NON-CONFORMING, ANISOTROPIC, RECTANGULAR FINITE ELEMENTS OF ARBITRARY ORDER FOR THE STOKES PROBLEM*

THOMAS APEL† AND GUNAR MATTHIES‡

Abstract. The numerical solution of the Stokes problem is analysed for four families of rectangular, non-conforming finite elements of arbitrary order. While two of them become unstable as the maximum aspect ratio of the finite elements increases, the other two remain stable. A complete convergence analysis is given for these two families on two-level triangulations with isotropic rectangular macro-elements and certain anisotropic refinement strategies within the macros.

Key words. non-conforming finite elements, anisotropic meshes, consistency error, inf-sup condition

AMS subject classifications. 65N30, 65N15

1. Introduction.

We consider the stationary Stokes problem

$$\begin{aligned} -\triangle \boldsymbol{u} + \nabla p &= \boldsymbol{f} && \text{in } \Omega, \\ \operatorname{div} \boldsymbol{u} &= 0 && \text{in } \Omega, \\ \boldsymbol{u} &= \boldsymbol{0} && \text{on } \partial\Omega, \end{aligned} \tag{1.1}$$

in a two-dimensional domain Ω. For the discretisation of (1.1), one has to make the fundamental decision of choosing either a continuous or a discontinuous pressure approximation. We advocate the latter due to a better mass conservation on element level. Furthermore, coupled multi-grid solvers work in general better for discontinuous pressure spaces [16, 17]. Moreover, one has to distinguish between conforming and non-conforming velocity spaces. Since the degrees of freedom in two-dimensional non-conforming methods are edge-oriented, these methods are advantageous for parallel implementation due to missing cross-point communication.

Well-known examples for low order non-conforming discretisations of (1.1) can be found in [9, 12, 14]. Families of non-conforming triangular elements of arbitrarily high polynomial order were given in [21, 22] where only in [21] the application to the Stokes problem was considered. Three families of quadrilateral elements and a family of hexahedral elements were derived recently in [20]. They generalise the well-known rotated Q_1 element [23].

The spaces for approximating the velocity and the pressure in Stokes and Navier–Stokes problems cannot be chosen independently but have to satisfy the discrete inf-sup stability condition. The solution of the Navier–Stokes problem includes in general layer phenomena. To capture this anisotropic behaviour, one usually employs adequately refined anisotropic meshes. The small element size in one direction is used to compensate for large derivatives in this direction. In such applications it is essential that the constant in the discrete inf-sup condition does not degenerate for increasing aspect ratio of the elements. We refer to [6] for the state of the art in about 2002. However, the above mentioned families were till now not investigated concerning their behaviour on meshes which contain elements with large aspect ratio. In this paper we

*Bericht Nr. 374, Fakultät für Mathematik, Ruhr-Universität Bochum, August 29, 2006.

†Institut für Mathematik und Bauinformatik, Universität der Bundeswehr München, 85577 Neubiberg, Germany, thomas.apel@unibw.de

‡Fakultät für Mathematik, Ruhr-Universität Bochum, 44780 Bochum, Germany, gunar.matthies@ruhr-uni-bochum.de

will analyse on anisotropic meshes the properties of the three quadrilateral families of [20] and a further new family which generalises the first order element developed in [5]. All considered elements are parametric. The first mentioned three families generalise the parametric rotated Q_1 element which itself is not inf-sup stable on anisotropic meshes. Although we investigate families of arbitrary polynomial degree, we focus on an h-version of the finite element method, this means that the dependence of the constants on the polynomial degree is not elaborated.

The paper is organised as follows. We start with basic notation in Sect. 2. All four considered families will be introduced in Sect. 3. The numerical results which are obtained by using these families on isotropic meshes, differ only slightly. However, a numerical study on selected anisotropic meshes shows that for two families the constant in the inf-sup condition degenerates to zero as the aspect ratio increases. We succeeded in analysing the remaining two families and present interpolation error estimates, the proof of the inf-sup condition, and an estimate of the consistency error in Sect. 5, 6, and 7, respectively. Since the proofs are technical, we summarise our results already in Sect. 4.

2. Notation. Let Ω be a bounded domain in $\mathbb{R}^2$. For a one- or two-dimensional domain $G \subset \Omega$ we will use the Sobolev space $W^{m,p}(G)$ with the norms $\|\cdot\|_{m,p,G}$ and semi-norms $|\cdot|_{m,p,G}$ where the last index will be omitted if $G = \Omega$. In the case $p = 2$ we write, as usual, $H^m(G) := W^{m,2}(G)$ and omit the index p in the norms. Note further that $L^p(G) := W^{0,p}(G)$. The vector-valued version of all spaces will be denoted by the corresponding bold-face letter. The inner product in $L^2(G)$ and its vector-valued and tensor-valued versions will be denoted by $(\cdot,\cdot)_G$. For a one-dimensional manifold $F \subset \partial G$, the inner product in $L^2(F)$ and its vector-valued versions is written as $\langle\cdot,\cdot\rangle_F$.

Let $\boldsymbol{X} := \boldsymbol{H}^1_0(\Omega)$ and $M := L^2_0(\Omega) = \{v \in L^2(\Omega) \,:\, (v,1) = 0\}$. A weak formulation of (1.1) reads

Find $(\boldsymbol{u}, p) \in \boldsymbol{X} \times M$ such that

$$\begin{aligned} a(\boldsymbol{u},\boldsymbol{v}) - b(\boldsymbol{v},p) &= (\boldsymbol{f},\boldsymbol{v}) \qquad && \forall \boldsymbol{v} \in \boldsymbol{X}, \\ b(\boldsymbol{u},q) &= 0 && \forall q \in M, \end{aligned} \tag{2.1}$$

where

$$a(\boldsymbol{u},\boldsymbol{v}) := (\nabla\boldsymbol{u}, \nabla\boldsymbol{v}), \qquad b(\boldsymbol{v},q) := (\operatorname{div}\boldsymbol{v}, q).$$

Let $\{\mathcal{T}_h\}$ be a family of admissible triangulations of Ω into rectangular elements. For a domain $G \subset \mathbb{R}^d$, $d = 1, 2$, let $P_k(G)$ denote the space of polynomials of degree less than or equal to k. For discretising the pressure, we will use the space

$$M_h := \{p \in L^2_0(\Omega) \,:\, p|_K \in P_{r-1}(K)\, \forall K \in \mathcal{T}_h\}$$

of generally discontinuous, piecewise polynomials of degree less than or equal to $r-1$. The velocity will be approximated by a non-conforming finite element space $\boldsymbol{X}_h := X_h^2$. The detailed definition for several choices of X_h will be given in Sect. 3.

Since the bilinear forms a and b are not defined for general functions from $\boldsymbol{X}_h$, we introduce the bilinear forms a_h and b_h by

$$a_h(\boldsymbol{u}_h,\boldsymbol{v}_h) := \sum_{K\in\mathcal{T}_h} (\nabla\boldsymbol{u}_h, \nabla\boldsymbol{v}_h)_K, \qquad b_h(\boldsymbol{v}_h, q_h) := \sum_{K\in\mathcal{T}_h} (\operatorname{div}\boldsymbol{v}_h, q_h)_K,$$

which coincide for regular arguments with a and b, i. e. $a_h(\boldsymbol{u},\boldsymbol{v}) = a(\boldsymbol{u},\boldsymbol{v})$ for $\boldsymbol{u},\boldsymbol{v} \in \boldsymbol{X}$, and $b_h(\boldsymbol{v},q) = b(\boldsymbol{v},q)$ for $\boldsymbol{v} \in \boldsymbol{X}$, $q \in M$. Furthermore, we define

$$|v_h|_{1,h} := \left(\sum_{K\in\mathcal{T}_h} |v_h|_{1,K}^2 \right)^{1/2}$$

which will be a norm on X_h, see (3.2) for the definition of X_h. The discretised Stokes problem reads

Find $(\boldsymbol{u}_h, p_h) \in \boldsymbol{X}_h \times M_h$ such that

$$\begin{aligned} a_h(\boldsymbol{u}_h,\boldsymbol{v}_h) - b_h(\boldsymbol{v}_h,p_h) &= (\boldsymbol{f},\boldsymbol{v}_h) \qquad && \forall \boldsymbol{v}_h \in \boldsymbol{X}_h, \\ b_h(\boldsymbol{u}_h,q_h) &= 0 && \forall q_h \in M_h. \end{aligned} \tag{2.2}$$

For the spaces $\boldsymbol{X}_h$ and M_h we define the quantity

$$\beta_h := \inf_{q_h\in M_h} \sup_{\boldsymbol{v}_h\in\boldsymbol{X}_h} \frac{b_h(\boldsymbol{v}_h,q_h)}{|\boldsymbol{v}|_{1,h}\,\|q_h\|_0}. \tag{2.3}$$

The positivity of β_h ensures that (2.2) is uniquely solvable. Moreover, the inverse of β_h enters the estimate of the discretisation error, see [9, Sect. II.2.2]. Since we are interested in anisotropic discretisations, the quantity β_h should be bounded away from zero not only for h tending to zero but also for increasing aspect ratio. We call this the inf-sup condition and shall see that it is not generally satisfied.

Within this paper, we will frequently use the notation $x \lesssim y$ for the case that $x \le C\,y$ holds true where C is a positive constant which is independent of the mesh parameters, i.e., independent of the element sizes and the aspect ratio.

3. Families of finite elements. For a non-negative integer k, we denote by L_k the k-th one-dimensional Legendre polynomial normalised such that $L_k(1) = 1$. Let

$$\delta_k := \int_{-1}^{1} L_k^2(t)\,dt = \frac{2}{2k+1}, \qquad k \ge 0.$$

The Legendre polynomials are orthogonal, i.e.,

$$\int_{-1}^{1} L_i(t)L_j(t)\,dt = \delta_{ij}\delta_i$$

where δ_{ij} denotes the Kronecker delta. The integrated Legendre polynomials $\widehat{L}_k$ which are defined by

$$\widehat{L}_k(t) := \int_{-1}^{t} L_k(s)\,ds$$

are related to the Legendre polynomials by

$$\widehat{L}_k(t) = \begin{cases} \dfrac{\delta_k}{2}\big(L_{k+1}(t) - L_{k-1}(t)\big) & k \ge 1, \\ L_1(t) + L_0(t) & k = 0. \end{cases} \tag{3.1}$$

For defining the non-conforming finite element space, we will use the triple notation in the sense of Ciarlet, [10]. Let $(\widehat{K},\widehat{V},\widehat{\mathcal{N}})$ be a finite element on the reference cell $\widehat{K}$. All quantities which are connected to the reference element will be indicated by a hat. Note that as an exception both the Legendre polynomials L_k and the integrated Legendre polynomials $\widehat{L}_k$ are always defined on reference interval $[-1,1]$. Let $F_K := \widehat{K} \to K$ be the affine transformation which maps $\widehat{K}$ onto K. Then, the function space $V(K)$ on K is defined by

$$V(K) := \big\{ v = \hat{v} \circ F_K^{-1} \,:\, \hat{v} \in \widehat{V} \big\}.$$

Using the notation $\mathcal{S}_h^i$ and $\mathcal{S}_h^b$ for the set of interior and boundary edges of $\mathcal{T}_h$, the global scalar finite element space X_h is given by

$$\begin{aligned} X_h := \Big\{ & v \in L^2(\Omega) \,:\, v|_K \in V(K)\,\forall K \in \mathcal{T}_h,\ \langle q, [v]_E\rangle_E = 0\,\forall q \in P_{r-1}(E)\,\forall E \in \mathcal{S}_h^i, \\ & \langle q, v\rangle_E = 0\,\forall q \in P_{r-1}(E)\,\forall E \in \mathcal{S}_h^b \Big\}. \end{aligned} \tag{3.2}$$

As usual, $[v]_E$ denotes the jump of the function v across the edge E. The vector-valued finite element space $\boldsymbol{X}_h$ consists of X_h in each component.

Now, we will define several finite elements on the reference cell $\widehat{K} = (-1,1)^2$. We introduce the nodal functionals

$$\widehat{N}_{ij}(\hat{v}) := \frac{1}{\delta_i\delta_j} \int_{-1}^{1}\int_{-1}^{1} \hat{v}(\hat{x}_1,\hat{x}_2)\, L_i(\hat{x}_1) L_j(\hat{x}_2)\, d\hat{x}_2\, d\hat{x}_1, \qquad i,j \geq 0,$$

and, for non-negative integers k,

$$\widehat{N}_{k,-}(\hat{v}) := \frac{1}{\delta_k} \int_{-1}^{1} \hat{v}(\hat{x}_1,-1)\, L_k(\hat{x}_1)\, d\hat{x}_1, \qquad \widehat{N}_{k,+}(\hat{v}) := \frac{1}{\delta_k} \int_{-1}^{1} \hat{v}(\hat{x}_1,+1)\, L_k(\hat{x}_1)\, d\hat{x}_1,$$

$$\widehat{N}_{-,k}(\hat{v}) := \frac{1}{\delta_k} \int_{-1}^{1} \hat{v}(-1,\hat{x}_2)\, L_k(\hat{x}_2)\, d\hat{x}_2, \qquad \widehat{N}_{+,k}(\hat{v}) := \frac{1}{\delta_k} \int_{-1}^{1} \hat{v}(+1,\hat{x}_2)\, L_k(\hat{x}_2)\, d\hat{x}_2.$$

For a fixed integer $r \geq 1$, the set $\widehat{\mathcal{N}}_r$ of nodal functionals is given by

$$\widehat{\mathcal{N}}_r := \big\{ \widehat{N}_{ij} \,:\, 0 \leq i+j \leq r-2 \big\} \cup \big\{ \widehat{N}_{k,-}, \widehat{N}_{k,+}, \widehat{N}_{-,k}, \widehat{N}_{+,k} \,:\, k = 0,\ldots,r-1 \big\}. \tag{3.3}$$

Now, we look for suitable function spaces $\widehat{V}_r \supset P_r(\widehat{K})$ with the property that the set $\widehat{\mathcal{N}}_r$ of nodal functionals is unisolvent with respect to the space $\widehat{V}_r$. The following four choices for $\widehat{V}_r$ fulfil these conditions. An illustration of these families for $r = 11$ is given in Figure 3.1 where the grey boxes indicate the space $P_r(\widehat{K})$, the dark boxes stand for further functions of form $L_i(\hat{x}_1)L_j(\hat{x}_2)$, and if a pair of bright boxes is connected by an arc then only the difference of the involved functions belongs to the space $\widehat{V}_r$.

A first choice was already given in Example 5 in [15]. We set

$$\widehat{V}_r^1 := P_r(\widehat{K}) \oplus \widehat{R}_r \oplus \widehat{S}_r \tag{3.4}$$

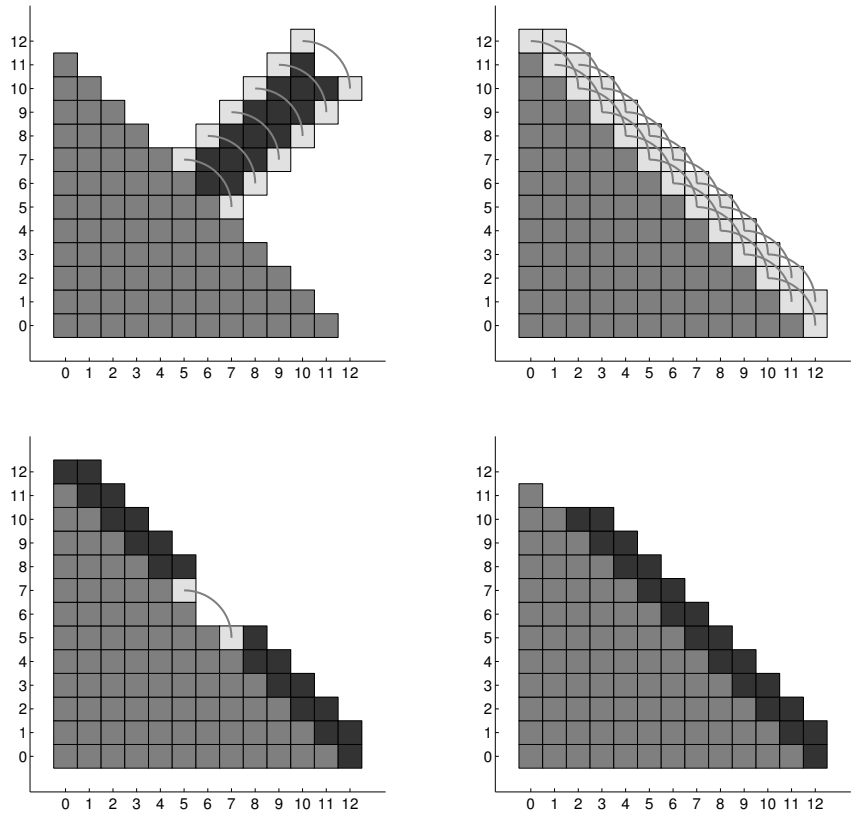


FIG. 3.1. *Illustration of the function spaces* $\widehat{V}_{11}^{\nu}$ *for* $\nu = 1, \ldots, 4$*, from top left to bottom right.*

where

$$\widehat{R}_r := \begin{cases} \operatorname{span}\Big\{ L_{i+1}(\hat{x}_1)L_i(\hat{x}_2),\ L_i(\hat{x}_1)L_{i+1}(\hat{x}_2) \\ \qquad L_{i+2}(\hat{x}_1)L_i(\hat{x}_2) - L_i(\hat{x}_1)L_{i+2}(\hat{x}_2),\ :\ i = r/2 \Big\}, & r \text{ even}, \\ \operatorname{span}\Big\{ L_{i+2}(\hat{x}_1)L_i(\hat{x}_2) - L_i(\hat{x}_1)L_{i+2}(\hat{x}_2)\ :\ i = (r-1)/2 \Big\}, & r \text{ odd}, \end{cases}$$

and

$$\widehat{S}_r := \operatorname{span}\Big\{ L_i(\hat{x}_1)L_i(\hat{x}_2),\ L_{i+1}(\hat{x}_1)L_i(\hat{x}_2),\ L_i(\hat{x}_1)L_{i+1}(\hat{x}_2), \\ L_{i+2}(\hat{x}_1)L_i(\hat{x}_2) - L_i(\hat{x}_1)L_{i+2}(\hat{x}_2)\ :\ r/2 < i \le r-1 \Big\}.$$

A second variant which is given by

$$\begin{aligned} \widehat{V}_r^2 := P_r(\widehat{K})\ &\oplus \operatorname{span}\Big\{ L_{i+2}(\hat{x}_1)L_{r-1-i}(\hat{x}_2) - L_i(\hat{x}_1)L_{r+1-i}(\hat{x}_2)\ ,\ i = 0, \ldots, r-1 \Big\} \\ &\oplus \operatorname{span}\Big\{ L_{i+2}(\hat{x}_1)L_{r-i}(\hat{x}_2) - L_i(\hat{x}_1)L_{r+2-i}(\hat{x}_2)\ ,\ i = 1, \ldots, r-1 \Big\} \end{aligned} \tag{3.5}$$

was introduced in [20]. There, also the following third version can be found:

$$\begin{aligned}\widehat{V}_r^3 :=& P_r(\widehat{K})\\ &\oplus \operatorname{span}\Big\{L_i(\hat{x}_1)L_{r+1-i}(\hat{x}_2),\ L_{r+1-i}(\hat{x}_1)L_i(\hat{x}_2)\ ,\ i=0,\dots,\lfloor r/2\rfloor-1\Big\}\\ &\oplus \operatorname{span}\Big\{L_i(\hat{x}_1)L_{r+2-i}(\hat{x}_2),\ L_{r+2-i}(\hat{x}_1)L_i(\hat{x}_2)\ ,\ i=1,\dots,\lfloor (r+1)/2\rfloor-1\Big\}\\ &\oplus \operatorname{span}\Big\{L_{\lfloor r/2\rfloor+2}(\hat{x}_1)L_{\lfloor r/2\rfloor}(\hat{x}_2)-L_{\lfloor r/2\rfloor}(\hat{x}_1)L_{\lfloor r/2\rfloor+2}(\hat{x}_2)\Big\}\end{aligned} \tag{3.6}$$

where $\lfloor s\rfloor$ is the largest integer which is less than or equal to s. Note that all these spaces generalise the parametric rotated Q_1 element, i. e. $\widehat{V}_1^1 = \widehat{V}_1^2 = \widehat{V}_1^3 = Q_1^{\mathrm{rot}}$. Finally, we introduce a new family,

$$\begin{aligned}\widehat{V}_r^4 :=& P_r(\widehat{K}) \oplus \operatorname{span}\big\{L_{r+1-i}(\hat{x}_1)L_i(\hat{x}_2),\ i=0,\dots,r-1\big\}\\ &\oplus \operatorname{span}\big\{L_{r+2-i}(\hat{x}_1)L_i(\hat{x}_2),\ i=1,\dots,r-1\big\}\end{aligned} \tag{3.7}$$

which recovers for $r = 1$ the space introduced in [5]. Note that this space is not symmetric with respect to $\hat{x}_1$ and $\hat{x}_2$. It is adapted to anisotropic elements which are elongated in the $\hat{x}_1$-direction. If the anisotropic element is stretched in the $\hat{x}_2$-direction, the roles of $\hat{x}_1$ and $\hat{x}_2$ must be exchanged in the definition of $\widehat{V}_r^4$. For isotropic elements one can use both variants or one of the families $\widehat{V}_r^\nu$, $\nu = 1, 2, 3$.

All four families are unisolvent with respect to the set $\mathcal{N}_r$ of nodal functionals defined in (3.3). The proof for $\widehat{V}_r^1$ can be found in [15]. The unisolvence results for $\widehat{V}_r^2$ and $\widehat{V}_r^3$ are given in [20]. The proof for the family $\widehat{V}_r^4$ will be given in the appendix, see Lemma A.1 on page 21.

TABLE 3.1
Results for second and third order elements on isotropic meshes.

pair	$\lvert\mathbf{u}-\mathbf{u}_h\rvert_{1,h}$ error	order	$\lVert\mathbf{u}-\mathbf{u}_h\rVert_0$ error	order	$\lVert p-p_h\rVert_0$ error	order
$V_2^1/P_1^{\mathrm{disc}}$	2.475-05	1.98	4.732-08	2.99	2.147-05	2.00
$V_2^2/P_1^{\mathrm{disc}}$	2.679-05	1.99	5.637-08	2.99	2.145-05	2.00
$V_2^3/P_1^{\mathrm{disc}}$	2.321-05	2.01	4.643-08	3.01	2.146-05	2.00
$V_2^4/P_1^{\mathrm{disc}}$	2.542-05	2.00	5.955-08	2.99	2.341-05	2.00
$V_3^1/P_2^{\mathrm{disc}}$	6.990-08	2.99	1.117-10	3.99	6.655-08	3.00
$V_3^2/P_2^{\mathrm{disc}}$	6.794-08	3.00	1.158-10	4.00	6.650-08	3.00
$V_3^3/P_2^{\mathrm{disc}}$	6.869-08	3.02	1.151-10	4.02	6.807-08	3.03
$V_3^4/P_2^{\mathrm{disc}}$	6.642-08	2.99	1.095-10	3.99	6.681-08	3.00

Table 3.1 shows that the results for different elements of the same order are comparable on isotropic meshes. The presented results were obtained on a uniform 64×64-decomposition of the unit square. The right hand side and the boundary conditions were chosen such that

$$u(x,y) = \begin{pmatrix}\sin(x)\sin(y)\\ \cos(x)\cos(y)\end{pmatrix}, \qquad p(x,y) = 2\cos(x)\sin(y) - p_0$$

is the given solution of (2.1) where the constant p_0 is determined by $p \in L_0^2(\Omega)$. The given orders were calculated from the data on the 64×64-mesh and the next coarser 32×32-mesh. All convergence rates are optimal.

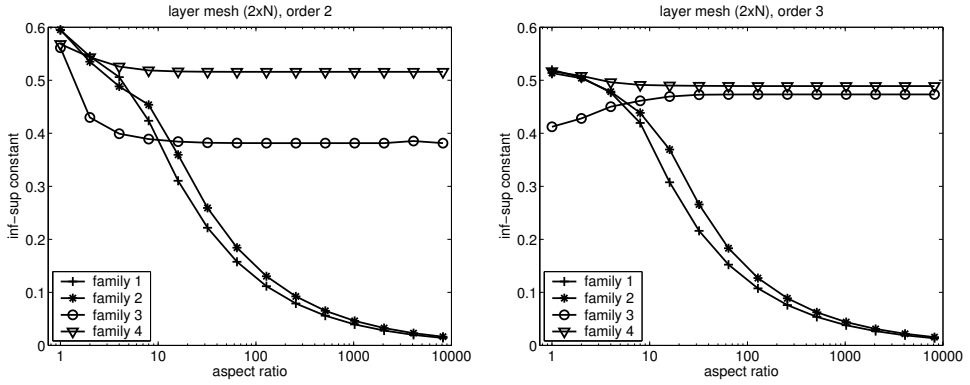


FIG. 3.2. *Plot of β_h against the aspect ratio for a family of layer meshes; left: $r = 2$, right: $r = 3$.*

In Figure 3.2 we display the inf-sup constants β_h for the finite element spaces $(\boldsymbol{X}_h, M_h)$ defined via $\widehat{V}_r^\nu$, $r = 2, 3$, $\nu = 1, 2, 3, 4$, on a family of meshes with increasing aspect ratio. These rectangular $2 \times N$-meshes are defined by their nodes (x_i, y_j), $x_i = i/2$, $i = 0, 1, 2$, $y_j = j/N$, $j = 0, \ldots, N$, yielding the aspect ratio $N/2$ for all elements. We see that the inf-sup constant β_h tends to zero for the families with $\nu = 1, 2$. This means that these families are not suited when anisotropic meshes are employed. For $\nu = 3, 4$, however, the inf-sup constant β_h remains bounded away from zero on a quite large level. Therefore we will investigate these two families further.

Since we will not focus on a p-version of the finite element method, the dependence of all constants on the parameter r is not elaborated within this paper. For $\nu = 3, 4$, we define by

$$\widehat{N}(\widehat{I}^\nu \hat{v}) = \widehat{N}(\hat{v}) \qquad \forall \widehat{N} \in \widehat{\mathcal{N}}$$

the canonical interpolation operators $\widehat{I}^\nu$ which map onto $\widehat{V}^\nu$. Note that these interpolation operators are well defined for all functions which belong to $W^{1,1}(\widehat{K})$. The global interpolant I_h^ν is defined elementwise as usual via $\big(I_h^\nu u\big)|_K = \big(\widehat{I}^\nu (u|_K \circ F_K)\big) \circ F_K^{-1}$. Since the edge functionals from $\widehat{N}$ are transformed into functionals which are associated with the edges of $\mathcal{T}_h$, the elementwise defined interpolation operators I_h^ν, $\nu = 3, 4$, map into the corresponding non-conforming space X_h. The vector-valued interpolation operators $\boldsymbol{I}_h^\nu : \boldsymbol{X} \to \boldsymbol{X}_h$ where $\boldsymbol{X}_h$ is based on $\widehat{V}^\nu$, $\nu = 3, 4$, are defined componentwise.

4. Main results. In the remaining part of this paper we will investigate the families 3 and 4 where we succeed to prove, on special meshes, the fulfilment of the inf-sup condition and estimates for the interpolation error and the consistency error. These three properties result in the final discretisation error estimate of order r. Since the proofs are quite lengthy, we summarise the results in this section. The proofs of Theorems 4.1 to 4.4 are given in Sections 5 to 7.

The first ingredient is an anisotropic local interpolation error estimate on an element K which is a rectangle with edge sizes $h_{K,1}$ and $h_{K,2}$. We assume that K is elongated in x_1-direction, i.e., $h_{K,1} \geq h_{K,2}$. We abbreviate $\partial_1 := \partial/\partial x_1$, $\partial_2 := \partial/\partial x_2$, $\hat{\partial}_1 := \partial/\partial \hat{x}_1$, $\hat{\partial}_2 := \partial/\partial \hat{x}_2$, and use the multi-index notation with $\alpha = (\alpha_1, \alpha_2)$, $\alpha_1, \alpha_2 \geq 0$, $|\alpha| = \alpha_1 + \alpha_2$, $D^\alpha = \partial_1^{\alpha_1} \partial_2^{\alpha_2}$, and $h_K^\alpha = h_{K,1}^{\alpha_1} h_{K,2}^{\alpha_2}$.

THEOREM 4.1 (Interpolation of smooth functions). *Let $u \in W^{\ell,p}(K)$ where $\ell \in \mathbb{N}$, $2 \le \ell \le r+1$, $p \in [1,\infty]$. Fix $m \in \{0,1\}$ and $q \in [1,\infty]$ such that $W^{\ell-m,p}(K) \hookrightarrow L^q(K)$. Then, the anisotropic interpolation error estimate*

$$|u - I_h^\nu u|_{m,q,K} \lesssim |K|^{1/q-1/p} \sum_{|\alpha|=\ell-m} h_K^\alpha |D^\alpha u|_{m,p,K}, \quad \nu = 3,4,$$

holds true.

Although the interpolation operators I_h^ν, $\nu = 3,4$, are well defined for functions from $W^{1,1}(K)$, the anisotropic interpolation error estimate holds only for functions from $W^{2,1}(K)$ if $r \ge 2$. The reason is that we need traces of derivatives in the proof, see also Example 5.5. Note that for $r = 1$ and $\ell = 1$, the anisotropic interpolation error estimate holds for I_h^4, see [4, Corollary 1], but not for I_h^3, see [7], since the parametric version is considered here.

Interpolation error estimates for regular functions are needed for the estimation of the best approximation error. However, for the proof of the fulfilment of the inf-sup condition, one needs an interpolation estimate for H^1-functions. Fortunately, the regularity assumption $u \in W^{2,1}(K)$ in Thm. 4.1 can be weakened if functions are considered which vanish on the short sides of K. To this end, we introduce the function space

$$W_E^{1,p}(K) := \big\{ v \in W^{1,p}(K) \,:\, v|_{E_{-,*}} = v|_{E_{+,*}} = 0 \big\}, \tag{4.1}$$

where we denote by $E_{-,*}$ and $E_{+,*}$ the edges on the left hand and the right hand side of the element K, respectively, see Figure 4.1. The anisotropic interpolation error estimate on $W_E^{1,p}(K)$ is stated in the following theorem.

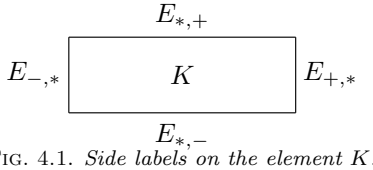


FIG. 4.1. *Side labels on the element K.*

THEOREM 4.2 (Interpolation of non-smooth functions). *Let $u \in W_E^{1,p}(K)$, $p \in [1,\infty]$. Fix $m \in \{0,1\}$ and $q \in [1,\infty]$ such that $W^{1-m,p}(K) \hookrightarrow L^q(K)$. Then, the anisotropic interpolation error estimate*

$$|u - I_h^\nu u|_{m,q,K} \lesssim |K|^{1/q-1/p} \sum_{|\alpha|=1-m} h^\alpha |D^\alpha u|_{m,p,K}, \quad \nu = 3,4,$$

holds true.

To be able to use this result for checking the inf-sup condition, we employ the macro-element technique from [8], see also [14, Sect. II.1.4]. To this end, let $\mathcal{T}_H$ be an admissible macro-triangulation of Ω into isotropic rectangular macro-elements Q,

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

The finally used triangulation $\mathcal{T}_h$ is obtained from $\mathcal{T}_H$ by applying local refinement strategies to the macro-elements Q in such a way that $\mathcal{T}_h$ is admissible. In particular,

we investigate here boundary layer and corner patches which are defined in the sequel. Other patches can be included provided a local inf-sup condition, as given by Lemmata 6.2 and 6.3, holds.

Let $Q = (x_-, x_+) \times (y_-, y_+)$ be a patch, and let $\tau_x = \{(x_{i-1}, x_i), i = 1, \dots, n\}$ and $\tau_y = \{(y_{j-1}, y_j), j = 1, \dots, m\}$ be decompositions of (x_-, x_+) and (y_-, y_+) into subintervals, respectively. In this way, the tensor-product triangulation $\tau_x \times \tau_y$ of Q is induced. If τ_x is the trivial decomposition of (x_-, x_+) into one interval then the obtained patch is called boundary layer patch. If $m = n$, $x_0 = x_-$, $y_0 = y_-$, $x_i = x_- + (x_+ - x_-)\sigma^{n-i}$, $i = 1, \dots, n$, and $y_j = y_- + (y_+ - y_-)\sigma^{n-j}$, $j = 1, \dots, n$, with a parameter $\sigma \in (0, 1)$, then the resulting patch is called corner patch. Typical boundary layer and corner patches are shown in Fig. 4.2. Note that the trivial decomposition of Q into one cell is allowed. The constants σ and n may vary from patch to patch. It is only required that the global mesh $\mathcal{T}_h$ is admissible.

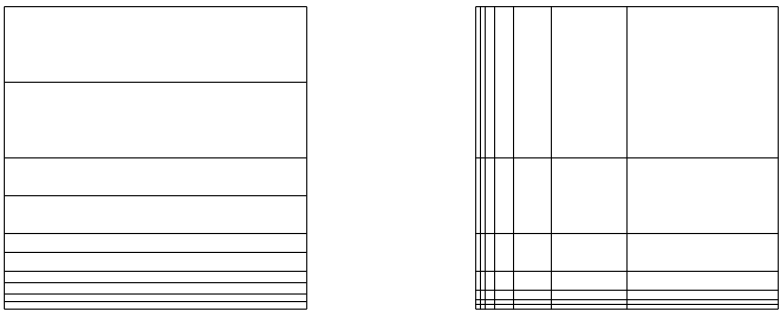


FIG. 4.2. *Typical boundary layer (left) and corner patches (right).*

THEOREM 4.3 (Inf-sup condition). *Let $\mathcal{T}_h$ be a triangulation of Ω as described above and consider $\mathbf{X}_h$ based on $\widehat{V}_r^3$ or $\widehat{V}_r^4$ with $r \geq 3$. Then, there exists a constant $\beta > 0$ such that $\beta_h \geq \beta$. As all constants in this paper, β is independent of the mesh parameter h and the maximal aspect ratio.*

Our method of proof works for all polynomial degrees $r \geq 3$, but not for $r = 2$. Since some numerical tests with $r = 2$ showed stability, the validity of a uniform inf-sup condition must be considered open for this case.

The consistency error is estimated in the following theorem.

THEOREM 4.4 (Consistency error). *Let $\boldsymbol{u} \in \boldsymbol{X} \cap \boldsymbol{H}^{r+1}(\Omega)$ and $p \in M \cap H^r(\Omega)$ be the solution of (2.1). Then, the consistency error estimate*

$$\begin{aligned}\sup_{\boldsymbol{v}_h \in \boldsymbol{X}_h} \frac{|a_h(\boldsymbol{u}, \boldsymbol{v}_h) - b_h(\boldsymbol{v}_h, p) - (\boldsymbol{f}, \boldsymbol{v}_h)|}{|\boldsymbol{v}_h|_{1,h}} \\ \lesssim \left(\sum_{K \in \mathcal{T}_h} \sum_{|\alpha|=r} h_K^{2\alpha} \left(\|D^\alpha(\partial_1 u_1 - p)\|_{0,K}^2 + \|D^\alpha \partial_1 u_2\|_{0,K}^2 \right) \right)^{1/2} \\ + \left(\sum_{K \in \mathcal{T}_h} h_{K,1}^2 \sum_{|\alpha|=r-1} h_K^{2\alpha} \|D^\alpha \boldsymbol{f}\|_{0,K}^2 \right)^{1/2}\end{aligned}$$

holds true. As above, the space $\mathbf{X}_h$ is based on $\widehat{V}_r^3$ or $\widehat{V}_r^4$.

Now, we can formulate our main result.

THEOREM 4.5 (Discretisation error estimate). *Let $(\boldsymbol{u},p)\in\boldsymbol{X}\times M$ and $(\boldsymbol{u}_h,p_h)\in\boldsymbol{X}_h\times M_h$ be the solutions of (2.1) and (2.2), respectively. The space $\mathbf{X}_h$ is based on $\widehat{V}_r^3$ or $\widehat{V}_r^4$ with $r\geq 3$. Then, the discretisation error is estimated by*

$$|\boldsymbol{u}-\boldsymbol{u}_h|_{1,h}+\|p-p_h\|_0\lesssim\left(\sum_{K\in\mathcal{T}_h}\sum_{|\alpha|=r}h_K^{2\alpha}\bigl(|D^\alpha\boldsymbol{u}|_{1,K}^2+\|D^\alpha p\|_{0,K}^2\bigr)\right)^{1/2}+\left(\sum_{K\in\mathcal{T}_h}h_{K,1}^2\sum_{|\alpha|=r-1}h_K^{2\alpha}\|D^\alpha\boldsymbol{f}\|_{0,K}^2\right)^{1/2}$$

provided that $\boldsymbol{u}\in\boldsymbol{H}^{r+1}(\Omega)$, $p\in H^r(\Omega)$, and $\boldsymbol{f}\in\boldsymbol{H}^{r-1}(\Omega)$.

Proof. Under the assumption that the uniform inf-sup condition $\beta_h\geq\beta>0$ is valid, Proposition II.2.16 of [9] gives

$$|\boldsymbol{u}-\boldsymbol{u}_h|_{1,h}+\|p-p_h\|_0\lesssim\inf_{\boldsymbol{v}_h\in\boldsymbol{X}_h}|\boldsymbol{u}-\boldsymbol{v}_h|_{1,h}+\inf_{q_h\in M_h}\|p-q_h\|_0+\sup_{\boldsymbol{v}_h\in\boldsymbol{X}_h}\frac{|a_h(\boldsymbol{u},\boldsymbol{v}_h)-b_h(\boldsymbol{v}_h,p)-(\boldsymbol{f},\boldsymbol{v}_h)|}{|\boldsymbol{v}_h|_{1,h}}.$$

The estimate for the error of the best approximation is obtained by combining the local estimates from Thm. 4.1. The consistency term is estimated in Thm. 4.4. The assertion is concluded by observing that parts of the right hand side of the consistency error estimate can be bounded by best approximation terms. □

Note that the estimate is standard in the sense that we have a method of order r for $\boldsymbol{u}\in\boldsymbol{H}^{r+1}(\Omega)$, $p\in H^r(\Omega)$, and $\boldsymbol{f}\in\boldsymbol{H}^{r-1}$. The advantage is that we can compensate large derivatives of $\boldsymbol{u}$ and p in x_2-direction by choosing small values for $h_{K,2}$.

5. Proof of the interpolation error estimates. Estimates for $\|u-I_hu\|_{0,q,K}$ and $\|\partial_1(u-I_hu)\|_{0,q,K}$ can be proved in the standard way since the element K is assumed to be stretched in x_1-direction. Even non-smooth functions $u\in W^{1,p}(K)$ can be treated by this approach.

LEMMA 5.1. *Let $u\in W^{\ell,p}(K)$ where $\ell\in\mathbb{N}$, $1\leq\ell\leq r+1$, $p\in[1,\infty]$. Fix $m\in\{0,1\}$ and $q\in[1,\infty]$ such that $W^{\ell-m,p}(K)\hookrightarrow L^q(K)$. Then, the anisotropic interpolation error estimate*

$$\|\partial_1^m(u-I_h^\nu u)\|_{0,q,K}\lesssim|K|^{1/q-1/p}\sum_{|\alpha|=\ell-m}h_K^\alpha|D^\alpha u|_{m,p,K},\quad\nu=3,4,$$

holds true.

Proof. We prove the estimate first on the reference element $\widehat{K}$. Since the interpolation operator $\widehat{I}^\nu$ preserves polynomials of degree r and since $\widehat{I}^\nu:W^{\ell,p}(\widehat{K})\to W^{m,q}(\widehat{K})$ is a bounded operator we have for all $\hat{w}\in P_{\ell-1}(\widehat{K})\subset P_r(\widehat{K})\subset\widehat{V}_r^\nu$

$$\|\hat{\partial}_1^m(\hat{u}-\widehat{I}^\nu\hat{u})\|_{0,q,\widehat{K}}=\|\hat{\partial}_1^m(\hat{u}-\hat{w})-\hat{\partial}_1^m\widehat{I}^\nu(\hat{u}-\hat{w})\|_{0,q,\widehat{K}}\lesssim\|\hat{u}-\hat{w}\|_{\ell,p,\widehat{K}}.$$

By using the Deny–Lions lemma, we obtain

$$\|\hat{\partial}_1^m(\hat{u}-\widehat{I}^\nu\hat{u})\|_{0,q,\widehat{K}}\lesssim|\hat{u}|_{\ell,p,\widehat{K}}.$$

The assertion is proved by transforming this estimate from $\widehat{K}$ to K using $h_{K,2} \le h_{K,1}$. □

The challenge is the proof for $\|\partial_2(u - I_h^\nu u)\|_{0,q,K}$, see [1, Ex. 2.1]. Due to [1, Lemma 2.2] we succeed if we find a set $\widehat{\mathcal{F}}_r$ of $\dim \hat{\partial}_2 \widehat{V}_r^\nu$ linear functionals, such that

$$\widehat{F} \in \big(W^{\ell-1,p}(\widehat{K})\big)' \quad \forall \widehat{F} \in \widehat{\mathcal{F}}_r, \tag{5.1}$$

$$\widehat{F}\big(\hat{\partial}_2(\hat{u} - \widehat{I}^\nu \hat{u})\big) = 0 \quad \forall \widehat{F} \in \widehat{\mathcal{F}}_r, \quad \forall \hat{u} \in W^{\ell,p}(\widehat{K}), \tag{5.2}$$

$$\hat{v} \in \hat{\partial}_2 \widehat{V}_r^\nu \quad \text{and} \quad \widehat{F}(\hat{v}) = 0 \quad \forall \widehat{F} \in \widehat{\mathcal{F}}_r \quad \Rightarrow \quad \hat{v} \equiv 0. \tag{5.3}$$

To this end, we define the set

$$\begin{aligned} \widehat{\mathcal{F}}_r := &\big\{\widehat{F}_{i,*} \,:\, i = 0, \dots, r-1\big\} \cup \big\{\widehat{F}_{ij} \,:\, 0 \le i + j \le r-2\big\} \\ &\cup \big\{\widehat{F}_{+,i}, \widehat{F}_{-,i} \,:\, i = 1, \dots, r-1\big\} \end{aligned} \tag{5.4}$$

with

$$\widehat{F}_{i,*}(\hat{v}) := \frac{1}{\delta_i \delta_0} \int_{-1}^{1}\int_{-1}^{1} L_i(\hat{x}_1) L_0(\hat{x}_2) \hat{v}(\hat{x}_1, \hat{x}_2)\, d\hat{x}_1\, d\hat{x}_2,$$

$$\widehat{F}_{ij}(\hat{v}) := \frac{2}{\delta_i \delta_j} \int_{-1}^{1}\int_{-1}^{1} L_i(\hat{x}_1) \widehat{L}_j(\hat{x}_2) \hat{v}(\hat{x}_1, \hat{x}_2)\, d\hat{x}_1\, d\hat{x}_2,$$

$$\widehat{F}_{+,i}(\hat{v}) := \frac{2}{\delta_i} \int_{-1}^{1} \widehat{L}_i(\hat{x}_2) \hat{v}(+1, \hat{x}_2)\, d\hat{x}_2,$$

$$\widehat{F}_{-,i}(\hat{v}) := \frac{2}{\delta_i} \int_{-1}^{1} \widehat{L}_i(\hat{x}_2) \hat{v}(-1, \hat{x}_2)\, d\hat{x}_2.$$

One easily checks that both the dimension of $\hat{\partial}_2 \widehat{V}_r^\nu$, $\nu = 3, 4$, and the number of functionals in $\widehat{\mathcal{F}}_r$ are equal to $r^2/2 + 5r/2 - 2$. It remains to prove that the set $\widehat{\mathcal{F}}_r$ satisfies the properties (5.1)–(5.3). We start with the unisolvence property (5.3) which needs the most complex proof, done for $\widehat{V}_r^4$ and $\widehat{V}_r^3$ separately.

LEMMA 5.2. *The set $\widehat{\mathcal{F}}_r$ of functionals is unisolvent on $\hat{\partial}_2 \widehat{V}_r^4$.*

Proof. The definition of $\widehat{V}_r^4$ gives immediately

$$\begin{aligned} \hat{\partial}_2 \widehat{V}_r^4 = P_{r-1}(\widehat{K}) &\oplus \operatorname{span}\big(L_{r-i}(\hat{x}_1) L_i(\hat{x}_2),\ i = 0, \dots, r-2\big) \\ &\oplus \operatorname{span}\big(L_{r+1-i}(\hat{x}_1) L_i(\hat{x}_2),\ i = 0, \dots, r-2\big). \end{aligned} \tag{5.5}$$

We will show that $\hat{v} \in \hat{\partial}_2 \widehat{V}_r$ and $\widehat{F}(\hat{v}) = 0$ for all $\widehat{F} \in \widehat{\mathcal{F}}_r$ gives $\hat{v} \equiv 0$. To this end, a general element $\hat{v} \in \hat{\partial}_2 \widehat{V}_r$ is written as

$$\hat{v} = \sum_{0 \le i+j \le r-1} \alpha_{ij} L_i(\hat{x}_1) L_j(\hat{x}_2) + \sum_{i=0}^{r-2} \beta_i L_{r-i}(\hat{x}_1) L_i(\hat{x}_2) + \sum_{i=0}^{r-2} \gamma_i L_{r+1-i}(\hat{x}_1) L_i(\hat{x}_2).$$

Applying $\widehat{F}_{n,*}$ to $\hat{v}$ and using the orthogonality of the Legendre polynomials it follows that

$$0 = \widehat{F}_{n,*}(\hat{v}) = \alpha_{n0}, \qquad n = 0, \dots, r-1. \tag{5.6}$$

Hence, the function $\hat{v}$ can be written as

$$\hat{v} = \sum_{i=0}^{r-2} \sum_{j=1}^{r-1-i} \alpha_{ij} L_i(\hat{x}_1) L_j(\hat{x}_2) + \sum_{i=0}^{r-2} \beta_i L_{r-i}(\hat{x}_1) L_i(\hat{x}_2) + \sum_{i=0}^{r-2} \gamma_i L_{r+1-i}(\hat{x}_1) L_i(\hat{x}_2).$$

Applying $\widehat{F}_{m,0}$, $m = 0, \ldots, r-2$, and using the orthogonality of the Legendre polynomials together with $2/\delta_0 = 1$ results in

$$\begin{aligned} 0 = \widehat{F}_{m,0}(\hat{v}) = &\sum_{i=0}^{r-2} \sum_{j=1}^{r-1-i} \alpha_{ij} \delta_{im} (\delta_{j1}\delta_1 + \delta_{j0}\delta_0) + \sum_{i=0}^{r-2} \beta_i \delta_{m,r-i} (\delta_{i1}\delta_1 + \delta_{i0}\delta_0) \\ &+ \sum_{i=0}^{r-2} \gamma_i \delta_{m,r+1-i} (\delta_{i1}\delta_1 + \delta_{i0}\delta_0) = \alpha_{m1}\delta_1 \end{aligned} \tag{5.7}$$

since all other terms vanish due to the range of i and j in the various sums. Hence, $\alpha_{m1} = 0$ for $m = 0, \ldots, r-2$. Now, we apply $\widehat{F}_{mn}$ for $n \geq 1$, i. e. for $n = 1, \ldots, r-2$, $m = 0, \ldots, r-2-n$. Applying relation (3.1) between Legendre polynomials and integrated Legendre polynomials, we obtain

$$\begin{aligned} 0 = \widehat{F}_{mn}(\hat{v}) = &\sum_{i=0}^{r-2} \sum_{j=2}^{r-1-i} \alpha_{ij} \delta_{im} (\delta_{j,n+1}\delta_{n+1} + \delta_{j,n-1}\delta_{n-1}) \\ &+ \sum_{i=0}^{r-2} \beta_i \delta_{m,r-i} (\delta_{i,n+1}\delta_{n+1} + \delta_{i,n-1}\delta_{n-1}) \\ &+ \sum_{i=0}^{r-2} \gamma_i \delta_{m,r+1-i} (\delta_{i,n+1}\delta_{n+1} + \delta_{i,n-1}\delta_{n-1}) \end{aligned}$$

where again the orthogonality of the Legendre polynomials was used. The first double sum can be simplified, and the sums with β_i and γ_i are again zero due to $m+n \leq r-2$. Thus, the equations become

$$0 = \alpha_{m,n+1}\delta_{n+1} - \alpha_{m,n-1}\delta_{n-1}.$$

For $n = 1$ one gets $\alpha_{m2} = 0$, $m = 0, \ldots, r-3$, since $\alpha_{m0} = 0$. By induction we obtain with (5.6) and (5.7)

$$\alpha_{ij} = 0, \qquad 0 \leq i + j \leq r - 1.$$

Hence, we have

$$\hat{v} = \sum_{i=0}^{r-2} \Big(\beta_i L_{r-i}(\hat{x}_1) + \gamma_i L_{r+1-i}(\hat{x}_1) \Big) L_i(\hat{x}_2).$$

Using $\widehat{F}_{+,n}(\hat{v}) = 0$ and $\widehat{F}_{-,n}(\hat{v}) = 0$, $n = 1, \ldots, r-1$, we get

$$\begin{aligned} 0 &= \sum_{i=0}^{r-2} \Big(\beta_i (+1)^{r-i} + \gamma_i (+1)^{r+1-i} \Big) \big(\delta_{n+1,i}\delta_{n+1} - \delta_{n-1,i}\delta_{n-1} \big), \\ 0 &= \sum_{i=0}^{r-2} \Big(\beta_i (-1)^{r-i} + \gamma_i (-1)^{r+1-i} \Big) \big(\delta_{n+1,i}\delta_{n+1} - \delta_{n-1,i}\delta_{n-1} \big). \end{aligned}$$

In the case $n = r-1$ and $n = r-2$ these equations reduce to

$$\begin{aligned} \beta_{r-2} + \gamma_{r-2} &= 0, \qquad \beta_{r-2} - \gamma_{r-2} = 0, \\ \beta_{r-3} + \gamma_{r-3} &= 0, \qquad \beta_{r-3} - \gamma_{r-3} = 0, \end{aligned}$$

i.e., we get

$$\beta_{r-2} = \gamma_{r-2} = \beta_{r-3} = \gamma_{r-3} = 0.$$

Using this result, we get by induction on $n = r-3, \ldots, 1$ the final result

$$\beta_i = \gamma_i = 0, \qquad i = 0, \ldots, r-2.$$

This means that $\hat{v} \equiv 0$. □

LEMMA 5.3. *The set $\widehat{\mathcal{F}}_r$ of functionals is unisolvent on $\hat{\partial}_2 \widehat{V}_r^3$.*

Proof. In order to get an expression for $\hat{\partial}_2 \widehat{V}_r^3$, we have to consider the difference term carefully. Setting

$$k := \lfloor r/2 \rfloor \tag{5.8}$$

and using the differentiated version of (3.1), we obtain

$$\begin{aligned} &\frac{\partial}{\partial \hat{x}_2}\big(L_{k+2}(\hat{x}_1)L_k(\hat{x}_2) - L_k(\hat{x}_1)L_{k+2}(\hat{x}_2)\big) \\ &\quad = L_{k+2}(\hat{x}_1)\left[L'_{k-2}(\hat{x}_2) + \frac{2}{\delta_{k-1}}L_{k-1}(\hat{x}_2)\right] - L_k(\hat{x}_1)\left[L'_k(\hat{x}_2) + \frac{2}{\delta_{k+1}}L_{k+1}(\hat{x}_2)\right] \\ &\quad = \frac{2}{\delta_{k+1}}\left[\frac{\delta_{k+1}}{\delta_{k-1}}L_{k+2}(\hat{x}_1)L_{k-1}(\hat{x}_2) - L_k(\hat{x}_1)L_{k+1}(\hat{x}_2)\right] + q_{r-1} \end{aligned}$$

with $q_{r-1} \in P_{r-1}(\widehat{K})$. Hence, the space $\hat{\partial}_2 \widehat{V}_r^3$ can be written as

$$\begin{aligned} \hat{\partial}_2 \widehat{V}_r^3 =& P_{r-1}(\widehat{K}) \\ &\oplus \operatorname{span}\Big\{L_{r-i}(\hat{x}_1)L_i(\hat{x}_2), L_{r+1-i}(\hat{x}_1)L_i \ : \ i = 0, \ldots, k-2 \text{ or } i = k+2, \ldots, r\Big\} \\ &\oplus \operatorname{span}\left\{\frac{\delta_{k+1}}{\delta_{k-1}}L_{k+2}(\hat{x}_1)L_{k-1}(\hat{x}_2) - L_k(\hat{x}_1)L_{k+1}(\hat{x}_2)\right\} \\ &\oplus \operatorname{span}\Big\{L_m(\hat{x}_1)L_n(\hat{x}_2)\Big\} \end{aligned} \tag{5.9}$$

where k is defined in (5.8) and

$$(m,n) := \begin{cases} (k+3, k-1) & \text{for } r \text{ odd}, \\ (k-1, k+1) & \text{for } r \text{ even}. \end{cases} \tag{5.10}$$

As usual, we will show that $\hat{v} \in \hat{\partial}_2 \widehat{V}_r^3$ with $\widehat{F}(\hat{v}) = 0$ for all $\widehat{F} \in \widehat{\mathcal{F}}_r$ results in $\hat{v} \equiv 0$. In order to do so, we write $\hat{v}$ as a linear combination of the basis functions given in (5.9) and the Legendre basis of $P_{r-1}(\widehat{K})$. Proceeding as in the beginning of the proof for Lemma 5.2, we get that all coefficients for functions from $P_{r-1}(\widehat{K})$

vanish. Hence, the function $\hat{v}$ can be represented as

$$\begin{aligned}\hat{v} =& \sum_{i=0}^{k-2} \Big(\beta_i L_{r-i}(\hat{x}_1)L_i(\hat{x}_2) + \gamma_i L_{r+1-i}(\hat{x}_1)L_i(\hat{x}_2)\Big) \\ &+ \sum_{i=k+2}^{r} \Big(\beta_i L_{r-i}(\hat{x}_1)L_i(\hat{x}_2) + \gamma_i L_{r+1-i}(\hat{x}_1)L_i(\hat{x}_2)\Big) \\ &+ \mu\left(\frac{\delta_{k+1}}{\delta_{k-1}} L_{k+2}(\hat{x}_1)L_{k-1}(\hat{x}_2) - L_k(\hat{x}_1)L_{k+1}(\hat{x}_2)\right) \\ &+ \vartheta L_m(\hat{x}_1)L_n(\hat{x}_2)\end{aligned}$$

with k from (5.8) and m, n from (5.10). Now, the cases of odd and even r have to be distinguished. Let r be even. Applying the nodal functionals $\widehat{F}_{\pm,k}$, we obtain

$$\begin{aligned}0 = \widehat{F}_{\pm,k}(\hat{v}) =& \mu\frac{\delta_{k+1}}{\delta_{k-1}} \int_{-1}^{1} (\pm 1)^{k+2} L_{k-1}(\hat{x}_2)\big(L_{k+1}(\hat{x}_2) - L_{k-1}(\hat{x}_2)\big)\, d\hat{x}_2 \\ &- \mu \int_{-1}^{1} (\pm 1)^{k} L_{k+1}(\hat{x}_2)\big(L_{k+1}(\hat{x}_2) - L_{k-1}(\hat{x}_2)\big)\, d\hat{x}_2 \\ &+ \vartheta \int_{-1}^{1} (\pm 1)^{k-1} L_{k+1}(\hat{x}_2)\big(L_{k+1}(\hat{x}_2) - L_{k-1}(\hat{x}_2)\big)\, d\hat{x}_2 \\ =& (\pm 1)^k \left[\frac{\delta_{k+1}}{\delta_{k-1}}(-\delta_{k-1})\mu - \delta_{k+1}\mu + (\pm 1)\delta_{k+1}\vartheta\right]\end{aligned}$$

by using again the orthogonality of the Legendre polynomials. Hence, we get

$$0 = -2\mu + \vartheta, \qquad 0 = -2\mu - \vartheta$$

which immediately gives $\mu = \vartheta = 0$. In a similar way, we obtain that $\mu = \vartheta = 0$ also for odd r. For showing that the coefficients β_i and γ_i will vanish, we proceed as in the end of the proof of Lemma 5.2. Here, we apply $\widehat{F}_{\pm,\ell}$ for $\ell = k+1, \ldots, r-1$ and $\ell = k-1, \ldots, 0$. Finally, we end up with $\hat{v} \equiv 0$ since all coefficients vanish. □

We are now prepared to prove the anisotropic interpolation error estimate for $\partial_2(u - I_h^\nu u)$, $\nu = 3, 4$. We start with the case of smooth functions $u \in W^{\ell,p}(K)$, $\ell \geq 2$.

LEMMA 5.4. *Let $u \in W^{\ell,p}(K)$ where $\ell \in \mathbb{N}$, $2 \leq \ell \leq r+1$, $p \in [1,\infty]$. Fix $q \in [1,\infty]$ such that $W^{\ell-1,p}(K) \hookrightarrow L^q(K)$. Then, the anisotropic interpolation error estimate*

$$\|\partial_2(u - I_h^\nu u)\|_{0,q,K} \lesssim |K|^{1/q-1/p} \sum_{|\alpha|=\ell-1} h_K^\alpha |D^\alpha u|_{1,p,K}, \quad \nu = 3, 4,$$

holds true.

Proof. We prove first an estimate on the reference element $\widehat{K}$. In order to be able to apply [1, Lemma 2.2], it remains to show that the functionals $\widehat{F} \in \widehat{\mathcal{F}}_r$ satisfy the properties (5.1) and (5.2) since the property (5.3) was already proved in Lemmata 5.2 and 5.3.

Property (5.1) is satisfied since $\ell - 1 \geq 1$ and all functionals are well defined for $\hat{u} \in W^{1,1}(\widehat{K})$. We note here that weaker functions cannot be used since their trace on an edge of $\widehat{K}$ must be integrable. This is the reason for the assumption $\ell \geq 2$.

Property (5.2) can also be proved simultaneously for both $\widehat{V}_r^3$ and $\widehat{V}_r^4$ since the nodal functionals $\widehat{\mathcal{N}}$ are the same. We have

$$\begin{aligned}
\widehat{F}_{i,*}\big(\hat{\partial}_2(\hat{u} - \widehat{I}^\nu \hat{u})\big) =& \frac{1}{\delta_i \delta_0} \int_{-1}^{1} \int_{-1}^{1} L_i(\hat{x}_1) L_0(\hat{x}_2) \hat{\partial}_2\big(\hat{u} - \widehat{I}^\nu \hat{u}\big)(\hat{x}_1, \hat{x}_2)\, d\hat{x}_1\, d\hat{x}_2 \\
=& \frac{1}{\delta_i \delta_0} \int_{-1}^{1} L_i(\hat{x}_1)\big(\hat{u} - \widehat{I}^\nu \hat{u}\big)(\hat{x}_1, +1)\, d\hat{x}_1 \\
& \qquad - \frac{1}{\delta_i \delta_0} \int_{-1}^{1} L_i(\hat{x}_1)\big(\hat{u} - \widehat{I}^\nu \hat{u}\big)(\hat{x}_1, -1)\, d\hat{x}_1 \\
=& \frac{1}{\delta_0}\big(\widehat{N}_{i,+}(\hat{u} - \widehat{I}^\nu \hat{u}) - \widehat{N}_{i,-}(\hat{u} - \widehat{I}^\nu \hat{u})\big) = 0, \qquad i = 0, \ldots, r-1, \\
\widehat{F}_{\pm,i}\big(\hat{\partial}_2(\hat{u} - \widehat{I}^\nu \hat{u})\big) =& \frac{2}{\delta_i} \int_{-1}^{1} \widehat{L}_i(\hat{x}_2) \hat{\partial}_2\big(\hat{u} - \widehat{I}^\nu \hat{u}\big)(\pm 1, \hat{x}_2)\, d\hat{x}_2 \\
=& -\frac{2}{\delta_i} \int_{-1}^{1} L_i(\hat{x}_2)\big(\hat{u} - \widehat{I}^\nu \hat{u}\big)(\pm 1, \hat{x}_2)\, d\hat{x}_2 \\
=& \widehat{N}_{\pm,i}\big(\hat{u} - \widehat{I}^\nu \hat{u}\big) = 0, \qquad i = 1, \ldots, r-1, \\
\widehat{F}_{ij}\big(\hat{\partial}_2(\hat{u} - \widehat{I}^\nu \hat{u})\big) =& \frac{2}{\delta_i \delta_j} \int_{-1}^{1} L_i(\hat{x}_1) \int_{-1}^{1} \widehat{L}_j(\hat{x}_2) \hat{\partial}_2\big(\hat{u} - \widehat{I}^\nu \hat{u}\big)(\hat{x}_1, \hat{x}_2)\, d\hat{x}_2\, d\hat{x}_1 \\
=& \delta_{j0} \frac{2}{\delta_i \delta_0} \int_{-1}^{1} L_i(\hat{x}_1) \widehat{L}_0(+1)\big(\hat{u} - \widehat{I}^\nu \hat{u}\big)(\hat{x}_1, 1)\, d\hat{x}_1 \\
& - \frac{2}{\delta_i \delta_j} \int_{-1}^{1} L_i(\hat{x}_1) \int_{-1}^{1} L_j(\hat{x}_2)\big(\hat{u} - \widehat{I}^\nu \hat{u}\big)(\hat{x}_1, \hat{x}_2)\, d\hat{x}_2\, d\hat{x}_1 \\
=& 2\delta_{j0} \widehat{N}_{i,+}\big(\hat{u} - \widehat{I}^\nu \hat{u}\big) - 2\widehat{N}_{ij}\big(\hat{u} - \widehat{I}^\nu \hat{u}\big) = 0, \qquad 0 \leq i + j \leq r - 2.
\end{aligned}$$

Consequently, Lemma 2.2 of [1] delivers

$$\|\hat{\partial}_2(\hat{u} - \widehat{I}^\nu \hat{u})\|_{0,q,\widehat{K}} \lesssim |\hat{\partial}_2 \hat{u}|_{\ell-1,p,\widehat{K}}. \tag{5.11}$$

Note that the condition $\hat{u} \in C(\widehat{K})$ in [1, Lemma 2.2] is not needed here but only for Lagrangian finite elements. The transformation of (5.11) from $\widehat{K}$ to K gives

$$\begin{aligned}
\|\partial_2(u - I_h^\nu u)\|_{0,q,K} &\lesssim |K|^{1/q-1/p} \sum_{|\alpha|=\ell-1} h_K^\alpha \|D^\alpha \partial_2 u\|_{0,p,K} \\
&\lesssim |K|^{1/q-1/p} \sum_{|\alpha|=\ell-1} h_K^\alpha |D^\alpha u|_{1,p,K}.
\end{aligned}$$

This is the desired estimate. □

EXAMPLE 5.5. *We show that the assumption $\ell \geq 2$ in Lemma 5.4 is essential. The validity of*

$$\|\partial_2(u - I_h^\nu u)\|_{0,q,K} \lesssim |K|^{1/q-1/p}|u|_{1,p,K}$$

is equivalent to

$$\|\hat\partial_2(\hat u - \widehat I^\nu \hat u)\|_{0,q,\widehat K} \lesssim \|\hat\partial_2 \hat u\|_{0,p,\widehat K}, \tag{5.12}$$

compare [1, Example 2.1]. In order to show that estimate (5.12) *is not valid, we adapt Example 2.3 of [1] and consider the sequence of functions*

$$\hat u_\varepsilon(\hat x_1, \hat x_2) = \hat x_2 \cdot \min\left\{1, \varepsilon \left|\ln \frac{\hat x_1 + 1}{2}\right|\right\}, \quad \varepsilon \to +0.$$

The limit function is defined pointwise by

$$\hat u_0(\hat x_1, \hat x_2) = \lim_{\varepsilon\to 0} \hat u_\varepsilon(\hat x_1, \hat x_2) = \begin{cases} \hat x_2 & \text{for } \hat x_1 = -1, \\ 0 & \text{for } \hat x_1 > -1. \end{cases}$$

Since, for $p < \infty$,

$$\lim_{\varepsilon\to 0} \|\hat\partial_2 \hat u_\varepsilon\|_{0,p,\widehat K} = 0 \quad \text{and} \quad \lim_{\varepsilon\to 0} \|\hat\partial_2(\hat u_\varepsilon - \widehat I^\nu \hat u_\varepsilon)\|_{0,q,\widehat K} = \|\hat\partial_2 \widehat I^\nu \hat u_0\|_{0,q,\widehat K} = C > 0,$$

the estimate (5.12) *cannot hold.*

The last equality is proved by contradiction: If $\|\hat\partial_2 \widehat I^\nu \hat u_0\|_{0,q,\widehat K} = 0$, then $I^\nu \hat u_0 =: \hat v(\hat x_1)$ is polynomial in $\hat x_1$, and $(\hat u_0 - \widehat I^\nu \hat u_0)(-1, \hat x_2) = \hat x_2 - \hat v(-1)$ is a non-vanishing first order polynomial in $\hat x_2$. On the other hand, we obtain from the definition of $\widehat I^\nu$ that this polynomial is orthogonal to all polynomials of order $r-1$. This is not possible for a non-vanishing first order polynomial if $r \geq 2$.

In order to prove an anisotropic interpolation error estimate for non-smooth functions $u \in W^{1,p}(K)$ we have to deal with the difficulty that $\widehat F_{-,i}$ and $\widehat F_{+,i}$, $i = 1, \ldots, r-1$, are in general not defined for functions $\hat v \in L^p(\widehat K)$. The back door is to consider functions from $W^{1,p}_E(K)$ which vanish on $E_{-,*}$ and $E_{+,*}$, see (4.1).

LEMMA 5.6. *Let $u \in W^{1,p}_E(K)$ defined in* (4.1)*, $p, q \in [1, \infty]$, $q \leq p$. Then, the estimate*

$$\|\partial_2(u - I_h^\nu u)\|_{0,q,K} \lesssim |K|^{1/q-1/p}|\partial_2 u|_{0,p,K}, \quad \nu = 3, 4,$$

holds true.

Proof. We start with an estimate on the reference element $\widehat K$. With respect to (5.3) and (5.4) we have the following equivalence of norms on the space $\hat\partial_2 \widehat V_r^\nu$, $\nu = 3, 4$,

$$\begin{aligned}\|\hat\partial_2 \widehat I^\nu \hat u\|_{0,q,\widehat K} \lesssim &\sum_{0\leq i+j\leq r-2} |\widehat F_{ij}(\hat\partial_2 \widehat I^\nu \hat u)| + \sum_{i=0}^{r-1} |\widehat F_{i,*}(\hat\partial_2 \widehat I^\nu \hat u)| \\ &+ \sum_{i=1}^{r-1} |\widehat F_{+,i}(\hat\partial_2 \widehat I^\nu \hat u)| + \sum_{i=1}^{r-1} |\widehat F_{+,i}(\hat\partial_2 \widehat I^\nu \hat u)|.\end{aligned}$$

For the functionals $\widehat{F}_{ij}$ and $\widehat{F}_{i,*}$, we use their property (5.2) and their boundedness for functions from $L^1(\widehat{K})$,

$$\begin{aligned}|\widehat{F}_{ij}(\hat{\partial}_2\widehat{I}^\nu\hat{u})| =&|\widehat{F}_{ij}(\hat{\partial}_2\hat{u})| \lesssim \|\hat{\partial}_2\hat{u}\|_{0,p,\widehat{K}}, \quad 0 \le i+j \le r-2,\\ |\widehat{F}_{i,*}(\hat{\partial}_2\widehat{I}^\nu\hat{u})| =&|\widehat{F}_{i,*}(\hat{\partial}_2\hat{u})| \lesssim \|\hat{\partial}_2\hat{u}\|_{0,p,\widehat{K}}, \quad i = 0,\dots,r-1.\end{aligned}$$

Moreover, we integrate by parts for the functionals $\widehat{F}_{\pm,i}$ and employ $\widehat{L}_i(\pm1) = 0$ for $i \ge 1$,

$$\begin{aligned}\left|\widehat{F}_{\pm,i}(\hat{\partial}_2\widehat{I}^\nu\hat{u})\right| =&\frac{2}{\delta_i}\left|\int_{-1}^{1}\widehat{L}_i(\hat{x}_2)(\hat{\partial}_2\widehat{I}^\nu\hat{u})(\pm1,\hat{x}_2)\,d\hat{x}_2\right| = \frac{2}{\delta_i}\left|\int_{-1}^{1}L_i(\hat{x}_2)\widehat{I}^\nu\hat{u}(\pm1,\hat{x}_2)\,d\hat{x}_2\right|\\ =&\left|\widehat{N}_{\pm,i}(\widehat{I}^\nu\hat{u})\right| = \left|\widehat{N}_{\pm,i}(\hat{u})\right| = 0.\end{aligned}$$

Therefore, we can conclude

$$\begin{aligned}\|\hat{\partial}_2\widehat{I}^\nu\hat{u}\|_{0,q,\widehat{K}} \lesssim&\|\hat{\partial}_2\hat{u}\|_{0,p,\widehat{K}},\\ \|\hat{\partial}_2(\hat{u}-\widehat{I}^\nu\hat{u})\|_{0,q,\widehat{K}} \lesssim&\|\hat{\partial}_2\hat{u}\|_{0,p,\widehat{K}}\end{aligned}$$

where we have used $\|\cdot\|_{0,q,\widehat{K}} \lesssim \|\cdot\|_{0,p,\widehat{K}}$ for $q \le p$. The assertion is proved by transforming this estimate from $\widehat{K}$ to K. □

Lemmata 5.1 and 5.4 yield Theorem 4.1, whereas Lemmata 5.1 and 5.6 give Theorem 4.2.

6. Checking the inf-sup condition. For showing that the inf-sup condition holds true, we will use the macro-element technique introduced in [8], see also [14, Sect. II.1.4]. As already described in Sect. 4, the domain Ω is decomposed into isotropic rectangular macro-elements Q such that

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

On the macro-elements Q from this macro-triangulation $\mathcal{T}_H$, we define subspaces of $\boldsymbol{X}_h$ and Q_h in the following way:

$$\begin{aligned}\boldsymbol{X}_h(Q) :=& \left\{\boldsymbol{v} \,:\, \boldsymbol{v}\in\boldsymbol{X}_h,\, \boldsymbol{v}\equiv\boldsymbol{0} \text{ in } \Omega\setminus\overline{Q}\right\},\\ M_h(Q) :=& \left\{q|_Q : q\in M_h\right\}\cap L^2_0(Q).\end{aligned}$$

Furthermore, let

$$\overline{M}_h := \left\{q\in L^2_0(\Omega) \,:\, q|_Q = \text{const.}\ \forall Q\in\mathcal{T}_H\right\}$$

be the space of piecewise constant functions with respect to the macro-triangulation $\mathcal{T}_H$. Now, we can give the uniform local inf-sup condition

$$\exists\beta_\ell > 0\ \forall Q\in\mathcal{T}_H \,:\, \inf_{q_h\in M_h(Q)}\sup_{\boldsymbol{v}_h\in\boldsymbol{X}_h(Q)}\frac{b_h(\boldsymbol{v}_h,q_h)}{|\boldsymbol{v}_h|_{1,h}\,\|q_h\|_{0,Q}} \ge \beta_\ell$$

Adapting the results from [8] to the non-conforming case, we can state the following lemma.

LEMMA 6.1. *Let the local inf-sup condition be fulfilled with a constant β_ℓ which is independent of the number of macro-elements and the mesh size h. Furthermore, we assume that there exists a subspace $\overline{\boldsymbol{X}}_h$ of $\boldsymbol{X}_h$ such that $(\overline{\boldsymbol{X}}_h, \overline{M}_h)$ satisfies an inf-sup condition with a constant $\overline{\beta} > 0$ which does not depend on h. Then, there exists a constant $\beta > 0$ independent of h such that $(\boldsymbol{X}_h, M_h)$ is inf-sup stable with the constant β.*

Proof. For conforming finite elements, the proof can be found in [8] and [14, Sect. II.1.4]. For the adaption of the idea to non-conforming finite elements, we refer to [11, 18, 19]. ∎

A careful inspection of the macro-element techniques shows that the global inf-sup constant is independent of all mesh parameters if the constants β_ℓ and $\overline{\beta}$ are independent of the mesh parameters.

For our purposes, we will use for $\overline{\boldsymbol{X}}_h$ the space of continuous functions which are piecewise biquadratic in $\mathcal{T}_H$. Note that this space is a subspace of $\boldsymbol{X}_h$ for $r \geq 3$ and forms together with $\overline{M}_h$ an inf-sup stable pair. Hence, is remains to show that the local inf-sup condition will be satisfied on boundary layer and corner patches. Let start with boundary layer patches.

LEMMA 6.2. *Let Q be a boundary layer patch. Then, there exists a constant $\beta > 0$ independent of the aspect ratio of the elements such that*

$$\beta_h := \inf_{q_h \in M_h(Q)} \sup_{\boldsymbol{v}_h \in \boldsymbol{X}_h(Q)} \frac{b_h(\boldsymbol{v}_h, q_h)}{|\boldsymbol{v}|_{1,Q} \, \|q_h\|_{0,Q}} > \beta \tag{6.1}$$

holds true provided $r \geq 2$.

Proof. Due to Fortin [13], the fulfilness of the inf-sup condition on a macro-element Q is equivalent to the existence of an interpolation operator $\boldsymbol{i}_Q : \boldsymbol{H}^1_0(Q) \to \boldsymbol{X}_h$ which has the following properties:

$$b_h(\boldsymbol{i}_Q \boldsymbol{v}, q_h) = b_h(\boldsymbol{v}, q_h) \qquad \forall q_h \in M_h(Q)\, \boldsymbol{v} \in \boldsymbol{H}^1_0(Q), \tag{6.2}$$

$$|\boldsymbol{i}_q \boldsymbol{v}|_{1,h} \lesssim |\boldsymbol{v}|_1 \qquad \forall \boldsymbol{v} \in \boldsymbol{H}^1_0(Q). \tag{6.3}$$

In the considered situation, we can choose the natural interpolation operator $\boldsymbol{I}^\nu$, $\nu = 3, 4$, for $\boldsymbol{i}_Q$. The condition (6.3) follows directly from the interpolation error estimate given in Thm. 4.2 for the case $m = 1$, $p = 2$. The condition (6.2) is a consequence of the natural interpolation operator which is defined via the nodal functionals from $\widehat{\mathcal{N}}_r$. Indeed, we have

$$\begin{aligned} b_h(\boldsymbol{I}^\nu \boldsymbol{v}, q_h) &= \sum_{K \subset Q} (\operatorname{div} \boldsymbol{I}^\nu \boldsymbol{v}, q_h)_K \\ &= \sum_{K \subset Q} \big[- (\boldsymbol{I}^\nu \boldsymbol{v}, \nabla q_h)_K + \langle \boldsymbol{n}^K \cdot \boldsymbol{I}^\nu \boldsymbol{v}, q_h \rangle_{\partial K} \big] \\ &= \sum_{K \subset Q} \big[- (\boldsymbol{v}, \nabla q_h)_K + \langle \boldsymbol{n}^K \cdot \boldsymbol{v}, q_h \rangle_{\partial K} \big] \\ &= \sum_{K \subset Q} (\operatorname{div} \boldsymbol{v}, q_h)_K = b_h(\boldsymbol{v}, q_h) \end{aligned}$$

where we have used that $\nabla q_h|_K \in \boldsymbol{P}_{r-2}(K)$ and $q_h|_E \in P_{r-1}(E)$ for all edges $E \in \partial K$. Since both conditions (6.2) and (6.3) are fulfilled and the constant in (6.3) is independent of all mesh parameters, we conclude that the local inf-sup condition is satisfied with a constant which does not depend on the mesh parameters. ∎

LEMMA 6.3. *Let Q be a corner patch. Then there exists a constant $\beta > 0$ independent of the aspect ratio of the elements such that* (6.1) *holds true provided $r \geq 3$.*

Proof. The proof follows the lines of the proof [24, Cor. 4.13], compare also [3, Lem. 5]. The proof uses also a macroelement technique where, as above, the space inclusion is valid for $r \geq 3$ only. □

Numerical tests for the case $r = 2$ have shown that all four families are stable on corner patches with $\sigma = \frac{1}{2}$. Nevertheless, our method of proof does not work for $r = 2$. Therefore the validity of a uniform inf-sup condition must be considered open for this case.

Now, we can finish this section with the

Proof. [of Thm. 4.3] Since we have shown all necessary ingredients for the macroelement technique, the stated inf-sup condition holds true with a constant which is independent of the mesh parameter, aspect ratio and mesh size. □

7. Proof of the consistency error estimate. For proving the estimate of the consistency error, we start with an estimate for an auxiliary problem.

LEMMA 7.1. *For arbitrary $\boldsymbol{\eta} \in \boldsymbol{H}^r(\Omega)$ and $g = -\operatorname{div} \boldsymbol{\eta}$ the estimate*

$$\sup_{v_h \in V_h} \frac{(\boldsymbol{\eta}, \nabla v_h)_h - (g, v_h)}{|v_h|_{1,h}} \lesssim \left(\sum_{K \in \mathcal{T}_h} \sum_{|\alpha|=r} h_K^{2\alpha} \|D^\alpha \boldsymbol{\eta}\|_{0,K}^2 \right)^{1/2} + \left(\sum_{K \in \mathcal{T}_h} h_{K,1}^2 \sum_{|\alpha|=r-1} h_K^{2\alpha} \|D^\alpha g\|_{0,K}^2 \right)^{1/2}$$

holds true where

$$(\boldsymbol{\eta}, \nabla v_h)_h := \sum_{K \in \mathcal{T}_h} (\boldsymbol{\eta}, \nabla v_h)_K.$$

Proof. By using $g = -\operatorname{div} \boldsymbol{\eta}$ and integration by parts, we can rewrite the numerator as

$$(\boldsymbol{\eta}, \nabla v_h)_h - (g, v_h) = \sum_{K \in \mathcal{T}_h} \langle \boldsymbol{\eta} \cdot \boldsymbol{n}_K, v_h \rangle_{\partial K} = T_1 + T_2$$

with

$$T_1 = \sum_{K \in \mathcal{T}_h} \sum_{E \in \mathcal{E}_S(K)} \int_E \boldsymbol{\eta} \cdot \boldsymbol{n}\, v_h \, d\gamma, \qquad T_2 = \sum_{K \in \mathcal{T}_h} \sum_{E \in \mathcal{E}_L(K)} \int_E \boldsymbol{\eta} \cdot \boldsymbol{n}\, v_h \, d\gamma,$$

where $\mathcal{E}_L(K)$ and $\mathcal{E}_S(K)$ denote the sets of long and short edges of K, respectively. Let us further assume that the long edges are parallel to the x_1-direction while the short edges are parallel to the x_2-direction. This gives that $n_2 = 0$ on short edges and $n_1 = 0$ on long edges. Hence, T_1 and T_2 simply to

$$T_1 = \sum_{E \in \mathcal{E}_S} \int_E \eta_1 n_1^E [v_h]_E \, d\gamma, \qquad T_2 = \sum_{E \in \mathcal{E}_L} \int_E \eta_2 n_2^E [v_h]_E \, d\gamma.$$

Here, $\mathcal{E}_L$ and $\mathcal{E}_S$ are the global sets of long and short edges, respectively. Let Π denote the special interpolation operator into the space of continuous Q_{r-1} elements

defined in [14, p. 108]. Due to the definition of V_h, the jump $[v_h]_E$ is orthogonal to all polynomials of degree $r-1$. Hence, we obtain

$$T_1 = \sum_{E\in\mathcal{E}_S} \int_E (\eta_1 - \Pi\eta_1) n_1^E [v_h]_E \, d\gamma, \qquad T_2 = \sum_{E\in\mathcal{E}_L} \int_E (\eta_2 - \Pi\eta_2) n_2^E [v_h]_E \, d\gamma$$

since the normal vectors are constant on the edges. Using the Gaussian theorem and the product rule, we get

$$\begin{aligned}
T_1 &= \sum_{K\in\mathcal{T}_h} \int_K \partial_1 \big[(\eta_1 - \Pi\eta_1) v_h\big] \, dx \\
&= \sum_{K\in\mathcal{T}_h} \int_K \partial_1(\eta_1 - \Pi\eta_1) v_h \, dx + \sum_{K\in\mathcal{T}_h} \int_K (\eta_1 - \Pi\eta_1) \partial_1 v_h \, dx, \\
T_2 &= \sum_{K\in\mathcal{T}_h} \int_K \partial_2 \big[(\eta_2 - \Pi\eta_2) v_h\big] \, dx \\
&= \sum_{K\in\mathcal{T}_h} \int_K \partial_2(\eta_2 - \Pi\eta_2) v_h \, dx + \sum_{K\in\mathcal{T}_h} \int_K (\eta_2 - \Pi\eta_2) \partial_2 v_h \, dx.
\end{aligned}$$

Due to the choice of the special interpolation operator Π, we have

$$\int_K \partial_i(\eta_i - \Pi\eta_i) \, dx = 0.$$

Hence, we get

$$\begin{aligned}
T_1 &= \sum_{K\in\mathcal{T}_h} \int_K \partial_1(\eta_1 - \Pi\eta_1)(v_h - v_h^K) \, dx + \sum_{K\in\mathcal{T}_h} \int_K (\eta_1 - \Pi\eta_1) \partial_1 v_h \, dx, \\
T_2 &= \sum_{K\in\mathcal{T}_h} \int_K \partial_2(\eta_2 - \Pi\eta_2)(v_h - v_h^K) \, dx + \sum_{K\in\mathcal{T}_h} \int_K (\eta_2 - \Pi\eta_2) \partial_2 v_h \, dx.
\end{aligned}$$

where v_h^K is the integral mean of v_h on the cell K. The interpolation operator Π fulfils

$$\|\partial_i(w - \Pi w)\|_{0,K} \lesssim \sum_{|\alpha|=r-1} h_K^\alpha \|D^\alpha \partial_i w\|_{0,K}, \qquad \|w - \Pi w\|_{0,K} \lesssim \sum_{|\alpha|=r} h_K^\alpha \|D^\alpha w\|_{0,K}$$

provided that $w \in H^r(K)$, see [2, 25]. We end up with

$$\begin{aligned}
|T_1| \lesssim \Bigg[&\left(\sum_{K\in\mathcal{T}_h} h_{K,1}^2 \sum_{|\alpha|=r-1} h_K^{2\alpha} \|D^\alpha \partial_1 \eta_1\|_{0,K}^2 \right)^{1/2} \\
&+ \left(\sum_{K\in\mathcal{T}_h} \sum_{|\alpha|=r} h_K^{2\alpha} \|D^\alpha \eta_1\|_{0,K}^2 \right)^{1/2} \Bigg] |v_h|_{1,h},
\end{aligned}$$

$$|T_2| \lesssim \left[\left(\sum_{K\in\mathcal{T}_h} h_{K,1}^2 \sum_{|\alpha|=r-1} h_K^{2\alpha} \|D^\alpha \partial_2 \eta_2\|_{0,K}^2 \right)^{1/2} + \left(\sum_{K\in\mathcal{T}_h} \sum_{|\alpha|=r} h_K^{2\alpha} \|D^\alpha \eta_2\|_{0,K}^2 \right)^{1/2} \right] |v_h|_{1,h},$$

where the Cauchy-Schwarz inequality was used. The desired estimate is obtained by using $\partial_2\eta_2 = -g - \partial_1\eta_1$. □

We finish this section with the

Proof. [of Thm. 4.4] The proof is accomplished by applying Lemma 7.1 to

$$\boldsymbol{\eta}^{(1)} = \begin{pmatrix} \partial_1 u_1 \\ \partial_2 u_1 \end{pmatrix}, \quad g^{(1)} = f_1 - \partial_1 p, \quad \text{and} \quad \boldsymbol{\eta}^{(2)} = \begin{pmatrix} \partial_1 u_2 \\ \partial_2 u_2 - p \end{pmatrix}, \quad g^{(2)} = f_2.$$

This choice leads to

$$\begin{aligned} &|a_h(\boldsymbol{u},\boldsymbol{v}_h) - b_h(\boldsymbol{v}_h,p) - (\boldsymbol{f},\boldsymbol{v}_h)| \\ &\quad = \big|(\boldsymbol{\eta}^{(1)}, \nabla v_{h,1})_h - (g^{(1)}, v_{h,1}) + (\boldsymbol{\eta}^{(2)}, \nabla v_{h,2})_h - (g^{(2)}, v_{h,2})\big| \\ &\quad \lesssim \left(\sum_{K\in\mathcal{T}_h} \sum_{|\alpha|=r} h_K^{2\alpha} \big(|D^\alpha \boldsymbol{u}|_{1,K}^2 + \|D^\alpha p\|_{0,K}^2 \big) \right)^{1/2} |\boldsymbol{v}_h|_{1,h} \\ &\qquad + \left(\sum_{K\in\mathcal{T}_h} h_{K,1}^2 \sum_{|\alpha|=r-1} h_K^{2\alpha} \big(\|D^\alpha \boldsymbol{f}\|_{0,K}^2 + \|D^\alpha \partial_1 p\|_{0,K}^2 \big) \right)^{1/2} |\boldsymbol{v}_h|_{1,h} \end{aligned}$$

The pressure terms in the last row are part of those in the second last row. This finishes the proof. □

Appendix A. Unisolvence. This appendix contains the proof of the unisolvence results for $\widehat{V}_r^4$.

LEMMA A.1. *The set $\widehat{\mathcal{N}}_r$ is unisolvent on $\widehat{V}_r^4$.*

Proof. We have to show that $\hat{v} \in \widehat{V}_r^4$ and $\widehat{N}(\hat{v}) = 0$ for all $\widehat{N} \in \widehat{\mathcal{N}}_r$ results in $\hat{v} \equiv 0$. To this end, we represent an arbitrary function $\hat{v} \in \widehat{V}_r^4$ in the form

$$\hat{v} = \sum_{0\le i+j\le r} \alpha_{ij} L_i(\hat{x}_1) L_j(\hat{x}_2) + \sum_{i=0}^{r-1} \beta_i L_{r+1-i}(\hat{x}_1) L_i(\hat{x}_2) + \sum_{i=1}^{r-1} \gamma_i L_{r+2-i}(\hat{x}_1) L_i(\hat{x}_2). \tag{A.1}$$

The orthogonality of the Legendre polynomials gives

$$0 = \widehat{N}_{mn}(\hat{v}) = \alpha_{mn}, \qquad 0 \le m+n \le r-2.$$

Hence, we have

$$\begin{aligned} \hat{v} = &\sum_{i=0}^{r-1} \alpha_{r-1-i,i} L_{r-1-i}(\hat{x}_1) L_i(\hat{x}_2) + \sum_{i=0}^{r} \alpha_{r-i,i} L_{r-i}(\hat{x}_1) L_i(\hat{x}_2) \\ &+ \sum_{i=0}^{r-1} \beta_i L_{r+1-i}(\hat{x}_1) L_i(\hat{x}_2) + \sum_{i=1}^{r-1} \gamma_i L_{r+2-i}(\hat{x}_1) L_i(\hat{x}_2). \end{aligned}$$

The application of $\widehat{N}_{-,n}$ and $\widehat{N}_{+,n}$, $0 \le n \le r-1$, results in

$$0 = \alpha_{r-1-n,n} + \alpha_{r-n,n} + \beta_n + \gamma_n, \qquad n = 0, \dots, r-1, \tag{A.2}$$
$$0 = \alpha_{r-1-n,n} - \alpha_{r-n,n} + \beta_n - \gamma_n, \qquad n = 0, \dots, r-1, \tag{A.3}$$

where again the orthogonality of the Legendre polynomials is used. The parameter $\gamma_0 = 0$ is introduced for the convenience of notation. The sum and the difference of (A.2) and (A.3) give

$$0 = \alpha_{r-1-n,n} + \beta_n, \qquad n = 0, \dots, r-1, \tag{A.4}$$
$$0 = \alpha_{r-n,n} + \gamma_n, \qquad n = 0, \dots, r-1, \tag{A.5}$$

respectively. If we apply $\widehat{N}_{m,-}$ and $\widehat{N}_{m,+}$, $m = 0, \dots, r-1$, then we get

$$0 = \alpha_{m,r-1-m} - \alpha_{m,r-m} + \beta_{r+1-m} - \gamma_{r+2-m}, \qquad m = 0, \dots, r-1, \tag{A.6}$$
$$0 = \alpha_{m,r-1-m} + \alpha_{m,r-m} + \beta_{r+1-m} + \gamma_{r+2-m}, \qquad m = 0, \dots, r-1, \tag{A.7}$$

where we set $\beta_r = \beta_{r+1} = 0$ and $\gamma_r = \gamma_{r+1} = \gamma_{r+2} = 0$. Taking the sum and the difference of (A.6) and (A.7), we obtain

$$0 = \alpha_{m,r-1-m} + \beta_{r+1-m}, \qquad m = 0, \dots, r-1, \tag{A.8}$$
$$0 = \alpha_{m,r-m} + \gamma_{r+2-m}, \qquad m = 0, \dots, r-1, \tag{A.9}$$

respectively. Using (A.4) for n and (A.8) for $m = r-1-n$, we obtain

$$\alpha_{r-1-n,n} + \beta_n = 0 = \alpha_{r-1-n,n} + \beta_{n+2}, \qquad n = 0, \dots, r-1.$$

Hence, we get $\beta_n = \beta_{n+2}, n = 0, \dots, r-1$. With $\beta_r = \beta_{r+1} = 0$ we conclude

$$\beta_n = 0, \qquad n = 0, \dots, r-1,$$

and with (A.4)

$$\alpha_{r-1-n,n} = 0, \qquad n = 0, \dots, r-1.$$

For $n = 1, \dots, r-1$ we obtain from (A.5) for n and from (A.9) for $m = r-n$

$$\alpha_{r-n,n} + \gamma_n = 0 = \alpha_{r-n,n} + \gamma_{n+2}, \qquad n = 1, \dots, r-1.$$

Hence, we have $\gamma_n = \gamma_{n+2}, n = 1, \dots, r-1$. Together with $\gamma_r = \gamma_{r+1} = 0$ we end up with

$$\gamma_n = 0, \qquad n = 1, \dots, r-1.$$

Putting this into (A.5) we get

$$\alpha_{r-n,n} = 0, \qquad n = 1, \dots, r-1.$$

It remains to consider $\alpha_{r,0}$ and $\alpha_{0,r}$. From (A.5) for $n = 0$ we get $\alpha_{r,0} = 0$ since $\gamma_0 = 0$. Setting $m = 0$ in (A.9) results in $\alpha_{0,r} = 0$ due to $\gamma_{r+2} = 0$. Finally, we have seen that all coefficients in the representation (A.1) vanish. Hence, $\hat{v} \equiv 0$ and the set of nodal functionals $\widehat{\mathcal{N}}_r$ is unisolvent on $\widehat{V}_r^4$. □

REFERENCES

[1] T. Apel, *Anisotropic finite elements: Local estimates and applications*, Advances in Numerical Mathematics, Teubner, Stuttgart, 1999.

[2] ———, *Anisotropic interpolation error estimates for a Q_r-element defined via moments*, in preparation, 2006.

[3] T. Apel and S. Nicaise, *The inf-sup condition for some low order elements on anisotropic meshes*, Calcolo, 41 (2004), pp. 89–113.

[4] T. Apel, S. Nicaise, and J. Schöberl, *Crouzeix-Raviart type finite elements on anisotropic meshes*, Numer. Math., 89 (2001), pp. 193–223.

[5] ———, *A non-conforming finite element method with anisotropic mesh grading for the Stokes problem in domains with edges*, IMA J. Numer. Anal., 21 (2001), pp. 843–856.

[6] T. Apel and H. M. Randrianarivony, *Stability of discretizations of the Stokes problem on anisotropic meshes*, Math. and Comput. Simulation, 61 (2003), pp. 437–447.

[7] R. Becker and R. Rannacher, *Finite element solution of the incompressible Navier-Stokes equations on anisotropically refined meshes*, in Fast solvers for flow problems (Kiel, 1994), vol. 49 of Notes Numer. Fluid Mech., Vieweg, Braunschweig, 1995, pp. 52–62.

[8] J. Boland and R. Nicolaides, *Stability of finite elements under divergence constraints*, SIAM J. Numer. Anal., 20 (1983), pp. 722–731.

[9] F. Brezzi and M. Fortin, *Mixed and Hybrid Finite Element Methods*, vol. 15 of Springer Series in Computational Mathematics, Springer-Verlag, 1991.

[10] P. G. Ciarlet, *The finite element method for elliptic problems*, North-Holland Publishing Company, Amsterdam - New York - Oxford, 1978.

[11] M. Crouzeix and R. S. Falk, *Nonconforming finite elements for the Stokes problem*, Math. Comp., 52 (1989), pp. 437–456.

[12] M. Crouzeix and P. A. Raviart, *Conforming and non-conforming finite elements for solving the stationary Stokes equations*, R.A.I.R.O. Anal. Numér., 7 (1973), pp. 33–76.

[13] M. Fortin, *An analysis of the convergence of mixed finite element methods*, RAIRO Anal. Numér., 11 (1977), pp. 341–354.

[14] V. Girault and P.-A. Raviart, *Finite Element Methods for Navier–Stokes equations*, Springer-Verlag, Berlin-Heidelberg-New York, 1986.

[15] J. P. Hennart, J. Jaffré, and J. E. Roberts, *A constructive method for deriving finite elements of nodal type*, Numer. Math., 53 (1988), pp. 701–738.

[16] V. John, *Higher order finite element methods and multigrid solvers in a benchmark problem for the 3d navier-stokes equations*, Int. J. Num. Meth. Fluids, 40 (2002), pp. 775–798.

[17] V. John and G. Matthies, *Higher order finite element discretizations in a benchmark problem for incompressible flows*, Int. J. Num. Meth. Fluids, 37 (2001), pp. 885–903.

[18] P. Knobloch, *On the application of the P_1^{mod} element to incompressible flow problems*, Comput. Visual. Sci., 6 (2004), pp. 185–195.

[19] ———, *On the inf-sup condition for the $P_3^{\mathrm{mod}}/P_2^{\mathrm{disc}}$ element*, Computing, 76 (2006), pp. 41–54.

[20] G. Matthies, *Inf-sup stable nonconforming finite elements of higher order on quadrilaterals and hexahedra*, Bericht 373, Fakultät für Mathematik, Ruhr-Universität Bochum, 2005.

[21] G. Matthies and L. Tobiska, *Inf-sup stable non-conforming finite elements of arbitrary order on triangles*, Numer. Math., 102 (2005), pp. 293–309.

[22] J. Maubach and P. Rabier, *Nonconforming finite elements of arbitrary degree over triangles*, RANA report 0328, Technical University of Eindhoven, 2003.

[23] R. Rannacher and S. Turek, *Simple nonconforming quadrilateral Stokes element*, Numer. Methods Partial Differential Equations, 8 (1992), pp. 97–111.

[24] D. Schötzau, C. Schwab, and R. Stenberg, *Mixed hp-FEM on anisotropic meshes. II. Hanging nodes and tensor products of boundary layer meshes*, Numer. Math., 83 (1999), pp. 667–697.

[25] M. Stynes and L. Tobiska, *Using rectangular Q_p elements in the SDFEM for a convection-diffusion problem with a boundary layer*, Preprint 06-08, Fakultät für Mathematik, Otto-von-Guericke-Universität Magdeburg, 2006.