Applications of Mathematics

Kenta Kobayashi; Takuya Tsuchiya

Error estimation for finite element solutions on meshes that contain thin elements

Applications of Mathematics, Vol. 69 (2024), No. 5, 571–588

Persistent URL: http://dml.cz/dmlcz/152631

Terms of use:

© Institute of Mathematics CAS, 2024

Institute of Mathematics of the Czech Academy of Sciences provides access to digitized documents strictly for personal use. Each copy of any part of this document must contain these *Terms of use*.



This document has been digitized, optimized for electronic delivery and stamped with digital signature within the project *DML-CZ: The Czech Digital Mathematics Library* http://dml.cz

69 (2024) APPLICATIONS OF MATHEMATICS No. 5, 571–588

ERROR ESTIMATION FOR FINITE ELEMENT SOLUTIONS ON MESHES THAT CONTAIN THIN ELEMENTS

Kenta Kobayashi, Kunitachi, Takuya Tsuchiya, Toyonaka

Received Feburary 28, 2024. Published online October 14, 2024.

Cordially dedicated to the memory of I. Babuška

Abstract. In an error estimation of finite element solutions to the Poisson equation, we usually impose the shape regularity assumption on the meshes to be used. In this paper, we show that even if the shape regularity condition is violated, the standard error estimation can be obtained if "bad" elements that violate the shape regularity or maximum angle condition are covered virtually by simplices that satisfy the minimum angle condition. A numerical experiment illustrates the theoretical result.

Keywords: finite element method; triangulation; minimum and maximum angle condition; shape regularity condition; bad triangles

MSC 2020: 65D05, 65N30

1. Introduction

The mathematical theory of finite element methods has been well established, and many well written textbooks on the subject are available [4], [6], [9]. In these textbooks, the central assumption of the theory is the *shape regularity condition* on the meshes that are used to define finite element solutions. The shape regularity condition requires that each element is not very degenerate. For example, if a mesh consists of triangles, each triangle should not be very "flat". It is known that the shape regularity condition is equivalent to the *minimum angle condition*.

The authors were supported by JSPS KAKENHI Grant Numbers 24K00538 and 21K03372.

DOI: 10.21136/AM.2024.0047-24

© Institute of Mathematics, Czech Academy of Sciences 2024.

However, suppose the problem domain has a highly complex shape, and/or the problem has an anisotropic nature, and we introduce an adaptive mesh refinement procedure. In that case, the meshes might contain some very thin elements and the error estimation of the finite element method may thus lose its theoretical validity.

There are two ways to overcome this issue. One way is to introduce a new geometric condition to meshes, where the condition ensures a standard convergence rate of finite element solutions. It is well known that if the maximum angles of triangles in a mesh are bounded by a fixed constant $\theta < \pi$, then the standard error estimation of finite element solutions holds [2]. This condition is known as the *maximum angle condition*. The approach has been extended in many ways. See, for example, [1], [11], [12], [14].

The other way was initiated by Kučera [16] and Duprez-Lleras-Lozinski [7]. Suppose that a mesh contains some "bad" elements. In [16], [7], if bad elements satisfy some conditions, the usual Lagrange interpolation is modified in certain ways on bad elements so that the standard error estimation holds.

This paper extends their approach. We show that if all bad elements are covered virtually by simplices that satisfy the minimum angle condition, the standard error estimation holds. Note that the simplices are virtual and do not need to be elements of the mesh considered.

In Section 2, we present the model problem, the precise definitions of conformal meshes, and the above-mentioned conditions. In addition, we give an example of bad meshes for which the standard error estimation holds. Section 3 presents detailed assumptions of meshes and the main theorem. Section 4 presents a simpler version of the assumption and the main theorem for the case $n = 2$. In Section 5, we compare our approach with approaches adopted in prior works [16], [7]. In Section 6, we present a numerical experiment that confirms the theoretical result.

2. Preliminary

2.1. Model problem. Let $\Omega \subset \mathbb{R}^n$ $(n = 2, 3)$ be a bounded polygonal or polyhedral domain. For simplicity, we consider the Poisson equation

$$-\Delta u = f \quad \text{in } \Omega, \quad u = 0 \quad \text{on } \partial\Omega$$

and its weak form

$$\int_\Omega \nabla u \cdot \nabla v \,\mathrm{d}\mathbf{x} = \int_\Omega f v \,\mathrm{d}\mathbf{x} \quad \forall\, v \in H_0^1(\Omega) \tag{2.1}$$

as the model problem, where $f \in L^2(\Omega)$ is a given function. Here, $L^p(\Omega)$, $H^m(\Omega)$, $H_0^1(\Omega)$, and $W^{m,p}(\Omega)$ $(m = 1, 2,\ 1 \leqslant p \leqslant \infty)$ are the standard Lebesgue and

Sobolev spaces with norms and seminorms $\|\cdot\|_{L^p(\Omega)}$, $\|\cdot\|_{H^m(\Omega)}$, $|\cdot|_{H^m(\Omega)}$, $\|\cdot\|_{W^{m,p}(\Omega)}$, and $|\cdot|_{W^{m,p}(\Omega)}$. In this paper, we assume that the exact solution u of (2.1) belongs to $H^2(\Omega)$.

2.2. Finite element method. We approximate the exact solution of (2.1) using the standard piecewise linear finite element method. To this end, we define the piecewise linear finite element spaces by

$$S_h := \{v_h \in C^0(\overline{\Omega}) \mid v_h|_\tau \in \mathcal{P}_1 \ \forall\, \tau \in \mathcal{T}_h\}, \quad S_{h0} := S_h \cap H_0^1(\Omega),$$

where $\mathcal{P}_1$ is the set of polynomials with degree at most 1. Here, $\mathcal{T}_h$ is a conformal simplicial mesh of Ω. That is, $\mathcal{T}_h$ is a finite set of n-simplices that satisfies the following conditions:

- ▷ $\bigcup_{\tau\in\mathcal{T}_h} \tau = \overline{\Omega}$.
- ▷ If $\tau_1 \cap \tau_2 \neq \emptyset$ for $\tau_1, \tau_2 \in \mathcal{T}_h$ with $\tau_1 \neq \tau_2$, then $\tau_1 \cap \tau_2$ is their common r-face ($r < n$), that is, there are no hanging nodes in $\mathcal{T}_h$.

As usual, we define $h_\tau := \operatorname{diam}\tau$ and $h := \max_{\tau\in\mathcal{T}_h} h_\tau$. The finite element solution $u_h \in S_{h0}$ is then defined as the unique solution of the equation

$$\int_\Omega \nabla u_h \cdot \nabla v_h \,\mathrm{d}\mathbf{x} = \int_\Omega f v_h \,\mathrm{d}\mathbf{x} \quad \forall\, v_h \in S_{h0}. \tag{2.2}$$

2.3. Céa's lemma and interpolations. For the error of u_h, Céa's lemma claims that there exists a constant $C > 0$ such that

$$|u - u_h|_{H^1(\Omega)} \leqslant C \inf_{w_h\in S_{h0}} |u - w_h|_{H^1(\Omega)} \leqslant C|u - \Pi^* u|_{H^1(\Omega)},$$

where $\Pi^* u \in S_h$ is *any* interpolation of the exact solution u. Therefore, if we have

$$|u - \Pi^* u|_{H^1(\Omega)} \leqslant Ch|u|_{H^2(\Omega)}, \tag{2.3}$$

we obtain the standard convergence rate

$$|u - u_h|_{H^1(\Omega)} \leqslant Ch|u|_{H^2(\Omega)}.$$

Usually, the Lagrange interpolation operator is used as Π^*. The piecewise linear Lagrange interpolation $\Pi_\tau^1 u \in \mathcal{P}_1$ is defined by

$$(\Pi_\tau^1 u)(\mathbf{x}_i) = u(\mathbf{x}_i) \text{ on } n\text{-simplex } \tau \text{ with vertices } \mathbf{x}_i,\ i = 1, \dots, n+1,$$
$$\Pi_h^1 u\,|_\tau = \Pi_\tau^1 u \quad \forall\, \tau \in \mathcal{T}_h.$$

In the next subsection, we mention conditions for which the Lagrange interpolation satisfies the standard error estimation

$$|u - \Pi_\tau^1 u|_{W^{1,p}(\tau)} \leqslant C_p h_\tau |u|_{W^{2,p}(\tau)} \quad \forall\, u \in W^{2,p}(\tau). \tag{2.4}$$

2.4. Conditions and constants related to meshes.

Definition 2.1. For an n-simplex τ, $\theta_\tau^{\min}$ and $\theta_\tau^{\max}$ denote the minimum and maximum inner angles of τ, respectively. Here, the inner angles of a tetrahedron mean, for $\theta_\tau^{\min}$, the inner angles of each face and the solid angles at vertices, and for $\theta_\tau^{\max}$, the inner angles of each face and the dihedral angles [3], [15].

The following conditions are well known for geometric conditions of simplices.

Definition 2.2.

(1) If $\theta_\tau^{\min} \geqslant \alpha > 0$, τ is said to satisfy the *minimum angle condition* with respect to α.
(2) If $\theta_\tau^{\max} \leqslant \beta < \pi$, τ is said to satisfy the *maximum angle condition* with respect to β.

If τ satisfies the minimum angle condition with respect to α, the error estimation (2.4) holds for any p, $1 \leqslant p \leqslant \infty$ with $C_p = C_{\alpha,p}^{\min}$. Moreover, there exists a constant D_α such that

$$\frac{h_\tau}{\varrho_\tau} \leqslant D_\alpha, \tag{2.5}$$

where ϱ_τ is the diameter of the inscribed ball in τ. This condition is known as the *shape regularity condition* with respect to D_α. It is well known that the shape regularity condition is equivalent to the minimum angle condition.

Suppose that τ satisfies the maximum angle condition with respect to β. If $n = 2$, (2.4) holds for any p, $1 \leqslant p \leqslant \infty$ with $C_p = C_{\beta,p}^{\max}$. If $n = 3$, (2.4) holds for p, $2 < p \leqslant \infty$ with $C_p = C_{\beta,p}^{\max}$; however, $C_{\beta,p}^{\max} \to \infty$ as $p \searrow 2$, see [8], [13], [15], [17].

To specify the characteristics of meshes, we introduce constants θ, ψ, where $0 < \theta \leqslant \pi/3$, $0 < \psi \leqslant \pi/3$, and fix them throughout this paper. We define

$$K_\theta := \begin{cases} \{\tau \subset \mathbb{R}^2 \mid \tau \text{ is a triangle and } \theta_\tau^{\max} \leqslant \pi - 2\theta\} & \text{for } n = 2, \\ \{\tau \subset \mathbb{R}^3 \mid \tau \text{ is a tetrahedron and } \theta_\tau^{\min} \geqslant \theta\} & \text{for } n = 3. \end{cases}$$

A simplex τ that does not belong to K_θ is regarded as "bad", a simplex τ that satisfies the minimum angle condition $\theta_\tau^{\min} \geqslant \theta$ is regarded as "good", and other types of simplex are regarded as "ordinary". Note that for $n = 2$, K_θ consists of "good" and "ordinary" triangles, and for $n = 3$, all tetrahedrons in K_θ are "good".

For a simplex τ that satisfies $\tau \in K_\theta$ or $\theta_\tau^{\min} \geqslant \psi$, we rewrite (2.4) as

$$|u - \Pi_\tau^1 u|_{H^1(\tau)} \leqslant A_{\theta,\psi} h_\tau |u|_{H^2(\tau)} \quad \forall\, u \in H^2(\tau), \tag{2.6}$$

where $A_{\theta,\psi}$ is a constant that depends only on θ and ψ. The existence of $A_{\theta,\psi}$ is assured by the minimum and maximum angle conditions. (The parameter θ is used to categorize elements in mesh whereas the parameter ψ relates to virtual simplices that appear later.)

2.5. Sobolev embedding theorem. Let $\hat\tau$ be a reference simplex with $h_{\hat\tau} = 1$. Then, by the Sobolev embedding theorem [5], we have

$$H^2(\hat\tau) \subset W^{1,6}(\hat\tau) \subset L^\infty(\hat\tau)$$

and

$$\|\hat u - \Pi_{\hat\tau}^1 \hat u\|_{L^\infty(\hat\tau)} \leqslant \widehat{B} |\hat u - \Pi_{\hat\tau}^1 \hat u|_{H^2(\hat\tau)} = \widehat{B} |\hat u|_{H^2(\hat\tau)}, \quad \hat u \in H^2(\hat\tau),$$

where $\widehat{B}$ is a constant that depends only on $\hat\tau$. Here, we used the fact that $(u - \Pi_{\hat\tau} u)(\mathbf{x}_i) = 0,\ i = 1, \dots, n+1,$ and $\|\hat u - \Pi_{\hat\tau}^1 \hat u\|_{H^2(\hat\tau)}$ may be bounded by $|\hat u - \Pi_{\hat\tau}^1 \hat u|_{H^2(\hat\tau)}$ up to a constant. For a general simplex τ with $\theta_\tau^{\min} \geqslant \psi$, we consider an affine map $\eta\colon\ \hat\tau \to \tau$, $\eta(\mathbf{x}) = A\mathbf{x} + \mathbf{b}$, where A is an $n \times n$ nonsingular matrix and $\mathbf{b} \in \mathbb{R}^n$. Each $u \in H^2(\tau)$ is pulled back to $\hat u$ as $\hat u(\mathbf{x}) := u(\eta(\mathbf{x}))$. We then have

$$\|u - \Pi_\tau^1 u\|_{L^\infty(\tau)}^2 = \|\hat u - \Pi_{\hat\tau}^1 \hat u\|_{L^\infty(\hat\tau)}^2 \leqslant \widehat{B}^2 |\hat u|_{H^2(\hat u)}^2 \leqslant \widehat{B}^2 |\det A|^{-1} \|A\|^4 |u|_{H^2(\tau)}^2,$$

where $\|A\|$ is the matrix norm of A with respect to the Euclidean norm. We denote the Lebesgue measure of τ by $|\tau|$. By [6], the coefficient of the last term is estimated as

$$\det A = \frac{|\tau|}{|\hat\tau|} \geqslant C h^n, \quad \|A\| \leqslant \frac{h_\tau}{\varrho_{\hat\tau}},$$

and we finally obtain the embedding inequality

$$\|u - \Pi_\tau^1 u\|_{L^\infty(\tau)} \leqslant B_\psi h_\tau^{2-n/2} |u|_{H^2(\tau)} \quad \forall\, u \in H^2(\tau), \tag{2.7}$$

where B_ψ is a constant that depends on ψ.

2.6. Example of error estimation on a bad mesh. The minimum or maximum angle conditions are sufficient but not necessary for the standard error estimation (2.4) or (2.6). Suppose that we have the two meshes $\mathcal{T}_h^1$ and $\mathcal{T}_h^2$ depicted in Figure 1 (where $\mathcal{T}_h^2$ has already been mentioned in [10]). The mesh $\mathcal{T}_h^1$ consists of congruent isosceles right triangles and satisfies the minimum angle (shape regularity)

condition. In contrast, $\mathcal{T}_h^2$ is a "refinement" of $\mathcal{T}_h^1$ and contains very flat triangles. Thus, $\mathcal{T}_h^2$ does not satisfy either the minimum or maximum angle condition when the flat triangles degenerate.

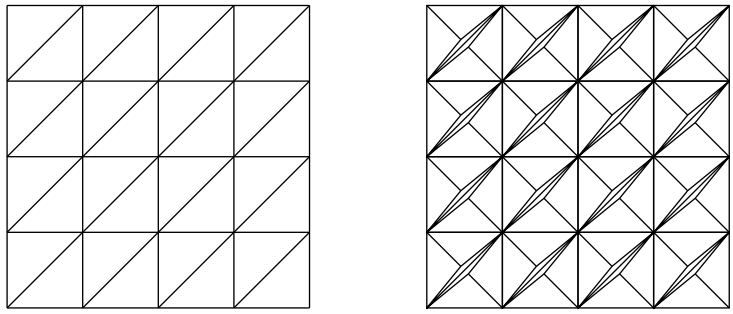


Figure 1. Mesh $\mathcal{T}_h^1$ (left) and its refinement $\mathcal{T}_h^2$ (right).

As each $\tau \in \mathcal{T}_h^2$ is a subset of some $\tau' \in \mathcal{T}_h^1$, we define the interpolation Π^* on $\mathcal{T}_h^2$ as follows. Let $\mathcal{N}_h^1$ and $\mathcal{N}_h^2$ be the sets of vertices of $\mathcal{T}_h^1$ and $\mathcal{T}_h^2$, respectively. Let $\{\varphi_i\} \subset S_h$ be the basis of S_h on $\mathcal{T}_h^2$ defined by

$$\varphi_i(\mathbf{x}_i) = 1, \quad \varphi_i(\mathbf{x}_j) = 0, \quad \mathbf{x}_i, \mathbf{x}_j \in \mathcal{N}_h^2, \quad i \neq j.$$

We define

$$w_i := \begin{cases} (\Pi^1_{\mathcal{T}_h^1} u)(\mathbf{x}_i) & \text{for } \mathbf{x}_i \in \mathcal{N}_h^2 \setminus \mathcal{N}_h^1, \\ u(\mathbf{x}_i) & \text{for } \mathbf{x}_i \in \mathcal{N}_h^1, \end{cases} \qquad \Pi^* u := \sum_i w_i \varphi_i,$$

where $\Pi^1_{\mathcal{T}_h^1}$ is the Lagrange interpolation on $\mathcal{T}_h^1$. As $\mathcal{T}_h^1$ is a good mesh, the standard error estimation (2.6) holds, and the finite element solutions converge to u with $\mathcal{O}(h)$. Moreover, Π^* inherits these properties because $\Pi^1_{\mathcal{T}_h^1} u = \Pi^* u$ on Ω.

This example reminds us that even if a mesh contains bad elements, we might be able to define an interpolation Π^* that satisfies (2.3) if bad elements are covered by fine simplices under some additional conditions.

3. Assumptions and the main theorem

In this section, we introduce assumptions that ensure the standard error estimation (2.3) on meshes that contain bad elements and explain the main result in detail. Recall that $0 < \theta$, $\psi \leqslant \pi/3$ are given fixed constants. Let M, N be positive integers and $C \geqslant 1$. These constants are also fixed throughout the paper.

For a simplex τ, let $v(\tau)$ be the set of vertices of τ. For n-simplices T_1 and T_2, we define

$$V(T_1, T_2) := \mathrm{Conv}(v(T_1) \cap v(T_2)), \quad V_{\partial\Omega}(T_1) := \mathrm{Conv}(v(T_1) \cap \partial\Omega),$$

where $\mathrm{Conv}(S)$ denotes the convex hull of the point set S. We present several examples in Figures 2 and 3. For triangles T_1 and T_2, $V(T_1, T_2)$ and $V_{\partial\Omega}(T_1)$ are drawn with thick lines and dots.

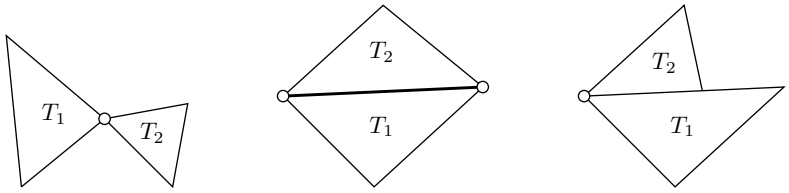


Figure 2. Examples of $V(T_1, T_2)$.

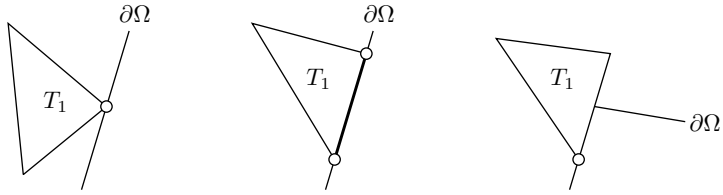


Figure 3. Examples of $V_{\partial\Omega}(T_1)$.

We impose the following conditions for the mesh $\mathcal{T}_h$.

Assumption 3.1.

(1) *There exist n-simplices $T_1, \ldots, T_m$ such that $\theta^{\min}_{T_k} \geqslant \psi$ and $h_{T_k} \leqslant Ch$ for $k = 1, \ldots, m$. We assume that $\overset{\circ}{T}_k \subset \Omega$, where $\overset{\circ}{T}_k$ is the interior set of T_k. The interior sets $\{\overset{\circ}{T}_k\}$ may overlap, and the multiplicity of the overlap of $\{\overset{\circ}{T}_k\}_{k=1}^m$ is at most M. (Note that we do NOT assume $T_k \in \mathcal{T}_h$ $(k = 1, \ldots, m)$.)*

(2) *The mesh $\mathcal{T}_h$ is expressed as a disjoint union $\mathcal{T}_h = \bigcup_{k=0}^{m} Q_k$ such that $Q_0 \subset \mathcal{T}_h \cap K_\theta$ and each "bad" element $\tau \in \mathcal{T}_h \setminus K_\theta$ belongs to one of $\{Q_k\}_{k=1}^m$. Furthermore, we assume that*

$$\bigcup_{\tau \in Q_k} \tau \subset \overline{T}_k, \quad \text{and} \quad \#Q_k \leqslant N \quad \text{for } k = 1, \ldots, m,$$

where $\#Q_k$ is the number of elements in Q_k. Note that Q_0 consists of "good" and "ordinary" elements, and Q_k $(k = 1, \ldots, m)$ may contain "good" or "ordinary" elements as well as "bad" elements.

(3) *If* $(v(\tau_0) \cap v(\tau_1)) \setminus v(T_k) \neq \emptyset$ *for* $\tau_0 \in Q_0$, $\tau_1 \in Q_k$, $1 \leqslant k \leqslant m$, *then we assume that* $\theta_{\tau_0}^{\min} \geqslant \theta$. *(Note that if* $n = 3$, *this condition is automatically satisfied because all elements belonging to* Q_0 *are "good".)*

(4) *If* $v(\tau_1) \cap v(\tau_2) \neq \emptyset$ *for* $\tau_1 \in Q_k$, $\tau_2 \in Q_l$, $1 \leqslant k < l \leqslant m$, *then we assume* $v(\tau_1) \cap v(\tau_2) \subset V(T_k, T_l)$.

(5) *If* $v(\tau_1) \cap \partial\Omega \neq \emptyset$ *for* $\tau_1 \in Q_k$, $1 \leqslant k \leqslant m$, *then we assume* $v(\tau_1) \cap \partial\Omega \subset V_{\partial\Omega}(T_k)$.

Let $\{\varphi_i\} \subset S_h$ be the basis of S_h on the mesh $\mathcal{T}_h$.

Definition 3.1. For $u \in H^2(\Omega)$, we define the interpolation $\Pi^* u$ on $\mathcal{T}_h$ by

$$w_i := \begin{cases} (\Pi_{T_k}^1 u)(\mathbf{x}_i) & \text{if } \mathbf{x}_i \in v(\tau) \text{ for some } \tau \in Q_k, \\ u(\mathbf{x}_i) & \text{otherwise,} \end{cases} \qquad \Pi^* u := \sum_i w_i \varphi_i,$$

where $\mathbf{x}_i$ are the vertices of the mesh $\mathcal{T}_h$ and $\Pi_{T_k}^1$ is the Lagrange interpolation on T_k.

R e m a r k 3.1. Assumption 3.1 (1), (2) requires that each bad element in $\mathcal{T}_h$ belongs to one of the disjoint element clusters Q_k $(k = 1, \ldots, m)$ and each Q_k is virtually covered by a simplex T_k $(k = 1, \ldots, m)$. The interpolation $\Pi^* u$ is defined by $\Pi_{T_k}^1 u$ on each T_k.

To explain the meaning of Assumption 3.1 (3), we consider the situation that $\tau_0 \in Q_0$ and $\tau_1 \in Q_k$ as is depicted in Figure 4. Then, τ_1 is virtually covered by T_k (with other elements in Q_k if $\{\tau_1\} \subsetneqq Q_k$). Let $\mathbf{x} \in v(\tau_0) \cap v(\tau_1) \setminus v(T_k)$. Figure 4 (left) depicts such a situation. As the interpolation $\Pi^* u$ is defined by $\Pi_{T_k}^1 u$, we have $(\Pi^* u)(\mathbf{x}) = (\Pi_{T_k}^1 u)(\mathbf{x}) \neq (\Pi_{\tau_0}^1 u)(\mathbf{x})$ in general in this case. To absorb the effect of $(\Pi^* u)(\mathbf{x}) \neq (\Pi_{\tau_0}^1 u)(\mathbf{x})$, we use the shape regularity condition (2.5) for τ_0 and the Sobolev embedding inequality (2.7) for T_k. Therefore, we require that τ_0 satisfies the minimum angle condition.

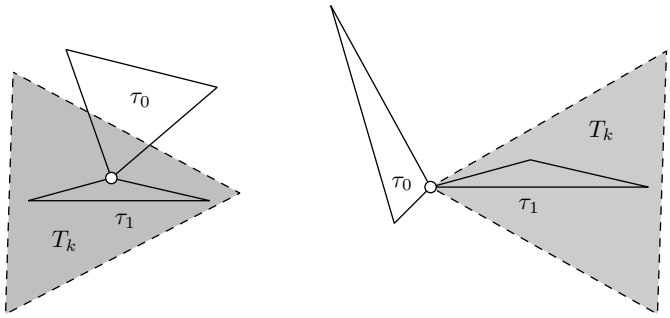


Figure 4. Two situations regarding Assumption 3.1 (3).

If $\mathbf{x} \in v(T_k)$, in contrast, we have $(\Pi^* u)(\mathbf{x}) = (\Pi_{T_k}^1 u)(\mathbf{x}) = (\Pi_{\tau_0}^1 u)(\mathbf{x})$, and we do not need to impose such a condition (see Figure 4 (right)).

Suppose now that $v(\tau_1) \cap v(\tau_2) \neq \emptyset$, $\tau_1 \in Q_k$, $\tau_2 \in Q_l$, $(1 \leqslant k < l \leqslant m)$ and $\mathbf{x} \in v(\tau_1) \cap v(\tau_2)$ is an inner point of Ω. If $(\Pi^1_{T_k} u)(\mathbf{x}) \neq (\Pi^1_{T_l} u)(\mathbf{x})$, then $\Pi^* u$ is not well-defined. This case may occur if $\mathbf{x}$, T_k, and T_l are in the situation depicted in Figure 5. Assumption 3.1 (4) is imposed to avoid such a situation. Three situations that satisfy Assumption 3.1 (4) are depicted in Figure 6.

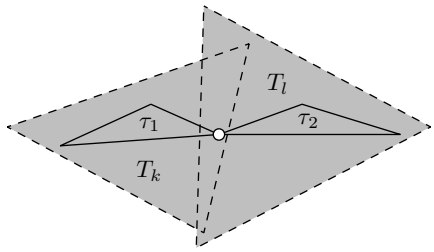


Figure 5. A situation that makes $\Pi^* u$ be not well-defined.

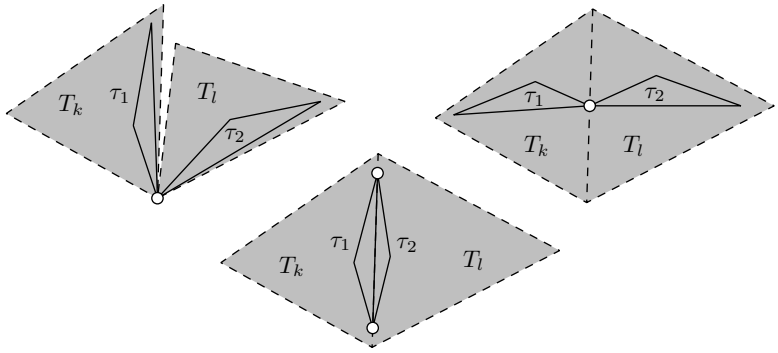


Figure 6. Three situations that satisfy Assumption 3.1(4).

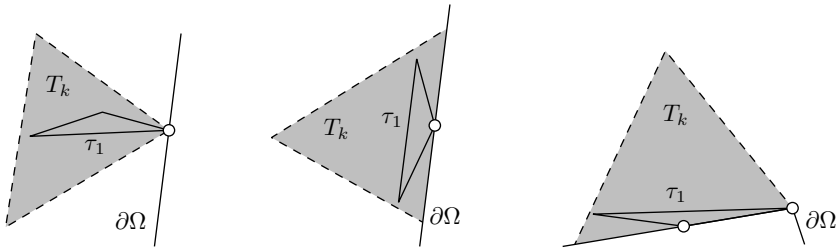


Figure 7. Three situations that satisfy Assumption 3.1 (5).

Suppose that $\tau_1 \in Q_k$ $(1 \leqslant k \leqslant m)$ with $v(\tau_1) \cap \partial\Omega \neq \emptyset$. The interpolation $\Pi^* u$ should then satisfy the boundary condition $\Pi^* u = 0$ on $\partial\Omega$. Assumption 3.1 (5) is imposed to ensure that this condition is fulfilled. Three situations that satisfy Assumption 3.1 (5) are depicted in Figure 7.

Theorem 3.1. *Suppose that the mesh $\mathcal{T}_h$ satisfies Assumption* 3.1. *The interpolation $\Pi^* u$ is then well defined for $u \in H^2(\Omega)$, and we have*

$$|u - \Pi^* u|_{H^1(\Omega)} \leqslant Eh|u|_{H^2(\Omega)} \quad \forall\, u \in H^2(\Omega), \tag{3.1}$$

$$E := \Big(A^2_{\theta,\psi}(n + 2 + C^2 M) + \frac{2(n+1)(n+2)\pi B^2_\psi C^{4-n} D_\theta MN}{n\theta}\Big)^{1/2}. \tag{3.2}$$

P r o o f. Define $W \subset Q_0$ by $W := \{\tau \in Q_0 \mid \theta^{\min}_\tau \geqslant \theta\}$. We decompose $|u - \Pi^* u|^2_{H^1(\Omega)}$ as

$$|u - \Pi^* u|^2_{H^1(\Omega)} = \sum_{\tau\in W} |u - \Pi^* u|^2_{H^1(\tau)} + \sum_{\tau\in Q_0\setminus W} |u - \Pi^* u|^2_{H^1(\tau)} + \sum_{k=1}^{m} \sum_{\tau\in Q_k} |u - \Pi^* u|^2_{H^1(\tau)}.$$

We estimate each term on the right-hand side. As

$$\Pi^* u = \Pi^1_\tau u - (\Pi^1_\tau u - \Pi^* u) = \Pi^1_\tau u - \sum_{\mathbf{x}_i \in v(\tau)} (u(\mathbf{x}_i) - w_i)\varphi_i$$

for $\tau \in W$, we have

$$u - \Pi^* u = u - \Pi^1_\tau u + \sum_{\mathbf{x}_i \in v(\tau)} (u(\mathbf{x}_i) - w_i)\varphi_i.$$

As $u(\mathbf{x}_i) \neq w_i$ occurs to at most $n + 1$ vertices of τ, by the triangle and Cauchy–Schwarz inequalities, we have

$$|u - \Pi^* u|^2_{H^1(\tau)} \leqslant (n + 2)|u - \Pi^1_\tau u|^2_{H^1(\tau)} + (n + 2) \sum_{\mathbf{x}_i \in v(\tau)} |u(\mathbf{x}_i) - w_i|^2 |\varphi_i|^2_{H^1(\tau)}.$$

Hence, we have

$$\begin{aligned}
\sum_{\tau\in W} |u - \Pi^* u|^2_{H^1(\tau)} &\leqslant (n + 2) \sum_{\tau\in W} |u - \Pi_\tau u|^2_{H^1(\tau)} \\
&\quad + (n + 2) \sum_{\tau\in W} \sum_{\mathbf{x}_i \in v(\tau)} |u(\mathbf{x}_i) - w_i|^2 |\varphi_i|^2_{H^1(\tau)} \\
&\leqslant (n + 2) A^2_{\theta,\psi} h^2 \sum_{\tau\in W} |u|^2_{H^2(\tau)} \\
&\quad + (n + 2) \sum_{\tau\in W} \sum_{\mathbf{x}_i \in v(\tau)} |u(\mathbf{x}_i) - w_i|^2 |\varphi_i|^2_{H^1(\tau)} \quad \text{(by (2.6))} \\
&\leqslant (n + 2) A^2_{\theta,\psi} h^2 \sum_{\tau\in W} |u|^2_{H^2(\tau)} \\
&\quad + (n + 2) \sum_{\tau\in W} \frac{D_\theta h^{n-2}_\tau}{n(n-1)} \sum_{\mathbf{x}_i \in v(\tau)} |u(\mathbf{x}_i) - w_i|^2.
\end{aligned}$$

Here, we used the fact that

$$|\nabla\varphi_i| \leqslant \frac{1}{e_\tau} = \frac{|L_\tau|}{n|\tau|} \leqslant \frac{h_\tau^{n-1}}{n(n-1)|\tau|}, \tag{3.3}$$
$$|\varphi_i|^2_{H^1(\tau)} \leqslant \frac{|\tau|}{e_\tau^2} \leqslant \frac{h_\tau^{n-1}}{n(n-1)e_\tau} \leqslant \frac{h_\tau^{n-1}}{n(n-1)\varrho_\tau} \leqslant \frac{D_\theta h_\tau^{n-2}}{n(n-1)} \quad \text{(by (2.5))},$$

where L_τ is the longest edge or widest facet of τ, and e_τ is the height of τ with respect to L_τ.

For the last term in the above inequality, we note that

$$\begin{aligned}
&(n+2)\sum_{\tau\in W}\frac{D_\theta h_\tau^{n-2}}{n(n-1)}\sum_{\mathbf{x}_i\in v(\tau)}|u(\mathbf{x}_i)-w_i|^2\\
&\quad\leqslant (n+2)\frac{D_\theta h^{n-2}}{n(n-1)}\sum_{\tau\in W}\sum_{k=1}^{m}\sum_{\beta\in Q_k}\sum_{\mathbf{x}_i\in v(\tau)\cap v(\beta)}|u(\mathbf{x}_i)-w_i|^2\\
&\quad\leqslant (n+2)\frac{D_\theta h^{n-2}}{n(n-1)}\sum_{\tau\in W}\sum_{k=1}^{m}\sum_{\beta\in Q_k}\sum_{\mathbf{x}_i\in v(\tau)\cap v(\beta)}\|u-\Pi^1_{T_k}u\|^2_{L^\infty(T_k)}\\
&\quad\leqslant (n+2)\frac{D_\theta h^{n-2}}{n(n-1)}\sum_{\tau\in W}\sum_{k=1}^{m}\sum_{\beta\in Q_k}\sum_{\mathbf{x}_i\in v(\tau)\cap v(\beta)}B_\psi^2 h_{T_k}^{4-n}|u|^2_{H^2(T_k)} \quad \text{(by (2.7))}\\
&\quad\leqslant (n+2)\frac{D_\theta h^{n-2}}{n(n-1)}\sum_{\tau\in W}\sum_{k=1}^{m}\sum_{\beta\in Q_k}\sum_{\mathbf{x}_i\in v(\tau)\cap v(\beta)}B_\psi^2 C^{4-n}h^{4-n}|u|^2_{H^2(T_k)}\\
&\qquad\qquad\qquad

$$\leqslant \frac{2(n+1)(n+2)\pi B_\psi^2 C^{4-n} D_\theta N h^2}{n\theta} \sum_{k=1}^{m} |u|^2_{H^2(T_k)} \quad \text{(by Assumption 3.1 (2))}$$
$$\leqslant \frac{2(n+1)(n+2)\pi B_\psi^2 C^{4-n} D_\theta M N h^2}{n\theta} |u|^2_{H^2(\Omega)} \quad \text{(by Assumption 3.1 (1))}.$$

Combining these inequalities, we obtain

$$\sum_{\tau\in W} |u-\Pi^* u|^2_{H^1(\tau)} \leqslant (n+2) A^2_{\theta,\psi} h^2 \sum_{\tau\in W} |u|^2_{H^2(\tau)} + \frac{2(n+1)(n+2)\pi B_\psi^2 C^{4-n} D_\theta M N h^2}{n\theta} |u|^2_{H^2(\Omega)}.$$

Suppose now that $\tau \in Q_0 \setminus W$. By Assumption 3.1(3), we have $\Pi^* u = \Pi^1_\tau u$ on τ as is pointed out in Remark 5. Hence, we have

$$\sum_{\tau\in Q_0\setminus W} |u-\Pi^* u|^2_{H^1(\tau)} = \sum_{\tau\in Q_0\setminus W} |u-\Pi^1_\tau u|^2_{H^1(\tau)} \leqslant A^2_{\theta,\psi} h^2 \sum_{\tau\in Q_0\setminus W} |u|^2_{H^2(\tau)} \quad \text{by (2.6)}.$$

As $\Pi^* u = \Pi^1_{T_k} u$ on $\tau \in Q_k$, $k = 1, \ldots, m$ by definition, we have

$$\begin{aligned}\sum_{k=1}^{m} \sum_{\tau\in Q_k} |u-\Pi^* u|^2_{H^1(\tau)} &\leqslant \sum_{k=1}^{m} |u-\Pi^1_{T_k} u|^2_{H^1(T_k)} \\ &\leqslant \sum_{k=1}^{m} A^2_{\theta,\psi} h^2_{T_k} |u|^2_{H^2(T_k)} \quad \text{(by (2.6))} \\ &\leqslant A^2_{\theta,\psi} C^2 M h^2 |u|^2_{H^2(\Omega)} \quad \text{(by Assumption 3.1 (1))}.\end{aligned}$$

Gathering these three inequalities, we obtain

$$\begin{aligned}&|u-\Pi^* u|^2_{H^1(\Omega)} \\ &\quad = \sum_{\tau\in W} |u-\Pi^* u|^2_{H^1(\tau)} + \sum_{\tau\in Q_0\setminus W} |u-\Pi^* u|^2_{H^1(\tau)} + \sum_{k=1}^{m} \sum_{\tau\in Q_k} |u-\Pi^* u|^2_{H^1(\tau)} \\ &\quad \leqslant (n+2) A^2_{\theta,\psi} h^2 \sum_{\tau\in W} |u|^2_{H^2(\tau)} + \frac{2(n+1)(n+2)\pi B_\psi^2 C^{4-n} D_\theta M N h^2}{n\theta} |u|^2_{H^2(\Omega)} \\ &\qquad + A^2_{\theta,\psi} h^2 \sum_{\tau\in Q_0\setminus W} |u|^2_{H^2(\tau)} + A^2_{\theta,\psi} C^2 M h^2 |u|^2_{H^2(\Omega)} \\ &\quad \leqslant \Big(A^2_{\theta,\psi}(n+2+C^2 M) + \frac{2(n+1)(n+2)\pi B_\psi^2 C^{4-n} D_\theta M N}{n\theta} \Big) h^2 |u|_{H^2(\Omega)},\end{aligned}$$

and Theorem 3.1 is proved. □

4. Simpler formulation for the case $n = 2$.

For the case $n = 2$, we introduce an assumption simpler than Assumption 3.1 that is still practical and applicable to many meshes.

Assumption 4.1.

(1) *There exist similar isosceles triangles $T_1, \ldots, T_m$ whose inner angles are ϕ, ϕ, and $\pi - 2\phi$. We suppose that $\overset{\circ}{T}_k \subset \Omega$, $k = 1, \ldots, m$, $\overset{\circ}{T}_i \cap \overset{\circ}{T}_j = \emptyset$ $(i \neq j,$ $i, j = 1, \ldots, m)$.*

(2) *If $\theta_\tau^{\max} > \pi - \phi$ for $\tau \in \mathcal{T}_h$, then $\overset{\circ}{\tau} \subset \overset{\circ}{T}_k$ with some k, and the longest edges of τ and T_k coincide. Let $\mathbf{x} \in v(\tau) \setminus v(T_k)$. All $\tau \in \mathcal{T}_h$ that share $\mathbf{x}$ must satisfy $\theta_\tau^{\min} \geqslant \phi/2$.*

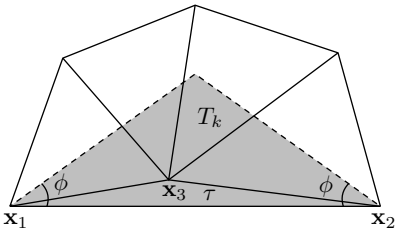


Figure 8. A bad element $\tau = \triangle \mathbf{x}_1\mathbf{x}_2\mathbf{x}_3$ is covered virtually by a triangle T_k. They share the edge $\mathbf{x}_1\mathbf{x}_2$. The minimum inner angles of all elements except τ that share $\mathbf{x}_3$ are greater than or equal to $\phi/2$.

Figure 8 depicts a typical situation. Assumption 4.1 is a special case of Assumption 3.1. In fact, if Assumption 4.1 is satisfied, Assumption 3.1 is satisfied with $M = N = 1$, $C = 1$, $\theta = \phi/2$, and $\psi = \phi$. Hence, we have the following corollary.

Corollary 4.1. *Suppose that the mesh $\mathcal{T}_h$ satisfies Assumption* 4.1. *Then the interpolation $\Pi^* u$ is well defined for $u \in H^2(\Omega)$, and we have*

$$|u - \Pi^* u|_{H^1(\Omega)} \leqslant Eh|u|_{H^2(\Omega)} \quad \forall\, u \in H^2(\Omega),$$

where E is a constant that depends only on ϕ and is independent of $h > 0$.

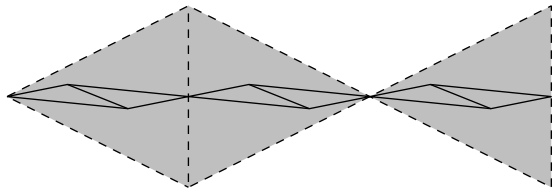


Figure 9. Bad elements that satisfy Assumption 3.1 but do not satisfy Assumption 4.1.

Figure 9 shows a mesh that satisfies Assumption 3.1 but does not satisfy Assumption 4.1.

5. Comparison of the results with the results of prior works

In this section, we compare our approach with approaches adopted in prior works. As stated, the approach adopted in this paper was initiated by Kučera [16] and Duprez, Lleras, and Lozinski [7].

First, we explain the approach of Kučera [16]. Suppose that there exists a bad triangle τ with nodes $\mathbf{x}_1$, $\mathbf{x}_2$, and $\mathbf{x}_3$ in the mesh at which the angle $\angle\mathbf{x}_1\mathbf{x}_3\mathbf{x}_2$ is close to π (Figure 10 (left)). The interpolation $\Pi^{**}u$ for $u \in W^{2,\infty}(\tau)$ is then defined by

$$(\Pi^{**}u)(\mathbf{x}_1) = u(\mathbf{x}_1), \quad (\Pi^{**}u)(\mathbf{x}_2) = u(\mathbf{x}_2), \quad (\nabla(\Pi^{**}u)(\mathbf{x}_0), \mathbf{v}) = (\nabla u(\mathbf{x}_0), \mathbf{v}),$$

where $\mathbf{x}_0$ is the foot on the edge $\overline{\mathbf{x}_1\mathbf{x}_2}$ from $\mathbf{x}_3$, and $\mathbf{v} := (\mathbf{x}_3 - \mathbf{x}_0)/|\mathbf{x}_3 - \mathbf{x}_0|$ (Figure 10 (right)).



Figure 10. A bad element and the modified Lagrange interpolation.

Let τ_1 be another element that shares the node $\mathbf{x}_3$. Obviously, $(\Pi^1_{\tau_1}u)(\mathbf{x}_3) \neq (\Pi^{**}u)(\mathbf{x}_3)$ in general. To adjust the difference, the *correction function* w is introduced so that

$$(u+w)(\mathbf{x}_3) = (\Pi^{**}u)(\mathbf{x}_3), \quad \operatorname{supp} w = B(\mathbf{x}_3, r_\tau),$$
$$B(\mathbf{x}_3, r_\tau) := \{\mathbf{x} : |\mathbf{x} - \mathbf{x}_3| \leqslant r_\tau\}, \quad r_\tau := \frac{1}{2}\min\{|\mathbf{x}_3 - \mathbf{x}_1|, |\mathbf{x}_3 - \mathbf{x}_2|\}.$$

Then $\Pi^{**}u$ is defined by $\Pi^1_{\tau_1}(u+w)$ on τ_1. To make $\Pi^{**}u$ well-defined globally, the supports of the correction functions should be disjoint (see Figure 11).

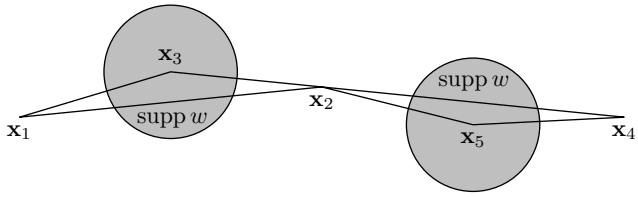


Figure 11. The supports of the correction functions should be disjount.

An advantage of the approach adopted in [16] is that the surrounding elements of a bad element need to satisfy only the maximum angle condition, whereas Theorem 3.1 requires that such surrounding elements satisfy the minimum angle condition (see Assumption 3.1(3)). Hence, [16], Theorem 5, can deal with the mesh depicted in Figure 12, whereas Theorem 3.1 is not applied in this case.

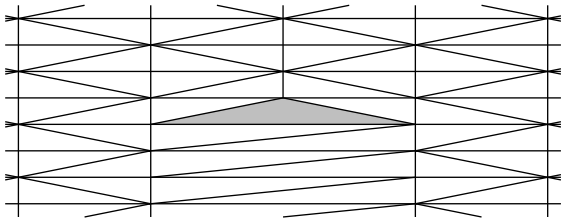


Figure 12. A bad element is surrounded by elements that satisfy the maximum angle condition.

In contrast, the condition that the supports of the correction functions are disjoint is rather restrictive. For example, [16], Theorem 5, cannot deal with the bad elements depicted in Figure 6 (center), whereas Theorem 3.1 can be applied. In [16], Section 3.4, Kučera presented a way to deal with clusters of bad elements. However, each such cluster must be sufficiently far from other clusters and the boundary $\partial\Omega$. Note that Theorem 3.1 can be applied in the cases of clusters of bad elements that are close or connected to other clusters or the boundary $\partial\Omega$ if Assumption 3.1(3) is satisfied.

Moreover, the approach in [16] deals with only triangular elements ($n = 2$) in the $W^{2,\infty}$ setting, whereas Theorem 3.1 is applicable to the more general case of n-simplices ($n = 2, 3$) in the H^2 setting.

Second, we explain the approach of Duprez, Lleras, and Lozinski [7], which is slightly different from that of [16]. Their assumption is that each cluster of simplices that contains bad elements should be star-shaped and completely surrounded by good simplices, see Figure 13.

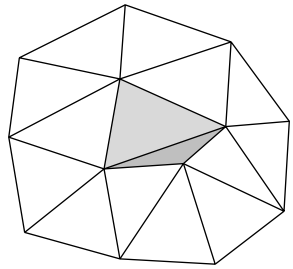


Figure 13. A bad element is surrounded by good elements.

The interpolation $\Pi^{***}u$ is then defined by Taylor polynomials on the cluster. On the surrounding good elements, the Lagrange interpolation $\Pi^1_\tau u$ is modified to ensure the continuity of $\Pi^{***}u$ between elements. To evaluate the influence of the procedure, Duprez, Lleras, and Lozinski used L^∞ error estimates of the Sobolev embedding, which is quite similar to our approach. Their main theorem [7], Proposition 2.7, can deal with both two- and three-dimensional cases in the H^2 setting.

Although [7], Proposition 2.7, is applicable to a wide range of configurations of bad elements, it seems that Theorem 3.1 can also deal with most of such cases. Furthermore, as Assumption 3.1 does not require that each cluster of bad elements is surrounded by good elements, Theorem 3.1 is applicable to a wider range of configurations than [7], Proposition 2.7. For example, the mesh given in the numerical experiment (Figure 14) cannot be dealt with using the cited approach.

6. Numerical experiment

We present a numerical experiment to confirm the theoretical results obtained in the previous sections. Let $\Omega := (0,1) \times (0,1)$. We consider the Poisson equation

$$-\Delta u = 2(x(1-x) + y(1-y)) \quad \text{in } \Omega, \quad u = 0 \quad \text{on } \partial\Omega, \tag{6.1}$$

whose exact solution is $u = x(1-x)y(1-y)$.

Let K be a positive integer and $k = 1/K$. To define the triangulation $\mathcal{T}_h$, we first divide Ω into K^2 small squares with side lengths k. Each small square is further subdivided into six triangles as depicted in Figure 14. Here, the coordinates of the points in each small square are (x,y), $(x+k,y)$, $(x+k,y+k)$, $(x,y+k)$, $(x+k/2,y+\alpha k)$, and $(x+k/2,y+(1-\alpha)k)$, where $0 < \alpha \ll 1$. Note that the smallest triangle becomes very bad when α approaches to zero. It is clear that this triangulation satisfies Assumption 4.1, and Corollary 4.1 is applicable to this triangulation $\mathcal{T}_h$.

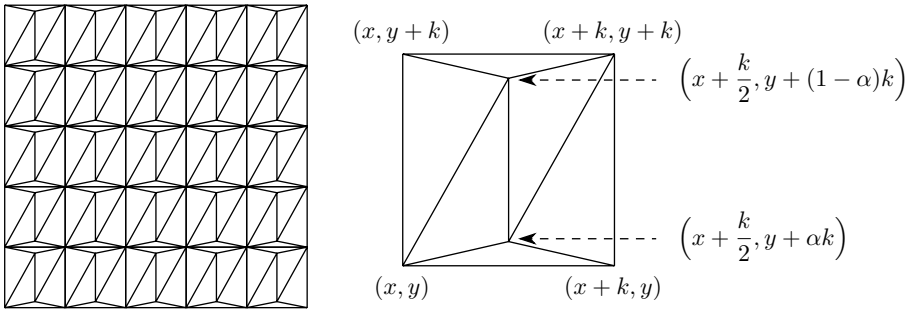


Figure 14. Triangulation $\mathcal{T}_h$ of Ω (left), and the subdivision of each small square (right).

We compute the piecewise linear finite element solutions $u_h \in S_{h0}$ of (6.1) on $\mathcal{T}_h$, and the error $|u-u_h|_{H^1(\Omega)}$ with varying K and α. Note that $h = (0.25+(1-\alpha)^2)^{1/2}k$ if $\alpha < (2-\sqrt{3})/2 = 0.1339\ldots$

Table 1 presents the results of the numerical experiment. Here, e_h stands for $|u-u_h|_{H^1(\Omega)}$. Although the meshes $\mathcal{T}_h$ contain many very bad elements, we observe that the error $|u-u_h|_{H^1(\Omega)}$ is of $\mathcal{O}(h)$ and consistent with Theorem 3.1 and Corollary 4.1.

	$\alpha = 0.1$		$\alpha = 0.01$		$\alpha = 0.0001$	
K	e_h	e_h/h	e_h	e_h/h	e_h	e_h/h
10	1.8002 e−2	0.17485	2.0839 e−2	0.18789	2.1237 e−2	0.18997
20	9.0151 e−3	0.17512	1.0440 e−2	0.18827	1.0641 e−2	0.19036
40	4.5093 e−3	0.17519	5.2229 e−3	0.18836	5.3231 e−3	0.19046
80	2.2548 e−3	0.17521	2.6118 e−3	0.18839	2.6619 e−3	0.19049
160	1.1274 e−3	0.17521	1.3059 e−3	0.18866	1.3310 e−3	0.19049

Table 1. Results of the numerical experiment. In the table, e_h stands for $|u - u_h|_{H^1(\Omega)}$.

7. Concluding remarks and future works

In this paper, we examined the error estimation of finite element solutions on meshes that contain many bad elements. Our research was motivated by the works of Kučera [16] and Duprez, Lleras, and Lozinski [7]. We showed that even if $\mathcal{T}_h$ contains many bad elements, finite element solutions $\{u_h\}$ converge to the exact solution $u \in H^2(\Omega)$ with order $\mathcal{O}(h)$ if each bad element is covered virtually by a simplex that satisfies the minimum angle condition. To be precise, we presented Assumption 3.1 that is very general and covers a wide range of triangulation patterns.

For the case $n = 2$, we gave Assumption 4.1, which is simpler than Assumption 3.1 but is still practical and applicable to many meshes that cannot be dealt with using the methods of [16], [7].

In our numerical experiment, the finite element solutions behaved exactly as predicted from Theorem 3.1 and Corollary 4.1.

The results of this paper can be extended in many ways. An immediate extension may be the consideration of finite element solutions using higher order elements, and the authors are now working on this extension.

References

[1] *T. Apel*: Anisotropic Finite Element: Local Estimates and Applications. Advances in Numerical Mathematics. B. G. Teubner, Leipzig, 1999. zbl MR

[2] *I. Babuška, A. K. Aziz*: On the angle condition in the finite element method. SIAM J. Numer. Anal. *13* (1976), 214–226. zbl MR doi

[3] *J. Brandts, S. Korotov, M. Křížek*: On the equivalence of regularity criteria for triangular and tetrahedral finite element partitions. Comput. Math. Appl. *55* (2008), 2227–2233. zbl MR doi

[4] *S. C. Brenner, L. R. Scott*: The Mathematical Theory of Finite Element Methods. Texts in Applied Mathematics 15. Springer, New York, 2008. zbl MR doi

[5] *H. Brezis*: Functional Analysis, Sobolev Spaces and Partial Differential Equations. Universitext. Springer, New York, 2011. zbl MR doi

[6] *P. G. Ciarlet*: The Finite Element Method for Elliptic Problems. Studies in Mathematics and Its Applications 4. North Holand, Amsterdam, 1978. zbl MR doi

[7] *M. Duprez, V. Lleras, A. Lozinski*: Finite element method with local damage of the mesh. ESAIM, Math. Model. Numer. Anal. *53* (2019), 1871–1891. zbl MR doi

[8] *R. G. Durán*: Error estimates for 3D narrow finite elements. Math. Comput. *68* (1999), 187–199. zbl MR doi
[9] *A. Ern, J.-L. Guermond*: Theory and Practice of Finite Elements. Applied Mathematical Sciences 159. Springer, New York, 2004. zbl MR doi
[10] *A. Hannukainen, S. Korotov, M. Křížek*: The maximum angle condition is not necessary for convergence of the finite element method. Numer. Math. *120* (2012), 79–88. zbl MR doi
[11] *P. Jamet*: Estimations d'erreur pour des éléments finis droits presque dégénérées. Rev. Franc. Automat. Inform. Rech. Operat., Analyse numer. *10* (1976), 43–60. (In French.) zbl MR doi
[12] *K. Kobayashi, T. Tsuchiya*: Extending Babuška-Aziz's theorem to higher-order Lagrange interpolation. Appl. Math., Praha *61* (2016), 121–133. zbl MR doi
[13] *K. Kobayashi, T. Tsuchiya*: Error analysis of Lagrange interpolation on tetrahedrons. J. Approx. Theory *249* (2020), Article ID 105302, 20 pages. zbl MR doi
[14] *M. Křížek*: On semiregular families of triangulations and linear interpolation. Appl. Math., Praha *36* (1991), 223–232. zbl MR doi
[15] *M. Křížek*: On the maximum angle condition for linear tetrahedral elements. SIAM J. Numer. Anal. *29* (1992), 513–520. zbl MR doi
[16] *V. Kučera*: On necessary and sufficient conditions for finite element convergence. Available at `https://arxiv.org/abs/1601.02942` (2016), 42 pages. doi
[17] *N. A. Shenk*: Uniform error estimates for certain narrow Lagrange finite elements. Math. Comput. *63* (1994), 105–119. zbl MR doi

Authors' addresses: *Kenta Kobayashi*, Hitotsubashi University, 2-1 Naka, Kunitachi, Tokyo 186-8601, Japan, e-mail: `kenta.k@r.hit-u.ac.jp`; *Takuya Tsuchiya* (corresponding author), Osaka University, Machikaneyamacho, 560-0043 Toyonaka, Japan, e-mail: `tsuchiya.takuya.plateau@kyudai.jp`.