Exponential convergence of the $hp$ Virtual Element Method with corner singularities

L. Beirão da Veiga $^*$†, A. Chernov ‡, L. Mascotto §, A. Russo ¶∥

Abstract

In the present work, we analyze the $hp$ version of Virtual Element methods for the 2D Poisson problem. We prove exponential convergence of the energy error employing sequences of polygonal meshes geometrically refined, thus extending the classical choices for the decomposition in the $hp$ Finite Element framework to very general decomposition of the domain. A new stabilization for the discrete bilinear form with explicit bounds in $h$ and $p$ is introduced. Numerical experiments validate the theoretical results. We also exhibit a numerical comparison between $hp$ Virtual Elements and $hp$ Finite Elements.

1 Introduction

The Virtual Element Method (VEM) is a very recent generalization of the Finite Element Method (FEM), see [13]. VEM utilizes polygonal/polyhedral meshes in lieu of the classical triangular/tetrahedral and quadrilateral/hexahedral meshes. This automatically includes nonconvex elements, hanging nodes (enabling natural handling of interface problems with nonmatching grids), easy construction of adaptive meshes and efficient approximations of geometric data features.

Among the properties of VEM, in addition to the employment of polytopal meshes, we recall the possibility of handling approximation spaces of arbitrary $C^k$ global regularity [23, 33] and approximation spaces that satisfy exactly the divergence-free constraint [5, 22].

The main idea of VEM consists in enriching the classical polynomial space with other functions, whose explicit knowledge is not needed for the construction of the method (this explains the name virtual).

We point out that the literature concerning methods based on polytopal meshes is not restricted to the Virtual Element Method. A (very short and incomplete) list of other polytopal-based methods follows: Hybrid High Order Methods [40], Mimetic Finite Difference [20, 32], Hybrid Discontinuous Galerkin Methods [39], Polygonal Finite Element Method [19, 31, 69], Polygonal Discontinuous Galerkin Methods [35], BEM-based Finite Element Methods [66].

Although VEM is a very recent technology, the associated bibliography is widespread. Among the treated topic, we recall the following works concerning implementation issues [18], general linear second-order elliptic problems [3, 15, 16, 38], Stokes problem [5, 22, 31, 67], Cahn-Hillard equation [6], locking-free linear elasticity problem [17, 43], small deformation problems in structural mechanics [12], plate bending problem [33], Steklov eigenvalue problem [63], residual based a-posteriori error estimation [24, 36, 52], serendipity Virtual Elements [14], application to discrete fracture network simulations [25, 27], contact problem [63], comparison with the Smoothed Finite Element Method [54], topology optimization [32], geomechanics problem [4], Helmholtz problem [55].

In all these works, the convergence of VEM approximations has been achieved by increasing the number of mesh elements while keeping the degree of local approximation fixed. In other words and according to the existing terminology, these methods utilize the $h$-version of VEM.

---

*$^*$Dip. di Matematica e Applicazioni, Università degli Studi di Milano-Bicocca, E-mail: lourenco.beirao@unimib.it
†IMATI-CNR, Pavia
‡Inst. für Mathematik, C. von Ossietzky Universität Oldenburg, E-mail: alexey.chernov@uni-oldenburg.de
§Dip. di Matematica, Università degli Studi di Milano, E-mail: lorenzo.mascotto@unimi.it and Inst. für Mathematik, C. von Ossietzky Universität Oldenburg, E-mail: mascotto.lorenzo@uni-oldenburg.de
¶Dip. di Matematica e Applicazioni, Università degli Studi di Milano-Bicocca, E-mail: alessandro.russo@unimib.it
∥IMATI-CNR, Pavia
An alternative avenue to construct convergent approximations is the p-version of VEM which is based on increasing the degree of local approximations while keeping the underlying mesh fixed. A combination of both methodologies is termed the hp-version of VEM.

The recent work [19] provides a mathematical ground for the p-version of VEM for the two-dimensional Poisson problem. In particular, it includes the a priori convergence theory for the p- and hp-version of VEM on quasiuniform polygonal meshes and for uniform distributions of local degrees of approximation. An exponential convergence has been established for analytic solutions and convergence at algebraic rates for solutions having finite Sobolev regularity.

The objectives of the present paper are the following. First, we generalize the results in [19] and in particular the definition of $H^1$ conforming Virtual Element to the case with varying local degree of accuracy from element to element. Such construction fits very naturally in the framework of VEM without additional complications. Furthermore, we extend the hp-VEM approach to nonquasiuniform approximations and prove their exponential convergence for nonsmooth solutions having typical corner singularities, see [7, 8].

Similarly to the hp-version of FEM (see [57] and the references therein) the approximation is based on geometrically refined meshes with appropriate (linearly varying) local degree of accuracy. In order to derive the proofs, we introduce a new stabilization for the method, which turns out to be more suitable for the p and hp version of VEM; in particular, explicit bounds of the new stabilizing term with respect to the $H^1$ seminorm in terms of the local degree of accuracy are shown. This proof requires a particular inverse estimate on polynomials, presented in the first appendix. To the best of the authors knowledge, such an inverse estimate has been never published before.

As a byproduct of this work, new tools for the approximation by means of functions in the Virtual Element Space are presented; such tools permits to avoid additional assumptions on the polygonal decomposition of the computational domain.

The structure of the paper is the following. After presenting the model problem and the Virtual Element Method in Section 2 and 3 respectively, we deal with explicit bounds in terms of the degree of accuracy of a new stabilization term in Section 4. In Section 5 we show the approximation results and the main theorem of the paper, namely the exponential convergence of the energy error in terms of the dimension of the virtual space, while in Section 6 we validate the theoretical results with numerical tests, including a set of experiments on the comparison between hp FEM and hp VEM. Finally, in the two Appendices A and B, we discuss two particular polynomial inverse estimates needed for the stability bounds of Section 4.

## 2 Model problem

In this section, we discuss the functional space setting and the model problem under consideration.

Firstly, we introduce the functional spaces that will be used throughout the paper. Let $\ell \in \mathbb{N}$ and let $\mathcal{D} \subseteq \mathbb{R}^2$ be a given domain whose closure contains the origin, i.e. $0 \in \mathcal{D}$. We denote with $L^\ell(\mathcal{D})$ the Lebesgue space of $\ell$-summable functions and we denote with $H^\ell(\omega)$ the Sobolev space of order $\ell$ on the domain $\mathcal{D}$, respectively; let $\| \cdot \|_{\ell,\mathcal{D}}$ and $| \cdot |_{\ell,\mathcal{D}}$ be the Sobolev norm and seminorm, see [2]. Let $H^1_0(\mathcal{D}) = \{ u \in H^1(\mathcal{D}) : u|_{\partial \mathcal{D}} = 0 \}$.

Let now $\beta \in (0, 1)$, $\Phi_\beta(x) = |x|^{\beta}$, where $| \cdot |$ represents the Euclidean norm in $\mathbb{R}^2$. Given $u : \omega \to \mathbb{R}$, $m, \ell \in \mathbb{N}$, $m \geq \ell$, we set

$$
|u|_{H^{m,\ell}_\beta(\mathcal{D})}^2 := \sum_{k=\ell}^{m} \| D^k u | \Phi_{\beta+k-\ell} \|_{0,\mathcal{D}}^2.
$$

(1)

where:

$$
|D^\alpha u|^2 = \sum_{\alpha \in \mathbb{N}^2, |\alpha| = n} |D^\alpha u|^2.
$$

We define the weighted Sobolev spaces

$$
H^{m,\ell}_\beta(\mathcal{D}) := \left\{ u \in L^2(\mathcal{D}) \left| \| u \|_{H^{m,\ell}_\beta(\mathcal{D})} < \infty \right. \right\},
$$

(2)

2
where the corresponding weighted Sobolev norm reads

\[
\|u\|_{H^m,D}^2 := \begin{cases} 
\|u\|_{L^2(D)}^2 + |u|^2_{H^m,D} & \text{if } \ell \geq 1, \\
\sum_{k=0}^m |D^k u|_{\partial D}^2 & \text{if } \ell = 0.
\end{cases}
\]  

Having this, we recall the countably normed spaces (also known as Babuška spaces), see \[57\] and the references therein:

\[B_\beta^\ell(D) := \{ u \mid u \in H^{m,\ell}(D), \forall m \geq \ell \text{ and } \|D^k u\|_{0,D} \leq c \cdot d^{k-\ell} \cdot (k-\ell)! \}, \forall k = \ell, \ell + 1, \ldots \]

where \(c \geq 0\) and \(d \geq 1\) are two constants depending on \(u\) and \(D\). We point out that space \(\mathbb{H}\) is not empty since it contains functions of the form \(u = |x|^\alpha\), for some \(\alpha > 0\).

Definition \((\mathbb{H})\) can be generalized to the case of functions with “multiple” singularities at the vertices of a polygonal domain i.e. adding in definition \((\mathbb{H})\) weights of the form \(\|x-x_0\|\), \(x_0\) being a vertex of the polygon different from \(0\); see \[57\]. In particular, defining \(N_V(D)\) and \(\{A_i(D)\}_{i=1}^{N_V(D)}\) the number and set of vertices of \(D\) respectively, we will denote the general space with \(H^{m,\ell}(D)\), where \(\beta = (\beta_1, \ldots, \beta_{N_V(D)})\) and \(\ell = (\ell_1(\beta_1), \ldots, \ell_{N_V(D)}(\beta_{N_V(D)}))\) are the vectors associated with the singularities at the vertices of \(\omega\). The associated weight function reads:

\[\Phi_\beta(x) = \Pi_{i=1}^{N_V(D)} r_i(x)^{\beta_i}, \quad r_i(x) = \min(1, |x-A_i(D)|).\]

Secondly, we introduce the model problem. Let \(\Omega\) be a open simply connected polygonal domain. Let \(f \in L^2(\Omega)\) be given. We consider the two dimensional Poisson problem:

\[-\Delta u = f \text{ in } \Omega, \quad u = 0 \text{ on } \partial \Omega \]

and its weak formulation:

\[
\text{find } u \in V := H^1_0(\Omega) \text{ such that } a(u,v) = (f,v)_{0,\Omega}, \quad \forall v \in V,
\]

where \((\cdot, \cdot)_{0,\Omega}\) is the \(L^2\) scalar product on \(\Omega\) and \(a(\cdot, \cdot) = (\nabla \cdot, \nabla \cdot)_{0,\Omega}\). The Lax-Milgram lemma guarantees the existence of a unique weak solution \(u \in V\).

We recall a regularity result for the solution of problem \((\mathbb{H})\). Let \(N_V\) and \(\{A_i\}_{i=1}^{N_V}\) be the number and the set of vertices of \(\Omega\) respectively; let \(\omega_i\) be the (internal) angle associated with vertex \(A_i\), \(i = 1, \ldots, N_V\). To each \(\omega_i\), we associate the set of the so-called singular exponents for Poisson problem with Dirichlet condition (see also \[57\], formula (4.2.2)):

\[\alpha_i = \frac{\pi}{\omega_i}, \quad \forall i = 1, \ldots, N_V.\]

Then, the following holds:

**Theorem 2.1.** Following the notation of problem \((\mathbb{H})\), let \(f \in H^{0,0}_\beta(\Omega), \beta \in [0,1)\); assume that the singular exponents \(\alpha_i\) defined in \((\mathbb{H})\) satisfy:

\[1 - \alpha_i < \beta_i < 1, \quad \text{if } \alpha_i < 1, \quad \forall i = 1, \ldots, N_V.\]

Then the solution of \((\mathbb{H})\) belongs to \(H^{2,2}_\beta(\Omega)\) and the a priori estimate:

\[\|u\|_{H^{2,2}_\beta(\Omega)} \leq c \|f\|_{H^{0,0}_\beta(\Omega)},\]

holds. Moreover, if \(f \in B^\ell_\beta\Omega\), then \(u \in B^\ell_\beta(\Omega)\).

**Proof.** See \[7][8].

For the sake of simplicity, we will assume in the rest of the paper that

\[0 \in \partial \Omega\]

is the only vertex at which the solution \(u\) of problem \((\mathbb{H})\) can be singular.

Finally, we point out that throughout the paper we will write \(a \approx b\) and \(a \lesssim b\) meaning that there exist \(c_1, c_2\) and \(c_3\) positive constants independent of the discretization parameters, such that \(c_1 a \leq b \leq c_2 a\) and \(a \leq c_3 b\) respectively; we will also denote by \(P_r(K)\) and \(P_r(E)\) the spaces of polynomials of degree \(r\) over a polygon \(K\) and an edge \(E\) respectively.
3 Virtual Element Spaces with non uniformly distributed degree of accuracy

In this section, we introduce a Virtual Element Method for problem (6) with nonuniform local degree of accuracy.

Let \{T_n\} be a sequence of polygonal decompositions of the domain \(\Omega\). The approximation will have a “geometric layer” structure; hence, in the sequel, the integer \(n\) will represent the number of layers used for the corner singularity refinement as in [31]; see Section 5.1 for the precise definition of layers.

Let \(V_n\) be the set of vertices of \(T_n\), \(\nu^b_n = \{\nu \in V_n \mid \nu \in \partial \Omega\}\) be the subset of boundary vertices, \(\mathcal{E}_n\) be the set of edges of \(T_n\), \(\mathcal{E}^b_n = \{e \in \mathcal{E}_n \mid e \subseteq \partial \Omega\}\) be the subset of boundary edges. To each \(K \in T_n\), we associate \(h_K = \text{diam}(K)\), \(\mathcal{V}^K = \{\nu \in V_n \mid \nu \in \partial K\}\) and \(\mathcal{E}^K = \{e \in \mathcal{E}_n \mid e \subseteq \partial K\}\). We require the two following basic assumptions on the regularity of the decomposition:

\begin{enumerate}
  \item \(\forall K \in T_n\), \(K\) is star-shaped with respect to a ball of radius greater than or equal to \(h_K\) \(\gamma\), where \(\gamma\) is a positive constant independent of the decompositions; see [31] for the definition of star-shapedness. We note that this condition can be satisfied by possibly many balls.
  \item \(\forall K \in T_n\), \(\forall E \in \mathcal{E}^K\), \(|E| \geq h_E \tilde{\gamma}\), where \(\tilde{\gamma}\) is a positive constant independent of the decompositions. Moreover, \(\forall K \in T_n\), \(\text{card}(\mathcal{E}^K)\) is uniformly bounded.
\end{enumerate}

More technical assumptions on the mesh will be introduced in Section 5 for the construction of proper geometric meshes.

Remark 1. Assuming that (D1) and (D2) hold true, then the following is also valid. The sub-triangulation \(\tilde{T}_n(K)\) of \(K\) obtained by joining the vertices of \(K\) to the center of the ball \(B(K)\) introduced in assumption (D1) is made of triangles that are star-shaped with respect to a ball of radius greater than or equal to \(\gamma_1 h_T\), \(h_T\) being the diameter of \(T\), \(\forall T \in \tilde{T}_n(K)\), and \(\gamma_1\) being a positive constant independent of the decompositions.

Given \(K \in T_n\), let \(i_K\) be the position of polygon \(K\) in the ordered sequence \(T_n\). Let \(p \in \mathbb{N}^{\text{card}(T_n)}\). We associate to each \(K \in T_n\) the local degree of accuracy \(p_{i_K} = (p)_{i_K}\). In order to simplify the notation, we write \(p_K := p_{i_K}\).

Henceforth, we assume that \(T_n\) is a conforming decomposition into polygons of \(\Omega\), i.e., for all \(E \in \mathcal{E}\), either \(E\) belongs to two polygons if it is an internal edge or it belongs to a single polygon if it is a boundary edge.

In the former case, it must hold that there exist \(K_1, K_2 \in T_n\) such that \(E \in \mathcal{E}^{K_1} \cap \mathcal{E}^{K_2}\); we associate to \(E\) the degree \(p_E = \max\{p_{K_1}, p_{K_2}\}\), that is we adopt the so-called maximum rule; see Remark 2 for further comments. In the latter case, let \(K \in T_n\) be the unique polygon in the decomposition such that \(E \in \mathcal{E}^K\); we associate to \(E\) the degree \(p_E = p_K\).

Let \(K \in T_n\). We firstly define the space of piecewise continuous polynomials on the boundary of \(K\):

\[\mathbb{B}(\partial K) := \{v_n \in C^0(\partial K) \mid v_n|_E \in \mathbb{P}_{p_E}(E), \forall E \in \mathcal{E}^K\}\].

The local virtual space on \(K\) reads:

\[V(K) := \{v_n \in H^1(K) \mid \Delta v_n \in \mathbb{P}_{p_K-2}(K) \text{ and } v_n \in \mathbb{B}(\partial K)\}\].

The convention \(\mathbb{P}_{-1}(K) = 0\) and where \(\mathbb{B}(\partial K)\) is defined in (9).

Definition (10) and the maximum rule immediately imply that \(\mathbb{P}_{p_K}(K) \subseteq V(K)\).

We associate with the local space the following set of degrees of freedom:

- the values at the vertices of \(K\);
- the values at \(p_E - 1\) internal nodes (e.g. Gauß-Lobatto nodes) for all \(E \in \mathcal{E}^K\);
- the scaled internal moments of the form \(\frac{1}{|K|} \int_K q_\alpha v_n\), where \(\{q_\alpha\}_{|\alpha|=0}^{p_K-2}\) is a properly chosen basis of \(\mathbb{P}_{p_K-2}(K)\); see [18] for a possible explicit choice of the basis.
This is in fact a set of degrees of freedom for the local space (10); see (13). If we set \( \text{dof}_i \) the \( i \)-th degree of freedom, \( i = 1, \ldots, \dim(V(K)) \), then we can define the local virtual canonical basis \( \{ \phi_j, j = 1, \ldots, \dim(V(K)) \} \) by:

\[
\text{dof}_i(\phi_j) = \delta_{i,j}, \quad \forall i, j = 1, \ldots, \dim(V(K)).
\] (11)

The global virtual space is obtained by matching the boundary degrees of freedom on each edge, i.e.:

\[
V_n := \{ v_n \in C^0(\overline{\Omega}) \mid v_n|_K \in V(K), \forall K \in \mathcal{T}_n; v_n|_{\partial\Omega} = 0 \}. \tag{12}
\]

We note that we can split the global continuous bilinear form \( a(\cdot, \cdot) \), introduced with the continuous problem (6), into a sum of local contributions as follows:

\[
a(u, v) = \sum_{K \in \mathcal{T}_n} a^K(u, v), \quad \text{where} \quad a^K(u, v) = (\nabla u, \nabla v)_{0,K}. \tag{13}
\]

We observe that we cannot compute the bilinear form \( a(\cdot, \cdot) \) applied on virtual functions since it is not possible in principle to know the values of such functions at any internal points of each polygon. The same argument applies to the computation of the right-hand side. For this reason, we must approximate both the stiffness matrix and the right-hand side.

Thus, the structure of VEM approximation is based on the two following ingredients which will be defined in what follows:

- a symmetric bilinear form \( a_n : V_n \times V_n \to \mathbb{R} \), which we decompose into a sum of local symmetric bilinear forms \( a^K_n : V(K) \times V(K) \to \mathbb{R} \) as follows:

\[
a_n(v_n, w_n) = \sum_{K \in \mathcal{T}_n} a^K_n(v_n, w_n), \quad \forall v_n, w_n \in V_n; \tag{14}
\]

- a piecewise discontinuous polynomial \( f_n \), which is piecewise of degree \( p_K \), and the associated linear functional \( (f_n, \cdot)_{0,\Omega} \).

The discrete bilinear form \( a_n(\cdot, \cdot) \) and the discrete right-hand side \( f_n \) are chosen in such a way that the discrete counterpart of (6)

\[
\text{find } u_n \in V_n \text{ such that } a_n(u_n, v_n) = (f_n, v_n)_{0,K}, \quad \forall v_n \in V_n
\] (15)

is well-posed and it is possible to recover local \( hp \)-estimates analogous to those proved in [19].

We start by discussing the construction of the discrete bilinear form. We require that \( a^K_n \) in [13] satisfy the two following assumptions:

(A1) polynomial consistency: \( \forall K \in \mathcal{T}_n \), it must hold:

\[
a^K_n(q, v_n) = a^K_n(q, v_n), \quad \forall q \in \mathbb{P}_{p_K}(K), \forall v_n \in V(K); \tag{16}
\]

(A2) local stability: \( \forall K \in \mathcal{T}_n \), it must hold

\[
\alpha_s(K)|v_n|^2_{1,K} \leq a^K_n(v_n, v_n) \leq \alpha^*(K)|v_n|^2_{1,K}, \quad \forall v_n \in V(K),
\] (17)

where \( 0 < \alpha_s(K) \leq \alpha^*(K) < +\infty \) are two constants, which may depend only on the local space \( V(K) \).

On each \( K \in \mathcal{T}_n \), we can introduce a local energy projector \( \Pi^K_{p_K} : V(K) \to \mathbb{P}_{p_K}(K) \) via

\[
\begin{align*}
& a^K(q, \Pi^K_{p_K}v_n - v_n) = 0 \quad \forall q \in \mathbb{P}_{p_K}(K), \forall v_n \in V(K), \tag{18a} \\
& \int_{\partial K} (\Pi^K_{p_K}v_n - v_n) ds = 0 \quad \forall v_n \in V(K). \tag{18b}
\end{align*}
\]

When no confusion occurs, we will write \( \Pi^K_{p_K} \) in lieu of \( \Pi^K_{p_K} \).
Note that condition \((18b)\) only fixes an additive constant of the projection and can be modified if necessary. Importantly, this local energy projector can be computed by means of the degrees of freedom of space \((10)\), see \([13, 18]\), without the need of knowing explicitly functions in the virtual space.

In \([13, 18]\), it was also shown that a computable candidate for \(a_n^K\) may have the following form:

\[
ad^K_n(u,v) = a^K(\Pi_{p_K}^\nu u_n, \Pi_{p_K}^\nu v_n) + S^K(u_n - \Pi_{p_K}^\nu u_n, v_n - \Pi_{p_K}^\nu w_n), \quad \forall v_n, w_n \in V(K),
\]

where \(S^K\) is any computable symmetric bilinear form on \(V(K)\) such that

\[
|c_*^K|v_n|^2_{1,K} \leq S^K(v_n, v_n) \leq c^*|v_n|^2_{1,K}, \quad \forall v_n \in V(K) \quad \text{with} \quad \Pi_{p_K}^\nu v_n = 0,
\]

where \(0 < c_*^K \leq c^*(K) < +\infty\) are two constants, depending possibly on the local space \(\text{ker}(\Pi_{p_K}^\nu)\). In \([13]\) it was shown that \((20)\) implies \((17)\) with:

\[
\alpha_*(K) = \min(1, c_*(K)), \quad \alpha^*(K) = \max(1, c^*(K)), \quad \forall K \in T_n.
\]

Now, we introduce a computable discrete loading term \(f_n\). Let \(S^{p,-1}(\Omega, T_n)\) be the set of piecewise discontinuous polynomials over the decomposition \(T_n\) of degree \(p_K\) on each \(K \in T_n\). For \(\ell \in \mathbb{N}\), let \(\Pi_{p,K}^\ell := \Pi_{p,K}^\ell \nu\) be the \(L^2(K)\) projector from the local space \((10)\) to \(P_\ell(K)\), the space of polynomials of degree \(\ell\) over \(K\); such a projector can easily be computed whenever \(\ell \leq p_K - 2\) by means of the internal degrees of freedom of the space \((10)\), see \([13]\).

We define the discrete loading term as follows: \(f_n \in \mathcal{S}^{p,-1}(\Omega, T_n)\) is such that

\[
(f_n, v_n)_{0,\Omega} = \sum_{K \in T_n} (f_n, v_n)_{0,K}, \quad \text{where} \quad (f_n, v_n)_{0,K} := \int_K \Pi_{p,K}^{-2}\nu f_n, \quad \forall v_n \in V_n.
\]

A deeper analysis on the discrete loading term can be found in \([3]\) and \([17]\).

We remark that in this paper we will not consider the case of approximation with \(p_K = 1\) in order to avoid technical discussions on the right-hand side.

**Remark 2.** We point out that in the definition of the local Virtual Space \((10)\), we fixed the degree of the edge to be the maximum of the degree of the two adjacent polygons (maximum-rule). One could also fix such an edge degree to be the minimum of the degree of the neighbouring polygons (minimum-rule). The first choice leads to \(P_{p,K}^\nu(K) \subseteq V(K)\); therefore, it is possible to recover local (i.e. on each polygon) classical \(hp\)-estimates, see \([19]\). On the other hand, in view of Section 5 also the choice of the minimum would yield the same convergence result.

Let \(\mathcal{F}_n^K, K \in T_n\), be the smallest positive constants such that:

\[
|(f_n, v_n)_{0,K} - (f, v_n)_{0,K}| \leq \mathcal{F}_n^K |v_n|_{1,K}, \quad \forall v_n \in V(K)
\]

and let

\[
\alpha(K) := \frac{1 + \alpha^*(K)}{\min_{K' \in T_n} \alpha_*(K')}, \quad \forall K \in T_n.
\]

where \(\alpha_*(K)\) and \(\alpha^*(K)\) are introduced in \((17)\).

We show how the energy error \(|u - u_n|_{1,\Omega}\) can be bounded, \(u\) and \(u_n\) being respectively the solutions of \((9)\) and \((15)\). We carry out, in particular, an abstract error analysis which is similar to the one presented in \([13]\); nevertheless, we decide to show the details, since assumption \((A2)\) is here weaker than its \(h\) counterpart in \([13]\), where the stability constants \(\alpha_*(K)\) and \(\alpha^*(K)\) are assumed to be independent of the local spaces.

**Lemma 3.1.** Assume that \((A1)\) and \((A2)\) hold. Let \(u\) and \(u_n\) be the solutions of problems \((9)\) and \((15)\) respectively. Then, for all \(u_1 \in V_n\) and for all \(u_\pi \in \mathcal{S}^{p,-1}(\Omega, T_n)\), it holds that

\[
|u - u_n|_{1,\Omega} \leq \sum_{K \in T_n} \alpha(K) \left\{ \mathcal{F}_n^K + |u - u_\pi|_{1,K} + |u - u_I|_{1,K} \right\},
\]

where \(\mathcal{F}_n^K\) and \(\alpha(K)\) are defined in \((23)\) and \((24)\) respectively.
Proof. Given any \( u_\varepsilon \in S^{p^{-1}}(\Omega, \mathcal{T}_n) \) and \( u_I \in V_n \):

\[
|u_n - u_I|^2_{1,\Omega} = \sum_{K \in \mathcal{T}_n} |u_n - u_I|^2_{1,K} \overset{\text{(A2)}}{\leq} \sum_{K \in \mathcal{T}_n} \alpha_n^{-1}(K) a_n^K(u_n - u_I, u_n - u_I)
\]

\[
\overset{\text{(A1), (23)}}{\leq} \left( \max_{K \in \mathcal{T}_n} \alpha_n^{-1}(K') \right) \sum_{K \in \mathcal{T}_n} \left\{ (f_n - f, u_n - u_I)_{0,K} - a_n^K(u_I - u_\varepsilon, u_n - u_I) - a^K(u_\varepsilon - u, u_n - u_I) \right\}
\]

\[
\overset{\text{(A1), (23)}}{\leq} \left( \max_{K \in \mathcal{T}_n} \alpha_n^{-1}(K') \right) \sum_{K \in \mathcal{T}_n} \left\{ (\mathcal{F}_n^K|u_I - u_n|_{1,K} + \alpha^*(K)|u_I - u_\varepsilon|_{1,K}|u_n - u_I|_{1,K} + |u - u_\varepsilon|_{1,K}|u_n - u_I|_{1,K} \right\}
\]

\[
\leq \left( \max_{K \in \mathcal{T}_n} \alpha_n^{-1}(K') \right) \left( \sum_{K \in \mathcal{T}_n} (\mathcal{F}_n^K + (1 + \alpha^*(K))|u - u_\varepsilon|_{1,K} + \alpha^*(K)|u - u_I|_{1,K})^2 \right)^{\frac{1}{2}} |u_n - u_I|_{1,\Omega}
\]

where the Cauchy-Schwarz inequality has been used in the last step. Applying a triangular inequality, we get:

\[
|u - u_n|_{1,\Omega} \leq \sum_{K \in \mathcal{T}_n} \frac{1 + \alpha^*(K)}{\min_{K \in \mathcal{T}_n} \alpha_n(K')} \left\{ \mathcal{F}_n^K + |u - u_\varepsilon|_{1,K} + |u - u_I|_{1,K} \right\}.
\]

This finishes the proof. \( \square \)

4 Stability

In this section, we present an explicit choice for the stabilizing bilinear form \( S^K \) introduced in (20) and we discuss the associated stability bounds (20) in terms of the local degree of accuracy. Our choice for the stabilization is the following:

\[
S^K(u_n, v_n) = \frac{p_K}{h_K}(u_n, v_n)_{0,\partial K} + \frac{p_K^2}{h_K^2}(\Pi_{p_K}^0 - 2u_n, \Pi_{p_K}^0 - 2v_n)_{0,K}.
\]

We note that this local stabilization term is explicitly computable by means of the local degrees of freedom, since on the boundary virtual functions are known polynomials and the \( L^2 \) projections are computable using only the internal degrees of freedom, see [15].

Following the guidelines of [57, formula (4.5.61)], that is the \( p \)-version of the Aubin-Nitsche duality argument, it holds for a convex \( K \):

\[
\|v_n - \Pi^V_{p_K} v_n\|_{0,K} \lesssim \frac{h_K}{p_K} |v_n - \Pi^V_{p_K} v_n|_{1,K}, \quad \forall v_n \in V(K).
\]

(27)

Note that in order to apply the Aubin-Nitsche argument we use the fact that \( v_n \in \ker(\Pi_{p_K}^\Sigma) \), which guarantees that \( v_n - \Pi^V v_n \) has zero average on the boundary.

Assume now that \( K \) is nonconvex. Let \( \pi < \omega_K < 2\pi \) the largest angle of \( K \). Then, the Aubin-Nitsche analysis in addition to interpolation theory, see [61,62], can be refined giving:

\[
\|v_n - \Pi^V v_n\|_{0,K} \lesssim \left( \frac{h_K}{p_K} \right)^{\frac{\omega_K}{\pi}} |v_n - \Pi^V_{p_K} v_n|_{1,K}, \quad \forall v_n \in V(K).
\]

(28)

We now prove the following result.

**Theorem 4.1.** Assume that \( p_K \), the degree of accuracy of the method on the element \( K \), coincides with the polynomial degrees \( p_E \), for all edges \( E \in \mathcal{E}^K \) of polygon \( K \). Then, using definition (20), the bounds in (20) hold with:

\[
c_*(K) \geq p_K^{-5}, \quad c^*(K) \leq \begin{cases} 1 & \text{if } K \text{ is convex}, \\ 2(1 - \frac{\omega_K}{\pi}) & \text{otherwise}, \end{cases}
\]

where \( \omega_K \) denotes the largest angle of \( K \).
Proof. We assume without loss of generality that the size of polygon $K$ is 1. The general result follows from a scaling argument.

We start by proving the estimate on $c_*(K)$. Integrating by parts, we obtain for $v_n \in \ker(\Pi_{pk}^\Sigma)$:

$$|v_n|_{1,K}^2 = \int_K \nabla v_n \cdot \nabla v_n = \int_K -\Delta v_n \Pi_{pk-2}^0 v_n + \int_{\partial K} \frac{\partial v_n}{\partial n} v_n. \tag{30}$$

We split our analysis into two parts. We firstly investigate the integral over $K$ in (30). For this purpose, we need a technical result, namely the following $hp$ polynomial inverse estimate in two dimensions, see Corollary A.5 (which can be applied thanks to Remark 1):

$$\|q\|_{1,K} \lesssim (p_K - 1) \|q\|_{-1,K} \lesssim p_K^2 \|q\|_{-1,K}, \quad \forall q \in \mathbb{P}_{pk-2}(K), \tag{31}$$

where we denote with $\| \cdot \|_{-1,K}^{\perp}$ the dual norm associated with $H_0^1(K)$, i.e.

$$\| \cdot \|_{-1,K}^{\perp} = \sup_{\Phi \in H_0^1(K) \setminus \{0\}} \frac{\langle \Phi, \cdot \rangle_{0,K}}{\|\Phi\|_{1,K}}.$$ 

Subsequently, we note that, owing to (31), we have:

$$\|\Delta v_n\|_{0,K} \lesssim p_K^2 \|\Delta v_n\|_{-1,K} = p_K^2 \sup_{\Phi \in H_0^1(K) \setminus \{0\}} \frac{\langle \Delta v_n, \Phi \rangle_{0,K}}{\|\Phi\|_{1,K}} = p_K^2 \sup_{\Phi \in H_0^1(K) \setminus \{0\}} \frac{\langle \nabla \Phi, \nabla v_n \rangle_{0,K}}{\|\Phi\|_{1,K}} \lesssim p_K^2 |v_n|_{1,K}. \tag{32}$$

As a consequence:

$$\int_K -\Delta v_n \Pi_{pk-2}^0 v_n \leq \|\Delta v_n\|_{0,K} \cdot \|\Pi_{pk-2}^0 v_n\|_{0,K} \leq p_K^2 \|\Pi_{pk-2}^0 v_n\|_{0,K}|v_n|_{1,K}. \tag{33}$$

Next, we turn our attention to the integral over $\partial K$ in (30). Applying a Neumann trace inequality (see e.g. [57, Theorem A33]):

$$\int_{\partial K} \frac{\partial v_n}{\partial n} \frac{\partial v_n}{\partial n} \leq \left\| \frac{\partial v_n}{\partial n} \right\|_{0,K} \lesssim (|v_n|_{1,K} + \|\Delta v_n\|_{0,K}) \left\| v_n \right\|_{1,K}. \tag{34}$$

Then, we use (32) on the second term in the first factor and a one dimensional $hp$ inverse estimate in addition to interpolation theory on the second factor (see [61, 62]), thus obtaining:

$$\int_{\partial K} \frac{\partial v_n}{\partial n} \frac{\partial v_n}{\partial n} \lesssim p_K^2 |v_n|_{1,K} \cdot p_K \| v_n \|_{0,\partial K}. \tag{35}$$

Plugging (33) and (34) in (30), we deduce:

$$|v_n|_{1,K}^2 \lesssim |v_n|_{1,K} \left\{ p_K^2 \|\Pi_{pk-2}^0 v_n\|_{0,K} + p_K^2 \|v_n\|_{0,\partial K} \right\},$$

whence

$$|v_n|_{1,K}^2 \lesssim p_K^2 \left( p_K^2 \|\Pi_{pk-2}^0 v_n\|_{0,K} \right) + p_K^5 \left( p_K \|v_n\|_{0,\partial K} \right) \lesssim p_K^5 S(K, v_n).$$}

Next, we estimate $c_*(K)$. Let $v_n \in \ker(\Pi_{pk}^\Sigma)$, then:

$$S(K, v_n) = p_K \|v_n\|_{0,\partial K} + p_K^2 \|\Pi_{pk-2}^0 v_n\|_{0,K} \lesssim p_K \|v_n\|_{0,\partial K} + p_K^2 \|v_n - \Pi_{pk}^\Sigma v_n\|_{0,K} \lesssim \frac{p_K}{p_K - 1} \|v_n\|_{1,K}^2. \tag{36}$$

We estimate the three terms separately. We begin with the first one. Applying the multiplicative trace inequality (see e.g. [31]), the $p$ version of the Aubin-Nitsche duality argument (27) for convex $K$ and (28) for nonconvex $K$:

$$p_K \|v_n\|^2_{0,\partial K} \lesssim p_K \left( \|v_n\|_{0,K} |v_n|_{1,K} + \|v_n\|_{0,K} \right) \lesssim p_K \left( |v_n - \Pi_{pk}^\Sigma v_n|_{0,K} |v_n|_{1,K} + \|v_n - \Pi_{pk}^\Sigma v_n\|^2_{0,K} \right) \lesssim \begin{cases} p_K \left( p_K^{-1} |v_n|_{1,K}^2 + p_K^{-2} \|v_n\|^2_{1,K} \right) \leq \frac{p_K^{-1}}{p_K} |v_n|_{1,K}^2, & \text{if } K \text{ is convex}, \\ p_K \left( \frac{p_K^{-1}}{p_K} |v_n|_{1,K}^2 + p_K^{-2} \frac{p_K^{-1}}{p_K} |v_n|_{1,K}^2 \right) \leq p^{-1} \frac{1}{p_K} |v_n|_{1,K}^2, & \text{otherwise}, \end{cases} \tag{37}$$

8
where we recall $\omega_K$ is the largest angle in $K$.

We now deal with the second term; using \cite[Lemma 4.1]{19}:

$$
p_K^2 \|v_n - \Pi_{p_K} v_n\|_{0,K}^2 \lesssim p_K^2 \|v_n - \Pi_{p_K} v_n\|_{1,K}^2 = \|v_n - \Pi_{p_K} v_n\|_{1,K}^2 \lesssim |v_n|_{1,K}^2,
$$

where in the last inequality we used that $v_n - \Pi_{p_K} v_n$ has zero average on $\partial K$.

Finally, we treat the third term; using Aubin-Nitsche argument \cite{27} and its modified version for nonconvex polygon \cite{28}:

$$
p_K^2 \|v_n - \Pi_{p_K} v_n\|_{0,K}^2 \lesssim \begin{cases} p_K^2 \|v_n - \Pi_{p_K} v_n\|_{1,K}^2 = |v_n|_{1,K}^2, & \text{if } K \text{ is convex}, \\ p_K^2 \|\Pi_{p_K} v_n\|_{1,K}^2 = p_K^2 (1 - \frac{\alpha}{4}) |v_n|_{1,K}^2, & \text{otherwise}. \end{cases}
$$

Collecting the three bounds, we obtain the claim. \hfill \Box

**Remark 3.** In order to keep the notation simpler, we proved Theorem \ref{thin} assuming that the polynomial degrees $p_E$ on each edge $E \in \mathcal{E}_K$ coincide with the degree of accuracy $p_K$ of the local space $V(K)$; the same result remains valid if $p_K \approx p_E$, for all $E \in \mathcal{E}_K$. In view of the forthcoming definition \cite{57}, the case of interest in the following will satisfy such condition and therefore, for the proof of the main result of this work, namely Theorem \ref{main}, we will not use directly Theorem \ref{thin} but its nonuniform degree version.

As a consequence of Theorem \ref{thin}, the quantity $\alpha(K)$, defined in \cite{24}, can be bounded in terms of $p_K$ as follows:

$$
\alpha(K) = \frac{1 + \alpha^*(K)}{\min_{K \in T_n} \alpha_*(K')} = \frac{1 + \max(1, c^*(K))}{\min_{K \in T_n} (\min(1, c(K')))} \lesssim \begin{cases} \max_{K \in T_n} p_K^2, & \text{if all } K \text{ are convex,} \\ 2(1 - \frac{\alpha}{4}) \max_{K \in T_n} p_K^2, & \text{otherwise}. \end{cases}
$$

**Remark 4.** Owing to \cite{29} formula (2.14)], we could replace the boundary term of $S^v_K$, defined in \cite{20}, with a spectrally equivalent algebraic expression employing Gauß-Lobatto nodes. In particular, let $\hat{I} = [-1, 1]$ and let $\{p_j^{p_j+1}\}_{j=0}^{p_j}$ and $\{\xi_j^{p_j+1}\}_{j=0}^{p_j}$ be the Gauß-Lobatto nodes and weights on $\hat{I}$ respectively. Then:

$$
c \sum_{j=0}^{p_j} q^2 (\xi_j^{p_j+1}) p_j^{p_j+1} \leq ||q||_{0, \hat{I}}^2 \leq \sum_{j=0}^{p_j} q^2 (\xi_j^{p_j+1}) p_j^{p_j+1}, \quad \forall q \in P_{p_j}(\hat{I}),
$$

where $c$ is a positive universal constant. We could replace in \cite{20} the $L^2$ integral on the boundary with a piecewise Gauß-Lobatto combination, mapping each edge on the reference interval $\hat{I}$ and using \cite{38}, the advantage of such a choice is that we can automatically use the nodal degrees of freedom on the skeleton, assuming that they have a Gauß-Lobatto distribution on each edge.

The boundary term of the new stabilization is now very close to the classical stabilization choice (see e.g. \cite{13} and \cite{19}) and its implementation is much easier than the implementation of \cite{20}, where one should reconstruct polynomials on each edge; in fact, it suffices to take instead of the Euclidean inner product of all the degrees of freedom only the boundary one with some Gauß-Lobatto weights.

For additional issues concerning the stabilization (only for the $h$ version of VEM) see \cite{21}, while for more details concerning the implementation of the method we refer to \cite{13}.

### 4.1 Numerical tests for the stability bounds

In Theorem \ref{thin} we proved the stability bounds \cite{20} for a possible choice of $S^v_K$. Such bounds, which also reflect on $\alpha_*(K)$ and $\alpha^*(K)$ introduced in \cite{17}, are rigorously proven but have a quite strage dependence on $p$. In the following, we check numerically whether the dependence on $p$ of the above-mentioned constants is sharp.

In order to do that, we note that finding $\alpha_*(K)$ and $\alpha^*(K)$ in \cite{17} is equivalent to find the minimum and maximum eigenvalues $\lambda_{\min}$ and $\lambda_{\max}$ of the generalized eigenvalue problem:

$$
A_n^K v_n = \lambda A^K v_n, \quad v_n \in V(K),
$$

where we recall $\omega_K$ is the largest angle in $K$.

We now deal with the second term; using \cite[Lemma 4.1]{19}:

$$
p_K^2 \|v_n - \Pi_{p_K} v_n\|_{0,K}^2 \lesssim p_K^2 \|v_n - \Pi_{p_K} v_n\|_{1,K}^2 = \|v_n - \Pi_{p_K} v_n\|_{1,K}^2 \lesssim |v_n|_{1,K}^2,
$$

where in the last inequality we used that $v_n - \Pi_{p_K} v_n$ has zero average on $\partial K$.

Finally, we treat the third term; using Aubin-Nitsche argument \cite{27} and its modified version for nonconvex polygon \cite{28}:

$$
p_K^2 \|v_n - \Pi_{p_K} v_n\|_{0,K}^2 \lesssim \begin{cases} p_K^2 \|v_n - \Pi_{p_K} v_n\|_{1,K}^2 = |v_n|_{1,K}^2, & \text{if } K \text{ is convex,} \\ p_K^2 \|\Pi_{p_K} v_n\|_{1,K}^2 = p_K^2 (1 - \frac{\alpha}{4}) |v_n|_{1,K}^2, & \text{otherwise}. \end{cases}
$$

Collecting the three bounds, we obtain the claim. \hfill \Box

**Remark 3.** In order to keep the notation simpler, we proved Theorem \ref{thin} assuming that the polynomial degrees $p_E$ on each edge $E \in \mathcal{E}_K$ coincide with the degree of accuracy $p_K$ of the local space $V(K)$; the same result remains valid if $p_K \approx p_E$, for all $E \in \mathcal{E}_K$. In view of the forthcoming definition \cite{57}, the case of interest in the following will satisfy such condition and therefore, for the proof of the main result of this work, namely Theorem \ref{main}, we will not use directly Theorem \ref{thin} but its nonuniform degree version.

As a consequence of Theorem \ref{thin}, the quantity $\alpha(K)$, defined in \cite{24}, can be bounded in terms of $p_K$ as follows:

$$
\alpha(K) = \frac{1 + \alpha^*(K)}{\min_{K \in T_n} \alpha_*(K')} = \frac{1 + \max(1, c^*(K))}{\min_{K \in T_n} (\min(1, c(K')))} \lesssim \begin{cases} \max_{K \in T_n} p_K^2, & \text{if all } K \text{ are convex,} \\ 2(1 - \frac{\alpha}{4}) \max_{K \in T_n} p_K^2, & \text{otherwise}. \end{cases}
$$

**Remark 4.** Owing to \cite{29} formula (2.14)], we could replace the boundary term of $S^v_K$, defined in \cite{20}, with a spectrally equivalent algebraic expression employing Gauß-Lobatto nodes. In particular, let $\hat{I} = [-1, 1]$ and let $\{p_j^{p_j+1}\}_{j=0}^{p_j}$ and $\{\xi_j^{p_j+1}\}_{j=0}^{p_j}$ be the Gauß-Lobatto nodes and weights on $\hat{I}$ respectively. Then:

$$
c \sum_{j=0}^{p_j} q^2 (\xi_j^{p_j+1}) p_j^{p_j+1} \leq ||q||_{0, \hat{I}}^2 \leq \sum_{j=0}^{p_j} q^2 (\xi_j^{p_j+1}) p_j^{p_j+1}, \quad \forall q \in P_{p_j}(\hat{I}),
$$

where $c$ is a positive universal constant. We could replace in \cite{20} the $L^2$ integral on the boundary with a piecewise Gauß-Lobatto combination, mapping each edge on the reference interval $\hat{I}$ and using \cite{38}, the advantage of such a choice is that we can automatically use the nodal degrees of freedom on the skeleton, assuming that they have a Gauß-Lobatto distribution on each edge.

The boundary term of the new stabilization is now very close to the classical stabilization choice (see e.g. \cite{13} and \cite{19}) and its implementation is much easier than the implementation of \cite{20}, where one should reconstruct polynomials on each edge; in fact, it suffices to take instead of the Euclidean inner product of all the degrees of freedom only the boundary one with some Gauß-Lobatto weights.

For additional issues concerning the stabilization (only for the $h$ version of VEM) see \cite{21}, while for more details concerning the implementation of the method we refer to \cite{13}.
Here, $A^K_n$ and $A^K \in \mathbb{R}^{\dim(V(K)) \times \dim(V(K))}$ are defined as

$$(A^K_n)_{i,j} = a^K_n(\varphi_i, \varphi_j), \quad (A^K)_{i,j} = a^K(\varphi_i, \varphi_j)$$

where $\{\varphi_i\}_{i=1}^{\dim(V(K))}$ denotes the virtual canonical basis of $V(K)$, see (11). We are adopting the usual notation, by calling $v_n \in \mathbb{R}^{\dim(V(K))}$ the vector of the degrees of freedom associated with $v_n \in V(K)$.

We note that we restrict our analysis on functions having zero average on $K$, since both $A^K_n$ and $A^K$ have constant functions in their kernel; this strategy allows to avoid the problems related to solving the generalized eigenvalue problem for singular matrices. Moreover, the entries of matrix $A^K$ are not computable exactly, since virtual functions are not known explicitly; therefore, we approximate them by solving numerically the associated diffusion problem, by means of a fine and high-order finite element approximation.

In Table 1, we present the results on three different types of polygon (namely, those which we will employ for the tests in the forthcoming Section 6): a square, a nonconvex decagon (like any of the polygons in the outer layer of Figure 1 b), a nonconvex hexagon (like any of the polygons in the outer layer of Figure 1 c).

| $p$ | sq. $\lambda_{\text{min}}$ | sq. $\lambda_{\text{max}}$ | dec. $\lambda_{\text{min}}$ | dec. $\lambda_{\text{max}}$ | hex. $\lambda_{\text{min}}$ | hex. $\lambda_{\text{max}}$ |
|-----|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| 2   | 7.8559e-01      | 1.0000e+00      | 7.9262e-02      | 5.5516e+00      | 1.6168e-01      | 1.1183e+00      |
| 3   | 4.6667e-01      | 1.0000e+00      | 1.0396e-01      | 8.6605e+00      | 1.3342e-01      | 1.4751e+00      |
| 4   | 3.3195e-01      | 1.0000e+00      | 4.5039e-02      | 1.0852e+01      | 1.0321e-01      | 1.6253e+00      |
| 5   | 2.7547e-01      | 1.0000e+00      | 3.4944e-02      | 1.0513e+01      | 7.4247e-02      | 1.8672e+00      |
| 6   | 2.1557e-01      | 1.0000e+00      | 2.3463e-02      | 1.1835e+01      | 5.5556e-02      | 1.6707e+00      |
| 7   | 1.8994e-01      | 1.0000e+00      | 2.0730e-02      | 9.7514e+00      | 3.5664e-02      | 1.9013e+00      |
| 8   | 1.4136e-01      | 1.0000e+00      | 1.6122e-02      | 1.0447e+01      | 2.7559e-02      | 1.8801e+00      |
| 9   | 1.2446e-01      | 1.0000e+00      | 1.8555e-02      | 7.9781e+00      | 2.1313e-02      | 1.8337e+00      |
| 10  | 9.2933e-02      | 1.0000e+00      | 1.3736e-02      | 3.9577e+01      | 1.7991e-02      | 5.6544e+00      |

Table 1: Minimum and maximum eigenvalues of the generalized eigenvalue problem (39) on: sq. = a square; dec. = a nonconvex decagon; hex. = a nonconvex hexagon.

As theoretically expected, the maximum generalized eigenvalue always scales like 1. On the contrary, the minimum eigenvalue behaves in all the three cases like $p^{-1}$. This means that in fact the bounds of Theorem 4.1 are abundant, whereas the actual behaviour of the stability bounds may be much milder. Unfortunately, currently we are not able to improve the stability bounds of Theorem 4.1. It is worth mentioning that this has no impact on the asymptotic exponential convergence results in the next section.

5 Exponential convergence for corner singularity on geometric meshes

In this section, we want to show that exponential convergence is achieved if geometric mesh refinement and degree of accuracy distribution are chosen appropriately.

In order to achieve such a convergence we employ geometrically graded polygonal meshes, which will be discussed in Section 5.1. Then, we show in Section 5.2 estimates for the first and the second terms in the error decomposition (24), in particular proving bounds for the local right-hand side approximation and for the local best approximation by means of polynomials. In Section 5.3 we obtain estimates for the third term in (25), in particular illustrating bounds for the best approximation by means of functions belonging to the virtual space defined in (12). Finally, in Section 5.4 under a proper choice for the polynomial degree vector $p$ introduced in Section 3 and the sequence $\{T_n\}_n$ of polygonal decompositions, we combine together the above error bounds; as a consequence, we guarantee exponential convergence for the error in the energy norm in terms of the number of degrees of freedom of the global virtual space $V_n$ defined in (12).
5.1 Geometric meshes

Here, we describe a class of sequences of nested geometric meshes which we will employ later in order to show error convergence. We recall we are assuming that the only “singular” corner is the origin $0 \in \partial \Omega$, see [5]. Let $\sigma \in (0, 1)$ be the grading parameter of the mesh.

The decomposition $T_n$ consists of $n + 1$ layers defined as follows. We set $L_0$ the 0-th layer as the set of all polygons $K$ in decomposition $T_n$ such that $0 \in \mathcal{V}_n^K$; next, we define by induction:

$$L_j = \left\{ K_1 \in T_n \mid \overline{K_1} \cap \overline{K_2} \neq \emptyset, \text{ for some } K_2 \in L_{j-1}, K_1 \notin \bigcup_{i=0}^{j-1} L_i \right\}, \quad j = 1, \ldots, n. \quad (40)$$

We set $T_0 = \{ \Omega \}$. Given $T_n$, the decomposition $T_{n+1}$ is obtained by refining $T_n$ only in the layer around the singularity (i.e. $L_0$). We require that at level $n$, the decomposition satisfies the following grading condition:

$$(D3) \quad h_K \approx \begin{cases} \frac{1 - \sigma}{\sigma} \text{dist}(0, K), & \text{if } K \notin L_0 \\ \sigma^n, & \text{otherwise.} \end{cases} \quad (41)$$

Furthermore, the number of elements in each layer is uniformly bounded with respect to the discretization parameters. We will also assume that $p_K \geq 2$. A more precise choice will be discussed in the forthcoming definition [57].

Assumption (D3) justifies the name geometric for the sequence; more specifically, the closer a polygon is to 0 the smaller its diameter is. Moreover, it is possible to check that the ratio between the size of two neighbouring layers is proportional to $\frac{1}{\sigma^n}$. As a consequence of assumption (D3), we also have, for $K \in L_j$, $h_K \approx \sigma^{n-j}$.

Example 5.1. A possible sequence satisfying (D1)-(D3) is the graded mesh of squares elements with hanging nodes on the $L$-shaped domain, that is used in [57, Definition 4.30], see Figure 1 (left). We note that in the VEM context, this mesh contains pentagons and squares, whereas in the Finite Element counterpart the very same mesh is “afflicted” by the presence of squares with hanging nodes.

Example 5.2. Another choice is depicted in Figure 1 (center). This mesh is obtained by merging all the elements that correspond to one layer in the mesh from Example 5.1 in a single large element. We observe that this mesh is made of $n$ decagons and one hexagon around 0. Moreover, we want to stress the fact that this mesh, that cannot be used in the conforming FEM environment, needs less then one third of the degrees of freedom of the previous one. Finally, we point out that such a mesh does not satisfy the star-shapedness assumption (D1).

Example 5.3. As a third example, see Figure 1 (right), we modify the mesh in Example 5.2 by adding an oblique cut on the “central” diagonal. This mesh still cannot be used for conforming FEM approximations. Obviously, it contains more elements and hence will result in more degrees of freedom than the mesh in Example 5.2. Notwithstanding, it satisfies (D1). Moreover, it satisfies also the technical assumption (D4) we will need in what follows, whereas the mesh in Example 5.2 does not.

We require the following additional assumption on the geometry of the decomposition. We will need it to state approximation results in Sections 5.2 and 5.3.

(D4) Let $T_n$ be a geometric polygonal decomposition; write $T_n = T_n^0 \cup T_n^1$, where $T_n^0 = L_0$ and $T_n^1 = \bigcup_{j=1}^n L_j$. Then, there exists a collection $C_n^1$ of squares such that:

- card($C_n^1$) = card($T_n^1$); for each $K \in T_n^1$, there exists $Q = Q(K) \in C_n^1$ such that $Q \supset K$ and $h_K \approx h_Q$; in addition, it must hold dist$(0, Q(K)) \approx h_K$;
- every $x \in \Omega$ belong at most to a fixed number of squares $Q$, independently on all the discretization parameters;
We point out that (D4) seems to be a rather technical requirement. Indeed, we will show in Section 6 that also meshes not satisfying (D4) may produce the expected convergence behaviour shown in Theorem 5.7.

We note that (D4) is in the spirit of the strategy of the overlapping square technique used in [19,35]. We here additionally require that squares covering polygons far from the singularity cannot cover also such a singularity (since in this case p approximation results would not hold, thus invalidating Theorem 5.7). We also stress that the decomposition in Example 5.2 does not satisfy neither (D1) nor (D4). Finally, we point out that instead of considering a decomposition of squares $C_n$, it is possible to consider in (D4) a decomposition in sufficiently regular quadrilaterals (e.g. parallelograms), since the same analysis by means of Legendre polynomials that follows (for instance in Lemmata 5.1 and 5.3) could be performed.

### 5.2 Local approximation by polynomials

Here, we deal with the approximation of the first and the second term in the right-hand side (25). What we are going to prove are $hp$ approximation properties by means of local polynomials on polygons. In $hp$-FEM literature, classical approximation of this type is not effectuated on general polygons but only on squares and triangles, see [9,10,46,49,57] and the references therein.

The basic tool behind this approach is the employment of orthogonal bases, namely tensor product of Legendre polynomials on the square, see [57], and Koornwinder polynomials (that is collapsed tensor product of Jacobi polynomials) on triangles, see [12,15]; with such basis, explicit computations can be performed, owing to properties of Legendre and Jacobi polynomials. On a generic polygon an explicit basis with good approximation properties is not available.

The error analysis follows the lines of [19,57] and is summarized below. Let $p$ be the vector of the local degree of accuracy on each polygon. We recall that we denote with $S^{p,-1}(T_n,\Omega)$ the space of piecewise discontinuous polynomials over the decomposition $T_n$ of degree $p_K$ on each polygon $K$.

The first result is a polynomial approximation estimate regarding regular functions on polygons far from the singularity. This result will be used for the approximation of the local second term in (25) for the elements $K$ separated from the singularity.

**Lemma 5.1.** Under assumptions (D1)-(D4), let $K \in L_j$, $j = 1, \ldots, n$. Let $Q(K)$ be defined in (D4) and let $u \in H^{s_K+3,2}_{\beta}(Q(K))$, $1 \leq s_K \leq p_K$. Then, there exists $\Phi \in \mathbb{P}_{p_K}(Q(K))$ such that:

$$
\|D^m(u - \Phi)\|_{0,K}^2 \lesssim \sigma^{2(n-j)(2-m-\beta)} \frac{\Gamma(p_K - s_K + 1)}{\Gamma(p_K + s_K + 3 - 2m)} \left( \frac{\rho}{2} \right)^{2s_K} |u|_{H^{s_K+3,2}_{\beta}(Q(K))}^2
$$

where $m = 0, 1, 2; 2 \leq j \leq n + 1; \rho = \max(1, \frac{1-\sigma}{\sigma})$, $\sigma$ is the grading parameter of the mesh and $\Gamma$ is the Gamma function.

---

**Figure 1:** Decomposition $T_n$, $n = 3$, for Example 5.1 (left), Example 5.2 (center), Example 5.3 (right).

- $\forall K \in T_n^0$, $K$ is star-shaped with respect to $0$; moreover, the subtriangulation of $K$ obtained by joining $0$ with the other vertices is uniformly shape regular ($\gamma$ being the shape-regularity constant).

We set $\Omega^\text{ext}_n = (\cup_{Q \subseteq C_1^2} Q) \cup (\cup_{K \in T_n^0} K)$.

We note that (D4) is in the spirit of the strategy of the overlapping square technique used in [19,35]. We here additionally require that squares covering polygons far from the singularity cannot cover also such a singularity (since in this case $p$ approximation results would not hold, thus invalidating Theorem 5.7). We also stress that the decomposition in Example 5.2 does not satisfy neither (D1) nor (D4). Finally, we point out that instead of considering a decomposition of squares $C_n$, it is possible to consider in (D4) a decomposition in sufficiently regular quadrilaterals (e.g. parallelograms), since the same analysis by means of Legendre polynomials that follows (for instance in Lemmata 5.1 and 5.3) could be performed.
Proof. The result follows from classical scaling arguments and \[57\] Lemma 4.53. Here, we only give the idea of the proof. Firstly one encapsulates polygon $K$ into the corresponding square $Q(K)$. It is possible to bound the left hand side of inequality (42) with the same (semi)norm on the square. After that, the square is mapped into the reference square $\hat{Q} = [-1,1]^2$ and a $p$ analysis by means of tensor product of Legendre polynomials is developed (see \[57\] Theorem 4.46). Subsequently, the reference square is pushed forward to square $\hat{Q}$. Using the property of the geometric mesh stated in assumption (D3) and \[57\] Lemma 5.40, the result follows.

Estimate on polygons around the singularity are discussed in the following lemma. We point out that for the error control in layer $L_0$ we can work directly on the element without the need of employing covering squares, as effectuated for the analysis on the polygons of the other layers, see Lemma \[57\]. The proof is an extension to polygonal domains of that in Theorem \[57\] Lemma 14.

Lemma 5.2. Under assumptions (D1)-(D4), let $K \in L_0$. Let $u \in H^2_\alpha(K)$, $\beta \in [0,1)$. Then, there exists $\Phi \in \mathbb{P}_1(K)$ such that:

$$|u - \Phi|^2_{2,K} \lesssim h^2_{K}(1 - \beta) \| |x|^\beta |D^2u| \|_{0,K}^2 \lesssim \sigma^2(1 - \beta)^n \| |x|^\beta |D^2u| \|_{0,K}^2,$$

(43)

where $\sigma$ is the grading factor from assumption (D3).

Proof. We start by proving the following Hardy inequality on polygons with a vertex at 0. Let $\alpha > 0$, let be given a function $u$ such that $\int_K |x|^{\alpha} |D^2u| < +\infty$ and $u \in C^0(\overline{K})$. Then:

$$\int_K |x|^{\alpha-2} |u - u(0)|^2 \leq c \int_K |x|^\alpha |D^1u|^2.$$

(44)

We consider the regular subtriangulation by joining 0 with the other vertices of $K$; the existence of such a decomposition is guaranteed by assumption (D4). Thanks to \[57\] Lemma 4.18, the “triangular” counterpart of (44) holds:

$$\int_T |x|^{\alpha-2} |u - u(0)|^2 \leq c \int_T |x|^\alpha |D^1u|^2, \quad \forall T \text{ in the subtriangulation of } K.$$

(45)

It suffices then to split the integral over $K$ into a sum of integrals over the triangles of the subtriangulation, apply (45) and collect all the terms.

Using (44) and applying the argument of \[57\] Lemma 4.19 to the polygon $K$, we observe that $H^2_\beta(K)$ is compactly embedded in $H^1(K)$. Using such a compact embedding and proceeding as in \[57\] Lemma 14.6, the following inequality holds true for a polygon $K$ star-shaped with respect to 0:

$$|U|^2_{1,K} \lesssim h^2_{K}(1 - \beta) \| |x|^\beta |D^2U| \|_{0,K}^2 + \sum_{i=1}^3 |U(A_i)|^2, \quad \forall U \in H^2_{\beta}(K),$$

(46)

where $\{A_i\}_{i=1}^3$ is a set of three arbitrary nonaligned vertices of $K$. Let $\Phi$ be the linear interpolant of $u$ at $A_i$, $i = 1, \ldots, 3$. Then, plugging $U = u - \Phi$ in (46), noting that $U(A_i) = 0$, $i = 1, 2, 3$, and using the geometric assumption (D3), we get the claim.

We note that (43) does not rely on $p$ approximation results, but only on scaling argument. This will be enough in order to prove the main result of this work, that is Theorem \[57\] and it is in accordance with the choice of the vector of local degrees of accuracy that will be effectuated in the forthcoming definition \[57\]. We emphasize that this is in the spirit of classical $hp$ refinement, see \[57\].

We now turn our attention to the approximation of the first local term in (25), i.e. to the local approximation of the loading term. Since we are approximating it with piecewise polynomials of local degree $p_K - 2$, we set $\overline{p} = p - 2$, i.e. $\forall K \in T_n$, $\overline{p}_K = p_K - 2$. We have, for all $v_n \in V_n$:

$$(f_n, v_n)_{0,K} - (f, v_n)_{0,K} = \sum_{K \in T_n} \int_K (\Pi^0_{p-2,K} f - f)(v_n - \Pi^0_{0,K} v_n) =: \sum_{K \in T_n} F_K(v_n),$$

(47)
where we recall we are assuming for the sake of simplicity $p_K \geq 2$, $\forall K \in T_n$, see Section 5.

As above, we develop a different analysis for polygons near and far from the singularity. We start with the “far” case.

**Lemma 5.3.** Under assumptions (D1)-(D4), let $K \in L_j$, $j = 1, \ldots, n$. Let $Q(K)$ be defined in (D4). Let $f \in H^{7\beta+3.2}_0(Q(K))$, $0 \leq \sigma_K \leq p_K$, with $p_K = p_{K'} - 2$. Then, for all $v_n \in V(K)$,

\[
F_K(v_n) \leq \|v_n\|_{1,K} \left( \frac{1}{\sigma} \right)^{\frac{1}{2}} \left( \frac{1}{2} \right)^\tau_p \left( \frac{\rho}{2} \right) \|f\|_{H^{7\beta+3.2}_0(Q(K))}
\]

with the same notation of Lemma 5.1.

**Proof.** It suffices to use a Cauchy-Schwarz inequality in (47), standard bounds for the projection errors and analogous estimate to those in Lemma 5.2. □

Assume now that $K$ is an element in the finest level $L_0$. We work here a bit differently from what we did in Lemma 5.2. In particular we get the following.

**Lemma 5.4.** Under assumptions (D1)-(D3), let $K \in L_0$. Assume $f \in L^2(K)$. Let $\beta \in [0,1]$. Then:

\[
F_K(v_n) \leq h_{K}^{1-\beta} \|v_n\|_{1,K} \|f\|_{0,K} \lesssim \sigma^{n(1-\beta)} \|v_n\|_{1,K}, \quad \forall v_n \in V(K),
\]

where $\sigma$ is the grading factor from assumption (D3).

**Proof.** Using a Cauchy-Schwarz inequality and Bramble Hilbert lemma (see (31)), we obtain:

\[
F_K(v_n) \lesssim h_{K} \|v_n\|_{1,K} \|f\|_{0,K} \lesssim h_{K}^{1-\beta} \|v_n\|_{1,K} \|f\|_{0,K} \lesssim \sigma^{n(1-\beta)} \|v_n\|_{1,K}, \quad \forall v_n \in V(K).
\]

We point out that for the proof of Lemmata 5.2 and 5.4 we work directly on the polygon without the need of using the covering squares technique of assumption (D4), like in Lemmata 5.1 and 5.3. This justifies the fact that in assumption (D4) we did not require the existence of a collection of squares $C_0^p$ associated with the finest layer $L_0$ but only the existence of collection $C_1^p$ associated with all the other layers.

### 5.3 Approximation by functions in the virtual space

Here, we treat the approximation of the third term in the right-hand side of (25). We observe that this term has two main differences with respect to the other two. The first difference is that we need an approximant $u_I$ which is globally continuous; the second one is that $u_I$ is not a piecewise polynomial but a function belonging to the virtual space $V_n$.

As done in Section 5.2, we split the analysis into two parts. Firstly, we work on polygons abutting the singularity, see Lemma 5.5; secondly, we work on elements $K$ in the first layer $L_0$, see Lemma 5.6.

**Lemma 5.5.** Let assumptions (D1)-(D4) hold. Let $K \in L_j$, $j = 1, \ldots, n$. Let $f$, the right-hand side of (6), belong to space $B^0_2(\Omega)$; consequently, $u$, the solution of problem (6), belongs to space $B^2_2(\Omega)$, see Theorem 2.1 and definition 1. Assume that $p_{K'} \approx p_{K'}$ if $E \in \partial K$. Assume moreover that if $K \in L_1$, then $p_K \approx 2$. Then, for all $1 \leq s_K \leq p_K$, there exists $u_I \in V(K)$ such that:

\[
|u - u_I|^2_{1,K} \lesssim \|f - \Pi^0_{p_{K'} - 2}f\|_{0,K}^2 \left( \frac{\rho \rho_{p_{K'},-1}}{2} \right) + \sigma^{n(3-2\beta)} p_K^{-2s_K-1} \left( \frac{\rho \rho_{p_{K'},-1}}{2} \right)^{2s_K+1} \sum_{E \in E_0} |u_I|^2_{H^{s_K+1}_0(E)},
\]

where we recall that $\Pi^0_{p_{K'} - 2}$ is the $L^2(K)$ orthogonal projection from $V(K)$ into $\Pi^0_{p_{K'} - 2}(K)$, $\sigma$ is the grading factor from assumption (D3) and $\rho = \max \left( 1, \frac{1-p_{K'}}{\sigma} \right)$. □
Proof. Before starting the proof, we observe that the boundary norm in the right-hand side of (48) exists, since \( u \in B^2_0(\Omega) \) implies that \( u \in H^1(K) \) for all \( t \in \mathbb{N} \) and polygons \( K \neq L_0 \).

We define \( u_I \) as the weak solution of the following problem:

\[
\begin{aligned}
-\Delta u_I &= \Pi_{pE-2}^0 f \quad \text{in } K, \\
u_I &= \pi u \quad \text{on } \partial K,
\end{aligned}
\]

where \( \pi u \in \mathbb{B}(\partial K) \), see [3], is defined in the following way. Assume for the time being that \( K \notin L_1 \). Let \( \tilde{I} = [-1, 1] \). Given an edge \( E \subseteq \partial K \), \( \pi u \) is defined as the push-forward of a function \( \pi u \in \mathbb{P}_p(\tilde{I}) \) which we fix as follows. Let \( \tilde{u} \) be the pull-back of \( u|_E \) on \( \tilde{I} \). Then, \( \pi u' \) is the Legendre expansion of \( \tilde{u} \) up to order \( pE - 1 \). In particular, we write:

\[
\tilde{u}'(\xi) = \sum_{i=0}^{\infty} c_i L_i(\xi), \quad \pi u' (\xi) = \sum_{i=0}^{pE-1} c_i L_i(\xi).
\]

Here \( \{ L_i(\xi) \}_{i=0}^\infty \) is the \( L^2(\tilde{I}) \) orthogonal basis of Legendre polynomials, with \( L_i(-1) = (-1)^i \) and \( L_i(1) = 1 \). Next, we define \( \pi u \) as:

\[
\pi u(\xi) = \int_{-1}^\xi \pi u'(\eta) d\eta + \tilde{u}(-1).
\]

It is possible to prove that \( \pi u \) interpolates \( \tilde{u} \) at the endpoints of \( \tilde{I} \) using the definition of \( \pi u \) and the fundamental theorem of calculus. Recalling [57] Theorem 3.14 and using simple algebra, the following holds true:

\[
\| \tilde{u} - \pi u \|_{\ell, \tilde{I}} \lesssim e^{s_k} p_{pE}^{-s_k-1+\ell} |u|_{j, K+1, \ell}, \quad \ell = 0, 1, \quad \forall 1 \leq s_k \leq p_K.
\]

Applying a scaling argument, interpolation theory (see [61,62]) and summing on all the edges, we get:

\[
\| u - \pi u \|_{\dot{H}^s_K} \lesssim e^{2s_k + 1} \left( \frac{h_E}{p_E} \right)^{2s_k + 1} \sum_{E \in \mathcal{E}^K} |u|_{j, K+1, E}, \quad \forall 1 \leq s_k \leq p_K.
\]

If now \( K \in L_1 \), we define \( \pi u|_E \) as above if \( E \) does not belong to the interface between \( L_0 \) and \( L_1 \), otherwise \( u_I \) is defined as the linear interpolant of \( u \) at the two endpoints of \( E \). We point out that (52) remains valid also if \( K \in L_1 \) paying an additional constant \( c^{2s_k+1} \), since \( p_K \approx 2 \) whenever \( K \in L_1 \). We also note that (52) implies, recalling that \( p_E \approx p_K \) if \( E \subseteq \partial K \) and following the ideas in [57] Lemma 3.39:

\[
\| u - u_I \|_{\dot{H}^{s_K+1/2}(K)} \approx \| u - \pi u \|_{\dot{H}^{s_K+1/2}(K)} \approx \sum_{E \in \mathcal{E}^K} |u|_{j, K+1, E}^2,
\]

where we recall that \( j \) denotes the number of the layer to which \( K \) belongs.

We are now ready to prove the error estimate. For arbitrary constants \( c_1 \) and \( c_2 \), there holds (also recalling that \( f - \Pi_{pK-2}^0 f \) is \( L^2 \)-orthogonal to constants):

\[
\| u - u_I \|_{1,K}^2 = \int_K |\nabla(u - u_I - c1)|^2 = \int_{\partial K} \frac{\partial (u - u_I)}{\partial n} (u - u_I - c1) - \int_K (f - \Pi_{pK-2}^0 f) (u - u_I - c2)
\]

\[
\leq \left\| \frac{\partial (u - u_I)}{\partial n} \right\|_{\dot{H}^{s_K+1/2}(K)} \| u - \pi u - c_1 \|_{\dot{H}^{s_K+1/2}(K)} + \| f - \Pi_{pK-2}^0 f \|_{0,K} \| u - u_I - c_2 \|_{0,K}.
\]

Applying the trace inequalities on Neumann and Dirichlet traces, choosing \( c_2 \) to be the average of \( u - \pi u \) on \( K \) and applying a Poincaré inequality, we get:

\[
\| u - u_I \|_{1,K}^2 \lesssim \| (u - u_I)|_{1,K} + \| f - \Pi_{pK-2}^0 f \|_{0,K} \| u - \pi u - c_1 \|_{\dot{H}^{s_K+1/2}(K)} + \| f - \Pi_{pK-2}^0 f \|_{0,K} \| u - u_I - c_2 \|_{0,K}.
\]

\[
\lesssim \| u - u_I - c_1 \|_{1,K} \left\{ \| f - \Pi_{pK-2}^0 f \|_{0,K} + \| u - u_I - c_1 \|_{\dot{H}^{s_K+1/2}(K)} \right\}.
\]
We deduce, picking $c_1$ to be the average of $u - u_I$ on $\partial K$ and applying a Poincaré inequality:

$$|u - u_I|^2_{1,K} \lesssim \|f - \Pi_{pK-2}^0 f\|_{0,K}^2 + \|u - \pi u\|^2_{2,\partial K}.$$  

In order to conclude, it suffices to apply \[53\].

We turn now our attention to the approximation on the polygons abutting the singularity.

**Lemma 5.6.** Let assumptions (D1)-(D4) hold. Let $f$, the right-hand side of (2), belong to space $B^0_{\beta}(\Omega)$; consequently, $u$, the solution of problem (2), belongs to space $B^2_{\beta}(\Omega)$, see Theorem 2.1 and definition [1]. Assume that $p_K = 2$ if $K \in L_0$ and $p_K \approx 2$ if $K \in L_1$. Then there exists $u_I \in V(K)$ such that:

$$|u - u_I|^2_{1,K} \lesssim \|f - \Pi_{pK-2}^0 f\|_{0,K}^2 + \|u - u_I - c_1\|^2_{2,\partial K},$$

where we recall that $\Pi_{pK-2}^0$ is the $L^2(\Omega)$ orthogonal projection from $V(K)$ into $P_{pK-2}(\Omega)$, $\sigma$ is the grading factor discussed in assumption (D3) and $n + 1$ is the number of layers.

**Proof.** We consider $u_I$ defined as in (19); in particular, we fix $\pi u$, the trace of $u_I$ on $\partial K$ to be the piecewise affine interpolant of $u$ at the vertices of $K$. From Lemma 5.5 we have:

$$|u - u_I|^2_{1,K} \lesssim \|f - \Pi_{pK-2}^0 f\|_{0,K}^2 + \|u - u_I - c_1\|^2_{2,\partial K},$$  

where $c_1$ is the average of $u - u_I$ on $\partial K$.

In order to get the claim, it suffices to bound the second term. As in Lemma 5.2 we consider the subtriaingular $\tilde{T}_n = \tilde{T}_n(K)$ of $K$ obtained by connecting all the vertices of $K$ to 0, see assumption (D4). In particular, every triangle $T \in \tilde{T}_n$ is star-shaped with respect to a ball of radius $\tilde{\gamma}_h$, where $\tilde{\gamma}$ is a positive universal constant. We define $\tilde{u}_K$ as the piecewise linear interpolant polynomials over the triangular subtriaangular, interpolating $u$ at the vertices of $T$, for every $T \in \tilde{T}_n$. Using [57] Lemma 4.16 and applying a Poincaré inequality, yield to:

$$\|u - u_I - c_1\|^2_{2,K} \lesssim \|u - \tilde{u}_K - c_1\|^2_{1,K} \lesssim \sum_{T \in \tilde{T}_n} |u - \tilde{u}_K|^2_{1,T} \lesssim \sum_{T \in \tilde{T}_n} h_T^{2(1-\beta)} \|\pi x|D^2 u\|^2_{0,T} \lesssim \sigma^{2(1-\beta)} \|x|D^2 u\|^2_{0,K}.$$  

We stress that the third inequality in (55) holds since $\tilde{u}_K|_T$ is a linear polynomial and therefore $D^2\tilde{u}_K = 0$ on all $T \in \tilde{T}_n$. \[55\]

We note that in Lemmata 5.3 and 5.4 the error between $f$ and its $L^2$ projection can be bounded using Lemmata 5.3 and 5.4. We also point out that the hypothesis concerning the distribution of the local degrees of accuracy, i.e. the fact that $p_K = 2$ if $K \in L_0$, $p_K \approx 2$ if $K \in L_1$, $p_K \approx p_K$ if $E \subset \partial K$, are in accordance with the forthcoming definition [57] that we will introduce for the proof of Theorem 5.7. Finally, we point out in Lemmata 5.5 and 5.6 we introduced a function $u_I$ which is locally in $V(K)$ and globally continuous; thus, $u_I$ is a function in the global Virtual Element Space $V_0$ introduced in [12].

### 5.4 Exponential convergence

We set $\Omega^\text{ext} = \bigcup_{n \in \mathbb{N}} \Omega_n^\text{ext} = \Omega_\text{ext}$, where the $\Omega_n^\text{ext}$ are introduced in (D4). We recall that we are assuming that $0 \notin \partial \Omega^\text{ext}$.

We observe that our error analysis needs regularity on $f$ and subsequently on $u$, the right-hand side and the solution of problem (1), respectively. In particular, we will require:

$$f \text{ can be extended to a function in } B^0_{\beta}(\Omega^\text{ext}), \quad u \text{ can be extended to a function in } B^2_{\beta}(\Omega^\text{ext}).$$  

With a little abuse of notation we will call this two functions $f$ and $u$. Assuming $f \in B^0_{\beta}(\Omega^\text{ext})$ automatically implies that $u$ is in $B^2_{\beta}(\Omega^\text{ext})$; this follows from classical elliptic regularity theory, see Theorem 2.1. In the classical $hp$ Finite Element Method, this regularity leads to exponential
convergence of the energy error, see [57]. In order to prove the same exponential convergence with
$h p$ VEM, we need [57] since the approximation by means of polynomials on the polygons not
abutting the singularity needs regularity of the target function on a square containing the polygon, see
Lemmata 5.1 and 5.3.

We recall the inflated domain $\Omega^{\text{ext}}$ has been built in such a way that the singularity is never
at the interior of $\Omega^{\text{ext}}$, see assumption (D3). We highlight also the fact that (56) can be easily
generalized to the case of multiple singularities, see e.g. [57].

In order to obtain exponential convergence of the energy error in terms of the number of degrees
of freedom, we will henceforth assume that the vector $p$ of the degrees of accuracy associated with
$\mathcal{T}_n$ is given by:

$$p_K = \begin{cases} 2 & \text{if } K \in \mathcal{L}_0, \\ \max (2, \lfloor \mu \cdot (j+1) \rfloor) & \text{if } K \in \mathcal{L}_j, \; j \geq 1, \end{cases}$$

(57)

where $\mu$ is a positive constant which will be determined in the proof of Theorem 5.7 and where
$\lfloor \cdot \rfloor$ is the ceiling function. Note that choice (57) could be modified asking for
$p_K = 1$ if $K \in \mathcal{L}_0$; under this requirement in fact Lemmata 5.2, 5.4 and 5.6 are still valid. Nonetheless, we prefer to
use (57) in order to avoid technical discussions on the construction of the right-hand side of the
method and keep the simple representation (22).

It is clear from (57) that if $K_1$ and $K_2$ belong to the $j$-th and the $(j+1)$-th layers respectively,
for some $j = 1, \ldots, n-1$, then $p_{K_1} \approx p_{K_2}$, independently on all the other discretization parameters.
Thus, owing to Section 4, we also have $\alpha(K_1) \approx \alpha(K_2)$, independently on all the other discretization
parameters. Besides, $p_E \approx p_K$ whenever $E \subset 0\hat{K}$.

**Theorem 5.7.** Let $\{\mathcal{T}_n\}_{n}$ be a sequence of polygonal decomposition satisfying (D1)-(D4). Let $u$ and $u_n$ be the solutions of problems (6) and (15) respectively; let $f$ be the right-hand side of
problem (6). Let $N = N(n) = \dim(V_n)$. Assume that $u$ and $f$ satisfy (58). Then, there exists
$\mu > 0$ such that $p$ defined in (57) guarantees the following exponential convergence of the $H^1$
error in terms of the number of degrees of freedom:

$$\|u - u_n\|_{1,\Omega} \lesssim \exp(-b\sqrt{N}),$$

(58)

with $b$ a constant independent of the discretization parameters.

**Proof.** It suffices to combine Lemma 3.1, the results of Section 4, Lemmata from 5.1 to 5.6 and to
use the same arguments of [57] Theorem 4.51, properly choosing the parameter $\mu$.

The basic idea behind the proof is that around the singularity, geometric mesh refinement are
employed, since $p$ approximation leads only to an algebraic decay of the error; on the other hand,
on polygons far from the singularity, it suffices to increase the degree of accuracy, since on such
polygons both the loading term and the exact solution of (6) are assumed to belong to the Babuska
space $B_{\text{ext}}^2(\Omega^{\text{ext}})$ defined in [41] and therefore $p$ approximation leads to exponential convergence of the local errors (see [19] Theorem 5.6).

Following [57] Theorem 4.51 and using Lemma 3.1 yield:

$$\|u - u_n\|_{1,\Omega} \leq c \max_{K \in \mathcal{T}_n} \alpha(K) \sigma^{2(1-\beta)(n+1)},$$

(59)

where $c$ is a constant independent of both the discretizations parameters and the number of layers.
Applying (37), we obtain:

$$\alpha(K) \lesssim p_K^2 n_{\max} K_{\in \mathcal{T}_n} p_{K_{\in \mathcal{T}_n}}^\beta \leq (n+1)^7, \quad \forall K \in \mathcal{T}_n,$$

(60)

where we recall $n+1$ denotes the number of layers. Plugging (60) in (59), we get:

$$\|u - u_n\|_{1,\Omega} \leq c(n+1)^7 \sigma^{2(1-\beta)(n+1)}.$$

We infer:

$$\|u - u_n\|_{1,\Omega} \lesssim \exp(-b(n+1)), \quad \text{for some } b > 0.$$

Now, we prove that $N \lesssim (n+1)^3$. In order to see this, we proceed as follows. In each layer
there exists a fixed maximum number of layers; this follows from the geometric assumptions (D1)
and (D3), applying for instance the arguments in [47, Section 4]. Using geometric assumption (D2) (which guarantees a maximum number of edges per each element), the definition of the local virtual space \( V(K) \) and the distribution of the local degrees of accuracy (57), it is straightforward to note that \( \forall K \in T_n \) the dimension of each local space is given by 
\[
\dim V(K) = p^2_K, \quad p^2_K \approx \ell^2
\]
for \( K \in L_\ell \).

Recalling again (57), we can now compute a bound for the dimension of the local space, viz. the number of the degrees of freedom:
\[
N \lessapprox \sum_{\ell=0}^L \ell^2 \leq L \max_{\ell=0}^L \ell^2 = L^3,
\]
where we stress that we are using that in each layer there is a fixed maximum number of elements. The result follows from Poincaré inequality.

6 Numerical results

We show here numerical experiments validating Theorem 5.7. Let \( u \), the solution of (6), given by the classical benchmark
\[
u(r, \theta) = r^{\frac{4}{3}} \sin \left( \frac{2}{3} (\theta + \frac{\pi}{2}) \right),
\]
on the \( L \)-shaped domain:
\[
\Omega = [-1, 1]^2 \setminus [-1, 0]^2.
\]

6.1 Tests on different meshes

We consider sequences of the meshes depicted in Figure 1 and we consider two different choices for the degree of accuracy distribution \( p \). As a first selection, we pick on all the elements a constant local degree of accuracy which is equal to the number of layers, i.e. \( p = (n+1, n+1, \ldots, n+1) \).

As a second selection, we pick \( p_K \) as in (57), with \( \mu = 1 \), \( \mu \) being the parameter introduced for the construction of the vector of the degrees of accuracy. In Figures 2, 3 and 4, the numerical results are shown.

On the \( y \)-axis, we plot a log scale of the relative energy error between \( u \), defined in (61), and the energy projection \( \Pi^p u_n \), defined in (18a), (18b), of the solution \( u_n \) of the discrete problem (15), i.e.
\[
|u - \Pi^p u_n|_{1, u, \Omega} := \sqrt{\sum_{K \in T_n} |u - \Pi^p u_n|^2_{V(K)}},
\]
on the other hand, in the \( x \)-axis we plot the cubic root of the number of the degrees of freedom of the relative virtual space. The reason for choice (63) is that it is not possible to compute the true energy error since virtual functions are not known explicitly.

We consider the behaviour of the error with three different \( \sigma \), grading parameter, namely \( \sigma = \frac{1}{2}, \sqrt{2} - 1 \), \( (\sqrt{2} - 1)^2 \) and we compare the three types of meshes.

As mentioned previously, the sequence of meshes in Figure 1 (center) does not satisfy assumptions (D1) and (D4). Nevertheless, the expected exponential convergence rate is attained in all cases and for all geometric parameters \( \sigma \).

6.2 A comparison between \( hp \) FEM and \( hp \) VEM

We want now to show a comparison between the performances of \( hp \) (quadrilateral) FEM and \( hp \) VEM. We stress that an analogous of Theorem 5.7 holds for \( hp \) FEM, see e.g. [57]. We consider again the benchmark with known solution (61) and we consider the quadrilateral mesh in Figure 5. In the following we will denote such mesh with d) whereas we denote with a), b) and c) the meshes depicted in Figure 1 (left), (centre) and (right) respectively. In particular, we pick in both cases \( p_K \) as in (57) \( \forall K \in T_n \), with \( \mu = 1 \). We discuss the case of sequences of meshes with grading parameter \( \sigma \) equal to \( \frac{1}{2}, \sqrt{2} - 1 \) and \( (\sqrt{2} - 1)^2 \).
Since we cannot compute the true energy error with the Virtual Element Method (it is not computable since functions in the virtual space are not known explicitly), in order to compare the two methods, we investigate the $L^2$ error on the skeleton $E_n$ (it is computable in all cases a),...d), since also the virtual functions are polynomials on $E_n$), i.e.

$$
\| u - u_n \|_{0,E_n}.
$$

The results are shown in Figure 5. It is possible to see that there is not a preferential choice; for instance, $hp$ VEM performs better than $hp$ FEM when $\sigma = \frac{1}{2}$, they perform almost the same when $\sigma = \sqrt{2} - 1$, performs much worse when $\sigma = (\sqrt{2} - 1)^2$.

In this sense, we can say that the two methods are comparable; nonetheless, the Virtual Element Methods leads to a huge flexibility in the choice of the domain meshing, thus implying the possibility of constructing spaces with a minor number of degrees of freedom.

As a final remark, we observe that we could perform the same analysis in Section 5 by modifying the definition of the local Virtual Spaces (10) into the serendipity local Virtual Spaces introduced in [14]; this would additionally decrease the number of the degrees of freedom of the space, leading as a final output of the method to very small-sized linear systems.
Appendix A

In this first appendix, we show the polynomial $hp$ inverse estimate and some technical background results. We will use the properties of some particular Jacobi polynomials $\{J_n^{\alpha,\beta}(x)\}_{n=0}^{\infty}$, $\alpha, \beta \geq 0$, namely Legendre and shifted-ultraspherical polynomials. Henceforth, we denote with $\tilde{I} = [-1,1]$ the reference interval.

The following result was firstly presented in [28]. We stress that Lemma A.1 holds for more general weights (i.e. $-1 < \alpha \leq \beta$), nonetheless we discuss here only the case $0 \leq \alpha \leq \beta$ which is sufficient for our purpose.

**Lemma A.1.** Let $0 \leq \alpha < \beta$. Then, $\forall q \in \mathbb{P}_p(\tilde{I})$ with $p \in \mathbb{N}$, it holds:

$$\int_{\tilde{I}} (1-x^2)^\alpha q(x)^2 dx \leq c(p+1)^{2(\beta-\alpha)} \int_{\tilde{I}} (1-x^2)^\beta q(x)^2 dx,$$

where $c$ is a positive constant depending on $\alpha$ and $\beta$, but not on $p$.

**Proof.** We split the proof into three parts. The first two are results dealing with shifted ultraspherical polynomials properties, while in the last one we show the assertion.

For the properties of shifted ultraspherical polynomials we refer to [11, 34, 44, 57, 58, 60]. We recall various facts that we will use throughout the proof about these polynomials.

* The $n$-th shifted ultraspherical polynomial $J_n^\alpha$, $\alpha \geq 0$, is the $n$-th Jacobi polynomial $J_n^{\alpha,\beta}$ with $\alpha = \beta \geq 0$; the sequence $\{J_n^{\alpha,\beta}\}_{n=0}^{\infty}$ forms an orthogonal (but not normal) basis for the weighted Lebesgue space:

$$L_{\rho_\alpha}(\tilde{I}) := \left\{ u \text{ Lebesgue-measurable on } \tilde{I} \mid \int_{\tilde{I}} \rho_\alpha(x)|u(x)|^2 dx < +\infty \right\},$$

where $\rho_\alpha$ is the weighted 1D function $\rho_\alpha(x) = (1-x^2)^\alpha$. 

---

Figure 4: Error $|u - \Pi_{n} u_{[1,n]}|$ for the meshes in Figure 3, $\sigma = (\sqrt{\pi} - 1)^2$. Left: the degree of accuracy is uniform and equal to the number of layers. Right: the degree of accuracy is varying over the mesh layers, $\mu = 1$ in (57).

Figure 5: Mesh used for the $hp$ FEM.
We want to show here: We start the proof of the theorem. As a last comment, the details of steps 1 and 2 are carried out

\[ \mu \] of accuracy (parameters \( h, p \)).

\[ \text{Figure 6: hp FEM vs hp FEM. } L^2 \text{ error on the skeleton } |u - u_h|_{0,E_a} \text{ for different sequence of meshes and different parameters } \sigma. \text{ Left: } \sigma = 1/3, \text{ middle: } \sigma = \sqrt{2} - 1, \text{ right: } \sigma = (\sqrt{2} - 1)^2, \text{ linearly varying over the mesh layers degrees of accuracy } (\mu = 1 \text{ in } (61)). \]

* Each \( J_n^\alpha \) is the \( n \)-th eigenfunction of the Sturm-Liouville problem:

\[
(r_{\alpha+1}(x)J_n^\alpha(x)')' + n(n + 2\alpha + 1)r_{\alpha}(x)J_n^\alpha(x) = 0, \quad \forall x \in \mathcal{I},
\]

with appropriate Dirichlet conditions at the endpoints of \( \mathcal{I} \):

\[
J_n^\alpha(\pm 1) = (-1)^n \left( \frac{n + \alpha}{n} \right).
\]

* The following orthogonality relation holds for \( n \geq 1 \), see e.g. [60, formula (4.3.3)]:

\[
\int_{-1}^{1} J_n^\alpha(x)J_m^\alpha(x)r_{\alpha}(x)dx = \delta_{n,m} \frac{2^{2\alpha+1} \Gamma(n + \alpha + 1)^2}{(2n + 2\alpha + 1)n!\Gamma(n + 2\alpha + 1)},
\]

where \( \delta_{n,m} \) is the Kronecker delta and \( \Gamma \) is the Gamma function.

* The following asymptotic behaviour of the Gamma function holds:

\[
\Gamma(t) = \sqrt{2\pi} e^{-t} t^{1/2} (1 + O(1/t)), \quad \text{for } t \to +\infty.
\]

* The following relation between shifted ultraspherical polynomials and their derivatives holds, see [30, Theorem 19.3]:

\[
(2n + 2\alpha + 1)J_n^\alpha(x) = \frac{n + 2\alpha + 1}{n + \alpha + 1} J_{n+1}^\alpha(x)' - \frac{n + \alpha}{n + 2\alpha} J_{n-1}^\alpha(x)'.
\]

We start the proof of the theorem. As a last comment, the details of steps 1 and 2 are carried out here (although they are already known), while step 3 is a detailed version of [28, Theorem 19.3].

**1st STEP** We want to show here:

\[
c_1 n \leq \int_{\mathcal{I}} (J_n^\alpha(x))'^2 r_{\alpha+1}(x)dx \leq c_2 n,
\]

where \( c_1 \) and \( c_2 \) are two positive constants independent on \( n \), but depending on \( \alpha \).

For the purpose, we observe that [65] and an integration by parts imply:

\[
\int_{\mathcal{I}} (J_n^\alpha(x))'^2 r_{\alpha+1}(x)dx = n(n + 2\alpha + 1) \int_{\mathcal{I}} (J_n^\alpha(x))^2 r_{\alpha}(x)dx.
\]

(70)

We stress that one could also show [70] by combining [66] with [60, formula (4.21.7)].

Next, we estimate \( \int_{\mathcal{I}} (J_n^\alpha(x))^2 r_{\alpha}(x)dx \). We set for the purpose:

\[
2^{-2\alpha-1}(2n + 2\alpha + 1) \int_{\mathcal{I}} (J_n^\alpha(x))^2 r_{\alpha}(x)dx =: g(n, \alpha) = \frac{\Gamma(n + \alpha + 1)^2}{\Gamma(n + 1)\Gamma(n + 2\alpha + 1)}.
\]
Function \( g(n, \alpha) \) is increasing in \( n \) for every fixed \( \alpha > -1 \). Besides, \( \lim_{n \to +\infty} g(n, \alpha) = 1 \). Thus:
\[
g(1, \alpha) \leq 2^{-2\alpha-1}(2n + 2\alpha + 1) \int_I J_n^\alpha(x)^2 \rho_\alpha(x) dx \leq 1,
\]
which implies:
\[
\frac{\tilde{c}_1}{n} \leq \int_I J_n^\alpha(x)^2 \rho_\alpha(x) dx \leq \frac{\tilde{c}_2}{n}, \quad n \geq 1, \tag{71}
\]
where \( \tilde{c}_1 \) and \( \tilde{c}_2 \) are two positive constants independent on \( n \), but dependent on \( \alpha \). Using that \( \alpha \geq 1 \), an explicit representation for the two constants \( \tilde{c}_1 \) and \( \tilde{c}_2 \) is given by:
\[
\tilde{c}_1 = \frac{2^{2\alpha+1}}{2\alpha + 3} \frac{\Gamma(2 + \alpha)^2}{\Gamma(2 + 2\alpha)}, \quad \tilde{c}_2 = 2^{2\alpha}. \tag{72}
\]
The claim, i.e. \( \text{(69)} \), follows combining \( \text{(71)} \) and \( \text{(70)} \). An explicit choice for the two constants \( c_1 \) and \( c_2 \) in \( \text{(71)} \) is given by:
\[
c_1 = \tilde{c}_1, \quad c_2 = (2\alpha + 2)2^{2\alpha}. \tag{73}
\]

2nd STEP We show secondly the following bound:
\[
\int_I (J_n^\alpha(x)')^2 \rho_\alpha(x) dx \leq b n^2, \tag{74}
\]
for some positive constant \( b \) independent on \( n \) but depending on \( \alpha \). We will prove this fact by induction. The cases \( n = 1, 2 \) are obvious since we have positive left and right-hand sides. Assume then that \( \text{(74)} \) holds up to \( n \) and we show the inequality for \( n + 1 \).

We observe that the following inequalities involving the coefficients in \( \text{(68)} \) are valid (we recall that \( \alpha, \beta \geq 0 \)):
\[
2n \leq 2n + 2\alpha + 1 \leq (2\alpha + 3)n, \quad 1 \leq \frac{n + 2\alpha + 1}{n + \alpha + 1} \leq 2, \quad \frac{1}{2} \leq \frac{n + \alpha}{n + 2\alpha} \leq 1. \tag{75}
\]
Then, using \( \text{(66)} \), \( \text{(68)} \) and \( \text{(75)} \), we get:
\[
\int_I (J_{n+1}^\alpha(x)')^2 \rho_\alpha(x) dx \leq \int_I \left( \frac{n + 2\alpha + 1}{n + \alpha + 1} J_{n+1}^\alpha(x) \right)^2 \rho_\alpha(x) dx
\]
\[
= \int_I \left( (2n + 2\alpha + 1) J_n^\alpha(x) \right)^2 \rho_\alpha(x) dx + \int_I \left( \frac{n + \alpha}{n + 2\alpha} J_{n-1}^\alpha(x) \right)^2 \rho_\alpha(x) dx \tag{76}
\]
\[
\leq (2\alpha + 3)^2 n^2 \int_I J_n^\alpha(x)^2 \rho_\alpha(x) dx + \int_I \left( J_{n-1}^\alpha(x)' \right)^2 \rho_\alpha dx.
\]
We apply \( \text{(71)} \) and the induction hypotesis to the first and second term in the right-hand side of \( \text{(76)} \) respectively, obtaining:
\[
\int_I (J_{n+1}^\alpha(x)')^2 \rho_\alpha(x) dx \leq c_0 c_2 n + b(n - 1)^2 =: \tilde{c}_2 n + b(n - 1)^2,
\]
where \( c_2 \) is defined in \( \text{(72)} \). Taking \( b \) large enough, for instance \( b \geq \tilde{c}_2 \), the following holds:
\[
\tilde{c}_2 n + b(n - 1)^2 \leq b(n + 1)^2. \tag{77}
\]
We point out that we have to take power 2 in the right-hand side of \( \text{(74)} \) because with smaller powers \( \text{(77)} \) would not be true.

3rd STEP We show \( \text{(64)} \). Let \( q \in \mathbb{P}_p(I) \). We expand it into a derivated Jacobi sum:
\[
g(x) = \sum_{n=1}^{p+1} a_n J_n^\alpha(x)'.
\]
where $\alpha$ is defined in (73). On the other hand, (74) implies:
\[
\int_I q(x)^2 \rho_\alpha(x) \, dx \leq \left( \sum_{n=1}^{p+1} |a_n|^2 \right)^{1/2} \leq b \left( \sum_{n=1}^{p+1} |a_n|^2 \right),
\]
where $b$ is introduced in (74). Combining (78) with (79) and using a Cauchy-Schwarz inequality for sequences, lead to:
\[
\int_I q(x)^2 \rho_\alpha(x) \, dx \leq b \left( \sum_{n=1}^{p+1} a_n^2 \right) \left( \sum_{n=1}^{p+1} n \right) \leq bc_1^{-1}(p+1)^2 \int_I q(x)^2 \rho_{\alpha+1}(x) \, dx.
\]

This is in fact the thesis when $\beta - \alpha = 1$. The case $\beta - \alpha \in \mathbb{N}$ is straightforward; it suffices in fact to iterate enough time the above computations.

Assume now $\alpha \in (\beta - 1, \beta)$. Then:
\[
\alpha = \frac{\beta - 1}{r} + \frac{\beta}{s}, \quad \text{with } \frac{1}{r} + \frac{1}{s} = 1 \quad \text{and } \frac{1}{r} = \beta - \alpha < 1.
\]

In order to conclude, using an Holder inequality and (80):
\[
\int_I q(x)^2 \rho_\alpha(x) \, dx = \int_I q(x)^2 \rho_\beta(x) \, dx \leq \left( \int_I q(x)^2 \rho_{\beta-1}(x) \, dx \right)^{1/\beta} \left( \int_I q(x)^2 \rho_{\beta}(x) \, dx \right)^{1/\beta} = (bc_1^{-1})^\beta (p+1)^{2(\beta-\alpha)} \int_I q(x)^2 \rho_\beta(x) \, dx.
\]

We point out that in order to prove the case $\alpha \in (\beta - 1, \beta)$ one could also use interpolation theory, see [51, 62]. Nonetheless, we believe that a direct computation is easily readable.

The following lemma is a quasi-one dimensional result on trapezoids. The idea is pretty similar to that in [50, Lemma D.3], although our result employs a different class of weight functions. We also point out that Lemma A.2 can be generalized to the case $-1 < \alpha < \beta$, see [50].

**Lemma A.2.** Let $d \in (0,1)$. Let $a, b \in \mathbb{R}$ such that $-1 + ad < 1 + bd$. We set the $(a, b, d)$-trapezoid as:
\[
D(a, b, d) = D = \{(x, y) \in \mathbb{R}^2 \mid y \in [0, d], \ -1 + ay \leq x \leq 1 + by\}.
\]
We associate to each $y^*$ the segment:
\[
I(y^*) = I^* = [-1 + ay^*, 1 + by^*].
\]

For every $\Phi \in C^0(\overline{D})$ such that:
\[
\Phi(\cdot, y^*) \in \mathbb{P}_3(I^*) \text{ is concave; } \Phi(x, y^*) \geq 0, \forall x \in I^*; \ \Phi = 0 \text{ only at the endpoints of } I^*, \ \forall y^* \in [0, d],
\]
the following quasi-one dimensional $p$ inverse estimate holds:
\[
\int_D \Phi^\alpha(x, y)q(x, y)^2 \, dx \, dy \leq c(p+1)^{2(\beta-\alpha)} \int_D \Phi^\beta(x, y)q(x, y)^2 \, dx \, dy, \ \forall q \in \mathbb{P}_p(D),
\]
where $c$ is a positive constant depending only on $\alpha$ and $\beta$, but not on $p$, and where $\beta > \alpha \geq 0$. 

---

23
Proof. 1st STEP Let \( \psi(x) = (1 - x^2) \) be the 1D bubble function associated to the reference interval \( \hat{T} := [-1, 1] \). Given \( y^* \in [0, d] \), we set \( F \) the affine function mapping \( \hat{T} \) in \( I^* \). Let \( \psi^*(x) = \psi(F^{-1}(x)) : I^* \to \mathbb{R} \).

Then, there exist two positive constants \( c_1 \) and \( c_2 \) depending only on \( a, b \) and \( y^* \) such that
\[
c_1 \psi^*(x) \leq \Phi(x, y^*) \leq c_2 \psi^*(x), \quad \forall x \in I^*.
\]

This follows from the fact that both \( \psi^*(\cdot) \) and \( \Phi(\cdot, y^*) \) are two positive quadratic/cubic concave polynomials annihilating only at the endpoints of the segment for all \( y^* \in [0, d] \), see [52].

Since \( \Phi \) is by hypothesis a continuous function in \( y^* \), then \( c_1 \) and \( c_2 \) depend continuously on \( y^* \). Having that \( y^* \) lives in the compact set \([0, d]\), then \( c_1 \) and \( c_2 \) attain maximum and minimum respectively. Further, such extremal points are strictly positive due to the positiveness of \( c_1 \) and \( c_2 \) seen as functions of \( y^* \), see [52]. Therefore, we can write:
\[
\overline{c}_1 \psi^*(x) \leq \Phi(x, y^*) \leq \overline{c}_2 \psi^*(x), \quad \forall x \in I^*,
\]

where \( 0 < \overline{c}_1 = \min_{y^* \in [0,d]}(c_1(y^*)) \) and \( \overline{c}_2 = \max_{y^* \in [0,d]}(c_2(y^*)) \) are now independent on \( y^* \).

2nd STEP We investigate a 1D inverse inequality. In particular, from Lemma A.1 and from Step 1, we have:
\[
\int_{I^*} \Phi(x, y)^\alpha q(x, y)^2 dx \leq \overline{c}_2 \int_{I^*} \psi^*(x) q(x, y)^2 dx \leq \frac{\overline{c}_2}{\overline{c}_1} c(p + 1)^{2(\beta - \alpha)} \int_{I^*} \Phi(x, y)^\beta q(x, y)^2 dx,
\]

where \( c \) is independent on \( y^* \).

3rd STEP The statement of the lemma is achieved by means of an integration of (84) over \( y^* \in [0, d] \).}

We show now a global inverse estimate on triangles. Again, the following result also holds for weight \(-1 < \alpha \leq \beta\), see [50].

Theorem A.3. Let \( 0 \leq \alpha \leq \beta \). Let \( \hat{T} \) be the reference triangle of vertices \((0,0), (1,0)\) and \((0,1)\). Let \( b_{\hat{T}} \) be the cubic bubble function associated with \( \hat{T} \); in particular, \( b_{\hat{T}} \in \mathbb{P}_3(\hat{T}) \) is such that \( b_{\hat{T}}|_{\partial\hat{T}} = 0 \). Then:
\[
\int_{\hat{T}} b_{\hat{T}}^\alpha q^2 \leq c(p + 1)^{2(\beta - \alpha)} \int_{\hat{T}} b_{\hat{T}}^\beta q^2, \quad \forall q \in \mathbb{P}_p(\hat{T}),
\]

where \( c \) is a positive constant independent on \( p \).

Proof. The proof is similar to that in [50] Theorem D2] and for this reason we only sketch it. The idea consists in partitioning \( \hat{T} \) into six (overlapping) trapezoids and applying some \( p \) inverse estimates analogous to those presented in Lemma A.2. In particular, the six overlapping trapezoids \( D_1, \ldots, D_6 \) are built, for instance, as in Figure 7.

![Figure 7: Overlapping trapezoids covering the reference triangle \( \hat{T} \)](image)

We observe that we can apply Lemma A.2 since its hypothesis are satisfied. In fact, on all the \( D_i, i = 1, \ldots, 6 \), \( b_{\hat{T}}|_{D_i} \) is a continuous function whose restriction on every segment parallel to the basis of the trapezoid is a cubic concave polynomial annihilating only at the endpoints of the basis.

Estimate (85) follows then from an overlapping argument applying Lemma A.2 on all the elements, noting that \( \hat{T} = \bigcup_{i=1}^6 D_i \).
We discuss also the following result, firstly presented in [11] Lemma B.3.

**Lemma A.4.** Let $T$ be a triangle, $b_T$ the associated cubic bubble function. Then:

\[ |qb_T|_{1,T} \leq c \frac{p+1}{h_T} \|qb_T\|_{0,T}, \quad \forall q \in P_p(T), \]

where $c$ is a positive constant independent on $h_T$ and $p$, $h_T = \text{diam}(T)$.

**Proof.** The proof is split into three parts; the first two of them are technical results dealing with Legendre-type approximations, while the third one deals with the proof of the lemma.

**1st STEP** The following estimate holds: for all $q \in P_p(T)$, $T := [-1,1]$

\[ \| (1 - x^2)q'(x) \|_{0,T} \leq c (p+1) \| (1 - x^2)^{\frac{3}{2}} q(x) \|_{0,T}, \quad (87) \]

where $c$ is a constant independent on $p$.

We recall that the following holds, see [57] formula (3.39)\[ \int_{T} (1 - x^2)^k L_n^{(k)}(x)^2 dx = \begin{cases} \frac{2}{2n+1} \frac{(n+1)!}{(n-k)!} & \text{if } n \geq k, \\ 0 & \text{otherwise}, \end{cases} \]

where $L_n^{(k)}(x)$ is the $k$-th derivative of the $n$-th Legendre polynomial. Then:

\[ \int_{T} (1 - x^2)^2 L_n''(x)^2 dx = \frac{2}{2n+1} \frac{(n+1)!}{(n-k)!} (n-1)(n+1) \leq \left( n + \frac{1}{2} \right)^2 \int_{T} (1 - x^2) L_n'(x)^2 dx. \quad (88) \]

Therefore, expanding $q$ into a derivated Legendre sum

\[ q(x) = \sum_{n=1}^{p+1} c_n L_n(x), \]

we have, owing to orthogonality of the second derivative of Legendre polynomials with respect to $L^2 (1 - x^2)^2$-weighted inner product and owing to (88),
\[
\int_{T} q'(x)(1 - x^2)^2 dx = \sum_{n=1}^{p+1} c_n^2 \int_{T} L_n''(x)^2 (1 - x^2)^2 dx \leq \sum_{n=1}^{p+1} c_n^2 \left( n + \frac{1}{2} \right)^2 \int_{T} L_n'(x)^2 (1 - x^2) dx \\
\leq \left( p + \frac{3}{2} \right)^2 \int_{T} q^2(x)(1 - x^2) dx \leq \frac{3}{2} (p+1)^2 \int_{T} q^2(x)(1 - x^2) dx.
\]

**2nd STEP** We show now the following 1D estimate. For $a < b$, let

\[ b_{[a,b]}(x) := \frac{(x-a)(b-x)}{(b-a)^2} \]

be the 1D quadratic bubble function, then:
\[
\| (b_{[a,b]}q)' \|_{0,[a,b]} \leq c \frac{p+1}{b-a} \|b_{[a,b]}q\|_{0,[a,b]}^{\frac{1}{2}}, \quad \forall q \in P_p([a,b]), \quad (89)
\]

where $c$ is a positive constant independent on $p$. It is sufficient to show (89) on the reference interval $[-1,1]$, since the general result follows from a scaling argument.

Owing to $\| b_{[-1,1]} \|_{\infty,[-1,1]} = \frac{1}{2} < 1$, the Leibniz derivation rule and a triangular inequality, we can write:
\[
\| (b_{[-1,1]}q)' \|_{0,[-1,1]} \leq \| b_{[-1,1]}q \|_{0,[-1,1]} + \| b_{[-1,1]}q'' \|_{0,[-1,1]} \leq \| q \|_{0,[-1,1]} + \| b_{[-1,1]}q'' \|_{0,[-1,1]} . \quad (90)
\]

Applying (64) (with $\alpha = 0$ and $\beta = 1$) and (67) to the first and second term of (90) respectively, we get (89).

**3rd STEP** We apply now (89) and we show the claim of the lemma.
Without loss of generality, we work on the reference triangle $\hat{T} = T$ of vertices $(0,0)$, $(1,0)$ and $(0,1)$. The statement follows from a scaling argument. The cubic bubble function on $\hat{T}$, which is given by the product of the barycentric coordinates, can be rewritten as:

$$b_{\hat{T}} = b_{[0,1-x]}(y)(1-x)b_{[0,1]}(x).$$

We only show the bound on the partial derivative with respect to $y$. The general case is an easy consequence.

$$\|\partial_y (b_{\hat{T}}q)\|^2_{0,\hat{T}} = \int_0^1 \int_0^{1-x} (\partial_y (b_{\hat{T}}(x,y)q(x,y)))^2 dy dx = \int_0^1 b_{[0,1]}(x)(1-x)^2 \int_0^{1-x} (\partial_y (b_{[0,1-x]}(y)q(x,y)))^2 dy dx. $$

We note that:

$$(1-x)^2 \int_0^{1-x} (\partial_y b_{[0,1-x]}(y)q(x,y))^2 dy = (1-x)^2 \|\partial_y (b_{[0,1-x]}(\cdot)q(x,\cdot))\|^2_{0,[0,1-x]} \leq c(p+1)^2 b_{[0,1-x]}^2(\cdot)q(x,\cdot)\|^2_{0,[0,1-x]}.$$ 

Since $b_{[0,1]} \leq 1-x$, we get $b_{[0,1]}^2(\cdot)b_{[0,1-x]}(y) \leq b_{\hat{T}}(x,y)$ and consequently:

$$\|\partial_2(b_{\hat{T}}q)\|^2_{0,\hat{T}} \leq c(p+1)^2 \int_0^1 \int_0^{1-x} b_{[0,1]}(x)b_{[0,1-x]}(y)q^2(x,y) dy dx \leq c(p+1)^2 b_{\hat{T}}^2 q^2.$$

We are now ready for the inverse estimate involving the $H^{-1}$ norm of polynomials.

**Theorem A.5.** Let $K \subseteq \mathbb{R}^2$ be a polygon. Assume that there exists $\mathcal{T}_n(K)$ subtriangulation of $K$ such that $h_K \approx h_T$, where $h_\omega = \text{diam}(\omega)$, $\omega \subseteq \mathbb{R}^2$. Let $q \in P_p(K)$, $p \in \mathbb{N}$. Then:

$$\|q\|_{0,K} \leq c\frac{(p+1)^2}{h_K} \|q\|_{-1,K},$$

where $\|q\|_{-1,K} := \|q\|_{(H^1_0(K))'}$.

**Proof.** Let $b_K$ be the “patch-bubble” function, defined on each $T \in \mathcal{T}_n(K)$ as the local cubic bubble function $b_T$ introduced in Lemma A.4 Then:

$$\|q\|_{-1,K} = \sup_{\Phi \in H^1_0(K), \Phi \neq 0} \frac{(q,\Phi)_{0,K}}{\|\Phi\|_{1,K}} \geq \frac{(q,b_K)_{0,K}}{\|b_K\|_{1,K}} = \frac{\|\sqrt{b_K}q\|_{0,K}}{\left(\sum_{T \in \mathcal{T}_n(K)} \|b_T\|^2_{0,T}\right)^{\frac{1}{2}}}.$$ 

Using now Lemma A.4 (91) and the geometric assumption (D2), we obtain:

$$\|q\|_{-1,K} \geq c \frac{\min_{T \in \mathcal{T}_n(K)} h_T}{p+1} \|q\sqrt{b_K}\|_{0,K} \geq c \frac{h_K}{p+1} \left(\sum_{T \in \mathcal{T}_n(K)} \|q\sqrt{b_T}\|^2_{0,T}\right)^{\frac{1}{2}}.$$ 

Finally, we apply Theorem A.3 with $\alpha = 0$ and $\beta = 1$ and get:

$$\|q\|_{-1,K} \geq c \frac{h_K}{(p+1)^2} \left(\sum_{T \in \mathcal{T}_n(K)} \|q\|^2_{0,T}\right)^{\frac{1}{2}} = c \frac{h_K}{(p+1)^2} \|q\|_{0,K}.$$
Appendix B

In this second appendix, we discuss the following standard $hp$ polynomial inverse estimate on triangles:

**Theorem B.1.** Let $T \subseteq \mathbb{R}^2$ be a triangle and let $h_T$ denote the diameter of $T$. Then:

$$|q|_{1,T} \leq c_{\text{inv}} \frac{p^2}{h_T} \|q\|_{0,T}, \quad \forall q \in P_p(T), \quad p \geq 1,$$

(93)

where $c_{\text{inv}}$ is a positive constant independent on $h_T$, $p$ and $q$.

We note that inequality (93) is a very well-known and widely used result. It is stated for instance in [57, Theorem 4.76]. Nonetheless, we were not able to find an explicit proof in literature.

**Proof of Theorem B.1.** We show the result on the reference triangle $\hat{T}$ of vertices $(0,0)$, $(1,0)$, $(0,1)$. The statement will follow from a scaling argument.

We consider a decomposition of $\hat{T}$ into the three overlapping parallelograms $P_1$, $P_2$ and $P_3$ depicted in Figure 8.

![Figure 8: Overlapping parallelograms covering the reference triangle $\hat{T}$](image)

We can write:

$$|q|_{1,\hat{T}} \leq |q|_{1,P_1} + |q|_{1,P_2} + |q|_{1,P_3}.$$  

(94)

We only have to prove that:

$$|q|_{1,P_i} \leq c_1 p^2 \|q\|_{0,P_i}, \quad i = 1, 2, 3.$$  

(95)

In particular, it suffices to prove the same inequality on the reference square $\hat{Q} = [-1,1]^2$ and then using an affine transformation in order to deduce the assertion of the theorem from (94). Thus, we must prove:

$$|q|_{1,\hat{Q}} \leq c_2 p^2 \|q\|_{0,\hat{Q}}, \quad \forall q \in P_p(\hat{Q}).$$  

(96)

For the purpose, we have to show:

$$\|\partial_i q\|_{0,\hat{Q}} \leq c_2 p^2 \|q\|_{0,\hat{Q}}, \quad i = 1, 2.$$  

Owing to [57, Theorem 3.96], we can write:

$$\|\partial_i q\|_{0,\hat{I}} \leq c_2 p^2 \|q\|_{0,\hat{I}},$$  

(97)

where

$$\hat{I} = \begin{cases} [-1,1] \times \{\tilde{y}\}, & \tilde{y} \in [-1,1] \quad \text{if } i = 1, \\
\{\tilde{x}\} \times [-1,1], & \tilde{x} \in [-1,1] \quad \text{if } i = 2.
\end{cases}$$

Here, the constant $c$ does not depend on $\tilde{y}$. Integrating (97) in $y$ (if $i = 1$) or in $x$ (if $i = 2$) from $-1$ to $1$, we get (96).

\[\square\]
Acknowledgement

The first author has received funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (grant agreement no. 681162).

References

[1] M. Abramowitz and I. A. Stegun. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. 1964.

[2] R. A. Adams and J. J. F. Fournier. *Sobolev Spaces*, volume 140. Academic Press, 2003.

[3] B. Ahmad, A. Alsaedi, F. Brezzi, L.D. Marini, and A. Russo. Equivalent Projectors for Virtual Element Method. *Computers & Mathematics with Applications*, 66(3):376–391, 2013.

[4] O. Andersen, H.M. Nilsen, and X. Raynaud. On the use of the Virtual Element Method for geomechanics on reservoir grids. [http://arxiv.org/abs/1606.09508](http://arxiv.org/abs/1606.09508).

[5] P. F. Antonietti, L. Beirão da Veiga, D. Mora, and M. Verani. A Stream Virtual Element formulation of the Stokes problem on polygonal meshes. *SIAM Journal of Numerical Analysis*, 52(1):386–404, 2014.

[6] P. F. Antonietti, L. Beirão da Veiga, S. Scacchi, and M. Verani. A $C^1$ Virtual Element Method for the Cahn–Hilliard equation with polygonal meshes. *SIAM Journal on Numerical Analysis*, 54(1):34–56, 2016.

[7] I. Babuška and B. Q. Guo. Regularity of the solution of elliptic problems with piecewise analytic data. Part I: Boundary value problems for linear elliptic equation of second order. *SIAM Journal on Mathematical Analysis*, 19(1):172–203, 1988.

[8] I. Babuška and B. Q. Guo. Regularity of the solution of elliptic problems with piecewise analytic data. Part II: The trace spaces and application to the boundary value problems with nonhomogeneous boundary conditions. *SIAM Journal on Mathematical Analysis*, 20(4):763–781, 1989.

[9] I. Babuška and M. Suri. The $hp$ version of the Finite Element Method with Quasiuniform Meshes. *Modélisation mathématique et analyse numérique*, 21(2):199–238, 1987.

[10] I. Babuška and M. Suri. The optimal convergence rate of the $p$-version of the Finite Element Method. *SIAM Journal on Numerical Analysis*, 24(4):750–776, 1987.

[11] R. E. Bank, A. Parsania, and S. Sauter. Saturation estimates for $hp$-finite element methods. *Computing and Visualization in Science*, 16(5):195–217, 2013.

[12] L. Beirão da Veiga, C. Lovadina, and D. Mora. A Virtual Element Method for elastic and inelastic problems on polytope meshes. *Computer Methods in Applied Mechanics and Engineering*, 295:327 – 346, 2015.

[13] L. Beirão da Veiga, F. Brezzi, A. Cangiani, G. Manzini, L.D. Marini, and A. Russo. Basic principles of Virtual Element Methods. *Mathematical Models and Methods in Applied Sciences*, 23(01):199–214, 2013.

[14] L. Beirão da Veiga, F. Brezzi, L. D. Marini, and A. Russo. Serendipity nodal VEM spaces. [http://arxiv.org/abs/1510.08477](http://arxiv.org/abs/1510.08477), 2015.

[15] L. Beirão da Veiga, F. Brezzi, L. D. Marini, and A. Russo. Mixed Virtual Element Methods for second order elliptic problems on polygonal meshes. *Mathematical Modelling and Numerical Analysis*, 50(3):727–747, 2016.

[16] L. Beirão da Veiga, F. Brezzi, L. D. Marini, and A. Russo. Virtual Element Method for general second-order elliptic problems on polygonal meshes. *Mathematical Models and Methods in Applied Sciences*, 26(4):729–750, 2016.

[17] L. Beirão da Veiga, F. Brezzi, and L.D. Marini. Virtual Elements for linear elasticity problems. *SIAM Journal of Numerical Analysis*, 51:794–812, 2013.

[18] L. Beirão da Veiga, F. Brezzi, L.D. Marini, and A. Russo. The Hitchhiker’s Guide to the Virtual Element Method. *Mathematical Models and Methods in Applied Sciences*, 24(8):1541–1573, 2014.

[19] L. Beirão da Veiga, A. Chernov, L. Mascotto, and A. Russo. Basic principles of $hp$ Virtual Elements on quasuniform meshes. *Mathematical Models and Methods in Applied Sciences*, 26(4):1567–1598, 2016.

[20] L. Beirão da Veiga, K. Lipnikov, and G. Manzini. The Mimetic Finite Difference Method for elliptic problems, volume 11. Springer, 2014.

[21] L. Beirão da Veiga, C. Lovadina, and A. Russo. Stability analysis for the Virtual Element Method. [http://arxiv.org/abs/1607.05988](http://arxiv.org/abs/1607.05988), 2016.

[22] L. Beirão da Veiga, C. Lovadina, and G. Vacca. Divergence free Virtual Elements for the Stokes problem on polygonal meshes. [http://arxiv.org/abs/1510.01655](http://arxiv.org/abs/1510.01655), 2015.

[23] L. Beirão da Veiga and G. Manzini. A Virtual Element Method with arbitrary regularity. *IMA Journal of Numerical Analysis*, 34(2):759–781, 2014.

[24] L. Beirão da Veiga and G. Manzini. Residual a posteriori error estimation for the Virtual Element Method for elliptic problems. *Mathematical Modelling and Numerical Analysis*, 49(2):577–599, 2015.

[25] M. F. Benedetto, S. Berrone, A. Borio, S. Pieraccini, and S. Scialò. A hybrid mortar Virtual Element Method for discrete fracture network simulations. *Journal of Computational Physics*, 306:148 – 166, 2016.
[26] M.F. Benedetto, S. Berrone, S. Pieraccini, and S. Scialò. The Virtual Element Method for Discrete Fracture Network simulations. *Comput. Meth. Appl. Mech. Engrg.*, 280:135–156, 2014.

[27] M.F. Benedetto, S. Berrone, and S. Scialò. A globally conforming method for solving flow in discrete fracture networks using the Virtual Element Method. *Finite Elem. in Anal. and Design*, 109:23 – 36, 2016.

[28] C. Bernardi, N. Fièff, and R. G. Owens. An error indicator for mortar element solutions to the Stokes problem. *IMA Journal of Numerical Analysis*, 21(4):857–886, 2001.

[29] C. Bernardi and Y. Maday. Polynomial interpolation results in Sobolev spaces. *Journal of Computational and Applied Mathematics*, 43(1):53–80, 1992.

[30] C. Bernardi and Y. Maday. Spectral Methods, in the Handbook of Numerical Analysis V, P.G Ciarlet & J.-L. Lions eds, 1997.

[31] S. C. Brenner and L. R. Scott. *The mathematical theory of Finite Element Methods*, volume 15. Texts in Applied Mathematics, Springer-Verlag, New York, third edition, 2008.

[32] F. Brezzi, K. Lipnikov, and M. Shashkov. Convergence of the mimetic finite difference method for diffusion problems on polyhedral meshes. *SIAM Journal on Numerical Analysis*, 43(5):1872–1896, 2005.

[33] F. Brezzi and L.D. Marini. Virtual Element Method for plate bending problems. *Computer Methods in Applied Mechanics and Engineering*, 253:455–462, 2013.

[34] E. Caceres and G.N. Gatica. A mixed Virtual Element Method for the pseudostress-velocity formulation of the Stokes problem. *IMA Journal of Numerical Analysis*, 2016. DOI: 10.1093/imanum/drw002.

[35] A. Cangiani, E. H Georgoulis, and P. Houston. hp-version Discontinuous Galerkin methods on polygonal and polyhedral meshes. *Mathematical Models and Methods in Applied Sciences*, 24(10):2009–2041, 2014.

[36] A. Cangiani, E. H. Georgoulis, T. Pryer, and O. J. Sutton. A posteriori error estimates for the Virtual Element Method. [http://arxiv.org/abs/1605.08555](http://arxiv.org/abs/1605.08555), 2016.

[37] A. Cangiani, V. Gyrya, and G. Manzini. The non-conforming Virtual Element Method for the Stokes equations. [http://arxiv.org/abs/1601.03770](http://arxiv.org/abs/1601.03770).

[38] A. Cangiani, G. Manzini, and O. J. Sutton. Conforming and nonconforming Virtual Element Methods for elliptic problems. [http://arxiv.org/abs/1507.03543](http://arxiv.org/abs/1507.03543), 2015.

[39] B. Cockburn, J. Gopalakrishnan, and R. Lazarov. Unified hybridization of Discontinuous Galerkin, mixed, and Continuous Galerkin methods for second order elliptic problems. *SIAM Journal on Numerical Analysis*, 47(2):1319–1365, 2009.

[40] D. A. Di Pietro and A. Ern. Hybrid High-order Methods for variable-diffusion problems on general meshes. *Comptes Rendus Mathématique*, 353(1):31–34, 2015.

[41] M. Dubiner. Spectral Methods on triangles and other domains. *Journal of Scientific Computing*, 6(4):345–390, 1991.

[42] A.L. Gain, G.H. Paulino, S.D. Leonardo, and I.F.M. Menezes. Topology optimization using polytopes. *Computer Methods in Applied Mechanics and Engineering*, 293:411–430, 2015.

[43] A.L. Gain, C. Talischi, and G.H. Paulino. On the Virtual Element Method for Three-Dimensional Elasticity Problems on Arbitrary Polyhedral Meshes. In press on CMAE. DOI: 10.1016/j.cma.2014.05.005.

[44] A. Ghizzetti and A. Ossicini. *Quadrature Formulae*. Birkh" auser, 1970.

[45] A. Gillette, A. Rand, and C. Bajaj. Error estimates for generalized barycentric interpolation. *Advances in computational mathematics*, 37(3):417–439, 2012.

[46] B. Guo and L. Wang. Jacobi approximations in non-uniformly Jacobi-weighted Sobolev spaces. *Journal of Approximation Theory*, 128(1):31–41, 2004.

[47] R. Hiptmair. A two-variable analogue of the classical orthogonal polynomials. *Theory and applications of special functions*, pages 435–495, 1975.

[48] H. Li and J. Shen. Optimal error estimates in Jacobi-weighted Sobolev spaces. *Mathematics of Computation*, 79(271):1621–1646, 2010.

[49] J. Melenk. hp-interpolation of non-smooth functions. *SIAM Journal of Numerical Analysis*, 43:127–155, 2005.

[50] D. Mora, G. Rivera, and R. Rodriguez. A posteriori error estimates for a virtual elements method for the steklov eigenvalue problem. [http://arxiv.org/abs/1609.07154](http://arxiv.org/abs/1609.07154).

[51] D. Mora, G. Rivera, and R. Rodríguez. A Virtual Element Method for the Steklov eigenvalue problem. *Mathematical Models and Methods in Applied Sciences*, 25(08):1421–1445, 2015.
[54] S. Natarajan, S.PA Bordas, and E.T. Ooi. Virtual and smoothed finite elements: a connection and its application to polygonal/polyhedral finite element methods. *International Journal for Numerical Methods in Engineering*, 104(13):1173–1199, 2015.

[55] Perugia, I., Pietra, P., and Russo, A. A plane wave Virtual Element Method for the Helmholtz problem. *Mathematical Modelling and Numerical Analysis*, 50(3):783–808, 2016.

[56] S. Rjasanow and S. Weißer. Higher order bem-based fem on polygonal meshes. *SIAM Journal on Numerical Analysis*, 50(5):2357–2378, 2012.

[57] C. Schwab. *p- and hp-Finite Element Methods: Theory and Applications in Solid and Fluid Mechanics*. Clarendon Press Oxford, 1998.

[58] J. Shen, T. Tang, and L.-L. Wang. *Spectral Methods: algorithms, analysis and applications*, volume 41. Springer Science & Business Media, 2011.

[59] N. Sukumar and A. Tabarraei. Conforming Polygonal Finite Elements. *International Journal for Numerical Methods in Engineering*, 61:2045–2066, 2004.

[60] G. Szegő. *Orthogonal polynomials*, volume 23. American Mathematical Soc., 1939.

[61] L. Tartar. *An introduction to Sobolev spaces and interpolation spaces*, volume 3. Springer Science & Business Media, 2007.

[62] H. Triebel. *Interpolation theory, function spaces, differential operators*. North-Holland, 1978.

[63] P. Wriggers, W. T. Rust, and B. D. Reddy. A virtual element method for contact. *Computational Mechanics*, pages 1–12, 2016. DOI: 10.1007/s00466-016-1331-x.