $\widehat{\alpha}$ tends to 0 (e.g. this error tends to 3.9×10^{-2} for $\varepsilon = 10^{-2}$ on the 160×160 mesh, or to 6.2×10^{-3} for $\varepsilon = 10^{-4}$ on the 320×320 mesh).

5. Conclusion

In this paper, we constructed a discrete gradient for piecewise constant functions. This discrete gradient exhibits several advantages: it is easy and cheap to compute, and it provides a simple scheme for the approximation of anisotropic convection-diffusion problems. We showed a weak convergence property of this discrete gradient to the gradient of the limit of the sequence of functions, together with a consistency property, both leading to the strong convergence of the discrete solution and of its discrete gradient in the case of a Dirichlet problem with full matrix diffusion.

Since this notion of admissible mesh includes Voronoï meshes, which are more and more used in practice, and which seem to remain tractable even in high space dimensions, applications to financial mathematics problems are being studied (Berton, 2006). Applications to finite-volume schemes for compressible Navier–Stokes equations are also expected to be succesful (Touazi, 2007). Further work includes a parametric study and the generalization to meshes without the orthogonality condition.

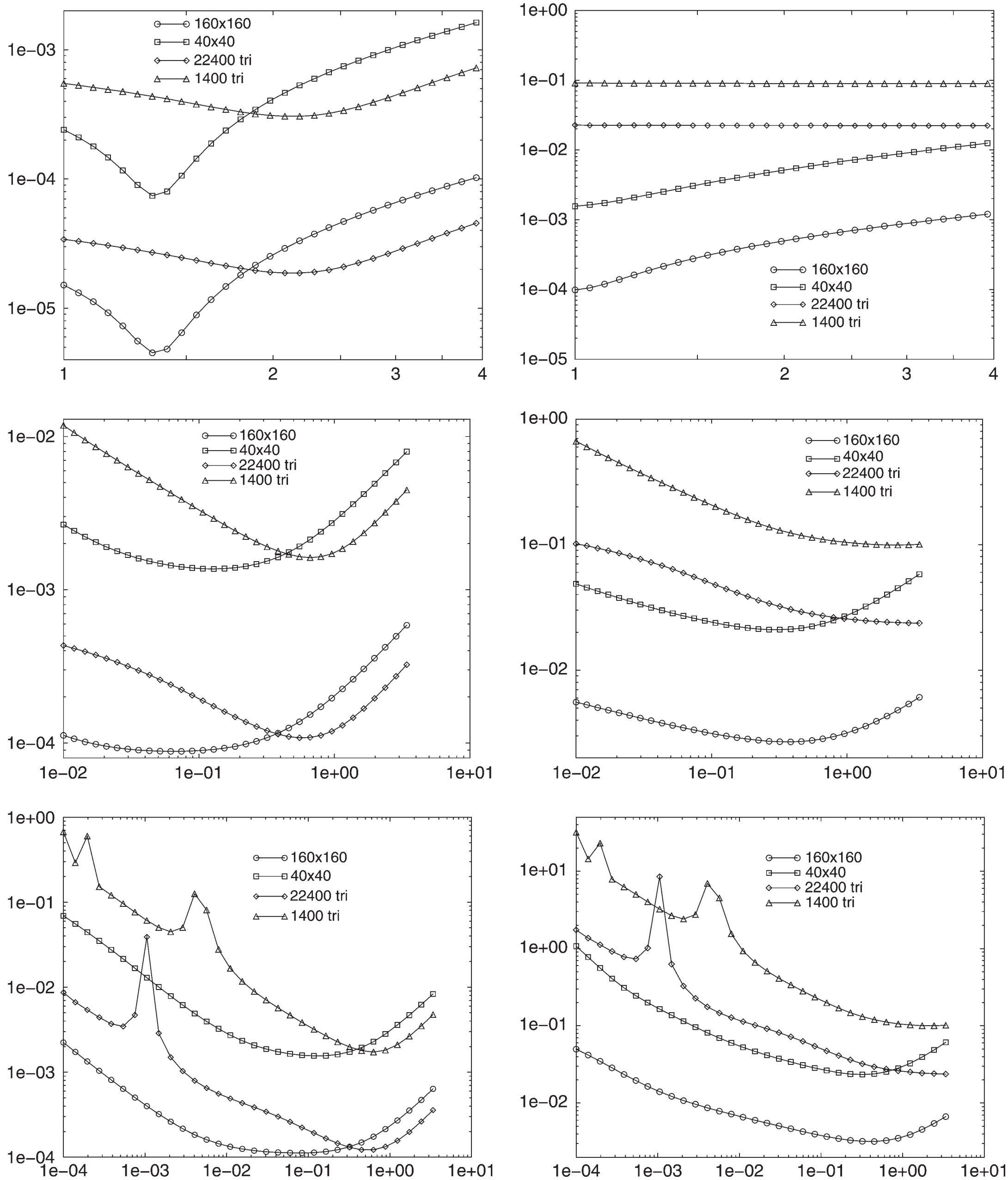

FIG. 2. Variation of the L^2 errors in the solution (left) and its gradient (right) for $\varepsilon = 1$ (top), $\varepsilon = 10^{-2}$ (middle), $\varepsilon = 10^{-4}$ (bottom) for the 40×40 and 160×160 rectangular meshes, and the 1400 and 22400 triangular meshes, with respect to the value of the parameter $\widehat{\alpha}$.

Acknowledgement

This work was supported by GDR MOMAS.

REFERENCES

AAVATSMARK, I., BARKVE, T., BOE, O. & MANNSETH, T. (1998a) Discretization on unstructured grids for inhomogeneous, anisotropic media. Part I: derivation of the methods. *SIAM J. Sci. Comput.*, **19**, 1700–1716.

AAVATSMARK, I., BARKVE, T., BOE, O. & MANNSETH, T. (1998b) Discretization on unstructured grids for inhomogeneous, anisotropic media. Part II: discussion and numerical results. *SIAM J. Sci. Comput.*, **19**, 1717–1736.

ANGERMANN, L. (2000) A finite element method for the numerical solution of convection-dominated anisotropic diffusion equations. *Numer. Math.*, **85**, 175–195.

ANGERMANN, L. (2003) Node-centered finite volume schemes and primal-dual mixed formulations. *Commun. Appl. Anal.*, **7**, 529–566.

BERTOLAZZI, E. & MANZINI, G. (2004) A cell-centered second-order accurate finite volume method for convection-diffusion problems on unstructured meshes. *Math. Models Methods Appl. Sci.*, **14**, 1235–1260.

BERTON, J. (2006) Comparaison de différentes méthodes pour apprécier les options américaines. *Thesis*, Marne-la-Vallée university, France, in preparation.

CHÉNIER, E., EYMARD, R. & NICOLAS, X. (2004) A finite volume scheme for the transport of radionucleides porous media: simulation of transport around a nuclear waste disposal site: the COUPLEX test cases. *Comput. Geosci.* (A. Bourgeat & M. Kern eds), **8**, 163–172.

COUDIÈRE, Y., VILA, J. P. & VILLEDIEU, P. (1999) Convergence rate of a finite volume scheme for a two-dimensional convection-diffusion problem. *M2AN Math. Model. Numer. Anal.*, **33**, 493–516.

DOMELEVO, K. & OMNES, P. (2005) A finite volume method for the Laplace equation on almost arbitrary two-dimensional grids. *ESAIM: Math. Model. and Num. Anal.* (to appear).

DRONIOU, J. & EYMARD, R. (2005) A mixed finite volume scheme for anisotropic diffusion problems on any grid. Preprint available at http://hal.ccsd.cnrs.fr/ccsd-00005565 (submitted).

EYMARD, R., GALLOUËT, T. & HERBIN, R. (1999) Convergence of finite volume approximations to the solutions of semilinear convection diffusion reaction equations. *Numer. Math.*, **82**, 91–116.

EYMARD, R., GALLOUËT, T. & HERBIN, R. (2000) Finite volume methods. *Handbook of Numerical Analysis*, vol. 7 (P. G. Ciarlet & J. L. Lions eds). North Holland, pp. 713–1020.

EYMARD, R., GALLOUËT, T. & HERBIN, R. (2001) Finite volume approximation of elliptic problems and convergence of an approximate gradient. *Appl. Numer. Math.*, **37**, 31–53.

EYMARD, R., GALLOUËT, T. & HERBIN, R. (2004a) A finite volume for anisotropic diffusion problems. *C. R. Acad. Sci.* , **339**, 299–302.

EYMARD, R., GALLOUËT, T., HERBIN, R. & MICHEL, A. (2002) Convergence of a finite volume scheme for nonlinear degenerate parabolic equations. *Numer. Math.*, **92**, 41–82.

EYMARD, R., HILHORST, D. & VOHRALÍK, M. (2004b) Combined finite volume-nonconforming/mixed-hybrid finite element scheme for degenerate parabolic problems (submitted).

FORSYTH, P. A. (1991) Control volume finite elements, a control volume finite element approach to NAPL groundwater contamination. *SIAM J. Sci. Stat. Comput.*, **12**, 1029–1057.

GALLOUËT, T., HERBIN, R. & VIGNAL, M. H. (2000) Error estimate for the approximate finite volume solutions of convection diffusion equations with general boundary conditions. *SIAM J. Numer. Anal.*, **37**, 1935–1972.

HERBIN, R. (1995) An error estimate for a finite volume scheme for a diffusion-convection problem on a triangular mesh. *Numer. Methods Partial Differ. Equ.*, **11**, 165–173.

HERBIN, R. (1996) Finite volume methods for diffusion convection equations on general meshes. *Finite Volumes for Complex Applications, Problems and Perspectives* (F. Benkhaldoun & R. Vilsmeier eds). Hermes, pp. 153–160.

HERBIN, R. & MARCHAND, E. (2001) Finite volume approximation of a class of variational inequalities. *IMA J. Numer. Anal.*, **21**, 553–585.

HERMELINE, F. (2000) A finite volume method for the approximation of diffusion operators on distorted meshes. *J. Comput. Phys.*, **160**, 481–499.

HU, X. H. & NICOLAIDES, R. A. (1992) Covolume techniques for anisotropic media. *Numer. Math.*, **61**, 215–234.

JAYANTHA, P. A. & TURNER, I. W. (2003) A second order finite volume technique for simulating transport in anisotropic media. *Int. J. Numer. Methods Heat Fluid Flow*, **13**, 31–56.

JAYANTHA, P. A. & TURNER, I. W. (2005) A second order control-volume finite-element least-squares strategy for simulating diffusion in strongly anisotropic media. *J. Comput. Math.*, **23**, 1–16.

LAMBERTON, D. & LAPEYRE, B. (1995) *An Introduction to Stochastic Calculus Applied to Finance*. Chapman and Hall.

LE POTIER, C. (2005) A finite volume method for the approximation of highly anisotropic diffusion operators on unstructured meshes. *Finite Volumes for Complex Applications IV, Marrakesh, Marocco*.

MISHEV, I. D. (1998) Finite volume methods on Voronoï meshes. *Numer. Methods. Partial Differ. Equ.*, **14**, 193–212.

NICOLAIDES, R. A. (1992) Direct discretization of planar div-curl problems. *SIAM J. Numer. Anal.*, **29**, 32–56.

PUTTI, M. & CORDES, C. (1998) Finite element approximation of the diffusion operator on tetrahedra. *SIAM J. Sci. Comput.*, **19**, 1154–1168.

TOUAZI, O. (2007) Mise en oeuvre d’un schéma de volumes finis pour les équations de Navier-Stokes compressibles. *Thesis*, Université de Marne-la-Vallée, France, in preparation.