A numerical verification method for solutions of nonlinear elliptic problems by an infinite dimensional Newton-type formulation

by
Kouji Hashimoto

Contents

1 Introduction **1**

2 Invertibility condition of linear elliptic operators **4**
2.1 A norm estimate of the inverse operator : Part 1 10
2.2 A norm estimate of the inverse operator : Part 2 13

3 A computational procedure for the constant $C(h)$ in nonconvex polygonal domains **15**

4 Applications **17**
4.1 Preliminary 17
4.2 Verification conditions 18

5 Numerical examples **20**
5.1 Verification results on the invertibility of the linearized operator 22
5.2 Verification results for the existence and local uniqueness of solutions of nonlinear problems 26

6 Conclusion **30**

Acknowledgments **31**

References **32**

A numerical verification method for solutions of nonlinear elliptic problems by an infinite dimensional Newton-type formulation

Kouji Hashimoto

Graduate School of Mathematics, Kyushu University
Fukuoka JAPAN

1 Introduction

In this dissertation, first we consider the solvability of the linear elliptic boundary value problem of the form

$$\begin{array}{rcll} \mathcal{L}u \equiv -\Delta u + b \cdot \nabla u + cu & = & g & \text{in} \quad \Omega, \\ u & = & 0 & \text{on} \quad \partial\Omega, \end{array} \tag{1.1}$$

that is equivalent to the invertibility of the operator $\mathcal{L}$ on a certain function space. Here, for $n = 1, 2, 3$, we assume that $b \in (W^1_\infty(\Omega))^n$, $c \in L^\infty(\Omega)$, where $\Omega \subset R^n$ is a bounded domain with piecewise smooth boundary. By using this result, we present a procedure to compute the operator norm corresponding to the inverse $\mathcal{L}^{-1}$.

Next, based upon the Newton-like method, we formulate a numerical verification condition for the existence of solutions of the following nonlinear elliptic problems:

$$\begin{array}{rcll} -\Delta u & = & f(x, u, \nabla u) & x \in \Omega, \\ u & = & 0 & x \in \partial\Omega. \end{array} \tag{1.2}$$

Recently, the rapid development of the computer enables us to analyze a more complicated problems in science and technology by numerical methods. Therefore, people working on various kinds of fields are interested in the guaranteed accuracy of the computed results. In such a situation, particularly, the numerical verification methods for solutions of differential equations become to be more and more important and interesting research area. As well known, the Newton type methods are frequently utilized in guaranteed computations for various problems. For finite dimensional problems, Alefeld showed a procedure for matrix equations, that is, which is sometimes referred as the Interval Gaussian Algorithm [2]. This result presented a basis of verified computations by using interval analysis, and, for example, it is applied to INTLAB by Rump [14] for self-validating algorithms in linear algebraic problems. Actually, by virtue of this numerical tools for guaranteed computations, we can also obtain exact and mathematical results by the computer for infinite dimensional problems such as the nonlinear boundary value problems treated in this dissertation.

Several works, based upon the principle originally found by Nakao, e.g., [5][7], have been presented for the numerical verification methods for solutions of (1.2). They use a method that consists of two procedures; one is a finite dimensional Newton-like iterative process, the other is the computation of the error bounds caused by the gap between the finite and infinite dimension in each iterative procedure. In general, the method for the finite dimensional part utilizes a kind of the interval Newton method. However, it has been recently observed that in the case of having the term with a first order derivative ∇u, this iterative process sometimes fails due to the divergence of the interval computations. In order to overcome this difficulty, an improvement is considered, in [8], which adopts a technique that avoids direct solving the interval system of equations for the finite dimensional part but they used some techniques estimating the norm of the inverse for the coefficient matrix.

In the present paper, we propose a new approach that utilizes the direct estimation of the norm of linearized inverse operators for (1.2) and yields further simplification of the verification procedures. This approach is in fact an extension of the method presented in [8]. Namely, we first verify the invertibility for linearized operators and compute guaranteed norm bounds for its inverse by applying the same principle as for the existing method. Next, we show the existence of solutions for (1.2) by proving the contractivity of the Newton-like operator with a residual form. Another direct computational method for bounds of the linearized operator has already been proposed by Plum (see, e.g.,[11], [13] etc.) using the eigenvalue enclosure methods with a homotopic technique. His method uses some homotopic steps with additional base functions and verified computations for relatively small matrix eigenvalue problems; this is considered a quite different approach from the present method. On the other hand, our verification procedure for nonlinear problems is very close to Plum's method based upon the infinite dimensional Newton's method for the residual type. Therefore, a comparison of these two methods, in respect to the total computational costs for verification of nonlinear problems, would very much depend on the individual problem.

In the below, we denote the L^2 inner product on Ω by $(\cdot,\cdot)$ and the norm by $\|\cdot\|_{L^2}$. And denote the usual Sobolev spaces on Ω by $H^k(\Omega)$ for any positive integer k. We also use the following Sobolev spaces:

$$\begin{array}{rcl} H_0^1(\Omega) & \equiv & \{v \in H^1(\Omega);\ v = 0 \text{ on } \partial\Omega\}, \\ V_\Delta^1(\Omega) & \equiv & \{v \in H_0^1(\Omega);\ \Delta v \in L^2(\Omega)\}. \end{array}$$

For $v \in H_0^1(\Omega)$, we define the H_0^1-norm by $\|v\|_{H_0^1} \equiv (\nabla v, \nabla v)^{1/2}$ and also define the H^2 semi-norm on Ω by, e.g., when $n = 2$,

$$|u|_{H^2} = \left(\|u_{xx}\|_{L^2}^2 + 2\,\|u_{xy}\|_{L^2}^2 + \|u_{yy}\|_{L^2}^2\right)^{\frac{1}{2}}.$$

For $n = 1$ or $n = 3$, analogously defined. And, $<\cdot,\cdot>$ denotes the duality pairing between $H_0^1(\Omega)$ and $H^{-1}(\Omega)$ which is the dual space of $H_0^1(\Omega)$. Moreover, we denote the finite dimensional subspace S_h of $H_0^1(\Omega)$ depending on the parameter h with nodal functions $\{\phi_i\}_{1\le i\le N}$. Notice that for all notations, we sometimes denote the notation with Ω when it depends on Ω.

2 Invertibility condition of linear elliptic operators

In the present section, we consider the numerical verification condition of invertibility for the operator $\mathcal{L}$ defined by (1.1), as well as we present a method to estimate the norm of the inverse operator corresponding to $\mathcal{L}^{-1}$.

First, for each $v \in H_0^1(\Omega)$, we define the H_0^1-projection $P_h v \in S_h$ by

$$(\nabla(v - P_h v), \nabla \phi_h) = 0, \qquad \forall \phi_h \in S_h.$$

Further, we assume the following a priori error estimation:

Assumption 1 *For an arbitrary $v \in V_\Delta^1(\Omega)$, there exists a constant $C(h)$ depending on h such that*

$$\|v - P_h v\|_{H_0^1} \leq C(h)\|\Delta v\|_{L^2}.$$

Here, $C(h)$ has to be numerically determined.

From this assumption, we obtain the following result.

Lemma 1 *Assumption 1 is equivalent to the following inequality:*

$$\|v - P_h v\|_{L^2} \leq C(h)\|v - P_h v\|_{H_0^1}. \tag{2.1}$$

This is called Aubin-Nitsche's trick.

Proof: First, we assume that Assumption 1 holds. Let $\phi \in V_\Delta^1(\Omega)$ be a solution of the following Poisson equation :

$$\begin{aligned} -\Delta \phi &= v - P_h v \quad \text{in} \quad \Omega, \\ \phi &= 0 \qquad\qquad \text{on} \quad \partial\Omega. \end{aligned}$$

Then, from $(\nabla v - \nabla P_h v, \nabla P_h \phi) = 0$, it follows that

$$\begin{aligned} \|v - P_h v\|_{L^2}^2 = (v - P_h v, v - P_h v) &= (\nabla \phi, \nabla v - \nabla P_h v) \\ &= (\nabla \phi - \nabla P_h \phi, \nabla v - \nabla P_h v) \\ &\leq \|\phi - P_h \phi\|_{H_0^1}\|v - P_h v\|_{H_0^1}. \end{aligned}$$

Hence, using Assumption 1, we can obtain

$$\|v - P_h v\|_{L^2} \leq C(h)\|v - P_h v\|_{H_0^1}.$$

Next, we assume that the inequality (2.1) holds. Then, from the definition of H_0^1-projection, we have

$$\begin{aligned}\|v - P_h v\|_{H_0^1}^2 &= (\nabla v - \nabla P_h v, \nabla v - \nabla P_h v)\\ &= (\nabla \phi, \nabla v - \nabla P_h v)\\ &= < \Delta\phi, v - P_h v > .\end{aligned}$$

If $v \in V_\Delta^1(\Omega)$ then, using (2.1), we have $\|v - P_h v\|_{H_0^1} \le C(h)\|\Delta\phi\|_{L^2}$. Therefore, this proof is completed. ∎

Notice that the invertibility of the elliptic operator $\mathcal{L}$ defined in (1.1) is equivalent to the unique solvability of the fixed point equation

$$u = Au, \tag{2.2}$$

where the compact operator $A : H_0^1 \longrightarrow H_0^1$ is defined by $Au := \Delta^{-1}(b\cdot\nabla u + cu)$ and where Δ^{-1} stands for the solution operator of the Poisson equation with homogeneous boundary condition.

Now, according to the usual verification principle, e.g.,[7], we formulate a sufficient condition for which the equation (1.2) has a unique solution. As the preliminary, we define the matrices $\mathbf{G} = (\mathbf{G}_{i,j})$ and $\mathbf{D} = (\mathbf{D}_{i,j})$ by :

$$\begin{aligned}\mathbf{G}_{i,j} &= (\nabla\phi_j, \nabla\phi_i) + (b\cdot\nabla\phi_j, \phi_i) + (c\phi_j, \phi_i),\\ \mathbf{D}_{i,j} &= (\nabla\phi_j, \nabla\phi_i),\end{aligned}$$

for $1 \le i, j \le N$. Let $\mathbf{L}$ be a lower triangular matrix satisfying the Cholesky decomposition: $\mathbf{D} = \mathbf{L}\mathbf{L}^T$. And we denote the matrix norm by $\|\cdot\|_E$ induced from the Euclidean 2-norm $\|\cdot\|_E$ in R^N.

Also, for $\kappa_1 = \|\ |b|_E\ \|_{L^\infty}$, $\kappa_2 = \|\mathrm{div}\ b\|_{L^\infty}$ and $\kappa_3 = \|c\|_{L^\infty}$, we define the following constants:

$$\begin{aligned}C_1 &= \kappa_1 + C_p\kappa_2, & C_3 &= \kappa_1 + C_p\kappa_3,\\ C_2 &= C_p\kappa_3, & C_4 &= \kappa_1 + C(h)\kappa_3,\end{aligned}$$

where $\|\cdot\|_{L^\infty}$ means L^∞ norm on Ω and C_p is a Poincaré constant such that $\|\phi\|_{L^2} \le C_p\|\phi\|_{H_0^1}$ for an arbitrary $\phi \in H_0^1(\Omega)$. Then we have the following main result of this paper.

Theorem 2.1 *If the matrix* $\mathbf{G}$ *is invertible and, for the constants defined above,*

$$C(h)\Big[C_3M(C_1+C_2)C(h)+C_4\Big]<1$$

holds, then the operator $\mathcal{L}$ *defined in (1.1) is invertible. Here,* $M \equiv \|\mathbf{L}^T\mathbf{G}^{-1}\mathbf{L}\|_E$ *and* $C(h)$ *is the same constant as in Assumption 1.*

Remark 1 *The main cost for checking the invertibility condition consists of the guaranteed estimation of* $\|\mathbf{L}^T\mathbf{G}^{-1}\mathbf{L}\|_E$*. First, we compute the matrix* $\mathbf{L}$ *by the interval Cholesky-decomposition. Next, by using the approximate LU decomposition of* $\mathbf{G}$ *and some error estimates, we enclose the guaranteed inverse* $\mathbf{G}^{-1}$*. Finally, we make a verified computation of the largest singular value for the matrix* $\mathbf{L}^T\mathbf{G}^{-1}\mathbf{L}$*, which is equal to the square root of the largest eigenvalue of* $\mathbf{L^TG^{-T}DG^{-1}L}$*, to obtain the desired estimation.*

Proof: First, as usual, we decompose the equation $u = Au$ as

$$\begin{aligned} P_h u &= P_h A u, \\ (I-P_h)u &= (I-P_h)Au, \end{aligned}$$

where I implies the identity map on $H_0^1(\Omega)$.

Next, according to the same formulation to that in [6],[7] etc., we define two operators by

$$N_h u \equiv P_h u - [I-A]_h^{-1} P_h (I-A)u$$

and

$$Tu \equiv N_h u + (I-P_h)Au,$$

respectively, where $[I-A]_h^{-1}$ means the inverse of $P_h(I-A)|_{S_h} : S_h \longrightarrow S_h$. Note that if we define the Galerkin approximation A_h on S_h of the operator A, then $[I-A]_h^{-1}$ coincides with $(I-A_h)^{-1}$ on S_h. The existence of the operator $[I-A]_h^{-1}$ is assumed, which is equivalent to the regularity of the corresponding matrix, and is numerically followed by the unique solvability of the linear system of equations in the verification process.
We now, for positive real numbers α and γ, define the set $U = U_h + U_\perp$ by

$$\begin{aligned} U_h &:= \left\{u_h \in S_h : \|u_h\|_{H_0^1} \le \gamma\right\}, \\ U_\perp &:= \left\{u_\perp \in S_h^\perp : \|u_\perp\|_{H_0^1} \le \alpha\right\}, \end{aligned}$$

where $S_h^{\perp}$ stands for the orthogonal complement of S_h in $H_0^1(\Omega)$. Then, by the fact that $u = Au$ is equivalent to $u = Tu$, in order to prove the unique existence of a solution to (2.2) in the set U, it suffices to show the inclusion $TU \overset{\circ}{\subset} U$ due to the linearity of the equation (e.g., [15]), where $TU \overset{\circ}{\subset} U$ implies that the closure of TU is included by the interior of U.

Further notice that a sufficient condition of this inclusion can be written as

$$\|N_h U\|_{H_0^1} \equiv \sup_{u\in U} \|N_h u\|_{H_0^1} < \gamma, \tag{2.3}$$

$$\begin{aligned}\|(I-P_h)AU\|_{H_0^1} &\equiv \sup_{u\in U} \|(I-P_h)Au\|_{H_0^1} \\ &\leq C(h) \sup_{u\in U} \|b\cdot\nabla u + cu\|_{L^2} < \alpha,\end{aligned} \tag{2.4}$$

where we have used the estimate in Assumption 1.

In the below, we estimate $\|N_h u\|_{H_0^1}$ and $\|b\cdot\nabla u + cu\|_{L^2}$ in (2.3) and (2.4), respectively.
First, for an arbitrary $u = u_h + u_\perp \in U_h + U_\perp$, setting $\psi_h := N_h(u_h + u_\perp)$, we have

$$\begin{aligned}\psi_h &= u_h - [I-A]_h^{-1} P_h (I-A)(u_h + u_\perp) \\ &= [I-A]_h^{-1} P_h A u_\perp.\end{aligned}$$

Now, note that for $v_h := P_h A u_\perp \in S_h$ we have

$$(\nabla\psi_h, \nabla\phi_h) + (b\cdot\nabla\psi_h, \phi_h) + (c\psi_h, \phi_h) = (\nabla v_h, \nabla\phi_h), \qquad \forall \phi_h \in S_h. \tag{2.5}$$

Denoting

$$\psi_h := \sum_{i=1}^{N} w_i \phi_i \qquad \text{and} \qquad v_h := \sum_{i=1}^{N} v_i \phi_i, \tag{2.6}$$

from (2.5), we have a matrix equation of the form

$$\mathbf{G}\vec{w} = \mathbf{D}\vec{v}. \tag{2.7}$$

Here, $\vec{w} = (w_1, w_2, \cdots, w_N)^T$ and $\vec{v} = (v_1, v_2, \cdots, v_N)^T$ are coefficient vectors of ψ_h and v_h, respectively. Therefore, from (2.6) and (2.7), it follows that

$$\begin{aligned}
\|\psi_h\|^2_{H^1_0} &= \vec{w}^T\mathbf{D}\vec{w} \\
&= \vec{w}^T\mathbf{D}\mathbf{G}^{-1}\mathbf{D}\vec{v} \\
&= (\mathbf{L}^T\vec{w})^T(\mathbf{L}^T\mathbf{G}^{-1}\mathbf{L})(\mathbf{L}^T\vec{v}) \\
&\leq \|\mathbf{L}^T\vec{w}\|_E\|\mathbf{L}^T\mathbf{G}^{-1}\mathbf{L}\|_E\|\mathbf{L}^T\vec{v}\|_E \\
&= \|\psi_h\|_{H^1_0}\|\mathbf{L}^T\mathbf{G}^{-1}\mathbf{L}\|_E\|v_h\|_{H^1_0}.
\end{aligned}$$

Note that, from the above fact, we have $\|\mathbf{L}^T\mathbf{G}^{-1}\mathbf{L}\|_E = \|[I-A]_h^{-1}\|_{H^1_0}$. Thus, defining $M \equiv \|\mathbf{L}^T\mathbf{G}^{-1}\mathbf{L}\|_E$, we obtain

$$\begin{aligned}
\|\psi_h\|_{H^1_0} &\leq M\|P_hAu_\perp\|_{H^1_0} \\
&= M\|P_h\Delta^{-1}(b\cdot\nabla u_\perp + cu_\perp)\|_{H^1_0} \\
&\leq M\|\Delta^{-1}(b\cdot\nabla u_\perp + cu_\perp)\|_{H^1_0}.
\end{aligned}\tag{2.8}$$

Next, letting $\psi_1 := \Delta^{-1}(b\cdot\nabla u_\perp)$, some simple calculations yields that

$$\begin{aligned}
\|\psi_1\|^2_{H^1_0} = (\nabla\psi_1, \nabla\psi_1) &= (-\Delta\psi_1, \psi_1) \\
&= (-b\cdot\nabla u_\perp, \psi_1) \\
&\leq \|u_\perp\|_{L^2}\|\mathrm{div}(b\psi_1)\|_{L^2} \\
&\leq C(h)(\kappa_1 + C_p\kappa_2)\alpha\|\psi_1\|_{H^1_0},
\end{aligned}\tag{2.9}$$

where we have used the fact $\|u_\perp\|_{L^2} \leq C(h)\alpha$. Furthermore, setting $\psi_2 := \Delta^{-1}(cu_\perp)$ and by applying the similar argument to the above, we have

$$\|\psi_2\|_{H^1_0} \leq C(h)C_p\kappa_3\alpha \tag{2.10}$$

Thus, by (2.8) – (2.10), we obtain the following estimate for the finite dimensional part

$$\|N_hU\|_{H^1_0} \leq M(C_1 + C_2)C(h)\alpha, \tag{2.11}$$

where $C_1 \equiv \kappa_1 + C_p\kappa_2$ and $C_2 \equiv C_p\kappa_3$. Next, observe that

$$\begin{aligned}
\|b\cdot\nabla u_h + cu_h\|_{L^2} &\leq \kappa_1\|u_h\|_{H^1_0} + C_p\kappa_3\|u_h\|_{H^1_0} \\
&\leq (\kappa_1 + C_p\kappa_3)\gamma, \\
\|b\cdot\nabla u_\perp + cu_\perp\|_{L^2} &\leq \kappa_1\|u_\perp\|_{H^1_0} + \kappa_3\|u_\perp\|_{L^2} \\
&\leq (\kappa_1 + C(h)\kappa_3)\alpha.
\end{aligned}$$

Therefore, by using (2.4) and the triangle inequality, we have

$$\|(I-P_h)AU\|_{H_0^1} \le C(h)(C_3\gamma + C_4\alpha), \tag{2.12}$$

where $C_3 \equiv \kappa_1 + C_p\kappa_3$, $C_4 \equiv \kappa_1 + C(h)\kappa_3$.

Now, from (2.11) and (2.12), the invertibility conditions (2.3) and (2.4) are reduced to

$$\begin{aligned} M(C_1+C_2)C(h)\alpha &< \gamma, & (2.13)\\ C(h)(C_3\gamma + C_4\alpha) &< \alpha. & (2.14)\end{aligned}$$

For an arbitrary small $\varepsilon > 0$, if we set $\gamma := M(C_1+C_2)C(h)\alpha + \varepsilon$, then the condition (2.13) clearly holds. Therefore, by substituting it to (2.14) we have

$$C(h)\Big[C_3(M(C_1+C_2)C(h)\alpha + \varepsilon) + C_4\alpha\Big] < \alpha,$$

which is equivalent to

$$C(h)\Big[C_3M(C_1+C_2)C(h) + C_4\Big] < 1.$$

Thus the desired conclusion is obtained. ∎

Remark 2 *The conditions (2.3) and (2.4) are equivalent to $|||T||| < 1$ in some scaled norm $|||\cdot|||$ in H_0^1, e.g., $|||v|||^2 \equiv ||P_hv||^2_{H_0^1}/\gamma^2 + ||(I-P_h)v||^2_{H_0^1}/\alpha^2$. Then, the invertibility of the operator $I-T$ follows by the convergence of the Neumann series.*

When the coefficient function b of the first order term is not differentiable, we have the following alternative condition.

Corollary 1 *For the operator $\mathcal{L}$ defined in (1.1), let $b \in (L^\infty(\Omega))^n$. If*

$$C(h)\Big[C_3M(\hat{C}_1 + C(h)C_2) + C_4\Big] < 1,$$

then the operator $\mathcal{L}$ defined in (1.1) is invertible. Here, $\hat{C}_1 = C_p\kappa_1$.

Proof: The difference from the proof of Theorem 2.1 is only the part concerning the estimates (2.9). Corresponding estimates are now

$$\begin{aligned}\|\psi_1\|^2_{H^1_0} = (-\Delta\psi_1, \psi_1) &= (-b\cdot\nabla u_\perp, \psi_1)\\ &\le \kappa_1\|u_\perp\|_{H^1_0}\|\psi_1\|_{L^2}\\ &\le C_p\kappa_1\alpha\|\psi_1\|_{H^1_0},\end{aligned}$$

which proves the corollary. ∎

2.1 A norm estimate of the inverse operator : Part 1

Now, our next purpose is the estimation of the operator norm $\|(I-A)^{-1}\|_{H^1_0}$ corresponding to the norm for $\mathcal{L}^{-1} : H^{-1} \to H^1_0$.

Theorem 2.2 *Under the same assumptions in Theorem 2.1, provided that*

$$\kappa \equiv C(h)\Big[C_3M(C_1+C_2)C(h)+C_4\Big] < 1,$$

then the following estimation holds

$$\|(I-A)^{-1}\|_{H^1_0} \le \|M_\perp + M_h\|_E^{1/2} =: \mathcal{M},$$

where the 2×2 *matrices* $M_\perp$, M_h *are defined by*

$$M_\perp = \begin{bmatrix} \tau_1^2 & \tau_1\tau_2 \\ \tau_1\tau_2 & \tau_2^2 \end{bmatrix}, \quad M_h = \begin{bmatrix} \tau_3^2 & \tau_3\tau_4 \\ \tau_3\tau_4 & \tau_4^2 \end{bmatrix}.$$

Here, τ_i, $(1\le i\le 4)$ *are given as follows*

$$\begin{aligned}\tau_1 &= \frac{MC(h)C_3}{1-\kappa}, & \tau_2 &= \frac{1}{1-\kappa},\\ \tau_3 &= M\Big[(C_1+C_2)C(h)\tau_1+1\Big], & \tau_4 &= M(C_1+C_2)C(h)\tau_2.\end{aligned}$$

Proof: Let ψ be an arbitrary element in $H^1_0(\Omega)$. Then, by the Fredholm alternative theorem, the invertibility of $(I-A)$ implies that there exists a unique element $u\in H^1_0(\Omega)$ satisfying $(I-A)u=\psi$. When we set

$$\begin{aligned}N_hu &:= P_hu - [I-A]_h^{-1}P_h((I-A)u-\psi),\\ Tu &:= N_hu + (I-P_h)(Au+\psi),\end{aligned}$$

$(I-A)u = \psi$ is equivalent to $Tu = u$. Using the unique decompositions $u = u_h + u_\perp$ and $\psi = \psi_h + \psi_\perp$ in $H_0^1(\Omega) = S_h \oplus S_h^\perp$, by some simple calculations, we have

$$\begin{aligned} u_h &= [I-A]_h^{-1}(P_h A u_\perp + P_h \psi), \\ u_\perp &= (I-P_h)A(u_h + u_\perp) + (I-P_h)\psi. \end{aligned}$$

Hence, taking notice of $M = \|[I-A]_h^{-1}\|_{H_0^1}$ and the estimates in the proof of Theorem 2.1, we have

$$\begin{aligned} \|u_h\|_{H_0^1} &\le M\|P_h A u_\perp + P_h\psi\|_{H_0^1} \\ &\le M(C_1 + C_2)C(h)\|u_\perp\|_{H_0^1} + M\|P_h\psi\|_{H_0^1}, \end{aligned} \tag{2.15}$$

$$\begin{aligned} \|u_\perp\|_{H_0^1} &\le \|(I-P_h)A(u_h + u_\perp)\|_{H_0^1} + \|(I-P_h)\psi\|_{H_0^1} \\ &\le C(h)(C_3\|u_h\|_{H_0^1} + C_4\|u_\perp\|_{H_0^1}) + \|(I-P_h)\psi\|_{H_0^1}. \end{aligned} \tag{2.16}$$

Substituting the estimate of $\|u_h\|_{H_0^1}$ in (2.15) into the last right-hand side of (2.16) and solving it with respect to $\|u_\perp\|_{H_0^1}$, we get

$$\begin{aligned} \|u_\perp\|_{H_0^1} &\le \frac{MC(h)C_3}{1-\kappa}\|P_h\psi\|_{H_0^1} + \frac{1}{1-\kappa}\|(I-P_h)\psi\|_{H_0^1} \\ &= \tau_1\|P_h\psi\|_{H_0^1} + \tau_2\|(I-P_h)\psi\|_{H_0^1}. \end{aligned} \tag{2.17}$$

Thus we also have by (2.15)

$$\begin{aligned} \|u_h\|_{H_0^1} &\le M(C_1 + C_2)C(h)\Big(\tau_1\|P_h\psi\|_{H_0^1} + \tau_2\|(I-P_h)\psi\|_{H_0^1}\Big) + M\|P_h\psi\|_{H_0^1} \\ &\le M\Big[(C_1 + C_2)C(h)\tau_1 + 1\Big]\|P_h\psi\|_{H_0^1} \\ &\qquad\qquad + M(C_1 + C_2)C(h)\tau_2\|(I-P_h)\psi\|_{H_0^1} \\ &= \tau_3\|P_h\psi\|_{H_0^1} + \tau_4\|(I-P_h)\psi\|_{H_0^1}. \end{aligned} \tag{2.18}$$

Therefore, we obtain the desired conclusion from (2.17) and (2.18). ∎

Moreover, we have the following estimates corresponding to Corollary 1.

Corollary 2 *Under the same assumption as in Corollary 1, if*

$$\hat{\kappa} \equiv C(h)\Big[C_3 M(\hat{C}_1 + C(h)C_2) + C_4\Big] < 1,$$

then

$$\|(I-A)^{-1}\|_{H_0^1} \;\leq\; \|\hat{M}_\perp + \hat{M}_h\|_E^{1/2} =: \hat{\mathcal{M}},$$

where the 2×2 *matrices* $M_\perp$, M_h *are defined by*

$$\hat{M}_\perp = \begin{bmatrix} \hat{\tau}_1^2 & \hat{\tau}_1\hat{\tau}_2 \\ \hat{\tau}_1\hat{\tau}_2 & \hat{\tau}_2^2 \end{bmatrix}, \quad \hat{M}_h = \begin{bmatrix} \hat{\tau}_3^2 & \hat{\tau}_3\hat{\tau}_4 \\ \hat{\tau}_3\hat{\tau}_4 & \hat{\tau}_4^2 \end{bmatrix}.$$

Here, $\hat{\tau}_i$, $(1\leq i\leq 4)$ *are given as follows*

$$\begin{aligned} \hat{\tau}_1 &= \frac{MC(h)C_3}{1-\hat{\kappa}}, & \hat{\tau}_2 &= \frac{1}{1-\hat{\kappa}}, \\ \hat{\tau}_3 &= M\Big[(\hat{C}_1 + C(h)C_2)\hat{\tau}_1 + 1\Big], & \hat{\tau}_4 &= M(\hat{C}_1 + C(h)C_2)\hat{\tau}_2. \end{aligned}$$

We now note that the following a priori estimate of the solution to (1.1) is obtained.

Theorem 2.3 *It follows that*

$$\|u\|_{H_0^1} \leq \|(I-A)^{-1}\|_{H_0^1}\|g\|_{H^{-1}},$$

where the H^{-1}*-norm* $\|\cdot\|_{H^{-1}}$ *is defined by*

$$\|g\|_{H^{-1}} \equiv \sup_{\phi\in H_0^1(\Omega)} \frac{< g,\phi >}{\|\phi\|_{H_0^1}}.$$

Particularly,

$$\|u\|_{H_0^1} \leq C_p\|(I-A)^{-1}\|_{H_0^1}\|g\|_{L^2} \quad \textit{for } g\in L^2(\Omega).$$

Indeed, defining $\psi := -\Delta^{-1}g$, then taking account that $(I-A)u = \psi$ and that

$$\|\psi\|_{H_0^1}^2 = (\nabla\psi,\nabla\psi) =< -\Delta\psi,\psi >=< g,\psi >\leq \|g\|_{H^{-1}}\|\psi\|_{H_0^1}.$$

The second part follows from the Poincaré inequality.

2.2 A norm estimate of the inverse operator : Part 2

In this subsection, we show the estimate for the solution of the linear equation (1.1) by using slightly different manner from the argument in the previous subsection.

Our method needs the constructive a priori error estimate between the function and its projection. In this estimation, the orthogonal property of the projection usually plays an important role. For the error analysis, there are many results which use this orthogonality([9][12][16][19]). However, in some case, e.g., [3], we need to get the desired estimates without this property. Therefore, we derive the following result in the general case.

Theorem 2.4 *Under the same assumptions in Theorem 2.1 and 2.2, provided that $\kappa < 1$ and let $u \in H_0^1(\Omega)$ be a unique solution for the linear equation (1.1), that is, $\mathcal{L}u = g$. Then, we have the following estimation:*

$$\|u\|_{H_0^1} \leq \mathcal{M}_h\|g\|_{H^{-1}} + \mathcal{M}_\perp\|(I - P_h)\Delta^{-1}g\|_{H^{-1}},$$

where $\mathcal{M}_h \equiv (\tau_1^2 + \tau_3^2)^{1/2}$ and $\mathcal{M}_\perp \equiv (\tau_2^2 + \tau_4^2)^{1/2}$.

Moreover, if $g \in L^2(\Omega)$, then

$$\|u\|_{H_0^1} \leq \mathcal{M}_*\|g\|_{L^2},$$

where $\mathcal{M}_* \equiv C_p\mathcal{M}_h + C(h)\mathcal{M}_\perp$. In this case, we obtain the following a priori estimation:

$$\|u - P_h u\|_{H_0^1} \leq C_{\mathcal{L}}(h)\|g\|_{L^2},$$

where the constant $C_{\mathcal{L}}(h)$ is taken as $C_{\mathcal{L}}(h) \equiv C(h)(MC_pC_3 + 1)\tau_2$.

Proof: The argument of the proof is very similar to that in Theorem 2.2. So, setting $\psi = -\Delta^{-1}g$, we rewrite (2.17) and (2.18) as

$$\begin{aligned}
\|u_\perp\|_{H_0^1} &\leq \tau_1\|P_h\psi\|_{H_0^1} + \tau_2\|(I - P_h)\psi\|_{H_0^1}, \\
\|u_h\|_{H_0^1} &\leq \tau_3\|P_h\psi\|_{H_0^1} + \tau_4\|(I - P_h)\psi\|_{H_0^1}.
\end{aligned}$$

From the above, we have

$$
\begin{aligned}
\|u\|_{H_0^1}^2 &= \|u_\perp\|_{H_0^1}^2 + \|u_h\|_{H_0^1}^2 \\
&\le \Big(\tau_1\|P_h\psi\|_{H_0^1} + \tau_2\|(I-P_h)\psi\|_{H_0^1}\Big)^2 \\
&\qquad\qquad + \Big(\tau_3\|P_h\psi\|_{H_0^1} + \tau_4\|(I-P_h)\psi\|_{H_0^1}\Big)^2 \\
&\le \Big[(\tau_1^2+\tau_3^2)^{1/2}\|P_h\psi\|_{H_0^1} + (\tau_2^2+\tau_4^2)^{1/2}\|(I-P_h)\psi\|_{H_0^1}\Big]^2 \\
&\le \Big[C_p(\tau_1^2+\tau_3^2)^{1/2} + C(h)(\tau_2^2+\tau_4^2)^{1/2}\Big]^2\|g\|_{L^2}^2,
\end{aligned}
$$

where we have used the fact that $\|P_h\psi\|_{H_0^1} \le \|\psi\|_{H_0^1}$ together with Theorem 2.3 and Assumption 1. Thus, the proof is completed. ∎

We now obtain the following estimates corresponding to Corollary 1 and 2.

Corollary 3 *Under the same assumptions in Corollary 1 or 2, provided that $\hat{\kappa} < 1$ and let $u \in H_0^1(\Omega)$ be a unique solution for the linear equation (1.1), that is, $\mathcal{L}u = g$. Then, we have the following estimation:*

$$
\|u\|_{H_0^1} \le \hat{\mathcal{M}}_h\|g\|_{H^{-1}} + \hat{\mathcal{M}}_\perp\|(I-P_h)\Delta^{-1}g\|_{H^{-1}},
$$

where $\hat{\mathcal{M}}_h \equiv (\hat{\tau}_1^2+\hat{\tau}_3^2)^{1/2}$ and $\hat{\mathcal{M}}_\perp \equiv (\hat{\tau}_2^2+\hat{\tau}_4^2)^{1/2}$.

Moreover if $g \in L^2(\Omega)$, then

$$
\|u\|_{H_0^1} \le \hat{\mathcal{M}}_*\|g\|_{L^2},
$$

where $\hat{\mathcal{M}}_* \equiv C_p\hat{\mathcal{M}}_h + C(h)\hat{\mathcal{M}}_\perp$. In this case, we obtain the following a priori estimation:

$$
\|u - P_hu\|_{H_0^1} \le \hat{C}_{\mathcal{L}}(h)\|g\|_{L^2},
$$

where the constant $\hat{C}_{\mathcal{L}}(h)$ is taken as $\hat{C}_{\mathcal{L}}(h) \equiv C(h)(MC_pC_3+1)\hat{\tau}_2$.

3 A computational procedure for the constant $C(h)$ in nonconvex polygonal domains

In this section, we introduce a computational procedure of the a priori error estimation for nonconvex nonsmooth domains.

Remark 3 *Notice that if Ω is rectangular then it follows that $C(h) = C_0 h$, where h is the maximum mesh size and the constant C_0 is given by the following.*

$$C_0 = \begin{cases} 1/\pi & \text{if} \quad \textit{the uniform piecewise bilinear rectangular element [16]}, \\ 1/(2\pi) & \text{if} \quad \textit{the uniform piecewise biquadratic rectangular element [10]}. \end{cases}$$

Moreover, if we consider the uniform piecewise linear triangular element, we have $C_0 = 1/2$.

Now, for nonconvex polygonal domains, a procedure is presented in [19]. We briefly describe this technique below.

Let Ω_* be a convex polygonal domain which includes Ω. We take Ω_{out} as a residual domain such that $\Omega_* = (\overline{\Omega} \cup \overline{\Omega_{\text{out}}}) \backslash \partial(\overline{\Omega} \cup \overline{\Omega_{\text{out}}})$ (see Figure 1).

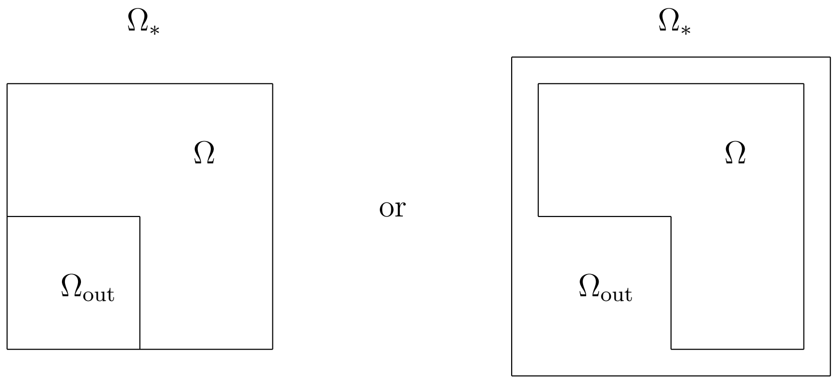

Figure 1: A convex extensional of the domain

Let S_h^* be a finite element subspace of $H_0^1(\Omega_*)$. Assume that finite element subspaces S_h and S_h^* satisfy $S_h \subset S_h^*$, and

$$P_h : H_0^1(\Omega) \to S_h, \qquad P_h^* : H_0^1(\Omega_*) \to S_h^*$$

are the H_0^1-projrction (see Figure 2). And we extend an arbitrary $v \in H_0^1(\Omega)$ to the function on Ω_* with $v = 0$ in $\Omega_{\rm out}$, which belongs to $H_0^1(\Omega_*)$ and is also denoted by v, then it follows that

$$\|v - P_h v\|_{L^2(\Omega)} \quad \leq \quad C_0 h \left(1 + K^2\right)^{1/2} \|v - P_h v\|_{H_0^1(\Omega)},$$

where C_0 is a constant depending on Ω_* and finite elements, and K is defined by

$$K \equiv \sup_{v \in H_0^1(\Omega)} \frac{\|P_h^* v - P_h v\|_{L^2(\Omega)}}{\|P_h^* v - P_h v\|_{L^2(\Omega_{\rm out})}}.$$

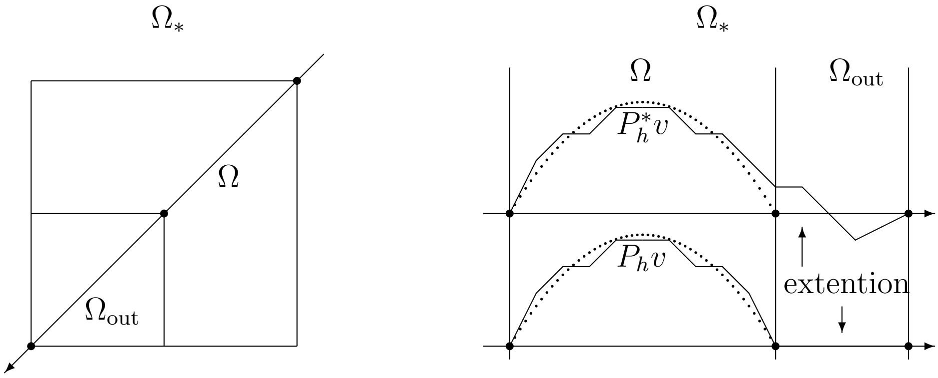

Figure 2: The image of $P_h v$ and $P_h^* v$ for $v \cdots$

Notice that K can be computed as the matrix eigenvalue problem, but, it has some negative order in the mesh size h. Combining this procedure and Lemma 1, we can obtain the a priori constant in Assumption 1 as $C(h) = C_0 h(1 + K^2)^{1/2}$ (see [19] for details).

4 Applications

In this section, we mention about the actual applications of the results obtained in the previous section to the verification of solutions for the nonlinear elliptic problem (1.2). We assume that the nonlinear map $f(u) \equiv f(\cdot, u, \nabla u)$ from $H_0^1(\Omega)$ into $L^2(\Omega)$ is continuous and bounded.

4.1 Preliminary

In this subsection, we transform the original boundary value problem (1.2) into the so-called residual equation by using an approximate solution $\tilde{u}_h \in V_h \subset H_0^1(\Omega)$ defined by

$$(\nabla \tilde{u}_h, \nabla \phi_h) = (f(\tilde{u}_h), \phi_h), \qquad \forall \phi_h \in V_h, \tag{4.1}$$

where V_h denotes the finite element subspace for the approximation. Usually, V_h coincides with S_h, but sometimes it does not (e.g. Section 5 in this paper). For the effective computation of the solution for (4.1) with guaranteed accuracy, refer, for example, [1], [14] etc.

Next, we define the $\bar{u} \in V_\Delta^1(\Omega)$ by the solution of Poisson's equation :

$$\begin{array}{rcll} -\Delta \bar{u} & = & f(\tilde{u}_h) & \text{in} \quad \Omega, \\ \bar{u} & = & 0 & \text{on} \quad \partial\Omega. \end{array}$$

Further, let define residues by

$$u - \tilde{u}_h = (u - \bar{u}) + (\bar{u} - \tilde{u}_h), \qquad w := u - \bar{u}, \quad v_0 := \bar{u} - \tilde{u}_h. \tag{4.2}$$

Note that v_0 is an unknown function but its norm can be computed by an a priori and a posteriori techniques (e.g., see [9], [17]). Thus, using the residues in (4.2), concerned problem is reduced to the following residual form

$$\begin{array}{rcll} -\Delta w & = & f(w + v_0 + \tilde{u}_h) - f(\tilde{u}_h) & \text{in} \quad \Omega, \\ w & = & 0 & \text{on} \quad \partial\Omega. \end{array} \tag{4.3}$$

Hence, denoting the Fréchet derivative at $\tilde{u}_h$ by $f'(\tilde{u}_h)$, the Newton-type residual equation for (4.3) is written as:

$$\begin{array}{rcll} -\Delta w - f'(\tilde{u}_h) w & = & g_r(w) & \text{in} \quad \Omega, \\ w & = & 0 & \text{on} \quad \partial\Omega, \end{array}$$

where $g_r(w) \equiv f(w + v_0 + \tilde{u}_h) - f(\tilde{u}_h) - f'(\tilde{u}_h)w$.

In the above, we assumed that the approximate solution $\tilde{u}_h$ is defined as an element in $H_0^1(\Omega)$, i.e., C^0-element. When we use the function satisfying $\tilde{u}_h \in V_h \subset H^2(\Omega)$, i.e., C^1-element, we can get more simpler residual Newton-type equation without v_0 of the form

$$\begin{array}{rcll} -\Delta w - f'(\tilde{u}_h)w &=& g_d(w) & \text{in} \quad \Omega, \\ w &=& 0 & \text{on} \quad \partial\Omega, \end{array}$$

where $w := u - \tilde{u}_h$ and $g_d(w) := f(w + \tilde{u}_h) + \Delta\tilde{u}_h - f'(\tilde{u}_h)w$. For another type of simple residual formulation for C^0-element, refer [6] or [13] in which some H^{-1} arguments are effectively used.

4.2 Verification conditions

We now write down again the nonlinear boundary value problem of the form:

$$\begin{array}{rcrcll} \mathcal{L}w &\equiv& -\Delta w - f'(\tilde{u}_h)w &=& g(w) & \text{in} \quad \Omega, \\ && w &=& 0 & \text{on} \quad \partial\Omega, \end{array} \tag{4.4}$$

where $g(w) \equiv g_r(w)$ or $g(w) \equiv g_d(w)$. If $\mathcal{L}$ is invertible, then (4.4) is rewritten as the fixed point form

$$w \quad = \quad F(w) \left(\equiv \mathcal{L}^{-1}g(w)\right). \tag{4.5}$$

Notice that the Newton-like operator F in (4.5) is compact on $H_0^1(\Omega)$ from the assumptions on f, and that it is expected to be a contraction map on some neighborhood of zero.
Therefore, we consider the set, which we often refer as the *candidate set*, of the form $W_\alpha \equiv \{w \in H_0^1(\Omega) : \|w\|_{H_0^1} \leq \alpha\}$.

First, for the existential condition of solutions, we need to choose the set W_α, which is equivalent to determine a positive number α, satisfying the following criterion based on the Schauder fixed point theorem:

$$F(W_\alpha) \subset W_\alpha. \tag{4.6}$$

And next, for the proof of local uniqueness within W_α, the following contraction property is needed on the same set W_α in (4.6):

$$\|F(w_1) - F(w_2)\|_{H_0^1} \leq k\|w_1 - w_2\|_{H_0^1}, \quad \forall w_1, w_2 \in W_\alpha, \tag{4.7}$$

for some constant $0 < k < 1$. Notice that, in the above case, the Schauder fixed point theorem can be replaced by the Banach fixed point theorem, which might yields an advantage if we apply our method to noncompact problems.
For (4.6), from the theorem 2.3, a sufficient condition can be written as

$$\begin{aligned} \|F(W_\alpha)\|_{H_0^1} &\equiv \sup_{w\in W_\alpha} \|F(w)\|_{H_0^1} \\ &\leq \mathcal{M}_1 \sup_{w\in W_\alpha} \|g(w)\|_{L^2} \leq \alpha, \end{aligned} \tag{4.8}$$

where $\mathcal{M}_1 \equiv \min\{C_p\mathcal{M}, \mathcal{M}_*\}$, and $\mathcal{M}$, $\mathcal{M}_*$ are the norms of the operator $\mathcal{L}^{-1} : H^{-1} \to H_0^1$ defined in the theorems 2.2 and 2.4.
On the other hand, for the verification of local uniqueness condition (4.7) on W_α, in general, we use the following deformation:

$$g(w_1) - g(w_2) = \Phi(w_1, w_2)(w_1 - w_2),$$

where $\Phi(w_1, w_2)$ denotes a function in w_1 and w_2, for example, if $g(w) = w^2$, then $\Phi(w_1, w_2) = w_1 + w_2$. Therefore, the condition (4.7) reduces to finding a constant $0 < k < 1$ satisfying the inequality of the form

$$\mathcal{M}_1\|\Phi(w_1, w_2)(w_1 - w_2)\|_{L^2} \leq k\|w_1 - w_2\|_{H_0^1}, \quad \forall w_1, w_2 \in W_\alpha. \tag{4.9}$$

5 Numerical examples

In this section, we present the following four examples, and we show numerical results on the invertibility of linearized operators and the existence of solutions of nonlinear problems in next two subsections.

Example 1 (Plum's example) *[12]*

$$\begin{aligned} -\Delta u &= u\left(\lambda - \frac{1}{2}|\nabla u|^2\right) && \text{in} \quad \Omega, \\ u &= 0 && \text{on} \quad \partial\Omega, \end{aligned} \tag{5.1}$$

where $\Omega = (0,1)^2$ *and* $\lambda > 0$ *is a parameter.*

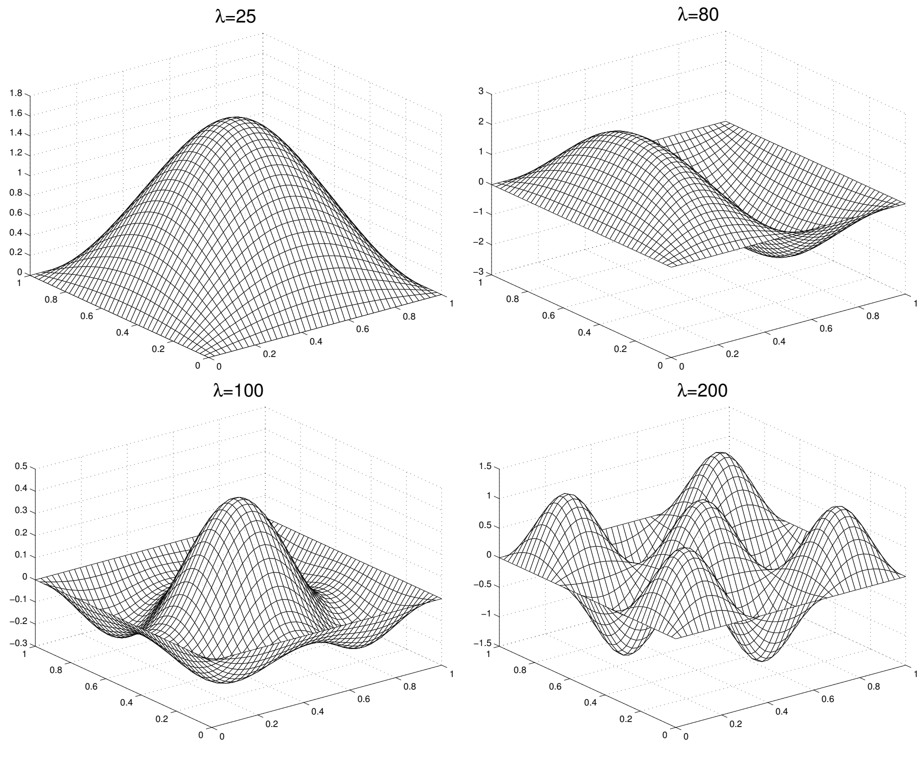

Figure 3: Approximate solutions to Example 1

Example 2 (Allen-Cahn equation)

$$\begin{aligned} -\Delta u &= \lambda u(u-a)(1-u) \quad \text{in} \quad \Omega, \\ u &= 0 \qquad\qquad\qquad\quad\ \text{on} \quad \partial\Omega, \end{aligned} \tag{5.2}$$

where $\Omega = (0,1)^2$ *and* $a > 0$, $\lambda > 0$ *are parameters.*

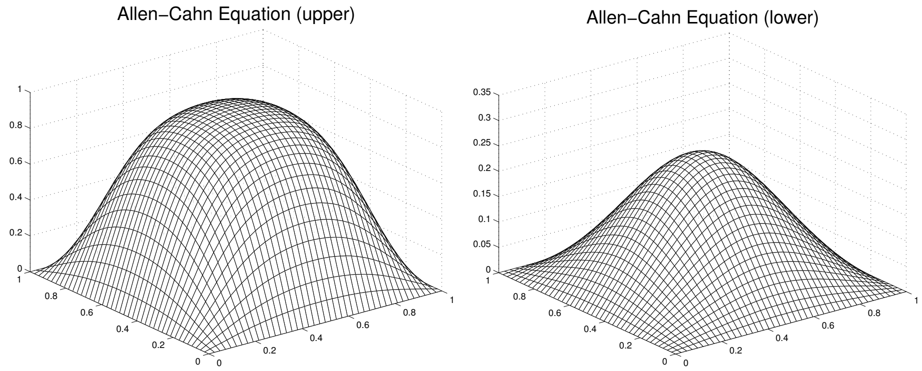

Figure 4: Approximate solutions to Example 2 for $a = 0.01$ and $\lambda = 150$

Example 3 (Emden's equation)

$$\begin{aligned} -\Delta u &= u^2 \quad \text{in} \quad \Omega, \\ u &= 0 \quad\ \text{on} \quad \partial\Omega, \end{aligned} \tag{5.3}$$

where $\Omega \equiv \Omega_{\text{co}} = (0,1)^2$ *or* $\Omega_{\text{non}} = (0,1)^2 \backslash [0, \frac{1}{3}]^2$.

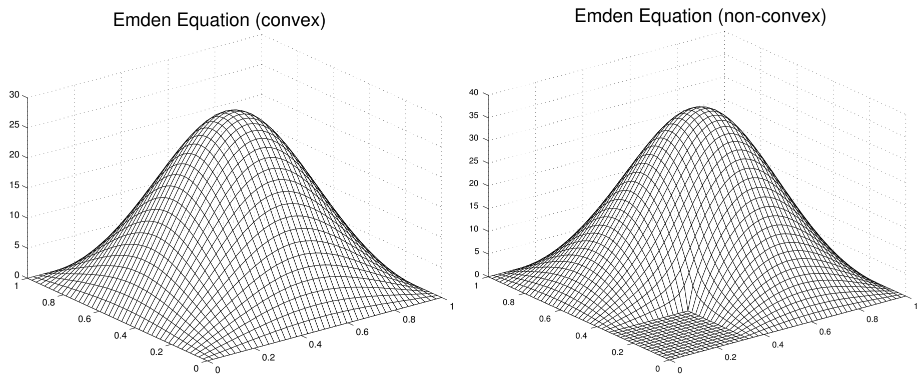

Figure 5: Approximate solutions to Example 3

Example 4 **(Burgers equation)**

$$\begin{array}{rcll} \Delta u &=& \lambda(u \cdot \nabla)u & \text{in} \quad \Omega, \\ u &=& \varphi(x,y) & \text{on} \quad \partial\Omega, \end{array} \tag{5.4}$$

where $\Omega = (0,1)^2$, $\varphi(x,y) = xy(1-y)$ *and* $\lambda > 0$ *is a parameter.*

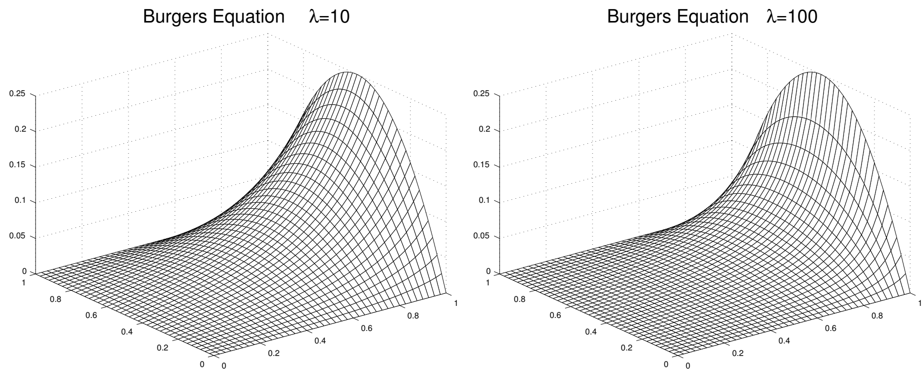

Figure 6: Approximate solutions to Example 4

5.1 Verification results on the invertibility of the linearized operator

First, we consider the invertibility of the linearized operators for examples which are shown above.

For the examples 1-2 and 3 in the nonconvex domain, the linearized operators $\mathcal{L}$ in (4.4) at the approximate solution $\tilde{u}_h \in V_h$ are given as follows :

$$\begin{array}{rcll} \mathcal{L}w &\equiv& -\Delta w + \tilde{u}_h(\nabla\tilde{u}_h \cdot \nabla w) - \left(\lambda - \dfrac{1}{2}|\nabla\tilde{u}_h|^2\right) w. & \text{(Example 1)} \\ \mathcal{L}w &\equiv& -\Delta w + \lambda\Big[a - 2(a+1)\tilde{u}_h + 3\tilde{u}_h^2\Big] w, & \text{(Example 2)} \\ \mathcal{L}w &\equiv& -\Delta w - 2\tilde{u}_h w, & \text{(Example 3)} \end{array}$$

Using piecewise biquadratic C^0-finite element space S_h with several mesh sizes, the constant C_0 for the a priori constant $C(h)$ in Assumption 1 can be taken as $1/(2\pi)$ (see Section 3 and [9]).

Here, for the linear operator $\mathcal{L}v = -\Delta v + b \cdot \nabla v + cv$, we use again notations κ_i, $(1 \le i \le 3)$ as

$$\kappa_1 = \|\ |b|_E\ \|_{L^\infty}, \qquad \kappa_2 = \|\text{div } b\|_{L^\infty}, \qquad \kappa_3 = \|c\|_{L^\infty}.$$

We show numerical verification results for the examples 1-2 and 3 in Tables 1-3 and 4, respectively. Notice that in Examples 1 and 4, for each $w \in H_0^1(\Omega)$, $g(w)$ in (4.4) does not belong to $L^2(\Omega)$. Namely, for the example 1, $g(w) = g_r(w)$ or $g(w) = g_d(w)$ is given as

$$\begin{aligned} g_r(w) \ &\equiv \ (w + v_0 + \tilde{u}_h)\left(\lambda - \frac{1}{2}|\nabla(w + v_0 + \tilde{u}_h)|^2\right) - \tilde{u}_h\left(\lambda - \frac{1}{2}|\nabla\tilde{u}_h|^2\right) \\ &\qquad + \tilde{u}_h(\nabla\tilde{u}_h \cdot \nabla w) - \left(\lambda - \frac{1}{2}|\nabla\tilde{u}_h|^2\right) w, \\ g_d(w) \ &\equiv \ (w + \tilde{u}_h)\left(\lambda - \frac{1}{2}|\nabla(w + \tilde{u}_h)|^2\right) + \Delta\tilde{u}_h \\ &\qquad + \tilde{u}_h(\nabla\tilde{u}_h \cdot \nabla w) - \left(\lambda - \frac{1}{2}|\nabla\tilde{u}_h|^2\right) w, \end{aligned}$$

for $\tilde{u}_h \in V_h = S_h \not\subset H^2(\Omega)$ or $\tilde{u}_h \in V_h \subset H^2(\Omega)$, respectively. So, for the term $w|\nabla w|^2$ in $g(w)$, we use the candidate set such as

$$W_\alpha \equiv \left\{w \in H_0^1(\Omega) \cap L^\infty(\Omega),\ \nabla w \in L^4(\Omega) : \max\{\|w\|_{L^\infty}, \|\nabla w\|_{L^4}\} \le \alpha\right\},$$

for solutions of (5.1).

Then, from the definition of the candidate set, we have to estimate $\|w\|_{L^\infty}$ and $\|\nabla w\|_{L^4}$. A procedure for these estimates is presented in [11][12], and we introduce this details in the below.

Now, for an arbitrary $v \in H^2(\Omega) \cap H_0^1(\Omega)$, we assume that there exist constants $\mathcal{M}_i$, $(0 \le i \le 3)$ such that

$$\begin{aligned} \|v\|_{L^2} \ &\le \ \mathcal{M}_0\|\mathcal{L}v\|_{L^2}, & \|v\|_{H_0^1} \ &\le \ \mathcal{M}_1\|\mathcal{L}v\|_{L^2}, \\ \|\Delta v\|_{L^2} \ &\le \ \mathcal{M}_2\|\mathcal{L}v\|_{L^2}, & |v|_{H^2} \ &\le \ \mathcal{M}_3\|\mathcal{L}v\|_{L^2}, \end{aligned}$$

where $|\cdot|_{H^2}$ denotes the H^2-seminorm. For example, when the constant $\mathcal{M}_1$ is obtained by our method, it follows that

$$\mathcal{M}_0 \le C_p\mathcal{M}_1, \quad \mathcal{M}_2 \le 1 + \kappa_1\mathcal{M}_1 + \kappa_3\mathcal{M}_0.$$

Moreover, if Ω is convex then we have $\mathcal{M}_3 \le \mathcal{M}_2$.

Then, from the constructive approach to the imbedding theory, it follows that

$$\begin{aligned}
\|v\|_{L^\infty} &\le K_0\|v\|_{L^2} + K_1\|v\|_{H_0^1} + K_2|v|_{H^2} &&\le \mathcal{M}_\infty\|\mathcal{L}v\|_{L^2}, \\
\|\nabla v\|_{L^4} &\le \Big[\|v\|_{L^\infty}(\|\Delta v\|_{L^2} + 2|v|_{H^2})\Big]^{1/2} &&\le \mathcal{M}_4\|\mathcal{L}v\|_{L^2},
\end{aligned}$$

where

$$\begin{aligned}
\mathcal{M}_\infty &:= K_0\mathcal{M}_0 + K_1\mathcal{M}_1 + K_2\mathcal{M}_3, \\
\mathcal{M}_4 &:= \Big[\mathcal{M}_\infty(\mathcal{M}_2 + 2\mathcal{M}_3)\Big]^{1/2},
\end{aligned}$$

and K_i, $(0 \le i \le 2)$ are positive constants depending on the domain. In particular, we can take these constants as

$$\begin{aligned}
K_0 &= \begin{cases} 0 & \text{if } \ n = 1 \\ 1 & \text{if } \ n = 2 \\ 1.0708 & \text{if } \ n = 3 \end{cases}, \\
K_1 &= \begin{cases} 1 & \text{if } \ n = 1 \\ 1.1548\sqrt{2/3} & \text{if } \ n = 2 \\ 1.6549\sqrt{3/3} & \text{if } \ n = 3 \end{cases}, \\
K_2 &= \begin{cases} 0 & \text{if } \ n = 1 \\ 0.22361\sqrt{28/45} & \text{if } \ n = 2 \\ 0.41413\sqrt{57/45} & \text{if } \ n = 3 \end{cases},
\end{aligned}$$

if $\Omega = (0,1)^n$.

Therefore, the criterion of the verification in (4.8) is reduced to the following inequality.

$$\|F(W_\alpha)\|_X \le \max\{\mathcal{M}_\infty, \mathcal{M}_4\} \sup_{w \in W_\alpha} \|g(w)\|_{L^2},$$

when $\|w\|_X = \max\{\|w\|_{L^\infty}, \|\nabla w\|_{L^4}\}$.

Table 1: Verification results for Example 1 for $V_h = S_h \not\subset H^2(\Omega)$ and $\lambda = 25$

$1/h$	$\mathcal{M}_1$	$\hat{\mathcal{M}}_1$	M	κ_1	κ_2	κ_3	C_p	$C(h)$
5	—	Fail	2.0337	6.1496	—	25.00	$1/(\sqrt{2}\pi)$	$h/(2\pi)$
10	—	1.5314	2.0351	6.3278	—	25.00	$1/(\sqrt{2}\pi)$	$h/(2\pi)$
20	—	0.6874	2.0353	6.2235	—	25.00	$1/(\sqrt{2}\pi)$	$h/(2\pi)$

$\hat{\mathcal{M}}_1 \equiv \min\{C_p\hat{\mathcal{M}}, \hat{\mathcal{M}}_*\}$.

Table 2: Verification results for Example 1 for $V_h \subset H^2(\Omega)$ and $\lambda = 25$

$1/h$	$\mathcal{M}_1$	$\hat{\mathcal{M}}_1$	M	κ_1	κ_2	κ_3	C_p	$C(h)$
5	3.7677	Fail	2.0338	6.2295	66.35	25.00	$1/(\sqrt{2}\pi)$	$h/(2\pi)$
10	0.6075	1.4491	2.0352	6.2133	71.14	25.00	$1/(\sqrt{2}\pi)$	$h/(2\pi)$
20	0.4912	0.6855	2.0353	6.1990	69.74	25.00	$1/(\sqrt{2}\pi)$	$h/(2\pi)$

Table 3: Verification results for Example 2 for $a = 0.01$ and $\lambda = 150$

	Upper					Lower				
$1/h$	$\mathcal{M}_1$	M	κ_1	κ_2	κ_3	$\mathcal{M}_1$	M	κ_1	κ_2	κ_3
5	Fail	1.4714	0	0	129.32	0.9571	2.5731	0	0	47.36
10	0.5758	1.4708	0	0	129.57	0.6570	2.5844	0	0	47.24
20	0.3786	1.4730	0	0	128.50	0.5994	2.5853	0	0	46.79

It can be taken as $C_p = 1/(\sqrt{2}\pi)$ and $C(h) = h/(2\pi)$. Here, Upper and Lower correspond to the upper and lower solutions of (5.2) in Figure 4.

Table 4: Verification results for Example 3 for Ω_{non}

$1/h$	$\mathcal{M}_1$	M	κ_1	κ_2	κ_3	C_p	$C(h)$
12	Fail	3.7029	0	0	74.4134	$\sqrt{8}/(3\pi)$	$2.5880{\cdot}h/(2\pi)$
24	8.7169	3.6992	0	0	74.1008	$\sqrt{8}/(3\pi)$	$3.2017{\cdot}h/(2\pi)$

We take the approximate finite element subspace V_h as $V_h = S_h$ for Tables 3 and 4.

5.2 Verification results for the existence and local uniqueness of solutions of nonlinear problems

We show verification results for the existence and local uniqueness of solutions for Examples 3 and 4. In this subsection, we take the same approximation subspace S_h as before. That is the piecewise biquadratic C^0-finite element space with the uniform and rectangular mesh.

In the case of the example 3, $\mathcal{L}$ and $g(w)$ in (4.4) are given as follows

$$\begin{aligned}\mathcal{L}w &\equiv -\Delta w - 2\tilde{u}_h w,\\ g_r(w) &\equiv w^2 + 2v_0 w + v_0^2 + 2\tilde{u}_h v_0,\\ g_d(w) &\equiv w^2 + \tilde{u}_h^2 + \Delta\tilde{u}_h.\end{aligned}$$

Therefore, for the candidate set

$$W_\alpha = \{w \in H_0^1(\Omega) : \|w\|_{H_0^1} \le \alpha\},$$

the condition (4.8) is given by

$$\begin{aligned}\mathcal{M}_1 \sup_{w\in W_\alpha} \|w^2 + 2v_0 w + v_0^2 + 2\tilde{u}_h v_0\|_{L^2} \;&\le\; \alpha,\\ \text{or}\qquad \mathcal{M}_1 \sup_{w\in W_\alpha} (\|w^2\|_{L^2} + \|w_0\|_{L^2}) \;&\le\; \alpha,\end{aligned}\tag{5.5}$$

where $w_0 = \Delta\tilde{u}_h + \tilde{u}_h^2$.
By (5.5) and some calculations using the several kinds of norms, e.g., [6], [17] etc., we obtain the existential condition (4.8) of the form

$$\mathcal{M}_1(E_2\alpha^2 + E_1\alpha + E_0) \le \alpha, \tag{5.6}$$

where $E_i,\ 0 \le i \le 2$, are constants dependent on the norms of $\tilde{u}_h$ and v_0. It implies that, for any positive number α satisfying the quadratic inequality (5.6), there exists at least one solution in the set of the form $\tilde{u}_h + v_0 + W_\alpha$. Note that such an α exists if and only if $\mathcal{M}_1(E_1 + 2\sqrt{E_0E_1}) \le 1$. Also, notice that a sufficient condition corresponding to the relation (4.9) can be similarly and readily treated, and it leads to a simple linear inequality in α such that $\mathcal{M}_1(2E_2\alpha + E_1) < 1$. Thus, we can determine two bounds for α, i.e., α_E and α_U, for which we assure the existence and the uniqueness of solutions, respectively.

Table 5: Verification results for Example 3 for $\Omega_{\rm co}$

$1/h$	$\mathcal{M}_1$	E_2	E_1	E_0	smallest α_E	largest α_U
5	1.5018	$1/\pi^2$	0.2441	4.0418	Fail	2.0809545
10	0.7428	$1/\pi^2$	0.0483	0.5195	0.4137281	6.4042950
20	0.6485	$1/\pi^2$	0.0087	0.0630	0.0412247	7.5655481
$1/h$	M	κ_1	κ_2	κ_3	$C(h)$	$\|v_0\|_{H_0^1}$
5	2.7265	0	0	58.97	$h/(2\pi)$	2.0748883
10	2.7455	0	0	58.91	$h/(2\pi)$	0.5480179
20	2.7467	0	0	58.69	$h/(2\pi)$	0.1345409

In this case, it can be taken as $C_p = 1/(\sqrt{2}\pi)$.

Table 5 shows computational results for the domain $\Omega_{\rm co} = (0,1)^2$. In the table, 'smallest α_E' and 'largest α_U' indicate the smallest and the largest bounds α satisfying the verification conditions (4.8) and (4.9), respectively. We take the approximate finite element subspace V_h as $V_h = S_h$.

Next, in the case of the example 4, we consider a modified candidate set of the form

$$W_\alpha \equiv \left\{ w \in H_0^1(\Omega) \cap L^\infty(\Omega) : \max\{\|w\|_{H_0^1}, \|w\|_{L^\infty}\} \le \alpha \right\}. \tag{5.7}$$

Namely, we enclose the solution of (5.4) in the Banach space $X \equiv H_0^1(\Omega) \cap L^\infty(\Omega)$ with norm $\|w\|_X \equiv \max\{\|w\|_{H_0^1}, \|w\|_{L^\infty}\}$. Further we need the inverse norm estimates in the following L^∞ sense:

$$\|v\|_{L^\infty} \le \mathcal{M}_\infty \|\mathcal{L}v\|_{L^2}, \quad \forall v \in H^2(\Omega) \cap H_0^1(\Omega),$$

where $\mathcal{M}_\infty$ can be computed by using $\mathcal{M}_1$ in the section 2 and the constructive approach to the imbedding theory described in [11], [12] (see the previous subsection).
Thus the condition for existence is written as

$$\max\{\mathcal{M}_1, \mathcal{M}_\infty\} \sup_{w \in W_\alpha} \|g(w)\|_{L^2} \le \alpha.$$

Then, the linearized operator $\mathcal{L}$ and the right-hand side $g(w)$ of (4.4) are as follows:

$$\begin{aligned} \mathcal{L}w &\equiv -\Delta w + \lambda(\tilde{u}_h \cdot \nabla)w + \lambda(w \cdot \nabla)\tilde{u}_h, \\ g_r(w) &\equiv -\lambda\left[((w+v_0)\cdot\nabla)(w+v_0) + (\tilde{u}_h \cdot \nabla)v_0 + (v_0 \cdot \nabla)\tilde{u}_h\right]. \end{aligned}$$

The verification conditions using α are similarly represented as in the previous example. That is, corresponding to the condition (5.6), it also leads to the inequality in α of the quadratic form such that

$$\max\{\mathcal{M}_1, \mathcal{M}_\infty\}(B_2\alpha^2 + B_1\alpha + B_0) \le \alpha, \tag{5.8}$$

where $B_i,\ 0 \le i \le 2$, are constants determined similarly as E_i in the previous example. Particularly, for the efficient computations, we used the L^∞ residual method for v_0 ([9]). And the uniqueness condition is also similarly given as before. The verification results for the parameter $\lambda = 10$ are shown in Table 6 with $V_h = S_h$.

Table 6: Verification results for Example 4 for $\lambda = 10$

$1/h$	$\mathcal{M}_1$	$\mathcal{M}_\infty$	B_2	B_1	B_0	smallest α_E	largest α_U
5	0.2907	0.8461	$10\sqrt{2}$	0.5616	0.0519	Fail	0.0219261
10	0.2489	0.7522	$10\sqrt{2}$	0.1355	0.0110	0.0105836	0.0422075
20	0.2358	0.7226	$10\sqrt{2}$	0.0304	0.0024	0.0018544	0.0478497
$1/h$	M	κ_1	κ_2	κ_3	$C(h)$	$\|v_0\|_{H_0^1}$	$\|v_0\|_{L^\infty}$
5	1.0052	3.5430	12.86	κ_2	$h/(2\pi)$	0.0119672	0.0277488
10	1.0054	3.5428	13.11	κ_2	$h/(2\pi)$	0.0028730	0.0067093
20	1.0055	3.5389	13.20	κ_2	$h/(2\pi)$	0.0006714	0.0014803

In this case, it can be taken as $C_p = 1/(\sqrt{2}\pi)$.

Remark 4 *The computational efficiency of the above results, in the example 3, was almost similar to that the existing methods up to now, e.g., comparing with [17]. But, the determination of the range for existence and/or uniqueness as shown in the tables might be impossible for those methods up to now. Particularly, we can find rather wide range which contains no solutions. For example, from the tables 5 and 6, we can conclude that there are no solutions at all for* α *in* $[0.0412247, 7.5655481]$ *and in* $[0.0018544, 0.0478497]$, *respectively. This property should be useful and powerful for the purpose to prove the nonexistence theorem in various kinds of problems.*

Remark 5 *For the present cases, we separately verified the existence and uniqueness by the criteria (4.6) and (4.7), respectively. We can also use another method to prove them simultaneously. Namely, the condition*

$$F(0) + F'(W)W \overset{\circ}{\subset} W$$

is satisfied for the candidate set W, then it implies that a locally unique solution is enclosed in W ([18]).

Remark 6 *All computations in Tables 1-6 are carried out on the Dell Precision 650 Workstation Xeon 3.20GHz Dual-CPU by using INTLAB 5, a toolbox in MATLAB 7.1 developed by Rump [14] for self-validating algorithms. Therefore, all numerical values in these tables are verified data in the sense of strictly rounding error control.*

6 Conclusion

We have proposed a new method and shown the actual effectiveness to the numerical verification of solutions of nonlinear elliptic boundary value problems with second order.

Though we have not yet applied to sufficiently many realistic examples, the results obtained should be sufficient to show the new algorithm actually seems to be superior to existing methods when concerning equations including a first-order derivative. On the other hand, our current method needs to estimate rigorously an upper bound of the largest eigenvalue for the complicated, usually very large, symmetric matrix $\mathbf{L^T G^{-T} D G^{-1} L}$. This should sometimes lead to very high computational costs compared with the previous methods, e.g.,[6]. Therefore, we should be careful in our choice of methods according to the properties of concerned problems. As the future work, it should be important to make many numerical experiments comparing our new method with other verification methods, including Plum's one etc., and to establish some criteria concerning the appropriate choice of the verification methods.

Acknowledgments

The author expresses his hearty thanks to Professor Mitsuhiro T. Nakao for his careful reading of the paper and giving many helpful and useful comments to improve the original manuscript.

References

[1] G. Alefeld, On the convergence of some interval-arithmetic modifications of Newton's method, *SIAM J. Numer. Anal.* **21** (1984), 363-372.

[2] G. Alefeld, J. Herzberger, *Introduction to Interval Computations*, Academic Press, New York, 1983.

[3] K. Hashimoto, K. Kobayashi, M.T. Nakao, Numerical verification methods for solutions of the free boundary problem, to appear in *Numer. Funct. Anal. Opt.*.

[4] K. Nagatou, N. Yamamoto & M.T. Nakao, An approach to the numerical verification of solutions for nonlinear elliptic problems with local uniqueness, *Numer. Funct. Anal. Opt.* **20** (1999), 543-565.

[5] M.T. Nakao, A numerical approach to the proof of existence of solutions for elliptic problems, *JPN J. Appl. Math.* **5** (1988), 313-332.

[6] M.T. Nakao, Solving nonlinear elliptic problems with result verification using an H^{-1} residual iteration, *Computing Suppl.* **9** (1993), 161-173.

[7] M.T. Nakao, Numerical verification methods for solutions of ordinary and partial differential equations, *Numer. Funct. Anal. Opt.* **22** [3&4] (2001), 321-356.

[8] M.T. Nakao & Y. Watanabe, An efficient approach to the numerical verification for solutions of elliptic differential equations, *Numer. Algorithms* **37** [1-4] (2004), 311-323.

[9] M.T. Nakao & N. Yamamoto, Numerical verification of solutions for nonlinear elliptic problems using L^{∞} residual method, *J. Math. Anal. Appl.* **217** (1998), 246-262.

[10] M.T. Nakao, N. Yamamoto, S. Kimura, On best constant in the optimal error estimates for the H_0^1-projection into piecewise polynomial spaces, *J. Approx. Theory* **93** (1998), 491-500.

[11] M. Plum, Explicit H_2-estimates and pointwise bounds for solutions of second-order elliptic boundary value problems, *J. Math. Anal. Appl.* **165** (1992), 36-61.

[12] M. Plum, Numerical existence proofs and explicit bounds for solutions of nonlinear elliptic boundary value problems, *Computing* **149** (1992), 25-44.

[13] M. Plum, Computer-assisted enclosure methods for elliptic differential equations, *Linear Algebra Appl.* **327** (2001), 147-187.

[14] S.M. Rump, *INTLAB-INTerval LABoratory*, a Matlab toolbox for verified computations, version 5. Inst. Infomatik, Technical University of Hamburg-Harburg.

[15] S.M. Rump, *Solving algebraic problems with high accuracy*, A new approach to scientific computation, (eds. Kulisch, U. & Miranker, W.L.), Academic Press, New York, 1983.

[16] M.H. Schultz, *Spline Analysis*, Prentice-Hall, London, 1973.

[17] N. Yamamoto & M.T. Nakao, Numerical verifications for solutions to elliptic equations using residual iterations with higher order finite element, *J. Comput. Appl. Math.* **60** (1995), 271-279.

[18] N. Yamamoto, A numerical verification method for solutions of boundary value problems with local uniqueness by Banach's fixed point theorem, *SIAM J. Numer. Anal.* **35** (1998), 2004-2013.

[19] N. Yamamoto, M.T. Nakao, Numerical verifications of solutions for elliptic equations in nonconvex polygonal domains, *Numer. Math.* **65** (1993), 503-521.