**DE GRUYTER** Comput. Methods Appl. Math. 2021; 21(3): 529–556

**Research Article**

Carsten Carstensen* and Jun Hu

# Hierarchical Argyris Finite Element Method for Adaptive and Multigrid Algorithms

https://doi.org/10.1515/cmam-2021-0083
Received April 21, 2021; revised May 12, 2021; accepted May 12, 2021

**Abstract:** The global arrangement of the degrees of freedom in a standard Argyris finite element method (FEM) enforces $C^2$ at interior vertices, while solely global $C^1$ continuity is required for the conformity in $H^2$. Since the Argyris finite element functions are *not* $C^2$ at the midpoints of edges in general, the bisection of an edge for mesh-refinement leads to non-nestedness: the standard Argyris finite element space $A'(\mathcal{T})$ associated to a triangulation $\mathcal{T}$ with a refinement $\widehat{\mathcal{T}}$ is *not* a subspace of the standard Argyris finite element space $A'(\widehat{\mathcal{T}})$ associated to the refined triangulation $\widehat{\mathcal{T}}$. This paper suggests an extension $A(\mathcal{T})$ of $A'(\mathcal{T})$ that allows for nestedness $A(\mathcal{T}) \subset A(\widehat{\mathcal{T}})$ under mesh-refinement. The extended Argyris finite element space $A(\mathcal{T})$ is called hierarchical, but is still based on the concept of the Argyris finite element as a triple $(T, P_5(T), (\Lambda_1, \dots, \Lambda_{21}))$ in the sense of Ciarlet. The other main results of this paper are the optimal convergence rates of an adaptive mesh-refinement algorithm via the abstract framework of the axioms of adaptivity and uniform convergence of a local multigrid V-cycle algorithm for the effective solution of the discrete system.

**Keywords:** Argyris Element, Adaptive Mesh-Refinement, Discrete Quasi-Interpolation, Multigrid V-Cycle Algorithm

**MSC 2010:** 65N30, 65N12, 65N15

## 1 Introduction

### 1.1 Motivation

One of the most well-known conforming plate elements dates back to Argyris [4, 8, 15] with a quintic polynomial space $P_5(T)$ on a triangle $T$ and 21 degrees of freedom. The practical application is much less popular amongst numerical analysts, because of the higher implementation efforts compared to piecewise quadratic nonconforming Morley or discontinuous Galerkin finite element schemes for instance; the reader may consider the finite element program in [11, Section 6.5] with less than 30 lines of Matlab, the clear analysis of the $C^0$ interior penalty method [9]. The point is that higher convergence rates of the quintic Argyris finite element space are rarely visible even in simple computational benchmarks with the biharmonic equation and a right-hand side in $L^2$ in a polygonal bounded Lipschitz domain $\Omega$. For a nonconvex domain $\Omega$ and a sequence of quasi-uniform triangulations, the convergence rates of the aforementioned schemes are the same and the extra effort for the quintic method appears not competitive.

Adaptive mesh-refinement algorithms are well established with optimal convergence rates [2, 10, 12, 13, 24] merely for the Morley (cf. [11, 20] and the references therein) amongst the aforementioned finite

***Corresponding author: Carsten Carstensen,** Humboldt-Universität zu Berlin, 10099 Berlin, Germany, e-mail: cc@math.hu-berlin.de
**Jun Hu,** LMAM and School of Mathematical Sciences, Peking University, Beijing 100871, P. R. China, e-mail: hujun@math.pku.edu.cn

element methods (FEM) for fourth-order problems. The a posteriori error control for the Argyris FEM is known for more than two decades and even included in the first book [26] on a posteriori error analysis. So it appears surprising that the important task is open to design a rate-optimal adaptive mesh-refinement algorithm and merely plain convergence has been achieved [23, Problem 44].

A closer look at the mathematical foundation of adaptive mesh-refinement reveals that the Argyris FEM leads to non-nestedness under mesh-refinements and this causes mathematical difficulties in the overall analysis in [2, 10, 12, 13, 24]. In fact, according to the knowledge of the authors, the only positive convergence result in the context of adaptive algorithms of conforming FEM that overcomes non-nestedness is [31] for a particular mesh-refinement strategy.

The treatment of non-nested subspaces has a longer tradition in the mathematical foundation of the multilevel solver, but uniform convergence of standard multigrid V-cycle algorithms has only been established for the Bogner–Fox–Schmit and the Powell–Sabin FEMs [25, 30]. Those are the conforming FEMs for fourth-order problems with the nestedness property under uniform mesh-refinement. See [28] for the Morley based preconditioner for the Argyris element for biharmonic equations.

The two examples at least indicate severe difficulties in the mathematical investigation of non-nested finite element spaces and this paper enforces the nestedness by a marginal extension of the global degrees of freedom.

## 1.2 Hierarchical Argyris FEM

The adaptive mesh-refinement algorithms and the multigrid methods are typically based on admissible triangulations obtained by a mesh-refinement strategy such as the newest-vertex bisection [24]. Given an initial triangulation $\mathcal{T}_0$ (and some initialization for the subordinated bisection rules), a sequence of successive (admissible) refinements $\mathcal{T}_0, \mathcal{T}_1, \mathcal{T}_2, \ldots$ defines an admissible triangulation $\mathcal{T}_\ell$ and the set $\mathbb{T}$ of admissible triangulations of all those shape-regular triangulations. Any such admissible triangulation $\mathcal{T}_\ell$ leads to a non-nested Argyris finite element space $A'(\mathcal{T}_\ell) \subset H_0^2(\Omega)$ in the sense that $A'(\mathcal{T}_\ell) \nsubseteq A'(\mathcal{T}_{\ell+1})$ (if $\mathcal{T}_{\ell+1} \neq \mathcal{T}_\ell$ is a strict refinement of $\mathcal{T}_\ell$). An immediate cure of this failure is the definition of the nested spaces

$$A(\mathcal{T}_\ell) := A'(\mathcal{T}_0) + A'(\mathcal{T}_1) + \cdots + A'(\mathcal{T}_\ell)$$

as the sum of all preceding finite element spaces: The nestedness of those sums $A(\mathcal{T}_\ell) \subset A(\mathcal{T}_{\ell+1})$ is immediate.

Suppose in addition that $\mathcal{T}_{\ell+1}$ is a one-level refinement of $\mathcal{T}_\ell$ for any $\ell \in \mathbb{N}_0$, i.e., any refined edge in $\mathcal{T}_\ell$ is bisected exactly once. Then, $A(\mathcal{T}_\ell) \subset H_0^2(\Omega)$ depends exclusively on $\mathcal{T} = \mathcal{T}_\ell \in \mathbb{T}$ but is independent of the finite sequence of successive (admissible) refinements that creates $\mathcal{T}$. This leads to the proposed finite element space $A(\mathcal{T})$ for each $\mathcal{T} \in \mathbb{T}$ with the degrees of freedom of the classical Argyris finite element space $A(\mathcal{T})$ plus one extra function per each interior vertex $z$ in the triangulation $\mathcal{T}$ that is not an initial vertex (i.e., $z$ is not a vertex in $\mathcal{T}_0$). In other words, the vertex $z$ had been created by bisection in some coarser triangulation $\mathcal{T}'$ as the midpoint $z = \operatorname{mid}(E)$ of an interior edge $E = \partial K_+ \cap \partial K_- = \operatorname{conv}\{A, B\}$ of length $|E|$ shared by the two triangles $K_+$ and $K_- \in \mathcal{T}'$ of respective area $|K_\pm|$. The additional nodal basis function $\psi_E \in H_0^2(K_+ \cup K_-) \cap A'(\mathcal{T}') \subset H_0^2(\Omega)$ reads

$$\psi_E := \mp \frac{2|K_\pm|}{|E|} \varphi_E^2 \varphi_{P_\pm} \quad \text{in } K_\pm.$$

(Herein $\varphi_A$, $\varphi_B$, $\varphi_{P_\pm}$ denote the barycentric coordinates of the vertices $A$, $B$, and $P_\pm$ in the triangle $K_\pm$ and $\varphi_E = 4\varphi_A\varphi_B$ abbreviates the quadratic edge-bubble function of the edge $E$ with vertices $A$ and $B$.) This is the nodal basis function in the coarser triangulation $\mathcal{T}'$ associated to the normal derivative at the midpoint $z$ of $E$ with a jump of the second-order normal-normal derivative $\partial^2_{\nu\nu}$ and has to be kept in the nested space $A(\mathcal{T})$; while the standard Argyris finite element space arranges the global degrees of freedom at the vertex $z$ in $\mathcal{T}$ such that all finite element functions in $A'(\mathcal{T})$ are $C^2$ at $z = \operatorname{mid}(E)$. The increase of unknowns in $A(\mathcal{T})$ compared to $A'(\mathcal{T})$ is by a factor less than $\frac{7}{6}$ only.

## 1.3 Contributions and Outline of This Paper

Section 2 gives all the details of the global scheme called hierarchical Argyris FEM because of one additional degree of freedom that is a one-sided second-order normal-normal derivative at a new interior node associated to the nodal basis function $\psi_E$ from above. Let us emphasize that the shape functions in each triangle are exactly those of the standard Argyris FEM; the scheme utilizes the Argyris finite element $(T, P_5(T), (\Lambda_1, \dots, \Lambda_{21}))$ in the sense of Ciarlet and modifies the global arrangements of the unknowns. This standard Argyris FEM leads to a finite element space $A'(\mathcal{T}) \subset H^2(\Omega)$ as a strict subspace of the $C^1$ conforming piecewise quintic polynomials $\widehat{A}(\mathcal{T}) := P_5(\mathcal{T}) \cap H^2(\Omega)$ characterized in [21] with singularities at cross-points (those are systematically generated in the newest vertex bisection). The hierarchical Argyris FEM leads to a finite element space $A(\mathcal{T}) \subset H^2(\Omega)$ with $A'(\mathcal{T}) \subset A(\mathcal{T}) \subset \widehat{A}(\mathcal{T})$ that allows for the nestedness property as the first main result of this paper: $A(\mathcal{T}) \subset A(\widehat{\mathcal{T}})$ holds for any (admissible) refinement $\widehat{\mathcal{T}}$ of $\mathcal{T} \in \mathbb{T}$.

The quasi-interpolation operator is one main tool in the a posteriori error analysis and in the split of multilevel schemes. Section 3 revisits [19] on the standard Argyris FEM and extends it to the discrete quasi-interpolation for the hierarchical version at hand in Theorem 2 with a discrete version that enables a proof of the discrete reliability from the axioms of adaptivity [10, 12] in Section 4. The second main result is the rate-optimality of an adaptive algorithm in Theorem 6 that in fact recovers the quartic convergence order of the scheme on polygons, even on non-convex ones, and clearly outperformed the aforementioned quadratic schemes in practice. Since there is no general convergence result available for the discontinuous Galerkin schemes, except for over-penalization as outlined in [3] for second-order problems, the presented adaptive hierarchical Argyris FEM is the most competitive numerical scheme for biharmonic equations and their relatives.

The third main result of this paper is the uniform convergence of a multigrid V-cycle algorithms with uniform convergence. Section 5 establishes a version for a sequence of successive one-level refinements based on newest-vertex bisection with local smoothing for the new nodes plus their neighbors. This is an effective algorithm known for second-order problems from [27] and Theorem 7 guarantees a bounded condition number for approximative inverse from the effective multigrid preconditioner for the biharmonic problem. The proofs in Sections 6 and 7 of this paper concern an auxiliary uniform triangulation that allow for standard approximation estimates from global interpolation of Sobolev norms. The link to the sequence of triangulations with local refinement utilized in the multigrid algorithm is through geometric arguments in [27] that *on the specific geometric structure of bisection grids* [14]. This paper replaces all geometry by a re-summation argument that is utterly algebraic and hence not restricted to 2D.

The analysis is outlined for the quintic version of the Argyris FEM for an easy and explicit presentation of the ideas. The results hold for the hierarchical conforming higher-order Argyris FEM as well. The motivating plate model problem is intrinsically 2D, and so the analysis is carried out in 2D, but the arguments does not rely on two space dimensions.

## 1.4 General Notation

Standard notation on Lebesgue and Sobolev spaces and norms applies throughout this paper, e.g., $\|\cdot\|_{L^2(\Omega)}$ denotes the $L^2$ norm. Sobolev functions are usually defined on open sets and the notation $W^{m,p}(T)$ (resp. $W^{m,p}(\mathcal{T})$) substitutes $W^{m,p}(\operatorname{int}(T))$ for a (compact) triangle $T$ and its interior $\operatorname{int}(T)$ (resp. $W^{m,p}(\operatorname{int}(\mathcal{T}))$) and their vector and matrix versions.

The context-sensitive measure $|\bullet|$ refers to (counting measure or cardinality) the number of elements of some finite set or the surface measure $|F|$ of a side $F$ or the volume $|T|$ of a tetrahedron $T$, etc., and not just to the modulus of a real number or the Euclidean length of a vector.

Throughout this paper, $A \lesssim B$ abbreviates $A \leq C\,B$ for some generic constant $C$ that solely depends on $\mathcal{T}_0$, the initial triangulation of the bounded polygonal Lipschitz domain $\Omega$ into triangles, and $A \approx B$ abbreviates $A \lesssim B \lesssim A$.

# 2 The Hierarchical Argyris FEM

Based on the class of admissible triangulations generated from successive mesh-refinement, the Argyris finite element $(T, P_5(T), (\Lambda_1, \dots, \Lambda_{21}))$ in the sense of Ciarlet is minimally extended in the global arrangements of the degrees of freedom.

## 2.1 Admissible Triangulations

Given any initial triangulation $\mathcal{T}_0$ of a bounded Lipschitz domain with polygonal boundary $\partial\Omega$ into triangles, let $\mathbb{T} = \mathbb{T}(\mathcal{T}_0)$ be the set of all regular triangulations obtained from $\mathcal{T}_0$ with a finite number of successive bisections of triangles [10, 24]. In fact, any triangle $T = \operatorname{conv}\{P_1, P_2, P_3\}$ is tagged and identified with the triple $(P_1, P_2, P_3)$ (the tag is a type and can be omitted in 2D).

The bisection $\operatorname{bisec}(T) := \{T_1, T_2\}$ of the finite element domain is simply $T_1 \equiv (P_2, \frac{1}{2}(P_1 + P_3), P_1)$ and $T_2 \equiv (P_3, \frac{1}{2}(P_1 + P_3), P_2)$, but the order of vertices inherited are steered by certain rules plus some initial conditions on $\mathcal{T}_0$. Further details on the newest-vertex bisection (NVB) can be found in [6, 24]. Define $\mathbb{T}(\mathcal{T})$ as the set of all admissible refinements $\widehat{\mathcal{T}}$ of $\mathcal{T}$.

Given any admissible triangulation $\mathcal{T} \in \mathbb{T}$, let $\mathcal{V}$ (resp. $\mathcal{E}$) denote the vertices (resp. edges) in $\mathcal{T}$ and let $\mathcal{V}(\Omega)$ (resp. $\mathcal{E}(\Omega)$) be the interior vertices (resp. edges); let $\mathcal{M} := \{\operatorname{mid}(E) : E \in \mathcal{E}\}$ denote the set of the edges' midpoints. For any $\mathcal{T} \in \mathbb{T}$ and a vertex $z \in \mathcal{V}$ or a midpoint $z = \operatorname{mid}(E) \in \mathcal{M}$ of an edge $E \in \mathcal{E}$, let

$$\mathcal{T}(z) := \{T \in \mathcal{T} : z \in T\}$$

denote the set of neighboring triangles that contain $z$. Let $\mathcal{V}(T)$ (resp. $\mathcal{E}(T)$) denotes the set of the three vertices (resp. edges) of a triangle $T$. For any $\mathcal{T} \in \mathbb{T}$ and a triangle $T \in \mathcal{T}$ let $\mathcal{R}_1(T) := \{K \in \mathcal{T} : \operatorname{dist}(T, K) = 0\}$ denote the neighboring triangles of $T$ (i.e., $T$ and one layer of triangles in $\mathcal{T}$ around it) and let

$$\Omega(T) := \operatorname{int}\Big(\bigcup \mathcal{R}_1(T)\Big)$$

denote the interior of the union of all neighboring triangles of $T$.

Within the definition of $\mathcal{T}$ via the NVB, each node $z \in \mathcal{V} \setminus \mathcal{V}_0$ has been created as a midpoint $z = \operatorname{mid}(E)$ of an edge $E$ with tangent vector $\tau_E$ and normal $\nu(z) := \nu_E$ (with fixed orientation such that it points outwards along the boundary $\partial\Omega$). In this way, any node $z \in \mathcal{V}(\Omega) \setminus \mathcal{V}_0$ is associated to one direction $\nu(z)$. Each triangle $T \in \mathcal{T}$ with vertex $z \in \mathcal{V} \setminus \mathcal{V}_0$ either belongs to the closed half-space $H_+(z) := \{x \in \mathbb{R}^2 : (x - z) \cdot \nu(z) \geq 0\}$ or to the closed half-space $H_-(z) := \{x \in \mathbb{R}^2 : (x - z) \cdot \nu(z) \leq 0\}$. This gives rise to the partition

$$\mathcal{T}_\pm(z) := \{T \in \mathcal{T}(z) : T \subset H_\pm(z)\},$$

which shall be relevant for the additional global degrees of freedom in the extended Argyris finite element space $A(\mathcal{T})$.

## 2.2 Extended Argyris Finite Element Space

The extended Argyris finite element space $A(\mathcal{T})$ is extended in the global degrees of freedom to make it hierarchical. Given an admissible triangulation $\mathcal{T} \in \mathbb{T}$, recall that any interior vertex $z \in \mathcal{V}(\Omega) \setminus \mathcal{V}_0$ is associated to some direction $\nu(z)$ and the separation $\mathcal{T}_+(z)$ and $\mathcal{T}_-(z)$ of the neighboring triangles $\mathcal{T}(z) := \{T \in \mathcal{T} : z \in \mathcal{V}(T)\}$ with vertex $z$.

The elementwise degrees of freedom of the Argyris finite element space are not changed except for a transformation of another global coordinate system in the direction $\nu(z)$ and its tangential direction $\tau(z) \perp \nu(z)$. The point is that the global degrees of freedom for the second-order normal-normal derivative $\partial^2_{\nu(z)\nu(z)}$ at $z$ is not uniquely defined at $z$, but allows for one value in $H_+(z)$ and a second value in $H_-(z)$.

The resulting global degrees of freedom are summarized for clarity for a given admissible triangulation $\mathcal{T} \in \mathbb{T}$. For each node $z$ let $m(z)$ denote the number of degrees of freedom associated to it while for each initial vertex $z \in \mathcal{V}_0$, the $m(z) = 6$ nodal degrees of freedom are the classical ones $\delta_z \partial^\alpha$ for a multi-index $\alpha \in \{0, 1, 2\}^2$ of order $|\alpha| := \alpha_1 + \alpha_2 = 0, 1, 2$ in terms of the Dirac evaluation functional $\delta_z$ at $z$ (i.e., $\delta_z f = f(z)$) of the derivative $\partial^\alpha := \partial^{\alpha_1}_{x_1^{\alpha_1}} \partial^{\alpha_2}_{x_2^{\alpha_2}}$, i.e.,

$$\delta_z \partial^\alpha f = (-1)^{|\alpha|} \partial^{\alpha_1}_{x_1^{\alpha_1}} \partial^{\alpha_2}_{x_2^{\alpha_2}} f(z) \quad \text{for any } f \in C^2.$$

In this lexicographic order of multi-indices and derivatives $\partial^\alpha$, those $m(z) = 6$ degrees of freedom at the vertex $z$ are also denoted as $\Lambda_{z,1}, \dots, \Lambda_{z,6}$. For any other interior vertex $z \in \mathcal{V}(\Omega) \setminus \mathcal{V}_0$, there are seven nodal degrees of freedom in terms of the coordinate system with $\nu(z)$ and $\tau(z)$, namely, $\delta_z$, $\delta_z \partial_{\nu(z)}$, $\delta_z \partial_{\tau(z)}$, $\delta_z \partial^2_{\tau(z)\tau(z)}$, $\delta_z \partial^2_{\tau(z)\nu(z)}$, and the two functionals $\delta_z^\pm \partial^2_{\nu(z)\nu(z)}$, where $\delta_z^\pm$ evaluates a function $f \in C^2(H_\pm(z))$ on the half-plane $H_\pm(z)$ and then takes the one-sided limit

$$\delta_z^\pm \partial^2_{\nu(z)\nu(z)} f = \lim_{H_\pm(x) \ni x \to z} (\partial^2_{\nu(z)\nu(z)} f)(x)$$

at $z$. In this order those $m(z) = 7$ degrees of freedom are also denoted as $\Lambda_{z,1}, \dots, \Lambda_{z,7}$. The split into two variables for the second-order normal derivatives in $H_\pm(z)$ makes the difference to the standard Argyris finite element space. The remaining degrees of freedom, namely, the $m(z) = 1$ normal derivative $\Lambda_{z,1} := \delta_{\mathrm{mid}(E)} \partial_{\nu_E}$ at any midpoint $z = \mathrm{mid}(E)$ of an interior edge $E$ with normal $\nu_E$, are not modified. All the degrees of freedom located at the boundary are set to zero except the $m(z) = 1$ second-order normal-normal derivative $\Lambda_{z,1} := \delta_z \partial^2_{\nu\nu}$ at a boundary node $z \in \mathcal{V}(\partial\Omega)$ that is not a corner of the boundary $\partial\Omega$ (i.e., the interior angle $\omega = \pi$ at a vertex $z \in \partial\Omega$ and the exterior normal $\nu$ is constant in a neighborhood of $z$ along $\partial\Omega$) to guarantee $A(\mathcal{T}) \subset P_5(\mathcal{T}) \cap H_0^2(\Omega)$.

**Remark 2.1** (Dimension). The dimension of the finite element space $A(\mathcal{T})$ is

$$N := 7|\mathcal{V}(\Omega)| - |\mathcal{V}_0(\Omega)| + |\mathcal{E}(\Omega)| + |\mathcal{V}(\partial\Omega)| - |V(\Omega)|$$

for the number $|\mathcal{V}(\Omega)|$ (resp. $|\mathcal{E}(\Omega)|$) of interior vertices $\mathcal{V}(\Omega)$ (resp. interior edges $\mathcal{E}(\Omega)$) and the number $|V(\Omega)|$ of corners $V(\Omega)$ of the polygon (or the finite union of polygons) $\partial\Omega$; $\mathcal{V}(\partial\Omega) := \mathcal{V} \setminus \mathcal{V}(\Omega)$ is the set of vertices on the boundary $\partial\Omega$.

**Notation 2.2** (Nodes). Given any admissible triangulation $\mathcal{T} \in \mathbb{T}$ and its set of vertices $\mathcal{V}$ and edges $\mathcal{E}$, let the set of nodes $\mathcal{N}$ be the set of all vertices or midpoints of edges that carry at least a degree of freedom in $A(\mathcal{T})$. Let $\mathcal{M} := \{\mathrm{mid}(E) : E \in \mathcal{E}\}$ denote the set of all midpoints of edges and $\mathcal{M}(\Omega) := \{z \in \mathcal{M} : z \in \Omega\}$ the set of all midpoints of interior edges (which are those shared by two triangles in $\mathcal{T}$). Then

$$\mathcal{N} := \mathcal{V}(\Omega) \cup \mathcal{M}(\Omega) \cup (\mathcal{V}(\partial\Omega) \setminus V(\Omega)).$$

**Notation 2.3** (Degrees of Freedom). There are $m(z) \in \{1, \dots, 7\}$ degrees of freedom $\Lambda_{z,1}, \dots, \Lambda_{z,m(z)}$ associated to any node $z \in \mathcal{N}$. Abbreviate the index set for all global degrees of freedom in $A(\mathcal{T})$ by

$$\mathcal{I} := \{(z, j) \in \mathcal{N} \times \{1, \dots, 7\} : z \in \mathcal{N} \text{ and } j = 1, \dots, m(z)\} = \bigcup \{\{z\} \times \{1, \dots, m(z)\} : z \in \mathcal{N}\}.$$

It is also convenient to use an alternative notation and to enumerate the set of degrees of freedom $\Lambda_1, \dots, \Lambda_N$ in $A(\mathcal{T})$,

$$\{\Lambda_{z,j} : z \in \mathcal{N},\ j = 1, \dots, m(z)\} \equiv \{\Lambda_1, \dots, \Lambda_N\}.$$

**Notation 2.4** (Nodal Basis). Given the $N$ global degrees of freedom $\Lambda_1, \dots, \Lambda_N$ of the extended Argyris finite element space $A(\mathcal{T})$, there exists a unique dual basis $\varphi_1, \dots, \varphi_N \in \widehat{A}(\mathcal{T}) \equiv P_5(\mathcal{T}) \cap H_0^2(\Omega)$ with the duality relation $\Lambda_k(\varphi_j) = \delta_{jk}$ for $j, k = 1, \dots, N$ and Kronecker's $\delta_{jk}$ and $A(\mathcal{T}) = \mathrm{span}\{\varphi_1, \dots, \varphi_N\}$. (This follows from the linear independence of $\Lambda_1, \dots, \Lambda_N$ in $\widehat{A}(\mathcal{T})$.) In the notation of the previous remark, the alternative notation $\psi_{z,j} := \varphi_k$ applies when $\Lambda_k = \Lambda_{z,j}$ and $(z, j) \in \mathcal{I}$ has the global number $k \in \{1, \dots, N\}$.

**Remark 2.5** (Standard Argyris Finite Element). The standard Argyris finite element space

$$A'(\mathcal{T}) = \{f \in A(\mathcal{T}) : \delta_z^+ \partial_{\nu(z)} \partial_{\nu(z)} f = \delta_z^- \partial_{\nu(z)} \partial_{\nu(z)} f \text{ for all } z \in \mathcal{V}(\Omega) \setminus \mathcal{V}_0\}$$

is a linear subspace of $A(\mathcal{T})$.

Unlike the classical Argyris finite element space $A'(\mathcal{T})$, the extended version $A(\mathcal{T})$ of this paper leads to nested subspaces. In fact, the desire for this nestedness property motivates the concept of the minimal extension of the Argyris finite element for adaptively refined triangulations.

**Theorem 1** (Nestedness). *Any triangulation $\mathcal{T} \in \mathbb{T}$ and its refinement $\widehat{\mathcal{T}} \in \mathbb{T}(\mathcal{T})$ allow for $A(\mathcal{T}) \subset A(\widehat{\mathcal{T}})$.*

*Proof.* It is straightforward to verify that any bisection of an interior edge $E = \partial T_+ \cap \partial T_- \in \mathcal{E}(\Omega)$ and the attached two triangles $T_\pm \in \mathcal{T}$ leads to seven new global degrees of freedom at the midpoint $z := \mathrm{mid}(E)$ of an edge $E \in \mathcal{E}$ ($\mathcal{E}$ is the set of edges in $\mathcal{T}$); whence $z \in \widehat{\mathcal{V}}(\Omega) \setminus \mathcal{V}$ ($\widehat{\mathcal{V}}(\Omega)$ denotes the set of interior vertices in $\widehat{\mathcal{T}}$). Given $v_A \in A(\mathcal{T})$, the polynomial $v_A|_{T_\pm} \in P_5(T_\pm)$ and its gradient is continuous along $E$; whence $v_A$, $\nabla v_A$, and $\partial_{\tau(z)} \nabla v_A$ are globally continuous at $z$. Note that the second order normal-normal derivative $\partial_{\nu(z)}\partial_{\nu(z)} v_A$ is in general discontinuous at $z$ which can be expressed by the global basis functions associated to functionals $\Lambda_{z,6}$ and $\Lambda_{z,7}$. Consequently, $v_A$ can be represented with the nodal basis functions $\widehat{\varphi}_1, \dots, \widehat{\varphi}_{\widehat{N}} \in P_5(\widehat{\mathcal{T}}) \cap H_0^2(\Omega)$ of Notation 2.4 on the refined triangulation $\widehat{\mathcal{T}}$; whence $v_A \in A(\widehat{\mathcal{T}})$. This proves the assertion $A(\mathcal{T}) \subset A(\widehat{\mathcal{T}})$ in case that $\widehat{\mathcal{T}}$ is refined from $\mathcal{T}$ by some bisection of an interior edge; the proof for an exterior edge can be adopted and is hence omitted.

The general assertion follows from the above, because any NVB is obtained by a sequence of successive bisections of marked edges. □

## 3 Discrete Quasi-Interpolation

This section is devoted to the design of a quasi-interpolation operator for the extended Argyris finite element spaces of this paper with properties (a)–(b) in this section as a tool for the proof of discrete reliability in Section 4 and (c)–(d) for the multigrid analysis in Sections 6–7. Given any triangle $T \in \mathcal{T}$ with the diameter $h_T := \mathrm{diam}(T)$ recall the patch $\Omega(T) := \mathrm{int}(\bigcup\{T' \in \mathcal{T} : \mathrm{dist}(T, T') = 0\})$ from Section 2.1.

**Theorem 2** (Discrete Quasi-Interpolation). *Given the initial triangulation $\mathcal{T}_0$ with initial conditions and the associated NVB refinement that defines $\mathbb{T}$, there exist positive constants $C_{\mathrm{apx}}$, $C_{\mathrm{c}}$, and $C_{\mathrm{d}}$ such that, for any triangulation $\mathcal{T} \in \mathbb{T}$ with any refinement $\widehat{\mathcal{T}} \in \mathbb{T}(\mathcal{T})$, there exists a linear operator $J : H^1(\Omega) \to A(\mathcal{T})$ satisfying (a)–(d) for any $T \in \mathcal{T}$,*
(a) *$\widehat{v}_A = J\widehat{v}_A$ holds in $T \in \mathcal{T} \cap \widehat{\mathcal{T}}$ for any $\widehat{v}_A \in A(\widehat{\mathcal{T}})$,*
(b) *$\sum_{m=0}^{2} h_T^{m-2} |f - Jf|_{H^m(T)} \le C_{\mathrm{apx}} |f|_{H^2(\Omega(T))}$ holds for any $f \in H_0^2(\Omega)$,*
(c) *$C_{\mathrm{c}}^{-1} \|Jf\|_{L^2(T)} \le \|f\|_{L^2(\Omega(T))} + h_T \|\nabla f\|_{L^2(\Omega(T))}$ holds for any $f \in H^1(\Omega(T))$,*
(d) *$\|f - Jf\|_{L^2(T)} \le C_{\mathrm{d}} h_T |f|_{H^1(\Omega(T))}$ holds for any $f \in H_0^1(\Omega)$.*

It is important to understand quasi-interpolation not as a single operator but rather as a class of operators with many choices [16] and the focus in this paper is on conditions from mesh-refinement. There are two conditions (C1)–(C2) in the first part of this paper on the rate-optimal adaptive algorithm (and (C3)–(C4) in the multigrid application in the second). Recall that the extended Argyris finite element space $A(\mathcal{T})$ is associated with the degrees of freedom $\{\Lambda_{z,j} : (z, j) \in \mathcal{J}\}$ of Notation 2.3 and the nodal basis functions $\{\psi_{z,j} : (z, j) \in \mathcal{J}\}$ of Notation 2.4. The definition of the quasi-interpolation in the proof of Theorem 2 below allows for a choice of $T_{z,j} \in \mathcal{T}$ for each $(z, j) \in \mathcal{J}$ in the design of a functional $M_{z,j}$. Recall that $\mathcal{T}(z)$ is the set of neighboring triangles $T \ni z$ and recall the half-planes $H_{\pm(z)}$ from Section 2.1 and $\mathcal{T}_\pm(z)$ from Section 2.2.

**Condition (C1).** The family $(T_{z,j} : (z, j) \in \mathcal{J})$ of triangles satisfies

$$T_{z,1}, \dots, T_{z,m(z)} \in \mathcal{T}(z) := \{T \in \mathcal{T} : z \in T\}$$

for all $z \in \mathcal{N}$ and, in addition, $T_{z,6} \in \mathcal{T}_+(z)$ and $T_{z,7} \in \mathcal{T}_-(z)$ at each $z \in \mathcal{V}(\Omega) \setminus \mathcal{V}_0(\Omega)$.

The next condition assures (a) for the admissible refinement $\widehat{\mathcal{T}}$ of $\mathcal{T}$. Define the set $\widehat{\mathcal{T}}(z) := \{\widehat{T} \in \widehat{\mathcal{T}} : z \in \widehat{T}\}$ of neighboring triangles $\widehat{T}$ from the refinement $\widehat{\mathcal{T}}$ with $z \in \widehat{T}$ and observe that $z \in \mathcal{N} \subset \widehat{\mathcal{N}}$ is also a node of the refined triangulation. Let $\widehat{\mathcal{T}}_\pm(z)$ denote the subset of triangles in $\widehat{\mathcal{T}}(z)$ for $z \in \mathcal{V}(\Omega) \setminus \mathcal{V}_0(\Omega)$ that belong to $H_\pm(z)$.

**Condition (C2).** The family $(T_{z,j} : (z,j) \in \mathcal{J})$ of triangles satisfies, for all $z \in \mathcal{N}$, that

$$\mathcal{T}(z) \cap \widehat{\mathcal{T}}(z) \neq \emptyset \implies T_{z,1}, \ldots, T_{z,m(z)} \in \mathcal{T}(z) \cap \widehat{\mathcal{T}}(z)$$

and in addition, for all $z \in V(\Omega) \setminus \mathcal{V}_0(\Omega)$ (with $m(z) = 7$), that

$$\begin{aligned}\mathcal{T}_+(z) \cap \widehat{\mathcal{T}}_+(z) \neq \emptyset &\implies T_{z,6} \in \mathcal{T}_+(z) \cap \widehat{\mathcal{T}}_+(z),\\ \mathcal{T}_-(z) \cap \widehat{\mathcal{T}}_-(z) \neq \emptyset &\implies T_{z,7} \in \mathcal{T}_-(z) \cap \widehat{\mathcal{T}}_-(z).\end{aligned}$$

If there is no coarse and fine triangle that contains $z$, i.e., if $\mathcal{T}(z) \cap \widehat{\mathcal{T}}(z) = \emptyset$, then (C2) puts no condition on the choice of $T_{z,j}$; the condition (C2) is fully redundant in the extreme case of $\mathcal{T} \cap \widehat{\mathcal{T}} = \emptyset$ of an overall refinement.

Conditions (C1)–(C2) can always be fulfilled by *some* choice of $(T_{z,j} : (z,j) \in \mathcal{J})$, which is non-unique in general, but *any* such choice with (C1) is sufficient for (b)–(d), while (C1)–(C2) are sufficient for (a). The proof of Theorem 2 is a modification of that in [19] on the classical Argyris finite element space to the finite element space $A(\mathcal{T})$ of this paper. Hence we merely sketch the arguments and comment on the necessary modifications in particular for (c)–(d). Homogeneous boundary conditions in this paper slightly simplify the presentation because they allow for merely weighted averaging over triangles (and avoid weighted averaging over edges) as pointed out already in [19, Remark 7.8].

*Definition of $J$ in Theorem 2.* Given $\mathcal{T} \in \mathbb{T}$ and Notations 2.2–2.3, any global degree of freedom $\Lambda_k = \Lambda_{z,j}$ in the finite element space $A(\mathcal{T})$ is associated to some node $z \in \mathcal{N}$ and is equal to some derivative $D_\alpha$ of order 0, 1, or 2 of a smooth function $f$ restricted to some domain $\omega_k$ with $z \in \overline{\omega_k}$;

$$\Lambda_k f = \Lambda_{z,j} f = D_\alpha(f|_{\omega_k})(z_k) \quad \text{for all } f \in C^2(\overline{\omega_k}).$$

Compared to the classical Argyris finite element space, the only modification here is that some nodes have 7 rather than 6 degrees of freedom of this form. The shape functions are the same but the global arrangement involves functionals with one-sided second-order normal-normal derivatives when either $\overline{\omega_k} \subset H_+(z)$ or $\overline{\omega_k} \subset H_-(z)$. Given $f \in H^1(\Omega)$, its approximation

$$Jf = \sum_{k=1}^{N} M_k(f)\varphi_k \in A(\mathcal{T}) \tag{3.1}$$

is defined in terms of the nodal basis functions $\varphi_k = \psi_{z,j} \in A(\mathcal{T})$ of Notation 2.4 and some linear functionals $M_k = M_{z,j}$ as in [19] (explicitly given in integral formulas in [19, equation (4.15)]). We recall from [19] that, given any $\Lambda_{z,j} = D_\alpha(\bullet|_{\omega_k})(z)$ one may select any $T_{z,j}$ with (C1). Given $(z, T_{z,j}, D_\alpha)$, the proof designs a weight function $\widetilde{\psi}_k$ (and below $\psi_k$) on $T_k$ and defines a derivative $\widetilde{D}_k$ with

$$M_k(f) = \int_{T_{z,j}} \tilde{\psi}_k(x)\widetilde{D}_k f(x)\,dx \quad \text{for all } f \in H^1(\omega_k) \tag{3.2}$$

for all $k = 1, \ldots, N$. The derivative $\widetilde{D}_k$ is equal to $D_k$ in case its order $|D_k|$ is 0 or 1, while an integration by parts in $T_{z,j}$ allows for the order $|\widetilde{D}_k| = 1$ when $|D_k| = 2$. We refer to [19, equations (4.8)–(4.16)] for all the details and the scaling properties of the associated functions and mention that [19] utilizes the notation $\kappa_k$ rather than $T_{z,j}$ (for the treatment of inhomogeneous boundary conditions, when $\kappa_k$ is a boundary edge).

The two one-sided second-order normal-normal derivatives $\partial_{\nu(z)}\partial_{\nu(z)}$ at $z = z_k \in \mathcal{V}(\Omega) \setminus \mathcal{V}_0(\Omega)$ deserves particular attention: Recall $T_{z,6} \subset H_+(z)$ from (C1) when $\Lambda_k = \Lambda_{z,6}$ is the version of $\partial_{\nu(z)}\partial_{\nu(z)}$ with values taken from $\omega_k \subset H_+(z)$ and $T_{z,7} \subset H_-(z)$ for $\Lambda_k = \Lambda_{z,7}$. The functions $\varphi_k$, $\psi_k$, and the functional (3.2) are established and analyzed in [19]; the only difference to the presented $A(\mathcal{T})$ is their global arrangement.

The design of the weight functions implies $\Lambda_k f = M_k(f)$ for all quintic polynomials $f \in P_5(\mathbb{R}^2)$; cf. [19, equation (4.11)] for the proof. This and a discussion of local and global degrees of freedom lead to the projection property in [19, equation (5.6)], i.e., the idempotence $J^2 = J$. □

*Proof of Theorem 2* (a). Given $T \in \mathcal{T} \cap \widehat{\mathcal{T}}$ and $\hat{v}_A \in A(\widehat{\mathcal{T}})$, the assertion $J\hat{v}_A = \hat{v}_A$ in $T$ is an identity of two quintic polynomials that follows if all 21 degrees of freedom of the Argyris finite element at $T$ coincide. It is important to understand that (in a certain local coordinate system) those degrees of freedom are present in each of the extended Argyris finite element spaces $A(\mathcal{T})$ and $A(\widehat{\mathcal{T}})$ at hand. Let $\Lambda_k = \Lambda_{z,j}$ be one degree of freedom at the node $z \in T$ that belongs to $T$ with global number $k \in \{1, \dots, N\}$ and $(z, j) \in \mathcal{I}$. Since $\widehat{\mathcal{T}}$ is a refinement, this degree of freedom $\Lambda_{z,j}$ is also contained in $A(\widehat{\mathcal{T}})$ with some global number $\widehat{k}$, but the same index $\Lambda_{z,j}$ (in precisely this notation) because $T \in \mathcal{T} \cap \widehat{\mathcal{T}}$.

Since $T \in \mathcal{T}(z) \cap \widehat{\mathcal{T}}(z) \neq \emptyset$, conditions (C1)–(C2) ensure $T_{z,j} \in \mathcal{T}(z) \cap \widehat{\mathcal{T}}(z)$. Moreover, if $z \in \mathcal{V}(\Omega) \setminus \mathcal{V}_0(\Omega)$ is a new interior vertex, then $T \in \mathcal{T}_+(z)$ and $j = 6$ (resp. $T \in \mathcal{T}_-(z)$ and $j = 7$) imply $T_{z,j} \in \mathcal{T}_+(z) \cap \widehat{\mathcal{T}}_+(z)$ (resp. $T_{z,j} \in \mathcal{T}_-(z) \cap \widehat{\mathcal{T}}_-(z)$). The two triangles $T$ and $T_{z,j}$ in $\mathcal{T}(z) \cap \widehat{\mathcal{T}}(z)$ are in general distinct, so let $f$ and $g$ in $P_5(\mathbb{R}^2)$ denote the extension of the quintic polynomials $\widehat{v}_A|_T$ and $\widehat{v}_A|_{T_{z,j}}$ from $T$ and $T_{z,j}$ to the entire plane $\mathbb{R}^2$, respectively. The degree of freedom $\Lambda_{z,j}$ is defined uniquely for the finite element space $A(\mathcal{T})$ and $A(\widehat{\mathcal{T}})$ and so $\Lambda_k f = \Lambda_k \widehat{v}_A = \Lambda_k g$. This follows for a new interior node $z \in \mathcal{V}(\Omega) \setminus \mathcal{V}_0(\Omega)$ and the two-sided second-order normal-normal derivatives (i.e., $j = 6, 7$) from the selection of $T$, $T_{z,j} \in \mathcal{T}_\pm(z) \cap \widehat{\mathcal{T}}_\pm(z)$ in the same half-plane $H_\pm(z)$.

The linear functionals $\Lambda_k$ and $M_k$ coincide in $P_5(\mathbb{R}^2)$ and so $\Lambda_k g = M_k(g) = M_k(\widehat{v}_A)$ with (3.2) in the last identity. The coefficient $M_k(\widehat{v}_A)$ arises in front of the nodal basis function $\varphi_k$ in the definition (3.1) of $J$ and so the duality relation (from Notation 2.4) shows $M_k(\widehat{v}_A) = \Lambda_k(J\widehat{v}_A)$. In conclusion, $\Lambda_k f = \Lambda_k(J\widehat{v}_A)$, which also reads $\Lambda_{z,j}(f - J\widehat{v}_A) = 0$. Recall that this holds for all $j = 1, \dots, m(z)$ and any node $z$ in $T$ except for any new interior vertex $z \in \mathcal{V}(\Omega) \setminus \mathcal{V}_0(\Omega)$ and $j \in \{6, 7\}$. In the later case $\Lambda_{z,6}(f - J\widehat{v}_A) = 0$ if $T \in \mathcal{T}_+(z)$ and $\Lambda_{z,7}(f - J\widehat{v}_A) = 0$ if $T \in \mathcal{T}_-(z)$. For a triangle $T$ with positive distance to the boundary this implies that $L_j(f - (J\widehat{v}_A)|_T) = 0$ for all $j = 1, \dots, 21$, when $(T, P_5(T), (L_1, \dots, L_{21}))$ denotes the quintic Argyris finite element (in the sense of Ciarlet).

In case $\operatorname{dist}(T, \partial\Omega) = 0$ that $T$ hits the boundary $\partial\Omega$ of the domain $\Omega$, not all of the 21 degrees of freedom $L_1, \dots, L_{21}$ are contained in the list $\Lambda_1, \dots, \Lambda_N$ because of the boundary conditions in $H_0^2(\Omega)$. But the realization of the boundary conditions in $J\widehat{v}_A \in A(\mathcal{T})$ and $\widehat{v}_A \in A(\widehat{\mathcal{T}})$ imply

$$D_\alpha(J\widehat{v}_A)(z) = 0 = D_\alpha(\widehat{v}_A)(z) \tag{3.3}$$

at all nodes $z \in \mathcal{V} \cup \mathcal{M}$ on the boundary $\partial\Omega$ with a derivative of order $|\alpha| = 0, 1$ and also for $|\alpha| = 2$ at a corner $z \in V(\Omega)$ of the domain. For a vertex $z \in \mathcal{V}(\partial\Omega) \setminus V(\Omega)$ on a flat part of the boundary $\partial\Omega$, (3.3) holds for the two second-order derivatives $D_\alpha = \partial^2_{\tau\tau}$ and $D_\alpha = \partial^2_{\tau\nu}$ at $z$ with one or two tangential derivatives. The remaining second-order normal-normal derivative $D_\alpha = \partial^2_{\nu\nu} = \Lambda_k = \Lambda_{z,1}$ at $z$ is a global degree of freedom with some number $k \in \{1, \dots, N\}$. Then (3.3) may fail to hold, but $\partial^2_{\nu\nu}(f - J\widehat{v}_A)(z) = \Lambda_{z,1}(f - J\widehat{v}_A) = 0$ follows from the previous analysis for $z \in T$. In summary, all 21 degrees of freedom vanish at $(f - (J\widehat{v}_A)|_T)$ in this case $\operatorname{dist}(T, \partial\Omega) = 0$ as well. Since $(T, P_5(T), (L_1, \dots, L_{21}))$ is a finite element in the sense of Ciarlet, the quintic polynomial $f - (J\widehat{v}_A)|_T$ vanishes in $T$. This concludes the proof of (a). □

*Proof of Theorem 2* (b). The arguments in the proof of [19, Theorem 7.1] show, for any $T$ with neighborhood $\Omega(T)$ and $m = 0, 1, 2$, that

$$|Jf|_{H^m(T)} \lesssim \sum_{k=0}^{2} h_T^{k-m} |f|_{H^k(\Omega(T))} \quad \text{for all } f \in H_0^2(\Omega). \tag{3.4}$$

It is a standard technique [8] to combine a stability property with an approximation property. If $T$ has a positive distance to the boundary $\partial\Omega$, then (3.1) leads to $Jg = g$ in $T$ for any affine $g \in P_1(\Omega(T))$. The choice of $g$ such that the integrals of $f - g$ and $\nabla(f - g)$ over $\Omega(T)$ vanish leads with (3.4) to

$$|f - Jf|_{H^m(T)} \le |f - g|_{H^m(T)} + |J(f - g)|_{H^m(T)} \lesssim \sum_{k=0}^{2} h_T^{k-m} |f - g|_{H^k(\Omega(T))} \lesssim h_T^{2-m} |f|_{H^2(\Omega(T))}$$

with Poincaré inequalities in $\Omega(T)$ in the final step. In the remaining case, $T$ hits the boundary $\partial\Omega$ and at least one edge $E \in \mathcal{E}$ on the boundary also belongs to $\partial\Omega(T) \supset E$. Friedrichs inequalities for $f \in H^2(\Omega(T))$ with $f$ and $\nabla f$ vanishing on $E$ and the shape-regularity of $\mathcal{T}$ imply $h_T^{-2}\|f\|_{L^2((\Omega(T))} + h_T^{-1}|f|_{H^1((\Omega(T))} \lesssim |f|_{H^2(\Omega(T))}$. Since (3.4) holds in this situation as well, this and triangle inequalities result in

$$|f - Jf|_{H^m(T)} \lesssim h_T^{2-m} |f|_{H^2(\Omega(T))} \quad \text{for all } f \in H_0^2(\Omega).$$

This concludes the proof of (b). □

*Proof of Theorem 2* (c). The point is that $Jf$ is well defined for $f \in H^1(\Omega(T))$ and the stability analysis in [19, equation (7.5)] in the present case of triangles $\kappa_j = T_j$ (with $\dim(T_j) = 2$) and for $m = 0$, the right-hand side in (3.4) is replaced by $\sum_{k=0}^{1} h_T^k |f|_{H^k(\Omega(T))}$ without the derivatives of second order. This concludes the proof of (c). □

*Proof of Theorem 2* (d). Given any $f \in H_0^1(\Omega)$, the substitution of $f$ by $f - g$ for some constant $g \in P_0(\Omega(T))$ with $Jg = g$ in $T$ in (c), results in

$$\|f - Jf\|_{L^2(T)} \lesssim \|f - g\|_{L^2(\Omega(T))} + h_T |f|_{H^1(\Omega(T))}. \tag{3.5}$$

If $T$ has a positive distance to the boundary $\partial\Omega$, then (3.1) shows $Jg = g$ in $T$ for the integral mean $g \in P_0(\Omega(T))$ of $f$ in $\Omega(T)$. This and a Poincaré inequality show

$$\|f - g\|_{L^2(\Omega(T))} \lesssim h_T |f|_{H^1(\Omega(T))}. \tag{3.6}$$

If $T$ hits the boundary $\partial\Omega$, then some edge $E \in \mathcal{E}(\partial\Omega)$ on the boundary also belongs to $\partial\Omega(T) \supset E$. Set $g = 0$ and deduce (3.6) from a Friedrichs inequality for $f \in H_0^1(\Omega)$ with $f|_E = 0$ and the shape-regularity of $\mathcal{T}$. The combination of (3.5)–(3.6) concludes the proof of (d). □

# 4 Adaptive Mesh-Refinement Algorithm

The axioms of adaptivity are four abstract properties (A1)–(A4) for an estimator $\eta$ and a distance function $\delta$ to provide optimal convergence rates [6, 10, 12].

## 4.1 Error Estimator and Distance

For any admissible triangulation $\mathcal{T} \in \mathbb{T}$ and the extended Argyris finite element space $A(\mathcal{T})$ and any right-hand side $f \in L^2(\Omega)$, let $u_A \in A(\mathcal{T}) \subset H_0^2(\Omega)$ solve

$$a(u_A, v_A) = F(v_A) \quad \text{for all } v_A \in A(\mathcal{T}) \tag{4.1}$$

with the scalar product $a$ and the linear form $F$ defined by

$$a(u, v) := \int_\Omega \Delta u \, \Delta v \, dx \quad \text{and} \quad F(v) := \int_\Omega f \, v \, dx \quad \text{for all } u, v \in H_0^2(\Omega).$$

The unique discrete solution $u_A \in A(\mathcal{T})$ to (4.1) gives rise to the error estimator $\eta(\mathcal{T}, \bullet) : \mathcal{T} \to (0, \infty)$ defined for any triangle $T \in \mathcal{T}$ with area $|T|$ by the square root of

$$\eta^2(\mathcal{T}, T) := \eta_T^2 := |T|^2 \|f - \Delta^2 u_A\|_{L^2(T)}^2 + \sum_{E \in \mathcal{E}(T) \cap \mathcal{E}(\Omega)} \Big( |T|^{\frac{1}{2}} \|[\Delta u_A]_E\|_{L^2(E)}^2 + |T|^{\frac{3}{2}} \Big\| \Big[ \frac{\partial \Delta u_A}{\partial \nu_E} \Big]_E \Big\|_{L^2(E)}^2 \Big). \tag{4.2}$$

Here and throughout the paper, $[\,\bullet\,]_E := (\bullet|_{T_+}) - (\bullet|_{T_-})$ denotes the jump across the interior edge $E = \partial T_+ \cap \partial T_-$ shared by the two neighboring triangles $T_\pm \in \mathcal{T}$ with unit normal $\nu_E$ (and a fixed orientation).

The analog formula applies to any admissible refinement $\widehat{\mathcal{T}} \in \mathbb{T}(\mathcal{T})$ with a discrete solution $\hat{u}_A$ and so defines $\eta^2(\widehat{\mathcal{T}}, T)$ for all $T \in \widehat{\mathcal{T}}$. The distance of $\mathcal{T}$ and $\widehat{\mathcal{T}}$ is

$$\delta(\mathcal{T}, \widehat{\mathcal{T}}) := |\hat{u}_A - u_A|_{H^2(\Omega)}.$$

## 4.2 Stability (A1) and Reduction (A2)

The first index $\mathcal{T}$ in the estimator $\eta(\mathcal{T}, \bullet) : \mathcal{T} \to (0, \infty)$ refers to an admissible triangulation $\mathcal{T} \in \mathbb{T}$, while the second in (4.2) is a triangle $T$ thereof. It is convenient to abbreviate $\ell^2$ norms over subsets of triangles $\mathcal{M} \subset \mathcal{T}$

by the sum convention

$$\eta(\mathcal{T},\mathcal{M}) := \Big(\sum_{T\in\mathcal{M}} \eta^2(\mathcal{T},T)\Big)^{\frac{1}{2}} \quad \text{for all } \mathcal{M}\subset\mathcal{T}.$$

The empty sum is zero and the full sum is abbreviated $\eta(\mathcal{T}) := \eta(\mathcal{T},\mathcal{T})$. With those conventions, the stability (A1) and the reduction (A2) read as follows.

**Theorem 3** (Discrete Stability and Reduction). *There exist universal constants $\Lambda_1, \Lambda_2$, which depend only on $\mathbb{T}$ through $\mathcal{T}_0$, such that the respective error estimators of* (4.2) *for any admissible refinement $\widehat{\mathcal{T}}$ of $\mathcal{T}\in\mathbb{T}$ satisfy*

$$|\eta(\widehat{\mathcal{T}},\widehat{\mathcal{T}}\cap\mathcal{T}) - \eta(\mathcal{T},\mathcal{T}\cap\widehat{\mathcal{T}})| \le \Lambda_1\,\delta(\mathcal{T},\widehat{\mathcal{T}}), \tag{A1}$$

$$\eta(\widehat{\mathcal{T}},\widehat{\mathcal{T}}\setminus\mathcal{T}) \le 2^{-\frac{1}{4}}\eta(\mathcal{T},\mathcal{T}\setminus\widehat{\mathcal{T}}) + \Lambda_2\,\delta(\mathcal{T},\widehat{\mathcal{T}}). \tag{A2}$$

*Proof.* Those results follow from triangle inequalities as outlined in second-order problems in [6, 10, 12, 13] followed by (discrete) trace inequalities; further details are omitted. □

## 4.3 Discrete Reliability (A3) and Quasi-Orthogonality (A4)

The discrete reliability is key to the a posteriori error analysis and follows with the discrete quasi-interpolation of Theorem 2 plus standard arguments.

**Theorem 4** (Discrete Reliability). *There exists a universal constant $\Lambda_3$, which depends only on $\mathbb{T}$ through $\mathcal{T}_0$, such that the respective error estimators of* (4.2) *for any admissible refinement $\widehat{\mathcal{T}}$ of $\mathcal{T}\in\mathbb{T}$ satisfy*

$$\delta(\mathcal{T},\widehat{\mathcal{T}}) \le \Lambda_3\,\eta(\mathcal{T},\mathcal{T}\setminus\widehat{\mathcal{T}}). \tag{A3}$$

*Proof.* Given the discrete solution $u_A\in A(\mathcal{T})$ (resp. $\hat{u}_A\in A(\widehat{\mathcal{T}})$) to (4.1) with respect to $A(\mathcal{T})$ (resp. $A(\widehat{\mathcal{T}})$) based on $\mathcal{T}$ (resp. $\widehat{\mathcal{T}}$), set $\hat{e} := \hat{u}_A - u_A \in A(\widehat{\mathcal{T}}) + A(\mathcal{T})$ and recall Theorem 1 to deduce $\hat{e}\in A(\widehat{\mathcal{T}})$. Given the quasi-interpolation of Theorem 2, set $e := J\hat{e}\in A(\mathcal{T})$ and $\hat{v} := \hat{e} - e = (1-J)\hat{e}\in A(\widehat{\mathcal{T}})$. The Galerkin property with $e\in A(\mathcal{T})\subset A(\widehat{\mathcal{T}})\subset H_0^2(\Omega)$ leads to

$$\delta^2 := \delta^2(\mathcal{T},\widehat{\mathcal{T}}) = |\hat{e}|^2_{H^2(\Omega)} = a(\hat{e},\hat{e}) = a(\hat{e},\hat{v}) = F(\hat{v}) - a(u_A,\hat{v}).$$

A piecewise integration by parts twice proves that the last residual term $F(\hat{v}) - a(u_A,\hat{v})$ is equal to

$$\sum_{T\in\mathcal{T}}\int_T (f-\Delta^2 u_A)\hat{v}\,dx - \sum_{E\in\mathcal{E}(\Omega)}\int_E\Big(\frac{\partial\hat{v}}{\partial\nu_E}[\Delta u_A]_E - \hat{v}\Big[\frac{\partial\Delta u_A}{\partial\nu_E}\Big]_E\Big)\,ds.$$

The definition of $\eta_T \equiv \sqrt{\eta_T^2}$ in (4.2), and Cauchy inequalities show that this is bounded from above by a generic constant $\lesssim 1$ times

$$\sum_{T\in\mathcal{T}}\eta_T\Big(|T|^{-1}\|\hat{v}\|_{L^2(T)} + |T|^{-\frac{1}{4}}\Big\|\frac{\partial\hat{v}}{\partial\nu_E}\Big\|_{L^2(\partial T)} + |T|^{-\frac{3}{4}}\|\hat{v}\|_{L^2(\partial T)}\Big).$$

A (discrete) trace inequality and the approximation properties for $\hat{v} := (I-J)\hat{e}$ show with the shape-regularity $|T|\approx h_T^2$ that

$$|T|^{-1}\|\hat{v}\|_{L^2(T)} + |T|^{-\frac{1}{4}}\Big\|\frac{\partial\hat{v}}{\partial\nu_E}\Big\|_{L^2(\partial T)} + |T|^{-\frac{3}{4}}\|\hat{v}\|_{L^2(\partial T)} \lesssim |\hat{e}|_{H^2(\Omega(T))}.$$

On the other hand, this contribution vanishes for all $T\in\widehat{\mathcal{T}}\cap\mathcal{T}$ because $\hat{v}=0$ therein. The finite overlap of the neighborhoods $(\Omega(T) : T\in\mathcal{T})$ leads to the bound $\delta^2 \lesssim (\sum_{T\in\mathcal{T}\setminus\widehat{\mathcal{T}}}\eta_T^2)^{\frac{1}{2}}|\hat{e}|_{H^2(\Omega)} = \eta(\mathcal{T},\mathcal{T}\setminus\widehat{\mathcal{T}})\,\delta$. □

The nestedness property of the extended Argyris finite element spaces implies orthogonality.

**Theorem 5** (Quasi-Orthogonality). *Any sequence $(\mathcal{T}_k)_{k\in\mathbb{N}_0}$ of successive admissible refinements of $\mathcal{T}_0$ satisfies, for all $\ell\in\mathbb{N}_0$, that*

$$\sum_{k=\ell}^{\infty}\delta^2(\mathcal{T}_k,\mathcal{T}_{k+1}) \le \Lambda_4\,\eta^2(\mathcal{T}_\ell) \tag{A4}$$

*with the universal constant $\Lambda_4 := \Lambda_3^2$ and $\Lambda_3$ from the previous theorem.*

*Proof.* Based on Theorem 1, the successive application of the Pythagoras theorem leads to

$$\delta^2(\mathcal{T}_\ell, \mathcal{T}_{\ell+m+1}) = \sum_{k=\ell}^{\ell+m} \delta^2(\mathcal{T}_k, \mathcal{T}_{k+1}).$$

This and Theorem 4 imply, for all $m, \ell \in \mathbb{N}_0$, that

$$\sum_{k=\ell}^{\ell+m} \delta^2(\mathcal{T}_k, \mathcal{T}_{k+1}) \leq \Lambda_3^2\, \eta^2(\mathcal{T}_\ell).$$

The proof concludes with $m \to \infty$. □

## 4.4 Adaptive Algorithm and Optimal Convergence Rates

The subsequent adaptive mesh-refinement algorithm is the nowadays standard strategy.

**Adaptive Algorithm.** The algorithm has the input parameter $0 < \Theta < 1$ and the initial mesh $\mathcal{T}_0$. Given any level $\ell = 0, 1, \dots,$ do

(a) compute $u_\ell \in A(\mathcal{T}_\ell)$ with $a(u_\ell, v_\ell) = F(v_\ell)$ for all $v_\ell \in A(\mathcal{T}_\ell)$.

(b) compute error estimator $\eta(\mathcal{T}_\ell, T)$ for any $T \in \mathcal{T}_\ell$ by (4.2) (with $u_\ell$ replacing $u_A$ on the right-hand side).

(c) mark $\mathcal{M}_\ell \subset \mathcal{T}_\ell$ with (almost) minimal cardinality such that

$$\Theta\, \eta(\mathcal{T}_\ell) \leq \eta(\mathcal{T}_\ell, \mathcal{M}_\ell).$$

(d) refine all $\mathcal{M}_\ell$ in $\mathcal{T}_\ell$ and create new triangulation $\mathcal{T}_{\ell+1}$ od.

Output are the discrete solutions $u_\ell$, the error estimators $\eta_\ell := \eta(\mathcal{T}_\ell)$, and the triangulations $(\mathcal{T}_\ell)$ for all $\ell = 0, 1, 2, \dots$ .

Let $\mathbb{T}(N)$ denote the set of all admissible triangulations $\mathcal{T} \in \mathbb{T}$ with number of triangles $|\mathcal{T}| \leq |\mathcal{T}_0| + N$. The data oscillation $\operatorname{osc}(f, \mathcal{T})$ is first defined by

$$\operatorname{osc}(f, T) := |T| \|f - f_T\|_{L^2(T)}$$

for the integral mean $f_T$ of $f \in L^2(\Omega)$ in $T \in \mathcal{T} \in \mathbb{T}$ and second as

$$\operatorname{osc}(f, \mathcal{T}) := \Big( \sum_{T \in \mathcal{T}} \operatorname{osc}(f, T)^2 \Big)^{\frac{1}{2}} \quad \text{for } \mathcal{T} \in \mathcal{T}.$$

The interpretation of the subsequent consequence is that of optimal convergence rates: If the right-hand side is smaller than infinity, then there is an optimal way to achieve the convergence rate $s$ and the following theorem asserts that also the outcome of the adaptive algorithm converges with the rate $s$.

**Theorem 6** (Optimal Rates). *There exists a universal constant $0 < \Theta_0 < 1$ such that for all $0 < s < \infty$ there exists some $\Lambda_{\mathrm{eq}}$ such that the output of the adaptive algorithm with $\Theta < \Theta_0$ satisfies*

$$\begin{aligned} &\sup_{\ell \in \mathbb{N}_0} (1 + |\mathcal{T}_\ell| - |\mathcal{T}_0|)^{-s} (\|u - u_\ell\|_{H^2(\Omega)} + \operatorname{osc}(f, \mathcal{T}_\ell)) \\ &\quad \leq \Lambda_{\mathrm{eq}} \sup_{N \in \mathbb{N}_0} (1 + N)^{-s} \min_{\mathcal{T} \in \mathbb{T}(N)} \Big( \min_{u_A \in A(\mathcal{T})} \|u - u_A\|_{H^2(\Omega)} + \operatorname{osc}(f, \mathcal{T}) \Big). \end{aligned}$$

*The constant $\Lambda_{\mathrm{eq}}$ solely depends on $\mathbb{T}$ through $\mathcal{T}_0$ and on $\Theta_0$ and $s$.*

*Proof.* This follows from the axioms of adaptivity [6, 10, 12] that (A1)–(A4) and sufficiently small $\Theta_0$ lead to

$$\sup_{\ell \in \mathbb{N}_0} (1 + |\mathcal{T}_\ell| - |\mathcal{T}_0|)^{-s}\, \eta_\ell \lesssim \sup_{N \in \mathbb{N}_0} (1 + N)^{-s} \min_{\mathcal{T} \in \mathbb{T}(N)} \eta(\mathcal{T}).$$

This, the best-approximation property of the scheme, and the equivalence of the error estimator plus data oscillations $\eta(\mathcal{T}) + \operatorname{osc}(f, \mathcal{T})$ to the error plus data oscillations $\|u - u_A\|_{H^2(\Omega)} + \operatorname{osc}(f, \mathcal{T})$ lead to the assertion. □

# 5 Multigrid V-Cycle Algorithm

The first subsection introduces the overall V-cycle algorithm with some smoothing operator that requires a closer look at the degrees of freedom for the extended Argyris FEM in the second subsection.

## 5.1 The Multigrid V-Cycle Algorithm

Let $\mathcal{T}_0, \mathcal{T}_1, \mathcal{T}_2, \dots$ be a sequence of nested conforming finite element triangulations of the domain $\Omega$ such that $\mathcal{T}_\ell$ is a refinement of $\mathcal{T}_{\ell-1}$ by the NVB algorithm and let $A(\mathcal{T}_\ell)$ be the extended Argyris finite element space of the previous section for any $\ell \in \mathbb{N}_0$. Define the discrete operator $\mathcal{A}_\ell : A(\mathcal{T}_\ell) \to A(\mathcal{T}_\ell)$ and the discrete right-hand side $\mathcal{F}_\ell \in A(\mathcal{T}_\ell)$ by

$$(\mathcal{A}_\ell w, v)_{L^2(\Omega)} := a(w, v) \quad \text{and} \quad (\mathcal{F}_\ell, v)_{L^2(\Omega)} := F(v) \quad \text{for all } w, v \in A(\mathcal{T}_\ell)$$

with the inner product $(\cdot, \cdot)_{L^2(\Omega)}$ in $L^2(\Omega)$. This defines the discrete problem on the level $\ell \in \mathbb{N}_0$: Seek $u_\ell \in A(\mathcal{T}_\ell)$ with

$$\mathcal{A}_\ell u_\ell = \mathcal{F}_\ell. \tag{5.1}$$

The prolongation and restriction operators are defined in terms of the $L^2$ inner product projection operator $\mathcal{Q}_\ell : H_0^2(\Omega) \to A(\mathcal{T}_\ell)$ and the energy inner product projection operator $\mathcal{P}_\ell : H_0^2(\Omega) \to A(\mathcal{T}_\ell)$ defined by

$$(\mathcal{Q}_\ell w, v_\ell)_{L^2(\Omega)} = (w, v_\ell)_{L^2(\Omega)} \quad \text{and} \quad a(\mathcal{P}_\ell w, v_\ell) = a(w, v_\ell) \tag{5.2}$$

for all $w \in H_0^2(\Omega)$ and all $v_\ell \in A(\mathcal{T}_\ell)$, $\ell \in \mathbb{N}_0$.

The local Gauss–Seidel relaxation of Section 5.2 defines a smoothing operator $\mathcal{R}_\ell : A(\mathcal{T}_\ell) \to A(\mathcal{T}_\ell)$ based on a selection of $\tilde{n}_\ell$ particular nodal basis functions $\phi_\ell^1, \dots, \phi_\ell^{\tilde{n}_\ell}$ in $A(\mathcal{T}_\ell)$ selected in Section 5.2 below. This specifies the local projection operator $P_\ell^j : H_0^2(\Omega) \to A(\mathcal{T}_\ell)$ onto the one-dimensional span of $\phi_\ell^j$ for all $j = 1, \dots, \tilde{n}_\ell$ and $\ell \in \mathbb{N}$ by

$$P_\ell^j w := \left( \frac{a(w, \phi_\ell^j)}{a(\phi_\ell^j, \phi_\ell^j)} \right) \phi_\ell^j \quad \text{for all } w \in H_0^2(\Omega). \tag{5.3}$$

The smoothing operator $\mathcal{R}_\ell : A(\mathcal{T}_\ell) \to A(\mathcal{T}_\ell)$ reads

$$\mathcal{R}_\ell := \left( I - \prod_{k=1}^{\tilde{n}_\ell} (I - P_\ell^k) \right) \mathcal{A}_\ell^{-1}. \tag{5.4}$$

Given an initial vector $u_\ell^{(0)} \in A(\mathcal{T}_\ell)$, the standard V-cycle multigrid algorithm [5] solves system (5.1) by the iterative method

$$u_\ell^{(k)} = u_\ell^{(k-1)} + \mathcal{B}_\ell(\mathcal{F}_\ell - \mathcal{A}_\ell u_\ell^{(k-1)}) \quad \text{for } k \in \mathbb{N} \tag{5.5}$$

with an approximate inverse $\mathcal{B}_\ell$ of $\mathcal{A}_\ell$ recursively defined in a V-cycle.

**Algorithm** (V-Cycle). Let $\mathcal{B}_0 := \mathcal{A}_0^{-1}$. For $\ell = 1, 2, 3, \dots$ and $g \in A(\mathcal{T}_\ell)$,
(1) Pre-smoothing: $w_1 := \mathcal{R}_\ell g$.
(2) Correction: $w_2 := w_1 + \mathcal{B}_{\ell-1}\mathcal{Q}_{\ell-1}(g - \mathcal{A}_\ell w_1)$.
(3) Post-smoothing: $\mathcal{B}_\ell g := w_2 + \mathcal{R}_\ell^T(g - \mathcal{A}_\ell w_2)$.
Then $\mathcal{B}_\ell$ is a uniform approximative inverse of $\mathcal{A}_\ell$ in the operator norm

$$\|I - \mathcal{B}_\ell \mathcal{A}_\ell\|_{\mathcal{A}} := \sqrt{\sup_{v_\ell \in A(\mathcal{T}_\ell)\setminus\{0\}} \frac{a(v_\ell - \mathcal{B}_\ell \mathcal{A}_\ell v_\ell, v_\ell)}{a(v_\ell, v_\ell)}}.$$

**Theorem 7.** *Under the assumptions of this subsection and with the selection of the basis functions $\phi_\ell^1, \dots, \phi_\ell^{\tilde{n}_\ell}$ in $A(\mathcal{T}_\ell)$ from Section 5.2 below, $\mathcal{A}_\ell$ and $\mathcal{B}_\ell$ satisfy*

$$\sup_{\ell \in \mathbb{N}_0} \|I - \mathcal{B}_\ell \mathcal{A}_\ell\|_{\mathcal{A}} \le \frac{c_0}{1 + c_0} < 1$$

*for some positive $c_0$ bounded in terms of universal constants and the reduced elliptic regularity of the polygonal domain $\Omega$.*

The proofs start after the description of the smoothing operator. Theorem 7 implies that the simple iterative scheme (5.5) converges with rate smaller than or equal to some $q < 1$ independent of the number of levels, triangles, or the dimension of the linear system. It also implies uniform convergence of preconditioned schemes, e.g., the conjugate gradient scheme.

### 5.2 Partial Gauss–Seidel Relaxation

The smoothing operator $\mathcal{R}_\ell$ is a Gauss–Seidel relaxation performed for *new* degrees of freedom and their neighbors, Recall that the finite element space $A(\mathcal{T}_\ell)$ on the level $\ell \in \mathbb{N}_0$ involves nodal basis functions $\psi^\ell_{z,j}$ and degrees of freedom $\Lambda^\ell_{z,j}$ associated to a node $z \in \mathcal{N}_\ell \subset \mathcal{M}_\ell \cup \mathcal{V}_\ell$. The set $\mathcal{V}_\ell$ (resp. $\mathcal{E}_\ell$ and $\mathcal{M}_\ell$) of vertices (resp. edges and edges' midpoints) in the triangulation $\mathcal{T}_\ell$ defines a superset of the set of nodes $\mathcal{N}_\ell$ as defined in Notation 2.2 (where $\mathcal{V}_\ell, \mathcal{E}_\ell, \mathcal{M}_\ell, \mathcal{N}_\ell, \mathcal{T}_\ell$ substitute $\mathcal{V}, \mathcal{E}, \mathcal{M}, \mathcal{N}, \mathcal{T}$). Each node $z \in \mathcal{N}_\ell$ locates $m(z, \ell) \in \{1, 6, 7\}$ degrees of freedom $\Lambda^\ell_{z,j}$ which is a derivative of order $\mu(z, j, \ell) \in \{0, 1, 2\}$ as introduced in Section 2.2 and a nodal basis function $\psi^\ell_{z,j}$ for $j = 1, \dots, m(z, \ell)$.

The neighboring triangles $\mathcal{T}_\ell(z) := \{T \in \mathcal{T}_\ell : z \in T\}$ are affected under mesh-refinement and the selection of $\psi^\ell_{z,j}$ to $\widetilde{A}(\mathcal{T}_\ell) \subset A(\mathcal{T}_\ell)$ below is characterized by the condition

$$\mathcal{T}_\ell(z) \neq \mathcal{T}_{\ell-1}(z). \tag{5.6}$$

In other words, (5.6) is the absence of $\mathcal{T}_\ell(z) = \mathcal{T}_{\ell-1}(z)$ and involves some slightly pathological situations: Given any $z \in \mathcal{M}_\ell \cup \mathcal{V}_\ell$ the condition $\mathcal{T}_\ell(z) = \mathcal{T}_{\ell-1}(z)$ is pointless if $\mathcal{T}_{\ell-1}(z)$ does not exist, i.e., either $\ell = 0$ or $z \in \mathcal{M}_\ell \setminus \mathcal{M}_{\ell-1}$ for $\ell \in \mathbb{N}$. In those cases the condition $\mathcal{T}_\ell(z) \neq \mathcal{T}_{\ell-1}(z)$ is regarded as fulfilled. In the remaining situations, i.e., $z \in \mathcal{M}_{\ell-1} \cup \mathcal{V}_{\ell-1}$ and $\ell \in \mathbb{N}$, the neighborhood $\mathcal{T}_{\ell-1}(z)$ is defined and then $\mathcal{T}_\ell(z) \neq \mathcal{T}_{\ell-1}(z)$ is equivalent to the failure of one of the inclusions $\mathcal{T}_\ell(z) \subset \mathcal{T}_{\ell-1}(z)$ and $\mathcal{T}_{\ell-1}(z) \subset \mathcal{T}_\ell(z)$. This concludes the comments on (5.6) and allows (in the notation of the previous subsections) for the definition

$$\widetilde{A}(\mathcal{T}_\ell) := \operatorname{span}\{\psi^\ell_{z,j} \in A(\mathcal{T}_\ell) : z \in \mathcal{M}_\ell \cup \mathcal{V}_\ell,\ 1 \le j \le m(z, \ell),\ (5.6) \text{ holds}\}$$

for $\ell \in \mathbb{N}$. (Notice that $\widetilde{A}(\mathcal{T}_\ell)$ depends on $\mathcal{T}_{\ell-1}$ and $\mathcal{T}_\ell$ by (5.6), but the dependence on $\mathcal{T}_{\ell-1}$ is suppressed in the notation for brevity.)

Throughout the present paper, let $\phi^1_\ell, \dots, \phi^{\tilde{n}_\ell}_\ell$ be some enumeration of the selected shape functions $\psi^\ell_{z,1}, \dots, \psi^\ell_{z,m(z,\ell)}$ for all $z \in \mathcal{N}_\ell$ for $\ell \in \mathbb{N}$ with (5.6) with the linear hull

$$\widetilde{A}(\mathcal{T}_\ell) = \operatorname{span}\{\phi^1_\ell, \dots, \phi^{\tilde{n}_\ell}_\ell\}.$$

The nodal basis functions $\phi^1_\ell, \dots, \phi^{\tilde{n}_\ell}_\ell$ enter the definition of the smoother in (5.3)–(5.4) for $\ell \in \mathbb{N}$ ($\ell = 0$ involves direct solve and $\widetilde{A}(\mathcal{T}_0)$ is not utilized).

The multigrid V-cycle algorithm is an adaptation of [27] (the Courant FEMs for second order problems) to the extended Argyris FEM. The multigrid V-cycle algorithm of [14] can also be generalized to the extended Argyris FEM.

## 6 Preliminaries in the Multigrid Analysis

This section starts with three preliminary subsections on the $L^2$ norm of (extended) Argyris finite element functions, on the NVB mesh-refinement, and on uniform meshes before the main result is established in Section 7.

### 6.1 Lumped Mass Matrix

Note that each degree of freedom $\Lambda_j \in A(\mathcal{T})^*$ in the extended Argyris finite element space $A(\mathcal{T})$ of dimension $N := \operatorname{diam}(A(\mathcal{T}))$ (and its dual $A(\mathcal{T})^*$) for an admissible triangulation $\mathcal{T} \in \mathbb{T}$ is associated with a node $z_j$

(a vertex of a triangle or the midpoint of an edge) and some (directional) derivative of order $\mu_j \in \{0, 1, 2\}$. The 21 degrees of freedom that act on some triangle $T \in \mathcal{T}$ form the entries in the index list

$$I(T) := \{j \in \{1, \dots, N\} : z_j \in T\}$$

of their global numbers; recall that the triangle $T$ is the compact convex hull of its vertices and so its edges' midpoints belong to $T$. Any function $v_A \in A(\mathcal{T})$ can be written in $T \in \mathcal{T}$ as

$$v_A = \sum_{j \in I(T)} (\Lambda_j v_A)\,\varphi_j \quad \text{a.e. in } T \tag{6.1}$$

with the real coefficient $\Lambda_j v_A \equiv \Lambda_j(v_A)$ of the (global) nodal basis function $\varphi_j$. Recall that the linear functionals $\Lambda_1, \dots, \Lambda_N$ (the degrees of freedom) and the (dual) nodal basis functions $\varphi_1, \dots, \varphi_N \in A(\mathcal{T})$ satisfy $\Lambda_j \varphi_k = \delta_{jk}$ for all $j, k = 1, \dots, N$.

The extended as well as the classical Argyris finite elements are *not* affine equivalent (although quasi-affine [15]) and so the scaling and the lumped representation of the $L^2$ norms are *not* immediate. For admissible triangulations, however, the subsequent equivalent constants solely depend on $\mathcal{T}_0$. Recall that $\mu_j$ denotes the order of the derivative $\Lambda_j$ at the node $z_j$.

**Lemma 1** (Lumped $L^2$ Norm). *Adopt the aforementioned notation for $v_A \in A(\mathcal{T})$ in $T \in \mathcal{T} \in \mathbb{T}$ with the representation* (6.1). *Then*

$$\|v_A\|^2_{L^2(T)} \approx \sum_{j \in I(T)} |T|^{1+\mu_j} (\Lambda_j v_A)^2.$$

*Proof.* Any triangle $T \in \mathcal{T} \in \mathbb{T}$ with edge $E$ of length $t := |E| > 0$ can be translated so that its first vertex $P$ coincides with the origin and so there is one reference triangle $T^{\text{ref}} := \operatorname{conv}\{0, Q_2, Q_3\}$ with $Q_2, Q_3 \in \mathbb{R}^2$ and $|Q_2| = 1$ so that $T = P + t\,T^{\text{ref}}$. This translation and scaling is possible for any triangle and leads to a set of possible reference triangles $\mathcal{T}^{\text{ref}}$ so that, for any $T \in \mathcal{T} \in \mathbb{T}$, there exists some $T^{\text{ref}} \in \mathcal{T}^{\text{ref}}$, $t > 0$, and $P \in \mathbb{R}^2$ with $T = P + t\,T^{\text{ref}}$. The point is that the NVB leads to admissible triangulations with at most $8|\mathcal{T}_0|$ interior different angles and the set $\mathcal{T}^{\text{ref}}$ has finite cardinality at most $5|\mathcal{T}_0|$.

The degrees of freedom simply translate, but the scaling depends on the order $\mu$ of the degree of freedom. Let the 21 degrees of freedom $\Lambda_1^{\text{ref}}, \dots, \Lambda_{21}^{\text{ref}}$ and the dual basis functions $\varphi_1^{\text{ref}}, \dots, \varphi_{21}^{\text{ref}}$ on some reference triangle $T^{\text{ref}} \in \mathcal{T}^{\text{ref}}$ be enumerated so that the order of the first three degrees of freedom is zero, the order of those with index four to twelve is one and the remaining are of order two. Without loss of generality, let $I(T) = \{1, \dots, 21\}$, so that the 21 degrees of freedom in (6.1) related to $T$ are given as for each $j$ as follows: $\Lambda_j$ is equal to $\Lambda_j^{\text{ref}}$ with an evaluation at a shifted node and the nodal basis function $\varphi_j$ reads $\varphi_j(x) = t^{\mu_j} \varphi_j^{\text{ref}}(\frac{x-P}{t})$ for all $x \in T$. Consequently, the $21 \times 21$ mass matrix $M(T)$ of $T$ has the components

$$M(T)_{jk} := \int_T \varphi_j(x)\varphi_k(x)\,dx = t^{\mu_j+\mu_k+2} M^{\text{ref}}_{jk} \tag{6.2}$$

with $M^{\text{ref}}_{jk} := \int_{T^{\text{ref}}} \varphi_j^{\text{ref}}(\xi)\varphi_k^{\text{ref}}(\xi)\,d\xi$ for all $j, k = 1, \dots, 21$. Each $T^{\text{ref}}$ is associated to one SPD mass matrix $M^{\text{ref}}$ with positive eigenvalues. (The definiteness follows from the $L^2$ norm representation exploited below.) Since there is solely a finite number of reference triangles in $\mathcal{T}^{\text{ref}}$, there exist the minimum $\underline{\lambda}$ (resp. maximum $\overline{\lambda}$) of all eigenvalues of a mass matrix $M^{\text{ref}}$ associated to some $T^{\text{ref}}$. The positive constants $\underline{\lambda}$ and $\overline{\lambda}$ solely depend on $\mathcal{T}_0$, hence they are generic and denoted as $\underline{\lambda} \approx 1 \approx \overline{\lambda}$.

To prove the lemma, suppose $T \in \mathcal{T} \in \mathbb{T}$ and let $T^{\text{ref}}$ be associated as before with $I(T) = \{1, \dots, 21\}$ and the representation (6.1). The Rayleigh quotients with $M^{\text{ref}}$ belong to the convex hull of its eigenvalues and so

$$\underline{\lambda}|x|^2 \le x \cdot M^{\text{ref}} x \le \overline{\lambda}|x|^2 \tag{6.3}$$

holds for all $x \in \mathbb{R}^{21}$ with the Euclid length $|x|$. Then $x_j := t^{1+\mu_j} \Lambda_j v_A$ for $j = 1, \dots, 21$ leads in (6.2)–(6.3) to some $\lambda$ with $\underline{\lambda} \le \lambda \le \overline{\lambda}$ and

$$\|v_A\|^2_{L^2(T)} = \sum_{j,k \in I(T)} (\Lambda_j v_A)(\Lambda_k v_A) M(T)_{jk} = x \cdot M^{\text{ref}} x = \lambda |x|^2.$$

Since the scaling factor $t$ satisfies $|T| = t^2 |T^{\text{ref}}|$ and each of the finite number of reference triangles satisfies $|T^{\text{ref}}| \approx 1$, this concludes the proof. $\square$

## 6.2 Re-summation

The analysis of multilevel methods requires global and local arguments. The re-summation lemma of this section links the locally refined triangulations from NVB with a structured uniform partition.

### 6.2.1 Three Constants in NVB

Given the initial triangulation $\mathcal{T}_0$ and the set $\mathbb{T}$ of admissible triangulations from Section 2.1, the following three bounds $M_1, M_2, M_3$ are universal in that they depend exclusively on $\mathcal{T}_0$.

Some concepts are required to define $M_1, M_2, M_3$. The first is the set of triangles $\mathcal{R}_j(\mathcal{T}, T)$ of $j$ layers around some $T \in \mathcal{T}$ in a triangulation $\mathcal{T} \in \mathbb{T}$: Given $\mathcal{T} \in \mathbb{T}$ define $\mathcal{R}_0(\mathcal{T}, T) := \{T\}$ and, for $j = 0, 1$, successively define

$$\mathcal{R}_{j+1}(\mathcal{T}, T) := \{K \in \mathcal{T} : \operatorname{dist}(T', K) = 0 \text{ for some } T' \in \mathcal{R}_j(\mathcal{T}, T)\} \tag{6.4}$$

and recall $\Omega(T) = \operatorname{int}(\bigcup \mathcal{R}_1(\mathcal{T}, T))$ in Theorem 2.

The second concept is the refinement level $\ell(T)$ of a triangle $T \in \bigcup \mathbb{T} \equiv \{T \text{ triangle} : \text{there exists } \mathcal{T} \in \mathbb{T}, T \in \mathcal{T}\}$. Since there exists a unique initial triangle $K \in \mathcal{T}_0$ with $T \subset K$, the quotient $\frac{|T|}{|K|} = 2^{-\ell(T)}$ of the areas $|T|$ and $|K|$ of $T$ and $K$ belongs to the set $\{1, \frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \dots\}$ and defines the refinement level $\ell(T)$ of $T$. For a set $\mathcal{M} \subset \bigcup \mathbb{T}$ of triangles, let $|\mathcal{M}|$ denote its cardinality and let $\ell(\mathcal{M}) := \{\ell(M) : M \in \mathcal{M}\}$ denote the set of the refinement levels attained in that set $\mathcal{M}$; $\min \ell(\mathcal{M}) \equiv \min_{M \in \mathcal{M}} \ell(M)$ denotes the minimal level.

**Lemma 2** (Universal Bounds). *The three positive constants*

$$\begin{aligned}
M_1 &:= \max_{\mathcal{T} \in \mathbb{T}} \max_{T \in \mathcal{T}} |\mathcal{R}_1(\mathcal{T}, T)|,\\
M_2 &:= \max_{\mathcal{T} \in \mathbb{T}} \max_{T \in \mathcal{T}} (\max \ell(\mathcal{R}_1(\mathcal{T}, T)) - \min \ell(\mathcal{R}_1(\mathcal{T}, T))),\\
M_3 &:= \max_{T \in \bigcup \mathbb{T}} |\{\mathcal{R}_2(\mathcal{T}, T) : \exists \mathcal{T} \in \mathbb{T},\ T \in \mathcal{T}\}|
\end{aligned}$$

*exclusively depend on $\mathcal{T}_0$; this is abbreviated by $M_1, M_2, M_3 \approx 1$.*

*Proof.* Recall that there are a finite number of different interior angles (at most $7|\mathcal{T}_0|$) possible in the triangles $\bigcup \mathbb{T}$; in particular there is a smallest $\omega_0 > 0$. This leads to a very coarse bound $M_1 \le \frac{5\pi}{\omega_0} - 2$. A closer look in the initial condition and NVB proves that two triangles $T_+, T_- \in \mathcal{T} \in \mathbb{T}$ that share an edge $E = \partial T_+ \cap \partial T_-$ have a level distance $\ell(T_+) - \ell(T_-) \in \{-1, 0, 1\}$ at most one (as already observed in [2]). Consequently, $M_2 \le M_1$.

For each fixed triangle $K \in \bigcup \mathbb{T}$, and any triangulation $\mathcal{T} \in \mathbb{T}$ containing $K \in \mathcal{T}$, the patch $\mathcal{R}_2(\mathcal{T}, K)$ is formed by at most $M_1^2$ triangles and each angle in each of them has a value from a finite set of (at most $7 \times |\mathcal{T}_0|$) different angles. Since $K$ is fixed, the number of different patches is bounded, i.e., the cardinality $|\{\mathcal{R}_2(\mathcal{T}, K) : \exists \mathcal{T} \in \mathbb{T}, K \in \mathcal{T}\}| \le M_3$ and $M_3 \approx 1$ is independent of $K$. □

### 6.2.2 Uniform Mesh-Refinement

Recall that the initial triangulation $\mathcal{T}_0$ from Section 2.1 satisfies the initial condition: The refinement edge $E(T) := \operatorname{conv}\{P_1, P_3\}$ associated to any triangle $T = \operatorname{conv}\{P_1, P_2, P_3\}$ given by $(P_1, P_2, P_3)$ (its vertices in a fixed order) satisfies that either $E = E(T_1) = E(T_2)$ or $E(T_1) \ne E \ne E(T_2)$ for any $T_1, T_2 \in \mathcal{T}_0$ with the common edge $E = \partial T_1 \cap \partial T_2$ (see [10, 24]). Uniform NVB means one bisection for each triangle and leads to the regular triangulation $\mathcal{T}_1^{\text{uniform}} := \{T' \in \operatorname{bisec}(T) : T \in \mathcal{T}_0\}$ of bisected triangles. It also follows that the triangles in $\mathcal{T}_1^{\text{uniform}} \in \mathbb{T}$ satisfy the initial condition as well and so the uniform refinement can be repeated:

$$\mathcal{T}_{m+1}^{\text{uniform}} := \{T' \in \operatorname{bisec}(T) : T \in \mathcal{T}_m^{\text{uniform}}\} \quad \text{for all } m \in \mathbb{N}_0$$

defines a sequence of quasi-uniform admissible triangulation. Since $\mathcal{T}_0$ is fixed throughout this paper, the shape-regular triangulation $\mathcal{T}_\ell^{\text{uniform}}$ is associated with a global mesh-size $\approx 2^{-\frac{m}{2}}$ in that each triangle

$T \in \mathcal{T}_m^{\mathrm{uniform}}$ of area $|T|$ and level $\ell(T) = m$ satisfies

$$h_0 2^{-\frac{m}{2}} \le |T|^{\frac{1}{2}} \le H_0 2^{-\frac{m}{2}} \quad \text{for all } T \in \mathcal{T}_m^{\mathrm{uniform}} \text{ and all } m \in \mathbb{N}_0$$

for the minimal resp. maximal initial mesh-size $h_0 := \min\{|T|^{\frac{1}{2}} : T \in \mathcal{T}_0\} \le H_0 := \max\{|T|^{\frac{1}{2}} : T \in \mathcal{T}_0\}$.

### 6.2.3 Re-summation

Adopt the notation of the previous two subsections and consider a sequence $(\mathcal{T}_\ell)_{\ell\in\mathbb{N}_0}$ of admissible triangulations obtained from $\mathcal{T}_0$ by successive admissible refinements $\mathcal{T}_{\ell+1} \in \mathbb{T}(\mathcal{T}_\ell)$ for any $\ell = 0, 1, 2, \dots$ . Recall the definition (6.4) of the neighborhood of a triangle $T \in \mathcal{T}_\ell$ and abbreviate

$$\mathcal{R}_{j,\ell}(T) := \mathcal{R}_j(\mathcal{T}_\ell, T) \quad \text{for all } T \in \mathcal{T}_\ell \text{ and all } \ell \in \mathbb{N}_0 \tag{6.5}$$

for $j = 1, 2$. For each $\ell \in \mathbb{N}_0$, define the two subsets $\mathcal{T}_\ell^\pm$ of $\mathcal{T}_\ell$ by

$$\mathcal{T}_\ell^+ := \{T \in \mathcal{T}_\ell : \mathcal{R}_{1,\ell}(T) \ne \mathcal{R}_{1,\ell+1}(T)\}, \tag{6.6}$$
$$\mathcal{T}_\ell^- := \{T \in \mathcal{T}_\ell : \mathcal{R}_{1,\ell}(T) \ne \mathcal{R}_{1,\ell-1}(T)\}. \tag{6.7}$$

For clarity set $\mathcal{T}_0^- := \mathcal{T}_0$ (for $\mathcal{T}_{-1}$ does not exist and $\mathcal{R}_{1,0}(T) = \mathcal{R}_{1,-1}(T)$ is impossible). The subsequent re-summation lemma allows the transfer of terms on the unstructured (and locally refined triangulations) $(\mathcal{T}_\ell)_{\ell\in\mathbb{N}_0}$ to uniform triangulations $(\mathcal{T}_m^{\mathrm{uniform}})_{m\in\mathbb{N}_0}$. It is written for fairly general summands to underline that it is only a re-arrangement of the sums: suppose that $s^2(j, T)$ is a non-negative real number defined for any triangle $T \in \bigcup\mathbb{T}$ and any $j \in \mathbb{N}_0$ with the sum convention

$$s^2(j, \mathcal{M}) := \sum_{M\in\mathcal{M}} s^2(j, M)$$

for any finite set $\mathcal{M}$ of admissible triangles; set $(\,\bullet\,)_+ := \max\{0, \bullet\}$.

**Lemma 3** (Re-summation). *Under the above notation,*

$$\sum_{\ell=0}^{L} \sum_{K\in\mathcal{T}_\ell^\pm} s^2(\min \ell(\mathcal{R}_{1,\ell}(K)), \mathcal{R}_{1,\ell}(K)) \le M_1 M_3 \sum_{m=0}^{\infty} \sum_{j=(m-M_2)_+}^{m} s^2(j, \mathcal{T}_m^{\mathrm{uniform}})$$

*holds for all $L \in \mathbb{N}_0$ ($\mathcal{T}_\ell^\pm$ in the second sum of the left-hand side denotes either $\mathcal{T}_\ell^+$ from* (6.6) *or $\mathcal{T}_\ell^-$ from* (6.7)*, the assertion holds in two versions).*

*Proof.* Notice that the right-hand side is independent of $L$ and so the upper index $L$ in the sum on the left-hand side could even be replaced by $\infty$. Given any $L \in \mathbb{N}_0$, however, the left hand-side is a finite sum unlike the right-hand side that could possibly be equal to infinity. In the latter case the assertion is trivial and so, without loss of generality, given $L \in \mathbb{N}_0$ suppose that only those $s^2(j, T)$ that arise on the left-hand side are possibly different from zero (for this makes the right-hand side as small as possible). Then only a finite number of summands are non-zero on both sides.

The proof departs with the definition of $M_2$ from Lemma 2 and $T \in \mathcal{R}_{1,\ell}(K)$ for some $K \in \mathcal{T}_\ell$ of level $\ell(T)$ and concludes that

$$(\ell(T) - M_2)_+ \le \min \ell(\mathcal{R}_{1,\ell}(K)) \le \ell(T).$$

Consequently,

$$s^2(\min \ell(\mathcal{R}_{1,\ell}(K)), T) \le \sum_{j=(\ell(T)-M_2)_+}^{\ell(T)} s^2(j, T).$$

Any $T \in \mathcal{R}_{1,\ell}(K)$ for some $K \in \mathcal{T}_\ell^\pm$ belongs to the set

$$\mathcal{T}_\ell' := \{T \in \mathcal{T}_\ell : \mathcal{R}_{2,\ell}(T) \ne \mathcal{R}_{2,\ell\pm1}(T)\}$$

and Lemma 2 asserts that there exist at most $M_1$ of those neighboring triangles $K$ in $\mathcal{R}_{1,\ell}(T)$. This leads for any $\ell \in \mathbb{N}_0$ to

$$\sum_{K \in \mathcal{T}_\ell^\pm} \sum_{T \in \mathcal{R}_{1,\ell}(K)} s^2(\min \ell(\mathcal{R}_{1,\ell}(K)), T) \le M_1 \sum_{T \in \mathcal{T}'_\ell} \sum_{j=(\ell(T)-M_2)_+}^{\ell(T)} s^2(j, T)$$
$$= M_1 \sum_{m=0}^{\infty} \sum_{T \in \mathcal{T}'_\ell \cap \mathcal{T}_m^{\text{uniform}}} \sum_{j=(\ell(T)-M_2)_+}^{\ell(T)} s^2(j, T).$$

The last identity utilizes that each admissible triangle $T$ belongs to $\mathcal{T}_m^{\text{uniform}}$ for the level $m = \ell(T) \in \mathbb{N}_0$. A rearrangement of the last inequality shows

$$\sum_{\ell=0}^{L} \sum_{K \in \mathcal{T}_\ell^\pm} \sum_{T \in \mathcal{R}_{1,\ell}(K)} s^2(\min \ell(\mathcal{R}_{1,\ell}(K)), T) \le M_1 \sum_{\ell=0}^{L} \sum_{m=0}^{\infty} \sum_{T \in \mathcal{T}'_\ell \cap \mathcal{T}_m^{\text{uniform}}} \sum_{j=(\ell(T)-M_2)_+}^{\ell(T)} s^2(j, T)$$
$$= M_1 \sum_{m=0}^{\infty} \sum_{j=(m-M_2)_+}^{m} \sum_{T \in \mathcal{T}_m^{\text{uniform}}} \mu(T) s^2(j, T)$$

for the multiplicity $\mu(T)$ defined as the cardinality $\mu(T) = |\mathcal{L}(T)| \le L + 1$ of the set

$$\mathcal{L}(T) := \{\ell_1, \dots, \ell_{\mu(T)}\} := \{\ell \in \{0, 1, \dots, L\} : T \in \mathcal{T}'_\ell\}$$

of indices $\ell_1 < \ell_2 < \dots < \ell_{\mu(T)}$. There is nothing to prove in the sequel for $\mathcal{L}(T) = \emptyset$ and $\mu(T) = 0$, so suppose $\mu(T) \in \{1, \dots, L\}$ for the moment. The definition of $\mathcal{T}'_\ell$ implies $\mathcal{R}_{2,\ell_1}(T) \ne \mathcal{R}_{2,\ell_2}(T) \ne \dots \ne \mathcal{R}_{2,\ell_{\mu(T)}}(T)$. Since $\mathcal{T}_{\ell_1}, \mathcal{T}_{\ell_2}, \dots, T_{\ell_{\mu(T)}}$ are successive refinements of each other, the sets $\mathcal{R}_{2,\ell_1}(T), \mathcal{R}_{2,\ell_2}(T), \dots, \mathcal{R}_{2,\ell_{\mu(T)}}(T)$ are pairwise distinct. In other words, the map $\ell \mapsto \mathcal{R}_{2,\ell}(T)$ is an injection of $\mathcal{L}(T)$ into $\{\mathcal{R}_2(\mathcal{T}, T) :$ there exists $\mathcal{T} \in \mathbb{T},\ T \in \mathcal{T}\}$. The cardinality of the latter set is at most $M_3$ from Lemma 2 and so $\mu(T) \le M_3$. The substitution of this in the previous upper bound proves

$$\sum_{\ell=0}^{L} \sum_{K \in \mathcal{T}_\ell^\pm} \sum_{T \in \mathcal{R}_{1,\ell}(K)} s^2(\min \ell(\mathcal{R}_{1,\ell}(K)), T) \le M_1 M_3 \sum_{m=0}^{\infty} \sum_{j=(m-M_2)_+}^{m} \sum_{T \in \mathcal{T}_m^{\text{uniform}}} s^2(j, T).$$

Recall the sum convention $s^2(j, \mathcal{T}_m^{\text{uniform}}) = \sum_{T \in \mathcal{T}_m^{\text{uniform}}} s^2(j, T)$ to conclude the proof. □

## 6.3 Analysis for Uniform Triangulations

All triangulations and the entire analysis throughout this subsection concern the uniform meshes from Section 6.2.2 and so the upper index is skipped here (but added once the reference is made outside of this subsection), so $\mathcal{T}_\ell \equiv \mathcal{T}_\ell^{\text{uniform}}$ and the mesh-size $h_\ell$ is $2^{-\frac{\ell}{2}}$ as $h_0$ and $H_0$ are viewed as global constants and hidden in the notation $1 \approx h_0 \approx H_0$ for simplicity. Recall that the shape regularity in $\mathcal{T}_\ell$ is inherited from $\mathcal{T}_0$ and so the diameter of $T$ is $\approx 2^{-\frac{\ell}{2}}$.

The strengthened Cauchy–Schwarz inequality reads as follows.

**Lemma 4.** *Any $\ell, m \in \mathbb{N}_0$, $v_\ell \in A(\mathcal{T}_\ell)$, and $v_m \in A(\mathcal{T}_m)$ satisfy*

$$2^{\frac{|\ell-m|}{4}} |a(v_\ell, v_m)| \lesssim h_\ell^{-2} \|v_\ell\|_{L^2(\Omega)} h_m^{-2} \|v_m\|_{L^2(\Omega)}.$$

*Proof.* The proof may follow the lines in [30, Lemma 5.1]; a simplified version is given below based on the trace inequality and inverse estimates. Their combination is sometimes called discrete trace inequality and asserts $\|f\|_{L^2(\partial T)} \lesssim |T|^{-\frac{1}{4}} \|f\|_{L^2(T)}$ for each polynomial $f \in P_5(T)$ in the triangle $T \in \mathcal{T} \in \mathbb{T}$ (see [17]). Without loss of generality let $\ell \le m$ and consider $K \in \mathcal{T}_\ell$ so that $v_\ell|_K$ is a quintic polynomial and $v_m|_K \in H^2(K)$ is a piecewise quintic polynomial. An integration by parts shows

$$\int_K \Delta v_\ell \Delta v_m \, dx = \int_{\partial K} (\Delta v_\ell)(v_K \cdot \nabla v_m) \, dx - \int_K \nabla v_m \cdot \nabla \Delta v_\ell \, dx$$
$$\lesssim \|\Delta v_\ell\|_{L^2(\partial K)} \|\nabla v_m\|_{L^2(\partial K)} + \|\nabla v_m\|_{L^2(K)} \|\nabla \Delta v_\ell\|_{L^2(K)}.$$

The discrete trace inequality leads to $\|\Delta v_\ell\|_{L^2(\partial K)} \lesssim h_\ell^{-\frac{1}{2}}\|\Delta v_\ell\|_{L^2(K)}$ and to $\|\nabla v_m\|_{L^2(\partial K\cap\partial T)} \lesssim h_m^{-\frac{1}{2}}\|\nabla v_m\|_{L^2(T)}$ for any $T \in \mathcal{T}_m$. Let $\mathcal{T}_k'(\partial K)$ denote the set of all triangles $T \in \mathcal{T}_m$ with $T \subset K$ and an intersection $\partial K \cap \partial T$ with the boundary of $K$ of positive length. The discrete trace estimate shows

$$\begin{aligned}\|\nabla v_m\|^2_{L^2(\partial K)} &= \sum_{T\in\mathcal{T}_k'(\partial K)} \|\nabla v_m\|^2_{L^2(\partial K\cap\partial T)} \leq \sum_{T\in\mathcal{T}_k'(\partial K)} \|\nabla v_m\|^2_{L^2(\partial T)}\\ &\lesssim h_m^{-1} \sum_{T\in\mathcal{T}_k'(\partial K)} \|\nabla v_m\|^2_{L^2(T)} \leq h_m^{-1}\|\nabla v_m\|^2_{L^2(K)}.\end{aligned}$$

The square root of this, the above estimates, and inverse estimates verify

$$\left|\int_K \Delta v_\ell \Delta v_m\,dx\right| \lesssim (h_\ell^{-\frac{1}{2}}h_m^{\frac{1}{2}} + h_\ell^{-1}h_m)h_\ell^{-2}\|v_\ell\|_{L^2(K)}h_m^{-2}\|v_\ell\|_{L^2(K)}.$$

The sum of all those estimates over $K \in \mathcal{T}_\ell$ and $h_\ell^{-\frac{1}{2}}h_m^{\frac{1}{2}} + h_\ell^{-1}h_m \lesssim 2^{\frac{\ell-m}{4}}$ conclude the proof. □

Let $P_\ell \equiv P_\ell^{\text{uniform}}$ and $Q_\ell \equiv Q_\ell^{\text{uniform}}$ denote the Galerkin and the $L^2$ projection onto $A(\mathcal{T}_\ell)$ are linear and bounded operators from $H_0^2(\Omega)$ into $H_0^2(\Omega)$ defined for any $v \in H_0^2(\Omega)$ by $P_\ell v, Q_\ell v \in A(\mathcal{T}_\ell)$ with

$$\begin{aligned} a(P_\ell v, w_A) &= a(v, w_A) && \text{for all } w_A \in A(\mathcal{T}_\ell),\\ (Q_\ell v, w_A)_{L^2(\Omega)} &= (v, w_A)_{L^2(\Omega)} && \text{for all } w_A \in A(\mathcal{T}_\ell).\end{aligned}$$

The quasi-uniform triangulations allow for approximation error estimates and for inverse estimates for functions and norms in the Sobolev spaces $H^s(\Omega)$ with semi-norm $|\bullet|_{H^s(\Omega)}$ for $0 \leq s \leq 2$ defined, for instance, by interpolation of $L^2(\Omega) \equiv H^0(\Omega)$ and $H^1(\Omega)$ resp. $H^2(\Omega)$.

The following approximation and stability properties are well known for the quasi-uniform triangulations at hand in this subsection.

**Lemma 5** ([30, Lemma 5.2]). *For any* $0 \leq r \leq 2$ *and* $\ell \in \mathbb{N}_0$,

$$\|v - Q_\ell v\|_{H^r(\Omega)} \lesssim 2^{(\frac{r}{2}-1)\ell}|v|_{H^2(\Omega)} \quad \text{and} \quad \|Q_\ell w\|_{H^r(\Omega)} \lesssim \|w\|_{H^r(\Omega)}$$

*hold for* $v \in H^2(\Omega)$ *and* $w \in H^r(\Omega)$. □

The analysis requires the known reduced elliptic regularity on polygons that involves some index $\frac{1}{2} < \sigma \leq 1$ throughout the remaining parts of this paper.

**Theorem 8** (Regularity Shift). *Given the bounded Lipschitz domain* $\Omega \subset \mathbb{R}^2$ *with polygonal boundary* $\partial\Omega$, *there exist some constants* $\sigma > \frac{1}{2}$ *and* $C > 0$ *such that any* $v \in H_0^2(\Omega)$ *with* $\Delta^2 v \in H^{\sigma-2}(\Omega)$ *satisfies* $v \in H_0^2(\Omega) \cap H^{2+\sigma}(\Omega)$ *and* $\|u\|_{H^{2+\sigma}(\Omega)} \leq C\|\Delta^2 u\|_{H^{\sigma-2}}$.

*Proof.* The result has been established by Blum and Rannacher for convex domains with $\sigma = 1$ and for the maximal interior angle $\omega$ smaller than $1.4\pi$ with $\sigma = 2$. The non-convex case leads to $\sigma < 1$ depending on the maximum angle and $\sigma > \frac{1}{2}$ follows for $\omega < 2\pi$. The result is frequently accepted and more details may be found in [1, 22]. It turns out that $v \in H_0^2(\Omega)$ with $\Delta^2 v \in H^{s-2}(\Omega)$ satisfies $v \in H_0^2(\Omega) \cap H^{2+s}(\Omega)$ and $\|u\|_{H^{2+\sigma}(\Omega)} \leq C\|\Delta^2 u\|_{H^{s-2}}$ for all $s$ with $0 \leq s \leq \sigma$ and this explains the description as a shift theorem. □

The reduced elliptic regularity and the Aubin–Nitsche duality technique provide error estimates in weaker Sobolev norms; the proof is outlined for completeness.

**Lemma 6** (Duality). *Any* $v \in H_0^2(\Omega)$ *and* $\ell \in \mathbb{N}_0$ *satisfy*

$$|v - P_\ell v|_{H^{2-\sigma}(\Omega)} \lesssim 2^{-\frac{\sigma\ell}{2}}|v - P_\ell v|_{H^2(\Omega)}.$$

*Proof.* Given $v - P_\ell v \in H_0^2(\Omega) \subset H^{2-\sigma}(\Omega)$, there exists a functional $F$ in the dual $H^{\sigma-2}(\Omega)$ with norm one such that $\|v - P_\ell v\|_{H^{2-\sigma}(\Omega)} = F(v - P_\ell v)$ from a corollary of the Hahn–Banach extension theorem. The weak solution $z \in H_0^2(\Omega)$ to $\Delta^2 z = F$ belongs to $H^{2+\sigma}(\Omega)$ by Theorem 8. Then $F(v - P_\ell v) = a(v - P_\ell v, z) = a(v - P_\ell v, z - z_A)$

for some approximation $z_A$ to $z$ and standard arguments on the (standard) Argyris finite element interpolation

$$\|v - P_\ell v\|_{H^{2-\sigma}(\Omega)} \le |v - P_\ell v|_{H^2(\Omega)} |z - z_A|_{H^2(\Omega)} \lesssim h_\ell^\sigma |v - P_\ell v|_{H^2(\Omega)} |z|_{H^{2+\sigma}(\Omega)}$$

conclude the proof with $\|z\|_{H^{2+\sigma}(\Omega)} \lesssim 1$ and $h_\ell \approx 2^{-\frac{\ell}{2}}$. □

**Lemma 7** (Stability). *Any $v \in H_0^2(\Omega)$ and $\ell, m \in \mathbb{N}_0$ with $v_\ell := (P_\ell - P_{\ell-1})v$ satisfy $(Q_m - Q_{m-1})v_\ell = 0$ for $\ell \le m-1$ and for $m \le \ell$ that*

$$|(Q_m - Q_{m-1})v_\ell|_{H^2(\Omega)} \lesssim 2^{-\frac{\sigma(\ell-m)}{2}} |v_\ell|_{H^2(\Omega)}.$$

*Proof.* The proof follows that in [30, Lemma 5.3] and is merely outlined with less typos. On uniform meshes, the interpolation theory provides an inverse estimate

$$|(Q_m - Q_{m-1})v_\ell|_{H^2(\Omega)} \lesssim 2^{\frac{\sigma m}{2}} |(Q_m - Q_{m-1})v_\ell|_{H^{2-\sigma}(\Omega)}$$

in fractional-order Sobolev spaces. The stability in Lemma 5 (b) leads to $|(Q_m - Q_{m-1})v_\ell|_{H^{2-\sigma}(\Omega)} \lesssim |v_\ell|_{H^{2-\sigma}(\Omega)}$. Lemma 6 applies to $v_\ell = (1 - P_\ell)P_{\ell-1}v$ and shows $|v_\ell|_{H^{2-\sigma}(\Omega)} \lesssim 2^{-\sigma\frac{\ell}{2}} |v_\ell|_{H^2(\Omega)}$. The combination of the aforementioned estimates concludes the proof. □

**Lemma 8.** *Any $v \in H_0^2(\Omega)$ and $0 \le r < 2$ satisfy*

$$\sum_{\ell=1}^{\infty} |(Q_\ell - Q_{\ell-1})v|_{H^2(\Omega)}^2 \lesssim |v|_{H^2(\Omega)}^2, \tag{a}$$

$$\sum_{\ell=0}^{\infty} 2^{(2-r)\ell} \|v - Q_\ell v\|_{H^r(\Omega)}^2 \lesssim |v|_{H^2(\Omega)}^2. \tag{b}$$

**Remark 6.1.** Lemma 8 (b) fails for $r = 2$. A counterexample concerns $L \in \mathbb{N}$ and some $v_L \in A(\mathcal{T}_L)$ of $H^2$ norm one, which is $L^2$ orthogonal to $A(\mathcal{T}_{L-1})$. Then the left-hand side of Lemma 8 (b) is $L$ and the right-hand side is one.

*Proof of Lemma 8* (a). The proof follows that of [30, Lemma 5.4]. Abbreviate $v_\ell := (P_\ell - P_{\ell-1})v$ for all $\ell \in \mathbb{N}_0$ with $P_{-1}v := 0$ and observe that the series $v = \sum_{\ell=0}^\infty v_\ell$ converges in $H_0^2(\Omega)$ by the convergence of the Galerkin scheme and $|v|_{H^2(\Omega)}^2 = \sum_{\ell=0}^\infty |v_\ell|_{H^2(\Omega)}^2$. Lemma 7 and triangle inequalities show

$$\begin{aligned}\sum_{m=1}^{\infty} |(Q_m - Q_{m-1})v|_{H^2(\Omega)}^2 &= \sum_{m=1}^{\infty} \Big| \sum_{\ell=m}^{\infty} (Q_m - Q_{m-1})v_\ell \Big|_{H^2(\Omega)}^2 \\ &\le \sum_{m=1}^{\infty} \Big( \sum_{\ell=m}^{\infty} |(Q_m - Q_{m-1})v_\ell|_{H^2(\Omega)} \Big)^2 \lesssim \sum_{m=1}^{\infty} \Big( \sum_{\ell=m}^{\infty} 2^{-\frac{\sigma(\ell-m)}{2}} |v_\ell|_{H^2(\Omega)} \Big)^2.\end{aligned}$$

With the abbreviations $a_\ell := |v_\ell|_{H^2(\Omega)}$ and $0 < q := 2^{-\frac{\sigma}{2}} < 1$, the last sum reads

$$\begin{aligned}\sum_{m=1}^{\infty} \Big( \sum_{\ell=m}^{\infty} a_\ell q^{\ell-m} \Big)^2 &= \sum_{m=1}^{\infty} \sum_{k,\ell=m}^{\infty} a_k a_\ell q^{k+\ell-2m} = \sum_{k,\ell=1}^{\infty} a_k a_\ell \Big( \sum_{m=1}^{\min\{k,\ell\}} q^{k+\ell-2m} \Big) \\ &\le \frac{1}{1-q^2} \sum_{k,\ell=1}^{\infty} a_k a_\ell q^{|k-\ell|} = \frac{1}{1-q^2} a \cdot Qa\end{aligned}$$

for the vector $a = (a_j : j \in \mathbb{N}_0)$ and the infinite matrix $Q$ with the entry $Q_{jk} = q^{|j-k|}$ for $j, k \in \mathbb{N}_0$. The Gershgorin circle theorem shows (for each finite submatrix of $Q$ and hence for $Q$) that all eigenvalues of the symmetric matrix $Q$ belong to the compact interval $[1 - \frac{q}{1-q}, 1 + \frac{q}{1-q}]$. Consequently, $a \cdot Qa \le \frac{1}{1-q} a \cdot a$. This and $a \cdot a = |v|_{H^2(\Omega)}^2$ conclude the proof. □

*Proof of Lemma 8* (b). Set $w_m := (Q_m - Q_{m-1})v$ and apply Lemma 5 (for $v$ and $\ell$ replaced by $w_m$ and $m$ with $w_m = w_m - Q_{m-1}w_m$) to verify

$$\|w_m\|_{H^r(\Omega)} \lesssim \gamma^m |w_m|_{H^2(\Omega)} \quad \text{for } 0 < \gamma := 2^{\frac{r}{2}-1} < 1.$$

After triangle inequalities this leads for $b_m := |w_m|_{H^2(\Omega)}$ to

$$\|v - Q_\ell v\|_{H^r(\Omega)} \le \sum_{m=\ell+1}^{\infty} \|w_m\|_{H^r(\Omega)} \lesssim \sum_{m=\ell+1}^{\infty} \gamma^m b_m.$$

The remaining calculations follow the arguments of part (a) of the proof and utilize the last estimate to establish

$$\sum_{\ell=0}^{\infty} 2^{(2-r)\ell} \|v - Q_\ell v\|_{H^r(\Omega)}^2 \lesssim \sum_{\ell=0}^{\infty} \sum_{k,m=\ell+1}^{\infty} \gamma^{m+k-2\ell} b_k b_m$$
$$= \sum_{k,m=1}^{\infty} b_k b_m \Bigg( \sum_{\ell=0}^{\min\{k,m\}-1} \gamma^{m+k-2\ell} \Bigg) \leq \frac{1}{1-\gamma^2} \sum_{k,m=1}^{\infty} \gamma^{|m-k|} b_k b_m.$$

The last sum is the product of the vector $(b_k : k \in \mathbb{N})$ times an infinite matrix $Q := (\gamma^{|m-k|} : k, m \in \mathbb{N})$ times the vector $(b_m : m \in \mathbb{N})$. The eigenvalues of $Q$ are bounded by $\frac{1}{1-\gamma}$ and hence

$$(1-\gamma)^2(1+\gamma) \sum_{\ell=0}^{\infty} 2^{(2-r)\ell} \|v - Q_\ell v\|_{H^r(\Omega)}^2 \leq \sum_{j=1}^{\infty} b_j^2 = \sum_{j=1}^{\infty} |(Q_j - Q_{j-1})v|_{H^2(\Omega)}^2.$$

This and Lemma 8 (a) conclude the proof. □

# 7 Convergence Analysis

## 7.1 Overview

The proof of Theorem 7 concludes in Section 7.6 with arguments of a general multilevel theory [5, 29] based on certain decompositions. Throughout the analysis of the multigrid algorithm of Section 5, there is a given sequence $(\mathcal{T}_\ell)_{\ell\in\mathbb{N}_0}$ of successive one-level mesh-refinements in $\mathbb{T}$ and $\ell \in \mathbb{N}_0$. Section 7.2 adapts the quasi-interpolation of Theorem 2 to the reduced space $\widetilde{A}(\mathcal{T}_\ell) = \operatorname{span}\{\phi_\ell^1, \dots, \phi_\ell^{\tilde{n}_\ell}\}$ from Section 5.2 and defines a sequence $(J_\ell)_{\ell\in\mathbb{N}_0}$ of quasi-interpolation operators. Sections 7.3–7.5 provide three estimates of partial sums in the split

$$v_L = J_0 v_L + \sum_{\ell=0}^{L-1} (J_{\ell+1} - J_\ell) v_L \tag{7.1}$$

of an arbitrary $v_L \in A(\mathcal{T}_L)$ and $L \in \mathbb{N}$. The key point in Propositions 1–3 below is that the generic constants do *neither* depend on $L \in \mathbb{N}$ *nor* on $(\mathcal{T}_\ell)_{\ell\in\mathbb{N}_0}$, but solely on $\mathcal{T}_0$ through the constants $M_1, M_2, M_3$ of Lemma 2 and $C_d$ from Theorem 9 (d) (resp. Theorem 2 (d)).

## 7.2 Quasi-Interpolation

The design of the quasi-interpolation operator in Theorem 2 *looks ahead* towards a refinement $\widehat{\mathcal{T}}$ of the current triangulation $\mathcal{T}$ under conditions (C1)–(C2). The multigrid convergence analysis relies on a sequence $(J_\ell)_{\ell\in\mathbb{N}_0}$ of quasi-interpolation operators, which *look backwards* in the sequence $(\mathcal{T}_\ell)_{\ell\in\mathbb{N}_0}$ of the successive mesh-refinements.

**Theorem 9.** *Given a sequence* $(\mathcal{T}_\ell)_{\ell\in\mathbb{N}_0}$ *of successive one-level mesh-refinements in* $\mathbb{T}$*, there exists a sequence* $(J_\ell)_{\ell\in\mathbb{N}_0}$ *of quasi-interpolation operators that satisfy properties* (b)–(d) *in Theorem 2 for all* $\ell \in \mathbb{N}_0$ *(when* $\mathcal{T}_\ell$ *and* $J_\ell$ *substitute* $\mathcal{T}$ *and* $J$*) and property* (a') *for all* $\ell \in \mathbb{N}_0$*:*
(a') $(J_{\ell+1} - J_\ell)v \in \widetilde{A}(\mathcal{T}_{\ell+1})$ *for all* $v \in H_0^2(\Omega)$.

The definition of $J$ is steered by a family $(T_{z,j} : (z,j) \in \mathcal{J})$ of triangles with (C1). Adapt all the notation $\mathcal{T}$, $\mathcal{N}$, $m(z)$ for $z \in \mathcal{N}$, $\psi_{z,j}$ for $(z,j) \in \mathcal{J}$ for the level $\ell \in \mathbb{N}$ and write $\mathcal{T}_\ell$, $\mathcal{N}_\ell$, $m(z,\ell)$ for $z \in \mathcal{N}_\ell$, $\psi_{z,j}^\ell$ for $(z,j) \in \mathcal{J}_\ell$, etc. In particular, $\mathcal{T}_\ell(z) := \{T \in \mathcal{T}_\ell : z \in T\}$ for all $z \in \mathcal{N}_\ell$ and $\mathcal{T}_{\ell,\pm}(z) = \mathcal{T}_\ell(z) \cap H_\pm(z)$ for all $z \in \mathcal{V}_\ell(\Omega) \setminus \mathcal{V}_0(\Omega)$.

Condition (C3) is nothing else than a rewriting of (C1) for each level.

**Condition (C3).** For any $\ell \in \mathbb{N}_0$, the family $(T_{z,j}^\ell : (z,j) \in \mathcal{J}_\ell)$ of triangles satisfies $T_{z,1}^\ell, \dots, T_{z,m(z,\ell)}^\ell \in \mathcal{T}_\ell(z)$ for all $z \in \mathcal{N}_\ell$ and, in addition, $T_{z,6}^\ell \in \mathcal{T}_{\ell,+}(z)$ and $T_{z,7}^\ell \in \mathcal{T}_{\ell,-}(z)$ at each $z \in \mathcal{V}_\ell(\Omega) \setminus \mathcal{V}_0(\Omega)$.

Note that the set of nodes increases with the level $\mathcal{N}_\ell \subset \mathcal{N}_{\ell+1}$ for any $\ell \in \mathbb{N}_0$ and

$$\widetilde{\mathcal{N}}_{\ell+1} := \{z \in \mathcal{N}_\ell : \mathcal{T}_\ell(z) \neq \mathcal{T}_{\ell+1}(z)\} \cup (\mathcal{N}_{\ell+1} \setminus \mathcal{N}_\ell) \subset \mathcal{N}_{\ell+1}$$

defines the space $\widetilde{A}(\mathcal{T}_{\ell+1}) = \operatorname{span}\{\psi_{z,j}^{\ell+1} : j = 1, \dots, m(z, \ell+1) \text{ for } z \in \widetilde{\mathcal{N}}_{\ell+1}\}$.

**Condition (C4).** For any $\ell \in \mathbb{N}_0$, the family $(T_{z,j}^{\ell+1} : (z, j) \in \mathcal{J}_{\ell+1})$ of triangles satisfies

$$T_{z,j}^{\ell+1} = T_{z,j}^{\ell} \quad \text{for all } z \in \mathcal{N}_{\ell+1} \setminus \widetilde{\mathcal{N}}_{\ell+1} \text{ and all } j = 1, \dots, m(z, \ell+1).$$

*Definition of* $(J_\ell)_{\ell\in\mathbb{N}_0}$ *in Theorem 9.* Mathematical induction proves that a family $(T_{z,j}^{\ell} : (z, j) \in \mathcal{J}_\ell, \ell \in \mathbb{N}_0)$ of triangles with (C3)–(C4) can be designed. (Note that $z \in \mathcal{N}_{\ell+1} \setminus \widetilde{\mathcal{N}}_{\ell+1}$ implies $z \in \mathcal{N}_\ell$ and $\mathcal{T}_\ell(z) = \mathcal{T}_{\ell+1}(z)$ so that $m(z, \ell+1) = m(z, \ell)$ and $T_{z,j}^{\ell} \in T_\ell(z)$ from (C3) enables the selection $T_{z,j}^{\ell+1} = T_{z,j}^{\ell} \in T_{\ell+1}(z)$ for (C3).) Condition (C3) means that (C1) holds on any level and this is sufficient for properties (b)–(d) asserted in Theorem 2 for all $\ell \in \mathbb{N}_0$ when $\mathcal{T}_\ell$ and $J_\ell$ substitute $\mathcal{T}$ and $J$. □

*Proof of Theorem 9* (a'). Consider $z \in \mathcal{N}_{\ell+1} \setminus \widetilde{\mathcal{N}}_{\ell+1}$ and $j = 1, \dots, m(z, \ell+1)$ and revisit of the design of the functionals and weight functions in (3.2) of the proof of Theorem 2. Since $\mathcal{T}_\ell(z) = \mathcal{T}_{\ell+1}(z)$, $m(z, \ell+1) = m(z, \ell)$, and $T_{z,j}^{\ell} = T_{z,j}^{\ell+1}$ from (C4), it holds $\psi_{z,j}^{\ell} = \psi_{z,j}^{\ell+1}$, $\Lambda_{z,j}^{\ell} = \Lambda_{z,j}^{\ell+1}$, and the weight functions, and the resulting respective functionals in (3.1)–(3.2) are equal. For instance $M_k^{\ell} = M_m^{\ell+1}$ provided $k$ is the number of the degree of freedom $\Lambda_{z,j}^{\ell} = \Lambda_k^{\ell}$ on the level $\ell$ and $m$ is the number of the same degree of freedom $\Lambda_{z,j}^{\ell} = \Lambda_{z,j}^{\ell+1} = \Lambda_m^{\ell+1}$ on the level $\ell+1$. The coefficient $M_k^{\ell} v$ of $\psi_{z,j}^{\ell} = \psi_{z,j}^{\ell+1}$ in the definition of $J_\ell v$ is the same coefficient $M_m^{\ell+1} v$ of $\psi_{z,j}^{\ell} = \psi_{z,j}^{\ell+1}$ in the definition of $J_{\ell+1} v$. This and the duality relation (from Notation 2.4) imply on the level $\ell+1$ that

$$\Lambda_{z,j}^{\ell+1}(J_{\ell+1} v - J_\ell v) = 0 \quad \text{for all } z \in \mathcal{N}_{\ell+1} \setminus \widetilde{\mathcal{N}}_{\ell+1} \text{ and } j = 1, \dots, m(z, \ell+1).$$

Consequently (with the duality in the first equality),

$$\begin{aligned} J_{\ell+1} v - J_\ell v &= \sum_{z\in\mathcal{N}_{\ell+1}} \sum_{j=1}^{m(z,\ell+1)} (\Lambda_{z,j}^{\ell+1}(J_{\ell+1} v - J_\ell v))\psi_{z,j}^{\ell+1} \\ &= \sum_{z\in\widetilde{\mathcal{N}}_{\ell+1}} \sum_{j=1}^{m(z,\ell+1)} (\Lambda_{z,j}^{\ell+1}(J_{\ell+1} v - J_\ell v))\psi_{z,j}^{\ell+1} \in \widetilde{A}(\mathcal{T}_{\ell+1}). \end{aligned}$$

The inclusion in the last step follows from the above characterization of $\widetilde{A}(\mathcal{T}_{\ell+1})$ by $\widetilde{\mathcal{N}}_{\ell+1}$ and concludes the proof of (a'). □

## 7.3 Stability of the Decomposition

The subsequent proposition is of general interest where $h_{\ell+1} \in P_0(\mathcal{T}_{\ell+1})$ is the piecewise constant mesh-size in the triangulation $\mathcal{T}_{\ell+1}$.

**Proposition 1** (Stability). *The decomposition* (7.1) *of* $v_L \in A(\mathcal{T}_L)$ *satisfies*

$$\sum_{\ell=0}^{L-1} (\|h_{\ell+1}^{-2}(J_{\ell+1} - J_\ell)v_L\|_{L^2(\Omega)}^2 + |(J_{\ell+1} - J_\ell)v_L|_{H^2(\Omega)}^2) \lesssim |v_L|_{H^2(\Omega)}^2.$$

*Proof.* The proof utilizes the $L^2$ projection $Q_m \equiv Q_m^{\text{uniform}}$ with respect to the uniform triangulations $\mathcal{T}_m^{\text{uniform}}$ of Section 6.3. Abbreviate $v_{\ell+1} := (J_{\ell+1} - J_\ell)v_L$ for $\ell = 0, \dots, L-1$. Recall the neighborhoods $\mathcal{R}_{j,\ell}(T) := \mathcal{R}_j(\mathcal{T}_\ell, T)$ for neighboring triangles of $T \in \mathcal{T}_\ell$ with respect to $\mathcal{T}_\ell$ from Section 6.2.1.

*Step one* provides for $\ell = 0, 1, \dots, L-1$ and $m(T) := \min \ell(\mathcal{R}_{1,\ell}(T))$ for $T \in \mathcal{T}_\ell$ with $\Omega(T) = \operatorname{int}(\bigcup \mathcal{R}_{1,\ell}(T))$ the estimate

$$\|v_{\ell+1}\|_{L^2(T)} \lesssim \begin{cases} 0 & \text{if } \mathcal{R}_{1,\ell}(T) = \mathcal{R}_{1,\ell+1}(T), \\ h_T\, |(1 - Q_{m(T)})v_L|_{H^1(\Omega(T))} & \text{else.} \end{cases} \tag{7.2}$$

The proof of (7.2) departs from the definition of $J_\ell$ and Theorem 2 and verifies $J_{\ell+1} v_L = J_\ell v_L$ a.e. in $T$ in case $\mathcal{R}_{1,\ell}(T) = \mathcal{R}_{1,\ell+1}(T)$; this proves the first alternative in (7.2).

Hence suppose $\mathcal{R}_{1,\ell}(T) \neq \mathcal{R}_{1,\ell+1}(T)$. The discrete function $v_{\ell+1}|_T = ((J_{\ell+1} - J_\ell)v_L)|_T$ is defined by the data $v_L|_{H^1(\Omega(T))}$ with stability as in Theorem 2 (d). The definition of $m(T) := \min \ell(\mathcal{R}_{1,\ell}(T))$ as a minimal level near $T$ in $\mathcal{T}_\ell$ guarantees that $w_{m(T)} := (Q_{m(T)}v_L)|_{\Omega(T)}$ is a quintic polynomial in each triangle in $\mathcal{R}_{1,\ell}(T)$ and in each triangle in $\mathcal{R}_{1,\ell+1}(\widehat{T})$ for $\widehat{T} \in \mathcal{T}_{\ell+1}$ with $\widehat{T} \subset T$. This and the definition of $J_\ell$ resp. $J_{\ell+1}$ show that $J_\ell w_{m(T)} = w_{m(T)} = J_{\ell+1} w_{m(T)}$ a.e. in $T$. Hence

$$\begin{aligned}\|v_{\ell+1}\|_{L^2(T)} &= \|(J_{\ell+1} - J_\ell)(v_L - Q_{m(T)}v_L)\|_{L^2(T)} \\ &\le \|(1 - J_{\ell+1})(v_L - Q_{m(T)}v_L)\|_{L^2(T)} + \|(1 - J_\ell)(v_L - Q_{m(T)}v_L)\|_{L^2(T)}.\end{aligned}$$

For $j = 0$, Theorem 2 (d) shows

$$\|(1 - J_{\ell+j})(v_L - Q_{m(T)}v_L)\|_{L^2(T)} \lesssim h_T\, |(1 - Q_{m(T)})v_L|_{H^1(\Omega(T))}. \tag{7.3}$$

Theorem 2 (d) also applies to each $\widehat{T} \in \mathcal{T}_{\ell+1}$ with $\widehat{T} \subset T$ as well and the sum of all those contributions proves (7.3) for $j = 1$. The combination with the second last displayed estimate concludes the proof of (7.2).

*Step two* recalls the abbreviations $\mathcal{T}_\ell^+$ from (6.6) and $m(K)$ from Step one and provides, for each $\ell \in \mathbb{N}_0$, that

$$\|h_{\ell+1}^{-2} v_{\ell+1}\|_{L^2(\Omega)}^2 + |v_{\ell+1}|_{H^2(\Omega)}^2 \lesssim \sum_{K\in\mathcal{T}_\ell^+} \sum_{T\in\mathcal{R}_{1,\ell}(K)} |T|^{-1} |v_L - Q_{m(K)}v_L|_{H^1(T)}^2. \tag{7.4}$$

The proof of (7.4) departs with piecewise (with respect to $\mathcal{T}_{\ell+1}$) inverse estimates and the piecewise constant mesh-sizes for the one-level refinement $\mathcal{T}_{\ell+1}$ of $\mathcal{T}_\ell$ with $h_{\ell+1} \le h_\ell \le 2h_{\ell+1}$ to verify

$$|v_{\ell+1}|_{H^2(\Omega)} \lesssim \|h_{\ell+1}^{-2} v_{\ell+1}\|_{L^2(\Omega)}.$$

Recall that any $K \in \mathcal{T}_\ell$ with $\mathcal{R}_{1,\ell}(K) \neq \mathcal{R}_{1,\ell+1}(K)$ belongs to $K \in \mathcal{T}_\ell^+$. This and (7.2) lead to

$$\|h_{\ell+1}^{-2} v_{\ell+1}\|_{L^2(\Omega)} + |v_{\ell+1}|_{H^2(\Omega)}^2 \lesssim \sum_{K\in\mathcal{T}_\ell^+} |K|^{-1}\, |v_L - Q_{m(K)}v_L|_{H^1(\Omega(K))}^2.$$

Since $|K| \approx |T|$ for $T \in \mathcal{R}_{1,\ell}(K)$ (from $M_2 \approx 1$), this implies (7.4).

*Step three* applies the re-summation of Lemma 3 with $s(j, T) := |T|^{-1}|v_L - Q_j v_L|_{H^1(T)}^2$ and so controls the right-hand side of (7.4). This and $|T| \approx 2^{-m}$ for $T \in \mathcal{T}_m^{\text{uniform}}$ result in an estimate

$$\text{LHS} := \sum_{\ell=0}^{L-1} (\|h_{\ell+1}^{-2} v_{\ell+1}\|_{L^2(\Omega)} + |v_{\ell+1}|_{H^2(\Omega)}^2) \lesssim \sum_{m=0}^{\infty} \sum_{j=(m-M_2)_+}^{m} 2^m |v_L - Q_j v_L|_{H^1(\Omega)}^2$$

for the left-hand side LHS in Proposition 1.

*Step four* finishes the proof. For any $j = (m - M_2)_+, \dots, m$, Lemma 2 and $2^m \le (2^{M_2})2^{(m-M_2)_+} \le 2^{M_2} 2^j \le 2^m$ show $2^m \approx 2^j$. Consequently,

$$\text{LHS} \lesssim \sum_{m=0}^{\infty} \sum_{j=(m-M_2)_+}^{m} 2^j |v_L - Q_j v_L|_{H^1(\Omega)}^2 \le (M_2 + 1) \sum_{m=0}^{\infty} 2^m |v_L - Q_m v_L|_{H^1(\Omega)}^2.$$

This and Lemma 8.b for $r = 1$ conclude the proof. □

## 7.4 Second Estimate for the Decomposition

The main argument for the following stability result requires Sobolev–Slobodeckij norms addressed before the proof concludes this subsection.

**Proposition 2** (Key Estimate). *Any $v_L \in A(\mathcal{T}_L)$ satisfies*

$$\sum_{\ell=0}^{L-1} \sum_{k=1}^{\tilde{n}_\ell} \frac{a(v_L - J_\ell v_L, \phi_\ell^k)^2}{|\phi_\ell^k|_{H^2(\Omega)}^2} \lesssim |v_L|_{H^2(\Omega)}^2.$$

The proof below requires Sobolev spaces $H^s(\omega)$ and their norms with a general index $\frac{1}{2} < s < 1$ on triangles. The possible choice $s = \frac{3}{4}$ underlines that the parameter $s$ has nothing to do with the regularity index $\sigma$. The most explicit definition for a subdomain $\omega$ of $\Omega$ in 2D is the Sobolev–Slobodeckij norm $\|\bullet\|_{H^s(\omega)} := (\|\bullet\|^2_{L^2(\omega)} + |\bullet|^2_{H^s(\omega)})^{\frac{1}{2}}$ with the semi-norm defined in two space dimensions for any $f \in H^1(\omega)$ by

$$|f|^2_{H^s(\omega)} := \iint_{\omega\times\omega} \frac{|f(x) - f(y)|^2}{|x - y|^{2+2s}}\,dx\,dy.$$

There are equivalent alternative definitions with equivalence constants that may depend on the domain $\omega$. The decomposition

$$\sum_{T\in\mathcal{T}} |f|^2_{H^s(T)} \le |f|^2_{H^s(\Omega)} \quad \text{for all } f \in H^s(\Omega) \text{ and } \mathcal{T} \in \mathbb{T} \tag{7.5}$$

is obvious for the (non-local) Sobolev–Slobodeckij semi-norm.

**Lemma 9** (Trace and Inverse Inequality). *Any $f \in H^s(T) \equiv H^s(\mathrm{int}(T))$ and any quartic polynomial $p_4 \in P_4(T)$ defined in a (compact) triangle $T$ with interior $\mathrm{int}(T)$ of diameter $h_T$ with minimal interior angle greater than or equal to $\omega_0 > 0$ and $\frac{1}{2} < s \le 1$ satisfy the trace (7.6) and inverse (7.7) inequality,*

$$c(s,\omega_0)h_T^{\frac{1}{2}}\|f\|_{L^2(\partial T)} \le \|f\|_{L^2(T)} + h_T^s|f|_{H^s(T)}, \tag{7.6}$$

$$c(s,\omega_0)h_T^s|p_4|_{H^s(T)} \le \|p_4\|_{L^2(T)}, \tag{7.7}$$

*for a universal positive constant $c(s,\omega_0) \approx 1$ that exclusively depends on $s$ and $\omega_0$.*

*Proof.* The trace and the inverse inequality are well established for a fixed domain such as the reference triangle $T_{\mathrm{ref}} := \mathrm{conv}\{(0,0),(1,0),(0,1)\}$ for various definitions of the semi-norm $|\bullet|_{H^s(T)}$. Any of which leads to the assertion for $T = T_{\mathrm{ref}}$. A standard scaling argument, i.e., the transformation of $T_{\mathrm{ref}}$ onto $T$ by an affine diffeomorphism and the substitution formula for integrals, then proves the lemma for the Sobolev–Slobodeckij semi-norm. The trace lemma has been utilized, e.g., in [7, (2.11)]; further details are found in [18, Lemma 7.2]. □

*Proof of Proposition 2.* Adopt the above notation on $(\mathcal{T}_\ell)_{\ell\in\mathbb{N}_0}$ and abbreviate $w_\ell := v_L - J_\ell v_L$ for $\ell \in \mathbb{N}_0$ (with $0 = w_L = w_{L+1} = \cdots$). Recall $\mathcal{T}_\ell^-$ from (6.7) and that $\phi_\ell^k = \psi_{z,j}^\ell$ has been selected in Section 5.2 with condition (5.6).

*Step one* establishes $\mathcal{T}_\ell(z) \subset \mathcal{T}_\ell^-$ for all $z \in \mathcal{M}_\ell \cup \mathcal{V}_\ell$ with (5.6) and $\ell \in \mathbb{N}_0$. There is nothing to prove for $\ell = 0$ and otherwise, if $z \in \mathcal{M}_\ell \cup \mathcal{V}_\ell$ satisfies (5.6) and $z \in T \in \mathcal{T}_\ell$, $\ell \in \mathbb{N}$, then there exists $K \in \mathcal{T}_\ell(z) \setminus \mathcal{T}_{\ell-1}$ and so $K \in \mathcal{R}_{1,\ell}(T) \setminus \mathcal{R}_{1,\ell-1}(T)$. Hence $T \in \mathcal{T}_\ell^-$.

*Step two comments on the finite overlap* of the supports of the nodal basis functions $\phi_\ell^1, \dots, \phi_\ell^{\tilde{n}_\ell}$. The nodal basis function $\phi_\ell^j$ is (uniquely) associated to a nodal degree of freedom $\Lambda_\ell^j \in A(\mathcal{T}_\ell)^*$ acting at a node $z_\ell^j \in \mathcal{M}_\ell \cup \mathcal{V}_\ell$ for each $j = 1,\dots,\tilde{n}_\ell$. The support $\operatorname{supp}\phi_\ell^j = \bigcup \mathcal{T}_\ell(z_\ell^j)$ of $\phi_\ell^j$ is covered by the triangles $\mathcal{T}_\ell(z_\ell^j)$ that contain $z_\ell^j$. Let $I_\ell \subset \{1,\dots,\tilde{n}_\ell\} \times \{1,\dots,\tilde{n}_\ell\}$ denote the subset of all index pairs $(j,k)$ with overlap $\operatorname{supp}\phi_\ell^j \cap \operatorname{supp}\phi_\ell^k$ of positive area, i.e., if $\mathcal{T}_\ell(z_\ell^j) \cap \mathcal{T}_\ell(z_\ell^k) \ne \emptyset$, for $j,k = 1,\dots,\tilde{n}_\ell$. Given any $j = 1,\dots,\tilde{n}_\ell$, let $I_\ell(j) := \{k = 1,\dots,n_\ell : (j,k) \in I_\ell\}$ denote the set of all $k$ such that the supports of the associated nodal basis functions overlap and let $|I_\ell(j)|$ denote its cardinality. Suppose $M_0$ is a universal upper bound of the number $|\mathcal{T}_\ell(z_\ell^j)| \le M_0$ of triangles in a nodal patch $\mathcal{T}_\ell(z_\ell^j)$ for $j = 1,\dots,\tilde{n}_\ell$ (and so $M_0 \le M_1$). Then at most $9M_0 + 7$ degrees of freedom with a number $k$ (at most $M_0 + 1$ vertices and $2M_0$ edges belong to the patch $\mathcal{T}_\ell(z_\ell^j)$ of an interior node $z_\ell^j$ carrying at most 7 and 1 degree of freedom in $A(\mathcal{T}_\ell)$, respectively) satisfy $(j,k) \in I_\ell$, i.e.,

$$M_0' := \max_{j=1,\dots,\tilde{n}_\ell} |I_\ell(j)| \le 9M_0 + 7 \lesssim 1.$$

*Step three* verifies that left-hand side LHS in Proposition 2 satisfies

$$\mathrm{LHS} \lesssim \sum_{\ell=0}^{L-1} \sum_{T\in\mathcal{T}_\ell^-} (|T|^{-1}|w_\ell|^2_{H^1(T)} + |T|^{s-1}|\nabla w_\ell|^2_{H^s(T)}).$$

Given any nodal basis function $\phi_\ell^k$ in $\widetilde{A}(\mathcal{T}_\ell)$ for $\ell \in \mathbb{N}_0$ and $k = 1, \dots, \tilde{n}_\ell$, a piecewise integration by parts shows

$$a(w_\ell, \phi_\ell^k) = \sum_{T\in\mathcal{T}_\ell} \bigg( \int_{\partial T} \frac{\Delta\phi_\ell^k \partial w_\ell}{\partial \nu_T}\, ds - \int_T \nabla\Delta\phi_\ell^k \cdot \nabla w_\ell \, dx \bigg).$$

Standard trace and inverse inequalities of the quintic polynomial $\phi_\ell^k$ on $T \in \mathcal{T}_\ell$ and (7.6) of Lemma 9 for $\frac{1}{2} < s < 1$ (for each component of $\nabla w \in H^s(T;\mathbb{R}^2)$) control each summand

$$\bigg| \int_{\partial T} \frac{\Delta\phi_\ell^k \partial w_\ell}{\partial \nu_T}\, ds - \int_T \nabla\Delta\phi_\ell^k \cdot \nabla w_\ell \, dx \bigg| \lesssim |\phi_\ell^k|_{H^2(\Omega)} (h_T^{-1}|w_\ell|_{H^1(T)} + h_T^{s-1}|\nabla w_\ell|_{H^s(T)}).$$

The sum is over all $T \in \mathcal{T}_\ell$ but the summands vanish for all triangles except for $T \in \mathcal{T}_\ell(z_\ell^k)$ in the support of $\phi_\ell^k$. Step one assures that all of those triangles $T$ in the support of at least one of the basis functions $\phi_\ell^1, \dots, \phi_\ell^{\tilde{n}_\ell}$ are contained in $\mathcal{T}_\ell^-$. Step two bounds the overlap. For any fixed $\ell \in \mathbb{N}$, this leads to

$$\sum_{k=1}^{\tilde{n}_\ell} \frac{a(w_\ell, \phi_\ell^k)^2}{|\phi_\ell^k|^2_{H^2(\Omega)}} \lesssim \sum_{T\in\mathcal{T}_\ell^-} (h_T^{-2}|w_\ell|^2_{H^1(T)} + h_T^{2(s-1)}|\nabla w_\ell|^2_{H^s(T)}).$$

The sum of the lower bounds for $\ell = 0, \dots, L-1$ is LHS and the sum of the upper bounds of the proceeding estimate is asserted bound.

*Step four* recalls the $L^2$ projection $Q_j$ onto $A(\mathcal{T}_j^{\text{uniform}})$ and considers $W_j := v_L - Q_j v_L \in H_0^2(\Omega)$ and defines

$$s^2(j, T) := |T|^{-2}\|W_j\|^2_{L^2(T)} + |T|^{-1}|W_j|^2_{H^1(T)} + |T|^{s-1}|\nabla W_j|^2_{H^s(T)}$$

for any admissible triangle $T \in \bigcup \mathbb{T}$ and for any $j \in \mathbb{N}_0$. Recall the sum convention $s^2(j, \mathcal{M}) := \sum_{M\in\mathcal{M}} s^2(j, M)$ for a finite set $\mathcal{M} \subset \bigcup \mathbb{T}$ of triangles.

*Step five* establishes, for each $T \in \mathcal{T}_\ell^-$, $\ell \in \mathbb{N}_0$, that

$$|T|^{-1}|w_\ell|^2_{H^1(T)} + |T|^{s-1}|\nabla w_\ell|^2_{H^s(T)} \lesssim s^2(\min \ell(\mathcal{R}_{1,\ell}(T)), \mathcal{R}_{1,\ell}(T)).$$

Any $T \in \mathcal{T}_\ell^-$, $\ell \in \mathbb{N}_0$, is admissible ($T \in \bigcup \mathbb{T}$) and has a level $\ell(T) \in \mathbb{N}_0$ with $|T|^{-1} \approx 2^{\ell(T)}$. The quasi-interpolation $(J_\ell v_L)|_T$ of Theorem 2 depends merely on $v_L|_{\Omega(T)}$ in the neighborhood $\Omega(T) := \operatorname{int}(\bigcup \mathcal{R}_{1,\ell}(T))$ of $T$ and satisfies

$$w_\ell \equiv (1 - J_\ell)v_L = (1 - J_\ell)W_m \quad \text{a.e. in } T$$

provided $m := \min \ell(\mathcal{R}_{1,\ell}(T))$ and $W_m \equiv v_L - Q_m v_L$ from step four. (Notice that $0 = (1 - J_\ell)Q_m v_L$ in $T$ follows as in step four of the proof of Proposition 1.) This and a triangle inequalities show

$$\begin{aligned} |T|^{-\frac{1}{2}}|w_\ell|_{H^1(T)} + |T|^{\frac{s-1}{2}}|\nabla w_\ell|_{H^s(T)} &= |T|^{-\frac{1}{2}}|(1 - J_\ell)W_m|_{H^1(T)} + |T|^{\frac{s-1}{2}}|\nabla(1 - J_\ell)W_m|_{H^s(T)} \\ &\leq s(T) + |T|^{-\frac{1}{2}}|J_\ell W_m|_{H^1(T)} + |T|^{\frac{s-1}{2}}|\nabla J_\ell W_m|_{H^s(T)} \end{aligned}$$

with $s(T) := |T|^{-\frac{1}{2}}|W_m|_{H^1(T)} + |T|^{\frac{s-1}{2}}|\nabla W_m|_{H^s(T)}$ as a temporary abbreviation. The inverse estimate (7.7) applies to each component of $\nabla J_\ell W_m \in P_4(T;\mathbb{R}^2)$. This and a standard inverse estimate show

$$|T|^{\frac{s-1}{2}}|\nabla J_\ell W_m|_{H^s(T)} \lesssim |T|^{-\frac{1}{2}}|J_\ell W_m|_{H^1(T)} \lesssim |T|^{-1}\|J_\ell W_m\|_{L^2(T)}.$$

The combination with the previous estimate plus Theorem 2 (c) with $f = W_m \equiv v_L - Q_m v_L$ in the neighborhood $\Omega(T) \equiv \operatorname{int}(\bigcup \mathcal{R}_{1,\ell}(T))$ of $T$ leads to

$$\begin{aligned} |T|^{-\frac{1}{2}}|w_\ell|_{H^1(T)} + |T|^{\frac{s-1}{2}}|\nabla w_\ell|_{H^s(T)} &\lesssim s(T) + |T|^{-1}\|J_\ell W_m\|_{L^2(T)} \\ &\lesssim s(T) + |T|^{-1}\|W_m\|_{L^2(\Omega(T))} + |T|^{-\frac{1}{2}}|W_m|_{H^1(\Omega(T))}. \end{aligned}$$

This and $|T| \approx |K|$ for $K \in \mathcal{R}_{1,\ell}(T)$ prove the assertion.

*Step six* concludes the proof. Steps two and four plus Lemma 3 show

$$\text{LHS} \lesssim \sum_{\ell=0}^{L-1} \sum_{T\in\mathcal{T}_\ell^-} s^2(\min \ell(\mathcal{R}_{1,\ell}(T)), \mathcal{R}_{1,\ell}(T)) \lesssim \sum_{m=0}^{\infty} \sum_{j=(m-M_2)_+}^{m} s^2(j, \mathcal{T}_m^{\text{uniform}}).$$

Step four in the proof of Proposition 1 explains $|T| \approx 2^{-m} \approx 2^{-j}$ for any $T \in \mathcal{T}_m^{\text{uniform}}$ in the sum over all $j = (m - M_2)_+, \dots, m$ with a bounded number of summands. This and (7.5) show the equivalence of the last upper bound to

$$\sum_{j=0}^{\infty} 2^{2j}\|W_j\|^2_{L^2(\Omega)} + \sum_{j=0}^{\infty} 2^{j}|W_j|^2_{H^1(\Omega)} + \sum_{j=0}^{\infty} 2^{(s-1)j}|\nabla W_j|^2_{H^s(\Omega)}.$$

Lemma 8.b applies to each of the three sums for $r = 0, 1, 1+s$ for $v$ replaced by $v_L$ and conclude the proof of LHS $\lesssim |v_L|^2_{H^2(\Omega)}$. □

## 7.5 Third Estimate of a Decomposition

This subsection concerns an arbitrary discrete function $\tilde{v}_\ell \in \widetilde{A}(\mathcal{T}_\ell)$ on fixed level $\ell \in \mathbb{N}$ (uniquely) rewritten as

$$\tilde{v}_\ell = \sum_{k=1}^{\tilde{n}_\ell} v_\ell^k \quad \text{with } v_\ell^k \in \operatorname{span}\{\phi_\ell^k\} \quad \text{for all } k = 1, \dots, \tilde{n}_\ell. \tag{7.8}$$

Recall that $P_\ell^j$ is the local Galerkin projection from (5.3).

**Proposition 3** (Remains). *Any function $\tilde{v}_\ell \in \widetilde{A}(\mathcal{T}_\ell)$ with (7.8) satisfies*

$$\sum_{j=1}^{\tilde{n}_\ell} \Big\| P_\ell^j \sum_{k=j+1}^{\tilde{n}_\ell} v_\ell^k \Big\|^2_{H^2(\Omega)} \lesssim \|h_\ell^{-2} v_\ell\|^2_{L^2(\Omega)}.$$

*Proof.* Recall the notation from step two in the proof of Proposition 2 on the finite overlap of the supports of the basis functions $\phi_\ell^1, \phi_\ell^2, \dots, \phi_\ell^{\tilde{n}_\ell}$ and the universal bound $M_0'$. Abbreviate $w_j := \sum_{k=j+1}^{\tilde{n}_\ell} v_\ell^k$ and evaluate the norm of $P_\ell^j w_j$ from (5.3) for $j = 1, \dots, \tilde{n}_\ell$. Triangle and Cauchy inequalities verify

$$|P_\ell^j w_j|_{H^2(\Omega)} \le \sum_{k \in I_\ell(j)} |v_\ell^k|_{H^2(\Omega)} \le \sqrt{M_0'} \sqrt{\sum_{k \in I_\ell(j)} |v_\ell^k|^2_{H^2(\Omega)}}.$$

The equivalence of norms $\|P_\ell^j w_j\|_{H^2(\Omega)} \approx |P_\ell^j w_j|_{H^2(\Omega)}$ in $H_0^2(\Omega)$, the square of the last displayed identity, and its sum over all $j = 1, \dots, \tilde{n}_\ell$ lead to an estimate for the left-hand side LHS in Proposition 3, namely,

$$\text{LHS} \equiv \sum_{j=1}^{\tilde{n}_\ell} \Big\| P_\ell^j \sum_{k=k+1}^{\tilde{n}_\ell} v_\ell^n \Big\|^2_{H^2(\Omega)} \approx \sum_{j=1}^{\tilde{n}_\ell} |P_\ell^j w_j|^2_{H^2(\Omega)} \le \sum_{j=1}^{\tilde{n}_\ell} \sum_{k \in I_\ell(j)} |v_\ell^k|^2_{H^2(\Omega)}.$$

The upper bound is a sum of $|v_\ell^k|^2_{H^2(\Omega)}$ over all pairs $(j, k) \in I_\ell$ and equal to

$$\sum_{k=1}^{\tilde{n}_\ell} \sum_{j \in I_\ell(k)} |v_\ell^k|^2_{H^2(\Omega)} = \sum_{k=1}^{\tilde{n}_\ell} |I_\ell(k)|\, |v_\ell^k|^2_{H^2(\Omega)} \le M_0' \sum_{k=1}^{\tilde{n}_\ell} |v_\ell^k|^2_{H^2(\Omega)}.$$

Consequently,

$$\text{LHS} \lesssim \sum_{k=1}^{\tilde{n}_\ell} |v_\ell^k|^2_{H^2(\Omega)} = \sum_{T \in \mathcal{T}_\ell} \sum_{k=1}^{\tilde{n}_\ell} |v_\ell^k|^2_{H^2(T)}.$$

For a fixed $T \in \mathcal{T}_\ell$, an inverse estimate

$$|v_\ell^k|_{H^2(T)} \lesssim h_T^{-2} \|v_\ell^k\|_{L^2(T)}$$

is followed by Lemma 1. Let $I(\ell, T) \subset \{1, \dots, \tilde{n}_\ell\}$ denote the set of at most 21 indices $j$ such that $z_\ell^j \in T$ and its associated degree of freedom $\Lambda_\ell^j \in A(\mathcal{T}_\ell)^*$ is a derivative of order $\mu_\ell^j \in \{0, 1, 2\}$ evaluated at $z_\ell^j$. The split (7.8) and the duality of $\phi_\ell^1, \dots, \phi_\ell^{\tilde{n}_\ell}$ and $\Lambda_\ell^1, \dots, \Lambda_\ell^{\tilde{n}_\ell}$ leads to $\Lambda_\ell^k(\tilde{v}_\ell)\phi_\ell^k = v_\ell^k$ and

$$\tilde{v}_\ell = \sum_{k=1}^{\tilde{n}_\ell} v_\ell^k = \sum_{k \in I(\ell,T)} v_\ell^k = \sum_{k \in I(\ell,T)} \Lambda_\ell^k(\tilde{v}_\ell)\phi_\ell^k \quad \text{a.e. in } T.$$

This is the identity (6.1) when $\phi_\ell^j$, $\Lambda_\ell^j$, $\tilde{n}_\ell$, $\tilde{v}_\ell$, and $I(\ell, T)$ replace $\psi_j$, $\Lambda_j$, $N$, $v_A$, and $I(T)$. In the current notation Lemma 1 therefore establishes

$$\|\tilde{v}_\ell\|_{L^2(T)}^2 \approx \sum_{k\in I(\ell,T)} |T|^{1+\mu_\ell^k} |\Lambda_\ell^k(\tilde{v}_\ell)|^2.$$

Replace $v_A$ by $v_\ell^k$ in Lemma 1 for the proof of $\|v_\ell^k\|_{L^2(T)}^2 \approx |T|^{1+\mu_\ell^k}|\Lambda_\ell^k(\tilde{v}_\ell)|^2$. The combination of the previous estimates verifies that

$$\sum_{k\in I(\ell,T)} \|v_\ell^k\|_{L^2(T)}^2 \approx \|\tilde{v}_\ell\|_{L^2(T)}^2.$$

This and the aforementioned inverse inequality result in

$$|T|^2 \sum_{k=1}^{\tilde{n}_\ell} |v_\ell^k|_{H^2(T)}^2 \lesssim \sum_{k\in I(\ell,T)} \|v_\ell^k\|_{L^2(T)}^2 \lesssim \|\tilde{v}_\ell\|_{L^2(T)}^2$$

for each $T \in \mathcal{T}_\ell$. The sum over all $T \in \mathcal{T}_\ell$ provides the second estimate in

$$\mathrm{LHS} \lesssim \sum_{k=1}^{\tilde{n}_\ell} \|v_\ell^k\|_{\mathcal{A}}^2 \lesssim \|h_\ell^{-2}\tilde{v}_\ell\|_{L^2(\Omega)}^2.$$

This concludes the proof. □

## 7.6 Proof of Theorem 7

The multigrid V-cycle operator $\mathcal{B}_L$ of Section 5.1 allows for a compact representation of the self-adjoint operator

$$I - \mathcal{B}_L\mathcal{A}_L = R^*R \quad \text{for } R := (I - \mathcal{P}_0)\prod_{\ell=1}^{L}\prod_{k=1}^{\tilde{n}_\ell}(I - P_\ell^k)$$

with adjoint $R^*$ in the Hilbert space $H_0^2(\Omega)$ with the energy inner product $a(\bullet,\bullet)$ and induced norm $\|\bullet\|_{\mathcal{A}} := a(\bullet,\bullet)^{\frac{1}{2}}$; cf., e.g., [5, equation (3.4)] for proofs and details. Recall that $\mathcal{P}_0$ is the Galerkin projection $H_0^2(\Omega)$ onto $A(\mathcal{T}_0)$ from (5.2), and $P_\ell^k$ is the local Galerkin projection from (5.3). The square of the operator norm

$$\|I - \mathcal{B}_L\mathcal{A}_L\|_{\mathcal{A}}^2 = \left\|(I - \mathcal{P}_0)\prod_{\ell=1}^{L}\prod_{k=1}^{\tilde{n}_\ell}(I - P_\ell^k)\right\|_{\mathcal{A}}^2$$

has been characterized in [29, Corolary 4.3] as $\frac{c_0}{1+c_0}$ for an expression $c_0$ that reads in the notation of this subsection as

$$c_0 = \sup_{\|v\|_{\mathcal{A}}=1} \inf_{v=v_0+\sum_{\ell=1}^{L}\sum_{k=1}^{\tilde{n}_\ell} v_\ell^k} \left(\left\|\mathcal{P}_0 \sum_{\ell=1}^{L}\sum_{k=1}^{\tilde{n}_\ell} v_\ell^k\right\|_{\mathcal{A}}^2 + \sum_{\ell=1}^{L}\sum_{k=1}^{\tilde{n}_\ell}\left\|P_\ell^k \sum_{(m,n)>(\ell,k)} v_m^n\right\|_{\mathcal{A}}^2\right).$$

Given any $L \in \mathbb{N}$ and any $v_L \in A(\mathcal{T}_L)$ with norm one in the semi-norm of $H_0^2(\Omega)$, the infimum in the definition of $c_0$ runs over all decompositions

$$v_L = v_0 + \sum_{\ell=1}^{L}\sum_{k=1}^{\tilde{n}_\ell} v_\ell^k \in A(\mathcal{T}_L)$$

with $v_\ell^k \in \operatorname{span}\{\phi_\ell^k\}$ in the direction of the nodal basis function $\phi_\ell^k$ selected in Section 5.2 for the smoother in (5.3)–(5.4). The symbol $>$ represents the lexicographic order and $(m,n) > (\ell,k)$ under the sum indicates a summation over all $m = \ell, \dots, L$ and $n = 1, \dots, \tilde{n}_m$ with either $m > \ell$ or ($m = \ell$ and $n > k$).

The bound of $c_0$ from above follows with the particular choice (7.1); this decomposition may not be infimal in the definition of $c_0$, but it leads to an upper bound $c_0 \lesssim 1$ in the end and so concludes the proof.

The details of the proof of $c_0 \lesssim 1$ depart with the definition as a supremum. The lower bound $\frac{c_0}{2}$ of the supremum $c_0 > 0$ (there is nothing left to prove if $c_0 = 0$) allows for $L \in \mathbb{N}_0$ and $v_L \in A(\mathcal{T}_L)$ with norm one in $H_0^2(\Omega)$ such that the infimum over all decompositions is greater than $\frac{c_0}{2}$. Theorem 9 (a’) guarantees that the split (7.1) is admissible in that

$$\tilde{v}_\ell := (J_\ell - J_{\ell-1})v_L \in \widetilde{A}(\mathcal{T}_\ell) = \operatorname{span}\{\phi_\ell^1, \dots, \phi_\ell^{\tilde{n}_\ell}\} \quad \text{for all } \ell = 1, \dots, L$$

can be written uniquely as a sum over $v_\ell^k \in \operatorname{span}\{\phi_\ell^k\}$ for all $k = 1, \dots, \tilde{n}_\ell$ and $\ell = 1, \dots, L$ so that $v_0 =: J_0 v_L$ and

$$v_L = v_0 + \sum_{\ell=1}^{L} \tilde{v}_\ell \quad \text{and} \quad \tilde{v}_\ell = \sum_{k=1}^{\tilde{n}_\ell} v_\ell^k \in \widetilde{A}(\mathcal{T}_\ell). \tag{7.9}$$

Abbreviate $\sum_{\ell,k} := \sum_{\ell=1}^{L} \sum_{k=1}^{\tilde{n}_\ell}$ in the sequel. Then (7.9) (and triangle and Young inequalities) leads to an upper bound

$$\frac{c_0}{2} \le \Big\| \mathcal{P}_0 \sum_{\ell,k} v_\ell^k \Big\|_{\mathcal{A}}^2 + 2 \sum_{\ell,k} \Big\| P_\ell^k \sum_{j=k+1}^{\tilde{n}_\ell} v_\ell^j \Big\|_{\mathcal{A}}^2 + 2 \sum_{\ell,k} \Big\| P_\ell^k \sum_{m=\ell+1}^{L} \sum_{n=1}^{\tilde{n}_m} v_m^n \Big\|_{\mathcal{A}}^2 .$$

The three sums in the upper bounds are named $S_1$, $S_2$, and $S_3$ in the following analysis that concludes the proof by $S_j \lesssim \|v_L\|_{H^2(\Omega)}^2 \approx 1$ for $j = 1, 2, 3$.

*Proof of $S_1 \lesssim 1$.* The first sum concerns $\sum_{\ell,k} v_\ell^k = v_L - J_0 v_L$. This, the stability of the Galerkin projection $\mathcal{P}_0$, and the stability of the quasi–interpolation operator $J_0$ of Theorem 2 show

$$S_1 \equiv \|\mathcal{P}_0 \sum_{\ell,k=1} v_\ell^k\|_{\mathcal{A}}^2 = \|\mathcal{P}_0(v_L - J_0 v_L)\|_{\mathcal{A}}^2 \lesssim \|v_L\|_{H^2(\Omega)}^2 = 1.$$

*Proof of $S_2 \lesssim 1$.* The second sum involves $\tilde{v}_\ell \in \widetilde{A}(\mathcal{T}_\ell)$ in Proposition 3 and shows

$$S_2 \equiv 2 \sum_{\ell,k}^{\tilde{n}_\ell} \|P_\ell^k \sum_{j=k+1}^{\tilde{n}_\ell} v_\ell^j\|_{\mathcal{A}}^2 \lesssim \sum_{\ell=1}^{L} \|h_\ell^{-2} \tilde{v}_\ell\|_{L^2(\Omega)}^2 = \sum_{\ell=1}^{L} \|h_\ell^{-2}(J_\ell - J_{\ell-1}) v_L\|_{L^2(\Omega)}^2 .$$

The last identity follows from the definition of $\tilde{v}_\ell$ and allows for Proposition 1 to conclude $S_2 \lesssim \|v_L\|_{H^2(\Omega)}^2 = 1$.

*Proof of $S_3 \lesssim 1$.* The third sum concerns (for $\ell = 1, \dots, L$)

$$\sum_{m=\ell+1}^{L} \sum_{n=1}^{\tilde{n}_m} v_m^n = \sum_{m=\ell+1}^{L} \tilde{v}_m = v_L - J_\ell v_L$$

and the norm $\|P_\ell^k(v_L - J_\ell v_L)\|_{\mathcal{A}}$ of its projection $P_\ell^k(v_L - J_\ell v_L)$ with

$$\|P_\ell^k(v - J_\ell v)\|_{\mathcal{A}}^2 = \frac{a(v_L - J_\ell v_L, \phi_\ell^k)^2}{|\phi_\ell^k|_{H^2(\Omega)}^2}.$$

This and Proposition 2 (note that $v_L - J_L v_L = 0$) show

$$S_3 \equiv 2 \sum_{\ell,k} \|P_\ell^k \sum_{m=\ell+1}^{L} \sum_{m=1}^{\tilde{n}_m} v_m^n\|_{\mathcal{A}}^2 = 2 \sum_{\ell,k} \frac{a(v_L - J_\ell v_L, \phi_\ell^k)^2}{|\phi_\ell^k|_{H^2(\Omega)}^2} \lesssim 1.$$

*Finish of the proof.* The summary of $\frac{c_0}{2} \le S_1 + S_2 + S_3 \lesssim 1$ concludes the proof. □

**Funding:** The first author was supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project-ID 255510958 – SPP 1748. The second author was supported by NSFC projects 11625101. The authors were partially supported by the Sino-German Center for Research Promotion under Grant GZ 1571.

## References

[1] C. Bacuta, J. H. Bramble and J. E. Pasciak, Shift theorems for the biharmonic Dirichlet problem, in: *Recent Progress in Computational and Applied PDEs*, Kluwer/Plenum, New York (2002), 1–26.

[2] P. Binev, W. Dahmen and R. DeVore, Adaptive finite element methods with convergence rates, *Numer. Math.* **97** (2004), no. 2, 219–268.

[3] A. Bonito and R. H. Nochetto, Quasi-optimal convergence rate of an adaptive discontinuous Galerkin method, *SIAM J. Numer. Anal.* **48** (2010), no. 2, 734–771.

[4] D. Braess, *Finite Elements. Theory, Fast Solvers, and Applications in Elasticity Theory*, Cambridge University, Cambridge, 2007.

[5] J. H. Bramble, *Multigrid Methods*, Pitman Res. Notes Math. 294, Longman, Harlow, 1993.
[6] S. C. Brenner and C. Carstensen, Finite Element Methods, in: *Encyclopedia of Computational Mechanics. Second Edition*, John Wiley & Sons, Chichester (2018), 1–47.
[7] S. C. Brenner, L. Owens and L.-Y. Sung, A weakly over-penalized symmetric interior penalty method, *Electron. Trans. Numer. Anal.* **30** (2008), 107–127.
[8] S. C. Brenner and L. R. Scott, *The Mathematical Theory of Finite Element Methods*, Texts Appl. Math. 15, Springer, New York, 2008.
[9] S. C. Brenner and L.-Y. Sung, $C^0$ interior penalty methods for fourth order elliptic boundary value problems on polygonal domains, *J. Sci. Comput.* **22/23** (2005), 83–118.
[10] C. Carstensen, M. Feischl, M. Page and D. Praetorius, Axioms of adaptivity, *Comput. Math. Appl.* **67** (2014), no. 6, 1195–1253.
[11] C. Carstensen, D. Gallistl and J. Hu, A discrete Helmholtz decomposition with Morley finite element functions and the optimality of adaptive finite element schemes, *Comput. Math. Appl.* **68** (2014), 2167–2181.
[12] C. Carstensen and H. Rabus, Axioms of adaptivity with separate marking for data resolution, *SIAM J. Numer. Anal.* **55** (2017), no. 6, 2644–2665.
[13] J. M. Cascon, C. Kreuzer, R. H. Nochetto and K. G. Siebert, Quasi-optimal convergence rate for an adaptive finite element method, *SIAM J. Numer. Anal.* **46** (2008), no. 5, 2524–2550.
[14] L. Chen, R. H. Nochetto and J. Xu, Optimal multilevel methods for graded bisection grids, *Numer. Math.* **120** (2012), no. 1, 1–34.
[15] P. G. Ciarlet, *The Finite Element Method for Elliptic Problems*, Stud. Math. Appl. 4, North-Holland, Amsterdam, 1978.
[16] P. Clément, Approximation by finite element functions using local regularization, *Rev. Franc. Automat. Inform. Rech. Operat. R* **9** (1975), 77–84.
[17] D. A. Di Pietro and A. Ern, *Mathematical Aspects of Discontinuous Galerkin Methods*, Math. Appl. (Berlin) 69, Springer, Heidelberg, 2012.
[18] A. Ern and J.-L. Guermond, Finite element quasi-interpolation and best approximation, *ESAIM Math. Model. Numer. Anal.* **51** (2017), no. 4, 1367–1385.
[19] V. Girault and L. R. Scott, Hermite interpolation of nonsmooth functions preserving boundary conditions, *Math. Comp.* **71** (2002), no. 239, 1043–1074.
[20] J. Hu, Z. Shi and J. Xu, Convergence and optimality of the adaptive Morley element method, *Numer. Math.* **121** (2012), no. 4, 731–752.
[21] J. Morgan and R. Scott, A nodal basis for $C^1$ piecewise polynomials of degree $n \geq 5$, *Math. Comp.* **29** (1975), 736–740.
[22] S. A. Nazarov and B. A. Plamenevsky, *Elliptic Problems in Domains with Piecewise Smooth Boundaries*, Walter de Gruyter, Berlin, 1994.
[23] R. H. Nochetto, K. G. Siebert and A. Veeser, Theory of adaptive finite element methods: an introduction, in: *Multiscale, Nonlinear and Adaptive Approximation*, Springer, Berlin (2009), 409–542.
[24] R. Stevenson, The completion of locally refined simplicial partitions created by bisection, *Math. Comp.* **77** (2008), no. 261, 227–241.
[25] S. Tang and X. Xu, Local multilevel methods with rectangular finite elements for the biharmonic problem, *SIAM J. Sci. Comput.* **39** (2017), no. 6, A2592–A2615.
[26] R. Verfürth, *A Review of a Posteriori Error Estimation and Adaptive Mesh-Refinement Techniques*, John Wiley & Sons, Chichester, 1996.
[27] H. Wu and Z. Chen, Uniform convergence of multigrid V-cycle on adaptively refined finite element meshes for second order elliptic problems, *Sci. China Ser. A* **49** (2006), no. 10, 1405–1429.
[28] J. Xu, The auxiliary space method and optimal multigrid preconditioning techniques for unstructured grids, *Computing* **56** (1996), 215–235.
[29] J. Xu and L. Zikatanov, The method of alternating projections and the method of subspace corrections in Hilbert space, *J. Amer. Math. Soc.* **15** (2002), no. 3, 573–597.
[30] X. Zhang, Multilevel Schwarz methods for the biharmonic Dirichlet problem, *SIAM J. Sci. Comput.* **15** (1994), 621–644.
[31] X. Zhao, J. Hu and Z. Shi, Convergence analysis of the adaptive finite element method with the red-green refinement, *Sci. China Math.* **53** (2010), no. 2, 499–512.