



HAL
open science

The Discrete Duality Finite Volume Method for Convection Diffusion Problems

Yves Coudière, Gianmarco Manzini

► **To cite this version:**

Yves Coudière, Gianmarco Manzini. The Discrete Duality Finite Volume Method for Convection Diffusion Problems. *SIAM Journal on Numerical Analysis*, 2010, 47 (6), pp.4163-4192. hal-00319254

HAL Id: hal-00319254

<https://hal.science/hal-00319254v1>

Submitted on 7 Sep 2008

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

THE DISCRETE DUALITY FINITE VOLUME METHOD FOR CONVECTION-DIFFUSION PROBLEMS

YVES COUDIÈRE * AND GIANMARCO MANZINI †

Abstract. In this paper we extend the Discrete Duality Finite Volume (DDFV) formulation to the steady convection-diffusion equation. The discrete gradients defined in DDFV are used to define a cell-based gradient for the control volumes of both the primal and dual meshes, in order to achieve a higher-order accurate numerical flux for the convection term. A priori analysis is carried out to show convergence of the approximation and a global first-order convergence rate is derived. The theoretical results are confirmed by some numerical experiments.

Key words. Convection-diffusion equation, discrete duality finite volume method, polynomial reconstruction, diamond scheme, unstructured meshes, polygonal meshes.

AMS subject classifications. 35J05, 35J25, 65N12, 65N15, 65N30

1. Introduction. The finite volume (FV) approximation to the steady convection-diffusion equation in divergence form is based on the discretization of the convection and diffusion fluxes. The convection and diffusion fluxes are numerically evaluated at the control volume interfaces of a suitable mesh partition of the computational domain. A quite common approach consists in combining an upwind numerical flux for the convection term with the numerical diffusion flux thus resulting in the *two-point flux scheme* [26] or the *diamond scheme* [13]. Several other approaches pertinent to this problem are also available, as, for example, the *finite volume element methods*, and the *mimetic finite difference methods*. Since there is a huge amount of literature concerning these latter techniques, we just refer the interested reader to [9, 17] for a recent overview.

The formula for the two-point flux at a given cell interface only involves the degrees-of-freedom of the two cells adjacent to that interface, thus offering the advantage of a very compact computational stencil. Nonetheless, a consistent formulation relies on the notion of *admissible meshes*, which requires some rather restrictive orthogonality constraint [21]. Combined with a first order convection flux, this approach is applied to the numerical discretization of non-coercive convection-diffusion equations in [19]. Several variants are also shown effective in the computation of anisotropic diffusion problems [20], in modeling groundwater flows in partially saturated porous media [25] and have been recently extended to general meshes by using some cell-based reconstruction of the solution gradient [22–24].

On its turn, the diamond scheme provides a discretization of the diffusion flux that is consistent for more general geometries. In such a case, the numerical diffusion flux is discretized by using a formula for the discrete gradient, which requires the vertex values. These latter are expressed by a linear interpolation of the primary unknown, i.e. the cell averages. Implementations based on linear least squares are known to be quite accurate [8, 14, 15, 34], and used for the numerical discretization of more complex model problems [33]. Recent extensions using non-linear averages have also been shown to provide discrete maximum and minimum principles [7, 30–32].

A breakthrough in the diamond scheme methodology comes from [28, 29], where it is proposed to treat the vertex values as numerical unknowns. The resulting scheme combines two distinct finite volume schemes on two overlapping meshes, the mesh of the primal cells

*Laboratoire de Mathématiques “Jean Leray”, UMR 6629 CNRS – Université de Nantes, 2, rue de la Houssinière - BP 92208 44322 Nantes Cédex 3, France e-mail: Yves.Coudiere@math.univ-nantes.fr

† Istituto di Matematica Applicata e Tecnologie Informatiche (IMATI) – CNR, via Ferrata 1, I – 27100 Pavia, Italy e-mail: Marco.Manzini@imati.cnr.it

of the original diamond scheme formulation, and the mesh of the dual control volumes built around the vertices of the primal mesh. The method can also be reformulated in the framework of mimetic discretizations [11, 12] by introducing discrete divergence and gradient operators. These operators satisfy a discrete duality relationship that mimics the integration-by-parts formula (Green formula), thus motivating the wording *Discrete Duality Finite Volume* (DDFV) method. The DDFV method is successfully applied to the Laplace equation in [18], and shown to provide a very accurate approximation to the solution gradient on distorted meshes in [27]. A generalization to nonlinear elliptic equations is found in [3], and to the approximation of the div-curl operator in [16].

Concerning the numerical convection flux, it is worth mentioning that upwinding is crucial to ensure numerical stability, particularly in the convection-dominated regime [4]. The degrees-of-freedom approximate either the average of the exact solution over the mesh control volumes or the solution value at a given set of points of the computational domain. Thus, the finite volume method provides a piecewise constant approximation to the scalar variable, which is intrinsically first-order accurate if the error is measured in the L^2 -norm. Higher order of accuracy can be obtained through a cell-based linear polynomial reconstruction, which requires some cell-based discretization of the solution gradient [6, 35].

To do so in the DDFV framework, we propose to express these cell-based gradients by means of the available DDFV gradients. We carry out the a priori analysis of the resulting method in the diffusive regime, show the consistency of the numerical fluxes, prove the convergence of the DDFV approximation to the scalar variable and its gradient, and derive an $\mathcal{O}(h)$ estimate for the convergence rate in a properly defined mesh-dependent norm. We also propose an original approach to prove the consistency of the numerical diffusion and convection flux, which is alternative to the more common approach based on cell averages of Taylor expansions. More precisely, we show how the analogy of the DDFV method with the mimetic finite differences method makes it possible to apply the error estimates for polynomial interpolation in Sobolev spaces in a finite volume context [10]. Due to its generality, we think that this approach may be of interest as a theoretical tool in the analysis of many other finite volume schemes.

The outline of the paper is as follows. In Section 2 we detail the DDFV formulation. In Section 3 we present the a priori analysis of the method. In Section 4 we show some numerical results that confirm accuracy and convergence of the method. In Section 5 we offer some final remarks and discuss the perspectives of future work.

2. The discrete duality finite volume method. We are concerned with the DDFV approximation to the scalar solution p of the steady convection-diffusion equation

$$\operatorname{div}(\mathbf{b}p - \mathbf{K}\nabla p) = f \quad \text{in } \Omega \quad (2.1a)$$

$$p = g \quad \text{on } \Gamma \quad (2.1b)$$

defined on the bounded polygonal domain $\Omega \subset \mathbb{R}^2$ with boundary Γ . We assume that $f \in L^2(\Omega)$, $g \in H^{1/2}(\Gamma)$, $\mathbf{b} \in (L^\infty(\Omega))^2$ and $\operatorname{div} \mathbf{b} \in L^\infty(\Omega)$, and $\mathbf{K} \in (L^\infty(\Omega))^{2 \times 2}$. We also assume that \mathbf{K} is a symmetric matrix and that there exists a couple of positive constants, namely κ_* and κ^* , such that for every $\mathbf{u} \in \mathbb{R}^2$ there holds

$$\kappa_* \|\mathbf{u}\|^2 \leq \|\mathbf{K}^{1/2}\mathbf{u}\|^2 \leq \kappa^* \|\mathbf{u}\|^2 \quad (2.2)$$

almost everywhere in Ω . Under such assumptions, it is possible to show that problem (2.1) admits a unique weak solution $p \in H^1(\Omega)$.

As pointed out in the introduction, the DDFV method is based on two finite volume schemes for the scalar unknowns, which are formulated on two distinct and overlapping meshes, e.g. \mathcal{T}_h and \mathcal{V}_h . These two finite volume schemes are coupled through the dis-

crete solution gradients that are defined over a third mesh denoted by \mathcal{D}_h . The mesh size parameter labelling \mathcal{T}_h , \mathcal{V}_h , and \mathcal{D}_h is given by

$$h = \max_{D \in \mathcal{D}_h} h_D,$$

where h_D is the diameter of the cell D of the mesh \mathcal{D}_h . Thus, we will consider a family of mesh triples, i.e. $\{(\mathcal{T}_h, \mathcal{V}_h, \mathcal{D}_h)\}_h$, and the DDFV numerical approximation is formulated on some sequence of these items for decreasing h . The three meshes of each triple are not independent, but \mathcal{V}_h and \mathcal{D}_h are derived from \mathcal{T}_h through the construction process of subsection 2.1. For this reason, we sometimes refer to \mathcal{T}_h as *the primary mesh*, and to the formers as *the secondary mesh* and *the diamond mesh*. Moreover, we require that each triple $(\mathcal{T}_h, \mathcal{V}_h, \mathcal{D}_h)$ satisfies some regularity conditions, c.f. subsection 2.2, to perform the convergence analysis.

2.1. Notations and mesh constructions.

The Primary Mesh. Each primary mesh \mathcal{T}_h is a conformal partition of Ω formed by $m_{\mathcal{T}_h}$ non-overlapping polygons T , $m_{\mathcal{V}_h}$ vertices V , and $m_{\mathcal{D}_h}$ edges e . For each polygon $T \in \mathcal{T}_h$, we denote its two-dimensional Lebesgue measure by $|T|$, its boundary by ∂T , and the position of its center of gravity by \mathbf{x}_T . The symbol ∂T also indicates the set of edges of the polygon T . Each edge is either the intersection between two adjacent polygons T_1 and T_2 , i.e. $e = T_1 \cap T_2$, or the intersection between a polygon T_1 and the domain boundary Γ , i.e. $e = T_1 \cap \Gamma$. The former case refers to an *internal* edge, the latter to a *boundary* edge. For each edge e , we denote its one-dimensional Lebesgue measure by $|e|$, and its midpoint by \mathbf{x}_e . We also distinguish between internal and boundary vertices, the formers being located in the interior of the domain Ω and the latters at the domain boundary Γ . We will find it useful to denote the set of boundary vertices by $\mathcal{V}_h^{\text{Bnd}}$ and the set of internal vertices by $\mathcal{V}_h^{\text{Int}}$.

The Secondary Mesh. To each vertex of the primary mesh \mathcal{T}_h , denoted by its position vector \mathbf{x}_V , we associate the polygonal control volume V , c.f. Figure 2.1. The collection of all the control volumes V forms the secondary mesh \mathcal{V}_h . The vertices of each cell V are the baricenters \mathbf{x}_T of the primary cells T surrounding \mathbf{x}_V , the midpoints \mathbf{x}_e of the edges connected to \mathbf{x}_V , and, additionally, the vertex \mathbf{x}_V itself if $\mathbf{x}_V \in \Gamma$. For each cell $V \in \mathcal{V}_h$, we denote its two-dimensional Lebesgue measure by $|V|$ and its boundary by ∂V . We emphasize that the polygonal cells of \mathcal{V}_h are generally non-convex subsets of Ω , even if the primary cells are convex. Each edge of the secondary mesh \mathcal{V}_h , usually denoted by the symbol f , is either an interface between two adjacent cells V_1 and V_2 , i.e. $f = V_1 \cap V_2$, or a subset of Γ that belongs to the boundary ∂V of some cell V .

The Diamond Mesh. The third mesh considered in the formulation of the DDFV method, namely the mesh \mathcal{D}_h , is formed by quadrilateral cells, *the diamonds*, which are denoted by D and have two-dimensional Lebesgue measure $|D|$. Each diamond cell D is uniquely referred to an edge e of the primal mesh \mathcal{T}_h , see Figure 2.1. Indeed,

- if e is an internal edge, the vertices of D are \mathbf{x}_{V_1} and \mathbf{x}_{V_2} , the end points of e , and \mathbf{x}_{T_1} and \mathbf{x}_{T_2} , the baricenters of the cells T_1 and T_2 sharing e ;
- if e is a boundary edge, the vertices of D are \mathbf{x}_{V_1} and \mathbf{x}_{V_2} , the end points of e , \mathbf{x}_{T_1} , the baricenter of the cell T_1 to which e belongs, and \mathbf{x}_e , the midpoint of e .

The local numbering of these items is determined in accordance with

$$\det(\mathbf{x}_{T_2} - \mathbf{x}_{T_1}, \mathbf{x}_{V_2} - \mathbf{x}_{V_1}) > 0,$$

and taking $\mathbf{x}_{T_2} := \mathbf{x}_e$ if e is a boundary edge.

Furthermore, if e is an internal edge, the diamond cell D , which is associated to e , contains exactly:

- two interfaces between the control volumes V_1 and V_2 , denoted by f_1 and f_2 , which respectively connect \mathbf{x}_{T_1} and \mathbf{x}_{T_2} to \mathbf{x}_e , and are such that $f_1 \cup f_2 = V_1 \cap V_2$. In such a case, we use the symbol f to denote the straight segment connecting \mathbf{x}_{T_1} to \mathbf{x}_{T_2} ; note that f is generally distinct from $f_1 \cup f_2$;
- the interface $e = T_1 \cap T_2$ of vertices \mathbf{x}_{V_1} and \mathbf{x}_{V_2} .

Instead, if e is a boundary edge, D contains exactly:

- one interface between the control volumes V_1 and V_2 , denoted by f , which connects \mathbf{x}_{T_1} and \mathbf{x}_e ; clearly, $f = V_1 \cap V_2$;
- the boundary edge $e \subset \Gamma$ that connects the vertices \mathbf{x}_{V_1} and \mathbf{x}_{V_2} ; we consider this edge as splitted in two segments that respectively connect \mathbf{x}_{V_1} and \mathbf{x}_{V_2} to \mathbf{x}_e .

In both cases, each diamond cell D is also uniquely associated to the interface f , and, when e is an internal edge, to the couple of internal interfaces (f_1, f_2) . Throughout the paper, we exploit this correspondance by using the symbols D_e and D_f to refer to the diamond cells related to e and f . We will also use expressions like $e \in \partial T$ and $f \in \partial V$ to index summations over the set of edges forming the boundaries of T and V .

We denote the unit normal vector to e , f , f_1 , and f_2 by \mathbf{n}_e , \mathbf{n}_f , \mathbf{n}_{f_1} , and \mathbf{n}_{f_2} . If e is an internal edge, the vector \mathbf{n}_e points from \mathbf{x}_{T_1} to \mathbf{x}_{T_2} ; when e is a boundary edge, \mathbf{n}_e is just the unit normal vector to $e \cap \Gamma$ pointing out of Ω . The other normal vectors are always oriented from \mathbf{x}_{V_1} to \mathbf{x}_{V_2} . For consistency of notation, $|\sigma|$ is the one-dimensional Lebesgue measure of $\sigma \in \{f_1, f_2, f\}$. It is easy to show that when e is internal there holds

$$|f|\mathbf{n}_f = |f_1|\mathbf{n}_{f_1} + |f_2|\mathbf{n}_{f_2}. \quad (2.3)$$

2.2. Mesh regularity assumptions. To formulate the mesh regularity assumptions needed by the convergence analysis, we consider the sub-triangulation $\tilde{\mathcal{T}}_{\mathcal{D}_h} = \{\tilde{T}\}$ of the cells of \mathcal{D}_h obtained by splitting the diamonds $D \in \mathcal{D}_h$ along their internal diagonal e . We emphasize that we never use $\tilde{\mathcal{T}}_{\mathcal{D}_h}$ in the practical implementation of the method, but just to introduce the mesh regularity assumptions that follow. Then, we assume that

- (M1) the triangulation $\tilde{\mathcal{T}}_{\mathcal{D}_h}$ is *non-degenerate* in the sense of [10], i.e. there exists a constant $\eta > 0$ such that, for all mesh size parameter h there holds:

$$\frac{\rho_{\tilde{T}}}{h_{\tilde{T}}} \geq \eta \quad \text{for every } \tilde{T} \in \tilde{\mathcal{T}}_{\mathcal{D}_h},$$

where $\rho_{\tilde{T}}$ and $h_{\tilde{T}}$ are the diameters of the inscribed and circumscribed circle for the triangle \tilde{T} ;

- (M2) there exist two constants $\theta > 0$ and $\gamma > 0$ such that for each $D \in \mathcal{D}_h$ we have

$$\theta \leq \theta_D \leq \pi - \theta \quad \text{and} \quad |f_1| + |f_2| \leq \gamma|f|$$

where θ_D is the angle between the direction of e and f , and f_1, f_2 are related to f through (2.3).

- (M3) there exists $\rho_1 > 0$ such that for all h there holds:

$$\frac{|D'|}{|D''|} \leq \rho_1 \quad \text{for every } D', D'' \in \mathcal{S}(\mathbf{g}_T), \text{ for every } T \in \mathcal{T}_h,$$

$$\frac{|D'|}{|D''|} \leq \rho_1, \quad \text{for every } D', D'' \in \mathcal{S}(\mathbf{g}_V), \text{ for every } V \in \mathcal{V}_h^{\text{int}},$$

where $\mathcal{S}(\mathbf{g}_T)$ and $\mathcal{S}(\mathbf{g}_V)$ are the stencils of the gradients of the piecewise reconstructions detailed in subsection 2.5.

The above requirements have several consequences that can be easily verified. Among them, we outline the following ones, which will be used later.

- From Assumption (M1), there holds:

$$\frac{h_D^2}{|D|} \leq \rho_0 \quad \text{for every } D \in \mathcal{D}_h, \quad (2.4)$$

where $\rho_0 = 2/(\pi\eta^2)$ is the second regularity constant used in the convergence analysis. Let us, indeed, consider the diamond cell $D \in \mathcal{D}_h$ and the two triangles $T_1, T_2 \in \mathcal{T}_h$ such that $|D \cap T_1| \neq 0$ and $|D \cap T_2| \neq 0$. By definition,

$$\begin{aligned} h_D &= \max_{\mathbf{x}, \mathbf{y} \in D} |\mathbf{x} - \mathbf{y}| \leq \max_{\substack{\mathbf{x} \in D \cap T_1 \\ \mathbf{y} \in D \cap T_2}} |\mathbf{x} - \mathbf{y}| \\ &\leq \max_{\mathbf{x}, \mathbf{y} \in D \cap T_1} |\mathbf{x} - \mathbf{y}| + \max_{\mathbf{x}, \mathbf{y} \in D \cap T_2} |\mathbf{x} - \mathbf{y}| \leq h_{T_1} + h_{T_2}. \end{aligned} \quad (2.5)$$

Noting that $|D| = |D \cap T_1| + |D \cap T_2|$, applying Assumption (M1), and using (2.5) gives the inequality chain

$$|D| = |D \cap T_1| + |D \cap T_2| \geq \pi(\rho_{T_1}^2 + \rho_{T_2}^2) \geq \pi\eta^2(h_{T_1}^2 + h_{T_2}^2) \geq \frac{\pi\eta^2}{2}h_D^2,$$

from which we get (2.4).

- Assumption (M1) also implies that there exist two positive constants, namely C^{Agm} and C^{Intp} , independent of h and such that for any $D \in \mathcal{D}_h$,

(i) for any $q \in H^1(D)$ and $\sigma \in \{e, f_1, f_1\}$ there holds

$$\|q\|_{L^2(\sigma)}^2 \leq C^{Agm} \sum_{i=0}^1 h_D^{2i-1} \|q\|_{H^i(D)}^2; \quad (2.6)$$

(ii) for any $q \in H^{k+1}(D)$ there holds

$$\inf_{q^{(k)} \in P_k(D)} \|q - q^{(k)}\|_{H^i(D)} \leq C^{Intp} h_D^{k+1-i} \|q\|_{H^{k+1}(D)}. \quad (2.7)$$

In (2.6)-(2.7), $H^i(D)$ is the Sobolev space of functions over D that are square integrable and whose partial derivatives of order up to the non-negative integer i are all square integrable, $i = 0$ corresponds to $L^2(D)$, and $P_k(D)$ is the space of polynomials of degree up to k . Sometimes, we refer to (2.6) as *the Agmon inequality*, see for a proof [2], and to (2.7) as *the interpolation inequality*, see for a proof [10].

- Assumption (M3) is only required to prove the convergence of the higher order formulation of the DDFV method.

2.3. The mimetic framework of the DDFV method. In this subsection, we formulate the DDFV method in the setting of the mimetic finite difference methods. More precisely, we proceed in accordance with these four steps:

(i) we consider two finite dimensional spaces, namely $Q_h \times \delta Q_h$ and X_h , whose elements represent discrete scalar and vector fields. The discrete scalars are the degrees-of-freedom of the DDFV method, while the discrete vectors are used to formulate the numerical convection and diffusion fluxes;

(ii) we equip Q_h and X_h with two suitable inner products;

(iii) we define a discrete divergence operator, a discrete gradient operator, and a boundary trace operator acting on $Q_h \times \delta Q_h$ and X_h , that satisfy a *discrete duality* relation mimicking the continuous Green formula;

(iv) we define some projection operators that will be used in the a priori analysis of the numerical approximation.

The discrete scalar fields are associated to the mesh cells in \mathcal{T}_h and \mathcal{V}_h through the finite dimensional space Q_h :

$$\text{for every } q \in Q_h, \quad q = \{\{q_T\}_{T \in \mathcal{T}_h}, \{q_V\}_{V \in \mathcal{V}_h}\} \quad \text{with } q_T, q_V \in \mathbb{R},$$

and to the boundary edges and vertices through the space δQ_h :

$$\text{for every } \delta q \in \delta Q_h, \quad \delta q = \{\{q_e\}_{e \subset \Gamma}\} \quad \text{with } q_e \in \mathbb{R}.$$

The discrete vector fields that form the finite dimensional space X_h are piecewise constant vectors over the mesh cells in \mathcal{D}_h :

$$\text{for every } \mathbf{u} \in X_h, \quad \mathbf{u} = \{\{\mathbf{u}_D\}_{D \in \mathcal{D}_h}\} \quad \text{with } \mathbf{u}_D \in \mathbb{R}^2.$$

In the formulation of the DDFV scheme, we also consider the affine space:

$$Q_{h,g} = \{q \in Q_h \text{ such that } q_V = g(\mathbf{x}_V) \text{ for every } V \in \mathcal{V}_h^{\text{Bnd}}\},$$

and the vector space $Q_{h,0}$ for $g = 0$ on Γ . To ease notation, we denote the elements of $Q_h \times \delta Q_h$ by $\bar{q} = (q, \delta q)$. We also use the symbol \bar{p}_h to denote (p_h, g^J) , where $p_h \in Q_{h,g}$ is the DDFV approximation to p and g^J is a suitable representation of the boundary data g in δQ_h that is defined later. The dimension of $Q_{h,g}$ equals the number of mesh cells in \mathcal{T}_h and $\mathcal{V}_h^{\text{int}}$, i.e. $m_{\mathcal{T}_h} + m_{\mathcal{V}_h^{\text{int}}}$, the dimension of δQ_h equals the number of boundary edges, i.e. $m_{\{e \subset \Gamma\}}$, and the dimension of X_h is twice the number of elements $m_{\mathcal{D}_h}$ of the mesh \mathcal{D}_h .

Then, we equip Q_h and X_h with the *inner products*

$$(p, q)_{Q_h} = \frac{1}{2} \left(\sum_{T \in \mathcal{T}_h} |T| p_T q_T + \sum_{V \in \mathcal{V}_h} |V| p_V q_V \right), \quad (2.8a)$$

$$(\mathbf{u}, \mathbf{v})_{X_h} = \sum_{D \in \mathcal{D}_h} |D| \mathbf{u}_D \cdot \mathbf{v}_D, \quad (2.8b)$$

and related norms $\|\cdot\|_{Q_h}$ and $\|\cdot\|_{X_h}$, and define the discrete bilinear form

$$(\mathbf{u}, \bar{q}) \in X_h \times (Q_h \times \delta Q_h) \rightarrow \langle \mathbf{u}, \mathbf{n}\bar{q} \rangle_{h,\Gamma} = \sum_{e \subset \Gamma} |e| \mathbf{n}_e \cdot \mathbf{u}_{D_e} \frac{1}{2} \left(q_e + \frac{q_{V_1} + q_{V_2}}{2} \right), \quad (2.9)$$

where the summation index e refers to the boundary edge with end points \mathbf{x}_{V_1} and \mathbf{x}_{V_2} and midpoint \mathbf{x}_e .

The *discrete gradient operator* $\nabla_h : Q_h \times \delta Q_h \rightarrow X_h$ is defined as

$$\nabla_h \bar{q} = \{\{\nabla_D \bar{q}\}_{D \in \mathcal{D}_h}\} \in X_h,$$

where

$$\nabla_D \bar{q} = \frac{1}{2|D|} ((q_{V_2} - q_{V_1})|f| \mathbf{n}_f + (q_{T_2} - q_{T_1})|e| \mathbf{n}_e). \quad (2.10)$$

In equation (2.10), we take $q_{T_2} := q_e$ when $e \subset \Gamma$, i.e. when the diamond cell D is adjacent to the boundary.

The *discrete divergence operator* $\text{div}_h : X_h \rightarrow Q_h$ is defined as

$$\text{div}_h \mathbf{u} = \{\{\text{div}_T \mathbf{u}\}_{T \in \mathcal{T}_h}, \{\text{div}_V \mathbf{u}\}_{V \in \mathcal{V}_h}\} \in Q_h \quad (2.11)$$

for every $\mathbf{u} \in X_h$, where

$$\operatorname{div}_T \mathbf{u} = \frac{1}{|T|} \sum_{e \in \partial T} |e| \mathbf{n}_e \cdot \mathbf{u}_{D_e}, \quad \text{for } T \in \mathcal{T}_h, \quad (2.12a)$$

$$\operatorname{div}_V \mathbf{u} = \frac{1}{|V|} \sum_{f \in \partial V} |f| \mathbf{n}_f \cdot \mathbf{u}_{D_f}, \quad \text{for } V \in \mathcal{V}_h, \quad (2.12b)$$

and D_e and D_f are the diamond cells related to e and f , respectively.

Lemma 2.1 below states the *discrete duality* relationship that holds between div_h and ∇_h with respect to the inner products (2.8a)-(2.8b) and the bilinear form (2.9). This relation mimics the Gauss-Green formula of the differential calculus through the discrete integration-by-parts formula (2.13).

LEMMA 2.1. *For all $\bar{q} \in Q_h \times \delta Q_h$ and $\mathbf{u} \in X_h$, there holds that*

$$(\operatorname{div}_h \mathbf{u}, q)_{Q_h} + (\mathbf{u}, \nabla_h \bar{q})_{X_h} = \langle \mathbf{u}, \mathbf{n} \bar{q} \rangle_{h, \Gamma}. \quad (2.13)$$

Proof. Note that

$$(\operatorname{div}_h \mathbf{u}, q)_{Q_h} = \frac{1}{2} \sum_{T \in \mathcal{T}_h} q_T \sum_{e \in \partial T} |e| \mathbf{n}_e \cdot \mathbf{u}_{D_e} + \frac{1}{2} \sum_{V \in \mathcal{V}_h} q_V \sum_{f \in \partial V} |f| \mathbf{n}_f \cdot \mathbf{u}_{D_f}.$$

Rearranging the summation terms in the previous formula and summing and subtracting the term with q_e yields:

$$\begin{aligned} \sum_{T \in \mathcal{T}_h} q_T \sum_{e \in \partial T} |e| \mathbf{n}_e \cdot \mathbf{u}_{D_e} &= \sum_{e \subset T_1 \cap T_2} |e| \mathbf{n}_e \cdot \mathbf{u}_{D_e} (q_{T_1} - q_{T_2}) \\ &\quad + \sum_{e \subset \Gamma} |e| \mathbf{n}_e \cdot \mathbf{u}_{D_e} (q_{T_1} - q_e) + \sum_{e \subset \Gamma} |e| \mathbf{n}_e \cdot \mathbf{u}_{D_e} q_e \\ &= \mathbb{T}_1 + \mathbb{T}_2 + \mathbb{T}_3. \end{aligned}$$

Likewise,

$$\begin{aligned} \sum_{V \in \mathcal{V}_h} q_V \sum_{f \in \partial V} |f| \mathbf{n}_f \cdot \mathbf{u}_{D_f} &= \sum_{f \subset V_1 \cap V_2} |f| \mathbf{n}_f \cdot \mathbf{u}_{D_f} (q_{V_1} - q_{V_2}) \\ &\quad + \sum_{e \subset \Gamma} |e| \mathbf{n}_e \cdot \mathbf{u}_{D_e} \frac{q_{V_1} + q_{V_2}}{2} = \mathbb{V}_1 + \mathbb{V}_2 \end{aligned}$$

The proof of the discrete Gauss-Green formula in (2.13) terminates by noting that:

$$(\mathbf{u}, \nabla_h \bar{q})_{X_h} = \mathbb{T}_1 + \mathbb{T}_2 + \mathbb{V}_1, \quad \text{and} \quad \langle \mathbf{u}, \mathbf{n} \bar{q} \rangle_{h, \Gamma} = \mathbb{T}_3 + \mathbb{V}_2.$$

□

For $q \in L^1(\Omega)$, the cell-average operator onto Q_h is given by:

$$q^{\mathbb{I}} = \left\{ \{q_T^{\mathbb{I}}\}_{T \in \mathcal{T}_h}, \{q_V^{\mathbb{I}}\}_{V \in \mathcal{V}_h} \right\} \in Q_h \quad (2.14)$$

where

$$q_T^{\mathbb{I}} = \frac{1}{|T|} \int_T q dV \quad \text{for all } T \in \mathcal{T}_h \quad \text{and} \quad q_V^{\mathbb{I}} = \frac{1}{|V|} \int_V q dV \quad \text{for all } V \in \mathcal{V}_h. \quad (2.15)$$

The pointwise projection from $C^0(\bar{\Omega})$ onto $Q_h \times \delta Q_h$ is given by:

$$q^{\mathbb{J}}|_{Q_h} = \left\{ \{q_T^{\mathbb{J}}\}_{T \in \mathcal{T}_h}, \{q_V^{\mathbb{J}}\}_{V \in \mathcal{V}_h} \right\} = \left\{ \{q(\mathbf{x}_T)\}_{T \in \mathcal{T}_h}, \{q(\mathbf{x}_V)\}_{V \in \mathcal{V}_h} \right\}, \quad (2.16a)$$

and

$$q^J|_{\delta Q_h} = \left\{ \left\{ q_e^J \right\}_{e \in \Gamma} \right\} = \left\{ \left\{ q(\mathbf{x}_e) \right\}_{e \in \Gamma} \right\}. \quad (2.16b)$$

For $\mathbf{u} \in (H^1(\Omega))^2$, the projection operator onto X_h is given by

$$\mathbf{u}^I = \left\{ \mathbf{u}_D^I \right\}_{D \in \mathcal{D}_h} \in X_h \quad (2.17)$$

where

$$\begin{aligned} \mathbf{n}_e \cdot \mathbf{u}_D^I &= \frac{1}{|e|} \int_e \mathbf{n}_e \cdot \mathbf{u} dS, \\ \mathbf{n}_f \cdot \mathbf{u}_D^I &= \frac{1}{|f|} \left(\int_{f_1} \mathbf{n}_{f_1} \cdot \mathbf{u} dS + \int_{f_2} \mathbf{n}_{f_2} \cdot \mathbf{u} dS \right). \end{aligned}$$

Definitions (2.11)-(2.12) are consistent with the Gauss divergence theorem since from (2.15)-(2.17) we easily obtain the result stated in Lemma 2.2 below.

LEMMA 2.2. *For every $\mathbf{u} \in (H^1(\Omega))^2$ there holds that:*

$$\begin{aligned} (\operatorname{div} \mathbf{u})_T^I &= \operatorname{div}_T \mathbf{u}^I \quad \text{for every } T \in \mathcal{T}_h, \\ (\operatorname{div} \mathbf{u})_V^I &= \operatorname{div}_V \mathbf{u}^I \quad \text{for every } V \in \mathcal{V}_h^{\text{int}}. \end{aligned}$$

Proof. For every $T \in \mathcal{T}_h$:

$$(\operatorname{div} \mathbf{u})_T^I = \frac{1}{|T|} \int_T \operatorname{div} \mathbf{u} dV = \frac{1}{|T|} \int_{\partial T} \mathbf{n}_T \cdot \mathbf{u} dS = \frac{1}{|T|} \sum_{e \in \partial T} |e| (\mathbf{u})_e^I = \operatorname{div}_T (\mathbf{u}^I).$$

Similarly, we find that $(\operatorname{div} \mathbf{u})_V^I = \operatorname{div}_V (\mathbf{u}^I)$ for every $V \in \mathcal{V}_h^{\text{int}}$. \square

Lemma 2.3 below focuses on the consistency of the discrete gradient with respect to the J-interpolation.

LEMMA 2.3. *For every $D \in \mathcal{D}_h$ and $q^{(1)} \in P_1(D)$ there holds:*

$$\nabla_D (q^{(1)})^J = (\nabla q^{(1)})_D^I.$$

Proof. As $q^{(1)}$ is a linear function over D , its gradient is a constant vector, thus implying that $(\nabla q)_\sigma^I = \mathbf{n}_\sigma \cdot \nabla q$ for $\sigma \in \{e, f_1, f_2\}$. The lemma is an obvious consequence of this fact. \square

2.4. The numerical diffusion flux. We assume that there exists a positive constant C_K independent of h such that for every $D \in \mathcal{D}_h$ we can find a constant matrix \mathbf{K}_D verifying

$$\max_{i,j=1,2} \operatorname{ess\,sup}_{\mathbf{x} \in D} |(\mathbf{K}_D)_{ij} - K_{ij}(\mathbf{x})| \leq C_K h_D, \quad (2.18)$$

where the constant C_K may depend on the regularity of the components of \mathbf{K} . For instance, if $K_{ij} \in W^{1,p}(\Omega)$ with $p > 2$ and D is convex or $K_{ij} \in W^{1,\infty}(\Omega)$ we can choose

$$(\mathbf{K}_D)_{ij} = \frac{1}{|D|} \int_D K(\mathbf{x}) dV.$$

Let $\mathbf{K}_h = \operatorname{diag} \{ \mathbf{K}_D \}_{D \in \mathcal{D}_h}$ be the block-diagonal matrix collecting the matrices \mathbf{K}_D . The numerical diffusion flux $\Phi_h^d(\bar{p}_h) \in X_h$ is given by:

$$\Phi_h^d(\bar{p}_h) = \mathbf{K}_h \nabla_h \bar{p}_h = \left\{ \mathbf{K}_D \nabla_D \bar{p}_h \right\}_{D \in \mathcal{D}_h}.$$

2.5. The numerical convection flux. In this section, we present the DDFV discretization of the convection term of equation (2.1a). For each internal diamond D with diagonals e and f , we consider the edge fluxes $\Phi_\sigma^{c,k}(\bar{p}_h)$ for $\sigma \in \{e, f_1, f_2\}$, $k = 0, 1$, and the total flux $\Phi_f^{c,k}(\bar{p}_h)$ given by

$$|f|\Phi_f^{c,k}(\bar{p}_h) = |f_1|\Phi_{f_1}^{c,k}(\bar{p}_h) + |f_2|\Phi_{f_2}^{c,k}(\bar{p}_h). \quad (2.19)$$

The numerical convection flux in X_h is given by

$$\Phi_h^{c,k}(\bar{p}_h) = \{ \{ \Phi_D^{c,k}(\bar{p}_h) \}_{D \in \mathcal{D}_h} \} \in X_h,$$

where, by definition,

$$\mathbf{n}_e \cdot \Phi_D^{c,k}(\bar{p}_h) = \Phi_e^{c,k}(\bar{p}_h), \quad \text{and} \quad \mathbf{n}_f \cdot \Phi_D^{c,k}(\bar{p}_h) = \Phi_f^{c,k}(\bar{p}_h).$$

To define the edge fluxes, we introduce the average flux integrals of the velocity field \mathbf{b}

$$b_\sigma = \frac{1}{|\sigma|} \int_\sigma \mathbf{n}_\sigma \cdot \mathbf{b} dS \quad \text{for } \sigma \in \{e, f_1, f_2\},$$

and the upwind (+) and downwind(-) velocities $b_\sigma^\pm = (b_\sigma \pm |b_\sigma|)/2$. The cell-wise reconstructions $R_T^k \bar{p}_h$ and $R_V^k \bar{p}_h$ of the unknown \bar{p}_h are piecewise constant ($k = 0$) or piecewise linear ($k = 1$) polynomials on, respectively, \mathcal{T}_h and \mathcal{V}_h . Now, let $\phi \in [1/2, 1]$ be a user-input parameter. The numerical convection flux is a convex combination of the upwind and downwind contributions evaluated using $R_T^k \bar{p}_h$ and $R_V^k \bar{p}_h$:

$$\begin{aligned} \Phi_e^{c,k}(\bar{p}_h) &= b_e^+ (\phi R_{T_1}^k \bar{p}_h(\mathbf{x}_e) + (1 - \phi) R_{T_2}^k \bar{p}_h(\mathbf{x}_e)) \\ &\quad + b_e^- (\phi R_{T_2}^k \bar{p}_h(\mathbf{x}_e) + (1 - \phi) R_{T_1}^k \bar{p}_h(\mathbf{x}_e)) \end{aligned} \quad (2.20a)$$

and this formula is modified by substituting the reconstructed value $R_{T_2}^k \bar{p}_h(\mathbf{x}_e)$ with $(\bar{p}_h)_e = g_e^j$ when e is on the boundary. The fluxes across f_i for $i = 1, 2$ are given by:

$$\begin{aligned} \Phi_{f_i}^{c,k}(\bar{p}_h) &= b_{f_i}^+ (\phi R_{V_1}^k \bar{p}_h(\mathbf{x}_{f_i}) + (1 - \phi) R_{V_2}^k \bar{p}_h(\mathbf{x}_{f_i})) \\ &\quad + b_{f_i}^- (\phi R_{V_2}^k \bar{p}_h(\mathbf{x}_{f_i}) + (1 - \phi) R_{V_1}^k \bar{p}_h(\mathbf{x}_{f_i})). \end{aligned} \quad (2.20b)$$

The choices $\phi = 1/2$ and $\phi = 1$ respectively correspond to the *central scheme* and to the *fully upwind scheme*, while the constraint $1/2 \leq \phi \leq 1$ is required to prove that the method is non-negative, c.f. Lemma 3.11 of Section 3.

Let $\bar{q} \in Q_h \times \delta Q_h$. The polynomial reconstructions $R_T^k \bar{q}$ and $R_V^k \bar{q}$ are given by:

$k = 0$ (piecewise constant)

$$R_T^0 \bar{q}(\mathbf{x}) = q_T \quad \text{for } \mathbf{x} \in T \text{ and every } T \in \mathcal{T}_h; \quad (2.21a)$$

$$R_V^0 \bar{q}(\mathbf{x}) = q_V \quad \text{for } \mathbf{x} \in V \text{ and every } V \in \mathcal{V}_h. \quad (2.21b)$$

$k = 1$ (piecewise linear)

$$R_T^1 \bar{q}(\mathbf{x}) = q_T + (\mathbf{x} - \mathbf{x}_T) \cdot \mathbf{g}_T(\nabla_h \bar{q}) \quad \text{for } \mathbf{x} \in T \text{ and every } T \in \mathcal{T}_h; \quad (2.22a)$$

$$R_V^1 \bar{q}(\mathbf{x}) = q_V + (\mathbf{x} - \mathbf{x}_V) \cdot \mathbf{g}_V(\nabla_h \bar{q}) \quad \text{for } \mathbf{x} \in V \text{ and every } V \in \mathcal{V}_h. \quad (2.22b)$$

Given a discrete vector $\mathbf{u} = \{ \mathbf{u}_D \}_{D \in \mathcal{D}_h} \in X_h$, the vectors $\mathbf{g}_T(\mathbf{u})$ and $\mathbf{g}_V(\mathbf{u})$ used in the piecewise linear reconstructions are convex combinations of the vector terms \mathbf{u}_D

$$\mathbf{g}_T(\mathbf{u}) = \sum_{D \in \mathcal{D}_h} \alpha_{T,D} \mathbf{u}_D \quad \text{and} \quad \mathbf{g}_V(\mathbf{u}) = \sum_{D \in \mathcal{D}_h} \alpha_{V,D} \mathbf{u}_D,$$

where the scalar coefficients $\boldsymbol{\alpha} = \{ \{ \alpha_{T,D} \}_{T \in \mathcal{T}_h, D \in \mathcal{D}_h}, \{ \alpha_{V,D} \}_{V \in \mathcal{V}_h, D \in \mathcal{D}_h} \}$ satisfy:

(G1) *non-negativity and normalization of the coefficients:*

$$\alpha_{T,D} \geq 0 \quad \text{for every } T \in \mathcal{T}_h, D \in \mathcal{D}_h \quad \text{and} \quad \sum_{D \in \mathcal{D}_h} \alpha_{T,D} = 1 \quad \text{for every } T \in \mathcal{T}_h;$$

$$\alpha_{V,D} \geq 0 \quad \text{for every } V \in \mathcal{V}_h, D \in \mathcal{D}_h \quad \text{and} \quad \sum_{D \in \mathcal{D}_h} \alpha_{V,D} = 1 \quad \text{for every } V \in \mathcal{V}_h;$$

(G2) *local supports:* there exists a positive integer number N_S independent of h such that

$$N_T(D) := \#\{T \in \mathcal{T}_h \text{ such that } D \in \mathcal{S}(\mathbf{g}_T)\} \leq N_S,$$

$$N_V(D) := \#\{V \in \mathcal{V}_h \text{ such that } D \in \mathcal{S}(\mathbf{g}_V)\} \leq N_S$$

for every $D \in \mathcal{D}_h$, where $\mathcal{S}(\mathbf{g}_T) = \cup_{\alpha_{T,D} > 0} D$ and $\mathcal{S}(\mathbf{g}_V) = \cup_{\alpha_{V,D} > 0} D$ are the *supports* of $\mathbf{g}_T(\cdot)$ and $\mathbf{g}_V(\cdot)$, and $\#\{\cdot\}$ is the cardinality of the set $\{\cdot\}$.

Assumptions (G1)-(G2) do not uniquely determine the coefficients α . In the numerical experiments of Section 4 we consider the following choice:

$$\alpha_{T,D} = |D \cap T|/|T| \quad \text{for every } T \in \mathcal{T}_h, D \in \mathcal{D}_h;$$

$$\alpha_{V,D} = |D \cap V|/|V| \quad \text{for every } V \in \mathcal{V}_h, D \in \mathcal{D}_h.$$

It is easy to show that both (G1) and (G2) are satisfied, and, in particular, that $N_T(D), N_V(D) \leq 2$ for every $D \in \mathcal{D}_h$.

LEMMA 2.4 (*k*-exactness). *Let us consider the function $q^{(k)} \in P_k(\Omega)$, and its pointwise interpolation $(q^{(k)})^J \in Q_h \times \delta Q_h$ defined in accordance with (2.16a)-(2.16b). Then,*

$$R_T^k(q^{(k)})^J(\mathbf{x}) = q^{(k)}(\mathbf{x}) \quad \text{for } \mathbf{x} \in T \text{ and every } T \in \mathcal{T}_h; \quad (2.23a)$$

$$R_V^k(q^{(k)})^J(\mathbf{x}) = q^{(k)}(\mathbf{x}) \quad \text{for } \mathbf{x} \in V \text{ and every } V \in \mathcal{V}_h. \quad (2.23b)$$

Proof. The case $k = 0$ is trivial. In the case $k = 1$, the gradient of $q^{(1)}$ is a constant vector on Ω and for every $D \in \mathcal{D}_h$ it must be equal to the discrete gradient of formula (2.10), i.e. $\nabla_D(q^{(1)})^J = \nabla q^{(1)}$. It is easy to see that for every $T \in \mathcal{T}_h$ and $V \in \mathcal{V}_h$ there holds $\mathbf{g}_T(\nabla_h(q^{(1)})^J) = \mathbf{g}_V(\nabla_h(q^{(1)})^J) = \nabla q^{(1)}$, due to the normalization of the coefficients α according to (G1). As $(q^{(1)})_T^J = q^{(1)}(\mathbf{x}_T)$, we have

$$R_T^1(q^{(1)})^J(\mathbf{x}) = q^{(1)}(\mathbf{x}_T) + (\mathbf{x} - \mathbf{x}_T) \cdot \nabla q^{(1)} = q^{(1)}(\mathbf{x})$$

for every $\mathbf{x} \in T$, and every $T \in \mathcal{T}_h$, thus proving (2.23a), and a similar calculation applies to $R_V^1(q^{(1)})^J$ to prove (2.23b). \square

2.6. The DDFV scheme. Let $f^I \in Q_h$ be the interpolation of f given by (2.14)-(2.15), and $g^J \in \delta Q_h$ the interpolation of g given by (2.16a)-(2.16b). Then, the formulation of the DDFV method reads as follows.

PROBLEM 2.1 (DDFV method).

Find $p_h \in Q_{h,g}$ such that

$$\operatorname{div}_h(\Phi_h^{c,k}(\bar{p}_h) - \Phi_h^d(\bar{p}_h)) = f^I \quad \text{in } Q_{h,0} \quad (2.24)$$

with $\bar{p}_h = (p_h, g^J) \in Q_h \times \delta Q_h$.

An alternative formulation, which is based on testing (2.24) on the elements of $Q_{h,0}$, reads as follows.

PROBLEM 2.2 (Discrete weak formulation).

Find $p_h \in Q_{h,g}$ such that:

$$\left(\Phi_h^d(\bar{p}_h) - \Phi_h^{c,k}(\bar{p}_h), \nabla_h(q, 0) \right)_{X_h} = (f^I, q)_{Q_h} \quad \text{for every } q \in Q_{h,0} \quad (2.25)$$

with $\bar{p}_h = (p_h, g^J) \in Q_h \times \delta Q_h$.

The equivalence of these formulations is a consequence of the discrete Green relation (2.13).

3. Convergence of the DDFV method. One of the major ingredients in the a priori analysis of finite volume methods consists in controlling the local consistency errors introduced by the numerical fluxes. To this purpose, the usual approach requires Taylor developments properly averaged on the mesh control volumes. In this work, we propose an alternative methodology based on the fact that the consistency errors are linear functionals vanishing for polynomial functions of degree k ($k = 0$ or $k = 1$ in our case). One might use the Bramble-Hilbert lemma and some suitably adapted scaling argument, like in the a priori analysis of finite element methods [10]. Nonetheless, a non-trivial issue is encountered along this line: the Bramble-Hilbert lemma holds true, but the scaling argument cannot be easily generalized as the supports of the local consistency errors are not affine-equivalent to a simple reference element. We circumvent this difficulty in Lemma 3.1 below by requiring that the euclidean norm of the error functional is uniformly bounded through a mesh-independent positive constant C , c.f. lemma's condition (ii). This constraint will be explicitly proved for both functionals expressing the local consistency error of diffusion and convection flux by applying lemmas 3.2-3.8. The main results of this section concerning the convergence of the DDFV approximation are eventually given in Theorems 3.15-3.17 for scalar solutions in $H^2(\Omega)$ and their gradients. It is worth noting that the continuous Sobolev embedding $H^2(\Omega) \hookrightarrow C^0(\bar{\Omega})$ [1] allows us to use the J interpolation, and also to derive point-wise estimates such as in Lemma 3.2 of subsection 3.1. An optimal convergence rate for the solution gradient and non-optimal convergence rate for the scalar unknown are also determined.

LEMMA 3.1. Let $\mathcal{S} \subset \mathcal{D}_h$, and take a diamond-shaped cell $D \subset \mathcal{S}$. Let also $\mathcal{J} : H^{k+1}(\mathcal{S}) \rightarrow \mathbb{R}^2$ be a linear vector-valued functional satisfying:

- (i) $\mathcal{J}(q^{(k)}) = 0$ for every $q^{(k)} \in P_k(\mathcal{S})$;
- (ii) there exists a constant C independent of h such that for every $q \in H^{k+1}(\mathcal{S})$ there holds

$$|D| \|\mathcal{J}(q)\|^2 \leq C \sum_{i=0}^k h_S^{2(i-\alpha)} \|q\|_{H^i(\mathcal{S})}^2,$$

where $\alpha = 0, 1$ and $h_S = \max_{D \subset \mathcal{S}} h_D$.

Then, for every $q \in H^{k+1}(\mathcal{S})$ we have:

$$|D| \|\mathcal{J}(q)\|^2 \leq C C^{Intp} h_S^{2(k+1-\alpha)} \|q\|_{H^{k+1}(\mathcal{S})}^2. \quad (3.1)$$

Proof. Let $q^{(k)} \in P_k(\mathcal{S})$. From Assumption (i)-(ii) we immediately have that

$$|D| \|\mathcal{J}(q)\|^2 = |D| \|\mathcal{J}(q - q^{(k)})\|^2 \leq C \sum_{i=0}^k h_S^{2(i-\alpha)} \|q - q^{(k)}\|_{H^i(\mathcal{S})}^2.$$

In view of Assumption (M1), we can apply interpolation inequality (2.7) to obtain:

$$\|q - q^{(k)}\|_{H^i(\mathcal{S})}^2 = \sum_{D \in \mathcal{S}} \|q - q^{(k)}\|_{H^i(D)}^2 \leq C^{Intp} h_S^{2(k+1-i)} \|q\|_{H^{k+1}(\mathcal{S})}^2,$$

from which inequality (3.1) follows. \square

3.1. Six easy pieces. In this subsection, we present six technical lemmas that prove some useful properties of scalar and vector functions defined over the diamond cells $D \in \mathcal{D}_h$. Remind that e and f indicate the diagonals of the current diamond cell D , and f_1 and f_2 are related to f by (2.3).

LEMMA 3.2. For any cell $D \in \mathcal{D}_h$, $q \in H^2(D)$ and $\mathbf{x} \in D$ there holds

$$|D||q(\mathbf{x})|^2 \leq 3 \sum_{i=0}^2 h_D^{2i} \|q\|_{H^i(D)}^2. \quad (3.2)$$

Proof. For any diamond cell D , let us consider $q \in C^\infty(\overline{D})$ and $\mathbf{x} \in D$. There exists a ball $\mathcal{B}_{(r,\mathbf{x})} = \{\mathbf{y} \in D, \text{ such that } \|\mathbf{y} - \mathbf{x}\| < r\}$ centered at \mathbf{x} of radius r that is entirely contained within D . The Taylor's development of q around \mathbf{x}

$$\begin{aligned} q(\mathbf{x}) &= q(\mathbf{y}) + \nabla q(\mathbf{y}) \cdot (\mathbf{x} - \mathbf{y}) \\ &\quad + \int_0^1 \nabla^2 q(\mathbf{y} + s(\mathbf{x} - \mathbf{y})) \cdot (\mathbf{x} - \mathbf{y})(\mathbf{x} - \mathbf{y})(1 - s) ds. \end{aligned}$$

holds for any $\mathbf{y} \in \mathcal{B}_{(r,\mathbf{x})}$. As a consequence, with $\|\mathbf{x} - \mathbf{y}\| \leq r \leq h_D$

$$|q(\mathbf{x})|^2 \leq 3 \left(|q(\mathbf{y})|^2 + h_D^2 |\nabla q(\mathbf{y})|^2 + h_D^4 \int_0^1 |\nabla^2 q(\mathbf{y} + s(\mathbf{x} - \mathbf{y}))|^2 (1 - s)^2 ds \right).$$

Integrating \mathbf{y} over D yields inequality (3.2) for the restriction of q to $\mathcal{B}_{(r,\mathbf{x})}$, and, thus, over the whole domain D since $\mathcal{B}_{(r,\mathbf{x})} \subset D$. The lemma's inequality is true for every $q \in H^2(D)$ by density and because of the continuous embedding of $H^2(D)$ in $C^0(D)$ [1]. \square

LEMMA 3.3. For any $D \in \mathcal{D}_h$ and any constant vector $\mathbf{u} \in \mathbb{R}^2$ such that $u_e = \mathbf{n}_e \cdot \mathbf{u}$ and $u_f = \mathbf{n}_f \cdot \mathbf{u}$ there holds:

$$|D|\|\mathbf{u}\|^2 \leq \frac{|e||f|}{\sin \theta_D} (|u_e|^2 + |u_f|^2) = \frac{2|D|}{\sin^2 \theta_D} (|u_e|^2 + |u_f|^2)$$

where θ_D is the angle between the direction of e and f introduced in Assumption (M2).

Proof. Let $\boldsymbol{\tau}_\sigma = \mathbf{n}_\sigma^\perp$ for $\sigma \in \{e, f\}$, where $(\cdot)^\perp$ indicates the counterclock-wise rotation, and observe that $\det[\mathbf{n}_e^T, \mathbf{n}_f^T] = \sin \theta_D$. Since

$$(u_e, u_f)^T = [\mathbf{n}_e^T, \mathbf{n}_f^T] \mathbf{u},$$

we immediately have that

$$\mathbf{u} = [\mathbf{n}_e^T, \mathbf{n}_f^T]^{-1} (u_e, u_f)^T = \frac{1}{\sin \theta_D} [-\boldsymbol{\tau}_f^T, \boldsymbol{\tau}_e^T] (u_e, u_f)^T.$$

As $\boldsymbol{\tau}_e \cdot \boldsymbol{\tau}_f = \mathbf{n}_e \cdot \mathbf{n}_f = \cos \theta_D$, a straightforward calculation yields:

$$\begin{aligned} \|\mathbf{u}\|^2 &= \frac{1}{\sin^2 \theta_D} (u_e, u_f) [-\boldsymbol{\tau}_f^T, \boldsymbol{\tau}_e^T]^T [-\boldsymbol{\tau}_f^T, \boldsymbol{\tau}_e^T] (u_e, u_f)^T \\ &= \frac{1}{\sin^2 \theta_D} (|u_e|^2 + |u_f|^2 - 2u_e u_f \cos \theta_D) \\ &\leq \frac{1 + |\cos \theta_D|}{\sin^2 \theta_D} (|u_e|^2 + |u_f|^2). \end{aligned}$$

The lemma follows by noting that $1 + |\cos \theta_D| \leq 2$ and multiplying by $|D| = (1/2)|e||f| \sin \theta_D$. \square

LEMMA 3.4. For any $D \in \mathcal{D}_h$ and any vector field $\mathbf{u} \in (H^1(D))^2$ there holds:

$$|D| \|(\mathbf{u})_D^{\mathbf{I}}\|^2 \leq \frac{(1+2\gamma)C^{Agm}}{\sin \theta_D} \sum_{i=0}^1 h_D^{2i} \|\mathbf{u}\|_{H^i(D)}^2.$$

Proof. We apply the Jensen's inequality and the Agmon inequality, c.f. (2.6), to derive

$$|u_\sigma|^2 = \left| \frac{1}{|\sigma|} \int_\sigma \mathbf{u} \cdot \mathbf{n}_\sigma dS \right|^2 \leq \frac{1}{|\sigma|} \int_\sigma |\mathbf{u}|^2 dS \leq \frac{C^{Agm}}{|\sigma|} \sum_{i=0}^1 h_D^{2i-1} \|\mathbf{u}\|_{H^i(D)}^2, \quad (3.3)$$

which holds for $\sigma \in \{e, f_1, f_2\}$. Equation (2.3) implies that $|f|u_f = |f_1|u_{f_1} + |f_2|u_{f_2}$, and then

$$|u_f|^2 \leq \frac{|f_1| + |f_2|}{|f|} \sum_{i=1}^2 \frac{|f_i|}{|f|} |u_{f_i}|^2 \leq \frac{2C^{Agm}\gamma}{|f|} \sum_{i=0}^1 h_D^{2i-1} \|\mathbf{u}\|_{H^i(D)}^2. \quad (3.4)$$

Starting from the result of Lemma 3.3 and bounding u_e and u_f through inequalities (3.3)-(3.4), we obtain

$$|D| \|\mathbf{u}_D^{\mathbf{I}}\|^2 \leq \frac{|e||f|}{\sin \theta_D} (|u_e|^2 + |u_f|^2) \leq C^{Agm} \frac{|e||f|}{\sin \theta_D} \left(\frac{1}{|e|} + \frac{2\gamma}{|f|} \right) \sum_{i=0}^1 h_D^{2i-1} \|\mathbf{u}\|_{H^i(D)}^2$$

from which the lemma follows since $|f| + 2\gamma|e| \leq (1+2\gamma)h_D$. \square

DEFINITION 3.5. For every $\bar{q} \in Q_h \times \delta Q_h$, let us introduce the quantity:

$$\mathcal{G}_D^2(\bar{q}) = \frac{1}{4|D|} \left(|e|^2 (q_{T_2} - q_{T_1})^2 + |f|^2 (q_{V_2} - q_{V_1})^2 \right) \quad (3.5)$$

with $q_{T_2} = q_e$ when $e \subset \Gamma$.

The quantity $\mathcal{G}_D^2(\bar{q})$ satisfies the useful properties stated in the two following lemmas.

LEMMA 3.6. For any $D \in \mathcal{D}_h$ and $\bar{q} \in Q_h \times \delta Q_h$ there holds

$$(1 - |\cos \theta_D|) \mathcal{G}_D^2(\bar{q}) \leq |D| \|\nabla_D \bar{q}\|^2 \leq (1 + |\cos \theta_D|) \mathcal{G}_D^2(\bar{q}).$$

Proof. By definition (2.10) and since $\mathbf{n}_e \cdot \mathbf{n}_f = \cos \theta_D$,

$$\begin{aligned} \|\nabla_D \bar{q}\|^2 &= \frac{1}{4|D|^2} \left(|e|^2 (q_{T_2} - q_{T_1})^2 + |f|^2 (q_{V_2} - q_{V_1})^2 \right. \\ &\quad \left. + 2|e||f|(q_{T_2} - q_{T_1})(q_{V_2} - q_{V_1}) \cos \theta_D \right). \end{aligned}$$

The result follows from $2|e||f||q_{T_2} - q_{T_1}||q_{V_2} - q_{V_1}| \leq 4|D|\mathcal{G}_D^2(\bar{q})$. \square

LEMMA 3.7. For any $D \in \mathcal{D}_h$, $q \in H^2(\Omega)$, and $q^{\mathbf{J}} \in Q_h \times \delta Q_h$ the interpolation of q defined in accordance with (2.16a)-(2.16b), there holds

$$|D| \mathcal{G}_D^2(q^{\mathbf{J}}) \leq 6\rho_0 \sum_{i=0}^2 h_D^{2i} \|q\|_{H^i(D)}^2. \quad (3.6)$$

Proof. By definition (3.5), recalling that $q_{T_i}^{\mathbf{J}} = q(\mathbf{x}_{T_i})$ and $q_{V_i}^{\mathbf{J}} = q(\mathbf{x}_{V_i})$ for $i = 1, 2$, noting that $\max(|e|, |f|) \leq h_D$, and using the result of Lemma 3.2 yields:

$$|D| \mathcal{G}_D^2(q^{\mathbf{J}}) \leq \frac{h_D^2}{2|D|} \sum_{i=1}^2 |D| \left(|q_{T_i}^{\mathbf{J}}|^2 + |q_{V_i}^{\mathbf{J}}|^2 \right) \leq \frac{6h_D^2}{|D|} \sum_{i=0}^2 h_D^{2i} \|q\|_{H^i(D)}^2.$$

We get inequality (3.6) by introducing the mesh regularity constant ρ_0 , c.f. (2.4). \square

LEMMA 3.8. Let $q \in H^2(\Omega)$, $q^J \in Q_h \times \delta Q_h$ its interpolation in accordance with (2.16a)-(2.16b), $C = 3$ if $k = 0$, and $C = 6 + 24\rho_0^2\rho_1^2$ if $k = 1$ where ρ_0 and ρ_1 are the mesh regularity constants defined in (M1) and (2.4). Then,
(i) for every $T \in \mathcal{T}_h$ and $e \in \partial T$ there holds

$$|D_e| |R_T^1 q^J(\mathbf{x}_e)|^2 \leq C \sum_{i=0}^2 h_S^{2i} \|q\|_{H^i(\mathcal{S}(\mathbf{g}_T))}^2,$$

where D_e is the diamond associated to e and $h_S = \max_{D \in \mathcal{S}(\mathbf{g}_T)} h_D$;

(ii) for every $V \in \mathcal{V}_h$ and $f_i \in \partial V$ with $i = 1, 2$ there holds

$$|D_{f_i}| |R_V^1 q^J(\mathbf{x}_{f_i})|^2 \leq C \sum_{i=0}^2 h_S^{2i} \|q\|_{H^i(\mathcal{S}(\mathbf{g}_V))}^2, \quad (3.7)$$

where D_{f_i} is the diamond associated to f_i , and $h_S = \max_{D \in \mathcal{S}(\mathbf{g}_V)} h_D$.

Proof. We start proving the inequality of item (i). Using the piecewise linear reconstruction of q^J , c.f. (2.22a), yields:

$$|R_T^1 q^J(\mathbf{x}_e)|^2 \leq 2 \left(|q_T^J|^2 + \|\mathbf{g}_T(\nabla_h q^J)\|^2 |\mathbf{x}_e - \mathbf{x}_T|^2 \right). \quad (3.8)$$

As, by definition, $q_T^J = q(\mathbf{x}_T)$, c.f. (2.16a), and $\mathbf{x}_T \in D$ we apply Lemma 3.2 to derive the inequality:

$$|D_e| |q_T^J|^2 \leq 3 \sum_{i=0}^2 h_{D_e}^{2i} \|q\|_{H^i(D_e)}^2, \quad (3.9)$$

which allows us to control the first term of the right-hand side of (3.8). To derive a similar bound for the other term in (3.8), we start from gradient definition (2.10), we note that $\alpha_{T,D} > 0$ from (G1), we introduce the symbol $\mathcal{G}_D^2(q^J)$, c.f. (3.5) and apply Lemma 3.6 (with $1 + |\cos \theta_D| \leq 2$):

$$\|\mathbf{g}_T(\nabla_h q^J)\|^2 \leq \sum_{D \in \mathcal{S}(\mathbf{g}_T)} \alpha_{T,D} \|\nabla_D q^J\|^2 \leq 2 \sum_{D \in \mathcal{S}(\mathbf{g}_T)} \frac{1}{|D|} \mathcal{G}_D^2(q^J). \quad (3.10)$$

Using inequality (3.6) from Lemma 3.7 into (3.10) yields

$$\|\mathbf{g}_T(\nabla_h q^J)\|^2 \leq 12\rho_0 \sum_{D \in \mathcal{S}(\mathbf{g}_T)} \frac{1}{|D|^2} \sum_{i=0}^2 h_D^{2i} \|q\|_{H^i(D)}.$$

Note that $|D_e| h_{D_e}^2 / |D|^2 = (h_{D_e}^2 / |D_e|) (|D_e| / |D|)^2 \leq \rho_0 \rho_1^2$, where ρ_1 is the constant of Assumption (M3). Therefore, by multiplying both sides of (3.10) by $|D_e| h_{D_e}^2$, rearranging the summations and introducing $h_S = \max_{D \in \mathcal{S}(\mathbf{g}_T)} h_D$ yields:

$$|D_e| h_{D_e}^2 \|\mathbf{g}_T(\nabla_h q^J)\|^2 \leq 12\rho_0^2 \rho_1^2 \sum_{i=0}^2 h_S^{2i} \|q\|_{H^i(\mathcal{S}(\mathbf{g}_T))}^2. \quad (3.11)$$

Combining (3.9) and (3.11) into (3.8), and noting that $|\mathbf{x}_e - \mathbf{x}_T| \leq h_{D_e}$ yields

$$|D_e| |R_T^1 q^J(\mathbf{x}_e)|^2 \leq 6 \sum_{i=0}^2 h_{D_e}^{2i} \|q\|_{H^i(D_e)}^2 + 24\rho_0^2 \rho_1^2 \sum_{i=0}^2 h_S^{2i} \|q\|_{H^i(\mathcal{S}(\mathbf{g}_T))}^2,$$

from which it is easy to derive the first lemma's inequality with the given expression of the constant C .

Lemma's inequality (3.7) follows by adapting the previous argument to the reconstruction $R_V^1 q^J$ that is defined over the dual cells $V \in \mathcal{V}_h$. \square

3.2. Consistency of numerical fluxes.

LEMMA 3.9. *There exists a positive constant C^d , which is independent of h , such that for every $q \in H^2(\Omega)$, there holds*

$$\|(\mathbf{K}\nabla q)^{\mathbf{I}} - \mathbf{K}_h \nabla_h q^{\mathbf{J}}\|_{L^2(\Omega)} \leq C^d h \|q\|_{H^2(\Omega)}. \quad (3.12)$$

Proof. Let us consider a diamond-shaped cell $D \in \mathcal{D}_h$. Adding and subtracting the term $\mathbf{K}_D(\nabla q)^{\mathbf{I}}$ to the local consistency error of the cell D yields

$$(\mathbf{K}\nabla q)_D^{\mathbf{I}} - \mathbf{K}_D \nabla_D q^{\mathbf{J}} \leq ((\mathbf{K} - \mathbf{K}_D)\nabla q)_D^{\mathbf{I}} - \mathbf{K}_D((\nabla q)^{\mathbf{I}} - \nabla_D q^{\mathbf{J}}).$$

Following the proof of Lemma 3.4 with $\mathbf{u} = \nabla q$, and recalling (2.18), it is easy to show that

$$\begin{aligned} |D| \|((\mathbf{K} - \mathbf{K}_D)\nabla q)_D^{\mathbf{I}}\|^2 &\leq \|\mathbf{K} - \mathbf{K}_D\|_{L^\infty(D)}^2 \frac{(1+2\gamma)C^{Agm}}{\sin \theta_D} \sum_{i=0}^1 h_D^{2i} \|\nabla q\|_{H^i(D)}^2 \\ &\leq \left(\frac{(1+2\gamma)C^{Agm}C_{\mathbf{K}}^2}{\sin \theta_D} \right) h_D^2 \left(\|q\|_{H^1(D)}^2 + h_D^2 \|q\|_{H^2(D)}^2 \right). \end{aligned} \quad (3.13)$$

Now, let us consider the vector functional $\mathcal{J} : H^2(D) \rightarrow \mathbb{R}^2$ defined by

$$\mathcal{J}(q) = \mathbf{K}_D((\nabla q)_D^{\mathbf{I}} - \nabla_D(q)^{\mathbf{J}}).$$

Lemma 2.3 implies that $\mathcal{J}(q^{(1)}) = 0$ for every $q^{(1)} \in P_1(D)$. Therefore, $\mathcal{J}(\cdot)$ satisfies condition (i) of Lemma 3.1 for $k = 1$. To check condition (ii), we start noting that:

$$|D| \|\mathcal{J}(q)\|^2 \leq 2|D| \left(\|(\mathbf{K}_D \nabla q)_D^{\mathbf{I}}\|^2 + \|\mathbf{K}_D \nabla_D(q)^{\mathbf{J}}\|^2 \right). \quad (3.14)$$

We control both terms in the right-hand side of (3.14) through condition (2.2) and Lemmas 3.4, 3.6 and 3.7. After translating the range of the summation index i to start from 1 in the final expression, the bound of the first term is:

$$|D| \|(\mathbf{K}_D \nabla q)_D^{\mathbf{I}}\|^2 \leq (\kappa^*)^2 |D| \|(\nabla q)_D^{\mathbf{I}}\|^2 \leq \frac{(1+2\gamma)C^{Agm}(\kappa^*)^2}{\sin \theta_D} \sum_{i=1}^2 h_D^{2(i-1)} \|q\|_{H^i(D)}^2. \quad (3.15)$$

Likewise, after noting that $|D|^{-1} \leq \rho_0 h_D^{-2}$, for the second term we get:

$$\begin{aligned} |D| \|\mathbf{K}_D \nabla_D(q)^{\mathbf{J}}\|^2 &\leq 2 \left(\frac{(\kappa^*)^2}{|D|} \right) (|D| \mathcal{G}_D^2(q^{\mathbf{J}})) \leq \frac{12(\kappa^*)^2 \rho_0}{|D|} \sum_{i=0}^2 h_D^{2i} \|q\|_{H^i(D)}^2 \\ &\leq 12(\kappa^* \rho_0)^2 \sum_{i=0}^2 h_D^{2(i-1)} \|q\|_{H^i(D)}^2. \end{aligned} \quad (3.16)$$

Finally, we verify the second hypothesis of Lemma 3.1 by substituting (3.15) and (3.16) into (3.14). The estimates are

$$|D| \|\mathcal{J}(q)\|^2 \leq (\kappa^*)^2 \left(\frac{(1+2\gamma)C^{Agm}}{\sin \theta_D} + 12(\rho_0)^2 \right) C^{Intp} h_D^2 \|q\|_{H^2(D)}^2,$$

and, combined with (3.13),

$$\|(\mathbf{K}\nabla q)^{\mathbf{I}} - \mathbf{K}_D \nabla_D q^{\mathbf{J}}\|_{L^2(D)} \leq C^d h_D \|q\|_{H^2(D)}, \quad (3.17)$$

with

$$(C^d)^2 = \left(\left((1+2\gamma)(\kappa^*)^2 + 3(C_{\mathbf{K}})^2 \right) \frac{C^{Agm}}{\sin \theta} + 12(\kappa^*)^2(\rho_0)^2 \right) C^{Intp},$$

since $\sin \theta_D \geq \sin \theta > 0$ from Assumption (M2). The constant C^d depends on K , ρ_0 , θ_D , and γ and on the constants C^{Agm} and C^{Intp} . Inequality (3.12) eventually follows from the summation over $D \in \mathcal{D}_h$ of (3.17). \square

LEMMA 3.10. *Let \mathbf{b} be a constant vector on Ω . Then, there exists a positive constant C^c which is independent of h such that for every $q \in H^{1+k}(\Omega)$ there holds:*

$$\|(\mathbf{b}q)^{\mathbb{I}} - \Phi_h^{c,k}(q^{\mathbb{J}})\|_{L^2(\Omega)} \leq C^c h^{1+k} \|q\|_{H^{1+k}(\Omega)},$$

where $k = 0, 1$ is the order of the polynomial reconstruction used to define the numerical convection flux.

Proof. For every $D \in \mathcal{D}_h$ we consider the functional

$$\mathcal{J}(q) = (\mathbf{b}q)_D^{\mathbb{I}} - \Phi_D^{c,k}(q^{\mathbb{J}}),$$

which is well defined for every $q \in H^{1+k}(\mathcal{S}) \cap C^0(\overline{\mathcal{S}})$ and $\mathcal{S} = D$ for $k = 0$ and $\mathcal{S} = \mathcal{S}(\mathbf{g}_{T_1}) \cup \mathcal{S}(\mathbf{g}_{T_2}) \cup \mathcal{S}(\mathbf{g}_{V_1}) \cup \mathcal{S}(\mathbf{g}_{V_2})$ for $k = 1$. Moreover, we conventionally take $\mathcal{S}(\mathbf{g}_{T_i}) = \mathcal{S}(\mathbf{g}_{V_i}) = D$ in the case $k = 0$ and $\mathcal{S}(\mathbf{g}_{T_2}) = D$ if $e \subset \Gamma$ for $k = 0, 1$. In view of Lemma 2.4, there holds that $\mathcal{J}(q^{(k)}) = 0$ for every $q^{(k)} \in P_k(\mathcal{S})$, implying that the first hypothesis of Lemma 3.1, i.e. condition (i), is satisfied. To verify the second hypothesis of Lemma 3.1, we start splitting the functional $\mathcal{J}(\cdot)$ in two terms:

$$|D| \|\mathcal{J}(q)\|^2 \leq 2|D| \left(\|(\mathbf{b}q)_D^{\mathbb{I}}\|^2 + \|\Phi_D^{c,k}(q^{\mathbb{J}})\|^2 \right). \quad (3.18)$$

We bound the first term in the right-hand side of (3.18) thanks to the result of Lemma 3.4:

$$|D| \|(\mathbf{b}q)_D^{\mathbb{I}}\|^2 \leq \|\mathbf{b}\|_{L^\infty(\Omega)}^2 (1 + 2\gamma) \frac{C^{Agm}}{\sin \theta_D} \sum_{i=0}^1 h_D^{2i} \|q\|_{H^i(D)}^2. \quad (3.19)$$

Using (2.19) and applying Lemma 3.3, we get

$$|D| \|\Phi_D^{c,k}(q^{\mathbb{J}})\|^2 \leq \frac{2|D|}{\sin^2 \theta_D} \left(|\Phi_e^{c,k}|^2 + \gamma \left(\frac{|f_1|}{|f|} |\Phi_{f_1}^{c,k}|^2 + \frac{|f_2|}{|f|} |\Phi_{f_2}^{c,k}|^2 \right) \right).$$

These terms are all defined in subsection 2.5. Since $0 \leq \theta_D \leq 1$ and using (2.20a),

$$\begin{aligned} |D| |\Phi_e^{c,k}(q^{\mathbb{J}})|^2 &\leq |D| |b_e|^2 (|R_{T_1}^k q^{\mathbb{J}}(\mathbf{x}_e)|^2 + |R_{T_2}^k q^{\mathbb{J}}(\mathbf{x}_e)|^2) \\ &\leq C |b_e|^2 \sum_{i=0}^2 h_S^{2i} \left(\|q\|_{H^i(\mathcal{S}(\mathbf{g}_{T_1}))}^2 + \|q\|_{H^i(\mathcal{S}(\mathbf{g}_{T_2}))}^2 \right). \end{aligned} \quad (3.20)$$

For $k = 1$, $|R_{T_1}^1 q^{\mathbb{J}}(\mathbf{x}_e)|$ and $|R_{T_2}^1 q^{\mathbb{J}}(\mathbf{x}_e)|$ are bounded by Lemma 3.8 for any internal edge, and $|R_{T_2}^1 q^{\mathbb{J}}(\mathbf{x}_e)|$ is bounded by Lemma 3.2 for any boundary edge, and the constant is $C = 6 + 24\rho_0^2\rho_1^2$. For $k = 0$, both terms are bounded using Lemma 3.2 and the constant is $C = 3$. The same argument applies to the other two terms, which refers to the convection flux across f_1 and f_2 , to obtain the similar bound

$$|D| |\Phi_{f_i}^{c,k}(q^{\mathbb{J}})|^2 \leq C |b_{f_i}|^2 \sum_{i=0}^2 h_S^{2i} \left(\|q\|_{H^i(\mathcal{S}(\mathbf{g}_{V_1}))}^2 + \|q\|_{H^i(\mathcal{S}(\mathbf{g}_{V_2}))}^2 \right). \quad (3.21)$$

Combining inequalities (3.20)-(3.21) yields

$$\begin{aligned} |D| \|\Phi_D^{c,k}(q^{\mathbb{J}})\|^2 &\leq \frac{2C \|\mathbf{b}\|_{L^\infty(\Omega)}^2}{\sin^2 \theta_D} \sum_{i=0}^2 h_S^{2i} \sum_{j=1}^2 \left(\|q\|_{H^i(\mathcal{S}(\mathbf{g}_{T_j}))}^2 + \gamma^2 \|q\|_{H^i(\mathcal{S}(\mathbf{g}_{V_j}))}^2 \right) \\ &\leq \frac{2C \|\mathbf{b}\|_{L^\infty(\Omega)}^2}{\sin^2 \theta_D} N_S \max(1, \gamma^2) \sum_{i=0}^2 h_S^{2i} \|q\|_{H^i(\mathcal{S})}^2, \end{aligned} \quad (3.22)$$

because

$$\begin{aligned} \sum_{j=1}^2 \left(\|q\|_{H^i(\mathcal{S}(\mathbf{g}_{T_j}))}^2 + \gamma^2 \|q\|_{H^i(\mathcal{S}(\mathbf{g}_{V_j}))}^2 \right) &\leq N_{\mathcal{S}} \max(1, \gamma^2) \sum_{D \in \mathcal{S}} \|q\|_{H^i(D)}^2 \\ &= N_{\mathcal{S}} \max(1, \gamma^2) \|q\|_{H^i(\mathcal{S})}^2. \end{aligned}$$

Finally, we get condition (ii) of Lemma 3.1 by substituting (3.19) and (3.22) into (3.18).

The global inequality of the lemma is eventually obtained from the summation over $D \in \mathcal{D}_h$ by using Assumption (M2) to bound $\sin \theta_D$ from below and Lemma 3.1 with $\alpha = 0$. Remark that the summation over D is obvious when $k = 0$ because $\mathcal{S} = D$. Instead, when $k = 1$ we take $\mathcal{S} = \mathcal{S}(\mathbf{g}_{T_1}) \cup \mathcal{S}(\mathbf{g}_{T_2}) \cup \mathcal{S}(\mathbf{g}_{V_1}) \cup \mathcal{S}(\mathbf{g}_{V_2})$, and the supports overlap in the summation process. However, Assumption (G2) implies that

$$\#\{D' \in \mathcal{D}_h \text{ such that } D' \subset \mathcal{S}(D)\} \leq N_{T_1}(D) + N_{T_2}(D) + N_{V_1}(D) + N_{V_2}(D) \leq 4N_{\mathcal{S}}$$

and, thus, for all integers $i \in [0, 2]$ we have

$$\sum_{D \in \mathcal{D}_h} \|q\|_{H^i(\mathcal{S})}^2 \leq 4N_{\mathcal{S}} \|q\|_{H^i(\Omega)}^2.$$

□

LEMMA 3.11. *Let $\bar{q} = (q, \delta q)$ be an element of $Q_h \times \delta Q_h$ with $\delta q = 0$. If $\phi \in [1/2, 1]$, then*

(i) *for $k = 0$ there holds:*

$$-\left(\Phi_h^{c,0}(\bar{q}), \nabla_h \bar{q}\right)_{X_h} \geq 0; \quad (3.23)$$

(ii) *for $k = 1$ there exists a positive constant C such that there holds:*

$$-\left(\Phi_h^{c,1}(\bar{q}), \nabla_h \bar{q}\right)_{X_h} \geq Ch \|\nabla_h \bar{q}\|_{X_h}^2.$$

Proof. Let us take $q_{T_2} = q_e = 0$ for $e \subset \Gamma$ and $q_{V_1} = q_{V_2} = 0$ for $V_1, V_2 \in \mathcal{V}_h^{\text{Bnd}}$. In this proof we will find it convenient to use the symbol $\mathcal{D}_h^{\text{Bnd}}$ to denote the collection of boundary diamond-shaped cells, i.e. those diamonds such that $e \subset \Gamma$ and that degenerate into a triangle. We will also use $\mathcal{D}_h^{\text{Int}} = \mathcal{D}_h \setminus \mathcal{D}_h^{\text{Bnd}}$, i.e. those diamond-shaped cells such that $e = T_1 \cap T_2$.

(i). For $k = 0$, a classical calculation yields:

$$\begin{aligned} &-\left(\Phi_h^{c,0}(\bar{q}), \nabla_h \bar{q}\right)_{X_h} \\ &= \left(\phi - \frac{1}{2}\right) \left(\sum_{T \in \mathcal{T}_h} |q_T|^2 \int_T \operatorname{div}(\mathbf{b}) dV + \sum_{V \in \mathcal{V}_h} |q_V|^2 \int_V \operatorname{div}(\mathbf{b}) dV \right. \\ &\quad \left. + \sum_{D \in \mathcal{D}_h} \left(|e| |b_e| (q_{T_1} - q_{T_2})^2 + \sum_{i=1}^2 |f_i| |b_{f_i}| (q_{V_1} - q_{V_2})^2 \right) \right), \end{aligned}$$

and the coercivity condition $\operatorname{div}(\mathbf{b}) \geq 0$ implies (3.23) for $1/2 \leq \phi \leq 1$.

(ii). Let us observe that the flux for $k = 1$ can be written by adding a correction term to the flux for $k = 0$, c.f. the linearity of the definitions given in subsection 2.5. Formally,

$$\Phi_h^{c,1}(\bar{q}) = \Phi_h^{c,0}(\bar{q}) + \mathbf{r}_h(\bar{q}).$$

Of course, the result proved in (i) implies the development:

$$-\left(\Phi_h^{c,1}(\bar{q}), \nabla_h \bar{q}\right)_{X_h} = -\left(\Phi_h^{c,0}(\bar{q}) + \mathbf{r}_h(\bar{q}), \nabla_h \bar{q}\right)_{X_h} \geq -(\mathbf{r}_h(\bar{q}), \nabla_h \bar{q})_{X_h},$$

and the Cauchy-Schwarz inequality gives:

$$|(\mathbf{r}_h(\bar{q}), \nabla_h \bar{q})_{X_h}| \leq \|\nabla_h \bar{q}\|_{X_h} \|\mathbf{r}_h(\bar{q})\|_{X_h}.$$

To derive a bound for $\|\mathbf{r}_h(\bar{q})\|_{X_h}$, we start applying Lemma 3.3,

$$\|\mathbf{r}_D(\bar{q})\|^2 \leq \frac{2}{\sin^2 \theta_D} \left(\|r_e(\bar{q})\|^2 + \gamma \left(\frac{|f_1|}{|f|} \|r_{f_1}(\bar{q})\|^2 + \frac{|f_2|}{|f|} \|r_{f_2}(\bar{q})\|^2 \right) \right), \quad (3.24)$$

and then we separately control the three terms in the right-hand side of (3.24) by the same argument. We obtain an expression for r_σ , $\sigma \in \{e, f_1, f_2\}$, by directly substituting the formulas for the reconstructions R_T^1 and R_V^1 in the formulas of the numerical convection flux, c.f. subsection 2.5. For example, when e is an internal edge, the correction term takes the form

$$\begin{aligned} \mathbf{r}_e(\bar{q}) &= \mathbf{b}_e^+ (\phi \mathbf{g}_{T_1}(\nabla_h \bar{q}) \cdot (\mathbf{x}_e - \mathbf{x}_{T_1}) + (1 - \phi) \mathbf{g}_{T_2}(\nabla_h \bar{q}) \cdot (\mathbf{x}_e - \mathbf{x}_{T_2})) \\ &\quad + \mathbf{b}_e^- (\phi \mathbf{g}_{T_2}(\nabla_h \bar{q}) \cdot (\mathbf{x}_e - \mathbf{x}_{T_2}) + (1 - \phi) \mathbf{g}_{T_1}(\nabla_h \bar{q}) \cdot (\mathbf{x}_e - \mathbf{x}_{T_1})), \end{aligned} \quad (3.25)$$

from which it is immediate to obtain:

$$\|\mathbf{r}_e(\bar{q})\|^2 \leq \|b_e\|^2 \left(\|\mathbf{g}_{T_1}(\nabla_h \bar{q})\|^2 |\mathbf{x}_e - \mathbf{x}_{T_1}|^2 + \|\mathbf{g}_{T_2}(\nabla_h \bar{q})\|^2 |\mathbf{x}_e - \mathbf{x}_{T_2}|^2 \right). \quad (3.26)$$

When e is a boundary edge, equations (3.25)-(3.26) are modified by simply eliminating all the terms that refer to T_2 .

Using (3.26) and the similar inequalities that are obtained by repeating the previous argument for r_{f_1} and r_{f_2} , gives

$$\begin{aligned} \sum_{D \in \mathcal{D}_h} |D| \|\mathbf{r}_D(\bar{q})\|^2 &\leq \|\mathbf{b}\|_{L^\infty(\Omega)}^2 \sum_{D \in \mathcal{D}_h^{\text{int}}} \frac{2|D|}{\sin^2 \theta_D} h_D^2 \sum_{i=1}^2 \left(\|\mathbf{g}_{T_i}(\nabla_h \bar{q})\|^2 + \gamma^2 \|\mathbf{g}_{V_i}(\nabla_h \bar{q})\|^2 \right) \\ &\quad + \|\mathbf{b}\|_{L^\infty(\Omega)}^2 \sum_{D \in \mathcal{D}_h^{\text{bnd}}} \frac{2|D|}{\sin^2 \theta_D} h_D^2 \left(\|\mathbf{g}_{T_1}(\nabla_h \bar{q})\|^2 + \gamma^2 \sum_{i=1}^2 \|\mathbf{g}_{V_i}(\nabla_h \bar{q})\|^2 \right). \end{aligned} \quad (3.27)$$

Now observe that Assumptions (G1)-(G2) imply

$$\|\mathbf{g}_T(\nabla_h \bar{q})\|^2 \leq \sum_{D \in \mathcal{S}(\mathbf{g}_T)} \alpha_{T,D} \|\nabla_D \bar{q}\|^2 \leq \sum_{D \in \mathcal{S}(\mathbf{g}_T)} \|\nabla_D \bar{q}\|^2,$$

and summing over \mathcal{D}_h yields

$$\begin{aligned} \sum_{D \in \mathcal{D}_h} |D| \left(\|\mathbf{g}_{T_1}(\nabla_h \bar{q})\|^2 + \|\mathbf{g}_{T_2}(\nabla_h \bar{q})\|^2 \right) &\leq \sum_{D \in \mathcal{D}_h} |D| \sum_{i=1}^2 \sum_{\hat{D} \in \mathcal{S}(\mathbf{g}_{T_i})} \|\nabla_{\hat{D}} \bar{q}\|^2 \\ &\leq \rho_1 \sum_{D \in \mathcal{D}_h} \sum_{i=1}^2 \sum_{\hat{D} \in \mathcal{S}(\mathbf{g}_{T_i})} |\hat{D}| \|\nabla_{\hat{D}} \bar{q}\|^2 \leq \rho_1 \sum_{\hat{D} \in \mathcal{S}(\mathbf{g}_{T_i})} N_T(\hat{D}) |\hat{D}| \|\nabla_{\hat{D}} \bar{q}\|^2 \leq \rho_1 N_S \|\nabla_h \bar{q}\|_{X_h}^2. \end{aligned}$$

A similar expression can be derived for the “ V ”-terms. Therefore, the desired bound is derived from (3.27) and Assumption (M2):

$$\|\mathbf{r}_h(\bar{q})\|_{X_h}^2 = \sum_{D \in \mathcal{D}_h} |D| \|\mathbf{r}_D(\bar{q})\|^2 \leq \left(\frac{2(1 + \gamma^2) \|\mathbf{b}\|_{L^\infty(\Omega)}^2 \rho_1 N_S}{\sin^2 \theta} \right) h^2 \|\nabla_h \bar{q}\|_{X_h}^2.$$

□

3.3. Convergence analysis. The following analysis is classical to compute error estimates in the linear case [21].

DEFINITION 3.12 (Consistency errors).

The consistency errors for the convection and diffusion flux are defined in X_h for any $q \in H^2(\Omega)$ by:

$$\mathcal{R}_h^c(q) = (\mathbf{b}q)^{\mathbf{I}} - \Phi_h^{c,k}(q^{\mathbf{J}}) \quad \text{and} \quad \mathcal{R}_h^d(q) = (\mathbf{K}\nabla q)^{\mathbf{I}} - \Phi_h^d(q^{\mathbf{J}}). \quad (3.28)$$

DEFINITION 3.13 (Approximation errors). The error $\bar{\mathcal{E}}_h \in Q_h \times \delta Q_h$ is given by

$$\bar{\mathcal{E}}_h = \bar{p}_h - p^{\mathbf{J}} = (\mathcal{E}_h, 0) \quad \text{with } \mathcal{E}_h \in Q_{h,0}$$

because we took $\bar{p}_h|_{\delta Q_h} = g^{\mathbf{J}} = p^{\mathbf{J}}|_{\delta Q_h}$ and $p_h \in Q_{h,g}$ implies that $p_V = g(\mathbf{x}_V) = p_V^{\mathbf{J}}$ for every $V \in \mathcal{V}_h^{\text{int}}$.

PROPOSITION 3.14 (Error equation).

$$\left(\Phi_h^d(\mathcal{E}_h) - \Phi_h^{c,k}(\mathcal{E}_h), \nabla_h \bar{\mathcal{E}}_h \right)_{X_h} = \left(\mathcal{R}_h^d(p) - \mathcal{R}_h^c(p), \nabla_h \bar{\mathcal{E}}_h \right)_{X_h}. \quad (3.29)$$

Proof. From (2.25) with $q = \mathcal{E}_h \in Q_{h,0}$ we have

$$\left(\Phi_h^d(p_h) - \Phi_h^{c,k}(p_h), \nabla_h \bar{\mathcal{E}}_h \right)_{X_h} = (f^{\mathbf{I}}, \mathcal{E}_h)_{Q_h} \quad (3.30)$$

Using the flux interpolation operators, the result of Lemma 2.2, and the discrete integration-by-part formula stated in Lemma 2.1 yield:

$$\begin{aligned} (f^{\mathbf{I}}, \mathcal{E}_h)_{Q_h} &= \left((\operatorname{div}(\mathbf{b}p - \mathbf{K}\nabla p))^{\mathbf{I}}, \mathcal{E}_h \right)_{Q_h} = \left(\operatorname{div}_h(\mathbf{b}p - \mathbf{K}\nabla p)^{\mathbf{I}}, \mathcal{E}_h \right)_{Q_h} \\ &= \left((\mathbf{K}\nabla p - \mathbf{b}p)^{\mathbf{I}}, \nabla_h \bar{\mathcal{E}}_h \right)_{X_h}. \end{aligned} \quad (3.31)$$

We terminate the derivation of (3.29) by comparing (3.30) and (3.31), subtracting from both sides the quantity $\left(\Phi_h^d(p^{\mathbf{J}}) - \Phi_h^{c,k}(p^{\mathbf{J}}), \nabla_h \bar{\mathcal{E}}_h \right)_{X_h}$, and using equations (3.28). □

THEOREM 3.15. Let $p \in H^2(\Omega)$ be the exact solution of problem (2.1) and $p_h \in Q_h$ its DDFV approximations on a family of meshes $\{\mathcal{T}_h\}_h$ satisfying the mesh regularity assumptions (M1)-(M3). Then, there exists a positive constant C independent of h such that

$$\|\nabla_h \bar{\mathcal{E}}_h\|_{X_h} \leq C (h\|p\|_{H^2(\Omega)} + h^{1+k}\|p\|_{H^{1+k}(\Omega)}) \quad \text{for } k = 0, 1. \quad (3.32)$$

Proof. Using the ellipticity of \mathbf{K} , the error equation, the Cauchy-Schwarz inequality and the non-negativity result of Lemma 3.11 yields:

$$\begin{aligned} \kappa_* \|\nabla_h \bar{\mathcal{E}}_h\|_{X_h}^2 &\leq (\mathbf{K}_h \nabla_h \bar{\mathcal{E}}_h, \nabla_h \bar{\mathcal{E}}_h)_{X_h} = (\Phi_h^d(\mathcal{E}_h), \nabla_h \bar{\mathcal{E}}_h)_{X_h} \\ &\leq (\mathcal{R}_h^d(p) - \mathcal{R}_h^c(p), \nabla_h \bar{\mathcal{E}}_h)_{X_h} + \left(\Phi_h^{c,k}(\mathcal{E}_h), \nabla_h \bar{\mathcal{E}}_h \right)_{X_h} \\ &\leq (\|\mathcal{R}_h^d(p) - \mathcal{R}_h^c(p)\|_{X_h} + Ckh\|\nabla_h \bar{\mathcal{E}}_h\|_{X_h}) \|\nabla_h \bar{\mathcal{E}}_h\|_{X_h}. \end{aligned}$$

From (3.32) we obtain that:

$$(\kappa_* - Ckh)\|\nabla_h \bar{\mathcal{E}}_h\|_{X_h} \leq \|\mathcal{R}_h^d(p) - \mathcal{R}_h^c(p)\|_{X_h}. \quad (3.33)$$

Then, we apply the triangle inequality to the right-hand side of (3.33), and use the results of Lemmas 3.9 and 3.10 to obtain (3.32), for all h if $k = 0$ and for all $h \leq \kappa_*/C$ if $k = 1$ with $C = \|\mathbf{b}\|_{L^\infty(\Omega)}(2(1 + \gamma^2)\rho_1 N_S)^{1/2}/\sin \theta_D$, which is the constant of Lemma 3.11. □

LEMMA 3.16. *Under Assumptions (M1)-(M2) there exists a positive constant C independent of h such that for every $\bar{q} = (q, 0) \in Q_h \times \delta Q_h$*

$$\|q\|_{Q_h} \leq C \|\nabla_h \bar{q}\|_{X_h}.$$

Proof. From 3.6 there holds:

$$|D| \|\nabla_D \bar{q}\|^2 \geq (1 - |\cos \theta_D|) \frac{1}{4|D|} \left(|e|^2 (q_{T_2} - q_{T_1})^2 + |f|^2 (q_{V_2} - q_{V_1})^2 \right).$$

The result follows from the classical discrete Poincaré inequality [21] applied separately to \mathcal{T}_h and \mathcal{V}_h , and Assumption (M2) to bound $\cos \theta_D$. \square

THEOREM 3.17. *Let us identify $\nabla_h \bar{p}_h \in X_h$ with the vector-valued function*

$$\nabla_h \bar{p}_h(\mathbf{x}) = \sum_{D \in \mathcal{D}_h} \nabla_D \bar{p}_h \chi_D(\mathbf{x}),$$

and consider the scalar function

$$\hat{p}_h(\mathbf{x}) = \frac{1}{2} \left(\sum_{T \in \mathcal{T}_h} p_T \chi_T(\mathbf{x}) + \sum_{V \in \mathcal{V}_h} p_V \chi_V(\mathbf{x}) \right)$$

for $\mathbf{x} \in \Omega$, where χ_D , χ_T , and χ_V are the characteristic functions of D , T , and V , respectively. Then, there exists a positive constant C such that

$$\|p - \hat{p}_h\|_{L^2(\Omega)} + \|\nabla p - \nabla_h \bar{p}_h\|_{L^2(\Omega)} \leq Ch \|p\|_{H^2(\Omega)}.$$

Proof. Following the proof of Lemma 3.9, we have

$$\|\nabla p - \nabla_h p^J\|_{L^2(\Omega)} \leq Ch \|p\|_{H^2(\Omega)}. \quad (3.34)$$

Thus, by adding and subtracting $\nabla_h p^J$, applying the triangle inequality, using (3.34), noting that $\|\nabla_h(p^J - \bar{p}_h)\|_{L^2(\Omega)} = \|\nabla_h \bar{\mathcal{E}}_h\|_{X_h}$ and applying (3.32), we obtain

$$\|\nabla p - \nabla_h \bar{p}_h\|_{L^2(\Omega)} \leq \|\nabla p - \nabla_h p^J\|_{L^2(\Omega)} + \|\nabla_h(p^J - \bar{p}_h)\|_{L^2(\Omega)} \leq Ch \|p\|_{H^2(\Omega)}.$$

Moreover,

$$\begin{aligned} \|p - \hat{p}_h\|_{L^2(\Omega)}^2 &\leq \int_{\Omega} \left| \sum_{T \in \mathcal{T}_h} (p(\mathbf{x}) - p_T) \chi_T(\mathbf{x}) \right|^2 dV + \int_{\Omega} \left| \sum_{V \in \mathcal{V}_h} (p(\mathbf{x}) - p_V) \chi_V(\mathbf{x}) \right|^2 dV \\ &\leq 2 \int_{\Omega} \left| p(\mathbf{x}) - \sum_{T \in \mathcal{T}_h} p(\mathbf{x}_T) \chi_T(\mathbf{x}) \right|^2 dV + 2 \int_{\Omega} \left| p(\mathbf{x}) - \sum_{V \in \mathcal{V}_h} p(\mathbf{x}_V) \chi_V(\mathbf{x}) \right|^2 dV \\ &\quad + 4 \|p^J - p_h\|_{Q_h}^2 \\ &\leq Ch^2 \|p\|_{H^2(\Omega)}^2. \end{aligned}$$

\square

4. Numerical experiments. The numerical experiments presented in this section aim to confirm the behavior of the current DDFV formulation expected from the theoretical considerations of the previous section. To this purpose, we solve (2.1) on the domain $\Omega =]0, 1[\times]0, 1[$ by applying the DDFV scheme with $\phi = 0.5$, i.e. the *central scheme*, and $\phi = 1$, i.e the *fully upwind scheme*, to equation (2.1) with diffusion tensor $\mathbf{K} = \mathbf{I}$ and velocity field $\mathbf{b}^T = (1, 3)^T$. The forcing term f and the boundary function g are taken according to the exact solution

$$p(x, y) = \left(x - e^{2(x-1)} \right) \left(y^2 - e^{3(y-1)} \right).$$

We consider two different sets of successively refined meshes, namely \mathcal{M}_1 and \mathcal{M}_2 . Mesh family \mathcal{M}_1 is taken from the FVCA-5 benchmark for anisotropic problems and its construction is detailed in [27]. Mesh family \mathcal{M}_2 is an example of a non-conforming refinement; more details are in [18]. The first two meshes of each mesh family are also shown in Figure 4.1, while Table 4.1 reports the number of cells in \mathcal{T}_h , \mathcal{V}_h , and \mathcal{D}_h , and the value of the mesh size h . These examples are solved by a C^{++} program based on a variant of P2MESH [5], a public domain library designed to manage data structures of two-dimensional unstructured meshes.

The approximation errors are measured by the following quantities:

$$\mathcal{E}_{Q_h}(p_h) = \frac{\|p^I - p_h\|_{Q_h}}{\|p^I\|_{Q_h}}$$

and

$$\mathcal{E}_{\mathcal{D}_h}(\nabla_h \bar{p}_h) = \frac{\left(\sum_{D \in \mathcal{D}_h} \int_D |\nabla p - \nabla_D \bar{p}_h|^2 dV \right)^{1/2}}{|p|_{H^1(\Omega)}}.$$

We report these errors and the corresponding convergence rates in Tables 4.2-4.5.

The results for the approximation of the scalar variable shows a quadratic convergence rate for the central scheme given by $\phi = 0.5$ and $k = 0, 1$, and for the second-order fully upwind scheme given by $\phi = 1$ and $k = 1$. This behavior is in agreement with that expected from a *second-order accurate* method, thus suggesting that the $\mathcal{O}(h)$ -estimate provided in Theorem 3.17 is not sharp. Instead, the numerical approximation to the scalar unknown provided by the *first-order* upwind scheme, i.e. by setting $\phi = 1$ and $k = 0$, is only linearly convergent, as expected.

Moreover, a super-convergence effect is visible in the approximation of the solution gradient when the second-order implementations of the DDFV scheme mentioned above are applied to the meshes of \mathcal{M}_1 , c.f. Tables 4.2 and 4.3. This effect disappears when the DDFV method is applied to the meshes in \mathcal{M}_2 . We explain this behavior as follows. If the diamond-shaped cell D is a parallelogram, and \mathbf{K} and \mathbf{b} are constant fields, the local consistency errors $\|(\mathbf{K}\nabla q)^I - \mathbf{K}_h \nabla_h q^J\|_{L^2(\Omega)}$ and $\|(\mathbf{b}q)^I - \Phi_h^{c,1}(q^J)\|_{L^2(\Omega)}$ can be shown to vanish for functions $q \in P_2(D)$. As a consequence, the global error estimate can be improved like in [18]:

$$\|\nabla_h \bar{\mathcal{E}}_h\|_{X_h} \leq C \left(h \|p\|_{H^2(\Omega_h^0)} + h^2 \|p\|_{H^3(\Omega_h^1)} \right),$$

where

$$\Omega_h^1 = \cup \{D \in \mathcal{D}_h \text{ such that } D \text{ is a parallelogram}\},$$

and $\Omega_h^0 = \Omega \setminus \Omega_h^1$. From Hölder inequality and the Sobolev injection $H^2(\Omega_h^0) \hookrightarrow L^p(\Omega_h^0)$, which holds for all $1 \leq p < \infty$ in 2-D [1], we get

$$\|p\|_{H^2(\Omega_h^0)} \leq C^{Sob} |\Omega_h^0|^{\frac{1}{2}(1-\frac{1}{p})} \|p\|_{H^3(\Omega)}.$$

Therefore, for every $\epsilon > 0$ there follows

$$\|\nabla_h \bar{\mathcal{E}}_h\|_{X_h} \leq C \left(h^2 + h |\Omega_h^0|^{\frac{1}{2}-\epsilon} \right) \|p\|_{H^3(\Omega)},$$

and requiring that $|\Omega_h^0| \leq \lambda h$ is sufficient to provide the observed superconvergence behavior for the gradient approximation:

$$\|\nabla_h \bar{\mathcal{E}}_h\|_{X_h} \leq C h^{\frac{3}{2}-\epsilon} \|p\|_{H^3(\Omega)}.$$

The above condition, which has already been pointed out in [18] for pure diffusion problems on triangular meshes, simply states that the size of the region covered by diamond-shaped

cells that are not parallelograms is decreasing at least like h . This occurs when some shape regularization in the mesh refinement process takes place, such as, for example, in mesh family \mathcal{M}_1 , or in the *homothetic* refinement of triangle-based meshes investigated in [18].

On the other hand, the refinement algorithm of \mathcal{M}_2 is such that $|\Omega_h^0| = \frac{7}{8}|\Omega|$, and the superconvergence effect is lost. Here, the remarkable fact is that a $\mathcal{O}(h^2)$ -convergence rate is still observed for the second-order DDFV approximations to the scalar variable, see the column $\mathcal{E}_{Q_h}(p_h)$ for $k = 1$ of Tables 4.4-4.5. Hence, we may conjecture that the second-order of accuracy of the DDFV method (and, likewise, of many other FV methods) is independent of the requirement that the diamond cells of the meshes are all becoming parallelograms when h tends to zero. Furthermore, a noteworthy consequence of this behavior is that a suitable post-processing of the numerical solution can still provide a piecewise-linear approximation to the scalar solution that is quadratically convergent in the $L^2(\Omega)$ -norm.

5. Conclusion. We considered a steady convection-diffusion problem with a full tensor coefficient discretized with the DDFV method. Under quite general assumptions on polygonal meshes, which can have non-conforming and non-convex elements, we proved the convergence of the scalar variable and its gradient. The numerical experiments confirm the theoretical prediction concerning the approximation of the gradient. In particular, the assumption that the domain is covered by diamond cells that asymptotically tend to be parallelograms seems necessary to obtain the superconvergent rate $h^{3/2}$. Instead, better accuracy is observed for the scalar solution even without this assumption, which is used in the proof proposed by [18]. How this condition could be relaxed to obtain a more general proof is still an open issue.

Acknowledgments. The authors are pleased to acknowledge the “Programma Professori Visitatori” held by “Gruppo Nazionale Calcolo Scientifico” (GNCS-InDAM), which financed the visit of Prof. Yves Coudière to IMATI-CNR in Pavia in July 2008, and the University of Nantes, which financed the visit of Dr. Gianmarco Manzini at the Laboratoire de Mathématiques “Jean Leray” in Nantes in June-July 2007.

REFERENCES

- [1] R. A. Adams. *Sobolev spaces*. Academic press, New York, 1975.
- [2] S. Agmon. *Lectures on Elliptic Boundary Value Problems*. Van Nostrand, 1965.
- [3] B. Andreianov, F. Boyer, and F. Hubert. Discrete duality finite volume schemes for Leray-Lions type elliptic problems on general 2d meshes. *Numer. Methods Partial Differential Equations*, 23(1):145–195, 2007.
- [4] T. J. Barth and D. C. Jespersen. The design and application of upwind schemes on unstructured meshes. AIAA Paper 89-0366, 1989.
- [5] E. Bertolazzi and G. Manzini. Algorithm 817 P2MESH: generic object-oriented interface between 2-D unstructured meshes and FEM/FVM-based PDE solvers. *ACM Trans. Math. Softw.*, 28(1):101–132, 2002.
- [6] E. Bertolazzi and G. Manzini. A cell-centered second-order accurate finite volume method for convection-diffusion problems on unstructured meshes. *Math. Models Methods Appl. Sci.*, 8:1235–1260, 2004.
- [7] E. Bertolazzi and G. Manzini. A second-order maximum principle preserving finite volume method for steady convection-diffusion problems. *SIAM, J. Numer. Anal.*, 43(5):2172–2199, 2006.
- [8] E. Bertolazzi and G. Manzini. On vertex reconstructions for cell-centered finite volume approximations of 2-D anisotropic diffusion problems. *Math. Models Methods Appl. Sci.*, 17(1):1–32, 2007.
- [9] P. Bochev and J. M. Hyman. Principles of mimetic discretizations. In D. Arnold, P. Bochev, R. Lehoucq, R. Nicolaides, and M. Shashkov, editors, *Compatible discretizations. Proceedings of IMA hot topics workshop on compatible discretizations*, IMA Volume 142. Springer-Verlag, 2006.
- [10] S. C. Brenner and L. R. Scott. *The Mathematical Theory of Finite Element Methods*, volume 15 of *Texts in Applied Mathematics*. Springer New York, 2008.

- [11] F. Brezzi, K. Lipnikov, and M. Shashkov. Convergence of the mimetic finite difference method for diffusion problems on polyhedral meshes. *SIAM J. Numer. Anal.*, 43(5):1872–1896, 2005.
- [12] A. Cangiani, G. Manzini, and A. Russo. Convergence analysis of a mimetic finite difference method for general second-order elliptic problems. Technical Report 15PV07/15/0, IMATI-CNR, 2007.
- [13] Y. Coudière. *Analyse de schémas volumes finis sur maillages non structurés pour des problèmes linéaires hyperboliques et elliptiques*. PhD thesis, Université "P. Sabatier" de Toulouse, Toulouse III, 1999.
- [14] Y. Coudière, J.-P. Vila, and P. Villedieu. Convergence rate of a finite volume scheme for a two-dimensional diffusion convection problem. *Math. Model. Numer. Anal.*, 33(3):493–516, 1999.
- [15] Y. Coudière and P. Villedieu. Convergence of a finite volume scheme for the linear convection-diffusion equation on locally refined meshes. *M2AN Math. Model. Numer. Anal.*, 34(6):1123–1149, 2000.
- [16] S. Delcourte, K. Domelevo, and P. Omnes. A discrete duality finite volume approach to Hodge decomposition and div-curl problems on almost arbitrary two-dimensional meshes. *SIAM J. Numer. Anal.*, 45(3):1142–1174, 2007.
- [17] K. Djadel and S. Nicaise. A non-conforming finite volume element method of weighted upstream type for the two-dimensional stationary Navier-Stokes system. *Appl. Numer. Math.*, 43(5):1872–1896, 2008.
- [18] K. Domelevo and P. Omnes. A finite volume method for the Laplace equation on almost arbitrary two-dimensional grids. *M2AN Math. Model. Numer. Anal.*, 39(6):1203–1249, 2005.
- [19] J. Droniou. Error estimates for the convergence of a finite volume discretization of convection-diffusion equations. *J. Numer. Math.*, 11(1):1–32, 2003.
- [20] J. Droniou and R. Eymard. A mixed finite volume scheme for anisotropic diffusion problems on any grid. *Numer. Math.*, 105(1):35–71, 2006.
- [21] R. Eymard, T. Gallouët, and R. Herbin. The finite volume method. In P. Ciarlet and J.L. Lions, editors, *Handbook for Numerical Analysis*, pages 715–1022. North Holland, 2000.
- [22] R. Eymard, T. Gallouët, and R. Herbin. A finite volume for anisotropic diffusion problems. *C. R. Math. Acad. Sci. Paris*, 339(4):299–302, 2004.
- [23] R. Eymard, T. Gallouët, and R. Herbin. A cell-centred finite volume approximation for anisotropic diffusion operators on unstructured meshes in any space dimension. *IMA J. Numer. Anal.*, 26(2):326–353, 2006.
- [24] R. Eymard, T. Gallouët, and R. Herbin. A new finite volume scheme for anisotropic diffusion problems on general grids: convergence analysis. *C. R. Math. Acad. Sci. Paris*, 6(344):403–406, 2007.
- [25] R. Eymard, M. Gutnic, and D. Hillhorst. The finite volume method for Richards equation. *Comput. Geosci.*, 3:259–294, 1999.
- [26] R. Herbin. An error estimate for a finite volume scheme for a diffusion-convection problem on a triangular mesh. *Numer. Methods Partial Differential Equations*, 11(2):165–173, 1995.
- [27] R. Herbin and F. Hubert. Benchmark on discretization schemes for anisotropic diffusion problems on general grids. In R. Eymard and J.-M. Hérard, editors, *Finite volumes for complex applications V*, Problems and Perspectives, pages 659–692, London, UK, 2008. Wiley.
- [28] F. Hermeline. A finite volume method for the approximation of diffusion operators on distorted meshes. *J. Comput. Phys.*, 160(2):481–499, 2000.
- [29] F. Hermeline. Approximation of diffusion operators with discontinuous tensor coefficients on distorted meshes. *Comput. Methods Appl. Mech. Engrg.*, 192(16):1939–1959, 2003.
- [30] C. Le Potier. Schéma volumes finis monotone pour des opérateurs de diffusion fortement anisotropes sur des maillages de triangles non structurés. *C. R. Math. Acad. Sci. Paris*, 341(12):787–792, 2005.
- [31] C. Le Potier. Schéma volumes finis pour des opérateurs de diffusion fortement anisotropes sur des maillages non structurés. *C. R. Math. Acad. Sci. Paris*, 340(12):921–926, 2005.
- [32] K. Lipnikov, M. Shashkov, D. Svyatskiy, and Yu. Vassilevski. Monotone finite volume schemes for diffusion equations on unstructured triangular and shape-regular polygonal meshes. *J. Comput. Phys.*, 227(1):492–512, 2007.
- [33] G. Manzini and S. Ferraris. Mass-conservative finite volume methods on 2-D unstructured grids for the Richards’ equation. *Adv. Water Res.*, 27(12):1199–1215, 2004.
- [34] G. Manzini and M. Putti. Mesh locking effects in the finite volume solution of 2-D anisotropic diffusion equations. *J. Comput. Phys.*, 220(2):751–771, 2007.
- [35] G. Manzini and A. Russo. A finite volume method for advection-diffusion problems in convection-dominated regimes. *Comput. Methods Appl. Mech. Engrg.*, 197(13-16):1242–1261, 2008.

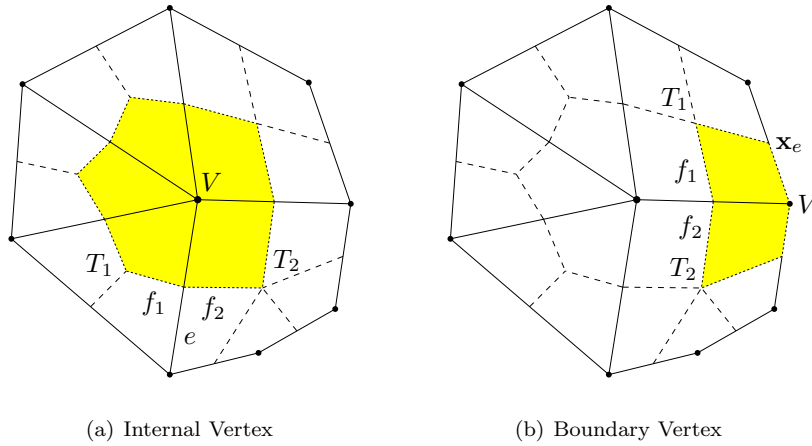


FIG. 2.1. Construction of the mesh \mathcal{V}_h (dashed and dotted lines) from mesh \mathcal{T}_h (solid lines). The plot on the left shows the case of an internal vertex, the plot on the right of a boundary vertex. The shaded region is the cell associated to the vertex V , its boundary ∂V being plotted by using dotted lines.

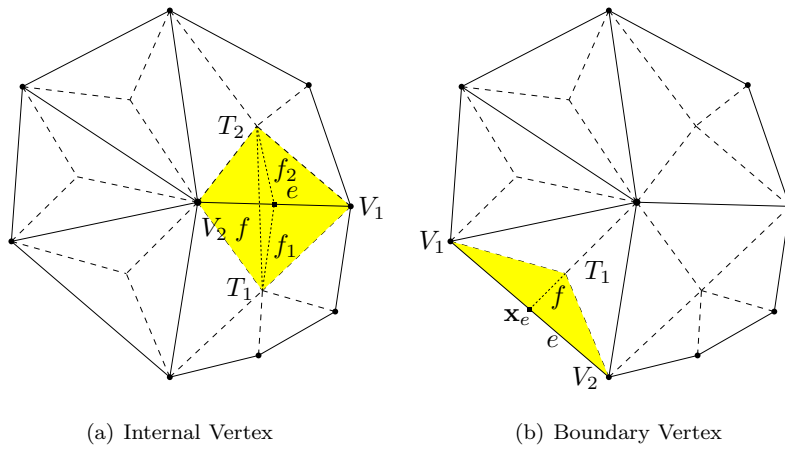
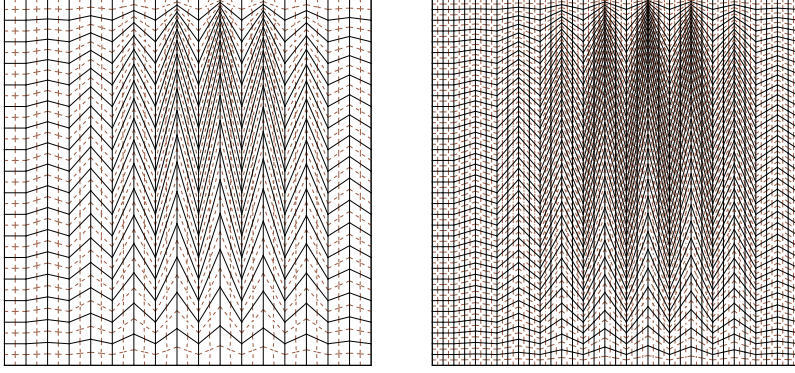
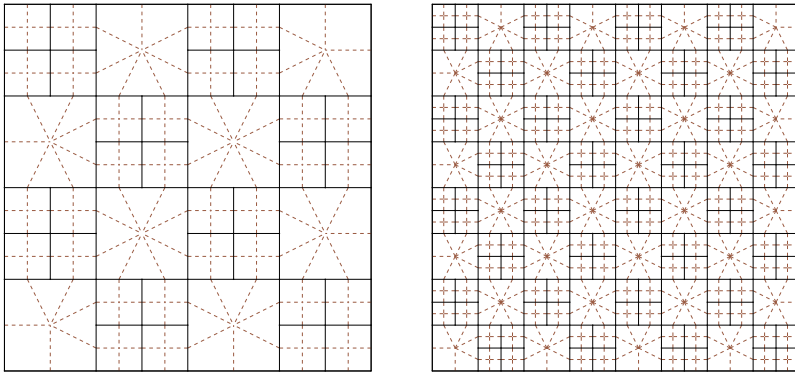


FIG. 2.2. Construction of the mesh \mathcal{D}_h (dashed lines) from mesh \mathcal{T}_h (solid lines). The plot on the left shows the case of an internal vertex, the plot on the right of a boundary vertex. The shaded region is the diamond cell D ; in both cases the diagonals f , f_1 , f_2 are plotted in dotted lines.



The first two meshes of mesh family \mathcal{M}_1



The first two meshes of mesh family \mathcal{M}_2

FIG. 4.1. The first two meshes of mesh family \mathcal{M}_1 (top) and \mathcal{M}_2 (bottom); each plot shows the primal mesh \mathcal{T}_h (solid line) and the dual mesh \mathcal{V}_h (dashed line).

TABLE 4.1

Mesh families \mathcal{M}_1 and \mathcal{M}_2 : $m_{\mathcal{T}_h}$ is the number of cells of \mathcal{T}_h , $m_{\mathcal{V}_h}$ is the number of cells of \mathcal{V}_h , $m_{\mathcal{D}_h}$ is the number of cells of \mathcal{D}_h , and h is the mesh size.

#	\mathcal{M}_1				\mathcal{M}_2			
	$m_{\mathcal{T}_h}$	$m_{\mathcal{V}_h}$	$m_{\mathcal{D}_h}$	h	$m_{\mathcal{T}_h}$	$m_{\mathcal{V}_h}$	$m_{\mathcal{D}_h}$	h
1	289	324	612	$2.711 \cdot 10^{-1}$	40	65	104	$2.500 \cdot 10^{-1}$
2	1156	1225	2380	$1.355 \cdot 10^{-1}$	160	241	400	$1.250 \cdot 10^{-1}$
3	2601	2704	5304	$9.036 \cdot 10^{-2}$	640	929	1568	$6.250 \cdot 10^{-2}$
4	4624	4761	9384	$6.777 \cdot 10^{-2}$	2560	3649	6208	$3.125 \cdot 10^{-2}$
5	7225	7396	14620	$5.422 \cdot 10^{-2}$	10240	14465	24704	$1.563 \cdot 10^{-2}$
6	10404	10609	21012	$4.518 \cdot 10^{-2}$	40960	57601	98560	$7.813 \cdot 10^{-3}$

TABLE 4.2

Relative errors and convergence rates for the scalar solution, $\mathcal{E}_{Q_h}(p_h)$, and the solution gradient, $\mathcal{E}_{\mathcal{D}_h}(\nabla_h \bar{p}_h)$; the central DDFV scheme for $\phi = 0.5$ and $k = 0$ and 1 is applied to the meshes of \mathcal{M}_1 .

#	$k = 0$				$k = 1$			
	$\mathcal{E}_{Q_h}(p_h)$	Rate	$\mathcal{E}_{\mathcal{D}_h}(\nabla_h \bar{p}_h)$	Rate	$\mathcal{E}_{Q_h}(p_h)$	Rate	$\mathcal{E}_{\mathcal{D}_h}(\nabla_h \bar{p}_h)$	Rate
1	$2.209 \cdot 10^{-2}$	--	$6.567 \cdot 10^{-2}$	--	$1.871 \cdot 10^{-2}$	--	$5.876 \cdot 10^{-2}$	--
2	$5.483 \cdot 10^{-3}$	2.010	$2.198 \cdot 10^{-2}$	1.579	$4.764 \cdot 10^{-3}$	1.973	$1.949 \cdot 10^{-2}$	1.592
3	$2.446 \cdot 10^{-3}$	1.990	$1.164 \cdot 10^{-2}$	1.567	$2.136 \cdot 10^{-3}$	1.978	$1.032 \cdot 10^{-2}$	1.568
4	$1.380 \cdot 10^{-3}$	1.991	$7.448 \cdot 10^{-3}$	1.553	$1.208 \cdot 10^{-3}$	1.981	$6.600 \cdot 10^{-3}$	1.552
5	$8.842 \cdot 10^{-4}$	1.993	$5.278 \cdot 10^{-3}$	1.543	$7.756 \cdot 10^{-4}$	1.985	$4.678 \cdot 10^{-3}$	1.542
6	$6.146 \cdot 10^{-4}$	1.995	$3.988 \cdot 10^{-3}$	1.536	$5.398 \cdot 10^{-4}$	1.987	$3.536 \cdot 10^{-3}$	1.535

TABLE 4.3

Relative errors and convergence rates for the scalar solution, $\mathcal{E}_{Q_h}(p_h)$, and the solution gradient, $\mathcal{E}_{\mathcal{D}_h}(\nabla_h \bar{p}_h)$; the fully upwind DDFV scheme for $\phi = 1$ and $k = 0$ and 1 is applied to the meshes of \mathcal{M}_1 .

#	$k = 0$				$k = 1$			
	$\mathcal{E}_{Q_h}(p_h)$	Rate	$\mathcal{E}_{\mathcal{D}_h}(\nabla_h \bar{p}_h)$	Rate	$\mathcal{E}_{Q_h}(p_h)$	Rate	$\mathcal{E}_{\mathcal{D}_h}(\nabla_h \bar{p}_h)$	Rate
1	$3.816 \cdot 10^{-2}$	--	$8.624 \cdot 10^{-2}$	--	$1.773 \cdot 10^{-2}$	--	$5.684 \cdot 10^{-2}$	--
2	$1.918 \cdot 10^{-2}$	0.992	$4.113 \cdot 10^{-2}$	1.067	$4.502 \cdot 10^{-3}$	1.977	$1.904 \cdot 10^{-2}$	1.577
3	$1.335 \cdot 10^{-2}$	0.894	$2.750 \cdot 10^{-2}$	0.993	$2.036 \cdot 10^{-3}$	1.957	$1.014 \cdot 10^{-2}$	1.554
4	$1.030 \cdot 10^{-2}$	0.901	$2.075 \cdot 10^{-2}$	0.978	$1.160 \cdot 10^{-3}$	1.955	$6.511 \cdot 10^{-3}$	1.539
5	$8.393 \cdot 10^{-3}$	0.915	$1.670 \cdot 10^{-2}$	0.975	$7.492 \cdot 10^{-4}$	1.959	$4.626 \cdot 10^{-3}$	1.531
6	$7.088 \cdot 10^{-3}$	0.926	$1.398 \cdot 10^{-2}$	0.975	$5.238 \cdot 10^{-4}$	1.963	$3.503 \cdot 10^{-3}$	1.525

TABLE 4.4

Relative errors and convergence rates for the scalar solution, $\mathcal{E}_{Q_h}(p_h)$, and the solution gradient, $\mathcal{E}_{\mathcal{D}_h}(\nabla_h \bar{p}_h)$; the central DDFV scheme for $\phi = 0.5$ and $k = 0$ and 1 is applied to the meshes of \mathcal{M}_2 .

#	$k = 0$				$k = 1$			
	$\mathcal{E}_{Q_h}(p_h)$	Rate	$\mathcal{E}_{\mathcal{D}_h}(\nabla_h \bar{p}_h)$	Rate	$\mathcal{E}_{Q_h}(p_h)$	Rate	$\mathcal{E}_{\mathcal{D}_h}(\nabla_h \bar{p}_h)$	Rate
1	$6.007 \cdot 10^{-2}$	--	$1.107 \cdot 10^{-1}$	--	$6.821 \cdot 10^{-2}$	--	$1.229 \cdot 10^{-1}$	--
2	$1.659 \cdot 10^{-2}$	1.856	$5.176 \cdot 10^{-2}$	1.096	$2.013 \cdot 10^{-2}$	1.760	$6.228 \cdot 10^{-2}$	0.980
3	$4.275 \cdot 10^{-3}$	1.956	$2.440 \cdot 10^{-2}$	1.085	$5.199 \cdot 10^{-3}$	1.952	$2.972 \cdot 10^{-2}$	1.067
4	$1.080 \cdot 10^{-3}$	1.984	$1.173 \cdot 10^{-2}$	1.056	$1.304 \cdot 10^{-3}$	1.995	$1.414 \cdot 10^{-2}$	1.071
5	$2.713 \cdot 10^{-4}$	1.993	$5.735 \cdot 10^{-3}$	1.032	$3.253 \cdot 10^{-4}$	2.002	$6.826 \cdot 10^{-3}$	1.050
6	$6.797 \cdot 10^{-5}$	1.997	$2.833 \cdot 10^{-3}$	1.017	$8.116 \cdot 10^{-5}$	2.002	$3.343 \cdot 10^{-3}$	1.029

TABLE 4.5

Relative errors and convergence rates for the scalar solution, $\mathcal{E}_{Q_h}(p_h)$, and the solution gradient, $\mathcal{E}_{\mathcal{D}_h}(\nabla_h \bar{p}_h)$; the fully upwind DDFV scheme for $\phi = 1$ and $k = 0$ and 1 is applied to the meshes of \mathcal{M}_2 .

#	$k = 0$				$k = 1$			
	$\mathcal{E}_{Q_h}(p_h)$	Rate	$\mathcal{E}_{\mathcal{D}_h}(\nabla_h \bar{p}_h)$	Rate	$\mathcal{E}_{Q_h}(p_h)$	Rate	$\mathcal{E}_{\mathcal{D}_h}(\nabla_h \bar{p}_h)$	Rate
1	$4.392 \cdot 10^{-2}$	--	$1.140 \cdot 10^{-1}$	--	$5.047 \cdot 10^{-2}$	--	$9.356 \cdot 10^{-2}$	--
2	$2.341 \cdot 10^{-2}$	0.907	$6.684 \cdot 10^{-2}$	0.769	$1.729 \cdot 10^{-2}$	1.545	$5.202 \cdot 10^{-2}$	0.846
3	$1.333 \cdot 10^{-2}$	0.812	$3.734 \cdot 10^{-2}$	0.840	$4.827 \cdot 10^{-3}$	1.840	$2.693 \cdot 10^{-2}$	0.949
4	$7.364 \cdot 10^{-3}$	0.856	$1.980 \cdot 10^{-2}$	0.914	$1.257 \cdot 10^{-3}$	1.941	$1.343 \cdot 10^{-2}$	1.004
5	$3.900 \cdot 10^{-3}$	0.917	$1.020 \cdot 10^{-2}$	0.957	$3.195 \cdot 10^{-4}$	1.976	$6.648 \cdot 10^{-3}$	1.014
6	$2.010 \cdot 10^{-3}$	0.956	$5.175 \cdot 10^{-3}$	0.978	$8.044 \cdot 10^{-5}$	1.989	$3.299 \cdot 10^{-3}$	1.011