Localization of global norms and robust a posteriori error control for transmission problems with sign-changing coefficients

Patrick Ciarlet, Martin Vohralík

Localization of global norms and robust a posteriori error control for transmission problems with sign-changing coefficients*

Patrick Ciarlet Jr.† Martin Vohralík‡

August 22, 2017

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, covering both conforming and nonconforming approximations. It combines a dual (residual) norm together with the distance to the correct functional space. Importantly, we show the equivalence of both these quantities defined globally over the entire computational domain with the Hilbertian sums of their localizations over patches of elements. In this framework, we then design a posteriori estimators which deliver simultaneously guaranteed error upper bound, global and local error lower bounds, and robustness with respect to the (sign-changing) diffusion tensor. Robustness with respect to the approximation polynomial degree is achieved as well. The estimators are given in a unified setting covering at once conforming, nonconforming, mixed, and discontinuous Galerkin finite element discretizations in two or three space dimensions. Numerical results illustrate the theoretical developments.

Key words: noncoercive problem, sign change, metamaterial, a posteriori error estimate, dual norm, distance to energy space, localization, equivalence local–global, minimization, best approximation, equilibrated flux, unified framework, robustness, finite element methods

1 Introduction

Let $\Omega \subset \mathbb{R}^d$, $1 \leq d \leq 3$, be an open polytope (polygon for $d = 2$, polyhedron for $d = 3$) with a Lipschitz-continuous boundary $\partial\Omega$, $\underline{\boldsymbol{\Sigma}}$ a tensor-valued diffusion tensor, and f a datum. We consider the following problem: find $u : \Omega \to \mathbb{R}$ such that

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

$$u = 0 \qquad \text{on } \partial\Omega. \tag{1.1b}$$

In contrast to the usual setting, cf. Ciarlet [23], we relax the assumption of $\underline{\boldsymbol{\Sigma}}$ being positive definite (and symmetric). Such a situation arises as a model problem in electromagnetism for interfaces between dielectrics and (negative) metamaterials or metals, see, e.g., Bonnet-BenDhia *et al.* [10] or Wallen *et al.* [53] 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 $\underline{\boldsymbol{\Sigma}}|_{\Omega_+} = \sigma_+\underline{\boldsymbol{I}}$ and $\underline{\boldsymbol{\Sigma}}|_{\Omega_-} = \sigma_-\underline{\boldsymbol{I}}$, where $\sigma_+ > 0$ and $\sigma_- < 0$ are two scalars and $\underline{\boldsymbol{I}}$ is the identity tensor.

We will call $u \in H^1_0(\Omega)$ a weak solution of (1.1) if

$$(\underline{\boldsymbol{\Sigma}}\nabla u, \nabla v) = (f, v) \qquad \forall v \in H^1_0(\Omega). \tag{1.2}$$

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

‡INRIA Paris, 2 rue Simone Iff, 75589 Paris, France & Université Paris-Est, CERMICS (ENPC), 77455 Marne-la-Vallée, France (martin.vohralik@inria.fr).

*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). It was also supported by the ERC-CZ project MORE "MOdelling REvisited + MOdel REduction", LL1202.

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 (also called Brezzi–Babuška or inf–sup) theorem, cf., e.g., Ern and Guermond [30, Theorem 2.6]. They have recently been revisited via the T-coercivity approach, see, e.g., Bonnet-BenDhia *et al.* [9] or Chesnel and Ciarlet [20] and the references therein. Conception of numerical approximations, their well-posedness, and a priori error estimates have been addressed in [20] in the conforming finite element context and in [21] in the nonconforming finite element and discontinuous Galerkin context.

A posteriori error analysis for problems of type (1.1) has likewise been started recently. In particular, Nicaise and Venel [39] bound the error between the known finite element approximation u_h and the unknown weak solution u given by (1.2) by a computable a posteriori indicator. The bound, however, features an unknown generic constant. The dependence of the quality of the estimator on the tensor $\underline{\boldsymbol{\Sigma}}$ (on the ratio, or contrast, σ_+/σ_- in the simplest setting) is, unfortunately, not traced; numerical experiments indicate deterioration of the behavior (so-called non-robustness) when the contrast is approaching the set of forbidden values given by an interval to which the value -1 always belongs. In [39], 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 on diffusion problems with jumping coefficients, see Bernardi and Verfürth [6], Ainsworth [1], or [52] and the references therein, only study the standard positive definite case.

In terms of a posteriori analysis, there are four goals of this contribution: firstly, we want to derive a posteriori error estimates which are *guaranteed*, certifying the maximal error and featuring no unknown constant. Secondly, we wish them to be *robust* with respect to the *jumps* and *sign changes* in the tensor $\underline{\boldsymbol{\Sigma}}$. The adaptive mesh refinement based on the a posteriori error estimators developed in this work produces in particular in our numerical experiments sequences of meshes leading to optimal decay rates for an arbitrarily singular solution. Thirdly, we want to develop a *unified framework* covering all standard numerical methods. We achieve this via the concept of *flux* and *potential reconstructions*, following Prager and Synge [42], Ladevèze and Leguillon [36], Kelly [34], Destuynder and Métivet [26], Luce and Wohlmuth [37], and Braess and Schöberl [12] for equilibrated fluxes, Prager and Synge [42], Destuynder and Métivet [25], Ainsworth [1], or Carstensen and Merdon [18] for the potentials, and the unifying frameworks in Nicaise *et al.* [40], Repin [43], Ainsworth [2], Carstensen *et al.* [16], Becker *et al.* [5], or [52, 31, 32], see also the references therein. Fourthly and lastly, in extension of Braess *et al.* [11] for conforming finite elements and of [32, 33] for nonconforming, mixed, and discontinuous Galerkin finite elements, we obtain *robustness* with respect to the approximation *polynomial degree*.

The key point for obtaining the above-discussed properties is a proper choice of the way the error is measured. For conforming (lying in the space $H^1_0(\Omega)$) approximations, Verfürth [49], Chaillou and Suri [19], Veeser and Verfürth [46], Kreuzer and Süli [35], and [52, 29, 31] used the *intrinsic problem-dependent norm* given by the dual norm of the residual stemming from the weak formulation. We articulate here two goals. We first identify a proper generalization of this concept to our setting, including nonconforming approximations $u_h \notin H^1_0(\Omega)$. The norm in which we measure the error is in particular given by

$$|||v|||^2 := \max_{\varphi\in H^1_0(\Omega);\,\|\nabla\varphi\|=1} (\underline{\boldsymbol{\Sigma}}\nabla_\theta v, \nabla\varphi)^2 + \min_{\zeta\in H^1_0(\Omega)} \|\nabla_\theta(v-\zeta)\|^2 + \sum_{e\in\mathcal{E}_h} h_e^{-1}\|\Pi^0_e[\![v]\!]\|^2_e. \tag{1.3}$$

Here v lies in $H^1(\mathcal{T}_h)$, the broken Sobolev space, see (2.4) below, and ∇_θ is the discrete gradient defined below by (2.6). For $v = u - u_h$, the first term above is the *dual norm of the residual*, the second one is the *distance to the energy space* in a gradient seminorm, and the last one evaluates the size of the *mean values of jumps* in the approximate solution u_h. Secondly, we prove that $|||\cdot|||$ as well as both its components $\|\cdot\|_*$ (first term in (1.3)) and $\|\cdot\|_\#$ (last two terms in (1.3)) are *equivalent* to the Hilbertian sums of their *localizations* on patches of elements. These results seem to be of independent interest, stating a local–global equivalence for norms that are only global at a first sight. For dual (residual) norms, a result of this type has probably first been shown in Babuška and Miller [4, Theorem 2.1.1], and may be deduced from more recent a posteriori analyses, see in particular Carstensen and Funken [17], Morin *et al.* [38], Verfürth [48, 50], Veeser and Verfürth [46], Cohen *et al.* [24], and the references therein, typically for piecewise polynomial approximations. It has recently been extended in [8] to any bounded linear functional on the Sobolev space $W^{1,p}_0(\Omega)$, $p > 1$. Galerkin orthogonality with respect to lowest-order modes turns out to be crucial here for one direction of the equivalence. For the distance to the energy space, our localization result seems to be new, although a clue can be again found in a posteriori error estimates for nonconforming finite element

methods on piecewise polynomial spaces, see, e.g., [16, Theorem 5.1], the survey [32], and the references therein. We also cite Veeser [45] who recently proved that local and global best-approximation errors in the energy norm are equivalent for piecewise polynomial spaces. Here, we derive the localization results on the entire broken Sobolev space $H^1(\mathcal{T}_h)$ and give direct and minimal proofs with clearly identified constants that only depend on mesh shape regularity and on the space dimension. This in particular gives robustness with respect to the tensor $\underline{\boldsymbol{\Sigma}}$ and does not request one to work with piecewise polynomial spaces. Computable upper bounds on the generic constants are also indicated. In a posteriori error analysis, these results allow to pass from merely global to actually local efficiency, namely in [49, 46, 52, 29, 31] and in the references therein.

Our paper is organized as follows. Section 2 sets the notation and assumptions and identifies and examines the intrinsic norm $|||\cdot|||$. Section 3 resumes our general findings about the localization of global norms. A posteriori estimates in an abstract framework for all standard numerical approximations of problem (1.1) then form the content of Section 4. Finally, Section 5 illustrates our theoretical developments on two numerical examples, whereas Section 6 gives some concluding remarks and outlook.

2 Setting

This section introduces the notation, assumptions, and discusses in detail the choice of the way we will measure the error in numerical approximations of problem (1.1).

2.1 Notation

Let $\{\mathcal{T}_h\}_h$ be a family of simplicial partitions of the domain Ω, i.e., $\cup_{K\in\mathcal{T}_h} K = \overline{\Omega}$ for all $\mathcal{T}_h$, any element $K \in \mathcal{T}_h$ for any mesh $\mathcal{T}_h$ is a closed simplex, and the intersection of two different simplices in one mesh $\mathcal{T}_h$ is either empty, a vertex, or their common l-dimensional face, $1 \le l \le d-1$. The set of vertices will be denoted by $\mathcal{V}_h$; it is composed of interior vertices $\mathcal{V}_h^{\rm int}$ and vertices located on the boundary $\mathcal{V}_h^{\rm ext}$. For element $K \in \mathcal{T}_h$, $\mathcal{V}_K$ denotes the set of its vertices. 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 $\mathbf{a}$, for $\omega_\mathbf{a}$ the corresponding open subdomain, and $\psi_\mathbf{a}$ for the continuous, piecewise affine "hat" function which takes value 1 at the vertex $\mathbf{a}$ and zero at the other vertices.

The mesh $(d-1)$-dimensional faces are collected in the set $\mathcal{E}_h$, with interior faces $\mathcal{E}_h^{\rm int}$ and boundary faces $\mathcal{E}_h^{\rm ext}$. 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 $e \in \mathcal{E}_h^{\rm int}$ and points outwards of Ω for $e \in \mathcal{E}_h^{\rm ext}$. We will use the jump operator $[\![\cdot]\!]$ yielding the difference evaluated along $\mathbf{n}_e$ of the traces of the argument from the two mesh elements that share $e \in \mathcal{E}_h^{\rm int}$ and the actual trace for $e \in \mathcal{E}_h^{\rm ext}$. Similarly, $\{\!\{\cdot\}\!\}$ stands for the mean value of the traces from adjacent mesh elements on faces from $\mathcal{E}_h^{\rm int}$ and the actual trace on $\mathcal{E}_h^{\rm ext}$. We denote by Π_e^0 the $L^2(e)$-orthogonal projection onto constants (mean value) on a face $e \in \mathcal{E}_h$.

For a d-dimensional subdomain ω of Ω, we use $(\cdot,\cdot)_\omega$ to denote the $L^2(\omega)$ or $[L^2(\omega)]^d$ scalar product and $\|\cdot\|_\omega$ for the associated norm; shall $\omega = \Omega$, the subscript is dropped. For $(d-1)$-dimensional subdomains, we similarly use $\langle\cdot,\cdot\rangle_\omega$ and $\|\cdot\|_\omega$.

2.2 Assumptions

Throughout the paper, we shall suppose the following:

Assumption 2.1 (Setting)**.** *We suppose that*

- *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 $\mathcal{T}_h$, $\max_{K\in\mathcal{T}_h} h_K/\varrho_K \le \kappa_\mathcal{T}$, where h_K is the diameter of K and ϱ_K is the diameter of the largest ball inscribed in K;*
- *$\underline{\boldsymbol{\Sigma}} \in [L^\infty(\Omega)]^{d\times d}$ is piecewise constant on each given $\mathcal{T}_h$;*
- *$f \in L^2(\Omega)$;*
- *there exists a linear bijective operator $\mathrm{T} : H^1_0(\Omega) \to H^1_0(\Omega)$, cf. [20, Definition 3], bounded in the sense that $\|\nabla(\mathrm{T}v)\| \le \|\mathrm{T}\|\|\nabla v\|$ for all $v \in H^1_0(\Omega)$, $\|\mathrm{T}\| < \infty$, and such that the bilinear form in (1.2) is T-coercive in the sense that $(\underline{\boldsymbol{\Sigma}}\nabla v, \nabla(\mathrm{T}v)) \ge \underline{\alpha}\|\nabla v\|^2$ for all $v \in H^1_0(\Omega)$, $\underline{\alpha} > 0$.*

Under Assumption 2.1, one immediately obtains:

Corollary 2.2 (Weak solution)**.** *There exists a well-posed solution of problem* (1.1) *in the sense* (1.2)*. It satisfies* $u \in H^1_0(\Omega)$ *and* $\boldsymbol{\sigma} := -\underline{\boldsymbol{\Sigma}}\nabla u \in \mathbf{H}(\mathrm{div},\Omega)$ *with* $\nabla\cdot\boldsymbol{\sigma} = f$*.*

2.3 Intrinsic norm in the conforming setting

Let, for the moment, $v \in H^1_0(\Omega)$ and $\nabla_\theta = \nabla$. The weak formulation (1.2) and Assumption 2.1 suggest the intrinsic norm

$$\|v\|_* := \max_{\varphi\in H^1_0(\Omega);\,\|\nabla\varphi\|=1} (\underline{\boldsymbol{\Sigma}}\nabla_\theta v, \nabla\varphi); \tag{2.1a}$$

this writing takes immediately the form we need in this paper, for v from the broken Sobolev space $H^1(\mathcal{T}_h)$ and the discrete gradient ∇_θ defined below. We define the local versions of (2.1a), for each vertex $\mathbf{a} \in \mathcal{V}_h$ and the corresponding patch subdomain $\omega_\mathbf{a}$, as

$$\|v\|_{*,\omega_\mathbf{a}} := \max_{\varphi\in H^1_0(\omega_\mathbf{a});\,\|\nabla\varphi\|_{\omega_\mathbf{a}}=1} (\underline{\boldsymbol{\Sigma}}\nabla_\theta v, \nabla\varphi)_{\omega_\mathbf{a}}. \tag{2.1b}$$

For $v \in H^1_0(\Omega)$, [–] the Cauchy–Schwarz inequality implies

$$\frac{(\underline{\boldsymbol{\Sigma}}\nabla v, \nabla(\mathrm{T}v))}{\|\nabla(\mathrm{T}v)\|} \le \|v\|_* \le \|\underline{\boldsymbol{\Sigma}}\nabla v\| \qquad \forall v \in H^1_0(\Omega), \tag{2.2}$$

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

$$\frac{\alpha}{\|\mathrm{T}\|}\|\nabla v\| \le \|v\|_* \le \|\underline{\boldsymbol{\Sigma}}\|_\infty\|\nabla v\| \qquad \forall v \in H^1_0(\Omega), \tag{2.3}$$

so that $\|v\|_*$ is indeed a norm on the space $H^1_0(\Omega)$, equivalent to the canonical norm $\|\nabla v\|$. Note, however, that the equivalence constants $\frac{\alpha}{\|\mathrm{T}\|}$ and $\|\underline{\boldsymbol{\Sigma}}\|_\infty$ are setting- and problem-dependent (not robust), see Remark 5.1 below for a discussion of a particular example. Remark also that $(\underline{\boldsymbol{\Sigma}}\nabla v, \nabla v)$ may become negative, which excludes the notion itself of an energy norm; on the other hand $\|v\|_* = \|\nabla v\|$ in the case where $\underline{\boldsymbol{\Sigma}} = \underline{\boldsymbol{I}}$, so that $\|v\|_*$ is a natural extension of the canonical norm of the Laplace operator.

2.4 Broken Sobolev space and broken and discrete gradients

In order to make our analysis as general as possible, we will henceforth often work with the broken Sobolev space $H^1(\mathcal{T}_h)$ related to the mesh $\mathcal{T}_h$,

$$H^1(\mathcal{T}_h) := \{v \in L^2(\Omega);\, v|_K \in H^1(K) \qquad \forall K \in \mathcal{T}_h\}. \tag{2.4}$$

The corresponding *broken gradient* ∇_h is given by, for $v \in H^1(\mathcal{T}_h)$,

$$(\nabla_h v)|_K = \nabla(v|_K) \qquad \forall K \in \mathcal{T}_h. \tag{2.5}$$

In order to, in particular, take into account in our analysis discontinuous Galerkin methods, we are lead to further generalize the concept of the broken gradient following, e.g., Di Pietro and Ern [27, Section 4.3] and the references therein. For each face $e \in \mathcal{E}_h$, let $\mathcal{T}_e$ regroup the (one or two) mesh elements sharing the face e. We let $\mathbf{V}^0(\mathcal{T}_e)$ stand for piecewise constant vectors on $\mathcal{T}_e$, i.e., $\mathbf{v}_h|_K \in [\mathbb{P}_0(K)]^d$ for all $K \in \mathcal{T}_e$. Alternatively, vectors $\mathbf{v}_h$ such that $\mathbf{v}_h|_K \in [\mathbb{P}_0(K)]^d + \mathbb{P}_0(K)\mathbf{x}$ for all $K \in \mathcal{T}_e$ (piecewise lowest-order Raviart–Thomas–Nédélec space) could also be used. In both cases, $\mathbf{v}_h\cdot\mathbf{n}_e$ is constant for $\mathbf{v}_h \in \mathbf{V}^0(\mathcal{T}_e)$. Let $v \in H^1(\mathcal{T}_h)$. We define the lifting operator $\mathfrak{l}_e : L^2(e) \to \mathbf{V}^0(\mathcal{T}_e)$ by

$$(\mathfrak{l}_e(\llbracket v\rrbracket), \mathbf{v}_h)_{\mathcal{T}_e} = \langle \{\!\{\mathbf{v}_h\}\!\}\cdot\mathbf{n}_e, \llbracket v\rrbracket\rangle_e \qquad \forall \mathbf{v}_h \in \mathbf{V}^0(\mathcal{T}_e).$$

We then extend $\mathfrak{l}_e(\llbracket v\rrbracket)$ by zero outside of $\mathcal{T}_e$. For a parameter $\theta \in \{-1, 0, 1\}$, the *discrete gradient* $\nabla_\theta v \in [L^2(\Omega)]^d$ is given by

$$\nabla_\theta v := \nabla_h v - \theta \sum_{e\in\mathcal{E}_h} \mathfrak{l}_e(\llbracket v\rrbracket). \tag{2.6}$$

We observe that $\nabla_\theta v = \nabla_h v$ when $\theta = 0$ or when the jumps of v are of mean value 0, i.e., $\langle [\![v]\!], 1\rangle_e = 0$ for all $e \in \mathcal{E}_h$. Similarly, both broken and discrete gradients are consistent extensions of the weak gradient ∇ in the sense that

$$\nabla_\theta v = \nabla_h v = \nabla v \qquad \forall v \in H^1_0(\Omega). \tag{2.7}$$

2.5 Nonconformity evaluation

An important observation is that $\|\cdot\|_*$ given by (2.1a) is merely a seminorm on the broken Sobolev space $H^1(\mathcal{T}_h)$. Consequently, it is not sufficient to evaluate the error therein, and we are lead to quantify the nonconformity $H^1(\mathcal{T}_h) \not\subset H^1_0(\Omega)$. An intrinsic measure here is simply the distance to the energy space $H^1_0(\Omega)$, $\min_{\zeta\in H^1_0(\Omega)}\|\nabla_\theta(v-\zeta)\|$ for $v \in H^1(\mathcal{T}_h)$. As in this expression, only the gradient seminorm appears, we are finally lead to evaluate the nonconformity as

$$\|v\|^2_\# := \min_{\zeta\in H^1_0(\Omega)} \|\nabla_\theta(v-\zeta)\|^2 + \sum_{e\in\mathcal{E}_h} h_e^{-1}\|\Pi^0_e [\![v]\!]\|^2_e \qquad v \in H^1(\mathcal{T}_h). \tag{2.8a}$$

The second term with the mean values of the jumps on the faces given by $\Pi^0_e[\![v]\!]$ ensures the validity of the broken Poincaré–Friedrichs inequality and plays a key role in Lemma 2.3 below. Note also that scaling both or one term in (2.8a) by generic constants is possible. As local versions of (2.8a), we define

$$\|v\|^2_{\#,\omega_{\mathbf{a}}} := \min_{\zeta\in H^1_\#(\omega_{\mathbf{a}})} \|\nabla_\theta(v-\zeta)\|^2_{\omega_{\mathbf{a}}} + \sum_{e\in\mathcal{E}_h,\, \mathbf{a}\in e} h_e^{-1}\|\Pi^0_e [\![v]\!]\|^2_e \tag{2.8b}$$

for each vertex $\mathbf{a} \in \mathcal{V}_h$ and the corresponding patch subdomain $\omega_{\mathbf{a}}$. Here

$$\begin{aligned} H^1_\#(\omega_{\mathbf{a}}) &:= H^1(\omega_{\mathbf{a}}), && \mathbf{a} \in \mathcal{V}_h^{\mathrm{int}}, \\ H^1_\#(\omega_{\mathbf{a}}) &:= \{v \in H^1(\omega_{\mathbf{a}});\, v = 0 \text{ on } \partial\omega_{\mathbf{a}} \cap \partial\Omega\}, && \mathbf{a} \in \mathcal{V}_h^{\mathrm{ext}}. \end{aligned} \tag{2.9}$$

2.6 Intrinsic norm

Combining (2.1a) and (2.8a), we define the *total intrinsic norm* as

$$|||v|||^2 = \|v\|^2_* + \|v\|^2_\# \qquad v \in H^1(\mathcal{T}_h). \tag{2.10}$$

We have the following simple but crucial result:

Lemma 2.3 (Intrinsic norm)**.** *Let the broken Sobolev space $H^1(\mathcal{T}_h)$ be given by (2.4) and the discrete gradient ∇_θ by (2.6) with $\theta \in \{-1, 0, 1\}$. Then $|||\cdot|||$ given by (2.10) defines a norm on $H^1(\mathcal{T}_h)$.*

Proof. Clearly, $|||\alpha v||| = |\alpha|\,|||v|||$ and $|||v + w||| \le |||v||| + |||w|||$ for any $\alpha \in \mathbb{R}$ and any $v, w \in H^1(\mathcal{T}_h)$. Let now $|||v||| = 0$. Then the second term in (2.8a) implies that the jumps of v are of mean value 0, $\langle [\![v]\!], 1\rangle_e = 0$ for all $e \in \mathcal{E}_h$, and thus $\nabla_\theta = \nabla_h$. Consequently, for $s := \arg\min_{\zeta\in H^1_0(\Omega)}\|\nabla_h(v-\zeta)\|$, the broken Poincaré–Friedrichs inequality

$$\|v - s\| \le C_{\mathrm{bPF},\Omega} h_\Omega \|\nabla_h(v - s)\|,$$

see Brenner [13] or [51], implies from the fact that the first term in (2.8a) vanishes that $v = s$ and thus $v \in H^1_0(\Omega)$. Finally, the equivalence (2.3) valid on the energy space $H^1_0(\Omega)$ shows that indeed $v = 0$. □

2.7 Evaluating the error by the dual norm of the residual and the distance to the energy space

When $\underline{\boldsymbol{\Sigma}} = \underline{\boldsymbol{I}}$, there holds, for arbitrary $u \in H^1_0(\Omega)$ and $u_h \in H^1(\mathcal{T}_h)$,

$$\|\nabla_\theta(u - u_h)\|^2 = \max_{\varphi\in H^1_0(\Omega);\, \|\nabla\varphi\|=1} (\nabla_\theta(u - u_h), \nabla\varphi)^2 + \min_{\zeta\in H^1_0(\Omega)} \|\nabla_\theta(u_h - \zeta)\|^2, \tag{2.11}$$

see Theorems 3.3 in [32, 28] and the references therein. Note that the present definition (2.8a) implies

$$\begin{aligned}\|u-u_h\|_{\#}^2 &= \min_{\zeta\in H_0^1(\Omega)} \|\nabla_\theta((u-u_h)-\zeta)\|^2 + \sum_{e\in\mathcal{E}_h} h_e^{-1}\|\Pi_e^0[\![u-u_h]\!]\|_e^2 \\ &= \min_{\zeta\in H_0^1(\Omega)} \|\nabla_\theta(u_h-\zeta)\|^2 + \sum_{e\in\mathcal{E}_h} h_e^{-1}\|\Pi_e^0[\![u_h]\!]\|_e^2,\end{aligned} \tag{2.12}$$

since $u\in H_0^1(\Omega)$ and since its jumps are zero. Thus $\|u-u_h\|_{\#}$ is a distance of u_h to the space $H_0^1(\Omega)$ and it simplifies to the energy distance $\min_{\zeta\in H_0^1(\Omega)}\|\nabla_\theta(u_h-\zeta)\| = \min_{\zeta\in H_0^1(\Omega)}\|\nabla_h(u_h-\zeta)\|$ whenever the jumps of u_h are of mean value zero, $\langle[\![u_h]\!],1\rangle_e = 0$ for all $e\in\mathcal{E}_h$. For $\underline{\boldsymbol{\Sigma}} = \underline{\boldsymbol{I}}$, our *intrinsic problem-dependent error* thus takes the form

$$|||u-u_h|||^2 = \|u-u_h\|_*^2 + \|u-u_h\|_{\#}^2 = \|\nabla_\theta(u-u_h)\|^2 + \sum_{e\in\mathcal{E}_h} h_e^{-1}\|\Pi_e^0[\![u_h]\!]\|_e^2,$$

so that in particular $|||u-u_h||| = \|\nabla_\theta(u-u_h)\|$ whenever the jumps of u_h are of mean value zero. In what concerns the first term $\|u-u_h\|_*$, using the dual norm definition (2.1a), equivalence (2.7) on $H_0^1(\Omega)$, and the weak solution characterization (1.2), it takes the form

$$\|u-u_h\|_* = \max_{\varphi\in H_0^1(\Omega);\,\|\nabla\varphi\|=1}\{(f,\varphi)-(\underline{\boldsymbol{\Sigma}}\nabla_\theta u_h,\nabla\varphi)\},$$

so that this is nothing but the dual norm of the residual. Note that only this term remains whenever $u_h\in H_0^1(\Omega)$; in this case $|||u-u_h||| = \|u-u_h\|_*$.

2.8 Orthogonality with respect to the hat functions

We conclude this introductory section by an assumption that will be crucial for some of the forthcoming results:

Assumption 2.4 (Galerkin orthogonality with respect to $\psi_{\mathbf{a}}$)**.** *There holds*

$$(\underline{\boldsymbol{\Sigma}}\nabla_\theta u_h,\nabla\psi_{\mathbf{a}})_{\omega_{\mathbf{a}}} = (f,\psi_{\mathbf{a}})_{\omega_{\mathbf{a}}} \qquad \forall\mathbf{a}\in\mathcal{V}_h^{\rm int}.$$

This assumption is naturally satisfied in most Galerkin numerical approximations of problem (1.1), namely in various conforming, nonconforming, and discontinuous Galerkin finite elements. Application to mixed finite elements can be achieved along the lines of [32, Section 4.4].

3 Equivalent localization of global dual and distance norms

This section shows that two types of global norms, dual norms on the space $H_0^1(\Omega)$ of the form $\|\cdot\|_*$ of (2.1a) and distance norms of the form $\|\cdot\|_{\#}$ of (2.8a), admit an equivalence with their local versions of the respective forms $\|\cdot\|_{*,\omega_{\mathbf{a}}}$ of (2.1b) and $\|\cdot\|_{\#,\omega_{\mathbf{a}}}$ of (2.8b). Let us note immediately that Assumption 2.4 is central for one direction in the first case. This may be seen as an extension of some previous results in [4, 17, 38, 48, 46, 24, 50, 31, 32] to the broken Sobolev space $H^1(\mathcal{T}_h)$ of (2.4). The presentation below is not necessarily linked to a posteriori error analysis and we find it of independent interest. We give direct and minimal proofs, with clearly identified constants that only depend on the mesh shape regularity $\kappa_{\mathcal{T}}$ and space dimension d. All results here actually hold for any space dimension $d\geq 1$.

3.1 Some useful local inequalities

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

$$\begin{aligned}H_*^1(\omega_{\mathbf{a}}) &:= \{v\in H^1(\omega_{\mathbf{a}});\,(v,1)_{\omega_{\mathbf{a}}}=0\}, &&\mathbf{a}\in\mathcal{V}_h^{\rm int},\\ H_*^1(\omega_{\mathbf{a}}) &:= \{v\in H^1(\omega_{\mathbf{a}});\,v=0 \text{ on } \partial\omega_{\mathbf{a}}\cap\partial\Omega\}, &&\mathbf{a}\in\mathcal{V}_h^{\rm ext}.\end{aligned} \tag{3.1}$$

It follows from [17, Theorem 3.1], [11, Section 3], see also [32, Lemma 3.12], that

$$\|\nabla(\psi_{\mathbf{a}}v)\|_{\omega_{\mathbf{a}}} \le C_{\mathrm{cont,PF}}\|\nabla v\|_{\omega_{\mathbf{a}}} \qquad \forall v \in H^1_*(\omega_{\mathbf{a}}),\ \forall \mathbf{a} \in \mathcal{V}_h, \tag{3.2}$$

where

$$C_{\mathrm{cont,PF}} := \max_{\mathbf{a}\in\mathcal{V}_h}\{1 + C_{\mathrm{PF},\omega_{\mathbf{a}}}h_{\omega_{\mathbf{a}}}\|\nabla\psi_{\mathbf{a}}\|_{\infty,\omega_{\mathbf{a}}}\} \tag{3.3}$$

only depends on the shape regularity parameter $\kappa_{\mathcal{T}}$ and possibly on the space dimension d. Here $C_{\mathrm{PF},\omega_{\mathbf{a}}}$ is the Poincaré–Friedrichs constant from

$$\|v\|_{\omega_{\mathbf{a}}} \le C_{\mathrm{PF},\omega_{\mathbf{a}}}h_{\omega_{\mathbf{a}}}\|\nabla v\|_{\omega_{\mathbf{a}}} \qquad \forall v \in H^1_*(\omega_{\mathbf{a}}),$$

see Payne and Weinberger [41] or Veeser and Verfürth [47].

Similarly, it follows as in [32, Lemma 3.13] and [33, Section 4] that

$$\|\nabla_h(\psi_{\mathbf{a}}v)\|_{\omega_{\mathbf{a}}} \le C_{\mathrm{cont,bPF}}\left(\|\nabla_h v\|_{\omega_{\mathbf{a}}} + \left\{\sum_{e\in\mathcal{E}_h,\,\mathbf{a}\in e} h_e^{-1}\|\Pi_e^0[\![v]\!]\|_e^2\right\}^{1/2}\right), \qquad \forall v \in H^1(\mathcal{T}_{\mathbf{a}}) \text{ with } (v,1)_{\omega_{\mathbf{a}}} = 0 \text{ when } \mathbf{a} \in \mathcal{V}_h^{\mathrm{int}},\ \forall \mathbf{a} \in \mathcal{V}_h, \tag{3.4}$$

where $C_{\mathrm{cont,bPF}} := \max_{\mathbf{a}\in\mathcal{V}_h}\{1 + C_{\mathrm{bPF},\omega_{\mathbf{a}}}h_{\omega_{\mathbf{a}}}\|\nabla\psi_{\mathbf{a}}\|_{\infty,\omega_{\mathbf{a}}}\}$ only depends on the shape regularity parameter $\kappa_{\mathcal{T}}$ and possibly on the space dimension d. Here $C_{\mathrm{bPF},\omega_{\mathbf{a}}}$ is the constant from the broken Poincaré–Friedrichs inequality

$$\|v\|_{\omega_{\mathbf{a}}} \le C_{\mathrm{bPF},\omega_{\mathbf{a}}}h_{\omega_{\mathbf{a}}}\left(\|\nabla_h v\|_{\omega_{\mathbf{a}}}^2 + \sum_{e\in\mathcal{E}_h,\,\mathbf{a}\in e} h_e^{-1}\|\Pi_e^0[\![v]\!]\|_e^2\right)^{1/2}, \qquad \forall v \in H^1(\mathcal{T}_{\mathbf{a}}) \text{ with } (v,1)_{\omega_{\mathbf{a}}} = 0 \text{ when } \mathbf{a} \in \mathcal{V}_h^{\mathrm{int}}, \tag{3.5}$$

see Brenner [13] or [51].

Finally, as the spaces $\mathbf{V}^0(\mathcal{T}_e)$ in the definition (2.6) of the discrete gradient consist of low-order polynomials, the inverse inequality gives

$$\|\mathbf{v}_h\cdot\mathbf{n}_e\|_e \le C_{\mathrm{inv}}h_e^{-1/2}\|\mathbf{v}_h\|_K \qquad \forall K \in \mathcal{T}_h,\ \forall e \in \mathcal{E}_K,\ \forall \mathbf{v}_h \in \mathbf{V}^0(\mathcal{T}_e), \tag{3.6}$$

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

3.2 Localization of dual (residual) norms

The following is our localization result for the dual norm of the residual $\|u-u_h\|_*$ defined by (2.1a), with the patchwise contributions $\|u-u_h\|_{*,\omega_{\mathbf{a}}}$ given by (2.1b):

Proposition 3.1 (Localization of the dual norm of the residual)**.** *Let u be the weak solution given by (1.2) and let $u_h \in H^1(\mathcal{T}_h)$ satisfying Assumption 2.4 be arbitrary. Then*

$$\|u-u_h\|_* \le (d+1)^{1/2}C_{\mathrm{cont,PF}}\left\{\sum_{\mathbf{a}\in\mathcal{V}_h}\|u-u_h\|_{*,\omega_{\mathbf{a}}}^2\right\}^{1/2}, \tag{3.7a}$$

$$\left\{\frac{1}{d+1}\sum_{\mathbf{a}\in\mathcal{V}_h}\|u-u_h\|_{*,\omega_{\mathbf{a}}}^2\right\}^{1/2} \le \|u-u_h\|_*. \tag{3.7b}$$

Remark 3.2 (Bound (3.7a) with patchwise constants)**.** *Using $C_{\mathrm{cont,PF},\omega_{\mathbf{a}}} := \{1 + C_{\mathrm{PF},\omega_{\mathbf{a}}}h_{\omega_{\mathbf{a}}}\|\nabla\psi_{\mathbf{a}}\|_{\infty,\omega_{\mathbf{a}}}\}$ in (3.7a) in place of $C_{\mathrm{cont,PF}}$, the slightly sharper bound*

$$\|u-u_h\|_* \le (d+1)^{1/2}\left\{\sum_{\mathbf{a}\in\mathcal{V}_h}C_{\mathrm{cont,PF},\omega_{\mathbf{a}}}^2\|u-u_h\|_{*,\omega_{\mathbf{a}}}^2\right\}^{1/2} \tag{3.8}$$

immediately follows.

This proposition is an immediate consequence of the following general theorem of independent interest. Recall that $C_{\mathrm{cont,PF}}$ is the constant from inequality (3.2):

Theorem 3.3 (Localization of a dual norm with $\psi_{\mathbf{a}}$-Galerkin orthogonality)**.** *Let* $\mathbf{v} \in [L^2(\Omega)]^d$. *Then, under the hat functions orthogonality condition*

$$(\mathbf{v}, \nabla\psi_{\mathbf{a}})_{\omega_{\mathbf{a}}} = 0 \qquad \forall \mathbf{a} \in \mathcal{V}_h^{\mathrm{int}}, \tag{3.9}$$

there holds

$$\max_{\varphi\in H_0^1(\Omega);\, \|\nabla\varphi\|=1} (\mathbf{v}, \nabla\varphi)^2 \le (d+1) C_{\mathrm{cont,PF}}^2 \sum_{\mathbf{a}\in\mathcal{V}_h} \max_{\varphi\in H_0^1(\omega_{\mathbf{a}});\, \|\nabla\varphi\|_{\omega_{\mathbf{a}}}=1} (\mathbf{v}, \nabla\varphi)_{\omega_{\mathbf{a}}}^2. \tag{3.10a}$$

There always holds

$$\sum_{\mathbf{a}\in\mathcal{V}_h} \max_{\varphi\in H_0^1(\omega_{\mathbf{a}});\, \|\nabla\varphi\|_{\omega_{\mathbf{a}}}=1} (\mathbf{v}, \nabla\varphi)_{\omega_{\mathbf{a}}}^2 \le (d+1) \max_{\varphi\in H_0^1(\Omega);\, \|\nabla\varphi\|=1} (\mathbf{v}, \nabla\varphi)^2. \tag{3.10b}$$

Proof. Let $\varphi \in H_0^1(\Omega)$ with $\|\nabla\varphi\| = 1$ be fixed. The partition of unity by the hat functions $\psi_{\mathbf{a}}$, $\sum_{\mathbf{a}\in\mathcal{V}_h} \psi_{\mathbf{a}} = 1$, and the Galerkin orthogonality with respect to $\psi_{\mathbf{a}}$ expressed by (3.9) give

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

where $\Pi_{0,\omega_{\mathbf{a}}}\varphi$ is the mean value of the function φ on the patch $\omega_{\mathbf{a}}$. There holds $(\varphi - \Pi_{0,\omega_{\mathbf{a}}}\varphi)|_{\omega_{\mathbf{a}}} \in H_*^1(\omega_{\mathbf{a}})$ for the space $H_*^1(\omega_{\mathbf{a}})$ given by (3.1) and $(\psi_{\mathbf{a}}(\varphi - \Pi_{0,\omega_{\mathbf{a}}}\varphi))|_{\omega_{\mathbf{a}}} \in H_0^1(\omega_{\mathbf{a}})$ for an interior vertex $\mathbf{a} \in \mathcal{V}_h^{\mathrm{int}}$. Similarly, $\varphi|_{\omega_{\mathbf{a}}} \in H_*^1(\omega_{\mathbf{a}})$ and $(\psi_{\mathbf{a}}\varphi)|_{\omega_{\mathbf{a}}} \in H_0^1(\omega_{\mathbf{a}})$ for a boundary vertex $\mathbf{a} \in \mathcal{V}_h^{\mathrm{ext}}$. Thus, passing to a maximum and using inequality (3.2) yields, for any interior vertex $\mathbf{a} \in \mathcal{V}_h^{\mathrm{int}}$,

$$\begin{aligned}(\mathbf{v}, \nabla(\psi_{\mathbf{a}}(\varphi - \Pi_{0,\omega_{\mathbf{a}}}\varphi)))_{\omega_{\mathbf{a}}} &= \|\nabla(\psi_{\mathbf{a}}(\varphi - \Pi_{0,\omega_{\mathbf{a}}}\varphi))\|_{\omega_{\mathbf{a}}} \Big(\mathbf{v}, \frac{\nabla(\psi_{\mathbf{a}}(\varphi - \Pi_{0,\omega_{\mathbf{a}}}\varphi))}{\|\nabla(\psi_{\mathbf{a}}(\varphi - \Pi_{0,\omega_{\mathbf{a}}}\varphi))\|_{\omega_{\mathbf{a}}}}\Big)_{\omega_{\mathbf{a}}} \\ &\le \|\nabla(\psi_{\mathbf{a}}(\varphi - \Pi_{0,\omega_{\mathbf{a}}}\varphi))\|_{\omega_{\mathbf{a}}} \max_{\varphi\in H_0^1(\omega_{\mathbf{a}});\, \|\nabla\varphi\|_{\omega_{\mathbf{a}}}=1} (\mathbf{v}, \nabla\varphi)_{\omega_{\mathbf{a}}} \\ &\le C_{\mathrm{cont,PF}} \|\nabla(\varphi - \Pi_{0,\omega_{\mathbf{a}}}\varphi)\|_{\omega_{\mathbf{a}}} \max_{\varphi\in H_0^1(\omega_{\mathbf{a}});\, \|\nabla\varphi\|_{\omega_{\mathbf{a}}}=1} (\mathbf{v}, \nabla\varphi)_{\omega_{\mathbf{a}}} \\ &= C_{\mathrm{cont,PF}} \|\nabla\varphi\|_{\omega_{\mathbf{a}}} \max_{\varphi\in H_0^1(\omega_{\mathbf{a}});\, \|\nabla\varphi\|_{\omega_{\mathbf{a}}}=1} (\mathbf{v}, \nabla\varphi)_{\omega_{\mathbf{a}}},\end{aligned}$$

finally employing that the gradient of a constant vanishes. A similar estimate holds for $\mathbf{a} \in \mathcal{V}_h^{\mathrm{ext}}$. Thus, the Cauchy–Schwarz inequality gives

$$(\mathbf{v}, \nabla\varphi)^2 \le C_{\mathrm{cont,PF}}^2 \sum_{\mathbf{a}\in\mathcal{V}_h} \|\nabla\varphi\|_{\omega_{\mathbf{a}}}^2 \sum_{\mathbf{a}\in\mathcal{V}_h} \max_{\varphi\in H_0^1(\omega_{\mathbf{a}});\, \|\nabla\varphi\|_{\omega_{\mathbf{a}}}=1} (\mathbf{v}, \nabla\varphi)_{\omega_{\mathbf{a}}}^2.$$

Now the fact that each simplex has $d+1$ vertices gives

$$\sum_{\mathbf{a}\in\mathcal{V}_h} \|\nabla\varphi\|_{\omega_{\mathbf{a}}}^2 = \sum_{\mathbf{a}\in\mathcal{V}_h} \sum_{K\in\mathcal{T}_{\mathbf{a}}} \|\nabla\varphi\|_K^2 = \sum_{K\in\mathcal{T}_h} \sum_{\mathbf{a}\in\mathcal{V}_K} \|\nabla\varphi\|_K^2 = (d+1)\|\nabla\varphi\|^2, \tag{3.11}$$

so that the premise $\|\nabla\varphi\| = 1$ finally yields (3.10a).

The converse estimate (3.10b) does not need the hypothesis (3.9). Let $\mathbf{a} \in \mathcal{V}_h$ and let $\zeta^{\mathbf{a}} \in H_0^1(\omega_{\mathbf{a}})$ be defined by the lifting

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

Then

$$(\mathbf{v}, \nabla\zeta^{\mathbf{a}})_{\omega_{\mathbf{a}}} = (\nabla\zeta^{\mathbf{a}}, \nabla\zeta^{\mathbf{a}})_{\omega_{\mathbf{a}}} = \max_{\varphi\in H_0^1(\omega_{\mathbf{a}});\, \|\nabla\varphi\|_{\omega_{\mathbf{a}}}=1} (\nabla\zeta^{\mathbf{a}}, \nabla\varphi)_{\omega_{\mathbf{a}}}^2 = \max_{\varphi\in H_0^1(\omega_{\mathbf{a}});\, \|\nabla\varphi\|_{\omega_{\mathbf{a}}}=1} (\mathbf{v}, \nabla\varphi)_{\omega_{\mathbf{a}}}^2.$$

Consequently, taking $\zeta := \sum_{\mathbf{a}\in\mathcal{V}_h} \zeta^{\mathbf{a}} \in H^1_0(\Omega)$,

$$\sum_{\mathbf{a}\in\mathcal{V}_h} \max_{\varphi\in H^1_0(\omega_\mathbf{a});\,\|\nabla\varphi\|_{\omega_\mathbf{a}}=1} (\mathbf{v},\nabla\varphi)^2_{\omega_\mathbf{a}} = \sum_{\mathbf{a}\in\mathcal{V}_h} (\mathbf{v},\nabla\zeta^{\mathbf{a}})_{\omega_\mathbf{a}} = (\mathbf{v},\nabla\zeta) \\ \le \max_{\varphi\in H^1_0(\Omega);\,\|\nabla\varphi\|=1} (\mathbf{v},\nabla\varphi)\|\nabla\zeta\|,$$

where we finally passed to the maximum. Noticing that

$$\|\nabla\zeta\|^2 = \sum_{K\in\mathcal{T}_h} \Big\| \sum_{\mathbf{a}\in\mathcal{V}_K} (\nabla\zeta^{\mathbf{a}})|_K \Big\|^2_K \le (d+1) \sum_{\mathbf{a}\in\mathcal{V}_h} \|\nabla\zeta^{\mathbf{a}}\|^2_{\omega_\mathbf{a}},$$

we arrive at (3.10b). □

Remark 3.4 (Further generalization)**.** *Theorem 3.3 has recently been extended to any bounded linear functional on the Sobolev space $W^{1,p}_0(\Omega)$, $p>1$, in [8].*

3.3 Localization of distances to the energy space

Recall that d is the space dimension, ∇_h is the broken gradient given by (2.5), ∇_θ is the discrete gradient given by (2.6) with the parameter $\theta\in\{-1,0,1\}$, and that the constants $C_{\mathrm{cont,bPF}}$ and C_{inv} are respectively given by (3.4) and (3.6). It appears that the distance $\|u-u_h\|_\#$ defined in (2.8a) admits a similar localization property for the contributions $\|u-u_h\|_{\#,\omega_\mathbf{a}}$ defined in (2.8b), as this was the case for $\|u-u_h\|_*$ in Proposition 3.1:

Proposition 3.5 (Localization of the distance to the energy space)**.** *Let $u\in H^1_0(\Omega)$ and $u_h\in H^1(\mathcal{T}_h)$ be arbitrary. Then*

$$\|u-u_h\|^2_\# \le C^2_{\mathrm{loc}} \sum_{\mathbf{a}\in\mathcal{V}_h} \|u-u_h\|^2_{\#,\omega_\mathbf{a}},$$
$$\sum_{\mathbf{a}\in\mathcal{V}_h} \|u-u_h\|^2_{\#,\omega_\mathbf{a}} \le (d+1)\|u-u_h\|^2_\#,$$

where

$$C^2_{\mathrm{loc}} := 8(d+1)C^2_{\mathrm{cont,bPF}} + \frac{1}{d}\big(2|\theta|^2(d+1)C^2_{\mathrm{inv}}+1\big) + \frac{1}{d}\big(8|\theta|^2(d+1)^3C^2_{\mathrm{cont,bPF}}C^2_{\mathrm{inv}}\big). \tag{3.12}$$

To prove this result, the following theorem of independent interest will be crucial:

Theorem 3.6 (Localization of a global distance for jumps of mean value zero)**.** *Let $v\in H^1(\mathcal{T}_h)$. Then, when the jumps of v have zero mean values, i.e.*

$$\langle [\![v]\!], 1\rangle_e = 0 \qquad \forall e\in\mathcal{E}_h, \tag{3.13}$$

there holds

$$\min_{\zeta\in H^1_0(\Omega)} \|\nabla_h(v-\zeta)\|^2 \le (d+1)C^2_{\mathrm{cont,bPF}} \sum_{\mathbf{a}\in\mathcal{V}_h} \min_{\zeta\in H^1_\#(\omega_\mathbf{a})} \|\nabla_h(v-\zeta)\|^2_{\omega_\mathbf{a}}. \tag{3.14a}$$

There always holds

$$\sum_{\mathbf{a}\in\mathcal{V}_h} \min_{\zeta\in H^1_\#(\omega_\mathbf{a})} \|\nabla_h(v-\zeta)\|^2_{\omega_\mathbf{a}} \le (d+1) \min_{\zeta\in H^1_0(\Omega)} \|\nabla_h(v-\zeta)\|^2; \tag{3.14b}$$

(3.14b) *also holds for the discrete gradient ∇_θ in place of the broken gradient ∇_h.*

Proof. The second claim (3.14b) is immediate, as local-best approximation is always subordinate to the global-best one:

$$\sum_{\mathbf{a}\in\mathcal{V}_h} \min_{\zeta\in H^1_\#(\omega_\mathbf{a})} \|\nabla_h(v-\zeta)\|^2_{\omega_\mathbf{a}} \le \min_{\zeta\in H^1_0(\Omega)} \sum_{\mathbf{a}\in\mathcal{V}_h} \|\nabla_h(v-\zeta)\|^2_{\omega_\mathbf{a}} = (d+1) \min_{\zeta\in H^1_0(\Omega)} \|\nabla_h(v-\zeta)\|^2.$$

In the inequality, we have used that restriction of any $\zeta \in H^1_0(\Omega)$ to the patch subdomain $\omega_{\mathbf{a}}$ for any vertex $\mathbf{a} \in \mathcal{V}_h$ lies in the space $H^1_{\#}(\omega_{\mathbf{a}})$ given by (2.9); in the equality, the fact that each element $K \in \mathcal{T}_h$ lies in $(d+1)$ patches has been employed as in (3.11). This estimate is obviously the same for $\nabla_h v$ replaced by $\nabla_\theta v$. The rest of the proof is thus dedicated to showing the first claim (3.14a).

Let, for a given vertex $\mathbf{a} \in \mathcal{V}_h$, $s^{\mathbf{a}}$ be defined by the orthogonal projection of the function $\psi_{\mathbf{a}} v$ onto the space $H^1_0(\omega_{\mathbf{a}})$,

$$s^{\mathbf{a}} := \arg \min_{\zeta \in H^1_0(\omega_{\mathbf{a}})} \|\nabla_h(\psi_{\mathbf{a}} v - \zeta)\|_{\omega_{\mathbf{a}}}; \tag{3.15}$$

equivalently, $s^{\mathbf{a}} \in H^1_0(\omega_{\mathbf{a}})$ solves

$$(\nabla s^{\mathbf{a}}, \nabla \zeta)_{\omega_{\mathbf{a}}} = (\nabla_h(\psi_{\mathbf{a}} v), \nabla \zeta)_{\omega_{\mathbf{a}}} \qquad \forall \zeta \in H^1_0(\omega_{\mathbf{a}}).$$

Extending $s^{\mathbf{a}}$ by zero outside of $\omega_{\mathbf{a}}$ and setting $s := \sum_{\mathbf{a} \in \mathcal{V}_h} s^{\mathbf{a}} \in H^1_0(\Omega)$, we have, also employing the partition of unity $\sum_{\mathbf{a} \in \mathcal{V}_K} \psi_{\mathbf{a}}|_K = 1|_K$,

$$\begin{aligned} \min_{\zeta \in H^1_0(\Omega)} \|\nabla_h(v - \zeta)\|^2 \le \|\nabla_h(v - s)\|^2 &= \sum_{K \in \mathcal{T}_h} \|\nabla_h(v - s)\|_K^2 \\ &= \sum_{K \in \mathcal{T}_h} \Big\| \sum_{\mathbf{a} \in \mathcal{V}_K} (\nabla_h(\psi_{\mathbf{a}} v - s^{\mathbf{a}}))|_K \Big\|_K^2 \le (d+1) \sum_{\mathbf{a} \in \mathcal{V}_h} \|\nabla_h(\psi_{\mathbf{a}} v - s^{\mathbf{a}})\|_{\omega_{\mathbf{a}}}^2. \end{aligned} \tag{3.16}$$

The fact that $\psi_{\mathbf{a}} \zeta \in H^1_0(\omega_{\mathbf{a}})$ for any $\zeta \in H^1_{\#}(\omega_{\mathbf{a}})$ gives from (3.15)

$$\|\nabla_h(\psi_{\mathbf{a}} v - s^{\mathbf{a}})\|_{\omega_{\mathbf{a}}} \le \inf_{\zeta \in H^1_{\#}(\omega_{\mathbf{a}})} \|\nabla_h(\psi_{\mathbf{a}}(v - \zeta))\|_{\omega_{\mathbf{a}}}. \tag{3.17}$$

Let $H^1_{\#,v}(\omega_{\mathbf{a}}) := H^1_{\#}(\omega_{\mathbf{a}})$ for $\mathbf{a} \in \mathcal{V}_h^{\mathrm{ext}}$ and $H^1_{\#,v}(\omega_{\mathbf{a}}) := \{\zeta \in H^1_{\#}(\omega_{\mathbf{a}});\ (\zeta, 1)_{\omega_{\mathbf{a}}} = (v, 1)_{\omega_{\mathbf{a}}}\}$ for $\mathbf{a} \in \mathcal{V}_h^{\mathrm{int}}$. Introducing this space allows us to restrain the arguments to mean value zero on vertices $\mathbf{a} \in \mathcal{V}_h^{\mathrm{int}}$, so that we can employ inequality (3.4). Therein the jumps actually disappear thanks to the present simplifying assumption (3.13). In combination with the Cauchy–Schwarz inequality, we obtain

$$\begin{aligned} \inf_{\zeta \in H^1_{\#}(\omega_{\mathbf{a}})} \|\nabla_h(\psi_{\mathbf{a}}(v - \zeta))\|_{\omega_{\mathbf{a}}} &\le \inf_{\zeta \in H^1_{\#,v}(\omega_{\mathbf{a}})} \|\nabla_h(\psi_{\mathbf{a}}(v - \zeta))\|_{\omega_{\mathbf{a}}} \\ &\le C_{\mathrm{cont,bPF}} \min_{\zeta \in H^1_{\#,v}(\omega_{\mathbf{a}})} \|\nabla_h(v - \zeta)\|_{\omega_{\mathbf{a}}} \\ &= C_{\mathrm{cont,bPF}} \min_{\zeta \in H^1_{\#}(\omega_{\mathbf{a}})} \|\nabla_h(v - \zeta)\|_{\omega_{\mathbf{a}}}; \end{aligned} \tag{3.18}$$

in the final equality, we have employed that the gradient of a constant on the patch $\omega_{\mathbf{a}}$ vanishes. Collecting these results finishes the proof. □

Proof of Proposition 3.5. The equality

$$\sum_{\mathbf{a} \in \mathcal{V}_h} \sum_{e \in \mathcal{E}_h,\, \mathbf{a} \in e} h_e^{-1} \|\Pi_e^0 [\![v]\!]\|_e^2 = d \sum_{e \in \mathcal{E}_h} h_e^{-1} \|\Pi_e^0 [\![v]\!]\|_e^2 \qquad \forall v \in H^1(\mathcal{T}_h) \tag{3.19}$$

follows immediately from the fact that each face is shared by d vertices. The second claim of Proposition 3.5 then follows from inequality (3.14b) employed for the discrete gradient ∇_θ and using definitions (2.8a) and (2.8b) together with property (2.12). We now turn to the proof of the first claim of Proposition 3.5.

Let $v \in H^1(\mathcal{T}_h)$ be arbitrary (not subject to condition (3.13)) and recall that by definition (2.8a), one has $\|v\|_{\#}^2 = \min_{\zeta \in H^1_0(\Omega)} \|\nabla_\theta(v - \zeta)\|^2 + \sum_{e \in \mathcal{E}_h} h_e^{-1} \|\Pi_e^0 [\![v]\!]\|_e^2$. From (3.19), the jump terms immediately take the local form requested in (2.8b). Denote $s_1 := \arg\min_{\zeta \in H^1_0(\Omega)} \|\nabla_\theta(v - \zeta)\|$ and $s_2 := \arg\min_{\zeta \in H^1_0(\Omega)} \|\nabla_h(v - \zeta)\|$. The minimization property of s_1 together with the discrete gradient definition (2.6) and the fact that the jumps of s_2 are zero give

$$\|\nabla_\theta(v - s_1)\|^2 \le \|\nabla_\theta(v - s_2)\|^2 = \Big\| \nabla_h(v - s_2) - \theta \sum_{e \in \mathcal{E}_h} \mathfrak{l}_e([\![v]\!]) \Big\|^2 \le 2\|\nabla_h(v - s_2)\|^2 + 2|\theta|^2 \Big\| \sum_{e \in \mathcal{E}_h} \mathfrak{l}_e([\![v]\!]) \Big\|^2. \tag{3.20}$$

Recall that $\mathcal{T}_e$ regroups the mesh elements sharing the face e. Thus, employing the inverse inequality (3.6) for the jump term above, we infer, as in [15, Proof of Theorem 6.3],

$$\left\| \sum_{e\in\mathcal{E}_h} \mathfrak{l}_e([\![v]\!]) \right\|^2 \le (d+1) \sum_{e\in\mathcal{E}_h} \|\mathfrak{l}_e([\![v]\!])\|_{\mathcal{T}_e}^2 \le (d+1) C_{\mathrm{inv}}^2 \sum_{e\in\mathcal{E}_h} h_e^{-1} \|\Pi_e^0 [\![v]\!]\|_e^2. \tag{3.21}$$

Finally, for the bound on $\|\nabla_h(v-s_2)\|$, we use that s_2 is the minimizer for the broken gradient ∇_h and proceed as in the proof of Theorem 3.6. In particular, both (3.16) and (3.17) hold true, whereas in (3.18), we need to employ inequality (3.4) without assumption (3.13), yielding

$$\inf_{\zeta\in H^1_{\#}(\omega_{\mathbf{a}})} \|\nabla_h(\psi_{\mathbf{a}}(v-\zeta))\|_{\omega_{\mathbf{a}}} \le C_{\mathrm{cont,bPF}} \left(\min_{\zeta\in H^1_{\#}(\omega_{\mathbf{a}})} \|\nabla_h(v-\zeta)\|_{\omega_{\mathbf{a}}} + \left\{ \sum_{e\in\mathcal{E}_h,\, \mathbf{a}\in e} h_e^{-1} \|\Pi_e^0 [\![v]\!]\|_e^2 \right\}^{1/2} \right). \tag{3.22}$$

Consequently,

$$\|\nabla_h(v-s_2)\|^2 \le 2(d+1) C_{\mathrm{cont,bPF}}^2 \sum_{\mathbf{a}\in\mathcal{V}_h} \left\{ \min_{\zeta\in H^1_{\#}(\omega_{\mathbf{a}})} \|\nabla_h(v-\zeta)\|_{\omega_{\mathbf{a}}}^2 + \sum_{e\in\mathcal{E}_h,\, \mathbf{a}\in e} h_e^{-1} \|\Pi_e^0 [\![v]\!]\|_e^2 \right\}.$$

This already gives an upper bound with a local minimization structure, and we are left to make reappear the discrete gradients in place of the broken ones. In order to do so, we proceed similarly to (3.20),

$$\begin{aligned} \sum_{\mathbf{a}\in\mathcal{V}_h} \min_{\zeta\in H^1_{\#}(\omega_{\mathbf{a}})} \|\nabla_h(v-\zeta)\|_{\omega_{\mathbf{a}}}^2 &\le 2 \sum_{\mathbf{a}\in\mathcal{V}_h} \min_{\zeta\in H^1_{\#}(\omega_{\mathbf{a}})} \|\nabla_\theta(v-\zeta)\|_{\omega_{\mathbf{a}}}^2 + 2|\theta|^2 \sum_{\mathbf{a}\in\mathcal{V}_h} \left\| \sum_{e\in\mathcal{E}_h} \mathfrak{l}_e([\![v]\!]) \right\|_{\omega_{\mathbf{a}}}^2 \\ &\le 2 \sum_{\mathbf{a}\in\mathcal{V}_h} \min_{\zeta\in H^1_{\#}(\omega_{\mathbf{a}})} \|\nabla_\theta(v-\zeta)\|_{\omega_{\mathbf{a}}}^2 + 2|\theta|^2 (d+1)^2 C_{\mathrm{inv}}^2 \sum_{e\in\mathcal{E}_h} h_e^{-1} \|\Pi_e^0 [\![v]\!]\|_e^2, \end{aligned} \tag{3.23}$$

where we have used $\sum_{\mathbf{a}\in\mathcal{V}_h} \|\sum_{e\in\mathcal{E}_h} \mathfrak{l}_e([\![v]\!])\|_{\omega_{\mathbf{a}}}^2 = (d+1)\|\sum_{e\in\mathcal{E}_h} \mathfrak{l}_e([\![v]\!])\|^2$ like in (3.11) and (3.21). Altogether,

$$\|v\|_{\#}^2 \le C_{\mathrm{loc}}^2 \sum_{\mathbf{a}\in\mathcal{V}_h} \left\{ \min_{\zeta\in H^1_{\#}(\omega_{\mathbf{a}})} \|\nabla_\theta(v-\zeta)\|_{\omega_{\mathbf{a}}}^2 + \sum_{e\in\mathcal{E}_h,\, \mathbf{a}\in e} h_e^{-1} \|\Pi_e^0 [\![v]\!]\|_e^2 \right\},$$

where C_{loc}^2 is given by (3.12), and the proof is concluded using the property (2.12) to apply the derived result to $v = u - u_h$. □

4 Guaranteed, robust, and locally efficient a posteriori estimates in a unified framework

We present in this section our a posteriori estimates on the error in a numerical approximation of problem (1.1). The estimates give guaranteed global error upper bound (global reliability). Crucially, we achieve all robustness with respect to the jumps and anisotropy of the diffusion tensor $\underline{\boldsymbol{\Sigma}}$, robustness with respect to the approximation polynomial degree, and local error lower bound (local efficiency); the latter in consequence of the localization results of Section 3. Our results are presented in an abstract framework, following [31, 32, 33]. This enables to cover at once basically any classical numerical method, in particular all types of conforming, nonconforming, mixed, and discontinuous Galerkin finite elements. The key idea is to build a piecewise polynomial $H_0^1(\Omega)$-conforming potential reconstruction and a piecewise polynomial $\mathbf{H}(\mathrm{div},\Omega)$-conforming equilibrated flux reconstruction, in extension of the methodology developed in [42, 36, 34, 25, 26, 37, 1, 40, 12, 43, 2, 52, 18, 5] and the references therein.

4.1 Flux and potential reconstruction

Let $\mathbb{P}_p(\mathcal{T}_h)$, $p \ge 0$, stand for piecewise polynomials on the mesh $\mathcal{T}_h$ of total degree at most p; we will denote by Π_p the $L^2(\Omega)$-orthogonal projection onto $\mathbb{P}_p(\mathcal{T}_h)$. For vector-valued functions, the Raviart–Thomas–Nédélec mixed finite element spaces will be used; $\mathbf{RTN}_p(\mathcal{T}_h) := \{\mathbf{v}_h \in [L^2(\Omega)]^d; \mathbf{v}_h|_K \in \mathbf{RTN}_p(K)\}$, $p \ge 0$,

with the local spaces $\mathbf{RTN}_p(K) := [\mathbb{P}_p(K)]^d + \mathbb{P}_p(K)\mathbf{x}$, $K \in \mathcal{T}_h$, see Brezzi and Fortin [14] or Roberts and Thomas [44].

To obtain an $\mathbf{H}(\mathrm{div},\Omega)$-conforming flux reconstruction, we solve homogeneous local Neumann (Neumann–Dirichlet close to the boundary) problems over patches of elements $\mathcal{T}_\mathbf{a}$ via the mixed finite element method:

Definition 4.1 (Equilibrated flux reconstruction)**.** *Let $u_h \in H^1(\mathcal{T}_h)$ satisfy Assumption 2.4. For all vertices $\mathbf{a} \in \mathcal{V}_h$, set*

$$\begin{aligned}
&\mathbf{V}_h^\mathbf{a} := \{\mathbf{v}_h \in \mathbf{RTN}_p(\mathcal{T}_\mathbf{a}) \cap \mathbf{H}(\mathrm{div},\omega_\mathbf{a});\ \mathbf{v}_h\cdot\mathbf{n}_{\omega_\mathbf{a}} = 0 \text{ on } \partial\omega_\mathbf{a}\},\\
&Q_h^\mathbf{a} := \{q_h \in \mathbb{P}_p(\mathcal{T}_\mathbf{a});\ (q_h,1)_{\omega_\mathbf{a}} = 0\},
\end{aligned} \qquad \mathbf{a} \in \mathcal{V}_h^{\mathrm{int}},$$

$$\begin{aligned}
&\mathbf{V}_h^\mathbf{a} := \{\mathbf{v}_h \in \mathbf{RTN}_p(\mathcal{T}_\mathbf{a}) \cap \mathbf{H}(\mathrm{div},\omega_\mathbf{a});\ \mathbf{v}_h\cdot\mathbf{n}_{\omega_\mathbf{a}} = 0 \text{ on } \partial\omega_\mathbf{a}\setminus\partial\Omega\},\\
&Q_h^\mathbf{a} := \mathbb{P}_p(\mathcal{T}_\mathbf{a}),
\end{aligned} \qquad \mathbf{a} \in \mathcal{V}_h^{\mathrm{ext}}.$$

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

$$(\boldsymbol{\sigma}_h^\mathbf{a}, \mathbf{v}_h)_{\omega_\mathbf{a}} - (\bar r_h^\mathbf{a}, \nabla\cdot\mathbf{v}_h)_{\omega_\mathbf{a}} = -(\psi_\mathbf{a}\underline{\boldsymbol{\Sigma}}\nabla_\theta u_h, \mathbf{v}_h)_{\omega_\mathbf{a}} \qquad \forall \mathbf{v}_h \in \mathbf{V}_h^\mathbf{a}, \tag{4.1a}$$

$$(\nabla\cdot\boldsymbol{\sigma}_h^\mathbf{a}, q_h)_{\omega_\mathbf{a}} = (\psi_\mathbf{a} f - \underline{\boldsymbol{\Sigma}}\nabla_\theta u_h\cdot\nabla\psi_\mathbf{a}, q_h)_{\omega_\mathbf{a}} \qquad \forall q_h \in Q_h^\mathbf{a} \tag{4.1b}$$

and define, after extension by zero outside of $\omega_\mathbf{a}$,

$$\boldsymbol{\sigma}_h := \sum_{\mathbf{a}\in\mathcal{V}_h} \boldsymbol{\sigma}_h^\mathbf{a}.$$

To obtain an $H_0^1(\Omega)$-conforming potential reconstruction, we solve homogeneous local Dirichlet problems over patches of elements $\mathcal{T}_\mathbf{a}$ via the finite element method:

Definition 4.2 (Potential reconstruction)**.** *Let $u_h \in H^1(\mathcal{T}_h)$. For all vertices $\mathbf{a} \in \mathcal{V}_h$, set*

$$W_h^\mathbf{a} := \mathbb{P}_{p+1}(\mathcal{T}_\mathbf{a}) \cap H_0^1(\omega_\mathbf{a}).$$

Then prescribe $s_h^\mathbf{a} \in W_h^\mathbf{a}$ by solving

$$(\nabla s_h^\mathbf{a}, \nabla\zeta_h)_{\omega_\mathbf{a}} = (\nabla_h(\psi_\mathbf{a} u_h), \nabla\zeta_h)_{\omega_\mathbf{a}} \qquad \forall \zeta_h \in W_h^\mathbf{a} \tag{4.2}$$

and define, after extension by zero outside of $\omega_\mathbf{a}$,

$$s_h := \sum_{\mathbf{a}\in\mathcal{V}_h} s_h^\mathbf{a}.$$

The two above constructions yield a piecewise vector-valued polynomial $\boldsymbol{\sigma}_h \in \mathbf{RTN}_p(\mathcal{T}_h)\cap\mathbf{H}(\mathrm{div},\Omega)$ and a piecewise scalar-valued polynomial $s_h \in \mathbb{P}_{p+1}\cap H_0^1(\Omega)$. In practice, the approximate solution u_h is a piecewise p-degree polynomial, see Assumption 4.6 below, and this fixes the degree p in Definitions 4.1 and 4.2. It is also easy to verify that $\nabla\cdot\boldsymbol{\sigma}_h = \Pi_p f$, see [12] or [32, Lemma 3.5]. Problems (4.1) and (4.2) actually admit local minimization characterizations, see, e.g., [32, Remarks 3.7 and 3.10] and [33, Corollaries 3.1 and 3.3]:

Remark 4.3 (Local minimizations)**.** *Problems (4.1) and (4.2) can be equivalently rewritten as*

$$\boldsymbol{\sigma}_h^\mathbf{a} := \arg \min_{\mathbf{v}_h\in\mathbf{V}_h^\mathbf{a},\ \nabla\cdot\mathbf{v}_h = \Pi_{Q_h^\mathbf{a}}(\psi_\mathbf{a} f - \underline{\boldsymbol{\Sigma}}\nabla_\theta u_h\cdot\nabla\psi_\mathbf{a})} \|\psi_\mathbf{a}\underline{\boldsymbol{\Sigma}}\nabla_\theta u_h + \mathbf{v}_h\|_{\omega_\mathbf{a}} \qquad \forall \mathbf{a} \in \mathcal{V}_h, \tag{4.3a}$$

$$s_h^\mathbf{a} := \arg \min_{\zeta_h\in W_h^\mathbf{a}} \|\nabla_h(\psi_\mathbf{a} u_h - \zeta_h)\|_{\omega_\mathbf{a}} \qquad \forall \mathbf{a} \in \mathcal{V}_h. \tag{4.3b}$$

Remark 4.4 (Discrete and broken gradients)**.** *In practice, one could also choose $s_h^\mathbf{a} := \arg\min_{\zeta_h\in W_h^\mathbf{a}} \|\nabla_\theta(\psi_\mathbf{a} u_h - \zeta_h)\|_{\omega_\mathbf{a}}$ instead of (4.3b). The current choice is motivated by the key property (4.5b) below which enables to prove the (local) efficiencies in Theorem 4.7.*

4.2 Guaranteed error control

We present here our a posteriori error estimate on the intrinsic error $|||u-u_h|||$ given by (2.10), still merely under Assumption 2.4, in the very abstract setting $u_h \in H^1(\mathcal{T}_h)$. Define the data oscillation estimators

$$\eta_{\mathrm{osc},K} := \frac{h_K}{\pi}\|f-\Pi_p f\|_K, \qquad K\in\mathcal{T}_h.$$

It follows as in [32] and the references therein that:

Theorem 4.5 (A posteriori estimate in a unified framework)**.** *Let u be the weak solution of problem (1.1) given by (1.2). Let $u_h \in H^1(\mathcal{T}_h)$ satisfying Assumption 2.4 be arbitrary. Consider the equilibrated flux reconstruction of Definition 4.1 and the potential reconstruction of Definition 4.2. Then the error $u-u_h$ measured in intrinsic norm (2.10) can be estimated by*

$$|||u-u_h|||^2 \le \sum_{K\in\mathcal{T}_h}(\|\underline{\boldsymbol{\Sigma}}\nabla_\theta u_h+\boldsymbol{\sigma}_h\|_K+\eta_{\mathrm{osc},K})^2+\sum_{K\in\mathcal{T}_h}\|\nabla_\theta(u_h-s_h)\|_K^2+\sum_{e\in\mathcal{E}_h}h_e^{-1}\|\Pi_e^0[\![u_h]\!]\|_e^2. \tag{4.4}$$

4.3 Robust (local) efficiency

We now prove the converse statement to Theorem 4.5, and this locally in the neighborhood of each mesh element. Results of Section 3 are of course crucial here, stating that the intrinsic norm (2.10) in which we measure the error indeed admits a local structure. For this local efficiency result, we need to suppose that the approximate solution u_h is a piecewise polynomial of the degree p that we have already used in Section 4.1 for the flux and potential reconstructions:

Assumption 4.6 (Piecewise polynomial approximation)**.** *The approximate solution u_h is a piecewise polynomial of degree $p\ge 1$, $u_h\in\mathbb{P}_p(\mathcal{T}_h)$.*

The crucial ingredient for local efficiency under Assumption 4.6 are the following two stability results for the problems (4.1) and (4.2), shown respectively in [11, Theorem 7] and in [32, Corollary 3.16] in two space dimensions and extended to three space dimensions in [33, Corollaries 3.1 and 3.3] (recall that the space $H^1_*(\omega_{\mathbf{a}})$ is given by (3.1)):

$$\|\psi_{\mathbf{a}}\underline{\boldsymbol{\Sigma}}\nabla_\theta u_h+\boldsymbol{\sigma}_h^{\mathbf{a}}\|_{\omega_{\mathbf{a}}} \le C_{\mathrm{st}}\max_{\varphi\in H^1_*(\omega_{\mathbf{a}});\,\|\nabla\varphi\|_{\omega_{\mathbf{a}}}=1}\{-(\psi_{\mathbf{a}}\underline{\boldsymbol{\Sigma}}\nabla_\theta u_h,\nabla\varphi)_{\omega_{\mathbf{a}}}+(\Pi_p(\psi_{\mathbf{a}}f)-\underline{\boldsymbol{\Sigma}}\nabla_\theta u_h\cdot\nabla\psi_{\mathbf{a}},\varphi)_{\omega_{\mathbf{a}}}\}, \tag{4.5a}$$

$$\min_{\zeta_h\in W_h^{\mathbf{a}}}\|\nabla_h(\psi_{\mathbf{a}}u_h-\zeta_h)\|_{\omega_{\mathbf{a}}} \le C_{\mathrm{st}}\min_{\zeta\in H^1_0(\omega_{\mathbf{a}})}\|\nabla_h(\psi_{\mathbf{a}}u_h-\zeta)\|_{\omega_{\mathbf{a}}}. \tag{4.5b}$$

Here C_{st} is a constant that only depends on the mesh shape regularity $\kappa_{\mathcal{T}}$ and on the space dimension d. A computable upper bound on C_{st} is given in [32, Lemma 3.23]. Note that (4.5b) is stated for the broken gradient (2.5).

Define a local efficiency data oscillation term

$$\tilde\eta_{\mathrm{osc},K} := \frac{h_K}{\pi}\|\psi_{\mathbf{a}}f-\Pi_p(\psi_{\mathbf{a}}f)\|_K, \qquad K\in\mathcal{T}_h, \tag{4.6}$$

together with $\tilde\eta_{\mathrm{osc}} := \{\sum_{K\in\mathcal{T}_h}(\tilde\eta_{\mathrm{osc},K})^2\}^{1/2}$. Recall that the dual norm of the residual $\|u-u_h\|_*$ is defined by (2.1a) and it localizes following Proposition 3.1; the distance to the energy space $\|u-u_h\|_\#$ is given by (2.8a) and it localizes following Proposition 3.5; the broken gradient is given by (2.5) and the discrete gradient by (2.6); the constants $C_{\mathrm{cont,PF}}$, $C_{\mathrm{cont,bPF}}$, and C_{loc} are respectively given by (3.2), (3.4), and (3.12). We then have:

Theorem 4.7 (Local and global efficiency and robustness for Theorem 4.5)**.** *Let u be the weak solution given by (1.2) and let u_h verify Assumptions 2.4 and 4.6. Then, for $\boldsymbol{\sigma}_h$ given by Definition 4.1,*

$$\|\underline{\boldsymbol{\Sigma}}\nabla_\theta u_h+\boldsymbol{\sigma}_h\|_K \le C_{\mathrm{st}}C_{\mathrm{cont,PF}}\sum_{\mathbf{a}\in\mathcal{V}_K}\|u-u_h\|_{*,\omega_{\mathbf{a}}}+C_{\mathrm{st}}\sum_{\mathbf{a}\in\mathcal{V}_K}\left\{\sum_{K'\in\mathcal{T}_{\mathbf{a}}}\tilde\eta_{\mathrm{osc},K'}^2\right\}^{1/2} \qquad \forall K\in\mathcal{T}_h, \tag{4.7a}$$

$$\|\underline{\boldsymbol{\Sigma}}\nabla_\theta u_h+\boldsymbol{\sigma}_h\| \le (d+1)C_{\mathrm{st}}C_{\mathrm{cont,PF}}\|u-u_h\|_*+(d+1)C_{\mathrm{st}}\tilde\eta_{\mathrm{osc}}. \tag{4.7b}$$

Similarly, for s_h given by Definition 4.2, when $\langle [\![u_h]\!], 1\rangle_e = 0$ for all faces $e \in \mathcal{E}_h$,

$$\|\nabla_h(u_h - s_h)\|_K \le C_{\rm st} C_{\rm cont,bPF} \sum_{\mathbf{a}\in\mathcal{V}_K} \|u - u_h\|_{\#,\omega_\mathbf{a}} \qquad \forall K \in \mathcal{T}_h, \tag{4.8a}$$

$$\|\nabla_h(u_h - s_h)\| \le (d+1) C_{\rm st} C_{\rm cont,bPF} \|u - u_h\|_{\#} \tag{4.8b}$$

and in general

$$\|\nabla_\theta(u_h - s_h)\|_K \le C_{\rm st} C_{\rm loc} \sum_{\mathbf{a}\in\mathcal{V}_K} \sum_{\mathbf{a}'\in\overline{\omega_\mathbf{a}}} \|u - u_h\|_{\#,\omega_{\mathbf{a}'}} \qquad \forall K \in \mathcal{T}_h, \tag{4.9a}$$

$$\|\nabla_\theta(u_h - s_h)\| \le (d+1)^{1/2} C_{\rm st} C_{\rm loc} \|u - u_h\|_{\#}. \tag{4.9b}$$

There always holds

$$h_e^{-1/2} \|\Pi_e^0 [\![u_h]\!]\|_e = h_e^{-1/2} \|\Pi_e^0 [\![u - u_h]\!]\|_e \qquad \forall e \in \mathcal{E}_h. \tag{4.10}$$

Proof. Assertion (4.7a) follows as in [11, Theorem 1], cf. also [32, Theorem 3.17], whereas inequality (4.7b) can be shown as in [32, Lemma 3.22]. As (4.10) is straightforward, we only prove inequalities (4.8) and (4.9).

Let first $\langle [\![u_h]\!], 1\rangle_e = 0$ for all $e \in \mathcal{E}_h$, so that in particular $\nabla_h = \nabla_\theta$. Fix an element $K \in \mathcal{T}_h$. Using Definition 4.2 of $s_h^\mathbf{a}$ that yields (4.3b), the potential reconstruction decomposition $s_h|_K = \sum_{\mathbf{a}\in\mathcal{V}_K} s_h^\mathbf{a}|_K$, the partition of unity by the hat functions $\sum_{\mathbf{a}\in\mathcal{V}_K} \psi_\mathbf{a}|_K = 1|_K$, the triangle inequality, and enlarging the integration set, we infer

$$\|\nabla_h(u_h - s_h)\|_K = \left\| \sum_{\mathbf{a}\in\mathcal{V}_K} (\nabla_h(\psi_\mathbf{a} u_h - s_h^\mathbf{a}))|_K \right\|_K \le \sum_{\mathbf{a}\in\mathcal{V}_K} \|\nabla_h(\psi_\mathbf{a} u_h - s_h^\mathbf{a})\|_{\omega_\mathbf{a}}. \tag{4.11}$$

Now, the stability (4.5b), inequalities (3.17)–(3.18) with $v = u_h$, the local norm definition (2.8b), and the fact that $\langle [\![u_h]\!], 1\rangle_e = 0$ for all $e \in \mathcal{E}_h$ imply $\|u_h\|_{\#,\omega_\mathbf{a}} = \|u - u_h\|_{\#,\omega_\mathbf{a}}$, proceeding as in (2.12). Thus

$$\|\nabla_h(\psi_\mathbf{a} u_h - s_h^\mathbf{a})\|_{\omega_\mathbf{a}} \le C_{\rm st} \min_{\zeta\in H_0^1(\omega_\mathbf{a})} \|\nabla_h(\psi_\mathbf{a} u_h - \zeta)\|_{\omega_\mathbf{a}} \le C_{\rm st} C_{\rm cont,bPF} \|u - u_h\|_{\#,\omega_\mathbf{a}}. \tag{4.12}$$

Thus (4.8a) follows. The global efficiency (4.8b) is then a consequence of the estimate of the form (3.16) together with (4.11), (4.12), (3.14b), and the norm definitions (2.8)

$$\begin{aligned}
\|\nabla_h(u_h - s_h)\|^2 &\le (d+1) \sum_{\mathbf{a}\in\mathcal{V}_h} \|\nabla_h(\psi_\mathbf{a} u_h - s_h^\mathbf{a})\|_{\omega_\mathbf{a}}^2 \\
&\le (d+1) C_{\rm st}^2 \sum_{\mathbf{a}\in\mathcal{V}_h} \min_{\zeta\in H_0^1(\omega_\mathbf{a})} \|\nabla_h(\psi_\mathbf{a} u_h - \zeta)\|_{\omega_\mathbf{a}}^2 \\
&\le (d+1) C_{\rm st}^2 C_{\rm cont,bPF}^2 \sum_{\mathbf{a}\in\mathcal{V}_h} \|u - u_h\|_{\#,\omega_\mathbf{a}}^2 \\
&\le (d+1)^2 C_{\rm st}^2 C_{\rm cont,bPF}^2 \min_{\zeta\in H_0^1(\Omega)} \|\nabla_h((u - u_h) - \zeta)\|^2.
\end{aligned} \tag{4.13}$$

In order to show (4.9a), remark first that, using the discrete gradient definition (2.6), the triangle inequality, and (4.11),

$$\|\nabla_\theta(u_h - s_h)\|_K = \left\| \nabla_h(u_h - s_h) - \theta \sum_{e\in\mathcal{E}_K} \mathfrak{l}_e([\![u_h]\!]) \right\|_K \le \sum_{\mathbf{a}\in\mathcal{V}_K} \|\nabla_h(\psi_\mathbf{a} u_h - s_h^\mathbf{a})\|_{\omega_\mathbf{a}} + |\theta| \left\| \sum_{e\in\mathcal{E}_K} \mathfrak{l}_e([\![u_h]\!]) \right\|_K.$$

First, as in (3.21), employing definition (2.8b),

$$\left\| \sum_{e\in\mathcal{E}_K} \mathfrak{l}_e([\![u_h]\!]) \right\|_K \le (d+1)^{1/2} C_{\rm inv} \left\{ \sum_{e\in\mathcal{E}_K} h_e^{-1} \|\Pi_e^0 [\![v]\!]\|_e^2 \right\}^{1/2} \le \frac{1}{d} (d+1)^{1/2} C_{\rm inv} \sum_{\mathbf{a}\in\mathcal{V}_K} \|u - u_h\|_{\#,\omega_\mathbf{a}}.$$

Next, the finite element stability (4.5b) together with (3.17) and (3.22) give

$$\|\nabla_h(\psi_{\mathbf{a}}u_h - s_h^{\mathbf{a}})\|_{\omega_{\mathbf{a}}} \le C_{\rm st} \min_{\zeta\in H^1_0(\omega_{\mathbf{a}})} \|\nabla_h(\psi_{\mathbf{a}}u_h - \zeta)\|_{\omega_{\mathbf{a}}} \le C_{\rm st} \inf_{\zeta\in H^1_\#(\omega_{\mathbf{a}})} \|\nabla_h(\psi_{\mathbf{a}}(u_h - \zeta))\|_{\omega_{\mathbf{a}}}$$
$$\le C_{\rm st}C_{\rm cont,bPF}\left(\min_{\zeta\in H^1_\#(\omega_{\mathbf{a}})} \|\nabla_h(u_h - \zeta)\|_{\omega_{\mathbf{a}}} + \left\{ \sum_{e\in\mathcal{E}_h,\, \mathbf{a}\in e} h_e^{-1}\|\Pi_e^0[\![u_h]\!]\|_e^2 \right\}^{1/2} \right).$$

Using once more the discrete gradient definition (2.6) and proceeding as in (3.23),

$$\min_{\zeta\in H^1_\#(\omega_{\mathbf{a}})} \|\nabla_h(u_h - \zeta)\|_{\omega_{\mathbf{a}}}^2 \le 2 \min_{\zeta\in H^1_\#(\omega_{\mathbf{a}})} \|\nabla_\theta(u_h - \zeta)\|_{\omega_{\mathbf{a}}}^2 + 2|\theta|^2 \left\| \sum_{e\in\mathcal{E}_h, e\in\overline{\omega_{\mathbf{a}}}} \mathfrak{l}_e([\![u_h]\!]) \right\|_{\omega_{\mathbf{a}}}^2$$
$$\le 2 \min_{\zeta\in H^1_\#(\omega_{\mathbf{a}})} \|\nabla_\theta(u_h - \zeta)\|_{\omega_{\mathbf{a}}}^2 + 2|\theta|^2(d+1)C_{\rm inv}^2 \sum_{e\in\mathcal{E}_h, e\in\overline{\omega_{\mathbf{a}}}} h_e^{-1}\|\Pi_e^0[\![u_h]\!]\|_e^2$$
$$\le \Big(2 + \frac{1}{d}2|\theta|^2(d+1)C_{\rm inv}^2\Big) \sum_{\mathbf{a}'\in\overline{\omega_{\mathbf{a}}}} \|u - u_h\|_{\#,\omega_{\mathbf{a}'}}^2,$$

where we have used estimate of the form (3.21) for such faces $e \in \mathcal{E}_h$ that lie in the closure of the patch subdomain $\overline{\omega_{\mathbf{a}}}$, whence

$$\min_{\zeta\in H^1_\#(\omega_{\mathbf{a}})} \|\nabla_h(u_h - \zeta)\|_{\omega_{\mathbf{a}}} \le \Big(2 + \frac{1}{d}2|\theta|^2(d+1)C_{\rm inv}^2\Big)^{1/2} \sum_{\mathbf{a}'\in\overline{\omega_{\mathbf{a}}}} \|u - u_h\|_{\#,\omega_{\mathbf{a}'}}.$$

Altogether,

$$\|\nabla_\theta(u_h - s_h)\|_K \le C_{\rm st}C_{\rm cont,bPF}\Big(2 + \frac{1}{d}2|\theta|^2(d+1)C_{\rm inv}^2\Big)^{1/2} \sum_{\mathbf{a}\in\mathcal{V}_K} \sum_{\mathbf{a}'\in\overline{\omega_{\mathbf{a}}}} \|u - u_h\|_{\#,\omega_{\mathbf{a}'}}$$
$$+ C_{\rm st}C_{\rm cont,bPF} \sum_{\mathbf{a}\in\mathcal{V}_K} \|u - u_h\|_{\#,\omega_{\mathbf{a}}}$$
$$+ \frac{1}{d}|\theta|(d+1)^{1/2}C_{\rm inv} \sum_{\mathbf{a}\in\mathcal{V}_K} \|u - u_h\|_{\#,\omega_{\mathbf{a}}},$$

which proves (4.9a), using that $C_{\rm st} \ge 1$ and definition (3.12) of the constant $C_{\rm loc}$.

Finally, for the global bound (4.9b), we first use, as in (3.20)–(3.21),

$$\|\nabla_\theta(u_h - s_h)\|^2 \le 2\|\nabla_h(u_h - s_h)\|^2 + 2|\theta|^2(d+1)C_{\rm inv}^2 \sum_{e\in\mathcal{E}_h} h_e^{-1}\|\Pi_e^0[\![u_h]\!]\|_e^2.$$

One next employs the first line of (4.13). From there, the conclusion follows as in the proof of Proposition 3.5. □

Remark 4.8 (Efficiency in the L^2 flux norm for Theorem 4.7)**.** *The Cauchy–Schwarz inequality gives* $\|v\|_{*,\omega_{\mathbf{a}}} \le \|\underline{\boldsymbol{\Sigma}}\nabla_\theta v\|_{\omega_{\mathbf{a}}}$ *and* $\|v\|_* \le \|\underline{\boldsymbol{\Sigma}}\nabla_\theta v\|$ *for all* $v \in H^1(\mathcal{T}_{\mathbf{a}})$ *from* (2.1)*, so that, immediately,*

$$\|\underline{\boldsymbol{\Sigma}}\nabla_\theta u_h + \boldsymbol{\sigma}_h\|_K \le C_{\rm st}C_{\rm cont,PF} \sum_{\mathbf{a}\in\mathcal{V}_K} \|\underline{\boldsymbol{\Sigma}}\nabla_\theta(u - u_h)\|_{\omega_{\mathbf{a}}} + C_{\rm st} \sum_{\mathbf{a}\in\mathcal{V}_K} \left\{ \sum_{K'\in\mathcal{T}_{\mathbf{a}}} \tilde\eta_{{\rm osc},K'}^2 \right\}^{1/2} \qquad \forall K \in \mathcal{T}_h,$$
$$\|\underline{\boldsymbol{\Sigma}}\nabla_\theta u_h + \boldsymbol{\sigma}_h\| \le (d+1)C_{\rm st}C_{\rm cont,PF}\|\underline{\boldsymbol{\Sigma}}\nabla_\theta(u - u_h)\| + (d+1)C_{\rm st}\tilde\eta_{\rm osc}.$$

Applications to conforming, nonconforming, mixed, and discontinuous Galerkin approximations are straightforward following [32, Section 4].

Remark 4.9 (More general diffusion tensors $\underline{\boldsymbol{\Sigma}}$)**.** *For Theorems 4.5 and 4.7, the requirement of piecewise constant diffusion tensor* $\underline{\boldsymbol{\Sigma}}$ *from Assumption 2.1 is unavoidable. It is namely crucial for inequality* (4.5a) *to hold. If* $\underline{\boldsymbol{\Sigma}}$ *is piecewise polynomial of degree* p' *and* u_h *is piecewise polynomial of degree* p*, then* $\mathbf{RTN}_{p+p'}(\mathcal{T}_{\mathbf{a}})$ *spaces would need to be chosen in Definition 4.1 to maintain the present form of the results; otherwise a supplementary oscillation term of the datum* $\underline{\boldsymbol{\Sigma}}$ *of the form of* (4.6) *would appear in Theorem 4.7.*

Remark 4.10 (Polynomial degree and cost of the reconstructions)**.** *The reconstruction of Definition 4.1 relies on solution of local problems with* $\mathbf{RTN}_p(\mathcal{T}_\mathbf{a})$*-spaces, whereas that of Definition 4.2 on solution of local problems with* $\mathbb{P}_{p+1}(\mathcal{T}_\mathbf{a})$ *spaces. Although these constructions are local, the associated computational burden may not be completely negligible. There exist various ways how to decrease it. First, the proofs in [11] and [33] actually show that the solves of local problems on each patch* $\mathcal{T}_\mathbf{a}$ *by finite elements can be replaced by an explicit run through* $\mathcal{T}_\mathbf{a}$ *and a local construction inside each mesh element. This explicit construction remarkably maintains the polynomial-degree robustness. Equilibrated reconstruction in* $\mathbf{RTN}_{p-1}(\mathcal{T}_\mathbf{a})$ *for* $u_h \in \mathbb{P}_p$ *has also been suggested in [11] and analyzed in [31, Section 6.2]; one does not know here, however, whether it leads to polynomial-degree robustness. A recent survey of cheaper (but possibly not polynomial-degree robust) a posteriori estimators via reconstructions can be found in [5].*

5 Numerical experiments

We report here the results of two numerical experiments, while relying on the conforming piecewise affine finite element approximation: find $u_h \in V_h := \mathbb{P}_1(\mathcal{T}_h) \cap H^1_0(\Omega)$ such that

$$(\underline{\boldsymbol{\Sigma}}\nabla u_h, \nabla v_h) = (f, v_h) \qquad \forall v_h \in V_h. \tag{5.1}$$

The experiments were implemented by Jan Blechta (Charles University in Prague) using the `dolfin-tape` [7] package built on top of the FEniCS Project [3].

We start by noting that that since $u_h \in H^1_0(\Omega)$, $|||u - u_h||| = \|u - u_h\|_*$, and the nonconformity error $\|u - u_h\|_\#$ is zero. We will focus on our a posteriori error estimates of Theorem 4.5, while tracing the error $\|u - u_h\|_*$ defined by (2.1a) and the estimate of (4.4) that simplifies to

$$\|u - u_h\|_* \leq \eta := \left\{ \sum_{K \in \mathcal{T}_h} \left(\|\underline{\boldsymbol{\Sigma}}\nabla u_h + \boldsymbol{\sigma}_h\|_K + \eta_{\mathrm{osc},K} \right)^2 \right\}^{1/2}; \tag{5.2}$$

indeed, $s_h = u_h$ and $[\![u_h]\!] = 0$ for all $e \in \mathcal{E}_h$ here, cf. (4.9). The efficiency of our estimates, proven by (4.7b) of Theorem 4.7, is in practice best appreciated by the effectivity index

$$\mathrm{Eff} := \frac{\eta}{\|u - u_h\|_*}. \tag{5.3}$$

In what follows, we compute $\|u - u_h\|_*$ approximately with relative accuracy 10^{-2}, see the details in [8, Section 5]. We will also display the canonical $H^1_0(\Omega)$-norm of the error $\|\nabla(u - u_h)\|$.

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 by the so-called newest-vertex bisection refinement algorithm.

5.1 A regular weak solution

We first consider the test case from [39, 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 $\underline{\boldsymbol{\Sigma}}|_{\Omega_+} = \sigma_+ \underline{\boldsymbol{I}}$, $\underline{\boldsymbol{\Sigma}}|_{\Omega_-} = \sigma_- \underline{\boldsymbol{I}}$ with $\sigma_+ = 1$ and $\sigma_- < 0$. The exact solution is given by

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

and the (inhomogeneous) source term f is prescribed accordingly. Note that this solution indeed leads to the homogeneous Dirichlet boundary condition. Together with its finite element approximation and the corresponding initial mesh for the setting $\sigma_- = -1/3$, it is presented in Figure 1. Higher values of the approximate solution can be noted in particular in the left subdomain Ω_-. Specifying the operator T as is in [39], see also Remark 5.1 below, one can see that the problem is well-posed when $\sigma_- \neq -1$.

The $\|u - u_h\|_*$ and $\|\nabla(u - u_h)\|$ errors and the estimate η of (5.2) are traced in the left parts of Figures 2–4, for three different choices of the parameter σ_-. The corresponding effectivity indices given

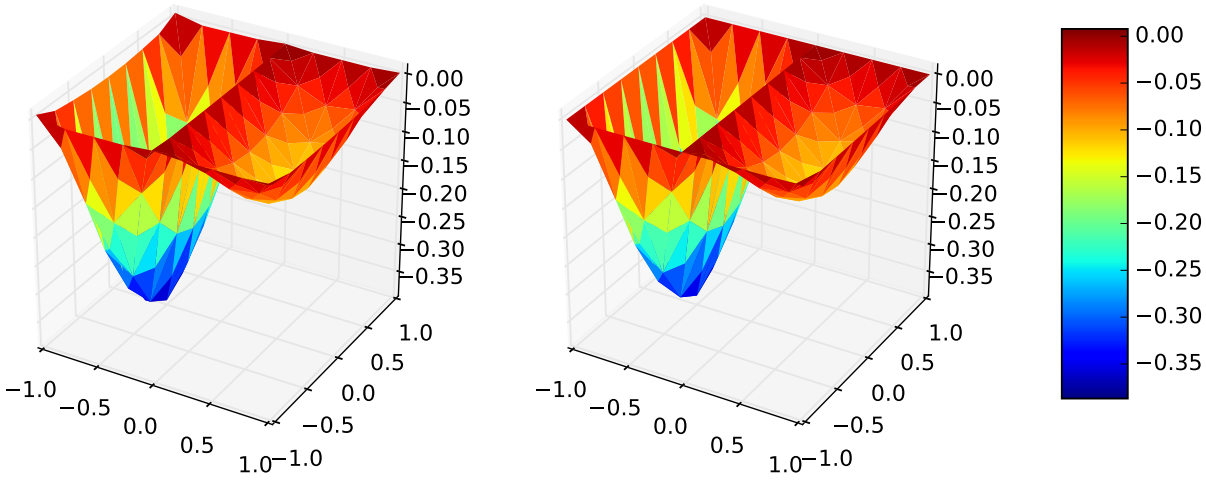

Figure 1: Exact (left) and approximate (right) solution with the corresponding initial mesh, the regular case with $\sigma_- = -1/3$

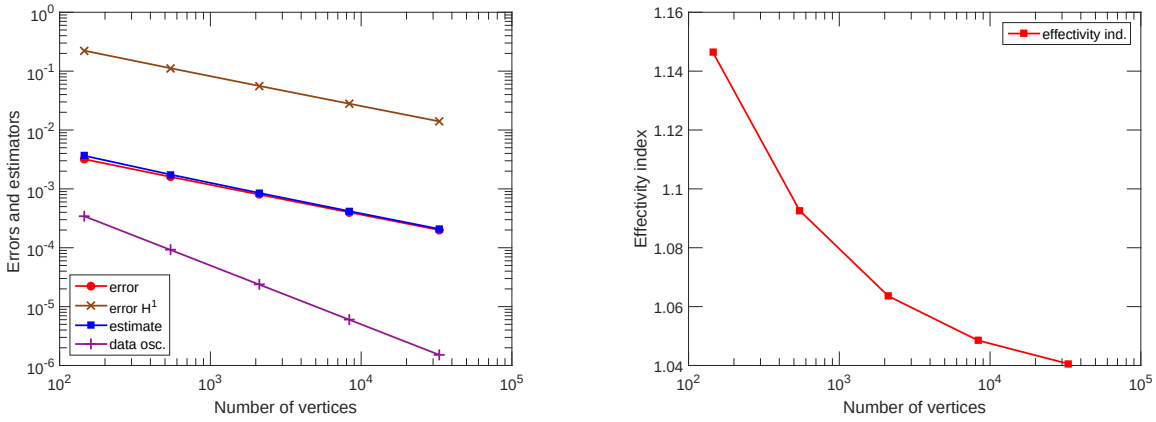

Figure 2: Estimates and errors for uniform mesh refinement (left) and the corresponding effectivity indices (right), the regular case with $\sigma_- = -0.01$

by (5.3) are plotted in the right parts of these figures. We observe a systematic first-order decrease of both errors, as predicted by the a priori error analysis, cf. [20]. The overall estimator η of (5.2), as well as its principal component given by $\{\sum_{K\in\mathcal{T}_h} \|\underline{\boldsymbol{\Sigma}}\nabla u_h + \boldsymbol{\sigma}_h\|_K^2\}^{1/2}$, also decrease with first order, in agreement with Theorems 4.5 and 4.7. On the other hand, the data oscillation estimator $\{\sum_{K\in\mathcal{T}_h} \eta_{\mathrm{osc},K}^2\}^{1/2}$ decreases with a slope of two and its influence rapidly diminishes.

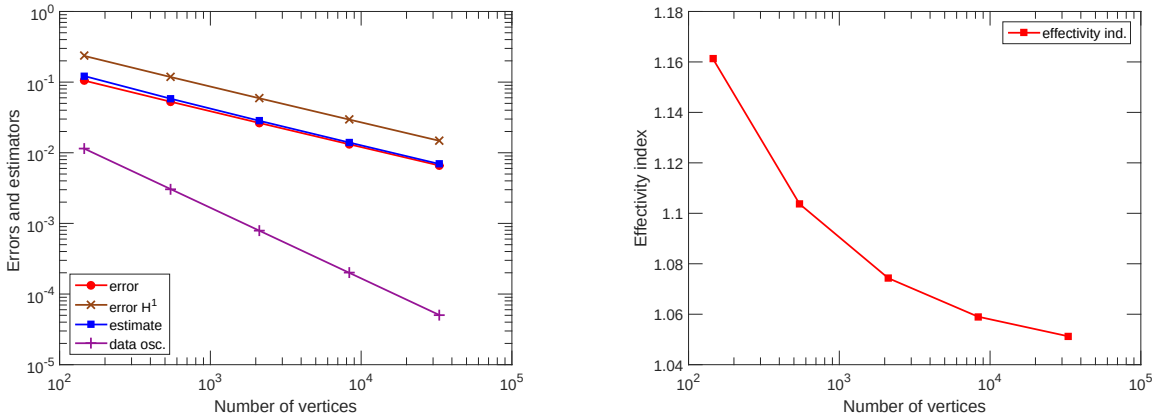

Figure 3: Estimates and errors for uniform mesh refinement (left) and the corresponding effectivity indices (right), the regular case with $\sigma_- = -1/3$

The effectivity indices in all the three settings are very close to the optimal value of one, including the last challenging case $\sigma_- = -0.99$ which is very close to the well-posedness limit. This clearly demonstrates the robustness of our estimates with respect to the jump and sign-change in the diffusion tensor $\underline{\boldsymbol{\Sigma}}$, if the error is measured in the intrinsic norm $\|u - u_h\|_*$. It can be noted from Figures 2–4 that such a robustness does not hold for the canonical norm $\|\nabla(u - u_h)\|$. Similarly, the upper bound $\|\underline{\boldsymbol{\Sigma}}\nabla(u - u_h)\|$ and the lower bound $\frac{(\underline{\boldsymbol{\Sigma}}\nabla(u-u_h),\nabla(\mathrm{T}(u-u_h)))}{\|\nabla(\mathrm{T}(u-u_h))\|}$ on the intrinsic error $\|u - u_h\|_*$ given by (2.2) seem rather $\underline{\boldsymbol{\Sigma}}$- and T-dependent, see in particular the numerical study in [22, Section 6]. Finally, Figure 5 illustrates that the distribution of the error is predicted very correctly by our estimators (plotting by a piecewise affine function is done as explained in [8, Section 5]).

We finish this section by a remark relative to the specific case $\sigma_- = -1/3$:

Remark 5.1 (Equivalence of the intrinsic norm with its upper and lower bounds)**.** *Consider the intrinsic norm $\|u-u_h\|_*$ given by (2.1a) together with its upper $\|\underline{\boldsymbol{\Sigma}}\nabla(u-u_h)\|$ and lower bounds $\frac{(\underline{\boldsymbol{\Sigma}}\nabla(u-u_h),\nabla(\mathrm{T}(u-u_h)))}{\|\nabla(\mathrm{T}(u-u_h))\|}$ that follow from (2.2). Interestingly enough, they all coincide in the case $\sigma_- = -1/3$. To explain this behavior, note first that $\|\underline{\boldsymbol{\Sigma}}\nabla(u - u_h)\|$ and $\|u - u_h\|_*$ will coincide whenever $\underline{\boldsymbol{\Sigma}}\nabla(u - u_h)$ is a gradient of some scalar field from $H^1_0(\Omega)$. This will happen when curl of $\underline{\boldsymbol{\Sigma}}\nabla(u - u_h) = 0$ on all $K \in \mathcal{T}_h$ and $[\![\mathrm{R}_{\frac{\pi}{2}}\underline{\boldsymbol{\Sigma}}\nabla(u - u_h)\cdot\mathbf{n}]\!] = 0$ on all $e \in \mathcal{E}_h$. These conditions are actually 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*

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

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 $\mathrm{T}u$ in Ω_- simplifies to $(\mathrm{T}u)|_{\Omega_-} = (-1 - 2\sigma_-)u_-$. With the

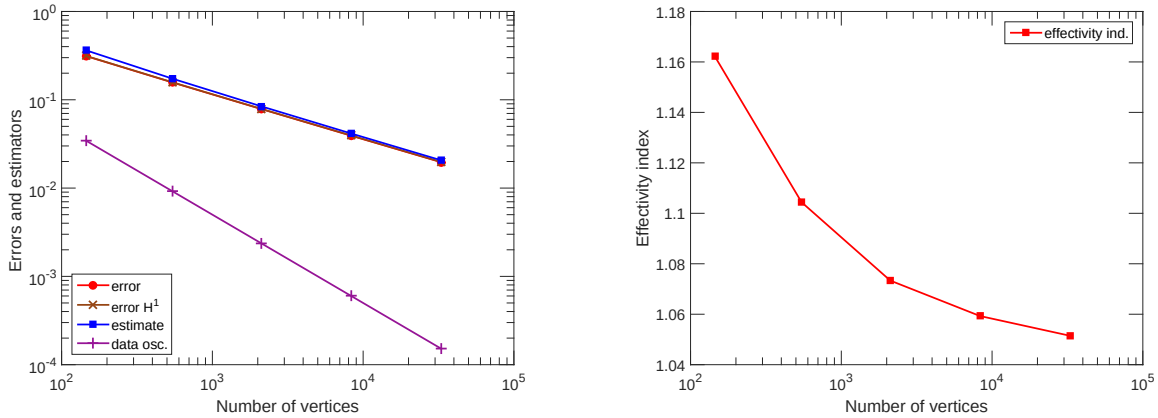

Figure 4: Estimates and errors for uniform mesh refinement (left) and the corresponding effectivity indices (right), the regular case with $\sigma_- = -0.99$

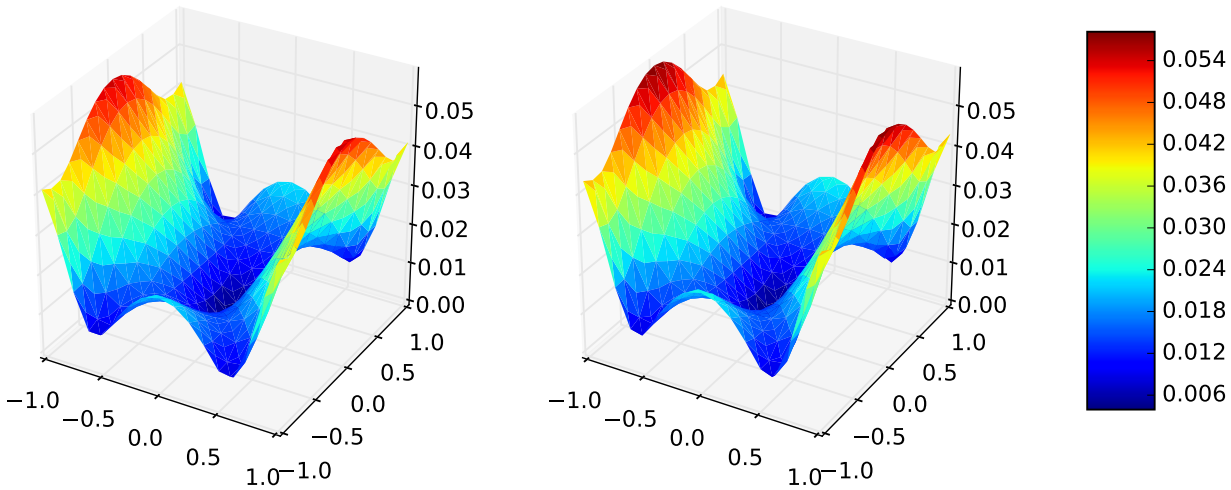

Figure 5: Exact (left) and estimated (right) error distribution, the regular case with $\sigma_- = -1/3$

help of these expressions, one may compute exactly

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

For $\sigma_- \in (-1/3, 0)$*, it holds that* $R(\sigma_-) \in (1/\sqrt{2}, 1)$ *and moreover* $R(-1/3) = 1$*, whereas* $\lim_{\sigma_- \to -1} R(\sigma_-) = 0$*. To obtain the same result for the ratio* $\frac{(\underline{\mathbf{\Sigma}}\nabla(u-u_h), \nabla(\mathrm{T}(u-u_h)))}{\|\nabla(\mathrm{T}(u-u_h))\|\,\|\underline{\mathbf{\Sigma}}\nabla(u-u_h)\|}$*, one needs to work with* symmetric *meshes*

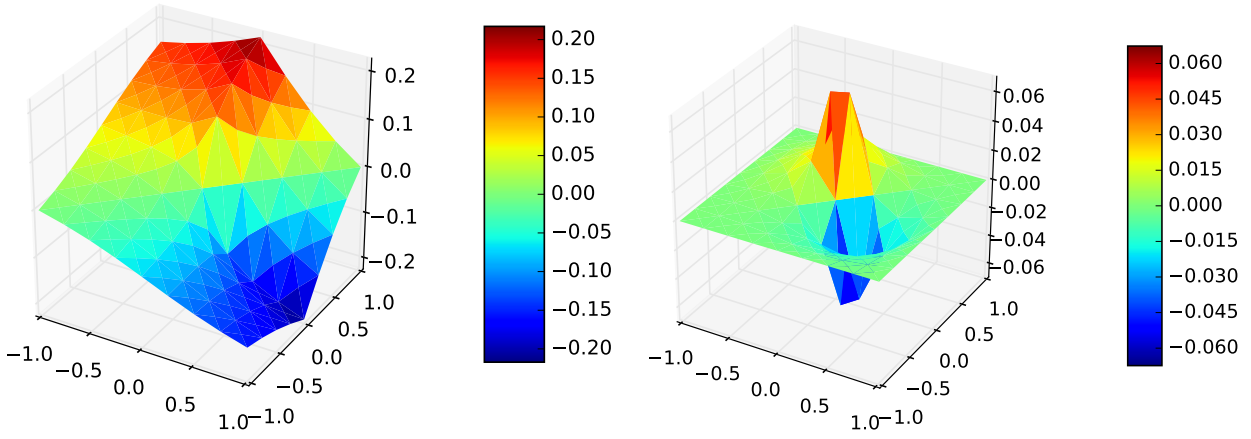

Figure 6: Approximate solution u_h (left) the pointwise error $u - u_h$ (right) on the corresponding initial mesh, the singular case with $\sigma_- = -3.1$

with respect to the line $\{x = 0\}$*, that is,* globally T-conform meshes *in the sense of [20]. In this case, one has* $\mathrm{T}V_h = V_h$*, so that the properties of* 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 of the formulation* (5.1).

5.2 A singular weak solution

We next consider the test case from [39, 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 $\underline{\boldsymbol{\Sigma}}|_{\Omega_+} = \sigma_+ \underline{\boldsymbol{I}}$, $\underline{\boldsymbol{\Sigma}}|_{\Omega_-} = \sigma_- \underline{\boldsymbol{I}}$ with $\sigma_+ = 1$ and $\sigma_- < 0$. The exact solution is according to Bonnet-BenDhia *et al.* [10] given by

$$\begin{aligned} u(x, y) &= r^\lambda (c_1 \sin(\lambda\theta) + c_2 \sin(\lambda(\pi/2 - \theta))) \text{ for } (x, y) \in \Omega_+, \\ u(x, y) &= r^\lambda (d_1 \sin(\lambda(\theta - \pi/2)) + d_2 \sin(\lambda(2\pi - \theta)) \text{ for } (x, y) \in \Omega_-. \end{aligned} \tag{5.4}$$

Here (r, θ) are the polar coordinates centered at the origin 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. The finite element approximation on the coarsest mesh for the case $\sigma_- = -3.1$ is presented in the left part of Figure 6. The steep gradient around the origin of the exact solution is largely missed by the approximation, as it can be seen from the right part of Figure 6. The inhomogeneous Dirichlet boundary condition is prescribed according to (5.4). It is naturally treated in the reconstruction of Definition 4.1, see [28, Definition 3.5], but we neglect here the additional quadrature estimator that theoretically appears in the upper and lower bounds, see [28, Theorems 3.3 and 3.12] and [33, Corollary 3.8]. The source term f corresponding to (5.4) is equal to 0; consequently, the data oscillation estimators $\eta_{\mathrm{osc},K}$ in (5.2) vanish. The operator T is specified in [39]; the problem is in particular well-posed when $\sigma_- < -3$ or $-1/3 < \sigma_- < 0$.

The intrinsic error norm $\|u - u_h\|_*$ together with the canonical error norm $\|\nabla(u - u_h)\|$ and the estimator η given by (5.2) are presented in the left parts of Figures 7–8. The corresponding effectivity indices are then given in the right parts of these figures. They are remarkably close to the optimal value of one in all the settings, illustrating numerically the robustness that has been proven in Section 4. The convergence orders of both errors and of the estimate for uniform mesh refinement correspond to the a priori analysis, being 0.46 and 0.12 respectively in the two settings. For adaptive mesh refinement (after a preliminary phase for the strongly singular setting), the convergence orders are optimal and close to 1. Finally, the predicted spatial distribution of the error still seems to be very accurate even in the close-to-the-limit singular case $\sigma_- = -3.1$, see Figure 9.

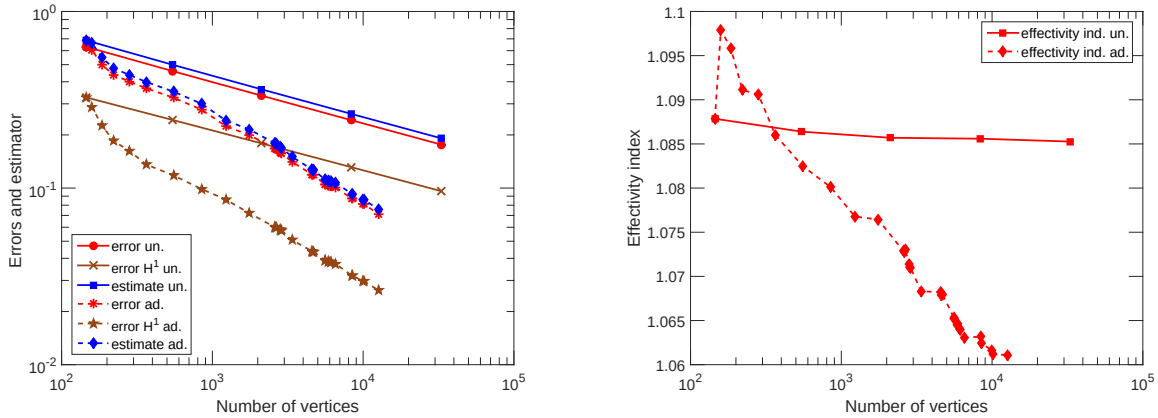

Figure 7: Estimates and errors for uniform and adaptive mesh refinement (left) and the corresponding effectivity indices (right), the singular case with $\sigma_- = -5$

6 Conclusions and outlook

We have shown in this work that globally defined dual norms as well as globally defined distance norms to the energy space admit an equivalent localization. Direct proofs with clearly identified constants are given. In the setting of the transmission problem with sign-changing coefficients (1.1), this suggests that the intrinsic global norm (2.10) is suitable for a posteriori error analysis. Indeed, relying on the concept of flux and potential reconstructions, we have obtained a guaranteed upper bound, as well as local lower bounds up to a generic constant independent of the jump or sign change in the diffusion coefficient and the approximation polynomial degree. This robustness is moreover obtained in a unified framework covering basically all classical numerical methods. Numerical experiments, in the conforming finite element setting, confirm these results [–]. Possible future developments include control of the error from a not completely converged linear solver (and corresponding stopping criteria), extension to nonlinear problems, or proposition of an adaptive operator T for self-adapting the method.

A Localization of the flux distance to the energy space

In extension of the discussion in Section 2.7, we can observe that

$$\|u - u_h\|_* \leq \min_{\boldsymbol{\sigma}\in\mathbf{H}(\mathrm{div},\Omega);\, \nabla\cdot\boldsymbol{\sigma}=f} \|\underline{\boldsymbol{\Sigma}}\nabla_\theta u_h + \boldsymbol{\sigma}\|$$

by the Green theorem, so that $\|u - u_h\|_*$ is linked to the *nonconformity* in the approximate *flux* $-\underline{\boldsymbol{\Sigma}}\nabla_\theta u_h$. We now present for this term a localization result like those of Section 3.2. Let $\mathbf{H}_*(\mathrm{div}, \omega_\mathbf{a})$ stand for $\mathbf{H}(\mathrm{div}, \omega_\mathbf{a})$ functions with zero normal trace in the appropriate sense on $\partial\omega_\mathbf{a}$ for $\mathbf{a} \in \mathcal{V}_h^{\mathrm{int}}$ and for $\mathbf{H}(\mathrm{div}, \omega_\mathbf{a})$ functions with zero normal trace in the appropriate sense on $\partial\omega_\mathbf{a} \setminus \partial\Omega$ for $\mathbf{a} \in \mathcal{V}_h^{\mathrm{ext}}$. One can show similarly as in Section 3.2, with the constant $C_{\mathrm{cont,PF}}$ of inequality (3.2) that:

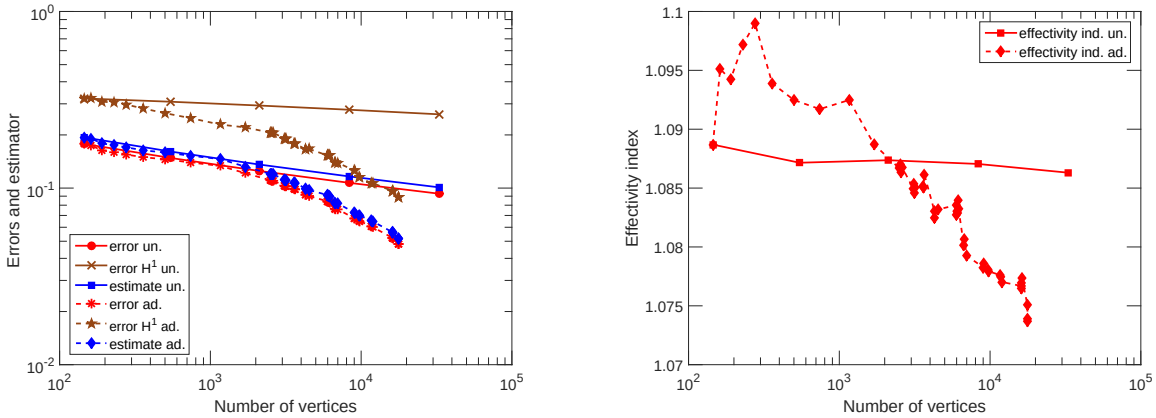

Figure 8: Estimates and errors for uniform and adaptive mesh refinement (left) and the corresponding effectivity indices (right), the singular case with $\sigma_- = -3.1$

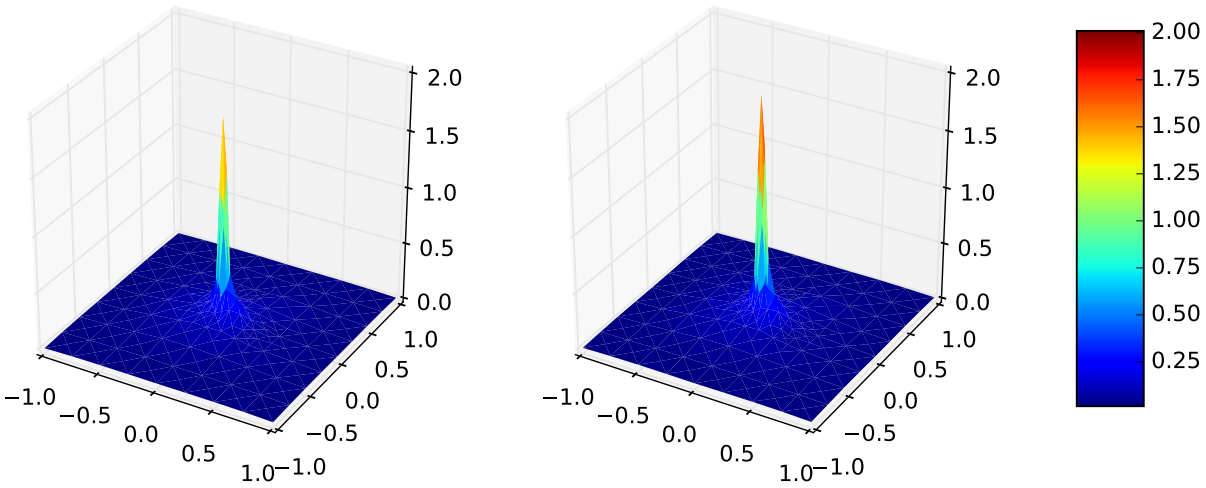

Figure 9: Exact (left) and estimated (right) error distribution on an adaptively refined mesh, the singular case with $\sigma_- = -3.1$

Theorem A.1 (Localization of the flux nonconformity evaluation)**.** *Let $u_h \in H^1(\mathcal{T}_h)$ satisfying Assump-*

tion 2.4 be arbitrary. Then

$$\min_{\boldsymbol{\sigma}\in\mathbf{H}(\mathrm{div},\Omega);\,\nabla\cdot\boldsymbol{\sigma}=f}\|\underline{\boldsymbol{\Sigma}}\nabla_\theta u_h+\boldsymbol{\sigma}\|^2\le(d+1)\sum_{\mathbf{a}\in\mathcal{V}_h}\min_{\boldsymbol{\sigma}^{\mathbf{a}}\in\mathbf{H}_*(\mathrm{div},\omega_{\mathbf{a}});\,\nabla\cdot\boldsymbol{\sigma}^{\mathbf{a}}=\psi_{\mathbf{a}}f-\underline{\boldsymbol{\Sigma}}\nabla_\theta u_h\cdot\nabla\psi_{\mathbf{a}}}\|\psi_{\mathbf{a}}\underline{\boldsymbol{\Sigma}}\nabla_\theta u_h+\boldsymbol{\sigma}^{\mathbf{a}}\|^2_{\omega_{\mathbf{a}}},$$

$$\sum_{\mathbf{a}\in\mathcal{V}_h}\min_{\boldsymbol{\sigma}^{\mathbf{a}}\in\mathbf{H}_*(\mathrm{div},\omega_{\mathbf{a}});\,\nabla\cdot\boldsymbol{\sigma}^{\mathbf{a}}=\psi_{\mathbf{a}}f-\underline{\boldsymbol{\Sigma}}\nabla_\theta u_h\cdot\nabla\psi_{\mathbf{a}}}\|\psi_{\mathbf{a}}\underline{\boldsymbol{\Sigma}}\nabla_\theta u_h+\boldsymbol{\sigma}^{\mathbf{a}}\|^2_{\omega_{\mathbf{a}}}\le(d+1)C^2_{\mathrm{cont,PF}}\min_{\boldsymbol{\sigma}\in\mathbf{H}(\mathrm{div},\Omega);\,\nabla\cdot\boldsymbol{\sigma}=f}\|\underline{\boldsymbol{\Sigma}}\nabla_\theta u_h+\boldsymbol{\sigma}\|^2.$$

Acknowledgement: The authors are very much grateful to Jan Blechta (Charles University in Prague) for kindly providing the numerical results.

References

[1] Ainsworth, M. Robust a posteriori error estimation for nonconforming finite element approximation. *SIAM J. Numer. Anal. 42*, 6 (2005), 2320–2341.

[2] Ainsworth, M. A framework for obtaining guaranteed error bounds for finite element approximations. *J. Comput. Appl. Math. 234*, 9 (2010), 2618–2632.

[3] Alnæs, M., Blechta, J., Hake, J., Johansson, A., Kehlet, B., Logg, A., Richardson, C., Ring, J., Rognes, M., and Wells, G. The FEniCS Project version 1.5. *Archive of Numerical Software 3*, 100 (2015).

[4] Babuška, I., and Miller, A. A feedback finite element method with a posteriori error estimation. I. The finite element method and some basic properties of the a posteriori error estimator. *Comput. Methods Appl. Mech. Engrg. 61*, 1 (1987), 1–40.

[5] Becker, R., Capatina, D., and Luce, R. Local flux reconstructions for standard finite element methods on triangular meshes. *SIAM J. Numer. Anal. 54*, 4 (2016), 2684–2706.

[6] Bernardi, C., and Verfürth, R. Adaptive finite element methods for elliptic equations with non-smooth coefficients. *Numer. Math. 85*, 4 (2000), 579–608.

[7] Blechta, J. Dolfin-tape, DOLFIN tools for a posteriori error estimation, version "paper-norms-nonlin-code-v1.0-rc3", 2016.

[8] Blechta, J., Málek, J., and Vohralík, M. Localization of the $W^{-1,q}$ norm for local a posteriori efficiency. HAL Preprint 01332481, submitted for publication, 2016.

[9] 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.

[10] 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.

[11] 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.

[12] Braess, D., and Schöberl, J. Equilibrated residual error estimator for edge elements. *Math. Comp. 77*, 262 (2008), 651–672.

[13] Brenner, S. C. Poincaré-Friedrichs inequalities for piecewise H^1 functions. *SIAM J. Numer. Anal. 41*, 1 (2003), 306–324.

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

[15] Cancès, E., Dusson, G., Maday, Y., Stamm, B., and Vohralík, M. Guaranteed and robust a posteriori bounds for Laplace eigenvalues and eigenvectors: a unified framework. HAL Preprint 01483461, submitted for publication, 2017.

[16] Carstensen, C., Eigel, M., Hoppe, R. H. W., and Löbhard, C. A review of unified a posteriori finite element error control. *Numer. Math. Theory Methods Appl. 5*, 4 (2012), 509–558.

[17] Carstensen, C., and Funken, S. A. Fully reliable localized error control in the FEM. *SIAM J. Sci. Comput. 21*, 4 (1999/00), 1465–1484.

[18] Carstensen, C., and Merdon, C. Computational survey on a posteriori error estimators for nonconforming finite element methods for the Poisson problem. *J. Comput. Appl. Math. 249* (2013), 74–94.

[19] Chaillou, A., and Suri, M. A posteriori estimation of the linearization error for strongly monotone nonlinear operators. *J. Comput. Appl. Math. 205*, 1 (2007), 72–87.

[20] 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.

[21] 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.

[22] Ciarlet, Jr., P., and Vohralík, M. Robust a posteriori error control for transmission problems with sign changing coefficients using localization of dual norms. HAL Preprint 01148476v1, 2015.

[23] Ciarlet, P. G. *The Finite Element Method for Elliptic Problems*, vol. 4 of *Studies in Mathematics and its Applications*. North-Holland, Amsterdam, 1978.

[24] Cohen, A., DeVore, R., and Nochetto, R. H. Convergence rates of AFEM with H^{-1} data. *Found. Comput. Math. 12*, 5 (2012), 671–718.

[25] Destuynder, P., and Métivet, B. Explicit error bounds for a nonconforming finite element method. *SIAM J. Numer. Anal. 35*, 5 (1998), 2099–2115.

[26] Destuynder, P., and Métivet, B. Explicit error bounds in a conforming finite element method. *Math. Comp. 68*, 228 (1999), 1379–1396.

[27] Di Pietro, D. A., and Ern, A. *Mathematical aspects of discontinuous Galerkin methods*, vol. 69 of *Mathématiques & Applications (Berlin) [Mathematics & Applications]*. Springer, Heidelberg, 2012.

[28] Dolejší, V., Ern, A., and Vohralík, M. *hp*-adaptation driven by polynomial-degree-robust a posteriori error estimates for elliptic problems. *SIAM J. Sci. Comput. 38*, 5 (2016), A3220–A3246.

[29] El Alaoui, 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.

[30] Ern, A., and Guermond, J.-L. *Theory and practice of finite elements*, vol. 159 of *Applied Mathematical Sciences*. Springer-Verlag, New York, 2004.

[31] Ern, A., and Vohralík, M. Adaptive inexact Newton methods with a posteriori stopping criteria for nonlinear diffusion PDEs. *SIAM J. Sci. Comput. 35*, 4 (2013), A1761–A1791.

[32] 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.

[33] Ern, A., and Vohralík, M. Stable broken H^1 and $\boldsymbol{H}(\mathrm{div})$ polynomial extensions for polynomial-degree-robust potential and flux reconstruction in three space dimensions. HAL Preprint 01422204, submitted for publication, 2016.

[34] 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.

[35] Kreuzer, C., and Süli, E. Adaptive finite element approximation of steady flows of incompressible fluids with implicit power-law-like rheology. *ESAIM Math. Model. Numer. Anal. 50*, 5 (2016), 1333–1369.

[36] Ladevèze, P., and Leguillon, D. Error estimate procedure in the finite element method and applications. *SIAM J. Numer. Anal. 20*, 3 (1983), 485–509.

[37] 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.

[38] Morin, P., Nochetto, R. H., and Siebert, K. G. Local problems on stars: a posteriori error estimators, convergence, and performance. *Math. Comp. 72*, 243 (2003), 1067–1097.

[39] 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.

[40] Nicaise, S., Witowski, K., and Wohlmuth, B. I. An a posteriori error estimator for the Lamé equation based on equilibrated fluxes. *IMA J. Numer. Anal. 28*, 2 (2008), 331–353.

[41] Payne, L. E., and Weinberger, H. F. An optimal Poincaré inequality for convex domains. *Arch. Rational Mech. Anal. 5* (1960), 286–292.

[42] Prager, W., and Synge, J. L. Approximations in elasticity based on the concept of function space. *Quart. Appl. Math. 5* (1947), 241–269.

[43] Repin, S. *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.

[44] 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.

[45] Veeser, A. Approximating gradients with continuous piecewise polynomial functions. *Found. Comput. Math. 16*, 3 (2016), 723–750.

[46] Veeser, A., and Verfürth, R. Explicit upper bounds for dual norms of residuals. *SIAM J. Numer. Anal. 47*, 3 (2009), 2387–2405.

[47] Veeser, A., and Verfürth, R. Poincaré constants for finite element stars. *IMA J. Numer. Anal. 32*, 1 (2012), 30–47.

[48] Verfürth, R. A posteriori error estimates for non-linear parabolic equations. Tech. report, Ruhr-Universität Bochum, 2004.

[49] Verfürth, R. Robust a posteriori error estimates for stationary convection-diffusion equations. *SIAM J. Numer. Anal. 43*, 4 (2005), 1766–1782.

[50] Verfürth, R. *A posteriori error estimation techniques for finite element methods*. Numerical Mathematics and Scientific Computation. Oxford University Press, Oxford, 2013.

[51] Vohralík, M. On the discrete Poincaré–Friedrichs inequalities for nonconforming approximations of the Sobolev space H^1. *Numer. Funct. Anal. Optim. 26*, 7-8 (2005), 925–952.

[52] 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.

[53] 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.