Japan Journal of Industrial and Applied Mathematics (2023) 40:475–512
https://doi.org/10.1007/s13160-022-00535-w

ORIGINAL PAPER

Anisotropic interpolation error estimates using a new geometric parameter

Hiroki Ishizaka[1] · Kenta Kobayashi[2] · Takuya Tsuchiya[1]

Received: 7 June 2021 / Revised: 2 August 2022 / Accepted: 11 August 2022 /
Published online: 25 August 2022

Abstract
We present precise anisotropic interpolation error estimates for smooth functions using a new geometric parameter and derive inverse inequalities on anisotropic meshes. In our theory, the interpolation error is bounded in terms of the diameter of a simplex and the geometric parameter. Imposing additional assumptions makes it possible to obtain anisotropic error estimates. This paper also includes corrections to an error in Theorem 2 of our previous paper, "General theory of interpolation error estimates on anisotropic meshes" (Japan Journal of Industrial and Applied Mathematics, 38 (2021) 163–191).

Keywords Finite element · Interpolation error estimates · Anisotropic meshes

Mathematics Subject Classification 65D05 · 65N30

1 Introduction

Analyzing the errors of interpolations on d-simplices is an important subject in numerical analysis. It is particularly crucial for finite element error analysis. Let us briefly outline the problems considered in this paper using the Lagrange interpolation operator.

Let $d \in \{1, 2, 3\}$. Let $\widehat{T} \subset \mathbb{R}^d$ and $T_0 \subset \mathbb{R}^d$ be a reference element and a simplex, respectively, that are affine equivalent. Let us consider two Lagrange finite

✉ Hiroki Ishizaka
h.ishizaka005@gmail.com

Kenta Kobayashi
kenta.k@r.hit-u.ac.jp

Takuya Tsuchiya
tsuchiya@math.sci.ehime-u.ac.jp

1 Graduate School of Science and Engineering, Ehime University, Matsuyama, Japan

2 Graduate School of Business Administration, Hitotsubashi University, Kunitachi, Japan

elements $\{\widehat{T}, \widehat{P} := \mathcal{P}^k, \widehat{\Sigma}\}$ and $\{T_0, P := \mathcal{P}^k, \Sigma\}$ with associated normed vector spaces $V(\widehat{T}) := \mathcal{C}(\widehat{T})$ and $V(T_0) := \mathcal{C}(T_0)$ with $k \in \mathbb{N}$, where $\mathcal{P}^m$ is the space of polynomials with degree at most $m \in \mathbb{N}_0 := \mathbb{N} \cup \{0\}$. For $\hat{\varphi} \in V(\widehat{T})$, we use the correspondences

$$(\varphi_0 : T_0 \to \mathbb{R}) \to (\hat{\varphi} := \varphi_0 \circ \boldsymbol{\Phi} : \widehat{T} \to \mathbb{R}),$$

where $\boldsymbol{\Phi}$ is an affine mapping. Let $I_{\widehat{T}}^k : V(\widehat{T}) \to \mathcal{P}^k$ and $I_{T_0}^k : V(T_0) \to \mathcal{P}^k$ be the corresponding Lagrange interpolation operators. Details can be found in Sect. 2.3.

We first consider the case in which $d = 1$. Let $\Omega := (0, 1) \subset \mathbb{R}$. For $N \in \mathbb{N}$, let $\mathbb{T}_h = \{0 = x_0 < x_1 < \cdots < x_N < x_{N+1} = 1\}$ be a mesh of $\overline{\Omega}$ such as

$$\overline{\Omega} := \bigcup_{i=1}^{N} T_0^i, \quad \mathrm{int} T_0^i \cap \mathrm{int} T_0^j = \emptyset \quad \text{for } i \neq j,$$

where $T_0^i := [x_i, x_{i+1}]$ for $0 \le i \le N$. We denote $h_i := x_{i+1} - x_i$ for $0 \le i \le N$. If we set $x_j := \frac{j}{N+1}$ for $j = 0, 1, \ldots, N, N+1$, the mesh $\mathbb{T}_h$ is said to be the uniform mesh. If we set $x_j := g\left(\frac{j}{N+1}\right)$ for $j = 1, \ldots, N, N+1$ with a grading function g, the mesh $\mathbb{T}_h$ is said to be the graded mesh with respect to $x = 0$; see [5]. In particular, when $g(y) := y^{\varepsilon}$ $(\varepsilon > 0)$, the mesh is called the radical mesh. To obtain the Lagrange interpolation error estimates, we impose standard assumptions and specify that $\ell, m \in \mathbb{N}_0$ and $p, q \in [1, \infty]$ such that

$$0 \le m \le \ell + 1 : \quad W^{\ell+1,p}(\widehat{T}) \hookrightarrow W^{m,q}(\widehat{T}). \tag{1}$$

Under these assumptions, the following holds for any $\varphi_0 \in W^{\ell+1,p}(T_0^i)$ with $\hat{\varphi} = \varphi_0 \circ \boldsymbol{\Phi}$:

$$|\varphi_0 - I_{T_0^i}^k \varphi_0|_{W^{m,q}(T_0^i)} \le c h_i^{\frac{1}{q} - \frac{1}{p} + \ell + 1 - m} |\varphi_0|_{W^{\ell+1,p}(T_0^i)}. \tag{2}$$

The proof of this statement is standard; see [12]. When $p = q$, it is possible to obtain optimal error estimates even if the scale is different for each element. When $q > p$, the order of convergence of the interpolation operator may deteriorate.

We now consider the cases in which $d = 2, 3$. Let $\Omega \subset \mathbb{R}^d$ be a bounded polyhedral domain. Let $\mathbb{T}_h = \{T_0\}$ be a simplicial mesh of $\overline{\Omega}$ made up of closed d-simplices, such as

$$\overline{\Omega} = \bigcup_{T_0 \in \mathbb{T}_h} T_0$$

with $h := \max_{T_0 \in \mathbb{T}_h} h_{T_0}$, where $h_{T_0} := \mathrm{diam}(T_0)$. For simplicity, we assume that $\mathbb{T}_h$ is conformal. That is, $\mathbb{T}_h$ is a simplicial mesh of $\overline{\Omega}$ without hanging nodes. Let $\widehat{T} \subset \mathbb{R}^d$ be the reference element defined in Sect. 2 and $\boldsymbol{\Phi}$ be the affine mapping defined in Eq. (13). For any $T_0 \in \mathbb{T}_h$, it holds that $T_0 = \boldsymbol{\Phi}(\widehat{T})$. Under the standard assumptions and Eq. (1), the following holds for any $\varphi_0 \in W^{\ell+1,p}(T_0)$ with $\hat{\varphi} = \varphi_0 \circ \boldsymbol{\Phi}$:

$$|\varphi_0 - I_{T_0}^k \varphi_0|_{W^{m,q}(T_0)} \leq c|T_0|^{\frac{1}{q}-\frac{1}{p}} \left(\frac{\alpha_{\max}}{\alpha_{\min}}\right)^m \left(\frac{H_{T_0}}{h_{T_0}}\right)^m h_{T_0}^{\ell+1-m} |\varphi_0|_{W^{\ell+1,p}(T_0)}, \tag{3}$$

where $|T_0|$ is the measure of T_0, the parameters $\alpha_{\max}$ and $\alpha_{\min}$ are defined in Eq. (50), and the parameter H_{T_0} is as proposed in a recent paper [16]; see Sect. 2.4 for a definition. The proof of estimate (3) can be found in Sect. 5. Compared with the one-dimensional case, the quantities $\alpha_{\max}/\alpha_{\min}$ and H_{T_0}/h_{T_0} negatively affect the order of convergence and do not appear in Eq.(2). The two quantities $\alpha_{\max}/\alpha_{\min}$ and H_{T_0}/h_{T_0} are considered in Sect. 7.1. As a mesh condition, the *shape-regularity condition* is widely used and well known. This condition states that there exists a constant $\gamma > 0$ such that

$$\rho_{T_0} \geq \gamma h_{T_0} \quad \forall \mathbb{T}_h \in \{\mathbb{T}_h\}, \quad \forall T_0 \in \mathbb{T}_h, \tag{4}$$

where ρ_{T_0} is the radius of the inscribed ball of T_0. Under this condition, it holds that

$$|\varphi_0 - I_{T_0}^k \varphi_0|_{W^{m,q}(T_0)} \leq c|T_0|^{\frac{1}{q}-\frac{1}{p}} h_{T_0}^{\ell+1-m} |\varphi_0|_{W^{\ell+1,p}(T_0)}; \tag{5}$$

see Sect. 7.1.1. If condition (4) is violated (i.e., the simplex becomes too flat as $h_{T_0} \to 0$), the quantity

$$\left(\frac{\alpha_{\max}}{\alpha_{\min}}\right)^m \left(\frac{H_{T_0}}{h_{T_0}}\right)^m h_{T_0}^{\ell+1-m}$$

may diverge even when $p = q$. The effect of the quantity $|T_0|^{\frac{1}{q}-\frac{1}{p}}$ on the interpolation error estimates is considered in Sect. 7.2.

In some cases, it is not necessary for condition (4) to hold to obtain Eq. (5). The shape-regularity condition can be relaxed to the *maximum-angle condition*, as stated in Eqs. (20) and (21), for both two-dimensional [4] and three-dimensional cases [20]. Anisotropic interpolation theory has also been developed [1, 2, 8]. The idea of Apel *et al.* is to construct a set of functionals satisfying conditions (54), (55), and (56). The introduction of these functionals makes it possible to remove the quantity $\alpha_{\max}/\alpha_{\min}$. Under the conditions of the maximum angle and coordinate system, anisotropic interpolation error estimates can then be deduced (e.g., see [1]).

In contrast, this paper proposes anisotropic interpolation error estimates using the new parameter under conditions (54), (55), and (56) and Assumption 1; i.e., we derive the following anisotropic error estimate (Theorem B, in particular, Corollary 1):

$$\begin{aligned} &|\varphi_0 - I_{T_0}^k \varphi_0|_{W^{m,q}(T_0)} \\ &\quad \leq c|T_0|^{\frac{1}{q}-\frac{1}{p}} \left(\frac{H_{T_0}}{h_{T_0}}\right)^m \sum_{|\gamma|=\ell-m} \mathscr{H}^\gamma |\partial^\gamma (\varphi_0 \circ \Phi_{T_0})|_{W^{m,p}(\Phi_{T_0}^{-1}(T_0))}, \end{aligned} \tag{6}$$

where $\boldsymbol{\Phi}_{T_0}$ is defined in Eq. (12), $\gamma := (\gamma_1, \dots, \gamma_d) \in \mathbb{N}_0^d$ is a multi-index, and $\mathscr{H}$ is specified in Definition 3. Theorem B applies to interpolations other than the Lagrange interpolation, and the basis for the proof of Theorem B is the *scaling argument* described in Sect. 3.

Because the new geometric parameter is used in the interpolation error analysis, the coefficient c used in the error estimation is independent of the geometry of the simplices, and the error estimations obtained may therefore be applied to arbitrary meshes, including very “flat” or anisotropic simplices. Furthermore, we are naturally able to consider the following geometric condition as being sufficient to obtain optimal order estimates (when $p = q$): there exists $\gamma_0 > 0$ such that

$$\frac{H_{T_0}}{h_{T_0}} \leq \gamma_0 \quad \forall \mathbb{T}_h \in \{\mathbb{T}_h\}, \quad \forall T_0 \in \mathbb{T}_h. \tag{7}$$

Condition (7) appears to be simpler than the maximum-angle condition. Furthermore, the quantity H_{T_0}/h_{T_0} can be easily calculated in the numerical process of finite element methods. Therefore, the new condition may be useful. A recent paper [18] showed that the new condition is satisfied if and only if the maximum-angle condition holds. We expect the new mesh condition to become an alternative to the maximum-angle condition.

Furthermore, under Assumption 1, component-wise inverse inequalities can be deduced as (see Sect. 7.3):

$$\|\partial^\gamma \varphi_h\|_{L^q(T)} \leq C^{IVC} |T|^{\frac{1}{q} - \frac{1}{p}} \mathscr{H}^{-\gamma} \|\varphi_h\|_{L^p(T)}.$$

In a previous paper [16], the present authors developed new interpolation error estimations in a general framework and derived Raviart–Thomas interpolations on d-simplices. However, the statement of Theorem 2 in [16] includes a mistake. That is, under standard assumptions, the quantity $\alpha_{\max}/\alpha_{\min}$ cannot be removed. We need to modify the statement of this theorem to correct this error. The current paper presents Theorems A (see Sect. 5) and B (see Sect. 6), which replace Theorem 2 of [16]. In Sect. 4, we explain the inaccuracies in the proof of Theorem 2 in [16] and describe how the results can be recovered using our Theorems A and B. Furthermore, the Babuška and Aziz technique is generally not applicable on anisotropic meshes in the proof of Theorem 3 in [16]. Details will be discussed in a coming paper [14].

When there is no ambiguity, we use the notation and definitions given in [16]. Throughout this paper, c denotes a constant independent of h (defined later), unless specified otherwise. These values may change in each context. $\mathbb{R}_+$ is the set of positive real numbers.

2 Strategy for constructing anisotropic interpolation theory

In standard interpolation theory, one introduces an affine mapping that connects the reference element to the mesh element. However, on anisotropic meshes, the interpolation errors may be overestimated. Therefore, our strategy is to divide the transformation into three affine mappings.

2.1 Standard positions of simplices

We recall [16, Section 3]. Let us first define a diagonal matrix $\widehat{A}^{(d)}$ as

$$\widehat{A}^{(d)} := \mathrm{diag}(\alpha_1, \dots, \alpha_d), \quad \alpha_i \in \mathbb{R}_+ \quad \forall i. \tag{8}$$

2.1.1 Two-dimensional case

Let $\widehat{T} \subset \mathbb{R}^2$ be the reference triangle with vertices $\hat{x}_1 := (0,0)^T$, $\hat{x}_2 := (1,0)^T$, and $\hat{x}_3 := (0,1)^T$.

Let $\widetilde{\mathfrak{T}}^{(2)}$ be the family of triangles

$$\widetilde{T} = \widehat{A}^{(2)}(\widehat{T})$$

with vertices $\tilde{x}_1 := (0,0)^T$, $\tilde{x}_2 := (\alpha_1, 0)^T$, and $\tilde{x}_3 := (0, \alpha_2)^T$.

We next define the regular matrices $\widetilde{A} \in \mathbb{R}^{2\times 2}$ by

$$\widetilde{A} := \begin{pmatrix} 1 & s \\ 0 & t \end{pmatrix}, \tag{9}$$

with parameters

$$s^2 + t^2 = 1, \quad t > 0.$$

For $\widetilde{T} \in \widetilde{\mathfrak{T}}^{(2)}$, let $\mathfrak{T}^{(2)}$ be the family of triangles

$$T = \widetilde{A}(\widetilde{T})$$

with vertices $x_1 := (0,0)^T$, $x_2 := (\alpha_1, 0)^T$, $x_3 := (\alpha_2 s, \alpha_2 t)^T$. We then have that $\alpha_1 = |x_1 - x_2| > 0$ and $\alpha_2 = |x_1 - x_3| > 0$.

2.1.2 Three-dimensional case

Let $\widehat{T}_1$ and $\widehat{T}_2$ be reference tetrahedra with the following vertices:

(i) $\widehat{T}_1$ has the vertices $\hat{x}_1 := (0,0,0)^T$, $\hat{x}_2 := (1,0,0)^T$, $\hat{x}_3 := (0,1,0)^T$, $\hat{x}_4 := (0,0,1)^T$;
(ii) $\widehat{T}_2$ has the vertices $\hat{x}_1 := (0,0,0)^T$, $\hat{x}_2 := (1,0,0)^T$, $\hat{x}_3 := (1,1,0)^T$, $\hat{x}_4 := (0,0,1)^T$.

Let $\widetilde{\mathfrak{T}}_i^{(3)}$, $i = 1, 2$, be the family of triangles

$$\widetilde{T}_i = \widehat{A}^{(3)}(\widehat{T}_i), \quad i = 1, 2,$$

with vertices

(i) $\tilde{x}_1 := (0,0,0)^T$, $\tilde{x}_2 := (\alpha_1, 0, 0)^T$, $\tilde{x}_3 := (0, \alpha_2, 0)^T$, and $\tilde{x}_4 := (0, 0, \alpha_3)^T$;
(ii) $\tilde{x}_1 := (0,0,0)^T$, $\tilde{x}_2 := (\alpha_1, 0, 0)^T$, $\tilde{x}_3 := (\alpha_1, \alpha_2, 0)^T$, and $\tilde{x}_4 := (0, 0, \alpha_3)^T$.

We next define the regular matrices $\widetilde{A}_1, \widetilde{A}_2 \in \mathbb{R}^{3\times 3}$ by

$$\widetilde{A}_1 := \begin{pmatrix} 1 & s_1 & s_{21} \\ 0 & t_1 & s_{22} \\ 0 & 0 & t_2 \end{pmatrix}, \quad \widetilde{A}_2 := \begin{pmatrix} 1 & -s_1 & s_{21} \\ 0 & t_1 & s_{22} \\ 0 & 0 & t_2 \end{pmatrix} \tag{10}$$

with parameters

$$\begin{cases} s_1^2 + t_1^2 = 1, \quad s_1 > 0, \quad t_1 > 0, \quad \alpha_2 s_1 \le \alpha_1/2, \\ s_{21}^2 + s_{22}^2 + t_2^2 = 1, \quad t_2 > 0, \quad \alpha_3 s_{21} \le \alpha_1/2. \end{cases}$$

For $\widetilde{T}_i \in \widetilde{\mathfrak{T}}_i^{(3)}$, $i = 1, 2$, let $\mathfrak{T}_i^{(3)}$, $i = 1, 2$, be the family of tetrahedra

$$T_i = \widetilde{A}_i(\widetilde{T}_i), \quad i = 1, 2,$$

with vertices

$$x_1 := (0,0,0)^T, \; x_2 := (\alpha_1, 0, 0)^T, \; x_4 := (\alpha_3 s_{21}, \alpha_3 s_{22}, \alpha_3 t_2)^T,$$
$$\begin{cases} x_3 := (\alpha_2 s_1, \alpha_2 t_1, 0)^T & \text{for case (i)}, \\ x_3 := (\alpha_1 - \alpha_2 s_1, \alpha_2 t_1, 0)^T & \text{for case (ii)}. \end{cases}$$

We then have $\alpha_1 = |x_1 - x_2| > 0$, $\alpha_3 = |x_1 - x_4| > 0$, and

$$\alpha_2 = \begin{cases} |x_1 - x_3| > 0 & \text{for case (i)}, \\ |x_2 - x_3| > 0 & \text{for case (ii)}. \end{cases}$$

In the following, we impose conditions for $T \in \mathfrak{T}^{(2)}$ in the two-dimensional case and $T \in \mathfrak{T}_1^{(3)} \cup \mathfrak{T}_2^{(3)} =: \mathfrak{T}^{(3)}$ in the three-dimensional case.

Condition 1 (Case in which $d = 2$) Let $T \in \mathfrak{T}^{(2)}$ with the vertices x_i ($i = 1, \dots, 3$) introduced in Sect. 2.1.1. We assume that $\overline{x_2 x_3}$ is the longest edge of T; i.e., $h_T := |x_2 - x_3|$. Recall that $\alpha_1 = |x_1 - x_2|$ and $\alpha_2 = |x_1 - x_3|$. We then assume that $\alpha_2 \le \alpha_1$. Note that $\alpha_1 = \mathcal{O}(h_T)$.

Condition 2 (Case in which $d = 3$) Let $T \in \mathfrak{T}^{(3)}$ with the vertices x_i ($i = 1, \dots, 4$) introduced in Sect. 2.1.2. Let L_i ($1 \le i \le 6$) be the edges of T. We denote by $L_{\min}$ the edge of T that has the minimum length; i.e., $|L_{\min}| = \min_{1 \le i \le 6} |L_i|$. We set $\alpha_2 := |L_{\min}|$ and assume that the end points of $L_{\min}$ are either $\{x_1, x_3\}$ or $\{x_2, x_3\}$.

Among the four edges that share an end point with $L_{\min}$, we take the longest edge $L_{\max}^{(\min)}$. Let x_1 and x_2 be the end points of edge $L_{\max}^{(\min)}$. Thus, we have that

$$\alpha_1 = |L_{\max}^{(\min)}| = |x_1 - x_2|.$$

Consider cutting $\mathbb{R}^3$ with the plane that contains the midpoint of edge $L_{\max}^{(\min)}$ and is perpendicular to the vector $x_1 - x_2$. We then have two cases:

(Type i) x_3 and x_4 belong to the same half-space;
(Type ii) x_3 and x_4 belong to different half-spaces.

In each respective case, we set

(Type i) x_1 and x_3 as the end points of $L_{\min}$, that is, $\alpha_2 = |x_1 - x_3|$;
(Type ii) x_2 and x_3 as the end points of $L_{\min}$, that is, $\alpha_2 = |x_2 - x_3|$.

Finally, we set $\alpha_3 = |x_1 - x_4|$. Note that we implicitly assume that x_1 and x_4 belong to the same half-space. In addition, note that $\alpha_1 = \mathcal{O}(h_T)$.

Remark 1 Let $\mathbb{T}_h$ be a conformal mesh. We assume that any simplex $T_0 \in \mathbb{T}_h$ is transformed into $T_1 \in \mathfrak{T}^{(2)}$ such that Condition 1 is satisfied (in the two-dimensional case) or $T_i \in \mathfrak{T}_i^{(3)}$, $i = 1, 2$, such that Condition 2 is satisfied (in the three-dimensional case) through appropriate rotation, translation, and mirror imaging. Note that none of the lengths of the edges of a simplex or the measure of the simplex is changed by the transformation.

Assumption 1 In anisotropic interpolation error analysis, we may impose the following geometric conditions for the simplex T:

1. If $d = 2$, there are no additional conditions;
2. If $d = 3$, there exists a positive constant M, independent of h_T, such that $|s_{22}| \le M \frac{\alpha_2 t_1}{\alpha_3}$. Note that if $s_{22} \neq 0$, this condition means that the order with respect to h_T of α_3 coincides with the order of α_2, whereas if $s_{22} = 0$, the order of α_3 may be different from that of α_2.

2.2 Affine mappings

In our strategy, we adopt the following affine mappings.

Definition 1 (*Affine mappings*) Let $\widetilde{T}, \widehat{T} \subset \mathbb{R}^d$ be the simplices defined in Sects. 2.1.1 and 2.1.2. That is,

$$\widetilde{T} = \widehat{\Phi}(\widehat{T}), \quad T = \widetilde{\Phi}(\widetilde{T}) \quad \text{with} \quad \tilde{x} := \widehat{\Phi}(\hat{x}) := \widehat{A}^{(d)}\hat{x}, \quad x := \widetilde{\Phi}(\tilde{x}) := \widetilde{A}\tilde{x}.$$

We then define an affine mapping $\Phi_T : \widehat{T} \to T$ as

$$\Phi_T := \widetilde{\Phi} \circ \widehat{\Phi} : \widehat{T} \to T, \ x := \Phi_T(\hat{x}) := A_T\hat{x}, \quad A_T := \widetilde{A}\widehat{A}^{(d)}. \tag{11}$$

Furthermore, let Φ_{T_0} be an affine mapping defined as

$$\Phi_{T_0} : T \ni x \mapsto A_{T_0}x + b_{T_0} \in T_0, \tag{12}$$

where $b_{T_0} \in \mathbb{R}^d$ and $A_{T_0} \in O(d)$ is a rotation and mirror imaging matrix. We then define an affine mapping $\Phi : \widehat{T} \to T_0$ as

$$\Phi := \Phi_{T_0} \circ \Phi_T : \widehat{T} \to T_0, \ x^{(0)} := \Phi(\hat{x}) = (\Phi_{T_0} \circ \Phi_T)(\hat{x}) = A\hat{x} + b_{T_0}, \tag{13}$$

where $A := A_{T_0}A_T$.

2.3 Finite element generation

We follow the procedure described in [12, Section 1.4.1 and 1.2.1]; see also [16, Section 3.5].

For the reference element $\widehat{T}$ defined in Sects. 2.1.1 and 2.1.2, let $\{\widehat{T}, \widehat{P}, \widehat{\Sigma}\}$ be a fixed reference finite element, where $\widehat{P}$ is a vector space of functions $\hat{p} : \widehat{T} \to \mathbb{R}^n$ for some positive integer n (typically $n = 1$ or $n = d$) and $\widehat{\Sigma}$ is a set of n_0 linear forms $\{\hat{\chi}_1, \dots, \hat{\chi}_{n_0}\}$ such that

$$\widehat{P} \ni \hat{p} \mapsto (\hat{\chi}_1(\hat{p}), \dots, \hat{\chi}_{n_0}(\hat{p}))^T \in \mathbb{R}^{n_0}$$

is bijective; i.e., $\widehat{\Sigma}$ is a basis for $\mathcal{L}(\widehat{P};\mathbb{R})$. Further, we denote by $\{\hat{\theta}_1, \dots, \hat{\theta}_{n_0}\}$ in $\widehat{P}$ the local ($\mathbb{R}^n$-valued) shape functions such that

$$\hat{\chi}_i(\hat{\theta}_j) = \delta_{ij}, \quad 1 \leq i,j \leq n_0.$$

Let $V(\widehat{T})$ be a normed vector space of functions $\hat{\varphi} : \widehat{T} \to \mathbb{R}^n$ such that $\widehat{P} \subset V(\widehat{T})$ and the linear forms $\{\hat{\chi}_1, \dots, \hat{\chi}_{n_0}\}$ can be extended to $V(\widehat{T})'$. The local interpolation operator $I_{\widehat{T}}$ is then defined by

$$I_{\widehat{T}} : V(\widehat{T}) \ni \hat{\varphi} \mapsto \sum_{i=1}^{n_0} \hat{\chi}_i(\hat{\varphi})\hat{\theta}_i \in \widehat{P}. \tag{14}$$

It is obvious that

$$\hat{\chi}_i(I_{\widehat{T}}\hat{\varphi}) = \hat{\chi}_i(\hat{\varphi}) \quad \forall \hat{\varphi} \in V(\widehat{T}), \quad i = 1, \dots, n_0, \tag{15}$$

$$I_{\widehat{T}}\hat{p} = \hat{p} \quad \forall \hat{p} \in \widehat{P}. \tag{16}$$

Let $\boldsymbol{\Phi}$ be the affine mapping defined in Eq. (13). For $T_0 = \boldsymbol{\Phi}(\widehat{T})$, we first define a Banach space $V(T_0)$ of $\mathbb{R}^n$-valued functions that is the counterpart of $V(\widehat{T})$ and define a linear bijection mapping by

$$\psi := \psi_{\widehat{T}} \circ \psi_{\widetilde{T}} \circ \psi_T : V(T_0) \ni \varphi \mapsto \hat{\varphi} := \psi(\varphi) := \varphi \circ \boldsymbol{\Phi} \in V(\widehat{T})$$

with the three linear bijection mappings

$$\begin{aligned}
\psi_T &: V(T_0) \ni \varphi_0 \mapsto \varphi := \psi_T(\varphi_0) := \varphi_0 \circ \boldsymbol{\Phi}_{T_0} \in V(T), \\
\psi_{\widetilde{T}} &: V(T) \ni \varphi \mapsto \tilde{\varphi} := \psi_{\widetilde{T}}(\varphi) := \varphi \circ \widetilde{\boldsymbol{\Phi}} \in V(\widetilde{T}), \\
\psi_{\widehat{T}} &: V(\widetilde{T}) \ni \tilde{\varphi} \mapsto \hat{\varphi} := \psi_{\widehat{T}}(\tilde{\varphi}) := \tilde{\varphi} \circ \widehat{\boldsymbol{\Phi}} \in V(\widehat{T}).
\end{aligned}$$

The triple $\{\widetilde{T}, \widetilde{P}, \widetilde{\Sigma}\}$ is defined as

$$\begin{cases}
\widetilde{T} = \widehat{\boldsymbol{\Phi}}(\widehat{T}); \\
\widetilde{P} = \{\psi_{\widehat{T}}^{-1}(\hat{p}); \hat{p} \in \widehat{P}\}; \\
\widetilde{\Sigma} = \{\{\tilde{\chi}_i\}_{1 \le i \le n_0}; \tilde{\chi}_i = \hat{\chi}_i(\psi_{\widehat{T}}(\tilde{p})), \forall \tilde{p} \in \widetilde{P}, \hat{\chi}_i \in \widehat{\Sigma}\}.
\end{cases}$$

The triples $\{T, P, \Sigma\}$ and $\{T_0, P_0, \Sigma_0\}$ are similarly defined. These triples are finite elements and the local shape functions are $\tilde{\theta}_i = \psi_{\widehat{T}}^{-1}(\hat{\theta}_i)$, $\theta_i = \psi_{\widetilde{T}}^{-1}(\tilde{\theta}_i)$, and $\theta_{0,i} := \psi_T^{-1}(\theta_i)$ for $1 \le i \le n_0$, and the associated local interpolation operators are respectively defined by

$$I_{\widetilde{T}} : V(\widetilde{T}) \ni \tilde{\varphi} \mapsto I_{\widetilde{T}}\tilde{\varphi} := \sum_{i=1}^{n_0} \tilde{\chi}_i(\tilde{\varphi})\tilde{\theta}_i \in \widetilde{P}, \tag{17}$$

$$I_T : V(T) \ni \varphi \mapsto I_T\varphi := \sum_{i=1}^{n_0} \chi_i(\varphi)\theta_i \in P, \tag{18}$$

$$I_{T_0} : V(T_0) \ni \varphi_0 \mapsto I_{T_0}\varphi_0 := \sum_{i=1}^{n_0} \chi_{0,i}(\varphi_0)\theta_i \in P_0. \tag{19}$$

Proposition 1 *The diagrams*

$$\begin{array}{ccccccc}
V(T_0) & \xrightarrow{\psi_T} & V(T) & \xrightarrow{\psi_{\widetilde{T}}} & V(\widetilde{T}) & \xrightarrow{\psi_{\widehat{T}}} & V(\widehat{T}) \\
\downarrow I_{T_0} & & \downarrow I_T & & \downarrow I_{\widetilde{T}} & & \downarrow I_{\widehat{T}} \\
P_0 & \xrightarrow[\psi_T]{} & P & \xrightarrow[\psi_{\widetilde{T}}]{} & \widetilde{P} & \xrightarrow[\psi_{\widehat{T}}]{} & \widehat{P}
\end{array}$$

commute.

Proof See, for example [12, Proposition 1.62]. □

2.4 New parameters

In a previous paper [16], we proposed two geometric parameters,

Definition 2 The parameter H_T is defined as

$$H_T := \frac{\prod_{i=1}^d \alpha_i}{|T|} h_T,$$

and the parameter H_{T_0} is defined as

$$H_{T_0} := \frac{h_{T_0}^2}{|T_0|} \min_{1 \le i \le 3} |L_i| \quad \text{if } d = 2, \quad H_{T_0} := \frac{h_{T_0}^2}{|T_0|} \min_{1 \le i,j \le 6, i \ne j} |L_i||L_j| \quad \text{if } d = 3$$

where L_i denotes the edges of the simplex $T_0 \subset \mathbb{R}^d$.

The following lemma shows the equivalence between H_T and H_{T_0}.

Lemma 1 *It holds that*

$$\frac{1}{2} H_{T_0} < H_T < 2 H_{T_0}.$$

Furthermore, in the two-dimensional case, H_{T_0} *is equivalent to the circumradius* R_2 *of* T_0.

Proof The proof can be found in [16, Lemma 3]. □

Remark 2 We set

$$H(h) := \max_{T_0 \in \mathbb{T}_h} H_{T_0}.$$

As we stated in the Introduction, if the maximum-angle condition is violated, the parameter $H(h)$ may diverge as $h \to 0$ on anisotropic meshes. Therefore, imposing the maximum-angle condition for mesh partitions guarantees the convergence of finite element methods [3]. Reference [4] studied cases in which the finite element solution may not converge to the exact solution.

We now state the following theorem concerning the new condition.

Theorem 1 *Condition* (7) *holds if and only if there exist* $0 < \gamma_1, \gamma_2 < \pi$ *such that*

$$d = 2: \quad \theta_{T_0,\max} \leq \gamma_1 \quad \forall \mathbb{T}_h \in \{\mathbb{T}_h\}, \quad \forall T_0 \in \mathbb{T}_h, \tag{20}$$

where $\theta_{T_0,\max}$ *is the maximum angle of* T_0, *and*

$$d = 3: \quad \theta_{T_0,\max} \leq \gamma_2, \quad \psi_{T_0,\max} \leq \gamma_2 \quad \forall \mathbb{T}_h \in \{\mathbb{T}_h\}, \quad \forall T_0 \in \mathbb{T}_h, \tag{21}$$

where $\theta_{T_0,\max}$ *is the maximum angle of all triangular faces of the tetrahedron* T_0 *and* $\psi_{T_0,\max}$ *is the maximum dihedral angle of* T_0. *Conditions* (20) *and* (21) *together constitute the maximum-angle condition.*

Proof In the case of $d = 2$, we use the previous result presented in [19]; i.e., there exists a constant $\gamma_3 > 0$ such that

$$\frac{R_2}{h_{T_0}} \leq \gamma_3 \quad \forall \mathbb{T}_h \in \{\mathbb{T}_h\}, \quad \forall T_0 \in \mathbb{T}_h,$$

if and only if condition (20) is satisfied. Combining this result with H_{T_0} being equivalent to the circumradius R_2 of T_0, we have the desired conclusion. In the case of $d = 3$, the proof can be found in a recent paper [18]. □

Lemma 2 *It holds that*

$$\|\widehat{A}^{(d)}\|_2 \leq h_T, \quad \|\widehat{A}^{(d)}\|_2 \|(\widehat{A}^{(d)})^{-1}\|_2 = \frac{\max\{\alpha_1, \dots, \alpha_d\}}{\min\{\alpha_1, \dots, \alpha_d\}}, \tag{22a}$$

$$\|\widetilde{A}\|_2 \leq \begin{cases} \sqrt{2} & \text{if } d = 2, \\ 2 & \text{if } d = 3, \end{cases} \quad \|\widetilde{A}\|_2 \|\widetilde{A}^{-1}\|_2 \leq \begin{cases} \frac{\alpha_1\alpha_2}{|T|} = \frac{H_T}{h_T} & \text{if } d = 2, \\ \frac{2}{3}\frac{\alpha_1\alpha_2\alpha_3}{|T|} = \frac{2}{3}\frac{H_T}{h_T} & \text{if } d = 3, \end{cases} \tag{22b}$$

$$\|A_{T_0}\|_2 = 1, \quad \|A_{T_0}^{-1}\|_2 = 1. \tag{22c}$$

Furthermore, we have

$$|\det(A_T)| = |\det(\widetilde{A})||\det(\widehat{A}^{(d)})| = d!|T|, \quad |\det(A_{T_0})| = 1. \tag{23}$$

Proof The proof of (22b) can be found in [16, (4.4), (4.5), (4.6), and (4.7)]. The inequality (22a) is easily proved. Because $A_{T_0} \in O(d)$, one easily finds that $A_{T_0}^{-1} \in O(d)$ and recovers Eq. (22c). The proof of equality (23) is standard. □

For matrix $A \in \mathbb{R}^{d \times d}$, we denote by $[A]_{ij}$ the (i, j)-component of A. We set $\|A\|_{\max} := \max_{1 \leq i,j \leq d} |[A]_{ij}|$. Furthermore, we use the inequality

$$\|A\|_{\max} \leq \|A\|_2. \tag{24}$$

3 Scaling argument

This section gives estimates related to a scaling argument corresponding to [12, Lemma 1.101]. The estimates play major roles in our analysis. Furthermore, we use the following inequality (see [12, Exercise 1.20]). Let $0 < r \leq s$ and $a_i \geq 0, i = 1, 2, \dots, n$ ($n \in \mathbb{N}$), be real numbers. Then, we have that

$$\left(\sum_{i=1}^{n} a_i^s \right)^{1/s} \leq \left(\sum_{i=1}^{n} a_i^r \right)^{1/r}. \tag{25}$$

Lemma 3 *Let $s \geq 0$ and $1 \leq p \leq \infty$. There exist positive constants c_1 and c_2 such that, for all $T_0 \in \mathbb{T}_h$ and $\varphi_0 \in W^{s,p}(T_0)$,*

$$c_1 |\varphi_0|_{W^{s,p}(T_0)} \leq |\varphi|_{W^{s,p}(T)} \leq c_2 |\varphi_0|_{W^{s,p}(T_0)} \tag{26}$$

with $\varphi = \varphi_0 \circ \Phi_{T_0}$.

Proof The following inequalities can be found in [12, Lemma 1.101]. There exists a positive constant c such that, for all $T_0 \in \mathbb{T}_h$ and $\varphi_0 \in W^{s,p}(T_0)$,

$$|\varphi|_{W^{s,p}(T)} \leq c \|A_{T_0}\|_2^s |\det(A_{T_0})|^{-\frac{1}{p}} |\varphi_0|_{W^{s,p}(T_0)}, \tag{27}$$

$$|\varphi_0|_{W^{s,p}(T_0)} \leq c \|A_{T_0}^{-1}\|_2^s |\det(A_{T_0})|^{\frac{1}{p}} |\varphi|_{W^{s,p}(T)} \tag{28}$$

with $\varphi = \varphi_0 \circ \Phi_{T_0}$. Using Eqs. (27) and (28) together with Eqs. (22c) and (23) yields Eq. (26). □

Lemma 4 *Let $m \in \mathbb{N}_0$ and $p \in [0, \infty)$. Let $\beta := (\beta_1, \dots, \beta_d) \in \mathbb{N}_0^d$ be a multi-index with $|\beta| = m$. Then, for any $\hat{\varphi} \in W^{m,p}(\widehat{T})$ with $\tilde{\varphi} = \hat{\varphi} \circ \widehat{\Phi}^{-1}$ and $\varphi = \tilde{\varphi} \circ \widetilde{\Phi}^{-1}$, it holds that*

$$|\varphi|_{W^{m,p}(T)} \leq C_1^{SA} |\det(A_T)|^{\frac{1}{p}} \|\widetilde{A}^{-1}\|_2^m \left(\sum_{|\beta|=m} (\alpha^{-\beta})^p \|\partial^\beta \hat{\varphi}\|_{L^p(\widehat{T})}^p \right)^{1/p}, \tag{29}$$

where C_1^{SA} is a constant that is independent of T and $\widetilde{T}$. When $p = \infty$, for any $\hat{\varphi} \in W^{m,\infty}(\widehat{T})$ with $\tilde{\varphi} = \hat{\varphi} \circ \widehat{\Phi}^{-1}$ and $\varphi = \tilde{\varphi} \circ \widetilde{\Phi}^{-1}$, it holds that

$$|\varphi|_{W^{m,\infty}(T)} \leq C_1^{SA,\infty} \|\widetilde{A}^{-1}\|_2^m \max_{|\beta|=m} \left(\alpha^{-\beta} \|\partial^\beta \hat{\varphi}\|_{L^\infty(\widehat{T})} \right), \tag{30}$$

where $C_1^{SA,\infty}$ is a constant that is independent of T and $\widetilde{T}$.

Proof Let $p \in [1, \infty)$. Because the space $\mathcal{C}^m(\widehat{T})$ is dense in the space $W^{m,p}(\widehat{T})$, we show that Eq. (29) holds for $\hat{\varphi} \in \mathcal{C}^m(\widehat{T})$ with $\tilde{\varphi} = \hat{\varphi} \circ \widehat{\Phi}^{-1}$ and $\varphi = \tilde{\varphi} \circ \widetilde{\Phi}^{-1}$. From $\hat{x}_j = \alpha_j^{-1} \tilde{x}_j$, we have that, for any multi-index β,

$$\partial^\beta \tilde{\varphi} = \alpha_1^{-\beta_1} \cdots \alpha_d^{-\beta_d} \partial^\beta \hat{\varphi} = \alpha^{-\beta} \partial^\beta \hat{\varphi}. \tag{31}$$

Through a change of variable, we obtain

$$|\tilde{\varphi}|^p_{W^{m,p}(\widetilde{T})} = \sum_{|\beta|=m} \|\partial^\beta \tilde{\varphi}\|^p_{L^p(\widetilde{T})} = |\det(\widehat{A}^{(d)})| \sum_{|\beta|=m} (\alpha^{-\beta})^p \|\partial^\beta \hat{\varphi}\|^p_{L^p(\widehat{T})}. \tag{32}$$

From the standard estimate in [12, Lemma 1.101], we have

$$|\varphi|_{W^{m,p}(T)} \le C_1^{SA} |\det(\widetilde{A})|^{\frac{1}{p}} \|\widetilde{A}^{-1}\|_2^m |\tilde{\varphi}|_{W^{m,p}(\widetilde{T})}. \tag{33}$$

Inequality (29) follows from Eqs. (32) and (33) with Eq. (23).

We consider the case in which $p = \infty$. A function $\hat{\varphi} \in W^{m,\infty}(\widehat{T})$ belongs to the space $W^{m,p}(\widehat{T})$ for any $p \in [1, \infty)$. Therefore, it holds that $\tilde{\varphi} \in W^{m,p}(\widetilde{T})$ for any $p \in [1, \infty)$ and, from Eq. (25), we obtain

$$\begin{aligned}
\|\partial^\gamma \tilde{\varphi}\|_{L^p(\widetilde{T})} &\le |\tilde{\varphi}|_{W^{|\gamma|,p}(\widetilde{T})} \\
&= |\det(\widehat{A}^{(d)})|^{\frac{1}{p}} \left(\sum_{|\beta|=|\gamma|} (\alpha^{-\beta})^p \|\partial^\beta \hat{\varphi}\|^p_{L^p(\widehat{T})} \right)^{1/p} \\
&\le \left(\sup_{1 \le p} |\det(\widehat{A}^{(d)})|^{\frac{1}{p}} \right) \sum_{|\beta|=|\gamma|} \alpha^{-\beta} \|\partial^\beta \hat{\varphi}\|_{L^p(\widehat{T})} \\
&\le c \left(\sup_{1 \le p} |\det(\widehat{A}^{(d)})|^{\frac{1}{p}} \right) \sum_{|\beta|=|\gamma|} \alpha^{-\beta} \|\partial^\beta \hat{\varphi}\|_{L^\infty(\widehat{T})} < \infty
\end{aligned} \tag{34}$$

for the multi-index $\gamma \in \mathbb{N}_0^d$ with $|\gamma| \le m$. This implies that the function $\partial^\gamma \tilde{\varphi}$ is in the space $L^\infty(\widetilde{T})$ for each $|\gamma| \le m$. Therefore, we have that $\tilde{\varphi} \in W^{m,\infty}(\widetilde{T})$. Taking the limit $p \to \infty$ in Eq. (34) and using $\lim_{p \to \infty} \|\cdot\|_{L^p(\widetilde{T})} = \|\cdot\|_{L^\infty(\widetilde{T})}$, we have

$$|\tilde{\varphi}|_{W^{m,\infty}(\widetilde{T})} \le c \max_{|\beta|=m} \left(\alpha^{-\beta} \|\partial^\beta \hat{\varphi}\|_{L^\infty(\widehat{T})} \right). \tag{35}$$

From the standard estimate in [12, Lemma 1.101], we have

$$|\varphi|_{W^{m,\infty}(T)} \le c \|\widetilde{A}^{-1}\|_2^m |\tilde{\varphi}|_{W^{m,\infty}(\widetilde{T})}. \tag{36}$$

Inequality (30) follows from Eqs. (35) and (36). □

We now introduce the following new notation.

Definition 3 We define a parameter $\mathscr{H}_i$, $i = 1, \ldots, d$, as

$$\begin{cases} \mathscr{H}_1 := \alpha_1, \quad \mathscr{H}_2 := \alpha_2 t \quad \text{if } d = 2, \\ \mathscr{H}_1 := \alpha_1, \quad \mathscr{H}_2 := \alpha_2 t_1, \quad \mathscr{H}_3 := \alpha_3 t_2 \quad \text{if } d = 3. \end{cases}$$

For a multi-index $\beta = (\beta_1, \dots, \beta_d) \in \mathbb{N}_0^d$, we use the notation

$$\mathscr{H}^\beta := \mathscr{H}_1^{\beta_1} \cdots \mathscr{H}_d^{\beta_d}, \quad \mathscr{H}^{-\beta} := \mathscr{H}_1^{-\beta_1} \cdots \mathscr{H}_d^{-\beta_d}.$$

We also define $\alpha^\beta := \alpha_1^{\beta_1} \cdots \alpha_d^{\beta_d}$ and $\alpha^{-\beta} := \alpha_1^{-\beta_1} \cdots \alpha_d^{-\beta_d}$.

Definition 4 We define vectors $r_n \in \mathbb{R}^d$, $n = 1, \dots, d$, as follows. If $d = 2$,

$$r_1 := (1, 0)^T, \quad r_2 := (s, t)^T,$$

and if $d = 3$,

$$r_1 := (1, 0, 0)^T, \quad r_3 := (s_{21}, s_{22}, t_2)^T,$$
$$\begin{cases} r_2 := (s_1, t_1, 0)^T & \text{for case (i)}, \\ r_2 := (-s_1, t_1, 0)^T & \text{for case (ii)}. \end{cases}$$

Furthermore, we define a directional derivative as

$$\frac{\partial}{\partial r_i^{(0)}} := (A_{T_0} r_i) \cdot \nabla_{x^{(0)}} = \sum_{j_0=1}^{d} (A_{T_0} r_i)_{j_0} \frac{\partial}{\partial x_{j_0}^{(0)}}, \quad i \in \{1 : d\},$$

where $A_{T_0} \in O(d)$ is the orthogonal matrix defined in Eq. (12). For a multi-index $\beta = (\beta_1, \dots, \beta_d) \in \mathbb{N}_0^d$, we use the notation

$$\partial_{r^{(0)}}^\beta := \frac{\partial^{|\beta|}}{(\partial r_1^{(0)})^{\beta_1} \dots (\partial r_d^{(0)})^{\beta_d}}.$$

Note 1 Recall that

$$\begin{aligned} &|s| \le 1, \quad \alpha_2 \le \alpha_1 && \text{if } d = 2, \\ &|s_1| \le 1, \quad |s_{21}| \le 1, \ \alpha_2 \le \alpha_3 \le \alpha_1 && \text{if } d = 3. \end{aligned}$$

When $d = 3$, if Assumption 1 is imposed, there exists a positive constant M, independent of h_T, such that $|s_{22}| \le M \frac{\alpha_2 t_1}{\alpha_3}$. Thus, if $d = 2$, we have

$$\alpha_1 |[\widetilde{A}]_{j1}| \le \mathscr{H}_j, \quad \alpha_2 |[\widetilde{A}]_{j2}| \le \mathscr{H}_j, \quad j = 1, 2,$$

and if $d = 3$, for $\widetilde{A} \in \{\widetilde{A}_1, \widetilde{A}_2\}$ and $j = 1, 2, 3$, we have

$$\alpha_1 |[\widetilde{A}]_{j1}| \le \mathscr{H}_j, \quad \alpha_2 |[\widetilde{A}]_{j2}| \le \mathscr{H}_j, \quad \alpha_3 |[\widetilde{A}]_{j3}| \le \max\{1, M\} \mathscr{H}_j, \quad j = 1, 2, 3.$$

Note 2 We use the following calculations in Lemma 5. For any multi-indices β and γ, we have

$$
\begin{aligned}
\partial_{\hat{x}}^{\beta+\gamma} &= \frac{\partial^{|\beta|+|\gamma|}}{\partial \hat{x}_1^{\beta_1} \cdots \partial \hat{x}_d^{\beta_d} \partial \hat{x}_1^{\gamma_1} \cdots \partial \hat{x}_d^{\gamma_d}} \\
&= \underbrace{\sum_{i_1^{(1)}=1}^{d} \alpha_1 [\widetilde{A}]_{i_1^{(1)} 1} \cdots \sum_{i_{\beta_1}^{(1)}=1}^{d} \alpha_1 [\widetilde{A}]_{i_{\beta_1}^{(1)} 1}}_{\beta_1 \text{times}} \cdots \underbrace{\sum_{i_1^{(d)}=1}^{d} \alpha_d [\widetilde{A}]_{i_1^{(d)} d} \cdots \sum_{i_{\beta_d}^{(d)}=1}^{d} \alpha_d [\widetilde{A}]_{i_{\beta_d}^{(d)} d}}_{\beta_d \text{times}} \\
&\quad \underbrace{\sum_{j_1^{(1)}=1}^{d} \alpha_1 [\widetilde{A}]_{j_1^{(1)} 1} \cdots \sum_{j_{\gamma_1}^{(1)}=1}^{d} \alpha_1 [\widetilde{A}]_{j_{\gamma_1}^{(1)} 1}}_{\gamma_1 \text{times}} \cdots \underbrace{\sum_{j_1^{(d)}=1}^{d} \alpha_d [\widetilde{A}]_{j_1^{(d)} d} \cdots \sum_{j_{\gamma_d}^{(d)}=1}^{d} \alpha_d [\widetilde{A}]_{j_{\gamma_d}^{(d)} d}}_{\gamma_d \text{times}} \\
&\quad \frac{\partial^{\beta_1}}{\underbrace{\partial x_{i_1^{(1)}} \cdots \partial x_{i_{\beta_1}^{(1)}}}_{\beta_1 \text{times}}} \cdots \frac{\partial^{\beta_d}}{\underbrace{\partial x_{i_1^{(d)}} \cdots \partial x_{i_{\beta_d}^{(d)}}}_{\beta_d \text{times}}} \frac{\partial^{\gamma_1}}{\underbrace{\partial x_{j_1^{(1)}} \cdots \partial x_{j_{\gamma_1}^{(1)}}}_{\gamma_1 \text{times}}} \cdots \frac{\partial^{\gamma_d}}{\underbrace{\partial x_{j_1^{(d)}} \cdots \partial x_{j_{\gamma_d}^{(d)}}}_{\gamma_d \text{times}}} .
\end{aligned}
$$

Let $\hat{\varphi} \in \mathcal{C}^{\ell}(\widehat{T})$ with $\tilde{\varphi} = \hat{\varphi} \circ \widehat{\Phi}^{-1}$ and $\varphi = \tilde{\varphi} \circ \widetilde{\Phi}^{-1}$. Then, for $1 \le i \le d$,

$$
\left| \frac{\partial \hat{\varphi}}{\partial \hat{x}_i} \right| \le \sum_{i_1^{(1)}=1}^{d} \alpha_i \left| [\widetilde{A}]_{i_1^{(1)} i} \right| \left| \frac{\partial \varphi}{\partial x_{i_1^{(1)}}} \right| \le \begin{cases} \alpha_i \|\widetilde{A}\|_{\max} \sum_{i_1^{(1)}=1}^{d} \left| \frac{\partial \varphi}{\partial x_{i_1^{(1)}}} \right| & \text{or,} \\ c \sum_{i_1^{(1)}=1}^{d} \mathscr{H}_{i_1^{(1)}} \left| \frac{\partial \varphi}{\partial x_{i_1^{(1)}}} \right|, \end{cases}
$$

and for $1 \le i,j \le d$,

$$
\begin{aligned}
\left| \frac{\partial^2 \hat{\varphi}}{\partial \hat{x}_i \partial \hat{x}_j} \right| &= \left| \sum_{i_1^{(1)}, j_1^{(1)}=1}^{d} \alpha_i \alpha_j [\widetilde{A}]_{i_1^{(1)} i} [\widetilde{A}]_{j_1^{(1)} j} \frac{\partial^2 \varphi}{\partial x_{i_1^{(1)}} \partial x_{j_1^{(1)}}} \right| \\
&\le \begin{cases} \alpha_i \alpha_j \|\widetilde{A}\|_{\max}^2 \sum_{i_1^{(1)}, j_1^{(1)}=1}^{d} \left| \frac{\partial^2 \varphi}{\partial x_{i_1^{(1)}} \partial x_{j_1^{(1)}}} \right| & \text{or,} \\ \alpha_j \sum_{j_1^{(1)}=1}^{d} |[\widetilde{A}]_{j_1^{(1)} j}| \left| \sum_{i_1^{(1)}=1}^{d} \alpha_i [\widetilde{A}]_{i_1^{(1)} i} \frac{\partial^2 \varphi}{\partial x_{i_1^{(1)}} \partial x_{j_1^{(1)}}} \right| & \\ \quad \le c \alpha_j \|\widetilde{A}\|_{\max} \sum_{j_1^{(1)}=1}^{d} \sum_{i_1^{(1)}=1}^{d} \mathscr{H}_{i_1^{(1)}} \left| \frac{\partial^2 \varphi}{\partial x_{i_1^{(1)}} \partial x_{j_1^{(1)}}} \right| & \text{or,} \\ c \sum_{i_1^{(1)}=1}^{d} \sum_{j_1^{(1)}=1}^{d} \mathscr{H}_{i_1^{(1)}} \mathscr{H}_{j_1^{(1)}} \left| \frac{\partial^2 \varphi}{\partial x_{i_1^{(1)}} \partial x_{j_1^{(1)}}} \right|. & \end{cases}
\end{aligned}
$$

Lemma 5 *Suppose that Assumption 1 is imposed. Let* $m \in \mathbb{N}_0$, $\ell \in \mathbb{N}_0$ *with* $\ell \ge m$ *and* $p \in [0, \infty]$. *Let* $\beta := (\beta_1, \ldots, \beta_d) \in \mathbb{N}_0^d$ *and* $\gamma := (\gamma_1, \ldots, \gamma_d) \in \mathbb{N}_0^d$

be multi-indices with $|\beta| = m$ *and* $|\gamma| = \ell - m$*. Then, for any* $\hat{\varphi} \in W^{\ell,p}(\widehat{T})$ *with* $\tilde{\varphi} = \hat{\varphi} \circ \widehat{\Phi}^{-1}$ *and* $\varphi = \tilde{\varphi} \circ \widetilde{\Phi}^{-1}$*, it holds that*

$$\|\partial^\beta \partial^\gamma \hat{\varphi}\|_{L^p(\widehat{T})} \leq C_2^{SA} |\det(A_T)|^{-\frac{1}{p}} \|\widetilde{A}\|_2^m \alpha^\beta \sum_{|\epsilon|=|\gamma|} \mathscr{H}^\epsilon |\partial^\epsilon \varphi|_{W^{m,p}(T)}, \tag{37}$$

where C_2^{SA} *is a constant that is independent of* T *and* $\widetilde{T}$*. Here, for* $p = \infty$ *and any positive real* x, $x^{-\frac{1}{p}} = 1$.

Proof Let $\varepsilon = (\varepsilon_1, \dots, \varepsilon_d) \in \mathbb{N}_0^d$ and $\delta = (\delta_1, \dots, \delta_d) \in \mathbb{N}_0^d$ be multi-indies with $|\varepsilon| = |\gamma|$ and $|\delta| = |\beta|$. Let $p \in [1, \infty)$. Because the space $\mathcal{C}^\ell(\widehat{T})$ is dense in the space $W^{\ell,p}(\widehat{T})$, we show that Eq. (37) holds for $\hat{\varphi} \in \mathcal{C}^\ell(\widehat{T})$ with $\tilde{\varphi} = \hat{\varphi} \circ \widehat{\Phi}^{-1}$ and $\varphi = \tilde{\varphi} \circ \widetilde{\Phi}^{-1}$. Through a simple calculation, we obtain

$$\begin{aligned}
|\partial^{\beta+\gamma} \hat{\varphi}| &= \left| \frac{\partial^\ell \hat{\varphi}}{\partial \hat{x}_1^{\beta_1} \cdots \partial \hat{x}_d^{\beta_d} \partial \hat{x}_1^{\gamma_1} \cdots \partial \hat{x}_d^{\gamma_d}} \right| \\
&\leq c \alpha^\beta \|\widetilde{A}\|_{\max}^{|\beta|} \underbrace{\sum_{i_1^{(1)}=1}^d \cdots \sum_{i_{\beta_1}^{(1)}=1}^d}_{\beta_1 \text{times}} \cdots \underbrace{\sum_{i_1^{(d)}=1}^d \cdots \sum_{i_{\beta_d}^{(d)}=1}^d}_{\beta_d \text{times}} \underbrace{\sum_{j_1^{(1)}=1}^d \cdots \sum_{j_{\gamma_1}^{(1)}=1}^d}_{\gamma_1 \text{times}} \cdots \underbrace{\sum_{j_1^{(d)}=1}^d \cdots \sum_{j_{\gamma_d}^{(d)}=1}^d}_{\gamma_d \text{times}} \\
&\quad \underbrace{\mathscr{H}_{j_1^{(1)}} \cdots \mathscr{H}_{j_{\varepsilon_1}^{(1)}}}_{\gamma_1 \text{times}} \cdots \underbrace{\mathscr{H}_{j_1^{(d)}} \cdots \mathscr{H}_{j_{\varepsilon_d}^{(d)}}}_{\gamma_d \text{times}} \\
&\quad \left| \underbrace{\frac{\partial^{\beta_1}}{\partial x_{i_1^{(1)}} \cdots \partial x_{i_{\beta_1}^{(1)}}}}_{\beta_1 \text{times}} \cdots \underbrace{\frac{\partial^{\beta_d}}{\partial x_{i_1^{(d)}} \cdots \partial x_{i_{\beta_d}^{(d)}}}}_{\beta_d \text{times}} \underbrace{\frac{\partial^{\gamma_1}}{\partial x_{j_1^{(1)}} \cdots \partial x_{j_{\gamma_1}^{(1)}}}}_{\gamma_1 \text{times}} \cdots \underbrace{\frac{\partial^{\gamma_d}}{\partial x_{j_1^{(d)}} \cdots \partial x_{j_{\gamma_d}^{(d)}}}}_{\gamma_d \text{times}} \varphi \right| \\
&\leq c \alpha^\beta \|\widetilde{A}\|_{\max}^{|\beta|} \sum_{|\delta|=|\beta|} \sum_{|\varepsilon|=|\gamma|} \mathscr{H}^\varepsilon |\partial^\delta \partial^\varepsilon \varphi|.
\end{aligned}$$

Using Eq. (24), we then have that

$$\begin{aligned}
\int_{\widehat{T}} |\partial^\beta \partial^\gamma \hat{\varphi}|^p d\hat{x} &\leq c \|\widetilde{A}\|_2^{mp} \alpha^{\beta p} \sum_{|\delta|=|\beta|} \sum_{|\epsilon|=|\gamma|} \mathscr{H}^{\varepsilon p} \int_{\widehat{T}} |\partial^\delta \partial^\varepsilon \varphi|^p d\hat{x} \\
&= c |\det(A_T)|^{-1} \|\widetilde{A}\|_2^{mp} \alpha^{\beta p} \sum_{|\delta|=|\beta|} \sum_{|\epsilon|=|\gamma|} \mathscr{H}^{\varepsilon p} \int_T |\partial^\delta \partial^\varepsilon \varphi|^p dx.
\end{aligned}$$

Therefore, using (25), we obtain

$$\|\partial^\beta \partial^\gamma \hat{\varphi}\|_{L^p(\widehat{T})} \leq c |\det(A_T)|^{-\frac{1}{p}} \|\widetilde{A}\|_2^m \alpha^\beta \sum_{|\epsilon|=|\gamma|} \mathscr{H}^\epsilon |\partial^\epsilon \varphi|_{W^{m,p}(T)},$$

which recovers Eq. (37).

We consider the case in which $p = \infty$. A function $\varphi \in W^{\ell,\infty}(T)$ belongs to the space $W^{\ell,p}(T)$ for any $p \in [1,\infty)$. Therefore, it holds that $\hat{\varphi} \in W^{\ell,p}(\widehat{T})$ for any $p \in [1,\infty)$, and thus,

$$\begin{aligned} \|\partial^\beta \partial^\gamma \hat{\varphi}\|_{L^p(\widehat{T})} &\leq c |\det(A_T)|^{-\frac{1}{p}} \|\widetilde{A}\|_2^m \alpha^\beta \sum_{|\epsilon|=|\gamma|} \mathscr{H}^\epsilon |\partial^\epsilon \varphi|_{W^{m,p}(T)} \\ &\leq c \|\widetilde{A}\|_2^m \alpha^\beta \sum_{|\epsilon|=|\gamma|} \mathscr{H}^\epsilon |\partial^\epsilon \varphi|_{W^{m,\infty}(T)} < \infty. \end{aligned} \tag{38}$$

This implies that the function $\partial^\beta \partial^\gamma \hat{\varphi}$ is in the space $L^\infty(\widehat{T})$. Inequality (37) for $p = \infty$ is obtained by taking the limit $p \to \infty$ in Eq. (38) on the basis that $\lim_{p\to\infty} \|\cdot\|_{L^p(\widehat{T})} = \|\cdot\|_{L^\infty(\widehat{T})}$. □

Remark 3 In inequality (37), it is possible to obtain the estimates in T_0 by specifically determining the matrix A_{T_0}.

Let $\ell = 2$, $m = 1$, and $p = q = 2$. Recall that

$$\Phi_{T_0} : T \ni x \mapsto x^{(0)} := A_{T_0} x + b_{T_0} \in T_0.$$

For $\varphi \in \mathcal{C}^2(T)$ with $\varphi_0 = \varphi \circ \Phi_{T_0}^{-1}$ and $1 \leq i,j \leq d$, we have

$$\left| \frac{\partial^2 \varphi}{\partial x_i \partial x_j}(x) \right| = \left| \sum_{i_1^{(1)}, j_1^{(1)}=1}^{2} [A_{T_0}]_{i_1^{(1)} i} [A_{T_0}]_{j_1^{(1)} j} \frac{\partial^2 \varphi_0}{\partial x^{(0)}_{i_1^{(1)}} \partial x^{(0)}_{j_1^{(1)}}}(x) \right|.$$

Let $d = 2$. We define the matrix A_{T_0} as

$$A_{T_0} := \begin{pmatrix} \cos\frac{\pi}{2} & -\sin\frac{\pi}{2} \\ \sin\frac{\pi}{2} & \cos\frac{\pi}{2} \end{pmatrix}.$$

Because $\|A_{T_0}\|_{\max} = 1$, we have

$$\left| \frac{\partial^2 \varphi}{\partial x_i \partial x_j}(x) \right| \leq \left| \frac{\partial^2 \varphi_0}{\partial x^{(0)}_{i+1} \partial x^{(0)}_{j+1}}(x) \right|,$$

where the indices i, $i+1$ and j, $j+1$ must be evaluated modulo 2. Because $|\det(A_{T_0})| = 1$, it holds that

$$\left\| \frac{\partial^2 \varphi}{\partial x_i \partial x_j} \right\|_{L^2(T)} \leq \left\| \frac{\partial^2 \varphi_0}{\partial x^{(0)}_{i_{i+1}} \partial x^{(0)}_{j+1}} \right\|_{L^2(T_0)}.$$

We then have

$$\sum_{j=1}^{2} \mathscr{H}_j \left| \frac{\partial \varphi}{\partial x_j} \right|_{H^1(T)} \leq \sum_{j=1}^{2} \mathscr{H}_j \left| \frac{\partial \varphi_0}{\partial x_{j+1}^{(0)}} \right|_{H^1(T_0)},$$

where the indices j, $j+1$ must be evaluated modulo 2.

We define the matrix A_{T_0} as

$$A_{T_0} := \begin{pmatrix} \cos\frac{\pi}{4} & -\sin\frac{\pi}{4} \\ \sin\frac{\pi}{4} & \cos\frac{\pi}{4} \end{pmatrix}.$$

We then have

$$\left| \frac{\partial^2 \varphi}{\partial x_i \partial x_j}(x) \right| \leq \frac{1}{\sqrt{2}} \sum_{i_1^{(1)}, j_1^{(1)}=1}^{2} \left| \frac{\partial^2 \varphi_0}{\partial x_{i_1^{(1)}}^{(0)} \partial x_{j_1^{(1)}}^{(0)}}(x) \right|,$$

which leads to

$$\left\| \frac{\partial^2 \varphi}{\partial x_i \partial x_j} \right\|_{L^2(T)}^2 \leq c \sum_{i_1^{(1)}, j_1^{(1)}=1}^{2} \left\| \frac{\partial^2 \varphi_0}{\partial x_{i_1^{(1)}}^{(0)} \partial x_{j_1^{(1)}}^{(0)}} \right\|_{L^2(T_0)}^2 \leq c |\varphi_0|_{H^2(T_0)}^2.$$

Using (25), we than have that

$$\sum_{j=1}^{2} \mathscr{H}_j \left| \frac{\partial \varphi}{\partial x_j} \right|_{H^1(T)} \leq \sum_{j=1}^{2} \mathscr{H}_j |\varphi_0|_{H^2(T_0)} \leq c h_{T_0} |\varphi_0|_{H^2(T_0)}.$$

In this case, anisotropic interpolation error estimates cannot be obtained.

Note 3 We use the following calculations in Lemma 6. For any multi-indices β and γ, we have

$$
\begin{aligned}
\partial_{\hat{x}}^{\beta+\gamma} &= \frac{\partial^{|\beta|+|\gamma|}}{\partial \hat{x}_1^{\beta_1} \cdots \partial \hat{x}_d^{\beta_d} \partial \hat{x}_1^{\gamma_1} \cdots \partial \hat{x}_d^{\gamma_d}} \\
&= \underbrace{\sum_{i_1^{(1)}, i_1^{(0,1)}=1}^{d} \alpha_1 [\widetilde{A}]_{i_1^{(1)} 1} [A_{T_0}]_{i_1^{(0,1)} i_1^{(1)}} \cdots \sum_{i_{\beta_1}^{(1)}, i_{\beta_1}^{(0,1)}=1}^{d} \alpha_1 [\widetilde{A}]_{i_{\beta_1}^{(1)} 1} [A_{T_0}]_{i_{\beta_1}^{(0,1)} i_{\beta_1}^{(1)}}}_{\beta_1 \text{times}} \cdots \\
&\quad \underbrace{\sum_{i_1^{(d)}, i_1^{(0,d)}=1}^{d} \alpha_d [\widetilde{A}]_{i_1^{(d)} d} [A_{T_0}]_{i_1^{(0,d)} i_1^{(d)}} \cdots \sum_{i_{\beta_d}^{(d)}, i_{\beta_d}^{(0,d)}=1}^{d} \alpha_d [\widetilde{A}]_{i_{\beta_d}^{(d)} d} [A_{T_0}]_{i_{\beta_d}^{(0,d)} i_{\beta_d}^{(d)}}}_{\beta_d \text{times}} \\
&\quad \underbrace{\sum_{j_1^{(1)}, j_1^{(0,1)}=1}^{d} \alpha_1 [\widetilde{A}]_{j_1^{(1)} 1} [A_{T_0}]_{j_1^{(0,1)} j_1^{(1)}} \cdots \sum_{j_{\gamma_1}^{(1)}, j_{\gamma_1}^{(0,1)}=1}^{d} \alpha_1 [\widetilde{A}]_{j_{\gamma_1}^{(1)} 1} [A_{T_0}]_{j_{\gamma_1}^{(0,1)} j_{\gamma_1}^{(1)}}}_{\gamma_1 \text{times}} \cdots \\
&\quad \underbrace{\sum_{j_1^{(d)}, j_1^{(0,d)}=1}^{d} \alpha_d [\widetilde{A}]_{j_1^{(d)} d} [A_{T_0}]_{j_1^{(0,d)} j_1^{(d)}} \cdots \sum_{j_{\gamma_d}^{(d)}, j_{\gamma_d}^{(0,d)}=1}^{d} \alpha_d [\widetilde{A}]_{j_{\gamma_d}^{(d)} d} [A_{T_0}]_{j_{\gamma_d}^{(0,d)} j_{\gamma_d}^{(d)}}}_{\gamma_d \text{times}} \\
&\quad \underbrace{\frac{\partial^{\beta_1}}{\partial x^{(0)}_{i_1^{(0,1)}} \cdots \partial x^{(0)}_{i_{\beta_1}^{(0,1)}}}}_{\beta_1 \text{times}} \cdots \underbrace{\frac{\partial^{\beta_d}}{\partial x^{(0)}_{i_1^{(0,d)}} \cdots \partial x^{(0)}_{i_{\beta_d}^{(0,d)}}}}_{\beta_d \text{times}} \underbrace{\frac{\partial^{\gamma_1}}{\partial x^{(0)}_{j_1^{(0,1)}} \cdots \partial x^{(0)}_{j_{\gamma_1}^{(0,1)}}}}_{\gamma_1 \text{times}} \cdots \underbrace{\frac{\partial^{\gamma_d}}{\partial x^{(0)}_{j_1^{(0,d)}} \cdots \partial x^{(0)}_{j_{\gamma_d}^{(0,d)}}}}_{\gamma_d \text{times}} .
\end{aligned}
$$

Let $\hat{\varphi} \in \mathcal{C}^{\ell}(\widehat{T})$ with $\tilde{\varphi} = \hat{\varphi} \circ \widehat{\Phi}^{-1}$, $\varphi = \tilde{\varphi} \circ \widetilde{\Phi}^{-1}$ and $\varphi_0 = \varphi \circ \Phi_{T_0}^{-1}$. Then, for $1 \leq i \leq d$,

$$
\begin{aligned}
\left| \frac{\partial \hat{\varphi}}{\partial \hat{x}_i} \right| &= \left| \sum_{i_1^{(1)}=1}^{d} \sum_{i_1^{(0,1)}=1}^{d} \alpha_i [\widetilde{A}]_{i_1^{(1)} i} [A_{T_0}]_{i_1^{(0,1)} i_1^{(1)}} \frac{\partial \varphi_0}{\partial x^{(0)}_{i_1^{(0,1)}}} \right| \\
&= \alpha_i \left| \sum_{i_1^{(1)}=1}^{d} \sum_{i_1^{(0,1)}=1}^{d} [A_{T_0}]_{i_1^{(0,1)} i_1^{(1)}} (r_i)_{i_1^{(1)}} \frac{\partial \varphi_0}{\partial x^{(0)}_{i_1^{(0,1)}}} \right| = \alpha_i \left| \frac{\partial \varphi_0}{\partial r_i^{(0)}} \right| \\
&\leq \alpha_i \| \widetilde{A} \|_{\max} \| A_{T_0} \|_{\max} \sum_{i_1^{(1)}=1}^{d} \sum_{i_1^{(0,1)}=1}^{d} \left| \frac{\partial \varphi_0}{\partial x^{(0)}_{i_1^{(0,1)}}} \right| ,
\end{aligned}
$$

and for $1 \leq i, j \leq d$,

$$
\begin{aligned}
\left|\frac{\partial^2 \hat{\varphi}}{\partial \hat{x}_i \partial \hat{x}_j}\right| &= \left| \sum_{i_1^{(1)}, j_1^{(1)}=1}^{d} \sum_{i_1^{(0,1)}, j_1^{(0,1)}=1}^{d} \alpha_i \alpha_j [\widetilde{A}]_{i_1^{(1)} i} [\widetilde{A}]_{j_1^{(1)} j} \right. \\
&\quad \left. [A_{T_0}]_{i_1^{(0,1)} i_1^{(1)}} [A_{T_0}]_{j_1^{(0,1)} j_1^{(1)}} \frac{\partial^2 \varphi_0}{\partial x^{(0)}_{i_1^{(0,1)}} \partial x^{(0)}_{j_1^{(0,1)}}} \right| = \alpha_i \alpha_j \left| \frac{\partial^2 \varphi_0}{\partial r_i^{(0)} \partial r_j^{(0)}} \right| \\
&\le \alpha_i \alpha_j \sum_{j_1^{(1)}=1}^{d} |[\widetilde{A}]_{j_1^{(1)} j}| \left| \sum_{j_1^{(0,1)}=1}^{d} [A_{T_0}]_{j_1^{(0,1)} j_1^{(1)}} \frac{\partial^2 \varphi_0}{\partial r_i^{(0)} \partial x^{(0)}_{j_1^{(0,1)}}} \right| \\
&\le \alpha_i \alpha_j \|\widetilde{A}\|_{\max} \|A_{T_0}\|_{\max} \sum_{j_1^{(0,1)}=1}^{d} \left| \frac{\partial^2 \varphi_0}{\partial r_i^{(0)} \partial x^{(0)}_{j_1^{(0,1)}}} \right| \\
&\le \alpha_i \alpha_j \|\widetilde{A}\|_{\max}^2 \|A_{T_0}\|_{\max}^2 \sum_{i_1^{(0,1)}, j_1^{(0,1)}=1}^{d} \left| \frac{\partial^2 \varphi_0}{\partial x^{(0)}_{i_1^{(0,1)}} \partial x^{(0)}_{j_1^{(0,1)}}} \right|.
\end{aligned}
$$

If Assumption 1 is not imposed, the estimates corresponding to Lemma 5 are as follows.

Lemma 6 *Let $m \in \mathbb{N}_0$, $\ell \in \mathbb{N}_0$ with $\ell \ge m$, and $p \in [0, \infty]$. Let $\beta := (\beta_1, \ldots, \beta_d) \in \mathbb{N}_0^d$ and $\gamma := (\gamma_1, \ldots, \gamma_d) \in \mathbb{N}_0^d$ be multi-indices with $|\beta| = m$ and $|\gamma| = \ell - m$. Then, for any $\hat{\varphi} \in W^{\ell,p}(\widehat{T})$ with $\tilde{\varphi} = \hat{\varphi} \circ \widehat{\Phi}^{-1}$, $\varphi = \tilde{\varphi} \circ \widetilde{\Phi}^{-1}$, and $\varphi_0 = \varphi \circ \Phi_{T_0}^{-1}$, it holds that*

$$
\|\partial^\beta \partial^\gamma \hat{\varphi}\|_{L^p(\widehat{T})} \le C_3^{SA} |\det(A_T)|^{-\frac{1}{p}} \|\widetilde{A}\|_2^m \alpha^\beta \sum_{|\varepsilon| = |\gamma|} \alpha^\varepsilon |\partial^\varepsilon_{r^{(0)}} \varphi_0|_{W^{m,p}(T_0)}, \tag{39}
$$

where C_3^{SA} is a constant that is independent of T_0 and $\widetilde{T}$. Here, for $p = \infty$ and any positive real x, $x^{-\frac{1}{p}} = 1$.

Proof We follow the proof of Lemma 5. Let $p \in [1, \infty)$. Because the space $\mathcal{C}^\ell(\widehat{T})$ is dense in the space $W^{\ell,p}(\widehat{T})$, we show that Eq. (39) holds for $\hat{\varphi} \in \mathcal{C}^\ell(\widehat{T})$ with $\tilde{\varphi} = \hat{\varphi} \circ \widehat{\Phi}^{-1}$, $\varphi = \tilde{\varphi} \circ \widetilde{\Phi}^{-1}$, and $\varphi_0 = \varphi \circ \Phi_{T_0}^{-1}$. For $1 \le i, k \le d$,

$$
\left|\partial^{\beta+\gamma} \hat{\varphi}\right| \le c \alpha^\beta \|\widetilde{A}\|_{\max}^{|\beta|} \|A_{T_0}\|_{\max}^{|\beta|} \sum_{|\delta| = |\beta|} \sum_{|\varepsilon| = |\gamma|} \alpha^\varepsilon \left| \partial^\delta \partial^\varepsilon_{r^{(0)}} \varphi_0 \right|.
$$

Using Eqs. (22c) and (24), we obtain Eq. (39) for $p \in [1, \infty]$ by an argument analogous to that used for Lemma 5. □

4 Remarks on anisotropic interpolation analysis

We use the following Bramble–Hilbert-type lemma on anisotropic meshes proposed in [1, Lemma 2.1].

Lemma 7 *Let $D \subset \mathbb{R}^d$ with $d \in \{2,3\}$ be a connected open set that is star-shaped with respect to a ball B. Let γ be a multi-index with $m := |\gamma|$ and $\varphi \in L^1(D)$ be a function with $\partial^\gamma \varphi \in W^{\ell-m,p}(D)$, where $\ell \in \mathbb{N}$, $m \in \mathbb{N}_0$, $0 \le m \le \ell$, and $p \in [1,\infty]$. Then, it holds that*

$$\|\partial^\gamma(\varphi - Q^{(\ell)}\varphi)\|_{W^{\ell-m,p}(D)} \le C^{BH}|\partial^\gamma \varphi|_{W^{\ell-m,p}(D)}, \tag{40}$$

where C^{BH} depends only on d, ℓ, $\mathrm{diam}D$, and $\mathrm{diam}B$, and $Q^{(\ell)}\varphi$ is defined as

$$(Q^{(\ell)}\varphi)(x) := \sum_{|\delta|\le \ell-1} \int_B \eta(y)(\partial^\delta \varphi)(y)\frac{(x-y)^\delta}{\delta!}dy \in \mathcal{P}^{\ell-1}, \tag{41}$$

where $\eta \in \mathcal{C}_0^\infty(B)$ is a given function with $\int_B \eta dx = 1$.

As explained in the Introduction, there exist some mistakes in the proof of Theorem 2 of [16], and the statement is not valid in its original form. To clarify the following description, we explain the errors in the proof. Let $\widehat{T} \subset \mathbb{R}^2$ be the reference element defined in Sect. 2.1.1. We set $k = m = 1$, $\ell = 2$, and $p = 2$. For $\hat{\varphi} \in H^2(\widehat{T})$, we set $\tilde{\varphi} := \hat{\varphi} \circ \widehat{\Phi}^{-1}$ and $\varphi := \tilde{\varphi} \circ \widetilde{\Phi}^{-1}$. Inequality (29) yields

$$|\varphi - I_T\varphi|_{H^1(T)} \le c|\det(A_T)|^{\frac{1}{2}} \|\widetilde{A}^{-1}\|_2 \left(\sum_{i=1}^{2} \alpha_i^{-2} \|\partial_{\hat{x}_i}(\hat{\varphi} - I_{\widehat{T}}\hat{\varphi})\|^2_{L^2(\widehat{T})} \right)^{\frac{1}{2}}. \tag{42}$$

The coefficient α_i^{-2} appears on the right-hand side of Eq. (42). In [16, Theorem 2], we wrongly claimed that α_i^{-2} could be canceled out. In fact, a further assumption is required for this. Using Eq. (41) and the triangle inequality, we have

$$\|\partial_{\hat{x}_i}(\hat{\varphi} - I_{\widehat{T}}\hat{\varphi})\|^2_{L^2(\widehat{T})} \le 2\|\partial_{\hat{x}_i}(\hat{\varphi} - Q^{(2)}\hat{\varphi})\|^2_{L^2(\widehat{T})} + 2\|\partial_{\hat{x}_i}(Q^{(2)}\hat{\varphi} - I_{\widehat{T}}\hat{\varphi})\|^2_{L^2(\widehat{T})}.$$

We use inequality (40) to obtain the target inequality [16, Theorem 2]. To this end, we have to show that

$$\|\partial_{\hat{x}_i}(Q^{(2)}\hat{\varphi} - I_{\widehat{T}}\hat{\varphi})\|_{L^2(\widehat{T})} \le c\|\partial_{\hat{x}_i}(\hat{\varphi} - Q^{(2)}\hat{\varphi})\|_{H^1(\widehat{T})}. \tag{43}$$

However, this is unlikely to hold because Eqs. (14) and (16) yield

$$\begin{aligned}\|\partial_{\hat{x}_i}(Q^{(2)}\hat{\varphi} - I_{\widehat{T}}\hat{\varphi})\|_{L^2(\widehat{T})} &= \|\partial_{\hat{x}_i}(I_{\widehat{T}}(Q^{(2)}\hat{\varphi}) - I_{\widehat{T}}\hat{\varphi})\|_{L^2(\widehat{T})} \\ &\le c\|Q^{(2)}\hat{\varphi} - \hat{\varphi}\|_{H^2(\widehat{T})} \le c|\hat{\varphi}|_{H^2(\widehat{T})}.\end{aligned}$$

Using the classical scaling argument (see [12, Lemma 1.101]), we have

$$|\hat{\varphi}|_{H^2(\widehat{T})} \leq c|\det(A_T)|^{-\frac{1}{2}}\|\widetilde{A}\|_2|\varphi|_{H^2(T)},$$

which does not include the quantity α_i. Therefore, the quantity α_i^{-1} in Eq. (42) remains. Thus, the proof of [16, Theorem 2] is incorrect.

To overcome this problem, we use some results from previous studies [1, 2]. That is, we assume that there exists a linear functional $\mathscr{F}_1$ such that

$$\begin{aligned}
&\mathscr{F}_1 \in H^1(\widehat{T})', \\
&\mathscr{F}_1(\partial_{\hat{x}_i}(\hat{\varphi} - I_{\widehat{T}}\hat{\varphi})) = 0 \quad i = 1,2, \quad \forall \hat{\varphi} \in \mathcal{C}(\widehat{T}) : \ \partial_{\hat{x}_i}\hat{\varphi} \in H^1(\widehat{T}), \\
&\hat{\eta} \in \mathcal{P}^1, \quad \mathscr{F}_1(\partial_{\hat{x}_i}\hat{\eta}) = 0 \quad i = 1,2, \quad \Rightarrow \quad \partial_{\hat{x}_i}\hat{\eta} = 0.
\end{aligned}$$

Because the polynomial spaces are finite-dimensional, all norms are equivalent; i.e., because $|\mathscr{F}_1(\partial_{\hat{x}_i}(\hat{\eta} - I_{\widehat{T}}\hat{\varphi}))|$ $(i = 1,2)$ is a norm on $\mathcal{P}^0$, we have that, for $i = 1,2$,

$$\begin{aligned}
\|\partial_{\hat{x}_i}(\hat{\eta} - I_{\widehat{T}}\hat{\varphi})\|_{L^2(\widehat{T})} &\leq c|\mathscr{F}_1(\partial_{\hat{x}_i}(\hat{\eta} - I_{\widehat{T}}\hat{\varphi}))| = c|\mathscr{F}_1(\partial_{\hat{x}_i}(\hat{\eta} - \hat{\varphi}))| \\
&\leq c\|\partial_{\hat{x}_i}(\hat{\eta} - \hat{\varphi})\|_{H^1(\widehat{T})}.
\end{aligned}$$

Setting $\hat{\eta} := Q^{(2)}\hat{\varphi}$, we obtain Eq. (43). Using inequality (40) yields

$$\|\partial_{\hat{x}_i}(\hat{\varphi} - I_{\widehat{T}}\hat{\varphi})\|^2_{L^2(\widehat{T})} \leq c|\partial_{\hat{x}_i}\hat{\varphi}|^2_{H^1(\widehat{T})},$$

and so inequality (42) together with Eq. (25) can be written as

$$|\varphi - I_T\varphi|_{H^1(T)} \leq c|\det(A_T)|^{\frac{1}{2}}\|\widetilde{A}^{-1}\|_2 \sum_{i,j=1}^{2} \alpha_i^{-1}\|\partial_{\hat{x}_i}\partial_{\hat{x}_j}\hat{\varphi}\|_{L^2(\widehat{T})}. \tag{44}$$

Inequality (37) yields

$$\|\partial_{\hat{x}_i}\partial_{\hat{x}_j}\hat{\varphi}\|_{L^2(\widehat{T})} \leq c|\det(A_T)|^{-\frac{1}{2}}\|\widetilde{A}\|_2\alpha_i \sum_{n=1}^{2} \mathscr{H}_n \left|\frac{\partial\varphi}{\partial x_n}\right|_{H^1(T)}. \tag{45}$$

Therefore, the quantity α_i^{-1} in Eq. (44) and the quantity α_i in Eq. (45) cancel out.

5 Classical interpolation error estimates

The following embedding results hold.

Theorem *Let $d \geq 2$, $s > 0$, and $p \in [1,\infty]$. Let $D \subset \mathbb{R}^d$ be a bounded open subset of $\mathbb{R}^d$. If D is a Lipschitz set, we have that*

$$W^{s,p}(D) \hookrightarrow \begin{cases} L^q(D) & \forall q \in [p, \frac{pd}{d-sp}] \text{ if } sp < d, \\ L^q(D) & \forall q \in [p,\infty), \text{if } sp = d, \\ L^\infty(D) \cap \mathcal{C}^{0,\xi}(\overline{D}) & \xi = 1 - \frac{d}{sp} \text{ if } sp > d. \end{cases} \tag{46}$$

Furthermore,

$$W^{s,p}(D) \hookrightarrow L^\infty(D) \cap \mathcal{C}^0(\overline{D}) \quad (\text{case } s = d \text{ and } p = 1). \tag{47}$$

Proof See, for example, [12, Corollary B.43, Theorem B.40], [13, Theorem 2.31], and the references therein. □

Remark 4 Let $s > 0$ and $p \in [1, \infty]$ be such that

$$s > \frac{d}{p} \quad \text{if } p > 1, \quad s \geq d \quad \text{if } p = 1.$$

Then, it holds that $W^{s,p}(D) \hookrightarrow \mathcal{C}^0(\overline{D})$.

Using the new geometric parameter H_{T_0}, it is possible to deduce the classical interpolation error estimates; e.g., see [12, Theorem 1.103] and [13, Theorem 11.13].

Theorem A *Let* $1 \leq p \leq \infty$ *and assume that there exists a nonnegative integer* k *such that*

$$\mathcal{P}^k \subset \widehat{P} \subset W^{k+1,p}(\widehat{T}) \subset V(\widehat{T}).$$

Let ℓ $(0 \leq \ell \leq k)$ *be such that* $W^{\ell+1,p}(\widehat{T}) \subset V(\widehat{T})$ *with the continuous embedding. Furthermore, assume that* $\ell, m \in \mathbb{N} \cup \{0\}$ *and* $p, q \in [1, \infty]$ *such that* $0 \leq m \leq \ell + 1$ *and*

$$W^{\ell+1,p}(\widehat{T}) \hookrightarrow W^{m,q}(\widehat{T}). \tag{48}$$

Then, for any $\varphi_0 \in W^{\ell+1,p}(T_0)$, *it holds that*

$$|\varphi_0 - I_{T_0}\varphi_0|_{W^{m,q}(T_0)} \leq C_*^I |T_0|^{\frac{1}{q}-\frac{1}{p}} \left(\frac{\alpha_{\max}}{\alpha_{\min}}\right)^m \left(\frac{H_{T_0}}{h_{T_0}}\right)^m h_{T_0}^{\ell+1-m} |\varphi_0|_{W^{\ell+1,p}(T_0)}, \tag{49}$$

where C_*^I *is a positive constant that is independent of* h_T *and* H_T, *and the parameters* $\alpha_{\max}$ *and* $\alpha_{\min}$ *are defined as*

$$\alpha_{\max} := \max\{\alpha_1, \ldots, \alpha_d\}, \quad \alpha_{\min} := \min\{\alpha_1, \ldots, \alpha_d\}. \tag{50}$$

Proof Let $\hat{\varphi} \in W^{\ell+1,p}(\widehat{T})$. Because $0 \leq \ell \leq k$, $\mathcal{P}^\ell \subset \mathcal{P}^k \subset \widehat{P}$. Therefore, for any $\hat{\eta} \in \mathcal{P}^\ell$, we have $I_{\widehat{T}}\hat{\eta} = \hat{\eta}$. Using Eqs. (16) and (48), we obtain

$$\begin{aligned} |\hat{\varphi} - I_{\widehat{T}}\hat{\varphi}|_{W^{m,q}(\widehat{T})} &\leq |\hat{\varphi} - \hat{\eta}|_{W^{m,q}(\widehat{T})} + |I_{\widehat{T}}(\hat{\eta} - \hat{\varphi})|_{W^{m,q}(\widehat{T})} \\ &\leq c\|\hat{\varphi} - \hat{\eta}\|_{W^{\ell+1,p}(\widehat{T})}, \end{aligned}$$

where we have used the stability of the interpolation operator $I_{\widehat{T}}$; i.e.,

$$|I_{\widehat{T}}(\hat{\eta} - \hat{\varphi})|_{W^{m,q}(\widehat{T})} \leq \sum_{i=1}^{n_0} |\hat{\chi}_i(\hat{\eta} - \hat{\varphi})||\hat{\theta}_i|_{W^{m,q}(\widehat{T})} \leq c\|\hat{\eta} - \tilde{\varphi}\|_{W^{\ell+1,p}(\widehat{T})}.$$

Using the classic Bramble–Hilbert-type lemma (e.g., [7, Lemma 4.3.8]), we obtain

$$|\tilde{\varphi} - I_{\widehat{T}}\hat{\varphi}|_{W^{m,q}(\widehat{T})} \leq c \inf_{\hat{\eta}\in\mathcal{P}^{\ell}} \|\hat{\eta} - \tilde{\varphi}\|_{W^{\ell+1,p}(\widehat{T})} \leq c|\hat{\varphi}|_{W^{\ell+1,p}(\widehat{T})}. \tag{51}$$

Inequalities (26), (29), (25), and (51) yield

$$\begin{aligned}
|\varphi_0 - I_{T_0}\varphi_0|_{W^{m,q}(T_0)} &\leq c|\varphi - I_T\varphi|_{W^{m,q}(T)} \\
&\leq c|\det(A_T)|^{\frac{1}{q}} \|\widetilde{A}^{-1}\|_2^m \left(\sum_{|\beta|=m} (\alpha^{-\beta})^q \|\partial^\beta(\hat{\varphi} - I_{\widehat{T}}\hat{\varphi})\|^q_{L^q(\widehat{T})} \right)^{1/q} \\
&\leq c|\det(A_T)|^{\frac{1}{q}} \|\widetilde{A}^{-1}\|_2^m \max\{\alpha_1^{-1}, \dots, \alpha_d^{-1}\}^{|\beta|} |\tilde{\varphi} - I_{\widehat{T}}\hat{\varphi}|_{W^{m,q}(\widehat{T})} \\
&\leq c|\det(A_T)|^{\frac{1}{q}} \|\widetilde{A}^{-1}\|_2^m \alpha_{\min}^{-|\beta|} |\hat{\varphi}|_{W^{\ell+1,p}(\widehat{T})}.
\end{aligned} \tag{52}$$

Using inequalities (25) and (39) together with Eq. (22c), we have

$$\begin{aligned}
&|\hat{\varphi}|_{W^{\ell+1,p}(\widehat{T})} \\
&\leq \sum_{|\gamma|=\ell+1-m} \sum_{|\beta|=m} \|\partial^\beta \partial^\gamma \hat{\varphi}\|_{L^p(\widehat{T})} \\
&\leq c|\det(A_T)|^{-\frac{1}{p}} \|\widetilde{A}\|_2^m \sum_{|\gamma|=\ell+1-m} \sum_{|\beta|=m} \alpha^\beta \sum_{|\epsilon|=|\gamma|} \alpha^\varepsilon |\partial^\epsilon_{r^{(0)}} \varphi_0|_{W^{m,p}(T_0)} \\
&\leq c|\det(A_T)|^{-\frac{1}{p}} \|\widetilde{A}\|_2^m \max\{\alpha_1, \dots, \alpha_d\}^{|\beta|} h_{T_0}^{\ell+1-m} |\varphi_0|_{W^{\ell+1,p}(T_0)} \\
&\leq c|\det(A_T)|^{-\frac{1}{p}} \|\widetilde{A}\|_2^m \alpha_{\max}^{|\beta|} h_{T_0}^{\ell+1-m} |\varphi_0|_{W^{\ell+1,p}(T_0)}.
\end{aligned} \tag{53}$$

From Eqs. (52) and (53) together with Eq. (23), we have the desired estimate (49). □

6 Anisotropic interpolation error estimates

6.1 Main theorem

Theorem A can be applied to standard isotropic elements as well as some classes of anisotropic elements. If we are concerned with anisotropic elements, it is desirable to remove the quantity $\alpha_{\max}/\alpha_{\min}$ from estimate (49). To this end, we employ the approach described in [1] and consider the case of a finite element with $V(\widehat{T}) := \mathcal{C}(\widehat{T})$ and $\widehat{P} := \mathcal{P}^k(\widehat{T})$ (Theorem B). However, one needs stronger assumptions to obtain the optimal estimate. When using finite elements that do not satisfy the assumptions of Theorem B (e.g., $\mathcal{P}^1$-bubble finite element), we have to use Theorem A. In these cases, it may not be possible to obtain optimal order estimates if the shape-regularity condition is violated.

Theorem B *Let $\{\widehat{T}, \widehat{P}, \widehat{\Sigma}\}$ be a finite element with the normed vector space $V(\widehat{T}) := \mathcal{C}(\widehat{T})$ and $\widehat{P} := \mathcal{P}^k(\widehat{T})$ with $k \geq 1$. Let $I_{\widehat{T}} : V(\widehat{T}) \to \widehat{P}$ be a linear operator. Fix $\ell \in \mathbb{N}$, $m \in \mathbb{N}_0$, and $p, q \in [1, \infty]$ such that $0 \leq m \leq \ell \leq k+1$, $\ell - m \geq 1$, and assume that the embeddings (46) and (47) with $s := \ell - m$ hold. Let β be a multi-index with $|\beta| = m$. We set $j := \dim(\partial^\beta \mathcal{P}^k)$. Assume that there exist linear functionals $\mathscr{F}_i$, $i = 1, \ldots, j$, such that*

$$\mathscr{F}_i \in W^{\ell-m,p}(\widehat{T})', \quad \forall i = 1, \ldots, j, \tag{54}$$

$$\mathscr{F}_i(\partial^\beta(\hat{\varphi} - I_{\widehat{T}}\hat{\varphi})) = 0 \quad \forall i = 1, \ldots, j, \quad \forall \hat{\varphi} \in \mathcal{C}(\widehat{T}) : \ \partial^\beta \hat{\varphi} \in W^{\ell-m,p}(\widehat{T}), \tag{55}$$

$$\hat{\eta} \in \mathcal{P}^k, \quad \mathscr{F}_i(\partial^\beta \hat{\eta}) = 0 \quad \forall i = 1, \ldots, j \quad \Rightarrow \quad \partial^\beta \hat{\eta} = 0. \tag{56}$$

For any $\hat{\varphi} \in W^{\ell,p}(\widehat{T}) \cap \mathcal{C}(\widehat{T})$, we set $\varphi_0 := \hat{\varphi} \circ \Phi^{-1}$. If Assumption 1 is imposed, it holds that

$$\begin{aligned} &|\varphi_0 - I_{T_0}\varphi_0|_{W^{m,q}(T_0)} \\ &\quad \leq C_1^{TB} |T_0|^{\frac{1}{q}-\frac{1}{p}} \left(\frac{H_{T_0}}{h_{T_0}} \right)^m \sum_{|\gamma|=\ell-m} \mathscr{H}^\gamma |\partial^\gamma (\varphi_0 \circ \Phi_{T_0})|_{W^{m,p}(\Phi_{T_0}^{-1}(T_0))}, \end{aligned} \tag{57}$$

where C_1^{TB} is a positive constant that is independent of h_{T_0} and H_{T_0}. Furthermore, if Assumption 1 is not imposed, it holds that

$$\begin{aligned} &|\varphi_0 - I_{T_0}\varphi_0|_{W^{m,q}(T_0)} \\ &\quad \leq C_2^{TB} |T_0|^{\frac{1}{q}-\frac{1}{p}} \left(\frac{H_{T_0}}{h_{T_0}} \right)^m \sum_{|\gamma|=\ell-m} \alpha^\gamma |\partial_{r^{(0)}}^\gamma \varphi_0|_{W^{m,p}(T_0)}, \end{aligned} \tag{58}$$

where C_2^{TB} is a positive constant that is independent of h_{T_0} and H_{T_0}.

Proof The introduction of the functionals $\mathscr{F}_i$ follows from [1]. In fact, under the same assumptions as made in Theorem B, we have (see [1, Lemma 2.2])

$$\|\partial^\beta(\hat{\varphi} - I_{\widehat{T}}\hat{\varphi})\|_{L^q(\widehat{T})} \leq C^B |\partial^\beta \hat{\varphi}|_{W^{\ell-m,p}(\widehat{T})}, \tag{59}$$

where $|\beta| = m$, $\hat{\varphi} \in \mathcal{C}(\widehat{T})$, and $\partial^\beta \hat{\varphi} \in W^{\ell-m,p}(\widehat{T})$.

Inequalities (26), (29), (25), and (59) yield

$$
\begin{aligned}
|\varphi_0 - I_{T_0}\varphi_0|_{W^{m,q}(T_0)} &\le c|\varphi - I_T\varphi|_{W^{m,q}(T)} \\
&\le c|\det(A_T)|^{\frac{1}{q}} \|\widetilde{A}^{-1}\|_2^m \left(\sum_{|\beta|=m} (\alpha^{-\beta})^q \|\partial^\beta(\hat{\varphi} - I_{\widehat{T}}\hat{\varphi})\|_{L^q(\widehat{T})}^q \right)^{1/q} \\
&\le c|\det(A_T)|^{\frac{1}{q}} \|\widetilde{A}^{-1}\|_2^m \sum_{|\beta|=m} (\alpha^{-\beta}) \|\partial^\beta(\hat{\varphi} - I_{\widehat{T}}\hat{\varphi})\|_{L^q(\widehat{T})} \\
&\le c|\det(A_T)|^{\frac{1}{q}} \|\widetilde{A}^{-1}\|_2^m \sum_{|\beta|=m} (\alpha^{-\beta}) |\partial^\beta \hat{\varphi}|_{W^{\ell-m,p}(\widehat{T})}.
\end{aligned}
\tag{60}
$$

If Assumption 1 is imposed, then using inequalities (25) and (37) leads to

$$
\begin{aligned}
&\sum_{|\beta|=m} (\alpha^{-\beta}) |\partial^\beta \hat{\varphi}|_{W^{\ell-m,p}(\widehat{T})} \\
&\quad\le \sum_{|\gamma|=\ell-m} \sum_{|\beta|=m} (\alpha^{-\beta}) \|\partial^{\beta+\gamma}\hat{\varphi}\|_{L^p(\widehat{T})} \\
&\quad\le c|\det(A_T)|^{-\frac{1}{p}} \|\widetilde{A}\|_2^m \sum_{|\gamma|=\ell-m} \sum_{|\beta|=m} (\alpha^{-\beta})\alpha^\beta \sum_{|\epsilon|=|\gamma|} \mathscr{H}^\epsilon |\partial^\epsilon \varphi|_{W^{m,p}(T)} \\
&\quad\le c|\det(A_T)|^{-\frac{1}{p}} \|\widetilde{A}\|_2^m \sum_{|\epsilon|=\ell-m} \mathscr{H}^\epsilon |\partial^\epsilon \varphi|_{W^{m,p}(T)}.
\end{aligned}
\tag{61}
$$

From Eqs. (60) and (61) together with Eqs. (22) and (23), we have the desired estimate (57) using $T = \Phi_{T_0}^{-1}(T_0)$ and $\varphi = \varphi_0 \circ \Phi_{T_0}$.

If Assumption 1 is not imposed, then an analogous argument using inequality (39) instead of (37) yields estimate (58). □

Example 1 Specific finite elements satisfying conditions (54), (55), and (56) are given in [2] and [1]; see also Sect. 6.2.

Remark 5 Finite elements that do not satisfy conditions (54), (55), and (56) can be found in [2, Table 3]; e.g., the $\mathcal{P}^1$-bubble finite element and the $\mathcal{P}^3$ Hermite finite element. In these cases, Theorem A can be applied.

6.2 Examples satisfying conditions (54), (55), and (56) in Theorem B

Corollary 1 *Let $\{\widehat{T}, \widehat{P}, \widehat{\Sigma}\}$ be the Lagrange finite element with $V(\widehat{T}) := \mathcal{C}(\widehat{T})$ and $\widehat{P} := \mathcal{P}^k(\widehat{T})$ for $k \ge 1$. Let $I_T : V(T) \to P$ be the corresponding local Lagrange interpolation operator. Let $m \in \mathbb{N}_0$, $\ell \in \mathbb{N}$, and $p \in \mathbb{R}$ be such that $0 \le m \le \ell \le k+1$ and*

$$d=2: \begin{cases} p \in (2,\infty] & \text{if } m=0, \ell=1, \\ p \in [1,\infty] & \text{if } m=0, \ell \ge 2 \text{ or } m \ge 1, \ell - m \ge 1, \end{cases}$$

$$d=3: \begin{cases} p \in \left(\frac{3}{\ell},\infty\right] & \text{if } m=0, \ell=1,2, \\ p \in (2,\infty] & \text{if } m \ge 1, \ell - m = 1, \\ p \in [1,\infty] & \text{if } m=0, \ell \ge 3 \text{ or } m \ge 1, \ell - m \ge 2. \end{cases}$$

We set $q \in [1,\infty]$ such that $W^{\ell-m,p}(\widehat{T}) \hookrightarrow L^q(\widehat{T})$. Then, for all $\hat{\varphi} \in W^{\ell,p}(\widehat{T})$ with $\varphi_0 := \hat{\varphi} \circ \Phi^{-1}$, we recover Eq. (57) if Assumption 1 is imposed, and Eq. (58) holds.

Furthermore, for any $\hat{\varphi} \in \mathcal{C}(\widehat{T})$ with $\varphi_0 := \hat{\varphi} \circ \Phi^{-1}$, it holds that

$$\|\varphi_0 - I_{T_0}\varphi_0\|_{L^\infty(T_0)} \le c\|\varphi_0\|_{L^\infty(T_0)}.$$

Proof The existence of functionals satisfying Eqs. (54), (55), and (56) is shown in the proof of [1, Lemma 2.4] for $d=2$ and in the proof of [1, Lemma 2.6] for $d=3$. Inequality (59) then holds. This implies that estimates (57) and (58) hold. ☐

Setting $V(T) := \mathcal{C}(T)$, we define the nodal Crouzeix–Raviart interpolation operators as

$$I_T^{CR,S} : V(T) \ni \varphi \mapsto I_T^{CR,S}\varphi := \sum_{i=1}^{d+1} \varphi(x_{F_i})\theta_i \in \mathcal{P}^1.$$

Corollary 2 *Let $\{\widehat{T}, \widehat{P}, \widehat{\Sigma}\}$ be the Crouzeix–Raviart finite element with $V(\widehat{T}) := \mathcal{C}(\widehat{T})$ and $\widehat{P} := \mathcal{P}^1(\widehat{T})$. Set $I_T; = I_T^{CR,S}$. Let $m \in \mathbb{N}_0$, $\ell \in \mathbb{N}$, and $p \in \mathbb{R}$ be such that*

$$d=2: \begin{cases} p \in (2,\infty] & \text{if } m=0, \ell=1, \\ p \in [1,\infty] & \text{if } m=0, \ell=2 \text{ or } m=1, \ell=2, \end{cases}$$

$$d=3: \begin{cases} p \in \left(\frac{3}{\ell},\infty\right] & \text{if } m=0, \ell=1,2, \\ p \in (2,\infty] & \text{if } m=1, \ell=2. \end{cases}$$

Set $q \in [1,\infty]$ such that $W^{\ell-m,p}(\widehat{T}) \hookrightarrow L^q(\widehat{T})$. Then, for all $\hat{\varphi} \in W^{\ell,p}(\widehat{T})$ with $\varphi_0 := \hat{\varphi} \circ \Phi^{-1}$, we recover Eq. (57) if Assumption 1 is imposed, and Eq. (58) holds.

Furthermore, for any $\hat{\varphi} \in \mathcal{C}(\widehat{T})$ with $\varphi_0 := \hat{\varphi} \circ \Phi^{-1}$, it holds that

$$\|\varphi_0 - I_{T_0}\varphi_0\|_{L^\infty(T_0)} \le c\|\varphi_0\|_{L^\infty(T_0)}.$$

Proof For $k=1$, we only introduce functionals $\mathscr{F}_i$ satisfying Eqs. (54)–(56) in Theorem B for each ℓ and m.

Let $m=0$. From the Sobolev embedding theorem, we have $W^{\ell,p}(\widehat{T}) \subset \mathcal{C}^0(\widehat{T})$ with $1 < p \le \infty$, $d < \ell p$ or $p=1$, $d \le \ell$. Under this condition, we use

$$\mathscr{F}_i(\hat{\varphi}) := \hat{\varphi}(\hat{x}_{\widehat{F}_i}), \quad \hat{\varphi} \in W^{\ell,p}(\widehat{T}), \quad i = 1, \dots, d+1.$$

Let $d = 2$ and $m = 1$ ($\ell = 2$). We set $\beta = (1, 0)$. Then, we have that $j = \dim(\partial^\beta \mathcal{P}^1) = 1$. We consider a functional

$$\mathscr{F}_1(\hat{\varphi}) := \int_0^{\frac{1}{2}} \hat{\varphi}(\hat{x}_1, 1/2) d\hat{x}_1, \quad \hat{\varphi} \in W^{2,p}(\widehat{T}).$$

By an analogous argument, we can set a functional for the case $\beta = (0, 1)$.

Let $d = 3$ and $m = 1$ ($\ell = 2$). We first consider Type (i) in Sect. 2.1.2. That is, the reference element is $\widehat{T} = \mathrm{conv}\{0, e_1, e_2, e_3\}$. Here, $e_1, \dots, e_3 \in \mathbb{R}^3$ form the canonical basis. We set $\beta = (1, 0, 0)$ and consider the functional

$$\mathscr{F}_1(\hat{\varphi}) := \int_0^{\frac{1}{3}} \hat{\varphi}(\hat{x}_1, 1/3, 1/3) d\hat{x}_1, \quad \hat{\varphi} \in W^{2,p}(\widehat{T}).$$

We now consider Type (ii) in Sect. 2.1.2. That is, the reference element is $\widehat{T} = \mathrm{conv}\{0, e_1, e_1 + e_2, e_3\}$. We set $\beta = (1, 0, 0)$ and consider the functional

$$\Phi_1(\hat{\varphi}) := \int_{\frac{1}{3}}^{\frac{2}{3}} \hat{\varphi}(\hat{x}_1, 1/3, 1/3) d\hat{x}_1, \quad \hat{\varphi} \in W^{2,p}(\widehat{T}).$$

By an analogous argument, we can set functionals for cases $\beta = (0, 1, 0), (0, 0, 1)$.

When $m = \ell = 0$ and $p = \infty$, we can easily check that

$$\|\hat{\varphi} - I_{\widehat{T}}^{CR,S} \hat{\varphi}\|_{L^\infty(\widehat{T})} \le c \|\hat{\varphi}\|_{L^\infty(\widehat{T})}.$$

□

7 Concluding remarks

As our concluding remarks, we identify several topics related to the results described in this paper.

7.1 Good elements or not for $d = 2, 3$?

In this subsection, we consider good elements on meshes. Here, we define "good elements" on meshes as those for which there exists some $\gamma_0 > 0$ satisfying Eq. (7). We treat a "Right-angled triangle," "Blade," and "Dagger" for $d = 2$, and a "Spire," "Spear," "Spindle," "Spike," "Splinter," and "Sliver" for $d = 3$, as introduced in [9]. We present the quantities $\alpha_{\max}/\alpha_{\min}$ and H_{T_0}/h_{T_0} for these elements.

7.1.1 Isotropic mesh

We consider the following condition. There exists a constant $\gamma_1 > 0$ such that, for any $\mathbb{T}_h \in \{\mathbb{T}_h\}$ and any simplex $T_0 \in \mathbb{T}_h$, we have

$$|T_0| \geq \gamma_1 h_{T_0}^d. \tag{62}$$

Condition (62) is equivalent to the shape-regularity condition; see [6, Theorem 1].

If geometric condition (62) is satisfied, it holds that

$$\frac{H_{T_0}}{h_{T_0}} \leq \frac{h_{T_0}^d}{|T_0|} \leq \frac{1}{\gamma_1}, \quad \frac{\alpha_{\max}}{\alpha_{\min}} \leq c\frac{h_{T_0}}{\alpha_2} \leq c\frac{h_{T_0}^d}{|T_0|} \leq \frac{c}{\gamma_1}.$$

If $p = q$ in Theorem A, one can obtain the optimal order $h_{T_0}^{\ell+1-m}$. In this case, elements satisfying geometric condition (62) are "good."

7.1.2 Anisotropic mesh: two-dimensional case

Let $S_0 \subset \mathbb{R}^2$ be a triangle. Let $0 < s \ll 1$, $s \in \mathbb{R}$, and $\varepsilon, \delta, \gamma \in \mathbb{R}$. A dagger has one short edge and a blade has no short edge.

Example 2 (Right-angled triangle) Let $S_0 \subset \mathbb{R}^2$ be the simplex with vertices $x_1 := (0,0)^T$, $x_2 := (s,0)^T$, and $x_3 := (0,s^\varepsilon)^T$ with $1 < \varepsilon$; see Fig. 1. Then, we have that $\alpha_1 = s$ and $\alpha_2 = s^\varepsilon$; i.e.,

$$\frac{\alpha_{\max}}{\alpha_{\min}} \leq s^{1-\varepsilon} \to \infty \quad \text{as } s \to 0, \quad \frac{H_{S_0}}{h_{S_0}} = 2.$$

In this case, the element S_0 is "good."

Example 3 (Dagger) Let $S_0 \subset \mathbb{R}^2$ be the simplex with vertices $x_1 := (0,0)^T$, $x_2 := (s,0)^T$, and $x_3 := (s^\delta, s^\varepsilon)^T$ with $1 < \varepsilon < \delta$, see Fig. 2. Then, we have that $\alpha_1 = \sqrt{(s-s^\delta)^2 + s^{2\varepsilon}}$ and $\alpha_2 = \sqrt{s^{2\delta} + s^{2\varepsilon}}$; i.e.,

$$\frac{\alpha_{\max}}{\alpha_{\min}} = \frac{\sqrt{(s-s^\delta)^2 + s^{2\varepsilon}}}{\sqrt{s^{2\delta} + s^{2\varepsilon}}} \leq cs^{1-\varepsilon} \to \infty \quad \text{as } s \to 0,$$
$$\frac{H_{S_0}}{h_{S_0}} = \frac{\sqrt{(s-s^\delta)^2 + s^{2\varepsilon}}\sqrt{s^{2\delta} + s^{2\varepsilon}}}{\frac{1}{2}s^{1+\varepsilon}} \leq c.$$

In this case, the element S_0 is "good."

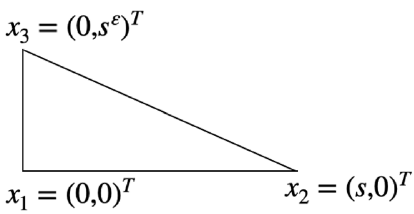


Fig. 1 Example 2: Right-angled triangle

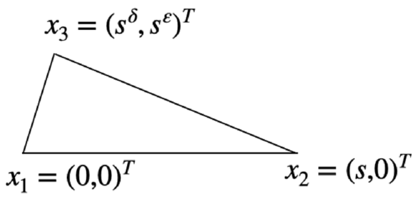


Fig. 2 Examples 3 and 5: Dagger

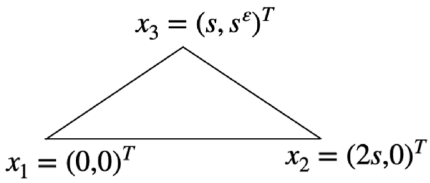


Fig. 3 Example 4: Blade

Remark 6 In the above examples, $\alpha_2 \approx \alpha_2 t = \mathscr{H}_2$ holds. That is, the good element $S_0 \subset \mathbb{R}^2$ satisfies conditions such as $\alpha_2 \approx \alpha_2 t = \mathscr{H}_2$.

Example 4 (Blade) Let $S_0 \subset \mathbb{R}^2$ be the simplex with vertices $x_1 := (0,0)^T$, $x_2 := (2s,0)^T$, and $x_3 := (s, s^\varepsilon)^T$ with $1 < \varepsilon$; see Fig. 3. Then, we have that $\alpha_1 = \alpha_2 = \sqrt{s^2 + s^{2\varepsilon}}$; i.e.,

$$\frac{\alpha_{\max}}{\alpha_{\min}} = 1, \quad \frac{H_{S_0}}{h_{S_0}} = \frac{s^2 + s^{2\varepsilon}}{s^{1+\varepsilon}} \to \infty \quad \text{as } s \to 0.$$

In this case, the element S_0 is "not good."

Example 5 (Dagger) Let $S_0 \subset \mathbb{R}^2$ be the simplex with vertices $x_1 := (0,0)^T$, $x_2 := (s,0)^T$, and $x_3 := (s^\delta, s^\varepsilon)^T$ with $1 < \delta < \varepsilon$; see Fig. 2. Then, we have that $\alpha_1 = \sqrt{(s - s^\delta)^2 + s^{2\varepsilon}}$ and $\alpha_2 = \sqrt{s^{2\delta} + s^{2\varepsilon}}$; i.e.,

$$\frac{\alpha_{\max}}{\alpha_{\min}} = \frac{\sqrt{(s - s^\delta)^2 + s^{2\varepsilon}}}{\sqrt{s^{2\delta} + s^{2\varepsilon}}} \leq c s^{1-\delta} \to \infty \quad \text{as } s \to 0,$$

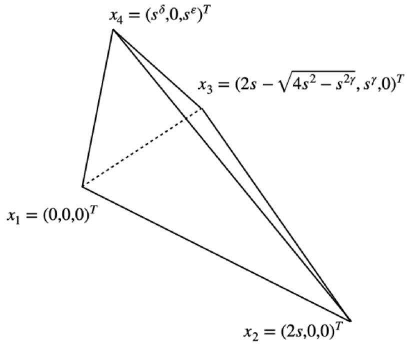
$$\frac{H_{S_0}}{h_{S_0}} = \frac{\sqrt{(s - s^\delta)^2 + s^{2\varepsilon}}\sqrt{s^{2\delta} + s^{2\varepsilon}}}{\frac{1}{2} s^{1+\varepsilon}} \leq c s^{\delta-\varepsilon} \to \infty \quad \text{as } s \to 0.$$

In this case, the element S_0 is "not good."

7.1.3 Anisotropic mesh: three-dimensional case

Example 6 Let $T_0 \subset \mathbb{R}^3$ be a tetrahedron. Let S_0 be the base of T_0; i.e., $S_0 = \triangle x_1 x_2 x_3$. Recall that

Fig. 4 Example 6



$$\frac{H_{T_0}}{h_{T_0}} = \frac{\alpha_1 \alpha_2 \alpha_3}{|T_0|} = \frac{\alpha_1 \alpha_2}{\frac{1}{2}\alpha_1 \alpha_2 t_1} \frac{\alpha_3}{\frac{1}{3}\alpha_3 t_2} \leq \frac{H_{S_0}}{h_{S_0}} \frac{\alpha_3}{\frac{1}{3}\mathscr{H}_3}. \tag{63}$$

If the triangle S_0 is "not good," such as in Examples 4 and 5, the quantity in Eq. (63) may diverge. In the following, we consider the case in which the triangle S_0 is "good."

Assume that there exists a positive constant M such that $\frac{H_{S_0}}{h_{S_0}} \leq M$. For simplicity, we set $x_1 := (0,0,0)^T$, $x_2 := (2s,0,0)^T$, and $x_3 := (2s - \sqrt{4s^2 - s^{2\gamma}}, s^\gamma, 0)^T$ with $1 < \gamma$; see Fig. 4. Then,

$$\alpha_1 = 2s, \ \alpha_2 = \sqrt{\frac{4s^{2\gamma}}{2 + \sqrt{4 - s^{2\gamma-2}}}},$$

and because $\alpha_{\max} \approx cs$,

$$\frac{\alpha_{\max}}{\alpha_{\min}} \leq \frac{cs}{\alpha_2} \leq cs^{1-\gamma} \to \infty \quad \text{as } s \to 0.$$

If we set $x_4 := (s, 0, s^\varepsilon)^T$ with $1 < \varepsilon$, the triangle $\triangle x_1 x_2 x_4$ is the blade (Example 4). Then,

$$\alpha_3 = \sqrt{s^2 + s^{2\varepsilon}}.$$

Thus, we have

$$\frac{H_{T_0}}{h_{T_0}} \leq c \frac{s^{2s+\gamma}}{s^{1+\gamma+\varepsilon}} \leq cs^{1-\varepsilon} \to \infty \quad \text{as } s \to 0.$$

In this case, the element T_0 is "not good."

If we set $x_4 := (s^\delta, 0, s^\varepsilon)^T$ with $1 < \delta < \varepsilon < \gamma$, the triangle $\triangle x_1 x_2 x_4$ is the dagger (Example 5). Then,

$$\alpha_3 = \sqrt{s^{2\delta} + s^{2\varepsilon}}.$$

Thus, we have

$$\frac{H_{T_0}}{h_{T_0}} \leq c\frac{s^{1+\gamma+\delta}}{s^{1+\gamma+\varepsilon}} \leq cs^{\delta-\varepsilon} \to \infty \quad \text{as } s \to 0.$$

In this case, the element T_0 is "not good."

If we set $x_4 := (s^\delta, 0, s^\varepsilon)^T$ with $1 < \varepsilon < \delta < \gamma$, the triangle $\triangle x_1x_2x_4$ is the dagger (Example 3). Then,

$$\alpha_3 = \sqrt{s^{2\delta} + s^{2\varepsilon}}.$$

Thus, we have

$$\frac{H_{T_0}}{h_{T_0}} \leq c\frac{s^{1+\gamma+\varepsilon}}{s^{1+\gamma+\varepsilon}} \leq c.$$

In this case, the element T_0 is "good" and $\alpha_3 \approx \alpha_3 t_2 = \mathscr{H}_3$ holds.

Example 7 In [9], the spire has a cycle of three daggers among its four triangles; see Fig. 5. The splinter has four daggers; see Fig. 9. The spear and spike have two daggers and two blades as triangles; see Figs. 6, 8. The spindle has four blades as triangles; see Fig. 7.

Fig. 5 Spire

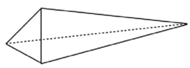


Fig. 6 Spear

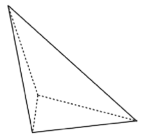
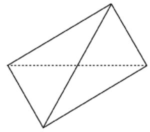


Fig. 7 Spindle



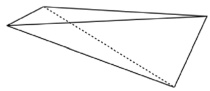


Fig. 8 Spike

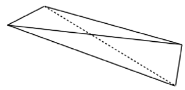


Fig. 9 Splinter

Remark 7 The above examples reveal that the good element $T_0 \subset \mathbb{R}^3$ satisfies conditions such as $\alpha_2 \approx \alpha_2 t_1 = \mathscr{H}_2$ and $\alpha_3 \approx \alpha_3 t_2 = \mathscr{H}_3$.

Example 8 Using an element T_0 called a *Sliver*, we compare the three quantities $\frac{h_{T_0}^3}{|T_0|}$, $\frac{H_{T_0}}{h_{T_0}}$, and $\frac{R_3}{h_{T_0}}$, where the parameter R_3 denotes the circumradius of T_0.

Let $T_0 \subset \mathbb{R}^3$ be the simplex with vertices $x_1 := (s^{\varepsilon_2}, 0, 0)^T$, $x_2 := (-s^{\varepsilon_2}, 0, 0)^T$, $x_3 := (0, -s, s^{\varepsilon_1})^T$, and $x_4 := (0, s, s^{\varepsilon_1})^T$ $(\varepsilon_1, \varepsilon_2 > 1)$, where $s := \frac{1}{N}$, $N \in \mathbb{N}$; see Fig. 10. Let L_i $(1 \le i \le 6)$ be the edges of T_0 with $\alpha_{\min} = L_1 \le L_2 \le \cdots \le L_6 = h_{T_0}$. Recall that $\alpha_{\max} \approx h_{T_0}$ and

$$\frac{\alpha_{\max}}{\alpha_{\min}} \le c \frac{L_6}{L_1}, \quad \frac{H_{T_0}}{h_{T_0}} = \frac{L_1 L_2}{|T_0|} h_{T_0}.$$

In Table 1, the angle between $\triangle x_1 x_2 x_3$ and $\triangle x_1 x_2 x_4$ tends to π as $s \to 0$, and the simplex T_0 is "not good." In Table 2, the angle between $\triangle x_1 x_3 x_4$ and $\triangle x_2 x_3 x_4$ tends

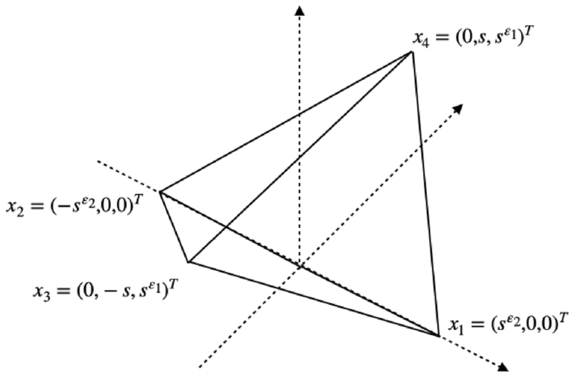


Fig. 10 Example 8: Sliver

Table 1 $h_{T_0}^3/|T_0|$, H_{T_0}/h_{T_0}, and R_3/h_{T_0} $(\varepsilon_1 = 1.5, \varepsilon_2 = 1.0)$

N	s	L_6/L_1	$h_{T_0}^3/\|T_0\|$	H_{T_0}/h_{T_0}	R_3/h_{T_0}
32	3.1250e−02	1.4033	6.7882e+01	3.4471e+01	5.0195e−01
64	1.5625e−02	1.4087	9.6000e+01	4.8375e+01	5.0098e−01
128	7.8125e−03	1.4115	1.3576e+02	6.8147e+01	5.0049e−01

Table 2 $h_{T_0}^3/|T_0|$, H_{T_0}/h_{T_0}, and R_3/h_{T_0} ($\varepsilon_1 = 1.0, \varepsilon_2 = 1.5$)

N	s	L_6/L_1	$h_{T_0}^3/\lvert T_0\rvert$	H_{T_0}/h_{T_0}	R_3/h_{T_0}
32	3.1250e−02	5.6569	6.7882e+01	8.5513	5.0006e−01
64	1.5625e−02	8.0000	9.6000e+01	8.5184	5.0002e−01
128	7.8125e−03	1.1314e+01	1.3576e+02	8.5018	5.0000e−01

Table 3 $h_{T_0}^3/|T_0|$, H_{T_0}/h_{T_0}, and R_3/h_{T_0} ($\varepsilon_1 = 1.5, \varepsilon_2 = 1.5$)

N	s	L_6/L_1	$h_{T_0}^3/\lvert T_0\rvert$	H_{T_0}/h_{T_0}	R_3/h_{T_0}
32	3.1250e−02	5.6569	3.8400e+02	3.4986e+01	1.4170
64	1.5625e−02	8.0000	7.6800e+02	4.8744e+01	2.0010
128	7.8125e−03	1.1314e+01	1.5360e+03	6.8411e+01	2.8288

to 0 as $s \to 0$, and the simplex T_0 is "good." In Table 3, from the numerical results, the simplex T_0 is "not good."

7.2 Effect of the quantity $|T_0|^{\frac{1}{q}-\frac{1}{p}}$ on the interpolation error estimates for $d = 2, 3$

We now consider the effect of the factor $|T_0|^{\frac{1}{q}-\frac{1}{p}}$.

7.2.1 Case in which $q > p$

When $q > p$, the factor may affect the convergence order. In particular, the interpolation error estimate may diverge on anisotropic mesh partitions.

Let $T_0 \subset \mathbb{R}^2$ be the triangle with vertices $x_1 := (0,0)^T$, $x_2 := (s,0)^T$, $x_3 := (0,s^\varepsilon)^T$ for $0 < s \ll 1$, $\varepsilon \geq 1$, $s \in \mathbb{R}$, and $\varepsilon \in \mathbb{R}$; see Fig. 1. Then,

$$\frac{\alpha_{\max}}{\alpha_{\min}} = s^{1-\varepsilon}, \quad |T_0| = \frac{1}{2}s^{1+\varepsilon}.$$

Let $k = 1$, $\ell = 2$, $m = 1$, $q = 2$, and $p \in (1,2)$. Then, $W^{1,p}(T_0) \hookrightarrow L^2(T_0)$ and Theorem B lead to

$$|\varphi_0 - I_{T_0}\varphi_0|_{H^1(T_0)} \leq cs^{-(1+\varepsilon)\frac{2-p}{2p}} \left(s\left|\frac{\partial \varphi_0}{\partial x_1}\right|_{W^{1,p}(T_0)} + s^\varepsilon \left|\frac{\partial \varphi_0}{\partial x_2}\right|_{W^{1,p}(T_0)} \right).$$

When $\varepsilon = 1$ (i.e., an isotropic element), we obtain

$$|\varphi_0 - I_{T_0}\varphi_0|_{H^1(T_0)} \leq ch_{T_0}^{\frac{2(p-1)}{p}} |\varphi_0|_{W^{2,p}(T_0)}, \quad \frac{2(p-1)}{p} > 0.$$

However, when $\varepsilon > 1$ (i.e., an anisotropic element), the estimate may diverge as $s \to 0$. Therefore, if $q > p$, the convergence order of the interpolation operator may deteriorate.

7.2.2 Case in which $q < p$

We consider Theorem B. Let $I_{T_0}^L : \mathcal{C}^0(T_0) \to \mathcal{P}^k$ $(k \in \mathbb{N})$ be the local Lagrange interpolation operator. Let $\varphi_0 \in W^{\ell,\infty}(T_0)$ be such that $\ell \in \mathbb{N}$, $2 \le \ell \le k+1$. Then, for any $m \in \{0, \dots, \ell - 1\}$ and $q \in [1, \infty]$, it holds that

$$|\varphi_0 - I_{T_0}^L \varphi_0|_{W^{m,q}(T_0)} \le c|T_0|^{\frac{1}{q}} \left(\frac{H_{T_0}}{h_{T_0}} \right)^m \sum_{|\gamma|=\ell-m} \mathscr{H}^\gamma |\partial^\gamma (\varphi_0 \circ \Phi_{T_0})|_{W^{m,\infty}(\Phi_{T_0}^{-1}(T_0))}. \tag{64}$$

Therefore, the convergence order is improved by $|T_0|^{\frac{1}{q}}$.

We can perform some numerical tests to confirm this. Let $k = 1$ and

$$\varphi_0(x, y, z) := x^2 + \frac{1}{4} y^2 + z^2.$$

(I) Let $T_0 \subset \mathbb{R}^3$ be the simplex with vertices $x_1 := (0,0,0)^T$, $x_2 := (s,0,0)^T$, $x_3 := (0, s^\varepsilon, 0)^T$, and $x_4 := (0,0,s^\delta)^T$ $(1 < \delta \le \varepsilon)$, and $0 < s \ll 1$, $s \in \mathbb{R}$; see Fig. 11. Then, we have that $\alpha_1 = \sqrt{s^2 + s^{2\varepsilon}}$, $\alpha_2 = s^\varepsilon$, and $\alpha_3 := \sqrt{s^{2\varepsilon} + s^{2\delta}}$; i.e.,

$$\frac{\alpha_{\max}}{\alpha_{\min}} \le c s^{1-\varepsilon}, \quad \frac{H_{T_0}}{h_{T_0}} \le c.$$

From Eq. (64) with $m = 1$, $\ell = 2$, and $q = 2$, because $|T_0| \approx s^{1+\varepsilon+\delta}$, we have the estimate

Fig. 11 Case in which $q < p$, Example (I)

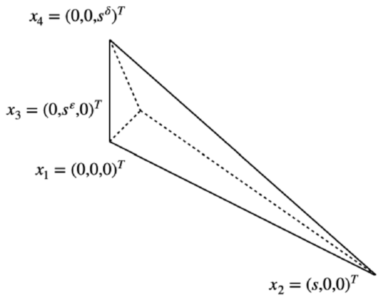


Table 4 Error of the local interpolation operator ($\varepsilon = 3.0, \delta = 2.0$)

N	s	$Err_s^{3.0}(H^1)$	r
64	1.5625e−02	2.4336e−08	
128	7.8125e−03	1.5209e−09	4.00
256	3.9062e−03	9.5053e−11	4.00

$$|\varphi_0 - I_{T_0}^L \varphi_0|_{H^1(T_0)} \le c h_{T_0}^{\frac{3+\varepsilon+\delta}{2}}.$$

Computational results for $\varepsilon = 3.0$ and $\delta = 2.0$ are presented in Table 4.

(II) Let $T_0 \subset \mathbb{R}^3$ be the simplex with vertices $x_1 := (0,0,0)^T$, $x_2 := (s,0,0)^T$, $x_3 := (s/2, s^\varepsilon, 0)^T$, and $x_4 := (0,0,s)^T$ $(1 < \varepsilon \le 6)$ and $0 < s \ll 1$, $s \in \mathbb{R}$; see Fig. 12. Then, we have that $\alpha_1 = s$, $\alpha_2 = \sqrt{s^2/4 + s^{2\varepsilon}}$, and $\alpha_3 := s$; i.e.,

$$\frac{\alpha_{\max}}{\alpha_{\min}} = \frac{s}{\sqrt{s^2/4 + s^{2\varepsilon}}} \le c, \quad \frac{H_{T_0}}{h_{T_0}} \le c s^{1-\varepsilon}.$$

From Eq. (64) with $m = 1$, $\ell = 2$, and $q = 2$, because $|T_0| \approx s^{2+\varepsilon}$, we have the estimate

$$|\varphi_0 - I_{T_0}^L \varphi_0|_{H^1(T_0)} \le c h_{T_0}^{3-\frac{\varepsilon}{2}}.$$

Computational results for $\varepsilon = 3.0, 6.0$ are presented in Table 5.

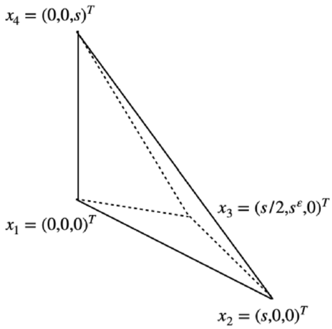


Fig. 12 Case in which $q < p$, Example (II)

Table 5 Error of the local interpolation operator ($\varepsilon = 3.0, 6.0$)

N	s	$Err_s^{3.0}(H^1)$	r	$Err_s^{6.0}(H^1)$	r
64	1.5625e−02	1.9934e−04		1.0206e−01	
128	7.8125e−03	7.0477e−05	1.50	1.0206e−01	0
256	3.9062e−03	2.4917e−05	1.50	1.0206e−01	0

7.3 Inverse inequalities

This section presents some limited results for the inverse inequalities. The results are only stated; the proofs can be found in [15].

Lemma 8 *Let $\widehat{P} := \mathcal{P}^k$ with $k \in \mathbb{N}$. If Assumption 1 is imposed, there exist positive constants $C_i^{IV,d}$, $i = 1, \dots, d$, independent of h_T and T, such that, for all $\varphi_h \in P = \{\hat{\varphi}_h \circ \Phi_T^{-1}; \hat{\varphi}_h \in \widehat{P}\}$,*

$$\left\| \frac{\partial \varphi_h}{\partial x_i} \right\|_{L^q(T)} \leq C_i^{IV,d} |T|^{\frac{1}{q} - \frac{1}{p}} \frac{1}{\mathscr{H}_i} \|\varphi_h\|_{L^p(T)}. \quad i = 1, \dots, d. \tag{65}$$

Remark 8 If Assumption 1 is not imposed, estimate (65) for $i = 3$ is

$$\left\| \frac{\partial \varphi_h}{\partial x_3} \right\|_{L^q(T)} \leq C_3^{IV,3} |T|^{\frac{1}{q} - \frac{1}{p}} \frac{H_T}{h_T} \frac{1}{\mathscr{H}_2} \|\varphi_h\|_{L^p(T)}. \tag{66}$$

This may not be sharp. We leave further arguments for future work.

Theorem D *Let $\widehat{P} := \mathcal{P}^k$ with $k \in \mathbb{N}_0$. Let $\gamma = (\gamma_1, \dots, \gamma_d) \in \mathbb{N}_0^d$ be a multi-index such that $0 \leq |\gamma| \leq k$. If Assumption 1 is imposed, there exists a positive constant C^{IVC}, independent of h_T and T, such that, for all $\varphi_h \in P = \{\hat{\varphi}_h \circ \Phi_T^{-1}; \hat{\varphi}_h \in \widehat{P}\}$,*

$$\|\partial^\gamma \varphi_h\|_{L^q(T)} \leq C^{IVC} |T|^{\frac{1}{q} - \frac{1}{p}} \mathscr{H}^{-\gamma} \|\varphi_h\|_{L^p(T)}. \tag{67}$$

Acknowledgements We would like to thank the anonymous referees and the editor of the journal for the valuable comments.

References

1. Apel, T.: Anisotropic Finite Elements: Local Estimates and Applications. Advances in Numerical Mathematics. Teubner, Stuttgart (1999)
2. Apel, T., Dobrowolski, M.: Anisotropic interpolation with applications to the finite element method. Computing **47**, 277–293 (1992)
3. Apel, T., Eckardt, L., Hauhner, C., Kempf, V.: The maximum angle condition on finite elements: useful or not? PAMM (2021)

4. Babuška, I., Aziz, A.K.: On the angle condition in the finite element method. SIAM J. Numer. Anal. **13**, 214–226 (1976)
5. Babuška, I., Suri, M.: The *p* and *h*–*p* versions of the finite elements method, basic principles and properties. SIAM Rev. **36**, 578–632 (1994)
6. Brandts, J., Korotov, S., Křížek, M.: On the equivalence of regularity criteria for triangular and tetrahedral finite element partitions. Comput. Math. Appl. **55**, 2227–2233 (2008)
7. Brenner, S.C., Scott, L.R.: The Mathematical Theory of Finite Element Methods, 3rd edn. Springer, New York (2008)
8. Chen, S., Shi, D., Zhao, Y.: Anisotropic interpolation and quasi-Wilson element for narrow quadrilateral meshes. IMA J. Numer. Anal. **24**, 77–95 (2004)
9. Cheng, S.-W., Dey, T.K., Edelsbrunner, H., Facello, M.A., Teng, S.-H.: Sliver exudation. J. ACM **47**, 883–904 (2000)
10. Ciarlet, P.G.: The Finite Element Method for Elliptic Problems. SIAM, New York (2002)
11. Dekel, S., Leviatan, D.: The Bramble–Hilbert lemma for convex domains. SIAM J. Math. Anal. **35**(5), 1203–1212 (2004)
12. Ern, A., Guermond, J.L.: Theory and Practice of Finite Elements. Springer, New York (2004)
13. Ern, A., Guermond, J.L.: Finite Elements I: Galerkin Approximation, Elliptic and Mixed PDEs. Springer, New York (2021)
14. Ishizaka, H.: Anisotropic Raviart–Thomas interpolation error estimates using a new geometric parameter (2021). arXiv:2110.02348
15. Ishizaka, H.: Anisotropic interpolation error analysis using a new geometric parameter and its applications. Ph.D. Thesis, Ehime University (2022)
16. Ishizaka, H., Kobayashi, K., Tsuchiya, T.: General theory of interpolation error estimates on anisotropic meshes. Jpn. J. Ind. Appl. Math. **38**(1), 163–191 (2021)
17. Ishizaka, H., Kobayashi, K., Tsuchiya, T.: Crouzeix–Raviart and Raviart–Thomas finite element error analysis on anisotropic meshes violating the maximum-angle condition. Jpn. J. Ind. Appl. Math. **38**(2), 645–675 (2021)
18. Ishizaka, H., Kobayashi, K., Suzuki, R., Tsuchiya, T.: A new geometric condition equivalent to the maximum angle condition for tetrahedrons. Comput. Math. Appl. **99**, 323–328 (2021)
19. Křížek, M.: On semiregular families of triangulations and linear interpolation. Appl. Math. Praha **36**, 223–232 (1991)
20. Křížek, M.: On the maximum angle condition for linear tetrahedral elements. SIAM J. Numer. Anal. **29**, 513–520 (1992)