NOLTA, IEICE

Paper

Verified computations to semilinear elliptic boundary value problems on arbitrary polygonal domains

Akitoshi Takayasu [1,3a)], *Xuefeng Liu* [2], *and Shin'ichi Oishi* [1,4]

[1] *Faculty of Science and Engineering, Waseda University*

[2] *Research Institute for Science and Engineering, Waseda University*

[3] *JSPS Research Fellow*

[4] *CREST/JST*

[a)] *takitoshi@aoni.waseda.jp*

Received April 20, 2012; Revised July 31, 2012; Published January 1, 2013

Abstract: In this paper, a numerical verification method is presented for second-order semilinear elliptic boundary value problems on arbitrary polygonal domains. Based on the Newton-Kantorovich theorem, our method can prove the existence and local uniqueness of the solution in the neighborhood of its approximation. In the treatment of polygonal domains with an arbitrary shape, which gives a singularity of the solution around the re-entrant corner, the computable error estimate of a projection into the finite-dimensional function space plays an essential role. In particular, the lack of smoothness of the solution makes classical error estimates fail on nonconvex domains. By using the *Hyper-circle equation*, an alternative error estimate of the projection has been proposed. Additionally, a new residual evaluation method based on the mixed finite element method works well. It yields more accurate evaluation than the existing method. The efficiency of our method is shown through illustrative numerical results on several polygonal domains.

Key Words: Computer-assisted proof, Semilinear elliptic equations, Finite element method, Verified numerical computations

1. Introduction

Let $\mathbb{R}$ and $\mathbb{N}$ be sets of real and natural numbers, respectively. Let Ω be a bounded polygonal domain in $\mathbb{R}^2$ with an arbitrary shape. We are concerned with the Dirichlet boundary value problem of the following semilinear elliptic equation:

$$\begin{cases} -\Delta u = f(\nabla u, u, x), & \text{in } \Omega, \\ u = 0, & \text{on } \partial\Omega, \end{cases} \tag{1}$$

where $f : H_0^1(\Omega) \to L^2(\Omega)$ is assumed to be Fréchet differentiable with respect to u, *e.g.*, for any bounded $N \in \mathbb{N}$, the following f

DOI: 10.1588/nolta.4.34

$$f(\nabla u, u, x) = (b \cdot \nabla)u + c_1 u + c_2 u^2 + c_3 u^3 + ... + c_N u^N + g$$

with $b \in (L^\infty(\Omega))^2$, $c_i \in L^\infty(\Omega)$, $(i = 1, ..., N)$ and $g \in L^2(\Omega)$ satisfies this condition. Here, the function spaces $H_0^1(\Omega)$, $L^2(\Omega)$ and $L^\infty(\Omega)$ are defined in Section 2. In this paper, we will propose a verified computation procedure for proving the existence of a solution for semilinear elliptic equations on arbitrary polygonal domains. If we have a *good* approximate solution in a certain function space, we will try to validate the existence of a solution with verified error bounds:

$$\|u - \hat{u}\|_{H_0^1} \leq \rho,$$

where u is the exact solution of (1) and $\hat{u}$ is its approximation. Our proposed method is based on the Newton-Kantorovich theorem.

Computer-assisted proofs are also known as verified computations for differential equations. The development of computer-assisted proofs to two-point boundary value problems (one-dimensional case) has pioneered by Kantorovich [8] and Urabe [27]. The works of McCarthy and Tapia [14] and Kedem [9] followed. In 1988, Nakao [15] presented a method of computer-assisted proof for the existence of solutions to elliptic problems including two-point boundary value problems. This method has been shown to be useful for generating a tight numerical inclusion of solutions [15, 17]. One of the features of his method is that a novel fixed-point formula is set up by decomposing the function space into the finite-dimensional part and its complement. In 1991, Plum [18] presented another method of proving the existence and uniqueness of solutions to elliptic boundary value problems. In his method, the norm of the inverse of a linearized operator is bounded by an eigenvalue-enclosing technique based on the homotopy method. In the last two decades, both Nakao's method and Plum's method have been demonstrated to be useful for developing a computer-assisted existence proof of solutions to various elliptic boundary value problems [15–19, 28].

1.1 Two previous works

This part is devoted to briefly describing two existing methods by Nakao [15] and Plum [19]. These two methods can be applied to the following operator equation. Let $\mathcal{A}$ be the linear operator $H_0^1(\Omega) \to H^{-1}(\Omega)$ and $\mathcal{N}$ be the nonlinear operator $H_0^1(\Omega) \to H^{-1}(\Omega)$, where $H^{-1}(\Omega)$ denotes the space of linear continuous functionals on $H_0^1(\Omega)$. We can define an operator equation as

$$\mathcal{F}(u) = \mathcal{A}u - \mathcal{N}(u) = 0, \tag{2}$$

where $\mathcal{F} : H_0^1(\Omega) \to H^{-1}(\Omega)$. Problem (1) is transformed into this operator equation. Assuming the invertibility of $\mathcal{A}$, in Nakao's method, (2) is transformed into the invariance form

$$\mathcal{A}^{-1}\mathcal{F}(u) = u - \mathcal{A}^{-1}\mathcal{N}(u) = 0. \tag{3}$$

Let $\mathcal{I}$ be the identity operator in $H_0^1(\Omega)$. The function space V_h is assumed to be a finite-dimensional subspace of $H_0^1(\Omega)$. Using the orthogonal projection $\mathcal{P}_h : H_0^1(\Omega) \to V_h \subset H_0^1(\Omega)$, Nakao's method transforms the equation

$$u = \mathcal{A}^{-1}\mathcal{N}(u),$$

which is equivalent to (3), into

$$\begin{aligned} \mathcal{P}_h u &= \mathcal{P}_h \mathcal{A}^{-1}\mathcal{N}(u), \\ (\mathcal{I} - \mathcal{P}_h)u &= (\mathcal{I} - \mathcal{P}_h)\mathcal{A}^{-1}\mathcal{N}(u). \end{aligned}$$

For a certain approximate solution $\hat{u} \in V_h$ of (3), Nakao's method further defines $N_h : H_0^1(\Omega) \to V_h$ as

$$N_h(u) := \mathcal{P}_h u - \left[(\mathcal{I} - \mathcal{P}_h \mathcal{A}^{-1}\mathcal{N}'[\hat{u}])|_{V_h}\right]^{-1} \mathcal{P}_h(u - \mathcal{A}^{-1}\mathcal{N}(u)).$$

Using this, the fixed-point formulation

$$T_N(u) = N_h(u) + (\mathcal{I} - \mathcal{P}_h)\mathcal{A}^{-1}\mathcal{N}(u)$$

is considered. Then Nakao's method searches for a nonempty bounded convex closed set $U \subset H_0^1(\Omega)$ satisfying $T_N(U) \subset U$. If we can find such a U, then Schauder's fixed-point theorem states that the set U includes at least one solution of (3). This is a simple outline of Nakao's method.

On the other hand, Plum's method considers (2) directly. In Plum's method, the constants δ and K_P are calculated explicitly such that

$$\|\mathcal{F}(\hat{u})\|_{H^{-1}} \leq \delta \tag{4}$$

and

$$\|u\|_{H_0^1} \leq K_P \|\mathcal{F}'[\hat{u}]u\|_{H^{-1}}, \quad \text{for all } u \in H_0^1(\Omega). \tag{5}$$

Furthermore, it is assumed that there exists a nondecreasing function $g : [0, \infty) \to [0, \infty)$ such that

$$\|\mathcal{F}'[\hat{u} + w] - \mathcal{F}'[\hat{u}]\|_{H_0^1, H^{-1}} \leq g(\|w\|_{H_0^1}), \ \forall w \in H_0^1(\Omega) \text{ with } g(t) \to 0 \text{ as } t \to 0. \tag{6}$$

In Plum's method, the existence of a solution for (2) is proved using the following theorem, which is similar to the Newton-Kantorovich theorem:

Theorem 1 (Plum [19])**.** *Let δ, K_P and g satisfy conditions (4)–(6). Suppose that a certain $\alpha_P > 0$ exists such that*

$$\delta \leq \frac{\alpha_P}{K_P} - G(\alpha_P),$$

where $G(t) := \int_0^t g(s)ds$, and

$$K_P g(\alpha_P) < 1$$

holds. Then, there exists a solution $u \in H_0^1(\Omega)$ of the equation $\mathcal{F}(u) = 0$ satisfying

$$\|u - \hat{u}\|_{H_0^1} \leq \alpha_P. \tag{7}$$

Moreover, the solution is locally unique under the side condition (7).

1.2 Features and challenges

The methods of Nakao, of Plum and the method to be proposed in this paper have no mathematical difference in the sense that each method uses the fixed-point theorem to prove the existence and uniqueness of solutions. One feature of our procedure is that it can treat (1) on arbitrary polygonal domains without any difficulty. On nonconvex domains, the solution of (1) lacks the H^2 regularity, which poses difficulty in deducing an explicit error estimate. Using the Newton-Kantorovich theorem, we present a verification theory specialized for the finite element method. Overcoming the difficulty, we will introduce a procedure to obtain an a posteriori error estimate for the finite element method. The error estimate is obtained only using the first derivative of the solution. This enables us to treat arbitrary domains.

In Section 2, we will prepare the notations and the framework of our proposed method. In Section 3, we are concerned with explicit error estimation and Sobolev's embedding constant evaluation. In Section 4, several constants needed in applying the Newton-Kantorovich theorem are evaluated. Furthermore, a new method of obtaining a residual bound of the operator equation using the Raviart-Thomas mixed finite element is proposed. Finally, in Section 5, illustrative numerical results are presented to show the usefulness of our procedure.

2. Preliminaries

Here, we introduce several notations used throughout this paper. An n-dimensional vector is denoted by $\mathrm{u} = (u_1, ..., u_n)^T \in \mathbb{R}^n$. Let $|\mathrm{u}|_{l^2}$ be the Euclidean norm

$$|\mathrm{u}|_{l^2} = \sqrt{u_1^2 + u_2^2 + ... + u_n^2}.$$

For a matrix $A \in \mathbb{R}^{n \times n}$, the norm $\|A\|_2$ denotes the spectral norm of matrix A.

Let Ω be a bounded polygonal domain in $\mathbb{R}^2$. Let $L^p(\Omega)$, $p \in [1, \infty)$ denote the space of pth power Lebesgue integrable functions on Ω. It follows that, for $u \in L^p(\Omega)$,

$$\int_\Omega |u(x)|^p dx < +\infty \quad \text{and} \quad \|u\|_{L^p} := \left(\int_\Omega |u(x)|^p dx \right)^{1/p}.$$

In particular, we mainly consider the case $p = 2$ for real functions. We denote the L^2 inner product and L^2 norm as

$$(u, v) := \int_\Omega u(x)v(x)dx, \qquad \|u\|_{L^2} := \sqrt{(u, u)},$$

respectively. For vector functions $\mathrm{u}, \mathrm{v} \in \left(L^2(\Omega)\right)^2$, the L^2 inner product of u and v is denoted by

$$(\mathrm{u}, \mathrm{v}) := \sum_{i=1}^{2} (u_i, v_i), \quad \text{for } \mathrm{u} = (u_1, u_2)^T,\ \mathrm{v} = (v_1, v_2)^T.$$

Let $L^\infty(\Omega)$ denote the space of functions essentially bounded on Ω with the norm

$$\|u\|_{L^\infty} := \operatorname*{ess\,sup}_{x \in \Omega} |u(x)|.$$

$H^r(\Omega)$ denotes the L^2 Sobolev space of order $r \in \mathbb{N}$ with the inner product

$$\langle u, v \rangle_r := \sum_{|k|=0}^{r} (D^{(k)}u, D^{(k)}v).$$

Here, $D^{(k)}$ denotes the partial differentiation with respect to the multi-index $k = (k_1, k_2)$ with $|k| = k_1 + k_2$:

$$D^{(k)}u := \frac{\partial^{|k|}u}{\partial x_1^{k_1} \partial x_2^{k_2}}.$$

The H^r norm and semi-norm are respectively defined for $u \in H^r(\Omega)$ by

$$\|u\|_{H^r} := \left(\sum_{|k| \le r} (D^{(k)}u, D^{(k)}u) \right)^{1/2}, \quad |u|_{H^r} := \left(\sum_{|k| = r} (D^{(k)}u, D^{(k)}u) \right)^{1/2}.$$

Let us further define $H_0^1(\Omega)$ as

$$H_0^1(\Omega) := \left\{ u \in H^1(\Omega) : u = 0 \ \text{ on } \partial\Omega \right\}$$

with the inner product

$$(\nabla u, \nabla v) = \int_\Omega \nabla u \cdot \nabla v dx$$

and the norm

$$\|u\|_{H_0^1} := |u|_{H^1} = \|\nabla u\|_{L^2} = \left(\int_\Omega |\nabla u|^2 dx \right)^{1/2}.$$

Here, '$u = 0$ on $\partial\Omega$' is in the trace sense. Generally, for $p \in [1, \infty]$, $W^{r,p}(\Omega)$ denotes the L^p Sobolev space of order $r \in \mathbb{N}$ with the norm,

$$\|u\|_{W^{r,p}} := \left(\sum_{|k| \le r} \int_\Omega |D^{(k)}u|^p dx \right)^{1/p} \quad \text{for} \quad p \in [1, \infty)$$

and

$$\|u\|_{W^{r,\infty}} := \sum_{|k| \le r} \operatorname*{ess\,sup}_{x \in \Omega} |D^{(k)}u(x)|.$$

Let $H^{-1}(\Omega)$ be the topological dual space of $H_0^1(\Omega)$, *i.e.*, the space of linear continuous functionals in $H_0^1(\Omega)$. Let $T \in H^{-1}(\Omega)$ and $u \in H_0^1(\Omega)$. We denote $Tu \in \mathbb{R}$ as $\langle T, u\rangle$. The norm of $T \in H^{-1}(\Omega)$ is defined by

$$\|T\|_{H^{-1}} := \sup_{0 \neq u \in H_0^1(\Omega)} \frac{|\langle T, u\rangle|}{\|u\|_{H_0^1}}.$$

Further, let X and Y be Banach spaces. The set of a bounded linear operator from X to Y is denoted by $\mathcal{L}(X, Y)$. For $\mathcal{T} \in \mathcal{L}(X, Y)$, its operator norm is denoted by

$$\|\mathcal{T}\|_{X,Y} := \sup_{0 \neq u \in X} \frac{\|\mathcal{T}u\|_Y}{\|u\|_X}.$$

Here, $\|\cdot\|_X$ is the norm in X and $\|\cdot\|_Y$ is the norm in Y.

Let us introduce Sobolev's embedding theorem. For the Banach spaces X and Y, the embedding $X \hookrightarrow Y$ means that a natural embedding map $u \in X \mapsto u \in Y$ is continuous, *i.e.*,

$$\|u\|_Y \leq C\|u\|_X$$

holds for a constant C. Using the Rellich-Kondrashov theorem [1], the following corollary is obtained.

Corollary 2. *Let $\Omega \subset \mathbb{R}^2$ be a bounded polygonal domain. The embedding $H^1(\Omega) \hookrightarrow L^p(\Omega)$ is compact for $\forall p \in [1, \infty)$. Then, it follows that, for $v \in H^1(\Omega)$ and $p \in [1, \infty)$,*

$$\|v\|_{L^p} \leq C_{e,p}|v|_{H^1}. \tag{8}$$

The constant $C_{e,p}$ depends on the shape of Ω. The method of obtaining its concrete value is introduced in Section 3.2. Now we use the notations $X := L^2(\Omega)$, $V := H_0^1(\Omega)$ and $V^* := H^{-1}(\Omega)$ for simplicity.

2.1 Framework of verified computations

This part is devoted to explaining the computer-assisted approach to solving the following abstract problem:

$$\text{Find } u \in V \text{ satisfying } \mathcal{F}(u) = 0 \text{ in } V^*, \tag{9}$$

where $\mathcal{F} : V \to V^*$ denotes a Fréchet differentiable mapping. Let V_h be a finite-dimensional subspace of V. Let $\hat{u} \in V_h \subset V$ be an approximate solution to (9). The Fréchet derivative of $\mathcal{F}$ at $\hat{u}$ is denoted by $\mathcal{F}'[\hat{u}] : V \to V^*$, *i.e.*,

$$\|\mathcal{F}(\hat{u} + \nu) - \mathcal{F}(\hat{u}) - \mathcal{F}'[\hat{u}]\nu\|_{V^*} = o(\|\nu\|_V).$$

To verify the existence and local uniqueness of the exact solution in the neighborhood of $\hat{u}$, we apply the Newton-Kantorovich theorem [6, 8] to (9).

Theorem 3 (Newton-Kantorovich's theorem)**.** *Assume that the Fréchet derivative $\mathcal{F}'[\hat{u}]$ is nonsingular and satisfies*

$$\|\mathcal{F}'[\hat{u}]^{-1}\mathcal{F}(\hat{u})\|_V \leq \alpha,$$

for a certain positive α. Then, let $\overline{B}(\hat{u}, 2\alpha) := \{v \in V : \|v - \hat{u}\|_V \leq 2\alpha\}$ be a closed ball centered at $\hat{u}$ with radius 2α. Also, let $D \supset \overline{B}(\hat{u}, 2\alpha)$ be an open ball in V. We assume that, for a certain positive ω, the following holds:

$$\|\mathcal{F}'[\hat{u}]^{-1}(\mathcal{F}'[v] - \mathcal{F}'[w])\|_{V,V} \leq \omega\|v - w\|_V, \quad \forall v, w \in D.$$

If

$$\alpha\omega \leq \frac{1}{2}$$

holds, then there is a solution $u \in V$ of (9) satisfying

$$\|u - \hat{u}\|_V \leq \rho := \frac{1 - \sqrt{1 - 2\alpha\omega}}{\omega}.$$

Furthermore, the solution u is locally unique in $\overline{B}(\hat{u}, \rho)$.

Remark 1. *To apply the Newton-Kantorovich theorem, we will calculate the following constants explicitly.*

$$\|\mathcal{F}'[\hat{u}]^{-1}\|_{V^*,V} \leq C_1, \tag{10}$$

$$\|\mathcal{F}(\hat{u})\|_{V^*} \leq C_{2,h}, \tag{11}$$

$$\|\mathcal{F}'[v] - \mathcal{F}'[w]\|_{V,V^*} \leq C_3\|v - w\|_V, \quad \forall v, w \in D \subset V. \tag{12}$$

Therefore, if $C_1^2 C_{2,h} C_3 \leq 1/2$ is confirmed by verified computations, then the existence and local uniqueness of the solution are proved numerically. Our main task in this paper is to calculate these constants explicitly.

Remark 2. *When $\alpha\omega \leq \frac{1}{2}$ is obtained, the uniqueness of the solution is also proved in the ball $\overline{B}(\hat{u}, 2\alpha)$ [6], so that there is a nonexistence area of the solution:*

$$\overline{B}(\hat{u}, 2\alpha) \setminus \overline{B}(\hat{u}, \rho) = \{v \in V : \rho < \|v - \hat{u}\|_V \leq 2\alpha\}.$$

2.2 Variational formulation

In this part, we provide variational formulations. We would like to deduce the form (9) from (1). Our verified computation approach proves the existence and local uniqueness of a weak solution of (1). Here, we rewrite $f(\nabla u, u, x)$ as $f(u)$ for simplicity. In the classical analysis of the variational theory, the weak solution to the Dirichlet boundary problem (1) is simply the solution of the following variational problem: Find $u \in V$ such that

$$(\nabla u, \nabla v) = (f(u), v), \quad \forall v \in V. \tag{13}$$

For $u, v \in V$, let us define the continuous bilinear form $A(\cdot,\cdot) : V \times V \to \mathbb{R}$ as

$$A(u, v) := (\nabla u, \nabla v).$$

For a fixed $u \in V$, $A(u, \cdot) \in V^*$ is a linear functional. It enables us to define the operator $\mathcal{A} : V \to V^*$ by

$$\langle \mathcal{A}u, v\rangle := A(u, v), \quad \forall v \in V.$$

It is obvious that $A(u, v)$ is an inner product in V. Then, for a given $T \in V^*$, Riesz's representation theorem states the existence of a unique solution $u \in V$ such that

$$A(u, v) = \langle T, v\rangle, \quad \forall v \in V,$$

in particular, $\|u\|_V = \|\mathcal{A}u\|_{V^*}$ holds. This shows the invertibility of $\mathcal{A}$. We denote the inverse of $\mathcal{A}$ as $\mathcal{A}^{-1} : V^* \to V$. Thus, the operator $\mathcal{A}$ becomes an isometric isomorphism. For a fixed $u \in V$, $(f(u), \cdot)$ becomes a linear functional. Then, we can define the nonlinear operator $\mathcal{N} : V \to V^*$ by

$$\langle \mathcal{N}(u), v\rangle = (f(u), v), \quad \forall v \in V.$$

Using these operators, the variational problem (13) can be transformed into $\mathcal{A}u = \mathcal{N}(u)$. Furthermore, we define the operator $\mathcal{F} : V \to V^*$ by $\mathcal{F}(u) := \mathcal{A}u - \mathcal{N}(u)$, which can be written as $\mathcal{F}(u) = 0$. This is simply the abstract problem (9).

To apply the Newton-Kantorovich theorem, the Fréchet derivative of $\mathcal{F}$ is needed. The Fréchet differentiability of $\mathcal{F}$ is derived by that of f. We now show that $\mathcal{F} : V \to V^*$ is Fréchet differentiable. For fixed $u, \hat{u} \in V$, $(f'(\hat{u})u, \cdot)$ is a linear functional on V. Here, $f'(\hat{u}) : V \to X$ is the Fréchet derivative of $f : V \to X$ at $\hat{u}$. Hence, we can define the operator $\mathcal{N}'[\hat{u}] : V \to V^*$ by

$$\langle \mathcal{N}'[\hat{u}]u, v\rangle := (f'(\hat{u})u, v), \quad \forall v \in V. \tag{14}$$

For a given $\hat{u} \in V$, the Fréchet derivative $\mathcal{F}'[\hat{u}] : V \to V^*$ of $\mathcal{F} : V \to V^*$ at $\hat{u}$ is given by

$$\mathcal{F}'[\hat{u}] = \mathcal{A} - \mathcal{N}'[\hat{u}].$$

In fact, we have

$$\begin{aligned}\|\mathcal{F}(\hat{u}+v)-\mathcal{F}(\hat{u})-(\mathcal{A}-\mathcal{N}'[\hat{u}])v\|_{V^*} &= \sup_{0\neq w\in V}\frac{|\langle \mathcal{N}(\hat{u}+v)-\mathcal{N}(\hat{u})-\mathcal{N}'[\hat{u}]v,w\rangle|}{\|w\|_V}\\ &= \sup_{0\neq w\in V}\frac{|(f(\hat{u}+v)-f(\hat{u})-f'(\hat{u})v,w)|}{\|w\|_V}\\ &\leq C_{e,2}\|\mu(\hat{u},v)\|_X,\end{aligned}$$

where $\hat{u}, v \in V$ and

$$\mu(\hat{u},v)=f(\hat{u}+v)-f(\hat{u})-f'(\hat{u})v.$$

From the Fréchet differentiability of $f: V \to X$, we have

$$\frac{\|\mu(\hat{u},v)\|_X}{\|v\|_V}\to 0, \quad (\|v\|_V \to 0).$$

This shows the Fréchet differentiability of $\mathcal{F}: V \to V^*$ at $\hat{u} \in V$.

Now, we define a natural embedding operator, $i_{(X\to V^*)}: X \to V^*$. For a fixed $w \in X$, $(w,\cdot) \in V^*$ is a linear functional. Then, we can define

$$\langle i_{(X\to V^*)}w, v\rangle := (w,v), \quad \forall\, v \in V.$$

Since the embedding operator $i_{(V\to X)}: V \to X$ is compact from Corollary 2, its adjoint operator $i_{(X\to V^*)}: X \to V^*$ becomes compact by Schauder's theorem [3]. The operator $i_{(X\to V^*)}: X \to V^*$ is compact and $f'(\hat{u}): V \to X$ is continuous so that the composite operator

$$\mathcal{N}'[\hat{u}] = i_{(X\to V^*)}\circ f'(\hat{u}): V \to V^* \tag{15}$$

is compact.

Remark 3. *The nonlinear operator $\mathcal{N}: V \to V^*$ is presented using this embedding operator s.t.*

$$\mathcal{N}(u) = i_{(X\to V^*)}\circ f(u) \in V^*, \quad \textit{for } f(u) \in X.$$

3. Explicit evaluations

Two constants play an important role in our verification framework. One is an error constant appearing in the error estimation of the finite element method. The other is Sobolev's embedding constant. Recently, an explicit value of error constants for the linear conforming finite element has been given in [10] and [12]. In this section, we will explain how to obtain explicit values of these constants.

3.1 Error constants of FEM

The evaluation of the error constant strongly depends on the shape of the domain. Here, let us define some notations corresponding to mesh triangulations. Let T^h be a mesh triangulation of Ω. The triangle element of T^h is denoted by K_h. Let us define the finite element space $V_h \subset V$ by

$$V_h := \mathrm{span}\{\phi_1, \phi_2, ..., \phi_n\} \subset V, \tag{16}$$

where ϕ_i is a base function of finite elements. If we consider a linear finite element space, n is the number of inner node points in T^h. Let us consider the following Poisson's equation for a given $f \in X$:

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

The weak formulation is presented by

$$\text{Find } u \in V \text{ satisfying } (\nabla u, \nabla v) = (f,v), \quad \forall v \in V. \tag{17}$$

Let us define the orthogonal projection that maps V to V_h by

$$(\nabla(u - \mathcal{P}_h u), \nabla v_h) = 0, \quad \forall v_h \in V_h. \tag{18}$$

The classical error estimation theory gives an a priori estimation of Poisson's equation for the projection $\mathcal{P}_h : V \to V_h$:

$$\|u - \mathcal{P}_h u\|_V \leq C_M \|f\|_X. \tag{19}$$

In the case of the Dirichlet boundary condition with convex domains, we know that the solution of (17) belongs to $H^2(\Omega)$. For such a solution, we say that it has H^2 regularity [7]. In this case, we can easily obtain a concrete value of C_M. Such regularity is not available for nonconvex domains. The lack of H^2 regularity causes a failure in the explicit evaluation of the constant C_M. To treat an arbitrary polygonal domain, we will adopt the technique developed by Liu and Oishi [13].

3.1.1 A priori error estimate with H^2 regularity

In this part, we assume that V_h consists of linear base functions. Here, we will introduce two constants $C_{h,i}$ $(i = 0, 1)$ that play an important role throughout this paper. These constants are related to function interpolations π_i $(i = 0, 1)$ over the triangle element $K_h \in T^h$. For $u \in L^2(K_h)$, $\pi_0 u$ is a constant function defined by

$$\pi_0 u := \left(\int_{K_h} u dx\right) / \left(\int_{K_h} 1 dx\right).$$

For $u \in H^2(K_h)$, the interpolation $\pi_1 u$ of u is a linear function defined by

$$(\pi_1 u)(x) := u(x) \text{ on the vertex of } K_h.$$

For $i = 0, 1$, let global interpolations $\pi_{h,i}$ be an extension of π_i to the entire domain. That is, $(\pi_{h,i} u)|_{K_h} = \pi_i(u|_{K_h})$. Here, we define $C_{h,i}$ over triangulation T^h by

$$C_{h,i} := \max_{K_h \in T^h} C_i(K_h), \quad i = 0, 1 \tag{20}$$

where

$$C_0(K_h) := \sup_{0 \neq v \in H^1(K_h)} \frac{\|\pi_0 u - u\|_X}{\|u\|_V}, \quad C_1(K_h) := \sup_{0 \neq v \in H^2(K_h)} \frac{|\pi_1 u - u|_{H^1}}{|u|_{H^2}}.$$

These constants $C_i(K_h)$ $(i = 0, 1)$ correspond to an eigenvalue of a differential operator. Kikuchi and Liu [10] give the upper bound of constants in the following lemma.

Lemma 1 (Kikuchi and Liu [10]). *For $\alpha \in (0, 1)$ and $\theta \in (0, \pi)$,*

$$C_0(K_h) \leq \frac{h}{\pi}\sqrt{\frac{\nu_+(\alpha, \theta)}{2}}, \quad C_1(K_h) \leq 0.493h \frac{\nu_+(\alpha, \theta)}{\sqrt{2\nu_-(\alpha, \theta)}}$$

with

$$\begin{aligned} \nu_-(\alpha, \theta) &= 1 + \alpha^2 - \sqrt{1 + 2\alpha^2 \cos 2\theta + \alpha^4}, \\ \nu_+(\alpha, \theta) &= 1 + \alpha^2 + \sqrt{1 + 2\alpha^2 \cos 2\theta + \alpha^4}. \end{aligned}$$

Here, $h = |OA|$, $\alpha = |OB|/|OA|$ and $\theta = \angle AOB$ (see Fig. 1).

In particular,

$$C_0 \leq \frac{1}{\pi}, \quad C_1 \leq 0.493$$

hold on the unit isosceles right-angle triangle. Using this lemma, the verified bound of constants $C_{h,i}$ $(i = 0, 1)$ is easy to obtain. Aside from this, other upper bounds for $C_i(K_h)$ $(i = 0, 1)$ are introduced by Kobayashi [12].

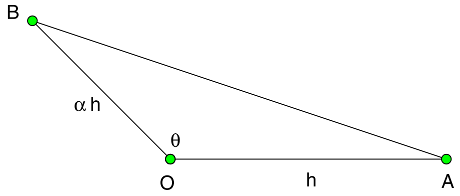

Fig. 1. Triangle element K_h for Lemma 1.

Lemma 2 (Kobayashi [12]). *For an arbitrary triangle element,*

$$C_0(K_h) < \sqrt{\frac{a^2+b^2+c^2}{28} - \frac{S^4}{a^2b^2c^2}}$$

and

$$C_1(K_h) < \sqrt{\frac{a^2b^2c^2}{16S^2} - \frac{a^2+b^2+c^2}{30} - \frac{S^2}{5}\left(\frac{1}{a^2}+\frac{1}{b^2}+\frac{1}{c^2}\right)}$$

hold, where $a = |BC|$, $b = |AC|$, $c = |AB|$ and S is the area of K_h in Fig. 2.

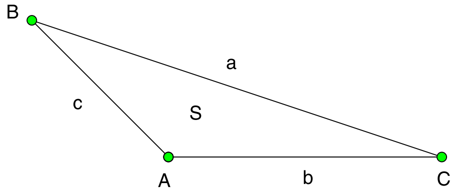

Fig. 2. Triangle element K_h for Lemma 2.

The classical a priori error estimate is given by the following theorem.

Theorem 4. *Let Ω be a* **convex** *polygonal domain. For a given $f \in X$, let u be the solution of the variational problem in (17). The error estimate between u and its approximation $\mathcal{P}_h u \in V_h$ is given by*

$$\|u - \mathcal{P}_h u\|_V \le C_{h,1}\|f\|_X, \quad \|u - \mathcal{P}_h u\|_X \le C_{h,1}\|u - \mathcal{P}_h u\|_V \le (C_{h,1})^2\|f\|_X.$$

Proof. Under the given assumptions, the solution u belongs to $H^2(\Omega)$. By using the interpolation error estimate for $\pi_{h,1}$, the minimization principle gives

$$\|u - \mathcal{P}_h u\|_V \le |u - \pi_{h,1}u|_{H^1} \le C_{h,1}|u|_{H^2} \le C_{h,1}\|f\|_X,$$

where the constant $C_{h,1}$ is the one defined in (20). Here, we use the fact [7] that, for $u \in H^2(\Omega) \cap H^1_0(\Omega)$ and $f = -\Delta u$, we have $|u|_{H^2} \le \|\Delta u\|_X = \|f\|_X$. Furthermore, by adopting Aubin-Nitsche's trick, we can deduce

$$\|u - \mathcal{P}_h u\|_X \le C_{h,1}\|u - \mathcal{P}_h u\|_V \le (C_{h,1})^2\|f\|_X.$$

□

Thus, one can take $C_M = C_{h,1}$ in (19) when we choose V_h as the linear finite element space.

3.1.2 A posteriori error estimate without H^2 regularity

For solutions with a singularity ($u \notin H^2(\Omega)$), it is difficult to give a computable a priori estimation. To solve this problem, Liu and Oishi [13] has proposed a new method based on the Prager-Synge theorem [20]. For readers' convenience, we give a sketch of the main result in [13] in the rest of Section 3.1. Let us define a function space corresponding to the lowest-order Raviart-Thomas mixed finite elements as

$$W_h := \left\{p_h \in H(\mathrm{div}, \Omega) : p_h = (a_k + c_k x, b_k + c_k y)^T \text{ in } K_h\right\},$$

where a_k, b_k and c_k are constants on element K_h and

$$H(\mathrm{div}, \Omega) := \left\{ \psi \in \left(L^2(\Omega) \right)^2 : \mathrm{div}\ \psi \in L^2(\Omega) \right\}.$$

Denoting the base function of W_h by ψ_i,

$$W_h = \mathrm{span}\{\psi_1, \psi_2, ..., \psi_l\}, \tag{21}$$

where l denotes the number of edges in T^h. The set of piecewise constant functions on T^h is defined by

$$M_h := \left\{ v \in L^2(T^h) : v \text{ is constant on each element of } T^h \right\}.$$

Let q_i be the constant function that has a support to be the ith element of T^h. We have

$$M_h = \mathrm{span}\{q_1, q_2, ..., q_m\}, \tag{22}$$

where m is the number of elements in T^h. Classical analysis shows that $\mathrm{div}(W_h) = M_h$ (see [21]). For each $f_h \in M_h$, we define a subset of W_h by

$$W_{f_h} := \{p_h \in W_h : \mathrm{div}\ p_h + f_h = 0, \text{ on each } K_h \in T^h\}.$$

We also define the orthogonal projection $\mathcal{P}_{h,0} : X \to M_h$ by

$$(u - \mathcal{P}_{h,0}u, \mu_h) = 0, \quad \forall \mu_h \in M_h.$$

The property of orthogonality indicates that

$$\|u\|_X^2 = \|\mathcal{P}_{h,0}u\|_X^2 + \|\mathcal{P}_{h,0}u - u\|_X^2, \quad \forall u \in X. \tag{23}$$

From definition (20), the error estimate of the approximation $\mathcal{P}_{h,0}u$ is given by

$$\|u - \mathcal{P}_{h,0}u\|_X \le C_{h,0}|u|_{H^1}, \quad \text{for } u \in H^1(\Omega).$$

To provide the error estimate for the projection $\mathcal{P}_h u$ without the second derivative of u, Liu and Oishi [13] introduce a new computable quantity κ such that

$$\kappa := \max_{0 \ne f_h \in M_h} \min_{v_h \in V_h} \min_{p_h \in W_{f_h}} \frac{\|p_h - \nabla v_h\|_X}{\|f_h\|_X}. \tag{24}$$

Lemma 3 (Liu and Oishi [13]). *For a given $f_h \in M_h$, let $\bar{u} \in H^1(\Omega)$ and $u_h \in V_h$ be solutions of the variational problems,*

$$(\nabla \bar{u}, \nabla v) = (f_h, v), \quad \forall v \in V \quad \textit{and} \quad (\nabla u_h, \nabla v_h) = (f_h, v_h), \quad \forall v_h \in V_h,$$

respectively. Then we have an error estimate using the quantity κ:

$$\|\bar{u} - u_h\|_V \le \kappa \|f_h\|_X. \tag{25}$$

Proof. From the Prager-Synge theorem [20], for $\bar{u}$, any $v_h \in V_h$ and $p_h \in W_{f_h}$, it follows that

$$\|\nabla \bar{u} - \nabla v_h\|_X^2 + \|\nabla \bar{u} - p_h\|_X^2 = \|p_h - \nabla v_h\|_X^2,$$

which is called the *hypercircle equation.* This can be checked by confirming the vanishing of cross terms. Then, the following inequality holds:

$$\|\nabla \bar{u} - \nabla v_h\|_X \le \|p_h - \nabla v_h\|_X, \quad \forall v_h \in V_h, \ \forall p_h \in W_{f_h}.$$

From the minimization principle, we obtain the error estimate between $\bar{u}$ and u_h:

$$\|\nabla \bar{u} - \nabla u_h\|_X \le \|\nabla \bar{u} - \nabla v_h\|_X \le \min_{p_h \in W_{f_h}} \|p_h - \nabla v_h\|_X.$$

Furthermore, the definition of κ yields

$$\|\nabla \bar{u} - \nabla u_h\|_X \le \kappa \|f_h\|_X.$$

□

Theorem 5 (Liu and Oishi [13]). *For $f \in X$, let $u \in V$ and $\mathcal{P}_h u \in V_h$ be the solutions of*

$$(\nabla u, \nabla v) = (f, v), \quad \forall v \in V \quad \textit{and} \quad (\nabla(\mathcal{P}_h u), \nabla v_h) = (f, v_h), \quad \forall v_h \in V_h,$$

respectively. Also, let $C_M := \sqrt{(C_{h,0})^2 + \kappa^2}$. We then have the following a posteriori estimation:

$$\|u - \mathcal{P}_h u\|_V \leq C_M \|f\|_X, \quad \|u - \mathcal{P}_h u\|_X \leq C_M \|u - \mathcal{P}_h u\|_V \leq (C_M)^2 \|f\|_X.$$

Proof. We follow the analogous framework by Kikuchi and Saito [11] to finish the proof. Let $\bar{u}$ and u_h be those defined in Lemma 3 with $f_h = \mathcal{P}_{h,0} f \in M_h$. The minimization principle leads to $\|u - \mathcal{P}_h u\|_V \leq \|u - u_h\|_V$. Decomposing $u - u_h$ by $(u - \bar{u}) + (\bar{u} - u_h)$, we have

$$\|u - \mathcal{P}_h u\|_V \leq \|u - u_h\|_V \leq \|u - \bar{u}\|_V + \|\bar{u} - u_h\|_V.$$

From the definitions of u and $\bar{u}$, it follows for $\forall v \in V$ that

$$(\nabla(u - \bar{u}), \nabla v) = (f - \mathcal{P}_{h,0} f, v) = ((I - \mathcal{P}_{h,0}) f, (I - \mathcal{P}_{h,0}) v).$$

Letting v be $u - \bar{u}$ and applying the error estimate for the projection $\mathcal{P}_{h,0}$, we have

$$\begin{array}{rcl} \|u - \bar{u}\|_V^2 & \leq & \left\|(I - \mathcal{P}_{h,0}) f\right\|_X \left\|(I - \mathcal{P}_{h,0})(u - \bar{u})\right\|_X \\ & \leq & \left\|(I - \mathcal{P}_{h,0}) f\right\|_X \cdot C_{h,0} \|u - \bar{u}\|_V. \end{array}$$

Hence, we have

$$\|u - \bar{u}\|_V \leq C_{h,0} \left\|(I - \mathcal{P}_{h,0}) f\right\|_X. \tag{26}$$

From (23), (25) and (26), the error $\|u - \mathcal{P}_h u\|_V$ is bounded by

$$\begin{array}{rcl} \|u - \mathcal{P}_h u\|_V & \leq & \|u - \bar{u}\|_V + \|\bar{u} - u_h\|_V \\ & \leq & \kappa \|\mathcal{P}_{h,0} f\|_X + C_{h,0} \left\|(I - \mathcal{P}_{h,0}) f\right\|_X \\ & \leq & \sqrt{(C_{h,0})^2 + \kappa^2} \ \|f\|_X. \end{array}$$

Furthermore, by adopting Aubin-Nitsche's trick, the estimate for $\|u - \mathcal{P}_h u\|_X$ can be obtained. Define $e := u - \mathcal{P}_h u \in X$ and $\zeta \in V$ satisfying

$$(\nabla \zeta, \nabla v) = (e, v), \quad \forall v \in V.$$

Then, we have

$$(e, e) = (\nabla \zeta, \nabla e) = (\nabla(\zeta - \mathcal{P}_h \zeta), \nabla e) \leq \|\nabla(\zeta - \mathcal{P}_h \zeta)\|_X \cdot \|\nabla e\|_X \leq C_M \|e\|_X \|\nabla e\|_X,$$

which leads to

$$\|u - \mathcal{P}_h u\|_X \leq C_M |u - \mathcal{P}_h u|_{H^1} \leq (C_M)^2 \|f\|_X.$$

□

Remark 4. *In [29], for an L-shaped domain, Yamamoto and Nakao proposed another method of giving an explicit a priori error estimation for Poisson's problem with a homogeneous Dirichlet boundary condition, where the technique of extending the L-shaped domain to a square one is used to deal with the solution singularity. Principally, the method in [29] can also be extended to solve problems on a general domain, but a complicated domain manipulation is hard work. Also, the numerical comparison in [13] shows that the method of Liu and Oishi gives a much sharper estimation of the constant C_M.*

Computation of κ

This part is devoted to evaluating the quantity κ in (24). The discussion will be divided into two steps. First, we derive the explicit forms of $u_h \in V_h$ and $p_h \in W_{f_h}$, which minimize $\|p_h - \nabla u_h\|_X$ for a fixed f_h. Then, we find $f_h \in M_h$ that maximizes $\|p_h - \nabla u_h\|_X / \|f_h\|_X$.

For a given $f_h \in M_h$, we consider the optimization problem

$$\inf_{u_h \in V_h} \inf_{p_h \in W_{f_h}} \|p_h - \nabla u_h\|_X. \tag{27}$$

The classical theory of the Raviart-Thomas finite element method [2, 4, 21] implies that the minimizer of (27) is given by solutions of the following two problems:

a) Find $p_h \in W_h$ and $\lambda_h \in M_h$ such that

$$\begin{cases} (p_h, q_h) + (\lambda_h, \text{div } q_h) = 0, & \forall q_h \in W_h, \\ (\text{div } p_h, \mu_h) + (f_h, \mu_h) = 0, & \forall \mu_h \in M_h. \end{cases}$$

b) Find $u_h \in V_h$ such that

$$(\nabla u_h, \nabla v_h) = (f_h, v_h), \quad \forall v_h \in V_h.$$

Let the base functions of the finite element spaces V_h, M_h, W_h be those in (16), (21) and (22). Define the matrices $P \in \mathbb{R}^{l \times l}$, $G \in \mathbb{R}^{n \times l}$, $S \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{n \times m}$, $M \in \mathbb{R}^{m \times m}$ and $N \in \mathbb{R}^{m \times l}$, whose i-j elements are given by

$$\begin{array}{ll} P_{i,j} = (\psi_i, \psi_j), & G_{i,j} = (\nabla \phi_i, \psi_j), \\ S_{i,j} = (\nabla \phi_i, \nabla \phi_j), & B_{i,j} = (\phi_i, q_j), \\ M_{i,j} = (q_i, q_j), & N_{i,j} = (q_i, \text{div } \psi_j). \end{array}$$

Additionally, suppose that $\mathrm{x} \in \mathbb{R}^l$, $\mathrm{y} \in \mathbb{R}^n$, $\mathrm{z} \in \mathbb{R}^m$ and $\mathrm{f} \in \mathbb{R}^m$ are vectors and let $p_h \in W_h$, $u_h \in V_h$, $\lambda_h \in M_h$ and $f_h \in M_h$ be the elements such that

$$\begin{array}{ll} \mathrm{x} = (x_1, ..., x_l)^T \in \mathbb{R}^l, & p_h = (\psi_1, ..., \psi_l) \cdot \mathrm{x} \in W_h, \\ \mathrm{y} = (u_1, ..., u_n)^T \in \mathbb{R}^n, & u_h = (\phi_1, ..., \phi_n) \cdot \mathrm{y} \in V_h, \\ \mathrm{z} = (z_1, ..., z_m)^T \in \mathbb{R}^m, & \lambda_h = (q_1, ..., q_l) \cdot \mathrm{z} \in M_h, \\ \mathrm{f} = (f_1, ..., f_m)^T \in \mathbb{R}^m, & f_h = (q_1, ..., q_l) \cdot \mathrm{f} \in M_h. \end{array}$$

By using matrix notations, problems **a)** and **b)** can be characterized by

$$\textbf{a)} \begin{cases} P\mathrm{x} + N^T\mathrm{z} = 0 \\ N\mathrm{x} + M\mathrm{f} = 0 \end{cases}, \quad \textbf{b)}\ S\mathrm{y} = B\mathrm{f}.$$

There are various methods of solving this system. By adopting block matrix arithmetic for this problem, the coefficient vectors of the minimizer, $p_h \in W_h$ and $u_h \in V_h$, are given by

$$\mathrm{x} = -P^{-1}N^T(NP^{-1}N^T)^{-1}M\mathrm{f} =: H\mathrm{f} \quad \text{and} \quad \mathrm{y} = S^{-1}B\mathrm{f} =: K\mathrm{f},$$

if $NP^{-1}N^T$ has an inverse matrix. Then, the following is obtained:

$$\begin{aligned} \|\nabla u_h - p_h\|_X^2 &= (\nabla u_h, \nabla u_h) + (p_h, p_h) - (\nabla u_h, p_h) - (p_h, \nabla u_h) \\ &= \mathrm{y}^T S\mathrm{y} + \mathrm{x}^T P\mathrm{x} - \mathrm{y}^T G\mathrm{x} - \mathrm{x}^T G^T \mathrm{y} \\ &= \mathrm{f}^T(K^T SK + H^T PH - K^T GH - H^T G^T K)\mathrm{f} \\ &= \mathrm{f}^T Q\mathrm{f}. \end{aligned}$$

Here, we put $Q = K^T SK + H^T PH - K^T GH - H^T GK \in \mathbb{R}^{m \times m}$. Note that Q is symmetric. Finally, κ^2 is given by

$$\kappa^2 \quad = \quad \max_{0 \neq f_h \in M_h} \min_{u_h \in V_h} \min_{p_h \in W_{f_h}} \frac{\|\nabla u_h - p_h\|_X^2}{\|f_h\|_X^2}$$

$$= \max_{0 \neq \mathrm{f} \in \mathbb{R}^m} \frac{\mathrm{f}^T Q \mathrm{f}}{\mathrm{f}^T M \mathrm{f}}.$$

This is simply the Rayleigh quotient form of a general matrix eigenvalue problem:

$$Q\mathrm{f} = \lambda M \mathrm{f}. \tag{28}$$

Thus, κ^2 is given by the maximum eigenvalue of (28).

3.2 Embedding constant

Another task of the explicit evaluation is to obtain Sobolev's embedding constant $H^1(\Omega) \hookrightarrow L^p(\Omega)$ on arbitrary polygonal domains. Sobolev's embedding constant in (8) is related to the minimal eigenvalue of the Laplacian $(-\Delta)$, which is discussed by Liu and Oishi [13]. The following lemma is introduced by Plum [19]. He pointed out, "*This is not always optimal but easy to compute*".

Lemma 4. *Let $\sigma \in [0, \infty)$ denote the minimal point of the spectrum corresponding to $-\Delta$ on V. Let $p \in [2, \infty)$ and ν denote the largest integer less than or equal to $p/2$. We have*

$$C_{e,p} := \left(\frac{1}{2}\right)^{\frac{1}{2} + \frac{2\nu - 3}{p}} \left[\frac{p}{2}\left(\frac{p}{2} - 1\right) \cdots \left(\frac{p}{2} - \nu + 2\right)\right]^{\frac{2}{p}} \sigma^{-\frac{1}{p}},$$

where the bracket term is set equal to 1 if $\nu = 1$.

Here, we need a verified lower bound of the minimal eigenvalue of $-\Delta$ on the treated domain. The following theorem gives a desired lower bound and was derived by Liu and Oishi [13].

Theorem 6 (Liu and Oishi [13]). *Let $\{\lambda_k\}$ be eigenvalues of $-\Delta$. λ_k^h is assumed to be its discretized approximation with verified computations. C_M is an error constant satisfying (19). Suppose*

$$1 - (C_M)^2 \lambda_k > 0.$$

Then, each eigenvalue of $-\Delta$ is bounded by

$$\frac{\lambda_k^h}{1 + (C_M)^2 \lambda_k^h} \leq \lambda_k \leq \lambda_k^h.$$

Using this result, we can take

$$\sigma \geq \frac{\lambda_1^h}{1 + (C_M)^2 \lambda_1^h}, \tag{29}$$

where λ_1^h is the first approximate eigenvalue in finite element discretized systems of the eigenvalue problem

$$-\Delta u = \lambda u$$

with the Dirichlet boundary condition $u = 0$ on $\partial\Omega$.

4. Verification theories

Our computer-assisted approach needs explicit values of (10)–(12) in Section 2.1. In this section, we discuss how to calculate each constant with verification.

4.1 Invertibility of linearized operator

Let $\hat{u} \in V_h$ be an approximate solution of (13). Here, we evaluate the upper bound of C_1 in (10), which corresponds to the inverse norm estimation of the Fréchet derivative operator $\mathcal{F}'[\hat{u}] = \mathcal{A} - \mathcal{N}'[\hat{u}]$. Let V_h be a finite element approximation of V and V_c be the orthogonal complement with an H_0^1 inner product. The theorem below is a modification of the main theorem of Nakao et al. [16] in 2005. Here, we give another proof. In Nakao et. al.'s original paper [16], Schauder's fixed-point theorem is used. Since the operator $\mathcal{N}'[\hat{u}]$ is compact from the statement in (15), Fredholm's alternative theorem can be applied to prove the invertibility of the linearized operator.

Theorem 7. *The operator $\mathcal{N}'[\hat{u}] : V \to V^*$ is the linear compact one defined in (14). The finite dimensional subspace V_h is that introduced in (16). Furthermore, $\mathcal{P}_h : V \to V_h$ is the orthogonal projection defined in (18). Let K_1, K_2 and K' be the constants to make the following inequalities hold:*

$$\|f'(\hat{u})u\|_X \leq K_1\|u\|_V, \quad \forall u \in V,$$
$$\|f'(\hat{u})u_c\|_X \leq K_2\|u_c\|_V, \quad \forall u_c \in V_c$$

and

$$\|\mathcal{P}_h\mathcal{A}^{-1}\mathcal{N}'[\hat{u}]u_c\|_V \leq K'\|u_c\|_V, \quad \forall u_c \in V_c.$$

Assume that there exists a positive constant $\tau > 0$ satisfying

$$\|u_h\|_V \leq \tau \left\|\mathcal{P}_h(\mathcal{I} - \mathcal{A}^{-1}\mathcal{N}'[\hat{u}])u_h\right\|_V, \quad \forall u_h \in V_h.$$

Moreover, as in (19), the error estimate of $\mathcal{P}_h$ is available for a given $f \in X$:

$$\|u - \mathcal{P}_h u\|_V \leq C_M\|f\|_X.$$

If $\nu_h := 1 - C_M(K_1\tau K' + K_2) > 0$, then $\mathcal{A} - \mathcal{N}'[\hat{u}] : V \to V^*$ is invertible and satisfies

$$\|(\mathcal{A} - \mathcal{N}'[\hat{u}])^{-1}\|_{V^*,V} \leq \|R\|_2,$$

where $\|R\|_2$ is the spectral norm of a matrix described by

$$R := \begin{bmatrix} \tau\left(\frac{K'}{\nu_h}C_M K_1\tau + 1\right) & \frac{\tau K'}{\nu_h} \\ \frac{C_M K_1 \tau}{\nu_h} & \frac{1}{\nu_h} \end{bmatrix} \in \mathbb{R}^{2\times 2}. \tag{30}$$

Proof. We fix $u \in V$. By putting $\varphi \in V^*$ as

$$(\mathcal{A} - \mathcal{N}'[\hat{u}])u = \varphi \tag{31}$$

and setting

$$u_h := \mathcal{P}_h u, \quad u_c := (\mathcal{I} - \mathcal{P}_h)u,$$
$$\varphi_h := \mathcal{P}_h\mathcal{A}^{-1}\varphi, \quad \varphi_c := (\mathcal{I} - \mathcal{P}_h)\mathcal{A}^{-1}\varphi,$$

the following are obtained:

$$u = u_h + u_c, \quad \mathcal{A}^{-1}\varphi = \varphi_h + \varphi_c.$$

Furthermore, the property of orthogonality indicates that

$$\|u_h\|_V^2 + \|u_c\|_V^2 = \|u\|_V^2, \quad \|\varphi_h\|_V^2 + \|\varphi_c\|_V^2 = \|\mathcal{A}^{-1}\varphi\|_V^2 = \|\varphi\|_{V^*}^2.$$

From (31), we have

$$\begin{aligned} & \mathcal{P}_h\mathcal{A}^{-1}(\mathcal{A} - \mathcal{N}'[\hat{u}])(u_h + u_c) = \varphi_h \\ \iff \quad & \mathcal{P}_h(\mathcal{I} - \mathcal{A}^{-1}\mathcal{N}'[\hat{u}])u_h = \mathcal{P}_h\mathcal{A}^{-1}\mathcal{N}'[\hat{u}]u_c + \varphi_h. \end{aligned}$$

From the assumption, the following inequality holds:

$$\begin{aligned} \|u_h\|_V &\leq \tau\|\mathcal{P}_h(\mathcal{I} - \mathcal{A}^{-1}\mathcal{N}'[\hat{u}])u_h\|_V \\ &= \tau\|\mathcal{P}_h\mathcal{A}^{-1}\mathcal{N}'[\hat{u}]u_c + \varphi_h\|_V \\ &\leq \tau\left(K'\|u_c\|_V + \|\varphi_h\|_V\right). \end{aligned} \tag{32}$$

On the other hand, from (31), it follows that

$$\begin{aligned} & (\mathcal{I} - \mathcal{P}_h)\mathcal{A}^{-1}(\mathcal{A} - \mathcal{N}'[\hat{u}])(u_h + u_c) = \varphi_c \\ \iff \quad & u_c = (\mathcal{I} - \mathcal{P}_h)\mathcal{A}^{-1}\mathcal{N}'[\hat{u}](u_h + u_c) + \varphi_c. \end{aligned}$$

For a given $f \in X$, we note that the solution of

$$(\nabla u, \nabla v) = (f, v), \quad \forall v \in V,$$

can be denoted by $u = \mathcal{A}^{-1} i_{(X\to V^*)} \circ f$. The error estimate (19) is rewritten by

$$\|u - \mathcal{P}_h u\|_V = \|(\mathcal{I} - \mathcal{P}_h)\mathcal{A}^{-1} i_{(X\to V^*)} \circ f\|_V \le C_M \|f\|_X.$$

The representation of $\mathcal{N}'[\hat{u}]$ in (15) yields

$$\begin{aligned} \|(\mathcal{I} - \mathcal{P}_h)\mathcal{A}^{-1}\mathcal{N}'[\hat{u}]u\|_V &= \|(\mathcal{I} - \mathcal{P}_h)\mathcal{A}^{-1} i_{(X\to V^*)} \circ f'(\hat{u})u\|_V \\ &\le C_M \|f'(\hat{u})u\|_X. \end{aligned}$$

Thus, it turns out from (32) that

$$\begin{aligned} \|u_c\|_V &= \left\|(\mathcal{I} - \mathcal{P}_h)\mathcal{A}^{-1}\mathcal{N}'[\hat{u}](u_h + u_c) + \varphi_c\right\|_V \\ &\le C_M \|f'(\hat{u})(u_h + u_c)\|_X + \|\varphi_c\|_V \\ &\le C_M \left(\|f'(\hat{u})u_h\|_X + \|f'(\hat{u})u_c\|_X\right) + \|\varphi_c\|_V \\ &\le C_M \left(K_1\|u_h\|_V + K_2\|u_c\|_V\right) + \|\varphi_c\|_V \\ &\le C_M \left(K_1\tau\left(K'\|u_c\|_V + \|\varphi_h\|_V\right) + K_2\|u_c\|_V\right) + \|\varphi_c\|_V \\ &= C_M(K_1\tau K' + K_2)\|u_c\|_V + C_M K_1 \tau \|\varphi_h\|_V + \|\varphi_c\|_V. \end{aligned}$$

If the assumption

$$\nu_h = 1 - C_M(K_1\tau K' + K_2) > 0 \tag{33}$$

holds, then we have

$$\|u_c\|_V \le \frac{1}{\nu_h}\left(C_M K_1 \tau \|\varphi_h\|_V + \|\varphi_c\|_V\right). \tag{34}$$

Under condition (33), by substituting (34) into (32), it follows that

$$\begin{aligned} \|u_h\|_V &\le \tau\left(\frac{K'}{\nu_h}\left(C_M K_1\tau\|\varphi_h\|_V + \|\varphi_c\|_V\right) + \|\varphi_h\|_V\right) \\ &= \tau\left(\frac{K'}{\nu_h} C_M K_1 \tau + 1\right)\|\varphi_h\|_V + \frac{\tau K'}{\nu_h}\|\varphi_c\|_V. \end{aligned}$$

Summing up the above arguments, the desired conclusion is obtained:

$$\|u\|_V \le \|R\|_2 \|(\mathcal{A} - \mathcal{N}'[\hat{u}])u\|_{V^*}, \tag{35}$$

where $R \in \mathbb{R}^{2\times 2}$ is simply the matrix in (30). From (35), if $(\mathcal{A} - \mathcal{N}'[\hat{u}])u = 0$ in V^*, it follows that $u = 0$. This implies that the operator $\mathcal{A} - \mathcal{N}'[\hat{u}] : V \to V^*$ is injective. Since the operator $\mathcal{A} - \mathcal{N}'[\hat{u}]$ is of the Fredholm type with an index of 0, it is also surjective. Thus, $\mathcal{A} - \mathcal{N}'[\hat{u}]$ is invertible and satisfies

$$\|(\mathcal{A} - \mathcal{N}'[\hat{u}])^{-1}\|_{V^*,V} \le \|R\|_2.$$

This completes the proof. □

Therefore, one can put $C_1 := \|R\|_2$ in (10).

4.1.1 Several constants

The constants K_1, K_2 and K' can be computed explicitly. For K_1 and K_2, we can choose

$$K_1 = \|f'(\hat{u})\|_{V,X} \quad \text{and} \quad K_2 = \|f'(\hat{u})\|_{V_c,X}.$$

Both depend on the concrete notation of the Fréchet derivative $f'(\hat{u})$. Furthermore, for K', let us estimate the norm of $\mathcal{P}_h\mathcal{A}^{-1}\mathcal{N}'[\hat{u}] : V_c \to V_h$ for $u_c \in V_c$:

$$\begin{aligned}\|\mathcal{P}_h\mathcal{A}^{-1}\mathcal{N}'[\hat{u}]u_c\|_V &= \sup_{0\neq v_h\in V_h}\frac{A\left(\mathcal{P}_h\mathcal{A}^{-1}\mathcal{N}'[\hat{u}]u_c, v_h\right)}{\|v_h\|_V}\\ &= \sup_{0\neq v_h\in V_h}\frac{A\left(\mathcal{A}^{-1}\mathcal{N}'[\hat{u}]u_c, v_h\right)}{\|v_h\|_V}\\ &= \sup_{0\neq v_h\in V_h}\frac{\langle\mathcal{N}'[\hat{u}]u_c, v_h\rangle}{\|v_h\|_V}\\ &= \sup_{0\neq v_h\in V_h}\frac{(f'(\hat{u})u_c, v_h)}{\|v_h\|_V}\\ &\le C_{e,2}\|f'(\hat{u})u_c\|_X\\ &\le C_{e,2}K_2\|u_c\|_X.\end{aligned}$$

Thus, one can put $K' = C_{e,2}K_2$. In Section 5, practical notations with respect to K_1 and K_2 are presented.

4.1.2 Method of calculating τ

The upper bound of τ will be evaluated as below. Putting $\mathcal{B}_h := \mathcal{P}_h(\mathcal{I} - \mathcal{A}^{-1}\mathcal{N}'[\hat{u}])|_{V_h} : V_h \to V_h$, it follows for $w_h \in V_h$ that

$$\begin{aligned}\|\mathcal{P}_h(\mathcal{I} - \mathcal{A}^{-1}\mathcal{N}'[\hat{u}])w_h\|_V &= \|\mathcal{B}_h w_h\|_V\\ &= \sup_{0\neq v_h\in V_h}\frac{A(\mathcal{B}_h w_h, v_h)}{\|v_h\|_V}\\ &\ge \left(\inf_{0\neq u_h\in V_h}\sup_{0\neq v_h\in V_h}\frac{A(\mathcal{B}_h u_h, v_h)}{\|u_h\|_V\|v_h\|_V}\right)\|w_h\|_V.\end{aligned}$$

Introduce a quantity η satisfying

$$\eta := \inf_{0\neq u_h\in V_h}\sup_{0\neq v_h\in V_h}\frac{A(\mathcal{B}_h u_h, v_h)}{\|u_h\|_V\|v_h\|_V}.$$

First, by selecting $v_h = \mathcal{B}_h u_h$ in the supremum in the definition, one can see that η is a non-negative quantity. If $\eta > 0$, then we can take $\tau = \eta^{-1}$. In the case of $\eta = 0$, the operator $\mathcal{B}_h$ is not invertible, which means that the bounded constant τ in Theorem 7 cannot be available. Thus, the verified procedure fails and other manipulation, such as mesh refinement, is necessary.

The verified evaluation of η is introduced as follows. Let $\mathrm{x}, \mathrm{y} \in \mathbb{R}^n$ be real vectors and $u_h, v_h \in V_h$ satisfying

$$\begin{aligned}\mathrm{x} = (u_1, ..., u_n)^T \in \mathbb{R}^n, &\qquad u_h = (\phi_1, ..., \phi_n)\cdot\mathrm{x} \in V_h\\ \mathrm{y} = (v_1, ..., v_n)^T \in \mathbb{R}^n, &\qquad v_h = (\phi_1, ..., \phi_n)\cdot\mathrm{y} \in V_h,\end{aligned}$$

respectively. We recall the definition of $S \in \mathbb{R}^{n\times n}$ on page 45 and redefine $B \in \mathbb{R}^{n\times n}$ whose i-j element is

$$B_{i,j} = (\nabla\phi_j, \nabla\phi_i) - (f'(\hat{u})\phi_j, \phi_i),$$

for $1 \le i, j \le n$. Therefore, we have

$$\begin{aligned}\eta &= \inf_{0\neq u_h\in V_h}\sup_{0\neq v_h\in V_h}\frac{A(\mathcal{B}_h u_h, v_h)}{\|u_h\|_V\|v_h\|_V}\\ &= \inf_{0\neq u_h\in V_h}\sup_{0\neq v_h\in V_h}\frac{A(\mathcal{P}_h(\mathcal{I} - \mathcal{A}^{-1}\mathcal{N}'[\hat{u}])u_h, v_h)}{\|u_h\|_V\|v_h\|_V}\\ &= \inf_{0\neq u_h\in V_h}\sup_{0\neq v_h\in V_h}\frac{A((\mathcal{I} - \mathcal{A}^{-1}\mathcal{N}'[\hat{u}])u_h, v_h)}{\|u_h\|_V\|v_h\|_V}\end{aligned}$$

$$
\begin{aligned}
&= \inf_{0\neq u_h\in V_h}\sup_{0\neq v_h\in V_h}\frac{(\nabla u_h,\nabla v_h)-(f'(\hat{u})u_h,v_h)}{\|u_h\|_V\|v_h\|_V}\\
&= \inf_{0\neq \mathrm{x}\in\mathbb{R}^n}\sup_{0\neq \mathrm{y}\in\mathbb{R}^n}\frac{\mathrm{x}^TB\mathrm{y}}{|\mathrm{x}^TS\mathrm{x}|^{1/2}|\mathrm{y}^TS\mathrm{y}|^{1/2}}.
\end{aligned}
$$

Since S is symmetric positive definite, there exists a lower triangular matrix L forming the Cholesky decomposition, $S = LL^T$. The Schwartz inequality $\mathrm{a}^T\mathrm{b} \leq |\mathrm{a}^T\mathrm{a}|^{1/2}|\mathrm{b}^T\mathrm{b}|^{1/2}$ holds for a, $\mathrm{b}\in\mathbb{R}^n$; in particular, the equal sign holds if a = b. Thus, for a fixed $\mathrm{a}\in\mathbb{R}^n$, it follows that

$$\sup_{\mathrm{b}\in\mathbb{R}^n}\frac{\mathrm{a}^T\mathrm{b}}{|\mathrm{b}^T\mathrm{b}|^{1/2}} = |\mathrm{a}^T\mathrm{a}|^{1/2}.$$

By using this equation for $\tilde{\mathrm{y}} := L^T\mathrm{y}$,

$$
\begin{aligned}
\eta &= \inf_{0\neq \mathrm{x}\in\mathbb{R}^n}\sup_{0\neq \mathrm{y}\in\mathbb{R}^n}\frac{\mathrm{x}^TB(L^{-T}L^T)\mathrm{y}}{|\mathrm{x}^TS\mathrm{x}|^{1/2}|\mathrm{y}^T(LL^T)\mathrm{y}|^{1/2}}\\
&= \inf_{0\neq \mathrm{x}\in\mathbb{R}^n}\sup_{0\neq \tilde{\mathrm{y}}\in\mathbb{R}^n}\frac{(\mathrm{x}^TBL^{-T})\tilde{\mathrm{y}}}{|\mathrm{x}^TS\mathrm{x}|^{1/2}|\tilde{\mathrm{y}}^T\tilde{\mathrm{y}}|^{1/2}}\\
&= \inf_{0\neq \mathrm{x}\in\mathbb{R}^n}\frac{|(L^{-1}B^T\mathrm{x})^T(L^{-1}B^T\mathrm{x})|^{1/2}}{|\mathrm{x}^TS\mathrm{x}|^{1/2}}\\
&= \inf_{0\neq \mathrm{x}\in\mathbb{R}^n}\frac{|\mathrm{x}^TBS^{-1}B^T\mathrm{x}|^{1/2}}{|\mathrm{x}^TS\mathrm{x}|^{1/2}}.
\end{aligned}
$$

This is simply the Rayleigh quotient form of a general matrix eigenvalue problem. Thus, η^2 is the smallest eigenvalue of

$$\text{Find } \lambda\in\mathbb{R},\ \mathrm{x}\in\mathbb{R}^n,\ \text{s.t. } BS^{-1}B^T\mathrm{x} = \lambda S\mathrm{x}.$$

We now discuss how to obtain a rigorous upper bound of τ by verified numerical computation. For a matrix $A\in\mathbb{R}^{n\times n}$, we define

$$\lambda_{\min}(A) := \min\{|\lambda| : \lambda\in\mathrm{Spec}(A)\},\quad \lambda_{\max}(A) := \max\{|\lambda| : \lambda\in\mathrm{Spec}(A)\},$$

where $\mathrm{Spec}(A)$ is the set of eigenvalues of A. Furthermore, let $\sigma_{\min}(A)$ be the minimum singular value of A. For $A, B\in\mathbb{R}^{n\times n}$,

$$\sigma_{\min}(A)\leq\lambda_{\min}(A),\quad \sigma_{\min}(AB)\geq\sigma_{\min}(A)\sigma_{\min}(B).$$

Since $\tau = \eta^{-1}$, the lower bound of η gives the upper bound of τ. As an efficient method of evaluating the lower bound of η by verified numerical computation, we use the following lemma, which effectively exploits the sparsity of B and S.

Lemma 5 (Rump 2011 [24]). *Let $\gamma > 0$ be an estimate of the lower bound $\sigma_{\min}(S^{-1}B)$. Check*

$$BB^T - \gamma^2S^2 \succeq 0, \tag{36}$$

where $A\succ 0$ ($\succeq 0$) means that $A\in\mathbb{R}^{n\times n}$ is symmetric positive (semi-)definite. If condition (36) is satisfied, then

$$\sigma_{\min}(S^{-1}B)\geq\gamma > 0. \tag{37}$$

Note that (36) can be checked using Rump's method (`isspd`) [23] by performing the sparse Cholesky decomposition once with the floating-point arithmetic. The sparse Cholesky decomposition algorithm is stable and efficient.

Now, let us consider the case of $B\in\mathbb{R}^{n\times n}$ being symmetric. In this case, from (37), we have

$$\eta = \lambda_{\min}(S^{-1}BS^{-1}B^T)^{1/2} = \lambda_{\min}(S^{-1}B)\geq\sigma_{\min}(S^{-1}B)\geq\gamma.$$

The upper bound of τ is evaluated as $\tau = \eta^{-1}\leq\gamma^{-1}$.

Next, let us consider the case of $B \in \mathbb{R}^{n\times n}$ being general. In this case, we have

$$\eta = \lambda_{\min}(S^{-1}BS^{-1}B^T)^{1/2} \geq \sigma_{\min}(S^{-1}BS^{-1}B^T)^{1/2} \geq \left(\sigma_{\min}(S^{-1}B)\sigma_{\min}(S^{-1}B^T)\right)^{1/2}. \tag{38}$$

If we further check

$$B^TB - \gamma'^2 S^2 \succeq 0$$

as above, then

$$\sigma_{\min}(S^{-1}B^T) \geq \gamma'$$

holds, so that it follows from (38) that

$$\tau = \eta^{-1} \leq \frac{1}{\sqrt{\gamma\gamma'}}.$$

4.2 Residual bounds

In this part, we consider how to perform residual evaluation (11) such that

$$\begin{aligned}\|\mathcal{F}(\hat{u})\|_{V^*} &= \sup_{0\neq v\in V}\frac{|\langle \mathcal{A}\hat{u} - \mathcal{N}(\hat{u}), v\rangle|}{\|v\|_V}\\ &= \sup_{0\neq v\in V}\frac{|(\nabla\hat{u}, \nabla v) - (f(\hat{u}), v)|}{\|v\|_V}\end{aligned}$$

using a smoothing technique with the Raviart-Thomas mixed finite element. Here, we introduce the Raviart-Thomas mixed finite element [2, 4, 21]. We follow the discussions in [2, 4]. Let $H(\mathrm{div}, \Omega)$ denote the space of vector functions such that

$$H(\mathrm{div}, \Omega) := \left\{\psi \in (L^2(\Omega))^2 : \mathrm{div}\ \psi \in L^2(\Omega)\right\}.$$

Let K_h be a triangle element in the triangulation of Ω. We define

$$P_k(K_h)\ :\ \text{the space of polynomials of degree less than or equal to } k \text{ on } K_h,$$

$$R_k(\partial K_h) := \{\varphi \in L^2(\partial K_h) : \varphi|_{e_i} \in P_k(e_i)\}, \quad \text{for any edge } e_i \text{ of } \partial K_h.$$

For $k \geq 0$, we define

$$RT_k(K_h) \ := \ \left\{ q \in (L^2(K_h))^2 \ : \ q = \begin{pmatrix} a_k \\ b_k \end{pmatrix} + c_k \cdot \begin{pmatrix} x \\ y \end{pmatrix},\ a_k, b_k, c_k \in P_k(K_h) \right\}.$$

The dimension of $RT_k(K_h)$ is $(k+1)(k+3)$. We now introduce basic results about $RT_k(K_h)$ spaces.

Proposition 1. *Let e_i be a subtense of vertex i $(= 1, 2, 3)$ and $\vec{n}_{|e_i} = (n_1^{(i)}, n_2^{(i)})^T$ be an outward unit normal vector on boundary e_i. For $q \in RT_k(K_h)$, it follows that*

$$\begin{cases} \mathrm{div}\ q \in P_k(K_h), \\ q \cdot \vec{n}_{|e_i} \in R_k(\partial K_h). \end{cases}$$

Moreover, the divergence operator from $RT_k(K_h)$ onto $P_k(K_h)$ is surjective, i.e.,

$$\mathrm{div}(RT_k(K_h)) = P_k(K_h).$$

For the entire domain Ω, the Raviart-Thomas finite element space RT_k is given by

$$\begin{aligned} RT_k \ := \ \Big\{ & p_h \in (L^2(\Omega))^2 \ : \ p_h|_{K_h} = \begin{pmatrix} a_k \\ b_k \end{pmatrix} + c_k \cdot \begin{pmatrix} x \\ y \end{pmatrix},\ a_k, b_k, c_k \in P_k(K_h), \\ & p_h \cdot \vec{n} \text{ is continuous on the interelement boundaries.} \Big\}. \end{aligned} \tag{39}$$

This is a finite-dimensional subspace of $H(\mathrm{div}, \Omega)$. Furthermore, let us define

$$M_h := \{v \in L^2(\Omega) \ : \ v|_{K_h} \in P_k(K_h)\}. \tag{40}$$

It follows that $\mathrm{div}(RT_k) = M_h$ (cf. Chapter IV.1 of [4]).

4.2.1 Proposal bounds with RT_k element

For residual bound estimation, some smoothing techniques have been proposed in [17, 19, 28]. This part is dedicated to proposing another smoothing technique using mixed finite elements. One feature of the proposed method is that we can use the basic property of the Raviart-Thomas element, $\mathrm{div}(RT_k) = M_h$. For a given $f_h \in M_h$, this property enables us to define a subspace of RT_k by

$$W_{f_h} = \{\ p_h \in RT_k : \mathrm{div}\ p_h + f_h = 0 \text{ for } f_h \in M_h\ \}.$$

Furthermore, we define $v_h \in M_h$ by an orthogonal projection of $v \in L^2(\Omega)$ such that

$$(v - v_h, w_h) = 0, \quad \forall w_h \in M_h.$$

Assume that the error estimate given by

$$\|v - v_h\|_X \le C_{M_h} \|v\|_V \text{ for } v_h \in M_h$$

holds. Also, we define $f_h(\hat{u}) \in M_h$ by the projection of $f(\hat{u}) \in L^2(\Omega)$. Finally, we obtain the following evaluation of the residual bound using $p_h \in W_{f_h(\hat{u})}$:

$$\begin{aligned}
\|\mathcal{F}(\hat{u})\|_{V^*} &= \sup_{0\neq v\in V} \frac{|(\nabla\hat{u}, \nabla v) - (f(\hat{u}), v)|}{\|v\|_V} \\
&= \sup_{0\neq v\in V} \frac{|(\nabla\hat{u} - p_h, \nabla v) + (p_h, \nabla v) - (f(\hat{u}), v)|}{\|v\|_V} \\
&\le \sup_{0\neq v\in V} \frac{|(\nabla\hat{u} - p_h, \nabla v)|}{\|v\|_V} + \sup_{0\neq v\in V} \frac{|(\mathrm{div}\ p_h + f(\hat{u}), v)|}{\|v\|_V} \\
&\le \|\nabla\hat{u} - p_h\|_X + \sup_{0\neq v\in V} \frac{|(\mathrm{div}\ p_h + f_h(\hat{u}) + f(\hat{u}) - f_h(\hat{u}), v)|}{\|v\|_V} \\
&\le \|\nabla\hat{u} - p_h\|_X + \sup_{0\neq v\in V} \frac{|(f(\hat{u}) - f_h(\hat{u}), v - v_h)|}{\|v\|_V} \\
&\le \|\nabla\hat{u} - p_h\|_X + C_{M_h}\ \|f(\hat{u}) - f_h(\hat{u})\|_X =: C_{2,h}.
\end{aligned} \tag{41}$$

Remark 5. *The proposed estimation (41) holds for $k \ge 0$. If the approximate solution $\hat{u}$ is obtained from V_h, which consists of a piecewise polynomial of degree $(k+1)$, an effective choice of the smoothing element is from RT_k and M_h is spanned by P_k elements. The rate of convergence can be expected to be $\|\nabla\hat{u} - p_h\|_X = O(h^{k+1})$ and $\|f(\hat{u}) - f_h(\hat{u})\|_X = O(h^{k+1})$ in the case that the solution has sufficient regularity. Further, C_{M_h} is bounded by $C_{h,0}$ defined in (20), although tighter evaluation will be expected for $k > 0$.*

4.2.2 How to determine p_h

This part is devoted to explaining a procedure for determining the smoothing element $p_h \in W_{f_h(\hat{u})}$. Using a verified numerical computation of linear equations, we have the interval function $\tilde{p}_h$. This includes the smoothing element $p_h \in \tilde{p}_h$ with verification. The mixed method for Poisson's equation is applied to our procedure. First, we write the original problem (1) as

$$\begin{cases} \nabla u = p, \\ -\mathrm{div}\ p = f(u). \end{cases}$$

This system leads directly to the following saddle point problem: Find $(p, u) \in H(\mathrm{div}, \Omega) \times X$ such that

$$\begin{cases} (p, q) + (u, \mathrm{div}\ q) &= 0, & \forall q \in H(\mathrm{div}, \Omega), \\ (\mathrm{div}\ p, v) &= -(f(u), v), & \forall v \in X. \end{cases} \tag{42}$$

Since the inf-sup condition of the general saddle point framework is obtained [2], this saddle point problem has the solution $(p, u) \in H(\mathrm{div}, \Omega) \times X$. Let M_h be defined in (40). As mentioned above, we determine $f_h(\hat{u}) \in M_h$ such that

$$(f(\hat{u}) - f_h(\hat{u}), v_h) = 0, \quad \forall v_h \in M_h.$$

To obtain $p_h \in W_{f_h(\hat{u})}$ for a given $f_h(\hat{u})$, we consider an approximation of the problem (42). We seek $(p_h, u_h) \in RT_k \times M_h$ defined in (39) and (40) satisfying

$$\left\{ \begin{array}{lcll} (p_h, q_h) + (u_h, \text{div } q_h) & = & 0, & \forall q_h \in RT_k, \\ (\text{div } p_h, v_h) & = & -(f(\hat{u}), v_h), & \forall v_h \in M_h. \end{array} \right. \tag{43}$$

Suppose $\{\psi_i\}$ and $\{q_i\}$ are the base functions of RT_k and M_h, respectively, that is,

$$RT_k = \text{span}\{\psi_1, ..., \psi_l\}, \ M_h = \text{span}\{q_1, ..., q_m\}.$$

Recall the matrices $P \in \mathbb{R}^{l \times l}$ and $N \in \mathbb{R}^{m \times l}$ on page 45. Additionally, suppose that $\mathrm{x} \in \mathbb{R}^l$, $\mathrm{z} \in \mathbb{R}^m$ and $\mathrm{f} \in \mathbb{R}^m$ are vectors. Using these notations, let $p_h \in RT_k$, $u_h \in M_h$ be elements described by

$$\begin{array}{rl} \mathrm{x} = (x_1, ..., x_l)^T \in \mathbb{R}^l, & p_h = (\psi_1, ..., \psi_l) \cdot \mathrm{x} \in RT_k, \\ \mathrm{z} = (z_1, ..., z_m)^T \in \mathbb{R}^m, & u_h = (q_1, ..., q_l) \cdot z \in M_h, \\ \mathrm{f} = [(f(\hat{u}), q_i)]_{i=1,...,m} \in \mathbb{R}^m. & \end{array}$$

By using matrix notations, problem (43) is finally characterized by

$$\left\{ \begin{array}{lcl} P\mathrm{x} + N^T \mathrm{z} & = & 0, \\ N\mathrm{x} & = & -\mathrm{f}. \end{array} \right.$$

To obtain $p_h \in W_{f_h(\hat{u})}$, we need to obtain the vector $\mathrm{x} \in \mathbb{R}^l$ with verified numerical computations. Here, we will use a basic algorithm to solve linear equations:

$$\text{Find } \mathrm{x}_z \in \mathbb{R}^{l+m} \text{ s.t. } \tilde{A}\mathrm{x}_z = \tilde{\mathrm{f}}, \quad \tilde{A} = \begin{pmatrix} P & N^T \\ N & 0 \end{pmatrix}, \quad \mathrm{x}_z = \begin{pmatrix} \mathrm{x} \\ \mathrm{z} \end{pmatrix}, \quad \tilde{\mathrm{f}} = \begin{pmatrix} 0 \\ -\mathrm{f} \end{pmatrix}.$$

The solution x_z is enclosed by verified numerical computations.

4.3 Lipschitz constant

Finally, we estimate the Lipschitz constant of $\mathcal{F}'[u] : V \to V^*$. Here, we assume that $f' : V \to \mathcal{L}(V, X)$ is Lipschitz continuous on the open ball $D \supset \overline{B}(\hat{u}, 2\alpha)$. Namely, there exists a positive constant C_L satisfying

$$|((f'(v) - f'(w))u, \psi)| \leq C_L \|v - w\|_V \|u\|_V \|\psi\|_V \tag{44}$$

for $v, w \in D$ and $u, \psi \in V$. Generally, the optimal estimation depends on the definition of f. We will discuss the estimation of C_L in Section 5 for a model case. For $v, w \in D$, we have

$$\begin{aligned} \|\mathcal{F}'[v] - \mathcal{F}'[w]\|_{V,V^*} &= \sup_{0 \neq u \in V} \sup_{0 \neq \psi \in V} \frac{|\langle (\mathcal{N}'[v] - \mathcal{N}'[w])u, \psi \rangle|}{\|u\|_V \|\psi\|_V} \\ &= \sup_{0 \neq u \in V} \sup_{0 \neq \psi \in V} \frac{|((f'(v) - f'(w))u, \psi)|}{\|u\|_V \|\psi\|_V} \\ &\leq C_L \|v - w\|_V. \end{aligned}$$

Therefore, one can put $C_3 := C_L$.

5. Computational results

To summarize this paper, we show our computational results. In the following, we present elliptic boundary problems on several polygonal domains. Firstly, let us consider a practical formulation of a certain example such that

$$\left\{ \begin{array}{ll} -\Delta u = f(u), & \text{in } \Omega, \\ u = 0, & \text{on } \partial\Omega \end{array} \right.$$

with

$$f(u) = b \cdot \nabla u + c_1 u + c_2 u^2 + c_3 u^3 + g.$$

Here, $b(x) \in (L^\infty(\Omega))^2$, $c_i \in L^\infty(\Omega)$, $(i = 1, 2, 3)$ and $g \in X$. To show the applicability of our verification theory to this problem, we must check whether f is Fréchet differentiable at $\hat{u} \in V_h$ as a map $f : V \to X$. This can be shown as follows. A candidate of $f'(\hat{u}) : V \to X$ is obviously

$$f'(\hat{u}) = b \cdot \nabla + c_1 + 2c_2\hat{u} + 3c_3\hat{u}^2.$$

Sobolev's embedding theorem states that $V \subset L^p(\Omega)$ for $p \in [1, \infty)$ with $\|v\|_{L^p} \le C_{e,p}\|v\|_V, \ \forall v \in V$. Recall the generalized Hölder inequality, cf. page 93 in [3], that is,

$$\|uvw\|_{L^{p_0}} \le \|u\|_{L^{p_1}}\|v\|_{L^{p_2}}\|w\|_{L^{p_3}}, \text{ for } \frac{1}{p_0} = \frac{1}{p_1} + \frac{1}{p_2} + \frac{1}{p_3} \le 1, \ 1 \le p_i \le \infty.$$

For $u, v, w \in V$, it follows that

$$\|uv\|_X \le \|u\|_{L^4}\|v\|_{L^4} \le C_{e,4}^2\|u\|_V\|v\|_V$$

and

$$\|uvw\|_X \le \|u\|_{L^6}\|v\|_{L^6}\|w\|_{L^6} \le C_{e,6}^3\|u\|_V\|v\|_V\|w\|_V.$$

Then, we have for $\nu \in V$

$$\begin{aligned}
\|f(\hat{u}+\nu) - f(\hat{u}) - f'(\hat{u})\nu\|_X &= \|(c_2 + 3c_3\hat{u})\nu^2 + c_3\nu^3\|_X \\
&\le \|c_2\|_{L^\infty}\|\nu^2\|_X + 3\|c_3\|_{L^\infty}\|\hat{u}\nu^2\|_X + \|c_3\|_{L^\infty}\|\nu^3\|_X \\
&\le \left(C_{e,4}^2\|c_2\|_{L^\infty} + C_{e,6}^3\|c_3\|_{L^\infty}(3\|\hat{u}\|_V + \|\nu\|_V)\right)\|\nu\|_V^2.
\end{aligned}$$

This shows the Fréchet differentiability of $f : V \to X$ at $\hat{u} \in V_h$.

For the inverse operator norm estimation, we need the following constants. We can assume that the computation result $\hat{u} \in V_h$ is essentially bounded so that $\hat{u} \in L^\infty(\Omega) \cap V$ is obtained.

$$\begin{aligned}
\|f'(\hat{u})\|_{V,X} &= \sup_{0 \ne v \in V} \frac{\|f'(\hat{u})v\|_X}{\|v\|_V} \\
&= \sup_{0 \ne v \in V} \frac{\|b \cdot \nabla v + c_1 v + 2c_2\hat{u}v + 3c_3\hat{u}^2 v\|_X}{\|v\|_V} \\
&\le \||b|_{l^2}\|_{L^\infty} + C_{e,2}\left(\|c_1\|_{L^\infty} + 2\|c_2\|_{L^\infty}\|\hat{u}\|_{L^\infty} + 3\|c_3\|_{L^\infty}\|\hat{u}\|_{L^\infty}^2\right) \\
&=: K_1,
\end{aligned}$$

where $b = (b_1, b_2)^T$ and $|b|_{l^2} = (b_1^2 + b_2^2)^{\frac{1}{2}}$. Furthermore, we have

$$\begin{aligned}
\|f'(\hat{u})\|_{V_c,X} &= \sup_{0 \ne v_c \in V_c} \frac{\|f'(\hat{u})v_c\|_X}{\|v_c\|_V} \\
&= \sup_{0 \ne v_c \in V_c} \frac{\|b \cdot \nabla v_c + c_1 v_c + 2c_2\hat{u}v_c + 3c_3\hat{u}^2 v_c\|_X}{\|v_c\|_V} \\
&\le \||b|_{l^2}\|_{L^\infty} + C_M\left(\|c_1\|_{L^\infty} + 2\|c_2\|_{L^\infty}\|\hat{u}\|_{L^\infty} + 3\|c_3\|_{L^\infty}\|\hat{u}\|_{L^\infty}^2\right) =: K_2.
\end{aligned}$$

Here, C_M is the quantity defined in (19). The explicit values of K_1 and K_2 are obtained by verified computations.

Let us describe the Lipschitz continuity of $\mathcal{F}'[\hat{u}] : V \to V^*$ by checking inequality (44). $\overline{B}(\hat{u}, 2\alpha)$ is assumed to be a closed ball centered at $\hat{u} \in V_h$ with radius $2\alpha := 2C_1C_{2,h}$. Select D as

$$D := \{v \in V : \|v - \hat{u}\|_V < 2\alpha + \varepsilon\} \supset \overline{B}(\hat{u}, 2\alpha)$$

with a small $\varepsilon > 0$. For $v, w \in D$ and $u, \psi \in V$, we have

$$|((f'(v) - f'(w))u, \psi)| = |(2c_2(v-w)u, \psi) + (3c_3(v+w)(v-w)u, \psi)|$$

$$
\begin{aligned}
\leq\quad & 2\|c_2\|_{L^\infty}\cdot\|v-w\|_{L^3}\cdot\|u\|_{L^3}\cdot\|\psi\|_{L^3}\\
& \qquad\qquad + 3\|c_3\|_{L^\infty}\cdot\|v+w\|_{L^4}\cdot\|v-w\|_{L^4}\cdot\|u\|_{L^4}\cdot\|\psi\|_{L^4}\\
\leq\quad & \left(2C_{e,3}^3\|c_2\|_{L^\infty}+3C_{e,4}^4\|c_3\|_{L^\infty}\|v+w\|_V\right)\|v-w\|_V\cdot\|u\|_V\cdot\|\psi\|_V.
\end{aligned}
$$

Since $v, w \in D$,

$$\|v+w\|_V < \ 2(\|\hat{u}\|_V + 2C_1C_{2,h} + \varepsilon).$$

Thus, it follows that

$$|((f'(v)-f'(w))u,\psi)| < C_L\|v-w\|_V\cdot\|u\|_V\cdot\|\psi\|_V, \quad \text{for } v,w\in D$$

with

$$C_L := 2\ C_{e,3}^3\|c_2\|_{L^\infty} + 6\ C_{e,4}^4\|c_3\|_{L^\infty}\left(\|\hat{u}\|_V + 2C_1C_{2,h}+\varepsilon\right).$$

5.1 On square domains

Next, we will present numerical results on square domains. All computations are carried out on a Cent OS (Linux), Quad-Core AMD Opteron(tm) Processor 8376 of 2.30 GHz with 512 GB RAM using MATLAB 2011a with INTLAB, a toolbox for verified numerical computations [22]. To obtain a triangular mesh, we use Gmsh [5] (`http://geuz.org/gmsh/`).

5.1.1 Example 1

Let us consider the following semilinear Dirichlet boundary value problem on $\Omega = (0,1)\times(0,1)$:

$$\begin{cases} -\Delta u = u^2, & \text{in } \Omega,\\ u = 0, & \text{on } \partial\Omega. \end{cases} \tag{45}$$

An approximate solution $\hat{u}\in V_h$ is calculated using the quadratic conforming finite elements on a nonuniform mesh triangulation. In the rest of this paper, C_M is evaluated by the method described in Section 3.1. Since the linear finite element space is the subspace of the quadratic one, the quadratic finite element provides absolutely better approximation of the exact solution. Therefore, for our current computation, we can use the projection error constant C_M corresponding to the linear finite element, which is easy to evaluate. We measure the mesh size using the maximum medium edge length for each element. For mesh sizes of 1/16 and 1/32, the maximum $\hat{u}$ is about $\|\hat{u}\|_\infty \approx 29.247$. Our verification procedure is applied to (45). When the mesh size is 1/32, it gives the following bounds:

$$C_1 \leq 3.121,\ C_{2,h}\leq 5.929\times 10^{-2},\ C_3\leq 7.165\times 10^{-2}.$$

Thus, it holds that

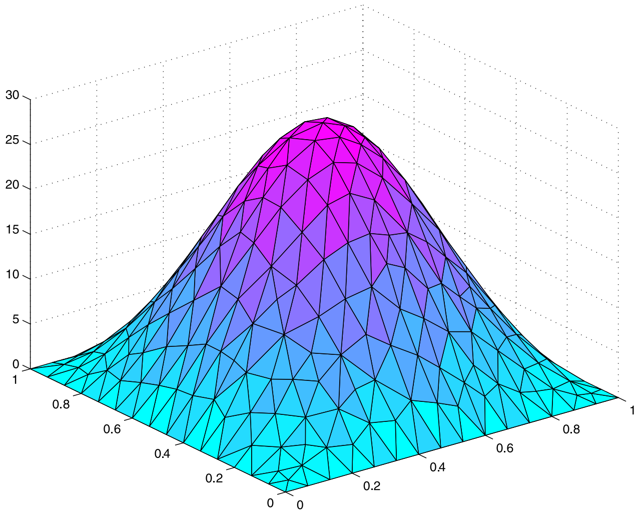

Fig. 3. Approximate solution $\hat{u}$ of (45), mesh size: 1/16.

$$C_1^2 C_{2,h} C_3 \leq 4.132 \times 10^{-2}.$$

It turns out that there exists a solution in the closed ball $\overline{B}(\hat{u}, \rho)$ with

$$\|u - \hat{u}\|_V \leq \rho = 1.889 \times 10^{-1}.$$

By increasing the number of grid points, guaranteed error bounds are improved. The convergence rate of the error depends on the ratio of $C_{2,h}$. Using the residual evaluation (41), it is expected to be $O(h^2)$. The guaranteed error bound is presented in Table I.

Table I. Verification results for (45).

$1/2^\gamma$	C_M	$C_{e,2}$	C_1	$C_{2,h}$	C_3	$C_1^2C_{2,h}C_3$	ρ
3	5.37713×10^{-2}	2.25079×10^{-1}	Failed	-	-	-	-
4	2.34505×10^{-2}	2.25079×10^{-1}	4.2711	2.91265×10^{-1}	7.16449×10^{-2}	3.80671×10^{-1}	1.67147
5	1.29561×10^{-2}	2.25079×10^{-1}	3.1191	5.92838×10^{-2}	7.16449×10^{-2}	4.13197×10^{-2}	1.88893×10^{-1}
6	6.37492×10^{-3}	2.25079×10^{-1}	2.8319	1.50966×10^{-2}	7.16449×10^{-2}	8.67379×10^{-3}	4.29384×10^{-2}

5.1.2 Example 2

Let us treat another semilinear Dirichlet boundary value problem on $\Omega = (0,1) \times (0,1)$:

$$\begin{cases} -\Delta u = u^3 + 5, & \text{in } \Omega, \\ u = 0, & \text{on } \partial\Omega. \end{cases} \tag{46}$$

The approximate solutions $\hat{u} \in V_h$ are calculated using the quadratic conforming finite element on a nonuniform mesh. We have three approximate solutions of (46): $\hat{u}_1$ ($\|\hat{u}_1\|_\infty \approx 6.263$), $\hat{u}_0$ ($\|\hat{u}_0\|_\infty \approx 0.371$) and $\hat{u}_{-1}$ ($\|\hat{u}_{-1}\|_\infty \approx 6.962$). Their shapes are shown in Fig. 4. For the approximation $\hat{u}_0$ with a mesh size of 1/8, our computer-assisted proof method yields the following bounds:

$$C_1^2 C_{2,h} C_3 \leq 1.348 \times 10^{-2}.$$

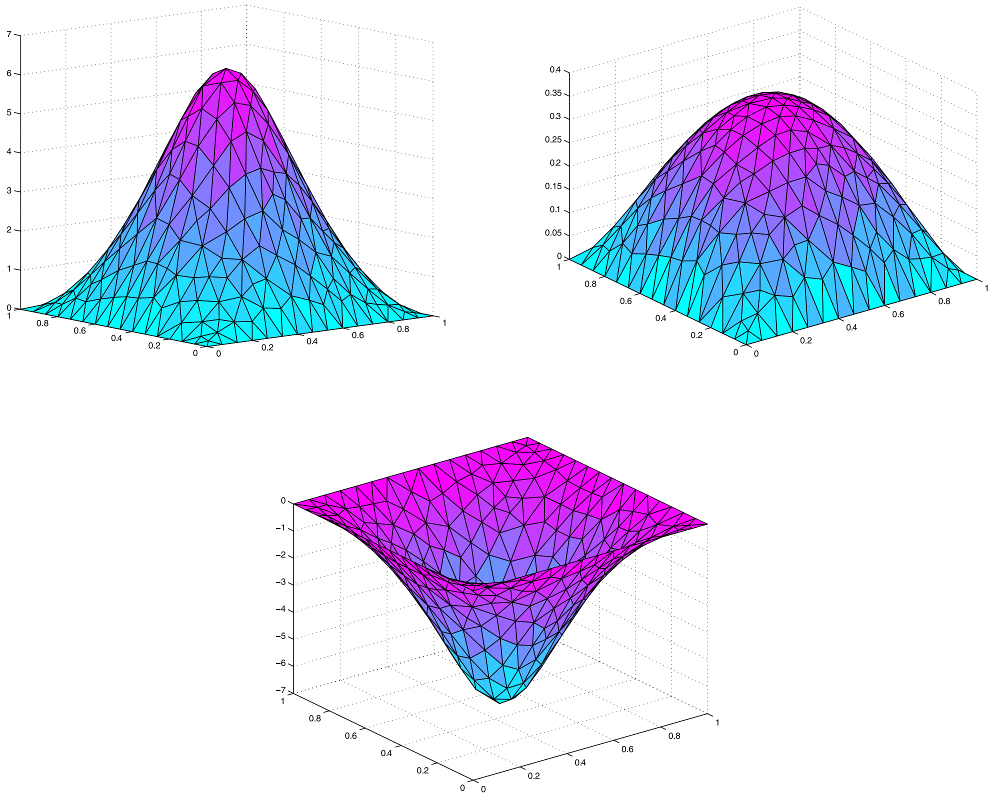

Fig. 4. Approximate solutions $\hat{u}_1$, $\hat{u}_0$ and $\hat{u}_{-1}$ of (46).

Thus, we can state that there exists an exact solution u in the closed ball $\overline{B}(\hat{u}_0, \rho_0)$ with

$$\|u_0 - \hat{u}_0\|_V \le \rho_0 = 2.231 \times 10^{-2}.$$

Guaranteed error bounds are improved by decreasing the mesh size h presented in Table II. Here, guaranteed error bounds are represented by $\|u_{-1} - \hat{u}_{-1}\|_V \le \rho_{-1}$, $\|u_0 - \hat{u}_0\|_V \le \rho_0$ and $\|u_1 - \hat{u}_1\|_V \le \rho_1$.

Table II. Verification results for (46).

$1/2^\gamma$	C_M	$C_{e,2}$	C_1	$C_{2,h}$	C_3	$C_1^2 C_{2,h} C_3$	ρ_{-1}
5	1.29082×10^{-2}	2.25079×10^{-1}	560.26	2.25048×10^{-2}	23.0453	1.62792×10^{5}	Failed
6	6.37492×10^{-3}	2.25079×10^{-1}	3.2627	5.99126×10^{-3}	7.73894	4.93563×10^{-1}	3.51111×10^{-2}

$1/2^\gamma$	C_M	$C_{e,2}$	C_1	$C_{2,h}$	C_3	$C_1^2 C_{2,h} C_3$	ρ_0
3	4.84064×10^{-2}	2.25079×10^{-1}	1.0155	2.18169×10^{-2}	5.98946×10^{-1}	1.34741×10^{-2}	2.23053×10^{-2}
4	2.34612×10^{-2}	2.25079×10^{-1}	1.0144	5.60044×10^{-3}	5.78994×10^{-1}	3.33647×10^{-3}	5.69042×10^{-3}
5	1.16517×10^{-2}	2.25079×10^{-1}	1.0141	1.47478×10^{-3}	5.73911×10^{-1}	8.70438×10^{-4}	1.49623×10^{-3}

$1/2^\gamma$	C_M	$C_{e,2}$	C_1	$C_{2,h}$	C_3	$C_1^2 C_{2,h} C_3$	ρ_1
5	1.29082×10^{-2}	2.25079×10^{-1}	2.9506	1.67181×10^{-2}	7.27778	1.05924	Failed
6	5.70970×10^{-3}	2.25079×10^{-1}	2.0967	4.17847×10^{-3}	7.22844	1.32777×10^{-1}	9.43554×10^{-3}

5.1.3 Example 3

For $\Omega = (0, 1) \times (0, 1)$, let us consider another example:

$$\begin{cases} -\Delta u = b \cdot \nabla u + u + u^2, & \text{in } \Omega, \\ u = 0, & \text{on } \partial\Omega, \end{cases} \tag{47}$$

where $b(x) = (1, 1)^T$. In this case, we have

$$\| |b(x)|_{l^2} \|_{L^\infty} = \sqrt{2}.$$

Figure 5 shows an approximate solution $\hat{u} \in V_h$ using the quadratic conforming finite element on a nonuniform mesh ($h = 1/16$). Our verification method yields

$$C_1 \le 4.559, \; C_{2,h} \le 6.488 \times 10^{-2}, \; C_3 \le 7.165 \times 10^{-2},$$

and

$$C_1^2 C_{2,h} C_3 \le 9.658 \times 10^{-2}.$$

The verified computation ensures that there exists an exact solution in the closed ball $\overline{B}(\hat{u}, \rho)$ with

$$\|u - \hat{u}\|_V \le \rho = 3.116 \times 10^{-1}.$$

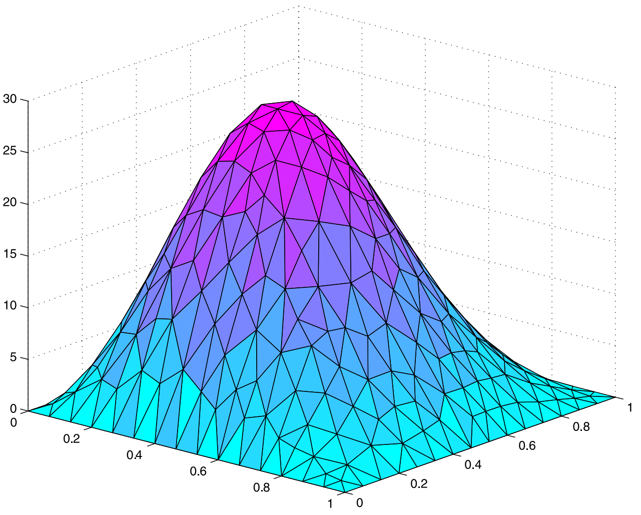

Fig. 5. Approximate solution $\hat{u}$ of (47).

Table III. Verification results for (47).

$1/2^\gamma$	C_M	$C_{e,2}$	C_1	$C_{2,h}$	C_3	$C_1^2C_{2,h}C_3$	ρ
4	2.60362×10^{-2}	2.25079×10^{-1}	11.517	2.87224×10^{-1}	7.16449×10^{-2}	2.72905	Failed
5	1.29082×10^{-2}	2.25079×10^{-1}	4.5584	6.48743×10^{-2}	7.16449×10^{-2}	9.65776×10^{-2}	3.11573×10^{-1}

5.2 On hexagonal domains

Let Ω be a hexagonal domain, whose coordinates of vertices are given by

$$(x_1,x_2)^T \in \left\{ \left(\sin\left(\frac{n\pi}{3}\right), \cos\left(\frac{n\pi}{3}\right) \right)^T \in \mathbb{R}^2 : n = 1,...,6 \right\}.$$

We consider the following Dirichlet boundary value problem

$$\begin{cases} -\Delta u = u^2 + 10, & \text{in } \Omega, \\ u = 0, & \text{on } \partial\Omega. \end{cases} \tag{48}$$

We pay attention to two approximate solutions $\hat{u}_1, \hat{u}_2 \in V_h$ given by the finite element method, which are shown in Fig. 6 and Fig. 7 with a mesh size of 1/16.

Here, we give a detailed description of an advantage of the residual evaluation (41). In Tables IV and V show the computational results of the approximate solution $\hat{u}_1 \in V_h$ based on the linear and quadratic conforming finite elements, respectively. In the residual evaluation (41), we adopt $p_h \in RT_0$ for $\hat{u}$ obtained using the linear finite element and $p_h \in RT_1$ for $\hat{u}$ obtained using the quadratic finite

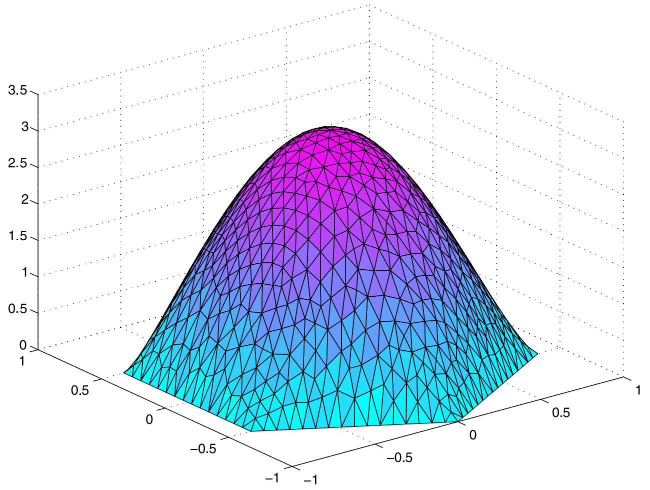

Fig. 6. Approximate solution $\hat{u}_1$ of (48).

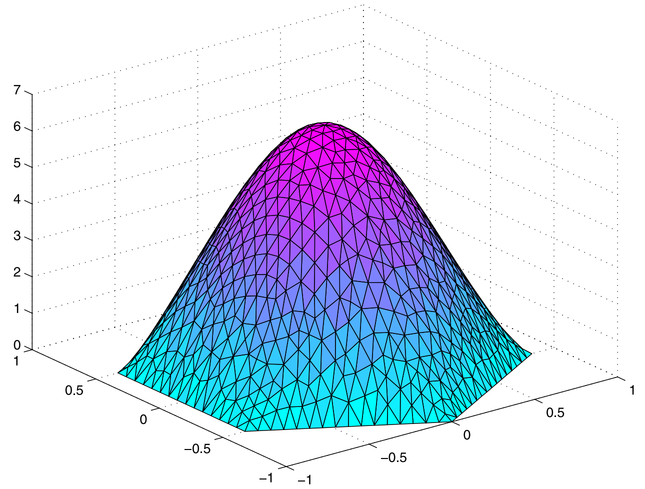

Fig. 7. Approximate solution $\hat{u}_2$ of (48).

Table IV. Results for $\hat{u}_1$ obtained using linear finite element with $p_h \in RT_0$.

$1/2^\gamma$	C_M	$C_{e,2}$	C_1	$C_{2,h}$	C_3	$C_1^2 C_{2,h} C_3$	ρ_1
3	6.288×10^{-2}	3.777×10^{-1}	3.667	9.003×10^{-1}	2.018×10^{-1}	2.441	Failed
4	3.231×10^{-2}	3.749×10^{-1}	3.477	4.609×10^{-1}	1.988×10^{-1}	1.107	Failed
5	1.886×10^{-2}	3.743×10^{-1}	3.404	2.248×10^{-1}	1.981×10^{-1}	5.155×10^{-1}	Failed
6	8.745×10^{-3}	3.741×10^{-1}	3.334	1.131×10^{-1}	1.978×10^{-1}	2.482×10^{-1}	4.403×10^{-1}
7	4.819×10^{-3}	3.739×10^{-1}	3.308	5.662×10^{-2}	1.977×10^{-1}	1.224×10^{-1}	2.004×10^{-1}

Table V. Results for $\hat{u}_1$ obtained using quadratic finite element with $p_h \in RT_1$.

$1/2^\gamma$	C_M	$C_{e,2}$	C_1	$C_{2,h}$	C_3	$C_1^2 C_{2,h} C_3$	ρ_1
3	5.776×10^{-2}	3.779×10^{-1}	3.821	1.255×10^{-1}	2.019×10^{-1}	3.699×10^{-1}	6.351×10^{-1}
4	2.823×10^{-2}	3.749×10^{-1}	3.496	4.479×10^{-2}	1.987×10^{-1}	1.088×10^{-1}	1.662×10^{-1}
5	2.138×10^{-2}	3.745×10^{-1}	3.437	1.491×10^{-2}	1.983×10^{-1}	3.491×10^{-2}	5.216×10^{-2}

element. Comparing two cases in Tables IV and V, we can observe that higher-order elements yield greatly improved results.

Next, we present results with respect to $\hat{u}_2$, which is from the quadratic finite element space. Table VI presents verified results for $\hat{u}_2$. Moreover, we compare three residual evaluation method. In Table VII, the first column denotes the result of the residual evaluation reported in [26]. The second column uses a refinement technique for residual evaluation, which was reported in [25, 28]. The third column denotes the method proposed in this paper. The comparison in Table VII implies that our proposed method enables much better estimation. Numerical values in the last column in Table VII express the upper bound of the absolute error ρ_2 using the residual bounds in (41).

Table VI. Results for $\hat{u}_2$ obtained using quadratic finite element with $p_h \in RT_1$.

$1/2^\gamma$	C_M	$C_{e,2}$	C_1	$C_{2,h}$	C_3	$C_1^2 C_{2,h} C_3$	ρ_2
3	5.776×10^{-2}	3.779×10^{-1}	5.167	1.966×10^{-1}	2.019×10^{-1}	1.061	Failed
4	2.823×10^{-2}	3.749×10^{-1}	3.805	6.715×10^{-2}	1.987×10^{-1}	1.931×10^{-1}	2.865×10^{-1}
5	2.138×10^{-2}	3.745×10^{-1}	3.642	2.204×10^{-2}	1.983×10^{-1}	5.794×10^{-2}	8.272×10^{-2}

Table VII. Comparison of residual evaluation methods.

$1/2^\gamma$	[26]	[25, 28]	(41)	$\|u - \hat{u}_2\|_V \le \rho_2$
3	11.18	0.7509	0.1966	Failed
4	5.467	0.3796	0.0672	0.2865
5	4.141	0.2776	0.0221	0.0828

5.3 On nonconvex domains

Another example is the case in which Ω is a nonconvex domain. Let us consider the Dirichlet boundary problem of the form

$$\begin{cases} -\Delta u = u^2 + 10, & \text{in } \Omega, \\ u = 0, & \text{on } \partial\Omega, \end{cases} \tag{49}$$

on $\Omega = (0,2)^2 \setminus [1,2]^2$ which is an L-shaped domain. An approximate solution $\hat{u} \in V_h$ of (49) is shown in Fig. 8 with a mesh size of 1/16. Verification results are shown in Table VIII.

Using the Raviart-Thomas mixed finite element, C_M is calculated by the procedure given by Theorem 5. The convergence rate of C_M becomes less than $O(h)$. This is caused by the lack of H^2 regularity. An undesirable situation with respect to the ratio of $C_{2,h}$ is similarly obtained for the same reason. Here, $C_{2,h}$ uses the evaluation in (41) by the P_2-RT_1 smoothing technique, which means the approximate solution is spanned by quadratic finite elements and p_h is chosen from RT_1. Although the convergence rate is low, there is a unique solution in the error bound ρ based on the Newton-Kantorovich theorem. In the case that the mesh size is 1/8, we have

$$C_1^2 C_{2,h} C_3 \le 1.541 \times 10^{-1}.$$

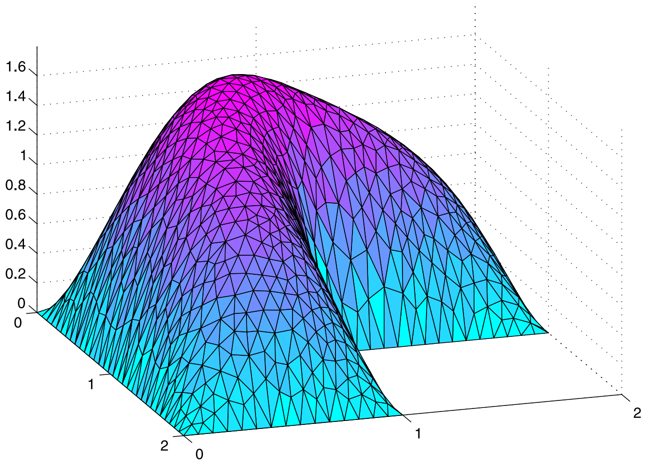

Fig. 8. Approximate solution $\hat{u}$ of (49).

Table VIII. Verification results for (49) on L-shaped domain.

$1/2^{\gamma}$	C_M	$C_{e,2}$	C_1	$C_{2,h}$	C_3	$C_1^2 C_{2,h} C_3$	ρ
3	7.33147×10^{-2}	3.29944×10^{-1}	1.4591	3.51574×10^{-1}	1.53955×10^{-1}	1.15242×10^{-1}	5.46550×10^{-1}
4	3.58873×10^{-2}	3.23931×10^{-1}	1.4216	2.17645×10^{-1}	1.48396×10^{-1}	6.52736×10^{-2}	3.20226×10^{-1}
5	1.89612×10^{-2}	3.22588×10^{-1}	1.4139	1.27178×10^{-1}	1.47167×10^{-1}	3.74146×10^{-2}	1.83309×10^{-1}

Thus, the radius of the ball $\bar{B}(\hat{u}, \rho)$ containing the exact solution is

$$\|u - \hat{u}\|_V \leq \rho = 1.153 \times 10^{-1}.$$

Acknowledgments

The authors express their sincere thanks to the referees for their valuable comments and their detailed review of our manuscript. This research was supported by a Grant-in-Aid for JSPS Fellows. The second author was partially supported by a Grant-in-Aid for Young Scientists (B) (No.23740092) from Japan Society for the Promotion of Science.

References

[1] R.A. Adams, *Sobolev Spaces*, Academic Press, New York, 1975.

[2] D. Braess, *Finite Elements*, 3rd edition, Cambridge University Press, Cambridge, 2007.

[3] H. Brézis, *Functional Analysis, Sobolev Spaces and Partial Differential Equations*, Universitext Series, Springer, 2010.

[4] F. Brezzi and M. Fortin, *Mixed and hybrid finite element methods*, vol. 15 of *Springer Series in Computational Mathematics*, Springer-Verlag, New York, 1991.

[5] C. Geuzaine and J.-F. Remacle, "Gmsh: A 3-d finite element mesh generator with built-in pre- and post-processing facilities," *Int. J. Numer. Meth. Eng.*, vol. 79, no. 11, pp. 1309–1331, 2009. `http://www.geuz.org/gmsh/`

[6] P. Deuflhard and G. Heindl, Affine invariant convergence theorems for newton's method and extensions to related methods, *SIAM Journal on Numerical Analysis*, vol. 16, no. 1, pp. 1–10, 1979.

[7] P. Grisvard, *Elliptic problems in nonsmooth domains*, Monographs and Studies in Mathematics, Pitman Advanced Pub. Program, 1985.

[8] L.V. Kantorovich and G.P. Akilov, *Functional analysis in normed spaces*, International Series of Monographs in Pure and Applied Mathematics, Pergamon Press, 1964.

[9] G. Kedem, "A posteriori error bounds for two-point boundary value problems," *SIAM Journal on Numerical Analysis*, vol. 18, no. 3, pp. 431–448, 1981.

[10] F. Kikuchi and X. Liu, “Estimation of interpolation error constants for the p0 and p1 triangular finite elements,” *Computer Methods in Applied Mechanics and Engineering*, vol. 196, no. 37–40, pp. 3750–3758, 2007.

[11] F. Kikuchi and H. Saito, “Remarks on a posteriori error estimation for finite element solutions,” *J. Comput. Appl. Math.*, vol. 199, no. 2, pp. 329–336, 2007.

[12] K. Kobayashi, “On the interpolation constants over triangular elements,” *RIMS Kokyuroku*, vol. 1733, pp. 58–77, 2011.

[13] X. Liu and S. Oishi, “Verified eigenvalue evaluation for the Laplacian over polygonal domains of arbitrary shape,” Submitted to SIAM on Numerical Analysis.

[14] M.A. McCarthy and R.A. Tapia, “Computable a posteriori l_∞-error bounds for the approximate solution of two-point boundary value problems,” *SIAM Journal on Numerical Analysis*, vol. 12, no. 6, pp. 919–937, 1975.

[15] M.T. Nakao, “A numerical approach to the proof of existence of solutions for elliptic problems,” *Japan J. Indust. Appl. Math.*, vol. 5, pp. 313–332, 1988.

[16] M.T. Nakao, K. Hashimoto, and Y. Watanabe, “A numerical method to verify the invertibility of linear elliptic operators with applications to nonlinear problems,” *Computing*, vol. 75, no. 1, pp. 1–14, 2005.

[17] M.T. Nakao and Y. Watanabe, “Numerical verification methods for solutions of semilinear elliptic boundary value problems,” *NOLTA*, vol. 2, no. 1, pp. 2–31, 2011.

[18] M. Plum, “Computer-assisted existence proofs for two-point boundary value problems,” *Computing*, vol. 46, pp. 19–34, 1991.

[19] M. Plum, “Computer-assisted proofs for semilinear elliptic boundary value problems,” *Japan J. Indust. Appl. Math.*, vol. 26, no. 2–3, pp. 419–442, 2009.

[20] W. Prager and J.L. Synge, “Approximations in elasticity based on the concept of function space,” *Quart. Appl. Math.*, vol. 5, pp. 241–269, 1947.

[21] P. Raviart and J. Thomas, “A mixed finite element method for 2-nd order elliptic problems,” In Ilio Galligani and Enrico Magenes, editors, *Mathematical Aspects of Finite Element Methods*, vol. 606 of *Lecture Notes in Mathematics*, pp. 292–315, Springer Berlin, Heidelberg, 1977.

[22] S.M. Rump, INTLAB - INTerval LABoratory, In Tibor Csendes, editor, *Developments in Reliable Computing*, pp. 77–104, Kluwer Academic Publishers, Dordrecht, 1999. `http://www.ti3.tu-harburg.de/rump/`

[23] S.M. Rump, “Verification of positive definiteness,” *BIT*, vol. 46, no. 2, pp. 433–452, 2006.

[24] S.M. Rump, “Verified bounds for singular values, in particular for the spectral norm of a matrix and its inverse,” *Bit Numerical Mathematics*, vol. 51, no. 2, pp. 367–384, 2011.

[25] A. Takayasu and S. Oishi, “A refinement technique to residual evaluation of computer assisted proofs for semilinear elliptic boundary value problems,” *RIMS Kokyuroku*, vol. 1733, pp. 118–126, 2011.

[26] A. Takayasu, S. Oishi, and T. Kubo, “Numerical existence theorem for solutions of two-point boundary value problems of nonlinear differential equations,” *NOLTA*, vol. 1, no. 1, pp. 105–118, 2011.

[27] M. Urabe, “Galerkin’s procedure for nonlinear periodic systems,” *Archive for Rational Mechanics and Analysis*, vol. 20, pp. 120–152, 1965.

[28] N. Yamamoto and M.T. Nakao, “Numerical verifications for solutions to elliptic equations using residual iterations with a higher order finite element,” *Journal of Computational and Applied Mathematics*, vol. 60, no. 1–2, pp. 271–279, 1995.

[29] N. Yamamoto and M.T. Nakao, “Numerical verifications of solutions for elliptic equations in nonconvex polygonal domains,” *Numerische Mathematik*, vol. 65, pp. 503–521, 1993.