L^2 Best-Approximation of the Elastic Stress in the Arnold-Winther FEM

C. Carstensen * D. Gallistl * M. Schedensack *

Abstract

The first part of this paper enfolds a medius analysis for mixed finite element methods (FEMs) and proves a best-approximation result in L^2 for the stress variable independent of the error of the Lagrange multiplier under the abstract conditions (LBB), condition (C) and efficiency (E). The second part applies the general result to the FEM of Arnold and Winther for linear elasticity: The stress error in L^2 is controlled by the L^2 best-approximation error of the true stress by *any* discrete function plus data oscillations. The analysis is valid without any extra regularity assumptions on the exact solution and also covers coarse meshes and Neumann boundary conditions. Further applications include Raviart-Thomas finite elements for the Poisson and the Stokes problem. The result has consequences for nonlinear approximation classes related to adaptive mixed finite element methods.

Keywords mixed finite element methods, medius analysis, linear elasticity, Arnold-Winther finite element Stokes problem, pseudostress
AMS subject classification 65M12, 65M60, 65N30

1 Introduction

Given bilinear forms $a : X \times X \to \mathbb{R}$ and $b : X \times Y \to \mathbb{R}$ in (real) Hilbert spaces X and Y and a given right-hand side $(G, F) \in X^\star \times Y^\star$, the typical mixed formulation seeks $(x, y) \in X \times Y$ with

$$\begin{aligned} a(x,\xi) + b(\xi, y) &= G(\xi) \qquad &&\text{for all } \xi \in X, \\ b(x,\eta) &= F(\eta) \qquad &&\text{for all } \eta \in Y. \end{aligned} \tag{1.1}$$

The discretisation of this formulation is preferred if constraints have to be enforced, but also if the main interest is in a proper approximation of the stress variable $x \in X$ [Bra07, BF91].

*Institut für Mathematik, Humboldt-Universität zu Berlin, Unter den Linden 6, D-10099 Berlin, Germany
This work was supported by the DFG Research Center Matheon and the Berlin Mathematical School.

Let $X \subseteq H$ be embedded into a Hilbert space H with a possibly weaker norm $\|\bullet\|_H$ (such as the L^2 norm rather than the $H(\mathrm{div})$ norm) and let there exist a (possibly mesh-dependent) norm $|\!|\!|\bullet|\!|\!|_h$ on a finite-dimensional subspace Y_h of Y. Given a finite-dimensional subspace $X_h \subseteq X$ with discrete solution $x_h \in X_h$, this paper identifies the abstract conditions (LBB), (C), and (E) that lead to the best-approximation result in the weaker norm $\| \bullet \|_H$ independent of the approximation error of the Lagrange multiplier,

$$\|x - x_h\|_H \lesssim \min_{\xi_h \in X_h} \|x - \xi_h\|_H + \text{data approximation}. \tag{1.2}$$

Here and throughout this paper, $A \lesssim B$ abbreviates $A \leq CB$ for some positive generic constant C which solely depends on the constants involved in (LBB), (C), and (E) and is, in particular, independent of the mesh-size or the Lamé parameter λ.

The condition (LBB) is the stability from the Brezzi splitting lemma [Bra07] for the discrete system with respect to the non-standard norms $\|\bullet\|_H$ and $|\!|\!|\bullet|\!|\!|_h$. The compatibility condition (C) for X_h, Y_h and b is condition (C) from [Bra07] and, in the case of the Stokes equations, equivalent to pointwise mass conservation of a finite element method. The condition (E) is the efficiency estimate

$$\sup_{\eta_h \in Y_h \setminus \{0\}} (F(\eta_h) - b(\xi_h, \eta_h)) \Big/ |\!|\!|\eta_h|\!|\!|_h \lesssim \|x - \xi_h\|_H + \text{data approximation}.$$

This combination of arguments from the a posteriori and a priori error analysis is called medius analysis [Gud10, CPS12] and leads to results which rely on no extra regularity of the weak solution $x \in X$ and hold on arbitrarily coarse meshes. The motivation for an L^2 theory of stress approximations stems from the concept of nonlinear approximation classes in the optimality analysis of adaptive mesh-refinement algorithms. Such a result was obtained by [HX12] for the Raviart-Thomas finite element discretisation of the Poisson equation for this purpose.

The prime goal of this paper is the medius analysis of the Arnold-Winther finite element method for the linear elasticity problem. The second part of this paper carries out the application of the abstract theory to this method. Let $\Omega \subseteq \mathbb{R}^n$ for $n = 2$ or $n = 3$ be a bounded polyhedral Lipschitz domain with (closed) Dirichlet boundary $\Gamma_D \subseteq \partial\Omega$ of positive surface measure and the (possibly empty) Neumann boundary $\Gamma_N = \partial\Omega \setminus \Gamma_D$ with outer unit normal ν. Given a volume force $f \in L^2(\Omega; \mathbb{R}^n)$, applied surface tractions $g \in L^2(\Gamma_N; \mathbb{R}^n)$, and boundary data $u_D \in H^1(\Omega; \mathbb{R}^n)$, the linear elasticity problem (in its strong form) seeks $u \in H^1(\Omega; \mathbb{R}^n)$ and $\sigma \in H(\mathrm{div}, \Omega; \mathbb{S}) := H(\mathrm{div}, \Omega)^n \cap L^2(\Omega; \mathbb{S})$ with

$$\begin{aligned} -\operatorname{div} \sigma &= f && \text{in } \Omega, \\ \sigma &= \mathbb{C}\varepsilon(u) && \text{in } \Omega, \\ u|_{\Gamma_D} &= u_D && \text{on } \Gamma_D, \\ \sigma\nu|_{\Gamma_N} &= g && \text{on } \Gamma_N. \end{aligned} \tag{1.3}$$

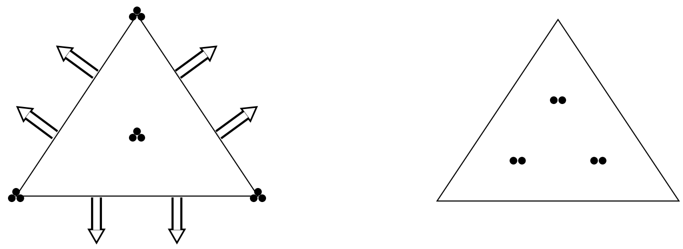

Figure 1: Illustration of the 30 degrees of freedom of the Arnold-Winther FEM for the lowest-order case $k = 1$ and $n = 2$ for the stress (left) and displacement (right). The three dots at the three vertices of the triangle (left) represent point evaluations while the three dots in the interior denote the integral means of the three components of τ_{AW}. The arrows represent the moments of order ≤ 1 of $\tau_{\mathrm{AW}}\nu_E$. The dots in the interior of the right triangle indicate the degrees of freedom for the two components of the P_1 displacement approximation.

Here and throughout the paper, $\mathbb{S} := \{A \in \mathbb{R}^{n\times n} \mid A = A^\top\}$ denotes the space of symmetric $n \times n$ matrices; the linear Green strain reads

$$\varepsilon(u) := (Du + (Du)^\top)/2,$$

and, for Lamé parameters $\mu > 0$ and $\lambda > 0$, the elasticity tensor $\mathbb{C}$ acts on $A \in \mathbb{R}^{n\times n}$ as

$$\mathbb{C}A := 2\mu A + \lambda(\operatorname{tr} A)1_{n\times n}.$$

For nearly incompressible materials, the Poisson ratio $\lambda/(2(\lambda+\mu))$ is close to $1/2$ and standard low-order conforming finite element methods (FEMs) are known to suffer from locking. The error shows the expected convergence rate only for very fine meshes with an impractically large number of degrees of freedom. The difficulty for mixed FEMs consists in finding an appropriate FEM space for the pointwise symmetric stress approximation. The simplest pointwise symmetric mixed FEM (conforming in $H(\mathrm{div})$) is the FEM of Arnold and Winther [AW02, AAW08] with numerical benchmark computations in [CEG11, CGP12, CGRT08]. For the exact solution (σ, u) of (1.3) and its approximation $(\sigma_{\mathrm{AW}}, u_k)$, the error analysis of [AW02, AAW08] proves quasi-optimal convergence

$$\begin{aligned}\|\sigma - \sigma_{\mathrm{AW}}\|_{H(\mathrm{div},\Omega)} &+ \|u - u_k\|_{L^2(\Omega)} \\ &\lesssim \min_{\tau_{\mathrm{AW}}\in \mathrm{AW}_k(\mathcal{T},g_{k+n})} \|\sigma - \tau_{\mathrm{AW}}\|_{H(\mathrm{div},\Omega)} + \inf_{v_k\in P_k(\mathcal{T};\mathbb{R}^n)} \|u - v_k\|_{L^2(\Omega)}.\end{aligned}$$

The constants hidden in the notation $\lesssim$ are independent of the mesh-size and the Lamé parameter λ but may depend on the parameter μ, on the domain Ω, as well as on the minimal angle of the shape-regular triangulation $\mathcal{T}$ with piecewise polynomials $P_k(\mathcal{T};\mathbb{R}^n)$. Details of the Arnold-Winther FEM space $\mathrm{AW}_k(\mathcal{T}, g_{k+n})$ and the embedded treatment of the applied traction forces follow in Section 4. The abstract result of this paper proves the quasi-optimal estimate with respect to the L^2 norm of the stress

$$\|\sigma - \sigma_{\mathrm{AW}}\|_{L^2(\Omega)} \lesssim \min_{\tau_{\mathrm{AW}}\in \mathrm{AW}_k(\mathcal{T},g_{k+n})} \|\sigma - \tau_{\mathrm{AW}}\|_{L^2(\Omega)} + \mathrm{osc}_k(f, \mathcal{T}). \tag{1.4}$$

For the definition of the (k-th order) data oscillations $\operatorname{osc}_k(f,\mathcal{T})$ see Subsection 3.3. The main result (1.4) may be contrasted with the well-established refined estimate

$$\|\sigma-\sigma_{\mathrm{AW}}\|_{\mathbb{C}^{-1}} \le \|\sigma - I_F\sigma\|_{\mathbb{C}^{-1}} \tag{1.5}$$

from [AW02, unlabeled displayed formula between (5.4) and (5.5)] with the Fortin operator I_F (called Π_h in [AW02, (3.3)–(3.5)]). It is not difficult to generalise this estimate (1.5) to

$$\|\sigma-\sigma_{\mathrm{AW}}\|_{\mathbb{C}^{-1}} \le \min_{\substack{\tau_{\mathrm{AW}}\in \mathrm{AW}_k(\mathcal{T},g_{k+n})\\ \operatorname{div}\tau_{\mathrm{AW}}=-\Pi_k f}} \|\sigma-\tau_{\mathrm{AW}}\|_{\mathbb{C}^{-1}} \tag{1.6}$$

with a minimum over all discrete test functions τ_{AW} with prescribed divergence $\operatorname{div}\tau_{\mathrm{AW}} = \operatorname{div}\sigma_{\mathrm{AW}} = -\Pi_k f$. The point is that the minimum in (1.4) is free from this restriction on the divergence in (1.6). On the expense of the extra data approximation term in (1.4) and the multiplicative generic constant, the estimate (1.4) states some approximation result as (1.6) without the side condition on the divergence of τ_{AW}. The main mathematical argument is the stability of the discrete mixed system with respect to the norms $\|\cdot\|_{L^2(\Omega)}$ and $|\!|\!|\cdot|\!|\!|_h$ rather than $\|\cdot\|_{H(\operatorname{div},\Omega)}$ and $\|\cdot\|_{L^2(\Omega)}$, where the mesh-dependent norm $|\!|\!|\cdot|\!|\!|_h$ involves the inter-element jumps of a piecewise smooth function. Similar mesh-dependent norms were previously considered in [BOP80, BV96, BDW99, LS06, HX12].

The analysis of this paper applies to the case $k = 1, 2$ and carefully explores the transformation behaviour of the local shape functions with a detailed analysis of the Piola transformation described in [AW02, AAW08]. The difficulty for a possible generalisation to $k \ge 3$ is that the local degrees of freedom are not preserved under the Piola transformation, cf. Subsection 5.2 for more details. Further applications include the mixed finite element discretisation of the Poisson and the Stokes equations.

The remaining parts of this paper are organised as follows. Section 2 states the precise abstract problem and proves the abstract result. Section 3 introduces the notation on the Navier-Lamé equations and regular triangulations. Section 4 gives an application of the abstract result to the Arnold-Winther FEM. Section 5 concludes the paper with some comments on extensions to the nonconforming Arnold-Winther FEM, a nonconforming treatment of Neumann data, the pure Dirichlet and Neumann problem, the application to the pseudostress FEM for the Stokes equations, and the equivalence of two approximation classes. Appendix A gives the necessary details on the basis functions and local degrees of freedom for the Arnold-Winther FEM.

2 Abstract Result

Let $(H, \|\bullet\|_H)$ and $(Y, \|\bullet\|_Y)$ be Hilbert spaces and $X \subseteq H$. Let $a : H\times H \to \mathbb{R}$ and $b : X \times Y \to \mathbb{R}$ be bilinear forms and $F : Y \to \mathbb{R}$ and $G : X \to \mathbb{R}$ be linear and bounded. Let $X(0) \subseteq X$ be some subspace of X and $X(g) = \sigma_g + X(0)$ for

$\sigma_g \in X$ an affine subspace of X to model boundary conditions. The abstract problem seeks $(x, y) \in X(g) \times Y$ with

$$\begin{aligned} a(x,\xi) + b(\xi, y) &= G(\xi) \qquad &&\text{for all } \xi \in X(0), \\ b(x,\eta) &= F(\eta) \qquad &&\text{for all } \eta \in Y. \end{aligned} \tag{2.1}$$

Let $X_h \subseteq X$, $X_h(0) := X_h \cap X(0)$ and $Y_h \subseteq Y$ denote some finite-dimensional subspaces of X, $X(0)$ and Y and set $X_h(g) := \sigma_{h,g} + X_h(0)$ for some $\sigma_{h,g} \in X_h$. The discrete problem seeks $(x_h, y_h) \in X_h(g) \times Y_h$ with

$$\begin{aligned} a(x_h,\xi_h) + b(\xi_h, y_h) &= G(\xi_h) \qquad &&\text{for all } \xi_h \in X_h(0), \\ b(x_h,\eta_h) &= F(\eta_h) \qquad &&\text{for all } \eta_h \in Y_h. \end{aligned} \tag{2.2}$$

The bilinear form $\mathcal{B} : (X_h \times Y_h) \times (X_h \times Y_h) \to \mathbb{R}$ associated with the system (2.2) reads, for any $x_h, \xi_h \in X_h$ and $y_h, \eta_h \in Y_h$, as

$$\mathcal{B}((x_h, y_h), (\xi_h, \eta_h)) := a(x_h, \xi_h) + b(x_h, \eta_h) + b(\xi_h, y_h).$$

Define the kernel

$$Z := \{\xi \in X(0) \mid \forall \eta \in Y,\, b(\xi, \eta) = 0\}$$

and its discrete counterpart

$$Z_h := \{\xi_h \in X_h(0) \mid \forall \eta_h \in Y_h,\ b(\xi_h, \eta_h) = 0\}. \tag{2.3}$$

Suppose that the subsequent hypotheses (LBB), (C) and (E) are valid. In applications, the term $\operatorname{dapx}(F, G)$ will measure the approximation of the given data F and G.

Let α_h, C_{cont}, M_h, γ_h and C_{eff} be positive real constants in the abstract conditions (LBB), (C), and (E) sufficient for best-approximation in H. The condition (LBB) consists of the standard conditions (LBB1a), (LBB1b), (LBB2a), (LBB2b) from the Brezzi splitting lemma [BF91, Bra07] which characterise the stability of the discrete system (2.2) with respect to the norms $\|\cdot\|_H$ and $|\!|\!|\cdot|\!|\!|_h$.

(LBB1a) $0 < \alpha_h = \inf_{\xi_h \in Z_h \setminus \{0\}} \sup_{\rho_h \in Z_h \setminus \{0\}} a(\rho_h, \xi_h) / (\|\rho_h\|_H \|\xi_h\|_H)$;

(LBB1b) $\forall \rho \in X \ \forall \xi_h \in X_h(0) \quad a(\rho, \xi_h) \le C_{\text{cont}} \, \|\rho\|_H \, \|\xi_h\|_H$;

(LBB2a) $\forall \xi_h \in X_h(0) \ \forall \eta_h \in Y_h \quad b(\xi_h, \eta_h) \le M_h \|\xi_h\|_H |\!|\!|\eta_h|\!|\!|_h$;

(LBB2b) $0 < \gamma_h = \inf_{\eta_h \in Y_h \setminus \{0\}} \sup_{\xi_h \in X_h(0) \setminus \{0\}} b(\xi_h, \eta_h) / (\|\xi_h\|_H |\!|\!|\eta_h|\!|\!|_h)$.

The compatibility of X_h and Y_h of [Bra07, Def. 4.6] reads

(C) $Z_h \subseteq Z$.

The efficiency of the residual $F - b(\xi_h, \cdot)$ in the dual norm of $(Y_h, |\!|\!|\cdot|\!|\!|_h)$ up to some data approximation term $\operatorname{dapx}(F, G)$ reads

(E) $\forall \xi_h \in X_h(g) \quad \sup_{\eta_h \in Y_h \setminus \{0\}} \dfrac{F(\eta_h) - b(\xi_h, \eta_h)}{\|\!|\eta_h|\!\|_h} \leq C_{\mathrm{eff}}(\|x - \xi_h\|_H + \mathrm{dapx}(F, G)).$

Theorem 2.1 (best-approximation in H)**.** *Suppose (LBB), (C), (E). Any solution* $(x, y) \in X(g) \times Y$ *to* (2.1) *and the discrete solution* $(x_h, y_h) \in X_h(g) \times Y_h$ *to* (2.2) *and any* $x_h^\star \in X_h(g)$ *satisfy*

$$\|x - x_h\|_H \leq C_{\mathrm{apx}}\big(\|x - x_h^\star\|_H + \mathrm{dapx}(F, G)\big).$$

The constant $C_{\mathrm{apx}} \approx 1$ *depends on* α_h*,* γ_h*,* M_h*,* C_{cont}*,* C_{eff} *from (LBB) and (E).*

The proof of Theorem 2.1 departs from an alternative formulation of condition (C).

Lemma 2.2. *Condition (C) is equivalent to*

$$\forall \eta \in Y \ \exists \eta_h^\star \in Y_h \ \forall \xi_h \in X_h(0) \quad b(\xi_h, \eta_h^\star) = b(\xi_h, \eta). \tag{2.4}$$

Proof. For the proof that condition (C) implies (2.4), consider the linear map $L : Y_h \to (X_h(0))^*$, $y_h \mapsto b(\bullet, y_h)$ with dual $L^* : X_h(0) \to Y_h^*$. The closed range theorem for the finite-dimensional image $L(Y_h)$ leads to

$$b(\bullet, Y_h) = \overline{L(Y_h)} = (\ker L^*)^0 \equiv \{\phi \in (X_h(0))^* \mid \forall \xi_h \in \ker L^*,\ \phi(\xi_h) = 0\}$$

(cf. [Bra07] for more details on the polar set $(\ker L^*)^0$). Hence, it suffices to show that for all $y \in Y$ and all $x_h \in \ker L^*$ it holds $b(x_h, y) = 0$. Since $\ker L^* = Z_h$, condition (C) states $\ker L^* \subseteq Z$. This proves (2.4).

For the proof of the converse implication, assume (2.4) and let $\xi_h \in Z_h$. For any $\eta \in Y$, (2.4) implies the existence of $\eta_h^\star$ such that $b(\xi_h, \eta) = b(\xi_h, \eta_h^\star) = 0$. Hence, $\xi_h \in Z$. □

Proof of Theorem 2.1. The condition (C) and Lemma 2.2 imply that there exists $y_h^\star \in Y_h$ with $b(\xi_h, y - y_h^\star) = 0$ for all $\xi_h \in X_h(0)$. Let $x_h^\star \in X_h(g)$. The condition (LBB) and the Brezzi splitting lemma [BF91, Bra07] imply the existence of $(\xi_h, \eta_h) \in X_h(0) \times Y_h$ with $\|\xi_h\|_H^2 + \|\!|\eta_h|\!\|_h^2 = 1$ and

$$\|x_h - x_h^\star\|_H + \|\!|y_h^\star - y_h|\!\|_h \leq C_{\mathrm{stab}} \mathcal{B}((x_h - x_h^\star, y_h - y_h^\star), (\xi_h, \eta_h)). \tag{2.5}$$

The constant $C_{\mathrm{stab}} \approx 1$ depends on $\alpha_h, \gamma_h, M_h, C_{\mathrm{cont}}$ from (LBB). The continuous and discrete equations (2.1)–(2.2) plus $X_h(0) \subseteq X(0)$ and $Y_h \subseteq Y$ show that

$$\begin{aligned} \mathcal{B}((x_h - x_h^\star, y_h - y_h^\star), (\xi_h, \eta_h)) &= \mathcal{B}((x - x_h^\star, y - y_h^\star), (\xi_h, \eta_h)) \\ &= a(x - x_h^\star, \xi_h) + F(\eta_h) - b(x_h^\star, \eta_h) + b(\xi_h, y - y_h^\star). \end{aligned} \tag{2.6}$$

The design of $y_h^\star$ implies that the last term vanishes. This, the conditions (LBB1b), (E), plus $\|\xi_h\|_H^2 + \|\!|\eta_h|\!\|_h^2 = 1$ imply

$$\mathcal{B}((x_h - x_h^\star, y_h - y_h^\star), (\xi_h, \eta_h)) \leq (C_{\mathrm{cont}} + C_{\mathrm{eff}})\big(\|x - x_h^\star\|_H + \mathrm{dapx}(F, G)\big).$$

The triangle inequality concludes the proof. □

Remark 2.3 (Nonconforming FEM). This remark explains why nonconforming methods with $X_h \not\subseteq X$ are excluded from the analysis of Theorem 2.1. For the sake of this exposition, assume homogeneous boundary conditions in the sense that $X(0) = X(g)$ and $X_h(0) = X_h(g)$ as well as $G = 0$ throughout this remark. Then some nonconforming variant of (2.5)–(2.6) reads

$$\begin{aligned} C_{\mathrm{stab}}^{-1}\big(\|x_h - x_h^\star\|_H + \|\!|y_h^\star - y_h|\!\|_h\big) &\le \mathcal{B}((x_h - x_h^\star, y_h - y_h^\star), (\xi_h, \eta_h)) \\ &= a(x - x_h^\star, \xi_h) + F(\eta_h) - b(x_h^\star, \eta_h) - a(x, \xi_h) - b(\xi_h, y). \end{aligned}$$

The nonconformity $\xi_h \notin X$ implies that the additional terms at the end, namely

$$a(x, \xi_h) + b(\xi_h, y) \tag{2.7}$$

do *not* vanish. Two examples for the Arnold-Winther FEM in Subsection 5.3 illustrate how (2.7) may lead to sub-optimal convergence rates in practical applications.

3 Preliminaries

This section provides the necessary preliminaries and notation on the Navier-Lamé equations and regular triangulations.

3.1 General Notation

Throughout the paper, standard notation on Lebesgue and Sobolev spaces and their norms applies and $(\bullet, \bullet)_{L^2(\Omega)}$ denotes the scalar product in $L^2(\Omega)$ (or $L^2(\Omega;\mathbb{R}^n)$ or $L^2(\Omega;\mathbb{R}^{n\times m})$). The subspace of $L^2(\Omega;\mathbb{R}^{m\times n})$ of functions with divergence in $L^2(\Omega;\mathbb{R}^m)$ is $H(\mathrm{div},\Omega;\mathbb{R}^{m\times n})$. For any measurable subset $\omega \subseteq \Omega$, $⨍_\omega \bullet\, dx$ denotes the integral mean on ω with respect to the n-dimensional Lebesgue measure. For $1 \le k < n$ and some k-dimensional face G, $⨍_G \bullet\, ds$ is the integral mean over G with respect to the k-dimensional surface measure. The duality pairing of $H^{-1/2}(\partial\Omega;\mathbb{R}^n)$ with $H^{1/2}(\partial\Omega;\mathbb{R}^n)$ is denoted by $\langle \bullet, \bullet\rangle_{\partial\Omega}$. The trace of a matrix $A \in \mathbb{R}^{n\times n}$ is denoted by $\mathrm{tr}(A)$ and the deviatoric part of A reads $\mathrm{dev}\, A := A - \mathrm{tr}(A)1_{n\times n}/n$. The dot denotes the product of two one-dimensional lists of the same length while the colon denotes the Euclidean product of matrices. The discrete norm $\|\!|\bullet|\!\|_h$ for piecewise polynomial L^2 functions is defined in (4.1) below. The measure $|\bullet|$ refers to the Euclidean norm of a vector or a matrix or, for a domain in $\mathbb{R}^n$, to its n-dimensional Lebesgue measure or its $(n-1)$-dimensional surface measure.

3.2 Weak Mixed Formulation of Linear Elasticity

For the sake of this exposition assume that both $\Gamma_D \neq \emptyset$ and $\Gamma_N \neq \emptyset$ have a positive $(n-1)$ dimensional surface measure; the results of this paper remain true both for the pure Dirichlet problem $\Gamma_D = \partial\Omega$ and for the pure Neumann problem $\Gamma_N = \partial\Omega$ with the necessary modifications from Section 5.1.

The weak mixed formulation of (1.3) is based on the spaces

$$\begin{aligned} X(g) &:= \{\tau \in H(\operatorname{div}, \Omega; \mathbb{S}) \mid \tau\nu|_{\Gamma_N} = g\} && \text{for the stress;} \\ Y &:= L^2(\Omega; \mathbb{R}^n) && \text{for the displacement.} \end{aligned}$$

Define the bilinear forms $a : H(\operatorname{div}, \Omega; \mathbb{S}) \times H(\operatorname{div}, \Omega; \mathbb{S}) \to \mathbb{R}$ and $b : H(\operatorname{div}, \Omega; \mathbb{S}) \times Y \to \mathbb{R}$ by

$$a(\tau, \rho) := (\mathbb{C}^{-1}\tau, \rho)_{L^2(\Omega)} \quad \text{and} \quad b(\tau, v) := (\operatorname{div}\tau, v)_{L^2(\Omega)}$$

for any $\tau, \rho \in H(\operatorname{div}, \Omega; \mathbb{S})$ and $v \in Y$. Recall the duality pairing $\langle\cdot,\cdot\rangle_{\partial\Omega}$ of $H^{-1/2}(\partial\Omega; \mathbb{R}^n)$ with $H^{1/2}(\partial\Omega; \mathbb{R}^n)$. The mixed formulation of the elasticity problem seeks $\sigma \in X(g)$ and $u \in Y$ with

$$\begin{aligned} a(\sigma, \tau) + b(\tau, u) &= \langle \tau\nu, u_D\rangle_{\partial\Omega} && \text{for all } \tau \in X(0), \\ b(\sigma, v) &= -(f, v)_{L^2(\Omega)} && \text{for all } v \in Y. \end{aligned} \tag{3.1}$$

The unique solvability of (3.1) is well known [BF91, Bra07].

3.3 Triangulations

Let $\mathcal{T}$ denote some shape-regular simplicial triangulation of the bounded polyhedral Lipschitz domain Ω, i.e., $\overline{\Omega} = \cup\mathcal{T}$ and $\mathcal{T}$ is regular in the sense that any two distinct simplices are either disjoint or share exactly one common node, edge or face (for $n = 2$ the edges coincide with the faces). For any simplex $T \in \mathcal{T}$, $\mathcal{F}(T)$ is the set of faces and $\mathcal{N}(T)$ is the set of vertices of T. Let $\mathcal{F}$ denote the set of faces of $\mathcal{T}$ and $\mathcal{N}$ the set of vertices. Assume that $\mathcal{T}$ resolves the boundary conditions in the sense that $\Gamma_D = \cup\{F \in \mathcal{F} \mid F \subseteq \Gamma_D\}$. For $\omega \subseteq \mathbb{R}^n$ define the sets $\mathcal{N}(\omega) := \mathcal{N} \cap \omega$ and $\mathcal{F}(\omega) := \{F \in \mathcal{F} \mid \operatorname{int}(F) \subseteq \omega\}$ for the relative interior $\operatorname{int}(F)$ of a face $F \in \mathcal{F}$ (e.g. the interior nodes are denoted by $\mathcal{N}(\Omega)$ and the interior faces by $\mathcal{F}(\Omega)$). The jump along an interior face $F \in \mathcal{F}(\Omega)$ with adjacent simplices T_+ and T_-, i.e., $F = T_+ \cap T_-$, is defined by $[v]_F := v|_{T_+} - v|_{T_-}$. The jump along boundary faces $F \in \mathcal{F}(\Gamma_D)$ reads $[v]_F := v|_{T_+}$ for that simplex $T_+ \in \mathcal{T}$ with $F \subseteq T_+$. In this situation, let ν_F denote the outer unit normal of T_+ on F. Given a simplex $T \in \mathcal{T}$, let ν_T denote the outer unit normal of T. For a face $F \in \mathcal{F}$, the patch ω_F reads $\omega_F := \operatorname{int}(\cup\{T \in \mathcal{T} \mid F \in \mathcal{F}(T)\})$.

Define the piecewise polynomials by

$$\begin{aligned} P_k(T; \mathbb{R}^m) &:= \{p \in L^2(T; \mathbb{R}^m) \mid \forall j = 1, \dots, m, \text{ the } j\text{-th component} \\ &\qquad\qquad p_j \text{ of } p \text{ is a polynomial of total degree } \le k\}, \\ P_k(\mathcal{T}; \mathbb{R}^m) &:= \{p_k \in L^2(\Omega; \mathbb{R}^m) \mid \forall T \in \mathcal{T},\ p_k|_T \in P_k(T; \mathbb{R}^m)\}. \end{aligned}$$

Furthermore, let $P_k(T; \mathbb{S})$ (resp. $P_k(\mathcal{T}; \mathbb{S})$) denote the polynomials (resp. piecewise polynomials) with values in $\mathbb{S}$. For any vertex $z \in \mathcal{N}$, the piecewise affine and globally continuous function with 1 at z and 0 at all other vertices of $\mathcal{N}$

defines the conforming P_1 nodal basis function φ_z. The piecewise affines (into the $\mathbb{R}^2$) with respect to $\mathcal{F}(\Gamma_N)$ read $P_1(\mathcal{F}(\Gamma_N);\mathbb{R}^2)$.

The L^2 projection $\Pi_k : L^2(\Omega;\mathbb{R}^m) \to P_k(\mathcal{T};\mathbb{R}^m)$ onto piecewise polynomials reads

$$\Pi_k f := \operatorname*{argmin}_{p_k \in P_k(\mathcal{T};\mathbb{R}^m)} \|f - p_k\|_{L^2(\Omega)}.$$

The oscillations of order k of f are defined as $\operatorname{osc}_k(f,\mathcal{T}) := \|h_\mathcal{T}(1-\Pi_k)f\|_{L^2(\Omega)}$ for the piecewise constant mesh-size $h_\mathcal{T} \in P_0(\mathcal{T};\mathbb{R})$, i.e., $h_\mathcal{T}|_T = \operatorname{diam}(T)$ for all $T \in \mathcal{T}$. The diameter of a face $F \in \mathcal{F}$ is denoted by h_F.

Let $\varepsilon_{\mathrm{NC}}$ and D_{NC} denote the piecewise action of $\varepsilon := \operatorname{sym} D$ and the derivative D with respect to the triangulation $\mathcal{T}$, i.e., for a piecewise polynomial function $p_k \in P_k(\mathcal{T};\mathbb{R}^n)$, let $\varepsilon_{\mathrm{NC}}(p_k)|_T := \varepsilon(p_k|_T)$ and $(D_{\mathrm{NC}}p_k)|_T := D(p_k|_T)$ for all $T \in \mathcal{T}$. For a piecewise polynomial tensor field $\tau_k \in P_k(\mathcal{T};\mathbb{R}^{n\times n})$ the piecewise action of the divergence is defined via $(\operatorname{div}_{\mathrm{NC}} \tau_k)|_T := \operatorname{div}_{\mathrm{NC}}(\tau_k|_T)$ for all $T \in \mathcal{T}$.

4 L^2 Best-Approximation of Arnold-Winther FEM

This section defines the Arnold-Winther FEM, and states the main result in Subsection 4.1 with the proof of (LBB), (C) and (E) in Subsections 4.2–4.4 for space dimensions $n = 2, 3$. Recall that $A \lesssim B$ abbreviates $A \leq CB$ for some positive generic constant C which may depend on the shape regularity of $\mathcal{T}$ and the Lamé parameter μ but neither on λ nor on the mesh-size.

4.1 Arnold-Winther FEM

For the degree $k = 1, 2$, the approximation of the displacement of the Arnold-Winther FEM reads

$$Y_h := P_k(\mathcal{T};\mathbb{R}^n).$$

The local Arnold-Winther finite element space reads

$$\mathrm{AW}_k(T) := \{\tau_{\mathrm{AW}} \in P_{k+n}(T;\mathbb{S}) \mid \operatorname{div} \tau_{\mathrm{AW}} \in P_k(T;\mathbb{R}^n)\}.$$

Let the function g_{k+n} define some suitable approximation of g by traces of piecewise $\mathrm{AW}_k(T)$ functions which are globally in $H(\operatorname{div},\Omega;\mathbb{S})$. In the case of smooth consistent Neumann data, g_{k+n} can be defined via the degrees of freedom of the Arnold-Winther finite element (see Figure 1 for $n = 2$ and $k = 0$). For inconsistent Neumann data, a modification at the vertices is required (as described in [CGRT08]). Throughout this paper, we merely impose the following condition.

(N) For all $v_1 \in P_1(\mathcal{F}(\Gamma_N);\mathbb{R}^2) \cap H^1(\Gamma_N;\mathbb{R}^2)$ it holds $\int_{\Gamma_N} (g - g_{k+n}) \cdot v_1 \, ds = 0$.

The stress is approximated in

$$\mathrm{AW}_k(\mathcal{T}, g_{k+n}) := \{\tau_{\mathrm{AW}} \in X(g_{k+n}) \mid \forall T \in \mathcal{T},\ \tau_{\mathrm{AW}}|_T \in \mathrm{AW}_k(T)\} \subseteq X(g_{k+n}).$$

The approximation properties of $\mathrm{AW}_k(\mathcal{T}, g_{k+n})$ depend in a delicate way on the choice of g_{k+n}. However, the main result in Theorem 4.1 requires only the condition (N) and $\mathrm{AW}_k(\mathcal{T}, g_{k+n}) \neq \emptyset$.

Define a discrete norm on Y_h by

$$|||v_h|||_h^2 := \|\varepsilon_{\mathrm{NC}}(v_h)\|_{L^2(\Omega)}^2 + \sum_{F\in\mathcal{F}(\Omega\cup\Gamma_D)} h_F^{-1}\|[v_h]_F\|_{L^2(F)}^2 \quad \text{for any } v_h \in Y_h. \tag{4.1}$$

(In fact, if $|||v_h|||_h = 0$, then v_h is continuous and a rigid body motion with $v_h|_{\Gamma_D} = 0$ and, hence, the Korn inequality [Bra07, p. 298] implies $v_h \equiv 0$.)

The discrete mixed problem seeks $\sigma_{\mathrm{AW}} \in \mathrm{AW}_k(\mathcal{T}, g_{k+n})$ and $u_k \in Y_h$ such that

$$\begin{aligned} a(\sigma_{\mathrm{AW}}, \tau_{\mathrm{AW}}) + b(\tau_{\mathrm{AW}}, u_k) &= \langle \tau_{\mathrm{AW}}\nu, u_D\rangle_{\partial\Omega} && \text{for all } \tau_{\mathrm{AW}} \in \mathrm{AW}_k(\mathcal{T}, 0),\\ b(\sigma_{\mathrm{AW}}, v_h) &= -(f, v_h)_{L^2(\Omega)} && \text{for all } v_h \in Y_h. \end{aligned} \tag{4.2}$$

The following theorem states the L^2 best-approximation result for the Arnold-Winther FEM with

$$\begin{aligned} X_h(0) &:= \mathrm{AW}_k(\mathcal{T}, 0) \subseteq X(0) \subseteq H := L^2(\Omega; \mathbb{S}),\\ X_h(g) &:= \mathrm{AW}_k(\mathcal{T}, g_{k+n}),\\ Y_h &:= P_k(\mathcal{T}; \mathbb{R}^n) \subseteq Y := L^2(\Omega; \mathbb{R}^n) \end{aligned}$$

and right-hand sides $G(\tau) := \langle \tau\nu, u_D\rangle_{\partial\Omega}$ and $F(v) = -(f, v)_{L^2(\Omega)}$ for all $\tau \in X(0)$ and $v \in L^2(\Omega; \mathbb{R}^n)$.

The following main result implies (1.4).

Theorem 4.1 (L^2 best-approximation for Arnold-Winther FEM)**.** *Suppose that g_{k+n} satisfies the condition (N). Then any $\tau_{\mathrm{AW}} \in \mathrm{AW}_k(\mathcal{T}, g_{k+n})$ satisfies*

$$\|\sigma - \sigma_{\mathrm{AW}}\|_{L^2(\Omega)} + |||\Pi_k u - u_k|||_h \lesssim \|\sigma - \tau_{\mathrm{AW}}\|_{L^2(\Omega)} + \mathrm{osc}_k(f, \mathcal{T}).$$

The condition (C) follows immediately from $\operatorname{div} \mathrm{AW}_k(\mathcal{T}, 0) \subseteq P_k(\mathcal{T}; \mathbb{R}^n)$. The conditions (LBB) and (E) are verified in the remaining subsections. This and Theorem 2.1 conclude the proof of Theorem 4.1.

4.2 Proof of (LBB1)

The following modification of [BF91, Proposition 7 in Section IV.3] follows from [CD98, Theorem 4.1].

Lemma 4.2 (tr-dev-div lemma)**.** *Let Σ_0 be a closed subspace of $H(\mathrm{div}, \Omega; \mathbb{R}^{n\times n})$ which does not contain the constant tensor $1_{n\times n}$. Then any $\tau \in \Sigma_0$ satisfies*

$$\|\mathrm{tr}(\tau)\|_{L^2(\Omega)} \lesssim \|\mathrm{dev}\,\tau\|_{L^2(\Omega)} + \|\mathrm{div}\,\tau\|_{L^2(\Omega)}.$$

Proof. The stated inequality is proven [CD98, Theorem 4.1] for the case $n = 2$. The generalisation to $n \geq 3$ follows directly with the arguments of [CD98] and [ASV88, ADM06]. □

Recall the definition of the discrete kernel Z_h from (2.3).

Lemma 4.3 ((LBB1))**.** *The bilinear form a is continuous and Z_h-elliptic with respect to the norm $\|\cdot\|_{L^2(\Omega)}$. The respective constants do not depend on λ or the mesh-sizes in $\mathcal{T}$, but possibly on the shape-regularity of $\mathcal{T}$.*

Proof. The continuity follows from $|\mathbb{C}^{-1}A| \lesssim |A|$ for all $A \in \mathbb{R}^{n\times n}$. Let $\tau_{\mathrm{AW}} \in \mathrm{AW}_k(\mathcal{T}, 0)$ with $\int_\Omega v_h \operatorname{div} \tau_{\mathrm{AW}}\, dx = 0$ for all $v_h \in Y_h$. Since $\operatorname{div} \tau_{\mathrm{AW}} \in Y_h$, this leads to $\operatorname{div} \tau_{\mathrm{AW}} = 0$. Since $\mathrm{AW}(\mathcal{T}, 0)$ is a closed subspace of $H(\operatorname{div}, \Omega; \mathbb{R}^{n\times n})$ which does not contain the constant tensor $1_{n\times n}$, Lemma 4.2 yields

$$\|\operatorname{tr} \tau_{\mathrm{AW}}\|_{L^2(\Omega)} \lesssim \|\operatorname{dev} \tau_{\mathrm{AW}}\|_{L^2(\Omega)} + \|\operatorname{div} \tau_{\mathrm{AW}}\|_{L^2(\Omega)}.$$

Since $\operatorname{div} \tau_{\mathrm{AW}} = 0$, this and $|\operatorname{dev} A|^2 \lesssim A : \mathbb{C}^{-1}A$ conclude the proof. □

4.3 Proof of (LBB2)

The stability involves the mesh-dependent norm $|\!|\!|\cdot|\!|\!|_h$ from (4.1).

Lemma 4.4 ((LBB2a))**.** *The discrete bilinear form b is continuous with respect to the norms $\|\cdot\|_{L^2(\Omega)}$ and $|\!|\!|\cdot|\!|\!|_h$, in the sense that*

$$b(\tau_{\mathrm{AW}}, v_h) \lesssim \|\tau_{\mathrm{AW}}\|_{L^2(\Omega)} |\!|\!| v_h |\!|\!|_h \quad \textit{for all } (\tau_{\mathrm{AW}}, v_h) \in \mathrm{AW}_k(\mathcal{T}, 0) \times Y_h.$$

Proof. An integration by parts and trace [DE12, Lemma 1.49] and inverse inequalities [DE12, Lemma 1.44] lead to

$$\begin{aligned}(\operatorname{div} \tau_{\mathrm{AW}}, v_h)_{L^2(\Omega)} &= -(\tau_{\mathrm{AW}}, D_{\mathrm{NC}} v_h)_{L^2(\Omega)} + \sum_{F\in\mathcal{F}(\Omega\cup\Gamma_D)} \int_F [v_h]_F \cdot (\tau_{\mathrm{AW}} \nu_F)\, ds \\ &\lesssim \|\varepsilon_{\mathrm{NC}}(v_h)\|_{L^2(\Omega)} \|\tau_{\mathrm{AW}}\|_{L^2(\Omega)} + \sum_{F\in\mathcal{F}(\Omega\cup\Gamma_D)} h_F^{-1/2} \|[v_h]_F\|_{L^2(F)} \|\tau_{\mathrm{AW}}\|_{L^2(\omega_F)} \\ &\lesssim \|\tau_{\mathrm{AW}}\|_{L^2(\Omega)} |\!|\!| v_h |\!|\!|_h.\end{aligned}$$

□

Theorem 4.5 ((LBB2b))**.** *The bilinear form b satisfies the inf-sup condition*

$$1 \lesssim \beta \leq \inf_{v_h\in Y_h\setminus\{0\}} \sup_{\tau_{\mathrm{AW}}\in\mathrm{AW}_k(\mathcal{T},0)\setminus\{0\}} \frac{b(\tau_{\mathrm{AW}}, v_h)}{\|\tau_{\mathrm{AW}}\|_{L^2(\Omega)} |\!|\!| v_h |\!|\!|_h}.$$

Proof. Let $v_h \in Y_h$ and define $\tau_{\mathrm{AW}} \in \mathrm{AW}_k(\mathcal{T}, 0)$ by the specification of the degrees of freedom in such a way that the volume moments of degree $\leq k - 1$ coincide with those of $-\varepsilon_{\mathrm{NC}}(v_h)$ and the moments of the faces of degree k coincide with the moments of weighted jumps of $[v_h]_F$. Precisely (with the notation of Appendix A) set the volume degrees of freedom as

$$L_{T,\ell}(\tau_{\mathrm{AW}}) = -L_{T,\ell}(\varepsilon_{\mathrm{NC}}(v_h)) \qquad \text{for } \ell \in I_{\mathrm{vol},1}(T),$$

and the face degrees of freedom as

$$L_{T,(F,\eta,\alpha)}(\tau_{\mathrm{AW}}) = h_F^{-1}\, e_\alpha \cdot \left(\fint_F \prod_{j=1}^{n-1} (\varphi_{z_j} - \varphi_{z_n})^{\eta_j}\, [v_h]_F\, ds \right)$$

for $(F,\eta,\alpha) \in I_{\mathrm{faces}}(T)$ with $F = \mathrm{conv}\{z_1,\dots,z_n\} \in F(\Omega \cup \Gamma_D)$ and set the remaining degrees of freedom as

$$L_{T,\ell}(\tau_{\mathrm{AW}}) = 0 \qquad \text{for } \ell \in I_{\mathrm{nodes}}(T) \cup I_{\mathrm{vol},2}(T) \cup I_{\mathrm{edges}}(T)$$

for all $T \in \mathcal{T}$. See Subsection A.1 for the precise definition of these functionals. An integration by parts yields

$$(v_h, \operatorname{div} \tau_{\mathrm{AW}})_{L^2(\Omega)} = -(D_{\mathrm{NC}} v_h, \tau_{\mathrm{AW}})_{L^2(\Omega)} + \sum_{F \in \mathcal{F}(\Omega \cup \Gamma_D)} \int_F [v_h]_F \cdot (\tau_{\mathrm{AW}} \nu_F)\, ds.$$

The definition of τ_{AW} proves

$$-(D_{\mathrm{NC}} v_h, \tau_{\mathrm{AW}})_{L^2(\Omega)} = \|\varepsilon_{\mathrm{NC}}(v_h)\|_{L^2(\Omega)}^2.$$

The hyper-face contributions satisfy, for any $F \in \mathcal{F}(\Omega \cup \Gamma_D)$, that

$$\int_F [v_h]_F \cdot (\tau_{\mathrm{AW}} \nu_F)\, ds = h_F^{-1} \|[v_h]_F\|_{L^2(F)}^2.$$

This leads to

$$b(\tau_{\mathrm{AW}}, v_h)_{L^2(\Omega)} = \|\!|v_h|\!\|_h^2. \tag{4.3}$$

Let $T \in \mathcal{T}$ and let $\sigma_\ell \in \mathrm{AW}_k(T)$ for $\ell \in I_{\mathrm{vol},1}(T) \cup I_{\mathrm{faces}}(T)$ denote the shape functions of the volume degrees of freedom and the face degrees of freedom on T. Then, the representation of $\tau_{\mathrm{AW}}|_T$ with respect to the shape functions reads

$$\tau_{\mathrm{AW}}|_T = \sum_{\ell \in I_{\mathrm{vol},1}(T) \cup I_{\mathrm{faces}}(T)} L_{T,\ell}(\tau_{\mathrm{AW}}) \sigma_\ell.$$

The triangle inequality and the scaling of the L^2 norm of the shape functions of Theorem A.7 show

$$\|\tau_{\mathrm{AW}}\|_{L^2(T)} \lesssim |T|^{1/2} \sum_{\ell \in I_{\mathrm{vol},1}(T) \cup I_{\mathrm{faces}}(T)} |L_{T,\ell}(\tau_{\mathrm{AW}})|.$$

Since $|\varphi_{z_j} - \varphi_{z_n}| \leq 1$ in Ω, a Cauchy inequality leads, for $\ell \in I_{\mathrm{vol},1}(T)$, to

$$|L_{T,\ell}(\tau_{\mathrm{AW}})| \leq \fint_T |\varepsilon_{\mathrm{NC}}(v_h)|\, dx \leq |T|^{-1/2}\, \|\varepsilon_{\mathrm{NC}}(v_h)\|_{L^2(T)}.$$

The same arguments reveals, for $\ell \in I_{\mathrm{faces}}(T)$, that

$$|L_{T,\ell}(\tau_{\mathrm{AW}})| \leq h_F^{-1}\, |F|^{-1/2}\, \|[v_h]_F\|_{L^2(F)}.$$

The shape regularity implies $|T|^{1/2}\, h_F^{-1}\, |F|^{-1/2} \approx h_F^{-1/2}$. The combination of the previous inequalities and the sum over all simplices results in

$$\|\tau_{\mathrm{AW}}\|_{L^2(\Omega)} \lesssim \|\!|v_h|\!\|_h.$$

This and (4.3) imply $\|\tau_{\mathrm{AW}}\|_{L^2(\Omega)} \|\!|v_h|\!\|_h \lesssim b(\tau_{\mathrm{AW}}, v_h)$. This concludes the proof. □

4.4 Proof of (E)

This subsection is devoted to the proof of (E) for the Arnold-Winther FEM based on the following three lemmas.

Lemma 4.6 (efficiency). *Any $\tau_{\mathrm{AW}} \in \mathrm{AW}_k(\mathcal{T}, g_{k+n})$ satisfies*

$$\|h_{\mathcal{T}}(f + \operatorname{div} \tau_{\mathrm{AW}})\|_{L^2(\Omega)} \lesssim \|\sigma - \tau_{\mathrm{AW}}\|_{L^2(\Omega)} + \operatorname{osc}_k(f, \mathcal{T}).$$

Proof. This follows from the well-established arguments of [Ver96]. □

Lemma 4.7 (discrete Korn inequality). *Any $v_h \in Y_h$ satisfies*

$$\|D_{\mathrm{NC}} v_h\|_{L^2(\Omega)} \lesssim |||v_h|||_h.$$

Proof. It follows from [Bre04, Eqn (1.19)] that

$$\|D_{\mathrm{NC}} v_h\|^2_{L^2(\Omega)} \lesssim \|v_h\|^2_{L^2(\Gamma_D)} + \|\varepsilon_{\mathrm{NC}}(v_h)\|^2_{L^2(\Omega)} + \sum_{F \in \mathcal{F}(\Omega)} h_F^{-1} \|[v_h]_F\|^2_{L^2(F)}.$$

Since $h_F \lesssim 1$, the boundary terms are controlled as

$$\sum_{F \in \mathcal{F}(\Gamma_D)} \|v_h\|^2_{L^2(F)} \lesssim \sum_{F \in \mathcal{F}(\Gamma_D)} h_F^{-1} \|v_h\|^2_{L^2(F)} \le |||v_h|||^2_h.$$ □

Let $\mathcal{T}(z) := \{T \in \mathcal{T} \mid z \in T\}$ and define the enrichment operator $J_1 : P_0(\mathcal{T}; \mathbb{R}^n) \to \big(P_1(\mathcal{T}; \mathbb{R}^n) \cap H^1(\Omega; \mathbb{R}^n)\big)$ for $v_0 \in P_0(\mathcal{T}; \mathbb{R}^n)$ by

$$J_1 v_0(z) = (\operatorname{card} \mathcal{T}(z))^{-1} \sum_{T \in \mathcal{T}(z)} v_0|_T(z)$$

for all $z \in \mathcal{N}(\Omega \cup \Gamma_N)$ and $J_1 v_0(z) = 0$ for all $z \in \mathcal{N}(\Gamma_D)$ followed by linear interpolation on all simplices.

Lemma 4.8 (approximation and stability estimates). *Any $v_h \in Y_h$ satisfies*

$$|||\Pi_0 v_h|||_h \lesssim |||v_h|||_h. \tag{4.4}$$

Any $v_0 \in P_0(\mathcal{T}; \mathbb{R}^n)$ satisfies

$$\|h_{\mathcal{T}}^{-1}(1 - J_1) v_0\|_{L^2(\Omega)} + \|D_{\mathrm{NC}} J_1 v_0\|_{L^2(\Omega)} \lesssim |||v_0|||_h. \tag{4.5}$$

Proof. For $F \in \mathcal{F}(\Omega)$ let $T_+, T_- \in \mathcal{T}$ with $F = T_+ \cap T_-$. Define the two seminorms ρ_1, ρ_2 on $P_k(\{T_+, T_-\}; \mathbb{R}^n)$ for $w_h \in P_k(\{T_+, T_-\}; \mathbb{R}^n)$ by

$$\rho_1(w_h) := h_F^{-1/2} \|[\Pi_0 w_h]_F\|_{L^2(F)},$$
$$\rho_2(w_h) := \|D_{\mathrm{NC}} w_h\|_{L^2(\omega_F)} + h_F^{-1/2} \|[w_h]_F\|_{L^2(F)}.$$

If $\rho_2(w_h) = 0$, then w_h is constant and continuous on ω_E. This implies $\Pi_0 w_h|_{T_+} = \Pi_0 w_h|_{T_-}$ and, hence, $\rho_1(w_h) = 0$. Since ρ_1 and ρ_2 are seminorms, there exists a constant $C > 0$ with $\rho_1 \leq C\rho_2$. A scaling argument proves that $C \approx 1$ is independent of the mesh-size. This proves

$$h_F^{-1/2} \|[\Pi_0 w_h]_F\|_{L^2(F)} \lesssim \|D_{\mathrm{NC}} w_h\|_{L^2(\omega_F)} + h_F^{-1/2} \|[w_h]_F\|_{L^2(F)}. \tag{4.6}$$

The same arguments apply for $F \in \mathcal{F}(\Gamma_D)$ and prove (4.6) for $\omega_F := \mathrm{int}(T_+)$ for the one simplex $T_+ \in \mathcal{T}$ with $F \subseteq T_+$. The sum over all edges and the bounded overlap of ω_F yield for any $v_h \in Y_h$ that

$$\sum_{F \in \mathcal{F}(\Omega \cup \Gamma_D)} h_F^{-1} \|[\Pi_0 v_h]_F\|^2_{L^2(F)} \lesssim \|D_{\mathrm{NC}} v_h\|^2_{L^2(\Omega)} + \sum_{F \in \mathcal{F}(\Omega \cup \Gamma_D)} h_F^{-1} \|[v_h]_F\|^2_{L^2(F)}.$$

This and Lemma 4.7 prove (4.4).

For the proof of (4.5), note that $v_0 \in P_0(\mathcal{T}; \mathbb{R}^n)$ implies for $z \in F \in \mathcal{F}(\Gamma_D)$ that

$$|v_0(z)|^2 \approx h_F^{-(n-1)} \|[v_0]_F\|^2_{L^2(F)}.$$

This and the arguments of [BS08, Lemma (10.6.6), p. 296] prove

$$\|h_{\mathcal{T}}^{-1}(v_0 - J_1 v_0)\|^2_{L^2(\Omega)} \lesssim \sum_{F \in \mathcal{F}(\Omega \cup \Gamma_D)} h_F^{-1} \|[v_0]_F\|^2_{L^2(F)}.$$

An inverse inequality proves

$$\|D_{\mathrm{NC}} J_1 v_0\|_{L^2(\Omega)} \lesssim \|h_{\mathcal{T}}^{-1}(v_0 - J_1 v_0)\|_{L^2(\Omega)}.$$

This concludes the proof. □

The proof of (E) concludes this subsection.

Lemma 4.9 ((E))**.** *Any $v_h \in Y_h$ with $\|\!|v_h|\!\|_h = 1$ and any $\tau_{\mathrm{AW}} \in \mathrm{AW}_k(\mathcal{T}, g_{k+n})$ satisfy*

$$(f + \operatorname{div} \tau_{\mathrm{AW}}, v_h)_{L^2(\Omega)} \lesssim \|\sigma - \tau_{\mathrm{AW}}\|_{L^2(\Omega)} + \mathrm{osc}_k(f, \mathcal{T}).$$

Proof. Let $v_h \in Y_h$ with $\|\!|v_h|\!\|_h = 1$ and consider

$$\begin{aligned}(f + \operatorname{div} \tau_{\mathrm{AW}}, v_h)_{L^2(\Omega)} \\ = (f + \operatorname{div} \tau_{\mathrm{AW}}, (1 - \Pi_0) v_h)_{L^2(\Omega)} + (f + \operatorname{div} \tau_{\mathrm{AW}}, \Pi_0 v_h)_{L^2(\Omega)}.\end{aligned} \tag{4.7}$$

The piecewise Poincaré inequality, Lemma 4.6 and the discrete Korn inequality from Lemma 4.7 control the first term of on the right-hand side of (4.7) as

$$\begin{aligned}(f + \operatorname{div} \tau_{\mathrm{AW}}, (1 - \Pi_0) v_h)_{L^2(\Omega)} &\lesssim \|h_{\mathcal{T}}(f + \operatorname{div} \tau_{\mathrm{AW}})\|_{L^2(\Omega)} \|D_{\mathrm{NC}} v_h\|_{L^2(\Omega)} \\ &\lesssim \|\sigma - \tau_{\mathrm{AW}}\|_{L^2(\Omega)} + \mathrm{osc}_k(f, \mathcal{T}).\end{aligned}$$

Lemma 4.8 implies for the second term on the right-hand side of (4.7) that

$$\begin{aligned}&(f + \operatorname{div} \tau_{\mathrm{AW}}, \Pi_0 v_h)_{L^2(\Omega)} \\ &\quad = (f + \operatorname{div} \tau_{\mathrm{AW}}, (1 - J_1)\Pi_0 v_h + J_1 \Pi_0 v_h)_{L^2(\Omega)} \\ &\quad \lesssim \|h_{\mathcal{T}}(f + \operatorname{div} \tau_{\mathrm{AW}})\|_{L^2(\Omega)} + (f + \operatorname{div} \tau_{\mathrm{AW}}, J_1 \Pi_0 v_h)_{L^2(\Omega)}.\end{aligned} \tag{4.8}$$

An integration by parts, $-\operatorname{div}\sigma = f$, and Lemma 4.8 reveal for the last contribution in (4.8) that

$$(f + \operatorname{div}\tau_{\mathrm{AW}}, J_1\Pi_0 v_h)_{L^2(\Omega)} = (\sigma - \tau_{\mathrm{AW}}, \varepsilon(J_1\Pi_0 v_h))_{L^2(\Omega)} \lesssim \|\sigma - \tau_{\mathrm{AW}}\|_{L^2(\Omega)}.$$

In the integration by parts, the boundary term $\int_{\Gamma_N}((\sigma - \tau_{\mathrm{AW}})\nu)\cdot J_1\Pi_0 v_h\,ds$ does not arise because g_{k+n} fulfills the condition (N). The combination of the foregoing displayed inequalities with Lemma 4.6 concludes the proof. □

5 Comments

This section discusses possible generalisations of Theorem 4.1 as well as applications to other finite element methods.

5.1 Pure Dirichlet and Pure Neumann Problem

The pure Dirichlet problem $\Gamma_D = \partial\Omega$ involves the stress space

$$X := \left\{\tau \in H(\operatorname{div},\Omega;\mathbb{S}) \;\middle|\; \int_\Omega \operatorname{tr}\tau\,dx = 0\right\}.$$

The arguments for the proof of Theorem 4.1 remain valid and, hence, Theorem 4.1 also holds in the case of pure Dirichlet boundary conditions.

The pure Neumann problem $\Gamma_N = \partial\Omega$ specifies the displacement up to rigid-body motions only. Hence, the space of the displacements reads

$$Y := L^2(\Omega;\mathbb{R}^n)/\mathrm{RM} \qquad \text{and} \qquad Y_h := P_k(\mathcal{T};\mathbb{R}^n)\cap Y$$

with the space of rigid-body motions RM on Ω. The discrete solution $(\sigma_{\mathrm{AW}}, u_k)$ to (4.2) and the exact solution (σ, u) satisfy the L^2 best-approximation of Theorem 4.1. The proof follows that of Section 4 with the analogue [Bre04, Eqn (1.18)] of Lemma 4.7.

5.2 Arnold-Winther FEM for $k \geq 3$

For $k \geq 3$, the analogues to the volume degrees of freedom from Section A.1.2 read $⨍_T \tau_{\mathrm{AW}} : \phi\,dx$ for $\phi \in \varepsilon(P_k(T;\mathbb{R}^n))$. For $k \geq 3$ it holds $\varepsilon(P_k(T;\mathbb{R}^n)) \subsetneq P_{k-1}(T;\mathbb{S})$. For a shape function σ_{AW} of a face degree of freedom with $⨍_T \sigma_{\mathrm{AW}} : \phi\,dx = 0$ for all $\phi \in \varepsilon(P_k(T;\mathbb{R}^n))$, it is not obvious that the term $⨍_{\widehat{T}} \widehat{\sigma}_{\mathrm{AW}} : \widehat{\phi}\,dx$ for $\widehat{\phi} \in \varepsilon(P_k(\widehat{T};\mathbb{R}^n))$ scales in the correct way or even vanishes. This disables the analysis of Theorem A.7 and, hence, $k \geq 3$ is excluded in Theorem 4.1.

5.3 Nonconforming Arnold-Winther FEM

This subsection supports the conjecture that the analysis of Theorem 4.1 does not hold for the nonconforming Arnold-Winther FEM. The nonconforming symmetric finite element spaces of [AW03] for the pure Dirichlet problem and $n = 2$ for the stresses read

$$X_{\mathrm{NC}} := \left\{ \tau \in P_2(\mathcal{T};\mathbb{S}) \;\middle|\; \begin{array}{l} \forall F \in \mathcal{F}, \nu_F \cdot (\tau\nu_F) \in P_1(F) \text{ and} \\ \forall F \in \mathcal{F}\ \forall v \in P_1(F;\mathbb{R}^2), \int_F v \cdot [\tau]_F \nu_F \, ds = 0 \end{array} \right\}$$

and $b_{\mathrm{NC}}(\tau_{\mathrm{NC}}, v_h) := (\operatorname{div}_{\mathrm{NC}} \tau_{\mathrm{NC}}, v_h)_{L^2(\Omega)}$ for $\tau_{\mathrm{NC}} \in X_{\mathrm{NC}}$ and $v_h \in P_1(\mathcal{T};\mathbb{R}^2)$. The discrete mixed system seeks $(\sigma_{\mathrm{NC}}, u_k) \in X_{\mathrm{NC}} \times P_1(\mathcal{T};\mathbb{R}^2)$ such that

$$\begin{aligned} a(\sigma_{\mathrm{NC}}, \tau_{\mathrm{NC}}) + b_{\mathrm{NC}}(\tau_{\mathrm{NC}}, u_k) &= 0 && \text{for all } \tau_{\mathrm{NC}} \in X_{\mathrm{NC}}, \\ b_{\mathrm{NC}}(\sigma_{\mathrm{NC}}, v_h) &= -(f, v_h)_{L^2(\Omega)} && \text{for all } v_h \in P_1(\mathcal{T};\mathbb{R}^2). \end{aligned} \tag{5.1}$$

The stability of the discrete system (5.1) with respect to the norms $\|\cdot\|_{L^2(\Omega)}$ and $|\!|\!|\cdot|\!|\!|_h$ follows from the arguments in Section 4. The additional term from (2.7) reads $|a(\sigma, \rho_h) + b_{\mathrm{NC}}(\rho_h, u)|$. This additional term disables a best-approximation result as in Theorem 4.1 for the nonconforming Arnold-Winther FEM. This is in accordance with the numerical experiments of [CEG11] where it was observed that the approximation of σ by σ_{NC} is only of first order. The same observation on reduced convergence rates is also valid for the reduced nonconforming finite element of [AW03].

5.4 Nonconforming Treatment of Neumann Data

Although the proof of (E) in Subsection 4.4 requires only the condition (N) on the Neumann data, the abstract result in Theorem 2.1 relies the conformity in the strong form $X_h(0) \subseteq X(0)$. One possibility of circumventing the nodal interpolation of possibly inconsistent Neumann data is to fix only the moments of order k of $\tau_{\mathrm{AW}}\nu|_F$ for all $F \in \mathcal{F}(\Gamma_N)$ as in [CGRT08]

$$\begin{aligned} \mathrm{AW}_{k,\mathrm{NC}}(\mathcal{T}, g) = \{\tau_{\mathrm{AW}} \in H(\operatorname{div}, \Omega; \mathbb{S}) \mid \forall T \in \mathcal{T},\ \tau_{\mathrm{AW}} \in \mathrm{AW}_k(T) \text{ and} \\ \forall v_k \in P_k(\mathcal{F}(\Gamma_N);\mathbb{R}^n),\ \int_{\Gamma_N} (g - \tau_{\mathrm{AW}}\nu) \cdot v_k \, ds = 0\}. \end{aligned}$$

The discrete problem seeks $\sigma_{\mathrm{AW}} \in \mathrm{AW}_{k,\mathrm{NC}}(\mathcal{T}, g)$ and $u_k \in P_k(\mathcal{T};\mathbb{R}^n)$ such that

$$\begin{aligned} a(\sigma_{\mathrm{AW}}, \tau_{\mathrm{AW}}) + b(\tau_{\mathrm{AW}}, u_k) &= \langle \tau_{\mathrm{AW}}\nu, u_D \rangle_{\partial\Omega} && \text{for all } \tau_{\mathrm{AW}} \in \mathrm{AW}_{k,\mathrm{NC}}(\mathcal{T}, 0), \\ b(\sigma_{\mathrm{AW}}, v_k) &= -(f, v_k)_{L^2(\Omega)} && \text{for all } v_k \in P_k(\mathcal{T};\mathbb{R}^n). \end{aligned}$$

The conditions (LBB), (C), and (E) can be verified with the methodology of Section 4. However, since $\mathrm{AW}_{k,\mathrm{NC}}(\mathcal{T}, 0) \not\subseteq X(0)$, a direct application of Theorem 2.1 is not possible. The following extension explains the sub-optimal convergence observed in the numerical experiments of rates in [CGRT08, Subsection 3.2]. For a smooth solution $u \in H^2(\Omega;\mathbb{R}^n)$ (resp. $H^3(\Omega;\mathbb{R}^n)$), the approximation error of the theorem below is merely $\mathcal{O}(h)$ (resp. $\mathcal{O}(h^{3/2})$) in accordance with all numerical experiments in [CGRT08].

Theorem 5.1. *Any $\tau_{\mathrm{AW}} \in \mathrm{AW}_{k,\mathrm{NC}}(\mathcal{T}, 0)$ and any $v_1 \in P_1(\mathcal{F}(\Gamma_N);\mathbb{R}^n)$ satisfy*

$$\|\sigma - \sigma_{\mathrm{AW}}\|_{L^2(\Omega)} \lesssim \|\sigma - \tau_{\mathrm{AW}}\|_{L^2(\Omega)} + \operatorname{osc}_k(f, \mathcal{T}) + \|h_{\mathcal{T}}^{-1/2}(u - v_1)\|_{L^2(\Gamma_N)}.$$

Proof. As pointed out in Remark 2.3, the nonconformity in the discretisation of the Neumann data results in the additional term (2.7) which reads (with $\xi_h = \rho_{\mathrm{AW}}$)

$$a(\sigma, \rho_{\mathrm{AW}}) + b(\rho_{\mathrm{AW}}, u)$$

for some $\rho_{\mathrm{AW}} \in \mathrm{AW}_{k,\mathrm{NC}}(\mathcal{T}, 0)$ with $\|\rho_{\mathrm{AW}}\|_{L^2(\Omega)} \leq 1$. The integration by parts and the boundary condition of ρ_{AW} show for any $v_1 \in P_1(\mathcal{F}(\Gamma_N); \mathbb{R}^2)$ that this equals

$$a(\sigma, \rho_{\mathrm{AW}}) + b(\rho_{\mathrm{AW}}, u) = \int_{\Gamma_N} u \cdot (\rho_{\mathrm{AW}} \nu)\, ds = \int_{\Gamma_N} (u - v_1) \cdot (\rho_{\mathrm{AW}} \nu)\, ds.$$

A trace and an inverse inequality show that this is controlled as

$$a(\sigma, \rho_{\mathrm{AW}}) + b(\rho_{\mathrm{AW}}, u) \lesssim \|h_{\mathcal{T}}^{-1/2}(u - v_1)\|_{L^2(\Gamma_N)}\, \|\rho_{\mathrm{AW}}\|_{L^2(\Omega)}.$$

The remaining terms are analysed in the proof of Theorem 2.1. Hence, further details are omitted. □

5.5 Equality of Approximation Classes

The notion of optimality of adaptive FEM in the literature is based on the concept of an approximation class [CKNS08, BDD04]. Given some $s > 0$ and an initial regular triangulation $\mathcal{T}_0$, the set $\mathbb{T}(N)$ of admissible triangulations $\mathcal{T}$ with $\mathrm{card}(\mathcal{T}) \leq \mathrm{card}(\mathcal{T}_0) + N$ leads to the following seminorms

$$|(\sigma, f)|_{\mathcal{A}_s,\mathrm{AWFEM}} := \sup_{N\in\mathbb{N}} \min_{\mathcal{T}\in\mathbb{T}(N)} N^s \sqrt{\|\sigma - \sigma_{\mathrm{AW}}\|^2 + \mathrm{osc}(f, \mathcal{T})^2}, \tag{5.2}$$

$$|(\sigma, f)|_{\mathcal{A}_s,\mathrm{bapx}} := \sup_{N\in\mathbb{N}} \min_{\mathcal{T}\in\mathbb{T}(N)} N^s \min_{\tau_{\mathrm{AW}}\in \mathrm{AW}(\mathcal{T}, g_{n+k})} \sqrt{\|\sigma - \tau_{\mathrm{AW}}\|^2 + \mathrm{osc}(f, \mathcal{T})^2}$$

with the approximation classes

$$\mathcal{A}_{s,g_{k+n}}^{\mathrm{AWFEM}} := \big\{(\sigma, f) \in H(\mathrm{div}, \Omega; \mathbb{S}) \times L^2(\Omega; \mathbb{R}^n)\, \big|\, |(\sigma, f)|_{\mathcal{A}_s,\mathrm{AWFEM}} < \infty \big\},$$
$$\mathcal{A}_{s,g_{k+n}}^{\mathrm{bapx}} := \big\{(\sigma, f) \in H(\mathrm{div}, \Omega; \mathbb{S}) \times L^2(\Omega; \mathbb{R}^n)\, \big|\, |(\sigma, f)|_{\mathcal{A}_{s,\mathrm{bapx}}} < \infty \big\}.$$

The set $\mathcal{A}_{s,g_{k+n}}^{\mathrm{AWFEM}}$ concerns the approximation by the Arnold-Winther FEM and σ_{AW} in (5.2) is the finite element solution with respect to the mesh $\mathcal{T}$ and the data f, while $\mathcal{A}_{s,g_{k+n}}^{\mathrm{bapx}}$ describes the approximability of σ by arbitrary functions in $\mathrm{AW}_k(\mathcal{T}, g_{n+k})$ (independent of any scheme). The approximation properties of functions in $\mathrm{AW}_k(\mathcal{T}, g_{n+k})$ may be sensitive to the utilised approximation of the Neumann data g by g_{n+k}.

Theorem 5.2 ($\mathcal{A}_s^{\mathrm{bapx}} = \mathcal{A}_s^{\mathrm{AWFEM}}$)**.** *Given any $f \in L^2(\Omega; \mathbb{R}^n)$, $g \in L^2(\Gamma_N; \mathbb{R}^n)$ and $u_D \in H^1(\Omega; \mathbb{R}^n)$ with the exact solution (σ, u) to (3.1), then, for any $s > 0$,*

$$(\sigma, f) \in \mathcal{A}_s^{\mathrm{AWFEM}} \quad \textit{if and only if} \quad (\sigma, f) \in \mathcal{A}_s^{\mathrm{bapx}}$$

and $|(\sigma, f)|_{\mathcal{A}_s,\mathrm{AWFEM}} \approx |(\sigma, f)|_{\mathcal{A}_{s,\mathrm{bapx}}}$. The equivalence constants depend on the domain Ω and on the shape-regularity of the triangulation $\mathcal{T}$.

Proof. This is an immediate consequence of the best-approximation result of Theorem 4.1. □

5.6 Stokes Equations

Theorem 2.1 immediately applies to the Arnold-Winther FEM of [CGP12] for the Stokes equations which corresponds to the formal limit $\lambda \to \infty$ with $\mathbb{C}^{-1}\sigma$ replaced by the deviatoric part $\operatorname{dev}\sigma$. This section therefore focuses on the pseudostress-velocity formulation of the Stokes problem of [CTVW10, CW07, CWZ10, CKP11],

$$\operatorname{div}\sigma + f = 0 \quad \text{and} \quad \operatorname{dev}\sigma = Du \text{ in } \Omega, \quad u = u_D \text{ on } \Gamma_D.$$

This leads to (2.1) with the spaces

$$\begin{aligned} H &:= L^2(\Omega;\mathbb{R}^{n\times n}),\\ X &:= \{\tau \in H(\operatorname{div},\Omega;\mathbb{R}^{n\times n}) \mid \textstyle\int_\Omega \operatorname{tr}\tau\,dx = 0\},\\ Y &:= L^2(\Omega;\mathbb{R}^n). \end{aligned}$$

For any $\tau,\rho \in X$ and $v \in Y$, the bilinear forms read

$$a(\tau,\rho) := (\operatorname{dev}\tau,\rho)_{L^2(\Omega)} \quad \text{and} \quad b(\tau,v) := (\operatorname{div}\tau,v)_{L^2(\Omega)}.$$

Given $u_D \in H^1(\Omega;\mathbb{R}^n)$ and $f \in L^2(\Omega;\mathbb{R}^n)$, set $G(\tau) := \langle \tau\nu, u_D\rangle_{\partial\Omega}$ and $F(v) := \int_\Omega f\cdot v\,dx$. It is well-established that the system (2.1) has a unique solution $(\sigma,u) \in X\times Y$. With the well-known Raviart-Thomas finite element space of degree $k \geq 0$ from [BF91, Bra07, BS08] let $X_h := \mathrm{RT}_k(\mathcal{T})^n \cap X$ and $Y_h := P_k(\mathcal{T};\mathbb{R}^n)$.

Theorem 5.3 (L^2 best-approximation for the Stokes equations)**.** *The pseudostress $\sigma \in X$ and its approximation $\sigma_h \in X_h$ satisfy the L^2 error estimate*

$$\|\sigma-\sigma_h\|_{L^2(\Omega)} \lesssim \inf_{\tau_h\in X_h}\|\sigma-\tau_h\|_{L^2(\Omega)} + \operatorname{osc}_k(f,\mathcal{T}).$$

Proof. The proof follows from (LBB), (C) and (E) and Theorem 2.1 for the norms $\|\bullet\|_H = \|\bullet\|_{L^2(\Omega)}$ and $|\!|\!|\bullet|\!|\!|_h$ given, for any $v_h \in Y_h$, by

$$|\!|\!|v_h|\!|\!|_h^2 := \|D_{\mathrm{NC}}(v_h)\|_{L^2(\Omega)}^2 + \sum_{F\in\mathcal{F}(\Omega\cup\Gamma_D)} h_F^{-1}\|[v_h]_F\|_{L^2(F)}^2. \tag{5.3}$$

The ellipticity condition (LBB1a) is a direct consequence of Lemma 4.2; the continuity (LBB1b) follows from the Cauchy inequality. For any $\tau_h \in X_h$ and $v_h \in Y_h$, a piecewise integration by parts proves that

$$b(\tau_h,v_h) = -(\tau_h, D_{\mathrm{NC}}v_h)_{L^2(\Omega)} + \sum_{F\in\mathcal{F}(\Omega\cup\Gamma_D)}\int_F [v_h]_F\cdot(\tau_h\nu_F)\,ds.$$

As in the proof of Lemma 4.4, the trace and inverse inequalities control the term $b(\tau_h,v_h)$ by $\|\tau_h\|_{L^2(\Omega)}\ |\!|\!|v_h|\!|\!|_h$. This leads to the proof of the continuity (LBB2a).

	$n=2$, $k=1$	$n=2$, $k=2$	$n=3$, $k=1$	$n=3$, $k=2$
nodal dofs	9	9	24	24
volume dofs	3	9	6	24
face dofs	12	18	36	72
add. volume dofs	-	1	6	21
add. edge dofs	-	-	90	120
sum	24	37	162	261

Table A.1: The numbers of degrees of freedom.

For the proof of the inf-sup condition (LBB2b), it is stated in [BV96, Lemma 2.1], [LS06, Lemma 2.1] that for any $v_h \in Y_h$ there exists some $\tilde\tau_h \in \mathrm{RT}_k(\mathcal{T})^n$ with

$$b(\tilde\tau_h, v_h) = \| v_h \|_h^2 \quad \text{and} \quad \|\tau_h\|_{L^2(\Omega)} \lesssim \| v_h \|_h.$$

Set $\tau_h := (\tilde\tau_h - n^{-1}(\int_\Omega \operatorname{tr}\tau_h\, dx)\, 1_{n\times n}) \in X_h$ and observe that $b(\tau_h, v_h) = \| v_j \|_h^2$ and $\|\tau_h\|_{L^2(\Omega)} \leq \|\tilde\tau_h\|_{L^2(\Omega)}$. This establishes the inf-sup condition (LBB2b). The bubble function technique due to [Ver96] allows the proof of the efficiency estimate

$$\|h_{\mathcal{T}}(f + \operatorname{div}\tau_h)\|_{L^2(\Omega)} \lesssim \|\sigma - \tau_h\|_{L^2(\Omega)} + \mathrm{osc}_k(f, \mathcal{T})$$

for all $\tau_h \in X_h$. The techniques of Lemmas 4.8–4.9 prove condition (E). □

Theorem 5.3 allows the following refinement of [CGS13, Thm. 3.5] where $n = 2$ and $k = 0$: That theorem holds for a nonlinear approximation class (in the notation of [CGS13])

$$\begin{aligned}\mathcal{A}_s'' := \big\{(\sigma, f, g) \in & H(\operatorname{div}, \Omega; \mathbb{R}^{2\times 2})/\mathbb{R} \times L^2(\Omega;\mathbb{R}^2) \\ & \times \big(H^1(\Omega;\mathbb{R}^2) \cap H^1(\mathcal{E}(\partial\Omega);\mathbb{R}^2)\big) \;\big|\; |(\sigma, f, g)|_{\mathcal{A}_s''} < \infty\big\}\end{aligned}$$

with $|(\sigma, f, g)|_{\mathcal{A}_s''} :=$

$$\sup_{N\in\mathbb{N}} N^s \inf_{\mathcal{T}\in\mathbb{T}(N)} \inf_{\tau_{\mathrm{PS}}\in\mathrm{PS}(\mathcal{T})} \left(\|\operatorname{dev}(\sigma - \tau_{\mathrm{PS}})\|_{L^2(\Omega)}^2 + \mathrm{osc}^2(f, \mathcal{T}) + \mathrm{osc}^2\left(\frac{\partial g}{\partial s}, \mathcal{E}(\partial\Omega)\right)\right)^{1/2}.$$

The refinement is that τ_{PS} in the preceding definition of the seminorm is optimal in the L^2 norm while in [CGS13, Thm. 3.5] it is the finite element solution with respect to $\mathcal{T}$.

A Local Degrees of Freedom of AW-MFEM

This appendix discusses the scaling of the shape functions of the Arnold-Winther-FEM. Subsection A.1 first recalls the local degrees of freedom.

A.1 Definition of the Local Degrees of Freedom

This subsection recalls the local degrees of freedom for the two-dimensional case $n = 2$ and the three-dimensional case $n = 3$ from [AW02, AAW08]. There it is proven that the local degrees of freedom are linearly independent functionals on the space $\mathrm{AW}_k(T)$. The number of the different degrees of freedom can be found in Table A.1.

Let $e_j := (\delta_{jk} | k \in \{1, \dots, n\}) \in \mathbb{R}^n$ denote the canonical unit vectors for $j \in \{1, \dots, n\}$ and set $e_{jk} := e_j \otimes e_k := e_j e_k^\top \in \mathbb{R}^{n\times n}$. The indices for the upper triangle part of an $n \times n$ matrix read as

$$\mathcal{H} := \{(j, k) \in \{1, \dots, n\}^2 \mid j \leq k\}.$$

For $T \in \mathcal{T}$, the local degrees of freedom are defined via the following groups of linear functionals for any $\tau_{\mathrm{AW}} \in \mathrm{AW}_k(T)$.

A.1.1 Nodal Degrees of Freedom

The first group of degrees of freedom are the nodal values of τ_{AW}. Define the index set $I_{\mathrm{nodes}}(T) := \mathcal{N}(T) \times \mathcal{H}$. The functional $L_{T,\ell}$ is defined by

$$L_{T,\ell}(\tau_{\mathrm{AW}}) := \tau_{\mathrm{AW}}(z) : e_{\alpha\beta} \qquad \text{for } \ell = (z, (\alpha, \beta)) \in I_{\mathrm{nodes}}(T).$$

A.1.2 Volume Degrees of Freedom

The second group of degrees of freedom concerns volume moments. Define

$$I_{\mathrm{vol},1}(T) := \begin{cases} \{0\} \times \mathcal{H} & \text{if } k = 1, \\ \{0, \dots, n\} \times \mathcal{H} & \text{if } k = 2. \end{cases}$$

For $T = \mathrm{conv}\{z_1, \dots, z_{n+1}\}$ the moments of degree zero read

$$L_{T,\ell}(\tau_{\mathrm{AW}}) := e_{\alpha\beta} : \fint_T \tau_{\mathrm{AW}}\, dx \qquad \text{for } \ell = (0, (\alpha, \beta)) \in I_{\mathrm{vol},1}(T)$$

and for $k = 2$ and $1 \leq j \leq n$, the moments of degree one read

$$L_{T,\ell}(\tau_{\mathrm{AW}}) := e_{\alpha\beta} : \fint_T (\varphi_{z_j} - \varphi_{z_{n+1}})\, \tau_{\mathrm{AW}}\, dx \qquad \text{for } \ell = (j, (\alpha, \beta)) \in I_{\mathrm{vol},1}(T).$$

Remark A.1. The volume degrees of freedom for $k \geq 2$ are defined in [AW02, AAW08] as the moments $\int_T \tau_{\mathrm{AW}} : \phi\, dx$ for $\phi \in \varepsilon(P_k(T; \mathbb{R}^n))$. Note that for $k = 2$ it holds $\varepsilon(P_2(T; \mathbb{R}^n)) = P_1(T; \mathbb{S})$ on each simplex T.

A.1.3 Face Degrees of Freedom

The third group of functionals consists of the face moments of τ_{AW}. Define

$$I_{\mathrm{faces}}(T) := \left\{ (F, \eta, \alpha) \in \mathcal{F}(T) \times \mathbb{N}_0^{n-1} \times \{1, \dots, n\} \;\middle|\; \sum_{j=1}^{n-1} \eta_j \leq k \right\}.$$

For $F = \mathrm{conv}\{z_1, \dots, z_n\}$ the moments of degree $0 \le \sum_{j=1}^{n-1} \eta_j \le k$ read

$$L_{T,\ell}(\tau_{\mathrm{AW}}) := e_\alpha \cdot \left(\fint_F \prod_{j=1}^{n-1} (\varphi_{z_j} - \varphi_{z_n})^{\eta_j} \tau_{\mathrm{AW}} \, ds \right) \nu_F, \text{ for } \ell = (F, \eta, \alpha) \in I_{\mathrm{faces}}(T).$$

A.1.4 Additional Volume Degrees of Freedom

Let $I_{\mathrm{vol},2}(T) := \{1, \dots, \dim(M_k(T))\}$ (with the convention $\{1, \dots, 0\} := \emptyset$) and

$$M_k(T) := \{\tau \in P_{k+n}(T; \mathbb{S}) \mid \operatorname{div} \tau = 0 \text{ and } \tau \nu_T = 0 \text{ on } \partial T\}.$$

In the lowest-order case $k = 1$ for $n = 2$, $M_k(T) = \{0\}$ [AW02, AAW08]. Let $(\phi_j)_{j \in I_{\mathrm{vol},2}}$ be a basis of $M_k(T)$ and define the functionals

$$L_{T,j}(\tau_{\mathrm{AW}}) := \fint_T \tau_{\mathrm{AW}} : \phi_j \, dx \qquad \text{for } j \in I_{\mathrm{vol},2}(T).$$

A.1.5 Additional Edge Degrees of Freedom

There exists additional degrees of freedom for the edges of T for $n = 3$. Let $\mathcal{E}(T)$ denote the set of edges of T and define

$$\begin{aligned} I_{\mathrm{edges}}(T) :=& \big\{ (E, j, \alpha, \beta) \in \mathcal{E}(T) \times \{0, \dots, k+1\} \times \{1,2\}^2 \bigm| (\alpha, \beta) \neq (2,1) \big\} \\ & \cup \big(\mathcal{F}(T) \times \mathcal{E}(T) \times \{0, \dots, k+1\} \big). \end{aligned}$$

For an edge $E = \mathrm{conv}\{a, b\}$ of $T \subseteq \mathbb{R}^3$ let ν_1 and ν_2 be linearly independent normal vectors of E. The moments of degree j of the normal directions read

$$L_{T,\ell}(\tau_{\mathrm{AW}}) := \fint_E (\varphi_a - \varphi_b)^j \, \nu_\alpha^\top \tau_{\mathrm{AW}} \, \nu_\beta \, ds \qquad \text{for } \ell = (E, j, \alpha, \beta) \in I_{\mathrm{edges}}(T).$$

For a face $F \in \mathcal{F}(T)$ with edge E, let s_E denote a unit tangent along E. The moments of degree j of the tangent-normal directions read

$$L_{T,\ell}(\tau_{\mathrm{AW}}) := \fint_E (\varphi_a - \varphi_b)^j \, s_E^\top \tau_{\mathrm{AW}} \, \nu_F \, ds \qquad \text{for } \ell = (F, E, j) \in I_{\mathrm{edges}}(T).$$

A.2 Scaling of Shape Functions

This subsection studies the dependence of the L^2 norm of shape functions of $\mathrm{AW}_k(T)$ on the mesh-size. Let $T \in \mathcal{T}$ and $\widehat{T}$ some reference simplex and let $\Psi : \widehat{T} \to T$ be an affine transformation with $\Psi(x) = Bx + b$. This transformation does not map $\mathrm{AW}_k(T)$ to $\mathrm{AW}_k(\widehat{T})$ in general. Therefore, the Piola transform is employed for $B^{-\top} := (B^\top)^{-1}$: Given $\sigma_{\mathrm{AW}} \in \mathrm{AW}_k(T)$, its Piola transformation reads

$$\widehat{\sigma}_{\mathrm{AW}} = B^{-1} (\sigma_{\mathrm{AW}} \circ \Psi) B^{-\top}.$$

The following Lemmas A.2–A.6 discuss the transformation of the local degrees of freedom under the Piola transform.

Lemma A.2. *Let $z \in \mathcal{N}(T)$ be a vertex of the simplex T. If $L_{T,\ell}(\sigma_{\mathrm{AW}}) = 0$ for all $\ell \in I_{\mathrm{nodes}}$, then $L_{\widehat{T},\ell}(\widehat{\sigma}_{\mathrm{AW}}) = 0$ for all $\ell \in I_{\mathrm{nodes}}(\widehat{T})$.*

Proof. This follows from $(\sigma_{\mathrm{AW}} \circ \Psi)(\Psi^{-1}(z)) = 0$. □

Lemma A.3. *Given $(j,(\alpha,\beta)) \in I_{\mathrm{vol},1}(\widehat{T})$, the transformation of the volume degrees of freedom reads*

$$L_{\widehat{T},(j,(\alpha,\beta))}(\widehat{\sigma}_{\mathrm{AW}}) = e_{\alpha\beta} : \Big(B^{-1}\Big(\sum_{(\gamma,\delta)\in\mathcal{H}} L_{T,(j,(\gamma,\delta))}(\sigma_{\mathrm{AW}}) e_{\gamma\delta}\Big)B^{-\top}\Big).$$

Proof. This follows from a transformation to T. □

Lemma A.4. *The transformation of the face degrees of freedom for $F = \Psi(\widehat{F})$ and $(\widehat{F},\eta,\alpha) \in I_{\mathrm{faces}}(\widehat{T})$ reads*

$$L_{\widehat{T},(\widehat{F},\eta,\alpha)}(\widehat{\sigma}_{\mathrm{AW}}) = |B^\top \nu_F|^{-1} e_\alpha \cdot \left(B^{-1}\sum_{\gamma=1}^{n} L_{T,(F,\eta,\gamma)}(\sigma_{\mathrm{AW}}) e_\gamma\right).$$

Proof. The transformation to $F = \mathrm{conv}\{z_1,\dots,z_n\}$ yields

$$L_{\widehat{T},(\widehat{F},\eta,\alpha)}(\widehat{\sigma}_{\mathrm{AW}}) = e_\alpha \cdot \left(\fint_F \prod_{j=1}^{n-1}(\varphi_{z_j} - \varphi_{z_n})^{\eta_j}(B^{-1}\sigma_{\mathrm{AW}})\,ds\right) B^{-\top}\nu_{\widehat{F}}.$$

Since $B^{-\top}\nu_{\widehat{F}}$ is orthogonal to F and $|B^\top \nu_F| = |B^{-\top}\nu_{\widehat{F}}|^{-1}$, this leads to

$$\begin{aligned} L_{\widehat{T},(\widehat{F},\eta,\alpha)}(\widehat{\sigma}_{\mathrm{AW}}) &= |B^\top \nu_F|^{-1}\, e_\alpha \cdot \left(\fint_F \prod_{j=1}^{n-1}(\varphi_{z_j} - \varphi_{z_n})^{\eta_j}(B^{-1}\sigma_{\mathrm{AW}})\,ds\right)\nu_F \\ &= |B^\top \nu_F|^{-1}\, e_\alpha \cdot \left(B^{-1}\left(\fint_F \prod_{j=1}^{n-1}(\varphi_{z_j} - \varphi_{z_n})^{\eta_j}\sigma_{\mathrm{AW}}\,ds\right)\nu_F\right) \\ &= |B^\top \nu_F|^{-1} e_\alpha \cdot \left(B^{-1}\sum_{\gamma=1}^{n} L_{T,(F,\eta,\gamma)}(\sigma_{\mathrm{AW}}) e_\gamma\right). \end{aligned}$$

□

Lemma A.5. *If $L_{T,j}(\sigma_{\mathrm{AW}}) = 0$ for all $j \in I_{\mathrm{vol},2}(T)$, then $L_{\widehat{T},j}(\widehat{\sigma}_{\mathrm{AW}}) = 0$ for all $j \in I_{\mathrm{vol},2}(\widehat{T})$.*

Proof. Let $\widehat{\phi}_j$ denote the basis function of $M_k(\widehat{T})$ associated with $L_{\widehat{T},j}$. A transformation and a calculation reveal

$$L_{\widehat{T},\mathrm{vol},j}(\widehat{\sigma}_{\mathrm{AW}}) = \fint_T \sigma_{\mathrm{AW}} : (B^{-\top}(\widehat{\phi}_j \circ \Psi^{-1})B^{-1})\,dx.$$

A further calculation reveals $\mathrm{div}((B^{-\top}(\widehat{\phi}_j \circ \Psi^{-1})B^{-1})) = 0$ if $\mathrm{div}\,\phi_j = 0$ [AW02, AAW08]. Let $F \in \mathcal{F}(T)$ be a face of T and $\widehat{F} = \Psi^{-1}(F) \in \mathcal{F}(\widehat{T})$ with normal $\widehat{\nu}_{\widehat{F}}$. Since $B^{-1}\widehat{\nu}_{\widehat{F}}$ is orthogonal to F, the symmetry of $(B^{-\top}(\widehat{\phi}_j \circ \Psi^{-1})B^{-1})$ proves $(B^{-\top}(\widehat{\phi}_j \circ \Psi^{-1})B^{-1}) \in M_k(T)$. This proves $L_{\widehat{T},j}(\widehat{\sigma}_{\mathrm{AW}}) = 0$. □

The following lemma concerns only the case $n = 3$.

Lemma A.6. *If $n = 3$ and $L_{T,\ell}(\sigma_{\mathrm{AW}}) = 0$ for all $\ell \in I_{\mathrm{edges}}(T)$, then $L_{\widehat{T},\ell}(\widehat{\sigma}_{\mathrm{AW}}) = 0$ for all $\ell \in I_{\mathrm{edges}}(\widehat{T})$.*

Proof. Let $(E, j, \alpha, \beta) \in I_{\mathrm{edges}}(T)$ and $\widehat{E} = \Psi^{-1}(E) = \mathrm{conv}\{c, d\}$ with normals $\widehat{\nu}_1, \widehat{\nu}_2$ associated to the edge degrees of freedom of $\widehat{T}$. The definition of the Piola transform and a transformation yield

$$\begin{aligned} L_{\widehat{T},(\Psi^{-1}(E),j,\alpha,\beta)}(\widehat{\sigma}_{\mathrm{AW}}) &= \fint_E (\varphi_{\Psi(c)} - \varphi_{\Psi(d)})^j \widehat{\nu}_\alpha^\top (B^{-1}(\sigma_{\mathrm{AW}} \circ \Psi) B^{-\top}) \widehat{\nu}_\beta \, ds \\ &= \fint_{\widehat{E}} (\varphi_{\Psi(c)} - \varphi_{\Psi(d)})^j (B^{-\top}\widehat{\nu}_\alpha)^\top \sigma_{\mathrm{AW}} (B^{-\top}\widehat{\nu}_\beta) \, ds. \end{aligned}$$

Since $B^{-\top}\widehat{\nu}_\gamma$ is orthogonal to E for $\gamma = \alpha, \beta$ and $L_{T,(E,j,\alpha,\beta)}(\sigma_{\mathrm{AW}}) = 0$, this vanishes. The same arguments prove (with the normal $\widehat{\nu}_{\Psi^{-1}(F)}$ of $\Psi^{-1}(F)$ and the tangent $\widehat{s}_{\widehat{E}}$ of $\widehat{E}$)

$$\begin{aligned} &L_{\widehat{T},(\Psi^{-1}(F),\Psi^{-1}(E),j)}(\widehat{\sigma}_{\mathrm{AW}}) \\ &\qquad = \fint_{\widehat{E}} (\varphi_{\Psi(c)} - \varphi_{\Psi(d)})^\ell (B^{-\top}\widehat{s}_{\widehat{E}})^\top \sigma_{\mathrm{AW}} (B^{-\top}\widehat{\nu}_{\Psi^{-1}(F)}) \, ds. \end{aligned} \tag{A.1}$$

Since $B^{-\top}\nu_F$ is orthogonal to $\Psi^{-1}(F)$ and $B^{-\top}\widehat{s}_{\widehat{E}} \in \mathrm{span}\{s_E, \nu_1, \nu_2\}$, (A.1) vanishes as well. □

For the remaining part of this section, let $\sigma_{\mathrm{AW}} \in \mathrm{AW}_k(T)$ be a shape function of a volume degree of freedom and $\tau_{\mathrm{AW}} \in \mathrm{AW}_k(T)$ a shape function of a face degree of freedom for some simplex $T \in \mathcal{T}$, i.e., there exists $\ell_0 \in I_{\mathrm{vol},1}(T)$ and $\ell_1 \in I_{\mathrm{faces}}(T)$ with

$$L_{T,\ell}(\sigma_{\mathrm{AW}}) = \delta_{\ell_0 \ell} \qquad \text{and} \qquad L_{T,\ell}(\tau_{\mathrm{AW}}) = \delta_{\ell_1 \ell}$$

for $\ell \in I_{\mathrm{nodes}}(T) \cup I_{\mathrm{vol},1}(T) \cup I_{\mathrm{faces}}(T) \cup I_{\mathrm{vol},2}(T) \cup I_{\mathrm{edges}}(T)$. The following theorem proves that the shape functions scale in the expected way.

Theorem A.7. *The above degrees of freedom with the aforementioned shape functions σ_{AW} and τ_{AW} on $T \in \mathcal{T}$ satisfy*

$$\|\sigma_{\mathrm{AW}}\|_{L^2(T)} + \|\tau_{\mathrm{AW}}\|_{L^2(T)} \lesssim |T|^{1/2}.$$

Proof. Let $\widehat{\sigma}_{\mathrm{AW}}$ and $\widehat{\tau}_{\mathrm{AW}}$ as above denote the Piola transform of σ_{AW} and τ_{AW}. Let $\sigma_\ell \in \mathrm{AW}_k(\widehat{T})$ for $\ell \in I_{\mathrm{vol},1}(\widehat{T}) \cup I_{\mathrm{faces}}(\widehat{T})$ denote the shape functions of the volume degrees of freedom and the face degrees of freedom on $\widehat{T}$. Since the Piola transform preserves the polynomial degree and the symmetry of σ_{AW} and τ_{AW} and the polynomial degree of their divergence, $\widehat{\sigma}_{\mathrm{AW}} \in \mathrm{AW}_k(\widehat{T})$

and $\widehat{\tau}_{\mathrm{AW}} \in \mathrm{AW}_k(\widehat{T})$ are Arnold-Winther functions on the reference simplex. Lemmas A.2–A.6 then prove

$$\begin{aligned}\widehat{\sigma}_{\mathrm{AW}} &= \sum_{\ell\in I_{\mathrm{vol},1}(\widehat{T})} L_{\widehat{T},\ell}(\widehat{\sigma}_{\mathrm{AW}})\sigma_\ell \\ &= \sum_{\ell=(j,(\alpha,\beta))\in I_{\mathrm{vol},1}(\widehat{T})} \left(e_{\alpha\beta} : \left(B^{-1}\left(\sum_{(\gamma,\delta)\in\mathcal{H}} L_{T,(j,(\gamma,\delta))}(\sigma_{\mathrm{AW}})e_{\gamma\delta}\right)B^{-\top}\right)\right)\sigma_\ell.\end{aligned}$$

This and $\|\sigma_\ell\|_{L^2(\widehat{T})} \lesssim 1$ yield

$$\|\widehat{\sigma}_{\mathrm{AW}}\|_{L^2(\widehat{T})} \lesssim |B^{-1}|^2.$$

The combination with $|B| \approx |B^{-1}|^{-1}$ plus a transformation lead to

$$\begin{aligned}\|\sigma_{\mathrm{AW}}\|_{L^2(T)} &= |T|^{1/2}\,\|\sigma_{\mathrm{AW}}\circ\Psi\|_{L^2(\widehat{T})} \\ &= |T|^{1/2}\,\|B\,\widehat{\sigma}_{\mathrm{AW}}\,B^\top\|_{L^2(\widehat{T})} \lesssim |T|^{1/2}.\end{aligned}$$

The same arguments and $|B^\top\nu_F|^{-1} \lesssim |B^{-1}|$ prove $\|\tau_{\mathrm{AW}}\|_{L^2(T)} \lesssim |T|^{1/2}$. □

Remark A.8. Theorem A.7 can alternatively be proven by the explicit computation of the local mass matrix in [CGRT08] for $n = 2$ and $k = 1$.

Acknowledgements

This work was supported by the DFG Research Center Matheon and the Berlin Mathematical School. The authors thank Professor Dietrich Braess for valuable comments.

References

[AAW08] Douglas N. Arnold, Gerard Awanou, and Ragnar Winther. Finite elements for symmetric tensors in three dimensions. *Math. Comp.*, 77(263):1229–1251, 2008.

[ADM06] Gabriel Acosta, Ricardo G. Durán, and María A. Muschietti. Solutions of the divergence operator on John domains. *Adv. Math.*, 206(2):373–401, 2006.

[ASV88] Douglas N. Arnold, L. Ridgway Scott, and Michael Vogelius. Regular inversion of the divergence operator with Dirichlet boundary conditions on a polygon. *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4)*, 15(2):169–192 (1989), 1988.

[AW02] Douglas N. Arnold and Ragnar Winther. Mixed finite elements for elasticity. *Numer. Math.*, 92(3):401–419, 2002.

[AW03] Douglas N. Arnold and Ragnar Winther. Nonconforming mixed elements for elasticity. *Math. Models Methods Appl. Sci.*, 13(3):295–307, 2003.

[BDD04] Peter Binev, Wolfgang Dahmen, and Ron DeVore. Adaptive finite element methods with convergence rates. *Numer. Math.*, 97(2):219–268, 2004.

[BDW99] Dietrich Braess, Wolfgang Dahmen, and Christian Wieners. A multigrid algorithm for the mortar finite element method. *SIAM J. Numer. Anal.*, 37(1):48–69, 1999.

[BF91] Franco Brezzi and Michel Fortin. *Mixed and hybrid finite element methods*, volume 15 of *Springer Series in Computational Mathematics*. Springer-Verlag, New York, 1991.

[BOP80] I. Babuška, J. Osborn, and J. Pitkäranta. Analysis of mixed methods using mesh dependent norms. *Math. Comp.*, 35(152):1039–1062, 1980.

[Bra07] Dietrich Braess. *Finite elements*. Cambridge University Press, Cambridge, third edition, 2007.

[Bre04] Susanne C. Brenner. Korn's inequalities for piecewise H^1 vector fields. *Math. Comp.*, 73(247):1067–1087, 2004.

[BS08] Susanne C. Brenner and L. Ridgway Scott. *The mathematical theory of finite element methods*, volume 15 of *Texts in Applied Mathematics*. Springer, New York, third edition, 2008.

[BV96] Dietrich Braess and Rüdiger Verfürth. A posteriori error estimators for the Raviart-Thomas element. *SIAM J. Numer. Anal.*, 33(6):2431–2444, 1996.

[CD98] Carsten Carstensen and Georg Dolzmann. A posteriori error estimates for mixed FEM in elasticity. *Numer. Math.*, 81(2):187–209, 1998.

[CEG11] C. Carstensen, M. Eigel, and J. Gedicke. Computational competition of symmetric mixed FEM in linear elasticity. *Comput. Methods Appl. Mech. Engrg.*, 200(41-44):2903–2915, 2011.

[CGP12] Carsten Carstensen, Joscha Gedicke, and Eun-Jae Park. Numerical experiments for the Arnold-Winther mixed finite elements for the Stokes problem. *SIAM J. Sci. Comput.*, 34(4):A2267–A2287, 2012.

[CGRT08] C. Carstensen, D. Günther, J. Reininghaus, and J. Thiele. The Arnold-Winther mixed FEM in linear elasticity. I. Implementation and numerical verification. *Comput. Methods Appl. Mech. Engrg.*, 197(33-40):3014–3023, 2008.

[CGS13] C. Carstensen, D. Gallistl, and M. Schedensack. Quasi-optimal adaptive pseudostress approximation of the Stokes Equations. *SIAM J. Numer. Anal.*, 51(3):1715–1734, 2013.

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

[CKP11] Carsten Carstensen, Dongho Kim, and Eun-Jae Park. A priori and a posteriori pseudostress-velocity mixed finite element error analysis for the Stokes problem. *SIAM J. Numer. Anal.*, 49(6):2501–2523, 2011.

[CPS12] C. Carstensen, D. Peterseim, and M. Schedensack. Comparison results of finite element methods for the Poisson model problem. *SIAM J. Numer. Anal.*, 50(6):2803–2823, 2012.

[CTVW10] Z. Cai, C. Tong, P. S. Vassilevski, and C. Wang. Mixed finite element methods for incompressible flow: stationary Stokes equations. *Numer. Methods Part. Diff. Eqs.*, 26(4):957 – 978, 2010.

[CW07] Zhiqiang Cai and Yanqiu Wang. A multigrid method for the pseudostress formulation of Stokes problems. *SIAM J. Sci. Comput.*, 29(5):2078–2095 (electronic), 2007.

[CWZ10] Zhiqiang Cai, Chunbo Wang, and Shun Zhang. Mixed finite element methods for incompressible flow: stationary Navier-Stokes equations. *SIAM J. Numer. Anal.*, 48(1):79–94, 2010.

[DE12] D. A. Di Pietro and A. Ern. *Mathematical aspects of discontinuous Galerkin methods*, volume 69 of *Mathématiques & Applications (Berlin) [Mathematics & Applications]*. Springer, Heidelberg, 2012.

[Gud10] Thirupathi Gudi. A new error analysis for discontinuous finite element methods for linear elliptic problems. *Math. Comp.*, 79(272):2169–2189, 2010.

[HX12] JianGuo Huang and YiFeng Xu. Convergence and complexity of arbitrary order adaptive mixed element methods for the Poisson equation. *Sci. China Math.*, 55(5):1083–1098, 2012.

[LS06] Carlo Lovadina and Rolf Stenberg. Energy norm a posteriori error estimates for mixed finite element methods. *Math. Comp.*, 75(256):1659–1674, 2006.

[Ver96] Rüdiger Verfürth. *A review of a posteriori estimation and adaptive mesh-refinement Techniques.* Advances in Numerical Mathematics. Wiley-Teubner, 1996.