A family of arbitrary-order mixed methods for heterogeneous anisotropic diffusion on general meshes

Daniele A. Di Pietro*[1] and Alexandre Ern†[2]

[1] Université de Montpellier 2, I3M, 34057 Montpellier CEDEX 5, France
[2]Université Paris-Est, CERMICS (ENPC), 6–8 avenue Blaise Pascal, 77455, Champs-sur-Marne, France

December 13, 2013

Abstract

In this work we propose a new family of arbitrary-order mixed methods for anisotropic heterogeneous diffusion on general polyhedral meshes. A key ingredient is the choice of flux degrees of freedom, which allows one to define a discrete divergence operator that satisfies the usual commuting diagram property. Based on this choice and on the discrete divergence operator, we define a flux reconstruction with suitable consistency and stability properties. The flux reconstruction and the discrete divergence operator are then used to define the discrete counterparts of the bilinear forms that appear in the continuous mixed formulation. A convergence analysis in the energy norm is carried out, and a supercloseness result for the L^2-norm of potential is proved. Several variations are considered, and the link with existing methods in the lowest-order case is discussed. Finally, the most relevant implementation issues are discussed, and some numerical tests are presented.

1 Introduction

Over the last few years, several discretization methods for diffusive problems have been proposed that support general meshes including polygonal or polyhedral elements and nonmatching interfaces. In most of the cases, such methods are obtained by preserving (or mimicking) to some extent the structure of the continuum operators at the discrete level. We mention the Mimetic Finite Difference (MFD) method of Kutznetsov, Lipnikov, and Shashkov [27] for which a convergence analysis has been carried out by Brezzi, Lipnikov, and Shashkov [12], see also [9] for the analysis of the nodal MFD method; the Mixed Finite Volume (MFV) method of Droniou and Eymard [22] and the Hybrid Finite Volume (HFV) method of Eymard, Gallouët, and Herbin [25]; the Discrete Geometric Approach of Codecasa, Specogna, and Trevisan [14] and the Compatible Discrete Operators (CDO) schemes recently introduced by Bonelle and Ern [6], both drawing on the seminal ideas of Bossavit [7, 8] and Tonti [28]. The similarities among the above-mentioned approaches and with other methods have been highlighted in

*daniele.di-pietro@univ-montp2.fr, corresponding author
†ern@cermics.enpc.fr

several papers. It has been shown by Droniou et al. [23] that generalized versions of the MFD, MFV, and HFV methods coincide at the algebraic level; a correspondence between CDO and nodal MFD methods has been established in [6]; the link between a generalized version of the HFV method and the classical Crouzeix–Raviart [15] element has been studied by Di Pietro and Lemaire [19] in the context of linear elasticity problems. A rather different point of view from the previous works is adopted in [16, 17], where the application of interior penalty strategies to consistent reconstructions of differential operators is considered for diffusive problems.

Until recently, the main focus has been on lowest-order methods. Very recent works consider, however, the extension to higher orders. An example is the arbitrary-order nodal MFD method of Beirão da Veiga, Lipnikov, and Manzini [5]. We also mention the Virtual Element (VE) method developed by Beirão da Veiga, Brezzi, and Marini for linear elasticity problems in [3]; cf. [4] for a presentation of the general ideas underlying the method applied to a model diffusion problem. The adjective "virtual" refers here to the fact that one defines a variational formulation in a finite element fashion, but without explicitly defining the underlying basis functions. While the present work was developed independently, analogies with VE methods are to be found in the general ideas underlying the reconstruction of differential operators. Key differences are that we develop a mixed approximation (hence face- rather than node-based) hinging on an explicit reconstruction of the diffusive flux. This can be interpreted as providing a full definition of basis functions. Finally, we mention the work of Brezzi, Buffa, and Manzini [10] on mimetic products of discrete differential forms, which also contains an extensive bibliographic section.

We focus here on the pure diffusion problem

$$\begin{aligned} -\boldsymbol{\nabla}\cdot(\mathbf{K}\boldsymbol{\nabla}u) &= f \qquad \text{in } \Omega \\ u &= 0 \qquad \text{on } \partial\Omega, \end{aligned} \tag{1}$$

where $\Omega \subset \mathbb{R}^d$, $d \geqslant 1$, denotes a bounded connected polygonal domain, $f \in L^2(\Omega)$ is a forcing term, and $\mathbf{K}$ is a piecewise constant, symmetric uniformly positive definite tensor-valued field whose eigenvalues are contained in the interval $[\lambda_\flat, \lambda_\sharp] \subset \mathbb{R}^+_*$. For $X \subset \Omega$, we denote by $(\cdot,\cdot)_X$ and $\|\cdot\|_X$ respectively the standard inner product and norm of $L^2(X)$, with the convention that the index is omitted if $X = \Omega$. Letting $\boldsymbol{\Sigma} := \boldsymbol{H}(\mathrm{div};\Omega)$ and $U := L^2(\Omega)$, the mixed variational formulation of problem (1) reads: Find $(\boldsymbol{s}, u) \in \boldsymbol{\Sigma} \times U$ such that

$$\begin{aligned} (\mathbf{K}^{-1}\boldsymbol{s}, \boldsymbol{t}) + (u, \boldsymbol{\nabla}\cdot\boldsymbol{t}) &= 0 \qquad && \forall \boldsymbol{t} \in \boldsymbol{\Sigma}, \\ -(\boldsymbol{\nabla}\cdot\boldsymbol{s}, v) &= (f, v) \qquad && \forall v \in U. \end{aligned} \tag{2}$$

Throughout this work, $\boldsymbol{s}$ and u will be termed flux and potential, respectively. It has been long known that mixed methods based on the weak formulation (2) perform well for problems such as (1) where the diffusion coefficient is possibly anisotropic and can exhibit large jumps across interfaces; cf., e.g., the discussion in [29, Section 5.4], where an interesting interpretation of this behaviour is proposed.

The key ideas of our method can be summarized as follows. Let $k \geqslant 0$ be a fixed polynomial degree and denote by U_h^k the space of broken polynomials of total degree $\leqslant k$ used to approximate the potential. The starting point is to define the vector space $\boldsymbol{\Sigma}_h^k$ of flux degrees of freedom (DOFs) so that an interpolator I_h^k on $\boldsymbol{\Sigma}_h^k$ and a discrete divergence operator $D_h^k : \boldsymbol{\Sigma}_h^k \to U_h^k$ can be defined that satisfy the usual commuting diagram property [20, 21].

The flux DOFs are polynomials of total degree $\leqslant k$ attached to faces and fluxes of polynomials of total degree $\leqslant k$ attached to cells (that is, gradients of such polynomials times the diffusion tensor). The commuting property of the discrete divergence operator and the choice of the flux degrees of freedom are used to define a (possibly nonconforming) flux reconstruction $\mathfrak{R}_h^k : \mathbf{\Sigma}_h^k \to L^2(\Omega)^d$ that (i) is exact when the potential is a polynomial of degree $(k+1)$ inside each element (*consistency*), and (ii) has coercivity properties on the kernel of D_h^k (*stability*), as well as *continuity* properties. The flux reconstruction is obtained by reasoning element-wise, and is composed of two contributions responsible for consistency and stability, respectively. It turns out that an important requirement in the analysis is that the two contributions are mutually $\mathbf{K}^{-1}$-orthogonal. We emphasize that the consistency property is designed so as to guarantee convergence also in the lowest-order case $k = 0$. We also observe, in passing, that the cell unknowns can be eliminated locally thereby yielding a global problem in the face unknowns only. Moreover, the use of fluxes of polynomial potentials (instead of fully vector-valued polynomials for the fluxes) results in a substantial reduction in the size of the local reconstruction problems with respect, e.g., to the Hybrid Mixed Discontinuous Galerkin method of Cockburn, Gopalakrishnan, and Lazarov [13]. Based on the flux reconstruction, we next define a bilinear form H on $\mathbf{\Sigma}_h^k \times \mathbf{\Sigma}_h^k$ which is the discrete counterpart of the inner product $(\mathbf{K}^{-1}\cdot, \cdot)$. Using the terminology of [6], the bilinear form H corresponds to the discrete Hodge operator. Finally, a discretization of (2) is obtained based on the bilinear form H and the discrete divergence operator D_h^k. Under the usual regularity assumptions for the exact solution, we prove that the error on the flux measured by the norm defined by H converges as h^{k+1} (h denotes here the meshsize). Additionally, a supercloseness result can be proved for the potential as in standard mixed finite elements (cf., e.g., [20, 21, 26]). Provided elliptic regularity holds, this means that the L^2-norm of the difference between the discrete potential and the L^2-orthogonal projection of u on U_h^k converges as h^{k+2}.

The material is organized as follows. In Section 2, after briefly recalling the notion of admissible mesh sequences, we define flux degrees of freedom, introduce the discrete divergence operator, and define the flux reconstruction upon which the method hinges. The discrete problem is stated at the end of this section and its well-posedness is established. In Section 3 we carry out the convergence analysis. Section 4 contains some variations of the method, in particular sufficient conditions to define virtual versions which have similar convergence properties as the original method, but for which the flux reconstruction is left undefined. In this section we also show that, in the lowest-order case $k = 0$, our method has fundamental similarities with both the HFV and the GDA methods (which are shown to coincide up to a different choice of the stabilization parameter). Finally, Section 5 addresses the most relevant implementation issues and contains some numerical tests.

2 Discretization

We introduce here the notion of admissible mesh sequences, recall some basic facts on broken functional spaces, and define the discrete divergence operator and the flux reconstruction upon which the discretization of (2) stated at the end of the section hinges.

2.1 Setting

2.1.1 Admissible mesh sequences

Closely following [18, Chapter 1] and [19, Section 2], we introduce the notion of admissible mesh sequences possibly including general polygonal or polyhedral elements and nonconforming interfaces. Let $\mathcal{H} \subset \mathbb{R}_*^+$ denote a countable set having 0 as its unique accumulation point. We consider mesh sequences $\mathcal{T}_\mathcal{H} := \{\mathcal{T}_h\}_{h\in\mathcal{H}}$ where, for all $h \in \mathcal{H}$, $\mathcal{T}_h$ denotes a finite collection of nonempty disjoint open polyhedra (called elements or cells) $\mathcal{T}_h = \{T\}$ such that $\overline{\Omega} = \bigcup_{T\in\mathcal{T}_h} \overline{T}$ and $h = \max_{T\in\mathcal{T}_h} h_T$ (h_T denotes here the diameter of T). We say that a hyperplanar closed connected subset F of $\overline{\Omega}$ is a mesh face if it has positive $(d-1)$-dimensional measure and if either there exist $T_1, T_2 \in \mathcal{T}_h$ such that $F \subset \partial T_1 \cap \partial T_2$ (and F is called an interface) or there exists $T \in \mathcal{T}_h$ such that $F \subset \partial T \cap \partial\Omega$ (and F is called a boundary face). Interfaces are collected in the set $\mathcal{F}_h^{\mathrm{i}}$, boundary faces in $\mathcal{F}_h^{\mathrm{b}}$ and we let $\mathcal{F}_h := \mathcal{F}_h^{\mathrm{i}} \cup \mathcal{F}_h^{\mathrm{b}}$. The diameter of a face $F \in \mathcal{F}_h$ is denoted by h_F. Moreover, we set, for all $T \in \mathcal{T}_h$, $\mathcal{F}_T := \{F \in \mathcal{F}_h \mid F \subset \partial T\}$, and, for all $F \in \mathcal{F}_T$, we denote by $\boldsymbol{n}_{TF}$ the normal to F pointing out of T. For every interface $F \subset \partial T_1 \cap \partial T_2$ we fix once and for all an orientation by means of a unit normal vector $\boldsymbol{n}_F$ and number the elements T_1 and T_2 so that $\boldsymbol{n}_F := \boldsymbol{n}_{T_1F}$. We also define $\epsilon_{TF} := \boldsymbol{n}_{TF}\cdot\boldsymbol{n}_F$ for all $T \in \mathcal{T}_h$ and all $F \in \mathcal{F}_T$.

It is assumed that, for all $h \in \mathcal{H}$, $\mathcal{T}_h$ admits a matching simplicial submesh $\mathfrak{T}_h$ and that the following holds:

(M1) *Shape-regularity.* There exists a real number $\varrho_1 > 0$ independent of h such that, for all $h \in \mathcal{H}$ and all simplex $S \in \mathfrak{T}_h$ of diameter h_S and inradius r_S, $\varrho_1 h_S \leqslant r_S$ holds.

(M2) *Contact-regularity.* There exists a real number $\varrho_2 > 0$ independent of h such that, for all $h \in \mathcal{H}$, all $T \in \mathcal{T}_h$, and all $S \in \mathfrak{T}_T := \{S \in \mathfrak{T}_h \mid S \subset T\}$, $\varrho_2 h_T \leqslant h_S$ holds.

(M3) *Star-shaped property.* For all $h \in \mathcal{H}$ and all $T \in \mathcal{T}_h$, there exists a point $\boldsymbol{x}_T \in T$ such that T is star-shaped with respect to $\boldsymbol{x}_T$ and, for all $F \in \mathcal{F}_T$, $d_{TF} \geqslant \varrho_3 h_T$ holds with d_{TF} orthogonal distance between $\boldsymbol{x}_T$ and F and $\varrho_3 > 0$ independent of h.

The star-shaped property is used here to define the flux reconstruction in Section 2.4.2. Additionally, owing to [18, Lemma 1.41], there exists $(d+1) \leqslant \overline{N_\partial} < +\infty$ such that

$$\forall h \in \mathcal{H}, \qquad \max_{T\in\mathcal{T}_h} \mathrm{card}(\mathcal{F}_T) \leqslant \overline{N_\partial}. \tag{3}$$

In what follows, we often abbreviate as $a \lesssim b$ the inequality $a \leqslant Cb$ with $C > 0$ independent of h and $\mathbf{K}$ (the dependence on $\mathbf{K}$ is explicitly tracked to pinpoint the effect of the local anisotropy ratio on the error estimates).

2.1.2 Basic results on broken functional spaces

In this section we recall some basic results for broken functional spaces that can be proved under assumptions **(M1)**–**(M2)**. For $h \in \mathcal{H}$ and an integer $k \geqslant 0$ we define the broken polynomial space

$$\mathbb{P}_d^k(\mathcal{T}_h) := \big\{v \in L^2(\Omega) \mid v_{|T} \in \mathbb{P}_d^k(T) \quad \forall T \in \mathcal{T}_h\big\},$$

where $\mathbb{P}_d^k(T)$ denotes the restriction to T of d-variate polynomials of total degree $\leqslant k$. We express local regularity in terms of the broken Sobolev spaces

$$H^l(\mathcal{T}_h) := \big\{v \in L^2(\Omega) \mid v_{|T} \in H^l(T) \quad \forall T \in \mathcal{T}_h\big\}.$$

The following trace inequalities hold for all $T \in \mathcal{T}_h$ and all $F \in \mathcal{F}_T$, cf. [18, Lemmata 1.46 and 1.49]:

$$\|v\|_F \leqslant C_{\mathrm{tr}} h_F^{-1/2} \|v\|_T \qquad \forall v \in \mathbb{P}_d^k(T), \tag{4}$$

$$\|v\|_F \leqslant C_{\mathrm{tr,c}} \Big(h_T^{-1} \|v\|_T^2 + h_T |v|_{H^1(T)}^2 \Big)^{1/2} \qquad \forall v \in H^1(T), \tag{5}$$

with C_{tr} and $C_{\mathrm{tr,c}}$ depending on ϱ_1 and ϱ_2 but independent of h. The following inverse inequality holds for all $T \in \mathcal{T}_h$ with C_{inv} again depending on ϱ_1, ϱ_2 but independent of h, cf. [18, Lemma 1.44],

$$\|\boldsymbol{\nabla} v\|_{L^2(T)^d} \leqslant C_{\mathrm{inv}} h_T^{-1} \|v\|_{L^2(T)} \qquad \forall v \in \mathbb{P}_d^k(T). \tag{6}$$

It follows from the mesh regularity assumptions together with [18, Lemma 1.40] and the results of [24] that the L^2-orthogonal projector π_h^k on $\mathbb{P}_d^k(\mathcal{T}_h)$ has optimal approximation properties: For all $T \in \mathcal{T}_h$, all $s \in \{0, \ldots, k+1\}$, and all $v \in H^s(T)$,

$$|v - \pi_h^k v|_{H^m(T)} \leqslant C_{\mathrm{app}} h_T^{s-m} |v|_{H^s(T)} \qquad \forall m \in \{0, \ldots, s\},$$

holds with C_{app} depending on ϱ_1 and ϱ_2 but independent of h. Finally, we recall the following Poincaré inequality valid for all $T \in \mathcal{T}_h$:

$$\|v - \overline{v}\|_T \leqslant C_{\mathrm{P}} h_T \|\boldsymbol{\nabla} v\|_T, \qquad \forall v \in H^1(T), \tag{7}$$

where $\overline{v} := (v,1)_T / |T|_d$ and C_{P} is independent of h ($C_{\mathrm{P}} = \pi^{-1}$ for convex elements [2]).

2.2 Flux degrees of freedom

We define in this section the degrees of freedom (DOFs) for the flux approximation. It is assumed from this point on that, for all $h \in \mathcal{H}$, $\mathcal{T}_h$ is $\mathbf{K}$-compliant, i.e., jumps in the diffusion coefficient do not occur inside elements. As a consequence,

$$\forall h \in \mathcal{H}, \qquad \mathbf{K}_T := \mathbf{K}_{|T} \in \mathbb{P}_d^0(T)^{d\times d} \qquad \forall T \in \mathcal{T}_h.$$

Let, for a fixed integer $k \geqslant 0$,

$$\mathbb{T}_T^k := \mathbf{K}_T \boldsymbol{\nabla} \mathbb{P}_d^{k,0}(T) \quad \forall T \in \mathcal{T}_h, \qquad \mathbb{F}_F^k := \mathbb{P}_{d-1}^k(F) \quad \forall F \in \mathcal{F}_h,$$

denote the flux DOFs attached to cells and faces, respectively, where, for $l \geqslant 0$, $\mathbb{P}_d^{l,0}(T)$ is spanned by scalar-valued polynomial functions of total degree $\leqslant l$ having zero average on T (concerning the zero-average condition, see Remark 4). For all $T \in \mathcal{T}_h$, we define the local space of DOFs for the flux approximation as

$$\boldsymbol{\Sigma}_T^k := \mathbb{T}_T^k \times \mathbb{F}_T^k, \qquad \mathbb{F}_T^k := \bigtimes_{F\in\mathcal{F}_T} \mathbb{F}_F^k.$$

In the lowest-order case $k = 0$ it is understood that the cell DOFs are unnecessary (the space $\mathbb{T}_T^0$ contains only the null function over T). The global space of DOFs for the flux is obtained from the local spaces $\{\boldsymbol{\Sigma}_T^k\}_{T\in\mathcal{T}_h}$ by patching interface values,

$$\boldsymbol{\Sigma}_h^k := \mathbb{T}_h^k \times \mathbb{F}_h^k, \qquad \mathbb{T}_h^k := \bigtimes_{T\in\mathcal{T}_h} \mathbb{T}_T^k, \qquad \mathbb{F}_h^k := \bigtimes_{F\in\mathcal{F}_h} \mathbb{F}_F^k. \tag{8}$$

To localize DOFs in $\boldsymbol{\Sigma}_h^k$ to $\boldsymbol{\Sigma}_T^k$ for a given element $T \in \mathcal{T}_h$, we introduce the restriction operator $L_T : \boldsymbol{\Sigma}_h^k \to \boldsymbol{\Sigma}_T^k$ which, for all $\boldsymbol{\tau}_h := (\{\boldsymbol{\tau}_T\}_{T\in\mathcal{T}_h}, \{\tau_F\}_{F\in\mathcal{F}_h}) \in \boldsymbol{\Sigma}_h^k$ is such that $L_T\boldsymbol{\tau}_h = (\boldsymbol{\tau}_T, \{\tau_F\}_{F\in\mathcal{F}_T}) \in \boldsymbol{\Sigma}_T^k$.

2.3 Discrete divergence operator

We next introduce a discrete divergence operator which is instrumental in the formulation of the method and use it to define the discrete counterpart of the $\boldsymbol{H}(\mathrm{div};\Omega)$-norm. The local discrete divergence operator

$$D_T^k : \boldsymbol{\Sigma}_T^k \to U_T^k := \mathbb{P}_d^k(T) \tag{9}$$

is such that, for all $\boldsymbol{\tau} = (\boldsymbol{\tau}_T, \{\tau_F\}_{F\in\mathcal{F}_T}) \in \boldsymbol{\Sigma}_T^k$ and all $v \in U_T^k$,

$$(D_T^k\boldsymbol{\tau}, v)_T = -(\boldsymbol{\nabla} v, \boldsymbol{\tau}_T)_T + \sum_{F\in\mathcal{F}_T} (v, \tau_F\epsilon_{TF})_F. \tag{10}$$

We equip the space $\boldsymbol{\Sigma}_T^k$ with the following $\boldsymbol{H}(\mathrm{div};\Omega)$-like norm:

$$\forall \boldsymbol{\tau} \in \boldsymbol{\Sigma}_T^k, \qquad |||\boldsymbol{\tau}|||_T^2 := \|\boldsymbol{\tau}_T\|_T^2 + h_T^2\|D_T^k\boldsymbol{\tau}\|_T^2 + \sum_{F\in\mathcal{F}_T} h_F\|\tau_F\|_F^2. \tag{11}$$

The global discrete divergence operator

$$D_h^k : \boldsymbol{\Sigma}_h^k \to U_h^k := \mathbb{P}_d^k(\mathcal{T}_h) \tag{12}$$

is such that, for all $\boldsymbol{\tau}_h \in \boldsymbol{\Sigma}_h^k$ and all $v_h \in U_h^k$,

$$(D_h^k\boldsymbol{\tau}_h, v_h) = \sum_{T\in\mathcal{T}_h} (D_T^k(L_T\boldsymbol{\tau}_h), v_h)_T. \tag{13}$$

We equip the space $\boldsymbol{\Sigma}_h^k$ with the following norm:

$$\forall \boldsymbol{\tau}_h \in \boldsymbol{\Sigma}_h^k, \qquad |||\boldsymbol{\tau}_h|||^2 := \sum_{T\in\mathcal{T}_h} |||L_T\boldsymbol{\tau}_h|||_T^2. \tag{14}$$

Let us study the properties of the local and global discrete divergence operators defined by (10) and (13), respectively. We let, for all $T \in \mathcal{T}_h$, $\boldsymbol{\Sigma}^+(T) := \{\boldsymbol{t} \in L^s(T) \mid \boldsymbol{\nabla}\cdot\boldsymbol{t} \in L^2(T)\}$ for a fixed $s > 2$. Classically, the moments of the normal components of functions in $\boldsymbol{\Sigma}^+(T)$ are meaningful on the faces of T; cf., e.g., [11, Section III.3.3]. We let $I_T^k : \boldsymbol{\Sigma}^+(T) \to \boldsymbol{\Sigma}_T^k$ denote the local interpolator such that, for all $\boldsymbol{t} \in \boldsymbol{\Sigma}^+(T)$, $I_T^k\boldsymbol{t} = (\boldsymbol{\tau}_T, \{\tau_F\}_{F\in\mathcal{F}_T})$ with

$$\boldsymbol{\tau}_T = \varpi_T^k\boldsymbol{t}, \qquad \tau_F = \pi_F^k(\boldsymbol{t}\cdot\boldsymbol{n}_F) \quad \forall F \in \mathcal{F}_T. \tag{15}$$

where π_F^k is the standard L^2-orthogonal projector on $\mathbb{F}_F^k$, while ϖ_T^k denotes the $(\mathbf{K}_T^{-1}\cdot,\cdot)_T$-orthogonal projector on $\mathbb{T}_T^k$ (in fact, an elliptic projector in terms of the potential) such that

$$(\mathbf{K}_T^{-1}\varpi_T^k\boldsymbol{t}, \boldsymbol{w})_T = (\mathbf{K}_T^{-1}\boldsymbol{t}, \boldsymbol{w})_T \qquad \forall \boldsymbol{w} \in \mathbb{T}_T^k. \tag{16}$$

Note that this projection is well-defined owing to the zero-average condition on $\mathbb{P}_d^{k,0}(T)$. Correspondingly, the global interpolator $I_h^k : \boldsymbol{\Sigma}^+ \to \boldsymbol{\Sigma}_h^k$ with $\boldsymbol{\Sigma}^+ := \{\boldsymbol{t} \in \boldsymbol{\Sigma} \mid \boldsymbol{t}_{|T} \in \boldsymbol{\Sigma}^+(T),\ \forall T \in \mathcal{T}_h\}$ is such that, for all $\boldsymbol{t} \in \boldsymbol{\Sigma}^+$, $I_h^k\boldsymbol{t} = (\{\boldsymbol{\tau}_T\}_{T\in\mathcal{T}_h}, \{\tau_F\}_{F\in\mathcal{F}})$ with

$$\boldsymbol{\tau}_T = \varpi_T^k\boldsymbol{t} \quad \forall T \in \mathcal{T}_h, \qquad \tau_F = \pi_F^k(\boldsymbol{t}\cdot\boldsymbol{n}_F) \quad \forall F \in \mathcal{F}_h. \tag{17}$$

Proposition 1 (Commuting property for discrete divergence operator)**.** *Denoting by π_T^k and π_h^k the L^2-orthogonal projectors on $\mathbb{P}_d^k(T)$ and $\mathbb{P}_d^k(\mathcal{T}_h)$, respectively, the following commuting diagrams hold:*

$$\begin{array}{ccc} \boldsymbol{\Sigma}^+(T) & \xrightarrow{\boldsymbol{\nabla}\cdot} & L^2(T) \\ \downarrow{\scriptstyle I_T^k} & & \downarrow{\scriptstyle \pi_T^k} \\ \boldsymbol{\Sigma}_T^k & \xrightarrow{D_T^k} & U_T^k \end{array} \qquad \begin{array}{ccc} \boldsymbol{\Sigma}^+ & \xrightarrow{\boldsymbol{\nabla}\cdot} & U \\ \downarrow{\scriptstyle I_h^k} & & \downarrow{\scriptstyle \pi_h^k} \\ \boldsymbol{\Sigma}_h^k & \xrightarrow{D_h^k} & U_h^k \end{array}$$

Proof. Consider the local commuting diagram. Let $T \in \mathcal{T}_h$, $\boldsymbol{t} \in \boldsymbol{\Sigma}^+(T)$, and set $\boldsymbol{\tau} := I_T^k \boldsymbol{t}$. We infer that, for all $v \in U_T^k$,

$$\begin{aligned}(\pi_T^k(\boldsymbol{\nabla}\cdot\boldsymbol{t}), v)_T = (\boldsymbol{\nabla}\cdot\boldsymbol{t}, v)_T &= -(\boldsymbol{\nabla} v, \boldsymbol{t})_T + \sum_{F\in\mathcal{F}_T} (v, (\boldsymbol{t}\cdot\boldsymbol{n}_F)\epsilon_{TF})_F \\ &= -(\boldsymbol{\nabla} v, \varpi_T^k \boldsymbol{t})_T + \sum_{F\in\mathcal{F}_T} (v, \pi_F^k(\boldsymbol{t}\cdot\boldsymbol{n}_F)\epsilon_{TF})_F \\ &= -(\boldsymbol{\nabla} v, \boldsymbol{\tau}_T)_T + \sum_{F\in\mathcal{F}_T} (v, \tau_F \epsilon_{TF})_F = (D_T^k \boldsymbol{\tau}, v)_T,\end{aligned}$$

where we have used that $\mathbf{K}_T \boldsymbol{\nabla} v \in \mathbb{T}_T^k$, $v_{|F} \in \mathbb{F}_F^k$, (15), and (10). For the global commuting diagram, we use that, for all $T \in \mathcal{T}_h$, $(D_h^k \boldsymbol{\tau}_h)_{|T} = D_T^k(L_T \boldsymbol{\tau}_h)$ for all $\boldsymbol{\tau}_h \in \boldsymbol{\Sigma}_h^k$ owing to (13), $L_T(I_h^k \boldsymbol{t}) = I_T^k(\boldsymbol{t}_{|T})$ for all $\boldsymbol{t} \in \boldsymbol{\Sigma}^+$ owing to (17), whence

$$(D_h^k(I_h^k \boldsymbol{t}))_{|T} = D_T^k(I_T^k(\boldsymbol{t}_{|T})) = \pi_T^k(\boldsymbol{\nabla}\cdot(\boldsymbol{t}_{|T})) = (\pi_h^k(\boldsymbol{\nabla}\cdot\boldsymbol{t}))_{|T},$$

owing to the local commuting diagram. □

An immediate consequence of Proposition 1 that I_h^k can play the role of a Fortin operator [11].

2.4 Flux reconstruction

We now introduce an arbitrary-order flux reconstruction $\mathfrak{R}_h^k$ inspired by the lowest-order Hybrid Finite Volume method of [23, Section 2.3]. The operator $\mathfrak{R}_h^k : \boldsymbol{\Sigma}_h^k \to L^2(\Omega)^d$ is obtained elementwise from local reconstructions $\mathfrak{R}_T^k : \boldsymbol{\Sigma}_T^k \to L^2(T)^d$ by setting, for all $T \in \mathcal{T}_h$,

$$\mathfrak{R}_h^k(\boldsymbol{\tau}_h)_{|T} := \mathfrak{R}_T^k(L_T \boldsymbol{\tau}_h) \qquad \forall \boldsymbol{\tau}_h \in \boldsymbol{\Sigma}_h^k. \tag{18}$$

Let H_T denote the local bilinear form on $\boldsymbol{\Sigma}_T^k \times \boldsymbol{\Sigma}_T^k$ such that, for all $\boldsymbol{\sigma}, \boldsymbol{\tau} \in \boldsymbol{\Sigma}_T^k$,

$$H_T(\boldsymbol{\sigma}, \boldsymbol{\tau}) := (\mathbf{K}_T^{-1} \mathfrak{R}_T^k \boldsymbol{\sigma}, \mathfrak{R}_T^k \boldsymbol{\tau})_T. \tag{19}$$

By construction, H_T is nonnegative. The global bilinear form H on $\boldsymbol{\Sigma}_h^k \times \boldsymbol{\Sigma}_h^k$ is such that, for all $\boldsymbol{\sigma}_h, \boldsymbol{\tau}_h \in \boldsymbol{\Sigma}_h^k$,

$$H(\boldsymbol{\sigma}_h, \boldsymbol{\tau}_h) := \sum_{T\in\mathcal{T}_h} H_T(L_T \boldsymbol{\sigma}_h, L_T \boldsymbol{\tau}_h) = (\mathbf{K}^{-1} \mathfrak{R}_h^k \boldsymbol{\sigma}_h, \mathfrak{R}_h^k \boldsymbol{\tau}_h), \tag{20}$$

where the last equality is a consequence of (18). We denote by $\|\cdot\|_H$ and $\|\cdot\|_{H,T}$ the seminorms defined by H on $\boldsymbol{\Sigma}_h^k$ and by H_T on $\boldsymbol{\Sigma}_T^k$, respectively. The use of a double bar notation is motivated by the fact that $\|\cdot\|_H$ (resp. $\|\cdot\|_{H,T}$) is a norm on $\ker(D_h^k)$ (resp. $\ker(D_T^k)$) as a result of requirement **(R2)** below.

We aim at satisfying the following properties for the local reconstructions:

(R1) *Consistency and orthogonality.* For all $T \in \mathcal{T}_h$, the space $\boldsymbol{\mathcal{R}}_T^k := \mathfrak{R}_T^k(\boldsymbol{\Sigma}_T^k)$ contains the space $\boldsymbol{\Gamma}_T^k := \mathbf{K}_T\boldsymbol{\nabla}\mathbb{P}_d^{k+1,0}(T)$, and, the local reconstruction operator can be decomposed as $\mathfrak{R}_T^k = \mathfrak{C}_T^k + \mathfrak{J}_T^k$ with $\mathfrak{C}_T^k : \boldsymbol{\Sigma}_T^k \to \boldsymbol{\Gamma}_T^k$ and $\mathfrak{J}_T^k : \boldsymbol{\Sigma}_T^k \to \boldsymbol{\mathcal{R}}_T^k$ such that

$$\mathfrak{C}_T^k(I_T^k\boldsymbol{w}) = \boldsymbol{w} \qquad \forall \boldsymbol{w} \in \boldsymbol{\Gamma}_T^k, \tag{21}$$

$$\mathfrak{J}_T^k(I_T^k\boldsymbol{w}) = 0, \qquad \forall \boldsymbol{w} \in \boldsymbol{\Gamma}_T^k, \tag{22}$$

$$(\mathbf{K}_T^{-1}\mathfrak{J}_T^k\boldsymbol{\tau}, \boldsymbol{w})_T = 0 \qquad \forall (\boldsymbol{\tau}, \boldsymbol{w}) \in \boldsymbol{\Sigma}_T^k \times \boldsymbol{\Gamma}_T^k. \tag{23}$$

Owing to properties (21) and (22), which imply that

$$\mathfrak{R}_T^k(I_T^k\boldsymbol{w}) = \boldsymbol{w} \qquad \forall \boldsymbol{w} \in \boldsymbol{\Gamma}_T^k. \tag{24}$$

$\mathfrak{C}_T^k$ is termed the consistent part of the reconstruction and $\mathfrak{J}_T^k$ is termed the residual. The residual is introduced to satisfy the stability property (25a) below. The orthogonality property (23) is used in Section 2.4.2 to prove (25a) and, via (29), in the proof of Theorem 11 (bound on $\mathfrak{T}_3$).

(R2) *Stability and continuity.* There is $\eta > 0$ independent of h and $\mathbf{K}$ such that, for all $T \in \mathcal{T}_h$,

$$\|\boldsymbol{\tau}\|_{H,T}^2 \geqslant \lambda_{\sharp,T}^{-1}\eta\|\boldsymbol{\tau}\|_T^2 \qquad \forall \boldsymbol{\tau} \in \ker(D_T^k), \tag{25a}$$

$$\|\boldsymbol{\tau}\|_{H,T}^2 \leqslant \eta^{-1}\lambda_{\flat,T}^{-1}\|\boldsymbol{\tau}\|_T^2 \qquad \forall \boldsymbol{\tau} \in \boldsymbol{\Sigma}_T^k, \tag{25b}$$

where $\lambda_{\sharp,T}$ and $\lambda_{\flat,T}$ denote the largest and smallest eigenvalue of $\mathbf{K}_T$, respectively. Inequality (25a) implies, in particular, that $\mathfrak{R}_T^k$ (resp. $\mathfrak{R}_h^k$) is injective on $\ker(D_T^k)$ (resp. $\ker(D_h^k)$). Accounting for (14) and (20), and summing over $T \in \mathcal{T}_h$, it is inferred from (25) that

$$\|\boldsymbol{\tau}_h\|_H^2 \geqslant \lambda_\sharp^{-1}\eta\|\boldsymbol{\tau}_h\|^2 \qquad \forall \boldsymbol{\tau}_h \in \ker(D_h^k), \tag{26a}$$

$$\|\boldsymbol{\tau}_h\|_H^2 \leqslant \eta^{-1}\lambda_\flat^{-1}\|\boldsymbol{\tau}_h\|^2 \qquad \forall \boldsymbol{\tau}_h \in \boldsymbol{\Sigma}_h^k, \tag{26b}$$

where we have used the fact that $\boldsymbol{\tau}_h \in \ker(D_h^k)$ implies $L_T\boldsymbol{\tau}_h \in \ker(D_T^k)$ for all $T \in \mathcal{T}_h$.

Proposition 2 (Consequences of **(R1)**)**.** *The following holds for all $T \in \mathcal{T}_h$:*

$$\forall \boldsymbol{\sigma}, \boldsymbol{\tau} \in \boldsymbol{\Sigma}_T^k, \qquad H_T(\boldsymbol{\sigma}, \boldsymbol{\tau}) = (\mathbf{K}_T^{-1}\mathfrak{C}_T^k\boldsymbol{\sigma}, \mathfrak{C}_T^k\boldsymbol{\tau})_T + (\mathbf{K}_T^{-1}\mathfrak{J}_T^k\boldsymbol{\sigma}, \mathfrak{J}_T^k\boldsymbol{\tau})_T, \tag{27}$$

hence,

$$\forall \boldsymbol{\tau} \in \boldsymbol{\Sigma}_T^k, \qquad \|\boldsymbol{\tau}\|_{H,T}^2 = \|\mathbf{K}_T^{-1/2}\mathfrak{R}_T^k\boldsymbol{\tau}\|_T^2 = \|\mathbf{K}_T^{-1/2}\mathfrak{C}_T^k\boldsymbol{\tau}\|_T^2 + \|\mathbf{K}_T^{-1/2}\mathfrak{J}_T^k\boldsymbol{\tau}\|_T^2. \tag{28}$$

Additionally, for all $\boldsymbol{w} \in \boldsymbol{\Gamma}_T^k$, letting $\boldsymbol{v} := I_T^k\boldsymbol{w}$ with I_T^k defined by (15)*, the following holds:*

$$H_T(\boldsymbol{v}, \boldsymbol{\tau}) = (\mathbf{K}_T^{-1}\boldsymbol{w}, \mathfrak{C}_T^k\boldsymbol{\tau})_T \qquad \forall \boldsymbol{\tau} \in \boldsymbol{\Sigma}_T^k. \tag{29}$$

2.4.1 Consistency

We first examine the consistency requirement expressed by (21) by generalizing the reasoning of [22, Lemma 6.1]. Let $\boldsymbol{t}, \boldsymbol{w} \in \boldsymbol{\Gamma}_T^k$ with $\boldsymbol{w} = \mathbf{K}_T\boldsymbol{\nabla}v$ for a specific $v \in \mathbb{P}_d^{k+1,0}(T)$. The following holds:

$$\sum_{F\in\mathcal{F}_T}(v, \boldsymbol{t}\cdot\boldsymbol{n}_{TF})_F = \int_T \boldsymbol{\nabla}\cdot(v\boldsymbol{t}) = (\boldsymbol{\nabla}v, \boldsymbol{t})_T + (v, \boldsymbol{\nabla}\cdot\boldsymbol{t})_T,$$

that is to say, owing to the symmetry of $\mathbf{K}_T$,

$$(\mathbf{K}_T^{-1}\boldsymbol{t},\boldsymbol{w})_T = -(v,\boldsymbol{\nabla}\cdot\boldsymbol{t})_T + \sum_{F\in\mathcal{F}_T}(v,\boldsymbol{t}\cdot\boldsymbol{n}_{TF})_F. \tag{30}$$

Let $\boldsymbol{\tau} = (\boldsymbol{\tau}_T,\{\tau_F\}_{F\in\mathcal{F}_T}) \in \boldsymbol{\Sigma}_T^k$ be such that $\boldsymbol{\tau} = I_T^k\boldsymbol{t}$ with I_T^k defined by (15). Owing to the local commutativity property from Proposition 1, we infer that $D_T^k\boldsymbol{\tau} = \pi_T^k(\boldsymbol{\nabla}\cdot\boldsymbol{t}) = \boldsymbol{\nabla}\cdot\boldsymbol{t}$, since $\boldsymbol{\nabla}\cdot\boldsymbol{t} \in \mathbb{P}_d^{k-1}(T) \subset \mathbb{P}_d^k(T)$. Moreover, since, for all $F\in\mathcal{F}_T$, $\boldsymbol{t}\cdot\boldsymbol{n}_F \in \mathbb{P}_{d-1}^k(F)$, we obtain $\tau_F = \boldsymbol{t}\cdot\boldsymbol{n}_F$, whence we infer from (30) that

$$(\mathbf{K}_T^{-1}\boldsymbol{t},\boldsymbol{w})_T = -(v,D_T^k\boldsymbol{\tau})_T + \sum_{F\in\mathcal{F}_T}(v,\tau_F\epsilon_{TF})_F. \tag{31}$$

Inspired by (31), we define the consistent part of the local flux reconstruction $\mathfrak{C}_T^k : \boldsymbol{\Sigma}_T^k \to \boldsymbol{\Gamma}_T^k$ such that, for all $\boldsymbol{\tau} = (\boldsymbol{\tau}_T,\{\tau_F\}_{F\in\mathcal{F}_T}) \in \boldsymbol{\Sigma}_T^k$ and all $\boldsymbol{w}\in\boldsymbol{\Gamma}_T^k$ with $\boldsymbol{w} = \mathbf{K}_T\boldsymbol{\nabla}v$ for a specific $v\in\mathbb{P}_d^{k+1,0}(T)$,

$$(\mathbf{K}_T^{-1}\mathfrak{C}_T^k\boldsymbol{\tau},\boldsymbol{w})_T = -(v,D_T^k\boldsymbol{\tau})_T + \sum_{F\in\mathcal{F}_T}(v,\tau_F\epsilon_{TF})_F. \tag{32}$$

Recalling that $\mathfrak{C}_T^k\boldsymbol{\tau}\in\boldsymbol{\Gamma}_T^k$ means that there exists $z\in\mathbb{P}_d^{k+1,0}(T)$ such that $\mathfrak{C}_T^k\boldsymbol{\tau} = \mathbf{K}_T\boldsymbol{\nabla}z$, we can reformulate (32) as the (well-posed) Neumann problem: Find $z\in\mathbb{P}_d^{k+1,0}(T)$ such that

$$(\mathbf{K}_T\boldsymbol{\nabla}z,\boldsymbol{\nabla}v)_T = -(v,D_T^k\boldsymbol{\tau})_T + \sum_{F\in\mathcal{F}_T}(v,\tau_F\epsilon_{TF})_F \qquad \forall v\in\mathbb{P}_d^{k+1,0}(T). \tag{33}$$

For further use, we note the following relation that stems from (32) after integrating by parts the left-hand side: For all $\boldsymbol{\tau}\in\boldsymbol{\Sigma}_T^k$ and all $v\in\mathbb{P}_d^{k+1,0}(T)$:

$$((\boldsymbol{\nabla}\cdot\mathfrak{C}_T^k - D_T^k)\boldsymbol{\tau},v)_T = \sum_{F\in\mathcal{F}_T}\epsilon_{TF}(\mathfrak{C}_T^k\boldsymbol{\tau}\cdot\boldsymbol{n}_F - \tau_F,v)_F. \tag{34}$$

Lemma 3 (Properties of $\mathfrak{C}_T^k$)**.** *Let $\mathfrak{C}_T^k$ be defined by* (32)*. Then, condition* (21) *holds. Additionally, there exists $\eta_1>0$ independent of h and $\mathbf{K}$ such that, for all $\boldsymbol{\tau}\in\boldsymbol{\Sigma}_T^k$,*

$$\|\mathbf{K}_T^{-1/2}\mathfrak{C}_T^k\boldsymbol{\tau}\|_T^2 \geqslant \lambda_{\sharp,T}^{-1}\|\boldsymbol{\tau}_T\|_T^2, \tag{35a}$$

$$\|\mathbf{K}_T^{-1/2}\mathfrak{C}_T^k\boldsymbol{\tau}\|_T^2 \leqslant \eta_1^{-1}\lambda_{\flat,T}^{-1}\|\boldsymbol{\tau}\|_T^2. \tag{35b}$$

Proof. Condition (21) follows from (31) and (32) since for all $\boldsymbol{t}\in\boldsymbol{\Gamma}_T^k$, $(\mathfrak{C}_T^kI_T^k\boldsymbol{t}-\boldsymbol{t})\in\boldsymbol{\Gamma}_T^k$ and $(\mathbf{K}_T^{-1}(\mathfrak{C}_T^kI_T^k\boldsymbol{t}-\boldsymbol{t}),\boldsymbol{w})_T = 0$ for all $\boldsymbol{w}\in\boldsymbol{\Gamma}_T^k$. Let us prove (35).

(i) *Proof of* (35a). Let $\boldsymbol{\tau} := (\boldsymbol{\tau}_T,\{\tau_F\}_{F\in\mathcal{F}_T}) \in \boldsymbol{\Sigma}_T^k$ be given with $\boldsymbol{\tau}_T = \mathbf{K}_T\boldsymbol{\nabla}v$ for a specific $v\in\mathbb{P}_d^{k,0}(T)$. Clearly, $\boldsymbol{\tau}_T\in\boldsymbol{\Gamma}_T^k$. Letting $\boldsymbol{w} = \boldsymbol{\tau}_T$ in (32) and using (10) to replace the first term on the right-hand side (this is possible since $v\in U_T^k$), it is inferred that

$$(\mathbf{K}_T^{-1}\mathfrak{C}_T^k\boldsymbol{\tau},\boldsymbol{\tau}_T)_T = (\boldsymbol{\nabla}v,\boldsymbol{\tau}_T)_T = \|\mathbf{K}_T^{-1/2}\boldsymbol{\tau}_T\|_T^2.$$

Using the Cauchy–Schwarz inequality to bound the left-hand side from above, we infer (35a) since

$$\lambda_{\sharp,T}^{-1/2}\|\boldsymbol{\tau}_T\|_T \leqslant \|\mathbf{K}_T^{-1/2}\boldsymbol{\tau}_T\|_T \leqslant \|\mathbf{K}_T^{-1/2}\mathfrak{C}_T^k\boldsymbol{\tau}\|_T.$$

(ii) *Proof of* (35b). Let $\boldsymbol{\tau} := (\boldsymbol{\tau}_T, \{\tau_F\}_{F\in\mathcal{F}_T}) \in \boldsymbol{\Sigma}_T^k$ with $\mathfrak{C}_T^k\boldsymbol{\tau} = \mathbf{K}_T\boldsymbol{\nabla}v$ for a specific $v \in \mathbb{P}_d^{k+1,0}(T)$. Recalling (10) and (32), and using $D_T^k\boldsymbol{\tau} \in U_T^k$, we infer that

$$\begin{aligned}\|\mathbf{K}_T^{-1/2}\mathfrak{C}_T^k\boldsymbol{\tau}\|_T^2 &= -(\pi_T^k v, D_T^k\boldsymbol{\tau})_T + \sum_{F\in\mathcal{F}_T}(v,\tau_F\epsilon_{TF})_F\\ &= (\boldsymbol{\tau}_T, \boldsymbol{\nabla}\pi_T^k v)_T + \sum_{F\in\mathcal{F}_T}(v-\pi_T^k v,\tau_F\epsilon_{TF})_F := \mathfrak{T}_1+\mathfrak{T}_2.\end{aligned}$$

We obtain

$$|\mathfrak{T}_1| \leqslant \|\boldsymbol{\tau}_T\|_T\|\boldsymbol{\nabla}\pi_T^k v\|_T \lesssim \|\boldsymbol{\tau}_T\|_T\|\boldsymbol{\nabla}v\|_T \lesssim \lambda_{\flat,T}^{-1}\|\boldsymbol{\tau}_T\|_T\|\mathbf{K}_T\boldsymbol{\nabla}v\|_T = \lambda_{\flat,T}^{-1}\|\boldsymbol{\tau}_T\|_T\|\mathfrak{C}_T^k\boldsymbol{\tau}\|_T,$$

where we have used the H^1-continuity of the L^2-orthogonal projection along with $\mathfrak{C}_T^k\boldsymbol{\tau} = \mathbf{K}_T\boldsymbol{\nabla}v$ to conclude. On the other hand, using the Cauchy–Schwarz inequality followed by the discrete trace and Poincaré's inequalities, it is inferred that, for all $F \in \mathcal{F}_T$,

$$\begin{aligned}|(v-\pi_T^k v,\tau_F)_F| &\leqslant C_{\mathrm{tr}}h_F^{-1/2}\|v-\pi_T^k v\|_T\|\tau_F\|_F\\ &\lesssim C_{\mathrm{tr}}C_{\mathrm{P}}h_F^{1/2}\|\boldsymbol{\nabla}v\|_T\|\tau_F\|_F \leqslant C_{\mathrm{tr}}C_{\mathrm{P}}\lambda_{\flat,T}^{-1}\|\mathfrak{C}_T^k\boldsymbol{\tau}\|_T h_F^{1/2}\|\tau_F\|_F,\end{aligned}$$

where we have used that $\|v-\pi_T^k v\|_T \leqslant \|v-\overline{v}\|_T \leqslant C_{\mathrm{P}}h_T\|\boldsymbol{\nabla}v\|_T \lesssim C_{\mathrm{P}}h_F\|\boldsymbol{\nabla}v\|_T$. Collecting the above estimates and using the discrete Cauchy–Schwarz inequality for the boundary terms together with (3) yields (35b). □

Remark 4 (Alternatives to the zero-average condition)**.** *The zero-average condition on the potential defining the elements of $\mathbb{T}_T^k$ and $\boldsymbol{\Gamma}_T^k$ ensures the well-posedness of the local Neumann problems* (33)*. An alternative condition commonly used in finite volume methods consists in considering functions that vanish at a given point inside the element (usually the point $\boldsymbol{x}_T$ of* **(M3)***). This corresponds to replacing the polynomial spaces $\mathbb{P}_d^{l,0}(T)$, $l \in \{k, k+1\}$, with the polynomial spaces $\mathbb{P}_d^{l,*}(T)$ of functions that vanish at $\boldsymbol{x}_T$. The local Neumann problems* (33) *remain well-posed since the right-hand side vanishes if v is a constant function.*

2.4.2 Construction of the residual on a pyramidal submesh

We now turn to devising a residual $\mathfrak{J}_T^k : \boldsymbol{\Sigma}_T^k \to \boldsymbol{\mathcal{R}}_T^k$; its design criteria are to match the orthogonality property (23) and the local stability property (25a). The latter means that $\|\mathbf{K}_T^{-1/2}\mathfrak{R}_T^k\boldsymbol{\tau}\|_T^2 = \|\mathbf{K}_T^{-1/2}\mathfrak{C}_T^k\boldsymbol{\tau}\|_T^2 + \|\mathbf{K}_T^{-1/2}\mathfrak{J}_T^k\boldsymbol{\tau}\|_T^2$ must control the first and the third terms on the right-hand side of (11) for all $\boldsymbol{\tau} \in \ker(D_T^k)$. Note that $\mathfrak{C}_T^k$ cannot achieve this control alone, since it only bounds the first term on the right-hand side of (11), see (35a).

Owing to **(M3)**, for every element $T \in \mathcal{T}_h$ it is possible to define a submesh composed of the face-based pyramids $\{\mathcal{P}_{TF}\}_{F\in\mathcal{F}_T}$ with apex $\boldsymbol{x}_T$. It has been proved in [19, Lemma 3] that the pyramidal submesh thus obtained inherits the shape- and contact-regularity properties of the original mesh. As a result, trace and inverse inequalities analogous to (4)–(6) hold.

For all $F \in \mathcal{F}_T$, we set $\boldsymbol{\Gamma}_{TF}^k := \mathbf{K}_T\boldsymbol{\nabla}\mathbb{P}_d^{k+1,0}(\mathcal{P}_{TF})$ extended by 0 outside $\mathcal{P}_{TF}$, and define the pyramidal residual $\mathfrak{J}_{TF}^k : \boldsymbol{\Sigma}_T^k \to \boldsymbol{\Gamma}_{TF}^k$ such that, for all $\boldsymbol{\tau} = (\boldsymbol{\tau}_T, \{\tau_F\}_{F\in\mathcal{F}_T}) \in \boldsymbol{\Sigma}_T^k$ and all $\boldsymbol{w} \in \boldsymbol{\Gamma}_{TF}^k$ with $\boldsymbol{w}_{|\mathcal{P}_{TF}} = \mathbf{K}_T\boldsymbol{\nabla}v$ for a specific $v \in \mathbb{P}_d^{k+1,0}(\mathcal{P}_{TF})$,

$$(\mathbf{K}_T^{-1}\mathfrak{J}_{TF}^k\boldsymbol{\tau},\boldsymbol{w})_{\mathcal{P}_{TF}} = ((\boldsymbol{\nabla}\cdot\mathfrak{C}_T^k - D_T^k)\boldsymbol{\tau}, v)_{\mathcal{P}_{TF}} - \epsilon_{TF}(\mathfrak{C}_T^k\boldsymbol{\tau}\cdot\boldsymbol{n}_F - \tau_F, v)_F. \tag{36}$$

Computing $\mathfrak{J}_{TF}^k$ amounts to solving a (well-posed) Neumann problem inside each pyramid. Note that here the zero mean condition on v is important since the right-hand side of (36) does not necessarily vanish for constant v. We define the residual by collecting all the pyramidal residuals,

$$\mathfrak{J}_T^k := \sum_{F\in\mathcal{F}_T} \mathfrak{J}_{TF}^k, \qquad \boldsymbol{\mathcal{R}}_T^k := \bigoplus_{T\in\mathcal{T}_h} \boldsymbol{\Gamma}_{TF}^k. \tag{37}$$

We can then reformulate (28) as follows:

$$\forall \boldsymbol{\tau}\in\boldsymbol{\Sigma}_T^k, \qquad \|\boldsymbol{\tau}\|_{H,T}^2 = \|\mathbf{K}_T^{-1/2}\mathfrak{R}_T^k\boldsymbol{\tau}\|_T^2 = \|\mathbf{K}_T^{-1/2}\mathfrak{C}_T^k\boldsymbol{\tau}\|_T^2 + \sum_{F\in\mathcal{F}_T} \|\mathbf{K}_T^{-1/2}\mathfrak{J}_{TF}^k\boldsymbol{\tau}\|_{\mathcal{P}_{TF}}^2. \tag{38}$$

Proposition 5 (Consistency and orthogonality of the residual)**.** *Conditions* (22) *and* (23) *hold.*

Proof. To prove (22), we observe that for all $\boldsymbol{t}\in\boldsymbol{\Gamma}_T^k$, $\mathfrak{C}_T^k(I_T^k\boldsymbol{t}) = \boldsymbol{t}$ so that $\boldsymbol{\nabla}\cdot\mathfrak{C}_T^k(I_T^k\boldsymbol{t}) = \boldsymbol{\nabla}\cdot\boldsymbol{t} = \pi_T^k(\boldsymbol{\nabla}\cdot\boldsymbol{t}) = D_T^k(I_T^k\boldsymbol{t})$. Moreover, for all $F\in\mathcal{F}_T$, $\mathfrak{C}_T^k(I_T^k\boldsymbol{t})\cdot\boldsymbol{n}_F = \boldsymbol{t}\cdot\boldsymbol{n}_F = \pi_F^k(\boldsymbol{t}\cdot\boldsymbol{n}_F)$. Hence, (36) yields that $\mathfrak{J}_{TF}^k(I_T^k\boldsymbol{t}) = 0$. To prove (23), we consider an arbitrary $\boldsymbol{t}\in\boldsymbol{\Gamma}_T^k$, sum (36) tested with $\boldsymbol{w} = \boldsymbol{t}_{|\mathcal{P}_{TF}}$ for all $F\in\mathcal{F}_T$, and use (34). □

Lemma 6 (Properties of $\mathfrak{R}_T^k$)**.** *Let $\mathfrak{C}_T^k$ and $\mathfrak{J}_T^k$ be defined by* (10)-(32) *and* (36)-(37)*, respectively. Let $\mathfrak{R}_T^k = \mathfrak{C}_T^k + \mathfrak{J}_T^k$. Then, properties* **(R1)**–**(R2)** *hold.*

Proof. The three properties composing **(R1)** have already been established. It only remains to check the properties (25) composing **(R2)**.

(i) *Proof of* (25a). Let $\boldsymbol{\tau}\in\ker(D_T^k)$, let $F\in\mathcal{F}_T$, and denote by $y_{TF}\in\mathbb{P}_d^k(\mathcal{P}_{TF})$ the function obtained by extending the function $h_F^{1/2}\epsilon_{TF}(\mathfrak{C}_T^k\boldsymbol{\tau}_{|F}\cdot\boldsymbol{n}_F - \tau_F)$ to $\mathcal{P}_{TF}$ by constant values along the direction of $\boldsymbol{x}_T - \overline{\boldsymbol{x}}_F$ (with $\overline{\boldsymbol{x}}_F$ denoting the barycenter of F). By construction and using mesh regularity, the following bound holds:

$$\|y_{TF}\|_{\mathcal{P}_{TF}} \lesssim h_F\|\mathfrak{C}_T^k\boldsymbol{\tau}\cdot\boldsymbol{n}_F - \tau_F\|_F. \tag{39}$$

Letting $\boldsymbol{w} = \mathbf{K}_T\boldsymbol{\nabla}y_{TF}\in\boldsymbol{\Gamma}_{TF}^k$ and $v = y_{TF}$ in (36) and using $D_T^k\boldsymbol{\tau} = 0$, it is inferred that

$$h_F^{1/2}\|\mathfrak{C}_T^k\boldsymbol{\tau}\cdot\boldsymbol{n}_F - \tau_F\|_F^2 = (\boldsymbol{\nabla}\cdot\mathfrak{C}_T^k\boldsymbol{\tau}, y_{TF})_{\mathcal{P}_{TF}} - (\mathfrak{J}_{TF}^k\boldsymbol{\tau}, \boldsymbol{\nabla}y_{TF})_{\mathcal{P}_{TF}} := \mathfrak{T}_1 + \mathfrak{T}_2. \tag{40}$$

Using the Cauchy–Schwarz inequality followed by the inverse inequality on $\mathcal{P}_{TF}$ together with (39) and using mesh regularity, it is inferred that

$$|\mathfrak{T}_1| \leqslant \|\boldsymbol{\nabla}\cdot\mathfrak{C}_T^k\boldsymbol{\tau}\|_{\mathcal{P}_{TF}}\|y_{TF}\|_{\mathcal{P}_{TF}} \lesssim \|\mathfrak{C}_T^k\boldsymbol{\tau}\|_{\mathcal{P}_{TF}}\|\mathfrak{C}_T^k\boldsymbol{\tau}\cdot\boldsymbol{n}_F - \tau_F\|_F.$$

Proceeding in a similar way for the second term, we obtain

$$|\mathfrak{T}_2| \leqslant \|\mathfrak{J}_{TF}^k\boldsymbol{\tau}\|_{\mathcal{P}_{TF}}\|\boldsymbol{\nabla}y_{TF}\|_{\mathcal{P}_{TF}} \lesssim \|\mathfrak{J}_{TF}^k\boldsymbol{\tau}\|_{\mathcal{P}_{TF}}\|\mathfrak{C}_T^k\boldsymbol{\tau}\cdot\boldsymbol{n}_F - \tau_F\|_F.$$

Using the above bounds to estimate the right-hand side of (40), it is inferred that

$$\lambda_{\sharp,T}^{-1/2}h_F^{1/2}\|\mathfrak{C}_T^k\boldsymbol{\tau}\cdot\boldsymbol{n}_F - \tau_F\|_F \lesssim \|\mathbf{K}_T^{-1/2}\mathfrak{C}_T^k\boldsymbol{\tau}\|_{\mathcal{P}_{TF}} + \|\mathbf{K}_T^{-1/2}\mathfrak{J}_{TF}^k\boldsymbol{\tau}\|_{\mathcal{P}_{TF}}. \tag{41}$$

Recalling (11) and the fact that $D_T^k\boldsymbol{\tau} = 0$, we then proceed as follows using (35a) together with the triangular inequality followed by (41) and the discrete trace inequality on $\mathcal{P}_{TF}$,

$$\begin{aligned}
|||\boldsymbol{\tau}|||_T^2 &= \|\boldsymbol{\tau}_T\|_T^2 + \sum_{F\in\mathcal{F}_T} h_F\|\tau_F\|_F^2\\
&\lesssim \lambda_{\sharp,T}\|\mathbf{K}_T^{-1/2}\mathfrak{C}_T^k\boldsymbol{\tau}\|_T^2 + \sum_{F\in\mathcal{F}_T} h_F\|\mathfrak{C}_T^k\boldsymbol{\tau}\cdot\boldsymbol{n}_F - \tau_F\|_F^2 + \sum_{F\in\mathcal{F}_T} h_F\|\mathfrak{C}_T^k\boldsymbol{\tau}\cdot\boldsymbol{n}_F\|_F^2\\
&\lesssim \lambda_{\sharp,T}\|\mathbf{K}_T^{-1/2}\mathfrak{C}_T^k\boldsymbol{\tau}\|_T^2 + \lambda_{\sharp,T}\sum_{F\in\mathcal{F}_T}\left\{\|\mathbf{K}_T^{-1/2}\mathfrak{C}_T^k\boldsymbol{\tau}\|_{\mathcal{P}_{TF}}^2 + \|\mathbf{K}_T^{-1/2}\mathfrak{J}_{TF}^k\boldsymbol{\tau}\|_{\mathcal{P}_{TF}}^2\right\}\\
&\lesssim \lambda_{\sharp,T}\|\mathbf{K}_T^{-1/2}\mathfrak{R}_T^k\boldsymbol{\tau}\|_T^2,
\end{aligned}$$

where the last inequality is a consequence of (38). This proves (25a).

(ii) *Proof of* (25b). Let $\boldsymbol{\tau}\in\boldsymbol{\Sigma}_T^k$. Recalling (35b) together with (38), it suffices to prove the continuity of the residuals in the $|||\cdot|||_T$-norm. Denote by $\mathfrak{T}_1$ and $\mathfrak{T}_2$ the terms on the right-hand side of (36). Using the Cauchy–Schwarz and triangle inequalities followed by the inverse and Poincaré's inequalities on $\mathcal{P}_{TF}$ and using mesh regularity, it is inferred for the first term that

$$\begin{aligned}
|\mathfrak{T}_1| &\leqslant \left(\|\boldsymbol{\nabla}\cdot\mathfrak{C}_T^k\boldsymbol{\tau}\|_{\mathcal{P}_{TF}} + \|D_T^k\boldsymbol{\tau}\|_{\mathcal{P}_{TF}}\right)\|v\|_{\mathcal{P}_{TF}}\\
&\lesssim \left(C_{\mathrm{inv}}h_T^{-1}\|\mathfrak{C}_T^k\boldsymbol{\tau}\|_{\mathcal{P}_{TF}} + \|D_T^k\boldsymbol{\tau}\|_{\mathcal{P}_{TF}}\right)C_{\mathrm{P}}h_T\|\boldsymbol{\nabla}v\|_{\mathcal{P}_{TF}}\\
&\lesssim \lambda_{\flat,T}^{-1/2}\left(\|\mathfrak{C}_T^k\boldsymbol{\tau}\|_{\mathcal{P}_{TF}} + h_T\|D_T^k\boldsymbol{\tau}\|_{\mathcal{P}_{TF}}\right)\|\mathbf{K}_T^{-1/2}\boldsymbol{w}\|_{\mathcal{P}_{TF}}.
\end{aligned}$$

For the second term, the Cauchy–Schwarz, triangle, discrete trace, and Poincaré's inequalities on $\mathcal{P}_{TF}$ yield

$$|\mathfrak{T}_2| \leqslant \lambda_{\flat,T}^{-1/2}\left(h_F^{1/2}\|\tau_F\|_F + \|\mathfrak{C}_T^k\boldsymbol{\tau}\|_{\mathcal{P}_{TF}}\right)\|\mathbf{K}_T^{-1/2}\boldsymbol{w}\|_{\mathcal{P}_{TF}}.$$

Collecting the above results leads to

$$\lambda_{\flat,T}^{1/2}\|\mathbf{K}_T^{-1/2}\mathfrak{J}_{TF}^k\boldsymbol{\tau}\|_{\mathcal{P}_{TF}} = \lambda_{\flat,T}^{1/2}\sup_{\boldsymbol{w}\in\boldsymbol{\Gamma}_{TF}^k\setminus\{\mathbf{0}\}}\frac{\mathfrak{T}_1+\mathfrak{T}_2}{\|\mathbf{K}_T^{-1/2}\boldsymbol{w}\|_{\mathcal{P}_{TF}}} \lesssim \|\mathfrak{C}_T^k\boldsymbol{\tau}\|_{\mathcal{P}_{TF}} + h_T\|D_T^k\boldsymbol{\tau}\|_{\mathcal{P}_{TF}} + h_F^{1/2}\|\tau_F\|_F.$$

The continuity of $\mathfrak{J}_T^k$ in the $|||\cdot|||_T$-norm then follows squaring the above inequality, summing over the element faces, and concluding with the discrete Cauchy–Schwarz inequality. □

Remark 7 (Penalty coefficient)**.** *A more general form for the residual can be obtained introducing a scalar penalty coefficient $\mu > 0$ and defining $\mathfrak{J}_{TF}^k$ as follows:*

$$\mu^{-1}(\mathbf{K}_T^{-1}\mathfrak{J}_{TF}^k\boldsymbol{\tau},\boldsymbol{w})_{\mathcal{P}_{TF}} = ((\boldsymbol{\nabla}\cdot\mathfrak{C}_T^k - D_T^k)\boldsymbol{\tau},v)_{\mathcal{P}_{TF}} - \epsilon_{TF}(\mathfrak{C}_T^k\boldsymbol{\tau}\cdot\boldsymbol{n}_F - \tau_F,v)_F. \tag{42}$$

2.5 Discrete problem and well-posedness

We consider the following discretization of (2) based on the spaces $\boldsymbol{\Sigma}_h^k$ and U_h^k defined by (8) and (12), respectively: Find $(\boldsymbol{\sigma}_h,u_h)\in\boldsymbol{\Sigma}_h^k\times U_h^k$ such that

$$H(\boldsymbol{\sigma}_h,\boldsymbol{\tau}_h) + (u_h, D_h^k\boldsymbol{\tau}_h) = 0 \qquad \forall\boldsymbol{\tau}_h\in\boldsymbol{\Sigma}_h^k, \tag{43a}$$

$$-(D_h^k\boldsymbol{\sigma}_h, v_h) = (f,v_h) \qquad \forall v_h\in U_h^k, \tag{43b}$$

with bilinear form H defined by (20). Owing to the choice for U_h^k, equation (43b) can be alternatively written, letting $f_h := \pi_h^k f$,

$$-D_h^k \boldsymbol{\sigma}_h = f_h. \tag{44}$$

Lemma 8 (Well-posedness)**.** *There exists $\beta > 0$ independent of h and $\mathbf{K}$ such that, for all $v_h \in U_h^k$, the following holds:*

$$\beta \|v_h\| \leqslant \sup_{\boldsymbol{\tau}_h \in \boldsymbol{\Sigma}_h^k \backslash \{\mathbf{0}\}} \frac{(D_h^k \boldsymbol{\tau}_h, v_h)}{\|\boldsymbol{\tau}_h\|}. \tag{45}$$

Additionally, problem (43) *is well-posed.*

Proof. The surjectivity of the operator D_h^k expressed by (45) is a classical consequence of the existence of an interpolation operator satisfying the commuting property of Proposition 1 and uniformly continuous in the sense that, for all $\boldsymbol{\tau} \in \boldsymbol{\Sigma}^+$, $\|I_h^k \boldsymbol{\tau}\| \leqslant C \|\boldsymbol{\tau}\|_{\boldsymbol{\Sigma}^+}$ holds with $C > 0$ independent of the meshsize. The well-posedness of problem (43) then results from (45) together with the $\|\cdot\|$-coercivity of the bilinear form H on $\ker(D_h^k)$, a consequence of (26a). Further details can be found in the book of Brezzi and Fortin [11], cf. in particular Section IV.1.2. □

3 Convergence analysis

In this section we prove an error estimate and, under regularity assumptions for the exact solution, identify convergence rates for the flux and potential approximations.

3.1 Basic error estimate

Lemma 9 (Basic error estimate)**.** *Let $(\boldsymbol{s}, u) \in \boldsymbol{\Sigma} \times U$ denote the unique solution to* (2) *and assume the additional regularity $\boldsymbol{s} \in \boldsymbol{\Sigma}^+$ with $\boldsymbol{\Sigma}^+$ defined in Section 2.3. Let $\widehat{\boldsymbol{\sigma}}_h := I_h^k \boldsymbol{s}$ and $\widehat{u}_h := \pi_h^k u$ where I_h^k is the interpolator defined by* (17) *and π_h^k is the L^2-orthogonal projector on $\mathbb{P}_d^k(\mathcal{T}_h)$. Denoting by $(\boldsymbol{\sigma}_h, u_h) \in \boldsymbol{\Sigma}_h^k \times U_h^k$ the unique solution to* (43)*, the following holds:*

$$\max\left(\tfrac{1}{2}\beta(\eta\lambda_\flat)^{1/2}\|\widehat{u}_h - u_h\|, \|\widehat{\boldsymbol{\sigma}}_h - \boldsymbol{\sigma}_h\|_H\right) \leqslant \sup_{\boldsymbol{\tau}_h \in \boldsymbol{\Sigma}_h^k,\, \|\boldsymbol{\tau}_h\|_H = 1} \mathcal{E}_h(\boldsymbol{\tau}_h). \tag{46}$$

with consistency error $\mathcal{E}_h(\boldsymbol{\tau}_h) := H(\widehat{\boldsymbol{\sigma}}_h, \boldsymbol{\tau}_h) + (\widehat{u}_h, D_h^k \boldsymbol{\tau}_h)$.

Proof. We denote by \$ the supremum on the right-hand side of (46). By virtue of (43a), the following holds for all $\boldsymbol{\tau}_h \in \boldsymbol{\Sigma}_h^k \backslash \{\mathbf{0}\}$:

$$H(\widehat{\boldsymbol{\sigma}}_h - \boldsymbol{\sigma}_h, \boldsymbol{\tau}_h) + (\widehat{u}_h - u_h, D_h^k \boldsymbol{\tau}_h) = \mathcal{E}_h(\boldsymbol{\tau}_h) = \mathcal{E}_h\left(\frac{\boldsymbol{\tau}_h}{\|\boldsymbol{\tau}_h\|_H}\right)\|\boldsymbol{\tau}_h\|_H \leqslant \$\|\boldsymbol{\tau}_h\|_H.$$

Hence, letting $\boldsymbol{\tau}_h = \widehat{\boldsymbol{\sigma}}_h - \boldsymbol{\sigma}_h$ in the above expression and using $(\widehat{\boldsymbol{\sigma}}_h - \boldsymbol{\sigma}_h) \in \ker(D_h^k)$ on the left-hand side, we obtain

$$\|\widehat{\boldsymbol{\sigma}}_h - \boldsymbol{\sigma}_h\|_H \leqslant \$ \tag{47}$$

Let us now estimate the error on the potential. Using (43a) together with the definition of the consistency error yields for all $\boldsymbol{\tau}_h \in \boldsymbol{\Sigma}_h^k$,

$$(D_h^k \boldsymbol{\tau}_h, \widehat{u}_h - u_h) = (D_h^k \boldsymbol{\tau}_h, \widehat{u}_h) - (D_h^k \boldsymbol{\tau}_h, u_h) = H(\boldsymbol{\sigma}_h - \widehat{\boldsymbol{\sigma}}_h, \boldsymbol{\tau}_h) + \mathcal{E}_h(\boldsymbol{\tau}_h).$$

Using the inf-sup condition (45) together with the above relation, the Cauchy–Schwarz inequality, and (26b), it is inferred that

$$\beta(\eta\lambda_\flat)^{1/2}\|\widehat{u}_h - u_h\| \leqslant \sup_{\boldsymbol{\tau}_h\in\boldsymbol{\Sigma}_h^k\setminus\{\mathbf{0}\}} \frac{(D_h^k\boldsymbol{\tau}_h, \widehat{u}_h - u_h)}{(\eta\lambda_\flat)^{-1/2}|||\boldsymbol{\tau}_h|||} \leqslant \|\boldsymbol{\sigma}_h - \widehat{\boldsymbol{\sigma}}_h\|_H + \$,$$

and the conclusion follows using (47). □

Corollary 10 ($|||\cdot|||$-error estimate)**.** *The following holds:*

$$(\eta/\lambda_\sharp)^{1/2}|||\widehat{\boldsymbol{\sigma}}_h - \boldsymbol{\sigma}_h||| \leqslant \sup_{\boldsymbol{\tau}_h\in\boldsymbol{\Sigma}_h^k,\, |||\boldsymbol{\tau}_h|||_H=1} \mathcal{E}_h(\boldsymbol{\tau}_h).$$

To estimate the convergence rate, we need to work with the unpatched space of flux DOFs

$$\check{\boldsymbol{\Sigma}}_h^k := \bigtimes_{T\in\mathcal{T}_h} \boldsymbol{\Sigma}_T^k. \tag{48}$$

The restriction operator L_T, for all $T\in\mathcal{T}_h$, can be extended to $\check{\boldsymbol{\Sigma}}_h^k$ and stills maps onto $\boldsymbol{\Sigma}_T^k$.

Theorem 11 (Convergence rate)**.** *Under the assumptions of Lemma 9, and assuming the additional regularity $u\in H_0^1(\Omega)\cap H^{k+2}(\mathcal{T}_h)$, the following holds:*

$$\beta(\eta\lambda_\flat)^{1/2}\|\widehat{u}_h - u_h\| + \|\widehat{\boldsymbol{\sigma}}_h - \boldsymbol{\sigma}_h\|_H \leqslant C\left(\sum_{T\in\mathcal{T}_h}\rho_{\mathbf{K},T}\lambda_{\sharp,T}h_T^{2(k+1)}\|u\|_{H^{k+2}(T)}^2\right)^{1/2}, \tag{49}$$

with C independent of h and $\mathbf{K}$ and anisotropy ratio $\rho_{\mathbf{K},T} := \lambda_{\sharp,T}/\lambda_{\flat,T}$.

Proof. We extend the bilinear form H defined by (20) to $\check{\boldsymbol{\Sigma}}_h^k\times\check{\boldsymbol{\Sigma}}_h^k$. With $\check{u}_h := \pi_h^{k+1}u$ and $\check{\boldsymbol{\sigma}}_h\in\check{\boldsymbol{\Sigma}}_h^k$ such that $L_T\check{\boldsymbol{\sigma}}_h = I_T^k(\mathbf{K}\boldsymbol{\nabla}\check{u}_h)_{|T}$ for all $T\in\mathcal{T}_h$, we infer that, for all $\boldsymbol{\tau}_h\in\boldsymbol{\Sigma}_h^k$,

$$\mathcal{E}_h(\boldsymbol{\tau}_h) = H(\widehat{\boldsymbol{\sigma}}_h - \check{\boldsymbol{\sigma}}_h, \boldsymbol{\tau}_h) + (\widehat{u}_h - \check{u}_h, D_h^k\boldsymbol{\tau}_h) + \Big\{H(\check{\boldsymbol{\sigma}}_h, \boldsymbol{\tau}_h) + (\check{u}_h, D_h^k\boldsymbol{\tau}_h)\Big\} := \mathfrak{T}_1 + \mathfrak{T}_2 + \mathfrak{T}_3.$$

We estimate the terms on the right-hand side.

(i) *Estimate of $\mathfrak{T}_1$.* Using the Cauchy–Schwarz inequality leads to

$$|\mathfrak{T}_1| \leqslant \left\{\sum_{T\in\mathcal{T}_h}\|L_T\widehat{\boldsymbol{\sigma}}_h - L_T\check{\boldsymbol{\sigma}}_h\|_{H,T}^2\right\}^{1/2} |||\boldsymbol{\tau}_h|||_H. \tag{50}$$

Owing to the local continuity assumption expressed by (25b), and recalling (11), we infer that, for all $T\in\mathcal{T}_h$,

$$\lambda_{\flat,T}\|L_T\widehat{\boldsymbol{\sigma}}_h - L_T\check{\boldsymbol{\sigma}}_h\|_{H,T}^2 \lesssim \|\widehat{\boldsymbol{\sigma}}_T - \check{\boldsymbol{\sigma}}_T\|_T^2 + \sum_{F\in\mathcal{F}_T} h_F\|\widehat{\sigma}_F - (\mathbf{K}\boldsymbol{\nabla}\check{u}_h)_{|T}\cdot\boldsymbol{n}_F\|_F^2 + h_T^2\|D_T^k(L_T\widehat{\boldsymbol{\sigma}}_h - L_T\check{\boldsymbol{\sigma}}_h)\|_T^2, \tag{51}$$

with $\sigma_F = \pi_F^k(\mathbf{K}\boldsymbol{\nabla}u)$. To estimate the first term on the right-hand side of (51), we observe that

$$\|\widehat{\boldsymbol{\sigma}}_T - \check{\boldsymbol{\sigma}}_T\|_T = \|\varpi_T^k(\mathbf{K}_T\boldsymbol{\nabla}(u - \check{u}_h))\|_T \leqslant \lambda_{\sharp,T}\|\boldsymbol{\nabla}(u - \check{u}_h)\|_T \lesssim \lambda_{\sharp,T}h_T^{k+1}\|u\|_{H^{k+2}(T)},$$

where we have used the approximation properties of π_h^{k+1} and, recalling the definition (16) of ϖ_T^k, the fact that $\|\varpi_T^k \boldsymbol{t}\|_T \leqslant \lambda_{\sharp,T}^{1/2} \|\mathbf{K}_T^{-1/2} \varpi_T^k \boldsymbol{t}\|_T \leqslant \lambda_{\sharp,T}^{1/2} \|\mathbf{K}_T^{-1/2} \boldsymbol{t}\|_T \leqslant \lambda_{\sharp,T} \|\mathbf{K}_T^{-1} \boldsymbol{t}\|_T$ for all $\boldsymbol{t} \in L^2(T)^d$. For the square root of the second term on the right-hand side of (51), a similar estimate can be obtained after using the trace inequality (5) together with the continuity and approximation properties of π_F^k and (3). Finally, using the commutativity of the left diagram in Proposition 1 and the continuity and approximation properties of π_T^k, it is inferred that

$$\begin{aligned} h_T \|D_T^k(L_T \widehat{\boldsymbol{\sigma}}_h - L_T \breve{\boldsymbol{\sigma}}_h)\|_T &= h_T \|\pi_T^k \left[\boldsymbol{\nabla}\cdot(\mathbf{K}_T \boldsymbol{\nabla}(u - \breve{u}_h))\right]\|_T \\ &\leqslant h_T \|\boldsymbol{\nabla}\cdot(\mathbf{K}_T \boldsymbol{\nabla}(u - \breve{u}_h))\|_T \lesssim \lambda_{\sharp,T} h_T^{k+1} \|u\|_{H^{k+2}(T)}. \end{aligned}$$

Using the above results to bound the right-hand side of (51) and recalling (50), we conclude by the discrete Cauchy–Schwarz inequality that

$$|\mathfrak{T}_1| \lesssim \sum_{T\in\mathcal{T}_h} \left\{ \rho_{\mathbf{K},T} \lambda_{\sharp,T} h_T^{2(k+1)} \|u\|_{H^{k+2}(T)}^2 \right\}^{1/2} \|\boldsymbol{\tau}_h\|_H. \tag{52}$$

(ii) *Estimate of* $\mathfrak{T}_2$. For all $T \in \mathcal{T}_h$, we obtain

$$\begin{aligned} |(\widehat{u}_h - \breve{u}_h, D_T^k L_T \boldsymbol{\tau}_h)_T| &\leqslant \|\pi_T^k(u - \breve{u}_h)\|_T \ \|D_T^k L_T \boldsymbol{\tau}_h\|_T \\ &\lesssim h_T^{k+1} \|u\|_{H^{k+2}(T)} \ \|L_T \boldsymbol{\tau}_h\|_T \qquad \text{eq. (11)} \\ &\lesssim \lambda_{\sharp,T}^{1/2} h_T^{k+1} \|u\|_{H^{k+2}(T)} \|L_T \boldsymbol{\tau}_h\|_{H,T}, \qquad \text{eq. (25a)} \end{aligned}$$

where in the first line we have used $\|\pi_T^k(u - \breve{u}_h)\|_T \leqslant \|u - \breve{u}_h\|_T \lesssim h_T^{k+2} \|u\|_{H^{k+2}(T)}$. Hence, the discrete Cauchy–Schwarz inequality yields,

$$|\mathfrak{T}_2| \lesssim \sum_{T\in\mathcal{T}_h} \left\{ \lambda_{\sharp,T} h_T^{2(k+1)} \|u\|_{H^{k+2}(T)}^2 \right\}^{1/2} \|\boldsymbol{\tau}_h\|_H. \tag{53}$$

(iii) *Estimate of* $\mathfrak{T}_3$. Letting $\overline{u}_h := \pi_h^0 u$ and observing that, for all $T \in \mathcal{T}_h$, $\boldsymbol{\nabla}\overline{u}_{h|T} = 0$ and $(\breve{u}_h - \overline{u}_h)_{|T} \in \mathbb{P}_d^{k+1,0}(T)$, we infer that

$$\begin{aligned} \mathfrak{T}_3 &= \sum_{T\in\mathcal{T}_h} (\boldsymbol{\nabla}(\breve{u}_h - \overline{u}_h), \mathfrak{C}_T^k L_T \boldsymbol{\tau}_h)_T + (\breve{u}_h, D_h^k \boldsymbol{\tau}_h) \qquad \text{eq. (29)} \\ &= \sum_{T\in\mathcal{T}_h} \left\{ -(\breve{u}_h - \overline{u}_h, D_T^k L_T \boldsymbol{\tau}_h)_T + \sum_{F\in\mathcal{F}_T} ((\breve{u}_h - \overline{u}_h)_{|T}, \tau_F \epsilon_{TF})_F \right\} + (\breve{u}_h, D_h^k \boldsymbol{\tau}_h) \qquad \text{eq. (32)} \\ &= \sum_{T\in\mathcal{T}_h} \left\{ -(\breve{u}_h - \overline{u}_h, D_T^k L_T \boldsymbol{\tau}_h)_T + \sum_{F\in\mathcal{F}_T} (\breve{u}_{h|T}, \tau_F \epsilon_{TF})_F - (\overline{u}_h, D_T^k L_T \boldsymbol{\tau}_h)_T \right\} \qquad \text{eq. (10)} \\ &\quad + (\breve{u}_h, D_h^k \boldsymbol{\tau}_h) \\ &= \sum_{T\in\mathcal{T}_h} \sum_{F\in\mathcal{F}_T} (\breve{u}_{h|T} - u, \tau_F \epsilon_{TF})_F, \qquad \text{eq. (13)} \end{aligned}$$

where in the last line we have used $\sum_{T\in\mathcal{T}_h} \sum_{F\in\mathcal{F}_T} (u, \tau_F \epsilon_{TF})_F = 0$, a consequence of the fact that u has single-valued traces on interfaces and vanishes on $\partial\Omega$. Hence it is inferred,

using the Cauchy–Schwarz inequality, (11) and (14) together with (25a), the continuous trace inequality (5), the approximation properties of π_h^{k+1}, and mesh regularity, that

$$\begin{aligned}|\mathfrak{T}_3| &\leqslant \left\{\sum_{T\in\mathcal{T}_h}\sum_{F\in\mathcal{F}_T}\lambda_{\sharp,T}h_F^{-1}\|\breve{u}_{h|T}-u\|_F^2\right\}^{1/2}\times\left\{\sum_{T\in\mathcal{T}_h}\sum_{F\in\mathcal{F}_T}\lambda_{\sharp,T}^{-1}h_F\|\tau_F\|_F^2\right\}^{1/2}\\ &\lesssim \left\{\sum_{T\in\mathcal{T}_h}\lambda_{\sharp,T}h_T^{2(k+1)}\|u\|_{H^{k+2}(T)}^2\right\}\|\boldsymbol{\tau}_h\|_H.\end{aligned}\tag{54}$$

The estimate (49) follows from Lemma 9, together with the bounds (52), (53), (54), and $1\leqslant\rho_{\mathbf{K},T}$. □

Remark 12. *The above proof of the bound on $\mathfrak{T}_1$ shows that, for all $T\in\mathcal{T}_h$,*

$$\|L_T(\widehat{\boldsymbol{\sigma}}_h-\breve{\boldsymbol{\sigma}}_h)\|_{H,T}=\|L_T\widehat{\boldsymbol{\sigma}}_h-I_T^k\mathbf{K}_T\boldsymbol{\nabla}\breve{u}_h\|_{H,T}\lesssim\rho_{\mathbf{K},T}\lambda_{\sharp,T}h_T^{k+1}\|u\|_{H^{k+2}(T)}.\tag{55}$$

3.2 Supercloseness of the potential

For the sake of simplicity, we assume throughout this section that

$$\mathbf{K}=\mathrm{Id}_d,\tag{56}$$

where Id_d denotes the identity matrix in $\mathbb{R}^{d\times d}$. The estimate of the error on the potential can be refined if elliptic regularity holds in the following form: For all $g\in L^2(\Omega)$, the unique solution $z\in H_0^1(\Omega)$ to

$$(\boldsymbol{\nabla}z,\boldsymbol{\nabla}v)=(g,v)\qquad\forall v\in H_0^1(\Omega),\tag{57}$$

satisfies the a priori estimate

$$\|z\|_{H^2(\Omega)}+\|\boldsymbol{w}\|_{H^1(\Omega)^d}\leqslant C_{\mathrm{ell}}\|g\|,\qquad\boldsymbol{w}:=\boldsymbol{\nabla}z,\tag{58}$$

with C_{ell} only depending on Ω.

Theorem 13 (Supercloseness of the potential)**.** *Under the assumptions of Theorem 11,* (56) *and elliptic regularity, the following holds:*

$$\|\widehat{u}_h-u_h\|\leqslant Ch^{k+2}\|u\|_{H^{k+2}(\Omega)},\tag{59}$$

with $C>0$ independent of h.

Proof. We sketch the proof. Let z solve (57) with $g=u_h-\widehat{u}_h$. We let $\widehat{\boldsymbol{\sigma}}_h$ and $\widehat{u}_h$ be defined as in Lemma 9 and set $\boldsymbol{s}_h:=\mathfrak{R}_h^k\boldsymbol{\sigma}_h$, $\widehat{\boldsymbol{s}}_h:=\mathfrak{R}_h^k\widehat{\boldsymbol{\sigma}}_h$, $\breve{u}_h:=\pi_h^{k+1}u$, $z_h:=\pi_h^{k+1}z$, $\boldsymbol{v}_h:=I_h^k\boldsymbol{w}$, and $\boldsymbol{w}_h:=\mathfrak{R}_h^k\boldsymbol{v}_h$. Owing to Proposition 1, and since, by definition, $\boldsymbol{\nabla}\cdot\boldsymbol{w}=\widehat{u}_h-u_h\in\mathbb{P}_d^k(\mathcal{T}_h)$, the equality $D_h^k\boldsymbol{v}_h=\boldsymbol{\nabla}\cdot\boldsymbol{w}$ holds. Hence, taking $\boldsymbol{\tau}_h=\boldsymbol{v}_h$ in (43a) and recalling (56),

$$-(\boldsymbol{\nabla}\cdot\boldsymbol{w},u_h)=-(D_h^k\boldsymbol{v}_h,u_h)=H(\boldsymbol{\sigma}_h,\boldsymbol{v}_h)=(\boldsymbol{s}_h,\boldsymbol{w}_h).$$

Using the above relation and denoting by $\boldsymbol{\nabla}_h$ the broken gradient on $\mathcal{T}_h$ leads to

$$\begin{aligned}\|\widehat{u}_h-u_h\|^2&=(\widehat{u}_h-u_h,\boldsymbol{\nabla}\cdot\boldsymbol{w})=(u-u_h,\boldsymbol{\nabla}\cdot\boldsymbol{w})\\&=-(\boldsymbol{s},\boldsymbol{w})+(\boldsymbol{s}_h,\boldsymbol{w}_h)\\&=(\boldsymbol{s}_h-\boldsymbol{s},\boldsymbol{w}-\boldsymbol{\nabla}_hz_h)+(\boldsymbol{s}_h-\boldsymbol{\nabla}\breve{u}_h,\boldsymbol{w}_h-\boldsymbol{w})\\&\quad+(\boldsymbol{s}_h-\boldsymbol{s},\boldsymbol{\nabla}_hz_h)+(\boldsymbol{\nabla}\breve{u}_h,\boldsymbol{w}_h-\boldsymbol{w}):=\mathfrak{T}_1+\ldots+\mathfrak{T}_4,\end{aligned}\tag{60}$$

where in the second line we have used $\boldsymbol{s} = \boldsymbol{\nabla} u$. Using the Cauchy–Schwarz inequality, it is inferred that

$$|\mathfrak{T}_1| \leqslant \|\boldsymbol{s}_h - \boldsymbol{s}\| \|\boldsymbol{w} - \boldsymbol{\nabla}_h z_h\| \lesssim h^{k+2} \|u\|_{H^{k+2}(\Omega)} \|z\|_{H^2(\Omega)}, \tag{61}$$

since $\|\boldsymbol{s}_h - \boldsymbol{s}\| \leqslant \|\boldsymbol{s}_h - \boldsymbol{\nabla} \breve{u}_h\| + \|\boldsymbol{\nabla} \breve{u}_h - \boldsymbol{s}\|$, the first term is bounded using (55) (observing that $\|\boldsymbol{s}_h - \boldsymbol{s}\|_T = \|\mathfrak{R}_T^k L_T \widehat{\boldsymbol{\sigma}}_h - \boldsymbol{\nabla} \breve{u}_h\|_T = \|L_T \widehat{\boldsymbol{\sigma}}_h - I_T^k \boldsymbol{\nabla} \breve{u}_h\|_{H,T}$) and the second using the approximation properties of π_h^{k+1}, while $\|\boldsymbol{w} - \boldsymbol{\nabla}_h z_h\| \lesssim h \|z\|_{H^2(\Omega)}$. Proceeding similarly, it is inferred that

$$|\mathfrak{T}_2| \leqslant \|\boldsymbol{s}_h - \boldsymbol{\nabla} \breve{u}_h\| \|\boldsymbol{w}_h - \boldsymbol{w}\| \leqslant (\|\boldsymbol{s}_h - \widehat{\boldsymbol{s}}_h\| + \|\widehat{\boldsymbol{s}}_h - \boldsymbol{\nabla} \breve{u}_h\|) \|\boldsymbol{w}_h - \boldsymbol{w}\| \lesssim h^{k+2} \|u\|_{H^{k+2}(\Omega)} \|\boldsymbol{w}\|_{H^1(\Omega)^d}. \tag{62}$$

The estimate on $\mathfrak{T}_3$ deserves being treated in more detail. Letting $\overline{z}_h := \pi_h^0 z_h$, so that $(z_h - \overline{z}_h)_{|T} \in \boldsymbol{\Gamma}_T^k$ for all $T \in \mathcal{T}_h$, it is inferred that

$$\begin{aligned}
(\boldsymbol{s}_h, \boldsymbol{\nabla}_h z_h) &= (\boldsymbol{s}_h, \boldsymbol{\nabla}_h (z_h - \overline{z}_h)) = \sum_{T \in \mathcal{T}_h} (\mathfrak{C}_T^k L_T \boldsymbol{\sigma}_h, \boldsymbol{\nabla} (z_h - \overline{z}_h))_T && \text{eq. (29)} \\
&= \sum_{T \in \mathcal{T}_h} \left\{ -(z_h - \overline{z}_h, D_T^k L_T \boldsymbol{\sigma}_h)_T + \sum_{F \in \mathcal{F}_T} ((z_h - \overline{z}_h)_{|T}, \sigma_F \epsilon_{TF})_F \right\} && \text{eq. (32)} \\
&= \sum_{T \in \mathcal{T}_h} \left\{ -(z_h, D_T^k L_T \boldsymbol{\sigma}_h)_T + \sum_{F \in \mathcal{F}_T} (z_{h|T}, \sigma_F \epsilon_{TF})_F, \right\}, && \text{eq. (10)}
\end{aligned}$$

where L_T is the restriction operator defined in Section 2.2. Hence, upon element-by-element integration by parts of the term $(\boldsymbol{s}, \boldsymbol{\nabla}_h z_h)$, we obtain

$$\mathfrak{T}_3 = (z_h, \boldsymbol{\nabla} \cdot \boldsymbol{s} - D_h^k \boldsymbol{\sigma}_h) + \sum_{T \in \mathcal{T}_h} \sum_{F \in \mathcal{F}_T} \epsilon_{TF} (z_{h|T}, \sigma_F - \boldsymbol{s} \cdot \boldsymbol{n}_F)_F := \mathfrak{T}_{3,1} + \mathfrak{T}_{3,2}.$$

Observing that the function $\boldsymbol{\nabla} \cdot \boldsymbol{s} - D_h^k \boldsymbol{\sigma}_h$ has zero average inside each element, we infer that $\mathfrak{T}_{3,1} = (z_h - \overline{z}_h, \boldsymbol{\nabla} \cdot \boldsymbol{s} - D_h^k \boldsymbol{\sigma}_h)$. Hence, using the approximation properties of the L^2-orthogonal projector leads to

$$|\mathfrak{T}_{3,1}| \lesssim h^{k+2} \|\boldsymbol{s}\|_{H^{k+1}(\Omega)^d} \|z\|_{H^1(\Omega)}.$$

Using the fact that both the trace of z and the normal trace of $\boldsymbol{s}$ are single-valued at interfaces and that z vanishes on $\partial\Omega$ it is inferred that

$$\mathfrak{T}_{3,2} = \sum_{T \in \mathcal{T}_h} \sum_{F \in \mathcal{F}_T} \epsilon_{TF} (z_{h|T} - z, \sigma_F - \boldsymbol{s} \cdot \boldsymbol{n}_F)_F.$$

As a consequence, using the Cauchy–Schwarz inequality followed by the approximation properties of the L^2-orthogonal projector and Corollary 10 together with (46) yields $|\mathfrak{T}_{3,2}| \lesssim h^{k+2} \|\boldsymbol{s}\|_{H^{k+1}(\Omega)^d} \|z\|_{H^2(\Omega)}$. Hence,

$$|\mathfrak{T}_3| \leqslant |\mathfrak{T}_{3,1}| + |\mathfrak{T}_{3,2}| \lesssim h^{k+2} \|\boldsymbol{s}\|_{H^{k+1}(\Omega)^d} \|z\|_{H^2(\Omega)}. \tag{63}$$

Proceeding in a similar manner as for $\mathfrak{T}_3$ it is inferred that

$$|\mathfrak{T}_4| = \left| \sum_{T \in \mathcal{T}_h} \sum_{F \in \mathcal{F}_T} \epsilon_{TF} (\breve{u}_{h|T} - u, v_F - \boldsymbol{w} \cdot \boldsymbol{n}_F)_F \right| \lesssim h^{k+2} \|u\|_{H^2(\Omega)} \|\boldsymbol{w}\|_{H^1(\Omega)}. \tag{64}$$

The desired result is then obtained using (61)–(64) to estimate the right-hand side of (60), using (58) with $g = \widehat{u}_h - u_h$ to infer that $\|z\|_{H^2(\Omega)} + \|\boldsymbol{w}\|_{H^1(\Omega)^d} \leqslant C_{\text{ell}} \|\widehat{u}_h - u_h\|$, and simplifying by $\|\widehat{u}_h - u_h\|$. □

Remark 14 (Elliptic regularity assumption)**.** *Elliptic regularity holds if, e.g., Ω is convex. If this is not the case, since Ω is polyhedral, one still has the additional regularity $u \in H^{1+t}(\Omega)$ and $\boldsymbol{s} \in H^t(\Omega)^d$ with $1/2 < t \leqslant 1$. Proceeding as in the above proof shows that the estimate* (59) *becomes $\|\widehat{u}_h - u_h\| \leqslant Ch^{k+t+1}\|u\|_{H^{k+2}(\Omega)}$.*

4 Variations

We consider in this section several variations of the method (43) and establish links with other methods in the lowest-order case.

4.1 Virtualization

We first discuss a modification of (43) that does not require to solve (36) inside each pyramid. As for VE methods, this can be interpreted as not specifying the basis functions underlying the reconstruction. Throughout this section, it is assumed that the bilinear form H is defined by (20) with reconstruction operator given by $\mathfrak{R}_T^k = \mathfrak{C}_T^k + \mathfrak{J}_T^k$. We consider a bilinear form H^{V} on $\boldsymbol{\Sigma}_h^k \times \boldsymbol{\Sigma}_h^k$ defined from local symmetric and nonnegative bilinear forms H_T^{V} on $\boldsymbol{\Sigma}_T^k \times \boldsymbol{\Sigma}_T^k$ such that, for all $\boldsymbol{\sigma}_h, \boldsymbol{\tau}_h \in \boldsymbol{\Sigma}_h^k$,

$$H^{\mathrm{V}}(\boldsymbol{\sigma}_h, \boldsymbol{\tau}_h) = \sum_{T \in \mathcal{T}_h} H_T^{\mathrm{V}}(L_T\boldsymbol{\sigma}_h, L_T\boldsymbol{\tau}_h). \tag{65}$$

We formulate the following requirements:

(V1) *Consistency.* The following holds, for all $T \in \mathcal{T}_h$ and all $\boldsymbol{t} \in \boldsymbol{\Gamma}_T^k$, letting $\boldsymbol{\tau} := I_T^k\boldsymbol{t}$:

$$H_T^{\mathrm{V}}(\boldsymbol{\tau}, \boldsymbol{\upsilon}) = H_T(\boldsymbol{\tau}, \boldsymbol{\upsilon}) \qquad \forall \boldsymbol{\upsilon} \in \boldsymbol{\Sigma}_T^k.$$

(V2) *Stability and continuity.* There exists $\gamma > 1$ independent of h and $\mathbf{K}$ such that, denoting for all $T \in \mathcal{T}_h$ by $\|\cdot\|_{H^{\mathrm{V}},T}$ the (semi-)norm induced on $\boldsymbol{\Sigma}_T^k$ by the bilinear form H_T^{V}, the following holds:

$$\forall T \in \mathcal{T}_h, \qquad \gamma^{-1}\|\boldsymbol{\tau}\|_{H^{\mathrm{V}},T} \leqslant \|\boldsymbol{\tau}\|_{H,T} \leqslant \gamma\|\boldsymbol{\tau}\|_{H^{\mathrm{V}},T}, \qquad \forall \boldsymbol{\tau} \in \boldsymbol{\Sigma}_T^k. \tag{66}$$

This implies, in particular, denoting by $\|\cdot\|_{H^{\mathrm{V}}}$ the (semi-)norm induced by the bilinear form H^{V} on $\boldsymbol{\Sigma}_h^k$, and recalling (65),

$$\gamma^{-1}\|\boldsymbol{\tau}_h\|_{H^{\mathrm{V}}} \leqslant \|\boldsymbol{\tau}_h\|_H \leqslant \gamma\|\boldsymbol{\tau}_h\|_{H^{\mathrm{V}}}, \qquad \forall \boldsymbol{\tau}_h \in \boldsymbol{\Sigma}_h^k. \tag{67}$$

Remark 15 (Consequence of **(V1)**)**.** *Extend the bilinear forms H and H^{V} to $\check{\boldsymbol{\Sigma}}_h^k \times \check{\boldsymbol{\Sigma}}_h^k$ (with $\check{\boldsymbol{\Sigma}}_h^k$ defined by* (48)*). For all $\boldsymbol{t} \in L^2(\Omega)^d$ such that $\boldsymbol{t}_{|T} \in \boldsymbol{\Gamma}_T^k$ for all $T \in \mathcal{T}_h$, and letting $\boldsymbol{\tau}_h \in \check{\boldsymbol{\Sigma}}_h^k$ be such that its restrictions satisfy $L_T\boldsymbol{\tau}_h = I_T^k\boldsymbol{t}_{|T}$ for all $T \in \mathcal{T}_h$, we infer from* **(V1)** *that*

$$H^{\mathrm{V}}(\boldsymbol{\tau}_h, \boldsymbol{\upsilon}_h) = H(\boldsymbol{\tau}_h, \boldsymbol{\upsilon}_h) \qquad \forall \boldsymbol{\upsilon}_h \in \boldsymbol{\Sigma}_h^k. \tag{68}$$

We consider the following variation of (43): Find $(\boldsymbol{\sigma}_h^{\mathrm{V}}, u_h^{\mathrm{V}}) \in \boldsymbol{\Sigma}_h^k \times U_h^k$ such that

$$H^{\mathrm{V}}(\boldsymbol{\sigma}_h^{\mathrm{V}}, \boldsymbol{\tau}_h) + (u_h^{\mathrm{V}}, D_h^k\boldsymbol{\tau}_h) = 0 \qquad \forall \boldsymbol{\tau}_h \in \boldsymbol{\Sigma}_h^k, \tag{69a}$$

$$-(D_h^k\boldsymbol{\sigma}_h^{\mathrm{V}}, v_h) = (f, v_h) \qquad \forall v_h \in U_h^k. \tag{69b}$$

The well-posedness of problem (69) is an immediate consequence of (67) and (26a) together with (45).

Theorem 16 (Convergence rate for the virtual version)**.** *Denote by* $(\boldsymbol{s},u)\in\boldsymbol{\Sigma}\times L^2(\Omega)$ *the exact solution to* (2) *and assume the additional regularity* $u\in H_0^1(\Omega)\cap H^{k+2}(\mathcal{T}_h)$*. Then, letting* $(\boldsymbol{\sigma}_h^{\mathrm{V}},u_h^{\mathrm{V}})\in\boldsymbol{\Sigma}_h^k\times U_h^k$ *denote the unique solution to* (69) *and with* $(\widehat{\boldsymbol{\sigma}}_h,\widehat{u}_h)\in\boldsymbol{\Sigma}_h^k\times U_h^k$ *defined as in Lemma 9, the following holds with* $C>0$ *independent of* h *and* $\mathbf{K}$ *and with* $\rho_{\mathbf{K},T}$ *defined as in Theorem 11:*

$$\beta(\eta\lambda_\flat)^{1/2}\|\widehat{u}_h-u_h^{\mathrm{V}}\|+\|\widehat{\boldsymbol{\sigma}}_h-\boldsymbol{\sigma}_h^{\mathrm{V}}\|_{H^{\mathrm{V}}}\leqslant C\left(\sum_{T\in\mathcal{T}_h}\rho_{\mathbf{K},T}\lambda_{\sharp,T}h_T^{2(k+1)}\|u\|_{H^{k+2}(T)}^2\right)^{1/2}.\tag{70}$$

Proof. We sketch the proof. As a consequence of (43b) and (69b), we infer that $D_h^k(\boldsymbol{\sigma}_h-\boldsymbol{\sigma}_h^{\mathrm{V}})=0$. Hence, letting $\boldsymbol{\tau}_h=(\boldsymbol{\sigma}_h-\boldsymbol{\sigma}_h^{\mathrm{V}})$ in (43a) and in (69a), respectively, it is inferred that

$$H(\boldsymbol{\sigma}_h,\boldsymbol{\sigma}_h-\boldsymbol{\sigma}_h^{\mathrm{V}})=H^{\mathrm{V}}(\boldsymbol{\sigma}_h^{\mathrm{V}},\boldsymbol{\sigma}_h-\boldsymbol{\sigma}_h^{\mathrm{V}})=0.\tag{71}$$

We extend both bilinear forms H and H^{V} to $\breve{\boldsymbol{\Sigma}}_h^k\times\breve{\boldsymbol{\Sigma}}_h^k$, and the corresponding norms to $\breve{\boldsymbol{\Sigma}}_h^k$. Letting $\breve{\boldsymbol{\sigma}}_h\in\breve{\boldsymbol{\Sigma}}_h^k$ be defined as in the proof of Theorem 11, and using (68), we obtain

$$H^{\mathrm{V}}(\breve{\boldsymbol{\sigma}}_h,\boldsymbol{\sigma}_h-\boldsymbol{\sigma}_h^{\mathrm{V}})-H(\breve{\boldsymbol{\sigma}}_h,\boldsymbol{\sigma}_h-\boldsymbol{\sigma}_h^{\mathrm{V}})=0=H^{\mathrm{V}}(\boldsymbol{\sigma}_h^{\mathrm{V}},\boldsymbol{\sigma}_h-\boldsymbol{\sigma}_h^{\mathrm{V}}).\tag{72}$$

Now we can proceed as follows:

$$\begin{aligned}
\|\boldsymbol{\sigma}_h-\boldsymbol{\sigma}_h^{\mathrm{V}}\|_{H^{\mathrm{V}}}^2&=H^{\mathrm{V}}(\boldsymbol{\sigma}_h,\boldsymbol{\sigma}_h-\boldsymbol{\sigma}_h^{\mathrm{V}})-H^{\mathrm{V}}(\boldsymbol{\sigma}_h^{\mathrm{V}},\boldsymbol{\sigma}_h-\boldsymbol{\sigma}_h^{\mathrm{V}})\\
&=H^{\mathrm{V}}(\boldsymbol{\sigma}_h-\breve{\boldsymbol{\sigma}}_h,\boldsymbol{\sigma}_h-\boldsymbol{\sigma}_h^{\mathrm{V}})+H(\breve{\boldsymbol{\sigma}}_h,\boldsymbol{\sigma}_h-\boldsymbol{\sigma}_h^{\mathrm{V}}) && \text{eq. (72)}\\
&=H^{\mathrm{V}}(\boldsymbol{\sigma}_h-\breve{\boldsymbol{\sigma}}_h,\boldsymbol{\sigma}_h-\boldsymbol{\sigma}_h^{\mathrm{V}})+H(\breve{\boldsymbol{\sigma}}_h-\boldsymbol{\sigma}_h,\boldsymbol{\sigma}_h-\boldsymbol{\sigma}_h^{\mathrm{V}}) && \text{eq. (71)}\\
&\leqslant\|\boldsymbol{\sigma}_h-\breve{\boldsymbol{\sigma}}_h\|_{H^{\mathrm{V}}}\,\|\boldsymbol{\sigma}_h-\boldsymbol{\sigma}_h^{\mathrm{V}}\|_{H^{\mathrm{V}}}+\|\breve{\boldsymbol{\sigma}}_h-\boldsymbol{\sigma}_h\|_H\,\|\boldsymbol{\sigma}_h-\boldsymbol{\sigma}_h^{\mathrm{V}}\|_H && \text{Cauchy–Schwarz}\\
&\leqslant 2\gamma\|\boldsymbol{\sigma}_h-\breve{\boldsymbol{\sigma}}_h\|_H\,\|\boldsymbol{\sigma}_h-\boldsymbol{\sigma}_h^{\mathrm{V}}\|_{H^{\mathrm{V}}}. && \text{eq. (67)}
\end{aligned}$$

Hence, $\|\boldsymbol{\sigma}_h-\boldsymbol{\sigma}_h^{\mathrm{V}}\|_{H^{\mathrm{V}}}\leqslant 2\gamma\|\boldsymbol{\sigma}_h-\breve{\boldsymbol{\sigma}}_h\|_H$. Therefore, using the triangle inequality and again (67), it is inferred that

$$\begin{aligned}
\|\widehat{\boldsymbol{\sigma}}_h-\boldsymbol{\sigma}_h^{\mathrm{V}}\|_{H^{\mathrm{V}}}&\leqslant\|\widehat{\boldsymbol{\sigma}}_h-\boldsymbol{\sigma}_h\|_{H^{\mathrm{V}}}+\|\boldsymbol{\sigma}_h-\boldsymbol{\sigma}_h^{\mathrm{V}}\|_{H^{\mathrm{V}}}\\
&\leqslant\gamma\|\widehat{\boldsymbol{\sigma}}_h-\boldsymbol{\sigma}_h\|_H+2\gamma\|\boldsymbol{\sigma}_h-\breve{\boldsymbol{\sigma}}_h\|_H\leqslant 3\gamma\|\widehat{\boldsymbol{\sigma}}_h-\boldsymbol{\sigma}_h\|_H+2\gamma\|\widehat{\boldsymbol{\sigma}}_h-\breve{\boldsymbol{\sigma}}_h\|_H.
\end{aligned}$$

The estimate for $\|\widehat{\boldsymbol{\sigma}}_h-\boldsymbol{\sigma}_h^{\mathrm{V}}\|_{H^{\mathrm{V}}}$ in (70) follows using (49) to estimate the first term on the right-hand side and using (55) for the second. The bound for $\|\widehat{u}_h-u_h^{\mathrm{V}}\|$ can be proved as in Lemma 9. □

A supercloseness result analogous to Theorem 13 can be proved also for the virtual method (69). The details are omitted for the sake of conciseness.

Example 17. *Assume that* $\mathbf{K}$ *is isotropic, i.e.,* $\mathbf{K}_T=\lambda_T\mathrm{Id}_d$*. A possible choice for* H^{V} *is to let, for all* $T\in\mathcal{T}_h$ *and all* $\boldsymbol{\sigma},\boldsymbol{\tau}\in\boldsymbol{\Sigma}_T^k$*,*

$$H_T^{\mathrm{V}}(\boldsymbol{\sigma},\boldsymbol{\tau}):=(\mathbf{K}_T^{-1}\mathfrak{C}_T^k\boldsymbol{\sigma},\mathfrak{C}_T^k\boldsymbol{\tau})_T+J_T^{\mathrm{V}}(\boldsymbol{\sigma},\boldsymbol{\tau}),\tag{73}$$

with stabilization bilinear form

$$J_T^{\mathrm{V}}(\boldsymbol{\sigma},\boldsymbol{\tau}):=\sum_{F\in\mathcal{F}_T}\lambda_T^{-1}h_F(\mathfrak{C}_T^k\boldsymbol{\sigma}\cdot\boldsymbol{n}_F-\sigma_F,\mathfrak{C}_T^k\boldsymbol{\tau}\cdot\boldsymbol{n}_F-\tau_F)_F.$$

Proposition 18. *Assume that* $\mathbf{K}$ *is isotropic. Properties* **(V1)**–**(V2)** *hold for the bilinear form* H_T^{V} *defined by* (73).

Proof. Property **(V1)** is readily verified since the second term on the right-hand side of (73) vanishes if $\boldsymbol{\sigma} = I_T^k \boldsymbol{t}$ with $\boldsymbol{t} \in \boldsymbol{\Gamma}_T^k$. The bound $\gamma^{-1}\|\boldsymbol{\tau}\|_{H^{\mathrm{V}},T} \leqslant \|\boldsymbol{\tau}\|_{H,T}$ in (66) is an immediate consequence of (41) together with (38). Let us prove the second bound $\|\boldsymbol{\tau}\|_{H,T} \leqslant \gamma\|\boldsymbol{\tau}\|_{H^{\mathrm{V}},T}$. Using (34) where we can take any $v \in \mathbb{P}_d^k(T)$, we infer that

$$\begin{aligned} h_T\|(\boldsymbol{\nabla}\cdot\mathfrak{C}_T^k - D_T^k)\boldsymbol{\tau}\|_T &= \sup_{v\in\mathbb{P}_d^k(T),\|v\|_T=1} h_T((\boldsymbol{\nabla}\cdot\mathfrak{C}_T^k - D_T^k)\boldsymbol{\tau}, v)_T \\ &= \sup_{v\in\mathbb{P}_d^k(T),\|v\|_T=1} \sum_{F\in\mathcal{F}_T} h_T\epsilon_{TF}(\mathfrak{C}_T^k\boldsymbol{\tau}\cdot\boldsymbol{n}_F - \tau_F, v)_F \\ &\lesssim \sup_{v\in\mathbb{P}_d^k(T),\|v\|_T=1} \sum_{F\in\mathcal{F}_T} h_F^{1/2}\|\mathfrak{C}_T^k\boldsymbol{\tau}\cdot\boldsymbol{n}_F - \tau_F\|_F\|v\|_{\mathcal{P}_{TF}} \leqslant \lambda_T^{1/2} J_T^{\mathrm{V}}(\boldsymbol{\tau},\boldsymbol{\tau})^{1/2}, \end{aligned}$$

where we have used the Cauchy–Schwarz, mesh regularity, discrete trace inequalities on the pyramid $\mathcal{P}_{TF}$, and the discrete Cauchy–Schwarz inequality to conclude. Recalling (36) and using discrete trace and Poincaré inequalities, we infer that, for all $\boldsymbol{\tau} \in \boldsymbol{\Sigma}_T^k$ and all $F \in \mathcal{F}_T$,

$$\begin{aligned} \|\mathbf{K}_T^{1/2}\mathfrak{J}_{TF}^k\boldsymbol{\tau}\|_{\mathcal{P}_{TF}} &\leqslant \sup_{v\in\mathbb{P}_d^{k+1,0}(T)} \frac{((\boldsymbol{\nabla}\cdot\mathfrak{C}_T^k - D_T^k)\boldsymbol{\tau}, v)_{\mathcal{P}_{TF}}}{\|\mathbf{K}_T^{1/2}\boldsymbol{\nabla}v\|_{\mathcal{P}_{TF}}} + \sup_{v\in\mathbb{P}_d^{k+1,0}(T)} \frac{(\mathfrak{C}_T^k\boldsymbol{\tau}\cdot\boldsymbol{n}_F - \tau_F, v)_F}{\|\mathbf{K}_T^{1/2}\boldsymbol{\nabla}v\|_{\mathcal{P}_{TF}}} \\ &\lesssim \lambda_T^{-1/2}\left\{h_T\|(\boldsymbol{\nabla}\cdot\mathfrak{C}_T^k - D_T^k)\boldsymbol{\tau}\|_{\mathcal{P}_{TF}} + h_F^{1/2}\|\mathfrak{C}_T^k\boldsymbol{\tau}\cdot\boldsymbol{n}_F - \tau_F\|_F\right\}. \end{aligned}$$

Hence, the above bounds together with the discrete Cauchy–Schwarz inequality imply $\|\mathbf{K}_T^{1/2}\mathfrak{J}_T^k\boldsymbol{\tau}\|_T \lesssim J_T^{\mathrm{V}}(\boldsymbol{\tau},\boldsymbol{\tau})^{1/2}$. The desired result then follows using this bound to estimate the second term on the right-hand side of (38) and using (3). This concludes the proof of **(V2)**. □

4.2 Lowest-order variations and links with other methods

In this section we derive explicit formulæ for the reconstruction of Section 2.4 in the lowest-order case $k = 0$ and discuss a variation leading to the methods of [22, 23, 14, 6].

4.2.1 An explicit formula for the lowest-order reconstruction

Let $T \in \mathcal{T}_h$ and let $\boldsymbol{w} \in \boldsymbol{\Gamma}_T^0$ be fixed. There exists a unique $v \in \mathbb{P}_d^{1,0}(T)$ such that $\boldsymbol{w} = \mathbf{K}_T\boldsymbol{\nabla}v$ and, setting $\boldsymbol{z} := \mathbf{K}_T^{-1}\boldsymbol{w} = \boldsymbol{\nabla}v$ leads to $v(\boldsymbol{x}) = \boldsymbol{z}\cdot(\boldsymbol{x} - \overline{\boldsymbol{x}}_T)$ for all $\boldsymbol{x} \in T$ with $\overline{\boldsymbol{x}}_T$ barycenter of T since v has zero mean in T. Hence, identifying constant functions with their value on T, and denoting by $\overline{\boldsymbol{x}}_F$ the barycenter of $F \in \mathcal{F}_T$, formula (32) yields

$$|T|_d(\mathfrak{C}_T^0\boldsymbol{\tau})\cdot\boldsymbol{z} = \underbrace{-|T|_d v(\overline{\boldsymbol{x}}_T)D_T^0\boldsymbol{\tau}}_{=0} + \sum_{F\in\mathcal{F}_T} |F|_{d-1}\boldsymbol{z}\cdot(\overline{\boldsymbol{x}}_F - \overline{\boldsymbol{x}}_T)\tau_F\epsilon_{TF},$$

where $|\cdot|_d$ and $|\cdot|_{d-1}$ denote the d- and $(d-1)$-dimensional Lebesgue measures, respectively. Since the vector $\boldsymbol{z}$ is arbitrary, we infer the following explicit formula for the consistent part of the reconstruction:

$$\mathfrak{C}_T^0\boldsymbol{\tau} = \frac{1}{|T|_d}\sum_{F\in\mathcal{F}_T} |F|_{d-1}(\overline{\boldsymbol{x}}_F - \overline{\boldsymbol{x}}_T)\tau_F\epsilon_{TF}, \tag{74}$$

which is precisely [22, eq. (9)] when the barycenter is chosen as the cell center (the measure of the face appears in (74) since we have interpreted face unknowns as average rather than integral values). Similarly, for all $F \in \mathcal{F}_T$, the general form of the pyramidal residual (42) is such that, for all $\boldsymbol{\tau} \in \boldsymbol{\Sigma}_T^0$ and all $\boldsymbol{w} \in \boldsymbol{\Gamma}_{TF}^0$,

$$\mathfrak{J}_{TF}^0 \boldsymbol{\tau} = \mu \frac{d}{d_{TF}} \epsilon_{TF} (\tau_F - \mathfrak{C}_T^0 \boldsymbol{\tau} \cdot \boldsymbol{n}_F)(\overline{\boldsymbol{x}}_F - \overline{\boldsymbol{x}}_{TF}), \tag{75}$$

where $\overline{\boldsymbol{x}}_{TF}$ denotes the barycenter of $\mathcal{P}_{TF}$ and $d_{TF} = (\overline{\boldsymbol{x}}_F - \overline{\boldsymbol{x}}_T) \cdot \boldsymbol{n}_{TF}$ as defined in **(M3)**.

4.2.2 Link with Mixed Finite Volumes

We assume, for the sake of simplicity, that for all $T \in \mathcal{T}_h$, we can take $\boldsymbol{x}_T = \overline{\boldsymbol{x}}_T$, the barycenter of T, in **(M3)**. Observing that $\overline{\boldsymbol{x}}_F - \overline{\boldsymbol{x}}_{TF} = \frac{d}{d+1}(\overline{\boldsymbol{x}}_F - \overline{\boldsymbol{x}}_T)$, (75) can be rewritten as

$$\mathfrak{J}_{TF}^0 \boldsymbol{\tau} = \widetilde{\mu} \frac{d}{d_{TF}} \epsilon_{TF} (\tau_F - \mathfrak{C}_T^0 \boldsymbol{\tau} \cdot \boldsymbol{n}_F)(\overline{\boldsymbol{x}}_F - \overline{\boldsymbol{x}}_T), \tag{76}$$

with $\widetilde{\mu} = \mu \frac{d}{d+1}$. This is the "weak" flux stabilization suggested in [23, Section 2.3] for a diagonal penalty matrix. The generalization corresponding to the inner product [23, eq. (2.31)] can be interpreted as a virtualization of the above method.

4.2.3 Link with the Discrete Geometric Approach

We next consider the Discrete Geometric Approach of [14], see also [6], for which the basis functions $\{\boldsymbol{\varphi}_{TF}\}_{F \in \mathcal{F}_T}$ for the flux reconstruction are piecewise constant on the pyramidal subdivision of T and such that, for all $F, G \in \mathcal{F}_T$,

$$\boldsymbol{\varphi}_{TG|\mathcal{P}_{TF}} = \frac{(\overline{\boldsymbol{x}}_G - \overline{\boldsymbol{x}}_T)}{d_{TG}|G|_{d-1}} \delta_{FG} + \left(\frac{1}{|T|_d} \mathrm{Id}_d - \frac{(\overline{\boldsymbol{x}}_F - \overline{\boldsymbol{x}}_T) \otimes \boldsymbol{n}_{TF}}{|T|_d d_{TF}} \right) (\overline{\boldsymbol{x}}_G - \overline{\boldsymbol{x}}_T),$$

where Id_d is the identity matrix of $\mathbb{R}^{d \times d}$ and δ_{FG} is equal to 1 if $F = G$, 0 otherwise. For all $\boldsymbol{\tau} \in \boldsymbol{\Sigma}_T^k$, the corresponding flux reconstruction $\boldsymbol{t}$ is piecewise constant on the pyramidal subdivision of T and such that, for all $F \in \mathcal{F}_T$,

$$\begin{aligned}
\boldsymbol{t}_{|\mathcal{P}_{TF}} &= \sum_{G \in \mathcal{F}_T} \tau_G \epsilon_{TG} \, |G|_{d-1} \boldsymbol{\varphi}_{TG|\mathcal{P}_{TF}} \\
&= \mathfrak{C}_T^0 \boldsymbol{\tau} + \frac{1}{d_{TF}} \left\{ \tau_F \epsilon_{TF} - \left(\frac{1}{|T|_d} \sum_{G \in \mathcal{F}_T} |G|_{d-1} \, (\overline{\boldsymbol{x}}_G - \overline{\boldsymbol{x}}_T) \, \tau_G \epsilon_{TG} \right) \cdot \boldsymbol{n}_{TF} \right\} (\overline{\boldsymbol{x}}_F - \overline{\boldsymbol{x}}_T) \\
&= \mathfrak{C}_T^0 \boldsymbol{\tau} + \frac{1}{d_{TF}} (\tau_F \epsilon_{TF} - \mathfrak{C}_T^0 \boldsymbol{\tau} \cdot \boldsymbol{n}_{TF}) \, (\overline{\boldsymbol{x}}_F - \overline{\boldsymbol{x}}_T).
\end{aligned}$$

The second term on the right-hand side allows us to identify the residual. Comparing its expression with (76) shows that the two residuals coincide provided the penalty coefficient in (76) is selected as $\widetilde{\mu} = 1/d$.

4.2.4 Link with lowest-order Raviart–Thomas finite elements

Consider a simplicial mesh. For all $T \in \mathcal{T}_h$, consider the lowest-order Raviart–Thomas finite element space $\mathbb{RT}_d^0(T) = \mathbb{P}_d^0(T)^d + \boldsymbol{x}\mathbb{P}_d^0(T)$ with local basis functions $\varphi_{TF}(\boldsymbol{x}) = \frac{1}{d|T|_d}(\boldsymbol{x} - \boldsymbol{x}_{v(F)})$ for all $F \in \mathcal{F}_T$, where $v(F)$ is the vertex of T opposite to F. Define the consitent part of the reconstruction by (74), and define a residual attached to each face $F \in \mathcal{F}_T$ as follows:

$$(\mathfrak{J}_{TF}^0\boldsymbol{\tau})(\boldsymbol{x}) = |F|_{d-1}\epsilon_{TF}(\tau_F - (\mathfrak{C}_T^0\boldsymbol{\tau})\cdot\boldsymbol{n}_F)\varphi_{TF}(\boldsymbol{x}), \qquad \forall \boldsymbol{x} \in T. \tag{77}$$

Observe that the residuals are now supported on T and no longer on the pyramid $\mathcal{P}_{TF}$ attached to the corresponding F. Then, set $\mathfrak{J}_T^0 := \sum_{F\in\mathcal{F}_T}\mathfrak{J}_{TF}^0$, which matches the consistency property (22) by construction. The orthogonality property (23) follows from the fact that, for all $(\boldsymbol{\tau}, \boldsymbol{w}) \in \boldsymbol{\Sigma}_T^0 \times \boldsymbol{\Gamma}_T^0$,

$$(\mathbf{K}_T^{-1}\mathfrak{J}_T^0\boldsymbol{\tau}, \boldsymbol{w})_T = \left\{\sum_{F\in\mathcal{F}_T}|F|_{d-1}\epsilon_{TF}(\tau_F - (\mathfrak{C}_T^0\boldsymbol{\tau})\cdot\boldsymbol{n}_F)(\overline{\boldsymbol{x}}_F - \overline{\boldsymbol{x}}_T)\right\}\cdot\mathbf{K}_T^{-1}\boldsymbol{w},$$

since $\int_T \varphi_{TF} = \overline{\boldsymbol{x}}_F - \overline{\boldsymbol{x}}_T$. The term between braces is equal to

$$|T|_d\mathfrak{C}_T^0\boldsymbol{\tau} - \left\{\sum_{F\in\mathcal{F}_T}|F|_{d-1}(\overline{\boldsymbol{x}}_F - \overline{\boldsymbol{x}}_T)\otimes\boldsymbol{n}_{TF}\right\}\mathfrak{C}_T^0\boldsymbol{\tau} = 0$$

since $\sum_{F\in\mathcal{F}_T}|F|_{d-1}(\overline{\boldsymbol{x}}_F - \overline{\boldsymbol{x}}_T)\otimes\boldsymbol{n}_{TF} = |T|_d\mathrm{Id}_d$. This yields the orthogonality property. Finally, the stability property (25a) results from classical properties of the mass matrix of Raviart–Thomas finite element functions, while the continuity property (25b) is straightforward to verify.

5 Implementation and numerical example

For the sake of completeness, we discuss some implementation aspects and present a numerical example to confirm the theoretical results. An essential step in the implementation consists in selecting suitable bases for the polynomial spaces that appear in the construction. Let $T \in \mathcal{T}_h$. An appropriate choice for the basis $\mathcal{B}_T^l$ of the polynomial space $\mathbb{P}_d^l(T)$ (the values $l = k$ and $l = k+1$ are used) is, letting $A^l := \{\boldsymbol{\alpha} = (\alpha_i)_{1\leqslant i\leqslant d} \in \mathbb{R}^d \mid \|\boldsymbol{\alpha}\|_{\ell^1} \leqslant l\}$,

$$\mathcal{B}_T^l := \left\{\prod_{i=1}^d \xi_{T,i}^{\alpha_i} \;\middle|\; \boldsymbol{\alpha} \in A^l, \quad \xi_{T,i} := \frac{x_i - \overline{x}_{T,i}}{h_T} \quad \forall 1 \leqslant i \leqslant d\right\}, \tag{78}$$

i.e., the basis $\mathcal{B}_T^l$ is spanned by monomials in the translated and scaled coordinate variables $(\xi_{T,i})_{1\leqslant i\leqslant d}$. Equation (78) defines hierarchical bases, so that we can construct and evaluate $\mathcal{B}_T^{k+1}$ at quadrature nodes and obtain $\mathcal{B}_T^k$ (used to solve (10)) by simply discarding the higher-order functions. Additionally, replacing the zero-average condition with the requirement that functions vanish at $\overline{\boldsymbol{x}}_T$ (cf. Remark 4), a basis for the polynomial space $\mathbb{P}_d^{k+1,*}(T)$ used in (33) is obtained by simply discarding the constant function in $\mathcal{B}_T^{k+1}$. The choice (78) performs well for moderate polynomial degrees and isotropic elements. For higher orders or anisotropic elements, one possibility is to perform an orthonormalization as described in [1] (in this case, the zero-average condition is the appropriate one). Orthonormalization was not necessary in

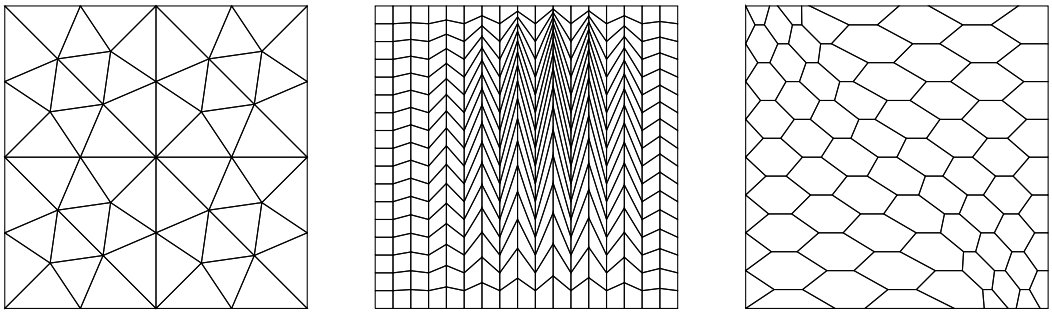


Figure 1: Triangular, skewed quadrangular, and hexagonal meshes for the numerical example of Section 5

the numerical example presented hereafter. Similarly, for the spaces $\mathbb{P}_d^k(F)$, $F \in \mathcal{F}_h$, we have chosen the basis of monomial functions that vanish at the face barycenter.

Concerning numerical integration, one possibility in the two-dimensional case is to exploit the decomposition of elements into pyramids and use standard quadrature rules inside each pyramid. In three space dimensions, this is also possible provided the faces of the elements are triangles or quadrangles yielding pyramidal subelements for which standard cubature rules are available. If this is not the case, a simplicial decomposition of the element can be considered at the price of increasing the number of quadrature nodes. Similarly, numerically integrating on the mesh faces is straightforward in two space dimensions and for elements with triangular or quadrangular faces in three space dimensions. For more general polygonal faces in three space dimensions, triangulating the face may be required.

To close this section, we present a numerical example for the homogeneous Dirichlet problem (2) on the unit square $\Omega = (0,1)^2$ with unit diagonal diffusion tensor $\mathbf{K}$ and exact solution $u = \sin(\pi x_1)\sin(\pi x_2)$. We consider the virtual bilinear form H^{V} defined by (65) and (73) and use mesh families obtained by homogeneous refinement of the meshes depicted in Figure 1. The convergence results of Figure 2 confirm the predicted orders of convergence. Further numerical experiments including the non-virtual version of the method will be presented in future work.

References

[1] F. Bassi, L. Botti, A. Colombo, D. A. Di Pietro, and P. Tesini. On the flexibility of agglomeration based physical space discontinuous Galerkin discretizations. *J. Comput. Phys.*, 231(1):45–65, 2012.

[2] M. Bebendorf. A note on the Poincaré inequality for convex domains. *Z. Anal. Anwendungen*, 22(4):751–756, 2003.

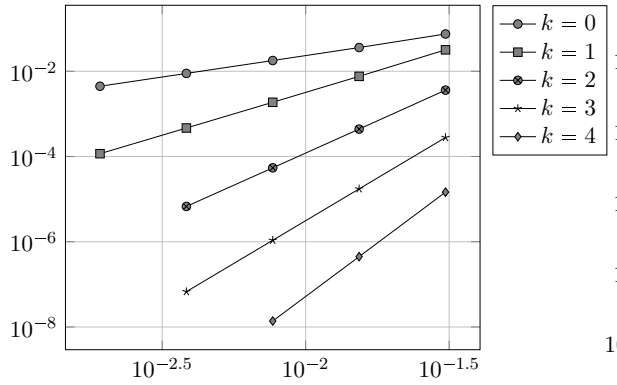
[3] L. Beirão da Veiga, F. Brezzi, A. Cangiani, G. Manzini, L. D. Marini, and A. Russo. Basic principles of virtual element methods. *M3AS Math. Models Methods Appl. Sci.*, 199(23):199–214, 2013.

[4] L. Beirão da Veiga, F. Brezzi, and L. D. Marini. Virtual elements for linear elasticity problems. *SIAM J. Numer. Anal.*, 2(51):794–812, 2013.

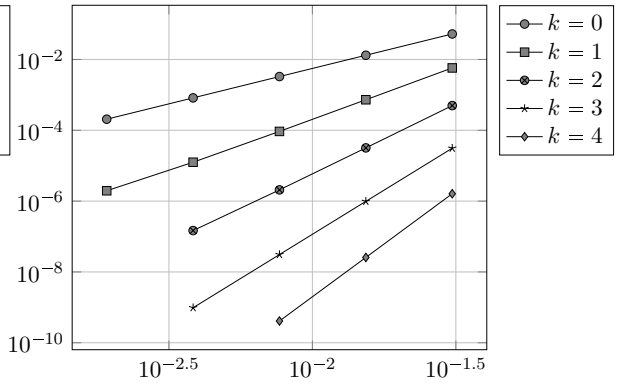
[5] L. Beirão da Veiga, K. Lipnikov, and G. Manzini. Arbitrary-order nodal mimetic discretizations of elliptic problems on general meshes. *SINUM*, 5(49):1737–1760, 2011.

[6] J. Bonelle and A. Ern. Analysis of compatible discrete operator schemes for elliptic problems on polyhedral meshes. *M2AN Math. Model. Numer. Anal.*, 2013. Accepted for publication.

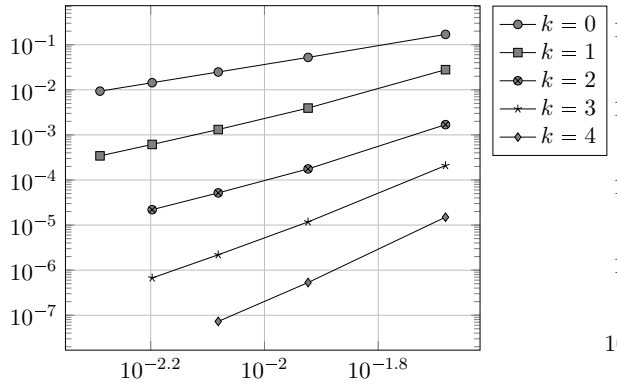
[7] A. Bossavit. On the geometry of electromagnetism. *J. Japan Soc. Appl. Electromagn. & Mech.*, 6:17–28 (no 1), 114–23 (no 2), 233–40(no 3), 318–26 (no 4), 1998.



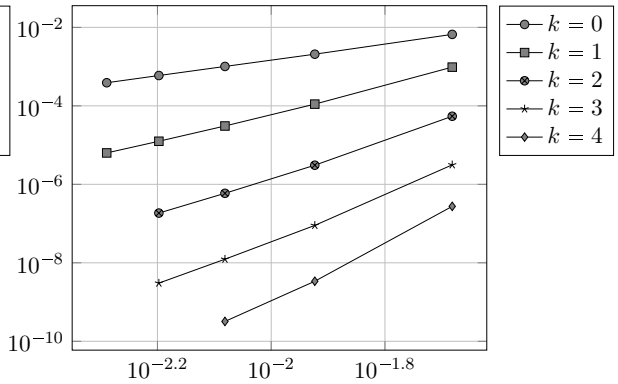
(a) $\|\widehat{\boldsymbol{\sigma}}_h - \boldsymbol{\sigma}_h\|_H$ vs. h, triangular mesh family



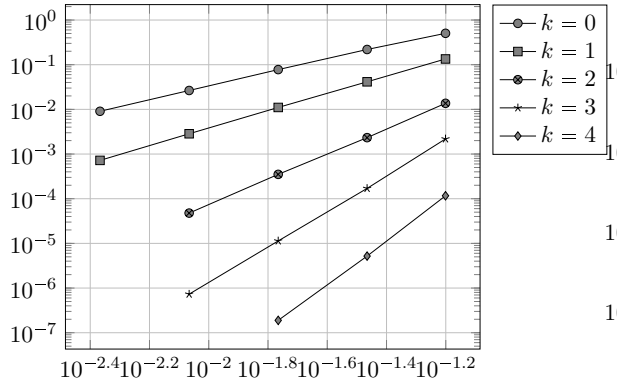
(b) $\|\widehat{u}_h - u_h\|$ vs. h, triangular mesh family



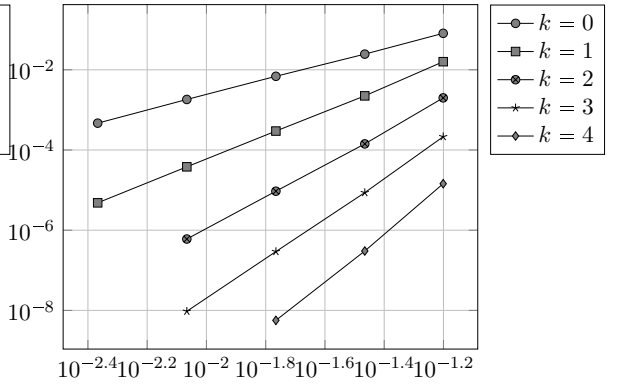
(c) $\|\widehat{\boldsymbol{\sigma}}_h - \boldsymbol{\sigma}_h\|_H$ vs. h, skewed quadrangular mesh family



(d) $\|\widehat{u}_h - u_h\|$ vs. h, skewed quadrangular mesh family



(e) $\|\widehat{\boldsymbol{\sigma}}_h - \boldsymbol{\sigma}_h\|_H$ vs. h, hexagonal mesh family



(f) $\|\widehat{u}_h - u_h\|$ vs. h, hexagonal mesh family

Figure 2: Convergence results for the numerical example of Section 5

[8] A. Bossavit. Computational electromagnetism and geometry. *J. Japan Soc. Appl. Electromagn. & Mech.*, 7–8:150–9 (no 1), 294–301 (no 2), 401–8 (no 3), 102–9 (no 4), 203–9 (no 5), 372–7 (no 6), 1999-2000.

[9] F. Brezzi, A. Buffa, and K. Lipnikov. Mimetic finite difference for elliptic problem. *M2AN Math. Model. Numer. Anal.*, 43:277–295, 2009.

[10] F. Brezzi, A. Buffa, and G. Manzini. Mimetic scalar products of discrete differential forms. Published online. DOI `http://dx.doi.org/10.1016/j.jcp.2013.08.017`, 2013.

[11] F. Brezzi and M. Fortin. *Mixed and hybrid finite element methods*, volume 15 of *Springer Series in Computational Mathematics*. Springer-Verlag, New York, 1991.

[12] 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.

[13] B. Cockburn, J. Gopalakrishnan, and R. Lazarov. Unified hybridization of discontinuous Galerkin, mixed, and continuous Galerkin methods for second order elliptic problems. *SIAM J. Numer. Anal.*, 47(2):1319–1365, 2009.

[14] L. Codecasa, R. Specogna, and F. Trevisan. A new set of basis functions for the discrete geometric approach. *J. Comput. Phys.*, 19(299):7401–7410, 2010.

[15] M. Crouzeix and P.-A. Raviart. Conforming and nonconforming finite element methods for solving the stationary Stokes equations. *RAIRO Modél. Math. Anal. Num.*, 7(3):33–75, 1973.

[16] D. A. Di Pietro. Cell centered Galerkin methods for diffusive problems. *M2AN Math. Model. Numer. Anal.*, 46(1):111–144, 2012.

[17] D. A. Di Pietro. On the conservativity of cell centered Galerkin methods. *C. R. Acad. Sci Paris, Ser. I*, 351:155–159, 2013.

[18] D. A. Di Pietro and A. Ern. *Mathematical aspects of discontinuous Galerkin methods*, volume 69 of *Mathématiques & Applications*. Springer-Verlag, Berlin, 2011.

[19] D. A. Di Pietro and S. Lemaire. An extension of the Crouzeix–Raviart space to general meshes with application to quasi-incompressible linear elasticity and Stokes flow. *Math. Comp.*, 2013. Accepted for publication. Preprint hal-00753660.

[20] J. Douglas and J. E. Roberts. Mixed finite element methods for second order elliptic problems. *Math. Appl. Comp.*, 1:91–103, 1982.

[21] J. Douglas and J. E. Roberts. Global estimates for mixed methods for second order elliptic equations. *Math. Comp.*, 44:39–52, 1985.

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

[23] J. Droniou, R. Eymard, T. Gallouët, and R. Herbin. A unified approach to mimetic finite difference, hybrid finite volume and mixed finite volume methods. *M3AS Mathematical Models and Methods in Applied Sciences*, 20(2):1–31, 2010.

[24] T. Dupont and R. Scott. Polynomial approximation of functions in Sobolev spaces. *Math. Comp.*, 34(150):441–463, 1980.

[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 J. Numer. Anal.*, 30(4):1009–1043, 2010.

[26] R. S. Falk and J. E. Osborn. Error estimates for mixed methods. *RAIRO Anal. numer.*, 14:309–324, 1980.

[27] Y. Kuznetsov, K. Lipnikov, and M. Shashkov. Mimetic finite difference method on polygonal meshes for diffusion-type problems. *Comput. Geosci.*, 8:301–324, 2004.

[28] E. Tonti. *On the formal structure of physical theories*. Istituto di Matematica del Politecnico di Milano, 1975.

[29] M. Vohralík. A posteriori error estimates for lowest-order mixed finite element discretizations of convection-diffusion-reaction equations. *SIAM J. Numer. Anal.*, 45(4):1570–1599, 2007.