

MOX-Report No. 37/2018

**Convergence analysis of a cell centered finite volume
diffusion operator on non-orthogonal polyhedral
meshes**

Bonaventura, L.; Della Rocca A.;

MOX, Dipartimento di Matematica
Politecnico di Milano, Via Bonardi 9 - 20133 Milano (Italy)

mox-dmat@polimi.it

<http://mox.polimi.it>

Convergence analysis of a cell centered finite volume diffusion operator on non-orthogonal polyhedral meshes

Luca Bonaventura⁽¹⁾, Alessandro Della Rocca^{(1),(2)}

June 6, 2018

⁽¹⁾ MOX – Modelling and Scientific Computing,
Dipartimento di Matematica, Politecnico di Milano
Via Bonardi 9, 20133 Milano, Italy
`luca.bonaventura@polimi.it`
`alessandro.dellarocca@polimi.it`

⁽²⁾ Tenova S.p.A.,
Global R&D,
Via Albareto 31, 16153 Genova, Italy
`alessandro.dellarocca@tenova.com`

Keywords: Finite volume, Cell centered methods, Convergence, Diffusion,
Unstructured meshes

AMS Subject Classification: 65M08, 65N08, 65N12, 65Z05, 76R50

Abstract

A simple but successful strategy for building a discrete diffusion operator in finite volume schemes of industrial use is to correct the standard two-point flux approximation with a term accounting for the local mesh non-orthogonality. Practical experience with a variety of different mesh typologies, including non-orthogonal tetrahedral, hexahedral and polyhedral meshes, has shown that this discrete diffusion operator is accurate and robust whenever the mesh is not too distorted and sufficiently regular. In this work, we show that this approach can be interpreted as equivalent to introducing an anisotropic operator that accounts for the preferential directions induced by the local mesh non-orthogonality. This allows to derive a convergence analysis of the corrected method under a quite weak global assumption on mesh distortion. This convergence proof, which is obtained for the first time for this finite volume method widely employed in industrial applications, provides a reference framework on how to interpret some of its variants commonly implemented in commercial finite volume codes. Numerical experiments are presented that confirm the accuracy and robustness of the results. Furthermore, we also show empirically that a least square approach to the gradient computation can provide second order convergence even when the mild non-orthogonality condition on the mesh is violated.

1 Introduction

Finite volume methods have been extremely popular in computational fluid dynamics (CFD) in the past and they still are an area of active research in numerical mathematics. Among the many different developments in this field, we recall the finite volume element scheme [9], the multi-point flux approximation schemes (MPFA) [1], [2], [3], or more recent variants, such as the mixed finite volume scheme (MFV) [16], the hybrid finite volume scheme (HFV) [23], [25] and the discrete duality finite volume schemes (DDFV) [4],[5], [11], [12], [32], [33], [34]. All these methods share indeed many common features, as discussed in [18].

In this work, however, we will focus on cell centered schemes, in which a single unknown is associated to each mesh cell. Cell centered finite volume methods are widely employed in industrial codes [7], [40] for a number of practical reasons. Indeed, they rely on relatively simple data structures, even for general unstructured meshes, and they allow for easy treatment of boundary conditions at singular boundary points, such as inner corners, while effectively handling general shapes of the computational domain. Cell centered methods can be naturally parallelized by domain decomposition techniques, guaranteeing minimal interprocessor communications, especially in their low order variants, due to the use of discrete operators built from local stencils. They allow an easy implementation of locally adaptive multilevel refinement strategies and they can be easily equipped with very efficient geometric multigrid procedures [43]. Finally, cell centered finite volume methods also allow an immediate extension to nonlinear coupled problems [24].

Known drawbacks of cell centered schemes are the reduced accuracy in strongly heterogeneous diffusion problems [25] with respect to MPFA, MFV or HFV schemes, as well as the only asymptotical recovery of the discrete Stokes formula, in contrast with the exact discrete property provided for example by DDFV schemes [11]. On the other hand, MPFA, MFV, HFV and DDFV schemes achieve such properties by introducing additional unknowns at selected mesh locations, thus implying an additional cost with respect to cell centered discretizations. It is still an open question if similar accuracy improvements can be obtained from cell centered schemes by introducing additional unknowns through local mesh refinement.

For these reasons, it is important to understand the analytical behaviour of cell centered finite volume discretizations on the typical non-orthogonal meshes practically required for industrial applications [7], [40]. For these applications, the so-called *Gauss cor-*

rected scheme, widely adopted by finite volume practitioners [30], [36], [37], [39], [42], appears to be a simple, robust and sufficiently accurate option. Notice that this scheme can also be interpreted as a specific realization of the recently introduced asymmetric gradient discretization method [17].

To the best of the authors' knowledge, the convergence properties of this finite volume method have never been analyzed in the case of non-orthogonal meshes. Indeed, convergence analyses of finite volume schemes for diffusion operators on unstructured mesh types are usually limited to polyhedral meshes satisfying an orthogonality condition [21], [22]. This is quite restrictive in practice, since none of the robust mesh generators usually adopted for pre-processing of industrial configurations are able to guarantee this condition.

In this work, we show that it is possible to prove the convergence of the *Gauss corrected* scheme on unstructured meshes satisfying a global and rather weak mesh regularity condition. This goal is achieved adapting the approach used in [22] for the convergence analysis of a cell-centered finite volume scheme for anisotropic diffusion problems on orthogonal meshes. A preliminary version of these results has been presented in [14]. Furthermore, we also show empirically that a least square approach to the gradient computation can provide second order convergence even when the mild non-orthogonality condition on the mesh is violated. It is to be remarked that existing convergence proofs for finite volume methods on non-orthogonal meshes either involve discretization schemes not guaranteeing local flux conservativity [24], [25], or DDFV schemes employing additional degrees of freedom [4], [5], or two-dimensional diamond schemes on meshes satisfying more restrictive regularity conditions [13]. We will focus here on the isotropic steady state diffusion equation

$$-\operatorname{div}(\alpha \nabla \underline{u}) = f, \quad \text{in } \Omega, \quad (1a)$$

$$\underline{u} = 0, \quad \text{on } \partial\Omega. \quad (1b)$$

We will assume that $\alpha : \Omega \rightarrow \mathbb{R}$ is a measurable function, $\alpha \in L^\infty(\Omega)$, such that $0 < \alpha_0 \leq \alpha(\mathbf{x})$ for a.e. $\mathbf{x} \in \mathbb{R}^d$, with $\alpha_0 \in \mathbb{R}$, and for $f \in L^2(\Omega)$. The classical weak formulation of problem (1) consists in finding $\underline{u} \in H_0^1(\Omega)$ such that

$$\int_{\Omega} \alpha(\mathbf{x}) \nabla \underline{u}(\mathbf{x}) \cdot \nabla v(\mathbf{x}) \, d\mathbf{x} = \int_{\Omega} f(\mathbf{x}) v(\mathbf{x}) \, d\mathbf{x}, \quad \forall v \in H_0^1(\Omega). \quad (2)$$

Rather than proving convergence directly for the finite volume scheme associated to the strong problem formulation (1), we will identify a

discrete weak formulation underlying the finite volume scheme and then to prove convergence of its solution to that of the continuous weak problem (26).

The paper is organized as follows. In section 2, several fundamental definitions of mesh related quantities and discrete functional spaces are introduced. In section 3, the cell centered finite volume method that is the focus of our analysis is presented. In section 4, the discrete weak formulation is recovered and in section 5, the convergence analysis of the *Gauss corrected* scheme is presented. In section 6, the results of some numerical experiments are reported. A proposal to overcome the constraints on the mesh for some specific three-dimensional mesh types is introduced in section 7. Finally, in section 8 some conclusions are drawn and some future developments are outlined.

2 Meshes and discrete spaces

The finite volume method is a mesh-based discretization technique suitable for any number of space dimensions, but in this work we only consider the $d = 3$ case. Since the computational domains of practical interest are usually of complex geometry, the focus here is on meshes composed of arbitrarily shaped polyhedral cells, in the sense of the formal definition below, see also [19], [25].

Definition 2.1. (*Polyhedral mesh*): Let Ω be a bounded, open polyhedral subset of \mathbb{R}^d . A polyhedral mesh for Ω is denoted by $\mathcal{D} = (\mathcal{M}, \mathcal{F}, \mathcal{P}, \mathcal{V})$, where the quadruple includes:

1. \mathcal{M} is a finite family of non-empty, connected, polyhedral, open, disjoint subsets of Ω called cells (or control volumes), such that $\overline{\Omega} = \cup_{K \in \mathcal{M}} \overline{K}$. For any $K \in \mathcal{M}$, $\partial K = \overline{K} \setminus K$ is the boundary of K , $|K| > 0$ denotes the measure of K , and $h_K = \text{diam}(K)$ is the diameter of K , that is the maximum distance between two points in K .
2. $\mathcal{F} = \mathcal{F}_{int} \cup \mathcal{F}_{ext}$ is a finite family of disjoint subsets of $\overline{\Omega}$ representing the faces. Let \mathcal{F}_{int} be the set of interior faces such that, for all $\sigma \in \mathcal{F}_{int}$, σ is a non-empty open subset of a hyperplane in \mathbb{R}^d with $\sigma \subset \Omega$, and let \mathcal{F}_{ext} be the set of boundary faces such that, for all $\sigma \in \mathcal{F}_{ext}$, σ is a non-empty open subset of $\partial\Omega$. It is assumed that, for any $K \in \mathcal{M}$, there exists a subset $\mathcal{F}_K \subset \mathcal{F}$ such that $\partial K = \cup_{\sigma \in \mathcal{F}_K} \sigma$. The set of cells sharing one face σ is $\mathcal{M}_\sigma = \{K \in \mathcal{M} : \sigma \in \mathcal{F}_K\}$. It is assumed that, for all $\sigma \in \mathcal{F}$, either \mathcal{M}_σ has exactly two

elements and then $\sigma \in \mathcal{F}_{int}$, or \mathcal{M}_K has exactly one element and then $\sigma \in \mathcal{F}_{ext}$. For all $\sigma \in \mathcal{F}$, $|\sigma| > 0$ denotes the $(d-1)$ -dimensional measure of σ , and $\bar{\mathbf{x}}_\sigma$ is the barycenter of σ .

3. $\mathcal{P} = (\mathbf{x}_K)_{K \in \mathcal{M}}$ is a family of points of Ω indexed by \mathcal{M} , such that for all $K \in \mathcal{M}$, $\mathbf{x}_K \in K$ and it is called the center of K , possibly corresponding to its barycenter. It is assumed that all cells $K \in \mathcal{M}$ are \mathbf{x}_K -star-shaped, in the sense that if $\mathbf{x} \in K$, then the line segment $[\mathbf{x}_K, \mathbf{x}] \subset K$.
4. \mathcal{V} is the finite set of vertices of the mesh. For $K \in \mathcal{M}$, \mathcal{V}_K collects all the vertices belonging to \bar{K} , while for $\sigma \in \mathcal{F}$, \mathcal{V}_σ collects all the vertices belonging to σ .

The size of the polyhedral mesh is defined as $h_{\mathcal{D}} = \sup\{h_K, K \in \mathcal{M}\}$.

Furthermore, for any $K \in \mathcal{M}$ and for any $\sigma \in \mathcal{F}_K$, $\mathbf{n}_{K,\sigma}$ is the constant unit vector normal to σ and outward to K . For any $K \in \mathcal{M}$, the set of neighbors of K is denoted by

$$\mathcal{N}_K = \{L \in \mathcal{M} \setminus \{K\}, \exists \sigma \in \mathcal{F}_{int}, \mathcal{M}_\sigma = \{K, L\}\}. \quad (3)$$

Additionally $d_{K,\sigma}$ denotes the orthogonal distance between \mathbf{x}_K and $\sigma \in \mathcal{F}_K$

$$d_{K,\sigma} = (\mathbf{x} - \mathbf{x}_K) \cdot \mathbf{n}_{K,\sigma}, \quad (4)$$

which is constant for all $\mathbf{x} \in \sigma$. From the assumption that K is \mathbf{x}_K -star-shaped, it follows that $d_{K,\sigma} > 0$ and that it also holds:

$$\sum_{\sigma \in \mathcal{F}_K} |\sigma| d_{K,\sigma} = d |K| \quad \forall K \in \mathcal{M}. \quad (5)$$

For all $K \in \mathcal{M}$ and $\sigma \in \mathcal{F}_K$, $D_{K,\sigma}$ denotes the cone with vertex \mathbf{x}_K and basis σ , also called half-diamond, that is the volume defined by

$$D_{K,\sigma} = \{t \mathbf{x}_K + (1-t) \mathbf{y}, t \in (0,1), \mathbf{y} \in \sigma\}. \quad (6)$$

For all $\sigma \in \mathcal{F}$, $D_\sigma = \cup_{K \in \mathcal{M}_\sigma} D_{K,\sigma}$ denotes the diamond associated to face σ , as in Figure 1.

Definition 2.1 covers a wide range of meshes, including meshes with non-convex cells, with non-planar faces requiring triangulation, or with hanging nodes. Furthermore, Definition 2.1 also includes tetrahedral and hexahedral meshes as particular cases, as well as meshes with wedge and pyramidal cells.

Finite volume methods are traditionally introduced in discrete functional spaces of piecewise constant functions [21]. In recent analyses [27], [28], associated inner products, norms and seminorms

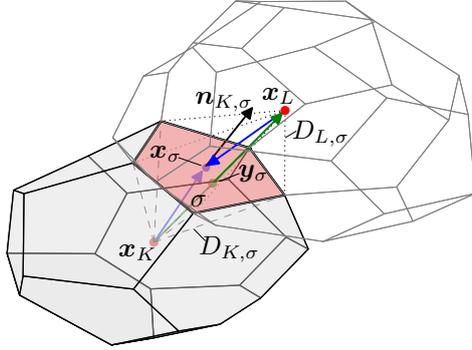


Figure 1: Non-orthogonal generally polyhedral mesh: the line segment $(\mathbf{x}_L - \mathbf{x}_K)$ is not aligned with the face normal unit vector $\mathbf{n}_{K,\sigma}$. Furthermore, the intersection point \mathbf{y}_σ between the face σ and the vector $(\mathbf{x}_L - \mathbf{x}_K)$ is not necessarily coincident with the face centroid \mathbf{x}_σ . The half diamonds $D_{K,\sigma}$ and $D_{L,\sigma}$ are represented with dashed and dotted lines, respectively.

are exploited to recast the discrete flux balance equations into an equivalent variational form, which naturally allows to derive stability estimates and to investigate the numerical convergence of specific schemes [25]. In the classical finite volume framework, the discrete flux balance equation corresponding to problem (1) takes the form

$$\sum_{\sigma \in \mathcal{F}_K} F_{K,\sigma}(u) = \int_K f(x) dx \quad \forall K \in \mathcal{M} \quad (7)$$

where the face flux is such that

$$F_{K,\sigma} \approx - \int_{\sigma} \alpha(x) \nabla u(x) \cdot \mathbf{n}_{K,\sigma} d\gamma(x)$$

and $d\gamma(x)$ denotes the infinitesimal face area element. A relevant feature of the scheme is the flux conservativity property

$$F_{K,\sigma}(u) + F_{L,\sigma}(u) = 0, \quad (8)$$

which is assumed to hold for all interior faces $\sigma \in \mathcal{F}_{int}$, where K and L are the cells sharing the face σ .

The convergence analysis of cell centered finite volume schemes on arbitrary polyhedral meshes [25] may also require to introduce the space $H_{\mathcal{D}}(\Omega) \subset L^p(\Omega)$, which consists of the functions that are piecewise constant on each cell $K \in \mathcal{M}$. For all $v \in H_{\mathcal{D}}(\Omega)$ and for all $K \in \mathcal{M}$, the constant value of v in K is denoted by v_K . Consequently, discrete functional analysis results for the convergence of finite volume schemes [21], [25] can be exploited.

In addition, in order to introduce proper test functions to check the convergence of the discrete solution to the continuous solution of the weak formulation, for all functions $\psi \in C(\Omega)$ a projection operator $P_{\mathcal{D}} : C(\Omega) \rightarrow H_{\mathcal{D}}(\Omega)$ is defined, such that $P_{\mathcal{D}}\psi = (\psi(\mathbf{x}_K))_{K \in \mathcal{M}}$.

3 A cell centered diffusion scheme for non-orthogonal meshes

The vast majority of finite volume schemes for diffusive problems are based on the application of the discrete Gauss theorem. The numerical approximation is derived as

$$\int_K \nabla \cdot (\alpha \nabla u) \, d\mathbf{x} = \sum_{\sigma \in \mathcal{F}_K} \int_{\sigma} \alpha \mathbf{n}_{K,\sigma} \cdot \nabla u \, d\gamma \approx \sum_{\sigma \in \mathcal{F}_K} F_{K,\sigma}^{(d)}(u) \quad (9)$$

where the numerical flux through face σ is computed as

$$F_{K,\sigma}^{(d)}(u) = |\sigma| \alpha_{K,\sigma} \mathbf{n}_{K,\sigma} \cdot \nabla_{K,\sigma} u, \quad \forall \sigma \in \mathcal{F}_K \quad (10)$$

and depends on the definition of the the face normal gradient. Usually, $\alpha_{K,\sigma}$ is approximated by the surface value interpolation $I_{\sigma}\alpha$, obtained from standard interpolation schemes. Linear interpolation is often chosen to preserve second order accuracy, while harmonic interpolation is sometimes selected, especially when the scalar diffusivity field α is strongly non-homogeneous [21]. A variety of alternative schemes can be constructed to approximate the face normal gradient $\mathbf{n}_{K,\sigma} \cdot \nabla_{K,\sigma} u$, each with its own specific features. Most of them are traditionally studied empirically, by directly testing them on specific meshes and representative flow problems [29], [41].

The simplest scheme for the face normal gradient is represented by the two-point flux approximation [21]

$$\mathbf{n}_{K,\sigma} \cdot \nabla_{K,\sigma} u = \frac{u_L - u_K}{d_{K,\sigma} + d_{L,\sigma}}. \quad (11)$$

Even though unconditionally monotone and coercive [15], it is of limited accuracy on unstructured meshes, where mesh non-orthogonality may lead to severe errors in the approximation of the diffusion fluxes [20]. In order to compensate for the unavoidable non-orthogonality of realistic unstructured meshes, a simple but effective solution is provided by the *Gauss corrected* scheme, which introduces a non-orthogonal correction term [36] in the two-point flux scheme, thus

obtaining the approximation

$$\begin{aligned} \mathbf{n}_{K,\sigma} \cdot \nabla_{K,\sigma} u &= \frac{u_L - u_K}{d_{K,\sigma} + d_{L,\sigma}} \\ &+ \left(\mathbf{n}_{K,\sigma} - \frac{\mathbf{x}_L - \mathbf{x}_K}{\mathbf{n}_{K,\sigma} \cdot (\mathbf{x}_L - \mathbf{x}_K)} \right) \cdot \nabla_\sigma u. \end{aligned} \quad (12)$$

Here, the first term corresponds to the two-point flux contribution in Eq.(11), expressing the diffusion flux component in the direction of the line segment $(\mathbf{x}_L - \mathbf{x}_K)$, while the second term accounts for the local mesh non-orthogonality across the face σ , expressed as the difference between the correct face normal diffusive flux estimated from a proper face gradient and the diffusive flux along the direction of $(\mathbf{x}_L - \mathbf{x}_K)$. In order to avoid oscillatory solutions [41], it is important that the gradient $\nabla_\sigma u$ at face σ is evaluated using a different scheme from the one employed in the first term of Eq.(12). Thus, the gradient $\nabla_\sigma u$ is usually estimated at face σ by interpolation of the neighbouring cells gradients $\nabla_K u$ and $\nabla_L u$ for $\mathcal{M}_\sigma = \{K, L\}$, which can be either the standard linear interpolation or, for increased simplicity, the midpoint rule. Indeed, if the cell derivatives are linear approximations, the diffusion flux will be more accurate than first order on very regular meshes [36]. The *Gauss corrected* scheme allows more accurate approximations than the two-point flux approximation (11), but it is not, in general, unconditionally coercive on arbitrary unstructured meshes. As a consequence, on irregular meshes it may become a source of numerical instability. On orthogonal grids, this scheme reduces to the classical two-point flux scheme, since the correction term vanishes.

Following [10], on a general unstructured polyhedral mesh like that of Definition 2.1, the centered discrete gradient operator $\nabla_{\mathcal{D}} : H_{\mathcal{D}}(\Omega) \rightarrow H_{\mathcal{D}}(\Omega)^d$ is defined as the piecewise constant function

$$\nabla_K u = \frac{1}{|K|} \sum_{\sigma \in \mathcal{F}_K} |\sigma| (I_\sigma u - u_K) \mathbf{n}_{K,\sigma} \quad (13)$$

for $u \in H_{\mathcal{D}}(\Omega)$. Since for any closed control volume the geometrical relations

$$\sum_{\sigma \in \mathcal{F}_K} |\sigma| \mathbf{n}_{K,\sigma} \cdot \mathbf{e}^{(i)} = \sum_{\sigma \in \mathcal{F}_K} |\sigma| n_{K,\sigma}^{(i)} = 0 \quad i = 1, \dots, d \quad (14)$$

hold, with $\mathbf{n}_{K,\sigma} = n_{K,\sigma}^{(i)} \mathbf{e}^{(i)}$, Eq.(13) is also equal to

$$\nabla_K u = \frac{1}{|K|} \sum_{\sigma \in \mathcal{F}_K} |\sigma| I_\sigma u \mathbf{n}_{K,\sigma}, \quad (15)$$

which is easily recognized as the finite volume discretization of the gradient based on the Gauss theorem [30]. For this reason, the gradient approximation in Eq.(13) is often identified as the Gauss gradient scheme.

The consistency of the discrete gradient in Eq.(13) has been analyzed in [25]. It stems directly from the geometrical identity

$$\sum_{\sigma \in \mathcal{F}_K} |\sigma| \mathbf{n}_{K,\sigma} (\mathbf{x}_\sigma - \mathbf{x}_K)^\top = |K| \mathbf{I}, \quad \forall K \in \mathcal{M}, \quad (16)$$

where $(\mathbf{x}_\sigma - \mathbf{x}_K)^\top$ is the transpose of the vector $(\mathbf{x}_\sigma - \mathbf{x}_K) \in \mathbb{R}^d$, see Figure 1, and $\mathbf{I} \in \mathbb{R}^d \times \mathbb{R}^d$ is the identity matrix. For any affine function $\psi : \Omega \rightarrow \mathbb{R}$ defined by $\psi(\mathbf{x}) = \mathbf{g} \cdot \mathbf{x} + c$, with $\mathbf{g} \in \mathbb{R}^d$ and $c \in \mathbb{R}$, assuming that $u_\sigma = \psi(\mathbf{x}_\sigma)$ and $u_K = \psi(\mathbf{x}_K)$, it results that $u_\sigma - u_K = (\mathbf{x}_\sigma - \mathbf{x}_K)^\top \mathbf{g} = (\mathbf{x}_\sigma - \mathbf{x}_K)^\top \nabla \psi$. Hence, expression (13) leads to $\nabla_K u = \nabla \psi$, which amounts to linear exactness for any affine function ψ on $K \in \mathcal{M}$, provided that $I_\sigma u = u_\sigma$, which is verified whenever $\mathbf{x}_\sigma = \mathbf{y}_\sigma$, $\forall \sigma \in \mathcal{F}_{int}$, see Figure 1.

Finally, if the face gradient $\nabla_\sigma u$ in Eq.(12) is computed using a linear interpolation operator applied to the cell gradients reconstructed via the Gauss scheme (15) from both cells sharing the face σ , the non-orthogonal correction term in Eq.(12) is associated to a large stencil which includes, besides cells K and L sharing face σ , all their neighbouring cells $M \in \mathcal{N}_K \cup \mathcal{N}_L$.

By applying the Gauss corrected scheme from Eq.(12) to the diffusion problem (1), one obtains the finite volume scheme

$$\sum_{L \in \mathcal{N}_K} F_{K,L} + \sum_{\sigma \in \mathcal{F}_{K,ext}} F_{K,\sigma} = \int_K f(\mathbf{x}) d\mathbf{x}, \quad \forall K \in \mathcal{M} \quad (17)$$

where the diffusive fluxes $F_{K,L} = -F_{K,\sigma}^{(d)}$ take the forms

$$F_{K,L} = \alpha_{K|L} \frac{|\sigma|}{d_{K,L}} (u_K - u_L) - \alpha_{K|L} |\sigma| \left(\mathbf{n}_{K,\sigma} - \frac{\mathbf{i}_{K,L}}{\mathbf{n}_{K,\sigma} \cdot \mathbf{i}_{K,L}} \right) \cdot \nabla_\sigma u, \quad \forall K|L \in \mathcal{F}_{int} \quad (18a)$$

$$F_{K,\sigma} = \alpha_{K,\sigma} \frac{|\sigma|}{d_{K,\sigma}} u_K - \alpha_{K,\sigma} |\sigma| \left(\mathbf{n}_{K,\sigma} - \frac{\mathbf{i}_{K,\sigma}}{\mathbf{n}_{K,\sigma} \cdot \mathbf{i}_{K,\sigma}} \right) \cdot \nabla_\sigma u, \quad \forall \sigma \in \mathcal{F}_{K,ext} \quad (18b)$$

with the shorthand notation $d_{K,L} = d_{K,\sigma} + d_{L,\sigma}$ and using the unit

vectors

$$\mathbf{i}_{K,L} = \frac{\mathbf{x}_L - \mathbf{x}_K}{|\mathbf{x}_L - \mathbf{x}_K|}, \quad \forall K|L \in \mathcal{F}_{int} \quad (19a)$$

$$\mathbf{i}_{K,\sigma} = \frac{\mathbf{x}_\sigma - \mathbf{x}_K}{|\mathbf{x}_\sigma - \mathbf{x}_K|}, \quad \forall \sigma \in \mathcal{F}_{K,ext}. \quad (19b)$$

The diffusivity in Eq.(18) is defined as

$$\alpha_{K|L} = \frac{1}{|D_\sigma|} \int_{D_\sigma} \alpha(\mathbf{x}) \, d\mathbf{x}, \quad \forall K|L \in \mathcal{F}_{int} \quad (20a)$$

$$\alpha_{K,\sigma} = \frac{1}{|D_{K,\sigma}|} \int_{D_{K,\sigma}} \alpha(\mathbf{x}) \, d\mathbf{x}, \quad \forall \sigma \in \mathcal{F}_{K,ext}, \quad (20b)$$

which define piecewise constant functions over the diamond cells D_σ and $D_{K,\sigma}$ dual to internal and external mesh faces respectively. Finally, it is important to notice that the fluxes (18) are locally conservative, since

$$F_{K,L} = -F_{L,K}, \quad \forall K|L \in \mathcal{F}_{int}. \quad (21)$$

In the definition of the fluxes, a reconstruction of the face gradient $\nabla_\sigma u$ must be employed. For this purpose, a linear interpolation operator is selected at internal faces

$$\nabla_\sigma u = I_\sigma \nabla u = \frac{d_{L,\sigma}}{d_{K,L}} \nabla_K u + \frac{d_{K,\sigma}}{d_{K,L}} \nabla_L u, \quad \forall \sigma \in \mathcal{F}_{int}, \quad (22)$$

while at boundary faces the simplest choice is $\nabla_\sigma u = \nabla_K u$, for all $\sigma \in \mathcal{F}_{K,ext}$. Here, we will use the approximation

$$\nabla_\sigma u = -\frac{u_K}{d_{K,\sigma}} \mathbf{n}_{K,\sigma} + (\nabla_K u - (\mathbf{n}_{K,\sigma} \cdot \nabla_K u) \mathbf{n}_{K,\sigma}), \quad \forall \sigma \in \mathcal{F}_{ext},$$

in order to recover $\mathbf{n}_{K,\sigma} \cdot \nabla_\sigma u = -u_K/d_{K,\sigma}$ at boundaries. The linear interpolation makes use of (15) with linear interpolation of the face values

$$I_\sigma u = \frac{d_{L,\sigma}}{d_{K,L}} u_K + \frac{d_{K,\sigma}}{d_{K,L}} u_L, \quad \forall \sigma \in \mathcal{F}_{int} \quad (23)$$

while the boundary face values follow directly from the homogeneous Dirichlet conditions in problem (1). Additionally, in practical implementations it is customary to compute the scalar diffusivity $\alpha_{K|L}$ at internal faces $K|L \in \mathcal{F}_{int}$ by a linear interpolation operator as $\alpha_{K|L} = I_\sigma \alpha$.

To allow for the treatment of non-orthogonal polyhedral meshes, it is useful to consider the associated isotropic diffusion problem

$$\Gamma_\alpha = \alpha \mathbf{I}, \quad (24)$$

where Γ_α is isotropic diffusivity tensor associated to the scalar diffusivity α . This allows to reformulate problem (1) as

$$-\operatorname{div}(\Gamma_\alpha \nabla \underline{u}) = f, \quad \text{in } \Omega, \quad (25a)$$

$$\underline{u} = 0, \quad \text{on } \partial\Omega \quad (25b)$$

with $\Gamma_\alpha(\mathbf{x})$ naturally verifying the usual assumptions [22]. Similarly, the associated weak formulation is given by

$$\begin{aligned} \underline{u} &\in H_0^1(\Omega), \\ \int_\Omega \Gamma_\alpha(\mathbf{x}) \nabla \underline{u}(\mathbf{x}) \cdot \nabla v(\mathbf{x}) \, d\mathbf{x} &= \int_\Omega f(x) v(x) \, d\mathbf{x}, \quad (26) \\ \forall v &\in H_0^1(\Omega). \end{aligned}$$

It is possible to derive a finite volume scheme for diffusion problems with tensorial diffusivity by constructing a local discrete gradient [22], in order to obtain at cell face σ a consistent approximation of the diffusive flux $-\int_\sigma (\Gamma_\alpha(\mathbf{x}) \nabla \underline{u}(\mathbf{x})) \cdot \mathbf{n}_\sigma \, d\gamma(\mathbf{x})$, with usual notation for finite volume schemes. To this purpose, it is beneficial to rewrite the diffusive flux for an internal face $K|L \in \mathcal{F}_{int}$ using the diffusivity tensor from Eq.(24). Since Γ_α is symmetric, it follows that

$$F_{K,L} = -|\sigma| (\Gamma_{K,L} \nabla_{K,L} u) \cdot \mathbf{n}_{K,\sigma} = -|\sigma| \nabla_{K,L} u \cdot (\Gamma_{K,L} \mathbf{n}_{K,\sigma}). \quad (27)$$

In order to allow for the treatment of non-orthogonal meshes, the following diffusivity tensor decomposition can be applied

$$\Gamma_{K,L} = \alpha_{K,L} \mathbf{I} = \Gamma_{K,L}^\parallel + \Gamma_{K,L}^\# \quad (28)$$

with anisotropic (directional) diffusivity tensors

$$\Gamma_{K,L}^\parallel = \alpha_{K,L} \frac{1}{\left(\mathbf{i}_{K,L}^\top \mathbf{n}_{K,\sigma}\right)^2} \mathbf{i}_{K,L} \mathbf{i}_{K,L}^\top \quad (29a)$$

$$\Gamma_{K,L}^\# = \alpha_{K,L} \left(\mathbf{I} - \frac{1}{\left(\mathbf{i}_{K,L}^\top \mathbf{n}_{K,\sigma}\right)^2} \mathbf{i}_{K,L} \mathbf{i}_{K,L}^\top \right) \quad (29b)$$

by following the natural directions locally identified from the non-orthogonal polyhedral mesh. Notice also that both diffusivity tensors are symmetric, since $\Gamma_{K,L}^\parallel = (\Gamma_{K,L}^\parallel)^\top$ and $\Gamma_{K,L}^\# = (\Gamma_{K,L}^\#)^\top$, and that

$$\Gamma_{K,L}^\parallel = \Gamma_{L,K}^\parallel \quad \text{and} \quad \Gamma_{K,L}^\# = \Gamma_{L,K}^\#. \quad (30)$$

A similar flux decomposition can be carried out at boundary faces $\sigma \in \mathcal{F}_{ext}$ by substituting the unit vector $\mathbf{i}_{K,L}$ with $\mathbf{i}_{K,\sigma}$.

By substituting the tensor decomposition from Eqs.(28)-(29) into the finite volume fluxes (27), one obtains that

$$\begin{aligned}\Gamma_{K,L}\mathbf{n}_{K,\sigma} &= \Gamma_{K,L}^{\parallel}\mathbf{n}_{K,\sigma} + \Gamma_{K,L}^{\#}\mathbf{n}_{K,\sigma} \\ &= \alpha_{K,L} \frac{1}{\mathbf{i}_{K,L}^{\top}\mathbf{n}_{K,\sigma}} \mathbf{i}_{K,L} \\ &\quad + \alpha_{K,L} \left(\mathbf{n}_{K,\sigma} - \frac{1}{\mathbf{i}_{K,L}^{\top}\mathbf{n}_{K,\sigma}} \mathbf{i}_{K,L} \right),\end{aligned}\tag{31}$$

which directly corresponds to the terms of the *Gauss corrected* scheme appearing in Eq.(18). In particular, the first term in Eq.(31), corresponding to the anisotropic diffusivity tensor $\Gamma_{K,L}^{\parallel}$, is amenable to approximation by a two-point flux scheme, in a manner similar to what is done in the perpendicular bisection method in [31]. This term, when inserted into the finite volume diffusive flux, yields

$$\begin{aligned}F_{K,L}^{\parallel} &= -|\sigma| \left(\Gamma_{K,L}^{\parallel} \nabla_{K,L} u \right) \cdot \mathbf{n}_{K,\sigma} = -|\sigma| \nabla_{K,L} u \cdot \left(\Gamma_{K,L}^{\parallel} \mathbf{n}_{K,\sigma} \right) \\ &= -\alpha_{K,L} \frac{|\sigma|}{\mathbf{i}_{K,L}^{\top}\mathbf{n}_{K,\sigma}} \mathbf{i}_{K,L} \cdot \nabla_{K,L} u \\ &= -\alpha_{K|L} \frac{|\sigma|}{\mathbf{i}_{K,L}^{\top}\mathbf{n}_{K,L}} \frac{u_K - u_L}{|\mathbf{x}_L - \mathbf{x}_K|},\end{aligned}\tag{32}$$

that generates a directional derivative which can be easily approximated via a two-point flux scheme. On the other hand, the second term in the diffusive flux corresponding to the anisotropic diffusivity tensor $\Gamma_{K,L}^{\#}$ must be treated via a reconstruction of the cell gradient. it is important also to notice that, from the tensor decomposition in Eqs.(28)-(29), the two-point flux portion increases its dominance for increasing mesh non-orthogonality, due to the increasing angle between the unit vectors $\mathbf{n}_{K,\sigma}$ and $\mathbf{i}_{K,L}$. This property is beneficial in guaranteeing diagonal dominance of the linear system matrix and thus numerical stability, as will be clear from the rest of the discussion.

4 Discrete weak formulation

Returning to the diffusive fluxes from *Gauss corrected* scheme (18), using the diffusivity tensor decomposition in Eqs.(28)-(29), the finite

volume fluxes can be rewritten in the form

$$F_{K,L} = \alpha_{K|L} \tau_{K|L} (u_K - u_L) - |\sigma| \nabla_{K|L} u \cdot \left(\Gamma_{K,L}^\# \mathbf{n}_{K,\sigma} \right), \quad (33a)$$

$$F_{K,\sigma} = \alpha_{K,\sigma} \tau_{K,\sigma} u_K - |\sigma| \nabla_{K,\sigma} u \cdot \left(\Gamma_{K,\sigma}^\# \mathbf{n}_{K,\sigma} \right), \quad (33b)$$

for the internal and external faces, respectively. In these formulae, the transmissivities

$$\tau_{K|L} = \frac{|\sigma|}{d_{K,L}}, \quad \forall K|L \in \mathcal{F}_{int} \quad \text{and} \quad \tau_{K,\sigma} = \frac{|\sigma|}{d_{K,\sigma}}, \quad \forall \sigma \in \mathcal{F}_{K,ext} \quad (34)$$

are introduced to simplify the notation and $\nabla_{K|L} u$ denotes a generic discrete gradient operator, still to be defined, that is piecewise constant on the diamond cells $D_{K|L}$ for all $K \in \mathcal{M}$ and $L \in \mathcal{N}_K$. If one defines the diamond cell gradient from the linear interpolation of cell gradients as in Eq.(22), then the fluxes become

$$F_{K,L} = \alpha_{K|L} \tau_{K|L} (u_K - u_L) + \left(-\nabla_K u \cdot \left(\Gamma_{K,L}^\# \mathbf{a}_{K,L} \right) + \nabla_L u \cdot \left(\Gamma_{K,L}^\# \mathbf{a}_{L,K} \right) \right), \quad (35a)$$

$$F_{K,\sigma} = \alpha_{K,\sigma} \tau_{K,\sigma} u_K - \nabla_K u \cdot \left(\Gamma_{K,\sigma}^\# \mathbf{a}_{K,\sigma} \right), \quad (35b)$$

where the vector quantities

$$\mathbf{a}_{K,L} = |\sigma| \frac{d_{L,\sigma}}{d_{K,L}} \mathbf{n}_{K,\sigma}, \quad \forall K|L \in \mathcal{F}_{int} \quad (36a)$$

$$\mathbf{a}_{K,\sigma} = |\sigma| \mathbf{n}_{K,\sigma}, \quad \forall \sigma \in \mathcal{F}_{K,ext}. \quad (36b)$$

have been introduced, which are such that $\mathbf{a}_{K,L} \neq \mathbf{a}_{L,K}$ generally. Notice also the approximation introduced in the boundary term $\nabla_{K,\sigma} u \approx \nabla_K u$. It is now possible to derive the weak formulation underlying the finite volume scheme (17). By multiplying Eq.(17) by the test function v_K and summing the result for all $K \in \mathcal{M}$, one obtains

$$\sum_{K \in \mathcal{M}} v_K \sum_{L \in \mathcal{N}_K} F_{K,L} + \sum_{K \in \mathcal{M}} v_K \sum_{\sigma \in \mathcal{F}_{K,ext}} F_{K,\sigma} = \sum_{K \in \mathcal{M}} v_K \int_K f(\mathbf{x}) \, d\mathbf{x}$$

that, after discrete integration by parts, produces

$$\begin{aligned} \sum_{K|L \in \mathcal{F}_{int}} (F_{K,L} v_K + F_{L,K} v_L) + \sum_{K \in \mathcal{M}} \sum_{\sigma \in \mathcal{F}_{K,ext}} F_{K,\sigma} v_K \\ = \sum_{K \in \mathcal{M}} v_K \int_K f(\mathbf{x}) \, d\mathbf{x} \end{aligned}$$

from which, due to flux conservativity (21), one obtains that

$$\begin{aligned} \sum_{K|L \in \mathcal{F}_{int}} F_{K,L}(v_K - v_L) + \sum_{K \in \mathcal{M}} \sum_{\sigma \in \mathcal{F}_{K,ext}} F_{K,\sigma} v_K \\ = \sum_{K \in \mathcal{M}} v_K \int_K f(\mathbf{x}) \, d\mathbf{x}. \end{aligned} \quad (37)$$

By substituting into Eq.(37) the fluxes (35) with the face gradient from the linear interpolation of the Gauss scheme (15), it is possible to identify two terms T_1 and $T_2 = T_{2,int} + T_{2,ext}$ in the expression

$$T_1 + T_{2,int} + T_{2,ext} = \sum_{K \in \mathcal{M}} v_K \int_K f(\mathbf{x}) \, d\mathbf{x}, \quad (38)$$

where

$$\begin{aligned} T_1 = \sum_{K|L \in \mathcal{F}_{int}} \alpha_{K|L} \tau_{K|L} (u_K - u_L) (v_K - v_L) \\ + \sum_{K \in \mathcal{M}} \sum_{\sigma \in \mathcal{F}_{K,ext}} \alpha_{K,\sigma} \tau_{K,\sigma} u_K v_K, \end{aligned} \quad (39a)$$

$$\begin{aligned} T_{2,int} = \sum_{K|L \in \mathcal{F}_{int}} \left(-\nabla_K u \cdot \left(\Gamma_{K,L}^{\parallel} \mathbf{a}_{K,L} \right) \right. \\ \left. + \nabla_L u \cdot \left(\Gamma_{K,L}^{\parallel} \mathbf{a}_{L,K} \right) \right) (v_K - v_L), \end{aligned} \quad (39b)$$

$$T_{2,ext} = - \sum_{K \in \mathcal{M}} \sum_{\sigma \in \mathcal{F}_{K,ext}} \nabla_K u \cdot \left(\Gamma_{K,\sigma}^{\parallel} \mathbf{a}_{K,\sigma} \right) v_K. \quad (39c)$$

Notice that term T_1 defines a symmetric bilinear form

$$\begin{aligned} [u, v]_{\mathcal{D}, \alpha, \parallel} = \sum_{K|L \in \mathcal{F}_{int}} \alpha_{K|L} \tau_{K|L} (u_K - u_L) (v_K - v_L) \\ + \sum_{K \in \mathcal{M}} \sum_{\sigma \in \mathcal{F}_{K,ext}} \alpha_{K,\sigma} \tau_{K,\sigma} u_K v_K \end{aligned} \quad (40)$$

which is a discretization of the term $\int_{\Omega} \nabla \underline{u}(\mathbf{x}) \cdot \left(\Gamma_{\alpha}^{\parallel}(\mathbf{x}) \nabla v(\mathbf{x}) \right) \, d\mathbf{x}$, directly corresponding to the portion of diffusive fluxes that can be ascribed to the anisotropic diffusion tensor $\Gamma_{\alpha}^{\parallel}$. This expresses the flux component that is parallel to the local vector $(\mathbf{x}_L - \mathbf{x}_K)$ associated to internal faces $\sigma = K|L$, or to the vector $(\mathbf{x}_{\sigma} - \mathbf{x}_K)$ associated to boundary faces. The term $T_2 = T_{2,int} + T_{2,ext}$ contains instead the vectors $\mathbf{a}_{K,L}$ and $\mathbf{a}_{K,\sigma}$, related to the diffusivity tensor

Γ_α^\sharp , but it is not yet in a form readily corresponding to a discrete weak formulation. To this purpose, it is convenient to rewrite T_2 as

$$\begin{aligned}
T_2 &= \sum_{K \in \mathcal{M}} \sum_{L \in \mathcal{N}_K} -\nabla_K u \cdot \left(\Gamma_{K,L}^\sharp \mathbf{a}_{K,L} \right) (v_K - v_L) \\
&\quad - \sum_{K \in \mathcal{M}} \sum_{\sigma \in \mathcal{F}_{K,ext}} \nabla_K u \cdot \left(\Gamma_{K,\sigma}^\sharp \mathbf{a}_{K,\sigma} \right) v_K \\
&= \sum_{K \in \mathcal{M}} \nabla_K u \cdot \left(\sum_{L \in \mathcal{N}_K} \Gamma_{K,L}^\sharp \mathbf{a}_{K,L} (v_L - v_K) \right) \\
&\quad - \sum_{K \in \mathcal{M}} \nabla_K u \cdot \left(\sum_{\sigma \in \mathcal{F}_{K,ext}} \Gamma_{K,\sigma}^\sharp \mathbf{a}_{K,\sigma} v_K \right).
\end{aligned}$$

The two summation terms between brackets contained in the last expression correspond to the internal faces and the boundary faces contributions, respectively. They can be interpreted as a discretization of the term $\int_K \Gamma_\alpha^\sharp \nabla v(\mathbf{x}) \, d\mathbf{x}$ for all $K \in \mathcal{M}$, which allows to introduce the piecewise constant function $(\Gamma_\alpha^\sharp \nabla v)_\mathcal{D}$ that is defined on each cell $K \in \mathcal{M}$ as

$$\begin{aligned}
(\Gamma_\alpha^\sharp \nabla v)_K &= \frac{1}{|K|} \left(\sum_{L \in \mathcal{N}_K} \Gamma_{K,L}^\sharp \mathbf{a}_{K,L} (v_L - v_K) \right. \\
&\quad \left. - \sum_{\sigma \in \mathcal{F}_{K,ext}} \Gamma_{K,\sigma}^\sharp \mathbf{a}_{K,\sigma} v_K \right), \tag{41}
\end{aligned}$$

expressing the fact that the test function gradient cannot be separated from the diffusivity tensor Γ_α^\sharp , since the latter is a face-based quantity defined from the local mesh non-orthogonality (i.e., from the angle between $\mathbf{i}_{K,L}$ and $\mathbf{n}_{K,\sigma}$ unit vectors). In this case, the term T_2 can be rewritten as a non-symmetric discrete bilinear form

$$\langle \nabla u, \nabla v \rangle_{\mathcal{D}, \alpha, \sharp} = \sum_{K \in \mathcal{M}} |K| \nabla_K u \cdot (\Gamma_\alpha^\sharp \nabla v)_K = T_2 \tag{42}$$

which is thus associated to the diffusivity tensor Γ_α^\sharp , expressing the contribution of diffusion from a local direction not aligned with $(\mathbf{x}_L - \mathbf{x}_K)$ at internal faces, or with $(\mathbf{x}_\sigma - \mathbf{x}_K)$ at boundary faces.

Thus, the discrete weak formulation implied when using the *Gauss corrected* scheme for the heterogeneous isotropic diffusion

problem (1) takes the form

$$\begin{aligned}
u &\in H_{\mathcal{D}}, \\
[u, v]_{\mathcal{D}, \alpha, \parallel} + \langle \nabla u, \nabla v \rangle_{\mathcal{D}, \alpha, \parallel} &= \sum_{K \in \mathcal{M}} v_K \int_K f(\mathbf{x}) \, d\mathbf{x}, \\
&\forall v \in H_{\mathcal{D}}.
\end{aligned} \tag{43}$$

Several remarks are in order on the basis of the previously introduced formulation. Similarly to [22], cell gradients can lead to a discrete inner product whenever the mesh geometry allows for a direct estimation of face normal fluxes, e.g., in the case of an orthogonal polyhedral mesh. When instead anisotropic effects (directional bias) emerge locally on cell faces due to mesh non-orthogonality, the construction of face gradients becomes inevitable, as done in [25]. In this latter case, the diffusivity tensor is necessarily defined on diamond cell support.

Secondly, the linear interpolation operator used to obtain the face gradient $\nabla_{\sigma} u$ in the *Gauss corrected* scheme from a linear combination of cell gradients calculated via the Gauss gradient scheme for all $K|L \in \mathcal{F}_{int}$ implies that

$$\nabla_{K,L} u = \frac{|D_{L,\sigma}|}{|D_{\sigma}|} \nabla_K u + \frac{|D_{K,\sigma}|}{|D_{\sigma}|} \nabla_L u, \tag{44}$$

which defines the face gradient as the diamond cell gradient obtained via an inverse volume weighting procedure. The same conclusion is also valid for the scalar diffusivity $\alpha_{K|L}$ defined from Eq.(20).

Finally, when the scalar diffusivity $\alpha_{K|L}$ appearing inside the fluxes (33) is computed by a linear interpolation procedure, the anisotropic diffusivity tensor $\Gamma_{K,L}^{\parallel}$ in the *Gauss corrected* scheme becomes

$$\begin{aligned}
\Gamma_{K|L}^{\parallel} &= \left(\frac{d_{L,\sigma}}{d_{K,L}} \alpha_K + \frac{d_{K,\sigma}}{d_{K,L}} \alpha_L \right) \left(\mathbf{I} - \frac{\mathbf{i}_{K,L} \mathbf{i}_{K,L}^{\top}}{\left(\mathbf{i}_{K,L}^{\top} \mathbf{n}_{K,\sigma} \right)^2} \right) \\
&= \left(\frac{d_{L,\sigma}}{d_{K,L}} \alpha_K + \frac{d_{K,\sigma}}{d_{K,L}} \alpha_L \right) \mathbf{J}_{K,L}
\end{aligned}$$

where the face non-orthogonality symmetric tensor $\mathbf{J}_{K,L}$ has been defined. As a consequence, the diffusive flux (33) can be recast into the form

$$\begin{aligned}
F_{K,L} &= \alpha_{K|L} \tau_{K|L} (u_K - u_L) - |\sigma| \left(\frac{d_{L,\sigma}}{d_{K,L}} \alpha_K + \frac{d_{K,\sigma}}{d_{K,L}} \alpha_L \right) \\
&\quad \times \left(\frac{d_{L,\sigma}}{d_{K,L}} \nabla_K u + \frac{d_{K,\sigma}}{d_{K,L}} \nabla_L u \right) \cdot (\mathbf{J}_{K,L} \mathbf{n}_{K,\sigma})
\end{aligned}$$

from which, after introducing the vectors

$$\mathbf{b}_{K,L} = |\sigma| \left(\frac{d_{L,\sigma}}{d_{K,L}} \right)^2 \mathbf{J}_{K,L} \mathbf{n}_{K,\sigma} \quad (45a)$$

$$\mathbf{b}_{L,K} = |\sigma| \left(\frac{d_{K,\sigma}}{d_{K,L}} \right)^2 \mathbf{J}_{L,K} \mathbf{n}_{L,\sigma}, \quad (45b)$$

one obtains that

$$\begin{aligned} F_{K,L} = & \alpha_{K|L} \tau_{K|L} (u_K - u_L) - \alpha_K \nabla_K u \cdot \mathbf{b}_{K,L} + \alpha_L \nabla_L u \cdot \mathbf{b}_{L,K} \\ & + |\sigma| (\alpha_K \nabla_L u + \alpha_L \nabla_K u) \cdot \left(\frac{d_{K,\sigma} d_{L,\sigma}}{d_{K,L}^2} \mathbf{J}_{K,L} \mathbf{n}_{K,\sigma} \right). \end{aligned} \quad (46)$$

Notice that, if the last term vanishes, the same structure of the anisotropic diffusion fluxes from [22] is recovered, similarly to the case of cell based diffusion coefficients, but with differently defined $\mathbf{b}_{K,L}$ and $\mathbf{b}_{L,K}$ vectors. This implies that a weak formulation similar to the one in [22] can also be obtained in this case. Nevertheless, on generally non-orthogonal meshes, the last term in Eq.(46) vanishes only when $\alpha_K \nabla_L u = -\alpha_L \nabla_K u$, i.e., only on internal faces where the flux is zero. In all the other meaningful cases, the last term in Eq.(46) is non zero and it is responsible for the cross terms inside the non-orthogonal correction $\langle \nabla u, \nabla v \rangle_{\mathcal{D},\alpha,\parallel}$ appearing in the weak formulation (43).

5 Convergence analysis

The term T_1 defined in Eq.(40) and appearing in the discrete weak formulation (43) exactly corresponds to the symmetric bilinear form appearing in [22] for isotropic diffusion operators on polyhedral meshes satisfying the additional orthogonality condition

$$(\mathbf{x}_L - \mathbf{x}_K) \perp \mathbf{n}_{K,\sigma}. \quad (47)$$

However, in the present analysis the same inner product corresponds only to the portion of the discrete bilinear form containing the contribution to the diffusive flux that is parallel to the local mesh direction, as identified from the cell-to-cell vector $(\mathbf{x}_L - \mathbf{x}_K)$. Formally, it is possible to define the discrete inner product

$$\begin{aligned} [u, v]_{\mathcal{D},\alpha,\parallel} = & \sum_{K|L \in \mathcal{F}_{int}} \alpha_{K|L} \tau_{K|L} (u_K - u_L)(v_K - v_L) \\ & + \sum_{K \in \mathcal{M}} \sum_{\sigma \in \mathcal{F}_{K,ext}} \alpha_{K,\sigma} \tau_{K,\sigma} u_K v_K \end{aligned} \quad (48)$$

from which the associated norm

$$\|u\|_{\mathcal{D}} = ([u, u]_{\mathcal{D},1,\|})^{1/2} \quad (49)$$

directly follows, where we have set $\alpha = 1$. Such norm verifies the discrete Poincaré inequality

$$\|w\|_{L^2(\Omega)} \leq \text{diam}(\Omega) \|w\|_{\mathcal{D}}, \quad \forall w \in H_{\mathcal{D}} \quad (50)$$

as from [21]. Furthermore, a relative compactness result in $L^2(\Omega)$ also holds.

Lemma 5.1 ([22], Lemma 2.1). *Let Ω be a bounded open connected polyhedral subset of \mathbb{R}^d , $d \in \mathbb{N}^*$ and let $(\mathcal{D}_n, u_n)_{n \in \mathbb{N}}$ be a sequence of discretizations such that, for all $n \in \mathbb{N}$, \mathcal{D}_n is an admissible finite volume mesh in sense of Definition 2.1 and $u_n \in H_{\mathcal{D}_n}(\Omega)$. Assume that $\lim_{n \rightarrow \infty} h_{\mathcal{D}_n} = 0$ and that there exists a constant $C_1 > 0$ such that $\|u\|_{\mathcal{D}_n} \leq C_1$, for all $n \in \mathbb{N}$. Then there exists a subsequence of $(\mathcal{D}_n, u_n)_{n \in \mathbb{N}}$, for simplicity denoted again by (\mathcal{D}_n, u_n) , and some $\underline{u} \in H_0^1(\Omega)$ such that u_n tends to \underline{u} in $L^2(\Omega)$ as $n \rightarrow \infty$, and the inequality*

$$\int_{\Omega} |\nabla \underline{u}(\mathbf{x})|^2 \, d\mathbf{x} \leq \liminf_{n \rightarrow \infty} \|u_n\|_{\mathcal{D}_n}^2 \quad (51)$$

holds. Furthermore, for all regular functions $L^\infty(\Omega)$, one has also that

$$\lim_{n \rightarrow \infty} [u_n, P_{\mathcal{D}_n} \varphi]_{\mathcal{D}_n, \alpha, \|} = \int_{\Omega} \Gamma_{\alpha}^{\|}(\mathbf{x}) \nabla \underline{u}(\mathbf{x}) \cdot \nabla \varphi(\mathbf{x}) \, d\mathbf{x}, \quad (52)$$

$$\forall \varphi \in C_c^\infty(\Omega).$$

with $P_{\mathcal{D}} : C(\Omega) \rightarrow H_{\mathcal{D}}(\Omega)$ the projection operator from Section 2.

Proof. The proof is similar to the one reported in [22], which is obtained for orthogonal meshes, even if orthogonality is not strictly required, after substitution of the scalar diffusivity α with the diffusivity tensor $\Gamma_{\alpha}^{\|}$.

From the discussion leading to the discrete weak form (43), it is useful to define a discrete gradient with anisotropic diffusivity biasing, see also Eq.(41).

Definition 5.1 (Discrete gradient with $\Gamma_{\alpha}^{\|}$ biasing). *Let Ω be a bounded open connected polyhedral subset of \mathbb{R}^d , $d \in \mathbb{N}^*$. Let \mathcal{D} be an admissible finite volume discretization in sense of Definition (2.1).*

The discrete gradient with Γ_α^\sharp anisotropic biasing $\nabla_{\mathcal{D},\alpha,\sharp} : H_{\mathcal{D}} \rightarrow H_{\mathcal{D}}^d$ is defined for any $u \in H_{\mathcal{D}}$ as the piecewise constant function

$$\begin{aligned} \nabla_{\mathcal{D},\alpha,\sharp} u(\mathbf{x}) &= (\Gamma_\alpha^\sharp \nabla u)_K \\ &= \frac{1}{|K|} \left(\sum_{L \in \mathcal{N}_K} \Gamma_{K,L}^\sharp \mathbf{a}_{K,L} (u_L - u_K) - \sum_{\sigma \in \mathcal{F}_{K,ext}} \Gamma_{K,\sigma}^\sharp \mathbf{a}_{K,\sigma} u_K \right), \\ &\text{for a.e. } \mathbf{x} \in K, \quad \forall K \in \mathcal{M}, \end{aligned} \tag{53}$$

where the discrete anisotropic diffusivity tensor $\Gamma_{K,L}^\sharp$ (and $\Gamma_{K,\sigma}^\sharp$) is defined in (29) and the vector quantities $\mathbf{a}_{K,L}$ (and $\mathbf{a}_{K,\sigma}$) are defined in (36).

From the diffusivity tensor decomposition in Eqs.(28)-(29), it is possible to split the Γ_α^\sharp -biased discrete gradient into two other discrete gradients.

Definition 5.2 (Decomposition of Γ_α^\sharp -biased discrete gradient). *Let Ω be a bounded open connected polyhedral subset of \mathbb{R}^d , $d \in \mathbb{N}^*$ and let \mathcal{D} be an admissible finite volume discretization in sense of Definition (2.1). Let $\nabla_{\mathcal{D},\alpha,\sharp}$ be the Γ_α^\sharp -biased discrete gradient, as from Definition 5.1, for any $u \in H_{\mathcal{D}}$. Then the Γ_α^\sharp -biased discrete gradient can be decomposed into the sum of two other discrete gradients*

$$\nabla_{\mathcal{D},\alpha,\sharp} u(\mathbf{x}) = \nabla_{\mathcal{D},\alpha} u(\mathbf{x}) - \nabla_{\mathcal{D},\alpha,\parallel} u(\mathbf{x}) \tag{54}$$

where

$$\begin{aligned} \nabla_{\mathcal{D},\alpha} u(\mathbf{x}) &= (\alpha \nabla u)_K \\ &= \frac{1}{|K|} \left(\sum_{L \in \mathcal{N}_K} \alpha_{K|L} \tau_{K|L} d_{L,\sigma} \mathbf{n}_{K,\sigma} (u_L - u_K) \right. \\ &\quad \left. - \sum_{\sigma \in \mathcal{F}_{K,ext}} \alpha_{K,\sigma} |\sigma| \mathbf{n}_{K,\sigma} u_K \right), \\ &\text{for a.e. } \mathbf{x} \in K, \quad \forall K \in \mathcal{M}, \end{aligned} \tag{55}$$

represents a diffusivity weighted discrete gradient, while

$$\begin{aligned} \nabla_{\mathcal{D},\alpha,\|} u(\mathbf{x}) &= (\Gamma_{\alpha}^{\|} \nabla u)_K \\ &= \frac{1}{|K|} \left(\sum_{L \in \mathcal{N}_K} \alpha_{K|L} \tau_{K|L} d_{L,\sigma} \frac{\mathbf{i}_{K,L}}{\mathbf{n}_{K,\sigma} \cdot \mathbf{i}_{K,L}} (u_L - u_K) \right. \\ &\quad \left. - \sum_{\sigma \in \mathcal{F}_{K,ext}} \alpha_{K,\sigma} |\sigma| \frac{\mathbf{i}_{K,L}}{\mathbf{n}_{K,\sigma} \cdot \mathbf{i}_{K,L}} u_K \right), \end{aligned} \quad (56)$$

for a.e. $\mathbf{x} \in K$, $\forall K \in \mathcal{M}$,

can be interpreted as a $\Gamma_{\alpha}^{\|}$ -biased discrete gradient.

For this finite volume diffusion scheme, the mesh regularity is measured by the factor

$$\tilde{\theta}_{\mathcal{D}} = \min \left\{ \min \left\{ \frac{d_{K,\sigma}}{d_{L,\sigma}}, \frac{h_K}{d_{K,\sigma}}, \mathbf{n}_{K,L} \cdot \mathbf{i}_{K,L} : \sigma \in \mathcal{F}_{int} \right\}, \min \left\{ \frac{h_K}{d_{K,\sigma}}, \mathbf{i}_{K,\sigma} \cdot \mathbf{n}_{K,\sigma} : \sigma \in \mathcal{F}_{ext} \right\} \right\}, \quad (57)$$

which expresses bounds in the empirical measures of mesh regularity that will be presented in Section 6. As a first result, one introduces the bound on the $L^2(\Omega)^d$ -norm of the $\Gamma_{\alpha}^{\|}$ -biased gradient on any element of $H_{\mathcal{D}}$.

Lemma 5.2 (Bound on $\nabla_{\mathcal{D},\alpha,\|} u$). *Let Ω be a bounded open connected polyhedral subset of \mathbb{R}^d , $d \in \mathbb{N}^*$. Let \mathcal{D} be an admissible finite volume discretization in sense of Definition 2.1 and let $0 < \theta \leq \tilde{\theta}_{\mathcal{D}}$. Then, there exists C_1 depending only on d , α and θ such that, for all $u \in H_{\mathcal{D}}$, one has*

$$\|\nabla_{\mathcal{D},\alpha,\|} u\|_{L^2(\Omega)^d} \leq C_1 \|u\|_{\mathcal{D}}. \quad (58)$$

Proof. Let $u \in H_{\mathcal{D}}$. Similarly as in [22], one introduces, for all $K \in \mathcal{M}$, $L \in \mathcal{N}_K$ and $\sigma = K|L$ the difference quantities $\delta_{K,\sigma} \mathbf{x} = (\mathbf{x}_L - \mathbf{x}_K)$ and $\delta_{K,\sigma} u = (u_L - u_K)$, and for all $\sigma \in \mathcal{F}_{K,\sigma}$ the quantities $\delta_{K,\sigma} \mathbf{x} = (\mathbf{x}_{\sigma} - \mathbf{x}_K)$ and $\delta_{K,\sigma} u = -u_K$. Then, the inner product norm in (49) leads for a given $K \in \mathcal{M}$ to

$$\begin{aligned} \|u\|_{\mathcal{D}}^2 &= [u, u]_{\mathcal{D},1,\|} \\ &= \sum_{K|L \in \mathcal{F}_{int}} \tau_{K|L} (u_L - u_K)^2 + \sum_{K \in \mathcal{M}} \sum_{\sigma \in \mathcal{F}_{K,ext}} \tau_{K,\sigma} (-u_K)^2 \\ &= \sum_{K \in \mathcal{M}} \frac{1}{2} \sum_{L \in \mathcal{N}_K} \tau_{K|L} (\delta_{K,L} u)^2 + \sum_{K \in \mathcal{M}} \sum_{\sigma \in \mathcal{F}_{K,ext}} \tau_{K,\sigma} (\delta_{K,\sigma} u)^2. \end{aligned}$$

Then, Definition 5.1 leads to

$$|K|(\Gamma_\alpha^\# \nabla u)_K = \sum_{\sigma \in \mathcal{F}_K} \alpha_\sigma \tau_\sigma d_{L,\sigma} \left(\mathbf{n}_{K,\sigma} - \frac{\delta_{K,\sigma} \mathbf{x}}{d_{K,L}} \right) \delta_{K,\sigma} u.$$

By using the Cauchy-Schwartz inequality, one obtains that

$$\begin{aligned} |K|^2 \left| (\Gamma_\alpha^\# \nabla u)_K \right|^2 &\leq \sum_{\sigma \in \mathcal{F}_K} \tau_\sigma \alpha_\sigma^2 \left| d_{L,\sigma} \left(\mathbf{n}_{K,\sigma} - \frac{\delta_{K,\sigma} \mathbf{x}}{d_{K,L}} \right) \right|^2 \\ &\quad \times \sum_{\sigma \in \mathcal{F}_K} \tau_\sigma (\delta_{K,\sigma} u)^2, \end{aligned}$$

from which, by introducing the upper bound for the scalar diffusivity $C_\alpha \geq \alpha_\sigma^2$, for all $\sigma \in \mathcal{F}$, and by noticing that, for $\sigma \in \mathcal{F}_K$, one has $\tau_\sigma \leq \frac{|\sigma|}{d_{K,\sigma}}$ and that $\delta_{K,\sigma} \mathbf{x} \leq (\mathbf{x}_\sigma - \mathbf{x}_K)$, it follows that

$$\begin{aligned} |K|^2 \left| (\Gamma_\alpha^\# \nabla u)_K \right|^2 &\leq C_\alpha \sum_{\sigma \in \mathcal{F}_K} d|D_{K,\sigma}| \left| \frac{d_{L,\sigma}}{d_{K,\sigma}} \mathbf{n}_{K,\sigma} - \frac{d_{L,\sigma}}{d_{K,\sigma}} \frac{\mathbf{x}_\sigma - \mathbf{x}_K}{d_{K,L}} \right|^2 \\ &\quad \times \sum_{\sigma \in \mathcal{F}_K} \tau_\sigma (\delta_{K,\sigma} u)^2 \\ &\leq C_\alpha d \sum_{\sigma \in \mathcal{F}_K} |D_{K,\sigma}| \left| \frac{d_{L,\sigma}}{d_{K,\sigma}} \mathbf{n}_{K,\sigma} \right|^2 \sum_{\sigma \in \mathcal{F}_K} \tau_\sigma (\delta_{K,\sigma} u)^2 \\ &\leq C_\alpha \frac{d}{\theta^2} |K| \sum_{\sigma \in \mathcal{F}_K} \tau_\sigma (\delta_{K,\sigma} u)^2. \end{aligned}$$

Finally, after summing over all $K \in \mathcal{M}$, one obtains that

$$\begin{aligned} \sum_{K \in \mathcal{M}} |K| \left| (\Gamma_\alpha^\# \nabla u)_K \right|^2 &\leq C_\alpha \frac{d}{\theta^2} \sum_{K \in \mathcal{M}} \sum_{\sigma \in \mathcal{F}_K} \tau_\sigma (\delta_{K,\sigma} u)^2 \\ &\leq C_\alpha \frac{d}{\theta^2} \sum_{K \in \mathcal{M}} \left(\sum_{\sigma \in \mathcal{F}_{K,int}} \tau_\sigma (\delta_{K,\sigma} u)^2 \right. \\ &\quad \left. + 2 \sum_{\sigma \in \mathcal{F}_{K,ext}} \tau_\sigma (\delta_{K,\sigma} u)^2 \right) \\ &= 2 C_\alpha \frac{d}{\theta^2} \|u\|_{\mathcal{D}}^2 \end{aligned}$$

from which (58) follows with $C_1 = (1/\theta)\sqrt{2C_\alpha d}$.

It is now possible to state a weak convergence property for the diffusion weighted discrete gradient.

Lemma 5.3 (Weak convergence of $\nabla_{\mathcal{D},\alpha}u$). *Let Ω be a bounded open connected polyhedral subset of \mathbb{R}^d , $d \in \mathbb{N}^*$. Let \mathcal{D} be an admissible finite volume discretization in sense of Definition (2.1) and let $0 < \theta \leq \tilde{\theta}_{\mathcal{D}}$. Assume that there exists $u \in H_{\mathcal{D}}$ and a function $\underline{u} \in H_0^1(\Omega)$ such that u tends to \underline{u} in $L^2(\Omega)$ as $h_{\mathcal{D}} \rightarrow 0$, while $\|u\|_{\mathcal{D}}$ remains bounded. Then $\nabla_{\mathcal{D},\alpha}u$ weakly converges to $\alpha \nabla \underline{u}$ in $L^2(\Omega)^d$ as $h_{\mathcal{D}} \rightarrow 0$. Additionally, $\nabla_{\mathcal{D},\alpha,\|}u$ weakly converges to $\Gamma_{\alpha}^{\|} \nabla \underline{u}$ as $h_{\mathcal{D}} \rightarrow 0$.*

Proof. In or Let $\varphi \in C_c^{\infty}(\Omega)$. Assume that $h_{\mathcal{D}}$ is small enough that, for all $K \in \mathcal{M}$ and $\mathbf{x} \in K$, if $\varphi(\mathbf{x}) \neq 0$ then $\mathcal{F}_{K,ext} = \emptyset$. Consider the term $T_1^{\mathcal{D}}$ defined as

$$\begin{aligned}
T_1^{\mathcal{D}} &= \int_{\Omega} P_{\mathcal{D}} \varphi(\mathbf{x}) \nabla_{\mathcal{D},\alpha,\|} u(\mathbf{x}) \, d\mathbf{x} = \sum_{K \in \mathcal{M}} |K| \varphi(\mathbf{x}_K) (\Gamma_{\alpha}^{\|} \nabla u)_K \\
&= \sum_{K|L \in \mathcal{F}_{int}} \left(\varphi(\mathbf{x}_K) d_{L,\sigma} \alpha_{K|L} \tau_{K|L} \right. \\
&\quad \times \left. \left(\mathbf{n}_{K,\sigma} - \frac{(\mathbf{x}_L - \mathbf{x}_K)}{d_{K,L}} \right) (u_L - u_K) \right) \\
&\quad + \sum_{K|L \in \mathcal{F}_{int}} \left(-\varphi(\mathbf{x}_L) d_{K,\sigma} \alpha_{K|L} \tau_{K|L} \right. \\
&\quad \times \left. \left(\mathbf{n}_{K,\sigma} - \frac{(\mathbf{x}_L - \mathbf{x}_K)}{d_{K,L}} \right) (u_K - u_L) \right) \\
&= \sum_{K|L \in \mathcal{F}_{int}} (\varphi(\mathbf{x}_K) d_{L,\sigma} + \varphi(\mathbf{x}_L) d_{K,\sigma}) \\
&\quad \times \alpha_{K|L} \tau_{K|L} \left(\mathbf{n}_{K,\sigma} - \frac{(\mathbf{x}_L - \mathbf{x}_K)}{d_{K,L}} \right) (u_L - u_K)
\end{aligned}$$

in which $\tau_{K|L} \left(\mathbf{n}_{K,\sigma} - \frac{(\mathbf{x}_L - \mathbf{x}_K)}{d_{K,L}} \right)$ defines a sort of non-orthogonal transmissivity, while the first term between brackets can be rewritten as

$$\begin{aligned}
&\varphi(\mathbf{x}_K) d_{L,\sigma} + \varphi(\mathbf{x}_L) d_{K,\sigma} \\
&= \varphi(\mathbf{x}_K) (\mathbf{x}_{\sigma} - \mathbf{x}_L) \cdot \mathbf{n}_{L,\sigma} + \varphi(\mathbf{x}_L) (\mathbf{x}_{\sigma} - \mathbf{x}_K) \cdot \mathbf{n}_{K,\sigma} \\
&= \left(\frac{\varphi(\mathbf{x}_K) + \varphi(\mathbf{x}_L)}{2} (\mathbf{x}_L - \mathbf{x}_K) \right. \\
&\quad \left. + (\varphi(\mathbf{x}_K) - \varphi(\mathbf{x}_L)) \left(\frac{\mathbf{x}_K + \mathbf{x}_L}{2} - \mathbf{x}_{\sigma} \right) \right) \cdot \mathbf{n}_{K,\sigma}.
\end{aligned}$$

The term $T_1^{\mathcal{D}}$ can be decomposed into a sum of two terms $T_1^{\mathcal{D}} =$

$T_2^{\mathcal{D}} + T_3^{\mathcal{D}}$, where

$$\begin{aligned}
T_2^{\mathcal{D}} &= \sum_{K|L \in \mathcal{F}_{int}} \mathbf{n}_{K,\sigma} \cdot (\mathbf{x}_L - \mathbf{x}_K) \frac{\varphi(\mathbf{x}_K) + \varphi(\mathbf{x}_L)}{2} \\
&\quad \times \alpha_{K|L} \tau_{K|L} \left(\mathbf{n}_{K,\sigma} - \frac{(\mathbf{x}_L - \mathbf{x}_K)}{d_{K,L}} \right) (u_L - u_K), \\
T_3^{\mathcal{D}} &= \sum_{K|L \in \mathcal{F}_{int}} \mathbf{n}_{K,\sigma} \cdot \left(\frac{\mathbf{x}_K + \mathbf{x}_L}{2} - \mathbf{x}_\sigma \right) (\varphi(\mathbf{x}_K) - \varphi(\mathbf{x}_L)) \\
&\quad \times \alpha_{K|L} \tau_{K|L} \left(\mathbf{n}_{K,\sigma} - \frac{(\mathbf{x}_L - \mathbf{x}_K)}{d_{K,L}} \right) (u_L - u_K).
\end{aligned}$$

Starting with the analysis of term $T_3^{\mathcal{D}}$, by Cauchy-Schwartz inequality one gets

$$\begin{aligned}
(T_3^{\mathcal{D}})^2 &\leq \sum_{K|L \in \mathcal{F}_{int}} \tau_{K|L} \alpha_{K|L}^2 \left| \frac{\mathbf{x}_K + \mathbf{x}_L}{2} - \mathbf{x}_\sigma \right|^2 \\
&\quad \times (\varphi(\mathbf{x}_K) - \varphi(\mathbf{x}_L))^2 \left| \mathbf{n}_{K,\sigma} - \frac{(\mathbf{x}_L - \mathbf{x}_K)}{d_{K,L}} \right|^2 \\
&\quad \times \sum_{K|L \in \mathcal{F}_{int}} \tau_{K|L} (u_L - u_K)^2,
\end{aligned}$$

in which, due to triangle inequality

$$\left| \frac{\mathbf{x}_K + \mathbf{x}_L}{2} - \mathbf{x}_\sigma \right| \leq \frac{1}{2} |\mathbf{x}_L - \mathbf{x}_\sigma| + \frac{1}{2} |\mathbf{x}_K - \mathbf{x}_\sigma| \leq h_{\mathcal{D}},$$

while due to mesh regularity

$$\begin{aligned}
&\left| \mathbf{n}_{K,\sigma} - \frac{(\mathbf{x}_L - \mathbf{x}_K)}{d_{K,L}} \right| \\
&= \left| \mathbf{n}_{K,\sigma} - \frac{\mathbf{i}_{K,L}}{\mathbf{n}_{K,\sigma} \cdot \mathbf{i}_{K,L}} \right| \leq 1 + \left| \frac{\mathbf{i}_{K,L}}{\mathbf{n}_{K,\sigma} \cdot \mathbf{i}_{K,L}} \right| \leq 1 + \frac{1}{\theta},
\end{aligned}$$

from which, after introducing $C_\alpha \geq \alpha_{K|L}^2$, it follows that

$$(T_3^{\mathcal{D}})^2 \leq C_\sigma C_2 \left(1 + \frac{1}{\theta} \right)^2 h_{\mathcal{D}}^2 |\Omega| \|u\|_{\mathcal{D}}^2$$

with C_2 only depending on d , Ω and φ . Thus one concludes that $\lim_{h_{\mathcal{D}} \rightarrow 0} T_3^{\mathcal{D}} = 0$. Successively, consider the term $T_2^{\mathcal{D}}$ that can be

rewritten as the sum of two terms

$$\begin{aligned}
T_2^{\mathcal{D}} &= \sum_{K|L \in \mathcal{F}_{int}} \alpha_{K|L} |\sigma| \frac{\varphi(\mathbf{x}_K) + \varphi(\mathbf{x}_L)}{2} \\
&\quad \times \left(\mathbf{n}_{K,\sigma} - \frac{\mathbf{i}_{K,L}}{\mathbf{n}_{K,\sigma} \cdot \mathbf{i}_{K,L}} \right) (u_L - u_K) \\
&= \sum_{K|L \in \mathcal{F}_{int}} \alpha_{K|L} |\sigma| \mathbf{n}_{K,\sigma} \frac{\varphi(\mathbf{x}_K) + \varphi(\mathbf{x}_L)}{2} (u_L - u_K) \\
&\quad - \sum_{K|L \in \mathcal{F}_{int}} \alpha_{K|L} |\sigma| \frac{\mathbf{i}_{K,L}}{\mathbf{n}_{K,\sigma} \cdot \mathbf{i}_{K,L}} \frac{\varphi(\mathbf{x}_K) + \varphi(\mathbf{x}_L)}{2} (u_L - u_K) \\
&= T_{2,1}^{\mathcal{D}} + T_{2,2}^{\mathcal{D}}.
\end{aligned}$$

Compare term $T_{2,1}^{\mathcal{D}}$ with the term

$$\begin{aligned}
T_4^{\mathcal{D}} &= - \int_{\Omega} \alpha(\mathbf{x}) u(\mathbf{x}) \nabla \varphi(\mathbf{x}) \, d\mathbf{x} \\
&= - \sum_{K \in \mathcal{M}} \sum_{\sigma \in \mathcal{F}_{K,int}} \alpha_{K|L} u_K \int_{K|L} \varphi(\mathbf{x}) \mathbf{n}_{K,\sigma} \, d\gamma(\mathbf{x}) \\
&= \sum_{K|L \in \mathcal{F}_{int}} \alpha_{K|L} (u_L - u_K) \int_{K|L} \varphi(\mathbf{x}) \mathbf{n}_{K,\sigma} \, d\gamma(\mathbf{x}),
\end{aligned}$$

which is such that

$$\lim_{h_{\mathcal{D}} \rightarrow 0} T_4^{\mathcal{D}} = - \int_{\Omega} \alpha(\mathbf{x}) \underline{u}(\mathbf{x}) \nabla \varphi(\mathbf{x}) \, d\mathbf{x} = \int_{\Omega} \alpha(\mathbf{x}) \varphi(\mathbf{x}) \nabla \underline{u}(\mathbf{x}) \, d\mathbf{x}.$$

Due to the fact that midpoint face interpolation is first order accurate

$$\left| \frac{1}{|\sigma|} \int_{K|L} \varphi(\mathbf{x}) \, d\gamma(\mathbf{x}) - \frac{\varphi(\mathbf{x}_K) + \varphi(\mathbf{x}_L)}{2} \right| \leq h_{\mathcal{D}} \|\nabla \varphi\|_{L^\infty(\Omega)},$$

one has that

$$\begin{aligned}
&(T_4^{\mathcal{D}} - T_{2,1}^{\mathcal{D}})^2 \\
&\leq \sum_{K|L \in \mathcal{F}_{int}} (\alpha_{K|L} |\sigma| \mathbf{n}_{K,\sigma})^2 (u_L - u_K)^2 \\
&\quad \times \sum_{K|L \in \mathcal{F}_{int}} \left| \frac{1}{|\sigma|} \int_{K|L} \varphi(\mathbf{x}) \, d\gamma(\mathbf{x}) - \frac{\varphi(\mathbf{x}_K) + \varphi(\mathbf{x}_L)}{2} \right|^2 \\
&\leq \sum_{K|L \in \mathcal{F}_{int}} (\alpha_{K|L} |\sigma| \mathbf{n}_{K,\sigma})^2 (u_L - u_K)^2 \sum_{K|L \in \mathcal{F}_{int}} h_{\mathcal{D}} \|\nabla \varphi\|_{L^\infty(\Omega)},
\end{aligned}$$

from which it follows that $\lim_{h_{\mathcal{D}} \rightarrow 0} (T_4^{\mathcal{D}} - T_{2,1}^{\mathcal{D}})^2 = 0$. Thus, $T_{2,1}^{\mathcal{D}} = T_1^{\mathcal{D}} - T_{2,2}^{\mathcal{D}}$ converges to $T_4^{\mathcal{D}}$ and, due to density of $C_c^\infty(\Omega)$ in $L^2(\Omega)$, $\nabla_{\mathcal{D},\alpha} u$ weakly converges to $\alpha \nabla \underline{u}$ as $h_{\mathcal{D}} \rightarrow 0$. Thus, the term $T_{2,2}^{\mathcal{D}}$ can be compared to

$$\begin{aligned} T_5^{\mathcal{D}} &= - \int_{\Omega} \left(\Gamma_{\alpha}^{\#}(\mathbf{x}) - \alpha(\mathbf{x}) \mathbf{I} \right) u(\mathbf{x}) \nabla \varphi(\mathbf{x}) \, d\mathbf{x} \\ &= - \sum_{K \in \mathcal{M}} \sum_{\sigma \in \mathcal{F}_{K,f}} \alpha_{K|L} \frac{\mathbf{i}_{K,L} \mathbf{i}_{K,L}^{\top}}{\left(\mathbf{i}_{K,L}^{\top} \mathbf{n}_{K,\sigma} \right)^2} u_K \int_{K|L} \varphi(\mathbf{x}) \mathbf{n}_{K,\sigma} \, d\gamma(\mathbf{x}) \end{aligned}$$

which is such that

$$\lim_{h_{\mathcal{D}} \rightarrow 0} T_5^{\mathcal{D}} = \int_{\Omega} \Gamma_{\alpha}^{\parallel}(\mathbf{x}) \underline{u}(\mathbf{x}) \nabla \varphi(\mathbf{x}) \, d\mathbf{x} = - \int_{\Omega} \Gamma_{\alpha}^{\parallel}(\mathbf{x}) \varphi(\mathbf{x}) \nabla \underline{u}(\mathbf{x}) \, d\mathbf{x}.$$

By a similar procedure, the term $T_{2,2}^{\mathcal{D}}$ converges to term $T_5^{\mathcal{D}}$ and so $\nabla_{\mathcal{D},\alpha} u$ weakly converges to $\Gamma_{\alpha}^{\parallel} \nabla \underline{u}$ as $h_{\mathcal{D}} \rightarrow 0$.

The diffusion weighted discrete gradient provides indeed a consistent gradient scheme.

Lemma 5.4 (Consistency of $\nabla_{\mathcal{D},\alpha}$). *Let Ω be a bounded open connected polyhedral subset of \mathbb{R}^d , $d \in \mathbb{N}^*$. Let \mathcal{D} be an admissible finite volume discretization in sense of Definition (2.1) and let $0 < \theta \leq \tilde{\theta}_{\mathcal{D}}$. Let $\underline{u} \in C^2(\bar{\Omega})$ be such that $\underline{u} = 0$ on $\partial\Omega$. Then there exists C_3 , depending only on Ω , θ , \underline{u} and α such that*

$$\|\nabla_{\mathcal{D},\alpha} P_{\mathcal{D}} \underline{u} - \alpha \nabla \underline{u}\|_{L^2(\Omega)^d} \leq C_3 h_{\mathcal{D}} \quad (59)$$

Proof. From Definition 5.2 for any $K \in \mathcal{M}$ one has

$$\begin{aligned} |K| (\nabla_{\mathcal{D},\alpha} P_{\mathcal{D}} \underline{u})_K &= \sum_{L \in \mathcal{N}_K} \alpha_{K|L} \tau_{K|L} d_{L,\sigma} \mathbf{n}_{K,\sigma} (\underline{u}(\mathbf{x}_L) - \underline{u}(\mathbf{x}_K)) \\ &\quad - \sum_{K \in \mathcal{M}} \sum_{\sigma \in \mathcal{F}_{K,ext}} \alpha_{K,\sigma} |\sigma| \mathbf{n}_{K,\sigma} \underline{u}(\mathbf{x}_K) \end{aligned}$$

Let $(\alpha \nabla \underline{u})_K$ be the mean value of $\alpha \nabla \underline{u}$ over K

$$(\alpha \nabla \underline{u})_K = \frac{1}{|K|} \int_K \alpha(\mathbf{x}) \nabla \underline{u}(\mathbf{x}) \, d\mathbf{x}.$$

Due to the regularity of \underline{u} and the homogeneous Dirichlet boundary conditions, the flux consistency error estimates include a constant

C_4 , only depending on L^∞ norm of second derivatives of \underline{u} (and of α), such that for all $\sigma = K|L \in \mathcal{F}_{int}$, one has

$$|e_\sigma| \leq C_4 h_{\mathcal{D}} \quad \text{with} \quad e_\sigma = (\alpha \nabla \underline{u})_K \cdot \mathbf{n}_{K,\sigma} - \alpha_{K|L} \frac{\underline{u}(\mathbf{x}_L) - \underline{u}(\mathbf{x}_K)}{d_{K,L}}$$

while for all $\sigma \in \mathcal{F}_{ext}$ one has

$$|e_\sigma| \leq C_4 h_{\mathcal{D}} \quad \text{with} \quad e_\sigma = (\alpha \nabla \underline{u})_K \cdot \mathbf{n}_{K,\sigma} - \alpha_{K,\sigma} \frac{-\underline{u}(\mathbf{x}_K)}{d_{K,\sigma}}$$

These flux consistency errors allow to recast $(\nabla_{\mathcal{D},\alpha} P_{\mathcal{D}} u)_K$ as

$$\begin{aligned} |K|(\nabla_{\mathcal{D},\alpha} P_{\mathcal{D}} u)_K &= \sum_{L \in \mathcal{N}_K} |\sigma| d_{L,\sigma} \mathbf{n}_{K,\sigma} (\alpha \nabla \underline{u})_K \cdot \mathbf{n}_{K,\sigma} \\ &\quad - \sum_{\sigma \in \mathcal{F}_{K,ext}} |\sigma| d_{K,\sigma} \mathbf{n}_{K,\sigma} (\alpha \nabla \underline{u})_K \cdot \mathbf{n}_{K,\sigma} + R_K \end{aligned}$$

where the consistency residual term is defined as

$$R_K = - \sum_{L \in \mathcal{N}_K} |\sigma| d_{L,\sigma} \mathbf{n}_{K,\sigma} e_\sigma - \sum_{\sigma \in \mathcal{F}_{K,ext}} |\sigma| d_{K,\sigma} \mathbf{n}_{K,\sigma} e_\sigma.$$

From the geometrical identity valid for any vector $\mathbf{x}_0, \mathbf{v} \in \mathbb{R}^d$ and for all $K \in \mathcal{M}$

$$\frac{1}{|K|} \sum_{\sigma \in \mathcal{F}_K} |\sigma| (\mathbf{x}_\sigma - \mathbf{x}_0) \mathbf{n}_{K,\sigma} \cdot \mathbf{v} = \mathbf{v}, \quad (60)$$

which is a direct consequence of Eq.(16), it follows that

$$|K|(\nabla_{\mathcal{D},\alpha} P_{\mathcal{D}} u)_K \leq \frac{1}{\theta} |K|(\alpha \nabla \underline{u})_K + R_K.$$

Due to flux consistency error estimates, it also follows that

$$\begin{aligned} |R_K| &\leq \sum_{L \in \mathcal{N}_K} |\sigma| d_{L,\sigma} |e_\sigma| + \sum_{\sigma \in \mathcal{F}_{K,ext}} |\sigma| d_{K,\sigma} |e_\sigma| \\ &\leq \frac{C_4}{\theta} h_{\mathcal{D}} \sum_{\sigma \in \mathcal{F}_K} |\sigma| d_{K,\sigma} = C_4 \frac{d}{\theta} |K| h_{\mathcal{D}}. \end{aligned} \quad (61)$$

As a consequence, one obtains that

$$\begin{aligned} &\sum_{K \in \mathcal{M}} |K| |(\nabla_{\mathcal{D},\alpha} P_{\mathcal{D}} u)_K - (\alpha \nabla \underline{u})_K|^2 \\ &\leq \sum_{K \in \mathcal{M}} C_4^2 \left(\frac{d}{\theta}\right)^2 h_{\mathcal{D}}^2 |K| = \left(C_4 \frac{d}{\theta}\right)^2 h_{\mathcal{D}}^2 |\Omega|. \end{aligned} \quad (62)$$

Using $\underline{u} \in C^2(\Omega)$ and α regularity, there exists C_5 , only dependent on L^∞ norm of the second derivatives of \underline{u} , such that

$$\sum_{K \in \mathcal{M}} \int_K |\alpha \nabla \underline{u} - (\alpha \nabla \underline{u})_K|^2 \leq C_5 h_{\mathcal{D}}^2. \quad (63)$$

From Eqs.(84) and (85), one gets the existence of C_c , only dependent on Ω , θ , \underline{u} and α , such that (59) holds.

In order to complete the convergence analysis, an upper bound in $\|\cdot\|_{\mathcal{D}}$ and the properties of weak consistency and convergence must also be proved for the Gauss gradient scheme appearing in the discrete bilinear form defined in Eq.(42). Remember that, as also argued in [24], even though consistent, the Gauss gradient scheme does not allow to obtain coercivity and hence uniqueness. Nevertheless, in practical implementations also the Gauss gradient scheme suffices in obtaining stable coercive diffusion operators, as widely verified in the numerical tests reported in Section 6. This is both a consequence of the limited non-orthogonality encountered in unstructured meshes used herein, and so a limited importance of the non-orthogonal correction term in the *Gauss corrected* scheme, and also a result of the action of $[u, v]_{\mathcal{D}, \alpha}$ term, that provides a consistent stabilization term to the bilinear form $\langle u, v \rangle_{\mathcal{D}, \alpha, \|\cdot\|}$.

It is now possible to prove convergence of the weak formulation associated to the *Gauss corrected* scheme, provided an assumption of the limited contribution from the mesh non-orthogonality correction term. In particular, one requires the condition of small gradient distortion

$$\sum_{K \in \mathcal{M}} |K| \nabla_{\mathcal{D}} u \cdot \nabla_{\mathcal{D}, 1} u \geq \sum_{K \in \mathcal{M}} |K| \nabla_{\mathcal{D}} u \cdot \nabla_{\mathcal{D}, 1, \|\cdot\|} u \quad (64)$$

which states that the diffusivity weighted discrete gradient with unit diffusivity $\nabla_{\mathcal{D}, 1} u$ has a preferential alignment with the Gauss (stabilized) gradient $\nabla_{\mathcal{D}} u$ with respect to the $\Gamma_{\alpha}^{\|\cdot\|}$ -weighted gradient $\nabla_{\mathcal{D}, 1, \|\cdot\|} u$ with unit diffusivity $\alpha = 1$. Notice that this sufficient condition for convergence is already known to finite volume practitioners, which usually require limited mesh non-orthogonality to have stable discretizations. In the present analysis, the role of condition (64) is made clear in proving the discrete $H^1(\Omega)$ estimate.

Lemma 5.5 (Discrete $H^1(\Omega)$ estimate). *Under assumption (64) and the hypotheses of the heterogeneous diffusion problem (1), let \mathcal{D} be an admissible finite volume discretization in sense of Definition (2.1) and let $0 < \theta \leq \hat{\theta}_{\mathcal{D}}$. Assume that $0 < \alpha_0 \leq \alpha(\mathbf{x})$ for a.e.*

$\mathbf{x} \in \Omega$ and also assume that $u \in H_{\mathcal{D}}$ is a solution of the discrete weak problem (43). Then the following estimate holds

$$\alpha_0 \|u\|_{\mathcal{D}} \leq \text{diam}(\Omega) \|f\|_{L^2(\Omega)}. \quad (65)$$

Proof. Consider the discrete weak formulation (43) and set $v = u$, to obtain

$$[u, u]_{\mathcal{D}, \alpha, \parallel} + \langle \nabla u, \nabla u \rangle_{\mathcal{D}, \alpha, \parallel} = \int_{\Omega} u(\mathbf{x}) f(\mathbf{x}) \, d\mathbf{x}.$$

After assumption (64) it follows that

$$\sum_{K \in \mathcal{M}} |K| \nabla_{\mathcal{D}} u \cdot \nabla_{\mathcal{D}, 1, \parallel} u \geq 0$$

which allows to conclude that

$$\langle \nabla u, \nabla u \rangle_{\mathcal{D}, \alpha, \parallel} = \sum_{K \in \mathcal{M}} \nabla_K u \cdot (\Gamma_{\alpha}^{\parallel} \nabla u)_K \geq \sum_{K \in \mathcal{M}} \alpha_0 \nabla_K u \cdot \nabla_{\mathcal{D}, 1, \parallel} u \geq 0,$$

from which one obtains

$$[u, u]_{\mathcal{D}, \alpha, \parallel} + \langle \nabla u, \nabla u \rangle_{\mathcal{D}, \alpha, \parallel} \geq [u, u]_{\mathcal{D}, \alpha, \parallel} \geq \alpha_0 \|u\|_{\mathcal{D}}^2. \quad (66)$$

From the Cauchy-Schwartz inequality and from the discrete Poincaré inequality (50) one also obtains

$$\begin{aligned} \int_{\Omega} u(\mathbf{x}) f(\mathbf{x}) \, d\mathbf{x} &\leq \|u\|_{L^2(\Omega)} \|f\|_{L^2(\Omega)} \\ &\leq \text{diam}(\Omega) \|u\|_{\mathcal{D}} \|f\|_{L^2(\Omega)}. \end{aligned} \quad (67)$$

Combining together Eqs.(66) and (67) allows to recover the discrete estimate (65).

Corollary 5.6 (Existence and uniqueness of a discrete solution). *Assume (64) and the hypotheses of problem (1). Let \mathcal{D} be an admissible finite volume discretization in sense of Definition (2.1) and let $0 < \theta \leq \tilde{\theta}_{\mathcal{D}}$. Then, there exists a unique solution to problem (43).*

Proof. Assume $f = 0$ in the finite dimensional system (43). From the discrete Poincaré inequality (50) one gets $u = 0$, thus proving that the linear problem (43) is uniquely solvable.

Finally, it is possible to state the convergence of the finite volume *Gauss corrected* scheme to the solution of the associated weak problem (43).

Theorem 5.7 (Convergence of *Gauss corrected* scheme). *Assuming the sufficient condition (64) and the hypotheses of the heterogeneous diffusion problem (1), let \mathcal{D} be an admissible finite volume discretization in sense of Definition (2.1) with $0 < \theta \leq \tilde{\theta}_{\mathcal{D}}$. Assume that $0 < \alpha_0 \leq \alpha(\mathbf{x})$ for a.e. $\mathbf{x} \in \Omega$ and also assume that $u \in H_{\mathcal{D}}$ is a solution to the discrete weak problem (43). Then u converges in $L^2(\Omega)$ to \underline{u} , that is the weak solution to problem (1) in the sense of (26), as $h_{\mathcal{D}} \rightarrow 0$.*

Proof. The convergence proof uses the compactness technique presented in [26]. Consider a subsequence of admissible discretizations $(\mathcal{D}_n)_{n \in \mathbb{N}}$ such that $h_{\mathcal{D}} \rightarrow 0$ as $n \rightarrow \infty$ while $\theta_{\mathcal{D}_n} \geq \theta$ for all $n \in \mathbb{N}$. Using Lemma 5.5 one can apply Lemma 5.1, which is a discrete counterpart of the Rellich theorem and gives the existence of a subsequence, for simplicity denoted again with $(\mathcal{D}_n)_{n \in \mathbb{N}}$, and of some $\underline{u} \in H_0^1(\Omega)$ such that the solution $u_{\mathcal{D}_n}$ to problem (43) tends to \underline{u} in $L^2(\Omega)$ as $n \rightarrow \infty$. Let $\varphi \in C_c^\infty(\Omega)$ and select $v = P_{\mathcal{D}}\varphi$ as test function in problem (43), from which one has

$$[u, P_{\mathcal{D}_n}\varphi]_{\mathcal{D}_n, \alpha, \parallel} + \langle \nabla u, \nabla P_{\mathcal{D}_n}\varphi \rangle_{\mathcal{D}_n, \alpha, \parallel} = \int_{\Omega} f(\mathbf{x}) P_{\mathcal{D}_n}\varphi \, d\mathbf{x}. \quad (68)$$

Let then $n \rightarrow \infty$ in Eq.(68). Thanks to Lemma 5.3 and Lemma 5.4, but also to their counterparts for the $\nabla_{\mathcal{D}}u$ in Appendix 8, considering the decomposition

$$\langle \nabla u, \nabla u \rangle_{\mathcal{D}, \alpha, \parallel} = \langle \nabla u, \nabla u \rangle_{\mathcal{D}, \alpha} - \langle \nabla u, \nabla u \rangle_{\mathcal{D}, \alpha, \parallel},$$

one obtains the convergence of the diffusivity weighted gradient portion $\langle \nabla u, \nabla u \rangle_{\mathcal{D}, \alpha}$ contained within $\langle \nabla u, \nabla u \rangle_{\mathcal{D}, \alpha, \parallel}$ term

$$\lim_{n \rightarrow \infty} \int_{\Omega} \nabla_{\mathcal{D}_n} u \cdot (\Gamma_{\alpha} \nabla P_{\mathcal{D}_n}\varphi) \, d\mathbf{x} = \int_{\Omega} \nabla \underline{u} \cdot (\Gamma_{\alpha}) \nabla \varphi \, d\mathbf{x}.$$

From Lemma 5.1 and Lemma 5.3 one also obtains that the sum of the remaining terms in problem (43) is such that

$$\lim_{n \rightarrow \infty} \int_{\Omega} \left([u, P_{\mathcal{D}_n}\varphi]_{\mathcal{D}_n, \alpha, \parallel} - \nabla_{\mathcal{D}_n} u \cdot (\Gamma_{\alpha}^{\parallel} \nabla P_{\mathcal{D}_n}\varphi) \, d\mathbf{x} \right) = 0.$$

Due to the fact that

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

one gets that any limit \underline{u} of a subsequence of solutions satisfies the weak problem (26) with $v = \varphi$. Uniqueness of the solution to (26) together with a classical density argument allow to deduce the convergence of the whole sequence u to the weak problem solution \underline{u} in $L^2(\Omega)$ as $h_{\mathcal{D}} \rightarrow 0$, since $\theta \leq \tilde{\theta}_{\mathcal{D}}$.

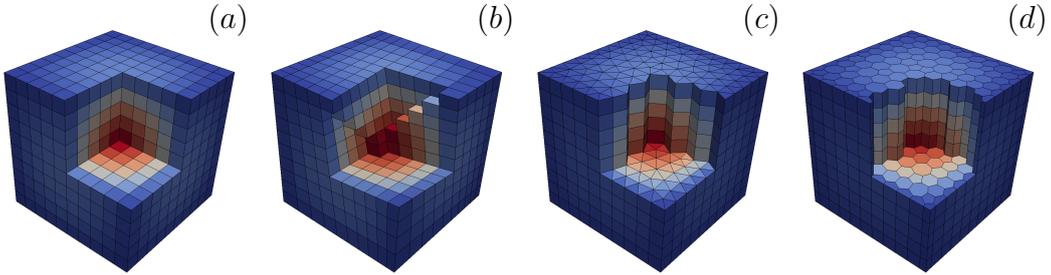


Figure 2: Different mesh types used in the numerical test: (a) orthogonal hexahedral, (b) skewed hexahedral, (c) triangular prismatic, (d) polygonal prismatic.

6 Numerical results

The *Gauss corrected* finite volume diffusion scheme in Eq.(12) was tested on a number of different mesh types commonly adopted in industrial applications. These include meshes composed of regular orthogonal hexahedra, skewed hexahedra, triangular prismatic and polygonal prismatic cells. All these meshes were constructed by means of a commercial finite volume mesh generator [6], which usually produces meshes of acceptable non-orthogonality, as commonly required in practical applications. In particular, the polygonal prismatic mesh was obtained after geometric dualization of the triangular prismatic one. Notice also that finer meshes are not produced by conformal refinement techniques, but generated *ex novo*. It should be remarked again that the *Gauss corrected* scheme reduces to the two-point flux approximation on orthogonal meshes.

The different mesh types are summarized in Table 1, where relevant quantities are reported. These include the parameters normally observed as quality indices after the mesh generation process, which are:

- *Non-orthogonality*, measured by the angle between the line segment $(\mathbf{x}_L - \mathbf{x}_K)$, joining cell centroids adjacent to face $\sigma \in \mathcal{F}_K$, and the face normal $\mathbf{n}_{K,\sigma}$, that is

$$\theta_\sigma = \arccos \left(\frac{(\mathbf{x}_L - \mathbf{x}_K) \cdot \mathbf{n}_{K,\sigma}}{|\mathbf{x}_L - \mathbf{x}_K|} \right),$$

$$\forall \sigma \in \mathcal{F}_{K,int}, \forall K \in \mathcal{M}. \quad (69)$$

A value close to 0 is optimal, since it reduces the amount of non-orthogonal correction with respect to the two-point flux approximation, see, e.g., Eq.(12). Here, both the mean non-orthogonality angle $\langle \theta_\sigma \rangle = \frac{1}{\#\mathcal{F}_{int}} \sum_{\sigma \in \mathcal{F}_{int}} \theta_\sigma$ and the maxi-

$Mesh$	$\langle d \rangle_\Omega$	$\langle \theta_\sigma \rangle$	θ_{\max}	AR_{\max}	S_{\max}
hex	1.0000×10^{-1}	0	0	1	0
	5.0000×10^{-2}	0	0	1	0
	2.5000×10^{-2}	0	0	1	0
	1.2500×10^{-2}	0	0	1	0
hexSkew	1.0033×10^{-1}	9.017	15.138	2.085	0.145
	5.0265×10^{-2}	8.982	15.392	2.352	0.162
	2.5160×10^{-2}	8.953	15.531	2.520	0.175
	1.2587×10^{-2}	8.935	15.588	2.615	0.181
triPrism	7.4413×10^{-2}	3.396	11.059	2.963	0.221
	3.7505×10^{-2}	2.862	10.651	2.969	0.218
	1.8748×10^{-2}	2.780	10.652	3.146	0.241
	9.3825×10^{-3}	2.574	11.008	3.176	0.239
polyPrism	9.7486×10^{-2}	5.246	14.912	2.790	0.741
	4.9855×10^{-2}	3.945	14.587	2.778	0.739
	2.5063×10^{-2}	3.416	15.082	2.777	0.739
	1.2582×10^{-2}	3.022	14.802	2.777	0.739

Table 1: Main geometric parameters for the different mesh types used in the accuracy test.

mum non-orthogonality angle $\theta_{\max} = \max_{\sigma \in \mathcal{F}_{int}} \theta_\sigma$ are considered.

- *Aspect ratio*, defined for each cell $K \in \mathcal{M}$ as

$$AR_K = \max \left\{ AR_{BB(K)}, \frac{\sum_{i=1}^d |\sigma_i|}{6|K|^{2/3}} \right\}, \quad \forall K \in \mathcal{M}, \quad (70)$$

where AR_K is the bounding box aspect ratio

$$AR_{BB(K)} = \frac{\max_{\sigma_i \in \mathcal{F}_{BB(K)}} \{|\sigma_i|\}}{\min_{\sigma_i \in \mathcal{F}_{BB(K)}} \{|\sigma_i|\}}, \quad \forall K \in \mathcal{M}, \quad (71)$$

defined in terms of the cell bounding box $BB(K)$ which encloses the cell K with a set of faces σ_i ($i = 1, \dots, d$) having normals oriented along the axes of the Cartesian reference frame used for the mesh definition. A value close to 1 indicates that the cell is isotropic. Mesh statistics generally consider the maximum value of the cell aspect ratio $AR_{\max} = \max_{K \in \mathcal{M}} AR_K$.

- *Skewness*, defined as the distance between the intersection point $\mathbf{y}_\sigma = [\mathbf{x}_L, \mathbf{x}_K] \cap \bar{\sigma}$ between the line segment $(\mathbf{x}_L - \mathbf{x}_K)$ connecting adjacent cell centroids and separating face $\sigma \in \mathcal{F}_{int}$ and the face centroid \mathbf{x}_σ , that is

$$S_\sigma = \frac{|(\mathbf{x}_\sigma - \mathbf{y}_\sigma)|}{f_\sigma}, \quad \forall \sigma \in \mathcal{F}, \quad (72)$$

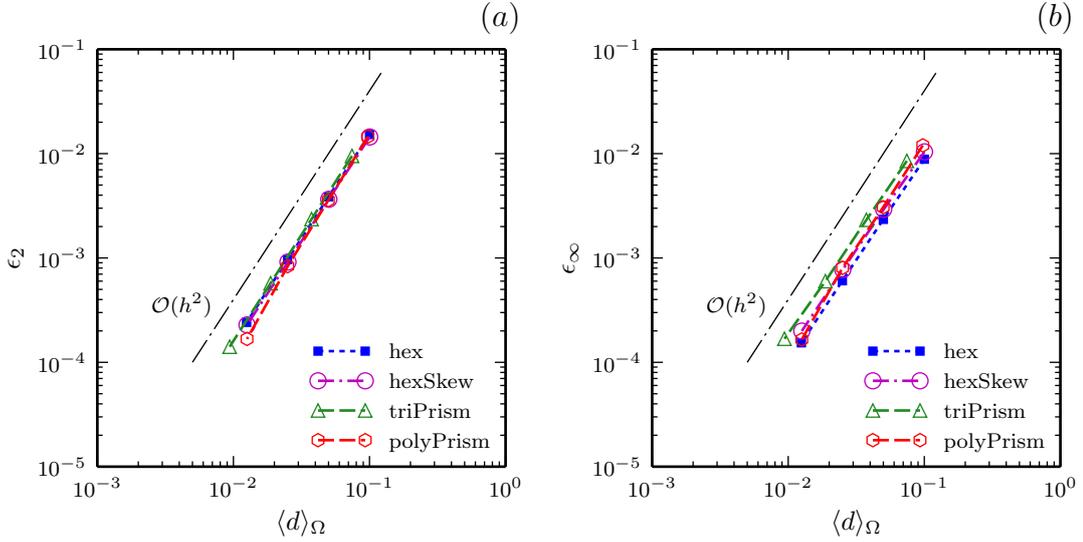


Figure 3: Relative error curves for different mesh types: (a) L^2 relative error $\epsilon_2 = \|e\|_2 / \|\underline{u}\|_2$, (b) L^∞ relative error $\epsilon_2 = \|e\|_\infty / \|\underline{u}\|_\infty$.

where the normalization factor f_σ is

$$f_\sigma = \max \{0.2 |(\mathbf{x}_L - \mathbf{x}_K)|, \max_{i \in \mathcal{V}_\sigma} \left| (\mathbf{x}_i - \mathbf{x}_\sigma) \cdot \frac{(\mathbf{x}_\sigma - \mathbf{y}_\sigma)}{|\mathbf{x}_\sigma - \mathbf{y}_\sigma|} \right| \}, \quad \forall \sigma \in \mathcal{F}_{int} \quad (73a)$$

$$f_\sigma = \max \{0.4 |(\mathbf{y}_\sigma - \mathbf{x}_K)|, \max_{i \in \mathcal{V}_\sigma} \left| (\mathbf{x}_i - \mathbf{x}_\sigma) \cdot \frac{(\mathbf{x}_\sigma - \mathbf{y}_\sigma)}{|\mathbf{x}_\sigma - \mathbf{y}_\sigma|} \right| \}, \quad \forall \sigma \in \mathcal{F}_{ext}. \quad (73b)$$

The optimal value for S_σ is 0, indicating that $\mathbf{y}_\sigma = \mathbf{x}_\sigma$, for which linear interpolation between adjacent cell values achieves second order consistency in face integral quantities. Mesh statistics generally take into account the maximum value of skewness $S_{\max} = \max_{\sigma \in \mathcal{F}} S_\sigma$.

The mesh resolution is measured by the mean magnitude of the cell to cell distance, that is $\langle d \rangle_\Omega = \langle |(\mathbf{x}_K - \mathbf{x}_L)| \rangle_{K|L \in \mathcal{F}_{int}}$.

Numerical experiments were carried out on $\Omega = (0, 1) \times (0, 1) \times (0, 1)$ assuming $\alpha(\mathbf{x}) = 1$ and considering the exact solution of problem (1) given by $\underline{u}(x_1, x_2, x_3) = x_1(1 - x_1)x_2(1 - x_2)x_3(1 - x_3)$. On each mesh, the error is measured as $e(\mathbf{x}_K) = u_K - \underline{u}(\mathbf{x}_K)$ for $K \in \mathcal{M}$ and it allows to estimate empirically the rate of convergence

$Mesh$	$\langle d \rangle_\Omega$	$\ e\ _2$	$\ e\ _\infty$	p_2	p_∞
hex	1.0000×10^{-1}	9.2721×10^{-5}	1.3410×10^{-4}	1.984	1.882
	5.0000×10^{-2}	2.3439×10^{-5}	3.6372×10^{-5}	1.996	1.946
	2.5000×10^{-2}	5.8771×10^{-6}	9.4379×10^{-6}	1.999	1.975
hexSkew	1.2500×10^{-2}	1.4704×10^{-6}	2.4011×10^{-6}	–	–
	1.0033×10^{-1}	8.7858×10^{-5}	1.5863×10^{-4}	1.988	1.809
	5.0265×10^{-2}	2.2232×10^{-5}	4.5431×10^{-5}	2.001	1.914
triPrism	2.5160×10^{-2}	5.5665×10^{-6}	1.2077×10^{-5}	2.004	1.957
	1.2587×10^{-2}	1.3890×10^{-6}	3.1135×10^{-6}	–	–
	7.4413×10^{-2}	5.7868×10^{-5}	1.3136×10^{-4}	2.036	1.886
polyPrism	3.7505×10^{-2}	1.4340×10^{-5}	3.6076×10^{-5}	2.040	1.946
	1.8748×10^{-2}	3.4852×10^{-6}	9.3566×10^{-6}	2.025	1.835
	9.3825×10^{-3}	8.5803×10^{-7}	2.6272×10^{-6}	–	–
polyPrism	9.7486×10^{-2}	8.8853×10^{-5}	1.8380×10^{-4}	2.096	2.011
	4.9855×10^{-2}	2.1796×10^{-5}	4.7717×10^{-5}	2.118	1.941
	2.5063×10^{-2}	5.0781×10^{-6}	1.2557×10^{-5}	2.330	2.292
	1.2582×10^{-2}	1.0198×10^{-6}	2.5873×10^{-6}	–	–

Table 2: Error behaviour in the numerical convergence test.

between two successive mesh sizes. The *Gauss corrected* scheme is implemented with the deferred correction approach, with the non-orthogonal correction term implemented explicitly and thus requiring outer iterations which are terminated with a tolerance level of 1.0×10^{-4} . The associated linear system is solved by a preconditioned conjugate gradient method with tolerance 1.0×10^{-16} and DIC preconditioning.

The error norms $\|e\|_2$ and $\|e\|_\infty$ together with the corresponding empirical orders of convergence p_2 and p_∞ are reported in Table 2, while the relative errors norms $\epsilon_2 = \|e\|_2/\|\underline{u}\|_2$ and $\epsilon_\infty = \|e\|_\infty/\|\underline{u}\|_\infty$ are shown in Figure 3. From both quantities, it is evident that second order accuracy is empirically verified for hexahedral, skewed hexahedral, triangular prismatic and polyhedral prismatic mesh types. It is remarkable that the accuracy of the *Gauss corrected* scheme appears insensitive to the cell shape, with only minimal differences in the infinity norm.

7 Beyond the Gauss gradient scheme

The Gauss discrete gradient operator $\nabla_{\mathcal{D}}$ that was introduced in Eq.(13) is bounded, weakly convergent and consistent, but generally it is not coercive. Thus, on strongly non-orthogonal meshes, the correction term in the *Gauss corrected* approach may not be coercive and consequently hamper the convergence of the finite volume

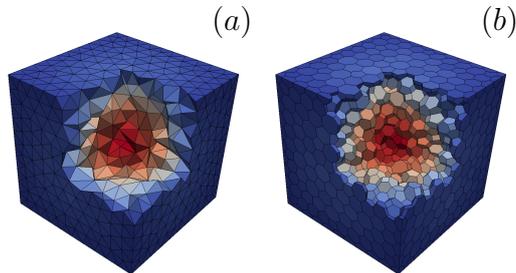


Figure 4: Strongly non-orthogonal mesh types used in the numerical test: (a) tetrahedral and (b) polyhedral meshes.

scheme.

To verify this point empirically, the same diffusion problem studied empirically in section 6 is now solved on a sequence of highly non-orthogonal tetrahedral and polyhedral meshes, see Figure 4, whose geometric parameters are summarized in Table 3. It is important to notice that the maximum non-orthogonality angle is such that $\theta_{\max} > \pi/4$ almost for every mesh, with the only exception of the two coarsest polyhedral meshes. This implies that the non-orthogonal correction term is the dominant term in the numerical flux.

The *Gauss corrected* scheme can still be applied, provided that a coercive gradient scheme is adopted for the non-orthogonal correction term. To this end, we assess here the performance a gradient approximation based on a least square fit, based on the fact that in linear upwind schemes it is known empirically to provide a coercive gradient discretization in the case of highly non-orthogonal tetrahedral meshes. Notice that, on orthogonal meshes it reduces to the Gauss scheme, hence becoming non-coercive. But this is of no concern as long as it is adopted only for the construction of the non-orthogonal correction term in the *Gauss corrected* fluxes.

Following [8], on a non-orthogonal mesh like that of Definition 2.1, we define the discrete gradient operator $\nabla_{\mathcal{D}}^{LS} : H_{\mathcal{D}}(\Omega) \rightarrow H_{\mathcal{D}}(\Omega)^d$ as the piecewise constant function

$$\nabla_K^{LS} u = \sum_{\sigma \in \mathcal{F}_{K,int}} (u_L - u_K) \mathbf{v}_{K,\sigma} + \sum_{\sigma \in \mathcal{F}_{K,ext}} (u_\sigma - u_K) \mathbf{v}_{K,\sigma} \quad (74)$$

for $u \in H_{\mathcal{D}}(\Omega)$, where the least squares vectors

$$\mathbf{v}_{K,\sigma} = w_{K,\sigma} \mathbf{W}_K^{-1} (\mathbf{x}_L - \mathbf{x}_K), \quad \forall \sigma \in \mathcal{F}_{K,int} \quad (75a)$$

$$\mathbf{v}_{K,\sigma} = w_{K,\sigma} \mathbf{W}_K^{-1} (\mathbf{x}_\sigma - \mathbf{x}_K), \quad \forall \sigma \in \mathcal{F}_{K,ext} \quad (75b)$$

<i>Mesh</i>	$\langle d \rangle_\Omega$	$\langle \theta_\sigma \rangle$	θ_{\max}	AR_{\max}	S_{\max}
tet	4.2582×10^{-2}	19.058	59.778	5.614	0.569
	2.1113×10^{-2}	19.307	63.201	7.166	0.743
	1.1023×10^{-2}	19.597	65.714	8.468	0.926
	5.5488×10^{-3}	19.741	66.510	8.370	0.901
poly	1.0481×10^{-1}	11.688	38.598	4.592	1.062
	5.3575×10^{-2}	11.801	40.406	3.188	1.134
	2.8269×10^{-2}	11.925	47.784	4.048	1.234
	1.4341×10^{-2}	11.944	50.781	4.188	1.439

Table 3: Main geometric parameters for the highly non-orthogonal mesh types used in the accuracy test.

are defined from the weighting tensor

$$\begin{aligned} \mathbf{W}_K = & \sum_{\sigma \in \mathcal{F}_{K,int}} w_{K,\sigma} (\mathbf{x}_L - \mathbf{x}_K) (\mathbf{x}_L - \mathbf{x}_K)^\top \\ & + \sum_{\sigma \in \mathcal{F}_{K,ext}} w_{K,\sigma} (\mathbf{x}_\sigma - \mathbf{x}_K) (\mathbf{x}_\sigma - \mathbf{x}_K)^\top, \quad \forall K \in \mathcal{M} \end{aligned} \quad (76)$$

as well as from the face weights $w_{K,\sigma}$, $\forall \sigma \in \mathcal{F}_K$ and $\forall K \in \mathcal{M}$. Different expressions for the face weights can be adopted. Here, we use the formulae

$$w_{K,\sigma} = \frac{d_{K,\sigma}}{d_{K,L}} \frac{|\sigma|}{|\mathbf{x}_L - \mathbf{x}_K|^2}, \quad \forall \sigma \in \mathcal{F}_{K,int} \quad (77a)$$

$$w_{K,\sigma} = \frac{|\sigma|}{|\mathbf{x}_L - \mathbf{x}_K|^2}, \quad \forall \sigma \in \mathcal{F}_{K,ext}. \quad (77b)$$

The least squares gradient scheme is empirically constructed from the approximate Taylor expansion at adjacent cell centroids and face centroids

$$u(\mathbf{x}_L) \approx u_K + \nabla_K^{LS} u \cdot (\mathbf{x}_L - \mathbf{x}_K), \quad \forall \sigma \in \mathcal{F}_{K,int}, \quad (78a)$$

$$u(\mathbf{x}_\sigma) \approx u_K + \nabla_K^{LS} u \cdot (\mathbf{x}_\sigma - \mathbf{x}_K), \quad \forall \sigma \in \mathcal{F}_{K,ext}, \quad (78b)$$

requiring the minimization of the piecewise constant mean-square-error objective function

$$\begin{aligned} G_K = & \sum_{\sigma \in \mathcal{F}_{K,int}} w_{K,\sigma} (u_L - u_K - \nabla_K^{LS} u \cdot (\mathbf{x}_L - \mathbf{x}_K))^2 \\ & + \sum_{\sigma \in \mathcal{F}_{K,ext}} w_{K,\sigma} (u_\sigma - u_K - \nabla_K^{LS} u \cdot (\mathbf{x}_\sigma - \mathbf{x}_K))^2. \end{aligned}$$

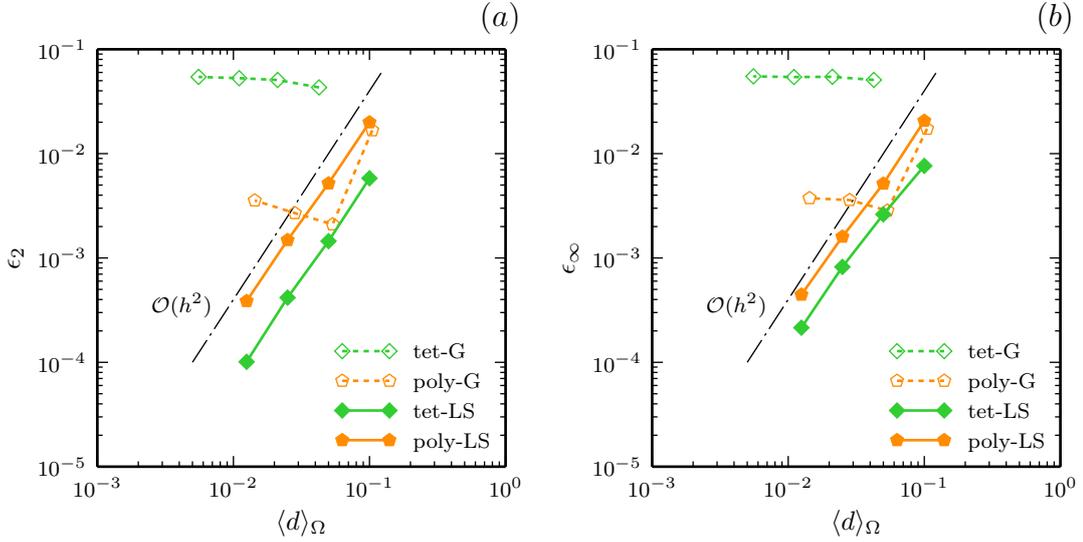


Figure 5: Relative error curves for non-orthogonal mesh types: (a) L^2 relative error $\epsilon_2 = \|e\|_2/\|\underline{u}\|_2$, (b) L^∞ relative error $\epsilon_2 = \|e\|_\infty/\|\underline{u}\|_\infty$. The *Gauss* (G) gradient scheme is assessed against the *leastSquares* (LS) scheme.

Mesh	Grad	$\langle d \rangle_\Omega$	$\ e\ _2$	$\ e\ _\infty$	p_2	p_∞
tet	G	4.2582×10^{-2}	2.6231×10^{-4}	7.9063×10^{-4}	-0.234	-0.104
	G	2.1113×10^{-2}	3.0907×10^{-4}	8.5032×10^{-4}	-0.062	0.007
	G	1.1023×10^{-2}	3.2172×10^{-4}	8.4668×10^{-4}	-0.039	-0.022
	G	5.5488×10^{-3}	3.3056×10^{-4}	8.5973×10^{-4}	-	-
tet	LS	4.2582×10^{-2}	3.5338×10^{-5}	1.1860×10^{-4}	1.978	1.518
	LS	2.1113×10^{-2}	8.8216×10^{-6}	4.0872×10^{-5}	1.917	1.782
	LS	1.1023×10^{-2}	2.5376×10^{-6}	1.2836×10^{-5}	2.064	1.955
	LS	5.5488×10^{-3}	6.1543×10^{-7}	3.3552×10^{-6}	-	-
poly	G	1.0481×10^{-1}	1.0196×10^{-4}	2.6907×10^{-4}	3.094	2.685
	G	5.3575×10^{-2}	1.2782×10^{-5}	4.4394×10^{-5}	-0.388	-0.366
	G	2.8269×10^{-2}	1.6381×10^{-5}	5.6089×10^{-5}	-0.411	-0.061
	G	1.4341×10^{-2}	2.1646×10^{-5}	5.8460×10^{-5}	-	-
poly	LS	1.0481×10^{-1}	1.2148×10^{-4}	3.2076×10^{-4}	2.009	2.061
	LS	5.3575×10^{-2}	3.1556×10^{-5}	8.0425×10^{-5}	1.957	1.830
	LS	2.8269×10^{-2}	9.0307×10^{-6}	2.4964×10^{-5}	1.983	1.886
	LS	1.4341×10^{-2}	2.3513×10^{-6}	6.9417×10^{-6}	-	-

Table 4: Error behaviour for *Gauss* (G) and *leastSquares* (LS) gradient schemes in the numerical convergence test on highly non-orthogonal meshes.

The results of the comparison between the Gauss and least squares

gradient schemes in the construction of the non-orthogonal term inside the *Gauss corrected* diffusion scheme are reported in Figure 5 and Table 4. On the strongly non-orthogonal tetrahedral meshes, the Gauss gradient does not allow to obtain a coercive numerical flux and leads to stagnation. The same discrete gradient scheme converges only on the first two coarser polyhedral meshes, while it diverges again on the two finest polyhedral meshes. On the contrary, when the non-orthogonal correction term of the *Gauss corrected* scheme is constructed from the least squares gradient scheme, a convergent second order behaviour is recovered on all the highly non-orthogonal meshes considered here. Even though in the present investigation no analytical results have been obtained for this discrete gradient scheme, it appears to be able to overcome the main limitation of the Gauss discrete gradient on strongly non-orthogonal meshes.

8 Conclusions

In this work, we have proven the convergence of the *Gauss corrected* scheme on unstructured meshes satisfying a global and rather weak mesh regularity condition. This goal has been achieved adapting the approach used in [22] for the convergence analysis of a cell-centered finite volume scheme for anisotropic diffusion problems on orthogonal meshes. We have also shown empirically that a least square approach to the gradient computation can provide second order convergence even when the mild mesh regularity condition is violated. To the best of the authors' knowledge, the convergence properties of the finite volume method analyzed here have never been studied rigorously in the case of non-orthogonal meshes. Indeed, convergence analyses of finite volume schemes for diffusion operators on unstructured mesh types are usually limited to polyhedral meshes satisfying an orthogonality condition [21], [22]. This is quite restrictive in practice, since none of the robust mesh generators usually adopted for pre-processing of industrial configurations are able to guarantee this condition.

From our development, it can be seen how the analysis of finite volume schemes is greatly simplified if it is approached from the associated discrete weak formulation using the functional tools defined in [26]. In particular, from this approach several interesting conclusions can be drawn, without recourse to the classical consistency analysis in terms of Taylor series expansion. In the case of industrial finite volume schemes, such as the one analyzed here, these conclusions are particularly interesting, because they shed some light over

the properties of techniques for which typically only empirical results are available. The role of the discrete gradient scheme is fundamental in many terms of the associated weak form, and its relevance is also reflected in the finite volume formulation, even though it may not be completely apparent when starting directly from the flux balance equations. By working with the weak form, conclusions about the coercivity of the finite volume scheme can be drawn, which are often considered inaccessible in the finite volume framework. Relevant discrete functional analysis results can be applied directly, in order to establish conditions for convergence and eventually to obtain error estimates under sufficient regularity assumptions. In particular, this approach allows to identify the mesh regularity requirements and the sufficient conditions for convergence, as well as to suggest possible future improvements.

The basic idea of correcting the two-point flux approximation with an additional term accounting for the local mesh non - orthogonality can also be found in many other finite volume schemes [13], [35], [36], [37], [38], [39], [41], and thus similar analysis techniques could also be applied to investigate sufficient conditions for convergence of other diffusion schemes. Connections with the asymmetric gradient schemes recently proposed in [17] also suggest a possible alternative to the present analysis. Finally, the diffusion operator analyzed in this work was used in [14] as the basis for the construction of accurate and efficient stabilized pressure correction methods for colocated finite volume schemes. The properties and advantages of these methods will be discussed in a series of forthcoming companion papers.

Acknowledgements

The results presented in this work are part of the doctoral thesis in Applied Mathematics [14] discussed by A.D.R. at Politecnico di Milano in 2018. The comments on the thesis by D. Di Pietro and J. Szmelter are kindly acknowledged. A.D.R. would also like to thank Tenova S.p.A. for sponsoring his Executive PhD at Politecnico di Milano and all the faculty members at MOX for their support.

Appendix: Properties of the Gauss gradient scheme

The Gauss gradient operator $\nabla_{\mathcal{D}}$ introduced in Eq.(13) is bounded, weakly convergent and consistent. In order to prove these properties, it is sufficient to rewrite it in the form

$$\nabla_K u = \frac{1}{|K|} \sum_{\sigma \in \mathcal{F}_K} |\sigma| \frac{d_{K,\sigma}}{d_{K,L}} (u_L - u_K) \mathbf{n}_{K,\sigma} \quad (79)$$

which allows to prove that is is bounded in the $L^2(\Omega)^d$ -norm.

Lemma 8.1 (Bound on $\nabla_{\mathcal{D}}u$). *Let Ω be a bounded open connected polyhedral subset of \mathbb{R}^d , $d \in \mathbb{N}^*$. Let \mathcal{D} be an admissible finite volume discretization in sense of Definition 2.1 and let $0 < \theta \leq \tilde{\theta}_{\mathcal{D}}$. Then, there exists C depending only on d , α and θ such that, for all $u \in H_{\mathcal{D}}$, one has*

$$\|\nabla_{\mathcal{D}}u\|_{L^2(\Omega)^d} \leq C \|u\|_{\mathcal{D}}. \quad (80)$$

Proof. Let $u \in H_{\mathcal{D}}$. As in Lemma 5.2, one introduces, for all $K \in \mathcal{M}$, $L \in \mathcal{N}_K$ and $\sigma = K|L$, the difference quantities $\delta_{K,\sigma} \mathbf{x}$ and $\delta_{K,\sigma} u$, from which the inner product norm in (49) leads for a given $K \in \mathcal{M}$ to

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

Then, from Eq.(79) one obtains that

$$|K| (\nabla u)_K = \sum_{\sigma \in \mathcal{F}_K} \tau_{\sigma} d_{K,\sigma} \delta_{K,\sigma} u.$$

By using the Cauchy-Schwartz inequality, one obtains that

$$\begin{aligned} |K|^2 |(\nabla u)_K|^2 &\leq \sum_{\sigma \in \mathcal{F}_K} \tau_{\sigma} |d_{K,\sigma} \mathbf{n}_{K,\sigma}|^2 \sum_{\sigma \in \mathcal{F}_K} \tau_{\sigma} (\delta_{K,\sigma} u)^2 \\ &\leq \sum_{\sigma \in \mathcal{F}_K} d |D_{K,\sigma}| \tau_{\sigma} \sum_{\sigma \in \mathcal{F}_K} \tau_{\sigma} (\delta_{K,\sigma} u)^2 \\ &= d |K| \sum_{\sigma \in \mathcal{F}_K} \tau_{\sigma} (\delta_{K,\sigma} u)^2. \end{aligned}$$

Summing over all $K \in \mathcal{M}$, one obtains that

$$\begin{aligned} \sum_{K \in \mathcal{M}} |K| |\nabla_K u|^2 &\leq d \sum_{K \in \mathcal{M}} \sum_{\sigma \in \mathcal{F}_K} \tau_\sigma (\delta_{K,\sigma} u)^2 \\ &\leq d \sum_{K \in \mathcal{M}} \left(\sum_{\sigma \in \mathcal{F}_{K,int}} \tau_\sigma (\delta_{K,\sigma} u)^2 + 2 \sum_{\sigma \in \mathcal{F}_{K,ext}} \tau_\sigma (\delta_{K,\sigma} u)^2 \right) \\ &= 2d \|u\|_{\mathcal{D}}^2 \end{aligned}$$

from which (58) follows with $C = \sqrt{2d}$.

First, the weak convergence of $\nabla_{\mathcal{D},\alpha} u$ will be studied, while successively a similar result will be obtained for $\nabla_{\mathcal{D},\alpha,\|} u$.

Lemma 8.2 (Weak convergence of $\nabla_{,\alpha} u$). *Let Ω be a bounded open connected polyhedral subset of \mathbb{R}^d , $d \in \mathbb{N}^*$. Let \mathcal{D} be an admissible finite volume discretization in sense of Definition (2.1) and let $0 < \theta \leq \tilde{\theta}_{\mathcal{D}}$. Assume that there exists $u \in H_{\mathcal{D}}$ and a function $\underline{u} \in H_0^1(\Omega)$ such that u tends to \underline{u} in $L^2(\Omega)$ as $h_{\mathcal{D}} \rightarrow 0$, while $\|u\|_{\mathcal{D}}$ remains bounded. Then $\nabla_{\mathcal{D}} u$ weakly converges to $\alpha \nabla \underline{u}$ in $L^2(\Omega)^d$ as $h_{\mathcal{D}} \rightarrow 0$.*

Proof. Let $\varphi \in C_c^\infty(\Omega)$. Assume that $h_{\mathcal{D}}$ is small enough that, for all $K \in \mathcal{M}$ and $\mathbf{x} \in K$, if $\varphi(\mathbf{x}) \neq 0$ then $\mathcal{F}_{K,ext} = \emptyset$. Consider the term $T_1^{\mathcal{D}}$ defined as

$$\begin{aligned} T_1^{\mathcal{D}} &= \int_{\Omega} P_{\mathcal{D}} \varphi(\mathbf{x}) \nabla_{\mathcal{D}} u(\mathbf{x}) \, d\mathbf{x} = \sum_{K \in \mathcal{M}} |K| \varphi(\mathbf{x}_K) \nabla_K u \\ &= \sum_{K|L \in \mathcal{F}_{int}} (\varphi(\mathbf{x}_K) d_{K,\sigma} + \varphi(\mathbf{x}_L) d_{L,\sigma}) \alpha_{K|L} \tau_{K|L} \mathbf{n}_{K,\sigma} (u_L - u_K). \end{aligned}$$

The first term between brackets can be rewritten as

$$\begin{aligned} &\varphi(\mathbf{x}_K) d_{K,\sigma} + \varphi(\mathbf{x}_L) d_{L,\sigma} \\ &= \varphi(\mathbf{x}_K) (\mathbf{x}_\sigma - \mathbf{x}_K) \cdot \mathbf{n}_{K,\sigma} + \varphi(\mathbf{x}_L) (\mathbf{x}_L - \mathbf{x}_\sigma) \cdot \mathbf{n}_{K,\sigma} \\ &= \left(\frac{\varphi(\mathbf{x}_K) + \varphi(\mathbf{x}_L)}{2} (\mathbf{x}_L - \mathbf{x}_K) \right. \\ &\quad \left. + (\varphi(\mathbf{x}_K) - \varphi(\mathbf{x}_L)) \left(\mathbf{x}_\sigma - \frac{\mathbf{x}_K + \mathbf{x}_L}{2} \right) \right) \cdot \mathbf{n}_{K,\sigma}. \end{aligned}$$

The term $T_1^{\mathcal{D}}$ can be decomposed into a sum of two terms $T_1^{\mathcal{D}} =$

$T_2^{\mathcal{D}} + T_3^{\mathcal{D}}$, where

$$T_2^{\mathcal{D}} = \sum_{K|L \in \mathcal{F}_{int}} \tau_{K|L} (u_L - u_K) \mathbf{n}_{K,\sigma} d_{K,L} \frac{\varphi(\mathbf{x}_K) + \varphi(\mathbf{x}_L)}{2},$$

$$T_3^{\mathcal{D}} = \sum_{K|L \in \mathcal{F}_{int}} \tau_{K|L} (u_L - u_K) \mathbf{n}_{K,\sigma} \left(\mathbf{x}_\sigma - \frac{\mathbf{x}_K + \mathbf{x}_L}{2} \right) \cdot \mathbf{n}_{K,\sigma}.$$

Starting with the analysis of term $T_3^{\mathcal{D}}$, by Cauchy-Schwartz and then triangle inequalities one gets

$$(T_3^{\mathcal{D}})^2 \leq \sum_{K|L \in \mathcal{F}_{int}} \tau_{K|L} (u_L - u_K)^2$$

$$\times \sum_{K|L \in \mathcal{F}_{int}} \tau_{K|L} (\varphi(\mathbf{x}_K) - \varphi(\mathbf{x}_L))^2 \times \left| \left(\mathbf{x}_\sigma - \frac{\mathbf{x}_K + \mathbf{x}_L}{2} \right) \cdot \mathbf{n}_{K,\sigma} \right|^2$$

in which, due to triangle inequality

$$\left| \left(\mathbf{x}_\sigma - \frac{\mathbf{x}_K + \mathbf{x}_L}{2} \right) \cdot \mathbf{n}_{K,\sigma} \right| \leq \frac{1}{2} |\mathbf{x}_\sigma - \mathbf{x}_K| + \frac{1}{2} |\mathbf{x}_\sigma - \mathbf{x}_L| \leq h_{\mathcal{D}},$$

while due to mesh regularity

$$\left| \mathbf{n}_{K,\sigma} - \frac{(\mathbf{x}_L - \mathbf{x}_K)}{d_{K,L}} \right|$$

$$= \left| \mathbf{n}_{K,\sigma} - \frac{\mathbf{i}_{K,L}}{\mathbf{n}_{K,\sigma} \cdot \mathbf{i}_{K,L}} \right| \leq 1 + \left| \frac{\mathbf{i}_{K,L}}{\mathbf{n}_{K,\sigma} \cdot \mathbf{i}_{K,L}} \right| \leq 1 + \frac{1}{\theta},$$

from which it follows that

$$(T_3^{\mathcal{D}})^2 \leq C h_{\mathcal{D}} |\Omega| \|u\|_{\mathcal{D}}^2$$

with C only depending on d , Ω and φ . Thus one concludes that $\lim_{h_{\mathcal{D}} \rightarrow 0} T_3^{\mathcal{D}} = 0$. Successively, compare $T_2^{\mathcal{D}}$ with the term

$$T_4^{\mathcal{D}} = - \int_{\Omega} u(\mathbf{x}) \nabla \varphi(\mathbf{x}) \, d\mathbf{x}$$

$$= \sum_{K|L \in \mathcal{F}_{int}} (u_L - u_K) \int_{K|L} \varphi(\mathbf{x}) \mathbf{n}_{K,\sigma} \, d\gamma(\mathbf{x}),$$

which is such that

$$\lim_{h_{\mathcal{D}} \rightarrow 0} T_4^{\mathcal{D}} = - \int_{\Omega} \underline{u}(\mathbf{x}) \nabla \varphi(\mathbf{x}) \, d\mathbf{x} = \int_{\Omega} \varphi(\mathbf{x}) \nabla \underline{u}(\mathbf{x}) \, d\mathbf{x}.$$

Due to the fact that midpoint face interpolation is first order accurate

$$\left| \frac{1}{|\sigma|} \int_{K|L} \varphi(\mathbf{x}) \, d\gamma(\mathbf{x}) - \frac{\varphi(\mathbf{x}_K) + \varphi(\mathbf{x}_L)}{2} \right| \leq h_{\mathcal{D}} \|\nabla \varphi\|_{L^\infty(\Omega)},$$

one has that

$$\begin{aligned} & (T_4^{\mathcal{D}} - T_2^{\mathcal{D}})^2 \\ & \leq \sum_{K|L \in \mathcal{F}_{int}} (|\sigma| \mathbf{n}_{K,\sigma})^2 (u_L - u_K)^2 \\ & \quad \times \sum_{K|L \in \mathcal{F}_{int}} \left| \frac{1}{|\sigma|} \int_{K|L} \varphi(\mathbf{x}) \, d\gamma(\mathbf{x}) - \frac{\varphi(\mathbf{x}_K) + \varphi(\mathbf{x}_L)}{2} \right|^2 \\ & \leq \sum_{K|L \in \mathcal{F}_{int}} |\sigma|^2 (u_L - u_K)^2 \sum_{K|L \in \mathcal{F}_{int}} h_{\mathcal{D}}^2 \|\nabla \varphi\|_{L^\infty(\Omega)}^2, \end{aligned}$$

from which it follows that $\lim_{h_{\mathcal{D}} \rightarrow 0} (T_4^{\mathcal{D}} - T_2^{\mathcal{D}})^2 = 0$. Thus, $T_2^{\mathcal{D}}$ converges to $T_4^{\mathcal{D}}$ and, due to density of $C_c^\infty(\Omega)$ in $L^2(\Omega)$, $\nabla_{\mathcal{D}} u$ weakly converges to $\nabla \underline{u}$ as $h_{\mathcal{D}} \rightarrow 0$. Since

$$\lim_{h_{\mathcal{D}} \rightarrow 0} T_4^{\mathcal{D}} = \int_{\Omega} \underline{u}(\mathbf{x}) \nabla \varphi(\mathbf{x}) \, d\mathbf{x} = - \int_{\Omega} \varphi(\mathbf{x}) \nabla \underline{u}(\mathbf{x}) \, d\mathbf{x},$$

by density of $C_c^\infty(\Omega)$ in $L^2(\Omega)$, one obtains the weak convergence of $\nabla_{\mathcal{D}} u(\mathbf{x})$ to $\nabla \underline{u}(\mathbf{x})$ as $h_{\mathcal{D}} \rightarrow 0$.

Lemma 8.3 (Consistency of $\nabla_{\mathcal{D}}$). *Let Ω be a bounded open connected polyhedral subset of \mathbb{R}^d , $d \in \mathbb{N}^*$. Let \mathcal{D} be an admissible finite volume discretization in sense of Definition (2.1) and let $0 < \theta \leq \tilde{\theta}_{\mathcal{D}}$. Let $\underline{u} \in C^2(\bar{\Omega})$ be such that $\bar{u} = 0$ on $\partial\Omega$. Then there exists C , depending only on Ω , θ , \bar{u} and α such that*

$$\|\nabla_{\mathcal{D}} P_{\mathcal{D}} \bar{u} - \nabla \bar{u}\|_{L^2(\Omega)^d} \leq C_3 h_{\mathcal{D}} \quad (81)$$

Proof. From Eq.(79) for any $K \in \mathcal{M}$ one has

$$\begin{aligned} |K|(\nabla_{\mathcal{D}} P_{\mathcal{D}} u)_K &= \sum_{L \in \mathcal{N}_K} \tau_{K|L} d_{K,\sigma} \mathbf{n}_{K,\sigma} (\underline{u}(\mathbf{x}_L) - \underline{u}(\mathbf{x}_K)) \\ &\quad - \sum_{K \in \mathcal{M}} \sum_{\sigma \in \mathcal{F}_{K,ext}} \tau_{K,\sigma} d_{K,\sigma} \mathbf{n}_{K,\sigma} \underline{u}(\mathbf{x}_K) \end{aligned}$$

Let $(\nabla \underline{u})_K$ be the mean value of $\nabla \underline{u}$ over K

$$(\nabla \underline{u})_K = \frac{1}{|K|} \int_K \nabla \underline{u}(\mathbf{x}) \, d\mathbf{x}.$$

Due to the regularity of \underline{u} and the homogeneous Dirichlet boundary conditions, the flux consistency error estimates include a constant C , only depending on L^∞ norm of second derivatives of \underline{u} (and of α), such that for all $\sigma = K|L \in \mathcal{F}_{int}$, one has

$$|e_\sigma| \leq Ch_{\mathcal{D}} \quad \text{with} \quad e_\sigma = (\nabla \underline{u})_K \cdot \mathbf{n}_{K,\sigma} - \frac{\underline{u}(\mathbf{x}_L) - \underline{u}(\mathbf{x}_K)}{d_{K,L}}$$

while for all $\sigma \in \mathcal{F}_{ext}$ one has

$$|e_\sigma| \leq Ch_{\mathcal{D}} \quad \text{with} \quad e_\sigma = (\nabla \underline{u})_K \cdot \mathbf{n}_{K,\sigma} - \frac{-\underline{u}(\mathbf{x}_K)}{d_{K,\sigma}}$$

These flux consistency errors allow to recast $(\nabla_{\mathcal{D}} P_{\mathcal{D}} u)_K$ as

$$|K|(\nabla_{\mathcal{D}} P_{\mathcal{D}} u)_K = \sum_{\sigma \in \mathcal{F}_{K,ext}} |\sigma| d_{K,\sigma} \mathbf{n}_{K,\sigma} (\nabla \underline{u})_K \cdot \mathbf{n}_{K,\sigma} + R_K$$

where the consistency residual term is defined as

$$R_K = - \sum_{\sigma \in \mathcal{F}_K} |\sigma| d_{K,\sigma} \mathbf{n}_{K,\sigma} e_\sigma.$$

From the geometrical identity valid for any vector $\mathbf{v} \in \mathbb{R}^d$ and for all $K \in \mathcal{M}$

$$\frac{1}{|K|} \sum_{\sigma \in \mathcal{F}_K} |\sigma| (\mathbf{x}_\sigma - \mathbf{x}_0) \mathbf{n}_{K,\sigma} \cdot \mathbf{v} = \mathbf{v}, \quad (82)$$

which is a direct consequence of the fact that each cell is a closed volume, it follows that

$$|K|(\nabla_{\mathcal{D}} P_{\mathcal{D}} u)_K = |K|(\nabla \underline{u})_K + R_K$$

Due to flux consistency error estimates, it also follows that

$$|R_K| \leq Ch_{\mathcal{D}} \sum_{\sigma \in \mathcal{F}_K} |\sigma| d_{K,\sigma} = Ch_{\mathcal{D}} d |K|. \quad (83)$$

As a consequence, one obtains that

$$\begin{aligned} & \sum_{K \in \mathcal{M}} |K| |(\nabla_{\mathcal{D}} P_{\mathcal{D}} u)_K - (\nabla \underline{u})_K|^2 \\ & \leq \sum_{K \in \mathcal{M}} |K| C^2 h_{\mathcal{D}}^2 d^2 = |\Omega| C^2 h_{\mathcal{D}}^2 d^2. \end{aligned} \quad (84)$$

Due to the regularity $\underline{u} \in C^2(\Omega)$, there exists another C , only dependent on L^∞ norm of the second derivatives of \underline{u} , such that

$$\sum_{K \in \mathcal{M}} \int_K |\nabla \underline{u} - (\nabla \underline{u})_K|^2 \leq Ch_{\mathcal{D}}^2. \quad (85)$$

From Eqs.(84) and (85), one gets the existence of C_c , only dependent on Ω and \underline{u} , such that (81) holds.

References

- [1] I. Aavatsmark, T. Barkve, O. Böe, and T. Mannseth. Discretization on unstructured grids for inhomogeneous, anisotropic media. Part I: Derivation of the methods. *SIAM Journal of Scientific Computing*, 19(5):1700–1716, 1998.
- [2] I. Aavatsmark, T. Barkve, O. Böe, and T. Mannseth. Discretization on unstructured grids for inhomogeneous, anisotropic media. Part II: Discussion and numerical results. *SIAM Journal of Scientific Computing*, 19(5):1717–1736, 1998.
- [3] I. Aavatsmark, G.T. Eigestad, B.T. Mallison, and J.M. Nordbotten. A compact multipoint flux approximation method with improved robustness. *Numerical Methods for Partial Differential Equations*, 24(5):1329–1360, 2008.
- [4] B. Andreianov, M. Bendahmane, and F. Hubert. On 3D DDFV discretization of gradient and divergence operators: discrete functional analysis tools and applications to degenerate parabolic problems. *Computational Methods in Applied Mathematics*, 13(4):369–410, 2013.
- [5] B. Andreianov, M. Bendahmane, F. Hubert, and S. Krell. On 3D DDFV discretization of gradient and divergence operators. I. Meshing, operators and discrete duality. *IMA Journal of Numerical Analysis*, 32(4):1574–1603, 2012.
- [6] ANSYS®. Gambit, Release 2.4, ANSYS, Inc., Canonsburg, PA, 2007.
- [7] ANSYS®. Fluent, Release 16.1, ANSYS, Inc., Canonsburg, PA, 2015.
- [8] T. Barth and M. Ohlberger. Finite Volume Methods: Foundation and Analysis. In E. Stein, R. de Borst, and T. Hughes,

- editors, *Encyclopedia of Computational Mechanics*, pages 439–473. Wiley, New York, NY, 2004.
- [9] Z. Cai. On the finite volume element method. *Numerische Mathematik*, 58(1):713–735, 1990.
- [10] E. Chénier, R. Eymard, and R. Herbin. A collocated finite volume scheme to solve free convection for general non-conforming grids. *Journal of Computational Physics*, 228(6):2296–2311, 2009.
- [11] Y. Coudière and F. Hubert. A 3D discrete duality finite volume method for nonlinear elliptic equations. *SIAM Journal of Scientific Computing*, 33(4):1739–1764, 2011.
- [12] Y. Coudière and G. Manzini. The discrete duality finite volume method for convection-diffusion problems. *SIAM Journal of Numerical Analysis*, 47(6):4163–4192, 2010.
- [13] Y. Coudière, J.P. Vila, and P. Villedieu. Convergence rate of a finite volume scheme for a two dimensional convection-diffusion problem. *ESAIM: Mathematical Modelling and Numerical Analysis*, 33(3):493–516, 1999.
- [14] A. Della Rocca. *Large-Eddy Simulations of Turbulent Reacting Flows with Industrial Applications*. PhD thesis, Politecnico di Milano, Milano, Italy, February, 2018.
- [15] J. Droniou. Finite volume schemes for diffusion equations: introduction to and review of modern methods. *Mathematical Models and Methods in Applied Sciences*, 24(8):1575–1619, 2014.
- [16] J. Droniou and R. Eymard. A mixed finite volume scheme for anisotropic diffusion problems on any grid. *Numerische Mathematik*, 105(1):35–71, 2006.
- [17] J. Droniou and R. Eymard. The asymmetric gradient discretisation method. In C. Cancès and P. Omnes, editors, *Finite Volumes for Complex Applications VIII - Methods and Theoretical Aspects*, pages 311–319, 2017.
- [18] J. Droniou, R. Eymard, T. Gallouët, and R. Herbin. A unified approach to mimetic finite difference, hybrid finite volume and mixed finite volume methods. *Mathematical Models and Methods in Applied Sciences*, 20(2):265–295, 2010.

- [19] J. Droniou, R. Eymard, and R. Herbin. Gradient schemes: generic tools for the numerical analysis of diffusion equations. *ESAIM: Mathematical Modelling and Numerical Analysis*, 50(3):749–781, 2016.
- [20] G.T. Eigestad and R.A. Klausen. On the convergence of the multi-point flux approximation O-method: numerical experiments for discontinuous permeability. *Numerical Methods for Partial Differential Equations*, 21(6):1079–1098, 2005.
- [21] R. Eymard, T. Gallouët, and R. Herbin. Finite Volume Methods. In P.G. Ciarlet and J.L. Lions, editors, *Handbook of Numerical Analysis*, chapter VII, pages 713–1020. North-Holland, Amsterdam, Netherlands, 2000.
- [22] 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 Journal of Numerical Analysis*, 26(2):326–353, 2006.
- [23] R. Eymard, T. Gallouët, and R. Herbin. A new finite volume scheme for anisotropic diffusion problems on general grids: convergence analysis. *Comptes Rendus Mathématique*, 344(6):403–406, 2007.
- [24] R. Eymard, T. Gallouët, and R. Herbin. Cell centred discretisation of non linear elliptic problems on general multidimensional polyhedral grids. *Journal of Numerical Mathematics*, 17(3):173–193, 2009.
- [25] R. Eymard, T. Gallouët, and R. Herbin. Discretization of heterogeneous and anisotropic diffusion problems on general nonconforming meshes SUSHI: a scheme using stabilization and hybrid interfaces. *IMA Journal of Numerical Analysis*, 30(4):1009–1043, 2010.
- [26] R. Eymard, T. Gallouët, R. Herbin, and J.C. Latché. Analysis tools for finite volume schemes. *Acta Mathematica Universitatis Comenianae*, 76(1):111–136, 2007.
- [27] R. Eymard, R. Herbin, and J.C. Latché. Convergence analysis of a collocated finite volume scheme for the incompressible Navier-Stokes equations on general 2D or 3D meshes. *SIAM Journal of Numerical Analysis*, 45(1):1–36, 2007.

- [28] R. Eymard, R. Herbin, J.C. Latché, and B. Piar. Convergence analysis of a locally stabilized collocated finite volume scheme for incompressible flows. *ESAIM: Mathematical Modelling and Numerical Analysis*, 43(5):889–927, 2009.
- [29] C. Farre, C.D. Perez-Segarra, M. Soria, and A. Oliva. Analysis of different numerical schemes for the resolution of convection-diffusion equations using finite-volume methods on three-dimensional unstructured grids. Part II: Numerical analysis. *Numerical Heat Transfer, Part B*, 49(4):351–375, 2006.
- [30] J.H. Ferziger and M. Perić. *Computational Methods for Fluid Dynamics*. Springer-Verlag, Berlin Heidelberg, Germany, third edition, 2002.
- [31] Z.E. Heinemann, C.W. Brand, M. Munka, and Y.M. Chen. Modeling reservoir geometry with irregular grids. *SPE Reservoir Engineering*, 6(2):225–232, 1991.
- [32] F. Hermeline. Approximation of 2-D and 3-D diffusion operators with variable full tensor coefficients on arbitrary meshes. *Computer Methods in Applied Mechanics and Engineering*, 196(21-24):2497–2526, 2007.
- [33] F. Hermeline. A finite volume method for approximating 3D diffusion operators on general meshes. *Journal of Computational Physics*, 228(16):5763–5786, 2009.
- [34] S. Krell and G. Manzini. The discrete duality finite volume method for Stokes equations on three-dimensional polyhedral meshes. *SIAM Journal of Numerical Analysis*, 50(2):808–837, 2012.
- [35] Y.G. Lai. Unstructured grid arbitrarily shaped element method for fluid flow simulation. *AIAA Journal*, 38(12):2246–2252, 2000.
- [36] S.R. Mathur and J.Y. Murthy. A pressure-based method for unstructured meshes. *Numerical Heat Transfer, Part B*, 31(2):195–215, 1997.
- [37] F. Moukalled, L. Mangani, and M. Darwish. *The Finite Volume Method in Computational Fluid Dynamics. An Advanced Introduction with OpenFOAM[®] and MATLAB[®]*. Springer, Cham, Switzerland, 2016.

- [38] V. Moureau, P. Domingo, and L. Vervisch. Design of a massively parallel CFD code for complex geometries. *Comptes Rendus Mécanique*, 339(2-3):141–148, 2011.
- [39] S. Muzaferija and D. Gosman. Finite-volume CFD procedure and adaptive error control strategy for grids of arbitrary topology. *Journal of Computational Physics*, 138(2):766–787, 1997.
- [40] OpenFOAM, The OpenFOAM Foundation. <https://openfoam.org/>.
- [41] C.D. Perez-Segarra, C. Farre, J. Cadafalch, and A. Oliva. Analysis of different numerical schemes for the resolution of convection-diffusion equations using finite-volume methods on three-dimensional unstructured grids. Part I: Discretization schemes. *Numerical Heat Transfer, Part B*, 49(4):333–350, 2006.
- [42] Y.Y. Tsui and Y.F. Pan. A pressure-correction method for incompressible flows using unstructured meshes. *Numerical Heat Transfer, Part B*, 49(1):43–65, 2006.
- [43] P. Wesseling and C.W. Oosterle. Geometric multigrid with applications to computational fluid dynamics. *Journal of Computational and Applied Mathematics*, 128(1-2):311–334, 2001.

MOX Technical Reports, last issues

Dipartimento di Matematica
Politecnico di Milano, Via Bonardi 9 - 20133 Milano (Italy)

- 36/2018** Agosti, A.; Ambrosi, D.; Turzi, S.
Strain energy storage and dissipation rate in active cell mechanics
- 32/2018** Dal Santo, N.; Deparis, S.; Manzoni, A.; Quarteroni, A.
An algebraic least squares reduced basis method for the solution of nonaffinely parametrized Stokes equations
- 35/2018** Possenti, L.; Casagrande, G.; Di Gregorio, S.; Zunino, P.; Costantino, M.L.
Numerical simulations of the microvascular fluid balance with a non-linear model of the lymphatic system
- 34/2018** Laurino, F.; Coclite, A.; Tiozzo, A.; Decuzzi, P.; Zunino, P.;
A multiscale computational approach for the interaction of functionalized nanoparticles with the microvasculature
- 33/2018** Pagani, S.; Manzoni, A.; Quarteroni, A.
Numerical approximation of parametrized problems in cardiac electrophysiology by a local reduced basis method
- 31/2018** Quarteroni, A.
The role of statistics in the era of big data: A computational scientist' perspective
- 30/2018** Ieva, F.; Palma, F.; Romo, J.
Bootstrap-based Inference for Dependence in Multivariate Functional Data
- 29/2018** Manzoni, A.; Bonomi, D.; Quarteroni, A.
Reduced order modeling for cardiac electrophysiology and mechanics: new methodologies, challenges & perspectives
- 28/2018** Gerbi, A.; Dede', L.; Quarteroni, A.
Segregated algorithms for the numerical simulation of cardiac electromechanics in the left human ventricle
- 26/2018** Vergara, C.; Zonca, S.
Extended Finite Elements method for fluid-structure interaction with an immersed thick non-linear structure