Humboldt-Universität zu Berlin
Mathematisch-Naturwissenschaftliche Fakultät
Institut für Mathematik

Conforming multilevel FEM for the biharmonic equation

Masterarbeit

zur Erlangung des akademischen Grades
Master of Science (M.Sc.)

eingereicht von: Benedikt Konstantin Gräßle
eingereicht am: 01.07.2021
Gutachter/innen: Prof. Carsten Carstensen
Dr. Hella Rabus

Abstract

A multigrid V-cycle with local smoothing is considered with an efficient and reliable estimator of the algebraic error. This gives rise to an efficient iterative solver for the *adaptive finite element method* (AFEM) with optimal time complexity. This thesis extends a posteriori error analysis for the hierarchical Argyris *finite element method* (FEM) to the biharmonic equation with inhomogeneous and mixed boundary conditions. Optimal convergence of the hierarchical Argyris AFEM with direct solve follows with the axioms of adaptivity by observing a best-approximation property for the Argyris interpolant of the essential boundary data. Numerical validation is presented for optimal rates of AFEM together with a comparison between a direct solver and the local multigrid solver. Model benchmarks include different boundary conditions, point loads and highlight the strength of the lowest-order conforming Argyris FEM. A conclusion underlines the rehabilitation of the Argyris element in conjunction with the extended Argyris space.

Contents

1 Introduction

Motivation: Developed in the past century, the *finite element method* (FEM) became one of the most widely used discretisation schemes for PDEs. It provides a variety of advantages over the classical *finite difference method* (FDM) to its user. From the engineer's perspective this may be the possibility to work on complex and unstructured grids whereas numerical analysts praise its vast mathematical foundation that requires less assumption to even prove optimality results. Since approximations obtained by FEM shall eventually be used to draw information about the original problem, the ability to estimate the error of the approximation is of huge importance. But more than that, a good error estimator also allows to geometrically locate parts of the underlying mesh where the error is strongest. Thus, in their role as a refinement indicator, they may be used in combination with adaptive mesh-refinement algorithms to drastically improve convergence. The resulting *adaptive finite element method* (AFEM) has been proven (and shown) to achieve optimal convergence rates for a variety of finite element schemes and differential equations.

Entry to the world of FEM is typically paved with elliptic second-order model problems that have been thoroughly analysed and for which a variety of finite element schemes, together with their dis- and advantages, have been established [BS08; BC17]. A PDE is always a model under certain, often idealistic, assumptions. So, more sophisticated models or models of more complex physical processes may not be formulated as second-order PDEs. One such example can be found in plate theory that provides different models for the deformation and stresses in thin elastic structures, called plates. The model problem here is the two-dimensional biharmonic equation [Cia88] – an elliptic fourth-order differential equation.

A conforming discretisation of the biharmonic problem with the FEM requires finite elements of class C^1. The possibly most popular triangular finite element of class C^1 is the Argyris element [AFS68] with its 21 *degrees of freedom* (dofs) for the full space $P_5(T)$ of quintic polynomials over a given triangle T. The high polynomial order is often referred to as a disadvantage of the Argyris element, arguing that optimal order of convergence requires strong smoothness assumptions that cannot be expected. So the implementational troubles would not be compensated by higher-order convergence. This motivated the design of many other elements such as the Bell [Bel69] triangle (obtained from the Argyris element by removing three dofs), composite elements such as the Clough-Tocher (CT) element [CT66] (C^1-piecewise cubic polynomials and 12 degrees of freedom), to only name a few. The reduced number of degrees of freedom is accompanied with the loss of 1 or 2 orders in convergence for the Bell and CT elements respectively. Severe difficulties due to the non-nestedness of the Argyris space are a mayor reason that hindered the development of Argyris AFEM until the preceding result of [CH21].

Nowadays, progress regarding the implementational aspects [DS08; Kir18] has been made such that many software packages, at least partially, support Argyris out of the box, whereas implementations for the more exotic composite elements, sometimes called macro elements, are notably less common. Furthermore, recent success in Carstensen and Hu [CH21] proves optimal rates for a minimal extension of the Argyris space by overcoming the obstacle of non-nestedness. This can be understood as the starting point of this thesis, motivated by a study of the resulting consequences. Though, a first consequence is already clear, the claimed disadvantage of the Argyris element's high order has been turned into a serious advantage as it directly contributes to the resulting convergence speed. Another important question is, how this result extends to other properties of the Argyris element.

Since the degrees of freedom allow direct interpolation of boundary data, the extension of [CH21], which considers homogeneous boundary conditions, to the incorporation of inhomogeneous boundary conditions will result in a strong justification for the Argyris element. One part of this thesis is therefore devoted to the treatment of inhomogeneous boundary conditions.

Model problem: The Kirchhoff-Love model is a prototypical model for the bending of thin plates and assumes that the full response of the plate subject to forces and moments can be represented by its mid-surface, leading to a two-dimensional model. While the theory, see, e.g., [Cia88; TW87], incorporates material parameters such as the Poisson ratio σ, they are not deemed of interest and the parameter-free description is presented.

Consider a two-dimensional domain $\Omega \subset \mathbb{R}^2$, the plate, whose boundary $\partial\Omega = \Gamma_{\mathrm{C}} \dot\cup \Gamma_{\mathrm{S}} \dot\cup \Gamma_{\mathrm{F}}$ is split into three disjoint parts. The *clamped boundary* Γ_C represents the portion of $\partial\Omega$ where displacement and bending of the plate are prescribed, whereas on the *simply-supported boundary* Γ_S only displacement is fixed. On the remaining *free boundary* $\Gamma_F = \partial\Omega \setminus (\Gamma_C \cup \Gamma_S)$ no such conditions are imposed. These requirements on $u : \Omega \to \mathbb{R}$, describing the global displacement of the plate, are called essential boundary conditions and are given by $g_{\mathrm{CS}} : \Gamma_C \cup \Gamma_S \to \mathbb{R}$ (for the displacement) and $g_{\mathrm{C}} : \Gamma_C \to \mathbb{R}$ (for the bending). Implicitly, these boundary data result in natural boundary conditions in the form of $t_{\mathrm{F}} : \Gamma_F \to \mathbb{R}$ and $t_{\mathrm{SF}} : \Gamma_F \cup \Gamma_S \to \mathbb{R}$.

The strong form of the plate bending problem, given force $f : \Omega \to \mathbb{R}$ and inhomogeneous *essential boundary data* $g_{\mathrm{CS}}, g_{\mathrm{S}}$ and *natural boundary data* $t_{\mathrm{F}}, t_{\mathrm{SF}}$, seeks the displacement u with

$$\begin{aligned}
\Delta^2 u &= f && && && \text{in } \Omega,\\
u &= g_{\mathrm{CS}} && \text{and} && \partial_\nu u = g_{\mathrm{C}} && \text{on } \Gamma_{\mathrm{C}},\\
u &= g_{\mathrm{CS}} && \text{and} && \partial^2_{\nu\nu} u = t_{\mathrm{SF}} && \text{on } \Gamma_{\mathrm{S}},\\
\partial^2_{\nu\nu} u &= t_{\mathrm{SF}} && \text{and} && \partial^3_{\tau\tau\nu} u + \partial_\nu \Delta u = t_{\mathrm{F}} && \text{on } \Gamma_{\mathrm{F}}.
\end{aligned} \tag{SF}$$

Adaptive FEM: Adaptive mesh refinement is one of the main ingredients that lead to superiority of the *adaptive finite element method* (AFEM) over the naive FEM with uniform

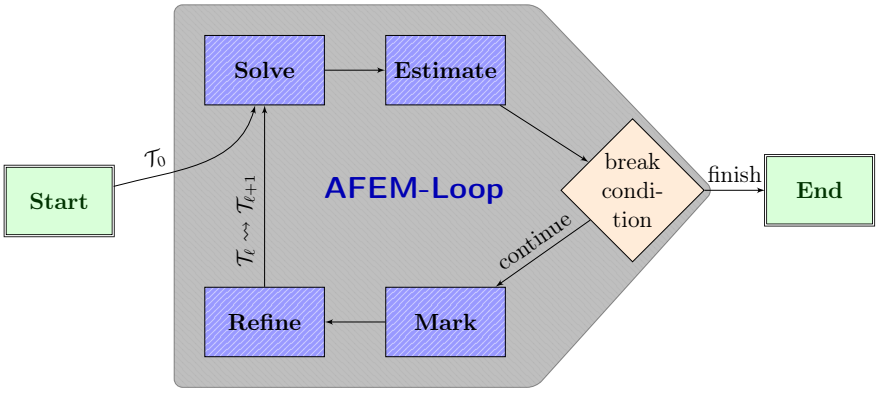

Figure 1: The AFEM loop with initial triangulation $\mathcal{T}_0$

refinement. The reason for this lies in the irregular behavior of solutions near singularities like, e.g., the corner singularity of the L-shaped domain. Whereas the adaptive algorithm locates singularities and concentrates the refinement towards them, uniform discretisations

would waste huge amounts of computational power and eventually lead to suboptimal convergence altogether. Figure 1 shows the four steps of the AFEM loop. A discrete approximation over the current triangulation $\mathcal{T} \equiv \mathcal{T}_\ell$ is computed in the *Solve* step. The next step, *Estimate*, uses this approximation and the aforementioned error estimator as refinement indicator to find parts of the mesh that have a high contribution towards the total error. In the step *Mark*, Dörfler marking is applied to mark a number of triangles for refinement. Then, after a necessary closure step to maintain regularity of the triangulations, the *Refine* step produces a refinement $\widehat{\mathcal{T}} \equiv \mathcal{T}_{\ell+1}$ of $\mathcal{T}$ with all the marked triangles refined. The applied method for refinement is the *newest vertex bisection* (NVB).

Moreover, with the *axioms of adaptivity* an abstract framework to prove optimal convergence rates exists [Car+14]. The computational compexity of one AFEM cycle can be as good as $\mathcal{O}(N)$, i.e., linear in the number of dofs N corresponding to a triangulation $\mathcal{T}$ if an appropriate solver is used [Car19].

Multilevel solver: Multilevel solvers build a class of iterative solvers that are specifically designed for discretisations of PDEs. The probably most known instance is the *multigrid* algorithm for the solution of self-adjoint elliptic PDEs. Multigrid is a linear iterative solver and can be seen as a prototype for a variety of multilevel solvers like, e.g., *domain decomposition methods* or the *method of subspace corrections* to only name a few [Xu92].

Consider the matrix equation $\mathbf{Ax} = \mathbf{b}$ for a regular matrix $\mathbf{A} \in \mathbb{R}^{N\times N}, \mathbf{b} \in \mathbb{R}^N$ and $N \in \mathbb{N}$. A linear iterative solver applies the iteration scheme

$$\mathbf{x}^{n+1} = \mathbf{x}^n - \mathbf{B}(\mathbf{A}\mathbf{x}^n - \mathbf{b}) \tag{1}$$

given some initial guess $\mathbf{x}^0$, cf. [Hac16], and computes a sequence of approximations $(\mathbf{x}^n)_n$ of $\mathbf{x}$. Thus, the iterative process is defined by the *approximate inverse* $\mathbf{B}$ of $\mathbf{A}$. The matrix $\mathbf{I} - \mathbf{BA}$ is called the *iteration matrix*. Considering (1) simply as an algebraic problem, classical iterative methods such as *Richardson* or *Gauß-Seidel* iterations come to mind. In the case of Gauß-Seidel, the approximative inverse $\mathbf{B} = (\mathbf{L} + \mathbf{D})^{-1}$ is given in terms of the lower triangular component $\mathbf{L}$ and diagonal $\mathbf{D}$ of $\mathbf{A}$ and can be efficiently computed using *forward substitution*. However, convergence of these basic iterative methods can be shown to heavily depend on the condition of the problem matrix $\mathbf{A}$.

In the case, that $\mathbf{A}$ is obtained by a discretisation of the plate problem, the condition number $\operatorname{cond}(\mathbf{A}) \approx h^{-4}$ for uniform meshes with mesh size h renders the application of these schemes virtually impossible. So the use of those alone for solving the system on a discretisation $\mathcal{T}$ is not fruitful. Conversely, these schemes do very well in reducing the high-oscillatory components of the error [Saa03; Hac16]. The difference to the purely algebraic setting is that for PDEs one typically has additional geometric information such as a sequence of triangulations $(\mathcal{T}_\ell)_{\ell=0}^L$ that result in $\mathcal{T} = \mathcal{T}_L$, e.g., by using AFEM. Multigrid takes advantage of previous discretisations $\mathcal{T}_0, \dots, \mathcal{T}_{L-1}$ to interpolate the approximation back and forth between subsequent discretisations. At each level $\mathcal{T}_\ell$ the application of the basic solvers damps the error (by removing the high oscillatory components at level ℓ) and results in an overall convergent scheme.

This incorporation of the geometric information is what made multigrid superior when it was developed (first for FDM and eventually for FEM). When adaptive mesh refinement was not as widely used, this algorithm came natural and reasonable. Now, with AFEM and locally instead of uniformly refined meshes, huge improvements can be made by modifying this algorithm such that the basic solvers at each level only act on the refined portion of the mesh. Hence, the resulting *local multigrid* solver perfectly fits into the AFEM loop.

Moreover, a second significant result in [CH21] is the proof of uniform convergence of the local multigrid method for the hierarchical Argyris AFEM.

It is worth noting, that multigrid cannot be only used as a solver but also renders a good preconditioner for other iterative solvers such as the *conjugate gradient method* (CG) and derivatives [Saa03].

Structure: After clarification of used notation and recitation of some auxillary results in section 2, the local multigrid method is introduced in section 3 in more detail. The focus here does not lie on a full picture nor replaces study of the literature, e.g., [Bra93; BZ95; Saa03; Hac16]. Instead a short, but contemporary, description of the multigrid method as representative for general multilevel methods is formulated. Hence, some assumptions are made that are either necessary for the result or the compactness of the presentation. One such assumption will be nestedness which holds for the extended Argyris spaces. Nevertheless, the introduction tries to be general so that it is directly applicable to other FEM as well and simultaneously providing an abstract introduction of the setting in that general PDE discretisations are formulated. Section 4 then introduces the Argyris finite element and the Argyris function space. The focus lies on the extended Argyris space from [CH21], but it will be clear, which results directly translate to the classical Argyris space. A reason to center around the extended Argyris space is that its nestedness will prove crutial in the a posteriori error analysis, presented in section 5. The theoretical results are then complemented by numerical results in section 6. Different benchmarks will provide numerical evidence to the results in [CH21] regarding optimal rates and uniform approximation of the multigrid algorithm as well as extensibility to other boundary situation as considered in this thesis.

2 Notation and Preliminaries

A domain $D \subset \mathbb{R}^d$ is an open and connected set. Throughout this thesis, a bounded Lipschitz domain $\Omega \subset \mathbb{R}^2$ with polynomial boundary $\partial\Omega$ is fixed. Its outer unit normal is denoted by ν and τ refers to the unit tangential vector with fixed orientation. Since Ω is a polygon, its boundary $\partial\Omega$ consists of $N_{\partial\Omega} \in \mathbb{N}_0 := \mathbb{N}\dot\cup\{0\}$ straight line segments $\Gamma_1, \dots, \Gamma_{N_{\partial\Omega}}$, called the edges $\mathcal{E}_\Omega = \{\Gamma_1, \dots, \Gamma_{N_{\partial\Omega}}\}$ of Ω. For simplicity it is assumed that the edges are enumerated counter-clockwise so that the corners of Ω are $\mathcal{V}_\Omega = \{P_j := \Gamma_j \cap \Gamma_{j+1} \mid j = 1, \dots, N_{\partial\Omega}\}$ and that these are actual corners of Ω, i.e., it is assumed that the inner angle ω_j between Γ_j and Γ_{j+1} is different from π (using cyclic notation, i.e., $\Gamma_{N_{\partial\Omega}+1} \equiv \Gamma_1$).

Remark. Broadly speaking, the domain Ω being Lipschitz means that the boundary $\partial\Omega$ can locally be seen as the graph of a Lipschitz function.

Furthermore, the boundary $\partial\Omega = \Gamma_{\mathrm{C}} \,\dot\cup\, \Gamma_{\mathrm{S}} \,\dot\cup\, \Gamma_{\mathrm{F}}$ is decomposed into the relatively open Γ_C, relatively closed Γ_S and into $\Gamma_F = \partial\Omega \setminus (\Gamma_C \cup \Gamma_S)$. To each of $\Gamma_C, \Gamma_S, \Gamma_F$ different boundary conditions are associated as motivated in their appearance in (SF).

A generic constant is denoted by C and standard notation $A \lesssim B$ hiding constants independent of h in inequalities $A \leq C\ B$ is used. Additionally $A \approx B$ is a shorthand notation for $A \lesssim B$ and $B \lesssim A$.

2.1 Triangulations

A triangulation $\mathcal{T}$ of the polygonal domain Ω is a set of triangles T that form an exact cover of Ω, i.e., the topological closure $\overline{\Omega}$ of Ω is given by the union

$$\overline{\Omega} = \bigcup_{T\in\mathcal{T}} T.$$

Denote the set of vertices by $\mathcal{V}$, edges by $\mathcal{E}$ and edge-midpoints by $\mathcal{M} := \{\mathrm{mid}\ E \mid E \in \mathcal{E}\}$. Only regular triangulations, cf. [Car19], will be considered, meaning that any two distinct triangles $T_1 \neq T_2 \in \mathcal{T}$ have either empty intersetion or have a vertex $T_1 \cap T_2 \in \mathcal{V}$ or an edge $T_1 \cap T_2 \in \mathcal{E}$ in common. The subset notation $\mathcal{S}(M) := \{S \in \mathcal{S} \mid S \subset \overline{M}\} \subseteq \mathcal{S}$ for $\mathcal{S} \in \{\mathcal{T}, \mathcal{E}, \mathcal{V}, \mathcal{M}\}$ is frequently used to refer to boundary edges $\mathcal{E}(\partial\Omega)$, boundary vertices $\mathcal{V}(\partial\Omega)$ or vertices $\mathcal{V}(T)$ of a triangle T among others. For a vertex $z \in \mathcal{V}$ (resp. edge $E \in \mathcal{E}$), also the vertex patch $\omega(z) := \mathrm{int}\big(\bigcup_{T\cap\{z\}\neq\emptyset} T\big)$ (resp. edge patch $\omega(E) := \mathrm{int}\big(\bigcup_{T\cap E\neq\emptyset} T\big)$) is defined and for a triangle $T \in \mathcal{T}$, the neighbour patch is given by $\Omega(T) := \bigcup_{z\in T} \omega(z)$.

Every triangle $T \in \mathcal{T}$ has an outer unit normal vector denoted by ν_T and for each edge $E \in \mathcal{E}$ a unit normal vector ν_E and unit tangent vector τ_E is fixed. The diameter of $T \in \mathcal{T}$ (resp. $E \in \mathcal{E}$), denoted by h_T (resp. h_E), defines a piecewise function $h_\mathcal{T}$ with $h_{\mathcal{T}|T} := h_T$ on Ω (resp. $h_\mathcal{E}$ with $h_{\mathcal{E}|E} := h_E$ on $\mathcal{E}$).

Under a uniform refinement of $\mathcal{T}$, a new triangulation $\widehat{\mathcal{T}}$ is understood that is obtained via red-refinement of the triangles $T \in \mathcal{T}$ (subdividing T into $T_1, \dots, T_4$ by connecting the vertices $z \in \mathcal{V}(T)$ with the edge-midpoints $\mathcal{M}(T)$ of the neighbouring edges). The newest vertex bisection (NVB) method allows for partial refinement of a triangulation $\mathcal{T}$ and the set of all admissible (obtained by NVB) refinements of $\mathcal{T}$ is denoted by $\mathbb{T}(\mathcal{T})$, see [Car19] for details.

2.2 Function spaces and derivatives

Consider an open subset $\omega \in \mathbb{R}^d$ for $d = 1, 2$. Let $C^k(\omega)$ denote the space of continuous functions with first k continuous derivatives, defined on ω. The space $C^k(\overline{\omega})$ denotes functions $f \in C^k(\omega)$ with first k derivatives that can be continuously extended up to the boundary $\partial\omega$ of ω. Then, $C^\infty(\omega) := \bigcup_{k\in\mathbb{N}} C^k(\omega)$ denotes the infinitely differentiable functions whereas $C_0^k(\omega) := \{f \in C^k(\omega) \mid \mathrm{supp} f \text{ compact}\}$ denotes the compactly supported $k \in \mathbb{N}_0 \cup \{\infty\}$-times differentiable functions. Standard notation for Lebesgue, $L^2(\omega)$, and Sobolev spaces, $H^k(\omega)$ for $k \in \mathbb{N}$, on ω applies and the reader is referred to standard literature [Gri85; Eva10]. The (weak) directional derivative of a (weakly) differentiable function $f : \omega \to \mathbb{R}$ in direction $v \in \mathbb{R}^d$ is denoted by $\partial_v f$ and $\partial^k_{v_1 v_2 \dots v_k} := \partial_{v_1}\partial_{v_2}\dots\partial_{v_k}$ abbreviates the k-th (weak) directional derivative in directions $v_1, v_2, \dots, v_k \in \mathbb{R}^d$. Using multi-index notation $a = (\alpha_1, \dots, \alpha_d) \in \mathbb{N}_0^d$ and $|\alpha| = \alpha_1 + \dots + \alpha_d$, the notation

$$D^\alpha f := \partial^{|\alpha|}_{x^\alpha} f := \frac{\partial^{|\alpha|} f}{\partial x^{\alpha_1} \dots \partial x^{\alpha_d}}$$

stands for the α-th (weak) partial derivative. Frequent abbreviations concern the gradient $\nabla := (\partial_{x_i})_{i=1,2}$, Hessian $D^2 := (\partial^2_{x_i x_j})_{i,j=1,2}$ and Laplacian $\Delta := \partial^2_{x_1 x_1} + \partial^2_{x_2 x_2}$.

Sobolev-Slobodeckij spaces, also known as fractional order Sobolev spaces, $H^{k+s}(w), k \in \mathbb{N}_0, s \in [0, 1)$, appear in the analysis at some point. For simplicity and the reduction of notational overhead, the notation is extended to include closed subsets $M \subset \mathbb{R}^d$ in that $H^{k+s}(M) := H^{k+s}(\mathrm{int}(M))$ for $k \in \mathbb{N}_0, s \in [0, 1)$. On any relatively open (or closed) line-segment γ in $\mathbb{R}^2$, Lebesgue and Sobolev spaces are defined by isometric identification of $\gamma \equiv \omega$ with an open (or closed) subset $\omega \subseteq \mathbb{R}^1$. In order to relate the fractional Sobolev spaces with the easier integer Sobolev spaces, the following Gagliardo-Nierenberg inequality is recalled. While this inequality holds for general Lipschitz domains, it will only be applied to line segments thereafter.

Theorem 2.1 ([BM18, Theorem 1])**.** *Let D be the $\mathbb{R}^d$, a half space or a Lipschitz domain and $0 \le s_1 \le s \le s_2, \theta \in (0, 1)$ be given. Then for all $f \in H^{s_2}(D)$,*

$$\|f\|_{H^s(D)} \lesssim \|f\|^{1-\theta}_{H^{s_1}(D)} \|f\|^{\theta}_{H^{s_2}(D)}.$$

For $\omega = T \in \mathcal{T}$ or any union of $T \in \mathcal{T}$, the L^2-scalar product is written as $(\bullet, \bullet)_{L^2(\omega)}$ and in case $\omega = E \in \mathcal{E}$ or any union of edges the notation $\langle \bullet, \bullet \rangle_{L^2(\omega)}$ is chosen. Let now γ denote a section $\gamma \in \{\Gamma_C, \Gamma_S, \Gamma_F, \partial\Omega\}$ of the boundary $\partial\Omega$ and define the spaces

$$\begin{aligned} H^1_\gamma(\Omega) &:= \{v \in H^1(\Omega) \mid v = 0 \text{ on } \gamma\} \\ H^2_\gamma(\Omega) &:= \{v \in H^2(\Omega) \cap H^1_\gamma \mid \partial_\nu v = 0 \text{ on } \gamma\}. \end{aligned}$$

The spaces $H^{-k}(\Omega) := (H^k_0(\Omega))'$ are defined as dual spaces of $H^k_0(\Omega)$, matching $H^k_0(\Omega) \equiv H^k_{\partial\Omega}(\Omega)$ for $k = 1, 2$ in this definition. Abbreviate $L^2(\mathcal{S}) := \prod_{E\in\mathcal{S}} L^2(E)$ for any subset of edges $\mathcal{S} \subseteq \mathcal{E}$.

The spaces of polynomials up to degree $k \in \mathbb{N}_0$ play an important role and the spaces of broken polynomials w.r.t. a triangulation $\mathcal{T}$ or edges $\mathcal{E}$ are defined in an analogous way,

$$\begin{aligned} P_k(\omega) &:= \{p \in L^2(\omega) \mid p \text{ is a polynomial of degree} \le k\}, \\ P_k(\mathcal{T}) &:= \left\{p \in L^2(\Omega) \;\middle|\; \forall T \in \mathcal{T}\ p_{|T} \in P_k(T)\right\}, \\ P_k(\mathcal{S}) &:= \left\{p \in L^2(\mathcal{S}) \;\middle|\; \forall E \in \mathcal{S}\ p_{|E} \in P_k(E)\right\}. \qquad (\mathcal{S} \subseteq \mathcal{E}) \end{aligned}$$

The piecewise L^2-projection onto piecewise polynomials on the edges $\mathcal{S} \subseteq \mathcal{E}$ or triangles $\mathcal{S} \subseteq \mathcal{T}$ is denoted $\Pi_{\mathcal{S},k} : L^2(\mathcal{S}) \to P_k(\mathcal{S})$ for $k \in \mathbb{N}_0$ and maps $L^2(\mathcal{S})$ onto $P_k(\mathcal{S})$.

2.3 Results on traces of Sobolev functions

Trace theorems are used to characterize the trace spaces of Sobolev functions and usually allow for a bounded right-inverse, which is then referred to as an extension operator. A general result [Gri85, Theorem 1.6.1.5] for traces on polygonal domains is recalled. While the full trace space is characterised in Grisvard, it is suffices to consider a subspace which can be characterized without introduction of extra notation.

Definition 2.1 (Trace space)**.** Consider the unit tangent vector τ_j on the segments $\Gamma_j = \operatorname{conv}\{P_{j-1}, P_j\} \in \mathcal{E}_\Omega$ of the polygonal boundary $\partial\Omega$ with corners $P_j \in \mathcal{V}_\Omega$. The subspace Tr of

$$\prod_{j=1}^{N_{\partial\Omega}} (C^1(\Gamma_j) \times C^0(\Gamma_j)),$$

defined by the compatibility conditions for $j = 1, \dots, N_{\partial\Omega}$, using cyclic notation,

$$\begin{aligned} \varphi_j(P_j) &= \varphi_{j+1}(P_j) \\ \begin{pmatrix} \partial_{\tau_j}\varphi_j(P_j) \\ \psi_j(P_j) \end{pmatrix} &= \begin{pmatrix} -\cos(\omega_j) & \sin(\omega_j) \\ -\sin(\omega_j) & -\cos(\omega_j) \end{pmatrix} \begin{pmatrix} \partial_{\tau_{j+1}}\varphi_{j+1}(P_j) \\ \psi_{j+1}(P_j) \end{pmatrix} \end{aligned} \tag{2}$$

for $(\varphi_j, \psi_j)_j \in \prod_{j=1}^{N_{\partial\Omega}} (C^1(\Gamma_j) \times C^0(\Gamma_j))$, is called *trace space.*

Theorem 2.2 ([Gri85, Theorem 1.6.1.5])**.** *Let Ω be a polygon with edges $\Gamma_1, \dots, \Gamma_{N_{\partial\Omega}}$. Then the trace map $\gamma_1 : H^2(\Omega) \to \prod_{j=1}^{N_{\partial\Omega}} (H^{3/2}(\Gamma_j) \times H^{1/2}(\Gamma_j))$*

$$\gamma_1 : v \mapsto (v_{|\Gamma_j}, (\partial_\nu v)_{|\Gamma_j})_j$$

maps $H^2(\Omega)$ onto a subspace $\gamma_1(H^2(\Omega)) \subseteq \prod_{j=1}^{N_{\partial\Omega}} (H^{3/2}(\Gamma_j) \times H^{1/2}(\Gamma_j))$ that contains Tr.

Remark. Note that the necessity of the conditions (2) comes to no surprise since any $v \in H^{2+\varepsilon}(\Omega), \varepsilon > 0$ allows for point evaluation of v and ∇v at the vertices P_j, which results in (2) for the image of v under the trace map. The full image space $\gamma_1(H^2(\Omega))$ can be characterised by some generalization of (2) to functions in

$$\prod_{j=1}^{N_{\partial\Omega}} (H^{3/2}(\Gamma_j) \times H^{1/2}(\Gamma_j))$$

with less regularity as given in [Gri92].

Furthermore the null space or kernel of the trace map γ_1 is known. It is clear that $H_0^2(\Omega) \subset \ker(\gamma_1)$ and also the converse holds.

Theorem 2.3 ([Gri92, Theorem 1.4.7])**.** *The kernel of the trace map γ_1 equals $H_0^2(\Omega) = \ker(\gamma_1)$.*

This shows that a bounded right-inverse $E : Tr \to H^2(\Omega)$ exists since the direct sum $H^2(\Omega) = H_0^2(\Omega) \oplus (H^2(\Omega)/H_0^2(\Omega))$ induces an isomorphism $H^2(\Omega)/H_0^2(\Omega) \to \gamma_1(H^2(\Omega))$ through γ_1 due to the open mapping theorem [RS80].

Theorem 2.4. *There exists a constant $C_{\mathrm{E}} > 0$ only depending on the domain Ω and a right-inverse $E : \gamma_1(H^2(\Omega)) \supseteq Tr \to H^2(\Omega)$ to the trace map from theorem 2.2 such that for any $t = (\varphi_j, \psi_j)_j \in \gamma_1(H^2(\Omega))$, it holds that $(\gamma_1 \circ E)t = t$ and*

$$\|Et\|_{H^2(\Omega)} \leq C_{\mathrm{E}} \sum_{j=1}^{N_{\partial\Omega}} \left(\|\varphi_j\|_{H^{3/2}(\Gamma_j)} + \|\psi_j\|_{H^{1/2}(\Gamma_j)} \right).$$

3 Abstract multilevel theory

The key idea is that very often solutions to elliptic partial differential equations (PDEs) can be broken down into components of different resolution or frequency. While the resolution of high frequencies requires a correspondingly fine discretisation, the low frequencies can be well approximatied by coarse meshes. This was the motivation for the application of Fourier analysis to these problems, sometimes also called *local mode analysis* [Bra77]. A typical motivation is the observation that the eigenfunctions of the one-dimensional Laplacian on the unit interval are given by $\psi_k(x) = \sin(k\pi x), k \in \mathbb{N}$, and form a orthogonal basis. Hence, any solution u to the Laplace equation $-\Delta u = f$ has an eigenexpansion in terms of the ψ_k and hence of different frequencies. Based on this interpretation, multilevel approaches have been developed in order to efficiently solve corresponding problems.

The vast success of these multilevel methods and their extension to various linear and nonlinear algebraic problems has made them one of the prominent choices for an efficient solve step. For a wide range of model problems, they provide convergence rates that are bounded independently of the problem size. Generally speaking, at least for a wide class of elliptic problems, these methods are the go-to methods for fast solve. Some have as low of a time complexity as linear, $\mathcal{O}(N)$, in the number of unknowns N (or ndof) of the algebraic system. This has to be put in contrast to direct solvers that only allow for polynomial but not linear, $\mathcal{O}(N^\alpha), \alpha > 1$, bounds, cf., e.g., [GN88].

3.1 Multigrid method

Consider the sequence $(\mathcal{T}_\ell)_{\ell\in\mathbb{N}_0}$ of regular triangulations of the polygonal domain Ω with an associated discrete function space $V(\mathcal{T}_\ell)$ for each level ℓ. Weak formulations of elliptic boundary value problems on Ω over some space $V \subset L^2(\Omega)$ are frequently written in operator notation as $\mathbb{A}u = \mathbb{F}$ with some differential operator $\mathbb{A} : V \to V'$ from some Banach space V to its dual V' and right-hand side $\mathbb{F} \in V'$. In this section, a general, self-adjoint and uniformly elliptic differential operator $\mathbb{A}$ is considered, where $(\mathbb{A}\bullet, \bullet)_{L^2(\Omega)} = a(\bullet, \bullet)$ defines a scalar-product a such that (V, a) for $V \subset L^2(\Omega)$ is a Hilbert space. Note, since only the algebraic viewpoint is of interest, the treatment of inhomogeneous (essential) boundary data does not appear explicitly.

Multilevel methods involve different spaces, like the $V(\mathcal{T}_\ell)$ in this setting. The question is, how does the continuous problem define associated discrete problems on the spaces $V(\mathcal{T}_\ell)$. With the sole assumption that all the discrete spaces $V(\mathcal{T}_\ell)$ are nested, i.e.,

$$V(\mathcal{T}_0) \subseteq V(\mathcal{T}_1) \subseteq \cdots \subseteq V(\mathcal{T}_\ell) \subseteq \cdots \subseteq V,$$

the continuous operators naturally define discrete counterparts. Define the discrete operator $\mathbb{A}_\ell : V(\mathcal{T}_\ell) \to V(\mathcal{T}_\ell)$ on each level ℓ via

$$(\mathbb{A}_\ell v_\ell, w_\ell)_{L^2(\Omega)} = a(v_\ell, w_\ell) \qquad \text{for all } v_\ell, w_\ell \in V(\mathcal{T}_\ell).$$

Given a discrete right-hand side $\mathbb{F}_\ell \in V(\mathcal{T}_\ell)$, e.g., induced by $\mathbb{F}$ in a similar way by

$$(\mathbb{F}_\ell, v_\ell)_{L^2(\Omega)} = \mathbb{F}(v_\ell) \qquad \text{for all } v_\ell, w_\ell \in V(\mathcal{T}_\ell), \tag{3}$$

this naturally defines an analogue to the continuous problem, namely the discrete problem

$$\mathbb{A}_\ell u_\ell = \mathbb{F}_\ell \quad \text{ on } V(\mathcal{T}_\ell). \tag{4}$$

The main idea of the multigrid algorithm is drawn from the fact that the discrete problems all approximate the same continuous problem and are thus related to each other by a single governing differential equation. To formulate this relation, some further operators are introduced. Let $\mathbb{Q}_\ell : V \to V(\mathcal{T}_\ell)$ be the L^2-projection and $\mathbb{P}_\ell : V \to V(\mathcal{T}_\ell)$ the energy inner product projection. Recall the defining relation of these operators, namely that for all $v \in V, w_\ell \in V(\mathcal{T}_\ell)$,

$$(\mathbb{Q}_\ell v, w_\ell)_{L^2(\Omega)} = (v, w_\ell)_{L^2(\Omega)} \quad \text{and} \quad a(\mathbb{P}_\ell v, w_\ell) = a(v, w_\ell).$$

Indeed, with these operators, the important identity

$$\mathbb{Q}_{\ell-1}\mathbb{A}_\ell = \mathbb{A}_{\ell-1}\mathbb{P}_{\ell-1}$$

on $V(\mathcal{T}_\ell)$ relates the discrete problems, since for $v_\ell, w_\ell \in V(\mathcal{T}_\ell)$

$$(\mathbb{Q}_{\ell-1}\mathbb{A}_\ell v_\ell, w_\ell)_{L^2(\Omega)} = a(v_\ell, \mathbb{Q}_{\ell-1} w_\ell) = (\mathbb{A}_{\ell-1}\mathbb{P}_{\ell-1} v_\ell, w_\ell)_{L^2(\Omega)}.$$

The global space V might be finite-dimensional, $V = V(\mathcal{T}_L)$ and the goal of the rest of the section is to solve for u_L in (4) for some level $L \in \mathbb{N}$ from the algebraic viewpoint. Thus, fix some level $L \in \mathbb{N}$ and restrict attention to the discretisations up to L. Said idea in the multigrid algorithm is an embedding, i.e., decomposition, of the discretization space $V(\mathcal{T}_L)$ into a sum of subspaces $V(\mathcal{T}_\ell), 0 \le \ell \le L$, by considering auxiliary problems, (4), on these subspaces $V(\mathcal{T}_\ell)$. This way, multilevel methods like multigrid can be interpreted as additive or multiplicative Schwarz methods and are closely related to subspace correction methods, cf. [GO95; Xu92; Zha92].

As motivated in the introduction, the final ingredient is the choice of an (approximate) solver of the auxiliary problems at each level ℓ. On each level ℓ a *smoothing operator* $\mathbb{S}_\ell : V(\mathcal{T}_\ell) \to V(\mathcal{T}_\ell)$ has to be chosen such that the error components corresponding to certain frequencies are reduced. Typical choices for the smoothing operator are additive or multiplicative Schwarz operators (corresponding to the classical Jacobi or Gauß-Seidel methods) and derivatives thereof, see [BP92] and [Saa03] for the bigger picture. A special choice of smoothing operators that perfectly fits into the setting of AFEM is introduced in the following section 3.2. The classical V-cycle multigrid algorithm [Bra93] with r pre-

Algorithm 1 (V(r,s)-cycle): The classical recursive definition.

Input: $v_\ell \in V(\mathcal{T}_\ell), r, s \in \mathbb{N}$

if $\ell = 0$ **then**

Exact solve: Solve exact on the coarsest grid $\ell = 0$, i.e.,
$\mathbb{B}_0 v_0 = \mathbb{A}_0^{-1} v_0$

else

Pre-smoothing: Start with $w_0 \coloneqq 0 \in V(\mathcal{T}_\ell)$ and smooth r times with $\mathbb{S}_\ell$, i.e.,
$w_{j+1} \coloneqq w_j + \mathbb{S}_\ell(v_\ell - \mathbb{A}_\ell w_j)$ for $j = 1, \dots, r$

Coarse-grid correction: Correct w_r on the next coarser level $\ell - 1$:
$w_{r+1} \coloneqq w_r + \mathbb{B}_{\ell-1}\mathbb{Q}_{\ell-1}(v_\ell - \mathbb{A}_\ell w_r)$

Post-smoothing: Start with w_{r+1} and perform s smoothing steps with $\mathbb{S}_\ell^\top$, i.e.,
$w_{j+1} \coloneqq w_j + \mathbb{S}_\ell^\top(v_\ell - \mathbb{A}_\ell w_j)$ for $j = r+2, \dots, r+s+1$

Set $\mathbb{B}_\ell v_\ell \coloneqq w_{r+s+1}$.

end if

Output: $\mathbb{B}_\ell v_\ell$

smoothing and s post-smoothing steps, algorithm 1, defines an approximate inverse $\mathbb{B}_\ell$ of $\mathbb{A}_\ell$ consisting of 3 steps performed on each level $\ell \geq 1$. In the post-smoothing step, the L^2-adjoint $\mathbb{S}_\ell^\top$ of $\mathbb{S}_\ell$ enters and on the coarsest level $\mathbb{B}_0 = \mathbb{A}_0^{-1}$ corresponds to direct solve.

Finally, this gives rise to the classical multigrid iteration as in [Bra93].

Definition 3.1 (Multigrid iteration)**.** Let the *approximative inverse* $\mathbb{B}_\ell$ of $\mathbb{A}_\ell$ be defined by algorithm 1 for some $r, s \in \mathbb{N}$ and $0 \leq \ell \leq L$. Given an initial solution $u_L^0 \in V(\mathcal{T}_L)$ to $\mathbb{A}_L u_L = \mathbb{F}_L$, the multigrid algorithm computes the iterates

$$u_L^k = u_L^{k-1} - \mathbb{B}_L(\mathbb{A}_L u_L^{k-1} - \mathbb{F}_L).$$

It is easily verified, that $\mathbb{B}_\ell$ defined in algorithm 1 is in fact a linear operator on $V(\mathcal{T}_\ell)$. Moreover, the following recurrence relation holds true with the identity $\mathbb{I}$.

Lemma 3.1 (Recurrence relation)**.** *The iteration matrix $\mathbb{I} - \mathbb{B}_\ell\mathbb{A}_\ell$ permits the recursion*

$$\mathbb{I} - \mathbb{B}_\ell\mathbb{A}_\ell\mathbb{P}_\ell = (\mathbb{I} - \mathbb{S}_\ell^\top \mathbb{A}_\ell\mathbb{P}_\ell)^s(\mathbb{I} - \mathbb{B}_{\ell-1}\mathbb{A}_{\ell-1}\mathbb{P}_{\ell-1})(\mathbb{I} - \mathbb{S}_\ell\mathbb{A}_\ell\mathbb{P}_\ell)^r.$$

This relation is well-known and a proof for $r = s = 1$ in [Bra93, Section 3] can easily be generalized.

3.2 Multigrid with local smoothing

Consider for a moment the situation that the $V(\mathcal{T}_\ell)$ of dimension $\dim V(\mathcal{T}_\ell) = N_\ell < \infty$ are generated by some finite element method (FEM) with adaptive mesh-refinement. Then each $V(\mathcal{T}_\ell)$ has a nodal basis $\{\varphi_\ell^1, \dots \varphi_\ell^{N_\ell}\}$ where the basis functions have support $\operatorname{supp}\varphi_\ell^j$ in some element $T \in \mathcal{T}_\ell$, edge-patch $\omega_\ell(E)$ for $E \in \mathcal{E}_\ell$ or node-patch $\omega_\ell(z)$ for $z \in \mathcal{V}_\ell$. Typically, the adaptive refinement will not lead to quasi-uniform meshes and, hence, a large portion of the mesh may be unrefined coming from $\mathcal{T}_{\ell-1}$ to $\mathcal{T}_\ell$. Consequently, $\{\varphi_\ell^1, \dots \varphi_\ell^{N_\ell}\} \cap V(\mathcal{T}_{\ell-1})$ may not be empty and in fact quite large. The idea of multigrid with local smoothing [WC06] is to take this into account by considering a smoother $\mathbb{S}_\ell : V(\mathcal{T}_\ell) \to V(\mathcal{T}_\ell)$ that, in some sense, only considers contributions from the subspace

$$\tilde{V}(\mathcal{T}_\ell) \coloneqq \operatorname{span}\Big(\{\varphi_\ell^1, \dots \varphi_\ell^{N_\ell}\} \setminus V(\mathcal{T}_{\ell-1})\Big) \subseteq V(\mathcal{T}_\ell), \tag{5}$$

delegating the other contributions to be handled on the coarser levels altogether. This observation also holds for other bases and at least the hierarchical bases for finite elements, cf., e.g., [Zie+81; Mit92], should be mentioned as natural extension to non-nested spaces.

For the general setting, any basis $\mathcal{B}(\mathcal{T}_\ell) = \{\varphi_\ell^1, \dots \varphi_\ell^{N_\ell}\}$ for $V(\mathcal{T}_\ell)$ may be fixed on each level and the spaces $\tilde{V}(\mathcal{T}_\ell)$ are defined by (5) for $\ell \leq 1$ and $\tilde{V}(\mathcal{T}_0) = V(\mathcal{T}_0)$. Denote the basis functions from $\mathcal{B}(\mathcal{T}_\ell)$ that remain in $\tilde{V}(\mathcal{T}_\ell)$ by $\{\tilde{\varphi}_\ell^1, \dots, \tilde{\varphi}_\ell^{\tilde{N}_\ell}\} = \mathcal{B}(\mathcal{T}_\ell) \setminus V(\mathcal{T}_{\ell-1})$ enumerated from 1 to $\tilde{N}_\ell \coloneqq \dim \tilde{V}(\mathcal{T}_\ell)$. It is clear that they form a basis of $\tilde{V}(\mathcal{T}_\ell)$ themselves.

Each basis function $\tilde{\varphi}_j$ of $\tilde{V}(\mathcal{T}_\ell)$ gives rise to the local projection operator $\mathbb{P}_\ell^j : V \to \operatorname{span}\{\tilde{\varphi}_\ell^j\} \subset V(\mathcal{T}_\ell)$ defined by the relation for $v \in V$,

$$a(\mathbb{P}_\ell^j v, \tilde{\varphi}_\ell^j) = a(v, \tilde{\varphi}_\ell^j).$$

This enables the following definition of the local Gauß-Seidel relaxation operator.

Definition 3.2 (local Gauß-Seidel relaxation)**.** The smoothing operator $\mathbb{S}_\ell : V(\mathcal{T}_\ell) \to V(\mathcal{T}_\ell)$,

$$\mathbb{S}_\ell \coloneqq \left(\mathbb{I} - (\mathbb{I} - \mathbb{P}_\ell^1)\cdots(\mathbb{I} - \mathbb{P}_\ell^{\tilde{N}_\ell})\right)\mathbb{A}_\ell^{-1} \tag{6}$$

is called *local Gauß-Seidel relaxation operator.*

Since the order of the basis functions is of no importance, the following understanding $\prod_{k=1}^{\tilde{N}_\ell}(\mathbb{I} - \mathbb{P}_\ell^k) \coloneqq (\mathbb{I} - \mathbb{P}_\ell^1)\cdots(\mathbb{I} - \mathbb{P}_\ell^{\tilde{N}_\ell})$ is used in the remainder of this section. By this definition, an alternative characterisation of $\mathbb{S}_\ell$ by $\mathbb{I} - \mathbb{S}_\ell\mathbb{A}_\ell = \prod_{k=1}^{\tilde{N}_\ell}(\mathbb{I} - \mathbb{P}_\ell^k)$ in $V(\mathcal{T}_\ell)$ follows immediately and clarifies the variant naming as multiplicative Schwarz operator used interchangeably as in, e.g., [BP92]. Recalling the recurrence relation, lemma 3.1, one notices that, instead of $\mathbb{I} - \mathbb{S}_\ell\mathbb{A}_\ell$, the term $\mathbb{I} - \mathbb{S}_\ell\mathbb{A}_\ell\mathbb{P}_\ell$ shows up. The fact, that a similar equation holds is due to properties of the projections involved in the definition of $\mathbb{S}_\ell$. Since this result is rather important for the subsequent argumentation, it is stated below.

Lemma 3.2 (Multiplicative identity for $\mathbb{S}_\ell$)**.** *The local Gauß-Seidel operator from definition 3.2 obeys to the identity*

$$\mathbb{I} - \mathbb{S}_\ell\mathbb{A}_\ell\mathbb{P}_\ell = \prod_{k=1}^{\tilde{N}_\ell}(\mathbb{I} - \mathbb{P}_\ell^k)$$

on all of V. Similarly for the L^2-adjoint $\mathbb{S}_\ell^\top$, it holds that $\mathbb{S}_\ell^\top\mathbb{A}_\ell\mathbb{P}_\ell = (\mathbb{S}_\ell\mathbb{A}_\ell\mathbb{P}_\ell)^*$ where the * in $(\mathbb{S}_\ell\mathbb{A}_\ell\mathbb{P}_\ell)^*$ denotes the adjoint in the Hilbert space (V, a). Hence, $\mathbb{I} - \mathbb{S}_\ell^\top\mathbb{A}_\ell\mathbb{P}_\ell = (\mathbb{I} - \mathbb{S}_\ell\mathbb{A}_\ell\mathbb{P}_\ell)^*$.

Proof (sketch). First note for $V = V(\mathcal{T}_L)$, that by the above arguments, the identity holds on $V(\mathcal{T}_L)$ simply by the definition of $\mathbb{S}_L$ and the fact that $\mathbb{P}_L$ is the identity on $V(\mathcal{T}_L)$. Expanding the product of the $(\mathbb{I} - \mathbb{P}_\ell^k)$ repeatedly produces the alternative representation

$$\mathbb{I} - \prod_{k=1}^{\tilde{N}_\ell}(\mathbb{I} - \mathbb{P}_\ell^k) = \sum_{k=1}^{\tilde{N}_\ell}\prod_{j=1}^{k-1}(\mathbb{I} - \mathbb{P}_\ell^j)\mathbb{P}_\ell^k.$$

The key argument is that $\mathbb{P}_\ell^k\mathbb{P}_\ell = \mathbb{P}_\ell^k$ for any $k = 1, \ldots, \tilde{N}_\ell$ as easily verified by the projection properties of $\mathbb{P}_\ell^k$ and $\mathbb{P}_\ell$. Thus,

$$\mathbb{S}_\ell\mathbb{A}_\ell\mathbb{P}_\ell = \left(\mathbb{I} - \prod_{k=1}^{\tilde{N}_\ell}(\mathbb{I} - \mathbb{P}_\ell^k)\right)\mathbb{P}_\ell = \sum_{k=1}^{\tilde{N}_\ell}\prod_{j=1}^{k-1}(\mathbb{I} - \mathbb{P}_\ell^j)\mathbb{P}_\ell^k = \mathbb{I} - \prod_{k=1}^{\tilde{N}_\ell}(\mathbb{I} - \mathbb{P}_\ell^k)$$

is already the first claim after rearranging terms. The second statement is shown by observing that by definition for arbitrary $v, w \in V$,

$$\begin{aligned} a(\mathbb{S}_\ell^\top\mathbb{A}_\ell\mathbb{P}_\ell v, w) &= a(\mathbb{S}_\ell^\top\mathbb{A}_\ell\mathbb{P}_\ell v, \mathbb{P}_\ell w) \\ &= (\mathbb{S}_\ell^\top\mathbb{A}_\ell\mathbb{P}_\ell v, \mathbb{A}_\ell\mathbb{P}_\ell w)_{L^2(\Omega)} \\ &= (\mathbb{A}_\ell\mathbb{P}_\ell v, \mathbb{S}_\ell\mathbb{A}_\ell\mathbb{P}_\ell w)_{L^2(\Omega)} \\ &= a(v, \mathbb{S}_\ell\mathbb{A}_\ell\mathbb{P}_\ell w). \end{aligned}$$

□

3.3 Convergence analysis

Classical convergence results for multigrid methods were either obtained by Fourier methods and thus restricted to simple (rectangular) domains, e.g., [Bra77], or required sufficient amounts of smoothing steps, e.g., [BH83], also [Bra84] and references therein. Using the abstract operator description of the multigrid algorithm of this section, a general convergence result can be obtained for more general domains, even for the V(1, 1)-cycle with only one pre- and post-smoothing. For simplicity in the convergence analysis, consider the case of smoothing once, i.e., $\mathbb{B}_\ell$ given by the V(1, 1)-cycle from algorithm 1. Using the explicit representation for the local Gauß-Seidel smoother, the following result for the multigrid iteration operator is obtained.

Lemma 3.3. *Abbreviate*

$$\mathbb{R}_L \coloneqq (\mathbb{I} - \mathbb{P}_0) \prod_{\ell=1}^{L} \prod_{k=1}^{\tilde{N}_\ell} (\mathbb{I} - \mathbb{P}_\ell^k).$$

Then, for every level $L \in \mathbb{N}_0$, the compact representation

$$\mathbb{I} - \mathbb{B}_L \mathbb{A}_L = \mathbb{R}_L^* \mathbb{R}_L$$

holds. Moreover, $\mathbb{I} - \mathbb{B}_L \mathbb{A}_L$ is self-adjoint in the Hilbert space (V, a).

Proof. A recursive application of lemma 3.1 combined with the identities for the special choice of smoothing operator, lemma 3.2, leads to the statement. □

By the above lemma, the multigrid method itself can be interpreted as some multiplicative method that consist of repeated a-orthogonal projections. As such, it can be understood as a *method of subspace corrections* (MSC), cf. [Bra+91], in the subspaces $V(\mathcal{T}_0), \operatorname{span}\{\tilde{\varphi}_\ell^k\} \subseteq V(\mathcal{T}_L)$ for $\ell = 1, \dots, L$ and $k = 1, \dots, \tilde{N}_\ell$. Since

$$V(\mathcal{T}_0) + \sum_{\ell=1}^{L} \sum_{k=1}^{\tilde{N}_\ell} \operatorname{span}\{\tilde{\varphi}_\ell^k\} = V(\mathcal{T}_0) + \sum_{\ell=1}^{L} \tilde{V}(\mathcal{T}_\ell) = V(\mathcal{T}_L),$$

the assumptions of [XZ02] apply and the following convergence result follows. Recall the lexicographic order $>$ for tuples in $\mathbb{R}^2$, i.e., $(m, n) > (\ell, k)$ holds if either $m > l$ or both, $m = \ell$ and $n > k$.

Theorem 3.4 ([XZ02, Corollary 4.3])**.** *Let $L \in \mathbb{N}_0$ and define*

$$c_L \coloneqq \sup_{\|v_L\|_a = 1} \inf_{v_L = v_0 + \sum_{\ell=1}^{L} \sum_{k=1}^{\tilde{N}_\ell} v_\ell^k} \left(\|\mathbb{P}_0(v_L - v_0)\|_a^2 + \sum_{\ell=1}^{L} \sum_{i=1}^{\tilde{N}_\ell} \left\| \mathbb{P}_\ell^k \sum_{(m,n) > (\ell,k)} v_m^k \right\|_a^2 \right)$$

where the infimum runs over all decompositions of $v_L \in V(\mathcal{T}_L)$ of the form

$$v_L = v_0 + \sum_{\ell=1}^{L} \sum_{k=1}^{\tilde{N}_\ell} v_\ell^k$$

with $v_\ell^k \in \operatorname{span}\{\tilde{\varphi}_\ell^k\}$. Then, with $\mathbb{R}_L$ from lemma 3.3 and operator norm $\| \bullet \|_a$ on $V(\mathcal{T}_\ell)$,

$$\|\mathbb{I} - \mathbb{B}_L \mathbb{A}_L\|_a = \|\mathbb{R}_L\|_a^2 = \frac{c_L}{1 + c_L} < 1.$$

Remark. Note, that theorem 3.4 itself does not provide uniform convergence in L, i.e., it may happen that $c_L \to \infty$ as $L \to \infty$. Conversely, if it is possible to bound $c_L \leq c_\infty$ for some constant c_∞, this implies uniform convergence since the function $x \mapsto x/(1+x)$ is monotone for $x > 0$ and, hence, $\|\mathbb{I} - \mathbb{B}_L \mathbb{A}_L\|_a \leq c_\infty/(1+c_\infty) < 1$ uniformly in L.

Proof (sketch). Use that $\|\mathbb{I} - \mathbb{B}_L \mathbb{A}_L\|_a = \|\mathbb{R}_L^* \mathbb{R}_L\|_a = \|\mathbb{R}_L\|_a^2$ and apply [XZ02, Corollary 4.3] verbatim. □

3.4 Estimation of the algebraic error

Consider the exact solution u_ℓ on level $\ell \in \mathbb{N}_0$ to $\mathbb{A}_\ell u_\ell = \mathbb{F}_\ell$ and an arbitrary approximation $v_\ell \in V(\mathcal{T}_\ell)$. In order to define a suitable stopping criterion for an iterative method, it would be favourable if the exact algebraic error $e_\ell = u_\ell - v_\ell$ is known. Unfortunately, computing it exactly has the same complexity as the original problem, due to linearity, and is thus not feasible. Instead, it is possible to compute a quantity that behaves equivalent to the error and is computable in linear time. From hereon, $\mathbb{B}_\ell$ is considered given by the $\mathrm{V}(r,s)$-cycle, i.e., algorithm 1, for fixed $r, s \in \mathbb{N}$.

Definition 3.3. For an arbitrary $v_\ell \in V(\mathcal{T}_\ell), \ell \in \mathbb{N}_0$, the *algebraic residual* $\mathrm{Res}_\ell \in V(\mathcal{T}_\ell)$ reads

$$\mathrm{Res}_\ell(v_\ell) \coloneqq \mathbb{A}_\ell v_\ell - \mathbb{F}_\ell.$$

Define the *algebraic error estimator* η_{alg} on $\mathcal{T}_\ell$ by

$$\eta_{\mathrm{alg}}^2(\mathcal{T}_\ell, v_\ell) \coloneqq (\mathrm{Res}_\ell, \mathbb{B}_\ell \mathrm{Res}_\ell)_{L^2(\Omega)}.$$

It is well known that integration of piecewise polynomials can be done in linear time by applying quadrature rules and may not even be necessary in an actual implementation, see section 6.1. Furthermore, in contrast to the exact algebraic error, the algebraic residual is a known quantity as it appears in the formulation of the multigrid iteration. The important property of the mentioned algebraic error estimator is the equivalence to the a-norm of the error stated below.

Lemma 3.5. *Let $\ell \in \mathbb{N}_0$ and set $C \coloneqq \|\mathbb{I} - \mathbb{B}_\ell \mathbb{A}_\ell\|_a$. Also consider the exact solution u_ℓ of $\mathbb{A}_\ell u_\ell = \mathbb{F}_\ell$ and an arbitrary $v_\ell \in V(\mathcal{T}_\ell)$. Then*

$$(1-C)\,\|u_\ell - v_\ell\|_a^2 \leq \eta_{\mathrm{alg}}^2(\mathcal{T}_\ell, v_\ell) \leq (1+C)\,\|u_\ell - v_\ell\|_a^2$$

with the a-norm $\|\bullet\|_a^2 \coloneqq a(\bullet,\bullet)$.

This lemma allows for an immediate consequence if the approximation property of $\mathbb{B}_\ell$ is uniform in ℓ.

Corollary 3.6. *Suppose that $\mathbb{B}_\ell$ is a uniform approximative inverse of $\mathbb{A}_\ell$, i.e.,*

$$\|\mathbb{I} - \mathbb{B}_\ell \mathbb{A}_\ell\|_a \leq C < 1$$

uniformly in $\ell \in \mathbb{N}_0$. Given u_ℓ, v_ℓ as in lemma 3.5, then

$$\|u_\ell - v_\ell\|_a \approx \eta_{\mathrm{alg}}(\mathcal{T}_\ell, v_\ell)$$

holds with equivalence constant independent on the level $\ell \in \mathbb{N}_0$.

Proof (of lemma 3.5). Abbreviate $e_\ell = u_\ell - v_\ell$. Linearity of a implies

$$\|u_\ell - v_\ell\|_a^2 \coloneqq a(e_\ell, e_\ell) = a(e_\ell, (\mathbb{I} - \mathbb{B}_\ell \mathbb{A}_\ell)e_\ell) + a(e_\ell, \mathbb{B}_\ell \mathbb{A}_\ell e_\ell).$$

The claim follows by rearranging terms accordingly, followed by a Cauchy inequality and definition of the operator norm so that for the first term

$$\begin{aligned}\|e_\ell\|_a^2 &\leq \|(\mathbb{I} - \mathbb{B}_\ell \mathbb{A}_\ell)e_\ell\|_a \|e_\ell\|_a + a(e_\ell, \mathbb{B}_\ell \mathbb{A}_\ell e_\ell) \\ &\leq C \|e_\ell\|_a^2 + a(e_\ell, \mathbb{B}_\ell \mathbb{A}_\ell e_\ell).\end{aligned}$$

Similarly for the second term and this finishes the proof by observing $a(e_\ell, \mathbb{B}_\ell \mathbb{A}_\ell e_\ell) = \eta_{\text{alg}}(\mathcal{T}_\ell, v_\ell)$.

□

Having a measure for the algebraic error gives rise to a stopping criterion in terms of error reduction in the energy norm. Given $\varepsilon > 0$, the multigrid algorithm, algorithm 2, computes the multigrid iterates, cf. definition 3.1, until $\eta_{\text{alg}}^2(\mathcal{T}_L, u_L^k) < \varepsilon \eta_{\text{alg}}^2(\mathcal{T}_L, u_L^0)$. While the reduction of the algebraic error after k multigrid iterations can be estimated by

$$\left\|u_L - u_L^k\right\|_a = \left\|(\mathbb{I} - \mathbb{B}_L \mathbb{A}_L)^k (u_L - u_L^0)\right\|_a \leq \|\mathbb{I} - \mathbb{B}_L \mathbb{A}_L\|_a^k \left\|u_L - u_L^0\right\|_a, \tag{7}$$

using the definition of the operator norm, this stopping criterion leads to bounds on the error in terms of ε.

Algorithm 2 ($\text{MG}_{r,s}$): Multigrid algorithm

Input: $\mathbb{A}_L, \mathbb{F}_L$, initial solution u_L^0 and stopping criterion $\varepsilon > 0$

```
Compute B_L by the V(r,s)-cycle, algorithm 1
k := 0
while η²_alg(T_L, u_L^k) ≥ ε η²_alg(T_L, u_L^0) do
    u_L^{k+1} := u_L^k − B_L(A_L u_L^k − F_L).
    k := k + 1
end while
u_L^ε := u_L^k
```

Output: u_L^ε

Lemma 3.7. *Set $\|\mathbb{I} - \mathbb{B}_L \mathbb{A}_L\|_a =: C < 1$. Let u_L be the solution to (4) and $u_L^\varepsilon = u_L^k$ the approximation obtained by the multigrid methdod, algorithm 2, for initial solution $u_L^0 \neq u_L$ and $\varepsilon > 0$ after k iterations. Then it holds that*

$$\|u_L - u_L^\varepsilon\|_a^2 < \frac{1+C}{1-C} \varepsilon \left\|u_L - u_L^0\right\|_a^2$$

and the number of iterations k is bounded from above by $\log_C((1-C)/(1+C)\varepsilon)/2$.

Proof. Using lemma 3.5 and the stopping criterion yields

$$\|u_L - u_L^\varepsilon\|_a^2 \leq \frac{1}{1-C} \eta_{\text{alg}}^2(\mathcal{T}_L, u_L^\varepsilon) < \frac{1}{1-C} \varepsilon \eta_{\text{alg}}^2(\mathcal{T}_L, u_L^0) \leq \frac{1+C}{1-C} \varepsilon \left\|u_L - u_L^0\right\|_a^2.$$

This shows the first claim. On the other hand, if $C^{2k} < (1-C)/(1+C)\varepsilon$, the stopping criterion is fulfilled and therefore bounds k, since with (7),

$$\eta_{\text{alg}}^2(\mathcal{T}_L, u_L^k) \leq (1+C)\left\|u_L - u_L^k\right\|_a^2 < (1-C)\varepsilon \left\|u_L - u_L^0\right\|_a^2 \leq \varepsilon \eta_{\text{alg}}^2(\mathcal{T}_L, u_L^0).$$

□

Hence, the algebraic error depends on the initial solution for the multigrid algorithm. Thus, it is only natural to perform nested iterations, e.g., use the previous approximation $u_{\ell-1}^{\varepsilon}$ as initial solution u_{ℓ}^{0} on a newer, finer triangulation $\mathcal{T}_{\ell}$. This is especially common, when $\mathbb{F}_{\ell}$ is given by (3). In the context of multigrid, the resulting scheme is frequently called *full multigrid*. For simplicity, the notation $u_{-1}^{\varepsilon} := 0$ is understood.

4 Argyris FEM for the biharmonic equation

The Argyris finite element [AFS68] is well-known in the context of the conforming plate finite element method (FEM) and has been studied in [Bra07; BS08] and others. It is possibly the most widely used triangular finite element of class C^1. In contrast to the Bogner-Fox-Schmidt rectangular element [BFS66], which is restricted to well-oriented rectangular grids, the Argyris element can be applied with any polygonal domain. While the classical Argyris finite element has 21 degrees of freedom that define the space of quintic (fifth order) polynomials over a triangle, generalisations to higher order (such as the 7-th order Argyris element) are known but nor further discussed here.

4.1 The Argyris element

In this work, both the classical Argyris FEM as well as the hierarchical Argyris FEM, originally proposed in [CH21], will be covered. Based on the space of quintic polynomials

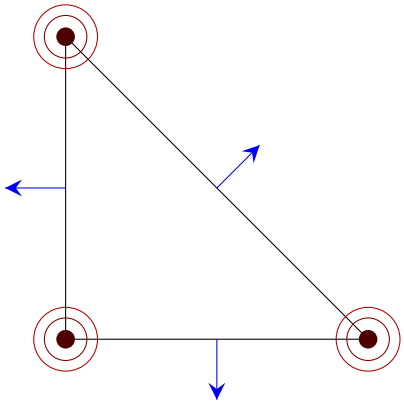

Figure 2: The usual graphic representation of the reference quintic Argyris finite element

$P_5(T)$ on a triangle, the Argyris finite element $(T, P_5(T), \{\Lambda_1, \dots, \Lambda_{21}\})$, in the sense of Ciarlet, is defined in terms of its 21 (local) degrees of freedom (dofs). These degrees of freedom $\{\Lambda_1, \dots, \Lambda_{21}\}$ are the evaluation functionals of partial derivatives of order $0, 1, 2$ at each vertex $z \in \mathcal{V}(T)$ and the normal derivative evaluation at the edge midpoint mid E of each edge $E \in \mathcal{E}(T)$. Those dofs are typically depicted with points (point evaluation), small circles (gradient evaluation) and large circles (hessian evaluation) at the vertices and an arrow in the normal direction (normal derivative evaluation) at the edge midpoints, see figure 2. Then, on the global level, these local degrees of freedom are identified as the same global dof (resticted to an element) if they are associated with the same node and represent the same derivative. Thus, consider now some admissible triangulation $\mathcal{T}$ with vertices $\mathcal{V}$, edges $\mathcal{E}$ and edge midpoints $\mathcal{M}$. The term node is used for an entity with an associated degree of freedom. In this case, the set of nodes $\mathcal{N} = \mathcal{V} \cup \mathcal{M}$ consists of the vertices $\mathcal{V}$ and edge midpoints $\mathcal{M}$. Building on the Argyris finite elements defined for each $T \in \mathcal{T}$, the classical Argyris function space $\mathfrak{A}(\mathcal{T})$ consists of functions $v_h \in P_5(\mathcal{T})$ that have continuous derivatives of orders $0, 1, 2$ at the vertices $z \in \mathcal{V}$ and continuous normal derivatives at the edge midpoints $\mathcal{M}$.

Definition 4.1 (classical Argyris space)**.** Consider an admissible triangulation $\mathcal{T}$ with

vertices $\mathcal{V}$ and edge midpoints $\mathcal{M}$. The *classical Argyris space* $\mathfrak{A}(\mathcal{T})$ over $\mathcal{T}$ is defined as

$$\mathfrak{A}(\mathcal{T}) \coloneqq \left\{ v_h \in P_5(\mathcal{T}) \;\middle|\; \begin{array}{rl} v_h, \nabla v_h, D^2 v_h & \text{are continuous at any } z \in \mathcal{V} \\ \partial_\nu v_h & \text{is continuous at any } z \in \mathcal{M} \end{array} \right\}.$$

Hence, the classical Argyris space has 6 (global) degrees of freedom associated with each vertex $z \in \mathcal{V}$, e.g., $\{\delta_z, \partial_x\delta_z, \partial_y\delta_z, \partial^2_{xx}\delta_z, \partial^2_{xy}\delta_z, \partial^2_{yy}\delta_z\}$ using standard Cartesian coordinates (x, y), and additionally one for the normal derivative, $\partial_{\nu_E}\delta_z$, at the edge midpoints $z = \operatorname{mid} E \in \mathcal{M}$. Here, δ_z denotes the Dirac functional (point evaluation) at point z.

A standard argument shows unisolvence and that these degrees of freedom uniquely determine the normal derivative of $v_h \in \mathfrak{A}(\mathcal{T})$ along the inner edges and, thus, lead to C^1 continuity of $\mathfrak{A}(\mathcal{T})$, cf. [BS08]. In other words, $\mathfrak{A}(\mathcal{T})$ is a subspace of the C^1 continuous piecewise quintic polynomials $P_5(\mathcal{T}) \cap C^1(\mathcal{T})$ and alternatively characterized as

$$\mathfrak{A}(\mathcal{T}) \coloneqq \{v_h \in P_5(\mathcal{T}) \cap C^1(\Omega) \mid D^2 v_h \text{ is continuous at } z \in \mathcal{V}\}. \tag{8}$$

4.2 The hierarchical Argyris finite element space

The question whether $\mathfrak{A}(\mathcal{T}) = P_5(\mathcal{T}) \cap C^1(\mathcal{T})$ holds has to be denied, except for few pathological cases. In fact, a notable property of the classical Argyris space lies in its non-nestedness with respect to mesh refinement. Consider some refinement $\widehat{\mathcal{T}}$ of $\mathcal{T}$, where some edge $E \in \mathcal{E}$ has been bisected and, thus, mid E is a new vertex in $\widehat{\mathcal{T}}$ together with the corresponding spaces $\mathfrak{A}(\mathcal{T})$ and $\mathfrak{A}(\widehat{\mathcal{T}})$. By construction of the (classical) Argyris space, any function $\widehat{v}_h \in \mathfrak{A}(\widehat{\mathcal{T}})$ has continuous second-order normal-normal derivative $\partial^2_{\nu_E\nu_E}\widehat{v}_h$ at mid E. Since for $\mathfrak{A}(\mathcal{T})$ by (8), only C^1-continuity along the interior of E holds, $\mathfrak{A}(\mathcal{T}) \not\subset \mathfrak{A}(\widehat{\mathcal{T}})$. Conversely, one finds that the remainder $\mathfrak{A}(\mathcal{T}) \setminus \mathfrak{A}(\widehat{\mathcal{T}})$ only consists of those functions $v_h \in \mathfrak{A}(\mathcal{T})$ that have discontinuous normal-normal derivative at such refined edge midpoints.

This observation leads to the design of the *extended Argyris space* $\mathscr{A}(\mathcal{T})$, originally proposed in [CH21]. The authors make use of a minimal extension to the finite element space that results in a hierarchical sequence of spaces, remedying the problematic that arises due to the non-nestedness of $\mathfrak{A}(\mathcal{T})$. The resulting *extended Argyris space* $\mathscr{A}(\mathcal{T})$ for an admissible mesh sequence generating $\mathcal{T} = \mathcal{T}_L$ $(\mathcal{T}_\ell)_{\ell=0}^L$, so that $\mathcal{T}_L = \mathcal{T}$ with initial triangulation $\mathcal{T}_0$, is a minimal extension of $\mathfrak{A}(\mathcal{T})$, in that the spaces coincide on the coarsest triangulation, i.e., $\mathscr{A}(\mathcal{T}_0) = \mathfrak{A}(\mathcal{T}_0)$, and given the sequence of successive refined meshes, $\mathscr{A}(\mathcal{T}) = \sum_{\ell=0}^L \mathfrak{A}(\mathcal{T}_\ell)$. The last representation already shows nestedness for the extended Argyris finite element space $\mathscr{A}(\mathcal{T})$.

An obvious consequence is that the resulting function spaces not only depend on the topology and geometry of the current mesh but also, to a certain extend that will be clarified later, on preceding triangulations. Although the design is tailored for the Argyris FEM, it may very well be adapted for other finite elements in the sense of Ciarlet.

The exact details in the construction shall not be of much interest here, the reader may find all the details in [CH21]. But, for sake of completeness, the important aspects will be recalled hereafter. In the design of the extended Argyris space the only modification is carried out to the global degrees of freedom. Neither the Argyris finite element, the shape functions on each triangle nor the local degrees of freedom will be modified. That is, except for fixing directions for the partial derivatives that make up the local degrees of freedom. The key lies in the consideration of adapted coordinate systems $\{\tau(z), \nu(z)\}$ at the vertices $z \in \mathcal{V}$ that are used to define the direction of differentiation for the degrees of

freedom (instead of partial derivatives in x and y direction as typically done). To make this more precise and simultaneously include the treatment of boundary conditions, which will be discussed next, the following definition summarises the requirements on the local coordinates $\{\tau(z), \nu(z)\}$. As pointed out before, the construction not only relies on a particular triangulation $\mathcal{T} = \mathcal{T}_L$ but also on the full sequence $(\mathcal{T}_\ell)_{\ell=0}^L$ starting with some initial triangulation $\mathcal{T}_0$ with vertices $\mathcal{V}_0$.

Definition 4.2 (Local vertex coordinates)**.** Let $\mathcal{T} = \mathcal{T}_L$, generated by the sequence $(\mathcal{T}_\ell)_{\ell=0}^L$ with initial triangulation $\mathcal{T}_0$ with vertices $\mathcal{V}_0$, be given. Define the local vertex coordinates $\{\tau(z), \nu(z)\}$ for $z \in \mathcal{V}$ with respect to the properties of z in the history of triangulations.
For $z \in \mathcal{V}(\Omega) \setminus \mathcal{V}_0$: Set $\tau(z) = \tau_E$ and $\nu(z) = \nu_E$ for an edge E with mid $E = z$ of some previous triangulation $\mathcal{T}_\ell, 0 \leq \ell < L$.
For $z \in \mathcal{V}_\Omega \cap \overline{(\Gamma_C \cup \Gamma_S)}$: Let $E_1, E_2 \in \mathcal{E}(\partial\Omega)$ be the (distinct and) unique edges on the boundary with $E_1 \cap E_2 = \{z\}$. If $z \in \overline{\Gamma_C}$, choose any edge $E \in \{E_1, E_2\} \cap \overline{\Gamma_C}$ of them that is part of the clamped boundary and set $\tau(z) = \tau_E, \nu(z) = \nu_E$. Otherwise $z \in \overline{\Gamma_S} \setminus \overline{\Gamma_C}$ and choose the coordinates to be both tangentials, i.e., $\tau(z) = \tau_{E_1}, \nu(z) = \tau_{E_2}$.
For the remaining cases, i.e., $z \in \mathcal{V}_0(\Omega) \cup \mathcal{V}_\Omega \setminus \overline{(\Gamma_C \cup \Gamma_S)}$: $\tau(z) = e_1$ and $\nu(z) = e_2$ are the standard Cartesian coordinates.

Now, one can write the degrees of freedom of the classical Argyris function space $\mathfrak{A}(\mathcal{T})$ associated with a vertex $z \in \mathcal{V}$ in terms of the local vertex coordinates, instead of Cartesian coordinates as before. These read $\delta_z, \partial_{\tau(z)}\delta_z, \partial_{\nu(z)}\delta_z, \partial^2_{\tau(z)\tau(z)}\delta_z, \partial^2_{\tau(z)\nu(z)}\delta_z, \partial^2_{\nu(z)\nu(z)}\delta_z$ for any $z \in \mathcal{V}$.

Recall the aforementioned observation that pinpoints the failure of nestedness to the continuity requirement of the second-order derivatives at new interior vertices $z \in \mathcal{V}(\Omega) \setminus \mathcal{V}_0$ across some historical edge $E \in \mathcal{E}_\ell \setminus \mathcal{E}$ with $z = \operatorname{mid} E$, where $\mathcal{E}_\ell$ denotes the edges of some previous triangulation $\mathcal{T}_\ell$ with $0 \leq \ell < L$. In order to drop this continuity requirement, the corresponding degree of freedom $\partial^2_{\nu(z)\nu(z)}\delta_z$ is split into two, namely $\partial^2_{\nu(z)\nu(z)}\delta_z^\pm$ that evaluate a function only on one side of E. In other words, the support of $\partial^2_{\nu(z)\nu(z)}\delta_z^\pm$ is one of the two halfspaces

$$H_\pm(z) \coloneqq \{x \in \mathbb{R}^2 \mid \pm (x - z) \cdot \nu(z) \geq 0\}$$

generated by E. Note also in definition 4.2, that the choice of such a historical edge E with $z = \operatorname{mid} E$ may not be unique but the generated halfspaces are, since in fact only $\{\tau_E, \nu_E\}$ are of importance.

With any node $z \in \mathcal{N}$ there are $m(z) \in \{1, 6, 7\}$ degrees of freedom $\Lambda_{z,1}, \ldots, \Lambda_{z,m(z)}$ associated. From here on the dependence on the refinement history will be suppressed as it remains only as a technical detail.

Definition 4.3 (Global degrees of freedom)**.** Consider a triangulation $\mathcal{T}$ and the local vertex coordinates from definition 4.2. The degrees of freedom of the extended Argyris space are the set

$$\{\Lambda_{z,j} \mid j = 1, \ldots, m(z), z \in \mathcal{N}\}$$

of functionals, defined by the following. Every vertex $z \in \mathcal{V}$ is associated with the $m(z) = 6$ or 7 degrees of freedom $\Lambda_{z,1}, \ldots, \Lambda_{z,m(z)}$. Namely, $\delta_z, \partial_{\tau(z)}\delta_z, \partial_{\nu(z)}\delta_z, \partial^2_{\tau(z)\tau(z)}\delta_z, \partial^2_{\tau(z)\nu(z)}\delta_z$, enumerated in this order as $\Lambda_{z,1}, \ldots, \Lambda_{z,5}$, and $\Lambda_{z,6} = \partial^2_{\nu(z)\nu(z)}\delta_z$ if $z \in \mathcal{V}(\partial\Omega) \cup \mathcal{V}_0$ or $\Lambda_{z,6} = \partial^2_{\nu(z)\nu(z)}\delta_z^+, \Lambda_{z,7} = \partial^2_{\nu(z)\nu(z)}\delta_z^-$ otherwise (that is $z \in \mathcal{V}(\Omega) \setminus \mathcal{V}_0$). To each edge $E \in \mathcal{E}$ with $z = \operatorname{mid} E$ the single, $m(z) = 1$, degree of freedom $\Lambda_{z,1} = \partial_{\nu_E}\delta_z$ is associated.

These global degrees of freedom define the extended Argyris space in the usual fashion.

Definition 4.4 (Extended Argyris space and nodal basis)**.** Given the N global degrees of freedom from definition 4.3 of the extended Argyris space, enumerated as $\Lambda_1, \dots, \Lambda_N$. Then the *nodal basis* $\{\varphi_1, \dots, \varphi_N\} \in P_5(\mathcal{T})$ is defined by the duality relation

$$\Lambda_i(\varphi_j) = \delta_{ij} \qquad \text{for all } i, j = 1, \dots, N$$

in $P_5(\mathcal{T})$ using the Kronecker delta δ_{ij}.

Standard arguments show that the nodal basis is in fact a basis in the classical sense.

Definition 4.5 (Extended Argyris space)**.** The *extended Argyris space* $\mathscr{A}(\mathcal{T})$ over $\mathcal{T}$ is defined by

$$\mathscr{A}(\mathcal{T}) \coloneqq \operatorname{span}\{\varphi_1, \dots, \varphi_N\}.$$

Given the global degrees of freedom there is a natural interpolation operator defined for functions that can be evaluated at the degrees of freedom.

Definition 4.6 (Interpolation operator)**.** Let $\Lambda_1, \dots, \Lambda_N$ denote the degrees of freedom of $\mathscr{A}(\mathcal{T})$ for a given triangulation $\mathcal{T}$ and let $\{\varphi_1, \dots, \varphi_N\}$ be the associated nodal basis. Then the *interpolation operator* $\mathcal{I} : C^2(\overline{\Omega}) \to \mathscr{A}(\mathcal{T})$ is defined by

$$\mathcal{I}v = \sum_{j=1}^{N} \Lambda_j(v)\varphi_j \in \mathscr{A}(\mathcal{T})$$

for all $v \in C^2(\overline{\Omega})$. The $\{\Lambda_j(v)\}_j$ are called the *coefficients* of $\mathcal{I}v$ with respect to the nodal basis.

Remark. For any $\varepsilon > 0$ the Sobolev embedding [Gri85] shows $H^{3+\varepsilon}(\Omega) \subset C^2(\overline{\Omega})$ by the natural embedding and hence, the interpolation operator is defined on $H^{3+\varepsilon}(\Omega)$ as well. By this definition, it is clear that $\mathcal{I}v_h = v_h$ for any $v_h \in \mathscr{A}(\mathcal{T})$.

Also this definition of the interpolation matches the global interpolant from [BS08] and while the Argyris finite element is not *affine-interpolation equivalent* it is *almost-affine* and allows the approximation result from [Cia02, Section 6] given in the global form for the extended Argyris space.

Theorem 4.1. *There exists a constant C only depending on the triangulation $\mathcal{T}$ such that for the Argyris interpolant from definition 4.6 the following approximation holds. For any $0 \le m \le 6$ and $v \in H^6(\Omega)$ it holds that*

$$|v - \mathcal{I}v|_{H^m(\Omega)} \le C|h_{\mathcal{T}}^{6-m} v|_{H^6(\Omega)}.$$

4.3 Boundary data

Before stating the hierarchical Argyris FEM for the plate problem, techniques for the incorporation of boundary conditions have to be discussed. Recall that the essential boundary data in the strong form of the plate problem (SF) is given by g_{CS} and g_{C}. The problem of approximating these boundary conditions in the (extended) Argyris space boils down to finding a good approximation $g_h \in \mathscr{A}(\mathcal{T})$ in the sense, that

$$g_{h|\Gamma_C \cup \Gamma_S} \cong g_{\mathrm{CS}} \quad \text{and} \quad (\partial_\nu g_h)_{|\Gamma_C} \cong g_{\mathrm{C}}.$$

A typical assumption is the existence of an extension $g \in H^2(\Omega)$ of g (typically denoted the same) to the whole domain such that $g_{\mathrm{CS}} = g_{|(\Gamma_C \cup \Gamma_S)}$ and $g_{\mathrm{C}} = (\partial_\nu g)_{|\Gamma_C}$. This only means slight regularity assumptions due to theorem 2.2. Hence, a natural choice for the approximation would be $g_h = \mathcal{I} g$ but this requires further regularity assumptions, e.g., $g \in H^{3+\varepsilon}(\Omega)$, even away from the boundary $\partial\Omega$. Since only the values (and normal derivatives) of g_h at the boundary are of importance, the behaviour of g inside Ω should have no influence.

In fact, the duality relation of the nodal basis with the degrees of freedom shows that most basis functions have vanishing trace on the boundary. So writing $\mathcal{I} g = \sum_{j=1}^N \Lambda_j(g)\varphi_j$ in the nodal basis shows that only a few coefficients $\Lambda_j(g)$ contribute to $g_{h|\Gamma_C \cup \Gamma_S}$ and $(\partial_\nu g_h)_{|\Gamma_C}$ for $g_h = \mathcal{I} g$. Thus, the other coefficients can be chosen as 0 and in turn the regularity assumption of g in the interior of Ω dropped. This leads to the definition of the *boundary interpolant* $\mathcal{I}_{\mathrm{b}}$, defined on the *boundary space*

$$B \coloneqq \left\{ v \in H^2(\Omega) \;\middle|\; \begin{array}{ll} \forall \Gamma \in \mathcal{E}_\Omega & v_{|\Gamma} \in H^3(\Gamma), (\partial_\nu v)_{|\Gamma} \in H^2(\Gamma) \text{ and} \\ \forall z \in \mathcal{V}_\Omega & \nabla v(z) \text{ and } D^2 v(z) \text{ exists} \end{array} \right\},$$

with slightly more regularity as necessary to define $\mathcal{I}_{\mathrm{b}}$ in order to obtain good approximation properties of $\mathcal{I}_{\mathrm{b}}$ in the following.

Definition 4.7 (Boundary interpolant)**.** Consider a regular triangulation $\mathcal{T}$ with vertices $\mathcal{V}$ and edge midpoints $\mathcal{M}$ and the associated extended Argyris space $\mathscr{A}(\mathcal{T})$. Define $\tilde{\mathcal{V}}_\Omega \coloneqq \mathcal{V}_\Omega \setminus \overline{\Gamma_F}$ and the index sets $I(z)$ for any vertex $z \in \mathcal{V}$ by

$$I(z) \coloneqq \begin{cases} \{1,2,3,4,5,6\} & \text{if } z \in \overline{\Gamma_C} \cap \tilde{\mathcal{V}}_\Omega, \\ \{1,2,3,4,5\} & \text{if } z \in \overline{\Gamma_C} \setminus \tilde{\mathcal{V}}_\Omega, \\ \{1,2,3,4,6\} & \text{if } z \in (\Gamma_S \setminus \overline{\Gamma_C}) \cap \tilde{\mathcal{V}}_\Omega, \\ \{1,2,4\} & \text{if } z \in (\Gamma_S \setminus \overline{\Gamma_C}) \setminus \tilde{\mathcal{V}}_\Omega, \\ \{\} & \text{else.} \end{cases}$$

Then the *boundary interpolant* $\mathcal{I}_{\mathrm{b}} : B \to \mathscr{A}(\mathcal{T})$ is defined as the mapping

$$v \mapsto \mathcal{I}_{\mathrm{b}} v = \sum_{z \in \mathcal{V}} \sum_{j \in I(z)} \Lambda_{z,j}(v)\varphi_{z,j} + \sum_{z \in \mathcal{M}(\Gamma_C)} \Lambda_{z,1}(v)\varphi_{z,1}. \tag{9}$$

Remark. The index set $I(z)$ is chosen such that only the relevant nodal values enter in the definition of $\mathcal{I}_{\mathrm{b}}$. Indeed, assume $E = \mathrm{conv}\{a,b\} \subset (\Gamma_C \cup \Gamma_S)$ and the local vertex coordinates read $\tau_E = \tau(a) = \tau(b)$ and $\nu_E = \nu(a) = \nu(b)$. Consider an arbitrary $v_h \in \mathscr{A}(\mathcal{T})$ as well as the local enumeration of the degrees of freedom from definition 4.3. Then $v_{h|E} \in P_5(E)$ is uniquely defined by the nodal values in tangent direction $\{\Lambda_{z,j}(v_h) \mid z \in \{a,b\}, j \in \{1,2,4\}\}$, in the sense that

$$v_{h|E} = \sum_{j \in \{1,2,4\}} (\Lambda_{a,j}(v)\varphi_{a,j|E} + \Lambda_{b,j}(v)\varphi_{b,j|E}).$$

Similarly, the normal component $(\partial_\nu v_h)_{|E} \in P_4(E)$ is defined by $\{\Lambda_{z,j}(v_h) \mid z \in \{a,b\}, j \in \{3,5\}\} \cup \{\Lambda_{\mathrm{mid}\, E,1}(v_h)\}$ corresponding to tangential derivatives of the normal component.

This is a well-known fact and required to show unisolvence of the local degrees of freedom. The assumption on the vertex coordinates holds if $z \notin \mathcal{V}_\Omega \setminus \Gamma_F$ is not a corner

vertex or a corner vertex with free boundary conditions on one side. Otherwise two distinct boundary edges $E_1, E_2 \in \mathcal{E}(\partial\Omega)$ meet at $z \in \mathcal{V}_\Omega \setminus \Gamma_F$. If one edge belongs to Γ_C and the other to $\Gamma_C \cup \Gamma_S$, this is case 1 in the definition of $I(z)$, then full information of the hessian is required. Otherwise both edges belong to Γ_S, this is case 3, and by definition 4.2 of the local vertex coordinates, $\tau(z) = \tau_{E_1}, \nu(z) = \tau_{E_2}$ and the first and second tangential derivatives at z on E_2 are given by $\{\Lambda_{z,3}, \Lambda_{z,6}\}$.

In summary, the boundary interpolant can be interpreted as the restriction of $\mathcal{I}$ to the nodal variables on the boundary $\partial\Omega$. Also the notion of boundary oscillations is important for the subsequent result of the best-approximation property of the boundary interpolant.

Definition 4.8 (Boundary oscillation)**.** Let a regular triangulation $\mathcal{T}$ of Ω into triangles with edges $\mathcal{E}$ and $v \in B$ be given. Define the boundary oscillation for v on a edge $E \in \mathcal{E}$ by

$$\operatorname{osc}_{CS}^2(E, v) \coloneqq \begin{cases} h_E^3 \|(1 - \Pi_{E,2})\partial_{\tau\tau\tau}^3 v\|_{L^2(E)}^2 & \text{if } E \in \mathcal{E}(\Gamma_C \cup \Gamma_S), \\ 0 & \text{else,} \end{cases} \tag{10}$$

$$\operatorname{osc}_{C}^2(E, v) \coloneqq \begin{cases} h_E^3 \|(1 - \Pi_{E,2})\partial_{\tau\tau\nu}^3 v\|_{L^2(E)}^2 & \text{if } E \in \mathcal{E}(\Gamma_C), \\ 0 & \text{else.} \end{cases} \tag{11}$$

Given a subset of boundary edges $\mathcal{S} \subseteq \mathcal{E}(\partial\Omega)$, the partial sums are frequently abbreviated by

$$\operatorname{osc}_{CS}(\mathcal{S}, v) \coloneqq \left(\sum_{E \in \mathcal{S}} \operatorname{osc}_{CS}^2(E, v) \right)^{1/2} \quad \text{and} \quad \operatorname{osc}_{C}(\mathcal{S}, v) \coloneqq \left(\sum_{E \in \mathcal{S}} \operatorname{osc}_{C}^2(E, v) \right)^{1/2}.$$

Set $\operatorname{osc}(\mathcal{S}, v) \coloneqq (\operatorname{osc}_{CS}^2(\mathcal{S}, v) + \operatorname{osc}_{C}^2(\mathcal{S}, v))^{1/2}$ for subsets $\mathcal{S} \subseteq \mathcal{E}(\partial\Omega)$.

In order to get the boundary oscillations, the best approximation property of $\mathcal{I}$ (of $\mathcal{I}_\mathrm{b}$) along an edge $E \in \mathcal{E}(\partial\Omega)$ has to be established first.

Lemma 4.2 (Boundary best-approximation)**.** *Let $v \in B$, such that $\mathcal{I}_\mathrm{b} v$ is defined, be given and define $c_\mathrm{b} = (1 - 720/(16\pi^4))^{-1/2}$. Then*

(1) $\forall E \in \mathcal{E}(\Gamma_C \cup \Gamma_S)$ $\quad \left\|\partial_{\tau\tau\tau}^3 (1 - \mathcal{I}_\mathrm{b}) v\right\|_{L^2(E)} = \left\|(1 - \Pi_{\mathcal{E},2})\partial_{\tau\tau\tau}^3 v\right\|_{L^2(E)}$

(2) $\forall E \in \mathcal{E}(\Gamma_S)$ $\quad \left\|\partial_{\tau\tau\nu}^3 (1 - \mathcal{I}_\mathrm{b}) v\right\|_{L^2(E)} \leq c_\mathrm{b} \left\|(1 - \Pi_{\mathcal{E},2})\partial_{\tau\tau\nu}^3 v\right\|_{L^2(E)}$

Remark. (1) means $\partial_{\tau\tau\tau}^3 \mathcal{I}_b v = \Pi_{\mathcal{E},2} \partial_{\tau\tau\tau}^3 v$ on E. The first 4 decimal digits of c_b read 1.3634.

Proof. (1) follows by repeated integration by parts with vanishing boundary terms due to interpolation at the endpoints. Consider an arbitrary quadratic polynomial $p_2 \in P_2(E)$, then

$$\langle p_2, \partial_{\tau\tau\tau}^3 (1 - \mathcal{I}_\mathrm{b}) v \rangle_{L^2(E)} = \langle \partial_{\tau\tau\tau}^3 p_2, (1 - \mathcal{I}_\mathrm{b}) v \rangle_{L^2(E)} = 0$$

vanishes due to $\partial_{\tau\tau\tau}^3 p_2 = 0$. This shows $\partial_{\tau\tau\tau}^3 \mathcal{I}_\mathrm{b} v = \Pi_{E,2} \partial_{\tau\tau\tau}^3 v$.

(2) is more involved, since equality does not hold. First observe, that due to the L^2-orthogonality, the Pythoagoras Theorem shows

$$\left\|\partial_{\tau\tau\nu}^3 (1 - \mathcal{I}_\mathrm{b}) v\right\|_{L^2(E)}^2 = \left\|(1 - \Pi_{\mathcal{E},2})\partial_{\tau\tau\nu}^3 v\right\|_{L^2(E)}^2 + \left\|\Pi_{\mathcal{E},2}\partial_{\tau\tau\nu}^3 v - \partial_{\tau\tau\nu}^3 \mathcal{I}_\mathrm{b} v\right\|_{L^2(E)}^2. \tag{12}$$

In order to show the estimate, the defect will be computed using the ansatz $\partial^2_{\tau\tau}(\partial_\nu \mathcal{I}_b v + d_m\varphi_m) = \Pi_{\mathcal{E},2}\partial^3_{\tau\tau\nu} v$ where $d_m \in \mathbb{R}$ and φ_m denotes the basis function in $P_4(E)$ that vanishes to first order at the endpoints and $\varphi_m(\text{mid } E) = 1$. This ansatz encapsulates that $\partial^3_{\tau\tau\nu}\mathcal{I}_b v$ fails to be the L^2-projection only by some multiple of $\partial^2_{\tau\tau}\varphi_m$. By the choice of φ_m and the interpolation at the endpoints, no boundary terms arise in the integration by parts formula; for $p_2 \in P_2(E)$ and $d_m \in \mathbb{R}$ it holds that

$$\langle p_2, \partial^3_{\tau\tau\nu}(1-\mathcal{I}_\mathrm{b})v - d_m\partial^2_{\tau\tau}\varphi_m\rangle_{L^2(E)} = \langle \partial^2_{\tau\tau}p_2, \partial_\nu(1-\mathcal{I}_\mathrm{b})v - d_m\varphi_m\rangle_{L^2(E)}.$$

Since $\partial^2_{\tau\tau}p_2 \in P_0(E)$, $\partial^2_{\tau\tau}(\partial_\nu\mathcal{I}_b v + d_m\varphi_m) = \Pi_{E,2}\partial^3_{\tau\tau\nu}v$ is the L^2-projection if and only if

$$\Pi_{E,0}(\partial_\nu(1-\mathcal{I}_\mathrm{b})v - d_m\varphi_m) = 0.$$

Therefore, with $d_m := \frac{\Pi_{E,0}(\partial_\nu(1-\mathcal{I}_\mathrm{b})v)}{\Pi_{E,0}\varphi_m}$ fixed, the identity for the deficiency $d_m\partial^2_{\tau\tau}\varphi_m = \Pi_{E,2}\partial^3_{\tau\tau\nu}v - \partial^3_{\tau\tau\nu}\mathcal{I}_b v$ holds true. In order to estimate its L^2 norm, choose a parametrisation of E in arc-length, i.e., identify $E \equiv [0, h_E]$. In terms of the two barycentric coordinates, φ_0 and φ_1, one has $\varphi_m = 16\varphi_0^2\varphi_1^2$. (With the identification $E \equiv [0,h_E]$, one has $\varphi_0(s) = 1 - s/h_E$ and $\varphi_1(s) = s/h_E$.) A direct calculation reveals $\Pi_{E,0}\varphi_m = \int_0^{h_E}\varphi_m\,\mathrm{d}s/h_E = 8/15$ and $\partial^4_{\tau\tau\tau\tau}\varphi_m = 384h_E^{-4}$. This leads to $d_m\partial^4_{\tau\tau\tau\tau}\varphi_m = 720h_F^{-4}\Pi_{E,0}(\partial_\nu(1-\mathcal{I}_\mathrm{b})v)$ and thus, for the L^2 norm of the deficiency,

$$\begin{aligned}
\left\|\Pi_{\mathcal{E},2}\partial^3_{\tau\tau\nu}v - \partial^3_{\tau\tau\nu}\mathcal{I}_b v\right\|^2_{L^2(E)} &= \left\langle d_m\partial^2_{\tau\tau}\varphi_m, \partial^3_{\tau\tau\nu}(1-\mathcal{I}_\mathrm{b})v\right\rangle_{L^2(E)} \\
&= \langle d_m\partial^4_{\tau\tau\tau\tau}\varphi_m, \partial_\nu(1-\mathcal{I}_\mathrm{b})v\rangle_{L^2(E)} \\
&= 720h_E^{-4}\langle\Pi_{\mathcal{E},0}(\partial_\nu(1-\mathcal{I}_\mathrm{b})v), \partial_\nu(1-\mathcal{I}_\mathrm{b})v\rangle_{L^2(E)} \\
&\le 720h_E^{-4}\left\|\partial_\nu(1-\mathcal{I}_\mathrm{b})v\right\|^2_{L^2(E)},
\end{aligned}$$

using $d_m\partial^2_{\tau\tau}\varphi_m \in P_2(E)$ and an integration by parts with vanishing boundary terms due to interpolation. A split of the domain of integration E into E_1, E_2 at the midpoint mid E of E allows the application of a Friedrichs inequality followed by a Poincare inequality with optimal constant $h_E/(2\pi)$ [Car19] on E_1 and E_2 separately (since $|E_1| = |E_2| = h_E/2$). Note that by definition of the interpolation, it holds that $\partial_\nu(1-\mathcal{I}_\mathrm{b})v$ vanishes at both endpoints and the midpoint of E and hence vanishes at the endpoints of both E_1 and E_2. Consequently, $\partial^2_{\tau\nu}(1-\mathcal{I}_\mathrm{b})v$ has zero integral mean over both E_1 and E_2. Hence

$$\begin{aligned}
\left\|\partial_\nu(1-\mathcal{I}_\mathrm{b})v\right\|^2_{L^2(E)} &\le \frac{h_E^4}{16\pi^4}\left(\left\|\partial^3_{\tau\tau\nu}(1-\mathcal{I}_\mathrm{b})v\right\|^2_{L^2(E_1)} + \left\|\partial^3_{\tau\tau\nu}(1-\mathcal{I}_\mathrm{b})v\right\|^2_{L^2(E_2)}\right) \\
&= \frac{h_E^4}{16\pi^4}\left\|\partial^3_{\tau\tau\nu}(1-\mathcal{I}_\mathrm{b})v\right\|^2_{L^2(E)}.
\end{aligned}$$

Collecting the terms leads to an estimate of the deficiency and an absorption by the left-hand side in (12) produces

$$\left(1 - \frac{720}{16\pi^4}\right)\left\|\partial^3_{\tau\tau\nu}(1-\mathcal{I}_\mathrm{b})v\right\|^2_{L^2(E)} \le \left\|(1-\Pi_{\mathcal{E},2})\partial^3_{\tau\tau\nu}v\right\|^2_{L^2(E)}$$

so that a division by $c_\mathrm{b}^2 > 0$ finishes the proof.

□

This shows that $\mathcal{I}_\mathrm{b}$ yields very good approximation results.

Lemma 4.3. *Let $\mathcal{T}$ be a regular triangulation of Ω. Given $v \in B$, for which the boundary interpolant $\mathcal{I}_\mathrm{b} v$ is defined, then there exists a constant $C_\mathrm{osc} > 0$, only depending on Ω, with*

$$\min_{w \in (1-\mathcal{I}_\mathrm{b})v + V} |w|^2_{H^2(\Omega)} \leq C^2_\mathrm{osc} \mathrm{osc}^2(\mathcal{E}(\partial\Omega), v).$$

Proof. The proof is a combination of the extension theorem with a best-approximation property of the boundary interpolant. First, consider the continuous problem and define a function $t = (\varphi_j, \psi_j)_j \in Tr$ in the trace space Tr from definition 2.1. Set $v_h = \mathcal{I}_\mathrm{b} v$ and define $t = (\varphi_j, \psi_j)_j \in Tr$ by setting on each edge $\Gamma_j \in \mathcal{E}_\Omega$ of Ω

$$\varphi_j = \begin{cases} v - v_h & \text{on } (\Gamma_C \cup \Gamma_S) \cap \Gamma_j \\ 0 & \text{else} \end{cases} \quad \text{and} \quad \psi_j = \begin{cases} \partial_\nu (v - v_h) & \text{on } \Gamma_C \cap \Gamma_j \\ 0 & \text{else} \end{cases}.$$

It is easily verified that by this choice (and the properties of the boundary interpolant $\mathcal{I}_\mathrm{b}$), that φ_j vanishes to second order and the ψ_j vanish to first order at the endpoints of each edge $E \in \mathcal{E}(\Omega)$. Thus, the compatibility conditions (2) are satisfied and $\varphi_j \in C^2(\Gamma_j), \psi_j \in C^1(\Gamma_j)$, by the Sobolev embedding $H^k(\Gamma_j) \subset C^{k-1}(\Gamma_j), k = 3, 2$ [Gri85]. Hence, $t \in Tr$ and even $\varphi_j \in H^3(\Gamma_j), \psi_j \in H^2(\Gamma_j)$ using the regularity of $\gamma_1(v) \in Tr$ and that $v_h \in C^2(\Gamma_j)$ is piecewise smooth. Clearly, $Et \in v - v_h + V$ and one finds for this particular choice by theorem 2.4,

$$\min_{w \in v - v_h + V} |w|^2_{H^2(\Omega)} \leq |Et|^2_{H^2(\Omega)} \leq \|Et\|^2_{H^2(\Omega)} \leq C_\mathrm{E} \sum_{j=1}^{n} \left(\|\varphi_j\|^2_{H^{3/2}(\Gamma_j)} + \|\psi_j\|^2_{H^{1/2}(\Gamma_j)} \right).$$

In order to estimate the broken Sobolev norms appearing here, the Gagliardo-Nierenberg inequality, theorem 2.1, is used to find constants $C_{\mathrm{GN},1}, C_{\mathrm{GN},2} > 0$ such that

$$\begin{aligned} &(1) \quad \|\varphi_j\|^2_{H^{3/2}(\Gamma_j)} \leq C_{\mathrm{GN},1} \|\varphi_j\|_{L^2(\Gamma_j)} \|\varphi_j\|_{H^3(\Gamma_j)}, \\ &(2) \quad \|\psi_j\|^2_{H^{1/2}(\Gamma_j)} \leq C_{\mathrm{GN},2} \|\psi_j\|^{3/2}_{L^2(\Gamma_j)} \|\psi_j\|^{1/2}_{H^2(\Gamma_j)}. \end{aligned}$$

Whereas φ_j (resp. ψ_j) is identically zero on Γ_F (resp. $\Gamma_F \cup \Gamma_S$), one finds, using multiple Friedrichs inequalities [Car19], that on the remaining edges E,

$$\begin{aligned} \|(\partial_\tau)^k \varphi_j\|_{L^2(E)} &\leq \left(\frac{h_E}{\pi}\right)^{3-k} \left\|\partial^3_{\tau\tau\tau}(1 - \mathcal{I}_\mathrm{b})v\right\|_{L^2(E)} && (k = 0, 1, 2), \\ \|(\partial_\tau)^k \psi_j\|_{L^2(E)} &\leq \left(\frac{h_E}{\pi}\right)^{2-k} \left\|\partial^3_{\tau\tau\nu}(1 - \mathcal{I}_\mathrm{b})v\right\|_{L^2(E)} && (k = 0, 1). \end{aligned}$$

Using the best-approximation property, lemma 4.2, this allows for

$$\|\varphi_j\|^2_{L^2(\Gamma_j)} = \sum_{E \in \mathcal{E}(\Gamma_j \cap (\Gamma_C \cup \Gamma_S))} \|\varphi_j\|^2_{L^2(E)} \leq \pi^{-6} \left\|h^3_{\mathcal{E}}(1 - \Pi_{\mathcal{E},2})\partial^3_{\tau\tau\tau} v\right\|^2_{L^2(\Gamma_j \cap (\Gamma_C \cup \Gamma_S))}$$

as well as $\|\psi_j\|^2_{L^2(\Gamma_j)} \leq \pi^{-4} c^2_\mathrm{b} \left\|h^2_{\mathcal{E}}(1 - \Pi_{\mathcal{E},2})\partial^3_{\tau\tau\nu} v\right\|^2_{L^2(\Gamma_j \cap \Gamma_C)}$. Similarly, with the trivial estimate $h_E \leq |\Omega|$ and $c_{\Omega,\mathrm{m}} \coloneqq \sum_{k=0}^{m} (|\Omega|/\pi)^{2k}$ for $m = 2, 3$, the full H^3 (resp. H^2) norm is estimated as

$$\begin{aligned} \|\varphi_j\|^2_{H^3(\Gamma_j)} &\leq c_{\Omega,3} \left\|(1 - \Pi_{\mathcal{E},2})\partial^3_{\tau\tau\tau} v\right\|^2_{L^2(\Gamma_j \cap (\Gamma_C \cup \Gamma_S))} \quad \text{and} \\ \|\psi_j\|^2_{H^2(\Gamma_j)} &\leq c_{\Omega,2} c^2_\mathrm{b} \left\|(1 - \Pi_{\mathcal{E},2})\partial^3_{\tau\tau\nu} v\right\|^2_{L^2(\Gamma_j \cap \Gamma_C)}. \end{aligned}$$

Eventually,

$$\min_{w\in(1-\mathcal{I}_{\mathrm{b}})v+V}|w|^2_{H^2(\Omega)} \leq C_{\mathrm{E}}\pi^{-3}\Big(C_{\mathrm{GN},1}c^{1/2}_{\Omega,3}\left\|h^{3/2}_{\mathcal{E}}(1-\Pi_{\mathcal{E},2})\partial^3_{\tau\tau\tau}v\right\|^2_{L^2(\Gamma_C\cup\Gamma_S)} + C_{\mathrm{GN},2}c^{1/4}_{\Omega,2}c^2_{\mathrm{b}}\left\|h^{3/2}_{\mathcal{E}}(1-\Pi_{\mathcal{E},2})\partial^3_{\tau\tau\nu}v\right\|^2_{L^2(\Gamma_C)}\Big),$$

using that $h_{\mathcal{E}}$ is a piecewise constant. The statement follows by setting

$$C^2_{\mathrm{osc}} \coloneqq C_{\mathrm{E}}\pi^{-3}\max\{C_{\mathrm{GN},1}c^{1/2}_{\Omega,3}, C_{\mathrm{GN},2}c^{1/4}_{\Omega,2}c^2_{\mathrm{b}}\}.$$

□

4.4 Discretisation of the biharmonic problem

The starting point of the FEM for the biharmonic or plate problem is the reformulation of (SF) as a variational or weak formulation. To simplify matters, assume that the essential boundary data is given by a single, globally defined function $g \in H^2(\Omega)$ with $g_{\mathrm{CS}} = g_{|(\Gamma_C\cup\Gamma_S)}$ and $g_{\mathrm{C}} = (\partial_\nu g)_{|\Gamma_C}$. Via definition, the essential boundary conditions are enforced by the space of test functions $V \coloneqq H^2_{\Gamma_C}(\Omega)\cap H^1_{\Gamma_S}$, leading to the space of admissible functions $\mathcal{A} \coloneqq g + V$. The weak formulation of the plate bending problem is obtained in the usual manner and reads; find $u \in \mathcal{A}$ such that

$$a(u,v) = L(v) \qquad \text{for all } v \in V. \tag{WF}$$

The source $F \in V'$ and the natural boundary data $t_{\mathrm{SF}} \in L^2(\Gamma_F\cup\Gamma_S), t_{\mathrm{F}} \in L^2(\Gamma_F)$ enter in the definition of the right-hand side $L \in V'$,

$$L(v) \coloneqq F(v) - \langle t_{\mathrm{SF}}, \partial_\nu v\rangle_{L^2(\Gamma_F\cup\Gamma_S)} + \langle t_{\mathrm{F}}, v\rangle_{L^2(\Gamma_F)}.$$

With Poisson's ratio $\sigma = 0$, the bilinear form $a : H^2(\Omega)\times H^2(\Omega)\to\mathbb{R}$ takes the form

$$a(u,v) \coloneqq (D^2u, D^2v)_{L^2(\Omega)}.$$

Consult [CM13; BS08] for modifications in case $\sigma \neq 0$. Solvability and uniqueness of (WF) depends on the boundary conditions. Clearly, the case $\Gamma_F = \partial\Omega$ (and thus $\Gamma_C = \Gamma_S = \emptyset$) leads to $V = H^2(\Omega)$ and solutions may only be unique up to the addition of affine functions.

Assumption 4.1. With Γ_C, Γ_S and the definition of $V = H^2_{\Gamma_C}(\Omega)\cap H^1_{\Gamma_S}$, the intersection $V \cap P_1(\Omega) = \{0\}$ of V with the affine functions is trivial.

This assumption is standard in the analysis of the biharmonic equation and not only guarantees that $a_{|V\times V} : V\times V\to\mathbb{R}$ becomes a scalar product on V, but also implies uniqueness and solvability of (WF) by application of the Lax-Milgram lemma, cf. [BS08]. The connection between (SF) and (WF) is given by the following lemma.

Lemma 4.4. *Suppose $u \in H^4(\Omega), g \in H^2(\Omega), t_{\mathrm{SF}} \in L^2(\Gamma_F\cup\Gamma_S)$ and $t_{\mathrm{F}} \in L^2(\Gamma_F)$, then u solves* (SF) *if and only if u solves* (WF) *with the representation $F(v) = (f,v)_{L^2(\Omega)}$ for some $f \in L^2(\Omega)$.*

The proof follows standard arguments, relying on the integration by parts formula. It is given for completeness.

Proof. Suppose that $u = g + u_0 \in H^4(\Omega)$ solves (WF) with $u_0 \in V$. The essential boundary conditions are satisfied by the definition of the space of admissible functions $\mathcal{A}$. Integration by parts with $v \in V \subset H^2(\Omega)$ yields

$$\begin{aligned} a(u,v) &= (\Delta^2 u, v) + \langle \nabla(\partial_\nu u), \nabla v \rangle_{L^2(\partial\Omega)} - \langle \partial_\nu \Delta u, v \rangle_{L^2(\partial\Omega)} \\ &= (\Delta^2 u, v) + \left\langle \partial^2_{\nu\nu} u, \partial_\nu v \right\rangle_{L^2(\Gamma_F \cup \Gamma_S)} - \left\langle \partial^3_{\tau\tau\nu} u + \partial_\nu \Delta u, v \right\rangle_{L^2(\Gamma_F)}, \end{aligned} \tag{13}$$

where the boundary conditions on V are taken into consideration in the last equality. Note, that the L^2-scalar product is independent of the underlying coordinate system, i.e.

$$\langle \nabla(\partial_\nu u), \nabla v \rangle_{L^2(\partial\Omega)} = \left\langle \partial^2_{\nu\nu} u, \partial_\nu v \right\rangle_{L^2(\partial\Omega)} + \left\langle \partial^2_{\tau\nu} u, \partial_\tau v \right\rangle_{L^2(\partial\Omega)}.$$

An additional integration by parts on $\partial\Omega$ shows

$$\left\langle \partial^2_{\tau\nu} u, \partial_\tau v \right\rangle_{L^2(\partial\Omega)} = - \left\langle \partial^3_{\tau\tau\nu} u, v \right\rangle_{L^2(\partial\Omega)},$$

since u has sufficient regularity and $\partial\Omega$ has empty boundary. This justifies (13). Testing with $v \in C_0^\infty(\Omega) \subset H_0^2(\Omega)$, dense in $L^2(\Omega)$, shows

$$\begin{aligned} F(v) &= a(u,v) + \langle t_{\mathrm{SF}}, \partial_\nu v \rangle_{L^2(\Gamma_F \cup \Gamma_S)} - \langle t_{\mathrm{F}}, v \rangle_{L^2(\Gamma_F)} \\ &= (\Delta^2 u, v)_{L^2(\Omega)} \end{aligned}$$

since all boundary terms vanish. Hence, $F(v) = (f, v)_{L^2(\Omega)}$ for $f = \Delta^2 u \in L^2(\Omega)$ by a density argument. Realisation of the natural boundary conditions in L^2 follow by similar arguments. Conversely, the direction (SF) $\Rightarrow$ (WF) follows by integration with test functions $v \in V$ and using integration by parts as in (13). □

In order to define a discrete counterpart of the space of admissible functions $\mathcal{A} = g + V \subset H^2(\Omega)$, a discrete version of the essential boundary data g is required. By the considerations in section 4.3, $g_h = \mathcal{I}_{\mathrm{b}} g$ is a reasonable choice. This leads to the assumption that $g \in B$, so that $\mathcal{I}_{\mathrm{b}} g$ is well-defined.

Consider a regular triangulation $\mathcal{T} = \mathcal{T}_L$ (generated by some admissible sequence $(\mathcal{T}_\ell)_{\ell=0}^L$) and $\mathcal{I}_{\mathrm{b}} : B \to \mathscr{A}(\mathcal{T})$. A straight-forward way to define the discrete space of test functions $V(\mathcal{T}) \subset V$ is by the definition

$$V(\mathcal{T}) \coloneqq \{ v_h \in \mathscr{A}(\mathcal{T}) \mid \mathcal{I}_{\mathrm{b}} v_h = 0 \} \subset V.$$

Let $\{\varphi_1, \dots, \varphi_N\}$ denote the nodal basis of $\mathscr{A}(\mathcal{T})$ and $\{\Lambda_1, \dots, \Lambda_N\}$ the degrees of freedom for $\mathscr{A}(\mathcal{T})$. By the above definition, this induces a nodal basis and degrees of freedom for the space $V(\mathcal{T})$, as follows

$$\begin{aligned} &\{\varphi_j \mid j = 1, \dots, N \text{ and } \varphi_j \text{ does not appear in (9)}\} && \text{(nodal basis)}, \\ &\{\Lambda_j \mid j = 1, \dots, N \text{ and } \Lambda_j \text{ does not appear in (9)}\} && \text{(degrees of freedom)}. \end{aligned}$$

The discrete space of test functions is the defined as $\mathcal{A}(\mathcal{T}) \coloneqq g_h + V(\mathcal{T}) \subset \mathscr{A}(\mathcal{T})$ with the boundary data approximation $g_h = \mathcal{I}_{\mathrm{b}} g \in \mathscr{A}(\mathcal{T})$. Since the extended Argyris test function space $V(\mathcal{T}) \subset V$ is conforming as well, the weak problem (WF) induces the corresponding discretisation; find $u_h \in \mathcal{A}(\mathcal{T})$ such that for all $v_h \in V(\mathcal{T})$

$$a(u_h, v_h) = L(v_h). \tag{DWF}$$

The solution to the discrete problem (DWF) can be obtained by a direct solver or with an iterative solver such as multigrid, cf. section 3.

4.5 Full multigrid AFEM

The rest of this section is devoted to the connection of AFEM and full multigrid. Consider a sequence of triangulations $\mathcal{T}_\ell$ with initial triangulation $\mathcal{T}_0$ and adopt the notation that an index ℓ refers to the corresponding discrete object at level ℓ. To connect the algebraic viewpoint from section 3 with the discrete weak formulation on level ℓ, one rewrites (DWF) as

$$a(u_{0,\ell}, v_\ell) = L(v_\ell) - a(g_\ell, v_\ell) \quad \text{for all } v_\ell \in V(\mathcal{T}_\ell), \tag{DWF$_\ell$}$$

by splitting the discrete solution $u_\ell = u_{0,\ell} + g_\ell$ on level ℓ into $u_{0,\ell} \in V(\mathcal{T}_\ell)$ and the boundary approximation $g_\ell = \mathcal{I}_\mathrm{b} g$. With this formulation, (DWF$_\ell$) takes the form of the algebraic setting (4) with modified right-hand side. The multigrid scheme is then applied to get an approximation $u_{0,\ell}^\varepsilon$ of $u_{0,\ell}$ which in turn yields an approximation $u_\ell^\varepsilon \coloneqq u_{0,\ell}^\varepsilon + g_\ell$ of the discrete problem u_ℓ.

Algorithm 3 FMGAFEM

Input: Initial triangulation $\mathcal{T}_0$, bulk parameter $\theta > 0$ and stopping criterion $\varepsilon > 0, r, s \in \mathbb{N}$

for $\ell = 0, 1, \dots$ **until** STOP **do**

Solve the algebraic problem for (DWF$_\ell$) with multigrid, i.e.,
compute $u_{0,\ell}^\varepsilon = MG_{r,s}(\varepsilon)$ from initial solution $u_{0,\ell}^0$ and set $u_\ell^\varepsilon \coloneqq u_{0,\ell}^\varepsilon + g_\ell$

Compute for all $T \in \mathcal{T}_\ell$ refinement indicators $\eta(T, u_\ell^\varepsilon)$ using u_ℓ^ε

Mark (almost) minimal subset $\tau_\ell \subset \mathcal{T}_\ell$ with

$$\theta \sum_{T \in \mathcal{T}_\ell} \eta(T, u_\ell^\varepsilon)^2 \leq \sum_{T \in \tau_\ell} \eta(T, u_\ell^\varepsilon)^2$$

Refine $\mathcal{T}_\ell \longrightarrow T_{\ell+1}$: smallest NVB refinement of $\mathcal{T}_\ell$ with $\tau_\ell \subseteq \mathcal{T}_\ell \setminus \mathcal{T}_{\ell+1}$

end for

Output: Sequences u_ℓ^ε and $\mathcal{T}_\ell$

In order to address the question of the initial solution to the multigrid algorithm, observe that u_ℓ are approximations to the exact solution u of (WF). Hence, the approximation $u_{\ell-1}^\varepsilon$ to $u_{\ell-1}$ is supposed to approximate u_ℓ^ε as well,, since

$$|u_\ell - u_{\ell-1}^\varepsilon|_{H^2(\Omega)} \leq |u_\ell - u_{\ell-1}|_{H^2(\Omega)} + |u_{\ell-1} - u_{\ell-1}^\varepsilon|_{H^2(\Omega)}$$

and the terms on the right-hand side converge. This leads to the choice $u_{0,\ell}^0 \coloneqq u_{\ell-1}^\varepsilon - g_\ell$ for the initial solution to the multigrid iterations on level ℓ. For simplicity, $u_{0,-1}^\varepsilon \coloneqq 0$ is understood.

This leads to the *full multigrid adaptive finite element method* (FMGAFEM), algorithm 3, as soon as a refinement indicator η to steer adaptive mesh refinement has been defined.

This section finishes by recalling that $\mathbb{B}_\ell$ is a uniform approximate inverse of $\mathbb{A}_\ell$ in $\ell \in \mathbb{N}_0$.

Theorem 4.5 ([CH21, Theorem 7]). *Let $\mathbb{B}_\ell$ and $\mathbb{A}_\ell$ denote the multigrid operators from section 3 for the hierarchical Argyris FMGAFEM, algorithm 3. Then there exists $c_\infty \in \mathbb{R}$ with*

$$\sup_{\ell \in \mathbb{N}_0} \|\mathbb{I} - \mathbb{B}_\ell \mathbb{A}_\ell\|_a \leq \frac{c_\infty}{1 + c_\infty} < 1.$$

The proof bounds the constant c_L in theorem 3.4 by choosing a decomposition for the infimum, producing a uniform bound in $L \in \mathbb{N}_0$ by c_∞. Recall (DWF$_\ell$) and the definition in section 3, to see that there is no dependence of the operators $\mathbb{B}_\ell$ and $\mathbb{A}_\ell$ on the inhomogeneous boundary data g_ℓ. So the only difference is that the spaces $V(\mathcal{T})$ here allow nodal bases on the boundary and, hence, the proof in [CH21] applies verbatim.

5 A posteriori error analysis

In this section, convergence of the error $e \coloneqq u - u_h$ of the exact solution u to (WF) and some u_h will be analysed. The discrete approximation u_h here is a solution to (DWF) on some triangulation $\mathcal{T}$. Typical a priori results lead to convergence results for the error e in terms on the maximum mesh size $h_{\max} = \max_{T\in\mathcal{T}} h_T$ only. With the requirement of sufficent smoothness, $u \in H^{2+s}(\Omega), 0 \leq s \leq 4$, for discretisations with the Argyris elements, rates of order $\mathcal{O}(h_{\max}^s)$ can be shown, cf. [BS08]. Local elliptic regularity asserts that, away from the boundary, u has 4 square-integrable weak derivatives more than the source term F. Hence, singularities of u are related to singularities of the boundary or of the source, cf. the decomposition of u in regular and singular components [Gri92] and shift theorems for the biharmonic operator [BBP02].

5.1 Error estimator

Adaptive mesh refinement is the method of choice when dealing with such localised singularites. A key ingredient for these mesh refinement strategies is an a posteriori error estimator. Moreover, in the framework of the *axioms of adaptivity* (AoA), cf. [Car+14; CR17], an estimator fulfilling the four axioms (A1)-(A4) immediately leads to optimal convergence rates of the adaptive finite element method (AFEM), introduced in section 1.

Before stating the estimator an assumption on the source has to be made. Recall $F \in V'$ and that $V \subset H^2(\Omega)$ is a Hilbert space with the standard H^2-scalar product

$$(\bullet,\bullet)_{H^2(\Omega)} \coloneqq \left(D^2\bullet, D^2\bullet\right)_{L^2(\Omega)} + (\nabla\bullet,\nabla\bullet)_{L^2(\Omega)} + (\bullet,\bullet)_{L^2(\Omega)}\,.$$

Hence, by the Riesz representation theorem, there exists $f \in V$ such that $F = (f,\bullet)_{H^2(\Omega)}$. In order to show the reduction property of the estimator, it will be necessary to assume that F can be represented in a similar form without appearance of the Hessian. Nevertheless, point evaluations (denoted δ_z for points z) do not amount into any problems if they are degrees of freedom for all discretisations, i.e., if they have support at the interior or free vertices $z \in \mathcal{V}_0 \setminus \mathcal{V}_0(\overline{\Gamma_C \cup \Gamma_S})$ of the initial triangulation $\mathcal{T}_0$.

Assumption 5.1. Fix an initial triangulation $\mathcal{T}_0$ with vertices $\mathcal{V}_0$ and set $\tilde{\mathcal{V}}_0 \coloneqq \mathcal{V}_0 \setminus \mathcal{V}_0(\overline{\Gamma_C \cup \Gamma_S})$. Then $F \in V'$ can be represented as

$$F = (f_1,\bullet)_{L^2(\Omega)} + (f_2,\nabla\bullet)_{L^2(\Omega)} + \sum_{z\in\tilde{\mathcal{V}}_0} \beta_z\delta_z \tag{14}$$

with $f_1 \in L^2(\Omega), f_2 \in L^2(\Omega)^2$ and $\beta_z \in \mathbb{R}$.

This assumption is fulfilled if $F \in V' \cap (H^{-1}(\Omega) \cup D_z)$ where $D_z = \operatorname{span}\{\delta_z \mid z \in \tilde{\mathcal{V}}_0\}$, using arguments similar to the above. For the rest of this section it will be assumed that assumption 5.1 holds and $F \in V'$ is given by (14) for fixed f_1 and f_2.

Residual based a posteriori error estimators do not rely on the knowledge of the exact solution. They merely exploit certain orthogonalities with respect to the error $e \coloneqq u - u_h$ and provide an equivalent measure of the error with respect to some norm. Most importantly these estimators are usually computable by means of the approximation u_h and related quantities.

Definition 5.1 (Error estimator)**.** Let $u_h \in \mathscr{A}(\mathcal{T})$ be the solution to the discrete problem (DWF) given by $F \in V'$, given by (14), and boundary data $g \in B$. Then the error

estimator η on a triangulation $\mathcal{T}$ is defined as the ℓ^2-norm of the local contributions on $T \in \mathcal{T}$,

$$\eta^2(\mathcal{T},T) = |T|^2\|f_1 - \Delta^2 u_h\|^2_{L^2(\Omega)} + |T|\|f_2\|^2_{L^2(\Omega)} + \sum_{E\in\mathcal{E}(T)} \Big(|T|^{1/2}\|\mathcal{R}_{\mathrm{FS}}u_h\|^2_{L^2(E)} + |T|^{3/2}\|\mathcal{R}_{\mathrm{F}}u_h\|^2_{L^2(E)}\Big) + \operatorname{osc}^2(\mathcal{E}(T),g). \tag{15}$$

with the edge residuals defined as

$$(\mathcal{R}_{\mathrm{FS}}u_h)_{|E} \coloneqq \begin{cases} [\partial^2_{\nu\nu}u_h]_E & \text{if } E \in \mathcal{E}(\Omega), \\ \partial^2_{\nu\nu}u_h - t_{\mathrm{SF}} & \text{if } E \in \mathcal{E}(\Gamma_F \cup \Gamma_S), \\ 0 & \text{else,} \end{cases}$$

$$(\mathcal{R}_{\mathrm{F}}u_h)_{|E} \coloneqq \begin{cases} [\partial^3_{\tau\tau\nu}u_h + \partial_\nu \Delta u_h]_E & \text{if } E \in \mathcal{E}(\Omega), \\ \partial^3_{\tau\tau\nu}u_h + \partial_\nu\Delta u_h - t_{\mathrm{F}} & \text{if } E \in \mathcal{E}(\Gamma_F), \\ 0 & \text{else.} \end{cases}$$

For convenience, the ℓ^2 norm over any subset of triangles $\mathcal{S} \subset \mathcal{T}$ is abbreviated by the sum convention

$$\eta(\mathcal{T},\mathcal{S}) \coloneqq \left(\sum_{T\in\mathcal{S}} \eta^2(\mathcal{T},T)\right)^{1/2}.$$

In case of $\mathcal{S} = \mathcal{T}$, the full estimator reads $\eta(\mathcal{T}) \coloneqq \eta(\mathcal{T},\mathcal{T})$.

This estimator depends on the triangulation $\mathcal{T}$ though the discrete solution u_h to (DWF). Furthermore, the local error contributions can be used to steer adaptive mesh refinement as refinement indicators and lead to AFEM, algorithm 4. The difference of

Algorithm 4 AFEM

Input: Initial triangulation $\mathcal{T}_0$, bulk parameter $0 < \theta < 1$

Solve the discrete problem (DWF) on $\mathcal{T}_\ell$

Compute for all $T \in \mathcal{T}_\ell$ the local estimations $\eta(\mathcal{T}_\ell, T)$ from (15)

Mark (almost) minimal subset $\tau_\ell \subset \mathcal{T}_\ell$ with

$$\theta\eta(\mathcal{T}_\ell)^2 \le \eta(\mathcal{T}_\ell,\tau_\ell)^2$$

Refine $\mathcal{T}_\ell \longrightarrow T_{\ell+1}$: smallest NVB refinement of $\mathcal{T}_\ell$ with $\tau_\ell \subseteq \mathcal{T}_\ell \setminus \mathcal{T}_{\ell+1}$

Output: Sequence of triangulations $\mathcal{T}_\ell$ and discrete solutions u_ℓ

two admissible triangulations $\mathcal{T}$ and $\widehat{\mathcal{T}}$ is a key notion in the statement of the axioms of adaptivity.

Definition 5.2. Consider two admissible triangulations $\mathcal{T}, \widehat{\mathcal{T}} \in \mathbb{T}(\mathcal{T}_0)$ with discrete solutions u_h and $\widehat{u_h}$, respectively. The distance between $\mathcal{T}$ and $\widehat{\mathcal{T}}$ is defined as

$$\delta(\mathcal{T},\widehat{\mathcal{T}}) \coloneqq |\widehat{u_h} - u_h|_{H^2(\Omega)}.$$

Before the statement and verification of the AoA, a technical tool is required first.

5.2 Discrete quasi-interpolation operator

As pointed out in the beginning of this section, a general solution u of the problem (WF) may not lie in the space $H^3(\Omega)$ but always in $H^2(\Omega)$ (by the choice of $V \subset H^2(\Omega)$). Neither does the error $e \coloneqq u - u_h$ of some discrete approximation $u_h \in V_h$ necessarily lie in $H^3(\Omega)$. Hence, the interpolation operator $\mathcal{I}$ may be defined neither for u nor for e. This makes the analysis somewhat complicated when comparing functions $w \in H^2(\Omega)$ with a discrete function on the discrete level. In other words, if one wants to consider a discrete representation of w and w has sufficient regularity only then it is possible to use $\mathcal{I}w$ which is guaranteed to have certain approximation properties as stated in theorem 4.1.

A typical tool in the analysis is thus a quasi-interpolation operator $J : H^2(\Omega) \to V(\mathcal{T})$ with approximation properties similar to $\mathcal{I}$, cf. [Clé75]. For C^1-conforming finite element spaces, the work [GS02] constructs a quasi-interpolant in a similar fashion to the well-known Scott-Zhang quasi-interpolation operator [SZ90]. In [CH21] a discrete quasi-interpolant is defined for the extended Argyris space (with obvious modification for the classical Argyris space). Even though the construction in [CH21] assumes homogeneous Dirichlet boundary conditions, the construction therein is easily generalized to the more general situation considered here. Actually, the changes due to non-homogeneous and mixed boundary conditions alone are clear from [GS02]. The following assumes, that $\mathscr{A}(\mathcal{T})$ and $V(\mathcal{T})$ refer to the extended Argyris space, defined in section 4 for triangulations $\mathcal{T} = \mathcal{T}_L$, generated by a sequence $(\mathcal{T}_\ell)_\ell$ of successive admissible refinements and initial triangulation $\mathcal{T}_0$.

Theorem 5.1 (discrete quasi-interpolation)**.** *There exists constants $c_{\mathrm{apx}}, c_{\mathrm{c}} > 0$ such that for any regular triangulation $\mathcal{T}$ and any admissible refinement thereof $\widehat{\mathcal{T}} \in \mathbb{T}(\mathcal{T})$ there exists a linear* ***quasi-interpolation operator*** *$J : H^2(\Omega) \to \mathscr{A}(\mathcal{T})$ satisfying $(a) - (e)$.*

(a) $J(V) = V(\mathcal{T})$ and $Jv(z) = v(z)$ for all $z \in \mathcal{V}$

(b) If $\widehat{v_h} \in \mathscr{A}(\widehat{\mathcal{T}})$: $\widehat{v_h}_{|T} = (J\widehat{v_h})_{|T}$ for any $T \in \mathcal{T} \cap \widehat{\mathcal{T}}$

(c) If $v \in H^2(\Omega)$ such that $\gamma_1(v) \in \gamma_1(\mathscr{A}(\mathcal{T}))$ then $v_{|\partial\Omega} = (Jv)_{|\partial\Omega}$ (boundary preserving)

(d) $|Jf|_{H^m(\Omega)} \leq c_{\mathrm{c}} |f|_{H^m(\Omega)}$ for any $f \in H^2(\Omega), m = 0, 1, 2$ (stability)

(e) $\sum_{m=0}^{2} h_T^{m-2} |(1 - J)f|_{H^m(T)} \leq c_{\mathrm{apx}} |f|_{H^2(\Omega(T))}$ for any $f \in V$ (approximation)

Proof (sketch). The difference to [CH21] in the requirement of J lies in the extension to the boundary and the exactness at the vertices. For this, the domain of definition is reduced to $H^2(\Omega)$. Recall the nodal basis $\{\varphi_1, \dots, \varphi_N\}$ and the degrees of freedom $\{\Lambda_1, \dots, \Lambda_N\}$ for $\mathscr{A}(\mathcal{T})$ with the identification $\Lambda_k \equiv \Lambda_{z,m(z)}$ as in definition 4.3. Given $f \in H^2(\Omega)$, define

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

as the discrete function with coefficients given by $M_k(f)$. The linear functionals M_k are chosen as in [CH21; GS02] to be of the integral form

$$M_k(f) = \int_{\kappa_k} \tilde{\psi}_k(x) \tilde{D}_k f(x) \, \mathrm{d}x$$

with derivative $\tilde{D}_k$ of order $\{0, 1\}$ and weights $\tilde{\psi}_k$. Except for $M_k \equiv M_{z,1}$ (under the same identification as above) for a vertex $z \in V$ where the point-evaluation $M_{z,1} = \Lambda_{z,1} = \delta_z$ is

chosen. The domain of integration for the other $M_k \equiv M_{z,j}$ is a triangle $\kappa_{z,j} = T_{z,j} \in \mathcal{T}$ with $\mathrm{int}(T_{z,j}) \cap \mathrm{supp}\Lambda_{z,j} \neq \emptyset$ if z is an interior vertex $z \in \mathcal{V}(\Omega)$, cf. [CH21] for details in the choice of $T_{z,j}$ that lead to (a) and (b). In the case of a boundary node $z \in \mathcal{N}(\partial\Omega) = \mathcal{V}(\partial\Omega) \cup \mathcal{M}(\partial\Omega)$, $\kappa_{z,j} = E \in \mathcal{E}(\partial\Omega)$ has to be an edge $E \in \mathcal{E}(\partial\Omega)$ on the boundary such that $\mathrm{int}(E) \cap \mathrm{supp}\Lambda_{z,j} \neq \emptyset$ and (c) follows by observing that the values of Jv at the boundary only depend on the values of $\gamma_1(v)$. The details of the construction are comprehensively studied in [GS02, Section 5 - 7] where the bounds (d)-(e) follow. By construction (i.e. the choice of the weights), the identity $M_k = \Lambda_k$ holds on the space of quintic polynomials $P_5(\Omega)$. □

With this quasi-interpolation operator, discrete versions of the extension operator, theorem 2.4, and boundary oscillations, theorem 4.3 follow.

Corollary 5.2. *Given a triangulation $\mathcal{T}$ with edges $\mathcal{E}$, there exists a discrete right-inverse $\tilde{E} : Tr(\mathcal{T}) \to \mathscr{A}(\mathcal{T})$ of the trace map γ_1 such that $\gamma_1 \circ \tilde{E} = id$ on $Tr(\mathcal{T})$ and*

$$\|\tilde{E}(\varphi_j, \psi_j)_j\|_{H^2(\Omega)} \leq C_{\mathrm{E}} c_{\mathrm{c}} \sum_{j=1}^{N_{\partial\Omega}} (\|\varphi_j\|_{H^{3/2}(\Gamma_j)} + \|\psi_j\|_{H^{1/2}(\Gamma_j)})$$

where $Tr(\mathcal{T}) \coloneqq \gamma_1(\mathscr{A}(\mathcal{T}))$ denotes the Argyris trace space.

Proof. With the extension operator E from theorem 2.4 and the quasi-interpolation J, one immediately observes, that for $\tilde{E} \coloneqq J \circ E$ and any $t = (\varphi_j, \psi_j)_j \in Tr(\mathcal{T})$, cf. [GS02, Corollary 7.7],

$$\left\|\tilde{E}t\right\|_{H^2(\Omega)} \leq c_{\mathrm{c}} \left\|Et\right\|_{H^2(\Omega)} \leq C_{\mathrm{E}} c_{\mathrm{c}} \sum_{j=1}^{N_{\partial\Omega}} (\|\varphi_j\|_{H^{3/2}(\Gamma_j)} + \|\psi_j\|_{H^{1/2}(\Gamma_j)}).$$

As a result of theorem 5.1 (c) and the fact, that Et has trace in $Tr(\mathcal{T})$, so does $\tilde{E}t$. □

Corollary 5.3. *Let $\mathcal{T}$ be a regular triangulation of Ω and $\widehat{\mathcal{T}}$ be an admissible refinement of $\mathcal{T}$. With the boundary interpolant $\mathcal{I}_{\mathrm{b}}$ for $\mathscr{A}(\mathcal{T})$ and $\widehat{\mathcal{I}_{\mathrm{b}}}$ for $\mathscr{A}(\widehat{\mathcal{T}})$, $C_{\mathrm{osc,d}} \coloneqq (1 + c_{\mathrm{b}}) c_{\mathrm{c}} C_{\mathrm{osc}}$,*

$$\min_{\widehat{w} \in (1 - \mathcal{I}_{\mathrm{b}})\widehat{\mathcal{I}_{\mathrm{b}}} v + V(\widehat{\mathcal{T}})} |\widehat{w}|^2_{H^2(\Omega)} \leq C^2_{\mathrm{osc,d}} \mathrm{osc}^2(\mathcal{E}(\partial\Omega) \setminus \widehat{\mathcal{E}}(\partial\Omega), v).$$

Proof. Define $\widehat{t}$ using the same construction for t in the proof of lemma 4.3 with v replaced by $\widehat{\mathcal{I}_{\mathrm{b}}} v$. Because $\mathcal{N}(\partial\Omega) \subset \widehat{\mathcal{N}}(\partial\Omega)$, $\mathcal{I}_{\mathrm{b}}\widehat{\mathcal{I}_{\mathrm{b}}} = \mathcal{I}_{\mathrm{b}}$ and $\widehat{t}$ has the same properties as before. This time, using the discrete extension operator $\tilde{E}$, one gets $\tilde{E}\widehat{t} \in (1 - \mathcal{I}_{\mathrm{b}})\widehat{\mathcal{I}_{\mathrm{b}}} v + V(\widehat{\mathcal{T}}) \subset (1 - \mathcal{I}_{\mathrm{b}})\widehat{\mathcal{I}_{\mathrm{b}}} v + V$ and it is thus possible to apply the continuous result to obtain

$$|\tilde{E}\widehat{t}|^2_{H^2(\Omega)} \leq c_{\mathrm{c}}^2 |E\widehat{t}|^2_{H^2(\Omega)} \leq c_{\mathrm{c}}^2 C^2_{\mathrm{osc}} \mathrm{osc}^2(\mathcal{E}(\partial\Omega), \widehat{\mathcal{I}_{\mathrm{b}}} v) = c_{\mathrm{c}}^2 C^2_{\mathrm{osc}} \mathrm{osc}^2(\mathcal{E}(\partial\Omega) \setminus \widehat{\mathcal{E}}(\partial\Omega), \widehat{\mathcal{I}_{\mathrm{b}}} v).$$

The last equality is obtained since on an unrefined edge $E \in \mathcal{E}(\partial\Omega) \cap \widehat{\mathcal{E}}(\partial\Omega)$ the contributions vanish due to $1 - \Pi_{\mathcal{E},2} \equiv 0$ on $P_2(E) \ni \widehat{\mathcal{I}_{\mathrm{b}}} v_{|E}$. To deduce the claim, observe that for $E \in \mathcal{E}(\partial\Omega) \setminus \widehat{\mathcal{E}}(\partial\Omega)$,

$$\left\|(1 - \Pi_{\mathcal{E},2})\partial^3_{\tau\tau\tau} \widehat{\mathcal{I}_{\mathrm{b}}} v\right\|_{L^2(E)} = \left\|\widehat{\Pi}_{\mathcal{E},2}(1 - \Pi_{\mathcal{E},2})\partial^3_{\tau\tau\tau} v\right\|_{L^2(E)} \leq \left\|(1 - \Pi_{\mathcal{E},2})\partial^3_{\tau\tau\tau} v\right\|_{L^2(E)}$$

and

$$\begin{aligned}\left\|(1-\Pi_{\mathcal{E},2})\partial^3_{\tau\tau\nu}\widehat{\mathcal{I}_{\mathrm{b}}}v\right\|_{L^2(E)} &\leq \left\|\partial^3_{\tau\tau\nu}(1-\widehat{\mathcal{I}_{\mathrm{b}}})v\right\|_{L^2(E)} + \left\|(1-\Pi_{\mathcal{E},2})\partial^3_{\tau\tau\nu}v\right\|_{L^2(E)}\\ &\leq (1+c_{\mathrm{b}})\left\|(1-\Pi_{\mathcal{E},2})\partial^3_{\tau\tau\nu}v\right\|_{L^2(E)}\end{aligned}$$

by a triangle inequality, lemma 4.2 and stability of L^2-projections. Thus, $\mathrm{osc}^2(\mathcal{E}(\partial\Omega)\setminus\widehat{\mathcal{E}}(\partial\Omega),\widehat{\mathcal{I}_{\mathrm{b}}}v)\leq(1+c_{\mathrm{b}})^2\mathrm{osc}^2(\mathcal{E}(\partial\Omega)\setminus\widehat{\mathcal{E}}(\partial\Omega),v)$ and collecting the terms gives the desired result. □

5.3 Optimal convergence

The aforementioned axioms of adaptivity (A1)-(A4) allow for optimal convergence rates as an immediate consequence [Car+14; CR17].

Theorem 5.4 (Rate optimality (AFEM))**.** *With the universal maximum threshold $0<\Theta<1$ there exists for all $0<s<\infty$ some constant Λ_{eq}, which depends only on $\mathbb{T}$ through $\mathcal{T}_0$, on Θ and on s, such that the sequence of triangulations $(\mathcal{T}_\ell)_\ell$ from the AFEM algorithm with $\theta\leq\Theta$ satisfies*

$$\sup_{\ell\in\mathbb{N}_0}(1+|\mathcal{T}_\ell|-|\mathcal{T}_0|)^s\eta(\mathcal{T}_\ell)\leq\Lambda_{\mathrm{eq}}\sup_{N\in\mathbb{N}_0}(1+N)^s\min_{\mathcal{T}\in\mathbb{T}(N)}\eta(\mathcal{T}).$$

The rest of this subsection is devoted to the proof of theorem 5.4, highlighting the modifications that are required for the treatment of inhomogeneous boundary conditions and more general source terms. Recall assumption 5.1 on the source. Axioms (A1) and (A2) prove stability and reduction of the estimator η using standard arguments.

Theorem 5.5. *For any admissible refinement $\widehat{\mathcal{T}}$ of $\mathcal{T}\in\mathbb{T}$, discrete stability and reduction hold true, i.e.*

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

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

The constants Λ_1,Λ_2 only depend on the initial triangulation $\mathcal{T}_0$.

Proof (sketch). These results follow from repeated application of (inverse) triangle inequalities and standard arguments, cf. [CH21] and references therein. One difference to the situation from [CH21] lies in the additional existence of boundary data oscillations. A quick look at the definition verifies that these oscillations do not provide any difficulties here. More precisely, on the unrefined portion of the mesh, $\mathrm{osc}(\mathcal{E},\widehat{\mathcal{E}}\cap\mathcal{E})=\mathrm{osc}(\widehat{\mathcal{E}},\widehat{\mathcal{E}}\cap\mathcal{E})$, and no extra difficulties arise in the treatment of (A1). On the other hand, at the refined portion of the mesh, $\mathrm{osc}(\widehat{\mathcal{E}},\widehat{\mathcal{E}}\cap\mathcal{E})\leq 2^{-3}\mathrm{osc}(\mathcal{E},\widehat{\mathcal{E}}\cap\mathcal{E})$ due to the best-approximation property of the L^2-projection and the fact, that each refined edge $\widehat{E}\in\widehat{\mathcal{E}}\setminus\mathcal{E}$ has length at most half, $|E|/2$, of its parent edge $E\in\mathcal{E}$ with $\widehat{E}\subset E$. The other difference is the additional term $|T|\,\|f_2\|^2_{L^2(T)}$. Since this term is independent of the discrete approximation it cancels out on the right hand side of (A1). Similarly, in (A2), the contribution from the fine triangulation $\widehat{\mathcal{T}}$ is absorbed by the coarse triangulation $\mathcal{T}$ since for $T\in\widehat{\mathcal{T}}\setminus\mathcal{T}$ it holds that $|\widehat{T}|\leq|T|/2$.

□

One important ingredient to the a posteriori error analysis is the reliability of the error estimator. In the context of optimal convergence rates and the axioms of adaptivity one is even interested in discrete reliability.

Theorem 5.6 (Discrete Reliability)**.** *A universal constant Λ_3, only depending on the initial triangulation $\mathcal{T}_0$ exists, such that for any admissible refinement $\widehat{\mathcal{T}}$ of $\mathcal{T} \in \mathbb{T}$,* (A3) *holds, i.e.,*

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

Proof. For the Dirichlet case, the proof is already contained in [CH21] so here the focus lies on necessary modifications to include inhomogeneous boundary conditions. The proof of discrete reliability relies on two triangulations. Let $u_h \in \mathcal{A}(\mathcal{T})$ and $\widehat{u_h} \in \mathcal{A}(\widehat{\mathcal{T}})$ be the discrete solutions to (DWF) based on the triangulations $\mathcal{T}$ and $\widehat{\mathcal{T}}$, respectively. Define the error $\widehat{e} \coloneqq \widehat{u_h} - u_h \in \mathcal{A}(\widehat{\mathcal{T}}) - \mathcal{A}(\mathcal{T}) = \widehat{g_h} - g_h + V(\widehat{\mathcal{T}})$ and note that the equality is due to the nestedness of the extended Argyris space. Because, in general, $\widehat{g_h} - g_h \notin V(\widehat{\mathcal{T}})$, it is not possible to deduce $J\widehat{e} \in V(\mathcal{T})$ and continue as in [CH21] by testing with $J\widehat{e}$. Instead, a split $\widehat{e} = \widehat{e_0} + \widehat{e_\mathrm{b}}$ into $\widehat{e_0} \in V(\widehat{\mathcal{T}})$ and $\widehat{e_\mathrm{b}} \in \mathscr{A}(\widehat{\mathcal{T}})$ allows to proceed as in [CH21] for one term and the other only contributes to boundary ocillations. More precisely, define $\widehat{e_\mathrm{b}}$ to be the biharmonic extension of the boundary data. That is, $\widehat{e_\mathrm{b}} \in \widehat{g_h} - g_h + V(\widehat{\mathcal{T}})$ is defined as the solution to

$$a(\widehat{e_\mathrm{b}}, \widehat{\phi}) = 0 \qquad \text{for all } \widehat{\phi} \in V(\widehat{\mathcal{T}}). \tag{16}$$

An alternative characterisation of $\widehat{e_\mathrm{b}}$ by the calculus of variations is

$$\widehat{e_\mathrm{b}} = \operatorname*{argmin}_{\widehat{\phi} \in \widehat{g_h} - g_h + V(\widehat{\mathcal{T}})} |\widehat{\phi}|^2_{H^2(\Omega)}. \tag{17}$$

Consequently, $\widehat{e_0} \coloneqq \widehat{e} - \widehat{e_\mathrm{b}} \in V(\widehat{\mathcal{T}})$ and hence the Galerkin property with $J\widehat{e_0} \in V(\mathcal{T}) \subset V(\widehat{\mathcal{T}}) \subset H^2_0(\Omega)$ leads to

$$\delta(\widehat{\mathcal{T}}, \mathcal{T})^2 = |\widehat{e}|^2_{H^2(\Omega)} = a(\widehat{e}, \widehat{e}) = L(\widehat{e_0} - J\widehat{e_0}) - a(u_h, \widehat{e_0} - J\widehat{e_0}) + a(\widehat{e}, \widehat{e_\mathrm{b}}).$$

Following [CH21], using integration by parts with $\widehat{v} \coloneqq \widehat{e_0} - J\widehat{e_0}$ twice, cf. (13), shows that

$$\begin{aligned} L(\widehat{v}) - a(u_h, \widehat{v}) =& F(\widehat{v}) - (\Delta^2 u_h, \widehat{v})_{L^2(\Omega)} \\ &+ \sum_{F \in \mathcal{E}} \left(\langle \mathcal{R}_\mathrm{F} u_h, \widehat{v} \rangle_{L^2(E)} - \langle \mathcal{R}_\mathrm{FS} u_h, \partial_\nu \widehat{v} \rangle_{L^2(E)} \right), \end{aligned}$$

where the terms $(t_\mathrm{SF}, \partial_\nu \widehat{v})_{L^2(\Gamma_F \cup \Gamma_S)} + (t_\mathrm{F}, \widehat{v})_{L^2(\Gamma_F)} = L(\widehat{v}) - F(\widehat{v})$ have been absorbed in the definition of the edge residuals.

Recall assumption 5.1, that leads to $f_1 \in L^2(\Omega)$ and $f_2 \in L^2(\Omega)^2$ such that

$$F(\widehat{v}) = (f_1, \widehat{v})_{L^2(\Omega)} + (f_2, \nabla \widehat{v})_{L^2(\Omega)} + \sum_{z \in \mathcal{V}} \beta_z \delta_z(\widehat{v})$$

with $\beta_z \in \mathbb{R}$ and translates the above into

$$\begin{aligned} L(\widehat{v}) - a(u_h, \widehat{v}) =& (f_1 - \Delta^2 u_h, \widehat{v})_{L^2(\Omega)} + (f_2, \nabla \widehat{v}) + \sum_{z \in \mathcal{V}} \beta_z \delta_z(\widehat{v}) \\ &+ \sum_{F \in \mathcal{E}} \left(\langle \mathcal{R}_\mathrm{F} u_h, \widehat{v} \rangle_{L^2(E)} - \langle \mathcal{R}_\mathrm{FS} u_h, \partial_\nu \widehat{v} \rangle_{L^2(E)} \right). \end{aligned}$$

Since, by theorem 5.1 (a), J interpolates exactly at $z \in \mathcal{V}$, so that $\delta_z \circ (1-J) = 0$ is zero for all $z \in \mathcal{V}$ and contributions from the Dirac functionals vanish. Standard arguments as in [CH21], employing Cauchy and trace inequalities as well as the approximation properties of J, theorem 5.1, result in

$$|L(\widehat{v}) - a(u_h, \widehat{v})| \lesssim \sum_{T \in \mathcal{T} \setminus \widehat{\mathcal{T}}} \eta(\mathcal{T}, T) |\widehat{e_0}|_{H^2(\Omega(T))} \lesssim \eta(\mathcal{T}, \mathcal{T} \setminus \widehat{\mathcal{T}}) |\widehat{e_0}|_{H^2(\Omega)}.$$

No contributions arise from $T \in \mathcal{T} \cap \widehat{\mathcal{T}}$ since here $\widehat{v} = \widehat{e_0} - J\widehat{e_0} = 0$.

This shows the first part of the claim, since $\widehat{e_\mathrm{b}}$ is by definition a-orthogonal to V, i.e.,

$$|\widehat{e}|^2_{H^2(\Omega)} = |\widehat{e_0} + \widehat{e_\mathrm{b}}|^2_{H^2(\Omega)} = |\widehat{e_0}|^2_{H^2(\Omega)} + |\widehat{e_\mathrm{b}}|^2_{H^2(\Omega)}$$

by the Pythagoras Theorem and hence $|\widehat{e_0}|_{H^2(\Omega)} \leq |\widehat{e}|_{H^2(\Omega)}$.

The estimation of the second term $a(\widehat{e}, \widehat{e_\mathrm{b}})$ continues with a Cauchy inequality and an application of corollary 5.3 that, by employing (17), verify

$$a(\widehat{e}, \widehat{e_\mathrm{b}}) \leq |\widehat{e_\mathrm{b}}|_{H^2(\Omega)} |\widehat{e}|_{H^2(\Omega)} \leq C_{\mathrm{osc,d}} \mathrm{osc}(\mathcal{E}(\partial\Omega) \setminus \widehat{\mathcal{E}}(\partial\Omega), g) |\widehat{e}|_{H^2(\Omega)}.$$

A combination of the given arguments provides

$$\delta(\mathcal{T}, \widehat{\mathcal{T}})^2 \lesssim \eta(\mathcal{T}, \mathcal{T} \setminus \widehat{\mathcal{T}}) \delta(\mathcal{T}, \widehat{\mathcal{T}})$$

by observing $\mathrm{osc}(\mathcal{E}(\partial\Omega) \setminus \widehat{\mathcal{E}}(\partial\Omega), g) \leq \eta(\mathcal{T}, \mathcal{T} \setminus \widehat{\mathcal{T}})$ and division by $\delta(\mathcal{T}, \widehat{\mathcal{T}})$, if positive, proves the existence of Λ_3. □

Next is the axiom of quasi-orthogonality $(A4)$ which is provided by proving the weakened version with an epsilon, $(\mathrm{A4}_\varepsilon)$. Together with axioms (A1) and (A2) the weakened quasi-orthogonality immediately leads to (A4), cf. [CR17].

Lemma 5.7 (quasi-orthogonality ($\varepsilon > 0$))**.** *The sequence of triangulations $(\mathcal{T}_\ell)_\ell$ of successive NVB refinements from the AFEM algorithm satisfies axiom $(A4_\varepsilon)$, i.e., for all $\varepsilon > 0$ there exists a constant $\Lambda_4(\varepsilon)$ such that for every $m, n \in \mathbb{N}_0$, the following holds*

$$\sum_{\ell=m}^{m+n} \delta^2(\mathcal{T}_{\ell+1}, \mathcal{T}_\ell) \leq \Lambda_4(\varepsilon)\, \eta^2(\mathcal{T}_m) + \varepsilon \sum_{\ell=m}^{m+n} \eta^2(\mathcal{T}_\ell). \tag{$\mathrm{A4}_\varepsilon$}$$

Corollary 5.8 (quasi-orthogonality)**.** *The sequence of triangulations $(\mathcal{T}_\ell)_\ell$ of successive NVB refinements from the AFEM algorithm satisfies axiom $(A4)$, i.e., there exists a constant Λ_4 such that for every $m, n \in \mathbb{N}_0$, the following holds*

$$\sum_{\ell=m}^{\infty} \delta^2(\mathcal{T}_{\ell+1}, \mathcal{T}_\ell) \leq \Lambda_4\, \eta^2(\mathcal{T}_m). \tag{A4}$$

Proof (of lemma 5.7). In the case of homogeneous essential boundary conditions, i.e., $g \equiv 0$, theorem 5.6 directly leads to quasi-orthogonality (A4), cf. [CH21, Theorem 5], using the Galerkin orthogonality, $a(u_L - u_\ell, v_\ell) = 0$ for all $v_\ell \in V(\mathcal{T}_\ell)$ and $L \geq \ell$.

Denote the discrete solution on level ℓ by u_ℓ and edges by $\mathcal{E}_\ell$. Furthermore, abbreviate the test spaces on each level by $V_\ell \coloneqq V(\mathcal{T}_\ell)$. Since, in general $u_{\ell+1} - u_\ell \notin V_{\ell+1}$

special treatment is necessary. With the aforementioned Galerkin orthogonality, a Cauchy inequality and corollary 5.3 show, for any $L \geq \ell$,

$$\begin{aligned} a(u_{L+1} - u_{\ell+1}, u_{\ell+1} - u_\ell) &= \min_{\phi_{\ell+1} \in u_\ell - u_{\ell+1} + V_{\ell+1}} a(u_{L+1} - u_{\ell+1}, \phi_{\ell+1}) \\ &\leq C_{\mathrm{osc,d}} \delta(\mathcal{T}_{L+1}, \mathcal{T}_{\ell+1}) \mathrm{osc}(\mathcal{E}_{\ell+1} \setminus \mathcal{E}_\ell, g). \end{aligned}$$

Since only contributions of $\mathcal{E}_{\ell+1} \setminus \mathcal{E}_\ell$ arise, each edge $E \in \bigcup_{\ell \geq m} \mathcal{E}_\ell$ appears at most once in the n-fold sum, $n \in \mathbb{N}$,

$$\sum_{\ell=m}^{m+n} \mathrm{osc}^2(\mathcal{E}_{\ell+1} \setminus \mathcal{E}_\ell, g).$$

Also each fine edge $\widehat{E} \in \mathcal{E}_\ell, \ell \geq m+1$, is generated by k-times, for some $k \in \mathbb{N}$, bisection of some coarse edge $E \in \mathcal{E}_m$ so that $h_{\widehat{E}} = h_E / 2^k$ and hence its contribution to the oscillation

$$h_{\widehat{E}}^3 \| (1 - \Pi_{\mathcal{E}_\ell, 2}) \partial_{\tau\tau\bullet}^3 g \|_{L^2(\widehat{E})}^2 \leq 2^{-3k} h_E^3 \| (1 - \Pi_{\mathcal{E}_m, 2}) \partial_{\tau\tau\bullet}^3 g \|_{L^2(\widehat{E})}^2, \tag{18}$$

with abbreviation $\partial_{\tau\tau\bullet}^3 = \partial_{\tau\tau\tau}^3$ or $\partial_{\tau\tau\nu}^3$, can be estimated by the oscillations on the coarse level m. The crucial estimate in this proof therefore utilises (18) to bound the aforementioned sum in terms of a geometric series of the oscillations at the coarse level m,

$$\sum_{\ell=m}^{m+n} \mathrm{osc}^2(\mathcal{E}_\ell \setminus \mathcal{E}_{\ell+1}, g) \leq \sum_{\ell=m}^{\infty} 2^{-3(\ell-m)} \mathrm{osc}^2(\mathcal{E}_m, g) = (1 - 2^{-3})^{-1} \mathrm{osc}^2(\mathcal{E}_m, g).$$

As a result, the oscillations decay sufficiently fast and using $\mathrm{osc}^2(\mathcal{E}_m, g) \leq \eta^2(\mathcal{T}_m)$, the generalised young inequality with $\alpha^2 = \varepsilon C_{\mathrm{osc,d}}^{-1} \Lambda_3^{-2} > 0$ and rearranging terms shows,

$$\begin{aligned} \sum_{\ell=m}^{m+n} \delta^2(\mathcal{T}_{\ell+1}, \mathcal{T}_\ell) &= \delta^2(\mathcal{T}_{m+n+1}, \mathcal{T}_m) - 2 \sum_{\ell=m}^{m+n} a(u_{n+m+1} - u_{\ell+1}, u_{\ell+1} - u_\ell) \\ &\leq \Lambda_3^2 \eta^2(\mathcal{T}_m) + C_{\mathrm{osc,d}} \sum_{\ell=m}^{m+n} 2\delta(\mathcal{T}_{n+m+1}, \mathcal{T}_{\ell+1}) \mathrm{osc}(\mathcal{E}_{\ell+1} \setminus \mathcal{E}_\ell, g) \\ &\leq \Lambda_3^2 \eta^2(\mathcal{T}_m) + C_{\mathrm{osc,d}} \sum_{\ell=m}^{m+n} \big(\alpha^2 \delta^2(\mathcal{T}_{n+m+1}, \mathcal{T}_{\ell+1}) + \alpha^{-2} \mathrm{osc}^2(\mathcal{E}_{\ell+1} \setminus \mathcal{E}_\ell, g) \big) \\ &\leq (1 + 8C_{\mathrm{osc,d}}^2 / (7\varepsilon)) \Lambda_3^2 \eta^2(\mathcal{T}_m) + \varepsilon \Lambda_3^{-2} \sum_{\ell=m}^{m+n} \delta^2(\mathcal{T}_{n+m+1}, \mathcal{T}_{\ell+1}). \end{aligned}$$

Since this result is true for any $\alpha > 0$, the claim follows by using theorem 5.6 to bound $\Lambda_3^{-2} \delta^2(\mathcal{T}_{n+m+1}, \mathcal{T}_{\ell+1}) \leq \eta^2(\mathcal{T}_{\ell+1})$ and $\Lambda_4(\varepsilon) \coloneqq (1 + 8C_{\mathrm{osc,d}}^2 / (7\varepsilon)) \Lambda_3^2$. □

This concludes the proof of optimal convergence rates by providing axioms (A1)-(A4), see [Car+14; CR17] for the implication of optimal rates.

6 Numerical results

This section is devoted to the numerical validation of the previous results and compares numerical results with the theoretical expectations. The main focus is on the evaluation of the Argyris FEM for the inhomogeneous biharmonic equation from section 4.4. Both, the classical and hierarchical Argyris FEM, will be examined with respect to convergence, while the focus lies on adaptive FEM and optimal rates as promised by the previous section. Furthermore, the hierarchical Argyris FEM will be utilized for a comparison between a standard direct solving attempt with AFEM, algorithm 4, and an iterative solver based on the multigrid method from section 3. The FMGAFEM, algorithm 3, is steered by the error estimator $\eta(\mathcal{T}_\ell)$ from (15), applied to the approximated solution instead of the (exact) solution on the discrete level.

Even though the Argyris FEM has been known for many years and is featured in many finite element reference literature, e.g., [BS08], it has mostly been treated as an orphan from the analytical viewpoint. The reasons therefore are manifold; firstly, the high order of the quintic polynomials is not only difficult to implement, but also results in a noticeable propagation of numerical errors due to the finite arithmetic. Moreover, the Argyris elements do not form an *affine-interpolation equivalent* family, meaning that the degrees of freedoms on two triangles T_1, T_2 are not preserved under the standard affine mapping $T_1 \to T_2$. Hence, an evaluation of the basis functions on some reference triangle is not enough. Nevertheless, in [DS08; Kir18] ways to compute the transformation – this can be done a priori – from the reference element have been established.

Secondly, the analysis appears troublesome as the corresponding function space misses favourable properties such as the nestedness that has been remedied by introducing the hierarchical Argyris FEM. The experiments start with an introductory benchmark on a convex domain with some smooth known solution. Initial triangulations for the first three benchmarks are shown in figure 3.

6.1 Implementation

The MATLAB implementation is based on the in-house software package AFEM and in particular the octAFEM package from [Bri21]. Though, the expressivity has been severely improved by switching to an object-oriented design (OOP) as in established large-scale finite element software such as, e.g., the FEniCS project [LMW12] and NETGEN/NGSOLVE [Sch14], where some inspiration for the abstract definition of objects was drawn. Thus, the framework allows for treating discrete functions, such as the computed solutions, as dedicated objects for which operations such as differentiation and integration is implemented. Integration is performed by application of a quadrature rule, see [Bri21, Appendix A.6], that is exact for polynomials up to degree 10, when used to compute L^2 quantities of quintic polynomials, or exact up to degree 15 when integrating possibly non-polyomial expressions such as given source f and boundary data g. Hence the *stiffness* and *mass matrices* can be computed exactly, whereas regarding integration of f and g, they are approximated by a higher order polynomial. The resulting data error will be neglected since the approximation is good enough for the benchmarks in the rest of this section.

Data structures for geometric and topological quantities are taken from AFEM and octAFEM. An additional data structure that keeps track of the parent entities for each entity (element, edge or node) of a triangulation allows to traverse in the refinement history. This makes the evaluation of a discrete function $v_\ell \in V(\mathcal{T}_\ell)$, defined on level ℓ, on refinements $\widehat{\mathcal{T}}_\ell$ of $\mathcal{T}_\ell$ possible and allows the interpolation of v_ℓ onto finer spaces $V(\widehat{\mathcal{T}}_\ell)$.

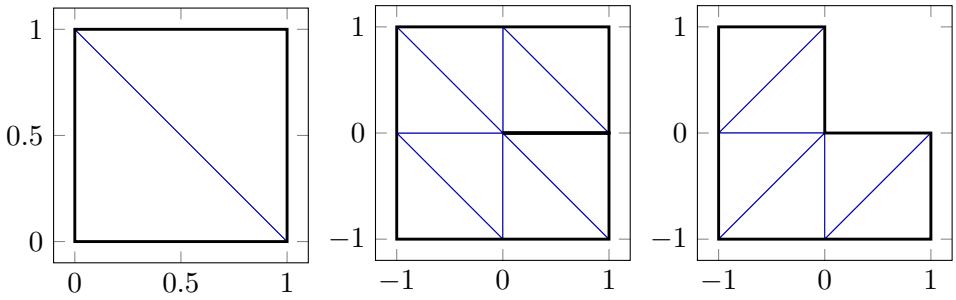

Figure 3: Initial triangulations $\mathcal{T}_0$ for the first three model problems; the unit square (left), the slit domain (middle) and the L-shaped domain (right).

Consequently, such an abstract setting allows for the immediate computation of the quantities that form the error estimator η from (15), the exact error $|u - u_\ell|_{H^2(\Omega)}$ if the exact solution u is available or an estimation thereof by computing $|\widehat{u} - u_\ell|_{H^2(\Omega)}$ for an approximation $\widehat{u}$ of u. This method will be applied in the last benchmark, where no exact solution is available – the approximation $\widehat{u}$ is obtained by performing further adaptive refinements and solving the problem on an even finer grid.

Computation of the stiffness matrix $\tilde{A}_\ell$, right-hand side $\tilde{F}_\ell$ and the other multigrid operators from section 3 follows the lines of Bramble [Bra93, Section 10]. Note, that as motivated in the reference; if A_ℓ, B_ℓ and F_ℓ are the matrix, resp. basis, representations of $\mathbb{A}_\ell$, $\mathbb{B}_\ell$ and $\mathbb{F}_\ell$, from section 3, then $\tilde{A}_\ell = \tilde{M}_\ell A$, $\tilde{F}_\ell = \tilde{M}_\ell F_\ell$ and $\tilde{B}_\ell \tilde{A}_\ell = B_\ell A_\ell$ with mass matrix $\tilde{M}_\ell$. As a consequence, the algebraic error estimator is computed as

$$\eta_{\mathrm{alg}}(\mathcal{T}_\ell, v_\ell) = (\tilde{A}_\ell \tilde{v}_\ell - \tilde{F}_\ell)^\top \tilde{B}_\ell (\tilde{A}_\ell \tilde{v}_\ell - \tilde{F}_\ell)$$

for any $v_\ell \in \mathscr{A}(\mathcal{T}_\ell)$ with basis representation $\tilde{v}_\ell$ without the need of an explicit integration. Additionally, the smoothing operator $\mathbb{S}_\ell$ is represented by the inverse of the lower triangular part of the submatrix of $\tilde{A}_\ell$, corresponding to the subspace $\tilde{V}(\mathcal{T}_\ell) \subset V(\mathcal{T}_\ell)$, cf. [BP92].

This results in two classes of solvers, a direct solver that computes u_ℓ by the MATLAB `mldivide` (commonly known as the \-operator) in algorithm 4 and a multigrid solver that computes u_ℓ^ε with the full multigrid algorithm 3. Furthermore, computation of $\|\mathbb{I} - \mathbb{B}_\ell \mathbb{A}_\ell\|_a$, when $\mathbb{B}_\ell$ is L^2-symmetric, e.g., defined by the $V(r,r)$-cycle for $r \in \mathbb{N}$, is possible via transformation into an eigenvalue problem. Leaving the details to the reader, this works by observing that, since $\mathbb{I} - \mathbb{B}_\ell \mathbb{A}_\ell$ is a-symmetric,

$$\|\mathbb{I} - \mathbb{B}_\ell \mathbb{A}_\ell\|_a = \sup_{v_\ell \in V(\mathcal{T}_\ell)} \frac{a((\mathbb{I} - \mathbb{B}_\ell \mathbb{A}_\ell) v_\ell, v_\ell)}{a(v_\ell, v_\ell)} = \max_{\tilde{v}_\ell \in \mathbb{R}^{N_\ell}} \frac{\tilde{v}_\ell^\top \tilde{A}_\ell (I - \tilde{B}_\ell \tilde{A}_\ell) \tilde{v}_\ell}{\tilde{v}_\ell^\top \tilde{A}_\ell \tilde{v}_\ell}$$

The last term involves so-called Rayleigh quotients and is maximized by the largest eigenvalue to the eigenvalue problem $(I - \tilde{B}_\ell \tilde{A}_\ell)\tilde{v}_\ell = \lambda \tilde{v}_\ell$ which is solved by MATLAB `eigs`.

The general approach for the implementation of the finite element method in the OOP context has been very well understood and a variety of different realizations are being used, so the reader is referred to the literature cited in this subsection and references therein.

6.2 An introductory benchmark on the unit square

Consider the purely clamped plate problem on the unit square domain $\Omega = (0,1)^2$ with source term $\Delta^2 u = f$ chosen to fit the exact solution

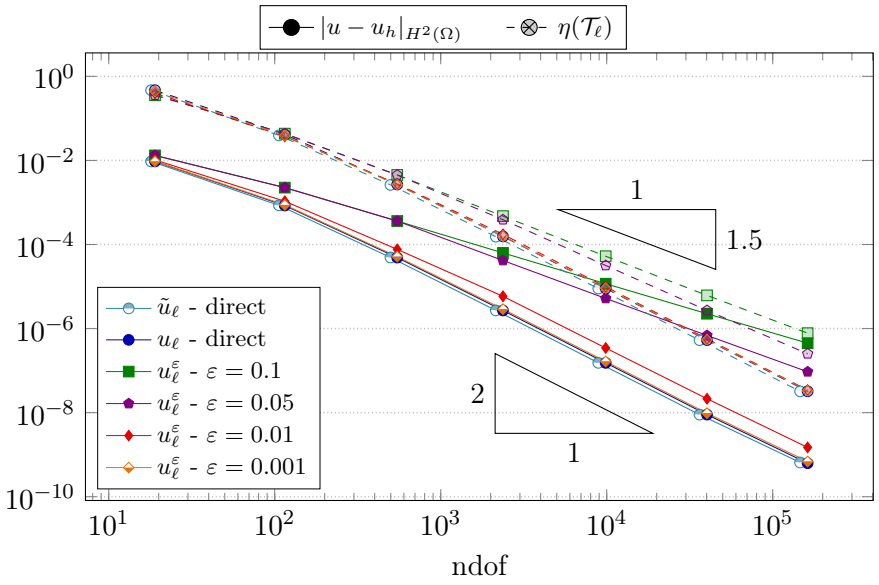

Figure 4: Convergence rates of the energy error $|u - u_h|_{H^2(\Omega)}$ (filled) and the estimated error η (opaque) for $u_h = \tilde{u}_\ell, u_\ell$ or u_ℓ^ε. The FMGAFEM employs the $V(1.1)$-cycle with stop criterion ε.

$$u(x, y) = x^2(1 - x)^2 y^2(1 - y)^2.$$

Purely clamped means in this notation that $\Gamma_C = \partial\Omega, \Gamma_S = \Gamma_F = \emptyset$ and the boundary data $u_D \equiv 0$ is chosen to match the exact solution. As the solution to this model problem is smooth, there is no notable difference between uniform and adaptive refinement. Hence, the discussion will focus on the uniform case. The solutions obtained with AFEM for the classical Argyris FEM are denoted by $\tilde{u}_\ell$ and solutions of the hierarchical Argyris FEM by u_ℓ for AFEM and u_ℓ^ε for FMGAFEM with different stopping criterions ε. The first observation in figure 4 is the confirmation of optimal convergence rates in both the energy norm $|u - u_h|_{H^2(\Omega)}$ and the error estimator η with direct solve, $u_h = \tilde{u}_\ell$ or u_ℓ. For the hierarchical Argyris FEM also the convergence rates of $u_h = u_\ell^\varepsilon$ using the iterative multigrid solver are plotted. For example, for $\varepsilon = 0.001$, no difference to direct solve is visible as these graphs overlap. Different choices of stopping criteria show that a correct choice here is crucial in order to not only obtain a good approximation of the discrete solution but also to obtain a reliable estimation of the error $|u - u_\ell^\varepsilon|_{H^2(\Omega)}$ by η. Setting the stopping criteria too high, shown in the case $\varepsilon = 0.1$, the error estimator η does not behave equivalent to the error. The cause of this behaviour lies in the fact that, e.g., for reliability of η, cf. theorem 5.6, a Galerkin orthogonality is used that only holds for the exact solution. Hence, the divergence for ε too big. Conversely, if ε is chosen small, the approximation u_ℓ^ε appears close enough for η to produce reliable results.

The convergence plots are only shown for the first seven iterations (excluding 0) as the condition of the problem increases in the order $\mathcal{O}(h^{-4}) \equiv \mathcal{O}(\text{ndof}^2)$. This leads to dominating numerical errors for higher discretisations and results eventually in divergence (not displayed). Whether this is an obstacle in the practical sense, may be debatable since this occurs only on sufficiently small meshes and the computed error squared is already close to machine precision (recall that η^2 is computed as a sum of local contributions and thus $\eta \approx 10^{-8}$ if and only if $\eta^2 \approx 10^{-16}$ which is already machine precision).

6.3 Reentrant corner on the L-shaped domain

The second model benchmark considers the L-shaped domain $\Omega = (-1,1)^2 \setminus [0,1]^2$ with homogeneous clamped boundary condition $g \equiv 0$ on $\Gamma_C = \partial\Omega$. Choose the exact singular solution to be (in polar coordinates)

$$u(r,\varphi) = (r^2\sin(\varphi)^2 - 1)^2(r^2\cos(\varphi)^2 - 1)^2 r^{1+\mu}\xi\left(\varphi - \frac{\pi}{2}\right), \quad \text{where}$$

$$\begin{aligned}\xi(\phi) = &\left(\frac{1}{\mu-1}\sin((\mu-1)\omega) - \frac{1}{\mu+1}\sin\left((\mu+1)\omega\right)\right)\left(\cos((\mu-1)\phi) - \cos\left((\mu+1)\phi\right)\right)\\ &- \left(\frac{1}{\mu-1}\sin\left((\mu-1)\phi\right) - \frac{1}{\mu+1}\sin\left((\mu+1)\phi\right)\right)\left(\cos\left((\mu-1)\omega\right) - \cos((\mu+1)\omega)\right)\end{aligned}$$

with interior angle $\omega = 3\pi/2$ at the singularity and solution $\mu \approx 0.54448$ to $\sin^2(\mu\omega) = \mu^2\sin^2(\omega)$. This is in fact the *singular solution* from Grisvard [Gri92] for the singular point $(0,0)$.

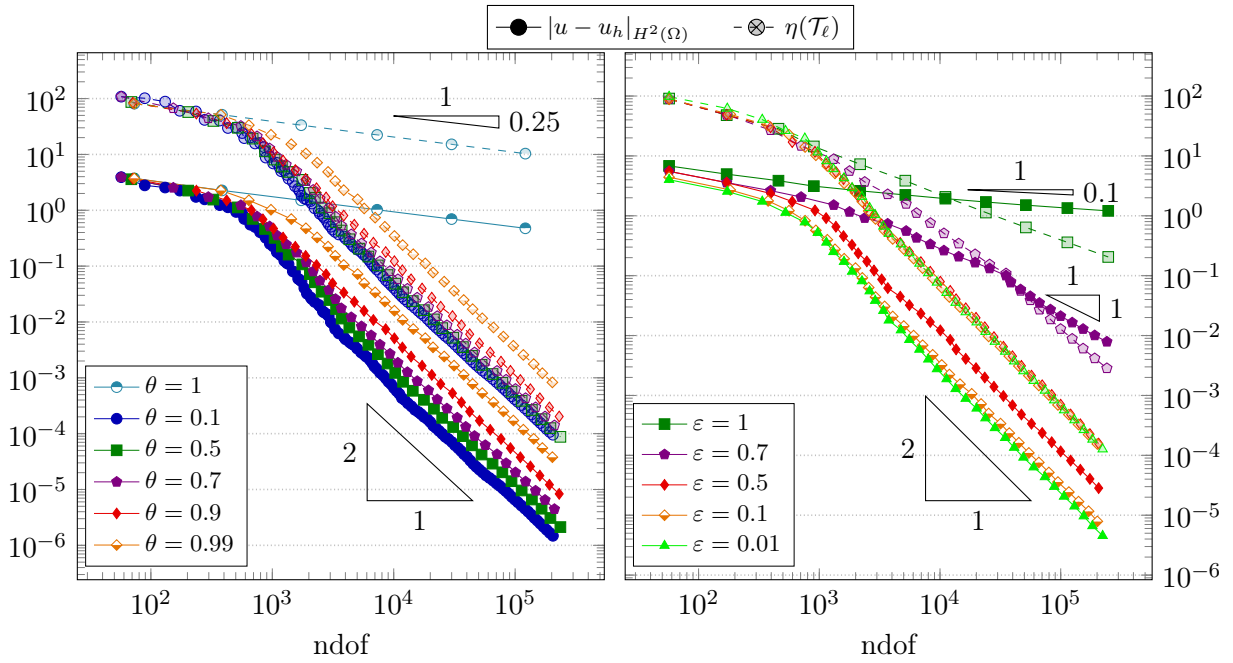

Figure 5: Convergence rates of the energy error $|u-u_h|_{H^2(\Omega)}$ (filled) and the estimated error η (opaque) for AFEM (left), $u_h = u_\ell$, and FMGAFEM (right), $u_h = u_\ell^\varepsilon$, $V(1,1)$-cycle.

The reduced regularity of the solution u leads to an observed convergence rate of order $\mathcal{O}(\mathrm{ndof}^{-1/4})$ for uniform refinement with respect to the mesh size h. As already underlined in the first benchmark, AFEM for the classical Argyris FEM and the hierarchical Argyris FEM behave qualitatively identical, in this and the following benchmarks the focus lies on the hierarchical Argyris FEM. Figure 5 shows the recovery of optimal convergence rates with the adaptive mesh refinement algorithm for both, direct solve and multigrid with proper stopping criterion. The FMGAFEM for this benchmark employs the bulk parameter $\theta = 0.7$. In case of minimal/trivial stopping criteria $\varepsilon = 1$, exactly one multigrid iteration is performed at each level. Even though the convergence of η appears better compared to AFEM with $\theta = 1$ (uniform refinement), the convergence in the energy error

		iterations				
cycle	$\|\mathbb{I} - \mathbb{B}_\ell \mathbb{A}_\ell\|_a$	$\varepsilon = 1$	$\varepsilon = 0.7$	$\varepsilon = 0.5$	$\varepsilon = 0.1$	$\varepsilon = 0.01$
$V(1,1)$	0.96	1	5	8	12-13	20
$V(2,2)$	0.91	1	3	4	6	9-10
$V(5,5)$	0.80	1	2	2	3	5

Table 1: Asymptotical behaviour of FMGAFEM employing the $V(r,r)$-cycle for $r = 1, 2, 5$. The table shows the number of iterations for different stopping criterion ε.

is actually worse in this case. This is further examined in the next benchmark, where the same behavior is observed. Nevertheless, it is already clear that sufficient accuracy for the discrete problem is required in order to obtain a working scheme.

The number of iterations for the FMGAFEM is shown in table 1, together with the reduction factor $\|\mathbb{I} - \mathbb{B}_\ell \mathbb{A}_\ell\|_a$ for different number of smoothing steps in the V-cycle, cf. algorithm 1. The rule of thumb that the number of iterations for a given accuracy is about antiproportional to the number of smoothing iterations in the V-cycle, seems to hold. Smoothing twice results in half as many iterations as for the $V(1,1)$-cycle with a single smoothing step.

It is worth noting that for AFEM huge values for the bulk parameter θ are viable. Even for $\theta = 0.99$ optimal rates are observed. This is unusual for lowest-order methods and the cause can be expected to relate to the relatively high polynomial degree of the hierarchical Argyris space, which causes the error contributions to practically vanish away of the singularity. A further observation concerns the refinement steps performed. For $0.5 \leq \theta \leq 0.99$, AFEM produces roughly the same number of iterations whereas for $\theta = 0.1$ about 3 times as many iterations are necessary.

6.4 Inhomogeneous boundary conditions on the slit domain

In this benchmark problem, the so called slit domain $(0,1)^2 \setminus (\{0\} \times [0,1))$ is considered with pure clamped boundary $\Gamma_C = \partial\Omega$. This is a classical non-Lipschitz domain used to illustrate that the analytical results extend beyond the Lipschitz case. In fact no critical result such as, e.g., extension operators, have been applied in the analysis so that analytical arguments apply here as well. Boundary data $g = u$ and source $f \equiv 1$ are chosen to match the exact solution (in polar coordinates)

$$u(r,\varphi) = -\frac{r^2}{16}\left(r^{1/2}\sin(\varphi/2) - \frac{r^2}{2}\sin^2(\varphi)\right).$$

A computation reveals $u \in C^2(\overline{\Omega}) \setminus C^3(\overline{\Omega})$ due to a singularity at the origin. Hence, the boundary interpolant, definition 4.7, is well-defined for u and the discrete boundary data $g_h = \mathcal{I}_\mathrm{b} u$ can be obtained by interpolation. (Recall from section 2 that $u \in C^k(\overline{\Omega})$ means that $u \in C^k(\Omega)$ and the derivatives of order up to k can be continuously extended up to the boundary $\partial\Omega$.)

Remark. When using a symbolic framework for the computation of $\mathcal{I}_\mathrm{b} u$, the singularity at the origin may result in a division by zero for the evaluation of the hessian of u at the origin. Hence, it may be useful to observe that u vanishes identically on $\{0\} \times [0,1]$ and the gradient and hessian vanish at the origin as well.

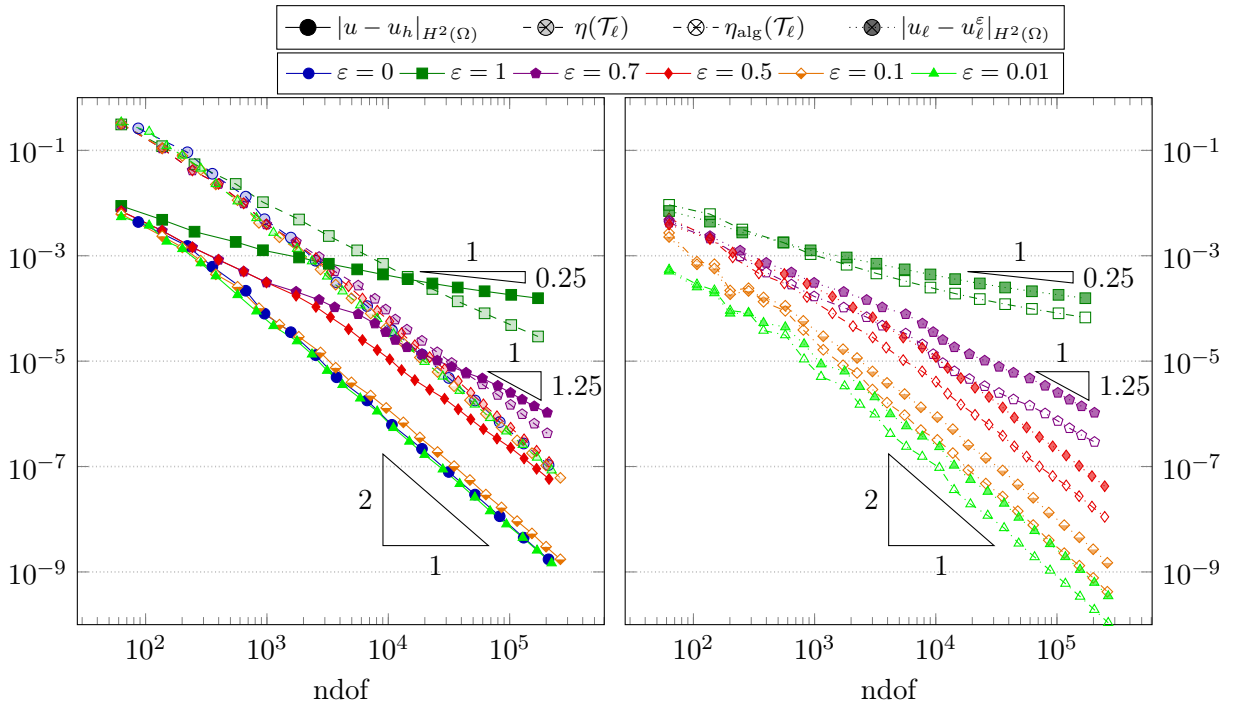

Figure 6: Convergence rates of different error quantities for different ε with $u_\ell^0 := u_\ell$. The energy error $|u - u_\ell^\varepsilon|_{H^2(\Omega)}$ (filled) and its estimator η (opaque) on the left. The algebraic error $|u_\ell - u_\ell^\varepsilon|_{H^2(\Omega)}$ (filled) and the algebraic error estimator η_{alg} (opaque) on the right.

Here, the inhomogeneous boundary data is not given by a piecewise polynomial and thus cannot be exactly resolved by an Argyris function. This produces an additional layer of complexity and leads to boundary data ocillations as discussed in section 4.

As before, the singularity results in decreased convergence of order $\mathcal{O}(\mathrm{ndof}^{-2/3})$ with uniform refinement, which is not as bad as before (undisplayed). In other words, the singularity in this problem can be interpreted as weaker compared to the previous benchmark on the Lshape. One reason for this is the required additional regularity $u \in C^2(\overline{\Omega})$ that allows $\mathcal{I}_\mathrm{b} u$ to be well defined.

Concerning the convergence, cf. figure 6, the same conclusions can be drawn as in the foregoing benchmark on the Lshape such as recovery of optimal rates for AFEM (with $\theta = 0.8$) shown on the left with the understanding $u_\ell^0 := u_\ell$. An interesting observation provides the cases when ε is too large, e.g., $\varepsilon = 1$ (exactly one multigrid iteration each level), as was already observed in the previous benchmark. In this case, there are too few iterations to obtain a meaningful approximation of the actual solution. Because the error estimator η requires the true solution (as it builds on the Galerkin property), meaningfulness of η decreases significantly as the approximation of the true solution reduces. In fact, for this extreme case, the convergence of η may not look too bad as it is better than for the uniform case. On the other hand, the true error is not at all reflected by the estimator and is actually worse than for uniform refinement. This leads to the conclusion that AFEM algorithm refines less towards the singularity than towards the rest, rendering the algorithm inefficient. This shows that a good approximation is mandatory for equivalence of $|u - u_h|_{H^2(\Omega)}$ and η.

	$\varepsilon = 1$	$\varepsilon = 0.7$	$\varepsilon = 0.5$	$\varepsilon = 0.1$	$\varepsilon = 0.01$
$\lvert u_\ell - u_\ell^\varepsilon\rvert_{H^2(\Omega)} / \lvert u - u_\ell\rvert_{H^2(\Omega)}$	1.00	1.00	0.99	0.82	0.31
$\eta_{\text{alg}} / \lvert u_\ell - u_\ell^\varepsilon\rvert_{H^2(\Omega)}$	0.42	0.28	0.26	0.28	0.36

Table 2: Estimates of ratios between error measures for the FMGAFEM employing the $V(1,1)$-cycle.

		iterations				
cycle	$\lVert \mathbb{I} - \mathbb{B}_\ell \mathbb{A}_\ell \rVert_a$	$\varepsilon = 1$	$\varepsilon = 0.7$	$\varepsilon = 0.5$	$\varepsilon = 0.1$	$\varepsilon = 0.01$
$V(1,1)$	0.93	1	4	6-7	7-8	6-7
$V(2,2)$	0.85	1	2-3	3	3-4	3-4
$V(5,5)$	0.72	1	2	2	2	2

Table 3: Asymptotical behaviour of FMGAFEM employing the $V(r,r)$-cycle for $r = 1, 2, 5$. The table shows the number of iterations for different stopping criterion ε.

An interesting observation concerns the convergence of $|u - u_\ell^{0.5}|_{H^2(\Omega)}$, which starts sub-optimal but improves to be seemingly optimal for ndof $> 10^5$. Conversely, the estimators η for $\varepsilon \leq 0.5$ behave almost the same and are not distinguishable.

Regarding the FMGAFEM algorithm, the energy error $|u - u_\ell^\varepsilon|_{H^2(\Omega)}$ can be roughly split into the discretisation error $|u - u_\ell|_{H^2(\Omega)}$ and the algebraic error $|u_\ell - u_\ell^\varepsilon|_{H^2(\Omega)}$. Figure 6 on the right compares the algebraic error with the algebraic error estimator η_{alg} and shows equivalence of both error measures. This is expected as a consequence of theorem 4.5, cf. corollary 3.6. Clearly, optimal overall convergence implies convergence of the algebraic error. Conversely, the behaviour for $\varepsilon = 0.5$ with optimal convergence of the algebraic error and (seemingly) optimal convergence of the energy error leads to the impression that this might as well be sufficient or at least enough for η to produce a meaningful estimation of $|u - u_\ell^\varepsilon|_{H^2(\Omega)}$. The conclusion is, that while η alone does not allow good judgement of the actual convergence of the energy error $|u - u_\ell^\varepsilon|_{H^2(\Omega)}$, in combination with η_{alg} some information about the overall convergence can be drawn. Table 2 shows that for $\varepsilon \geq 0.5$ the algebraic error dominates the energy error and gives the ratio $\eta_{\text{alg}} / |u_\ell - u_\ell^\varepsilon|_{H^2(\Omega)}$, obtained from the last iteration displayed in figure 6. Equivalence of η_{alg} with $|u_\ell - u_\ell^\varepsilon|_{H^2(\Omega)}$ means that this ratio can be used to calibrate the stopping criterion ε.

For completeness, the number of iterations of the FMGAFEM per level ℓ is given in table 3. Notable is, that the $V(5,5)$-cycle does not lead to much of an improvement over the other cycles as a minimum of 2 iterations is performed (except for $\varepsilon = 1$). Hence, shifting the computation into the V-cycle by raising the number of smoothing steps does not need to reduce the number of iterations by the same amount. Conversely, the $V(5,5)$-cycle for $\varepsilon = 0.5$ produces the same computations than for $\varepsilon = 0.01$, at least asymptotically. Another notable observation concerns the behaviour of the $V(1,1)$-cycle for $\varepsilon = 0.01$, where less iterations are necessary than for $\varepsilon = 0.1$. It turns out, that, from the first iterate on, $\varepsilon = 0.01$ leads to a significantly lower algebraic error than for $\varepsilon = 0.1$, cf. figure 6, which initially only performs about 2 iterations.

6.5 Multiple singularities and point load

In this final benchmark the situation at hand is given by the motivation of considering the biharmonic equation in the context of plate bending. Consider a quadratic floor in

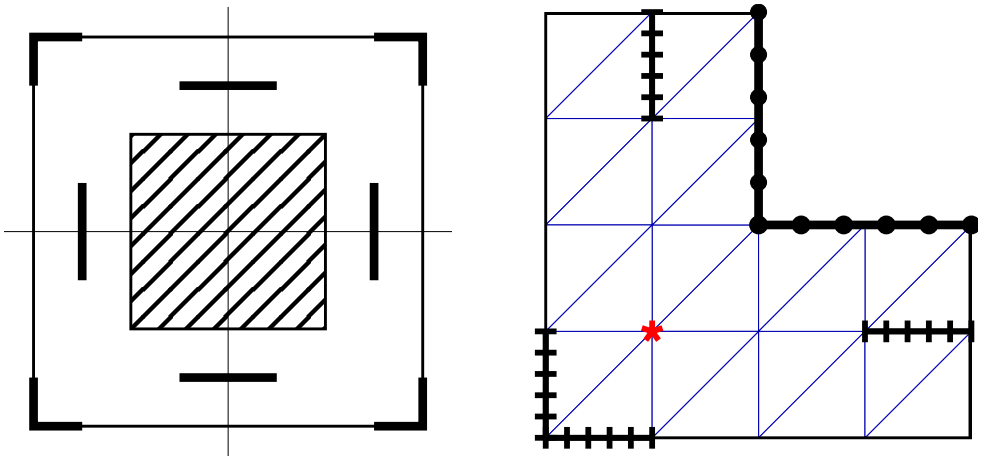

Figure 7: Layout of the setting (left) and initial triangulation $\mathcal{T}_0$ for the discretisation of the lower left quarter. The boundary conditions are given by Γ_C (●—), Γ_S (+), and Γ_F (—). The support of the point force is located at the red star (✱).

some building, e.g., skyscraper, made out of reinforced concrete, depicted in figure 7. The core of the building carries all the weight, so suppose that the floor is embedded into the skeleton on a central square and thus immovable around the center. Furthermore there are supports that hold the floor in place but do not fix tilting. Since the layout is symmetric, the further considerations are reduced to the lower left quarter so that the domain of consideration is again the L-shaped domain $\Omega = (-1,1)^2 \setminus [0,1]^2$. Displacement and bending is prescribed on $\Gamma_C = \{0\}\times[0,1)\cup[0,1)\times\{0\}$ whereas only displacement is fixed on $\Gamma_S = \{-1\}\times[1/2,1]\cup[-1,-1/2]\times\{-1\}\cup\widetilde{\Gamma_S}$, shown in figure 7 on the right. This introduces a break with the convention $\Gamma_S \subseteq \partial\Omega$, since $\widetilde{\Gamma_S} = \{-1/2\}\times[1/2,1]\cup[1/2,1]\times\{-1/2\} \not\subset \partial\Omega$ lies inside the domain, $\operatorname{int}(\widetilde{\Gamma_S}) \subset \Omega$. Hence this problem does not fit into the setting from (SF) – it can be understood in the form (SF) with the additional constraint that the displacement is fixed on $\widetilde{\Gamma_S}$.

But, since the fact $\Gamma_C, \Gamma_S, \Gamma_F \subseteq \partial\Omega$ never played a role in the preceding analysis other than to imply $\Gamma_C, \Gamma_S, \Gamma_F$ being of co-dimension greater or equal than 1, this problem is perfectly realised by (WF) and the results from the previous sections immediately generalise to this and similar situations.

For this benchmark a point load $F = \delta_z \in H^{-2}(\Omega)$ at $z = (-1/2, -1/2)$ leads to a situation with no known solution u to the biharmonic equation

$$\Delta^2 u = F$$

with homogeneous boundary conditions, $g \equiv 0$. Note, that the methods from this thesis allow for immediate application to $g \in B$.

Different boundary conditions and the fact that $F \notin H^{-1}(\Omega)$ (but $F \in H^{-1-\varepsilon}(\Omega)$ for $0 < \varepsilon < 1$) lead to a problem with multiple singularities. A priori mesh generation is not at all clear due to the existence of multiple singularities, which require different amounts of local refinement. Figure 8 shows that the estimator η from (15) localizes the error correctly as AFEM recovers optimal convergence rates, cf. figure 9. In the convergence graph, a clear pre-asymptotic regime up to about 10^3 ndof (also depending on the bulk parameter θ) can be observed. This matches the observation that $\mathcal{T}_{14}$, produced by AFEM with $\theta = 0.1$, has only slightly improved resolution around the singularity, whereas $\mathcal{T}_{38}$ from the same

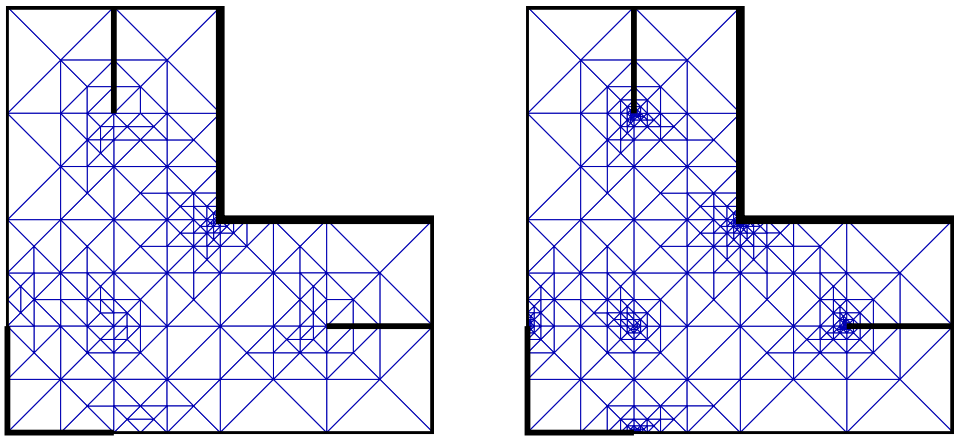

Figure 8: Output triangulations $\mathcal{T}_{14}$ with 1019 ndof (left) and $\mathcal{T}_{38}$ with 3325 ndof (right) of the AFEM algorithm with $\theta = 0.1$.

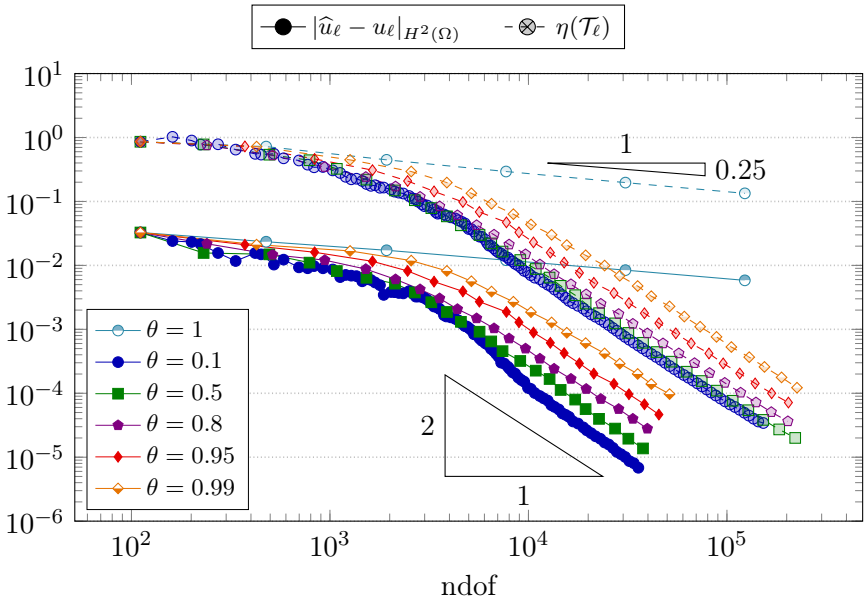

Figure 9: Convergence plot of the approximate energy error $|\widehat{u}_\ell - u_\ell|_{H^2(\Omega)}$ (filled) and η (opaque) for AFEM with different bulk parameter θ.

algorithm already shows clear local refinement towards these singularities. In fact, large parts, away from the singularities, were not further refined at all.

In order to compare the estimation from η with the energy error $|u - u_\ell|_{H^2(\Omega)}$, even though the exact solution u is unavailable, an approximation $\widehat{u}_\ell$ of u is used to get an approximative answer. The approximation $\widehat{u}_\ell$ on $\mathscr{A}(\widehat{\mathcal{T}}_\ell)$ is computed by the AFEM algorithm, starting from $\mathcal{T}_\ell$, on a mesh $\widehat{\mathcal{T}}_\ell$ with 1.5 as many ndof as for $\mathcal{T}_\ell$. Due to the additional overhead, this can be understood as an overkill approach, but turns out to be enough to derive the desired information. The choice of the factor 1.5 has to be understood as a heuristic since even though a higher resolution leads to higher accuracy, numerical errors due to the high extra refinement dominate earlier. This is also the reason why convergence of $|\widehat{u}_\ell - u_\ell|_{H^2(\Omega)}$ is only displayed until about $4 \cdot 10^4$ ndof; on finer meshes this approach does not yield usable results. Nevertheless, it can be seen that the estimator η converges further. Hence, the conclusion from this benchmark is that AFEM is able to recover optimal rates for various complex problems and therefore has to be seen as the method of choice.

Similar to previous observations, huge values for the bulk parameter θ are possible, even though for $0.5 \leq \theta \leq 0.99$, AFEM roughly perform thes same number of iterations. Only the very pre-asymptotic regime ($\text{ndof} < 2 \cdot 10^3$) shows significantly less iterations for high bulk parameter. Convergence of the estimator η is of optimal order for all presented values of θ. Regarding the (approximated) energy error, the situation is not so clear, as $\theta = 0.99$ seems to converge slightly suboptimal with rate $\mathcal{O}(\text{ndof}^{-1.9})$. It is not clear whether this is a pre-asymptotic behaviour or not. Nevertheless, $\theta \leq 0.95$ apparently produces optimal rates, already in the given segment. The approach of approximating $|u - u_\ell|_{H^2(\Omega)}$ by $|\widehat{u}_\ell - u_\ell|_{H^2(\Omega)}$ can also be used to obtain estimates on coarser meshes for the over-estimation, also called efficiency index, $EF(\eta, \mathcal{T}_\ell) \coloneqq \eta / |u - u_\ell|_{H^2(\Omega)}$ of η. In this case $EF(\eta, \mathcal{T}_\ell) \approx 30$ for $\theta = 0.8$. The efficiency index can be used to obtain an estimation for the energy error on the finer levels, since the over-estimation is roughly constant through the convergence, due to equivalence.

To wrap things up, a comparison of the time complexity for the direct solver (AFEM) and the iterative solver (FMGAFEM), employing the $V(1,1)$-cycle with the local Gauß-Seidel (GS) relaxation operator/smoother, $\mathbb{S}_\ell$ from definition 3.2, is given. With the interpretation provided in section 3.2, $\mathbb{S}_\ell$ only acts on the refined portion of the mesh. For comparison, the *global Gauß-Seidel relaxation operator*, that is defined by identity (6) with projections $(\mathbb{P}_\ell^j)_j$ for the full set of nodal bases functions and thus acts on the whole space on each level ℓ, is also considered. This is the widely used multiplicative Schwarz operator, cf. [BP92]. Figure 10 compares the CPU time in seconds, measured on an Intel© Xeon© Platinum 8280L CPU at 2.70GHz with 3TB RAM. The bulk parameter is chosen to be $\theta = 0.1$ and leads to many refinement steps, so that the tendency for the convergence rates are clearly visible. As expected, the direct solver is not able to achieve linear rates, even though the aproximated rate $\mathcal{O}(\text{ndof}^{1.3})$ is not much worse. Interestingly, both versions of FMGAFEM, with local and global smoothing, are even slightly better than linear, $\mathcal{O}(\text{ndof})$, but clearly, this has to be seen as pre-asymptotic behaviour. The number of multigrid iterations necessary to archieve optimal rates has been measured as 3, while in the beginning, for $\text{ndof} < 10^4$, even a single iteration was sufficient. This holds for both, local and global GS smoothing. The reason for even a single iteration of multigrid to suffice in this situation is that $\theta = 0.1$ leads to less distance $\delta(\mathcal{T}_\ell, \mathcal{T}_{\ell+1})$ compared to the meshes considered before.

Another observation is that one iteration with global GS smoothing is roughly 3 times

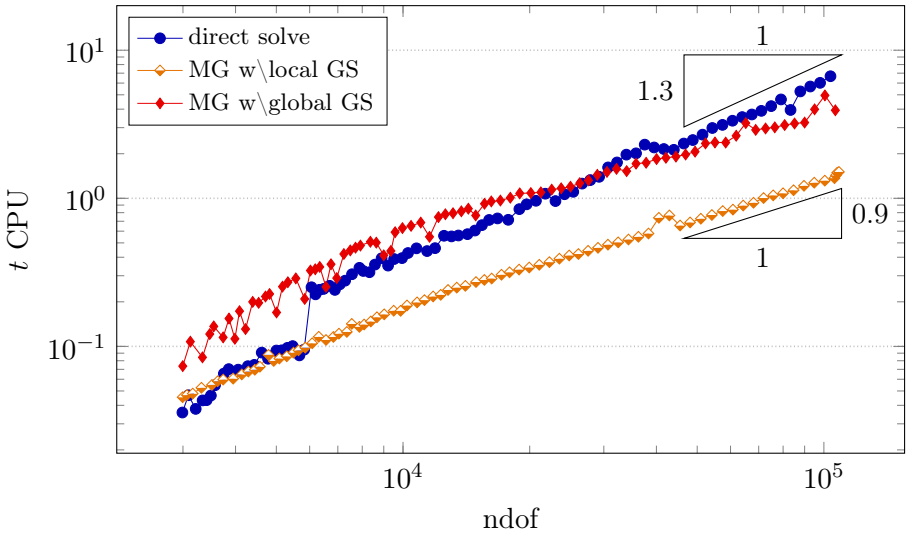

Figure 10: Convergence of CPU time t in seconds for the solution step in AFEM (blue) and one multigrid iteration in FMGAFEM with local GS smoother (yellow) and global GS smoother (red). Measured on Intel© Xeon© Platinum 8280L CPU at 2.70GHz with 3TB RAM.

as expensive as the local version. Since this extra amount of work does not manifest in (asymptotically) less iterations necessary, local GS smoothing has to be seen superior.

6.6 Conclusion

The conclusions to be drawn from this experimental section are manifold. First of all comes a comparison between the classical and the extended Argyris space. Qualitatively, both methods perform identically and no difference could be made out for this set of benchmark problems. Together with the benefit of theoretical results for the hierarchical Argyris FEM, this is the reason, the classical Argyris FEM is undisplayed for the most part. As shown in figure 4, the extended Argyris space has only a few additional degrees of freedom on the same mesh, compared with the classical Argyris space. All in all, the extra work required by the additional degrees of freedom seems negligible, so that the extended Argyris FEM emerges as the method of choice for most situations, where theoretical results are to be applied. Furthermore, the hierarchical Argyris FEM allows for an iterative version of AFEM using multigrid iterations.

The FMGAFEM scheme performs extremely well, with only a few iterations required. It has been shown that the number of iterations also depends on the amount of smoothing iterations in the $V(r,r)$-cycle using the local Gauß-Seidel smoother from section 3.2, that only works on the refined portion of the mesh. Replacing the local Gauß-Seidel smoother with the global version (smoothing on the full set of nodal variables), i.e., from [BP92], does not result in any kind of qualitative improvements (tested on these model problems). The global smoother only leads to insignificantly better approximations, with an asymptotically close number of iterations required.

The question of how much the iterative solver can outperform a direct solver cannot be quantitatively answered with the MATLAB implementation provided, since MATLAB *mldivide* results in a call to a highly optimized low-level procedure. Notably, figure 10 already implies competitive temporal results for the MATLAB implementation at hand.

For quantitative results, also implementing the multigrid solver in a low-level programming language would be required. Furthermore, it is standard practice to use direct solvers for these basic model problems, since their performance is satisfactory for limited ndof and result in only slight suboptimal convergence rates [GN88]. Therefore, only for more complicated problems, the advantages of multigrid will be more clearly visible. Nevertheless, it is shown, by the observations provided, that the multigrid solver works in linear time, since the number of iterations is bounded, cf. figure 10, tables 1 and 3. Also, choosing the local Gauß-Seidel smoother over the global version can result in reduced computational cost.

It was shown that adaptive mesh refinement with an adopted version of the error estimator η from (15) leads to FMGAFEM with very good performance. Also, in conjunction with the reliable and efficient estimator η_{alg} from definition 3.3, cf. lemma 3.5, valuable information about the overall convergence can be drawn. Furthermore, it has been understood that high accuracy of the discrete approximation at the very beginning propagates through subsequent iterations, cf. $\theta = 0.01$ in figure 6 on the right. Interestingly enough, this resulted in slightly less multigrid iterations than for the bigger stopping criterion $\theta = 0.1$, cf. table 3. As expected, lower values of ε are leading to less dominance of the algebraic error in the total error, cf. table 2. But, even if the algebraic error is the mayor contribution of the total error, optimal convergence can still be reached. Of course, convergence of the algebraic error has to be of the same order as the discretisation error. This implies that for η to give a useful refinement indicator, the requirement on the accuracy of the approximation to the exact error is not too demanding for only a few local multigrid iterations.

Other applications of multigrid are also possible. The approximate inverse $\mathbb{B}_\ell$ is also a very popular preconditioner and could possibly be used in conjunction with, e.g., Krylov subspace methods [Saa03]. In this situation, the decrease of $\|\mathbb{I}-\mathbb{B}_\ell\mathbb{A}_\ell\|_a$ for more smoothing iterations, as shown in tables 1 and 3, suggests that increasing smoothing iterations leads to a better preconditioner.

Given $v \in H^2(\Omega)$, the Argyris FEM allows for direct interpolation of $v_{|E}$ and $(\partial_\nu v)_{|E}$ separately along arbitrary edges E of the triangulation as well as point evaluations of v. As shown in the example from section 6.5, this immediately allows to treat certain constrained problems as well.

References

[AFS68] J. H. Argyris, I. Fried, and D. W. Scharpf. "The TUBA Family of Plate Elements for the Matrix Displacement Method". *The Aeronautical Journal* 72.692 (Aug. 1968), pp. 701–709.

[BBP02] C. Bacuta, J. H. Bramble, and J. E. Pasciak. "Shift Theorems for the Biharmonic Dirichlet Problem". In: *Recent Progress in Computational and Applied PDES*. Springer US, 2002, pp. 1–26.

[BC17] S. C. Brenner and C. Carstensen. "Finite Element Methods". In: *Encyclopedia of Computational Mechanics Second Edition*. John Wiley & Sons, Ltd, Dec. 2017, pp. 1–47.

[Bel69] K. Bell. "A refined triangular plate bending finite element". *Int. J. Numer. Meth. Engng.* 1.1 (Jan. 1969), pp. 101–122.

[BFS66] F. K. Bogner, R. L. Fox, and L. A. Schmidt. *The Generation of Inter-Element-Compatible Stiffness and Mass Matrices by the Use of Interpolation Formulas*. Tech. rep. AFFDL-TR-66-80. Nov. 1966, pp. 395–443.

[BH83] D. Braess and W. Hackbusch. "A New Convergence Proof for the Multigrid Method Including the V-Cycle". *SIAM J. Numer. Anal.* 20.5 (Oct. 1983), pp. 967–975.

[BM18] H. Brezis and P. Mironescu. "Gagliardo-Nirenberg inequalities and non-inequalities: the full story". *Annales de l'Institut Henri Poincaré (C) Non Linear Analysis* 35.5 (2018), pp. 1355–1376.

[BP92] J. H. Bramble and J. E. Pasciak. "The analysis of smoothers for multigrid algorithms". *Math. Comp.* 58.198 (May 1992), pp. 467–467.

[Bra+91] J. H. Bramble et al. "Convergence estimates for multigrid algorithms without regularity assumptions". *Mathematics of Computation* 57 (July 1991), pp. 23–45.

[Bra07] D. Braess. *Finite Elements: Theory, Fast Solvers, and Applications in Elasticity Theory*. 3rd ed. Cambridge University Press, 2007.

[Bra77] A. Brandt. "Multi-level adaptive solutions to boundary-value problems". *Math. Comp.* 31.138 (May 1977).

[Bra84] D. Braess. "The Covergence Rate of a Multigrid Method with Gauss-Seidel Relaxation for the Poisson Equation". *Mathematics of Computation* 42.166 (Apr. 1984).

[Bra93] J. H. Bramble. *Multigrid methods*. Pitman research notes in mathematics series 294. Longman Scientific & Technical; Wiley, 1993.

[Bri21] P. Bringmann. "Adaptive least-squares finite element method with optimal convergence rates". PhD thesis. Jan. 2021.

[BS08] S. C. Brenner and L. R. Scott. *The Mathematical Theory of Finite Element Methods*. Vol. 15. Texts in Applied Mathematics. Springer New York, 2008.

[BZ95] J. H. Bramble and X. Zhang. "Multigrid methods for the biharmonic problem discretized by conforming C^1 finite elements on nonnested meshes". *Numerical Functional Analysis and Optimization* 16.7-8 (Jan. 1995), pp. 835–846.

[Car+14] C. Carstensen et al. "Axioms of adaptivity". *Computers & Mathematics with Applications* 67.6 (Apr. 2014), pp. 1195–1253.

[Car19] C. Carstensen. *CPDE I - Lecture notes*. 2019.

[CH21] C. Carstensen and J. Hu. "Hierarchical Argyris Finite Element Method for Adaptive and Multigrid Algorithms". *Computational Methods in Applied Mathematics* 21.3 (July 2021), pp. 529–556.

[Cia02] P. G. Ciarlet. *The Finite Element Method for Elliptic Problems*. Classics in Applied Mathematics. Society for Industrial and Applied Mathematics, Jan. 2002.

[Cia88] P. G. Ciarlet. *Mathematical elasticity*. Studies in mathematics and its applications v. 20, 27, 29. Elsevier Science Pub. Co, 1988.

[Clé75] P. Clément. "Approximation by finite element functions using local regularization". *R.A.I.R.O. Analyse Numérique* 9.R2 (1975), pp. 77–84.

[CM13] C. Carstensen and C. Merdon. "Computational survey on a posteriori error estimators for nonconforming finite element methods for the Poisson problem". *Journal of Computational and Applied Mathematics* 249 (Sept. 2013), pp. 74–94.

[CR17] C. Carstensen and H. Rabus. "Axioms of Adaptivity with Separate Marking for Data Resolution". *SIAM J. Numer. Anal.* 55.6 (Jan. 2017), pp. 2644–2665.

[CT66] R. W. Clough and J. L. Tocher. *Finite Element Stiffness Matrices for Analysis of Plate Bending*. Tech. rep. AFFDL-TR-66-80. Nov. 1966, pp. 515–545.

[DS08] V. Domínguez and F.-J. Sayas. "Algorithm 884: A Simple Matlab Implementation of the Argyris Element". *ACM Trans. Math. Softw.* 35.2 (July 2008), pp. 1–11.

[Eva10] L. C. Evans. *Partial differential equations*. 2nd ed. Graduate studies in mathematics v. 19. OCLC: ocn465190110. American Mathematical Society, 2010.

[GN88] A. George and E. Ng. "On the Complexity of Sparse QR and LU Factorization of Finite-Element Matrices". *SIAM J. Sci. and Stat. Comput.* 9.5 (Sept. 1988), pp. 849–861.

[GO95] M. Griebel and P. Oswald. "On the abstract theory of additive and multiplicative Schwarz algorithms". *Numer. Math.* 70.2 (Mar. 1995), pp. 163–180.

[Gri85] P. Grisvard. *Elliptic Problems in Nonsmooth Domains*. Society for Industrial and Applied Mathematics, 1985.

[Gri92] P. Grisvard. *Singularities in boundary value problems*. Research notes in applied mathematics 22. Masson, 1992.

[GS02] V. Girault and L. R. Scott. "Hermite interpolation of nonsmooth functions preserving boundary conditions". *Math. Comp.* 71.239 (Jan. 2002), pp. 1043–1074.

[Hac16] W. Hackbusch. *Iterative Solution of Large Sparse Systems of Equations*. Vol. 95. Applied Mathematical Sciences. Springer International Publishing, 2016.

[Kir18] R. C. Kirby. "A general approach to transforming finite elements". *The SMAI journal of computational mathematics* 4 (Apr. 2018), pp. 197–224.

[LMW12] *Automated Solution of Differential Equations by the Finite Element Method.* Vol. 84. Lecture Notes in Computational Science and Engineering. Springer Berlin Heidelberg, 2012.

[Mit92] W. F. Mitchell. "Optimal Multilevel Iterative Methods for Adaptive Grids". *SIAM J. Sci. and Stat. Comput.* 13.1 (Jan. 1992), pp. 146–167.

[RS80] M. Reed and B. Simon. *Methods of modern mathematical physics.* Rev. and enl. ed. Academic Press, 1980.

[Saa03] Y. Saad. *Iterative Methods for Sparse Linear Systems.* Second. Society for Industrial and Applied Mathematics, Jan. 2003.

[Sch14] J. Schöberl. "C++11 Implementation of Finite Elements in NGSolve". In: Sept. 2014.

[SZ90] L. R. Scott and S. Zhang. "Finite element interpolation of nonsmooth functions satisfying boundary conditions". *Math. Comp.* 54.190 (1990), pp. 483–493.

[TW87] S. P. Timošenko and S. Woinowsky-Krieger. *Theory of plates and shells.* 2. ed., reissued. Engineering societies monographs. McGraw-Hill, 1987.

[WC06] 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* 49.10 (Oct. 2006), pp. 1405–1429.

[Xu92] J. Xu. "Iterative Methods by Space Decomposition and Subspace Correction". *SIAM Rev.* 34.4 (Dec. 1992), pp. 581–613.

[XZ02] J. Xu and L. Zikatanov. "The method of alternating projections and the method of subspace corrections in Hilbert space". *J. Amer. Math. Soc.* 15.3 (Apr. 2002), pp. 573–597.

[Zha92] X. Zhang. "Multilevel Schwarz methods". *Numer. Math.* 63.1 (Dec. 1992), pp. 521–539.

[Zie+81] O. C. Zienkiewicz et al. "Hierarchical Finite Element Approaches Error Estimates and Adaptive Refinement" (July 1981).