Comput Visual Sci (2013) 16:195–217
DOI 10.1007/s00791-015-0234-2

Saturation estimates for *hp*-finite element methods

Randolph E. Bank · Asieh Parsania · Stefan Sauter

Received: 24 March 2014 / Accepted: 28 May 2014 / Published online: 27 February 2015

Abstract In this paper we will prove *saturation estimates* for the adaptive *hp*-finite element method for linear, second order partial differential equations. More specifically we will consider a sequence of nested finite element discretizations where we allow for both, local mesh refinement and locally increasing the polynomial order. We will prove that the energy norm of the error on the finer level can be estimated by the sum of a *contraction* of the old error and data oscillations. We will derive estimates of the contraction factor which are explicit with respect to the local mesh width and the local polynomial degree. In order to cover *p*-refinement of finite element spaces new polynomial projection operators will be introduced and new polynomial inverse estimates will be derived.

Keywords *hp* Finite elements · Saturation estimate · A posteriori error estimate · Linear elliptic equation · Polynomial inverse estimates

Mathematics Subject Classification 65N15 · 65N30 · 65N35 · 65N50

Communicated by Gabriel Wittum.

Dedicated to Dietrich Braess on the occasion of his 75th birthday and Wolfgang Hackbusch on the occasion of his 65th birthday.

Part of this work has been carried out during a visit of the third author at the Department of Mathematics, University of California, San Diego, La Jolla. This support is gratefully acknowledged.

R. E. Bank
Department of Mathematics, University of California, San Diego, La Jolla, CA 92093-0112, USA
e-mail: rbank@ucsd.edu

A. Parsania · S. Sauter (✉)
Institut für Mathematik, Universität Zürich, Winterthurerstrasse 190, 8057 Zurich, Switzerland
e-mail: stas@math.uzh.ch

A. Parsania
e-mail: asieh.parsania@math.uzh.ch

1 Introduction

In this paper, we will consider the discretization of linear, second order elliptic partial differential equations by finite elements. Nowadays, adaptive techniques based on a posteriori error estimation have been established to set up a sequence of finite element approximations which should converge towards the exact solution. The advantage compared to uniform mesh refinement is that the finite element spaces are enriched from level to level in a problem oriented way.

A posteriori error estimation and adaptivity are well established methodologies for the numerical solution of partial differential equations by finite elements (cf. [2,4–6,11,18,25,32,36,38]).

Some types of error estimators as, e.g., hierarchical error estimators (see, e.g., [9,10,12]) require explicitly or implicitly the *saturation* assumption which states that the error on the refined mesh and/or with higher polynomial degree is strictly smaller than the error on the previous mesh/polynomial degree. In the pioneering paper [19] the saturation assumption is proved for the $\mathbb{P}_1$-finite element method for the Poisson problem in two spatial dimensions under the assumption that the data oscillations are small. In [25] the convergence of adaptive finite element methods (AFEM) for general (nonsymmetric) second order linear elliptic partial differential equations is proved, where the

term "adaptivity" is understood in the sense of adaptive *mesh* refinement and the polynomial degree stays fixed. The theory in [25] also generalizes the proof of the saturation property to quite general 2nd order elliptic problems and estimate the error on the refined mesh by the error of the coarser mesh plus a data oscillation term.

In this paper, we will focus on adaptive *hp*-refinement, i.e., the finite element space is enriched by increasing locally the polynomial degree of the ansatz functions while we allow also for conventional local *h*-refinement, where the elements of the finite element mesh are geometrically subdivided. We will show (and quantify) that, for *residual* a posteriori error estimation, the saturation property, i.e., the error contraction from level to level behaves like $\left(1-\frac{C}{p^{5/2}}\right)$ provided the data oscillations are sufficiently well resolved. Hence, *p*-refinement should be combined with *h*-refinement in order to guarantee that the numerical solution converges towards the exact solution. Common strategies for *hp*-refinement are based on the estimation of the local regularity of the solution on a triangle by using error estimators for different local polynomial orders in order to decide for *h*- or *p*-refinement; for details we refer to [3,14,17,21,29,33,34,37]. The *hp*-refinement indicator which is implemented in the software package PLTMG (cf. [7]) is based on the superconvergence result that recovered derivatives for elements of degree *p* have higher order accuracy, provided the true solution has the required smoothness (cf. [8]).

Our a posteriori error estimation takes into account data oscillations but does not incorporate errors due to numerical quadrature [1], to iterative approximations of the solution of the linear system [22], and to approximations of the domain [15,20].

The paper is organised as follows. In Sect. 2 we will introduce the elliptic boundary value problem and formulate appropriate assumptions to ensure the well-posedness of this problem.

The *hp*-finite element method will be defined in Sect. 3 and standard assumptions on mesh refinment, shape regularity, and the polynomial degree distribution will be introduced.

In Sect. 4 we will recall the definition of the residual a posteriori error estimator for *hp*-finite elements and its reliability estimate.

In Sect. 5 we will introduce some polynomial projection operator which maps *global* polynomials on triangle patches to *piecewise* polynomials of lower degree. This allows to localize projected residuals by multiplying the resulting piecewise polynomials with appropriate bubble functions. We will investigate the stability constant of the projection operator while its explicit dependence on the polynomial degree for *p*-refinement will be analysed numerically in "Appendix 1".

The saturation estimate will be proved in Sect. 6.

In "Appendix 2" we will derive polynomial inverse estimates containing those bubble functions as weights which have been used in Sect. 6 to prove the saturation property.

Remark 1 The theory in [25] indicates how an adaptive finite element procedure should be defined such that the sequence of finite element solutions converges. Note that the *rate* of convergence for adaptive finite elements is investigated in, e.g., [11,35,36].

Besides the estimates derived for the saturation property, the *convergence* theory requires a reduction of the data oscillations which, for *h*-refinement, is (essentially) related to the fact that the local mesh width shrinks by a fixed factor for the marked elements. For *p*-refinement, the analogue condition is that the *hp*-weight of the data oscillations term also shrinks by a factor smaller than one. Due to the non-robust *p*-dependence of polynomial inverse estimates this cannot be expected in a straightforward way. In order not to overload this paper we decided to leave the convergence of an adaptive *hp*-finite element method as well as the detailed description of the *hp*-adaptive refinement strategy to a forthcoming paper. We also emphasize that a posteriori error estimators which are based on the hypercircle method (cf. [30]) such as the *equilibrated residual error estimates* are *p*-robust (see [13]).

2 Setting

Let $\Omega \subset \mathbb{R}^d$ be a bounded Lipschitz domain. Consider the Dirichlet problem for given $f \in L^2(\Omega)$:

$$\begin{aligned} -\operatorname{div}(\mathbf{A}\nabla u) + \langle \mathbf{b}, \nabla u\rangle + cu &= f \quad \text{in } \Omega, \\ u &= 0 \quad \text{on } \partial\Omega \end{aligned} \tag{2.1}$$

with variational formulation: Find $u \in H_0^1(\Omega)$ such that

$$\begin{aligned} a(u,v) &:= \int_\Omega \langle \mathbf{A}\nabla u, \nabla v\rangle + (\langle \mathbf{b}, \nabla u\rangle + cu)\, v = \int_\Omega fv \\ &=: F(v) \qquad \forall v \in H_0^1(\Omega). \end{aligned} \tag{2.2}$$

Assumption 1 The coefficients in (2.2) satisfy $\mathbf{A} \in C^{0,1}\left(\Omega, \mathbb{R}^{d\times d}_{\text{sym}}\right)$, $\mathbf{b} \in C^{0,1}\left(\Omega, \mathbb{R}^d\right)$, $c \in L^\infty(\Omega)$, and

$$\begin{aligned} 0 < \alpha &:= \inf_{x\in\Omega} \inf_{v\in\mathbb{R}^d\setminus\{0\}} \frac{\langle \mathbf{A}v, v\rangle}{\langle v, v\rangle} \le \sup_{x\in\Omega} \sup_{v\in\mathbb{R}^d\setminus\{0\}} \frac{\langle \mathbf{A}v, v\rangle}{\langle v, v\rangle} \\ &:= \beta < \infty \end{aligned}$$

$$0 \le \inf_{x\in\Omega} \left(c(x) - \frac{1}{2}\operatorname{div}\mathbf{b}(x)\right).$$

We set $c_\infty := \|c\|_{L^\infty(\Omega)}$ and $b_\infty := \max\left\{\|\mathbf{b}\|_{L^\infty(\Omega,\mathbb{R}^d)}, \|\operatorname{div}\mathbf{b}\|_{L^\infty(\Omega)}\right\}$. The energy norm is denoted by

$$\|v\|_{\mathrm{PDE}} := a(v,v)^{1/2},$$

where Assumption 1 implies that $\|\cdot\|_{\mathrm{PDE}}$ is a norm and Friedrichs inequality implies

$$\|v\|_{\mathrm{PDE}}^2 \geq \int_\Omega \langle \mathbf{A}\nabla v, \nabla v\rangle \geq \alpha \|\nabla v\|_{L^2(\Omega)}^2 \geq \frac{\alpha}{c_{\mathrm{F}}} \|v\|_{H^1(\Omega)}^2, \tag{2.3}$$

where c_{F} denotes the Friedrichs constant. In fact, the norms $\|\cdot\|_{\mathrm{PDE}}$ and $\|\cdot\|_{H^1(\Omega)}$ are equivalent since also

$$\|v\|_{\mathrm{PDE}}^2 \leq C_a \|v\|_{H^1(\Omega)}^2 \text{ with } C_a := \frac{b_\infty}{2} + \max\{c_\infty, \beta\}. \tag{2.4}$$

For a subdomain $\omega \subset \Omega$ we set

$$\|v\|_{\mathrm{PDE},\omega}^2 := a_\omega(v,v) := \int_\omega \left(\langle \mathbf{A}\nabla v, \nabla v\rangle + \langle \mathbf{b}, \nabla v\rangle v + cv^2\right).$$

Remark 2 The constants in the estimates below possibly depend (continuously) on α, β, c_∞, and b_∞ and might tend to infinity with increasing β, b_∞, c_∞, α^{-1}. We suppress the dependence in the notation.

Note that these conditions ensure that problem (2.2) is well posed and the coercivity estimate holds trivially

$$a(v,v) = \|v\|_{\mathrm{PDE}}^2 \qquad \forall v \in H_0^1(\Omega).$$

Assumption 1 implies the continuity of $a(\cdot,\cdot)$, i.e.,

$$a(u,v) \leq C_S \|v\|_{\mathrm{PDE}} \|v\|_{\mathrm{PDE}} \qquad \forall v \in H_0^1(\Omega) \tag{2.5}$$

with $C_S := 1 + c_{\mathrm{F}} \frac{b_\infty}{\alpha}$.

3 Conforming *hp*-finite elements

Let $\Omega \subset \mathbb{R}^2$ be a polygonal domain and let $\mathcal{T} := \{K_i : 1 \leq i \leq N\}$ denote a conforming simplicial finite element mesh. With each element $K \in \mathcal{T}$ we associate of polynomial degree $p_K \in \mathbb{N}_{\geq 1}$ which are collected into the polynomial degree vector $\mathbf{p} = (p_K)_{K\in\mathcal{T}}$. Then, we define the conforming *hp*-finite element space for the mesh $\mathcal{T}$ with local polynomials of degree p_K by

$$S_{\mathcal{T}}^{\mathbf{p}} := \left\{u \in H_0^1(\Omega) \mid \forall K \in \mathcal{T} \quad u|_K \in \mathbb{P}_{p_K}\right\}. \tag{3.1}$$

Here $\mathbb{P}_p$ denote the space of bivariate polynomials of maximal total degree p. For a subset $\omega \subset \Omega$, we write $\mathbb{P}_p(\omega)$ to indicate explicitly that we consider $u \in \mathbb{P}_p(\omega)$ as a polynomial on ω. Formally we define $\mathbb{P}_{-1} := \{0\}$. We set

$$p_{\mathcal{T}} := \max\{p_K : K \in \mathcal{T}\}.$$

By convention the triangles $K \in \mathcal{T}$ are open sets. The boundaries of the triangles $K \in \mathcal{T}$ consist of one-dimensional edges which are collected in the set $\mathcal{E}$. Furthermore, let $\mathcal{E}_\Omega := \{E \in \mathcal{E} \mid E \subset \Omega\}$. The union $\mathfrak{S}_\Omega := \bigcup_{E\in\mathcal{E}_\Omega} E$ forms the inner skeleton of the mesh $\mathcal{T}$. For each $E \in \mathcal{E}$ we fix one unit vector $\mathbf{n}_E$ which is perpendicular to E. If $E \subset \partial\Omega$, the orientation is chosen such that $\mathbf{n}_E$ points to the exterior of Ω. The $\mathcal{E}$-piecewise constant vector field $\mathbf{n}$ is given by $\mathbf{n}|_E := \mathbf{n}_E$. Finally we define the jump of some piecewise smooth function $g \in \prod_{K\in\mathcal{T}} H^1(K)$ across $E \in \mathcal{E}_\Omega$ by

$$[g]_E(x) := \lim_{\varepsilon \searrow 0} (g(x+\varepsilon\mathbf{n}_E) - g(x-\varepsilon\mathbf{n}_E)) \quad \forall x \in \overset{\circ}{E}.$$

This defines the jump function $[g]|_E := [g]_E$ for all $E \in \mathcal{E}_\Omega$ almost everywhere.

Let $\mathcal{N}_\Omega^1$ denote the set of inner vertices of $\mathcal{T}$. For $z \in \mathcal{N}_\Omega^1$, we denote by $b_z^1 \in S_{\mathcal{T}}^1$ the canonical continuous, piecewise affine basis function. The *volume star* for the node z is given by $\omega_z := \operatorname{supp} b_z^1$ and its measure is denoted by $|\omega_z|$. For $z \in \mathcal{N}_\Omega^1$, we set $\mathcal{E}_z := \left\{E \in \mathcal{E} : E \subset \mathfrak{S}_\Omega \cap \overset{\circ}{\omega}_z\right\}$ and $\mathcal{T}_z := \{K \in \mathcal{T} : K \subset \omega_z\}$. Let $\mathcal{V}_K$ denote the set of inner vertices of K and let $\omega_K := \bigcup_{z\in\mathcal{V}_K} \omega_z$.

We denote by $\nabla_{\mathcal{T}}$ the trianglewise gradient and by $\operatorname{div}_{\mathcal{T}}$ the trianglewise divergence operator. Let $\mathfrak{h}_{\mathcal{T}}$ denote the $\mathcal{T}$-piecewise constant function with values $\mathfrak{h}_{\mathcal{T}}|_K := \operatorname{diam} K$ for all $K \in \mathcal{T}$. Similarly we define $\mathfrak{h}_{\mathcal{E}} : \mathfrak{S}_\Omega \to \mathbb{R}$ as the $\mathcal{E}$-piecewise constant function $\mathfrak{h}_{\mathcal{E}}|_E := \operatorname{diam} E$ for all $E \in \mathcal{E}_\Omega$. The maximal mesh width in $\mathcal{T}$ is defined by

$$h_{\mathcal{T},\max} := \max\{\operatorname{diam} K : K \in \mathcal{T}\}.$$

If $\mathcal{T}$ is clear from the context we write h short for $h_{\mathcal{T},\max}$. The shape regularity of $\mathcal{T}$ is described by the constant

$$\rho_{\mathcal{T}} := \max\left\{\frac{\operatorname{diam} K}{\operatorname{diam} \mathcal{B}_K} : K \in \mathcal{T}\right\}, \tag{3.2}$$

where $\mathcal{B}_K$ is the maximal inscribed ball in K. Since $\mathcal{T}$ contains finitely many simplices the constant $\rho_{\mathcal{T}}$ is always bounded but becomes large if the simplices are degenerate, e.g., are flat or needle-shaped. The constants in the following estimates depend on the mesh via the constant $\rho_{\mathcal{T}}$—they are bounded for any fixed $\rho_{\mathcal{T}}$ but, possibly, become large for large $\rho_{\mathcal{T}}$.

Concerning the polynomial degree distribution we assume throughout the paper that the polynomial degrees of neighbouring elements are comparable[1]:

$$\rho_{\mathcal{T}}^{-1}(p_K+1) \leq p_{K'}+1 \leq \rho_{\mathcal{T}}(p_K+1) \quad \forall K, K' \in \mathcal{T} \text{ with } \overline{K} \cap \overline{K'} \neq \emptyset. \tag{3.3}$$

[1] We use here the same constant ρ as for the shape regularity to simplify the notation.

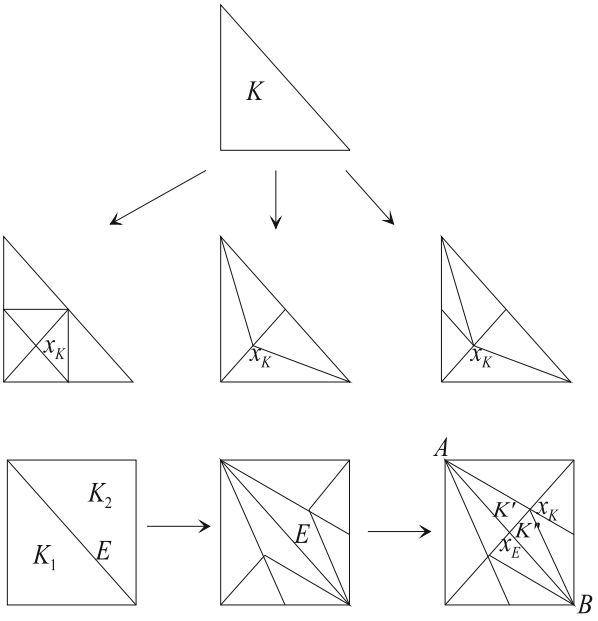

Fig. 1 Refinement patterns of a triangle which satisfies the interior node property. *Second row*, from left to right: Pic. 1: Regular refinement. Pic. 2,3: Refinement patterns for the elimination of hanging nodes. *Third row* if two triangles K_1, K_2 share an edge E and they will be both h-refined, then the common edge E must get an interior point x_E

The finite element solution is defined by:

$$\text{Find } u_{\mathcal{T}}^{\mathbf{p}} \in S_{\mathcal{T}}^{\mathbf{p}} \text{ such that } a\left(u_{\mathcal{T}}^{\mathbf{p}}, v\right) = F(v) \quad \forall v \in S_{\mathcal{T}}^{\mathbf{p}}. \tag{3.4}$$

In view of an adaptive solution process we generate a sequence $S_\ell := S_{\mathcal{T}_\ell}^{\mathbf{p}_\ell}$, $\ell \in \mathbb{N}_0$, of finite element spaces, where we require that all meshes $\mathcal{T}_\ell$ are conforming and the constants ρ_ℓ corresponding to the shape regularity of the mesh $\mathcal{T}_\ell$ and the polynomial degree vector $\mathbf{p}_\ell$ are uniformly bounded from above by some positive constant ρ. We also assume that $\mathcal{T}_{\ell+1}$ is a refinement of $\mathcal{T}_\ell$ in the sense that for any $K \in \mathcal{T}_\ell$ there is a subset $\text{sons}(K) \subset \mathcal{T}_{\ell+1}$ such that

$$\overline{K} = \bigcup_{K' \in \text{sons}(K)} \overline{K'}.$$

To reduce technicalities we make the following assumption concerning the concrete refinement method (cf. Fig. 1). As usual for conforming h-refinement, there exists two types of refinements. Some triangles are *marked* for refinement while this marking induces some additional refinement of neighbouring triangles in order to avoid *hanging nodes*.

Assumption 2 (a) A triangle K, which is *marked* for refinement, is *regularly refined* by connecting the midpoint of the edges as well as the midpoint of the longest side with the opposite vertex in K (cf. Fig. 1, Pic 1) so that the set $\text{sons}(K)$ contains six new triangles.

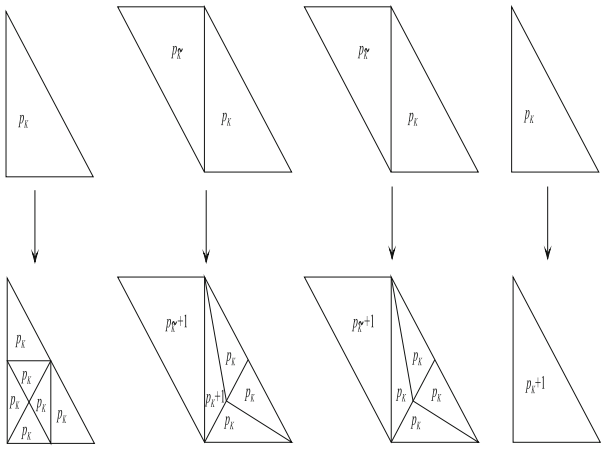

Fig. 2 Definition of the polynomial degrees. From left to right: Pic1: regular refinement. Pic 2: K is h-refined, $\widetilde{K}$ is p-refined and $p_{\widetilde{K}} \geq p_K$. Pic 3: K is h-refined, $\widetilde{K}$ is p-refined and $p_{\widetilde{K}} < p_K$. Pic 4: p-refinement

(b) To eliminate hanging nodes neighbouring triangles are refined by inserting a line L from one hanging node to the opposite vertex and connecting the vertices of K with the midpoint of L (cf. Fig. 1, Pic 2). If there is a further hanging node then this node is connected also with the midpoint of L (cf. Fig. 1, Pic 3). If K contains three hanging nodes or the shape regularity of the new triangles exceeds some threshold it will be regularly refined.

(c) For any triangle $K \in \mathcal{T}_\ell$, one of the following conditions are satisfied (cf. Fig. 2):

(i) K will be p-refined, i.e., $K \in \mathcal{T}_{\ell+1}$ and the polynomial degree is raised by 1.

(ii) K will be h-refined, i.e., there exists a set of sons $\sigma(K) \subset \mathcal{T}_{\ell+1}$ with $\overline{K} = \bigcup_{K' \in \text{sons}(K)} \overline{K'}$ and at least one vertex of each K' lies in the interior of K. The polynomial degree p_K defines the polynomial degree on $K' \in \text{sons}(K)$ as follows

$$p_{K'} := \begin{cases} p_K & \text{if } K \text{ is regularly refined,} \\ p_K + 1 \text{ if } \left(\begin{array}{l} \exists \widetilde{K} \in \mathcal{T}_\ell : \overline{K'} \cap \overline{\widetilde{K}} \text{ is a full} \\ \text{edge of } \widetilde{K} \\ \widetilde{K} \text{ is } p-\text{refined} \\ p_{\widetilde{K}} \geq p_K \end{array} \right), \\ p_K & \text{otherwise.} \end{cases}$$

Assumption 2 implies the *interior node property* (cf. [25, Sec. 3.4]).

Definition 1 (*interior node property*) Any $K \in \mathcal{T}_\ell$ which will be regularly h-refined and the three adjacent triangles $T' \in \mathcal{T}_\ell$ as well as their common sides contain a node of the finer mesh $\mathcal{T}_{\ell+1}$ in their interior and the resulting triangulation $\mathcal{T}_{\ell+1}$ has no hanging nodes.

Remark 3 Let $K_1, K_2 \in \mathcal{T}_\ell$ denote two triangles which share an edge E and let $p_m := p_{K_m}$, $m = 1, 2$. The condition

$u \in H_0^1(\Omega)$ in the definition of $S^{\mathbf{p}_\ell}_{\mathcal{T}_\ell}$ in (3.1) implies that for any $u \in S_\ell$ the one-dimensional polynomial degree of u along E satisfies

$$\deg(u|_E) \le p_E := \min\{p_1, p_2\}.$$

Notation 3 *To reduce the number of indices we write* u_ℓ *short for* $u^{\mathbf{p}_\ell}_{\mathcal{T}_\ell}$, $\mathfrak{h}_\ell$ *short for* $\mathfrak{h}_{\mathcal{T}_\ell}$, div_ℓ *short for* $\mathrm{div}_{\mathcal{T}_\ell}$, p_ℓ *short for* $p_{\mathcal{T}_\ell}$, $\mathcal{N}^1_\ell$ *for the inner triangle vertices for the mesh* $\mathcal{T}_\ell$, *etc. The star* ω_z *corresponds always to the triangulation* $\mathcal{T}_\ell$ *while we suppress this additional index in the notation of* ω_z.

Definition 2 The *saturation estimate* for a sequence of finite element solutions $(u_\ell)_\ell$ is an estimate of the form

$$\|u_{\ell+1} - u\|_{\mathrm{PDE}} \le \kappa_\ell \|u_\ell - u\|_{\mathrm{PDE}}$$

for some $\kappa_\ell < 1$ such that $\prod_{\ell=1}^{\infty} \kappa_\ell = 0$.

It was proved in [19] that the saturation estimate holds for the case of a bounded, two-dimensional domain Ω with coefficients

$$\mathbf{A} = \mathbf{I}, \mathbf{b} = \mathbf{0} \quad \text{and} \quad c = 0, \tag{3.5}$$

where $\mathbf{I}$ is the 2×2 unit matrix and the analysis was restricted to $\mathbb{P}_1$ finite elements with h-refinement. It was proved that it is necessary and sufficient for the saturation estimate that the *data oscillations* [which will be introduced in (6.5)] are controlled. Here, we generalize this result to the setting described in Sect. 2 and also derive p-explicit estimates for the contraction factor κ_ℓ.

4 Residual a posteriori error estimation

The Galerkin error is denoted by $e_\ell := u - u_\ell$. In the following, we will investigate under which condition the *saturation* estimate of the form

$$\|e_{\ell+1}\|_{\mathrm{PDE}} \le \kappa_\ell \|e_\ell\|_{\mathrm{PDE}}, \tag{4.1}$$

hold for some $\kappa_\ell \in]0, 1[$ depending only on the polynomial degree p and the shape-regularity of the mesh but not on the mesh width.

For the proof of the saturation estimate, we will use tools from residual a posteriori error estimation which we briefly recall: To obtain an a posteriori error estimate we obtain by Galerkin's orthogonality for every $v \in S_\ell$

$$\begin{aligned}\|e_\ell\|^2_{\mathrm{PDE}} = a(e_\ell, e_\ell - v) &= \int_\Omega \mathrm{res}(u_\ell)(e_\ell - v) \\ &+ \int_{\mathfrak{S}_\Omega} \mathrm{Res}(u_\ell)(e_\ell - v),\end{aligned} \tag{4.2}$$

where the volume residual $\mathrm{res} : S_\ell \to L^2(\Omega)$ is given by

$$\mathrm{res}(v) := f + \mathrm{div}_\ell(\mathbf{A}\nabla v) - \langle \mathbf{b}, \nabla v\rangle - cv$$

and the edge residual $\mathrm{Res} : S_\ell \to L^2(\mathfrak{S}_\Omega)$ is given by

$$\mathrm{Res}(v) := -\langle \mathbf{An}, [\nabla v]\rangle \quad \text{a.e. in } \mathfrak{S}_\Omega.$$

By choosing $v \in S_\ell$ as the Clément interpolation of e_ℓ and using a trace inequality for the last term in (4.2), results in the classical residual a posteriori error estimation. In [26,27] the local and global residual a posteriori error estimator is defined by

$$\begin{aligned}\eta^2_K(v) := &\left\| \frac{h_K}{p_K} \mathrm{res}(v) \right\|^2_{L^2(K)} \\ &+ \sum_{E\subset\partial K\cap\Omega} \left\| \sqrt{\frac{h_E}{2p_K}} \mathrm{Res}(v) \right\|^2_{L^2(E)} \quad \forall v \in S_\ell \quad \forall K \in \mathcal{T}_\ell.\end{aligned} \tag{4.3}$$

The global error estimator is given by

$$\eta_\ell(v) := \sqrt{\sum_{K\in\mathcal{T}_\ell} \eta^2_K(v)}.$$

Due to the finite overlap of the stars ω_z, the error estimator (4.3) is equivalent to

$$\begin{aligned}\eta^{\mathrm{star}}_\ell(v) &:= \sqrt{\sum_{z\in\mathcal{N}^1_\ell} \eta^2_z(v)} \quad \text{with} \\ \eta^2_z(v) &:= \left\| \frac{h_z}{p_z} \mathrm{res}(v) \right\|^2_{L^2(\omega_z)} \\ &\quad + \sum_{E\in\mathcal{E}_z} \left\| \sqrt{\frac{h_z}{p_z}} \mathrm{Res}(v) \right\|^2_{L^2(E)}\end{aligned} \tag{4.4}$$

and

$$\begin{aligned}p_z &:= \min\{p_K : K \subset \omega_z\} \quad \text{and} \\ h_z &:= \max\{h_K : K \subset \omega_z\}.\end{aligned} \tag{4.5}$$

Theorem 4 (Melenk, Wohlmuth) *Let* $\Omega \subset \mathbb{R}^2$ *be a bounded Lipschitz domain. Let* $a(\cdot,\cdot)$ *in (2.2) satisfy Assumption 1 and let* $f \in L^2(\Omega)$. *The solution of (2.2) is denoted by* u *and its Galerkin approximation by* u_ℓ *[see (3.4)]. There exists a constant* C_{rel} *independent of the local mesh width and the local polynomial degree but, possibly, depending on the constants in Assumption 1 such that*

$$\|u - u_\ell\|_{\mathrm{PDE}} \le C_{\mathrm{rel}} \eta_\ell(u_\ell) \le C_{\mathrm{rel}} \eta^{\mathrm{star}}_\ell(u_\ell).$$

The proof of this theorem is a slight modification of [27, Theorem 3.6] and we include it here for completeness.

Proof The error $u - u_\ell$ can be estimated by using (4.2) and by setting $w = e_\ell - Ie_\ell$ with the hp-Clément interpolation operator I as in [27, Section 2.1]:

$$\|e_\ell\|^2_{\mathrm{PDE}} = \int_\Omega w \ \mathrm{res}\ (u_\ell) + \int_{\mathfrak{S}_\Omega} w \ \mathrm{Res}\,(u_\ell)$$
$$\leq C_{\mathrm{rel}} \eta_\ell\,(u_\ell)\,\|e_\ell\|_{\mathrm{PDE}}$$

for all $v \in S_\ell$. Clearly we have

$$\eta_\ell\,(v) \leq \eta_\ell^{\mathrm{star}}\,(v) \leq C_\sharp \eta_\ell\,(v)\,,$$

where $C_\sharp$ depends only on the constant ρ_ℓ in (3.2) and (3.3). □

5 Projection of polynomials onto piecewise polynomials

The proof of the saturation estimate is based on estimates of some projection of the volume residual to the space of piecewise polynomials locally on stars ω_z. In this section, we will derive stability estimates for this projection.

We start with a result of a weighed L^2 projection of *global* polynomials of maximal total degree p to *piecewise* polynomials of lower degree. The setting is as follows.

Let $z \in \mathbb{R}^2$ and let $\mathcal{T}_z := \{K_i : 1 \leq i \leq q\}$ denote a triangle patch around z, i.e., $\mathcal{T}_z$ is a set of (open) triangles which

- are pairwise disjoint,
- share z as a common vertex.
- For all $1 \leq i \leq q$, the triangles K_{i-1} and K_i share one common edge.[2]

Let[3] $\omega_z := \mathrm{int}\left(\bigcup_{i=1}^q \overline{K_i}\right)$ and let $\mathfrak{S} := \omega_z \cap \bigcup_{i=1}^q \partial K_i$ denote the inner mesh skeleton. We denote by $\mathbb{P}_p\,(\mathcal{T}_z)$ the space of *piecewise* polynomials, i.e.,

$$\mathbb{P}_p\,(\mathcal{T}_z) := \left\{f : \omega_z \backslash \mathfrak{S} \to \mathbb{R} \mid \forall 1 \leq i \leq q; \right.$$
$$\left. f|_{K_i} \in \mathbb{P}_p\,(K_i)\right\}. \tag{5.1}$$

Next, we will introduce weighted scalar products and associated norms. The weights are defined triangle- and edgewise and depend whether the triangle will be h-refined of p-refined.

Definition 3 (a) p-refinement.
If K will be p-refined, then, the *cubic weight function* $\Phi_K^{(3)}$ and *quadratic edge bubble* Φ_E are given, on the reference element $\widehat{K} := \mathrm{conv}\left(\binom{0}{0}, \binom{1}{0}, \binom{0}{1}\right)$ and on the reference interval $\widehat{E} := (0, 1)$, by

$$\Phi_{\widehat{K}}^{(3)}\,(x_1, x_2) = (1 - x_1 - x_2)\,x_1 x_2 \quad \text{and}$$
$$\Phi_{\widehat{E}}\,(x) = x\,(1 - x)\,, \tag{5.2}$$

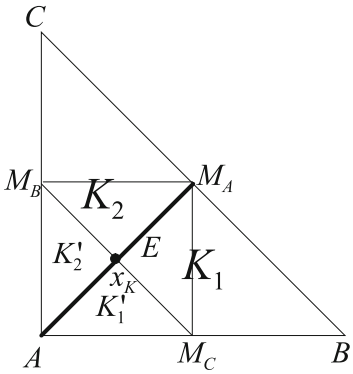

Fig. 3 Illustration for the notation of a regulary h-refined triangle K. The edge $E = \overline{AM_A}$ splits K into the two triangles K_1 and K_2. The subtriangle K_1' has vertices A, M_C, x_K while the vertices of K_2' are A, x_K, M_B

while on K and E we set

$$\Phi_K^{(3)} := \Phi_{\widehat{K}}^{(3)} \circ \Lambda_K^{-1} \quad \text{and} \quad \Phi_E := \Phi_{\widehat{E}} \circ \Lambda_E^{-1}, \tag{5.3}$$

where $\Lambda_K : \widehat{K} \to K$ and $\Lambda_E : \widehat{E} \to E$ are affine pullbacks.[4]

(b) h-refinement.
The *edge bubble* Φ_E for h-refinement is the same as for p-refinement.

(b1) Let K be regularly refined (cf. Fig. 3). Then, $\Phi_{K,K}^{(1)}$ is the piecewise linear function on the submesh sons (K) which has value 1 at x_K and value 0 at all other vertices of the refined mesh. Let E denote the edge as indicated in Fig. 3 which splits K into the triangles K_1 and K_2.
Then $\Phi_K^{(2)}$ is the product of the barycentric coordinates for the two endpoints of E with respect to the two triangles K_1 and K_2.

(b2) If K is non-regularly h-refined (cf. Fig. 1, Pic. 2,3), then the weight function for K is the piecewise *linear* bubble function $\Phi_K^{(1)}$ which interpolates $\Phi_K^{(3)}$ at the vertices of the submesh sons (K).

The *weight function for a triangle* K is

$$\Phi_K := \begin{cases} \Phi_K^{(3)} & \text{if } K \text{ will be } p\text{-refined,} \\ \Phi_K^{(1)} & \text{if } K \text{ will be non-regularly } h\text{-refined,} \\ \Phi_{K,K}^{(1)} + \Phi_K^{(2)} & \text{if } K \text{ will be regularly } h\text{-refined.} \end{cases} \tag{5.4}$$

For $z \in \mathcal{N}_\Omega^1$, the function $\Phi_z : \omega_z \to \mathbb{R}$ is given by

$$\Phi_z|_K := \Phi_K \qquad \forall K \in \omega_z \tag{5.5}$$

and extended by zero to Ω.

[2] We use here the convention $K_0 := K_q$. Clearly $q \geq 3$ holds.

[3] For a subset $\omega \subset \mathbb{R}^2$, we denote by $\mathrm{int}\,(\omega)$ the open interior of ω.

[4] Note that the scalings compared to the scalings in [39, p. 83] differ by fixed constants of order 1.

These weight functions induce bilinear forms $(\cdot,\cdot)_K$ and $(\cdot,\cdot)_z$ via

$$(u,v)_K := \int_K \Phi_K uv \quad \text{and}$$
$$(u,v)_z := \sum_{K\subset\omega_z} (u,v)_K = \int_{\omega_z} \Phi_z uv$$

and a corresponding norm $\|\cdot\|_z := (\cdot,\cdot)_z^{1/2}$.

Next we define a projection
$\Pi_K^p : \mathbb{P}_p(K) \to \mathbb{P}_{p-1}(K)$
by

$$\int_K \Phi_K \left(\Pi_K^p v\right) w = \int_K \Phi_K vw \qquad \forall w \in \mathbb{P}_{p-1}(K), \quad (5.6)$$

where the definition of Φ_K is as in (5.4), i.e., depends on how K will be refined.

Definition 4 For a triangle patch $\mathcal{T}_z$, let p_z be as in (4.5). The *star-wise polynomial projection* Π_z is applied to polynomials $v \in \mathbb{P}_{p_z-1}(\omega_z)$ and given by

$$(\Pi_z v)|_K := \begin{cases} \Phi_K^{(3)} \Pi_K^{p_z-1} v & \text{if } K \text{ is } p\text{-refined,} \\ \Phi_K^{(1)} v & \text{if } K \text{ is non-regularly } h\text{-refined,} \\ \Phi_{K,K}^{(1)} v + \Phi_K \Pi_K^{p_z-1} v & \text{if } K \text{ is regularly } h\text{-refined,} \end{cases}$$

Theorem 5 *Let $p \geq 1$. For all $u \in \mathbb{P}_p(\omega_z)$, the condition*

$$\sum_{K\in\mathcal{T}_z} \int_K \Phi_K^{(3)} uw = 0 \qquad \forall w \in \mathbb{P}_{p-1}(\mathcal{T}_z). \quad (5.7)$$

implies $u = 0$.

For a proof we refer to [24, Theorem 1.1]. A consequence of Theorem 5 is the following corollary. To reduce technicalities we make an assumption on the minimal local polynomial degree.

Assumption 6 For all ℓ and $z \in \mathcal{N}_\ell^1$ it holds: If all $K \subset \omega_z$ will be p-refined then $p_z \geq 2$ otherwise $p_z \geq 1$.

Corollary 1 *Let Assumption 6 be valid. The projection Π_z is injective.*

Proof If all triangles in ω_z are p-refined, then the injectivity follows from Theorem 5.

If, at least, one triangle is h-refined we distinguish between two cases:

(a) K is non-regularly h-refined. Then, the positivity of $\Phi_K^{(1)}$ implies $(\Pi_z v)|_K = 0 \implies v = 0$.
(b) K is regularly h-refined. We use the notation as introduced in Fig. 3. Note that $(\Pi_z v)|_{K_1'\cup K_2'} := \left(\Phi_{K,K}^{(1)}\left(v + \Pi_K^{p_z-1} v\right)\right)\Big|_{K_1'\cup K_2'}$.
(b1) If the degree of v satisfies $\deg v = p_z - 2$, it holds $\Pi_K^{p_z-1} v = v|_K$ Then, $\Pi_z v|_{K_1'\cup K_2'} = 2\,\Phi_{K,K}^{(1)} v\Big|_{K_1'\cup K_2'}$. The positivity of $\Phi_{K,K}^{(1)}$ on $K_1'\cup K_2'$ together with the analytic continuation principle, i.e., $v|_{K_1'\cup K_2'} = 0 \implies v = 0$, imply the injectivity of Π_z for this case.
(b2) If $\deg v = p_z - 1$, it holds $v + \Pi_K^{p_z-1} v \neq 0$. The positivity of $\Phi_{K,K}^{(1)}$ again implies $(\Pi_z v)|_{K_1'\cup K_2'} \neq 0$. □

Corollary 2 *Let Assumption 6 be valid. For all $z \in \mathcal{N}_\Omega^1$, the estimates*

$$\inf_{v\in\mathbb{P}_{p-1}(\omega_z)\setminus\{0\}} \frac{(v,\Pi_z v)_{L^2(\omega_z)}}{\|v\|_z^2} \geq c_\pi. \quad (5.8)$$

$$c\left\|\Phi_z^{-1/2}\Pi_z v\right\|_{L^2(\omega_z)} \leq (v,\Pi_z v)_{L^2(\omega_z)}^{1/2} \leq \|v\|_z \leq \|v\|_{L^2(\omega_z)}. \quad (5.9)$$

hold. The constant c_π in (5.8) satisfies $0 < c_\pi \leq 1$ and depends, possibly, on the polynomial degree p and the shape regularity of the mesh.

Proof For the proof of (5.8), we distinguish between the following cases.

(a) If all triangles in ω_z are p-refined, estimate (5.8) for some constant $c_\pi > 0$ follows from the injectivity of Π_z via the compactness argument in [24, Theorem 6.4] and the equivalence of norms on the finite dimensional space $\mathbb{P}_{p_z-1}(\omega_z)$.
(b) At least one triangle in ω_z is h-refined. Let $K \subset \omega_z$.
(b1) K is non-regularly h-refined. Then, the positivity of $\Phi_K^{(1)}$ implies $(v,\Pi_z v)_{L^2(K)} = \left\|\Phi_K^{1/2} v\right\|_{L^2(K)}^2 > 0$ for all $v \in \mathbb{P}_{p_z-1}(K)\setminus\{0\}$.
(b2) K is regularly h-refined. We use the notation as introduced in Fig. 3. Then,

$$\begin{aligned}
&(v,\Pi_z v)_{L^2(K)} \\
&= \left(v, \Phi_{K,K}^{(1)} v\right)_{L^2(K)} + \left(v, \Phi_K \Pi_K^{p_z-1} v\right)_{L^2(K)} \\
&= \left(v, \Phi_{K,K}^{(1)} v\right)_{L^2(K)} + \left(\Pi_K^{p_z-1} v, \Phi_K \Pi_K^{p_z-1} v\right)_{L^2(K)} \\
&= \left\|\sqrt{\Phi_{K,K}^{(1)}} v\right\|_{L^2(K)}^2 + \left\|\sqrt{\Phi_K}\Pi_K^{p_z-1} v\right\|_{L^2(K)}^2 \\
&\geq \left\|\sqrt{\Phi_{K,K}^{(1)}} v\right\|_{L^2(K)}^2.
\end{aligned}$$

Again, the positivity of $\Phi_{K,K}^{(1)}$ on $K_1'\cup K_2'$ implies $(v,\Pi_z v)_{L^2(K)} > 0$ for all $v \in \mathbb{P}_{p_z-1}(K)\setminus\{0\}$.

For the estimate (5.9) we again consider the different refinement options separately. It is easy to check that pointwise on K, we have $\Phi_K \le 1$ so that $\Phi_K^2 \le \Phi_K$.

(a) K is p-refined, i.e., $\Phi_K := \Phi_K^{(3)}$. Estimate

$$(v, \Pi_z v)_{L^2(K)} = \int_K \Phi_K v \Pi_K^{p_z-1} v \le \left\| \sqrt{\Phi_K} v \right\|^2_{L^2(K)} \tag{5.10a}$$

holds since $\Pi_K^{p_z-1}$ is a projection. On the other hand,

$$\begin{aligned}(v, \Pi_z v)_{L^2(K)} &= \int_K \Phi_K \left(\Pi_K^{p_z-1} v\right)^2 \\ &= \left\| \Phi_K^{-1/2} \Pi_z v \right\|^2_{L^2(K)}. \end{aligned} \tag{5.11a}$$

(b) K is non-regularly h-refined. Then,

$$(v, \Pi_z v)_{L^2(K)} = \left\| \sqrt{\Phi_K} v \right\|^2_{L^2(K)}. \tag{5.10b}$$

From (5.10b) we get

$$\begin{aligned}(v, \Pi_z v)_{L^2(K)} &= \left\| \Phi_K^{-1/2} \Phi_K v \right\|^2_{L^2(K)} \\ &= \left\| \Phi_K^{-1/2} \Pi_z v \right\|^2_{L^2(K)}. \end{aligned} \tag{5.11b}$$

(c) K is regularly h-refined. Then,

$$\begin{aligned}&(v, \Pi_z v)_{L^2(K)} \\ &= \left\| \sqrt{\Phi_{K,K}^{(1)}} v \right\|^2_{L^2(K)} + \left\| \sqrt{\Phi_K} \Pi_K^{p_z-1} v \right\|^2_{L^2(K)} \\ &\le \left\| \sqrt{\Phi_{K,K}^{(1)}} v \right\|^2_{L^2(K)} + \left\| \sqrt{\Phi_K} v \right\|^2_{L^2(K)} \\ &\le 2 \left\| \sqrt{\Phi_K} v \right\|^2_{L^2(K)}. \end{aligned} \tag{5.10c}$$

For the first estimate in (5.9) we use the pointwise estimate on K

$$\begin{aligned}(\Pi_z v)^2 &= \left(\Phi_{K,K}^{(1)} v + \Phi_K \Pi_K^{p_z-1} v\right)^2 \\ &\le 2 \left(\left(\Phi_{K,K}^{(1)} v\right)^2 + \left(\Phi_K \Pi_K^{p_z-1} v\right)^2 \right) \\ &\le 2 \left(\Phi_{K,K}^{(1)} + \Phi_K\right) \left(\Phi_{K,K}^{(1)} v^2 + \Phi_K \left(\Pi_K^{p_z-1} v\right)^2 \right) \\ &\le 4 \Phi_K \left(\Phi_{K,K}^{(1)} v^2 + \Phi_K \left(\Pi_K^{p_z-1} v\right)^2 \right) \end{aligned}$$

to get

$$\begin{aligned}&\int_K \Phi_K^{-1} (\Pi_z v)^2 \\ &\le 4 \left(\left\| \sqrt{\Phi_{K,K}^{(1)}} v \right\|^2_{L^2(K)} + \left\| \sqrt{\Phi_K} \Pi_K^{p_z-1} v \right\|^2_{L^2(K)} \right) \\ &= 4 (v, \Pi_z v)_{L^2(K)}. \end{aligned} \tag{5.11c}$$

The second estimate in (5.9) follows by summing the inequality (5.11c) over all $K \subset \omega_z$ while the third one is a consequence of $\|v\|_z \le \|v\|_{L^2(\omega_z)}$ since $0 \le \Phi_z \le 1$.

The first estimate in (5.9) also follows by summation over all $K \subset \omega_z$ the inequalities (5.11). □

The derivation of a sharp positive lower bound for c_π seems to rather involved. Instead we have performed numerical experiments (cf. "Appendix 1") to support the following conjecture.

Conjecture 1 *The constant c_π is bounded from below by a constant $c_0 > 0$ which only depends on the shape regularity of the mesh but neither on the mesh width nor on the polynomial degree p.*

For $z \in \mathcal{N}_\ell^1$, we introduce the subspaces for $K \in \mathcal{T}_\ell$ (recall Notation 3)

$$\begin{aligned}&S_{\ell+1,K} := \{u \in S_{\ell+1} \mid \operatorname{supp} u \subset K\} \quad \text{and} \\ &S_{\ell+1,z} := \{u \in S_{\ell+1} \mid \operatorname{supp} u \subset \omega_z\}. \end{aligned} \tag{5.12}$$

Theorem 7 *Let Assumptions 2 and 6 be satisfied. For $z \in \mathcal{N}_\ell^1$, $E \in \mathcal{E}_z$, and $w \in S_\ell$, let $J_E(w) := \langle \mathbf{An}_E, [\nabla w]_E \rangle$. Set $p := p_z$ [cf. (4.5)]. For any $v \in \mathbb{P}_{p-1}(\omega_z)$, there exists a $\varphi_{\ell+1,z} \in S_{\ell+1,z}$ such that*

$$\sum_{E \in \mathcal{E}_z} \int_E J_E b_z^1 = \sum_{E \in \mathcal{E}_z} \int_E J_E \varphi_{\ell+1,z}, \tag{5.13a}$$

$$c_\pi \left\| \mathfrak{h}_\ell v \right\|_z \le \left| \int_{\omega_z} v \left(b_z^1 - \varphi_{\ell+1,z} \right) \right|, \tag{5.13b}$$

$$\begin{aligned}&\left\| \Phi_z^{-1/2} \left(b_z^1 - \varphi_{\ell+1,z} \right) \right\|_{L^2(\omega_z)} \\ &\quad + c_2 \frac{h_z}{p} \left\| \left(b_z^1 - \varphi_{\ell+1,z} \right) \right\|_{\mathrm{PDE},\omega_z} \le C_1 h_z. \end{aligned} \tag{5.13c}$$

The constant $c_2 > 0$ only depends on $\alpha, \beta, b_\infty, c_\infty$, and the shape-regularity of the mesh while C_1 is a number.

Proof We make the ansatz

$$\varphi_{\ell+1,z} = b_z^1 - \psi_{\ell+1,z},$$

for some $\psi_{\ell+1,z}$ with $\psi_{\ell+1,z}\big|_K \in S_{\ell+1,K}$ for all $K \subset \omega_z$. Hence $\psi_{\ell+1,z}\big|_K$ vanishes on all edges and condition (5.13a) trivially is satisfied.

Statement (5.13b) is trivial for $v = 0$ and we consider here $v \in \mathbb{P}_{p-1}(\omega_z) \setminus \{0\}$. Let

$$\psi_{\ell+1,z} = h_z \frac{\Pi_z v}{\|v\|_z}$$

and observe that $\psi_{\ell+1,z} \in S_{\ell+1,z}$. Hence, by Corollary 2 we obtain

$$\left| \int_{\omega_z} v \left(b_z^1 - \varphi_{\ell+1,z} \right) \right| = h_z \frac{\left| (v, \Pi_z v)_{L^2(\omega_z)} \right|}{\|v\|_z} \geq c_\pi h_z \|v\|_z .$$

Finally, we consider estimate (5.13c) and get

$$\left\| \Phi_z^{-1/2} \left(b_z^1 - \varphi_{\ell+1,z} \right) \right\|_{L^2(\omega_z)} = h_z \frac{\left\| \Phi_z^{-1/2} \Pi_z v \right\|_{L^2(\omega_z)}}{\|v\|_z} \overset{(5.9)}{\leq} \frac{h_z}{c} \frac{(v, \Pi_z v)_{L^2(\omega_z)}^{1/2}}{\|v\|_z} \overset{(5.9)}{\leq} \tilde{C} h_z. \tag{5.14}$$

For the H^1-seminorm we get

$$\left\| \nabla \left(b_z^1 - \varphi_{\ell+1,z} \right) \right\|_{L^2(\omega_z)} = h_z \left\| \nabla \left(\frac{\Pi_z v}{\|v\|_z} \right) \right\|_{L^2(\omega_z)} . \tag{5.15}$$

We distinguish again three cases.

Case a Let $K \subset \omega_z$ be a triangle which will be p-refined. Hence, $\Phi_K = \Phi_K^{(3)}$ (cf. Definition 3).

We apply Lemma 3 to obtain

$$\|\nabla (\Pi_z v)\|_{L^2(K)}^2 \leq C_0' \frac{p^2}{h_z^2} \left\| \Phi_K^{1/2} \Pi_K^{p_z - 1} v \right\|_{L^2(K)}^2 \leq C_0' \frac{p^2}{h_z^2} \left\| \Phi_K^{1/2} v \right\|_{L^2(K)}^2 . \tag{5.16a}$$

The last inequality in (5.16a) is trivial for $p_K > p$ since $(\Pi_z v)|_K = v$, while, for $p_K = p$, we employ (5.10a) and (5.11a).

Case b Let $K \subset \omega_z$ be a triangle which is non-regularly h-refined. Hence, $\Phi_K = \Phi_K^{(1)}$. We introduce the function[5] $d_K : K \to \mathbb{R}$ by

$$d_K = d_{\widehat{K}} \circ \Lambda_K^{-1} \quad \text{with} \quad d_{\widehat{K}}(x) := \operatorname{dist}\left(x, \partial \widehat{K}\right),$$

where Λ_K is as in (5.3). Since both, $\Phi_K^{(1)}$ and d_K are piecewise linear bubble functions with maximal value $\mathcal{O}(1)$ in the interior it is easy to verify that the pointwise estimates hold

$$\left.\begin{array}{c} c d_K \leq \Phi_K^{(1)} \leq C d_K \\ c \|\nabla d_K\| \leq \left\| \nabla \Phi_K^{(1)} \right\| \leq C \|\nabla d_K\| \end{array}\right\} \quad \text{a.e.} \tag{5.17}$$

with fixed constants $0 < c, C = O(1)$. Estimates (5.17) imply the pointwise estimate

$$\frac{1}{2} \|\nabla (v \Phi_K)\|^2 \leq \Phi_K^2 \|\nabla v\|^2 + \|\nabla \Phi_K\|^2 v^2 \leq C^2 \left(d_K^2 \|\nabla v\|^2 + v^2 \|\nabla d_K\|^2 \right).$$

[5] The function d_K differs from the function Φ_K in [27, (27)] only by a scaling constant which is of order 1.

Hence, we may use [27, (23) with $\delta = 1$ and (22) with $\alpha = 0$ and $\beta = 1$] to obtain

$$\frac{1}{2} \|\nabla (\Pi_z v)\|_{L^2(K)}^2 = \frac{1}{2} \|\nabla (v \Phi_K)\|_{L^2(K)}^2 \leq \frac{C^2}{h_K^2} \left(p^2 \left\| \Phi_K^{1/2} v \right\|_{L^2(K)}^2 + \|v\|_{L^2(K)}^2 \right) \leq \tilde{C}^2 \frac{p^2}{h_K^2} \left\| \Phi_K^{1/2} v \right\|_{L^2(K)}^2 . \tag{5.16b}$$

Case c Let $K \subset \omega_z$ be a triangle which is regularly h-refined. We employ the notation as explained in Definition 3(b1); illustrated in Fig. 3. It holds

$$\nabla \left((\Pi_z v)|_K \right) = \nabla \left(\Phi_{K,K}^{(1)} v + \Phi_K \Pi_K^{p_z - 1} v \right) = \nabla \left(\Phi_{K,K}^{(1)} \left(v + \Pi_K^{p_z - 1} v \right) \right) + \nabla \left(\Phi_K^{(2)} \Pi_K^{p_z - 1} v \right).$$

For the first term and the piecewise linear bubble $\Phi_{K,K}^{(1)}$ we can argue as in Case b to obtain

$$\left\| \nabla \left(\Phi_{K,K}^{(1)} \left(v + \Pi_K^{p_z - 1} v \right) \right) \right\|_{L^2(K)}^2 \leq \tilde{C}^2 \frac{p^2}{h_K^2} \left\| \sqrt{\Phi_{K,K}^{(1)}} \left(v + \Pi_K^{p_z - 1} v \right) \right\|_{L^2(K)}^2 , \tag{5.18a}$$

while we employ Lemma 4 for

$$\left\| \nabla \left(\Phi_K^{(2)} \Pi_K^{p_z - 1} v \right) \right\|_{L^2(K)}^2 \leq \hat{C}^2 \frac{p^2}{h_K^2} \left\| \sqrt{\Phi_{K,K}^{(1)}} \left(\Pi_K^{p_z - 1} v \right) \right\|_{L^2(K)}^2 . \tag{5.18b}$$

Thus,

$$\begin{aligned} &\left\| \nabla \left((\Pi_z v)|_K \right) \right\|_{L^2(K)}^2 \\ &\leq 2 \left(\left\| \nabla \left(\Phi_{K,K}^{(1)} \left(v + \Pi_K^{p_z - 1} v \right) \right) \right\|_{L^2(K)}^2 + \left\| \nabla \left(\Phi_K^{(2)} \Pi_K^{p_z - 1} v \right) \right\|_{L^2(K)}^2 \right) \\ &\overset{(5.18)}{\leq} 2 \left(\tilde{C}^2 + \hat{C}^2 \right) \frac{p^2}{h_K^2} \left(\left\| \sqrt{\Phi_{K,K}^{(1)}} \left(v + \Pi_K^{p_z - 1} v \right) \right\|_{L^2(K)}^2 + \left\| \sqrt{\Phi_K^{(2)}} \left(\Pi_K^{p_z - 1} v \right) \right\|_{L^2(K)}^2 \right) \\ &\leq 4 \left(\tilde{C}^2 + \hat{C}^2 \right) \frac{p^2}{h_K^2} \\ &\quad \times \left(\left\| \sqrt{\Phi_{K,K}^{(1)}} v \right\|_{L^2(K)}^2 + \left\| \sqrt{\Phi_{K,K}^{(1)}} \Pi_K^{p_z - 1} v \right\|_{L^2(K)}^2 \right. \end{aligned} \tag{5.16c}$$

$$
\begin{aligned}
&+ \left\| \sqrt{\Phi_K^{(2)}} \left(\Pi_K^{p_z-1} v \right) \right\|_{L^2(K)}^2 \Bigg) \\
&= 4 \left(\tilde{C}^2 + \hat{C}^2 \right) \frac{p^2}{h_K^2} \\
&\quad \times \left(\left\| \sqrt{\Phi_{K,K}^{(1)}} v \right\|_{L^2(K)}^2 + \left\| \sqrt{\Phi_K} \Pi_K^{p_z-1} v \right\|_{L^2(K)}^2 \right) \\
&\overset{(5.16a)}{\leq} 4 \left(\tilde{C}^2 + \hat{C}^2 \right) \frac{p^2}{h_K^2} \\
&\quad \times \left(\left\| \sqrt{\Phi_{K,K}^{(1)}} v \right\|_{L^2(K)}^2 + \left\| \sqrt{\Phi_K} v \right\|_{L^2(K)}^2 \right) \\
&\leq 8 \left(\tilde{C}^2 + \hat{C}^2 \right) \frac{p^2}{h_K^2} \left\| \sqrt{\Phi_K} v \right\|_{L^2(K)}^2 .
\end{aligned}
$$

The combination of (5.16) with (5.15) leads to

$$\left\| \nabla \varphi_{\ell+1,z} \right\|_{L^2(\omega_z)} \leq \tilde{C}_0 p.$$

□

6 The saturation property

Note that the Pythagoras theorem

$$\|e_\ell\|_{\mathrm{PDE}}^2 = \|e_{\ell+1}\|_{\mathrm{PDE}}^2 + \|u_\ell - u_{\ell+1}\|_{\mathrm{PDE}}^2$$

only holds for symmetric bilinear forms, i.e., $\mathbf{b} = \mathbf{0}$ in (2.2). For non-symmetric bilinear forms one can prove a quasi-orthogonality and we follow here [25, Proof of Lemma 2.1.]. One ingredient in the proof is an Aubin-Nitsche argument (see, e.g., [16]) which we recall here. For $0 < s \leq 1$, we say that the adjoint problem

$$\text{For given } g \in L^2(\Omega) \quad \text{find } \psi_g \in H_0^1(\Omega) \quad \text{such that}$$
$$a(v, \psi_g) := \int_\Omega g v \qquad \forall v \in H_0^1(\Omega)$$

is $H^{1+s}(\Omega)$-*regular* if, for any right-hand side $g \in L^2(\Omega)$, the solution ψ_g is in $H^{1+s}(\Omega)$ and there exists a constant C_s independent of $g \in L^2(\Omega)$ such that

$$\left\| \psi_g \right\|_{H^{1+s}(\Omega)} \leq C_s \|g\|_{L^2(\Omega)} .$$

We introduce the adjoint approximation property for a subspace $S \subset H_0^1(\Omega)$ by

$$\eta(S) := \sup_{g \in L^2(\Omega)\setminus\{0\}} \inf_{v \in S} \frac{\left\| \psi_g - v \right\|_{\mathrm{PDE}}}{\|g\|_{L^2(\Omega)}}.$$

In our context, we obtain, e.g., from [16] the estimate

$$\|e_{\ell+1}\|_{L^2(\Omega)} \leq C_S \eta(S_{\ell+1}) \|e_{\ell+1}\|_{\mathrm{PDE}} .$$

If the adjoint problem is $H^{1+s}(\Omega)$-regular, standard approximation results for finite elements lead to

$$
\begin{aligned}
\eta(S_\ell) &\leq C_a \sup_{g \in L^2(\Omega)\setminus\{0\}} \inf_{v \in S_\ell} \frac{\left\| \psi_g - v \right\|_{H^1(\Omega)}}{\|g\|_{L^2(\Omega)}} \\
&\leq C_a C_{\mathrm{approx}} h_\ell^s \sup_{g \in L^2(\Omega)\setminus\{0\}} \frac{\left\| \psi_g \right\|_{H^{1+s}(\Omega)}}{\|g\|_{L^2(\Omega)}} \\
&\leq C_a C_{\mathrm{approx}} C_s h_\ell^s ,
\end{aligned}
$$

where C_{approx} only depends on the shape regularity of the mesh. Hence,

$$
\begin{aligned}
&\|e_{\ell+1}\|_{L^2(\Omega)} \leq C_{\mathrm{dual}} h_{\ell+1}^s \|e_{\ell+1}\|_{\mathrm{PDE}} \quad \text{with} \\
&C_{\mathrm{dual}} := C_S C_a C_{\mathrm{approx}} C_s .
\end{aligned} \tag{6.1}
$$

Lemma 1 *Let Assumption 1 be satisfied and let the adjoint problem be $H^{1+s}(\Omega)$ regular for some $0 < s \leq 1$. Then, there exists some $C_\star > 0$ depending only on $\alpha, \beta, b_\infty, c_\infty$, and the shape regularity of the mesh such that, for any finite element mesh $\mathcal{T}_{\ell+1}$ with maximal mesh width $h_{\ell+1} < C_\star^{-s}$, the quasi-orthogonality*

$$
\begin{aligned}
&\|e_{\ell+1}\|_{\mathrm{PDE}}^2 \leq \Lambda_{\ell+1}^2 \|e_\ell\|_{\mathrm{PDE}}^2 - \|u_{\ell+1} - u_\ell\|_{\mathrm{PDE}}^2 \\
&\text{with} \qquad \Lambda_{\ell+1}^2 := \frac{1}{1 - C_\star h_{\ell+1}^s}
\end{aligned} \tag{6.2}
$$

holds.

The proof is adapted from [25, Lem. 4.1] and included here for completeness.

Proof We set $\widehat{\varepsilon}_\ell := u_{\ell+1} - u_\ell$. It is easy to conclude from Galerkin's orthogonality that

$$\|e_\ell\|_{\mathrm{PDE}}^2 = \|e_{\ell+1}\|_{\mathrm{PDE}}^2 + \|\widehat{\varepsilon}_\ell\|_{\mathrm{PDE}}^2 + a(\widehat{\varepsilon}_\ell, e_{\ell+1})$$

holds. Then, integration by parts yields

$$
\begin{aligned}
a(\widehat{\varepsilon}_\ell, e_{\ell+1}) &= a(e_{\ell+1}, \widehat{\varepsilon}_\ell) + \int_\Omega (\langle \mathbf{b}, \nabla \widehat{\varepsilon}_\ell \rangle e_{\ell+1} \\
&\qquad - \langle \mathbf{b}, \nabla e_{\ell+1} \rangle \widehat{\varepsilon}_\ell) \\
&= \int_\Omega (2 \langle \mathbf{b}, \nabla \widehat{\varepsilon}_\ell \rangle + (\operatorname{div} \mathbf{b}) \widehat{\varepsilon}_\ell) e_{\ell+1}.
\end{aligned}
$$

Hence,

$$
\begin{aligned}
\|e_{\ell+1}\|_{\mathrm{PDE}}^2 = \|e_\ell\|_{\mathrm{PDE}}^2 - \|\widehat{\varepsilon}_\ell\|_{\mathrm{PDE}}^2 - \int_\Omega (2 \langle \mathbf{b}, \nabla \widehat{\varepsilon}_\ell \rangle \\
+ (\operatorname{div} \mathbf{b}) \widehat{\varepsilon}_\ell) e_{\ell+1}.
\end{aligned} \tag{6.3}
$$

The integral can be estimated by Young's inequality and estimate (6.1)

$$
\begin{aligned}
&- \int_\Omega (2 \langle \mathbf{b}, \nabla \widehat{\varepsilon}_\ell \rangle + (\operatorname{div} \mathbf{b}) \widehat{\varepsilon}_\ell) e_{\ell+1} \\
&\leq \delta \|e_{\ell+1}\|_{L^2(\Omega)}^2 + \frac{(3 b_\infty)^2}{2\delta} \|\widehat{\varepsilon}_\ell\|_{H^1(\Omega)}^2
\end{aligned}
$$

$$\overset{(2.3)}{\leq} \delta \|e_{\ell+1}\|^2_{L^2(\Omega)} + \frac{9b_\infty^2 c_{\mathrm{F}}}{2\delta\alpha} \|\widehat{\varepsilon}_\ell\|^2_{\mathrm{PDE}}$$
$$\leq \delta C^2_{\mathrm{dual}} h^{2s}_{\ell+1} \|e_{\ell+1}\|^2_{\mathrm{PDE}} + \frac{9b_\infty^2 c_{\mathrm{F}}}{2\delta\alpha} \|\widehat{\varepsilon}_\ell\|^2_{\mathrm{PDE}}.$$

Inserting this into (6.3) leads to

$$\left(1 - \delta C^2_{\mathrm{dual}} h^{2s}_{\ell+1}\right) \|e_{\ell+1}\|^2_{\mathrm{PDE}} \leq \|e_\ell\|^2_{\mathrm{PDE}}$$
$$- \left(1 - \frac{9b_\infty^2 c_{\mathrm{F}}}{2\delta\alpha}\right) \|\widehat{\varepsilon}_\ell\|^2_{\mathrm{PDE}}.$$

We choose δ such that both parenthesis have the same value and obtain

$$\|e_{\ell+1}\|^2_{\mathrm{PDE}} \leq \frac{\|e_\ell\|^2_{\mathrm{PDE}}}{1 - C_\star h^s_{\ell+1}} - \|\widehat{\varepsilon}_\ell\|^2_{\mathrm{PDE}} \quad \text{with}$$
$$C_\star := 3C_{\mathrm{dual}} b_\infty \sqrt{\frac{c_{\mathrm{F}}}{2\alpha}}.$$

Let the mesh width $h_{\ell+1}$ of $\mathcal{T}_{\ell+1}$ satisfy $h^s_{\ell+1} < C_\star^{-1}$. Then the assertion holds with $\Lambda^2_{\ell+1}$ as in (6.2). □

The proof of the saturation estimate requires conditions on the *data oscillations*. First, we will introduce some edge bubble for triangles with a common edge. For $E \in \mathcal{E}_\Omega$, let $K_1, K_2 \in \mathcal{T}_\ell$ denote the triangles which share E as the common edge.

Case a Both, K_1, K_2 will be p-refined.

In this case, let $\varphi^{(2)}_{K_1,K_2} \in S_{\ell+1}$ be the *quadratic edge bubble*, i.e., the product of the barycentric coordinates in K_1, K_2 for the endpoints of E.

Case b Both K_1, K_2 will be h-refined.

Let $x_E \in E$ denote the interior vertex on the edge E (cf. Fig. 1) and let $K \in \{K_1, K_2\}$ be an adjacent triangle with inner vertex x_K. Let $K' := \mathrm{conv}\{x_E, x_K, A\}$ and $K'' := \mathrm{conv}\{x_E, B, x_K\}$ with A, B denoting the endpoints of E. Then, the piecewise affine edge bubble $\varphi^{(1)}_{K_1,K_2}$, restricted to K, has value 1 at x_E and vanishes at all other vertices of triangles in $\mathrm{sons}(K_1)$. Assumption 2 ensures that $\varphi^{(1)}_{K_1,K_2} \in S_{\ell+1}$.

Case c K_1 will be p-refined and K_2 will be h-refined. Let $p_m := p_{K_m}$, $m = 1, 2$ and define $p_E := \min\{p_1, p_2\}$. Let $K' \in \mathrm{sons}(K_2)$ be the triangle which contains E as an edge. Then, $\varphi^{(2)}_{K_1,K'}\big|_{K_1}$ (resp. $\varphi^{(2)}_{K_1,K'}\big|_{K'}$) is the product of the barycentric coordinates in K_1 (resp. K') for the endpoints of E and zero outside $\overline{K_1 \cup K'}$.

Case d K_1 will be h-refined and K_2 will be p-refined. Then $\varphi^{(2)}_{K',K_2}$ is defined as in Case c by interchanging the roles of K_1 and K_2.

We define

$$\varphi_E := \begin{cases} \varphi^{(2)}_{K_1,K_2} & \text{in Case a,} \\ \varphi^{(1)}_{K_1,K_2} & \text{in Case b,} \\ \varphi^{(2)}_{K_1,K'} & \text{in Case c,} \\ \varphi^{(2)}_{K',K_2} & \text{in Case d.} \end{cases} \tag{6.4}$$

For $g \in L^2(\Omega)$, we define averages $g_z \in \mathbb{P}_{p_z-1}(\omega_z)$ [with p_z as in (4.5)] as the $L^2(\omega_z)$-orthogonal projection onto $\mathbb{P}_{p_z-1}(\omega_z)$.

The data oscillations are defined by

$$\mathrm{osc}(v) := \sqrt{\sum_{z \in \mathcal{N}_\ell^1} \mathrm{osc}_z^2(v)} \quad \text{with}$$
$$\mathrm{osc}_z(v) := \left\| \frac{h_z}{p_z} \Phi^{1/2}_{\mathrm{osc},z} (\mathrm{res}(v) - \mathrm{res}_z(v)) \right\|_{L^2(\omega_z)} \tag{6.5}$$

with

$$\Phi_{\mathrm{osc},z} := \frac{p_z^5}{c_\pi^2} \Phi_z + p_z^3 \Phi_{\mathcal{E},z} + 1 \quad \text{and} \quad \Phi_{\mathcal{E},z} := \sum_{E \in \mathcal{E}_z} \varphi_E$$

and $\mathrm{res}_z(v)$ is a shorthand for $(\mathrm{res}(v))_z$.

Theorem 8 *Let Assumptions 1, 2, and 6 be satisfied. We assume that the adjoint problem is $H^{1+s}(\Omega)$ regular for some $0 < s \leq 1$. Further we assume that the maximal mesh width of $\mathcal{T}_{\ell+1}$ satisfies $h_{\ell+1} < C_\star^{-s}$ with $C_\star$ as in Lemma 1. Let c_π be as in (5.13b) and C_{rel} as in (4).*

There exists a constant $C_2 > 0$ depending on $\alpha, \beta, b_\infty, c_\infty$, and ρ but independent of $\mathfrak{h}_\ell$, $\mathbf{p}_\ell$, u, and f such that for any $0 \leq \mu \leq 1$ and any $C_3 > C_2 C_{\mathrm{rel}}$ the condition

$$\mathrm{osc}(u_\ell) \leq \frac{\mu}{C_3} \|e_\ell\|_{\mathrm{PDE}} \tag{6.6}$$

implies the error reduction

$$\|e_{\ell+1}\|_{\mathrm{PDE}} \leq \kappa_\ell \|e_\ell\|_{\mathrm{PDE}} \quad \text{with}$$
$$\kappa_\ell := \sqrt{\left(\Lambda^2_{\ell+1} - \left(\frac{c_\pi}{C_3 p_\ell^{5/2}}\right)^2 (1 - \mu^2)\right)}.$$

Remark 4 The condition on C_3 implies that $\kappa_\ell > 0$. From the definition of $\Lambda^2_{\ell+1} = \frac{1}{1 - C_\star h^s_{\ell+1}}$ with $h_{\ell+1} < (2C_\star)^{-1/s}$ as in (6.2) it follows that the condition

$$h_{\ell+1} \leq H(p_\ell) := C_4 \left(\frac{c_\pi}{p_\ell^{5/2}}\right)^{2/s} \quad \text{with}$$
$$C_4 := \left(\frac{1}{C_\star (2C_3 C_{\mathrm{rel}})^2}\right)^{1/s}$$

implies

$$\Lambda^2_{\ell+1} - \left(\frac{c_\pi}{C_3 C_{\mathrm{rel}} p_\ell^{5/2}}\right)^2 \leq 1 - \frac{1}{2}\left(\frac{c_\pi}{C_3 C_{\mathrm{rel}} p_\ell^{5/2}}\right)^2$$

and, for $0 \le \mu < 1/\sqrt{2}$, it holds

$$\kappa_\ell \le \sqrt{1 - \left(\frac{c_\pi}{C_3 C_{\mathrm{rel}} p_\ell^{5/2}}\right)^2 \left(\frac{1}{2} - \mu^2\right)} < 1.$$

Proof of Theorem 8 Since $u_{\ell+1} - u_\ell \in S_{\ell+1}$, the quasi-orthogonality (cf. Lemma 1) implies

$$\Lambda_{\ell+1}^2 \|e_\ell\|_{\mathrm{PDE}}^2 \ge \|e_{\ell+1}\|_{\mathrm{PDE}}^2 + \|u_{\ell+1} - u_\ell\|_{\mathrm{PDE}}^2 \tag{6.7}$$

with $\Lambda_{\ell+1}^2$ as in (6.2). Hence it is sufficient to prove a lower bound for $\|u_{\ell+1} - u_\ell\|_{\mathrm{PDE}}^2$ in terms of $\|e_\ell\|_{\mathrm{PDE}}^2$ and data oscillations. The residual a posteriori error estimate can be recast in the form of stars [cf. (4.4)]: By a triangle inequality we obtain

$$\begin{aligned}\|e_\ell\|_{\mathrm{PDE}}^2 \le 2C_{\mathrm{rel}}^2 \sum_{z \in \mathcal{N}_\ell^1} \Bigg(& \left\| \frac{h_z}{p_z} \operatorname{res}_z(u_\ell) \right\|_{L^2(\omega_z)}^2 \\ & + \sum_{E \in \mathcal{E}_z} \left\| \sqrt{\frac{h_z}{p_z}} \operatorname{Res}(u_\ell) \right\|_{L^2(E)}^2 \\ & + \left\| \frac{h_z}{p_z} \left(\operatorname{res}_z(u_\ell) - \operatorname{res}(u_\ell)\right) \right\|_{L^2(\omega_z)}^2 \Bigg).\end{aligned} \tag{6.8}$$

Hence, it is sufficient to bound the jumps and projected volume residuals from above by $\|u_{\ell+1} - u_\ell\|_{\mathrm{PDE}}$ and to control the last term by the oscillation condition (6.6).

We start with the jump term and employ the same arguments as in [25, Proof of Lemma 3.1, Step 2]. Since u_ℓ is continuous, $[\nabla u_\ell]_E$ is parallel to $\mathbf{n}_E$, i.e., $[\nabla u_\ell]_E = j_E \mathbf{n}_E$ and $j_E = \left[\frac{\partial u_\ell}{\partial \mathbf{n}_E}\right]_E \in \mathbb{P}_{p_E - 1}$ with $p_E := \min\{p_1, p_2\}$ (cf. Remark 3). The continuity of the coefficient matrix $\mathbf{A}$ implies

$$J_E := \langle \mathbf{A}\mathbf{n}_E, [\nabla u_\ell]_E \rangle = \langle \mathbf{A}\mathbf{n}_E, \mathbf{n}_E \rangle \, j_E =: a_E j_E, \tag{6.9}$$

where $\alpha \le a_E(x) \le \beta$ (cf. Assumption 1). Consequently

$$\begin{aligned}\left\| \sqrt{\varphi_E} J_E \right\|_{L^2(E)}^2 &= \int_E \alpha_E (j_E \varphi_E) J_E \le \beta \int_E (j_E \varphi_E) J_E \\ \left\| \sqrt{\varphi_E} J_E \right\|_{L^2(E)}^2 &= \int_E \alpha_E^2 j_E^2 \varphi_E \ge \alpha^2 \int_E j_E^2 \varphi_E \\ &\overset{[26,\,\text{Lem. 2.4}]}{\ge} c \frac{\alpha^2}{p_E^2} \|j_E\|_{L^2(E)}^2 \\ &\ge c \left(\frac{\alpha}{\beta p_E}\right)^2 \|J_E\|_{L^2(E)}^2.\end{aligned} \tag{6.10}$$

Thus

$$\left\| \sqrt{h_E \varphi_E} J_E \right\|_{L^2(E)}^2 \le \beta h_E \int_E \langle \mathbf{A}\mathbf{n}, [\nabla u_\ell]_E \rangle (j_E \varphi_E).$$

Next we extend j_E to ω_E. For $K \subset \omega_E$, let Λ_K, Λ_E be chosen such that [cf. (5.3)] $\Lambda_K|_{x_2=0} = \Lambda_E$ holds. Let $\hat{Z}(x) := (x_1, 0)^\intercal$ for $x = (x_1, x_2) \in \mathbb{R}^2$. We define $j_E^\star : \omega_E \to \mathbb{R}$ trianglewise by

$$j_E^\star|_K := j_E \circ \Lambda_K \circ \hat{Z} \circ \Lambda_K^{-1}. \tag{6.11}$$

Note that $j_E^\star$ is a polynomial of degree $p_E - 1$ on both triangles which share E as the common edge. The construction of φ_E along the definition of the polynomial degrees on the refined mesh (cf. Assumption 2) imply $\varphi_E j_E^\star \in S_{\ell+1}$. By using partial integration and the fact that $u_{\ell+1}$ is the Galerkin solution we get for any $E \subset \omega_z$

$$\frac{1}{\beta} \left\| \sqrt{h_E \varphi_E} J_E \right\|_{L^2(E)}^2 \le h_E \int_E \langle \mathbf{A}\mathbf{n}, [\nabla u_\ell]_E \rangle (j_E \varphi_E) \tag{6.12a}$$

$$= h_E \int_{\omega_E} \left\langle \mathbf{A} \nabla u_\ell, \nabla\left(j_E^\star \varphi_E\right) \right\rangle + \operatorname{div}_\ell(\mathbf{A}\nabla_\ell u_\ell)\left(j_E^\star \varphi_E\right) \tag{6.12b}$$

$$\begin{aligned}= h_E \Bigg\{ & a\left(u_\ell - u_{\ell+1}, j_E^\star \varphi_E\right) + \int_{\omega_E} \operatorname{res}_z(u_\ell)\left(j_E^\star \varphi_E\right) \\ & + \left(\operatorname{res}(u_\ell) - \operatorname{res}_z(u_\ell)\right)\left(j_E^\star \varphi_E\right) \Bigg\}.\end{aligned} \tag{6.12c}$$

From Lemma 5 and Corollary 6 we conclude by an affine pullback to the reference element that

$$\begin{aligned}\left\| \sqrt{\varphi_E} j_E^\star \right\|_{L^2(\omega_E)} &\le C_5 \left\| \sqrt{h_E \varphi_E} j_E \right\|_{L^2(E)} \\ &\overset{(6.9)}{\le} \frac{C_5}{\alpha} \left\| \sqrt{h_E \varphi_E} J_E \right\|_{L^2(E)},\end{aligned} \tag{6.13a}$$

$$\left\| \varphi_E j_E^\star \right\|_{\mathrm{PDE}} \le C_6 p_E \left\| \sqrt{\frac{\varphi_E}{h_E}} J_E \right\|_{L^2(E)}, \tag{6.13b}$$

where C_5, C_6 only depend on $\alpha, \beta, b_\infty, c_\infty$ and the shape regularity of the mesh. The combination of (6.12) with (6.13) and (6.10) leads to

$$\begin{aligned}\left\| \sqrt{\frac{h_E}{p_E}} \operatorname{Res}(u_\ell) \right\|_{L^2(E)} \le C_7 p_E^{3/2} \Bigg\{ & \|u_\ell - u_{\ell+1}\|_{\mathrm{PDE},\omega_E} \\ & + \left\| \frac{h_E}{p_E} \varphi_E^{1/2} \operatorname{res}_z(u_\ell) \right\|_{L^2(\omega_E)} \\ & + \left\| \frac{h_E}{p_E} \varphi_E^{1/2} \left(\operatorname{res}(u_\ell) - \operatorname{res}_z(u_\ell)\right) \right\|_{L^2(\omega_E)} \Bigg\}\end{aligned} \tag{6.14}$$

with C_7 depending only on $\alpha, \beta, b_\infty, c_\infty$, and ρ. A summation of the squared inequality (6.14) over all $E \in \mathcal{E}_z$ yields

$$\begin{aligned}\sum_{E \in \mathcal{E}_z} \left\| \sqrt{\frac{h_E}{p_E}} \operatorname{Res}(u_\ell) \right\|_{L^2(E)}^2 \le C_8 p_z^3 & \\ \times \Bigg\{ \|u_\ell - u_{\ell+1}\|_{\mathrm{PDE},\omega_z}^2 + \left\| \frac{h_z}{p_z} \Phi_{\mathcal{E},z}^{1/2} \operatorname{res}_z(u_\ell) \right\|_{L^2(\omega_z)}^2 & \\ + \left\| \frac{h_z}{p_z} \Phi_{\mathcal{E},z}^{1/2} \left(\operatorname{res}(u_\ell) - \operatorname{res}_z(u_\ell)\right) \right\|_{L^2(\omega_z)}^2 \Bigg\}. &\end{aligned} \tag{6.15}$$

Hence, we are left with the estimate of the volume residual.

Partial integration and the fact that u_ℓ solves the Galerkin equations leads to

$$\begin{aligned}\sum_{E\in\mathcal{E}_z}\int_E J_E b_z^1 &= \int_{\omega_z}\left\langle \mathbf{A}\nabla u_\ell, \nabla b_z^1\right\rangle + \operatorname{div}_\ell(\mathbf{A}\nabla_\ell u_\ell)\, b_z^1\\ &= \int_{\omega_z}\operatorname{res}(u_\ell)\, b_z^1 \qquad (6.16)\\ &= \int_{\omega_z}\operatorname{res}_z(u_\ell)\, b_z^1 + \int_{\omega_z}(\operatorname{res}(u_\ell)-\operatorname{res}_z(u_\ell))\, b_z^1.\end{aligned}$$

We choose $\varphi_{\ell+1,z}$ as in Theorem 7 such that (5.13) holds and obtain as in (6.12c)

$$\begin{aligned}\sum_{E\in\mathcal{E}_z}\int_E J_E\varphi_{\ell+1,z} &= \sum_{E\in\mathcal{E}_z}\int_E \langle \mathbf{An}, [\nabla u_\ell]_E\rangle\,\varphi_{\ell+1,z}\\ &= \sum_{K\subset\omega_z}\int_K \left\langle \mathbf{A}\nabla u_\ell, \nabla\varphi_{\ell+1,z}\right\rangle + \operatorname{div}_\ell(\mathbf{A}\nabla u_\ell)\,\varphi_{\ell+1,z}\\ &= a\left(u_\ell-u_{\ell+1},\varphi_{\ell+1,z}\right) + \int_{\omega_z}\operatorname{res}_z(u_\ell)\,\varphi_{\ell+1,z}\\ &\quad + \int_{\omega_z}(\operatorname{res}(u_\ell)-\operatorname{res}_z(u_\ell))\,\varphi_{\ell+1,z}. \qquad (6.17)\end{aligned}$$

The combination of (6.16) and (6.17) with (5.13a) allows to eliminate the jump residuals and we obtain

$$\begin{aligned}\int_{\omega_z}\operatorname{res}_z(u_\ell)\left(b_z^1-\varphi_{\ell+1,z}\right) &= a\left(u_\ell-u_{\ell+1},\varphi_{\ell+1,z}-b_z^1\right)\\ &\quad + \int_{\omega_z}(\operatorname{res}(u_\ell)-\operatorname{res}_z(u_\ell))\left(\varphi_{\ell+1,z}-b_z^1\right). \qquad (6.18)\end{aligned}$$

Recall the definition of $\varphi_{\ell+1,z}$ as in the proof of Theorem 7

$$b_z^1-\varphi_{\ell+1,z}=v_z \quad\text{with}\quad v_z:=h_z\frac{\Pi_z v}{\|v\|_z}$$

and we apply this definition for $v=\operatorname{res}_z(u_\ell)$. From this and (5.13b) we obtain a bound of the averaged volume residual on stars

$$\begin{aligned}c_\pi\left\|\Phi_z^{1/2}h_z\operatorname{res}_z(u_\ell)\right\|_{L^2(\omega_z)} &\le C_S\|u_\ell-u_{\ell+1}\|_{\mathrm{PDE},\omega_z}\left\|\varphi_{\ell+1,z}-b_z^1\right\|_{\mathrm{PDE},\omega_z}\\ &\quad + \left\|h_z\Phi_z^{1/2}(\operatorname{res}(u_\ell)-\operatorname{res}_z(u_\ell))\right\|_{L^2(\omega_z)}h_z^{-1}\\ &\quad\times\left(\left\|\Phi_z^{-1/2}\left(\varphi_{\ell+1,z}-b_z^1\right)\right\|_{L^2(\omega_z)}\right). \qquad (6.19)\end{aligned}$$

Hence, from (5.13) we conclude

$$\begin{aligned}c_\pi\left\|\Phi_z^{1/2}h_z\operatorname{res}_z(u_\ell)\right\|_{L^2(\omega_z)} &\le C_S\frac{C_1}{c_2}p_z\|u_\ell-u_{\ell+1}\|_{\mathrm{PDE},\omega_z}\\ &\quad + 2C_1\left\|h_z\Phi_z^{1/2}(\operatorname{res}(u_\ell)-\operatorname{res}_z(u_\ell))\right\|_{L^2(\omega_z)}. \qquad (6.20a)\end{aligned}$$

Note that[6]

$$\begin{aligned}\left\|\Phi_z^{1/2}h_z\operatorname{res}_z(u_\ell)\right\|_{L^2(\omega_z)} &\ge \frac{c}{p_z^{1-s}}\\ &\quad\times\left\|\frac{h_z}{p_z}\Phi_{\mathcal{E},z}^{s/2}\operatorname{res}_z(u_\ell)\right\|_{L^2(\omega_z)} \quad\text{for } s\in\{0,1\}. \qquad (6.20b)\end{aligned}$$

The first and second term in the right-hand side in (6.8) can be estimated by means of (6.14) and (6.20)

$$\begin{aligned}\|e_\ell\|_{\mathrm{PDE}}^2 &\le 2C_{\mathrm{rel}}^2C\sum_{z\in\mathcal{N}_\ell^1}\left(\frac{p_z^5}{c_\pi^2}\|u_\ell-u_{\ell+1}\|_{\mathrm{PDE},\omega_z}^2\right.\\ &\quad\left.+\left\|\frac{h_z}{p_z}\Phi_{\mathrm{osc},z}^{1/2}(\operatorname{res}(u_\ell)-\operatorname{res}_z(u_\ell))\right\|_{L^2(\omega_z)}^2\right), \qquad (6.21)\end{aligned}$$

where, again, C_9 only depends on α, β, b_∞, c_∞, and ρ.

Taking into account the finite overlap of the supports ω_z we end up with

$$\|e_\ell\|_{\mathrm{PDE}}^2 \le (C_2C_{\mathrm{rel}})^2\left(\frac{p_\ell^5}{c_\pi^2}\|u_\ell-u_{\ell+1}\|_{\mathrm{PDE}}^2+\operatorname{osc}^2(u_\ell)\right) \qquad (6.22)$$

where C_2 only depends on α, β, b_∞, c_∞, and ρ.

Choose $C_3 > C_2C_{\mathrm{rel}}$. The assumption that the data oscillations are small, i.e., $\operatorname{osc}(u_\ell)\le\frac{\mu}{C_3}$, implies for any $0\le\mu<1$

$$\sum_{z\in\mathcal{N}_\ell^1}\|u_\ell-u_{\ell+1}\|_{\mathrm{PDE},\omega_z}^2 \ge \frac{c_\pi^2}{C_3^2p_\ell^5}\left(1-\mu^2\right)\|e_\ell\|_{\mathrm{PDE}}^2.$$

The combination with (6.7) finally leads to

$$\begin{aligned}\|e_{\ell+1}\|_{\mathrm{PDE}}^2 &\le \Lambda_{\ell+1}^2\|e_\ell\|_{\mathrm{PDE}}^2-\|u_{\ell+1}-u_\ell\|_{\mathrm{PDE}}^2\\ &\le\left(\Lambda_{\ell+1}^2-\left(\frac{c_\pi}{C_3p_\ell^{5/2}}\right)^2\left(1-\mu^2\right)\right)\|e_\ell\|_{\mathrm{PDE}}^2\end{aligned}$$

and this is the assertion. □

[6] For $s=1$ this follows from Corollary 6. For $s=0$, we conclude from [39, Prop. 3.37, Cor. 3.40, Prop. 3.46] that

$$\left\|\sqrt{\Phi_K^{(3)}}v\right\|_{L^2(K)} \ge \frac{c}{p^2}\|v\|_{L^2(K)} \qquad \forall v\in\mathbb{P}_p(K),\quad p\ge1$$

holds and from [27, (22) with $\alpha=0$ and $\beta=1$]

$$\left\|\sqrt{\Phi_K^{(1)}}v\right\|_{L^2(K)} \ge \frac{c}{p}\|v\|_{L^2(K)} \qquad \forall v\in\mathbb{P}_p(K),\quad p\ge1$$

for some constant $c>0$ which is independent of p and h_K. Finally, for $\Phi_{K,K}^{(1)}+\Phi_K^{(2)}$ we employ

$$\Phi_K^{(3)}\le\Phi_{K,K}^{(1)}+\Phi_K^{(2)} \qquad \text{pointwise}$$

to obtain

$$\left\|\sqrt{\Phi_{K,K}^{(1)}+\Phi_K^{(2)}}v\right\|_{L^2(K)} \ge \left\|\sqrt{\Phi_K^{(3)}}v\right\|_{L^2(K)} \ge \frac{c}{p^2}\|v\|_{L^2(K)} \qquad \forall v\in\mathbb{P}_p(K),\quad p\ge1.$$

Corollary 3 *Let the assumptions of Theorem 8 be satisfied. Condition (6.6) follows from the computable condition*

$$\operatorname{osc}(u_\ell) \le \frac{\hat{\mu} c_\pi}{C_{10} p_\ell^{5/2}} \left(\sum_{E \in \mathcal{E}_\Omega} \left\| \sqrt{\frac{h_E}{p_E}} \operatorname{Res}(u_\ell) \right\|_{L^2(E)}^2 \right)^{1/2} \tag{6.23}$$

for sufficiently small $0 < \hat{\mu} \le \hat{\mu}_0$, where $\hat{\mu}_0$ depends on $\alpha, \beta, b_\infty, c_\infty, \mu, C_3$, and ρ.

Proof Observe that (6.12c) and (6.17) remain true if $u_{\ell+1}$ is replaced by u. Hence, we may also replace $u_{\ell+1}$ by u in (6.14) and (6.20). By doing so, the combination of (6.14) and (6.20) yields after summing the squared norms over all $z \in \mathcal{N}_\ell^1$ the estimate

$$\left(\sum_{E \in \mathcal{E}_\Omega} \left\| \sqrt{\frac{h_E}{p_E}} \operatorname{Res}(u_\ell) \right\|_{L^2(E)}^2 \right)^{1/2} \le C_{10} \left(\frac{p_\ell^{5/2}}{c_\pi} \|e_\ell\|_{\mathrm{PDE}} + \operatorname{osc}(u_\ell) \right),$$

where C_{10} only depends $\alpha, \beta, b_\infty, c_\infty$, and ρ. The condition (6.23) implies (since $0 \le c_\pi \le 1$ [cf. Cor. (2)]

$$\begin{aligned} \operatorname{osc}(u_\ell) &\le \hat{\mu} \frac{1}{C_{10}} \frac{c_\pi}{p_\ell^{5/2}} \left(\sum_{E \in \mathcal{E}_\Omega} \left\| \sqrt{\frac{h_E}{p_E}} \operatorname{Res}(u_\ell) \right\|_{L^2(E)}^2 \right)^{1/2} \\ &\le \widehat{\mu} \|e_\ell\|_{\mathrm{PDE}} + \widehat{\mu} \operatorname{osc}(u_\ell). \end{aligned}$$

For sufficiently small $0 < \hat{\mu} \le \hat{\mu}_0$, this implies (6.6). □

Corollary 4 *Assume that the sequence of meshes and polynomial distributions are chosen such that the oscillation condition (6.6) holds on every level ℓ. Let Conjecture 1 be satisfied. Then, the contraction of the error on level ℓ is given by*

$$\zeta_\ell := \prod_{k=1}^{\ell} \left(1 - \frac{C}{p_k^{5/2}} \right)$$

for a constant $0 < C \le 1$ which is independent of the polynomial degrees, i.e.,

$$\|e_\ell\|_{\mathrm{PDE}} \le \zeta_\ell \|e_0\|_{\mathrm{PDE}}.$$

Recall that p_k denotes the maximal polynomial degree at level k which is monotonously increasing. Define the sequence $(n_i)_{i \in \mathbb{N}}$ recursively by $n_0 = 0$ and, for $i = 1, 2, \ldots$, by the condition

$$p_k = i \quad \text{for } n_{i-1} + 1 \le k \le n_i,$$

i.e., the maximal polynomial degree stays fix for $\delta_i := n_i - n_{i-1}$ consecutive levels.

1. *If, for some $k_0 > 0$, it holds $n_k = \infty$ for all $k \ge k_0$, then $\zeta_\ell \le \left(1 - C p_{k_0}^{-5/2}\right)^\ell \to 0$ as $\ell \to \infty$.*
2. *If, $\sum_{k=1}^{\infty} p_k^{-5/2} = +\infty$, then $\lim_{\ell \to \infty} \zeta_\ell = 0$.*
3. *If $\delta_i \ge c i^{3/2}$, then $\lim_{\ell \to \infty} \zeta_\ell = 0$ as $\ell \to \infty$, while $\lim_{k \to \infty} p_k = \infty$.*

Proof The first statement is trivial. For the second statement we employ for $s > 0$ and $C < 1$

$$\zeta_\ell \le \exp\left(\sum_{k=1}^{\ell} \log\left(1 - \frac{C}{p_k^s} \right) \right).$$

Note that, for $0 < \varepsilon < 1$,

$$\log(1 - \varepsilon) \le -\varepsilon$$

so that

$$\zeta_\ell \le \exp\left(-C \sum_{k=1}^{\ell} \frac{1}{p_k^s} \right).$$

From this, the second statement follows. For the third one we use

$$\lim_{\ell \to \infty} \zeta_\ell \le \exp\left(-C \sum_{i=1}^{\infty} \frac{\delta_i}{i^s} \right).$$

Hence, for $\delta_i \ge c i^{s-1}$ we have $\lim_{\ell \to \infty} \zeta_\ell = 0$. □

7 Appendix 1: Lower bound for the constant c_π: numerical experiments

In this appendix we will invest the dependence of the stability constant c_π of the polynomial projection operator Π_z [cf. (5.8)] on the polynomial degree p. We consider mainly two cases: pure p-refinement and h-refinement.

7.1 p-Refinement

First, we will rewrite the definition of c_π as an algebraic eigenvalue problem which we will solve numerically. We have performed numerical experiments for the two-dimensional setting on stars as described in this paper but also considered the one-dimensional case where ω_z consists of the two intervals which have z as a common endpoint.

7.1.1 Equivalent formulation

The goal is to investigate the dependence of the constant

$$c_\pi := \inf_{v \in \mathbb{P}_p(\omega_z) \setminus \{0\}} \frac{\left(v, \Pi_z^p v\right)_{L^2(\omega_z)}}{\left\| \Phi_z^{1/2} v \right\|_{L^2(\omega_z)}^2} \tag{7.1}$$

on the polynomial degree p numerically. Let d denote the spatial dimension. Let ω_z consists of $q \ge d$ simplices K_i, $1 \le i \le q$.

By employing a global affine map we can pull back the star ω_z to a reference configuration, where $K_1 = \widehat{K}$ is the unit simplex, on the expense that c_π in (7.1) depends additionally on the shape regularity of K_1. Let $\chi_i : \widehat{K} \to K_i$ denote affine bijection with the special choice $\chi_1 = \mathrm{id}$. Then,

$$\left\| \Phi_z^{1/2} v \right\|^2_{L^2(\omega_z)} = \sum_{i=1}^{q} \frac{|K_i|}{|\widehat{K}|} \int_{\widehat{K}} \widehat{\Phi_K} \widehat{v_i}^2,$$

where $\widehat{v_i} = v \circ \chi_i$ and $\Phi_{\widehat{K}}$ denotes the product of barycentric coordinates. Let $(P_n)_{n \in \iota_p}$ denote a basis of $\mathbb{P}_p(\omega_z)$ for a suitable index set ι_p. We write

$$v = \sum_{n=0}^{p} v_n P_n \tag{7.2}$$

and obtain

$$\left\| \Phi_z^{1/2} v \right\|^2_{L^2(\omega_z)} = \mathbf{v}^\mathsf{T} \mathbf{M}^{(p)} \mathbf{v},$$

where

$$\left(\mathbf{M}_i^{(p)}\right)_{n,m} := \frac{|K_i|}{|\widehat{K}|} \int_{\widehat{K}} \widehat{\Phi_K} \left(P_n \circ \chi_i\right) \left(P_m \circ \chi_i\right), \quad n, m \in \iota_p$$

$$\text{and} \quad \mathbf{M}^{(p)} := \sum_{i=1}^{q} \mathbf{M}_i^{(p)}.$$

For the special case that $\widehat{\Phi_{\widehat{K}}}$ is the polynomial bubble function we can choose an orthogonal basis for $\mathbb{P}_p\left(\widehat{K}\right)$ (cf. [23,31]) so that $\mathbf{M}_1^{(p)}$ is a diagonal matrix.

In order to invest (7.1) we introduce a matrix representation of $\Pi_z^p v$ with v as in (7.2) via the ansatz

$$\Pi_z^p v|_{K_i} = \sum_{m \in \iota_p} w_{m,i} P_m \circ \chi_i^{-1}$$

The coefficients $\mathbf{w}_i = (w_{m,i})_{m \in \iota_p}$ are determined via

$$\mathbf{w}_i = \left(\mathbf{M}_1^{(p-1)}\right)^{-1} \mathbf{W}_i \mathbf{v} \quad \text{with}$$

$$(\mathbf{W}_i)_{m,n} := \frac{|K_i|}{|\widehat{K}|} \int_{\widehat{K}} \widehat{\Phi_K} P_n \circ \chi_i P_m \quad \text{for} \quad m \in \iota_{p-1}, n \in \iota_p.$$

Hence,

$$\left(v, \Pi_z^p v\right)_z = \mathbf{v}^\mathsf{T} \mathbf{B} \mathbf{v} \quad \text{with} \quad \mathbf{B} := \sum_{i=1}^{q} \mathbf{W}_i^\mathsf{T} \left(\mathbf{M}_{1,1}^{(p-1)}\right)^{-1} \mathbf{W}_i$$

so that the constant c_π has the algebraic representation

$$c_\pi = \inf_{\mathbf{v} \in \mathbb{R}^{\iota_p}} \frac{\mathbf{v}^\mathsf{T} \mathbf{B} \mathbf{v}}{\mathbf{v}^\mathsf{T} \mathbf{M}^{(p)} \mathbf{v}}.$$

Hence, c_π is the smallest eigenvalue of

$$\left(\mathbf{M}^{(p)}\right)^{-1/2} \mathbf{B} \left(\mathbf{M}^{(p)}\right)^{-1/2}.$$

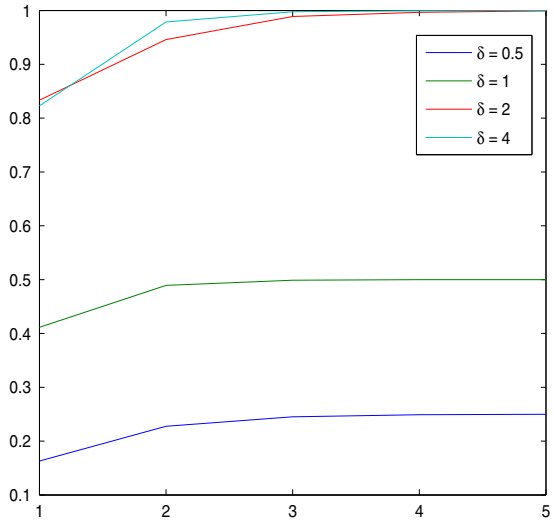

Fig. 4 Performance of c_π versus p for the one-dimensional case

7.1.2 The one-dimensional case

In this case we have $\widehat{K} = [-1, 1]$ and P_n are the Jacobi polynomials $P_n^{(1,1)}$ which are defined as follows

$$P_n^{(\alpha,\beta)}(x) = \frac{(2)_n}{n!} \, {}_2F_1\left(\begin{matrix} -n, n+\alpha+\beta+1 \\ \alpha+1 \end{matrix}; \frac{1-x}{2}\right),$$

where $(\cdot)_n$ is Pochhammer's symbol and ${}_2F_1$ is the terminating Gauss hypergeometric function

$${}_2F_1\left(\begin{matrix} -n, b \\ c \end{matrix}; z\right) = \sum_{k=0}^{n} \frac{(-n)_k (b)_k}{(c)_k k!} z^k.$$

We consider $K_1 = \widehat{K}$ and $K_2 = [1, 1+\delta]$ for some $\delta > 0$. Note that $\mathbf{M}_1^{(p)}$ in this case is given by

$$\mathbf{M}_1^{(p)} = \mathrm{diag}\left[8 \frac{(n+1)}{(2n+3)(n+2)} : n \in \iota_p\right].$$

The mapping χ_2 is defined by

$$\chi_2\left(\hat{x}\right) = \frac{1-\hat{x}}{2} + \frac{1+\hat{x}}{2}(1+\delta).$$

To observe the behaviour of c_π with respect to p and δ, we consider three different cases: $\delta = 0.5$, $\delta = 1$, $\delta = 2$, $\delta = 4$. The following observations can be obtained from Fig. 4:

- c_π converges to a positive constant with respect to p,
- c_π is properly bounded from below,
- c_π is decreasing as δ goes to zero.

7.1.3 The two-dimensional case

Now we consider Jacobi bivariate polynomials as our basis functions on the reference triangle, which are defined as follows:

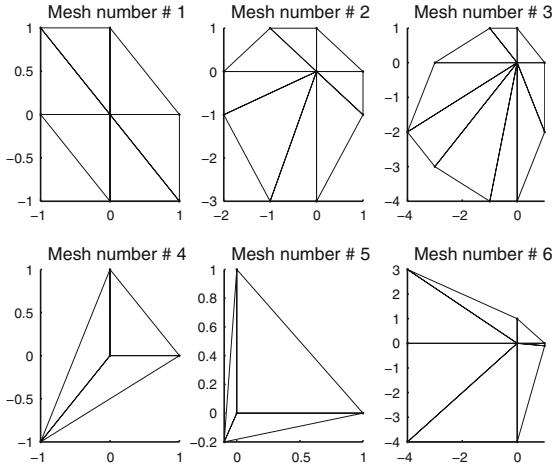

Fig. 5 Illustration of the geometric configuations described in (7.3)

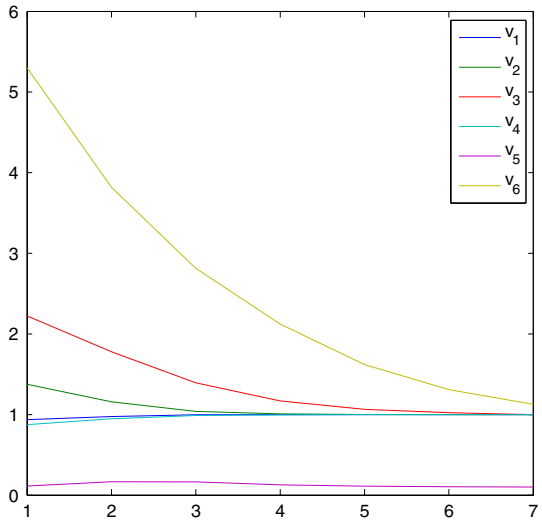

Fig. 6 Performance of c_π versus p for the two-dimensional cases

$$P_{n,k}^{1,1,1}(x,y) := (1-x)^k P_{n-k}^{(1,3+2k)}(1-2x) P_k^{(1,1)}\left(1-\frac{2y}{1-x}\right),$$

which is a polynomial of degree n in x and y.

We study different triangulations. Again we assume that K_1 is the unit simplex and the common point of all triangles in the patch is $(0,0)$. The meshes consist of the following nodes and are illustrated in Fig. 5:

$$\begin{aligned}
v_1 &= \{(0,0),(1,0),(0,1),(-1,1),(-1,0),(0,-1),(1,-1)\},\\
v_2 &= \{(0,0),(1,0),(0,1),(-1,1),(-2,0),(-2,-1),(-1,-3),\\
&\quad (0,-3),(1,-1)\},\\
v_3 &= \{(0,0),(1,0),(0,1),(-1,1),(-3,0),(-4,-2),(-3,-3),\\
&\quad (-1,-4),(0,-4),(1,-2)\},\\
v_4 &= \{(0,0),(1,0),(0,1),(-1,-1)\},\\
v_5 &= \{(0,0),(1,0),(0,1),(-0.1,-0.2)\},\\
v_6 &= \{(0,0),(1,0),(0,1),(-4,3),(-4,0),(-4,-4),(0,-4),\\
&\quad (1,-0.1)\}.
\end{aligned} \tag{7.3}$$

Figure 6 shows the behaviour of c_π with respect to p in each case and we summarize the main observations.

(a) In the first three cases, i.e., the number of triangles (at least six) is varying while the shape regularity constant is always moderately bounded, the lower bound of c_π is approximately 1. It also shows that the constant c_π is robust with respect to the elongation of the triangles which is in analogy to the one-dimensional observation (δ increases).
(b) If we consider the minimal number (three) elements, again, with moderate shape regular constant, we still get a proper lower bound. Recall that the dimension of the image space $\mathbb{P}_{p-1}(\mathcal{T}_z)$ [in (5.1)] increases with the number of triangles so that we expect that the constant c_π becomes larger with increasing number of triangles.
(c) On the other hand, if we consider the minimal configuration with only three triangles *and* large shape regularity constant (the area of the triangles is highly varying) as described by v_5, then the constant c_π becomes smaller as expected.
(d) Configuration v_6 supports the statement that, if the space $\mathbb{P}_{p-1}(\mathcal{T}_z)$ is large enough, then a few tiny elements can be still harmless. We can see that these numerical examples confirm our hypothesis that c_π depends on the shape regularity of our meshes but does not depend on p.

7.2 h-Refinement

In this section we study the similar problem as in previous section but with h-refinement instead of p-refinement. In other word, we apply one level of regular h-refinement on each mesh and observe the behaviour of the constant c_π with respect to p on the refined mesh. To be able to make a comparison between the results, we take the same patches as in previous section. From the definition of Φ_z for this case we have

$$c_\pi = \inf_{v\in\mathbb{P}_{p-1}(K)} \frac{\sum_{K\subset\omega_z}\int_K v\Phi_{K,K}^{(1)}v+\int_K v\left(\Phi_{K,K}^{(1)}+\Phi_K^{(2)}\right)\prod_K^{p-1}v}{\sum_{K\subset\omega_z}\int_K v\left(\Phi_{K,K}^{(1)}+\Phi_K^{(2)}\right)v},$$

where $\Phi_{K,K}^{(1)}$ and $\Phi_K^{(2)}$ are piecewise linear and quadratic functions defined as in (5.4). Figure 7 shows the behaviour of c_π for the same patches with respect to p. It supports our hypothesis and shows the similar behaviour as in p-version.

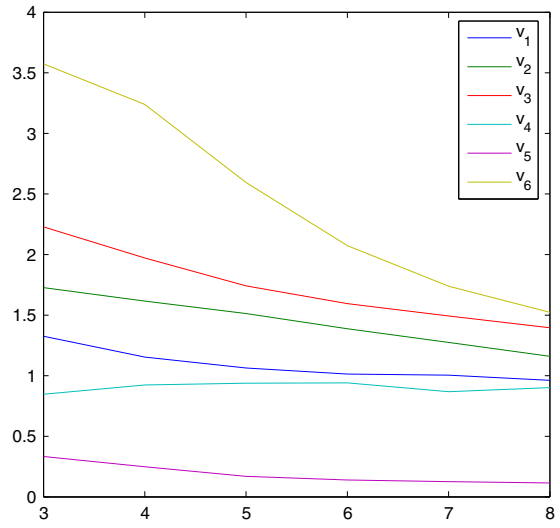

Fig. 7 Behaviour of c_π versus p with one level of h-refinement

Also here we observe that c_π does not depend on p, but it only depends on the shape regularity of the mesh.

8 Appendix 2: Polynomial inverse estimates

We start with a one-dimensional estimate.

Lemma 2 *For $a < b$, let $\Phi_{[a,b]}(x) = \frac{(x-a)(b-x)}{(b-a)^2}$ denote the one-dimensional bubble function. Then,*

$$\left\|\left(\Phi_{[a,b]}v\right)'\right\|_{L^2([a,b])} \leq C\frac{p+1}{b-a}\left\|\Phi_{[a,b]}^{1/2}v\right\|_{L^2([a,b])} \quad \forall v\in\mathbb{P}_p([a,b]).$$

Proof We first prove the result for $(a,b)=(0,1)$. Observe that $\left\|\Phi'_{[0,1]}\right\|_{L^\infty([0,1])}=1$ so that Leibniz rule gives us

$$\begin{aligned}\left\|\left(\Phi_{[0,1]}v\right)'\right\|_{L^2([0,1])} &\leq \left\|\Phi'_{[0,1]}v\right\|_{L^2([0,1])}+\left\|\Phi_{[0,1]}v'\right\|_{L^2([0,1])}\\ &\leq \|v\|_{L^2([0,1])}+\left\|\Phi_{[0,1]}v'\right\|_{L^2([0,1])}.\end{aligned}$$

For the first term, we apply [27, Lemma 2.4 with $\alpha=0$ and $\beta=1$] and for the second term [27, Lemma 2.4 with $\delta=1$] to obtain

$$\left\|\left(\Phi_{[0,1]}v\right)'\right\|_{L^2([0,1])} \leq C(p+1)\left\|\Phi_{[0,1]}^{1/2}v\right\|_{L^2([0,1])}. \quad (8.1)$$

The result then follows via a scaling argument. □

Corollary 5 *Let $a<b$ and $\Phi_{[a,b]}$ be as in Lemma 2. Let $\Psi_{[a,b]}\in W^{1,\infty}([a,b])$ be a function with the properties*

$$\left|\Psi_{[a,b]}\right| \leq C_{11}\Phi_{[a,b]} \quad \textit{pointwise and}$$
$$\left\|\Psi'_{[a,b]}\right\|_{L^\infty([a,b])} \leq \frac{C_{12}}{b-a}.$$

Then

$$\left\|\left(\Psi_{[a,b]}v\right)'\right\|_{L^2([a,b])} \leq C(C_{11}+C_{12})\frac{p+1}{b-a}\left\|\Phi_{[a,b]}^{1/2}v\right\|_{L^2([a,b])} \quad \forall v\in\mathbb{P}_p([a,b]).$$

Proof Leibniz' rule gives us

$$\begin{aligned}\left\|\left(\Psi_{[a,b]}v\right)'\right\|_{L^2([a,b])} &\leq \left\|\Psi'_{[a,b]}v\right\|_{L^2([a,b])}+\left\|\Psi_{[a,b]}v'\right\|_{L^2([a,b])}\\ &\leq \frac{C_{12}}{b-a}\|v\|_{L^2([a,b])}+C_{11}\left\|\Phi_{[a,b]}v'\right\|_{L^2([a,b])}\\ &\leq C(C_{11}+C_{12})\frac{p+1}{b-a}\left\|\Phi_{[a,b]}^{1/2}v\right\|_{L^2([a,b])},\end{aligned}$$

where the last inequality follows as (8.1). □

The two-dimensional version is formulated next. The estimates are similar to those in [39, Sec. 3.6] but differ by powers of the weight functions in the right-hand side and also by the choice of the weight function in Lemma 4. The proofs follow the lines of the proofs in [39, Prop. 3.46] and also employs tools from [28, Appendix D].

Lemma 3 *Let K denote a triangle and let Φ_K be the cubic bubble function as defined in (5.3). Then, it holds for all $v\in\mathbb{P}_p(K)$*

$$\|\nabla(\Phi_K v)\|_{L^2(K)} \leq C\frac{p+1}{h_K}\left\|\Phi_K^{1/2}v\right\|_{L^2(K)}.$$

Proof Let $K=\widehat{K}$ be the two-dimensional reference triangle. Note that

$$\Phi_{\widehat{K}}(x_1,x_2)=\Phi_{[0,1-x_1]}(x_2)(1-x_1)\Phi_{[0,1]}(x_1)$$

with $\Phi_{[a,b]}$ as in Lemma 2. First, we consider the derivative with respect to x_2 and obtain

$$\begin{aligned}&\left\|\partial_2\left(\Phi_{\widehat{K}}v\right)\right\|^2_{L^2(\widehat{K})}\\ &=\int_0^1\left(\int_0^{1-x_1}\left(\partial_2\left(\Phi_{\widehat{K}}(x_1,x_2)\,v(x_1,x_2)\right)\right)^2dx_2\right)dx_1\\ &=\int_0^1\Phi^2_{[0,1]}(x_1)(1-x_1)^2\\ &\times\left(\int_0^{1-x_1}\left(\frac{\partial}{\partial x_2}\left(\Phi_{[0,1-x_1]}(x_2)\,v(x_1,x_2)\right)\right)^2dx_2\right)dx_1.\end{aligned}$$

We then get

$$\begin{aligned}&(1-x_1)^2\int_0^{1-x_1}\left(\frac{\partial}{\partial x_2}\left(\Phi_{[0,1-x_1]}(x_2)\,v(x_1,x_2)\right)\right)^2dx_2\\ &=(1-x_1)^2\left\|\left(\Phi_{[0,1-x_1]}v(x_1,\cdot)\right)'\right\|^2_{L^2(0,1-x_1)}\\ &\overset{\text{Lem. 2}}{\leq} C(p+1)^2\left\|\Phi^{1/2}_{[0,1-x_1]}v(x_1,\cdot)\right\|^2_{L^2(0,1-x_1)}.\end{aligned}$$

Since $\Phi^2_{[0,1]}(x_1)\,\Phi_{[0,1-x_1]}(x_2) \le \Phi_{\widehat{K}}(x_1,x_2)$ we end up with

$$\begin{aligned}\left\|\partial_2\left(\Phi_{\widehat{K}}v\right)\right\|^2_{L^2(\widehat{K})} &\le C\,(p+1)^2\\ &\times\int_0^1\int_0^{1-x_1}\Phi^2_{[0,1]}(x_1)\,\Phi_{[0,1-x_1]}(x_2)\,v^2(x_1,x_2)\,dx_2dx_1\\ &\le C\,(p+1)^2\left\|\Phi^{1/2}_{\widehat{K}}v\right\|^2_{L^2(\widehat{K})}.\end{aligned}$$

Since $\widehat{K}$, $\psi_{\widehat{K}}$, and the integral are invariant under permutations of the coordinates, the same estimate holds for the other partial derivatives. □

Lemma 4 *Let K be regularly h-refined and let $\Phi^{(2)}_K$ be as explained in Definition 3 and illustrated in Fig. 3. Then, it holds for all $v\in\mathbb{P}_p(K)$*

$$\left\|\nabla\left(\Phi^{(2)}_K v\right)\right\|_{L^2(K)} \le C\frac{p+1}{h_K}\left\|\sqrt{\Phi^{(2)}_K}v\right\|_{L^2(K)}.$$

Proof Via an affine transformation it suffices to prove the result for the reference element $K_1=\widehat{K}$ and $K_2=\operatorname{conv}\left\{\binom{0}{0},\binom{0}{1},\binom{-1}{0}\right\}$. The common edge is $E=\{0\}\times(0,1)$. Let $K=K_1\cup K_2$. The edge bubble $\Phi^{(2)}_K$ [cf. (5.4)] is given by

$$\Phi^{(2)}_K(x_1,x_2)=x_2\,(1-|x_1|-x_2).$$

We first consider the derivative with respect to x_2. Let

$$\Phi_{[0,1-|x_1|]}(x_2)=\frac{x_2\,(1-|x_1|-x_2)}{(1-|x_1|)^2},$$

i.e., $\Phi_{[0,1-|x_1|]}$ is the one-dimensional bubble function for $[0,1-|x_1|]$ and satisfies $\Phi^{(2)}_K=(1-|x_1|)^2\,\Phi_{[0,1-|x_1|]}$. Hence,

$$\begin{aligned}\left\|\partial_2\left(\Phi^{(2)}_K v\right)\right\|^2_{L^2(K)} &= \int_{-1}^1(1-|x_1|)^4\\ &\times\int_0^{1-|x_1|}\left(\partial_2\left(\Phi_{[0,1-|x_1|]}(x_2)\,v(x_1,x_2)\right)\right)^2dx_2dx_1\\ &\overset{\text{Lem. 2}}{\le} C\,(p+1)^2\int_{-1}^1(1-|x_1|)^2\\ &\times\int_0^{1-|x_1|}\Phi_{[0,1-|x_1|]}(x_2)\,v^2(x_1,x_2)\,dx_2dx_1\\ &= C\,(p+1)^2\int_K\Phi^{(2)}_K v^2.\end{aligned}\tag{8.2}$$

Next, we will estimate the derivative with respect to x_1. We split the triangle into the two regions

$$D_1:=\left\{\binom{x_1}{x_2}\in K: x_2\le\frac12\right\}\quad\text{and}$$

$$D_2:=\operatorname{conv}\left\{\binom{1/2}{1/2},\binom{0}{1},\binom{-1/2}{1/2}\right\}.$$

In addition, we will need

$$D_3:=\operatorname{conv}\left\{\binom{0}{0},\binom{1/2}{1/2},\binom{0}{1},\binom{-1/2}{1/2}\right\}.$$

On D_1 we obtain

$$\begin{aligned}\left\|\partial_1\left(\Phi^{(2)}_K v\right)\right\|^2_{L^2(D_1)} &\le \int_0^{1/2}x_2^2\,(1-x_2)^2\\ &\times\int_{x_2-1}^{1-x_2}\left(\partial_1\left(\frac{1-|x_1|-x_2}{1-x_2}v(x_1,x_2)\right)\right)^2dx_1dx_2\\ &\overset{\text{Lem. 5}}{\le} C\,(p+1)^2\int_0^{1/2}\int_{x_2-1}^{1-x_2}\frac{x_2}{1-x_2}\\ &\times\Phi^{(2)}_K(x_1,x_2)\,v^2(x_1,x_2)\,dx_1dx_2\\ &\le C\,(p+1)^2\left\|\sqrt{\Phi^{(2)}_K}v\right\|^2_{L^2(D_1)},\end{aligned}\tag{8.3}$$

since $x_2/(1-x_2)\le 1$ on D_1.

On D_2, we observe that

$$\frac12\left\|\partial_1\left(\Phi^{(2)}_K v\right)\right\|^2_{L^2(D_2)}\le\left\|v\partial_1\Phi^{(2)}_K\right\|^2_{L^2(D_2)}+\left\|\Phi^{(2)}_K\partial_1 v\right\|^2_{L^2(D_2)}.\tag{8.4}$$

Let $d:D_3\to\mathbb{R}$ be defined by

$$d(x_1,x_2)=c_3\operatorname{dist}\left((x_1,x_2)^{\mathsf T},\partial D_3\right)$$

where the scaling c_3 is chosen such that d interpolates $\Phi^{(2)}_K$ at the vertices of the two triangles $K_m\cap D_3$, $m=1,2$. Note that

$$d\le\Phi^{(2)}_K\le 1 \text{ pointwise in } D_3 \text{ and } \Phi^{(2)}_K\le 2d \text{ pointwise in } D_2.$$

Since $\left\|\partial_1\Phi^{(2)}_K\right\|_{L^\infty(K)}\le C$ we obtain for the first term in (8.4) as in (5.16b)

$$\begin{aligned}\left\|v\partial_1\Phi^{(2)}_K\right\|_{L^2(D_2)} &\le C\,\|v\|_{L^2(D_2)}\\ &\le C\,\|v\|_{L^2(D_3)}\le C\,(p+1)\left\|d^{1/2}v\right\|_{L^2(D_3)}\\ &\le C\,(p+1)\left\|\sqrt{\Phi^{(2)}_K}v\right\|_{L^2(D_3)}.\end{aligned}\tag{8.5a}$$

For the second term in (8.4) we get, again, as in (5.16b)

$$\begin{aligned}\left\|\Phi^{(2)}_K\partial_1 v\right\|_{L^2(D_2)} &\le 2\,\|d\partial_1 v\|_{L^2(D_3)}\\ &\overset{\text{[26, (23) with }\delta=1]}{\le} C\,(p+1)\left\|\sqrt{d}v\right\|_{L^2(D_3)}\\ &\le C\,(p+1)\left\|\sqrt{\Phi^{(2)}_K}v\right\|_{L^2(D_3)}.\end{aligned}\tag{8.5b}$$

The combination of (8.4) and (8.5) yields

$$\left\|\partial_1\left(\Phi^{(2)}_K v\right)\right\|^2_{L^2(D_2)}\le C\,(p+1)^2\left\|\sqrt{\Phi^{(2)}_K}v\right\|^2_{L^2(D_3)}.\tag{8.6}$$

The combination of (8.2), (8.3), and (8.6) yields the assertion. □

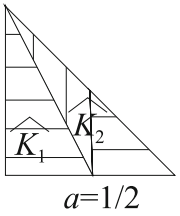

Fig. 8 Reference triangle $\widehat{K}$ which is split into $\widehat{K}_1$ and $\widehat{K}_2$. The *shaded regions* illustrate the integration domains in the splitting of the integral in (8.8)

The following lemma is illustrated in Fig. 8.

Lemma 5 *Let $\widehat{K}$ be the reference triangle split into $\widehat{K}_1 = \operatorname{conv}\left(\binom{0}{0}, \binom{a}{0}, \binom{0}{1}\right)$ and $\widehat{K}_2 = \operatorname{conv}\left(\binom{a}{0}, \binom{1}{0}, \binom{0}{1}\right)$ for some $a \in]0,1[$. Let φ_E^{lin} denote the continuous, piecewise linear function which has value 1 at $\binom{a}{0}$ and vanishes at $\partial\widehat{K}\backslash E_1$ with $E_1 = [0,1]\times\{0\}$. Then, for any polynomial $v \in \mathbb{P}_p$ which is constant with respect to x_2 it holds*

$$\left\|\sqrt{\varphi_E^{\mathrm{lin}}}v\right\|_{L^2(\widehat{K})} \le C\left\|\sqrt{\varphi_E^{\mathrm{lin}}}v\right\|_{L^2(E_1)},$$
$$\left\|\nabla\left(\varphi_E^{\mathrm{lin}}v\right)\right\|_{L^2(\widehat{K})} \le C\,(p+1)\left\|\sqrt{\varphi_E^{\mathrm{lin}}}v\right\|_{L^2(E_1)}.$$

Proof We prove this lemma only for $a = 1/2$ to reduce technicalities. The arguments apply verbatim for the general case. The function φ_E^{lin} and its partial derivatives are given by

$$\varphi_E^{\mathrm{lin}}(x_1,x_2) = \begin{cases} 2x_1 & (x_1,x_2)\in\widehat{K}_1,\\ 2(1-x_1-x_2) & (x_1,x_2)\in\widehat{K}_2,\end{cases}$$
$$\partial_1\varphi_E^{\mathrm{lin}}(x_1,x_2) = \begin{cases} 2 & (x_1,x_2)\in\widehat{K}_1,\\ -2 & (x_1,x_2)\in\widehat{K}_2.\end{cases}$$
$$\partial_2\varphi_E^{\mathrm{lin}}(x_1,x_2) = \begin{cases} 0 & (x_1,x_2)\in\widehat{K}_1,\\ -2 & (x_1,x_2)\in\widehat{K}_2,\end{cases}$$

Since $v \in \mathbb{P}_p$ is constant with respect to x_2 we write, with a slight abuse of notation, $v(x_1,x_2) = v(x_1)$. Hence,

$$\left\|\sqrt{\varphi_E^{\mathrm{lin}}}v\right\|^2_{L^2(\widehat{K})} = \int_0^1 v^2(x_1)\left(\int_0^{1-x_1}\varphi_E^{\mathrm{lin}}(x_1,x_2)\,dx_2\right)dx_1.$$

The result of the inner integration is

$$r(x_1) := \int_0^{1-x_1}\varphi_E^{\mathrm{lin}}(x_1,x_2)\,dx_2$$
$$= \begin{cases} 2\left(\int_0^{1-2x_1} x_1\,dx_2 + \int_{1-2x_1}^{1-x_1}(1-x_1-x_2)\,dx_2\right)\\ 2\int_0^{1-x_1}(1-x_1-x_2)\,dx_2\end{cases}$$
$$= \begin{cases} x_1(2-3x_1) & x_1\le 1/2,\\ (1-x_1)^2 & x_1 > 1/2.\end{cases}$$

Since $r \le \varphi_E^{\mathrm{lin}}(\cdot,0)$ pointwise on $[0,1]$, the first assertion follows.

Next, we investigate the derivative with respect to x_2. It holds

$$\partial_2\left(\varphi_E^{\mathrm{lin}}v\right) = v\partial_2\varphi_E^{\mathrm{lin}} = v\times\begin{cases}0 & (x_1,x_2)\in\widehat{K}_1,\\ -2 & (x_1,x_2)\in\widehat{K}_2.\end{cases}$$

Thus,

$$\int_{\widehat{K}}\left(\partial_2\left(\varphi_E^{\mathrm{lin}}v\right)\right)^2 = \int_{\widehat{K}}\left(v\partial_2\varphi_E^{\mathrm{lin}}\right)^2 \le 4\int_{\widehat{K}}v^2 \le 4\int_0^1 v^2$$
$$\overset{\text{[26, Lemma 2.4 with }\alpha=0\text{ and }\beta=1]}{\le} 4(p+1)^2\int_0^1\Phi_{[0,1]}v^2$$
$$\le 4C(p+1)^2\int_0^1\varphi_E^{\mathrm{lin}}v^2. \tag{8.7}$$

For the derivative with respect to x_1, we get

$$q(x_1,x_2) := \partial_1\left(\varphi_E^{\mathrm{lin}}(x_1,x_2)\,v(x_1)\right)$$
$$= 2\begin{cases}(x_1v(x_1))' & \text{in }\widehat{K}_1,\\ (1-x_1-x_2)v'(x_1) - v(x_1) & \text{in }\widehat{K}_2.\end{cases}$$

The function q is on $\widehat{K}_1$ and on $\widehat{K}_2$, an affine function with respect to x_2. We split the integral into

$$\int_0^1\int_0^{1-x_1}\cdots$$
$$= \int_0^{1/2}\int_0^{1-2x_1}\cdots + \int_0^{1/2}\int_{1-2x_1}^{1-x_1}\cdots + \int_{1/2}^1\int_0^{1-x_1}$$
$$=: W_1 + W_2 + W_3 \tag{8.8}$$

and obtain for the summands

$$W_1 = 4\int_0^{1/2}(1-2x_1)\left((x_1v(x_1))'\right)^2dx_1$$
$$\le \int_0^{1/2}\left((2x_1v(x_1))'\right)^2dx_1$$
$$\le \int_0^1\left(\left(\varphi_E^{\mathrm{lin}}(x_1,0)\,v(x_1)\right)'\right)^2dx_1$$
$$\overset{\text{Cor. 5}}{\le} C(p+1)^2\left\|\sqrt{\varphi_E^{\mathrm{lin}}}v\right\|^2_{L^2([0,1])}.$$

For W_2 and W_3 we use the fact that the Simpson rule is exact for quadratic polynomials and (8.7) to obtain

$$W_2 = 4\int_0^{1/2}\int_{1-2x_1}^{1-x_1}\left((1-x_1-x_2)v'(x_1) - v(x_1)\right)^2dx_2dx_1$$
$$= \frac{2}{3}\int_0^{1/2}x_1\left(\left(\underbrace{x_1v'(x_1)-v(x_1)}_{=(x_1v(x_1))'-2v(x_1)}\right)^2 + 4\left(\underbrace{\frac{x_1}{2}v'(x_1)-v(x_1)}_{\frac{(x_1v(x_1))'}{2}-\frac{3}{2}v(x_1)}\right)^2 + v^2(x_1)\right)dx_1$$

$$\leq \frac{2}{3}\int_0^{1/2} x_1\left(4\left(\left(x_1 v\left(x_1\right)\right)'\right)^2+27v^2\left(x_1\right)\right)dx_1$$
$$\leq \frac{4}{3}\int_0^1\left(\left(\left(\varphi_E^{\text{lin}}\left(x_1,0\right)v\left(x_1\right)\right)'\right)^2\right)dx_1+9\left\|v\right\|^2_{L^2([0,1])}$$
$$\overset{\text{Cor. 5}}{\leq} C\left(p+1\right)^2\left\|\sqrt{\varphi_E^{\text{lin}}}v\right\|^2_{L^2([0,1])}.$$

Finally, for W_3 we obtain

$$\begin{aligned}
W_3 &= 4\int_{1/2}^1\int_0^{1-x_1}\left(\left(1-x_1-x_2\right)v'\left(x_1\right)-v\left(x_1\right)\right)^2dx_2dx_1\\
&= 4\int_{1/2}^1\frac{1-x_1}{6}\left(\left(\underbrace{\left(1-x_1\right)v'\left(x_1\right)-v\left(x_1\right)}_{\left(\left(1-x_1\right)v\left(x_1\right)\right)'}\right)^2\right.\\
&\quad\left.+4\left(\underbrace{\left(\frac{1-x_1}{2}\right)v'\left(x_1\right)-v\left(x_1\right)}_{\frac{1}{2}\left(\left(\left(1-x_1\right)v\left(x_1\right)\right)'-v\left(x_1\right)\right)}\right)^2+v^2\left(x_1\right)\right)dx_1\\
&\leq 4\int_{1/2}^1\frac{1-x_1}{6}\left(3\left(\left(\left(1-x_1\right)v\left(x_1\right)\right)'\right)^2+9v\left(x_1\right)\right)dx_1\\
&\leq \frac{1}{12}\int_{1/2}^1 3\left(\left(2\left(1-x_1\right)v\left(x_1\right)\right)'\right)^2+36v^2\left(x_1\right)dx_1\\
&\leq \frac{1}{4}\int_0^1\left(\left(\varphi_E^{\text{lin}}\left(x_1\right)v\left(x_1\right)\right)'\right)^2dx_1+3\left\|v\right\|^2_{L^2[0,1]}\\
&\overset{\text{Cor. 5}}{\leq} C\left(p+1\right)^2\left\|\sqrt{\varphi_E^{\text{lin}}}v\right\|^2_{L^2([0,1])}.
\end{aligned}$$

□

Corollary 6 *Let $\widehat{K}$ be the reference triangle and let $\varphi_E\left(x_1,x_2\right)=x_1\left(1-x_1-x_2\right)$ denote the quadratic edge bubble on $\widehat{K}$ for the edge $E_1=[0,1]\times\{0\}$. Then, for any polynomial $v\in\mathbb{P}_p$ which is constant with respect to x_2 it holds*

$$\left\|\sqrt{\varphi_E}v\right\|_{L^2(\widehat{K})}\leq C\left\|\sqrt{\varphi_E}v\right\|_{L^2(E_1)},$$
$$\left\|\nabla\left(\varphi_E v\right)\right\|_{L^2(\widehat{K})}\leq C\left(p+1\right)\left\|\sqrt{\varphi_E}v\right\|_{L^2(E_1)}.$$

The proof follows by a simple repetition of the arguments of the proof of Lemma 5.

Lemma 6 *Let K be a triangle and E one of its edges. Then, for any of the functions φ_E (6.4) and corresponding version Φ_K as in (5.4) it holds*

$$\left\|\varphi_E^{1/2}v\right\|_{L^2(D)}\leq C\left(p+1\right)\left\|\Phi_K^{1/2}v\right\|_{L^2(D)}\qquad\forall v\in\mathbb{P}_p\left(D\right).$$

The proof requires two preparatory lemmata and follows the ideas in [28, Appendix D].

Lemma 7 *Let $I=[a,b]$ for some $a<b$ and let $\omega:I\to\mathbb{R}$ be a weight function which satisfies*

$$\exists A,B,D\geq 0\ \textit{with}\quad\begin{cases}\omega\ \textit{is positive in}\ ]a,b[\,,\\ \omega\left(x\right)\leq A\varphi_a\left(x\right)+B\varphi_b\left(x\right)+D\Phi_{[a,b]}\left(x\right),\end{cases}\qquad P_I$$

where $\varphi_b\left(x\right)=\frac{x-a}{b-a}$, $\varphi_a=1-\varphi_b$, and $\Phi_{[a,b]}=\varphi_a\varphi_b$ as in (2). Then, it holds

$$\left\|\omega^{1/2}v\right\|_{L^2(I)}\leq C\left(p+1\right)\left\|\sqrt{\omega\Phi_{[a,b]}}v\right\|_{L^2(I)}\qquad\forall v\in\mathbb{P}_p\left(I\right),$$

where C is independent of p, v, ω, a, b.

Proof By employing an affine transform it is sufficient to prove the assertion for the unit interval $I=[0,1]$.

(a) $\omega\left(x\right)=\Phi_{[0,1]}\left(x\right)=x\left(1-x\right)$. We may apply standard inverse estimates to obtain

$$\left\|\Phi_{[0,1]}^{1/2}v\right\|_{L^2(I)}\overset{\text{[26, with }\alpha=1,\beta=2]}{\leq} C\left(p+1\right)\left\|\Phi_{[0,1]}v\right\|_{L^2(I)}=C\left(p+1\right)\left\|\sqrt{\omega\Phi_{[0,1]}}v\right\|_{L^2(I)}.$$

(b) For $\omega\left(x\right)=\varphi_b\left(x\right)=x$ we observe that $\omega\left(x\right)\leq 2\Phi_{[0,1]}\left(x\right)$ holds for all $0\leq x\leq 1/2$ so that

$$\begin{aligned}\left\|\omega^{1/2}v\right\|^2_{L^2(I)}&\leq 2\left\|\Phi_{[0,1]}^{1/2}v\right\|^2_{L^2(I)}+\left\|v\right\|^2_{L^2\left(\left[\frac{1}{2},1\right]\right)}\\&\overset{\text{[26, Lem. 2.4]}}{\leq} C\left(p+1\right)^2\\&\quad\times\left(\left\|\Phi_{[0,1]}v\right\|^2_{L^2(I)}+\left\|\Phi_{\left[\frac{1}{2},1\right]}^{1/2}v\right\|^2_{L^2\left(\left[\frac{1}{2},1\right]\right)}\right).\end{aligned}$$

The result now follows from $\Phi_{[0,1]}^{1/2}\leq\omega^{1/2}$ pointwise in $[0,1]$ and $\sqrt{2\omega}\geq 1$ pointwise on $\left[\frac{1}{2},1\right]$.

(c) The $\omega\left(x\right)=\varphi_a\left(x\right)$ follows from Case b by symmetry.

(d) Let ω be a general weight function which satisfies the assumptions of the lemma. Hence, from Part a,b,c we conclude that

$$\begin{aligned}\left\|\omega^{1/2}v\right\|^2_{L^2(I)}&=A\left\|\sqrt{\varphi_0}v\right\|^2_{L^2(I)}+B\left\|\sqrt{\varphi_1}v\right\|^2_{L^2(I)}\\&\quad+D\left\|\sqrt{\Phi_{[0,1]}}v\right\|^2_{L^2(I)}\leq C'\left(p+1\right)^2\left\|\sqrt{\omega\Phi_{[0,1]}}v\right\|^2_{L^2(I)}\end{aligned}$$

holds. □

The following lemma is a weighted version of [28, Lem. D3].

Lemma 8 *Let $d \in (0,1)$, a, b be given such that $-1+ad < 1+bd$ and define the trapezoid*

$$D := D(a,b,d) := \Big\{(x_1,x_2) \in \mathbb{R}^2 \mid x_2 \in (0,d) \quad \text{and} \quad -1+ax_2 < x_1 < 1+bx_2\Big\}.$$

Let $\omega \in \mathbb{P}_2(D)$ be a polynomial such that for any $0 \le x_2 \le d$, $\omega(\cdot, x_2)$ has property $P_{[-1+ax_2,1+bx_2]}$.

On D we define the weight function

$$\Phi_{a,b,d}(x_1,x_2) := \min\{|x_1-(-1+ax_2)|, |x_1-(1+bx_2)|\}$$

which measures the distance of the point (x_1,x_2) from the lateral edges of D. Then, there exists a constant $C = C(a,b,d)$ such that for all $p \in \mathbb{N}$ and all polynomials $v \in \mathbb{P}_p(D)$ it holds

$$\left\|\omega^{1/2} v\right\|_{L^2(D)} \le C(p+1)\left\|\sqrt{\omega \Phi_{a,b,d}}\, v\right\|_{L^2(D)}.$$

Proof Note that

$$C_{13}\Phi_{[-1+ax_2,1+bx_2]}(x_1) \le \Phi_{a,b,d}(x_1,x_2) \le C_{14}\Phi_{[-1+ax_2,1+bx_2]}(x_1)$$

for positive constants C_{13}, C_{14} which only depends on a, b, d. Hence, the one-dimensional case (Lemma 7) implies

$$\int_{-1+ax_2}^{1+bx_2} \omega^{1/2}(x_1,x_2)\, v^2(x_1,x_2)\, dx_1 \le C(p+1)^2 \times \int_{-1+ax_2}^{1+bx_2} \sqrt{\omega(x_1,x_2)\,\Phi_{a,b,d}(x_1,x_2)}\, v^2(x_1,x_2)\, dx_1.$$

Integrating this estimate over $x_2 \in (0,d)$ completes the proof. □

Proof of Lemma 6 By using an affine pullback we may restrict to the case that K is the equi-sided triangle conv $\left(\binom{0}{0}, \binom{1}{0}, \frac{1}{2}\binom{1}{\sqrt{3}}\right)$ and $E = (0,1) \times \{0\}$.

It turns out that the proofs for the different cases in (6.4) for φ_E and in (5.4) for Φ_K uses the same arguments and we work them out exemplarily for the case of the quadratic edge bubble

$$\varphi_E(x_1,x_2) = \left(x_1 - \frac{x_2}{\sqrt{3}}\right)\left(1 - x_1 - \frac{x_2}{\sqrt{3}}\right)$$

and for $\Phi_K = \Phi_K^{(3)}$ being the cubic bubble on K.

First, we will cover K with 4 trapezoids and one triangle: Let $v = \left(\cos\frac{\pi}{4}, \sin\frac{\pi}{4}\right)^\intercal = 2^{-1/2}(1,1)^\intercal$. Then,

1. $T_1 := \left\{\binom{\hat{x}_1}{0} + sv : \begin{pmatrix} 0 \le \hat{x}_1 \le 1/2 \\ 0 \le s \le L_1(\hat{x}_1)\end{pmatrix}\right\}$ with

$$L_1(\hat{x}_1) := \frac{\sqrt{6}}{1+\sqrt{3}}(1-\hat{x}_1).$$

2. T_2 : mirror image of T_1 with respect to the angle bisector at $(0,0)^\intercal$.
3. T_3 : counter-clockwise rotations of T_1 by $\frac{3\pi}{4}$ about the barycenter of K.
4. T_4 : mirror image of T_3 with respect to the angle bisector at $(1,0)^\intercal$.
5. $T_5 := \{(x_1,x_2)^\intercal \in K \mid x_2 \ge 1/2\}$.

Case T_1: We introduce

$$\chi : \begin{pmatrix} 0 \le \hat{x}_1 \le 1/2 \\ 0 \le s \le L_1(\hat{x}_1)\end{pmatrix} \to T_1 \quad \text{by} \quad \chi(\hat{x}_1, s) := \binom{\hat{x}_1}{0} + sv$$

The bubble function φ_E restricted to the line $\binom{\hat{x}_1}{0} + sv$ results in

$$\psi_{\hat{x}_1}(s) := \varphi_E \circ \chi(\hat{x}_1, s) = (L_1(\hat{x}_1) - s)\left(\frac{\sqrt{3}+1}{\sqrt{6}}\hat{x}_1 + \frac{s}{3}\right) \quad \forall \begin{pmatrix} 0 \le \hat{x}_1 \le 1/2 \\ 0 \le s \le L_1(\hat{x}_1)\end{pmatrix}.$$

Note that the function $\psi_{\hat{x}_1}$ satisfies the assumptions of Lemma 7 and $\hat{v} := v \circ \chi$ is a polynomial of maximal degree p. Hence,

$$\int_{T_1} \varphi_E v^2 = \int_0^{1/2}\left(\int_0^{L_1(\hat{x}_1)} \psi_{\hat{x}_1}(s)\, \hat{v}^2 ds\right) dx_1$$
$$\le C(p+1)^2 \int_0^{1/2}\left(\int_0^{L_1(\hat{x}_1)} \psi_{\hat{x}_1}(s)\, \Phi_{[0,L_1(\hat{x}_1)]}(s)\, \hat{v}^2 ds\right) d\hat{x}_1.$$

Composing $\Phi_{[0,L_1(\hat{x}_1)]}(s)$ with χ^{-1} yields the function

$$d(x_1,x_2) = \frac{1+\sqrt{3}}{\sqrt{3}} \frac{x_2\left(1 - x_1 - \frac{x_2}{\sqrt{3}}\right)}{(1-x_1+x_2)^2}.$$

Note that the distance function

$$\Phi_K^1(x_1,x_2) = \operatorname{dist}\left((x_1,x_2)^\intercal, \partial K\right)$$

is piecewise linear on K. It is easy to verify that $d(x_1,x_2) \le C\Phi_K^1(x_1,x_2)$ pointwise on T_1 for some $C = O(1)$ so that

$$\int_{T_1} \varphi_E v^2 \le C'(p+1)^2 \int_{T_1} \varphi_E \Phi_K^1 v^2.$$

Since $\varphi_E \Phi_K^1 \le \tilde{C}\Phi_K$ pointwise on K we have proved the assertion for T_1.

Case T_3: The proof for the trapezoid T_3 follows by symmetry.

Case T_2: Next, we will consider the trapezoid T_2 and first note that by interchanging the x_1, x_2-variables the case becomes equivalent to the estimate

$$\int_{T_1} \varphi_{\tilde{E}} v^2 \le C'(p+1)^2 \int_{T_1} \Phi_K^{(3)} v^2 \qquad v \in \mathbb{P}_p(T_1),$$

where $\varphi_{\tilde{E}}$ is the qudratic edge bubble for the edge $\tilde{E} = \overline{\binom{0}{0}, \binom{1/2}{\sqrt{3}/2}}$ with explicit form

$$\varphi_{\tilde{E}}(x_1, x_2) = \frac{2}{\sqrt{3}} x_2 \left(1 - x_1 - \frac{x_2}{\sqrt{3}}\right).$$

This time, the bubble function $\varphi_{\tilde{E}}$, restricted to the line $\binom{\hat{x}_1}{0} + sv$, is given by

$$\tilde{\psi}_{\hat{x}_1}(s) := \varphi_{\tilde{E}} \circ \chi(\hat{x}_1, s) = \frac{\sqrt{3}+1}{3} s\left(L(\hat{x}_1) - s\right) \quad \forall \begin{pmatrix} 0 \le \hat{x}_1 \le 1/2 \\ 0 \le s \le L_1(\hat{x}_1) \end{pmatrix}.$$

The function $\tilde{\psi}_{\hat{x}_1}$ satisfies the assumptions of Lemma 7 so that

$$\int_{T_1} \varphi_{\tilde{E}} v^2 = \int_0^{1/2} \left(\int_0^{L_1(\hat{x}_1)} \tilde{\psi}_{\hat{x}_1}(s)\, \hat{v}^2 ds \right) dx_1$$
$$\le C\,(p+1)^2 \int_0^{1/2} \left(\int_0^{L_1(\hat{x}_1)} \tilde{\psi}_{\hat{x}_1}(s)\, \Phi_{[0, L_1(\hat{x}_1)]}(s)\, \hat{v}^2 ds \right) d\hat{x}_1.$$

Now we can argue as for the Case of T_1 to obtain

$$\int_{T_1} \varphi_{\tilde{E}} v^2 \le C'\,(p+1)^2 \int_{T_1} \varphi_{\tilde{E}} \Phi_K^1 v^2.$$

Since $\varphi_{\tilde{E}} \Phi_K^1 \le \tilde{C} \Phi_K$ pointwise on K the assertion follows for T_2.

Case T_4: The proof for the trapezoid T_4 again follows by symmetry from the case T_2.

Case T_5: On T_5 we have the pointwise estimate $\varphi_E \le C \Phi_K^{(3)}$ and the estimate for T_5 is trivial. □

References

1. Ainsworth, M., Oden, J.T.: A posteriori error estimation in finite element analysis. Comput. Methods Appl. Mech. Eng. **142**(1–2), 1–88 (1997)
2. Ainsworth, M., Oden, J.T.: A Posteriori Error Estimation in Finite Element Analysis. Wiley, New York (2000)
3. Ainsworth, M., Senior, B.: Aspects of an adaptive *hp*-finite element method: adaptive strategy, conforming approximation and efficient solvers. Comput. Methods Appl. Mech. Eng. **150**(1–4), 65–87 (1997). Symposium on Advances in Computational Mechanics, Vol. 2 (Austin, TX, 1997)
4. Babuška, I., Rheinboldt, W.C.: A-posteriori error estimates for the finite element method. Int. J. Numer. Methods Eng. **12**, 1597–1615 (1978)
5. Babuška, I., Rheinboldt, W.C.: Error estimates for adaptive finite element computations. SIAM J. Numer. Anal. **15**, 736–754 (1978)
6. Bangerth, W., Rannacher, R.: Adaptive Finite Element Methods for Differential Equations. Birkhäuser, Basel (2003)
7. Bank, R.: PLTMG: A Software Package for Solving Elliptic Partial Differential Equations. Users' Guide 11.0. Department of Mathematics, University of California at San Diego, La Jolla (2012)
8. Bank, R.E., Nguyen, H.: *hp* adaptive finite elements based on derivative recovery and superconvergence. Comput. Vis. Sci. **14**, 287–299 (2012)
9. Bank, R.E., Weiser, A.: Some a posteriori error estimators for elliptic partial differential equations. Math. Comput. **44**, 283–301 (1985)
10. Bank, R.E., Xu, J.: Asymptotically exact a posteriori error estimators. II. General unstructured grids. SIAM J. Numer. Anal. **41**(6), 2313–2332 (2003)
11. Binev, P., Dahmen, W., DeVore, R.: Adaptive finite element methods with convergence rates. Numer. Math. **97**(2), 219–268 (2004)
12. Bornemann, F.A., Erdmann, B., Kornhuber, R.: A posteriori error estimates for elliptic problems in two and three space dimensions. SIAM J. Numer. Anal. **33**(3), 1188–1204 (1996)
13. Braess, D., Pillwein, V., Schöberl, J.: Equilibrated residual error estimates are *p*-robust. Comput. Methods Appl. Mech. Eng. **198**(13–14), 1189–1197 (2009)
14. Bürg, M., Dörfler, W.: Convergence of an adaptive *hp* finite element strategy in higher space-dimensions. Appl. Numer. Math. **61**(11), 1132–1146 (2011)
15. Carstensen, C., Sauter, S.: A posteriori error analysis for elliptic PDEs on domains with complicated structures. Numer. Math. **96**, 691–712 (2004)
16. Ciarlet, P.: The Finite Element Method for Elliptic Problems. North-Holland, Amsterdam (1987)
17. Demkowicz, L.: Computing with *hp*-adaptive finite elements. Vol. 1. Chapman & Hall/CRC Applied Mathematics and Nonlinear Science Series. Chapman & Hall/CRC, Boca Raton (2007)
18. Dörfler, W.: A convergent adaptive algorithm for Poisson's equation. SIAM J. Numer. Anal. **33**(3), 1106–1124 (1996)
19. Dörfler, W., Nochetto, R.H.: Small data oscillation implies the saturation assumption. Numer. Math. **91**(1), 1–12 (2002)
20. Dörfler, W., Rumpf, M.: An adaptive strategy for elliptic problems including a posteriori controlled boundary approximation. Math. Comput. **67**(224), 1361–1382 (1998)
21. Eibner, T., Melenk, J.: An adaptive strategy for hp-fem based on testing for analyticity. Comput. Mech. **39**, 575–595 (2007)
22. Hackbusch, W., Wappler, J.U.: Remarks on a posteriori error estimation for inaccurate finite element solutions. Computing **60**(2), 175–191 (1998)
23. Koornwinder, T.: Two-variable analogues of the classical orthogonal polynomials. In Theory and application of special functions. In: Proc. Advanced Sem., Math. Res. Center, Univ. Wisconsin, Madison, Wis., pp. 435–495 (1975). Math. Res. Center, Univ. Wisconsin, Publ. No. 35. Academic Press, New York (1975)
24. Koornwinder, T.H., Sauter, S.A.: The intersection of bivariate orthogonal polynomials on triangle patches. Math. Comput. (in press). arXiv:1307.8429v3 [math.NA]
25. Mekchay, K., Nochetto, R.H.: Convergence of adaptive finite element methods for general second order linear elliptic PDEs. SIAM J. Numer. Anal. **43**(5), 1803–1827 (2005)
26. Melenk, J.M.: *hp*-interpolation of nonsmooth functions and an application to *hp*-a posteriori error estimation. SIAM J. Numer. Anal. **43**(1), 127–155 (2005)
27. Melenk, J.M., Wohlmuth, B.I.: On residual-based a posteriori error estimation in *hp*-FEM. Adv. Comput. Math., **15**(1–4), 311–331 (2002), 2001
28. Melenk, M.: HP-Interpolation of Non-smooth Functions (extended version). Technical Report NI03050, Isaac Newton Institute for Mathematical Sciences (2003)
29. Mitchell, W.F., McClain, M.A.: A survey of *hp*-adaptive strategies for elliptic partial differential equations. In: Recent Advances in Computational and Applied Mathematics, pp. 227–258. Springer, Dordrecht (2011)
30. Prager, W., Synge, J.L.: Approximations in elasticity based on the concept of function space. Quart. Appl. Math. **5**, 241–269 (1947)

31. Proriol, J.: Sur une famille de polynomes à deux variables orthogonaux dans un triangle. C. R. Acad. Sci. Paris **245**, 2459–2461 (1957)
32. Repin, S.: A Posteriori Estimates for Partial Differential Equations, vol. 4. Walter de Gruyter GmbH & Co. KG, Berlin (2008)
33. Schmidt, A., Siebert, K.G.: A posteriori estimators for the *h*-*p* version of the finite element method in 1D. Appl. Numer. Math. **35**(1), 43–66 (2000)
34. Solin, P., Dubcova, L., Dolezel, I.: Adaptive *hp*-FEM with arbitrary-level hanging nodes for Maxwell's equations. Adv. Appl. Math. Mech. **2**(4), 518–532 (2010)
35. Stevenson, R.: An optimal adaptive finite element method. SIAM J. Numer. Anal. **42**(5), 2188–2217 (2005)
36. Stevenson, R.: Optimality of a standard adaptive finite element method. Found. Comput. Math. **7**(2), 245–269 (2007)
37. Süli, E., Houston, P., Schwab, C.: *hp*-finite element methods for hyperbolic problems. In: The Mathematics of Finite Elements and Applications. X, MAFELAP 1999 (Uxbridge), pp. 143–162. Elsevier, Oxford (2000)
38. Verfürth, R.: A Review of A Posteriori Error Estimation and Adaptive Mesh Refinement. Wiley and Teubner, New York (1996)
39. Verfürth, R.: A Posteriori Error Estimation Techniques for Finite Element Methods. Oxford University Press, Oxford (2013)