



MOX-Report No. 70/2021

## **Adaptive VEM: Stabilization-Free A Posteriori Error Analysis**

Beirao da Veiga, L.; Canuto, C.; Nochetto, R.H.; Vacca, G.;  
Verani, M.

MOX, Dipartimento di Matematica  
Politecnico di Milano, Via Bonardi 9 - 20133 Milano (Italy)

[mox-dmat@polimi.it](mailto:mox-dmat@polimi.it)

<http://mox.polimi.it>

# ADAPTIVE VEM: STABILIZATION-FREE A POSTERIORI ERROR ANALYSIS

L. BEIRÃO DA VEIGA\*, C. CANUTO †, R. H. NOCHETTO ‡, G. VACCA §, AND M. VERANI ¶

**Abstract.** In the present paper we initiate the challenging task of building a mathematically sound theory for Adaptive Virtual Element Methods (AVEMs). Among the realm of polygonal meshes, we restrict our analysis to triangular meshes with hanging nodes in 2d – the simplest meshes with a systematic refinement procedure that preserves shape regularity and optimal complexity. A major challenge in the a posteriori error analysis of AVEMs is the presence of the stabilization term, which is of the same order as the residual-type error estimator but prevents the equivalence of the latter with the energy error. Under the assumption that any chain of recursively created hanging nodes has uniformly bounded length, we show that the stabilization term can be made arbitrarily small relative to the error estimator provided the stabilization parameter of the scheme is sufficiently large. This quantitative estimate leads to stabilization-free upper and lower a posteriori bounds for the energy error. This novel and crucial property of VEMs hinges on the largest subspace of continuous piecewise linear functions and the delicate interplay between its coarser scales and the finer ones of the VEM space. Our results apply to  $H^1$ -conforming (lowest order) VEMs of any kind, including the classical and enhanced VEMs.

**Key words.** Virtual element method, non conforming meshes, a posteriori error analysis, stabilization

**AMS subject classifications.** 65N30, 65N50

**1. Introduction.** A posteriori error estimates have become over the last four decades an indispensable tool for realistic and intricate computations in both science and engineering. They are computable quantities in terms of the discrete solution and data that control the approximation error, typically in the energy norm  $|\cdot|_{1,\Omega}$ , from both above and below. Such estimators can be split into local contributions and exploited to drive adaptive algorithms that equidistribute the approximation error and so the computational effort. This has made simulation of complex phenomena accessible with modest computational resources.

Practice and theory of a posteriori error analysis and ensuing adaptive algorithms is a relatively mature research area for linear elliptic partial differential equations (PDEs), especially with the finite element method (FEM). They give rise to the so-called adaptive FEMs (or AFEMs for short). We refer to the survey papers [42, 43] for an account of the state-of-the-art on the following two fundamental and complementary aspects of this endeavor:

- *Derivation of a posteriori error estimates:* Residual estimators are the first and simplest estimators; see Babuška and Miller [4] and [5]. They exhibit upper and lower bounds (up to data oscillation) with stability constants of moderate size that depend on interpolation constants and thus on the geometry of the underlying meshes. Other estimators have been developed over the years with the goal of getting more precise or even constant free estimates; examples are local problems on elements [6] and stars [40], gradient recovery estimators [47, 44], and flux equilibration estimators [21, 20]. It turns out that they are all equivalent to the energy error. In addition, low order approximation [4, 5, 6] has evolved into high-order methods such as the  $hp$ -FEM [37].
- *Proof of convergence and optimality of AFEMs:* The study of adaptive loops of the form

$$(1.1) \quad \text{SOLVE} \longrightarrow \text{ESTIMATE} \longrightarrow \text{MARK} \longrightarrow \text{REFINE}$$

is an outstanding problem in numerical analysis of PDEs. The issue at stake is that discrete solutions at different level of resolution, typically on nested meshes, must be compared. This, in conjunction

---

\*Dipartimento di Matematica e Applicazioni, Università degli Studi di Milano Bicocca, Via Roberto Cozzi 55 - 20125 Milano, Italy ([lourenco.beirao@unimib.it](mailto:lourenco.beirao@unimib.it))

†Dipartimento di Scienze Matematiche G.L. Lagrange, Politecnico di Torino, Corso Duca degli Abruzzi 24 - 10129 Torino, Italy ([claudio.canuto@polito.it](mailto:claudio.canuto@polito.it))

‡Department of Mathematics and Institute for Physical Science and Technology, University of Maryland, College Park - 20742, MD, USA ([rh@math.umd.edu](mailto:rh@math.umd.edu))

§Dipartimento di Matematica, Università degli Studi di Bari, Via Edoardo Orabona 4 - 70125 Bari, Italy ([giuseppe.vacca@uniba.it](mailto:giuseppe.vacca@uniba.it))

¶MOX-Laboratory for Modeling and Scientific Computing, Dipartimento di Matematica, Politecnico di Milano, Piazza Leonardo da Vinci 32 - 20133 Milano, Italy ([marco.verani@polimi.it](mailto:marco.verani@polimi.it))

with the upper bound and Dörfler marking, yields a contraction property for every step of (1.1). Optimality entails further understanding of how the a posteriori estimator changes with the discrete solution and mesh refinement, as well as whether it can be localized to the refined region and yet provide control of the error between discrete solutions. This, combined with marking minimal sets and complexity estimates for mesh refinement strategies, leads to optimality of AFEM in the sense that the energy error decreases with optimal rate (up to a multiplicative constant) in terms of degrees of freedom. Theory for fixed polynomial degree [42, 43] extends somewhat to variable order [26, 27].

**Virtual element methods (VEMs).** They are a relatively new discretization paradigm which allows for general polytopal meshes, any polynomial degree, and yet conforming  $H^1$ -approximations for second order problems [10, 11]. This geometric flexibility is very useful in some applications (a few examples being [15, 29, 3, 17, 14]), but comes at a price for the design and practical use of adaptive VEMs (or AVEMs for short).

Two natural, but yet open, questions arise:

- *Procedure:* Is it possible to systematically refine general polytopes and preserve shape regularity? Beirão da Veiga and Manzini [8] proposed a first residual based error estimator and introduced a simple refinement rule for any convex polygon. The rule is to connect the barycenter of the polygon with mid-points of edges, where the word “edge” needs to be intended disregarding the existence of hanging nodes generated during the refinement procedure. It is not difficult to check that such procedure guarantees to generate a sequence of shape regular meshes. A refinement procedure that guarantees shape regularity under more general mesh assumptions is lacking. Note that shape regularity is critical to have robust interpolation estimates regardless of the resolution level.
- *Complexity:* Is it possible to prove that the number of elements generated by REFIN is proportional to the number of elements marked collectively for refinement by MARK? On the one hand, the answer is affirmative if the refinement is completely local. This in turn comes at the expense of unlimited growth of nodes per element, which may be hard to handle computationally and does not add enhanced accuracy. On the other hand, restricting the number of hanging nodes per edge makes the question very delicate, and generally false at every step of (1.1). This is altogether crucial to show that iteration of (1.1) leads to an error decay comparable with the best approximation in terms of degrees of freedom.

The development of a posteriori error estimates for VEMs mimics that of FEMs. The estimator of residual type most relevant to us is that proposed by Cangiani et al. [23]. The upper and lower bounds derived in [23] involve stabilization terms, but are valid for arbitrary polygonal elements, any polynomial degree, and general (coercive) second order operators with variable coefficients. Paper [16] by Berrone and Borio is a first attempt to remove the stabilization term from the a posteriori error estimator, although the proof hinges on certain properties that are only numerically checked. We achieve this goal in this paper with a fully rigorous but completely different approach from [16]. Estimators for the  $hp$ -version of VEMs are developed in [9], for anisotropic VEMs in [2], and for mixed VEMs in [25, 41]. Gradient recovery estimators are derived in [28] whereas those based on equilibrated fluxes are studied in [30, 31].

A key constituent of VEMs to deal with general polytopes is stabilization. Even though the role of stabilization is clear and precise in the a priori error analysis of VEMs to make the discrete bilinear form coercive, it remains elusive in the a posteriori counterpart. The main contribution of this paper is to show that such a role is not vital.

**Setting.** Our approach to adaptivity for VEMs is twofold. In this paper, we consider residual estimators and derive *stabilization-free* a posteriori upper and lower bounds. The removal of the stabilization term is essential to study (1.1), and thereby prove its convergence and optimal complexity in [13]. To achieve these goals, we put ourselves in the simplest possible but relevant setting consisting of the following four simplifying assumptions.

- *Meshes:* We consider partitions  $\mathcal{T}$  of a polygonal domain  $\Omega$  for  $d = 2$  made of elements  $E$ , which are triangles with hanging nodes and refined via the newest vertex bisection (NVB). In contrast to FEMs, the hanging nodes carry degrees of freedom in the VEM philosophy. The NVB dictates

a unique infinite binary tree with roots in the initial mesh, in which every triangle  $E$  is uniquely determined and traceable back to the roots. This geometric rigidity is crucial to optimal complexity (see [18, 43] for  $d = 2$  and [46, 42] for  $d > 2$ ), and plays an essential role in the study of AFEMs [42, 43] as well as in the sequel paper [13] on AVEMs. Quadrilateral partitions with hanging nodes are practical in the VEM context and amenable to analysis, but general polygonal elements are currently out of reach.

- *Polynomial degree:* We restrict our analysis to piecewise linear elements on the skeleton  $\mathcal{E}$  of  $\mathcal{T}$ . This is not just for convenience to simplify the presentation. It enters in the notion of global index (see Definition 2.1) and the scaled Poincaré inequality (see Proposition 3.1). They lead to the stabilization-free a posteriori error estimators discussed below. Extensions to higher polynomial degree are conceivable but not obvious.
- *Global index:* This is a natural number  $\lambda(\mathbf{x})$  that characterizes the level of a hanging node  $\mathbf{x}$  generated by successive NVB of an element  $E \in \mathcal{T}$ . We make the key assumption that, for all hanging nodes  $\mathbf{x}$  of all meshes  $\mathcal{T}$ , there exists a universal constant  $\Lambda > 0$  such that

$$(1.2) \quad \lambda(\mathbf{x}) \leq \Lambda.$$

This novel notion has profound geometric consequences. First, any chain of recursively created hanging nodes has uniformly bounded length, second a side of a triangle  $E$  can contain at most  $2^\Lambda - 1$  hanging nodes, and third any edge  $e$  of  $E$  has a size comparable with that of  $E$ , namely  $h_e \simeq h_E$ . These properties are instrumental to prove the scaled Poincaré inequality, but do not prevent deep refinement in the interior of  $E$ .

- *Data:* We consider  $\Omega$  to be polygonal and the symmetric elliptic PDE

$$(1.3) \quad -\nabla \cdot (A\nabla u) + cu = f \quad \text{in } \Omega,$$

with piecewise constant data  $\mathcal{D} = (A, c, f)$  and vanishing Dirichlet boundary condition. This choice simplifies the presentation and avoids approximation of  $\Omega$  and data oscillation terms. We extend the main estimates to variable data  $\mathcal{D}$  at the end of the paper. However, piecewise constant data play a fundamental role in the design of AVEMs in [13] because we approximate adaptively  $\mathcal{D}$  to a desired level of accuracy before we reduce the PDE error to a comparable level. Therefore, the analysis of [13] hinges on having  $\mathcal{D}$  piecewise constant when dealing with a posteriori error estimators for (1.3).

Our approach is a first attempt to develop mathematically sound AVEMs. This simplest setting serves to highlight similarities and striking differences with respect to AFEMs. We later embark on several extensions, such as variable data and quadrilaterals meshes and, but leave open the questions of general polygonal elements and higher order polynomial degree.

**Contributions.** We now describe our main contributions. Let  $\mathbb{V}_{\mathcal{T}}$  be a general  $H^1$ -conforming (lowest order) VEM space over  $\mathcal{T}$  (see for instance [10, 1]), which entails a suitable continuous extension of piecewise linear functions on the skeleton  $\mathcal{E}$  to  $\Omega$ . Let  $a_{\mathcal{T}}$  and  $m_{\mathcal{T}}$  be the VEM bilinear forms corresponding to the second-order and zeroth-order terms in (1.3), and let  $\mathcal{F}_{\mathcal{T}}$  be the linear form corresponding to the forcing term; we refer to Section 2.3 for details. If  $S_{\mathcal{T}}$  denotes the stabilization term, then the discrete solution  $u_{\mathcal{T}} \in \mathbb{V}_{\mathcal{T}}$  satisfies

$$(1.4) \quad a_{\mathcal{T}}(u_{\mathcal{T}}, v) + m_{\mathcal{T}}(u_{\mathcal{T}}, v) + \gamma S_{\mathcal{T}}(u_{\mathcal{T}}, v) = \mathcal{F}_{\mathcal{T}}(v) \quad \forall v \in \mathbb{V}_{\mathcal{T}}.$$

Problem (1.4) admits a unique solution  $u_{\mathcal{T}}$  for all values of the stabilization parameter  $\gamma > 0$ . Let  $\eta_{\mathcal{T}}(u_{\mathcal{T}}, \mathcal{D})$  be the residual a posteriori error estimator for piecewise constant data  $\mathcal{D}$  studied in Section 4. Our global a posteriori error estimates read as follows:

$$(1.5) \quad c_{\text{apost}} \eta_{\mathcal{T}}^2(u_{\mathcal{T}}, \mathcal{D}) - S_{\mathcal{T}}(u_{\mathcal{T}}, u_{\mathcal{T}}) \leq |u - u_{\mathcal{T}}|_{1,\Omega}^2 \leq C_{\text{apost}} (\eta_{\mathcal{T}}^2(u_{\mathcal{T}}, \mathcal{D}) + S_{\mathcal{T}}(u_{\mathcal{T}}, u_{\mathcal{T}})),$$

for suitable constants  $c_{\text{apost}} < C_{\text{apost}}$ , see Proposition 4.1 and Corollary 4.3. We stress that, in contrast to [24], the stabilization term  $S_{\mathcal{T}}$  appears without the constant  $\gamma$  in (1.5). Our main result is

Proposition 4.4: there exists a constant  $C_B > 0$  depending on  $\Lambda$  but independent of  $\mathcal{T}$ ,  $u_{\mathcal{T}}$  and  $\gamma$  such that

$$(1.6) \quad \gamma^2 S_{\mathcal{T}}(u_{\mathcal{T}}, u_{\mathcal{T}}) \leq C_B \eta_{\mathcal{T}}^2(u_{\mathcal{T}}, \mathcal{D}).$$

Computations reveal that (1.6) is sharp provided the number of hanging nodes is large relative to the total, and confirm that the stabilization term  $S_{\mathcal{T}}(u_{\mathcal{T}}, u_{\mathcal{T}})$  is of the same asymptotic order as the estimator  $\eta_{\mathcal{T}}^2(u_{\mathcal{T}}, \mathcal{D})$  and cannot be neglected. The significance of (1.6) is that it gives the quantitative condition  $\gamma^2 > C_B/c_{\text{apost}}$  on  $\gamma$  for  $S_{\mathcal{T}}(u_{\mathcal{T}}, u_{\mathcal{T}})$  to be absorbed within  $\eta_{\mathcal{T}}^2(u_{\mathcal{T}}, \mathcal{D})$  and, combined with (1.5), yields the stabilization-free a posteriori error estimates

$$(1.7) \quad \left(c_{\text{apost}} - \frac{C_B}{\gamma^2}\right) \eta_{\mathcal{T}}^2(u_{\mathcal{T}}, \mathcal{D}) \leq |u - u_{\mathcal{T}}|_{1,\Omega}^2 \leq C_{\text{apost}} \left(1 + \frac{C_B}{\gamma^2}\right) \eta_{\mathcal{T}}^2(u_{\mathcal{T}}, \mathcal{D}).$$

In contrast to a priori error estimates, this new estimate sheds light on the secondary role played by stabilization in a posteriori error analysis. Moreover, the relation between discrete solutions on different meshes involves the stabilization terms on each mesh, which complicates the theory of (1.1). Applying once again (1.6) in [13], we derive convergence and optimality of (1.1), where  $\gamma$  is to be chosen perhaps a bit larger than in (1.7).

We conclude this introduction with a heuristic explanation of the idea behind (1.6). It is inspired by the analysis of adaptive discontinuous Galerkin methods (dG) by Karakashian and Pascal [34, 35] and Bonito and Nochetto [19]. It turns out that to control the penalty term of dG, which is also of the same order as the estimator, a suitable estimate involving the penalty parameter  $\gamma$  similar to (1.6) is derived in [35] to prove convergence and is further exploited in [19] to show convergence under minimal regularity and optimality of (1.1). This hinges on the subspace  $\mathbb{V}_{\mathcal{T}}^0$  of all continuous, piecewise linear functions over  $\mathcal{T}$ . It turns out that the stabilization term  $S_{\mathcal{T}}$  vanishes on the subspace  $\mathbb{V}_{\mathcal{T}}^0$ , namely  $S_{\mathcal{T}}(w, v) = 0$  for all  $v \in \mathbb{V}_{\mathcal{T}}^0, w \in \mathbb{V}_{\mathcal{T}}$ . The delicate issue at stake is to relate the coarser scales of  $\mathbb{V}_{\mathcal{T}}^0$  with the finer ones of  $\mathbb{V}_{\mathcal{T}}$ , which is made possible by the restriction (1.2) on the global index  $\lambda$ . This leads to the following two fundamental estimates for VEMs.

The first key estimate is the following scaled Poincaré inequality proved in Section 5:

$$(1.8) \quad \sum_{E \in \mathcal{T}} h_E^{-2} \|v\|_{0,E}^2 \lesssim |v|_{1,\Omega}^2 \quad \forall v \in \mathbb{V}_{\mathcal{T}}$$

so that  $v$  vanishes at all nodes of  $\mathcal{T}^0$ , the so-called *proper nodes*. The second key estimate relates the interpolation errors in  $\mathbb{V}_{\mathcal{T}}^0$  and  $\mathbb{V}_{\mathcal{T}}$  due to the corresponding piecewise linear Lagrange interpolation operators  $\mathcal{I}_{\mathcal{T}}^0$  and  $\mathcal{I}_{\mathcal{T}}$ :

$$(1.9) \quad |v - \mathcal{I}_{\mathcal{T}}^0 v|_{1,\Omega} \lesssim |v - \mathcal{I}_{\mathcal{T}} v|_{1,\mathcal{T}} \quad \forall v \in \mathbb{V}_{\mathcal{T}};$$

note that in general  $\mathcal{I}_{\mathcal{T}} v$  is discontinuous in  $\Omega$  and  $|\cdot|_{1,\mathcal{T}}$  stands for the broken  $H^1$ -seminorm. This estimate is proved in Section 6 along with (1.6). The constants hidden in both (1.8) and (1.9) depend on  $\Lambda$  in (1.2) and blow-up as  $\Lambda \rightarrow \infty$ . This extends, upon suitably modifying the VEM, to rectangular elements but not to general polygons.

**Outline.** The paper is organized as follows. In Section 2 we introduce the bilinear forms associated with (1.3) and the VEM discretization with piecewise linear functions on the skeleton  $\mathcal{E}$ . We also discuss the notion of global index  $\lambda$  and the main restriction (1.2). In Section 3 we present some technical estimates such as (1.8). The a posteriori error analysis is carried out in Section 4. Inequality (1.8) is instrumental to derive (1.5) without the parameter  $\gamma$ , which combined with (1.6) yields (1.7) immediately. We also verify computationally that the dependence on  $\gamma$  in (1.6) is generically sharp. We postpone the proofs of (1.8) and (1.6) to Sections 5 and 6, respectively. In Section 7 we extend our results to variable coefficients. We conclude in Section 8 with an insightful numerical experiment on a highly singular problem that illustrates, in particular, the ability of VEM to capture the local solution structure.

**2. The problem and its discretization.** In a polygonal domain  $\Omega \subset \mathbb{R}^2$ , consider the second-order Dirichlet boundary-value problem

$$(2.1) \quad -\nabla \cdot (A \nabla u) + cu = f \quad \text{in } \Omega, \quad u = 0 \quad \text{on } \partial\Omega,$$

where  $A \in (L^\infty(\Omega))^{2 \times 2}$  is symmetric and uniformly positive-definite in  $\Omega$ ,  $c \in L^\infty(\Omega)$  is non-negative in  $\Omega$ , and  $f \in L^2(\Omega)$ . The variational formulation of the problem is

$$(2.2) \quad u \in \mathbb{V} : \mathcal{B}(u, v) = (f, v)_\Omega \quad \forall v \in \mathbb{V},$$

with  $\mathbb{V} := H_0^1(\Omega)$  and  $\mathcal{B}(u, v) := a(u, v) + m(u, v)$  where

$$a(u, v) := \int_\Omega (A \nabla u) \cdot \nabla v, \quad m(u, v) := \int_\Omega c u v$$

are the bilinear forms associated with Problem (2.1).

**2.1. Triangulations.** In view of the adaptive discretization of the problem, let us fix an initial conforming partition  $\mathcal{T}_0$  of  $\bar{\Omega}$  made of triangular elements. Let us denote by  $\mathcal{T}$  any refinement of  $\mathcal{T}_0$  obtained by a finite number of successive newest-vertex element bisections; the triangulation  $\mathcal{T}$  need not be conforming, since hanging nodes may be generated by the refinement. Let  $\mathcal{N}$  denote the set of nodes of  $\mathcal{T}$ , i.e., the collection of all vertices of the triangles in  $\mathcal{T}$ ; a node  $z \in \mathcal{N}$  is *proper* if it is a vertex of all triangles containing it; otherwise, it is a *hanging node*. Thus,  $\mathcal{N} = \mathcal{P} \cup \mathcal{H}$  is partitioned into the union of the set  $\mathcal{P}$  of proper nodes and the set  $\mathcal{H}$  of hanging nodes.

Given an element  $E \in \mathcal{T}$ , let  $\mathcal{N}_E$  be the set of nodes sitting on  $\partial E$ ; it contains the three vertices and, possibly, some hanging node. If the cardinality  $|\mathcal{N}_E| = 3$ ,  $E$  is said a *proper triangle* of  $\mathcal{T}$ ; if  $|\mathcal{N}_E| > 3$ , then according to the VEM philosophy  $E$  is not viewed as a triangle, but as a polygon having  $|\mathcal{N}_E|$  edges, some of which are placed consecutively on the same line. Any such edge  $e \subset \partial E$  is called an *interface* (with the neighboring element, or with the exterior of the domain); the set of all edges of  $E$  is denoted by  $\mathcal{E}_E$ . Note that if  $e \subset \partial E \cap \partial E'$ , then it is an edge for both elements; consequently, it is meaningful to define the *skeleton* of the triangulation  $\mathcal{T}$  by setting  $\mathcal{E} := \bigcup_{E \in \mathcal{T}} \mathcal{E}_E$ . Throughout the paper, we will set  $h_E = |E|^{1/2}$  for an element and  $h_e = |e|$  for an edge.

The concept of *global index* of a hanging node will be crucial in the sequel. To define it, let us first observe that any hanging node  $\mathbf{x} \in \mathcal{H}$  has been obtained through a newest-vertex bisection by halving an edge of a triangle in the preceding triangulation; denoting by  $\mathbf{x}', \mathbf{x}'' \in \mathcal{N}$  the endpoints of such edge, let us set  $\mathcal{B}(\mathbf{x}) = \{\mathbf{x}', \mathbf{x}''\}$ .

**DEFINITION 2.1** (Global index of a node). *The global index  $\lambda$  of a node  $\mathbf{x} \in \mathcal{N}$  is recursively defined as follows:*

- If  $\mathbf{x} \in \mathcal{P}$ , then set  $\lambda(\mathbf{x}) := 0$ ;
- If  $\mathbf{x} \in \mathcal{H}$ , with  $\mathbf{x}', \mathbf{x}'' \in \mathcal{B}(\mathbf{x})$ , then set  $\lambda(\mathbf{x}) := \max(\lambda(\mathbf{x}'), \lambda(\mathbf{x}'')) + 1$ .

We require that the largest global index in  $\mathcal{T}$ , defined as

$$\Lambda_{\mathcal{T}} := \max_{\mathbf{x} \in \mathcal{N}} \lambda(\mathbf{x}),$$

does not blow-up when we take successive refinements of the initial triangulation  $\mathcal{T}_0$ .

**Assumption 2.2** (Boundedness of the global index). There exists a constant  $\Lambda \geq 1$  such that for all considered triangulations  $\mathcal{T}$ , one has

$$\Lambda_{\mathcal{T}} \leq \Lambda.$$

**Remark 2.3.** This assumption forces the length of any chain of recursively created hanging nodes to be uniformly bounded. In addition, it has the following implications for each element  $E \in \mathcal{T}$ :

- i) If  $L \subset \partial E$  is one of the three sides of the triangle  $E$ , then  $L$  may contain at most  $2^\Lambda - 1$  hanging nodes; consequently,  $|\mathcal{N}_E| \leq 3 \cdot 2^\Lambda$ .
- ii) If  $e \subset \partial E$  is any edge, then  $h_e \simeq h_E$ , where the hidden constants only depend on the shape of the initial triangulation and possibly on  $\Lambda$ .

Fig. 1 displays examples that illustrate the dynamic change of  $\lambda(\mathbf{x})$  for a given  $\mathbf{x} \in \mathcal{N}$ .

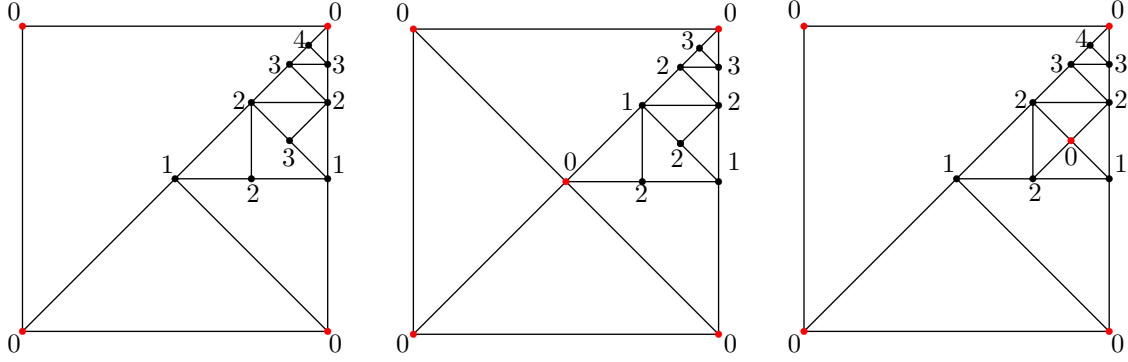


Fig. 1: Three examples of distributions of proper nodes (red) and hanging nodes (black), with associated global indices  $\lambda$ . The bisection added in the middle picture converts the centered node into proper, and induces nonlocal changes of global indices on chains associated with it. If  $\Lambda = 3$ , then the leftmost mesh is not admissible and this procedure is instrumental to restore admissibility. The right picture illustrates the creation of a proper node without nonlocal effects on global indices.

**2.2. VEM spaces and projectors.** In order to define a space of discrete functions in  $\Omega$  associated with  $\mathcal{T}$ , for each element  $E \in \mathcal{T}$  let us first introduce the space of continuous, piecewise affine functions on  $\partial E$

$$(2.3) \quad \mathbb{V}_{\partial E} := \{v \in \mathcal{C}^0(\partial E) : v|_e \in \mathbb{P}_1(e) \ \forall e \in \mathcal{E}_E\}.$$

Then, one needs to introduce a finite dimensional space  $\mathbb{V}_E \subset \mathcal{C}^0(E)$  satisfying the three following properties:

$$(2.4) \quad \dim \mathbb{V}_E = |\mathcal{N}_E|, \quad \mathbb{P}_1(E) \subseteq \mathbb{V}_E, \quad \tau_{\partial E}(\mathbb{V}_E) = \mathbb{V}_{\partial E},$$

where  $\tau_{\partial E}$  is the trace operator on the boundary of  $E$ . Obviously, if  $E$  is a proper triangle, then  $\mathbb{V}_E = \mathbb{P}_1(E)$  is the usual space of affine functions in  $E$ ; otherwise, note that a function in  $\mathbb{V}_E$  is uniquely identified by its trace on  $\partial E$ , but its value in the interior of  $E$  must be defined.

The results of the present paper apply to any generic VEM space satisfying the conditions above and a suitable stability property introduced below. The well known examples are the basic VEM space of [10]

$$(2.5) \quad \mathbb{V}_E := \{v \in H^1(E) : v|_{\partial E} \in \mathbb{V}_{\partial E}, \ \Delta v = 0\}$$

and the more advanced “enhanced” space from [1, 12]

$$(2.6) \quad \mathbb{V}_E := \left\{v \in H^1(E) : v|_{\partial E} \in \mathbb{V}_{\partial E}, \ \Delta v \in \mathbb{P}_1(E), \ \int_E (v - \Pi_E^\nabla v) q_1 = 0 \ \forall q_1 \in \mathbb{P}_1(E)\right\},$$

where the projector  $\Pi_E^\nabla : H^1(E) \rightarrow \mathbb{P}_1(E)$  is defined by the conditions

$$(2.7) \quad (\nabla(v - \Pi_E^\nabla v), \nabla w)_E = 0 \quad \forall w \in \mathbb{P}_1(E), \quad \int_{\partial E} (v - \Pi_E^\nabla v) = 0.$$

It is easy to check that the above spaces are well defined and satisfy conditions (2.4).

Once the local spaces  $\mathbb{V}_E$  are defined, we introduce the global discrete space

$$(2.8) \quad \mathbb{V}_{\mathcal{T}} := \{v \in \mathbb{V} : v|_E \in \mathbb{V}_E \ \forall E \in \mathcal{T}\}.$$

Note that functions in  $\mathbb{V}_{\mathcal{T}}$  are piecewise affine on the skeleton  $\mathcal{E}$ , and indeed they are uniquely determined by their values therein and are globally continuous. Introducing the spaces of piecewise polynomial functions on  $\mathcal{T}$

$$(2.9) \quad \mathbb{W}_{\mathcal{T}}^k := \{w \in L^2(\Omega) : w|_E \in \mathbb{P}_k(E) \quad \forall E \in \mathcal{T}\}, \quad k = 0, 1,$$

we also define the subspace of continuous, piecewise affine functions on  $\mathcal{T}$

$$(2.10) \quad \mathbb{V}_{\mathcal{T}}^0 := \mathbb{V}_{\mathcal{T}} \cap \mathbb{W}_{\mathcal{T}}^1,$$

which will play a key role in the sequel.

The discretization of Problem (2.1) will involve certain projection operators, that we are going to define locally and then globally. To this end, let  $\Pi_{\mathcal{T}}^{\nabla} : \mathbb{V}_{\mathcal{T}} \rightarrow \mathbb{W}_{\mathcal{T}}^1$  be the operator that restricts to  $\Pi_E^{\nabla}$  on each  $E \in \mathcal{T}$ . Similarly, let  $\mathcal{I}_E : \mathbb{V}_E \rightarrow \mathbb{P}_1(E)$  be the Lagrange interpolation operator at the vertices of  $E$ , and let  $\mathcal{I}_{\mathcal{T}} : \mathbb{V}_{\mathcal{T}} \rightarrow \mathbb{W}_{\mathcal{T}}^1$  be the Lagrange interpolation operator that restricts to  $\mathcal{I}_E$  on each  $E \in \mathcal{T}$ . Note that  $\Pi_{\mathcal{T}}^{\nabla} v = v$  and  $\mathcal{I}_{\mathcal{T}} v = v$  for all  $v \in \mathbb{V}_{\mathcal{T}}^0$ . Finally, let  $\Pi_E^0 : L^2(E) \rightarrow \mathbb{P}_1(E)$ , resp.  $\Pi_{\mathcal{T}}^0 : L^2(\Omega) \rightarrow \mathbb{W}_{\mathcal{T}}^1$ , be the local, resp. global,  $L^2$ -orthogonal projection operator.

Using an integration by parts, it is easy to check that the  $\Pi_E^{\nabla}$  operator is directly computable from the boundary values of  $v \in \mathbb{V}_E$ , and the same clearly holds for  $\mathcal{I}_E$ . On the contrary, on a general VEM space the operator  $\Pi_E^0$  may be not computable. A notable exception is given by the space (2.6), since by definition of the space it easily follows the following property:

$$(2.11) \quad \text{For the local space choice (2.6) the operators } \Pi_E^0 \text{ and } \Pi_E^{\nabla} \text{ coincide.}$$

**2.3. The discrete problem.** Next, we introduce the discrete bilinear forms to be used in a Galerkin discretization of our problem. Here we make a simplifying assumption on the coefficients of the equation, in order to arrive at the core of our contribution without too much technical burden. In Sect. 7 we will discuss the general situation.

*Assumption 2.4* (Coefficients and right-hand side of the equation). The coefficients  $A$  and  $c$  and the right-hand side  $f$  in (2.1) are constant in each element of  $\mathcal{T}$ .

Under this assumption, define  $a_{\mathcal{T}}, m_{\mathcal{T}} : \mathbb{V}_{\mathcal{T}} \times \mathbb{V}_{\mathcal{T}} \rightarrow \mathbb{R}$  by

$$(2.12) \quad \begin{aligned} a_{\mathcal{T}}(v, w) &:= \sum_{E \in \mathcal{T}} \int_E (A_E \nabla \Pi_E^{\nabla} v) \cdot \nabla \Pi_E^{\nabla} w =: \sum_{E \in \mathcal{T}} a_E(v, w), \\ m_{\mathcal{T}}(v, w) &:= \sum_{E \in \mathcal{T}} c_E \int_E \Pi_E^{\nabla} v \Pi_E^{\nabla} w =: \sum_{E \in \mathcal{T}} m_E(v, w) \end{aligned}$$

with  $A_E = A|_E \in \mathbb{R}^{2 \times 2}$  and  $c_E = c|_E \in \mathbb{R}$ .

Next, for any  $E \in \mathcal{T}$ , we introduce the symmetric bilinear form  $s_E : \mathbb{V}_E \times \mathbb{V}_E \rightarrow \mathbb{R}$

$$(2.13) \quad s_E(v, w) = \sum_{i=1}^{\mathcal{N}_E} v(\mathbf{x}_i) w(\mathbf{x}_i),$$

with  $\{\mathbf{x}_i\}_{i=1}^{\mathcal{N}_E}$  denoting the vertexes of  $E$ . Such form will take the role of a stabilization in the numerical method; other choices for the stabilizing form are available in the literature and the results presented here easily extend to such cases. We assume the following condition, stating that the local virtual spaces  $\mathbb{V}_E$  constitute a “stable lifting” of the element boundary values:

$$(2.14) \quad c_s |v|_{1,E}^2 \leq s_E(v, v) \leq C_s |v|_{1,E}^2 \quad \forall v \in \mathbb{V}_E / \mathbb{R},$$

for constants  $C_s \geq c_s > 0$  independent of  $E$ . For a proof of (2.14) for some typical choices of  $\mathbb{V}_E$  and  $s_E$  we refer to [7, 22]; in particular, the result holds for the choices (2.5) or (2.6), and (2.13). With the local form  $s_E$  at hand, we define the local stabilizing form

$$(2.15) \quad S_E(v, w) := s_E(v - \mathcal{I}_E v, w - \mathcal{I}_E w) \quad \forall v, w \in \mathbb{V}_E,$$

as well as the global stabilizing form

$$(2.16) \quad S_{\mathcal{T}}(v, w) := \sum_{E \in \mathcal{T}} S_E(v, w) \quad \forall v, w \in \mathbb{V}_{\mathcal{T}}.$$

Note that from (2.14) we obtain

$$(2.17) \quad S_{\mathcal{T}}(v, v) \simeq |v - \mathcal{I}_{\mathcal{T}}v|_{1, \mathcal{T}}^2 \quad \forall v \in \mathbb{V}_{\mathcal{T}},$$

where  $|\cdot|_{1, \mathcal{T}}$  denotes the broken  $H^1$ -seminorm over the mesh  $\mathcal{T}$ .

Finally, for all  $v, w \in \mathbb{V}_{\mathcal{T}}$  we define the complete bilinear form

$$(2.18) \quad \mathcal{B}_{\mathcal{T}} : \mathbb{V}_{\mathcal{T}} \times \mathbb{V}_{\mathcal{T}} \rightarrow \mathbb{R}, \quad \mathcal{B}_{\mathcal{T}}(v, w) := a_{\mathcal{T}}(v, w) + \gamma S_{\mathcal{T}}(v, w) + m_{\mathcal{T}}(v, w),$$

where  $\gamma \geq \gamma_0$  for some fixed  $\gamma_0 > 0$  is a stabilization constant independent of  $\mathcal{T}$ , which will be chosen later on.

The following properties are an easy consequence of the definitions and bounds outlined above.

LEMMA 2.5 (Properties of bilinear forms). *i) For any  $v \in \mathbb{V}_{\mathcal{T}}$  and any  $w \in \mathbb{V}_{\mathcal{T}}^0$ , it holds*

$$(2.19) \quad a_{\mathcal{T}}(v, w) = a(v, w), \quad S_{\mathcal{T}}(v, w) = 0.$$

*ii) The form  $\mathcal{B}_{\mathcal{T}}$  satisfies*

$$(2.20) \quad \beta |v|_{1, \Omega}^2 \leq \mathcal{B}_{\mathcal{T}}(v, v), \quad |\mathcal{B}_{\mathcal{T}}(v, w)| \leq B |v|_{1, \Omega} |w|_{1, \Omega}, \quad \forall v, w \in \mathbb{V}_{\mathcal{T}},$$

with continuity and coercivity constants  $B \geq \beta > 0$  independent of the triangulation  $\mathcal{T}$ . The constant  $\beta$  may be chosen independent of the parameter  $\gamma$ .

*Proof.* The first condition in Property (i) follows easily recalling Assumption 2.4 and noting that  $\nabla \Pi_E^{\nabla} v$  corresponds to the  $L^2(E)$  projection of  $\nabla v$  on the constant vector fields (living on  $E$ ). Property (ii) follows from (2.17) with trivial arguments.  $\square$

Regarding the approximation of the loading term, we here consider

$$(2.21) \quad \mathcal{F}_{\mathcal{T}}(v) := \sum_{E \in \mathcal{T}} \int_E f \Pi_E^{\nabla} v \quad \forall v \in H_0^1(\Omega).$$

We now have all the ingredients to set the Galerkin discretization of Problem (2.1): find  $u_{\mathcal{T}} \in \mathbb{V}_{\mathcal{T}}$  such that

$$(2.22) \quad \mathcal{B}_{\mathcal{T}}(u_{\mathcal{T}}, v) = \mathcal{F}_{\mathcal{T}}(v) \quad \forall v \in \mathbb{V}_{\mathcal{T}}.$$

Lemma 2.5 guarantees existence, uniqueness and stability of the Galerkin solution. We now establish a useful version of Galerkin orthogonality.

LEMMA 2.6 (Quasi-Galerkin orthogonality). *The solutions  $u$  of (2.2) and  $u_{\mathcal{T}}$  of (2.22) satisfy*

$$(2.23) \quad \mathcal{B}(u - u_{\mathcal{T}}, v) = \sum_{E \in \mathcal{T}} c_E \int_E (\Pi_{\mathcal{T}}^{\nabla} u_{\mathcal{T}} - u_{\mathcal{T}}) v \quad \forall v \in \mathbb{V}_{\mathcal{T}}^0.$$

In particular, the choice (2.6) of enhanced VEM space further implies

$$(2.24) \quad \mathcal{B}(u - u_{\mathcal{T}}, v) = 0 \quad \forall v \in \mathbb{V}_{\mathcal{T}}^0.$$

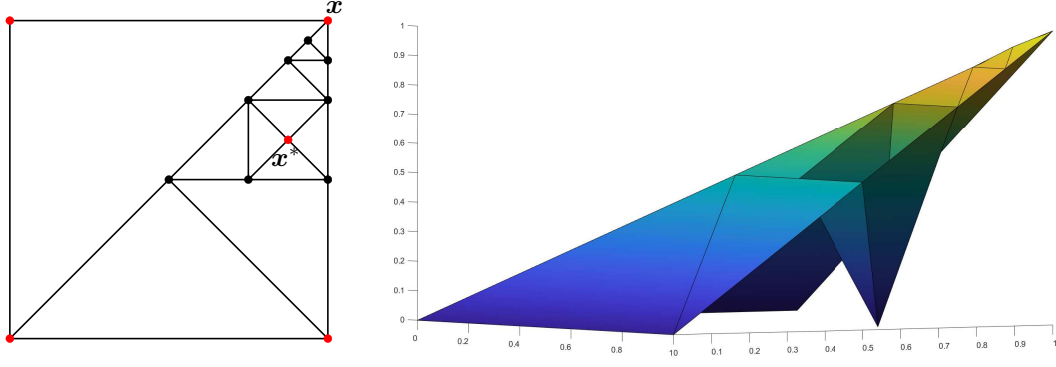


Fig. 2: Left: detail of a mesh  $\mathcal{T}$ , in which red nodes  $\mathbf{x}$  and  $\mathbf{x}^*$  are proper nodes. Right: basis function  $\psi_{\mathbf{x}} \in \mathbb{V}_{\mathcal{T}}^0$ ; notice that  $\psi_{\mathbf{x}}(\mathbf{x}^*) = 0$  and the basis function  $\psi_{\mathbf{x}^*} \in \mathbb{V}_{\mathcal{T}}^0$  is the usual hat function supported in the square centered at  $\mathbf{x}^*$  (not depicted).

*Proof.* The definitions (2.2) and (2.22) imply

$$\mathcal{B}(u - u_{\mathcal{T}}, v) = ((f, v)_{\Omega} - \mathcal{F}_{\mathcal{T}}(v)) + (\mathcal{B}_{\mathcal{T}}(u_{\mathcal{T}}, v) - \mathcal{B}(u_{\mathcal{T}}, v)) \quad \forall v \in \mathbb{V}_{\mathcal{T}}.$$

If  $v \in \mathbb{V}_{\mathcal{T}}^0$ , then  $\Pi_{\mathcal{T}}^{\nabla} v = v$  and  $\mathcal{F}_{\mathcal{T}}(v) = (f, v)_{\Omega}$  according to (2.21). On the other hand, (2.19) yields

$$\mathcal{B}_{\mathcal{T}}(u_{\mathcal{T}}, v) - \mathcal{B}(u_{\mathcal{T}}, v) = m_{\mathcal{T}}(u_{\mathcal{T}}, v) - m(u_{\mathcal{T}}, v) = \sum_{E \in \mathcal{T}} c_E \int_E (\Pi_{\mathcal{T}}^{\nabla} u_{\mathcal{T}} - u_{\mathcal{T}}) v \quad \forall v \in \mathbb{V}_{\mathcal{T}}^0,$$

which in turn leads to (2.23). Finally, for the choice (2.6) the right-hand side of (2.23) vanishes because  $v \in \mathbb{P}_1(E)$  for all  $E \in \mathcal{T}$ . This completes the proof.  $\square$

**3. Preparatory results.** In preparation for the subsequent a posteriori error analysis, we collect here some useful results involving functions in  $\mathbb{V}_{\mathcal{T}}$  or in  $\mathbb{V}_{\mathcal{T}}^0$ .

The first result is a scaled Poincaré inequality in  $\mathbb{V}_{\mathcal{T}}$ , which will be crucial in the sequel. We recall that  $\mathcal{P}$  denotes the set of proper nodes in  $\mathcal{T}$ .

**PROPOSITION 3.1** (Scaled Poincaré inequality in  $\mathbb{V}_{\mathcal{T}}$ ). *There exists a constant  $C_{\Lambda} > 0$  depending on  $\Lambda$  but independent of  $\mathcal{T}$ , such that*

$$(3.1) \quad \sum_{E \in \mathcal{T}} h_E^{-2} \|v\|_{0,E}^2 \leq C_{\Lambda} |v|_{1,\Omega}^2 \quad \forall v \in \mathbb{V}_{\mathcal{T}} \text{ such that } v(\mathbf{x}) = 0 \ \forall \mathbf{x} \in \mathcal{P}.$$

Due to the technical nature of the proof, we postpone it to Sect. 5.

Next, we go back to the space  $\mathbb{V}_{\mathcal{T}}^0$  introduced in (2.10). We note that functions in this space are uniquely determined by their values at the proper nodes of  $\mathcal{T}$ . Indeed, a function  $v \in \mathbb{V}_{\mathcal{T}}^0$  is affine in each element of  $\mathcal{T}$ , hence, it is uniquely determined by its values at the three vertices of the element: if the vertex  $\mathbf{x}$  is a hanging node, with  $\mathcal{B}(\mathbf{x}) = \{\mathbf{x}', \mathbf{x}''\}$ , then  $v(\mathbf{x}) = \frac{1}{2}(v(\mathbf{x}') + v(\mathbf{x}''))$ .

In particular,  $\mathbb{V}_{\mathcal{T}}^0$  is spanned by the Lagrange basis

$$(3.2) \quad \forall \mathbf{x} \in \mathcal{P} : \quad \psi_{\mathbf{x}} \in \mathbb{V}_{\mathcal{T}}^0 \quad \text{satisfies} \quad \psi_{\mathbf{x}}(\mathbf{z}) = \begin{cases} 1 & \text{if } \mathbf{z} = \mathbf{x}, \\ 0 & \text{if } \mathbf{z} \in \mathcal{P} \setminus \{\mathbf{x}\} \end{cases}$$

(see Fig. 2 for an example of such a basis function, which looks different from the standard pyramidal basis functions on conforming meshes). Thus, it is natural to introduce the operator

$$\mathcal{I}_{\mathcal{T}}^0 : \mathbb{V}_{\mathcal{T}} \rightarrow \mathbb{V}_{\mathcal{T}}^0$$

defined as the Lagrange interpolation operator at the nodes in  $\mathcal{P}$ .

We will also need some Clément quasi-interpolation operators. Precisely, let us denote by  $\tilde{\mathcal{I}}_{\mathcal{T}}^0 : \mathbb{V} \rightarrow \mathbb{V}_{\mathcal{T}}^0$  the classical Clément operator on the finite-element space  $\mathbb{V}_{\mathcal{T}}^0$ ; similarly, let  $\tilde{\mathcal{I}}_{\mathcal{T}} : \mathbb{V} \rightarrow \mathbb{V}_{\mathcal{T}}$  be the Clément operator on the virtual-element space  $\mathbb{V}_{\mathcal{T}}$ , as defined in [38].

LEMMA 3.2 (Clement interpolation estimate). *The following inequality holds*

$$(3.3) \quad \sum_{E \in \mathcal{T}} h_E^{-2} \|v - \tilde{\mathcal{I}}_{\mathcal{T}}^0 v\|_{0,E}^2 \lesssim |v|_{1,\Omega}^2 \quad \forall v \in \mathbb{V},$$

where the hidden constant depends on the maximal index  $\Lambda$  but not on  $\mathcal{T}$ .

*Proof.* Let  $v_{\mathcal{T}} = \tilde{\mathcal{I}}_{\mathcal{T}} v \in \mathbb{V}_{\mathcal{T}}$ . Since

$$v - \tilde{\mathcal{I}}_{\mathcal{T}}^0 v = (v - v_{\mathcal{T}}) + (v_{\mathcal{T}} - \tilde{\mathcal{I}}_{\mathcal{T}}^0 v_{\mathcal{T}}) + (\tilde{\mathcal{I}}_{\mathcal{T}}^0 v_{\mathcal{T}} - \tilde{\mathcal{I}}_{\mathcal{T}}^0 v)$$

and  $\tilde{\mathcal{I}}_{\mathcal{T}}^0$  is locally stable in  $L^2$ , we deduce

$$\begin{aligned} \sum_{E \in \mathcal{T}} h_E^{-2} \|v - \tilde{\mathcal{I}}_{\mathcal{T}}^0 v\|_{0,E}^2 &\lesssim \sum_{E \in \mathcal{T}} h_E^{-2} \|v - v_{\mathcal{T}}\|_{0,E}^2 + h_E^{-2} \|v_{\mathcal{T}} - \tilde{\mathcal{I}}_{\mathcal{T}}^0 v_{\mathcal{T}}\|_{0,E}^2 \\ &\lesssim |v|_{1,\Omega}^2 + \sum_{E \in \mathcal{T}} h_E^{-2} \|v_{\mathcal{T}} - \tilde{\mathcal{I}}_{\mathcal{T}}^0 v_{\mathcal{T}}\|_{0,E}^2. \end{aligned}$$

Thus, we need to prove

$$\sum_{E \in \mathcal{T}} h_E^{-2} \|v_{\mathcal{T}} - \tilde{\mathcal{I}}_{\mathcal{T}}^0 v_{\mathcal{T}}\|_{0,E}^2 \lesssim |v_{\mathcal{T}}|_{1,\Omega}^2.$$

To show this bound, write

$$v_{\mathcal{T}} - \tilde{\mathcal{I}}_{\mathcal{T}}^0 v_{\mathcal{T}} = (v_{\mathcal{T}} - \mathcal{I}_{\mathcal{T}}^0 v_{\mathcal{T}}) + \tilde{\mathcal{I}}_{\mathcal{T}}^0 (\mathcal{I}_{\mathcal{T}}^0 v_{\mathcal{T}} - v_{\mathcal{T}})$$

because  $\tilde{\mathcal{I}}_{\mathcal{T}}^0$  is invariant in  $\mathbb{V}_{\mathcal{T}}^0$ , i.e.  $\mathcal{I}_{\mathcal{T}}^0 v_{\mathcal{T}} = \tilde{\mathcal{I}}_{\mathcal{T}}^0 (\mathcal{I}_{\mathcal{T}}^0 v_{\mathcal{T}})$ . We finally use again the stability of  $\tilde{\mathcal{I}}_{\mathcal{T}}^0$  in  $L^2$  together with Proposition 3.1 to obtain

$$\sum_{E \in \mathcal{T}} h_E^{-2} \|v_{\mathcal{T}} - \tilde{\mathcal{I}}_{\mathcal{T}}^0 v_{\mathcal{T}}\|_{0,E}^2 \lesssim \sum_{E \in \mathcal{T}} h_E^{-2} \|v_{\mathcal{T}} - \mathcal{I}_{\mathcal{T}}^0 v_{\mathcal{T}}\|_{0,E}^2 \lesssim C_{\Lambda} |v_{\mathcal{T}}|_{1,\Omega}^2.$$

This concludes the proof.  $\square$

**4. A posteriori error analysis.** Since we are interested in building adaptive discretizations, we rely on a posteriori error control. Following [24], we first introduce a residual-type a posteriori estimator. To this end, recalling that  $\mathcal{D} = (A, c, f)$  denotes the set of piecewise constant data, for any  $v \in \mathbb{V}_{\mathcal{T}}$  and any element  $E$  let us define the internal residual over  $E$

$$(4.1) \quad r_{\mathcal{T}}(E; v, \mathcal{D}) := f_E - c_E \Pi_E^{\nabla} v.$$

Similarly, for any two elements  $E_1, E_2 \in \mathcal{T}$  sharing an edge  $e \in \mathcal{E}_{E_1} \cap \mathcal{E}_{E_2}$ , let us define the jump residual over  $e$

$$(4.2) \quad j_{\mathcal{T}}(e; v, \mathcal{D}) := \llbracket A_E \nabla \Pi_{\mathcal{T}}^{\nabla} v \rrbracket_e = (A_{E_1} \nabla \Pi_{E_1}^{\nabla} v|_{E_1}) \cdot \mathbf{n}_1 + (A_{E_2} \nabla \Pi_{E_2}^{\nabla} v|_{E_2}) \cdot \mathbf{n}_2,$$

where  $\mathbf{n}_i$  denotes the unit normal vector to  $e$  pointing outward with respect to  $E_i$ ; set  $j_{\mathcal{T}}(e; v) = 0$  if  $e \subset \partial\Omega$ . Then, taking into account Remark 2.3, we define the local residual estimator associated with  $E$

$$(4.3) \quad \eta_{\mathcal{T}}^2(E; v, \mathcal{D}) := h_E^2 \|r_{\mathcal{T}}(E; v, \mathcal{D})\|_{0,E}^2 + \frac{1}{2} \sum_{e \in \mathcal{E}_E} h_E \|j_{\mathcal{T}}(e; v, \mathcal{D})\|_{0,e}^2,$$

as well as the global residual estimator

$$(4.4) \quad \eta_{\mathcal{T}}^2(v, \mathcal{D}) := \sum_{E \in \mathcal{T}} \eta_{\mathcal{T}}^2(E; v, \mathcal{D}).$$

An upper bound of the energy error is provided by the following result. The proof follows [24, Theorem 13], with the remarkable technical difference that the stabilization term  $S_{\mathcal{T}}(u_{\mathcal{T}}, u_{\mathcal{T}})$  is not scaled by the constant  $\gamma$  in (4.5).

**PROPOSITION 4.1** (Upper bound). *There exists a constant  $C_{\text{apost}} > 0$  depending on  $\Lambda$  and  $\mathcal{D}$  but independent of  $u$ ,  $\mathcal{T}$ ,  $u_{\mathcal{T}}$  and  $\gamma$ , such that*

$$(4.5) \quad |u - u_{\mathcal{T}}|_{1,\Omega}^2 \leq C_{\text{apost}} (\eta_{\mathcal{T}}^2(u_{\mathcal{T}}, \mathcal{D}) + S_{\mathcal{T}}(u_{\mathcal{T}}, u_{\mathcal{T}})).$$

*Proof.* We let  $v \in H_0^1(\Omega)$  and proceed as in [24, Theorem 13] to write

$$\mathcal{B}(u - u_{\mathcal{T}}, v) = ((f, v - v_{\mathcal{T}})_{\Omega} - \mathcal{B}(u_{\mathcal{T}}, v - v_{\mathcal{T}})) + \mathcal{B}(u - u_{\mathcal{T}}, v_{\mathcal{T}}) =: I + II,$$

except that we choose  $v_{\mathcal{T}} = \tilde{\mathcal{I}}_{\mathcal{T}}^0 v \in \mathbb{V}_{\mathcal{T}}^0$ , where  $\tilde{\mathcal{I}}_{\mathcal{T}}^0$  is the Clément quasi-interpolation operator on  $\mathbb{V}_{\mathcal{T}}^0$ . This choice has a remarkable impact on (4.5) compared with [24, Theorem 13].

We start by estimating the term  $I$ : one has

$$\begin{aligned} I &= \sum_{E \in \mathcal{T}} \{ (f, v - v_{\mathcal{T}})_E - (A_E \nabla \Pi_E^{\nabla} u_{\mathcal{T}}, \nabla(v - v_{\mathcal{T}}))_E - c_E (\Pi_E^{\nabla} u_{\mathcal{T}}, v - v_{\mathcal{T}})_E \} \\ &\quad + \sum_{E \in \mathcal{T}} \{ (A_E \nabla (\Pi_E^{\nabla} u_{\mathcal{T}} - u_{\mathcal{T}}), \nabla(v - v_{\mathcal{T}}))_E + c_E ((\Pi_E^{\nabla} u_{\mathcal{T}} - u_{\mathcal{T}}), v - v_{\mathcal{T}})_E \} =: I_1 + I_2. \end{aligned}$$

Integrating by parts, employing Lemma 3.2 and proceeding as in [24], we get

$$(4.6) \quad |I_1| \lesssim \eta_{\mathcal{T}}(u_{\mathcal{T}}, \mathcal{D}) |v|_{1,\Omega}$$

$$(4.7) \quad |I_2| \lesssim \left( \sum_{E \in \mathcal{T}} \|\nabla(\Pi_E^{\nabla} u_{\mathcal{T}} - u_{\mathcal{T}})\|_{0,E}^2 + h_E \|\Pi_E^{\nabla} u_{\mathcal{T}} - u_{\mathcal{T}}\|_{0,E}^2 \right)^{1/2} |v|_{1,\Omega} \lesssim S_{\mathcal{T}}(u_{\mathcal{T}}, u_{\mathcal{T}})^{1/2} |v|_{1,\Omega}.$$

We now deal with term  $II$ . We first apply Lemma 2.6 to obtain

$$\mathcal{B}(u - u_{\mathcal{T}}, v_{\mathcal{T}}) = \sum_{E \in \mathcal{T}} c_E \int_E (\Pi_E^{\nabla} u_{\mathcal{T}} - u_{\mathcal{T}}) v_{\mathcal{T}},$$

because  $v_{\mathcal{T}} \in \mathbb{V}_{\mathcal{T}}^0$ ; this is the key difference with [24, Theorem 13]. We next recall the definition (2.7) of  $\Pi_E^{\nabla}$ , and corresponding scaled Poincaré inequality  $\|\Pi_E^{\nabla} u_{\mathcal{T}} - u_{\mathcal{T}}\|_{0,E} \lesssim h_E |\Pi_E^{\nabla} u_{\mathcal{T}} - u_{\mathcal{T}}|_{1,E}$  for all  $E \in \mathcal{T}$ , to arrive at

$$\mathcal{B}(u - u_{\mathcal{T}}, v_{\mathcal{T}}) \lesssim h_E S_{\mathcal{T}}(u_{\mathcal{T}}, u_{\mathcal{T}})^{1/2} |v_{\mathcal{T}}|_{1,\Omega}.$$

Finally, taking  $v = u - u_{\mathcal{T}} \in H_0^1(\Omega)$ , employing the coercivity of  $\mathcal{B}(\cdot, \cdot)$  and combining the above estimates yield the assertion.  $\square$

We state the following result, which is proven in [24] for the choice (2.6) but it holds with the same proof for any other admissible choice of  $\mathbb{V}_E$ .

**PROPOSITION 4.2** (Local lower bound). *There holds*

$$(4.8) \quad \eta_{\mathcal{T}}^2(E; u_{\mathcal{T}}, \mathcal{D}) \lesssim \sum_{E' \in \omega_E} (|u - u_{\mathcal{T}}|_{1,E'}^2 + S_{E'}(u_{\mathcal{T}}, u_{\mathcal{T}}))$$

where  $\omega_E := \{E' : |\partial E \cap \partial E'| \neq \emptyset\}$ . The hidden constant is independent of  $\gamma, h, u$  and  $u_{\mathcal{T}}$ .

COROLLARY 4.3 (Global lower bound). *There exists a constant  $c_{\text{apost}} > 0$ , depending on  $\Lambda$  but independent of  $u$ ,  $\mathcal{T}$ ,  $u_{\mathcal{T}}$  and  $\gamma$ , such that*

$$(4.9) \quad c_{\text{apost}} \eta_{\mathcal{T}}^2(u_{\mathcal{T}}, \mathcal{D}) \leq |u - u_{\mathcal{T}}|_{1,\Omega}^2 + S_{\mathcal{T}}(u_{\mathcal{T}}, u_{\mathcal{T}}).$$

We now state the main result contained in this paper. Due to the technical nature of the proof, we postpone it to Sect. 6.

PROPOSITION 4.4 (Bound of the stabilization term by the residual). *There exists a constant  $C_B > 0$  depending on  $\Lambda$  but independent of  $\mathcal{T}$ ,  $u_{\mathcal{T}}$  and  $\gamma$  such that*

$$(4.10) \quad \gamma^2 S_{\mathcal{T}}(u_{\mathcal{T}}, u_{\mathcal{T}}) \leq C_B \eta_{\mathcal{T}}^2(u_{\mathcal{T}}, \mathcal{D}).$$

Combining Proposition 4.1, Corollary 4.3 and Proposition 4.4, we get the following stabilization-free (global) upper and lower bounds.

COROLLARY 4.5 (Stabilization-free a posteriori error estimates). *Assume that the parameter  $\gamma$  is chosen to satisfy  $\gamma^2 > \frac{C_B}{c_{\text{apost}}}$ . Then it holds*

$$(4.11) \quad (c_{\text{apost}} - C_B \gamma^{-2}) \eta_{\mathcal{T}}^2(u_{\mathcal{T}}, \mathcal{D}) \leq |u - u_{\mathcal{T}}|_{1,\Omega}^2 \leq c_{\text{apost}} (1 + C_B \gamma^{-2}) \eta_{\mathcal{T}}^2(u_{\mathcal{T}}, \mathcal{D}).$$

**4.1. First numerical results.** The aim of the present section is to confirm the theoretical predictions of Proposition 4.4 and in particular to assess the sharpness of inequality (4.10). To this end we solve a Poisson problem (2.1) employing in the VEM setting (2.22) the classic adaptive loop, already mentioned in (1.1),

$$(4.12) \quad \text{SOLVE} \longrightarrow \text{ESTIMATE} \longrightarrow \text{MARK} \longrightarrow \text{REFINE}$$

For the marking strategy, we use the Dörfler criterion [32]: determine the minimal set  $\mathcal{M} \subset \mathcal{T}$  such that

$$(4.13) \quad \theta \eta_{\mathcal{T}}^2(u_{\mathcal{T}}, \mathcal{D}) \leq \sum_{E \in \mathcal{M}} \eta_{\mathcal{T}}^2(E; u_{\mathcal{T}}, \mathcal{D})$$

for  $\theta \in (0, 1)$ . We stress that in VEM framework clearly the meshes generated by the adaptive algorithm need not be conforming. However, in accordance with Assumption 2.2, we require that the global index is uniformly bounded by  $\Lambda$ . We iterate the loop (4.12) with the stopping criterion based on the total number of degrees of freedom NDoFs

$$(4.14) \quad \text{NDoFs} \geq \text{N\_Max}.$$

In the present test we consider the L-shaped domain  $\Omega = (-1, 1) \times (-1, 1) \setminus [0, 1] \times [-1, 0]$  and solve the Poisson problem (2.1) with  $A = I$ ,  $c = 0$ ,  $f = 1$  and vanishing boundary conditions. The exact solution has a singular behaviour at the re-entrant corner. In the test we adopt the loop (4.12) with Dörfler parameter  $\theta = 0.5$ , and  $\text{N\_Max} = 2000$ , furthermore we pick  $\Lambda = 10$ . We adopt the **dofi-dofi** stabilization (2.13). To assess the effectiveness of bound (4.10) we consider the following quantity

$$\text{ratio} := \frac{\gamma^2 S_{\mathcal{T}}(u_{\mathcal{T}}, u_{\mathcal{T}})}{\eta_{\mathcal{T}}^2(u_{\mathcal{T}}, \mathcal{D})}.$$

In Fig. 3 we display the quantity **ratio** for different values of the stabilization parameter  $\gamma$  obtained with the adaptive algorithm (4.12). Notice that for all the proposed values of  $\gamma$  the quantity **ratio** at the first iteration of the algorithm is zero since the starting mesh  $\mathcal{T} = \mathcal{T}_0$  is made of triangular elements consequently  $S_{\mathcal{T}}(u_{\mathcal{T}}, u_{\mathcal{T}}) = 0$ . Fig. 3 shows that the estimate of Proposition 4.4 is sharp and for the proposed problem the constant  $C_B$  in (4.10) is bounded from above by 0.1.

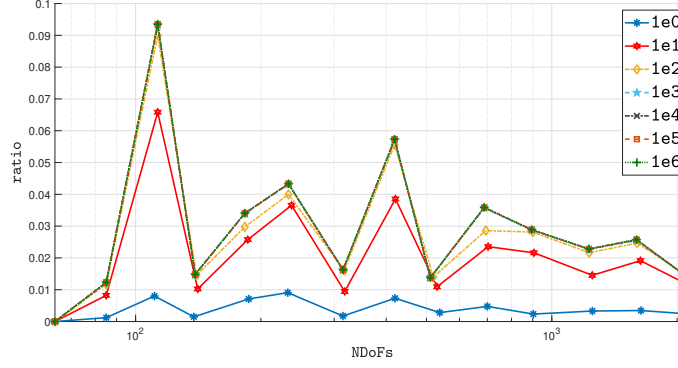


Fig. 3: Test 1. Sharpness of bound (4.10): **ratio** between the term  $\gamma^2 S_{\mathcal{T}}(u_{\mathcal{T}}, u_{\mathcal{T}})$  and the term  $\eta_{\mathcal{T}}^2(u_{\mathcal{T}}, \mathcal{D})$  obtained with the adaptive algorithm (4.12).

**5. Scaled Poincaré inequality in  $\mathbb{V}_{\mathcal{T}}$ : proof of Proposition 3.1.** We first introduce some useful definitions. Let  $\mathbb{T}$  denote the infinite binary tree obtained by newest-vertex bisection from the initial partition  $\mathcal{T}_0$ . If  $T \in \mathbb{T}$  is not a root, denote its parent by  $A(T)$ , and let  $\mathcal{A}(T)$  the chain of its ancestors, i.e.,

$$\mathcal{A}(T) = \{A_1(T) = A(T), A_j(T) = A(A_{j-1}(T)) \text{ for } j \geq 2 \text{ until the root is reached}\}.$$

Given an integer  $m \geq 1$ , let  $\mathcal{A}_m(T)$  be the subchain containing the first  $m$  ancestors of  $T$ .

Given any  $T \in \mathcal{T}$ , with vertices  $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3$ , define the cumulative index of  $T$  to be

$$\lambda(T) := \sum_{i=1}^3 \lambda(\mathbf{x}_i),$$

where  $\lambda(\mathbf{x}_i)$  is the global index of the node  $\mathbf{x}_i$ .

Let  $v \in \mathbb{V}_{\mathcal{T}}$  satisfy  $v(\mathbf{x}) = 0$  for all  $\mathbf{x} \in \mathcal{P}$ . We divide the proof into several steps.

*Step 1. Local bounds of norms.* Let  $E \in \mathcal{T}$  be fixed. If one of its vertices is a proper node, we immediately have

$$(5.1) \quad h_E^{-2} \|v\|_{0,E}^2 \lesssim |v|_{1,E}^2.$$

So, from now on, we assume that none of the vertices of  $E$  is a proper node. Since  $v$  need not vanish in  $E$ , we use the inequality

$$(5.2) \quad h_E^{-2} \|v\|_{0,E}^2 \lesssim (|v(\mathbf{x}_0)|^2 + |v|_{1,E}^2),$$

where  $\mathbf{x}_0$  is any point in  $E$ . Let us choose  $\mathbf{x}_0$  as the newest vertex of  $E$ . In the two previous inequalities, the hidden constants only depend on  $\dim \mathbb{V}_E$ , which by (2.4) and Remark 2.3 can be bounded by  $3 \cdot 2^A$ .

*Step 2. Path to a proper node.* Denote by  $i, j, k$  the global indices of the vertices of  $E$ , with  $i$  being the global index of  $\mathbf{x}_0$ ; by assumption, they are all  $> 0$ . Consider the parent  $T = A(E)$  of  $E$ , and let  $\ell \geq 0$  be the global index of the vertex of  $T$  not belonging to  $E$ . We claim that

$$(5.3) \quad \ell < i.$$

To prove this, observe that  $\mathbf{x}_0$  is the midpoint of an edge  $e$  of  $T$ , whose endpoints have global indices  $\ell$  and (say)  $k$  (see Fig. 4). By definition of global index of a hanging node, it holds  $i = \max(k, \ell) + 1$ : if  $k \leq \ell$ , then  $i = \ell + 1$  whence  $\ell = i - 1 < i$ , whereas if  $\ell < k$ , then  $i = k + 1$  whence  $\ell < i - 1 < i$ .

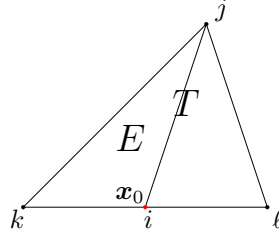


Fig. 4: Sample of element  $E \in \mathcal{T}$ , where  $T = A(E)$  is the parent of  $E$ ,  $i, j, k$  are the global indices of the vertices of  $E$  ( $i$  being the global index of  $\mathbf{x}_0$ ), and  $\ell$  is global index of the remaining vertex of  $T$ .

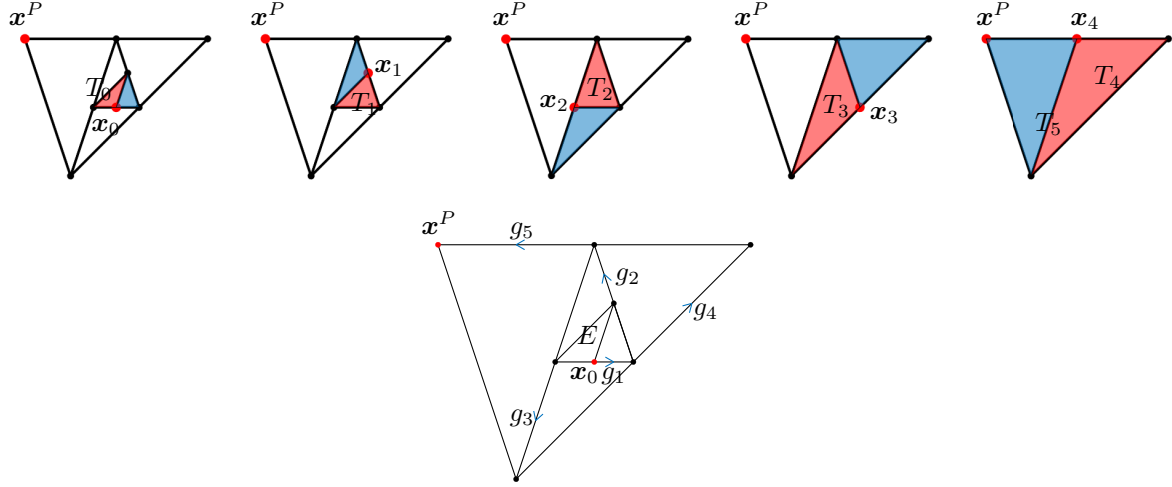


Fig. 5: The chain of ancestors of  $E = T_0$  leading to  $T_5$  having a proper node as a vertex (above). The path  $g_1 \rightarrow g_2 \rightarrow \dots \rightarrow g_5$  connecting  $\mathbf{x}_0 \in E$  to the proper node  $\mathbf{x}^P \in T_5$  (below)

Inequality (5.3) implies that the cumulative index decreases in going from  $E$  to  $T$ :

$$\lambda(E) > \lambda(T).$$

By repeating the argument above with  $E$  replaced by  $T$ , and then arguing recursively, we realize that when we move along the chain  $\mathcal{A}(E)$  of ancestors of  $E$ , the cumulative index strictly decreases by at least 1 unit each time, until it becomes  $< 3$ , indicating the presence of a proper node. In this way, after observing that  $\lambda(E) \leq 3\Lambda$  by Assumption 2.2, we obtain the existence of a subchain of ancestors

$$(5.4) \quad \mathcal{A}_M(E) = \{T_0 = E, T_1, \dots, T_M\}$$

with the following properties:

1.  $T_M$  is the first element in the chain which has a proper node, say  $\mathbf{x}^P$ , as a vertex;
2.  $M < 3\Lambda$ ;
3. for  $m > 0$ , each  $T_m$  has an edge  $g_m$  whose midpoint is the newest vertex of  $T_{m-1}$ ; the path  $g_1 \rightarrow g_2 \rightarrow \dots \rightarrow g_M$  connects the node  $\mathbf{x}_0 \in E$  (the midpoint of  $g_1$ ) to the proper node  $\mathbf{x}^P \in T_M$  (the endpoint of  $g_M$ ) (see Fig. 5).

*Step 3. Properties of edges with hanging nodes.* Consider an edge  $g$  shared by two triangles  $T, T' \in \mathbb{T}$ , with  $T' \in \mathcal{T}$ ; suppose that the midpoint  $\hat{\mathbf{x}}$  of  $g$  is a hanging node for  $T'$ , created by a refinement of  $T$  to produce elements in  $\mathcal{T}$ . Then,  $g$  cannot contain proper nodes except possibly the endpoints, since

their presence would be possible only by a refinement of  $T'$ , which is ruled out by the assumption on  $\hat{\mathbf{x}}$ .

Consequently, the edge  $g$  is partitioned by the hanging nodes into a number of edges  $\tilde{e}$  of elements  $\tilde{E} \in \mathcal{T}$  contained in  $T$ ; recalling Remark 2.3, the number of such edges is bounded by  $2^\Lambda$ .

*Step 4. Sequence of elements along the path.* We apply the conclusions of *Step 3.* to each edge  $g_m$  of the path defined in *Step 2.* We obtain the existence of a sequence of edges  $e_n$  ( $1 \leq n \leq N_E$  for some integer  $N_E$ ) and corresponding elements  $E_n \in \mathcal{T}$ , such that (see Fig. 6)

1.  $e_n \subset \partial E_n$ ;
2.  $e_n \subseteq g_m$  for some  $m$ , and correspondingly  $E_n \subseteq T_m$ , with

$$|E_n| \simeq |e_n|^2 \geq 2^{-2\Lambda} |g_m|^2 \simeq 2^{-2\Lambda} |T_m|,$$

where the hidden constants only depend on the shape of the initial triangulation but not on  $\Lambda$ ;

3. the number  $N_E$  of such elements is bounded by  $M2^\Lambda < 3\Lambda 2^\Lambda$ ;
4. writing  $e_n = [\mathbf{x}_{n-1}, \mathbf{x}_n]$ , then  $\mathbf{x}_0$  is the newest vertex of  $E$ , whereas  $\mathbf{x}_N$  is the proper node  $\mathbf{x}^P$ .

*Step 5. Bound of  $|v(\mathbf{x}_0)|$ .* Let us write

$$v(\mathbf{x}_0) = v(\mathbf{x}_0) - v(\mathbf{x}^P) = \sum_{n=1}^{N_E} (v(\mathbf{x}_{n-1}) - v(\mathbf{x}_n)) = \sum_{n=1}^{N_E} \nabla v|_{E_n} \cdot (\mathbf{x}_{n-1} - \mathbf{x}_n)$$

Then,

$$|v(\mathbf{x}_0)|^2 \leq N_E \sum_{n=1}^{N_E} \|\nabla v|_{E_n}\|^2 \|\mathbf{x}_{n-1} - \mathbf{x}_n\|^2 \lesssim N_E \sum_{n=1}^{N_E} \|\nabla v|_{E_n}\|^2 |E_n| = N_E \sum_{n=1}^{N_E} |v|_{1,E_n}^2.$$

Inserting this into (5.2) yields

$$(5.5) \quad h_E^{-2} \|v\|_{0,E}^2 \lesssim \left( |v|_{1,E}^2 + N_E \sum_{n=1}^{N_E} |v|_{1,E_n}^2 \right).$$

Taking into account also (5.1), we end up with the bound

$$(5.6) \quad \sum_{E \in \mathcal{T}} h_E^{-2} \|v\|_{0,E}^2 \lesssim \left( \sum_{E \in \mathcal{T}} |v|_{1,E}^2 + 3\Lambda 2^\Lambda \sum_{E \in \mathcal{T}^h} \sum_{n=1}^{N_E} |v|_{1,E_n}^2 \right),$$

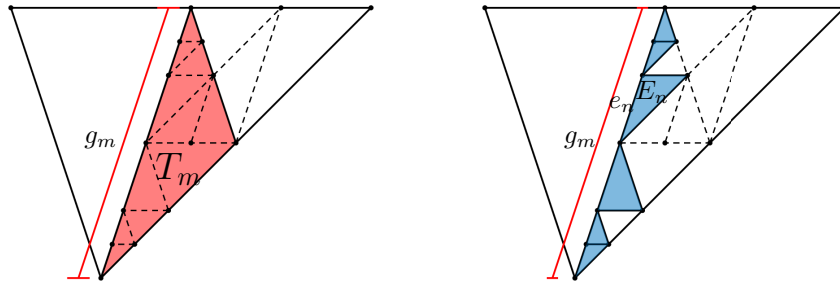


Fig. 6: The edge  $g_m$  of the ancestor  $T_m$  of  $E$  (left). The partition of  $g_m$  into edges  $e_n$  of elements  $E_n \in \mathcal{T}$  (right)

where  $\mathcal{T}^h$  denotes the set of elements in  $\mathcal{T}$  whose vertices are all hanging nodes. Thus, we will arrive at the desired result (3.1) if we show that an element  $E_n \in \mathcal{T}$  may occur in the double summation on the right-hand side a number of times bounded by some constant depending only on  $\Lambda$ .

*Step 6. Combinatorial count.* Let  $E_\star = E_n$  be such an element, which is contained in some triangle  $T_\star = T_m$  according to *Step 4.*, where  $T_\star$  belongs to a subchain of ancestors  $\mathcal{A}_M(E)$  defined in (5.4), for some  $E \in \mathcal{T}^h$ .

Notice that  $T_\star \in \mathcal{A}(E_\star)$  and its measure satisfies  $|T_\star| \lesssim 2^{2\Lambda}|E_\star|$  by *Step 4.* Thus, the number  $K$  of admissible  $T_\star$ 's satisfies

$$2^K |E_\star| = |T_\star| \lesssim 2^{2\Lambda} |E_\star|,$$

whence  $K \leq 2\Lambda + c$  for some  $c$  independent of  $\Lambda$ .

On the other hand, a triangle  $T_\star$  may belong to a subchain of ancestors  $\mathcal{A}_M(E)$ , for at most  $2^{3\Lambda}$  descendants  $E$ , since we have seen in *Step 2.* that  $M < 3\Lambda$ .

We conclude that  $E_\star$  may occur in the double summation on the right-hand side of (5.6) at most  $(2\Lambda + c)2^{3\Lambda}$  times. This concludes the proof of Proposition 3.1.

**6. Bounding the stabilization term by the residual.** This section is devoted to the proof of Proposition 4.4. The proof relies on an inequality between the interpolation errors in  $\mathbb{V}_{\mathcal{T}}^0$  and  $\mathbb{V}_{\mathcal{T}}$ , which will be the object of the next subsection.

**6.1. Interpolation errors.** We wish to bound the interpolation error relative to the ‘conforming’ operator  $\mathcal{I}_{\mathcal{T}}^0$  by the interpolation error relative to the ‘non-conforming’ operator  $\mathcal{I}_{\mathcal{T}}$ , namely we aim at providing a bound of the kind

$$|v - \mathcal{I}_{\mathcal{T}}^0 v|_{1,\Omega} \lesssim |v - \mathcal{I}_{\mathcal{T}} v|_{1,\mathcal{T}} \quad \forall v \in \mathbb{V}_{\mathcal{T}}.$$

Note that by the triangle inequality

$$|v - \mathcal{I}_{\mathcal{T}}^0 v|_{1,\Omega} = |v - \mathcal{I}_{\mathcal{T}}^0 v|_{1,\mathcal{T}} \leq |v - \mathcal{I}_{\mathcal{T}} v|_{1,\mathcal{T}} + |\mathcal{I}_{\mathcal{T}} v - \mathcal{I}_{\mathcal{T}}^0 v|_{1,\mathcal{T}},$$

it is enough to prove the bound

$$(6.1) \quad |\mathcal{I}_{\mathcal{T}} v - \mathcal{I}_{\mathcal{T}}^0 v|_{1,\mathcal{T}} \lesssim |v - \mathcal{I}_{\mathcal{T}} v|_{1,\mathcal{T}}.$$

To this end, we need several preparatory results that allow us to express both semi-norms as sums of hierarchical details. Let us start by considering the right-hand side.

Let  $E$  be any element in  $\mathcal{T}$ . Define

$$\begin{aligned} \mathcal{N}_E &= \{\mathbf{x} : \mathbf{x} \text{ is a node of } \mathcal{T} \text{ sitting on } \partial E\}, \\ \mathcal{V}_E &= \{\mathbf{x} : \mathbf{x} \text{ is a vertex of } E\}, \\ \mathcal{H}_E &= \mathcal{N}_E \setminus \mathcal{V}_E = \{\mathbf{x} : \mathbf{x} \text{ is a hanging node for } E\}. \end{aligned}$$

To each function  $v \in \mathbb{V}_E$  we associate a vector  $d(v) = \{d(v, \mathbf{z})\}_{\mathbf{z} \in \mathcal{N}_E}$  that collects the following values, so called *hierarchical details* of  $v$ :

$$(6.2) \quad d(v, \mathbf{z}) = \begin{cases} v(\mathbf{z}) & \text{if } \mathbf{z} \in \mathcal{V}_E, \\ v(\mathbf{z}) - \frac{1}{2}(v(\mathbf{z}') + v(\mathbf{z}'')) & \text{if } \mathbf{z} \in \mathcal{H}_E, \end{cases}$$

where for  $\mathbf{z} \in \mathcal{H}_E$  we denote by  $\mathbf{z}', \mathbf{z}'' \in \mathcal{B}(\mathbf{z})$  the endpoints of the edge halved to create  $\mathbf{z}$ . While the collection  $\{v(\mathbf{z})\}_{\mathbf{z} \in \mathcal{N}_E}$  represents the coefficients expressing  $v \in \mathbb{V}_E$  in terms of the (local) dual basis associated to the degrees of freedom, the collection  $\{d(v, \mathbf{z})\}_{\mathbf{z} \in \mathcal{N}_E}$  represents the coefficients with respect to a hierarchical-type basis.

The following lemma introduces a relation between the  $H^1$  semi-norm of a function  $v \in \mathbb{V}_E$  and the Euclidean norm of  $d(v)$ .

LEMMA 6.1 (Local interpolation error vs hierarchical details). *Under Assumption 2.2, for all  $E \in \mathcal{T}$  the relation*

$$(6.3) \quad \sum_{\mathbf{x} \in \mathcal{H}_E} d^2(v, \mathbf{x}) \lesssim |v - \mathcal{I}_E v|_{1,E}^2 \lesssim \sum_{\mathbf{x} \in \mathcal{H}_E} d^2(v, \mathbf{x}), \quad \forall v \in \mathbb{V}_E$$

holds with hidden constants only depending on  $\Lambda$ .

*Proof.* We recall that, according to [7, 22], the stability property (2.14) holds true for the particular choice (2.13) of stabilization (which need not be the stabilization used in our Galerkin scheme). Hence, we obtain

$$(6.4) \quad |v - \mathcal{I}_E v|_{1,E}^2 \simeq \sum_{\mathbf{x} \in \mathcal{H}_E} |v(\mathbf{x}) - \mathcal{I}_E(\mathbf{x})|^2 \quad \forall v \in \mathbb{V}_E,$$

where the symbol  $\simeq$  denotes an equivalence up to uniform constants. Noting that  $d(v, \mathbf{x}) = d(v - \mathcal{I}_E v, \mathbf{x})$ , equation (6.4) yields that (6.3) is equivalent to

$$\sum_{\mathbf{x} \in \mathcal{H}_E} d^2(v - \mathcal{I}_E v, \mathbf{x}) \simeq \sum_{\mathbf{x} \in \mathcal{H}_E} |v(\mathbf{x}) - \mathcal{I}_E(\mathbf{x})|^2 \quad \forall v \in \mathbb{V}_E,$$

which in turn corresponds to

$$(6.5) \quad \sum_{\mathbf{x} \in \mathcal{H}_E} d^2(w, \mathbf{x}) \simeq \sum_{\mathbf{x} \in \mathcal{H}_E} |w(\mathbf{x})|^2 \quad \forall w \in \tilde{\mathbb{V}}_E,$$

where  $\tilde{\mathbb{V}}_E = \{v \in \mathbb{V}_E : v(\mathbf{x}) = 0 \ \forall \mathbf{x} \in \mathcal{V}_E\}$ . The two quantities appearing in (6.5) are norms on the finite dimensional space  $\tilde{\mathbb{V}}_E$ . Furthermore, since both depend only on  $\{w(\mathbf{x})\}_{\mathbf{x} \in \mathcal{H}_E}$ , the values of  $w$  at the hanging nodes, such norms do not depend on the shape of  $E$ . On the other hand, an inspection of (6.2) reveals that the norm at the left hand side depends on the particular node “pattern” on  $E$ , i.e. on which sequential edge subdivision led to the appearance of hanging nodes on  $\partial E$ . Nevertheless, due to Assumption 2.2, not only the number of hanging nodes is uniformly bounded, but also the number of possible patterns is finite. As a consequence, property (6.5) easily follows from the equivalence of norms in finite dimensional spaces, with hidden constants only depending on  $\Lambda$ .  $\square$

In order to get the global equivalence, we observe that the set  $\mathcal{H} = \bigcup_{E \in \mathcal{T}} \mathcal{H}_E$  of all hanging nodes of  $\mathcal{T}$  is a disjoint union, i.e.,  $\mathbf{x} \in \mathcal{H}$  if and only if there exists a unique  $E \in \mathcal{T}$  such that  $\mathbf{x} \in \mathcal{H}_E$ . Combining this with (6.3), and recalling that  $(v - \mathcal{I}_\mathcal{T} v)|_E = v|_E - \mathcal{I}_E v|_E$  for any  $E \in \mathcal{T}$ , we obtain the following result.

COROLLARY 6.2 (Global interpolation error vs hierarchical details). *The following semi-norm equivalence holds:*

$$(6.6) \quad c_D \sum_{\mathbf{x} \in \mathcal{H}} d^2(v, \mathbf{x}) \leq |v - \mathcal{I}_\mathcal{T} v|_{1,\mathcal{T}}^2 \leq C_D \sum_{\mathbf{x} \in \mathcal{H}} d^2(v, \mathbf{x}), \quad \forall v \in \mathbb{V}_\mathcal{T},$$

where the constants  $C_D \geq c_D > 0$  depend on  $\Lambda$  but are independent of the triangulation  $\mathcal{T}$ .

Let us now focus on the left-hand side of (6.1). Since  $\mathcal{I}_\mathcal{T} v - \mathcal{I}_\mathcal{T}^0 v$  is affine on each element of  $\mathcal{T}$ , one has

$$|\mathcal{I}_\mathcal{T} v - \mathcal{I}_\mathcal{T}^0 v|_{1,\mathcal{T}}^2 = \sum_{E \in \mathcal{T}} |\mathcal{I}_\mathcal{T} v - \mathcal{I}_\mathcal{T}^0 v|_{1,E}^2 \simeq \sum_{E \in \mathcal{T}} \sum_{\mathbf{x} \in \mathcal{V}_E} (\mathcal{I}_E v - \mathcal{I}_\mathcal{T}^0 v)^2(\mathbf{x}).$$

Note that  $(\mathcal{I}_E v)(\mathbf{x}) = v(\mathbf{x})$  if  $\mathbf{x} \in \mathcal{V}_E$ . Furthermore,  $(\mathcal{I}_\mathcal{T}^0 v)(\mathbf{x}) = v(\mathbf{x})$  if  $\mathbf{x}$  is a proper node of  $\mathcal{T}$ . Thus, for any  $\mathbf{x} \in \mathcal{H}$ , let us define the detail

$$\delta(v, \mathbf{x}) = v(\mathbf{x}) - (\mathcal{I}_\mathcal{T}^0 v)(\mathbf{x}),$$

so that

$$(6.7) \quad |\mathcal{I}_{\mathcal{T}}v - \mathcal{I}_{\mathcal{T}}^0v|_{1,\mathcal{T}}^2 \simeq \sum_{\mathbf{x} \in \mathcal{H}} \delta^2(v, \mathbf{z}).$$

Recalling (6.6), the desired result (6.1) follows if we prove the bound

$$(6.8) \quad \sum_{\mathbf{x} \in \mathcal{H}} \delta^2(v, \mathbf{x}) \lesssim \sum_{\mathbf{x} \in \mathcal{H}} d^2(v, \mathbf{x}) \quad \forall v \in \mathbb{V}_{\mathcal{T}}.$$

From now on, to ease the notation, we assume  $v$  fixed and we write  $d(\mathbf{x}) := d(v, \mathbf{x})$ ,  $\delta(\mathbf{x}) := \delta(v, \mathbf{x})$ , and  $v^* := \mathcal{I}_{\mathcal{T}}^0v$ . Setting

$$\boldsymbol{\delta} = (\delta(\mathbf{x}))_{\mathbf{x} \in \mathcal{H}}, \quad \mathbf{d} = (d(\mathbf{x}))_{\mathbf{x} \in \mathcal{H}},$$

the desired inequality (6.8) is equivalent to

$$(6.9) \quad \|\boldsymbol{\delta}\|_{l^2(\mathcal{H})} \lesssim \|\mathbf{d}\|_{l^2(\mathcal{H})}.$$

We can relate  $\boldsymbol{\delta}$  to  $\mathbf{d}$  as follows: let  $\mathbf{x} \in \mathcal{H}$ , and let  $\mathbf{x}', \mathbf{x}'' \in \mathcal{B}(\mathbf{x})$ ; since  $v^*$  is linear on the segment  $[\mathbf{x}', \mathbf{x}'']$ , one has

$$(6.10) \quad \begin{aligned} \delta(\mathbf{x}) &= v(\mathbf{x}) - v^*(\mathbf{x}) = v(\mathbf{x}) - \frac{1}{2}(v^*(\mathbf{x}') + v^*(\mathbf{x}'')) \\ &= v(\mathbf{x}) - \frac{1}{2}(v(\mathbf{x}') + v(\mathbf{x}'')) + \frac{1}{2}(v(\mathbf{x}') - v^*(\mathbf{x}')) + \frac{1}{2}(v(\mathbf{x}'') - v^*(\mathbf{x}'')) \\ &= d(\mathbf{x}) + \frac{1}{2}\delta(\mathbf{x}') + \frac{1}{2}\delta(\mathbf{x}''). \end{aligned}$$

Thus, we have  $\boldsymbol{\delta} = \mathbf{W}\mathbf{d}$  for a suitable matrix of weights  $\mathbf{W} : l^2(\mathcal{H}) \rightarrow l^2(\mathcal{H})$ , and (6.9) holds true for any  $\mathbf{d}$  if and only if

$$\|\mathbf{W}\|_2 \lesssim 1.$$

To establish this bound, it is convenient to organize the hanging nodes in a block-wise manner according to the values of the global index  $\lambda \in [1, \Lambda_{\mathcal{T}}]$ . Let

$$\mathcal{H} = \bigcup_{1 \leq \lambda \leq \Lambda_{\mathcal{T}}} \mathcal{H}_{\lambda} \quad \text{with } \mathcal{H}_{\lambda} = \{\mathbf{x} \in \mathcal{H} : \lambda(\mathbf{x}) = \lambda\},$$

and let  $\boldsymbol{\delta} = (\delta_{\lambda})_{1 \leq \lambda \leq \Lambda_{\mathcal{T}}}$ ,  $\mathbf{d} = (d_{\lambda})_{1 \leq \lambda \leq \Lambda_{\mathcal{T}}}$  be the corresponding decompositions on the vectors  $\boldsymbol{\delta}$  and  $\mathbf{d}$ ; then, the matrix  $\mathbf{W}$ , considered as a block matrix, can be factorized as

$$(6.11) \quad \mathbf{W} = \mathbf{W}_{\Lambda_{\mathcal{T}}} \mathbf{W}_{\Lambda_{\mathcal{T}}-1} \cdots \mathbf{W}_{\lambda} \cdots \mathbf{W}_2 \mathbf{W}_1,$$

where the lower-triangular matrix  $\mathbf{W}_{\lambda}$  realizes the transformation (6.10) for the hanging nodes of index  $\lambda$ , leaving unchanged the other ones. In particular,  $\mathbf{W}_1 = \mathbf{I}$  since  $\delta(\mathbf{x}') = \delta(\mathbf{x}'') = 0$  when  $\mathbf{x}'$  and  $\mathbf{x}''$  are proper nodes; on the other hand, any other  $\mathbf{W}_{\lambda}$  differs from the identity matrix only in the rows corresponding to the block  $\lambda$ : each such row contains at most two non-zero entries, equal to  $\frac{1}{2}$ , in the off-diagonal positions, and 1 on the diagonal (see Fig. 7). In order to estimate the norm of  $\mathbf{W}_{\lambda}$ , we use Hölder's inequality  $\|\mathbf{W}_{\lambda}\|_2 \leq (\|\mathbf{W}_{\lambda}\|_1 \|\mathbf{W}_{\lambda}\|_{\infty})^{1/2}$ . Easily one has

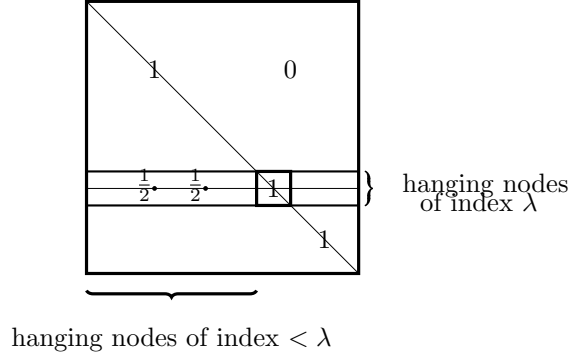
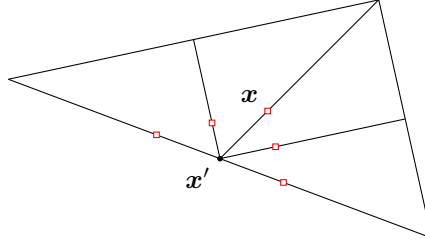
$$\|\mathbf{W}_{\lambda}\|_{\infty} \leq \frac{1}{2} + \frac{1}{2} + 1 = 2, \quad \|\mathbf{W}_{\lambda}\|_1 \leq 5\frac{1}{2} + 1 = \frac{7}{2},$$

the latter inequality stemming from the fact that a hanging node of index  $< \lambda$  may appear on the right-hand side of (6.10) at most 5 times (since at most 5 edges meet at a node, see Fig. 8). Hence,  $\|\mathbf{W}_{\lambda}\|_2 \leq 7^{1/2}$ , which yields

$$\|\mathbf{W}\|_2 \leq \prod_{2 \leq \lambda \leq \Lambda_{\mathcal{T}}} \|\mathbf{W}_{\lambda}\|_2 \leq 7^{(\Lambda_{\mathcal{T}}-1)/2} \leq 7^{(\Lambda-1)/2}$$

as desired.

In conclusion, concatenating (6.7), (6.8) and (6.6), we obtain the following result, which is the announced inequality between interpolation errors.

Fig. 7: The structure of one factor  $\mathbf{W}_\lambda$  in the factorization (6.11)Fig. 8: Given  $\mathbf{x}'$  with  $\lambda(\mathbf{x}') < \lambda$ , at most 5 hanging nodes  $\mathbf{x}$  of index  $\lambda$  may be such that  $\mathbf{x}' \in B(\mathbf{x})$ 

PROPOSITION 6.3 (Comparison between interpolation operators). *Under Assumption 2.2, there exists a constant  $C_I > 0$ , depending on  $\Lambda$  but independent of  $\mathcal{T}$ , such that*

$$(6.12) \quad |v - \mathcal{I}_\mathcal{T}^0 v|_{1,\Omega} \leq C_I |v - \mathcal{I}_\mathcal{T} v|_{1,\mathcal{T}} \quad \forall v \in \mathbb{V}_\mathcal{T}.$$

**6.2. Proof of Proposition 4.4.** By (2.19) and the definition (2.18), for all  $w \in \mathbb{V}_\mathcal{T}^0$  we obtain

$$\gamma S_\mathcal{T}(u_\mathcal{T}, u_\mathcal{T}) = \gamma S_\mathcal{T}(u_\mathcal{T}, u_\mathcal{T} - w) = \mathcal{B}_\mathcal{T}(u_\mathcal{T}, u_\mathcal{T} - w) - a_\mathcal{T}(u_\mathcal{T}, u_\mathcal{T} - w) - m_\mathcal{T}(u_\mathcal{T}, u_\mathcal{T} - w).$$

By (2.22), (4.1), the scaled Poincaré inequality and the continuity of  $\Pi_\mathcal{T}^\nabla$  with respect to the  $H^1$  broken seminorm we get

$$\begin{aligned} \mathcal{B}_\mathcal{T}(u_\mathcal{T}, u_\mathcal{T} - w) - m_\mathcal{T}(u_\mathcal{T}, u_\mathcal{T} - w) &= (f - c\Pi_\mathcal{T}^\nabla u_\mathcal{T}, \Pi_\mathcal{T}^\nabla(u_\mathcal{T} - w))_\Omega \\ &\leq \sum_{E \in \mathcal{T}} h_E \|r_\mathcal{T}(E; u_\mathcal{T}, \mathcal{D})\|_{0,E} h_E^{-1} (\|u_\mathcal{T} - w\|_{0,E} + h_E |u_\mathcal{T} - w|_{1,E}). \end{aligned}$$

On the other hand, by (2.12), (2.7) and (4.1)

$$\begin{aligned} a_\mathcal{T}(u_\mathcal{T}, u_\mathcal{T} - w) &= \sum_{E \in \mathcal{T}} (A_E \nabla \Pi_E^\nabla u_\mathcal{T}, \nabla \Pi_E^\nabla(u_\mathcal{T} - w))_E = \sum_{E \in \mathcal{T}} (A_E \nabla \Pi_E^\nabla u_\mathcal{T}, \nabla(u_\mathcal{T} - w))_E \\ &= \sum_{E \in \mathcal{T}} ((A_E \nabla \Pi_E^\nabla u_\mathcal{T}) \cdot \mathbf{n}, u_\mathcal{T} - w)_{\partial E} = \sum_{e \in \mathcal{E}} (j_\mathcal{T}(e; u_\mathcal{T}, \mathcal{D}), u_\mathcal{T} - w)_e \\ &\leq \frac{1}{2} \sum_{E \in \mathcal{T}} \sum_{e \in \mathcal{E}_E} h_E^{1/2} \|j_\mathcal{T}(e; u_\mathcal{T}, \mathcal{D})\|_{0,e} h_E^{-1/2} \|u_\mathcal{T} - w\|_{0,e}. \end{aligned}$$

Recalling (4.3) and (4.4), we thus obtain for any  $\delta > 0$

$$(6.13) \quad \gamma S_\mathcal{T}(u_\mathcal{T}, u_\mathcal{T}) \leq \frac{1}{2\delta} \eta_\mathcal{T}^2(u_\mathcal{T}, \mathcal{D}) + \frac{\delta}{2} \Phi_\mathcal{T}(u_\mathcal{T} - w) \quad \forall w \in \mathbb{V}_\mathcal{T}^0,$$

with

$$\begin{aligned}\Phi_{\mathcal{T}}(u_{\mathcal{T}} - w) &= \sum_{E \in \mathcal{T}} \left( h_E^{-2} \|u_{\mathcal{T}} - w\|_{0,E}^2 + |u_{\mathcal{T}} - w|_{1,E}^2 + \sum_{e \in \mathcal{E}_E} h_E^{-1} \|u_{\mathcal{T}} - w\|_{0,e}^2 \right) \\ &\lesssim \sum_{E \in \mathcal{T}} (h_E^{-2} \|u_{\mathcal{T}} - w\|_{0,E}^2 + |u_{\mathcal{T}} - w|_{1,E}^2) .\end{aligned}$$

At this point, we choose  $w = \mathcal{I}_{\mathcal{T}}^0 u_{\mathcal{T}}$  in (6.13) and apply (3.1) to  $u_{\mathcal{T}} - \mathcal{I}_{\mathcal{T}}^0 u_{\mathcal{T}}$ , getting

$$\Phi_{\mathcal{T}}(u_{\mathcal{T}} - \mathcal{I}_{\mathcal{T}}^0 u_{\mathcal{T}}) \lesssim |u_{\mathcal{T}} - \mathcal{I}_{\mathcal{T}}^0 u_{\mathcal{T}}|_{1,\Omega}^2 .$$

Then recalling (6.12) and (2.17), we derive the existence of a constant  $C_{\Phi} > 1$  independent of  $\mathcal{T}$  and  $u_{\mathcal{T}}$ , such that

$$\Phi_{\mathcal{T}}(u_{\mathcal{T}} - \mathcal{I}_{\mathcal{T}}^0 u_{\mathcal{T}}) \leq C_{\Phi} S_{\mathcal{T}}(u_{\mathcal{T}}, u_{\mathcal{T}}) .$$

We obtain the desired result by choosing  $\delta = \frac{\gamma}{C_{\Phi}}$  in (6.13), and setting  $C_B := C_{\Phi}$ .

**7. Variable data.** We briefly consider the extension of Propositions 4.1 and 4.2 to the case of variable data  $\mathcal{D} = (A, c, f)$ . To this end, we denote by  $\widehat{\mathcal{D}} = (\widehat{A}, \widehat{c}, \widehat{f})$  a piecewise constant approximation to  $\mathcal{D}$ . The discrete virtual problem is obtained from (2.22) by taking  $A_E = \widehat{A}_E$ ,  $c_E = \widehat{c}_E$  and  $f_E = \widehat{f}_E$  in (2.12) and (2.21), respectively. Similarly, we define  $\eta_{\mathcal{T}}^2(u_{\mathcal{T}}, \widehat{\mathcal{D}})$  from (4.1)-(4.2). The following result generalizes Proposition 4.1 to variable data.

**PROPOSITION 7.1** (Global upper bound). *There exists a constant  $\widehat{C}_{apost} > 0$  depending on  $\Lambda$  and  $\mathcal{D}$ , but independent of  $u$ ,  $\mathcal{T}$ ,  $u_{\mathcal{T}}$  and  $\gamma$ , such that*

$$(7.1) \quad |u - u_{\mathcal{T}}|_{1,\Omega}^2 \leq \widehat{C}_{apost} \left( \eta_{\mathcal{T}}^2(u_{\mathcal{T}}, \widehat{\mathcal{D}}) + S_{\mathcal{T}}(u_{\mathcal{T}}, u_{\mathcal{T}}) + \Psi_{\mathcal{T}}^2(u_{\mathcal{T}}, \widehat{\mathcal{D}}) \right)$$

where

$$\begin{aligned}\Psi_{\mathcal{T}}^2(u_{\mathcal{T}}, \widehat{\mathcal{D}}) &= \sum_{E \in \mathcal{T}} \Psi_{\mathcal{T}}^2(E; u_{\mathcal{T}}, \widehat{\mathcal{D}}) , \\ \Psi_{\mathcal{T}}^2(E; u_{\mathcal{T}}, \widehat{\mathcal{D}}) &= h_E^2 \|f - \widehat{f}_E\|_{0,E}^2 + \|(A - \widehat{A}_E) \nabla \Pi_E^{\nabla} u_{\mathcal{T}}\|_{0,E}^2 + \|(c - \widehat{c}_E) \nabla \Pi_E^{\nabla} u_{\mathcal{T}}\|_{0,E}^2 .\end{aligned}$$

*Proof.* We proceed as in the proof of Proposition 4.1. We set  $v_{\mathcal{T}} = \widetilde{\mathcal{I}}_{\mathcal{T}}^0 v$  and we get

$$\begin{aligned}\mathcal{B}(u - u_{\mathcal{T}}, v) &= \sum_{E \in \mathcal{T}} \left\{ (\widehat{f}, v - v_{\mathcal{T}})_E - (\widehat{A}_E \nabla \Pi_E^{\nabla} u_{\mathcal{T}}, \nabla(v - v_{\mathcal{T}}))_E - (\widehat{c}_E \Pi_E^{\nabla} u_{\mathcal{T}}, v - v_{\mathcal{T}})_E \right\} \\ &\quad + \sum_{E \in \mathcal{T}} \left\{ ((\widehat{A}_E - A) \nabla \Pi_E^{\nabla} u_{\mathcal{T}}, \nabla v)_E + ((\widehat{c}_E - c) \Pi_E^{\nabla} u_{\mathcal{T}}, v)_E \right\} \\ &\quad + \sum_{E \in \mathcal{T}} \left\{ (A \nabla (\Pi_E^{\nabla} - I) u_{\mathcal{T}}, \nabla v)_E + (c (\Pi_E^{\nabla} - I) u_{\mathcal{T}}, v)_E \right\} \\ &\quad + \gamma S_{\mathcal{T}}(u_{\mathcal{T}}, v_{\mathcal{T}}) + (f - \widehat{f}, v)\end{aligned}$$

from which the thesis easily follows using standard arguments.  $\square$

*Remark 7.2.* In the a posteriori bound (7.1) we highlight the presence of the oscillation term  $\Psi_{\mathcal{T}}(u_{\mathcal{T}}, \widehat{\mathcal{D}})$  measuring the impact of data approximation on the error.

The following is a generalization of Proposition 4.2 to variable coefficients.

**PROPOSITION 7.3** (Local lower bound). *There holds*

$$(7.2) \quad \eta_{\mathcal{T}}^2(E; u_{\mathcal{T}}, \widehat{\mathcal{D}}) \lesssim \sum_{E' \in \omega_E} \left( \|u - u_{\mathcal{T}}\|_{1,E'}^2 + S_{E'}(u_{\mathcal{T}}, u_{\mathcal{T}}) + \Psi_{\mathcal{T}}^2(E; u_{\mathcal{T}}, \widehat{\mathcal{D}}) \right)$$

where  $\omega_E := \{E' : |\partial E \cap \partial E'| \neq \emptyset\}$ . The hidden constant is independent of  $\gamma, h, u$  and  $u_{\mathcal{T}}$ .

*Proof.* Using standard arguments of a posteriori analysis the thesis follows as in [24].  $\square$

With these results at hand, we can extend the validity of Corollary (4.5) to the variable-coefficient case, as follows.

**COROLLARY 7.4** (Stabilization-free a posteriori error estimates). *There exist constants  $\hat{C}_{apost} \geq \hat{c}_{apost} > 0$  depending on  $\Lambda$  and  $\mathcal{D}$ , but independent of  $u$ ,  $\mathcal{T}$ ,  $u_{\mathcal{T}}$  and  $\gamma$ , such that if  $\gamma$  is chosen to satisfy  $\gamma^2 > \frac{C_B}{\hat{c}_{apost}}$ , it holds*

$$(7.3) \quad |u - u_{\mathcal{T}}|_{1,\Omega}^2 \leq \hat{C}_{apost} \left( (1 + C_B \gamma^{-2}) \eta_{\mathcal{T}}^2(u_{\mathcal{T}}, \mathcal{D}) + \Psi_{\mathcal{T}}^2(u_{\mathcal{T}}, \hat{\mathcal{D}}) \right).$$

$$(7.4) \quad (\hat{c}_{apost} - C_B \gamma^{-2}) \eta_{\mathcal{T}}^2(u_{\mathcal{T}}, \mathcal{D}) \leq |u - u_{\mathcal{T}}|_{1,\Omega}^2 + \Psi_{\mathcal{T}}^2(u_{\mathcal{T}}, \hat{\mathcal{D}}).$$

*Remark 7.5.* So far, we have considered homogeneous Dirichlet boundary conditions. Following [45], it is possible to extend our analysis to the case of non-homogeneous conditions, which amounts to incorporate the oscillation of the boundary data measured in  $H^{1/2}(\partial\Omega)$  into the term  $\Psi_{\mathcal{T}}^2(u_{\mathcal{T}}, \hat{\mathcal{D}})$ . Since this endeavor is similar to AFEMs, we omit the technical details and refer to [45].

**8. Numerical results.** In this section we present a numerical experiment to confirm the theoretical results in Proposition 4.4 and Corollary 4.5 and to test the practical performances of the proposed VEM scheme (2.22) coupled with the adaptive algorithm (4.12). In order to compute the VEM error between the exact solution  $u_{\text{ex}}$  and the VEM solution  $u_{\mathcal{T}}$ , we consider the computable  $H^1$ -like error quantity:

$$\mathbf{H}^1\text{-error} := \frac{(\sum_{E \in \mathcal{T}} \|\nabla(u_{\text{ex}} - \Pi_E^{\nabla} u_{\mathcal{T}})\|_{0,E}^2)^2}{\|\nabla u_{\text{ex}}\|_{0,\Omega}^2}.$$

Notice that  $\mathbf{H}^1\text{-error}$  (as well as the discrete problem (2.22)) depends only on the DoFs values of the discrete solution  $u_{\mathcal{T}}$ , hence, it is independent of the choice of the VEM space  $\mathbb{V}_E$  in (2.8). If the mesh  $\mathcal{T}$  does not contain hanging nodes, obviously  $\mathbf{H}^1\text{-error}$  coincides with the ‘true’  $H^1$ -relative error. In the numerical test we use the **dofi-dofi** stabilization (2.13) with stabilization parameter  $\gamma = 1$  (cf. (2.18)), we pick  $\Lambda = 10$  (cf. Assumption 2.2) and Dörfler parameter  $\theta = 0.5$  (cf. (4.13)), whereas the stopping parameter **NMax** (cf. (4.14)) will be specified later.

We consider from [39, Example 5.3] the Poisson problem (2.1) with piecewise constant coefficients and vanishing load with the following data:  $\Omega = (-1, 1)^2$ ,  $A = aI$ , with  $a = 161.4476387975881$  in the first and third quadrant and  $a = 1$  in the second and fourth quadrant,  $c = 0$  and  $f = 0$ . Employing the Kellogg formula [36] the exact solution is given in polar coordinates by  $u_{\text{ex}}(r, \alpha) = r^{\delta} \nu(\alpha)$  where

$$\nu(\alpha) := \begin{cases} \cos\left(\left(\frac{\pi}{2} - \sigma\right)\delta\right) \cos\left(\left(\alpha - \frac{\pi}{2} + \rho\right)\delta\right) & \text{if } 0 \leq \alpha \leq \pi/2, \\ \cos(\rho\delta) \cos\left(\left(\alpha - \pi + \sigma\right)\delta\right) & \text{if } \pi/2 \leq \alpha \leq \pi, \\ \cos(\sigma\delta) \cos\left(\left(\alpha - \pi - \rho\right)\delta\right) & \text{if } \pi \leq \alpha \leq 3\pi/2, \\ \cos\left(\left(\frac{\pi}{2} - \rho\right)\delta\right) \cos\left(\left(\alpha - \frac{3\pi}{2} - \sigma\right)\delta\right) & \text{if } 3\pi/2 \leq \alpha \leq 2\pi, \end{cases}$$

and where the numbers  $\delta$ ,  $\rho$ ,  $\sigma$  satisfy suitable nonlinear relations. In particular we pick  $\delta = 0.1$ ,  $\rho = \pi/4$  and  $\sigma = -14.92256510455152$ . Notice that the exact solution  $u_{\text{ex}}$  is in the Sobolev space  $H^{1+\epsilon}$  only for  $\epsilon < 0.1$  and thus is very singular at the origin.

In Fig. 9 (plot on the left) we display  $\mathbf{H}^1\text{-error}$ , the estimator  $\eta_{\mathcal{T}}(u_{\mathcal{T}}, \mathcal{D})$  and the stabilization term  $S_{\mathcal{T}}(u_{\mathcal{T}}, u_{\mathcal{T}})^{1/2}$  obtained with the adaptive algorithm (4.12) with stopping parameter **NMax** = 25000. The predictions of Proposition 4.4 and Corollary 4.5 are confirmed: the estimator bounds from above both the energy error and the stabilization term. Furthermore, one can appreciate that, after a fairly long transient due to the highly singular structure of the solution, the error decay reaches asymptotically the theoretical optimal rate  $\mathbf{NDoFs}^{-0.5}$  (whereas the estimator decays with this rate along the whole refinement history).

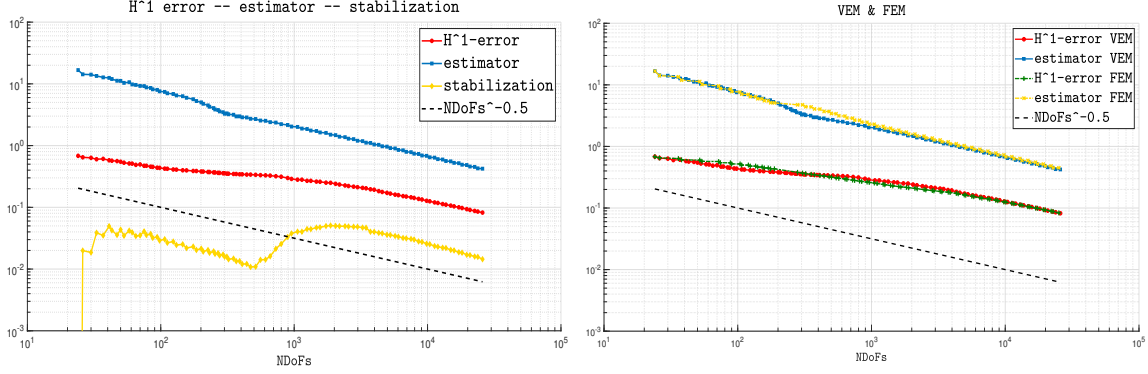


Fig. 9: Left:  $H^1$ -error, estimator  $\eta_{\mathcal{T}}(u_{\mathcal{T}}, \mathcal{D})$ , stabilization term  $S_{\mathcal{T}}(u_{\mathcal{T}}, u_{\mathcal{T}})^{1/2}$ . Right:  $H^1$ -error and estimator  $\eta_{\mathcal{T}}(u_{\mathcal{T}}, \mathcal{D})$  obtained with VEM and FEM. In both figures the optimal decay is indicated by the dashed line with slope  $-0.5$ .

To validate the practical performances of the proposed numerical scheme (2.22), we compare the results obtained with our VEM to those obtained with a standard  $\mathbb{P}_1$  FEM, implemented as a VEM with  $\Lambda = 0$  (cf. Assumption 2.2). The results for both methods are obtained with an “in-house” code, yet the FEM outcomes coincide (up to machine precision) with those obtained with the code developed in [33]. In Fig. 9 (plot on the right) we display  $H^1$ -error and estimator  $\eta_{\mathcal{T}}(u_{\mathcal{T}}, \mathcal{D})$  obtained with VEM and FEM coupled with the adaptive algorithm (4.12) and stopping parameter  $N_{\text{Max}} = 25000$ . Notice that for FEM  $H^1$ -error is the “true”  $H^1$ -relative error. Both methods yield very similar results in terms of behaviour of the error and the estimator.

However, a deeper analysis shows important differences between VEM and FEM approximations in terms of the final grids denoted respectively with  $\mathcal{T}_{\text{VEM}}$  and  $\mathcal{T}_{\text{FEM}}$ . In Fig. 10 we display the meshes  $\mathcal{T}_{\text{VEM}}$  and  $\mathcal{T}_{\text{FEM}}$  obtained with stopping parameter  $N_{\text{Max}} = 5000$ . The number of nodes  $N_{\text{vertices}}$  and elements  $N_{\text{elements}}$  are  $N_{\text{vertices}}(\mathcal{T}_{\text{VEM}}) = 5259$ ,  $N_{\text{elements}}(\mathcal{T}_{\text{VEM}}) = 8725$ ,  $N_{\text{vertices}}(\mathcal{T}_{\text{FEM}}) = 5070$ ,  $N_{\text{elements}}(\mathcal{T}_{\text{FEM}}) = 10094$ , i.e. the mesh  $\mathcal{T}_{\text{FEM}}$  has 16% more elements than the mesh  $\mathcal{T}_{\text{VEM}}$ . Furthermore the number of polygons in  $\mathcal{T}_{\text{VEM}}$  (elements with more than three vertices) is 1653: 1563 quadrilaterals, 86 pentagons, 2 hexagons, 1 heptagon and 1 nonagon.

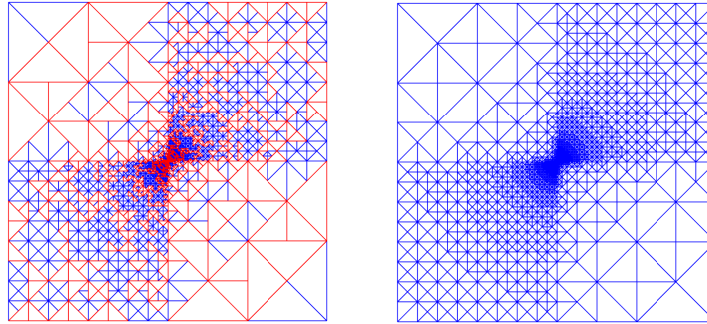


Fig. 10: Left: final grid  $\mathcal{T}_{\text{VEM}}$  obtained with VEM. Right: final grid  $\mathcal{T}_{\text{FEM}}$  obtained with FEM. Mesh elements having more than three vertices are drawn in red.

The grids  $\mathcal{T}_{\text{VEM}}$  and  $\mathcal{T}_{\text{FEM}}$  are highly graded at the origin along the bisector of the first and third quadrants. However from Fig. 11 we can appreciate how grid  $\mathcal{T}_{\text{VEM}}$  still exhibits a rather strong grading also for the zoom scaled to  $10^{-9}$ .

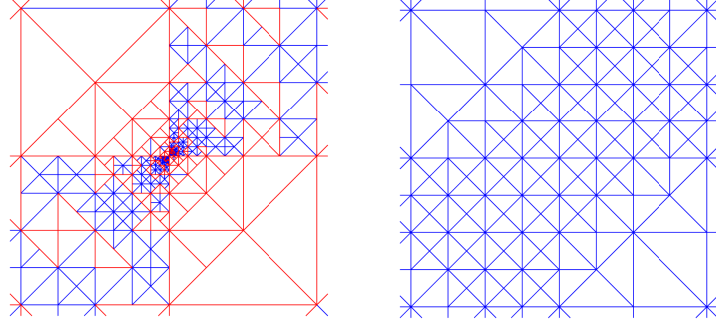


Fig. 11: Left: final grid  $\mathcal{T}_{\text{VEM}}$ . Right: final grid  $\mathcal{T}_{\text{FEM}}$ . Zoom to  $(-10^{-9}, 10^{-9})^2$ . Mesh elements having more than three vertices are drawn in red.

Finally, in Fig. 12 we plot the zoom to  $(-10^{-10}, 10^{-10})^2$  for the grid  $\mathcal{T}_{\text{VEM}}$  and the plot of the discrete solution for the finer grid. We highlight the presence of the nonagon with two nodes having global index  $\lambda = 3$ . It is worth noting that the largest global index is  $\lambda = 3$ , whence the threshold  $\Lambda = 10$  is never reached by AVEM. Therefore, the condition  $\lambda \leq \Lambda$  is not restrictive in practice.

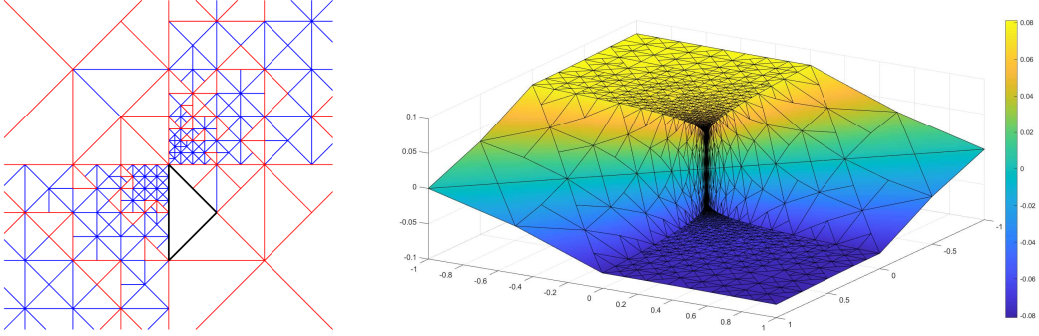


Fig. 12: Left: final grid  $\mathcal{T}_{\text{VEM}}$ , zoom to  $(-10^{-10}, 10^{-10})^2$  (mesh elements having more than three vertices are drawn in red, and the mesh element drawn in black is a nonagon). Right: graph of the discrete solution.

**9. Conclusions.** The analysis of this paper relies crucially on the existence of a subspace  $\mathbb{V}_{\mathcal{T}}^0 \subseteq \mathbb{V}_{\mathcal{T}}$  satisfying the following properties:

- The discrete forms satisfy the consistency property (2.19) on  $\mathbb{V}_{\mathcal{T}}^0$ ;
- There exists a subset  $\mathcal{P}$  of mesh nodes such that the collection of linear operators  $v \rightarrow \{v(\mathbf{x})\}_{\mathbf{x} \in \mathcal{P}}$  constitutes a set of degrees of freedom for  $\mathbb{V}_{\mathcal{T}}^0$ ;
- Propositions 3.1 and 6.3 hold for the above choice of  $\mathbb{V}_{\mathcal{T}}^0$  and  $\mathcal{P}$ .

We have established these properties for meshes made of triangles, which is the most common situation in finite element methods. Yet, there are other cases in which the above construction can be easily applied. One notable example is that of square meshes, where we assume a standard quadtree element refinement procedure that subdivides each square into four squares; in this framework the advantage of allowing hanging nodes is evident. In such case, the space  $\mathbb{V}_{\mathcal{T}}^0$  is chosen as

$$\mathbb{V}_{\mathcal{T}}^0 = \{v \in \mathbb{V}_{\mathcal{T}} \text{ such that } v|_E \in \mathbb{Q}_1(E) \ \forall E \in \mathcal{T}\},$$

where  $\mathbb{Q}_1$  denotes the space of bilinear polynomials, and  $\mathcal{P}$  is the set of proper (i.e., non-hanging) nodes of the mesh. It is not difficult to adapt to the new framework the arguments given in the paper,

and prove the validity of the three conditions above, thereby arriving at the same conclusions obtained for triangles. Obviously, heterogeneous meshes formed by triangles and squares could be handled as well.

The extension of the techniques presented above to general polygonal meshes (for which even a deep understanding of the refinement strategies is currently missing) seems highly non-trivial. However, we believe that the main result of this paper, namely the bound of the stabilization term by the a posteriori error estimator, should hold in a wide variety of situations. We also hope that some of the ideas that we have elaborated here will turn useful to attack the challenging problems of providing a sound mathematical framework to adaptive virtual element methods (AVEM).

In any case, as anticipated in the Introduction, the results presented here will serve as a basis in the sequel paper [13] for the design of an AVEM on triangular partitions admitting hanging nodes, which will be proven to converge and possess optimal approximation properties.

### Acknowledgements

LBdV, CC and MV were partially supported by the Italian MIUR through the PRIN grants n. 201744KLJL (LBdV, MV) and n. 201752HKH8 (CC). CC was also supported by the DISMA Excellence Project (CUP: E11G18000350001). RHN has been supported in part by NSF grant DMS-1908267. These supports are gratefully acknowledged. LBdV, CC, MV and GV are members of the INdAM research group GNCS.

### REFERENCES

- [1] B. Ahmad, A. Alsaedi, F. Brezzi, L. D. Marini, and A. Russo. Equivalent projectors for virtual element methods. *Comput. Math. Appl.*, 66(3):376–391, 2013.
- [2] P.F. Antonietti, S. Berrone, A. Borio, A. D’Auria, M. Verani, and S. Weisser. Anisotropic a posteriori error estimate for the virtual element method. *arXiv:2001.00381v1*.
- [3] E. Artioli, S. Marfia, and E. Sacco. Vem-based tracking algorithm for cohesive/frictional 2d fracture. *Comp. Meth. in Appl. Mech. and Engrng.*, 365:112956, 2020.
- [4] I. Babuška and A. Miller. 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):1–40, 1987.
- [5] I. Babuška and W. C. Rheinboldt. Error estimates for adaptive finite element computations. *SIAM J. Numer. Anal.*, 15(4):736–754, 1978.
- [6] R. E. Bank and A. Weiser. Some a posteriori error estimators for elliptic partial differential equations. *Math. Comp.*, 44(170):283–301, 1985.
- [7] L. Beirão da Veiga, C. Lovadina, and A. Russo. Stability analysis for the virtual element method. *Math. Mod. and Meth. in Appl. Sci.*, 27(13):2557–2594, 2017.
- [8] L. Beirão da Veiga and G. Manzini. Residual a posteriori error estimation for the virtual element method for elliptic problems. *ESAIM Math. Model. Numer. Anal.*, 49(2):577–599, 2015.
- [9] L. Beirão da Veiga, G. Manzini, and L. Mascotto. A posteriori error estimation and adaptivity in *hp* virtual elements. *Numer. Math.*, 143(1):139–175, 2019.
- [10] L. Beirão da Veiga, F. Brezzi, A. Cangiani, G. Manzini, L. D. Marini, and A. Russo. Basic principles of virtual element methods. *Math. Models Methods Appl. Sci.*, 23(1):199–214, 2013.
- [11] L. Beirão da Veiga, F. Brezzi, L. D. Marini, and A. Russo. The Hitchhiker’s Guide to the Virtual Element Method. *Math. Models Methods Appl. Sci.*, 24(8):1541–1573, 2014.
- [12] L. Beirão da Veiga, F. Brezzi, L. D. Marini, and A. Russo. Virtual Element Method for general second-order elliptic problems on polygonal meshes. *Math. Models Methods Appl. Sci.*, 24(4):729–750, 2016.
- [13] L. Beirão da Veiga, C. Canuto, R. H. Nochetto, G. Vacca, and M. Verani. Adaptive VEM: Convergence and optimality. (*in preparation*), 2021.
- [14] L. Beiro da Veiga, C. Canuto, R. H. Nochetto, and G. Vacca. Equilibrium analysis of an immersed rigid leaflet by the virtual element method. *Math. Mod. and Meth. in Appl. Sci.*, 31(07):1323–1372, 2021.
- [15] M.F. Benedetto, S. Berrone, S. Pieraccini, and S. Scialo. The virtual element method for discrete fracture network simulations. *Comp. Meth. in Appl. Mech. and Engrng.*, 280:135–156, 2014.
- [16] S. Berrone and A. Borio. A residual a posteriori error estimate for the Virtual Element Method. *Math. Models Methods Appl. Sci.*, 27(8):1423–1458, 2017.
- [17] C. Bhm, B. Hudobivnik, M. Marino, and P. Wriggers. Electro-magneto-mechanically response of polycrystalline materials: Computational homogenization via the virtual element method. *Comp. Meth. in Appl. Mech. and Engrng.*, 380:113775, 2021.
- [18] P. Binev, W. Dahmen, and R. DeVore. Adaptive finite element methods with convergence rates. *Numer. Math.*, 97(2):219–268, 2004.

- [19] A. Bonito and R. H. Nochetto. Quasi-optimal convergence rate of an adaptive discontinuous Galerkin method. *SIAM J. Numer. Anal.*, 48(2):734–771, 2010.
- [20] D. Braess, V. Pillwein, and J. Schöberl. Equilibrated residual error estimates are  $p$ -robust. *Comput. Methods Appl. Mech. Engrg.*, 198(13-14):1189–1197, 2009.
- [21] D. Braess and J. Schöberl. Equilibrated residual error estimator for edge elements. *Math. Comp.*, 77(262):651–672, 2008.
- [22] S. C. Brenner and L. Y. Sung. Virtual element methods on meshes with small edges or faces. *Math. Models Methods Appl. Sci.*, 28(7):1291–1336, 2018.
- [23] A. Cangiani, E. H. Georgoulis, T. Pryer, and O. J. Sutton. A posteriori error estimates for the virtual element method. *Numer. Math.*, 137(4):857–893, 2017.
- [24] A. Cangiani, E. H. Georgoulis, T. Pryer, and O. J. Sutton. A posteriori error estimates for the virtual element method. *Numer. Math.*, 137(4):857–893, 2017.
- [25] A. Cangiani and M. M. Munar. A posteriori error estimates for mixed virtual element methods. *arXiv:1904.10054v1*.
- [26] C. Canuto, R. H. Nochetto, R. Stevenson, and M. Verani. Adaptive spectral Galerkin methods with dynamic marking. *SIAM J. Numer. Anal.*, 54(6):3193–3213, 2016.
- [27] C. Canuto, R. H. Nochetto, R. Stevenson, and M. Verani. Convergence and optimality of  $hp$ -AFEM. *Numer. Math.*, 135(4):1073–1119, 2017.
- [28] H. Chi, L. Beirão da Veiga, and G. H. Paulino. A simple and effective gradient recovery scheme and a *a posteriori* error estimator for the virtual element method (VEM). *Comput. Methods Appl. Mech. Engrg.*, 347:21–58, 2019.
- [29] H. Chi, A. Pereira, I. F. M. Menezes, and G. H. Paulino. Virtual element method (vem)-based topology optimization: an integrated framework. *Struct. Multidisc. Optim.*, 62:1089–1114, 2020.
- [30] F. Dassi, J. Gedicke, and L. Mascotto. Adaptive virtual element methods with equilibrated fluxes. *arXiv:2004.11220v2*.
- [31] F. Dassi, J. Gedicke, and L. Mascotto. Adaptive virtual elements based on hybridized, reliable, and efficient flux reconstructions. *arXiv:2107.03716v2*.
- [32] W. Dörfler. A convergent adaptive algorithm for Poisson’s equation. *SIAM J. Numer. Anal.*, 33(3):1106–1124, 1996.
- [33] S. Funken, D. Praetorius, and P. Wissgott. Efficient implementation of adaptive P1-FEM in Matlab. *Comput. Methods Appl. Math.*, 11(4):460–490, 2011.
- [34] O. A. Karakashian and F. Pascal. A posteriori error estimates for a discontinuous Galerkin approximation of second-order elliptic problems. *SIAM J. Numer. Anal.*, 41(6):2374–2399, 2003.
- [35] O. A. Karakashian and F. Pascal. Convergence of adaptive discontinuous Galerkin approximations of second-order elliptic problems. *SIAM J. Numer. Anal.*, 45(2):641–665, 2007.
- [36] R. B. Kellogg. On the Poisson equation with intersecting interfaces. *Appl. Anal.*, 4:101–129, 1975.
- [37] J. M. Melenk and B. I. Wohlmuth. On residual-based a posteriori error estimation in  $hp$ -FEM. *Adv. Comput. Math.*, 15(1-4):311–331 (2002), 2001.
- [38] D. Mora, G. Rivera, and R. Rodríguez. A virtual element method for the Steklov eigenvalue problem. *Math. Models Methods Appl. Sci.*, 25(8):1421–1445, 2015.
- [39] P. Morin, R. H. Nochetto, and K. G. Siebert. Data oscillation and convergence of adaptive FEM. *SIAM J. Numer. Anal.*, 38(2):466–488, 2000.
- [40] P. Morin, R. H. Nochetto, and K. G. Siebert. Local problems on stars: a posteriori error estimators, convergence, and performance. *Math. Comp.*, 72(243):1067–1097, 2003.
- [41] M. Munar and F. A. Sequeira. A posteriori error analysis of a mixed virtual element method for a nonlinear Brinkman model of porous media flow. *Comput. Math. Appl.*, 80(5):1240–1259, 2020.
- [42] R. H. Nochetto, K. G. Siebert, and A. Veiser. Theory of adaptive finite element methods: an introduction. In *Multiscale, nonlinear and adaptive approximation*, pages 409–542. Springer, Berlin, 2009.
- [43] R. H. Nochetto and A. Veiser. Primer of adaptive finite element methods. In *Multiscale and adaptivity: modeling, numerics and applications*, volume 2040 of *Lecture Notes in Math.*, pages 125–225. Springer, Heidelberg, 2012.
- [44] R. Rodríguez. Some remarks on Zienkiewicz-Zhu estimator. *Numer. Methods Partial Differential Equations*, 10(5):625–635, 1994.
- [45] R. Sacchi and A. Veiser. Locally efficient and reliable *a posteriori* error estimators for Dirichlet problems. *Math. Models Methods Appl. Sci.*, 16(3):319–346, 2006.
- [46] R. Stevenson. The completion of locally refined simplicial partitions created by bisection. *Math. Comp.*, 77(261):227–241, 2008.
- [47] O. C. Zienkiewicz and J. Z. Zhu. A simple error estimator and adaptive procedure for practical engineering analysis. *Internat. J. Numer. Methods Engrg.*, 24(2):337–357, 1987.

## MOX Technical Reports, last issues

Dipartimento di Matematica  
Politecnico di Milano, Via Bonardi 9 - 20133 Milano (Italy)

- 69/2021** Antonietti, P.F.; Caldana, M.; Dede', L.  
*Accelerating Algebraic Multigrid Methods via Artificial Neural Networks*
- 65/2021** Mazzieri, I.; Muhr, M.; Stupazzini, M.; Wohlmuth, B.  
*Elasto-acoustic modelling and simulation for the seismic response of structures: The case of the Tahtali dam in the 2020 Izmir earthquake*
- 66/2021** Antonietti, P.F.; Botti, M.; Mazzieri, I.  
*On mathematical and numerical modelling of multiphysics wave propagation with polygonal Discontinuous Galerkin methods*
- 68/2021** Regazzoni, F.; Salvador, M.; Dede', L.; Quarteroni, A.  
*A machine learning method for real-time numerical simulations of cardiac electromechanics*
- 67/2021** Salvador, M.; Regazzoni, F.; Pagani, S.; Dede', L.; Trayanova, N.; Quarteroni, A.  
*The role of mechano-electric feedbacks and hemodynamic coupling in scar-related ventricular tachycardia*
- 64/2021** Clarotto, L.; Allard, D.; Menafooglio, A.  
*A new class of alpha-transformations for the spatial analysis of Compositional Data*
- 63/2021** Rosafalco, L.; Torzoni, M.; Manzoni, A.; Mariani, S.; Corigliano, A.  
*Online structural health monitoring by model order reduction and deep learning algorithms*
- 62/2021** Lupo Pasini, M.; Burcul, M.; Reeve, S.; Eisenbach, M.; Perotto, S.  
*Fast and accurate predictions of total energy for solid solution alloys with graph convolutional neural networks*
- 59/2021** Stella, S.; Regazzoni, F.; Vergara, C.; Dede', L.; Quarteroni, A.  
*A fast cardiac electromechanics model coupling the Eikonal and the nonlinear mechanics equations*
- 60/2021** Rosafalco, L.; Manzoni, A.; Mariani, S.; Corigliano, A.  
*Fully convolutional networks for structural health monitoring through multivariate time series classification*