Hypercircle-based local error estimation for the finite element solution of partial differential equations

Taiga NAKANO

Doctoral Program in Fundamental Sciences
Graduate School of Science and Technology
Niigata University

February 2023

ACKNOWLEDGEMENT

This dissertation is the result of three years of study in the Doctoral Program in the Graduate School of Science and Technology, Niigata University. I would like to express my deep gratitude to my advisor, Professor Xuefeng LIU, for his insightful advice and encouragement. I also express my sincere thanks to NAPSON Co., Ltd. for providing the support for this study. In addition, members of the laboratory participated in this study and active discussions were held. I am deeply grateful to them. Finally, I would like to express my gratitude to my family for their continuous support and cooperation in the preparation of this dissertation. This study was supported by JST SPRING, Grant Number JPMJSP2121, and NAPSON Co., Ltd.

February 2023.

Taiga NAKANO

ABSTRACT

High-precision numerical methods are required to solve differential equations that govern physical models for measurement problems appearing in the semiconductor industry, in order to provide reliable and accurate measurement results. In the field of numerical analysis, recent studies are concerning on methods that provide guaranteed error estimation for numerical results obtained by finite element methods (FEM). In particular, Kikuchi and Liu have proposed the hypercircle based *a posteriori* and *a priori* error estimation. As required by the four-probe method for resistivity measurement, the local error estimation for FEM solutions of Poisson equations plays an important role in improving the precision of the measurement results. This study aims to extend the hypercircle method to estimate local errors for boundary value problems of the Poisson equation. Meanwhile, the application of the hypercircle method to the error estimation theory for the non-homogeneous Neumann boundary value problems of the modified Helmholtz equation is considered.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Background

The motivation of this research originates from the error analysis of the four-probe method, which has been used for the resistivity measurement of semiconductors over the past century [18]. The image of the four-probe method is illustrated in Figure 1.1: four probes A, B, C, and D are aligned on the surface of the sample; a constant current I_{AD} is applied between A and D and the potential difference V_{BC} between B and C is measured. The resistivity ρ is then calculated by $\rho = F_c V_{BC}/I_{AD}$, where F_c is the correction factor. As an important quantity for high-precision measurement, F_c is evaluated theoretically by considering the governing equation of the distribution of the potential. A well-used model for the potential distribution u is described by the following boundary value problem of Poisson's equation (see, e.g., [18, 49]):

$$-\Delta u = 2\rho\, I_{AD}\, (\delta(A;x) - \delta(D;x)) \text{ in } \Omega; \quad \frac{\partial u}{\partial \mathbf{n}} = 0 \text{ on } \partial\Omega\,, \tag{1.1}$$

where $\delta(A;x)$ and $\delta(D;x)$ are Dirac's delta functions located at A and D, respectively. Note that this model regards the current I_{AD} as a point charge on the surface of the sample. By setting $\rho\, I_{AD} = 1$, the value of F_c can be evaluated by

$$F_c = \frac{1}{u(B) - u(C)}\,.$$

The calculation of F_c only utilizes the potential u at the probes B and C, i.e., the local information of the solution around the probes. To have a sharp estimation of the correction factor F_c, the local error around the probes is of interest and the local error estimation for the FEM approximation to u is wanted. Note that the right-hand side of (1.1) does not belong to the L^2 space. In this study, instead of the equation (1.1), a model problem $-\Delta u = f$ with $f \in L^2(\Omega)$is considered in Chapter 2.

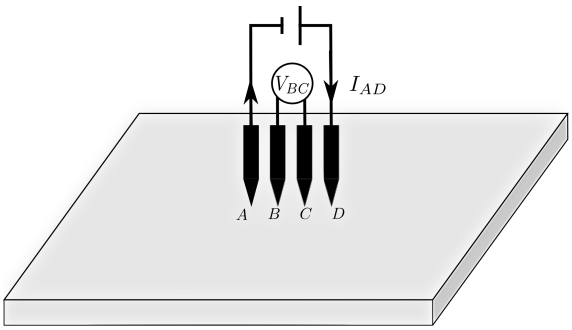


Figure 1.1: The four-probe method

In case that probe A and D have non-zero contact part $\Gamma_A, \Gamma_D \subset \partial\Omega$, respectively, the model boundary value problem (1.1) is approximated as the following non-homogeneous Neumann boundary value problem:

$$-\Delta u = 0 \text{ in } \Omega, \frac{\partial u}{\partial \mathbf{n}} = 0 \text{ on } \partial\Omega \setminus (\Gamma_A \cup \Gamma_D)\,, \frac{\partial u}{\partial \mathbf{n}} = -\frac{1}{|\Gamma_A|} \text{ on } \Gamma_A, \frac{\partial u}{\partial \mathbf{n}} = \frac{1}{|\Gamma_D|} \text{ on } \Gamma_D.$$

The guaranteed error estimation for the non-homogeneous Neumann boundary value problem is also essential to give a precise measurement method. In this dissertation, as an extension of the hypercircle method, the quantitative error estimation for the Neumann boundary value problems will be shown in Chapter 3, and it will be applied to give the eigenvalue bounds of the Steklov eigenvalue problems.

1.2 Survey

1.2.1 Hypercircle method

The hypercircle method, namely the Prager–Synge theorem, was developed more than fifty years ago by [41] for elastic analysis. Around the same time, based on the T^*T theory of Kato [19], Fujita developed a method similar to the hypercircle method [15], which has been applied to develop the point-wise estimation method for boundary value problems with specially constructed base functions. The application of the hypercircle method to the *a posteriori* error estimation can also be found in [6, 20, 27, 33, 36, 38]. In particular, Kikuchi's approach in [20] is to construct the hypercircle by utilizing $\mathbf{p}_h \in H(\mathrm{div}\,;\Omega)$ such that $\mathrm{div}\,\mathbf{p}_h + f_h = 0$ holds exactly instead of developing $\mathbf{p}_h \in H(\mathrm{div}\,;\Omega)$ by processing the discontinuity of ∇u_h across the edges of elements in [6, 38], where f_h is the projection of f to totally discontinuous piecewise polynomials. Further, Liu succeeded in deriving the *a priori* error estimation in [33] by inheriting Kikuchi's approach.

1.2.2 Guaranteed local error estimation

The local error estimation was studied in early time by Nitsche and Schatz [39]. In the studies of [39, 45], for the subdomain Ω_0 of interest, an intermediate subdomain Ω_1 such that $\Omega_0 \subset\subset \Omega_1 \subset\subset \Omega$ is utilized to deduce the following error estimation: for $u \in H^l(\Omega)$ $(l \geq 1)$,

$$\|u - u_h\|_{s,\Omega_0} \leq C(h^{l-s}\|u\|_{l,\Omega_1} + \|u - u_h\|_{-p,\Omega_1}) ,$$

for $s = 0$ or 1 and p as a fixed integer. In [39], an estimation for quasi-uniform meshes was provided. In [47, 48], based on the knowledge of the local error distribution, Xu and Zhou proposed a parallel technique that uses a coarse grid to approximate the low frequencies part of the residual error and then uses a refined grid for the high-frequency part. Discussions on relaxing the assumption in [39] applied to the mesh can be found in [11, 14]. In adaptive finite element methods, indicators based on the local error estimation have been well studied; see, e.g., [12, 13, 29, 47, 48]. The above results on local error estimation mainly focus on the qualitative analysis (e.g., convergence rate) of local error terms, while the explicit bound for the local error estimation is not available.

1.2.3 The error estimation for Neumann boundary value problems and the Steklov eigenvalue problems

For the finite element approximation of the Neumann boundary value problem, the past relevant studies mainly focused on the qualitative (e.g., convergence rate) error estimation. Recently, Li and Liu proposed an a posteriori quantitative error esti-

mation by utilizing the hypercircle method in [27] and it is motivated by giving the eigenvalue bound of the Steklov eigenvalue problem.

The Steklov eigenvalue problem is one of the important eigenvalue problems for differential operators; see [1, 4, 23] for a systematic introduction of background and applications. Below is a review of the numerical approaches to the eigenvalues of Steklov eigenvalue problems. The qualitative error estimation by conforming FEM for Steklov eigenvalue problems are discussed in [7,9,25,30], based on which [5,28,46] study more efficient algorithms such as two-grid and multilevel methods to solve Steklov eigenvalue problems. The *a posteriori* error estimates with conforming FEM and nonconforming Crouzeix-Raviart FEM are discussed in [3] and [42], respectively. Especially, in [26, 50], the asymptotic lower bounds for Steklov eigenvalue problems are discussed along with nonconforming finite elements. In [51], explicit lower bounds for the Steklov eigenvalues are obtained by using the Crouzeix-Raviart finite element along with an extension of the lower bound theorem of [31].

1.3 Contribution

In this study, we propose a quantitative error estimation method for the local error of the finite element solutions and a quantitative global error estimation for the Neumann boundary value problems. Such a method is regarded as an extension of the explicit error estimation theorem developed by Liu [33], which inherits the idea of Kikuchi [22] to utilize the hypercircle method.

Contributions of this study are summarized as follows:

(1) In this study, we successfully developed a new local error estimation method by incorporating cutoff functions in the calculation of inner product and norm in the hypercircle formula. By combining the conventional method by Kikuchi and Liu with our new approach, the following quantitative local error estimation was obtained (refer to Theorem 2.3.6 in Chapter 2).

$$\|\nabla(u - u_h)\|_S \leq \sqrt{E_1^2 + E_2^2} + 2\mathrm{Osc}(f).$$

Here, the left-hand side of the inequality represents the local error, while the quantity on the right-hand side is computable from the approximate solution. The numerical results presented in §4.2 of Chapter 4 demonstrate that the proposed method naturally handles cases requiring complex processing in previous studies and provides sharper error estimates compared to conventional global error estimation. This result is published in [37].

(2) This study proposed a new quantitative error estimation for finite element solutions of the non-homogeneous Neumann boundary value problem for the modified Helmholtz equation, which has not been extensively studied before (see Chapter 3, Theorem 3.3.3). The proposed error estimation has the following expression:

$$\|u - u_h\|_a \leq M_h \|f\|_b, \quad \|u - u_h\|_b \leq M_h^2 \|f\|_b.$$

where the left-hand side of inequalities is the error and M_h is the computable error constant. This study also presented computable the Steklov eigenvalue

bounds based on Liu's method [31] by utilizing newly developed error estimation (see Chapter 3, Theorem 3.4.1, Chapter 4 §4.2). The result is submitted to "Computational Methods in Applied Mathematics" (preprint is available in [35]).

1.4 Structure of dissertation

The rest of the paper is organized as follows: In Chapter 2, we discuss the hypercircle method and our newly developed guaranteed local error estimation referring to [37]. In Chapter 3, The hypercircle method is extended for the global error estimation for the Neumann boundary value problems of the modified Helmholtz equation, and we give the eigenvalue bound of the Steklov eigenvalue problems referring to [35]. In Chapter 4, we show numerical examples based on theoretical results in Chapter 2 and Chapter 3. Finally, we summarize the conclusions and discuss future studies in Chapter 5.

Chapter 2

Guaranteed local error estimation for boundary value problems

2.1 Preliminary

2.1.1 Problem settings

Throughout this study, the domain Ω is assumed to be a bounded polygonal domain of $\mathbb{R}^2$. Thus, Ω can be completely triangulated without any gap near the boundary. Standard symbols are used for the Sobolev spaces $H^m(\Omega)(m > 0)$. The norm of $L^2(\Omega)$ is written as $\|\cdot\|_{L^2(\Omega)}$ or $\|\cdot\|_\Omega$. Symbols $|\cdot|_{H^m(\Omega)}, \|\cdot\|_{H^m(\Omega)}$ denote the semi-norm and norm of $H^m(\Omega)$, respectively. Let $(\cdot,\cdot)$ be the inner product of $L^2(\Omega)$ or $(L^2(\Omega))^2$. Sobolev space $W^{1,\infty}(\Omega)$ is a function space where weak derivatives up to the first order are essentially bounded on Ω. The standard vector valued function

space $H(\mathrm{div};\Omega)$ is defined as follows:

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

In this chapter, the finite element solution for the following model boundary value problem will be discussed:

$$-\Delta u = f \text{ in } \Omega, \quad \frac{\partial u}{\partial \mathbf{n}} = g_N \text{ on } \Gamma_N, \quad u = g_D \text{ on } \Gamma_D. \tag{2.1}$$

Here, Γ_N and Γ_D are disjoint subsets of $\partial\Omega$ satisfying $\Gamma_N \cup \Gamma_D = \partial\Omega$; $\mathbf{n}$ is the unit outer normal direction on the boundary and $\frac{\partial}{\partial \mathbf{n}}$ is the directional derivative along $\mathbf{n}$ on $\partial\Omega$.

Let S be a subdomain of Ω of interest. Suppose that u_h is an approximate solution to the problem (2.1). The error of $(\nabla u - \nabla u_h)$ in the subdomain S will be evaluated in this study.

The weak form for the aforementioned problem is given by:

$$\text{Find } u \in V \text{ s.t. } (\nabla u, \nabla v) = (f, v) + (g_N, v)_{\Gamma_N} \quad \forall v \in V_0. \tag{2.2}$$

where

$$(g_N, v)_{\Gamma_N} := \int_{\Gamma_N} g_N v \ ds.$$

In case Γ_D is not an empty set, the function space V of the trial function and the function space V_0 of the test function are defined by

$$V := \left\{ v \in H^1(\Omega); v = g_D \text{ on } \Gamma_D \right\},\ V_0 := \left\{ v \in H^1(\Omega); v = 0 \text{ on } \Gamma_D \right\}.$$

For an empty Γ_D, the definition of V and V_0 are modified as follows:

$$V = V_0 = \left\{ v \in H^1(\Omega); \int_\Omega v \, dx = 0 \right\}.$$

2.1.2 Finite element space setting

To prepare for the discussion on the newly developed local error estimation in §2.3, we review the standard FEM approaches to (2.1). To simplify the discussion, assume g_D, g_N in the boundary conditions of the model problem (2.1) to be piecewise linear and piecewise constant at the boundary edges of $\mathcal{T}_h$, respectively. Let $\mathcal{T}_h$ be a proper triangulation of the domain Ω. Given an element $K \in \mathcal{T}_h$, let h_K denote the length of longest edge of K. The mesh size h of $\mathcal{T}_h$ is defined as follows:

$$h := \max_{K \in \mathcal{T}_h} h_K.$$

On each element $K \in \mathcal{T}_h$, the set of polynomials with degree up to d is denoted by $P_d(K)$. Let $V_h, V_{h,0}$ denote the conforming finite element spaces consisting of piecewise linear and continuous functions, the boundary conditions of which follow the settings of V and V_0, respectively. The finite element formulation of (2.2) is given by

$$\text{Find } u_h \in V_h \text{ s.t. } (\nabla u_h, \nabla v_h) = (f, v_h) + (g_N, v_h)_{\Gamma_N} \quad \forall v_h \in V_{h,0}. \tag{2.3}$$

To provide the local error estimation for $(\nabla u - \nabla u_h)$ over the subdomain S, let us introduce the following finite element spaces.

(a) Piecewise constant function space:

$$X_h := \left\{ v_h \in L^2(\Omega) \ : \ v_h|_K \in P_0(K), \ \forall K \in \mathcal{T}_h \right\}.$$

In case Γ_D is empty, it is further required that $\int_\Omega v_h \, dx = 0$ for $v_h \in X_h$.

(b) The Raviart–Thomas finite element space:

$$RT_h := \{ p_h \in H(\mathrm{div}; \Omega) \ : \ p_h|_K = (a_K + c_K x, b_K + c_K y) \text{ for } K \in \mathcal{T}_h \}.$$

$$RT_{h,0} := \{ p_h \in RT_h \ : \ p_h \cdot \mathbf{n} = 0 \text{ on } \Gamma_N \}.$$

Here, $a_K, b_K, c_K \in P_0(K)$ for $K \in \mathcal{T}_h$.

The standard mixed finite element formulation of (2.1) reads: Find $(p_h, \mu_h) \in RT_h \times X_h, \ p_h \cdot \mathbf{n} = g_N$ on Γ_N, s.t.

$$(p_h, q_h) + (\mathrm{div} \ q_h, \mu_h) + (\mathrm{div} \ p_h, \eta_h) = \int_{\Gamma_D} g_D (q_h \cdot \mathbf{n}) \, ds - (f, \eta_h) \tag{2.4}$$

for $(q_h, \eta_h) \in RT_{h,0} \times X_h$.

Define the projection $\pi_h : L^2(\Omega) \to X_h$ such that for $f \in L^2(\Omega)$,

$$(f - \pi_h f, \eta_h) = 0 \quad \forall \eta_h \in X_h.$$

The following error estimation holds for π_h along with an error constant C_0,

$$\| f - \pi_h f \|_\Omega \le C_0 h |f|_{H^1(\Omega)} \quad \forall f \in H^1(\Omega). \tag{2.5}$$

To give a concrete value of C_0 in (2.5), let us define $C_0(K)$ as a constant that depends on the shape of the triangle $K \in \mathcal{T}_h$ and satisfies

$$\|f - \pi_h f\|_K \leq C_0(K)|f|_{H^1(K)} \quad \forall f \in H^1(K).$$

By using $C_0(K)$, the constant C_0 that depends on triangulation can be defined by

$$C_0 := \max_{K \in \mathcal{T}_h} \frac{C_0(K)}{h}. \tag{2.6}$$

The previous studies [20, 21, 24, 32] reported that the optimal value of $C_0(K)$ is given by $C_0(K) := h_K/j_{1,1} (\leq 0.261 h_K)$ using positive minimum root $j_{1,1} \approx 3.83171$ of the first kind Bessel's function J_1.

2.2 Global *a priori* error estimation for the finite element solutions

In this section, we introduce the global error estimation developed in [27, 33, 40], which will be used in Lemma 2.3.5 for local error estimation. We focus on the global *a priori* error estimation for problems with homogeneous boundary value conditions, which fits the needs in the proof for Lemma 2.3.5. For global *a priori* error estimation of non-homogeneous boundary value problems, refer to [27].

As a preparation for Lemma 2.3.5, let us consider the following boundary value problem.

$$-\Delta\phi = f \text{ in } \Omega, \quad \frac{\partial \phi}{\partial \mathbf{n}} = 0 \text{ on } \Gamma_N, \quad \phi = 0 \text{ on } \Gamma_D. \tag{2.7}$$

The weak formulation of (2.7) seeks $\phi \in V_0$, s.t.,

$$(\nabla\phi, \nabla v) = (f, v), \quad \forall v \in V_0. \tag{2.8}$$

The Galerkin projection operator $P_h : V_0 \to V_{h,0}$ satisfies, for $v \in V_0$

$$(\nabla(v - P_h v), \nabla v_h) = 0 \quad \forall v_h \in V_{h,0}\,. \tag{2.9}$$

In [33], the following quantity κ_h is introduced for the purpose of *a priori* error estimation to the Galerkin projection $P_h\phi$:

$$\kappa_h := \max_{f_h \in X_h} \min_{\substack{v_h \in V_{h,0},\ q_h \in RT_{h,0},\\ \text{div } q_h + f_h = 0}} \frac{\|\nabla v_h - q_h\|}{\|f_h\|}.$$

The theorem below provides an *a priori* error estimation using κ_h and the Prager–Synge theorem.

Theorem 2.2.1 (Global *a priori* error estimation [33])**.** *Given $f \in L^2(\Omega)$, let ϕ be the solution to* (2.8)*. Then, the following error estimation holds.*

$$|\phi - P_h\phi|_{H^1(\Omega)} \le C(h)\|f\|_\Omega, \tag{2.10}$$

$$\|\phi - P_h\phi\|_\Omega \le C(h)|\phi - P_h\phi|_{H^1(\Omega)} \le C(h)^2\|f\|_\Omega\,, \tag{2.11}$$

where $C(h) := \sqrt{\kappa_h^2 + (C_0 h)^2}$; C_0 is the quantity defined in (2.6)*.*

Remark 2.2.2. We point out that the result in Theorem 2.2.1 is applicable to non-convex domains, for which the solution may not belong to $H^2(\Omega)$. When the exact solution ϕ belongs $H^2(\Omega)$, the constant $C(h)$ appearing in (2.10), (2.11) can be

replaced for the error constant of the Lagrange interpolation. For example, in §4.1.2, it is possible to use $C_h = 0.493h$ as $C(h)$ for the right-angled triangle mesh [20, 21].

Remark 2.2.3. (Calculation of κ_h) Given $f_h \in X_h$, let $R_h : X_h \to V_h$, $T_h : X_h \to RT_h$ be the linear operators that map f_h to the Lagrange FEM approximation of $\nabla\phi$ and the Raviart–Thomas FEM approximation of $\nabla\phi$, respectively. Then, κ_h is characterized by the following maximum formulation:

$$\kappa_h = \max_{f_h \in X_h} \frac{\|(R_h - T_h)f_h\|_\Omega}{\|f_h\|_\Omega},$$

which is determined by solving a matrix eigenvalue problem. In [27, 33], κ_h is calculated by solving two FEM solutions involving the Lagrange FEM and the Raviart–Thomas FEM. To reduce the computational cost, one can utilize the relationship between the Crouzeix–Raviart FEM and the Raviart–Thomas FEM, which has been studied in [17, 34]. That is, the Raviart–Thomas solution can be obtained by a post-processing of the Lagrange FEM solution to (2.8) (see, e.g., [6, 44]). Such discussion will be considered in a future study.

2.3 Weighted hypercircle formula and the guaranteed local error estimation

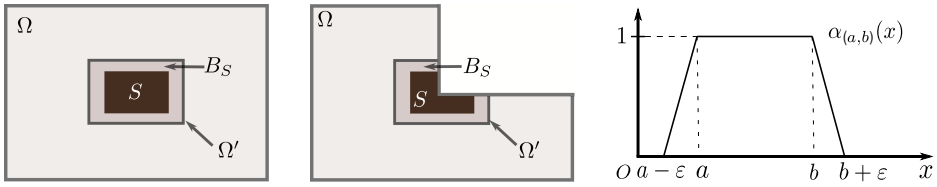
In this section, we propose guaranteed local error estimation for the finite element solutions. Let $S(\subset \Omega)$ be the subdomain of interest. In §2.3.1, the weighted inner product and weighted norm corresponding to S will be introduced through a cutoff function α. In §2.3.2, we show the weighted hypercircle formula as an extension of

the Prager–Synge theorem. The result of the local error estimation will be provided in §2.3.3.

2.3.1 The weight function

Let Ω' be the extended domain of S with a band of width ε, that is, $\Omega' := \{x \in \Omega \ ; \ \mathrm{dist}(x, S) < \varepsilon\}$. Denote the band surrounding S by B_S. Refer to Figure 2.1-(a),(b) for two examples of S and B_S. The weight function $\alpha \in W^{1,\infty}(\Omega)$ is defined as a piecewise polynomial with the following property.

$$\alpha(x,y) = \begin{cases} 1 & (x,y) \in S \\ 0 & (x,y) \in (\Omega')^c \end{cases} , \quad 0 \le \alpha(x,y) \le 1, \text{ for } (x,y) \in B_S \, .$$



(a) Square domain. (b) L-shaped domain. (c) Graph of $\alpha_{(a,b)}(x)$.

Figure 2.1: Definition of the weight function α.

To construct a concrete weight function $\alpha(x, y)$, let us define $\alpha_{(a,b)}$ over interval (a, b) as follows.

$$\alpha_{(a,b)}(x) := \begin{cases} 1 + (x-a)/\varepsilon & x \in (a-\varepsilon, a] \\ 1 & x \in (a, b) \\ 1 - (x-b)/\varepsilon & x \in [b, b+\varepsilon) \\ 0 & \text{otherwise} \end{cases} .$$

Refer to Figure 2.1-(c) for the graph of $\alpha_{(a,b)}$. For S being a rectangular subdomain constructed by the Cartesian product of two open intervals (x_a, x_b), (y_a, y_b), the weight function α can be defined by $\alpha(x, y) = \min\{\alpha_{(x_a,x_b)}(x), \alpha_{(y_a,y_b)}(y)\}$.

The weighted inner product and norm are defined by using α as follows.

(a) Weighted inner product $(\cdot,\cdot)_\alpha$: For $f, g \in L^2(\Omega)$ or $f, g \in (L^2(\Omega))^2$,

$$(f, g)_\alpha := \int_{\Omega'} \alpha^2 f \cdot g \; dx.$$

(b) Weighted norm $\|\cdot\|_\alpha$: For $f \in L^2(\Omega)$,

$$\|f\|_\alpha := \sqrt{(f, f)_\alpha} = \sqrt{\int_{\Omega'} \alpha^2 f^2 \; dx} \quad (= \|\alpha f\|_\Omega).$$

The following inequalities hold.

$$\|f\|_S \le \|f\|_\alpha \le \|f\|_{\Omega'} \le \|f\|_\Omega. \tag{2.12}$$

2.3.2 Weighted hypercircle formula

In this subsection, a weighted hypercircle formula is proposed, which can be regarded as an extension to the classical Prager–Synge theorem below.

Prager–Synge's theorem [41] Let ϕ be the solution of (2.7). For any $v \in V_0$ and $\widetilde{p} \in H(div;\Omega)$ satisfying

$$div\ \widetilde{p} + f = 0 \text{ in } \Omega,\ \widetilde{p} \cdot \mathbf{n} = 0 \text{ on } \Gamma_N\ ,$$

we have,

$$\|\nabla\phi - \nabla v\|_\Omega^2 + \|\nabla\phi - \widetilde{p}\|_\Omega^2 = \|\nabla v - \widetilde{p}\|_\Omega^2. \tag{2.13}$$

The quantities $\{\nabla\phi, \nabla v, \widetilde{p}\}$ in the above equation formulate a hypercircle as illustrated in Figure 2.2.

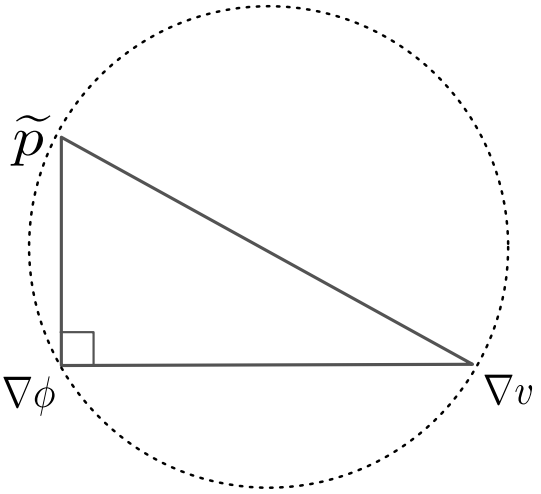


Figure 2.2: Hypercircle for $\{\nabla\phi, \nabla v, \widetilde{p}\}$

For weighted norms introduced in the previous subsection, we have the following extended formulation of the hypercircle (2.13).

Theorem 2.3.1. *Let u be the solution of (2.1). For any $v \in V_h$ and $p \in H(\mathrm{div};\Omega)$ satisfying*

$$\mathrm{div}\ p + f = 0\ \textit{in}\ \Omega, \quad p \cdot \mathbf{n} = g_N\ \textit{on}\ \Gamma_N.$$

Then,

$$\|\nabla u - \nabla v\|_\alpha^2 \leq \|\nabla v - p\|_\alpha^2 + 8\,\|\nabla\alpha\|_{L^\infty(\Omega)}^2 \|u - v\|_{\Omega'}^2.$$

Here,

$$\|\nabla\alpha\|_{L^\infty(\Omega)} = \max\left\{ \left\| \frac{\partial\alpha}{\partial x} \right\|_{L^\infty(\Omega)}, \ \left\| \frac{\partial\alpha}{\partial y} \right\|_{L^\infty(\Omega)} \right\}.$$

Proof. The expansion of $\|\nabla v - p\|_\alpha^2 = \|(\nabla v - \nabla u) + (\nabla u - p)\|_\alpha^2$ tells that

$$\|\nabla v - p\|_\alpha^2 = \|\nabla v - \nabla u\|_\alpha^2 + \|\nabla u - p\|_\alpha^2 + 2(\nabla v - \nabla u, \nabla u - p)_\alpha. \tag{2.14}$$

Let $w := v - u$. Below, we show the estimation for the cross-term of (2.14), i.e., $(\nabla w, \nabla u - p)_\alpha$.

To deal with $(\nabla w, \nabla u)_\alpha$, let us take the test function as $\alpha^2 w$ in (2.2) and apply the chain rule to $\alpha^2 w$, i.e., $\nabla(\alpha^2 w) = w\nabla\alpha^2 + \alpha^2\nabla w$. Then, we have

$$(\nabla w, \nabla u)_\alpha = -\int_\Omega w\nabla\alpha^2 \cdot \nabla u \; dx + \int_\Omega f\,(\alpha^2 w)\; dx + \int_{\Gamma_N} g_N\,(\alpha^2 w) ds. \tag{2.15}$$

For $(\nabla w, p)_\alpha$, the fact $w = 0$ on Γ_D and Green's formula tell that

$$\begin{aligned}
(\nabla w, p)_\alpha &= \int_\Omega \nabla w \cdot (\alpha^2 p)\; dx = \int_{\Gamma_N} g_N\,(\alpha^2 w) ds - \int_\Omega w\ \mathrm{div}(\alpha^2 p)\; dx \\
&= \int_{\Gamma_N} g_N\,(\alpha^2 w) ds - \int_\Omega (\alpha^2 w)\ \mathrm{div}\ p\; dx - \int_\Omega w\nabla\alpha^2 \cdot p\; dx. \\
&= -\int_\Omega w\nabla\alpha^2 \cdot p\; dx + \int_\Omega f\,(\alpha^2 w)\; dx + \int_{\Gamma_N} g_N\,(\alpha^2 w) ds.
\end{aligned} \tag{2.16}$$

By taking (2.15)-(2.16) and noticing that $\alpha = 0$ on $\Omega \setminus \Omega'$, we have

$$(\nabla v - \nabla u, \nabla u - p)_\alpha = -\int_{\Omega'} (v - u)\nabla\alpha^2 \cdot (\nabla u - p)\; dx \tag{2.17}$$

Move the cross term (2.17) in (2.14) to the left-hand side, From, (2.17) and (2.14), we have

$$\|\nabla v - p\|_\alpha^2 \geq \|\nabla v - \nabla u\|_\alpha^2 + \|\nabla u - p\|_\alpha^2 - 2\left|\int_{\Omega'} (v-u)\nabla\alpha^2 \cdot (\nabla u - p)\ dx\right| \quad (2.18)$$

From the relation $\nabla\alpha^2 = 2\alpha\nabla\alpha$ and the inequality $2|ab| \leq 2a^2 + b^2/2$, it holds

$$\begin{aligned}
&\left|\int_{\Omega'} 2(u-v)\nabla\alpha \cdot \alpha(\nabla u - \mathbf{p})\ dx\right| \\
&\leq \int_{\Omega'} 2|(u-v)\nabla\alpha|^2 + |\alpha(\nabla u - \mathbf{p})|^2/2\ dx \\
&\leq 2\left(\|\frac{\partial\alpha}{\partial x}\|_{L^\infty(\Omega)}^2 + \|\frac{\partial\alpha}{\partial y}\|_{L^\infty(\Omega)}^2\right)\ \|u-v\|_{\Omega'}^2 + \ \|\nabla u - \mathbf{p}\|_\alpha^2/2 \\
&\leq 4\|\nabla\alpha\|_{L^\infty(\Omega)}^2 \|u-v\|_{\Omega'}^2 + \ \|\nabla u - \mathbf{p}\|_\alpha^2/2. \qquad (2.19)
\end{aligned}$$

By the substitution of (2.19) to (2.18), we draw the conclusion.

Remark 2.3.2. Theorem 2.3.1 holds no matter $\partial S \cap \partial\Omega = \emptyset$ or not, as can be confirmed in the proof.

2.3.3 Guaranteed local error estimation for finite element solutions

As a preparation to the argument of the main result, let us follow the idea of Kikuchi [22] to introduce auxiliary functions $\overline{u} \in V$ and $\overline{u}_h \in V_h$ as the solutions to the

following equations.

$$(\nabla\overline{u}, \nabla v) = (\pi_h f, v) + (g_N, v)_{\Gamma_N},\ \forall v \in V_0\,; \tag{2.20}$$

$$(\nabla\overline{u}_h, \nabla v_h) = (\pi_h f, v_h) + (g_N, v_h)_{\Gamma_N},\ \forall v_h \in V_{h,0}\,. \tag{2.21}$$

Both solutions are introduced only for error analysis in a theoretical way, and the above equations do not need to be solved explicitly.

Lemma 2.3.3. *For $\overline{u}$ of (2.20) and $\overline{u}_h$ of (2.21), the following estimations hold.*

$$|u - \overline{u}|_{H^1(\Omega)} \leq \mathrm{Osc}(f), \tag{2.22}$$

$$|u_h - \overline{u}_h|_{H^1(\Omega)} \leq \mathrm{Osc}(f). \tag{2.23}$$

Here, h_K is the diameter of triangle K and

$$Osc(f) := \left\{ \sum_{K\in\mathcal{T}_h} C_0(K)^2 \|f - \pi_h f\|_K^2 \right\}^{\frac{1}{2}}. \tag{2.24}$$

Proof. According to the definitions of u and $\overline{u}$,

$$(\nabla(u-\overline{u}), \nabla v) = (f-\pi_h f, v) = (f-\pi_h f, v-\pi_h v) = \sum_{K\in\mathcal{T}_h} (f-\pi_h f, v-\pi_h v)_K, \quad v \in V_0.$$

By taking $v = u-\overline{u}$ in the above equation and using the error estimation of projection π_h in (2.5), we have

$$|u-\overline{u}|^2_{H^1(\Omega)} = (\nabla(u-\overline{u}), \nabla(u-\overline{u})) \leq \sum_{K\in\mathcal{T}_h} \|f-\pi_h f\|_K \cdot C_0(K)|u-\overline{u}|_{H^1(K)} \leq \mathrm{Osc}(f)|u-\overline{u}|_{H^1(\Omega)}.$$

Hence,

$$|u - \overline{u}|_{H^1(\Omega)} \leq \mathrm{Osc}(f).$$

Estimation (2.23) is obtained in the same way.

Remark 2.3.4. The introduction $\mathrm{Osc}(f)$ in (2.24) takes advantage of the non-uniform mesh. Define $\widehat{\mathrm{Osc}}(f)$ by

$$\widehat{\mathrm{Osc}}(f) := C_0 h \|f - \pi_h f\|_\Omega \quad (C_0 \leq 0.261).$$

Then, $\mathrm{Osc}(f) \leq \widehat{\mathrm{Osc}}(f)$ holds for general meshes. In the case of uniform meshes, $\mathrm{Osc}(f) = \widehat{\mathrm{Osc}}(f) = O(h^2)$ for a smooth f.

Below, we apply Theorem 2.3.1 to the current function settings.

To state the results in Lemma 2.3.5 and 2.3.6, let us define the following four quantities $\overline{E}_1,\ \overline{E}_2,\ E_1$ and E_2. Note that the explicit value of E_1, E_2 will be utilized in the final error estimation.

$$\overline{E}_1 := \|\nabla \overline{u}_h - p_h\|_\alpha, \quad \overline{E}_2 := 2\sqrt{2}C(h)\ \|\nabla\alpha\|_{L^\infty(\Omega)} \|\nabla \overline{u}_h - p_h\|_\Omega,$$
$$E_1 := \|\nabla u_h - p_h\|_\alpha + \mathrm{Osc}(f), \quad E_2 := 2\sqrt{2}C(h)\ \|\nabla\alpha\|_{L^\infty(\Omega)}\ \|\nabla u_h - p_h\|_\Omega\ .$$

Lemma 2.3.5 (Local error estimation for u using $\overline{u}_h$)**.** *Let u and $\overline{u}_h$ be the solutions of* (2.2), (2.21)*, respectively. For $p_h \in RT_h$ satisfying*

$$\mathrm{div}\ p_h + \pi_h f = 0\ \textit{in}\ \Omega,\ p_h \cdot \mathbf{n} = g_N\ \textit{on}\ \Gamma_N\ , \tag{2.25}$$

the following local error estimation holds.

$$\|\nabla u - \nabla \overline{u}_h\|_S \leq \sqrt{\overline{E}_1^2 + \overline{E}_2^2} + \mathrm{Osc}(f). \tag{2.26}$$

Proof. With $\overline{u}$ defined in (2.20) and the triangle inequality, we obtain:

$$\|\nabla(u - \overline{u}_h)\|_S \leq \|\nabla(u - \overline{u})\|_S + \|\nabla(\overline{u} - \overline{u}_h)\|_S \leq \|\nabla(u - \overline{u})\|_\Omega + \|\nabla(\overline{u} - \overline{u}_h)\|_\alpha.$$

By applying the estimation of $\|\nabla(u - \overline{u})\|_\Omega$ in Lemma 2.3.3 and Theorem 2.3.1 to $\|\nabla(\overline{u} - \overline{u}_h)\|_\alpha$, the following estimation holds.

$$\|\nabla(u - \overline{u}_h)\|_S \quad \leq \quad \left\{\|\nabla \overline{u}_h - p_h\|_\alpha^2 + 8\|\nabla\alpha\|_\infty^2 \|\overline{u} - \overline{u}_h\|_{\Omega'}^2\right\}^{\frac{1}{2}} + \mathrm{Osc}(f)\,. \tag{2.27}$$

Next, we give the estimation for $\|\overline{u} - \overline{u}_h\|_{\Omega'}$ in (2.27) by considering its bound $\|\overline{u} - \overline{u}_h\|_\Omega$. For this purpose, let us define the dual problem.

$$\text{Find } \phi \in V_0 \text{ s.t. } (\nabla\phi, \nabla v) = (\overline{u} - \overline{u}_h, v) \quad \forall v \in V_0.$$

By applying P_h defined in (2.9) along with the *a priori* estimation (2.10) in Theorem 2.2.1, we have,

$$\begin{aligned} \|\overline{u} - \overline{u}_h\|_\Omega^2 &\leq \|\nabla(\phi - P_h\phi)\|_\Omega \, \|\nabla(\overline{u} - \overline{u}_h)\|_\Omega \\ &\leq C(h) \, \|\overline{u} - \overline{u}_h\|_\Omega \, \|\nabla(\overline{u} - \overline{u}_h)\|_\Omega. \end{aligned}$$

From the Prager–Synge theorem, the hypercircle below is available for $\overline{u}$ defined in

(2.20) and p_h in (2.25),

$$\|\nabla\overline{u} - \nabla v_h\|_\Omega^2 + \|\nabla\overline{u} - p_h\|_\Omega^2 = \|\nabla v_h - p_h\|_\Omega^2, \quad \forall v_h \in V_h\,. \tag{2.28}$$

By taking $v_h := \overline{u}_h$, we obtain the following estimations:

$$\|\nabla\overline{u} - p_h\|_\Omega \le \|\nabla\overline{u}_h - p_h\|_\Omega\,, \quad \|\nabla\overline{u} - \nabla\overline{u}_h\|_\Omega \le \|\nabla\overline{u}_h - p_h\|_\Omega\,. \tag{2.29}$$

Thus, we have the estimation of $\|\overline{u} - \overline{u}_h\|_\Omega$:

$$\|\overline{u} - \overline{u}_h\|_\Omega \le C(h)\|\nabla(\overline{u} - \overline{u}_h)\|_\Omega \le C(h)\|\nabla\overline{u}_h - p_h\|_\Omega. \tag{2.30}$$

Finally, applying (2.29) and (2.30) to the first term of the right-hand side of (2.27), we have

$$\begin{aligned}
&\|\nabla\overline{u}_h - p_h\|_\alpha^2 + 8\,\|\nabla\alpha\|_\infty^2\|\overline{u} - \overline{u}_h\|_\Omega^2 \\
&\le \|\nabla\overline{u}_h - p_h\|_\alpha^2 + 8\,\|\nabla\alpha\|_\infty^2\,C(h)^2\,\|\nabla\overline{u}_h - p_h\|_\Omega^2 \\
&= \overline{E}_1^2 + \overline{E}_2^2.
\end{aligned}$$

Now, we draw the conclusion by sorting the estimation of (2.27).

Theorem 2.3.6. *Under the assumptions of Lemma 2.3.5, the following estimation holds.*

$$\|\nabla u - \nabla u_h\|_S \le \sqrt{E_1^2 + E_2^2} + 2\,\mathrm{Osc}(f)\ (=: \widehat{E}_L)\,. \tag{2.31}$$

Proof. First, we apply the triangle inequality to $(u - \overline{u}_h) + (\overline{u}_h - u_h)$ and the

estimation (2.23) in Lemma 2.3.3 to have,

$$\|\nabla(u-u_h)\|_S \le \|\nabla(u-\overline{u}_h)\|_S + \|\nabla(\overline{u}_h-u_h)\|_S \le \|\nabla(u-\overline{u}_h)\|_S + \mathrm{Osc}(f)\ .$$

Next, we apply the result in Lemma 2.3.5 to $\|\nabla(u-\overline{u}_h)\|_S$ and process the term $\overline{u}_h$ in $\overline{E}_1$ and $\overline{E}_2$. For the term $\|\nabla\overline{u}_h - p_h\|_\alpha$ in $\overline{E}_1$, apply the triangle inequality and (2.23) to obtain

$$\overline{E}_1 = \|\nabla\overline{u}_h - p_h\|_\alpha \le \|\nabla\overline{u}_h - \nabla u_h\|_\alpha + \|\nabla u_h - p_h\|_\alpha \le E_1\ .$$

For the term $\|\nabla\overline{u}_h - p_h\|_\Omega$ in $\overline{E}_2$, we utilize the minimization principle for the approximation to $\overline{u}$ in V_h, i.e., $\|\nabla\overline{u} - \nabla\overline{u}_h\|_\Omega \le \|\nabla\overline{u} - \nabla u_h\|_\Omega$. This inequality and the two equations obtained by substituting $v_h := \overline{u}_h$ and $v_h := u_h$ in the hypercircle (2.28) lead to

$$\|\nabla\overline{u}_h - p_h\|_\Omega \le \|\nabla u_h - p_h\|_\Omega\ .$$

Hence, $\overline{E}_2 \le E_2$. Now, we draw the conclusion as in (2.31).

Remark 2.3.7. The estimation (2.22) along with the hypercircle (2.28) leads to an *a posteriori* estimation of the global error.

$$\|\nabla(u-u_h)\|_\Omega \le \|\nabla u_h - p_h\|_\Omega + \mathrm{Osc}(f)\ . \tag{2.32}$$

Here, p_h can be chosen freely to approximate ∇u under the condition (2.25). Such an estimation can be regarded as a revision of Kikuchi's result [20] for non-uniform meshes.

2.3.4 Convergence analysis and application to non-uniform meshes

In this subsection, we have an analysis on the convergence behavior for the proposed error estimation and show its application in efficient computing with non-uniform meshes.

For a solution $u \in H^2$ solved by FEM over a uniform mesh with mesh size h, the global error terms $\|\nabla u_h - p_h\|_\Omega$ and $C(h)$ have the convergence rate as $O(h)$. Thus, the following convergence rate is available from Theorem 2.3.6.

$$E_1 = O(h),\ E_2 = O(h^2),\ \widehat{E}_L = O(h)\ . \tag{2.33}$$

Application to non-uniform mesh. Theoretical analysis on local error estimation tells that for $u \in H^2(\Omega)$ ([47, 48]):

$$\|u - u_h\|_{1,S} = O(h_{\Omega'} + h_G^2). \tag{2.34}$$

Here, $h_{\Omega'}$ denotes the mesh size of Ω'; h_G the one for the mesh outside of Ω'. Estimation (2.34) implies an asymptotically optimal error with rate $O(h_{\Omega'})$ in H^1 norm locally by taking $h_G = O(\sqrt{h_{\Omega'}})$.

Such *a priori* estimation motivates us to apply our proposed error estimator to non-uniform meshes to have more efficient computation and error estimation. For the error term E_1, it is expected that $\|\nabla u_h - p_h\|_\alpha = O(h_{\Omega'} + h_G^2)$, when p_h also provides a good approximation to ∇u as in (2.34). However, such a result is not discussed yet in the existing literature. In this dissertation, rather than theoretical

analysis, we perform a numerical experiment with the mesh size of the non-uniform mesh selected as $h_G = \sqrt{h_{\Omega'}}$, and confirm that (see details in Chapter 4, §4.1)

$$E_1,\ E_2,\ \widehat{E}_L = O(h_{\Omega'}) = O(h_G^2)\ .$$

Chapter 3

Guaranteed error estimation for the non-homogeneous Neumann boundary condition

3.1 Finite element approximation of the Neumann boundary value problem

3.1.1 Objective problem

In this chapter, we consider the following Neumann boundary value problem appearing in the Steklov eigenvalue problems:

$$-\Delta u + cu = 0 \text{ in } \Omega; \quad \frac{\partial u}{\partial \mathbf{n}} = f \text{ on } \Gamma = \partial\Omega\,. \tag{3.1}$$

Note that in case $c = 0$, f is further required to satisfy $\int_{\partial\Omega} f \mathrm{d}s = 0$. For a positive c, we take $V = H^1(\Omega)$. If $c = 0$, then the function obtained by adding a constant to one solution of (3.1) also satisfies the same boundary value problem, which means that the solution of (3.1) is not unique. Upon this property, let us take $V := \{v \in H^1(\Omega) \ : \ \int_\Gamma v \mathrm{d}s = 0\}$ when $c = 0$.

A weak formulation of the above problem is as follows: Find $u \in V$ such that

$$a(u, v) = b(f, v) \quad \forall v \in V, \tag{3.2}$$

where

$$a(u, v) := \int_\Omega \nabla u \cdot \nabla v + cuv \ \mathrm{d}x, \quad b(u, v) := \int_{\partial\Omega} uv \ \mathrm{d}s.$$

Evidently the bilinear form $a(\cdot, \cdot)$ is symmetric, continuous and coercive over V. The norm induced by $a(\cdot, \cdot)$ (resp. $b(\cdot, \cdot)$) is denoted by $\|u\|_a := \sqrt{a(u, u)}$ (resp. $\|u\|_b := \sqrt{b(u, u)}$).

3.1.2 Finite element approximation

Let $\mathcal{T}_h$ be a shape regular triangulation of the domain Ω. For each element $K \in \mathcal{T}_h$, denote by h_K the longest edge length of K and define the mesh size h by the maximal value of h_K. Particularly, it is assumed that, at corners of the domain, each boundary edge of the triangulation is only shared by one triangle. Such an assumption is utilized in the proof of Lemma 3.3.2 to have a sharper error estimation.

The piecewise linear H^1-conforming finite element space V^h is defined by

$$V^h \ := \ \{v_h \in V : \ v_h|_K \in P_1(K) \ \ \forall K \in \mathcal{T}_h\} \ ,$$

where $P_1(K)$ is the space of polynomials of degree ≤ 1 on K.

The conforming finite element approximation of (3.2) is defined as follows: Find $u_h \in V^h$ such that

$$a(u_h, v_h) = b(f, v_h) \quad \forall v_h \in V^h. \tag{3.3}$$

In this chapter, the following classical finite element spaces will be used in constructing the *a priori* error estimate for the FEM solution. Let E_h be the set of edges of the triangulation, and $E_{h,\Gamma}$ the set of edges on the boundary of the domain. Let $\mathcal{T}_h^b$ be the set of elements of $\mathcal{T}_h$ having at least one edge on $\partial\Omega$.

(i) Piecewise function spaces X^h and X_Γ^h:

$$\begin{aligned} X^h &:= \{v \in L^2(\Omega) : v|_K \in P_1(K) \quad \forall K \in \mathcal{T}_h\} \\ X_\Gamma^h &:= \{v \in L^2(\Gamma) : v|_e \in P_1(e) \quad \forall e \in E_{h,\Gamma}\} \end{aligned}$$

where $P_1(e)$ is the space of polynomials of degree ≤ 1 on the edge e. In case that $c = 0$, we further assume that $\int_\Gamma v \, \mathrm{d}s = 0$ for $v \in X_\Gamma^h$.

(ii) The Raviart–Thomas FEM space W^h with order one ([8]):

$$W^h := \Big\{ p_h \in H(\mathrm{div}, \Omega) \quad | \quad p_h = (a_K, b_K) + c_K \, (x, y) \, , \\ a_K, b_K, c_K \in P_1(K) \text{ for } K \in \mathcal{T}_h \Big\}.$$

The freedoms of the Raviart–Thomas FEM space can be defined by the normal trace of p_h on the edges of the triangulation. Hence, $\{(p_h \cdot \mathbf{n})|_\Gamma \mid p_h \in W^h\} = X^h$. The space $W_{f_h}^h$ is a subset of W^h corresponding to $f_h \in X_\Gamma^h$:

$$W_{f_h}^h := \{p_h \in W^h \mid p_h \cdot \mathbf{n} = f_h \text{ on } \Gamma\}.$$

In particular, $W_0^h := \{p_h \in W^h \mid p_h \cdot \mathbf{n} = 0 \text{ on } \Gamma\}$.

Under current space settings, the following relations are available.

$$V^h \subset X^h, \quad \mathrm{div}(W^h) = X^h, \quad \gamma(V^h) \subset X_\Gamma^h \ .$$

3.2 Hypercircle method for the modified Helmholtz equations

In this section, we introduce the hypercircle to be used to facilitate the error estimate in solving the eigenvalue problem. Let us introduce the following semi-norm (or norm if $c > 0$) for $p \in H(\mathrm{div}; \Omega)$:

$$\|p\|^2_{H(\mathrm{div}),c} := \int_\Omega |\mathrm{div}\ p|^2 + c|p|^2 \mathrm{d}\Omega \ .$$

Theorem 3.2.1. *Given $f_h \in X_\Gamma^h$, let u be the solution of (3.2) with $f := f_h$. For $v_h \in V^h$ and $p_h \in W_{f_h}^h$ satisfying div $p_h = cv_h$, the following hypercircle holds:*

$$\|u - v_h\|_a^2 + \|\nabla u - p_h\|^2_{H(div),c} = \|\nabla v_h - p_h\|^2_{L^2} \ . \tag{3.4}$$

Proof. Rewriting $\nabla v_h - p_h$ by $(\nabla v_h - \nabla u) + (\nabla u - p_h)$, we have

$$\|\nabla v_h - p_h\|^2_{L^2} = \|\nabla v_h - \nabla u\|^2_{L^2} + \|\nabla u - p_h\|^2_{L^2} + 2(\nabla v_h - \nabla u, \nabla u - p_h).$$

Furthermore, the Green theorem and the Neumann boundary conditions setting lead

to

$$\begin{aligned}(\nabla u_h - \nabla u, \nabla u - p_h) &= (v_h - u, -cu + \operatorname{div} p_h) \\ &= (v_h - u, -cu + cv_h) = c\|u - v_h\|_{L^2}^2.\end{aligned}$$

Noticing that $\|\nabla u - p_h\|_{H(\mathrm{div}),c}^2 = \|\nabla u - p_h\|_{L^2}^2 + c\|u - v_h\|_{L^2}^2$, we obtain the hypercircle in (3.4).

Next, let us introduce the quantity κ_h such that

$$\kappa_h := \max_{f_h \in X_\Gamma^h \setminus \{0\}} \min_{\substack{v_h \in V^h,\ p_h \in W_{f_h}^h \\ \operatorname{div} p_h = c v_h}} \frac{\|\nabla v_h - p_h\|_{L^2}}{\|f_h\|_b}. \tag{3.5}$$

Lemma 3.2.2. *Given $f_h \in X_\Gamma^h$, let $\tilde{u} \in V$ and $\tilde{u}_h \in V^h$ be the solutions to the following variational problems, respectively,*

$$\begin{aligned}a(\tilde{u}, v) &= b(f_h, v) \quad \forall v \in V, \\ a(\tilde{u}_h, v_h) &= b(f_h, v_h) \quad \forall v_h \in V^h.\end{aligned} \tag{3.6}$$

Then, the following error estimate holds:

$$\|\tilde{u} - \tilde{u}_h\|_a \le \kappa_h \|f_h\|_b \ . \tag{3.7}$$

Proof. In Theorem 3.2.1, take $v_h := \tilde{u}_h$, $u := \tilde{u}$ and $p_h \in W_{f_h}^h$ such that $\operatorname{div} p_h = c\tilde{u}_h$, then we have

$$\|\tilde{u} - \tilde{u}_h\|_a \le \|\nabla \tilde{u}_h - p_h\|_{L^2}. \tag{3.8}$$

By further considering the minimization of p_h and the variation of f_h in X_Γ^h, we draw the conclusion in (3.7).

Remark 3.2.3. In Theorem 3.3 of [27], a general case such that div $p_h - c\tilde{u}_h \neq 0$ is discussed, for which the formulation of κ_h is little complicated with a free parameter to be adjusted properly. Since the Raviart–Thomas space W^h in this dissertation has a higher order, one can find $p_h \in W^h$ such that div $p_h = c\tilde{u}_h$ holds for $\tilde{u}_h \in V^h$. As a defect of the current setting, the Raviart–Thomas space W^h with a higher order will cause larger matrices in the computation. In (3.16) of §3.3.2, a new quantity $\bar{\kappa}_h$, which can be solved with improved computation efficiency, is proposed to produce a reasonable upper bound of κ_h.

3.3 Guaranteed *a priori* error estimation

3.3.1 *A priori* error estimation

To provide a guaranteed *a priori* error estimation of the FEM solution, we first quote an explicit bound for the constant in the trace theorem. A direct estimation of $C_e(K)$ with FEM approximations is also provided in §4.2.4.

Lemma 3.3.1 ([51]). *Let e be an edge of triangle element K. Define function space*

$$V_e(K) := \{v \in H^1(K) \ \ | \ \ \int_e v \ ds = 0\} .$$

Given $u \in V_e(K)$, we have the following inequality related to the trace theorem:

$$\|u\|_{L^2(e)} \leq C_e(K)|u|_{H^1(K)}, \quad C_e(K) := 0.574\sqrt{\frac{|e|}{|K|}}h_K \leq 0.8118\frac{h_K}{\sqrt{H_K}} . \tag{3.9}$$

Here, H_K denotes the height of triangle K with respect to edge e.

Given an element K of $\mathcal{T}^h$ with e as one of its edges, let $\pi_{0,e}$ be the linear operator that takes the average of a function on edge e. Let I be the identity operator. Note that $\pi_{0,e}v$ is defined over the element K. For function $v \in H^1(\Omega)$, $(I - \pi_{0,e})v|_K$ is regarded as a shift of v, that is,

$$(I - \pi_{0,e})v|_K = v|_K - \frac{1}{|e|}\int_e v \mathrm{d}s \in H^1(K) \ .$$

Since $(I - \pi_{0,e})v|_K$ has zero integral on the boundary edge e, the following error estimation holds:

$$\|(I - \pi_{0,e})v\|_{L^2(e)} \leq C_e(K)|v|_{H^1(K)} \ . \tag{3.10}$$

Let us introduce a piecewise L^2 projection operator $\pi_{h,\Gamma} : L^2(\Gamma) \mapsto X_\Gamma^h$ on the boundary faces: Given $f \in L^2(\Gamma)$, $\pi_{h,\Gamma}f \in X_\Gamma^h$ satisfies

$$b(f - \pi_{h,\Gamma}f, v_h) = 0 \quad \forall v_h \in X_\Gamma^h.$$

It is easy to see that on a boundary edge e of $\mathcal{T}^h$,

$$\int_e (f - \pi_{h,\Gamma}f)|_e \ \pi_{0,e}v \ \mathrm{d}s = 0 \quad \forall v \in H^1(\Omega) \ .$$

Lemma 3.3.2. *Let u and $\tilde{u}$ be solutions to (3.2) and (3.6), respectively, with f_h taken as $f_h := \pi_{h,\Gamma}f$. Then, the following error estimate holds:*

$$\|u - \tilde{u}\|_a \leq C_{e,h}\|(I - \pi_{h,\Gamma})f\|_b, \tag{3.11}$$

where $C_{e,h}$ takes the maximum of $C_e(K)$ over the boundary elements:

$$C_{e,h} := \max_{K \in \mathcal{T}_h^b} C_e(K) = O(h^{1/2}) \, .$$

Proof. Setting $v = u - \tilde{u}$ in (3.2) and (3.6), we have

$$\begin{aligned} a(u-\tilde{u}, u-\tilde{u}) &= b(f - f_h, u - \tilde{u}) = \sum_{e \subset E_{h,\Gamma}} \int_e (I - \pi_{h,\Gamma}) f \cdot (I - \pi_{0,e})(u - \tilde{u}) \mathrm{d}s \\ &\leq \|(I - \pi_{h,\Gamma}) f\|_b \left\{ \sum_{e \in E_{h,\Gamma}} \|(I - \pi_{0,e})(u - \tilde{u})\|^2_{L^2(e)} \right\}^{1/2} . \end{aligned} \tag{3.12}$$

By applying the estimation (3.10), we have

$$\sum_{e \in E_{h,\Gamma}} \|(I - \pi_{0,e})(u - \tilde{u})\|^2_{L^2(e)} \leq \sum_{K \in \mathcal{T}_h^b} C_e(K)^2 |u - \tilde{u}|^2_{H^1(K)} \leq C^2_{e,h} \|u - \tilde{u}\|^2_a \, . \tag{3.13}$$

Note that, the first inequality of the above estimation holds under the assumption that each boundary edge of the triangulation is only shared by one triangle. For a general mesh without such an assumption, the coefficient in the estimation should be doubled. The estimations (3.12) and (3.13) lead to the estimation (3.11). The convergence rate of $C_{e,h}$ as $C_{e,h} = O(h^{1/2})$ for regular meshes is obvious from the estimation (3.9).

Now, we are ready to propose the explicit *a priori* error estimation.

Theorem 3.3.3. *Let u and u_h be solutions to (3.2) and (3.3), respectively. The following error estimates hold.*

$$\|u - u_h\|_a \leq M_h \|f\|_b, \quad \|u - u_h\|_b \leq M_h \|u - u_h\|_a \leq M_h^2 \|f\|_b, \tag{3.14}$$

where $M_h := \sqrt{C_{e,h}^2 + \kappa_h^2}$.

Proof. Take $f_h := \pi_{h,\Gamma} f$ and consider the decomposition $f = f_h + (f - f_h)$. Let $\tilde{u}_h$ be the one defined in Lemma 3.3.2 corresponding to f_h. The minimization principle for the FEM solution u_h tells that $\|u - u_h\|_a \leq \|u - \tilde{u}_h\|_a$. By further applying (3.7) of Lemma 3.2.2 and (3.11) of Lemma 3.3.2, we have

$$\begin{aligned} \|u - u_h\|_a &\leq \|u - \tilde{u}_h\|_a \leq \|u - \tilde{u}\|_a + \|\tilde{u} - \tilde{u}_h\|_a \\ &\leq C_{e,h}\|(I - \pi_{h,\Gamma})f\|_b + \kappa_h \|f_h\|_b \\ &\leq \sqrt{C_{e,h}^2 + \kappa_h^2}\|f\|_b = M_h \|f\|_b \,. \end{aligned}$$

The error estimate (3.14) can be obtained by applying the standard Aubin–Nitsche duality technique.

Remark 3.3.4. The analysis of $C_{e,h}$ tells that $C_{e,h} = O(h^{1/2})$, and numerical results in Chapter 4, §4.2.1 imply that κ_h has the convergence rate as $O(h^{1/2})$ even for convex domains and high-order FEM spaces. Hence, the proposed *a priori* error estimation with the quantity M_h has the convergence rate as $O(h^{1/2})$, which will lead to a lower eigenvalue bound given by (3.21) with a degenerated convergence rate as $O(h)$. From classical discussions of the solution regularity of Neumann boundary condition, it is known that the solution has the regularity as $u \in H^{1+r}(\Omega)$ for a general $f \in L^2(\partial\Omega)$ with $r \in [0, 1/2)$; see, e.g., [43, Theorem 4] and [16, Theorem 31.34]. Therefore, such a convergence rate of M_h is reasonable, as the *a priori* error estimation has to manipulate the worst case of the solution regularity. Meanwhile, the FEM approximations of the leading eigenvalues over the unit square domain demonstrate the $O(h^2)$ convergence rate (see the discussion in Chapter 4). Thus, as

the defect of the proposed lower eigenvalue bounds in this dissertation, the estimation (3.21) using $M_h = O(h^{1/2})$ is sub-optimal for smooth eigenfunctions.

Remark 3.3.5. It is worth pointing out that Theorem 3.3.3 is also available for general $\mathbb{R}^n$ $(n \geq 2)$ spaces by providing explicit values for the involved quantities. The value of κ_h can be computed by using the hypercircle for standard FEM spaces on $\mathbb{R}^n$ domain. For the constant $C_{e,h}$ appearing in Lemma 3.3.1, the method used in [51] to evaluate $C_{e,h}$ can be easily extended to a $\mathbb{R}^n$ simplex; see such a discussion in, e.g., the corrigendum of [2, Lemma 1].

3.3.2 Computation of κ_h

This subsection is dedicated to a description of the algorithm to evaluate κ_h defined in (3.5).

First, for a fixed $f_h \in X_\Gamma^h$, we consider the following minimization problem:

$$\min_{u_h \in V^h} \min_{\substack{p_h \in W_{f_h}^h \\ \operatorname{div} p_h = c u_h}} \|\nabla u_h - p_h\|_{L^2}^2 \ .$$

The above problem is reformulated as finding the stationary point for the following objective function: for $(u_h, p_h, x_h) \in V^h \times W_{f_h}^h \times X^h$,

$$\mathcal{F}(u_h, p_h, x_h) := \frac{\|\nabla u_h - p_h\|_{L^2}^2}{2} + (x_h, \operatorname{div} p_h - c u_h).$$

Then, stationary point (u_h, p_h, x_h) satisfies

$$\begin{cases} (\nabla u_h, \nabla v_h) & -(p_h, \nabla v_h) & -c(x_h, v_h) & = 0 \\ -(\nabla u_h, q_h) & +(p_h, q_h) & +(x_h, \operatorname{div} q_h) & = 0 \\ -c(u_h, y_h) & +(\operatorname{div} p_h, y_h) & & = 0 \end{cases} \tag{3.15}$$

for all $(v_h, q_h, y_h) \in V^h \times W_0^h \times X^h$.

To confirm the existence and uniqueness of (u_h, p_h, x_h) of the system (3.15), we cite the following result from [8]. Note that the notation below is restricted to the discussion of Proposition 3.3.6 in the rest of current subsection.

Proposition 3.3.6 (Proposition 1.1 of [8], p.38)**.** *Let V and Q be Hilbert spaces, the dual spaces of which are denoted by V' and Q', respectively. Let $B : V \to Q'$ be a linear operator. Let $g \in \operatorname{Im}(B)$ and let the bilinear form $a(\cdot, \cdot)$ be coercive on $\operatorname{Ker}(B)$, that is, there exists α_0 such that*

$$a(v_0, v_0) \geq \alpha_0 \|v_0\|^2 \quad \forall v_0 \in \operatorname{Ker}(B).$$

Then, given $f \in V'$, there exists a unique $u \in V$ solution of the equations:

$$Bu = g; \quad a(u, v_0) = \langle f, v_0 \rangle_{V' \times V} \quad \forall v_0 \in \operatorname{Ker}(B) \ .$$

To apply Proposition 3.3.6, we consider a reformulation of (3.15). Let $\hat{p}_h$ be a fixed function of $W_{f_h}^h$ and introduce $p_{h,0} := p_h - \hat{p}_h \in W_0^h$. The equations in (3.15)

becomes

$$\left\{\begin{array}{llll} (\nabla u_h, \nabla v_h) & -(p_{h,0}, \nabla v_h) & -c(x_h, v_h) & = (\hat{p}_h, \nabla v_h) \\ -(\nabla u_h, q_h) & +(p_{h,0}, q_h) & +(x_h, \mathrm{div}\, q_h) & = -(\hat{p}_h, q_h) \\ -c(u_h, y_h) & +(\mathrm{div}\, p_{h,0}, y_h) & & = -(\mathrm{div}\, \hat{p}_h, y_h) \end{array}\right. .$$

Let us consider the following function settings.

$$\begin{aligned} & V := V^h \times W_0^h, \quad Q := X^h, \\ & \langle f, \{v_h, q_h\}\rangle_{V' \times V} := (\hat{p}_h, \nabla v_h - q_h)_\Omega, \quad \langle g, \cdot\rangle_{Q' \times Q} := (-\mathrm{div}\, \hat{p}_h, \cdot)_\Omega \ , \\ & a(\{u_h, p_{h,0}\}, \{v_h, q_h\}) := (\nabla u_h - p_{h,0}, \nabla v_h - q_h)_\Omega, \\ & \langle B(\{u_h, p_{h,0}\}), \cdot\rangle_{Q' \times Q} := (\mathrm{div}\, p_{h,0} - c u_h, \cdot)_\Omega \ . \end{aligned}$$

The inner product of V is defined by

$$\langle \{u_h, p_h\}, \{v_h, q_h\}\rangle_V := (\nabla u_h, \nabla v_h) + c(u_h, v_h) + (p_h, q_h) + (\mathrm{div}\ p_h, \mathrm{div}\ q_h) \ ,$$

which induces the norm as $\|\{u_h, p_h\}\|_V = \{\|\nabla u_h\|_\Omega^2 + c\|u_h\|_\Omega^2 + \|p_h\|_{H(\mathrm{div})}^2\}^{\frac{1}{2}}$. Since the involved spaces are finite dimensional, $\mathrm{Im}(B)$ is the closed subspace of $V^h \times W_0^h$. The positive-definiteness and boundedness of $a(\cdot, \cdot)$ are easy to confirm.

The coercivity of $a(\cdot, \cdot)$ over $\mathrm{Ker}(B)$ can be confirmed by the following equality:

for $\{u_h, p_{h,0}\} \in \mathrm{Ker}(B)$, by applying Green's formula,

$$\begin{aligned} a(\{u_h, p_{h,0}\}, \{u_h, p_{h,0}\}) &= \|\nabla u_h\|^2 - 2(\nabla u_h, p_{h,0}) + \|p_{h,0}\|^2 \\ &= \|\nabla u_h\|^2 + 2c\|u_h\|^2 + \|p_{h,0}\|^2 \\ &= \|\nabla u_h\|^2 + c\|u_h\|^2 + \|\mathrm{div}\ p_{h,0}\|^2 + \|p_{h,0}\|^2 \\ &= \|\{u_h, p_{h,0}\}\|_V^2. \end{aligned}$$

Therefore, Proposition 3.3.6 makes certain that the functional $\mathcal{F}$ has a unique saddle point $(u_h, p_{h,0} + \hat{p}_h, x_h)$ in $V^h \times W^h_{f_h} \times X^h$, giving a solution to the problem. The evaluation of κ_h can be done by further considering the maximization of $\|\nabla u_h - p_h\|^2_{L^2} / \|f_h\|_b{}^2$ for all $f_h \in X^h_\Gamma$.

In the practical computation, we propose an efficient way that provides an upper bound for κ_h. Given an $f_h \in X^h_\Gamma$, let us consider the following formulation that determines $\tilde{u}_h \in V^h$ and $p_h \in W^h_{f_h}$ subsequently.

(a) Find $\tilde{u}_h \in V^h$ s.t.

$$a(\tilde{u}_h, v_h) = b(f_h, v_h) \quad \forall v_h \in V^h.$$

(b) Let $\tilde{u}_h$ be the solution of (a). Find $p_h \in W^h_{f_h}$ and $\rho_h \in X^h$, $r \in \mathbb{R}$ s.t.

$$\begin{cases} (p_h, q_h) + (\rho_h, \mathrm{div}\ q_h) + (\rho_h, s) = 0 & \forall q_h \in W^h_0,\ \forall s \in \mathbb{R} \\ (\mathrm{div}\ p_h, \eta_h) + (r, \eta_h) = c(\tilde{u}_h, \eta_h) & \forall \eta_h \in X^h \end{cases}.$$

For each given f_h, there exist unique solution $\tilde{u}_h$ and p_h to the sub-problems (a)

and (b). By using the mapping from f_h to $\tilde{u}_h$ and p_h, let us introduce the quantity $\bar{\kappa}_h$, which works as an upper bound of κ_h:

$$\bar{\kappa}_h := \max_{f_h \in X_\Gamma^h \setminus \{0\}} \frac{\|\nabla \tilde{u}_h - p_h\|_0}{\|f_h\|_b}. \tag{3.16}$$

According to the definition of $\bar{\kappa}_h$, it is required to find f_h that maximizes the value of $\|\nabla \tilde{u}_h - p_h\|_0 / \|f_h\|_b$, which can be achieved by solving an eigenvalue problem for matrices. Since $\tilde{u}_h \in V^h$ and $p_h \in W_{f_h}^h$ are determined subsequently, the matrices involved in setting up the linear system will have a quite smaller size than the ones in solving (3.3.2). For a detailed description of the evaluation of κ_h and $\tilde{\kappa}_h$, refer to ([33]), where an analogous problem is considered.

Remark 3.3.7. The introduction of variable r in the setting of problem (b) is to make certain a regular matrix in solving the linear systems. By setting $v_h = 1$ in the problem (a), we have

$$c \int_\Omega \tilde{u}_h \, \mathrm{d}x = \int_{\partial\Omega} f_h ds = \int_{\partial\Omega} p_h \cdot \mathbf{n} \, \mathrm{d}s = \int_\Omega \mathrm{div}\ p_h \, \mathrm{d}x \, .$$

The above relation implies that $(\mathrm{div}\ p_h - c\tilde{u}_h, \cdot)$ has a kernel space with constant function.

3.4 Application to the Steklov eigenvalue problem

As an application of the error estimation (3.14), we are concerned with the following model Steklov eigenvalue problem:

$$-\Delta u + cu = 0 \quad \text{in } \Omega \, ; \quad \frac{\partial u}{\partial \mathbf{n}} = \lambda u \quad \text{on } \Gamma = \partial\Omega \, , \tag{3.17}$$

In case $c = 0$, the eigenvalue problem (3.17) has the zero eigenvalue and the eigenfunctions associated to the non-zero eigenvalues have zero integral on the boundary of the domain.

A weak formulation of the above problem is as follows: Find $\lambda \in \mathbb{R}$ and $u \in V$ such that $\|u\|_b = 1$ and

$$a(u, v) = \lambda b(u, v) \quad \forall v \in V \, . \tag{3.18}$$

Let us consider the operator $\mathcal{D}^{-1} : L^2(\Gamma) \to V$ such that for $f \in L^2(\Gamma)$, $\mathcal{D}^{-1} f = u$ satisfies the variational equation

$$a(\mathcal{D}^{-1} f, v) = b(f, v) \quad \forall v \in V.$$

As a compatibility condition for the definition of $\mathcal{D}^{-1}$, it is required that $\int_\Gamma f = 0$ in case $c = 0$. Let γ be the trace operator $\gamma : V \to L^2(\Gamma)$. Under the current assumption that the domain has a polygonal boundary, $\mathcal{D}^{-1} \circ \gamma : V \to V$ is a compact operator [10]. The operator $\mathcal{D}^{-1} \circ \gamma$ has the zero eigenvalue, for which the associated eigenspace is just $H_0^1(\Omega)$. The rest eigenvalues of $\mathcal{D}^{-1} \circ \gamma$ form a sequence $\{\mu_k\}$ as follows:

$$\mu_k > 0, \quad \mu_1 \geq \mu_2 \geq \cdots , \quad \lim_{k \to \infty} \mu_k = 0.$$

The trace operator γ will be omitted if there is no ambiguity. The weak formulation of the eigenvalue problem for $\mathcal{D}^{-1} \circ \gamma$ is given by: Find $u \in V$ and $\mu \geq 0$ such that,

$$b(u, v) = \mu a(u, v) \quad \forall v \in V \, . \tag{3.19}$$

The eigenfunctions of (3.19) form a complete orthonormal basis of V.

As for the relation between the eigenvalue problem of $\mathcal{D}^{-1} \circ \gamma$ and the one defined in (3.18), we have that the non-zero eigenvalues μ_k's are given by the reverse of λ_k, i.e., $\mu_k = 1/\lambda_k$.

From the above argument, the eigenvalue problem (3.18) has an eigenvalue sequence $\{\lambda_k\}$:

$$0 < \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_k \leq \cdots , \quad \lim_{k \to \infty} \lambda_k = \infty \ .$$

Finite element approximation The conforming finite element approximation of (3.18) is defined as follows: Find $\lambda_h (> 0) \in \mathbb{R}$ and $u_h \in V^h$ such that $\|u_h\|_b = 1$ and

$$a(u_h, v_h) \quad = \quad \lambda_h b(u_h, v_h) \quad \forall v_h \in V^h. \tag{3.20}$$

Let $n := \dim(V^h)$ and $n_0 := n - \dim(V^h \cap H_0^1(\Omega))$. The eigenvalue problem (3.20) has n_0 positive eigenvalues

$$0 < \lambda_{1,h} \leq \lambda_{2,h} \leq \cdots \leq \lambda_{n_0,h} < \infty \quad (n_0 \leq n) \ .$$

Define the projection $P_h : V \to V^h$ by

$$a(u - P_h u, v_h) = 0 \quad \forall v_h \in V^h \ .$$

Below is the result from [51] that provides lower eigenvalue bounds.

Theorem 3.4.1. *Suppose the following inequality holds for the projection error:*

$$\|(I - P_h)u\|_b \leq M_h \|(I - P_h)u\|_a \quad \forall u \in V \ .$$

Let $\lambda_{k,h}$ be the k-th eigenvalue of (3.20)*. A lower bound of the eigenvalue λ_k of* (3.18) *is given by*

$$\lambda_k \geq \frac{\lambda_{k,h}}{1 + M_h^2 \lambda_{k,h}}, \quad k = 1, \cdots, n_0. \tag{3.21}$$

Chapter 4

Numerical Experiments

4.1 Guaranteed local error estimation

4.1.1 Preparation

The selection of bandwidth of the B_S is important in the local error estimation. A large bandwidth of B_S leads to a large value of E_1, while a small bandwidth of B_S results in a large value of $\|\nabla\alpha\|_{L^\infty(\Omega)}$ in E_2. Therefore, in each example, we first investigate the impact of the bandwidth of B_S, and then take an appropriate width of B_S for subsequent computation.

Besides the symbols E_1, E_2 in (2.31), we introduce new symbols as follows:

- The local error and its estimation are denoted by

$$E_L := \|\nabla u - \nabla u_h\|_S, \quad \widehat{E}_L := \sqrt{E_1^2 + E_2^2} + 2\,\mathrm{Osc}(f)\,.$$

- The global error and its estimation in (2.32) are denoted by

$$E_G := \|\nabla u - \nabla u_h\|_\Omega, \quad \widehat{E}_G := \|\nabla u_h - p_h\|_\Omega + C_0 h \|f - \pi_h f\|.$$

Here, $u_h \in V_h$ and $p_h \in RT_h$ are finite element solutions of the objective problems; p_h also satisfies the condition (2.25).

4.1.2 Square domain

The error estimation proposed in this dissertation is applicable to problems with different boundary conditions. To illustrate this feature, let us consider the following Poisson equations over the unit square domain $\Omega = (0,1)^2$, where the subdomain is selected as $S = (0.375, 0.625)^2$.

(a) Dirichlet boundary condition (exact solution $u = \sin(\pi x)\sin(\pi y)$).

$$-\Delta u = 2\pi^2 \sin(\pi x)\sin(\pi y) \text{ in } \Omega, \quad u = 0 \text{ on } \partial\Omega. \tag{4.1}$$

(b) Neumann boundary condition (exact solution $u = \cos(\pi x)\cos(\pi y)$).

$$-\Delta u = 2\pi^2 \cos(\pi x)\cos(\pi y) \text{ in } \Omega, \quad \frac{\partial u}{\partial \mathbf{n}} = 0 \text{ on } \partial\Omega, \quad \int_\Omega u dx = 0\,. \tag{4.2}$$

The finite element solutions u_h, p_h are computed with uniform meshes, and the mesh size h here is chosen as the leg length of the triangle element for a uniform mesh.

Asymptotic behavior of the proposed local error estimator over a uniform mesh. For Dirichlet and Neumann boundary conditions, the dependencies of the

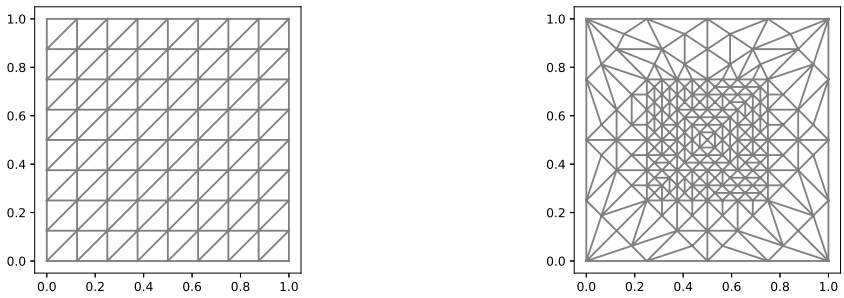


Figure 4.1: Uniform and non-uniform mesh (rectangle domain).

local error estimator $\widehat{E}_L$ on the bandwidth of B_S are shown in Figure 4.2 and Figure 4.3, respectively. The relative variation of the local error estimator with respect to bandwidth selection is displayed for two problems. It is noteworthy that the local error estimation is not significantly sensitive to variations in bandwidth. For example, in Figure 4.2, for $h = 1/64$, the relative variation in error estimation with respect to a bandwidth in the range $[0.075, 0.125]$ is less than 20%.

In the following discussion, the bandwidth of B_S is selected as 0.1 for the Dirichlet boundary condition and 0.075 for the Neumann boundary condition.

Table 4.1: Error estimate for Dirichlet BVP (square domain, uniform mesh)

h	κ_h	$C(h)$	E_L	E_1	E_2	$\widehat{E}_L$	$\widehat{E}_G$
1/16	0.030	0.036	0.060	0.090	0.255	0.300	0.264
1/32	0.015	0.018	0.030	0.041	0.065	0.084	0.129
1/64	0.008	0.009	0.015	0.020	0.016	0.027	0.064
1/128	0.004	0.005	0.007	0.010	0.004	0.011	0.032
1/256	0.002	0.002	0.004	0.005	0.001	0.005	0.016

A detailed discussion on each component of the error estimators is also presented;

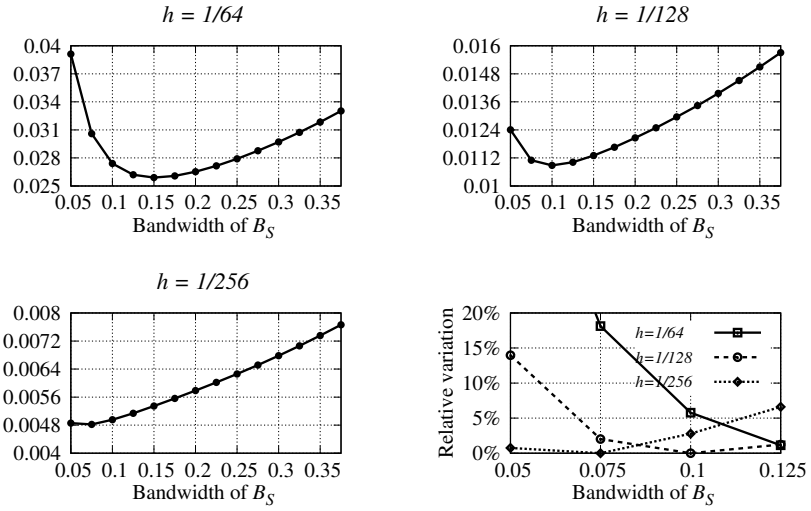


Figure 4.2: Dependency of local error estimation on the bandwidth of B_S (Dirichlet BVP, square domain, uniform mesh).

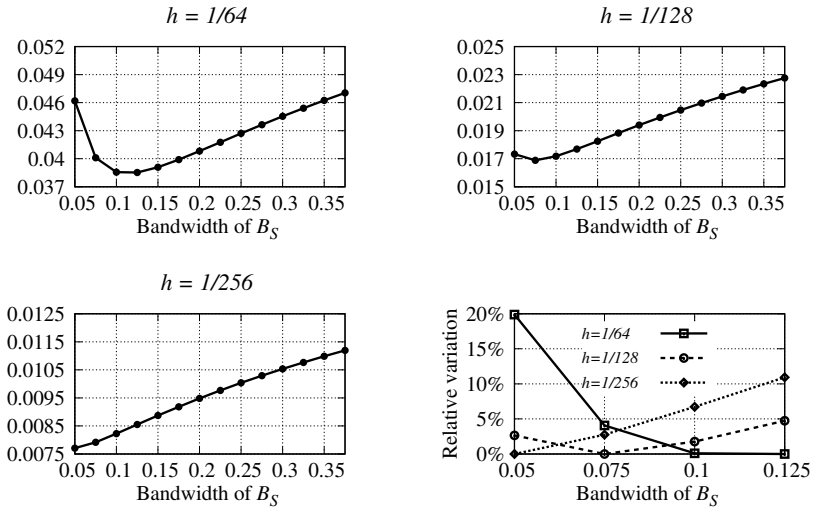


Figure 4.3: Dependency of local error estimation on the bandwidth of B_S (Neumann BVP, square domain, uniform mesh).

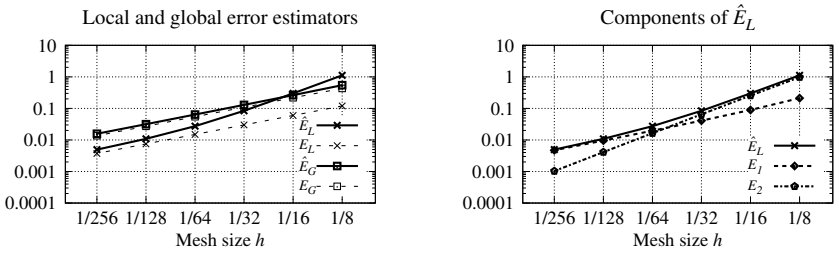


Figure 4.4: Error estimators for Dirichlet BVP (square domain, uniform mesh).

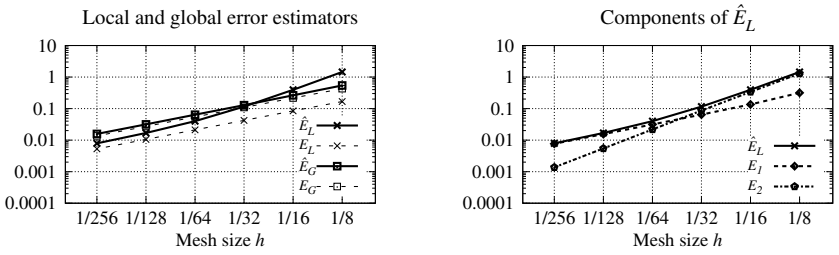


Figure 4.5: Error estimators for Neumann BVP (square domain, uniform mesh).

see Table 4.1 and Figure 4.4 for Dirichlet boundary condition, Table 4.2 and Figure 4.5 for Neumann boundary condition. From the numerical results, we confirm that for both the problems the main term E_1 of the error estimation (2.31) becomes dominant when $h \leq 1/64$, which agrees with the analysis in §2.3.4.

Convergence behavior for non-uniform meshes. Based on numerical results, we investigate the behavior of our proposed estimator (2.31) for non-uniform meshes under the setting $h_G = O(\sqrt{h_{\Omega'}})$; see a sample non-uniform mesh in Figure 4.1. The subdomain S and the bandwidth B_S are set to $(0.375, 0.625)^2$ and 0.125, respectively.

The numerical results in Figure 4.6 and Table 4.3 show that the convergence rates of $\|\nabla u_h - p_h\|_\alpha$ and E_1 are almost $O(h_{\Omega'})$. The numerical results support the

Table 4.2: Error estimate for Neumann BVP (square domain, uniform mesh).

h	κ_h	$C(h)$	E_L	E_1	E_2	$\widehat{E}_L$	$\widehat{E}_G$
1/16	0.030	0.036	0.084	0.135	0.339	0.394	0.263
1/32	0.015	0.018	0.042	0.065	0.086	0.115	0.129
1/64	0.008	0.009	0.021	0.031	0.022	0.040	0.064
1/128	0.004	0.005	0.011	0.015	0.005	0.017	0.032
1/256	0.002	0.002	0.005	0.008	0.001	0.008	0.016

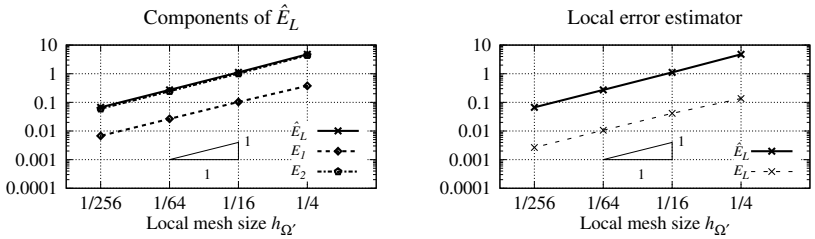


Figure 4.6: Local error estimator for lowest order FEM over non-uniform mesh (squared domain).

Table 4.3: Convergence rate of $\|\nabla u_h - p_h\|_\alpha$ over non-uniform mesh (squared domain).

$h_{\Omega'}$	$\|\nabla u_h - p_h\|_\alpha$	Order
0.177	0.132	-
0.044	0.033	0.996
0.011	0.008	1.022
0.003	0.002	1.067

expectation that $\|\nabla u_h - p_h\|_\alpha = O(h_{\Omega'})$ under current mesh configuration (i.e., $h_G = \sqrt{h_{\Omega'}}$). In case that the lowest degree Raviart–Thomas FEM is employed to compute the global term E_2, the convergence rate of the estimator $\widehat{E}_L$ is approximately $O(h_{\Omega'})$. The theoretical convergence rate of $\|\nabla u_h - p_h\|_\alpha$ will be considered in our succeeding

research.

As a conclusion, our proposed local error estimator is dominated by the local error term $E_1 = O(h_{\Omega'})$. Thus, it is possible to increase the efficiency of computation by using a non-uniform mesh with a raw triangulation for the subdomain outside of the part of interest.

4.1.3 L-shaped domain

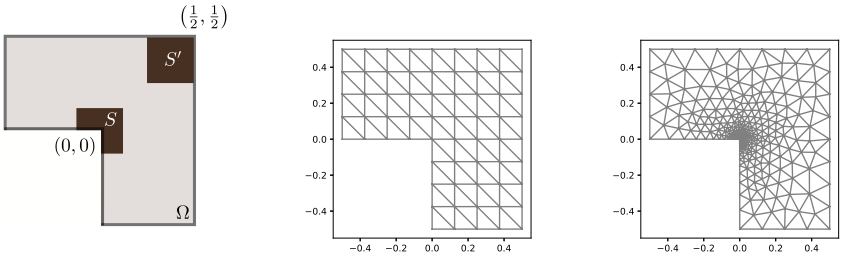


Figure 4.7: Uniform and non-uniform mesh for an L-shaped domain Ω with two subdomains S, S'.

The proposed error estimation (2.31) is applicable to problems with a singular solution and the even case in which the subdomain S and Ω share a common part of the boundary. In this sub-section, we consider the boundary value problem over an L-shaped domain $\Omega := (-0.5, 0.5)^2 \setminus [-0.5, 0]^2$; see Figure 4.7. The error estimation on two subdomains $S = \Omega \cap (-0.125, 0.125)^2$ and $S' = (0.25,\ 0.5)^2$ will be considered.

Let $u = r^{\frac{2}{3}} \sin\left(\frac{2}{3}(\theta + \frac{\pi}{2})\right) \cos(\pi x) \cos(\pi y)$, where r and θ are the variables under the polar coordinates. Define $f = -\Delta u$. Then u is the solution of the following

equation.

$$-\Delta u = f \text{ in } \Omega, \quad u = 0 \text{ on } \partial\Omega. \tag{4.3}$$

It is easy to confirm that $u \notin H^2(\Omega)$ due to the singularity around the re-entry corner point of the domain.

Selection of the bandwidth of B_S. The dependency of $\widehat{E}_L$ on the bandwidth of the B_S is shown in Figure 4.8. In the following computation, the bandwidth of $B_S, B_{S'}$ is selected as 0.225, 0.25, respectively.

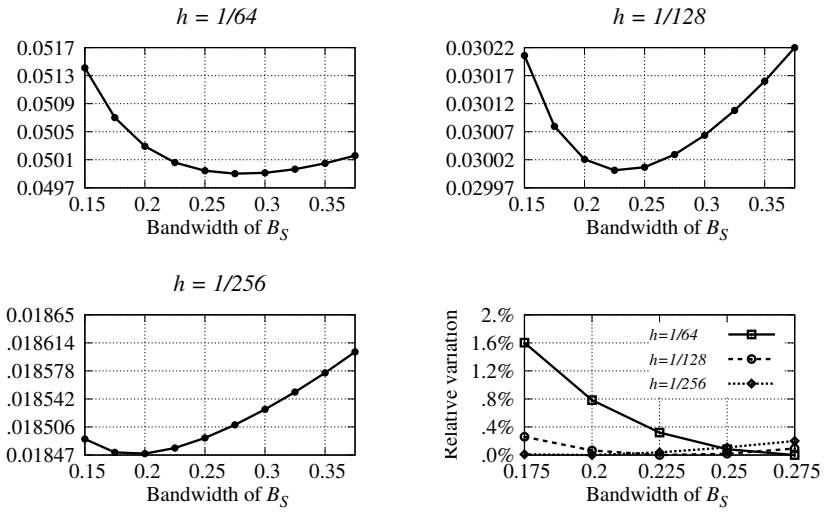


Figure 4.8: Dependency of local error estimation on the bandwidth of B_S (L-shaped domain).

For subdomain S, the asymptotic behavior of $\widehat{E}_L$ with respect to mesh size h is shown in Table 4.4 and Figure 4.9. The numerical results tell that the local error component E_1 in $\widehat{E}_L$ gradually becomes dominant as the mesh is refined.

Table 4.4: Error estimators for subdomain S (uniform mesh of L-shaped domain).

h	κ_h	$C(h)$	E_L	E_1	E_2	$\widehat{E}_L$	$\widehat{E}_G$
1/16	0.046	0.050	0.080	0.129	0.102	0.184	0.172
1/32	0.028	0.029	0.050	0.077	0.034	0.089	0.095
1/64	0.017	0.018	0.032	0.047	0.012	0.050	0.055
1/128	0.011	0.011	0.020	0.029	0.004	0.030	0.032
1/256	0.007	0.006	0.013	0.018	0.002	0.018	0.020

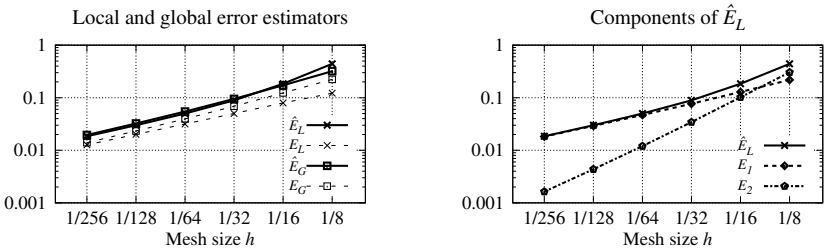


Figure 4.9: Error estimators for subdomain S (uniform mesh of L-shaped domain).

We also compare E_L, $\widehat{E}_L$ with E_G, $\widehat{E}_G$ in Table 4.5. Denote $\beta := E_L/E_G$ and $\widehat{\beta} := \widehat{E}_L/\widehat{E}_G$. It is observed that the approximation error concentrates in the subdomain S around the re-entry corner as the mesh is refined. For $h = 1/256$, the local error in S is about 91% of the global error in the whole domain. Finally, we consider the local error estimation for a non-uniform mesh; see computation results Table 4.6 and Figure 4.10. It is observed that for the subdomain S, both β and $\widehat{\beta}$ become smaller compared to the results in the case of uniform meshes, which implies that a denser mesh around the re-entry corner improves the quality of local approximation.

Remark 4.1.1. The computation codes and results in this section are available on the following website.

`https://ganjin.online/nakano/Guaranteed_local_error_estimation`

Table 4.5: Comparison of the estimated local error on S and S'.

	subdomain S		subdomain S'			
h	$\beta(\%)$	$\widehat{\beta}(\%)$	E_L	$\widehat{E}_L$	$\beta(\%)$	$\widehat{\beta}(\%)$
1/16	64	105	0.041	0.134	33	78
1/32	73	93	0.021	0.050	30	52
1/64	80	91	0.010	0.020	26	38
1/128	86	92	0.005	0.009	22	28
1/256	91	94	0.003	0.004	18	21

Table 4.6: Error estimators for subdomain S (non-uniform mesh of L-shaped domain)

h	κ_h	$C(h)$	E_L	E_1	E_2	$\widehat{E}_L$	$\widehat{E}_G$	$\beta(\%)$	$\widehat{\beta}(\%)$
0.141	0.039	0.054	0.023	0.048	0.182	0.213	0.166	21	129
0.081	0.021	0.030	0.012	0.023	0.051	0.063	0.081	21	78
0.041	0.011	0.015	0.007	0.012	0.013	0.020	0.040	23	49
0.020	0.006	0.008	0.004	0.006	0.003	0.008	0.020	26	39
0.010	0.003	0.004	0.002	0.004	0.001	0.004	0.010	30	38

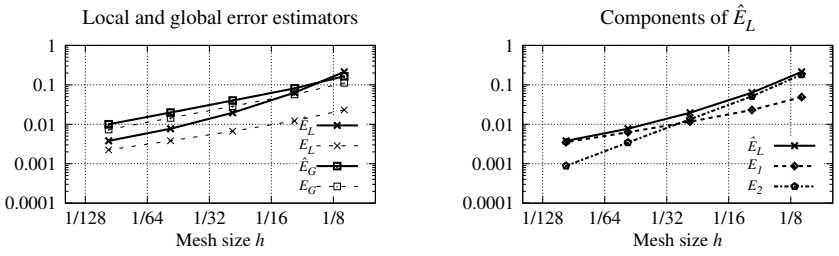


Figure 4.10: Error estimators for subdomain S (non-uniform mesh of L-shaped domain).

4.2 Numerical experiments for the Steklov eigenvalue estimation

In this section, we apply the eigenvalue estimation (3.21) along with the explicit *a priori* error estimation solve the eigenvalue problem (3.17) on both the unit square domain $\Omega = (0,1) \times (0,1)$ and the L-shaped domain $\Omega = (0,2) \times (0,2) \setminus [1,2] \times [1,2]$. Here, we select c appearing in (3.17) as 1. Also, the existing method of [51] based on the nonconforming FEM is utilized to compare the efficiency with each other.

4.2.1 Evaluation of κ_h and $\bar{\kappa}_h$

We adopt two different methods in subsection 3.3.2 to evaluate κ_h and $\bar{\kappa}_h$ and display the computation results in Table 4.7-4.8. It is observed that the $\bar{\kappa}_h$ gives very close upper bound of κ_h; for the square domain, the leading 4 significant digits of $\bar{\kappa}_h$ and κ_h are the same to each other. Thus, $\bar{\kappa}_h$ will be utilized instead of κ_h in the following computation examples. It is worth to point out that the value of κ_h has a convergence rate, denoted by $\gamma(\kappa_h)$ in the tables, as $O(h^{1/2})$ for both the square domain and the L-shaped domain. To confirm the dependency of the convergence rate of κ_h on the order of FEM spaces, the hypercircle using FEM spaces (i.e., $V^h, W^h, X^h, X^h_\Gamma$) of order 2 is used to evaluate κ_h, denoted by $\kappa_{h,2}$, is also displayed in Table 4.7. Numerical results tell that $\gamma(\kappa_{h,2})$ is still 0.5.

It is of great interest when the worst case of the projection error happens. To confirm for which f_h the value of κ_h is reached, we draw the figures of such an f_h and its corresponding conforming FEM solution u_h. Since f_h is defined on the boundary

Table 4.7: Quantities $\kappa_h, \bar{\kappa}_h$ and $\kappa_{h,2}$ for the unit square domain (γ: convergence rate)

h	$\sqrt{2}/4$	$\sqrt{2}/8$	$\sqrt{2}/16$	$\sqrt{2}/32$
κ_h	0.2891	0.2042	0.1443	0.1021
$\gamma(\kappa_h)$	-	0.50	0.50	0.50
$\bar{\kappa}_h$	0.2891	0.2042	0.1443	0.1021
$\gamma(\bar{\kappa}_h)$	-	0.50	0.50	0.50
$\kappa_{h,2}$	0.2291	0.1621	0.1146	0.0811
$\gamma(\kappa_{h,2})$	-	0.50	0.50	0.50

Table 4.8: Quantities κ_h and $\bar{\kappa}_h$ for the L-shaped domain domain (γ: convergence rate)

h	$\sqrt{2}/2$	$\sqrt{2}/4$	$\sqrt{2}/8$	$\sqrt{2}/16$
κ_h	0.5075	0.3624	0.2588	0.1846
$\gamma(\kappa_h)$	-	0.49	0.49	0.49
$\bar{\kappa}_h$	0.5106	0.3633	0.2591	0.1847
$\gamma(\bar{\kappa}_h)$	-	0.49	0.49	0.49

of the domain, let us introduce a parameter L to measure the arc length from the vertex located at the origin point; see Figure 4.11. The graphs of f_h and the contour lines of u_h for the square domain and the L-shaped domain are displayed in Figure 4.12 and 4.13, respectively. Note that f_h is normalized by the L^∞ norm in each figure. The numerical results imply that when the value of f is concentrated at the corner of the domain, the worst case of the projection error happens. For the square domain, there is large variation of both f_h and the conforming FEM solution u_h around the four corners, while for the L-shaped domain, the variation of both f_h and u_h is concentrated at the re-entry corner. A theoretical investigation of the worst cases for the Neumann boundary conditions is of interest and will be considered in future work.

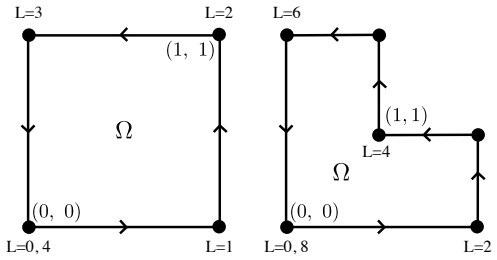


Figure 4.11: Parameter L for the arc length of domain boundary

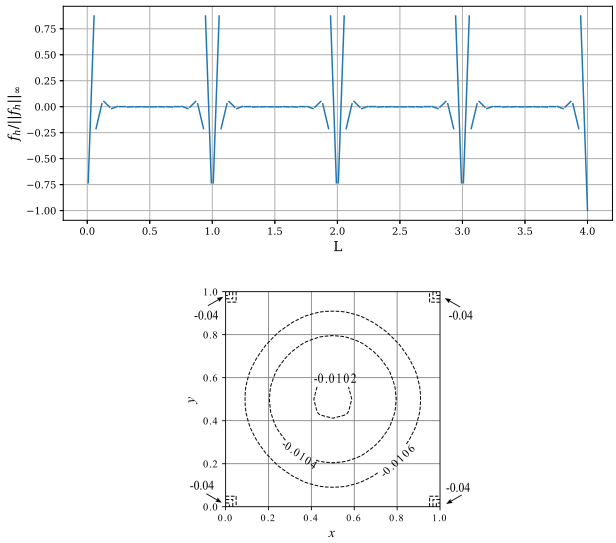


Figure 4.12: The worst f_h (left) and u_h (right) that determine κ_h (square domain)

4.2.2 Preparation for eigenvalue estimation

The explicit values of the exact eigenvalues for both domains are not available. For the unit square domain, the following high-precision estimation with reliable significant digits are used as a nice approximation to true eigenvalues ([50]).

$$\text{(unit square) } \lambda_1 \approx 0.240079, \quad \lambda_2 = \lambda_3 \approx 1.49230\,.$$

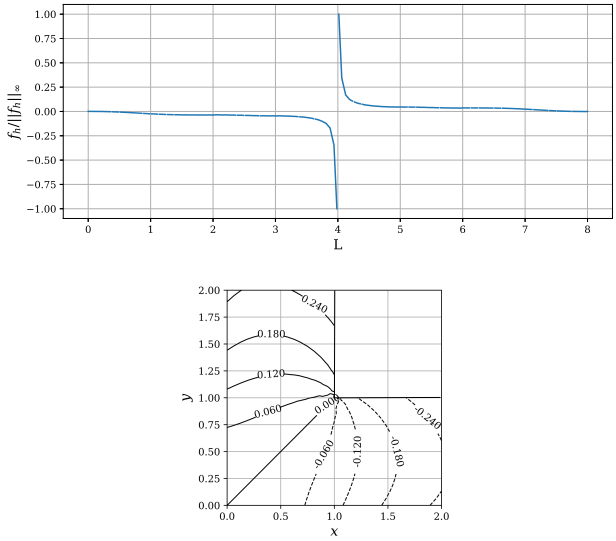


Figure 4.13: The worst f_h (left) and u_h (right) that determine κ_h (L-shaped domain)

In case of the L-shaped domain, the cubic conforming FEM with the mesh size $h = \sqrt{2}/256$ provides a high-precision approximation to eigenvalues:

$$\text{(L-shaped domain) } \lambda_1 \approx 0.3414160, \quad \lambda_2 \approx 0.6168667, \quad \lambda_3 \approx 0.9842784\,.$$

For both domains, the uniform meshes are adopted. The eigenvalue estimation

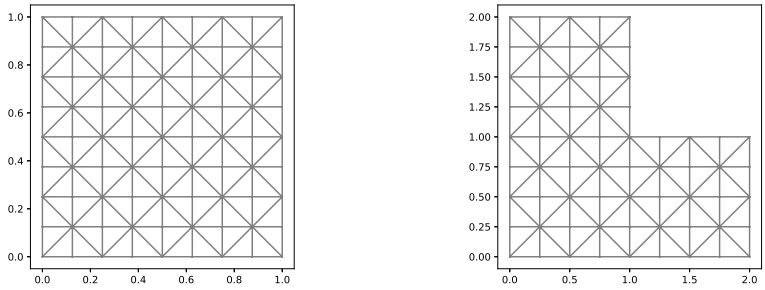


Figure 4.14: The unit square and L-shaped domains

(3.21) provides a guaranteed lower eigenvalue bound:

$$\underline{\lambda}_{k,h} := \frac{\lambda_{k,h}}{1 + M_h^2 \lambda_{k,h}}, \qquad M_h = \sqrt{C_{e,h}^2 + \kappa_h^2}\ , \tag{4.4}$$

where $\lambda_{k,h}$ denotes the k-th approximate eigenvalue from the conforming FEM and the quantity $C_{e,h}$ in estimating M_h is given by

$$C_{e,h} := 0.8118 \max_{K \in \mathcal{T}_h^b} \frac{h_K}{\sqrt{H_K}} (= 0.9654\sqrt{h_K})\ .$$

Note that $h_K = \sqrt{2} H_K$. The eigenvalue estimation from Theorem 3.8 of [51] has the formula as follows.

$$\underline{\lambda}_{k,h}^{\rm nc} := \frac{\lambda_{k,h}^{\rm nc}}{1 + \widehat{C}_h^2 \lambda_{k,h}^{\rm nc}}\ , \tag{4.5}$$

where $\lambda_{k,h}^{\rm nc}$ denotes the k-th approximate eigenvalue from the Crouzeix-Raviart FEM. Particularly, for the uniform mesh used here, $\widehat{C}_{e,h}$ is estimated by

$$\begin{aligned} \widehat{C}_{e,h} &= 0.6711 \max_{K \in \mathcal{T}_h^b} \frac{h_K}{\sqrt{H_K}} + \frac{0.1893}{\sqrt{\lambda_{1,h}^{\rm nc}}} \max_{K \in \mathcal{T}_h} h_K \\ &= 0.7981\sqrt{h_K} + \frac{0.1893}{\sqrt{\lambda_{1,h}^{\rm nc}}} h_K\ . \end{aligned}$$

4.2.3 Computation results for two domains

Sample uniform triangular meshes for two domains are displayed in Figure 4.14, where the mesh size for the unit square is $h = \sqrt{2}/8$ and the one for the L-shaped domain is $h = \sqrt{2}/4$.

For the unit square domain, the eigenvalue estimations (3.21) for the leading 3

eigenvalues are displayed in Table 4.9, while the results based on the nonconforming FEM ([51]) are displayed in Table 4.10. The results for the L-shaped domain are displayed in Table 4.11 and 4.12. Figure 4.15 and Figure 4.16 describe the relation between the absolute errors and the degrees of freedom (DOF) over the unit square and L-shaped domains, respectively. Here, the DOF of (3.21) is counted as the the dimension of the linear conforming FEM space V^h, while the one for [51] is the dimension of the Crouzeix-Raviart FEM space.

Let us also introduce the total errors by

$$\begin{aligned} \text{Error-(4.4)} &:= |\lambda_1 - \underline{\lambda}_{1,h}| + |\lambda_2 - \underline{\lambda}_{2,h}| + |\lambda_3 - \underline{\lambda}_{3,h}| \ , \\ \text{Error-(4.5)} &:= |\lambda_1 - \underline{\lambda}^{\rm nc}_{1,h}| + |\lambda_2 - \underline{\lambda}^{\rm nc}_{2,h}| + |\lambda_3 - \underline{\lambda}^{\rm nc}_{3,h}| \ . \end{aligned}$$

The relation between the total errors and the degrees of freedom is displayed in Figure 4.17.

Different from the nonconforming FEM in [51] which merely provides the guaranteed lower eigenvalue bounds, the conforming FEM produces both the upper bounds and the lower bounds of the eigenvalues. From the computational results for the two domains and the comparison between the bound (3.21) and the one from [51], we draw the conclusion that

(1) Both the lower eigenvalue bounds proposed in this dissertation and the one in [51] have a sub-optimal convergence rate for the leading Steklov eigenvalues, compared with the convergence rate estimated by the numerical results themselves.

(2) With the same degree of freedom, the lower bound in (3.21) (or (4.4)) gives

slightly better estimation than the one from the nonconforming FEM. However, to obtain the bound (3.21), one has to pay more effort to solve a matrix problem to obtain $\bar{\kappa}_h$.

Table 4.9: Quantities in the eigenvalue estimation (4.4) (γ: convergence rate; unit square domain)

h	$\sqrt{2}/4$	$\sqrt{2}/8$	$\sqrt{2}/16$	$\sqrt{2}/32$	γ
$\bar{\kappa}_h$	0.2891	0.2042	0.1443	0.1021	0.51
$C_{e,h}$	0.5740	0.4059	0.2870	0.2029	0.50
M_h	0.6427	0.4544	0.3208	0.2272	0.51
$\lambda_{1,h}$	0.2404841	0.2401798	0.2401042	0.2400854	2.01
$\underline{\lambda}_{1,h}$	0.218753	0.228833	0.2343144	0.2371468	0.95
$\lambda_{2,h}$	1.527151	1.502305	1.494918	1.492966	1.92
$\underline{\lambda}_{2,h}$	0.936415	1.146662	1.295596	1.386153	0.72

(Note: $\lambda_{2,h} = \lambda_{3,h}$, $\underline{\lambda}_{2,h} = \underline{\lambda}_{3,h}$)

Table 4.10: Quantities in the eigenvalue estimation (4.5) (γ: convergence rate; unit square domain)

h	$\sqrt{2}/4$	$\sqrt{2}/8$	$\sqrt{2}/16$	$\sqrt{2}/32$	γ
$\widehat{C}_{e,h}$	0.6110176	0.4038323	0.2714162	0.1848489	0.61
$\lambda_{1,h}^{\mathrm{nc}}$	0.2404829	0.2401793	0.2401041	0.2400853	2.0
$\underline{\lambda}_{1,h}^{\mathrm{nc}}$	0.2206705	0.2311264	0.235931	0.2381318	1.13
$\lambda_{2,h}^{\mathrm{nc}}$	1.460229	1.483297	1.489892	1.491678	1.88
$\underline{\lambda}_{2,h}^{\mathrm{nc}}$	0.9450309	1.19438	1.342541	1.419335	0.95

(Note: $\lambda_{2,h}^{\mathrm{nc}} = \lambda_{3,h}^{\mathrm{nc}}$, $\underline{\lambda}_{2,h}^{\mathrm{nc}} = \underline{\lambda}_{3,h}^{\mathrm{nc}}$)

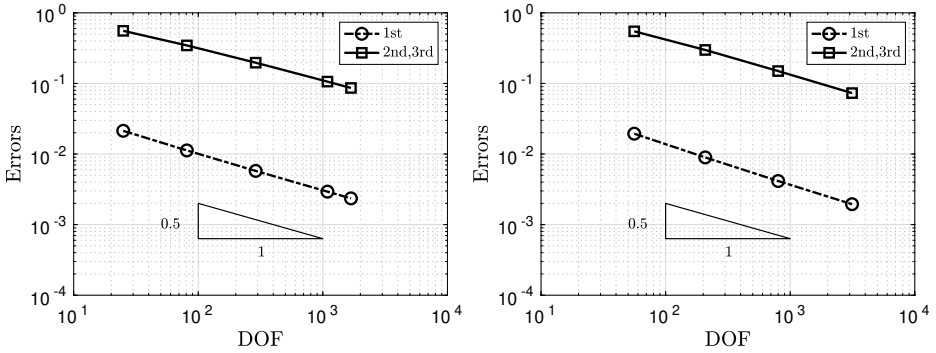


Figure 4.15: Errors of eigenvalue bounds v.s. DOF (the unit square domain) (Left: $|\lambda_i - \underline{\lambda}_{i,h}|$; Right: $|\lambda_i - \underline{\lambda}^{\mathrm{nc}}_{i,h}|$ $(i = 1, 2, 3)$)

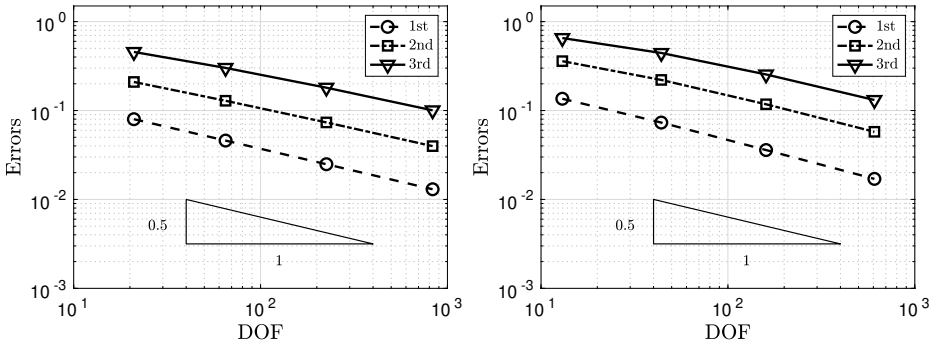


Figure 4.16: Errors of eigenvalue bounds v.s. DOF (the L-shaped domain) (Left: $|\lambda_i - \underline{\lambda}_{i,h}|$, Right: $|\lambda_i - \underline{\lambda}^{\mathrm{nc}}_{i,h}|$ $(i = 1, 2, 3)$)

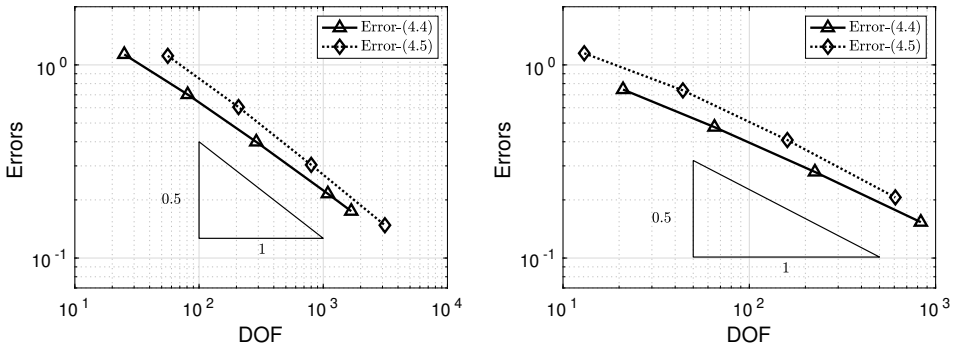


Figure 4.17: The total errors for the eigenvalue bounds v.s. DOF (Left: the unit square; Right: the L-shaped domain)

Table 4.11: Quantities in the eigenvalue estimation (4.4) (γ: convergence rate; L-shaped domain)

h	$\sqrt{2}/2$	$\sqrt{2}/4$	$\sqrt{2}/8$	$\sqrt{2}/16$	γ
$\bar{\kappa}_h$	0.5106	0.3633	0.2591	0.1847	0.48
$C_{e,h}$	0.8118	0.5740	0.4059	0.2870	0.50
M_h	0.9590	0.6793	0.4815	0.3413	0.50
$\lambda_{1,h}$	0.3443305	0.3421498	0.3416010	0.3414626	2.06
$\underline{\lambda}_{1,h}$	0.2615119	0.2954914	0.3165279	0.3283997	0.93
$\lambda_{2,h}$	0.6513041	0.6299816	0.6217140	0.6186763	1.45
$\underline{\lambda}_{2,h}$	0.4073133	0.4880800	0.5433766	0.5770854	0.89
$\lambda_{3,h}$	1.0278736	0.9968693	0.9876317	0.9851393	2.02
$\underline{\lambda}_{3,h}$	0.5283698	0.6827630	0.8035932	0.8837230	0.85

Table 4.12: Quantities in the eigenvalue estimation (4.5) (γ: convergence rate; L-shaped domain)

h	$\sqrt{2}/2$	$\sqrt{2}/4$	$\sqrt{2}/8$	$\sqrt{2}/16$	γ
$\widehat{C}_{e,h}$	0.8997886	0.5890361	0.3928155	0.2659045	0.63
$\lambda_{1,h}^{\mathrm{nc}}$	0.3425959	0.3416846	0.3414799	0.3414316	2.08
$\underline{\lambda}_{1,h}^{\mathrm{nc}}$	0.2682036	0.3054704	0.3243874	0.3333834	1.07
$\lambda_{2,h}^{\mathrm{nc}}$	0.5829704	0.6039094	0.6120116	0.6150436	1.42
$\underline{\lambda}_{2,h}^{\mathrm{nc}}$	0.3960439	0.4992908	0.5592028	0.5894119	0.99
$\lambda_{3,h}^{\mathrm{nc}}$	0.9608929	0.9769290	0.9821661	0.9837098	1.76
$\underline{\lambda}_{3,h}^{\mathrm{nc}}$	0.5404476	0.7296185	0.8529063	0.9197389	0.88

4.2.4 Comparison with the optimal $C_e(K)$ and proposed bound in (3.9)

In this subsection, we estimate the trace constant $C_e(K)$ over several triangle K's directly, and compare with its bound in (3.9). For $i = 1, 2, 3$, denote the i-th edge of

K by e_i. Let us introduce the function space V_{e_i} after V_e in Lemma 3.3.1.

$$V_{e_i}(K) = \{u \in H^1(K) | \int_{e_i} u \, ds = 0\}.$$

The trace constant $C_{e_i}(K)$ is the quantity that makes certain the following estimation holds.

$$\|u\|_{L^2(e_i)} \leq C_{e_i}(K)|u|_{H^1(K)} \quad \forall u \in V_{e_i}(K) \ .$$

The determination of $C_{e_i}(K)$ reduces to finding the minimal positive eigenvalue of the following Steklov eigenvalue problem:

$$-\Delta u = 0 \text{ in } K, \ \ \frac{\partial u}{\partial \mathbf{n}} = \lambda u \text{ on } e_i, \quad \frac{\partial u}{\partial \mathbf{n}} = 0 \text{ on } \partial K \setminus e_i. \tag{4.6}$$

By taking $a(u,v) := (\nabla u, \nabla v)_K, \ b(u,v) := (u,v)_{e_i}$, the weak formulation of (4.6) is given as follows:

$$\text{Find } (\lambda, u) \in \mathbb{R} \times V_{e_i} \text{ s.t.} \quad a(u,v) = \lambda b(u,v) \quad \forall v \in V_{e_i}(K).$$

The strict lower eigenvalue bound for the above eigenvalue problem can be obtained by an analogous argument as performed in this dissertation, the detail of which is omitted here.

We consider three types of triangles (see Figure 4.18) and evaluate $C_{e_i}(K)$ by solving the corresponding Steklov eigenvalue problems using the linear conforming FEM. The results are shown in Table 4.13. It is observed that the bound in (3.9) is not too rough and a direct estimation of $C_{e_i}(K)$ by solving the Steklov eigenvalue problem can obtain a sharper bound for the constant.

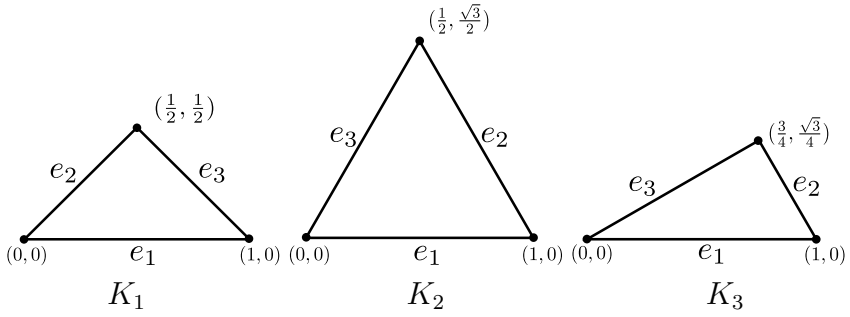


Figure 4.18: Three types of triangles

Table 4.13: Evaluation of $C_{e_i}(K)$ (mesh size $h = 1/256$)

	Approximation of $C_{e_i}(K)$			Upper bound of $C_{e_i}(K)$			Upper bound in (3.9)		
	e_1	e_2	e_3	e_1	e_2	e_3	e_1	e_2	e_3
K_1	0.7071	0.5516	0.5516	0.7198	0.5571	0.5571	1.1481	0.9654	0.9654
K_2	0.6361	0.6361	0.6361	0.6446	0.6446	0.6446	0.8723	0.8723	0.8723
K_3	0.7700	0.4285	0.7071	0.7843	0.4320	0.7169	1.2337	0.8723	1.1480

Remark 4.2.1. The computation codes and results in this section are available on the following website.

`https://ganjin.online/nakano/Guaranteed_error_estimation_for_modified_Helmholtz_eq`

Chapter 5

Conclusion

In this dissertation, we investigate "quantitative error estimation" for two model problems and develop new methods to compute explicit upper bounds for the error of finite element solutions using the hypercircle method. The results of this study are summarized as follows:

(1) A new quantitative local error estimation for finite element solutions of the boundary value problem of the Poisson equation is presented by utilizing the extended hypercircle method (Chapter 3, Theorem 2.3.6). This result is published in [37].

(2) By using the hypercircle method, the quantitative *a priori* error estimation for finite element solutions of the non-homogeneous Neumann boundary value problem of the modified Helmholtz equation is proposed; see Chapter 3, Theorem 3.3.3. The proposed *a priori* error estimation is further combined with Liu's method [31] to provide computable eigenvalue bounds for the

Steklov eigenvalue problem; see Theorem 3.4.1 of §3 and the discussion in §4.2. This result will appear in *Computational Methods in Applied Mathematics* (the preprint is available [35]).

The proposed error estimations in this dissertation have a distinct advantage over previous studies in the field of numerical analysis.

(1) While most of the existing literature focuses solely on the qualitative error analysis of the FEM solution, this study presents the first approach of quantitative error estimation (i.e., error bounds with explicit values) for FEM solutions. Theorem 2.3.6 in Chapter 2 provides the local error estimator for homogeneous boundary value problem, and Theorem 3.3.3 in Chapter 3 provides the global error estimator for non-homogeneous boundary value problem. It is worth pointing out that the convergence rate of obtained local quantitative error estimation agrees with the result from qualitative error analysis even for non-uniform meshes.

(2) The proposed method in this study is capable of addressing the challenging issue of solution singularity that arises around re-entry corners of non-convex boundaries. Different from existing approaches, the proposed *a priori* error estimation does not require higher regularity of the solution and can thus be applied to non-convex domains.

Future work includes the application of the local error estimation to the four-probe method used in resistivity measurement. A promising approach is to combine the idea of [15] and the hypercircle method with the finite element method.

Bibliography

[1] D. S. A. Bermúdez, R. Rodríguez. A finite element solution of an added mass formulation for coupled fluid-solid vibrations. *Numer. Math.*, 87(2):201–227, 2000.

[2] M. Ainsworth and T. Vejchodský. Robust error bounds for finite element approximation of reaction–diffusion problems with non-constant reaction coefficient in arbitrary space dimension. *Comput. Methods Appl. Mech. Eng.*, 281:184–199, 2014.

[3] M. G. Armentano and C. Padra. A posteriori error estimates for the steklov eigenvalue problem. *Appl. Numer. Math.*, 58(5):593–601, 2008.

[4] S. Bergman and M. Schiffer. *Kernel functions and elliptic differential equations in mathematical physics*. Academic Press, New York, 1953.

[5] H. Bi, Y. Zhang, and Y. Yang. Two-grid discretizations and a local finite element scheme for a non-selfadjoint stekloff eigenvalue problem. *Comput. Math. Appl.*, 2018.

[6] D. Braess. *Finite Elements. Theory, Fast Solvers and Applications in Solid Mechanics*. Cambridge University Press, 2007.

[7] J. H. Bramble and J. E. Osborn. Approximation of steklov eigenvalues of non-selfadjoint second order elliptic operators. In *The mathematical foundations of the finite element method with applications to partial differential equations*, pages 387–408. Elsevier, 1972.

[8] F. Brezzi and M. Fortin. *Mixed and Hybrid Finite Element Methods*, volume 15 of *Springer Series in Computational Mathematics*. Springer, 1991.

[9] F. Cakoni, D. Colton, S. Meng, and P. Monk. Stekloff eigenvalues in inverse scattering. *SIAM J. Appl. Math.*, 76(4):1737–1763, 2016.

[10] F. Demengel and G. Demengel. *Functional spaces for the theory of elliptic partial differential equations*. Springer, 2012. translated by R. Erné.

[11] A. Demlow. Local a posteriori estimates for pointwise gradient errors in finite element methods for elliptic problems. *Math. Comp.*, 76:19–42, 01 2007.

[12] A. Demlow. Convergence of an adaptive finite element method for controlling local energy error. *SIAM J. Numer. Anal.*, 48:470–497, 2010.

[13] A. Demlow. Quasi-optimality of adaptive finite element methods for controlling local energy errors. *Numer. Math.*, 134:27–60, 2016.

[14] A. Demlow, J. Guzman, and A. H. Schatz. Local energy estimates for the finite element method on sharply varying grids. *Math. Comp.*, 80:1–9, 2011.

[15] H. Fujita. Contribution to the theory of upper and lower bounds in boundary value problems. *J. Phys. Soc. Japan*, 10(1):1–8, 1955.

[16] J.-L. Guermond and A. Ern. *Finite Elements II: Galerkin Approximation, Elliptic and Mixed PDEs*. Springer, 2021.

[17] J. Hu and R. Ma. The enriched Crouzeix–Raviart elements are equivalent to the Raviart–Thomas elements. *J. Sci. Comput.*, 63(2):410–425, may 2015.

[18] F. E. I. Miccoli, H. Pfnür, and C. Tegenkamp. The 100th anniversary of the four-point probe technique: the role of probe geometries in isotropic and anisotropic systems. *J. Phys. Condens. Matter*, 27(22):223201, 2015.

[19] T. Kato. On some approximate methods concerning the operators T^*T. *Mathematische Annalen*, 126:253–262, 1953.

[20] F. Kikuchi and X. Liu. Determination of the Babuška-Aziz constant for the linear triangular finite element. *Jpn. J. Ind. Appl. Math.*, 23(1):75–82, 2006.

[21] F. Kikuchi and X. Liu. Estimation of interpolation error constants for the P_0 and P_1 triangular finite elements. *Comput. Methods. Appl. Mech. Eng.*, 196(37):3750–3758, 2007.

[22] F. Kikuchi and H. Saito. Remarks on a posteriori error estimation for finite element solutions. *J. Comput. Appl. Mech.*, 199(2):329–336, 2007.

[23] N. Kuznetsov, T. Kulczycki, M. Kwaśnicki, A. Nazarov, S. Poborchi, I. Polterovich, and B. Siudeja. The legacy of vladimir andreevich steklov. *Notices of the AMS*, 61(1):190, 2014.

[24] R. S. Laugesen and B. A. Siudeja. Minimizing Neumann fundamental tones of triangles: An optimal Poincaré inequality. *J. Differ. Equ.*, 249(1):118–135, 2010.

[25] M. Li, Q. Lin, and S. Zhang. Extrapolation and superconvergence of the steklov eigenvalue problem. *Adv. Comput. Math.*, 33(1):25–44, 2010.

[26] Q. Li, Q. Lin, and H. Xie. Nonconforming finite element approximations of the steklov eigenvalue problem and its lower bound approximations. *Appl. Math.*, 58(2):129–151, 2013.

[27] Q. Li and X. Liu. Explicit finite element error estimates for nonhomogeneous neumann problems. *Appl. Math.*, 63:1–13, 2018.

[28] Q. Li and Y. Yang. A two-grid discretization scheme for the steklov eigenvalue problem. *J. Appl. Math. Comput.*, 36(1-2):129–139, 2011.

[29] X. Liao and R. Notchetto. Local a posteriori error estimates and adaptive control of pollution effects. *Numer. Methods Partial Differential Equations*, 19(4):421–442, 2003.

[30] J. Liu, J. Sun, and T. Turner. Spectral indicator method for a non-selfadjoint steklov eigenvalue problem. *J. Sci. Comput.*, 79(3):1814–1831, 2019.

[31] X. Liu. A framework of verified eigenvalue bounds for self-adjoint differential operators. *Appl. Math. Comput.*, 267:341–355, 2015.

[32] X. Liu and F. Kikuchi. Analysis and estimation of error constants for P_0 and P_1 interpolations over triangular finite elements. *J. Math. Sci. Univ. Tokyo*, 17(1):27–78, 2010.

[33] X. Liu and S. Oishi. Verified eigenvalue evaluation for the laplacian over polygonal domains of arbitrary shape. *SIAM J. Numer. Anal.*, 51(3):1634–1654, 2013.

[34] L. D. Marini. An inexpensive method for the evaluation of the solution of the lowest order Raviart–Thomas mixed method. *SIAM J. Numer. Anal.*, 22(3):493–496, 1985.

[35] T. Nakano, Q. Li, , M. Yue, and X. Liu. Guaranteed lower eigenvalue bound of steklov operator with conforming finite element methods. 2022. `https://arxiv.org/abs/2001.09820` (Submitted to Computational Methods in Applied Mathematics).

[36] T. Nakano and X. Liu. Explicit a posteriori local error estimation for finite element solutions. *Transactions of the Japan Society for Industrial and Applied Mathematics*, 29(4):362–382, 2019. in Japanese.

[37] T. Nakano and X. Liu. Guaranteed local error estimation for finite element solutions of boundary value problems. *J. Comput. Appl. Math.*, 425:115061, 2023.

[38] P. Neittaanmäki and S. Repin. *Reliable methods for computer simulation: error control and a posteriori estimates*. Elsevier. Amsterdam, 2004.

[39] J. A. Nitsche and A. H. Schatz. Interior estimates for Ritz-Galerkin methods. *Math Comput*, 28(128):937–958, 1974.

[40] S. Oishi, T. Ogita, M. Kashiwagi, X. Liu, et al. *Principle of verified numerical computations*. CORONA publisher, 2018. in Japanese.

[41] W. Prager and J. L. Synge. Approximations in elasticity based on the concept of function space. *Q. Appl. Math.*, 5(3):241–269, 1947.

[42] A. D. Russo and A. E. Alonso. A posteriori error estimates for nonconforming approximations of steklov eigenvalue problems. *Comput. Math. Appl.*, 62(11):4100–4117, 2011.

[43] G. Savaré. Regularity results for elliptic equations in lipschitz domains. *Journal of Functional Analysis*, 152(1):176–201, 1998.

[44] T. Vejchodský. Flux reconstructions in the Lehmann–Goerisch method for lower bounds on eigenvalues. *J. Comput. Appl. Math.*, 340:676–690, 2018.

[45] L. B. Wahlbin. *Local behavior in finite element methods*, volume 2 of *Handbook of Numerical Analysis*, pages 353–522. Elsevier, 1991.

[46] H. Xie. A type of multilevel method for the steklov eigenvalue problem. *IMA J. Numer. Anal.*, 34(2):592–608, 2014.

[47] J. Xu and A. Zhou. Local and parallel finite element algorithms based on two-grid discretizations. *Math. Comp.*, 69(231):881–909, 2000.

[48] J. Xu and A. Zhou. Local and parallel finite element algorithms based on two-grid discretizations for nonlinear problems. *Adv. Comput. Math.*, 14:293–327, 2001.

[49] M. Yamashita and M. Agu. Geometrical correction factor for semiconductor resistivity measurements by four-point probe method. *Jpn. J. Appl. Phys.*, 23(11, Part 1):1499–1504, 1984.

[50] Y. Yang, Q. Li, and S. Li. Nonconforming finite element approximations of the steklov eigenvalue problem. *Appl. Numer. Math.*, 59(10):2388–2401, 2009.

[51] C. You, H. Xie, and X. Liu. Guaranteed eigenvalue bounds for the steklov eigenvalue problem. *SIAM J. Numer. Anal.*, 57(3):1395–1410, 2019.