



HAL
open science

Robust a posteriori error control for transmission problems with sign-changing coefficients using localization of dual norms

Patrick Ciarlet, Martin Vohralík

► **To cite this version:**

Patrick Ciarlet, Martin Vohralík. Robust a posteriori error control for transmission problems with sign-changing coefficients using localization of dual norms. 2015. hal-01148476v1

HAL Id: hal-01148476

<https://inria.hal.science/hal-01148476v1>

Preprint submitted on 4 May 2015 (v1), last revised 16 Oct 2018 (v3)

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.

Robust a posteriori error control for transmission problems with sign-changing coefficients using localization of dual norms

Patrick Ciarlet Jr.[†] Martin Vohralík[‡]

May 4, 2015

Abstract

We present a posteriori error analysis of diffusion problems where the diffusion tensor is not necessarily symmetric and positive definite and can in particular change its sign. We first identify the correct intrinsic error norm for such problems, including both conforming and nonconforming approximations. This involves global dual (residual) norms. Importantly, we show their equivalence with the Hilbertian sums of their localizations. We then design estimators which deliver simultaneously guaranteed error upper bound, local and global error lower bounds, and robustness with respect to the diffusion tensor as well as with respect to the approximation polynomial degree. The estimators are given in a unified setting covering at once conforming, nonconforming, mixed, and discontinuous Galerkin finite element discretizations. Numerical results illustrate the theoretical developments.

Key words: noncoercive problem, sign change, metamaterial, a posteriori error estimate, intrinsic dual error norm, localization, equilibrated flux, unified framework, robustness, finite element methods

1 Introduction

Let $\Omega \subset \mathbb{R}^d$, $d \geq 2$, be an open polytope (polygon for $d = 2$, polyhedron for $d = 3$), $\underline{\Sigma}$ a tensor-valued diffusion tensor, and f a datum. We consider the following problem: find $u : \Omega \rightarrow \mathbb{R}$ such that

$$-\nabla \cdot (\underline{\Sigma} \nabla u) = f \quad \text{in } \Omega, \quad (1.1a)$$

$$u = 0 \quad \text{on } \partial\Omega. \quad (1.1b)$$

In contrast to the usual setting, cf. Ciarlet [12], we relax the assumption of $\underline{\Sigma}$ being positive definite (and symmetric). Such a situation arises as a model problem in electromagnetism for interfaces between dielectrics and (negative) metamaterials, see, e.g., Bonnet-BenDhia *et al.* [3] or Wallen *et al.* [30] and the references therein. The exemplar situation is the case where Ω is composed of two subdomains Ω_+ and Ω_- of nonzero measure such that $\underline{\Sigma}|_{\Omega_+} = \sigma_+ \mathbf{I}$ and $\underline{\Sigma}|_{\Omega_-} = \sigma_- \mathbf{I}$, where $\sigma_+ > 0$ and $\sigma_- < 0$ are two scalars and \mathbf{I} is the identity tensor.

Let (\cdot, \cdot) denote the $L^2(\Omega)$ or $[L^2(\Omega)]^d$ scalar product; $\|\cdot\|$ stands for the associated norm. We will call $u \in H_0^1(\Omega)$ a weak solution of (1.1) if

$$(\underline{\Sigma} \nabla u, \nabla v) = (f, v) \quad \forall v \in H_0^1(\Omega). \quad (1.2)$$

Conditions for well-posedness (existence, uniqueness, and continuous dependence on the data) of the general problem (1.2) follow from the celebrated Banach–Nečas–Babuška theorem, cf., e.g., Ern and Guermond [15, Theorem 2.6]. They have recently been revisited via the T-coercivity approach, see, e.g., Bonnet-BenDhia *et al.* [2] or Chesnel and Ciarlet [10] and the references therein. Conception of numerical approximations, their well-posedness, and a priori error estimates have been addressed in [10] in the conforming finite element context and in [11] for nonconforming and discontinuous Galerkin context.

[†]Laboratoire POEMS, UMR 7231 CNRS/ENSTA/INRIA, ENSTA ParisTech, 828 Boulevard des Maréchaux, 91762 Palaiseau, France (patrick.ciarlet@ensta-paristech.fr).

[‡]INRIA Paris-Rocquencourt, B.P. 105, 78153 Le Chesnay, France (martin.vohralik@inria.fr).

A posteriori error analysis for problems of type (1.1) has likewise been started recently. In particular, Nicaise and Venel [23] bound the error between the known finite element approximation and the unknown weak solution u of (1.2) by a computable a posteriori indicator. This bound features an unknown constant, so that the error control is not guaranteed. More seriously, the dependence of the quality of the estimator on the tensor $\underline{\Sigma}$ (on the ratio, or contrast, σ_+/σ_- in the simplest setting) is not traced; numerical experiments indicate deterioration of the behavior (so-called non-robustness) when the contrast is approaching the set of forbidden values, equal to an interval to which -1 always belongs. There is also a need for a discrete version of the trace lifting operator, both in the analysis and in the implementation. The previous contributions, see Bernardi and Verfürth [1] or [29] and the references therein, only study the standard positive definite case.

The goal of this contribution is threefold: firstly, we want to derive a posteriori error estimates which are *guaranteed*, certifying the maximal error. Secondly, we wish them to be *robust* with respect to the jumps and *sign changes* in the tensor $\underline{\Sigma}$. The adaptive mesh refinement based on the a posteriori error estimators developed in this work produces in our numerical experiments sequences of meshes leading to optimal decay rates for an arbitrarily singular solution. Finally, we want to develop a *unified framework* covering all standard numerical methods. The guaranteed upper bound will be achieved via the equilibrated fluxes concept following Prager and Synge [24], Ladevèze and Leguillon [19], Kelly [18], Destuynder and Métivet [13], Luce and Wohlmuth [20], Braess and Schöberl [5], Repin [25], or [29, 16].

The key point for obtaining the above-discussed robustness is a proper choice of the way the error is measured. For conforming (lying in the space $H_0^1(\Omega)$) approximations, following Verfürth [27], Chaillou and Suri [9], and [29, 14], we introduce the *intrinsic problem-dependent norm*

$$\|v\|_{\#} := \sup_{\varphi \in H_0^1(\Omega); \|\nabla\varphi\|=1} (\underline{\Sigma}\nabla v, \nabla\varphi) \quad v \in H_0^1(\Omega). \quad (1.3)$$

Importantly, we prove that $\|v\|_{\#}$ is *equivalent* with the Hilbertian sum of its *localizations* on patches of elements. The unified framework we develop also includes nonconforming approximations and turns out to be subtle in the present non-coercive setting; it requires a proper generalization of (1.3) to $v \notin H_0^1(\Omega)$. We elaborate it in extension of the recent work [16]. In addition to the robustness with respect to the tensor $\underline{\Sigma}$, we then also obtain *robustness* with respect to the approximation *polynomial degree*.

Our paper is organized as follows. Section 2 sets the notation and assumptions on both continuous and discrete levels. Section 3 discusses conforming approximations while relying on the intrinsic norm (1.3). Section 4 then extends these concepts to an abstract setting and presents our unified framework for all standard numerical methods. Section 5 resumes our general findings about the localization of the dual norms that we find new and of independent interest. Finally, Section 6 illustrates our theoretical developments on two numerical examples and Section 7 gives some concluding remarks and outlook.

2 Setting

This section introduces the assumptions, notation, and discusses in detail the choice of the way we will measure the error.

2.1 Assumptions

Throughout the paper, we shall suppose the following:

Assumption 2.1 (Setting). *We suppose that*

- $\{\mathcal{T}_h\}_h$ is a family of simplicial partitions of Ω , i.e., $\cup_{K \in \mathcal{T}_h} \overline{K} = \overline{\Omega}$ for all \mathcal{T}_h , any $K \in \mathcal{T}_h$ for any \mathcal{T}_h is a closed simplex, and the intersection of two different simplices in one \mathcal{T}_h is either empty, a vertex, or an l -dimensional face, $1 \leq l \leq d-1$;
- the family $\{\mathcal{T}_h\}_h$ is shape regular in the sense that there exists a constant $\kappa_{\mathcal{T}} > 0$ such that, for all triangulations \mathcal{T}_h , $\max_{K \in \mathcal{T}_h} h_K / \varrho_K \leq \kappa_{\mathcal{T}}$, where h_K is the diameter of K and ϱ_K is the diameter of the largest ball inscribed in K ;
- $\underline{\Sigma} \in [L^\infty(\Omega)]^{d \times d}$ is piecewise constant on each given \mathcal{T}_h ;

- $f \in L^2(\Omega)$;
- there exists a linear bijective operator $T : H_0^1(\Omega) \rightarrow H_0^1(\Omega)$, cf. [10, Definition 3], bounded in the sense that $\|\nabla(Tv)\| \leq \|T\|\|\nabla v\|$ for all $v \in H_0^1(\Omega)$, $\|T\| < \infty$, and T -coercive in the sense that $(\underline{\Sigma}\nabla v, \nabla(Tv)) \geq \underline{\alpha}\|\nabla v\|^2$ for all $v \in H_0^1(\Omega)$, $\underline{\alpha} > 0$.

Corollary 2.2 (Weak solution). *Under Assumption 2.1, there exists a well-posed solution of (1.1) in the sense (1.2). It satisfies, $u \in H_0^1(\Omega)$ and $\sigma := -\underline{\Sigma}\nabla u \in \mathbf{H}(\text{div}, \Omega)$ with $\nabla \cdot \sigma = f$.*

2.2 Intrinsic norm

Under Assumption 2.1, the intrinsic norm given by (1.3) satisfies, for $v \in H_0^1(\Omega)$, $\|v\|_{\#} = 0$ if and only if $v = 0$. Moreover, the Cauchy–Schwarz inequality implies

$$\frac{(\underline{\Sigma}\nabla v, \nabla(Tv))}{\|\nabla(Tv)\|} \leq \|v\|_{\#} \leq \|\underline{\Sigma}\nabla v\| \quad \forall v \in H_0^1(\Omega), \quad (2.1)$$

and the boundedness and coercivity of T and the boundedness of $\underline{\Sigma}$ allow to further confine

$$\frac{\underline{\alpha}}{\|T\|} \|\nabla v\| \leq \|v\|_{\#} \leq \|\underline{\Sigma}\|_{\infty} \|\nabla v\| \quad \forall v \in H_0^1(\Omega), \quad (2.2)$$

so that $\|v\|_{\#}$ is equivalent to the canonical norm $\|\nabla v\|$ on the space $H_0^1(\Omega)$. Note that $(\underline{\Sigma}\nabla v, \nabla v)$ may become negative, which excludes the notion itself of an energy norm; on the other hand $\|v\|_{\#} = \|\nabla v\|$ in the case where $\underline{\Sigma} = \underline{I}$, so that $\|v\|_{\#}$ is a natural extension of the canonical norm of the Laplace equation.

Let $H^1(\mathcal{T}_h) := \{v \in L^2(\Omega); v|_K \in H^1(K) \text{ for all } K \in \mathcal{T}_h\}$ be the broken Sobolev space and ∇ the corresponding broken (elementwise) gradient. We will need to extend definition (1.3) to more general arguments; if $v \in H^1(\mathcal{T}_h)$ but $v \notin H_0^1(\Omega)$, $\|v\|_{\#}$ is not anymore a norm and seems insufficient. In fact, when $\underline{\Sigma} = \underline{I}$, there holds, for arbitrary $u \in H_0^1(\Omega)$ and $u_h \in H^1(\mathcal{T}_h)$, cf. [16, Theorem 3.3] and the references therein,

$$\|\nabla(u - u_h)\|^2 = \sup_{\varphi \in H_0^1(\Omega); \|\nabla\varphi\|=1} (\nabla(u - u_h), \nabla\varphi)^2 + \inf_{s \in H_0^1(\Omega)} \|\nabla(u_h - s)\|^2.$$

Thus $\|u - u_h\|_{\#}$ is only one part of the distance $\|\nabla(u - u_h)\|$ in this case. We are led to measure the error between the weak solution $u \in H_0^1(\Omega)$ of (1.2) and an arbitrary $u_h \in H^1(\mathcal{T}_h)$ by

$$\|u - u_h\|_{\#}^2 := \|u - u_h\|_{\#}^2 + \inf_{s \in H_0^1(\Omega)} \|\nabla(u_h - s)\|^2. \quad (2.3)$$

Of course, the second term disappears whenever $u_h \in H_0^1(\Omega)$; in this case, $\|u - u_h\|_{\#} = \|u - u_h\|_{\#}$. In the general case $u_h \in H^1(\mathcal{T}_h)$, there holds

$$\|u - u_h\|_{\#} = \sup_{\varphi \in H_0^1(\Omega); \|\nabla\varphi\|=1} \{(f, \varphi) - (\underline{\Sigma}\nabla u_h, \nabla\varphi)\}$$

using (1.2). Thus $\|u - u_h\|_{\#}$ is the dual norm of the residual. Moreover,

$$\|u - u_h\|_{\#} \leq \inf_{\sigma \in \mathbf{H}(\text{div}, \Omega); \nabla \cdot \sigma = f} \|\underline{\Sigma}\nabla u_h + \sigma\|$$

by the Green theorem, so that $\|u - u_h\|_{\#}$ is linked to the *nonconformity* in the approximate flux $-\underline{\Sigma}\nabla u_h$. The second term in (2.3) then expresses the possible nonconformity in the discrete *potential* u_h .

Finally, in what follows, we will need to localize (1.3). Let ω be a subdomain of Ω . For a function v piecewise H^1 on ω , we define

$$\|v\|_{\#, \omega} := \sup_{\varphi \in H_0^1(\omega); \|\nabla\varphi\|_{H_0^1(\omega)}=1} (\underline{\Sigma}\nabla v, \nabla\varphi)_{\omega}, \quad (2.4)$$

where $(\cdot, \cdot)_{\omega}$ is the $L^2(\omega)$ or $[L^2(\omega)]^d$ scalar product and $\|\cdot\|_{\omega}$ the associated norm. Note that the Cauchy–Schwarz inequality still gives, for all piecewise H^1 functions v on ω ,

$$\|v\|_{\#, \omega} \leq \|\underline{\Sigma}\nabla v\|_{\omega}. \quad (2.5)$$

2.3 Notation

We denote a generic element of \mathcal{T}_h by K ; \mathbf{n}_K then stands for the outward unit normal vector to K . The mesh $(d-1)$ -dimensional faces are collected in the set \mathcal{E}_h , with interior faces $\mathcal{E}_h^{\text{int}}$ and boundary faces $\mathcal{E}_h^{\text{ext}}$. A generic face is denoted by e and its diameter by h_e . For any $e \in \mathcal{E}_h$, \mathbf{n}_e stands for the unit normal vector to e ; the orientation is arbitrary but fixed for $e \in \mathcal{E}_h^{\text{int}}$ and points outwards of Ω for $e \in \mathcal{E}_h^{\text{ext}}$. We will use the jump operator $[\![\cdot]\!]_e$ yielding the difference evaluated along \mathbf{n}_e of the traces of the argument from the two mesh elements that share $e \in \mathcal{E}_h^{\text{int}}$ and the actual trace for $e \in \mathcal{E}_h^{\text{ext}}$. For $K \in \mathcal{T}_h$, $\mathcal{E}_K^{\text{int}}$ stands for all faces e of K not lying on $\partial\Omega$. Finally, \mathcal{T}_K denotes all the elements in \mathcal{T}_h sharing a face with K and ω_K the corresponding subdomain. The set of vertices will be denoted by \mathcal{V}_h ; it is composed of interior vertices $\mathcal{V}_h^{\text{int}}$ and vertices located on the boundary $\mathcal{V}_h^{\text{ext}}$. For a vertex $\mathbf{a} \in \mathcal{V}_h$, $\mathcal{T}_{\mathbf{a}}$ stands for the patch of the elements of \mathcal{T}_h which share \mathbf{a} , $\omega_{\mathbf{a}}$ the corresponding subdomain, and $\psi_{\mathbf{a}}$ the continuous, piecewise affine ‘‘hat’’ function which takes value 1 at the vertex \mathbf{a} and zero at the other vertices. For $K \in \mathcal{T}_h$, \mathcal{V}_K denotes the set of vertices of K .

We use $\mathbb{P}_p(\mathcal{T}_h)$, $p \geq 0$, to denote piecewise polynomials on \mathcal{T}_h of total degree at most p ; Π_p is the $L^2(\Omega)$ -orthogonal projection onto $\mathbb{P}_p(\mathcal{T}_h)$. We will rely heavily on the Raviart–Thomas–Nédélec mixed finite element spaces $\mathbf{RTN}_p(\mathcal{T}_h) := \{\mathbf{v}_h \in \mathbf{H}(\text{div}, \Omega); \mathbf{v}_h|_K \in \mathbf{RTN}_p(K)\}$, $p \geq 0$, with the local spaces $\mathbf{RTN}_p(K) := [\mathbb{P}_p]^d + \mathbb{P}_p \mathbf{x}$, $K \in \mathcal{T}_h$, see see Brezzi and Fortin [6] or Roberts and Thomas [26]. The abstract notations $\mathbf{V}_h \times Q_h$ for $\mathbf{RTN}_p(\mathcal{T}_h) \times \mathbb{P}_p(\mathcal{T}_h)$, as well as $\mathbf{V}_h(\omega) \times Q_h(\omega)$ for their restriction to $\omega \subset \Omega$, will be often employed. The $L^2(\Omega)$ -orthogonal projection onto Q_h will then be denoted by Π_{Q_h} .

3 Conforming discretizations

This section adapts to the present $\underline{\Sigma}$ -setting and summarizes the a posteriori error estimates that have been previously obtained in the literature for conforming discretizations of (1.1) for symmetric and positive definite $\underline{\Sigma}$. It in particular covers conforming finite element, vertex- and cell-centered finite volume, and finite difference discretizations.

3.1 Residual estimates

Set $V_h := \mathbb{P}_p(\mathcal{T}_h) \cap H_0^1(\Omega)$, $p \geq 1$. The finite element method for problem (1.1) reads: find $u_h \in V_h$ such that

$$(\underline{\Sigma} \nabla u_h, \nabla v_h) = (f, v_h) \quad \forall v_h \in V_h. \quad (3.1)$$

For $p = 1$ and piecewise constant f , (3.1) may also coincide with a finite volume or finite difference discretization, see [29, Section 3] and below. Following Verfürth [28], the residual-based a posteriori error estimator is, for $K \in \mathcal{T}_h$, given by

$$\eta_{\text{res}, K} := h_K \|f + \nabla \cdot (\underline{\Sigma} \nabla u_h)\|_K + \left\{ \sum_{e \in \mathcal{E}_K^{\text{int}}} h_e \|[\![\underline{\Sigma} \nabla u_h]\!] \cdot \mathbf{n}_e\|_e^2 \right\}^{1/2}. \quad (3.2)$$

Define the data oscillation term

$$\eta_{\text{osc}, K} := h_K \|f - \Pi_p f\|_K, \quad (3.3)$$

as well as the global versions of these estimators $\eta := \{\sum_{K \in \mathcal{T}_h} (\eta_{\cdot, K})^2\}^{1/2}$. The following results follow from [22, 27, 9, 29, 14, 28] after a close verification:

Theorem 3.1 (Residual estimate for conforming discretizations). *Let u be the weak solution given by (1.2). Let $u_h \in V_h$ be its finite element approximation given by (3.1). Then there exists a constant $C_{\text{up}} > 0$ only depending on the space dimension d and on the shape-regularity parameter $\kappa_{\mathcal{T}}$ such that*

$$\|u - u_h\|_{\#} \leq C_{\text{up}} \eta_{\text{res}}. \quad (3.4)$$

Theorem 3.2 (Local and global efficiency and robustness for Theorem 3.1). *Let u be the weak solution given by (1.2) and let $u_h \in V_h$ be arbitrary. Then there exists a constant $C_{\text{low}} > 0$ only depending on d , $\kappa_{\mathcal{T}}$, and the polynomial degree p such that*

$$\eta_{\text{res},K} \leq C_{\text{low}} \left(\|u - u_h\|_{\#, \omega_K} + \left\{ \sum_{K' \in \mathcal{T}_K} \eta_{\text{osc},K'}^2 \right\}^{1/2} \right) \quad \forall K \in \mathcal{T}_h, \quad (3.5a)$$

$$\eta_{\text{res}} \leq C_{\text{low}} (\|u - u_h\|_{\#} + \eta_{\text{osc}}). \quad (3.5b)$$

Remark 3.3 (Theorems 3.1 and 3.2). *The estimate (3.4) follows from [28, Theorem 3.57] (cf. also Proposition 1.3 and equation (1.17) in this reference). The local lower bound (3.5a) is a consequence of [28, Theorem 3.59] (cf. also Proposition 3.37 in this reference) while noting that $\underline{\Sigma} \nabla u_h$ is a piecewise $(p-1)$ -degree polynomial by the requirement on $\underline{\Sigma}$ in Assumption 2.1 and while relying on the definition (2.4) of the weak norm $\|\cdot\|_{\#, \omega}$ in the spirit of [9, 14]. The global lower bound (3.5b) then follows by the dual norm technique of [27] while proceeding exactly as in the proof of [29, Theorem 5.4]. Note that neither C_{up} nor C_{low} depends on $\underline{\Sigma}$, so that, for the intrinsic norm $\|\cdot\|_{\#}$ of (1.3) and its localization $\|\cdot\|_{\#, \omega}$ of (2.4), the residual estimates are indeed robust with respect to $\underline{\Sigma}$. On the contrary, the dependence of $C_{\text{up}} C_{\text{low}}$ on the polynomial degree p seems unavoidable, see [22].*

Remark 3.4 (Efficiency in the L^2 flux norm for Theorem 3.1). *Owing to (2.5), we immediately have from (3.5)*

$$\eta_{\text{res},K} \leq C_{\text{low}} \left(\|\underline{\Sigma} \nabla (u - u_h)\|_{\omega_K} + \left\{ \sum_{K' \in \mathcal{T}_K} \eta_{\text{osc},K'}^2 \right\}^{1/2} \right) \quad \forall K \in \mathcal{T}_h,$$

$$\eta_{\text{res}} \leq C_{\text{low}} (\|\underline{\Sigma} \nabla (u - u_h)\| + \eta_{\text{osc}}).$$

3.2 Equilibrated fluxes estimates

The inconvenience of the estimates of the previous section is that the constants C_{up} from (3.4) and C_{low} from (3.5) are unknown. This is in particular crucial for the upper bound, since Theorem 3.1 does not really allow to control the error. A remedy can be achieved via the equilibrated flux concept.

Consider the first-order case only, i.e., $p = 1$, and look for $u_h \in V_h$ such that

$$-\langle \underline{\Sigma} \nabla u_h \cdot \mathbf{n}_D, 1 \rangle_{\partial D} = (f, 1)_D \quad \forall D \in \mathcal{D}_h^{\text{int}}. \quad (3.6)$$

The formulations (3.1) and (3.6) are equivalent for the source term f piecewise constant on \mathcal{T}_h , when \mathcal{D}_h is a dual mesh where to each vertex of \mathcal{V}_h , one associates a dual volume D by joining the surrounding l -dimensional, $1 \leq l \leq d-1$, face and element barycenters as indicated in Figure 1, left, for $d = 2$. The notation $\mathcal{D}_h^{\text{int}}$ stands for the dual volumes associated with the interior vertices of \mathcal{T}_h , and $\mathcal{D}_h^{\text{ext}}$ for those associated with the boundary vertices of \mathcal{T}_h . We will also need a second simplicial mesh \mathcal{S}_h , a submesh of both \mathcal{T}_h and \mathcal{D}_h , constructed using the barycenters of the elements and l -faces, $1 \leq l \leq d-1$, of the mesh \mathcal{T}_h as indicated in Figure 1, right, for $d = 2$. We will use the notation \mathcal{S}_D for the submesh of the dual volume D by the simplices of \mathcal{S}_h . Note that (3.6) is a local conservation equation, describing at once the conforming finite element, vertex-centered finite volume, cell-centered finite volume, and finite difference (for \mathcal{D}_h being rectangular) discretizations, see [29, Section 3] and the references therein.

To obtain the equilibrated fluxes, we will start from (3.6). We will solve some local Neumann problems on the interior dual volumes $D \in \mathcal{D}_h^{\text{int}}$, and Dirichlet–Neumann problems for $D \in \mathcal{D}_h^{\text{ext}}$, by means of lowest-order Raviart–Thomas–Nédélec mixed finite elements. Following Luce and Wohlmuth [20] and [29], define the space

$$\mathbf{RTN}_0^N(\mathcal{S}_D) := \{\mathbf{v}_h \in \mathbf{RTN}_0(\mathcal{S}_D); \mathbf{v}_h \cdot \mathbf{n}_D = -\underline{\Sigma} \nabla u_h \cdot \mathbf{n}_D \text{ on } \partial D \setminus \partial \Omega\} \quad (3.7)$$

for any $D \in \mathcal{D}_h$. Note that $-\underline{\Sigma} \nabla u_h \cdot \mathbf{n}_D$ is univalued everywhere in (3.7) and that the function $-\underline{\Sigma} \nabla u_h \cdot \mathbf{n}_D$ on the boundary of each $D \in \mathcal{D}_h^{\text{int}}$ by (3.6) satisfies the Neumann compatibility condition with the source term f . Denote by $\Pi_{0, \mathcal{S}_h} f$ the elementwise mean values of f on the fine mesh \mathcal{S}_h . Then we can define:

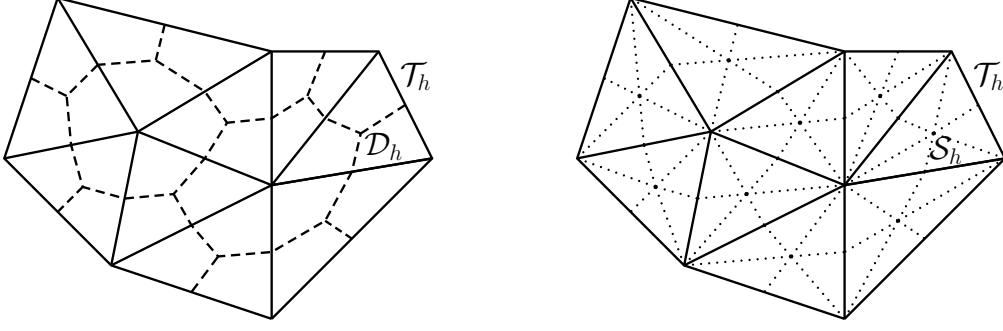


Figure 1: Simplicial mesh \mathcal{T}_h and the dual mesh \mathcal{D}_h (left); simplicial submesh \mathcal{S}_h (right)

Definition 3.5 (Equilibrated flux reconstruction on dual mesh). *Let u_h be given by (3.6). Prescribe $\sigma_h \in \mathbf{RTN}_0(\mathcal{S}_h)$ separately on each $D \in \mathcal{D}_h$ by*

$$\sigma_h|_D := \arg \inf_{\mathbf{v}_h \in \mathbf{RTN}_0^N(\mathcal{S}_D), \nabla \cdot \mathbf{v}_h = \Pi_{0, \mathcal{S}_h} f} \|\underline{\Sigma} \nabla u_h + \mathbf{v}_h\|_D. \quad (3.8)$$

It is classical, see Brezzi and Fortin [6] or Roberts and Thomas [26], that (3.8) is equivalent to finding $\sigma_h \in \mathbf{RTN}_0^N(\mathcal{S}_D)$ and $\bar{r}_h \in \mathbb{P}_0^*(\mathcal{S}_D)$ such that

$$(\sigma_h, \mathbf{v}_h)_D - (\bar{r}_h, \nabla \cdot \mathbf{v}_h)_D = -(\underline{\Sigma} \nabla u_h, \mathbf{v}_h)_D \quad \forall \mathbf{v}_h \in \mathbf{RTN}_0^{N,0}(\mathcal{S}_D), \quad (3.9a)$$

$$(\nabla \cdot \sigma_h, q_h)_D = (f, q_h)_D \quad \forall q_h \in \mathbb{P}_0^*(\mathcal{S}_D). \quad (3.9b)$$

Here $\mathbf{RTN}_0^{N,0}(\mathcal{S}_D)$ is as the space $\mathbf{RTN}_0^N(\mathcal{S}_D)$ of (3.7), with the homogeneous normal flux condition $\mathbf{v}_h \cdot \mathbf{n}_D = 0$ on $\partial D \setminus \partial \Omega$. Moreover, for $D \in \mathcal{D}_h^{\text{ext}}$, $\mathbb{P}_0^*(\mathcal{S}_D)$ is spanned by piecewise constants on \mathcal{S}_D , whereas for $D \in \mathcal{D}_h^{\text{int}}$, $\mathbb{P}_0^*(\mathcal{S}_D)$ is spanned by piecewise constants on \mathcal{S}_D with imposed zero mean value on D .

Define the data oscillation term

$$\eta_{\text{osc}, K} := \frac{h_K}{\pi} \|f - \Pi_{0, \mathcal{S}_h} f\|_K, \quad (3.10)$$

for $K \in \mathcal{S}_h$, and its global version $\eta_{\text{osc}} := \left\{ \sum_{K \in \mathcal{S}_h} (\eta_{\text{osc}, K})^2 \right\}^{1/2}$. Checking carefully the proofs in [29], we can summarize the following results:

Theorem 3.6 (Equilibrated flux estimate for lowest-order conforming discretizations). *Let u be the weak solution given by (1.2). Let $u_h \in V_h$ be its finite element / finite volume / finite difference approximation given by (3.6). Consider the equilibrated flux reconstruction of Definition 3.5. Then*

$$\|u - u_h\|_{\#} \leq \left\{ \sum_{K \in \mathcal{S}_h} (\|\underline{\Sigma} \nabla u_h + \sigma_h\|_K + \eta_{\text{osc}, K})^2 \right\}^{1/2}. \quad (3.11)$$

Theorem 3.7 (Local and global efficiency and robustness for Theorem 3.6). *Let u be the weak solution given by (1.2). Let σ_h be given by Definition 3.5. Then there exists a constant $\bar{C}_{\text{low}} > 0$ only depending on d and $\kappa_{\mathcal{T}}$ such that*

$$\|\underline{\Sigma} \nabla u_h + \sigma_h\|_D \leq \bar{C}_{\text{low}} \left(\|u - u_h\|_{\#, D} + \left\{ \sum_{K \in \mathcal{S}_D} \eta_{\text{osc}, K}^2 \right\}^{1/2} \right) \quad \forall D \in \mathcal{D}_h, \quad (3.12a)$$

$$\|\underline{\Sigma} \nabla u_h + \sigma_h\| \leq \bar{C}_{\text{low}} (\|u - u_h\|_{\#} + \eta_{\text{osc}}). \quad (3.12b)$$

Remark 3.8 (Efficiency in the L^2 flux norm for Theorem 3.6). *Owing to (2.5), we also have*

$$\|\underline{\Sigma} \nabla u_h + \sigma_h\|_D \leq \bar{C}_{\text{low}} \left(\|\underline{\Sigma} \nabla (u - u_h)\|_D + \left\{ \sum_{K \in \mathcal{S}_D} \eta_{\text{osc}, K}^2 \right\}^{1/2} \right) \quad \forall D \in \mathcal{D}_h,$$

$$\|\underline{\Sigma} \nabla u_h + \sigma_h\| \leq \bar{C}_{\text{low}} (\|\underline{\Sigma} \nabla (u - u_h)\| + \eta_{\text{osc}}).$$

Remark that the local efficiency is satisfied on each single dual volume $D \in \mathcal{D}_h$, without any neighborhood, see (3.12a). Extensive numerical comparisons in Carstensen and Merdon [8] show that these dual mesh equilibrated flux estimates are one of the most precise. This may be linked to the fact that the estimators of (3.11) themselves are minimized by (3.8) under the constraint $\nabla \cdot \boldsymbol{\sigma}_h = \Pi_{0, \mathcal{S}_h} f$. The generalization of this type of estimates to arbitrary approximation degree p is a bit technical but can be done, see [17, Section 9.2].

4 Unified framework

We present here an abstract framework, following [16], covering at once basically all classical numerical methods, in particular conforming, nonconforming, mixed, and discontinuous Galerkin finite elements. For a technical reason linked to the treatment of the nonconformity, we restrict ourselves to $d = 2$ in this section.

4.1 Guaranteed error control

The error upper bound in this section will be obtained under the following abstract assumption (recall that $\psi_{\mathbf{a}}$ is the $\mathbb{P}_1(\mathcal{T}_h) \cap H^1(\Omega)$ hat function associated with a vertex $\mathbf{a} \in \mathcal{V}_h$):

Assumption 4.1 (Setting and Galerkin orthogonality with respect to $\psi_{\mathbf{a}}$). *There holds*

$$u_h \in H^1(\mathcal{T}_h) \quad (4.1a)$$

$$(\boldsymbol{\Sigma} \nabla u_h, \nabla \psi_{\mathbf{a}})_{\omega_{\mathbf{a}}} = (f, \psi_{\mathbf{a}})_{\omega_{\mathbf{a}}} \quad \forall \mathbf{a} \in \mathcal{V}_h^{\text{int}}. \quad (4.1b)$$

In order to prove efficiency, a piecewise polynomial requirement and a mean value continuity of the jumps will be added below in Assumption 4.5.

In order to reconstruct a flux, we, similarly to the procedure of Section 3.2, will solve local Neumann or Neumann–Dirichlet mixed finite element problems. The Neumann condition will now be homogeneous, and the problems will be posed over patches of elements of \mathcal{T}_h , with overlap:

Definition 4.2 (Equilibrated flux reconstruction). *Let u_h satisfy Assumption 4.1. For $\mathbf{a} \in \mathcal{V}_h$, set*

$$\begin{aligned} \mathbf{V}_h^{\mathbf{a}} &:= \{\mathbf{v}_h \in \mathbf{V}_h(\omega_{\mathbf{a}}); \mathbf{v}_h \cdot \mathbf{n}_{\omega_{\mathbf{a}}} = 0 \text{ on } \partial\omega_{\mathbf{a}}\}, & \mathbf{a} \in \mathcal{V}_h^{\text{int}}, \\ Q_h^{\mathbf{a}} &:= \{q_h \in Q_h(\omega_{\mathbf{a}}); (q_h, 1)_{\omega_{\mathbf{a}}} = 0\}, \end{aligned} \quad (4.2a)$$

$$\begin{aligned} \mathbf{V}_h^{\mathbf{a}} &:= \{\mathbf{v}_h \in \mathbf{V}_h(\omega_{\mathbf{a}}); \mathbf{v}_h \cdot \mathbf{n}_{\omega_{\mathbf{a}}} = 0 \text{ on } \partial\omega_{\mathbf{a}} \setminus \partial\Omega\}, & \mathbf{a} \in \mathcal{V}_h^{\text{ext}}, \\ Q_h^{\mathbf{a}} &:= Q_h(\omega_{\mathbf{a}}), \end{aligned} \quad (4.2b)$$

Then prescribe $\boldsymbol{\varsigma}_h^{\mathbf{a}} \in \mathbf{V}_h^{\mathbf{a}}$ and $\bar{r}_h^{\mathbf{a}} \in Q_h^{\mathbf{a}}$ by solving

$$(\boldsymbol{\varsigma}_h^{\mathbf{a}}, \mathbf{v}_h)_{\omega_{\mathbf{a}}} - (\bar{r}_h^{\mathbf{a}}, \nabla \cdot \mathbf{v}_h)_{\omega_{\mathbf{a}}} = -(\psi_{\mathbf{a}} \boldsymbol{\Sigma} \nabla u_h, \mathbf{v}_h)_{\omega_{\mathbf{a}}} \quad \forall \mathbf{v}_h \in \mathbf{V}_h^{\mathbf{a}}, \quad (4.3a)$$

$$(\nabla \cdot \boldsymbol{\varsigma}_h^{\mathbf{a}}, q_h)_{\omega_{\mathbf{a}}} = (\psi_{\mathbf{a}} f - \boldsymbol{\Sigma} \nabla u_h \cdot \nabla \psi_{\mathbf{a}}, q_h)_{\omega_{\mathbf{a}}} \quad \forall q_h \in Q_h^{\mathbf{a}} \quad (4.3b)$$

and

$$\boldsymbol{\sigma}_h := \sum_{\mathbf{a} \in \mathcal{V}_h} \boldsymbol{\varsigma}_h^{\mathbf{a}}. \quad (4.4)$$

To cope with the nonconformity in u_h , we proceed similarly. Let $\mathbf{R}_{\frac{\pi}{2}} := \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ be the matrix of rotation by $\frac{\pi}{2}$ and denote by $\mathbf{R}_{\frac{\pi}{2}} \nabla$ the (broken) weak curl, i.e., the (broken) rotated gradient: for $v \in H^1(\mathcal{T}_h)$, $(\mathbf{R}_{\frac{\pi}{2}} \nabla v)|_K = (-\partial_y v|_K, \partial_x v|_K)$, $K \in \mathcal{T}_h$.

Definition 4.3 (Potential reconstruction). *Let u_h satisfy (4.1a). For $\mathbf{a} \in \mathcal{V}_h$, define $\mathbf{V}_h^{\mathbf{a}} \times Q_h^{\mathbf{a}}$ by (4.2a), and prescribe $\boldsymbol{\varsigma}_h^{\mathbf{a}} \in \mathbf{V}_h^{\mathbf{a}}$ and $\bar{r}_h^{\mathbf{a}} \in Q_h^{\mathbf{a}}$ by solving*

$$(\boldsymbol{\varsigma}_h^{\mathbf{a}}, \mathbf{v}_h)_{\omega_{\mathbf{a}}} - (\bar{r}_h^{\mathbf{a}}, \nabla \cdot \mathbf{v}_h)_{\omega_{\mathbf{a}}} = -(\mathbf{R}_{\frac{\pi}{2}} \nabla(\psi_{\mathbf{a}} u_h), \mathbf{v}_h)_{\omega_{\mathbf{a}}} \quad \forall \mathbf{v}_h \in \mathbf{V}_h^{\mathbf{a}}, \quad (4.5a)$$

$$(\nabla \cdot \boldsymbol{\varsigma}_h^{\mathbf{a}}, q_h)_{\omega_{\mathbf{a}}} = 0 \quad \forall q_h \in Q_h^{\mathbf{a}}. \quad (4.5b)$$

Then, set

$$-\mathbf{R}_{\frac{\pi}{2}} \nabla s_h^{\mathbf{a}} = \boldsymbol{\zeta}_h^{\mathbf{a}}, \quad (4.6a)$$

$$s_h^{\mathbf{a}} = 0 \text{ on } \partial\omega_{\mathbf{a}}, \quad (4.6b)$$

$$s_h := \sum_{\mathbf{a} \in \mathcal{V}_h} s_h^{\mathbf{a}}. \quad (4.6c)$$

It is classical that the two above problems are well-posed. As all $\boldsymbol{\zeta}_h^{\mathbf{a}}$ of Definition 4.2, extended by zero outside of $\omega_{\mathbf{a}}$, lie in $\mathbf{H}(\text{div}, \Omega)$, and similarly all $s_h^{\mathbf{a}}$ of Definition 4.3 lie in $H_0^1(\Omega)$, we obtain that $\boldsymbol{\sigma}_h$ is a piecewise polynomial in $\mathbf{H}(\text{div}, \Omega)$ (in \mathbf{V}_h actually) and s_h is a piecewise polynomial in $H_0^1(\Omega)$ (in a space W_h such that $\mathbf{R}_{\frac{\pi}{2}} \nabla W_h = \mathbf{V}_h$ actually). It is easy to verify that $\nabla \cdot \boldsymbol{\sigma}_h = \Pi_{Q_h} f$, see [5] or [16, Lemma 3.5]; both problems also admit local minimization characterizations similar to (3.8). Define the data oscillation terms

$$\eta_{\text{osc}, K} := \frac{h_K}{\pi} \|f - \Pi_{Q_h} f\|_K, \quad K \in \mathcal{T}_h, \quad (4.7)$$

and recall the definition (2.3) of $\|\cdot\|$. It follows from the proofs in [16] that:

Theorem 4.4 (Equilibrated flux estimate in a unified framework). *Let u be the weak solution given by (1.2). Let u_h satisfying Assumption 4.1 be arbitrary. Consider the equilibrated flux reconstruction of Definition 4.2 and the potential reconstruction of Definition 4.3. Then*

$$\|u - u_h\|^2 \leq \sum_{K \in \mathcal{T}_h} (\|\boldsymbol{\Sigma} \nabla u_h + \boldsymbol{\sigma}_h\|_K + \eta_{\text{osc}, K})^2 + \sum_{K \in \mathcal{T}_h} \|\nabla(u_h - s_h)\|_K^2. \quad (4.8)$$

4.2 Robust efficiency

In order to prove efficiency, we will need the following assumption:

Assumption 4.5 (Weak jumps continuity and piecewise polynomial form). *Let u_h be piecewise polynomial such that, with suitable spaces $\mathbf{V}_h \times Q_h$ and for all $\mathbf{a} \in \mathcal{V}_h$ and all $K \in \mathcal{T}_{\mathbf{a}}$,*

$$(\psi_{\mathbf{a}} \boldsymbol{\Sigma} \nabla u_h)|_K \in \mathbf{V}_h(K), \quad (4.9a)$$

$$(\boldsymbol{\Sigma} \nabla u_h \cdot \nabla \psi_{\mathbf{a}})|_K \in Q_h(K), \quad (4.9b)$$

$$(\mathbf{R}_{\frac{\pi}{2}} \nabla(\psi_{\mathbf{a}} u_h))|_K \in \mathbf{V}_h(K). \quad (4.9c)$$

Moreover, let there hold

$$\langle \llbracket u_h \rrbracket, 1 \rangle_e = 0 \quad \forall e \in \mathcal{E}_h. \quad (4.10)$$

Remark 4.6 (Practical discrete spaces). *Typically, $u_h \in \mathbb{P}_p(\mathcal{T}_h)$ for some $p \geq 1$. Then, under Assumption 2.1, $\mathbf{V}_h \times Q_h := \mathbf{RTN}_p \times \mathbb{P}_p(\mathcal{T}_h)$ is the appropriate choice for the spaces \mathbf{V}_h and Q_h .*

Some more tools will now be needed. Let the patchwise Sobolev spaces be given by

$$\begin{aligned} H_*^1(\omega_{\mathbf{a}}) &:= \{v \in H^1(\omega_{\mathbf{a}}); (v, 1)_{\omega_{\mathbf{a}}} = 0\}, & \mathbf{a} \in \mathcal{V}_h^{\text{int}}, \\ H_*^1(\omega_{\mathbf{a}}) &:= \{v \in H^1(\omega_{\mathbf{a}}); v = 0 \text{ on } \partial\omega_{\mathbf{a}} \cap \partial\Omega\}, & \mathbf{a} \in \mathcal{V}_h^{\text{ext}}. \end{aligned}$$

It follows from [7, Theorem 3.1], [4, Section 3], see also [16, Lemma 3.12], that

$$\|\nabla(\psi_{\mathbf{a}} v)\|_{\omega_{\mathbf{a}}} \leq C_{\text{cont}, \text{PF}} \|\nabla v\|_{\omega_{\mathbf{a}}} \quad \forall v \in H_*^1(\omega_{\mathbf{a}}), \forall \mathbf{a} \in \mathcal{V}_h, \quad (4.11)$$

with a Poincaré–Friedrichs constant $C_{\text{cont}, \text{PF}}$ only depending on the shape regularity parameter $\kappa_{\mathcal{T}}$. Similarly, let the patchwise broken Sobolev spaces be given by

$$\begin{aligned} H_{\sharp}^1(\mathcal{T}_{\mathbf{a}}) &:= \{v \in H^1(\mathcal{T}_{\mathbf{a}}); \langle \llbracket v \rrbracket, 1 \rangle_e = 0 \quad \forall e \in \mathcal{E}_h^{\text{int}}, \mathbf{a} \in e, \quad (v, 1)_{\omega_{\mathbf{a}}} = 0\} & \mathbf{a} \in \mathcal{V}_h^{\text{int}}, \\ H_{\sharp}^1(\mathcal{T}_{\mathbf{a}}) &:= \{v \in H^1(\mathcal{T}_{\mathbf{a}}); \langle \llbracket v \rrbracket, 1 \rangle_e = 0 \quad \forall e \in \mathcal{E}_h, \mathbf{a} \in e\}, & \mathbf{a} \in \mathcal{V}_h^{\text{ext}}. \end{aligned}$$

It follows from [16, Lemma 3.13] that

$$\|\nabla(\psi_{\mathbf{a}}v)\|_{\omega_{\mathbf{a}}} \leq C_{\text{cont,bPF}}\|\nabla v\|_{\omega_{\mathbf{a}}} \quad \forall v \in H_{\#}^1(\mathcal{T}_{\mathbf{a}}), \forall \mathbf{a} \in \mathcal{V}_h, \quad (4.12)$$

with a broken Poincaré–Friedrichs constant $C_{\text{cont,bPF}}$ only depending on $\kappa_{\mathcal{T}}$. Let also

$$H_b^1(\omega_{\mathbf{a}}) := \{v \in H^1(\omega_{\mathbf{a}}); (v, 1)_{\omega_{\mathbf{a}}} = 0\}, \quad \mathbf{a} \in \mathcal{V}_h, \quad (4.13)$$

and

$$H_{\#}^1(\omega_{\mathbf{a}}) := H^1(\omega_{\mathbf{a}}), \quad \mathbf{a} \in \mathcal{V}_h^{\text{int}}, \quad (4.14a)$$

$$H_{\#}^1(\omega_{\mathbf{a}}) := \{v \in H^1(\omega_{\mathbf{a}}); v = 0 \text{ on } \partial\omega_{\mathbf{a}} \cap \partial\Omega\}, \quad \mathbf{a} \in \mathcal{V}_h^{\text{ext}}. \quad (4.14b)$$

Let C_{st} be the constant from the mixed finite element stability

$$\|\psi_{\mathbf{a}}\underline{\Sigma}\nabla u_h + \mathfrak{S}_h^{\mathbf{a}}\|_{\omega_{\mathbf{a}}} \leq C_{\text{st}} \sup_{v \in H_{\#}^1(\omega_{\mathbf{a}}); \|\nabla v\|_{\omega_{\mathbf{a}}}=1} \{-(\psi_{\mathbf{a}}\underline{\Sigma}\nabla u_h, \nabla v)_{\omega_{\mathbf{a}}} + (\Pi_{Q_h}(\psi_{\mathbf{a}}f) - \underline{\Sigma}\nabla u_h, \nabla \psi_{\mathbf{a}}, v)_{\omega_{\mathbf{a}}}\}, \quad (4.15a)$$

$$\|\mathbb{R}_{\frac{\pi}{2}}\nabla(\psi_{\mathbf{a}}u_h) + \mathfrak{S}_h^{\mathbf{a}}\|_{\omega_{\mathbf{a}}} \leq C_{\text{st}} \sup_{v \in H_b^1(\omega_{\mathbf{a}}); \|\nabla v\|_{\omega_{\mathbf{a}}}=1} \{-(\mathbb{R}_{\frac{\pi}{2}}\nabla(\psi_{\mathbf{a}}u_h), \nabla v)_{\omega_{\mathbf{a}}}\}, \quad (4.15b)$$

for respectively the problems (4.3) and (4.5), see [4, Theorem 7] and [16, Corollary 3.16]; a computable upper bound for C_{st} is given in [16, Lemma 3.23]. Define finally a modified data oscillation term

$$\tilde{\eta}_{\text{osc},K} := \frac{h_K}{\pi} \|\psi_{\mathbf{a}}f - \Pi_{Q_h}(\psi_{\mathbf{a}}f)\|_K, \quad K \in \mathcal{T}_h, \quad (4.16)$$

together with $\tilde{\eta}_{\text{osc}} := \{\sum_{K \in \mathcal{T}_h} (\tilde{\eta}_{\text{osc},K})^2\}^{1/2}$. We have:

Theorem 4.7 (Local and global efficiency and robustness for Theorem 4.4). *Let u be the weak solution given by (1.2) and let u_h verify Assumption 4.5. Then, for σ_h given by Definition 4.2,*

$$\|\underline{\Sigma}\nabla u_h + \sigma_h\|_K \leq C_{\text{st}}C_{\text{cont,PF}} \sum_{\mathbf{a} \in \mathcal{V}_K} \|u - u_h\|_{\#, \omega_{\mathbf{a}}} + C_{\text{st}} \sum_{\mathbf{a} \in \mathcal{V}_K} \left\{ \sum_{K' \in \mathcal{T}_{\mathbf{a}}} \tilde{\eta}_{\text{osc},K'}^2 \right\}^{1/2} \quad \forall K \in \mathcal{T}_h, \quad (4.17a)$$

$$\|\underline{\Sigma}\nabla u_h + \sigma_h\| \leq 3C_{\text{st}}C_{\text{cont,PF}}\|u - u_h\|_{\#} + 3C_{\text{st}}\tilde{\eta}_{\text{osc}}, \quad (4.17b)$$

and, for s_h given by Definition 4.3,

$$\|\nabla(u_h - s_h)\|_K \leq C_{\text{st}}C_{\text{cont,bPF}} \sum_{\mathbf{a} \in \mathcal{V}_K} \inf_{s^{\mathbf{a}} \in H_{\#}^1(\omega_{\mathbf{a}})} \|\nabla(u_h - s^{\mathbf{a}})\|_{\omega_{\mathbf{a}}} \quad \forall K \in \mathcal{T}_h, \quad (4.18a)$$

$$\|\nabla(u_h - s_h)\| \leq 3C_{\text{st}}C_{\text{cont,bPF}} \inf_{s \in H_0^1(\Omega)} \|\nabla(u_h - s)\|. \quad (4.18b)$$

Proof. Assertion (4.17a) follows as in [16, Theorem 3.17], whereas inequality (4.17b) can be shown as in [16, Lemma 3.22]. We now present the proof of (4.18a). Using Definition 4.3, the potential reconstruction decomposition $s_h|_K = \sum_{\mathbf{a} \in \mathcal{V}_K} s_h^{\mathbf{a}}|_K$, the partition of unity $\sum_{\mathbf{a} \in \mathcal{V}_K} \psi_{\mathbf{a}}|_K = 1|_K$, and the triangle inequality,

$$\begin{aligned} \|\nabla(u_h - s_h)\|_K &= \left\| \sum_{\mathbf{a} \in \mathcal{V}_K} (\nabla(\psi_{\mathbf{a}}u_h - s_h^{\mathbf{a}}))|_K \right\|_K \leq \sum_{\mathbf{a} \in \mathcal{V}_K} \|\nabla(\psi_{\mathbf{a}}u_h - s_h^{\mathbf{a}})\|_K \\ &= \sum_{\mathbf{a} \in \mathcal{V}_K} \|\mathbb{R}_{\frac{\pi}{2}}\nabla(\psi_{\mathbf{a}}u_h - s_h^{\mathbf{a}})\|_K = \sum_{\mathbf{a} \in \mathcal{V}_K} \|\mathbb{R}_{\frac{\pi}{2}}\nabla(\psi_{\mathbf{a}}u_h) + \mathfrak{S}_h^{\mathbf{a}}\|_K \\ &\leq \sum_{\mathbf{a} \in \mathcal{V}_K} \|\mathbb{R}_{\frac{\pi}{2}}\nabla(\psi_{\mathbf{a}}u_h) + \mathfrak{S}_h^{\mathbf{a}}\|_{\omega_{\mathbf{a}}}. \end{aligned}$$

We next observe that, for an arbitrary $s^{\mathbf{a}} \in H_{\#}^1(\omega_{\mathbf{a}})$,

$$(\mathbb{R}_{\frac{\pi}{2}}\nabla(\psi_{\mathbf{a}}s^{\mathbf{a}}), \nabla v)_{\omega_{\mathbf{a}}} = 0 \quad \forall v \in H_b^1(\omega_{\mathbf{a}}),$$

with the spaces $H_b^1(\omega_{\mathbf{a}})$ and $H_{\#}^1(\omega_{\mathbf{a}})$ defined respectively in (4.13) and (4.14). Let $H_{\#,u_h}^1(\omega_{\mathbf{a}}) := H_{\#}^1(\omega_{\mathbf{a}})$ for $\mathbf{a} \in \mathcal{V}_h^{\text{ext}}$ and $H_{\#,u_h}^1(\omega_{\mathbf{a}}) := \{v \in H_{\#}^1(\omega_{\mathbf{a}}); (v, 1)_{\omega_{\mathbf{a}}} = (u_h, 1)_{\omega_{\mathbf{a}}}\}$ for $\mathbf{a} \in \mathcal{V}_h^{\text{int}}$. Then, using the mixed finite element stability (4.15b), the Cauchy–Schwarz inequality, the fact that $u_h - s^{\mathbf{a}} \in H_{\#}^1(\mathcal{T}_{\mathbf{a}})$ for an arbitrary $s^{\mathbf{a}} \in H_{\#,u_h}^1(\omega_{\mathbf{a}})$, and (4.12),

$$\begin{aligned} \|\mathbb{R}_{\frac{\pi}{2}} \nabla(\psi_{\mathbf{a}} u_h) + \mathfrak{S}_h^{\mathbf{a}}\|_{\omega_{\mathbf{a}}} &\leq C_{\text{st}} \sup_{v \in H_b^1(\omega_{\mathbf{a}}); \|\nabla v\|_{\omega_{\mathbf{a}}}=1} \inf_{s^{\mathbf{a}} \in H_{\#,u_h}^1(\omega_{\mathbf{a}})} -(\mathbb{R}_{\frac{\pi}{2}} \nabla(\psi_{\mathbf{a}}(u_h - s^{\mathbf{a}})), \nabla v)_{\omega_{\mathbf{a}}} \\ &\leq C_{\text{st}} \inf_{s^{\mathbf{a}} \in H_{\#,u_h}^1(\omega_{\mathbf{a}})} \|\mathbb{R}_{\frac{\pi}{2}} \nabla(\psi_{\mathbf{a}}(u_h - s^{\mathbf{a}}))\|_{\omega_{\mathbf{a}}} \\ &\leq C_{\text{st}} C_{\text{cont,bPF}} \inf_{s^{\mathbf{a}} \in H_{\#,u_h}^1(\omega_{\mathbf{a}})} \|\nabla(u_h - s^{\mathbf{a}})\|_{\omega_{\mathbf{a}}} \\ &= C_{\text{st}} C_{\text{cont,bPF}} \inf_{s^{\mathbf{a}} \in H_{\#}^1(\omega_{\mathbf{a}})} \|\nabla(u_h - s^{\mathbf{a}})\|_{\omega_{\mathbf{a}}}. \end{aligned}$$

The global efficiency is shown as follows: As in [16, Lemma 3.22], we develop

$$\begin{aligned} \|\nabla(u_h - s_h)\|^2 &= \sum_{K \in \mathcal{T}_h} \|\nabla(u_h - s_h)\|_K^2 = \sum_{K \in \mathcal{T}_h} \left\| \sum_{\mathbf{a} \in \mathcal{V}_K} (\nabla(\psi_{\mathbf{a}} u_h - s_h^{\mathbf{a}}))|_K \right\|_K^2 \\ &\leq 3 \sum_{K \in \mathcal{T}_h} \sum_{\mathbf{a} \in \mathcal{V}_K} \|\nabla(\psi_{\mathbf{a}} u_h - s_h^{\mathbf{a}})\|_K^2 = 3 \sum_{\mathbf{a} \in \mathcal{V}_h} \|\nabla(\psi_{\mathbf{a}} u_h - s_h^{\mathbf{a}})\|_{\omega_{\mathbf{a}}}^2. \end{aligned}$$

Using $\|\nabla(\psi_{\mathbf{a}} u_h - s_h^{\mathbf{a}})\|_{\omega_{\mathbf{a}}} = \|\mathbb{R}_{\frac{\pi}{2}} \nabla(\psi_{\mathbf{a}} u_h) + \mathfrak{S}_h^{\mathbf{a}}\|_{\omega_{\mathbf{a}}}$ and the above estimate for the latter term proves

$$\|\nabla(u_h - s_h)\|^2 \leq 3 C_{\text{st}}^2 C_{\text{cont,bPF}}^2 \sum_{\mathbf{a} \in \mathcal{V}_h} \inf_{s^{\mathbf{a}} \in H_{\#}^1(\omega_{\mathbf{a}})} \|\nabla(u_h - s^{\mathbf{a}})\|_{\omega_{\mathbf{a}}}^2.$$

Finally,

$$\sum_{\mathbf{a} \in \mathcal{V}_h} \inf_{s^{\mathbf{a}} \in H_{\#}^1(\omega_{\mathbf{a}})} \|\nabla(u_h - s^{\mathbf{a}})\|_{\omega_{\mathbf{a}}}^2 \leq \inf_{s \in H_0^1(\Omega)} \sum_{\mathbf{a} \in \mathcal{V}_h} \|\nabla(u_h - s)\|_{\omega_{\mathbf{a}}}^2 \leq 3 \inf_{s \in H_0^1(\Omega)} \|\nabla(u_h - s)\|^2 \quad (4.19)$$

shows the assertion (4.18b). □

Remark 4.8 (Efficiency in the L^2 flux norm for Theorem 4.7). *Owing to (2.5), we also have*

$$\begin{aligned} \|\underline{\Sigma} \nabla u_h + \boldsymbol{\sigma}_h\|_K &\leq C_{\text{st}} C_{\text{cont,PF}} \sum_{\mathbf{a} \in \mathcal{V}_K} \|\underline{\Sigma} \nabla(u - u_h)\|_{\omega_{\mathbf{a}}} + C_{\text{st}} \sum_{\mathbf{a} \in \mathcal{V}_K} \left\{ \sum_{K' \in \mathcal{T}_{\mathbf{a}}} \tilde{\eta}_{\text{osc},K'}^2 \right\}^{1/2} \quad \forall K \in \mathcal{T}_h, \\ \|\underline{\Sigma} \nabla u_h + \boldsymbol{\sigma}_h\| &\leq 3 C_{\text{st}} C_{\text{cont,PF}} \|\underline{\Sigma} \nabla(u - u_h)\| + 3 C_{\text{st}} \tilde{\eta}_{\text{osc}}. \end{aligned}$$

Applications to conforming, nonconforming, mixed, and discontinuous Galerkin approximations are straightforward following [16, Section 4].

5 Equivalent localization of global dual norms

This section presents some relations between local and global versions of dual (residual) norms, not necessarily linked to a posteriori error analysis, that we find of independent interest. In particular, the results presented below allow to pass from merely global to actually local efficiency in [27, 29, 14]. Here again, the general space dimension case $d \geq 2$ is treated. To present our results, recall first the inequalities (4.11) and (4.12) for the spaces $H_*^1(\omega_{\mathbf{a}})$ and $H_{\#}^1(\mathcal{T}_{\mathbf{a}})$, respectively.

Theorem 5.1 (Localization of a dual norm with $\psi_{\mathbf{a}}$ -Galerkin orthogonality). *Let $\boldsymbol{\xi} \in [L^2(\Omega)]^d$ such that*

$$(\boldsymbol{\xi}, \nabla \psi_{\mathbf{a}})_{\omega_{\mathbf{a}}} = 0 \quad \forall \mathbf{a} \in \mathcal{V}_h^{\text{int}} \quad (5.1)$$

be arbitrary. Then

$$\sup_{\varphi \in H_0^1(\Omega); \|\nabla\varphi\|=1} (\boldsymbol{\xi}, \nabla\varphi)^2 \leq (d+1)C_{\text{cont,PF}}^2 \sum_{\mathbf{a} \in \mathcal{V}_h} \sup_{\varphi \in H_0^1(\omega_{\mathbf{a}}); \|\nabla\varphi\|_{\omega_{\mathbf{a}}}=1} (\boldsymbol{\xi}, \nabla\varphi)_{\omega_{\mathbf{a}}}^2, \quad (5.2a)$$

$$\sum_{\mathbf{a} \in \mathcal{V}_h} \sup_{\varphi \in H_0^1(\omega_{\mathbf{a}}); \|\nabla\varphi\|_{\omega_{\mathbf{a}}}=1} (\boldsymbol{\xi}, \nabla\varphi)_{\omega_{\mathbf{a}}}^2 \leq (d+1) \sup_{\varphi \in H_0^1(\Omega); \|\nabla\varphi\|=1} (\boldsymbol{\xi}, \nabla\varphi)^2. \quad (5.2b)$$

Proof. Let $\varphi \in H_0^1(\Omega)$ with $\|\nabla\varphi\| = 1$ be fixed. The partition of unity of $\psi_{\mathbf{a}}$, $\sum_{\mathbf{a} \in \mathcal{V}_h} \psi_{\mathbf{a}} = 1$, and the Galerkin orthogonality with respect to $\psi_{\mathbf{a}}$ (5.1) give

$$\begin{aligned} (\boldsymbol{\xi}, \nabla\varphi) &= \sum_{\mathbf{a} \in \mathcal{V}_h} (\boldsymbol{\xi}, \nabla(\psi_{\mathbf{a}}\varphi)) = \sum_{\mathbf{a} \in \mathcal{V}_h} (\boldsymbol{\xi}, \nabla(\psi_{\mathbf{a}}\varphi))_{\omega_{\mathbf{a}}} \\ &= \sum_{\mathbf{a} \in \mathcal{V}_h^{\text{int}}} (\boldsymbol{\xi}, \nabla(\psi_{\mathbf{a}}(\varphi - \Pi_{0,\omega_{\mathbf{a}}}\varphi)))_{\omega_{\mathbf{a}}} + \sum_{\mathbf{a} \in \mathcal{V}_h^{\text{ext}}} (\boldsymbol{\xi}, \nabla(\psi_{\mathbf{a}}\varphi))_{\omega_{\mathbf{a}}}, \end{aligned}$$

where $\Pi_{0,\omega_{\mathbf{a}}}\varphi$ is the mean value of φ on $\omega_{\mathbf{a}}$. There holds $(\varphi - \Pi_{0,\omega_{\mathbf{a}}}\varphi)|_{\omega_{\mathbf{a}}} \in H_*^1(\omega_{\mathbf{a}})$ and $(\psi_{\mathbf{a}}(\varphi - \Pi_{0,\omega_{\mathbf{a}}}\varphi))|_{\omega_{\mathbf{a}}} \in H_0^1(\omega_{\mathbf{a}})$ for $\mathbf{a} \in \mathcal{V}_h^{\text{int}}$ and $\varphi|_{\omega_{\mathbf{a}}} \in H_*^1(\omega_{\mathbf{a}})$ and $(\psi_{\mathbf{a}}\varphi)|_{\omega_{\mathbf{a}}} \in H_0^1(\omega_{\mathbf{a}})$ for $\mathbf{a} \in \mathcal{V}_h^{\text{ext}}$. Thus, passing to a supremum and using inequality (4.11) yields, for $\mathbf{a} \in \mathcal{V}_h^{\text{int}}$,

$$\begin{aligned} (\boldsymbol{\xi}, \nabla(\psi_{\mathbf{a}}(\varphi - \Pi_{0,\omega_{\mathbf{a}}}\varphi)))_{\omega_{\mathbf{a}}} &\leq \|\nabla(\psi_{\mathbf{a}}(\varphi - \Pi_{0,\omega_{\mathbf{a}}}\varphi))\|_{\omega_{\mathbf{a}}} \sup_{v \in H_0^1(\omega_{\mathbf{a}}); \|\nabla v\|_{\omega_{\mathbf{a}}}=1} (\boldsymbol{\xi}, \nabla v)_{\omega_{\mathbf{a}}} \\ &\leq C_{\text{cont,PF}} \|\nabla(\varphi - \Pi_{0,\omega_{\mathbf{a}}}\varphi)\|_{\omega_{\mathbf{a}}} \sup_{v \in H_0^1(\omega_{\mathbf{a}}); \|\nabla v\|_{\omega_{\mathbf{a}}}=1} (\boldsymbol{\xi}, \nabla v)_{\omega_{\mathbf{a}}} \\ &= C_{\text{cont,PF}} \|\nabla\varphi\|_{\omega_{\mathbf{a}}} \sup_{v \in H_0^1(\omega_{\mathbf{a}}); \|\nabla v\|_{\omega_{\mathbf{a}}}=1} (\boldsymbol{\xi}, \nabla v)_{\omega_{\mathbf{a}}}. \end{aligned}$$

A similar estimate holds for $\mathbf{a} \in \mathcal{V}_h^{\text{ext}}$. Thus, the Cauchy–Schwarz inequality gives

$$(\boldsymbol{\xi}, \nabla\varphi)^2 \leq C_{\text{cont,PF}}^2 \sum_{\mathbf{a} \in \mathcal{V}_h} \|\nabla\varphi\|_{\omega_{\mathbf{a}}}^2 \sum_{\mathbf{a} \in \mathcal{V}_h} \sup_{v \in H_0^1(\omega_{\mathbf{a}}); \|\nabla v\|_{\omega_{\mathbf{a}}}=1} (\boldsymbol{\xi}, \nabla v)_{\omega_{\mathbf{a}}}^2$$

and the fact that each simplex has $d+1$ vertices with $\|\nabla\varphi\| = 1$ yield (5.2a).

The converse estimate (5.2b) in fact does not need the hypothesis (5.1). Let $\mathbf{a} \in \mathcal{V}_h$ and let $v^{\mathbf{a}} \in H_0^1(\omega_{\mathbf{a}})$ be defined by the lifting

$$(\nabla v^{\mathbf{a}}, \nabla\varphi)_{\omega_{\mathbf{a}}} = (\boldsymbol{\xi}, \nabla\varphi)_{\omega_{\mathbf{a}}} \quad \forall \varphi \in H_0^1(\omega_{\mathbf{a}}).$$

Then

$$\|\nabla v^{\mathbf{a}}\|_{\omega_{\mathbf{a}}}^2 = (\boldsymbol{\xi}, \nabla v^{\mathbf{a}})_{\omega_{\mathbf{a}}} = \sup_{\varphi \in H_0^1(\omega_{\mathbf{a}}); \|\nabla\varphi\|_{\omega_{\mathbf{a}}}=1} (\boldsymbol{\xi}, \nabla\varphi)_{\omega_{\mathbf{a}}}^2$$

and

$$\begin{aligned} \sum_{\mathbf{a} \in \mathcal{V}_h} \sup_{\varphi \in H_0^1(\omega_{\mathbf{a}}); \|\nabla\varphi\|_{\omega_{\mathbf{a}}}=1} (\boldsymbol{\xi}, \nabla\varphi)_{\omega_{\mathbf{a}}}^2 &= \sum_{\mathbf{a} \in \mathcal{V}_h} (\boldsymbol{\xi}, \nabla v^{\mathbf{a}})_{\omega_{\mathbf{a}}} = (\boldsymbol{\xi}, \nabla v) \\ &\leq \sup_{\varphi \in H_0^1(\Omega); \|\nabla\varphi\|=1} (\boldsymbol{\xi}, \nabla\varphi) \|\nabla v\|, \end{aligned}$$

where we have set $v := \sum_{\mathbf{a} \in \mathcal{V}_h} v^{\mathbf{a}} \in H_0^1(\Omega)$ and passed to the supremum. Noticing that

$$\|\nabla v\|^2 = \sum_{K \in \mathcal{T}_h} \left\| \sum_{\mathbf{a} \in \mathcal{V}_K} (\nabla v^{\mathbf{a}})|_K \right\|_K^2 \leq (d+1) \sum_{\mathbf{a} \in \mathcal{V}_h} \|\nabla v^{\mathbf{a}}\|_{\omega_{\mathbf{a}}}^2,$$

we arrive at (5.2b). \square

A particular consequence, letting $\boldsymbol{\xi} = \underline{\Sigma}\nabla(u - u_h)$ and noticing that (1.2) and (4.1b) give (5.1), is:

Corollary 5.2 (Localization of the dual norm of the residual). *Let u be the weak solution given by (1.2) and let u_h satisfying Assumption 4.1 be arbitrary. Then*

$$\begin{aligned} \|u - u_h\|_{\#}^2 &\leq (d+1)C_{\text{cont,PF}}^2 \sum_{\mathbf{a} \in \mathcal{V}_h} \|u - u_h\|_{\#, \omega_{\mathbf{a}}}^2, \\ \sum_{\mathbf{a} \in \mathcal{V}_h} \|u - u_h\|_{\#, \omega_{\mathbf{a}}}^2 &\leq (d+1)\|u - u_h\|_{\#}^2. \end{aligned}$$

It appears that the second term in the total error measure (2.3) admits a similar localization property:

Theorem 5.3 (Localization of the potential nonconformity evaluation). *Let $u_h \in H^1(\mathcal{T}_h)$ satisfying (4.10) be arbitrary. Then*

$$\begin{aligned} \inf_{s \in H_0^1(\Omega)} \|\nabla(u_h - s)\|^2 &\leq (d+1)C_{\text{cont,bPF}}^2 \sum_{\mathbf{a} \in \mathcal{V}_h} \inf_{s^{\mathbf{a}} \in H_{\#}^1(\omega_{\mathbf{a}})} \|\nabla(u_h - s^{\mathbf{a}})\|_{\omega_{\mathbf{a}}}^2, \\ \sum_{\mathbf{a} \in \mathcal{V}_h} \inf_{s^{\mathbf{a}} \in H_{\#}^1(\omega_{\mathbf{a}})} \|\nabla(u_h - s^{\mathbf{a}})\|_{\omega_{\mathbf{a}}}^2 &\leq (d+1) \inf_{s \in H_0^1(\Omega)} \|\nabla(u_h - s)\|^2. \end{aligned}$$

Proof. The second claim follows exactly as in (4.19). To show the first one, let, for $\mathbf{a} \in \mathcal{V}_h$, $v^{\mathbf{a}}$ be given by the orthogonal projection of $\psi_{\mathbf{a}}u_h$ onto $H_0^1(\omega_{\mathbf{a}})$,

$$v^{\mathbf{a}} := \arg \inf_{v \in H_0^1(\omega_{\mathbf{a}})} \|\nabla(\psi_{\mathbf{a}}u_h - v)\|_{\omega_{\mathbf{a}}};$$

equivalently, $v^{\mathbf{a}} \in H_0^1(\omega_{\mathbf{a}})$ solves

$$(\nabla v^{\mathbf{a}}, \nabla \varphi)_{\omega_{\mathbf{a}}} = (\nabla(\psi_{\mathbf{a}}u_h), \nabla \varphi)_{\omega_{\mathbf{a}}} \quad \forall \varphi \in H_0^1(\omega_{\mathbf{a}}).$$

Setting $v := \sum_{\mathbf{a} \in \mathcal{V}_h} v^{\mathbf{a}} \in H_0^1(\Omega)$, we have

$$\begin{aligned} \inf_{s \in H_0^1(\Omega)} \|\nabla(u_h - s)\|^2 &\leq \|\nabla(u_h - v)\|^2 = \sum_{K \in \mathcal{T}_h} \left\| \sum_{\mathbf{a} \in \mathcal{V}_K} (\nabla(\psi_{\mathbf{a}}u_h - v^{\mathbf{a}})) \right\|_K^2 \\ &\leq (d+1) \sum_{\mathbf{a} \in \mathcal{V}_h} \|\nabla(\psi_{\mathbf{a}}u_h - v^{\mathbf{a}})\|_{\omega_{\mathbf{a}}}^2. \end{aligned}$$

The fact that $\psi_{\mathbf{a}}s^{\mathbf{a}} \in H_0^1(\omega_{\mathbf{a}})$ for any $s^{\mathbf{a}} \in H_{\#}^1(\omega_{\mathbf{a}})$ gives

$$\|\nabla(\psi_{\mathbf{a}}u_h - v^{\mathbf{a}})\|_{\omega_{\mathbf{a}}} \leq \inf_{s^{\mathbf{a}} \in H_{\#}^1(\omega_{\mathbf{a}})} \|\nabla(\psi_{\mathbf{a}}(u_h - s^{\mathbf{a}}))\|_{\omega_{\mathbf{a}}}.$$

Finally, inequality (4.12) while proceeding as in the proof of Theorem 4.7 concerning the mean value of $u_h - s^{\mathbf{a}}$ gives

$$\inf_{s^{\mathbf{a}} \in H_{\#}^1(\omega_{\mathbf{a}})} \|\nabla(\psi_{\mathbf{a}}(u_h - s^{\mathbf{a}}))\|_{\omega_{\mathbf{a}}} \leq C_{\text{cont,bPF}} \inf_{s^{\mathbf{a}} \in H_{\#}^1(\omega_{\mathbf{a}})} \|\nabla(u_h - s^{\mathbf{a}})\|_{\omega_{\mathbf{a}}},$$

and collecting the results finishes the proof. \square

Let $\mathbf{H}_*(\text{div}, \omega_{\mathbf{a}})$ stand for $\mathbf{H}(\text{div}, \omega_{\mathbf{a}})$ functions with zero normal trace in the appropriate sense on $\partial\omega_{\mathbf{a}}$ for $\mathbf{a} \in \mathcal{V}_h^{\text{int}}$ and for $\mathbf{H}(\text{div}, \omega_{\mathbf{a}})$ functions with zero normal trace in the appropriate sense on $\partial\omega_{\mathbf{a}} \setminus \partial\Omega$ for $\mathbf{a} \in \mathcal{V}_h^{\text{ext}}$. One can show similarly as above that:

Theorem 5.4 (Localization of the flux nonconformity evaluation). *Let u_h satisfying Assumption 4.1 be arbitrary. Then*

$$\begin{aligned} \inf_{\sigma \in \mathbf{H}(\text{div}, \Omega); \nabla \cdot \sigma = f} \|\underline{\Sigma} \nabla u_h + \sigma\|^2 &\leq (d+1) \sum_{\mathbf{a} \in \mathcal{V}_h} \inf_{\varsigma^{\mathbf{a}} \in \mathbf{H}_*(\text{div}, \omega_{\mathbf{a}}); \nabla \cdot \varsigma^{\mathbf{a}} = \psi_{\mathbf{a}}f - \underline{\Sigma} \nabla u_h \cdot \nabla \psi_{\mathbf{a}}} \|\psi_{\mathbf{a}} \underline{\Sigma} \nabla u_h + \varsigma^{\mathbf{a}}\|_{\omega_{\mathbf{a}}}^2, \\ \sum_{\mathbf{a} \in \mathcal{V}_h} \inf_{\varsigma^{\mathbf{a}} \in \mathbf{H}_*(\text{div}, \omega_{\mathbf{a}}); \nabla \cdot \varsigma^{\mathbf{a}} = \psi_{\mathbf{a}}f - \underline{\Sigma} \nabla u_h \cdot \nabla \psi_{\mathbf{a}}} \|\psi_{\mathbf{a}} \underline{\Sigma} \nabla u_h + \varsigma^{\mathbf{a}}\|_{\omega_{\mathbf{a}}}^2 &\leq (d+1)C_{\text{cont,PF}}^2 \inf_{\sigma \in \mathbf{H}(\text{div}, \Omega); \nabla \cdot \sigma = f} \|\underline{\Sigma} \nabla u_h + \sigma\|^2. \end{aligned}$$

6 Numerical experiments

We report here the results of some numerical experiments, while relying on the conforming piecewise affine finite element approximation in the form (3.6) (with $V_h := \mathbb{P}_1(\mathcal{T}_h) \cap H_0^1(\Omega)$) and on the estimators of Section 3.2.

We will trace the error upper bound $\|\underline{\Sigma}\nabla(u - u_h)\|$ and the error lower bound $\frac{(\underline{\Sigma}\nabla(u - u_h), \nabla(\mathbb{T}(u - u_h)))}{\|\nabla(\mathbb{T}(u - u_h))\|}$ (see (2.1)), the canonical $H_0^1(\Omega)$ -norm error $\|\nabla(u - u_h)\|$, and the overall estimate of (3.11)

$$\eta := \left\{ \sum_{K \in \mathcal{T}_h} (\|\underline{\Sigma}\nabla u_h + \sigma_h\|_K + \eta_{\text{osc}, K})^2 \right\}^{1/2}.$$

The effectivity index for our dual error norm $\|u - u_h\|_{\#}$,

$$\text{Eff} := \frac{\eta}{\|u - u_h\|_{\#}},$$

is not calculable as it is not straightforward to compute $\|u - u_h\|_{\#}$. We thus present the upper (for error upper bound) and lower (for error lower bound) effectivity indices given respectively as

$$\text{Eff}_+ := \frac{\eta}{\|\underline{\Sigma}\nabla(u - u_h)\|}, \quad \text{Eff}_- := \frac{\eta}{\frac{(\underline{\Sigma}\nabla(u - u_h), \nabla(\mathbb{T}(u - u_h)))}{\|\nabla(\mathbb{T}(u - u_h))\|}}.$$

Thanks to (2.1), the effectivity index Eff needs to lie between these two indices. Note, however, that the upper effectivity index Eff_+ can become smaller than 1, which is not in contradiction with the guaranteed upper bound of Theorem 3.6 on $\|u - u_h\|_{\#}$.

Two test cases, one with a regular solution and one with a singular solution, are considered. Only uniform mesh refinement is used in the first case, whereas mesh adaptivity is employed in the second one. Here all elements where the estimator exceeds 50% of the maximal estimator value on the given mesh are refined regularly into four sub-elements and the so-called longest edge refinement is used so as to recover a matching mesh using the Matlab [21] PDE toolbox `refinemesh` algorithm.

In order to approximately compute the various errors, we use in each triangle the 7-point quadrature formula, see, e.g., [31, Section 9.10]. The additional error is negligible in the regular test case but not in the singular one. There, we proceed as follows. Recalling that $\underline{\Sigma}$ is piecewise constant on \mathcal{T}_h , we first develop

$$\begin{aligned} \|\underline{\Sigma}\nabla(u - u_h)\|^2 &= \|\underline{\Sigma}\nabla u\|^2 - 2(\underline{\Sigma}\nabla(u - u_h), \underline{\Sigma}\nabla u_h) - \|\underline{\Sigma}\nabla u_h\|^2 \\ &= \|\underline{\Sigma}\nabla u\|^2 - 2 \sum_{K \in \mathcal{T}_h} (f, \underline{\Sigma}u_h)_K - 2 \sum_{K \in \mathcal{T}_h} \langle \underline{\Sigma}\nabla(u - u_h) \cdot \mathbf{n}_K, \underline{\Sigma}u_h \rangle_{\partial K} - \|\underline{\Sigma}\nabla u_h\|^2. \end{aligned}$$

The second and third terms do not vanish as usual in the energy setting, since there is no Galerkin orthogonality for the flux error. As, however, the source term in the singular test case is zero, the second, volumetric, term does disappear. We are left with approximating the skeleton integrals in the third term that we undertake using the 5-point Gauss–Legendre quadrature formula on each half-edge of the given triangulation. We only use the 7-point quadrature formula to approximate the first term above on the finest adaptively refined mesh. This approximate error computation turns out to be much more precise. The errors $\frac{(\underline{\Sigma}\nabla(u - u_h), \nabla(\mathbb{T}(u - u_h)))}{\|\nabla(\mathbb{T}(u - u_h))\|}$ and $\|\nabla(u - u_h)\|$ are treated similarly.

6.1 A regular weak solution

We first consider the test case from [23, Section 5.1] with a regular solution. We set $\Omega := (-1, 1) \times (-1, 1)$ with $\Omega_+ := (0, 1) \times (-1, 1)$ and $\Omega_- := (-1, 0) \times (-1, 1)$ and let $\sigma_+ = 1$ and $\sigma_- < 0$. The exact solution is given by

$$\begin{aligned} u(x, y) &= \sigma_- x(x + 1)(x - 1)(y + 1)(y - 1) & \text{for } (x, y) \in \Omega_+, \\ u(x, y) &= x(x + 1)(x - 1)(y + 1)(y - 1) & \text{for } (x, y) \in \Omega_-, \end{aligned}$$

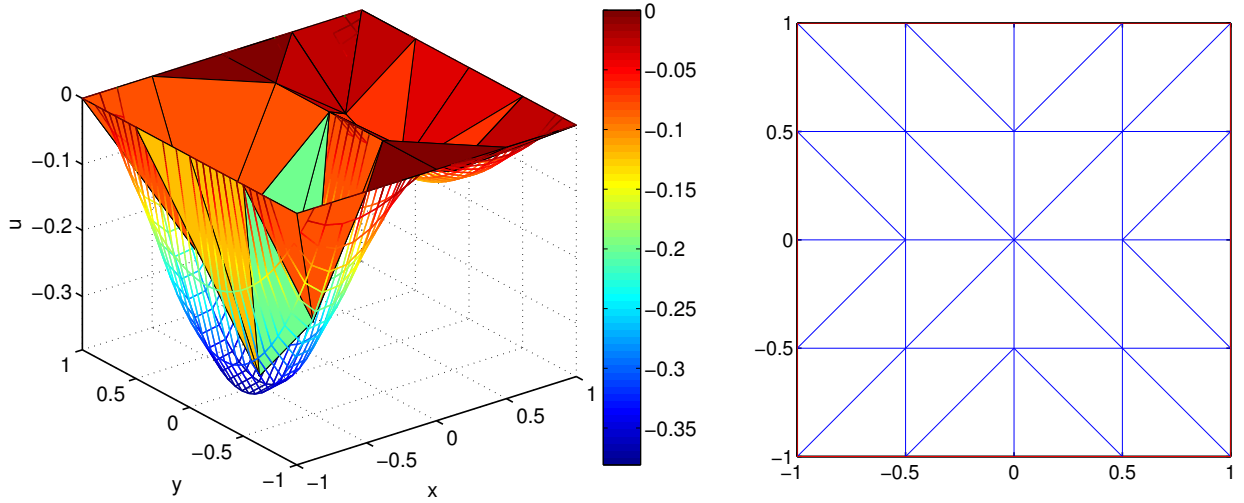


Figure 2: Exact and approximate solution (left) and the corresponding initial mesh (right), the regular case with $\sigma_- = -1/3$

and the source term (inhomogeneous) and the Dirichlet boundary condition (homogeneous) are prescribed accordingly. The exact solution together with its finite element approximation and the corresponding initial mesh for the setting $\sigma_- = -1/3$ are presented in Figure 2. The operator T is specified in [23], see also Remark 6.1 below; the problem is in particular well-posed when $\sigma_- \neq -1$.

The various errors and the estimate η of (3.11) are traced in the left parts of Figures 3–5, for three different choices of the parameter σ_- . The corresponding upper and lower effectivity indices are plotted in the right parts of these figures. We observe a systematic first-order decrease of all the errors, as predicted by the a priori error analysis, cf. [10]. The overall estimator η of (3.11), as well as its principal component given by $\{\sum_{K \in \mathcal{T}_h} \|\underline{\Sigma} \nabla u_h + \sigma_h\|_K^2\}^{1/2}$, also decrease with first order, in agreement with Theorems 3.6 and 3.7. On the other hand, the data oscillation estimator $\{\sum_{K \in \mathcal{S}_h} \eta_{\text{osc}, K}^2\}^{1/2}$, cf. (3.10), decreases with a slope of two and its influence rapidly diminishes.

The upper effectivity indices are in all settings very close to the optimal value of one. The lower effectivity index in the last test setting, very close to the well-posedness limit, however, deteriorates considerably. This is in no contradiction with the robustness of our estimates with respect to the intrinsic norm $\|u - u_h\|_{\#}$. It is the lower bound $\frac{(\underline{\Sigma} \nabla(u - u_h), \nabla(T(u - u_h)))}{\|\nabla(T(u - u_h))\|}$ which gets far from $\|u - u_h\|_{\#}$. Finally, Figure 6 illustrates that the distribution of the error is predicted very correctly; the comparison is done between the local estimators $\|\underline{\Sigma} \nabla u_h + \sigma_h\|_K$ and the upper error bounds $\|\underline{\Sigma} \nabla(u - u_h)\|_K$.

Remark 6.1 (Equivalence of the dual norm with its upper and lower bounds). *Interestingly enough, all $\|\underline{\Sigma} \nabla(u - u_h)\|$, $\|u - u_h\|_{\#}$, and $\frac{(\underline{\Sigma} \nabla(u - u_h), \nabla(T(u - u_h)))}{\|\nabla(T(u - u_h))\|}$ (and thus all the effectivity indices) coincide for the setting $\sigma_- = -1/3$. To explain this behavior, note first that $\|\underline{\Sigma} \nabla(u - u_h)\|$ and $\|u - u_h\|_{\#}$ will coincide whenever $\underline{\Sigma} \nabla(u - u_h)$ is a gradient of some scalar field from $H_0^1(\Omega)$. This will happen when curl of $\underline{\Sigma} \nabla(u - u_h) = 0$ on all $K \in \mathcal{T}_h$ and $\llbracket R_{\frac{\pi}{2}} \underline{\Sigma} \nabla(u - u_h) \cdot \mathbf{n} \rrbracket = 0$ on all $e \in \mathcal{E}_h$: this condition is satisfied in this test case, for all values of σ_- . To account for the other equality, one then notices that, when $\sigma_- \in (-1, 0)$, the operator T may be defined as*

$$Tu(x, y) = \begin{cases} u_+(x, y) & \text{for } (x, y) \in \Omega_+, \\ -u_-(x, y) + 2u_+(-x, y) & \text{for } (x, y) \in \Omega_- \end{cases}$$

for this test case, with $u_+ := u|_{\Omega_+}$ and $u_- := u|_{\Omega_-}$. The chosen exact solution being such that $u_+(x, y) = -\sigma_- u_-(-x, y)$ for $(x, y) \in \Omega_+$, the formula for Tu in Ω_- simplifies to $(Tu)|_{\Omega_-} = (-1 - 2\sigma_-)u_-$. With the

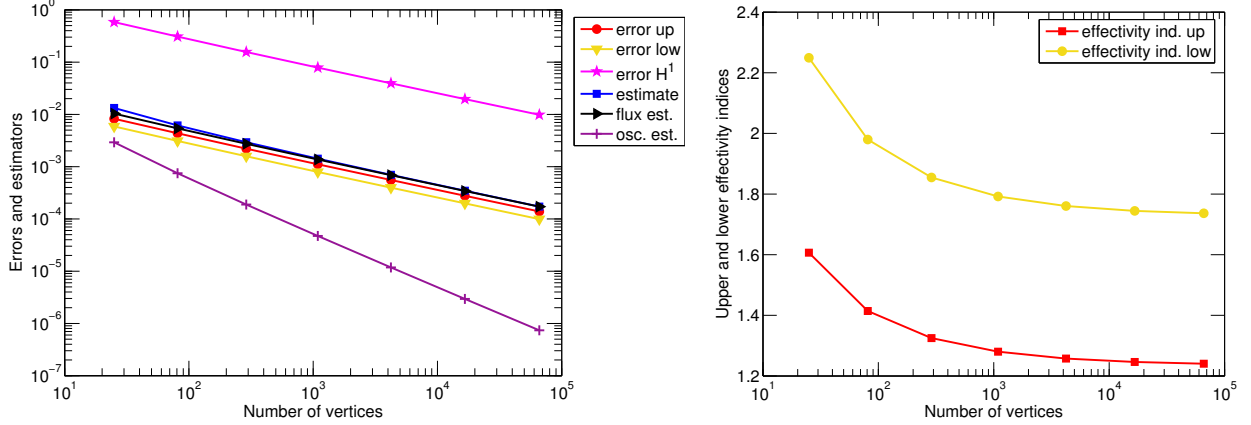


Figure 3: Estimates and errors for uniform mesh refinement (left) and the corresponding effectivity indices (right), the regular case with $\sigma_- = -0.01$

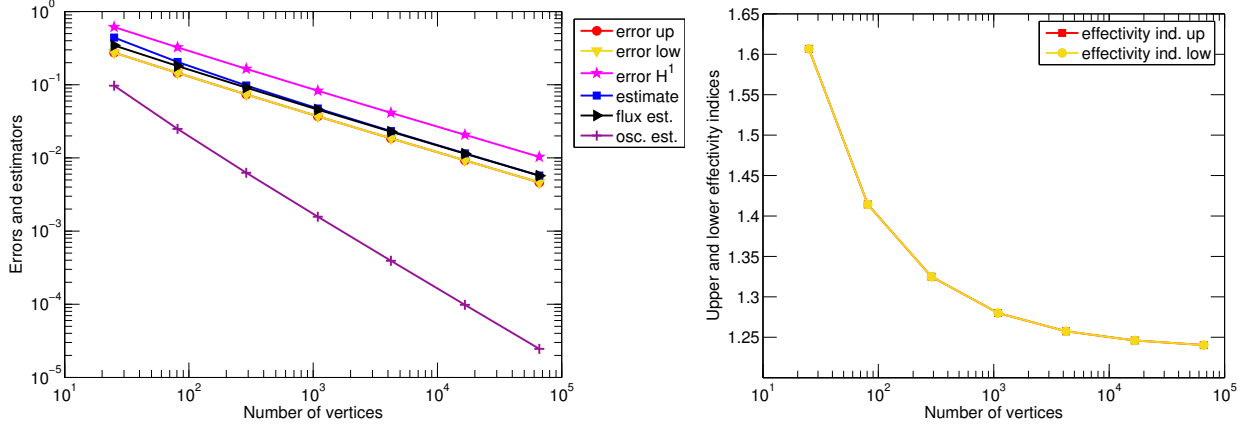


Figure 4: Estimates and errors for uniform mesh refinement (left) and the corresponding effectivity indices (right), the regular case with $\sigma_- = -1/3$

help of these expressions, one may compute exactly

$$R(\sigma_-) := \frac{(\underline{\Sigma}\nabla u, \nabla(\mathbb{T}u))}{\|\nabla(\mathbb{T}u)\| \|\underline{\Sigma}\nabla u\|} = \frac{(1 + 2\sigma_- + |\sigma_-|)}{(|1 + 2\sigma_-|^2 + |\sigma_-|^2)^{1/2} \sqrt{2}}.$$

For $\sigma_- \in (-1/3, 0)$, it holds that $R(\sigma_-) \in (1/\sqrt{2}, 1)$ and moreover $R(-1/3) = 1$, whereas $\lim_{\sigma_- \rightarrow -1} R(\sigma_-) = 0$. To obtain the same result for the ratio $\frac{(\underline{\Sigma}\nabla(u-u_h), \nabla(\mathbb{T}(u-u_h)))}{\|\nabla(\mathbb{T}(u-u_h))\| \|\underline{\Sigma}\nabla(u-u_h)\|}$, one needs to work with symmetric meshes with respect to the line $\{x = 0\}$, that is, globally \mathbb{T} -conform meshes in the sense of [10]. In this case, one has $\text{TV}_h = V_h$, so that the properties of \mathbb{T} at the continuous level carry over to the discrete level, whereas $u_h|_{\Omega_+}(x, y) = -\sigma_- u_h|_{\Omega_-}(-x, y)$ for $(x, y) \in \Omega_+$, by direct inspection on the formulation (3.1).

6.2 A singular weak solution

We next consider the test case from [23, Section 5.2] with a singular solution. We set $\Omega := (-1, 1) \times (-1, 1)$ with $\Omega_+ := (0, 1) \times (0, 1)$ and $\Omega_- := \Omega \setminus \overline{\Omega_+}$ and let again $\sigma_+ = 1$ and $\sigma_- < 0$. The exact solution is

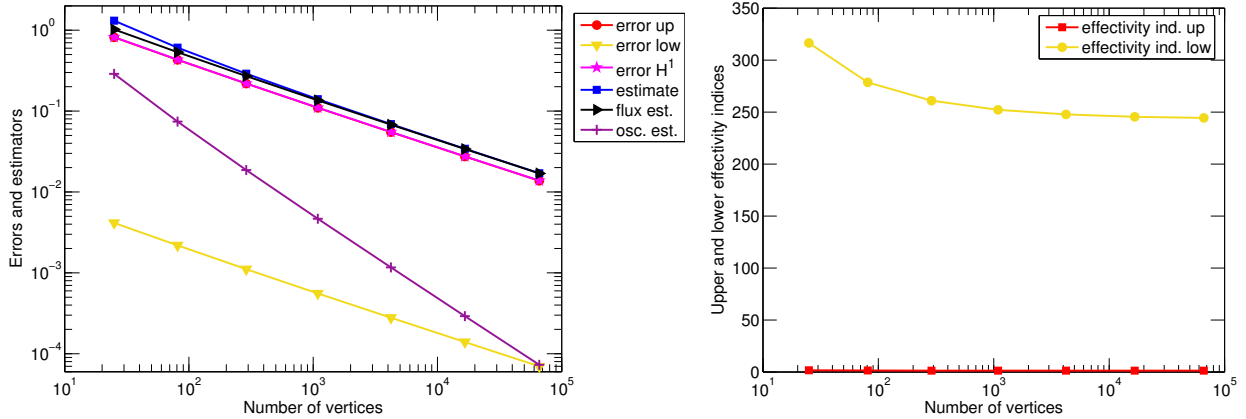


Figure 5: Estimates and errors for uniform mesh refinement (left) and the corresponding effectivity indices (right), the regular case with $\sigma_- = -0.99$

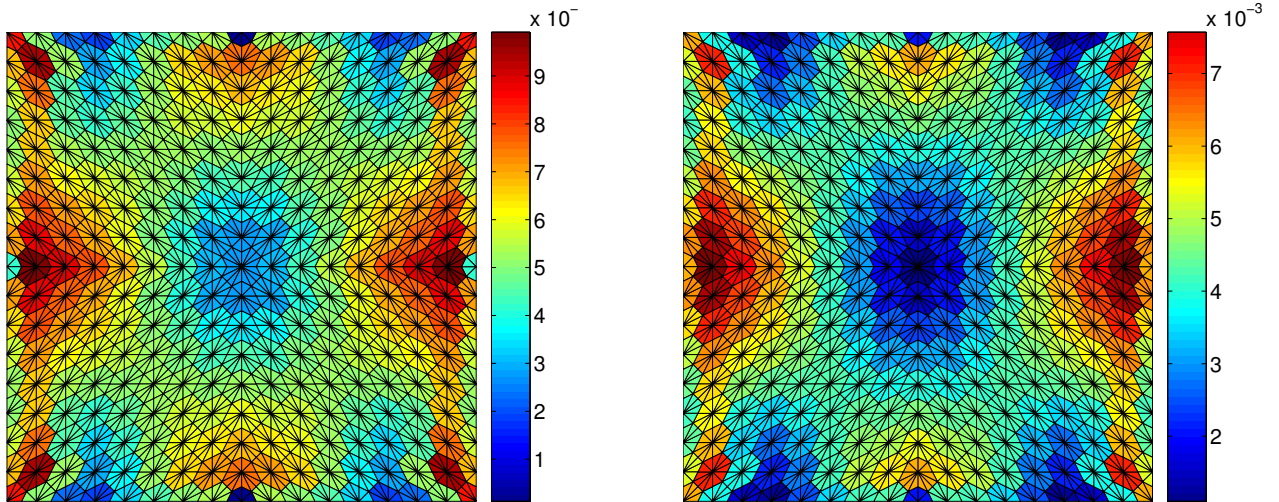


Figure 6: Estimated (left) and exact (right) error distribution, the regular case with $\sigma_- = -1/3$

according to Bonnet-BenDhia *et al.* [3] given by

$$\begin{aligned}
 u(x, y) &= r^\lambda (c_1 \sin(\lambda\theta) + c_2 \sin(\lambda(\pi/2 - \theta))) \text{ for } (x, y) \in \Omega_+, \\
 u(x, y) &= r^\lambda (d_1 \sin(\lambda(\theta - \pi/2)) + d_2 \sin(\lambda(2\pi - \theta))) \text{ for } (x, y) \in \Omega_-.
 \end{aligned}$$

Here (r, θ) are the polar coordinates and $\lambda = 2/\pi \arccos((1 - \sigma_-)/(2|1 + \sigma_-|))$. We consider two test settings with $\sigma_- = -5$ and $\sigma_- = -3.1$ leading respectively to $c_1 = 1, c_2 = -1, d_1 = -0.8, d_2 = -0.8, \lambda \approx 0.4601069123$ and $c_1 = 1, c_2 = -1, d_1 \approx -0.3556451613, d_2 \approx 0.3556451613, \lambda \approx 0.1391989493$. Classically, $u \in H^{1+\lambda}(\Omega)$ only, with a singularity at the origin. These exact solutions together with their finite element approximations on coarse adaptively refined meshes (the same initial mesh as in the previous section was used) are presented in Figure 7. The inhomogeneous Dirichlet boundary condition is prescribed accordingly, whereas the source term f is set to 0. The operator T is specified in [23]; the problem is particular well-posed when $\sigma_- < -3$ or $-1/3 < \sigma_- < 0$.

The various errors, the estimators, and the corresponding effectivity indices are presented in Figures 8–10. The convergence rates for uniform mesh refinement correspond to the a priori analysis, being 0.46 and

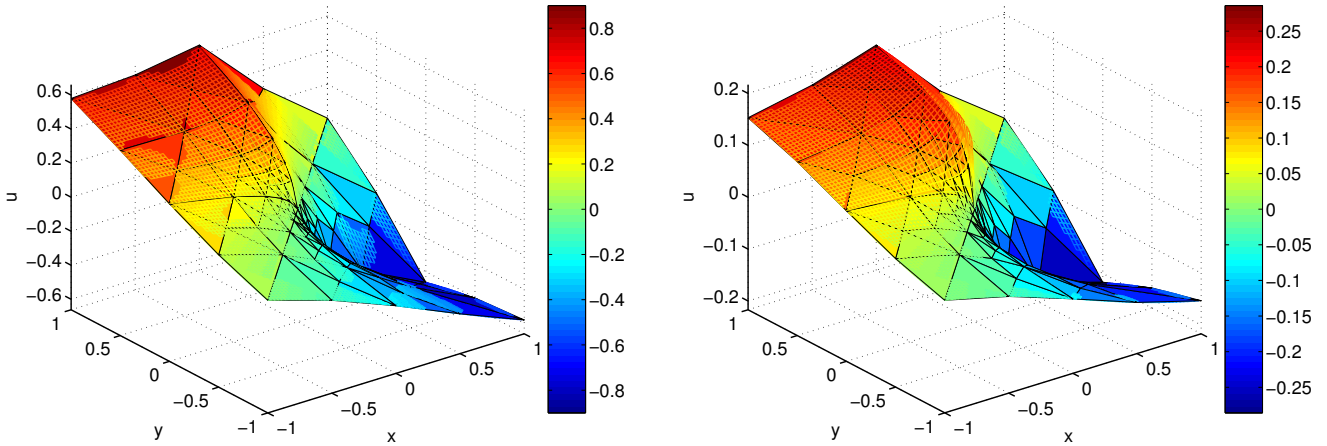


Figure 7: Exact and approximate solution, the singular case with $\sigma_- = -5$ (left) and $\sigma_- = -3.1$ (right)

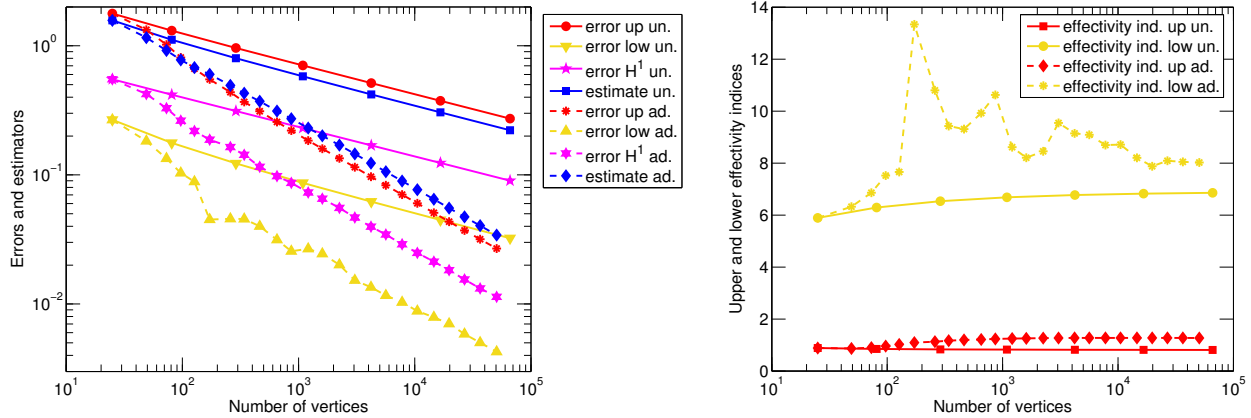


Figure 8: Estimates and errors for uniform and adaptive mesh refinement (left) and the corresponding effectivity indices (right), the singular case with $\sigma_- = -5$

0.12 respectively in the two settings for both $\|\underline{\Sigma}\nabla(u - u_h)\|$ and $\|\nabla(u - u_h)\|$. The convergence rate for $\frac{(\underline{\Sigma}\nabla(u - u_h), \nabla(\mathbb{T}(u - u_h)))}{\|\nabla(\mathbb{T}(u - u_h))\|}$ in the less singular setting is still 0.47, but for $\sigma_- = -3.1$, we have obtained 0.22. The estimators for uniform mesh refinement converge at order 0.46 and 0.20, respectively. Adaptive mesh refinement seems to optimize the situation, after a preliminary phase for the strongly singular setting. The convergence orders are respectively 1.01 and 0.97 for both $\|\underline{\Sigma}\nabla(u - u_h)\|$ and $\|\nabla(u - u_h)\|$, 1.00 and 1.39 for $\frac{(\underline{\Sigma}\nabla(u - u_h), \nabla(\mathbb{T}(u - u_h)))}{\|\nabla(\mathbb{T}(u - u_h))\|}$, and 1.01 and 0.98 for the estimators. Here, more markedly, the gaps in (2.1), namely between $\frac{(\underline{\Sigma}\nabla(u - u_h), \nabla(\mathbb{T}(u - u_h)))}{\|\nabla(\mathbb{T}(u - u_h))\|}$ and $\|u - u_h\|_{\#}$ (and consequently $\|\underline{\Sigma}\nabla(u - u_h)\|$), appear. For this reason, the upper and lower effectivity indices present a wider gap. The predicted spatial distribution of the error, though, still seems to be very accurate even in the close-to-the-limit singular case $\sigma_- = -3.1$, see Figure 11.

7 Conclusions and outlook

We have presented a posteriori error estimates for transmission problems with sign-changing coefficients. In the setting of the intrinsic dual norm (2.3) and equilibrated fluxes, we obtain a guaranteed upper bound, as well as a lower bound up to a generic constant independent of the jump in the diffusion coefficient.

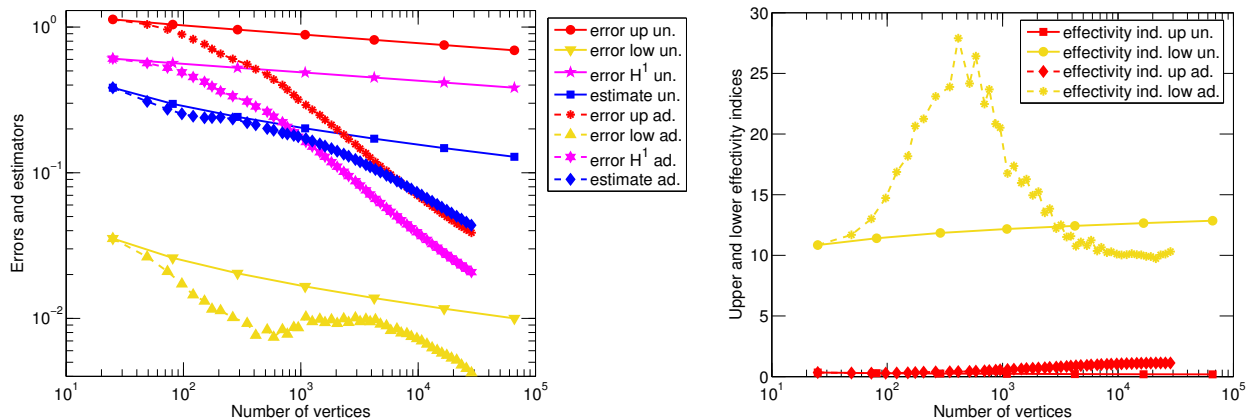


Figure 9: Estimates and errors for uniform and adaptive mesh refinement (left) and the corresponding effectivity indices (right), the singular case with $\sigma_- = -3.1$

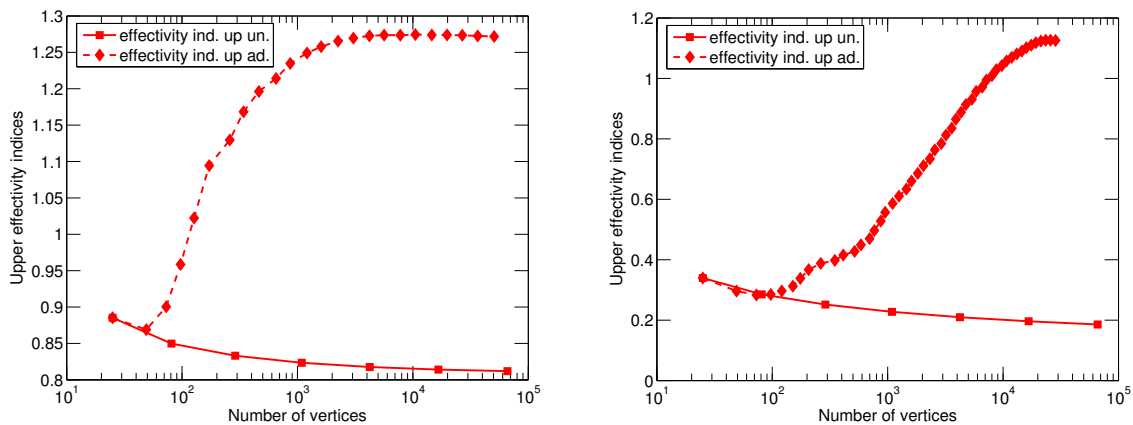


Figure 10: Upper effectivity indices, the singular case with $\sigma_- = -5$ (left) and $\sigma_- = -3.1$ (right)

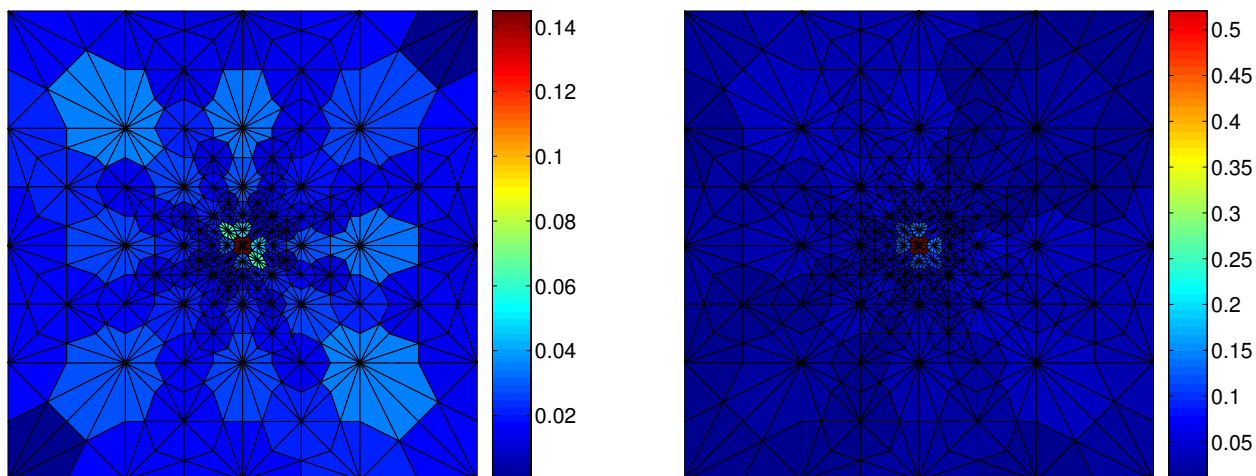


Figure 11: Estimated (left) and exact (right) error distribution, the singular case with $\sigma_- = -3.1$

This robustness result is moreover obtained in a unified framework covering basically all classical numerical methods. Numerical experiments, in the conforming finite element setting, confirm these results, but also show that if one rather uses the easily computable upper and lower bounds for the intrinsic dual norm, the robustness does not carry over. Possible future developments include the control of the error from a not completely converged linear solver (and corresponding stopping criteria), an extension to nonlinear problems, or a proposition of an adaptive operator T for self-adapting the method.

References

- [1] BERNARDI, C., AND VERFÜRTH, R. Adaptive finite element methods for elliptic equations with non-smooth coefficients. *Numer. Math.* 85, 4 (2000), 579–608.
- [2] BONNET-BEN DHIA, A.-S., CHESNEL, L., AND CIARLET JR., P. T-coercivity for scalar interface problems between dielectrics and metamaterials. *M2AN Math. Model. Numer. Anal.* 46 (2012), 1363–1387.
- [3] BONNET-BENDHIA, A.-S., DAUGE, M., AND RAMDANI, K. Analyse spectrale et singularités d’un problème de transmission non coercif. *C. R. Acad. Sci. Paris Sér. I Math.* 328, 8 (1999), 717–720.
- [4] BRAESS, D., PILLWEIN, V., AND SCHÖBERL, J. Equilibrated residual error estimates are p -robust. *Comput. Methods Appl. Mech. Engrg.* 198, 13-14 (2009), 1189–1197.
- [5] BRAESS, D., AND SCHÖBERL, J. Equilibrated residual error estimator for edge elements. *Math. Comp.* 77, 262 (2008), 651–672.
- [6] BREZZI, F., AND FORTIN, M. *Mixed and hybrid finite element methods*, vol. 15 of *Springer Series in Computational Mathematics*. Springer-Verlag, New York, 1991.
- [7] CARSTENSEN, C., AND FUNKEN, S. A. Fully reliable localized error control in the FEM. *SIAM J. Sci. Comput.* 21, 4 (1999/00), 1465–1484.
- [8] CARSTENSEN, C., AND MERDON, C. Estimator competition for Poisson problems. *J. Comput. Math.* 28, 3 (2010), 309–330.
- [9] CHAILLOU, A. L., AND SURI, M. A posteriori estimation of the linearization error for strongly monotone nonlinear operators. *J. Comput. Appl. Math.* 205, 1 (2007), 72–87.
- [10] CHESNEL, L., AND CIARLET, JR., P. T-coercivity and continuous Galerkin methods: application to transmission problems with sign changing coefficients. *Numer. Math.* 124, 1 (2013), 1–29.
- [11] CHUNG, E. T., AND CIARLET, JR., P. A staggered discontinuous Galerkin method for wave propagation in media with dielectrics and meta-materials. *J. Comput. Appl. Math.* 239 (2013), 189–207.
- [12] CIARLET, P. G. *The Finite Element Method for Elliptic Problems*, vol. 4 of *Studies in Mathematics and its Applications*. North-Holland, Amsterdam, 1978.
- [13] DESTUYNDER, P., AND MÉTIVET, B. Explicit error bounds in a conforming finite element method. *Math. Comp.* 68, 228 (1999), 1379–1396.
- [14] EL ALAOU, L., ERN, A., AND VOHRALÍK, M. Guaranteed and robust a posteriori error estimates and balancing discretization and linearization errors for monotone nonlinear problems. *Comput. Methods Appl. Mech. Engrg.* 200, 37-40 (2011), 2782–2795.
- [15] ERN, A., AND GUERMOND, J.-L. *Theory and practice of finite elements*, vol. 159 of *Applied Mathematical Sciences*. Springer-Verlag, New York, 2004.
- [16] ERN, A., AND VOHRALÍK, M. Polynomial-degree-robust a posteriori estimates in a unified setting for conforming, nonconforming, discontinuous Galerkin, and mixed discretizations. *SIAM J. Numer. Anal.* 53, 2 (2015), 1058–1081.

- [17] HANNUKAINEN, A., STENBERG, R., AND VOHRALÍK, M. A unified framework for a posteriori error estimation for the Stokes problem. *Numer. Math.* 122, 4 (2012), 725–769.
- [18] KELLY, D. W. The self-equilibration of residuals and complementary a posteriori error estimates in the finite element method. *Internat. J. Numer. Methods Engrg.* 20, 8 (1984), 1491–1506.
- [19] LADEVÈZE, P., AND LEGUILLON, D. Error estimate procedure in the finite element method and applications. *SIAM J. Numer. Anal.* 20, 3 (1983), 485–509.
- [20] LUCE, R., AND WOHLMUTH, B. I. A local a posteriori error estimator based on equilibrated fluxes. *SIAM J. Numer. Anal.* 42, 4 (2004), 1394–1414.
- [21] MATLAB. *Version 7.0.4.365 (R14)*. The MathWorks Inc., Natick, Massachusetts, 2005.
- [22] MELENK, J. M., AND WOHLMUTH, B. I. On residual-based a posteriori error estimation in *hp*-FEM. *Adv. Comput. Math.* 15, 1-4 (2001), 311–331 (2002). A posteriori error estimation and adaptive computational methods.
- [23] NICAISE, S., AND VENEL, J. A posteriori error estimates for a finite element approximation of transmission problems with sign changing coefficients. *J. Comput. Appl. Math.* 235 (2011), 4272–4282.
- [24] PRAGER, W., AND SYNGE, J. L. Approximations in elasticity based on the concept of function space. *Quart. Appl. Math.* 5 (1947), 241–269.
- [25] REPIN, S. I. *A posteriori estimates for partial differential equations*, vol. 4 of *Radon Series on Computational and Applied Mathematics*. Walter de Gruyter GmbH & Co. KG, Berlin, 2008.
- [26] ROBERTS, J. E., AND THOMAS, J.-M. Mixed and hybrid methods. In *Handbook of Numerical Analysis, Vol. II*. North-Holland, Amsterdam, 1991, pp. 523–639.
- [27] VERFÜRTH, R. Robust a posteriori error estimates for stationary convection-diffusion equations. *SIAM J. Numer. Anal.* 43, 4 (2005), 1766–1782.
- [28] VERFÜRTH, R. *A posteriori error estimation techniques for finite element methods*. Numerical Mathematics and Scientific Computation. Oxford University Press, Oxford, 2013.
- [29] VOHRALÍK, M. Guaranteed and fully robust a posteriori error estimates for conforming discretizations of diffusion problems with discontinuous coefficients. *J. Sci. Comput.* 46, 3 (2011), 397–438.
- [30] WALLEN, H., KETTUNEN, H., AND SIHVOLA, A. Surface modes of negative-parameter interfaces and the importance of rounding sharp corners. *Metamaterials* 2 (2008), 113–121.
- [31] ZIENKIEWICZ, O. C., AND TAYLOR, R. L. *The finite element method. The basis*, fifth ed. Butterworth-Heinemann, Oxford, 2000.