A mixed finite volume scheme for anisotropic diffusion problems on any grid

Jérôme Droniou [1] and Robert Eymard [2]

08/08/2006

Abstract

We present a new finite volume scheme for anisotropic heterogeneous diffusion problems on unstructured irregular grids, which simultaneously gives an approximation of the solution and of its gradient. The approximate solution is shown to converge to the continuous one as the size of the mesh tends to 0, and an error estimate is given. An easy implementation method is then proposed, and the efficiency of the scheme is shown on various types of grids and for various diffusion matrices.

Keywords. Finite volume scheme, unstructured grids, irregular grids, anisotropic heterogeneous diffusion problems.

1 Introduction

The computation of an approximate solution for equations involving a second order elliptic operator is needed in so many physical and engineering areas, where the efficiency of some discretization methods, such as finite difference, finite element or finite volume methods, has been proved The use of finite volume methods is particularly popular in the oil engineering field, since it allows for coupled physical phenomena in the same grids, for which the conservation of various extensive quantities appears to be a main feature. However, it is more challenging to define convergent finite volume schemes for second-order elliptic operators on refined, distorted or irregular grids, designed for the purpose of another problem.

For example, in the framework of geological basin simulation, the grids are initially fitted on the geological layers boundaries, which is a first reason for the loss of orthogonality. Then, these grids are modified during the simulation, following the compaction of these layers (see [15]), thus leading to irregular grids, as those proposed by [17]. As a consequence, it is no longer possible to compute the fluxes resulting from a finite volume scheme for a second order operator by a simple two-point difference across each interface between two neighboring control volumes. Such a two-point scheme is consistent only in the case of an isotropic operator, using a grid such that the lines connecting the centers of the control volumes are orthogonal to the edges of the mesh. The problem of finding a consistent expression using only a small number of points, for the finite volume fluxes in the general case of any grid and any anisotropic second order operator, has led to many works (see [1], [2], [3], [15] and references therein; see also [21]). A recent finite volume scheme has been proposed [12, 13], permitting to obtain a convergence property in the case of an

[1] Département de Mathématiques, UMR CNRS 5149, CC 051, Université Montpellier II, Place Eugène Bataillon, 34095 Montpellier cedex 5, France. email: droniou@math.univ-montp2.fr

[2] Laboratoire d'Analyse et de Mathématiques Appliquées, UMR 8050, Université de Marne-la-Vallée, 5 boulevard Descartes, Champs-sur-Marne, 77454 Marne-la-Vallée Cedex 2, France. email:Robert.Eymard@univ-mlv.fr

anisotropic heterogeneous diffusion problem on unstructured grids, which all the same satisfy the above orthogonality condition. In the case where such an orthogonality condition is not satisfied, a classical method is the mixed finite element method which also gives an approximation of the fluxes and of the gradient of the unknown (see [4], [5], [6], [24] for example, among a very large literature). Note that, although the Raviart-Thomas basis is not directly available on control volumes which are not simplices or regular polyhedra, such a basis can be built on more general irregular grids. In [18], such a construction is completed using decomposition into simplices and a local elimination of the unknowns at the internal edges. In [10] and [14], such basis functions are obtained from the resolution of a Neumann elliptic problem in each grid block. However, it has been observed that the use of mixed finite element method could demand high refined grids on some highly heterogeneous and anisotropic cases (see [20] and the numerical results provided in the present paper). An improvement of the mixed finite element scheme is the expanded mixed finite element scheme [7], where different discrete approximations are proposed for the unknown, its gradient and the product of the diffusion matrix by the gradient of the unknown; however, this last scheme seems to present the same restrictions on the meshes as the mixed finite element scheme. [19] gives a review of different "mixed" methods, albeit mostly on structured (or not very general) grids.

We thus propose in this paper an original finite volume method, called the mixed finite volume method, which can be applied on any type of grids in any space dimension, with very few restrictions on the control volumes. The implementation of this scheme is proved to be easy, and no geometric complex shape function has to be computed. Accurate results are obtained on coarse irregular grids in the case of highly heterogeneous anisotropic problems. In order to show the mathematical and numerical properties of this scheme, we study here the following problem: find an approximation of $\bar{u}$, weak solution to the following problem:

$$\begin{aligned} &-\mathrm{div}(\Lambda\nabla\bar{u}) = f \text{ in } \Omega,\\ &\bar{u} = 0 \text{ on } \partial\Omega, \end{aligned} \tag{1}$$

under the following assumptions:

$$\Omega \text{ is an open bounded connected polygonal subset of } \mathbb{R}^d,\ d \geq 1, \tag{2}$$

$$\begin{gathered} \Lambda : \Omega \to M_d(\mathbb{R}) \text{ is a bounded measurable function such that}\\ \text{there exists } \alpha_0 > 0 \text{ satisfying } \Lambda(x)\xi\cdot\xi \geq \alpha_0|\xi|^2 \text{ for a.e. } x \in \Omega \text{ and all } \xi \in \mathbb{R}^d, \end{gathered} \tag{3}$$

(where $M_d(\mathbb{R})$ stands for the space of $d \times d$ real matrices) and

$$f \in L^2(\Omega). \tag{4}$$

Thanks to Lax-Milgram theorem, there exists a unique weak solution to (1) in the sense that $\bar{u} \in H^1_0(\Omega)$ and the equation is satisfied in the sense of distributions on Ω.

The principle of the mixed finite volume scheme, described in Section 2, is the following. We simultaneously look for approximations u_K and $\mathbf{v}_K$ of $\bar{u}$ and $\nabla\bar{u}$ in each control volume K, and for approximations F_σ of $\int_\sigma \Lambda(x)\nabla\bar{u}(x)\cdot\mathbf{n}_\sigma\,\mathrm{d}\gamma(x)$ at each edge σ of the mesh, where $\mathbf{n}_\sigma$ is a unit vector normal to σ. The values F_σ must then satisfy the conservation equation in each control volume, and consistency relations are imposed on u_K, $\mathbf{v}_K$ and F_σ. After having investigated in Section 3 the properties of a space associated with the scheme, we show in Section 4 that it leads to a linear system which has one and only one approximate solution u, $\mathbf{v}$ and F, and we provide the mathematical analysis of its convergence and give an error estimate. In Section 5, we propose an easy implementation procedure for the scheme, and we use it for the study of some numerical examples. We thus obtain acceptable results on some grids for which it would be complex to use other methods, or to which empirical methods apply but no mathematical result of convergence nor stability has yet been obtained.

2 Definition of the mixed finite volume scheme and main results

We first present the notion of admissible discretization of the domain Ω, which is necessary to give the expression of the mixed finite volume scheme.

Definition 2.1 [Admissible discretization] *Let Ω be an open bounded polygonal subset of $\mathbb{R}^d$ ($d \geq 1$), and $\partial\Omega = \overline{\Omega} \setminus \Omega$ its boundary. An admissible finite volume discretization of Ω is given by $\mathcal{D} = (\mathcal{M}, \mathcal{E}, \mathcal{P})$, where:*

- *$\mathcal{M}$ is a finite family of non empty open polygonal convex disjoint subsets of Ω (the "control volumes") such that $\overline{\Omega} = \cup_{K \in \mathcal{M}} \overline{K}$.*
- *$\mathcal{E}$ 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 an affine hyperplane E of $\mathbb{R}^d$ and $K \in \mathcal{M}$ with $\sigma \subset \partial K \cap E$ and σ is a non empty open convex subset of E. We assume that, for all $K \in \mathcal{M}$, there exists a subset $\mathcal{E}_K$ of $\mathcal{E}$ such that $\partial K = \cup_{\sigma \in \mathcal{E}_K} \overline{\sigma}$. We also assume that, for all $\sigma \in \mathcal{E}$, either $\sigma \subset \partial\Omega$ or $\overline{\sigma} = \overline{K} \cap \overline{L}$ for some $(K, L) \in \mathcal{M} \times \mathcal{M}$.*
- *$\mathcal{P}$ is a family of points of Ω indexed by $\mathcal{M}$, denoted by $\mathcal{P} = (\mathbf{x}_K)_{K \in \mathcal{M}}$ and such that, for all $K \in \mathcal{M}$, $\mathbf{x}_K \in K$.*

Some examples of admissible meshes in the sense of the above definition are shown in Figures 1 and 2 in Section 5.

Remark 2.1 *Though the elements of $\mathcal{E}_K$ may not be the real edges of a control volume K (each $\sigma \in \mathcal{E}_K$ may be only a part of a full edge, see figure 2), we will in the following call "edges of K" the elements of $\mathcal{E}_K$. Notice that we could also cut each intersection $\overline{K} \cap \overline{L}$ into more than one edge without changing neither our study nor our results.*

Remark 2.2 *The whole mathematical study done in this paper applies whatever the choice of the point $\mathbf{x}_K$ in each $K \in \mathcal{M}$. In particular, we do not impose any orthogonality condition connecting the edges and the points $\mathbf{x}_K$. However, the magnitude of the numerical error (and, for some regular or structured types of mesh, its order) does depend on this choice.*
We could also extend our definition to non-planar edges, under some curvature condition. In this case, it remains possible to use the mixed finite volume scheme and to prove its convergence.

The following notations are used. The measure of a control volume K is denoted by $\mathrm{m}(K)$; the $(d-1)$-dimensional measure of an edge σ is $\mathrm{m}(\sigma)$. In the case where $\sigma \in \mathcal{E}$ is such that $\overline{\sigma} = \overline{K} \cap \overline{L}$ for $(K, L) \in \mathcal{M} \times \mathcal{M}$, we denote $\sigma = K|L$. For all $\sigma \in \mathcal{E}$, $\mathbf{x}_\sigma$ is the barycenter of σ. If $\sigma \in \mathcal{E}_K$ then $\mathbf{n}_{K,\sigma}$ is the unit normal to σ outward to K. The set of interior (resp. boundary) edges is

denoted by $\mathcal{E}_{\rm int}$ (resp. $\mathcal{E}_{\rm ext}$), that is $\mathcal{E}_{\rm int} = \{\sigma \in \mathcal{E};\ \sigma \not\subset \partial\Omega\}$ (resp. $\mathcal{E}_{\rm ext} = \{\sigma \in \mathcal{E};\ \sigma \subset \partial\Omega\}$). For all $K \in \mathcal{M}$, we denote by $\mathcal{N}_K$ the subset of $\mathcal{M}$ of the neighboring control volumes (that is, the L such that $\overline{K} \cap \overline{L}$ is an edge of the discretization).
To study the convergence of the scheme, we will need the following two quantities: the size of the discretization

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

and the regularity of the discretization

$$\mathrm{regul}(\mathcal{D}) = \sup\left\{\max\left(\frac{\mathrm{diam}(K)^d}{\rho_K^d}, \mathrm{Card}(\mathcal{E}_K)\right)\ ;\ K \in \mathcal{M}\right\} \tag{5}$$

where, for $K \in \mathcal{M}$, ρ_K is the supremum of the radius of the balls contained in K. Notice that, for all $K \in \mathcal{M}$,

$$\mathrm{diam}(K)^d \le \mathrm{regul}(\mathcal{D})\rho_K^d \le \frac{\mathrm{regul}(\mathcal{D})}{\omega_d}\mathrm{m}(K) \tag{6}$$

where ω_d is the volume of the unit ball in $\mathbb{R}^d$. Note also that $\mathrm{regul}(\mathcal{D})$ does not increase in a local refinement procedure, which will allow the scheme to handle such procedures.

We now define the mixed finite volume scheme. Let $\mathcal{D}$ be an admissible discretization of Ω in the sense of Definition 2.1. Denote by $H_\mathcal{D}$ the set of real functions on Ω which are constant on each control volume $K \in \mathcal{M}$ (if $h \in H_\mathcal{D}$, we let h_K be its value on K).
As said in the introduction, the idea is to consider three sets of unknowns, namely $u \in H_\mathcal{D}$ which approximates $\bar{u}$, $\mathbf{v} \in H_\mathcal{D}^d$ which approximates $\nabla\bar{u}$ and a family of real numbers $F = (F_{K,\sigma})_{K\in\mathcal{M},\sigma\in\mathcal{E}_K}$ (we denote by $\mathcal{F}_\mathcal{D}$ the set of such families) which approximates $(\int_\sigma \Lambda(x)\nabla\bar{u}(x)\cdot \mathbf{n}_{K,\sigma}\, d\gamma(x))_{K\in\mathcal{M},\sigma\in\mathcal{E}_K}$.
Taking $\nu = (\nu_K)_{K\in\mathcal{M}}$ a family of nonnegative numbers, we define $L_\nu(\mathcal{D})$ as the space of $(u, \mathbf{v}, F) \in H_\mathcal{D} \times H_\mathcal{D}^d \times \mathcal{F}_\mathcal{D}$ such that

$$\begin{aligned} &\mathbf{v}_K \cdot (\mathbf{x}_\sigma - \mathbf{x}_K) + \mathbf{v}_L \cdot (\mathbf{x}_L - \mathbf{x}_\sigma) + \nu_K \mathrm{m}(K) F_{K,\sigma} - \nu_L \mathrm{m}(L) F_{L,\sigma} = u_L - u_K, \\ &\qquad\qquad\qquad\qquad\qquad\qquad \forall K \in \mathcal{M},\ \forall L \in \mathcal{N}_K,\ \text{with } \sigma = K|L, \\ &\mathbf{v}_K \cdot (\mathbf{x}_\sigma - \mathbf{x}_K) + \nu_K \mathrm{m}(K) F_{K,\sigma} = -u_K, \quad \forall K \in \mathcal{M},\ \forall \sigma \in \mathcal{E}_K \cap \mathcal{E}_{\rm ext} \end{aligned} \tag{7}$$

and we define the mixed finite volume scheme as: find $(u, \mathbf{v}, F) \in L_\nu(\mathcal{D})$ such that

$$F_{K,\sigma} + F_{L,\sigma} = 0, \quad \forall \sigma = K|L \in \mathcal{E}_{\rm int}, \tag{8}$$

$$\mathrm{m}(K)\Lambda_K \mathbf{v}_K = \sum_{\sigma\in\mathcal{E}_K} F_{K,\sigma}(\mathbf{x}_\sigma - \mathbf{x}_K), \quad \forall K \in \mathcal{M} \tag{9}$$

(where $\Lambda_K = \frac{1}{\mathrm{m}(K)}\int_K \Lambda(x)\,\mathrm{d}x$) and

$$-\sum_{\sigma\in\mathcal{E}_K} F_{K,\sigma} = \int_K f(x)\,\mathrm{d}x, \quad \forall K \in \mathcal{M}. \tag{10}$$

The origin of each of these equations is quite easy to understand. Since u and $\mathbf{v}$ stand for approximate values of $\bar{u}$ and $\nabla\bar{u}$, equation (7) simply states, if we assume $\nu_K = 0$, that $\mathbf{v}$ is a discrete gradient of u: it is the discrete counterpart of $u(\mathbf{x}_L) - u(\mathbf{x}_K) = u(\mathbf{x}_L) - u(\mathbf{x}_\sigma) +$

$u(\mathbf{x}_\sigma) - u(\mathbf{x}_K) \approx \nabla u(\mathbf{x}_L) \cdot (\mathbf{x}_L - \mathbf{x}_\sigma) + \nabla u(\mathbf{x}_K) \cdot (\mathbf{x}_\sigma - \mathbf{x}_K)$. This equation is slightly penalized with the fluxes to ensure existence and estimates on the said fluxes (to study the convergence of the scheme, we will assume $\nu_K > 0$; see the theorems below). Notice that the boundary condition $\bar{u} = 0$ is contained in the second line of (7). As $F_{K,\sigma}$ stands for an approximate value of $\int_\sigma \Lambda\nabla(x)\bar{u}(x) \cdot \mathbf{n}_{K,\sigma}\, d\gamma(x)$, it is natural to ask for the conservation property (8), and the balance (10) simply comes from the integration of (1) on a control volume. Last, the link (9) between $\Lambda\mathbf{v}$ and its fluxes is justified by Lemma 6.1 in the appendix, which shows that one can reconstruct a vector from its fluxes through the edges of a control volume.

Our main results on the mixed finite volume scheme are the following. The first one states that there exists a unique solution to the scheme. The second one gives the convergence of this solution to the solution of the continuous problem, as the size of the mesh tends to 0, and the third one provides an error estimate in the case of smooth data.

Theorem 2.1 *Let us assume Assumptions (2)-(4). Let $\mathcal{D}$ be an admissible discretization of Ω in the sense of Definition 2.1. Let $(\nu_K)_{K\in\mathcal{M}}$ be a family of positive real numbers. Then there exists one and only one $(u, \mathbf{v}, F)$ solution to ((7),(8),(9),(10)).*

Theorem 2.2 *Let us assume Assumptions (2)-(4). Let $(\mathcal{D}_m)_{m\geq 1}$ be admissible discretizations of Ω in the sense of Definition 2.1, such that $\mathrm{size}(\mathcal{D}_m) \to 0$ as $m \to \infty$ and $(\mathrm{regul}(\mathcal{D}_m))_{m\geq 1}$ is bounded. Let $\nu_0 > 0$ and $\beta \in (2-2d, 4-2d)$ be fixed. For all $m \geq 1$, let $(u_m, \mathbf{v}_m, F_m)$ be the solution to ((7),(8),(9),(10)) for the discretization $\mathcal{D}_m$, setting $\nu_K = \nu_0 \mathrm{diam}(K)^\beta$ for all $K \in \mathcal{M}_m$. Let $\bar{u}$ be the weak solution to (1).*
Then, as $m \to \infty$, $\mathbf{v}_m \to \nabla\bar{u}$ strongly in $L^2(\Omega)^d$ and $u_m \to \bar{u}$ weakly in $L^2(\Omega)$ and strongly in $L^q(\Omega)$ for all $q < 2$.

Theorem 2.3 *Let us assume Assumptions (2)-(4). Let $\mathcal{D}$ be an admissible discretization of Ω in the sense of Definition 2.1, such that $\mathrm{size}(\mathcal{D}) \leq 1$ and $\mathrm{regul}(\mathcal{D}) \leq \theta$ for some $\theta > 0$. We take $\nu_0 > 0$ and $\beta \in (2-2d, 4-2d)$ and, for all $K \in \mathcal{M}$, we let $\nu_K = \nu_0 \mathrm{diam}(K)^\beta$. Let $(u, \mathbf{v}, F)$ be the solution to ((7),(8),(9),(10)). Let $\bar{u}$ be the weak solution to (1). We assume that $\Lambda \in C^1(\overline{\Omega}; M_d(\mathbb{R}))$ and $\bar{u} \in C^2(\overline{\Omega})$.*
Then there exists C_1 only depending on d, Ω, $\bar{u}$, Λ, θ and ν_0 such that

$$\|\mathbf{v} - \nabla\bar{u}\|_{L^2(\Omega)^d} \leq C_1 \mathrm{size}(\mathcal{D})^{\frac{1}{2}\min(\beta+2d-2,4-2d-\beta)} \tag{11}$$

and

$$\|u - \bar{u}\|_{L^2(\Omega)} \leq C_1 \mathrm{size}(\mathcal{D})^{\frac{1}{2}\min(\beta+2d-2,4-2d-\beta)} \tag{12}$$

(note that the maximum value of $\frac{1}{2}\min(\beta + 2d - 2, 4 - 2d - \beta)$ is $\frac{1}{2}$, obtained for $\beta = 3 - 2d$).

Remark 2.3 *These error estimates are not sharp, and the numerical results in Section 5 show a much better order of convergence.*

3 The discretization space

We investigate here some properties of the space $L_\nu(\mathcal{D})$, which will be useful to study the mixed finite volume scheme. Recall that $L_\nu(\mathcal{D})$ is the space of $(u, \mathbf{v}, F)$ which satisfy (7).

Lemma 3.1 [Poincaré's Inequality] *Let us assume Assumption (2). Let $\mathcal{D}$ be an admissible discretization of Ω in the sense of Definition 2.1, such that* $\text{regul}(\mathcal{D}) \le \theta$ *for some $\theta > 0$. Let $(\nu_K)_{K\in\mathcal{M}}$ be a family of nonnegative real numbers. Then there exists C_2 only depending on d, Ω and θ such that, for all $(u, \mathbf{v}, F) \in L_\nu(\mathcal{D})$,*

$$\|u\|_{L^2(\Omega)} \le C_2 \left(\|\mathbf{v}\|_{L^2(\Omega)^d} + N_2(\mathcal{D}, \nu, F) \right), \tag{13}$$

where we have noted $N_2(\mathcal{D}, \nu, F) = \left(\sum_{K\in\mathcal{M}} \sum_{\sigma\in\mathcal{E}_K} \text{diam}(K)^{2d-2} \nu_K^2 F_{K,\sigma}^2 \text{m}(K) \right)^{1/2}$.

Proof.
Let $R > 0$ and $x_0 \in \Omega$ be such that $\Omega \subset B(x_0, R)$ (the open ball of center x_0 and radius R). We extend u by the value 0 in $B(x_0, R) \setminus \Omega$, and we consider $w \in H_0^1(B(x_0, R)) \cap H^2(B(x_0, R))$ such that $-\Delta w(x) = u(x)$ for a.e. $x \in B(x_0, R)$. We multiply each equation of (7) by $\int_\sigma \nabla w(x) \cdot \mathbf{n}_{K,\sigma}\, \mathrm{d}\gamma(x)$, and we sum on $\sigma \in \mathcal{E}$; since $\mathbf{n}_{K,\sigma} = -\mathbf{n}_{L,\sigma}$ whenever $\sigma = K|L$, we find

$$\begin{aligned}
&\sum_{\sigma\in\mathcal{E}_{\text{int}}, \sigma=K|L} \mathbf{v}_K \cdot (\mathbf{x}_\sigma - \mathbf{x}_K) \int_\sigma \nabla w(x) \cdot \mathbf{n}_{K,\sigma}\, \mathrm{d}\gamma(x) + \mathbf{v}_L \cdot (\mathbf{x}_\sigma - \mathbf{x}_L) \int_\sigma \nabla w(x) \cdot \mathbf{n}_{L,\sigma}\, \mathrm{d}\gamma(x) \\
&\quad + \sum_{\sigma\in\mathcal{E}_{\text{ext}}, \sigma\in\mathcal{E}_K} \mathbf{v}_K \cdot (\mathbf{x}_\sigma - \mathbf{x}_K) \int_\sigma \nabla w(x) \cdot \mathbf{n}_{K,\sigma}\, \mathrm{d}\gamma(x) \\
&\quad + \sum_{\sigma\in\mathcal{E}_{\text{int}}, \sigma=K|L} \nu_K \text{m}(K) F_{K,\sigma} \int_\sigma \nabla w(x) \cdot \mathbf{n}_{K,\sigma}\, \mathrm{d}\gamma(x) + \nu_L \text{m}(L) F_{L,\sigma} \int_\sigma \nabla w(x) \cdot \mathbf{n}_{L,\sigma}\, \mathrm{d}\gamma(x) \\
&\quad + \sum_{\sigma\in\mathcal{E}_{\text{ext}}, \sigma\in\mathcal{E}_K} \nu_K \text{m}(K) F_{K,\sigma} \int_\sigma \nabla w(x) \cdot \mathbf{n}_{K,\sigma}\, \mathrm{d}\gamma(x) \\
= &- \sum_{\sigma\in\mathcal{E}_{\text{int}}, \sigma=K|L} u_K \int_\sigma \nabla w(x) \cdot \mathbf{n}_{K,\sigma}\, \mathrm{d}\gamma(x) + u_L \int_\sigma \nabla w(x) \cdot \mathbf{n}_{L,\sigma}\, \mathrm{d}\gamma(x) \\
&\quad - \sum_{\sigma\in\mathcal{E}_{\text{ext}}, \sigma\in\mathcal{E}_K} u_K \int_\sigma \nabla w(x) \cdot \mathbf{n}_{K,\sigma}\, \mathrm{d}\gamma(x).
\end{aligned}$$

Gathering by control volumes, we find

$$\begin{aligned}
&\sum_{K\in\mathcal{M}} \mathbf{v}_K \cdot \sum_{\sigma\in\mathcal{E}_K} (\mathbf{x}_\sigma - \mathbf{x}_K) \int_\sigma \nabla w(x) \cdot \mathbf{n}_{K,\sigma}\, \mathrm{d}\gamma(x) \\
&+ \sum_{K\in\mathcal{M}} \sum_{\sigma\in\mathcal{E}_K} \nu_K \text{m}(K) F_{K,\sigma} \int_\sigma \nabla w(x) \cdot \mathbf{n}_{K,\sigma}\, \mathrm{d}\gamma(x) &&= - \sum_{K\in\mathcal{M}} u_K \sum_{\sigma\in\mathcal{E}_K} \int_\sigma \nabla w(x) \cdot \mathbf{n}_{K,\sigma}\, \mathrm{d}\gamma(x) \\
& &&= - \sum_{K\in\mathcal{M}} u_K \int_K \Delta w(x)\, dx \\
& &&= \sum_{K\in\mathcal{M}} \text{m}(K) u_K^2 = ||u||^2_{L^2(\Omega)}.
\end{aligned} \tag{14}$$

Let T_1 and T_2 be the two terms in the left-hand side of this equation.

Define $T_3 = \int_\Omega \mathbf{v}(x) \cdot \nabla w(x)\, \mathrm{d}x$; we have

$$|T_3| \le \|\mathbf{v}\|_{L^2(\Omega)^d} \|w\|_{H^1(\Omega)} \tag{15}$$

and we want to compare T_1 with T_3. In order to do so, we apply Lemma 6.1 in the appendix to the vector $\mathbf{G}_K = \frac{1}{\mathrm{m}(K)}\int_K \nabla w(x)\,\mathrm{d}x$, which gives

$$\int_K \nabla w(x)\,\mathrm{d}x = \mathrm{m}(K)\mathbf{G}_K = \sum_{\sigma\in\mathcal{E}_K} \mathrm{m}(\sigma)\mathbf{G}_K\cdot\mathbf{n}_{K,\sigma}(\mathbf{x}_\sigma - \mathbf{x}_K)$$

and therefore

$$T_3 = \sum_{K\in\mathcal{M}} \mathbf{v}_K \cdot \sum_{\sigma\in\mathcal{E}_K} \mathrm{m}(\sigma)\mathbf{G}_K\cdot\mathbf{n}_{K,\sigma}(\mathbf{x}_\sigma - \mathbf{x}_K).$$

Hence, setting $\mathbf{G}_\sigma = \frac{1}{\mathrm{m}(\sigma)}\int_\sigma \nabla w(x)\,\mathrm{d}\gamma(x)$, we get

$$|T_1 - T_3| \le \sum_{K\in\mathcal{M}} |\mathbf{v}_K| \sum_{\sigma\in\mathcal{E}_K} \mathrm{m}(\sigma)\ |\mathbf{G}_K - \mathbf{G}_\sigma|\ \mathrm{diam}(K).$$

Thanks to the Cauchy-Schwarz inequality, we find

$$(T_1 - T_3)^2 \le \left(\sum_{K\in\mathcal{M}} |\mathbf{v}_K|^2 \sum_{\sigma\in\mathcal{E}_K} \mathrm{m}(\sigma)\mathrm{diam}(K)\right)\left(\sum_{K\in\mathcal{M}}\sum_{\sigma\in\mathcal{E}_K} \mathrm{m}(\sigma)\mathrm{diam}(K)|\mathbf{G}_K - \mathbf{G}_\sigma|^2\right).$$

We now apply Lemma 6.3 in the appendix, which gives C_3 only depending on d and θ such that

$$|\mathbf{G}_K - \mathbf{G}_\sigma|^2 \le C_3 \frac{\mathrm{diam}(K)}{\mathrm{m}(\sigma)} \|w\|^2_{H^2(K)} \tag{16}$$

(notice that $\alpha := \frac{1}{2}\theta^{-1/d} < \mathrm{regul}(\mathcal{D})^{-1/d} \le \rho_K/\mathrm{diam}(K)$ is valid in Lemma 6.3). We also have, for $\sigma \in \mathcal{E}_K$, $\mathrm{m}(\sigma) \le \omega_{d-1}\mathrm{diam}(K)^{d-1}$, where ω_{d-1} is the volume of the unit ball in $\mathbb{R}^{d-1}$. Therefore, according to (6) and since $\mathrm{regul}(\mathcal{D}) \ge \mathrm{card}(\mathcal{E}_K)$ for all $K \in \mathcal{M}$,

$$\begin{aligned}
(T_1 - T_3)^2 &\le \left(\sum_{K\in\mathcal{M}} |\mathbf{v}_K|^2 \sum_{\sigma\in\mathcal{E}_K} \mathrm{m}(\sigma)\ \mathrm{diam}(K)\right)\left(\sum_{K\in\mathcal{M}}\sum_{\sigma\in\mathcal{E}_K} C_3\mathrm{diam}(K)^2 ||w||^2_{H^2(K)}\right)\\
&\le \left(\omega_{d-1}\mathrm{regul}(\mathcal{D})\sum_{K\in\mathcal{M}} |\mathbf{v}_K|^2\mathrm{diam}(K)^d\right)\left(C_3\mathrm{size}(\mathcal{D})^2\mathrm{regul}(\mathcal{D})\|w\|^2_{H^2(\Omega)}\right)\\
&\le \frac{\omega_{d-1}\mathrm{regul}(\mathcal{D})^2}{\omega_d}||\mathbf{v}||^2_{L^2(\Omega)^d} C_3\mathrm{diam}(\Omega)^2\mathrm{regul}(\mathcal{D})\|w\|^2_{H^2(\Omega)}.
\end{aligned} \tag{17}$$

Turning to T_2, we have $T_2 = \sum_{K\in\mathcal{M}}\sum_{\sigma\in\mathcal{E}_K} \nu_K\mathrm{m}(K)F_{K,\sigma}\mathrm{m}(\sigma)\mathbf{G}_\sigma\cdot\mathbf{n}_{K,\sigma}$, which we compare with $T_4 = \sum_{K\in\mathcal{M}}\sum_{\sigma\in\mathcal{E}_K} \nu_K\mathrm{m}(K)F_{K,\sigma}\mathrm{m}(\sigma)\mathbf{G}_K\cdot\mathbf{n}_{K,\sigma}$ thanks to (16):

$$\begin{aligned}
&(T_2 - T_4)^2\\
&\quad\le \left(\sum_{K\in\mathcal{M}}\sum_{\sigma\in\mathcal{E}_K} \mathrm{diam}(K)\mathrm{m}(\sigma)\nu_K^2 F_{K,\sigma}^2\mathrm{m}(K)^2\right)\left(\sum_{K\in\mathcal{M}}\sum_{\sigma\in\mathcal{E}_K} \frac{\mathrm{m}(\sigma)}{\mathrm{diam}(K)}|\mathbf{G}_K - \mathbf{G}_\sigma|^2\right)\\
&\quad\le \left(\omega_{d-1}\omega_d \sum_{K\in\mathcal{M}}\sum_{\sigma\in\mathcal{E}_K} \mathrm{diam}(K)^{2d}\nu_K^2 F_{K,\sigma}^2\mathrm{m}(K)\right)\mathrm{regul}(\mathcal{D})C_3||w||^2_{H^2(\Omega)}\\
&\quad\le \omega_{d-1}\omega_d\mathrm{diam}(\Omega)^2 N_2(\mathcal{D},\nu,F)^2\mathrm{regul}(\mathcal{D})C_3||w||^2_{H^2(\Omega)}.
\end{aligned} \tag{18}$$

On the other hand, we can write

$$\begin{aligned} T_4^2 &\leq \left(\sum_{K\in\mathcal{M}}\sum_{\sigma\in\mathcal{E}_K}\mathrm{m}(\sigma)^2\nu_K^2F_{K,\sigma}^2\mathrm{m}(K)\right)\left(\sum_{K\in\mathcal{M}}\sum_{\sigma\in\mathcal{E}_K}\mathrm{m}(K)|\mathbf{G}_K|^2\right)\\ &\leq \omega_{d-1}^2N_2(\mathcal{D},\nu,F)^2\left(\mathrm{regul}(\mathcal{D})\sum_{K\in\mathcal{M}}\mathrm{m}(K)|\mathbf{G}_K|^2\right)\\ &\leq \omega_{d-1}^2N_2(\mathcal{D},\nu,F)^2\mathrm{regul}(\mathcal{D})\|w\|_{H^1(\Omega)}^2. \end{aligned} \tag{19}$$

Thanks to (15), (17), (18) and (19), we can come back in (14) to find

$$\begin{aligned} ||u||_{L^2(\Omega)}^2 &= T_1+T_2\\ &\leq |T_1-T_3|+|T_3|+|T_2-T_4|+|T_4|\\ &\leq \sqrt{\frac{\omega_{d-1}C_3\theta^3}{\omega_d}}\mathrm{diam}(\Omega)\|\mathbf{v}\|_{L^2(\Omega)^d}||w||_{H^2(\Omega)}+||\mathbf{v}||_{L^2(\Omega)^d}||w||_{H^1(\Omega)}\\ &\quad+\sqrt{\omega_{d-1}\omega_dC_3\theta}\,\mathrm{diam}(\Omega)N_2(\mathcal{D},\nu,F)\|w\|_{H^2(\Omega)}+\omega_{d-1}\sqrt{\theta}\,N_2(\mathcal{D},\nu,F)\|w\|_{H^1(\Omega)}. \end{aligned}$$

Since there exists C_4 only depending on d and $B(x_0,R)$ (the ball chosen at the beginning of the proof) such that $||w||_{H^2(\Omega)}\leq C_4||u||_{L^2(\Omega)}$, this concludes the proof. □

Lemma 3.2 [Equicontinuity of the translations] *Let us assume Assumption (2). Let $\mathcal{D}$ be an admissible discretization of Ω in the sense of Definition 2.1, such that* $\mathrm{regul}(\mathcal{D})\leq\theta$ *for some $\theta>0$. Let $(\nu_K)_{K\in\mathcal{M}}$ be a family of nonnegative real numbers. Then there exists C_5 only depending on d, Ω and θ such that, for all $(u,\mathbf{v},F)\in L_\nu(\mathcal{D})$ and all $\xi\in\mathbb{R}^d$,*

$$\|u(\cdot+\xi)-u\|_{L^1(\mathbb{R}^d)}\leq C_5\left(\|\mathbf{v}\|_{L^1(\Omega)^d}+N_1(\mathcal{D},\nu,F)\right)|\xi|, \tag{20}$$

where $N_1(\mathcal{D},\nu,F)=\sum_{K\in\mathcal{M}}\sum_{\sigma\in\mathcal{E}_K}\mathrm{diam}(K)^{d-1}\nu_K|F_{K,\sigma}|\mathrm{m}(K)$ (and u has been extended by 0 outside Ω).

Proof.
For all $\sigma\in\mathcal{E}$, let us define $D_\sigma u=|u_L-u_K|$ if $\sigma=K|L$ and $D_\sigma u=|u_K|$ if $\sigma\in\mathcal{E}_K\cap\mathcal{E}_{\mathrm{ext}}$. For $(x,\xi)\in\mathbb{R}^d\times\mathbb{R}^d$ and $\sigma\in\mathcal{E}$, we define $\chi(x,\xi,\sigma)$ by 1 if $\sigma\cap[x,x+\xi]\neq\emptyset$ and by 0 otherwise. We then have, for all $\xi\in\mathbb{R}^d$ and a.e. $x\in\mathbb{R}^d$ (the x's such that x and $x+\xi$ do not belong to $\cup_{K\in\mathcal{M}}\partial K$, and $[x,x+\xi]$ does not intersect the relative boundary of any edge),

$$|u(x+\xi)-u(x)|\leq\sum_{\sigma\in\mathcal{E}}\chi(x,\xi,\sigma)D_\sigma u.$$

Applying (7), we get $|u(x+\xi)-u(x)|\leq T_5(x)+T_6(x)$ with

$$\begin{aligned} T_5(x) &= \sum_{\sigma\in\mathcal{E}_{\mathrm{int}},\sigma=K|L}\chi(x,\xi,\sigma)(|\mathbf{v}_K||\mathbf{x}_\sigma-\mathbf{x}_K|+|\mathbf{v}_L||\mathbf{x}_L-\mathbf{x}_\sigma|)\\ &\quad+\sum_{\sigma\in\mathcal{E}_{\mathrm{ext}},\sigma\in\mathcal{E}_K}\chi(x,\xi,\sigma)|\mathbf{v}_K||\mathbf{x}_\sigma-\mathbf{x}_K|\\ &\leq \sum_{K\in\mathcal{M}}\sum_{\sigma\in\mathcal{E}_K}\chi(x,\xi,\sigma)\mathrm{diam}(K)|\mathbf{v}_K| \end{aligned}$$

and

$$\begin{aligned} T_6(x) &= \sum_{\sigma\in\mathcal{E}_{\rm int},\sigma=K|L} \chi(x,\xi,\sigma)\left(\nu_K {\rm m}(K)|F_{K,\sigma}| + \nu_L {\rm m}(L)|F_{L,\sigma}|\right) \\ &\quad + \sum_{\sigma\in\mathcal{E}_{\rm ext},\sigma\in\mathcal{E}_K} \chi(x,\xi,\sigma)\nu_K {\rm m}(K)|F_{K,\sigma}| \\ &= \sum_{K\in\mathcal{M}}\sum_{\sigma\in\mathcal{E}_K} \chi(x,\xi,\sigma)\nu_K {\rm m}(K)|F_{K,\sigma}|. \end{aligned}$$

In order that $\chi(x,\xi,\sigma)\neq 0$, x must lie in the set $\sigma - [0,1]\xi$ which has measure ${\rm m}(\sigma)|\mathbf{n}_\sigma\cdot\xi|$ (where $\mathbf{n}_\sigma$ is a unit normal to σ). Hence,

$$\int_{\mathbb{R}^d} \chi(x,\xi,\sigma)\,dx \le {\rm m}(\sigma)|\mathbf{n}_\sigma\cdot\xi| \le \omega_{d-1}{\rm diam}(K)^{d-1}|\xi| \qquad \text{if } \sigma\in\mathcal{E}_K.$$

Since ${\rm Card}(\mathcal{E}_K)\le {\rm regul}(\mathcal{D})$, this gives

$$\int_{\mathbb{R}^d} T_5(x)\,dx \le \omega_{d-1}{\rm regul}(\mathcal{D})|\xi| \sum_{K\in\mathcal{M}} {\rm diam}(K)^d|\mathbf{v}_K|$$

and

$$\int_{\mathbb{R}^d} T_6(x)\,{\rm d}x \le \omega_{d-1}|\xi| \sum_{K\in\mathcal{M}}\sum_{\sigma\in\mathcal{E}_K} {\rm diam}(K)^{d-1}\nu_K|F_{K,\sigma}|{\rm m}(K),$$

which concludes the proof thanks to (6). □

Remark 3.1 *We could prove that* $\|u(\cdot+\xi)-u\|^2_{L^2(\mathbb{R}^d)} \le C(\|\mathbf{v}\|^2_{L^2(\Omega)^d} + N_2(\mathcal{D},\nu,F)^2)\ |\xi|\ (|\xi| + {\rm size}(\mathcal{D}))$ *by assuming that* $\max\{{\rm diam}(K)/\rho_L\ ,\ (K,L)\in\mathcal{M}\times\mathcal{M}\}$ *remains bounded. This would give in Theorem 2.2 the strong convergence of u_m in $L^2(\Omega)$, but this would also prevent from considering locally refined mesh, so we prefer not to add this assumption. Notice however that Theorem 2.3 states a strong convergence in $L^2(\Omega)$ of the approximate solution u.*

Lemma 3.3 [Compactness property] *Let us assume Assumption (2). Let $(\mathcal{D}_m)_{m\ge1}$ be admissible discretizations of Ω in the sense of Definition 2.1, such that* ${\rm size}(\mathcal{D}_m)\to 0$ *as $m\to\infty$ and* $({\rm regul}(\mathcal{D}_m))_{m\ge1}$ *is bounded. Let $(u_m,\mathbf{v}_m,F_m,\nu_m)_{m\ge1}$ be such that $(u_m,\mathbf{v}_m,F_m)\in L_{\nu_m}(\mathcal{D}_m)$, $(\mathbf{v}_m)_{m\ge1}$ is bounded in $L^2(\Omega)^d$ and $N_2(\mathcal{D}_m,\nu_m,F_m)\to 0$ as $m\to\infty$ (N_2 has been defined in Lemma 3.1).*
Then there exists a subsequence of $(\mathcal{D}_m)_{m\ge1}$ (still denoted by $(\mathcal{D}_m)_{m\ge1}$) and $\bar u\in H^1_0(\Omega)$ such that the corresponding sequence $(u_m)_{m\ge1}$ converges to $\bar u$ weakly in $L^2(\Omega)$ and strongly in $L^q(\Omega)$ for all $q<2$, and such that $(\mathbf{v}_m)_{m\ge1}$ converges to $\nabla\bar u$ weakly in $L^2(\Omega)^d$.

Proof.
Notice first that, for all discretization $\mathcal{D}$, for all $\nu=(\nu_K)_{K\in\mathcal{M}}$ nonnegative numbers and for all $F=(F_{K,\sigma})_{K\in\mathcal{M},\,\sigma\in\mathcal{E}_K}$,

$$\begin{aligned} N_1(\mathcal{D},\nu,F) &= \sum_{K\in\mathcal{M}}\sum_{\sigma\in\mathcal{E}_K} {\rm diam}(K)^{d-1}\nu_K|F_{K,\sigma}|{\rm m}(K) \\ &\le \left(\sum_{K\in\mathcal{M}}\sum_{\sigma\in\mathcal{E}_K} {\rm diam}(K)^{2d-2}\nu_K^2F_{K,\sigma}^2{\rm m}(K)\right)^{1/2}\left(\sum_{K\in\mathcal{M}}\sum_{\sigma\in\mathcal{E}_K}{\rm m}(K)\right)^{1/2} \\ &\le N_2(\mathcal{D},\nu,F){\rm regul}(\mathcal{D})^{1/2}{\rm m}(\Omega)^{1/2}. \end{aligned} \tag{21}$$

Hence, if $N_2(\mathcal{D},\nu,F)$ and $\mathrm{regul}(\mathcal{D})$ are bounded, so is $N_1(\mathcal{D},\nu,F)$. Owing to this, the hypotheses, Lemmas 3.1, 3.2 and Kolmogorov's compactness theorem allow to extract a subsequence such that $\mathbf{v}_m \to \bar{\mathbf{v}}$ weakly in $L^2(\Omega)^d$ and $u_m \to \bar{u}$ weakly in $L^2(\Omega)$ and strongly in $L^1(\Omega)$ (which implies the strong convergence in $L^q(\Omega)$ for all $q<2$). We now extend u_m, $\bar{u}$, $\mathbf{v}_m$ and $\bar{\mathbf{v}}$ by 0 outside Ω and we prove that $\bar{\mathbf{v}} = \nabla\bar{u}$ in the distributional sense on $\mathbb{R}^d$. This will conclude that $\bar{u}\in H^1(\mathbb{R}^d)$ and, since $\bar{u}=0$ outside Ω, that $\bar{u}\in H^1_0(\Omega)$.

Let $\mathbf{e}\in\mathbb{R}^d$ and $\varphi\in C^\infty_c(\mathbb{R}^d)$. For simplicity, we drop the index m for $\mathcal{D}_m$, $\mathbf{v}_m$ and u_m. We multiply each equation of (7) by $\int_\sigma \varphi(x)\,\mathrm{d}\gamma(x)\mathbf{e}\cdot\mathbf{n}_{K,\sigma}$, we sum all these equations and we gather by control volumes, getting $T_7+T_8=T_9$ with

$$T_7 = \sum_{K\in\mathcal{M}} \mathbf{v}_K\cdot\sum_{\sigma\in\mathcal{E}_K}\int_\sigma \varphi(x)\,\mathrm{d}\gamma(x)\ \mathbf{e}\cdot\mathbf{n}_{K,\sigma}(\mathbf{x}_\sigma-\mathbf{x}_K),$$

$$T_8 = \sum_{K\in\mathcal{M}}\sum_{\sigma\in\mathcal{E}_K}\nu_K\mathrm{m}(K)F_{K,\sigma}\int_\sigma\varphi(x)\,\mathrm{d}\gamma(x)\mathbf{e}\cdot\mathbf{n}_{K,\sigma}$$

and

$$T_9 = -\sum_{K\in\mathcal{M}} u_K\sum_{\sigma\in\mathcal{E}_K}\int_\sigma\varphi(x)\,\mathrm{d}\gamma(x)\ \mathbf{e}\cdot\mathbf{n}_{K,\sigma} = -\int_\Omega u(x)\mathrm{div}(\varphi(x)\mathbf{e})\,\mathrm{d}x.$$

We want to compare T_7 with T_{10} defined by

$$T_{10} = \sum_{K\in\mathcal{M}}\mathbf{v}_K\cdot\sum_{\sigma\in\mathcal{E}_K}\frac{1}{\mathrm{m}(K)}\int_K\varphi(x)\,\mathrm{d}x\ \mathrm{m}(\sigma)\ \mathbf{e}\cdot\mathbf{n}_{K,\sigma}(\mathbf{x}_\sigma-\mathbf{x}_K).$$

Since there exists C_6 only depending on φ such that, for all $\sigma\in\mathcal{E}_K$,

$$\left|\frac{1}{\mathrm{m}(\sigma)}\int_\sigma\varphi(x)\,\mathrm{d}\gamma(x)-\frac{1}{\mathrm{m}(K)}\int_K\varphi(x)\,\mathrm{d}x\right|\le C_6\mathrm{size}(\mathcal{D}),$$

we get that

$$|T_7-T_{10}|\le C_6|\mathbf{e}|\mathrm{size}(\mathcal{D})\sum_{K\in\mathcal{M}}|\mathbf{v}_K|\sum_{\sigma\in\mathcal{E}_K}\mathrm{m}(\sigma)|\mathbf{x}_\sigma-\mathbf{x}_K|.$$

But $\mathrm{m}(\sigma)|\mathbf{x}_\sigma-\mathbf{x}_K|\le\omega_{d-1}\mathrm{diam}(K)^d\le\frac{\omega_{d-1}\mathrm{regul}(\mathcal{D})}{\omega_d}\mathrm{m}(K)$ and, since $\mathrm{card}(\mathcal{E}_K)\le\mathrm{regul}(\mathcal{D})$, we obtain

$$|T_7-T_{10}|\le C_6|\mathbf{e}|\mathrm{size}(\mathcal{D})\frac{\omega_{d-1}\mathrm{regul}(\mathcal{D})^2}{\omega_d}\|\mathbf{v}\|_{L^1(\Omega)}$$

and thus $\lim_{\mathrm{size}(\mathcal{D})\to0}|T_7-T_{10}|=0$. Moreover, thanks to Lemma 6.1, we get $T_{10}=\int_\Omega\varphi(x)\mathbf{v}(x)\cdot\mathbf{e}\,\mathrm{d}x$ and so $\lim_{\mathrm{size}(\mathcal{D})\to0}T_{10}=\int_\Omega\varphi(x)\bar{\mathbf{v}}(x)\cdot\mathbf{e}\,\mathrm{d}x=\int_{\mathbb{R}^d}\varphi(x)\bar{\mathbf{v}}(x)\cdot\mathbf{e}\,\mathrm{d}x$ ($\bar{\mathbf{v}}$ has been extended by 0 outside Ω). This proves that

$$\lim_{\mathrm{size}(\mathcal{D})\to0}T_7=\int_{\mathbb{R}^d}\varphi(x)\bar{\mathbf{v}}(x)\cdot\mathbf{e}\,\mathrm{d}x. \tag{22}$$

Since φ is bounded, by (21) we find C_7 only depending on φ and $\mathbf{e}$ such that

$$\begin{aligned}|T_8| &\le C_7\sum_{K\in\mathcal{M}}\sum_{\sigma\in\mathcal{E}_K}\mathrm{m}(\sigma)\nu_K|F_{K,\sigma}|\mathrm{m}(K)\\ &\le C_7\omega_{d-1}N_1(\mathcal{D},\nu,F)\\ &\le C_7\omega_{d-1}\mathrm{regul}(\mathcal{D})^{1/2}\mathrm{m}(\Omega)^{1/2}N_2(\mathcal{D},\nu,F)\end{aligned}$$

and therefore, by the assumptions,

$$\lim_{\mathrm{size}(\mathcal{D})\to 0} T_8 = 0. \tag{23}$$

We clearly have

$$\lim_{\mathrm{size}(\mathcal{D})\to 0} T_9 = -\int_\Omega \bar{u}(x)\mathrm{div}(\varphi(x)\mathbf{e})\,\mathrm{d}x = -\int_{\mathbb{R}^d} \bar{u}(x)\mathrm{div}(\varphi(x)\mathbf{e})\,\mathrm{d}x$$

(recall that $\bar{u}$ has been extended by 0 outside Ω). Gathering this limit with (22) and (23) in $T_7 + T_8 = T_9$, we obtain

$$\int_{\mathbb{R}^d} \varphi(x)\bar{\mathbf{v}}(x)\cdot\mathbf{e}\,dx = -\int_{\mathbb{R}^d} \bar{u}(x)\mathrm{div}(\varphi(x)\mathbf{e})\,\mathrm{d}x,$$

which concludes the proof that $\bar{\mathbf{v}} = \nabla\bar{u}$ in the distributional sense on $\mathbb{R}^d$. □

4 Study of the mixed finite volume scheme

We first prove an *a priori* estimate on the solution to the scheme. This estimate shows in particular that, if $f = 0$, then $F = 0$ and $\mathbf{v} = 0$, and thus $u = 0$ by Lemma 3.1; since ((7),(8),(9),(10)) is square and linear in $(u, \mathbf{v}, F)$, the existence and uniqueness of the solution to the mixed finite volume scheme (i.e. Theorem 2.1) is an immediate consequence of this lemma.

Lemma 4.1 *Let us assume Assumptions (2)-(4). Let $\mathcal{D}$ be an admissible discretization of Ω in the sense of Definition 2.1. Let $(\nu_K)_{K\in\mathcal{M}}$ be a family of positive real numbers and $(u, \mathbf{v}, F)$ be a solution of ((7),(8),(9),(10)). Then, for all $\nu_0 > 0$ and all $\beta_0 \geq \beta \geq 2-2d$ such that $\nu_K \leq \nu_0 \mathrm{diam}(K)^\beta$ ($\forall K \in \mathcal{M}$), and for all $\theta \geq \mathrm{regul}(\mathcal{D})$, this solution satisfies*

$$\|\mathbf{v}\|^2_{L^2(\Omega)^d} + \sum_{K\in\mathcal{M}}\sum_{\sigma\in\mathcal{E}_K} \nu_K F^2_{K,\sigma}\mathrm{m}(K) \leq C_8 ||f||^2_{L^2(\Omega)} \tag{24}$$

where C_8 only depends on d, Ω, α_0, θ, ν_0 and β_0.

Proof.
Multiply (10) by u_K, sum on the control volumes and gather by edges using (8):

$$\sum_{\sigma\in\mathcal{E}_{\mathrm{int}},\sigma=K|L} F_{K,\sigma}(u_L - u_K) + \sum_{\sigma\in\mathcal{E}_{\mathrm{ext}},\sigma\in\mathcal{E}_K} -F_{K,\sigma}u_K = \int_\Omega f(x)u(x)\,dx.$$

Using (7) and (8), and gathering by control volumes, this gives

$$\begin{aligned}
&\int_\Omega f(x)u(x)\,dx \\
&\quad= \sum_{\sigma\in\mathcal{E}_{\mathrm{int}},\sigma=K|L} F_{K,\sigma}\mathbf{v}_K\cdot(\mathbf{x}_\sigma - \mathbf{x}_K) + F_{L,\sigma}\mathbf{v}_L\cdot(\mathbf{x}_\sigma - \mathbf{x}_L) + \sum_{\sigma\in\mathcal{E}_{\mathrm{ext}},\sigma\in\mathcal{E}_K} F_{K,\sigma}\mathbf{v}_K\cdot(\mathbf{x}_\sigma - \mathbf{x}_K) \\
&\qquad+ \sum_{\sigma\in\mathcal{E}_{\mathrm{int}},\sigma=K|L} \nu_K\mathrm{m}(K)F^2_{K,\sigma} + \nu_L\mathrm{m}(L)F^2_{L,\sigma} + \sum_{\sigma\in\mathcal{E}_{\mathrm{ext}},\sigma\in\mathcal{E}_K} \nu_K\mathrm{m}(K)F^2_{K,\sigma} \\
&\quad= \sum_{K\in\mathcal{M}} \mathbf{v}_K\cdot\sum_{\sigma\in\mathcal{E}_K} F_{K,\sigma}(\mathbf{x}_\sigma - \mathbf{x}_K) + \sum_{K\in\mathcal{M}}\sum_{\sigma\in\mathcal{E}_K} \nu_K\mathrm{m}(K)F^2_{K,\sigma}.
\end{aligned}$$

Applying (9), we obtain

$$\begin{aligned}\int_\Omega \mathbf{v}(x)\cdot\Lambda(x)\mathbf{v}(x)\,\mathrm{d}x + \sum_{K\in\mathcal{M}}\sum_{\sigma\in\mathcal{E}_K}\nu_K F_{K,\sigma}^2\mathrm{m}(K) &= \int_\Omega f(x)u(x)\,\mathrm{d}x \\ &\le ||f||_{L^2(\Omega)}||u||_{L^2(\Omega)}.\end{aligned}\tag{25}$$

Using Young's inequality and Lemma 3.1, we deduce that, for all $\varepsilon>0$,

$$\begin{aligned}\alpha_0||\mathbf{v}||^2_{L^2(\Omega)^d} + \sum_{K\in\mathcal{M}}\sum_{\sigma\in\mathcal{E}_K}\nu_K F_{K,\sigma}^2\mathrm{m}(K) \le \frac{1}{2\varepsilon}||f||^2_{L^2(\Omega)} + \varepsilon C_2^2||\mathbf{v}||^2_{L^2(\Omega)^d} \\ +\varepsilon C_2^2\sum_{K\in\mathcal{M}}\sum_{\sigma\in\mathcal{E}_K}\mathrm{diam}(K)^{2d-2}\nu_K^2F_{K,\sigma}^2\mathrm{m}(K).\end{aligned}\tag{26}$$

Since $\nu_K\le\nu_0\mathrm{diam}(K)^\beta$, we have $\nu_K\mathrm{diam}(K)^{2d-2}\le\nu_0\mathrm{diam}(K)^{\beta+2d-2}\le\nu_0\mathrm{diam}(\Omega)^{\beta+2d-2}\le\nu_0\sup(1,\mathrm{diam}(\Omega)^{\beta_0+2d-2})$ (recall that $\beta+2d-2\ge0$). Hence, (26) gives

$$\begin{aligned}\alpha_0||\mathbf{v}||^2_{L^2(\Omega)^d} + \sum_{K\in\mathcal{M}}\sum_{\sigma\in\mathcal{E}_K}\nu_K F_{K,\sigma}^2\mathrm{m}(K) \le \frac{1}{2\varepsilon}||f||^2_{L^2(\Omega)} + \varepsilon C_2^2||\mathbf{v}||^2_{L^2(\Omega)^d} \\ +\varepsilon\nu_0\sup(1,\mathrm{diam}(\Omega)^{\beta_0+2d-2})C_2^2\sum_{K\in\mathcal{M}}\sum_{\sigma\in\mathcal{E}_K}\nu_K F_{K,\sigma}^2\mathrm{m}(K).\end{aligned}$$

Taking $\varepsilon=\min(\frac{\alpha_0}{2C_2^2},\frac{1}{2\nu_0\sup(1,\mathrm{diam}(\Omega)^{\beta_0+2d-2})C_2^2})$ concludes the proof of the lemma. □

We now prove the convergence of the approximate solution toward the weak solution of (1).
Proof of Theorem 2.2.
For the simplicity of the notations, we omit the index m as in the proof of Lemma 3.3. We first note that, thanks to Estimate (24) and since $\nu_K=\nu_0\mathrm{diam}(K)^\beta$,

$$\begin{aligned}N_2(\mathcal{D},\nu,F)^2 &= \sum_{K\in\mathcal{M}}\sum_{\sigma\in\mathcal{E}_K}\mathrm{diam}(K)^{2d-2}\nu_K^2F_{K,\sigma}^2\mathrm{m}(K) \\ &= \nu_0\sum_{K\in\mathcal{M}}\sum_{\sigma\in\mathcal{E}_K}\mathrm{diam}(K)^{\beta+2d-2}\nu_KF_{K,\sigma}^2\mathrm{m}(K) \\ &\le \nu_0\mathrm{size}(\mathcal{D})^{\beta+2d-2}C_9\end{aligned}$$

where C_9 does not depend on the discretization $\mathcal{D}$ (recall that regul($\mathcal{D}$) is bounded). Since $\beta+2d-2>0$, this last quantity tends to 0, and so does $N_2(\mathcal{D},\nu,F)$, as size($\mathcal{D}$) $\to 0$. Hence, still using (24), we see that the assumptions of Lemma 3.3 are satisfied; there exists thus $\bar{u}\in H_0^1(\Omega)$ such that, up to a subsequence and as size($\mathcal{D}$) $\to 0$, $\mathbf{v}\to\nabla\bar{u}$ weakly in $L^2(\Omega)^d$ and $u\to\bar{u}$ weakly in $L^2(\Omega)$ and strongly in $L^q(\Omega)$ for $q<2$.
We now prove that the limit function $\bar{u}$ is the weak solution to (1). Since any subsequence of $(u,\mathbf{v})$ has a subsequence which converges as above, and since the reasoning we are going to make proves that any such limit of a subsequence is the (unique) weak solution to (1), this will conclude the proof, except for the strong convergence of $\mathbf{v}$.

Let $\varphi\in C_c^\infty(\Omega)$. We multiply (10) by $\varphi(\mathbf{x}_K)$ and we sum on K. Gathering by edges thanks to (8), we get

$$\sum_{\sigma\in\mathcal{E}_{\rm int},\sigma=K|L}F_{K,\sigma}(\varphi(\mathbf{x}_L)-\varphi(\mathbf{x}_K)) = \sum_{K\in\mathcal{M}}\int_K\varphi(\mathbf{x}_K)f(x)\,\mathrm{d}x$$

as long as size($\mathcal{D}$) is small enough (so that $\varphi = 0$ on the control volumes K such that $\partial K \cap \partial\Omega \neq \emptyset$). We set, for $\sigma = K|L$,

$$\varphi(\mathbf{x}_L) - \varphi(\mathbf{x}_K) = \frac{1}{\mathrm{m}(K)}\int_K \nabla\varphi(x)\,\mathrm{d}x \cdot (\mathbf{x}_\sigma - \mathbf{x}_K) + \frac{1}{\mathrm{m}(L)}\int_L \nabla\varphi(x)\,\mathrm{d}x \cdot (\mathbf{x}_L - \mathbf{x}_\sigma) + R_{KL}$$

and we have $|R_{KL}| \leq C_\varphi(\mathrm{diam}(K)^2 + \mathrm{diam}(L)^2)$. We then obtain, gathering by control volumes and using (9) (and the fact that $\varphi = 0$ on the control volumes on the boundary of Ω),

$$\int_\Omega \Lambda_{\mathcal{D}}\mathbf{v}(x) \cdot \nabla\varphi(x)\,\mathrm{d}x = \int_\Omega f(x)\varphi_{\mathcal{D}}(x)\,\mathrm{d}x + T_{11}, \tag{27}$$

where $\Lambda_{\mathcal{D}}$ and $\varphi_{\mathcal{D}}$ are constant respectively equal to Λ_K and $\varphi(\mathbf{x}_K)$ on each mesh K, and

$$|T_{11}| \leq C_\varphi \sum_{\sigma\in\mathcal{E}_{\mathrm{int}},\sigma=K|L} |F_{K,\sigma}|(\mathrm{diam}(K)^2 + \mathrm{diam}(L)^2) \leq C_\varphi \sum_{K\in\mathcal{M}}\sum_{\sigma\in\mathcal{E}_K} \mathrm{diam}(K)^2|F_{K,\sigma}|.$$

Let us estimate this term. We have

$$\begin{aligned}|T_{11}|^2 &\leq C_\varphi^2 \left(\sum_{K\in\mathcal{M}}\sum_{\sigma\in\mathcal{E}_K} \nu_K F_{K,\sigma}^2 \mathrm{m}(K)\right)\left(\sum_{K\in\mathcal{M}}\sum_{\sigma\in\mathcal{E}_K} \frac{\mathrm{diam}(K)^4}{\nu_K \mathrm{m}(K)}\right)\\ &\leq C_{10} \sum_{K\in\mathcal{M}}\sum_{\sigma\in\mathcal{E}_K} \frac{\mathrm{diam}(K)^4}{\nu_K \mathrm{m}(K)^2}\mathrm{m}(K) \end{aligned}\tag{28}$$

where, according to (24), C_{10} does not depend on the mesh since regul($\mathcal{D}$) stays bounded. But $\nu_K = \nu_0 \mathrm{diam}(K)^\beta$ and $\mathrm{diam}(K)^d \leq \frac{\mathrm{regul}(\mathcal{D})}{\omega_d}\mathrm{m}(K)$, so that

$$\frac{\mathrm{diam}(K)^4}{\nu_K \mathrm{m}(K)^2} \leq \frac{\mathrm{regul}(\mathcal{D})^2 \mathrm{diam}(K)^{4-\beta}}{\omega_d^2 \nu_0 \mathrm{diam}(K)^{2d}} = \frac{\mathrm{regul}(\mathcal{D})^2}{\omega_d^2\nu_0}\mathrm{diam}(K)^{4-2d-\beta}.$$

Since $4 - 2d - \beta > 0$, we deduce from (28) that

$$|T_{11}|^2 \leq C_{10}\frac{\mathrm{regul}(\mathcal{D})^2}{\omega_d^2\nu_0}\mathrm{size}(\mathcal{D})^{4-2d-\beta} \sum_{K\in\mathcal{M}}\sum_{\sigma\in\mathcal{E}_K} \mathrm{m}(K) \leq \frac{C_{10}\mathrm{regul}(\mathcal{D})^3\mathrm{m}(\Omega)}{\omega_d^2\nu_0}\mathrm{size}(\mathcal{D})^{4-2d-\beta}$$

and this quantity tends to 0 as size($\mathcal{D}$) → 0. Hence, we can pass to the limit in (27) to see that

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

which proves that $\bar{u}$ is the weak solution to (1).

The strong convergence of $\mathbf{v}$ to $\nabla\bar{u}$ is a consequence of (25). From this equation, and defining $N(\mathbf{w})^2 = \int_\Omega \Lambda(x)\mathbf{w}(x)\cdot\mathbf{w}(x)\,\mathrm{d}x$, we have $N(\mathbf{v})^2 \leq \int_\Omega f(x)u(x)\,\mathrm{d}x$ and thus

$$\limsup_{\mathrm{size}(\mathcal{D})\to 0} N(\mathbf{v})^2 \leq \lim_{\mathrm{size}(\mathcal{D})\to 0}\int_\Omega f(x)u(x)\,\mathrm{d}x = \int_\Omega f(x)\bar{u}(x)\,\mathrm{d}x = N(\nabla\bar{u})^2 \tag{29}$$

(we use the fact that $u \to \bar{u}$ weakly in $L^2(\Omega)$ and that $\bar{u}$ is the weak solution to (1)). But N is a norm on $L^2(\Omega)^d$, equivalent to the usual norm and coming from the scalar product $\langle \mathbf{w}, \mathbf{z}\rangle = \int_\Omega \frac{\Lambda(x)+\Lambda(x)^T}{2}\mathbf{w}(x)\cdot\mathbf{z}(x)\,\mathrm{d}x$; since $\mathbf{v} \to \nabla\bar{u}$ weakly in $L^2(\Omega)^d$ as size($\mathcal{D}$) → 0, we therefore also have $N(\nabla\bar{u}) \leq \liminf_{\mathrm{size}(\mathcal{D})\to 0} N(\mathbf{v})$. We conclude with (29) that $N(\mathbf{v}) \to N(\nabla\bar{u})$ as size($\mathcal{D}$) → 0, and thus that the weak convergence of $\mathbf{v}$ to $\nabla\bar{u}$ in $L^2(\Omega)^d$ is in fact strong. □

Remark 4.1 *As a consequence of (25) and the strong convergence of* $\mathbf{v}$ *to* $\nabla\bar{u}$*, we see that* $\sum_{K\in\mathcal{M}}\sum_{\sigma\in\mathcal{E}_K}\nu_K F_{K,\sigma}^2\mathrm{m}(K)\to 0$ *as* $\mathrm{size}(\mathcal{D})\to 0$*. This strengthens Lemma 4.1 which only states that this quantity is bounded.*

To conclude this section, we prove the error estimates. Note that these estimates could be extended, for $d\le 3$, to the case $\bar{u}\in H^2(\Omega)$ following some arguments of [16].

PROOF OF THEOREM 2.3.

In this proof, we denote by C_i (for all integer i) various real numbers which can depend on d, Ω, $\bar{u}$, Λ and θ, but not on $\mathrm{size}(\mathcal{D})$. We also denote, for all $K\in\mathcal{M}$ and $\sigma\in\mathcal{E}_K$, $\bar{u}_K=\bar{u}(\mathbf{x}_K)$, $\bar{u}_\sigma=\bar{u}(\mathbf{x}_\sigma)$,

$$\bar{F}_{K,\sigma}=\int_\sigma \Lambda(x)\nabla\bar{u}(x)\cdot\mathbf{n}_{K,\sigma}\,\mathrm{d}\gamma(x),$$

$$\bar{\mathbf{v}}_K=\frac{1}{\mathrm{m}(K)}\Lambda_K^{-1}\sum_{\sigma\in\mathcal{E}_K}\bar{F}_{K,\sigma}(\mathbf{x}_\sigma-\mathbf{x}_K)$$

(notice that Λ_K is indeed invertible since, from (3), $\Lambda_K\ge\alpha_0\mathrm{Id}$). Thanks to Lemma 6.1, we have

$$|\bar{\mathbf{v}}_K-\nabla\bar{u}(x)|\le C_{11}\mathrm{diam}(K),\quad \forall x\in K\,,\ \forall K\in\mathcal{M},\tag{30}$$

which implies

$$\bar{\mathbf{v}}_K\cdot(\mathbf{x}_\sigma-\mathbf{x}_K)=\bar{u}_\sigma-\bar{u}_K+R_{K,\sigma},\quad \forall K\in\mathcal{M},\ \forall\sigma\in\mathcal{E}_K,$$

with $|R_{K,\sigma}|\le C_{12}\mathrm{diam}(K)^2$ for all $K\in\mathcal{M}$ and $\sigma\in\mathcal{E}_K$. Since $\bar{u}$ is a classical solution to (1), we have

$$-\sum_{\sigma\in\mathcal{E}_K}\bar{F}_{K,\sigma}=\int_K f(x)\,\mathrm{d}x,\quad \forall K\in\mathcal{M}.$$

Denoting, for all $K\in\mathcal{M}$ and all $\sigma\in\mathcal{E}_K$, $\widehat{u}_K=u_K-\bar{u}_K$, $\widehat{\mathbf{v}}_K=\mathbf{v}_K-\bar{\mathbf{v}}_K$ and $\widehat{F}_{K,\sigma}=F_{K,\sigma}-\bar{F}_{K,\sigma}$, we see that

$$-\sum_{\sigma\in\mathcal{E}_K}\widehat{F}_{K,\sigma}=0,\quad \forall K\in\mathcal{M},\tag{31}$$

$$\widehat{F}_{K,\sigma}+\widehat{F}_{L,\sigma}=0,\quad \forall\sigma=K|L\in\mathcal{E}_{\mathrm{int}},\tag{32}$$

$$\mathrm{m}(K)\Lambda_K\widehat{\mathbf{v}}_K=\sum_{\sigma\in\mathcal{E}_K}\widehat{F}_{K,\sigma}(\mathbf{x}_\sigma-\mathbf{x}_K),\quad \forall K\in\mathcal{M},\tag{33}$$

$$\begin{aligned}
&\widehat{\mathbf{v}}_K\cdot(\mathbf{x}_\sigma-\mathbf{x}_K)+\widehat{\mathbf{v}}_L\cdot(\mathbf{x}_L-\mathbf{x}_\sigma)+\nu_K\mathrm{m}(K)\widehat{F}_{K,\sigma}+\nu_K\mathrm{m}(K)\bar{F}_{K,\sigma}+R_{K,\sigma}\\
&\qquad\qquad -\nu_L\mathrm{m}(L)\widehat{F}_{L,\sigma}-\nu_L\mathrm{m}(L)\bar{F}_{L,\sigma}-R_{L,\sigma}=\widehat{u}_L-\widehat{u}_K,\\
&\qquad\qquad\qquad \forall K\in\mathcal{M},\ \forall L\in\mathcal{N}_K,\ \text{with } \sigma=K|L,
\end{aligned}\tag{34}$$

$$\begin{aligned}
&\widehat{\mathbf{v}}_K\cdot(\mathbf{x}_\sigma-\mathbf{x}_K)+\nu_K\mathrm{m}(K)\widehat{F}_{K,\sigma}+\nu_K\mathrm{m}(K)\bar{F}_{K,\sigma}+R_{K,\sigma}=-\widehat{u}_K,\\
&\qquad\qquad\qquad \forall K\in\mathcal{M},\ \forall\sigma\in\mathcal{E}_K\cap\mathcal{E}_{\mathrm{ext}}.
\end{aligned}$$

We then get, multiplying (31) by $\widehat{u}_K$, (34) by $\widehat{F}_{K,\sigma}$ and using (32), (33),

$$\sum_{K\in\mathcal{M}}\mathrm{m}(K)\Lambda_K\widehat{\mathbf{v}}_K\cdot\widehat{\mathbf{v}}_K+\sum_{K\in\mathcal{M}}\sum_{\sigma\in\mathcal{E}_K}\nu_K\mathrm{m}(K)\widehat{F}_{K,\sigma}^2=T_{12}+T_{13},\tag{35}$$

where

$$T_{12} = -\sum_{K\in\mathcal{M}}\sum_{\sigma\in\mathcal{E}_K} \nu_K \mathrm{m}(K)\bar{F}_{K,\sigma}\widehat{F}_{K,\sigma},$$

$$T_{13} = -\sum_{K\in\mathcal{M}}\sum_{\sigma\in\mathcal{E}_K} R_{K,\sigma}\widehat{F}_{K,\sigma}.$$

Using Young's inequality and the fact that $|\bar{F}_{K,\sigma}| \leq C_{13}\mathrm{diam}(K)^{d-1}$, we have

$$\begin{aligned} |T_{12}| &\leq \frac{1}{2}\sum_{K\in\mathcal{M}}\sum_{\sigma\in\mathcal{E}_K} \nu_K \mathrm{m}(K)\bar{F}^2_{K,\sigma} + \frac{1}{2}\sum_{K\in\mathcal{M}}\sum_{\sigma\in\mathcal{E}_K} \nu_K \mathrm{m}(K)\widehat{F}^2_{K,\sigma} \\ &\leq C_{14}\mathrm{size}(\mathcal{D})^{\beta+2d-2} + \frac{1}{2}\sum_{K\in\mathcal{M}}\sum_{\sigma\in\mathcal{E}_K} \nu_K \mathrm{m}(K)\widehat{F}^2_{K,\sigma}. \end{aligned}$$

Similarly, since $|R_{K,\sigma}| \leq C_{12}\mathrm{diam}(K)^2$,

$$\begin{aligned} |T_{13}| &\leq \frac{1}{2}\sum_{K\in\mathcal{M}}\sum_{\sigma\in\mathcal{E}_K} \frac{R^2_{K,\sigma}}{\nu_K \mathrm{m}(K)} + \frac{1}{2}\sum_{K\in\mathcal{M}}\sum_{\sigma\in\mathcal{E}_K} \nu_K \mathrm{m}(K)\widehat{F}^2_{K,\sigma} \\ &\leq C^2_{12}\sum_{K\in\mathcal{M}}\sum_{\sigma\in\mathcal{E}_K} \frac{\mathrm{diam}(K)^4}{\nu_K \mathrm{m}(K)^2}\mathrm{m}(K) + \frac{1}{2}\sum_{K\in\mathcal{M}}\sum_{\sigma\in\mathcal{E}_K} \nu_K \mathrm{m}(K)\widehat{F}^2_{K,\sigma} \\ &\leq C_{15}\mathrm{size}(\mathcal{D})^{4-2d-\beta} + \frac{1}{2}\sum_{K\in\mathcal{M}}\sum_{\sigma\in\mathcal{E}_K} \nu_K \mathrm{m}(K)\widehat{F}^2_{K,\sigma}. \end{aligned}$$

Gathering these two estimates in (35), the terms involving $\widehat{F}_{K,\sigma}$ in the left-hand side and the right-hand side compensate and we obtain

$$\alpha_0||\widehat{\mathbf{v}}||^2_{L^2(\Omega)^d} \leq C_{16}\left(\mathrm{size}(\mathcal{D})^{\beta+2d-2} + \mathrm{size}(\mathcal{D})^{4-2d-\beta}\right). \tag{36}$$

Estimate (11) follows, using the fact that $\mathrm{size}(\mathcal{D}) \leq 1$ and that $||\bar{\mathbf{v}} - \nabla\bar{u}||_{L^\infty(\Omega)^d} \leq C_{11}\mathrm{size}(\mathcal{D})$. We now set $\widetilde{F}_{K,\sigma} = \widehat{F}_{K,\sigma} + \bar{F}_{K,\sigma} + \frac{R_{K,\sigma}}{\nu_K \mathrm{m}(K)} = F_{K,\sigma} + \frac{R_{K,\sigma}}{\nu_K \mathrm{m}(K)}$ for all $K \in \mathcal{M}$ and $\sigma \in \mathcal{E}_K$, and we estimate $N_2(\mathcal{D},\nu,\widetilde{F})$ the following way:

$$\begin{aligned} N_2(\mathcal{D},\nu,\widetilde{F})^2 &= \sum_{K\in\mathcal{M}}\sum_{\sigma\in\mathcal{E}_K} \mathrm{diam}(K)^{2d-2}\nu^2_K\widetilde{F}^2_{K,\sigma}\mathrm{m}(K) \\ &\leq 2\sum_{K\in\mathcal{M}}\sum_{\sigma\in\mathcal{E}_K} \mathrm{diam}(K)^{2d-2}\nu^2_K F^2_{K,\sigma}\mathrm{m}(K) \\ &\quad +2\sum_{K\in\mathcal{M}}\sum_{\sigma\in\mathcal{E}_K} \mathrm{diam}(K)^{2d-2}\nu^2_K\mathrm{m}(K)\frac{C^2_{12}\mathrm{diam}(K)^4}{(\nu_K\mathrm{m}(K))^2} \\ &\leq C_{17}(\mathrm{size}(\mathcal{D})^{\beta+2d-2} + \mathrm{size}(\mathcal{D})^2) \end{aligned} \tag{37}$$

(we have used (24)). Since (34) implies that $(\widehat{u},\widehat{\mathbf{v}},\widetilde{F}) \in L_\nu(\mathcal{D})$, Lemma 3.1 gives

$$\|\widehat{u}\|_{L^2(\Omega)} \leq C_2\left(\|\widehat{\mathbf{v}}\|_{L^2(\Omega)^d} + N_2(\mathcal{D},\nu,\widetilde{F})\right)$$

and (12) follows from (36), (37) and an easy estimate between $\bar{u}_K$ and the values of $\bar{u}$ on K. □

5 Implementation

We present the practical implementation in the case where $\Lambda(x)$ is symmetric for a.e. $x \in \Omega$, though it is valid for any Λ (notice that, in the physical problems given in the introduction of this paper, the diffusion tensor is always symmetric).

5.1 Resolution procedure

The size of System ((7),(8),(9),(10)) is equal to $(d+1)\mathrm{Card}(\mathcal{M}) + 2\mathrm{Card}(\mathcal{E}_{\mathrm{int}}) + \mathrm{Card}(\mathcal{E}_{\mathrm{ext}})$. However, it is possible to proceed to an algebraic elimination which leads to a symmetric positive definite sparse linear system with $\mathrm{Card}(\mathcal{E}_{\mathrm{int}})$ unknowns, following the same principles as in the hybrid resolution of a mixed finite element problem (see for example [23]). Indeed, for all $(u, \mathbf{v}, F)$ such that (7) and (9) hold, we define $(u_\sigma)_{\sigma\in\mathcal{E}}$ by

$$\mathbf{v}_K \cdot (\mathbf{x}_\sigma - \mathbf{x}_K) + \nu_K F_{K,\sigma} \mathrm{m}(K) = u_\sigma - u_K, \quad \forall K \in \mathcal{M},\ \forall \sigma \in \mathcal{E}_K.$$

We thus have $u_\sigma = 0$ for all $\sigma \in \mathcal{E}_{\mathrm{ext}}$. We can then express $(\mathbf{v}, F)$ as a function of $(u_\sigma)_{\sigma\in\mathcal{E}}$ and of u, since we have

$$\frac{1}{\mathrm{m}(K)} \sum_{\sigma'\in\mathcal{E}_K} F_{K,\sigma'} \Lambda_K^{-1}(\mathbf{x}_{\sigma'} - \mathbf{x}_K)\cdot(\mathbf{x}_\sigma - \mathbf{x}_K) + \nu_K F_{K,\sigma}\mathrm{m}(K) = u_\sigma - u_K, \qquad \forall K \in \mathcal{M},\ \forall \sigma \in \mathcal{E}_K,$$

which is, for all $K \in \mathcal{M}$, an invertible linear system with unknown $(F_{K,\sigma})_{\sigma\in\mathcal{E}_K}$, under the form $B_K(F_{K,\sigma})_{\sigma\in\mathcal{E}_K} = (u_\sigma - u_K)_{\sigma\in\mathcal{E}_K}$ where B_K is a symmetric positive definite matrix (thanks to the condition $\nu_K > 0$). We can then write

$$F_{K,\sigma} = \sum_{\sigma'\in\mathcal{E}_K} (B_K^{-1})_{\sigma\sigma'}(u_{\sigma'} - u_K), \quad \forall K \in \mathcal{M},\ \forall \sigma \in \mathcal{E}_K. \tag{38}$$

We then obtain from (10), denoting $b_{K,\sigma'} = \sum_{\sigma\in\mathcal{E}_K}(B_K^{-1})_{\sigma\sigma'}$ and $b_K = \sum_{\sigma'\in\mathcal{E}_K} b_{K,\sigma'}$, that u_K satisfies the relation

$$-\sum_{\sigma'\in\mathcal{E}_K} b_{K,\sigma'} u_{\sigma'} + b_K u_K = \int_K f(x)\,\mathrm{d}x. \tag{39}$$

We have $(b_{K,\sigma'})_{\sigma'\in\mathcal{E}_K} = B_K^{-1}(1)_{\sigma'\in\mathcal{E}_K}$ and therefore we get $b_K = (1)_{\sigma'\in\mathcal{E}_K} \cdot B_K^{-1}(1)_{\sigma'\in\mathcal{E}_K} > 0$ since B_K^{-1} is symmetric positive definite. Reporting the previous linear relations in (8), we find

$$\begin{aligned}\sum_{\sigma'\in\mathcal{E}_K}\left((B_K^{-1})_{\sigma\sigma'} - \frac{b_{K,\sigma}b_{K,\sigma'}}{b_K}\right)u_{\sigma'} + \sum_{\sigma'\in\mathcal{E}_L}\left((B_L^{-1})_{\sigma\sigma'} - \frac{b_{L,\sigma}b_{L,\sigma'}}{b_L}\right)u_{\sigma'} = \\ \frac{b_{K,\sigma}}{b_K}\int_K f(x)\,\mathrm{d}x + \frac{b_{L,\sigma}}{b_L}\int_L f(x)\,\mathrm{d}x, \quad \forall \sigma = K|L \in \mathcal{E}_{\mathrm{int}},\end{aligned} \tag{40}$$

which is a symmetric linear system, whose unknowns are $(u_\sigma)_{\sigma\in\mathcal{E}_{\mathrm{int}}}$. Let us show that its matrix M is positive. We can write, for all family of real numbers $(u_\sigma)_{\sigma\in\mathcal{E}_{\mathrm{int}}}$,

$$(u_\sigma)_{\sigma\in\mathcal{E}_{\mathrm{int}}} \cdot M\ (u_\sigma)_{\sigma\in\mathcal{E}_{\mathrm{int}}} = \sum_{K\in\mathcal{M}}\left(\sum_{\sigma\in\mathcal{E}_K}\sum_{\sigma'\in\mathcal{E}_K}(B_K^{-1})_{\sigma\sigma'}u_\sigma u_{\sigma'} - \frac{(\sum_{\sigma\in\mathcal{E}_K} b_{K,\sigma}u_\sigma)^2}{b_K}\right).$$

Thanks to the fact that B_K^{-1} is symmetric positive definite, we get, using the Cauchy-Schwarz inequality,

$$\left((1)_{\sigma\in\mathcal{E}_K}\cdot B_K^{-1}(u_\sigma)_{\sigma\in\mathcal{E}_K}\right)^2 \le \left((1)_{\sigma\in\mathcal{E}_K}\cdot B_K^{-1}(1)_{\sigma\in\mathcal{E}_K}\right)\left((u_\sigma)_{\sigma\in\mathcal{E}_K}\cdot B_K^{-1}(u_\sigma)_{\sigma\in\mathcal{E}_K}\right),$$

which is exactly

$$\left(\sum_{\sigma\in\mathcal{E}_K} b_{K,\sigma}u_\sigma\right)^2 \le b_K\sum_{\sigma\in\mathcal{E}_K}\sum_{\sigma'\in\mathcal{E}_K}(B_K^{-1})_{\sigma\sigma'}u_\sigma u_{\sigma'}.$$

In order to show that M is definite, we simply remark that the preceding reasoning shows that the systems ((7),(8),(9),(10)) and (40) are equivalent. Hence, since ((7),(8),(9),(10)) has a unique solution, so must (40), which means that M is invertible.

Hence, we can first solve $(u_\sigma)_{\sigma\in\mathcal{E}_{\rm int}}$ from (40), and then compute (u,F) thanks to relations (39) and (38) and finally $\mathbf{v}$ by (9).

5.2 Numerical results

Taking $\nu_K = 0$ for all $K \in \mathcal{M}$, we could prove in the symmetric case, via a minimization technique, that there exists at least one $(u,\mathbf{v},F) \in L_\nu(\mathcal{D})$ solution of ((7),(8),(9),(10)). In this case, $(u,\mathbf{v})$ is unique, but this is no longer true for F in the general case (see however section 6.2). Within such a choice, the proof of convergence of $(u,\mathbf{v})$ to the continuous solution remains an open problem. Nevertheless, this gives an indication that very small values of $(\nu_K)_{K\in\mathcal{M}}$ can be considered. Hence we take $\nu_K = 10^{-9}/\mathrm{m}(K)$ in all the following computations. The inversion of matrices B_K arising in (38) and the solving of System (40) are then realized using direct methods.

5.2.1 Case of a homogeneous isotropic problem

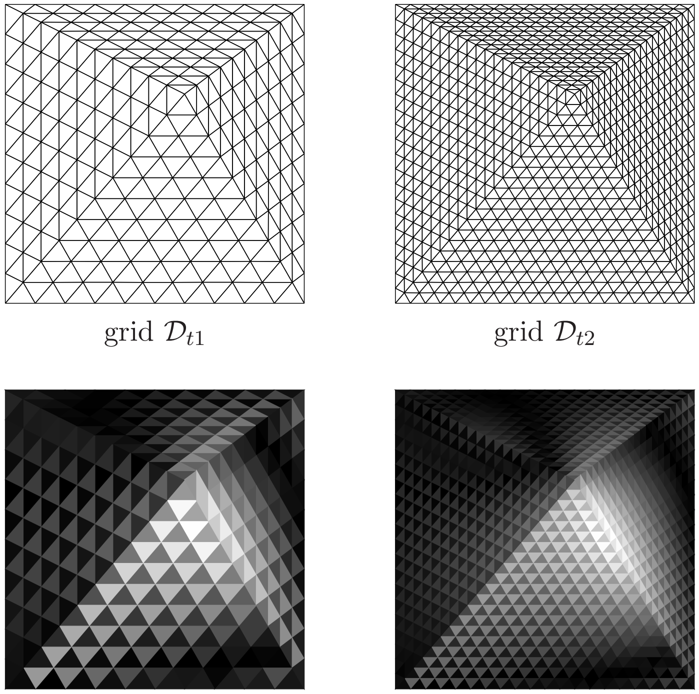
We consider here the case $d = 2$, $\Omega = (0,1)\times(0,1)$, $\Lambda = \mathrm{I}_d$ and $\bar{u}(x) = x_1(1-x_1)x_2(1-x_2)$ for all $x = (x_1,x_2)\in\Omega$.

We first present in Figure 1 two different triangular discretizations $\mathcal{D}_{t1}$ and $\mathcal{D}_{t2}$ used for the computation of an approximate solution for the problem. We also show in Figure 1 the error $e_\mathcal{D}$, defined by

$$e_K = \frac{|u_K - \bar{u}(\mathbf{x}_K)|}{\|\bar{u}\|_{L^\infty(\Omega)}},\quad \forall K\in\mathcal{M},$$

using discretizations $\mathcal{D}_{t1}$ and $\mathcal{D}_{t2}$. Note that these discretizations do not respect the Delaunay condition on a sub-domain of Ω, and that the 4-point finite volume scheme (see [11]) cannot be used on these grids. The grids $\mathcal{D}_{t2}$ and $\mathcal{D}_{t3}$ (which is not represented here) have been obtained from $\mathcal{D}_{t1}$ (containing 400 control volumes) by the respective divisions by 2 and 4 of each edge (there are 1600 control volumes in $\mathcal{D}_{t2}$ and 6400 in $\mathcal{D}_{t3}$). For all these discretizations, the points $\mathbf{x}_K$ have been located at the center of gravity of the control volumes. The errors in L^2 norm obtained with these grids are given in Table 1.

We observe that the numerical orders of convergence for $\|u-\bar{u}\|_{L^2(\Omega)}$ and $\|\mathbf{v}-\nabla\bar{u}\|_{L^2(\Omega)^d}$ both seem to be near 1, and therefore no super-convergence property can reasonably be expected in this case.



grid $\mathcal{D}_{t1}$

grid $\mathcal{D}_{t2}$

error on $\mathcal{D}_{t1}$
black = 0, white = $2.2\ 10^{-2}$

error on $\mathcal{D}_{t2}$
black = 0, white = $8.9\ 10^{-3}$

Figure 1: Non-Delaunay triangular grids and error $e_{\mathcal{D}}$.

Grid	$\|u - \bar{u}\|_{L^2(\Omega)}$	$\|\mathbf{v} - \nabla\bar{u}\|_{L^2(\Omega)^d}$
$\mathcal{D}_{t1}$	$5.1\ 10^{-4}$	$1.8\ 10^{-2}$
$\mathcal{D}_{t2}$	$1.9\ 10^{-4}$	$9.0\ 10^{-3}$
$\mathcal{D}_{t3}$	$8.2\ 10^{-5}$	$4.5\ 10^{-3}$
order of convergence	1.3	1

Table 1: Errors on non-Delaunay triangular grids.

We then present in Figure 2 discretizations $\mathcal{D}_{q1}$ and $\mathcal{D}_{q2}$ and error $e_{\mathcal{D}}$ using these grids. Such grids could be obtained using a refinement procedure: for example, in the case of coupled systems, the grid might have been refined in order to improve the convergence on another equation (thanks to some *a posteriori* estimates maybe) and must then be used to solve (1) which is the second part of the system. The grid $\mathcal{D}_{q2}$ has been obtained from $\mathcal{D}_{q1}$ by a uniform division of each edge by 2, and $\mathcal{D}_{q3}$ (not represented here) has been obtained from $\mathcal{D}_{q2}$ in the same way. The respective errors in L^2 norm obtained with these grids are given in Table 2.

Grid	$\|u - \bar{u}\|_{L^2(\Omega)}$	$\|\mathbf{v} - \nabla\bar{u}\|_{L^2(\Omega)^d}$
$\mathcal{D}_{q1}$	$8.7\ 10^{-4}$	$5.8\ 10^{-3}$
$\mathcal{D}_{q2}$	$1.7\ 10^{-4}$	$1.3\ 10^{-3}$
$\mathcal{D}_{q3}$	$3.9\ 10^{-5}$	$4.0\ 10^{-4}$
order of convergence	2.2	1.8

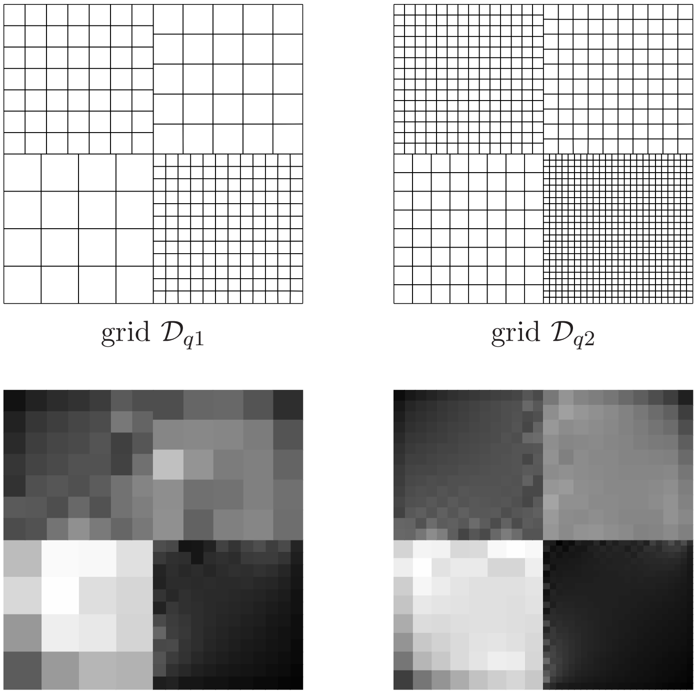
Table 2: Errors on rectangular locally refined grids.

We then observe that the numerical order convergence is better than 2 for $\|u - \bar{u}\|_{L^2(\Omega)}$, which corresponds to a case of a mainly structured grid (there is no significant additional error located at the internal boundaries between the differently gridded subdomains, see Figure 2).
Finally, in Figure 3, we represent grids $\mathcal{D}_\flat$ and $\mathcal{D}_\sharp$ and the error $e_{\mathcal{D}}$ thus obtained. These meshes (which have the same number of control volumes) could correspond to the case of moving meshes (for example, due to a phenomenon of compaction, see [15]). The respective errors in L^2 norm obtained with these grids are given in Table 3.

Grid	$\|u - \bar{u}\|_{L^2(\Omega)}$	$\|\mathbf{v} - \nabla\bar{u}\|_{L^2(\Omega)^d}$
$\mathcal{D}_\flat$	$2.0\ 10^{-4}$	$6.7\ 10^{-4}$
$\mathcal{D}_\sharp$	$4.6\ 10^{-4}$	$1.8\ 10^{-3}$

Table 3: Errors on “compacted” grids.

We observe that the error is mainly connected to the size of the control volumes, and maybe to some effect of loss of regularity of the mesh.



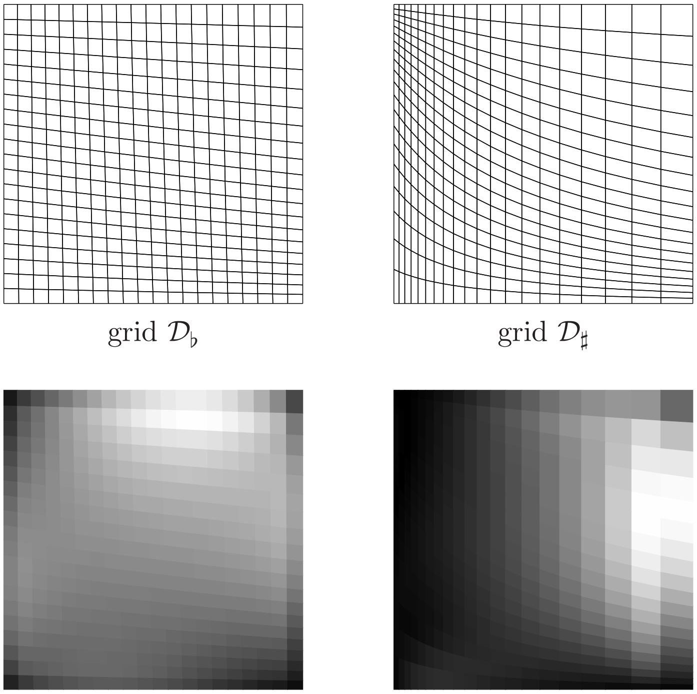
grid $\mathcal{D}_{q1}$

grid $\mathcal{D}_{q2}$

error on $\mathcal{D}_{q1}$
black $= 0$, white $= 2.7\ 10^{-2}$

error on $\mathcal{D}_{q2}$
black $= 0$, white $= 5.3\ 10^{-3}$

Figure 2: Rectangular locally refined grids and error $e_{\mathcal{D}}$.



grid $\mathcal{D}_\flat$

grid $\mathcal{D}_\sharp$

error on $\mathcal{D}_\flat$
black = 0, white = 5.4 10^{-3}

error on $\mathcal{D}_\sharp$
black = 0, white = 1.5 10^{-2}

Figure 3: “Compacted” grids and error $e_{\mathcal{D}}$.

5.2.2 Case of a heterogeneous anisotropic problem, comparison with mixed finite element schemes

Let us now give some numerical results in a highly heterogeneous and anisotropic case, inspired by [20]. With $\Omega = (0,1) \times (0,1)$, let us define $\bar{x} = (-0.1, -0.1)$ and $\varepsilon = 10^{-4}$, and let us set

$$\Lambda(x) = \begin{pmatrix} (x_2 - \bar{x}_2)^2 + \varepsilon(x_1 - \bar{x}_1)^2 & -(1-\varepsilon)(x_1 - \bar{x}_1)(x_2 - \bar{x}_2) \\ \\ -(1-\varepsilon)(x_1 - \bar{x}_1)(x_2 - \bar{x}_2) & (x_1 - \bar{x}_1)^2 + \varepsilon(x_2 - \bar{x}_2)^2 \end{pmatrix}, \ \forall x \in \Omega.$$

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 $1/\varepsilon = 10^4$ in the whole domain. Note that, thanks to the choice $\bar{x} = (-0.1, -0.1)$, we have $\inf_{x \in \Omega} \underline{\lambda}(x) = |\bar{x}|^2 \varepsilon = 0.02\varepsilon$ and $\sup_{x \in \Omega} \underline{\lambda}(x) = |\widehat{x} - \bar{x}|^2 \varepsilon = 2.42\varepsilon$ with $\widehat{x} = (1,1)$. Therefore $\underline{\lambda}(x)/\underline{\lambda}(y)$ and $\overline{\lambda}(x)/\overline{\lambda}(y)$ are in the range $[1/121, 121]$ for all $x, y \in \Omega$ (note that in [20], these ratios are in the range $(0, +\infty)$ since the author takes $\bar{x} = (0,0)$, but then (3) does not hold). Since the directions of anisotropy are not constant, one cannot solve this problem by a classical finite volume method on a tilted rectangular mesh. We assume that the solution of Problem (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 the function f satisfies:

$$\begin{aligned} f(x) \quad = \quad & \pi^2(1+\varepsilon)\sin(\pi x_1)\sin(\pi x_2)|x - \bar{x}|^2 \\ & +\pi(1-3\varepsilon)\cos(\pi x_1)\sin(\pi x_2)(x_1 - \bar{x}_1) \\ & +\pi(1-3\varepsilon)\sin(\pi x_1)\cos(\pi x_2)(x_2 - \bar{x}_2) \\ & +2\pi^2(1-\varepsilon)\cos(\pi x_1)\cos(\pi x_2)(x_1 - \bar{x}_1)(x_2 - \bar{x}_2), \ \forall x \in \Omega. \end{aligned}$$

We then compare on this problem the numerical solution given by the mixed finite volume scheme (denoted by MFV below) with the one obtained using the low degree mixed finite element scheme (denoted by MFE below) in the case of triangles or rectangles. We compute the solutions with both schemes on the following grids: $\mathcal{D}_{t4}$, including 5600 acute triangles, $\mathcal{D}_{t5}$, including $4 \times 5600 = 22400$ acute triangles, $\mathcal{D}_{t6}$, including $16 \times 5600 = 89600$ acute triangles, $\mathcal{D}_{q4}$, including 1600 rectangles (in fact, squares), $\mathcal{D}_{q5}$, including $4 \times 1600 = 6400$ rectangles, $\mathcal{D}_{q6}$, including $25 \times 1600 = 40000$ rectangles. For the triangular grids $\mathcal{D}_{t4}$, $\mathcal{D}_{t5}$ and $\mathcal{D}_{t6}$, the points $\mathbf{x}_K$ have been located at the circumcenter of the triangles.

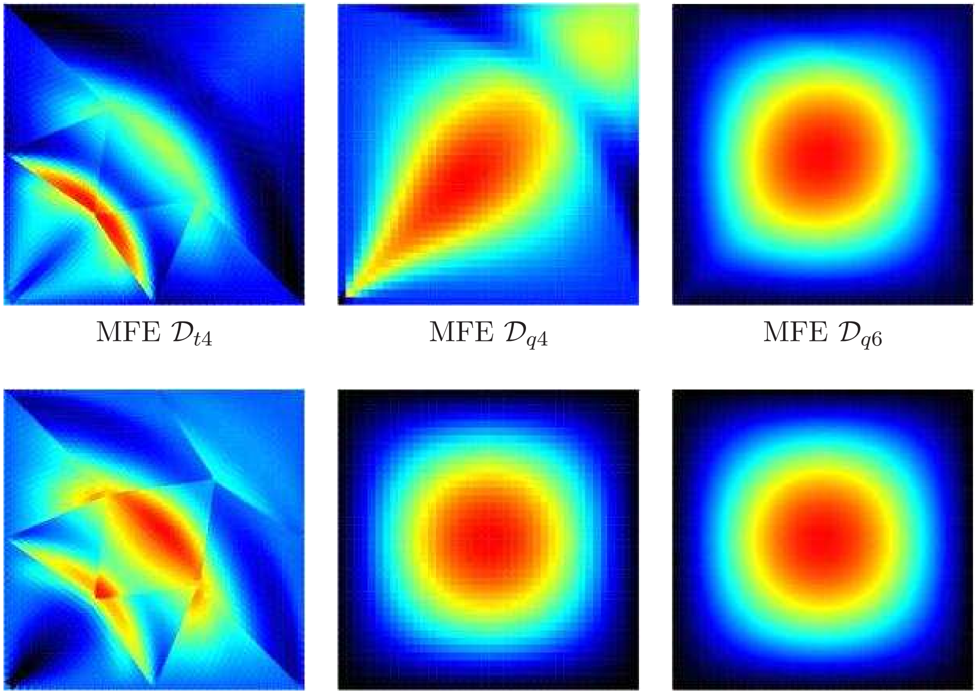
Remark 5.1 *Choosing for $\mathbf{x}_K$ the circumcenter of the triangle instead of the center of gravity leads to an error about ten percent lower on the grids $\mathcal{D}_{t4}$, $\mathcal{D}_{t5}$ and $\mathcal{D}_{t6}$.*

For the rectangular grids, the points $\mathbf{x}_K$ have been located at the center of gravity of the control volumes. We provide in Table 4 the error $\|u - \bar{u}\|_{L^2(\Omega)}$, as well as the minimum value $u_{\min} = \min_{K \in \mathcal{M}} u_K$ and the maximum value $u_{\max} = \max_{K \in \mathcal{M}} u_K$ of the approximate solution (note that the exact solution $\bar{u}$ varies between 0 at the edges of Ω and 1 at its center), using both schemes.

These results show a surprisingly bad performance for the MFE and MFV schemes on triangular grids (this was pointed out for the MFE scheme in [20]). An order of convergence close to 2 is nevertheless observed for the $L^2(\Omega)$ norm, with a very high multiplicative constant. But this similarity between both schemes does no longer hold on the other grids: on the regular rectangular grids (on which the MFE solution can be computed using the classical RT basis), the MFV method provides accurate results where the MFE scheme is far from the exact solution. Moreover, in the case of the MFV scheme, the bounds on the approximate solution are close to

Grid	MFE	MFE	MFE	MFV	MFV	MFV
	$\|u - \bar{u}\|_{L^2(\Omega)}$	$u_{\min}$	$u_{\max}$	$\|u - \bar{u}\|_{L^2(\Omega)}$	$u_{\min}$	$u_{\max}$
$\mathcal{D}_{t4}$	1.53	-1.32	6.35	1.20	-2.46	4.68
$\mathcal{D}_{t5}$	0.397	-0.344	2.20	0.315	-0.633	1.99
$\mathcal{D}_{t6}$	0.101	-0.0867	1.20	0.0807	-0.163	1.25
order of convergence	1.96	—	—	1.95	—	—
$\mathcal{D}_{q4}$	0.795	-1.03	2.62	0.000912	0.000566	0.997
$\mathcal{D}_{q5}$	0.200	-0.259	1.38	0.000162	0.000141	0.999
$\mathcal{D}_{q6}$	0.0320	-0.0415	1.06	0.0000202	0.0000229	1.00
order of convergence	2.3	—	—	2.75	—	—

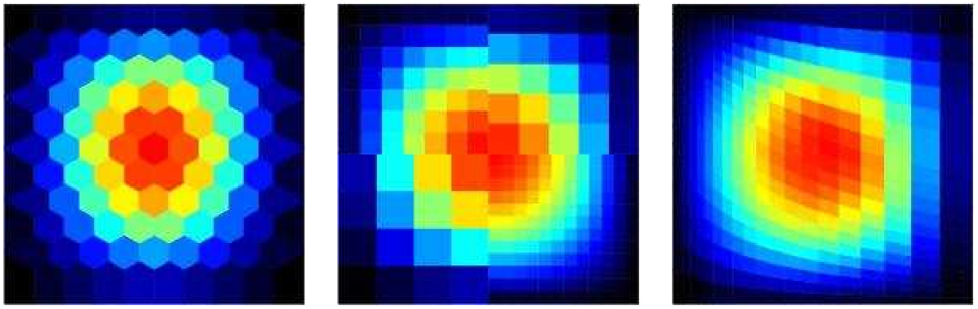
Table 4: Comparison between the mixed finite element and mixed finite volume schemes on triangular and rectangular grids for Le Potier's test case.



MFE $\mathcal{D}_{t4}$ MFE $\mathcal{D}_{q4}$ MFE $\mathcal{D}_{q6}$

MFV $\mathcal{D}_{t4}$ MFV $\mathcal{D}_{q4}$ MFV $\mathcal{D}_{q6}$

Figure 4: Solutions of mixed finite element and mixed finite volume schemes for Le Potier's test case (black $= u_{\min}$, red $= u_{\max}$).



MFV $\mathcal{D}_v$ MFV $\mathcal{D}_{q1}$ MFV $\mathcal{D}_\sharp$

Figure 5: Solutions of the mixed finite volume scheme for Le Potier's test case on irregular grids (black = $u_{\min}$, red = $u_{\max}$).

that of the exact solution. These results are confirmed by Figure 4, where some of the numerical solutions considered are plotted.
We give in Table 5 the values $\|u - \bar{u}\|_{L^2(\Omega)}$, $u_{\min}$ and $u_{\max}$ in the case where the MFV scheme is used on three irregular grids: the grid $\mathcal{D}_v$ which is a Voronoï tessellation with 105 control volumes, the grid $\mathcal{D}_{q1}$, already considered above, including $16 + 144 + 49 + 25 = 234$ rectangles (in fact again, squares) and the grid $\mathcal{D}_\sharp$ with 400 quadrangles, also considered above.

Grid	MFV $\|u - \bar{u}\|_{L^2(\Omega)}$	MFV $u_{\min}$	MFV $u_{\max}$
$\mathcal{D}_v$	0.0929	0.0126	0.980
$\mathcal{D}_{q1}$	0.0232	0.00259	1.00
$\mathcal{D}_\sharp$	0.0217	-0.00890	0.999

Table 5: Errors and minima/maxima values of the MFV solution for Le Potier's test case on irregular grids.

These results show an acceptable convergence, confirmed by Figure 5 in which the corresponding approximate solutions are drawn.

6 Appendix

6.1 Technical lemmas

Lemma 6.1 justifies the link (9) between the approximate gradient and the approximate fluxes.

Lemma 6.1 *Let K be a non empty open convex polygonal set in $\mathbb{R}^d$. For $\sigma \in \mathcal{E}_K$ (the edges of K, in the sense given in Definition 2.1), we let $\mathbf{x}_\sigma$ be the center of gravity of σ; we also denote $\mathbf{n}_{K,\sigma}$ the unit normal to σ outward to K. Then, for all vector $\mathbf{e} \in \mathbb{R}^d$ and for all point $\mathbf{x}_K \in \mathbb{R}^d$, we have*

$$\mathrm{m}(K)\mathbf{e} = \sum_{\sigma \in \mathcal{E}_K} \mathrm{m}(\sigma)\mathbf{e} \cdot \mathbf{n}_{K,\sigma}(\mathbf{x}_\sigma - \mathbf{x}_K).$$

PROOF.
We denote by a superscript i the i-th coordinate of vectors and points in $\mathbb{R}^d$. By Stokes formula, we have

$$\mathrm{m}(K)\mathbf{e}^i = \int_K \mathrm{div}((x^i - \mathbf{x}_K^i)\mathbf{e})\,\mathrm{d}x = \sum_{\sigma\in\mathcal{E}_K}\int_\sigma (x^i - \mathbf{x}_K^i)\mathbf{e}\cdot\mathbf{n}_{K,\sigma}\,\mathrm{d}\gamma(x)$$

and the proof is concluded since, by definition of the center of gravity, $\int_\sigma (x^i - \mathbf{x}_K^i)\,\mathrm{d}\gamma(x) = \int_\sigma x^i\,d\gamma(x) - \mathrm{m}(\sigma)\mathbf{x}_K^i = \mathrm{m}(\sigma)\mathbf{x}_\sigma^i - \mathrm{m}(\sigma)\mathbf{x}_K^i$. □

The following lemma is quite similar to [9, Lemma 7.2], but since the proof of Lemma 6.3 uses this result with slightly more general hypotheses than in [9], we include the full proof of Lemma 6.2 for sake of completeness.

Lemma 6.2 *Let K be a non empty open polygonal convex set in $\mathbb{R}^d$. Let E be an affine hyperplane of $\mathbb{R}^d$ and σ be a non empty open subset of E contained in $\partial K\cap E$. We assume that there exists $\alpha>0$ and $\mathbf{p}_K\in K$ such that $B(\mathbf{p}_K,\alpha\mathrm{diam}(K))\subset K$. We denote $\triangle_{K,\sigma}$ the convex hull of σ and $\mathbf{p}_K$. Then there exists C_{18} only depending on d and α such that, for all $v\in H^1(K)$,*

$$\left(\frac{1}{\mathrm{m}(\triangle_{K,\sigma})}\int_{\triangle_{K,\sigma}} v(x)\,\mathrm{d}x - \frac{1}{\mathrm{m}(\sigma)}\int_\sigma v(\xi)\,\mathrm{d}\gamma(\xi)\right)^2 \le \frac{C_{18}\mathrm{dist}(\mathbf{p}_K,E)^2}{\mathrm{m}(\triangle_{K,\sigma})}\int_{\triangle_{K,\sigma}}|\nabla v(x)|^2\,\mathrm{d}x.$$

PROOF.
The regular functions being dense in $H^1(K)$ (since K is convex), it is sufficient to prove the lemma for $v\in C^1(\mathbb{R}^d)$. By translation and rotation, we can assume that $E=\{0\}\times\mathbb{R}^{d-1}$, $\sigma=\{0\}\times\widetilde{\sigma}$ with $\widetilde{\sigma}\subset\mathbb{R}^{d-1}$ and that $\mathbf{p}_K=(p_1,0)$ with $p_1=\mathrm{dist}(\mathbf{p}_K,E)$.
Notice that, since K is convex and $\partial K\cap E$ contains a non empty open subset of E, K is on one side of E. In particular, $B(\mathbf{p}_K,\alpha\mathrm{diam}(K))$ is also on one side of E (it is contained in K) and

$$p_1 = \mathrm{dist}(\mathbf{p}_K,E)\ge\alpha\mathrm{diam}(K). \tag{41}$$

For $a\in[0,p_1]$, we denote $\widetilde{\sigma}_a=\{z\in\mathbb{R}^{d-1}\mid(a,z)\in\triangle_{K,\sigma}\}$. By definition, $(a,z)\in\triangle_{K,\sigma}$ if and only if there exists $t\in[0,1]$ and $y\in\widetilde{\sigma}$ such that $t(p_1,0)+(1-t)(0,y)=(a,z)$; this is equivalent to $t=\frac{a}{p_1}$ and $z=(1-t)y=\left(1-\frac{a}{p_1}\right)y$. Thus, $\widetilde{\sigma}_a=\left(1-\frac{a}{p_1}\right)\widetilde{\sigma}$.

For all $y\in\widetilde{\sigma}$ and all $a\in[0,p_1]$, we have

$$v(0,y)-v\left(a,\left(1-\frac{a}{p_1}\right)y\right)=\int_0^1\nabla v\left(ta,\left(1-t\frac{a}{p_1}\right)y\right)\cdot\left(-a,\frac{a}{p_1}y\right)\,\mathrm{d}t.$$

Integrating on $y\in\widetilde{\sigma}$ and using the change of variable $z=\left(1-\frac{a}{p_1}\right)y$, we find

$$\int_\sigma v(\xi)\,\mathrm{d}\gamma(\xi)-\frac{1}{\left(1-\frac{a}{p_1}\right)^{d-1}}\int_{\widetilde{\sigma}_a}v(a,z)\,\mathrm{d}z=\int_{\widetilde{\sigma}}\int_0^1\nabla v\left(ta,\left(1-t\frac{a}{p_1}\right)y\right)\cdot\left(-a,\frac{a}{p_1}y\right)\,\mathrm{d}t\mathrm{d}y.$$

Multiplying by $\left(1-\frac{a}{p_1}\right)^{d-1}$ and integrating on $a\in[0,p_1]$, we obtain

$$\begin{aligned}&\int_\sigma v(\xi)\,\mathrm{d}\gamma(\xi)\int_0^{p_1}\left(1-\frac{a}{p_1}\right)^{d-1}\mathrm{d}a-\int_0^{p_1}\int_{\widetilde{\sigma}_a}v(a,z)\,\mathrm{d}z\mathrm{d}a\\&\quad=\int_0^{p_1}\left(1-\frac{a}{p_1}\right)^{d-1}\int_{\widetilde{\sigma}}\int_0^1\nabla v\left(ta,\left(1-t\frac{a}{p_1}\right)y\right)\cdot\left(-a,\frac{a}{p_1}y\right)\,\mathrm{d}t\mathrm{d}y\mathrm{d}a.\end{aligned}$$

But $\int_0^{p_1}\left(1-\frac{a}{p_1}\right)^{d-1}\,\mathrm{d}a=\frac{p_1}{d}$ and $\mathrm{m}(\triangle_{K,\sigma})=\frac{\mathrm{m}(\sigma)p_1}{d}$; therefore, dividing by $\mathrm{m}(\triangle_{K,\sigma})$, we find

$$\begin{aligned}
&\frac{1}{\mathrm{m}(\sigma)}\int_\sigma v(\xi)\,\mathrm{d}\gamma(\xi)-\frac{1}{\mathrm{m}(\triangle_{K,\sigma})}\int_{\triangle_{K,\sigma}}v(x)\,\mathrm{d}x\\
&\quad=\ \frac{1}{\mathrm{m}(\triangle_{K,\sigma})}\int_0^{p_1}\left(1-\frac{a}{p_1}\right)^{d-1}\int_{\widetilde{\sigma}}\int_0^1\nabla v\left(ta,\left(1-t\frac{a}{p_1}\right)y\right)\cdot\left(-a,\frac{a}{p_1}y\right)\,\mathrm{d}t\mathrm{d}y\mathrm{d}a. \qquad (42)
\end{aligned}$$

For all $y\in\widetilde{\sigma}$, we have $|y|=|(0,y)|\le|(0,y)-\mathbf{p}_K|+|\mathbf{p}_K|\le\mathrm{diam}(K)+p_1$ (because $(0,y)$ and $\mathbf{p_K}$ belong to $\overline{K}$). By (41), this implies $|y|\le(\frac{1}{\alpha}+1)p_1$ and thus

$$\begin{aligned}
&\left|\int_0^{p_1}\left(1-\frac{a}{p_1}\right)^{d-1}\int_{\widetilde{\sigma}}\int_0^1\nabla v\left(ta,\left(1-t\frac{a}{p_1}\right)y\right)\cdot\left(-a,\frac{a}{p_1}y\right)\,\mathrm{d}t\mathrm{d}y\mathrm{d}a\right|\\
&\quad\le\ C_{19}\int_0^{p_1}\left(1-\frac{a}{p_1}\right)^{d-1}\int_{\widetilde{\sigma}}\int_0^1\left|\nabla v\left(ta,\left(1-t\frac{a}{p_1}\right)y\right)\right|a\,\mathrm{d}t\mathrm{d}y\mathrm{d}a\\
&\quad\le\ C_{19}\int_0^{p_1}\int_{\widetilde{\sigma}}\int_0^1\left|\nabla v\left(ta,\left(1-t\frac{a}{p_1}\right)y\right)\right|a\left(1-\frac{ta}{p_1}\right)^{d-1}\,\mathrm{d}t\mathrm{d}y\mathrm{d}a \qquad (43)
\end{aligned}$$

where C_{19} only depends on α (we have used the obvious fact that, for $t\in]0,1[$, $1-\frac{a}{p_1}\le 1-\frac{ta}{p_1}$). But, for all $a\in]0,p_1[$, the change of variable

$$\varphi_a:(t,y)\in]0,1[\times\widetilde{\sigma}\to z=\left(ta,\left(1-t\frac{a}{p_1}\right)y\right)\in\varphi_a(]0,1[\times\widetilde{\sigma})$$

has Jacobian determinant equal to $a\left(1-\frac{ta}{p_1}\right)^{d-1}$ and therefore

$$\int_{\widetilde{\sigma}}\int_0^1\left|\nabla v\left(ta,\left(1-t\frac{a}{p_1}\right)y\right)\right|a\left(1-\frac{ta}{p_1}\right)^{d-1}\,\mathrm{d}t\mathrm{d}y=\int_{\varphi_a(]0,1[\times\widetilde{\sigma})}|\nabla v(z)|\,\mathrm{d}z.$$

Moreover, $(ta,(1-t\frac{a}{p_1})y)=\frac{ta}{p_1}(p_1,0)+(1-\frac{ta}{p_1})(0,y)$ with $\frac{ta}{p_1}\in]0,1[$; hence, $\varphi_a(]0,1[\times\widetilde{\sigma})\subset\triangle_{K,\sigma}$ and we obtain

$$\int_0^{p_1}\int_{\widetilde{\sigma}}\int_0^1\left|\nabla v\left(ta,\left(1-t\frac{a}{p_1}\right)y\right)\right|a\left(1-\frac{ta}{p_1}\right)^{d-1}\,\mathrm{d}t\mathrm{d}y\mathrm{d}a\le p_1\int_{\triangle_{K,\sigma}}|\nabla v(z)|\,\mathrm{d}z.$$

We introduce this inequality in (43) and use the resulting estimate in (42) to obtain

$$\left|\frac{1}{\mathrm{m}(\triangle_{K,\sigma})}\int_{\triangle_{K,\sigma}}v(x)\,\mathrm{d}x-\frac{1}{\mathrm{m}(\sigma)}\int_\sigma v(\xi)\,\mathrm{d}\gamma(\xi)\right|\le\frac{C_{19}p_1}{\mathrm{m}(\triangle_{K,\sigma})}\int_{\triangle_{K,\sigma}}|\nabla v(x)|\,\mathrm{d}x$$

and the conclusion follows from the Cauchy-Schwarz inequality, recalling that $p_1=\mathrm{dist}(\mathbf{p}_K,E)$.
□

Lemma 6.3 *Let K be a non empty open polygonal convex set in $\mathbb{R}^d$ such that, for some $\alpha>0$, there exists a ball of radius $\alpha\mathrm{diam}(K)$ contained in K. Let E be an affine hyperplane of $\mathbb{R}^d$ and σ be a non empty open subset of E contained in $\partial K\cap E$. Then there exists C_{20} only depending on d and α such that, for all $v\in H^1(K)$,*

$$\left(\frac{1}{\mathrm{m}(K)}\int_K v(x)\,\mathrm{d}x-\frac{1}{\mathrm{m}(\sigma)}\int_\sigma v(x)\,\mathrm{d}\gamma(x)\right)^2\le\frac{C_{20}\mathrm{diam}(K)}{\mathrm{m}(\sigma)}\int_K|\nabla v(x)|^2\,\mathrm{d}x.$$

Proof.
Let $B(\mathbf{p}_K, \alpha \mathrm{diam}(K)) \subset K$ and $\triangle_{K,\sigma}$ be the convex hull of $\mathbf{p}_K$ and σ. By Lemma 6.2, we have

$$\left(\frac{1}{\mathrm{m}(\triangle_{K,\sigma})}\int_{\triangle_{K,\sigma}} v(x)\,\mathrm{d}x - \frac{1}{\mathrm{m}(\sigma)}\int_\sigma v(x)\,\mathrm{d}\gamma(x)\right)^2 \le \frac{C_{18}\mathrm{dist}(\mathbf{p}_K,E)^2}{\mathrm{m}(\triangle_{K,\sigma})}\int_K |\nabla v(x)|^2\,\mathrm{d}x.$$

But $\mathrm{m}(\triangle_{K,\sigma}) = \frac{\mathrm{m}(\sigma)\mathrm{dist}(\mathbf{p}_K,E)}{d}$ and $\mathrm{dist}(\mathbf{p}_K,E) \le \mathrm{dist}(\mathbf{p}_K,\sigma) \le \mathrm{diam}(K)$. Therefore,

$$\left(\frac{1}{\mathrm{m}(\triangle_{K,\sigma})}\int_{\triangle_{K,\sigma}} v(x)\,\mathrm{d}x - \frac{1}{\mathrm{m}(\sigma)}\int_\sigma v(x)\,\mathrm{d}\gamma(x)\right)^2 \le \frac{C_{18}d\,\mathrm{diam}(K)}{\mathrm{m}(\sigma)}\int_K |\nabla v(x)|^2\,\mathrm{d}x. \quad (44)$$

Using Lemma 7.1 in [9], we get C_{21} only depending on d such that

$$\left(\frac{1}{\mathrm{m}(\triangle_{K,\sigma})}\int_{\triangle_{K,\sigma}} v(x)\,\mathrm{d}x - \frac{1}{\mathrm{m}(K)}\int_K v(x)\,\mathrm{d}x\right)^2 \le \frac{C_{21}\mathrm{diam}(K)^{d+2}}{\mathrm{m}(\triangle_{K,\sigma})\mathrm{m}(K)}\int_K |\nabla v(x)|^2\,\mathrm{d}x,$$

which implies

$$\left(\frac{1}{\mathrm{m}(\triangle_{K,\sigma})}\int_{\triangle_{K,\sigma}} v(x)\,\mathrm{d}x - \frac{1}{\mathrm{m}(K)}\int_K v(x)\,\mathrm{d}x\right)^2 \le \frac{C_{21}d\,\mathrm{diam}(K)^{d+2}}{\mathrm{m}(\sigma)\mathrm{dist}(\mathbf{p}_K,E)\mathrm{m}(K)}\int_K |\nabla v(x)|^2\,\mathrm{d}x.$$

But, as in the proof of Lemma 6.2, we have $\mathrm{dist}(\mathbf{p}_K,E) \ge \alpha\mathrm{diam}(K)$ (see (41)). Since $\mathrm{m}(K) \ge \omega_d\alpha^d\mathrm{diam}(K)^d$, we deduce that

$$\left(\frac{1}{\mathrm{m}(\triangle_{K,\sigma})}\int_{\triangle_{K,\sigma}} v(x)\,\mathrm{d}x - \frac{1}{\mathrm{m}(K)}\int_K v(x)\,\mathrm{d}x\right)^2 \le \frac{C_{21}d\,\mathrm{diam}(K)}{\omega_d\alpha^{d+1}\mathrm{m}(\sigma)}\int_K |\nabla v(x)|^2\,\mathrm{d}x. \quad (45)$$

The lemma follows from (44) and (45). □

6.2 Simplicial meshes

For some meshes, it is possible to completely drop the penalization on the fluxes, that is to say to take $\nu_K = 0$ in (7). This is for example the case if each control volume K of the mesh is a simplex, i.e. if K is the interior of the convex hull of $d+1$ points of $\mathbb{R}^d$ such that no affine hyperplane of $\mathbb{R}^d$ contains all of them and if $\mathrm{Card}(\mathcal{E}_K) = d+1$. In this situation, the following lemma is the key ingredient to the study of the mixed finite volume scheme with $\nu_K = 0$.

Lemma 6.4 *Let us assume Assumptions (2)-(4). Let $\mathcal{D}$ be an admissible discretization of Ω in the sense of Definition 2.1, such that $\mathrm{regul}(\mathcal{D}) \le \theta$ for some $\theta > 0$ and $\mathcal{M}$ is made of simplicial control volumes. Let $\mathbf{v} \in H^d_{\mathcal{D}}$ and a family of real numbers $(F_{K,\sigma})_{K\in\mathcal{M},\,\sigma\in\mathcal{E}_K}$ be given such that (9) and (10) hold. Then there exists C_{22} only depending on d, Ω, Λ and θ such that*

$$\sum_{K\in\mathcal{M}}\sum_{\sigma\in\mathcal{E}_K}\mathrm{diam}(K)^{2-d}F_{K,\sigma}^2 \le C_{22}(||f||^2_{L^2(\Omega)} + ||\mathbf{v}||^2_{L^2(\Omega)^d}). \quad (46)$$

PROOF.
For $K \in \mathcal{M}$, let A_K be the $(d+1)\times(d+1)$ matrix whose columns are $(1, \mathbf{x}_\sigma - \mathbf{x}_K)^T_{\sigma\in\mathcal{E}_K}$ (since K is simplicial, it has $d+1$ edges and A_K is indeed a square matrix). The equations (9)-(10) can be written $A_K F_K = E_K$, where $F_K = (F_{K,\sigma})_{\sigma\in\mathcal{E}_K}$ and

$$E_K = \begin{pmatrix} -\int_K f(x)\,\mathrm{d}x \\ \mathrm{m}(K)\Lambda_K \mathbf{v}_K \end{pmatrix}.$$

We now want to estimate $||A_K^{-1}||$ (the matrix norm being induced by the euclidean norm on $\mathbb{R}^{d+1}$) and, in order to achieve this, we divide the rest of the proof in several steps.

Step 1: this step is devoted to allow the assumption $\mathrm{diam}(K)=1$ in Steps 2 and 3.
Let $K_0 = \mathrm{diam}(K)^{-1}K$. Then $\mathbf{x}_{K,0} = \mathrm{diam}(K)^{-1}\mathbf{x}_K \in K_0$ and the barycenters of the edges of K_0 are $\mathbf{x}_{\sigma,0} = \mathrm{diam}(K)^{-1}\mathbf{x}_\sigma$. Notice also that, if $\rho_{K,0}$ is the supremum of the radius of the balls included in K_0, then

$$\frac{1}{\rho_{K,0}} = \frac{\mathrm{diam}(K_0)}{\rho_{K,0}} = \frac{\mathrm{diam}(K)}{\rho_K} \le \mathrm{regul}(\mathcal{D})^{1/d} \le \theta^{1/d}. \tag{47}$$

Let $A_{K,0}$ be the $(d+1)\times(d+1)$ matrix corresponding to K_0, that is to say whose columns are $(1, \mathbf{x}_{\sigma,0} - \mathbf{x}_{K,0})^T_{\sigma\in\mathcal{E}_K} = (1, \mathrm{diam}(K)^{-1}(\mathbf{x}_\sigma - \mathbf{x}_K))^T_{\sigma\in\mathcal{E}_K}$. Since

$$A_K = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & \mathrm{diam}(K) & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \mathrm{diam}(K) \end{pmatrix} A_{K,0},$$

we have $||A_K^{-1}|| \le \sup(1, \mathrm{diam}(K)^{-1})||A_{K,0}^{-1}||$. Hence, an estimate on $||A_{K,0}^{-1}||$ gives an estimate on $||A_K^{-1}||$.

Step 2: estimate on $A_{K,0}$.
By (47), K_0 contains a closed ball of radius $\frac{1}{2}\theta^{-1/d}$. Up to a translation (which does not change the vectors $\mathbf{x}_{\sigma,0} - \mathbf{x}_{K,0}$, and hence does not change $A_{K,0}$), we can assume that this ball is centered at 0. Since $\mathrm{diam}(K_0)=1$, we have then $\overline{B}(0, \frac{1}{2}\theta^{-1/d}) \subset K_0 \subset \overline{B}(0,1)$.
Let Z_θ be the set of couples $(L, \mathbf{x}_L)$, where L is a simplex such that $\overline{B}(0, \frac{1}{2}\theta^{-1/d}) \subset \overline{L} \subset \overline{B}(0,1)$ and $x_L \in \overline{L}$. Each simplex is defined by $d+1$ vertices in $\mathbb{R}^d$ so Z_θ can be considered as a subset of $P = (\mathbb{R}^d)^{d+1}/S_{d+1} \times \mathbb{R}^d$, where S_{d+1} is the symmetric group acting on $(\mathbb{R}^d)^{d+1}$ by permuting the vectors. As such, Z_θ is compact in P: it is straightforward if we express the condition "the adherence of a simplex contains $\overline{B}(0, \frac{1}{2}\theta^{-1/d})$" as "any point of $\overline{B}(0, \frac{1}{2}\theta^{-1/d})$ is a convex combination of the vertices of the simplex", which is a closed condition with respect to the vertices of the simplex.
For $(L, \mathbf{x}_L) \in Z_\theta$, let $M(L, \mathbf{x}_L)$ be the set of $(d+1)\times(d+1)$ matrices whose columns are, up to permutations, $(1, \mathbf{x}_\sigma - \mathbf{x}_L)^T_{\sigma\in\mathcal{E}_L}$ ($\mathcal{E}_L$ being the set of edges of L and $\mathbf{x}_\sigma$ being the barycenter of σ). $M(L, \mathbf{x}_L)$ can be considered as an element of $M_{d+1}(\mathbb{R})/S_{d+1}$ (S_{d+1} acting by permuting the columns) and the application $(L, \mathbf{x}_L) \in Z_\theta \to M(L, \mathbf{x}_L) \in M_{d+1}(\mathbb{R})/S_{d+1}$ is continuous: to see this, just recall that the barycenter of an edge $\sigma \in \mathcal{E}_L$ is $\mathbf{x}_\sigma = \frac{1}{d}\sum_{i=1}^d \mathbf{x}_i$, where $\mathbf{x}_i$ are the vertices of σ (i.e. all vertices but one of L).

If $(L,\mathbf{x}_L)\in Z_\theta$, all the matrices of $M(L,\mathbf{x}_L)$ are invertible. Indeed, assume that such a matrix has a non-trivial element $(\lambda_1,\dots,\lambda_{d+1})$ in its kernel; this leads (denoting $(\sigma_1,\dots,\sigma_{d+1})$ the edges of L) to $\sum_{i=1}^{d+1}\lambda_i=0$ and $\sum_{i=1}^{d+1}\lambda_i(\mathbf{x}_{\sigma_i}-\mathbf{x}_L)=\sum_{i=1}^{d+1}\lambda_i\mathbf{x}_{\sigma_i}=0$. Assuming $\lambda_{d+1}\neq 0$, we then can write $\mathbf{x}_{\sigma_{d+1}}=\sum_{i=1}^{d}\mu_i\mathbf{x}_{\sigma_i}$ with $\sum_{i=1}^{d}\mu_i=1$ (since $\mu_i=-\lambda_i/\lambda_{d+1}$). This means that $\mathbf{x}_{\sigma_{d+1}}$ is in the affine hyperplane $\mathcal{H}$ generated by the other barycenters of edges. Note that $\mathcal{H}$ is parallel to σ_{d+1} (this is a straightforward consequence of Thales' theorem at the vertex which does not belong to σ_{d+1}, and of the fact that the barycenters $(\mathbf{x}_{\sigma_1},\dots,\mathbf{x}_{\sigma_d})$ of the edges are in fact the barycenters of the vertices of the corresponding edge). Therefore $\mathcal{H}$ contains the whole edge σ_{d+1}, because it contains $\mathbf{x}_{\sigma_{d+1}}\in\sigma_{d+1}$. Let $\mathbf{a}$ be the vertex of L which does not belong to σ_{d+1}; $\mathbf{a}$ belongs to σ_1 and we denote $(\mathbf{b}_1,\dots,\mathbf{b}_{d-1})$ the other vertices of σ_1 (which also belong to $\sigma_{d+1}\subset\mathcal{H}$). We have $\mathbf{x}_{\sigma_1}=\frac{1}{d}(\mathbf{a}+\sum_{i=1}^{d-1}\mathbf{b}_i)$, and therefore $\mathbf{a}=d\mathbf{x}_{\sigma_1}-\sum_{i=1}^{d-1}\mathbf{b}_i$; but $d-\sum_{i=1}^{d-1}1=1$ and thus $\mathbf{a}$ belongs to the affine hyperplane generated by $(\mathbf{x}_{\sigma_1},\mathbf{b}_1,\dots,\mathbf{b}_{d-1})$. Since all these points belong to $\mathcal{H}$, we have $\mathbf{a}\in\mathcal{H}$ and, since $\sigma_{d+1}\subset\mathcal{H}$, all the vertices of L in fact belong to $\mathcal{H}$; L is thus contained in an hyperplane, which is a contradiction with the fact that it contains a non-trivial ball. Thus, for $(L,\mathbf{x}_L)\in Z_\theta$, $M(L,\mathbf{x}_L)$ is in fact an element of $Gl_{d+1}(\mathbb{R})/S_{d+1}$.
The inversion $\mathrm{inv}:Gl_{d+1}(\mathbb{R})\to Gl_{d+1}(\mathbb{R})$ is continuous; hence, $||\mathrm{inv}(\cdot)||:Gl_{d+1}(\mathbb{R})\to\mathbb{R}$ is also continuous. Permuting the columns of a matrix comes down to permuting the lines of its inverse, which does not change the norm; therefore $||\mathrm{inv}(\cdot)||:Gl_{d+1}(\mathbb{R})/S_{d+1}\to\mathbb{R}$ is well defined and also continuous.
We can now conclude this step. The application $Z_\theta\to Gl_{d+1}(\mathbb{R})/S_{d+1}\to\mathbb{R}$ defined by $(L,\mathbf{x}_L)\to M(L,\mathbf{x}_L)\to||\mathrm{inv}(M(L,\mathbf{x}_L))||$ is continuous. Since Z_θ is compact, this application is bounded by some C_{23} only depending on d and θ. As $(K_0,\mathbf{x}_{K,0})\in Z_\theta$, this shows that $||A_{K,0}^{-1}||\le C_{23}$.

Step 3: conclusion.
Using the preceding steps, we find $||F_K||\le||A_K^{-1}||\,||E_K||\le C_{23}\sup(1,\mathrm{diam}(K)^{-1})||E_K||$. Hence,

$$\sum_{K\in\mathcal{M}}\mathrm{diam}(K)^{2-d}||F_K||^2\le C_{23}^2\sup(\mathrm{diam}(\Omega)^2,1)\sum_{K\in\mathcal{M}}\mathrm{diam}(K)^{-d}||E_K||^2.$$

But $||E_K||^2\le\mathrm{m}(K)\int_K|f(x)|^2\,\mathrm{d}x+C_{24}\mathrm{m}(K)^2|\mathbf{v}_K|^2$ with C_{24} only depending on Λ. Since $\mathrm{m}(K)\le\omega_d\mathrm{diam}(K)^d$, this concludes the proof of (46). □

Let us now consider ((7),(8),(9),(10)) with $\nu_K=0$; notice that the results of Section 3 still hold in this situation.
Equation (25) gives, if $\nu_K=0$, an estimate on $\mathbf{v}$ in $L^2(\Omega)^d$ which, thanks to Lemma 6.4, translates into an estimate on the fluxes (this estimate replaces the one obtained before thanks to the penalization), provided that the control volumes are simplicial. This gives, as in the penalized case, existence and uniqueness of a solution to the non-penalized mixed finite volume scheme (i.e. ((7),(8),(9),(10)) with $\nu_K=0$). From the estimate on the fluxes, it is straightforward to see that the term T_{11} in the proof of Theorem 2.2 still tends to 0 as $\mathrm{size}(\mathcal{D})\to 0$. Hence, in the case of simplicial control volumes, the solution to the mixed finite volume scheme ((7),(8),(9),(10)) with $\nu_K=0$ still converges toward the weak solution of (1).
It is also quite easy to establish, in this situation, error estimates in the case of smooth data Λ and $\bar{u}$; these estimates are in fact quite better than the ones of Theorem 2.3: we can prove that

$$\|\mathbf{v}-\nabla\bar{u}\|_{L^2(\Omega)^d}\le C_{25}\mathrm{size}(\mathcal{D})\qquad\text{and}\qquad\|u-\bar{u}\|_{L^2(\Omega)}\le C_{25}\mathrm{size}(\mathcal{D}).$$

To obtain such rates of convergence, one must simply bound T_{13} in (35) by using Lemma 6.4 with $F = \widehat{F}$, $\mathbf{v} = \widehat{\mathbf{v}}$ and $f = 0$.

In the particular case where $\mathcal{D}$ is made of simplicial control volumes, and, for all $K \in \mathcal{M}$, $\nu_K = 0$ and $\mathbf{x}_K$ is the center of gravity of K, then the solution $(u, \mathbf{v}, F)$ of ((7),(8),(9),(10)) is also the solution of the following generalization of the expanded mixed finite element scheme [7]: find $(u, \mathbf{v}, \mathbf{w} = \sum_{K\in\mathcal{M}} \sum_{\sigma\in\mathcal{E}_K} F_{K,\sigma}\mathbf{W}_{K,\sigma}) \in H_{\mathcal{D}} \times H^d_{\mathcal{D}} \times RT^0$ (RT^0 denotes here the lowest degree Raviart-Thomas basis $(\mathbf{W}_\sigma)_{\sigma\in\mathcal{E}}$ on the mesh $\mathcal{M}$, such that, choosing for an internal edge $\sigma = K|L$ the orientation from K to L, then $\mathbf{W}_\sigma$ restricted to K is $\mathbf{W}_{K,\sigma}$ and $\mathbf{W}_\sigma$ restricted to L is $-\mathbf{W}_{L,\sigma}$ — note that $\mathbf{w} \in RT^0$ thanks to (8)) such that

$$\int_\Omega \Lambda(x)\mathbf{v}(x) \cdot \mathbf{v}'(x)\,\mathrm{d}x = \int_\Omega \mathbf{w}(x) \cdot \mathbf{v}'(x)\,\mathrm{d}x,\ \forall \mathbf{v}' \in H^d_{\mathcal{D}},$$

which gives (9),

$$\int_\Omega \mathbf{v}(x) \cdot \mathbf{w}'(x)\,\mathrm{d}x + \int_\Omega u(x)\mathrm{div}\mathbf{w}'(x)\,\mathrm{d}x = 0, \forall \mathbf{w}' \in RT^0,$$

which gives (7) with $\nu_K = 0$, and

$$-\int_\Omega u'(x)\mathrm{div}\mathbf{w}(x)\,\mathrm{d}x = \int_\Omega u'(x)f(x)\,\mathrm{d}x,\ \forall u' \in H_{\mathcal{D}},$$

which gives (10). This formulation (an expanded version of [8]) differs from that of [7], in which the restrictions of $\mathbf{v}$ and $\mathbf{w}$ on each control volume must belong to the same space. The proof of convergence of the mixed finite volume scheme therefore gives at the same time that of this particular version of the expanded mixed finite element scheme.

References

[1] I. Aavatsmark, An introduction to multipoint flux approximations for quadrilateral grids. Locally conservative numerical methods for flow in porous media. Comput. Geosci. 6, 405–432 (2002).

[2] I. Aavatsmark, T. Barkve, O. Boe and T. Mannseth, Discretization on unstructured grids for inhomogeneous, anisotropic media. Part I: Derivation of the methods. Journal on Scientific Computing, 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 Scientific Computing, 19, 1717–1736 (1998).

[4] T. Arbogast, L.C. Cowsar, M.F. Wheeler and I. Yotov, Mixed finite element methods on nonmatching multiblock grids. SIAM J. Numer. Anal. 37, No.4, 1295–1315 (2000).

[5] T. Arbogast, M.F. Wheeler and I. Yotov, Mixed finite elements for elliptic problems with tensor coefficients as cell-centered finite differences. SIAM J. Numer. Anal. 34, No.2, 828–852 (1997).

[6] G. Chavent, G. Cohen and J. Jaffré, Discontinuous upwinding and mixed finite elements for two-phase flows in reservoir simulation. Comput. Methods Appl. Mech. Eng. 47, 93–118 (1984).

[7] Z. Chen, Expanded mixed finite element methods for linear second-order elliptic problems. I. RAIRO, Modélisation Math. Anal. Numér. 32, No.4, 479–499 (1998).

Expanded mixed finite element methods for quasilinear second order elliptic problems. II. RAIRO, Modélisation Math. Anal. Numér. 32, No.4, 501–520 (1998).

[8] J-P. Croisille, Finite volume box-schemes and mixed methods, Math. Model. and Numer. Anal., 34, No.5, 1087–1106 (2000).

[9] J. Droniou, Error estimates for the convergence of a finite volume discretization of convection-diffusion equations. J. Numer. Math. 11, 1–32 (2003).

[10] J. Droniou, R. Eymard, D. Hilhorst and X. D. Zhou, Convergence of a finite volume - mixed finite element method for a system of a hyperbolic and an elliptic equations. IMA Journal of Numerical Analysis 23, 507–538 (2003).

[11] R. Eymard, T. Gallouët and R. Herbin, Finite Volume Methods. Handbook of Numerical Analysis, Edited by P.G. Ciarlet and J.L. Lions, North Holland 7, 713–1020 (2000).

[12] R. Eymard, T. Gallouët and R. Herbin, A finite volume for anisotropic diffusion problems. Comptes Rendus de l'Académie des Sciences 339, 299–302 (2004).

[13] R. Eymard, T. Gallouët and R. Herbin, A cell-centred finite-volume approximation for anisotropic diffusion operators on unstructured meshes in any space dimension. IMA J. of Num. Anal. Advance Access (2005). doi:10.1093/imanum/dri036

[14] R. Eymard, T. Gallouët and R. Herbin, Finite volume approximation of elliptic problems and convergence of an approximate gradient. Appl. Num. Math. 37, 31–53 (2001).

[15] I. Faille, A control volume method to solve an elliptic equation on a two- dimensional irregular mesh. Comput. Methods Appl. Mech. Eng. 100, 275–290 (1992).

[16] T. Gallouët, R. Herbin and M.H. Vignal, Error estimate for the approximate finite volume solutions of convection diffusion equations with Dirichlet, Neumann or Fourier boundary conditions, SIAM J. Numer. Anal., **37**, No. 6, 1935–1972 (2000).

[17] D.S. Kershaw, Differencing of the diffusion equation in Lagrangian hydrodynamic codes. J. Comput. Phys. 39, 375–395 (1981).

[18] Y. Kuznetsov and S. Repin, Convergence analysis and error estimates for mixed finite element method on distorted meshes. J. Numer. Math. **13**, No.1, 33–51 (2005).

[19] R.A. Klausen and T.F. Russell, Relationships among some locally conservative discretization methods which handle discontinuous coefficients, Computational Geosciences, 8, 341–377 (2004).

[20] C. Le Potier, A finite volume method for the approximation of highly anisotropic diffusion operators on unstructured meshes, Finite Volumes for Complex Applications IV, Marrakesh, Marocco, (2005).

[21] K. Lipnikov, J. Morel and M. Shashkov, Mimetic finite difference methods for diffusion equations on non-orthogonal non-conformal meshes. (English) J. Comput. Phys. 199, 589–597 (2004).

[22] P.A. Raviart and J.M. Thomas, A mixed finite element method for 2nd order elliptic problems. Math. Aspects Finite Elem. Meth., Proc. Conf. Rome 1975, Lect. Notes Math. 606, 292–315 (1977).

[23] J.E. Roberts and J.M. Thomas, Mixed and hybrid methods. Ciarlet, P. G. (ed.) et al. Handbook of numerical analysis. North-Holland 2, 523–639 (1991).

[24] A. Younes, P. Ackerer and G. Chavent, From mixed finite elements to finite volumes for elliptic PDEs in two and three dimensions. Int. J. Numer. Methods Eng. 59, 365-388 (2004).