Carsten Carstensen*, Benedikt Gräßle, and Neela Nataraj

Unifying a posteriori error analysis of five piecewise quadratic discretisations for the biharmonic equation

https://doi.org/10.1515/jnma-2022-0092
Received September 24, 2022; revised January 13, 2023; accepted January 23, 2023

Abstract: An abstract property (H) is the key to a complete a priori error analysis in the (discrete) energy norm for several nonstandard finite element methods in the recent work [Lowest-order equivalent nonstandard finite element methods for biharmonic plates, Carstensen and Nataraj, M2AN, 2022]. This paper investigates the impact of (H) to the a posteriori error analysis and establishes known and novel explicit residual-based a posteriori error estimates. The abstract framework applies to Morley, two versions of discontinuous Galerkin, C^0 interior penalty, as well as weakly over-penalized symmetric interior penalty schemes for the biharmonic equation with a general source term in $H^{-2}(\Omega)$.

Keywords: a posteriori, residual-based, biharmonic problem, smoother, best-approximation, companion operator, C^0 interior penalty, discontinuous Galerkin, WOPSIP, Morley

Classification: 65N30, 65N12, 65N50

1 Introduction

The concept of a quasi-optimal smoother and the key assumption (H) from [24] allow for an abstract a posteriori error analysis for five lowest-order schemes for the biharmonic problem. This paper unifies and completes [2, 4, 5, 34, 36, 40] and provides novel reliable and efficient a posteriori error estimators for a right-hand side $F \in H^{-2}(\Omega)$.

1.1 Overview

The traditional view on a posteriori error control is that the well-posedness of the linear problem on the continuous level directly leads from the error to residuals and their dual norms. In the simplest setting of a Hilbert space (V, a) with induced norm $\|\!|\cdot\|\!| := a(\cdot,\cdot)^{1/2}$, the weak solution $u \in V$ is the Riesz representation of a given source $F \in V^*$: $u \in V$ solves

$$a(u, v) = F(v) \qquad \forall\, v \in V. \tag{1.1}$$

Given any conforming companion $J_h u_h \in V$ to some discrete approximation $u_h \in V_h$, where $V_h \not\subseteq V$ is typically not a subset of V and $J_h u_h \in V$ is a postprocessing of u_h, the norm of the error $e := u - J_h u_h \in V$ is the norm of the residual $F - a(J_h u_h, \cdot) \in V^*$: The Riesz isometry between the residual and its Riesz representation $e \in V$ reads

$$\|\!|e\|\!| = \|\!|F - a(J_h u_h, \cdot)\|\!|_* := \sup_{v \in V\setminus\{0\}} \frac{|F(v) - a(J_h u_h, v)|}{\|\!|v\|\!|}. \tag{1.2}$$

***Corresponding author: Carsten Carstensen,** Department of Mathematics, Humboldt-Universität zu Berlin, 10099 Berlin, Germany. Distinguished Visiting Professor, Department of Mathematics, Indian Institute of Technology Bombay, Powai, Mumbai-400076, India. Email: cc@math.hu-berlin.de
Benedikt Gräßle, Department of Mathematics, Humboldt-Universität zu Berlin, 10099 Berlin, Germany.
Neela Nataraj, Department of Mathematics, Indian Institute of Technology Bombay, Powai, Mumbai 400076, India.

Tab. 1: Five discretizations.

	Morley	C^0IP	dG I	dG II	WOPSIP	Reference
V_h	$\mathrm{M}(\mathcal{T})$	$S_0^2(\mathcal{T})$	$P_2(\mathcal{T})$	$P_2(\mathcal{T})$	$P_2(\mathcal{T})$	(4.1)
$I = I_h I_\mathrm{M} : V \to V_h$	id	$I_\mathrm{C} I_\mathrm{M}$	I_M	I_M	I_M	I_M in Def. 4.1, I_C in (4.6)
$J_h = J I_\mathrm{M} : V_h \to V$	J	$J I_\mathrm{M}$	$J I_\mathrm{M}$	$J I_\mathrm{M}$	$J I_\mathrm{M}$	J in Lemma 4.1

The a posteriori error control is left with the task of deriving computable upper and lower bounds of the dual norm $|\!|\!|F - a(J_h u_h, \cdot)|\!|\!|_*$. The known data are $F \in V^*$ and $J_h u_h \in V$ and the techniques to derive bounds are very different from those of an a priori error analysis.

The paradigm change in this paper employs a recent tool (H) (stated in Section 2.2 below) from the *a priori error analysis* [24] to arrive at an *a posteriori error bound*

$$|\!|\!|e|\!|\!|^2 \leqslant C\big(\|u_h - J_h u_h\|_h^2 + \mathrm{Res}((1 - J_h I)e) + \text{data approximation error}\big) \tag{1.3}$$

with some operator $I \in L(V; V_h)$ and a norm $\|\cdot\|_h$ on $V + V_h$. The main advantage of the master estimate (1.3) over the error identity (1.2) is the known structure $(1 - J_h I)e \in V$ of the test function. The a posteriori error analysis based on (1.3) then only requires to study the properties of the operators $(1 - J_h I) \in L(V; V)$. This allows explicit estimates of the error term $\mathrm{Res}((1 - J_h I)e)$ with universal arguments for generic $u_h \in V_h$ and, most importantly, independent of the discrete system that defines $u_h \in V_h$.

The application to the biharmonic equation (1.1) provides novel simultaneous insight in the residuals and estimators for the piecewise quadratic discrete solution $u_h \in P_2(\mathcal{T})$ to the Morley, two variants of discontinuous Galerkin (dG), the C^0 interior penalty (C^0IP), and the weakly over-penalized symmetric interior penalty (WOPSIP) method. Table 1 below displays the discrete spaces V_h and operators I, J_h introduced in Section 4. The multiplicative constant C in (1.3) exclusively depends on the shape regularity of the underlying triangulation.

The discussion includes the standard and modified schemes that come with and without a smoother J_h on the right-hand side. This paper completes the a posteriori error analysis for these lowest-order discretisations and provides novel reliable and efficient a posteriori error estimators for a rather general class of general sources $F \in V^*$.

1.2 Outline

Section 2 introduces the abstract discretisation scheme with the key assumption (H) for the a priori analysis in [24]. Section 3 discusses a known abstract error identity and its application in the a posteriori error analysis. This is followed by the concept of a quasi-optimal smoother and the a priori key property (H) that lead to an explicit a posteriori error bound with a particular structure of the test function as in (1.3). Section 4 provides examples for the abstract setting in terms of five lowest-order schemes for the biharmonic equation. Section 5 establishes explicit estimates for the error contributions of the a posteriori error bound from Section 3. Section 6 presents a unified a posteriori error control for five lowest-order schemes for the biharmonic equation in a simplified setting with a right-hand side $F \in L^2(\Omega)$ and recovers [2, 4, 5, 34, 36, 40]. The restriction to sources in L^2 underlines the state of the art before this paper and thereby highlights the new paradigm through comparison with known results. The emphasis in Section 7 is on a class of general sources $F \in V^*$ with a novel a posteriori error estimator of the residual that is reliable and efficient up to data-oscillations. Appendix A shades a different light on the discussion in Section 7 and provides lower and upper bounds for the dual norm of functionals $F \in V^*$.

The presentation is laid out in two dimensions with shape-regular triangulations into triangles and second-order discretizations for simplicity; but the arguments apply to 3D as well, cf. [26] for a companion operator J_h in 3D. The abstract results of this paper will be applied to an a posteriori error analysis of semilinear problems [18], where a linearisation enforces (piecewise polynomial) $F \in H^{-2}(\Omega) \setminus L^2(\Omega)$ in future research.

1.3 General notation

Standard notation on Lebesgue and Sobolev spaces, their norms, and L^2 scalar products applies throughout the paper such as the abbreviation $\|\cdot\|$ for $\|\cdot\|_{L^2(\Omega)}$. Recall that the energy norm $|||\cdot||| := \|D^2\cdot\|$ is a norm on $H^2_0(\Omega)$. Throughout this paper, $\mathcal{T}$ denotes a shape-regular triangulation of a polygonal and bounded (possibly multiply-connected) Lipschitz domain $\Omega \subset \mathbb{R}^2$ into triangles. Let $\mathcal{V}(\Omega)$ and $\mathcal{E}(\Omega)$ denote the set of interior vertices and edges in the triangulation $\mathcal{T}$ and let $\mathcal{V}(\partial\Omega)$ and $\mathcal{E}(\partial\Omega)$ denote the boundary vertices and edges. The gradient and Hessian operators $\nabla_{\mathrm{pw}} := D_{\mathrm{pw}}$ and D^2_{pw} act piecewise on the space $H^m(\mathcal{T}) := \prod_{T\in\mathcal{T}} H^m(T)$ of piecewise Sobolev functions for $m = 1, 2$ with the abbreviation $H^m(K) := H^m(\operatorname{int} K)$ for a triangle or edge $K \in \mathcal{T}\cup\mathcal{E}$ with relative interior $\operatorname{int}(K)$. The space $P_k(K)$ of polynomials of total degree at most $k \in \mathbb{N}_0$ on $K \in \mathcal{T}\cup\mathcal{E}$ with diameter h_K defines the space of piecewise polynomials

$$P_k(\mathcal{T}) := \{p \in L^\infty(\Omega) : p|_T \in P_k(T)\ \forall T \in \mathcal{T}\}.$$

The mesh-size $h_\mathcal{T} \in P_0(\mathcal{T})$ is the piecewise constant function with $h_\mathcal{T}|_T \equiv h_T := \operatorname{diam}(T)$ for all $T \in \mathcal{T}$. Throughout this paper, let $H^k(\Omega; X)$, $H^k(\mathcal{T}; X)$, resp. $P_k(\mathcal{T}; X)$ denote the space of (piecewise) Sobolev functions resp. polynomials with values in $X = \mathbb{R}^2, \mathbb{R}^{2\times2}, \mathbb{S}$ for $k \in \mathbb{N}_0$; $\mathbb{S} \subset \mathbb{R}^{2\times2}$ is the set of symmetric 2×2 matrices. The spaces $H^{-k}(\Omega) := (H^k_0(\Omega))^*$ are the dual spaces of $H^k_0(\Omega)$ for $k \in \mathbb{N}$. Given any function $v \in L^2(E)$ on an edge $E \in \mathcal{E}$, define the integral mean $⨍_E v\,dx := h_E^{-1}\int_E v\,dx$. The notation $A \lesssim B$ abbreviates $A \leqslant CB$ for some positive generic constant C, which exclusively depends on the shape-regularity of the underlying triangulation $\mathcal{T}$; $A \approx B$ abbreviates $A \lesssim B \lesssim A$.

2 Unified a priori error analysis

Nonstandard schemes compute discontinuous approximations in general and require a smoother to map the discrete functions into the continuous space V.

2.1 Discretisation

Given the Hilbert space (V, a) from the continuous problem (1.1), consider some bigger Hilbert space $(\widehat{V}, \widehat{a})$ that contains $V \subset \widehat{V}$ as well as the discrete spaces $V_h, V_{\mathrm{nc}} \subset \widehat{V}$. Let $\widehat{a} := a_{\mathrm{pw}} + j_h$ be the sum of the semi-scalar products $a_{\mathrm{pw}}, j_h : \widehat{V}\times\widehat{V} \to \mathbb{R}$ where, a_{pw} extends $a = a_{\mathrm{pw}}|_{V\times V}$ and is a scalar product with induced norm $|||\cdot|||_{\mathrm{pw}} := a_{\mathrm{pw}}(\cdot,\cdot)^{1/2}$ in $V + V_{\mathrm{nc}}$. The semi-scalar product $j_h : \widehat{V}\times\widehat{V} \to \mathbb{R}$ represents jumps that vanish in $V + V_{\mathrm{nc}}$, i.e., $j_h(v,\cdot) = 0$ for any $v \in V + V_{\mathrm{nc}}$. The induced norm on $\widehat{V}$ reads

$$\|\cdot\|_h := \left(|||\cdot|||^2_{\mathrm{pw}} + j_h(\cdot,\cdot)\right)^{1/2} \quad \text{and satisfies} \quad |||\cdot|||_{\mathrm{pw}} = \|\cdot\|_h \text{ in } V + V_{\mathrm{nc}}. \tag{2.1}$$

The discretisation consists of a finite-dimensional trial and test space V_h with respect to a shape-regular triangulation $\mathcal{T}$ of Ω and the (possibly unsymmetric) bilinear form

$$a_h : (V + V_h + V_{\mathrm{nc}}) \times (V + V_h + V_{\mathrm{nc}}) \to \mathbb{R}.$$

We assume that a_h is V_h-elliptic and bounded on V_h with respect to $\|\cdot\|_h$ in the sense that some universal constants $0 < \alpha \leqslant M < \infty$ satisfy, for all $v_h, w_h \in V_h$, that

$$\alpha\|v_h\|^2_h \leqslant a_h(v_h, v_h), \qquad a_h(v_h, w_h) \leqslant M\|v_h\|_h\|w_h\|_h. \tag{2.2}$$

Since $V_h \subsetneq V$ is *not* a subset of V, the evaluation $F(v_h)$ at $v_h \in V_h$ is *not* well-defined for general $F \in V^*$. Therefore many of the earlier contributions, in particular to the a posteriori error control, merely consider $F \in L^2(\Omega)$ whenever $\widehat{V} \subset L^2(\Omega)$. The series of papers [42–44] advertise a smoother $Q \in L(V_h; V)$ to evaluate the modified source $F(Qv_h)$ on the discrete level. This paper complements those contributions on the a priori error analysis by reliable and efficient a posteriori error estimates. This is itself highly relevant in scientific computing and a first step towards adaptive mesh-refining.

To be more general, this paper considers a rather general class of sources that allow an extension $\widehat{F} \in \widehat{V}^*$ of $F = \widehat{F}|_V$. The Lax–Milgram lemma ensures the existence of a unique discrete solution $u_h \in V_h$ to

$$a_h(u_h, v_h) = \widehat{F}(Qv_h) \qquad \forall v_h \in V_h \tag{2.3}$$

for the two cases $Q = \mathrm{id}$ (no smoother, but depending on $\widehat{F}$) and $Q = J_h$ for a smoother $J_h \in L(V_h; V)$. The history of J_h is related to averaging techniques and dates back to the analysis of the Crouzeix–Raviart method [16, 25, 44] for the reliable error control [22]. An earlier motivation was the construction of intergrid transfer operators in the convergence analysis of multigrid methods for nonconforming schemes [9].

The first results will be derived for $\widehat{F} \equiv f \in L^2(\Omega)$ to recover known results in a unified framework, while Section 7 specifies a large class of extended sources $\widehat{F}$ and provides novel a posteriori error estimates with and without smoother.

2.2 Quasi-best approximation

The abstract framework from [24] provides a tool for the a priori analysis therein.

Definition 2.1 (quasi-optimal smoother). An operator $J_h \in L(V_h; V)$ is called a quasi-optimal smoother if there exists a constant $C_\mathrm{J} \geqslant 0$ such that

$$\|v_h - J_h v_h\|_h \leqslant C_\mathrm{J} \min_{v \in V} \|v - v_h\|_h \qquad \forall v_h \in V_h. \tag{2.4}$$

All the examples in [42–44] discuss $J_h \in L(V_h; V)$ with $J_h = \mathrm{id}$ in $V_h \cap V$. The framework in [24] introduces a smoother that satisfies (2.4) and is quasi-optimal with a constant $C_\mathrm{J} \approx 1$. The interpretation is that $J_h v_h \in V$ is a good approximation of $v_h \in V_h$ and provides a bridge between the discrete objects in V_h and V.

The key assumption (H) connects the bilinear forms a from (1.1) and a_h from (2.3) and requires the existence of $\Lambda_\mathrm{H} \geqslant 0$ with

$$a_h(w_h, v_h) - a(J_h w_h, J_h v_h) \leqslant \Lambda_\mathrm{H} \|w_h - J_h w_h\|_h \|v_h\|_h \qquad \forall w_h, v_h \in V_h. \tag{H}$$

This assumption leads to quasi-optimality of u_h in the discrete norm $\|\cdot\|_h$ and holds for a class of problems including the examples in [24] except WOPSIP. A key step is therefore the design of a quasi-optimal smoother, e.g., $J_h = J \circ I_\mathrm{nc}$ with the conforming companion J and a generalised interpolation operator I_nc.

Theorem 2.1 (quasi-best approximation). *Given an operator $J_h \in L(V_h; V)$ with* (2.4) *and* (H)*, there exists a constant $C_\mathrm{qo} > 0$ (that exclusively depends on α, M, C_J, Λ_H, and $\|J_h\|$) such that the exact solution $u \in V$ to* (1.1) *and the discrete solution u_h to* (2.3) *satisfy*

$$\|u - u_h\|_h \leqslant C_\mathrm{qo} \min_{v_h \in V_h} \|u - v_h\|_h. \tag{QO}$$

Proof. This is proven in [24, Thm. 2.3] inspired by the seminal work [42]. □

A stronger version $(\widehat{\mathrm{H}})$ of (H) in [24, Sect. 6] even leads to a priori error bounds in weaker (piecewise) Sobolev norms.

2.3 Transfer operators

The error analysis requires transfer operators with certain approximation properties between the three subspaces V, V_h, V_nc of $\widehat{V}$. Throughout this paper, assume there are three linear operators $I_h \in L(V_\mathrm{nc}; V_h), I_\mathrm{nc} \in$

I_{nc}	$\in L(V + V_{\mathrm{nc}} + V_h; V_{\mathrm{nc}})$
I_h	$\in L(V_{\mathrm{nc}}; V_h)$
J	$\in L(V_{\mathrm{nc}}; V)$
J_h	$\in L(V + V_{\mathrm{nc}} + V_h; V)$

Tab. 2: Operators.

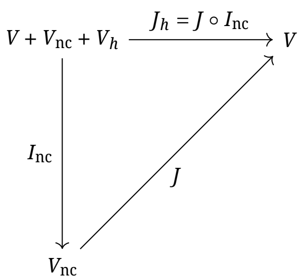

Fig. 1: Definition of J_h.

$L(V + V_h + V_{\mathrm{nc}}; V_{\mathrm{nc}})$, and the conforming companion operator $J \in L(V_{\mathrm{nc}}; V)$ and constants $\Lambda_h, \Lambda_{\mathrm{nc}}, \Lambda_{\mathrm{J}} \geqslant 0$ such that

$$\|v_{\mathrm{nc}} - I_h v_{\mathrm{nc}}\|_h \leqslant \Lambda_h \min_{v \in V} |||v - v_{\mathrm{nc}}|||_{\mathrm{pw}} \qquad \forall\, v_{\mathrm{nc}} \in V_{\mathrm{nc}} \tag{2.5}$$

$$\|v_h - I_{\mathrm{nc}} v_h\|_h \leqslant \Lambda_{\mathrm{nc}} \min_{v \in V} \|v - v_h\|_h \qquad \forall\, v_h \in V_h \tag{2.6}$$

$$|||v_{\mathrm{nc}} - J v_{\mathrm{nc}}|||_{\mathrm{pw}} \leqslant \Lambda_{\mathrm{J}} \min_{v \in V} |||v - v_{\mathrm{nc}}|||_{\mathrm{pw}} \qquad \forall\, v_{\mathrm{nc}} \in V_{\mathrm{nc}}. \tag{2.7}$$

Two immediate consequences on the abstract level at hand shall be utilized below.

Lemma 2.1 (intermediate bound). *Given any* $v \in V$ *and* $v_{\mathrm{nc}} \in V_{\mathrm{nc}}$, (2.5)–(2.7) *imply*

$$|||v - J I_{\mathrm{nc}} I_h v_{\mathrm{nc}}||| \leqslant (1 + \Lambda_{\mathrm{J}})(1 + \Lambda_{\mathrm{nc}})(1 + \Lambda_h)|||v - v_{\mathrm{nc}}|||_{\mathrm{pw}}. \tag{2.8}$$

Proof. Let $w_{\mathrm{nc}} := I_{\mathrm{nc}} I_h v_{\mathrm{nc}}$ and $w_h := I_h v_{\mathrm{nc}}$. The triangle inequality and (2.7) show

$$|||v - J w_{\mathrm{nc}}||| \leqslant |||v - w_{\mathrm{nc}}|||_{\mathrm{pw}} + |||(1 - J) w_{\mathrm{nc}}|||_{\mathrm{pw}} \leqslant (1 + \Lambda_{\mathrm{J}})|||v - w_{\mathrm{nc}}|||_{\mathrm{pw}}.$$

Note $|||v - w_{\mathrm{nc}}|||_{\mathrm{pw}} = \|v - w_{\mathrm{nc}}\|_h$ from (2.1). The triangle inequality and (2.5)–(2.6) show

$$\|v - w_{\mathrm{nc}}\|_h \leqslant \|v - w_h\|_h + \|(1 - I_{\mathrm{nc}}) w_h\|_h \leqslant (1 + \Lambda_{\mathrm{nc}})\|v - w_h\|_h$$
$$\|v - w_h\|_h \leqslant \|v - v_{\mathrm{nc}}\|_h + \|(1 - I_h) v_{\mathrm{nc}}\|_h \leqslant (1 + \Lambda_h)|||v - v_{\mathrm{nc}}|||_{\mathrm{pw}}.$$

The combination of those estimates establishes (2.8). □

The above transfer operators (see Fig. 1 and Table 2) lead to a quasi-optimal smoother $J_h := J \circ I_{\mathrm{nc}} \in L(V_h; V)$. Although J_h maps $V + V_{\mathrm{nc}} + V_h \to V$, its restriction to V_h plays a central role in the sequel.

Lemma 2.2 (quasi-optimal smoother). *Given any* $v_h \in V_h$, *and* $J_h := J \circ I_{\mathrm{nc}} \in L(V_h; V)$, (2.6)–(2.7) *show* (2.4) *with* $C_{\mathrm{J}} := \Lambda_{\mathrm{nc}} + \Lambda_{\mathrm{J}} + \Lambda_{\mathrm{J}}\Lambda_{\mathrm{nc}}$.

Proof. A triangle inequality with $v_{\mathrm{nc}} := I_{\mathrm{nc}} v_h$, and (2.7) verify

$$\|v_h - J v_{\mathrm{nc}}\|_h \leqslant \|v_h - v_{\mathrm{nc}}\|_h + \Lambda_{\mathrm{J}}\left(\|v - v_h\|_h + \|v_h - v_{\mathrm{nc}}\|_h\right)$$

for an arbitrary $v \in V$. This and (2.6) conclude the proof. □

Lemma 2.2 shows that J_h is a quasi-optimal smoother with the following property.

Theorem 2.2 (quasi-best approximation [24]). *Let* $u \in V$ *resp.* $u_h \in V_h$ *solve* (1.1) *resp.* (2.3). *Suppose* (H), (2.1)–(2.2), *and* (2.5)–(2.7). *Then*

$$|||u - J_h u_h||| + \|u - u_h\|_h \lesssim \min_{v_{\mathrm{nc}} \in V_{\mathrm{nc}}} |||u - v_{\mathrm{nc}}|||_{\mathrm{pw}}.$$

Proof. Lemma 2.2 and Theorem 2.1 verify (QO) for J_h. A triangle inequality, (2.5), and $|||\cdot|||_{\mathrm{pw}} = \|\cdot\|_h$ in $V + V_{\mathrm{nc}}$ verify

$$\|u - u_h\|_h \leqslant C_{\mathrm{qo}}\|u - I_h v_{\mathrm{nc}}\|_h \leqslant C_{\mathrm{qo}}(1 + \Lambda_h)|||u - v_{\mathrm{nc}}|||_{\mathrm{pw}}$$

for arbitrary $v_{\mathrm{nc}} \in V_{\mathrm{nc}}$. The proof of Lemma 2.1 shows $|||v - J_h w_h||| \leqslant (1+\Lambda_{\mathrm{J}})(1+\Lambda_{\mathrm{nc}})\|v - w_h\|_h$ for an arbitrary $v \in V$, $w_h \in V_h$. The combination with the previously displayed inequality concludes the proof. □

3 Abstract a posteriori error analysis

The abstract error identity in Subsection 3.1 reveals that $|||\mathrm{Res}|||_*$ is a contribution to the error. Subsection 3.2 revisits the Crouzeix–Raviart and Morley FEM and recalls known bounds thereof. Subsection 3.3 explains a paradigm shift towards a universal error analysis that is explicit in the structure of the test function through a quasi-optimal smoother and the property (H).

3.1 Abstract error identity for $F \in V^*$

Given the exact solution $u \in V$ to (1.1) and the discrete solution $u_h \in V_h$ to (2.3), the natural error $u - u_h \in V + V_h \subset \widehat{V}$ can be measured in the norm $\|\cdot\|_h$ from Subsection 2.1. This allows a well-known split with the residual $\mathrm{Res} := F - a_{\mathrm{pw}}(u_h, \cdot) \in V^*$ [19].

Theorem 3.1 (error identity). *The exact solution $u \in V$ to* (1.1) *and the discrete solution $u_h \in V_h$ to* (2.3) *satisfy*

$$\|u - u_h\|_h^2 = |||\mathrm{Res}|||_*^2 + \min_{v \in V}\|v - u_h\|_h^2. \tag{3.1}$$

Proof. Let $w \in V$ be the Riesz representation of the linear and bounded functional $a_{\mathrm{pw}}(u_h, \cdot) \in V^*$ in the Hilbert space (V, a_{pw}), so that $a_{\mathrm{pw}}(u_h - w, \cdot) = 0$ in V. This orthogonality shows that $w \in V$ is the best-approximation of $u_h \in V_h \subset \widehat{V}$ in the complete subspace V, i.e.,

$$\delta := |||w - u_h|||_{\mathrm{pw}} = \min_{v \in V}|||v - u_h|||_{\mathrm{pw}} \tag{3.2}$$

and allows for the Pythagoras identity

$$|||u - u_h|||_{\mathrm{pw}}^2 = |||u - w|||^2 + |||w - u_h|||_{\mathrm{pw}}^2. \tag{3.3}$$

The orthogonality also shows, for all $v \in V$, that

$$a(u - w, v) = a(u, v) - a_{\mathrm{pw}}(u_h, v) = \mathrm{Res}(v)$$

with $a(u, \cdot) = F$ in V in the last step. In other words, $u - w$ is the Riesz representation of $\mathrm{Res} \in V^*$ in the Hilbert space (V, a) and the Riesz isomorphism reveals

$$|||u - w||| = |||\mathrm{Res}|||_* := \sup_{v \in V\setminus\{0\}} \frac{\mathrm{Res}(v)}{|||v|||}. \tag{3.4}$$

The summary of (3.2)–(3.4) reads $|||u - u_h|||_{\mathrm{pw}}^2 = |||\mathrm{Res}|||_*^2 + \delta^2$. Since $j_h(\cdot, v) = j_h(v, \cdot) = 0$, the proof concludes with $\|v - u_h\|_h^2 = |||v - u_h|||_{\mathrm{pw}}^2 + j_h(u_h, u_h)$ for any $v \in V$. □

Remark 3.1 (explicit a posteriori bounds). The proof of Theorem 3.1 is nothing but a Pythagoras identity and serves as an idealisation: While $j_h(u_h, u_h)$ comes for free, the computation of $|||\mathrm{Res}|||_*$ or of $\delta = \min_{v \in V}\|v - u_h\|_h^2 - j_h(u_h, u_h)$ is far too costly. Instead, the error identity rather serves as a guide to design individual upper bounds of δ and $|||\mathrm{Res}|||_*$. The a priori error analysis of Section 2.1 provides a quasi-optimal smoother $J_h \in L(V_h; V)$. Then (2.4) shows

$$\min_{v \in V}\|v - u_h\|_h \leqslant \|u_h - J_h u_h\|_h \leqslant C_{\mathrm{J}} \min_{v \in V}\|v - u_h\|_h. \tag{3.5}$$

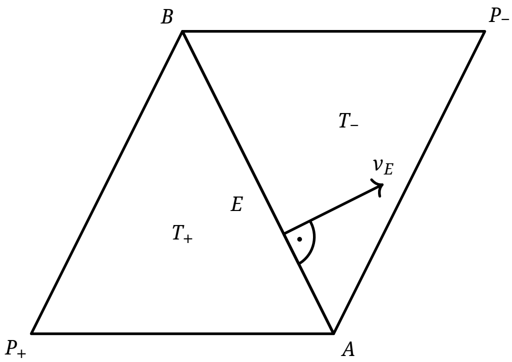

Fig. 2: The interior edge patch $\omega(E)$ and normal $\nu_E = \pm\nu_{T_\pm}$ of $E = \partial T_+ \cap \partial T_-$.

In the language of a posteriori error control, (3.5) asserts the reliability and efficiency of the a posteriori estimator $\|u_h - J_h u_h\|_h$ of the error $\min_{v\in V} \|v - u_h\|_h$. This ends the discussion of $\|u_h - J_h u_h\|_h$ and motivates the focus on bounds of Res below.

In order to understand the difference between the classical and the current treatment, the two simplest nonconforming schemes will be discussed in the subsequent subsection.

3.2 Crouzeix–Raviart and Morley FEM

This subsection motivates the abstract a posteriori error analysis by a recollection [10, 11, 14, 19, 20, 29, 41] for $m = 1$ and [2, 13, 35, 36] for $m = 2$ of the simplest nonconforming schemes for the m-harmonic equation $(-\Delta)^m u = f$ for $m = 1, 2$ with right-hand function $f \in L^2(\Omega)$. The weak solution seeks $u \in V := H^m_0(\Omega) \subset \widehat{V} := H^m(\mathcal{T})$ to

$$a(u, v) = (f, v)_{L^2(\Omega)} \qquad \forall\, v \in V \tag{3.6}$$

with the energy scalar product $a := a_{\mathrm{pw}}|_{V\times V}$ and $a_{\mathrm{pw}}(\cdot,\cdot) := (D^m\cdot, D^m\cdot)_{L^2(\Omega)}$ in $\widehat{V}$.

3.2.1 Crouzeix–Raviart FEM

Let $u \in V := H^1_0(\Omega)$ be the weak solution to the Poisson model problem, i.e., u solves (3.6) for $m = 1$. The Crouzeix–Raviart finite element space requires the definition of jumps across an edge $E \in \mathcal{E}$ in the triangulation $\mathcal{T}$. Let ν_T be the unit outer normal of $T \in \mathcal{T}$ and fix the orientation of the unit normal ν_E on every edge $E \in \mathcal{E}$ with midpoint $\mathrm{mid}(T)$. Every interior edge $E = \partial T_+ \cap \partial T_- \in \mathcal{E}(\Omega)$ has exactly two neighbouring triangles $T_+, T_- \in \mathcal{T}$ as in Fig. 2, labelled such that $\nu_E = \pm\nu_{T_\pm}|_E$, and the jump of a piecewise Sobolev function $v \in H^1(\mathcal{T})$ across E reads $[v]_E := v|_{T_+} - v|_{T_-} \in H^1(E)$. On a boundary edge $E \in \mathcal{E}(\partial\Omega)$, the jump $[v]_E := v$ is the unique trace of the function $v \in H^1(\mathcal{T})$. Define the space

$$\mathrm{CR}^1_0(\mathcal{T}) := \big\{p \in P_1(\mathcal{T}) \,\big|\, [p]_E(\mathrm{mid}\,E) = 0 \text{ vanishes for every edge } E \in \mathcal{E}\big\}$$

of piecewise affine polynomials over a given shape-regular triangulation $\mathcal{T}$ with continuity at the midpoints of the edges. This space $V_{\mathrm{nc}} := \mathrm{CR}^1_0(\mathcal{T})$ comes with the natural interpolation operator $I_{\mathrm{CR}} : V + V_{\mathrm{nc}} \to V_{\mathrm{nc}}$ that maps $v \in V + V_{\mathrm{nc}}$ to the unique function $I_{\mathrm{CR}} v \in V_{\mathrm{nc}}$ with $\int_E (v - I_{\mathrm{CR}} v)\,\mathrm{d}s = 0$ for every edge $E \in \mathcal{E}$. The classical formulation of the lowest-order nonconforming Crouzeix–Raviart FEM approximates the weak solution $u \in H^1_0(\Omega)$ of (3.6) with the discrete solution $u_{\mathrm{CR}} \in \mathrm{CR}^1_0(\mathcal{T}) \equiv V_{\mathrm{nc}}$ to

$$a_{\mathrm{pw}}(u_{\mathrm{CR}}, v_{\mathrm{CR}}) \equiv \int_\Omega \nabla_{\mathrm{pw}} u_{\mathrm{CR}} \cdot \nabla_{\mathrm{pw}} v_{\mathrm{CR}}\,\mathrm{d}x = (f, v_{\mathrm{CR}})_{L^2(\Omega)} \qquad \forall\, v_{\mathrm{CR}} \in \mathrm{CR}^1_0(\mathcal{T}). \tag{3.7}$$

This is exactly (2.3) for the natural choice $\widehat{F} := F \equiv f \in L^2(\Omega)$ and without smoother $Q :=$ id. The semi-scalar product a_{pw} induces the piecewise energy norm $|||\cdot|||_{\mathrm{pw}} := \|\nabla_{\mathrm{pw}} \cdot\|$ in $V + V_{\mathrm{nc}} \equiv H^1_0(\Omega) + \mathrm{CR}^1_0(\mathcal{T})$ [23]. In this particular example, the residual from Section 3.1 reads

$$\mathrm{Res} := (f, \cdot)_{L^2(\Omega)} - a_{\mathrm{pw}}(u_{\mathrm{CR}}, \cdot) \in V^*.$$

3.2.2 Classical residual-based explicit error estimator

This approach follows [19] and is closely related to the analysis of conforming schemes. Let $I_{\mathrm{C}} : H^1_0(\Omega)+\mathrm{CR}^1_0(\mathcal{T}) \to S^1_0(\mathcal{T})$ denote a quasi-interpolation operator onto the continuous piecewise affine polynomials $S^1_0(\mathcal{T}) := P_1(\mathcal{T}) \cap H^1_0(\Omega)$ with homogeneous boundary conditions. Since (3.7) holds, the definition of the residual shows $\mathrm{Res}(w_{\mathrm{C}}) = 0$ for any $w_{\mathrm{C}} \in S^1_0(\mathcal{T}) \subset \mathrm{CR}^1_0(\mathcal{T})$, i.e., $S^1_0(\mathcal{T}) \subset \ker \mathrm{Res}$ lies in the kernel of $\mathrm{Res} \in V^*$, and an integration by parts with the test function $w := v - I_{\mathrm{C}} v$ shows, for $f \in L^2(\Omega)$, that

$$\begin{aligned}\mathrm{Res}(v) = \mathrm{Res}(w) &= (f, w)_{L^2(\Omega)} - \sum_{T\in\mathcal{T}} \sum_{E\in\mathcal{E}(T)} \int_E \nabla_{\mathrm{pw}} u_{\mathrm{CR}} \cdot \nu_E w \,\mathrm{d}s\\ &= (f, w)_{L^2(\Omega)} - \sum_{E\in\mathcal{E}(\Omega)} \int_E [\nabla_{\mathrm{pw}} u_{\mathrm{CR}}]_E \cdot \nu_E w \,\mathrm{d}s.\end{aligned}$$

The last step is a careful resummation over the edges: Each interior edge $E \in \mathcal{E}(\Omega)$ has two contributions (from T_+ and T_- as in Fig. 2) with opposite signs from $\nu_{T_+} = -\nu_{T_-}$ on E. No contributions arise from the boundary edges $E \in \mathcal{E}(\partial\Omega)$ because of $w|_{\partial\Omega} = 0$. Cauchy inequalities show

$$\mathrm{Res}(v) \leqslant \|h_{\mathcal{T}} f\| \|h_{\mathcal{T}}^{-1} w\| + \sum_{E\in\mathcal{E}(\Omega)} h_E^{1/2} \|[\nabla_{\mathrm{pw}} u_{\mathrm{CR}}]_E \cdot \nu_E\|_{L^2(E)} h_E^{-1/2} \|w\|_{L^2(E)}. \tag{3.8}$$

The quasi-interpolation operator I_{C} from [28, 39] satisfies the stability estimates

$$h_T^{-1} \|v - I_{\mathrm{C}} v\|_{L^2(T)} \leqslant C_{\mathrm{apx}} \|\nabla v\|_{L^2(\omega(T))} \quad \text{in } T \in \mathcal{T}$$

with a constant $C_{\mathrm{apx}} > 0$ that exclusively depends on the shape regularity of $\mathcal{T}$. Here $\omega(T)$ denotes the layer-1 patch around $T \in \mathcal{T}$. The trace inequality [30, Eq. (12.17)]:

$$h_E^{-1/2} \|v\|_{L^2(E)} \leqslant C_{\mathrm{tr}} \left(h_T^{-1} \|v\|_{L^2(T(E))} + \|\nabla v\|_{L^2(T(E))} \right) \qquad \forall\, v \in V$$

bounds the norms on the edge $E \subset \partial T(E)$ by norms of some adjacent triangle $T(E) \in \mathcal{T}$ with a constant $C_{\mathrm{tr}} > 0$ that exclusively depends on the shape-regularity of $\mathcal{T}$. This and a final Cauchy inequality in ℓ^2 for the sum in (3.8) show

$$|||\mathrm{Res}|||_* := \sup_{v\in V\setminus\{0\}} \frac{\mathrm{Res}(v)}{|||v|||} \lesssim \|h_{\mathcal{T}} f\| + \sqrt{\sum_{E\in\mathcal{E}(\Omega)} h_E \|[\nabla u_{\mathrm{CR}}]_E \cdot \nu_E\|^2_{L^2(E)}}. \tag{3.9}$$

The jump term in (3.9) can be bounded by $\|h_{\mathcal{T}} f\|$ and the simpler

$$|||\mathrm{Res}|||_* \lesssim \|h_{\mathcal{T}} f\| \tag{3.10}$$

estimate without normal jumps is possible. For any interior edge $E \in \mathcal{E}(\Omega)$, the edge-patch $\omega(E) := \mathrm{int}(T_+ \cup T_-)$ depicted in Fig. 2 is the union of the two neighboring triangles $T_+, T_- \in \mathcal{T}$.

Lemma 3.1 (bound without jumps). *The normal jumps from* (3.9) *satisfy*

$$h_E^{1/2} \|[\nabla u_{\mathrm{CR}}]_E \cdot \nu_E\|_{L^2(E)} \lesssim \|h_{\mathcal{T}} f\|_{L^2(\omega(E))} \qquad \forall E \in \mathcal{E}(\Omega).$$

Proof. Recall the edge-oriented basis function $\psi_E \in CR^1_0(\mathcal{T})$ as the unique function in $\mathrm{CR}^1_0(\mathcal{T})$ with $\psi_E(\mathrm{mid}\ E) = 1$ and $\psi_E(\mathrm{mid}\ F) = 0$ for every other edge $F \in \mathcal{E}\setminus\{E\}$. Since $\psi_E \in \mathrm{CR}^1_0(\mathcal{T})$ is piecewise affine, its support $\overline{\omega(E)}$ is the edge-patch $\omega(E)$ with $\psi_E \equiv 1$ on $E = \partial T_+ \cap \partial T_-$. This, an integration by parts for the interior edge $E \in \mathcal{E}(\Omega)$, and (3.7) prove for $\beta := [\nabla u_{\mathrm{CR}}]_E \cdot \nu_E \in \mathbb{R}$,

$$\|[\nabla u_{\mathrm{CR}}]_E \cdot \nu_E\|^2_{L^2(E)} = \beta \int_E [\nabla u_{\mathrm{CR}}]_E \cdot \nu_E \, \psi_E \, \mathrm{d}s = \beta \int_{\omega(E)} \nabla u_{\mathrm{CR}} \cdot \nabla \psi_E \, \mathrm{d}x = \beta \, (f, \psi_E)_{L^2(\omega(E))}.$$

The midpoint quadrature rule shows $\|\psi_E\|^2_{L^2(T)} = |T|/3 \approx h_T^2 \approx h_E^2$ in 2D by shape-regularity. Since $\|\psi_E\|^2_{L^2(E)} = |E| = h_E$, the previous displayed identity, a Cauchy inequality, and the definition of β verify

$$\|[\nabla u_{\mathrm{CR}}]_E \cdot \nu_E\|^2_{L^2(E)} \lesssim \|h_{\mathcal{T}} f\|_{L^2(\omega(E))} h_E^{-1/2} \|[\nabla u_{\mathrm{CR}}]_E \cdot \nu_E\|_{L^2(E)}.$$

This concludes the proof of (3.10). □

3.2.3 Bound from Crouzeix–Raviart interpolation

The integration by parts formula on $T \in \mathcal{T}$ and the definition of the natural interpolation $I_{\mathrm{CR}} : V + V_{\mathrm{nc}} \to V_{\mathrm{nc}}$ reveal

$$\int_T \nabla(v - I_{\mathrm{CR}} v) \cdot \nabla p_1 \, \mathrm{d}x = \sum_{E \in \mathcal{E}(T)} [\nabla_{\mathrm{pw}} p_1]_E \cdot \nu_E \int_E (v - I_{\mathrm{CR}} v) \, \mathrm{d}s = 0$$

for any $v \in V + V_{\mathrm{nc}}$, $p_1 \in P_1(\mathcal{T})$ and $(v - I_{\mathrm{CR}} v) \perp P_1(\mathcal{T})$ is a_{pw}-orthogonal to $P_1(\mathcal{T}) \supset \mathrm{CR}_0^1(\mathcal{T})$. This, (3.7), and the interpolation error estimate $\|h_{\mathcal{T}}^{-1}(v - I_{\mathrm{CR}} v)\| \leqslant \varkappa_{\mathrm{CR}} |||v|||$ from [12, Sect. 4] with $\varkappa_{\mathrm{CR}} = (1/48 + j_{1,1}^2)^{1/2} \leqslant 0.2983$ for the first positive root $j_{1,1}$ of the Bessel function of the first kind shows

$$|||\mathrm{Res}|||_* := \sup_{v \in V \setminus \{0\}} \frac{\mathrm{Res}(v)}{|||v|||} = \sup_{v \in V \setminus \{0\}} \frac{(f, v - I_{\mathrm{CR}} v)_{L^2(\Omega)}}{|||v|||} \leqslant \varkappa_{\mathrm{CR}} \|h_{\mathcal{T}} f\|. \tag{3.11}$$

The difference to the bound in Lemma 3.1 is not only the explicit control in terms of the smaller constant $\varkappa_{\mathrm{CR}}$, but above all, that the methodology directly controls $|||\mathrm{Res}|||_*$ as in [13, p. 317] without jump terms. The latter also follows from (3.9) and Lemma 3.1.

The key observation is that this technique does *not* need any conforming subspace $S_0^1(\mathcal{T}) \subset \mathrm{CR}_0^1(\mathcal{T})$ and this is a relevant advance for the application to the Morley FEM.

3.2.4 Morley FEM

Let $u \in V := H_0^2(\Omega)$ be the weak solution to the biharmonic equation $\Delta^2 u = f \in L^2(\Omega)$, i.e., u solves (3.6) for $m = 2$. Define the normal jump $[\partial v / \partial \nu_E]_E := [\nabla v \cdot \nu_E]_E$ of a function $v \in H^2(\mathcal{T})$ along an edge $E \in \mathcal{E}$. The Morley function space

$$\mathrm{M}(\mathcal{T}) := \left\{ p \in P_2(\mathcal{T}) \;\middle|\; \begin{array}{l} p(z) \text{ is continuous at every } z \in \mathcal{V}(\Omega) \text{ and } p|_{\mathcal{V}(\partial\Omega)} = 0 \\ [\nabla_{\mathrm{pw}} p \cdot \nu_E]_E(\mathrm{mid}\, E) = 0 \text{ vanishes for every edge } E \in \mathcal{E} \end{array} \right\} \tag{3.12}$$

comes with a natural interpolation operator $I_{\mathrm{M}} : H_0^2(\Omega) + \mathrm{M}(\mathcal{T}) \to \mathrm{M}(\mathcal{T})$.

Definition 3.1 (classical Morley interpolation [8, 12]). Given any function $v \in H_0^2(\Omega) + \mathrm{M}(\mathcal{T})$, the Morley interpolation operator $I_{\mathrm{M}} : H_0^2(\Omega) + \mathrm{M}(\mathcal{T}) \to \mathrm{M}(\mathcal{T})$ defines $I_{\mathrm{M}} v \in \mathrm{M}(\mathcal{T})$ by

$$(v - I_{\mathrm{M}} v)(z) = 0 \quad \text{for } z \in \mathcal{V}(\Omega) \quad \text{and} \quad ⨍_E \frac{\partial (v - I_{\mathrm{M}} v)}{\partial \nu_E} \, \mathrm{d}s = 0 \quad \text{for } E \in \mathcal{E}(\Omega).$$

This interpolation operator possesses the a_{pw}-orthogonality property $v - I_{\mathrm{M}} v \perp_{a_{\mathrm{pw}}} P_2(\mathcal{T})$ for any $v \in H_0^2(\Omega) + \mathrm{M}(\mathcal{T})$. The nonconforming Morley FEM approximates $u \in H_0^2(\Omega)$ with the unique discrete solution $u_{\mathrm{M}} \in \mathrm{M}(\mathcal{T}) =: V_{\mathrm{nc}}$ to

$$a_{\mathrm{pw}}(u_{\mathrm{M}}, v_{\mathrm{M}}) := \int_\Omega D^2_{\mathrm{pw}} u_{\mathrm{M}} : D^2_{\mathrm{pw}} v_{\mathrm{M}} \, \mathrm{d}x = (f, v_{\mathrm{M}})_{L^2(\Omega)} \qquad \forall\, v_{\mathrm{M}} \in \mathrm{M}(\mathcal{T}). \tag{3.13}$$

This represents (2.3) for $\widehat{F} := F \equiv f \in L^2(\Omega)$ and $Q = \mathrm{id}$ while Section 6.2 considers $Q = \mathrm{id}$ and $Q = J_h$ simultaneously in a new a posteriori analysis and Section 7 discusses general sources $F \in V^*$. Note that a_{pw} is a scalar-product in $V + V_{\mathrm{nc}} \equiv H_0^2(\Omega) + \mathrm{M}(\mathcal{T})$ [23]. The residual from Section 3.1 reads $\mathrm{Res} := (f, \cdot)_{L^2(\Omega)} - a_{\mathrm{pw}}(u_{\mathrm{M}}, \cdot) \in V^*$.

3.2.5 Bounds from Morley interpolation

An approach similar to Subsection 3.2.2 for the Crouzeix–Raviart FEM fails immediately because $S_0^2(\mathcal{T}) \cap H_0^2(\Omega)$ is not rich enough: For many triangulations $S_0^2(\mathcal{T}) \cap H_0^2(\Omega) = \{0\}$ is trivial, however not in general [38, Sect. 3.3].

However, the a_{pw}-orthogonality $v - I_{\mathrm{M}} v \perp_{a_{\mathrm{pw}}} P_2(\mathcal{T})$ for all $v \in V$ with the Morley interpolation I_{M} allows the arguments from Subsection 3.2.3 that lead in [2, 36] to

$$|||\mathrm{Res}|||_* := \sup_{v \in V} \frac{\mathrm{Res}(v)}{|||v|||} = \sup_{v \in V} \frac{(f, v - I_{\mathrm{M}} v)_{L^2(\Omega)}}{|||v|||} \leqslant \varkappa_{\mathrm{M}} \|h_{\mathcal{T}}^2 f\|. \tag{3.14}$$

The interpolation error estimate $\|h_{\mathcal{T}}^{-2}(v - I_{\mathrm{M}} v)\| \leqslant \varkappa_{\mathrm{M}} |||v|||$ holds with constant $\varkappa_{\mathrm{M}} \leqslant 0.2575$ [12, Sect. 4].

3.3 Paradigm of unified a posteriori error analysis

The discussion in this subsection departs from the error identity of Theorem 3.1 that includes the dual norm $|||\mathrm{Res}|||_*$ of the residual $\mathrm{Res} \in V^*$. Recall that $u \in V$ solves (1.1) in V and $u_h \in V_h$ solves (2.3).

Subsection 3.1 discussed the error identity (3.1) with the dual norm of the residual given as a supremum over all continuous test functions. Since $u_h \notin V$ in general, the computable (conforming) post-processing $J_h u_h \in V$ serves as its approximation and motivates the error definition $e := u - J_h u_h \in V$ on the continuous level and $I_h I_{\mathrm{nc}} u - u_h \in V_h$ on the discrete level. The efficient error estimator $\|u_h - J_h u_h\|_h$ from (3.5) is computable and a triangle inequality in the norm $\|\cdot\|_h$ and (2.1) lead to

$$\|u - u_h\|_h \leqslant |||e||| + \|u_h - J_h u_h\|_h, \qquad |||e||| \leqslant \|u - u_h\|_h + \|u_h - J_h u_h\|_h. \tag{3.15}$$

Recall $\widehat{F}|_V = F$ from (2.3). The first argument to establish an alternative abstract error bound applies the continuous (resp. discrete) equation (1.1) (resp. (2.3)) to the test function $J_h e_h \in V$ (resp. $e_h := I_h I_{\mathrm{nc}} e \in V_h$), namely,

$$a_h(u_h, e_h) = \widehat{F}(Q e_h) = a(u, J_h e_h) - \widehat{F}(J_h e_h - Q e_h). \tag{3.16}$$

For $Q = J_h$, the last term vanishes and (3.16) becomes the key identity $a_h(u_h, e_h) = a(u, J_h e_h)$. The second argument is the link of $a_h(u_h, e_h)$ to $a(J_h u_h, J_h e_h)$ by (H),

$$a_h(u_h, e_h) - a(J_h u_h, J_h e_h) \leqslant \Lambda_{\mathrm{H}} \|u_h - J_h u_h\|_h \|e_h\|_h. \tag{3.17}$$

The (generalized) key identity (3.16) shows that the left-hand side of (3.17) is equal to $a(e, J_h e_h) - \widehat{F}(J_h e_h - Q e_h)$. This and the abbreviation $w := e - J_h e_h$ show

$$|||e|||^2 = a(e, w) + a(e, J_h e_h) \leqslant F(w) - a(J_h u_h, w) + \widehat{F}((J_h - Q) e_h) + \Lambda_{\mathrm{H}} \|u_h - J_h u_h\|_h \|e_h\|_h$$

with $a(u, w) = F(w)$ in the last step. This, the Cauchy inequality $a_{\mathrm{pw}}(u_h - J_h u_h, w) \leqslant \|u_h - J_h u_h\|_h |||w|||$ using (2.1), and the residual $\mathrm{Res} := F - a_{\mathrm{pw}}(u_h, \cdot) \in V^*$ reveal

$$|||e|||^2 \leqslant (|||w||| + \Lambda_{\mathrm{H}} \|e_h\|_h) \, \|u_h - J_h u_h\|_h + \mathrm{Res}(w) + \widehat{F}(J_h e_h - Q e_h). \tag{3.18}$$

Theorem 3.2 (alternative abstract error bound). *Let $J_h \in L(V_h; V)$ be a quasi-optimal smoother and suppose* (2.5)–(2.7) *and* (H). *Then there exists a constant $C_1 > 0$ such that the error $e := u - J_h u_h \in V$ for the solution $u \in V$ to* (1.1) *and $u_h \in V_h$ to* (2.3) *satisfies*

$$\|u - u_h\|_h^2 + |||e|||^2 \leqslant C_1^2 \left(\|u_h - J_h u_h\|_h^2 + \mathrm{Res}(e - J_h I_h I_{\mathrm{nc}} e) + \widehat{F}(J_h e_h - Q e_h) \right). \tag{3.19}$$

Proof. Abbreviate $w := e - J_h e_h \in V$ with $e_h := I_h I_{\mathrm{nc}} e \in V_h$. Lemma 2.1 leads to $C_2^{-1} |||w||| \leqslant |||e - I_{\mathrm{nc}} e|||_{\mathrm{pw}} \leqslant (1 + \|I_{\mathrm{nc}}\|) |||e|||$ for $C_2 := (1 + \Lambda_{\mathrm{J}})(1 + \Lambda_{\mathrm{nc}})(1 + \Lambda_h)$ and the operator norms control $\|e_h\|_h \leqslant \|I_{\mathrm{nc}}\| \|I_h\| |||e|||$. This, a Young inequality, and (3.18) show

$$\frac{1}{2} |||e|||^2 \leqslant \frac{1}{2} C_3^2 \|u_h - J_h u_h\|_h^2 + \mathrm{Res}(w) + \widehat{F}(J_h e_h - Q e_h)$$

with $C_3 := C_2(1 + \|I_{\mathrm{nc}}\|) + \Lambda_{\mathrm{H}} \|I_{\mathrm{nc}}\| \|I_h\|$. This and (3.15) conclude the proof of (3.19) for $C_1^2 := \max\{2 + 3C_3^2, 6\}$. □

The equivalence $\|u_h - J_h u_h\|_h \approx \min_{v \in V} \|v - u_h\|_h$ from (2.4) provides

$$\|u - u_h\|_h^2 + |||e|||^2 \lesssim \mathrm{Res}(e - J_h I_h I_{\mathrm{nc}} e) + \widehat{F}(J_h e_h - Q e_h) + \min_{v \in V} \|v - u_h\|_h^2$$

as an equivalent formulation of (3.19). The remaining parts of this paper discuss explicit bounds of the right-hand side of (3.19) for a simultaneous a posteriori analysis of five nonstandard FEMs for the biharmonic equation.

4 Examples of lowest-order finite element schemes

This section introduces the spaces and transfer operators for five lowest-order methods for the biharmonic equation.

4.1 Three second-order finite element spaces

Recall the space of piecewise polynomials $P_k(\mathcal{T})$ of total degree at most $k \in \mathbb{N}$ from Subsection 1.3. Let $S^k(\mathcal{T}) := P_k(\mathcal{T}) \cap C^0(\Omega)$ and $S_0^k(\mathcal{T}) := \{p \in S^k(\mathcal{T}) \mid p_{|\partial\Omega} = 0\} = P_k(\mathcal{T}) \cap H_0^1(\Omega)$. The associated L^2 projection $\Pi_k : L^2(\Omega) \to P_k(\mathcal{T})$ is defined by the L^2 orthogonality $(1 - \Pi_k)v \perp P_k(\mathcal{T})$ for all $v \in L^2(\Omega)$. Recall the nonconforming Morley space $\mathrm{M}(\mathcal{T})$ from (3.12). Throughout the remaining parts of this paper on the biharmonic equation, specify $V_{\mathrm{nc}} := \mathrm{M}(\mathcal{T})$, $V := H_0^2(\Omega) \subset \widehat{V} := H^2(\mathcal{T})$, and

$$V_h := \begin{cases} \mathrm{M}(\mathcal{T}) & \text{for Morley} \\ P_2(\mathcal{T}) & \text{for dG or WOPSIP} \\ S_0^2(\mathcal{T}) & \text{for } C^0\text{IP}. \end{cases} \tag{4.1}$$

4.2 Hilbert space of piecewise H^2 functions

The semi-scalar product $a_{\mathrm{pw}} := (D_{\mathrm{pw}}^2 \cdot, D_{\mathrm{pw}}^2 \cdot)_{L^2(\Omega)}$ in $\widehat{V} := H^2(\mathcal{T})$ extends the energy scalar product $a := a_{\mathrm{pw}}|_{V \times V}$ and the subspace $(\mathrm{M}(\mathcal{T}), a_{\mathrm{pw}})$ is a Hilbert space. Recall the jump $[v]_E$ resp. normal jump $[\partial v / \partial \nu_E]_E$ across an edge $E \in \mathcal{E}$ of a piecewise function $v \in H^1(\mathcal{T})$ resp. $v \in H^2(\mathcal{T})$ from Subsections 3.2.3 and 3.2.4. Let $\mathcal{V}(E)$ denote the vertices of the edge $E \in \mathcal{E}$. Define the semi-scalar product $j_h : \widehat{V} \times \widehat{V}$, for any $v, w \in \widehat{V}$, by

$$j_h(v, w) := \sum_{E \in \mathcal{E}} \left(\sum_{z \in \mathcal{V}(E)} \frac{[v]_E(z)}{h_E} \frac{[w]_E(z)}{h_E} + \fint_E \left[\frac{\partial v}{\partial \nu_E}\right]_E \mathrm{d}s \fint_E \left[\frac{\partial w}{\partial \nu_E}\right]_E \mathrm{d}s \right). \tag{4.2}$$

Since $j_h(v, \cdot) = 0$ vanishes for any $v \in V + \mathrm{M}(\mathcal{T})$, $(H^2(\mathcal{T}), a_{\mathrm{pw}} + j_h)$ is a Hilbert space with the induced norm $\|\cdot\|_h$ from (2.1).

Remark 4.1 (completeness of $(\widehat{V}, a_{\mathrm{pw}} + j_h)$). It is clear [24, Sect. 4.1] that $(\widehat{V}, \|\cdot\|_h)$ is a normed linear space. Recall that $(H^2(\mathcal{T}), \|\cdot\|_{H^2(\mathcal{T})})$ equipped with the piecewise H^2 norm $\|\cdot\|_{H^2(\mathcal{T})}^2 := \sum_{T \in \mathcal{T}} \|\cdot\|_{H^2(T)}^2$ is a Banach space.

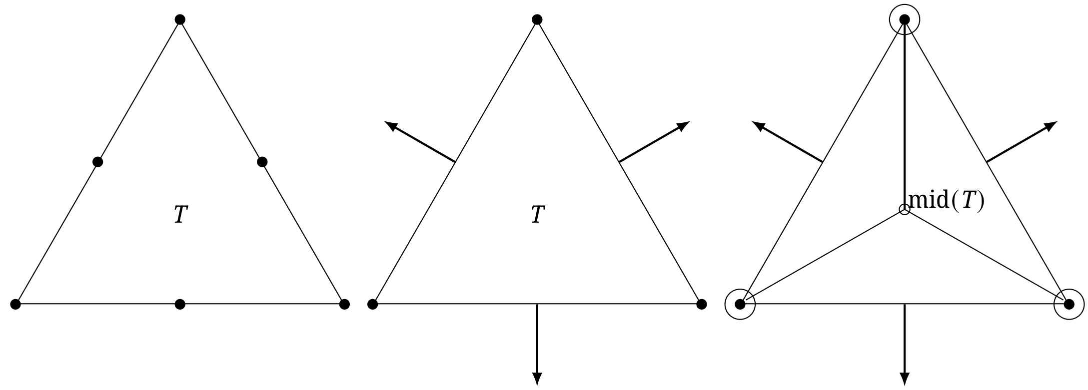

Fig. 3: The Lagrange P_2, the Morley, and the HCT finite element (left to right).

Let $Q : H^2(\mathcal{T}) \to P_1(\mathcal{T})$ denote the H^2 orthogonal projection onto the finite dimensional space $P_1(\mathcal{T}) \subset H^2(\mathcal{T})$ and set $X := (1 - Q)H^2(\mathcal{T})$. The Bramble–Hilbert lemma (see [30, Lem. 11.9]) asserts that $|||\cdot|||_{\mathrm{pw}}$ is a norm on X stronger than the piecewise H^2 norm $\|\cdot\|_{H^2(\mathcal{T})} \lesssim \|\cdot\|_h$. Since [24, Thm. 4.1] shows that $\|\cdot\|_h \lesssim \|\cdot\|_{H^2(\mathcal{T})}$ is also weaker than the piecewise H^2 norm, both norms are equivalent on X and X is complete. The direct sum $\widehat{V} = X \oplus P_1(\mathcal{T})$ of two complete spaces is complete.

4.3 Classical and averaged Morley interpolation

The classical Morley interpolant from Subsection 3.2.4 is defined for functions in $V + \mathrm{M}(\mathcal{T})$ and has an extension to piecewise H^2 functions. Define the average $\langle\varphi\rangle_E := \frac{1}{2}\left(\varphi|_{T_+} + \varphi|_{T_-}\right)$ of $\varphi \in H^1(\mathcal{T})$ across an interior edge $E = \partial T_+ \cap \partial T_- \in \mathcal{E}(\Omega)$ of the adjacent triangles T_+ and $T_- \in \mathcal{T}$ as in Fig. 2 and $\langle\varphi\rangle_E := \varphi|_E$ along a boundary edge $E \in \mathcal{E}(\partial\Omega)$. Let $\mathcal{T}(z) := \{T \in \mathcal{T} \mid z \in T\}$ denote the $|\mathcal{T}(z)| \in \mathbb{N}$ many neighbouring triangles of $z \in T \in \mathcal{T}$.

Definition 4.1 (Morley interpolation [24]). Given any piecewise function $v_{\mathrm{pw}} \in \widehat{V}$, the Morley interpolation operator $I_{\mathrm{M}} : \widehat{V} \to V_{\mathrm{nc}}$ sets the degrees of freedom of the Morley finite element function $I_{\mathrm{M}} v_{\mathrm{pw}} \in V_{\mathrm{nc}} := \mathrm{M}(\mathcal{T})$ by

$$I_{\mathrm{M}} v_{\mathrm{pw}}(z) := |\mathcal{T}(z)|^{-1} \sum_{T \in \mathcal{T}(z)} (v_{\mathrm{pw}}|_T)(z) \qquad \text{for } z \in \mathcal{V}(\Omega)$$
$$⨍_E \frac{\partial I_{\mathrm{M}} v_{\mathrm{pw}}}{\partial \nu_E}\,\mathrm{d}s := ⨍_E \left\langle \frac{\partial v_{\mathrm{pw}}}{\partial \nu_E} \right\rangle_E \mathrm{d}s \qquad \text{for } E \in \mathcal{E}(\Omega).$$

It is well known that there is a unique quadratic polynomial $I_{\mathrm{M}} v_{\mathrm{pw}}|_T \in P_2(T)$ that assumes the above values $(I_{\mathrm{M}} v_{\mathrm{pw}})(z)$ and $⨍_E \partial I_{\mathrm{M}} v_{\mathrm{pw}}/\partial \nu_E\,\mathrm{d}s$ at $z \in \mathcal{V}(T)$ and for all $E \in \mathcal{E}(T)$. Explicit formulas for the basis functions can be found in [14]. This definition extends the classical Morley interpolation from Definition 3.1 to piecewise H^2 functions in $\widehat{V} \equiv H^2(\mathcal{T})$. For any $v \in H^2_0(\Omega) + \mathrm{M}(\mathcal{T})$, the a_{pw}-orthogonality

$$a_{\mathrm{pw}}(v - I_{\mathrm{M}} v, w_2) = 0 \qquad \forall\, w_2 \in P_2(\mathcal{T}) \tag{4.3}$$

verifies the best-approximation property

$$|||v - I_{\mathrm{M}} v|||_{\mathrm{pw}} = \min_{v_2 \in P_2(\mathcal{T})} |||v - v_2|||_{\mathrm{pw}}. \tag{4.4}$$

This does not extend to discontinuous functions $v_h \in H^2(\mathcal{T})$ in general. Recall $\|\cdot\|_h$ from (2.1).

Theorem 4.1 (interpolation error [24, Thm. 4.3]). *Any piecewise smooth function $v_{\rm pw} \in H^2(\mathcal{T})$ and its Morley interpolation $I_{\rm M} v_{\rm pw} \in {\rm M}(\mathcal{T})$ from Definition 4.1 satisfy*

$$(a) \qquad \|v_{\rm pw} - I_{\rm M} v_{\rm pw}\|_h \lesssim \|(1-\Pi_0) D^2_{\rm pw} v_{\rm pw}\| + j_h(v_{\rm pw}, v_{\rm pw})^{1/2}$$

$$(b) \qquad \sum_{m=0}^{2} h_{\mathcal{T}}^{m-2} |v_{\rm pw} - I_{\rm M} v_{\rm pw}|_{H^m(\mathcal{T})} \approx \min_{w_{\rm M} \in {\rm M}(\mathcal{T})} \|v_{\rm pw} - w_{\rm M}\|_h \approx \|v_{\rm pw} - I_{\rm M} v_{\rm pw}\|_h.$$

Since $I_{\rm M} \in L(\widehat{V}; {\rm M}(\mathcal{T}))$ is a bounded operator, the Cauchy inequality and the best-approximation property (4.4) verify (2.6) for $I_{\rm nc} := I_{\rm M}$ with $\Lambda_{\rm nc} := \Lambda_{\rm M} := 2 + \|I_{\rm M}\|_h$. Indeed, for arbitrary $v_2 \in P_2(\mathcal{T})$ and $v \in V$,

$$\|v_2 - I_{\rm M} v_2\|_h \leqslant \|v_2 - v\|_h + \|\!|v - I_{\rm M} v|\!\|_{\rm pw} + \|I_{\rm M}(v - v_2)\|_h \leqslant \Lambda_{\rm M} \|v - v_2\|_h. \tag{4.5}$$

4.4 Transfer operator I_h

The abstract setting from Section 2.1 requires a transfer operator I_h with (2.5) from $V_{\rm nc} := {\rm M}(\mathcal{T})$ into V_h defined in (4.1) for the different schemes. The natural choice $I_h := {\rm id}$ for the Morley, dG, and WOPSIP method with $V_{\rm nc} \subseteq V_h$ fulfils (2.5) with $\Lambda_h = 0$. The situation is different for the C^0IP method with $V_{\rm nc} \nsubseteq V_h := S^2_0(\mathcal{T})$ and requires the Lagrange interpolation $I_h := I_{\rm C} : {\rm M}(\mathcal{T}) \to S^2_0(\mathcal{T})$ defined, for all $v_{\rm M} \in {\rm M}(\mathcal{T})$, by

$$(I_{\rm C} v_{\rm M})(z) = \begin{cases} v_{\rm M}(z) & \forall\, z \in \mathcal{V} \\ \langle v_{\rm M} \rangle_E\,(z) & \text{for } z = {\rm mid}(E),\ E \in \mathcal{E}(\Omega) \\ 0 & \text{for } z = {\rm mid}(E),\ E \in \mathcal{E}(\partial\Omega). \end{cases} \tag{4.6}$$

(It is well known that there exists a unique $I_{\rm C} v_{\rm M}|_T \in P_2(T)$ with prescribed values at the vertices and edge midpoints from the unisolvence of the P_2 Lagrange finite element.) Lemma 3.2 in [15] establishes (2.5) for the operator $I_h = I_{\rm C}$ with $\Lambda_h \approx 1$.

4.5 Companion operator J

A conforming finite-dimensional subspace of $H^2_0(\Omega)$ is the Hsieh–Clough–Tocher (HCT) [27, Ch. 6] space ${\rm HCT}(\mathcal{T}) := \{v \in H^2_0(\Omega) :\ v|_T \in P_3(\mathcal{K}(T)) \text{ for all } T \in \mathcal{T}\}$ with the subtriangulation $\mathcal{K}(T) := \{{\rm conv}\{E, {\rm mid}(T)\} : E \in \mathcal{E}(T)\}$ of $T \in \mathcal{T}$ obtained by joining the vertices of T with ${\rm mid}(T)$. Figure 3 shows the degrees of freedom of the HCT finite element that extend those of the Morley element and facilitate the design of a right-inverse to $I_{\rm M} : \widehat{V} \to {\rm M}(\mathcal{T})$.

Lemma 4.1 (right-inverse [25, 33, 42]). *There exists a linear right-inverse $J : {\rm M}(\mathcal{T}) \to HCT(\mathcal{T}) + P_8(\mathcal{T}) \cap H^2_0(\Omega)$ for $I_{\rm M} : V \to {\rm M}(\mathcal{T})$ and a constant $\Lambda_{\rm J}$, that exclusively depends on the shape regularity, such that any $v_{\rm M} \in {\rm M}(\mathcal{T})$ satisfies*

$$\|\!|v_{\rm M} - J v_{\rm M}|\!\|_{\rm pw} \leqslant \Lambda_{\rm J} \min_{v \in V} \|\!|v_{\rm M} - v|\!\|_{\rm pw}.$$

See [25, Sect. 5] for the definition of $J \in L(V_{\rm nc}; V)$. Note that Lemma 4.1 verifies (2.7) for the conforming companion J. Recall from the previous subsections that $I_{\rm nc} := I_{\rm M} \in L(\widehat{V}; V_{\rm nc})$ and $I_h \in L(V_{\rm nc}; V_h)$ verify (2.6)–(2.5). An immediate consequence of Lemma 2.2 is that $J_h := J I_{\rm M} \in L(\widehat{V}; V)$ is a quasi-optimal smoother. We refer to [25] for a 3D version.

5 Building blocks for explicit residual-based a posteriori error estimators

This section establishes bounds on the error contributions in the right-hand side of (3.19). Recall the residual $\mathrm{Res} := F - a_{\mathrm{pw}}(u_h, \cdot) \in V^*$ from Section 3 and set $V_{\mathrm{nc}} := \mathrm{M}(\mathcal{T})$ with interpolation operator $I_{\mathrm{nc}} \equiv I_{\mathrm{M}}$ and quasi-optimal smoother $J_h \equiv JI_{\mathrm{M}}$ throughout the remaining parts of this paper.

5.1 Estimates for $1 - J_h I_h I_{\mathrm{M}}$ and $(1 - J_h) I_h I_{\mathrm{M}}$

The linear operators $1 - J_h I_h I_{\mathrm{M}} : V \to V$ and $I_h I_{\mathrm{M}} - J_h I_h I_{\mathrm{M}} : V \to \widehat{V}$ are stable in the energy norm.

Lemma 5.1 (stability). *Any $v \in V = H_0^2(\Omega)$ with $\widehat{w} = (1 - J_h I_h I_{\mathrm{M}})v \in V$ or $\widehat{w} = (1 - J_h) I_h I_{\mathrm{M}} v \in \widehat{V}$ satisfies*

$$\sum_{m=0}^{2} |h_{\mathcal{T}}^{m-2} \widehat{w}|_{H^m(\Omega)}^2 + \sum_{E \in \mathcal{E}(\Omega)} \left(\|h_E^{-3/2} \widehat{w}\|_{L^2(E)}^2 + \|h_E^{-1/2} \nabla \widehat{w}\|_{L^2(E)}^2 \right) \leqslant C_4^2 \|\!|v|\!\|^2.$$

Proof. Since J is a right-inverse of I_{M}, the functions $v, v_{\mathrm{M}} := I_{\mathrm{M}} v, v_h := I_h v_{\mathrm{M}}, I_{\mathrm{M}} v_h$ and $J_h v_h$ in $H^2(\mathcal{T})$ are continuous at any vertex $z \in \mathcal{V}$ and coincide at $z \in \mathcal{V}$. Hence, $\widehat{w}|_T \in H^2(T)$ vanishes at the three vertices of the triangle $T \in \mathcal{T}$. It is textbook analysis [3, 6, 27, 30] to derive the bounds

$$\sum_{m=0}^{2} |h_{\mathcal{T}}^{m-2} \widehat{w}|_{H^m(T)}^2 \leqslant C_{\mathrm{BH}}^2 |\widehat{w}|_{H^2(T)}^2 \tag{5.1}$$

from an application of the Bramble–Hilbert lemma with a constant $C_{\mathrm{BH}} > 0$ and we refer to [17, Sect. 3] for explicit constants in terms of the maximal angles in the triangle $T \in \mathcal{T}$. The sum of all those estimates (5.1) results in

$$\sum_{m=0}^{2} |h_{\mathcal{T}}^{m-2} \widehat{w}|_{H^m(\Omega)}^2 \leqslant C_{\mathrm{BH}}^2 \|\!|\widehat{w}|\!\|_{\mathrm{pw}}^2.$$

The previous estimate, $\|\!|\widehat{w}|\!\|_{\mathrm{pw}} \leqslant \|\widehat{w}\|_h \leqslant C_5 \|\!|v|\!\|$ with $C_5 := \max\{1, C_{\mathrm{J}}\}(1+\Lambda_{\mathrm{J}})(1+\Lambda_{\mathrm{nc}})(1+\Lambda_h)$ from Lemma 2.1–2.2 and (4.4) conclude the proof of

$$\sum_{m=0}^{2} |h_{\mathcal{T}}^{m-2} \widehat{w}|_{H^m(\Omega)}^2 \leqslant C_{\mathrm{BH}}^2 C_5^2 \|\!|v|\!\|^2. \tag{5.2}$$

Given any interior edge $E \in \mathcal{E}(\Omega)$ with adjacent triangle $T(E) \in \mathcal{T}$, the trace inequality [30, Eq. (12.17)] provides a constant $C_{\mathrm{tr}} > 0$ exclusively depending on the shape-regularity with

$$h_E^{-3/2} \|\widehat{w}\|_{L^2(E)} + h_E^{-1/2} \|\nabla \widehat{w}\|_{L^2(E)} \leqslant C_{\mathrm{tr}} \left(h_T^{-2} \|\widehat{w}\|_{L^2(T(E))} + h_T^{-1} |\widehat{w}|_{H^1(T(E))} + |\widehat{w}|_{H^2(T(E))} \right).$$

This and the sum over the interior edges $E \in \mathcal{E}(\Omega)$ result in

$$\sum_{E \in \mathcal{E}(\Omega)} \left(h_E^{-3/2} \|\widehat{w}\|_{L^2(E)} + h_E^{-1/2} \|\nabla \widehat{w}\|_{L^2(E)} \right)^2 \leqslant 3 C_{\mathrm{tr}}^2 \sum_{m=0}^{2} \sum_{E \in \mathcal{E}(\Omega)} |h_T^{m-2} \widehat{w}|_{H^m(T(E))}^2.$$

Since every triangle $T(E) \in \mathcal{T}$ is counted at most 3 times (once for every edge $E \in \mathcal{E}(T(E))$) in the last sum, the claim follows with $C_4 := (3C_{\mathrm{tr}} + 1) C_{\mathrm{BH}} C_5$. □

Corollary 5.1 (bound for $F \in L^2(\Omega)$). Any $F = f \in L^2(\Omega)$ and $v \in V = H_0^2(\Omega)$ with $\widehat{w} := (1 - J_h I_h I_{\mathrm{M}})v \in V$ or $\widehat{w} := (1 - J_h) I_h I_{\mathrm{M}} v \in \widehat{V}$ satisfy

$$\int_{\Omega} f\, \widehat{w}\ \mathrm{d}x \leqslant C_6 \|h_{\mathcal{T}}^2 f\|\ \|\!|v|\!\|. \tag{5.3}$$

Proof. This follows from Lemma 5.1 and a Cauchy inequality in $L^2(\Omega)$ in

$$\int_\Omega f\,\widehat{w}\;\mathrm{d}x \leqslant \|h_{\mathcal{T}}^2 f\|\|h_{\mathcal{T}}^{-2}\widehat{w}\| \leqslant C_6\|h_{\mathcal{T}}^2 f\|\;|||v|||. \qquad \square$$

Define the oscillations of $f \in L^2(\Omega)$ by $\mathrm{osc}_2(f, T) := \|h_{\mathcal{T}}^2(f - \Pi_2 f)\|_{L^2(T)}$ and abbreviate

$$\mathrm{osc}_2(f, \mathcal{S}) := \sqrt{\sum_{T\in\mathcal{S}} \mathrm{osc}_2^2(f, T)}$$

for a subset $\mathcal{S} \subseteq \mathcal{T}$ of triangles in $\mathcal{T}$. The efficiency of the term $\|h_{\mathcal{T}}^2 f\|_{L^2(T)}$ is known, e.g., from [5, Lem. 4.2, Rem. 4.4]; Section 7 treats a more general source $F \in V^*$.

Lemma 5.2 (efficiency up to oscillations [5]). *Let $u \in V$ be the weak solution to* (1.1) *for a right-hand side $F = f \in L^2(\Omega)$. Then $\|h_{\mathcal{T}}^2 f\|_{L^2(T)} \lesssim |u - I_{\mathrm{M}}u|_{H^2(T)} + \mathrm{osc}_2(f, T)$.*

5.2 Error estimates for $a_{\mathrm{pw}}(v_h, w)$

Recall the abbreviation $w := v - J_h I_h I_{\mathrm{M}} v$ for $v \in V$. Since J from Subsection 4.5 is a right-inverse of the Morley interpolation I_{M} from Subsection 4.3, the key observation for the situation $I_h = \mathrm{id}$ is

$$I_{\mathrm{M}}w = I_{\mathrm{M}}v - I_{\mathrm{M}}J_h I_h I_{\mathrm{M}}v = I_{\mathrm{M}}(v - I_h I_{\mathrm{M}}v) = 0. \tag{5.4}$$

This is the case for the Morley, dG, and WOPSIP methods and, hence, the a-orthogonality of the Morley interpolation of $w \in V$ and $I_{\mathrm{M}}w = 0$ imply $a_{\mathrm{pw}}(u_h, w) = 0$. For the C^0IP method with $V_h = S_0^2(\mathcal{T})$ and $I_h = I_{\mathrm{C}} \neq \mathrm{id}$ from Subsection 4.4 the situation differs and is the content of the remaining part of this subsection.

Lemma 5.3 (bound for $a_{\mathrm{pw}}(v_h, w)$). *Any $v_h \in V_h$ and $v \in V$ with $w := v - J_h I_h I_{\mathrm{M}} v$ satisfies*

$$|a_{\mathrm{pw}}(v_h, w)| \leqslant \begin{cases} 0 & \text{if } I_h = \mathrm{id} \\ C_4\sqrt{\sum_{E\in\mathcal{E}(\Omega)} h_E\|[\partial^2_{\nu\nu}v_h]_E\|^2_{L^2(E)}}\;|||v||| & \text{if } I_h = I_{\mathrm{C}}. \end{cases}$$

Proof. With the remark succeeding (5.4), the claim holds for $I_h = \mathrm{id}$ and it remains the case $I_h = I_{\mathrm{C}}$. Since the piecewise Hessian $D^2_{\mathrm{pw}}v_h$ of $v_h \in S_0^2(\mathcal{T})$ is piecewise constant, no volume contributions arise in a piecewise integration by parts with the conforming test function $w \in V$. A careful re-arrangement of the contributions along the boundary ∂T of $T \in \mathcal{T}$ reveals

$$a_{\mathrm{pw}}(v_h, w) = \sum_{E\in\mathcal{E}(\Omega)} \int_E \nabla w \cdot [D^2_{\mathrm{pw}}v_h]_E \nu_E \,\mathrm{d}s. \tag{5.5}$$

Recall from the proof of Lemma 5.1 that $w(z) = 0$ vanishes at any vertex, whence $\int_E \partial w/\partial s\;\mathrm{d}s = 0$ on any edge $E \in \mathcal{E}$. Since the matrix $[D^2_{\mathrm{pw}}v_h]_E \in P_0(E; \mathbb{S})$ is constant, the split $\nabla w = (\partial w/\partial s)\tau_E + (\partial w/\partial\nu_E)\nu_E$ along $E \in \mathcal{E}$ and the Cauchy inequality show

$$\int_E \nabla w \cdot [D^2_{\mathrm{pw}}v_h]_E \nu_E \,\mathrm{d}s = \int_E \frac{\partial w}{\partial \nu_E}[\partial^2_{\nu\nu}v_h]_E \,\mathrm{d}s \leqslant h_E^{-1/2}\left\|\frac{\partial w}{\partial \nu_E}\right\|_{L^2(E)} h_E^{1/2}\left\|[\partial^2_{\nu\nu}v_h]_E\right\|_{L^2(E)}.$$

Notice that the trace of $\nabla w \cdot \nu_E$ along E is continuous for $w \in V$. This, a Cauchy inequality in ℓ^2, and $\|\partial w/\partial\nu_E\|_{L^2(E)} \leqslant \|\nabla w\|_{L^2(E)}$ verify

$$a_{\mathrm{pw}}(v_h, w) \leqslant \sqrt{\sum_{E\in\mathcal{E}(\Omega)} h_E\|[\partial^2_{\nu\nu}v_h]_E\|^2_{L^2(E)}}\sqrt{\sum_{E\in\mathcal{E}(\Omega)} h_E^{-1}\|\nabla w\|^2_{L^2(E)}}.$$

This and Lemma 5.1 conclude the proof. $\square$

The efficiency estimate of the jump contributions in Lemma 5.3 is known, e.g., from the C^0IP method [4]. For any edge $E \in \mathcal{E}$, the sub-triangulation $\mathcal{T}(\omega(E)) := \{T \in \mathcal{T} \mid E \subset \partial T\}$ in the edge-patch $\omega(E) := \operatorname{int}(T_+ \cup T_-)$ consists of one or two triangles.

Lemma 5.4 (see [4, Lem. 4.3]). *Let $u \in V$ solve (1.1) for $F = f \in L^2(\Omega)$. Any $v_h \in P_2(\mathcal{T})$ and any edge $E \in \mathcal{E}$ satisfy*

$$h_E^{1/2}\|[\partial^2_{\nu\nu} v_h]_E\|_{L^2(E)} \lesssim |u - v_h|_{H^2(\mathcal{T}(\omega(E)))} + \operatorname{osc}_2(f, \mathcal{T}(\omega(E))).$$

Proof. The proof of [4, Lem. 4.3] for the jump $[\partial^2_{\nu\nu} v_h]_E$ of any $v_h \in P_2(\mathcal{T})$ shows that

$$h_E^{1/2}\|[\partial^2_{\nu\nu} v_h]_E\|_{L^2(E)} \lesssim |u - v_h|_{H^2(\mathcal{T}(\omega(E)))} + \|h^2_{\mathcal{T}} f\|_{L^2(\omega(E))}.$$

Lemma 5.2 and $|u - I_{\mathrm{M}} u|_{H^2(T)} = \min_{v_h \in P_2(T)} |u - v_h|_{H^2(T)}$ as in (4.4) conclude the proof. □

5.3 Estimate of $\|v_h - J_h v_h\|_h$

This subsection discusses reliable and efficient bounds of $\|v_h - J_h v_h\|_h$ in terms of two different jump terms that appear in the a posteriori analysis, e.g., in [2, 4, 5, 36].

Theorem 5.1 (reliability and efficiency of $\|v_h - J_h v_h\|_h$). *Any $v_h \in V_h$ satisfies*

$$\min_{v \in V} \|v - v_h\|_h^2 \approx \|v_h - J_h v_h\|_h^2 \approx \sum_{E \in \mathcal{E}} h_E \|[D^2_{\mathrm{pw}} v_h]_E \tau_E\|^2_{L^2(E)} + j_h(v_h, v_h)$$
$$\approx \sum_{E \in \mathcal{E}} \left(h_E^{-3} \|[v_h]_E\|^2_{L^2(E)} + h_E^{-1} \left\| \left[\frac{\partial v_h}{\partial \nu_E} \right]_E \right\|^2_{L^2(E)} \right).$$

The remaining parts of this subsection are devoted to the proof and depart with the following generalization of [36, Thm. 2.1].

Lemma 5.5 (bound for $\|v_h - J_h v_h\|_h$). *Any $v_h \in V_h$ satisfies*

$$C_4^{-1}\|v_h - J_h v_h\|_h^2 \leqslant \sum_{E \in \mathcal{E}} h_E \|[D^2_{\mathrm{pw}} v_h]_E \tau_E\|^2_{L^2(E)} + j_h(v_h, v_h).$$

Proof. Given any $v_h \in V_h$, set $v_{\mathrm{M}} := I_{\mathrm{M}} v_h \in \mathrm{M}(\mathcal{T})$. A triangle inequality and (2.1) verify $\|v_h - J_h v_h\|_h \leqslant \|v_h - v_{\mathrm{M}}\|_h + |||v_{\mathrm{M}} - J v_{\mathrm{M}}|||_{\mathrm{pw}}$. It follows from [25, Lem. 5.1] that

$$|||v_{\mathrm{M}} - J v_{\mathrm{M}}|||^2_{\mathrm{pw}} \lesssim \sum_{E \in \mathcal{E}} h_E \|[D^2_{\mathrm{pw}} v_{\mathrm{M}}]_E \tau_E\|^2_{L^2(E)}.$$

This, a triangle inequality, and the discrete trace inequality $h_E^{1/2}\|D^2_{\mathrm{pw}}(v_h - v_{\mathrm{M}})\|_{L^2(E)} \lesssim \|D^2(v_h - v_{\mathrm{M}})\|_{L^2(T)}$ from [30, Lem. 12.8] result in

$$|||v_{\mathrm{M}} - J v_{\mathrm{M}}|||^2_{\mathrm{pw}} \lesssim \sum_{E \in \mathcal{E}} h_E \|[D^2_{\mathrm{pw}} v_h]_E \tau_E\|^2_{L^2(E)} + |||v_h - v_{\mathrm{M}}|||^2_{\mathrm{pw}}.$$

This and $|||v_h - v_{\mathrm{M}}|||_{\mathrm{pw}} \leqslant \|v_h - v_{\mathrm{M}}\|_h \lesssim j_h(v_h, v_h)^{1/2}$ from (2.1) and Theorem 4.1.a with $D^2_{\mathrm{pw}} v_h \in P_0(\mathcal{T})$ conclude the proof. □

The inverse inequality leads to an alternative upper bound in Lemma 5.5.

Lemma 5.6 (alternative bound). *Any $v_h \in V_h$ and any edge $E \in \mathcal{E}$ satisfy*

$$h_E \|[D^2_{\mathrm{pw}} v_h]_E \tau_E\|^2_{L^2(E)} + \sum_{z \in \mathcal{V}(E)} \frac{|[v_h]_E(z)|^2}{h_E^2} + \left| \fint_E \left[\frac{\partial v_h}{\partial \nu_E} \right]_E \mathrm{d}s \right|^2$$
$$\leqslant C_7 \left(h_E^{-3} \|[v_h]_E\|^2_{L^2(E)} + h_E^{-1} \|[\partial v_h / \partial \nu_E]_E\|^2_{L^2(E)} \right).$$

Proof. The split $D^2 v_h \cdot \tau_E = (\partial^2 v_h/\partial s\partial s)\tau_E + (\partial^2 v_h/\partial s\partial \nu_E)\nu_E$, the Cauchy inequality, and the linearity of the jump show

$$\|[D^2_{\mathrm{pw}} v_h]_E \cdot \tau_E\|_{L^2(E)} \leqslant \left\| \frac{\partial^2}{\partial s\partial s}\,[v_h]_E \right\|_{L^2(E)} + \left\| \frac{\partial}{\partial s}\left[\frac{\partial v_h}{\partial \nu_E}\right]_E \right\|_{L^2(E)}.$$

The inverse inequality [30, Lem. 12.1] states the existence of a constant $C_{\mathrm{inv}} > 0$ with $|p|_{H^m(E)} \leqslant C_{\mathrm{inv}} h_E^{-m} \|p\|_{L^2(E)}$ and $\|p\|_{L^p(E)} \leqslant C_{\mathrm{inv}} h_E^{1/p-1/q} \|p\|_{L^q(E)}$ for any $p \in P_2(E)$ and $m \in \mathbb{N}_0$, $1 \leqslant p, q \leqslant \infty$. Since $[v_h]_E$ and $[\partial v_h/\partial \nu_E]_E$ are quadratic polynomials on E, this shows

$$h_E^{1/2}\|[D^2_{\mathrm{pw}} v_h]_E \cdot \tau_E\|_{L^2(E)} \leqslant C_{\mathrm{inv}}\left(h_E^{-3/2}\,\|[v_h]_E\|_{L^2(E)} + h_E^{-1/2}\|[\partial v_h/\partial\nu_E]_E\|_{L^2(E)}\right)$$

$$\sum_{z\in\mathcal{V}(E)} \frac{|[v_h]_E(z)|}{h_E} \leqslant 2h_E^{-1}\|[v_h]_E\|_{L^\infty(E)} \leqslant 2C_{\mathrm{inv}} h_E^{-3/2}\,\|[v_h]_E\|_{L^2(E)}$$

$$\left| \fint_E \left[\frac{\partial v_h}{\partial \nu_E}\right]_E \mathrm{d}s \right| \leqslant h_E^{-1}\|[\partial v_h/\partial\nu_E]_E\|_{L^1(E)} \leqslant C_{\mathrm{inv}} h_E^{-1/2}\|[\partial v_h/\partial\nu_E]_E\|_{L^2(E)}.$$

The sum of these terms squared and the Cauchy inequality $(A+B)^2 \leqslant 2A^2 + 2B^2$ for $A, B \in \mathbb{R}$ conclude the proof with $C_7 := 6C_{\mathrm{inv}}^2$. □

Proof of Theorem 5.1. The reliability of the first estimator follows from Lemma 5.5. This and Lemma 5.6 provide

$$\begin{aligned}\|v_h - J_h v_h\|_h^2 &\lesssim \sum_{E\in\mathcal{E}} h_E\|[D^2_{\mathrm{pw}} v_h]_E \tau_E\|^2_{L^2(E)} + j_h(v_h, v_h)\\ &\lesssim \sum_{E\in\mathcal{E}} \left(h_E^{-3}\,\|[v_h]_E\|^2_{L^2(E)} + h_E^{-1}\left\| \left[\frac{\partial v_h}{\partial \nu_E}\right]_E \right\|^2_{L^2(E)} \right).\end{aligned}$$

Since the jumps $[J_h v_h]_E$ and $[\partial J_h v_h/\partial\nu_E]_E$ vanish for a conforming function $J_h v_h \in V$ on any edge $E \in \mathcal{E}$ and $|||v_h - J_h v_h|||_{\mathrm{pw}} \geqslant 0$, the last term is bounded by

$$|||v_h - J_h v_h|||^2_{\mathrm{pw}} + \sum_{E\in\mathcal{E}} \left(h_E^{-3}\,\|[v_h - J_h v_h]_E\|^2_{L^2(E)} + h_E^{-1}\left\| \left[\frac{\partial (v_h - J_h v_h)}{\partial \nu_E}\right]_E \right\|^2_{L^2(E)} \right) \approx \|v_h - J_h v_h\|_h^2$$

with the equivalence of norms in $V + P_2(\mathcal{T})$ from [15, Thm. 4.1] in the last step. This proves the equivalence of both estimators to $\|v_h - J_h v_h\|_h \leqslant C_J \min_{v\in V}\|v - v_h\|_h$ by the quasi-optimality (2.4) of J_h. The trivial estimate $\min_{v\in V}\|v - v_h\|_h \leqslant \|v_h - J_h v_h\|_h$ concludes the proof. □

6 Unified a posteriori error control

This section reconsiders the biharmonic equation (3.6) with weak solution $u \in V := H^2_0(\Omega)$ and the discrete solution $u_h \in V_h$ of the Morley, dG, C^0IP, and WOPSIP schemes defined in Subsections 6.2–6.6. The presentation unifies the a posteriori error analysis of the well-known discretization schemes with orignal and modified right-hand side.

6.1 Discretisation of the biharmonic equation

Recall that $\widehat{V} \equiv H^2(\mathcal{T})$ is a Hilbert space with scalar product $a_{\mathrm{pw}} + j_h$. Recall the discrete spaces $V_{\mathrm{nc}} := \mathrm{M}(\mathcal{T})$ and V_h from Section 4. The weak solution $u \in V := H^2_0(\Omega)$ to the biharmonic equation $\Delta^2 u = F \in V^*$ solves (1.1) with the energy scalar product $a := a_{\mathrm{pw}}|_{V\times V}$ on V and $a_{\mathrm{pw}} : \widehat{V} \times \widehat{V} \to \mathbb{R}$ given in Subsection 4.2.

Recall $J_h := JI_{\mathrm{M}} \in L(\widehat{V}; V)$ from Section 4. Each method defines its particular discrete bilinear form $a_h : (V_h + \mathrm{M}(\mathcal{T})) \times (V_h + \mathrm{M}(\mathcal{T})) \to \mathbb{R}$ in the subsequent subsections. The discrete solution $u_h \in V_h$ solves

$$a_h(u_h, v_h) = (f, Qv_h)_{L^2(\Omega)} \qquad \forall\, v_h \in V_h \tag{6.1}$$

with $f \in L^2(\Omega)$ and $Q \in \{\mathrm{id}, J_h\}$ in this section. The discrete problem (6.1) is a rewriting of (2.3) for $\widehat{F} := F \equiv f \in L^2(\Omega)$ without smoother $Q := \mathrm{id}$ or with the quasi-optimal (by Lemma 2.2) smoother $Q := J_h$. Section 7 discusses more general right-hand sides $F \in V^*$ with a natural extension $\widehat{F} \in H^2(\mathcal{T})^*$. The key assumption (H) from [24] holds for the Morley, dG, and C^0IP discretisations. Hence, the a priori estimate from Theorem 2.2 holds for these methods and leads to the quasi-best approximation property

$$\|u - u_h\|_h \lesssim \|\!|u - I_M u|\!\|_{\mathrm{pw}} = \min_{v_2 \in P_2(\mathcal{T})} \|\!|u - v_2|\!\|_{\mathrm{pw}}. \tag{6.2}$$

In particular, this shows equivalence of these methods from an a priori point of view.

6.2 Morley FEM

The Morley FEM for the biharmonic equation (6.1) comes with $a_h := a_{\mathrm{pw}}$. The subsequent result recovers the equivalent a posteriori estimates from [36, Thm. 2.2] and [2, Eq. (3.2)].

Theorem 6.1 (a posteriori estimate). *The discrete Morley solution $u_h \in V_h$ to* (6.1) *and the exact solution $u \in V$ to* (1.1) *with source $f \in L^2(\Omega)$ satisfy*

$$\begin{aligned}\|\!|u - u_h|\!\|_{\mathrm{pw}}^2 + \mathrm{osc}_2^2(f) &\approx \|h_{\mathcal{T}}^2 f\|^2 + \sum_{E \in \mathcal{E}} h_E \|[D_{\mathrm{pw}}^2 u_h]_E \tau_E\|_{L^2(E)}^2 \\ &\approx \|h_{\mathcal{T}}^2 f\|^2 + \sum_{E \in \mathcal{E}} \left(h_E^{-3} \|[u_h]_E\|_{L^2(E)}^2 + h_E^{-1} \left\| \left[\frac{\partial u_h}{\partial \nu_E} \right]_E \right\|_{L^2(E)}^2 \right).\end{aligned}$$

The equivalence constants exclusively depend on the shape-regularity of $\mathcal{T}$.

Proof. Set $w := e - J_h e_h \in V$, $e_h := I_M e \in \mathrm{M}(\mathcal{T})$ and $\widehat{w} := (1 - J_h) e_h \in V + \mathrm{M}(\mathcal{T})$ for $e := u - J_h u_h$ and recall $I_h = \mathrm{id}$ from Subsection 4.4. Since $a_{\mathrm{pw}}(u_h, w) = 0$ from Lemma 5.3, the definition of the residual and $F(w) \lesssim \|h_{\mathcal{T}}^2 f\| \|\!|e|\!\|$ from Corollary 5.1 show

$$\mathrm{Res}(w) := F(w) - a_{\mathrm{pw}}(u_h, w) = F(w) \lesssim \|h_{\mathcal{T}}^2 f\| \|\!|e|\!\|.$$

Since $(f, J_h e_h - Q e_h)_{L^2(\Omega)} = 0$ vanishes for $Q = J_h$, Corollary 5.1 provides

$$\int_\Omega f(J_h e_h - Q e_h)\,\mathrm{d}x \lesssim \|h_{\mathcal{T}}^2 f\| \|\!|e|\!\| \tag{6.3}$$

for $Q = \mathrm{id}$ and $Q = J_h$. The two previously displayed estimates and (3.15) verify

$$\mathrm{Res}(w) + \int_\Omega f(J_h e_h - Q e_h)\,\mathrm{d}x \lesssim \|h_{\mathcal{T}}^2 f\| (\|\!|u - u_h|\!\|_{\mathrm{pw}} + \|u_h - J_h u_h\|_h).$$

The stability of the L^2 projection shows $\mathrm{osc}_2(f) \leqslant \|h_{\mathcal{T}}^2 f\|$. Hence, Theorem 3.2 plus a weighted Young inequality result in

$$\|\!|u - u_h|\!\|_{\mathrm{pw}}^2 + \mathrm{osc}_2^2(f) \lesssim \|h_{\mathcal{T}}^2 f\|^2 + \|u_h - J_h u_h\|_h^2. \tag{6.4}$$

Since $j_h(u_h, u_h) = 0$ for $u_h \in \mathrm{M}(\mathcal{T})$, Theorem 5.1 bounds $\|u_h - J_h u_h\|_h^2$ in (6.4) by either of the jump terms. This proves the reliability for both estimators. The efficiency of $\|h_{\mathcal{T}}^2 f\|$ follows from Lemma 5.2 while Theorem 5.1 verifies the efficiency for all jump terms. □

6.3 Discontinuous Galerkin 1

Recall the definition of the jump $[\cdot]_E$ and average $\langle\cdot\rangle_E$ (applied componentwise to matrix-valued functions) along an edge $E \in \mathcal{E}$ from Subsection 3.2 and 4.3.

The bilinear form

$$a_h(\cdot,\cdot) = a_{\mathrm{pw}}(\cdot,\cdot) + b_h(\cdot,\cdot) + c_{\mathrm{dG}}(\cdot,\cdot) \tag{6.5}$$

for the discontinuous Galerkin method (dG) [1, 32] depends on $-1 \leqslant \Theta \leqslant 1$ and parameters $\sigma_1, \sigma_2 > 0$. For every $v_2, w_2 \in P_2(\mathcal{T}) \supset V_{\mathrm{nc}} + V_h$,

$$b_h(v_2, w_2) := -\Theta\mathcal{J}(v_2, w_2) - \mathcal{J}(w_2, v_2) \tag{6.6a}$$

$$\mathcal{J}(v_2, w_2) := \sum_{E\in\mathcal{E}} \int_E [\nabla_{\mathrm{pw}} v_2]_E \cdot \left\langle D^2_{\mathrm{pw}} w_2 \right\rangle_E \nu_E \,\mathrm{d}s \tag{6.6b}$$

$$c_{\mathrm{dG}}(v_2, w_2) := \sum_{E\in\mathcal{E}} \left(\frac{\sigma_1}{h_E^3} \int_E [v_2]_E\, [w_2]_E \,\mathrm{d}s + \frac{\sigma_2}{h_E} \int_E \left[\frac{\partial v_2}{\partial \nu_E}\right]_E \left[\frac{\partial w_2}{\partial \nu_E}\right]_E \mathrm{d}s \right). \tag{6.6c}$$

This is the symmetric (resp. non symmetric) interior penalty Galerkin formulation for $\Theta = 1$ (resp. $\Theta = -1$). An appropriate choice [32, 40] of the parameters σ_1, σ_2 guarantees V_h-ellipticity (2.2). Throughout this paper, (2.2) is assumed for $\sigma_1 = \sigma_2 \approx 1$. The following theorem recovers the known a posteriori error estimator from [21] for the linear part.

Theorem 6.2 (a posteriori estimate). *The discrete dG solution $u_h \in V_h$ to* (6.1) *with a_h from* (6.5) *and the exact solution $u \in V$ to* (1.1) *with $f \in L^2(\Omega)$ satisfy*

$$\begin{aligned}\|u-u_h\|_h^2 + \mathrm{osc}_2^2(f) &\approx \|h_{\mathcal{T}}^2 f\|^2 + \sum_{E\in\mathcal{E}} h_E \|[D^2_{\mathrm{pw}} u_h]_E \tau_E\|^2_{L^2(E)} + j_h(u_h, u_h)\\ &\approx \|h_{\mathcal{T}}^2 f\|^2 + \sum_{E\in\mathcal{E}} \left(h_E^{-3}\, \|[u_h]_E\|^2_{L^2(E)} + h_E^{-1} \left\| \left[\frac{\partial u_h}{\partial \nu_E}\right]_E \right\|^2_{L^2(E)} \right).\end{aligned}$$

The equivalence constants exclusively depend on the shape-regularity of $\mathcal{T}$.

Proof. Recall $I_h = \mathrm{id}$ from Subsection 4.4 and, thus, the proof of the reliability and efficiency follows the proof of Theorem 6.1 verbatim except for $j_h(u_h, u_h) \neq 0$ in general. The additional term $j_h(u_h, u_h)$ from the reliability estimate of $\|u_h - J_h u_h\|_h$ in Lemma 5.5 enters the right-hand side of the first estimator. Since $\|u-u_h\|_h^2$ bounds the efficient jump terms $j_h(u_h, u_h) = j_h(u-u_h, u-u_h)$ by definition in (2.1), this concludes the proof. □

Corollary 6.1. The discrete dG solution $u_h \in V_h$ to (6.1) with a_h from (6.5) and the exact solution $u \in V$ to (1.1) and $f \in L^2(\Omega)$ satisfy

$$|||u-u_h|||^2_{\mathrm{pw}} + c_{\mathrm{dG}}(u_h, u_h) + \mathrm{osc}_2^2(f) \approx \|h_{\mathcal{T}}^2 f\|^2 + c_{\mathrm{dG}}(u_h, u_h).$$

Proof. Since $\sigma_1 = \sigma_2 \approx 1$, the jump contributions in the second estimator in Theorem 6.2 are equivalent to $c_{\mathrm{dG}}(u_h, u_h)$. Because $c_{\mathrm{dG}}(v, \cdot) = 0$ vanishes for any $v \in V$, the statement follows with the equivalence $\|u-u_h\|_h^2 \approx |||u-u_h|||^2_{\mathrm{pw}} + c_{\mathrm{dG}}(u_h, u_h)$ from [15, Thm. 4.1]. □

6.4 Discontinuous Galerkin 2

The identity $a(v, w) = (\Delta v, \Delta w)_{L^2(\Omega)}$ for $v, w \in V$ motivates the alternative discontinuous Galerkin method from [34, 40] with discrete bilinear form

$$a_h = (\Delta_{\mathrm{pw}}\cdot, \Delta_{\mathrm{pw}}\cdot)_{L^2(\Omega)} + b_h + c_{\mathrm{dG}}. \tag{6.7}$$

The semi-scalar product c_{dG} is (6.6c) and b_h reads, for any $v_2, w_2 \in P_2(\mathcal{T}) \supset V_{\mathrm{nc}} + V_h$,

$$b_h(v_2, w_2) := -\Theta\mathcal{J}(v_2, w_2) - \mathcal{J}(w_2, v_2) \tag{6.8a}$$

$$\mathcal{J}(v_2, w_2) := \sum_{E\in\mathcal{E}} \int_E \left[\frac{\partial v_2}{\partial \nu_E}\right]_E \langle \Delta_{\mathrm{pw}} w_2 \rangle_E \,\mathrm{d}s \tag{6.8b}$$

for $-1 \leqslant \Theta \leqslant 1$. Appropriate parameters σ_1, σ_2 in c_{dG} guarantee V_h-ellipticity (2.2) of a_h [40]. The bilinear form (6.7) allows for (H).

Lemma 6.1 (quasi-best approximation). *The discontinuous Galerkin method with a_h from* (6.7) *satisfies* (H) *and the quasi-best approximation property* (6.2).

Proof. Given $v_h, w_h \in V_h$, abbreviate $v := J_h v_h, w := J_h w_h \in V$ and $v_{\mathrm{M}} := I_{\mathrm{M}} v_h, w_{\mathrm{M}} := I_{\mathrm{M}} w_h \in \mathrm{M}(\mathcal{T})$. Algebraic manipulations as in [24, Eq. (6.15)] reveal

$$\begin{aligned} a_h(v_h, w_h) - a(v, w) &= (\Delta_{\mathrm{pw}}(v_h - v_{\mathrm{M}}), \Delta_{\mathrm{pw}} w_h)_{L^2(\Omega)} + b_h(v_h - v_{\mathrm{M}}, w_h) \\ &\quad + (\Delta_{\mathrm{pw}} v_{\mathrm{M}}, \Delta_{\mathrm{pw}}(w_h - w_{\mathrm{M}}))_{L^2(\Omega)} + b_h(v_{\mathrm{M}}, w_h - w_{\mathrm{M}}) \\ &\quad + c_{\mathrm{dG}}(v_h, w_h) + (\Delta_{\mathrm{pw}} v_{\mathrm{M}}, \Delta_{\mathrm{pw}} w_{\mathrm{M}})_{L^2(\Omega)} - a(v, w). \end{aligned} \tag{6.9}$$

Cauchy inequalities, $\|\Delta_{\mathrm{pw}} \cdot\| \leqslant \sqrt{2} ||| \cdot |||_{\mathrm{pw}}$, the boundedness of b_h, and (4.5) provide

$$(\Delta_{\mathrm{pw}}(v_h - v_{\mathrm{M}}), \Delta_{\mathrm{pw}} w_h)_{L^2(\Omega)} + b_h(v_h - v_{\mathrm{M}}, w_h) \leqslant (2 + \|b_h\|)\Lambda_{\mathrm{M}} \|v - v_h\|_h \|w_h\|_h. \tag{6.10}$$

Recall the definition of the jump $[\cdot]_E$ and average $\langle\cdot\rangle_E$ along an edge $E \in \mathcal{E}$ from Subsection 3.2 and 4.3 and the product rule for jump terms $[ab]_E = \langle a\rangle_E [b]_E + [a]_E \langle b\rangle_E$ for any $a, b \in H^1(\mathcal{T})$. This and an integration by parts verify

$$\begin{aligned} &(\Delta_{\mathrm{pw}} v_{\mathrm{M}}, \Delta_{\mathrm{pw}}(w_h - w_{\mathrm{M}}))_{L^2(\Omega)} + b_h(v_{\mathrm{M}}, w_h - w_{\mathrm{M}}) \\ &\quad = \sum_{E\in\mathcal{E}} \int_E \left([\Delta_{\mathrm{pw}} v_{\mathrm{M}}]_E \left\langle \frac{\partial (w_h - w_{\mathrm{M}})}{\partial \nu_E} \right\rangle_E - \Theta \left[\frac{\partial v_{\mathrm{M}}}{\partial \nu_E} \right]_E \langle \Delta_{\mathrm{pw}}(w_h - w_{\mathrm{M}})\rangle_E \right) \mathrm{d}s = 0 \end{aligned} \tag{6.11}$$

with $\int_E \langle \partial(w_2 - I_{\mathrm{M}} w_2)/\partial\nu_E\rangle_E \,\mathrm{d}s = \int_E [\partial v_{\mathrm{M}}/\partial \nu_E]_E \,\mathrm{d}s = 0$ for any edge $E \in \mathcal{E}$ from the definition of I_{M} in the last step. Since the Morley interpolation I_{M} exactly interpolates the integral mean over an edge $E \in \mathcal{E}$ of the normal derivative of $w \equiv J w_{\mathrm{M}} \in V$ (from $I_{\mathrm{M}} J = 1$), an integration by parts for any $p_2 \in P_2(\mathcal{T})$ shows the orthogonality

$$(\Delta_{\mathrm{pw}} p_2, \Delta_{\mathrm{pw}}(w - w_{\mathrm{M}}))_{L^2(\Omega)} = \sum_{E\in\mathcal{E}} \langle \Delta_{\mathrm{pw}} p_2\rangle_E \int_E \partial((1 - I_{\mathrm{M}}) J w_{\mathrm{M}})/\partial\nu_E \,\mathrm{d}s = 0.$$

Since $a(v, w) = (\Delta v, \Delta w)_{L^2(\Omega)}$, this, a Cauchy inequality, and $\|\Delta_{\mathrm{pw}} \cdot\| \leqslant \sqrt{2} ||| \cdot |||_{\mathrm{pw}}$ imply

$$\begin{aligned} (\Delta_{\mathrm{pw}} v_{\mathrm{M}}, \Delta_{\mathrm{pw}} w_{\mathrm{M}})_{L^2(\Omega)} - a(v, w) &= (\Delta_{\mathrm{pw}}(1 - J) v_{\mathrm{M}}, \Delta_{\mathrm{pw}} J w_{\mathrm{M}})_{L^2(\Omega)} \\ &\leqslant 2(1 + \Lambda_{\mathrm{M}}) \|J\|_h \|I_{\mathrm{M}}\|_h \|v - v_h\|_h \|w_h\|_h. \end{aligned}$$

This, the combination of (6.9) with (6.10)–(6.11), and $c_{\mathrm{dG}}(v_h, w_h) \leqslant \Lambda_{\mathrm{c}} \|v - v_h\|_h \|w_h\|_h$ for $\Lambda_{\mathrm{c}} \lesssim 1$ from [24, Sect. 7] conclude the proof of (H). The quasi-best approximation property (6.2) is a consequence of (H) and Theorem 2.2. □

Since the dG formulations from Subsections 6.3–6.4 allow for (H) and utilize the same space $V_h = P_2(\mathcal{T})$, the a posteriori results from Subsection 6.3 follow verbatim for the alternative dG formulation in this subsection.

Theorem 6.3 (a posteriori estimate). *The discrete dG solution $u_h \in V_h$ to* (6.1) *with a_h from* (6.7) *and the exact solution $u \in V$ to* (1.1) *with $f \in L^2(\Omega)$ satisfy*

$$\begin{aligned} \|u - u_h\|_h^2 + \mathrm{osc}_2^2(f) &\approx \|h_{\mathcal{T}}^2 f\|^2 + \sum_{E\in\mathcal{E}} h_E \|[D^2_{\mathrm{pw}} u_h]_E \tau_E\|^2_{L^2(E)} + j_h(u_h, u_h) \\ &\approx \|h_{\mathcal{T}}^2 f\|^2 + \sum_{E\in\mathcal{E}} \left(h_E^{-3} \|[u_h]_E\|^2_{L^2(E)} + h_E^{-1} \left\| \left[\frac{\partial u_h}{\partial \nu_E} \right]_E \right\|^2_{L^2(E)} \right). \end{aligned}$$

The equivalence constants exclusively depend on the shape-regularity of $\mathcal{T}$.

The following corollary provides an improvement to the a posteriori error estimator that comes without the jump term $\|[\Delta_{\mathrm{pw}} u_h]_E\|_{L^2(E)}$ over an interior edge $E \in \mathcal{E}(\Omega)$ and so refines the a posteriori result in [34].

Corollary 6.2 (see [34]). The discrete dG solution $u_h \in V_h$ to (6.1) with a_h from (6.7) and the exact solution $u \in V$ to (1.1) with $f \in L^2(\Omega)$ satisfy

$$|||u - u_h|||_{\mathrm{pw}}^2 + c_{\mathrm{dG}}(u_h, u_h) + \mathrm{osc}_2^2(f) \approx \|h_{\mathcal{T}}^2 f\|^2 + c_{\mathrm{dG}}(u_h, u_h).$$

6.5 C^0 interior penalty (C^0IP)

The bilinear form $a_h = a_{\mathrm{pw}} + b_h + c_{\mathrm{IP}}$ for C^0IP [7, 21] utilizes b_h from (6.6a) and depends on the parameter $\sigma_{\mathrm{IP}} > 0$ in

$$c_{\mathrm{IP}}(v_2, w_2) := \sum_{E\in\mathcal{E}} \frac{\sigma_{\mathrm{IP}}}{h_E} \int_E \left[\frac{\partial v_2}{\partial \nu_E}\right]_E \left[\frac{\partial w_2}{\partial \nu_E}\right]_E \mathrm{d}s \tag{6.12a}$$

for $v_2, w_2 \in V_h := P_2(\mathcal{T})$. The scheme is a modification of the dG method in Section 6.3 with trial and test functions restricted to the continuous piecewise polynomials $V_h := S_0^2(\mathcal{T})$. For $\sigma_{\mathrm{IP}} \approx 1$ sufficiently large but bounded, the bilinear form is coercive. The abstract framework applies the transfer operator $I_h = I_{\mathrm{C}} \in L(V_{\mathrm{nc}}; V_h)$ from Subsection 4.4.

Theorem 6.4 (a posteriori estimate). *The discrete solution $u_h \in V_h$ of the C^0IP method to* (6.1) *and the exact solution $u \in V$ to* (1.1) *with $f \in L^2(\Omega)$ satisfy*

$$\begin{aligned}
&\|u - u_h\|_h^2 + \mathrm{osc}_2^2(f)\\
&\quad\approx \|h_{\mathcal{T}}^2 f\|^2 + \sum_{E\in\mathcal{E}} h_E \|[D^2_{\mathrm{pw}} u_h]_E \tau_E\|^2_{L^2(E)} + \sum_{E\in\mathcal{E}(\Omega)} h_E \|[\partial^2_{\nu\nu} u_h]_E\|^2_{L^2(E)} + j_h(u_h, u_h)\\
&\quad\approx \|h_{\mathcal{T}}^2 f\|^2 + \sum_{E\in\mathcal{E}} h_E^{-1} \left\|\left[\frac{\partial u_h}{\partial \nu_E}\right]_E\right\|^2_{L^2(E)} + \sum_{E\in\mathcal{E}(\Omega)} h_E \|[\partial^2_{\nu\nu} u_h]_E\|^2_{L^2(E)}.
\end{aligned}$$

The equivalence constants exclusively depend on the shape-regularity of $\mathcal{T}$.

Proof. Set $w := (1 - J_h I_{\mathrm{C}} I_{\mathrm{M}})e \in V$ and $\widehat{w} := (Q - J_h) I_{\mathrm{C}} I_{\mathrm{M}} e \in V + S_0^2(\mathcal{T})$ for $e := u - J_h u_h \in V$. The definition of the residual, Corollary 5.1, and Lemma 5.3 with $I_h = I_{\mathrm{C}}$ show

$$\mathrm{Res}(w) := F(w) - a_{\mathrm{pw}}(u_h, w) \lesssim \left(\|h_{\mathcal{T}}^2 f\| + \sqrt{\sum_{E\in\mathcal{E}(\Omega)} h_E \|[\partial^2_{\nu\nu} u_h]_E\|^2_{L^2(E)}}\right) |||e|||.$$

Theorem 3.2 and the definition of the residual result in $\|u - u_h\|_h^2 \lesssim \|u_h - J_h u_h\|^2 + \mathrm{Res}(w) - \widehat{F}(\widehat{w})$. Since the stability of the L^2 projection shows $\mathrm{osc}_2(f) \leq \|h_{\mathcal{T}}^2 f\|$, this, the bound $\widehat{F}(\widehat{w}) \lesssim \|h_{\mathcal{T}}^2 f\| |||e|||$ from (6.3), and a weighted Young inequality reveal

$$\|u - u_h\|_h^2 + \mathrm{osc}_2^2(f) \lesssim \|h_{\mathcal{T}}^2 f\|^2 + \|u_h - J_h u_h\|_h^2 + \sum_{E\in\mathcal{E}(\Omega)} h_E \|[\partial^2_{\nu\nu} u_h]_E\|^2_{L^2(E)}.$$

Theorem 5.1 bounds $\|u_h - J_h u_h\|^2$ either in terms of $\sum_{E\in\mathcal{E}} h_E \|[D^2_{\mathrm{pw}} u_h]_E \tau_E\|^2_{L^2(E)}$ plus $j_h(u_h, u_h)$ or in terms of $\sum_{E\in\mathcal{E}} h_E^{-1} \|[\partial u_h / \partial \nu_E]_E\|^2_{L^2(E)}$ (because $[u_h]_E \equiv 0$ for $u_h \in S_0^2(\mathcal{T})$). This concludes the proof of the reliability. Lemma 5.4 provides the efficiency of the normal-normal jumps. The efficiency of the remaining terms follows verbatim as in the proof of Theorem 6.2. □

The following corollary recovers the a posteriori result from [4, Sects. 3 and 4].

Corollary 6.3 (see [4]). The discrete C^0IP solution $u_h \in V_h$ to (6.1) and the exact solution $u \in V$ to (1.1) with $f \in L^2(\Omega)$ satisfy

$$|||u - u_h|||_{\mathrm{pw}}^2 + c_{\mathrm{IP}}(u_h, u_h) + \mathrm{osc}_2^2(f) \approx \|h_{\mathcal{T}}^2 f\|^2 + c_{\mathrm{IP}}(u_h, u_h) + \sum_{E\in\mathcal{E}(\Omega)} h_E \|[\partial^2_{\nu\nu} u_h]_E\|^2_{L^2(E)}.$$

Proof. Since $[v_h]_E = 0$ for any $v_h \in S_0^2(\mathcal{T})$, $c_{\mathrm{IP}} = c_{\mathrm{dG}}$ coincide in $S_0^2(\mathcal{T}) \times S_0^2(\mathcal{T})$ and the proof follows verbatim that of Corollary 6.1; further details are omitted. □

6.6 WOPSIP

The weakly over-penalized symmetric interior penalty (WOPSIP) scheme [5] is a penalty method with the stabilisation term

$$c_{\mathrm{P}}(v,w) := \sum_{E\in\mathcal{E}} h_E^{-2}\left(\sum_{z\in\mathcal{V}(E)} \frac{[v]_E(z)}{h_E}\frac{[w]_E(z)}{h_E} + \fint_E \left[\frac{\partial v}{\partial \nu_E}\right]_E \mathrm{d}s \fint_E \left[\frac{\partial w}{\partial \nu_E}\right]_E \mathrm{d}s\right) \tag{6.13}$$

for piecewise smooth functions $v, w \in H^2(\mathcal{T})$. The difference of c_{P} in (6.13) to j_h from (4.2) is the over-pernalisation by an additional negative power of the mesh size h_E. This and $h_{\max} := \max_{T\in\mathcal{T}} h_T$ establish

$$j_h(v,v) \leqslant h_{\max}^2 c_{\mathrm{P}}(v,v) \qquad \forall\, v \in \widehat{V} := H^2(\mathcal{T}). \tag{6.14}$$

Hence $\|\cdot\|_{\mathrm{P}} := \big(|\!|\!|\cdot|\!|\!|_{\mathrm{pw}}^2 + c_{\mathrm{P}}(\cdot,\cdot)\big)^{1/2}$ is a norm in $\widehat{V}$ stronger than $\|\cdot\|_h$. The WOPSIP method computes the discrete solution $u_h \in V_h := P_2(\mathcal{T})$ to (6.1) with the bilinear form $a_h := a_{\mathrm{pw}} + c_{\mathrm{P}}$ and fits into the abstract setting with $V_{\mathrm{nc}} := \mathrm{M}(\mathcal{T})$.

The main difference to the methods under consideration above is the missing quasi-best approximation property due to the penalisation. Instead of this, the following a priori estimate for the energy norm

$$|\!|\!|u - u_h|\!|\!|_{\mathrm{pw}}^2 + c_{\mathrm{P}}(u_h, u_h) \lesssim |\!|\!|u - I_{\mathrm{M}} u|\!|\!|_{\mathrm{pw}}^2 + |\!|\!|h_{\mathcal{T}} I_{\mathrm{M}} u|\!|\!|_{\mathrm{pw}}^2$$

holds with the extra term $|\!|\!|h_{\mathcal{T}} I_{\mathrm{M}} u|\!|\!|_{\mathrm{pw}}^2$ (see [24, Thm. 9.1]). This suggests that (H) does not hold, but the methodology of the a posteriori analysis of Subsection 3.3 is still applicable. Indeed, the key assumption only enters in the error bound from Theorem 3.2 and a careful analysis with $I_h = \mathrm{id}$ leads to (3.19). This allows the application of the developed tool chain and leads to a new a posteriori estimate *without* the WOPSIP stabilisation term (6.13) but still with the weaker stabilization j_h.

Theorem 6.5 (a posteriori estimate). *The WOPSIP solution $u_h \in P_2(\mathcal{T})$ to* (6.1) *and the exact solution $u \in V$ to* (1.1) *with $f \in L^2(\Omega)$ satisfy*

$$\begin{aligned}\|u - u_h\|_h^2 + \mathrm{osc}_2^2(f) &\approx \|h_{\mathcal{T}}^2 f\|^2 + \sum_{E\in\mathcal{E}} h_E \|[D_{\mathrm{pw}}^2 u_h]_E \tau_E\|_{L^2(E)}^2 + j_h(u_h,u_h)\\ &\approx \|h_{\mathcal{T}}^2 f\|^2 + \sum_{E\in\mathcal{E}} \left(h_E^{-3}\, \|[u_h]_E\|_{L^2(E)}^2 + h_E^{-1} \left\|\left[\frac{\partial u_h}{\partial \nu_E}\right]_E\right\|_{L^2(E)}^2\right).\end{aligned}$$

The equivalence constants exclusively depend on the shape-regularity of $\mathcal{T}$.

Proof. Let $e := u - J_h u_h \in V$ and recall $J_h := J I_{\mathrm{M}}$ as well as $I_{\mathrm{M}} J = \mathrm{id}$ on V_{nc}. The key assumption (H) enters the proof of Theorem 3.2 with (3.17). This proof exploits that the transfer operator $I_h := \mathrm{id} : V_{\mathrm{nc}} \to V_h$ is the identity and deduces (3.17) directly (and so circumvents (H)). Indeed, since $c_{\mathrm{P}}(\cdot, v_{\mathrm{nc}}) = 0$ for any $v \in V_{\mathrm{nc}}$ and $a_{\mathrm{pw}}(u_h, e_h - J_h e_h) = 0$ by the orthogonality (4.3) for $e_h := I_h I_{\mathrm{M}} e = I_{\mathrm{M}} J_h e_h \in V_{\mathrm{nc}}$,

$$\begin{aligned} a_h(u_h, e_h) &= a_{\mathrm{pw}}(u_h, e_h) = a_{\mathrm{pw}}(u_h, J_h e_h)\\ a_h(u_h, e_h) - a(J_h u_h, J_h e_h) &= a_{\mathrm{pw}}(u_h - J_h u_h, J_h e_h) \leqslant \|J_h\| |\!|\!|u_h - J_h u_h|\!|\!|_{\mathrm{pw}} |\!|\!|e_h|\!|\!|_{\mathrm{pw}}\end{aligned}$$

follow with a Cauchy inequality in the last step. Hence (3.17) even holds with the weaker norm $|\!|\!|\cdot|\!|\!|_{\mathrm{pw}} \leqslant \|\cdot\|_h$. The remaining parts of the proof for Theorem 3.2 apply analogously and verify

$$\|u - u_h\|_h^2 \lesssim \|u_h - J_h u_h\|_h^2 + \mathrm{Res}(w) - \int_\Omega f \widehat{w}\, \mathrm{d}x = \|u_h - J_h u_h\|_h^2 + \int_\Omega f(w - \widehat{w})\, \mathrm{d}x$$

with $a_{\mathrm{pw}}(u_h, w) = 0$ from Lemma 5.3 for $w := e - J_h I_h I_{\mathrm{M}} e$ and $\widehat{w} := Q e_h - J_h e_h \in V + V_h$. The remaining arguments follow the proofs of Theorem 6.2 and Theorem 6.1 verbatim. □

The inclusion of the stabilisation term c_{P} on both sides of the error estimate in Theorem 6.5 recovers the a posteriori estimate from [5, Sect. 6].

Corollary 6.4 (see [5]). The discrete WOPSIP solution $u_h \in V_h$ to (6.1) and the exact solution $u \in V$ to (1.1) with $f \in L^2(\Omega)$ satisfy

$$\begin{aligned}|||u - u_h|||_{\mathrm{pw}}^2 &+ c_{\mathrm{P}}(u_h, u_h) + \mathrm{osc}_2^2(f)\\ &\approx \|h_{\mathcal{T}}^2 f\|^2 + \sum_{E\in\mathcal{E}} \left(h_E^{-3} \|[u_h]_E\|_{L^2(E)}^2 + h_E^{-1} \left\| \left[\frac{\partial u_h}{\partial \nu_E} \right]_E \right\|_{L^2(E)}^2 \right) + c_{\mathrm{P}}(u_h, u_h).\end{aligned}$$

The equivalence constants exclusively depend on the shape-regularity of $\mathcal{T}$. □

Proof. This follows from Corollary 6.1 with $c_{\mathrm{P}}(\cdot, v) = 0$ for all $v \in V$ and $j_h(u_h, u_h) \lesssim c_{\mathrm{P}}(u_h, u_h)$ from (6.14); further details are omitted. □

7 More general sources

This section considers a class of rather general right-hand sides $F \in V^*$ and introduces an estimator for the residual that is reliable and efficient up to a data approximation error.

7.1 A general class of source terms

Every functional in $F \in V^* \equiv H^{-2}(\Omega)$ has (non-unique) representations by volume loads $f_\alpha \in L^2(\Omega)$ for all 6 multi-indices $\alpha = (\alpha_1, \alpha_2) \in \mathbb{N}_0^2$ of order $|\alpha| := \alpha_1 + \alpha_2 \leqslant 2$, written $(f_\alpha)_{|\alpha|\leqslant 2} \in L^2(\Omega)^6$, with

$$F(\varphi) \equiv \langle F, \varphi\rangle = \sum_{|\alpha|\leqslant 2} (f_\alpha, \partial^\alpha \varphi)_{L^2(\Omega)} \qquad \forall\, \varphi \in H_0^2(\Omega). \tag{7.1}$$

Theorem 7.1 (characterization). *Given any $F \in H^{-2}(\Omega)$ there exist $(f_\alpha)_{|\alpha|\leqslant 2} \in L^2(\Omega)^6$ such that* (7.1) *holds. The norm of F in $H^{-2}(\Omega)$ (the dual of $H_0^2(\Omega)$ endowed with the full Sobolev norm of $H^2(\Omega)$) is the minimum*

$$\|F\|_{H^{-2}(\Omega)} = \min\left\{ \sqrt{\textstyle\sum_{|\alpha|\leqslant 2} \|f_\alpha\|_{L^2(\Omega)}^2} \ :\ (f_\alpha)_{|\alpha|\leqslant 2} \in L^2(\Omega)^6 \text{ satisfies } (7.1)\right\}.$$

Proof. This is a natural generalization of the corresponding result for functionals in $H^{-1}(\Omega)$, e.g., [31, Sect. 5.9, Thm. 1]; hence further details are omitted. □

Remark 7.1 (characterization for semi-norm $|||\cdot|||$). The norm representation of Theorem 7.1 is given in the (full) norm $\|\cdot\|_{H^2(\Omega)}$ of $H^2(\Omega)$. A corresponding assertion

$$|||F|||_* := \sup_{v\in V} F(v)/|||v||| = \min_{\sigma\in L^2(\Omega;\mathbb{S})} \{\|\sigma\|_{L^2(\Omega)} \ :\ F = (\sigma, D^2\cdot)_{L^2(\Omega)}\} \tag{7.2}$$

follows from the Riesz representation theorem for the H^2 seminorm $|||\cdot||| \equiv |\cdot|_{H^2(\Omega)}$ as well. The minimizer $\sigma = D^2 u \in L^2(\Omega;\mathbb{S})$ in (7.2) is the Hessian of the weak solution $u \in V$ to (1.1).

A more general source term may include point forces $\delta_z \in V^*$ at finitely many points $z \in A \subset \overline{\Omega}$ and line loads $(g_0, \cdot)_{L^2(\Gamma_0)}, (g_1, \partial_\nu \cdot)_{L^2(\Gamma_1)}$ along the hypersurfaces $\Gamma_0, \Gamma_1 \subset \Omega$ in addition to (7.1). The Dirac delta distribution $\delta_z(f) = f(z)$ evaluates $f \in V \subset C(\overline{\Omega})$ at the atom z and we suppose for simplicity that the mesh is adapted in that $A \subset \mathcal{V}(\Omega)$ consists of interior vertices. Recall the set $\mathcal{T}(z) := \{T \in \mathcal{T} \ :\ z \in T\}$ of neighbouring triangles from Subsection 4.3 and suppose that the mesh resolves $\Gamma_j = \bigcup \mathcal{E}(\Gamma_j)$ with $\mathcal{E}(\Gamma_j) := \{E \in \mathcal{E} \ :\ \mathrm{int}(E) \subset \Gamma_j\}$ for $j = 0, 1$.

This section considers sources $F := \widehat{F}|_V \in V^*$ in terms of an extended source $\widehat{F} \in \widehat{V}^* \equiv H^2(\mathcal{T})^*$, defined, for $\widehat{v} \in H^2(\mathcal{T})$, by

$$\widehat{F}(\widehat{v}) := \sum_{|\alpha|\leqslant 2} (f_\alpha, \partial_{\mathrm{pw}}^\alpha \widehat{v})_{L^2(\Omega)} + \sum_{j=0,1} (g_j, \langle \partial_\nu^j \widehat{v}\rangle_{\Gamma_j})_{L^2(\Gamma_j)} + \sum_{z\in A} \sum_{T\in\mathcal{T}(z)} \beta_{T,z} \widehat{v}|_T(z). \tag{7.3}$$

The given data in (7.3) are Lebesgue functions $(f_\alpha)_{|\alpha|\leqslant 2} \in L^2(\Omega)^6$, line loads $g_j \in L^2(\Gamma_j)$ along the hypersurface $\Gamma_j \subset \bigcup \mathcal{E}(\Omega)$ for $j = 0, 1$, and point forces of intensity $\beta_z = \sum_{T\in\mathcal{T}(z)} \beta_{T,z} \in \mathbb{R}$ at $z \in A \subset \mathcal{V}(\Omega)$.

Remark 7.2 (influence of $\widehat{F}$). Since $\widehat{F} \circ J_h = F \circ J_h$ holds (for the five schemes from Section 6) with the smoother $Q = J_h$, the discrete solution $u_h \in V_h$ to (2.3) depends on F but is independent of its representation (7.3). The classical scheme without smoother $Q = \mathrm{id}$, however, depends on the chosen data for the representation $\widehat{F}$.

Throughout this section, we suppose that we have piecewise smooth approximations $G_j \in L^2(\Gamma_j)$ of g_j for $j = 0, 1$ and $F_\alpha \in H^{|\alpha|}(\mathcal{T})$ of f_α in (7.3) for $|\alpha| \leqslant 2$ to define an approximation $\widehat{F}_{\mathrm{apx}}$ of $\widehat{F}$ with piecewise smooth data. The reason for this approximation is that we shall integrate by parts with piecewise smooth functions to reveal an efficient a posteriori upper error bound in the subsequent subsection.

Definition 7.1 (data approximation error). The approximated source term reads

$$\widehat{F}_{\mathrm{apx}}(\widehat{v}) := \sum_{|\alpha|\leqslant 2} (F_\alpha, \partial^\alpha_{\mathrm{pw}} \widehat{v})_{L^2(\Omega)} + \sum_{j=0,1} \sum_{E\in\mathcal{E}(\Gamma_j)} (G_j, \langle \partial^j_\nu \widehat{v}\rangle_E)_{L^2(E)} \tag{7.4}$$

for all $\widehat{v} \in H^2(\mathcal{T})$. The data approximation error $\mathrm{apx}(F, \mathcal{T}) := \big(\sum_{T\in\mathcal{T}} \mathrm{apx}^2(F, T)\big)^{1/2}$ has, on the triangle $T \in \mathcal{T}$, the contribution

$$\mathrm{apx}^2(F, T) := \sum_{|\alpha|\leqslant 2} \|h_T^{2-|\alpha|}(f_\alpha - F_\alpha)\|^2_{L^2(T)} + \sum_{j=0,1} \sum_{E\in\mathcal{E}(\Gamma_j)\cap\mathcal{E}(T)} \|h_E^{3/2-j}(g_j - G_j)\|^2_{L^2(E)}.$$

The data approximation error generalizes data oscillations. Let $\Pi_{E,k} : L^2(E) \to P_k(E)$ denote the L^2 projection onto $P_k(E)$ on the edge $E \in \mathcal{E}$.

Example 7.1 (data oscillations). The natural candidates for $(F_\alpha)_{|\alpha|\leqslant 2}$ and G_0, G_1 in (7.4) are L^2 projections onto polynomials of degree at most $k \in \mathbb{N}_0$. Then the data approximation error $\mathrm{apx}^2(F, \mathcal{T})$ becomes an oscillation term

$$\mathrm{osc}^2(F, \mathcal{T}) := \sum_{|\alpha|\leqslant 2} \|h_{\mathcal{T}}^{2-|\alpha|}(1 - \Pi_k) f_\alpha\|^2_{L^2(\Omega)} + \sum_{j=0,1} \sum_{E\in\mathcal{E}(\Gamma_j)} \|h_E^{3/2-j}(1 - \Pi_{E,k}) g_j\|^2_{L^2(E)}.$$

Lemma 7.1 (data approximation error). *With the linear operators J, I_{M}, I_h from Table 2,*

$$\max\left\{ |||(F - \widehat{F}_{\mathrm{apx}})(1 - J_h I_h I_{\mathrm{M}})|||_*, |||(\widehat{F} - \widehat{F}_{\mathrm{apx}})(1 - J_h) I_h I_{\mathrm{M}}|||_* \right\} \leqslant C_4\, \mathrm{apx}(F, \mathcal{T}).$$

Proof. Recall that $w := (1 - J_h I_h I_{\mathrm{M}})v$ vanishes at the vertices for all $v \in V$ and for all five schemes under consideration. This shows

$$(F - \widehat{F}_{\mathrm{apx}})(w) = \sum_{|\alpha|\leqslant 2} (f_\alpha - F_\alpha, \partial^\alpha w)_{L^2(\Omega)} + \sum_{j=0,1} (g_j - G_j, \partial^j_\nu w)_{L^2(\Gamma_j)} \leqslant C_4\, \mathrm{apx}(F, \mathcal{T}) |||v|||$$

with a Cauchy inequality and the constant C_4 from Lemma 5.1 in the last step. Analog arguments provide the asserted bound of $|||(\widehat{F} - \widehat{F}_{\mathrm{apx}}) \circ (1 - J_h) I_h I_{\mathrm{M}}|||_*$. □

7.2 Estimator for the residual

The paradigm shift in this paper is that Theorem 3.2 provides an upper error bound with a specific structure of the test function as an element in $(1 - J_h I_h I_{\mathrm{M}})V$ for the residual part. This subsection designs an estimator $\mu(\mathcal{T})$ for the dual norm $|||\mathrm{Res} \circ (1 - J_h I_h I_{\mathrm{M}})|||_*$ of the residual that is reliable and efficient up to the data approximation error $\mathrm{apx}(F, \mathcal{T})$

$$|||\mathrm{Res} \circ (1 - J_h I_h I_{\mathrm{M}})|||_* \lesssim \mu(\mathcal{T}) + \mathrm{apx}(F, \mathcal{T}) \lesssim |||u - u_h|||_{\mathrm{pw}} + \mathrm{apx}(F, \mathcal{T}). \tag{7.5}$$

The residual $\mathrm{Res} := F - a_{\mathrm{pw}}(u_h, \cdot) \in V^*$ includes the discrete solution $u_h \in V_h$ to (2.3) with or without smoother $Q \in \{\mathrm{id}, J_h\}$. The analysis in this section for an upper bound of the dual norm $|||\mathrm{Res} \circ (1 - J_h I_h I_{\mathrm{M}})|||_*$ allows for a general discrete object $u_h \in V_h$; said differently, $u_h \in V_h$ is arbitrary in (7.5).

To define the estimator contributions in $\mu(\mathcal{T})$, abbreviate $F_0 := F_{(0,0)} \in L^2(\Omega)$,

$$F_1 := \begin{pmatrix} F_{(1,0)} \\ F_{(0,1)} \end{pmatrix} \in H^1(\mathcal{T};\mathbb{R}^2), \qquad F_2 := \begin{pmatrix} F_{(2,0)} & \frac{1}{2}F_{(1,1)} \\ \frac{1}{2}F_{(1,1)} & F_{(0,2)} \end{pmatrix} \in H^2(\mathcal{T};\mathbb{S}). \tag{7.6}$$

The extra factor 1/2 in the definition of F_2 allows the simplification $(F_j, D^j_{\mathrm{pw}}\cdot)_{L^2(\Omega)} = \sum_{|\alpha|=j}(F_\alpha, \partial^\alpha_{\mathrm{pw}}\cdot)_{L^2(\Omega)}$ for $j = 0, 1, 2$. Let the divergence

$$\mathrm{div}_{\mathrm{pw}} F_2 := \begin{pmatrix} \mathrm{div}_{\mathrm{pw}}(F_2)_1 \\ \mathrm{div}_{\mathrm{pw}}(F_2)_2 \end{pmatrix} \in H^1(\mathcal{T};\mathbb{R}^2)$$

of the matrix-valued function $F_2 \equiv ((F_2)_1; (F_2)_2) \in H^2(\mathcal{T};\mathbb{S})$ apply row-wise. Recall J, I_{M}, I_h from Section 4 and the special treatment of $I_h = \mathrm{id}$ in Subsection 5.2. Define

$$\begin{aligned}
\mu_1^2(\mathcal{T}) &:= \|h_{\mathcal{T}}^2(F_0 - \mathrm{div}_{\mathrm{pw}}F_1 + \mathrm{div}^2_{\mathrm{pw}}F_2)\|^2\\
\mu_2^2(\mathcal{T}) &:= \sum_{E\in\mathcal{E}(\Omega)} h_E^3\|G_0 + [F_1 - \mathrm{div}_{\mathrm{pw}}F_2 - \partial(F_2\tau_E)/\partial s]_E\cdot\nu_E\|^2_{L^2(E)}\\
\mu_3^2(\mathcal{T}) &:= \sum_{E\in\mathcal{E}(\Omega)} \begin{cases} h_E\|(1-\Pi_{E,0})(G_1 + [F_2\nu_E]_E\cdot\nu_E)\|^2_{L^2(E)} & \text{if } I_h = \mathrm{id}\\ h_E\|G_1 + [(F_2 - D^2_{\mathrm{pw}}u_h)\nu_E]_E\cdot\nu_E\|^2_{L^2(E)} & \text{if } I_h = I_{\mathrm{C}}\end{cases}\\
\mu^2(\mathcal{T}) &:= \mu_1^2(\mathcal{T}) + \mu_2^2(\mathcal{T}) + \mu_3^2(\mathcal{T}).
\end{aligned}$$

Here $G_j \in L^2(\Gamma_j) \subset L^2(\bigcup\mathcal{E})$ is extended by zero to the entire skeleton for $j = 0, 1$.

Proposition 7.1 (reliability). The estimator $\mu(\mathcal{T}) \equiv \mu^2(\mathcal{T})^{1/2}$ of the residual is reliable

$$C_4^{-1}|||\mathrm{Res}\circ(1 - J_hI_hI_{\mathrm{M}})|||_* \leqslant \mu(\mathcal{T}) + \mathrm{apx}(F,\mathcal{T}).$$

Proof. Given any $v \in V$, the function $w := v - J_hI_hI_{\mathrm{M}}v \in V$ vanishes at the vertices $z \in \mathcal{V}$. The split $\nabla w = (\partial w/\partial\nu_E)\nu_E + (\partial w/\partial s)\tau_E$ along an edge $E = \mathrm{conv}\{A, B\} \in \mathcal{E}$ and an integration by parts with $w(A) = w(B) = 0$ verify

$$\begin{aligned}
([F_2\nu_E]_E, \nabla w)_{L^2(E)} &= ([F_2\nu_E]_E, \nu_E\partial w/\partial\nu_E)_{L^2(E)} + ([F_2\nu_E]_E, \tau_E\partial w/\partial s)_{L^2(E)}\\
&= ([F_2\nu_E]_E, \nu_E\partial w/\partial\nu_E)_{L^2(E)} - (\partial[F_2\tau_E]_E/\partial s, \nu_E w)_{L^2(E)}
\end{aligned} \tag{7.7}$$

with $\tau_E\cdot F_2\nu_E = \nu_E\cdot F_2\tau_E$ for all symmetric matrix-valued $F_2 \in H^2(\mathcal{T};\mathbb{S})$ in the last step. An integration by parts and (7.7) lead to

$$\begin{aligned}
\widehat{F}_{\mathrm{apx}}(w) = (F_0 - \mathrm{div}_{\mathrm{pw}}F_1 + \mathrm{div}^2_{\mathrm{pw}}F_2, w)_{L^2(\Omega)} &+ \sum_{E\in\mathcal{E}(\Omega)}(G_1 + [F_2\nu_E]_E\nu_E, \partial w/\partial\nu_E)_{L^2(E)}\\
&+ \sum_{E\in\mathcal{E}(\Omega)}(G_0 + [F_1 - \mathrm{div}_{\mathrm{pw}}F_2 - \partial(F_2\tau_E)/\partial s]_E\cdot\nu_E, w)_{L^2(E)}.
\end{aligned} \tag{7.8}$$

Since $I_{\mathrm{M}}w = 0$ for $I_h = \mathrm{id}$, the integral mean $\Pi_{E,0}(\partial w/\partial\nu_E) \equiv 0$ vanishes along any edge $E \in \mathcal{E}$. Hence, $(p_0, \partial w/\partial\nu_E)_{L^2(E)} = 0$ is zero for any constant $p_0 \in P_0(E)$. An integration by parts with (5.5) for the piecewise constant Hessian $D^2_{\mathrm{pw}}u_h \in P_0(\mathcal{T};\mathbb{S})$, $q_0 := \Pi_{E,0}(G_2 + [F_2\nu_E]_E\cdot\nu_E) \in P_0(E)$, and the split of ∇w as in (7.7) result in

$$a_{\mathrm{pw}}(u_h, w) = (D^2_{\mathrm{pw}}u_h, D^2w)_{L^2(\Omega)} = \sum_{E\in\mathcal{E}(\Omega)}\begin{cases}(q_0, \partial w/\partial\nu_E)_{L^2(E)} & \text{if } I_h = \mathrm{id}\\ ([D^2_{\mathrm{pw}}u_h\nu_E]_E\cdot\nu_E, \partial w/\partial\nu_E)_{L^2(E)} & \text{otherwise.}\end{cases}$$

This and the Cauchy inequality reveal

$$\begin{aligned}
\widehat{F}_{\mathrm{apx}}(w) - a_{\mathrm{pw}}(u_h, w) &\leqslant \mu(\mathcal{T})\sqrt{\|h_{\mathcal{T}}^{-2}w\|^2 + \sum_{E\in\mathcal{E}(\Omega)}\left(\|h_E^{-3/2}w\|^2_{L^2(E)} + \left\|h_E^{-1/2}\frac{\partial w}{\partial\nu_E}\right\|^2_{L^2(E)}\right)}\\
&\leqslant C_4\mu(\mathcal{T})|||v|||
\end{aligned}$$

with the constant C_4 from Lemma 5.1 in the last step. This and Lemma 7.1 provide $\mathrm{Res}(w) = (F - \widehat{F}_{\mathrm{apx}})(w) + \widehat{F}_{\mathrm{apx}}(w) - a_{\mathrm{pw}}(u_h, w) \lesssim \mu(\mathcal{T}) + \mathrm{apx}(F, \mathcal{T})$. □

Proposition 7.2 (efficiency up to data approximation). Let $u \in V$ solve (1.1) with the right-hand side $F \equiv \widehat{F}|_V \in V^*$ given by (7.3). If $G_0, G_1 \in P_k(\mathcal{E})$ and $(F_\alpha)_{|\alpha|\leqslant 2} \in P_k(\mathcal{T})^6$ are piecewise polynomials of degree at most $k \in \mathbb{N}_0$, then the estimator $\mu(\mathcal{T})$ of the residual is efficient up to the data approximation error

$$C_8^{-1}\mu(\mathcal{T}) \leqslant |||u - u_h|||_{\mathrm{pw}} + \mathrm{apx}(F, \mathcal{T}).$$

The constant C_8 exclusively depends on the shape-regularity of $\mathcal{T}$ and on $k \in \mathbb{N}_0$.

Before the technical proof of Proposition 7.2 follows in Subsection 7.4, the extension of the a posteriori analysis from Section 6 to $F \in V^*$ is in order.

7.3 Application to lowest-order schemes

This subsection extends the a posteriori error control from Section 6 for the right-hand side $F \in L^2(\Omega)$ to a general source $F \equiv \widehat{F}|_V \in V^*$ from (7.3). In fact, the efficient bounds of $\|u_h - J_h u_h\|_h$ from Theorem 5.1 imply the following novel result generalizing [23, Thm. 6.2] for $Q = J_h$. Let $(F_\alpha)_{|\alpha|\leqslant 2} \in P_k(\mathcal{T})^6$, $(G_0, G_1) \in P_k(\mathcal{E})^2$ be piecewise polynomials of degree at most $k \in \mathbb{N}_0$ that enter Definition 7.1 for the data approximation error $\mathrm{apx}(F, \mathcal{T})$.

Theorem 7.2 (a posteriori for $Q = J_h$). *Let $u_h \in V_h$ solve* (2.3) *with $Q = J_h$ for any of the five discrete schemes from Section* 6 *and let $u \in V$ solve* (1.1). *Then*

$$\begin{aligned}\|u - u_h\|_h^2 + \mathrm{apx}^2(F, \mathcal{T}) &\approx \mu^2(\mathcal{T}) + \sum_{E\in\mathcal{E}} h_E \|[D^2_{\mathrm{pw}} u_h]_E \tau_E\|^2_{L^2(E)} + j_h(u_h, u_h) + \mathrm{apx}^2(F, \mathcal{T})\\ &\approx \mu^2(\mathcal{T}) + \sum_{E\in\mathcal{E}} \left(h_E^{-3}\, \|[u_h]_E\|^2_{L^2(E)} + h_E^{-1} \left\| \left[\frac{\partial u_h}{\partial \nu_E}\right]_E \right\|^2_{L^2(E)} \right) + \mathrm{apx}^2(F, \mathcal{T}).\end{aligned}$$

The hidden equivalence constants exclusively depend on the shape-regularity of $\mathcal{T}$ and on $k \in \mathbb{N}_0$.

Proof. Theorem 3.2 provides $\|u - u_h\|_h \lesssim \mathrm{Res}(w) + \|u_h - J_h u_h\|_h$ for $w = v - J_h I_h I_{\mathrm{M}} v$ and some $v \in V$ for Morley, dG, and C^0IP. Recall from the proof of Theorem 6.5 that this error bound also holds for the WOPSIP scheme even without the validity of (H) in full generality. The efficient bound $\mathrm{Res}(w) \lesssim \mu(\mathcal{T}) \lesssim |||u - u_h|||_{\mathrm{pw}} + \mathrm{apx}(F, \mathcal{T})$ of the residual by the estimator $\mu(\mathcal{T})$ from Proposition 7.1–7.2 and the efficient a posteriori control of $\|u_h - J_h u_h\|_h$ from Theorem 5.1 conclude the proof. □

The original formulation (2.3) without a smoother, $Q = \mathrm{id}$, leads to an additional term

$$\widehat{F}(e_h - J_h e_h) = (\widehat{F} - \widehat{F}_{\mathrm{apx}})(e_h - J_h e_h) + \widehat{F}_{\mathrm{apx}}(e_h - J_h e_h)$$

in the a posteriori error bound from Theorem 3.2 and reflects the particular choice of the extended data $\widehat{F}$ in the definition (7.3). While the difference $\widehat{F} - \widehat{F}_{\mathrm{apx}}$ is bounded by the data approximation error $\mathrm{apx}(F, \mathcal{T})$, the non-conforming test function $e_h - J_h e_h \not\subset V$ prevents an efficient control of the higher-order volume sources in (7.4) by residual terms through an integration by parts.

For Morley, dG, and WOPSIP, the critical terms are the intermediate sources f_α for $|\alpha| = 1$ and the proof below explains why those are omitted in the (reduced) model class of right-hand sides in [23]. The following theorem generalizes [23, Thm. 6.1] for $Q = \mathrm{id}$.

Theorem 7.3 (a posteriori for $Q = \mathrm{id}$). *Suppose*

$$F_\alpha := 0 \quad \forall\, |\alpha| = 1, \qquad F_\alpha \in P_0(\mathcal{T}) \quad \forall\, |\alpha| = 2 \tag{7.9}$$

for Morley, dG, WOPSIP, and

$$F_\alpha := 0 \quad \forall\, |\alpha| = 2, \quad G_1 := 0 \tag{7.10}$$

for C^0IP. Let $u_h \in V_h$ solve (2.3) *without smoother, $Q = \mathrm{id}$, for any of the five discrete schemes from Section* 6 *and let $u \in V$ solve* (1.1). *Then*

$$\|u - u_h\|_h^2 + \mathrm{apx}^2(F, \mathcal{T}) \approx \mu^2(\mathcal{T}) + \sum_{E \in \mathcal{E}} h_E \|[D^2_{\mathrm{pw}} u_h]_E \tau_E\|^2_{L^2(E)} + j_h(u_h, u_h) + \mathrm{apx}^2(F, \mathcal{T})$$
$$\approx \mu^2(\mathcal{T}) + \sum_{E \in \mathcal{E}} \left(h_E^{-3} \|[u_h]_E\|^2_{L^2(E)} + h_E^{-1} \left\| \left[\frac{\partial u_h}{\partial \nu_E} \right]_E \right\|^2_{L^2(E)} \right) + \mathrm{apx}^2(F, \mathcal{T}).$$

The hidden equivalence constants exclusively depend on the shape-regularity of $\mathcal{T}$ and on $k \in \mathbb{N}_0$.

Proof. For Morley, dG, and C^0IP, Theorem 3.2 with $e := u - J_h u_h \in V$ and the split $\widehat{F}(v_h) = \widehat{F}_{\mathrm{apx}}(v_h) + (\widehat{F} - \widehat{F}_{\mathrm{apx}})(v_h)$ for $v_h := (1 - J_h) I_h I_{\mathrm{M}} e \in V_h$ provide

$$\|u - u_h\|_h^2 \lesssim \|u_h - J_h u_h\|_h^2 + \mathrm{Res}(e - J_h I_h I_{\mathrm{M}} e) - (\widehat{F} - \widehat{F}_{\mathrm{apx}})(v_h) - \widehat{F}_{\mathrm{apx}}(v_h) \tag{7.11}$$
$$\lesssim \|u_h - J_h u_h\|_h^2 + (\mu(\mathcal{T}) + \mathrm{apx}(F, \mathcal{T}))\, |||u - J_h u_h||| - \widehat{F}_{\mathrm{apx}}(v_h) \tag{7.12}$$

with Lemma 7.1 and Proposition 7.1 in the last step. The discussion in the proof of Theorem 6.5 implies (7.11)–(7.12) also for the WOPSIP method. The triangle inequality $|||u - J_h u_h||| \leqslant \|u - u_h\|_h + \|u_h - J_h u_h\|_h$, (7.12), and a Young inequality verify

$$\|u - u_h\|_h^2 \lesssim \|u_h - J_h u_h\|_h^2 + \mu^2(\mathcal{T}) + \mathrm{apx}^2(F, \mathcal{T}) - \widehat{F}_{\mathrm{apx}}(v_h). \tag{7.13}$$

It remains to bound the extra term $\widehat{F}_{\mathrm{apx}}(v_h)$. Recall the abbreviations F_0, F_1, F_2 from (7.6).

The key step towards an efficient control of $\widehat{F}_{\mathrm{apx}}(v_h)$ is an integration by parts in (7.8) that collects the volume loads F_0, F_1, F_2 in the single residual term $\mu_1(\mathcal{T})$ (resp. the jumps in $\mu_2(\mathcal{T}), \mu_3(\mathcal{T})$). A similar approach for the efficient bound of $\widehat{F}_{\mathrm{apx}}(v_h)$ with the non-conforming test function $v_h \notin V$ leads to additional terms from the product rule for jumps on the edge $E \in \mathcal{E}$, namely,

$$[F_1 \cdot \nu_E v_h]_E = \langle F_1 \cdot \nu_E \rangle_E \, [v_h]_E + [F_1 \cdot \nu_E]_E \, \langle v_h \rangle_E$$
$$[F_2 \nu_E \cdot \nabla v_h]_E = \langle F_2 \nu_E \rangle_E \, [\nabla v_h]_E + [F_2 \nu_E]_E \, \langle \nabla v_h \rangle_E .$$

However, the average terms $\langle F_1 \cdot \nu_E \rangle_E$ and $\langle F_2 \nu_E \rangle_E$ over the edges $E \in \mathcal{E}$ are *no* residuals and their efficiency is open; cf. the partial efficiency result (excluding the average terms) in [37, Thm. 7.2] or the omission of the efficiency analysis in [21]. Instead, the assumptions (7.9)–(7.10) and the additional information on the structure of the test function $v_h \in (1 - J_h) I_h I_{\mathrm{M}} V$ allows the efficient control of $\widehat{F}_{\mathrm{apx}}(v_h)$.

Case $I_h = \mathrm{id}$: Since $F_2 \in P_0(\mathcal{T}; \mathbb{S}) = D^2_{\mathrm{pw}} P_2(\mathcal{T})$ is piecewise constant, $I_{\mathrm{M}} v_h = 0$ from $I_{\mathrm{M}} J_h = I_{\mathrm{M}}$ and (4.3) verify the L^2 orthogonality $v_h \perp F_2$. This and (7.9) lead to

$$\widehat{F}_{\mathrm{apx}}(v_h) = (F_0, v_h)_{L^2(\Omega)} + (G_0, v_h)_{L^2(\Gamma_0)} + \sum_{E \in \mathcal{E}(\Gamma_1)} ((1 - \Pi_{E,0}) G_1, \partial_\nu v_h)_{L^2(E)} \tag{7.14}$$

with $\Pi_{E,0} \partial_\nu v_h \, ds = 0$ for any $E \in \mathcal{E}$ from $I_{\mathrm{M}} v_h = 0$.

Case $I_h = I_C$: Since the test function $v_h \in V + S^2_0(\mathcal{T})$ is H^1 conforming, (7.10) and an integration by parts show

$$\widehat{F}_{\mathrm{apx}}(v_h) = (F_0 - \mathrm{div}_{\mathrm{pw}} F_1, v_h)_{L^2(\Omega)} + (G_0 + [F_1]_E \cdot \nu_E, v_h)_{L^2(\Gamma_j)}. \tag{7.15}$$

Cauchy inequalities, Lemma 5.1, and (7.14) for Morley, dG, WOPSIP and (7.15) for C^0IP result in $|\widehat{F}_{\mathrm{apx}}(v_h)| \lesssim \mu(\mathcal{T}) |||e|||$. This, (7.13), and a Young inequality provides

$$\|u - u_h\|_h^2 \lesssim \|u_h - J_h u_h\|_h^2 + \mu^2(\mathcal{T}) + \mathrm{apx}^2(F, \mathcal{T}).$$

Theorem 5.1 and the efficiency of $\mu(\mathcal{T})$ from Proposition 7.2 conclude the proof. □

Remark 7.3 (apx$(F, \mathcal{T})$ in Theorem 7.3). Since Theorem 7.3 requires $(F_\alpha)_{|\alpha|=1} \equiv 0$ to vanish for the Morley, dG, and WOPSIP methods, the data approximation error apx$(F, \mathcal{T})$ includes the term $\sqrt{\sum_{|\alpha|=1} \|h_\mathcal{T} f_\alpha\|^2_{L^2(\Omega)}}$. This term is linear in the mesh-size and converges with the expected rate for lowest-order schemes. This term may even be of higher order if the triangulation is quasi-uniform and Ω is non-convex with a reduced convergence rate $\|u - u_h\|_h = \mathcal{O}(h^\sigma_{\max})$ of the schemes. However, it is *not* a classical (higher-oder) data oscillation term if $f_\alpha \neq 0$ does not vanish for all $|\alpha| = 1$. The assumption (7.10) for C^0IP leads to the term $\sqrt{\sum_{|\alpha|=2} \|f_\alpha\|^2_{L^2(\Omega)}}$ independent of the mesh-size in the data approximation error apx$(F, \mathcal{T})$. Hence a meaningful interpretation of the a posteriori estimate in Theorem 7.3 for C^0IP requires $\|f_\alpha\|_{L^2(\Omega)}$ to be small for all $|\alpha| = 2$.

Remark 7.4 (smoother vs. no smoother). Since Theorem 7.2 for the smoother $Q = J_h$ applies to any choice of data approximations, Remark 7.1 shows that the data approximation error apx$(F, \mathcal{T})$ can be replaced by data oscillations osc$(F, \mathcal{T})$ of arbitrary order. This provides a novel reliable and efficient a posteriori error bound for any right-hand side $F \in V^*$ of the form (7.3) up to data oscillations.

For no smoother $Q = \mathrm{id}$, additional requirements on the data approximations (7.9) for Morley, dG, and WOPSIP (resp. (7.10) for C^0IP) in Theorem 7.3 seem necessary for an efficient error control. However, Remark 7.3 explains that this either restricts the admissible data in (7.3) or leads to terms in the data approximation error apx$(F, \mathcal{T})$ that are no oscillations.

Remark 7.5 ($F \in L^2(\Omega)$). Theorems 7.2–7.3 for source terms $F \equiv f \in L^2(\Omega)$ with $F_0 := \Pi_2 f \in P_2(\mathcal{T})$ (and $F_\alpha = f_\alpha \equiv 0$ for all $|\alpha| = 1, 2$ as well as $G_0 = G_1 \equiv 0$) imply the a posteriori results of Theorems 6.1–6.5. Indeed, the Pythagoras theorem $\|f\|^2_{L^2(T)} = \|f - \Pi_2 f\|^2_{L^2(T)} + \|\Pi_2 f\|^2_{L^2(T)}$ for the triangle $T \in \mathcal{T}$ verifies

$$\mu(\mathcal{T})^2 + \mathrm{apx}^2(F, \mathcal{T}) = \|h^2_\mathcal{T} f\|^2 + \sum_{E \in \mathcal{E}(\Omega)} \begin{cases} 0 & \text{if } I_h = \mathrm{id} \\ h_E \|[D^2_{\mathrm{pw}} u_h \nu_E]_E \cdot \nu_E\|^2_{L^2(E)} & \text{otherwise.} \end{cases}$$

Since $j_h(u_h, u_h) = 0$ vanishes for all Morley solutions $u_h \in \mathrm{M}(\mathcal{T})$ and every C^0IP solution $u_h \in S^2_0(\mathcal{T})$ has zero jump $[u_h]_E \equiv 0$ along an edge $E \in \mathcal{E}$, Theorems 7.2–7.3 recover the corresponding results from Section 6.

7.4 Proof of Proposition 7.2

This proof applies the bubble-function methodology [45]. Recall

$$\|D^2(u - I_\mathrm{M} u)\|_{L^2(T)} = \min_{v_h \in P_2(\mathcal{T})} \|D^2(u - v_h)\|_{L^2(T)} \leqslant \|D^2(u - u_h)\|_{L^2(T)} \tag{7.16}$$

for any $T \in \mathcal{T}$ from the best-approximation property (4.4).

Step 1 (efficiency of the volume contribution). Let $\varpi := F_0 - \mathrm{div}_\mathrm{pw} F_1 + \mathrm{div}^2_\mathrm{pw} F_2 \in P_k(T)$ abbreviate the volume contribution of $\mu(\mathcal{T})$ for some $T \in \mathcal{T}$. The element bubble-function $b_T = 27\varphi_1\varphi_2\varphi_3 \in P_3(T) \cap H^1_0(T)$ with $\|b_T\|_{L^\infty(T)} = 1$ is given in terms of the three barycentric coordinates $\varphi_j \in P_1(T)$ for $j = 1, 2, 3$. Since $a_\mathrm{pw}(I_\mathrm{M} u, b^2_T \varpi) = 0$ from (4.3) and $I_\mathrm{M}(b^2_T \varpi) \equiv 0$, the equivalence of the weighted norm $\|b_T \varpi\|_{L^2(T)} \approx \|\varpi\|_{L^2(T)}$ and an integration by parts without boundary terms from $b^2_T \varpi \in H^2_0(T)$ show

$$\begin{aligned} \|\varpi\|^2_{L^2(T)} &\approx (F_0 - \mathrm{div}_\mathrm{pw} F_1 + \mathrm{div}^2_\mathrm{pw} F_2, b^2_T \varpi)_{L^2(T)} \\ &= \widehat{F}_\mathrm{apx}(b^2_T \varpi) = a_\mathrm{pw}(u - I_\mathrm{M} u, b^2_T \varpi) + (\widehat{F}_\mathrm{apx}(b^2_T \varpi) - F(b^2_T \varpi)) \\ &\lesssim \Big(\|D^2(u - I_\mathrm{M} u)\|_{L^2(T)} + \mathrm{apx}(F, T)\Big)\, \|D^2(b^2_T \varpi)\|_{L^2(T)} \end{aligned}$$

with $(F - \widehat{F}_\mathrm{apx})(v) = (F - \widehat{F}_\mathrm{apx})(v - J_h I_h I_\mathrm{M} v) \lesssim \mathrm{apx}(F, T)\|D^2 v\|_{L^2(T)}$ for $v \in H^2_0(T)$ from $I_\mathrm{M} v \equiv 0$ plus Lemma 7.1 and a Cauchy inequality in the last step. This and the inverse inequality $h^2_T \|D^2(b^2_T \varpi)\|_{L^2(T)} \lesssim \|b^2_T \varpi\|_{L^2(T)} \leqslant \|\varpi\|_{L^2(T)}$ from [30, Lem. 12.1] conclude the proof of the local efficiency of the volume contributions, namely,

$$h^2_T \|F_0 - \mathrm{div}_\mathrm{pw} F_1 + \mathrm{div}^2_\mathrm{pw} F_2\|_{L^2(T)} \lesssim \|D^2(u - I_\mathrm{M} u)\|_{L^2(T)} + \mathrm{apx}(F, T). \tag{7.17}$$

Step 2 (set-up for an interior edge). For any interior edge $E = \operatorname{conv}\{A, B\} = T_+ \cap T_- \in \mathcal{E}(\Omega)$, let $\varphi_{P_+}, \varphi_{A,+}, \varphi_{B,+} \in P_1(\mathbb{R}^2)$ (resp. $\varphi_{P_-}, \varphi_{A,-}, \varphi_{B,-} \in P_1(\mathbb{R}^2)$) denote the barycentric coordinates of $T_+ = \operatorname{conv}\{P_+, A, B\}$ (resp. $T_- := \operatorname{conv}\{P_-, A, B\}$) seen as globally defined affine functions. The edge bubble-function reads $b_E := 16\varphi_{A,+}\varphi_{A,-}\varphi_{B,+} \times\varphi_{B,-} \in P_4(\mathbb{R}^2) \cap H^1_0(\omega(E))$ and $b_{T_\pm} := 27\varphi_{A,\pm}\varphi_{B,\pm}\varphi_{P_\pm} \in P_3(\mathbb{R}^2) \cap H^1_0(T_\pm)$ denotes the element bubble-function in $T_\pm \in \mathcal{T}$ with $\nu_{T_\pm}|_E = \pm\nu_E$. Let $\partial_\nu := \partial/\partial\nu_E$ abbreviate the normal derivative and recall that the gradient $\nabla\varphi_{P_+} = -\varrho_E^{-1}\nu_E$ of the barycentric coordinate φ_{P_+} scales like h_E^{-1} with the height $\varrho_E := 2|T_+|/|E| \approx h_E$ from shape-regularity. The function b_E^2 has been utilised in the literature before, e.g., in [4, p. 788] with its scaling properties; the usage of $b_{T_\pm}^2$ is standard. The product rule and $\varphi_{T_+}|_E \equiv 0$ verify

$$\partial_\nu(\varphi_{T_+} b_E^2) = -\varrho_E^{-1} b_E^2 \quad \text{on } E. \tag{7.18}$$

Given $p \in \mathbb{N}_0$, any polynomial $q \in P_p(E)$ on the edge E defines a unique polynomial on the straight line L that extends $E \subset L$. The extension of q from L to $\mathbb{R}^2$ by constant values along the normal ν_E defines a polynomial $\widehat{q} \in P_p(\mathbb{R}^2)$ on $\mathbb{R}^2$ of the same degree. Let $\Pi_L P_\pm \in L$ denote the projection of the vertex $P_\pm \in T_\pm$ opposite to E onto L along the normal direction ν_E. The maximal value $\|\widehat{q}\|_{L^\infty(\omega(E))}$ is attained on the line segment $\widehat{L} := \operatorname{conv}\{E, \Pi_L P_\pm\} \subset L$ and the shape-regularity controls the ratio $|\widehat{L}|/|E| \geqslant 1$. Hence

$$\|\widehat{q}\|_{L^\infty(w(E))} = \|q\|_{L^\infty(\widehat{L})} \leqslant C\|q\|_{L^\infty(E)} \tag{7.19}$$

follows with some constant $C \approx 1$ that exclusively depends on the shape-regularity of the triangulation $\mathcal{T}$ and on p.

Step 3 (efficiency of the first jump contribution). This step establishes the local efficiency of the term $\vartheta_E := G_1 + [F_1 - \operatorname{div}_{\mathrm{pw}} F_2 - \partial F_2\tau_E/\partial s]_E \cdot \nu_E \in P_k(E)$ in the form

$$h_E^{3/2}\|\vartheta_E\|_{L^2(E)} \lesssim \|D^2_{\mathrm{pw}}(u - I_{\mathrm{M}}u)\|_{L^2(\omega(E))} + \operatorname{apx}(F, T_+) + \operatorname{apx}(F, T_-). \tag{7.20}$$

Let $\xi_E \in P_{2k}(E)$ denote the (unique) Riesz representation of the functional $\varrho_E(\partial_\nu b_E^2, \cdot)_{L^2(E)}$ in the vector space $P_{2k}(E)$ with respect to the weighted scalar product $(b_E^2\cdot, \cdot)_{L^2(E)}$, i.e.,

$$(b_E^2\xi_E, p_{2k})_{L^2(E)} = \varrho_E(\partial_\nu b_E^2, p_{2k})_{L^2(E)} \qquad \forall\, p_{2k} \in P_{2k}(E). \tag{7.21}$$

This, the equivalence of the weighted norm $\|b_E\xi_E\|_{L^2(E)} \approx \|b_E\|_{L^2(E)}$, and $h_E \approx \varrho_E$ show

$$\|\xi_E\|^2_{L^2(E)} \approx \|b_E\xi_E\|^2_{L^2(E)} = h_E(\partial_\nu b_E^2, \xi_E)_{L^2(E)} \lesssim h_E\|\partial_\nu b_E^2\|_{L^2(E)}\|\xi_E\|_{L^2(E)}$$

with a Cauchy inequality in the last step. Hölder's inequality and an inverse estimate [30, Lem. 12.1] lead to

$$\|\partial_\nu b_E^2\|_{L^2(E)} \leqslant h_E^{1/2}\|\partial_\nu b_E^2\|_{L^\infty(E)} \leqslant h_E^{1/2}\|\nabla b_E^2\|_{L^\infty(E)} \lesssim h_E^{-1/2}\|b_E^2\|_{L^\infty(T)} \leqslant h_E^{-1/2}.$$

This proves $\|\xi_E\|_{L^2(E)} \lesssim h_E^{1/2}$ and another inverse inequality provides $\|\xi_E\|_{L^\infty(E)} \lesssim h_E^{-1/2}\|\xi_E\|_{L^2(E)} \lesssim 1$. Let $\widehat{\vartheta}_E \in P_k(\mathbb{R}^2)$ and $\widehat{\xi}_E \in P_{2k}(\mathbb{R}^2)$ denote the extension of $\vartheta_E \in P_k(E)$ and $\xi_E \in P_{2k}(E)$ to $\mathbb{R}^2$ as in Step 2. This, (7.18), and (7.21) verify the L^2 orthogonality

$$\partial_\nu((b_E^2 + \varphi_{T_+} b_E^2\widehat{\xi}_E)\widehat{\vartheta}_E) = (\partial_\nu(b_E^2) - \varrho_E^{-1} b_E^2\widehat{\xi}_E)\widehat{\vartheta}_E \perp P_k(E) \qquad \text{in } L^2(E).$$

Let $\xi_{T_\pm} \in P_{2k}(T_\pm)$ be the unique solution to

$$(b_{T_\pm}^2\xi_{T_\pm}, p_{2k})_{L^2(T_\pm)} = (b_E^2 + \varphi_{T_+} b_E^2\widehat{\xi}_E, p_{2k})_{L^2(T_\pm)} \qquad \forall\, p_{2k} \in P_{2k}(T_\pm).$$

An inverse inequality and (7.19) show $\|\xi_{T_\pm}\|_{L^\infty(T_\pm)} \lesssim 1$. The definition of $\xi_{T_\pm}$ verifies that the function $\psi_E := (b_E^2 + \varphi_{T_+} b_E^2\widehat{\xi}_E - b_{T_+}^2\chi_{T_+}\widehat{\xi}_{T_+} - b_{T_-}^2\chi_{T_-}\widehat{\xi}_{T_-})\widehat{\vartheta}_E \in H^2_0(\omega(E))$ is $L^2(T_\pm)$ orthogonal to $P_k(T_\pm)$. Since $b_{T_\pm}^2 \in H^2_0(T_\pm)$ vanishes on E, the normal derivative $\partial_\nu\psi_E|_E \equiv \partial_\nu((b_E^2 + \varphi_{T_+} b_E^2\widehat{\xi}_E)\widehat{\vartheta}_E)|_E \perp P_k(E)$ is $L^2(E)$ orthogonal to $P_k(E)$. This, (7.7), and an integration by parts show

$$\begin{aligned}
0 &= (F_0 - \operatorname{div}_{\mathrm{pw}} F_1 + \operatorname{div}^2_{\mathrm{pw}} F_2, \psi_E)_{L^2(\omega(E))} + (G_2 - [F_2\nu_E]_E \cdot \nu_E, \partial_\nu\psi_E)_{L^2(E)} \\
&= (F_0, \psi_E)_{L^2(\omega(E))} + (F_1 - \operatorname{div}_{\mathrm{pw}} F_2, \nabla\psi_E)_{L^2(\omega(E))} \\
&\quad - ([F_1 - \operatorname{div}_{\mathrm{pw}} F_2]_E \cdot \nu_E, \psi_E)_{L^2(E)} + (G_2 - [F_2\nu_E]_E \cdot \nu_E, \partial_\nu\psi_E)_{L^2(E)} \\
&= \widehat{F}_{\mathrm{apx}}(\psi_E) - (\vartheta_E, \psi_E)_{L^2(E)}.
\end{aligned}$$

The Morley interpolation $I_{\mathrm{M}}\psi_E \equiv 0$ of $\psi_E \in H^2_0(\omega(E))$ vanishes from $\Pi_{E,0}\partial_\nu\psi_E = 0$ and $a_{\mathrm{pw}}(I_{\mathrm{M}}u, \psi_E) = 0$ follows from (4.3). Since $(\psi_E - b_E^2\vartheta_E)|_E \equiv 0$ is zero on E, the equivalence $\|\vartheta_E\|_{L^2(E)} \approx \|b_E\vartheta_E\|_{L^2(E)}$ results in

$$\|\vartheta_E\|^2_{L^2(E)} \approx (\vartheta_E, b_E^2\vartheta_E)_{L^2(E)} = (\vartheta_E, \psi_E)_{L^2(E)} = \widehat{F}_{\mathrm{apx}}(\psi_E).$$

With $a(u,\psi_E) = F(\psi_E)$ from (1.1), this shows

$$\begin{aligned}\|\vartheta_E\|^2_{L^2(E)} &= a_{\mathrm{pw}}(u - I_{\mathrm{M}}u, \psi_E) + (\widehat{F}_{\mathrm{apx}}(\psi_E) - F(\psi_E))\\ &\leqslant \Big(\|D^2_{\mathrm{pw}}(u - I_{\mathrm{M}}u)\|_{L^2(\omega(E))} + \operatorname{apx}(F, T_-) + \operatorname{apx}(F, T_+)\Big)\,\|D^2\psi_E\|_{L^2(\omega(E))}.\end{aligned}$$

The inverse inequality, $\|\psi_E\|_{L^2(\omega(E))} \lesssim \|\widehat{\vartheta}_E\|_{L^2(\omega(E))}$, and (7.19) provide

$$h_E^2\|D^2\psi_E\|_{L^2(\omega(E))} \lesssim \|\psi_E\|_{L^2(\omega(E))} \lesssim \|\widehat{\vartheta}_E\|_{L^2(\omega(E))} \lesssim h_E^{1/2}\|\vartheta_E\|_{L^2(E)}. \tag{7.22}$$

This verifies the efficiency (7.20) of the jump contributions ϑ_E.

Step 4 (efficiency of the second jump contribution). The local efficiency of the remaining term follows with similar arguments. Since the Hessian $D^2_{\mathrm{pw}}u_h$ of $u_h \in P_2(\mathcal{T})$ is piecewise constant, the stability of the L^2 projection results in

$$\|(1 - \Pi_{E,0})(G_1 + [F_2\nu_E]_E \cdot \nu_E)\|_{L^2(E)} \leqslant \|G_1 + [(F_2 - D^2_{\mathrm{pw}}u_h)\nu_E]_E \cdot \nu_E\|_{L^2(E)}.$$

It is therefore sufficient to prove the local efficiency of the term $\zeta_E := G_1 + [(F_2 - D^2_{\mathrm{pw}}u_h)\nu_E]_E \cdot \nu_E \in P_k(E)$, namely,

$$h_E^{1/2}\|\zeta_E\|_{L^2(E)} \lesssim \|D^2_{\mathrm{pw}}(u - u_h)\|_{L^2(\omega(E))} + \operatorname{apx}(F, T_+) + \operatorname{apx}(F, T_-). \tag{7.23}$$

Indeed, let $\varrho_{T_\pm} \in P_{2k}(T_\pm)$ be the unique solution to

$$(b^2_{T_\pm}\varrho_{T_\pm}, p_{2k})_{L^2(T_\pm)} = (\varphi_{T_+}b_E^2, p_{2k})_{L^2(T_\pm)} \qquad \forall\, p_{2k} \in P_{2k}(T_\pm).$$

Observe that $\psi_2 := -(\varphi_{T_+}b_E^2 - b^2_{T_+}\chi_{T_+}\varrho_{T_+} - b^2_{T_-}\chi_{T_-}\varrho_{T_-})\widehat{\zeta}_E \in H^2_0(\omega(E))$ is L^2 perpendicular to $P_k(T_\pm)$ with zero trace $\psi_2|_E \equiv 0$ on E. This and an integration by parts show

$$\begin{aligned}0 &= (F_0 - \operatorname{div}_{\mathrm{pw}}F_1 + \operatorname{div}^2_{\mathrm{pw}}F_2, \psi_2)_{L^2(\omega(E))} + (\vartheta_E, \psi_2)_{L^2(E)}\\ &= \widehat{F}_{\mathrm{apx}}(\psi_2) - (G_1 + [F_2\nu_E]_E \cdot \nu_E, \partial_\nu\psi_2)_{L^2(E)} = \widehat{F}_{\mathrm{apx}}(\psi_2) - a_{\mathrm{pw}}(u_h, \psi_2) - (\zeta_E, \partial_\nu\psi_2)_{L^2(E)}\end{aligned}$$

with $([D^2_{\mathrm{pw}}u_h\nu_E]_E \cdot \nu_E, \partial_\nu\psi_2)_{L^2(E)} = a_{\mathrm{pw}}(u_h, \psi_2)$ in the last step. The equivalences $\|\zeta_E\|_{L^2(E)} \approx \|b_E\zeta_E\|_{L^2(E)}$ and $h_E \approx \varrho_E$, (7.18), and $\partial_\nu(\psi_2 + \varphi_{T_+}b_E^2\widehat{\zeta}_E)|_E \equiv 0$ provide

$$\begin{aligned}h_E^{-1}\|\zeta_E\|^2_{L^2(E)} &\approx \varrho_E^{-1}(\zeta_E, b_E^2\zeta_E)_{L^2(E)} = -(\zeta_E, \partial_\nu(\varphi_{T_+}b_E^2\widehat{\zeta}_E))_{L^2(E)} = (\zeta_E, \partial_\nu\psi_2)_{L^2(E)}\\ &= \widehat{F}_{\mathrm{apx}}(\psi_2) - a_{\mathrm{pw}}(u_h, \psi_2) = a_{\mathrm{pw}}(u - u_h, \psi_2) + (\widehat{F}_{\mathrm{apx}}(\psi_2) - F(\psi_2))\end{aligned}$$

with $a(u, \psi_2) = F(\psi_2)$ from (1.1) in the last step. The remaining steps follow Step 3 and utilize $\|D^2\psi_2\|_{L^2(\omega(E))} \lesssim h_E^{-3/2}\|\zeta_E\|_{L^2(E)}$ from an inverse inequality as in (7.22); further details are omitted. The combination of the local efficiency results (7.17), (7.20), and (7.23) with (7.16) concludes the proof. □

Funding: The research of the first two authors have been supported by the Deutsche Forschungsgemeinschaft in the Priority Program 1748 under the project 'Foundation and Application of Generalized Mixed FEM Towards Nonlinear Problems in Solid Mechanics' (CA 151/22-2). This paper has been supported by SPARC project (id 235) 'The Mathematics and Computation of Plates' and SERB POWER Fellowship SPF/2020/000019. The second author is supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – The Berlin Mathematics Research Center MATH+ (EXC-2046/1, project ID: 390685689).

References

[1] G. A. Baker, Finite element methods for elliptic equations using nonconforming elements, *Math. Comp.* **31** (1977), No. 137, 45–59.

[2] L. Beirão da Veiga, J. Niiranen, and R. Stenberg, A posteriori error estimates for the Morley plate bending element, *Numer. Math.* **106** (2007), No. 2, 165–179.

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

[4] S. C. Brenner, T. Gudi, and L.-Y. Sung, An a posteriori error estimator for a quadratic C^0-interior penalty method for the biharmonic problem, *IMA J. Numer. Anal.* **30** (2010), No. 3, 777–798.

[5] S. C. Brenner, T. Gudi, and L.-Y. Sung, A weakly over-penalized symmetric interior penalty method for the biharmonic problem, *Electron. Trans. Numer. Anal.* **37** (2010), 214–238.

[6] S. C. Brenner and L. R. Scott, *The Mathematical Theory of Finite Element Methods*, 3rd ed., Springer, 2007.

[7] S. C. Brenner and L.-Y. Sung, C^0 interior penalty methods for fourth order elliptic boundary value problems on polygonal domains, *J. Sci. Comput.* **22/23** (2005), 83–118.

[8] S. C. Brenner, L.-Y. Sung, H. Zhang, and Y. Zhang, A Morley finite element method for the displacement obstacle problem of clamped Kirchhoff plates, *J. Comput. Appl. Math.* **254** (2013), 31–42.

[9] S. C. Brenner, Convergence of nonconforming multigrid methods without full elliptic regularity, *Math. Comp.* **68** (1999), No. 225, 25–53.

[10] C. Carstensen, S. Bartels, and S. Jansche, A posteriori error estimates for nonconforming finite element methods, *Numer. Math.* **92** (2002), No. 2, 233–256.

[11] C. Carstensen, M. Eigel, R. H. W. Hoppe, and C. Löbhard, A review of unified a posteriori finite element error control, *Numer. Math. Theory Methods Appl.* **5** (2012), No. 4, 509–558.

[12] C. Carstensen and D. Gallistl, Guaranteed lower eigenvalue bounds for the biharmonic equation, *Numer. Math.* **126** (2014), No. 1, 33–51.

[13] C. Carstensen, D. Gallistl, and J. Hu, A posteriori error estimates for nonconforming finite element methods for fourth-order problems on rectangles, *Numer. Math.* **124** (2013), No. 2, 309–335.

[14] C. Carstensen, D. Gallistl, and J. Hu, A discrete Helmholtz decomposition with Morley finite element functions and the optimality of adaptive finite element schemes, *Comput. Math. Appl.* **68** (2014), No. 12, part B, 2167–2181.

[15] C. Carstensen, D. Gallistl, and N. Nataraj, Comparison results of nonstandard P_2 finite element methods for the biharmonic problem, *ESAIM Math. Model. Numer. Anal.* (2015), 977–990.

[16] C. Carstensen, D. Gallistl, and M. Schedensack, Adaptive nonconforming Crouzeix–Raviart FEM for eigenvalue problems, *Math. Comp.* **84** (2015), 1061–1087.

[17] C. Carstensen, J. Gedicke, and D. Rim, Explicit error estimates for Courant, Crouzeix–Raviart and Raviart–Thomas finite element methods, *J. Comput. Math.* **30** (2012), No. 4, 337–353.

[18] C. Carstensen, B. Gräßle, and N. Nataraj, A posteriori error control for fourth-order semilinear problems with quadratic nonlinearity, *SIAM J. Numer. Anal.* **62** (2024) (to appear).

[19] C. Carstensen and J. Hu, A unifying theory of a posteriori error control for nonconforming finite element methods, *Numer. Math.* **107** (2007), No. 3, 473–502.

[20] C. Carstensen, J. Hu, and A. Orlando, Framework For The A Posteriori Error Analysis Of Nonconforming Finite Element, *SIAM J. Numer. Anal.* **45** (2007), No. 1, 68–82.

[21] C. Carstensen, G. Mallik, and N. Nataraj, A priori and a posteriori error control of discontinuous Galerkin finite element methods for the von Kármán equations, *IMA J. Numer. Anal.* **39** (2019), 167–200.

[22] C. Carstensen and C. Merdon, Computational survey on a posteriori error estimators for the Crouzeix–Raviart nonconforming finite element method for the Stokes problem, *Comput. Methods Appl. Math.* **14** (2014), No. 1, 35–54.

[23] C. Carstensen and N. Nataraj, A priori and a posteriori error analysis of the Crouzeix–Raviart and Morley FEM with original and modified right-hand sides, *Comput. Methods Appl. Math.* **21** (2021), No. 2, 289–315.

[24] C. Carstensen and N. Nataraj, Lowest-order equivalent nonstandard finite element methods for biharmonic plates, *ESAIM: M2AN* **56** (2022), No. 1, 41–78.

[25] C. Carstensen and S. Puttkammer, How to prove the discrete reliability for nonconforming finite element methods, *J. Comput. Math* **38** (2020), No. 1, 142–175.

[26] C. Carstensen and S. Puttkammer, Direct guaranteed lower eigenvalue bounds with optimal a priori convergence rates for the bi-Laplacian, *SIAM Journal on Numerical Analysis*, **61** (2023), No. 2, 812–836.

[27] P. G. Ciarlet, *The Finite Element Method for Elliptic Problems*, North-Holland, Amsterdam, 1978.

[28] P. Clément, Approximation by finite element functions using local regularization, *Rev. Française Automat. Informat. Recherche Opérationnelle Sér. Rouge Anal. Numér.* **9** (1975), No. R-2, 77–84.

[29] E. Dari, R. Duran, C. Padra, and V. Vampa, A posteriori error estimators for nonconforming finite element methods, *RAIRO Modél. Math. Anal. Numér.* **30** (1996), No. 4, 385–400.

[30] A. Ern and J. L. Guermond, *Finite Elements, I: Approximation and Interpolation*, Texts in Applied Mathematicsi, Vol. 72, Springer International Publishing, Cham, 2021.

[31] L. C. Evans, *Partial Differential Equations*, 2nd ed., Graduate Studies in Mathematics, Vol. 19, American Mathematical Society, Providence, R.I, 2010.

[32] X. Feng and O. A. Karakashian, Fully discrete dynamic mesh discontinuous Galerkin methods for the Cahn–Hilliard equation of phase transition, *Math. Comp.* **76** (2007), No. 259, 1093–1117.

[33] D. Gallistl, Morley finite element method for the eigenvalues of the biharmonic operator, *IMA J. Numer. Anal.* **35** (2015), No. 4, 1779–1811.

[34] E. H. Georgoulis, P. Houston, and J. Virtanen, An a posteriori error indicator for discontinuous Galerkin approximations of fourth-order elliptic problems, *IMA J. Numer. Anal.* **31** (2011), No. 1, 281–298.

[35] J. Hu, Z. Shi, and J. Xu, Convergence and optimality of the adaptive Morley element method, *Numer. Math.* **121** (2012), No. 4, 731–752.

[36] J. Hu and Z. C. Shi, A new a posteriori error estimate for the Morley element, *Numer. Math.* **112** (2009), No. 1, 25–40.

[37] D. Kim, A. K. Pani, and E.-J. Park, Morley finite element methods for the stationary quasi-geostrophic equation, *Computer Methods in Applied Mechanics and Engineering* **375** (2021), 113639.

[38] L. R. Scott, *C Piecewise Polynomials Satisfying Boundary Conditions*, Research Report UC/CS TR-2019-18, Dept. Comp. Sci., Univ. Chicago, Report, 2019.

[39] L. R. Scott and S. Zhang, Finite element interpolation of nonsmooth functions satisfying boundary conditions, *Math. Comp.* **54** (1990), No. 190, 483–493.

[40] E. Süli and I. Mozolevski, *hp*-version interior penalty DGFEMs for the biharmonic equation, *Comput. Methods Appl. Mech. Engrg.* **196** (2007), No. 13-16, 1851–1863.

[41] R. Vanselow, New results concerning the DWR method for some nonconforming FEM, *Appl. Math.* **57** (2012), No. 6, 551–568.

[42] A. Veeser and P. Zanotti, Quasi-optimal nonconforming methods for symmetric elliptic problems, I — Abstract theory, *SIAM J. Numer. Anal.* **56** (2018), No. 3, 1621–1642.

[43] A. Veeser and P. Zanotti, Quasi-optimal nonconforming methods for symmetric elliptic problems, III — Discontinuous Galerkin and other interior penalty methods, *SIAM J. Numer. Anal.* **56** (2018), No. 5, 2871–2894.

[44] A. Veeser and P. Zanotti, Quasi-optimal nonconforming methods for symmetric elliptic problems, II — Overconsistency and classical nonconforming elements, *SIAM J. Numer. Anal.* **57** (2019), No. 1, 266–292.

[45] R. Verfürth, *A Review of a Posteriori Error Estimation and Adaptive Mesh-Refinement Techniques*, Wiley-Teubner, 1996.

A A posteriori error control of a piecewise polynomial source in $H^{-2}(\Omega)$

This appendix provides an alternative view on the reliable and efficient estimator from Section 7 as lower and upper bounds for the dual norm of a piecewise polynomial source in $H^{-2}(\Omega)$. Suppose the piecewise polynomials $\Lambda_0 \in P_k(\mathcal{T})$, $\Lambda_1 \in P_k(\mathcal{T};\mathbb{R}^2)$, and $\Lambda_2 \in P_k(\mathcal{T};\mathbb{S})$ define the linear functional $\Lambda \in H^{-2}(\Omega)$ by

$$\Lambda(v) := \int_\Omega (\Lambda_0 v + \Lambda_1 \cdot \nabla v + \Lambda_2 : D^2 v)\,\mathrm{d}x \qquad \forall\, v \in H^2_0(\Omega). \tag{A.1}$$

Recall the transfer operators I_{M}, I_h, J_h for the five quadratic discretization schemes of Section 4 listed in Table 1. A reliable and efficient estimator $\mu^2(\mathcal{T}) := \mu_1^2(\mathcal{T}) + \mu_2^2(\mathcal{T}) + \mu_3^2(\mathcal{T})$ of the functional Λ is given by

$$\begin{aligned}
\mu_1^2(\mathcal{T}) &:= \|h_{\mathcal{T}}^2(\Lambda_0 - \mathrm{div}_{\mathrm{pw}}\Lambda_1 + \mathrm{div}^2_{\mathrm{pw}}\Lambda_2)\|^2\\
\mu_2^2(\mathcal{T}) &:= \sum_{E\in\mathcal{E}(\Omega)} h_E^3 \|[\Lambda_1 - \mathrm{div}_{\mathrm{pw}}\Lambda_2 - \partial(\Lambda_2\tau_E)/\partial s]_E \cdot \nu_E\|^2_{L^2(E)}\\
\mu_3^2(\mathcal{T}) &:= \sum_{E\in\mathcal{E}(\Omega)} \begin{cases} h_E\|(1-\Pi_{E,0})[\Lambda_2\nu_E]_E\cdot\nu_E\|^2_{L^2(E)} & \text{if } I_h = \mathrm{id}\\ h_E\|[\Lambda_2\nu_E]_E\cdot\nu_E\|^2_{L^2(E)} & \text{if } I_h = I_{\mathrm{C}}.\end{cases}
\end{aligned}$$

Theorem A.1 (reliability and efficiency). *There exist positive constants $C_{\mathrm{rel}}, C_{\mathrm{eff}} > 0$ that exclusively depend on the shape regularity of $\mathcal{T}$ and on the polynomial degree $k \in \mathbb{N}_0$ such that*

$$C_{\mathrm{rel}}^{-1}|||\Lambda \circ (1 - J_h I_h I_{\mathrm{M}})|||_* \leqslant \mu(\mathcal{T}) \leqslant C_{\mathrm{eff}}|||\Lambda|||_*.$$

Proof. The discussion in Subsection 7.2 applies to $\mathrm{Res} := F := \Lambda$ and $u_h := 0$ with $\mathrm{apx}(F,\mathcal{T}) = 0$. In this particular case, Proposition 7.1 provides the first inequality

$$|||\Lambda \circ (1 - J_h I_h I_{\mathrm{M}})|||_* \leqslant C_{\mathrm{rel}}\mu(\mathcal{T})$$

with $C_{\mathrm{rel}} = C_4$. Let $u \in H_0^2(\Omega)$ denote the Riesz representation of $a(u,\cdot) = \Lambda \in H^{-2}(\Omega)$ with the isometry $|||\Lambda|||_* = |||u|||$ in the Hilbert space $(H_0^2(\Omega), a)$ and $|||\cdot||| \equiv a(\cdot,\cdot)^{1/2}$. Then the efficiency estimate

$$\mu(\mathcal{T}) \leqslant C_{\mathrm{eff}} |||\Lambda|||_*$$

follows from Proposition 7.2 with $C_{\mathrm{eff}} = C_8$. □

Theorem A.1 allows for a direct application to the linearization of semilinear problems in [18]. It can be further generalized in various directions, e.g., in the spirit of Section 7 that considers the a posteriori error analysis of the linear biharmonic problem for a more general class of functionals in $H^{-2}(\Omega)$ including line and point loads. The reliability requires only piecewise smoothness of $\Lambda_0, \Lambda_1, \Lambda_2$ so that the traces and derivatives in μ_1, μ_2, μ_3 exist, while the efficiency may require extra oscillation terms (as in (7.5)).