HAL open science

Equivalence of local-and global-best approximations, a simple stable local commuting projector, and optimal hp approximation estimates in H(div)

Alexandre Ern, Thirupathi Gudi, Iain Smears, Martin Vohralík

To cite this version:

Alexandre Ern, Thirupathi Gudi, Iain Smears, Martin Vohralík. Equivalence of local-and global-best approximations, a simple stable local commuting projector, and optimal hp approximation estimates in H(div). 2019. hal-02268960v1

HAL Id: hal-02268960

https://inria.hal.science/hal-02268960v1

Preprint submitted on 22 Aug 2019 (v1), last revised 16 Mar 2021 (v4)

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

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

Equivalence of local- and global-best approximations, a simple stable local commuting projector, and optimal hp approximation estimates in $\boldsymbol{H}(\mathrm{div})$*

Alexandre Ern[†‡] Thirupathi Gudi[§] Iain Smears[¶] Martin Vohralík[‡†]

August 22, 2019

Abstract

Given an arbitrary function in $\boldsymbol{H}(\mathrm{div})$, we show that the error attained by the global-best approximation by $\boldsymbol{H}(\mathrm{div})$-conforming piecewise polynomial Raviart–Thomas–Nédélec elements under additional constraints on the divergence and normal flux on the boundary, is, up to a generic constant, equivalent to the sum of independent local-best approximation errors over individual mesh elements, without constraints on the divergence or normal fluxes. The generic constant only depends on the shape-regularity of the underlying simplicial mesh, the space dimension, and the polynomial degree of the approximations. The analysis also gives rise to a stable, local, commuting projector in $\boldsymbol{H}(\mathrm{div})$, delivering an approximation error that is equivalent to the local-best approximation. We next present a variant of the equivalence result, where robustness of the constant with respect to the polynomial degree is attained for unbalanced approximations. These two results together further enable us to derive rates of convergence of global-best approximations that are fully optimal in both the mesh size h and the polynomial degree p, for vector fields that only feature elementwise the minimal necessary Sobolev regularity. We finally show how to apply our findings to derive optimal a priori hp-error estimates for mixed and least-squares finite element methods applied to a model diffusion problem.

Keywords. best approximation, piecewise polynomial, localization, $\boldsymbol{H}(\mathrm{div})$ Sobolev space, Raviart–Thomas–Nédélec space, minimal regularity, optimal error bound, commuting projector, mixed finite element method, least-squares method, a priori error estimate.

1 Introduction

Interpolation operators that approximate a given function with weak gradient, curl, or divergence by a piecewise polynomial of degree p are fundamental in numerical analysis. Typically, this has to be done over a computational domain Ω covered by a mesh $\mathcal{T}$ with characteristic size h. Probably the most widespread are the canonical interpolation operators associated with the canonical degrees of freedom of the finite elements from the discrete de Rham sequence, which in particular include the Nédélec and Raviart–Thomas finite elements. The advantage of these operators is that they are local (that is, defined independently on each element K of the mesh $\mathcal{T}$) and that they commute with the appropriate differential operators. They are also projectors, i.e., they leave the interpolated function invariant if it is already a piecewise polynomial, and they lead to optimal approximation error bounds with respect to the mesh size h. However, the canonical interpolation operators have two main deficiencies. Firstly, these operators can act on a given function only if it possesses more regularity beyond the minimal H^1, $\boldsymbol{H}(\mathrm{div})$, and $\boldsymbol{H}(\mathrm{curl})$ regularity. Secondly, they are not well-suited to derive approximation error bounds that are explicit in the polynomial degree p.

[†]Université Paris-Est, CERMICS (ENPC), 77455 Marne-la-Vallée, France.
[‡]Inria, 2 rue Simone Iff, 75589 Paris, France.
Emails: alexandre.ern@enpc.fr, martin.vohralik@inria.fr.
[§]Department of Mathematics, Indian Institute of Science Bangalore, India 560012. Email: gudi@iisc.ac.in.
[¶]Department of Mathematics, University College London, London WC1E 6BT, UK. Email: i.smears@ucl.ac.uk.
*This project has received funding from the European Research Council (ERC) under the European Unions Horizon 2020 research and innovation program (grant agreement No 647134 GATIPOR).

1.1 Interpolation operators and hp-approximation

The projection-based interpolation operators, see Demkowicz and Buffa [17], Demkowicz [16], and the references therein, lead to *optimal approximation properties* in the *mesh size* h and *quasi-optimal approximation properties* in the *polynomial degree* p (up to logarithmic factors). They were derived under a conjecture of existence of commuting and polynomial-preserving extension operators from the boundary of the given element K to its interior which was later established by Demkowicz *et al.* in [18, 19, 20]; the approximation results are summarized in [20, Theorem 8.1]. Thus, these operators essentially lift the second drawback of the canonical interpolation operators described above (up to logarithmic factors), while still sharing the same important properties, i.e., they are defined locally, they are projectors, and they commute with the appropriate differential operators. However, these operators again require more regularity beyond the minimal H^1, $\boldsymbol{H}(\mathrm{div})$, and $\boldsymbol{H}(\mathrm{curl})$ regularity, so that the first drawback remains.

In the particular case of $\boldsymbol{H}(\mathrm{div})$, which constitutes the focus of the present work, the normal component of the interpolate on each mesh face is fully dictated by the normal component of the interpolated function on that face, which requires $\boldsymbol{H}^s(\mathrm{div})$ regularity with $s > 0$, which is slightly more than $\boldsymbol{H}(\mathrm{div})$ regularity. Some further refinements can be found in Bespalov and Heuer [6] in the $\boldsymbol{H}(\mathrm{div})$ case. Recently, building on [17, 16], a commuting projector that fully removes the second drawback above in that it has fully *optimal p-approximation properties* (does not feature the logarithmic factors) has been devised in [34]. To define the projector, though, higher regularity is needed, with in particular $\boldsymbol{H}^s(\mathrm{div})$, $s \geq 1$, in the case of interest here.

The issue of constructing (quasi-)interpolation projectors under the minimal regularities H^1, $\boldsymbol{H}(\mathrm{div})$, and $\boldsymbol{H}(\mathrm{curl})$ has been addressed before, cf., e.g., Clément [14], Scott and Zhang [42], and Bernardi and Girault [4] in the H^1 case, Nochetto and Stamm [36] in the $\boldsymbol{H}(\mathrm{div})$ case, and Bernardi and Hecht [5] in the $\boldsymbol{H}(\mathrm{curl})$ case; see also the references therein. Stability and h-optimal approximation estimates in any L^p-norm, $1 \leq p \leq \infty$, has recently been achieved by Ern and Guermond in [24] in a unified setting for a wide range of finite elements encompassing the whole discrete de Rham sequence. The arguments used in [24] are somewhat different from those in the previous references: a projection onto the fully discontinuous (broken) piecewise polynomial space is applied first, followed by an averaging operator to ensure the appropriate H^1, $\boldsymbol{H}(\mathrm{div})$, or $\boldsymbol{H}(\mathrm{curl})$ trace continuity. Unfortunately, all of the above quasi-interpolation projectors do not commute with the appropriate differential operators and, moreover, they are only shown to be optimal in h but not in p.

1.2 Stable local commuting projectors under minimal regularity

Constructing projectors applicable under the *minimal regularities* H^1, $\boldsymbol{H}(\mathrm{div})$, and $\boldsymbol{H}(\mathrm{curl})$ that would in addition be *commuting*, *stable*, and *locally defined* represents a long-standing effort. Stability, commutativity, and the projection property were obtained by Christiansen and Winther in [13] by composing the canonical interpolation operators with mollification, following some earlier ideas in particular from Schöberl [40, 41], cf. also Ern and Guermond [23] for a shrinking technique avoiding the need of extensions outside of the domain and Licht [32] for essential boundary conditions only prescribed on the part of the boundary of Ω. These operators are, however, not locally defined. This last remaining issue was finally remedied in [29], where a patch-based construction resembling that of the Clément operator [14] is introduced. However, no approximation properties are discussed, and stability is achieved only in the graph space of the appropriate differential operator, e.g., $\boldsymbol{H}(\mathrm{div})$ but not in $\boldsymbol{L}^2$ for the case of interest here.

1.3 Equivalence of local-best and global-best approximations

Finally, in a seemingly rather unconnected recent result, Veeser [43] showed that the error in the *best approximation* of a given scalar-valued function in H^1 by *continuous* piecewise polynomials is *equivalent* up to a generic constant to that by *discontinuous* piecewise polynomials. This result is termed equivalence of global- and local-best approximations. A predecessor result in the lowest-order case $p = 1$ and up to data oscillation can be easily deduced from Carstensen *et al.* [12, Theorem 2.1 and inequalities (3.2), (3.5), and (3.6)], see also the references therein; equivalences between approximations by different numerical methods are studied in [12]. A similar result is also given in Aurada *et al.* [1, Proposition 3.1], and an improvement of the dependence of the equivalence constant on the polynomial degree in two space dimensions is developed in [11, Theorem 4]. This equivalence result might be surprising at a first glance, since the local-best error is clearly smaller than the global-best one. The twist comes from the fact that the function to be approximated is continuous in the sense of traces because of its H^1-regularity, so one does

not gain in approximating it by discontinuous piecewise polynomials. For finite element discretizations of coercive problems, this result in particular allows one to obtain a priori error estimates without the passage through the Bramble–Hilbert lemma. Another important application is for approximation classes in the theory of a-posteriori-based convergence and optimality [43].

1.4 The present manuscript

Let $\Omega \subset \mathbb{R}^d$, $d \in \{2,3\}$, be a Lipschitz polygon or polyhedron with boundary Γ. Let $\Gamma_{\mathrm{N}} \subset \Gamma$ be a (possibly empty) subset and let $\boldsymbol{H}_{0,\Gamma_{\mathrm{N}}}(\mathrm{div},\Omega) := \{\boldsymbol{v} \in \boldsymbol{H}(\mathrm{div},\Omega),\ \boldsymbol{v}\cdot\boldsymbol{n} = 0 \text{ on } \Gamma_{\mathrm{N}}\}$; precise definitions are given below. Let $\mathcal{T}$ be a simplicial and shape-regular (possibly locally refined) mesh of Ω with elements K of diameter h_K, with $h := \max_{K\in\mathcal{T}} h_K$, on which we consider Raviart–Thomas–Nédélec (RTN) piecewise polynomials of order $p \geq 0$. Variable polynomial degrees can be taken into account by proceeding as in, e.g., [22]. We avoid it here for the sake of clarity of exposition. Our main results can be divided into four groups.

1) Equivalence of local- and global-best approximations in $\boldsymbol{H}(\mathrm{div})$ under minimal regularity

Consider the best-approximation error $E_{\mathcal{T},p}(\boldsymbol{v})$, in a dimensionally consistent weighted $\boldsymbol{H}(\mathrm{div})$-norm defined in (3.1) below, of an arbitrary function $\boldsymbol{v} \in \boldsymbol{H}_{0,\Gamma_{\mathrm{N}}}(\mathrm{div},\Omega)$ by piecewise RTN functions that are $\boldsymbol{H}_{0,\Gamma_{\mathrm{N}}}(\mathrm{div},\Omega)$-conforming and satisfy a constraint on the divergence. Because of the $\boldsymbol{H}_{0,\Gamma_{\mathrm{N}}}(\mathrm{div},\Omega)$-conformity (normal trace continuity) constraint, this is a *global* minimization problem. In our first main result, Theorem 3.1, we show that $E_{\mathcal{T},p}(\boldsymbol{v})$ is, up to a generic constant, *equivalent* to the *local*-best approximation errors $e_{K,p}(\boldsymbol{v})$ defined by elementwise minimizations *without any constraint* on the inter-element continuity of the normal trace or on the divergence. We stress that $e_{K,p}(\boldsymbol{v})$ only depends on $\boldsymbol{v}|_K$ on each mesh element $K \in \mathcal{T}$; we simply write $\boldsymbol{v}$ to alleviate the notation. The generic constant entering the equivalence result only depends on the shape-regularity of the simplicial mesh $\mathcal{T}$, the space dimension d, and the polynomial degree p. This extends the results of [1, 11, 12, 43] to the $\boldsymbol{H}(\mathrm{div})$ case. A variant of this result, Theorem 3.4, bounds, up to a *polynomial-degree robust constant*, the global-best error $E_{\mathcal{T},p}(\boldsymbol{v})$ by the local-best approximation errors $e_{K,p-1}(\boldsymbol{v})$ of one degree lower.

Our main tool to achieve the above results is the *equilibrated flux reconstruction*. This allows us to transform locally (on patches of elements) a discontinuous piecewise polynomial with a suitable elementwise divergence constraint into a normal-trace-continuous piecewise polynomial with the expected elementwise divergence constraint. This has been traditionally used in *a posteriori* error analysis of *primal* finite element methods derived from H^1-formulations, see Destuynder and Métivet [21], Luce and Wohlmuth [33], Braess and Schöberl [8], Ern and Vohralík [25, 26], Becker *et al.* [3], and the references therein. We now employ it here in the context of *a priori* error analysis of *dual* approximations in $\boldsymbol{H}(\mathrm{div})$. In order to prove Theorem 3.1, we combine the equilibrated flux reconstruction with the bubble function technique of Verfürth [44] (see also Gudi [30] for using this tool in a priori error analysis). It is this step that leads to the dependence of the generic constant in Theorem 3.1 on the polynomial degree p. As for Theorem 3.4 that leads to a p-robust bound, as in several references above, we rather rely on the p-robust stability of the right inverse of the divergence operator of Costabel and McIntosh [15, Corollary 3.4] on each single element. We then combine this result with the p-robust patch stability bounds of Braess *et al.* [9, Theorem 7] in two space dimensions and Ern and Vohralík [27, Corollaries 3.3 and 3.8] in three space dimensions, which are themselves based on the stable polynomial extension operators of Demkowicz *et al.* [20, Theorem 7.1].

2) A simple stable local commuting projector defined under the minimal $\boldsymbol{H}(\mathrm{div})$ regularity

It turns out that the methodology we develop in point 1) above in order to prove the local-best global-best equivalence results stated in Theorem 3.1 immediately leads to the definition of a projector that is defined over the entire $\boldsymbol{H}_{0,\Gamma_{\mathrm{N}}}(\mathrm{div},\Omega)$ space without any additional regularity, enjoys a commuting property with the divergence operator, is locally defined over patches of elements, and is stable in $\boldsymbol{L}^2$ for all functions $\boldsymbol{v} \in \boldsymbol{H}_{0,\Gamma_{\mathrm{N}}}(\mathrm{div},\Omega)$ with piecewise p-polynomial divergence. This operator is also stable in $\boldsymbol{L}^2$ up to a hp data oscillation term for the divergence for all functions $\boldsymbol{v} \in \boldsymbol{H}_{0,\Gamma_{\mathrm{N}}}(\mathrm{div},\Omega)$, and it is stable in $\boldsymbol{H}_{0,\Gamma_{\mathrm{N}}}(\mathrm{div},\Omega)$ for all functions in that space.

This may be compared to the result of [29], where stability in $\boldsymbol{H}(\mathrm{div})$ is achieved but not in $\boldsymbol{L}^2$. Moreover, our projector has a very simple construction, with elementwise local-best approximations combined patch by patch to the final projector via the flux equilibration technique. The essential (no-flux) boundary condition on a part of the computational domain only is here taken into account without

any difficulty. All these results are summarized in Definition 3.7 and Theorem 3.8 and improve on [13] in that they are local, on [29] in that they lead to the above-described $\boldsymbol{L}^2$-stability up to data oscillation, and on [24] in that they satisfy a commuting property with the divergence operator.

3) Optimal *hp*-approximation estimates in $\boldsymbol{H}(\mathrm{div})$

Our third main result is Theorem 3.5 where we derive *hp-approximation estimates*. These estimates feature the following four properties: i) they request *no global regularity* of the approximated function $\boldsymbol{v}$ beyond $\boldsymbol{H}_{0,\Gamma_\mathrm{N}}(\mathrm{div},\Omega)$; ii) only the *minimal local regularity* is needed, that is, $\boldsymbol{v}|_K \in \boldsymbol{H}^s(K)$ for all $K \in \mathcal{T}$, for any $s \geq 0$; iii) the convergence rates are *fully optimal* in both the *mesh-size* h and the *polynomial degree* p, in particular featuring no logarithmic factor of the polynomial degree p; iv) *no higher order norms of the divergence* of $\boldsymbol{v}$ appear in the bound whenever $s \geq 1$. This improves on [17, 16] in removing the suboptimality with respect to the polynomial degree, on [17, 16, 34] in reducing the regularity requirements, and on approximations using Clément-type operators in removing the need for regularity assumptions over the (overlapping) elemental patches while reducing it instead to (nonoverlapping) elements. The proof of these fully optimal *hp*-approximation estimates relies on the elementwise local-best approximation errors described in point 1) and crucially combines Theorems 3.1 and 3.4.

4) Applications to mixed finite element and least-squares mixed finite element methods

We finally showcase how the above results can be turned into fully optimal *hp-a priori error estimates* for two popular classes of numerical methods for second-order elliptic partial differential equations. In mixed finite element methods, c.f. the original contributions of Raviart and Thomas [38] and Nédélec [35], or the textbook Boffi *et al.* [7], the error $\|\boldsymbol{\sigma} - \boldsymbol{\sigma}_\mathrm{M}\|$ between the exact flux $\boldsymbol{\sigma}$ and its mixed approximation $\boldsymbol{\sigma}_\mathrm{M}$ immediately takes the form of the $\boldsymbol{L}^2$-norm term in the constrained global-best approximation error of Theorems 3.1 and 3.4 (cf. Lemma 5.1), so the application is immediate. For the family of least-squares mixed finite element methods, see Pehlivanov *et al.* [37], Cai and Ku [10], Ku [31], and the references therein, the application is a little less immediate, and for completeness we establish it in Lemmas 5.2 and 5.3. These results allow us in particular to circumvent the typical use of interpolation or quasi-interpolation operators to obtain error estimates that hinge upon increased regularity assumptions. Note also that an immediate application of the above commuting projector in the context of mixed finite elements is the construction of a Fortin operator under the minimal $\boldsymbol{H}(\mathrm{div})$ regularity.

The rest of the manuscript is organized as follows. In Section 2, we introduce the discrete setting and the main notation. In Section 3, we state our main results, namely Theorems 3.1 and 3.4 stating the relations between the local- and global-best approximations, Theorem 3.5 stating the optimal *hp*-approximation estimates, and Theorem 3.8 about the simple stable local commuting projector. We prove these results in Section 4. Finally, we present an application of our main results to the a priori error analysis of mixed finite element and least-squares mixed finite methods in Section 5.

2 Discrete setting and notation

Let $\Omega \subset \mathbb{R}^d$ for $d \in \{2,3\}$ be an open, bounded, connected polygon or polyhedron with Lipschitz boundary Γ. Let $\mathcal{T}$ be a given conforming, simplicial, possibly locally refined mesh of Ω, i.e. $\overline{\Omega} = \cup_{K\in\mathcal{T}} K$, where any K is a closed simplex and the intersection of two different simplices is either an empty set or their common vertex, edge, or face. Let Γ_D be a (possibly empty) closed subset of Γ, and let $\Gamma_\mathrm{N} := \Gamma \setminus \Gamma_\mathrm{D}$ be its (relatively open) complement in Γ, with the assumption that $\mathcal{T}$ matches Γ_D and Γ_N in the sense that every boundary face of the mesh $\mathcal{T}$ is fully contained either in Γ_D or in $\overline{\Gamma_\mathrm{N}}$. Let $\boldsymbol{L}^2(\Omega) := L^2(\Omega;\mathbb{R}^d)$, and $\boldsymbol{H}(\mathrm{div},\Omega) := \{\boldsymbol{v} \in \boldsymbol{L}^2(\Omega), \nabla\cdot\boldsymbol{v} \in L^2(\Omega)\}$. Furthermore, we define the space $\boldsymbol{H}_{0,\Gamma_\mathrm{N}}(\mathrm{div},\Omega) := \{\boldsymbol{v} \in \boldsymbol{H}(\mathrm{div},\Omega), \boldsymbol{v}\cdot\boldsymbol{n} = 0 \text{ on } \Gamma_\mathrm{N}\}$, where $\boldsymbol{v}\cdot\boldsymbol{n} = 0$ on Γ_N means that $\langle \boldsymbol{v}\cdot\boldsymbol{n}, \varphi\rangle_\Gamma = 0$ for all functions $\varphi \in H^1(\Omega)$ that have vanishing trace on Γ_D.

For an open subset $\omega \subset \Omega$, let $\boldsymbol{L}^2(\omega) := L^2(\omega;\mathbb{R}^d)$ and $\boldsymbol{H}(\mathrm{div},\omega) := \{\boldsymbol{v} \in \boldsymbol{L}^2(\omega), \nabla\cdot\boldsymbol{v} \in L^2(\omega)\}$. We also denote by $(\cdot,\cdot)_\omega$ and $\|\cdot\|_\omega$ the L^2-inner product and norm for scalar- or vector-valued functions on ω. In the special case where $\omega = \Omega$, we drop the subscript, i.e. $(\cdot,\cdot) := (\cdot,\cdot)_\Omega$ and $\|\cdot\| := \|\cdot\|_\Omega$. The diameter of ω is denoted by h_ω, and its outward unit normal as $\boldsymbol{n}_\omega$.

For any mesh element $K \in \mathcal{T}$, its diameter is denoted by h_K, and we set $h := \max_{K\in\mathcal{T}} h_K$. Let $\mathcal{V}_\Omega$ denote the set of interior vertices of $\mathcal{T}$, i.e. the vertices contained in Ω. Let $\mathcal{V}_\Gamma$ denote the set of vertices of $\mathcal{T}$ on the boundary Γ, and set $\mathcal{V} := \mathcal{V}_\Omega \cup \mathcal{V}_\Gamma$. We divide $\mathcal{V}_\Gamma$ into two disjoint sets $\mathcal{V}_\mathrm{D}$ and

$\mathcal{V}_{\rm N}$, where $\mathcal{V}_{\rm D}$ contains all vertices in $\Gamma_{\rm D}$ (recalling that $\Gamma_{\rm D}$ is assumed to be closed) and $\mathcal{V}_{\rm N}$ consists of all vertices in $\Gamma_{\rm N}$. For each vertex $\boldsymbol{a} \in \mathcal{V}$, define the patch $\mathcal{T}_{\boldsymbol{a}} := \{K \in \mathcal{T},\, \boldsymbol{a} \text{ is a vertex of } K\}$ and the corresponding open subdomain $\omega_{\boldsymbol{a}} := \{\cup_{K\in\mathcal{T}_{\boldsymbol{a}}} K\}^\circ$. The piecewise affine Lagrange finite element basis function associated with a vertex $\boldsymbol{a} \in \mathcal{V}$ is denoted by $\psi_{\boldsymbol{a}}$. Let $\mathcal{F}$ denote the set of all $(d-1)$-dimensional faces of $\mathcal{T}$. By convention, we consider faces to be closed sets. Let $\mathcal{F}_\Omega$ stand for the set of interior faces of the mesh, i.e. those faces in $\mathcal{F}$ that are intersections of the boundaries of two mesh elements. Let $\mathcal{F}_{\rm D}$, respectively $\mathcal{F}_{\rm N}$, denote the set of all faces in $\mathcal{F}$ that are contained in $\Gamma_{\rm D}$, respectively $\overline{\Gamma_{\rm N}}$, and let $\mathcal{F}_\Gamma := \mathcal{F}_{\rm D} \cup \mathcal{F}_{\rm N}$ be the set of all boundary faces. The assumption that $\mathcal{T}$ is a conforming mesh that matches the boundary components $\Gamma_{\rm D}$ and $\Gamma_{\rm N}$ implies that $\Gamma_{\rm D} = \cup_{F\in\mathcal{F}_{\rm D}} F$ and $\overline{\Gamma_{\rm N}} = \cup_{F\in\mathcal{F}_{\rm N}} F$. For an element $K \in \mathcal{T}$, we denote the set of all faces of K by $\mathcal{F}_{\rm K}$, and the set of all vertices of K by $\mathcal{V}_K$. For each interior vertex $\boldsymbol{a} \in \mathcal{V}_\Omega$, we let $\mathcal{F}^{\rm in}_{\boldsymbol{a}}$ denote the set of all faces that contain the vertex $\boldsymbol{a}$ (and thus do not lie on the boundary of $\omega_{\boldsymbol{a}}$). For boundary vertices $\boldsymbol{a} \in \mathcal{V}_\Gamma$, let $\mathcal{F}^{\rm in}_{\boldsymbol{a}}$ collect the faces that contain the vertex $\boldsymbol{a}$ but do not lie on the Dirichlet boundary Γ_D. The mesh shape-regularity parameter is defined as $\kappa_{\mathcal{T}} := \max_{K\in\mathcal{T}} h_K/\varrho_K$, where ϱ_K is the diameter of the largest ball inscribed in K.

Let $p \geq 0$ be a nonnegative integer. For $S \in \{K, F\}$, where $K \in \mathcal{T}$ is an element and $F \in \mathcal{F}$ is a face, we define $\mathcal{P}_p(S)$ as the space of all polynomials of total degree at most p on S. If $\widetilde{\mathcal{T}}$ denotes a subset of elements of $\mathcal{T}$, $\mathcal{P}_p(\widetilde{\mathcal{T}}) := \{v_h \in L^2(\Omega),\, v_h|_K \in \mathcal{P}_p(K)\ \forall K \in \widetilde{\mathcal{T}}\}$ is the space of piecewise polynomials of degree at most p over $\widetilde{\mathcal{T}}$. Typically, $\widetilde{\mathcal{T}}$ will be either the whole mesh $\mathcal{T}$ or the patch $\mathcal{T}_{\boldsymbol{a}}$ as defined above. We define the piecewise Raviart–Thomas–Nédélec space $\boldsymbol{RTN}_p(\mathcal{T}) := \{\boldsymbol{v}_{\mathcal{T}} \in \boldsymbol{L}^2(\Omega),\ \boldsymbol{v}_{\mathcal{T}}|_K \in \boldsymbol{RTN}_p(K)\}$, where $\boldsymbol{RTN}_p(K) := \mathcal{P}_p(K;\mathbb{R}^d) + \boldsymbol{x}\mathcal{P}_p(K)$ and $\mathcal{P}_p(K;\mathbb{R}^d)$ denote the space of $\mathbb{R}^d$-valued functions defined on K with each component being a polynomial of degree at most p in $\mathcal{P}_p(K)$. Note that with this choice of notation, functions in the space $\boldsymbol{RTN}_p(\mathcal{T})$ do not necessarily belong to $\boldsymbol{H}(\mathrm{div},\Omega)$; thus, $\boldsymbol{RTN}_p(\mathcal{T}) \cap \boldsymbol{H}(\mathrm{div},\Omega)$ is a proper subspace of $\boldsymbol{RTN}_p(\mathcal{T})$ which is classically characterized as those functions in $\boldsymbol{RTN}_p(\mathcal{T})$ having a continuous normal component across interior faces $F \in \mathcal{F}_\Omega$. Moreover, $\boldsymbol{RTN}_p(\mathcal{T})$ is a subspace of $\boldsymbol{C}^1(\mathcal{T}) := \{\boldsymbol{v} \in \boldsymbol{L}^2(\Omega),\ \boldsymbol{v}|_K \in \boldsymbol{C}^1(K) \text{ for all } K \in \mathcal{T}\}$, the space of piecewise (broken) first-order component-wise differentiable vector-valued fields over $\mathcal{T}$. To avoid confusion between piecewise smooth and globally smooth functions, we denote the elementwise gradient and the elementwise divergence by $\nabla_{\mathcal{T}}$ and by $\nabla_{\mathcal{T}}\cdot$, respectively.

If $F \in \mathcal{F}_\Omega$ is an interior face, then there exist two mesh elements K_+ and K_- in $\mathcal{T}$ such that $F = \partial K_+ \cap \partial K_-$. We then let $\boldsymbol{n}_F = \boldsymbol{n}_{K_-} = -\boldsymbol{n}_{K_+}$ be the unit normal to F that points outward K_- and inward K_+. We define for any $\boldsymbol{v}_{\mathcal{T}} \in \boldsymbol{RTN}_p(\mathcal{T})$ the jump $[\![\boldsymbol{v}_{\mathcal{T}}]\!]_F$ and mean $\{\!\{\boldsymbol{v}_{\mathcal{T}}\}\!\}_F$ on F by

$$[\![\boldsymbol{v}_{\mathcal{T}}]\!]_F := \boldsymbol{v}_{\mathcal{T}}|_{K_+} - \boldsymbol{v}_{\mathcal{T}}|_{K_-}, \quad \{\!\{\boldsymbol{v}_{\mathcal{T}}\}\!\}_F := \frac{1}{2}\left(\boldsymbol{v}_{\mathcal{T}}|_{K_+} + \boldsymbol{v}_{\mathcal{T}}|_{K_-}\right). \tag{2.1}$$

If $F \in \mathcal{F}_\Gamma$ is a boundary face, then, for notational convenience, we define $[\![\boldsymbol{v}_{\mathcal{T}}]\!]_F := \boldsymbol{v}_{\mathcal{T}}|_F$ as well as $\{\!\{\boldsymbol{v}_{\mathcal{T}}\}\!\}_F := \boldsymbol{v}_{\mathcal{T}}|_F$.

For each polynomial degree $p \geq 0$, let $\Pi^p_{\mathcal{T}}\colon L^2(\Omega) \to \mathcal{P}_p(\mathcal{T})$ denote the L^2-orthogonal projection of order p. Similarly, let Π^p_F denote the L^2-orthogonal projection of order p on a face $F \in \mathcal{F}$, which maps $L^2(F)$ to $\mathcal{P}_p(F)$. Let $\boldsymbol{I}^p_{\mathcal{T}}\colon \boldsymbol{C}^1(\mathcal{T}) \to \boldsymbol{RTN}_p(\mathcal{T})$ be the piecewise Raviart–Thomas–Nédélec interpolant. The domain of $\boldsymbol{I}^p_{\mathcal{T}}$ can be taken (much) larger than $\boldsymbol{C}^1(\mathcal{T})$, but not as large as piecewise $\boldsymbol{H}(\mathrm{div})$ fields; the present choice is sufficient for our purposes. For any $\boldsymbol{v} \in \boldsymbol{C}^1(\mathcal{T})$, the interpolant $\boldsymbol{I}^p_{\mathcal{T}}\boldsymbol{v}$ is defined separately on each element $K \in \mathcal{T}$ by the conditions

$$\begin{aligned} ((\boldsymbol{I}^p_{\mathcal{T}}\boldsymbol{v})|_K\cdot\boldsymbol{n}_K, q_K)_F &= (\boldsymbol{v}|_K\cdot\boldsymbol{n}_K, q_K)_F \quad && \forall q_K \in \mathcal{P}_p(F),\ \forall F \in \mathcal{F}_K, \\ (\boldsymbol{I}^p_{\mathcal{T}}\boldsymbol{v}, \boldsymbol{r}_K)_K &= (\boldsymbol{v}, \boldsymbol{r}_K)_K && \forall \boldsymbol{r}_K \in \mathcal{P}_{p-1}(K;\mathbb{R}^d), \end{aligned} \tag{2.2}$$

where $\boldsymbol{v}|_K\cdot\boldsymbol{n}_K$ denotes the normal trace of $\boldsymbol{v}|_K$, the restriction of $\boldsymbol{v}$ to K. Note that (2.2) implies that $((\boldsymbol{I}^p_{\mathcal{T}}\boldsymbol{v})|_K\cdot\boldsymbol{n}_K)|_F = \Pi^p_F((\boldsymbol{v}|_K\cdot\boldsymbol{n}_K)|_F)$ for all faces $F \subset \mathcal{F}_{\rm K}$. A useful property of the operator $\boldsymbol{I}^p_{\mathcal{T}}$ is the commuting identity:

$$\nabla_{\mathcal{T}}\cdot(\boldsymbol{I}^p_{\mathcal{T}}\boldsymbol{v}) = \Pi^p_{\mathcal{T}}(\nabla_{\mathcal{T}}\cdot\boldsymbol{v}) \quad \forall \boldsymbol{v} \in \boldsymbol{C}^1(\mathcal{T}). \tag{2.3}$$

In the spirit of Braess *et al.* [9] and [25, 26, 28], we define the local mixed finite element spaces $\boldsymbol{V}_p(\omega_{\boldsymbol{a}})$ by

$$\boldsymbol{V}_p(\omega_{\boldsymbol{a}}) := \begin{cases} \{\boldsymbol{v}_{\boldsymbol{a}} \in \boldsymbol{RTN}_p(\mathcal{T}_{\boldsymbol{a}}) \cap \boldsymbol{H}(\mathrm{div};\omega_{\boldsymbol{a}}),\ \boldsymbol{v}_{\boldsymbol{a}}\cdot\boldsymbol{n}_{\omega_{\boldsymbol{a}}} = 0 \text{ on } \partial\omega_{\boldsymbol{a}}\} & \text{if } \boldsymbol{a} \in \mathcal{V}_\Omega \cup \mathcal{V}_{\rm N}, \\ \{\boldsymbol{v}_{\boldsymbol{a}} \in \boldsymbol{RTN}_p(\mathcal{T}_{\boldsymbol{a}}) \cap \boldsymbol{H}(\mathrm{div};\omega_{\boldsymbol{a}}),\ \boldsymbol{v}_{\boldsymbol{a}}\cdot\boldsymbol{n}_{\omega_{\boldsymbol{a}}} = 0 \text{ on } \partial\omega_{\boldsymbol{a}} \setminus \Gamma^{\boldsymbol{a}}_{\rm D}\} & \text{if } \boldsymbol{a} \in \mathcal{V}_{\rm D}, \end{cases} \tag{2.4}$$

where $\Gamma^{\boldsymbol{a}}_{\rm D}$ contains those boundary faces from $\Gamma_{\rm D}$ that share the vertex $\boldsymbol{a}$. In particular, we observe that when $\partial\omega_{\boldsymbol{a}} \cap \Gamma_{\rm N} \neq \emptyset$, then $\boldsymbol{v}_{\boldsymbol{a}}\cdot\boldsymbol{n} = 0$ on $\Gamma_{\rm N}$ for any $\boldsymbol{v}_{\boldsymbol{a}} \in \boldsymbol{V}_p(\omega_{\boldsymbol{a}})$. As a result of the above definitions, it follows that the zero extension to all of Ω of any $\boldsymbol{v}_{\boldsymbol{a}} \in \boldsymbol{V}_p(\omega_{\boldsymbol{a}})$ belongs to $\boldsymbol{RTN}_p(\mathcal{T}) \cap \boldsymbol{H}_{0,\Gamma_{\rm N}}(\mathrm{div},\Omega)$.

3 Main results

3.1 Equivalence of local- and global-best approximations in $\boldsymbol{H}_{0,\Gamma_{\mathrm{N}}}(\mathbf{div},\Omega)$

For any function $\boldsymbol{v} \in \boldsymbol{H}_{0,\Gamma_{\mathrm{N}}}(\mathrm{div},\Omega)$, we consider the global-best approximation error $E_{\mathcal{T},p}(\boldsymbol{v})$ defined as the best approximation, in a weighted norm, from $\boldsymbol{RTN}_p(\mathcal{T}) \cap \boldsymbol{H}_{0,\Gamma_{\mathrm{N}}}(\mathrm{div},\Omega)$, subject to a constraint on the divergence:

$$[E_{\mathcal{T},p}(\boldsymbol{v})]^2 := \min_{\substack{\boldsymbol{v}_{\mathcal{T}} \in \boldsymbol{RTN}_p(\mathcal{T}) \cap \boldsymbol{H}_{0,\Gamma_{\mathrm{N}}}(\mathrm{div},\Omega) \\ \nabla\cdot\boldsymbol{v}_{\mathcal{T}} = \Pi^p_{\mathcal{T}}(\nabla\cdot\boldsymbol{v})}} \|\boldsymbol{v} - \boldsymbol{v}_{\mathcal{T}}\|^2_{\Omega} + \sum_{K\in\mathcal{T}} \left[\frac{h_K}{p+1}\|\nabla\cdot\boldsymbol{v} - \Pi^p_{\mathcal{T}}(\nabla\cdot\boldsymbol{v})\|_K\right]^2. \tag{3.1}$$

We further consider the local approximation errors defined on each element $K \in \mathcal{T}$ by

$$[e_{K,p}(\boldsymbol{v})]^2 := \min_{\boldsymbol{v}_K \in \boldsymbol{RTN}_p(K)} \|\boldsymbol{v} - \boldsymbol{v}_K\|^2_K + \left[\frac{h_K}{p+1}\|\nabla\cdot\boldsymbol{v} - \Pi^p_{\mathcal{T}}(\nabla\cdot\boldsymbol{v})\|_K\right]^2. \tag{3.2}$$

Note that the minimization in (3.2) does not involve a constraint on the divergence nor on the normal component on Γ_{N} (whenever relevant). Furthermore, since $\Pi^p_{\mathcal{T}}$ is the L^2-orthogonal projection onto the broken polynomial space $\mathcal{P}_p(\mathcal{T})$, we have $\|\nabla\cdot\boldsymbol{v} - \Pi^p_{\mathcal{T}}(\nabla\cdot\boldsymbol{v})\|_K = \min_{q\in\mathcal{P}_p(K)}\|\nabla\cdot\boldsymbol{v} - q\|_K$. Thus the local approximation errors $e_{K,p}(\boldsymbol{v})$ involve the local-best approximation errors in $\boldsymbol{L}^2$ plus a weighted L^2 best approximation error of the divergence.

Theorem 3.1 (Equivalence of local- and global-best approximations)**.** *There exists a constant C depending only on the space dimension d, the shape-regularity parameter $\kappa_{\mathcal{T}}$ of $\mathcal{T}$, and the polynomial degree $p \geq 0$, such that, for any $\boldsymbol{v} \in \boldsymbol{H}_{0,\Gamma_{\mathrm{N}}}(\mathrm{div},\Omega)$,*

$$[E_{\mathcal{T},p}(\boldsymbol{v})]^2 \leq C \sum_{K\in\mathcal{T}} [e_{K,p}(\boldsymbol{v})]^2 \leq C\,[E_{\mathcal{T},p}(\boldsymbol{v})]^2. \tag{3.3}$$

Theorem 3.1 shows that the global-best approximation error $E_{\mathcal{T},p}(\boldsymbol{v})$, is, up to a constant, equivalent to the local-best approximation error, i.e. the sum of the individual best approximation errors $e_{K,p}(\boldsymbol{v})$ over all mesh elements K. The first inequality in (3.3) is clearly nontrivial since the local-best approximations do not generally lead to a $\boldsymbol{H}_{0,\Gamma_{\mathrm{N}}}(\mathrm{div},\Omega)$-conforming object. Since the minimization in (3.2) is done over the space $\boldsymbol{RTN}_p(K)$ without any normal trace or divergence constraint, the second inequality in (3.3) holds trivially.

Remark 3.2 (Necessity of the divergence error terms)**.** The terms $\frac{h_K}{p+1}\|\nabla\cdot\boldsymbol{v} - \Pi^p_{\mathcal{T}}(\nabla\cdot\boldsymbol{v})\|_K$ might seem superfluous at the first sight, since they take an identical form in both $E_{\mathcal{T},p}(\boldsymbol{v})$ and $e_{K,p}(\boldsymbol{v})$. However, they cannot be removed from the local contributions $e_{K,p}(\boldsymbol{v})$. Otherwise, it would be possible to choose a sequence of functions $\boldsymbol{v}$ in $\boldsymbol{H}_{0,\Gamma_{\mathrm{N}}}(\mathrm{div},\Omega)$ approaching a function $\boldsymbol{\tau}_{\mathcal{T}} \in \boldsymbol{RTN}_p(\mathcal{T})$ but $\boldsymbol{\tau}_{\mathcal{T}} \notin \boldsymbol{H}_{0,\Gamma_{\mathrm{N}}}(\mathrm{div},\Omega)$ such that the middle term in (3.3) would tend to zero but $E_{\mathcal{T},p}(\boldsymbol{v})$ would remain uniformly bounded away from zero.

Remark 3.3 (Equivalence with constraint on the right-hand side)**.** Theorem 3.1 also straightforwardly implies that

$$\begin{aligned}[E_{\mathcal{T},p}(\boldsymbol{v})]^2 &\leq C \sum_{K\in\mathcal{T}} \left\{ \min_{\substack{\boldsymbol{v}_K \in \boldsymbol{RTN}_p(K) \\ \nabla\cdot\boldsymbol{v}_K = \Pi^p_{\mathcal{T}}(\nabla\cdot\boldsymbol{v})|_K}} \|\boldsymbol{v} - \boldsymbol{v}_K\|^2_K + \left[\frac{h_K}{p+1}\|\nabla\cdot\boldsymbol{v} - \Pi^p_{\mathcal{T}}(\nabla\cdot\boldsymbol{v})\|_K\right]^2 \right\} \\ &\leq C\,[E_{\mathcal{T},p}(\boldsymbol{v})]^2 \end{aligned} \tag{3.4}$$

with the same constant C, where the minimization problems in the middle term include a constraint on the divergence to mirror the divergence constraint in $E_{\mathcal{T},p}(\boldsymbol{v})$.

3.2 Polynomial-degree-robust one-sided bound

Our second result states a bound where the global-best approximation error is bounded in terms of the sums of local-best approximation errors with a constant that is robust with respect to the polynomial degree, but where the polynomial degree in the local approximation errors is $(p-1)$ instead of p. As a result, this is a one-sided inequality and not an equivalence, and it is valid only for $p \geq 1$.

Theorem 3.4 (Polynomial-degree-robust bound)**.** *There exists a constant C, depending only on the space dimension d and the shape-regularity parameter $\kappa_{\mathcal{T}}$ of $\mathcal{T}$, such that, for any $\boldsymbol{v} \in \boldsymbol{H}_{0,\Gamma_{\mathrm{N}}}(\mathrm{div},\Omega)$ and any $p \geq 1$,*

$$[E_{\mathcal{T},p}(\boldsymbol{v})]^2 \leq C \sum_{K\in\mathcal{T}} [e_{K,p-1}(\boldsymbol{v})]^2. \tag{3.5}$$

3.3 Optimal-order *hp*-approximation estimates

We now focus on functions with some additional elementwise regularity. For any $s \geq 0$ and any mesh element $K \in \mathcal{T}$, let $\boldsymbol{H}^s(K)$ denote the space of vector fields in $\boldsymbol{L}^2(K)$ with each component in $H^s(K)$. Our third main result delivers *hp*-optimal convergence rates for vector fields in $\boldsymbol{H}_{0,\Gamma_{\mathrm{N}}}(\mathrm{div},\Omega)$ with the minimally necessary additional elementwise regularity. Recall the definition (3.1) of $E_{\mathcal{T},p}(\boldsymbol{v})$.

Theorem 3.5 (*hp*-optimal convergence rates)**.** *Let $s \geq 0$ and let $\boldsymbol{v} \in \boldsymbol{H}_{0,\Gamma_{\mathrm{N}}}(\mathrm{div},\Omega)$ be such that*

$$\boldsymbol{v}|_K \in \boldsymbol{H}^s(K) \quad \forall K \in \mathcal{T}.$$

Let the polynomial degree $p \geq 0$. Then there exists a constant C, depending only on the regularity exponent s, the space dimension d, and the shape-regularity parameter $\kappa_{\mathcal{T}}$ of $\mathcal{T}$, such that

$$[E_{\mathcal{T},p}(\boldsymbol{v})]^2 \leq C\Big\{ \sum_{K\in\mathcal{T}} \Big[\frac{h_K^{\min(s,p+1)}}{(p+1)^s}\|\boldsymbol{v}\|_{\boldsymbol{H}^s(K)}\Big]^2 + \delta_{s<1}\Big[\frac{h_K}{p+1}\|\nabla\cdot\boldsymbol{v}\|_K\Big]^2\Big\}, \tag{3.6}$$

where $\delta_{s<1} := 1$ if $s < 1$ and $\delta_{s<1} := 0$ if $s \geq 1$.

Remark 3.6 (Full *hp*-optimality)**.** Theorem 3.5 shows that optimal order convergence rates with respect to both the mesh-sizes h_K and the polynomial degree p can be obtained despite the unfavorable dependence of the constant C on the polynomial degree p in Theorem 3.1 and unbalanced polynomial degrees in Theorem 3.4.

3.4 A simple stable local commuting projector in $\boldsymbol{H}_{0,\Gamma_{\mathrm{N}}}(\mathrm{div},\Omega)$

Our fourth main result is the construction of a local and stable commuting projector defined over the entire $\boldsymbol{H}_{0,\Gamma_{\mathrm{N}}}(\mathrm{div},\Omega)$ that leads to an approximation error that is equivalent to the local-best approximation error. Recall the definition of the broken Raviart–Thomas–Nédélec interpolant $\boldsymbol{I}^p_{\mathcal{T}}$ from (2.2) and that of the piecewise polynomial patchwise $\boldsymbol{H}(\mathrm{div};\omega_{\boldsymbol{a}})$-conforming spaces $\boldsymbol{V}_p(\omega_{\boldsymbol{a}})$ from (2.4). Recall also that zero extensions of elements of $\boldsymbol{V}_p(\omega_{\boldsymbol{a}})$ belong to $\boldsymbol{RTN}_p(\mathcal{T}) \cap \boldsymbol{H}_{0,\Gamma_{\mathrm{N}}}(\mathrm{div},\Omega)$, and that $\psi_{\boldsymbol{a}}$ is the piecewise affine Lagrange finite element basis function associated with the vertex $\boldsymbol{a}$.

Definition 3.7 (A simple locally-defined mapping from $\boldsymbol{H}_{0,\Gamma_{\mathrm{N}}}(\mathrm{div},\Omega)$ to $\boldsymbol{RTN}_p(\mathcal{T}) \cap \boldsymbol{H}_{0,\Gamma_{\mathrm{N}}}(\mathrm{div},\Omega)$)**.** *Let $\boldsymbol{v} \in \boldsymbol{H}_{0,\Gamma_{\mathrm{N}}}(\mathrm{div},\Omega)$ be arbitrary. Let $\boldsymbol{\tau}_{\mathcal{T}} \in \boldsymbol{RTN}_p(\mathcal{T})$ be defined elementwise by*

$$\boldsymbol{\tau}_{\mathcal{T}}|_K := \operatorname*{arg\,min}_{\substack{\boldsymbol{v}_K\in\boldsymbol{RTN}_p(K)\\ \nabla\cdot\boldsymbol{v}_K=\Pi^p_{\mathcal{T}}(\nabla\cdot\boldsymbol{v})}} \|\boldsymbol{v}-\boldsymbol{v}_K\|_K \qquad \forall K \in \mathcal{T}. \tag{3.7}$$

For each mesh vertex $\boldsymbol{a} \in \mathcal{V}$, let $\boldsymbol{\sigma}_{\boldsymbol{a}} \in \boldsymbol{V}_p(\omega_{\boldsymbol{a}})$ be defined by

$$\boldsymbol{\sigma}_{\boldsymbol{a}} := \operatorname*{arg\,min}_{\substack{\boldsymbol{v}_{\boldsymbol{a}}\in\boldsymbol{V}_p(\omega_{\boldsymbol{a}})\\ \nabla\cdot\boldsymbol{v}_{\boldsymbol{a}}=\Pi^p_{\mathcal{T}}(\psi_{\boldsymbol{a}}\nabla\cdot\boldsymbol{v}+\nabla\psi_{\boldsymbol{a}}\cdot\boldsymbol{\tau}_{\mathcal{T}})}} \|\boldsymbol{v}_{\boldsymbol{a}} - \boldsymbol{I}^p_{\mathcal{T}}(\psi_{\boldsymbol{a}}\boldsymbol{\tau}_{\mathcal{T}})\|_{\omega_{\boldsymbol{a}}}. \tag{3.8}$$

Extending the functions $\boldsymbol{\sigma}_{\boldsymbol{a}}$ from the patch domains $\omega_{\boldsymbol{a}}$ to the rest of Ω by zero, we define $P^p_{\mathcal{T}}(\boldsymbol{v}) \in \boldsymbol{RTN}_p(\mathcal{T}) \cap \boldsymbol{H}_{0,\Gamma_{\mathrm{N}}}(\mathrm{div},\Omega)$ by

$$P^p_{\mathcal{T}}(\boldsymbol{v}) := \boldsymbol{\sigma}_{\mathcal{T}} := \sum_{\boldsymbol{a}\in\mathcal{V}} \boldsymbol{\sigma}_{\boldsymbol{a}}. \tag{3.9}$$

The justification that the construction of $P^p_{\mathcal{T}}(\boldsymbol{v})$ is well-defined in given in Section 4.1 below. The first step (3.7) in Definition 3.7 considers the elementwise $\boldsymbol{L}^2$-norm local-best approximation that defines the discontinuous piecewise RTN polynomial $\boldsymbol{\tau}_{\mathcal{T}}$ closest to $\boldsymbol{v}$ under the divergence constraint. The second step in (3.8) can be seen as smoothing $\boldsymbol{\tau}_{\mathcal{T}}$ over the patch subdomains $\omega_{\boldsymbol{a}}$ to obtain an $\boldsymbol{H}(\mathrm{div})$-conforming approximation $\boldsymbol{\sigma}_{\boldsymbol{a}}$ over each vertex patch with a suitably prescribed divergence. These approximations $\boldsymbol{\sigma}_{\boldsymbol{a}}$ are then summed into $P^p_{\mathcal{T}}(\boldsymbol{v})$. The overall procedure is motivated by equilibrated flux reconstructions coming from *a posteriori* error estimation [21, 8, 25]. Here we adapt those techniques to the purpose of *a priori* error analysis. Our main result is the following.

Theorem 3.8 (Commutativity, projection, approximation, and stability of $P^p_{\mathcal{T}}$)**.** *Let a mesh $\mathcal{T}$ of Ω and a polynomial degree $p \geq 0$ be fixed. Let $\boldsymbol{v} \in \boldsymbol{H}_{0,\Gamma_{\mathrm{N}}}(\mathrm{div},\Omega)$ be arbitrary. Then, $P^p_{\mathcal{T}}(\boldsymbol{v})$ from Definition 3.7 satisfies*

$$\nabla\cdot P^p_{\mathcal{T}}(\boldsymbol{v}) = \Pi^p_{\mathcal{T}}(\nabla\cdot\boldsymbol{v}), \tag{3.10}$$

$$P^p_{\mathcal{T}}(\boldsymbol{v}) = \boldsymbol{v} \qquad \text{for all } \boldsymbol{v} \in \boldsymbol{RTN}_p(\mathcal{T}) \cap \boldsymbol{H}_{0,\Gamma_{\mathrm{N}}}(\mathrm{div},\Omega), \tag{3.11}$$

$$\|\boldsymbol{v} - P^p_{\mathcal{T}}(\boldsymbol{v})\|^2_K + \Big[\frac{h_K}{p+1}\|\nabla\cdot(\boldsymbol{v} - P^p_{\mathcal{T}}(\boldsymbol{v}))\|_K\Big]^2 \leq C \sum_{K'\in\mathcal{T}_K} [e_{K',p}(\boldsymbol{v})]^2 \qquad \forall K \in \mathcal{T}, \tag{3.12}$$

where $\mathcal{T}_K := \cup_{\boldsymbol{a}\in\mathcal{V}_K}\mathcal{T}_{\boldsymbol{a}}$ are the neighboring elements of K. Furthermore, we have

$$\|P^p_{\mathcal{T}}(\boldsymbol{v})\|^2_K \leq C \sum_{K'\in\mathcal{T}_K} \left\{\|\boldsymbol{v}\|^2_{K'} + \Big[\frac{h_{K'}}{p+1}\|\nabla\cdot\boldsymbol{v} - \Pi^p_{\mathcal{T}}(\nabla\cdot\boldsymbol{v})\|_{K'}\Big]^2\right\} \qquad \forall K \in \mathcal{T}, \tag{3.13}$$

$$\|P^p_{\mathcal{T}}(\boldsymbol{v})\|^2_K + h^2_\Omega\|\nabla\cdot P^p_{\mathcal{T}}(\boldsymbol{v})\|^2_K \leq C \sum_{K'\in\mathcal{T}_K} \left\{\|\boldsymbol{v}\|^2_{K'} + h^2_\Omega\|\nabla\cdot\boldsymbol{v}\|^2_{K'}\right\} \qquad \forall K \in \mathcal{T}, \tag{3.14}$$

recalling that h_Ω denotes the diameter of Ω and where the constant C only depends on the space dimension d, the shape-regularity parameter $\kappa_{\mathcal{T}}$ of $\mathcal{T}$, and the polynomial degree p.

Property (3.13) readily implies that $P^p_{\mathcal{T}}$ is globally L^2-stable up to hp data oscillation of the divergence, since summing over the mesh elements leads to

$$\|P^p_{\mathcal{T}}(\boldsymbol{v})\|^2 \leq C\left\{\|\boldsymbol{v}\|^2 + \sum_{K\in\mathcal{T}}\Big[\frac{h_K}{p+1}\|\nabla\cdot\boldsymbol{v} - \Pi^p_{\mathcal{T}}(\nabla\cdot\boldsymbol{v})\|_K\Big]^2\right\} \qquad \forall \boldsymbol{v} \in \boldsymbol{H}_{0,\Gamma_{\mathrm{N}}}(\mathrm{div},\Omega).$$

Similarly, from (3.14), we infer that $P^p_{\mathcal{T}}$ is $\boldsymbol{H}(\mathrm{div})$-stable, since

$$\|P^p_{\mathcal{T}}(\boldsymbol{v})\|^2 + h^2_\Omega\|\nabla\cdot P^p_{\mathcal{T}}(\boldsymbol{v})\|^2 \leq C\big[\|\boldsymbol{v}\|^2 + h^2_\Omega\|\nabla\cdot\boldsymbol{v}\|^2\big] \qquad \forall \boldsymbol{v} \in \boldsymbol{H}_{0,\Gamma_{\mathrm{N}}}(\mathrm{div},\Omega).$$

We note that, for the divergence term, (3.13) improves the bound (5.2) of [29, Theorem 5.2] (in particular, we have $\|\nabla\cdot\boldsymbol{v} - \Pi^p_{\mathcal{T}}(\nabla\cdot\boldsymbol{v})\|_K$ in place of $\|\nabla\cdot\boldsymbol{v}\|_K$), whereas (3.14) is similar to the combination of the bounds (5.2) and (5.3) of [29, Theorem 5.2].

4 Proofs of the main results

This section, whose goal is to prove our main results, namely Theorem 3.1, Theorem 3.4, Theorem 3.5, and Theorem 3.8, is organized as follows. First, in Section 4.1, we provide some justifications on the construction of the mapping $P^p_{\mathcal{T}}$ from Definition 3.7 and we establish the statement (3.10) from Theorem 3.8 (on commuting with the divergence). Then, in Section 4.2, we prove the statement (3.12) from Theorem 3.8 on the approximation properties of $P^p_{\mathcal{T}}$. This is the more technical part of the proofs. Next, in Section 4.3, we finish the proof of Theorem 3.8 by proving the remaining three statements (3.11), (3.13), and (3.14), and we also prove Theorem 3.1. The proof of Theorem 3.4 is then presented in Section 4.4, where we adapt slightly the arguments from the previous sections. Finally the proof of Theorem 3.5, which hinges on Theorem 3.1 and on Theorem 3.4, is contained in Section 4.5.

4.1 Justifications on the construction of $P^p_{\mathcal{T}}$

We start by showing that the projector $P^p_{\mathcal{T}}$ of Definition 3.7 is well-defined on $\boldsymbol{H}_{0,\Gamma_{\mathrm{N}}}(\mathrm{div},\Omega)$

Lemma 4.1 (Discrete weak divergence of $\boldsymbol{L}^2$-projection)**.** *For any function $\boldsymbol{v} \in \boldsymbol{H}_{0,\Gamma_{\mathrm{N}}}(\mathrm{div},\Omega)$, let $\boldsymbol{\tau}_{\mathcal{T}}$ be defined elementwise in (3.7). Then*

$$(\nabla\cdot\boldsymbol{v}, \psi_{\boldsymbol{a}})_{\omega_{\boldsymbol{a}}} + (\boldsymbol{\tau}_{\mathcal{T}}, \nabla\psi_{\boldsymbol{a}})_{\omega_{\boldsymbol{a}}} = 0 \quad \forall \boldsymbol{a} \in \mathcal{V}_\Omega \cup \mathcal{V}_{\mathrm{N}}. \tag{4.1}$$

Proof. First, observe that for any vertex $\boldsymbol{a} \in \mathcal{V}_\Omega \cup \mathcal{V}_{\mathrm{N}}$, the hat function $\psi_{\boldsymbol{a}}$ belongs to $H^1_{\Gamma_{\mathrm{D}}}(\Omega)$ owing to the conformity of $\mathcal{T}$ with respect to the Dirichlet and Neumann boundary sets. Therefore, $(\nabla\cdot\boldsymbol{v}, \psi_{\boldsymbol{a}})_{\omega_{\boldsymbol{a}}} + (\boldsymbol{v}, \nabla\psi_{\boldsymbol{a}})_{\omega_{\boldsymbol{a}}} = 0$, where we use the fact that $\omega_{\boldsymbol{a}}$ is the support of $\psi_{\boldsymbol{a}}$. Since $\nabla\psi_{\boldsymbol{a}}$ is a constant vector on each element K, the Euler–Lagrange equations for (3.7) imply that

$$(\boldsymbol{\tau}_{\mathcal{T}}, \nabla\psi_{\boldsymbol{a}})_K = (\boldsymbol{v}, \nabla\psi_{\boldsymbol{a}})_K \qquad \forall K \in \mathcal{T}_{\boldsymbol{a}}. \tag{4.2}$$

Consequently, $(\boldsymbol{\tau}_{\mathcal{T}}, \nabla\psi_{\boldsymbol{a}})_{\omega_{\boldsymbol{a}}} = (\boldsymbol{v}, \nabla\psi_{\boldsymbol{a}})_{\omega_{\boldsymbol{a}}}$, and (4.1) follows. □

We now show that the local minimization problems (3.8) give well-defined local contributions $\boldsymbol{\sigma}_{\boldsymbol{a}}$.

Lemma 4.2 (Existence and uniqueness of local problems)**.** *For each vertex $\boldsymbol{a} \in \mathcal{V}$, there exists a unique $\boldsymbol{\sigma}_{\boldsymbol{a}} \in \boldsymbol{V}_p(\omega_{\boldsymbol{a}})$ satisfying (3.8).*

Proof. The minimization problem (3.8) is equivalent to a mixed finite element problem in the patch subdomain $\omega_{\boldsymbol{a}}$. For Dirichlet boundary vertices $\boldsymbol{a} \in \mathcal{V}_{\mathrm{D}}$, this problem is well-posed with a unique minimizer since the space $\boldsymbol{V}_p(\omega_{\boldsymbol{a}})$ of (2.4) does not impose the normal constraint everywhere on $\partial\omega_{\boldsymbol{a}}$. For interior and Neumann vertices $\boldsymbol{a} \in \mathcal{V}_\Omega \cup \mathcal{V}_{\mathrm{N}}$, the source term in the divergence constraint satisfies the compatibility condition

$$(\Pi^p_{\mathcal{T}}(\psi_{\boldsymbol{a}}\nabla\cdot\boldsymbol{v} + \nabla\psi_{\boldsymbol{a}}\cdot\boldsymbol{\tau}_{\mathcal{T}}), 1)_{\omega_{\boldsymbol{a}}} = (\nabla\cdot\boldsymbol{v}, \psi_{\boldsymbol{a}})_{\omega_{\boldsymbol{a}}} + (\boldsymbol{\tau}_{\mathcal{T}}, \nabla\psi_{\boldsymbol{a}})_{\omega_{\boldsymbol{a}}} = 0, \tag{4.3}$$

where the second equality follows from Lemma 4.1. Therefore, $\boldsymbol{\sigma}_{\boldsymbol{a}}$ is also well-defined for interior and Neumann vertices $\boldsymbol{a} \in \mathcal{V}_\Omega \cup \mathcal{V}_{\mathrm{N}}$. □

It follows from Lemma 4.2 that $P^p_{\mathcal{T}}(\boldsymbol{v}) \in \boldsymbol{RTN}_p(\mathcal{T}) \cap \boldsymbol{H}_{0,\Gamma_{\mathrm{N}}}(\mathrm{div},\Omega)$ is well-defined for every $\boldsymbol{v} \in \boldsymbol{H}_{0,\Gamma_{\mathrm{N}}}(\mathrm{div},\Omega)$. We are now ready to establish the statement (3.10) from Theorem 3.8 concerning the commuting of $P^p_{\mathcal{T}}$ with the divergence operator.

Lemma 4.3 (Statement (3.10) from Theorem 3.8)**.** *$P^p_{\mathcal{T}}$ satisfies the commuting property (3.10).*

Proof. Since the functions $\{\psi_{\boldsymbol{a}}\}_{\boldsymbol{a}\in\mathcal{V}}$ form a partition of unity over Ω, i.e., $\sum_{\boldsymbol{a}\in\mathcal{V}} \psi_{\boldsymbol{a}} = 1$, and consequently $\sum_{\boldsymbol{a}\in\mathcal{V}} \nabla\psi_{\boldsymbol{a}} = \boldsymbol{0}$, we find that

$$\nabla\cdot P^p_{\mathcal{T}}(\boldsymbol{v}) = \sum_{\boldsymbol{a}\in\mathcal{V}} \nabla\cdot\boldsymbol{\sigma}_{\boldsymbol{a}} = \sum_{\boldsymbol{a}\in\mathcal{V}} \left\{\Pi^p_{\mathcal{T}}(\psi_{\boldsymbol{a}}\nabla\cdot\boldsymbol{v} + \nabla\psi_{\boldsymbol{a}}\cdot\boldsymbol{\tau}_{\mathcal{T}})\right\} = \Pi^p_{\mathcal{T}}(\nabla\cdot\boldsymbol{v}). \tag{4.4}$$

□

We will also use below the following polynomial-degree-robust equivalence for constrained and unconstrained local-best approximations, where the second inequality may come as a surprise.

Lemma 4.4 (Local constrained/unconstrained equivalence)**.** *Let a simplex $K \in \mathcal{T}$ and $\boldsymbol{v} \in \boldsymbol{H}(\mathrm{div}; K)$ be fixed. Let $\boldsymbol{\tau}_{\mathcal{T}}$ be defined elementwise in (3.7). Then, there exists a constant C, depending only on the space dimension d and the mesh shape-regularity parameter $\kappa_{\mathcal{T}}$, such that*

$$e_{K,p}(\boldsymbol{v}) \le \|\boldsymbol{v} - \boldsymbol{\tau}_{\mathcal{T}}\|_K + \frac{h_K}{p+1}\|\nabla\cdot(\boldsymbol{v} - \boldsymbol{\tau}_{\mathcal{T}})\|_K \le C e_{K,p}(\boldsymbol{v}). \tag{4.5}$$

Proof. Since $\nabla\cdot\boldsymbol{\tau}_{\mathcal{T}} = \Pi^p_{\mathcal{T}}(\nabla\cdot\boldsymbol{v})$ from (3.7), the first inequality in (4.5) is obvious, so we show the second one. First, $\frac{h_K}{p+1}\|\nabla\cdot(\boldsymbol{v} - \boldsymbol{\tau}_{\mathcal{T}})\|_K \le e_{K,p}(\boldsymbol{v})$ trivially holds true for the same reason, so it remains only to bound $\|\boldsymbol{v} - \boldsymbol{\tau}_{\mathcal{T}}\|_K$. Let $\widetilde{\boldsymbol{\tau}}_{\mathcal{T}}$ be the elementwise $\boldsymbol{L}^2$-projection of $\boldsymbol{v}$ into $\boldsymbol{RTN}_p(\mathcal{T})$, so that

$$[e_{K,p}(\boldsymbol{v})]^2 = \|\boldsymbol{v} - \widetilde{\boldsymbol{\tau}}_{\mathcal{T}}\|^2_K + \frac{h_K^2}{(p+1)^2}\|\nabla\cdot\boldsymbol{v} - \Pi^p_{\mathcal{T}}(\nabla\cdot\boldsymbol{v})\|^2_K.$$

It follows from [15, Corollary 3.4] that there exists $\boldsymbol{v}_K \in \boldsymbol{RTN}_p(K)$ such that $\nabla\cdot\boldsymbol{v}_K = \Pi^p_{\mathcal{T}}(\nabla\cdot\boldsymbol{v})$ and

$$\|\boldsymbol{v}_K - \widetilde{\boldsymbol{\tau}}_{\mathcal{T}}\|_K \le C \sup_{\substack{\varphi\in H^1_0(K)\\ \|\nabla\varphi\|_K=1}} \left\{(\Pi^p_{\mathcal{T}}(\nabla\cdot\boldsymbol{v}) - \nabla\cdot\widetilde{\boldsymbol{\tau}}_{\mathcal{T}}, \varphi)_K\right\}, \tag{4.6}$$

where C only depends on d and $\kappa_{\mathcal{T}}$. Since $(\nabla\cdot\boldsymbol{v}, \varphi)_K + (\boldsymbol{v}, \nabla\varphi)_K = 0$ and since $(\nabla\cdot\widetilde{\boldsymbol{\tau}}_{\mathcal{T}}, \varphi)_K + (\widetilde{\boldsymbol{\tau}}_{\mathcal{T}}, \nabla\varphi)_K = 0$ for all $\varphi \in H^1_0(K)$, we see that

$$(\Pi^p_{\mathcal{T}}(\nabla\cdot\boldsymbol{v}), \varphi)_K - (\nabla\cdot\widetilde{\boldsymbol{\tau}}_{\mathcal{T}}, \varphi)_K = (\Pi^p_{\mathcal{T}}(\nabla\cdot\boldsymbol{v}) - \nabla\cdot\boldsymbol{v}, \varphi - \Pi^p_{\mathcal{T}}(\varphi))_K - (\boldsymbol{v} - \widetilde{\boldsymbol{\tau}}_{\mathcal{T}}, \nabla\varphi)_K.$$

Therefore, the inequality (4.6) combined with the approximation bound $\|\varphi - \Pi^p_{\mathcal{T}}(\varphi)\|_K \le C\frac{h_K}{p+1}\|\nabla\varphi\|_K$, with a constant C depending only on d and $\kappa_{\mathcal{T}}$, implies that

$$\|\boldsymbol{v}_K - \widetilde{\boldsymbol{\tau}}_{\mathcal{T}}\|_K \le C\left\{\|\boldsymbol{v} - \widetilde{\boldsymbol{\tau}}_{\mathcal{T}}\|^2_K + \left[\frac{h_K}{p+1}\|\nabla\cdot\boldsymbol{v} - \Pi^p_{\mathcal{T}}(\nabla\cdot\boldsymbol{v})\|_K\right]^2\right\}^{\frac{1}{2}} = C e_{K,p}(\boldsymbol{v}).$$

Owing to the triangle inequality $\|\boldsymbol{v} - \boldsymbol{v}_K\|_K \le \|\boldsymbol{v} - \widetilde{\boldsymbol{\tau}}_{\mathcal{T}}\|_K + \|\widetilde{\boldsymbol{\tau}}_{\mathcal{T}} - \boldsymbol{v}_K\|_K$, we infer that $\|\boldsymbol{v} - \boldsymbol{v}_K\|_K \le C e_{K,p}(\boldsymbol{v})$. Finally, the elementwise definition of $\boldsymbol{\tau}_{\mathcal{T}}$ as the minimizer in (3.7) implies that $\|\boldsymbol{v} - \boldsymbol{\tau}_{\mathcal{T}}\|_K \le \|\boldsymbol{v} - \boldsymbol{v}_K\|_K$, and this yields the second bound (4.5). □

4.2 Proof of the statement (3.12) from Theorem 3.8

Let us start with two useful technical results. For a given vertex $\boldsymbol{a} \in \mathcal{V}$, let the space $H^1_*(\omega_{\boldsymbol{a}})$ be defined by

$$H^1_*(\omega_{\boldsymbol{a}}) := \begin{cases} \{\varphi \in H^1(\omega_{\boldsymbol{a}}), \quad (\varphi, 1)_{\omega_{\boldsymbol{a}}} = 0\} & \text{if } \boldsymbol{a} \in \mathcal{V}_\Omega \cup \mathcal{V}_\mathrm{N}, \\ \{\varphi \in H^1(\omega_{\boldsymbol{a}}), \quad \varphi|_{\partial\omega_{\boldsymbol{a}} \cap \Gamma_\mathrm{D}^{\boldsymbol{a}}} = 0\} & \text{if } \boldsymbol{a} \in \mathcal{V}_\mathrm{D}, \end{cases} \tag{4.7}$$

where we recall that $\Gamma_\mathrm{D}^{\boldsymbol{a}}$ contains those boundary faces from Γ_D that share the vertex $\boldsymbol{a}$. Recall also the discrete spaces $\boldsymbol{V}_p(\omega_{\boldsymbol{a}})$ defined in (2.4). The following result has been shown in Braess *et al.* [9, Theorem 7] in two space dimensions and [27, Corollaries 3.3, 3.6, and 3.8] in three space dimensions.

Lemma 4.5 (Stability of patchwise flux equilibration)**.** *Let a vertex $\boldsymbol{a} \in \mathcal{V}$ be fixed, and let $g_{\boldsymbol{a}} \in \mathcal{P}_p(\mathcal{T}_{\boldsymbol{a}})$ and $\boldsymbol{\tau}_{\boldsymbol{a}} \in \boldsymbol{RTN}_p(\mathcal{T}_{\boldsymbol{a}})$ be given discontinuous piecewise polynomials with the condition $(g_{\boldsymbol{a}}, 1)_{\omega_{\boldsymbol{a}}} = 0$ if $\boldsymbol{a} \in \mathcal{V}_\Omega \cup \mathcal{V}_\mathrm{N}$. Then, there exists a constant C, depending only on the space dimension d and the mesh shape-regularity parameter $\kappa_{\mathcal{T}}$, such that*

$$\min_{\substack{\boldsymbol{v}_{\boldsymbol{a}} \in \boldsymbol{V}_p(\omega_{\boldsymbol{a}}) \\ \nabla\cdot\boldsymbol{v}_{\boldsymbol{a}} = g_{\boldsymbol{a}}}} \|\boldsymbol{v}_{\boldsymbol{a}} - \boldsymbol{\tau}_{\boldsymbol{a}}\|_{\omega_{\boldsymbol{a}}} \le C \sup_{\substack{\varphi \in H^1_*(\omega_{\boldsymbol{a}}) \\ \|\nabla\varphi\|_{\omega_{\boldsymbol{a}}} = 1}} \left\{ (g_{\boldsymbol{a}}, \varphi)_{\omega_{\boldsymbol{a}}} + (\boldsymbol{\tau}_{\boldsymbol{a}}, \nabla\varphi)_{\omega_{\boldsymbol{a}}} \right\}. \tag{4.8}$$

We shall also use the following auxiliary bound for jump terms based on the bubble function technique of Verfürth, cf. [44].

Lemma 4.6 (Bound on face terms)**.** *Let a face $F \in \mathcal{F}$ be fixed, and let $\mathcal{T}_F$ be the set of one or two elements $K \in \mathcal{T}$ to which F belongs, with ω_F the corresponding open subdomain. Let h_F denote the diameter of F. Then, there exists a constant C, depending on the space dimension d, the mesh shape-regularity parameter $\kappa_{\mathcal{T}}$, and the polynomial degree p, such that*

$$h_F^{1/2} \|q_h\|_F \le C \sup_{\substack{\varphi \in H^1(\omega_F) \\ \varphi = 0 \text{ on } \partial\omega_F \setminus F \\ \|\nabla\varphi\|_{\omega_F} = 1}} (q_h, \varphi)_F \qquad \forall q_h \in \mathcal{P}_p(F). \tag{4.9}$$

We are now ready to prove the statement (3.12) from Theorem 3.8. Let $\boldsymbol{v} \in \boldsymbol{H}_{0,\Gamma_\mathrm{N}}(\mathrm{div}, \Omega)$ be arbitrary. Since it follows from $\nabla\cdot P^p_{\mathcal{T}}(\boldsymbol{v}) = \Pi^p_{\mathcal{T}}(\nabla\cdot\boldsymbol{v})$ that

$$\frac{h_K}{p+1} \|\nabla\cdot\boldsymbol{v} - \nabla\cdot P^p_{\mathcal{T}}(\boldsymbol{v})\|_K \le e_{K,p}(\boldsymbol{v}), \tag{4.10}$$

it only remains to prove that

$$\|\boldsymbol{v} - P^p_{\mathcal{T}}(\boldsymbol{v})\|_K \le C \left\{ \sum_{K' \in \mathcal{T}_K} e_{K',p}(\boldsymbol{v})^2 \right\}^{\frac{1}{2}} \qquad \forall K \in \mathcal{T}. \tag{4.11}$$

We proceed for this purpose in two steps.

Step 1. Bound on $\boldsymbol{\sigma}_{\boldsymbol{a}}$. Recall that $\boldsymbol{\sigma}_{\boldsymbol{a}}$ is defined in (3.8) with $\boldsymbol{\tau}_{\mathcal{T}}$ defined elementwise in (3.7).

Lemma 4.7 (Bound on $\boldsymbol{\sigma}_{\boldsymbol{a}}$)**.** *There exists a constant C, depending only on d, $\kappa_{\mathcal{T}}$, and p, such that*

$$\|\boldsymbol{\sigma}_{\boldsymbol{a}} - \boldsymbol{I}^p_{\mathcal{T}}(\psi_{\boldsymbol{a}}\boldsymbol{\tau}_{\mathcal{T}})\|_{\omega_{\boldsymbol{a}}} \le C \left\{ \sum_{K \in \mathcal{T}_{\boldsymbol{a}}} [e_{K,p}(\boldsymbol{v})]^2 \right\}^{\frac{1}{2}} \qquad \forall \boldsymbol{a} \in \mathcal{V}. \tag{4.12}$$

Proof. First, since $\boldsymbol{I}^p_{\mathcal{T}}(\psi_{\boldsymbol{a}}\boldsymbol{\tau}_{\mathcal{T}}) \in \boldsymbol{RTN}_p(\mathcal{T}_{\boldsymbol{a}})$, we can apply Lemma 4.5 to $\boldsymbol{\sigma}_{\boldsymbol{a}}$, with the choices $\boldsymbol{\tau}_{\boldsymbol{a}} := \boldsymbol{I}^p_{\mathcal{T}}(\psi_{\boldsymbol{a}}\boldsymbol{\tau}_{\mathcal{T}})$ and $g_{\boldsymbol{a}} := \Pi^p_{\mathcal{T}}(\psi_{\boldsymbol{a}}\nabla\cdot\boldsymbol{v} + \nabla\psi_{\boldsymbol{a}}\cdot\boldsymbol{\tau}_{\mathcal{T}}) \in \mathcal{P}_p(\mathcal{T}_{\boldsymbol{a}})$ to obtain

$$\|\boldsymbol{\sigma}_{\boldsymbol{a}} - \boldsymbol{I}^p_{\mathcal{T}}(\psi_{\boldsymbol{a}}\boldsymbol{\tau}_{\mathcal{T}})\|_{\omega_{\boldsymbol{a}}} \le C \sup_{\substack{\varphi \in H^1_*(\omega_{\boldsymbol{a}}) \\ \|\nabla\varphi\|_{\omega_{\boldsymbol{a}}} = 1}} \left\{ (g_{\boldsymbol{a}}, \varphi)_{\omega_{\boldsymbol{a}}} + (\boldsymbol{I}^p_{\mathcal{T}}(\psi_{\boldsymbol{a}}\boldsymbol{\tau}_{\mathcal{T}}), \nabla\varphi)_{\omega_{\boldsymbol{a}}} \right\}, \tag{4.13}$$

where the space $H^1_*(\omega_{\boldsymbol{a}})$ is defined in (4.7). Let $h_{\omega_{\boldsymbol{a}}}$ denote the diameter of $\omega_{\boldsymbol{a}}$ and recall the Poincaré inequality $\|v\|_{\omega_{\boldsymbol{a}}} \le C h_{\omega_{\boldsymbol{a}}} \|\nabla v\|_{\omega_{\boldsymbol{a}}}$ on $H^1_*(\omega_{\boldsymbol{a}})$, with a constant C depending only on the dimension d and

on $\kappa_{\mathcal{T}}$. Moreover, note that the shape-regularity of the mesh implies that $h_{\omega_{\boldsymbol{a}}} \approx h_K \approx h_F$ for all $K \in \mathcal{T}_{\boldsymbol{a}}$ and all $F \in \mathcal{F}_{\boldsymbol{a}}^{\mathrm{in}}$.

To bound the right-hand side of (4.13), consider an arbitrary $\varphi \in H^1_*(\omega_{\boldsymbol{a}})$ such that $\|\nabla\varphi\|_{\omega_{\boldsymbol{a}}} = 1$. Then, using integration by parts elementwise, we find that

$$\begin{aligned}\big(\boldsymbol{I}_{\mathcal{T}}^p(\psi_{\boldsymbol{a}}\boldsymbol{\tau}_{\mathcal{T}}), \nabla\varphi\big)_{\omega_{\boldsymbol{a}}} &= \sum_{F\in\mathcal{F}_{\boldsymbol{a}}^{\mathrm{in}}} \big([\![\boldsymbol{I}_{\mathcal{T}}^p(\psi_{\boldsymbol{a}}\boldsymbol{\tau}_{\mathcal{T}})]\!]\cdot\boldsymbol{n}_F, \varphi\big)_F - \sum_{K\in\mathcal{T}_{\boldsymbol{a}}} \big(\nabla\cdot\boldsymbol{I}_{\mathcal{T}}^p(\psi_a\boldsymbol{\tau}_{\mathcal{T}}), \varphi\big)_K \\ &= \sum_{F\in\mathcal{F}_{\boldsymbol{a}}^{\mathrm{in}}} \big(\Pi_F^p(\psi_{\boldsymbol{a}}[\![\boldsymbol{\tau}_{\mathcal{T}}]\!]\cdot\boldsymbol{n}_F), \varphi\big)_F - \big(\Pi_{\mathcal{T}}^p(\nabla_{\mathcal{T}}\cdot(\psi_a\boldsymbol{\tau}_{\mathcal{T}})), \varphi\big)_{\omega_{\boldsymbol{a}}}.\end{aligned}$$

Here, in the first identity, the set of faces can be restricted to $\mathcal{F}_{\boldsymbol{a}}^{\mathrm{in}}$; indeed, for interior vertices, this follows from the fact that $\psi_{\boldsymbol{a}}$ vanishes on $\partial\omega_{\boldsymbol{a}}$, whereas for boundary vertices, $\varphi \in H^1_*(\omega_{\boldsymbol{a}})$ vanishes on $\Gamma_{\mathrm{D}}^{\boldsymbol{a}}$. The second identity is then obtained from the definition of $\boldsymbol{I}_{\mathcal{T}}^p$ in (2.2) and the commutation identity (2.3). Expanding $\nabla_{\mathcal{T}}\cdot(\psi_a\boldsymbol{\tau}_{\mathcal{T}}) = \nabla\psi_{\boldsymbol{a}}\cdot\boldsymbol{\tau}_{\mathcal{T}} + \psi_{\boldsymbol{a}}\nabla_{\mathcal{T}}\cdot\boldsymbol{\tau}_{\mathcal{T}}$ and simplifying gives

$$\begin{aligned}(g_{\boldsymbol{a}}, \varphi)_{\omega_{\boldsymbol{a}}} + (\boldsymbol{I}_{\mathcal{T}}^p(\psi_{\boldsymbol{a}}\boldsymbol{\tau}_{\mathcal{T}}), \nabla\varphi)_{\omega_{\boldsymbol{a}}} &= (\Pi_{\mathcal{T}}^p(\psi_{\boldsymbol{a}}\nabla_{\mathcal{T}}\cdot(\boldsymbol{v}-\boldsymbol{\tau}_{\mathcal{T}})), \varphi)_{\omega_{\boldsymbol{a}}} \\ &\quad + \sum_{F\in\mathcal{F}_{\boldsymbol{a}}^{\mathrm{in}}} \big(\Pi_F^p(\psi_{\boldsymbol{a}}[\![\boldsymbol{\tau}_{\mathcal{T}}]\!]\cdot\boldsymbol{n}_F), \varphi\big)_F.\end{aligned} \tag{4.14}$$

We now bound the two terms on the right-hand side of (4.14) separately. To bound the first term, we consider first the case $p \geq 1$: using the divergence constraint on $\boldsymbol{\tau}_{\mathcal{T}}$ in (3.7), the orthogonality of the L^2-projections, the approximation bound $\|\varphi - \Pi_{\mathcal{T}}^{p-1}\varphi\|_K \leq C\frac{h_K}{p+1}\|\nabla\varphi\|_K$ (note that $\frac{1}{p} \leq \frac{2}{p+1}$ for all $p \geq 1$), along with $\|\psi_{\boldsymbol{a}}\|_{\infty,\omega_{\boldsymbol{a}}} = 1$ and $\|\nabla\varphi\|_{\omega_{\boldsymbol{a}}} = 1$, we find that there is a constant C, depending only on d and $\kappa_{\mathcal{T}}$, such that

$$\begin{aligned}&(\Pi_{\mathcal{T}}^p(\psi_{\boldsymbol{a}}\nabla_{\mathcal{T}}\cdot(\boldsymbol{v}-\boldsymbol{\tau}_{\mathcal{T}})), \varphi)_{\omega_{\boldsymbol{a}}} \\ &= \Big(\nabla\cdot\boldsymbol{v} - \Pi_{\mathcal{T}}^p(\nabla\cdot\boldsymbol{v}), \psi_{\boldsymbol{a}}\Pi_{\mathcal{T}}^p(\varphi - \Pi_{\mathcal{T}}^{p-1}\varphi)\Big)_{\omega_{\boldsymbol{a}}} \\ &\leq C\left\{\sum_{K\in\mathcal{T}_{\boldsymbol{a}}} \frac{h_K^2}{(p+1)^2}\|\nabla\cdot\boldsymbol{v} - \Pi_{\mathcal{T}}^p(\nabla\cdot\boldsymbol{v})\|_K^2\right\}^{\frac{1}{2}} \leq C\left\{\sum_{K\in\mathcal{T}_{\boldsymbol{a}}} [e_{K,p}(\boldsymbol{v})]^2\right\}^{\frac{1}{2}}.\end{aligned} \tag{4.15}$$

For $p = 0$, we instead apply the Cauchy–Schwarz inequality, the stability of the L^2-projection, the Poincaré inequality on $H^1_*(\omega_{\boldsymbol{a}})$, and $\|\nabla\varphi\|_{\omega_{\boldsymbol{a}}} = 1$ to get

$$\begin{aligned}|(\Pi_{\mathcal{T}}^p(\psi_{\boldsymbol{a}}\nabla_{\mathcal{T}}\cdot(\boldsymbol{v}-\boldsymbol{\tau}_{\mathcal{T}})), \varphi)_{\omega_{\boldsymbol{a}}}| &\leq C\|\nabla_{\mathcal{T}}\cdot(\boldsymbol{v}-\boldsymbol{\tau}_{\mathcal{T}})\|_{\omega_{\boldsymbol{a}}} h_{\omega_{\boldsymbol{a}}}\|\nabla\varphi\|_{\omega_{\boldsymbol{a}}} \\ &\leq C\left\{\sum_{K\in\mathcal{T}_{\boldsymbol{a}}} [e_{K,p}(\boldsymbol{v})]^2\right\}^{\frac{1}{2}},\end{aligned}$$

where C depends only on d and $\kappa_{\mathcal{T}}$.

To bound the second term on the right-hand side of (4.14), we recall the trace inequality

$$\|\varphi\|_F^2 \leq C\left(\|\nabla\varphi\|_K\|\varphi\|_K + h_K^{-1}\|\varphi\|_K^2\right),$$

for any $\varphi \in H^1(K)$ and $F \in \mathcal{F}_K$, where C depends only on d and $\kappa_{\mathcal{T}}$. Combined with the Poincaré inequality on $H^1_*(\omega_{\boldsymbol{a}})$ and $\|\nabla\varphi\|_{\omega_{\boldsymbol{a}}} = 1$, this gives

$$\sum_{F\in\mathcal{F}_{\boldsymbol{a}}^{\mathrm{in}}} |\big(\Pi_F^p(\psi_{\boldsymbol{a}}[\![\boldsymbol{\tau}_{\mathcal{T}}]\!]\cdot\boldsymbol{n}_F), \varphi\big)_F| \leq C\left\{\sum_{F\in\mathcal{F}_{\boldsymbol{a}}^{\mathrm{in}}} h_F\|[\![\boldsymbol{\tau}_{\mathcal{T}}]\!]\cdot\boldsymbol{n}_F\|_F^2\right\}^{\frac{1}{2}}, \tag{4.16}$$

with C depending only on d and $\kappa_{\mathcal{T}}$. Finally, we invoke Lemma 4.6, yielding, for each $F \in \mathcal{F}_{\boldsymbol{a}}^{\mathrm{in}}$,

$$h_F^{1/2}\|[\![\boldsymbol{\tau}_{\mathcal{T}}]\!]\cdot\boldsymbol{n}_F\|_F \leq C \sup_{\substack{w\in H^1(\omega_F)\\ w=0 \text{ on } \partial\omega_F\setminus F\\ \|\nabla w\|_{\omega_F}=1}} ([\![\boldsymbol{\tau}_{\mathcal{T}}]\!]\cdot\boldsymbol{n}_F, w)_F, \tag{4.17}$$

where now the constant C depends on the polynomial degree p in addition to d and $\kappa_{\mathcal{T}}$. Fix $w \in H^1(\omega_F)$ such that $w = 0$ on $\partial\omega_F \setminus F$ and $\|\nabla w\|_{\omega_F} = 1$. By definition, $F \in \mathcal{F}_{\boldsymbol{a}}^{\mathrm{in}}$ means that F is either an internal

face shared by two simplices, or a Neumann boundary face. Then, the zero extension of w to Ω belongs to $H^1_{\Gamma_\mathrm{D}}(\Omega)$. Since $\boldsymbol{v} \in \boldsymbol{H}_{0,\Gamma_\mathrm{N}}(\mathrm{div},\Omega)$, we infer from the definition of the weak divergence that

$$(\nabla\cdot\boldsymbol{v}, w)_{\omega_F} + (\boldsymbol{v}, \nabla w)_{\omega_F} = 0.$$

Consequently, developing $([\![\boldsymbol{\tau}_\mathcal{T}]\!]\cdot\boldsymbol{n}_F, w)_F$ shows that

$$\begin{aligned}
|([\![\boldsymbol{\tau}_\mathcal{T}]\!]\cdot\boldsymbol{n}_F, w)_F| &= |(\nabla_\mathcal{T}\cdot\boldsymbol{\tau}_\mathcal{T}, w)_{\omega_F} + (\boldsymbol{\tau}_\mathcal{T}, \nabla w)_{\omega_F}| \\
&\le |(\nabla_\mathcal{T}\cdot(\boldsymbol{\tau}_\mathcal{T} - \boldsymbol{v}), w - \Pi^p_\mathcal{T} w)_{\omega_F}| + |(\boldsymbol{\tau}_\mathcal{T} - \boldsymbol{v}, \nabla w)_{\omega_F}| \\
&\le \|\nabla_\mathcal{T}\cdot(\boldsymbol{\tau}_\mathcal{T} - \boldsymbol{v})\|_{\omega_F}\|w - \Pi^p_\mathcal{T} w\|_{\omega_F} + \|\boldsymbol{\tau}_\mathcal{T} - \boldsymbol{v}\|_{\omega_F}\|\nabla w\|_{\omega_F} \\
&\le C \sum_{K\in\mathcal{T}_F} \left\{ \|\boldsymbol{v} - \boldsymbol{\tau}_\mathcal{T}\|^2_K + \frac{h_K^2}{(p+1)^2}\|\nabla\cdot(\boldsymbol{v} - \boldsymbol{\tau}_\mathcal{T})\|^2_K \right\}^{1/2},
\end{aligned}$$

owing to the Cauchy–Schwarz inequality, the orthogonality of the L^2-projection, and the Poincaré–Friedrichs inequality $\|w\|_{\omega_F} \le C h_F \|\nabla w\|_{\omega_F}$. Hence, Lemma 4.4 implies that

$$\sum_{F\in\mathcal{F}^{\mathrm{in}}_{\boldsymbol{a}}} \left(\Pi^p_F(\psi_{\boldsymbol{a}}[\![\boldsymbol{\tau}_\mathcal{T}]\!]\cdot\boldsymbol{n}_F), \varphi\right)_F \le C \left\{ \sum_{K\in\mathcal{T}_{\boldsymbol{a}}} [e_{K,p}(\boldsymbol{v})]^2 \right\}^{\frac{1}{2}},$$

where the constant C depends only on d, $\kappa_\mathcal{T}$, and the polynomial degree p via (4.17). Combining these bounds implies (4.12). □

Step 2. Bound on $\|\boldsymbol{v} - P^p_\mathcal{T}(\boldsymbol{v})\|_K$. Let $K \in \mathcal{T}$. In this second and last step, we first show that

$$\|P^p_\mathcal{T}(\boldsymbol{v}) - \boldsymbol{\tau}_\mathcal{T}\|_K \le C \left\{ \sum_{K'\in\mathcal{T}_K} [e_{K',p}(\boldsymbol{v})]^2 \right\}^{\frac{1}{2}}. \tag{4.18}$$

Recalling that $\mathcal{V}_K$ denotes the set of vertices of the element K, using the partition of unity $\sum_{\boldsymbol{a}\in\mathcal{V}_K} \psi_{\boldsymbol{a}}|_K = 1$ and the linearity of $\boldsymbol{I}^p_\mathcal{T}$ as well as definition (3.9) of $P^p_\mathcal{T}(\boldsymbol{v})$ and the fact that $\boldsymbol{\tau}_\mathcal{T} = \boldsymbol{I}^p_\mathcal{T}(\boldsymbol{\tau}_\mathcal{T})$, we find that

$$(P^p_\mathcal{T}(\boldsymbol{v}) - \boldsymbol{\tau}_\mathcal{T})|_K = (P^p_\mathcal{T}(\boldsymbol{v}) - \boldsymbol{I}^p_\mathcal{T}(\boldsymbol{\tau}_\mathcal{T}))|_K = \sum_{\boldsymbol{a}\in\mathcal{V}_K} (\boldsymbol{\sigma}_{\boldsymbol{a}} - \boldsymbol{I}^p_\mathcal{T}(\psi_{\boldsymbol{a}}\boldsymbol{\tau}_\mathcal{T}))|_K.$$

Thus,

$$\|P^p_\mathcal{T}(\boldsymbol{v}) - \boldsymbol{\tau}_\mathcal{T}\|^2_K = \left\| \sum_{\boldsymbol{a}\in\mathcal{V}_K} (\boldsymbol{\sigma}_{\boldsymbol{a}} - \boldsymbol{I}^p_\mathcal{T}(\psi_{\boldsymbol{a}}\boldsymbol{\tau}_\mathcal{T})) \right\|^2_K \le (d+1) \sum_{\boldsymbol{a}\in\mathcal{V}_K} \|\boldsymbol{\sigma}_{\boldsymbol{a}} - \boldsymbol{I}^p_\mathcal{T}(\psi_{\boldsymbol{a}}\boldsymbol{\tau}_\mathcal{T})\|^2_{\omega_{\boldsymbol{a}}},$$

and Lemma 4.7 then yields (4.18). Having obtained (4.18), the main bound (4.11) then follows from the triangle inequality and Lemma 4.4, since

$$\|\boldsymbol{v} - P^p_\mathcal{T}(\boldsymbol{v})\|_K \le \|\boldsymbol{v} - \boldsymbol{\tau}_\mathcal{T}\|_K + \|\boldsymbol{\tau}_\mathcal{T} - P^p_\mathcal{T}(\boldsymbol{v})\|_K \le C \left\{ \sum_{K'\in\mathcal{T}_K} [e_{K',p}(\boldsymbol{v})]^2 \right\}^{\frac{1}{2}}.$$

This completes the proof of the statement (3.12) from Theorem 3.8.

4.3 Proofs of Theorem 3.1 and of Theorem 3.8

Consider an arbitrary function $\boldsymbol{v} \in \boldsymbol{H}_{0,\Gamma_\mathrm{N}}(\mathrm{div},\Omega)$.

Proof of Theorem 3.1. The projection $P^p_\mathcal{T}(\boldsymbol{v})$, whose existence and commutativity properties were shown in Section 4.1 above, satisfies the constraints of the global minimization set in (3.1). Therefore, Theorem 3.1 follows by picking the function $P^p_\mathcal{T}(\boldsymbol{v})$ from the minimization set, applying the local approximation property (3.12), summing over all mesh elements, and invoking the shape-regularity of the mesh which implies that the number of neighbors a mesh cell can have is uniformly bounded from above.

Proof of Theorem 3.8. It remains to prove the statements (3.11), (3.13), and (3.14).

To prove (3.11), we observe that if $\boldsymbol{v} \in \boldsymbol{RTN}_p(\mathcal{T}) \cap \boldsymbol{H}_{0,\Gamma_\mathrm{N}}(\mathrm{div},\Omega)$, then it follows from the definition (3.2) that $e_{K,p}(\boldsymbol{v}) = 0$ for all $K \in \mathcal{T}$, and thus (3.11) follows immediately from (3.12).

To prove (3.13), we observe that, for any $K \in \mathcal{T}$, the triangle inequality yields

$$\|P^p_{\mathcal{T}}(\boldsymbol{v})\|_K \le \|\boldsymbol{v}\|_K + \|\boldsymbol{v} - P^p_{\mathcal{T}}(\boldsymbol{v})\|_K.$$

The first term is trivially contained in the right-hand side of (3.13). Bounding the second one by (3.12), the definition (3.2) of $e_{K,p}(\boldsymbol{v})$ implies that

$$e_{K,p}(\boldsymbol{v}) \le \|\boldsymbol{v}\|_K + \frac{h_K}{(p+1)}\|\nabla\cdot\boldsymbol{v} - \Pi^p_{\mathcal{T}}(\nabla\cdot\boldsymbol{v})\|_K.$$

This shows that (3.13) holds true.

Finally, from (3.13), (3.14) follows immediately since $\frac{h_K}{p+1} \le h_\Omega$ and since both $\|\Pi^p_{\mathcal{T}}(\nabla\cdot\boldsymbol{v})\|_K$ and $\|\nabla\cdot\boldsymbol{v} - \Pi^p_{\mathcal{T}}(\nabla\cdot\boldsymbol{v})\|_K$ are bounded by $\|\nabla\cdot\boldsymbol{v}\|_K$.

4.4 Proof of Theorem 3.4

We now consider $p \ge 1$. The proof of Theorem 3.4 is done in the same spirit as that of Theorem 3.1 in Section 4.3, with a change in the polynomial degree in Definition 3.7 in order to obtain a polynomial-degree-robust constant. Let $\boldsymbol{v} \in \boldsymbol{H}_{0,\Gamma_\mathrm{N}}(\mathrm{div},\Omega)$. As before, in order to show (3.5), it is enough to prove that there exists $\boldsymbol{\sigma}_{\mathcal{T}} \in \boldsymbol{RTN}_p(\mathcal{T}) \cap \boldsymbol{H}_{0,\Gamma_\mathrm{N}}(\mathrm{div},\Omega)$ such that $\nabla\cdot\boldsymbol{\sigma}_{\mathcal{T}} = \Pi^p_{\mathcal{T}}(\nabla\cdot\boldsymbol{v})$ and

$$\|\boldsymbol{v} - \boldsymbol{\sigma}_{\mathcal{T}}\| \le C\left\{\sum_{K\in\mathcal{T}}[e_{K,p-1}(\boldsymbol{v})]^2\right\}^{\frac{1}{2}}, \tag{4.19}$$

where C is a constant depending only on d and $\kappa_{\mathcal{T}}$. To this purpose, we adapt Definition 3.7 as follows.

Definition 4.8 (A simple locally-defined mapping from $\boldsymbol{H}_{0,\Gamma_\mathrm{N}}(\mathrm{div},\Omega)$ to $\boldsymbol{RTN}_p(\mathcal{T}) \cap \boldsymbol{H}_{0,\Gamma_\mathrm{N}}(\mathrm{div},\Omega)$)**.** *Let $\boldsymbol{v} \in \boldsymbol{H}_{0,\Gamma_\mathrm{N}}(\mathrm{div},\Omega)$ be arbitrary. Let $\boldsymbol{\tau}_{\mathcal{T}}$ be defined elementwise by*

$$\boldsymbol{\tau}_{\mathcal{T}}|_K := \underset{\substack{\boldsymbol{v}_K\in\boldsymbol{RTN}_{p-1}(K)\\ \nabla\cdot\boldsymbol{v}_K=\Pi^{p-1}_{\mathcal{T}}(\nabla\cdot\boldsymbol{v})}}{\arg\min} \|\boldsymbol{v} - \boldsymbol{v}_K\|_K \qquad \forall K \in \mathcal{T}. \tag{4.20}$$

For each mesh vertex $\boldsymbol{a} \in \mathcal{V}$, the patchwise contributions $\boldsymbol{\sigma}_{\boldsymbol{a}}$ are now defined as

$$\boldsymbol{\sigma}_{\boldsymbol{a}} := \underset{\substack{\boldsymbol{v}_{\boldsymbol{a}}\in\boldsymbol{V}_p(\omega_{\boldsymbol{a}})\\ \nabla\cdot\boldsymbol{v}_{\boldsymbol{a}}=\Pi^p_{\mathcal{T}}(\psi_{\boldsymbol{a}}\nabla\cdot\boldsymbol{v})+\nabla\psi_{\boldsymbol{a}}\cdot\boldsymbol{\tau}_{\mathcal{T}}}}{\arg\min} \|\boldsymbol{v}_{\boldsymbol{a}} - \psi_{\boldsymbol{a}}\boldsymbol{\tau}_{\mathcal{T}}\|_{\omega_{\boldsymbol{a}}}, \tag{4.21}$$

with the spaces $\boldsymbol{V}_p(\omega_{\boldsymbol{a}})$ still defined in (2.4). Finally, the equilibrated flux reconstruction $\boldsymbol{\sigma}_{\mathcal{T}} \in \boldsymbol{RTN}_p(\mathcal{T}) \cap \boldsymbol{H}_{0,\Gamma_\mathrm{N}}(\mathrm{div},\Omega)$, after extending each $\boldsymbol{\sigma}_{\boldsymbol{a}}$ by zero, is defined as

$$\boldsymbol{\sigma}_{\mathcal{T}} := \sum_{\boldsymbol{a}\in\mathcal{V}} \boldsymbol{\sigma}_{\boldsymbol{a}}. \tag{4.22}$$

Note that the elementwise minimization in (4.20) is done over $(p-1)$-degree Raviart–Thomas–Nédélec spaces, in contrast to (3.7), and the RTN interpolation $\boldsymbol{I}^p_{\mathcal{T}}$ is not used in (4.21), in contrast to (3.8).

Since the orthogonality property (4.2) also holds here, we infer that (4.1) still holds with the above definitions. This in turn gives the necessary compatibility condition yielding existence and uniqueness for the local minimization problems (4.21) in the spirit of Lemma 4.2. Finally, just as in (4.4), we deduce that $\nabla\cdot\boldsymbol{\sigma}_{\mathcal{T}} = \Pi^p_{\mathcal{T}}(\nabla\cdot\boldsymbol{v})$. It thus remains to prove that

$$\|\boldsymbol{\tau}_{\mathcal{T}} - \boldsymbol{\sigma}_{\mathcal{T}}\| \le C\left\{\sum_{K\in\mathcal{T}}[e_{K,p-1}(\boldsymbol{v})]^2\right\}^{\frac{1}{2}}, \tag{4.23}$$

with C only depending on d and $\kappa_{\mathcal{T}}$. Then (4.19) follows from (4.23) by the triangle inequality $\|\boldsymbol{v}-\boldsymbol{\sigma}_{\mathcal{T}}\| \le \|\boldsymbol{v} - \boldsymbol{\tau}_{\mathcal{T}}\| + \|\boldsymbol{\tau}_{\mathcal{T}} - \boldsymbol{\sigma}_{\mathcal{T}}\|$, where the divergence-constrained minimization in $\|\boldsymbol{v} - \boldsymbol{\tau}_{\mathcal{T}}\|_K$ is subordinate to the unconstrained one in $e_{K,p-1}(\boldsymbol{v})$ by Lemma 4.4 applied with $(p-1)$ in place of p.

Lemma 4.9 (Bound on $\boldsymbol{\sigma}_{\boldsymbol{a}}$)**.** *There exists a constant C, depending only on d and $\kappa_{\mathcal{T}}$, such that*

$$\|\boldsymbol{\sigma}_{\boldsymbol{a}} - \psi_{\boldsymbol{a}}\boldsymbol{\tau}_{\mathcal{T}}\|_{\omega_{\boldsymbol{a}}} \le C \left\{ \sum_{K\in\mathcal{T}_{\boldsymbol{a}}} [e_{K,p-1}(\boldsymbol{v})]^2 \right\}^{\frac{1}{2}} \qquad \forall \boldsymbol{a} \in \mathcal{V}. \tag{4.24}$$

Proof. Fix a vertex $\boldsymbol{a} \in \mathcal{V}$. We rely on Lemma 4.5, where we take $\boldsymbol{\tau}_{\boldsymbol{a}} := \psi_{\boldsymbol{a}}\boldsymbol{\tau}_{\mathcal{T}}$ and $g_{\boldsymbol{a}} := \Pi^p_{\mathcal{T}}(\psi_{\boldsymbol{a}}\nabla\cdot\boldsymbol{v}) + \nabla\psi_{\boldsymbol{a}}\cdot\boldsymbol{\tau}_{\mathcal{T}}$ in order to apply it to our construction (4.21) from Definition 4.8. This yields

$$\|\boldsymbol{\sigma}_{\boldsymbol{a}} - \psi_{\boldsymbol{a}}\boldsymbol{\tau}_{\mathcal{T}}\|_{\omega_{\boldsymbol{a}}} \le C \sup_{\substack{v\in H^1_*(\omega_{\boldsymbol{a}})\\ \|\nabla\varphi\|_{\omega_{\boldsymbol{a}}}=1}} \left\{ (g_{\boldsymbol{a}},\varphi)_{\omega_{\boldsymbol{a}}} + (\boldsymbol{\tau}_{\boldsymbol{a}},\nabla\varphi)_{\omega_{\boldsymbol{a}}} \right\}.$$

Let $\varphi \in H^1_*(\omega_{\boldsymbol{a}})$ with $\|\nabla\varphi\|_{\omega_{\boldsymbol{a}}} = 1$ be fixed, where we recall that the space $H^1_*(\omega_{\boldsymbol{a}})$ is defined in (4.7). Then, the product $\psi_{\boldsymbol{a}}\varphi \in H^1_{\Gamma_{\mathrm{D}}}(\Omega)$ for any $\boldsymbol{a} \in \mathcal{V}$ and thus the definition of the weak divergence implies that

$$\big(\boldsymbol{v}, \nabla(\psi_{\boldsymbol{a}}\varphi)\big)_{\omega_{\boldsymbol{a}}} + \big(\nabla\cdot\boldsymbol{v}, \psi_{\boldsymbol{a}}\varphi\big)_{\omega_{\boldsymbol{a}}} = 0.$$

Then, the product rule and the orthogonality of the L^2-projection give

$$\begin{aligned}
&(g_{\boldsymbol{a}},\varphi)_{\omega_{\boldsymbol{a}}} + (\boldsymbol{\tau}_{\boldsymbol{a}},\nabla\varphi)_{\omega_{\boldsymbol{a}}}\\
&= \big(\Pi^p_{\mathcal{T}}(\psi_a\nabla\cdot\boldsymbol{v}),\varphi\big)_{\omega_{\boldsymbol{a}}} + (\nabla\psi_{\boldsymbol{a}}\cdot\boldsymbol{\tau}_{\mathcal{T}},\varphi)_{\omega_{\boldsymbol{a}}} + \big(\psi_{\boldsymbol{a}}\boldsymbol{\tau}_{\mathcal{T}},\nabla\varphi\big)_{\omega_{\boldsymbol{a}}}\\
&= \big(\nabla\cdot\boldsymbol{v}, \psi_{\boldsymbol{a}}\Pi^p_{\mathcal{T}}(\varphi)\big)_{\omega_{\boldsymbol{a}}} + \big(\boldsymbol{\tau}_{\mathcal{T}}, \nabla(\psi_{\boldsymbol{a}}\varphi)\big)_{\omega_{\boldsymbol{a}}}\\
&= \big(\nabla\cdot\boldsymbol{v}, \psi_{\boldsymbol{a}}(\Pi^p_{\mathcal{T}}(\varphi)-\varphi)\big)_{\omega_{\boldsymbol{a}}} + \big(\boldsymbol{\tau}_{\mathcal{T}} - \boldsymbol{v}, \nabla(\psi_{\boldsymbol{a}}\varphi)\big)_{\omega_{\boldsymbol{a}}}\\
&= \big(\psi_{\boldsymbol{a}}(\nabla\cdot\boldsymbol{v} - \Pi^{p-1}_{\mathcal{T}}(\nabla\cdot\boldsymbol{v})), \Pi^p_{\mathcal{T}}(\varphi)-\varphi\big)_{\omega_{\boldsymbol{a}}} + \big(\boldsymbol{\tau}_{\mathcal{T}} - \boldsymbol{v}, \nabla(\psi_{\boldsymbol{a}}\varphi)\big)_{\omega_{\boldsymbol{a}}},
\end{aligned}$$

since $\psi_{\boldsymbol{a}}\Pi^{p-1}_{\mathcal{T}}(\nabla\cdot\boldsymbol{v})$ is a piecewise polynomial of degree at most p. Therefore, we have

$$\begin{aligned}
|(g_{\boldsymbol{a}},\varphi)_{\omega_{\boldsymbol{a}}} + (\boldsymbol{\tau}_{\boldsymbol{a}},\nabla\varphi)_{\omega_{\boldsymbol{a}}}| &\le C \sum_{K\in\mathcal{T}_{\boldsymbol{a}}} \Big[\frac{h_K}{p}\|\nabla\cdot\boldsymbol{v} - \Pi^{p-1}_{\mathcal{T}}(\nabla\cdot\boldsymbol{v})\|_K\Big]\|\nabla\varphi\|_{\omega_{\boldsymbol{a}}}\\
&\quad + \|\boldsymbol{v} - \boldsymbol{\tau}_{\mathcal{T}}\|_{\omega_{\boldsymbol{a}}}\|\nabla(\psi_{\boldsymbol{a}}\varphi)\|_{\omega_{\boldsymbol{a}}}\\
&\le C\,(1 + \|\nabla(\psi_{\boldsymbol{a}}\varphi)\|_{\omega_{\boldsymbol{a}}}) \left\{ \sum_{K\in\mathcal{T}_{\boldsymbol{a}}} [e_{K,p-1}(\boldsymbol{v})]^2 \right\}^{\frac{1}{2}},
\end{aligned}$$

where we used $\|\psi_{\boldsymbol{a}}\|_{\infty,\omega_{\boldsymbol{a}}} = 1$, the hp approximation bound $\|\varphi - \Pi^p_{\mathcal{T}}(\varphi)\|_K \le C\frac{h_K}{p+1}\|\nabla\varphi\|_K \le C\frac{h_K}{p}\|\nabla\varphi\|_K$, the Cauchy–Schwarz inequality, the scaling $\|\nabla\varphi\|_{\omega_{\boldsymbol{a}}} = 1$, and Lemma 4.4. Finally, bound (4.24) follows from the inequality $\|\nabla(\psi_{\boldsymbol{a}}\varphi)\|_{\omega_{\boldsymbol{a}}} \le C\|\nabla\varphi\|_{\omega_{\boldsymbol{a}}} \le C$ for all $\varphi \in H^1_*(\omega_{\boldsymbol{a}})$, owing to the Poincaré inequality on $H^1_*(\omega_{\boldsymbol{a}})$ and $\|\nabla\varphi\|_{\omega_{\boldsymbol{a}}} = 1$. □

Finally, we obtain (4.23) from Lemma 4.9 and the estimate

$$\|\boldsymbol{\sigma}_{\mathcal{T}} - \boldsymbol{\tau}_{\mathcal{T}}\|^2 = \sum_{K\in\mathcal{T}} \Big\| \sum_{\boldsymbol{a}\in\mathcal{V}_K} \big(\boldsymbol{\sigma}_{\boldsymbol{a}} - \psi_{\boldsymbol{a}}\boldsymbol{\tau}_{\mathcal{T}}\big) \Big\|_K^2 \le (d+1) \sum_{\boldsymbol{a}\in\mathcal{V}} \|\boldsymbol{\sigma}_{\boldsymbol{a}} - \psi_{\boldsymbol{a}}\boldsymbol{\tau}_{\mathcal{T}}\|^2_{\omega_{\boldsymbol{a}}}.$$

As explained above, (4.23) then implies (3.5) and completes the proof of Theorem 3.4.

4.5 Proof of Theorem 3.5

The proof of Theorem 3.5 hinges on the bounds from Theorems 3.1 and 3.4. dy Recall the definitions (3.1) of $E_{\mathcal{T},p}(\boldsymbol{v})$ and (3.2) of $e_{K,p}(\boldsymbol{v})$. Recall also the notation $\delta_{s<1} := 1$ if $s < 1$ and $\delta_{s<1} := 0$ if $s \ge 1$. To prove (3.6), we proceed in two steps.

Step 1. We first suppose that $p \le s$ and let $t := \min(s, p+1)$. Here, we will employ Theorem 3.1. Since $\mathcal{P}_p(K;\mathbb{R}^d) \subset \boldsymbol{RTN}_p(K)$, well-known hp-approximation bounds, see e.g. [2, Lemma 4.1], imply that

$$[e_{K,p}(\boldsymbol{v})]^2 \le C\Big\{ \Big[\frac{h_K^t}{(p+1)^s}\|v\|_{\boldsymbol{H}^s(K)}\Big]^2 + \delta_{s<1}\Big[\frac{h_K}{p+1}\|\nabla\cdot\boldsymbol{v}\|_K\Big]^2 \Big\}, \tag{4.25}$$

for each $K \in \mathcal{T}$, with C depending only on s, d, $\kappa_{\mathcal{T}}$. Note that for $s < 1$, we applied here the trivial bound $\|\nabla\cdot\boldsymbol{v} - \Pi^p_{\mathcal{T}}(\nabla\cdot\boldsymbol{v})\|_K \le \|\nabla\cdot\boldsymbol{v}\|_K$ as $\boldsymbol{v}|_K \in \boldsymbol{H}^s(K)$ is insufficient to improve the bound on the error of the divergence. Combining (4.25) with the first bound in (3.3) of Theorem 3.1 then implies that there exists a constant $C_{s,d,\kappa_{\mathcal{T}},p}$ depending only on s, d, $\kappa_{\mathcal{T}}$, and p, such that

$$[E_{\mathcal{T},p}(\boldsymbol{v})]^2 \le C_{s,d,\kappa_{\mathcal{T}},p} \sum_{K\in\mathcal{T}} \left\{ \Big[\frac{h_K^t}{(p+1)^s}\|v\|_{\boldsymbol{H}^s(K)}\Big]^2 + \delta_{s<1}\Big[\frac{h_K}{p+1}\|\nabla\cdot\boldsymbol{v}\|_K\Big]^2 \right\}.$$

Define then the constant $C^\star_{s,d,\kappa_{\mathcal{T}}} := \max_{0\le p\le s} C_{s,d,\kappa_{\mathcal{T}},p}$, so that, for all $p \le s$,

$$[E_{\mathcal{T},p}(\boldsymbol{v})]^2 \le C^\star_{s,d,\kappa_{\mathcal{T}}} \sum_{K\in\mathcal{T}} \left\{ \Big[\frac{h_K^t}{(p+1)^s}\|v\|_{\boldsymbol{H}^s(K)}\Big]^2 + \delta_{s<1}\Big[\frac{h_K}{p+1}\|\nabla\cdot\boldsymbol{v}\|_K\Big]^2 \right\}.$$

This implies (3.6) for any $p \le s$ with constant $C = C^\star_{s,d,\kappa_{\mathcal{T}}}$.

Step 2. Now consider the case $p > s$; since p is an integer, this implies that $p \ge 1$. Here we rely on Theorem 3.4. The approximation bounds, similarly to in (4.25), imply that there exists a constant C, depending only on s, d, and $\kappa_{\mathcal{T}}$, such that

$$[e_{K,p-1}(\boldsymbol{v})]^2 \le C\left\{ \Big[\frac{h_K^s}{p^s}\|\boldsymbol{v}\|_{\boldsymbol{H}^s(K)}\Big]^2 + \delta_{s<1}\Big[\frac{h_K}{p}\|\nabla\cdot\boldsymbol{v}\|_K\Big]^2 \right\},$$

for all $K \in \mathcal{T}$. Note that $p+1 \le 2p$ for all $p \ge 1$, so that the terms p^s in the denominators above can be replaced by $(p+1)^s$ at the cost of an extra s-dependent constant, and similarly for $1/p \le 2/(p+1)$. Hence, the inequality (3.5) of Theorem 3.4 and summation over the elements of $\mathcal{T}$ shows that there exists a constant $C^\sharp_{s,d,\kappa_{\mathcal{T}}}$ depending only on s, d, and $\kappa_{\mathcal{T}}$ such that (3.6) holds with constant $C = C^\sharp_{s,d,\kappa_{\mathcal{T}}}$ for all $p > s$.

Conclusion. Combining Steps 1 and 2 shows that (3.6) holds for general $s \ge 0$ and $p \ge 0$ with a constant C that can be taken as $\max\{C^\star_{s,d,\kappa_{\mathcal{T}}}, C^\sharp_{s,d,\kappa_{\mathcal{T}}}\}$, which then depends only on s, d, and $\kappa_{\mathcal{T}}$.

5 Application to a priori error estimates

In this section we show how to apply the results of Section 3 to the a priori error analysis of mixed finite element methods and least-squares mixed finite element methods for a model diffusion problem.

5.1 Mixed finite element methods

Let us consider the dual mixed finite element method for the Poisson model problem, following Raviart and Thomas [38], Nédélec [35], Roberts and Thomas [39], or Boffi *et al.* [7]. Let $f \in L^2(\Omega)$ and $\Gamma_{\mathrm{N}} = \emptyset$ for simplicity, so that $\boldsymbol{H}_{0,\Gamma_{\mathrm{N}}}(\mathrm{div},\Omega)$ becomes $\boldsymbol{H}(\mathrm{div},\Omega)$. Consider the Laplace problem of finding $u : \Omega \to \mathbb{R}$ such that

$$-\Delta u = f \qquad \text{in } \Omega, \tag{5.1a}$$
$$u = 0 \qquad \text{on } \partial\Omega. \tag{5.1b}$$

The primal weak formulation of (5.1) reads: find $u \in H^1_0(\Omega)$ such that

$$(\nabla u, \nabla v) = (f, v) \qquad \forall v \in H^1_0(\Omega). \tag{5.2}$$

The dual weak formulation of (5.1) then reads: find $\boldsymbol{\sigma} \in \boldsymbol{H}(\mathrm{div},\Omega)$ and $u \in L^2(\Omega)$ such that

$$(\boldsymbol{\sigma}, \boldsymbol{v}) - (u, \nabla\cdot\boldsymbol{v}) = 0 \qquad \forall \boldsymbol{v} \in \boldsymbol{H}(\mathrm{div},\Omega), \tag{5.3a}$$
$$(\nabla\cdot\boldsymbol{\sigma}, q) = (f, q) \qquad \forall q \in L^2(\Omega). \tag{5.3b}$$

Classically, u from (5.2) and (5.3) coincide and $\boldsymbol{\sigma} = -\nabla u$. The dual mixed finite element method of order $p \ge 0$ for the problem (5.3) then looks for the pair $\boldsymbol{\sigma}_{\mathrm{M}} \in \boldsymbol{RTN}_p(\mathcal{T}) \cap \boldsymbol{H}(\mathrm{div},\Omega)$ and $u_{\mathrm{M}} \in \mathcal{P}_p(\mathcal{T})$ such that

$$(\boldsymbol{\sigma}_{\mathrm{M}}, \boldsymbol{v}_{\mathcal{T}}) - (u_{\mathrm{M}}, \nabla\cdot\boldsymbol{v}_{\mathcal{T}}) = 0 \qquad \forall \boldsymbol{v}_{\mathcal{T}} \in \boldsymbol{RTN}_p(\mathcal{T}) \cap \boldsymbol{H}(\mathrm{div},\Omega), \tag{5.4a}$$
$$(\nabla\cdot\boldsymbol{\sigma}_{\mathrm{M}}, q_{\mathcal{T}}) = (f, q_{\mathcal{T}}) \qquad \forall q_{\mathcal{T}} \in \mathcal{P}_p(\mathcal{T}). \tag{5.4b}$$

It is immediate to check from (5.3b) and (5.4b) that $\nabla\cdot\boldsymbol{\sigma}_{\mathrm{M}} = \Pi^p_{\mathcal{T}}(\nabla\cdot\boldsymbol{\sigma})$. Furthermore, the following *a priori* error characterization is classical, cf. [7]. We include its proof to highlight the precise arguments.

Lemma 5.1 (A priori bound for mixed finite element methods)**.** *Let $\boldsymbol{\sigma}_{\mathrm{M}}$ be the first component of the dual mixed finite solution solving (5.4), approximating $\boldsymbol{\sigma}$ from (5.3). Then*

$$\|\boldsymbol{\sigma}-\boldsymbol{\sigma}_{\mathrm{M}}\| = \min_{\substack{\boldsymbol{v}_{\mathcal{T}}\in\boldsymbol{RTN}_p(\mathcal{T})\cap\boldsymbol{H}(\mathrm{div},\Omega)\\ \nabla\cdot\boldsymbol{v}_{\mathcal{T}}=\Pi^p_{\mathcal{T}}(\nabla\cdot\boldsymbol{\sigma})}} \|\boldsymbol{\sigma}-\boldsymbol{v}_{\mathcal{T}}\|.$$

Proof. Subtracting (5.4a) from (5.3a), we have

$$(\boldsymbol{\sigma}-\boldsymbol{\sigma}_{\mathrm{M}},\boldsymbol{v}_{\mathcal{T}}) - (u-u_{\mathrm{M}},\nabla\cdot\boldsymbol{v}_{\mathcal{T}}) = 0 \quad \forall \boldsymbol{v}_{\mathcal{T}}\in\boldsymbol{RTN}_p(\mathcal{T})\cap\boldsymbol{H}(\mathrm{div},\Omega). \tag{5.5}$$

Let $\boldsymbol{\sigma}_{\mathcal{T}}\in\boldsymbol{RTN}_p(\mathcal{T})\cap\boldsymbol{H}(\mathrm{div},\Omega)$ be such that $\nabla\cdot\boldsymbol{\sigma}_{\mathcal{T}}=\Pi^p_{\mathcal{T}}(\nabla\cdot\boldsymbol{\sigma})$. Taking $\boldsymbol{v}_{\mathcal{T}}=\boldsymbol{\sigma}_{\mathcal{T}}-\boldsymbol{\sigma}_{\mathrm{M}}$ in (5.5), we obtain, since $\nabla\cdot\boldsymbol{v}_{\mathcal{T}}=0$,

$$(\boldsymbol{\sigma}-\boldsymbol{\sigma}_{\mathrm{M}},\boldsymbol{\sigma}_{\mathcal{T}}-\boldsymbol{\sigma}_{\mathrm{M}}) = 0.$$

Now clearly

$$\|\boldsymbol{\sigma}-\boldsymbol{\sigma}_{\mathrm{M}}\|^2 = (\boldsymbol{\sigma}-\boldsymbol{\sigma}_{\mathrm{M}},\boldsymbol{\sigma}-\boldsymbol{\sigma}_{\mathrm{M}}) = (\boldsymbol{\sigma}-\boldsymbol{\sigma}_{\mathrm{M}},\boldsymbol{\sigma}-\boldsymbol{\sigma}_{\mathcal{T}}) \le \|\boldsymbol{\sigma}-\boldsymbol{\sigma}_{\mathrm{M}}\|\,\|\boldsymbol{\sigma}-\boldsymbol{\sigma}_{\mathcal{T}}\|,$$

and hence $\|\boldsymbol{\sigma}-\boldsymbol{\sigma}_{\mathrm{M}}\|\le\|\boldsymbol{\sigma}-\boldsymbol{\sigma}_{\mathcal{T}}\|$. Since $\boldsymbol{\sigma}_{\mathcal{T}}$ is arbitrary subject to the divergence constraint and can be taken as $\boldsymbol{\sigma}_{\mathrm{M}}$, we obtain the assertion. □

Thus, $\|\boldsymbol{\sigma}-\boldsymbol{\sigma}_{\mathrm{M}}\|$ can be readily estimated by using Theorems 3.1, 3.4, and 3.5.

5.2 Least-squares mixed finite element methods

In this subsection, we showcase the application of our results to the least-squares mixed finite element method discussed in Pehlivanov *et al.* [37], Cai and Ku [10], and Ku [31], see also the references therein.

Let again $\Gamma_{\mathrm{N}}=\emptyset$ for simplicity and $f\in L^2(\Omega)$. Let $\boldsymbol{\sigma}\in\boldsymbol{H}(\mathrm{div},\Omega)$ and $u\in H^1_0(\Omega)$ be such that

$$(\boldsymbol{\sigma},u) \coloneqq \arg\min_{(\boldsymbol{p},v)\in\boldsymbol{H}(\mathrm{div},\Omega)\times H^1_0(\Omega)} \left\{h_\Omega^2\|\nabla\cdot\boldsymbol{p}-f\|^2+\|\boldsymbol{p}+\nabla v\|^2\right\},$$

where we recall that h_Ω is a length scale equal to the diameter of Ω. Then $\boldsymbol{\sigma}\in\boldsymbol{H}(\mathrm{div},\Omega)$ and $u\in H^1_0(\Omega)$ solve the following system of equations:

$$(\boldsymbol{\sigma}+\nabla u,\nabla v) = 0 \qquad \forall v\in H^1_0(\Omega), \tag{5.6a}$$

$$h_\Omega^2(\nabla\cdot\boldsymbol{\sigma},\nabla\cdot\boldsymbol{p}) + (\boldsymbol{\sigma}+\nabla u,\boldsymbol{p}) = h_\Omega^2(f,\nabla\cdot\boldsymbol{p}) \qquad \forall\boldsymbol{p}\in\boldsymbol{H}(\mathrm{div},\Omega). \tag{5.6b}$$

Again, $\boldsymbol{\sigma}$ and u coincide with the solutions of (5.2) and (5.3). Let $p\ge 0$ and $q\ge 1$ denote two fixed polynomial degrees. The least-squares mixed finite element method for the problem (5.6) consists of finding $\boldsymbol{\sigma}_{\mathrm{LS}}\in\boldsymbol{RTN}_p(\mathcal{T})\cap\boldsymbol{H}(\mathrm{div},\Omega)$ and $u_{\mathrm{LS}}\in\mathcal{P}^q(\mathcal{T})\cap H^1_0(\Omega)$ such that

$$(\boldsymbol{\sigma}_{\mathrm{LS}}+\nabla u_{\mathrm{LS}},\nabla v_{\mathcal{T}}) = 0 \qquad \forall v_{\mathcal{T}}\in\mathcal{P}^q(\mathcal{T})\cap H^1_0(\Omega), \tag{5.7a}$$

$$h_\Omega^2(\nabla\cdot\boldsymbol{\sigma}_{\mathrm{LS}},\nabla\cdot\boldsymbol{p}_{\mathcal{T}}) + (\boldsymbol{\sigma}_{\mathrm{LS}}+\nabla u_{\mathrm{LS}},\boldsymbol{p}_{\mathcal{T}}) = h_\Omega^2(f,\nabla\cdot\boldsymbol{p}_{\mathcal{T}}) \ \forall\boldsymbol{p}_{\mathcal{T}}\in\boldsymbol{RTN}_p(\mathcal{T})\cap\boldsymbol{H}(\mathrm{div},\Omega). \tag{5.7b}$$

Similarly to Lemma 5.1, we can obtain the following *a priori* error characterization.

Lemma 5.2 (A priori bound for least-squares mixed finite element methods)**.** *Let $(\boldsymbol{\sigma}_{\mathrm{LS}},u_{\mathrm{LS}})$ be the least-squares mixed finite solution pair solving (5.7), approximating $(\boldsymbol{\sigma},u)$ from (5.6). Then there exists a constant C, only depending on the domain Ω, such that*

$$\begin{aligned}&\|\boldsymbol{\sigma}-\boldsymbol{\sigma}_{\mathrm{LS}}\|+\|\nabla(u-u_{\mathrm{LS}})\|\\ &\le C\left(\min_{\substack{\boldsymbol{v}_{\mathcal{T}}\in\boldsymbol{RTN}_p(\mathcal{T})\cap\boldsymbol{H}(\mathrm{div},\Omega)\\ \nabla\cdot\boldsymbol{v}_{\mathcal{T}}=\Pi^p_{\mathcal{T}}(\nabla\cdot\boldsymbol{\sigma})}}\|\boldsymbol{\sigma}-\boldsymbol{v}_{\mathcal{T}}\| + \min_{v_{\mathcal{T}}\in\mathcal{P}^q(\mathcal{T})\cap H^1_0(\Omega)}\|\nabla(u-v_{\mathcal{T}})\|\right).\end{aligned}$$

Proof. Define the bilinear form $\mathcal{A}$ on $(\boldsymbol{H}(\mathrm{div},\Omega)\times H^1_0(\Omega))\times(\boldsymbol{H}(\mathrm{div},\Omega)\times H^1_0(\Omega))$ by

$$\mathcal{A}(\boldsymbol{\sigma},u;\boldsymbol{p},v) \coloneqq (\boldsymbol{\sigma}+\nabla u,\nabla v) + h_\Omega^2(\nabla\cdot\boldsymbol{\sigma},\nabla\cdot\boldsymbol{p}) + (\boldsymbol{\sigma}+\nabla u,\boldsymbol{p}).$$

We have the following orthogonality from (5.6) and (5.7):

$$\mathcal{A}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_{\mathrm{LS}}, u - u_{\mathrm{LS}}; \boldsymbol{p}_{\mathcal{T}}, v_{\mathcal{T}}) = 0 \tag{5.8}$$

for all $\boldsymbol{p}_{\mathcal{T}} \in \boldsymbol{RTN}_p(\mathcal{T}) \cap \boldsymbol{H}(\mathrm{div}, \Omega)$ and for all $v_{\mathcal{T}} \in \mathcal{P}^q(\mathcal{T}) \cap H_0^1(\Omega)$. Moreover, the following coercivity is known from [37]: there exists a constant C such that

$$\mathcal{A}(\boldsymbol{p}, v; \boldsymbol{p}, v) \geq C \left(\|\boldsymbol{p}\|^2 + h_\Omega^2 \|\nabla\cdot\boldsymbol{p}\|^2 + \|\nabla v\|^2 \right) \qquad \forall (\boldsymbol{p}, v) \in \boldsymbol{H}(\mathrm{div}, \Omega) \times H_0^1(\Omega)). \tag{5.9}$$

Indeed, owing to the Cauchy–Schwarz and Young inequalities, we have, for any $0 < \varepsilon < 2$,

$$\begin{aligned}
&\mathcal{A}(\boldsymbol{p}, v; \boldsymbol{p}, v) \\
&= \|\nabla v\|^2 + (2-\varepsilon)(\boldsymbol{p}, \nabla v) + h_\Omega^2 \|\nabla\cdot\boldsymbol{p}\|^2 + \|\boldsymbol{p}\|^2 + \varepsilon(\nabla v, \boldsymbol{p}) \\
&\geq \|\boldsymbol{p}\|^2 + \|\nabla v\|^2 - \frac{2-\varepsilon}{2}(\|\boldsymbol{p}\|^2 + \|\nabla v\|^2) + h_\Omega^2 \|\nabla\cdot\boldsymbol{p}\|^2 - \varepsilon \|\nabla\cdot\boldsymbol{p}\| \|v\| \\
&\geq \frac{\varepsilon}{2}(\|\boldsymbol{p}\|^2 + \|\nabla v\|^2) + h_\Omega^2 \|\nabla\cdot\boldsymbol{p}\|^2 - \varepsilon C_{\mathrm{PF}} h_\Omega \left(C_{\mathrm{PF}} h_\Omega \|\nabla\cdot\boldsymbol{p}\|^2 + \frac{1}{4 C_{\mathrm{PF}} h_\Omega} \|\nabla v\|^2 \right) \\
&= \frac{\varepsilon}{2} \|\boldsymbol{p}\|^2 + \frac{\varepsilon}{4} \|\nabla v\|^2 + \|\nabla\cdot\boldsymbol{p}\|^2 (h_\Omega^2 - \varepsilon C_{\mathrm{PF}}^2 h_\Omega^2),
\end{aligned}$$

where we have also employed the Green theorem $(\nabla v, \boldsymbol{p}) = -(\nabla\cdot\boldsymbol{p}, v)$ and the Poincaré–Friedrichs inequality $\|v\| \leq C_{\mathrm{PF}} h_\Omega \|\nabla v\|$ (here h_Ω is the diameter of Ω and $C_{\mathrm{PF}} \leq 1$ a generic constant). The assertion (5.9) follows by choosing, e.g., $\varepsilon = h_\Omega^2/(2C_{\mathrm{PF}}^2 h_\Omega^2)$. Note that, fixing $h_\Omega = h_\Omega$ and employing $C_{\mathrm{PF}} = 1$, the constant in (5.9) can be taken as $1/8$.

Let now $\boldsymbol{v}_{\mathcal{T}} \in \boldsymbol{RTN}_p(\mathcal{T}) \cap \boldsymbol{H}(\mathrm{div}, \Omega)$ be such that $\nabla\cdot\boldsymbol{v}_{\mathcal{T}} = \Pi_{\mathcal{T}}^p(\nabla\cdot\boldsymbol{\sigma})$ and $v_{\mathcal{T}} \in \mathcal{P}^q(\mathcal{T}) \cap H_0^1(\Omega)$ be an arbitrary function. Set $q_{\mathcal{T}} = v_{\mathcal{T}} - u_{\mathrm{LS}}$ and $\boldsymbol{p}_{\mathcal{T}} = \boldsymbol{v}_{\mathcal{T}} - \boldsymbol{\sigma}_{\mathrm{LS}}$. Then using (5.8) and (5.9), we find

$$\begin{aligned}
C \left(\|\boldsymbol{p}_{\mathcal{T}}\|^2 + \|\nabla q_{\mathcal{T}}\|^2 \right) &\leq \mathcal{A}(\boldsymbol{v}_{\mathcal{T}} - \boldsymbol{\sigma}_{\mathrm{LS}}, v_{\mathcal{T}} - u_{\mathrm{LS}}; \boldsymbol{p}_{\mathcal{T}}, q_{\mathcal{T}}) \\
&= \mathcal{A}(\boldsymbol{v}_{\mathcal{T}} - \boldsymbol{\sigma}, v_{\mathcal{T}} - u; \boldsymbol{p}_{\mathcal{T}}, q_{\mathcal{T}}) \\
&= (\boldsymbol{v}_{\mathcal{T}} - \boldsymbol{\sigma} + \nabla(v_{\mathcal{T}} - u), \nabla q_{\mathcal{T}}) + h_\Omega^2 (\nabla\cdot(\boldsymbol{v}_{\mathcal{T}} - \boldsymbol{\sigma}), \nabla\cdot\boldsymbol{p}_{\mathcal{T}}) \\
&\quad + (\boldsymbol{v}_{\mathcal{T}} - \boldsymbol{\sigma} + \nabla(v_{\mathcal{T}} - u), \boldsymbol{p}_{\mathcal{T}}).
\end{aligned}$$

Since $\nabla\cdot\boldsymbol{v}_{\mathcal{T}} = \Pi_{\mathcal{T}}^p(\nabla\cdot\boldsymbol{\sigma})$ and $\nabla\cdot\boldsymbol{p}_{\mathcal{T}} \in \mathcal{P}^p(\mathcal{T})$, we have $(\nabla\cdot(\boldsymbol{v}_{\mathcal{T}} - \boldsymbol{\sigma}), \nabla\cdot\boldsymbol{p}_{\mathcal{T}}) = 0$. Using the Cauchy–Schwarz and the Young inequality, we then obtain

$$\|\boldsymbol{p}_{\mathcal{T}}\|^2 + \|\nabla q_{\mathcal{T}}\|^2 \leq C \left(\|\boldsymbol{\sigma} - \boldsymbol{v}_{\mathcal{T}}\|^2 + \|\nabla(u - v_{\mathcal{T}})\|^2 \right),$$

which proves the claim owing to the triangle inequality and since $\boldsymbol{v}_{\mathcal{T}}$ and $v_{\mathcal{T}}$ are arbitrary. □

The two terms in the error bound from Lemma 5.2 are uncoupled. For the first one, we can again straightforwardly use Theorems 3.1, 3.4, and 3.5. For the second one, the result of Veeser [43] yields

$$\min_{v_{\mathcal{T}} \in \mathcal{P}^q(\mathcal{T}) \cap H_0^1(\Omega)} \|\nabla(u - v_{\mathcal{T}})\|^2 \leq C \sum_{K \in \mathcal{T}} \min_{q_K \in \mathcal{P}^q(K)} \|\nabla(u - q_K)\|_K^2,$$

where the constant C depends only on the space dimension d, the shape-regularity parameter $\kappa_{\mathcal{T}}$ of $\mathcal{T}$, and the polynomial degree q, which is again optimal.

Finally, a localized estimate for the error $\nabla\cdot(\boldsymbol{\sigma} - \boldsymbol{\sigma}_{\mathrm{LS}})$ follows by the combination of the above results with the following lemma.

Lemma 5.3 (A priori bound on the divergence for least-squares mixed finite element methods)**.** *Let* $(\boldsymbol{\sigma}_{\mathrm{LS}}, u_{\mathrm{LS}})$ *be the least-squares mixed finite solution solving* (5.7)*, approximating* $(\boldsymbol{\sigma}, u)$ *from* (5.6)*. Then*

$$\begin{aligned}
h_\Omega^2 \|\nabla\cdot(\boldsymbol{\sigma} - \boldsymbol{\sigma}_{\mathrm{LS}})\|^2 \leq h_\Omega^2 \|\nabla\cdot\boldsymbol{\sigma} - \Pi_{\mathcal{T}}^p(\nabla\cdot\boldsymbol{\sigma})\|^2 + \|\nabla(u - u_{\mathrm{LS}})\|^2 \\
+ \min_{\substack{\boldsymbol{v}_{\mathcal{T}} \in \boldsymbol{RTN}_p(\mathcal{T}) \cap \boldsymbol{H}(\mathrm{div}, \Omega) \\ \nabla\cdot\boldsymbol{v}_{\mathcal{T}} = \Pi_{\mathcal{T}}^p(\nabla\cdot\boldsymbol{\sigma})}} \|\boldsymbol{\sigma} - \boldsymbol{v}_{\mathcal{T}}\|^2.
\end{aligned}$$

Proof. Again let $\boldsymbol{\sigma}_{\mathcal{T}} \in \boldsymbol{RTN}_p(\mathcal{T}) \cap \boldsymbol{H}(\mathrm{div},\Omega)$ be such that $\nabla\cdot\boldsymbol{\sigma}_{\mathcal{T}} = \Pi^p_{\mathcal{T}}(\nabla\cdot\boldsymbol{\sigma})$. Using (5.6b) and (5.7b), we have

$$
\begin{aligned}
&h_\Omega^2\|\nabla\cdot(\boldsymbol{\sigma}-\boldsymbol{\sigma}_{\mathrm{LS}})\|^2\\
&= h_\Omega^2(\nabla\cdot(\boldsymbol{\sigma}-\boldsymbol{\sigma}_{\mathrm{LS}}),\nabla\cdot(\boldsymbol{\sigma}-\boldsymbol{\sigma}_{\mathrm{LS}}))\\
&= h_\Omega^2(\nabla\cdot(\boldsymbol{\sigma}-\boldsymbol{\sigma}_{\mathrm{LS}}),\nabla\cdot(\boldsymbol{\sigma}-\boldsymbol{\sigma}_{\mathcal{T}})) + h_\Omega^2(\nabla\cdot(\boldsymbol{\sigma}-\boldsymbol{\sigma}_{\mathrm{LS}}),\nabla\cdot(\boldsymbol{\sigma}_{\mathcal{T}}-\boldsymbol{\sigma}_{\mathrm{LS}}))\\
&= h_\Omega^2(\nabla\cdot(\boldsymbol{\sigma}-\boldsymbol{\sigma}_{\mathrm{LS}}),\nabla\cdot(\boldsymbol{\sigma}-\boldsymbol{\sigma}_{\mathcal{T}})) + h_\Omega^2(\nabla\cdot(\boldsymbol{\sigma}-\boldsymbol{\sigma}_{\mathrm{LS}}),\nabla\cdot(\boldsymbol{\sigma}_{\mathcal{T}}-\boldsymbol{\sigma}_{\mathrm{LS}}))\\
&\quad - \mathcal{A}(\boldsymbol{\sigma}-\boldsymbol{\sigma}_{\mathrm{LS}}, u-u_{\mathrm{LS}};\boldsymbol{\sigma}_{\mathcal{T}}-\boldsymbol{\sigma}_{\mathrm{LS}},0)\\
&= h_\Omega^2(\nabla\cdot(\boldsymbol{\sigma}-\boldsymbol{\sigma}_{\mathrm{LS}}),\nabla\cdot(\boldsymbol{\sigma}-\boldsymbol{\sigma}_{\mathcal{T}})) - ((\boldsymbol{\sigma}-\boldsymbol{\sigma}_{\mathrm{LS}})+\nabla(u-u_{\mathrm{LS}}),\boldsymbol{\sigma}_{\mathcal{T}}-\boldsymbol{\sigma}_{\mathrm{LS}})\\
&= h_\Omega^2(\nabla\cdot(\boldsymbol{\sigma}-\boldsymbol{\sigma}_{\mathcal{T}}),\nabla\cdot(\boldsymbol{\sigma}-\boldsymbol{\sigma}_{\mathcal{T}})) - ((\boldsymbol{\sigma}-\boldsymbol{\sigma}_{\mathrm{LS}})+\nabla(u-u_{\mathrm{LS}}),\boldsymbol{\sigma}_{\mathcal{T}}-\boldsymbol{\sigma}_{\mathrm{LS}}),
\end{aligned}
$$

where we used that $(\nabla\cdot\boldsymbol{\sigma}_{\mathrm{LS}},\nabla\cdot(\boldsymbol{\sigma}-\boldsymbol{\sigma}_{\mathcal{T}})) = (\nabla\cdot\boldsymbol{\sigma}_{\mathcal{T}},\nabla\cdot(\boldsymbol{\sigma}-\boldsymbol{\sigma}_{\mathcal{T}})) = 0$ since both $\nabla\cdot\boldsymbol{\sigma}_{\mathrm{LS}}$ and $\nabla\cdot\boldsymbol{\sigma}_{\mathcal{T}}$ belong to $\mathcal{P}_p(\mathcal{T})$ in the last equality. Adding and subtracting $\boldsymbol{\sigma}$ in the second term on the right-hand side above and applying the Cauchy–Schwarz and Young inequalities implies that

$$
\begin{aligned}
&h_\Omega^2\|\nabla\cdot(\boldsymbol{\sigma}-\boldsymbol{\sigma}_{\mathrm{LS}})\|^2 + \|\boldsymbol{\sigma}-\boldsymbol{\sigma}_{\mathrm{LS}}\|^2\\
&= h_\Omega^2\|\nabla\cdot\boldsymbol{\sigma} - \Pi^p_{\mathcal{T}}(\nabla\cdot\boldsymbol{\sigma})\|^2 - (\nabla(u-u_{\mathrm{LS}}),\boldsymbol{\sigma}-\boldsymbol{\sigma}_{\mathrm{LS}})\\
&\quad - (\nabla(u-u_{\mathrm{LS}}),\boldsymbol{\sigma}_{\mathcal{T}}-\boldsymbol{\sigma}) - (\boldsymbol{\sigma}-\boldsymbol{\sigma}_{\mathrm{LS}},\boldsymbol{\sigma}_{\mathcal{T}}-\boldsymbol{\sigma})\\
&\le h_\Omega^2\|\nabla\cdot\boldsymbol{\sigma} - \Pi^p_{\mathcal{T}}(\nabla\cdot\boldsymbol{\sigma})\|^2 + \|\nabla(u-u_{\mathrm{LS}})\|^2 + \|\boldsymbol{\sigma}-\boldsymbol{\sigma}_{\mathcal{T}}\|^2 + \|\boldsymbol{\sigma}-\boldsymbol{\sigma}_{\mathrm{LS}}\|^2.
\end{aligned}
$$

We infer that

$$
h_\Omega^2\|\nabla\cdot(\boldsymbol{\sigma}-\boldsymbol{\sigma}_{\mathrm{LS}})\|^2 \le h_\Omega^2\|\nabla\cdot\boldsymbol{\sigma} - \Pi^p_{\mathcal{T}}(\nabla\cdot\boldsymbol{\sigma})\|^2 + \|\nabla(u-u_{\mathrm{LS}})\|^2 + \|\boldsymbol{\sigma}-\boldsymbol{\sigma}_{\mathcal{T}}\|^2.
$$

This finishes the proof since $\boldsymbol{\sigma}_{\mathcal{T}}$ is arbitrary. □

References

[1] M. Aurada, M. Feischl, J. Kemetmüller, M. Page, and D. Praetorius. Each $H^{1/2}$-stable projection yields convergence and quasi-optimality of adaptive FEM with inhomogeneous Dirichlet data in $\mathbb{R}^d$. *ESAIM Math. Model. Numer. Anal.*, 47(4):1207–1235, 2013. ISSN 0764-583X. doi: 10.1051/m2an/2013069. URL https://doi.org/10.1051/m2an/2013069.

[2] I. Babuška and M. Suri. The *h-p* version of the finite element method with quasi-uniform meshes. *RAIRO Modél. Math. Anal. Numér.*, 21(2):199–238, 1987. ISSN 0764-583X. doi: 10.1051/m2an/1987210201991. URL http://dx.doi.org/10.1051/m2an/1987210201991.

[3] R. Becker, D. Capatina, and R. Luce. Local flux reconstructions for standard finite element methods on triangular meshes. *SIAM J. Numer. Anal.*, 54(4):2684–2706, 2016. ISSN 0036-1429. doi: 10.1137/16M1064817. URL http://dx.doi.org/10.1137/16M1064817.

[4] C. Bernardi and V. Girault. A local regularization operator for triangular and quadrilateral finite elements. *SIAM J. Numer. Anal.*, 35(5):1893–1916, 1998. ISSN 0036-1429. doi: 10.1137/S0036142995293766. URL https://doi.org/10.1137/S0036142995293766.

[5] C. Bernardi and F. Hecht. Quelques propriétés d'approximation des éléments finis de Nédélec, application à l'analyse a posteriori. *C. R. Math. Acad. Sci. Paris*, 344(7):461–466, 2007. ISSN 1631-073X. doi: 10.1016/j.crma.2007.02.010. URL https://doi.org/10.1016/j.crma.2007.02.010.

[6] A. Bespalov and N. Heuer. A new $\mathbf{H}(\mathrm{div})$-conforming p-interpolation operator in two dimensions. *ESAIM Math. Model. Numer. Anal.*, 45(2):255–275, 2011. ISSN 0764-583X. doi: 10.1051/m2an/2010039. URL https://doi.org/10.1051/m2an/2010039.

[7] D. Boffi, F. Brezzi, and M. Fortin. *Mixed finite element methods and applications*, volume 44 of *Springer Series in Computational Mathematics*. Springer, Heidelberg, 2013. ISBN 978-3-642-36518-8; 978-3-642-36519-5. doi: 10.1007/978-3-642-36519-5. URL https://doi.org/10.1007/978-3-642-36519-5.

[8] D. Braess and J. Schöberl. Equilibrated residual error estimator for edge elements. *Math. Comp.*, 77 (262):651–672, 2008. ISSN 0025-5718. doi: 10.1090/S0025-5718-07-02080-7. URL http://dx.doi.org/10.1090/S0025-5718-07-02080-7.

[9] D. Braess, V. Pillwein, and J. Schöberl. Equilibrated residual error estimates are p-robust. *Comput. Methods Appl. Mech. Engrg.*, 198(13-14):1189–1197, 2009. ISSN 0045-7825. doi: 10.1016/j.cma.2008.12.010. URL http://dx.doi.org/10.1016/j.cma.2008.12.010.

[10] Z. Cai and J. Ku. Optimal error estimate for the div least-squares method with data $f \in L^2$ and application to nonlinear problems. *SIAM J. Numer. Anal.*, 47(6):4098–4111, 2010. ISSN 0036-1429. doi: 10.1137/080738350. URL https://doi.org/10.1137/080738350.

[11] C. Canuto, R. H. Nochetto, R. Stevenson, and M. Verani. Convergence and optimality of hp-AFEM. *Numer. Math.*, 135(4):1073–1119, 2017. ISSN 0029-599X. doi: 10.1007/s00211-016-0826-x. URL https://doi.org/10.1007/s00211-016-0826-x.

[12] C. Carstensen, D. Peterseim, and M. Schedensack. Comparison results of finite element methods for the Poisson model problem. *SIAM J. Numer. Anal.*, 50(6):2803–2823, 2012. ISSN 0036-1429. doi: 10.1137/110845707. URL https://doi.org/10.1137/110845707.

[13] S. H. Christiansen and R. Winther. Smoothed projections in finite element exterior calculus. *Math. Comp.*, 77(262):813–829, 2008. ISSN 0025-5718. doi: 10.1090/S0025-5718-07-02081-9. URL http://dx.doi.org/10.1090/S0025-5718-07-02081-9.

[14] P. Clément. Approximation by finite element functions using local regularization. *RAIRO Anal. Numer.*, 9(R-2):77–84, 1975. ISSN 0399-0516.

[15] M. Costabel and A. McIntosh. On Bogovskiĭ and regularized Poincaré integral operators for de Rham complexes on Lipschitz domains. *Math. Z.*, 265(2):297–320, 2010. ISSN 0025-5874. doi: 10.1007/s00209-009-0517-8. URL http://dx.doi.org/10.1007/s00209-009-0517-8.

[16] L. Demkowicz. Polynomial exact sequences and projection-based interpolation with application to Maxwell equations. In D. Boffi, F. Brezzi, L. F. Demkowicz, R. G. Durán, R. S. Falk, and M. Fortin, editors, *Mixed finite elements, compatibility conditions, and applications*, volume 1939 of *Lecture Notes in Mathematics*, pages x+235. Springer-Verlag, Berlin; Fondazione C.I.M.E., Florence, 2008. ISBN 978-3-540-78314-5. doi: 10.1007/978-3-540-78319-0_3. URL https://doi.org/10.1007/978-3-540-78319-0_3. Lectures given at the C.I.M.E. Summer School held in Cetraro, June 26–July 1, 2006, Edited by D. Boffi and L. Gastaldi.

[17] L. Demkowicz and A. Buffa. H^1, $H(\mathrm{curl})$ and $H(\mathrm{div})$-conforming projection-based interpolation in three dimensions. Quasi-optimal p-interpolation estimates. *Comput. Methods Appl. Mech. Engrg.*, 194(2-5):267–296, 2005. ISSN 0045-7825. doi: 10.1016/j.cma.2004.07.007. URL https://doi.org/10.1016/j.cma.2004.07.007.

[18] L. Demkowicz, J. Gopalakrishnan, and J. Schöberl. Polynomial extension operators. Part I. *SIAM J. Numer. Anal.*, 46(6):3006–3031, 2008. ISSN 0036-1429. doi: 10.1137/070698786. URL http://dx.doi.org/10.1137/070698786.

[19] L. Demkowicz, J. Gopalakrishnan, and J. Schöberl. Polynomial extension operators. Part II. *SIAM J. Numer. Anal.*, 47(5):3293–3324, 2009. ISSN 0036-1429. doi: 10.1137/070698798. URL http://dx.doi.org/10.1137/070698798.

[20] L. Demkowicz, J. Gopalakrishnan, and J. Schöberl. Polynomial extension operators. Part III. *Math. Comp.*, 81(279):1289–1326, 2012. ISSN 0025-5718. doi: 10.1090/S0025-5718-2011-02536-6. URL http://dx.doi.org/10.1090/S0025-5718-2011-02536-6.

[21] P. Destuynder and B. Métivet. Explicit error bounds in a conforming finite element method. *Math. Comp.*, 68(228):1379–1396, 1999. ISSN 0025-5718. doi: 10.1090/S0025-5718-99-01093-5. URL http://dx.doi.org/10.1090/S0025-5718-99-01093-5.

[22] V. Dolejší, A. Ern, and M. Vohralík. hp-adaptation driven by polynomial-degree-robust a posteriori error estimates for elliptic problems. *SIAM J. Sci. Comput.*, 38(5):A3220–A3246, 2016. ISSN 1064-8275. doi: 10.1137/15M1026687. URL http://dx.doi.org/10.1137/15M1026687.

[23] A. Ern and J.-L. Guermond. Mollification in strongly Lipschitz domains with application to continuous and discrete de Rham complexes. *Comput. Methods Appl. Math.*, 16(1):51–75, 2016. ISSN 1609-4840. doi: 10.1515/cmam-2015-0034. URL https://doi.org/10.1515/cmam-2015-0034.

[24] A. Ern and J.-L. Guermond. Finite element quasi-interpolation and best approximation. *ESAIM Math. Model. Numer. Anal.*, 51(4):1367–1385, 2017. ISSN 0764-583X. URL https://doi.org/10.1051/m2an/2016066.

[25] A. Ern and M. Vohralík. Adaptive inexact Newton methods with a posteriori stopping criteria for nonlinear diffusion PDEs. *SIAM J. Sci. Comput.*, 35(4):A1761–A1791, 2013. ISSN 1064-8275. doi: 10.1137/120896918. URL http://dx.doi.org/10.1137/120896918.

[26] A. Ern and M. Vohralík. Polynomial-degree-robust a posteriori estimates in a unified setting for conforming, nonconforming, discontinuous Galerkin, and mixed discretizations. *SIAM J. Numer. Anal.*, 53(2):1058–1081, 2015. ISSN 0036-1429. doi: 10.1137/130950100. URL http://dx.doi.org/10.1137/130950100.

[27] A. Ern and M. Vohralík. Stable broken H^1 and $\boldsymbol{H}(\mathrm{div})$ polynomial extensions for polynomial-degree-robust potential and flux reconstruction in three space dimensions. *Math. Comp.*, 2019. doi: 10.1090/mcom/3482. URL https://hal.inria.fr/hal-01422204. DOI 10.1090/mcom/3482.

[28] A. Ern, I. Smears, and M. Vohralík. Discrete p-robust $\boldsymbol{H}(\mathrm{div})$-liftings and a posteriori estimates for elliptic problems with H^{-1} source terms. *Calcolo*, 54(3):1009–1025, 2017. doi: 10.1007/s10092-017-0217-4. URL http://link.springer.com/article/10.1007/s10092-017-0217-4.

[29] R. S. Falk and R. Winther. Local bounded cochain projections. *Math. Comp.*, 83(290):2631–2656, 2014. ISSN 0025-5718. doi: 10.1090/S0025-5718-2014-02827-5. URL http://dx.doi.org/10.1090/S0025-5718-2014-02827-5.

[30] T. Gudi. A new error analysis for discontinuous finite element methods for linear elliptic problems. *Math. Comp.*, 79(272):2169–2189, 2010. ISSN 0025-5718. doi: 10.1090/S0025-5718-10-02360-4. URL http://dx.doi.org/10.1090/S0025-5718-10-02360-4.

[31] J. Ku. A comment on least-squares finite element methods with minimum regularity assumptions. *Int. J. Numer. Anal. Model.*, 10(4):899–903, 2013. ISSN 1705-5105.

[32] M. W. Licht. Smoothed projections and mixed boundary conditions. *Math. Comp.*, 88(316):607–635, 2019. ISSN 0025-5718. doi: 10.1090/mcom/3330. URL https://doi.org/10.1090/mcom/3330.

[33] R. Luce and B. I. Wohlmuth. A local a posteriori error estimator based on equilibrated fluxes. *SIAM J. Numer. Anal.*, 42(4):1394–1414, 2004. ISSN 0036-1429. doi: 10.1137/S0036142903433790. URL http://dx.doi.org/10.1137/S0036142903433790.

[34] J. M. Melenk and C. Rojik. On commuting p-version projection-based interpolation on tetrahedra. *Math. Comp.*, 2019. doi: 10.1090/mcom/3454. URL https://doi.org/10.1090/mcom/3454. DOI 10.1090/mcom/3454.

[35] J.-C. Nédélec. Mixed finite elements in $\mathbb{R}^3$. *Numer. Math.*, 35(3):315–341, 1980. ISSN 0029-599X.

[36] R. H. Nochetto and B. Stamm. A posteriori error estimates for the electric field integral equation on polyhedra. In *Contributions to partial differential equations and applications*, volume 47 of *Comput. Methods Appl. Sci.*, pages 371–394. Springer, Cham, 2019.

[37] A. I. Pehlivanov, G. F. Carey, and R. D. Lazarov. Least-squares mixed finite elements for second-order elliptic problems. *SIAM J. Numer. Anal.*, 31(5):1368–1377, 1994. ISSN 0036-1429. doi: 10.1137/0731071. URL https://doi.org/10.1137/0731071.

[38] P.-A. Raviart and J.-M. Thomas. A mixed finite element method for 2nd order elliptic problems. In *Mathematical aspects of finite element methods (Proc. Conf., Consiglio Naz. delle Ricerche (C.N.R.), Rome, 1975)*, pages 292–315. Lecture Notes in Math., Vol. 606. Springer, Berlin, 1977.

[39] J. E. Roberts and J.-M. Thomas. Mixed and hybrid methods. In *Handbook of Numerical Analysis, Vol. II*, pages 523–639. North-Holland, Amsterdam, 1991.

[40] J. Schöberl. Commuting quasi-interpolation operators for mixed finite elements. Tech. rep. ISC-01-10-MATH, Texas A&M University, 2001. URL https://isc.tamu.edu/resources/preprints/2001/2001-10.pdf.

[41] J. Schöberl. A multilevel decomposition result in $\boldsymbol{H}(\mathrm{curl})$. In P. Wesseling, C. Oosterlee, and P. Hemker, editors, *Multigrid, Multilevel and Multiscale Methods (EMG 2005)*, 2005. ISBN 90-9020969-7.

[42] L. R. Scott and S. Zhang. Finite element interpolation of nonsmooth functions satisfying boundary conditions. *Math. Comp.*, 54(190):483–493, 1990. ISSN 0025-5718. doi: 10.2307/2008497. URL https://doi.org/10.2307/2008497.

[43] A. Veeser. Approximating gradients with continuous piecewise polynomial functions. *Found. Comput. Math.*, 16(3):723–750, 2016. ISSN 1615-3375. doi: 10.1007/s10208-015-9262-z. URL http://dx.doi.org/10.1007/s10208-015-9262-z.

[44] R. Verfürth. *A posteriori error estimation techniques for finite element methods*. Numerical Mathematics and Scientific Computation. Oxford University Press, Oxford, 2013. ISBN 978-0-19-967942-3. doi: 10.1093/acprof:oso/9780199679423.001.0001. URL http://dx.doi.org/10.1093/acprof:oso/9780199679423.001.0001.