

$L^2$-ORTHOGONAL PROJECTIONS ONTO FINITE ELEMENTS ON LOCALLY REFINED MESHES ARE $H^1$-STABLE

MICHAEL KARKULIK, CARL-MARTIN PFEILER, AND DIRK PRAETORIUS

Abstract. We merge and extend recent results which prove the $H^1$-stability of the $L^2$-orthogonal projection onto standard finite element spaces, provided that the underlying simplicial triangulation is appropriately graded. For lowest-order Courant finite elements $S^1(T)$ in $\mathbb{R}^d$, $d \geq 2$, we prove that such a grading is always ensured for adaptive meshes generated by newest vertex bisection. For higher-order finite elements $S^p(T)$, $p \geq 1$, we extend existing bounds on the polynomial degree with a computer-assisted proof. We also consider $L^2$-orthogonal projections onto certain subspaces of $S^p(T)$ which incorporate zero Dirichlet boundary conditions resp. an integral mean zero property.

1. Introduction

Let $T$ be a simplicial mesh of a $d$-dimensional domain (or manifold) $\Omega$ and $\Pi(T)$ be the $L^2$-orthogonal projection onto the finite element space $S^p(T)$ of $T$-piecewise polynomials of degree $\leq p$ which are globally continuous; see [5] for the formal definition. Some of the various theoretical and practical applications of $\Pi(T)$, cf. [BY14, BPS02, KPP13], require uniform $H^1$-stability, i.e.,

$$\|\Pi(T)v\|_{H^1(\Omega)} \leq C_1 \|v\|_{H^1(\Omega)} \quad \text{for all } v \in H^1(\Omega)$$

with a constant $C_1 > 0$ independent of $T$. While the proof is well-known in the case of globally quasi-uniform meshes [BX91], it is rather demanding in the case of locally refined meshes. Existing results are based on two different approaches:

(I) Imposing a-priori bounds on the grading of the mesh. Works based on this approach include [BPS02, Car02, CT87, EJ95, Ste02].

(II) Considering an arbitrary coarse mesh and a fixed refinement strategy. This approach was carried out in [BY14, Car04, KPP13].

Both approaches are substantial: The first one can be used for arbitrary sequences of meshes, as long as the grading fulfills the given a-priori bound. The second approach can be used for arbitrary coarse meshes but a fixed refinement strategy. This will imply an a-posteriori bound on the grading which may be higher than the a-priori bound of the first approach (e.g., if already the coarse mesh violates the a-priori bound). Hence, the advantage of either approach is the drawback of the other.

In order to combine these two approaches, we use ideas from [BY14]. The latter work links a level-function $\text{level} : T \to \mathbb{N}_0$ to the grading parameter $\mu > 1$ of a mesh via $\mu^-\text{level}(T) \simeq \text{diam}(T)$. If $\text{level}$ does not change too much between neighboring elements, a certain bound on $\mu$ then implies $H^1$-stability. The connection between level-function, grading parameter, and element diameter clearly allows for highly nonuniform meshes. The advantage of a fixed refinement strategy is that one can intrinsically define

\[\text{Date: July 3, 2014.}\]

\[2000 \text{ Mathematics Subject Classification. 65N30, 65N50.}\]

\[\text{Key words and phrases. newest vertex bisection, adaptive mesh-refinement, } L^2\text{-projection, } H^1\text{-stability.}\]
a level-function through the number of local refinements and hide the grading parameter in a constant. This concept is also used in the works [BY14, Car04, KPP13]. The works [Car04, KPP13] consider lowest-order elements and the two-dimensional case \( d = 2 \) only and are restricted to newest vertex bisection (NVB) or variants; see Section 4 below for the precise refinement rules of NVB. Up to now, NVB is the only local refinement strategy for simplicial meshes which can be used in mathematically justified adaptive finite element/boundary element methods; see [CKNS08, FKMP13, CFPP14] and the references therein for the fine properties of NVB used.

The work [BY14] is the first one to consider finite elements of higher order for \( d \in \{2, 3\} \) (i.e., the authors use a computer-assisted proof to show \( H^1 \)-stability for \( p \leq 12 \) for \( d = 2 \) resp. \( p \leq 7 \) for \( d = 3 \)); see [CT87] for \( d = 1 \) and general \( p \geq 1 \). In [BY14], it is assumed that \( \mu = 2 \) and that level changes at most by one on an vertex patch. However, these assumptions do not apply to common local mesh-refinement strategies such as NVB, where \( \mu = 2^{1/d} \), and [BY14] does not discuss in detail which refinement strategies are admissible and covered by its analysis.

The work at hand merges and extends the mentioned results in different ways: We do not restrict ourselves to a certain refinement strategy as we did in [KPP13], but rely on the more general connection between level-function and grading parameter as in [BY14]. We sharpen the results of the latter work by leaving the mesh-grading parameter \( \mu \) as well as the level difference \( L \) of neighboring elements variable. For the lowest-order case \( p = 1 \), this allows us to extend [Car04] and our own work [KPP13], which were restricted to NVB in \( d = 2 \), to arbitrary dimensions \( d \geq 2 \). Second, the bounds on the polynomial degrees from [BY14] can be improved, where we use a computer-assisted proof as in [BY14]. Finally, we also discuss immediate consequences like stability in positive and negative fractional-order Sobolev spaces and weighted \( L^2 \)-spaces. The present manuscript is also the first one in this context to consider \( H^1 \)-stability for subspaces which account for, e.g., Dirichlet or Neumann boundary conditions, i.e., all the spaces that are used in basic finite element theory.

The outline of this paper reads as follows: Section 2 introduces the mandatory notation and then formulates our main results (Theorem 4 and Theorem 5) as well as several implications. Section 3 provides a proof of Theorem 4 where we refine the analysis of [BY14]. Section 4 gives a short introduction to NVB and recalls its specific properties used for the proof of Theorem 5.

Throughout the proofs, we use the abbreviate notation \( A \lesssim B \) which means \( A \leq cB \) with some multiplicative constant \( c > 0 \) which is clear from the context. Moreover, we write \( A \simeq B \) to abbreviate \( A \lesssim B \leq A \).

2. General Notation & Main Results

2.1. Simplicial mesh. A \textit{d-simplex} \( T \) in \( \mathbb{R}^D \) is the convex hull

\[
T = \operatorname{conv}\{x_0, \ldots, x_d\} := \left\{ \sum_{j=0}^{d} \lambda_j x_j : \lambda_j \geq 0 \text{ with } \sum_{j=0}^{d} \lambda_j = 1 \right\}
\]

of \( d+1 \) vertices \( x_0, \ldots, x_d \in \mathbb{R}^D \) that do not lie on a \((d-1)\)-dimensional hyperplane, i.e., a triangle for \( d = 2 \) resp. tetrahedron for \( d = 3 \). By \( \mathcal{N}(T) := \{x_0, \ldots, x_d\} \), we denote the set of vertices of a \( d \)-simplex \( T \).

A finite set \( \mathcal{T} \) of \( d \)-simplices in \( \mathbb{R}^D \) is said to be a \textit{partition} of a \( d \)-dimensional manifold \( \Omega \subset \mathbb{R}^D \), if \( \Omega \) is the (relative) interior of the union of these \( d \)-simplices, and if the intersection of any two different simplices \( T, T' \in \mathcal{T} \) has \( d \)-dimensional measure zero. A
partition $\mathcal{T}$ is said to be **conforming**, if the intersection of any two different simplices $T, T' \in \mathcal{T}$ is either empty or a hyperface of both $T$ and $T'$. Throughout this work, a conforming partition $\mathcal{T}$ is called a **mesh**. We denote by $\mathcal{F}(\mathcal{T})$ the set of its $(d - 1)$-dimensional faces. We note that all faces $F \in \mathcal{F}(\mathcal{T})$ are $(d - 1)$-dimensional simplices. The set $\mathcal{N}(\mathcal{T}) = \bigcup_{T \in \mathcal{T}} \mathcal{N}(T) = \{z_1, \ldots, z_N\}$ is the collection of all vertices of $\mathcal{T}$.

Recall that $|T|^{1/d} \leq \text{diam}(T)$, where $\text{diam}(T)$ denotes the Euclidean diameter of the $d$-simplex $T$ and where $|\cdot|$ is the $d$-dimensional measure. A mesh $\mathcal{T}$ of $d$-simplices is said to be $\gamma$-**shape regular** if

$$
\max_{T \in \mathcal{T}} \frac{\text{diam}(T)^d}{|T|} \leq \gamma < \infty.
$$

With a mesh $\mathcal{T}$, we associate the local mesh-width function

$$
h \in L^\infty(\Omega), \quad h_T := \text{diam}(T) \quad \text{for all} \ T \in \mathcal{T}.
$$

Let $\Gamma_D \subset \partial\Omega$ be a (possibly empty) relatively open subset of the boundary $\partial\Omega$ which is resolved by the mesh $\mathcal{T}$, i.e.,

$$
\Gamma_D = \bigcup \{T \cap \partial\Omega : T \cap \Gamma_D \neq \emptyset\}.
$$

Under this assumption, the set $\mathcal{T}_D := \{F \in \mathcal{F}(\mathcal{T}) : F \subseteq \Gamma_D\}$ provides a mesh of $\Gamma_D$ of $(d - 1)$-simplices.

2.2. **Finite element spaces.** For some mesh $\mathcal{T}$ which resolves the Dirichlet boundary $\Gamma_D$, we consider the finite element space

$$
S_p(\mathcal{T}) := \{v \in C(\Omega) : \forall T \in \mathcal{T} \quad v|_T \text{ is a polynomial of degree } \leq p\}.
$$

We let

$$
S_p^D(\mathcal{T}) := \{v \in S_p(\mathcal{T}) : v|_{\Gamma_D} = 0\}
$$

and note that $S_p^D(\mathcal{T}) = S_p(\mathcal{T})$ for $\Gamma_D = \emptyset$. We define the subspaces

$$
S_p^\circ(\mathcal{T}) := \{w \in S_p(\mathcal{T}) : \exists T \in \mathcal{T} \quad \text{supp}(w) \subseteq T\},
$$

$$
S_p^{\perp}(\mathcal{T}) := \{v \in S_p(\mathcal{T}) : \forall w \in S_p^\circ(\mathcal{T}) \quad \int_{\Omega} vw \, dx = 0\}.
$$

For each vertex $z_i \in \mathcal{N}(\mathcal{T}) = \{z_1, \ldots, z_N\}$, we denote by $\omega_i := \bigcup \{T \in \mathcal{T} : z_i \in T\}$ the vertex patch and

$$
S_p(\mathcal{T})_i := \{v \in S_p^{\perp}(\mathcal{T}) : \text{supp}(v) \subseteq \omega_i\}.
$$

Let $K_1, K_2 > 0$ satisfy the following two assumptions

$$(10a) \quad \forall v \in S_p^{\perp}(\mathcal{T})_i \exists v_i \in S_p^\circ(\mathcal{T})_i \quad v = \sum_{i=1}^N v_i \quad \text{and} \quad \sum_{i=1}^N \|v_i\|^2_{L^2(\Omega)} \leq K_1 \|v\|^2_{L^2(\Omega)};$$

$$(10b) \quad \forall v_i \in S_p^\circ(\mathcal{T})_i \quad \sum_{i=1}^N \|v_i\|^2_{L^2(\Omega)} \leq K_2 \sum_{i=1}^N \|v_i\|^2_{L^2(\Omega)}.$$

The precise values of $K_1$ and $K_2$ play a crucial role in due course. In [BY14], they have been computed numerically for $d = 2, 3$ and certain polynomial degrees $p$. The following lemma, where the $\mathcal{T}$-independence of the constants follows from a simple scaling argument, provides the main ingredient for such a computer-assisted proof.
Lemma 1. Suppose that $K_1, K_2 > 0$ satisfy the following two assumptions

\[(11a) \quad \forall v \in S^0(\mathcal{T})^\perp \exists v_i \in S^0(\mathcal{T})^\perp \forall T \in \mathcal{T} \quad v = \sum_{i=1}^{N} v_i \text{ and } \sum_{i=1}^{N} \|v_i\|^2_{L^2(T)} \leq K_1 \|v\|^2_{L^2(T)} , \]

\[(11b) \quad \forall v_i \in S^0(\mathcal{T})^\perp \forall T \in \mathcal{T} \quad \left\|\sum_{i=1}^{N} v_i\right\|^2_{L^2(T)} \leq K_2 \sum_{i=1}^{N} \|v_i\|^2_{L^2(T)} . \]

Then, (11a) follows even with the same constants $K_1, K_2$, which are in particular independent of $\mathcal{T}$. \hfill \Box

With the help of Lemma 1, the following bounds can be proved, where (ii) follows from (11b) and the fact that each element $T \in \mathcal{T}$ belongs to at most $(d + 1)$ node patches.

Proposition 2.  
(i) For $p = 1$ and $d \geq 2$, \((10)\) holds with $K_1 = 2$ and $K_2 = (d+2)/2$.
(ii) For $p \geq 1$ and $d \geq 2$, \((10b)\) holds with $K_2 = d + 1$.

The proof of (i) builds on the following eigenvalue result from \cite[Prop. 6.1]{BPS02}.

Lemma 3. Let $h_1, \ldots, h_{d+1} \in \mathbb{R}\setminus \{0\}$ and define

\[\hat{B} \in \mathbb{R}^{(d+1)\times (d+1)} , \quad \hat{B}_{jk} = (1 + \delta_{jk}) \left( \frac{h_i}{h_k} \right) \left( \frac{h_j}{h_k} \right) \]

with Kronecker’s delta $\delta_{jk}$. Then, the eigenvalues of $\hat{B}$ belong to $\{\lambda_+, \lambda_-, 2\}$ with

\[\lambda_{\pm} = d + 3 \pm \left( \sum_{j,k=1}^{d+1} \frac{h_j^2}{h_k^2} \right)^{1/2} . \]

Proof of Proposition 2 (i). Note that $S^1(\mathcal{T})^\perp = S^1(\mathcal{T})$. If we choose $h_1 = \cdots = h_{d+1} = 1$ in Lemma 3 then the matrix $\hat{B}$ is in fact a mass-matrix on some fixed reference element $\hat{T} \subseteq \mathbb{R}^d$. According to Lemma 3 the eigenvalues of this matrix can only take the values $\lambda_- = 2$ or $\lambda_+ = 2(d + 2)$. The Rayleigh quotient thus proves

\[2 \mathbf{x} \cdot \mathbf{x} \leq \mathbf{x} \cdot \hat{B} \mathbf{x} \leq 2(d + 2) \mathbf{x} \cdot \mathbf{x} \quad \text{for all } \mathbf{x} \in \mathbb{R}^{d+1} . \]

For the nodal hat functions $\hat{\varphi}_j \in P^1(\hat{T})$ associated to the $j$-th node $\hat{z}_j$ of $\hat{T}$ (i.e., $\hat{\varphi}_j(\hat{z}_k) = \delta_{jk}$) and the $j$-th unit vector $\mathbf{e}_j \in \mathbb{R}^{d+1}$, it holds $\|\hat{\varphi}_j\|^2_{L^2(\hat{T})} = \mathbf{e}_j \cdot \hat{B} \mathbf{e}_j = \hat{B}_{jj} = 4$ and hence

\[\frac{1}{2} \sum_{j=1}^{d+1} \mathbf{x}_j^2 \|\hat{\varphi}_j\|^2_{L^2(\hat{T})} = 2 \mathbf{x} \cdot \mathbf{x} \leq \left\|\sum_{j=1}^{d+1} \mathbf{x}_j \hat{\varphi}_j\right\|^2_{L^2(\hat{T})} \leq 2(d + 2) \mathbf{x} \cdot \mathbf{x} = \frac{d + 2}{2} \sum_{j=1}^{d+1} \mathbf{x}_j^2 \|\hat{\varphi}_j\|^2_{L^2(\hat{T})} . \]

For each vertex $z_i \in \mathcal{N}(\mathcal{T})$, let $\varphi_i \in S^1_0(\mathcal{T})$ denote the corresponding hat function. Let $T \in \mathcal{T}$. Since only $d + 1$ hat functions are non-trivial on $T$, a scaling argument thus proves

\[\frac{1}{2} \sum_{j=1}^{N} \mathbf{x}_j^2 \|\varphi_j\|^2_{L^2(T)} \leq \left\|\sum_{j=1}^{N} \mathbf{x}_j \varphi_j\right\|^2_{L^2(T)} \leq \frac{d + 2}{2} \sum_{j=1}^{N} \mathbf{x}_j^2 \|\varphi_j\|^2_{L^2(T)} \quad \text{for all } \mathbf{x} \in \mathbb{R}^N . \]

This proves (11) with $K_1 = 2$ and $K_2 = (d+2)/2$ and hence concludes the proof. \hfill \Box
2.3. Main results. For a subspace \( S^p_+ (T) \) of \( S^p (T) \) (e.g., \( S^p_+ (T) = S^p_0 (T) \), but further examples are found in Section 4 below), we consider the \( L^2 \)-orthogonal projection \( \Pi_+ : L^2 (\Omega) \to S^p_+ (T) \) which is uniquely defined through

\[
\int_{\Omega} (v - \Pi_+ (T) v) \, dx = 0 \quad \text{for all } v \in L^2 (\Omega) \text{ and } V \in S^p_+ (T).
\]

Since \( S^p_+ (T) \) is a discrete subspace of \( H^1 (\Omega) \), one may ask for stability of this projection with respect to the \( H^1 \)-norm. The following two theorems are the main results of this work, where the first one generalizes a corresponding result of [BY14].

**Theorem 4.** Suppose that \( T \) is a mesh which resolves the Dirichlet boundary \( \Gamma_D \). Let \( \text{level} : T \to \mathbb{N}_0 \) be an element level function and \( \mu > 1 \) and \( L \in \mathbb{N} \) with

\[
\begin{alignat}{2}
&|T|^{-\mu \text{level}(T) - d} \leq C_2 \\
&\text{diam}(T) \leq C_3 |T|^{-\mu \text{level}(T)} \\
&|\text{level}(T) - \text{level}(T')| \leq L \quad \text{for all } T, T' \in T \text{ with } T \cap T' \in F(T), \tag{15a}
\end{alignat}
\]

\[
1 < \mu L < \frac{\sqrt{K_1 K_2} + 1}{\sqrt{K_1 K_2} - 1}. \tag{15c}
\]

with arbitrary constants \( C_2, C_3 > 0 \). Then, there exists a constant \( C_4 > 0 \) which depends only on \( C_2, C_3, L, \mu, d, \) and \( p \) such that the \( L^2 \)-orthogonal projection onto \( S^p_D (T) \) is \( H^1_D \)-stable, i.e.,

\[
\| \nabla \Pi_D (T) v \|_{L^2 (\Omega)} \leq C_4 \| \nabla v \|_{L^2 (\Omega)} \quad \text{for all } v \in H^1_D (\Omega).
\]

In particular, it follows \( \| \Pi_D (T) v \|_{H^1 (\Omega)} \leq C_4 \| v \|_{H^1 (\Omega)} \) with \( C_4 = (1 + C_2)^{1/2} \), and \( \Box \) is just the special case of \( \Gamma_D = \emptyset \).

We remark that \((15a)\) implies \( \gamma \)-shape regularity \( \Box \) of \( T \) with \( \gamma = C_4 / C_2 \). For newest vertex bisection (NVB), we have the following theorem which guarantees the assumptions of Theorem \( \Box \) for lowest-order elements \( p = 1 \), but arbitrary dimension \( d \geq 2 \). We refer to Section 4 for the precise statement of NVB, but stress that NVB guarantees uniform \( \gamma \)-shape regularity \( \Box \) for all refinements \( T \) of an initial mesh \( T_0 \).

**Theorem 5.** Suppose that the mesh \( T \) is an NVB refinement of an admissible initial mesh \( T_0 \). In particular, \( T \) is \( \gamma \)-shape regular \( \Box \), where \( \gamma \) depends only on \( T_0 \). Define the constants

\[
C_2 = \min_{T \in T_0} |T|, \quad C_3 = \gamma^{1/d} \max_{T \in T_0} |T|^{1/d}, \quad L = 1, \text{ and } \quad \mu = 2^{1/d}.
\]

Then, there exists a level function \( \text{level} : T \to \mathbb{N}_0 \) such that the assumptions \((15)\) of Theorem \( \Box \) hold true for \( p = 1 \). Consequently, the \( L^2 \)-orthogonal projection onto \( S^p_D (T) \) is \( H^1_D \)-stable \( \Box \), and \( C_4 \) depends only on \( T_0 \) and \( d \).

While Theorem \( \Box \) and Theorem \( \Box \) are rigorously proved by mathematical analysis, the proofs of the following corollaries involve a computer-assisted step in the spirit of [BY14].

**Corollary 6.** Suppose that the mesh \( T \) satisfies \((15a)-(15b)\) with \( \mu = 2^{1/d} \) and \( L = 2 \), e.g., since NVB is used with an admissible initial mesh \( T_0 \). For quadratic polynomials \( p = 2 \) and dimensions \( d = 1, \ldots, 23 \), the \( L^2 \)-orthogonal projection onto \( S^p_D (T) \) is \( H^1_D \)-stable \( \Box \), and \( C_4 \) depends only on \( T_0 \) and \( d \).

The following corollary slightly improves the corresponding result of [BY14], where stability of the \( L^2 \)-orthogonal projection is only proved for polynomial degrees \( p = 1, \ldots, 12 \) in 2D.
Corollary 7. Suppose that the mesh $\mathcal{T}$ satisfies (15a)–(15b) with $\mu = 2^{1/d}$ and $L = 2$, e.g., since NVB is used with an admissible initial mesh $\mathcal{T}_0$. For fixed dimension $d = 2$ and polynomial degree $p = 1, \ldots, 20$, the $L^2$-orthogonal projection onto $S^0_D(\mathcal{T})$ is $H^1_D$-stable (16), and $C^\alpha_D$ depends only on $\mathcal{T}_0$ and $p$.

The final corollary slightly improves the corresponding result of [BY14], where stability of the $L^2$-orthogonal projection is only proved for polynomial degrees $p = 1, \ldots, 7$ in 3D.

Corollary 8. Suppose that the mesh $\mathcal{T}$ satisfies (15a)–(15b) with $\mu = 2^{1/d}$ and $L = 2$, e.g., since NVB is used with an admissible initial mesh $\mathcal{T}_0$. For fixed dimension $d = 3$ and polynomial degree $p = 1, \ldots, 8$, the $L^2$-orthogonal projection onto $S^0_D(\mathcal{T})$ is $H^1_D$-stable (16), and $C^\alpha_D$ depends only on $\mathcal{T}_0$ and $p$.

2.4. Remarks and discussion. (i) In [KPP13], an inductive implementation of 2D NVB is proposed which is well-defined even if the initial mesh $\mathcal{T}_0$ does not satisfy the usual admissibility condition (see Section 4.1). [KPP13] Prop. 6] proves that each refinement $\mathcal{T}$ of $\mathcal{T}_0$ then satisfies (15a)–(15b) with $\mu = 2^{1/2}$ and $L = 2$. For $p = 1$ and $d = 2$, it follows $\mu^2 = 2 < 3 = \sqrt{K_1K_2} + 1$. Hence, Theorem 4 proves $H^1_D$-stability (16) of $\Pi_D(\mathcal{T}) : L^2(\Omega) \to S^1_D(\mathcal{T})$ and thus gives an alternative proof of [KPP13, Thm. 3].

(ii) More generally, we observe for $p = 1$ and $\mu = 2^{1/d}$ that the level difference $L$ in (15b) may grow as $d \to \infty$. In this case, the criterion (15c) becomes

$$
\mu^L = 2^{L/d} < \frac{\sqrt{K_1K_2} + 1}{\sqrt{K_1K_2} - 1} = \frac{\sqrt{d + 2} + 1}{\sqrt{d + 2} - 1},
$$

whence

$$
L < d \log_2 \left( \frac{\sqrt{d + 2} + 1}{\sqrt{d + 2} - 1} \right) = \frac{d}{\sqrt{d + 2} - 1} \log_2 \left( \frac{2}{\sqrt{d + 2} - 1} \right) \sqrt{d + 2} - 1 \approx \sqrt{d},
$$

i.e., the upper bound on $L$ grows with the square root of the dimension $d$. Table I provides some numerical results for $S := \sqrt{d + 2} + 1$ and $\mu = 2^{1/d}$ as well as the maximal level $L > 0$ which guarantees validity of (15c) and hence stability (16) of the $L^2$-orthogonal projection onto $S^1_D(\mathcal{T})$.

(iii) Corollary 6 is at least valid for $p = 2$ and all $d = 1, \ldots, 23$ (and $L = 2$ in this range). The computer-assisted proof (by use of MAPLE) led to enormous runtimes so that we did not compute larger dimensions $d \geq 24$.

(iv) Corollary 7 is at least valid for $d = 2$ and all $p = 1, \ldots, 20$ (and $L = 2$ resp. $L = 3$ for $p \geq 3$). The computer-assisted proof (by use of MAPLE) led to enormous runtimes so that we did not compute larger polynomial degrees $p \geq 21$.

(iii) Corollary 8 is at least valid for $d = 2$ and all $p = 1, \ldots, 8$ (and $L = 2$ resp. $L = 3$ for $p \leq 7$). The computer-assisted proof (by use of MAPLE) led to enormous runtimes so that we did not compute larger polynomial degrees $p \geq 9$.

3. Proof of Theorem 4

Our proof of Theorem 4 is split into three propositions which break down the question of $H^1_D$-stability (16) of $\Pi_D(\mathcal{T})$ to certain properties of the level function $\text{level}()$. The first proposition is a criterion for $H^1$-stability of the $L^2$-projection which essentially goes back to [BY14] Thm. 4.1, Thm. 4.2]. We formulate the result in a slightly extended way by letting the grading parameter $\mu > 1$ and the level difference $L \in \mathbb{N}$ be variable, while
\[ S := \sqrt{\frac{K_1 K_2 + 1}{K_1 K_2}} \] for \( p = 1 \) and \( \mu = 2^{1/d} \) and

| \( d \) | \( S \) | \( \mu \) | \( L \) |
|---|---|---|---|
| 2 | 3.000000000000000 | 1.414213562373095 | 3 |
| 3 | 2.618033988749895 | 1.259921049894873 | 4 |
| 4 | 2.377955897113271 | 1.189207115002721 | 5 |
| 5 | 2.215250437021530 | 1.14899394709735 | 5 |
| 6 | 2.093836321356054 | 1.122462048309373 | 6 |
| 7 | 2.000000000000000 | 1.104089513673812 | 7 |
| 8 | 1.924950591148529 | 1.090507732665258 | 7 |
| 9 | 1.863324958071080 | 1.071773462536293 | 8 |
| 10 | 1.811654839115955 | 1.065041089439963 | 8 |
| 11 | 1.767591879243998 | 1.059463094359295 | 9 |
| 12 | 1.729485751811376 | 1.054766076481647 | 9 |
| 13 | 1.696140478029631 | 1.050756638653219 | 10 |
| 14 | 1.666666666666667 | 1.047294122820627 | 10 |
| 15 | 1.640388203202207 | 1.044273782427414 | 11 |
| 16 | 1.61781257308151 | 1.042373824274114 | 11 |
| 17 | 1.595433215948964 | 1.041616010650844 | 11 |
| 18 | 1.576014311052587 | 1.039752603184313 | 12 |
| 19 | 1.558257109438972 | 1.037155044446192 | 12 |
| 20 | 1.541944358078422 | 1.035264923841378 | 12 |
| 21 | 1.526893774846611 | 1.03357783007028 | 12 |
| 22 | 1.512957473875335 | 1.032008279734210 | 13 |
| 23 | 1.500000000000000 | 1.03056544752009 | 13 |
| 24 | 1.487921961587423 | 1.029302236643492 | 13 |
| 25 | 1.476627109438972 | 1.02813826656067 | 14 |

Table 1. Numerical values for lowest-order elements \( p = 1 \) and variable dimension \( d = 2, \ldots, 25 \) for \( S := \sqrt{\frac{K_1 K_2 + 1}{K_1 K_2}} \), \( \mu = 2^{1/d} \), and the maximal level \( L \in \mathbb{N} \) which ensures (15c), i.e. \( \mu^L < S \), and hence stability (16) of the \( L^2 \)-orthogonal projection onto \( S^p_D(T) \).

\( \mu = 2 \) and \( L = 1 \) in [BY14]. We will only sketch the proof for traceability, and we refer to the respective results in [BY14].

**Proposition 9.** Suppose that \( T \) is a mesh which satisfies (1), and that \( \text{level}' : T \to \mathbb{N}_0 \) and \( \mu > 1 \) are such that

\[
C_5 \mu^{-\text{level}(T)-d} \leq |T| \quad \text{and} \quad \text{diam}(T) \leq C_6 \mu^{-\text{level}(T)}
\]

for all \( T \in T \),

\[
|\text{level}'(T) - \text{level}'(T')| \leq L \quad \text{for all } T, T' \in T \text{ with } T \cap T' \neq \emptyset,
\]

\[
1 < \mu^L < \sqrt{\frac{K_1 K_2 + 1}{K_1 K_2}}.
\]

Then, there is a constant \( C_4 > 0 \) which depends only on \( C_5, C_6, L, \mu, d, \) and \( p \) such that the \( L^2 \)-orthogonal projection onto \( S^p_D(T) \) is \( H^1 \)-stable (16).

**Remark 10.** Note that assumption (18a) and \( |T|^{1/d} \leq \text{diam}(T) \) imply \( \gamma \)-shape regularity (2) of \( T \) with \( \gamma = C_5^d |C_6|^{-d} \). Moreover, we note that (18a) is slightly stronger than the corresponding assumption (15a) of Theorem 4. \( \square \)

**Remark 11.** The analysis of [BY14] assumes (18) with \( \mu = 2 \); see [BY14, eq. (1.2)]. For \( L = 1 \), [BY14, Thm. 4.2] proves stability of the \( L^2 \)-orthogonal projection \( \Pi(T) : H^1(\Omega) \to \mathbb{R} \).
\( S^1(\mathcal{T}) \) provided that \( q := \frac{\sqrt{K_1K_2} - 1}{\sqrt{K_1K_2} + 1} \leq 1/2, \) i.e., \( \mu = 2 < \frac{\sqrt{K_1K_2} + 1}{\sqrt{K_1K_2} - 1}. \) Note that \( \mu = 2^{1/d} \) for bisection-based mesh-refinement; see Section 4.

Sketch of proof of Proposition 2. Let \( u \in L^2(\Omega) \). For \( z_i \in \mathcal{N}(\mathcal{T}) \), denote by \( P_i \) the \( L^2 \)-orthogonal projection onto the space \( S^0_i(\mathcal{T}) \). With an arbitrary \( u^{(0)} \in S^0_i(\mathcal{T}) \) and the iteration \( u^{(n+1)} := u^{(n)} + \sum_{i=1}^{N} P_i(u - u^{(n)}) \), we define an approximation to \( \Pi_D(\mathcal{T})u \) via

\[
\ell := \sum_{\ell=0}^{\ell} \alpha_{\ell} u^{(\ell)}, \quad \sum_{\ell=0}^{\ell} \alpha_{\ell} = 1,
\]

where \( \alpha_{\ell} \) are appropriately scaled coefficients of the Chebyshev polynomial of degree \( \ell \).

Then, \([BY14\text{ Lem. 2.3}]\) states

\[
\|\Pi_D(\mathcal{T})u - w^{(\ell)}\|_{L^2(\Omega)} \leq \frac{2q^\ell}{1 + q^\ell} \|\Pi_D(\mathcal{T})u - w^{(0)}\|_{L^2(\Omega)},
\]

where

\[
0 < q := \frac{\sqrt{K_1K_2} - 1}{\sqrt{K_1K_2} + 1} < 1.
\]

This estimate is used to show a decay property of the \( L^2 \)-orthogonal projection. More specifically, denote by \( \Omega_i := \bigcup \{ T \in \mathcal{T} : \text{level}(T) = i \} \) the collection of elements \( T \in \mathcal{T} \) with \( \text{level}(T) = i \). With the characteristic function \( \chi_{\Omega_k} \) of \( \Omega_k \), we define \( u_k := u \chi_{\Omega_k} \).

Arguing as in \([BY14\text{ Lem. 3.1}]\), we see that \([18d]\) and \([19]\) imply

\[
\|\Pi_D(\mathcal{T})u_k\|_{L^2(\Omega_i)} \leq \min \{ 1, 2q^{(|i-k|-1)/L} \} \|u_k\|_{L^2(\Omega)}.
\]

As in \([BY14\text{ Thm. 4.1}]\), it follows from \([18c]\) and hence \( \mu^k q < 1 \) that

\[
\sum_{i=1}^{\infty} \mu^i \|\Pi_D(\mathcal{T})u\|_{L^2(\Omega_i)}^2 \lesssim \sum_{i=1}^{\infty} \mu^i \|u\|_{L^2(\Omega_i)}^2;
\]

some details are sketched in the proof of Theorem \([15]\) below. According to \([18d]\) and \( |T|^{1/d} \leq \text{diam}(T) \), it holds \( \mu^i \simeq \text{diam}(T) \) for all \( T \in \mathcal{T} \) with \( T \subseteq \Omega_i \). Therefore, \( [21] \) shows stability of \( \Pi_D(\mathcal{T}) \) in the weighted \( L^2 \)-norm

\[
\|h^{-1}\Pi_D(\mathcal{T})u\|_{L^2(\Omega)} \lesssim \|h^{-1}u\|_{L^2(\Omega)},
\]

where the hidden constant depends only on \( C_5 \), \( C_6 \), and \( \mu^k q < 1 \). Note that \( \Pi_D(\mathcal{T}) = \Pi_D(\mathcal{T})^+ + \Pi_D(\mathcal{T})^0 \), where \( \Pi_D(\mathcal{T})^0 \) is the \( L^2 \)-orthogonal projection onto \( S^0(\mathcal{T})^0 \). Since functions in \( S^0(\mathcal{T})^0 \) are only supported on one single element, it holds

\[
\|h^{-1}\Pi_D(\mathcal{T})^0u\|_{L^2(\Omega)} \leq \|h^{-1}u\|_{L^2(\Omega)}.
\]

Combining the last two estimates, we infer

\[
\|h^{-1}\Pi_D(\mathcal{T})u\|_{L^2(\Omega)} \lesssim \|h^{-1}u\|_{L^2(\Omega)} \quad \text{for all } u \in L^2(\Omega),
\]

while the hidden constant depends only on \( C_5 \), \( C_6 \), and \( \mu^k q < 1 \). From estimate \([22]\), \( H^1 \)-stability of \( \Pi_D(\mathcal{T}) \) follows with standard arguments, cf. \([BY14\text{ Thm. 4.2}]\): Let \( J(\mathcal{T}) : H^1(\Omega) \to S^p(\mathcal{T}) \) denote the Scott-Zhang projection \([SZ90]\) and recall that

\[
\|h^{\alpha}(1 - J(\mathcal{T}))v\|_{L^2(\Omega)} + \|h^{1+\alpha}\nabla J(\mathcal{T})v\|_{L^2(\Omega)} \lesssim \|h^{1+\alpha}\nabla v\|_{L^2(\Omega)}
\]

for all \( v \in H^1(\Omega) \) and \( \alpha \in \mathbb{R} \) as well as \( J(\mathcal{T})v = v \) for all \( v \in S^p(\mathcal{T}) \). The hidden constant depends only on \( \gamma \)-shape regularity \([3]\) of \( \mathcal{T} \) and on the polynomial degree \( p \) and hence only on \( C_5 \), \( C_6 \), \( d \), and \( p \). Moreover, \( J(\mathcal{T}) \) can be chosen such that \( J(\mathcal{T})v \in S^p(\mathcal{T}) \) for all
\[ v \in H_D^1(\mathcal{T}); \text{ see e.g. } \text{[AFK+13].} \] Let \( v \in H_D^1(\Omega) \). With an inverse estimate and \( \alpha = -1 \), we therefore obtain
\[
\| \nabla \Pi_D(\mathcal{T}) v \|_{L^2(\Omega)} \leq \| \nabla (\Pi_D(\mathcal{T}) - J(\mathcal{T})) v \|_{L^2(\Omega)} + \| \nabla J(\mathcal{T}) v \|_{L^2(\Omega)} \]
\[
\lesssim \| h^{-1}(\Pi_D(\mathcal{T}) - J(\mathcal{T})) v \|_{L^2(\Omega)} + \| \nabla v \|_{L^2(\Omega)}.
\]
With \([22]\) and the projection property \( \Pi_D(\mathcal{T}) J(\mathcal{T}) v = J(\mathcal{T}) v \in S_D^p(\mathcal{T}) \), we conclude
\[
\| h^{-1}(\Pi_D(\mathcal{T}) - J(\mathcal{T})) v \|_{L^2(\Omega)} = \| h^{-1}\Pi_D(\mathcal{T})(1 - J(\mathcal{T})) v \|_{L^2(\Omega)} \lesssim \| \nabla v \|_{L^2(\Omega)}.
\]
This proves \( H_D^1 \)-stability \([11]\) of \( S_D^p(\mathcal{T}) \), and the overall constant \( C_{11} > 0 \) depends only on \( C_7, C_8, L, \mu, d, \) and \( p \). \( \square \)

For each node \( z_\ell \in \mathcal{N}(\mathcal{T}) \), let \( h_\ell > 0 \) denote some positive scalar (to be fixed later) which behaves like the local element-size, i.e., \( h_\ell \simeq \text{diam}(T) \) for all \( T \in \mathcal{T} \) with \( z_\ell \in T \).

The next proposition provides a stability criterion in terms of these nodal values \( h_\ell \). For 2D, a similar result is first found in \([\text{Car04}]\) for (a slightly modified) red-green-blue refinement and \( p = 1 \) and adapted in \([\text{KPP13}]\) to NVB in 2D.

**Proposition 12.** Assume that \( \mathcal{T} \) is a \( \gamma \)-shape regular mesh which satisfies \([4]\). Let \( \mu > 1 \) and \( L \in \mathbb{N} \). For all \( T \in \mathcal{T} \) and all \( z_j, z_k \in \mathcal{N}(T) \), we suppose that the chosen scalars \( h_j, h_k > 0 \) satisfy
\[
\frac{h_j}{h_k} \leq \mu^L \leq \frac{\sqrt{K_1K_2} + 1}{\sqrt{K_1K_2} - 1}.
\]
as well as
\[
C_7 \text{diam}(T) \leq h_j \leq C_8 \text{diam}(T)
\]
with constants \( C_7, C_8 > 0 \). Then, there exists a level function \( \text{level}^1 : \mathcal{T} \to \mathbb{N}_0 \) such that the assumptions \([13]\) of Proposition \([5]\) are satisfied with \( C_9 = \frac{C_7}{\gamma} \) and \( C_{10} = C_8^{-1} \mu \). In particular, the \( L^2 \)-orthogonal projection \( \Pi_D(\mathcal{T}) \) onto \( S_D^p(\mathcal{T}) \) is \( H_D^1 \)-stable \([10]\), and the constant \( C_{11} > 0 \) depends only on \( C_7, C_8, L, \mu, \gamma, d, \) and \( p \).

**Proof.** Without loss of generality, we may assume \( 0 < h_j < 1 \) for all \( j = 1, \ldots, N \) by multiplicative scaling. Recall \( \mu > 1 \). For each node \( z_\ell \in \mathcal{N}(\mathcal{T}) \), we fix \( \ell_\ell \in \mathbb{N}_0 \) such that
\[
\mu^{-\ell_\ell} \leq h_j < \mu^{-\ell_\ell + 1}.
\]
Let \( T \in \mathcal{T} \) and \( z_j, z_k \in \mathcal{N}(T) \). From \( \mu > 1 \) and
\[
\mu^{\ell_k - \ell_j - 1} = \frac{\mu^{-\ell_j}}{\mu^{-\ell_k + 1}} \leq \frac{h_j}{h_k} \leq \mu^L,
\]
we get \( \ell_k - \ell_j < L + 1 \) and hence \( \ell_k - \ell_j \leq L \) from \( \ell_j, \ell_k, L \in \mathbb{N} \). Symmetry of the argument thus yields
\[
|\ell_j - \ell_k| \leq L \quad \text{for all } T \in \mathcal{T} \text{ and } z_j, z_k \in \mathcal{N}(T).
\]
Define
\[
\text{level}^1(T) := \min_{z_\ell \in \mathcal{N}(T)} \ell_\ell
\]
To see \([13B]\), let \( T, T' \in \mathcal{T} \) with \( T \cap T' \neq \emptyset \). Since \( \mathcal{T} \) is conforming, there exists \( z_n \in \mathcal{N}(T) \cap \mathcal{N}(T') \). Let \( z_j \in \mathcal{N}(T) \) and \( z_k \in \mathcal{N}(T') \) satisfy \( \text{level}^1(T) = \ell_j \) resp. \( \text{level}^1(T') = \ell_k \). Due to \([25]\), it holds \( \ell_j \leq \ell_n \leq \ell_j + L \) as well as \( \ell_k \leq \ell_n \leq \ell_k + L \). This implies \( \ell_j \leq \ell_n \leq \ell_k + L \) and \( \ell_k \leq \ell_n \leq \ell_j + L \) and hence results in
\[
|\text{level}^1(T) - \text{level}^1(T')| = |\ell_j - \ell_k| \leq L.
\]
To prove (18a), let $z_j \in N(T)$ with $\ell_j = \text{level}'(T)$. Then,

$$
\mu^{-\text{level}'(T)d} \leq h_j^d \leq C_5 \delta(T)^d \leq C_5 \gamma |T|
$$

and

$$
C_4 \delta(T) \leq h_j \mu^{-\ell_j+1} = \mu \mu^{-\text{level}'(T)}.
$$

This proves (18a) with $C_5 = C_d^{-d}/\gamma$ and $C_6 = C_d^{-1}\mu$. The upper bound (18a) on $\mu$ holds by assumption (15c). Altogether, we have thus verified (18). Consequently, Proposition 9 applies and completes the proof. □

In view of Proposition 12, it only remains to define the nodal values $h_j > 0$ for the nodes $z_j \in N(T)$ of $T$. The following proposition completes the proof of Theorem 4.

**Proposition 13.** Let $T$ be a mesh which fulfills the assumptions (4) and (15) of Theorem 4. On the set $\mathcal{N}(T) = \{z_1, \ldots, z_N\}$ of nodes, we define the following nodal distance $\delta(\cdot, \cdot) \in \mathbb{N}_0$

$$
\delta(z_j, z_k) := \begin{cases} 
0 & \text{for } z_j = z_k, \\
1 & \text{if there exists } T \in \mathcal{T} \text{ with } z_j, z_k \in T, \\
\infty & \text{for the minimal number } n \in \mathbb{N} \text{ of elements } T_1, \ldots, T_n \in \mathcal{T} \text{ such that } z_j \in T_1, z_k \in T_n, \text{ and } T_i \cap T_{i+1} \in \mathcal{F}(T) \text{ for all } i = 1, \ldots, n-1.
\end{cases}
$$

For $z_j \in \mathcal{N}(T)$, we let

$$
(26) \quad h_j := \min_{T_j' \in T} \min_{z_j' \in \mathcal{N}(T_j')} \mu^{L_d(z_j, z_j')-\text{level}(T_j')}
$$

Then, $T$ is $\gamma$-shape regular (2) with $\gamma = C_d / C_2$ and the assumptions (23) of Proposition 12 are satisfied with $C_5 = C_d^{-1/d}$ and some constant $C_7 > 0$ which depends only on $C_3, L, \mu$, $\gamma$-shape regularity of $T$, and the dimension $d \geq 2$. In particular, the $L^2$-orthogonal projection $\Pi_D(T)$ onto $S^d_0(T)$ is $H^1_D$-stable (16), and the constant $C_4 > 0$ depends only on $C_2, C_3, L, \mu, d, p$.

**Proof.**

**Step 1.** We validate the lower estimate in (23a): Let $T \in \mathcal{T}$ and $z_j, z_k \in \mathcal{N}(T)$. Choose $T_k' \in T$ and $z_k' \in \mathcal{N}(T_k')$ which attain the minima in the definition of $h_k = \mu^{L_d(z_k, z_k')-\text{level}(T_k')}$. Together with $|\delta(z_j, z_k') - \delta(z_k, z_k')| \leq 1$ and $\mu > 1$, we obtain

$$
\frac{h_j}{h_k} \leq \frac{\mu^{L_d(z_j, z_k')-\text{level}(T_j')}}{\mu^{L_d(z_k, z_k')-\text{level}(T_k')}} = \mu^{L_d(z_j, z_k')-\text{level}(T_j')} \leq \mu^L.
$$

The upper bound in (23a) holds by assumption (15c).

**Step 2.** We validate the upper bound in (23b) with $C_8 = C_2^{-1/d}$: Let $T \in \mathcal{T}$ and $z_j \in \mathcal{N}(T)$. By definition of $h_j$, we see

$$
h_j \leq \mu^{-\text{level}(T)} \leq C_2^{-1/d}|T|^{1/d} \leq C_2^{-1/d} \delta(T)
$$

**Step 3.** It only remains to verify the lower bound in (23b): Let $T \in \mathcal{T}$ and $z_j \in \mathcal{N}(T)$. Choose $T_j' \in T$ and $z_j' \in \mathcal{N}(T_j')$ which attain the minima in the definition of $h_j = \mu^{L_d(z_j, z_j')-\text{level}(T_j')}$. If $\text{level}(T_j') < \text{level}(T)$, we would see

$$
h_j \leq \mu^{-\text{level}(T)} < \mu^{-\text{level}(T_j')} \leq \mu^{L_d(z_j, z_j')-\text{level}(T_j')} = h_j.
$$
Consequently, it holds $\text{level}(T) \leq \text{level}(T'_j)$. We obtain

$$h_j = \mu^L \delta(z_j, z'_j) - \text{level}(T'_j) = \mu^{-\text{level}(T)} \mu^L \delta(z_j, z'_j) - |\text{level}(T'_j) - \text{level}(T)|$$

$$\geq C^\frac{1}{3} \text{diam}(T) \mu^L \delta(z_j, z'_j) - |\text{level}(T'_j) - \text{level}(T)|.$$  

For $\delta(z_j, z'_j) = n$, there exist elements $T_1, \ldots, T_n \in T$ with $z_j \in T_1$, $z'_j \in T_n$ and $T_i \cap T_{i+1} \in \mathcal{F}(T)$ for all $i = 1, \ldots, n - 1$. Due to $\gamma$-shape regularity of $T$, the number of elements in the node patches of $z_j$ resp. $z'_j$ is uniformly bounded by some $\gamma$-dependent constant $C_9 > 0$. Hence, there exists a sequence of elements $\tilde{T}_1, \ldots, \tilde{T}_m \in T$ with $m \leq n + 2 C_9$ such that $\tilde{T}_1 = T$, $\tilde{T}_m = T'_j$, and $\tilde{T}_i \cap \tilde{T}_{i+1} \in \mathcal{F}(T)$ for all $i = 1, \ldots, m - 1$. By assumption (15b), it holds $|\text{level}(\tilde{T}_i) - \text{level}(\tilde{T}_{i+1})| \leq L$. The triangle inequality yields

$$|\text{level}(T'_j) - \text{level}(T)| \leq \sum_{i=1}^{m-1} |\text{level}(\tilde{T}_i) - \text{level}(\tilde{T}_{i+1})| \leq L(m - 1) \leq L((n - 1) + 2 C_9).$$

This yields

$$\mu^L \delta(z_j, z'_j) - |\text{level}(T'_j) - \text{level}(T)| = L n - |\text{level}(T'_j) - \text{level}(T)| \geq L(1 - 2 C_9)$$

and hence

$$h_j \geq C^\frac{1}{3} \mu L^{1-2} \text{diam}(T).$$

Altogether, we thus see the lower bound in (23b) with $C^\frac{1}{3} = C^\frac{1}{3} \mu L^{1-2} > 0$. □

4. Bisection of Simplicial Meshes and Proof of Theorem 5

The bisection of a simplicial mesh of $d$-simplices in $\mathbb{R}^D$ can be done in different ways, e.g. [Sew72] for $d = 2$, [Kos94] for $d = 3$, and [Mau95, Tra97] for $d \geq 3$. We mainly follow the presentation in [Ste08]: Each simplex $T \in T_0$ of the initial mesh is identified with an ordered sequence of its vertices and associated with the type $\gamma = 0$, i.e.,

$$T = (x_0, \ldots, x_d)_{\gamma}.$$

The edge between $x_0$ and $x_d$ is the so-called refinement edge of $T$, which is denoted by $e(T)$ in the following. Bisection of a simplex $T = (x_0, \ldots, x_d)_{\gamma}$ of type $\gamma \in \{0, \ldots, d - 1\}$ provides the sons

$$(28a) \quad T' = (x_0, \frac{x_0 + x_d}{2}, x_1, \ldots, x_\gamma, x_{\gamma+1}, \ldots, x_{d-1})_{(\gamma+1) \mod d}$$

and

$$(28b) \quad T'' = (x_d, \frac{x_0 + x_d}{2}, x_1, \ldots, x_\gamma, x_{d-1}, \ldots, x_{\gamma+1})_{(\gamma+1) \mod d}$$

where the sequences $(x_{\gamma+1}, \ldots, x_{d-1})$ and $(x_1, \ldots, x_\gamma)$ are void for $\gamma = d - 1$ resp. $\gamma = 0$. It holds $|T' | = |T|/2 = |T''|$ with the $d$-dimensional measure $|\cdot|$. Overall, the choice of the refinement edges for all $T \in T_0$ determines the refinement strategy. Bisection is a binary refinement rule and thus gives rise to some level function: For a coarse-mesh simplex $T \in T_0$, we define $\text{level}_{\text{nvb}}(T) = 0$. If $T$ is bisected into two sons $T', T''$, we define $\text{level}_{\text{nvb}}(T') = \text{level}_{\text{nvb}}(T) + 1 = \text{level}_{\text{nvb}}(T'')$. In this case, it also holds $|T| = |T' |/2 = |T''|$ and hence $\text{level}_{\text{nvb}}(T) = \log_2(|\widehat{T}|/|T|)$ if $\widehat{T} \in T_0$ is the unique coarse-mesh element with $T \subseteq \widehat{T}$.

Note that bisection does not lead to conforming partitions in general. By recursive refinement, called newest vertex bisection (NVB) in the following, it is usually guaranteed that the refined partition is in fact conforming and hence a mesh. To ensure
that this recursion terminates, one requires properties on the initial mesh \( \mathcal{T}_0 \). Following Stevenson [Ste08], we call the initial mesh \( \mathcal{T}_0 \) admissible if

- \( \mathcal{T}_0 \) is conforming,
- for all \( T, T' \in \mathcal{T}_0 \) which share a \((d-1)\)-dimensional face \( T \cap T' \), one of the following assertions holds:
  (a) \( \varepsilon(T) = \varepsilon(T') \subset T \cap T' \) and \( T, T' \) are reflected neighbors
  (b) there exist son simplices \( t \subset T \) and \( t' \subset T' \) with \( T \cap T' = t \cap t' \) which are reflected neighbors.

Here, reflected neighbors means that the ordered vertices of \( T' \) coincide with either \( T := (x_0, \ldots, x_d)_\gamma \) or \( T_R := (x_0, x_1, \ldots, x_{\gamma-1}, x_{\gamma+1}, \ldots, x_d)_\gamma \) on all but one position, and we note that \( T \) and \( T_R \) are equivalent in the sense that bisection leads to the very same two sons [28]. Due to [Ste08, Thm. 4.3], admissibility of \( \mathcal{T}_0 \) ensures that, for all \( n \in \mathbb{N} \), each partition \( \mathcal{T} \) obtained from bisection of \( \mathcal{T}_0 \) with level_{nvb}(T) = n for all \( T \in \mathcal{T} \) is already conforming. Moreover, the admissibility condition is not only sufficient, but also necessary to ensure this. Finally, admissibility of \( \mathcal{T}_0 \) guarantees that the recursive implementation of NVB terminates.

Moreover, NVB of \( \mathcal{T} \) implicitly leads to NVB of \( \mathcal{T}_D = \{ F \in \mathcal{F} \mathcal{T} : F \subset \Gamma_D \} \) if \( \Gamma_D \) is resolved [14]. Such an observation is for instance used for adaptive finite element methods, where the given Dirichlet data are discretized by means of the \( L^2(\Gamma_D) \)-orthogonal projection onto the discrete trace space [AFK+13], and stability of the latter is required in \( H^{1/2}(\Gamma_D) \).

The following lemma has first been proved for \( d = 3 \) in [Kos94, Lem. 3] and (implicitly) generalized for \( d \geq 2 \) in [Ste08, Cor. 4.6].

**Lemma 14.** Suppose that the mesh \( \mathcal{T} \) is an NVB refinement of an admissible initial mesh \( \mathcal{T}_0 \). Then, the natural level function level = level_{nvb} satisfies \((15b)\) with \( L = 1 \).

**Sketch of proof.** Stevenson [Ste08, Cor. 4.6] proves that for \( \varepsilon(T) \subset F \) it holds either \( \text{level}(T) = \text{level}(T') \) with \( \varepsilon(T) = \varepsilon(T') \) or \( \text{level}(T) = \text{level}(T') + 1 \) with \( \varepsilon(T) = \varepsilon(T') \) for one of the two children \( t' \) of \( T' \). The claim \((15b)\) is the first step of his proof. \( \square \)

**Proof of Theorem 2.** The proof will be concluded by application of Theorem 2, hence we will check \((15a)-(15c)\). First, Lemma 14 shows \((15b)\). The estimates \((15a)\) can be shown as follows. For each \( T \in \mathcal{T} \), let \( \hat{T} \in \mathcal{T}_0 \) be the unique ancestor with \( T \subset \hat{T} \). By definition of NVB and \( \mu := 2^{1/d} \), it holds

\[ |T| = |\hat{T}| 2^{-\text{level}(T)} = |\hat{T}| \mu^{-\text{level}(T)d}. \]

This proves

\[ C_2 \mu^{-\text{level}(T)d} \leq |T| \quad \text{with} \quad C_2 := \min_{T_0 \in \mathcal{T}_0} |T_0|. \]

Since NVB only leads to finitely many shapes of simplices, all NVB generated meshes are uniformly \( \gamma \)-shape regular [2], where \( 0 < \gamma < \infty \) depends only on the initial mesh \( \mathcal{T}_0 \). This proves

\[ \text{diam}(T)^d \leq \gamma |T| \leq C_d \mu^{-\text{level}(T)d} \quad \text{with} \quad C_d := \gamma^{1/d} \max_{T_0 \in \mathcal{T}_0} |T_0|^{1/d}. \]

This proves \((15a)\), and it remains to show \((15c)\). According to Proposition 2, it holds \( K_1 = 2 \) and \( K_2 = (d + 2)/2 \) for the lowest-order case \( p = 1 \). We consider the scalar function \( f(t) := 1 + 4/t - 4^{1/t} \). We prove that \( f(t) > 0 \) for \( t \geq 3 \). Since \( \lim_{t \to \infty} f(t) = 0 \),
it remains to show that $f(\cdot)$ is strictly decreasing on the interval $[3, \infty)$. This follows from

$$f'(t) = -\frac{4}{t^2} + \frac{\ln 4}{t^2} 4^{1/t} = \frac{1}{t^2}(4^{1/t} \ln 4 - 4) \leq \frac{1}{t^2}(4^{1/3} \ln 4 - 4) < 0 \quad \text{for } t \geq 3.$$  

With $f(2) = 1 > 0$, we infer that $f(d) > 0$ for all dimensions $d \in \mathbb{N}$ with $d \geq 2$. Hence,

$$\mu^2 = 4^{1/d} < 4^{1/d} + f(d) = \frac{d + 4}{d} < \frac{d + 2 + 2\sqrt{d + 2} + 1}{d + 2 - 2\sqrt{d + 2} + 1} = \left(\frac{\sqrt{K_1K_2 + 1}}{\sqrt{K_1K_2 - 1}}\right)^2$$

which shows (15c). Here, the second estimate (marked with a !-symbol) follows from elementary calculations. □

5. Computer-Assisted Proof of Corollary 6, 8

As in [BY14], we rely on a computer-assisted proof to show Corollaries 6, 8. For different dimensions $d \geq 2$ and polynomial degrees $p \geq 1$, we will numerically compute the optimal constants $K_1$, $K_2$ satisfying (11a) and (11b) in Lemma 1. Then, we will utilize Theorem 4 to check if the assumptions (13) hold true with $\mu = 2^{1/d}$ and some $L \in \mathbb{N}$.

First, note that Lemma 1 allows to restrict the considerations to one single reference element, e.g., $T = \text{conv}\{0, e_1, \ldots, e_d\}$, where $e_j$ is the $j$-th standard unit vector. Then, the computation of $K_1$ and $K_2$ corresponds to two generalized eigenvalue problems

$$(29) \quad A_{1,2}x = \lambda_{1,2}M_{1,2}x.$$  

Here, $A_{1,2}$ is a positive semi-definite and symmetric matrix, and $M_{1,2}$ is even positive definite and symmetric. Hence, the constants $K_{1,2} > 0$ from (11) turn out to be the respective maximal eigenvalues $\lambda_{1,2}$ of (29). We compute these maximal eigenvalues numerically in MAPLE with 40-digit floating point precision.

It remains to specify the decomposition of $v \in S^p(T)^\perp$ in (11a). To do so, denote by $\varphi_i \in S^1(T)$ the hat function corresponding to the vertex $z_i \in \mathcal{N}(T) = \{z_1, \ldots, z_N\}$. Choose $v_i = \mathcal{L}(v\varphi_i)$, where $\mathcal{L}$ is some linear interpolation operator on $S^p(T)^\perp$. As $v\varphi_i$ vanishes outside the vertex patch $\omega_i$, it holds $\mathcal{L}(v\varphi_i) \in S^p(T)^\perp$, and as $\sum_{i=2}^N \varphi_i = 1$, it holds $\sum_{i=1}^N v_i = v$. Note that this uniquely identifies a function in $S^p(T)^\perp$.

We will employ two different interpolation operators: The first one, $\mathcal{L} = \mathcal{L}^\text{unif}$ will be employed for the proof of Corollaries 6 and 8. It is chosen as the standard interpolation operator in the usual degrees of freedom of $S^p(T)$ which are uniformly distributed on the boundaries of the elements $T \in \mathcal{T}$. The second interpolation operator $\mathcal{L} = \mathcal{L}^\text{gauss}$ will be employed only for the proof of Corollary 7; hence we need to define it only for $d = 2$. It is chosen to interpolate in all vertices of $T$ as well as in $p - 1$ points on every edge; these points are chosen to be the first $p - 1$ Gauss points on $[-1, 1]$, mapped accordingly to all edges.

**Proof of Corollary 6.** For fixed $p = 2$ and variable dimension $d$, Table 2 provides our numerical results for $K_1$, $K_2$, $S := \frac{\sqrt{K_1K_2 + 1}}{\sqrt{K_1K_2 - 1}}$, $\mu = 2^{1/d}$ as well as the maximal level $L > 0$ which guarantees validity of (15c). At least for $d = 2, \ldots, 23$, we hence prove stability (1) of the $L^2$-orthogonal projection onto $S^p_0(T)$ (even with $L = 2$).

**Proof of Corollary 7.** For fixed $d = 2$ and variable polynomial degree $p$, Table 3 provides our numerical results for $K_1$, $K_2$, and $S := \frac{\sqrt{K_1K_2 + 1}}{\sqrt{K_1K_2 - 1}}$. With $\mu = 21/2$, this guarantees validity of (15c) and hence stability (1) of the $L^2$-orthogonal projection onto $S^p_0(T)$ for all shown values of $p = 1, \ldots, 20$ and $L = 2$ (resp. $L = 3$ for $p \geq 3$). □
The Sections 6.2–6.3 generalize works of Steinbach \cite{Ste01,Ste02} on the

This section briefly collects some additional observations and generalizations of Theo-

H \leq L \text{ weighted validity of (15c) and hence stability (1) of the}

Proof of Corollary 5. For fixed degree 3 and variable polynomial degree \( p \), Table 2 provides our numerical results for \( K_1, K_2 \), and \( S := \frac{\sqrt{K_1K_2+1}}{\sqrt{K_1K_2-1}} \). With \( \mu = 2^{1/3} \), this guarantees validity of (15c) and hence stability (1) of the \( L^2 \)-orthogonal projection onto \( S^2_0(\mathcal{T}) \) for all shown values of \( p = 1, \ldots, 8 \) and \( L = 2 \) (resp. \( L = 3 \) for \( p \leq 7 \)).

6. Extensions

This section briefly collects some additional observations and generalizations of Theorem 4. The Sections 6.2–6.3 generalize works of Steinbach \cite{Ste01, Ste02} on the \( L^2 \)-orthogonal projection in fractional-order Sobolev spaces \( H^{s}(\Omega) \) and \( \tilde{H}^{s}(\Omega) \) for \( 0 < s \leq 1 \). In Section 6.4, we prove that stability in \( H^s(\Omega) \) and \( \tilde{H}^{-s}(\Omega) \) is preserved if the \( L^2 \)-orthogonal projection onto \( S^p(\mathcal{T}) \)-functions with integral mean zero is used. Section 6.5 shows that stability of the \( L^2 \)-orthogonal projection is also guaranteed for locally weighted \( L^2 \)-norms. The main analytical tool is interpolation between Hilbert spaces which is briefly recalled in the following subsection.

6.1. Preliminaries on interpolation spaces. Let \( X_0 \) and \( X_1 \) be Hilbert spaces with \( X_0 \supseteq X_1 \) and continuous inclusion, i.e., there exists some constant \( C > 0 \) such that \( \|x\|_{X_0} \leq C \|x\|_{X_1} \) for all \( x \in X_1 \). Interpolation theory, e.g. \cite{BL76,Tar07}, provides a
means to define intermediate normed spaces

(30) \[ X_1 \subseteq X_s := [X_0; X_1], s \subseteq X_0 \quad \text{for all } 0 < s < 1, \]
where \([\cdot;\cdot]_s\) denotes the interpolation operator of, e.g., the real \(K\)-method. The norm related to the intermediate interpolation space \(X_s\) satisfies

\[
\|x\|_{X_s} \leq \|x\|_{X_{0,s}}^{1-s} \|x\|_{X_1}^s \quad \text{for all } x \in X_1.
\]

We shall use the so-called interpolation estimate: Let \(X_0 \supseteq X_1\) and \(Y_0 \supseteq Y_1\) be Hilbert spaces with continuous inclusions. Let \(\Pi : X_0 \to Y_0\) be a linear operator with \(\Pi(X_1) \subseteq Y_1\).

Assume that \(\Pi : X_0 \to Y_0\) as well as \(\Pi : X_1 \to Y_1\) are continuous, i.e.,

\[
\|\Pi x\|_{Y_0} \leq c_0 \|x\|_{X_0} \quad \text{for all } x \in X_0,
\]

\[
\|\Pi x\|_{Y_1} \leq c_1 \|x\|_{X_1} \quad \text{for all } x \in X_1,
\]

with respective operator norms \(c_0, c_1 > 0\). Let \(0 < s < 1\) and \(X_s = [X_0; X_1]_s\) and \(Y_s = [Y_0; Y_1]_s\). Then, \(\Pi : X_s \to Y_s\) is a well-defined, linear, and continuous operator with

\[
\|\Pi x\|_{Y_s} \leq c_0^{1-s} c_1^s \|x\|_{X_s} \quad \text{for all } x \in X_s.
\]

Note that for other interpolation methods than the real \(K\)-method, the previous estimates \(31\) and \(33\) hold only up to some additional generic constant.

6.2. \(L^2\)-orthogonal projection on positive-order Sobolev spaces \(H^*_*(\Omega)\) and \(\tilde{H}^*_*(\Omega)\).

For \(0 < s < 1\), the Sobolev spaces of fractional order can be defined by interpolation \([\text{McL00}]\)

\[
H^*_*(\Omega) = [L^2(\Omega); H^1(\Omega)]_s \quad \text{and} \quad \tilde{H}^*_*(\Omega) = [L^2(\Omega); H^1_0(\Omega)]_s.
\]

Here, \(H^1_0(\Omega) = \{v \in H^1(\Omega) : v|_{\partial\Omega} = 0\}\) denotes the space of \(H^1\)-functions with zero trace.

For \(\Gamma_D = \emptyset\), Theorem \(\text{[\text{McL00}]}\) states that the \(L^2\)-orthogonal projection \(\Pi(\mathcal{T}) : L^2(\Omega) \to S^p(\mathcal{T})\) is \(H^1\)-stable. With \(c_0 = 1\) and \(c_1 = C_0\) the interpolation estimate \(33\) proves \(H^*_*(\Omega)\)-stability

\[
\|\Pi(\mathcal{T})v\|_{H^*_*(\Omega)} \leq C_0 \|v\|_{H^*_*(\Omega)} \quad \text{for all } v \in H^*_*(\Omega).
\]

For \(\Gamma_D = \partial\Omega\), Theorem \(\text{[\text{McL00}]}\) states that the \(L^2\)-orthogonal projection \(\Pi_0(\mathcal{T}) : L^2(\Omega) \to S^0(\mathcal{T})\) onto \(S^0(\mathcal{T}) := S^p(\mathcal{T}) \cap H^1_0(\Omega)\) is \(H^1_0\)-stable. With \(c_0 = 1\) and \(c_1 = C_0\) the interpolation estimate \(33\) proves \(\tilde{H}^*_*(\Omega)\)-stability

\[
\|\Pi_0(\mathcal{T})v\|_{\tilde{H}^*_*(\Omega)} \leq C_0 \|v\|_{\tilde{H}^*_*(\Omega)} \quad \text{for all } v \in \tilde{H}^*_*(\Omega).
\]

6.3. \(L^2\)-orthogonal projection on negative-order Sobolev spaces \(H^{-\ast}(\Omega)\) and \(\tilde{H}^{-\ast}(\Omega)\).

For \(0 < s \leq 1\), the Sobolev spaces of negative fractional order are defined as dual spaces

\[
H^{-s}(\Omega) = \tilde{H}^*_*(\Omega)^* \quad \text{and} \quad \tilde{H}^{-s}(\Omega) = H^*_*(\Omega)^*.
\]

where duality is understood with respect to the extended \(L^2\)-scalar product \([\text{McL00}]\). It is known that this implies that \(L^2(\Omega)\) is a dense subspace of both \(H^{-s}(\Omega)\) or \(\tilde{H}^{-s}(\Omega)\) with respect to the corresponding dual norms

\[
\|\psi\|_{H^{-s}(\Omega)} = \sup_{v \in H^s(\Omega) \setminus \{0\}} \frac{\langle \psi; v \rangle}{\|v\|_{H^s(\Omega)}} \quad \text{resp.} \quad \|\psi\|_{\tilde{H}^{-s}(\Omega)} = \sup_{v \in H^s(\Omega) \setminus \{0\}} \frac{\langle \psi; v \rangle}{\|v\|_{H^s(\Omega)}}.
\]

We adopt the notation from the previous subsection and let \(\psi \in L^2(\Omega) \subset \tilde{H}^{-s}(\Omega)\) and \(v \in H^s(\Omega)\). By definition of duality and the dual norm, we obtain

\[
\langle \Pi(\mathcal{T})\psi; v \rangle = \int_\Omega v \Pi(\mathcal{T})\psi \, dx = \int_\Omega \psi \Pi(\mathcal{T})v \, dx \leq \|\psi\|_{\tilde{H}^{-s}(\Omega)} \|\Pi(\mathcal{T})v\|_{H^s(\Omega)}
\]

16
Together with the $H^s$-stability (35) of $\Pi(T)$, this yields
\[ \|\Pi(T)\psi\|_{H^{-s}(\Omega)} \leq \|\psi\|_{H^{-s}(\Omega)} \sup_{v \in H^{s}(\Omega) \setminus \{0\}} \frac{\|\Pi(T)v\|_{H^{s}(\Omega)}}{\|v\|_{H^{s}(\Omega)}} \leq C_s\|\psi\|_{H^{-s}(\Omega)} \text{ for all } \psi \in L^2(\Omega), \]
i.e., $\Pi(T)$ is $H^{-s}$-stable for $L^2$-functions. Since $L^2(\Omega)$ is a dense subspace of $H^{-s}(\Omega)$, functional analysis guarantees a unique linear and continuous extension $\Pi(T) : H^{-s}(\Omega) \to S^p(T)$ which even has the same operator norm, i.e.,
\[ \|\Pi(T)\psi\|_{H^{-s}(\Omega)} \leq C_s\|\psi\|_{H^{-s}(\Omega)} \text{ for all } \psi \in H^{-s}(\Omega). \]

Arguing along the same lines for $\Gamma_D = \partial \Omega$, one sees that the $L^2$-orthogonal projection $\Pi_0(T)$ onto $S^p_0(T)$ admits a unique extension to a linear operator on $H^{-s}(\Omega)$ such that
\[ \|\Pi_0(T)\psi\|_{H^{-s}(\Omega)} \leq C_s\|\psi\|_{H^{-s}(\Omega)} \text{ for all } \psi \in H^{-s}(\Omega). \]

6.4. $L^2$-orthogonal projection onto $S^p(T)$ with zero integral mean. Let $L^2_0(\Omega) = \{v \in L^2(\Omega) : \int_\Omega v \, dx = 0\}$ as well as $H^1_0(\Omega) = H^1(\Omega) \cap L^2_0(\Omega)$ and $S^p_0(T) = S^p(T) \cap L^2_0(\Omega)$. We aim to prove that the $L^2$-orthogonal projection $\Pi_\star(T) : L^2(\Omega) \to S^p_0(T)$ is $H^1$-stable, i.e.,
\[ \|\nabla \Pi_\star(T)v\|_{L^2(\Omega)} \leq C_s\|\nabla v\|_{L^2(\Omega)} \text{ for all } v \in H^1(\Omega). \]
To see this, let $\Pi_0 : L^2(\Omega) \to \mathbb{R}$ be defined by $\Pi_0v = \frac{1}{|\Omega|} \int_\Omega v \, dx$. Identifying $\mathbb{R}$ with the constant functions on $\Omega$, it is known that $(1 - \Pi_0)$ is the $L^2$-orthogonal projection onto $L^2_0(\Omega)$. Since $\nabla \Pi_0 v = 0$, it follows that $(1 - \Pi_0)$ is also the $H^1$-orthogonal projection onto $H^1_0(\Omega)$. From nestedness $S^p_0(T) \subset H^1_0(\Omega) \subset L^2_0(\Omega)$, we obtain $\Pi_\star(T) = \Pi_\star(T)(1 - \Pi_0)$. Note that
\[ \int_\Omega (v - \Pi_\star(T)v) \, dx = 0 \text{ for all } v \in L^2(\Omega) \text{ and } V \in S^p_0(T), \]
\[ \int_\Omega (v - \Pi_\star(T)v) \, dx = 0 \text{ for all } v \in L^2_0(\Omega). \]
Since $S^p(T) = \text{span}(S^p_0(T) \cup \{1\})$, this implies
\[ \int_\Omega (v - \Pi_\star(T)v) \, dx = 0 \text{ for all } v \in L^2_0(\Omega) \text{ and } V \in S^p(T) \]
and hence $\Pi_\star(T)v = \Pi(T)v$ for all $v \in L^2_0(\Omega)$. Consequently, the stability (35) follows from (16) and $\Pi_\star(T) = \Pi_\star(T)(1 - \Pi_0) = \Pi(T)(1 - \Pi_0)$. Arguing along the lines of Section 6.2 6.3 one also obtains for all $0 \leq s \leq 1$
\[ \|\Pi_\star(T)v\|_{H^s(\Omega)} \leq C_s\|v\|_{H^s(\Omega)} \text{ for all } v \in H^s(\Omega), \]
\[ \|\Pi_\star(T)\psi\|_{H^{-s}(\Omega)} \leq C_s\|\psi\|_{H^{-s}(\Omega)} \text{ for all } \psi \in H^{-s}(\Omega), \]
for the unique continuous extension of $\Pi_\star(T)$ from $L^2(\Omega)$ to $H^{-s}(\Omega)$.

6.5. $L^2$-orthogonal projection on weighted $L^2$-spaces. The stability of the $L^2$-orthogonal projection $\Pi_D(T)$ onto $S^p_D(T)$ can also be established with respect to certain mesh-size weighted $L^2$-norms. The proof of Theorem 4 reveals that the assumptions (13) also ensure stability
\[ \|h^{-s}\Pi_D(T)v\|_{L^2(\Omega)} \leq C_\theta\|h^{-s}v\|_{L^2(\Omega)} \text{ for all } v \in L^2(\Omega) \]
in some weighted $L^2$-norm with $s = 1$, where $h \in L^\infty(\Omega)$ denotes the local mesh-width function (5). The constant $C_\theta > 0$ depends only on $C_\theta$, $C_3$, $L$, $\mu$, and $d$; see Section 6.
and, in particular, the proof of Proposition 9. For some positive weight-function \( \omega \in L^\infty(\Omega) \), we consider the weighted \( L^2 \)-norm

\[
\|v\|^2_{L^2(\omega;\Omega)} = \int_\Omega \omega(x)|v(x)|^2 \, dx,
\]

and \( L^2(\omega;\Omega) \) denotes the space of all measurable functions for which this norm is finite. It is known [BL76, Tar07] that interpolation of weighted \( L^p \)-spaces with \( \omega_0 \lesssim \omega_1 \) leads to

\[
[L^2(\omega_0;\Omega); L^2(\omega_1;\Omega)]_s = L^2(\omega_0^{1-s}\omega_1^s;\Omega) \quad \text{for all } 0 < s < 1.
\]

With \( \omega_0 = 1 \) and \( \omega_1 = h^{-2} \) as well as \( c_0 = 1 \) and \( c_1 = C_{11} \), the interpolation estimate [33] and (45) yield

\[
\|\Pi_D(\mathcal{T})v\|_{L^2(h^{-2s};\Omega)} \leq C_{11}\|v\|_{L^2(h^{-2s};\Omega)} \quad \text{for all } v \in L^2(h^{-2s};\Omega) = L^2(\Omega) \text{ and } 0 < s < 1.
\]

This, however, is equivalent to (44) for all \( 0 < s < 1 \). A deeper look into the proof of Theorem 4 resp. [BY14, Thm. 4.1] reveals the following improved result:

**Theorem 15.** Let \( s \in \mathbb{R} \). Suppose the assumptions of Theorem 7 as well as additionally \( \mu[s][L] < \frac{\sqrt{K_1K_2+1}}{\sqrt{K_1K_2-1}} \). Then, the \( L^2 \)-orthogonal projection onto \( S^p_D(\mathcal{T}) \) satisfies

\[
\|h^s\Pi_D(\mathcal{T})v\|_{L^2(\Omega)} \leq C_{11}\|h^sv\|_{L^2(\Omega)} \quad \text{for all } v \in L^2(\Omega).
\]

The constant \( C_{11} > 0 \) depends only on \( C_2, C_3, L, \mu, d, p, \) and \( |s| \).

**Sketch of proof.** Due to Proposition 13 and Proposition 12 the assumptions of Proposition 9 are satisfied with \( \mu[s][L] < \frac{\sqrt{K_1K_2+1}}{\sqrt{K_1K_2-1}} \) instead of the particular case \( s = -1 \). We use the notation of the proof of Proposition 9. Let \( N \) denote the maximal level of all elements \( T \in \mathcal{T} \). We define the norm

\[
\|v\|_{2-s}^2 := \sum_{i=0}^N \mu^{2si}\|v\|^2_{L^2(\Omega_i)} \simeq \|h^{-s}v\|^2_{L^2(\Omega)}.
\]

As above in the proof of Theorem 4 it suffices to prove

\[
\|\Pi_D(\mathcal{T})^v\|_{2-s} \lesssim \|v\|_{2-s} \quad \text{for all } v \in L^2(\Omega)
\]

to prove (46). For \( v \in L^2(\Omega) \), let \( v_k := v\chi_{\Omega_k} \). By means of (20) and with \( \kappa(n) := \min\{1,2^q[|n|-1]/L\} \), elementary calculation shows

\[
\|\Pi_D(\mathcal{T})^v\|_{2-s}^2 = \sum_{k=0}^N \|\Pi_D(\mathcal{T})^v_k\|_{2-s}^2 = \sum_{k=0}^N \mu^{2si}\|\sum_{k=0}^N \Pi_D(\mathcal{T})^v_k\|_{L^2(\Omega_i)}^2
\]

\[
= \sum_{i,k,\ell=0}^N \mu^{2si} \int_{\Omega_i} (\Pi_D(\mathcal{T})^v_k)(\Pi_D(\mathcal{T})^v_\ell) \, dx
\]

\[
\leq \sum_{i,k,\ell=0}^N \mu^{2si} K(|i-k|)K(|i-\ell|)\|v_k\|_{L^2(\Omega_i)}\|v_\ell\|_{L^2(\Omega)}
\]

\[
= \sum_{k,\ell=0}^N \left( \sum_{i=0}^N \mu^{n(i-k)}K(|i-k|)\mu^{n(i-\ell)}K(|i-\ell|) \right) \mu^{|k|}\|v\|_{L^2(\Omega_k)}\mu^{|\ell|}\|v\|_{L^2(\Omega_\ell)}.
\]

\[=: A_{ke} \]
It hence remains to bound $\Lambda$. For any eigenvalue $\lambda$, the matrix $A$ of all eigenvalues of $C$ definition of $\kappa$ holds for all $|s| < s_*$. Then, stability (46) guarantees weighted-$L^2$ stability (46) of the $L^2$-orthogonal projection onto $S_p^r(T)$ for all $|s| < s_*$. The matrix $A \in \mathbb{R}^{(N+1) \times (N+1)}$ is symmetric. With $\Lambda$ being the maximum absolute value of all eigenvalues of $A$, it hence follows

$$
\| \Pi_D(T)^{1/2}v \|_{L^2}^2 \leq \Lambda \sum_{k=0}^{N} \mu^{2k} \| v \|_{L^2(\Omega_s)}^2 = \Lambda \| v \|_{L^2}^2
$$

for all $v \in L^2(\Omega)$. It hence remains to bound $\Lambda$. For any eigenvalue $\lambda$ of $A$, it holds $|\lambda| \leq \max_{k=0,\ldots,N} \sum_{\ell=0}^{N} |A_{k\ell}|$. Define $C_s = \sum_{n=-N}^{N} \mu^n \kappa(n)$. Since all matrix coefficients $A_{k\ell}$ are positive, it holds

$$
\sum_{\ell=0}^{N} |A_{k\ell}| = \sum_{i=0}^{N} \mu^{s(i-k)} \kappa(|i-k|) \sum_{\ell=0}^{N} \mu^{s(i-\ell)} \kappa(|i-\ell|) \leq C_s \sum_{i=0}^{N} \mu^{s(i-k)} \kappa(|i-k|) \leq C_s^2
$$

Note that $C_s = C_{-s}$. Without loss of generality, we may hence assume $s > 0$. With the definition of $\kappa(\cdot)$ and the geometric series, we obtain

$$
C_s \leq \sum_{n=0}^{N} \mu^{-sn} + 2 \sum_{n=1}^{N} \mu^s q^{(n-1)/L} \leq \infty (1/\mu^s)^n + 2 \mu^s \sum_{n=1}^{\infty} ((\mu Ls q)^{1/L})^{n-1} \leq M < \infty.
$$

The upper bound $M$ depends only on $\mu > 1$ and $(\mu Ls q)^{1/L} < 1$.

**Remark 16.** Suppose that newest vertex bisection is used for mesh-refinement, i.e., $L = 1$ and $\mu = 2^{1/d}$. Then, stability (46) holds for all $|s| < s_* := \log_2 S/\log_2 \mu = d \log_2 S$, where $S := \sqrt{KN_2^2 + K_1^2}/\sqrt{KN_2^2 - 1}$; see Table 5 for $p = 1, 2$ and $d = 2, \ldots, 20$. 

Acknowledgement. DP acknowledges support through the Austrian Science Fund (FWF) under grant P21732 Adaptive boundary element method and P27005 Optimal adaptivity for BEM and FEM-BEM coupling. MK acknowledges support by CONICYT projects Anillo ACT1118 (ANANUM) and Fondecyt 3140614. The authors thank the anonymous referees for their suggestion to relate the current analysis to the results of Bank and Yserentant [BY14] which greatly improved this work from the lowest-order case to general $p \geq 1$.

References

[AFK+13] Markus Aurada, Michael Feischl, Josef F. Kemetmüller, Marcus Page, and Dirk Praetorius. Each $H^{1/2}$-stable projection yields convergence and quasi-optimality of adaptive FEM with inhomogeneous Dirichlet data in $\mathbb{R}^d$. Math. Model. Numer. Anal., 47:1207–1235, 2013.

[BL76] Jörn Bergh and Jörgen Löfström. Interpolation spaces. An introduction. Springer-Verlag, Berlin, 1976. Grundlehrer der Mathematischen Wissenschaften, No. 223.

[BPS02] James H. Bramble, Joseph E. Pasciak, and Olaf Steinbach. On the stability of the $L^2$ projection in $H^1(\Omega)$. Math. Comp., 71(237):147–156, 2002.

[BX91] James H. Bramble and Jinchao Xu. Some estimates for a weighted $L^2$ projection. Math. Comp., 56(194):463–476, 1991.

[BY14] Randolph E. Bank and Harry Yserentant. On the $H^1$-stability of the $L^2$-projection onto finite element spaces. Numer. Math., 126(2):361–381, 2014.

[Car02] Carsten Carstensen. Merging the Bramble-Pasciak-Steinbach and the Crouzeix-Thomée criterion for $H^1$-stability of the $L^2$-projection onto finite element spaces. Numer. Math., 91(2):157–163, 2002.

[Car04] Carsten Carstensen. An adaptive mesh-refining algorithm allowing for an $H^1$ stable $L^2$ projection onto Courant finite element spaces. Constr. Approx., 20(4):549–564, 2004.

[CFPP14] C. Carstensen, M. Feischl, M. Page, and D. Praetorius. Axioms of adaptivity. Comput. Math. Appl., 67(6):1195–1253, 2014.

[CKNS08] J. Manuel Cascon, Christian Kreuzer, Ricardo H. Nochetto, and Kunibert G. Siebert. Quasi-optimal convergence rate for an adaptive finite element method. SIAM J. Numer. Anal., 46(5):2524–2550, 2008.

[CT87] Michel Crouzeix and Vidar Thomée. The stability in $L^p$ and $W^{1,p}$ of the $L^2$-projection onto finite element function spaces. Math. Comp., 48(178):521–532, 1987.

[EJ95] Kenneth Eriksson and Claes Johnson. Adaptive finite element methods for parabolic problems. II. Optimal error estimates in $L^\infty L^2$ and $L^\infty L^\infty$. SIAM J. Numer. Anal., 32(3):706–740, 1995.

[FKMP13] M. Feischl, M. Karkulik, J. M. Melenk, and D. Praetorius. Quasi-optimal Convergence Rate for an Adaptive Boundary Element Method. SIAM J. Numer. Anal., 51(2):1327–1348, 2013.

[Kos94] Igor Kossaczký. A recursive approach to local mesh refinement in two and three dimensions. J. Comput. Appl. Math., 55(3):275–288, 1994.

[KPP13] Michael Karkulik, David Pavliček, and Dirk Praetorius. On 2D newest vertex bisection: Optimality of mesh-closure and $H^1$-stability of $L^2$-projection. Constr. Approx., 38:213–234, 2013.

[Mau95] Joseph M. Maubach. Local bisection refinement for $n$-simplicial grids generated by reflection. SIAM J. Sci. Comput., 16(1):210–227, 1995.

[McL00] William McLean. Strongly elliptic systems and boundary integral equations. Cambridge University Press, Cambridge, 2000.

[Sew72] Edward Granville Sewell. Automatic generation of triangulations for piecewise polynomial approximation. PhD thesis, Purdue University, 1972.

[Ste01] Olaf Steinbach. On the stability of the $L_2$ projection in fractional Sobolev spaces. Numer. Math., 88(2):367–379, 2001.

[Ste02] Olaf Steinbach. On a generalized $L_2$ projection and some related stability estimates in Sobolev spaces. Numer. Math., 90(4):775–786, 2002.

[Ste08] Rob Stevenson. The completion of locally refined simplicial partitions created by bisection. Math. Comp., 77(261):227–241, 2008.
[SZ90] L. Ridgway Scott and Shangyou Zhang. Finite element interpolation of nonsmooth functions satisfying boundary conditions. Math. Comp., 54(190):483–493, 1990.

[Tar07] Luc Tartar. An introduction to Sobolev spaces and interpolation spaces, volume 3 of Lecture Notes of the Unione Matematica Italiana. Springer, Berlin, 2007.

[Tra97] Christoph T. Traxler. An algorithm for adaptive mesh refinement in n dimensions. Computing, 59(2):115–137, 1997.

Pontificia Universidad Católica de Chile, Facultad de Matemáticas, Avenida Vícuta Mackenna 4860, Santiago, Chile
E-mail address: mkarkulik@mat.puc.cl

Vienna University of Technology, Institute for Analysis and Scientific Computing, Wiedner Hauptstrasse 8-10, A-1040 Wien, Austria
E-mail address: e1027222@student.tuwien.ac.at
E-mail address: Dirk.Praetorius@tuwien.ac.at (corresponding author)