Invited Paper

Numerical verification methods for solutions of semilinear elliptic boundary value problems

Mitsuhiro T. Nakao^{1a)} and Yoshitaka Watanabe^{2b)}

¹ Sasebo National College of Technology

1-1 Okishin-cho, Sasebo, Nagasaki 857-1193, Japan

² Research Institute for Information Technology, Kyushu University

6-10-1 Hakozaki, Higashi-ku, Fukuoka 812-8581, Japan

^{a)} mtnakao@post.cc.sasebo.ac.jp

^{b)} watanabe@cc.kyushu-u.ac.jp

Abstract: This article describes a survey on numerical verification methods for second-order semilinear elliptic boundary value problems introduced by authors and their colleagues. Here “numerical verification” means a computer-assisted numerical method for proving the existence of a solution in a close and explicit neighborhood of an approximate solution. Three kinds of methods based on the infinite dimensional fixed-point theorems using Newton-like operator will be presented. In each verification method, a projection into a finite dimensional subspace and constructive error estimates of the projection play an important and essential role. It is shown that these methods are really useful for actual problems by illustrating numerical examples.

Key Words: numerical verification method, nonlinear differential equation, computer-assisted proof

1. Introduction

Consider the following nonlinear elliptic boundary value problem

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

where Ω is a bounded convex domain in $\mathbb{R}^n$ ($n = 1, 2, 3$) with piecewise smooth boundary $\partial\Omega$ and the map f is assumed to satisfy appropriate conditions described later. The equation (1) has many applications in, for example, mathematical biology/physics and hydrodynamics field, and various analytical or numerical methods and techniques have been proposed for getting the solution u .

In this article, we would like to report *three* numerical verification methods for the problem (1) developed by authors and their colleagues. The aim of numerical verification is to verify the existence of the solution u of (1) around a given approximate solution u_h with mathematical rigorous error bounds in computer.

We now give a brief introduction to each method.

FN-Int — Newton-like iteration with finite dimensional interval coefficients

The verification principle of FN-Int is originated by one of the authors [35] in 1988 and several improvements have been done up to now. The problem (1) is rewritten as the infinite dimensional fixed-point form, and then the equation is decomposed into the finite dimensional part and the infinite dimensional error part. If the both of finite dimensional and error part lead to the contraction maps under suitable assumptions, an infinite dimensional fixed-point theorem implies the existence of the solution in a certain function set [40, 54, 89].

In the self-validating process in computer, the finite dimensional part is taken to be a set of linear combinations of base functions of an approximation subspace with interval coefficients, and Newton-like iteration is executed. Therefore the computation comes down to solving interval linear systems. On the other hand, for the infinite dimensional part, a successive iteration using a priori error estimate of the projection for the solution of Poisson's equations is applied. These two processes are connected with each other at every iteration step.

Many verification results show that FN-Int is sufficiently useful [54]. This is based on a principle of contractility, namely, local contraction property of Newton operator for the finite dimensional part and the same property for the error part which comes from the uniform approximation of a compact operator when the dimension of the finite dimensional approximation subspace is sufficiently large.

FN-Norm — Newton-like iteration with finite dimensional norm bounds

In case that there is the first-order derivative ∇u in f , it was proved that there exist *some cases* for which the effect of error estimations no longer work in the Newton-like iteration, and as a result, the verification process fails even if very finer approximation subspaces are used.

In order to overcome this difficulty, a modified verification method (FN-Norm) using norm estimates instead of solving the interval linear system in the Newton-like iteration is proposed [60]. In FN-Norm the finite dimensional part is replaced by a ball with small radius and main computation centers on validated computations of singular or eigenvalue value problems of an interval matrix.

IN-Linz — infinite dimensional Newton method with linearized operators

One of peculiar features of FN-Int and FN-Norm is that they can avoid the computation of a bound for inverse operator (denoted by L^{-1}) of linearized problems for (1) which is needed in another verification approaches [9].

In the meantime, we proposed, as IN-Linz, an approach that utilizes the direct estimation of the norm of L^{-1} [61] and yields further simplification of the verification procedures. IN-Linz is a direct extension of FN-Norm. We first verify the invertibility for L and compute guaranteed norm bounds for L^{-1} by applying the same principle as in the previous method, i.e., FN-Norm. After that, we confirm the existence of solutions by proving the contractility of the infinite dimensional Newton-like operator with a residual form.

We now give some additionally brief remarks on another direct computational methods of bounds for L^{-1} has been proposed by Plum [72–75] and Oishi [70], for example. From the standpoint based upon the infinite dimensional Newton's method of the residual type, IN-Linz is very close to Plum's and Oishi's method. Plum's method is called the Rayleigh-Ritz and the Lehmann-Goerisch method which uses some homotopic steps with additional base functions and verified computations for relatively small matrix eigenvalue problems. This is considered as a quite different approach from our method. His approach needs not infinite dimensional projection error estimates and is applicable to unbounded domains. However when L is non self-adjoint, higher-order differential base functions are needed in order to work well LL^* -type homotopic steps. Oishi's method bounds the operator norm of L^{-1} by estimations based on numerical computation of the matrix norm of its Galerkin approximation. Although an application result [70] is ordinary differential equations, his procedure describes on general Banach spaces and operators, and has a common point from our method. We also cite some estimations of L^{-1} for parabolic and hyperbolic equations by Minamoto and the first authors of the present paper [16, 17, 19–21].

A comparison of FN-Int, FN-Norm and IN-Linz, with respect to the total computational costs for

verification of nonlinear problems, would very much depend on the individual problem and introduced approximation subspaces. Up to now, we only say: “One must put the right technique in the right place.”

Other applications

Now we list other applications of our verification principle which, due to the page limitation, we would not discuss in more detail in this article.

- elliptic equations with Neumann boundary conditions and systems [98, 99, 103].
- numerical verifications in nonconvex domains [14, 62, 100, 105, 106]
- Stokes [47, 48] and Navier-Stokes equations [31, 62, 91, 97].
- Rayleigh-Bénard heat convection problems [10, 59, 92, 95].
- eigenvalue enclosing and excluding for self-adjoint and non-self-adjoint operators [26–30, 32–34, 52, 55, 87, 96].
- parabolic [39, 44, 67] and hyperbolic equations [45],
- initial value problems for ODEs [15, 107].
- verification of the solution curve, turning point and bifurcation point [18, 22, 23, 49, 68, 88].
- variational inequalities [51, 53, 80–85], singularly perturbed problem [6], and free boundary problems [7, 58]

The contents of this paper are as follows. In Section 2 a fixed-point formulation using Newton-like iteration in certain appropriate function spaces are considered. By virtue of constructive a priori error estimates of the H_0^1 -projection for Poisson’s equations, a computable verification algorithm (FN-Int) and improved computable verification conditions (FN-Norm) are given in Section 3 and 4, respectively. In Section 5 we describe an another fixed-point formulation and existence theorems using the linearized operator (IN-Linz).

For the readability, we sometimes redefine symbols F , T , $\hat{T}$, N_h , G , ρ , etc. in each section.

2. Fixed-point formulation using Poisson’s equations

2.1 Function spaces and weak form

For some integer m , let $H^m(\Omega)$ denote the L^2 -Sobolev space of order m on Ω .

We define

$$H_0^1(\Omega) := \{u(x) \in H^1(\Omega) \mid u(x) = 0, x \in \partial\Omega\}$$

with the inner product $(\nabla u, \nabla v)_{L^2}$ and the norm $\|u\|_{H_0^1(\Omega)} := \|\nabla u\|_{L^2(\Omega)}$, where $(u, v)_{L^2}$ implies L^2 -inner product on Ω . The nonlinear operator $f : H_0^1(\Omega) \rightarrow L^2(\Omega)$ is supposed to be, for the simplicity, continuous, Fréchet differentiable on $H_0^1(\Omega)$ and also maps bounded sets in $H_0^1(\Omega)$ into bounded sets in $L^2(\Omega)$. We rewrite the problem (1) in the following weak form: find $u \in H_0^1(\Omega)$ such that

$$(\nabla u, \nabla v)_{L^2} = (f(\cdot, u, \nabla u), v)_{L^2}, \quad \forall v \in H_0^1(\Omega). \quad (2)$$

2.2 Fixed-point formulation

It is well known [5] that for every $\xi \in L^2(\Omega)$ Poisson’s equation

$$\begin{cases} -\Delta \psi &= \xi, & \text{in } \Omega, \\ \psi &= 0, & \text{on } \partial\Omega \end{cases} \quad (3)$$

has a unique solution $\psi \in H^2(\Omega) \cap H_0^1(\Omega)$ and the estimates

$$|\psi|_{H^2(\Omega)} \leq \|\xi\|_{L^2(\Omega)} \quad (4)$$

hold, where $|\cdot|_{H^2(\Omega)}$ means the semi-norm on $H^2(\Omega)$ defined by

$$|v|_{H^2(\Omega)}^2 = \sum_{i,j=1}^n \|\partial^2 v / \partial x_i \partial x_j\|_{L^2(\Omega)}^2.$$

For $\xi \in L^2(\Omega)$ let $(-\Delta)^{-1}\xi$ be the solution of Poisson's equation (3), then the operator $(-\Delta)^{-1} : L^2(\Omega) \rightarrow H_0^1(\Omega)$ is compact because of the compactness of the embedding $H^2(\Omega) \hookrightarrow H^1(\Omega)$ [1].

From assumptions of f , the nonlinear operator F defined by

$$F := (-\Delta)^{-1}f \quad (5)$$

is also a compact map on $H_0^1(\Omega)$. Then the weak form (2) can be equivalently rewritten as the fixed-point form

$$u = F(\cdot, u, \nabla u) \quad (6)$$

in $H_0^1(\Omega)$. In the below, for the simplicity, we denote $f(u) := f(\cdot, u, \nabla u)$ and $Fu := F(\cdot, u, \nabla u)$.

2.3 Approximation subspace and projection error

Let S_h be a finite dimensional approximation subspace of $H_0^1(\Omega)$ dependent on the parameter $h > 0$. For example, S_h is taken to be a finite element subspace with mesh size h .

Now let $P_h : H_0^1(\Omega) \rightarrow S_h$ denote the H_0^1 -projection defined by

$$(\nabla(\phi - P_h\phi), \nabla v)_{L^2} = 0, \quad \forall v \in S_h, \quad (7)$$

and suppose that P_h has the following approximation properties.

$$\|v - P_h v\|_{H_0^1(\Omega)} \leq C(h)|v|_{H^2(\Omega)}, \quad \forall v \in H_0^1(\Omega) \cap H^2(\Omega), \quad (8)$$

$$\|v - P_h v\|_{L^2(\Omega)} \leq C(h)\|v - P_h v\|_{H_0^1(\Omega)}, \quad \forall v \in H_0^1(\Omega), \quad (9)$$

where $C(h) > 0$ is a positive constant which is *numerically* determined with the property that $C(h) \rightarrow 0$ as $h \rightarrow 0$. We emphasize that especially the estimate (8) is indispensable in our argument and the compactness of the embedding $H^2(\Omega) \hookrightarrow H^1(\Omega)$ is essential in getting the constant $C(h)$ with desired property. Usually the second estimation (9) for P_h is derived by using a technique so called Aubin-Nitsche's trick [4].

These assumptions (8), (9) hold for many finite element subspaces of $H_0^1(\Omega)$ [4, 50, 54, 56, 57, 63] or function spaces of Fourier series with finite truncation [92]. For example it can be taken as $C(h) = h/\pi$ and $h/(2\pi)$ for bilinear and biquadratic element, respectively, for the rectangular mesh on the square domain [50], and $C(h) = 0.493h$ for the linear and uniform triangular mesh of the convex polygonal domain [12]. Furthermore, a constructive a priori L^∞ error estimate for the projection P_h can also be obtained [43, 46].

Remark 1 In case of nonconvex polygonal domain, since the solution of (3) does not belong to $H^2(\Omega)$, we have to estimate a constant $\tilde{C} > 0$ of the form

$$\|v - P_h v\|_{H_0^1(\Omega)} \leq \tilde{C}\|\Delta v\|_{L^2(\Omega)}, \quad \forall v \in \{\phi \in H_0^1(\Omega) \mid \Delta \phi \in L^2(\Omega)\}. \quad (10)$$

There are some useful techniques and consideration to obtain mathematically rigorous upper bounds for the constant $\tilde{C}$ satisfying (10) with adequate order for such nonconvex domains [8, 14, 62, 100, 105, 106].

Since S_h is the closed subspace of $H_0^1(\Omega)$, each element of $H_0^1(\Omega)$ can be uniquely represented as the direct sum of the element of S_h and $S_h^\perp$. Here $S_h^\perp$ stands for the orthogonal complement subspace of S_h in $H_0^1(\Omega)$. Therefore, the fixed-point equation $u = Fu$ in $H_0^1(\Omega)$ can also be uniquely decomposed as the finite dimensional (projection) part and the infinite dimensional (error) part such that

$$\begin{cases} P_h u &= P_h F u, \\ (I - P_h)u &= (I - P_h)F u, \end{cases} \quad (11)$$

which is the basic principle of our verification theory.

2.4 Newton-like operator

In order to apply our verification method for general problems, we introduce the Newton-like method [40]. Note that we can also formulate simple iteration method for F and succeeded the verification for several realistic problems [36] although it needs that the map F is retractive in some neighborhood of the fixed-point to be verified.

For a solution satisfying (11), we fix an approximate solution $u_h \in S_h$ of (2) and define the nonlinear operator $N_h : H_0^1(\Omega) \rightarrow S_h$ by

$$N_h u = P_h u - [I - F'(u_h)]_h^{-1} P_h(u - Fu),$$

where $F'(u_h)$ stands for the Fréchet derivative of F at u_h , and $[I - F'(u_h)]_h^{-1}$ the inverse on S_h of the restriction operator $P_h(I - F'(u_h))|_{S_h}$.

Remark 2 Existence of $[I - F'(u_h)]_h^{-1}$ is equivalent to the invertibility of a matrix G defined by (21) in next section, which is able to be numerically checked in the actual verification process.

Remark 3 We can replace $P_h F'(u_h)$ by some appropriate operators, which means f is not necessarily Fréchet differentiable (see the verification example later).

Now we define the infinite dimensional operator T on $H_0^1(\Omega)$ by

$$Tu = N_h u + (I - P_h)Fu. \quad (12)$$

Then T is considered as the Newton-like operator for the former part of (11) but the simple iterative operator for the latter part. Moreover T is compact on $H_0^1(\Omega)$ by the compactness of F .

Furthermore, it can be seen that

Proposition 1 The fixed-point equation

$$u = Tu \quad (13)$$

is equivalent to (6).

Proof. If $u = Fu$, namely, (11) holds, then we have, using the former part,

$$0 = [I - F'(u_h)]_h^{-1} P_h(u - Fu), \quad (14)$$

which yields

$$P_h u = P_h N_h u. \quad (15)$$

Therefore, $u = Tu$ follows by adding (15) to each side of the latter part of (11). Conversely, if $u = Tu$, then we immediately obtain (15), and thus (14) follows. The conclusion would be straightforward. $\square$

Since T is the compact map on $H_0^1(\Omega)$, the Schauder fixed-point theorem asserts that if for a nonempty, bounded, convex and closed set $U \subset H_0^1(\Omega)$,

$$TU \subset U$$

holds, then there exists a fixed-point of T in U . We call such a set U , expected to be $TU \subset U$, as a *candidate set*.

If a candidate set U is chosen such as

$$U := u_h + U_h + U_*, \quad U_h \subset S_h, \quad U_* \subset S_h^\perp, \quad (16)$$

then the verification condition $TU \subset U$ can be written by

$$\begin{cases} N_h U - u_h & \subset U_h, \\ (I - P_h)FU & \subset U_*. \end{cases} \quad (17)$$

Note that when the approximate solution $u_h \in S_h$ of (2) is sufficiently good, the finite dimensional part will be possibly contractive. Moreover, the magnitude of the infinite dimensional part of T , i.e., $(I - P_h)Fu$, is expected to be contractive when the parameter h of S_h are taken to be sufficiently small because of the approximation property (8) with (4) of P_h .

3. Verification method FN-Int

This section is devoted to the construction of the first verification algorithm FN-Int which generates candidate sets expected to satisfy the condition (17). The direct solving a finite dimensional problem with additional error estimates is the essential point of this method.

3.1 Infinite dimensional part

The following general criterion is led by the constructive a priori error estimates for Poisson's equations [35].

Theorem 1 Let the infinite dimensional term of the candidate set U_* in (16) be a ball with radius $\alpha > 0$ such as

$$U_* := \{u_* \in S_h^\perp \mid \|u_*\|_{H_0^1(\Omega)} \leq \alpha\}. \quad (18)$$

If for a bounded closed subset $U_h \subset S_h$ and the candidate set $U = u_h + U_h + U_*$,

$$C(h) \sup_{u \in U} \|f(u)\|_{L^2(\Omega)} \leq \alpha \quad (19)$$

holds, then the second term of the condition (17) is satisfied.

Proof. For each $u \in U$, from (4), (8), and (19), we have

$$\begin{aligned} \|(I - P_h)Fu\|_{H_0^1(\Omega)} &\leq C(h)|Fu|_{H^2(\Omega)} \\ &\leq C(h)\|f(u)\|_{L^2(\Omega)} \\ &\leq \alpha, \end{aligned}$$

which shows $u \in U_*$ thus $(I - P_h)FU \subset U_*$. $\square$

3.2 Finite dimensional part using interval coefficients

We present a procedure to make up the finite dimensional candidate set. Let K be the dimension of S_h , $\{\phi_i\}_{1 \leq i \leq K}$ a basis of S_h , and $\mathbb{IR}^K$ the set of K -dimensional interval vector.

The set U_h is taken to be a set of linear combinations of base functions in S_h with interval coefficients such as

$$U_h := \sum_{i=1}^K B_i \phi_i, \quad (20)$$

where B_i has upper and lower bounds such that $B_i = [\underline{B}_i, \overline{B}_i]$. Here $\sum_{i=1}^K B_i \phi_i$ is interpreted as the set of functions in which each element is linear combination of $\{\phi_i\}_{1 \leq i \leq K}$ whose coefficient of ϕ_i belongs to the corresponding interval $[\underline{B}_i, \overline{B}_i]$ for each $1 \leq i \leq K$.

Now, defining $K \times K$ matrix G by

$$[G]_{ij} = (\nabla \phi_j, \nabla \phi_i)_{L^2} - (f'(u_h) \phi_j, \phi_i)_{L^2}, \quad (21)$$

we obtain a sufficient condition of the first part of (17) [54].

Theorem 2 Assume that the candidate set $U \subset H_0^1(\Omega)$ is defined by (18) and (20) with (16), and that any element $u \in U$ is represented by

$$u = u_h + \hat{u}_h + u_*, \quad \hat{u}_h \in U_h, \quad u_* \in U_*.$$

Let $\mathbf{d} = (d_i) \in \mathbb{IR}^K$ denote an interval enclosure of the set whose i -th component consists of

$$\{(f(u) - f'(u_h)\hat{u}_h, \phi_i)_{L^2} - (\nabla u_h, \nabla \phi_i)_{L^2} \in \mathbb{R} \mid u \in U\}, \quad 1 \leq i \leq K. \quad (22)$$

If, for an interval vector $\mathbf{v} = (v_i) \in \mathbb{IR}^K$ enclosing the solution $\mathbf{x} \in \mathbb{IR}^K$ for the linear equation

$$G\mathbf{x} = \mathbf{d}, \quad (23)$$

the conditions

$$v_i \subset B_i, \quad 1 \leq i \leq K \quad (24)$$

hold, then the former half of the condition (17) is satisfied.

Proof. For all $u = u_h + \hat{u}_h + u_* \in U$, since

$$\begin{aligned} N_h u - u_h &= \hat{u}_h - [I - F'(u_h)]_h^{-1} P_h(u - Fu) \\ &= [I - F'(u_h)]_h^{-1} (P_h(I - F'(u_h)\hat{u}_h) - P_h(u - Fu)) \\ &= [I - F'(u_h)]_h^{-1} P_h(Fu - F'(u_h)\hat{u}_h - u_h), \end{aligned}$$

by setting

$$v_h := N_h u - u_h = \sum_{i=1}^K v_i \phi_i, \quad \mathbf{v} := (v_i) \in \mathbb{R}^K,$$

we have

$$P_h(I - F'(u_h))v_h = P_h(Fu - F'(u_h)\hat{u}_h - u_h). \quad (25)$$

From the definition of P_h , the equation (25) is equivalent to

$$\sum_{i=1}^K \{(\nabla \phi_i, \nabla \phi_j)_{L^2} - (f'(u_h)\phi_i, \phi_j)_{L^2}\} v_i = (f(u) - f'(u_h)\hat{u}_h, \phi_j)_{L^2} - (\nabla u_h, \nabla \phi_j)_{L^2}, \quad (26)$$

$$1 \leq j \leq K.$$

Then the condition (24) assures

$$v_h = N_h u - u_h \in U_h,$$

which proves the theorem. $\square$

Based on Theorem 1 and 2, we give the following verification algorithm FN-Int.

Verification algorithm FN-Int

- $k = 0$

Set initial values $B_i^{(0)} \in \mathbb{IR}$ ($1 \leq i \leq K$) and $\alpha^{(0)} > 0$.

- $k \geq 1$

1. For a fixed small constant $\varepsilon > 0$ set

$$\hat{B}_i^{(k)} := (1 + \varepsilon) B_i^{(k-1)} \quad (1 \leq i \leq K), \quad \hat{\alpha}^{(k)} := (1 + \varepsilon) \alpha^{(k-1)}.$$

2. The k -th candidate set $U^{(k)}$ is defined by

$$\begin{aligned} U_h^{(k)} &:= \sum_{i=1}^K \hat{B}_i \phi_i \subset S_h, \\ U_*^{(k)} &:= \{v_* \in S_h^\perp \mid \|v_*\|_{H_0^1(\Omega)} \leq \hat{\alpha}^{(k)}\}, \\ U^{(k)} &:= u_h + U_h^{(k)} + U_*^{(k)}. \end{aligned}$$

3. Compute values of k -th iteration by

$$\begin{aligned} B_i^{(k)} &:= v_i \text{ of (23) in Theorem 2 for } U^{(k)}, \\ \alpha^{(k)} &:= C(h) \sup_{u \in U^{(k)}} \|f(u)\|_{L^2(\Omega)}. \end{aligned}$$

4. If $B_i^{(k)} \subset \hat{B}_i^{(k)}$ ($1 \leq i \leq K$) and $\alpha^{(k)} \leq \hat{\alpha}^{(k)}$ then stop, and there exists a desired solution in $U^{(k)} \subset H_0^1(\Omega)$.
5. Setting $k := k + 1$ and return to the step 1. If k reaches a maximum iteration number or some norm of $B_i^{(k)}$ and $\alpha^{(k)}$ exceed a criterion then stop, and the verification fails.

Remark 4 The extension procedure involving ε occurring in this algorithm is called “ ε -inflation” which is a kind of acceleration technique [78]. The concrete value of $\varepsilon > 0$ should be adapted to the actual problem.

Remark 5 The interval vector $\mathbf{v} \in \mathbb{IR}^K$ in Theorem 2 is automatically obtained in computer by verified solution of some linear systems with matrix G (for example see Rump [76]).

Remark 6 In the actual computation, we have to treat the infinite dimensional term U_* . Moreover, it is impossible to avoid the effect of rounding error of floating point arithmetic. However, by norm estimates and interval arithmetic software taking into account effects of rounding error, we can obtain mathematically rigorous enclosing or upper bounds for $B_i^{(k)}$ and $\alpha^{(k)}$ in the *over-estimated sense* using a priori norm estimates and the interval arithmetic.

3.3 A verification example of FN-Int

We verified a nontrivial solution of the following simplifying Grad-Shafranov equation on the unit square $\Omega = (0, 1) \times (0, 1)$ [90].

$$\begin{cases} -\Delta u &= \lambda \max\{0, u\} & \text{in } \Omega, \\ u &= -1 & \text{on } \partial\Omega, \end{cases} \quad (27)$$

where $\lambda \geq 0$. The Eq. (27) arises in equilibrium analysis of confined MHD (magnetohydrodynamics) plasmas in an infinite cylindrical domain. If $u(x)$ is a nontrivial solution of Eq. (27) for some λ , in a domain $\Omega_p \equiv \{x \in \Omega; u(x) \geq 0\}$, $w := u|_{\Omega_p}$ is an eigenfunction of $-\Delta w = \lambda w$ in Ω_p , $w = 0$ on $\partial\Omega_p$. Then Eq. (27) can be considered of a nonlinear eigenvalue problem, and the portion of Ω where $u(x) = 0$ can be regarded as a free boundary corresponding to the boundary of plasmas.

Although the nonlinear term $\max\{0, u\}$ is non-differentiable, we can make an extension of the derivative of f , and a suitable modification of the Fréchet derivative of F . Let $\delta_x : 0 = x_0 < x_1 < \dots < x_N = 1$ be a uniform partition in x -direction, and let δ_y be the same partition as δ_x for y direction. We define the partition of Ω by $\delta \equiv \delta_x \otimes \delta_y$. Further, we define the finite element subspace S_h by $S_h \equiv \mathcal{M}_0^1(x) \otimes \mathcal{M}_0^1(y)$, where $\mathcal{M}_0^1(x)$, $\mathcal{M}_0^1(y)$ are sets of piecewise linear polynomials on $(0, 1)$ in the variables x and y , respectively. We can take the constant $C(h)$ of (8) as h/π [35].

For $\lambda = 30$, number of partitions $N = 50$, the algorithm FN-Int succeeded in the verification with iteration number 7 and $\alpha^{(7)} = 0.10956$. Figure 1 shows the shape of approximate solution for $u + 1$ and the range of a free boundary, respectively.

Fig. 1. Shape of approximate solution and range of the free boundary.

3.4 Accuracy improvement by the residual method

For the case in which the value of right-hand side $f(u)$ of the equation (1) becomes very large, $\alpha^{(k)}$ in the iteration of FN-Int may become very large too, which causes the interval explosion and therefore the verification scheme fails.

Some residual techniques are useful to overcome this difficulty. First improvement is usual “direct” residual method. When $\Delta u_h \in L^2(\Omega)$, the problem (1) is equivalent to finding a solution $w = u - u_h$ satisfying

$$\begin{cases} -\Delta w &= f(w + u_h) + \Delta u_h, & \text{in } \Omega, \\ w &= 0, & \text{on } \partial\Omega. \end{cases} \quad (28)$$

If u_h is chosen as a sufficiently good approximation of the exact solution, then the norm of the right-hand side in Eq. (28) and $\alpha^{(k)}$ in FN-Int become small [89]. Also we can take an another approximate solution $\hat{u}_h$ from the original u_h satisfying $\Delta \hat{u}_h \in L^2(\Omega)$ and consider the residual problem (28) using $\hat{u}_h$, where $\hat{u}_h$ itself does not necessarily belong to S_h .

Second one is “ H^{-1} -type” residual method [42]. Assume that the approximate solution $u_h \in S_h$ satisfies

$$(\nabla u_h, \nabla \phi_h)_{L^2} = (f(u_h), \phi_h)_{L^2}, \quad \forall \phi_h \in S_h. \quad (29)$$

It is rather natural assumption for the existence of an approximate solution satisfying Eq. (29), for it is a finite dimensional problem. It is not so difficult to obtain strict enclosures of u_h by some interval methods [2, 24, 69]. Further, to calculate of $F'(u_h)$, it will be sufficient to take a point element, i.e., an approximation of u_h , instead of interval enclosure of u_h itself.

Now we take an element $\bar{u} \in H_0^1(\Omega) \cap H^2(\Omega)$ as the solution to the following Poisson equation

$$\begin{cases} -\Delta \bar{u} &= f(u_h), & \text{in } \Omega, \\ \bar{u} &= 0, & \text{on } \partial\Omega, \end{cases} \quad (30)$$

and try to find an exact solution u around $\bar{u}$. The fact that $v_0 := \bar{u} - u_h \in S_h^\perp$ implies $P_h \bar{u}$ coincides with u_h . Therefore, while the explicit form of v_0 is unknown, the norm can be estimated as follows, by using the constructive a priori error estimates (8), (9) for P_h ,

$$\|v_0\|_{H_0^1(\Omega)} \leq C(h)\|f(u_h)\|_{L^2(\Omega)}, \quad \|v_0\|_{L^2(\Omega)} \leq C(h)^2\|f(u_h)\|_{L^2(\Omega)}. \quad (31)$$

Thus, in this case, we consider the candidate set U of the form

$$U = u_h + v_0 + \sum_{i=1}^K B_i \phi_i + U_*.$$

By setting $w = u - \bar{u}$, Eq. (1) is rewritten as the following residual form in w .

$$\begin{cases} -\Delta w &= f(w + u_h + v_0) - f(u_h), & \text{in } \Omega, \\ w &= 0, & \text{on } \partial\Omega. \end{cases} \quad (32)$$

Here, v_0 stands for an element in $H_0^1(\Omega)$ which is determined by the Riesz representation theorem from the residual functional $\Delta u_h + f(u_h) \in H^{-1}$ [54]. Then we need not the condition $\Delta u_h \in L^2(\Omega)$. Moreover, we can obtain an a posteriori error estimation for the norm of v_0 . Let $S_h^* \subset H^1(\Omega)$ be a finite element subspace whose basis consists of the basis of S_h and the base functions having nonzero values on the boundary $\partial\Omega$. We define the two dimensional vector valued function $\bar{\nabla} u_h \in S_h^* \times S_h^*$ by the L^2 projection of $\nabla u_h \in L^2(\Omega) \times L^2(\Omega)$ into $S_h^* \times S_h^*$ and set $\bar{\Delta} u_h := \nabla \cdot \bar{\nabla} u_h \in L^2(\Omega)$. Then we can easily obtain the following identity:

$$(\bar{\nabla} u_h, \nabla v)_{L^2} + (\bar{\Delta} u_h, v)_{L^2} = 0, \quad \forall v \in H_0^1(\Omega), \quad (33)$$

therefore

$$(\nabla v_0, \nabla v)_{L^2} = (\bar{\nabla} u_h - \nabla u_h, \nabla v)_{L^2} + (\bar{\Delta} u_h + f(u_h), v)_{L^2} = 0, \quad \forall v \in H_0^1(\Omega). \quad (34)$$

Here choosing $v = v_0$ in (34) and using the Aubin-Nitsche inequality $\|v_0\|_{L^2(\Omega)} \leq C(h)\|v_0\|_{H_0^1(\Omega)}$, we get the following a posteriori estimate:

$$\|v_0\|_{H_0^1(\Omega)} \leq \|\bar{\nabla} u_h - \nabla u_h\|_{L^2(\Omega)} + C(h)\|\bar{\Delta} u_h + f(u_h)\|_{L^2(\Omega)}. \quad (35)$$

If we use the higher order element, the right-hand side of (35) will converge to 0 with higher order than $C(h)$ [46, 101]. Thus we can expect that the a posteriori error estimates for the right-hand side can be actually smaller than the a priori estimates (31). Therefore, it will be possible that, if we use a higher order finite element to obtain the approximation u_h , then it will enable us to verify with high accuracy even for rather coarse mesh.

We finally notice that this technique can also be extended for the nonconvex and nonsmooth domain such as L-shape domain for which the solution has low regularity and we could no longer apply the usual a priori estimates [100].

3.5 Verification with local uniqueness

For the fixed-point form (13): $u = Tu$ on $H_0^1(\Omega)$, set $w = u - u_h$ and consider a residual equation

$$w = \hat{T}w,$$

where

$$\hat{T}w := T(w + u_h) - u_h.$$

Let $W \subset H_0^1(\Omega)$ be a bounded closed and convex set which contains 0 as an inner point satisfying

$$aW \subset W, \quad \forall a \in \mathbb{R}, \quad |a| \leq 1.$$

Then by the Banach fixed-point theorem, if the inclusion

$$\{\hat{T}(0) + \hat{T}'(w_1)w_2 \in H_0^1(\Omega) \mid w_1, w_2 \in W\} \subset \overset{\circ}{W} \quad (36)$$

holds, we can find that there exists a fixed-point of $\hat{T}$ in $\hat{T}W$ and it is unique within W . Here $\hat{T}'$ stands for the Fréchet derivative of $\hat{T}$ and $\overset{\circ}{W}$ the set of interior points of W .

Based on the verification technique described above, we can modify the verification algorithm FN-Int for proving the existence and local uniqueness condition (36) [25, 102, 108].

4. Verification method FN-Norm

In this section, we describe the second verification algorithm FN-Norm.

4.1 Some difficulties in FN-Int

A lot of verification results prove the actual effectiveness of FN-Int if $f(\cdot)$ in (1) has *no* first-order derivative [36–38, 41, 54]. However, in the case that there is a first-order derivative in f , the verification algorithm does not always work well [40, 87]. We now give a simple example [60]. When $f(u) \equiv \nabla u$, the set in (22) becomes

$$-(u_h + \nabla u_h, \nabla \phi_i)_{L^2} + (\nabla u_*, \phi_i)_{L^2}. \quad (37)$$

Here, the first term $-(u_h + \nabla u_h, \nabla \phi_i)_{L^2}$ of (37) could be estimated by a narrow interval using usual interval arithmetic. Conversely, the second term $(\nabla u_*, \phi_i)_{L^2}$ would be estimated as $|(\nabla u_*, \phi_i)_{L^2}| \leq \alpha \|\phi_i\|_{L^2(\Omega)}$ which generally causes an overestimation of d_i in (23) as the interval enclosing

$$-(u_h + \nabla u_h, \nabla \phi_i)_{L^2} + [-1, 1]\alpha \|\phi_i\|_{L^2(\Omega)}$$

unless basis have any specific properties. For example, if we take $\Omega = (0, 1)$ and S_h the set of piecewise linear function with uniform mesh size h , then $\|\phi_i\|_{L^2(\Omega)}$ is $(2h/3)^{1/2}$, and note that d_i is $O(h^{1/2})$. Taking account that the dimension of the matrix G is proportional to $1/h$, one can deduce that $B_i^{(k)} \subset \hat{B}_i^{(k)}$ might not occur even if h tends to be small. The situation should be the same even if we use estimate such as $|(\nabla u_*, \phi_i)_{L^2}| \leq \|\nabla u_*\|_{L^2(\Omega)} \|\phi_i\|_{L^2(\Omega)}$. Moreover, there might be some possibilities that the norm of $\|f(u)\|_{L^2(\Omega)}$ could be explosive because of the property of interval arithmetic.

4.2 Improvement of computation for finite dimensional part

In order to overcome these difficulties described in the previous subsection, a slightly modified method FN-Norm is introduced to compute the finite dimensional part of the candidate set using by norm estimates. Let U_h be the finite dimensional term of a candidate set in (16) defined as a ball with radius $\gamma > 0$ such that

$$U_h := \left\{ \hat{u}_h \in S_h \mid \|\hat{u}_h\|_{H_0^1(\Omega)} \leq \gamma \right\}. \quad (38)$$

We define the $K \times K$ positive definite matrix A_1 by

$$[A_1]_{mn} = (\nabla \phi_n, \nabla \phi_m)_{L^2}, \quad (39)$$

the decomposed factor L_1 for A_1 by $A_1 = L_1 L_1^T$, and the matrix norm $\rho > 0$ by

$$\rho = \|L_1^T G^{-1} L_1\|_E, \quad (40)$$

where G is defined in (21) and $\|\cdot\|_E$ is the Euclidean norm. Then a different verification condition from (24) in Theorem 2 for finite dimensional part can be obtained.

Theorem 3 Let define the candidate set $U = u_h + U_h + U_*$ by (38) and (18), and let any element $u \in U$ be represented by

$$u = u_h + \hat{u}_h + u_*, \quad \hat{u}_h \in S_h, \quad u_* \in S_h^\perp.$$

If it holds that

$$\rho \sup_{u \in U} \|P_h F u - P_h F'(u_h) \hat{u}_h - u_h\|_{H_0^1(\Omega)} \leq \gamma, \quad (41)$$

then the former half of the condition (17) is satisfied.

Remark 7 In (41), $P_h F u - P_h F'(u_h) \hat{u}_h - u_h$ corresponds to the higher-order residual for Taylor expansion of $P_h F u$ at u_h , hence the inequality (41) would be expected to hold if the radius of U_h and U_* tend to be small when ρ is bounded independently of h . In the actual calculation on a computer, we make an over-estimation $\|P_h F u - P_h F'(u_h) \hat{u}_h - u_h\|_{H_0^1(\Omega)}$ using α , γ and a priori constant $C(h)$.

Remark 8 The estimate of ρ in (40) is, generally, singular value problem of a matrix. There are some verification algorithms to estimate a rigorous bound for the maximum (or minimum) singular value [77, 93, 104]. Also, the interval Cholesky decomposition algorithm is usually feasible because of the positive definiteness of A_1 [3].

Proof. From (17), it is sufficient to show that

$$\|N_h u - u_h\|_{H_0^1(\Omega)} \leq \rho \|P_h F u - P_h F'(u_h) \hat{u}_h - u_h\|_{H_0^1(\Omega)} \quad (42)$$

for each $u = u_h + \hat{u}_h + u_* \in U$.

It is easily shown that for $v_h = \sum_{i=1}^K v_i \phi_i \in S_h$ and a K dimensional vector $\mathbf{v} := (v_i)$, $\|v_h\|_{H_0^1(\Omega)} = \|L_1^T \mathbf{v}\|_E$ holds. Since

$$N_h u - u_h = [I - F'(u_h)]_h^{-1} P_h (F u - F'(u_h) \hat{u}_h - u_h), \quad (43)$$

by setting

$$v_h := N_h u - u_h = \sum_{i=1}^K v_i \phi_i, \quad \mathbf{v} := (v_i) \in \mathbb{R}^K,$$

$$w_h := P_h (F u - F'(u_h) \hat{u}_h - u_h) = \sum_{i=1}^K w_i \phi_i, \quad \mathbf{w} := (w_i) \in \mathbb{R}^K,$$

the equation (43) is written as

$$P_h (I - F'(u_h)) v_h = w_h. \quad (44)$$

From the definition of P_h , the equation (44) is equivalent to

$$\sum_{i=1}^K \{(\nabla \phi_i, \nabla \phi_j)_{L^2} - (f'(u_h) \phi_i, \phi_j)_{L^2}\} v_i = \sum_{i=1}^K (\nabla \phi_i, \nabla \phi_j)_{L^2} w_i, \quad 1 \leq j \leq K. \quad (45)$$

Then the equation (45) can be represented by the matrix and vector form as

$$\mathbf{v} = G^{-1} A_1 \mathbf{w}.$$

Consequently,

$$\begin{aligned} \|N_h u - u_h\|_{H_0^1(\Omega)} &= \|L_1^T \mathbf{v}\|_E \\ &= \|L_1^T G^{-1} A_1 \mathbf{w}\|_E \\ &\leq \|L_1^T G^{-1} L_1\| \|L_1^T \mathbf{w}\|_E \\ &= \rho \|P_h F u - P_h F'(u_h) \hat{u}_h - u_h\|_{H_0^1(\Omega)}, \end{aligned}$$

which proves the theorem. $\square$

Based on Theorem 3, we obtain the following verification algorithm of FN-Norm.

Verification algorithm FN-Norm

- $k = 0$

Set initial values $\gamma^{(0)} > 0$ and $\alpha^{(0)} > 0$.

- $k \geq 1$

1. For a fixed small constant $\varepsilon > 0$ set

$$\hat{\gamma}^{(k)} := (1 + \varepsilon) \gamma^{(k-1)}, \quad \hat{\alpha}^{(k)} := (1 + \varepsilon) \alpha^{(k-1)}.$$

2. The k -th candidate set $U^{(k)}$ is defined by

$$U_h^{(k)} := \{\hat{v}_h \in S_h \mid \|\hat{v}_h\|_{H_0^1(\Omega)} \leq \hat{\gamma}^{(k)}\},$$

$$U_*^{(k)} := \{v_* \in S_h^\perp \mid \|v_*\|_{H_0^1(\Omega)} \leq \hat{\alpha}^{(k)}\},$$

$$U^{(k)} := u_h + U_h^{(k)} + U_*^{(k)}$$

3. Compute values of k -th iteration by

$$\gamma^{(k)} := \sup_{u \in U^{(k)}} \rho \|P_h F u - P_h F'(u_h) \hat{u}_h - u_h\|_{H_0^1(\Omega)},$$

$$\alpha^{(k)} := C(h) \sup_{u \in U^{(k)}} \|f(u)\|_{L^2(\Omega)}.$$

4. If $\gamma^{(k)} \leq \hat{\gamma}^{(k)}$ and $\alpha^{(k)} \leq \hat{\alpha}^{(k)}$ hold then stop, and there exists a desired solution in $U^{(k)} \subset H_0^1(\Omega)$.

5. Setting $k := k + 1$ and return to the step 1. If k reaches a maximum iteration number or $\gamma^{(k)}$ and $\alpha^{(k)}$ exceed a criterion then stop, and the verification fails.

The essential difference between the algorithm FN-Int and FN-Norm is that the finite dimensional part U_h of the candidate set is taken as a ball U_h in $H_0^1(\Omega)$ in the latter. Using the norm estimates, one can expect to avoid the drawbacks of FN-Int, namely, the local over-estimates in the calculation of the vector $\mathbf{d}$ and the risk of explosive enlargement of the candidate set caused by interval computations.

Of course, we need some additional task to estimate the norm of the composite matrix $L_1^T G^{-1} L_1$, which is sometimes higher than the computational cost in the FN-Int.

4.3 A comparison of FN-Int and FN-Norm

We now give some numerical examples which confirm us the effectiveness of the algorithm FN-Norm.

In the verification step, interval arithmetic is used to take account of the effects of rounding errors in the floating point computations. We use Sun ONE Studio 7, Compiler Collection Fortran 95 [86] on FUJITSU GP7000F model 900 (CPU: SPARC64-GP 400MHz, OS: Solaris7).

Consider the following two dimensional problem

$$\begin{cases} -\Delta u = k \cdot \nabla u + u + \omega & \text{in } \Omega, \\ u = 0 & \text{on } \partial\Omega, \end{cases} \quad (46)$$

where $\Omega = (0, 1) \times (0, 1)$, $k = (k_1, k_2)^T \in \mathbb{R}^2$ is a parameter, and ω is chosen such that $u = \sin(\pi x) \sin(\pi y)$ is the exact solution of (46). The approximate subspace S_h is taken as the following set of double finite Fourier series:

$$S_h = \left\{ \sum_{m,n=1}^N a_{mn} \sin(\pi m x) \sin(\pi n y); a_{mn} \in \mathbb{R} \right\}.$$

Then $\dim S_h = N^2$, $h = 1/N$ and a priori constant $C(h)$ can be taken as $1/(\pi \sqrt{((N+1)^2 + 1)})$ [87]. Table I shows the verification results.

Table I. Verification results for the two-dimension case $k = k_1 = k_2$.

k	FN-Int				FN-Norm			
	$\max_i B_i $		$\ U_*\ _{H_0^1(\Omega)}$		$\ U_h\ _{H_0^1(\Omega)}$		$\ U_*\ _{H_0^1(\Omega)}$	
	$N = 10$	$N = 30$	$N = 10$	$N = 30$	$N = 10$	$N = 30$	$N = 10$	$N = 30$
0.0	0.00089	0.00012	0.28468	0.10131	0.00198	0.00026	0.28468	0.10131
0.5	0.03883	0.02083	0.29325	0.10263	0.05153	0.01755	0.29208	0.10222
1.0	0.09043	0.03809	0.31275	0.10660	0.10350	0.03496	0.30204	0.10341
1.5	0.15381	0.06118	0.36187	0.11753	0.15899	0.05251	0.31492	0.10489
2.0	0.30024	0.10475	0.51491	0.15025	0.22030	0.07068	0.33127	0.10667
2.5	2.85612	0.36382	3.83422	0.40620	0.29205	0.08969	0.35260	0.10878
3.0	$\times$	$\times$	$\times$	$\times$	0.37641	0.10966	0.38025	0.11125
4.0	$\times$	$\times$	$\times$	$\times$	0.60983	0.15309	0.46481	0.11740
5.0	$\times$	$\times$	$\times$	$\times$	1.02990	0.20575	0.63113	0.12604

Though this example is artificial, one can see that FN-Int fails when the terms of the first-order derivative are dominating, while FN-Norm works well for all cases.

4.4 A verification example of FN-Norm

As a more realistic application of FN-Norm we examined the following two-point boundary value problem for Duffing's equation.

$$\begin{cases} -w'' = kw' + w^3 - B \cos x, & x \in (0, 2\pi), \\ w(0) = w(2\pi) = p, \end{cases} \quad (47)$$

where $k = 0.1$, $B = 0.4464$ and $p = -0.6644$. Setting $u = w - \hat{w}$ with $\hat{w} := \sin(x/2) + p$, the Eq. (47) can be rewritten as

$$\begin{cases} -u'' = k(u + \hat{w})' + (u + \hat{w})^3 - B \cos x + \hat{w}'', & x \in (0, 2\pi), \\ u(0) = u(2\pi) = 0. \end{cases} \quad (48)$$

Figure 2 shows the shape of approximate solution for $w = u + \hat{w}$. We divide the interval $(0, 2\pi)$ into $K + 1$ equal parts and take the finite element subspace S_h whose elements are piecewise linear polynomials on each subinterval. The constant $C(h)$ of (8) can be taken as $1/((K - 1)\pi)$. We use interval arithmetic toolbox INTLAB [79] Version 6 with MATLAB 7.7.0.471 (R2008b) on Intel Core 2 Duo 2.6GHz (OS: Mac OS X 10.5.4).

Fig. 2. Shape of approximate solution for Duffing equation.

For $K = 250$, FN-Norm succeeded in the verification with iteration number 6 by

$$\gamma^{(6)} = 0.000621257124644, \quad \alpha^{(6)} = 0.008635570049880.$$

Remark 9 For one-dimensional case covering Duffing's equation (47), an effective improvement of FN-Int based on estimations of the infinite dimensional error part α on each subinterval have been proposed [40]. This refinement technique should be comparable to FN-Norm.

Remark 10 The accuracy improvement by residual formulation and verification with local uniqueness described in previous section are also applicable to FN-Norm [94, 108].

5. Verification method IN-Linz

In this section, we propose the third verification method IN-Linz based upon the infinite dimensional Newton's method of the residual type. This approach utilizes the direct estimation of the norm of linearized inverse operators and yield further simplification of the verification procedures compared with FN-Int and FN-Norm. IN-Linz is an extension of FN-Norm. We first verify the invertibility for linearized operators and compute guaranteed norm bounds for its inverse by applying the same principle as for the FN-Norm. Next, we confirm the existence of solutions by proving the contractility of the infinite dimensional Newton-like operator with a residual form.

5.1 Infinite dimensional Newton method

Now we consider the following *residual* equation

$$\begin{cases} -\Delta w &= g(x, w, \nabla w), & x \in \Omega, \\ w &= 0, & x \in \partial\Omega, \end{cases} \quad (49)$$

and verify a solution w around 0. Assume that the right-hand side $g(w) := g(\cdot, w, \nabla w)$ is already rewritten as a residual form and g has the same assumptions on f in Section 2. For example, g is given of the form

$$g(w) = f(w + u_h) + \Delta u_h \quad (50)$$

which is same as in (28) or

$$g(w) = f(w + u_h + v_0) - f(u_h) \quad (51)$$

in (32). In the below, an approximate solution of Eq. (49) is denoted by w_h .

Remark 11 We have some alternatives for the choice for w_h . For example, if we want to have the same Newton matrix G in (21), namely, to get $g'(w_h)w = f'(u_h)w$, we can take $w_h = 0$ in (50) or $w_h = -v_0$ in (51).

Denoting the Fréchet derivative of g at w_h by $g'(w_h)$, the Newton-type residual equation for Eq. (49) is written as

$$\begin{cases} -\Delta w - g'(w_h)w &= g(w) - g'(w_h)w, & \text{in } \Omega, \\ w &= 0, & \text{on } \partial\Omega. \end{cases} \quad (52)$$

Here assume that the linear operator

$$\hat{L}w := -\Delta w - g'(w_h)w : H^2(\Omega) \cap H_0^1(\Omega) \longrightarrow L^2(\Omega) \quad (53)$$

has the inverse $\hat{L}^{-1} : L^2(\Omega) \rightarrow H^2(\Omega) \cap H_0^1(\Omega)$ and there exists $\hat{M} > 0$ such that

$$\|\hat{L}^{-1}\phi\|_{H_0^1(\Omega)} \leq \hat{M}\|\phi\|_{L^2(\Omega)}, \quad \forall \phi \in L^2(\Omega). \quad (54)$$

Then the problem (52) is rewritten as the fixed-point form

$$w = \hat{T}w \quad \text{on } H_0^1(\Omega), \quad (55)$$

where $\hat{T}w := \hat{L}^{-1}(g(w) - g'(w_h)w)$. This infinite dimensional Newton-like operator $\hat{T}$ is compact, and it is expected to be a contraction map on some neighborhood of zero.

Therefore, we consider the candidate set of the form

$$W := \left\{ w \in H_0^1(\Omega) \mid \|w\|_{H_0^1(\Omega)} \leq \alpha \right\} \quad (56)$$

for an $\alpha > 0$.

For the existential condition of solutions, by applying the Schauder fixed-point theorem, we need to choose the candidate set W satisfying the criterion

$$\hat{T}W \subset W. \quad (57)$$

From (54), a sufficient condition for the criterion (57) is

$$\sup_{w \in W} \|\hat{T}w\|_{H_0^1(\Omega)} \leq \hat{M} \sup_{w \in W} \|g(w) - g'(w_h)w\|_{L^2(\Omega)} \leq \alpha. \quad (58)$$

Furthermore, for the proof of local uniqueness of the solution within W , the following contraction property is needed on the same set W satisfying (58):

$$\|\hat{T}w_1 - \hat{T}w_2\|_{H_0^1(\Omega)} \leq \hat{k}\|w_1 - w_2\|_{H_0^1(\Omega)}, \quad \forall w_1, w_2 \in W, \quad (59)$$

for some constant $\hat{k} < 1$. In the above case, the Schauder fixed-point theorem can be replaced by the Banach fixed-point theorem, which might yield an advantage if we apply our method to noncompact problems. Note that, by using the norm estimation of $\hat{L}^{-1}$ in (54), it is possible to adopt other fixed-point theorems [71, 108, 109].

5.2 Invertibility condition of $\hat{L}$

By defining

$$Lw := w - (-\Delta)^{-1}g'(w_h)w : H_0^1(\Omega) \longrightarrow H_0^1(\Omega), \quad (60)$$

it is shown that

Proposition 2 If L is invertible, then $\hat{L}$ is also invertible.

Proof. Assume that L has an inverse. Then for every $\phi \in L^2(\Omega)$, there exists $v \in H_0^1(\Omega)$ such that $v = (-\Delta)^{-1}\phi$, and there exists $w \in H_0^1(\Omega)$ such that $Lw = v$, which means

$$w - (-\Delta)^{-1}g'(w_h)w = (-\Delta)^{-1}\phi \quad \Rightarrow \quad w = (-\Delta)^{-1}(g'(w_h)w + \phi).$$

Therefore by the definition of $(-\Delta)^{-1}$, $w \in H^2(\Omega) \cap H_0^1(\Omega)$ and $-\Delta w = g'(w_h)w + \phi$ implies $\hat{L}w = \phi$. $\square$

Hence we will describe a computable condition assuring the invertibility of the linear operator L on $H_0^1(\Omega)$. Basically, this verification method is an extension of the one for solutions of second-order elliptic boundary value problems [60, 61].

Here, we set three positive constants ν_1 , ν_2 and ν_3 satisfying

$$\|P_h(-\Delta)^{-1}g'(w_h)w_*\|_{H_0^1(\Omega)} \leq \nu_1 \|w_*\|_{H_0^1(\Omega)}, \quad \forall w_* \in S_h^\perp, \quad (61)$$

and

$$\|g'(w_h)w\|_{L^2(\Omega)} \leq \nu_2 \|P_h w\|_{H_0^1(\Omega)} + \nu_3 \|(I - P_h)w\|_{H_0^1(\Omega)}, \quad \forall w \in H_0^1(\Omega). \quad (62)$$

An example of upper bounds for these constants ν_1 , ν_2 and ν_3 are as follows [61]. The term $g'(w_h)w$ can be written as

$$g'(w_h)w = (b \cdot \nabla + c)w,$$

where $b \in L^\infty(\Omega)^n$ and $c \in L^\infty(\Omega)$. For $w_* \in S_h^\perp$, since $P_h(-\Delta)^{-1}(b \cdot \nabla w_* + cw_*) \in S_h$ and

$$\begin{aligned} \|\nabla P_h(-\Delta)^{-1}(b \cdot \nabla w_* + cw_*)\|_{L^2(\Omega)}^2 &= (\nabla P_h(-\Delta)^{-1}(b \cdot \nabla w_* + cw_*), \nabla P_h(-\Delta)^{-1}(b \cdot \nabla w_* + cw_*))_{L^2} \\ &= (\nabla(-\Delta)^{-1}(b \cdot \nabla w_* + cw_*), \nabla P_h(-\Delta)^{-1}(b \cdot \nabla w_* + cw_*))_{L^2} \\ &= (b \cdot \nabla w_* + cw_*, P_h(-\Delta)^{-1}(b \cdot \nabla w_* + cw_*))_{L^2} \\ &\leq \|b \cdot \nabla w_* + cw_*\|_{L^2(\Omega)} \|P_h(-\Delta)^{-1}(b \cdot \nabla w_* + cw_*)\|_{L^2(\Omega)} \\ &\leq \|b \cdot \nabla w_* + cw_*\|_{L^2(\Omega)} C_p \|\nabla P_h(-\Delta)^{-1}(b \cdot \nabla w_* + cw_*)\|_{L^2(\Omega)}, \end{aligned}$$

from (9), we have

$$\begin{aligned} \|P_h \nabla(-\Delta)^{-1}(b \cdot \nabla w_* + cw_*)\|_{L^2(\Omega)} &\leq C_p \|b \cdot \nabla w_* + cw_*\|_{L^2(\Omega)} \\ &\leq C_p (\|b \cdot \nabla w_*\|_{L^2(\Omega)} + \|cw_*\|_{L^2(\Omega)}) \\ &\leq C_p (\|b\|_{L^\infty(\Omega)^n} \|\nabla w_*\|_{L^2(\Omega)} + \|c\|_{L^\infty(\Omega)} \|w_*\|_{L^2(\Omega)}) \\ &\leq C_p (\|b\|_{L^\infty(\Omega)^n} + C(h)\|c\|_{L^\infty(\Omega)}) \|\nabla w_*\|_{L^2(\Omega)}, \end{aligned}$$

where $\|b\|_{L^\infty(\Omega)^n}^2 := \sum_{i=1}^n \|b_i\|_{L^\infty(\Omega)}^2$. Therefore we can take

$$\begin{aligned} \nu_1 &= C_p (\|b\|_{L^\infty(\Omega)^n} + C(h)\|c\|_{L^\infty(\Omega)}), \\ \nu_2 &= \|b\|_{L^\infty(\Omega)^n} + C_p \|c\|_{L^\infty(\Omega)}, \\ \nu_3 &= \|b\|_{L^\infty(\Omega)^n} + C(h)\|c\|_{L^\infty(\Omega)}. \end{aligned}$$

Here, $C_p > 0$ denotes the Poincaré or Rayleigh-Ritz constants which satisfies

$$\|u\|_{L^2(\Omega)} \leq C_p \|\nabla u\|_{L^2(\Omega)}, \quad u \in H_0^1(\Omega). \quad (63)$$

For example, if $\Omega = (0, 1) \times (0, 1)$, $C_p = 1/(\pi\sqrt{2})$.

Note that if P_h and $(-\Delta)^{-1}$ are commutative [92], or b is differentiable [61], we can derive more accurate estimates for the above. Recently, Kinoshita, et al. also show it is possible to obtain a similar kind of accurate estimate for ν_i , even though b is not differentiable [13, 64, 65].

5.2.1 Newton-like method

We now give a condition that the problem $Lw = 0$ has only unique solution $w = 0$. By defining $F : H_0^1(\Omega) \rightarrow H_0^1(\Omega)$ by

$$Fw = (-\Delta)^{-1}g'(w_h)w, \quad (64)$$

the problem $Lw = 0$ can be rewritten equivalently in the fixed-point form

$$w = Fw. \quad (65)$$

As described up to now, from the uniqueness of decomposition of $H_0^1(\Omega)$ into S_h and $S_h^\perp$, the fixed-point equation $w = Fw$ on $H_0^1(\Omega)$ is equivalently rewritten as

$$\begin{cases} P_h w &= P_h Fw, \\ (I - P_h)w &= (I - P_h)Fw. \end{cases} \quad (66)$$

Now, as in the Section 2.4, let us define the Newton-like operator $N_h : H_0^1(\Omega) \rightarrow S_h$ by

$$N_h w = P_h w - [I - F]_h^{-1} P_h (I - F)w.$$

Setting the map T on $H_0^1(\Omega)$ defined by

$$Tw = N_h w + (I - P_h)Fw,$$

we find that the two fixed-point problems: $w = Fw$ and $w = Tw$ are equivalent (see Proposition 1).

Next, for positive constants $\hat{\gamma}$ and $\hat{\alpha}$, we define the candidate set $W \subset H_0^1(\Omega)$ by

$$W_h := \left\{ w_h \in S_h \mid \|w_h\|_{H_0^1(\Omega)} \leq \hat{\gamma}, \right\} \subset S_h, \quad W_* := \left\{ w_* \in S_h^\perp \mid \|w_*\|_{H_0^1(\Omega)} \leq \hat{\alpha}, \right\} \subset S_h^\perp,$$

and

$$W := W_h + W_*.$$

Then a sufficient condition for the invertibility result is as follows [25].

Lemma 1 When an inclusion

$$\overline{TW} \subset \overset{\circ}{W} \quad (67)$$

holds, then L is invertible.

Proof. If there exists $w \in H_0^1(\Omega)$ such that $Lw = 0$ and $w \neq 0$, then w also satisfies $w = Tw$. Since T is linear operator, for each $t \in \mathbb{R}$, we have

$$\begin{aligned} T(tw) &= tT(w) \\ &= tw. \end{aligned}$$

Then, we can choose $\hat{t} \in \mathbb{R}$ satisfying

$$\hat{t}w \in \partial W.$$

However, this contradicts with $\overline{TW} \subset \overset{\circ}{W}$ and $T(tw) = tw$. Therefore $w = 0$. That is, $w = 0$ is a unique solution of $Lw = 0$. $\square$

We now describe a procedure to construct the candidate set W of $H_0^1(\Omega)$ which is expected to satisfy the inclusion (67). From the unique decomposability of $w \in W$ into S_h and $S_h^\perp$, we shall check for finite and infinite part separately.

The finite dimensional part of the inclusion, $\overline{N_h W} \subset \overset{\circ}{W}_h$, can be written as

$$\sup_{w \in W} \|N_h w\|_{H_0^1(\Omega)} < \hat{\gamma}.$$

On the other hand, the infinite dimensional part of the inclusion, $\overline{(I - P_h)\hat{F}W} \subset \overset{\circ}{W}_*$, means

$$(I - P_h)(-\Delta)^{-1}g'(w_h)w \in \overset{\circ}{W}_*$$

for each $w \in W$. Therefore, from (8), the condition

$$C(h) \sup_{w \in W} \|g'(w_h)w\|_{L^2(\Omega)} < \hat{\alpha}$$

is sufficient.

From the above facts we can derive the following lemma.

Lemma 2 If one can check the conditions:

$$\sup_{w \in W} \|N_h w\|_{H_0^1(\Omega)} < \hat{\gamma}, \quad C(h) \sup_{w \in W} \|g'(w_h)w\|_{L^2(\Omega)} < \hat{\alpha},$$

then L is invertible.

5.2.2 Computable criterion for the invertibility of L

In order to check the verification condition in Lemma 2, for given positive parameters $\hat{\alpha}$ and $\hat{\gamma}$, we have to compute

$$\gamma := \sup_{\hat{w} \in W} \|N_h \hat{w}\|_{H_0^1(\Omega)}, \quad \alpha := C(h) \sup_{\hat{w} \in W} \|g'(w_h) \hat{w}\|_{L^2(\Omega)},$$

and confirm that

$$\gamma < \hat{\gamma}, \quad \alpha < \hat{\alpha}.$$

For each $w \in W$ such that $w = w_h + w_*$, $w_h \in W_h$, $w_* \in W_*$, we obtain

$$\begin{aligned} N_h w &= P_h w - [I - F]_h^{-1} P_h (I - F) w \\ &= [I - F]_h^{-1} P_h (-\Delta)^{-1} (g'(w_h) w - g'(w_h) w_h) \\ &= [I - F]_h^{-1} P_h (-\Delta)^{-1} g'(w_h) w_* \end{aligned}$$

from the linearity of $g'(w_h)$. Here by setting

$$\begin{aligned} s_h &:= P_h (-\Delta)^{-1} g'(w_h) w_* = \sum_{n=1}^K s_{h,n} \phi_n \in S_h, \quad \mathbf{s} := [s_{h,n}] \in \mathbb{R}^K, \\ N_h w &:= \sum_{n=1}^K t_{h,n} \phi_n \in S_h, \quad \mathbf{t} := [t_{h,n}] \in \mathbb{R}^K, \end{aligned}$$

the definition of $[I - F]_h^{-1}$ implies

$$(\nabla(I - F)N_h w, \nabla v_h)_{L^2} = (\nabla s_h, \nabla v_h)_{L^2}, \quad \forall v_h \in S_h. \quad (68)$$

Equation (68) is equivalent to

$$\sum_{n=1}^K t_{h,n} ((\nabla \phi_n, \nabla \phi_m)_{L^2} - (g'(w_h) \phi_n, \phi_m)_{L^2}) = \sum_{n=1}^K s_{h,n} (\nabla \phi_n, \nabla \phi_m)_{L^2}, \quad 1 \leq m \leq K,$$

Therefore defining

$$[G]_{mn} := (\nabla \phi_n, \nabla \phi_m)_{L^2} - (g'(w_h) \phi_n, \phi_m)_{L^2},$$

we have

$$G\mathbf{t} = A_1 \mathbf{s},$$

where A_1 is defined in (39), and

$$\begin{aligned} \|N_h w\|_{H_0^1(\Omega)} &= \|L_1^T \mathbf{t}\|_E \\ &= \|L_1^T G^{-1} A_1 \mathbf{s}\|_E \\ &\leq \|L_1^T G^{-1} L_1\|_E \|L_1^T \mathbf{s}\|_E \\ &\leq \rho \|s_h\|_{H_0^1(\Omega)}, \end{aligned} \quad (69)$$

where the matrix L_1 and $\rho > 0$ are same as in (40). Using assumption (61), we have

$$\|s_h\|_{H_0^1(\Omega)} \leq \nu_1 \|w_*\|_{H_0^1(\Omega)} \leq \nu_1 \hat{\alpha}, \quad (70)$$

hence from (70) and (69),

$$\|N_h w\|_{H_0^1(\Omega)} \leq \rho \nu_1 \hat{\alpha}, \quad \forall w \in W. \quad (71)$$

Moreover, from the assumption (62), for each $w \in W$ it holds that

$$\begin{aligned} \|g'(w_h) w\|_{L^2(\Omega)} &\leq \nu_2 \|w_h\|_{H_0^1(\Omega)} + \nu_3 \|w_*\|_{H_0^1(\Omega)} \\ &\leq \nu_2 \hat{\gamma} + \nu_3 \hat{\alpha}. \end{aligned} \quad (72)$$

Therefore, the following criterion for verification holds.

Theorem 4 If

$$\kappa := C(h)(\rho\nu_1\nu_2 + \nu_3) < 1 \quad (73)$$

holds, then the operator L has the inverse.

Proof. Since

$$\gamma \leq \rho\nu_1\hat{\alpha}, \quad \alpha \leq C(h)(\nu_2\hat{\gamma} + \nu_3\hat{\alpha}),$$

in order to obtain conditions $\gamma < \hat{\gamma}$ and $\alpha < \hat{\alpha}$, we have to check inequalities

$$\rho\nu_1\hat{\alpha} < \hat{\gamma} \quad \text{and} \quad C(h)(\nu_2\hat{\gamma} + \nu_3\hat{\alpha}) < \hat{\alpha} \quad (74)$$

for some $\hat{\alpha}, \hat{\gamma} > 0$.

Using the assumption (73), we have

$$1 - C(h)(\rho\nu_1\nu_2 + \nu_3) > 0,$$

then for any fixed $\delta > 0$, a positive number $\hat{\alpha}$ can be taken satisfying

$$\hat{\alpha} [1 - C(h)(\rho\nu_1\nu_2 + \nu_3)] > C(h)\nu_2\delta. \quad (75)$$

Therefore, by setting

$$\hat{\gamma} = \rho\nu_1\hat{\alpha} + \delta, \quad (76)$$

the first inequality of (74) is hold, and substituting δ in (76) into (75), we get

$$\hat{\alpha} - \hat{\alpha}C(h) [\rho\nu_1\nu_2 + \nu_3] > C(h)\nu_2(\hat{\gamma} - \rho\nu_1\hat{\alpha}),$$

then

$$\hat{\alpha} > C(h)(\hat{\alpha}\nu_3 + \nu_2\hat{\gamma})$$

which is the second inequality of (74). $\square$

5.3 Upper bound for L^{-1}

In this subsection, under the invertibility criterion for L , we will present an upper bound $M > 0$ satisfying

$$\|L^{-1}\phi\|_{H_0^1(\Omega)} \leq M\|\phi\|_{H_0^1(\Omega)}, \quad \forall \phi \in H_0^1(\Omega). \quad (77)$$

Theorem 5 Under the same assumption as (73), an upper bound $M > 0$ for (77) can be taken as

$$M = \|\mathcal{M}\|_E,$$

where

$$\mathcal{M} = \begin{bmatrix} \rho \left(1 + \frac{\nu_1 C(h) \nu_2 \rho}{1 - \kappa} \right) & \frac{\rho \nu_1}{1 - \kappa} \\ \frac{C(h) \nu_2 \rho}{1 - \kappa} & \frac{1}{1 - \kappa} \end{bmatrix} \in \mathbb{R}^{2 \times 2}. \quad (78)$$

Proof. For each $\phi \in H_0^1(\Omega)$, there exists unique $\psi \in H_0^1(\Omega)$ such that

$$L\psi = (I - F)\psi = \phi. \quad (79)$$

This problem (79) can be regarded as the fixed-point problem

$$\psi = F\psi + \phi \quad (80)$$

for ψ , and using the $H_0^1(\Omega)$ -projection $P_h : H_0^1(\Omega) \rightarrow S_h$, the problem (80) can be decomposed into

$$\begin{cases} P_h \psi &= P_h(F\psi + \phi), \\ (I - P_h)\psi &= (I - P_h)(F\psi + \phi). \end{cases} \quad (81)$$

Therefore by introducing a finite and infinite dimensional Newton-like operators such as

$$\hat{N}_h \psi := P_h \psi - [I - F]_h^{-1} P_h ((I - F) \psi - \phi),$$

and

$$\hat{T} \psi := \hat{N}_h \psi + (I - P_h)(F \psi + \phi),$$

respectively, it is easily seen that the fixed-point problem $\psi = \hat{T} \psi$ is equivalent to Eq. (80).

Now, decomposing ψ and ϕ as follows

$$\psi = \psi_h + \psi_*, \quad \phi = \phi_h + \phi_*, \quad \psi_h, \phi_h \in S_h, \quad \psi_*, \phi_* \in S_h^\perp,$$

we will estimate $\|\psi_h\|_{H_0^1(\Omega)}$ and $\|\psi_*\|_{H_0^1(\Omega)}$ using $\|\phi_h\|_{H_0^1(\Omega)}$ and $\|\phi_*\|_{H_0^1(\Omega)}$.

Since

$$\begin{aligned} \psi_h &= \hat{N}_h \psi \\ &= P_h \psi - [I - F]_h^{-1} P_h ((I - F) \psi - \phi) \\ &= [I - F]_h^{-1} (-P_h F P_h \psi + P_h F \psi + \phi_h) \\ &= [I - F]_h^{-1} (P_h F \psi_* + \phi_h), \end{aligned}$$

using (61), we have

$$\begin{aligned} \|\psi_h\|_{H_0^1(\Omega)} &= \|[I - F]_h^{-1} (P_h F \psi_* + \phi_h)\|_{H_0^1(\Omega)} \\ &\leq \rho \|P_h F \psi_* + \phi_h\|_{H_0^1(\Omega)} \\ &\leq \rho (\|P_h (-\Delta)^{-1} g'(w_h) \psi_*\|_{H_0^1(\Omega)} + \|\phi_h\|_{H_0^1(\Omega)}) \\ &\leq \rho (\nu_1 \|\psi_*\|_{H_0^1(\Omega)} + \|\phi_h\|_{H_0^1(\Omega)}). \end{aligned} \tag{82}$$

Moreover, because of

$$\psi_* = (I - P_h) F (\psi_h + \psi_*) + \phi_*,$$

we obtain, from (8) and (62),

$$\begin{aligned} \|\psi_*\|_{H_0^1(\Omega)} &= \|(I - P_h) F (\psi_h + \psi_*) + \phi_*\|_{H_0^1(\Omega)} \\ &\leq C(h) \|g'(w_h) (\psi_h + \psi_*)\|_{L^2(\Omega)} + \|\phi_*\|_{H_0^1(\Omega)} \\ &\leq C(h) \nu_2 \|\psi_h\|_{H_0^1(\Omega)} + C(h) \nu_3 \|\psi_*\|_{H_0^1(\Omega)} + \|\phi_*\|_{H_0^1(\Omega)}. \end{aligned} \tag{83}$$

By substituting (82) into (83) it follows that

$$\begin{aligned} \|\psi_*\|_{H_0^1(\Omega)} &\leq C(h) \nu_2 \rho \nu_1 \|\psi_*\|_{H_0^1(\Omega)} + C(h) \nu_2 \rho \|\phi_h\|_{H_0^1(\Omega)} + C(h) \nu_3 \|\psi_*\|_{H_0^1(\Omega)} + \|\phi_*\|_{H_0^1(\Omega)} \\ &= C(h) [\nu_2 \rho \nu_1 + \nu_3] \|\psi_*\|_{H_0^1(\Omega)} + C(h) \nu_2 \rho \|\phi_h\|_{H_0^1(\Omega)} + \|\phi_*\|_{H_0^1(\Omega)} \\ &= \kappa \|\psi_*\|_{H_0^1(\Omega)} + C(h) \nu_2 \rho \|\phi_h\|_{H_0^1(\Omega)} + \|\phi_*\|_{H_0^1(\Omega)}. \end{aligned}$$

Then we have

$$\|\psi_*\|_{H_0^1(\Omega)} \leq \frac{C(h) \nu_2 \rho}{1 - \kappa} \|\phi_h\|_{H_0^1(\Omega)} + \frac{1}{1 - \kappa} \|\phi_*\|_{H_0^1(\Omega)}. \tag{84}$$

Also by substituting (84) into (82) it holds that

$$\begin{aligned} \|\psi_h\|_{H_0^1(\Omega)} &\leq \rho \nu_1 \frac{C(h) \nu_2 \rho}{1 - \kappa} \|\phi_h\|_{H_0^1(\Omega)} + \rho \nu_1 \frac{1}{1 - \kappa} \|\phi_*\|_{H_0^1(\Omega)} + \rho \|\phi_h\|_{H_0^1(\Omega)} \\ &= \rho \left(1 + \frac{\nu_1 C(h) \nu_2 \rho}{1 - \kappa} \right) \|\phi_h\|_{H_0^1(\Omega)} + \frac{\rho \nu_1}{1 - \kappa} \|\phi_*\|_{H_0^1(\Omega)}. \end{aligned}$$

Consequently, we have

$$\begin{bmatrix} \|\psi_h\|_{H_0^1(\Omega)} \\ \|\psi_*\|_{H_0^1(\Omega)} \end{bmatrix} \leq \begin{bmatrix} \rho \left(1 + \frac{\nu_1 C(h) \nu_2 \rho}{1 - \kappa} \right) & \frac{\rho \nu_1}{1 - \kappa} \\ \frac{C(h) \nu_2 \rho}{1 - \kappa} & \frac{1}{1 - \kappa} \end{bmatrix} \begin{bmatrix} \|\phi_h\|_{H_0^1(\Omega)} \\ \|\phi_*\|_{H_0^1(\Omega)} \end{bmatrix},$$

which implies the desired result. $\square$

Remark 12 In Theorem 5, if we can take κ , $C(h)$ and ν_1 converging to 0 as h to 0, it implies $\|\mathcal{M}\|_E \rightarrow \max\{\rho, 1\}$, which means that lower bound of M for (77) is *not less than* 1.

By virtue of Theorem 5, an upper bound for $\hat{M} > 0$ of (54) is obtained as follows.

Theorem 6 Under the same assumption (73), an upper bound $\hat{M} > 0$ for (54) can be taken as

$$\hat{M} = C_p M,$$

where $C_p > 0$ is the Poincaré constant same as in (63).

Proof. For each $\phi \in L^2(\Omega)$ by setting $\psi = \hat{L}^{-1}\phi \in H^2(\Omega) \cap H_0^1(\Omega)$, the definition of $\hat{L}$ implies

$$\begin{aligned}\hat{L}\psi &= -\Delta\psi - g'(w_h)\psi \\ &= (-\Delta)L\psi = \phi,\end{aligned}$$

then

$$\psi = L^{-1}(-\Delta)^{-1}\phi.$$

Therefore from (77) and (63) we obtain

$$\begin{aligned}\|\psi\|_{H_0^1(\Omega)} &\leq M\|(-\Delta)^{-1}\phi\|_{H_0^1(\Omega)} \\ &\leq C_p M\|\phi\|_{L^2(\Omega)}.\end{aligned}$$

□

5.4 Direct computation of upper bound for $\hat{L}^{-1}$

In this subsection we will show an alternative estimation for $\hat{M} > 0$ of (54) instead of Theorem 6 under the invertibility criterion for $\hat{L}$.

We define $K \times K$ positive definite matrices A_2 and L_2 by

$$[A_2]_{mn} := (\phi_n, \phi_m)_{L^2}, \quad (85)$$

and the decomposed factor of A_2 such that $A_2 = L_2 L_2^T$, respectively. We also define the positive number $\hat{\rho}$ by

$$\hat{\rho} := \|L_1^T G^{-1} L_2\|_E. \quad (86)$$

Then the following theorem can be shown.

Theorem 7 Under the assumption (73), if

$$\hat{\kappa} := C(h)\nu_3(1 + \hat{\rho}\nu_2) < 1$$

holds, then $\hat{M} > 0$ in (54) can be taken as

$$\hat{M} = \frac{\sqrt{\hat{\rho}^2 + C(h)^2(1 + \nu_2\hat{\rho})^2}}{1 - \hat{\kappa}}. \quad (87)$$

Proof. Since $\hat{L}$ is invertible, for each $\phi \in L^2(\Omega)$, there exists $\psi \in H^2(\Omega) \cap H_0^1(\Omega)$ such that

$$\hat{L}\psi = -\Delta\psi - g'(w_h)\psi = \phi,$$

which implies

$$-\Delta\psi = g'(w_h)\psi + \phi. \quad (88)$$

Because Eq. (88) is equivalent to the fixed-point form

$$\psi = (-\Delta)^{-1}(g'(w_h)\psi + \phi), \quad (89)$$

by using the $H_0^1(\Omega)$ -projection $P_h : H_0^1(\Omega) \rightarrow S_h$, the equation (89) can be decomposed as before, namely,

$$\begin{cases} P_h \psi &= P_h(-\Delta)^{-1}(g'(w_h)\psi + \phi), \\ (I - P_h)\psi &= (I - P_h)(-\Delta)^{-1}(g'(w_h)\psi + \phi). \end{cases} \quad (90)$$

Now, by using the decomposition of ψ such as

$$\psi = \psi_h + \psi_*, \quad \psi_h \in S_h, \quad \psi_* \in S_h^\perp,$$

we aim to estimate $\|\psi_h\|_{H_0^1(\Omega)}$ and $\|\psi_*\|_{H_0^1(\Omega)}$ by $\|\phi\|_{L^2(\Omega)}$. From the uniqueness of decomposition, we have

$$\psi_h = P_h(-\Delta)^{-1}(g'(w_h)\psi + \phi),$$

which implies

$$\begin{aligned} (\nabla \psi_h, \nabla v_h)_{L^2} &= (\nabla(-\Delta)^{-1}(g'(w_h)\psi + \phi), \nabla v_h)_{L^2}, \\ &= (g'(w_h)\psi + \phi, v_h)_{L^2}, \\ &= (g'(w_h)\psi_h + g'(w_h)\psi_* + \phi, v_h)_{L^2}, \quad \forall v_h \in S_h, \end{aligned}$$

namely,

$$(\nabla \psi_h, \nabla v_h)_{L^2} - (g'(w_h)\psi_h, v_h)_{L^2} = (g'(w_h)\psi_* + \phi, v_h)_{L^2}, \quad \forall v_h \in S_h. \quad (91)$$

By setting

$$\psi_h := \sum_{n=1}^K \hat{t}_{h,n} \phi_n \in S_h, \quad \hat{\mathbf{t}} := [\hat{t}_{h,n}] \in \mathbb{R}^K, \quad \hat{\mathbf{f}} := [(g'(w_h)\psi_* + \phi, \phi_n)_{L^2}] \in \mathbb{R}^K,$$

Eq. (91) can be written as

$$G\hat{\mathbf{t}} = \hat{\mathbf{f}}.$$

Now, defining the L^2 -projection $P_0 : L^2(\Omega) \rightarrow S_h$ as

$$(v - P_0 v, v_h)_{L^2} = 0, \quad \forall v_h \in S_h,$$

it is easily seen that

$$\|P_0(g'(w_h)\psi_* + \phi)\|_{L^2(\Omega)} = \|L_2^{-1}\hat{\mathbf{f}}\|_E. \quad (92)$$

Therefore, from (92), (86) and (62) it follows that

$$\begin{aligned} \|\psi_h\|_{H_0^1(\Omega)} &= \|L_1^T \hat{\mathbf{t}}\|_E \\ &= \|L_1^T G^{-1} \hat{\mathbf{f}}\|_E \\ &\leq \|L_1^T G^{-1} L_2\|_E \|L_2^{-1} \hat{\mathbf{f}}\|_E \\ &\leq \hat{\rho} \|P_0(g'(w_h)\psi_* + \phi)\|_{L^2(\Omega)} \\ &\leq \hat{\rho} \|g'(w_h)\psi_* + \phi\|_{L^2(\Omega)} \\ &\leq \hat{\rho} (\|g'(w_h)\psi_*\|_{L^2(\Omega)} + \|\phi\|_{L^2(\Omega)}), \end{aligned}$$

and using (62), we obtain

$$\|\psi_h\|_{H_0^1(\Omega)} \leq \hat{\rho} \nu_3 \|\psi_*\|_{H_0^1(\Omega)} + \hat{\rho} \|\phi\|_{L^2(\Omega)}. \quad (93)$$

Also, from the fact that

$$\psi_* = (I - P_h)(-\Delta)^{-1}(g'(w_h)\psi + \phi),$$

we have, by (8) and (62)

$$\begin{aligned} \|\psi_*\|_{H_0^1(\Omega)} &\leq C(h) \|g'(w_h)(\psi_h + \psi_*) + \phi\|_{L^2(\Omega)} \\ &\leq C(h) (\|g'(w_h)(\psi_h + \psi_*)\|_{L^2(\Omega)} + \|\phi\|_{L^2(\Omega)}) \\ &\leq C(h) \nu_2 \|\psi_h\|_{H_0^1(\Omega)} + C(h) \nu_3 \|\psi_*\|_{H_0^1(\Omega)} + C(h) \|\phi\|_{L^2(\Omega)}. \end{aligned} \quad (94)$$

By substituting (93) into (94), it follows that

$$\begin{aligned}
\|\psi_*\|_{H_0^1(\Omega)} &\leq C(h)\nu_2\hat{\rho}\nu_3\|\psi_*\|_{H_0^1(\Omega)} + C(h)\nu_2\hat{\rho}\|\phi\|_{L^2(\Omega)} + C(h)\nu_3\|\psi_*\|_{H_0^1(\Omega)} + C(h)\|\phi\|_{L^2(\Omega)} \\
&= C(h)[\nu_2\hat{\rho}\nu_3 + \nu_3]\|\psi_*\|_{H_0^1(\Omega)} + C(h)\nu_2\hat{\rho}\|\phi\|_{L^2(\Omega)} + C(h)\|\phi\|_{L^2(\Omega)} \\
&= \hat{\kappa}\|\psi_*\|_{H_0^1(\Omega)} + C(h)(\nu_2\hat{\rho} + 1)\|\phi\|_{L^2(\Omega)},
\end{aligned}$$

thus

$$\|\psi_*\|_{H_0^1(\Omega)} \leq C(h)\frac{1 + \nu_2\hat{\rho}}{1 - \hat{\kappa}}\|\phi\|_{L^2(\Omega)}. \quad (95)$$

Also by substituting (95) into (93), we have

$$\begin{aligned}
\|\psi_h\|_{H_0^1(\Omega)} &\leq C(h)\frac{(1 + \nu_2\hat{\rho})\hat{\rho}\nu_3}{1 - \hat{\kappa}}\|\phi\|_{L^2(\Omega)} + \hat{\rho}\|\phi\|_{L^2(\Omega)} \\
&= \hat{\rho}\left(1 + C(h)\frac{(1 + \nu_2\hat{\rho})\nu_3}{1 - \hat{\kappa}}\right)\|\phi\|_{L^2(\Omega)} \\
&= \hat{\rho}\left(1 + \frac{\hat{\kappa}}{1 - \hat{\kappa}}\right)\|\phi\|_{L^2(\Omega)} \\
&= \frac{\hat{\rho}}{1 - \hat{\kappa}}\|\phi\|_{L^2(\Omega)}.
\end{aligned}$$

Consequently, we have completed the proof by

$$\begin{aligned}
\|\psi\|_{H_0^1(\Omega)}^2 &= \|\psi_h\|_{H_0^1(\Omega)}^2 + \|\psi_*\|_{H_0^1(\Omega)}^2 \\
&\leq \left(\frac{\hat{\rho}}{1 - \hat{\kappa}}\right)^2 \|\phi\|_{L^2(\Omega)}^2 + \left(C(h)\frac{1 + \nu_2\hat{\rho}}{1 - \hat{\kappa}}\right)^2 \|\phi\|_{L^2(\Omega)}^2 \\
&= \frac{1}{(1 - \hat{\kappa})^2} (\hat{\rho}^2 + C(h)^2(1 + \nu_2\hat{\rho})^2) \|\phi\|_{L^2(\Omega)}^2 \\
&= \hat{M}^2 \|\phi\|_{L^2(\Omega)}^2.
\end{aligned}$$

□

Remark 13 If we can take $\hat{\kappa} \rightarrow 0$ and $C(h) \rightarrow 0$ as $h \rightarrow 0$ in (87), we have $\hat{M} \rightarrow \hat{\rho}$. Comparing with Theorem 6 ($C_p M$), the estimation (87) has an advantage because it has no restricted lower bound 1 (cf. Remark 12). However, in order to obtain $\hat{\rho}$ of (86), the additional computation is needed, and the difference in these two kinds of estimates for $\hat{M}$ possibly depends on given problems and approximation subspace S_h , especially the size of $h > 0$.

5.5 A verification example of IN-Linz

Consider the following Emden equation on the unit square Ω :

$$\begin{cases} -\Delta u &= u^2 & \text{in } \Omega, \\ u &= 0 & \text{on } \partial\Omega. \end{cases} \quad (96)$$

We define the finite dimensional space S_h as the tensor product space of $\mathbb{P}_N$, where $\mathbb{P}_N$ stands for the set of homogeneous Legendre polynomials on $(0, 1)$ such that

$$\frac{(-1)^n \sqrt{2n-1}}{(n-1)!} \left(\frac{d}{dx}\right)^{n-2} (1-x)^{n-1} x^{n-1}, \quad n \geq 2.$$

We can take the constant $C(h)$ of (8) as $1/\sqrt{2(2N+1)(2N+5)}$ for $N \geq 4$ [66].

The approximate solution $u_h \in S_h$ satisfying

$$(\nabla u_h, \nabla v_h)_{L^2} = (u_h^2, v_h)_{L^2}, \quad \forall v_h \in S_h \quad (97)$$

is obtained by the usual Newton method with some appropriate initial value. Because $\Delta u_h \in L^2(\Omega)$, we can use the direct residual formulation (28), then it is sufficient to compute a solution of the above

nonlinear equation (97) by the usual floating point arithmetic. Moreover, the Galerkin approximation with higher degree polynomials enables us to get very small residual errors.

Table II shows the verification results by the infinite dimensional Newton method described in Section 5.1 using bi- N degree polynomials. In the table, “—” means no calculation was done due to the failure of the invertibility condition (73). “Existence” and “Uniqueness” mean $H_0^1(\Omega)$ norm sense. The L^∞ error $\|u - u_h\|_{L^\infty}$ is a posteriori error by using the explicit Sobolev inequality by Plum [72] after enclosing a solution.

Table II. Verification results for Emden equation.

N	$\ u_h\ _{L^2}$	$\ u_h\ _{L^\infty}$	$\ u_h\ _{H_0^1}$	$\ \Delta u_h + u_h^2\ _{L^2}$	
8	13.0736	29.2334	58.8264	6.49785E+00	
16	13.0734	29.2571	58.8259	7.61524E-03	
24	13.0734	29.2571	58.8259	4.17816E-06	
N	M	κ	Existence	Uniqueness	$\ u - u_h\ _{L^\infty}$
8	2.74009	2.82214	—	—	—
16	2.74682	0.82836	7.43965E-02	0.154966	1.39722E+00
24	2.74719	0.38951	6.86979E-06	0.608198	1.32133E-04
N	$\tilde{M}$	$\hat{\kappa}$	Existence	Uniqueness	$\ u - u_h\ _{L^\infty}$
8	0.37933	—	—	—	—
16	0.37993	0.38705	5.54067E-03	1.37835	1.10836E+00
24	0.37994	0.18198	2.10263E-06	1.98713	4.33409E-05

In the verification step, we use interval arithmetic library INTLIB [11] with Intel Fortran 10.1 compiler on Intel Core2 Quad Q6700 (OS: Ubuntu Linux 7.10).

Acknowledgments

This work was supported by a Grant-in-Aid from the Ministry of Education, Culture, Sports, Science and Technology of Japan (No.20224001, No.21540134).

References

- [1] R.A. Adams, *Sobolev Spaces*, Academic Press, New York, 1975.
- [2] G. Alefeld and J. Herzberger, *Introduction to Interval Computations*, Academic Press, New York, 1993.
- [3] G. Alefeld and G. Mayer, “The Cholesky method for interval data,” *Linear Algebra and its Applications*, vol. 194, pp. 161–182, November 1993.
- [4] P.G. Ciarlet, *The Finite Element Method for Elliptic Problems*, North-Holland, Amsterdam, 1978.
- [5] P. Grisvard, *Elliptic Problems in Nonsmooth Domains*, Pitman, Boston, 1985.
- [6] K. Hashimoto, R. Abe, M.T. Nakao, and Y. Watanabe, “A numerical verification method for solutions of singularly perturbed problems with nonlinearity,” *Japan Journal of Industrial and Applied Mathematics*, vol. 22, no. 1, pp. 111–131, February 2005.
- [7] K. Hashimoto, K. Kobayashi, and M.T. Nakao, “Numerical verification methods for solutions of the free boundary problem,” *Numerical Functional Analysis and Optimization*, vol. 26, no. 4-5, pp. 523–542, June 2005.
- [8] K. Hashimoto, K. Nagatou, and M.T. Nakao, “A computational approach to constructive a priori error estimate for finite element approximations of bi-harmonic problems in nonconvex polygonal domains,” *Information*, vol. 9, no. 4, pp. 573–580, July 2006.
- [9] J.G. Heywood, W. Nagata, and W. Xie, “A numerically based existence theorem for the Navier-Stokes equations,” *Journal of Mathematical Fluid Mechanics*, vol. 1, no. 1, pp. 5–23, April 1999.

- [10] M. Kim, M.T. Nakao, Y. Watanabe, and T. Nishida, "A numerical verification method of bifurcating solutions for 3-dimensional Rayleigh-Bénard problems," *Numerische Mathematik*, vol. 111, no. 3, pp. 389–406, January 2009.
- [11] R.B. Kearfott, "Algorithm 763: INTERVAL_ARITHMETIC: A Fortran 90 module for an interval data type," *ACM Transactions on Mathematical Software*, vol. 22, no. 4, pp. 385–392, December 1996.
- [12] F. Kikuchi and L. Xuefeng, "Determination of the Babuska-Aziz constant for the linear triangular finite element," *Japan Journal of Industrial and Applied Mathematics*, vol. 23, no. 1, pp. 75–82, February 2006.
- [13] T. Kinoshita, K. Hashimoto, and M.T. Nakao, "On the L^2 a priori error estimates to the finite element solution of elliptic problems with singular adjoint operator," *Numerical Functional Analysis and Optimization*, vol. 30, no. 3-4, pp. 289–305, March 2009.
- [14] K. Kobayashi, "A constructive a priori error estimation for finite element discretizations in a non-convex domain using singular functions," *Japan Journal of Industrial and Applied Mathematics*, vol. 26, no. 2, pp. 493–516, October 2009.
- [15] T. Komori and N. Yamamoto, "A numerical verification method for initial value problems of ODEs using Nakao's method for boundary value problems," (in Japanese), *Transaction of the Japan Society for Industrial and Applied Mathematics*, vol. 18, no. 3, pp. 303–319, September 2008.
- [16] T. Minamoto, and M.T. Nakao, "Numerical verifications of solutions for nonlinear parabolic equations in one-space dimensional case," *Reliable Computing*, vol. 3, no. 2, pp. 137–147, May 1997.
- [17] T. Minamoto, "Numerical verification of solutions for nonlinear hyperbolic equations," *Applied Mathematics Letters*, vol. 10, no. 6, pp. 91–96, November 1997.
- [18] T. Minamoto, N. Yamamoto, and M.T. Nakao, "Numerical verification method for solution of the perturbed Gelfand equation," *Methods and Applications of Analysis*, vol. 7, no. 1, pp. 251–262, March 2000.
- [19] T. Minamoto, "Numerical existence and uniqueness proof for solutions of semilinear parabolic equations," *Applied Mathematics Letters*, vol. 14, no. 6, pp. 707–714, August 2001.
- [20] T. Minamoto, "Numerical verification method for solutions of nonlinear hyperbolic equations" in *Symbolic Algebraic Methods and Verification Methods*, ed. G. Alefeld, J. Rohn, S. Rump, and T. Yamamoto, pp. 173–181, Springer, Wien NewYork, 2001.
- [21] T. Minamoto, "Numerical existence and uniqueness proof for solutions of nonlinear hyperbolic equations," *Journal of Computational and Applied Mathematics*, vol. 135, no. 1, pp. 79–90, October 2001.
- [22] T. Minamoto, "Numerical method with guaranteed accuracy of a double turning point for a radially symmetric solution of the perturbed Gelfand equation," *Journal of Computational and Applied Mathematics*, vol. 169, no. 1, pp. 151–160, August 2004.
- [23] T. Minamoto and M.T. Nakao, "Numerical method for verifying the existence and local uniqueness of a double turning point for a radially symmetric solution of the perturbed Gelfand equation," *Journal of Computational and Applied Mathematics*, vol. 202, no. 2, pp. 177–185, May 2007.
- [24] R.E. Moore, R.B. Kearfott, and M.J. Cloud, *Introduction to Interval Analysis*, Society for Industrial and Applied Mathematics, 2009.
- [25] K. Nagatou, N. Yamamoto, and M.T. Nakao, "An approach to the numerical verification of solutions for nonlinear elliptic problems with local uniqueness," *Numerical Functional Analysis and Optimization*, vol. 20, no. 5-6, pp. 543–565, April 1999.
- [26] K. Nagatou, "A numerical method to verify the elliptic eigenvalue problems including a uniqueness property," *Computing*, vol. 63, no. 2, pp. 109–130, 1999.
- [27] K. Nagatou and M.T. Nakao, "An enclosure method of eigenvalues for the elliptic operator linearized at an exact solution of nonlinear problems," *Linear Algebra and its Applications*, vol. 324, no. 1-3, pp. 81–106, February 2001.

- [28] K. Nagatou, “An enclosure method for real eigenvalues of non-selfadjoint elliptic operators,” *Proceedings of the Fifth China-Japan Joint Seminar on Numerical Mathematics*, pp. 57–64, 2002.
- [29] K. Nagatou, M.T. Nakao, and M. Wakayama, “Verified numerical computations for eigenvalues of non-commutative harmonic oscillators,” *Numerical Functional Analysis and Optimization*, vol. 23, no. 5-6, pp. 633–650, 2002.
- [30] K. Nagatou, “A computer-assisted proof on the stability of the Kolmogorov flows of incompressible viscous fluid,” *Journal of Computational and Applied Mathematics*, vol. 169, no. 1, pp. 33–44, August 2004.
- [31] K. Nagatou, K. Hashimoto, and M.T. Nakao, “Numerical verification of stationary solutions for Navier-Stokes problems,” *Journal of Computational and Applied Mathematics*, vol. 199, no. 2, pp. 445–451, February 2007.
- [32] K. Nagatou, “Validated computation for infinite dimensional eigenvalue problems,” *12th GAMM - IMACS International Symposium on Scientific Computing, Computer Arithmetic and Validated Numerics (SCAN 2006)*, Order Number E2821, p. 3, 2007.
- [33] K. Nagatou, “Validated computations for fundamental solutions of linear ordinary differential equation” in *Inequalities and Applications*, International Series of Numerical Mathematics, vol. 157, eds. C. Bandle, A. Gilányi, L. Losonczy, Z. Páles, and M. Plum, 2009.
- [34] K. Nagatou, “Numerical verification method for infinite dimensional eigenvalue problems,” *Japan Journal of Industrial and Applied Mathematics*, vol. 26, no. 2, pp. 477–491, October 2009.
- [35] M.T. Nakao, “A numerical approach to the proof of existence of solutions for elliptic problems,” *Japan Journal of Applied Mathematics*, vol. 5, no. 2, pp. 313–332, June 1988.
- [36] M.T. Nakao, “A computational verification method of existence of solutions for nonlinear elliptic equations” in *Recent topics in nonlinear PDE IV*, Lecture Notes in Numerical and Applied Analysis, vol. 10, eds. M. Mimura and T. Nishida, pp. 101–120, Kinokuniya, Tokyo, 1989.
- [37] M.T. Nakao, “A numerical approach to the proof of existence of solutions for elliptic problems II,” *Japan Journal of Applied Mathematics*, vol. 7, no. 3, pp. 477–488, October 1990.
- [38] M.T. Nakao and N. Yamamoto, “Numerical verifications of solutions for elliptic equations with strong nonlinearity,” *Numerical Functional Analysis and Optimization*, vol. 12, no. 5-6, pp. 535–543, 1991.
- [39] M.T. Nakao, “Solving nonlinear parabolic problems with result verification. Part I: one-space dimensional case,” *Journal of Computational and Applied Mathematics*, vol. 38, no. 1-3, pp. 323–334, December 1991.
- [40] M.T. Nakao, “A numerical verification method for the existence of weak solutions for nonlinear boundary value problems,” *Journal of Mathematical Analysis and Applications*, vol. 164, no. 2, pp. 489–507, May 1992.
- [41] M.T. Nakao, “Computable error estimates for FEM and numerical verification of solutions for nonlinear PDEs” in *Computational and Applied Mathematics I, Algorithms and Theory*, eds. C. Brezinski and U. Kulisch, pp. 357–366, North-Holland, 1992.
- [42] M.T. Nakao, “Solving nonlinear elliptic problems with result verification using an H^{-1} type residual iteration” in *Validation Numerics: theory and applications*, eds. R. Albrecht, G. Alefeld, and H.J. Stetter, Computing Supplementum, vol. 9, pp. 161–173, Springer-Verlag, Wien, 1993.
- [43] M.T. Nakao, “Computable L^∞ error estimates in the finite element method with application to nonlinear elliptic problems” in *Contributions in Numerical Mathematics*, ed. R.P. Agarwal, pp. 309–319, World Scientific, Singapore, 1993.
- [44] M.T. Nakao and Y. Watanabe, “On computational proofs of the existence of solutions to nonlinear parabolic problems,” *Journal of Computational and Applied Mathematics*, vol. 50, no. 1-3, pp. 401–410, May 1994.

- [45] M.T. Nakao, “Numerical verifications of solutions for nonlinear hyperbolic equations,” *Interval Computations*, vol. 4, no. 4, pp. 64–77, 1994.
- [46] M.T. Nakao and N. Yamamoto, “Numerical verification of solutions for nonlinear elliptic problems using L^∞ residual method,” *Journal of Mathematical Analysis and Applications*, vol. 217, no. 1, pp. 246–262, January 1998.
- [47] M.T. Nakao, N. Yamamoto, and Y. Watanabe, “A posteriori and constructive a priori error bounds for finite element solutions of the Stokes equations,” *Journal of Computational and Applied Mathematics*, vol. 91, no. 1, pp. 137–158, April 1998.
- [48] M.T. Nakao, N. Yamamoto, and Y. Watanabe, “Constructive L^2 error estimates for finite element solutions of the Stokes equations,” *Reliable Computing*, vol. 4, no. 2, pp. 115–124, May 1998.
- [49] M.T. Nakao, N. Yamamoto, and Y. Nishimura, “Numerical verification of the solution curve for some parametrized nonlinear elliptic problem,” *Proceedings of the Third China-Japan Seminar on Numerical Mathematics*, pp. 238–245, Science Press, Beijing, 1998.
- [50] M.T. Nakao, N. Yamamoto, and S. Kimura, “On best constant in the error bound for the H_0^1 -projection into piecewise polynomial spaces,” *Journal of Approximation Theory*, vol. 93, no. 3, pp. 491–500, June 1998.
- [51] M.T. Nakao and C.S. Ryoo, “Numerical verifications of solutions for variational inequalities using Newton-like method,” *Information*, vol. 2, no. 1, pp. 27–35, January 1999.
- [52] M.T. Nakao, N. Yamamoto, and K. Nagatou, “Numerical verifications for eigenvalues of second-order elliptic operators,” *Japan Journal of Industrial and Applied Mathematics*, vol. 16, no. 3, pp. 307–320, December 1999.
- [53] M.T. Nakao, S.H. Lee, and C.S. Ryoo, “Numerical verification of solutions for elasto-plastic torsion problems,” *Computers and Mathematics with Applications*, vol. 39, no. 3-4, pp. 195–204, February 2000.
- [54] M.T. Nakao, “Numerical verification methods for solutions of ordinary and partial differential equations,” *Numerical Functional Analysis and Optimization*, vol. 22, no. 3 & 4, pp. 321–356, June 2001.
- [55] M.T. Nakao, N. Yamamoto, and Y. Watanabe, “Verified numerical computations for an inverse elliptic eigenvalue problem with finite data,” *Japan Journal of Industrial and Applied Mathematics*, vol. 18, no. 2, pp. 587–602, June 2001.
- [56] M.T. Nakao, N. Yamamoto, “A guaranteed bound of the optimal constant in the error estimates for linear triangular element” in *Topics in Numerical Analysis With Special Emphasis on Nonlinear Problems*, Computing Supplement, vol. 15, eds. G. Alfeld and X. Chen, pp. 163–173, Springer, Wien, New York, 2001.
- [57] M.T. Nakao and N. Yamamoto, “A guaranteed bound of the optimal constant in the error estimates for linear triangular element Part II: Details” in *Perspectives on Enclosure Methods*, eds. U. Kulisch, R. Lohner, and A. Facius, pp. 265–276, Springer, Wien, New York, 2001.
- [58] M.T. Nakao and C.S. Ryoo, “Numerical verification methods for solutions of free boundary problems,” *Mathematical Modeling and Numerical Simulation in Continuum Mechanics*, Lecture Notes in Computational Science and Engineering, vol. 19, eds. I. Babuska, P.G. Ciarlet, and T. Miyoshi, pp. 195–208, Springer Verlag, 2002.
- [59] M.T. Nakao, Y. Watanabe, N. Yamamoto, and T. Nishida, “Some computer assisted proofs for solutions of the heat convection problems,” *Reliable Computing*, vol. 9, no. 5, pp. 359–372, October 2003.
- [60] M.T. Nakao, and Y. Watanabe, “An efficient approach to the numerical verification for solutions of elliptic differential equations,” *Numerical Algorithms*, vol. 37, no. 1-4, pp. 311–323, December 2004.
- [61] 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, July 2005.

- [62] M.T. Nakao, K. Hashimoto, and K. Kobayashi, “Verified numerical computation of solutions for the stationary Navier-Stokes equation in nonconvex polygonal domains,” *Hokkaido Mathematical Journal*, vol. 36, no. 4, pp. 777–799, November 2007.
- [63] M.T. Nakao, K. Hashimoto, and K. Nagatou, “A computational approach to constructive a priori and a posteriori error estimates for finite element approximations of bi-harmonic problems” in *Proceedings of the 4th JSIAM-SIMAI Seminar on Industrial and Applied Mathematics*, pp. 139–148, 2008.
- [64] M.T. Nakao and K. Hashimoto, “Guaranteed error bounds for finite element approximations of noncoercive elliptic problems and their applications,” *Journal of Computational and Applied Mathematics*, vol. 218, no. 1, pp. 106–115, August 2008.
- [65] M.T. Nakao and T. Kinoshita, “Some remarks on the behaviour of the finite element solution in nonsmooth domains,” *Applied Mathematics Letters*, vol. 21, no. 12, pp. 1310–1314, December 2008.
- [66] M.T. Nakao and T. Kinoshita, “On very accurate verification of solutions for boundary value problems by using spectral methods,” *JSIAM Letters*, vol. 1, pp. 21–24, 2009.
- [67] M.T. Nakao and K. Hashimoto, “A numerical verification method for solutions of nonlinear parabolic problems,” *Journal of Math-for-Industry*, vol. 1, 2009A-9, pp. 69–72, 2009.
- [68] M.T. Nakao, Y. Watanabe, N. Yamamoto, T. Nishida, and M. Kim, “Computer assisted proofs of bifurcating solutions for nonlinear heat convection problems,” *Journal of Scientific Computing*, vol. 43, no. 3, pp. 388–401, June 2010.
- [69] A. Neumaier, *Interval Methods for Systems of Equations*, Cambridge University Press, Cambridge, 1990.
- [70] S. Oishi, “Numerical verification of existence and inclusion of solutions for nonlinear operator equations,” *Journal of Computational and Applied Mathematics*, vol. 60, no. 1-2, pp. 171–185, June 1995.
- [71] J.M. Ortega, “The Newton-Kantorovich theorem,” *The American Mathematical Monthly*, vol. 75, no. 6, pp. 658–660, June–July 1968.
- [72] M. Plum, “Explicit H_2 -estimates and pointwise bounds for solutions of second-order elliptic boundary value problems,” *Journal of Mathematical Analysis and Applications*, vol. 165, no. 1, pp. 36–61, March 1992.
- [73] M. Plum, “Existence and enclosure results for continua of solutions of parameter-dependent nonlinear boundary value problems,” *Journal of Computational and Applied Mathematics*, vol. 60, no. 1-2, pp. 187–200, June 1995.
- [74] M. Plum, “Computer-assisted enclosure methods for elliptic differential equations,” *Linear Algebra and its Applications*, vol. 324, no. 1-3, pp. 147–187, February 2001.
- [75] M. Plum, “Computer-assisted proofs for semilinear elliptic boundary value problems,” *Japan Journal of Industrial and Applied Mathematics*, vol. 26, no. 2-3, pp. 419–442, October 2009.
- [76] S.M. Rump, “On the solution of interval linear systems,” *Computing*, vol. 47, no. 3-4, pp. 337–353, September 1992.
- [77] S.M. Rump, “Verification methods for dense and sparse systems of equations” in *Topics in Validated Computations*, ed. J. Herzberger, pp. 63–135, Elsevier Science, North Holland, Amsterdam, 1994.
- [78] S.M. Rump, “A note on epsilon-inflation,” *Reliable Computing*, vol. 4, no. 4, pp. 371–375, November 1998.
- [79] S.M. Rump, “INTLAB – INTerval LABoratory,” in *Developments in Reliable Computing*, ed. T. Csendes, pp. 77–104, Kluwer Academic Publishers, Dordrecht, 1999, <http://www.ti3.tu-harburg.de/rump/>.
- [80] C.S. Ryoo and M.T. Nakao, “Numerical verification of solutions for variational inequalities,” *Numerische Mathematik*, vol. 81, no. 2, pp. 305–320, December 1998.
- [81] C.S. Ryoo, “Numerical verification of solutions for a simplified Signorini problem,” *Computers and Mathematics with Applications*, vol. 40, no. 8-9, pp. 1003–1013, October–November 2000.

- [82] C.S. Ryoo and M.T. Nakao, “Numerical verification of solutions for variational inequalities of the second kind,” *Computers and Mathematics with Applications*, vol. 43, no. 10-11, pp. 1371–1380, June 2002.
- [83] C.S. Ryoo, H. Song, and S.D. Kim, “Numerical verification of solutions for some unilateral boundary value problems,” *Computers and Mathematics with Applications*, vol. 44, no. 5-6, pp. 787–797, September 2002.
- [84] C.S. Ryoo and M.T. Nakao, “Numerical verification of solutions for obstacle problems,” *Journal of Computational and Applied Mathematics*, vol. 161, no. 2, pp. 405–416, December 2003.
- [85] C.S. Ryoo, “Numerical verification of solutions for Signorini problems using Newton-like method,” *International Journal for Numerical Methods in Engineering*, vol. 73, no. 8, pp. 1181–1196, February 2008.
- [86] Sun Microsystems, “Fortran 95 Interval Arithmetic Programming Reference,” <http://docs.sun.com/source/816-2462/index.html>.
- [87] K. Toyonaga, M.T. Nakao, and Y. Watanabe, “Verified numerical computations for multiple and nearly multiple eigenvalues of elliptic operators,” *Journal of Computational and Applied Mathematics*, vol. 147, no. 1, pp. 175–190, October 2002.
- [88] T. Tsuchiya and M.T. Nakao, “Numerical verification of solutions of parametrized nonlinear boundary value problems with turning points,” *Japan Journal of Industrial and Applied Mathematics*, vol. 14, no. 3, pp. 357–372, October 1997.
- [89] Y. Watanabe and M.T. Nakao, “Numerical verifications of solutions for nonlinear elliptic equations,” *Japan Journal of Industrial and Applied Mathematics*, vol. 10, no. 1, pp. 165–178, February 1993.
- [90] Y. Watanabe, N. Yamamoto, and M.T. Nakao, “Verified computations of solutions for non-differentiable elliptic equations related to MHD equilibria,” *Nonlinear Analysis, Theory, Methods & Applications*, vol. 28, no. 3, pp. 577–587, February 1997.
- [91] Y. Watanabe, N. Yamamoto, and M.T. Nakao, “A numerical verification for the Navier-Stokes equations,” *Reliable Computing*, vol. 5, no. 3, pp. 347–357, August 1999.
- [92] Y. Watanabe, N. Yamamoto, M.T. Nakao, and T. Nishida, “A numerical verification of non-trivial solutions for the heat convection problem,” *Journal of Mathematical Fluid Mechanics*, vol. 6, no. 1, pp. 1–20, March 2004.
- [93] Y. Watanabe, N. Yamamoto, and M.T. Nakao, “Verification methods of generalized eigenvalue problems and its applications,” (in Japanese), *Transaction of the Japan Society for Industrial and Applied Mathematics*, vol. 9, no. 3, pp. 137–150, September 1999.
- [94] Y. Watanabe, N. Yamamoto, and M.T. Nakao, “An efficient approach to the numerical verification for solutions of elliptic differential equations with local uniqueness,” (in Japanese), *Transaction of the Japan Society for Industrial and Applied Mathematics*, vol. 15, no. 4, pp. 509–520, December 2005.
- [95] Y. Watanabe and M.T. Nakao, “Numerical verification method of solutions for elliptic equations and its application to the Rayleigh-Bénard problem,” *Japan Journal of Industrial and Applied Mathematics*, vol. 26, no. 2-3, pp. 443–463, October 2009.
- [96] Y. Watanabe, M. Plum, and M.T. Nakao, “A computer-assisted instability proof for the Orr-Sommerfeld problem with Poiseuille flow,” *Zeitschrift für Angewandte Mathematik und Mechanik*, vol. 89, no. 1, pp. 5–18, January 2009.
- [97] Y. Watanabe, “A computer-assisted proof for the Kolmogorov flows of incompressible viscous fluid,” *Journal of Computational and Applied Mathematics*, vol. 223, no. 2, pp. 953–966, January 2009.
- [98] Y. Watanabe, “A numerical verification method for two-coupled elliptic partial differential equations,” *Japan Journal of Industrial and Applied Mathematics*, vol. 26, no. 2-3, pp. 233–247, October 2009.
- [99] S. Yamamoto, “Validated computations for a system of reaction diffusion equations of FitzHugh-Nagumo type with Neumann boundary,” (in Japanese), Master Thesis of Kyushu University, 2008.

- [100] N. Yamamoto and M.T. Nakao, “Numerical verifications of solutions for elliptic equations in nonconvex polygonal domains,” *Numerische Mathematik*, vol. 65, no. 1, pp. 503–521, December 1993.
- [101] 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, June 1995.
- [102] N. Yamamoto, “A numerical verification method for solutions of boundary value problems with local uniqueness by Banach’s fixed point theorem,” *SIAM Journal on Numerical Analysis*, vol. 35, no. 5, pp. 2004–2013, October 1998.
- [103] N. Yamamoto, M.T. Nakao, and Y. Watanabe, “Validated computation for a linear elliptic problem with a parameter” in *Advances in Numerical Mathematics*, pp. 155–162, Gakkotosho, Tokyo, 1999.
- [104] N. Yamamoto, “A simple method for error bounds of eigenvalues of symmetric matrices,” *Linear Algebra and its Applications*, vol. 324, no. 1-3, pp. 227–234, February 2001.
- [105] N. Yamamoto and K. Hayakawa, “Error estimation with guaranteed accuracy of finite element method in nonconvex polygonal domains,” *Journal of Computational and Applied Mathematics*, vol. 159, no. 1, pp. 173–183, October 2003.
- [106] N. Yamamoto and K. Gemma, “On error estimation of finite element approximations to the elliptic equations in nonconvex polygonal domains,” *Journal of Computational and Applied Mathematics*, vol. 199, no. 2, pp. 286–296, February 2007.
- [107] N. Yamamoto and T. Komori, “An application of Taylor models to the Nakao method on ODEs,” *Japan Journal of Industrial and Applied Mathematics*, vol. 26, no. 2, pp. 365–392, October 2009.
- [108] N. Yamamoto, M.T. Nakao, and Y. Watanabe, “A theorem for numerical verification on local uniqueness of solutions to fixed-point equations,” in preparation.
- [109] E. Zeidler (translated by P.R. Wadsack), *Nonlinear Functional Analysis and Its Applications I: Fixed-Point Theorems*, Springer-Verlag, New-York, 1986.