ESAIM: M2AN 56 (2022) 41–78
https://doi.org/10.1051/m2an/2021085

ESAIM: Mathematical Modelling and Numerical Analysis
www.esaim-m2an.org

LOWEST-ORDER EQUIVALENT NONSTANDARD FINITE ELEMENT METHODS FOR BIHARMONIC PLATES

Carsten Carstensen[1,2] and Neela Nataraj[2,*]

Abstract. The popular (piecewise) quadratic schemes for the biharmonic equation based on triangles are the nonconforming Morley finite element, the discontinuous Galerkin, the C^0 interior penalty, and the WOPSIP schemes. Those methods are modified in their right-hand side $F \in H^{-2}(\Omega)$ replaced by $F \circ (JI_{\mathrm{M}})$ and then are quasi-optimal in their respective discrete norms. The smoother JI_{M} is defined for a piecewise smooth input function by a (generalized) Morley interpolation I_{M} followed by a companion operator J. An abstract framework for the error analysis in the energy, weaker and piecewise Sobolev norms for the schemes is outlined and applied to the biharmonic equation. Three errors are also equivalent in some particular discrete norm from [Carstensen, Gallistl, Nataraj, *ESAIM: M2AN* **49** (2015) 977–990.] without data oscillations. This paper extends the work [Veeser and Zanotti, *SIAM J. Numer. Anal.* **56** (2018) 1621–1642] to the discontinuous Galerkin scheme and adds error estimates in weaker and piecewise Sobolev norms.

Mathematics Subject Classification. 65N30, 65N12, 65N50.

Received August 19, 2021. Accepted December 14, 2021.

1. Introduction

The paper contributes to lower-order nonstandard finite element methods for a biharmonic plate problem in a real Hilbert space (V, a). Given $F \in L^2(\Omega)$, [23] compares the errors for nonstandard finite element methods (FEM) of the clamped biharmonic plate problem based on piecewise quadratic polynomials, namely the nonconforming Morley FEM [26], the symmetric interior penalty discontinuous Galerkin FEM (dGFEM) [28], and the C^0 interior penalty method (C^0IP) [6], with respective solutions u_{M}, u_h, and u_{IP}; Table 1 displays details of the respective schemes. For $F \in L^2(\Omega)$, dGFEM and hp-dGFEM for biharmonic and fourth-order problems, are extensively studied in [27, 28, 30, 31, 36, 37, 39].

For a general right-hand side $F \in H^{-2}(\Omega)$, the standard right-hand side $F(v_h)$ remains undefined for nonstandard finite element methods. A postprocessing procedure in [6] enables to introduce a new C^0IP method for right-hand sides in $H^{-2}(\Omega)$. In [41–43], the discrete test functions are transformed into conforming functions (J is called smoother in those works) before applying the load functional and quasi-optimal energy norm

Keywords and phrases. Biharmonic problem, best-approximation, *a priori* error estimates, companion operator, C^0 interior penalty, discontinuous Galerkin method, WOPSIP, Morley, comparison.

[1] Department of Mathematics, Humboldt-Universität zu Berlin, 10099 Berlin, Germany.
[2] Department of Mathematics, Indian Institute of Technology Bombay, Powai, Mumbai 400076, India.
*Corresponding author: neela@math.iitb.ac.in; nataraj.neela@gmail.com

© The authors. Published by EDP Sciences, SMAI 2022

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

TABLE 1. Overview of notation for four discrete schemes with discrete bilinear form (5.12)

	Schemes			
Notation	Morley FEM	dGFEM	C^0IP	WOPSIP
Section reference	Section 3.2	Section 7	Section 8	Section 10
V_h	$\mathrm{M}(\mathcal{T})$	$P_2(\mathcal{T})$	$S_0^2(\mathcal{T})$	$P_2(\mathcal{T})$
A_h	a_{pw} in (3.3)	A_{dG} in (7.1)	A_{IP} in (8.1)	A_{P} in (10.2)
b_h	0	$-\Theta\,\mathcal{J} - \mathcal{J}^*$ in (6.2)	$-\Theta\,\mathcal{J} - \mathcal{J}^*$ in (6.2)	0
c_h	0	c_{dG} in (4.3)	c_{IP} in (8.2)	c_{P} in (10.1)
$I_h : \mathrm{M}(\mathcal{T}) \to V_h$	id	id	I_{C} in (8.4)	id

estimates

$$\|u - u_{\mathrm{M}}\|_{\mathrm{pw}} \approx \min_{v_{\mathrm{M}} \in \mathrm{M}(\mathcal{T})} \|u - v_{\mathrm{M}}\|_{\mathrm{pw}}, \ \|u - u_{\mathrm{IP}}\|_{\mathrm{IP}} \approx \min_{v_{\mathrm{IP}} \in \mathrm{IP}(\mathcal{T})} \|u - v_{\mathrm{IP}}\|_{\mathrm{IP}}$$

are derived for the Morley FEM and C^0IP method.

The papers [41–43] discuss minimal conditions on a smoother for each problem, while this paper presents one smoother JI_{M} for *all* schemes; the best-approximation for the dGFEM is a new result. The smoother JI_{M} also allows a post-processing with *a priori* error estimates in weaker and piecewise Sobolev norms.

Table 1 summarizes the notation of spaces, bilinear forms, and an operator for the four second-order methods for the biharmonic problem detailed in Section 3.2 and in Sections 7, 8, and 10.

Contributions

The main contributions of this paper are

(a) the design and analysis of a *generalized Morley interpolation* operator I_{M} for piecewise smooth functions in $H^2(\mathcal{T})$,
(b) the design of modified schemes for Morley FEM, dGFEM, C^0IP method, and a weakly over-penalized symmetric interior penalty (WOPSIP) method for the biharmonic problem for data in $F \in H^{-2}(\Omega)$,
(c) an abstract framework for the best-approximation property and weaker (piecewise) Sobolev norm estimates,
(d) *a priori* error estimates in (piecewise) Sobolev norms for the lowest-order nonstandard finite element methods for biharmonic plates,
(e) an extension of the results of [23] to an equivalence

$$\|u - u_{\mathrm{M}}\|_h \approx \|u - u_{\mathrm{IP}}\|_h \approx \|u - u_{\mathrm{dG}}\|_h \approx \big\|(1 - \Pi_0) D^2 u\big\|_{L^2(\Omega)} \tag{1.1}$$

without data oscillations for the modified schemes,
(f) the proof of the best approximation for the modified dGFEM that extends [42] and Theorem 4.3 of [23].

Remark 1.1 (Medius analysis)**.** The quasi-optimality of nonconforming and discontinuous Galerkin methods was established in the seminal paper [34] for the original method up to data oscillations for $F \in L^2(\Omega)$. Arguments from *a posteriori* error analysis [44] give new insight in the consistency term from the Strang–Fix lemmas. The techniques in this paper circumvent any *a posteriori* error analysis and take advantage of the extra benefits of the companion operator J.

Remark 1.2 (Smoother)**.** The fundamental series of contributions [41–43] on the quasi-optimality concerns best-approximation for a modified scheme with $F_h := F \circ J$ for a smoother J. Elementary algebra indicates a key identity (of (5.5) below) that is already mentioned in (5.15) of [6] and makes the source term in the consistency disappear.

Remark 1.3 (Extension to higher order)**.** A (general) Morley interpolation allows for a simultaneous analysis of four lowest-order schemes, but appears to be restricted to piecewise quadratics at first glance. But the combination of JI_{M} as an averaging smoother with higher-order bubble-smoothers shall enable applications to higher-order schemes as indicated in [42].

Remark 1.4 (Extension to 3D)**.** Although the plate problem is intrinsic two-dimensional, there are three-dimensional Morley finite elements with a recent companion operator [18] that guarantees the fundamental properties in 3D such that the abstract framework applies.

Organization

The remaining parts of this paper are organised as follows. Section 2 provides an abstract characterization of the best-approximation property that applies to the various applications considered in this and future papers [25]. Section 3 presents preliminaries, a nonconforming discretisation, introduces a *novel* generalized Morley interpolation operator for discontinuous functions, and states a best-approximation [16] result for nonconforming discretisations with data $F \in H^{-2}(\Omega)$. Section 4 proves a crucial equivalence result of two discrete norms for a piecewise $H^2(\mathcal{T})$ function, and proves approximation properties for the generalized Morley interpolation operator. Section 5 provides a framework for dG methods and the proof of a best-approximation property under a set of general assumptions. Section 6 develops the abstract result for *a priori* error estimates in weaker and piecewise Sobolev norms. Sections 7 and 8 recall the dG and C^0IP schemes and verify the assumptions of Sections 5 and 6 for the best-approximation result in the energy norm as well as weaker and piecewise Sobolev norms without data oscillations. The paper concludes with the equivalence (1.1) of errors in Section 9 and a proof of quasi-optimality up to penalty for the WOPSIP scheme in Section 10.

General notation

Standard notation of Lebesgue and Sobolev spaces, their norms, and L^2 scalar products applies throughout the paper such as the abbreviation $\|\bullet\|$ for $\|\bullet\|_{L^2(\Omega)}$. For real s, $H^s(\Omega)$ denotes the Sobolev space associated with the Sobolev–Slobodeckii semi-norm $|\bullet|_{H^s(\Omega)}$ [33]; $H^s(T) := H^s(\mathrm{int}(T))$ abbreviates the Sobolev space with respect to the interior $\mathrm{int}(T) \neq \emptyset$ of a (compact) triangle T. The closure of $D(\Omega)$ in $H^s(\Omega)$ is denoted by $H^s_0(\Omega)$ and $H^{-s}(\Omega)$ is the dual of $H^s_0(\Omega)$. The triple norm $|||\bullet||| := |\bullet|_{H^2(\Omega)}$ is the energy norm and $|||\bullet|||_{\mathrm{pw}} := |\bullet|_{H^2(\mathcal{T})} := \|D^2_{\mathrm{pw}}\bullet\|$ is its piecewise version with the piecewise Hessian D^2_{pw}. Given any function $v \in L^2(\omega)$, define the integral mean $\fint_\omega v\,\mathrm{dx} := 1/|\omega| \int_\omega v\,\mathrm{dx}$; where $|\omega|$ denotes the area of ω. The notation $A \lesssim B$ abbreviates $A \leq CB$ for some positive generic constant C, which depends only on Ω and the shape regulatity of $\mathcal{T}$; $A \approx B$ abbreviates $A \lesssim B \lesssim A$.

2. Prologue

This section characterizes the best-approximation property of a class of non-conforming finite element methods. The biharmonic problem is put in an abstract framework in real Hilbert spaces X and Y and a bounded bilinear form $a : X \times Y \to \mathbb{R}$ satisfying an inf-sup condition. Given a right-hand side $F \in Y^*$, the *exact problem* seeks $x \in X$ with

$$a(x, \bullet) = F \quad \text{in } Y.$$

The *discrete problem* is put in an analog framework with finite-dimensional real Hilbert spaces X_h and Y_h and a bilinear form $a_h : X_h \times Y_h \to \mathbb{R}$. The discrete space X_h (resp. Y_h) is *not* a subspace of X (resp. Y) in general, but X and X_h (resp. Y and Y_h) belong to one common bigger vector space that gives rise to the sum $\widehat{X} = X + X_h$ (resp. $\widehat{Y} = Y + Y_h$). It is not supposed that this is a direct sum, so the intersection $X \cap X_h$ (resp. $Y \cap Y_h$) may be non-trivial. We suppose that $\widehat{X}$ and $\widehat{Y}$ are real Hilbert spaces with (complete) subspaces X, X_h and Y, Y_h. The linear and bounded map $Q \in L(Y_h; Y)$ links the right-hand side $F \in Y^*$ of the exact problem to the right-hand side $F_h := F \circ Q \in Y_h^*$ of the discrete problem. The map Q is called smoother in [41–43] because

it maps a (possibly) discontinuous function $y_h \in Y_h$ to a smooth function Qy_h in applications. The resulting *discrete problem* seeks $x_h \in X_h$ with

$$a_h(x_h, \bullet) = F(Q\bullet) \text{ in } Y_h. \tag{2.1}$$

We also suppose that the exact and discrete problems are well-posed and this means in particular that $\dim X_h = \dim Y_h < \infty$ and that the bounded bilinear forms a and a_h satisfy inf-sup conditions with positive constants α and α_h and are non-degenerate such that the associated linear and bounded operators $A \in L(X; Y^*)$ and $A_h \in L(X_h; Y_h^*)$ are bijective; the associated linear operators are defined by $Ax := a(x, \bullet) \in Y^*$ for all $x \in X$ and by $A_h x_h := a_h(x_h, \bullet) \in Y_h^*$ for all $x_h \in X_h$.

The general discussion in [41–43], leads to an optimal smoothing $Q = \Pi_Y|_{Y_h}$ for the orthogonal projection $\Pi_Y \in L(\widehat{Y})$ onto Y. This is a global operation in general and hence infeasible for practical computations. Notice carefully that all examples in [41–43] discuss $Q \in L(Y_h; Y)$ with $Qz = z$ for all $z \in Y_h \cap Y$, abbreviated by

$$Q = \text{id in } Y_h \cap Y. \tag{2.2}$$

This paper introduces a smoother JI_{M} for the discontinuous Galerkin schemes that satisfies (2.2) and is quasi-optimal in the following sense with a constant $\Lambda_{\mathrm{Q}} \geq 0$ that is exclusively bounded in terms of the shape regularity of the underlying triangulations.

Definition 2.1 (Quasi-optimal smoother)**.** A linear bounded operator $Q \in L(Y_h; Y)$ is called a quasi-optimal smoother if there exists some $\Lambda_{\mathrm{Q}} \geq 0$ such that

$$\|y_h - Qy_h\|_{\widehat{Y}} \leq \Lambda_{\mathrm{Q}} \|y_h - y\|_{\widehat{Y}} \text{ for all } \; y_h \in Y_h \text{ and all } \; y \in Y. \tag{2.3}$$

The proofs of Lemma 2.2 and 2.4 below rely on compactness arguments whence in the Appendix A, the constants Λ_{P} and Λ_{Q} depend on the discrete space. *The point is that this paper designs a smoother in Section 4.3 with a constant that does not depend on the mesh-size.*

Lemma 2.2. *The operator $Q \in L(Y_h; Y)$ is a quasi-optimal smoother if and only if (2.2) holds.*

The (nonconforming) finite element method is characterized by the operator $M \in L(X; X_h)$ that maps any $x \in X$ to a right-hand side $F := a(x, \bullet) \in Y^*$ and then to the solution $Mx := x_h = A_h^{-1}(Q^*F)$ to (2.1), *i.e.*, $Mx := A_h^{-1}(a(x, Q\bullet)) \in X_h$ for all $x \in X$, or, in operator form,

$$M := A_h^{-1} Q^* A \in L(X; X_h).$$

In other words, the subsequent diagram commutes.

$$\begin{tikzcd}
X \arrow[r, "A"] \arrow[d, "M"] & Y^* \arrow[d, "Q^*"] \\
X_h \arrow[u, dashed, "P"] & Y_h^* \arrow[l, "A_h^{-1}"']
\end{tikzcd}$$

This diagram also depicts some (linear and bounded) operator $P : X_h \to X$ that will become a quasi-optimal smoother in the context of the best-approximation property of M below. A synonym to the best-approximation property of M is to say M is quasi-opimal in the following sense.

Definition 2.3 (Quasi-optimal)**.** The above operator M is said to be quasi-optimal if

$$\textbf{(QO)} \quad \exists\, C_{\rm qo} > 0 \;\; \forall x \in X \quad \|x - Mx\|_{\widehat{X}} \le C_{\rm qo} \min_{x_h \in X_h} \|x - x_h\|_{\widehat{X}}.$$

A first characterisation of **(QO)** has been given in [41] in terms of

$$M = \mathrm{id} \text{ in } X_h \cap X. \tag{2.4}$$

Lemma 2.4 ([41])**.** *Under the present notation,* **(QO)** *is equivalent to* (2.4)*.*

The above lemmas characterize $C_{\rm qo}$ and $\Lambda_{\rm Q}$ by a compactness argument and it remains to control $C_{\rm qo}$ and $\Lambda_{\rm Q}$ in terms of mesh-size independent bounds in applications. This paper designs in Section 4.3 a smoother in the spirit of [41–43] based on earlier work in the context of *a posteriori* error control [9, 12, 22] and adaptive mesh-refinement [14, 17–19]. The outcome is a quasi-optimal smoother $P \in L(X_h; X)$ with a constant $\Lambda_{\rm P}$ that depends only on the shape regularity of the underlying finite element mesh and

$$\|x_h - Px_h\|_{\widehat{X}} \le \Lambda_{\rm P} \|x_h - x\|_{\widehat{X}} \text{ for all } \; x_h \in X_h \text{ and all } \; x \in X. \tag{2.5}$$

The proof of the following characterization of best-approximation shall be given in the Appendix A.

Theorem 2.5. *Suppose $P \in L(X_h; X)$ and $\Lambda_{\rm P}$ satisfy* (2.5)*. Then* **(QO)** *is equivalent to the existence of $\Lambda_{\rm H} > 0$ with*

$$\textbf{(H)} \quad a_h(x_h, y_h) - a(Px_h, Qy_h) \le \Lambda_{\rm H} \|x_h - Px_h\|_{\widehat{X}} \|y_h\|_{Y_h} \text{ for all } \; x_h \in X_h \text{ and } \; y_h \in Y_h.$$

In particular, if **(H)** holds, then **(QO)** follows with a constant $C_{\rm qo}$ that depends solely on α_h, $\Lambda_{\rm H}$, $\Lambda_{\rm P}$, $\|Q\|$, and $\|A\|$.

The next theorem presents a key estimate that is crucial for goal-oriented error control and duality arguments for weaker norm estimates. The proof and the dependence of contants are presented in the Appendix A. The motivation for $\widehat{\textbf{(QO)}}$ is exemplified in Theorem 2.7 below.

Theorem 2.6. *Suppose P and Q are quasi-optimal smoothers with* (2.3)–(2.5) *and suppose* **(QO)***. Then the existence of $\widehat{C_{\rm qo}} > 0$ with*

$$\widehat{\textbf{(QO)}} \;\; a(x - PMx, y) \le \widehat{C_{\rm qo}} \|x - Mx\|_{\widehat{X}} \|y - y_h\|_{\widehat{Y}} \text{ for all } \; x \in X, \; y \in Y, \text{ and } \; y_h \in Y_h$$

is equivalent to the existence of $\widehat{\Lambda_{\rm H}} > 0$ with

$$\widehat{\textbf{(H)}} \; a_h(x_h', y_h) - a(Px_h', Qy_h) \le \widehat{\Lambda_{\rm H}} \|x_h' - Px_h'\|_{\widehat{X}} \|y_h - Qy_h\|_{\widehat{Y}} \text{ for all } \; x_h' \in X_h', \text{ and } \; y_h \in Y_h.$$

In particular, if $\widehat{\textbf{(H)}}$ holds, $\widehat{\textbf{(QO)}}$ follows with a constant $\widehat{C_{\rm qo}}$ that depends solely on $\|a\|$, Λ_2', $\Lambda_{\rm P}$, and $\Lambda_{\rm Q}$.

The *a priori* error estimates in weaker Sololev norms (weaker than the energy norm) are a corollary of Theorem 2.6 and the elliptic regularity, the latter is written in an abstract form by the assumption that X_s and Y_s are two Hilbert spaces with $X \subset X_s$ and $Y_s \subset Y$ such that

$$\textbf{(R)} \;\; \exists C_{\rm reg} > 0 \; \forall F \in X_s^* \;\; \|A^{-*}F\|_{Y_s} \le C_{\rm reg} \|F\|_{X_s^*}$$

for the solution $y := A^{-*}F \in Y_s \subset Y$ to $a(\bullet, y) = F \in X_s^* \subset X^*$.

Theorem 2.7 (Weak *a priori*)**.** *Under the assumptions of Theorem 2.6,* $\widehat{\textbf{(QO)}}$ *and* **(R)** *imply*

$$\|x - PMx\|_{X_s} \le \widehat{C_{\rm qo}} \|x - Mx\|_{\widehat{X}} \sup_{\substack{y \in Y_s \\ \|y\|_{Y_s} \le C_{\rm reg}}} \inf_{y_h \in Y_h} \|y - y_h\|_{\widehat{Y}} \text{ for all } x \in X.$$

Proof. Given $x - PMx \in X \subset X_s$, a corollary of the Hahn–Banach extension theorem leads to some $F \in X_s^* \subset X^*$ with norm $\|F\|_{X_s^*} \le 1$ in X_s^* and $\|x - PMx\|_{X_s} = F(x - PMx)$. The dual solution $y \in Y$ to $F = a(\bullet, y) \in X^*$ satisfies **(R)** and $\widehat{\textbf{(QO)}}$ leads to

$$\|x - PMx\|_{X_s} = a(x - PMx, y) \le \widehat{C_{\rm qo}} \|x - Mx\|_{\widehat{X}} \|y - y_h\|_{\widehat{Y}}$$

for any $y_h \in Y_h$. This and $\|y\|_{Y_s} \le C_{\rm reg} \|F\|_{X_s^*} \le C_{\rm reg}$ conclude the proof. □

Example 2.8 (Standard)**.** For the m-harmonic operator $A = (-1)^m \Delta^m$ and $X = H_0^m(\Omega) = Y$, **(R)** holds for $X_s = H_0^{m-s}(\Omega)$, $Y_s = H^{m+s}(\Omega)$ and $1/2 \le s \le 1$, $m = 1$ or 2. Typical first-order approximation properties of the discrete finite element spaces result in

$$\sup_{\substack{y \in Y_s \\ \|y\|_{Y_s} \le C_{\rm reg}}} \inf_{y_h \in Y_h} \|y - y_h\|_{\widehat{Y}} = O(h_{\max}^s)$$

in terms of the maximal mesh-size $h_{\max}$ of the underlying finite element mesh.

Remark 2.9 (Best-approximation constant)**.** The paper [41] gives a formula for the best-approximation constant $C_{\rm qo}$ for some slightly simpler problem in one Hilbert space.

Remark 2.10 (Injective smoother)**.** Under the above notation $Q \in L(Y_h; Y)$ is injective if and only if $M \in L(X; X_h)$ is surjective [41]. Then there exists a right-inverse $P \in L(X_h; X)$ to M and **(H)** holds with $\Lambda_{\rm H} = 0$ (this follows with the arguments of the proof of Theorem 2.7 for P' that is in fact a quasi-optimal smoother owing to (A.5)). Consequently, the discrete scheme is equivalent to a conforming Petrov–Galerkin scheme.

Remark 2.11 (Non injective smoother)**.** In case $Q \in L(Y_h; Y)$ is *not* injective, the discrete problem may reduced to the range $X_h' := \mathcal{R}(M)$ of M and the orthogonal complement Y_h' of the kernel of Q in Y_h. However, the explicit computation of the reduced discrete spaces X_h' and Y_h' may be costly and hence this paper outlines a general analysis that allows non-injective quasi-optimal smoothers.

Example 2.12 (Smoother for Morley)**.** For the standard Morley interpolation operator $I_{\rm M} : H_0^2(\Omega) \to {\rm M}(\mathcal{T})$ and a companion operator $J : {\rm M}(\mathcal{T}) \to H_0^2(\Omega)$ (*cf.*, Lem. 3.7 below for details) the smoother $Q = JI_{\rm M}$ is injective because J is a right-inverse of $I_{\rm M}$.

Example 2.13 (Smoother for dG)**.** This paper advertises a smoother $Q := JI_{\rm M}$ for a (generalized) Morley interpolation $I_{\rm M} : (P_2(\mathcal{T}) + H_0^2(\Omega)) \to {\rm M}(\mathcal{T})$ (*cf.*, (3.5) below for details) followed by a companion operator J from the previous example for the dG FEM. Then $\dim P_2(\mathcal{T}) = 6|\mathcal{T}|$ is strictly larger than $\dim {\rm M}(\mathcal{T}) = |\mathcal{V}(\Omega)| + |\mathcal{E}(\Omega)|$; whence Q cannot be injective.

The situation for the C^0 IP with the discrete space $S_0^2(\mathcal{T})$ (of the same dimension as ${\rm M}(\mathcal{T})$) is more involved and is discussed in more details in Section 8 below.

3. Preliminaries

3.1. Continuous model problem

Suppose $u \in V := H_0^2(\Omega)$ solves the biharmonic equation $\Delta^2 u = F$ for a given right-hand side $F \in V^* \equiv H^{-2}(\Omega)$ in a planar bounded Lipschitz domain Ω with polygonal boundary $\partial\Omega$. The weak form of this equation reads

$$a(u, v) = F(v) \quad \text{for all } v \in V \tag{3.1}$$

with the scalar product $a(v, w) := \int_\Omega D^2 v : D^2 w \, {\rm dx}$ for all $v, w \in V$. It is well known that (3.1) has a unique solution u and elliptic regularity [1, 3, 32, 38] holds in the sense that $F \in H^{-s}(\Omega)$ implies $u \in V \cap H^{4-s}(\Omega)$ for

all s with $2-\sigma_{\rm reg}\le s\le 2$ with the index of elliptic regularity $\sigma_{\rm reg}>0$. The lowest-order nonconforming finite element schemes suggest a linear convergence rate in the energy norm for a solution $u\in H^t(\Omega)$ at most for all $t\ge 3$. Therefore $\sigma:=\min\{1,\sigma_{\rm reg}\}$ is fixed throughout this paper and exclusively depends on Ω. The regularity is frequently employed in the following formulation.

Example 3.1 (Regularity)**.** There exists a constant $0<\sigma\le 1$ such that $F\in H^{-s}(\Omega)$ with $2-\sigma\le s\le 2$ satisfies $u\in V\cap H^{4-s}(\Omega)$ and

$$\|u\|_{H^{4-s}(\Omega)}\le C_{\rm reg}(s)\|F\|_{H^{-s}(\Omega)} \tag{3.2}$$

for some constant $C_{\rm reg}(s)<\infty$, which depends on Ω and s. (The dependence on s results from the equivalence of Sobolev norms that may depend on the index s in general.)

It is true that pure Dirichlet boundary conditions in the model example lead to $\sigma>1/2$ and then allow for a control of the traces D^2u in the jump terms. This paper circumvents this argument and all the results hold for $\sigma\ge 0$. The new discrete analysis is therefore much more flexible and allows for generalizations of the model problem *e.g.*, for mixed and boundary conditions of less smoothness.

3.2. Nonconforming discretisation

Throughout the rest of this article, the following notations are adopted. Let $\mathcal{T}$ denote a shape regular triangulation of the polygonal Lipschitz domain into compact triangles. Associate its piecewise constant mesh-size $h_{\mathcal{T}}\in P_0(\mathcal{T})$ with $h_T:=h_{\mathcal{T}}|_T:=\operatorname{diam}(T)\approx|T|^{1/2}$ in any triangle $T\in\mathcal{T}$ of area $|T|$ and its maximal mesh-size $h_{\max}:=\max h_{\mathcal{T}}$. Let $\mathcal{V}$ (resp. $\mathcal{V}(\Omega)$ or $\mathcal{V}(\partial\Omega)$) denote the set of all (resp. interior or boundary) vertices in $\mathcal{T}$. Let $\mathcal{E}$ (resp. $\mathcal{E}(\Omega)$ or $\mathcal{E}(\partial\Omega)$) denote the set of all (resp. interior or boundary) edges. The length of an edge E is denoted by h_E. Let Π_k denote the $L^2(\Omega)$ orthogonal projection onto the piecewise polynomials $P_k(\mathcal{T}):=\{v\in L^2(\Omega):\forall T\in\mathcal{T},\ v|_T\in P_k(T)\}$ of degree at most $k\in\mathbb{N}_0$. Let the Hilbert space $H^1(\mathcal{T})\equiv\prod_{T\in\mathcal{T}}H^1(T)$. Define the jump $[\varphi]_E:=\varphi|_{T_+}-\varphi|_{T_-}$ and the average $\langle\varphi\rangle_E:=\frac12\big(\varphi|_{T_+}+\varphi|_{T_-}\big)$ across the interior edge $E=\partial T_+\cap\partial T_-\in\mathcal{E}(\Omega)$ of $\varphi\in H^1(\mathcal{T})$ of the adjacent triangles T_+ and $T_-\in\mathcal{T}$ in an order such that the unit normal vector $\nu_{T_+}|_E=\nu_E=-\nu_{T_-}|_E$ along the edge E has a fixed orientation and points outside T_+ and inside T_-; ν_T is the outward unit normal of T along ∂T. The edge-patch $\omega(E):=\operatorname{int}(T_+\cup T_-)$ of the interior edge $E=\partial T_+\cap\partial T_-\in\mathcal{E}(\Omega)$ is the interior of union $T_+\cup T_-$ of the neighboring triangles T_+ and T_-. Extend the definition of the jump and the average to an edge $E\in\mathcal{E}(\partial\Omega)$ on the boundary by $[\varphi]_E:=\varphi|_E$ and $\langle\varphi\rangle_E:=\varphi|_E$ owing to the homogeneous boundary conditions. Jump and average are understood componentwise for any vector function. The edge-patch $\omega(E):=\operatorname{int}(T_+)$ of an edge $E\in\mathcal{E}(\partial\Omega)$ on the boundary is simply the interior of the one triangle T_+ with the edge E in the triangulation $\mathcal{T}$.

The nonconforming *Morley* finite element space [26] reads

$$\mathrm{M}'(\mathcal{T}):=\left\{v_{\rm M}\in P_2(\mathcal{T})\,\middle|\,\begin{array}{l}v_{\rm M}\text{ is continuous at the vertices and its normal derivatives}\\ \nu_E\cdot\nabla_{\rm pw}v_{\rm M}\text{ are continuous at the midpoints of interior edges}\end{array}\right\},$$
$$\mathrm{M}(\mathcal{T}):=\left\{v_{\rm M}\in \mathrm{M}'(\mathcal{T})\,\middle|\,\begin{array}{l}v_{\rm M}\text{ vanishes at the vertices of }\partial\Omega\text{ and its normal derivatives}\\ \nu_E\cdot\nabla_{\rm pw}v_{\rm M}\text{ vanish at the midpoints of boundary edges}\end{array}\right\}.$$

Figure 1a depicts the degrees of freedom of the Morley finite element

$$\left(T,P_2(T),(\delta_z:z\in\mathcal{V}(T))\cup\left(\fint_E\partial_{\nu_E}\bullet\,\mathrm{ds}:E\in\mathcal{E}(T)\right)\right)$$

(in the sense of Ciarlet) in the triangle T with set of vertices $\mathcal{V}(T)$ and set of edges $\mathcal{E}(T)$.

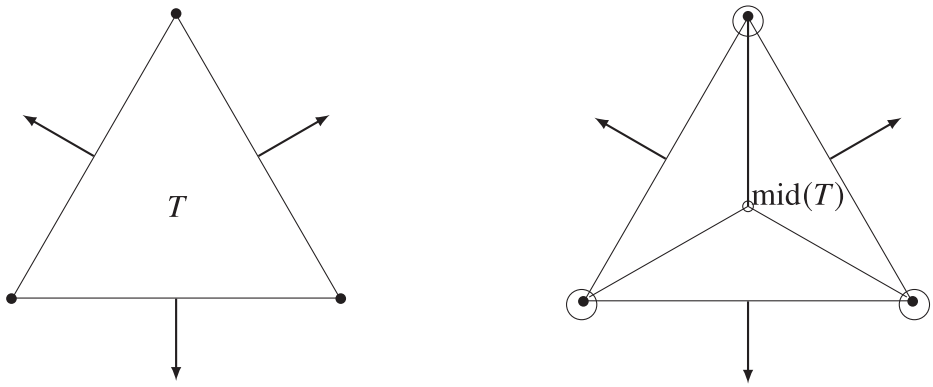

FIGURE 1. (a) Morley (*left*) and (b) HCT (*right*) finite element.

The semi-scalar product $a_{\rm pw}$ is defined by the piecewise differential operator $D^2_{\rm pw}$ and

$$a_{\rm pw}(v_{\rm pw}, w_{\rm pw}) := \sum_{T\in\mathcal{T}} \int_T D^2 v_{\rm pw} : D^2 w_{\rm pw}\,{\rm dx} \quad \text{for all } v_{\rm pw}, w_{\rm pw} \in H^2(\mathcal{T}). \tag{3.3}$$

It induces a piecewise H^2 seminorm $\|\bullet\|_{\rm pw} = a_{\rm pw}(\bullet,\bullet)^{1/2}$ that is also a norm in $\mathrm{M}(\mathcal{T})$. Then $(\mathrm{M}(\mathcal{T}), a_{\rm pw})$ is a (finite-dimensional) Hilbert space so that, given any $F_h \in \mathrm{M}(\mathcal{T})^*$, there exists a unique discrete solution $u_{\rm M} \in \mathrm{M}(\mathcal{T})$ to

$$a_{\rm pw}(u_{\rm M}, v_{\rm M}) = F_h(v_{\rm M}) \quad \text{for all } v_{\rm M} \in \mathrm{M}(\mathcal{T}). \tag{3.4}$$

3.3. Interpolation of discontinuous functions

Lemma 3.2 (Interpolation estimates I [8, 10])**.** *The Morley interpolation operator $I_{\rm M} : V \to \mathrm{M}(\mathcal{T})$ is defined by $(I_{\rm M}v)(z) = v(z)$ and $\fint_E \frac{\partial I_{\rm M} v}{\partial \nu_E}\,{\rm ds} = \fint_E \frac{\partial v}{\partial \nu_E}\,{\rm ds}$ for any $z \in \mathcal{V}(\Omega)$ and $E \in \mathcal{E}(\Omega)$. It satisfies*

(a) *the integral mean property of the Hessian, $D^2_{\rm pw} I_{\rm M} = \Pi_0 D^2$,*
(b) $\sum_{m=0}^{2} h_T^{m-2} |v - I_{\rm M} v|_{H^m(T)} \le 2\|(1-\Pi_0)D^2 v\|_{L^2(T)}$ *for all $v \in H^2(T)$ and any $T \in \mathcal{T}$,*
(c) $\|v - I_{\rm M} v\|_{\rm pw} \lesssim h_{\max}^s \|v\|_{H^{2+s}(\Omega)}$ *for all $v \in H^{2+s}(\Omega)$ and all $0 \le s \le 1$.*

A reformulation of Lemma 3.2a is the best-approximation property

$$a_{\rm pw}(v - I_{\rm M} v, w_2) = 0 \quad \text{for all } v \in V \text{ and all } w_2 \in P_2(\mathcal{T}). \tag{3.5}$$

A reformulation of Lemma 3.2b is the existence of a universal constant $\kappa > 0$ with

$$\|h_{\mathcal{T}}^{-2}(v - I_{\rm M} v)\| \le \kappa \|v - I_{\rm M} v\|_{\rm pw} \quad \text{for all } v \in V. \tag{3.6}$$

(In fact $\kappa = 0.25745784465$ from [10] is independent of the shape of the triangle T.)

Remark 3.3 (Pythagoras)**.** The functions $v_{\rm pw}, w_{\rm pw} \in H^2(\mathcal{T}) := \{v_{\rm pw} \in L^2(\Omega) : \forall T \in \mathcal{T},\ v_{\rm pw}|_T \in H^2(T)\}$ are orthogonal iff $a_{\rm pw}(v_{\rm pw}, w_{\rm pw}) = 0$ holds and then the Pythagoras theorem leads to

$$\|v - w_{\rm M}\|^2_{\rm pw} = \|v - I_{\rm M} v\|^2_{\rm pw} + \|w_{\rm M} - I_{\rm M} v\|^2_{\rm pw} \tag{3.7}$$

for all $v \in V$ and $w_{\rm M} \in \mathrm{M}(\mathcal{T})$. In particular, $\|v - I_{\rm M} v\|_{\rm pw} = \min_{v_2 \in P_2(T)} \|v - v_2\|_{\rm pw}$.

Definition 3.4 ((local) Morley interpolation)**.** Given any $T \in \mathcal{T}$ and $v_{\rm pw} \in H^2(T)$, the (local) Morley interpolation $I_{\rm M}^{\rm loc} v_{\rm pw}|_T \in P_2(T)$ is defined by the degrees of freedom of the Morley finite element such that, for all $z \in \mathcal{V}(T)$ and for all $E \in \mathcal{E}(T)$,

$$\big(I_{\rm M}^{\rm loc} v_{\rm pw} - v_{\rm pw}\big)|_T(z) = 0 \quad \text{and} \quad \fint_E \big(\partial\big(I_{\rm M}^{\rm loc} v_{\rm pw} - v_{\rm pw}\big)|_T/\partial\nu_E\big)\,{\rm ds} = 0.$$

The Morley interpolation allows for an extension (still denoted by $I_{\rm M}$) to piecewise H^2 functions in $H^2(\mathcal{T}) \equiv \prod_{T\in\mathcal{T}} H^2(T)$ by averaging the degrees of freedom.

Definition 3.5 (Morley interpolation)**.** Given any $v_{\rm pw} \in H^2(\mathcal{T})$, define $I_{\rm M} v_{\rm pw} := v_{\rm M} \in {\rm M}(\mathcal{T})$ by the degrees of freedom as follows. For any interior vertex $z \in \mathcal{V}(\Omega)$ with set of attached triangles $\mathcal{T}(z)$ that has cardinality $|\mathcal{T}(z)| \in \mathbb{N}$ and any interior edge $E = \partial T_+ \cap \partial T_- \in \mathcal{E}(\Omega)$ and its mean value operator $\langle\bullet\rangle_E$ (the arithmetic mean of the two traces from the triangles T_+ and $T_- \in \mathcal{T}$ along their common edge $E = \partial T_+ \cap \partial T_-$), set

$$v_{\rm M}(z) := |\mathcal{T}(z)|^{-1} \sum_{T\in\mathcal{T}(z)} (v_{\rm pw}|_T)(z) \quad \text{and} \quad \fint_E \frac{\partial v_{\rm M}}{\partial \nu_E}\,{\rm ds} := \fint_E \left\langle \frac{\partial v_{\rm pw}}{\partial \nu_E} \right\rangle_E {\rm ds}.$$

(The remaining degrees of freedom at vertices and edges on the boundary are zero for homogeneous boundary conditions.)

Remark 3.6 (Standard Morley interpolation *vs.* Def. 3.5)**.** The interpolation operator $I_{\rm M}$ of Definition 3.5 extends that of standard Morley interpolation operator in the sense that the two definitions coincide for functions in $H_0^2(\Omega)$. This justifies the use of the same symbol $I_{\rm M}$.

3.4. Companion operator and best-approximation for the Morley FEM

A conforming finite-dimensional subspace of $H_0^2(\Omega)$ is provided by the Hsieh–Clough–Tocher (HCT) FEM ([26], Chap. 6). For any $T \in \mathcal{T}$, let $\mathcal{K}(T) := \{T_E : \ E \in \mathcal{E}(T)\}$ denote the triangulation of T into three sub-triangles $T_E := {\rm conv}\{E, {\rm mid}(T)\}$ with edges $E \in \mathcal{E}(T)$ and common vertex ${\rm mid}(T)$ depicted in Figure 1b. Then,

$$\mathrm{HCT}(\mathcal{T}) := \big\{v \in H_0^2(\Omega) : \ v|_T \in P_3(\mathcal{K}(T)) \text{ for all } T \in \mathcal{T}\big\}. \tag{3.8}$$

The degrees of freedom in a triangle $T \in \mathcal{T}$ are the nodal values $\psi(z)$ and its derivative $\nabla\psi(z)$ of the function $\psi \in \mathrm{HCT}(\mathcal{T})$ at any vertex $z \in \mathcal{V}(T)$ and the values $\partial\psi/\partial\nu_E({\rm mid}(E))$ of the normal derivatives at the midpoint ${\rm mid}(E)$ of any edge $E \in \mathcal{E}(T)$.

Lemma 3.7 (Right-inverse [17, 29, 41])**.** *There exists a linear map $J : {\rm M}(\mathcal{T}) \to (\mathrm{HCT}(\mathcal{T}) + P_8(\mathcal{T})) \cap H_0^2(\Omega)$ and a constant $\Lambda_{\rm J}$ (that exclusively depends on the shape regularity of $\mathcal{T}$) such that any $v_{\rm M} \in {\rm M}(\mathcal{T})$ satisfies (a)–(e).*

(a) $Jv_{\rm M}(z) = v_{\rm M}(z)$ *for any* $z \in \mathcal{V}$*;*
(b) $\nabla(Jv_{\rm M})(z) = |\mathcal{T}(z)|^{-1} \sum_{T\in\mathcal{T}(z)} (\nabla v_{\rm M}|_T)(z)$ *for* $z \in \mathcal{V}(\Omega)$*;*
(c) $\fint_E \partial Jv_{\rm M}/\partial\nu_E\,{\rm d}s = \fint_E \partial v_{\rm M}/\partial\nu_E\,{\rm d}s$ *for any* $E \in \mathcal{E}$*;*
(d) $\|v_{\rm M} - Jv_{\rm M}\|_{\rm pw} \le \Lambda_{\rm J} \min_{v\in V} \|v_{\rm M} - v\|_{\rm pw}$*;*
(e) $v_{\rm M} - Jv_{\rm M} \perp P_2(\mathcal{T})$ *in* $L^2(\Omega)$*.*

The operator J of Lemma 3.7 with (a)–(c) is a right-inverse for $I_{\rm M} : V \to {\rm M}(\mathcal{T})$, *i.e.*,

$$I_{\rm M} J = {\rm id} \text{ in } {\rm M}(\mathcal{T}). \tag{3.9}$$

Examples are provided in [17, 29, 41]. For earlier references in the literature, see [5, 13, 20]. A right-inverse with benefits like (d) and (e) is called companion operator and [17] defines $J : \mathrm{M}(\mathcal{T}) \to V$ so that (a)–(e) of Lemma 3.7 hold (*cf.*, in particular [17], Lem. 5.1 for the analysis of (d) and (e)).

Given $F \in L^2(\Omega)$, we may choose $F_h \equiv F$ in the discrete scheme (3.4); but otherwise $F_h = F \circ J$ is the option throughout this paper; other choices are proposed in [6, 42]. Given any Lebesgue function $F \equiv F_h \in L^2(\Omega)$ with its L^2 projection $\Pi_2 F$ onto $P_2(\mathcal{T})$, define its oscillations $\mathrm{osc}_2(F, \mathcal{T}) := \big\|h_{\mathcal{T}}^2(F - \Pi_2 F)\big\|$.

Theorem 3.8 (Best-approximation up to data approximation [23, 34])**.** *The constant* $C_1 := \max\{\kappa\Lambda_{\mathrm{J}}, 1 + \Lambda_{\mathrm{J}}\}$, *the solution* $u \in V$ *to* (3.1) *with* $F \in L^2(\Omega)$, *and the solution* $u_{\mathrm{M}} \in \mathrm{M}(\mathcal{T})$ *to* (3.4) *with* $F_h \equiv F$ *satisfy* $C_1^{-1}|\!|\!|u - u_{\mathrm{M}}|\!|\!|_{\mathrm{pw}} \le |\!|\!|u - I_{\mathrm{M}} u|\!|\!|_{\mathrm{pw}} + \mathrm{osc}_2(F, \mathcal{T})$.

The discrete scheme (3.4) requires a discrete right-hand side F_h for a general $F \in H^{-2}(\Omega)$. The evaluation of $F_h := F \circ J$ is feasible with $F_h(v_{\mathrm{M}}) := F(Jv_{\mathrm{M}})$ for all $v_{\mathrm{M}} \in \mathrm{M}(\mathcal{T})$ and the *(modified) nonconforming scheme* seeks the solution $u_{\mathrm{M}} \in \mathrm{M}(\mathcal{T})$ to

$$a_{\mathrm{pw}}(u_{\mathrm{M}}, v_{\mathrm{M}}) = F(Jv_{\mathrm{M}}) \quad \text{for all } v_{\mathrm{M}} \in \mathrm{M}(\mathcal{T}). \tag{3.10}$$

Let Λ_0 denote the norm of $1 - J$, where the right-inverse $J \in L(\mathrm{M}(\mathcal{T}); V)$ is regarded as a linear map between $\mathrm{M}(\mathcal{T})$ and V,

$$\Lambda_0 := \sup_{v_{\mathrm{M}} \in \mathrm{M}(\mathcal{T}) \setminus \{0\}} |\!|\!|v_{\mathrm{M}} - Jv_{\mathrm{M}}|\!|\!|_{\mathrm{pw}} / |\!|\!|v_{\mathrm{M}}|\!|\!|_{\mathrm{pw}} \le \Lambda_{\mathrm{J}}. \tag{3.11}$$

Theorem 3.9 (Best-approximation [16, 43])**.** *The solution* $u \in V$ *to* (3.1) *with* $F \in V^*$ *and the solution* $u_{\mathrm{M}} \in \mathrm{M}(\mathcal{T})$ *to* (3.10) *satisfy* $|\!|\!|u - u_{\mathrm{M}}|\!|\!|_{\mathrm{pw}} \le \sqrt{1 + \Lambda_0^2}\, |\!|\!|u - I_{\mathrm{M}} u|\!|\!|_{\mathrm{pw}}$. *The constant* $\sqrt{1 + \Lambda_0^2}$ *is optimal.*

Remark 3.10 (Extra orthogonality in Lem. 3.7e)**.** The L^2 orthogonality in Lemma 3.7e allows control over dual norm estimates of the form $\|v_{\mathrm{M}} - Jv_{\mathrm{M}}\|_{H^{-s}(\Omega)} \lesssim \|h_{\mathcal{T}}^s (v_{\mathrm{M}} - Jv_{\mathrm{M}})\|$ for $0 \le s \le 2$. This is critical in eigenvalue analysis or problems with low-order terms; for *e.g.*, in [14, 17]. The L^2 orthogonality in Lemma 3.7e also allows a direct proof of Theorem 3.8 that circumvents the *a posteriori* error analysis of the consistency term as part of the medius analysis [34]. Notice that the proof of the best-approximation of Theorem 3.9 for the modified scheme does not require the L^2 orthogonality in Lemma 3.7e.

Remark 3.11 (Minimal assumptions on the smoother)**.** The series of papers [41–43] addresses the question on the minimal assumptions on the smoother (partly as a right inverse only). This paper utilizes a smoother J with the properties of Lemma 3.7a–d.

The point in the subsequent example is that the smoother JI_{M} may be more costly than averaging in other examples but it is at almost no extra costs for the case of point forces, which are of practical importance in civil engineering.

Example 3.12 (Point forces)**.** Let m denote the point forces in the right-hand side, *i.e.*, let

$$F = \sum_{j=1}^{m} \alpha_j \delta_{a_j}, \tag{3.12}$$

the triangulation can be adopted such that the concentration point a_j becomes a vertex in the triangulation. The right-inverse property of I_{M} displays that $(JI_{\mathrm{M}} v_{\mathrm{M}})(z) = v_{\mathrm{M}}(z)$ holds at any vertex $z \in \mathcal{V}$ and for any Morley function $v_{\mathrm{M}} \in \mathrm{M}(\mathcal{T})$. Hence the evaluation of the modified right-hand side $F_h := F \circ J$ leads for (3.12) to $F_h v_{\mathrm{M}} = \sum_{j=1}^m \alpha_j v_{\mathrm{M}}(a_j)$. The averaging of I_{M} in Definition 3.5 shows in the more general case $V_h \subset P_2(\Omega)$ that $F_h := F \circ JI_{\mathrm{M}}$ leads to $F_h v_{\mathrm{M}} = \sum_{j=1}^m \alpha_j / |\mathcal{T}(a_j)| \sum_{T \in \mathcal{T}(a_j)} (v_h|_T)(a_j)$. The same formula applies to other smoothers like the enrichment in [6, 41–43].

4. Interpolation of piecewise H^2 functions

4.1. Equivalent norms

The Hilbert space $H^2(\mathcal{T}) \equiv \prod_{T\in\mathcal{T}} H^2(T)$ is endowed with a norm $\| \bullet \|_h$ from [23] defined by

$$\|v_{\rm pw}\|_h^2 := |\!|\!| v_{\rm pw} |\!|\!|_{\rm pw}^2 + j_h(v_{\rm pw})^2 \quad \text{for all } v_{\rm pw} \in H^2(\mathcal{T}). \tag{4.1}$$

The homogeneous boundary conditions in $H_0^2(\Omega)$ are included in the the jump contributions

$$j_h(v_{\rm pw})^2 := \sum_{E\in\mathcal{E}} \sum_{z\in\mathcal{V}(E)} h_E^{-2} |[v_{\rm pw}]_E(z)|^2 + \sum_{E\in\mathcal{E}} \left| \fint_E [\partial v_{\rm pw}/\partial\nu_E]_E \,{\rm ds} \right|^2 \tag{4.2}$$

by $[v_{\rm pw}]_E(z) = v_{\rm pw}|_{\omega(E)}(z)$ for $z \in \mathcal{V}(\partial\Omega)$ and $\left[\frac{\partial v_{\rm pw}}{\partial\nu_E}\right]_E = \frac{\partial v_{\rm pw}}{\partial\nu_E}|_E$ for $E \in \mathcal{E}(\partial\Omega)$ at the boundary with jump partner zero owing to the homogeneous boundary conditions in (3.1).

The discontinuous Galerkin schemes of [2,28] are associated with a another family of norms $\|\bullet\|_{\rm dG}$ depending on the two positive parameters $\sigma_1, \sigma_2 > 0$ in the semi-norm scalar product

$$c_{\rm dG}(v_{\rm pw}, w_{\rm pw}) := \sum_{E\in\mathcal{E}} \frac{\sigma_1}{h_E^3} \int_E [v_{\rm pw}]_E [w_{\rm pw}]_E \,{\rm ds} + \frac{\sigma_2}{h_E} \int_E \left[\frac{\partial v_{\rm pw}}{\partial\nu_E}\right]_E \left[\frac{\partial w_{\rm pw}}{\partial\nu_E}\right]_E {\rm ds} \tag{4.3}$$

for all $v_{\rm pw}, w_{\rm pw} \in H^2(\mathcal{T})$. The DG norm $\|\bullet\|_{\rm dG}$ is the square root of

$$\|v_{\rm pw}\|_{\rm dG}^2 := |\!|\!| v_{\rm pw} |\!|\!|_{\rm pw}^2 + c_{\rm dG}(v_{\rm pw}, v_{\rm pw}) \tag{4.4}$$

for all $v_{\rm pw} \in H^2(\mathcal{T})$. It depends on the parameters $\sigma_1, \sigma_2 > 0$ and so do all constants in the sequel; in particular those suppressed in the abbreviations $\lesssim$ and $\approx$. The conditions on the ellipticity of the scheme in Lemma 7.1 below will assert that σ_1 and σ_2 are sufficiently large. The analysis of this paper assumes this and simplifies the notation $\sigma_1 \approx 1 \approx \sigma_2$.

One result in Theorem 4.1 of [23] shows that $\|\bullet\|_h \approx \|\bullet\|_{\rm dG}$ in $H_0^2(\Omega) + P_2(\mathcal{T})$; but the two norms are equivalent in the larger vector space $H^2(\mathcal{T})$.

Theorem 4.1 ($\| \bullet \|_h \approx \|\bullet\|_{\rm dG}$)**.** *The function $\| \bullet \|_h$ from (4.1) and $\|\bullet\|_{\rm dG}$ from (4.4) define norms in $H^2(\mathcal{T})$ with*

$$\|v_{\rm pw}\|_h \approx \|v_{\rm pw}\|_{\rm dG} \lesssim \sum_{m=0}^{2} |h_{\mathcal{T}}^{m-2} v_{\rm pw}|_{H^m(\mathcal{T})} \quad \textit{for all } v_{\rm pw} \in H^2(\mathcal{T}).$$

Remark 4.2 ($\{j_h = 0\} \cap P_2(\mathcal{T}) = {\rm M}(\mathcal{T})$)**.** For any $v_2 \in P_2(\mathcal{T})$, the condition $j_h(v_2) = 0$ is equivalent to $v_2 \in {\rm M}(\mathcal{T})$. (This follows from the definitions of ${\rm M}(\mathcal{T})$ and j_h.)

Proof of $\| \bullet \|_h \lesssim \|\bullet\|_{\rm dG}$. The (possibly discontinuous) piecewise affine interpolation $v_1 \in P_1(\mathcal{T})$ of $v_{\rm pw} \in H^2(\mathcal{T})$ is defined by nodal interpolation $v_1|_T(z) = v_{\rm pw}|_T(z)$ at the three vertices $z \in \mathcal{V}(T)$ in each triangle $T \in \mathcal{T}$. It is well known from standard finite element interpolation [4,6,26] that the error $w := v_{\rm pw} - v_1 \in H^2(T)$ satisfies

$$\sum_{m=0}^{2} h_T^{m-2} |w|_{H^m(T)} \lesssim |v_{\rm pw}|_{H^2(T)} \tag{4.5}$$

for each triangle $T \in \mathcal{T}$ with explicit constants [21] that exclusively depend on the maximal angle in the triangulation. The nodal interpolation implies $[v_{\rm pw}]_E(z) = [v_1]_E(z)$ at each vertex $z \in \mathcal{V}(E)$ of an edge $E \in \mathcal{E}$. Since $[v_1]_E$ is an affine function along the edge $E \in \mathcal{E}$, an inverse estimate shows

$$h_E/6 \sum_{z\in\mathcal{V}(E)} |[v_1]_E(z)|^2 \le \|[v_1]_E\|_{L^2(E)}^2 \le 2\|[v_{\rm pw}]_E\|_{L^2(E)}^2 + 2\|[w]_E\|_{L^2(E)}^2 \tag{4.6}$$

with a triangle inequality in the last step for $w = v_{\rm pw} - v_1$. (The constant $h_E/6$ in the first inequality of (4.6) stems from the eigenvalues $h_E/2$ and $h_E/6$ of the 2×2 mass matrix of piecewise linear functions in 1D.) This implies an estimate for the first term of the definition of $j_h(v_{\rm pw})$:

$$\sum_{E\in\mathcal{E}}\sum_{z\in\mathcal{V}(E)} h_E^{-2}\big|[v_{\rm pw}]_E(z)\big|^2 \le 12\sum_{E\in\mathcal{E}} h_E^{-3}\Big(\big\|[v_{\rm pw}]_E\big\|^2_{L^2(E)} + \|[w]_E\|^2_{L^2(E)}\Big).$$

A typical contribution $(\fint_E [\partial v_{\rm pw}/\partial\nu_E]_E\,{\rm ds})^2$ for the second term (in the definition of j_h) is controlled with a Cauchy inequality by $h_E^{-1}\|[\partial v_{\rm pw}/\partial\nu_E]_E\|^2_{L^2(E)}$. This results in

$$j_h(v_{\rm pw})^2 \le \sum_{E\in\mathcal{E}} h_E^{-1}12\Big(h_E^{-2}\Big(\|[v_{\rm pw}]_E\|^2_{L^2(E)} + \|[w]_E\|^2_{L^2(E)}\Big) + \big\|[\partial v_{\rm pw}/\partial\nu_E]_E\big\|^2_{L^2(E)}\Big).$$

A triangle inequality $\|[w]_E\|_{L^2(E)} \le \|w|_{T_+}\|_{L^2(E)} + \|w|_{T_-}\|_{L^2(E)}$ for an interior edge $E = \partial T_+\cap\partial T_- \in \mathcal{E}(\Omega)$ shared by the two triangles $T_\pm\in\mathcal{T}$ plus trace inequalities on $T_\pm$ show

$$h_E^{1/2}\|[w]_E\|_{L^2(E)} \lesssim \|w\|_{L^2(\omega(E))} + h_E\|\nabla_{\rm pw} w\|_{L^2(\omega(E))} \lesssim h_E^2\big\|D^2_{\rm pw} v_{\rm pw}\big\|_{L^2(\omega(E))} \tag{4.7}$$

with (4.5) in the end. The omission of T_- in the above arguments for an edge $E\in\mathcal{E}(\partial\Omega)$ on the boundary provide (4.7) with $\overline{\omega(E)} = T_+$.

This and the finite overlap show $\sum_{E\in\mathcal{E}} h_E^{-3}\|[w]_E\|^2_{L^2(E)} \lesssim |\!|\!| v_{\rm pw}|\!|\!|^2_{\rm pw}$. In conclusion,

$$j_h(v_{\rm pw})^2 \lesssim |\!|\!| v_{\rm pw}|\!|\!|^2_{\rm pw} + \sum_{E\in\mathcal{E}} h_E^{-3}\|[v_{\rm pw}]_E\|^2_{L^2(E)} + \sum_{E\in\mathcal{E}} h_E^{-1}\big\|[\partial v_{\rm pw}/\partial\nu_E]_E\big\|^2_{L^2(E)}.$$

The upper bound in the latter estimate is $\|v_{\rm pw}\|^2_{\rm dG}$ up to the weights $\sigma_1\approx 1\approx\sigma_2$. □

Proof of $\|\bullet\|_{\rm dG} \lesssim \|\bullet\|_h$. Recall the piecewise affine interpolation $v_1\in P_1(\mathcal{T})$ of $v_{\rm pw}\in H^2(\mathcal{T})$ and $w := v_{\rm pw} - v_1 \in H^2(\mathcal{T})$ with (4.5) from the previous part of the proof. Standard trace inequalities as in (4.7) for the first term (and an analog for the second term $h_E^{-1/2}\|[\partial w/\partial\nu_E]_E\|_{L^2(E)}$) for $E\in\mathcal{E}$ provide

$$h_E^{-3/2}\|[w]_E\|_{L^2(E)} + h_E^{-1/2}\|[\partial w/\partial\nu_E]_E\|_{L^2(E)} \lesssim \big\|D^2_{\rm pw} v_{\rm pw}\big\|_{L^2(\omega(E))}. \tag{4.8}$$

This and triangle inequalities result in

$$\begin{aligned} &h_E^{-3/2}\|[v_{\rm pw}]_E\|_{L^2(E)} + h_E^{-1/2}\|[\partial v_{\rm pw}/\partial\nu_E]_E\|_{L^2(E)} \\ &\quad\lesssim h_E^{-3/2}\|[v_1]_E\|_{L^2(E)} + h_E^{-1/2}\|[\partial v_1/\partial\nu_E]_E\|_{L^2(E)} + \big\|D^2_{\rm pw} v_{\rm pw}\big\|_{L^2(\omega(E))}. \end{aligned}$$

The constant factor 1/2 in the upper bound of the first subsequent inequality (displayed as 2 in the lower bound) stems from the eigenvalues $h_E/2$ and $h_E/6$ of the 2×2 mass matrix of piecewise linear functions in 1D,

$$2h_E^{-3}\|[v_1]_E\|^2_{L^2(E)} \le h_E^{-2}\sum_{z\in\mathcal{V}(E)} |[v_1]_E(z)|^2 = h_E^{-2}\sum_{z\in\mathcal{V}(E)} |[v_{\rm pw}]_E(z)|^2$$

with the nodal interpolation property $v_1|_T(z) = v_{\rm pw}|_T(z)$ for $z\in\mathcal{V}(T)$, $T\in\mathcal{T}$, in the last step. The jump $[\partial v_1/\partial\nu_E]_E$ is constant along the edge E and so

$$h_E^{-1/2}\|[\partial v_1/\partial\nu_E]_E\|_{L^2(E)} = \Big|\fint_E [\partial v_1/\partial\nu_E]_E\,{\rm ds}\Big| \le \Big|\fint_E [\partial v_{\rm pw}/\partial\nu_E]_E\,{\rm ds}\Big| + \Big|\fint_E [\partial w/\partial\nu_E]_E\,{\rm ds}\Big|$$

with a triangle inequality in the last step. A Cauchy inequality $\|\nabla w_{\rm pw}\|_{L^1(E)} \le h_E^{1/2}\|\nabla w_{\rm pw}\|_{L^2(E)}$ and a trace inequality show (as above in (4.8)) that

$$\left|\fint_E [\partial w/\partial\nu_E]_E \,{\rm ds}\right| \le h_E^{-1/2}\|[\partial w/\partial\nu_E]_E\|_{L^2(E)} \lesssim \left\|D^2_{\rm pw} v_{\rm pw}\right\|_{L^2(\omega(E))}.$$

The combination of all aforementioned estimates reads

$$h_E^{-3}\left\|[v_{\rm pw}]_E\right\|^2_{L^2(E)} + h_E^{-1}\left\|[\partial v_{\rm pw}/\partial\nu_E]_E\right\|^2_{L^2(E)} \lesssim \left\|D^2_{\rm pw} v_{\rm pw}\right\|^2_{L^2(\omega(E))} + h_E^{-2}\sum_{z\in\mathcal{V}(E)}\left|[v_{\rm pw}]_E(z)\right|^2 + \left|\fint_E [\partial v_{\rm pw}/\partial\nu_E]_E \,{\rm ds}\right|^2.$$

The sum of all those estimates over $E\in\mathcal{E}$ plus $\|v_{\rm pw}\|^2_{\rm pw}$ leads to an estimate with the lower bound $\|v_{\rm pw}\|^2_{\rm dG}$ up to the weights $\sigma_1\approx 1\approx\sigma_2$. The finite overlap of the edge-patches $(\omega(E) : E\in\mathcal{E})$ shows that the resulting upper bound is $\lesssim \|v_{\rm pw}\|_h$. □

Proof of the upper bound. The proof of the asserted inequality starts with triangle inequalities for the jumps of $v_{\rm pw}\in H^2(\mathcal{T})$ and the shape regularity for $h_E\approx h_T$ for $E\in\mathcal{E}(T)$. This and a Cauchy inequality $\|\nabla v_{\rm pw}\|_{L^1(E)} \le h_E^{1/2}\|\nabla v_{\rm pw}\|_{L^2(E)}$ lead to

$$j_h(v_{\rm pw})^2 \lesssim \sum_{T\in\mathcal{T}}\left(h_T^{-2}\sum_{z\in\mathcal{V}(T)}|(v_{\rm pw}|_T)(z)|^2 + h_T^{-1}\sum_{E\in\mathcal{E}(T)}\|\nabla v_{\rm pw}\|^2_{L^2(E)}\right).$$

A one-dimensional trace inequality (with a factor 1 that follows from 1D integration)

$$|(v_{\rm pw}|_T)(z)| \le h_E^{-1/2}\|v_{\rm pw}\|_{L^2(E)} + h_E^{1/2}\|\nabla v_{\rm pw}\|_{L^2(E)}$$

along the edge $E\in\mathcal{E}(T)$ of the triangle $T\in\mathcal{T}$ with vertex $z\in\mathcal{V}(E)$ results in

$$j_h(v_{\rm pw})^2 \lesssim \sum_{T\in\mathcal{T}}\left(h_T^{-3}\|v_{\rm pw}\|^2_{L^2(\partial T)} + h_T^{-1}\|\nabla v_{\rm pw}\|^2_{L^2(\partial T)}\right) \lesssim \sum_{T\in\mathcal{T}}\sum_{m=0}^{2}\left|h_{\mathcal{T}}^{m-2} v_{\rm pw}\right|^2_{H^m(T)}$$

with standard trace inequalities on ∂T for $v_{\rm pw}$ and $\nabla v_{\rm pw}$ in the last step. The right-hand side is $\sum_{m=0}^2 |h_{\mathcal{T}}^{m-2}v_{\rm pw}|^2_{H^m(\mathcal{T})}$ as asserted. The remaining details are omitted for brevity. □

4.2. Interpolation errors

The interpolation error estimates are summarised in one theorem.

Theorem 4.3 (Interpolation)**.** *Any $v_{\rm pw}\in H^2(\mathcal{T})$ and its Morley interpolation $I_{\rm M}v_{\rm pw}\in {\rm M}(\mathcal{T})$ from Definition 3.5 satisfy*

(a) $\sum_{m=0}^2 |h_{\mathcal{T}}^{m-2}(v_{\rm pw}-I_{\rm M}v_{\rm pw})|_{H^m(\mathcal{T})} \lesssim \left\|(1-\Pi_0)D^2_{\rm pw}v_{\rm pw}\right\| + j_h(v_{\rm pw}) \le \|v_{\rm pw}\|_h$;

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

Proof of (a). The first step reduces the analysis to piecewise quadratic functions by the piecewise Morley interpolation $I_{\rm M}^{\rm loc}$. Definition 3.4 shows $\int_E \nabla\big(I_{\rm M}^{\rm loc}v_{\rm pw}-v_{\rm pw}\big)|_T\,{\rm ds}=0$ for an edge $E\in\mathcal{E}(T)$ of the triangle T and therefore $D^2\big(I_{\rm M}^{\rm loc}v_{\rm pw}\big)|_T = \Pi_0 D^2 v_{\rm pw}|_T$ *a.e.* in $T\in\mathcal{T}$. Notice that the piecewise defined Morley interpolation

$v_2 := I_{\mathrm{M}}^{\mathrm{loc}} v_{\mathrm{pw}} \in P_2(\mathcal{T})$ is discontinuous (and shares none of the compatibility or boundary conditions) in general. The interpolation error estimates of Lemma 3.2b read

$$\sum_{m=0}^{2} h_T^{m-2} |v_2 - v_{\mathrm{pw}}|_{H^m(T)} \le 2 \big\| (1-\Pi_0) D^2 v_{\mathrm{pw}} \big\|_{L^2(T)}.$$

This and a triangle inequality show that it remains to prove that $v_{\mathrm{M}} := I_{\mathrm{M}} v_{\mathrm{pw}}$ satisfies

$$\sum_{m=0}^{2} h_T^{m-2} |v_2 - v_{\mathrm{M}}|_{H^m(T)} \lesssim j_h(v_{\mathrm{pw}}, T) \tag{4.9}$$

for the jump terms localised to a neighbourhood $\Omega(T)$ of $T \in \mathcal{T}$ as follows. The neighbourhood $\Omega(T)$ is the interior of the union $\cup\{K \in \mathcal{T} : \mathrm{dist}(T,K) = 0\}$ of $T \in \mathcal{T}$ plus one layer of triangles in $\mathcal{T}$ around. Then

$$j_h(v_{\mathrm{pw}}, T)^2 := \sum_{z \in \mathcal{V}(T)} \sum_{F \in \mathcal{E}(z)} h_F^{-2} |[v_{\mathrm{pw}}]_F(z)|^2 + \sum_{E \in \mathcal{E}(T)} \left| \fint_E \left[\frac{\partial v_{\mathrm{pw}}}{\partial \nu_E} \right]_E \mathrm{ds} \right|^2$$

is the contribution from T and its neighbourhood $\Omega(T)$ to the full jump term $j_h(v_{\mathrm{pw}})^2$ with the spider $\mathcal{E}(z) := \{F \in \mathcal{E} : z \in \mathcal{V}(F)\}$ of edges with one end-point $z \in \mathcal{V}(T)$.

The second step reduces the analysis to piecewise quadratic functions. The first obervation is that the averaging of the degrees of freedom in the definition of I_{M} merely employs the data of $v_2 = I_{\mathrm{M}}^{\mathrm{loc}} v_{\mathrm{pw}}$ in the sense that $v_{\mathrm{M}} = I_{\mathrm{M}} v_{\mathrm{pw}} = I_{\mathrm{M}} v_2$. This explains why $j_h(v_{\mathrm{pw}}, T) = j_h(v_2, T)$ in the asserted estimate (4.9). The second observation is that the left-hand side of (4.9) involves the polyonomial $(v_2 - v_{\mathrm{M}})|_T \in P_2(T)$ that allows for inverse estimates

$$\sum_{m=0}^{2} h_T^{m-2} |v_2 - I_{\mathrm{M}} v_2|_{H^m(T)} \lesssim h_T^{-2} \|v_2 - I_{\mathrm{M}} v_2\|_{L^2(T)}.$$

The overall conclusion is that it suffices to prove, for all $v_2 \in P_2(\mathcal{T})$, that

$$h_T^{-4} \|v_2 - I_{\mathrm{M}} v_2\|_{L^2(T)}^2 \lesssim j_h(v_2, T)^2. \tag{4.10}$$

In fact, equation (4.10) and the aforementioned arguments lead to a localised form of the assertion. The sum over all $T \in \mathcal{T}$ and the bounded overlap of $(\Omega(T) : T \in \mathcal{T})$ then conclude the proof of the theorem.

The third step reduces the proof of (4.10) to six coefficients. The six degrees of freedom on a triangle $T \in \mathcal{T}$ are the three point evaluations δ_z at the three vertices $z \in \mathcal{V}(T)$ and the three integral means of the normal derivatives $\fint_E \partial_{\nu_E} \bullet \, \mathrm{ds}$ along the three edges $E \in \mathcal{E}(T)$. The six dual basis functions ψ_z for $z \in \mathcal{V}(T)$ and ψ_E for $E \in \mathcal{E}(T)$ in $P_2(T)$ are defined by the duality relations $\psi_E(z) = 0 = \fint_E \partial_{\nu_E} \psi_z \, \mathrm{ds}$ and $\psi_z(z) = 1 = \fint_E \partial_{\nu_E} \psi_E \, \mathrm{ds}$ for all $z \in \mathcal{V}(T)$ and $E \in \mathcal{E}(T)$, while $\psi_y(z) = 0 = \fint_E \partial_{\nu_E} \psi_F \, \mathrm{ds}$ for all vertices $z \ne y \in \mathcal{V}(T)$ and edges $E \ne F \in \mathcal{E}(T)$. Those functions are known and given explicitly (*e.g.*, in [22] in the context of a short implementation of the Morley FEM in 30 lines of Matlab) with a scaling (which is generally understood and follows from the explicit formulas)

$$\|\psi_z\|_{L^2(T)} \approx |T|^{1/2} \approx h_T \quad \text{and} \quad \|\psi_E\|_{L^2(T)} \approx h_T |T|^{1/2} \approx h_T^2$$

for all $z \in \mathcal{V}(T)$ and $E \in \mathcal{E}(T)$. On the other hand, given the dual basis of $P_2(T)$, any function $w_2 := v_2 - v_{\mathrm{M}} \in P_2(T)$ for $v_{\mathrm{M}} := I_{\mathrm{M}} v_2$ allows for a representation

$$w_2 = v_2 - v_{\mathrm{M}} = \sum_{z \in \mathcal{V}(T)} w(z)\, \psi_z + \sum_{E \in \mathcal{E}(T)} w(E)\, \psi_E \quad \text{in } T$$

with the real coefficients $w(z) := v_2|_T(z) - v_{\mathrm{M}}(z)$ and $w(E) := \fint_E \partial_{\nu_E} v_2|_T \, \mathrm{ds} - \fint_E \partial_{\nu_E} v_{\mathrm{M}} \, \mathrm{ds}$ for $z \in \mathcal{V}(T)$ and $E \in \mathcal{E}(T)$. Notice that the contributions of the piecewise quadratic v_2 are taken from $T \in \mathcal{T}$ and this is

written explicitly by $v_2|_T$ in the coefficients, while the corresponding values of the Morley function $v_{\mathrm{M}} \in \mathrm{M}(\mathcal{T})$ are independent of T as long as $z \in \mathcal{V}(T)$ or $E \in \mathcal{E}(T)$. Given the coefficients $w(z)$ and $w(E)$, the triangle inequality in $L^2(T)$ and the scaling of the dual basis functions lead to

$$\|v_2 - v_{\mathrm{M}}\|_{L^2(T)} \lesssim h_T \sum_{z\in\mathcal{V}(T)} |w(z)| + h_T^2 \sum_{E\in\mathcal{E}(T)} |w(E)|. \tag{4.11}$$

The fourth step analyses the coefficients in (4.11). Let the triangles $\mathcal{T}(z) := \{T \in \mathcal{T} : z \in \mathcal{V}(T)\} = \{T(1), \dots, T(J)\}$ at the vertex $z \in \mathcal{V}$ be enumerated such that $T(j)$ and $T(j+1)$ share an edge $\partial T(j) \cap \partial T(j+1) =: E(j) \in \mathcal{E}(z)$ for $j = 1, \dots, J$. For an interior vertex $z \in \mathcal{V}(\Omega)$, the patch is closed and then $T(1)$ and $T(J)$ share an edge $\partial T(1) \cap \partial T(J) =: E(J) \in \mathcal{E}(z)$ as well. Define $x_j := ((v_2 - v_{\mathrm{M}})|_{T(j)})(z)$ for $j = 1, \dots, J$ and observe for an interior vertex $z \in \mathcal{V}(\Omega)$ that $\sum_{j=1}^J x_j = 0$ (from the choice of $v_{\mathrm{M}}(z)$ as the arithmetic mean of the $v_2|_{T(j)}$) and that

$$\sum_{E\in\mathcal{E}(z)} |[v_2]_E(z)| = \sum_{j=1}^J |x_{j+1} - x_j|$$

with $x_{J+1} := x_1$ (recall $z \in \mathcal{V}(\Omega)$ here). Since the arithmetic mean of the real numbers $x_1, \dots, x_J$ vanishes, zero belongs to their convex hull; whence $\underline{m} := \min_{j=1,\dots,J} x_j \le 0 \le \max_{j=1,\dots,J} x_j =: \overline{m}$. A triangle inequality in this sequence $x_1, \dots, x_J$ shows that $\overline{m} - \underline{m} \le \sum_{j=1}^J |x_{j+1} - x_j|$ (even with an omitted factor 1/2). It follows $|x_1|, \dots, |x_J| \le \sum_{j=1}^J |x_{j+1} - x_j|$ and so, for a triangle $T \in \mathcal{T}(z)$ in the notation of (4.11),

$$|w(z)| \le \sum_{E\in\mathcal{E}(z)} |[v_2]_E(z)| \le J^{1/2} \sqrt{\sum_{E\in\mathcal{E}(z)} |[v_2]_E(z)|^2} \tag{4.12}$$

follows (with a Cauchy inequality in $\mathbb{R}^J$ in the end). This is suboptimal and the best constant in a squared version of this argument is contained in Appendix C of [17]. Observe that $J \lesssim 1$ is bounded from above by the shape regularity of the triangulation $\mathcal{T}$.

In the remaining case of a vertex $z \in \mathcal{V}(\partial\Omega)$ on the boundary, $v_{\mathrm{M}}(z) = 0$ and, in the above notation $\mathcal{T}(z) = \{T \in \mathcal{T} : z \in \mathcal{V}(T)\} = \{T(1), \dots, T(J)\}$ and $x_j = ((v_2 - v_{\mathrm{M}})|_{T(j)})(z) = (v_2|_{T(j)})(z)$ for $j = 1, \dots, J$. The homogeneous boundary conditions enter in the jump terms for $E(1) := T_1 \cap \partial\Omega$ and $E(J) := T_J \cap \partial\Omega$ and

$$\sum_{E\in\mathcal{E}(z)} |[v_2]_E(z)| = |x_1| + |x_J| + \sum_{j=1}^{J-1} |x_{j+1} - x_j|.$$

Triangle inequalities show $|x_1|, \dots, |x_J| \le |x_1| + |x_J| + \sum_{j=1}^{J-1} |x_{j+1} - x_j|$ (even with an omitted factor 1/2) and the above arguments lead to (4.12) as well. (The optimal constant for this argument may be found in Lem. 4.2 of [11].)

Recall the design of the Morley interpolation in Definition 3.5 with the arithmetic mean $\fint_E \partial_{\nu_E} v_{\mathrm{M}}\,\mathrm{ds} = \fint_E \left\langle \frac{\partial v_2}{\partial \nu_E} \right\rangle_E \mathrm{ds}$ of the two normal traces for an interior edge $E = \partial T_+ \cap \partial T_- \in \mathcal{E}(\Omega)$. This leads to the edge contribution

$$w(E) = \fint_E \partial_{\nu_E} v_2|_T\,\mathrm{ds} - \fint_E \left\langle \frac{\partial v_2}{\partial \nu_E} \right\rangle_E \mathrm{ds} = \pm\frac{1}{2} \fint_E [\partial_{\nu_E} v_2]_E\,\mathrm{ds}$$

in (4.11) with a sign $\pm$ for $T = T_\pm$. The boundary conditions for a boundary edge $E \in \mathcal{E}(\partial\Omega)$ and the jump convention for $[\bullet]_E$ (recall that ν_E points outwards for $E \subset \partial\Omega$) directly show $w(E) = \fint_E [\partial_{\nu_E} v_2]_E\,\mathrm{ds}$. It follows

$$|w(E)| \le \left| \fint_E [\partial_{\nu_E} v_2]_E\,\mathrm{ds} \right| \text{ for any } E \in \mathcal{E}(T). \tag{4.13}$$

The fitfh step finishes the proof. Recall that the coefficients $w(z)$ for $z \in \mathcal{V}(z)$ and $w(E)$ for any $E \in \mathcal{E}(T)$ in (4.11) satisfy (4.12) and (4.13). The resulting estimate reads

$$h_T^{-4}\|v_2 - v_{\mathrm{M}}\|_{L^2(T)}^2 \lesssim h_T^{-2} \sum_{z\in\mathcal{V}(T)} \sum_{E\in\mathcal{E}(z)} |[v_2]_E(z)|^2 + \sum_{E\in\mathcal{E}(T)} \left| \fint_E [\partial_{\nu_E} v_2]_E \,\mathrm{ds}\right|^2 \approx j_h(v_2, T)^2$$

with the shape regularity $h_F \approx h_T$ for $F \in \mathcal{E}(z)$ and $z \in \mathcal{V}(T)$ in the end. This concludes the proof of (4.10) and thus that of (a) as outlined at the end of the second step. □

Proof of (b). Given any $w_{\mathrm{M}} \in \mathrm{M}(\mathcal{T})$, part (a) shows that the first term T_1 in the equivalence (b) is $T_1 \lesssim |\!|\!|v_{\mathrm{pw}} - w_{\mathrm{M}}|\!|\!|_{\mathrm{pw}} + j_h(v_{\mathrm{pw}} - w_{\mathrm{M}}) \le \|v_{\mathrm{pw}} - w_{\mathrm{M}}\|_h =: T_2$ with $j_h(v_{\mathrm{pw}}) = j_h(v_{\mathrm{pw}} - w_{\mathrm{M}})$ in the last step. Theorem 4.1 applies to $v_{\mathrm{pw}} - w_{\mathrm{M}} \in H^2(\mathcal{T})$ and proves $T_2 \lesssim \sum_{m=0}^2 |h_{\mathcal{T}}^{m-2}(v_{\mathrm{pw}} - w_{\mathrm{M}})|_{H^m(\mathcal{T})} =: T_3$. The estimates $T_1 \lesssim T_2 \lesssim T_3$ hold for all $w_{\mathrm{M}} \in \mathrm{M}(\mathcal{T})$ and so for the respective minima as well. Since $I_{\mathrm{M}} v_{\mathrm{pw}} \in \mathrm{M}(\mathcal{T})$, the remaining estimate $\min_{w_{\mathrm{M}}\in\mathrm{M}(\mathcal{T})} T_3 \le T_1$ is obvious. □

Remark 4.4 ($I_{\mathrm{M}} J I_{\mathrm{M}} = I_{\mathrm{M}}$ in $H^2(\mathcal{T})$)**.** Let J be any right-inverse of I_{M} in the sense of (3.9). Since $I_{\mathrm{M}} J$ is identity in $\mathrm{M}(\mathcal{T})$, $I_{\mathrm{M}} J I_{\mathrm{M}} v_{\mathrm{pw}} = I_{\mathrm{M}} v_{\mathrm{pw}}$ holds for any $v_{\mathrm{pw}} \in H^2(\mathcal{T})$.

4.3. Approximation errors

The subsequent theorem discusses the approximation properties of $J \circ I_{\mathrm{M}}$ for piecewise smooth and piecewise quadratic functions. It is formulated in terms of $\|\bullet\|_h \approx \|\bullet\|_{\mathrm{dG}}$ and the norm equivalence implies an (undisplayed) analog for $\|\bullet\|_{\mathrm{dG}}$ as well.

Theorem 4.5 (Approximation)**.** *Any $v_{\mathrm{pw}} \in H^2(\mathcal{T})$ and $v_2 \in P_2(\mathcal{T})$ satisfy (a)–(d).*

(a) $\|v_{\mathrm{pw}} - J I_{\mathrm{M}} v_{\mathrm{pw}}\|_h \lesssim \|(1-\Pi_0) D_{\mathrm{pw}}^2 v_{\mathrm{pw}}\|_{L^2(\Omega)} + \min_{v\in H_0^2(\Omega)} \|v_{\mathrm{pw}} - v\|_h$;

(b) $\sum_{m=0}^2 |h_{\mathcal{T}}^{m-2}(v_{\mathrm{pw}} - J I_{\mathrm{M}} v_{\mathrm{pw}})|_{H^m(\mathcal{T})} \lesssim \|(1-\Pi_0) D_{\mathrm{pw}}^2 v_{\mathrm{pw}}\|_{L^2(\Omega)} + \min_{v\in H_0^2(\Omega)} \sum_{m=0}^2 |h_{\mathcal{T}}^{m-2}(v_{\mathrm{pw}} - v)|_{H^m(\mathcal{T})}$;

(c) $\|v_2 - J I_{\mathrm{M}} v_2\|_h \approx \min_{v\in H_0^2(\Omega)} \|v_2 - v\|_h \approx \sum_{m=0}^2 |h_{\mathcal{T}}^{m-2}(v_2 - J I_{\mathrm{M}} v_2)|_{H^m(\mathcal{T})} \approx \min_{v\in H_0^2(\Omega)} \sum_{m=0}^2 |h_{\mathcal{T}}^{m-2}(v_2 - v)|_{H^m(\mathcal{T})}$;

(d) $\|v_2 - J I_{\mathrm{M}} v_2\|_{H^s(\mathcal{T})} \lesssim h_{\max}^{2-s} \min_{v\in H_0^2(\Omega)} \|v_2 - v\|_h$ *holds for any* $0 \le s \le 2$.

Remark 4.6. Theorem 4.5 implies that $P = Q = J \circ I_{\mathrm{M}}$ is a quasi-optimal smoother with constant $\Lambda_{\mathrm{P}} = \Lambda_{\mathrm{Q}}$ that depends only on the shape regularity of the triangulation.

Remark 4.7 (Remainder in (a) and (b))**.** The extra term $\big\|(1-\Pi_0) D_{\mathrm{pw}}^2 v_{\mathrm{pw}}\big\|_{L^2(\Omega)}$ in the upper bound will vanish for piecewise quadratic functions but cannot be omitted in (a) and (b). For a proof of the latter statement by contradiction consider some $v_{\mathrm{pw}} \in H_0^2(\Omega) \setminus (\mathrm{HCT}(\mathcal{T}) + P_8(\mathcal{T}))$. Since $J I_{\mathrm{M}} v_{\mathrm{pw}} \in \mathrm{HCT}(\mathcal{T}) + P_8(\mathcal{T})$, the left-hand side in (a) and (b) is positive, while $v = v_{\mathrm{pw}} \in H_0^2(\Omega)$ leads to a right-hand side zero if the term $\|(1-\Pi_0) D_{\mathrm{pw}}^2 v_{\mathrm{pw}}\|_{L^2(\Omega)}$ was neglected.

Proof of (a). Theorem 4.1 implies the first estimate (4.14a) below and Theorem 4.3a asserts the second (4.14b) for the Morley interpolation $v_{\mathrm{M}} := I_{\mathrm{M}} v_{\mathrm{pw}}$ of $v_{\mathrm{pw}} \in H^2(\mathcal{T})$ in

$$\|v_{\mathrm{pw}} - v_{\mathrm{M}}\|_h \lesssim \sum_{m=0}^2 \big|h_{\mathcal{T}}^{m-2}(v_{\mathrm{pw}} - v_{\mathrm{M}})\big|_{H^m(\mathcal{T})} \lesssim \big\|(1-\Pi_0) D_{\mathrm{pw}}^2 v_{\mathrm{pw}}\big\|_{L^2(\Omega)} + j_h(v_{\mathrm{pw}}). \tag{4.14}$$

Notice that $j_h(v_{\mathrm{M}} - J v_{\mathrm{M}}) = 0$ implies $\|v_{\mathrm{M}} - J v_{\mathrm{M}}\|_h = |\!|\!|v_{\mathrm{M}} - J v_{\mathrm{M}}|\!|\!|_{\mathrm{pw}} \lesssim |\!|\!|v_{\mathrm{M}} - v|\!|\!|_{\mathrm{pw}}$ for any $v \in V$ from Lemma 3.7d in the last step. This and a triangle inequality, (4.14), and $|\!|\!|v_{\mathrm{pw}} - v|\!|\!|_{\mathrm{pw}} + j_h(v_{\mathrm{pw}}) \le \sqrt{2}\|v_{\mathrm{pw}} - v\|_h$ show

$$\|v_{\mathrm{M}} - J v_{\mathrm{M}}\|_h \lesssim |\!|\!|v_{\mathrm{pw}} - v|\!|\!|_{\mathrm{pw}} + |\!|\!|v_{\mathrm{pw}} - v_{\mathrm{M}}|\!|\!|_{\mathrm{pw}} \lesssim \|(1-\Pi_0) D_{\mathrm{pw}}^2 v_{\mathrm{pw}}\|_{L^2(\Omega)} + \|v_{\mathrm{pw}} - v\|_h. \tag{4.15}$$

This and a triangle inequality $\|v_{\mathrm{pw}} - J I_{\mathrm{M}} v_{\mathrm{pw}}\|_h \le \|v_{\mathrm{pw}} - v_{\mathrm{M}}\|_h + \|v_{\mathrm{M}} - J v_{\mathrm{M}}\|_h$ and (4.14), (4.15) conclude the proof of (a). □

Proof of (b). Adapt the notation of part (a) and recall that Theorem 4.3a provides (4.14b), the second estimate in (4.14). Since $v_{\mathrm{M}} - Jv_{\mathrm{M}} = I_{\mathrm{M}}Jv_{\mathrm{M}} - Jv_{\mathrm{M}}$ (from (3.9)), Lemma 3.2 controls this interpolation error of $Jv_{\mathrm{M}} \in H_0^2(\Omega)$ and shows

$$\sum_{m=0}^{2} \big|h_{\mathcal{T}}^{m-2}(v_{\mathrm{M}} - Jv_{\mathrm{M}})\big|_{H^m(\mathcal{T})} \le 2|\!|\!|v_{\mathrm{M}} - Jv_{\mathrm{M}}|\!|\!|_{\mathrm{pw}} \lesssim \big\|(1-\Pi_0)D^2_{\mathrm{pw}}v_{\mathrm{pw}}\big\|_{L^2(\Omega)} + \|v_{\mathrm{pw}} - v\|_h$$

with (4.15) in the last step. Theorem 4.1 applies to $v_{\mathrm{pw}} - v \in H^2(\mathcal{T})$. The combination of the resulting estimate with the previous one concludes the proof of (b). □

Proof of (c). The assertions (a) and (b) apply to $v_{\mathrm{pw}} := v_2 \in P_2(\mathcal{T})$ and the extra term $\|(1-\Pi_0)D^2_{\mathrm{pw}}v_{\mathrm{pw}}\|_{L^2(\Omega)}$ vanishes. The resulting estimates allow for obvious converse inequalities and so prove, for $v_2 \in P_2(\mathcal{T})$ and $v_{\mathrm{M}} := I_{\mathrm{M}}v_2 \in \mathrm{M}(\mathcal{T})$, that

$$\|v_2 - Jv_{\mathrm{M}}\|_h \approx \min_{v\in H_0^2(\Omega)} \|v_2 - v\|_h \lesssim \sum_{m=0}^{2} \big|h_{\mathcal{T}}^{m-2}(v_2 - Jv_{\mathrm{M}})\big|_{H^m(\mathcal{T})} \approx \min_{v\in H_0^2(\Omega)} \sum_{m=0}^{2} \big|h_{\mathcal{T}}^{m-2}(v_2 - v)\big|_{H^m(\mathcal{T})}$$

with Theorem 4.1 in between the two equivalences. A triangle inequality, the estimate (4.14b), the estimate for $(1-J)v_{\mathrm{M}}$ in the proof of (b) and (4.1) applies to $v_{\mathrm{pw}} := v_2 \in P_2(\mathcal{T})$ and shows $\sum_{m=0}^{2} |h_{\mathcal{T}}^{m-2}(v_2 - Jv_{\mathrm{M}})|_{H^m(\mathcal{T})} \lesssim j_h(v_2) + \|v_2 - v\|_h = j_h(v_2 - v) + \|v_2 - v\|_h \le 2\|v_2 - v\|_h$ for any $v \in H_0^2(\Omega)$. This concludes the proof of (c). □

Proof of (d). The equivalence of the Sobolev–Slobodeckii norm and the norm by interpolation of Sobolev spaces ([35], Rem. 9.1), for instance for a fixed reference triangle $T = T_{\mathrm{ref}}$ with $C_2(s) = C_2(s, T_{\mathrm{ref}})$, provides for $w := (v_2 - JI_{\mathrm{M}}v_2)|_T \in H^2(T)$ the estimate

$$\|w\|_{H^s(T)} \le C_2(s)\, \|w\|^{2-s}_{H^1(T)} \|w\|^{s-1}_{H^2(T)} \quad \text{for } 1 < s < 2. \tag{4.16}$$

A straightforward transformation of Sobolev norms ([26], Thm. 3.1.2) show (4.16) for any triangle $T \in \mathcal{T}$ with $C_2(s) = C_2(s, T) = \kappa^{1+s}C_2(s, T_{\mathrm{ref}})$ for the condition number $\kappa = \sigma_1/\sigma_2$ of the affine transformation $a + Bx$ of T_{ref} to T with the 2×2 matrix B and its positive singular values $\sigma_2 \le \sigma_1$. A more detailed analysis [15] reveals that $C_2(s)$ exclusively depends on s (but exploits singularities as s approaches the end-points 0 and 1). The estimate (4.16) shows the first inequality in

$$C_2(s)^{-2}\, \|w\|^2_{H^s(T)} \le \|w\|^{2(2-s)}_{H^1(T)} \|w\|^{2(s-1)}_{H^2(T)} \le \|w\|^2_{H^1(T)} + \|w\|^{2(2-s)}_{H^1(T)} |w|^{2(s-1)}_{H^2(T)}$$

with the subadditivity $(a+b)^p \le a^p + b^p$ for $a, b \ge 0$ and $0 < p = s - 1 < 1$ (*e.g.*, from the concavity of $x \mapsto x^p$ for non-negative x) in the last step. An elementary estimate is followed by the Young inequality $ab \le a^p/p + b^q/q$ for $p = (2-s)^{-1}$, $q = (s-1)^{-1}$, $a = \big\|h_T^{-1}w\big\|^{2(2-s)}_{H^1(T)}$, and $b = |w|^{2(s-1)}_{H^2(T)}$ to prove

$$h_{\max}^{2(s-2)} \|w\|^{2(2-s)}_{H^1(T)} |w|^{2(s-1)}_{H^2(T)} \le \big\|h_T^{-1}w\big\|^{2(2-s)}_{H^1(T)} |w|^{2(s-1)}_{H^2(T)} \le \big\|h_T^{-1}w\big\|^2_{H^1(T)} + |w|^2_{H^2(T)}.$$

This and the trivial estimate $h_{\max}^{2(s-1)} \le \mathrm{diam}(\Omega)^{2(s-1)}$ leads to

$$\|w\|^2_{H^s(T)} \le C_3 h_{\max}^{2(2-s)} \Big(\big\|h_T^{-1}w\big\|^2_{H^1(T)} + |w|^2_{H^2(T)} \Big)$$

for $C_3 = C_2(s)^2(1 + \mathrm{diam}(\Omega)^{2(s-1)})$. The sum over all those contributions over $T \in \mathcal{T}$ proves

$$\|v_2 - JI_{\mathrm{M}}v_2\|_{H^s(\mathcal{T})} \le C_3^{1/2} h_{\max}^{2-s} \Big(\big\|h_{\mathcal{T}}^{-1}(v_2 - JI_{\mathrm{M}}v_2)\big\|_{H^1(\mathcal{T})} + |\!|\!|v_2 - JI_{\mathrm{M}}v_2|\!|\!|_{\mathrm{pw}} \Big)$$

$$\lesssim h_{\max}^{2-s} \min_{v\in V} \|v_2 - v\|_h$$

with Theorem 4.5c in the last step. This concludes the proof of (d) for $1 < s < 2$. The assertion (d) is included in Theorem 4.5c for $s = 0, 1, 2$. The remaining case $0 < s < 1$ is similar to the above analysis with $\|w\|_{H^s(T)} \le C_2(s)\|w\|_{L^2(T)}^{1-s}\|w\|_{H^1(T)}^{s}$ replacing (4.16) and analogous arguments; hence further details are omitted. □

5. Abstract framework for best-approximation of lower-order methods

5.1. Discretisation

Suppose that $V_h \subset H^2(\mathcal{T})$ is the finite-dimensional trial and test space of an abstract (discontinuous Galerkin) scheme with a bilinear form

$$A_h : (V_h + \mathrm{M}(\mathcal{T})) \times (V_h + \mathrm{M}(\mathcal{T})) \to \mathbb{R}$$

that is coercive and continuous with respect to some norm $\|\bullet\|_h$ in $H^2(\mathcal{T})$ in the sense that, for all $v_h, w_h \in V_h$,

$$\alpha\|v_h\|_h^2 \le A_h(v_h, v_h) \text{ and } A_h(v_h, w_h) \le M\|v_h\|_h\|w_h\|_h \tag{5.1}$$

hold for some universal constants $0 < \alpha, M < \infty$. Suppose that $\|\bullet\|_h$ is a norm in $H^2(\mathcal{T})$ and equal to the norm $|||\bullet|||_{\mathrm{pw}} := a_{\mathrm{pw}}(\bullet,\bullet)^{1/2}$ in $V + \mathrm{M}(\mathcal{T})$ and stronger in general, *i.e.*,

$$\text{(a) } |||\bullet|||_{\mathrm{pw}} \le \|\bullet\|_h \text{ in } H^2(\mathcal{T}) \quad \text{and} \quad \text{(b) } |||\bullet|||_{\mathrm{pw}} = \|\bullet\|_h \text{ in } V + \mathrm{M}(\mathcal{T}). \tag{5.2}$$

Given a linear operator $JI_{\mathrm{M}} : V_h \to V$ with the companion operator J from Lemma 3.7 and the (extended) linear interpolation operator I_{M} from Section 3.3 the discrete problem reads: Given $F \in V^* = H^{-2}(\Omega)$ seek the discrete solution $u_h \in V_h$ to

$$A_h(u_h, v_h) = F(JI_{\mathrm{M}}v_h) \quad \text{for all } v_h \in V_h. \tag{5.3}$$

The Lax-Milgram lemma assures the existence of a unique discrete solution u_h to (5.3).

Remark 5.1 ($\|\bullet\|_h$)**.** The examples of Sections 7 and 8 utilize $\|\bullet\|_h$ given in (4.1) and (4.2), but the abstract framework allows more general A_h and $\|\bullet\|_h$ with (5.1) and (5.2) in (5.3).

5.2. First glance at the analysis

This subsection motivates the abstract conditions and emphasises the relevance of the discrete consistency condition (dcc)

$$a_{\mathrm{pw}}(I_{\mathrm{M}}u, e_h - I_{\mathrm{M}}e_h) + b_h(I_{\mathrm{M}}u, e_h - I_{\mathrm{M}}e_h) \le \Lambda_{\mathrm{dc}}|||u - I_{\mathrm{M}}u|||_{\mathrm{pw}}\|e_h\|_h \tag{5.4}$$

that leads to the best-approximation in terms of $|||u - I_{\mathrm{M}}u|||_{\mathrm{pw}} = \min_{v_2\in P_2(T)} |||u - v_2|||_{\mathrm{pw}}$ from (3.7). The test function $e_h := I_hI_{\mathrm{M}}u - u_h \in V_h \subset H^2(\mathcal{T})$ is the discrete approximation of the error $u - u_h$ with $I_{\mathrm{M}} : H^2(\mathcal{T}) \to \mathrm{M}(\mathcal{T})$ from Definition 3.5 and a transfer operator $I_h : \mathrm{M}(\mathcal{T}) \to V_h$ from Section 5.3 below. For the dGFEM of Section 7 and the WOPSIP scheme of Section 10, I_h is the identity 1 and otherwise it is controlled nicely (*cf.*, (5.11) below for details) [23]. So we may neglect the difference $1 - I_h$ for the sake of this first look at the analysis and suppose $I_h = 1$. The key identity from the continuous problem (3.1) and the discrete one (5.3) reads

$$a(u, JI_{\mathrm{M}}e_h) = F(JI_{\mathrm{M}}e_h) = A_h(u_h, e_h). \tag{5.5}$$

The stability of the scheme $\alpha\|e_h\|_h^2 \le A_h(e_h, e_h) \le M\|e_h\|_h^2$ motivates the investigation of

$$A_h(e_h, e_h) = a_{\mathrm{pw}}(e_h, e_h) + b_h(e_h, e_h) + c_h(e_h, e_h) \tag{5.6}$$

for the three bilinear forms that define the class of problems in (5.12) displayed in Table 1. The stability term $c_h(\bullet,\bullet)$ is controlled nicely in harmony with the discrete norm $\|\bullet\|_h$, while $b_h(\bullet,\bullet)$ drives the method and completes the leading term $a_{\rm pw}(\bullet,\bullet)$.

The definition of e_h in (5.6) leads to

$$A_h(e_h,e_h) = a_{\rm pw}(I_{\rm M}u,e_h) - A_h(u_h,e_h) + b_h(I_{\rm M}u,e_h) + c_h(I_{\rm M}u,e_h) \tag{5.7}$$

Since J is a the right-inverse of $I_{\rm M}$, (3.5) implies $a_{\rm pw}(I_{\rm M}u, I_{\rm M}e_h - JI_{\rm M}e_h) = 0$. This and elementary algebra show

$$a_{\rm pw}(I_{\rm M}u,e_h) = a_{\rm pw}(I_{\rm M}u,e_h - I_{\rm M}e_h) + a_{\rm pw}(I_{\rm M}u, JI_{\rm M}e_h).$$

This in combination with (5.5) leads in (5.7) to

$$A_h(e_h,e_h) = a_{\rm pw}(I_{\rm M}u,e_h - I_{\rm M}e_h) + a_{\rm pw}(I_{\rm M}u - u, JI_{\rm M}e_h) + b_h(I_{\rm M}u,e_h) + c_h(I_{\rm M}u,e_h). \tag{5.8}$$

The second term in the right-hand side of (5.8) is equal to $a_{\rm pw}(I_{\rm M}u - u, JI_{\rm M}e_h - I_{\rm M}e_h)$ and the stabilisation term is equal to $c_h(I_{\rm M}u - u, e_h)$. They are controlled by $|\!|\!|u - I_{\rm M}u|\!|\!|_{\rm pw}\|e_h\|_h$. The bilinear form b_h enjoys the miraculous property $b_h(I_{\rm M}u, I_{\rm M}e_h) = 0$ for the discontinuous Galerkin schemes of this paper. The remaining term on the right-hand side of (5.8) is $a_{\rm pw}(I_{\rm M}u, e_h - I_{\rm M}e_h) + b_h(I_{\rm M}u, e_h - I_{\rm M}e_h)$ and in fact controlled by the dcc (5.4). The proof of dcc in Section 8 is one key argument in this paper.

Remark 5.2. The arguments in this section applies to the case where $A_h(\bullet,\bullet)$ satisfies an inf-sup condition; (and not the coercivity condition). The key idea is to estimate the consistency error $F(JI_{\rm M}v_h) - A_h(I_{\rm M}u, v_h)$ using (5.5) and (5.6) and the orthogonality of the interpolation operator.

5.3. Transfer operators between V_h and $\mathrm{M}(\mathcal{T})$

Recall $I_{\rm M} : H^2(\mathcal{T}) \to \mathrm{M}(\mathcal{T})$ from Definition 3.5, and suppose the existence of some constant $\Lambda_{\rm M} \ge 0$ with

$$\|v_h - I_{\rm M}v_h\|_h \le \Lambda_{\rm M}\|v_h - v\|_h \qquad \text{for all } v_h \in V_h \text{ and all } v \in V. \tag{5.9}$$

Suppose the existence of constants $\Lambda'_{\rm M}, M_{\rm M} \ge 0$, and boundedness in the sense that

$$\begin{aligned} \|v_h - I_{\rm M}v_h\|_h &\le \Lambda'_{\rm M}\|v_h\|_h \qquad \text{for all } v_h \in V_h,\\ |\!|\!|I_{\rm M}v_h|\!|\!|_{\rm pw} &\le M_{\rm M}\|v_h\|_h \qquad \text{for all } v_h \in V_h. \end{aligned} \tag{5.10}$$

Apparently $\Lambda'_{\rm M} \le \Lambda_{\rm M}$ (with $v = 0$) and $M_{\rm M} \le 1 + \Lambda'_{\rm M}$ (with (5.2a) and a triangle inequality). The possibly smaller constant $M_{\rm M}$ enters in Theorem 5.4a, while $\Lambda_{\rm M}$ appears in Theorems 5.4b and 6.2 below. The above conditions control the transfer from V_h into $\mathrm{M}(\mathcal{T})$ via $I_{\rm M} : V + V_h + \mathrm{M}(\mathcal{T}) \to \mathrm{M}(\mathcal{T})$.

The transfer from $\mathrm{M}(\mathcal{T})$ into V_h is modeled by some linear map $I_h : \mathrm{M}(\mathcal{T}) \to V_h$ that is bounded in the sense that there exists some constant $\Lambda_h > 0$ such that

$$\|v_{\rm M} - I_h v_{\rm M}\|_h \le \Lambda_h |\!|\!|v_{\rm M} - v|\!|\!|_{\rm pw} \quad \text{for all } v_{\rm M} \in \mathrm{M}(\mathcal{T}) \text{ and for all } v \in V. \tag{5.11}$$

The examples of this paper concern the discrete norm from (4.1), (4.2) and then the estimates of this subsection follow for piecewise quadratic discrete spaces.

Example 5.3 (Equations (5.9) and (5.10) hold for $V_h \subseteq P_2(\mathcal{T})$ and (4.1), (4.2))**.** Suppose that the discrete norm $\|\bullet\|_h$ is defined by (4.1), (4.2) and $V_h \subseteq P_2(\mathcal{T})$. Then (5.2) and (5.9), (5.10) follow.

Proof of (5.9) *and* (5.10). Given any $v_2 \in P_2(\mathcal{T})$ and any $v \in V$, a triangle inequality shows

$$\|v_2 - I_{\mathrm{M}} v_2\|_h \le \|v_2 - JI_{\mathrm{M}} v_2\|_h + \|I_{\mathrm{M}} v_2 - JI_{\mathrm{M}} v_2\|_h =: t_1 + t_2.$$

Theorem 4.5c controls the first term $t_1 := \|v_2 - JI_{\mathrm{M}} v_2\|_h \lesssim \|v_2 - v\|_h$ on the right-hand side. Since $j_h(v_{\mathrm{M}}) = 0$ in (4.2) vanishes for $v_{\mathrm{M}} := I_{\mathrm{M}} v_2 \in \mathrm{M}(\mathcal{T})$, the second term $t_2 := |\!|\!| v_{\mathrm{M}} - Jv_{\mathrm{M}} |\!|\!|_{\mathrm{pw}} \lesssim |\!|\!| v_{\mathrm{M}} - v |\!|\!|_{\mathrm{pw}}$ with Lemma 3.7d in the last step. Since $\|(1 - \Pi_0) D^2_{\mathrm{pw}} v_2\| = 0$ vanishes for $v_2 \in P_2(\mathcal{T})$, Theorem 4.3a shows $|\!|\!| v_{\mathrm{M}} - v_2 |\!|\!|_{\mathrm{pw}} \lesssim j_h(v_2) = j_h(v_2 - v) \le \|v_2 - v\|_h$ with (4.1) and (4.2) in the last two steps. This and a triangle inequality prove $t_2 \lesssim \|v_2 - v\|_h$. The combination of the estimates for $t_1 + t_2 \lesssim \|v_2 - v\|_h$ proves (5.9); and (5.9) immediately implies (5.10) as discussed above. □

5.4. Sufficient conditions for best-approximation

The bilinear forms $A_h, a_{\mathrm{pw}}, b_h, c_h : (V_h + \mathrm{M}(\mathcal{T})) \times (V_h + \mathrm{M}(\mathcal{T})) \to \mathbb{R}$ in the discrete problem (5.3), (all bounded because $V_h + \mathrm{M}(\mathcal{T})$ is finite dimensional) read

$$A_h(\widehat{v}, \widehat{w}) := a_{\mathrm{pw}}(\widehat{v}, \widehat{w}) + b_h(\widehat{v}, \widehat{w}) + c_h(\widehat{v}, \widehat{w}) \quad \text{for all } \widehat{v}, \widehat{w} \in V_h + \mathrm{M}(\mathcal{T}). \tag{5.12}$$

The key assumption in abstract form is the *discrete consistency condition* with a constant $0 < \Lambda_{\mathrm{dc}} < \infty$: All functions $v_{\mathrm{M}} \in \mathrm{M}(\mathcal{T})$, $w_h \in V_h$, and all $v, w \in V$ satisfy

$$a_{\mathrm{pw}}(v_{\mathrm{M}}, w_h - I_{\mathrm{M}} w_h) + b_h(v_{\mathrm{M}}, w_h - I_{\mathrm{M}} w_h) \le \Lambda_{\mathrm{dc}} |\!|\!| v_{\mathrm{M}} - v |\!|\!|_{\mathrm{pw}} \|w_h - w\|_h. \tag{5.13}$$

(This is a straightforward generalization of (5.4) from Sect. 5.2.) Assume that $b_h : (V_h + \mathrm{M}(\mathcal{T})) \times (V_h + \mathrm{M}(\mathcal{T})) \to \mathbb{R}$ is bounded in $V_h + \mathrm{M}(\mathcal{T})$ by a constant $0 < M_{\mathrm{b}} < \infty$ and vanishes in $\mathrm{M}(\mathcal{T}) \times \mathrm{M}(\mathcal{T})$ in the sense that all $v_h, w_h \in V_h$ and all $v_{\mathrm{M}}, w_{\mathrm{M}} \in \mathrm{M}(\mathcal{T})$ satisfy

$$b_h(v_{\mathrm{M}}, w_{\mathrm{M}}) = 0, \tag{5.14}$$

$$b_h(v_h + v_{\mathrm{M}}, w_h + w_{\mathrm{M}}) \le M_{\mathrm{b}} \|v_h + v_{\mathrm{M}}\|_h \|w_h + w_{\mathrm{M}}\|_h. \tag{5.15}$$

Suppose that the bilinear form $c_h : (V_h + \mathrm{M}(\mathcal{T})) \times (V_h + \mathrm{M}(\mathcal{T})) \to \mathbb{R}$ and a constant $0 < \Lambda_{\mathrm{c}} < \infty$ satisfy

$$c_h(v_h, w_h) \le \Lambda_{\mathrm{c}} \|v - v_h\|_h \|w - w_h\|_h \quad \text{for all } v_h, w_h \in V_h \text{ and } v, w \in V. \tag{5.16}$$

Theorem 5.4 (Best-approximation)**.** *Suppose* (5.1), (5.2) *and* (5.9)–(5.16). *Let* $u \in V$ *solve* (3.1) *and let* $u_h \in V_h$ *solve* (5.3). *Then*

$$(a)\ \|u - u_h\|_h \le C_{\mathrm{qo}} |\!|\!| u - I_{\mathrm{M}} u |\!|\!|_{\mathrm{pw}} \text{ and } \ (b)\ |\!|\!| u - JI_{\mathrm{M}} u_h |\!|\!| \le (1 + \Lambda_{\mathrm{M}})(1 + \Lambda_{\mathrm{J}}) \|u - u_h\|_h$$

hold with the constant $C_{\mathrm{qo}} := 1 + \Lambda_h + \alpha^{-1}(\Lambda_0 M_{\mathrm{M}} + \Lambda_h(1 + M_{\mathrm{b}}) + \Lambda_{\mathrm{dc}} + \Lambda_{\mathrm{c}}(1 + \Lambda_h))$.

Remark 5.5. The Morley FEM is included in the (non-symmetric) abstract framework of Theorem 5.4 and leads to a sub-optimal best-approximation constant $C_{\mathrm{qo}} = 1 + \Lambda_0$.

The error analysis of a post-processing dates back at least to [6] with a design of an enrichment operator for C^0IP functions replaced here by the smoother JI_{M}. For $F \in H^{-s}(\Omega)$ with $2 - \sigma \le s \le 2$ (and $u \in H^{4-s}(\Omega)$ from elliptic regularity), Theorem 5.4 verifies

$$|\!|\!| u - JI_{\mathrm{M}} u_h |\!|\!| \le C_{\mathrm{qo}}(1 + \Lambda_{\mathrm{M}})(1 + \Lambda_{\mathrm{J}}) |\!|\!| u - I_{\mathrm{M}} u |\!|\!|_{\mathrm{pw}} \lesssim h_{\max}^{2-s} \|F\|_{H^{-s}(\Omega)}.$$

5.5. Proofs

Abbreviate $e_h := I_h I_{\mathrm{M}} u - u_h \in V_h$ and $(a_{\mathrm{pw}} + b_h)(\bullet,\bullet) := a_{\mathrm{pw}}(\bullet,\bullet) + b_h(\bullet,\bullet)$ for the sum of the bilinear forms.

Lemma 5.6 (Key identity)**.** *It holds*

$$\begin{aligned} A_h(e_h, e_h) &= a_{\mathrm{pw}}(u, (1-J)I_{\mathrm{M}} e_h) + (a_{\mathrm{pw}} + b_h)((I_h - 1)I_{\mathrm{M}} u, e_h) \\ &\quad + (a_{\mathrm{pw}} + b_h)(I_{\mathrm{M}} u, e_h - I_{\mathrm{M}} e_h) + c_h(I_h I_{\mathrm{M}} u, e_h). \end{aligned}$$

Proof. The test function $v := J I_{\mathrm{M}} e_h \in V$ in (3.1) and the test function $v_h := e_h \in V_h$ in (5.3) lead to

$$a(u, J I_{\mathrm{M}} e_h) = F(J I_{\mathrm{M}} e_h) = A_h(u_h, e_h).$$

This and the definition $e_h = I_h I_{\mathrm{M}} u - u_h$ result in

$$A_h(e_h, e_h) = A_h(I_h I_{\mathrm{M}} u, e_h) - a(u, J I_{\mathrm{M}} e_h).$$

The identity $a_{\mathrm{pw}}(u, I_{\mathrm{M}} e_h) = a_{\mathrm{pw}}(I_{\mathrm{M}} u, I_{\mathrm{M}} e_h)$ from (3.5) shows that this is equal to

$$a_{\mathrm{pw}}(u, (1-J)I_{\mathrm{M}} e_h) + A_h(I_h I_{\mathrm{M}} u, e_h) - a_{\mathrm{pw}}(I_{\mathrm{M}} u, I_{\mathrm{M}} e_h).$$

The last term $a_{\mathrm{pw}}(I_{\mathrm{M}} u, I_{\mathrm{M}} e_h)$ is part of $A_h(I_{\mathrm{M}} u, I_{\mathrm{M}} e_h)$ by (5.12), while $b_h(I_{\mathrm{M}} u, I_{\mathrm{M}} e_h) = 0$ owing to (5.14). This and elementary algebra conclude the proof. □

Lemma 5.7. *The assumptions* (3.5)*,* (3.9)*,* (3.11)*, and* (5.10) *imply*

$$a_{\mathrm{pw}}(u, (1-J)I_{\mathrm{M}} e_h) \le \Lambda_0 M_{\mathrm{M}} |\!|\!| u - I_{\mathrm{M}} u |\!|\!|_{\mathrm{pw}} \|e_h\|_h.$$

Proof. Set $v_{\mathrm{M}} := I_{\mathrm{M}} e_h$ and recall $I_{\mathrm{M}}(v_{\mathrm{M}} - J v_{\mathrm{M}}) = 0$ from (3.9). The orthogonality (3.5) and the Cauchy inequality with respect to $a_{\mathrm{pw}}(\bullet,\bullet)$ imply

$$a_{\mathrm{pw}}(u, (1-J)v_{\mathrm{M}}) = a_{\mathrm{pw}}(u - I_{\mathrm{M}} u, (1-J)v_{\mathrm{M}}) \le |\!|\!| u - I_{\mathrm{M}} u |\!|\!|_{\mathrm{pw}} |\!|\!| (1-J)v_{\mathrm{M}} |\!|\!|_{\mathrm{pw}}.$$

The boundedness of $1 - J$ in (3.11) and the boundedness of I_{M} in (5.10) show

$$|\!|\!| (1-J)v_{\mathrm{M}} |\!|\!|_{\mathrm{pw}} \le \Lambda_0 |\!|\!| v_{\mathrm{M}} |\!|\!|_{\mathrm{pw}} \le \Lambda_0 M_{\mathrm{M}} \|e_h\|_h.$$

The combination of the two displayed inequalities concludes the proof. □

Lemma 5.8. *The assumptions* (5.2)*,* (5.11)*, and* (5.15) *imply*

$$(a_{\mathrm{pw}} + b_h)((I_h - 1)I_{\mathrm{M}} u, e_h) \le \Lambda_h (1 + M_{\mathrm{b}}) |\!|\!| u - I_{\mathrm{M}} u |\!|\!|_{\mathrm{pw}} \|e_h\|_h.$$

Proof. For $e_h \in V_h$ and $I_{\mathrm{M}} u \in \mathrm{M}(\mathcal{T})$, the Cauchy inequality plus (5.2a) show

$$a_{\mathrm{pw}}((I_h - 1)I_{\mathrm{M}} u, e_h) \le \|(1 - I_h)I_{\mathrm{M}} u\|_h \|e_h\|_h.$$

This and the boundedness of b_h in (5.15) result in

$$(a_{\mathrm{pw}} + b_h)((I_h - 1)I_{\mathrm{M}} u, e_h) \le (1 + M_{\mathrm{b}}) \|(1 - I_h)I_{\mathrm{M}} u\|_h \|e_h\|_h.$$

The inequality (5.11) with $v_{\mathrm{M}} = I_{\mathrm{M}} u$ and $v = u$ reads

$$\|(1 - I_h)I_{\mathrm{M}} u\|_h \le \Lambda_h |\!|\!| u - I_{\mathrm{M}} u |\!|\!|_{\mathrm{pw}}.$$

The combination of the last two displayed inequalities concludes the proof. □

Lemma 5.9. *The assumptions* (5.2), (5.11), (5.13), *and* (5.16) *imply*

$$(a_{\rm pw}+b_h)(I_{\rm M}u,e_h-I_{\rm M}e_h)+c_h(I_hI_{\rm M}u,e_h)\le(\Lambda_{\rm dc}+\Lambda_{\rm c}(1+\Lambda_h))|\!|\!|u-I_{\rm M}u|\!|\!|_{\rm pw}\|e_h\|_h.$$

Proof. The discrete consistency condition (5.13) for $v_{\rm M}:=I_{\rm M}u$, $w_h:=e_h$, $v:=u$, and $w:=0$ lead to the upper bound

$$(a_{\rm pw}+b_h)(I_{\rm M}u,e_h-I_{\rm M}e_h)\le\Lambda_{\rm dc}|\!|\!|u-I_{\rm M}u|\!|\!|_{\rm pw}\|e_h\|_h$$

for the first term on the left-hand side of the asserted estimate. The remaining contribution $c_h(I_hI_{\rm M}u,e_h)$ is controlled in (5.16) with $v_h:=I_hI_{\rm M}u$, $w_h:=e_h$, $v=u$, and $w=0$ by

$$c_h(I_hI_{\rm M}u,e_h)\le\Lambda_{\rm c}\|u-I_hI_{\rm M}u\|_h\|e_h\|_h.$$

A triangle inequality in $\|\bullet\|_h$, (5.11) with $v_{\rm M}:=I_{\rm M}u$ and $v=u$, and (5.2b) show

$$\|u-I_hI_{\rm M}u\|_h\le(1+\Lambda_h)|\!|\!|u-I_{\rm M}u|\!|\!|_{\rm pw}.$$

The combination of the resulting inequalities concludes the proof. □

Proof of best-approximation in Theorem 5.4a. The discrete ellipticity (5.1) is followed by Lemma 5.6 with terms controlled in Lemmas 5.7–5.9. This leads (after a division by $\|e_h\|_h$, if positive) to

$$\alpha\|e_h\|_h\le(\Lambda_0M_{\rm M}+\Lambda_h(1+M_{\rm b})+\Lambda_{\rm dc}+\Lambda_{\rm c}(1+\Lambda_h))|\!|\!|u-I_{\rm M}u|\!|\!|_{\rm pw}.$$

On the other hand, $\|(I_h-1)I_{\rm M}u\|_h\le\Lambda_h|\!|\!|u-I_{\rm M}u|\!|\!|_{\rm pw}$ from (5.11) for $v_{\rm M}:=I_{\rm M}u$ and $v:=u$. Triangle inequalities in $\|\bullet\|_h$, (5.2b), and the last two inequalities result in

$$\|u-u_h\|_h\le|\!|\!|u-I_{\rm M}u|\!|\!|_{\rm pw}+\|I_{\rm M}u-I_hI_{\rm M}u\|_h+\|e_h\|_h\le C_{\rm qo}|\!|\!|u-I_{\rm M}u|\!|\!|_{\rm pw}$$

with the constant $C_{\rm qo}$ displayed in the assertion. □

Proof for post-processing in Theorem 5.4b. The assertion (b) is formulated in terms of u and u_h but holds for general $v_h\in V_h$ and $v\in V$ and the abbreviation $v_{\rm M}:=I_{\rm M}v_h\in M(\mathcal{T})$. A triangle inequality and Lemma 3.7d prove

$$|\!|\!|v-Jv_{\rm M}|\!|\!|\le|\!|\!|v-v_{\rm M}|\!|\!|_{\rm pw}+|\!|\!|(1-J)v_{\rm M}|\!|\!|_{\rm pw}\le(1+\Lambda_{\rm J})|\!|\!|v-v_{\rm M}|\!|\!|_{\rm pw}.$$

A triangle inequality, (5.2a) twice, and (5.9) show

$$|\!|\!|v-v_{\rm M}|\!|\!|_{\rm pw}\le|\!|\!|v-v_h|\!|\!|_{\rm pw}+\|v_h-v_{\rm M}\|_h\le(1+\Lambda_{\rm M})\|v-v_h\|_h.$$

The combination of the two displayed estimates reads

$$|\!|\!|v-Jv_{\rm M}|\!|\!|\le(1+\Lambda_{\rm J})(1+\Lambda_{\rm M})\|v-v_h\|_h.$$

□

6. Weaker and piecewise Sobolev norm error estimates

6.1. Assumptions and result

This subsection presents one further condition sufficient for a lower-order *a priori* error estimate for the discrete problem (5.3) beyond the hypotheses of Sections 5.1–5.4: The *dual discrete consistency* with a constant $0\le\Lambda_{\rm ddc}<\infty$ asserts that any $v_h\in V_h$, $w_{\rm M}\in{\rm M}(\mathcal{T})$, and any $v,w\in V$ satisfy

$$a_{\rm pw}(v_h-I_{\rm M}v_h,w_{\rm M})+b_h(v_h,w_{\rm M})\le\Lambda_{\rm ddc}\|v-v_h\|_h|\!|\!|w-w_{\rm M}|\!|\!|_{\rm pw}.\tag{6.1}$$

Remark 6.1 (Symmetry)**.** If the bilinear form $(a_{\rm pw}+b_h)(\bullet,\bullet)$ is symmetric, then (5.13) and (5.14) imply (6.1). (Rewrite the left-hand side in (6.1) with (5.14) and symmetry into the left-hand side of (5.13) with v_h and $w_{\rm M}$ replacing w_h and $v_{\rm M}$ to establish (6.1).)

Since $u_h \in V_h$ may not belong to $H^s(\Omega)$ for $2-\sigma \le s \le 2$ in general, the post-processing $JI_{\rm M}u_h \in V$ arises in the duality argument with $0<\sigma\le 1$ from Example 3.1.

Theorem 6.2 (Lower-order error estimates)**.** *Suppose the assumptions of Theorem 5.4, (6.1), and $2-\sigma \le s \le 2$. Then there exist constants $C_4(s), C_5(s) > 0$ such that (a) and (b) hold for any $F \in H^{-s}(\Omega)$ with solution $u \in V$ to (3.1) and the solution $u_h \in V_h$ to (5.3). (a) $\|u-JI_{\rm M}u_h\|_{H^s(\Omega)} \le C_4](s)h_{\max}^{2-s}\|u-u_h\|_h$ and (b), if $u_h \in P_2(\mathcal{T})$, then $\|u-u_h\|_{H^s(\mathcal{T})} \le C_5(s)h_{\max}^{2-s}\|u-u_h\|_h$.*

6.2. Duality and algebra

The duality of $H^{-s}(\Omega)$ and $H_0^s(\Omega)$ reveals for the exact solution $u \in V$ to (3.1) and the post-processing $v := JI_{\rm M}u_h \in V$ of the discrete solution $u_h \in V_h$ to (5.3) that

$$\|u-v\|_{H^s(\Omega)} = \sup_{0\ne G\in H^{-s}(\Omega)} \frac{G(u-v)}{\|G\|_{H^{-s}(\Omega)}} = G(u-v).$$

The supremum is attained for some $G \in H^{-s}(\Omega) \subset V^*$ with norm $\|G\|_{H^{-s}(\Omega)} = 1$ owing to a corollary of the Hahn–Banach theorem. The functional $a(z,\bullet) = G \in V^*$ has a unique Riesz representation $z \in V$ in the Hilbert space (V,a); $z \in V$ is the weak solution to the PDE $\Delta^2 z = G$. The elliptic regularity (as in Example 3.1) leads to $z \in V \cap H^{4-s}(\Omega)$ with $2 \le 4-s \le 2+\sigma$ and (3.2); hence

$$\|u-v\|_{H^s(\Omega)} = a(u-v,z) \quad \text{and} \quad \|z\|_{H^{4-s}(\Omega)} \le C_{\rm reg}.$$

The proof of Theorem 6.2 consists of a series of lemmas to establish an upper bound of $a(u-v,z)$ for the above $z \in V \cap H^{4-s}(\Omega)$. The notation

$$v := JI_{\rm M}u_h \in V, \qquad z_h := I_hI_{\rm M}z \in V_h \quad \text{and} \quad \zeta := JI_{\rm M}z_h \in V$$

for the discrete, exact, and dual solution u_h, u, and z applies throughout this section.

Lemma 6.3 (Key identity)**.** *It holds*

$$\begin{aligned} a(u-v,z) &= a(u-v,z-\zeta) + a_{\rm pw}(u_h-v,\zeta-z_h) + a_{\rm pw}(I_{\rm M}u_h-v,z_h-I_{\rm M}z_h) \\ &\quad + a_{\rm pw}(u_h-I_{\rm M}u_h,I_{\rm M}z_h-\zeta) + A_h(u_h,z_h) - a_{\rm pw}(u_h,I_{\rm M}z_h) + a_{\rm pw}(u_h-I_{\rm M}u_h,z_h). \end{aligned}$$

Proof. Let $\zeta \equiv JI_{\rm M}z_h \in V$ substitute the test function v in (3.1). This and the test function $v_h := z_h$ in (5.3) lead to

$$a(u,\zeta) = F(\zeta) = A_h(u_h,z_h).$$

This identity and elementary algebra result in

$$\begin{aligned} a(u-v,\zeta) &= a_{\rm pw}(u_h-v,\zeta-z_h) + a_{\rm pw}(I_{\rm M}u_h-v,z_h) + a_{\rm pw}(u_h,(1-J)I_{\rm M}z_h) \\ &\quad + A_h(u_h,z_h) - a_{\rm pw}(u_h,I_{\rm M}z_h) + a_{\rm pw}(u_h-I_{\rm M}u_h,z_h). \end{aligned}$$

The formulas $a_{\rm pw}((1-J)I_{\rm M}u_h,I_{\rm M}z_h) = 0 = a_{\rm pw}(I_{\rm M}u_h,(1-J)I_{\rm M}z_h)$ (from (3.5) and (3.9)) and elementary algebra conclude the proof. □

6.3. Elementary bounds

Lemma 6.4. *Each of the following terms (a) $a(u-v, z-\zeta)$, (b) $a_{\rm pw}(u_h-v, \zeta-z_h)$, (c) $a_{\rm pw}(I_{\rm M}u_h-v, z_h-I_{\rm M}z_h)$, and (d) $a_{\rm pw}(u_h-I_{\rm M}u_h, I_{\rm M}z_h-\zeta)$ is bounded in modulus by a constant $\le (1+(1+\Lambda_{\rm J})(1+\Lambda_{\rm M}))^2(1+\Lambda_h)$ times $\|u-u_h\|_h |\!|\!| z-I_{\rm M}z |\!|\!|_{\rm pw}$.*

Proof. The assumption (5.9) (with (v, v_h) replaced by (u, u_h) and (z, z_h)) implies

$$\|u_h - I_{\rm M}u_h\|_h \le \Lambda_{\rm M}\|u-u_h\|_h \quad \text{and} \quad \|z_h - I_{\rm M}z_h\|_h \le \Lambda_{\rm M}\|z-z_h\|_h. \tag{6.2}$$

Recall $v \equiv JI_{\rm M}u_h$ and $\zeta \equiv JI_{\rm M}z_h \in V$ to deduce from Lemma 3.7d that

$$|\!|\!| v - I_{\rm M}u_h |\!|\!|_{\rm pw} \le \Lambda_{\rm J} |\!|\!| u - I_{\rm M}u_h |\!|\!|_{\rm pw} \quad \text{and} \quad |\!|\!| \zeta - I_{\rm M}z_h |\!|\!|_{\rm pw} \le \Lambda_{\rm J} |\!|\!| z - I_{\rm M}z_h |\!|\!|_{\rm pw}. \tag{6.3}$$

The combination of (6.2)-(6.3) with (5.2a) and triangle inequalities lead to

$$\begin{aligned} |\!|\!| u-v |\!|\!| &\le |\!|\!| u - I_{\rm M}u_h |\!|\!|_{\rm pw} + |\!|\!| v - I_{\rm M}u_h |\!|\!|_{\rm pw} \le (1+\Lambda_{\rm J}) |\!|\!| u - I_{\rm M}u_h |\!|\!|_{\rm pw} \\ &\le (1+\Lambda_{\rm J})(|\!|\!| u-u_h |\!|\!|_{\rm pw} + \|u_h - I_{\rm M}u_h\|_h) \le (1+\Lambda_{\rm J})(1+\Lambda_{\rm M})\|u-u_h\|_h. \end{aligned} \tag{6.4}$$

The above arguments have not utilized any solution property and hence also apply for (z, ζ, z_h) replacing (u, v, u_h) to reveal (instead of (6.4))

$$|\!|\!| z - \zeta |\!|\!| \le (1+\Lambda_{\rm J})(1+\Lambda_{\rm M})\|z-z_h\|_h. \tag{6.5}$$

Consider $v_{\rm M} := I_{\rm M}z \in {\rm M}(\mathcal{T})$ with $z_h \equiv I_h I_{\rm M} z = I_h v_{\rm M}$ in (5.11) to show

$$\|z_h - I_{\rm M}z\|_h \le \Lambda_h |\!|\!| z - I_{\rm M}z |\!|\!|_{\rm pw}. \tag{6.6}$$

This, a triangle inequality, and (5.2b) result in

$$\|z-z_h\|_h \le (1+\Lambda_h) |\!|\!| z - I_{\rm M}z |\!|\!|_{\rm pw}. \tag{6.7}$$

The combination of (6.5) and (6.7) proves

$$|\!|\!| z - \zeta |\!|\!| \le (1+\Lambda_{\rm J})(1+\Lambda_{\rm M})(1+\Lambda_h) |\!|\!| z - I_{\rm M}z |\!|\!|_{\rm pw}. \tag{6.8}$$

Proof of (a). This follows from a Cauchy inequality plus (6.4) and (6.8). □

Proof of (b). A triangle inequality, (5.2a), and (6.4) verify

$$|\!|\!| u_h - v |\!|\!|_{\rm pw} \le \|u_h - v\|_h \le (1+(1+\Lambda_{\rm J})(1+\Lambda_{\rm M}))\|u-u_h\|_h.$$

The triangle inequality with (5.2a) and (6.7)-(6.8) show

$$|\!|\!| \zeta - z_h |\!|\!|_{\rm pw} \le \|\zeta - z_h\|_h \le (1+\Lambda_h)(1+(1+\Lambda_{\rm J})(1+\Lambda_{\rm M})) |\!|\!| z - I_{\rm M}z |\!|\!|_{\rm pw}.$$

A Cauchy inequality and the preceding estimates conclude the proof of (b). □

Proof of (c). The estimate (6.3), a triangle inequality, (5.2a), and (6.2) show

$$|\!|\!| v - I_{\rm M}u_h |\!|\!|_{\rm pw} \le \Lambda_{\rm J} |\!|\!| u - I_{\rm M}u_h |\!|\!|_{\rm pw} \le \Lambda_{\rm J}(1+\Lambda_{\rm M})\|u-u_h\|_h. \tag{6.9}$$

The combination of (6.2) and (6.7) after (5.2a) leads to

$$|\!|\!| z_h - I_{\rm M}z_h |\!|\!|_{\rm pw} \le \|z_h - I_{\rm M}z_h\|_h \le \Lambda_{\rm M}(1+\Lambda_h) |\!|\!| z - I_{\rm M}z |\!|\!|_{\rm pw}. \tag{6.10}$$

A Cauchy inequality and the preceding estimates conclude the proof of (c). □

Proof of (d). This follows from a Cauchy inequality with (5.2a) and the estimates for $\|u_h - I_{\rm M}u_h\|_h$ in (6.2) and $|\!|\!| \zeta - I_{\rm M}z_h |\!|\!|_{\rm pw}$ in (6.3). □

6.4. Discrete consistency bounds

Lemma 6.5. *It holds*

$$\begin{aligned} &A_h(u_h, z_h) - a_{\rm pw}(u_h, I_{\rm M} z_h) + a_{\rm pw}(u_h - I_{\rm M} u_h, z_h) \\ &\quad \le (1+\Lambda_h)((2+M_{\rm b})\Lambda_{\rm M}^2 + (\Lambda_{\rm dc}+\Lambda_{\rm ddc})(1+\Lambda_{\rm M})\Lambda_{\rm J} + \Lambda_{\rm c})\|u-u_h\|_h |\!|\!| z - I_{\rm M} z |\!|\!|_{\rm pw}. \end{aligned}$$

Proof. Recall $b_h(I_{\rm M} u_h, I_{\rm M} z_h) = 0$ from (5.14) and exploit (5.12) with elementary (but lengthy) algebra to check that the left-hand side LHS of the assertion is equal to

$$\begin{aligned} \text{LHS} = {}& (2a_{\rm pw} + b_h)(u_h - I_{\rm M} u_h, z_h - I_{\rm M} z_h) + (a_{\rm pw} + b_h)(I_{\rm M} u_h, z_h - I_{\rm M} z_h) && (6.11) \\ &+ (a_{\rm pw}(u_h - I_{\rm M} u_h, I_{\rm M} z_h) + b_h(u_h, I_{\rm M} z_h)) + c_h(u_h, z_h) && (6.12) \end{aligned}$$

with the short notation, *e.g.*, $(2a_{\rm pw} + b_h)(\bullet,\bullet) := 2a_{\rm pw}(\bullet,\bullet) + b_h(\bullet,\bullet)$, for the sum of the bilinear forms announced in Section 5.5. The two lines (6.11) and (6.12) of expressions for the LHS give rise to four estimates. The continuity of $a_{\rm pw}(\bullet,\bullet)$ and $b_h(\bullet,\bullet)$ in (5.15), (6.2), and (6.10) prove

$$(2a_{\rm pw} + b_h)(u_h - I_{\rm M} u_h, z_h - I_{\rm M} z_h) \le (2+M_{\rm b})\Lambda_{\rm M}^2(1+\Lambda_h)\|u-u_h\|_h |\!|\!| z - I_{\rm M} z |\!|\!|_{\rm pw}.$$

The discrete consistency (5.13) leads in the last term in (6.11) to a product of $|\!|\!| v - I_{\rm M} u_h |\!|\!|_{\rm pw}$ controlled in (6.9) and $\|z - z_h\|_h$ controlled in (6.7). This results in

$$(a_{\rm pw} + b_h)(I_{\rm M} u_h, z_h - I_{\rm M} z_h) \le \Lambda_{\rm dc}\Lambda_{\rm J}(1+\Lambda_{\rm M})(1+\Lambda_h)\|u-u_h\|_h |\!|\!| z - I_{\rm M} z |\!|\!|_{\rm pw}.$$

The dual discrete consistency in (6.1) applies to the first two terms in (6.12) and leads to $\Lambda_{\rm ddc}\|u-u_h\|_h$ times $|\!|\!| \zeta - I_{\rm M} z_h |\!|\!|_{\rm pw}$ controlled with (5.2b) in (6.3). This with (6.2b) and (6.7) result in

$$a_{\rm pw}(u_h - I_{\rm M} u_h, I_{\rm M} z_h) + b_h(u_h, I_{\rm M} z_h) \le \Lambda_{\rm ddc}\Lambda_{\rm J}(1+\Lambda_{\rm M})(1+\Lambda_h)\|u-u_h\|_h |\!|\!| z - I_{\rm M} z |\!|\!|_{\rm pw}.$$

The last term in (6.12) is controlled in (5.16). This and (6.7) show

$$c_h(u_h, z_h) \le \Lambda_{\rm c}\|u-u_h\|_h\|z-z_h\|_h \le \Lambda_{\rm c}(1+\Lambda_h)\|u-u_h\|_h |\!|\!| z - I_{\rm M} z |\!|\!|_{\rm pw}.$$

A combination of the preceding four estimates with (6.11) and (6.12) concludes the proof. □

6.5. Proof of Theorem 6.2

Given $2-\sigma \le s \le 2$, there exists a constant $0 < C_{\rm int}(s) < \infty$ (which exclusively depends on the shape regularity of $\mathcal{T}$ and s) such that the solution $z \in V$ of the dual problem in Section 6.2 satisfies (with Lem. 3.2c) that

$$|\!|\!| z - I_{\rm M} z |\!|\!|_{\rm pw} \le C_{\rm int}(s) h_{\max}^{2-s}\|z\|_{H^{4-s}(\Omega)} \le C_{\rm int}(s) C_{\rm reg}(s) h_{\max}^{2-s}\|G\|_{H^{-s}(\Omega)}. \tag{6.13}$$

Proof of (a). Recall $\|u-v\|_{H^s(\Omega)} = a(u-v,z)$ from Section 6.2 and its formula in Lemma 6.3. Lemma 6.4 applies to the first four terms and Lemma 6.5 to the remaining three. The resulting estimate reads

$$\|u-v\|_{H^s(\Omega)} \le C_6\|u-u_h\|_h |\!|\!| z - I_{\rm M} z |\!|\!|_{\rm pw}$$

with $C_6 = (1+\Lambda_h)\big(4(1+(1+\Lambda_{\rm J})(1+\Lambda_{\rm M}))^2 + (2+M_{\rm b})\Lambda_{\rm M}^2 + (\Lambda_{\rm dc}+\Lambda_{\rm ddc})(1+\Lambda_{\rm M})\Lambda_{\rm J} + \Lambda_{\rm c}\big)$. This and $|\!|\!| z - I_{\rm M} z |\!|\!|_{\rm pw} \le C_{\rm int}(s) C_{\rm reg}(s) h_{\max}^{2-s}$ from (6.13) prove Theorem 6.2a. □

Proof of (b). The norm in $H^s(\mathcal{T}) = \prod_{T\in\mathcal{T}} H^s(T)$ is the ℓ^2 norm of those contributions $\|\bullet\|_{H^s(T)}$ for all $T \in \mathcal{T}$. The Sobolev–Slobodeckii semi-norm over Ω involves double integrals over $\Omega \times \Omega$ and so is larger than or equal

to the sum of the contributions over $T \times T$ for all the triangles $T \in \mathcal{T}$, *i.e.*, $\sum_{T\in\mathcal{T}} |\bullet|^2_{H^s(T)} \le |\bullet|^2_{H^s(\Omega)}$ for any $1 < s < 2$. Hence Theorem 6.2a implies

$$\|u - JI_{\mathrm{M}}u_h\|_{H^s(\mathcal{T})} \le C_4(s)\, h^{2-s}_{\max}\|u - u_h\|_h \quad \text{for all } s \text{ with } 2-\sigma \le s \le 2. \tag{6.14}$$

Since $u_h \in P_2(\mathcal{T})$, Theorem 4.5d provides the estimate

$$\|u_h - JI_{\mathrm{M}}u_h\|_{H^s(\mathcal{T})} \lesssim h^{2-s}_{\max}\|u - u_h\|_h.$$

The triangle inequality in the norm of $H^s(\mathcal{T})$ concludes the proof of Theorem 6.2b. □

6.6. Verification of (H) and $\widehat{\mathbf{(H)}}$

For the choice $P = Q = J \circ I_{\mathrm{M}}$, Theorem 4.5c shows that (2.3)–(2.5) hold for all the lowest-order schemes $V_h \subseteq P_2(\mathcal{T})$ considered in this paper. This subsection verifies **(H)** and $\widehat{\mathbf{(H)}}$.

Lemma 6.6 (Verification of **(H)** and $\widehat{\mathbf{(H)}}$)**.** *Suppose the assumptions of Theorems 5.4 and 6.2. Any $v_h \in V_h$, $v := JI_{\mathrm{M}}v_h \in V$, $w_h \in V_h$, and $w = JI_{\mathrm{M}}w_h \in V$ satisfy*

$$\begin{aligned}
&\mathbf{(H)}\ A_h(v_h, w_h) - a(v, w) \le \Lambda_1\|v_h - v\|_h\|w_h\|_h,\\
&\widehat{\mathbf{(H)}}\ A_h(v_h, w_h) - a(v, w) \le \Lambda_2\|v_h - v\|_h\|w_h - w\|_h
\end{aligned}$$

with $\Lambda_1 := (1+\Lambda_{\mathrm{M}})(\Lambda_{\mathrm{dc}} + \Lambda_{\mathrm{J}}M_{\mathrm{M}}\|J\|) + (1+M_{\mathrm{b}})\Lambda_{\mathrm{M}} + \Lambda_{\mathrm{c}}$, *and* $\Lambda_2 := \Lambda_{\mathrm{ddc}} + (1+M_{\mathrm{b}})\Lambda_{\mathrm{M}}^2 + (1+\Lambda_{\mathrm{M}})(1+\Lambda_{\mathrm{M}} + \Lambda_{\mathrm{dc}}) + \Lambda_{\mathrm{c}}$.

Proof of **(H)**. For $w_{\mathrm{M}} := I_{\mathrm{M}}w$, and $v_{\mathrm{M}} := I_{\mathrm{M}}v$, (3.9) implies $w_{\mathrm{M}} = I_{\mathrm{M}}w_h$, and $v_{\mathrm{M}} = I_{\mathrm{M}}v_h$. The definition of $A_h(\bullet,\bullet)$, algebraic manipulations, and (5.14) result in

$$\begin{aligned}
A_h(v_h, w_h) - a(v, w) &= a_{\mathrm{pw}}(v_h, w_h) + b_h(v_h, w_h) + c_h(v_h, w_h) - a(v, w)\\
&= a_{\mathrm{pw}}(v_h - v_{\mathrm{M}}, w_h) + b_h(v_h - v_{\mathrm{M}}, w_h) + a_{\mathrm{pw}}(v_{\mathrm{M}}, w_h - w_{\mathrm{M}}) + b_h(v_{\mathrm{M}}, w_h - w_{\mathrm{M}})\\
&\quad + c_h(v_h, w_h) + a_{\mathrm{pw}}(v_{\mathrm{M}}, w_{\mathrm{M}}) - a(v, w).
\end{aligned} \tag{6.15}$$

The boundedness of $a_{\mathrm{pw}}(\bullet,\bullet)$, (5.15), (5.9), and (5.2) prove

$$a_{\mathrm{pw}}(v_h - v_{\mathrm{M}}, w_h) + b_h(v_h - v_{\mathrm{M}}, w_h) \le (1+M_{\mathrm{b}})\Lambda_{\mathrm{M}}\|v_h - v\|_h\|w_h\|_h.$$

The discrete consistency condition (5.13) (with $w = 0$), a triangle inequality, (5.9), and (5.2) show

$$a_{\mathrm{pw}}(v_{\mathrm{M}}, w_h - w_{\mathrm{M}}) + b_h(v_{\mathrm{M}}, w_h - w_{\mathrm{M}}) \le \Lambda_{\mathrm{dc}}(1+\Lambda_{\mathrm{M}})\|v_h - v\|_h\|w_h\|_h.$$

The bound in (5.16) with the choice $w = 0$ implies

$$c_h(v_h, w_h) \le \Lambda_{\mathrm{c}}\|v_h - v\|_h\|w_h\|_h.$$

The orthogonality condition (3.5), Lemma 3.7d, (5.9), and (5.10), result in

$$a_{\mathrm{pw}}(v_{\mathrm{M}}, w_{\mathrm{M}}) - a(v, w) = a_{\mathrm{pw}}((1-J)v_{\mathrm{M}}, Jw_{\mathrm{M}}) \le \Lambda_{\mathrm{J}}(1+\Lambda_{\mathrm{M}})M_{\mathrm{M}}\|J\|\|v_h - v\|_h\|w_h\|_h.$$

A combination of the last four displayed estimates in (6.15) leads to the desired result. □

Proof of $\widehat{\mathbf{(H)}}$. An alternate split of the left-hand side of the desired estimate leads to

$$\begin{aligned} A_h(v_h, w_h) - a(v, w) = {} & a_{\mathrm{pw}}(v_h - v_{\mathrm{M}}, w_{\mathrm{M}}) + b_h(v_h, w_{\mathrm{M}}) + a_{\mathrm{pw}}(v_h - v_{\mathrm{M}}, w_h - w_{\mathrm{M}}) \\ & + b_h(v_h - v_{\mathrm{M}}, w_h - w_{\mathrm{M}}) + a_{\mathrm{pw}}(v_{\mathrm{M}}, w_h - w_{\mathrm{M}}) + b_h(v_{\mathrm{M}}, w_h - w_{\mathrm{M}}) \\ & + a_{\mathrm{pw}}(v_{\mathrm{M}}, w_{\mathrm{M}}) - a(v, w) + c_h(v_h, w_h). \end{aligned} \tag{6.16}$$

The discrete consistency condition (6.1) shows

$$a_{\mathrm{pw}}(v_h - v_{\mathrm{M}}, w_{\mathrm{M}}) + b_h(v_h, w_{\mathrm{M}}) \le \Lambda_{\mathrm{ddc}} \|v - v_h\|_h |\!|\!| w - w_{\mathrm{M}} |\!|\!|_{\mathrm{pw}}.$$

The boundedness of $a_{\mathrm{pw}}(\bullet, \bullet)$, (5.15), and (5.9) prove

$$a_{\mathrm{pw}}(v_h - v_{\mathrm{M}}, w_h - w_{\mathrm{M}}) + b_h(v_h - v_{\mathrm{M}}, w_h - w_{\mathrm{M}}) \le (1 + M_{\mathrm{b}}) \Lambda_{\mathrm{M}}^2 \|v_h - v\|_h \|w_h - w\|_h.$$

The discrete consistency condition (5.13) and (5.2) lead to

$$a_{\mathrm{pw}}(v_{\mathrm{M}}, w_h - w_{\mathrm{M}}) + b_h(v_{\mathrm{M}}, w_h - w_{\mathrm{M}}) \le \Lambda_{\mathrm{dc}}(1 + \Lambda_{\mathrm{M}}) \|v_h - v\|_h \|w_h - w\|_h.$$

The orthogonality condition (3.5) and (5.9) result in

$$a_{\mathrm{pw}}(v_{\mathrm{M}}, w_{\mathrm{M}}) - a(v, w) = -a_{\mathrm{pw}}(v - v_{\mathrm{M}}, w - w_{\mathrm{M}}) \le (1 + \Lambda_{\mathrm{M}})^2 \|v_h - v\|_h \|w_h - w\|_h.$$

The bound in (5.16) implies

$$c_h(v_h, w_h) \le \Lambda_{\mathrm{c}} \|v_h - v\|_h \|w_h - w\|_h.$$

A substitution of the last five displayed estimates in (6.16) leads to the desired result. □

Remark 6.7 (Thm. 2.6 implies Thm. 6.2)**.** For $v := JI_{\mathrm{M}} u_h \in V$, recall $\|u - v\|_{H^s(\Omega)} = a(u - v, z)$ from Section 6.2. Theorem 2.6 applies as Lemma 6.6 holds and with (5.2) leads to $a(u - v, z) \le \widehat{C_{\mathrm{qo}}} \|u - u_h\|_h \|z - I_{\mathrm{M}} z\|_h$.

7. Modified dGFEM

The bilinear form $A_h(\bullet, \bullet) := A_{\mathrm{dG}}(\bullet, \bullet)$ [2, 28] is defined, for all $v_2, w_2 \in V_h := P_2(\mathcal{T})$, by

$$A_{\mathrm{dG}}(v_2, w_2) := a_{\mathrm{pw}}(v_2, w_2) + b_h(v_2, w_2) + c_{\mathrm{dG}}(v_2, w_2), \tag{7.1}$$

$$b_h(v_2, w_2) := -\Theta \mathcal{J}(v_2, w_2) - \mathcal{J}(w_2, v_2), \tag{7.2a}$$

$$\mathcal{J}(v_2, w_2) := \sum_{E \in \mathcal{E}} \int_E [\nabla_{\mathrm{pw}} v_2]_E \cdot \langle D_{\mathrm{pw}}^2 w_2 \rangle_E \nu_E \, \mathrm{ds} \tag{7.2b}$$

with $c_{\mathrm{dG}}(\bullet, \bullet)$ from (4.3) and given $-1 \le \Theta \le 1$. Let the jumps $[\bullet]_E$ across and the averages $\langle \bullet \rangle_E$ at an edge $E \in \mathcal{E}$ from Section 4.1 act componentwise. Recall from Theorem 4.1 the equivalent discrete norms $\| \bullet \|_h \approx \| \bullet \|_{\mathrm{dG}}$ in $H^2(\mathcal{T})$ defined in (4.1), (4.2) and (4.4). Set $\Theta = 1$ (resp. $\Theta = -1$) to obtain the symmetric (resp. non symmetric) interior penalty Galerkin formulation; see [39] for an alternative formulation. Appropriate positive parameters σ_1, σ_2 in (4.3) guarantee (5.1).

Lemma 7.1 (Boundedness and ellipticity of $A_{\mathrm{dG}}(\bullet, \bullet)$ [28, 39])**.** *(a) Any $v_2, w_2 \in P_2(\mathcal{T})$ satisfy $A_{\mathrm{dG}}(v_2, w_2) \lesssim \|v_2\|_{\mathrm{dG}} \|w_2\|_{\mathrm{dG}}$. (b) For $\Theta = -1$ and any $\sigma_{\mathrm{dG}} = \sigma_1 = \sigma_2 > 0$, $\|v_2\|_{\mathrm{dG}}^2 \le A_{\mathrm{dG}}(v_2, v_2)$ holds for all $v_2 \in P_2(\mathcal{T})$. (c) For $-1 < \Theta \le 1$ and a sufficiently large parameter $\sigma_{\mathrm{dG}} = \sigma_1 = \sigma_2 > 0$, there exists $\alpha > 0$ (which depends on σ_{dG} and the shape regularity of $\mathcal{T}$) such that $\alpha \|v_2\|_{\mathrm{dG}}^2 \le A_{\mathrm{dG}}(v_2, v_2)$ for all $v_2 \in P_2(\mathcal{T})$.*

Throughout this paper, the parameter $\sigma_{\rm dG}$ is chosen to guarantee the ellipticity of $A_{\rm dG}(\bullet,\bullet)$ in Lemma 7.1 with the short notation $\sigma_{\rm dG}\approx 1\approx\alpha$. The modified dGFEM (5.3) seeks the solution $u_{\rm dG}\in P_2(\mathcal{T})$ to

$$A_{\rm dG}(u_{\rm dG},v_2)=F(JI_{\rm M}v_2)\quad\text{for all } v_2\in P_2(\mathcal{T}). \tag{7.3}$$

Theorem 7.2 (Error estimates)**.** *The solution $u\in V$ to (3.1) and the solution $u_{\rm dG}\in P_2(\mathcal{T})$ to (7.3) satisfy (a) $\|u-u_{\rm dG}\|_h\lesssim |\!|\!| u-I_{\rm M}u|\!|\!|_{\rm pw}$ and (b) if $\Theta=1$ and $F\in H^{-s}(\Omega)$ for $2-\sigma\le s\le 2$, then $\|u-JI_{\rm M}u_{\rm dG}\|_{H^s(\Omega)}+\|u-u_{\rm dG}\|_{H^s(\mathcal{T})}\lesssim h_{\max}^{2-s}\|u-u_{\rm dG}\|_h$.*

Overview of the proof. The assertion (a) follows from Theorem 5.4 for the particular spaces, operators, norms, and bilinear forms defined below. The application of Theorem 5.4 requires the proof of the abstract conditions (5.1) and (5.2), (5.9)–(5.16). The assertion (b) follows from Theorem 6.2 provided (6.1) holds.

Setting and first consequences. Recall $V_h:=P_2(\mathcal{T})$ and the norms $\|\bullet\|_h$ and $\|\bullet\|_{\rm dG}$ in (4.1), (4.2) and (4.4). Recall the Morley interpolation operator $I_{\rm M}$ from Definition 3.5 and the companion operator J from Lemma 3.7. Recall that Lemma 7.1 guarantees (5.1) and (5.2). The dGFEM in (7.3) corresponds to (5.3) with the solution $u_h:=u_{\rm dG}$. Example 5.3 implies (5.9) and (5.10). Set $I_h:=\mathrm{id}$ and observe (5.11) holds for $\Lambda_h=0$. Recall $A_h(\bullet,\bullet):=A_{\rm dG}(\bullet,\bullet)$, $b_h(\bullet,\bullet):=-\Theta\mathcal{J}(\bullet,\bullet)-\mathcal{J}^*(\bullet,\bullet)$, and $c_h(\bullet,\bullet):=c_{\rm dG}(\bullet,\bullet)$ in (5.12).

Proof of (5.13). Since the integral $\int_E[\nabla_{\rm pw}v_{\rm M}]_E\,{\rm ds}=0$ vanishes for $v_{\rm M}\in{\rm M}(\mathcal{T})$ and since $\langle D^2_{\rm pw}w_2\rangle_E$ is constant on any edge $E\in\mathcal{E}$ for any $w_2\in P_2(\mathcal{T})$,

$$\mathcal{J}(v_{\rm M},w_2)=\sum_{E\in\mathcal{E}}\int_E[\nabla_{\rm pw}v_{\rm M}]_E\cdot\langle D^2_{\rm pw}w_2\rangle_E\nu_E\,{\rm ds}=0. \tag{7.4}$$

Hence the term $\Theta\mathcal{J}(v_{\rm M},w_2-I_{\rm M}w_2)$ disappears below in definitions of $(a_{\rm pw}+b_h)(\bullet,\bullet)$, written in the short notation of Section 5.5; $(a_{\rm pw}+b_h)(v_{\rm M},w_2-I_{\rm M}w_2)$ is equal to

$$\sum_{T\in\mathcal{T}}\int_K D^2_{\rm pw}v_{\rm M}:D^2_{\rm pw}(w_2-I_{\rm M}w_2)\,{\rm dx}-\sum_{E\in\mathcal{E}}\int_E[\nabla_{\rm pw}(w_2-I_{\rm M}w_2)]_E\cdot\langle D^2_{\rm pw}v_{\rm M}\rangle_E\nu_E\,{\rm ds}.$$

A piecewise integration by parts of the term $a_{\rm pw}(v_{\rm M},w_2-I_{\rm M}w_2)$ shows equality to

$$\sum_{E\in\mathcal{E}}\int_E\Big([\nabla_{\rm pw}(w_2-I_{\rm M}w_2)\cdot(D^2_{\rm pw}v_{\rm M}\,\nu_E)]_E-[\nabla_{\rm pw}(w_2-I_{\rm M}w_2)]_E\cdot\langle D^2_{\rm pw}v_{\rm M}\rangle_E\nu_E\Big)\,{\rm ds}.$$

The product rule for the jump terms results in

$$(a_{\rm pw}+b_h)(v_{\rm M},w_2-I_{\rm M}w_2)=\sum_{E\in\mathcal{E}}\int_E\langle\nabla_{\rm pw}(w_2-I_{\rm M}w_2)\rangle_E\cdot\big[D^2_{\rm pw}v_{\rm M}\big]_E\nu_E\,{\rm ds}. \tag{7.5}$$

The further analysis concerns the split of the vector $\langle\nabla_{\rm pw}(w_2-I_{\rm M}w_2)\rangle_E\in P_1(E;\mathbb{R}^2)\equiv P_1(E)^2$ into normal and tangential components,

$$\langle\nabla_{\rm pw}(w_2-I_{\rm M}w_2)\rangle_E=\langle\partial(w_2-I_{\rm M}w_2)/\partial\nu_E\rangle_E\nu_E+\langle\partial(w_2-I_{\rm M}w_2)/\partial s\rangle_E\tau_E.$$

The integral of the normal component $\langle\partial(w_2-I_{\rm M}w_2)/\partial\nu_E\rangle_E$ over an edge $E\in\mathcal{E}$ vanishes by definition of $I_{\rm M}w_2$ in Definition 3.5. Since the jump $\big[\partial^2_{\nu_E\nu_E}v_{\rm M}\big]_E:=\nu_E\cdot\big[D^2_{\rm pw}v_{\rm M}\big]_E\nu_E$ is constant along E, the integral $\int_E\langle\partial(w_2-I_{\rm M}w_2)/\partial\nu_E\rangle_E\big[\partial^2_{\nu_E\nu_E}v_{\rm M}\big]_E\,{\rm ds}=0$ vanishes. The tangential components with $\big[\partial^2_{\tau_E\nu_E}v_{\rm M}\big]_E:=\tau_E\cdot\big[D^2v_{\rm M}\big]_E\nu_E$ remain in

$$(a_{\rm pw}+b_h)(v_{\rm M},w_2-I_{\rm M}w_2)=\sum_{E\in\mathcal{E}}\int_E\langle\partial(w_2-I_{\rm M}w_2)/\partial s\rangle_E\big[\partial^2_{\tau_E\nu_E}v_{\rm M}\big]_E\,{\rm ds}.$$

The Hadamard jump condition asserts that the jump in the derivative of a globally continuous function that is smooth up to the boundary on either side of an interface E points merely in the normal direction ν_E only. The function $Jv_{\rm M}$ has a continuous gradient $\nabla Jv_{\rm M}$ and $\nabla Jv_{\rm M}$ is smooth on each triangle $\widehat{T}$ in the HCT refinement of $\mathcal{T}$. Hence $\big[\partial^2_{\tau_E\nu_E} Jv_{\rm M}\big]_E = 0$ along E. Consequently,

$$(a_{\rm pw} + b_h)(v_{\rm M}, w_2 - I_{\rm M} w_2) = \sum_{E\in\mathcal{E}} \int_E \langle \partial(w_2 - I_{\rm M} w_2)/\partial s\rangle_E \big[\partial^2_{\tau_E\nu_E}(1-J)v_{\rm M}\big]_E \,{\rm ds}. \tag{7.6}$$

For an interior edge $E = \partial T_+ \cap \partial T_- = \partial\widehat{T}_+ \cap \partial\widehat{T}_-$ with the neighbouring triangles $T_\pm \in \mathcal{T}$ and the two neighbouring sub-triangles $\widehat{T}_\pm := {\rm conv}(E, {\rm mid}(T_\pm))$ from the HCT refinement of $\mathcal{T}$ with patches $\widehat{\omega}(E) = {\rm int}(\widehat{T}_+ \cup \widehat{T}_-) \subset \omega(E) = {\rm int}(T_+ \cup T_-)$, Cauchy and triangle inequalities show

$$\begin{aligned} I(E) &:= \int_E \langle \partial(w_2 - I_{\rm M} w_2)/\partial s\rangle_E \big[\partial^2_{\tau_E\nu_E}(1-J)v_{\rm M}\big]_E \,{\rm ds} \\ &\le \frac{1}{2}\big(\|\nabla(w_2 - I_{\rm M} w_2)|_{T_+}\|_{L^2(E)} + \|\nabla(w_2 - I_{\rm M} w_2)|_{T_-}\|_{L^2(E)}\big) \\ &\quad \times \Big(\|D^2(1-J)v_{\rm M}|_{\widehat{T}_+}\|_{L^2(E)} + \|D^2(1-J)v_{\rm M}|_{\widehat{T}_-}\|_{L^2(E)}\Big). \end{aligned}$$

Since $(w_2 - I_{\rm M} w_2)|_{\widehat{T}_\pm}$ resp. $(1-J)v_{\rm M}|_{\widehat{T}_\pm}$ is a polynomial of degree at most 2 resp. 3 in the triangle $\widehat{T}_\pm$, the discrete trace inequalities

$$\begin{aligned} \big\|\nabla(w_2 - I_{\rm M} w_2)|_{T_\pm}\big\|_{L^2(E)} &\le h_E^{1/2} C_7 \big\|h_{T_\pm}^{-1}\nabla_{\rm pw}(w_2 - I_{\rm M} w_2)\big\|_{L^2(\widehat{T}_\pm)} \\ \Big\|D^2(1-J)v_{\rm M}|_{\widehat{T}_\pm}\Big\|_{L^2(E)} &\le h_E^{-1/2} C_7 |(1-J)v_{\rm M}|_{H^2(\widehat{T}_\pm)} \end{aligned}$$

hold for a constant $C_7 \approx 1$ that solely depends on the shape regularity of $\widehat{T}_\pm$ (and so on the shape regularity of $\mathcal{T}$). This leads to

$$I(E) \le C_7^2 \big\|h_{\mathcal{T}}^{-1}\nabla_{\rm pw}(w_2 - I_{\rm M} w_2)\big\|_{L^2(\widehat{\omega}(E))} \big\|D^2_{\rm pw}(1-J)v_{\rm M}\big\|_{L^2(\widehat{\omega}(E))}$$

for any interior edge $E \in \mathcal{E}(\Omega)$ with the reduced edge-patch $\widehat{\omega}(E)$. The same estimate follows for a boundary edge $E \in \mathcal{E}(\partial\Omega)$ (the proof omits T_-, $\widehat{T}_-$, and some factor 1/2 above). Since the reduced edge-patches $(\widehat{\omega}(E) : E \in \mathcal{E})$ have no overlap, the sum of all the above estimates of $I(E)$ in (7.6) and Cauchy inequalities prove

$$(a_{\rm pw} + b_h)(v_{\rm M}, w_2 - I_{\rm M} w_2) \le C_7^2 \big|h_{\mathcal{T}}^{-1}(w_2 - I_{\rm M} w_2)\big|_{H^1(\Omega)} |\!|\!|(1-J)v_{\rm M}|\!|\!|_{\rm pw}. \tag{7.7}$$

Recall from Theorem 4.3a (with $\|(1-\Pi_0)D^2_{\rm pw} w_2\| = 0$ for $w_2 \in P_2(\mathcal{T})$) that

$$\big|h_{\mathcal{T}}^{-1}(w_2 - I_{\rm M} w_2)\big|_{H^1(\mathcal{T})} \lesssim j_h(w_2 - w) \lesssim \|w_2 - w\|_h$$

for all $w \in V$. Lemma 3.7d shows $|\!|\!|(1-J)v_{\rm M}|\!|\!|_{\rm pw} \le \Lambda_{\rm J} |\!|\!|v_{\rm M} - v|\!|\!|_{\rm pw}$ for any $v \in V$. The combination of this with (7.7) concludes the proof of (5.13). □

Proof of (5.14). This follows from (7.4). □

Proof of (5.15). This follows from the boundedness of $b_h(\bullet,\bullet)$ (see [28, 39]). □

Proof of (5.16). The jump contributions in (4.3) vanish for arguments in V, $c_{\rm dG}(v_2, w_2) = c_{\rm dG}(v - v_2, w - w_2)$ on the left-hand side of (5.16). Recall that $c_{\rm dG}(\bullet,\bullet)$ is a semi-norm scalar product and the Cauchy inequality with the induced semi-norm $|\bullet|_{c_{\rm dG}} := c_{\rm dG}(\bullet,\bullet)^{1/2} \le \|\bullet\|_{\rm dG}$ is a part of the discrete norm $\|\bullet\|_{\rm dG}$. This leads to (5.16) with $\Lambda_{\rm c} = 1$. □

Proof of (6.1) *for* $\Theta = 1$. This follows from Remark 6.1. □

8. Modified C^0IP method

For the right-hand side $F \in H^{-2}(\Omega)$, the modified C^0IP method is based on the continuous Lagrange P_2 finite element space $V_h := S_0^2(\mathcal{T}) := P_2(\mathcal{T}) \cap H_0^1(\Omega)$ and penalty terms along edges. The scheme is a modification of the dGFEM in Section 7 but with trial and test functions restricted to $S_0^2(\mathcal{T}) := P_2(\mathcal{T}) \cap H_0^1(\Omega)$. The norm $\|\bullet\|_{\mathrm{IP}}$ is $\|\bullet\|_{\mathrm{dG}}$ with restriction to $S_0^2(\mathcal{T})$ and excludes one of the penalty parameters of the modified dGFEM.

Given $\sigma_{\mathrm{IP}} > 0$, the bilinear forms [6,24] for $v_{\mathrm{IP}}, w_{\mathrm{IP}} \in S_0^2(\mathcal{T})$ are defined by

$$A_{\mathrm{IP}}(v_{\mathrm{IP}}, w_{\mathrm{IP}}) := a_{\mathrm{pw}}(v_{\mathrm{IP}}, w_{\mathrm{IP}}) + b_h(v_{\mathrm{IP}}, w_{\mathrm{IP}}) + c_{\mathrm{IP}}(v_{\mathrm{IP}}, w_{\mathrm{IP}}), \text{ where} \tag{8.1}$$

$$c_{\mathrm{IP}}(v_{\mathrm{IP}}, w_{\mathrm{IP}}) := \sum_{E \in \mathcal{E}} \frac{\sigma_{\mathrm{IP}}}{h_E} \int_E \left[\frac{\partial v_{\mathrm{IP}}}{\partial \nu_E}\right]_E \left[\frac{\partial w_{\mathrm{IP}}}{\partial \nu_E}\right]_E \mathrm{ds}, \tag{8.2}$$

and $b_h(\bullet,\bullet) = b_h(\bullet,\bullet)|_{S_0^2(\mathcal{T})}$ from (7.2a).

The modified C^0IP method is of the form (5.3) and seeks $u_{\mathrm{IP}} \in S_0^2(\mathcal{T})$ such that

$$A_{\mathrm{IP}}(u_{\mathrm{IP}}, v_{\mathrm{IP}}) = F(JI_{\mathrm{M}} v_{\mathrm{IP}}) \quad \text{for all } v_{\mathrm{IP}} \in S_0^2(\mathcal{T}). \tag{8.3}$$

For all $v + v_{\mathrm{IP}} \in V + S_0^2(\mathcal{T})$, the discrete norm reads $\|v + v_{\mathrm{IP}}\|_{\mathrm{IP}} := (|\!|\!| v + v_{\mathrm{IP}} |\!|\!|_{\mathrm{pw}}^2 + c_{\mathrm{IP}}(v_{\mathrm{IP}}, v_{\mathrm{IP}}))^{1/2}$ and $j_h(v_{\mathrm{IP}}) = (\sum_{E \in \mathcal{E}} (⨍_E [\partial v_{\mathrm{IP}}/\partial \nu_E]_E \,\mathrm{ds})^2)^{1/2}$. Theorem 4.1 shows $\|v_{\mathrm{IP}}\|_h \approx \|v_{\mathrm{IP}}\|_{\mathrm{IP}}$. The coercivity $\|\cdot\|_{\mathrm{IP}}^2 \lesssim A_{\mathrm{IP}}(\cdot,\cdot)$ on $S_0^2(\mathcal{T})$ holds provided σ_{IP} is sufficiently large [6,24]. The boundedness $A_{\mathrm{IP}}(v_{\mathrm{IP}}, w_{\mathrm{IP}}) \lesssim \|v_{\mathrm{IP}}\|_{\mathrm{IP}} \|w_{\mathrm{IP}}\|_{\mathrm{IP}}$ holds for all $v_{\mathrm{IP}}, w_{\mathrm{IP}}$ on $S_0^2(\mathcal{T})$ and (8.3) has a unique solution $u_{\mathrm{IP}} \in S_0^2(\mathcal{T})$.

Theorem 8.1 (Error estimates)**.** *The solution $u \in V$ to (3.1) and the solution $u_{\mathrm{IP}} \in S_0^2(\mathcal{T})$ to (8.3) satisfy (a) $\|u - u_{\mathrm{IP}}\|_h \lesssim |\!|\!| u - I_{\mathrm{M}} u |\!|\!|_{\mathrm{pw}}$ and (b) if $\Theta = 1$ and $F \in H^{-s}(\Omega)$ for $2 - \sigma \le s \le 2$, then $\|u - JI_{\mathrm{M}} u_{\mathrm{IP}}\|_{H^s(\Omega)} + \|u - u_{\mathrm{IP}}\|_{H^s(\mathcal{T})} \lesssim h_{\max}^{2-s} \|u - u_{\mathrm{IP}}\|_h$.*

Remark 8.2. A C^0IP discrete scheme is analysed in [6] for a general $F \in H^{\sigma-2}(\Omega)$. The consistency of the scheme allows a best approximation ([6], Lem. 8) (since V_h subset $H^{2-\sigma}(\Omega)$ in the pure Dirichlet problem for C^0IP).

For $F \in H^{-2}(\Omega)$, a modifed scheme and error estimates for the post-processed solution are derived in (4.17) and Theorem 4 of [6].

Overview of the proof of Theorem 8.1. The proof follows the lines of that of Theorem 7.2 and partly from the analysis provided there. The bilinear forms in the C^0IP are exactly the respective bilinear forms of the dGFEM when restricted to the subspace $S_0^2(\mathcal{T}) + M(\mathcal{T})$. With the single exception of (5.11), all the estimates in (5.9)–(5.16) and (6.1) for $\Theta = 1$ follow for $V_h = S_0^2(\mathcal{T})$ in the C^0IP from the respective properties verified in Section 7 for $V_h = P_2(\mathcal{T})$ in the dGFEM. The remaining detail is the analysis of the operator $I_h \equiv I_{\mathrm{C}} : \mathrm{M}(\mathcal{T}) \to S_0^2(\mathcal{T})$ (denoted by I_2^* in Lem. 3.2 of [23]) defined by averaging the values of a function $v_{\mathrm{M}} \in \mathrm{M}(\mathcal{T})$ at the midpoint of an interior edge E,

$$(I_{\mathrm{C}} v_{\mathrm{M}})(z) = \begin{cases} v_{\mathrm{M}}(z) & \text{for all } z \in \mathcal{V}, \\ \langle v_{\mathrm{M}} \rangle_E(z) & \text{for } z = \mathrm{mid}(E),\ E \in \mathcal{E}(\Omega), \\ 0 & \text{for } z = \mathrm{mid}(E),\ E \in \mathcal{E}(\partial\Omega). \end{cases} \tag{8.4}$$

Proof of (5.11). This is included in Lemma 3.2f of [23] in a slightly different notation. In the notation of this paper, Lemma 3.2 of [23] shows

$$\|v_{\mathrm{M}} - I_{\mathrm{C}} v_{\mathrm{M}}\|_h^2 \lesssim \sum_{E \in \mathcal{E}} h_E^{-2} |[v_{\mathrm{M}}]_E(\mathrm{mid}(E))|^2 \tag{8.5}$$

for any $v_{\mathrm{M}} \in \mathrm{M}(\mathcal{T})$. Lemma 3.3 of [23] controls the upper bound of (8.5) by the a posteriori terms $\sum_{E\in\mathcal{E}} h_E \|\big[D^2_{\mathrm{pw}} v_{\mathrm{M}}\big]_E \tau_E\|^2_{L^2(E)}$. The latter is efficient, *i.e.*, $\lesssim |\!|\!| v_{\mathrm{M}} - v |\!|\!|_{\mathrm{pw}}$ for any $v \in H^2_0(\Omega)$. This leads to (5.11). Theorem 4.5 allows for an alternative proof that departs at (8.5) with the quadratic function $[v_{\mathrm{M}}]_E$ along the edge $E \in \mathcal{E}$. Since $[v_{\mathrm{M}}]_E$ vanishes at each end point $z \in \mathcal{V}(E)$ (owing to the continuity of the Morley function $v_{\mathrm{M}} \in \mathrm{M}(\mathcal{T})$ at the vertices), the (exact) Simpson's quadrature rule asserts

$$[v_{\mathrm{M}}]_E(\mathrm{mid}(E)) = 3/2 \fint_E [v_{\mathrm{M}}]_E \,\mathrm{ds}.$$

A Cauchy inequality, the continuity of $Jv_{\mathrm{M}} \in V$, triangle and trace inequalities lead to

$$h_E^{-2} |[v_{\mathrm{M}}]_E(\mathrm{mid}(E))|^2 \le \frac{9}{4} h_E^{-3} \|[v_{\mathrm{M}} - Jv_{\mathrm{M}}]_E\|^2_{L^2(E)} \lesssim \sum_{\ell=0}^{1} \big| h_{\mathcal{T}}^{\ell-2}(v_{\mathrm{M}} - Jv_M)\big|^2_{H^\ell(\omega(E))}.$$

This estimate and the finite overlap of the edge-patches $(\omega(E) : E \in \mathcal{E})$ lead to an upper bound in (8.5) as in Theorem 4.5c and so to $\|v_{\mathrm{M}} - I_{\mathrm{C}} v_{\mathrm{M}}\|_h \lesssim \|v_{\mathrm{M}} - v\|_h = |\!|\!| v_{\mathrm{M}} - v |\!|\!|_{\mathrm{pw}}$. □

Remark 8.3 (Q is not injective)**.** An illustration shall be given for a triangulation $\mathcal{T} = \{T_1, T_2\}$ of a convex quadrilateral $\bar{\Omega} = T_1 \cup T_2 = \mathrm{conv}\{P_1, P_2, P_3, P_4\}$. There is exactly one basis function $b_E \in S^2_0(\mathcal{T}) = V_h$ defined on $T_1 = \mathrm{conv}\,\{P_1, P_2, P_4\}$ and on $T_2 = \mathrm{conv}\,\{P_2, P_3, P_4\}$ by $b_E = 4\varphi_2\varphi_4$ for the nodal basis functions $\varphi_1, \varphi_2, \varphi_3, \varphi_4$. Given b_E for the edge $E = \mathrm{conv}\,\{P_2, P_4\} = \partial T_1 \cap \partial T_2$, the normal derivative of b_E along E on T_1 reads in its integral

$$\begin{aligned}\int_E \partial b_E|_{T_1}/\partial\nu_E \,\mathrm{ds} &= \nu_E \cdot \int_E \nabla b_E|_{T_1} \,\mathrm{ds} = 4\nu_E \cdot \int_E (\varphi_2 \nabla\varphi_4|_{T_1} + \varphi_4 \nabla\varphi_1|_{T_1}) \,\mathrm{ds} \\ &= 2|E|\nu_E \cdot (\nabla\varphi_2|_{T_1} + \nabla\varphi_4|_{T_1}) = -2|E|\nu_E \cdot \nabla\varphi_1|_{T_1}\end{aligned}$$

with $\nabla(\varphi_1 + \varphi_2 + \varphi_4) = 0$ in T_1 in the last step. Elementary geometry shows $\nabla\varphi_1|_{T_1} = \nu_E/\rho_{E,1}$ for the height $\rho_{E,1} = \frac{2|T_1|}{|E|}$ of E in T_1 and ν_E pointing from T_1 into T_2. Consequently, $\int_E \partial b_E|_{T_1}/\partial\nu_E \,\mathrm{ds} = \frac{|E|^2}{|T_1|}$. The analogous calculation for T_2 leads to $\int_E \partial b_E|_{T_2}/\partial\nu_E \,\mathrm{ds} = -\frac{|E|^2}{|T_2|}$ with a change of sign because $\nabla\varphi_3|_{T_2} = \nu_E/\rho_{E,2}$. The definition of $I_{\mathrm{M}} b_E$ takes the average of the two integral means

$$\frac{1}{2}\Big(\fint_E \partial b_E|_{T_1}/\partial\nu_E + \fint_E \partial b_E|_{T_2}/\partial\nu_E\Big) \,\mathrm{ds} = \frac{|E|}{2}\big(|T_1|^{-1} - |T_2|^{-1}\big)$$

as the value for $\fint_E \partial I_{\mathrm{M}} b_E/\partial\nu_E \,\mathrm{ds}$. Since this is the only degree of freedom in $\mathrm{M}(\mathcal{T})$ for the triangulation $\mathcal{T} = \{T_1, T_2\}$, it follows that $I : S^2_0(\mathcal{T}) \to \mathrm{M}(\mathcal{T})$ is injective if and only of $|T_1| = |T_2|$. (This condition is independent of shape-regularity of $\mathcal{T}$ and thus more involved.)

9. Comparison

The paper [23] has established equivalence of discrete solutions to Morley FEM, C^0IP and dGFEM up to oscillations for $F \in L^2(\Omega)$ and for the original schemes with $F_h \equiv F$. The subsequent theorem establishes the three modified schemes with $F_h = F \circ J$ without extra oscillation terms. Throughout this section, the norm $\|\cdot\|_h$ is defined in (4.1) and (4.2).

Theorem 9.1. *The discrete solutions u_{M}, u_{IP} and u_{dG} of the Morley FEM, C^0IP and dGFEM satisfy*

$$\|u - u_{\mathrm{M}}\|_h \approx \|u - u_{\mathrm{dG}}\|_h \approx \|u - u_{\mathrm{IP}}\|_h \approx \|(1 - \Pi_0) D^2 u\|_{L^2(\Omega)}.$$

The equivalence constants $\approx$ depend on shape regularity and on the stabilisation parameters $\sigma_{\mathrm{dG}}, \sigma_{\mathrm{IP}} \approx 1$.

Remark 9.2 (Discrete dG norm equivalence, Theorem 4.1, [23])**.** The norm $\|\cdot\|_h$ satisfies

$$\begin{aligned} \|\bullet\|_h &= |\!|\!| \bullet |\!|\!|_{\mathrm{pw}} && \text{on } V + \mathrm{M}(\mathcal{T}),\\ \|\bullet\|_h &\approx \|\bullet\|_{\mathrm{dG}} && \text{on } V + P_2(\mathcal{T}),\\ \|\bullet\|_h &\approx \|\bullet\|_{\mathrm{IP}} && \text{on } V + S_0^2(\mathcal{T}). \end{aligned}$$

Proof. Lemma 3.2a, the Pythogoras identity, and Theorem 2.2 of [16] show

$$\big\|(1-\Pi_0)D^2u\big\|_{L^2(\Omega)} = |\!|\!|u - I_{\mathrm{M}}u|\!|\!|_{\mathrm{pw}} \le |\!|\!|u - u_{\mathrm{M}}|\!|\!|_{\mathrm{pw}} \lesssim |\!|\!|u - I_{\mathrm{M}}u|\!|\!|_{\mathrm{pw}}.$$

The L^2 best-approximation property of $\Pi_0 u$, (5.2a), Theorem 7.2, and Lemma 3.2a lead to

$$\begin{aligned} \|(1-\Pi_0)D^2u\|_{L^2(\Omega)} &= \min_{v_h\in P_2(\mathcal{T})} \|u - v_h\|_h \le \|u - u_{\mathrm{dG}}\|_h\\ &\lesssim |\!|\!|u - I_{\mathrm{M}}u|\!|\!|_{\mathrm{pw}} = \|(1-\Pi_0)D^2u\|_{L^2(\Omega)}. \end{aligned}$$

Theorem 8.1 leads to similar results for $\|u - u_{\mathrm{IP}}\|_h$. A combination of the above displayed inequalities concludes the proof. □

10. Modified WOPSIP method

The weakly over-penalized symmetric interior penalty (WOPSIP) scheme [7] is a penalty method with the stabilisation term

$$\begin{aligned} c_{\mathrm{P}}(v_{\mathrm{pw}}, w_{\mathrm{pw}}) := &\sum_{E\in\mathcal{E}} \sum_{z\in\mathcal{V}(E)} h_E^{-4}([v_{\mathrm{pw}}]_E(z))([w_{\mathrm{pw}}]_E(z))\\ &+ \sum_{E\in\mathcal{E}} h_E^{-2}\Big(\fint_E [\partial v_{\mathrm{pw}}/\partial\nu_E]_E \,\mathrm{ds}\Big)\Big(\fint_E [\partial w_{\mathrm{pw}}/\partial\nu_E]_E \,\mathrm{ds}\Big) \end{aligned} \tag{10.1}$$

for piecewise smooth functions $v_{\mathrm{pw}}, w_{\mathrm{pw}} \in H^2(\mathcal{T})$. This semi-norm scalar product $c_{\mathrm{P}}(\bullet,\bullet)$ is an analog to that one behind the jump j_h from (4.2) with different powers of the mesh-size. It follows as in Theorem 4.1 that

$$A_{\mathrm{P}}(v_{\mathrm{pw}}, w_{\mathrm{pw}}) := a_{\mathrm{pw}}(v_{\mathrm{pw}}, w_{\mathrm{pw}}) + c_{\mathrm{P}}(v_{\mathrm{pw}}, w_{\mathrm{pw}}) \quad \text{for all } v_{\mathrm{pw}}, w_{\mathrm{pw}} \in H^2(\mathcal{T}) \tag{10.2}$$

defines a scalar product and so $\|\bullet\|_{\mathrm{P}} := A_{\mathrm{P}}(\bullet,\bullet)^{1/2}$ is a norm in $H^2(\mathcal{T})$. Consequently, there exists a unique solution $u_{\mathrm{P}} \in V_h := P_2(\mathcal{T})$ to

$$A_{\mathrm{P}}(u_{\mathrm{P}}, v_2) = F(JI_{\mathrm{M}}v_2) \quad \text{for all } v_2 \in P_2(\mathcal{T}). \tag{10.3}$$

The increased condition number in the over-penalization of the jumps by the negative powers of the mesh-size in (10.1) can be compensated by some preconditioner ([7], p. 218f) and the entire WOPSIP linear algebra with (10.3) becomes intrinsically parallel.

Theorem 10.1 (Error estimate)**.** *Any $F \in H^{-s}(\Omega)$ with $2-\sigma \le s \le 2$, the solution $u \in V$ to (3.1) and the solution $u_{\mathrm{P}} \in P_2(\mathcal{T})$ to (10.3) satisfy*

$$\begin{aligned} |\!|\!|u - u_{\mathrm{P}}|\!|\!|_{\mathrm{pw}}^2 + c_{\mathrm{P}}(u_{\mathrm{P}}, u_{\mathrm{P}}) &\le (1+\Lambda_{\mathrm{P}}^2)|\!|\!|u - I_{\mathrm{M}}u|\!|\!|_{\mathrm{pw}}^2 + \Lambda_{\mathrm{P}}^2 |\!|\!|h_{\mathcal{T}} I_{\mathrm{M}}u|\!|\!|_{\mathrm{pw}}^2;\\ \|u - JI_{\mathrm{M}}u_{\mathrm{P}}\|_{H^s(\Omega)} + \|u - u_{\mathrm{P}}\|_{H^s(\mathcal{T})} &\le C_8(s) h_{\max}^{2-s} \|u - u_{\mathrm{P}}\|_{\mathrm{P}}. \end{aligned}$$

The constant Λ_{P} exclusively depends on the shape regularity of $\mathcal{T}$, while C_8 depends on the shape regularity of $\mathcal{T}$ and on s.

The subsequent lemma specifies the constant $\Lambda_{\rm P}$ in the best-approximation estimate.

Lemma 10.2. *There exists some positive $\Lambda_{\rm P} < \infty$, that exclusively depends on the shape regularity of $\mathcal{T}$, such that $\|\!| h_{\mathcal{T}}^{-1}(v_2 - I_{\rm M} v_2)|\!\|_{\rm pw}^2 + \|\!|(1-J) I_{\rm M} v_2|\!\|_{\rm pw}^2 \le \Lambda_{\rm P}^2 \|v_2 - v\|_{\rm P}^2$ holds for all $v_2 \in P_2(\mathcal{T})$ and all $v \in V$.*

Proof of Lemma 10.2. The analysis of $\|\!| h_{\mathcal{T}}^{-1}(v_2 - I_{\rm M} v_2)|\!\|_{\rm pw}$ returns to the proof of Theorem 4.3 that eventually provides (4.11) for one fixed triangle $T \in \mathcal{T}$ with its neighourhood $\Omega(T)$ for any $v_2 \in P_2(\mathcal{T})$. The substitution of v_2 by $h_T^{-1} v_2$ (with a fixed scaling factor h_T) in (4.11) after a standard inverse estimate shows, for all $v_2 \in P_2(\mathcal{T})$, that

$$h_T^{-2} |v_2 - I_{\rm M} v_2|_{H^2(T)}^2 \lesssim h_T^{-6} \|v_2 - I_{\rm M} v_2\|_{L^2(T)}^2 \lesssim j_h\big(h_T^{-1} v_2, T\big)^2.$$

The shape regularity of $\mathcal{T}$ implies that all edge-sizes in the sub-triangulation $\mathcal{T}(\Omega(T))$ that covers the neighbouhood $\Omega(T)$ (of T and one layer of triangles around T) are equivalent to h_T. Hence $j_h\big(h_T^{-1} v_2, T\big)$ is equivalent to the respective contributions in $c_{\rm P}(v_2, v_2)$:

$$j_h(h_T^{-1} v_2, T)^2 \lesssim \sum_{z \in \mathcal{V}(T)} \sum_{F \in \mathcal{E}(z)} h_F^{-4} ([v_2]_F(z))^2 + \sum_{E \in \mathcal{E}(T)} h_E^{-2} \Big(\fint_E [\partial v_{\rm pw} / \partial \nu_E]_E \, {\rm ds} \Big)^2.$$

The combination of this estimate with the previous one and the sum over all those estimates lead to

$$\|\!| h_{\mathcal{T}}^{-1}(v_2 - I_{\rm M} v_2)|\!\|_{\rm pw}^2 \lesssim c_{\rm P}(v_2, v_2) \tag{10.4}$$

owing to the finite overlap of the family $(\Omega(T) : T \in \mathcal{T})$. The second term $\|\!|(1-J) I_{\rm M} v_2|\!\|_{\rm pw}$ is controlled with Lemma 3.7d and a triangle inequality in

$$\Lambda_{\rm J}^{-1} \|\!|(1-J) I_{\rm M} v_2|\!\|_{\rm pw} \le \|\!| v - I_{\rm M} v_2|\!\|_{\rm pw} \le \|\!| v - v_2|\!\|_{\rm pw} + \|\!| v_2 - I_{\rm M} v_2|\!\|_{\rm pw}.$$

Since the the last term $\|\!| v_2 - I_{\rm M} v_2|\!\|_{\rm pw} \le h_{\max} \|\!| h_{\mathcal{T}}^{-1}(v_2 - I_{\rm M} v_2)|\!\|_{\rm pw}$ is bounded in (10.4), the summary of the aforementioned estimates concludes the proof of the lemma. □

Proof of energy norm estimate in Theorem 10.1. The equations (3.1) and (10.3) show the key identity

$$a(u, J I_{\rm M} e_{\rm P}) = F(J I_{\rm M} e_{\rm P}) = a_{\rm pw}(u_{\rm P}, e_{\rm P}) + c_{\rm P}(u_{\rm P}, e_{\rm P}) \quad \text{for } e_{\rm P} := I_{\rm M} u - u_{\rm P} \in P_2(\mathcal{T}).$$

Remark 4.2 applies verbatim and provides $c_{\rm P}(I_{\rm M} u, e_{\rm P}) = 0$. This, the key identity, and the definition of the norm $\| \bullet \|_{\rm P} := A_{\rm P}(\bullet, \bullet)^{1/2}$ lead to

$$\|e_{\rm P}\|_{\rm P}^2 = a_{\rm pw}(I_{\rm M} u, e_{\rm P}) - a(u, J I_{\rm M} e_{\rm P}) = a_{\rm pw}(u, e_{\rm P} - J I_{\rm M} e_{\rm P}) \tag{10.5}$$

with (3.5) in the last step. Set $e_{\rm M} := I_{\rm M} e_{\rm P} \in {\rm M}(\mathcal{T})$ and split $e_{\rm P} - J I_{\rm M} e_{\rm P} = (e_{\rm P} - e_{\rm M}) + (1-J) e_{\rm M}$. The last term in (10.5) is equal to

$$\begin{aligned} a_{\rm pw}(u, e_{\rm P} - e_{\rm M}) + a_{\rm pw}(u, e_{\rm M} - J e_{\rm M}) &= a_{\rm pw}(I_{\rm M} u, e_{\rm P} - e_{\rm M}) + a_{\rm pw}(u - I_{\rm M} u, e_{\rm M} - J e_{\rm M}) \\ &\le \|\!| h_{\mathcal{T}} I_{\rm M} u|\!\|_{\rm pw} \|\!| h_{\mathcal{T}}^{-1}(e_{\rm P} - e_{\rm M})|\!\|_{\rm pw} + \|\!| u - I_{\rm M} u|\!\|_{\rm pw} \|\!| e_{\rm M} - J e_{\rm M}|\!\|_{\rm pw} \\ &\le \Lambda_{\rm P} (\|\!| h_{\mathcal{T}} I_{\rm M} u|\!\|_{\rm pw}^2 + \|\!| u - I_{\rm M} u|\!\|_{\rm pw}^2)^{1/2} \|e_{\rm P}\|_{\rm P} \end{aligned}$$

with (3.5) twice in the first equality, (weighted) Cauchy inequalities for the inequality in the third line, and Lemma 10.2 (with $v = 0$) in the end. The combination with (10.5) proves

$$\|e_{\rm P}\|_{\rm P}^2 \le \Lambda_{\rm P}^2 \Big(\|\!| h_{\mathcal{T}} I_{\rm M} u|\!\|_{\rm pw}^2 + \|\!| u - I_{\rm M} u|\!\|_{\rm pw}^2 \Big).$$

This and the Pythagoras theorem (3.7) conclude the proof. □

Proof of error estimates in weaker (piecewise) Sobolev norms in Theorem 10.1. The error analysis in weaker norms adapts the notation of the beginning of Section 6.2 on $v := JI_{\mathrm{M}}u_{\mathrm{P}} \in V$ and $\zeta := JI_{\mathrm{M}}z \in V$ for the dual solution $z \in V$ with $\|z\|_{H^{4-s}(\Omega)} \le C_{\mathrm{reg}}$ and

$$\|u - v\|_{H^s(\Omega)} = a(u - v, z) = a(u, z - \zeta) + a_{\mathrm{pw}}(u_{\mathrm{P}} - v, z)$$

with the key identity $a(u, \zeta) = F(\zeta) = A_{\mathrm{P}}(u_{\mathrm{P}}, I_{\mathrm{M}}z) = a_{\mathrm{pw}}(u_{\mathrm{P}}, I_{\mathrm{M}}z) = a_{\mathrm{pw}}(u_{\mathrm{P}}, z)$ (from Rem. 4.2 and (3.5)) in the last step. Since $a_{\mathrm{pw}}(I_{\mathrm{M}}u, z - \zeta) = 0$ (from (3.5) with $I_{\mathrm{M}}z = I_{\mathrm{M}}\zeta$ from (3.9)), the first term

$$a(u, z - \zeta) = a_{\mathrm{pw}}(u - I_{\mathrm{M}}u, z - \zeta) \le (1 + \Lambda_{\mathrm{J}})(1 + \Lambda_{\mathrm{M}})|\!|\!|u - I_{\mathrm{M}}u|\!|\!|_{\mathrm{pw}}|\!|\!|z - I_{\mathrm{M}}z|\!|\!|_{\mathrm{pw}}$$

is controlled by (6.5) (with $z_h = I_{\mathrm{M}}z$, $I_h = \mathrm{id}$). The analysis of the second term $a_{\mathrm{pw}}(u_{\mathrm{P}} - v, \zeta)$ follows the corresponding lines of the proof of the best-approximation in Theorem 10.1 with $u_{\mathrm{M}} := I_{\mathrm{M}}u_{\mathrm{P}}$, $u_{\mathrm{P}} - v = u_{\mathrm{P}} - u_{\mathrm{M}} + (1 - J)u_{\mathrm{M}}$. This shows

$$\begin{aligned}
a_{\mathrm{pw}}(u_{\mathrm{P}} - v, z) &= a_{\mathrm{pw}}(u_{\mathrm{P}} - u_{\mathrm{M}}, z) + a_{\mathrm{pw}}(u_{\mathrm{M}} - Ju_{\mathrm{M}}, z) \\
&= a_{\mathrm{pw}}(u_{\mathrm{P}} - u_{\mathrm{M}}, I_{\mathrm{M}}z) + a_{\mathrm{pw}}(u_{\mathrm{M}} - Ju_{\mathrm{M}}, z - I_{\mathrm{M}}z) \\
&\le |\!|\!|h_{\mathcal{T}} I_{\mathrm{M}}z|\!|\!|_{\mathrm{pw}} |\!|\!|h_{\mathcal{T}}^{-1}(u_{\mathrm{P}} - u_{\mathrm{M}})|\!|\!|_{\mathrm{pw}} + |\!|\!|z - I_{\mathrm{M}}z|\!|\!|_{\mathrm{pw}} |\!|\!|u_{\mathrm{M}} - Ju_{\mathrm{M}}|\!|\!|_{\mathrm{pw}} \\
&\le \Lambda_{\mathrm{P}} \Big(|\!|\!|h_{\mathcal{T}} I_{\mathrm{M}}z|\!|\!|_{\mathrm{pw}}^2 + |\!|\!|z - I_{\mathrm{M}}z|\!|\!|_{\mathrm{pw}}^2 \Big)^{1/2} \|u - u_{\mathrm{P}}\|_{\mathrm{P}}
\end{aligned}$$

from Lemma 10.2 with $v = u$. The final argument is the regularity of z and the approximation estimates $|\!|\!|z - I_{\mathrm{M}}z|\!|\!|_{\mathrm{pw}} \le C_{\mathrm{int}}(s) C_{\mathrm{reg}}(s) h_{\max}^{2-s}$ from Section 6.5. The combination of the above arguments shows

$$\|u - JI_{\mathrm{M}}u_{\mathrm{P}}\|_{H^s(\Omega)} \lesssim h_{\max}^{2-s} \|u - u_{\mathrm{P}}\|_{\mathrm{P}}.$$

Theorem 4.5d applies to $u_{\mathrm{P}} \in P_2(\mathcal{T})$ and the remaining arguments follow the last lines in the proof Theorem 6.2b with a triangle inequality in $H^s(\mathcal{T})$ in the end. □

Appendix A.

Proof of Lemma 2.2. For $z \in Y_h \cap Y$, (2.2) implies $\|z - Qz\|_{\widehat{Y}} = 0$, and hence (2.3) holds. The converse is a consequence of the finite dimension of Y_h. In the context of Peetre–Tartar lemma [40], let $\widetilde{Y} := (1 - Q)(Y_h) \subset \widehat{Y}$ denote the range of $1 - Q$ and abbreviate $A := 1 - \Pi_Y$ and $B := \Pi_Y$. Then $A \in L(\widetilde{Y}; \widehat{Y})$ is injective because $A\widetilde{y} = 0$ means $\widetilde{y} \in \widetilde{Y} \cap Y$ for some $y_h \in Y_h$ with $\widetilde{y} = (1 - Q)y_h$, whence $y_h \in Y_h \cap Y$ and $\widetilde{y} = 0$ from (2.2). Notice that $B \in L(\widetilde{Y}; Y)$ is compact (for $\widetilde{Y}$ has finite dimension). Since $\widetilde{y} = A\widetilde{y} + B\widetilde{y}$ implies $\|\widetilde{y}\|_{\widehat{Y}} \le \|A\widetilde{y}\|_{\widehat{Y}} + \|B\widetilde{y}\|_{\widehat{Y}}$ for all $\widetilde{y} \in \widehat{Y}$, the Peetre–Tartar lemma proves $\gamma\|\widetilde{y}\|_{\widehat{Y}} \le \|A\widetilde{y}\|_{\widehat{Y}}$ for all $\widetilde{y} \in \widehat{Y}$ and some constant $\gamma > 0$. This implies for all $y_h \in Y_h$ that

$$\gamma\|y_h - Qy_h\|_{\widehat{Y}} \le \|A(y_h - Qy_h)\|_{\widehat{Y}} = \|y_h - \Pi_Y y_h\|_{\widehat{Y}} \le \|y_h - y\|_{\widehat{Y}} \text{ for all } \ y \in Y.$$

This is (2.3) with $\Lambda_{\mathrm{Q}} := 1/\gamma > 0$. The point in those compactness arguments is that we do know that $\gamma = 1/\Lambda_{\mathrm{Q}} > 0$ with (2.3) exists, but we do not know how it depends, *e.g.*, on $\dim Y_h$. □

Proof of Lemma 2.4. It is obvious that **(QO)** implies (2.4). The converse follows from a compactness argument from $\dim X_h < \infty$. The subspace $\widetilde{X} := (1 - M)X \subset \widehat{X}$ is complete because $1 - M : (X_h \cap X)^{\perp} \cap X \to \widetilde{X}$ is linear, bounded, and bijective for the (complete) orthogonal complement $(X_h \cap X)^{\perp} \cap X$ of $X_h \cap X$ in X. Then $A := 1 - \Pi_{X_h} \in L(\widetilde{X}; \widehat{X})$ is injective (as $A\widetilde{x} = 0$ implies $\widetilde{x} = (1 - M)x \in X_h$, whence $x \in X_h \cap X$ and $\widetilde{x} = 0$ in $\widetilde{X}$ from (2.4)). Since $B := \Pi_{X_h} \in L(\widetilde{X}; \widehat{X}_h)$ is compact and $\widetilde{x} = A\widetilde{x} + B\widetilde{x}$ implies

$$\|\widetilde{x}\|_{\widehat{X}} \le \|A\widetilde{x}\|_{\widehat{X}} + \|B\widetilde{x}\|_{\widehat{X}} \text{ for all } \ \widetilde{x} \in \widehat{X},$$

the Peetre–Tartar lemma [40] leads to some $\gamma > 0$ with

$$\gamma\|(1-M)x\|_{\widehat{X}} \le \|A(1-M)x\|_{\widehat{X}} = \|x - \Pi_{X_h}x\|_{\widehat{X}} \le \|x - x_h\|_{\widehat{X}}$$

for all $x \in X$ and $(1-M)x \in \widetilde{X}$ and for all $x_h \in X_h$. This is **(QO)** with $C_{\rm qo} := 1/\gamma$. □

Proof of Theorem 2.5. [Proof of "$\Longrightarrow$"] Suppose M satisfies **(QO)** with constant $C_{\rm qo}$. Then for all x_h, y_h, the definition of MPx_h leads to the identity

$$\begin{aligned} a_h(x_h, y_h) - a(Px_h, Qy_h) &= a_h(x_h - MPx_h, Qy_h) = \langle Q^*A_h(x_h - MPx_h), y_h\rangle_{Y_h^*\times Y_h} \\ &\le \|Q^*A_h\|\,\|y_h\|_{Y_h}\|x_h - MPx_h\|_{X_h}. \end{aligned}$$

It remains to prove that $\|x_h - MPx_h\|_{X_h} \le (1 + C_{\rm qo})\|x_h - Px_h\|_{\widehat{X}}$. This follows from a triangle inequality $\|x_h - MPx_h\|_{X_h} \le \|x_h - Px_h\|_{\widehat{X}} + \|Px_h - MPx_h\|_{\widehat{X}}$ and **(QO)** in $\|Px_h - MPx_h\|_{\widehat{X}} \le C_{\rm qo}\|Px_h - x_h\|_{\widehat{X}}$. In conclusion, **(H)** holds with $\Lambda_{\rm H} := \|Q^*A_h\|(1 + C_{\rm qo})$.

[Proof of "$\Longleftarrow$"] Let $\Pi_{X_h} \in L(\widehat{X})$ denote the orthogonal projection onto the closed subset X_h in the Hilbert space $\widehat{X}$. Given any $x \in X$, let $x_h^* := \Pi_{X_h}x = \arg\min_{\xi_h\in X_h}\|x - \xi_h\|_{\widehat{X}}$ denote the best-approximation of x in X_h in the Hilbert space $\widehat{X}$ and set $e_h := x_h^* - Mx \in X_h$. The inf-sup condition for $a_h(\bullet,\bullet)$ leads to $y_h \in Y_h$ with norm $\|y_h\|_{Y_h} \le 1$ such that

$$\alpha_h\|e_h\|_{X_h} = a_h(e_h, y_h) = a_h(x_h^*, y_h) - a_h(Mx, y_h).$$

Recall the definition of $x_h := Mx = A_h^{-1}Q^*Ax$ as the discrete solution for the right-hand side $a(x, Q\bullet)$ to verify

$$a_h(Mx, y_h) = a_h(x_h, y_h) = a(x, Qy_h).$$

This leads to the identity

$$\alpha_h\|e_h\|_{X_h} = a_h(x_h^*, y_h) - a(x, Qy_h). \tag{A.1}$$

Hypothesis **(H)** and $\|y_h\|_{Y_h} \le 1$ lead to the first and (2.5) to the last inequality in

$$a_h(x_h^*, y_h) - a(Px_h^*, Qy_h) \le \Lambda_{\rm H}\|x_h^* - Px_h^*\|_{\widehat{X}} \le \Lambda_{\rm H}\Lambda_{\rm P}\|x - x_h^*\|_{\widehat{X}}. \tag{A.2}$$

A triangle inequality and (2.5) imply

$$\|x - Px_h^*\|_X \le \|x - x_h^*\|_{\widehat{X}} + \|x_h^* - Px_h^*\|_{\widehat{X}} \le\ (1 + \Lambda_{\rm P})\|x - x_h^*\|_{\widehat{X}}.$$

With the operator norm $||Q^*A||$ in $L(X; Y_h^*)$ and the duality brackets $\langle\bullet,\bullet\rangle_{Y_h^*\times Y_h}$ in $Y_h^*\times Y_h$, this and $\|y_h\|_{Y_h} \le 1$ show

$$a(Px_h^* - x, Qy_h) = \langle Q^*A(Px_h^* - x), y_h\rangle_{Y_h^*\times Y_h} \le \|Q^*A\|(1 + \Lambda_{\rm P})\|x - x_h^*\|_{\widehat{X}}. \tag{A.3}$$

The combination of (A.1)–(A.3) reads

$$\alpha_h\|e_h\|_{X_h} \le (\Lambda_{\rm H}\Lambda_{\rm P} + \|Q^*A\|(1 + \Lambda_{\rm P}))\|x - x_h^*\|_{\widehat{X}}.$$

Define $C_{\rm qo} := 1 + \alpha_h^{-1}(\Lambda_{\rm H}\Lambda_{\rm P} + \|Q^*A\|(1 + \Lambda_{\rm P}))$ and rewrite the last estimate as

$$\|Mx - x_h^*\|_{X_h} = \|e_h\|_{X_h} \le (C_{\rm qo} - 1)\|x - x_h^*\|_{\widehat{X}}. \tag{A.4}$$

A triangle inequality $\|x - Mx\|_{\widehat{X}} \le \|x - x_h^*\|_{\widehat{X}} + \|Mx - x_h^*\|_{X_h}$ and (A.4) prove **(QO)** because of $\|x - x_h^*\|_{\widehat{X}} = \min_{x_h\in X_h}\|x - x_h\|_{\widehat{X}}$. □

Proof of Theorem 2.6. [Proof of "$\Longrightarrow$"] Given $x'_h \in X'_h$, let $N := \operatorname{Ker} M \subset X$; and let $N^\perp$ denote the orthogonal complement of N in the Hilbert space X. The restriction $M|_{N^\perp} : N^\perp \to X'_h$ of M is linear, bounded, and bijective and hence has a linear and bounded inverse $S := (M|_{N^\perp})^{-1} : X'_h \to N^\perp$. Since N is closed in $\widehat{X} = X + X_h$, the orthogonal projection $\Pi_N \in L(\widehat{X})$ onto N is well-defined and so is its restriction $\Pi_N|_{X'_h} \in L(X'_h; X)$. Given $S \in L(X'_h; X)$ and $\Pi_N|_{X'_h}$, define $P' := \Pi_N + S \in L(X'_h; X)$. Let $x := P'x'_h = \Pi_N x'_h + Sx'_h \in X$ and observe $MX = M|_{N^\perp}(SX'_h) = x_h$ and $MP' = \mathrm{id}$ in X'_h. Let $\xi := \Pi_X x'_h \in X$ be the best-approximation of x_h in X with respect to the norm of $\widehat{X}$. Since $\xi - P'M\xi \in N \perp x'_h - P'x'_h \in N^\perp$, the Pythogoras theorem in $\widehat{X}$ reads

$$\|\xi - x'_h + P'(x'_h - M\xi)\|^2_{\widehat{X}} = \|x'_h - Px'_h\|^2_{\widehat{X}} + \|\xi - P'M\xi\|^2_{\widehat{X}}.$$

The left-hand side of the above displayed equality is an upper bound of $\|x'_h - Px'_h\|^2_{\widehat{X}}$ and is smaller than or equal to the square of

$$\|\xi - x'_h + P'(x'_h - M\xi)\|_{\widehat{X}} \le \|(1 - \Pi_X)x'_h\|_{\widehat{X}} + \|P'\|\, \|x'_h - M\xi\|_{X_h}$$

with the operator norm $\|P'\|$ of $P' \in L(X'_h; X)$. Consequently,

$$\|x'_h - Px'_h\|_{\widehat{X}} \le \|(1 - \Pi_X)x'_h\|_{\widehat{X}} + \|P'\|\, \|x'_h - M\xi\|_{X_h}.$$

A triangle inequality and **(QO)** with $\|\xi - M\xi\|_{\widehat{X}} \le C_{\mathrm{qo}}\|\xi - x'_h\|_{\widehat{X}}$ show

$$\|x'_h - M\xi\|_{X_h} \le \|\xi - x'_h\|_{\widehat{X}} + C_{\mathrm{qo}}\|\xi - x'_h\|_{\widehat{X}} = (1 + C_{\mathrm{qo}})\|\xi - x'_h\|_{\widehat{X}}.$$

The combination of the previous two displayed estimates with $\|(1 - \Pi_X)x'_h\|_{\widehat{X}} \le \|x'_h - Px'_h\|_{\widehat{X}}$ (from $Px'_h \in X$ and the definition of $\Pi_X x_h$) and $\Lambda_{\mathrm{P}'} := 1 + \|P'\|(1 + C_{\mathrm{qo}})$ proves

$$\|x'_h - P'x'_h\|_{\widehat{X}} \le \Lambda_{\mathrm{P}'}\|x'_h - Px'_h\|_{\widehat{X}}. \tag{A.5}$$

This and $y := Qy_h \in Y$ lead in $\widehat{\mathbf{(QO)}}$ to

$$a(P'x'_h - PMP'x'_h, Qy_h) \le \widehat{C_{\mathrm{qo}}}\|x'_h - P'x'_h\|_{\widehat{X}}\|y_h - Qy_h\|_{\widehat{Y}}.$$

Recall the definition of $x_h = Mx = A_h^{-1}Q^*Ax$ as the discrete solution to the right-hand side $a(x, Q\bullet)$ to verify $a_h(Mx, y_h) = a_h(x_h, y_h) = a(x, Qy_h)$. The combination with the last displayed inequality with $MP'x'_h = x'_h$ leads to

$$a(P'x'_h - PMP'x'_h, Qy_h) = a_h(x'_h, y_h) - a(Px'_h, Qy_h) \le \widehat{C_{\mathrm{qo}}}\|x'_h - Px'_h\|_{\widehat{X}}\|y_h - Qy_h\|_{\widehat{Y}}.$$

This and (A.5) prove $\widehat{\mathbf{(H)}}$ with $\widehat{\Lambda_{\mathrm{H}}} = \widehat{C_{\mathrm{qo}}}\Lambda_{\mathrm{P}'} = \widehat{C_{\mathrm{qo}}}(1 + \|P'\|(1 + C_{\mathrm{qo}}))$.

[Proof of "$\Longleftarrow$"] Given any $x \in X$ and $y_h \in Y_h$, let $x'_h := Mx \in X_h$ with $a_h(x'_h, y_h) = a(x, Qy_h)$. This shows in $\widehat{\mathbf{(H)}}$ that

$$a(x - PMx, Qy_h) \le \Lambda'_2\|x'_h - Px'_h\|_{\widehat{X}}\|y_h - Qy_h\|_{\widehat{Y}}.$$

This and the operator norm $\|a\|$ of $a(\bullet, \bullet)$ show

$$\begin{aligned} a(x - PMx, y) &= a(x - PMx, y - Qy_h) + a(x - PMx, Qy_h) \\ &\le \|a\|\|x - PMx\|_X\|y - Qy_h\|_Y + \Lambda'_2\|Mx - PMx\|_{\widehat{X}}\|y_h - Qy_h\|_{\widehat{Y}}. \end{aligned}$$

This and the elementary inequalities

$$\|x - PMx\|_X \le \|x - Mx\|_{\widehat{X}} + \|Mx - PMx\|_{\widehat{X}},\ \|Mx - PMx\|_{\widehat{X}} \le \Lambda_{\mathrm{P}}\|x - Mx\|_{\widehat{X}},$$

and

$$\|y - Qy_h\|_Y \le \|y - y_h\|_{\widehat{Y}} + \|y_h - Qy_h\|_{\widehat{Y}},\ \|y_h - Qy_h\|_{\widehat{Y}} \le \Lambda_{\mathrm{Q}}\|y - y_h\|_{\widehat{Y}}$$

conclude the proof of $\widehat{\mathbf{(QO)}}$ with $\widehat{C_{\mathrm{qo}}} := \|a\|(1 + \Lambda_{\mathrm{P}})(1 + \Lambda_{\mathrm{Q}}) + \Lambda'_2\Lambda_{\mathrm{P}}\Lambda_{\mathrm{Q}}$. □

Acknowledgements. The research of the first author has 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). The finalization of this paper has been supported by SPARC project (id 235) entitled *the mathematics and computation of plates* and SERB POWER Fellowship SPF/2020/000019.

References

[1] S. Agmon, Lectures on Elliptic Boundary Value Problems. AMS Chelsea Publishing, Providence, RI (2010).

[2] G.A. Baker, Finite element methods for elliptic equations using nonconforming elements. *Math. Comp.* **31** (1977) 45–59.

[3] H. Blum and R. Rannacher, On the boundary value problem of the biharmonic operator on domains with angular corners. *Math. Methods Appl. Sci.* **2** (1980) 556–581.

[4] D. Braess, Finite Elements, Theory, Fast Solvers, and Applications in Elasticity Theory, 3rd edition. Cambridge (2007).

[5] S.C. Brenner, Convergence of nonconforming multigrid methods without full elliptic regularity. *Math. Comp.* **68** (1999) 25–53.

[6] 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.

[7] 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.

[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] C. Carstensen, A unifying theory of a posteriori finite element error control. *Numer. Math.* **100** (2005) 617–637.

[10] C. Carstensen and D. Gallistl, Guaranteed lower eigenvalue bounds for the biharmonic equation. *Numer. Math.* **126** (2014) 33–51.

[11] C. Carstensen and F. Hellwig, Constants in discrete Poincaré and Friedrichs inequalities and discrete quasi-interpolation. *CMAM* **18** (2017) 433–450.

[12] C. Carstensen and J. Hu, A unifying theory of a posteriori error control for nonconforming finite element methods. *Numer. Math.* **107** (2007) 473–502.

[13] C. Carstensen and C. Merdon, Computational survey on a posteriori error estimators for nonconforming finite element methods for the Poisson problem. *J. Comput. Appl. Math.* **249** (2013) 74–94.

[14] C. Carstensen and N. Nataraj, Adaptive Morley FEM for the von Kármán equations with optimal convergence rates. *SIAM J. Numer. Anal.* **59** (2021) 696–719.

[15] C. Carstensen and N. Nataraj, Mathematics and computation of plates. Under preparation (2021).

[16] 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) 289–315.

[17] C. Carstensen and S. Puttkammer, How to prove the discrete reliability for nonconforming finite element methods. *J. Comput. Math.* **38** (2020) 142–175.

[18] C. Carstensen and S. Puttkammer, Direct guaranteed lower eigenvalue bounds with optimal a priori convergence rates for the bi-Laplacian. Preprint arXiv.org:2105.01505 (2021).

[19] C. Carstensen and H. Rabus, Axioms of adaptivity with separate marking for data resolution. *SIAM J. Numer. Anal.* **55** (2017) 2644–2665.

[20] 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) 509–558.

[21] 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) 337–353.

[22] C. Carstensen, D. Gallistl and J. Hu, A discrete Helmholtz decomposition with Morley finite element functions and the optimality of adaptive finite element schemes. *Comput. Math. Appl.* **68** (2014) 2167–2181.

[23] C. Carstensen, D. Gallistl and N. Nataraj, Comparison results of nonstandard P_2 finite element methods for the biharmonic problem. *ESAIM: M2AN* **49** (2015) 977–990.

[24] 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.

[25] C. Carstensen, N. Nataraj, C.R. Gopikrishnan and S. Devika, Unifying a priori and a posteriori error analysis for the lowest-order FEMs in fourth-order semi-linear problems with trilinear nonlinearity. Under preparation (2021).

[26] P.G. Ciarlet, The Finite Element Method for Elliptic Problems. North-Holland, Amsterdam (1978).

[27] G. Engel, K. Garikipati, T.J.R. Hughes, M.G. Larson, L. Mazzei and R.L. Taylor, Continuous/discontinuous finite element approximations of fourth-order elliptic problems in structural and continuum mechanics with applications to thin beams and plates, and strain gradient elasticity. *Comput. Methods Appl. Mech. Eng.* **191** (2002) 3669–3750.

[28] 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) 1093–1117 (electronic).

[29] D. Gallistl, Morley finite element method for the eigenvalues of the biharmonic operator. *IMA J. Numer. Anal.* **35** (2015) 1779–1811.

[30] E.H. Georgoulis and P. Houston, Discontinuous Galerkin methods for the biharmonic problem. *IMA J. Numer. Anal.* **29** (2009) 573–594.

[31] 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) 281–298.

[32] D. Gilbarg and N.S. Trudinger, Elliptic Partial Differential Equations of Second Order. *Classics in Mathematics*. Springer-Verlag, Berlin (2001). Reprint of the 1998 edition.

[33] P. Grisvard, Singularities in Boundary Value Problems. Vol. RMA 22. Masson & Springer-Verlag (1992).

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

[35] J.-L. Lions and E. Magenes, Non-homogeneous Boundary Value Problems and Applications. Vol. I. Springer-Verlag, New York-Heidelberg (1972). Translated from the French by P. Kenneth, Die Grundlehren der mathematischen Wissenschaften, Band 181.

[36] I. Mozolevski and E. Süli, A priori error analysis for the *hp*-version of the discontinuous Galerkin finite element method for the biharmonic equation. *Comput. Methods Appl. Math.* **3** (2003) 596–607.

[37] I. Mozolevski, E. Süli and P.R. Bösing, *hp*-version a priori error analysis of interior penalty discontinuous Galerkin finite element approximations to the biharmonic equation. *J. Sci. Comput.* **30** (2007) 465–491.

[38] J. Nečas, Les méthodes directes en théorie des équations elliptiques. Masson et Cie, Éditeurs, Paris; Academia, Éditeurs, Prague (1967).

[39] E. Süli and I. Mozolevski, *hp*-version interior penalty DGFEMs for the biharmonic equation. *Comput. Methods Appl. Mech. Eng.* **196** (2007) 1851–1863.

[40] L. Tartar, An Introduction to Sobolev Spaces and Interpolation Spaces. Springer, Berlin; Heidelberg (2010).

[41] A. Veeser and P. Zanotti, Quasi-optimal nonconforming methods for symmetric elliptic problems. I – Abstract theory. *SIAM J. Numer. Anal.* **56** (2018) 1621–1642.

[42] 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) 2871–2894.

[43] 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) 266–292.

[44] R. Verfürth, A Posteriori Error Estimation Techniques for Finite Element Methods. *Numerical Mathematics and Scientific Computation*. Oxford University Press, Oxford (2013).

Subscribe to Open (S2O)
A fair and sustainable open access model

This journal is currently published in open access under a Subscribe-to-Open model (S2O). S2O is a transformative model that aims to move subscription journals to open access. Open access is the free, immediate, online availability of research articles combined with the rights to use these articles fully in the digital environment. We are thankful to our subscribers and sponsors for making it possible to publish this journal in open access, free of charge for authors.

Please help to maintain this journal in open access!

Check that your library subscribes to the journal, or make a personal donation to the S2O programme, by contacting subscribers@edpsciences.org

More information, including a list of sponsors and a financial transparency report, available at: https://www.edpsciences.org/en/maths-s2o-programme