Bilinear Immersed Finite Elements For Interface Problems

Xiaoming He

Dissertation submitted to the Faculty of the
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
in
Mathematics

Tao Lin, Chair
Joseph Wang, Co-chair
Slimane Adjerid
Christopher Beattie
Shu-Ming Sun

April 20, 2009
Blacksburg, Virginia

Keywords: Interface Problems, Immersed Finite Elements, Error Estimates, Galerkin Method, Finite Volume Element Method, Discontinuous Galerkin Method

Bilinear Immersed Finite Elements For Interface Problems

Xiaoming He

(ABSTRACT)

In this dissertation we discuss bilinear immersed finite elements (IFE) for solving interface problems. The related research works can be categorized into three aspects: (1) the construction of the bilinear immersed finite element spaces; (2) numerical methods based on these IFE spaces for solving interface problems; and (3) the corresponding error analysis. All of these together form a solid foundation for the bilinear IFEs.

The research on immersed finite elements is motivated by many real world applications, in which a simulation domain is often formed by several materials separated from each other by curves or surfaces while a mesh independent of interface instead of a body-fitting mesh is preferred. The bilinear IFE spaces are nonconforming finite element spaces and the mesh can be independent of interface. The error estimates for the interpolation of a Sobolev function in a bilinear IFE space indicate that this space has the usual approximation capability expected from bilinear polynomials, which is $O(h^2)$ in L^2 norm and $O(h)$ in H^1 norm. Then the immersed spaces are applied in Galerkin, finite volume element (FVE) and discontinuous Galerkin (DG) methods for solving interface problems. Numerical examples show that these methods based on the bilinear IFE spaces have the same optimal convergence rates as those based on the standard bilinear finite element for solutions with certain smoothness. For the symmetric selective immersed discontinuous Galerkin method based on bilinear IFE, we have established its optimal convergence rate. For the Galerkin method based on bilinear IFE, we have also established its convergence.

One of the important advantages of the discontinuous Galerkin method is its flexibility for both p and h mesh refinement. Because IFEs can use a mesh independent of interface, such as a structured mesh, the combination of a DG method and IFEs allows a flexible adaptive mesh independent of interface to be used for solving interface problems. That is, a mesh independent of interface can be refined wherever needed, such as around the interface and the singular source. We also develop an efficient selective immersed discontinuous Galerkin method. It uses the sophisticated discontinuous Galerkin formulation only around the locations needed, but uses the simpler Galerkin formulation everywhere else. This selective formulation leads to an algebraic system with far less unknowns than the immersed DG method without scarifying the accuracy; hence it is far more efficient than the conventional discontinuous Galerkin formulations.

This work received support from the NSF grant DMS-0713763 and NSERC (Canada).

Acknowledgments

First of all, I would like to thank my advisor, Dr. Tao Lin, for your inspiring guidance in the past four years. Without your help, patience and support, I could not appreciate mathematics this far. I believe that I will benefit from the knowledge, capability and experience I obtained from you in my whole life.

I also want to thank the co-chair of my committee, Dr. Joseph Wang, for the opportunity and guidance you gave me to study applications in aerospace engineering with you. It is my first opportunity to look into the dynamic relationship between applied mathematics and engineering and I highly appreciate it.

I would also like to thank my committee, Dr. Slimane Adjerid, Dr. Christopher Beattie, and Dr. Shu-Ming Sun. I have learned a lot from you and I'm grateful to your effort and support on my dissertation.

I would like to thank my parents, grandparents and all of my family and friends. Your love is the most important part of my life, which always encourages me to bravely chase my dream.

Lan, my love, without you I don't know if I can finish this work in four years. I won't forget those extra dinners you prepared for me at 3:00am when I was busy on my work with an empty stomach. I won't forget the happiness and energy you gave me when I felt depressed about my failures. I won't forget anything you did for me so that I can concentrate on my research. I love you and I will love you forever.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

We consider the following model interface problem:

$$-\nabla \cdot \big(\beta \nabla u\big) = f, \quad (x,y) \in \Omega, \tag{1.1}$$

$$u|_{\partial\Omega} = g, \tag{1.2}$$

together with the jump conditions on the interface Γ:

$$[u]\,|_{\Gamma} = 0, \tag{1.3}$$

$$\left[\beta \frac{\partial u}{\partial n}\right]|_{\Gamma} = 0. \tag{1.4}$$

As illustrated in Figure 1.1, without loss of generality, we assume that $\Omega \subset I\!R^2$ is a rectangular domain, the interface Γ is a curve separating Ω into two sub-domains Ω^-, Ω^+ such that $\overline{\Omega} = \overline{\Omega^-} \cup \overline{\Omega^+} \cup \Gamma$, and the coefficient $\beta(x,y)$ is a piecewise constant function defined by

$$\beta(x,y) = \begin{cases} \beta^-, & (x,y) \in \Omega^-, \\ \beta^+, & (x,y) \in \Omega^+. \end{cases}$$

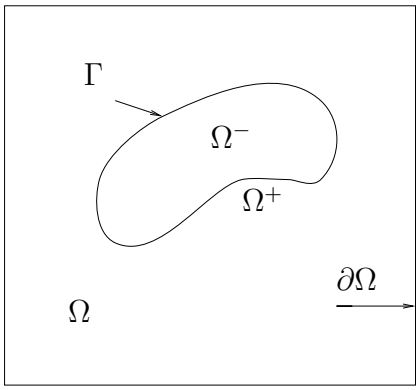

Figure 1.1: A sketch of the domain for the interface problem.

First we consider the homogeneous jump condition for introducing immersed finite elements. In Chapter 9, we will discuss how to extend the IFE spaces for dealing with the non-homogeneous jump conditions.

It is well-known that efficiently solving this kind of interface problem is critical in many applications of engineering and sciences, including electromagnetic problems [6, 32, 151, 152, 165, 166, 167, 170, 171, 191], flow problems [55, 56, 57, 58, 22, 23, 75, 120, 138], shape/topology optimization problems [25, 26, 27, 74, 95, 108, 109, 125, 188], and the modeling of nonlinear phenomena [118, 194], to name just a few. In this chapter, we will first discuss a few representative applications and then review some previous work on the model interface problem.

1.1 Applications of the model interface problem

In this section, we will discuss three interesting applications that involve the model interface problem: charging in space, projection method for Navier-Stokes equation and the shape/topology optimization.

1.1.1 Charging in space

Charging is the physics underlying many problem associated with either natural phenomenon or engineering applications. Charging in space appears in a series of important problems in aerospace engineering, such as spacecraft charging under different types of solar winds, ion propulsion and electrostatic levitation of lunar dust, to name just a few. In the following, we will discuss the electrostatic levitation of lunar dust.

It is general accepted that lunar dust levitation is the primary mechanism for lunar horizontal glow and lunar dust fountain, which were observed by Surveyor 5, 6, 7 and Apollo 17 [181]. Also, lunar dust can cause a wide range of serious problems for spacecraft and astronauts such as vision obscuration, false instrument readings, clogging of equipment, seal failures, contamination of surface, abrasion of space suits, breathing problems, and long term lung problems for astronauts, etc. Figure 1.2 shows some pictures of lunar dust. The shape and charging of lunar dust are actually the major reasons to cause those problems. From the following remarks by Gene Cernan, the commander of Apollo 17, at Apollo 17 technical debriefing, we can see the dramatic influence of lunar dust on spacecraft and astronauts.

Remark 1.1.1 *...I think probably one of the most aggravating, restricting facets of lunar surface exploration is the dust and its adherence to everything no matter what kind of material, whether it be skin, suit material, metal, no matter what it be and it's restrictive friction-like action to everything it gets on....We tried to dust them and bang the dust off*

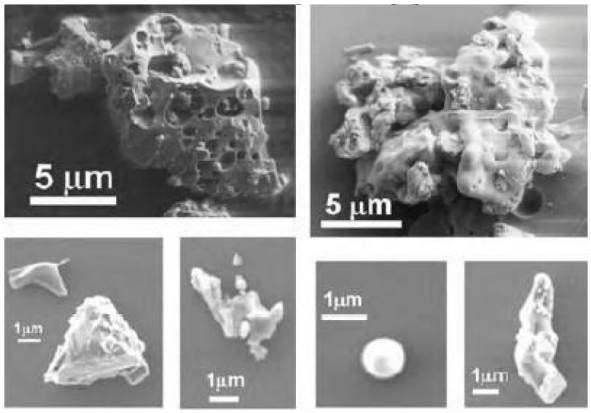

Figure 1.2: SEM images of Apollo 17 lunar dust 70051

and clean them, and there was just no way...I think dust is probably one of our greatest inhibitors to a nominal operation on the Moon. I think we can overcome other physiological or physical or mechanical problems except dust...

In a word, it is very important to study the electrostatic levitation of lunar dust. Because lunar surface, lunar dust, spacecraft and astronauts are charged under different types of solar winds, they all interact with each other and affect the electric field. First, we need to simulate the solar winds to generate the electric field. Then we can simulate the movement of the lunar dust and the interaction between space and different objects, such as lunar surface, lunar dust, spacecraft and astronauts. One of the efficient methods for plasma simulation is called Particle-In-Cell (PIC) method.

In PIC method, first we set up a mesh for the domain of the simulation problem, which is called PIC mesh. It will be used to locate the simulation particles, collect the information from particles, and realize the effect of the electromagnetic field on the particles. Second, because the charging is always changing, we need a partition in time. When the partition is fine enough, we can assume that in each time interval, the charge is fixed. Third, at the beginning of each time interval, we deposit the charge of all the particles to the PIC mesh nodes according to the position of the particles. Fourth, using the charge, we can compute the electric potential, hence the electric field. Then, we can compute the acceleration, velocity and movement of the particles in this time interval. At the end of this time interval, we get the new positions of all the particles, which will be used at the start of the next time interval.

Therefore, the basic steps in the PIC simulation are

$$
\begin{aligned}
& \text{Position of particles and charge on objects} \\
\Longrightarrow\ & \text{Charge on nodes} \\
\Longrightarrow\ & \text{Electric potential} \\
\Longrightarrow\ & \text{Electric field} \\
\Longrightarrow\ & \text{Acceleration, velocity and movement of particles} \\
\Longrightarrow\ & \text{New position of particles and charge on objects} \\
\Longrightarrow\ & \cdots
\end{aligned}
$$

One of the key steps in the simulation is

$$\text{Charge on nodes} \implies \text{Electric potential}$$

for which we need to numerically solve the following partial differential equation (PDE)

$$-\nabla \cdot \big(\beta \nabla \phi\big) = \rho, \quad \text{in } \Omega,$$

with certain boundary conditions. Here ϕ is the potential, ρ is the charge density and β is the dielectric parameter. This is the same as our model interface problem.

In general, PIC itself needs two meshes. One is to locate the simulation particles, collect the information from particles, and realize the effect of the electromagnetic field on the particles. The other one is for numerically solving a partial differential equation for the electric potential. Since we generally need thousands of simulation steps to reach the stable status, the two meshes should be well structured and easily communicate with each other. Otherwise, the computational expense is formidable. However, the problem usually has complicated interfaces between the space and different objects, such as lunar surfaces, spacecraft and astronauts. Therefore, when we use the standard finite element, finite difference or finite volume methods to solve the partial differential equation, we have to use body-fitting meshes, which are unstructured and not suitable for PIC. Meanwhile, when the spacecraft and astronauts are moving, the interfaces are moving. Hence the body-fitting meshes have to be reformed again and again, which will increase computational cost even further by significant amount.

This example is just one of the numerous plasma particle simulation applications. Many other kinds of plasma particle simulations, such as ion optics plasma dynamics [130], are in the same situation as above. Therefore, a structured mesh instead of a body-fitting mesh is necessary for solving the interface partial differential equation for PIC method. As we can see later, the immersed finite element method is one of the most efficient methods to solve interface partial differential equations with a mesh independent of interface.

1.1.2 Projection method for solving Navier-Stokes equations

It is well known that the Navier-Stokes equation is highly critical to flow problems. Therefore, how to solve it numerically leads to a lot of interesting work. One of the popular methods is the projection method, which was developed by A. J. Chorin first at the end of 1960's, see [55, 56, 57, 58]. Since then, people have generalized his idea to develop many kinds of projection methods, see [22, 23, 120, 138] and references therein.

Consider the following Navier-Stokes equation

$$\begin{aligned} &\rho(U_t + U \cdot \nabla U) = -\nabla p + \nu \Delta U + F, \\ &-\nabla \cdot U = 0, \end{aligned}$$

where U is the velocity, p is the pressure, ρ is the density and ν is the viscosity. Basically, Chorin's projection method is an iteration method as follows. First, we set up a partition $0 = t_0 < t_1 < \cdots < t_m < \cdots$ of the time with a fixed step size Δt. Let U^n be the approximation of U at time t_n, then the first step is to calculate an intermediate velocity U^* which satisfies

$$\frac{U^* - U^n}{\Delta t} = -U^n \cdot \nabla U^n + \frac{1}{\rho}(\nu \Delta U + F)^n.$$

Then we correct U^* by the pressure term with

$$\frac{U^{n+1} - U^*}{\Delta t} = -\frac{\nabla p}{\rho}, \tag{1.5}$$

where

$$-\nabla \cdot U^{n+1} = 0. \tag{1.6}$$

Then by (1.5) and (1.6), we get

$$\nabla \cdot (\frac{1}{\rho} \nabla p) = \frac{\nabla \cdot U^*}{\Delta t}.$$

Finally, we get the iteration method as follows.

$$\begin{aligned} \frac{U^* - U^n}{\Delta t} &= -U^n \cdot \nabla U^n + \frac{1}{\rho}(\nu \Delta U + F)^n, \\ \nabla \cdot (\frac{1}{\rho} \nabla p) &= \frac{\nabla \cdot U^*}{\Delta t}, \\ \frac{U^{n+1} - U^*}{\Delta t} &= -\frac{\nabla p}{\rho}. \end{aligned}$$

Obviously, the second step is the most time consuming part of this method. Therefore, an efficient method for solving the PDE in the second step is crucial for the performance of the whole method. If there are at least two fluids in the simulation domain separated from each other, then ρ has jump at the interface of different flows, which leads to the model interface problem considered in this dissertation.

1.1.3 Shape/Topology optimization

From our daily experience, we know that the efficiency and reliability of manufactured products depend on geometrical aspects. Therefore, it is not surprising that optimal shape design problems have attracted the attentions of many mathematicians and engineers. Many shape optimization problems have been studied for more than 15 years, such as shape optimization of linearly elastic structures, see [25] and references therein. There are also some quite new problems, such as topology optimization for fluid problems, see [95] and references therein.

In general, we want to minimize a cost function in an optimization problem. However, the minimization problem is usually subjected to some partial differential equations which are similar to our model interface problem. For example, consider the following topology optimization of heat conduction problems [95]. For the design of an optimal heat-conducting device, we want to find the conductivity distribution $k(x)$ that produces the least heat when the amount of high conduction material is limited. Let Ω be the domain. T is the temperature and f is the volumetric heat source. The cost function is

$$C = \int_{\Omega} \nabla T \cdot (k \nabla T),$$

subject to

$$\begin{aligned} &\nabla \cdot (k \nabla T) + f = 0, \quad \text{in } \Omega, \\ &T = 0, \quad \text{on } \Gamma_D, \\ &(k \nabla T) \cdot n = 0, \quad \text{on } \Gamma_N, \end{aligned} \tag{1.7}$$

where Γ_D is the Dirichlet boundary, Γ_N is the Neumann boundary, n is the unit normal vector of Γ_N. The cost function C represents the amount of the heat to be minimized. When the problem domain consists of at least two materials, k is discontinuous, then (1.7) becomes the model interface problem considered in this dissertation.

1.2 Methods for solving interface problems

In this section, we will survey some methods for solving interface problems.

Conventional finite difference (FD) methods [117, 186] and finite element (FE) methods [12, 34, 54] can be used to solve the model interface problem. However, in order to guarantee their convergence, they have to use a body-fitting mesh. That is, the mesh must be tailored to resolve the interface, see Figure 1.3. Otherwise, because of the lack of smoothness of the exact solution across the interface, these methods may not have optimal convergence rates or even don't converge at all [17, 34, 54].

This restriction leads to many drawbacks. For the applications with moving interfaces, the mesh has to be reformed again and again according to the varying interface, which prevents

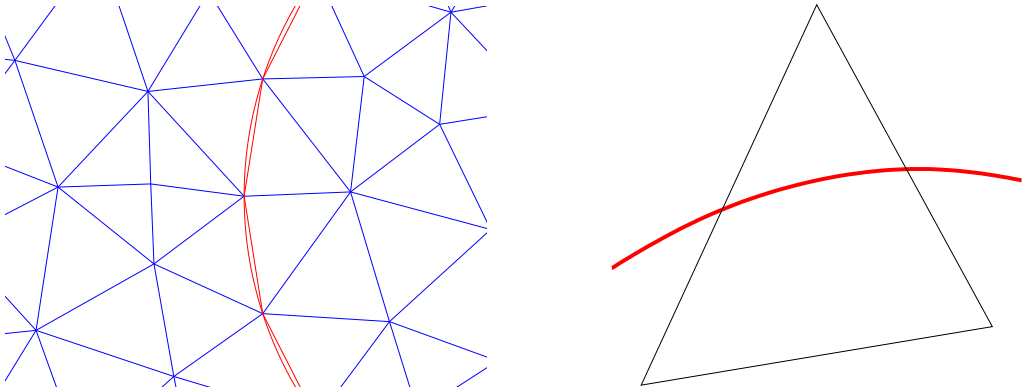

Figure 1.3: The plot on the left shows how elements should be placed on one side of an interface in a standard FE method. An element not allowed in a standard FE method is illustrated by the plot on the right.

the conventional finite difference and finite element method from working efficiently. For large scale problems, the computational expense for reforming the mesh again and again is formidable. Additionally there are many applications and methods which prefer structured meshes and work best with Cartesian meshes, such as Particle-In-Cell method for plasma particle simulation, multigrid, level set method for moving interface/boundary problems, and construction of super convergent approximations to important quantities. Last but not least, many efficient solvers and packages, are available for structured meshes, not body-fitting meshes.

Therefore, many methods have been developed to get rid of this limitation so that interface problems can be solved with a mesh independent of interface. In finite difference formulation, we first note the early work of Peskin's immersed boundary method [174, 175]. Since then, finite difference or finite volume methods such as the Cartesian grid method [4, 39], embedded boundary method [119], immersed interface method [93, 136, 137], cut-cell method [123], matched interface and boundary method [205, 206], etc., have been developed.

In finite element formulation, Babuška et al. [13, 16, 15] developed the generalized finite element method. Their basic idea is to form the local basis functions in an element by solving the interface problem locally. The local basis functions in their method can capture important features of the exact solution and they can even be non-polynomials. Examples of such methods in this framework are the partition of unity method [14], the extended finite element methods (X-FEMs) [24, 163, 189], and the Eulerian-Lagrangian localized adjoint methods [86, 87].

Immersed finite element method also falls into this framework, but it doesn't solve the interface problem locally to form the local basis functions. Instead, it uses the jump conditions to form the reference basis functions and then maps them onto the local elements.

1.3 Organization of the dissertation

In this dissertation, we will discuss the three fundamental aspects for the development of bilinear immersed finite element(IFE): the construction of the bilinear immersed finite element spaces, the implementation of numerical methods based on these new IFE spaces for interface problems, and the corresponding error estimation. The rest of this dissertation is organized as follows.

In Chapter 2, we recall the history of immersed finite elements.

In Chapter 3, we recall a bilinear immersed finite element space and present its properties.

In Chapter 4, we prove the optimal approximation capability of a bilinear immersed finite element space. Some numerical examples are also provided to verify the interpolation error estimates.

In Chapter 5, we discuss the Galerkin method based on the bilinear immersed finite element space. Numerical examples show that this method possesses the optimal convergence rates in both L^2 and H^1 norms. Convergence of this method is also proved.

In Chapter 6, we apply the bilinear immersed finite element space to the finite volume element(FVE) method. Optimal convergence rates in L^2 and H^1 norm are demonstrated numerically.

In Chapter 7, we introduce an immersed discontinuous Galerkin (DG) method which combines the interior penalty DG method together with a bilinear immersed finite element space. Numerical examples are provided to illustrate the features of the two methods, including the optimal convergence rate in energy norm, the effect of local mesh refinement, and the dependence and independence on the penalty constant, etc. The convergence analysis of the immersed DG method is just a special case of the convergence analysis in Chapter 8.

In Chapter 8, we propose a selective immersed discontinuous Galerkin method. This method only applies the discontinuous Galerkin formulation wherever necessary, but the standard Galerkin formulation everywhere else. This method possesses the easy local refinement feature of DG method while keeping the computational cost as close to that of Galerkin method as possible. Hence, this new method has greatly alleviated the shortcoming of the usual DG methods, which is their much larger computational cost than Galerkin method. We also prove the optimal convergence rate in energy norm for the symmetric selective immersed DG method with bilinear IFE.

In Chapter 9, we introduce a new bilinear IFE space to deal with the non-homogeneous flux jump condition based on the bilinear IFE space discussed in Chapter 3. Some numerical examples are provided.

In Chapter 10, we draw some conclusions for this dissertation and discuss some real world applications of immersed finite elements. Some future plans are also discussed in this chapter.

Chapter 2

Review for immersed finite elements(IFE)

In this chapter, we review some representative work for IFE, including the basic idea, construction of different IFE spaces, implementation of different numerical methods based on these IFE spaces, and the corresponding error analysis.

2.1 Basic idea of IFE

The recently developed immersed finite element (IFE) methods [2, 3, 40, 84, 98, 112, 113, 129, 130, 141, 142, 143, 144, 149, 150, 187, 199] falls into the general framework of Babuška and J.E. Osborn [16, 15] to adapt finite element methods for interface problems by employing local basis functions formed according to the interface jump conditions while their meshes can be independent of the interface. However, IFE methods do not locally solve the interface problem. The main idea in IFE methods is more similar to that used for the Hsieh-Clough-Tocher macro C^1 element [33] where each local basis function in an element is defined piecewisely by cubic polynomials on three sub-triangles such that the required continuity can be satisfied.

In IFE, we can use a mesh independent of interface, such as Cartesian mesh, and allow the interface to go through the interior of the elements. Therefore, the mesh in an IFE method consists of interface elements whose interiors are cut through by the interfaces and the rest called non-interface elements. An IFE method uses standard finite element functions in all non-interface elements. Special piecewise polynomials satisfying interface jump conditions are employed only in interface elements. This is the basic idea of IFE.

2.2 Construction of IFE spaces

In 1998, Z. Li [141] introduced an IFE space for a 1D interface problem as follows. Consider

$$\begin{aligned} &-(\beta(x)u')' + q(x)u(x) = f(x), \quad 0 \le x \le 1, \\ &u(0) = 0, \quad u(1) = 0, \end{aligned}$$

where

$$\begin{aligned} &0 = \alpha_0 < \alpha_1 < \alpha_2 = 1, \\ &\beta = \beta_i \text{ in } [\alpha_i, \alpha_{i+1}), \quad i = 0, 1, \\ &[u]_{\alpha_1} = 0, \quad [\beta u_x]_{\alpha_1} = 0. \end{aligned}$$

Here $[u]_{\alpha_1}$ is the jump of function u at α_1.

In a uniform partition $0 = x_0 < x_1 < \cdots < x_n = 1$ of $[0, 1]$ with step size h, assume that the element $[x_j, x_{j+1}]$ has the interface point α_1 inside it. Two piecewise linear basis functions at x_j and x_{j+1} were constructed as follows.

$$\phi_j(x) = \begin{cases} \dfrac{x - x_{j-1}}{h}, & x_{j-1} \le x < x_j, \\ \dfrac{x_j - x}{D} + 1, & x_j \le x < \alpha_1, \\ \dfrac{\rho(x_{j+1} - x)}{D}, & \alpha_1 \le x < x_{j+1}, \\ 0, & otherwise, \end{cases}$$

$$\phi_{j+1}(x) = \begin{cases} \dfrac{x - x_j}{D}, & x_j \le x < \alpha_1 \\ \dfrac{\rho(x - x_{j+1})}{D} + 1, & \alpha_1 \le < x_{j+1}, \\ \dfrac{x_{j+2} - x}{h}, & x_{j+1} \le x < x_{j+2}, \\ 0, & otherwise, \end{cases}$$

where

$$\begin{aligned} &\rho = \frac{\beta_0}{\beta_1}, \\ &D = h - \frac{\beta_1 - \beta_0}{\beta_1}(x_{j+1} - \alpha_1). \end{aligned}$$

See Figure 2.1 for the graph of the two basis functions.

Remark 2.2.1 *We would like to note that these basis functions satisfy the interface jump conditions. Moreover, for the cases with more interface points, we can use the same way to construct more basis functions as above.*

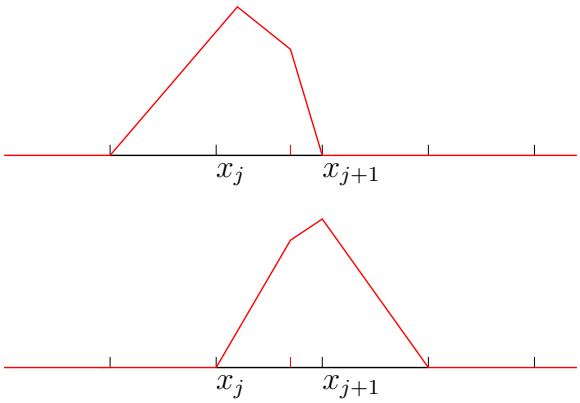

Figure 2.1: Two 1D linear IFE basis functions

These two functions are used as the global finite element basis at nodes x_j and x_{j+1} while standard finite element basis are used at other nodes. Then the new finite element space is the span of all these global nodal basis functions.

Remark 2.2.2 *Basically, the above construction falls into the general framework of Babuška and J.E. Osborn [17, 16] from which we can deduce the construction above.*

In the following, we will briefly recall the definitions of the 2D linear IFE space discussed in [143, 144]. For any subset T of Ω, we let

$$T^s = T \cap \Omega^s, \quad s = -, +.$$

For any function $f(x, y)$ defined in $T \subset \Omega$, we can restrict it to $T^s, s = -, +$ to obtain two functions as

$$f^s(x, y) = f(x, y), \quad \text{if } (x, y) \in T^s, s = -, +.$$

We use $\overline{DE}$ to denote the line segment between two points $D, E \in \Omega$. Consider a typical triangular element $T \in \mathcal{T}_h$. Here, $\mathcal{T}_h, h > 0$ is a family of triangular meshes of the solution domain Ω. Assume that the three vertices of T are $A_i, i = 1, 2, 3$, with $A_i = (x_i, y_i)^t$. If T is an interface element, then we use $D = (x_{_D}, y_{_D})^T$ and $E = (x_{_E}, y_{_E})^T$ to denote the interface points on its edges, see the sketch in Figure 2.2.

Then the 2D linear IFE functions, which satisfy the interface jump conditions, are defined as follows:

$$\phi(x, y) = \begin{cases} \phi^-(x, y) = a^- x + b^- y + c^-, & (x, y) \in T^-, \\ \phi^+(x, y) = a^+ x + b^+ y + c^+, & (x, y) \in T^+, \\ \phi^-(D) = \phi^+(D), \quad \phi^-(E) = \phi^+(E), & \\ \beta^- \dfrac{\partial \phi^-}{\partial \mathbf{n}_{\overline{DE}}} = \beta^+ \dfrac{\partial \phi^+}{\partial \mathbf{n}_{\overline{DE}}}, & \end{cases} \tag{2.1}$$

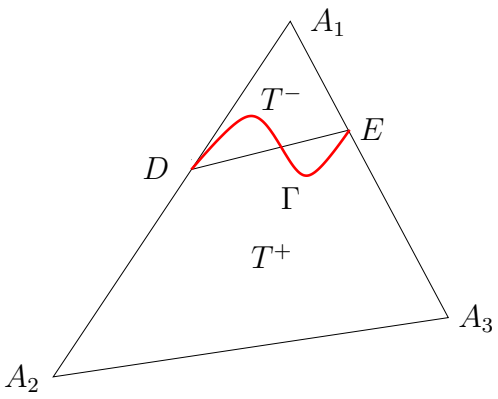

Figure 2.2: A typical triangular interface element.

where $\mathbf{n}_{\overline{DE}}$ is the unit normal vector of the line $\overline{DE}$. We let $\phi_i(X)$ be the piecewise linear function described by (2.1) such that

$$\phi_i(x_j, y_j) = \begin{cases} 1, & \text{if } i = j, \\ 0, & \text{if } i \neq j, \end{cases}$$

for $1 \leq i, j \leq 3$, and we call them the 2D linear IFE nodal basis functions on an interface element T. Now we use the mesh $\mathcal{T}_h$ to define the 2D linear immersed finite element (IFE) space $S_h(\Omega)$. First, for every element $T \in \mathcal{T}_h$, we let $S_h(T) = span\{\phi_i,\ i = 1, 2, 3\}$, where $\phi_i,\ i = 1, 2, 3$ are the standard bilinear nodal basis functions for a non-interface element T; otherwise, $\phi_i,\ i = 1, 2, 3$ are the immersed bilinear basis functions defined above. Then, we define a continuous piecewise bilinear global nodal basis function $\phi_N(x, y)$ for each node $(x_N, y_N)^t$ of $\mathcal{T}_h$ such that $\phi_N(x_N, y_N) = 1$ but zero at other nodes, and $\phi_N|_T \in S_h(T)$ for any rectangle $T \in \mathcal{T}_h$. Finally, we define $S_h(\Omega)$ as the span of these global nodal basis functions. For more details about this space, such as its properties, we refer reader to [143, 144].

In addition to the above two IFE spaces, T. Lin, Y. Lin, R. C. Rogers and L. M. Ryan [149] developed a bilinear IFE space in 2004. R. Kafafy, T. Lin, Y. Lin and J. Wang [129] developed a 3D linear IFE space in 2005. B. Camp, T. Lin, Y. Lin and W. W. Sun [40] developed 1D quadratic IFE space and a special 2D quadratic IFE space in 2006. S. Adjerid and T. Lin [3] developed 1D IFE spaces with any degree in 2009.

2.3 IFE methods for interface problems

Galerkin method is a natural application of IFEs and all the IFE spaces mentioned above have been applied to this method. For example, the Galerkin scheme using the 2D linear IFE space can be described as follows: find $u_h \in S_{h,0}$ satisfying

$$\sum_{T \in \mathcal{T}_h} \int_T \beta \nabla u_h \cdot \nabla v_h \, dxdy \;=\; \int_\Omega f v_h \, dxdy, \forall\ v_h \in S_{h,0},$$

where $S_{h,0}(\Omega) \subset S_h(\Omega)$ consists of functions of $S_h(\Omega)$ vanishing on $\mathcal{N}_h \cap \partial\Omega$.

For more details about Galerkin method using IFE spaces, such as the the numerical examples demonstrating the performance, we refer the reader to see [3, 40, 141, 143, 144, 149, 150].

In addition to Galerkin method, in 1999, R. E. Ewing, Z. Li, T. Lin and Y. Lin [84] applied 2D linear IFE space to finite volume element method. In 2007, S. Adjerid and T. Lin [2] applied 1D IFE spaces to a discontinuous Galerkin method, which transfer a higher-order PDE to a system of first-order PDEs. Also, J. Wang, T. Lin and R. Kafafy applied the 3D linear IFE space to some electromagnetic problems, such as plasma particle simulation and ion optics modeling, see [130, 151, 152].

2.4 Error estimation of IFE

After we develop an IFE space, it is natural for us to analyze its approximation capability. For the 1D linear IFE space, Z. Li [141] obtained the interpolation error estimation and the finite element solution error estimation in 1998. Z. Li, T. Lin, Y. Lin and R. C. Rogers [143] obtained the IFE interpolation error estimation for the 2D liner IFE space in 2004. For 1D IFE spaces with any degree, S. Adjerid and T. Lin[3] derived the interpolation error estimate and the finite element solution error estimate in 2009. We would like to note that the analysis in 2D and higher dimension is much more difficult than that of 1D case since the IFE functions are continuous in 1D but discontinuous in higher dimensional cases.

Let h be the step size of a mesh. For the approximation capability of the standard 2D linear finite element space, we have the following well known theorem, see [178] and references therein.

Theorem 2.4.1 *There exists a constant C independent of mesh such that*

$$\begin{aligned} \|I_h u - u\|_{0,\Omega} &\leq Ch^2 |u|_{2,\Omega}, \\ |I_h u - u|_{1,\Omega} &\leq Ch |u|_{2,\Omega}, \end{aligned}$$

for any $u \in H^2(\Omega)$ and $h > 0$ small enough, where $I_h u$ is the finite element interpolation of u.

One popular way for the interpolation error estimation is the scaling technique. However, when we apply this technique to the 2D or 3D IFE spaces, it's not clear that if the constant C in the error bound depends on the interface and the optimal order cannot be achieved, which cause the error estimation to fail. Therefore, multi-point Taylor expansions were used to estimate the interpolation error of 2D linear IFE space [143]. The following is the final conclusion from [143].

Theorem 2.4.2 *There exists a constant C independent of interface and mesh such that*

$$\begin{aligned} \|I_h u - u\|_{0,\Omega} &\leq Ch^2 \|u\|_{2,\Omega}, \\ |I_h u - u|_{1,\Omega} &\leq Ch \|u\|_{2,\Omega}, \end{aligned}$$

for any $u \in PH^2_{int}(\Omega)$ and $h > 0$ small enough, where $I_h u$ is the immersed finite element interpolation of u. Here

$$\begin{aligned} &PH^2_{int}(\Omega) = \left\{ u \in C(\Omega), u|_{\Omega^s} \in H^2(\Omega^s), s = -, +, \left[\beta \frac{\partial u}{\partial \mathbf{n_\Gamma}}\right] = 0 \quad on\ \Gamma \cap \Omega \right\}, \\ &\|I_h u - u\|^2_{0,\Omega} = \|I_h u - u\|^2_{0,\Omega^+} + \|I_h u - u\|^2_{0,\Omega^-}, \\ &|I_h u - u|^2_{1,\Omega} = |I_h u - u|^2_{1,\Omega^+} + |I_h u - u|^2_{1,\Omega^-}. \end{aligned}$$

We can see that the interpolation error of 2D linear IFE space has the same accuracy order as that of the standard 2D linear finite element space.

For the 1D p^{th} degree IFE spaces, S. Adjerid and T. Lin [3] obtained error estimates for the interpolation error and IFE solution error as follow:

Consider

$$\begin{aligned} &-(\beta(x)u')' + q(x)u(x) = f(x), \quad x \in I = (a, b), \\ &u(a) = u_0, \quad u(b) = u_1, \end{aligned}$$

where

$$\begin{aligned} &\beta = \begin{cases} \beta^-, & x \in I^- = (a, \alpha), \\ \beta^+, & x \in I^+ = [\alpha, b), \end{cases} \\ &[u]_\alpha = 0, \\ &[\beta u_x]_\alpha = 0. \end{aligned}$$

Here $[u]_\alpha$ is the jump of function u at α.

Let

$$\tilde{H}^{p+1}(I) = \left\{ u \in C(I), u|_{I^s} \in H^{p+1}(\Omega^s), s = -, +, \left[\beta \frac{\partial^j u}{\partial x^j}\right]_\alpha = 0, \quad j = 1, 2, \cdots, p \right\}.$$

Then from [3], we have the following two theorems.

Theorem 2.4.3 *There exists a constant C independent of α such that for all $u \in \tilde{H}^{p+1}(\Omega)$, we have*

$$\begin{aligned} \|I_p u - u\|_0 &\leq C \frac{4^p}{(p-1)!} h^{p+1} |u|_{p+1}, \\ |I_p u - u|_1 &\leq C \frac{4^p}{(p-1)!} h^p |u|_{p+1}. \end{aligned}$$

Here $I_p u$ is the p^{th} degree IFE interpolation of u.

Theorem 2.4.4 *There exists a constant C independent of α such that the p^{th} degree IFE solution u_h satisfies*

$$\|u_h - u\|_0 \leq C\frac{4^p}{(p-1)!}h^{p+1}\,|u|_{p+1}\,,$$
$$|u_h - u|_1 \leq C\frac{4^p}{(p-1)!}h^{p}\,|u|_{p+1}\,.$$

Chapter 3

Bilinear immersed finite element

From this chapter, we start to discuss the bilinear immersed finite element which was originally introduced in [149]. We will first recall the definition of bilinear immersed finite element basis functions and slightly modify it. Then we use them to construct a bilinear immersed finite element space and describe its basic properties [110, 111, 112, 113, 149].

3.1 A bilinear IFE space

In this section, we will recall the construction of a bilinear immersed finite element space from [149] and slightly modify it. Without loss of generality, we assume in the discussion from now on that the elements in a rectangular mesh have the following features when the mesh size is small enough:

(H_1): An interface Γ will not intersect an edge of any element at more than two points unless this edge is part of Γ.

(H_2): If Γ intersects the boundary of a rectangle at two points, then these two points must be on different edges of this rectangle.

Now we recall the following definitions. For any subset Λ of Ω, we let

$$\Lambda^s = \Lambda \cap \Omega^s, \quad s = -, +.$$

For any function $f(x, y)$ defined in $\Lambda \subset \Omega$, we can restrict it to $\Lambda^s, s = -, +$ to obtain two functions as

$$f^s(x, y) = f(x, y), \quad \text{if } (x, y) \in \Lambda^s, s = -, +.$$

We use $\overline{DE}$ to denote the line segment between two points $D, E \in \Omega$. Assume $\mathcal{T}_h, h > 0$ is a family of rectangular meshes of the solution domain Ω that can be a union of rectangles.

The mesh consists of interface elements whose interiors are cut through by the interfaces and the rest called non-interface elements. We use standard bilinear finite element functions in all the non-interface elements. Special piecewise bilinear polynomials satisfying interface jump conditions are employed only in interface elements as follows [112, 149].

We consider a typical rectangle element $T \in \mathcal{T}_h$. Assume that the four vertices of T are $A_i, i = 1, 2, 3, 4$, with $A_i = (x_i, y_i)^t$. If T is an interface element, then we use $D = (x_{_D}, y_{_D})^T$ and $E = (x_{_E}, y_{_E})^T$ to denote the interface points on its edges. When the mesh is fine enough, there are two types of rectangular interface elements. Type I are those for which the interface intersects with two of its adjacent edges; Type II are those for which the interface intersects with two of its opposite edges, see the sketch in Figure 3.1.

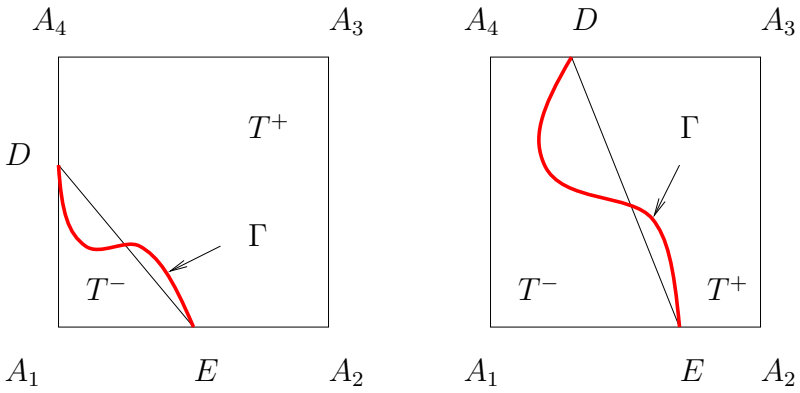

Figure 3.1: Two typical interface elements. The element on the left is of Type I while the one on the right is of Type II.

Our main concern is the finite element functions on an interface rectangle $T \in \mathcal{T}_h$. For our interface problems, the interface Γ divides an interface element T into two subsets T^- and T^+, we naturally can try to form a piecewise function by two bilinear polynomials defined in T^- and T^+, respectively. The challenge is obviously how to put them together so that the jump conditions across the interface are maintained.

Note that each bilinear polynomial has four freedoms (coefficients). The values of the finite element function at the vertices of T provide four restrictions. The normal derivative jump condition on $\overline{DE}$ provides another. Then we can have three more restrictions by requiring the continuity of the finite element function at interface points D,E and $\frac{D+E}{2}$. Intuitively, these eight conditions can yield the desired piecewise bilinear polynomial in an interface

rectangle. This idea leads us to consider the bilinear IFE functions defined as follows:

$$\phi(x,y)=\begin{cases}\phi^-(x,y)=a^-x+b^-y+c^-+d^-xy, & (x,y)\in T^-,\\ \phi^+(x,y)=a^+x+b^+y+c^++d^+xy, & (x,y)\in T^+,\\ \phi^-(D)=\phi^+(D),\quad \phi^-(E)=\phi^+(E),\\ \phi^-(\frac{D+E}{2})=\phi^+(\frac{D+E}{2}),\\ \int_{\overline{DE}}\left(\beta^-\frac{\partial\phi^-}{\partial\mathbf{n}_{\overline{DE}}}-\beta^+\frac{\partial\phi^+}{\partial\mathbf{n}_{\overline{DE}}}\right)ds=0,\end{cases} \tag{3.1}$$

where $\mathbf{n}_{\overline{DE}}$ is the unit vector perpendicular to the line $\overline{DE}$. This is the definition of bilinear IFE functions introduced in [112, 149].

However, when $\overline{DE}$ is vertical or horizontal, $\phi^-(D)=\phi^+(D)$ and $\phi^-(E)=\phi^+(E)$ automatically imply $\phi^-(\frac{D+E}{2})=\phi^+(\frac{D+E}{2})$, hence the bilinear IFE functions cannot be uniquely defined. This was observed by M. L. Ryan and T. Lin before, but they showed that the above definition can still uniquely determine the bilinear IFE functions in a limit sense. Here, we directly propose one more condition to replace $\phi^-(\frac{D+E}{2})=\phi^+(\frac{D+E}{2})$ so that the piecewise bilinear IFE functions can be uniquely determined.

In fact, for general cases in which $\overline{DE}$ is neither vertical nor horizontal, we can show the following lemma, which indicates that the mixed 2nd derivative of a bilinear IFE function is continuous.

Lemma 3.1.1

$$d^- \;=\; d^+, \tag{3.2}$$

Proof. Plugging $\phi^-(D)=\phi^+(D)$ and $\phi^-(E)=\phi^+(E)$ into $\phi^-(\frac{D+E}{2})=\phi^+(\frac{D+E}{2})$, we finish the proof.

■

Also, in Chapter 4, $d^-=d^+$ will be one of the key tools for the interpolation error estimation. Therefore, we use $d^-=d^+$ to replace $\phi^-(\frac{D+E}{2})=\phi^+(\frac{D+E}{2})$ to define the bilinear IFE functions as follows:

$$\phi(x,y)=\begin{cases}\phi^-(x,y)=a^-x+b^-y+c^-+d^-xy, & (x,y)\in T^-,\\ \phi^+(x,y)=a^+x+b^+y+c^++d^+xy, & (x,y)\in T^+,\\ \phi^-(D)=\phi^+(D),\quad \phi^-(E)=\phi^+(E),\quad d^-=d^+,\\ \int_{\overline{DE}}\left(\beta^-\frac{\partial\phi^-}{\partial\mathbf{n}_{\overline{DE}}}-\beta^+\frac{\partial\phi^+}{\partial\mathbf{n}_{\overline{DE}}}\right)ds=0.\end{cases} \tag{3.3}$$

In the above discussion, we use the original curve Γ to separate the two pieces T^- and T^+. Actually, we can also use $\overline{DE}$, the linear approximation to Γ, to separate them. Suppose the

line $\overline{DE}$ separates T into two subelements: $\tilde{T}^-$ and $\tilde{T}^+$. Here $\tilde{T}^-$ is the polygon contained in T sharing at least one vertex of T^- inside Ω^- and the other subelement is $\tilde{T}^+$. Then we can define the bilinear IFE functions as follows:

$$\phi(x,y) = \begin{cases} \phi^-(x,y) = a^-x + b^-y + c^- + d^-xy, & (x,y) \in \tilde{T}^-, \\ \phi^+(x,y) = a^+x + b^+y + c^+ + d^+xy, & (x,y) \in \tilde{T}^+, \\ \phi^-(D) = \phi^+(D), \quad \phi^-(E) = \phi^+(E), \quad d^- = d^+, & \\ \int_{\overline{DE}} \left(\beta^- \frac{\partial \phi^-}{\partial \mathbf{n}_{\overline{DE}}} - \beta^+ \frac{\partial \phi^+}{\partial \mathbf{n}_{\overline{DE}}} \right) ds = 0. & \end{cases} \tag{3.4}$$

In both (3.3) and (3.4), we obtain the same formulas $\phi^-(x,y)$ and $\phi^+(x,y)$. The only difference is the domain of the piecewise function. The two ways have their own advantages, see Section 3.3 for more detail. We would like to note that we now prefer to use $\overline{DE}$ to separate the two pieces even though we use Γ in [112, 113, 149].

We let $\phi_i(X)$ be the piecewise linear function described by (3.4) such that

$$\phi_i(x_j, y_j) = \begin{cases} 1, & \text{if } i = j, \\ 0, & \text{if } i \neq j, \end{cases}$$

for $1 \leq i, j \leq 4$, and we call them the bilinear IFE nodal basis functions on an interface element T.

Now we use the mesh $\mathcal{T}_h$ to define the bilinear immersed finite element (IFE) space $S_h(\Omega)$. First, for every element $T \in \mathcal{T}_h$, we let $S_h(T) = span\{\phi_i,\, i = 1,2,3,4\}$, where ϕ_i, $i = 1,2,3,4$ are the standard bilinear nodal basis functions for a non-interface element T; otherwise, ϕ_i, $i = 1,2,3,4$ are the immersed bilinear basis functions defined above. Then, we define a continuous piecewise bilinear global nodal basis function $\phi_N(x,y)$ for each node $(x_N, y_N)^t$ of $\mathcal{T}_h$ such that $\phi_N(x_N, y_N) = 1$ but zero at other nodes, and $\phi_N|_T \in S_h(T)$ for any rectangle $T \in \mathcal{T}_h$. Finally, we define $S_h(\Omega)$ as the span of these global nodal basis functions.

3.2 The existence and uniqueness of the bilinear IFE basis functions

In this section, we discuss the existence and uniqueness of the bilinear IFE basis functions. They were briefly proved in [149] and we will show more details in this section. As usual, we only need to define the nodal bilinear IFE basis functions $\hat{\phi}_i(\hat{X})$, $i = 1,2,3,4$ in the reference element $\hat{T}$ with vertices $\hat{A}_i = (\hat{x}_i, \hat{y}_i)^T$, $i = 1,2,3,4$:

$$\hat{A}_1 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \hat{A}_2 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \hat{A}_3 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \hat{A}_4 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

The interface element T is related to the reference element by the usual affine mapping:

$$X = F(\hat{X}) = B + M\hat{X}, \quad X = \begin{pmatrix} x \\ y \end{pmatrix}, \quad \hat{X} = \begin{pmatrix} \hat{x} \\ \hat{y} \end{pmatrix}. \tag{3.5}$$

This affine mapping and $\hat{\phi}_i(\hat{X})$, $i = 1, 2, 3, 4$ can be used to defined $\phi_i(X)$, $i = 1, 2, 3, 4$ through the standard procedure.

Assume that under the affine mapping $\Gamma \cap T$ becomes $\hat{\Gamma}$, D becomes $\hat{D}$, E becomes $\hat{E}$, and $\overline{\hat{D}\hat{E}}$ separates $\hat{T}$ into $\hat{\hat{T}}^+$ and $\hat{\hat{T}}^-$. Accordingly, there are two types of reference interface elements, see the sketch in Figure 3.2.

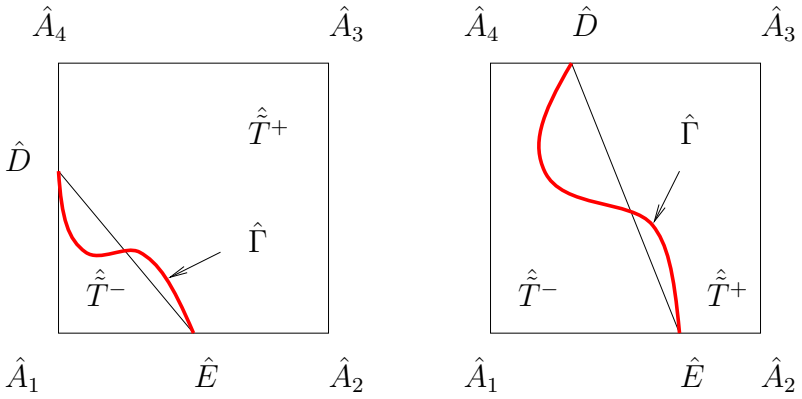

Figure 3.2: Two reference interface elements. The element on the left is of Type I while the one on the right is of Type II.

By choosing proper B and M, i.e., a proper affine mapping, we can assume

$$\hat{D} = \begin{pmatrix} 0 \\ \hat{b} \end{pmatrix}, \hat{E} = \begin{pmatrix} \hat{a} \\ 0 \end{pmatrix}$$

for Type I reference element, and

$$\hat{D} = \begin{pmatrix} \hat{b} \\ 1 \end{pmatrix}, \hat{E} = \begin{pmatrix} \hat{a} \\ 0 \end{pmatrix}$$

for Type II reference element. Obviously, we can assume $0 < \hat{a}, \hat{b} \leq 1$ for Type I reference element, and $0 \leq \hat{a}, \hat{b} \leq 1$ for Type II reference element.

Assume $\hat{\phi}_i(\hat{X}), i = 1, 2, 3, 4$ are the bilinear IFE nodal basis on the reference element $\hat{T}$ such that

$$\hat{\phi}_i(\hat{x}, \hat{y}) = \begin{cases} \hat{a}_i^- + \hat{b}_i^- \hat{x} + \hat{c}_i^- \hat{y} + \hat{d}_i^- \hat{x}\hat{y}, & \text{if } (\hat{x}, \hat{y}) \in \hat{\hat{T}}^-, \\ \hat{a}_i^+ + \hat{b}_i^+ \hat{x} + \hat{c}_i^+ \hat{y} + \hat{d}_i^+ \hat{x}\hat{y}, & \text{if } (\hat{x}, \hat{y}) \in \hat{\hat{T}}^+. \end{cases}$$

Then $\hat{\phi}_i(\hat{X}), i = 1, 2, 3, 4$ should satisfy

$$\begin{cases} \hat{\phi}_i(\hat{A}_j) = \begin{cases} 1, & \text{if } i = j, \\ 0, & \text{if } i \neq j, \end{cases} \\ \hat{\phi}_i^-(\hat{D}) = \hat{\phi}_i^+(\hat{D}), \quad \hat{\phi}_i^-(\hat{E}) = \hat{\phi}_i^+(\hat{E}), \quad \hat{d}_i^- = \hat{d}_i^+ \\ \int_{\overline{\hat{D}\hat{E}}} \left(\beta^+ \frac{\partial \hat{\phi}_i^-}{\partial \mathbf{n}_{\overline{\hat{D}\hat{E}}}} - \beta^- \frac{\partial \hat{\phi}_i^+}{\partial \mathbf{n}_{\overline{\hat{D}\hat{E}}}} \right) ds = 0. \end{cases} \tag{3.6}$$

Basis functions in Type I elements: For Type I element, we first note that the nodal value constraints at $\hat{A}_1$ imply that

$$\hat{a}_i^- = \begin{cases} 1, & i = 1, \\ 0, & i = 2, 3, 4. \end{cases}$$

Also, the nodal value constraints at $\hat{A}_i$, $i = 2, 3, 4$ allow us to express $\hat{b}_i^+, \hat{c}_i^+$ and $\hat{d}_i^+$ as linear functions of $\hat{a}_i^+$. Then, the four conditions across the interface lead to a linear system about $\hat{b}_i^-, \hat{c}_i^-, \hat{d}_i^-, \hat{a}_i^+$. Solving this linear system, we can see that

$$\hat{b}_i^- = \frac{P_{i,1}(\hat{a}, \hat{b})}{W}, \hat{c}_i^- = \frac{P_{i,2}(\hat{a}, \hat{b})}{W}, \hat{d}_i^- = \frac{P_{i,3}(\hat{a}, \hat{b})}{W}, \hat{a}_i^+ = \frac{P_{i,4}(\hat{a}, \hat{b})}{W},$$

$$W = \begin{cases} \Big[\hat{a}\hat{b}^2(2 - \hat{b}) + \hat{a}^2\hat{b}(2 - \hat{a})\Big] \\ +R\Big[2\hat{b}^2(1 - \hat{a}) + 2\hat{a}^2(1 - \hat{b}) + \hat{a}^3\hat{b} + \hat{a}\hat{b}^3\Big], & \text{if } R = \beta^-/\beta^+ \geq 1, \\ \\ R\Big[\hat{a}\hat{b}^2(2 - \hat{b}) + \hat{a}^2\hat{b}(2 - \hat{a})\Big] \\ +\Big[2\hat{b}^2(1 - \hat{a}) + 2\hat{a}^2(1 - \hat{b}) + \hat{a}^3\hat{b} + \hat{a}\hat{b}^3\Big], & \text{if } R = \beta^+/\beta^- \geq 1, \end{cases}$$

where $P_{i,j}(\hat{a}, \hat{b}), j = 1, 2, 3, 4$ are polynomials of $\hat{a}$ and $\hat{b}$. Moreover, $P_{i,j}(\hat{a}, \hat{b}), j = 1, 2, 3, 4$ are linear combinations of the following terms:

$$\hat{a}^2, \hat{b}^2, \hat{a}\hat{b}, \hat{a}^3, \hat{b}^3, \hat{a}^2\hat{b}, \hat{a}\hat{b}^2, \hat{a}^3\hat{b}, \hat{a}\hat{b}^3. \tag{3.7}$$

In fact, the whole 8×8 linear system arising from (3.6) for the Type I reference interface element is the following.

$$A \begin{pmatrix} \hat{a}_i^- \\ \hat{b}_i^- \\ \hat{c}_i^- \\ \hat{d}_i^- \\ \hat{a}_i^+ \\ \hat{b}_i^+ \\ \hat{c}_i^+ \\ \hat{d}_i^+ \end{pmatrix} = \mathbf{b_i},$$

where

$$A = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & \hat{b} & 0 & -1 & 0 & -\hat{b} & 0 \\ 1 & \hat{a} & 0 & 0 & -1 & -\hat{a} & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & -1 \\ 0 & R\hat{b} & R\hat{a} & R\frac{\hat{a}^2+\hat{b}^2}{2} & 0 & -\hat{b} & -\hat{a} & -\frac{\hat{a}^2+\hat{b}^2}{2} \end{pmatrix},$$

and

$$\mathbf{b_1} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \mathbf{b_2} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \mathbf{b_3} = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \mathbf{b_4} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

Here, $R = \beta^-/\beta^+$. Direct calculations give us

$$\begin{aligned} det(A) &= -\frac{1}{2}\hat{a}^3\hat{b} + \frac{1}{2}R\hat{a}^3\hat{b} + \hat{a}^2\hat{b} - R\hat{a}^2\hat{b} + R\hat{a}^2 - \frac{1}{2}\hat{a}\hat{b}^3 + \frac{1}{2}R\hat{a}\hat{b}^3 + \hat{a}\hat{b}^2 - R\hat{a}\hat{b}^2 + R\hat{b}^2 \\ &= \hat{a}^2\hat{b}(1-\frac{1}{2}\hat{a}) + \hat{a}\hat{b}^2(1-\frac{1}{2}\hat{b}) + R\hat{a}^2(1-\hat{b}) + R\hat{b}^2(1-\hat{a}) + \frac{1}{2}R\hat{a}^3\hat{b} + \frac{1}{2}R\hat{a}\hat{b}^3 \\ &> 0, \end{aligned}$$

which shows that the matrix A is non-singular for all $\hat{a}, \hat{b} \in [0,1]$. Solving this linear system, we have

$$\hat{a}_i^- = \begin{cases} 1, & i = 1, \\ 0, & i = 2, \\ 0, & i = 3, \\ 0, & i = 4, \end{cases}$$

$$\hat{b}_i^- = \begin{cases} \dfrac{-(-2R\hat{b}\hat{a} + R\hat{a}^2\hat{b} + R\hat{b}^3 + 2\hat{b}^2 + 2\hat{a}\hat{b} - \hat{a}^2\hat{b} - \hat{b}^3 + 2R\hat{a}^2)}{W}, & i = 1, \\ \dfrac{-(R\hat{a}^2\hat{b} - R\hat{b}^3 - 2\hat{b}^2 - \hat{a}^2\hat{b} + \hat{b}^3 - 2R\hat{a}^2)}{W}, & i = 2, \\ \dfrac{\hat{b}(\hat{a}^2 + R\hat{b}^2\hat{a} + R\hat{a}^3 + \hat{b}^2 - R\hat{a}^2 - R\hat{b}^2 - \hat{a}^3 - \hat{a}\hat{b}^2)}{W}, & i = 3, \\ \dfrac{-\hat{b}(2R\hat{a} - 3R\hat{a}^2 - R\hat{b}^2 - 2\hat{a} + 3\hat{a}^2 + \hat{b}^2 + R\hat{a}^3 + R\hat{b}^2\hat{a} - \hat{a}^3 - \hat{b}^2\hat{a})}{W}, & i = 4, \end{cases}$$

$$\hat{c}_i^- = \begin{cases} \dfrac{-(-2R\hat{b}\hat{a} + R\hat{a}^3 + R\hat{b}^2\hat{a} + 2\hat{a}\hat{b} + 2\hat{a}^2 - \hat{a}^3 - \hat{b}^2\hat{a} + 2R\hat{b}^2)}{W}, & i = 1, \\ \dfrac{-\hat{a}(2R\hat{b} - R\hat{a}^2 - 3R\hat{b}^2 - 2\hat{b} + \hat{a}^2 + 3\hat{b}^2 + R\hat{a}^2\hat{b} + R\hat{b}^3 - \hat{a}^2\hat{b} - \hat{b}^3)}{W}, & i = 2, \\ \dfrac{\hat{a}(R\hat{a}^2\hat{b} + R\hat{b}^3 + \hat{a}^2 + \hat{b}^2 - R\hat{a}^2 - R\hat{b}^2 - \hat{a}^2\hat{b} - \hat{b}^3)}{W}, & i = 3, \\ \dfrac{R\hat{a}^3 - R\hat{b}^2\hat{a} + 2\hat{a}^2 - \hat{a}^3 + \hat{b}^2\hat{a} + 2R\hat{b}^2}{W}, & i = 4, \end{cases}$$

$$\hat{d}_i^- = \begin{cases} \dfrac{2R(\hat{a}^2 + \hat{b}^2)}{W}, & i = 1, \\ \dfrac{2(-R\hat{a}^2 - R\hat{b}^2 + R\hat{a}^2\hat{b} - \hat{a}^2\hat{b})}{W}, & i = 2, \\ \dfrac{-2(-R\hat{b}^2 + R\hat{b}^2\hat{a} - R\hat{a}^2 + R\hat{a}^2\hat{b} - \hat{b}^2\hat{a} - \hat{a}^2\hat{b})}{W}, & i = 3, \\ \dfrac{-2(R\hat{b}^2 + \hat{b}^2\hat{a} - R\hat{b}^2\hat{a} + R\hat{a}^2)}{W}, & i = 4, \end{cases}$$

$$\hat{a}_i^+ = \begin{cases} \dfrac{2R(\hat{a}^2 + \hat{b}^2)}{W}, & i = 1, \\ \dfrac{\hat{a}b(-2R\hat{b} + R\hat{a}^2 + R\hat{b}^2 + 2\hat{b} - \hat{a}^2 - \hat{b}^2)}{W}, & i = 2, \\ \dfrac{-\hat{a}\hat{b}(-\hat{b}^2 + R\hat{a}^2 + R\hat{b}^2 - \hat{a}^2)}{W}, & i = 3, \\ \dfrac{\hat{a}\hat{b}(-2R\hat{a} + R\hat{a}^2 + R\hat{b}^2 + 2\hat{a} - \hat{a}^2 - \hat{b}^2)}{W}, & i = 4, \end{cases}$$

$$\hat{b}_i^+ = \begin{cases} \dfrac{-2R(\hat{a}^2 + \hat{b}^2)}{W}, & i = 1, \\ \dfrac{-2(-R\hat{a}^2 - R\hat{b}^2 + R\hat{a}^2\hat{b} - \hat{a}^2\hat{b})}{W}, & i = 2, \\ \dfrac{\hat{a}\hat{b}(-\hat{b}^2 + R\hat{a}^2 + R\hat{b}^2 - \hat{a}^2)}{W}, & i = 3, \\ \dfrac{-\hat{a}\hat{b}(-2R\hat{a} + R\hat{a}^2 + R\hat{b}^2 + 2\hat{a} - \hat{a}^2 - \hat{b}^2)}{W}, & i = 4, \end{cases}$$

$$\hat{c}_i^+ = \begin{cases} \dfrac{-2R(\hat{a}^2 + \hat{b}^2)}{W}, & i = 1, \\ \dfrac{-\hat{a}\hat{b}(-2R\hat{b} + R\hat{a}^2 + R\hat{b}^2 + 2\hat{b} - \hat{a}^2 - \hat{b}^2)}{W}, & i = 2, \\ \dfrac{\hat{a}\hat{b}(-\hat{b}^2 + R\hat{a}^2 + R\hat{b}^2 - \hat{a}^2)}{W}, & i = 3, \\ \dfrac{2(R\hat{b}^2 + \hat{b}^2\hat{a} - R\hat{b}^2\hat{a} + R\hat{a}^2)}{W}, & i = 4, \end{cases}$$

$$\hat{d}_i^+ = \begin{cases} \dfrac{2R(\hat{a}^2+\hat{b}^2)}{W}, & i=1, \\ \dfrac{2(-R\hat{a}^2 - R\hat{b}^2 + R\hat{a}^2\hat{b} - a^2\hat{b})}{W}, & i=2, \\ \dfrac{-2(-R\hat{b}^2 + R\hat{b}^2\hat{a} - R\hat{a}^2 + R\hat{a}^2\hat{b} - \hat{b}^2\hat{a} - \hat{a}^2\hat{b})}{W}, & i=3, \\ \dfrac{-2(R\hat{b}^2 + b^2\hat{a} - R\hat{b}^2\hat{a} + R\hat{a}^2)}{W}, & i=4. \end{cases}$$

Basis functions in Type II elements: Similarly, the nodal value constraints require that

$$\hat{a}_i^- = \begin{cases} 1, & i=1, \\ 0, & i=2,3,4. \end{cases} \quad \hat{c}_i^- = \begin{cases} -1, & i=1, \\ 0, & i=2,3, \\ 1, & i=4. \end{cases}$$

Also, the nodal value constraints imply that $\hat{b}_i^+, \hat{c}_i^+$ are linear functions of $\hat{a}_i^+$ and $\hat{d}_i^+$. Then, the conditions across the interface lead to a linear system about $\hat{b}_i^-, \hat{d}_i^-, \hat{a}_i^+$ and $\hat{d}_i^+$. Solving this linear system, we have

$$\hat{b}_i^- = \frac{P_{i,1}(\hat{a},\hat{b})}{W}, \hat{c}_i^- = \frac{P_{i,2}(\hat{a},\hat{b})}{W}, \hat{d}_i^- = \frac{P_{i,3}(\hat{a},\hat{b})}{W}, \hat{a}_i^+ = \frac{P_{i,4}(\hat{a},\hat{b})}{W},$$

$$W = \begin{cases} \text{if } R = \beta^-/\beta^+ \geq 1: \\ \Big[\hat{a}(1-\hat{a}^2) + \hat{b}(1-\hat{b}^2) + 2(\hat{a}-\hat{b})^2 + \hat{a}^2\hat{b} + \hat{a}\hat{b}^2\Big] + R\Big[(2-\hat{a}-\hat{b}) + (\hat{a}+\hat{b})(\hat{a}-\hat{b})^2\Big], \\ \text{if } R = \beta^+/\beta^- \geq 1: \\ R\Big[\hat{a}(1-\hat{a}^2) + \hat{b}(1-\hat{b}^2) + 2(\hat{a}-\hat{b})^2 + \hat{a}^2\hat{b} + \hat{a}\hat{b}^2\Big] + \Big[(2-\hat{a}-\hat{b}) + (\hat{a}+\hat{b})(\hat{a}-\hat{b})^2\Big], \end{cases}$$

where $P_{i,j}(\hat{a},\hat{b}), j=1,2,3,4$ are polynomials of $\hat{a}$ and $\hat{b}$. Moreover, $P_{i,j}(\hat{a},\hat{b}), j=1,2,3,4$ are linear combinations of the following terms:

$$1, \hat{a}, \hat{b}, \hat{a}^2, \hat{b}^2, \hat{a}\hat{b}, \hat{a}^3, \hat{b}^3, \hat{a}^2\hat{b}, \hat{a}\hat{b}^2, \hat{a}^3\hat{b}, \hat{a}\hat{b}^3. \tag{3.8}$$

In fact, the whole 8×8 linear system arising from (3.6) for the Type II reference interface element is the following.

$$A \begin{pmatrix} \hat{a}_i^- \\ \hat{b}_i^- \\ \hat{c}_i^- \\ \hat{d}_i^- \\ \hat{a}_i^+ \\ \hat{b}_i^+ \\ \hat{c}_i^+ \\ \hat{d}_i^+ \end{pmatrix} = \mathbf{b_i},$$

where

$$A=\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & \hat{b} & 1 & \hat{b} & -1 & -\hat{b} & -1 & -\hat{b} \\ 1 & \hat{a} & 0 & 0 & -1 & -\hat{a} & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & -1 \\ 0 & R & R(\hat{a}-\hat{b}) & R\frac{1+\hat{a}^2-\hat{b}^2}{2} & 0 & -1 & -(\hat{a}-\hat{b}) & -\frac{1+\hat{a}^2-\hat{b}^2}{2} \end{pmatrix},$$

and

$$\mathbf{b_1}=\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \mathbf{b_2}=\begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \mathbf{b_3}=\begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \mathbf{b_4}=\begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

Here, $R=\beta^-/\beta^+$. Direct calculations give us

$$\begin{aligned} det(A) &= -\frac{1}{2}\hat{a}^3+\frac{1}{2}R\hat{a}^3-\frac{1}{2}R\hat{a}^2\hat{b}+\frac{1}{2}\hat{a}^2\hat{b}+\hat{a}^2-2\hat{a}\hat{b}-\frac{1}{2}R\hat{a}+\frac{1}{2}\hat{a}-\frac{1}{2}R\hat{a}\hat{b}^2+\frac{1}{2}\hat{a}\hat{b}^2 \\ &\quad +\frac{1}{2}R\hat{b}^3+R-\frac{1}{2}\hat{b}^3+\hat{b}^2-\frac{1}{2}R\hat{b}+\frac{1}{2}\hat{b} \\ &= \frac{1}{2}\hat{a}(1-\hat{a}^2)+\frac{1}{2}\hat{b}(1-\hat{b}^2)+(\hat{a}-\hat{b})^2+\frac{1}{2}\hat{a}^2\hat{b}+\frac{1}{2}\hat{a}\hat{b}^2+\frac{1}{2}R(2-\hat{a}-\hat{b}) \\ &\quad +\frac{1}{2}R(\hat{a}-\hat{b})^2(\hat{a}+\hat{b}) \\ &> 0, \end{aligned}$$

which shows that the matrix A is non-singular for all $\hat{a},\hat{b}\in[0,1]$. Solving this linear system, we have

$$\hat{a}_i^- = \begin{cases} 1, & i=1, \\ 0, & i=2, \\ 0, & i=3, \\ 0, & i=4, \end{cases}$$

$$\hat{b}_i^- = \begin{cases} \underline{\frac{\hat{b}R\hat{a}^2 - 3R\hat{a}^2 - \hat{b}\hat{a}^2 + \hat{a}^2 + 2\hat{a}R\hat{b} + 2R\hat{a} - 2\hat{a} + 2\hat{a}\hat{b} - R}{W}} \\ +\frac{-R\hat{b} + R\hat{b}^2 - R\hat{b}^3 + \hat{b}^3 - 3\hat{b}^2 + \hat{b} - 1}{W}, & i = 1, \\ \frac{-(-R\hat{a}^2 + \hat{b}R\hat{a}^2 - \hat{b}\hat{a}^2 - \hat{a}^2 + 4\hat{a}\hat{b} + R\hat{b}^2 - R\hat{b}^3 + R\hat{b} - R + \hat{b}^3 - \hat{b} - 3\hat{b}^2 - 1)}{W}, & i = 2, \\ \frac{-\hat{a}^3 + R\hat{a}^3 - R\hat{a}^2 + \hat{a}^2 + R\hat{a} - \hat{a}R\hat{b}^2 - \hat{a} + \hat{a}\hat{b}^2 - R + R\hat{b}^2 + 1 - \hat{b}^2}{W}, & i = 3, \\ \underline{\frac{-(-\hat{a}^3 + R\hat{a}^3 - 3R\hat{a}^2 + 3\hat{a}^2 + 3R\hat{a} + 2\hat{a}R\hat{b} - \hat{a}R\hat{b}^2 - 3\hat{a} - 2\hat{a}\hat{b} + \hat{a}\hat{b}^2 - R)}{W}} \\ +\frac{-(R\hat{b}^2 - 2R\hat{b} + 1 - \hat{b}^2 + 2\hat{b})}{W}, & i = 4, \end{cases}$$

$$\hat{c}_i^- = \begin{cases} -1, & i = 1, \\ 0, & i = 2, \\ 0, & i = 3, \\ 1, & i = 4, \end{cases}$$

$$\hat{d}_i^- = \begin{cases} \frac{2(R\hat{a}^2 - 2\hat{a}R\hat{b} + R - R\hat{b} + R\hat{b}^2 + \hat{b})}{W}, & i = 1, \\ \frac{-2(\hat{a}^2 - 2\hat{a}\hat{b} + R - R\hat{b} + \hat{b}^2 + \hat{b})}{W}, & i = 2, \\ \frac{2(\hat{a}^2 - R\hat{a} + \hat{a} - 2\hat{a}\hat{b} + R + \hat{b}^2)}{W}, & i = 3, \\ \frac{-2(R\hat{a}^2 - R\hat{a} - 2\hat{a}R\hat{b} + \hat{a} + R\hat{b}^2 + R)}{W}, & i = 4, \end{cases}$$

$$\hat{a}_i^+ = \begin{cases} \frac{-(-\hat{b} + 2\hat{a}R\hat{b} - 2R\hat{a}^2 + R\hat{b} - 2\hat{b}^2 - 2R - \hat{b}\hat{a}^2 - R\hat{b}^3 + 2\hat{a}\hat{b} + \hat{b}^3 + \hat{b}R\hat{a}^2)}{W}, & i = 1, \\ \frac{\hat{a}(1 - R + R\hat{a}^2 + \hat{b}^2 - \hat{a}^2 - R\hat{b}^2)}{W}, & i = 2, \\ \frac{-\hat{a}(-1 + R + R\hat{a}^2 + \hat{b}^2 - \hat{a}^2 - R\hat{b}^2)}{W}, & i = 3, \\ \frac{\hat{a}(-1 + 2R\hat{b} + R - 2R\hat{a} + 2\hat{a} + R\hat{a}^2 - 2\hat{b} + \hat{b}^2 - \hat{a}^2 - R\hat{b}^2)}{W}, & i = 4, \end{cases}$$

$$\hat{b}_i^+ = \begin{cases} \frac{-\hat{b} + 2\hat{a}R\hat{b} - 2R\hat{a}^2 + R\hat{b} - 2\hat{b}^2 - 2R - \hat{b}\hat{a}^2 - R\hat{b}^3 + 2\hat{a}\hat{b} + \hat{b}^3 + \hat{b}R\hat{a}^2}{W}, & i = 1, \\ \frac{-(-\hat{b} + R\hat{b} - 2\hat{a}^2 - 2\hat{b}^2 - 2R - \hat{b}\hat{a}^2 - R\hat{b}^3 + 4\hat{a}\hat{b} + \hat{b}^3 + \hat{b}R\hat{a}^2)}{W}, & i = 2, \\ \frac{\hat{a}(-1 + R + R\hat{a}^2 + \hat{b}^2 - \hat{a}^2 - R\hat{b}^2)}{W}, & i = 3, \\ \frac{-\hat{a}(-1 + 2R\hat{b} + R - 2R\hat{a} + 2\hat{a} + R\hat{a}^2 - 2\hat{b} + \hat{b}^2 - \hat{a}^2 - R\hat{b}^2)}{W}, & i = 4, \end{cases}$$

$$\hat{c}_i^+ = \begin{cases} \dfrac{-2(R\hat{a}^2 - 2\hat{a}R\hat{b} + R - R\hat{b} + R\hat{b}^2 + \hat{b})}{W}, & i = 1, \\ \dfrac{-(R\hat{a}^3 - \hat{a}^3 - \hat{b}R\hat{a}^2 + \hat{b}\hat{a}^2 - \hat{a}R\hat{b}^2 - R\hat{a} + \hat{a} + \hat{a}\hat{b}^2 + R\hat{b}^3 + R\hat{b} - \hat{b}^3 - \hat{b})}{W}, & i = 2, \\ \dfrac{R\hat{a} - \hat{a} + R\hat{a}^3 - \hat{a}^3 - \hat{b}R\hat{a}^2 + \hat{b}\hat{a}^2 - aR\hat{b}^2 + \hat{a}\hat{b}^2 + R\hat{b}^3 - R\hat{b} - \hat{b}^3 + \hat{b}}{W}, & i = 3, \\ \dfrac{2(R\hat{a}^2 - R\hat{a} - 2\hat{a}R\hat{b} + \hat{a} + R\hat{b}^2 + R)}{W}, & i = 4, \end{cases}$$

$$\hat{d}_i^+ = \begin{cases} \dfrac{2(R\hat{a}^2 - 2\hat{a}R\hat{b} + R - R\hat{b} + R\hat{b}^2 + \hat{b})}{W}, & i = 1, \\ \dfrac{-2(\hat{a}^2 - 2\hat{a}\hat{b} + R - R\hat{b} + \hat{b}^2 + \hat{b})}{W}, & i = 2, \\ \dfrac{2(\hat{a}^2 - R\hat{a} + \hat{a} - 2\hat{a}\hat{b} + R + \hat{b}^2)}{W}, & i = 3, \\ \dfrac{-2(R\hat{a}^2 - R\hat{a} - 2\hat{a}R\hat{b} + \hat{a} + R\hat{b}^2 + R)}{W}, & i = 4. \end{cases}$$

3.3 Some comments on immersed finite element spaces

The word "immersed" is used for this kind of finite element space just to emphasize the fact that the mesh can be independent of interface such that the interface can be immersed inside elements of this mesh. Figure 3.3 illustrates the difference between a bilinear IFE local nodal basis function and a standard bilinear local nodal basis function.

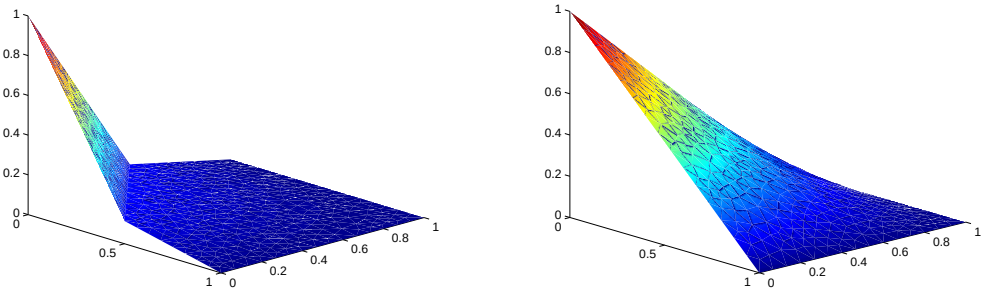

Figure 3.3: The plot on the left is for one of the bilinear IFE local nodal basis functions, the plot on the right is the corresponding regular bilinear local nodal basis function on the same element.

In the following we discuss the difference between the two ways to separate the two pieces which are mentioned in Section 3.1. See Figure 3.4 for two typical bilinear IFE basis functions of Type I and Type II whose two pieces are separated by $\overline{DE}$ and see Figure 3.5 for two typical bilinear IFE basis functions of Type I and Type II whose two pieces are separated

by the original interface Γ. Note that the basis functions defined by Γ are discontinuous inside the element, but the basis functions defined by $\overline{DE}$ are continuous inside the element. Computationally we always use $\overline{DE}$ to locate the two pieces since it's more convenient and doesn't decrease the accuracy order. Theoretically, the two different ways to locate the two pieces of an IFE function ϕ have their own advantages. It is more natural to use Γ to define the two pieces since the original problem domain is separated by the Γ. However, the definition of ϕ using $\overline{DE}$ gives us the continuity of the basis function and its flux inside the whole element, which are important properties to be used in the convergence analysis in Chapter 5 and Chapter 8. The error estimation for the bilinear IFE interpolation in Chapter 4 can go through for both of the two kinds of definitions. Therefore, we now prefer to use $\overline{DE}$ to separate the two pieces even though we use Γ in [112, 113]. We would like to note that for all the other 2D or 3D IFE spaces, this issue can be handled similarly.

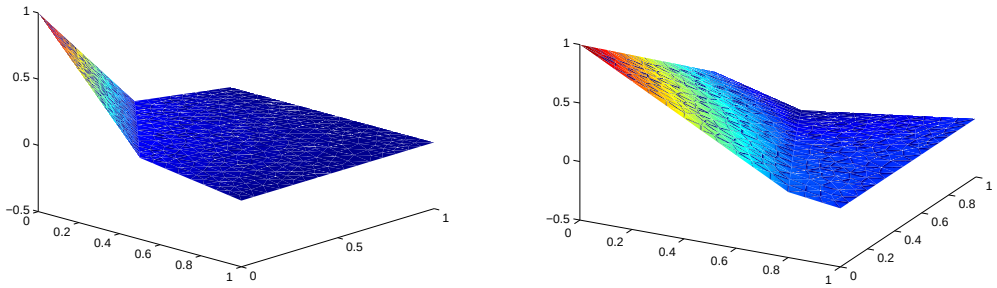

Figure 3.4: The plot on the left is a bilinear IFE basis on a Type I interface element and the plot on the right is a bilinear IFE basis on a Type II interface element. Both of them use $\overline{DE}$ to separate the two pieces.

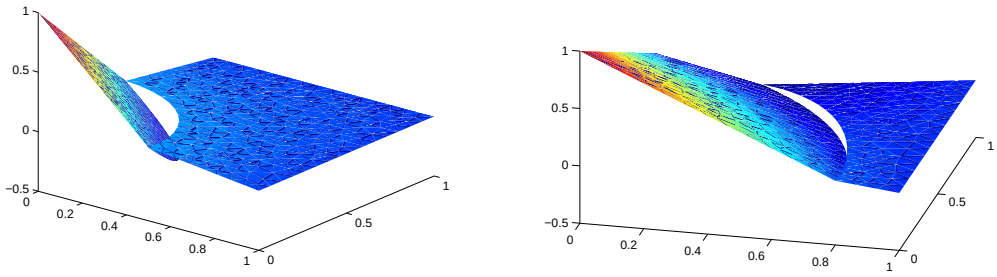

Figure 3.5: The plot on the left is a typical bilinear IFE basis on a Type I interface element and the plot on the right is a typical bilinear IFE basis on a Type II interface element. Both of them use the original interface curve to separate the two pieces.

Finally we would like to point out the following fact. If a node X_N is a vertex of at least one interface element, then its corresponding global nodal basis function $\phi_N \in S_h(\Omega)$ is always discontinuous on the elements edges cut by the interface, no matter which way mentioned above is chosen to separate the two pieces of the immersed basis functions on interface elements. Therefore, immersed finite element spaces are non-conforming finite element spaces. Figure 3.6 provides a sketch of the surface of a global bilinear IFE basis function over its support, from which we can see its discontinuity on the edge cut by the interface. Figure 3.7 illustrates the difference between a typical bilinear IFE global nodal basis function and a standard bilinear global nodal basis functions.

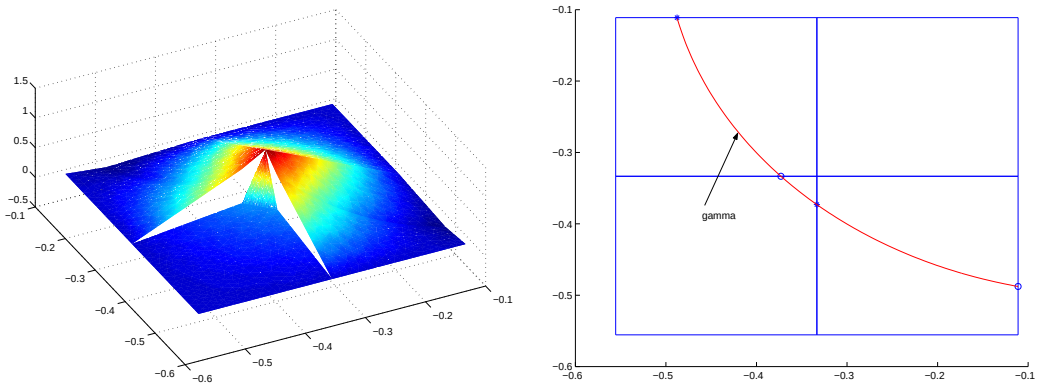

Figure 3.6: The plot on the left is the surface of one global bilinear IFE basis over its support, the plot on the right shows the elements forming the support and the interface.

3.4 Basic properties of the bilinear IFE space

In this section, we present basic properties of the bilinear IFE space discussed in [110, 111, 112, 113, 149]. It is easy to see that $S_h(\Omega)$ has the following properties [112, 149]:

- The IFE space $S_h(\Omega)$ has the same number of nodal basis functions as that formed by the usual bilinear polynomials on the same partition of Ω.
- For a mesh $\mathcal{T}_h$ fine enough, most of its rectangles are non-interface rectangles, and most of the nodal basis functions of the IFE space $S_h(\Omega)$ are just the usual bilinear nodal basis functions except for few nodes in the vicinity of the interface Γ.
- For any $\phi \in S_h(\Omega)$, we have

$$\phi|_{\Omega\backslash\Omega'} \in H^1(\Omega\backslash\Omega'),$$

  where Ω' is the union of interface rectangles.

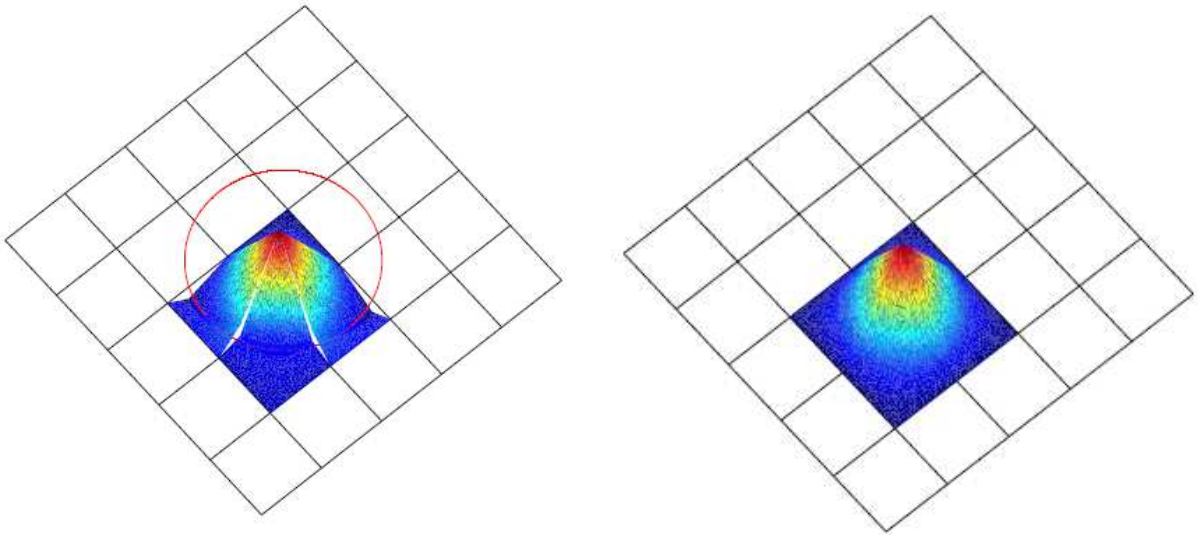

Figure 3.7: The plot on the left is for a bilinear IFE global nodal basis function, the plot on the right is the corresponding regular bilinear global nodal basis function.

In the discussion from now on, we denote any $v(x,y) \in S_h(T)$ as follows

$$v(x,y) = \begin{cases} v^-(x,y) = a^- + b^- x + c^- y + d^- xy, & (x,y) \in \tilde{T}^-, \\ v^+(x,y) = a^+ + b^+ x + c^+ y + d^+ xy, & (x,y) \in \tilde{T}^+. \end{cases} \tag{3.9}$$

The results in the following two lemmas are related to the continuity of functions in $S_h(\Omega)$ across the interface. First, the following lemma shows that every function $v \in S_h(T)$ where T is an interface element is continuous across $\overline{DE}$.

Lemma 3.4.1 *For any $v \in S_h(T)$ written as (3.9), we have*

$$v^- \equiv v^+ \quad on \quad \overline{DE}. \tag{3.10}$$

Proof. Note that any $v \in S_h(T)$ is continuous at D, E and $\frac{D+E}{2}$. Since each piece of v is a bilinear polynomial , then v is continuous on $\overline{DE}$. Hence $v^- \equiv v^+$ on $\overline{DE}$.

■

Lemma 3.4.2 *Assume $T \in \mathcal{T}_h$ is an interface element.*

1. *If $\Gamma \cap T$ is a line segment, then*

$$\phi^-|_{\Gamma\cap T} = \phi^+|_{\Gamma\cap T}, \quad \forall \phi \in S_h(\Omega).$$

2. *Every function $\phi \in S_h(T)$ satisfies the flux jump condition on $\Gamma \cap T$ exactly in the following weak sense:*

$$\int_{\Gamma\cap T} (\beta^-\nabla\phi^- - \beta^+\nabla\phi^+)\cdot \mathbf{n}_\Gamma ds = 0.$$

Proof. Property 1 follows directly from (3.10). For any $\phi \in S_h(T)$, it is obvious that $\phi^s \in H^2(T^s), \quad s = -, +$. Also, because ϕ is a piecewise bilinear polynomial satisfying (3.1), Green's formula leads to

$$\int_{\Gamma\cap T} (\beta^-\nabla\phi^- - \beta^+\nabla\phi^+)\cdot \mathbf{n}_\Gamma ds \;\; = \;\; -\int_{\overline{DE}} (\beta^-\nabla\phi^- - \beta^+\nabla\phi^+)\cdot \mathbf{n}_{\overline{DE}} ds = 0.$$

■

As stated in the following theorem, the local basis functions in this bilinear IFE space has the property of partition of unity.

Theorem 3.4.1 *Let $T \in \mathcal{T}_h$ be an interface element and let $\phi_i(X) \in S_h(T)$, $i = 1, 2, 3, 4$ be the local bilinear IFE basis functions defined above. Then,*

$$\phi_1(X) + \phi_2(X) + \phi_3(X) + \phi_4(X) = 1, \forall X \in T.$$

Proof. We only need to verify this for the corresponding basis functions on the reference element $\hat{T}$. For either Type I or Type II element, by direct calculations, we can see that

$$\sum_{i=1}^4 \hat{a}_i^s = 1, \quad \sum_{i=1}^4 \hat{b}_i^s = 0, \quad \sum_{i=1}^4 \hat{c}_i^s = 0, \quad \sum_{i=1}^4 \hat{d}_i^s = 0, \quad s = \pm.$$

These imply that the partition of unity holds for the basis functions on the reference element and the result of this theorem follows.

■

The following lemma indicates that the bilinear IFE functions are consistent with the standard bilinear finite element functions.

Lemma 3.4.3 *Assume that $T \in \mathcal{T}_h$ is and interface element and $\phi \in S_h(T)$. If $\beta^- = \beta^+$, then*

$$\phi^- = \phi^+$$

and ϕ becomes a bilinear polynomial.

Proof. By direct calculations we can see that the result is true for $\hat{\phi}_i$, $i = 1, 2, 3, 4$ and then for ϕ_i, $i = 1, 2, 3, 4$. Since $\phi \in S_h(T)$ is a linear combination of ϕ_i, $i = 1, 2, 3, 4$, we know that the result of this lemma is also true for every $\phi \in S_h(T)$.

■

In the discussion below, we need another assumption on the mesh $\mathcal{T}_h$.

(H_3): The family of partitions $\mathcal{T}_h$ with $h > 0$ is regular. (See Definition 3.4.1 of [178] for regular partitions)

We use C to represent a generic constant whose value might be different from line to line. Unless otherwise specified, all the generic constants C in the presentation below are independent of interface and mesh. The following theorem establishes bounds for the bilinear IFE basis functions.

Theorem 3.4.2 *Let $T \in \mathcal{T}_h$ be an interface element and let $\phi_i(X) \in S_h(T)$, $i = 1, 2, 3, 4$, be the local bilinear IFE basis functions defined above. Then, there exists constants C such that*

$$|\phi_i(X)| \leq C, i = 1, 2, 3, 4, \tag{3.11}$$

$$\|\nabla \phi_i(X)\| \leq Ch^{-1}, i = 1, 2, 3, 4, \tag{3.12}$$

$$\left| \frac{\partial^2 \phi_i(X)}{\partial x \partial y} \right| \leq Ch^{-2}, i = 1, 2, 3, 4. \tag{3.13}$$

Proof. Without loss of generality, we assume that $R = \beta^-/\beta^+ \geq 1$. Similar arguments hold for $R = \beta^+/\beta^- \geq 1$.

First, we show that the coefficients $\hat{a}_i^s, \hat{b}_i^s, \hat{c}_i^s, \hat{d}_i^s, s = \pm$, $i = 1, 2, 3, 4$ of $\hat{\phi}_i$ are bounded. Note that these coefficients are linear combinations of $\frac{\hat{a}^k \hat{b}^l}{W}$ with the values of k and l listed in (3.7) for Type I interface element and (3.8) for Type II interface element. For Type I interface

element, since $0 \le \hat{a}, \hat{b} \le 1$, we have

$$
\begin{aligned}
\frac{\hat{a}^2}{W} &\le \begin{cases} \dfrac{\hat{a}^2}{R2\hat{a}^2(1-\hat{b})} \le \dfrac{1}{R} \le 1, & \text{if } 0 \le \hat{b} \le 1/2, \\ \dfrac{\hat{a}^2}{\hat{a}^2\hat{b}(2-\hat{a})} \le 2, & \text{if } 1/2 \le \hat{b} \le 1, \end{cases} \\
\frac{\hat{b}^2}{W} &\le \begin{cases} \dfrac{\hat{b}^2}{R2\hat{b}^2(1-\hat{a})} \le \dfrac{1}{R} \le 1, & \text{if } 0 \le \hat{a} \le 1/2, \\ \dfrac{\hat{b}^2}{\hat{a}\hat{b}^2(2-\hat{a})} \le 2, & \text{if } 1/2 \le \hat{a} \le 1, \end{cases} \\
\frac{\hat{a}\hat{b}}{W} &\le \begin{cases} \dfrac{\hat{a}\hat{b}}{\hat{a}^2\hat{b}(2-\hat{a})} = \dfrac{1}{\hat{a}(2-\hat{a})} \le 2, & \text{if } 1/2 \le \hat{a} \le 1, \\ \dfrac{\hat{a}\hat{b}}{\hat{a}\hat{b}^2(2-\hat{b})} = \dfrac{1}{\hat{b}(2-\hat{b})} \le 2, & \text{if } 1/2 \le \hat{b} \le 1, \\ \dfrac{\hat{a}\hat{b}}{R\big[2\hat{b}^2(1-\hat{a}) + 2\hat{a}^2(1-\hat{b})\big]} \le \dfrac{ab}{R(\hat{a}^2+\hat{b}^2)} \le \dfrac{1}{2R} \le \dfrac{1}{2}, & \text{if } 0 \le \hat{a}, \hat{b} \le 1/2, \end{cases} \\
\frac{\hat{a}^3}{W} &\le \begin{cases} \dfrac{\hat{a}^3}{R2\hat{a}^2(1-\hat{b})} \le \dfrac{\hat{a}}{R} \le \dfrac{1}{R} \le 1, & \text{if } 0 \le \hat{b} \le 1/2, \\ \dfrac{\hat{a}^3}{\hat{a}^2\hat{b}(2-\hat{a})} = \dfrac{\hat{a}}{\hat{b}(2-\hat{a})} \le 2\hat{a} \le 2, & \text{if } 1/2 \le \hat{b} \le 1, \end{cases} \\
\frac{\hat{b}^3}{W} &\le \begin{cases} \dfrac{\hat{b}^3}{R2\hat{b}^2(1-\hat{a})} \le \dfrac{\hat{b}}{R} \le \dfrac{1}{R} \le 1, & \text{if } 0 \le \hat{a} \le 1/2, \\ \dfrac{\hat{b}^3}{\hat{a}\hat{b}^2(2-\hat{b})} = \dfrac{\hat{b}}{\hat{a}(2-\hat{b})} \le 2\hat{b} \le 2, & \text{if } 1/2 \le \hat{a} \le 1, \end{cases} \\
\frac{\hat{a}^2\hat{b}}{W} &\le \frac{\hat{a}^2\hat{b}}{\hat{a}^2\hat{b}(2-\hat{a})} = \frac{1}{(2-\hat{a})} \le 1, \\
\frac{\hat{a}\hat{b}^2}{W} &\le \frac{\hat{a}\hat{b}^2}{\hat{a}\hat{b}^2(2-\hat{b})} = \frac{1}{(2-\hat{b})} \le 1, \\
\frac{\hat{a}^3\hat{b}}{W} &\le \frac{\hat{a}^3\hat{b}}{R\hat{a}^3\hat{b}} = \frac{1}{R} \le 1, \\
\frac{\hat{a}\hat{b}^3}{W} &\le \frac{\hat{a}\hat{b}^3}{R\hat{a}\hat{b}^3} = \frac{1}{R} \le 1.
\end{aligned}
$$

These properties imply that there exists a constant C such that $0 \le \left|\dfrac{\hat{a}^k\hat{b}^l}{W}\right| \le C$ for the values of k and l listed in (3.7).

For Type II interface element, by the following direct calculations, we can see that there exists a constant C independent of interface and partition such that $0 \le \left|\dfrac{\hat{a}^k\hat{b}^l}{W}\right| \le C$ for the

values of k and l listed in (3.8) since $0 \le \hat{a}, \hat{b} \le 1$.

$$\frac{1}{W} \le \begin{cases} \frac{1}{\hat{a}^2\hat{b}} \le 8, & \text{if } 1/2 \le \hat{a}, \hat{b} \le 1, \\ \frac{1}{R(2-\hat{a}-\hat{b})} \le \frac{2}{R} \le 2, & \text{otherwise.} \end{cases}$$

These inequalities lead to the boundedness of $\hat{a}_i, \hat{b}_i, \hat{c}_i, \hat{d}_i,\ i = 1, 2, 3, 4$, which imply the boundedness of $\hat{\phi}_i$, $i = 1, 2, 3, 4$. Then (3.11) follows because the affine transformation (3.5) is used to define ϕ_i, $i = 1, 2, 3, 4$ from $\hat{\phi}_i$.

Since the partition is regular, we have

$$\left\|M^{-T}\right\| \le Ch^{-1}. \tag{3.14}$$

Then, (3.12) follows from

$$\nabla\phi_i = M^{-T}\nabla\hat{\phi}_i,$$

and the boundedness of the coefficients of $\nabla\hat{\phi}_i$.

As for (3.13), we first let

$$m_{11} = \frac{\partial\hat{x}}{\partial x},\ , m_{12} = \frac{\partial\hat{y}}{\partial x},\ , m_{21} = \frac{\partial\hat{x}}{\partial y},\ , m_{22} = \frac{\partial\hat{y}}{\partial y},$$

then

$$M^{-T} = \begin{pmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{pmatrix}.$$

Hence by $\frac{\partial^2\hat{\phi}_i(X)}{\partial\hat{x}^2} = 0$ and $\frac{\partial^2\hat{\phi}_i(X)}{\partial\hat{y}^2} = 0$, we get

$$\begin{aligned} \frac{\partial^2\phi_i(X)}{\partial x\partial y} &= \frac{\partial}{\partial y}\left(\frac{\partial\phi_i(X)}{\partial x}\right) \\ &= \frac{\partial}{\partial y}\left(m_{11}\frac{\partial\hat{\phi}_i(X)}{\partial\hat{x}} + m_{12}\frac{\partial\hat{\phi}_i(X)}{\partial\hat{y}}\right) \\ &= m_{11}\frac{\partial^2\hat{\phi}_i(X)}{\partial\hat{x}^2}\frac{\partial\hat{x}}{\partial y} + m_{11}\frac{\partial^2\hat{\phi}_i(X)}{\partial\hat{x}\partial\hat{y}}\frac{\partial\hat{y}}{\partial y} + m_{12}\frac{\partial^2\hat{\phi}_i(X)}{\partial\hat{y}\partial\hat{x}}\frac{\partial\hat{x}}{\partial y} + m_{12}\frac{\partial^2\hat{\phi}_i(X)}{\partial\hat{y}^2}\frac{\partial\hat{y}}{\partial y} \\ &= \frac{\partial^2\hat{\phi}_i(\hat{X})}{\partial\hat{x}\partial\hat{y}}(m_{12}m_{21} + m_{11}m_{22}). \end{aligned}$$

By (3.14), we have

$$|m_{ij}| \le Ch^{-1},\ i, j = 1, 2.$$

Hence

$$|m_{12}m_{21} + m_{11}m_{22}| \leq Ch^{-2}.$$

Note that $\dfrac{\partial^2 \hat{\phi}_i(\hat{X})}{\partial \hat{x}\partial \hat{y}}$ is a bounded constant $\hat{d}^-(=\hat{d}^+)$, then

$$\left|\frac{\partial^2 \phi_i(X)}{\partial x \partial y}\right| \ \leq \ Ch^{-2}.$$

■

For any function u defined on a rectangular element $T = \square A_1A_2A_3A_4$, we let $\hat{u}$ be the corresponding function on $\hat{T}$ induced by u with

$$\hat{u}(\hat{X}) = u(F(\hat{X})),$$

where $F : \hat{T} \longrightarrow T$ is the affine mapping defined by (3.5) in section 3.2. Let J_F be the Jacobian of F. Without loss of generality, we assume that $A_1 \in \tilde{T}^-,\ A_2, A_3, A_4 \in \tilde{T}^+$ for a Type I rectangular interface element, $A_1, A_4 \in \tilde{T}^-,\ A_2, A_3 \in \tilde{T}^+$ for a Type II rectangular interface element for the following lemma.

Lemma 3.4.4 *On each interface element T, we have following results:*

1. *If $T = \square A_1A_2A_3A_4$ is a rectangular interface element of Type I, then every $\tilde{u}_h \in S_h(T)$ can be uniquely represented as follows:*

$$\tilde{u}_h(X) = \begin{cases} \tilde{u}_h^-(X) \ = \ \tilde{u}_h(A_1)\psi_1(X) + \sum_{i=2}^{4} w_i\psi_i(X), & X \in \tilde{T}^-, \\ \\ \tilde{u}_h^+(X) \ = \ w_1\psi_1(X) + \sum_{i=2}^{4} \tilde{u}_h(A_i)\psi_i(X), & X \in \tilde{T}^+, \end{cases} \tag{3.15}$$

2. *If $T = \square A_1A_2A_3A_4$ is a rectangular interface element of Type II, then for every $\tilde{u}_h \in S_h(T)$ can be uniquely represented as follows:*

$$\tilde{u}_h(X) = \begin{cases} \tilde{u}_h^-(X) \ = \ \tilde{u}_h(A_1)\psi_1(X) + \sum_{i=2}^{3} w_i\psi_i(X) + \tilde{u}_h(A_4)\psi_4(X), & X \in \tilde{T}^-, \\ \\ \tilde{u}_h^+(X) \ = \ w_1\psi_1(X) + \sum_{i=2}^{3} \tilde{u}_h(A_i)\psi_i(X) + w_4\psi_4(X), & X \in \tilde{T}^+, \end{cases} \tag{3.16}$$

3. There exist constants C_1 and C_2 such that

$$C_1 \left\|\vec{u}^+\right\| \ \leq \ \left\|\vec{u}^-\right\| \ \leq \ C_2 \left\|\vec{u}^+\right\|, \tag{3.17}$$

where

$$\vec{u}^- = \begin{pmatrix} \tilde{u}_h(A_1) \\ w_2 \\ w_3 \\ w_4 \end{pmatrix}, \quad \vec{u}^+ = \begin{pmatrix} w_1 \\ \tilde{u}_h(A_2) \\ \tilde{u}_h(A_3) \\ \tilde{u}_h(A_4) \end{pmatrix}.$$

for rectangular interface elements of Type I and

$$\vec{u}^- = \begin{pmatrix} \tilde{u}_h(A_1) \\ w_2 \\ w_3 \\ \tilde{u}_h(A_4) \end{pmatrix}, \quad \vec{u}^+ = \begin{pmatrix} w_1 \\ \tilde{u}_h(A_2) \\ \tilde{u}_h(A_3) \\ w_4 \end{pmatrix}.$$

for rectangular interface elements of Type II.

4. There exist constants C_3 and C_4 such that

$$C_3 \left\|\vec{w}\right\| \ \leq \ \left\|\vec{u}\right\| \ \leq \ C_4 \left\|\vec{w}\right\|, \tag{3.18}$$

where

$$\vec{u} = \begin{bmatrix} \tilde{u}_h(A_1) \\ \tilde{u}_h(A_2) \\ \tilde{u}_h(A_3) \\ \tilde{u}_h(A_4) \end{bmatrix}, \quad \vec{w} = \begin{bmatrix} w_1 \\ w_2 \\ w_3 \\ w_4 \end{bmatrix}.$$

Proof. Without loss of generality, we only need to discuss on the reference interface element. For a local interface element $[0,h]\times[0,h]$, if we define $\hat{a} = \frac{a}{h}$ and $\hat{b} = \frac{b}{h}$, then we can get the same conclusions. First, we discuss rectangular interface element of Type I. Applying jump conditions specified in (3.3) to $\hat{\tilde{u}}_h$ in (3.15), we obtain a linear system about coefficients in (3.15) which can be written as

$$A_l\vec{w} = A_r\vec{u}, \quad \vec{w} = (w_1, w_2, w_3, w_4)^t, \quad \vec{u} = (\tilde{u}_h(A_1), \tilde{u}_h(A_2), \tilde{u}_h(A_3), \tilde{u}_h(A_4))^t.$$

where

$$A_l = \begin{pmatrix} 1-\hat{b} & 0 & 0 & -\hat{b} \\ 1-\hat{a} & -\hat{a} & 0 & 0 \\ 1 & 1 & -1 & 1 \\ -2\hat{a}-2\hat{b}+\hat{a}^2+\hat{b}^2 & R(-2\hat{b}+\hat{a}^2+\hat{b}^2) & -R(\hat{a}^2+\hat{b}^2) & R(-2\hat{a}+\hat{a}^2+\hat{b}^2) \end{pmatrix},$$

$$A_r = \begin{pmatrix} 1-\hat{b} & 0 & 0 & -\hat{b} \\ 1-\hat{a} & -\hat{a} & 0 & 0 \\ 1 & 1 & -1 & 1 \\ R(-2\hat{a}-2\hat{b}+\hat{a}^2+\hat{b}^2) & -2\hat{b}+\hat{a}^2+\hat{b}^2 & -(\hat{a}^2+\hat{b}^2) & -2\hat{a}+\hat{a}^2+\hat{b}^2 \end{pmatrix}.$$

Here $R = \beta^-/\beta^+$. Direct calculations give us

$$\begin{aligned} det(A_l) &= -2R\hat{a}^2 + 2R\hat{a}^2\hat{b} - 2R\hat{b}^2 + 2R\hat{a}\hat{b}^2 - R\hat{a}^3\hat{b} - R\hat{a}\hat{b}^3 - 2\hat{a}^2\hat{b} - 2\hat{a}\hat{b}^2 + \hat{a}^3\hat{b} + \hat{a}\hat{b}^3 \\ &= -2R\hat{a}^2(1-\hat{b}) - 2R\hat{b}^2(1-\hat{a}) - R\hat{a}^3\hat{b} - R\hat{a}\hat{b}^3 - \hat{a}^2\hat{b}(2-\hat{a}) - \hat{a}\hat{b}^2(2-\hat{b}) \\ &< 0, \end{aligned}$$

which shows that the matrix A_l is non-singular for all $\hat{a}, \hat{b} \in [0,1]$. Hence $\vec{C}$ can be uniquely determined by $\vec{u}$ and this leads to the unique representation of $\tilde{u}_h \in S_h(T)$ in (3.15).

Rearrange the linear system above we can have

$$A_- \vec{u}^- = A_+ \vec{u}^+.$$

where

$$A_- = \begin{pmatrix} 1-\hat{b} & 0 & 0 & \hat{b} \\ 1-\hat{a} & \hat{a} & 0 & 0 \\ 1 & -1 & 1 & -1 \\ R(-2\hat{a}-2\hat{b}+\hat{a}^2+\hat{b}^2) & -R(-2\hat{b}+\hat{a}^2+\hat{b}^2) & R(\hat{a}^2+\hat{b}^2) & -R(-2\hat{a}+\hat{a}^2+\hat{b}^2) \end{pmatrix},$$

$$A_+ = \begin{pmatrix} 1-\hat{b} & 0 & 0 & \hat{b} \\ 1-\hat{a} & \hat{a} & 0 & 0 \\ 1 & -1 & 1 & -1 \\ -2\hat{a}-2\hat{b}+\hat{a}^2+\hat{b}^2 & -(-2\hat{b}+\hat{a}^2+\hat{b}^2) & \hat{a}^2+\hat{b}^2 & -(-2\hat{a}+\hat{a}^2+\hat{b}^2) \end{pmatrix}.$$

By direct calculations, we can show that A_-^{-1} and A_+^{-1} exist such that the entries of $A_-^{-1}A_+$ and $A_+^{-1}A_-$ are either polynomials of $\hat{a}$ and $\hat{b}$ or linear combination of functions in following forms:

$$\frac{\hat{a}^{r_a}\hat{b}^{r_b}}{\hat{a}^2+\hat{b}^2}, r_a \geq 0, r_b \geq 0, r_a + r_b \geq 2.$$

It can be shown that all of these functions of $\hat{a}$ and $\hat{b}$ are bounded by a constant independent of $\hat{a}$ and $\hat{b}$. Therefore, there exits a constant C such that $\left\|A_-^{-1}A_+\right\| \leq C$, $\left\|A_+^{-1}A_-\right\| \leq C$, $\forall \hat{a}, \hat{b} \in [0,1]$ and this leads to (3.17) for rectangular interface elements of Type I.

Similar arguments can be applied to obtain results for Type II rectangular and triangular interface elements. For rectangular interface elements of Type II, applying jump conditions specified in (3.3) to $\hat{\tilde{u}}_h$ in (3.16), we obtain a linear system about coefficients in (3.16) which can be written as

$$A_l\vec{w} = A_r\vec{u}, \quad \vec{C} = (w_1, w_2, w_3, w_4)^t, \quad \vec{u} = (\tilde{u}_h(A_1), \tilde{u}_h(A_2), \tilde{u}_h(A_3), \tilde{u}_h(A_4))^t.$$

where

$$A_l = \begin{pmatrix} 0 & 0 & -\hat{b} & 1-\hat{b} \\ 1-\hat{a} & -\hat{a} & 0 & 0 \\ 1 & 1 & -1 & -1 \\ \hat{a}^2-\hat{b}^2-1-2\hat{a}+2\hat{b} & R(\hat{a}^2-\hat{b}^2-1) & R(\hat{b}^2-\hat{a}^2-1) & \hat{b}^2-\hat{a}^2-1+2\hat{a}-2\hat{b} \end{pmatrix},$$

$$A_r = \begin{pmatrix} 0 & 0 & -\hat{b} & 1-\hat{b} \\ 1-\hat{a} & -\hat{a} & 0 & 0 \\ 1 & 1 & -1 & -1 \\ R(\hat{a}^2-\hat{b}^2-1-2\hat{a}+2\hat{b}) & \hat{a}^2-\hat{b}^2-1 & \hat{b}^2-\hat{a}^2-1 & R(\hat{b}^2-\hat{a}^2-1+2\hat{a}-2\hat{b}) \end{pmatrix}.$$

Here $R = \beta^-/\beta^+$. Direct calculations give us

$$\begin{aligned} det(A_l) &= -R\hat{a}+\hat{a}-R\hat{a}\hat{b}^2-Ra^2\hat{b}-R\hat{b}+2R+\hat{b}+2\hat{a}^2+2\hat{b}^2-\hat{b}^3+R\hat{b}^3+\hat{a}^2\hat{b}-4\hat{a}\hat{b} \\ &\quad +\hat{a}\hat{b}^2+R\hat{a}^3-\hat{a}^3 \\ &= \hat{a}(1-\hat{a}^2)+\hat{b}(1-\hat{b}^2)+2(\hat{a}-\hat{b})^2+\hat{a}^2\hat{b}+\hat{a}\hat{b}^2+R(1-\hat{a})+R(1-\hat{b}) \\ &\quad +R(\hat{a}-\hat{b})^2(\hat{a}+\hat{b}) \\ &> 0, \end{aligned}$$

which shows that the matrix A_l is non-singular for all $\hat{a}, \hat{b} \in [0,1]$. Hence $\vec{C}$ can be uniquely determined by $\vec{u}$ and this leads to the unique representation of $\tilde{u}_h \in S_h(T)$ in (3.16).

Rearrange the linear system above we can have

$$A_- \vec{u}^- = A_+ \vec{u}^+.$$

where

$$A_- = \begin{pmatrix} 0 & 0 & \hat{b} & 1-\hat{b} \\ 1-\hat{a} & \hat{a} & 0 & 0 \\ 1 & -1 & 1 & -1 \\ R(\hat{a}^2-\hat{b}^2-1-2\hat{a}+2\hat{b}) & R(1-\hat{a}^2+\hat{b}^2) & R(1+\hat{a}^2-\hat{b}^2) & R(\hat{b}^2-\hat{a}^2-1+2\hat{a}-2\hat{b}) \end{pmatrix},$$

$$A_+ = \begin{pmatrix} 0 & 0 & \hat{b} & 1-\hat{b} \\ 1-\hat{a} & \hat{a} & 0 & 0 \\ 1 & -1 & 1 & -1 \\ \hat{a}^2-\hat{b}^2-1-2\hat{a}+2\hat{b} & 1-\hat{a}^2+\hat{b}^2 & 1+\hat{a}^2-\hat{b}^2 & \hat{b}^2-\hat{a}^2-1+2\hat{a}-2\hat{b} \end{pmatrix}.$$

By direct calculations, we can show that A_-^{-1} and A_+^{-1} exist such that the entries of $A_-^{-1}A_+$ and $A_+^{-1}A_-$ are either polynomials of $\hat{a}$ and $\hat{b}$ or linear combination of functions in following forms:

$$\frac{\hat{a}^{r_a}\hat{b}^{r_b}}{(\hat{a}-\hat{b})^2+1}, r_a \geq 0, r_b \geq 0, r_a + r_b \geq 0.$$

It can be shown that all of these functions of $\hat{a}$ and $\hat{b}$ are bounded by a constant independent of $\hat{a}$ and $\hat{b}$. Therefore, there exits a constant C such that $\left\|A_-^{-1}A_+\right\| \leq C,\ \left\|A_+^{-1}A_-\right\| \leq C,\ \forall \hat{a}, \hat{b} \in [0,1]$ and this leads to (3.17) for rectangular interface elements of Type II.

Finally similar arguments can be used to prove (3.18).

■

The results in the next lemma indicate that the bilinear IFE functions also satisfy certain inverse inequalities. For an interface element T of Type I, we let $T_{1/2} \subset T$ be the half of T that is formed by a diagonal line not intersecting with $\overline{DE}$. For an interface element T of Type II, we can first divide T into 4 small rectangles by connecting the midpoints of opposite edges of T. Note that $\overline{DE}$ can intersect with three of these small rectangles at most. Then, we let $T_{1/4}$ be one of these small rectangles not intersecting with $\overline{DE}$. Also, for each element T, we let h_x and h_y be the edge lengthes in x-direction and y-direction, respectively, and we let $h = \max\{h_x, h_y\}$. In the following discussion, we will use the notation $T_{1/t}$ with $t = 2$ for a Type I interface element and $t = 4$ for a Type II interface element. It's easy to see that $T_{1/t} \subset \tilde{T}^s, s = +$ or $-$.

Lemma 3.4.5 *There exists a constant C such that for all $v \in S_h(T)$, we have*

$$\begin{cases} \|v\|_{\infty,\tilde{T}^s} \leq \dfrac{C}{h}\|v\|_{0,T}, \\ \\ \|v_x\|_{\infty,\tilde{T}^s} \leq \dfrac{C}{h}\|v_x\|_{0,T}, \quad \|v_y\|_{\infty,\tilde{T}^s} \leq \dfrac{C}{h}\|v_y\|_{0,T}, \end{cases} \tag{3.19}$$

$$|v|_{1,T} \geq Ch|v|_{2,T}, \tag{3.20}$$

$$|v|_{0,T} \geq Ch^2|v|_{2,T}, \tag{3.21}$$

$$|v|_{0,T} \geq Ch|v|_{1,T^s}, \tag{3.22}$$

provided that $\tilde{T}^s \supset T_{1/t}, s = +$ or $-, t = 2, 4$.

Proof. For each $v \in S_h(T)$, its restriction on $\tilde{T}^s$ is just a bilinear polynomial, we denote it by v^s. Let v^s_{ext} be the extension of v^s to the whole interface element T. Then, we can obtain the first inequality in (3.19) as follows:

$$\begin{aligned} \|v\|_{\infty,\tilde{T}^s} &= \|v^s\|_{\infty,\tilde{T}^s} = \left\|\widehat{v^s}\right\|_{\infty,\hat{T}^s} \leq \left\|\widehat{v^s_{ext}}\right\|_{\infty,\hat{T}} \leq C\left\|\widehat{v^s_{ext}}\right\|_{0,\hat{T}_{1/t}} = C\left\|\widehat{v^s}\right\|_{0,\hat{T}_{1/t}} \\ &\leq C\left|J_F\right|^{-1/2}\|v^s\|_{0,T_{1/t}} \leq Ch^{-1}\|v\|_{0,T_{1/t}} \leq Ch^{-1}\|v\|_{0,T}. \end{aligned}$$

The other inequalities in (3.19) can be shown similarly.

$\forall v \in S_h(T)$, we have

$$v(x,y) = \begin{cases} v^-(x,y) = a^-x + b^-y + c^- + d^-xy, & (x,y) \in \tilde{T}^-, \\ v^+(x,y) = a^+x + b^+y + c^+ + d^+xy, & (x,y) \in \tilde{T}^+, \end{cases}$$

By Lemma 3.4.1, we have $d^- = d^+$. Therefore,

$$\begin{aligned} |v|_{2,T}^2 &= \|v_{xx}\|_{0,T}^2 + \|v_{yy}\|_{0,T}^2 + \|v_{xy}\|_{0,T}^2 = \|v_{xy}\|_{0,T}^2 \\ &= \int_{\tilde{T}^-} (d^-)^2 dxdy + \int_{\tilde{T}^+} (d^+)^2 dxdy = \int_T (d^-)^2 dxdy \\ &= (d^-)^2 h_x h_y, \end{aligned}$$

and

$$\begin{aligned} |v|_{2,T_{1/t}}^2 &= \|v_{xx}\|_{0,T_{1/t}}^2 + \|v_{yy}\|_{0,T_{1/t}}^2 + \|v_{xy}\|_{0,T_{1/t}}^2 \\ &= \|v_{xy}\|_{0,T_{1/t}}^2 = \int_{T_{1/t}} (d^-)^2 dxdy \\ &= \frac{1}{t}(d^-)^2 h_x h_y. \end{aligned}$$

Then, using the standard inverse inequality, we have

$$\begin{aligned} |v|_{1,T}^2 &\geq |v|_{1,T_{1/t}}^2 \geq Ch^2 |v|_{2,T_{1/t}}^2 = Ch^2 \frac{1}{t}(d^-)^2 h_x h_y \\ &= Ch^2 |v|_{2,T}^2, \end{aligned}$$

which leads to (3.20). With the same idea, we can prove (3.21) and (3.22) as follows.

$$\begin{aligned} |v|_{0,T}^2 &\geq |v|_{0,T_{1/t}}^2 \geq Ch^4 |v|_{2,T_{1/t}}^2 = Ch^4 \frac{1}{t}(d^-)^2 h_x h_y = Ch^4 |v|_{2,T}^2, \\ |v|_{0,T}^2 &\geq |v^s|_{0,T_{1/t}}^2 \geq Ch^2 |v^s|_{1,T_{1/t}}^2 \geq Ch^2 |v_{ext}^s|_{1,T}^2 \geq Ch^2 |v|_{1,\tilde{T}^s}^2. \end{aligned}$$

■

Chapter 4

Approximation capability of the bilinear IFE space

In this chapter we will analyze the approximation capability of the bilinear IFE space discussed in Chapter 3 by estimating the finite element interpolation error and present some numerical examples to verify the theoretical conclusion [111, 112]. One popular way for the interpolation error estimation is the scaling technique. However, when we apply this technique to the 2D or 3D IFE spaces, it's not clear that if the constant C in the error bound depends on the interface and the optimal order cannot be concluded, which cause the error estimation to fail. Therefore, we follow the idea in [143] to estimate the interpolation error of bilinear IFE space by using multi-point Taylor expansions.

4.1 Error estimation for the bilinear IFE interpolation

Our goal in this section is to estimate the bilinear IFE interpolation error, which is a critical step in error estimation of a finite element (or finite volume element) method. We will basically follow the framework developed in [143] which deals with a triangular IFE space. However, the local bilinear IFE basis functions have a second degree term involving xy which leads to new difficulties demanding different techniques to analyze the interpolation error. In addition, we note that, topologically, there are two types of interface elements for a mesh formed by rectangles in contrast with a triangular mesh in which there is basically only one type of interface element, and the two types of interface elements need to be discussed separately. We focus on the bilinear IFE interpolation of a function from a suitable Sobolev space, and will derive error estimates in the corresponding Sobolev norms.

In [112], we analyze the interpolation error for the bilinear IFE space defined by the functions in (3.1). In this dissertation, we choose the functions in (3.4) to construct the bilinear IFE space, so we'll analyze the interpolation error with similar arguments.

4.1.1 Some preliminaries

For any subset $\Lambda \subset \Omega$ whose interior is cut through by the interface Γ, we let

$$PH^2_{int}(\Lambda) = \left\{ u \in C(\Lambda), u|_{\Lambda^s} \in H^2(\Lambda^s), s = -,+, \left[\beta \frac{\partial u}{\partial \mathbf{n}_\Gamma}\right] = 0 \ \text{ on } \Gamma \cap \Lambda \right\},$$
$$PC^m_{int}(\Lambda) = \left\{ u \in C(\Lambda), u|_{\Lambda^s} \in C^m(\Lambda^s), s = -,+, \left[\beta \frac{\partial u}{\partial \mathbf{n}_\Gamma}\right] = 0 \ \text{ on } \Gamma \cap \Lambda \right\},$$

Obviously, we have $PC^2_{int}(\Lambda) \subset PH^2_{int}(\Lambda)$. For any function $u \in PH^2_{int}(\Lambda)$, we define

$$\begin{aligned} \|u\|^2_{s,\Lambda} &= \|u\|^2_{s,\Lambda^+} + \|u\|^2_{s,\Lambda^-}, \ s = 0,1,2, \\ |u|^2_{s,\Lambda} &= |u|^2_{s,\Lambda^+} + |u|^2_{s,\Lambda^-}, \ s = 0,1,2. \end{aligned}$$

Consider an interface element $T \in \mathcal{T}_h$. For any function $w_h \in S_h(T)$ and $u \in PH^2_{int}(T)$, we define

$$\begin{aligned} \|w_h + u\|^2_{s,T} &= \|w_h + u\|^2_{s,\tilde{T}^+ \cap T^+} + \|w_h + u\|^2_{s,\tilde{T}^+ \cap T^-} + \|w_h + u\|^2_{s,\tilde{T}^- \cap T^+} + \|w_h + u\|^2_{s,\tilde{T}^- \cap T^-}, \\ |w_h + u|^2_{s,T} &= |w_h + u|^2_{s,\tilde{T}^+ \cap T^+} + |w_h + u|^2_{s,\tilde{T}^+ \cap T^-} + |w_h + u|^2_{s,\tilde{T}^- \cap T^+} + |w_h + u|^2_{s,\tilde{T}^- \cap T^-}, \\ & \quad s = 0,1,2. \end{aligned}$$

Here note that one of $\tilde{T}^+ \cap T^-$ and $\tilde{T}^- \cap T^+$ might be empty. In that case we can remove the norm and semi-norm on the empty set from the above definitions.

In this chapter, we assume that the interface curve Γ and the mesh $\mathcal{T}_h$ satisfy the following assumptions:

(H_4): The interface curve Γ is defined by a piecewise C^2 function, and the mesh $\mathcal{T}_h$ is formed such that the subset of Γ in every interface element is C^2.

(H_5): The interface Γ is smooth enough so that $PC^3_{int}(T)$ is dense in $PH^2_{int}(T)$ for every interface element T of $\mathcal{T}_h$.

We note that (H_5) will hold if Γ is sufficiently smooth, see the results of [161, 201] on the transmission problems.

For a function $u \in PH^2_{int}(T), \ \ T \in \mathcal{T}_h$, we let $I_{h,T}u \in S_h(T)$ be its interpolation such that $I_{h,T}u(X) = u(X)$ when X is a vertex of T. For an element T with vertices A_1, A_2, A_3, A_4, we have

$$I_{h,T}u(X) = u(A_1)\phi_1(X) + u(A_2)\phi_2(X) + u(A_3)\phi_3(X) + u(A_4)\phi_4(X).$$

Accordingly, for a function $u \in PH^2_{int}(\Omega)$, we let $I_h u \in S_h(\Omega)$ be its interpolation such that $I_h u|_T = I_{h,T}(u|_T)$ for any $T \in \mathcal{T}_h$.

The purpose of this section is to derive error estimates for the interpolation of $u \in PH^2_{int}(\Omega)$, and we will carry out the discussion piecewisely for each element T in the mesh $\mathcal{T}_h$. Recall that the error estimate of $I_h u$ in any non-interface element T is well known, see for example [178]:

$$\|I_h u - u\|_{0,T} + h \|I_h u - u\|_{1,T} \leq Ch^2 \|u\|_{2,T}.$$

Therefore, in the discussion from now on, we focus on interface elements of $\mathcal{T}_h$.

We call a point $X = (x, y)^T$ in an interface element T an *obscure point* if one of the four line segments connecting X and the vertices of T intersects the interface more than once. Without loss of generality, we discuss interface elements that do not contain any obscure point because the arguments used below can be readily extended to handle the interface elements with obscure points.

Let $\tilde{\rho} = \frac{\beta^+}{\beta^-}$, $\rho = \frac{\beta^-}{\beta^+}$. For any point $\widetilde{A} \in \Gamma$, let $\widetilde{A}^{\perp}$ be the orthogonal projection of $\tilde{A}$ onto $\overline{DE}$. Now let us recall the following three lemmas from [143].

Lemma 4.1.1 *Assume* $\mathbf{n}(\widetilde{A}) = (n_x(\widetilde{A}), n_y(\widetilde{A}))^T$ *is the unit normal vector of* Γ *at* $\widetilde{A}$, $\mathbf{n}(\overline{DE}) = (\overline{n}_x, \overline{n}_y)^T$ *is the unit normal vector of* $\overline{DE}$, *and* $X_{\overline{DE}}$ *is a point on* $\overline{DE}$. *Then, for every function* $u(x, y)$ *satisfying the interface jump conditions (1.3) and (1.4), we have*

$$\nabla u^+(\widetilde{A}) = N^-(\widetilde{A}) \nabla u^-(\widetilde{A}), N^-(\widetilde{A}) = \begin{pmatrix} n_y(\widetilde{A})^2 + \rho n_x(\widetilde{A})^2 & (\rho - 1) n_x(\widetilde{A}) n_y(\widetilde{A}) \\ (\rho - 1) n_x(\widetilde{A}) n_y(\widetilde{A}) & n_x(\widetilde{A})^2 + \rho n_y(\widetilde{A})^2 \end{pmatrix}, \quad (4.1)$$

$$\nabla u^-(\widetilde{A}) = N^+(\widetilde{A}) \nabla u^+(\widetilde{A}), N^+(\widetilde{A}) = \begin{pmatrix} n_y(\widetilde{A})^2 + \tilde{\rho} n_x(\widetilde{A})^2 & (\tilde{\rho} - 1) n_x(\widetilde{A}) n_y(\widetilde{A}) \\ (\tilde{\rho} - 1) n_x(\widetilde{A}) n_y(\widetilde{A}) & n_x(\widetilde{A})^2 + \tilde{\rho} n_y(\widetilde{A})^2 \end{pmatrix}, \quad (4.2)$$

and for every $v \in S_h(T)$ *we have*

$$\nabla v^+(X_{\overline{DE}}) = N^-_{\overline{DE}} \nabla v^-(X_{\overline{DE}}) \quad , \quad N^-_{\overline{DE}} = \begin{pmatrix} \overline{n}_y^2 + \rho \overline{n}_x^2 & (\rho - 1) \overline{n}_x \overline{n}_y \\ (\rho - 1) \overline{n}_x \overline{n}_y & \overline{n}_x^2 + \rho \overline{n}_y^2 \end{pmatrix}, \quad (4.3)$$

$$\nabla v^-(X_{\overline{DE}}) = N^+_{\overline{DE}} \nabla v^+(X_{\overline{DE}}) \quad , \quad N^+_{\overline{DE}} = \begin{pmatrix} \overline{n}_y^2 + \tilde{\rho} \overline{n}_x^2 & (\tilde{\rho} - 1) \overline{n}_x \overline{n}_y \\ (\tilde{\rho} - 1) \overline{n}_x \overline{n}_y & \overline{n}_x^2 + \tilde{\rho} \overline{n}_y^2 \end{pmatrix}. \quad (4.4)$$

Proof. We only carry out the proof for (4.1). The proof for the other three conclusions is similar. Because of (1.3), we have $u^+(x, y) - u^-(x, y) = 0$ on Γ. Assume Γ can be parametrized as follows:

$$\Gamma : \begin{cases} x = x(t), \\ y = y(t), \end{cases}$$

then $u^+(x(t),y(t))-u^-(x(t),y(t))=0$. Therefore, we have $0=u_x^+x'(t)+u_y^+y'(t)-u_x^-x'(t)-u_y^-y'(t)$. Let $\vec{T}(t)=\begin{pmatrix} x'(t) \\ y'(t) \end{pmatrix}$, then $\vec{T}(t)$ is tangent to Γ. Assume $\vec{T}(t)=\vec{T}$ at $\widetilde{A}$, then $\vec{T}\cdot\mathbf{n}(\widetilde{\mathbf{A}})=\mathbf{0}$. Therefore, we get $\vec{T}=\begin{pmatrix} -n_y(\widetilde{A}) \\ n_x(\widetilde{A}) \end{pmatrix}$. Then

$$0=-u_x^+(\widetilde{A})n_y(\widetilde{A})+u_y^+(\widetilde{A})n_x(\widetilde{A})+u_x^-(\widetilde{A})n_y(\widetilde{A})-u_y^-(\widetilde{A})n_x(\widetilde{A}). \tag{4.5}$$

Because of (1.4), we have

$$\begin{aligned} 0 &= \beta^+\frac{\partial u^+}{\partial n(\widetilde{A})}-\beta^-\frac{\partial u^-}{\partial n(\widetilde{A})} \\ &= \beta^+\nabla u^+(\widetilde{A})\cdot n(\widetilde{A})-\beta^-\nabla u^-(\widetilde{A})\cdot n(\widetilde{A}) \\ &= \beta^+u_x^+(\widetilde{A})n_x(\widetilde{A})+\beta^+u_y^+(\widetilde{A})n_y(\widetilde{A})-\beta^-u_x^-(\widetilde{A})n_x(\widetilde{A})-\beta^-u_y^-(\widetilde{A})n_y(\widetilde{A}). \end{aligned} \tag{4.6}$$

Solve (4.5) and (4.6) for $u_x^+(\widetilde{A})$ and $u_y^+(\widetilde{A})$, then we get (4.1).

■

Since $\Gamma\cap T$ is a C^2 curve, when the mesh $\mathcal{T}_h$ is fine enough, we can introduce a local coordinate system centered at point D with one axis in the direction of $\overline{DE}$. For any point $(x,y)^T$, let (ξ,η) be its coordinates in this local coordinate system where ξ-axis is in the direction of $\overline{DE}$. Then we have

$$\begin{pmatrix} x \\ y \end{pmatrix}=\begin{pmatrix} x_D \\ y_D \end{pmatrix}+\begin{pmatrix} \cos(\theta_{DE}) & -\sin(\theta_{DE}) \\ \sin(\theta_{DE}) & \cos(\theta_{DE}) \end{pmatrix}\begin{pmatrix} \xi \\ \eta \end{pmatrix}, \tag{4.7}$$

where $(x_D,y_D)^T$ is the coordinates of point D and θ_{DE} is the angle between $\overline{DE}$ and the x axis. As in [91], we have the following Lemma.

Lemma 4.1.2 *Let $\bar{\xi}$ be the length of $\overline{DE}$, then Γ has the following equation in this local system:*

$$\eta=\phi(\xi),\quad \xi\in[0,\bar{\xi}],$$

with

$$|\phi(\xi)|\le Ch^2, \tag{4.8}$$
$$|\phi'(\xi)|\le Ch. \tag{4.9}$$

Proof. Note that Γ passes both D and E, whose local coordinates are $(0,0)$ and $(\bar{\xi},0)$ separately. Therefore, in the local coordinate system, we have $\phi(0)=\phi(\bar{\xi})=0$ and $\phi\in C^2$. By using Mean Value Theorem, there exits a $\xi_1\in(0,\bar{\xi})$ such that $\phi'(\xi_1)=0$. Then

$$\phi'(\xi)=\phi'(\xi_1)+\int_{\xi_1}^{\xi}\phi''(s)\,ds\ =\int_{\xi_1}^{\xi}\phi''(s)\,ds\ ,\forall\xi\in(0,\bar{\xi}).$$

Because $\phi\in C^2$ implies ϕ'' is bounded in $[0,\bar{\xi}]$, we get

$$|\phi'(\xi)|\le\left|\int_{\xi_1}^{\xi}\phi''(s)\,ds\ \right|\le C\,|\xi-\xi_1|\le Ch.$$

By Taylor expansion, $\forall\xi\in(0,\bar{\xi})$ there exists $\xi_2\in(0,\xi)$ such that

$$\phi(\xi)=\phi(0)+\xi\phi'(\xi_2)=\xi\phi'(\xi_2).$$

Then

$$|\phi(\xi)|\le|\xi\phi'(\xi_2)|\le Ch^2.$$

■

From now on, if necessary, for any point P, we will use

$$\begin{pmatrix}x_P\\y_P\end{pmatrix}\ \text{ and }\ \begin{pmatrix}\xi_P\\\eta_P\end{pmatrix}$$

to denote its coordinates in the $x-y$ and $\xi-\eta$ systems, respectively.

Lemma 4.1.3 *There exist constants $C>0$ such that for any point $\widetilde{A}\in\Gamma$, we have*

$$\|\widetilde{A}-\widetilde{A}^{\perp}\|\le Ch^2, \tag{4.10}$$

$$\|N^s_{\overline{DE}}-N^s(\widetilde{A})\|\le Ch,\quad s=-,+. \tag{4.11}$$

Proof. The proof is the same as the proof of Lemma 3.1 in [143]. We just repeat it here. We only need to prove these in the local coordinate system for one type of vector norm because the transformation (4.7) preserves the vector length and all the finite dimensional vector norms are equivalent to each other. In the local system, $\forall\widetilde{A}\in\Gamma$, there exists $\widetilde{\xi}\in(0,\bar{\xi})$ such that

$$\widetilde{A}=\begin{pmatrix}\tilde{\xi}\\\phi(\tilde{\xi})\end{pmatrix},\widetilde{A}_{\perp}=\begin{pmatrix}\tilde{\xi}\\0\end{pmatrix}.$$

Hence (4.10) is just the consequence of (4.8). Also, we have

$$\mathbf{n}(\overline{DE}) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \mathbf{n}(\widetilde{A}) = \frac{1}{\sqrt{1+\left(\phi'(\tilde{\xi})\right)^2}} \begin{pmatrix} -\phi'(\tilde{\xi}) \\ 1 \end{pmatrix}.$$

Then, by (4.9), we have

$$\begin{aligned}
\left\|\mathbf{n}(\overline{DE}) - \mathbf{n}(\widetilde{A})\right\| &= \left|\frac{\phi'(\tilde{\xi})}{\sqrt{1+\left(\phi'(\tilde{\xi})\right)^2}}\right| + \left|1 - \frac{1}{\sqrt{1+\left(\phi'(\tilde{\xi})\right)^2}}\right| \\
&\leq \left|\phi'(\tilde{\xi})\right| + \left|\frac{\sqrt{1+\left(\phi'(\tilde{\xi})\right)^2}-1}{\sqrt{1+\left(\phi'(\tilde{\xi})\right)^2}}\right| \\
&\leq \left|\phi'(\tilde{\xi})\right| + \left|\frac{\left(\phi'(\tilde{\xi})\right)^2}{\sqrt{1+\left(\phi'(\tilde{\xi})\right)^2}\left(\sqrt{1+\left(\phi'(\tilde{\xi})\right)^2}+1\right)}\right| \\
&\leq \left|\phi'(\tilde{\xi})\right| + \left|\left(\phi'(\tilde{\xi})\right)^2\right| \\
&\leq Ch,
\end{aligned}$$

which together with definition of $N^s_{\overline{DE}}$ and $N^s(\widetilde{A})$, $s = -, +$, lead to (4.11).

■

The following lemma gives the straight forward Taylor expansion of a bilinear function.

Lemma 4.1.4 *If $f(x,y) = a + bx + cy + dxy$, $X = (x,y)$, $Z = (x_z, y_z)$, then*

$$f(Z) = f(X) + \nabla f(X) \cdot (Z - X) + d(x_z - x)(y_z - y).$$

Proof.

$$\begin{aligned}
\nabla f(X) \cdot (Z - X) &= (bx_z - bx + cy_z - cy) + d(yx_z - yx + xy_z - xy) \\
&= (a + bx_z + cy_z + dx_z y_z) - (a + bx + cy + dxy) \\
&\quad + dxy - dx_z y_z + d(yx_z - yx + xy_z - xy) \\
&= f(x_z, y_z) - f(x,y) + d(yx_z + xy_z - xy - x_z y_z) \\
&= f(Z) - f(X) - d(x_z - x)(y_z - y).
\end{aligned}$$

Then $f(Z) = f(X) + \nabla f(X) \cdot (Z - X) + d(x_z - x)(y_z - y)$.

■

Now we separately discuss the IFE interpolation error estimates for the two types of interface elements, i.e, Type I elements and Type II elements.

4.1.2 Interpolation error on a Type I interface element

In this section, we will discuss the bilinear IFE interpolation error on a Type I interface element. Without loss of generality, we assume $T \in \mathcal{T}_h$ is a Type I interface element with vertices $A_i = (x_i, y_i)$, $i = 1, 2, 3, 4$, such that $A_1 \in T^+$ and $A_i \in T^-$, $i = 2, 3, 4$, see Figure 4.1.

We start with the estimation on $\tilde{T}^- \cap T^-$. Consider a point $X = (x, y)^T \in \tilde{T}^- \cap T^-$ and assume that line segments $\overline{XA_i}$, $i = 2, 3, 4$ do not intersect with the interface and $\overline{DE}$, while line segment $\overline{XA_1}$ meets Γ at $\widetilde{A}_1$ (see Figure 4.1) with

$$\widetilde{A}_1 = \tilde{t}A_1 + (1 - \tilde{t})X = (\widetilde{x}_1, \widetilde{y}_1)^T \tag{4.12}$$

for a certain $\tilde{t}$. Note that $\widetilde{A}_1^{\perp}$ is the orthogonal projection of $\widetilde{A}_1 \in \Gamma$ onto $\overline{DE}$ (see Figure 4.1).

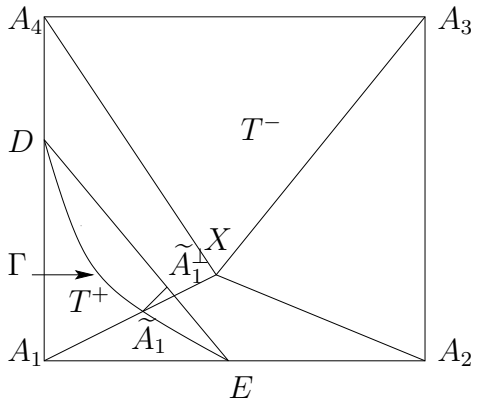

Figure 4.1: An interface rectangle element with no obscure point. A point X $\in \tilde{T}^- \cap T^-$ is connected to the four vertices by line segments in a Type I interface element

In all the discussion from now on, for a given point $X = (x, y)^T$, we let $X^s = (y, x)^T$. $\forall X = (x, y)^T \in T^-, A = (x_A, y_A)^T \in T^+$, we let $\widetilde{A} = (\widetilde{x}, \widetilde{y})^T$ be the intersection point of Γ and $\overline{AX}$.

The lemma below establishes an expansion of a bilinear IFE function across the interface.

Lemma 4.1.5 *Assume that* $v \in S_h(T), X = (x, y)^T \in \tilde{T}^-, A = (x_A, y_A)^T \in \tilde{T}^+$. *Then*

$$\begin{aligned} v(A) &= v(X) + \nabla v(X) \cdot (A - X) + (N^-_{\overline{DE}} - I)\nabla v(X) \cdot (A - \widetilde{A}) \\ &\quad + (N^-_{\overline{DE}} - I)\nabla v(X) \cdot (\widetilde{A} - X_{\overline{DE}}) + d^- N^-_{\overline{DE}}(X^s_{\overline{DE}} - X^s) \cdot (A - X_{\overline{DE}}) \\ &\quad + d^-(x_A - \bar{x})(y_A - \bar{y}) + d^-(\bar{x} - x)(\bar{y} - y), \end{aligned}$$

where $X_{\overline{DE}} = (\bar{x}, \bar{y})^T$ *is an arbitrary point on* $\overline{DE}$.

Proof. By Lemma 4.1.4, Lemma 3.4.1 and (4.3), we have

$$\begin{aligned} v(A) &= v^+(A) \\ &= v^+(X_{\overline{DE}}) + \nabla v^+(X_{\overline{DE}}) \cdot (A - X_{\overline{DE}}) + d^+(x_A - \bar{x})(y_A - \bar{y}) \\ &= v^-(X_{\overline{DE}}) + N^-_{\overline{DE}} \nabla v^-(X_{\overline{DE}}) \cdot (A - X_{\overline{DE}}) + d^-(x_A - \bar{x})(y_A - \bar{y}) \\ &= v^-(X) + \nabla v^-(X) \cdot (X_{\overline{DE}} - X) + d^-(\bar{x} - x)(\bar{y} - y) \\ &\quad + N^-_{\overline{DE}} \nabla v^-(X_{\overline{DE}}) \cdot (A - X_{\overline{DE}}) + d^-(x_A - \bar{x})(y_A - \bar{y}) \\ &= v^-(X) + \nabla v^-(X) \cdot (A - X) + (N^-_{\overline{DE}} - I)\nabla v^-(X) \cdot (A - X_{\overline{DE}}) \\ &\quad + N^-_{\overline{DE}} \Big[\nabla v^-(X_{\overline{DE}}) - \nabla v^-(X) \Big] \cdot (A - X_{\overline{DE}}) \\ &\quad + d^-(x_A - \bar{x})(y_A - \bar{y}) + d^-(\bar{x} - x)(\bar{y} - y). \end{aligned}$$

Because

$$\nabla v^-(X_{\overline{DE}}) = \begin{pmatrix} b^- + d^- \bar{y} \\ c^- + d^- \bar{x} \end{pmatrix}, \nabla v^-(X) = \begin{pmatrix} b^- + d^- y \\ c^- + d^- x \end{pmatrix},$$

we have $\nabla v^-(X_{\overline{DE}}) - \nabla v^-(X) = d^-(X^s_{\overline{DE}} - X^s)$. Hence,

$$\begin{aligned} v(A) &= v^-(X) + \nabla v^-(X) \cdot (A - X) + (N^-_{\overline{DE}} - I)\nabla v^-(X) \cdot (A - X_{\overline{DE}}) + \\ &\quad N^-_{\overline{DE}} d^-(X^s_{\overline{DE}} - X^s) \cdot (A - X_{\overline{DE}}) + d^-(x_A - \bar{x})(y_A - \bar{y}) + d^-(\bar{x} - x)(\bar{y} - y) \\ &= v(X) + \nabla v(X) \cdot (A - X) + (N^-_{\overline{DE}} - I)\nabla v(X) \cdot (A - \widetilde{A}) \\ &\quad + (N^-_{\overline{DE}} - I)\nabla v(X) \cdot (\widetilde{A} - X_{\overline{DE}}) + d^- N^-_{\overline{DE}}(X^s_{\overline{DE}} - X^s) \cdot (A - X_{\overline{DE}}) \\ &\quad + d^-(x_A - \bar{x})(y_A - \bar{y}) + d^-(\bar{x} - x)(\bar{y} - y). \end{aligned}$$

■

Now we prove the following two lemmas which will be used to prove the expansion of the interpolation error.

Lemma 4.1.6 *Assume that $v \in S_h(T), X = (x, y)^T \in \tilde{T}^-$. Then we have*

$$\begin{aligned}
&\nabla v(X) \cdot \sum_{i=1}^{4}(A_i - X)\phi_i(X) \\
= \; & -(N_{\overline{DE}}^- - I)\nabla v(X) \cdot (A_1 - \widetilde{A}_1)\phi_1(X) - (N_{\overline{DE}}^- - I)\nabla v(X) \cdot (\widetilde{A}_1 - X_{\overline{DE}})\phi_1(X) \\
& -d^- \Big[N_{\overline{DE}}^-(X_{\overline{DE}}^s - X^s) \cdot (A_1 - X_{\overline{DE}})\phi_1(X) + (x_1 - \bar{x})(y_1 - \bar{y})\phi_1(X) \\
& +(\bar{x} - x)(\bar{y} - y)\phi_1(X) + \sum_{i=2}^{4} \Big[(x_i - x)(y_i - y)\phi_i(X) \Big] \Big].
\end{aligned}$$

Proof. By using Lemma 4.1.4 and Lemma 4.1.5, we can get

$$\begin{aligned}
v(A_i) &= v(X) + \nabla v(X) \cdot (A_i - X) + d^-(x_i - x)(y_i - y), i = 2, 3, 4, \\
v(A_1) &= v(X) + \nabla v(X) \cdot (A_1 - X) + (N_{\overline{DE}}^- - I)\nabla v(X) \cdot (A_1 - \widetilde{A}_1) \\
&\quad +(N_{\overline{DE}}^- - I)\nabla v(X) \cdot (\widetilde{A}_1 - X_{\overline{DE}}) + d^- N_{\overline{DE}}^-(X_{\overline{DE}}^s - X^s) \cdot (A_1 - X_{\overline{DE}}) \\
&\quad +d^-(x_1 - \bar{x})(y_1 - \bar{y}) + d^-(\bar{x} - x_1)(\bar{y} - y_1).
\end{aligned}$$

Because $v \in S_h(T)$,

$$\begin{aligned}
v(X) &= I_{h,T}v(X) \\
&= \sum_{i=1}^{4} v(A_i)\phi_i(X) \\
&= v(X)\sum_{i=1}^{4}\phi_i(X) + \nabla v(X) \cdot \sum_{i=1}^{4}(A_i - X)\phi_i(X) \\
&\quad +(N_{\overline{DE}}^- - I)\nabla v(X) \cdot (A_1 - \widetilde{A}_1)\phi_1(X) + (N_{\overline{DE}}^- - I)\nabla v(X) \cdot (\widetilde{A}_1 - X_{\overline{DE}})\phi_1(X) \\
&\quad +d^- \Big[N_{\overline{DE}}^-(X_{\overline{DE}}^s - X^s) \cdot (A_1 - X_{\overline{DE}})\phi_1(X) + (x_1 - \bar{x})(y_1 - \bar{y})\phi_1(X) \\
&\quad +(\bar{x} - x)(\bar{y} - y)\phi_1(X) + \sum_{i=2}^{4}(x_i - x)(y_i - y)\phi_i(X) \Big],
\end{aligned}$$

which leads to the result of this lemma because of Theorem 3.4.1.

■

Lemma 4.1.7 *Given a two-dimensional vector* $\mathbf{q}$*, a point* $X \in \tilde{T}^-$ *and two real numbers* r, d^-*, then there exists a function* $v \in S_h(T)$ *such that* $\nabla v(X) = \mathbf{q}, v(X) = r, \dfrac{\partial^2 v^-(X)}{\partial x \partial y} = d^-$

and

$$\begin{aligned}
& \mathbf{q}\cdot\sum_{i=1}^{4}(A_i-X)\phi_i(X)\\
= \;& -(N^-_{\overline{DE}}-I)\mathbf{q}\cdot(A_1-\widetilde{A}_1)\phi_1(X)-(N^-_{\overline{DE}}-I)\mathbf{q}\cdot(\widetilde{A}_1-X_{\overline{DE}})\phi_1(X)\\
& -d^-\Big[N^-_{\overline{DE}}(X^s_{\overline{DE}}-X^s)\cdot(A_1-X_{\overline{DE}})\phi_1(X)+(x_1-\bar{x})(y_1-\bar{y})\phi_1(X)\\
& +(\bar{x}-x)(\bar{y}-y)\phi_1(X)+\sum_{i=2}^{4}(x_i-x)(y_i-y)\phi_i(X)\Big], \qquad (4.13)
\end{aligned}$$

where $X_{\overline{DE}}$ is an arbitrary point on $\overline{DE}$.

Proof. Let $v(Y)$ be a piecewise bilinear function in term of $Y=(x,y)$ as follows.

$$v(x,y)=\begin{cases} v^-(x,y)=a^-+b^-x+c^-y+d^-xy, & (x,y)\in\tilde{T}^-,\\ v^+(x,y)=a^++b^+x+c^+y+d^+xy, & (x,y)\in\tilde{T}^+. \end{cases}$$

First, $\nabla v(X)=\mathbf{q}, v(X)=r, \dfrac{\partial^2 v^-(X)}{\partial x\partial y}=d^-$ uniquely determine $v^-(Y)$. Then the interface conditions

$$\int_{\overline{DE}}\left(\beta^+\frac{\partial v^+}{\partial \mathbf{n}_{\overline{DE}}}-\beta^-\frac{\partial v^-}{\partial \mathbf{n}_{\overline{DE}}}\right)\,ds\;=0,$$

$v^-(D)=v^+(D), v^-(E)=v^+(E)$, and $d^-=d^+$ uniquely determine $v^+(Y)$. These conditions also imply that $v(Y)$ is a function in the local bilinear IFE space $S_h(T)$. The proof is finished by replacing $\nabla v(X)$ by $\mathbf{q}$ in Lemma 4.1.6.

■

We can now derive an expansion for the bilinear IFE interpolation error at any point $X\in\tilde{T}^-\cap T^-$ of a Type I interface element T.

Theorem 4.1.1 *For any $u \in PC^2_{int}(T)$ and $X = (x, y)^T \in \tilde{T}^- \cap T^-$, we have*

$$
\begin{aligned}
& I_{h,T}u(X) - u(X) \\
= \; & (N^-(\widetilde{A}_1) - N^-_{\overline{DE}})\nabla u(X)\cdot(A_1 - \widetilde{A}_1)\phi_1(X) - (N^-_{\overline{DE}} - I)\nabla u(X)\cdot(\widetilde{A}_1 - X_{\overline{DE}})\phi_1(X) \\
& -\frac{\partial^2 u(X)}{\partial x \partial y}\Big[N^-_{\overline{DE}}(X^s_{\overline{DE}} - X^s)\cdot(A_1 - X_{\overline{DE}})\phi_1(X) + (x_1 - \bar{x})(y_1 - \bar{y})\phi_1(X) \\
& +(\bar{x} - x)(\bar{y} - y)\phi_1(X) + \sum_{i=2}^{4}(x_i - x)(y_i - y)\phi_i(X)\Big] \\
& +(N^-(\widetilde{A}_1) - I)\int_0^1 \frac{d[\nabla u(t\widetilde{A}_1 + (1-t)X)]}{dt}\cdot(A_1 - \widetilde{A}_1)\; dt\; \phi_1(X) \\
& +\int_0^{\tilde{t}}(1-t)\frac{d^2u(tA_1 + (1-t)X)}{dt^2}\; dt\; \phi_1(X) + \int_{\tilde{t}}^1 (1-t)\frac{d^2u(tA_1 + (1-t)X)}{dt^2}\; dt\; \phi_1(X) \\
& +\sum_{i=2}^{4}\int_0^1 (1-t)\frac{d^2u(tA_i + (1-t)X)}{dt^2}\; dt\; \phi_i(X). \qquad (4.14)
\end{aligned}
$$

where $X_{\overline{DE}}$ is an arbitrary point on $\overline{DE}$.

Proof. Since $t \mapsto u(tA_i + (1-t)X))$, $i = 2, 3, 4$ are C^2 functions in terms of t, using integration by parts, we have

$$
\begin{aligned}
u(A_i) &= u(X) + \int_0^1 \frac{d\; u(tA_i + (1-t)X)}{dt}\; dt \\
&= u(X) - \int_0^1 \frac{d\; u(tA_i + (1-t)X)}{dt}\; d(1-t) \\
&= u(X) - \frac{d\; u(tA_i + (1-t)X)}{dt}(1-t)\Big|_{t=0}^{t=1} + \int_0^1 (1-t)\frac{d^2u(tA_i + (1-t)X)}{dt^2}\; dt \\
&= u(X) + \nabla u(X)\cdot(A_i - X) + \int_0^1 (1-t)\frac{d^2u(tA_i + (1-t)X)}{dt^2}\; dt. \qquad (4.15)
\end{aligned}
$$

Using integration by parts, (4.12) and (4.1), we have

$$
\begin{aligned}
u(A_1) &= u(X) + \int_0^{\tilde{t}} \frac{d\ u(tA_1+(1-t)X)}{dt}\ dt\ + \int_{\tilde{t}}^1 \frac{d\ u(tA_1+(1-t)X)}{dt}\ dt \\
&= u(X) - \int_0^{\tilde{t}} \frac{d\ u(tA_1+(1-t)X)}{dt}\ d(1-t)\ - \int_{\tilde{t}}^1 \frac{d\ u(tA_1+(1-t)X)}{dt}\ d(1-t) \\
&= u(X) - \frac{d\ u(tA_1+(1-t)X)}{dt}(1-t)\Big|_{t=0}^{t=\tilde{t}} + \int_0^{\tilde{t}} (1-t)\frac{d^2u(tA_1+(1-t)X)}{dt^2}\ dt \\
&\quad -\frac{d\ u(tA_1+(1-t)X)}{dt}(1-t)\Big|_{t=\tilde{t}}^{t=1} + \int_{\tilde{t}}^1 (1-t)\frac{d^2u(tA_1+(1-t)X)}{dt^2}\ dt \\
&= u(X) - (1-\tilde{t})\nabla u^-(\widetilde{A}_1)\cdot(A_1-X) + \nabla u(X)\cdot(A_1-X) \\
&\quad + \int_0^{\tilde{t}} (1-t)\frac{d^2u(tA_1+(1-t)X)}{dt^2}\ dt + (1-\tilde{t})\nabla u^+(\widetilde{A}_1)\cdot(A_1-X) \\
&\quad + \int_{\tilde{t}}^1 (1-t)\frac{d^2u(tA_1+(1-t)X)}{dt^2}\ dt \\
&= u(X) - \nabla u^-(\widetilde{A}_1)\cdot(A_1-\widetilde{A}_1) + \nabla u(X)\cdot(A_1-X) \\
&\quad + \int_0^{\tilde{t}} (1-t)\frac{d^2u(tA_1+(1-t)X)}{dt^2}\ dt + \nabla u^+(\widetilde{A}_1)\cdot(A_1-\widetilde{A}_1) \\
&\quad + \int_{\tilde{t}}^1 (1-t)\frac{d^2u(tA_1+(1-t)X)}{dt^2}\ dt \\
&= u(X) + \nabla u(X)\cdot(A_1-X) + (N^-(\widetilde{A}_1)-I)\nabla u^-(\widetilde{A}_1)\cdot(A_1-\widetilde{A}_1) \\
&\quad + \int_0^{\tilde{t}} (1-t)\frac{d^2u(tA_1+(1-t)X)}{dt^2}\ dt\ + \int_{\tilde{t}}^1 (1-t)\frac{d^2u(tA_1+(1-t)X)}{dt^2}\ dt \\
&= u(X) + \nabla u(X)\cdot(A_1-X) + (N^-(\widetilde{A}_1)-I)\nabla u(X)\cdot(A_1-\widetilde{A}_1) \\
&\quad +(N^-(\widetilde{A}_1)-I)\int_0^1 \frac{d[\nabla u(t\widetilde{A}_1+(1-t)X)]}{dt}\cdot(A_1-\widetilde{A}_1)\ dt \\
&\quad + \int_0^{\tilde{t}} (1-t)\frac{d^2u(tA_1+(1-t)X)}{dt^2}\ dt \\
&\quad + \int_{\tilde{t}}^1 (1-t)\frac{d^2u(tA_1+(1-t)X)}{dt^2}\ dt. \qquad (4.16)
\end{aligned}
$$

Then

$$\begin{aligned} I_{h,T}u(X) &= \sum_{i=1}^{4} u(A_i)\phi_i(X) \\ &= u(X)\sum_{i=1}^{4}\phi_i(X) + \nabla u(X)\cdot\sum_{i=1}^{4}(A_i - X)\phi_i(X) \\ &\quad +(N^-(\widetilde{A}_1) - I)\nabla u(X)\cdot(A_1 - \widetilde{A}_1)\phi_1(X) \\ &\quad +(N^-(\widetilde{A}_1) - I)\int_0^1 \frac{d[\nabla u(t\widetilde{A}_1 + (1-t)X)]}{dt}\cdot(A_1 - \widetilde{A}_1)\ dt\ \phi_1(X) \\ &\quad + \int_0^{\tilde{t}} (1-t)\frac{d^2u(tA_1 + (1-t)X)}{dt^2}\ dt\ \phi_1(X) \\ &\quad + \int_{\tilde{t}}^1 (1-t)\frac{d^2u(tA_1 + (1-t)X)}{dt^2}\ dt\ \phi_1(X) \\ &\quad + \sum_{i=2}^{4}\int_0^1 (1-t)\frac{d^2u(tA_i + (1-t)X)}{dt^2}\ dt\ \phi_i(X). \end{aligned} \tag{4.17}$$

Now letting $\mathbf{q} = \nabla u(X), r = u(X), d^- = \dfrac{\partial^2 u(X)}{\partial x \partial y}$ in Lemma 4.1.7, we have

$$\begin{aligned} &\nabla u(X)\cdot\sum_{i=1}^{4}(A_i - X)\phi_i(X) \\ = \ &-(N^-_{\overline{DE}} - I)\nabla u(X)\cdot(A_1 - \widetilde{A}_1)\phi_1(X) - (N^-_{\overline{DE}} - I)\nabla u(X)\cdot(\widetilde{A}_1 - X_{\overline{DE}})\phi_1(X) \\ &-\frac{\partial^2 u(X)}{\partial x\partial y}\Big[N^-_{\overline{DE}}(X^s_{\overline{DE}} - X^s)\cdot(A_1 - X_{\overline{DE}})\phi_1(X) + (x_1 - \bar{x})(y_1 - \bar{y})\phi_1(X) \\ &+(\bar{x} - x)(\bar{y} - y)\phi_1(X) + \sum_{i=2}^{4}(x_i - x)(y_i - y)\phi_i(X)\Big]. \end{aligned} \tag{4.18}$$

Finally, (4.14) follows from (4.17), (4.18) and Theorem 3.4.1.

■

The following theorem establish a bound in L^2 norm for the bilinear IFE interpolation error.

Theorem 4.1.2 *There exits a constant C independent of interface and mesh such that*

$$\|I_{h,T}u - u\|_{0,\tilde{T}^-\cap T^-} \le Ch^2\left(|u|_{1,\tilde{T}^-\cap T^-} + |u|_{2,\tilde{T}^-\cap T^-}\right) \le Ch^2\|u\|_{2,T} \tag{4.19}$$

for any $u \in PH^2_{int}(T)$, where $T \in \mathcal{T}_h$ is a Type I interface element.

Proof. Let Q_i, $i = 1, 2, \cdots, 9$ be the 9 terms on the right hand side of (4.14), and we proceed by estimating their L^2 norms. By Lemma 4.1.3, Theorem 3.4.2, and by letting $X_{\overline{DE}} = \widetilde{A}_1^{\perp}$ in (4.14), we have the following estimate for the L^2 norms of the first three terms:

$$\begin{aligned}
& \|Q_1\|_{0,\tilde{T}^-\cap T^-} + \|Q_2\|_{0,\tilde{T}^-\cap T^-} + \|Q_3\|_{0,\tilde{T}^-\cap T^-} \\
= \; & \Big\| \big(N^-(\widetilde{A}_1) - N^-_{\overline{DE}}\big)\nabla u(X)\cdot(A_1 - \widetilde{A}_1)\phi_1(X) \Big\|_{0,\tilde{T}^-\cap T^-} \\
& + \Big\| (N^-_{\overline{DE}} - I)\nabla u(X)\cdot(\widetilde{A}_1 - \widetilde{A}_1^{\perp})\phi_1(X) \Big\|_{0,\tilde{T}^-\cap T^-} \\
& + \Big\| \frac{\partial^2 u(X)}{\partial x \partial y}\Big[N^-_{\overline{DE}}(X^s_{\overline{DE}} - X^s)\cdot(A_1 - X_{\overline{DE}})\phi_1(X) + (x_1 - \bar{x})(y_1 - \bar{y})\phi_1(X) \\
& + (\bar{x} - x)(\bar{y} - y)\phi_1(X) + \sum_{i=2}^{4}(x_i - x)(y_i - y)\phi_i(X)\Big] \Big\|_{0,\tilde{T}^-\cap T^-} \\
\leq \; & Ch^2 \left|u\right|_{1,\tilde{T}^-\cap T^-} + Ch^2 \left|u\right|_{2,\tilde{T}^-\cap T^-} \\
\leq \; & Ch^2 \left\|u\right\|_{2,T}.
\end{aligned}$$

For the fourth term, we first note that

$$\begin{aligned}
& \frac{d[\nabla u(t\widetilde{A}_1 + (1-t)X)]}{dt}\cdot(A_1 - \widetilde{A}_1) \\
= \; & u_{\xi\xi}(\xi,\eta)(\widetilde{x}_1 - x)(x_1 - \widetilde{x}_1) + 2u_{\xi\eta}(\xi,\eta)[(\widetilde{y}_1 - y)(x_1 - \widetilde{x}_1) + (\widetilde{x}_1 - x)(y_1 - \widetilde{y}_1)] \\
& + u_{\eta\eta}(\xi,\eta)(\widetilde{y}_1 - y)(y_1 - \widetilde{y}_1)
\end{aligned}$$

with $\xi = t\widetilde{x}_1 + (1-t)x, \eta = t\widetilde{y}_1 + (1-t)y$. Then,

$$\begin{aligned}
Q_4^2 \; \leq \; & C\Big(\int_0^1 \big[u_{\xi\xi}(\xi,\eta)(\widetilde{x}_1 - x)(x_1 - \widetilde{x}_1) + \\
& 2u_{\xi\eta}(\xi,\eta)\Big((\widetilde{y}_1 - y)(x_1 - \widetilde{x}_1) + (\widetilde{x}_1 - x)(y_1 - \widetilde{y}_1)\Big) + u_{\eta\eta}(\xi,\eta)(\widetilde{y}_1 - y)(y_1 - \widetilde{y}_1)\big]dt\Big)^2 \\
\leq \; & Ch^4\left(\int_0^1 \big[u_{\xi\xi}(\xi,\eta) + 2u_{\xi\eta}(\xi,\eta) + u_{\eta\eta}(\xi,\eta)\big]dt\right)^2 \\
\leq \; & Ch^4 \int_0^1 \left[u^2_{\xi\xi}(\xi,\eta) + u^2_{\xi\eta}(\xi,\eta) + u^2_{\eta\eta}(\xi,\eta)\right] dt,
\end{aligned}$$

here C stands for a generic constant. Since both X and $\widetilde{A}_1$ are in T^-, then $(\xi,\eta) \in T^-$.

Therefore,

$$\begin{aligned}
\|Q_4\|^2_{0,\tilde{T}^-\cap T^-} &= \int_{\tilde{T}^-\cap T^-} Q_4^2 d\xi d\eta \\
&\leq Ch^4 \int_{\tilde{T}^-\cap T^-} \int_0^1 \left[u^2_{\xi\xi}(\xi,\eta) + u^2_{\xi\eta}(\xi,\eta) + u^2_{\eta\eta}(\xi,\eta)\right] dt d\xi d\eta \\
&\leq Ch^4 \int_{\tilde{T}^-\cap T^-} \left[u^2_{\xi\xi}(\xi,\eta) + u^2_{\xi\eta}(\xi,\eta) + u^2_{\eta\eta}(\xi,\eta)\right] d\xi d\eta \\
&\leq Ch^4 \left|u\right|^2_{2,\tilde{T}^-\cap T^-},
\end{aligned}$$

hence

$$\|Q_4\|_{0,\tilde{T}^-\cap T^-} \leq Ch^2 \left|u\right|_{2,\tilde{T}^-\cap T^-} \leq Ch^2 \left\|u\right\|_{2,T}.$$

For the fifth term, we have

$$\begin{aligned}
Q_5^2 &\leq C\left(\int_0^{\tilde{t}} (1-t)\left[u_{\xi\xi}(\xi,\eta)(x_1-x)^2 + 2u_{\xi\eta}(\xi,\eta)(x_1-x)(y_1-y) + u_{\eta\eta}(\xi,\eta)(y_1-y)^2\right] dt\right)^2 \\
&\leq Ch^4\left(\int_0^{\tilde{t}} (1-t)\left[u_{\xi\xi}(\xi,\eta) + 2u_{\xi\eta}(\xi,\eta) + u_{\eta\eta}(\xi,\eta)\right] dt\right)^2 \\
&\leq Ch^4 \int_0^{\tilde{t}} (1-t)^2\left[u^2_{\xi\xi}(\xi,\eta) + u^2_{\xi\eta}(\xi,\eta) + u^2_{\eta\eta}(\xi,\eta)\right] dt
\end{aligned}$$

with $\xi = tx_1 + (1-t)x, \eta = ty_1 + (1-t)y$. Then

$$\begin{aligned}
\|Q_5\|^2_{\tilde{T}^-\cap T^-} &\leq Ch^4 \int_{\tilde{T}^-\cap T^-} \int_0^{\tilde{t}} (1-t)^2\left[u^2_{\xi\xi}(\xi,\eta) + u^2_{\xi\eta}(\xi,\eta) + u^2_{\eta\eta}(\xi,\eta)\right] dt d\xi d\eta \\
&\leq Ch^4 \int_{\tilde{T}^-\cap T^-} \left[u^2_{\xi\xi}(\xi,\eta) + u^2_{\xi\eta}(\xi,\eta) + u^2_{\eta\eta}(\xi,\eta)\right] d\xi d\eta \leq Ch^4 \left|u\right|^2_{2,\tilde{T}^-\cap T^-},
\end{aligned}$$

hence

$$\|Q_5\|_{\tilde{T}^-\cap T^-} \leq Ch^2 \left|u\right|_{2,\tilde{T}^-\cap T^-} \leq Ch^2 \left\|u\right\|_{2,T}.$$

Similarly, we can show that $\|Q_i\|_{\tilde{T}^-\cap T^-} \leq Ch^2 \|u\|_{2,T}$, $i = 6,7,8,9$. Finally, (4.19) follows from the estimates for Q_i, $i = 1,2,\cdots,9$ above.

■

We now turn to the estimate of bilinear IFE interpolation error in H^1 norm on the subelement T^-. In the discussion below, we let I_i, $i = 1,2,3,4$ be the integral terms involving the vertices A_i, $i = 1,2,3,4$ in (4.14).

Theorem 4.1.3 *For any $u \in PC^3_{int}(T)$ and $X = (x, y)^T \in \tilde{T}^- \cap T^-$, we have*

$$\begin{aligned}
&\frac{\partial(I_{h,T}u(X) - u(X))}{\partial x} \\
= \;& (N^-(\widetilde{A}_1) - N^-_{\overline{DE}})\nabla u(X)\cdot(A_1 - \widetilde{A}_1)\frac{\partial \phi_1(X)}{\partial x} - (N^-_{\overline{DE}} - I)\nabla u(X)\cdot(\widetilde{A}_1 - X_{\overline{DE}})\frac{\partial \phi_1(X)}{\partial x} \\
& -\frac{\partial^2 u(X)}{\partial x \partial y}\Big[N^-_{\overline{DE}}(X^s_{\overline{DE}} - X^s)\cdot(A_1 - X_{\overline{DE}})\frac{\partial \phi_1(X)}{\partial x} + N^-_{\overline{DE}}(0,-1)^T\cdot(A_1 - X_{\overline{DE}})\phi_1(X) \\
& +(x_1 - \bar{x})(y_1 - \bar{y})\frac{\partial \phi_1(X)}{\partial x} - (\bar{y} - y)\phi_1(X) + (\bar{x} - x)(\bar{y} - y)\frac{\partial \phi_1(X)}{\partial x} \\
& +\sum_{i=2}^{4}\Big[-(y_i - y)\phi_i(X) + (x_i - x)(y_i - y)\frac{\partial \phi_i(X)}{\partial x}\Big]\Big] + \sum_{i=1}^{4} I_i \frac{\partial \phi_i(X)}{\partial x}, \qquad (4.20)
\end{aligned}$$

$$\begin{aligned}
&\frac{\partial(I_{h,T}u(X) - u(X))}{\partial y} \\
= \;& (N^-(\widetilde{A}_1) - N^-_{\overline{DE}})\nabla u(X)\cdot(A_1 - \widetilde{A}_1)\frac{\partial \phi_1(X)}{\partial y} - (N^-_{\overline{DE}} - I)\nabla u(X)\cdot(\widetilde{A}_1 - X_{\overline{DE}})\frac{\partial \phi_1(X)}{\partial y} \\
& -\frac{\partial^2 u(X)}{\partial x \partial y}\Big[N^-_{\overline{DE}}(X^s_{\overline{DE}} - X^s)\cdot(A_1 - X_{\overline{DE}})\frac{\partial \phi_1(X)}{\partial y} + N^-_{\overline{DE}}(-1,0)^T\cdot(A_1 - X_{\overline{DE}})\phi_1(X) \\
& +(x_1 - \bar{x})(y_1 - \bar{y})\frac{\partial \phi_1(X)}{\partial y} - (\bar{x} - x)\phi_1(X) + (\bar{x} - x)(\bar{y} - y)\frac{\partial \phi_1(X)}{\partial y} \\
& +\sum_{i=2}^{4}\Big[-(x_i - x)\phi_i(X) + (x_i - x)(y_i - y)\frac{\partial \phi_i(X)}{\partial y}\Big]\Big] + \sum_{i=1}^{4} I_i \frac{\partial \phi_i(X)}{\partial y}, \qquad (4.21)
\end{aligned}$$

where $X_{\overline{DE}}$ is an arbitrary point on $\overline{DE}$.

Proof. We give a proof only for (4.20), similar arguments can be used to show (4.21). Taking

derivative for x on both sides of (4.14), we can get

$$\begin{aligned}
&\frac{\partial(I_{h,T}u(X)-u(X))}{\partial x}\\
=\;&\frac{\partial}{\partial x}\Big[(N^-(\widetilde{A}_1)-N^-_{\overline{DE}})\nabla u(X)\cdot(A_1-\widetilde{A}_1)\Big]\phi_1(X)\\
&+(N^-(\widetilde{A}_1)-N^-_{\overline{DE}})\nabla u(X)\cdot(A_1-\widetilde{A}_1)\frac{\partial\phi_1(X)}{\partial x}\\
&-\frac{\partial}{\partial x}\Big[(N^-_{\overline{DE}}-I)\nabla u(X)\cdot(\widetilde{A}_1-X_{\overline{DE}})\Big]\phi_1(X)-(N^-_{\overline{DE}}-I)\nabla u(X)\cdot(\widetilde{A}_1-X_{\overline{DE}})\frac{\partial\phi_1(X)}{\partial x}\\
&-\frac{\partial^3u(X)}{\partial x^2\partial y}\Big[N^-_{\overline{DE}}(X^s_{\overline{DE}}-X^s)\cdot(A_1-X_{\overline{DE}})\phi_1(X)+(x_1-\bar{x})(y_1-\bar{y})\phi_1(X)\\
&+(\bar{x}-x)(\bar{y}-y)\phi_1(X)+\sum_{i=2}^{4}(x_i-x)(y_i-y)\phi_i(X)\Big]\\
&-\frac{\partial^2u(X)}{\partial x\partial y}\Big[N^-_{\overline{DE}}(X^s_{\overline{DE}}-X^s)\cdot(A_1-X_{\overline{DE}})\frac{\partial\phi_1(X)}{\partial x}+N^-_{\overline{DE}}(0,-1)^T\cdot(A_1-X_{\overline{DE}})\phi_1(X)\\
&+(x_1-\bar{x})(y_1-\bar{y})\frac{\partial\phi_1(X)}{\partial x}-(\bar{y}-y)\phi_1(X)+(\bar{x}-x)(\bar{y}-y)\frac{\partial\phi_1(X)}{\partial x}\\
&+\sum_{i=2}^{4}\Big[-(y_i-y)\phi_i(X)+(x_i-x)(y_i-y)\frac{\partial\phi_i(X)}{\partial x}\Big]\Big]\\
&+\sum_{i=1}^{4}I_i\frac{\partial\phi_i(X)}{\partial x}+\sum_{i=1}^{4}\frac{\partial I_i}{\partial x}\phi_i(X). \qquad (4.22)
\end{aligned}$$

Taking the first derivative with respect to x on both sides of (4.15) and (4.16), we can get

$$\begin{aligned}
\frac{\partial I_i}{\partial x} &= -P\cdot(A_i-X), i=2,3,4,\\
\frac{\partial I_1}{\partial x} &= -P\cdot(A_1-X)-\frac{\partial}{\partial x}\Big[(N^-(\widetilde{A}_1)-I)\nabla u(X)(A_1-\widetilde{A}_1)\Big],
\end{aligned}$$

where

$$P=\frac{\partial}{\partial x}\nabla u(X)=(\frac{\partial^2u(X)}{\partial x^2},\frac{\partial^2u(X)}{\partial x\partial y})^T.$$

Hence,

$$\sum_{i=1}^{4}\frac{\partial I_i}{\partial x}\phi_i(X) = -P\cdot\sum_{i=1}^{4}(A_i-X)\phi_i(X)-\frac{\partial}{\partial x}\Big[(N^-(\widetilde{A}_1)-I)\nabla u(X)(A_1-\widetilde{A}_1)\Big]\phi_1(X).$$

Applying Lemma 4.1.7 to the first term on the right hand side above, letting $\mathbf{q}=P, d^-=$

$\frac{\partial^3 u(X)}{\partial x^2 \partial y}$, we have

$$\begin{aligned}
&\sum_{i=1}^{4} \frac{\partial I_i}{\partial x}\phi_i(X) \\
= \; & -\frac{\partial}{\partial x}\Big[(N^-(\widetilde{A}_1) - I)\nabla u(X)\cdot(A_1 - \widetilde{A}_1)\Big]\phi_1(X) + (N^-_{\overline{DE}} - I)P\cdot(A_1 - \widetilde{A}_1)\phi_1(X) \\
& + (N^-_{\overline{DE}} - I)P\cdot(\widetilde{A}_1 - X_{\overline{DE}})\phi_1(X) + \frac{\partial^3 u(X)}{\partial x^2\partial y}\Big[N^-_{\overline{DE}}(X^s_{\overline{DE}} - X^s)\cdot(A_1 - X_{\overline{DE}})\phi_1(X) \\
& + (x_1 - \bar{x})(y_1 - \bar{y})\phi_1(X) + (\bar{x} - x)(\bar{y} - y)\phi_1(X) + \sum_{i=2}^{4}(x_i - x)(y_i - y)\phi_i(X)\Big]. \quad (4.23)
\end{aligned}$$

By direct calculations, we also have

$$\begin{aligned}
& \frac{\partial}{\partial x}\left[(N^-(\widetilde{A}_1) - N^-_{\overline{DE}})\nabla u(X)\cdot(A_1 - \widetilde{A}_1)\right]\phi_1(X) \\
& -\frac{\partial}{\partial x}\left[(N^-_{\overline{DE}} - I)\nabla u(X)\cdot(\widetilde{A}_1 - X_{\overline{DE}})\right]\phi_1(X) \\
& -\frac{\partial}{\partial x}\left[(N^-(\widetilde{A}_1) - I)\nabla u(X)\cdot(A_1 - \widetilde{A}_1)\right]\phi_1(X) \\
& + (N^-_{\overline{DE}} - I)P\cdot(A_1 - \widetilde{A}_1)\phi_1(X) + (N^-_{\overline{DE}} - I)P\cdot(\widetilde{A}_1 - X_{\overline{DE}})\phi_1(X) \\
= \; & -\frac{\partial}{\partial x}\left[(N^-_{\overline{DE}} - I)\nabla u(X)\cdot(A_1 - \widetilde{A}_1)\right]\phi_1(X) \\
& -\frac{\partial}{\partial x}\left[(N^-_{\overline{DE}} - I)\nabla u(X)\cdot(\widetilde{A}_1 - X_{\overline{DE}})\right]\phi_1(X) + (N^-_{\overline{DE}} - I)P\cdot(A_1 - X_{\overline{DE}})\phi_1(X) \\
= \; & -\frac{\partial}{\partial x}\left[(N^-_{\overline{DE}} - I)\nabla u(X)\cdot(A_1 - X_{\overline{DE}})\right]\phi_1(X) + (N^-_{\overline{DE}} - I)P\cdot(A_1 - X_{\overline{DE}})\phi_1(X) \\
= \; & -(N^-_{\overline{DE}} - I)P\cdot(A_1 - X_{\overline{DE}})\phi_1(X) + (N^-_{\overline{DE}} - I)P\cdot(A_1 - X_{\overline{DE}})\phi_1(X) \\
= \; & 0. \qquad (4.24)
\end{aligned}$$

Plugging (4.23) and (4.24) into (4.22), we finish the proof of (4.20).

■

Based on the above expansion, we get the following theorem.

Theorem 4.1.4 *There exits a constant C independent of interface and mesh such that*

$$|I_{h,T}u - u|_{1,\tilde{T}^-\cap T^-} \le Ch\left(|u|_{1,\tilde{T}^-\cap T^-} + |u|_{2,\tilde{T}^-\cap T^-}\right) \le Ch\,\|u\|_{2,T}$$

for any $u \in PH^2_{int}(T)$, where T is a Type I interface element.

Proof. The result follows by letting $X_{\overline{DE}} = \widetilde{A}_1^{\perp}$ in (4.20) and (4.21), applying Theorem 3.4.2, and applying arguments similar to those used in the proof of Theorem 4.1.2. Note that (3.12) in Theorem 3.4.2 has to be used here.

■

We now turn to the estimate of bilinear IFE interpolation error in H^2 norm on the subelement $\tilde{T}^- \cap T^-$.

Theorem 4.1.5 *For any $u \in PC^3_{int}(T)$ and $X = (x, y)^T \in \tilde{T}^- \cap T^-$, we have*

$$\begin{aligned}
&\frac{\partial^2 (I_{h,T}u(X) - u(X))}{\partial x \partial y} \\
= \;& (N^-(\widetilde{A}_1) - N^-_{\overline{DE}})\nabla u(X) \cdot (A_1 - \widetilde{A}_1)\frac{\partial^2 \phi_1(X)}{\partial x \partial y} \\
& -(N^-_{\overline{DE}} - I)\nabla u(X) \cdot (\widetilde{A}_1 - X_{\overline{DE}})\frac{\partial^2 \phi_1(X)}{\partial x \partial y} \\
& -\frac{\partial^2 u(X)}{\partial x \partial y}\Big[N^-_{\overline{DE}}(X^s_{\overline{DE}} - X^s) \cdot (A_1 - X_{\overline{DE}})\frac{\partial^2 \phi_1(X)}{\partial x \partial y} \\
& +N^-_{\overline{DE}}(-1, 0)^T \cdot (A_1 - X_{\overline{DE}})\frac{\partial \phi_1(X)}{\partial x} + N^-_{\overline{DE}}(0, -1)^T \cdot (A_1 - X_{\overline{DE}})\frac{\partial \phi_1(X)}{\partial y} \\
& +(x_1 - \bar{x})(y_1 - \bar{y})\frac{\partial^2 \phi_1(X)}{\partial x \partial y} + \phi_1(X) - (\bar{y} - y)\frac{\partial \phi_1(X)}{\partial y} - (\bar{x} - x)\frac{\partial \phi_1(X)}{\partial x} \\
& +(\bar{x} - x)(\bar{y} - y)\frac{\partial^2 \phi_1(X)}{\partial x \partial y} + \sum_{i=2}^{4}\Big[\phi_i(X) - (y_i - y)\frac{\partial \phi_i(X)}{\partial y} - (x_i - x)\frac{\partial \phi_i(X)}{\partial x} \\
& +(x_i - x)(y_i - y)\frac{\partial^2 \phi_i(X)}{\partial x \partial y}\Big]\Big] + \sum_{i=1}^{4} I_i \frac{\partial^2 \phi_i(X)}{\partial x \partial y}, \qquad (4.25)
\end{aligned}$$

where $X_{\overline{DE}}$ is an arbitrary point on $\overline{DE}$.

Proof. Taking derivative for y on both sides of (4.20), we can get

$$
\begin{aligned}
&\frac{\partial^2(I_{h,T}u(X)-u(X))}{\partial x\partial y}\\
=\;&\frac{\partial}{\partial y}\Big[(N^-(\widetilde{A}_1)-N^-_{\overline{DE}})\nabla u(X)\cdot(A_1-\widetilde{A}_1)\Big]\frac{\partial\phi_1(X)}{\partial x}\\
&+(N^-(\widetilde{A}_1)-N^-_{\overline{DE}})\nabla u(X)\cdot(A_1-\widetilde{A}_1)\frac{\partial^2\phi_1(X)}{\partial x\partial y}\\
&-\frac{\partial}{\partial y}\Big[(N^-_{\overline{DE}}-I)\nabla u(X)\cdot(\widetilde{A}_1-X_{\overline{DE}})\Big]\frac{\partial\phi_1(X)}{\partial x}\\
&-(N^-_{\overline{DE}}-I)\nabla u(X)\cdot(\widetilde{A}_1-X_{\overline{DE}})\frac{\partial^2\phi_1(X)}{\partial x\partial y}\\
&-\frac{\partial^3u(X)}{\partial x\partial y^2}\Big[N^-_{\overline{DE}}(X^s_{\overline{DE}}-X^s)\cdot(A_1-X_{\overline{DE}})\frac{\partial\phi_1(X)}{\partial x}\\
&+N^-_{\overline{DE}}(0,-1)^T\cdot(A_1-X_{\overline{DE}})\phi_1(X)+(x_1-\bar{x})(y_1-\bar{y})\frac{\partial\phi_1(X)}{\partial x}\\
&-(\bar{y}-y)\phi_1(X)+(\bar{x}-x)(\bar{y}-y)\frac{\partial\phi_1(X)}{\partial x}\\
&+\sum_{i=2}^{4}\Big[-(y_i-y)\phi_i(X)+(x_i-x)(y_i-y)\frac{\partial\phi_i(X)}{\partial x}\Big]\Big]\\
&-\frac{\partial^2u(X)}{\partial x\partial y}\Big[N^-_{\overline{DE}}(X^s_{\overline{DE}}-X^s)\cdot(A_1-X_{\overline{DE}})\frac{\partial^2\phi_1(X)}{\partial x\partial y}\\
&+N^-_{\overline{DE}}(-1,0)^T\cdot(A_1-X_{\overline{DE}})\frac{\partial\phi_1(X)}{\partial x}+N^-_{\overline{DE}}(0,-1)^T\cdot(A_1-X_{\overline{DE}})\frac{\partial\phi_1(X)}{\partial y}\\
&+(x_1-\bar{x})(y_1-\bar{y})\frac{\partial^2\phi_1(X)}{\partial x\partial y}+\phi_1(X)-(\bar{y}-y)\frac{\partial\phi_1(X)}{\partial y}\\
&-(\bar{x}-x)\frac{\partial\phi_1(X)}{\partial x}+(\bar{x}-x)(\bar{y}-y)\frac{\partial^2\phi_1(X)}{\partial x\partial y}\\
&+\sum_{i=2}^{4}\Big[\phi_i(X)-(y_i-y)\frac{\partial\phi_i(X)}{\partial y}-(x_i-x)\frac{\partial\phi_i(X)}{\partial x}+(x_i-x)(y_i-y)\frac{\partial^2\phi_i(X)}{\partial x\partial y}\Big]\Big]\\
&+\sum_{i=1}^{4}\frac{\partial I_i}{\partial y}\frac{\partial\phi_i(X)}{\partial x}+\sum_{i=1}^{4}I_i\frac{\partial^2\phi_i(X)}{\partial x\partial y}. \qquad (4.26)
\end{aligned}
$$

Taking the first derivative with respect to y on both sides of (4.15) and (4.16), we can get

$$
\begin{aligned}
&\frac{\partial I_i}{\partial y}=-P\cdot(A_i-X),i=2,3,4,\\
&\frac{\partial I_1}{\partial y}=-P\cdot(A_1-X)-\frac{\partial}{\partial y}\Big[(N^-(\widetilde{A}_1)-I)\nabla u(X)(A_1-\widetilde{A}_1)\Big],
\end{aligned}
$$

where

$$P = \frac{\partial}{\partial y}\nabla u(X) = (\frac{\partial^2 u(X)}{\partial x \partial y}, \frac{\partial^2 u(X)}{\partial y^2})^T.$$

Hence,

$$\begin{aligned}
&\sum_{i=1}^{4} \frac{\partial I_i}{\partial y}\frac{\partial \phi_i(X)}{\partial x} \\
= \; & -P \cdot \sum_{i=1}^{4}(A_i - X)\frac{\partial \phi_i(X)}{\partial x} - \frac{\partial}{\partial y}\Big[(N^-(\widetilde{A}_1) - I)\nabla u(X)(A_1 - \widetilde{A}_1)\Big]\frac{\partial \phi_1(X)}{\partial x}. \quad (4.27)
\end{aligned}$$

Taking the derivative for x on both sides of (4.13), we get

$$\begin{aligned}
&\mathbf{q} \cdot \sum_{i=1}^{4}(A_i - X)\frac{\partial \phi_i(X)}{\partial x} \\
= \; & -(N^-_{\overline{DE}} - I)\mathbf{q} \cdot (A_1 - \widetilde{A}_1)\frac{\partial \phi_1(X)}{\partial x} - (N^-_{\overline{DE}} - I)\mathbf{q} \cdot (\widetilde{A}_1 - X_{\overline{DE}})\frac{\partial \phi_1(X)}{\partial x} \\
& -d^-\Big[N^-_{\overline{DE}}(0,-1)^T \cdot (A_1 - X_{\overline{DE}})\phi_1(X) + N^-_{\overline{DE}}(X^s_{\overline{DE}} - X^s) \cdot (A_1 - X_{\overline{DE}})\frac{\partial \phi_1(X)}{\partial x} \\
& +(x_1 - \bar{x})(y_1 - \bar{y})\frac{\partial \phi_1(X)}{\partial x} - (\bar{y} - y)\phi_1(X) + (\bar{x} - x)(\bar{y} - y)\frac{\partial \phi_1(X)}{\partial x} \\
& -\sum_{i=2}^{4}(y_i - y)\phi_i(X) + \sum_{i=2}^{4}(x_i - x)(y_i - y)\frac{\partial \phi_1(X)}{\partial x}\Big]. \quad (4.28)
\end{aligned}$$

Applying (4.28) to the first term on the right hand side of (4.27) and letting $\mathbf{q} = P, d^- = \frac{\partial^3 u(X)}{\partial x \partial y^2}$, we have

$$\begin{aligned}
&\sum_{i=1}^{4} \frac{\partial I_i}{\partial y}\frac{\partial \phi_i(X)}{\partial x} \\
= \; & -\frac{\partial}{\partial y}\Big[(N^-(\widetilde{A}_1) - I)\nabla u(X)(A_1 - \widetilde{A}_1)\Big]\frac{\partial \phi_1(X)}{\partial x} + (N^-_{\overline{DE}} - I)P \cdot (A_1 - \widetilde{A}_1)\frac{\partial \phi_1(X)}{\partial x} \\
& +(N^-_{\overline{DE}} - I)P \cdot (\widetilde{A}_1 - X_{\overline{DE}})\frac{\partial \phi_1(X)}{\partial x} \\
& -\frac{\partial^3 u(X)}{\partial x \partial y^2}\Big[N^-_{\overline{DE}}(0,-1)^T \cdot (A_1 - X_{\overline{DE}})\phi_1(X) + N^-_{\overline{DE}}(X^s_{\overline{DE}} - X^s) \cdot (A_1 - X_{\overline{DE}})\frac{\partial \phi_1(X)}{\partial x} \\
& +(x_1 - \bar{x})(y_1 - \bar{y})\frac{\partial \phi_1(X)}{\partial x} - (\bar{y} - y)\phi_1(X) + (\bar{x} - x)(\bar{y} - y)\frac{\partial \phi_1(X)}{\partial x} \\
& +\sum_{i=2}^{4}\Big[-(y_i - y)\phi_i(X) + (x_i - x)(y_i - y)\frac{\partial \phi_i(X)}{\partial x}\Big]\Big]. \quad (4.29)
\end{aligned}$$

By direct calculations similar to (4.24), we also have

$$\begin{aligned}
&\frac{\partial}{\partial y}(N^-(\widetilde{A}_1) - N^-_{\overline{DE}})\nabla u(X)\cdot(A_1-\widetilde{A}_1)\frac{\partial \phi_1(X)}{\partial x}\\
&-\frac{\partial}{\partial y}\Big[(N^-_{\overline{DE}}-I)\nabla u(X)\cdot(\widetilde{A}_1 - X_{\overline{DE}})\Big]\frac{\partial \phi_1(X)}{\partial x}\\
&-\frac{\partial}{\partial y}\Big[(N^-(\widetilde{A}_1)-I)\nabla u(X)(A_1-\widetilde{A}_1)\Big]\frac{\partial \phi_1(X)}{\partial x}\\
&+(N^-_{\overline{DE}}-I)P\cdot(A_1-\widetilde{A}_1)\frac{\partial \phi_1(X)}{\partial x} + (N^-_{\overline{DE}}-I)P\cdot(\widetilde{A}_1 - X_{\overline{DE}})\frac{\partial \phi_1(X)}{\partial x}\\
=\ & 0. \qquad (4.30)
\end{aligned}$$

Plugging (4.29) and (4.30) into (4.26), we finish the proof of (4.25).

■

Based on the above expansion, we get the following theorem.

Theorem 4.1.6 *There exits a constant C independent of interface and mesh such that*

$$|I_{h,T}u - u|_{2,\tilde{T}^-\cap T^-} \le C\left(|u|_{1,\tilde{T}^-\cap T^-} + |u|_{2,\tilde{T}^-\cap T^-}\right) \le C\,\|u\|_{2,T} \qquad (4.31)$$

for any $u \in PH^2_{int}(T)$, where T is a Type I interface element.

Proof. Since $\frac{\partial^2(I_{h,T}u)}{\partial x^2} = \frac{\partial^2(I_{h,T}u)}{\partial y^2} = 0$, then we complete the proof by applying the same techniques of Theorem 4.1.2 to Theorem 4.1.5. Note that (3.13) is used here.

■

Now we discuss the estimation on $\tilde{T}^+\cap T^+$ for Type I interface elements. The estimate on $\tilde{T}^+\cap T^+$ is rather similar to that on $\tilde{T}^-\cap T^-$, so we only state the results as follows. Let $X = (x,y)^T$ be a point in $\tilde{T}^+\cap T^+$. Without loss of generality, we can assume that line segments $\overline{XA_1}$ does not intersect with the interface and $\overline{DE}$, while line segment $\overline{XA_i}, i = 2,3,4$ meet Γ at $\widetilde{A}_i$, $i = 2,3,4$, see Figure 4.2. Also, we assume that $A_i = (x_i, y_i)^T$, $i = 1,2,3,4$ and

$$\widetilde{A}_i = \tilde{t}_i A_i + (1-\tilde{t}_i)X = (\widetilde{x}_i, \widetilde{y}_i)^T, i = 2,3,4.$$

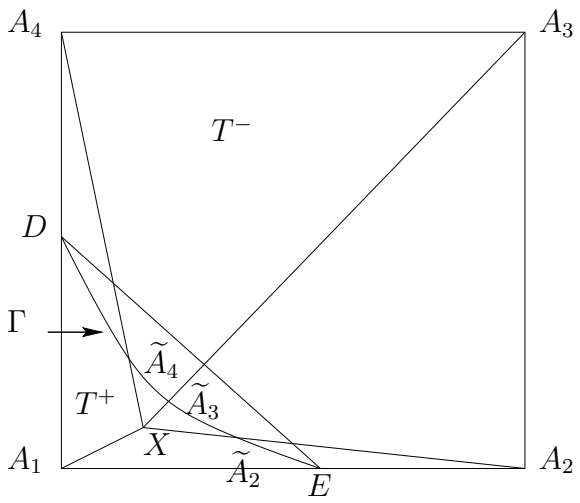

Figure 4.2: A point X $\in \tilde{T}^+ \cap T^+$ is connected to the four vertices by line segments in a Type I interface element

Lemma 4.1.8 *Given a two-dimensional vector* $\mathbf{q}$*, a point* $X \in \tilde{T}^+$ *and two real numbers* r, d^+*, then there exists a* $v \in S_h(T)$ *such that* $\nabla v(X) = \mathbf{q}, v(X) = r, \dfrac{\partial^2 v^+(X)}{\partial x \partial y} = d^+$ *and*

$$
\begin{aligned}
& \mathbf{q} \cdot \sum_{i=1}^{4} \Big[(A_i - X)\phi_i(X) \Big] \\
= \; & \sum_{i=2}^{4} \Big[-(N^+_{\overline{DE}} - I)\mathbf{q} \cdot (A_i - \widetilde{A}_i)\phi_i(X) - (N^+_{\overline{DE}} - I)\mathbf{q} \cdot (\widetilde{A}_i - X^{(i)}_{\overline{DE}})\phi_i(X) \Big] \\
& -d^+ \Big[\sum_{i=2}^{4} \Big[N^+_{\overline{DE}}(X^{(i)s}_{\overline{DE}} - X^s) \cdot (A_i - X^{(i)}_{\overline{DE}})\phi_i(X) + (x_i - \bar{x}_i)(y_i - \bar{y}_i)\phi_i(X) + \\
& (\bar{x}_i - x)(\bar{y}_i - y)\phi_i(X) \Big] + (x_1 - x)(y_1 - y)\phi_1(X) \Big],
\end{aligned}
$$

where $X^{(i)}_{\overline{DE}} = (\bar{x}_i, \bar{y}_i)^T$*,* $i = 2, 3, 4$ *are arbitrary points on* $\overline{DE}$*.*

Theorem 4.1.7 *For any $u \in PC^2_{int}(T)$, $X = (x, y)^T \in \tilde{T}^+ \cap T^+$, we have*

$$
\begin{aligned}
& I_{h,T}u(X) - u(X) \\
= \; & \sum_{i=2}^{4} \Big[(N^+(\widetilde{A}_i) - N^+_{\overline{DE}}) \nabla u(X) \cdot (A_i - \widetilde{A}_i)\phi_i(X) - (N^+_{\overline{DE}} - I)\nabla u(X) \cdot (\widetilde{A}_i - X^{(i)}_{\overline{DE}})\phi_i(X) \Big] \\
& - \frac{\partial^2 u(X)}{\partial x \partial y} \Big[\sum_{i=2}^{4} \Big[N^+_{\overline{DE}}(X^{(i)s}_{\overline{DE}} - X^s) \cdot (A_i - X^{(i)}_{\overline{DE}})\phi_i(X) + (x_i - \bar{x}_i)(y_i - \bar{y}_i)\phi_i(X) \\
& + (\bar{x}_i - x)(\bar{y}_i - y)\phi_i(X) \Big] + (x_1 - x)(y_1 - y)\phi_1(X) \Big] \\
& + \sum_{i=2}^{4} \Big[(N^+(\widetilde{A}_i) - I) \int_0^1 \frac{d[\nabla u(t\widetilde{A}_i + (1-t)X)]}{dt} \cdot (A_i - \widetilde{A}_i) \; dt \; \phi_i(X) \\
& + \int_0^{\tilde{t}_i} (1-t) \frac{d^2 u(tA_i + (1-t)X)}{dt^2} \; dt \; \phi_i(X) + \int_{\tilde{t}_i}^1 (1-t) \frac{d^2 u(tA_i + (1-t)X)}{dt^2} \; dt \; \phi_i(X) \Big] \\
& + \int_0^1 (1-t) \frac{d^2 u(tA_1 + (1-t)X)}{dt^2} \; dt \; \phi_1(X), \qquad (4.32)
\end{aligned}
$$

where $X^{(i)}_{\overline{DE}} = (\bar{x}_i, \bar{y}_i)^T$, $i = 2, 3, 4$ are arbitrary points on $\overline{DE}$.

■

We now let I_i, $i = 1, 2, 3, 4$ be the integral terms involving the vertices A_i, $i = 1, 2, 3, 4$ in (4.32).

Theorem 4.1.8 *For any $u \in PC^3_{int}(T)$, $X = (x, y)^T \in \tilde{T}^+ \cap T^+$, we have*

$$
\begin{aligned}
& \frac{\partial (I_{h,T}u(X) - u(X))}{\partial x} \\
= \; & \sum_{i=2}^{4} \Big[(N^+(\widetilde{A}_i) - N^+_{\overline{DE}}) \nabla u(X) \cdot (A_i - \widetilde{A}_i) \frac{\partial \phi_i(X)}{\partial x} \\
& - (N^+_{\overline{DE}} - I)\nabla u(X) \cdot (\widetilde{A}_i - X^{(i)}_{\overline{DE}}) \frac{\partial \phi_i(X)}{\partial x} \Big] \\
& - \frac{\partial^2 u(X)}{\partial x \partial y} \Big[\sum_{i=2}^{4} \Big[N^+_{\overline{DE}}(X^{(i)s}_{\overline{DE}} - X^s) \cdot (A_i - X^{(i)}_{\overline{DE}}) \frac{\partial \phi_i(X)}{\partial x} \\
& + N^+_{\overline{DE}}(0, -1)^T \cdot (A_i - X^{(i)}_{\overline{DE}})\phi_i(X) + (x_i - \bar{x}_i)(y_i - \bar{y}_i) \frac{\partial \phi_i(X)}{\partial x} - (\bar{y}_i - y)\phi_i(X) \\
& + (\bar{x}_i - x)(\bar{y}_i - y) \frac{\partial \phi_i(X)}{\partial x} \Big] - (y_1 - y)\phi_1(X) + (x_1 - x)(y_1 - y) \frac{\partial \phi_1(X)}{\partial x} \Big] \\
& + \sum_{i=1}^{4} I_i \frac{\partial \phi_i(X)}{\partial x},
\end{aligned}
$$

$$
\begin{aligned}
&\frac{\partial(I_{h,T}u(X)-u(X))}{\partial y}\\
=\ &\sum_{i=2}^{4}\Big[(N^+(\widetilde{A}_i)-N^+_{\overline{DE}})\nabla u(X)\cdot(A_i-\widetilde{A}_i)\frac{\partial\phi_i(X)}{\partial y}\\
&-(N^+_{\overline{DE}}-I)\nabla u(X)\cdot(\widetilde{A}_i-X^{(i)}_{\overline{DE}})\frac{\partial\phi_i(X)}{\partial y}\Big]\\
&-\frac{\partial^2u(X)}{\partial x\partial y}\Big[\sum_{i=2}^{4}\Big[N^+_{\overline{DE}}(X^{(i)s}_{\overline{DE}}-X^s)\cdot(A_i-X^{(i)}_{\overline{DE}})\frac{\partial\phi_i(X)}{\partial y}\\
&+N^+_{\overline{DE}}(-1,0)^T\cdot(A_i-X^{(i)}_{\overline{DE}})\phi_i(X)+(x_i-\bar{x}_i)(y_i-\bar{y}_i)\frac{\partial\phi_i(X)}{\partial y}-(\bar{x}_i-x)\phi_i(X)\\
&+(\bar{x}_i-x)(\bar{y}_i-y)\frac{\partial\phi_i(X)}{\partial y}\Big]-(x_1-x)\phi_1(X)+(x_1-x)(y_1-y)\frac{\partial\phi_1(X)}{\partial y}\Big]\\
&+\sum_{i=1}^{4}I_i\frac{\partial\phi_i(X)}{\partial y},
\end{aligned}
$$

where $X^{(i)}_{\overline{DE}}=(\bar{x}_i,\bar{y}_i)^T$, $i=2,3,4$ *are arbitrary points on* $\overline{DE}$.

■

Theorem 4.1.9 *For any* $u\in PC^3_{int}(T)$ *and* $X=(x,y)^T\in\tilde{T}^+\cap T^+$, *we have*

$$
\begin{aligned}
&\frac{\partial^2(I_{h,T}u(X)-u(X))}{\partial x\partial y}\\
=\ &\sum_{i=2}^{4}\Big[(N^+(\widetilde{A}_i)-N^+_{\overline{DE}})\nabla u(X)\cdot(A_i-\widetilde{A}_i)\frac{\partial^2\phi_i(X)}{\partial x\partial y}\\
&-(N^+_{\overline{DE}}-I)\nabla u(X)\cdot(\widetilde{A}_i-X^{(i)}_{\overline{DE}})\frac{\partial^2\phi_i(X)}{\partial x\partial y}\Big]\\
&-\frac{\partial^2u(X)}{\partial x\partial y}\Big[\sum_{i=2}^{4}\Big[N^+_{\overline{DE}}(X^{(i)s}_{\overline{DE}}-X^s)\cdot(A_i-X^{(i)}_{\overline{DE}})\frac{\partial^2\phi_i(X)}{\partial x\partial y}\\
&+N^+_{\overline{DE}}(-1,0)^T\cdot(A_i-X^{(i)}_{\overline{DE}})\frac{\partial\phi_i(X)}{\partial x}+N^+_{\overline{DE}}(0,-1)^T\cdot(A_i-X^{(i)}_{\overline{DE}})\frac{\partial\phi_i(X)}{\partial y}\\
&+(x_i-\bar{x}_i)(y_i-\bar{y}_i)\frac{\partial^2\phi_i(X)}{\partial x\partial y}+\phi_i(X)-(\bar{y}_i-y)\frac{\partial\phi_i(X)}{\partial y}-(\bar{x}_i-x)\frac{\partial\phi_i(X)}{\partial x}\\
&+(\bar{x}_i-x)(\bar{y}_i-y)\frac{\partial^2\phi_i(X)}{\partial x\partial y}\Big]+\phi_1(X)-(y_1-y)\frac{\partial\phi_1(X)}{\partial y}-(x_1-x)\frac{\partial\phi_1(X)}{\partial x}\\
&+(x_1-x)(y_1-y)\frac{\partial^2\phi_1(X)}{\partial x\partial y}\Big]+\sum_{i=1}^{4}I_i\frac{\partial^2\phi_i(X)}{\partial x\partial y},
\end{aligned}
$$

where $X_{\overline{DE}}^{(i)} = (\bar{x}_i, \bar{y}_i)^T$, $i = 2, 3, 4$ *are arbitrary points on* $\overline{DE}$.

■

Theorem 4.1.10 *There exits a constant* C *independent of interface and mesh such that*

$$\begin{aligned}
&\|I_{h,T}u - u\|_{0,\tilde{T}^+\cap T^+} \le Ch^2 \left(|u|_{1,\tilde{T}^+\cap T^+} + |u|_{2,\tilde{T}^+\cap T^+}\right) \le Ch^2 \|u\|_{2,T}, \\
&|I_{h,T}u - u|_{1,\tilde{T}^+\cap T^+} \le Ch \left(|u|_{1,\tilde{T}^+\cap T^+} + |u|_{2,\tilde{T}^+\cap T^+}\right) \le Ch \|u\|_{2,T}, \\
&|I_{h,T}u - u|_{2,\tilde{T}^+\cap T^+} \le C \left(|u|_{1,\tilde{T}^+\cap T^+} + |u|_{2,\tilde{T}^+\cap T^+}\right) \le C \|u\|_{2,T},
\end{aligned}$$

for any $u \in PH^2_{int}(T)$, *where* T *is a Type I interface element.*

■

If $\tilde{T}^+ \cap T^-$ is not empty, then we can use (4.17) and Lemma 4.1.8 to obtain a expansion for the interpolation error, which is similar to (4.14) and (4.32). Then we can follow the same arguments for $\tilde{T}^- \cap T^-$ to obtain the following theorem.

Theorem 4.1.11 *There exits a constant* C *independent of interface and mesh such that*

$$\begin{aligned}
&\|I_{h,T}u - u\|_{0,\tilde{T}^+\cap T^-} \le Ch^2 \left(|u|_{1,\tilde{T}^+\cap T^-} + |u|_{2,\tilde{T}^+\cap T^-}\right) \le Ch^2 \|u\|_{2,T}, \\
&|I_{h,T}u - u|_{1,\tilde{T}^+\cap T^-} \le Ch \left(|u|_{1,\tilde{T}^+\cap T^-} + |u|_{2,\tilde{T}^+\cap T^-}\right) \le Ch \|u\|_{2,T}, \\
&|I_{h,T}u - u|_{2,\tilde{T}^+\cap T^-} \le C \left(|u|_{1,\tilde{T}^+\cap T^-} + |u|_{2,\tilde{T}^+\cap T^-}\right) \le C \|u\|_{2,T},
\end{aligned}$$

for any $u \in PH^2_{int}(T)$, *where* T *is a Type I interface element.*

Similarly, if $\tilde{T}^- \cap T^+$ is not empty, we can also obtain the following theorem.

Theorem 4.1.12 *There exits a constant* C *independent of interface and mesh such that*

$$\begin{aligned}
&\|I_{h,T}u - u\|_{0,\tilde{T}^-\cap T^+} \le Ch^2 \left(|u|_{1,\tilde{T}^-\cap T^+} + |u|_{2,\tilde{T}^-\cap T^+}\right) \le Ch^2 \|u\|_{2,T}, \\
&|I_{h,T}u - u|_{1,\tilde{T}^-\cap T^+} \le Ch \left(|u|_{1,\tilde{T}^-\cap T^+} + |u|_{2,\tilde{T}^-\cap T^+}\right) \le Ch \|u\|_{2,T}, \\
&|I_{h,T}u - u|_{2,\tilde{T}^-\cap T^+} \le C \left(|u|_{1,\tilde{T}^-\cap T^+} + |u|_{2,\tilde{T}^-\cap T^+}\right) \le C \|u\|_{2,T},
\end{aligned}$$

for any $u \in PH^2_{int}(T)$, *where* T *is a Type I interface element.*

Finally, combining Theorems 4.1.2, 4.1.4, 4.1.6, 4.1.10, 4.1.11 and 4.1.12, we have the following theorem for the interpolation error on each Type I interface element.

Theorem 4.1.13 *There exits a constant C independent of interface and mesh such that*

$$\begin{aligned}
&\|I_{h,T}u - u\|_{0,T} \le Ch^2 \left(|u|_{1,T} + |u|_{2,T}\right) \le Ch^2 \|u\|_{2,T}, \\
&|I_{h,T}u - u|_{1,T} \le Ch \left(|u|_{1,T} + |u|_{2,T}\right) \le Ch \|u\|_{2,T}, \\
&|I_{h,T}u - u|_{2,T} \le C \left(|u|_{1,T} + |u|_{2,T}\right) \le C \|u\|_{2,T},
\end{aligned}$$

for any $u \in PH^2_{int}(T)$, where T is a Type I interface element.

4.1.3 Interpolation error on a Type II interface element

In this section, we will discuss the bilinear IFE interpolation error on a Type II interface element. The estimate for Type II interface elements is similar to that for Type I interface elements. Without loss of generality, we assume $T \in \mathcal{T}_h$ is a Type II interface element with vertices $A_i = (x_i, y_i)$, $i = 1, 2, 3, 4$, such that $A_1, A_2 \in T^+$ and $A_3, A_4 \in T^-$, see Figure 4.3.

We start with the estimation on $\tilde{T}^- \cap T^-$. Let $X = (x, y)^T$ be a point in $\tilde{T}^- \cap T^-$. Without loss of generality, we can assume that line segments $\overline{XA_i}, i = 3, 4$ do not intersect with the interface Γ and $\overline{DE}$, while line segment $\overline{XA_i}, i = 1, 2$ meet Γ at $\widetilde{A}_i, i = 1, 2$, see Figure 4.3. Also, we assume that

$$\widetilde{A}_i = \tilde{t}_i A_i + (1 - \tilde{t}_i)X = (\widetilde{x}_i, \widetilde{y}_i)^T, i = 1, 2. \tag{4.33}$$

With arguments similar to those used for Lemma 4.1.5, we have the following lemma. The only difference is that T is a Type II interface element here.

Lemma 4.1.9 *Assume that $v \in S_h(T), X = (x, y)^T \in \tilde{T}^-, A = (x_A, y_A)^T \in \tilde{T}^+$. Then*

$$\begin{aligned}
v(A) \;=\; & v(X) + \nabla v(X) \cdot (A - X) + (N^-_{\overline{DE}} - I)\nabla v(X) \cdot (A - \widetilde{A}) \\
& + (N^-_{\overline{DE}} - I)\nabla v(X) \cdot (\widetilde{A} - X_{\overline{DE}}) + d^- N^-_{\overline{DE}}(X^s_{\overline{DE}} - X^s) \cdot (A - X_{\overline{DE}}) \\
& + d^-(x_A - \bar{x})(y_A - \bar{y}) + d^-(\bar{x} - x)(\bar{y} - y),
\end{aligned}$$

where $X_{\overline{DE}} = (\bar{x}, \bar{y})^T$ is an arbitrary point on $\overline{DE}$.

The following lemma is similar to Lemma 4.1.6.

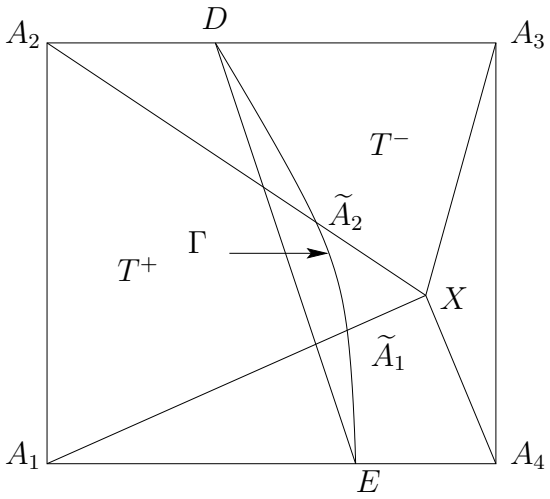

Figure 4.3: A point X $\in \tilde{T}^- \cap T^-$ is connected to the four vertices by line segments in a Type II interface element

Lemma 4.1.10 *For any $v \in S_h(T)$ and $X = (x, y)^T \in \tilde{T}^-$,*

$$\begin{aligned}
&\nabla v(X) \cdot \sum_{i=1}^{4}(A_i - X)\phi_i(X) \\
= &\sum_{i=1}^{2}\Big[-(N^-_{\overline{DE}} - I)\nabla v(X)\cdot(A_i - \widetilde{A}_i)\phi_i(X) - (N^-_{\overline{DE}} - I)\nabla v(X)\cdot(\widetilde{A}_i - X^{(i)}_{\overline{DE}})\phi_i(X)\Big] \\
&-d^-\Big[\sum_{i=1}^{2}\Big[N^-_{\overline{DE}}(X^{(i)s}_{\overline{DE}} - X^s)\cdot(A_i - X^{(i)}_{\overline{DE}})\phi_i(X) + (x_i - \bar{x}_i)(y_i - \bar{y}_i)\phi_i(X) \\
&+(\bar{x}_i - x)(\bar{y}_i - y)\phi_i(X)\Big] + \sum_{i=3}^{4}\Big[(x_i - x)(y_i - y)\phi_i(X)\Big]\Big],
\end{aligned}$$

where $X^{(i)}_{\overline{DE}} = (\bar{x}_i, \bar{y}_i)^T$, $i = 1, 2$ are arbitrary points on $\overline{DE}$.

Proof. By using Lemma 4.1.4 and Lemma 4.1.9,we can get

$$\begin{aligned}
v(A_i) &= v(X) + \nabla v(X)\cdot(A_i - X) + d^-(x_i - x)(y_i - y), \quad i = 3, 4, \\
v(A_i) &= v(X) + \nabla v(X)\cdot(A_i - X) + (N^-_{\overline{DE}} - I)\nabla v(X)\cdot(A_i - \widetilde{A}_i) + \\
&\quad (N^-_{\overline{DE}} - I)\nabla v(X)\cdot(\widetilde{A}_i - X^{(i)}_{\overline{DE}}) + d^- N^-_{\overline{DE}}(X^{(i)s}_{\overline{DE}} - X^s)\cdot(A_i - X^{(i)}_{\overline{DE}}) \\
&\quad +d^-(x_i - \bar{x}_i)(y_i - \bar{y}_i) + d^-(\bar{x}_i - x)(\bar{y}_i - y), \quad i = 1, 2.
\end{aligned}$$

Because $v \in S_h(T)$,

$$\begin{aligned} v(X) &= I_{h,T}v(X) \\ &= \sum_{i=1}^{4} v(A_i)\phi_i(X) \\ &= v(X)\sum_{i=1}^{4}\phi_i(X) + \nabla v(X)\cdot\sum_{i=1}^{4}(A_i - X)\phi_i(X) \\ &\quad + \sum_{i=1}^{2}\Big[(N_{\overline{DE}}^{-} - I)\nabla v(X)\cdot(A_i - \widetilde{A}_i)\phi_i(X) + (N_{\overline{DE}}^{-} - I)\nabla v(X)\cdot(\widetilde{A}_i - X_{\overline{DE}}^{(i)})\phi_i(X)\Big] \\ &\quad + d^{-}\Big[\sum_{i=1}^{2}\Big[N_{\overline{DE}}^{-}(X_{\overline{DE}}^{(i)s} - X^s)\cdot(A_i - X_{\overline{DE}}^{(i)})\phi_i(X) + (x_i - \bar{x}_i)(y_i - \bar{y}_i)\phi_i(X) \\ &\quad + (\bar{x}_i - x)(\bar{y}_i - y)\phi_i(X)\Big] + \sum_{i=3}^{4}(x_i - x)(y_i - y)\phi_i(X)\Big]. \end{aligned}$$

Because of Theorem 3.4.1,the proof is completed.

■

By arguments similar to those used for Lemma 4.1.7, we get the following lemma.

Lemma 4.1.11 *Given a two-dimensional vector* $\mathbf{q}$*, a point* $X \in \tilde{T}^{-}$ *and two real numbers* r, d^{-}*, then there exists a function* $v \in S_h(T)$ *such that* $\nabla v(X) = \mathbf{q}, v(X) = r, \dfrac{\partial^2 v^{-}(X)}{\partial x \partial y} = d^{-}$ *and*

$$\begin{aligned} &\mathbf{q}(X)\cdot\sum_{i=1}^{4}(A_i - X)\phi_i(X) \\ =\ &\sum_{i=1}^{2}\Big[-(N_{\overline{DE}}^{-} - I)\mathbf{q}\cdot(A_i - \widetilde{A}_i)\phi_i(X) - (N_{\overline{DE}}^{-} - I)\mathbf{q}\cdot(\widetilde{A}_i - X_{\overline{DE}}^{(i)})\phi_i(X)\Big] \\ &- d^{-}\Big[\sum_{i=1}^{2}\Big[N_{\overline{DE}}^{-}(X_{\overline{DE}}^{(i)s} - X^s)\cdot(A_i - X_{\overline{DE}}^{(i)})\phi_i(X) + (x_i - \bar{x}_i)(y_i - \bar{y}_i)\phi_i(X) \\ &+ (\bar{x}_i - x)(\bar{y}_i - y)\phi_i(X)\Big] + \sum_{i=3}^{4}\Big[(x_i - x)(y_i - y)\phi_i(X)\Big]\Big], \end{aligned} \tag{4.34}$$

where $X_{\overline{DE}}^{(i)} = (\bar{x}_i, \bar{y}_i)^T$*,* $i = 1, 2$ *are arbitrary points on* $\overline{DE}$*.*

■

We can now derive an expansion of the bilinear IFE interpolation error at any point $X \in T^-$ of a Type II interface element T. The proof here is similar to the proof of (4.14).

Theorem 4.1.14 *For any $u \in PC^2_{int}(T)$ and $X = (x, y)^T \in \tilde{T}^- \cap T^-$, we have*

$$\begin{aligned}
& I_{h,T}u(X) - u(X) \\
= & \sum_{i=1}^{2} \Big[(N^-(\widetilde{A}_i) - N^-_{\overline{DE}}) \nabla u(X) \cdot (A_i - \widetilde{A}_i)\phi_i(X) - (N^-_{\overline{DE}} - I)\nabla u(X) \cdot (\widetilde{A}_i - X^{(i)}_{\overline{DE}})\phi_i(X) \Big] \\
& - \frac{\partial^2 u(X)}{\partial x \partial y} \Big[\sum_{i=1}^{2} \Big[N^-_{\overline{DE}}(X^{(i)s}_{\overline{DE}} - X^s) \cdot (A_i - X^{(i)}_{\overline{DE}})\phi_i(X) + (x_i - \bar{x}_i)(y_i - \bar{y}_i)\phi_i(X) \\
& + (\bar{x}_i - x)(\bar{y}_i - y)\phi_i(X) \Big] + \sum_{i=3}^{4} (x_i - x)(y_i - y)\phi_i(X) \Big] \\
& + \sum_{i=1}^{2} \Big[(N^-(\widetilde{A}_i) - I) \int_0^1 \frac{d[\nabla u^-(t\widetilde{A}_i + (1-t)X)]}{dt} \cdot (A_i - \widetilde{A}_i)\ dt\ \phi_i(X) \\
& + \int_0^{\tilde{t}_i} (1-t) \frac{d^2 u(tA_i + (1-t)X)}{dt^2}\ dt\ \phi_i(X) + \int_{\tilde{t}_i}^1 (1-t) \frac{d^2 u(tA_i + (1-t)X)}{dt^2}\ dt\ \phi_i(X) \Big] \\
& + \sum_{i=3}^{4} \int_0^1 (1-t) \frac{d^2 u(tA_i + (1-t)X)}{dt^2}\ dt\ \phi_i(X), \qquad (4.35)
\end{aligned}$$

where $X^{(i)}_{\overline{DE}} = (\bar{x}_i, \bar{y}_i)^T$, $i = 1, 2$ are arbitrary points on $\overline{DE}$.

Proof. Since $t \mapsto u(tA_i + (1-t)X))(i = 3, 4)$ is a C^2 function,for i=3,4, we have

$$\begin{aligned}
u(A_i) &= u(X) + \int_0^1 \frac{d\ u(tA_i + (1-t)X)}{dt}\ dt \\
&= u(X) - \int_0^1 \frac{d\ u(tA_i + (1-t)X)}{dt}\ d(1-t) \\
&= u(X) - \frac{d\ u(tA_i + (1-t)X)}{dt}(1-t)\Big|_{t=0}^{t=1} + \int_0^1 (1-t) \frac{d^2 u(tA_i + (1-t)X)}{dt^2}\ dt \\
&= u(X) + \nabla u(X) \cdot (A_i - X) + \int_0^1 (1-t) \frac{d^2 u(tA_i + (1-t)X)}{dt^2}\ dt. \qquad (4.36)
\end{aligned}$$

For $i = 1, 2$, using integration by parts, (4.33) and (4.1), we have

$$
\begin{aligned}
u(A_i) &= u(X) + \int_0^{\tilde{t}_i} \frac{d\ u(tA_i + (1-t)X)}{dt}\ dt\ + \int_{\tilde{t}_i}^1 \frac{d\ u(tA_i + (1-t)X)}{dt}\ dt \\
&= u(X) - \int_0^{\tilde{t}_i} \frac{d\ u(tA_i + (1-t)X)}{dt}\ d(1-t)\ - \int_{\tilde{t}_i}^1 \frac{d\ u(tA_i + (1-t)X)}{dt}\ d(1-t) \\
&= u(X) - \frac{d\ u(tA_i + (1-t)X)}{dt}(1-t)\Big|_{t=0}^{t=\tilde{t}_i} + \int_0^{\tilde{t}} (1-t)\frac{d^2u(tA_i + (1-t)X)}{dt^2}\ dt \\
&\quad -\frac{d\ u(tA_i + (1-t)X)}{dt}(1-t)\Big|_{t=\tilde{t}_i}^{t=1} + \int_{\tilde{t}}^1 (1-t)\frac{d^2u(tA_i + (1-t)X)}{dt^2}\ dt \\
&= u(X) - (1-\tilde{t}_i)\nabla u^-(\widetilde{A}_i)\cdot(A_i - X) + \nabla u(X)\cdot(A_i - X) \\
&\quad + \int_0^{\tilde{t}} (1-t)\frac{d^2u(tA_i + (1-t)X)}{dt^2}\ dt + (1-\tilde{t}_i)\nabla u^+(\widetilde{A}_i)\cdot(A_i - X) \\
&\quad + \int_{\tilde{t}}^1 (1-t)\frac{d^2u(tA_i + (1-t)X)}{dt^2}\ dt \\
&= u(X) - \nabla u^-(\widetilde{A}_i)\cdot(A_i - \widetilde{A}_i) + \nabla u(X)\cdot(A_i - X) \\
&\quad + \int_0^{\tilde{t}} (1-t)\frac{d^2u(tA_i + (1-t)X)}{dt^2}\ dt + \nabla u^+(\widetilde{A}_i)\cdot(A_i - \widetilde{A}_i) \\
&\quad + \int_{\tilde{t}}^1 (1-t)\frac{d^2u(tA_i + (1-t)X)}{dt^2}\ dt \\
&= u(X) + \nabla u(X)\cdot(A_i - X) + (N^-(\widetilde{A}_i) - I)\nabla u^-(\widetilde{A}_i)\cdot(A_i - \widetilde{A}_i) \\
&\quad + \int_0^{\tilde{t}_i} (1-t)\frac{d^2u(tA_i + (1-t)X)}{dt^2}\ dt\ + \int_{\tilde{t}_i}^1 (1-t)\frac{d^2u(tA_i + (1-t)X)}{dt^2}\ dt \\
&= u(X) + \nabla u(X)\cdot(A_i - X) + (N^-(\widetilde{A}_i) - I)\nabla u(X)\cdot(A_i - \widetilde{A}_i) \\
&\quad + (N^-(\widetilde{A}_i) - I)\int_0^1 \frac{d[\nabla u(t\widetilde{A}_i + (1-t)X)]}{dt}\cdot(A_i - \widetilde{A}_i)\ dt \\
&\quad + \int_0^{\tilde{t}_i} (1-t)\frac{d^2u(tA_i + (1-t)X)}{dt^2}\ dt \\
&\quad + \int_{\tilde{t}_i}^1 (1-t)\frac{d^2u(tA_i + (1-t)X)}{dt^2}\ dt. \qquad (4.37)
\end{aligned}
$$

Then

$$\begin{aligned} I_{h,T}u(X) &= \sum_{i=1}^{4} u(A_i)\phi_i(X) \\ &= u(X)\sum_{i=1}^{4}\phi_i(X) + \nabla u(X)\cdot\sum_{i=1}^{4}(A_i - X)\phi_i(X) \\ &\quad + \sum_{i=1}^{2}\Big[(N^-(\widetilde{A}_i) - I)\nabla u(X)\cdot(A_i - \widetilde{A}_i)\phi_i(X) \\ &\quad + (N^-(\widetilde{A}_i) - I)\int_0^1 \frac{d[\nabla u(t\widetilde{A}_i + (1-t)X)]}{dt}\cdot(A_i - \widetilde{A}_i)\ dt\ \phi_i(X) \\ &\quad + \int_0^{\tilde{t}_i}(1-t)\frac{d^2u(tA_i + (1-t)X)}{dt^2}\ dt\ \phi_i(X) \\ &\quad + \int_{\tilde{t}_i}^1(1-t)\frac{d^2u(tA_i + (1-t)X)}{dt^2}\ dt\ \phi_i(X)\Big] \\ &\quad + \sum_{i=3}^{4}\int_0^1(1-t)\frac{d^2u(tA_i + (1-t)X)}{dt^2}\ dt\ \phi_i(X). \end{aligned} \tag{4.38}$$

Now let $\mathbf{q} = \nabla u(X), r = u(X), d^- = \dfrac{\partial^2 u(X)}{\partial x \partial y}$ in Lemma 4.1.11, we have

$$\begin{aligned} &\nabla u(X)\cdot\sum_{i=1}^{4}(A_i - X)\phi_i(X) \\ = &\sum_{i=1}^{2}\Big[-(N^-_{\overline{DE}} - I)\nabla u(X)\cdot(A_i - \widetilde{A}_i)\phi_i(X) - (N^-_{\overline{DE}} - I)\nabla u(X)(X)\cdot(\widetilde{A}_i - X^{(i)}_{\overline{DE}})\phi_i(X)\Big] \\ &-\frac{\partial^2 u(X)}{\partial x\partial y}\Big[\sum_{i=1}^{2}\Big[N^-_{\overline{DE}}(X^{(i)s}_{\overline{DE}} - X^s)\cdot(A_i - X^{(i)}_{\overline{DE}})\phi_i(X) + (x_i - \bar{x}_i)(y_i - \bar{y}_i)\phi_i(X) \\ &+(\bar{x}_i - x)(\bar{y}_i - y)\phi_i(X)\Big] + \sum_{i=3}^{4}(x_i - x)(y_i - y)\phi_i(X)\Big] \end{aligned} \tag{4.39}$$

Finally, (4.35) follows from (4.38), (4.39) and Theorem 3.4.1.

■

Let I_i, $i = 1, 2, 3, 4$ be the integral terms involving vertices $A_i, i = 1, 2, 3, 4$ in (4.35). Then we can prove the following two theorems similar to Theorem 4.1.3 and Theorem 4.1.5.

Theorem 4.1.15 *For any $u \in PC^3_{int}(T)$ and $X = (x, y)^T \in \tilde{T}^- \cap T^-$, we have*

$$\begin{aligned}
&\frac{\partial(I_{h,T}u(X) - u(X))}{\partial x} \\
= &\sum_{i=1}^{2}\Big[(N^-(\widetilde{A}_i) - N^-_{\overline{DE}})\nabla u(X)\cdot(A_i - \widetilde{A}_i)\frac{\partial\phi_i(X)}{\partial x} \\
&-(N^-_{\overline{DE}} - I)\nabla u(X)\cdot(\widetilde{A}_i - X^{(i)}_{\overline{DE}})\frac{\partial\phi_i(X)}{\partial x}\Big] \\
&-\frac{\partial^2 u(X)}{\partial x\partial y}\Big[\sum_{i=1}^{2}\Big[N^-_{\overline{DE}}(X^{(i)s}_{\overline{DE}} - X^s)\cdot(A_i - X^{(i)}_{\overline{DE}})\frac{\partial\phi_i(X)}{\partial x} \\
&+N^-_{\overline{DE}}(0,-1)^T\cdot(A_i - X^{(i)}_{\overline{DE}})\phi_i(X) + (x_i - \bar{x}_i)(y_i - \bar{y}_i)\frac{\partial\phi_i(X)}{\partial x} - (\bar{y}_i - y)\phi_i(X) \\
&+(\bar{x}_i - x)(\bar{y}_i - y)\frac{\partial\phi_i(X)}{\partial x}\Big] + \sum_{i=3}^{4}\Big[-(y_i - y)\phi_i(X) + (x_i - x)(y_i - y)\frac{\partial\phi_i(X)}{\partial x}\Big]\Big] \\
&+\sum_{i=1}^{4} I_i\frac{\partial\phi_i(X)}{\partial x}, \qquad (4.40)
\end{aligned}$$

$$\begin{aligned}
&\frac{\partial(I_{h,T}u(X) - u(X))}{\partial y} \\
= &\sum_{i=1}^{2}\Big[(N^-(\widetilde{A}_i) - N^-_{\overline{DE}})\nabla u(X)\cdot(A_i - \widetilde{A}_i)\frac{\partial\phi_i(X)}{\partial y} \\
&-(N^-_{\overline{DE}} - I)\nabla u(X)\cdot(\widetilde{A}_i - X^{(i)}_{\overline{DE}})\frac{\partial\phi_i(X)}{\partial y}\Big] \\
&-\frac{\partial^2 u(X)}{\partial x\partial y}\Big[\sum_{i=1}^{2}\Big[N^-_{\overline{DE}}(X^{(i)s}_{\overline{DE}} - X^s)\cdot(A_i - X^{(i)}_{\overline{DE}})\frac{\partial\phi_i(X)}{\partial y} \\
&+N^-_{\overline{DE}}(-1,0)^T\cdot(A_i - X^{(i)}_{\overline{DE}})\phi_i(X) + (x_i - \bar{x}_i)(y_i - \bar{y}_i)\frac{\partial\phi_i(X)}{\partial y} - (\bar{x}_i - x)\phi_i(X) \\
&+(\bar{x}_i - x)(\bar{y}_i - y)\frac{\partial\phi_i(X)}{\partial y}\Big] + \sum_{i=3}^{4}\Big[-(x_i - x)\phi_i(X) + (x_i - x)(y_i - y)\frac{\partial\phi_i(X)}{\partial y}\Big]\Big] \\
&+\sum_{i=1}^{4} I_i\frac{\partial\phi_i(X)}{\partial y}, \qquad (4.41)
\end{aligned}$$

where $X^{(i)}_{\overline{DE}} = (\bar{x}_i, \bar{y}_i)^T$, $i = 1, 2$ are arbitrary points on $\overline{DE}$.

Proof. We give a proof only for (4.40). (4.41) can be carried out similarly. Taking derivative

for x on both sides of (4.35), we can get

$$\begin{aligned}
&\frac{\partial(I_{h,T}u(X)-u(X))}{\partial x}\\
=\ &\sum_{i=1}^{2}\Big[\frac{\partial}{\partial x}\Big[(N^-(\widetilde{A}_i)-N^-_{\overline{DE}})\nabla u(X)\cdot(A_i-\widetilde{A}_i)\Big]\phi_i(X)+\\
&(N^-(\widetilde{A}_i)-N^-_{\overline{DE}})\nabla u(X)\cdot(A_i-\widetilde{A}_1)\frac{\partial\phi_i(X)}{\partial x}\\
&-\frac{\partial}{\partial x}\Big[(N^-_{\overline{DE}}-I)\nabla u(X)\cdot(\widetilde{A}_i-X^{(i)}_{\overline{DE}})\Big]\phi_i(X)-(N^-_{\overline{DE}}-I)\nabla u(X)\cdot(\widetilde{A}_i-X^{(i)}_{\overline{DE}})\frac{\partial\phi_i(X)}{\partial x}\Big]\\
&-\frac{\partial^3u(X)}{\partial x^2\partial y}\Big[\sum_{i=1}^{2}\Big[N^-_{\overline{DE}}(X^{(i)s}_{\overline{DE}}-X^s)\cdot(A_i-X^{(i)}_{\overline{DE}})\phi_i(X)+(x_i-\bar{x}_i)(y_i-\bar{y}_i)\phi_i(X)\\
&+(\bar{x}_i-x)(\bar{y}_i-y)\phi_i(X)+\sum_{i=3}^{4}(x_i-x)(y_i-y)\phi_i(X)\Big]-\\
&\frac{\partial^2u}{\partial x\partial y}\Big[\sum_{i=1}^{2}\Big[N^-_{\overline{DE}}(X^{(i)s}_{\overline{DE}}-X^s)\cdot(A_i-X^{(i)}_{\overline{DE}})\frac{\partial\phi_i(X)}{\partial x}+N^-_{\overline{DE}}(0,-1)^T\cdot(A_i-X^{(i)}_{\overline{DE}})\phi_i(X)\\
&+(x_i-\bar{x}_i)(y_i-\bar{y}_i)\frac{\partial\phi_i(X)}{\partial x}-(\bar{y}_i-y)\phi_i(X)+(\bar{x}_i-x)(\bar{y}_i-y)\frac{\partial\phi_i(X)}{\partial x}+\\
&\sum_{i=3}^{4}\Big[-(y_i-y)\phi_i(X)+(x_i-x)(y_i-y)\frac{\partial\phi_i(X)}{\partial x}\Big]\Big]\\
&+\sum_{i=1}^{4}I_i\frac{\partial\phi_i(X)}{\partial x}+\sum_{i=1}^{4}\frac{\partial I_i}{\partial x}\phi_i(X). \qquad (4.42)
\end{aligned}$$

Taking the derivative with respect to x on both sides of (4.36) and (4.37), we can get

$$\begin{aligned}
&\frac{\partial I_i}{\partial x}=-P\cdot(A_i-X),\quad i=3,4,\\
&\frac{\partial I_i}{\partial x}=-P\cdot(A_i-X)-\frac{\partial}{\partial x}\Big[(N^-(\widetilde{A}_i)-I)\nabla u(X)(A_i-\widetilde{A}_i)\Big],\quad i=1,2,
\end{aligned}$$

where

$$P=\frac{\partial}{\partial x}\nabla u(X)=(\frac{\partial^2u(X)}{\partial x^2},\frac{\partial^2u(X)}{\partial x\partial y})^T.$$

Hence

$$\begin{aligned}
&\sum_{i=1}^{4}\frac{\partial I_i}{\partial x}\phi_i(X)\\
=\ &-P\cdot\sum_{i=1}^{4}(A_i-X)\phi_i(X)-\sum_{i=1}^{2}\frac{\partial}{\partial x}\Big[(N^-(\widetilde{A}_i)-I)\nabla u(X)(A_i-\widetilde{A}_i)\Big]\phi_i(X).
\end{aligned}$$

Applying Lemma 4.1.11 to the first term on the right hand side above, letting $\mathbf{q} = P, d^- = \frac{\partial^3 u(X)}{\partial x^2 \partial y}$, we have

$$\begin{aligned}
&\sum_{i=1}^{4} \frac{\partial I_i}{\partial x}\phi_i(X) \\
= \ &\sum_{i=1}^{2} \Big[\frac{\partial}{\partial x}\Big[(N^-(\widetilde{A}_i) - I)\nabla u(X)\cdot(A_i - \widetilde{A}_i)\Big]\phi_i(X) + \\
&(N^-_{\overline{DE}} - I)P\cdot(A_i - \widetilde{A}_i)\phi_i(X) + (N^-_{\overline{DE}} - I)P\cdot(\widetilde{A}_i - X^{(i)}_{\overline{DE}})\phi_i(X)\Big] \\
&+\frac{\partial^3 u}{\partial x^2\partial y}\Big[\sum_{i=1}^{2}\Big[N^-_{\overline{DE}}(X^{(i)s}_{\overline{DE}} - X^s)\cdot(A_i - X^{(i)}_{\overline{DE}})\phi_i(X) + (x_i - \bar{x}_i)(y_i - \bar{y}_i)\phi_i(X) \\
&+(\bar{x}_i - x)(\bar{y}_i - y)\phi_i(X)\Big] + \sum_{i=3}^{4}\Big[(x_i - x)(y_i - y)\phi_i(X)\Big]\Big]. \qquad (4.43)
\end{aligned}$$

For $i = 1, 2$, by direct calculations, we also have

$$\begin{aligned}
&\frac{\partial}{\partial x}\Big[(N^-(\widetilde{A}_i) - N^-_{\overline{DE}})\nabla u(X)\cdot(A_i - \widetilde{A}_i)\Big]\phi_i(X) \\
&-\frac{\partial}{\partial x}\Big[(N^-_{\overline{DE}} - I)\nabla u(X)\cdot(\widetilde{A}_i - X^{(i)}_{\overline{DE}})\Big]\phi_i(X) \\
&-\frac{\partial}{\partial x}\Big[(N^-(\widetilde{A}_i) - I)\nabla u(X)\cdot(A_i - \widetilde{A}_i)\Big]\phi_i(X) \\
&+(N^-_{\overline{DE}} - I)P\cdot(A_i - \widetilde{A}_i)\phi_i(X) + (N^-_{\overline{DE}} - I)P\cdot(\widetilde{A}_i - X^{(i)}_{\overline{DE}})\phi_i(X) \\
= \ &-\frac{\partial}{\partial x}\Big[(N^-_{\overline{DE}} - I)\nabla u(X)\cdot(A_i - \widetilde{A}_i)\Big]\phi_i(X) \\
&-\frac{\partial}{\partial x}\Big[(N^-_{\overline{DE}} - I)\nabla u(X)\cdot(\widetilde{A}_i - X^{(i)}_{\overline{DE}})\Big]\phi_i(X) + (N^-_{\overline{DE}} - I)P\cdot(A_i - X^{(i)}_{\overline{DE}})\phi_i(X) \\
= \ &-\frac{\partial}{\partial x}\Big[(N^-_{\overline{DE}} - I)\nabla u(X)\cdot(A_i - X^{(i)}_{\overline{DE}})\Big]\phi_i(X) + (N^-_{\overline{DE}} - I)P\cdot(A_i - X^{(i)}_{\overline{DE}})\phi_i(X) \\
= \ &-(N^-_{\overline{DE}} - I)P\cdot(A_i - X_{\overline{DE}})\phi_i(X) + (N^-_{\overline{DE}} - I)P\cdot(A_i - X^{(i)}_{\overline{DE}})\phi_i(X) \\
= \ &0. \qquad (4.44)
\end{aligned}$$

Plugging (4.43) and (4.44) into (4.42),we finish the proof of (4.40).

■

Theorem 4.1.16 *For any $u \in PC^3_{int}(T)$ and $X = (x, y)^T \in \tilde{T}^- \cap T^-$, we have*

$$\begin{aligned}
&\frac{\partial^2(I_{h,T}u(X) - u(X))}{\partial x \partial y} \\
= \ & \sum_{i=1}^{2} \Big[(N^-(\widetilde{A}_i) - N^-_{\overline{DE}}) \nabla u(X) \cdot (A_i - \widetilde{A}_i) \frac{\partial^2 \phi_i(X)}{\partial x \partial y} \\
& -(N^-_{\overline{DE}} - I) \nabla u(X) \cdot (\widetilde{A}_i - X^{(i)}_{\overline{DE}}) \frac{\partial^2 \phi_i(X)}{\partial x \partial y} \Big] \\
& -\frac{\partial^2 u(X)}{\partial x \partial y} \Big[\sum_{i=1}^{2} N^-_{\overline{DE}} (X^{(i)s}_{\overline{DE}} - X^s) \cdot (A_i - X^{(i)}_{\overline{DE}}) \frac{\partial^2 \phi_i(X)}{\partial x \partial y} \\
& +N^-_{\overline{DE}}(-1, 0)^T \cdot (A_i - X^{(i)}_{\overline{DE}}) \frac{\partial \phi_i(X)}{\partial x} + N^-_{\overline{DE}}(0, -1)^T \cdot (A_i - X^{(i)}_{\overline{DE}}) \frac{\partial \phi_i(X)}{\partial y} \\
& +(x_i - \bar{x}_i)(y_i - \bar{y}_i) \frac{\partial^2 \phi_i(X)}{\partial x \partial y} + \phi_i(X) - (\bar{y}_i - y) \frac{\partial \phi_i(X)}{\partial y} - (\bar{x}_i - x) \frac{\partial \phi_i(X)}{\partial x} \\
& +(\bar{x}_i - x)(\bar{y}_i - y) \frac{\partial^2 \phi_i(X)}{\partial x \partial y} \Big] + \sum_{i=3}^{4} \Big[\phi_i(X) - (y_i - y) \frac{\partial \phi_i(X)}{\partial y} - (x_i - x) \frac{\partial \phi_i(X)}{\partial x} \\
& +(x_i - x)(y_i - y) \frac{\partial^2 \phi_i(X)}{\partial x \partial y} \Big] \Big] + \sum_{i=1}^{4} I_i \frac{\partial^2 \phi_i(X)}{\partial x \partial y}, \qquad (4.45)
\end{aligned}$$

where $X^{(i)}_{\overline{DE}} = (\bar{x}_i, \bar{y}_i)^T$, $i = 1, 2$ are arbitrary points on $\overline{DE}$.

Proof. Taking derivative for y on both sides of (4.40), we can get

$$
\begin{aligned}
&\frac{\partial^2(I_{h,T}u(X)-u(X))}{\partial x\partial y}\\
=\ &\sum_{i=1}^{2}\Big\{\frac{\partial}{\partial y}\Big[(N^-(\widetilde{A}_i)-N^-_{\overline{DE}})\nabla u(X)\cdot(A_i-\widetilde{A}_i)\Big]\frac{\partial\phi_i(X)}{\partial x}\\
&+(N^-(\widetilde{A}_i)-N^-_{\overline{DE}})\nabla u(X)\cdot(A_i-\widetilde{A}_i)\frac{\partial^2\phi_i(X)}{\partial x\partial y}\\
&-\frac{\partial}{\partial y}\Big[(N^-_{\overline{DE}}-I)\nabla u(X)\cdot(\widetilde{A}_i-X^{(i)}_{\overline{DE}})\Big]\frac{\partial\phi_i(X)}{\partial x}\\
&-(N^-_{\overline{DE}}-I)\nabla u(X)\cdot(\widetilde{A}_i-X^{(i)}_{\overline{DE}})\frac{\partial^2\phi_i(X)}{\partial x\partial y}\Big\}\\
&-\frac{\partial^3u(X)}{\partial x\partial y^2}\Big\{\sum_{i=1}^{2}\Big[N^-_{\overline{DE}}(X^{(i)s}_{\overline{DE}}-X^s)\cdot(A_i-X^{(i)}_{\overline{DE}})\frac{\partial\phi_i(X)}{\partial x}\\
&+N^-_{\overline{DE}}(0,-1)^T\cdot(A_i-X^{(i)}_{\overline{DE}})\phi_i(X)+(x_i-\bar{x}_i)(y_i-\bar{y}_i)\frac{\partial\phi_i(X)}{\partial x}\\
&-(\bar{y}_i-y)\phi_i(X)+(\bar{x}_i-x)(\bar{y}_i-y)\frac{\partial\phi_i(X)}{\partial x}\Big]\\
&+\sum_{i=3}^{4}\Big[-(y_i-y)\phi_i(X)+(x_i-x)(y_i-y)\frac{\partial\phi_i(X)}{\partial x}\Big]\Big\}\\
&-\frac{\partial^2u(X)}{\partial x\partial y}\Big\{\sum_{i=1}^{2}\Big[N^-_{\overline{DE}}(X^{(i)s}_{\overline{DE}}-X^s)\cdot(A_i-X^{(i)}_{\overline{DE}})\frac{\partial^2\phi_i(X)}{\partial x\partial y}\\
&+N^-_{\overline{DE}}(-1,0)^T\cdot(A_i-X^{(i)}_{\overline{DE}})\frac{\partial\phi_i(X)}{\partial x}+N^-_{\overline{DE}}(0,-1)^T\cdot(A_i-X^{(i)}_{\overline{DE}})\frac{\partial\phi_i(X)}{\partial y}\\
&+(x_i-\bar{x}_i)(y_i-\bar{y}_i)\frac{\partial^2\phi_i(X)}{\partial x\partial y}+\phi_i(X)-(\bar{y}_i-y)\frac{\partial\phi_i(X)}{\partial y}\\
&-(\bar{x}_i-x)\frac{\partial\phi_i(X)}{\partial x}+(\bar{x}_i-x)(\bar{y}_i-y)\frac{\partial^2\phi_i(X)}{\partial x\partial y}\Big]\\
&+\sum_{i=3}^{4}\Big[\phi_i(X)-(y_i-y)\frac{\partial\phi_i(X)}{\partial y}-(x_i-x)\frac{\partial\phi_i(X)}{\partial x}+(x_i-x)(y_i-y)\frac{\partial^2\phi_i(X)}{\partial x\partial y}\Big]\Big\}\\
&+\sum_{i=1}^{4}\frac{\partial I_i}{\partial y}\frac{\partial\phi_i(X)}{\partial x}+\sum_{i=1}^{4}I_i\frac{\partial^2\phi_i(X)}{\partial x\partial y}. \qquad (4.46)
\end{aligned}
$$

Taking the derivative with respect to y on both sides of (4.36) and (4.37), then we can get

$$\begin{aligned}
&\frac{\partial I_i}{\partial y} = -P\cdot(A_i - X), \quad i = 3,4,\\
&\frac{\partial I_i}{\partial y} = -P\cdot(A_i - X) - \frac{\partial}{\partial x}\Big[(N^-(\widetilde{A}_i) - I)\nabla u(X)(A_i - \widetilde{A}_i)\Big], \quad i = 1,2,
\end{aligned}$$

where

$$P = \frac{\partial}{\partial y}\nabla u(X) = (\frac{\partial^2 u(X)}{\partial x \partial y}, \frac{\partial^2 u(X)}{\partial y^2})^T.$$

Hence,

$$\begin{aligned}
\sum_{i=1}^{4}\frac{\partial I_i}{\partial y}\frac{\partial \phi_i(X)}{\partial x} &= -P\cdot\sum_{i=1}^{4}(A_i - X)\frac{\partial \phi_i(X)}{\partial x}\\
&\quad -\sum_{i=1}^{2}\frac{\partial}{\partial y}\Big[(N^-(\widetilde{A}_i) - I)\nabla u(X)(A_i - \widetilde{A}_i)\Big]\frac{\partial \phi_i(X)}{\partial x}. \qquad (4.47)
\end{aligned}$$

Taking the derivative for x on both sides of (4.34), we get

$$\begin{aligned}
&\mathbf{q}\cdot\sum_{i=1}^{4}(A_i - X)\frac{\partial \phi_i(X)}{\partial x}\\
=\ &\sum_{i=1}^{2}\Big[-(N^-_{\overline{DE}} - I)\mathbf{q}\cdot(A_i - \widetilde{A}_i)\frac{\partial \phi_i(X)}{\partial x} - (N^-_{\overline{DE}} - I)\mathbf{q}\cdot(\widetilde{A}_i - X^{(i)}_{\overline{DE}})\frac{\partial \phi_i(X)}{\partial x}\Big]\\
&-d^-\Big[\sum_{i=1}^{2}\Big[N^-_{\overline{DE}}(0,-1)^T\cdot(A_i - X^{(i)}_{\overline{DE}})\phi_i(X) + N^-_{\overline{DE}}(X^{(i)s}_{\overline{DE}} - X^s)\cdot(A_i - X^{(i)}_{\overline{DE}})\frac{\partial \phi_i(X)}{\partial x}\\
&+(x_i - \bar{x}_i)(y_i - \bar{y}_i)\frac{\partial \phi_i(X)}{\partial x} - (\bar{y}_i - y)\phi_i(X) + (\bar{x}_i - x)(\bar{y}_i - y)\frac{\partial \phi_i(X)}{\partial x}\Big]\\
&+\sum_{i=3}^{4}\Big[-(y_i - y)\phi_i(X) + (x_i - x)(y_i - y)\frac{\partial \phi_1(X)}{\partial x}\Big]\Big]. \qquad (4.48)
\end{aligned}$$

Applying (4.48) to the first term on the right hand side of (4.47) and letting

$$\mathbf{q} = P, d^- = \frac{\partial^3 u(X)}{\partial x \partial y^2},$$

we have

$$\begin{aligned}
&\sum_{i=1}^{4}\frac{\partial I_i}{\partial y}\frac{\partial \phi_i(X)}{\partial x}\\
=\ &\sum_{i=1}^{2}\Big\{-\frac{\partial}{\partial y}\Big[(N^-(\widetilde{A}_i)-I)\nabla u(X)(A_i-\widetilde{A}_i)\Big]\frac{\partial \phi_i(X)}{\partial x}+(N^-_{\overline{DE}}-I)P\cdot(A_i-\widetilde{A}_i)\frac{\partial \phi_i(X)}{\partial x}\\
&+(N^-_{\overline{DE}}-I)P\cdot(\widetilde{A}_i-X_{\overline{DE}})\frac{\partial \phi_i(X)}{\partial x}\Big\}\\
&-\frac{\partial^3 u(X)}{\partial x\partial y^2}\Big\{\sum_{i=1}^{2}\Big[N^-_{\overline{DE}}(0,-1)^T\cdot(A_i-X^{(i)}_{\overline{DE}})\phi_i(X)+N^-_{\overline{DE}}(X^{(i)s}_{\overline{DE}}-X^s)\cdot(A_i-X_{\overline{DE}})\frac{\partial \phi_i(X)}{\partial x}\\
&+(x_i-\bar{x}_i)(y_i-\bar{y}_i)\frac{\partial \phi_i(X)}{\partial x}-(\bar{y}_i-y)\phi_i(X)+(\bar{x}_i-x)(\bar{y}_i-y)\frac{\partial \phi_i(X)}{\partial x}\Big]\\
&+\sum_{i=3}^{4}\Big[-(y_i-y)\phi_i(X)+(x_i-x)(y_i-y)\frac{\partial \phi_i(X)}{\partial x}\Big]\Big\}. \qquad (4.49)
\end{aligned}$$

For $i=1,2$, by direct calculations similar to (4.44), we also have

$$\begin{aligned}
&\frac{\partial}{\partial y}(N^-(\widetilde{A}_i)-N^-_{\overline{DE}})\nabla u(X)\cdot(A_i-\widetilde{A}_i)\frac{\partial \phi_i(X)}{\partial x}\\
&-\frac{\partial}{\partial y}\Big[(N^-_{\overline{DE}}-I)\nabla u(X)\cdot(\widetilde{A}_i-X^{(i)}_{\overline{DE}})\Big]\frac{\partial \phi_i(X)}{\partial x}\\
&-\frac{\partial}{\partial y}\Big[(N^-(\widetilde{A}_i)-I)\nabla u(X)(A_i-\widetilde{A}_i)\Big]\frac{\partial \phi_i(X)}{\partial x}\\
&+(N^-_{\overline{DE}}-I)P\cdot(A_i-\widetilde{A}_i)\frac{\partial \phi_i(X)}{\partial x}+(N^-_{\overline{DE}}-I)P\cdot(\widetilde{A}_1-X^{(i)}_{\overline{DE}})\frac{\partial \phi_1(X)}{\partial x}\\
=\ &0. \qquad (4.50)
\end{aligned}$$

Plugging (4.49) and (4.50) into (4.46), we finish the proof of (4.45).

■

Based on the above expansions, we get the following theorem.

Theorem 4.1.17 *There exits a constant C independent of interface and mesh such that*

$$\|I_{h,T}u-u\|_{0,\tilde{T}^-\cap T^-}\le Ch^2\left(|u|_{1,\tilde{T}^-\cap T^-}+|u|_{2,\tilde{T}^-\cap T^-}\right)\le Ch^2\|u\|_{2,T}, \qquad (4.51)$$

$$|I_{h,T}u-u|_{1,\tilde{T}^-\cap T^-}\le Ch\left(|u|_{1,\tilde{T}^-\cap T^-}+|u|_{2,\tilde{T}^-\cap T^-}\right)\le Ch\|u\|_{2,T}, \qquad (4.52)$$

$$|I_{h,T}u-u|_{2,\tilde{T}^-\cap T^-}\le C\left(|u|_{1,\tilde{T}^-\cap T^-}+|u|_{2,\tilde{T}^-\cap T^-}\right)\le C\|u\|_{2,T}, \qquad (4.53)$$

for any $u\in PH^2_{int}(T)$ where T is a Type II interface element.

Proof. Because of (H_6), we need only show that this is true for any $u \in PC^3_{int}(T)$. For (4.51), the result follows by letting $X^{(i)s}_{\overline{DE}} = \widetilde{A}_i^{\perp}(i = 1, 2)$ in (4.35) and applying arguments similar to those used in the proof of Theorem 4.1.2. For (4.52), the result follows by letting $X^{(i)s}_{\overline{DE}} = \widetilde{A}_i^{\perp}(i = 1, 2)$ in (4.40) and (4.41) and applying arguments similar to those used in the proof of Theorem 4.1.2. Note that (3.12) in Theorem 3.4.2 is used here. Finally, since $\frac{\partial^2(I_{h,T}u)}{\partial x^2} = \frac{\partial^2(I_{h,T}u)}{\partial y^2} = 0$, then we complete the proof of (4.53) by applying the same techniques of Theorem 4.1.2 to (4.45). Note that (3.13) in Theorem 3.4.2 is used here.

■

The estimate on $\tilde{T}^+ \cap T^+$ is rather similar to that on $\tilde{T}^- \cap T^-$, so we only state the results in this section. Let $X = (x, y)^T$ be a point in $\tilde{T}^+ \cap T^+$. Without loss of generality, we assume that line segments $\overline{XA_i}, i = 1, 2$ do not intersect with the interface and $\overline{DE}$, while line segment $\overline{XA_i}, i = 3, 4$ meet Γ at $\widetilde{A}_i, i = 3, 4$, see Figure 4.4. Also, we assume that $A_i = (x_i, y_i)^T, i = 1, 2, 3, 4$ and

$$\widetilde{A}_i = \tilde{t}_i A_i + (1 - \tilde{t}_i)X = (\widetilde{x}_i, \widetilde{y}_i)^T, i = 3, 4.$$

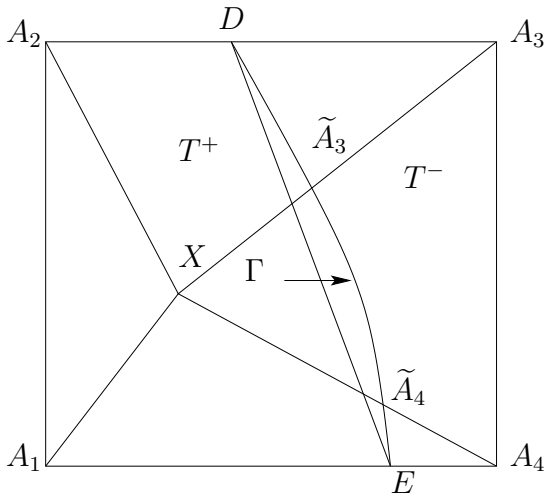

Figure 4.4: A point X $\in \tilde{T}^+ \cap T^+$ is connected to the four vertices by line segments in a Type II interface element

Lemma 4.1.12 *Given a two-dimensional vector* $\mathbf{q}$*, a point* $X \in \tilde{T}^+$ *and two real numbers* r, d^+*, then there exists a function* $v \in S_h(T)$ *such that* $\nabla v(X) = \mathbf{q}, v(X) = r, \frac{\partial^2 v^+(X)}{\partial x \partial y} = d^+$

and

$$\begin{aligned}
&\mathbf{q}(X)\cdot\sum_{i=1}^{4}(A_i-X)\phi_i(X)\\
=\;&\sum_{i=3}^{4}\Big[-(N^+_{\overline{DE}}-I)\mathbf{q}\cdot(A_i-\widetilde{A}_i)\phi_i(X)-(N^+_{\overline{DE}}-I)\mathbf{q}\cdot(\widetilde{A}_i-X^{(i)}_{\overline{DE}})\phi_i(X)\Big]\\
&-d^+\Big[\sum_{i=3}^{4}\Big[N^+_{\overline{DE}}(X^{(i)s}_{\overline{DE}}-X^s)\cdot(A_i-X^{(i)}_{\overline{DE}})\phi_i(X)+(x_i-\bar{x}_i)(y_i-\bar{y}_i)\phi_i(X)\\
&+(\bar{x}_i-x)(\bar{y}_i-y)\phi_i(X)\Big]+\sum_{i=1}^{2}\Big[(x_i-x)(y_i-y)\phi_i(X)\Big]\Big],
\end{aligned}$$

where $X^{(i)}_{\overline{DE}}=(\bar{x}_i,\bar{y}_i)^T$, $i=3,4$ *are arbitrary points on* $\overline{DE}$.

■

Theorem 4.1.18 *For any* $u\in PC^2_{int}(T)$ *and* $X=(x,y)^T\in\tilde{T}^+\cap T^+$, *we have*

$$\begin{aligned}
&I_{h,T}u(X)-u(X)\\
=\;&\sum_{i=3}^{4}\Big[(N^+(\widetilde{A}_i)-N^+_{\overline{DE}})\nabla u(X)\cdot(A_i-\widetilde{A}_i)\phi_i(X)-(N^+_{\overline{DE}}-I)\nabla u(X)\cdot(\widetilde{A}_i-X^{(i)}_{\overline{DE}})\phi_i(X)\Big]\\
&-\frac{\partial^2u(X)}{\partial x\partial y}\Big[\sum_{i=3}^{4}\Big[N^+_{\overline{DE}}(X^{(i)s}_{\overline{DE}}-X^s)\cdot(A_i-X^{(i)}_{\overline{DE}})\phi_i(X)+(x_i-\bar{x}_i)(y_i-\bar{y}_i)\phi_i(X)\\
&+(\bar{x}_i-x)(\bar{y}_i-y)\phi_i(X)\Big]+\sum_{i=1}^{2}(x_i-x)(y_i-y)\phi_i(X)\Big]\\
&+\sum_{i=3}^{4}\Big[(N^+(\widetilde{A}_i)-I)\int_0^1\frac{d[\nabla u^+(t\widetilde{A}_i+(1-t)X)]}{dt}\cdot(A_i-\widetilde{A}_i)\;dt\;\phi_i(X)\\
&+\int_0^{\tilde{t}_i}(1-t)\frac{d^2u(tA_i+(1-t)X)}{dt^2}\;dt\;\phi_i(X)+\int_{\tilde{t}_i}^1(1-t)\frac{d^2u(tA_i+(1-t)X)}{dt^2}\;dt\;\phi_i(X)\Big]\\
&+\sum_{i=1}^{2}\int_0^1(1-t)\frac{d^2u(tA_i+(1-t)X)}{dt^2}\;dt\;\phi_i(X),
\end{aligned}\tag{4.54}$$

where $X^{(i)}_{\overline{DE}}=(\bar{x}_i,\bar{y}_i)^T$, $i=3,4$ *are arbitrary points on* $\overline{DE}$.

■

Let I_i, $i = 1, 2, 3, 4$ be the integral terms involving vertices A_i, $i = 1, 2, 3, 4$ in (4.54).

Theorem 4.1.19 *For any $u \in PC^3_{int}(T)$ and $X = (x, y)^T \in \tilde{T}^+ \cap T^+$, we have*

$$\begin{aligned}
&\frac{\partial(I_{h,T}u(X) - u(X))}{\partial x} \\
= \ & \sum_{i=3}^{4} \Big[(N^+(\widetilde{A}_i) - N^+_{\overline{DE}}) \nabla u(X) \cdot (A_i - \widetilde{A}_i) \frac{\partial \phi_i(X)}{\partial x} \\
& - (N^+_{\overline{DE}} - I) \nabla u(X) \cdot (\widetilde{A}_i - X^{(i)}_{\overline{DE}}) \frac{\partial \phi_i(X)}{\partial x} \Big] \\
& - \frac{\partial^2 u(X)}{\partial x \partial y} \Big[\sum_{i=3}^{4} \Big[N^+_{\overline{DE}} (X^{(i)s}_{\overline{DE}} - X^s) \cdot (A_i - X^{(i)}_{\overline{DE}}) \frac{\partial \phi_i(X)}{\partial x} \\
& + N^+_{\overline{DE}} (0, -1)^T \cdot (A_i - X^{(i)}_{\overline{DE}}) \phi_i(X) + (x_i - \bar{x}_i)(y_i - \bar{y}_i) \frac{\partial \phi_i(X)}{\partial x} \\
& - (\bar{y}_i - y) \phi_i(X) + (\bar{x}_i - x)(\bar{y}_i - y) \frac{\partial \phi_i(X)}{\partial x} \Big] \\
& + \sum_{i=1}^{2} \Big[- (y_i - y) \phi_i(X) + (x_i - x)(y_i - y) \frac{\partial \phi_i(X)}{\partial x} \Big] \Big] + \sum_{i=1}^{4} I_i \frac{\partial \phi_i(X)}{\partial x},
\end{aligned}$$

$$\begin{aligned}
&\frac{\partial(I_{h,T}u(X) - u(X))}{\partial y} \\
= \ & \sum_{i=3}^{4} \Big[(N^+(\widetilde{A}_i) - N^+_{\overline{DE}}) \nabla u(X) \cdot (A_i - \widetilde{A}_i) \frac{\partial \phi_i(X)}{\partial y} \\
& - (N^+_{\overline{DE}} - I) \nabla u(X) \cdot (\widetilde{A}_i - X^{(i)}_{\overline{DE}}) \frac{\partial \phi_i(X)}{\partial y} \Big] \\
& - \frac{\partial^2 u(X)}{\partial x \partial y} \Big[\sum_{i=3}^{4} \Big[N^+_{\overline{DE}} (X^{(i)s}_{\overline{DE}} - X^s) \cdot (A_i - X^{(i)}_{\overline{DE}}) \frac{\partial \phi_i(X)}{\partial y} \\
& + N^+_{\overline{DE}} (-1, 0)^T \cdot (A_i - X^{(i)}_{\overline{DE}}) \phi_i(X) + (x_i - \bar{x}_i)(y_i - \bar{y}_i) \frac{\partial \phi_i(X)}{\partial y} \\
& - (\bar{x}_i - x) \phi_i(X) + (\bar{x}_i - x)(\bar{y}_i - y) \frac{\partial \phi_i(X)}{\partial y} \Big] \\
& + \sum_{i=1}^{2} \Big[- (x_i - x) \phi_i(X) + (x_i - x)(y_i - y) \frac{\partial \phi_i(X)}{\partial y} \Big] \Big] + \sum_{i=1}^{4} I_i \frac{\partial \phi_i(X)}{\partial y},
\end{aligned}$$

where $X^{(i)}_{\overline{DE}} = (\bar{x}_i, \bar{y}_i)^T$, $i = 3, 4$ are arbitrary points on $\overline{DE}$.

Theorem 4.1.20 *For any $u \in PC^3_{int}(T)$ and $X = (x, y)^T \in \tilde{T}^+ \cap T^+$, we have*

$$\begin{aligned}
&\frac{\partial^2 (I_{h,T}u(X) - u(X))}{\partial x \partial y} \\
= \ & \sum_{i=3}^{4} \Big[(N^+(\widetilde{A}_i) - N^+_{\overline{DE}}) \nabla u(X) \cdot (A_i - \widetilde{A}_i) \frac{\partial^2 \phi_i(X)}{\partial x \partial y} \\
& - (N^+_{\overline{DE}} - I) \nabla u(X) \cdot (\widetilde{A}_i - X^{(i)}_{\overline{DE}}) \frac{\partial^2 \phi_i(X)}{\partial x \partial y} \Big] \\
& - \frac{\partial^2 u(X)}{\partial x \partial y} \Big[\sum_{i=3}^{4} N^+_{\overline{DE}} (X^{(i)s}_{\overline{DE}} - X^s) \cdot (A_i - X^{(i)}_{\overline{DE}}) \frac{\partial^2 \phi_i(X)}{\partial x \partial y} \\
& + N^+_{\overline{DE}} (-1, 0)^T \cdot (A_i - X^{(i)}_{\overline{DE}}) \frac{\partial \phi_i(X)}{\partial x} + N^+_{\overline{DE}} (0, -1)^T \cdot (A_i - X^{(i)}_{\overline{DE}}) \frac{\partial \phi_i(X)}{\partial y} \\
& + (x_i - \bar{x}_i)(y_i - \bar{y}_i) \frac{\partial^2 \phi_i(X)}{\partial x \partial y} + \phi_i(X) - (\bar{y}_i - y) \frac{\partial \phi_i(X)}{\partial y} - (\bar{x}_i - x) \frac{\partial \phi_i(X)}{\partial x} \\
& + (\bar{x}_i - x)(\bar{y}_i - y) \frac{\partial^2 \phi_i(X)}{\partial x \partial y} \Big] + \sum_{i=1}^{2} \Big[\phi_i(X) - (y_i - y) \frac{\partial \phi_i(X)}{\partial y} - (x_i - x) \frac{\partial \phi_i(X)}{\partial x} \\
& + (x_i - x)(y_i - y) \frac{\partial^2 \phi_i(X)}{\partial x \partial y} \Big] \Big] + \sum_{i=1}^{4} I_i \frac{\partial^2 \phi_i(X)}{\partial x \partial y},
\end{aligned}$$

where $X^{(i)}_{\overline{DE}} = (\bar{x}_i, \bar{y}_i)^T$, $i = 1, 2$ are arbitrary points on $\overline{DE}$.

Theorem 4.1.21 *There exits a constant C independent of interface and mesh such that*

$$\begin{aligned}
&\|I_{h,T}u - u\|_{0,\tilde{T}^+ \cap T^+} \le Ch^2 \left(|u|_{1,\tilde{T}^+ \cap T^+} + |u|_{2,\tilde{T}^+ \cap T^+} \right) \le Ch^2 \|u\|_{2,T}, \\
&|I_{h,T}u - u|_{1,\tilde{T}^+ \cap T^+} \le Ch \left(|u|_{1,\tilde{T}^+ \cap T^+} + |u|_{2,\tilde{T}^+ \cap T^+} \right) \le Ch \|u\|_{2,T}, \\
&|I_{h,T}u - u|_{2,\tilde{T}^+ \cap T^+} \le C \left(|u|_{1,\tilde{T}^+ \cap T^+} + |u|_{2,\tilde{T}^+ \cap T^+} \right) \le C \|u\|_{2,T},
\end{aligned}$$

for any $u \in PH^2_{int}(T)$ where T is a Type II interface rectangle.

Following the idea of Theorem 4.1.11, we can obtain the following two lemmas.

Theorem 4.1.22 *If $\tilde{T}^+ \cap T^-$ is not empty, then there exits a constant C independent of interface and mesh such that*

$$\begin{aligned}
&\|I_{h,T}u - u\|_{0,\tilde{T}^+ \cap T^-} \le Ch^2 \left(|u|_{1,\tilde{T}^+ \cap T^-} + |u|_{2,\tilde{T}^+ \cap T^-} \right) \le Ch^2 \|u\|_{2,T}, \\
&|I_{h,T}u - u|_{1,\tilde{T}^+ \cap T^-} \le Ch \left(|u|_{1,\tilde{T}^+ \cap T^-} + |u|_{2,\tilde{T}^+ \cap T^-} \right) \le Ch \|u\|_{2,T}, \\
&|I_{h,T}u - u|_{2,\tilde{T}^+ \cap T^-} \le C \left(|u|_{1,\tilde{T}^+ \cap T^-} + |u|_{2,\tilde{T}^+ \cap T^-} \right) \le C \|u\|_{2,T},
\end{aligned}$$

for any $u \in PH^2_{int}(T)$, where T is a Type II interface element.

Theorem 4.1.23 *If $\tilde{T}^- \cap T^+$ is not empty, there exits a constant C independent of interface and mesh such that*

$$\begin{aligned}
&\|I_{h,T}u - u\|_{0,\tilde{T}^- \cap T^+} \leq Ch^2 \left(|u|_{1,\tilde{T}^- \cap T^+} + |u|_{2,\tilde{T}^- \cap T^+} \right) \leq Ch^2 \|u\|_{2,T}, \\
&|I_{h,T}u - u|_{1,\tilde{T}^- \cap T^+} \leq Ch \left(|u|_{1,\tilde{T}^- \cap T^+} + |u|_{2,\tilde{T}^- \cap T^+} \right) \leq Ch \|u\|_{2,T}, \\
&|I_{h,T}u - u|_{2,\tilde{T}^- \cap T^+} \leq C \left(|u|_{1,\tilde{T}^- \cap T^+} + |u|_{2,\tilde{T}^- \cap T^+} \right) \leq C \|u\|_{2,T},
\end{aligned}$$

for any $u \in PH^2_{int}(T)$, where T is a Type II interface element.

Finally, combining Theorems 4.1.17, 4.1.21, 4.1.22 and 4.1.23, we have the following theorem for the interpolation error on each Type II interface element.

Theorem 4.1.24 *There exits a constant C independent of interface and mesh such that*

$$\begin{aligned}
&\|I_{h,T}u - u\|_{0,T} \leq Ch^2 \left(|u|_{1,T} + |u|_{2,T} \right) \leq Ch^2 \|u\|_{2,T}, \\
&|I_{h,T}u - u|_{1,T} \leq Ch \left(|u|_{1,T} + |u|_{2,T} \right) \leq Ch \|u\|_{2,T}, \\
&|I_{h,T}u - u|_{2,T} \leq C \left(|u|_{1,T} + |u|_{2,T} \right) \leq C \|u\|_{2,T},
\end{aligned}$$

for any $u \in PH^2_{int}(T)$, where T is a Type II interface element.

4.1.4 Interpolation error on Ω

We now ready to derive the error estimates for the interpolation $I_h u$ in $S_h(\Omega)$.

Theorem 4.1.25 *There exists a constant C independent of interface and mesh such that*

$$\|I_h u - u\|_{0,\Omega} \leq Ch^2 \|u\|_{2,\Omega}, \tag{4.55}$$
$$|I_h u - u|_{1,\Omega} \leq Ch \|u\|_{2,\Omega}, \tag{4.56}$$
$$|I_h u - u|_{2,\Omega} \leq C \|u\|_{2,\Omega}, \tag{4.57}$$

for any $u \in PH^2_{int}(\Omega)$ and $h > 0$ small enough.

Proof. First we have

$$\|I_h u - u\|^2_{0,\Omega} = \sum_{T \in \mathcal{T}_h} \|I_{h,T}u - u\|^2_{0,T}.$$

If T is a Type I interface element, by Theorem 4.1.13, we have

$$\|I_{h,T}u - u\|_{0,T}^2 \leq Ch^4 \|u\|_{2,T}^2 .$$

Similarly, if T is a Type II interface element, by Theorem 4.1.24, we have

$$\|I_{h,T}u - u\|_{0,T}^2 \leq Ch^4 \|u\|_{2,T}^2 .$$

If T is a non-interface element, by the standard finite element interpolation error theory, we can get

$$\|I_{h,T}u - u\|_{0,T}^2 \leq Ch^4 \|u\|_{2,T}^2$$

Therefore, we have

$$\|I_h u - u\|_{0,\Omega}^2 \leq \sum_{T\in\mathcal{T}_h} Ch^4 \|u\|_{2,T}^2 = Ch^4 \|u\|_{2,\Omega}^2 ,$$

which leads to (4.55). Similar derivation can be carried out to obtain (4.56) and (4.57).

■

4.2 Numerical examples

We now present a group of numerical results to illustrate the approximation features of the bilinear IFE space.

For simplicity, we only present results obtained by using the bilinear IFE space based on uniformly rectangular Cartesian partitions in the rectangular domain $\Omega = (-1, 1) \times (-1, 1)$. The interface curve Γ is a circle with radius $r_0 = \pi/6.28$ which separates Ω into two subdomains Ω^- and Ω^+ with $\Omega^- = \{(x, y) \mid x^2 + y^2 \leq r_0^2\}$. Here we show numerical results for the bilinear IFE interpolation $I_h u$ of the following function

$$u(x,y) = \begin{cases} \dfrac{r^\alpha}{\beta^-}, & \text{if } r \leq r_0, \\ \dfrac{r^\alpha}{\beta^+} + \left(\dfrac{1}{\beta^-} - \dfrac{1}{\beta^+}\right) r_0^\alpha, & \text{otherwise,} \end{cases}$$

with $\alpha = 5, r = \sqrt{x^2 + y^2}$. In the following tables, $\|\cdot\|_0$ represents the usual L^2 norm, $|\cdot|_1$ is the usual semi-H^1 norm, and of course, they are computed numerically according to the mesh used.

Table 4.1 contains actual errors of the IFE interpolation $I_h u$ with various partition sizes h for $\beta^- = 1, \beta^+ = 10$ which represents a moderate discontinuity in the coefficient. By simple calculations, we can easily see that the data in this table satisfy

$$\|I_h u - u\|_0 \approx \frac{1}{4} \|I_{\hat{h}} u - u\|_0 , |I_h u - u|_1 \approx \frac{1}{2} |I_{\hat{h}} u - u|_1 ,$$

for $h = \hat{h}/2$. Using linear regression, we can also see that the data in this table obey

$$\|I_h u - u\|_0 \approx 0.3750\ h^{1.996}, |I_h u - u|_1 \approx 0.9405\ h^{1.002},$$

which clearly indicates that the interpolation converges to u with convergence rates $O(h^2)$ and $O(h)$ in the L^2 norm and H^1 norm, respectively, as predicted by Theorem 4.1.25.

Table 4.2 contains actual errors of the IFE interpolation $I_h u$ with various partition size h for $\beta^- = 1, \beta^+ = 10000$ which represents a large discontinuity in the coefficient. Using linear regression again, we can see that

$$\|I_h u - u\|_0 \approx 0.09557\ h^{1.954}, |I_h u - u|_1 \approx 0.3582\ h^{1.030},$$

which are also in agreement with the error estimates given in Theorem 4.1.25. From Figure 4.5 for the linear regressions above, we can see that the data in Table 4.1 and 4.2 match the linear regression lines very well.

h	$\|I_h u - u\|_0$	$\|I_h u - u\|_1$
1/16	0.001479	0.05848
1/32	3.715×10^{-4}	0.02918
1/64	9.321×10^{-5}	0.01453
1/128	2.334×10^{-5}	0.007264
1/256	5.840×10^{-6}	0.003635

Table 4.1: Errors in the interpolation $I_h u$ when $\beta^- = 1, \beta^+ = 10$

h	$\|I_h u - u\|_0$	$\|I_h u - u\|_1$
1/16	4.159×10^{-4}	0.02089
1/32	1.104×10^{-4}	0.01011
1/64	2.878×10^{-5}	0.004832
1/128	7.323×10^{-6}	0.002401
1/256	1.848×10^{-6}	0.001209

Table 4.2: Errors in the interpolation $I_h u$ when $\beta^- = 1, \beta^+ = 10000$

Remark 4.2.1 *When we compute the errors in L^2 norm and H^1 semi-norm in the tables above, we use Gauss quadratures. Because the integrands of these errors usually lack the smoothness which is required by Gauss quadratures, some extra errors are caused by the lack of smoothness. However, in each local interface element, these extra errors only exist in the parts bounded by the original interface Γ and its approximation $\overline{DE}$, whose area is $O(h^3)$. Since the integrands are usually bounded, then the extra error in that area is $O(h^3)$, which doesn't change the interpolation error order. All the errors calculated for the numerical examples in this dissertation have similar situation.*

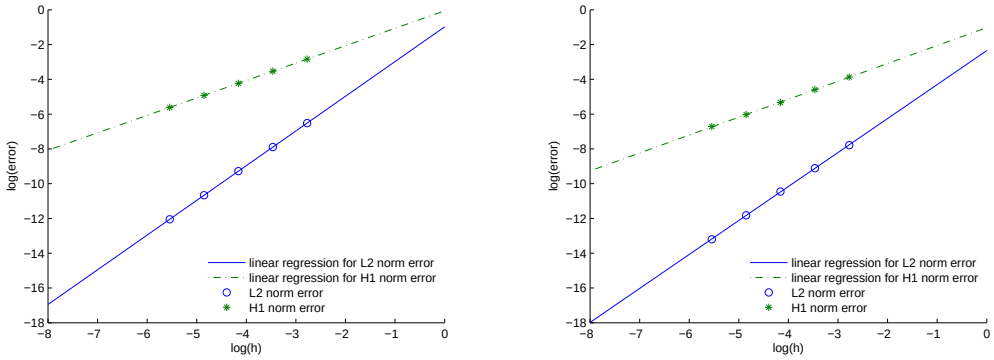

Figure 4.5: The plot on the left is for the linear regression of the data in Table 4.1 and the plot on the right is for the linear regression of the data in Table 4.2.

Chapter 5

Galerkin method with bilinear IFE

In Chapter 3 and Chapter 4, we construct the bilinear IFE space $S_h(\Omega)$ and analyze its approximation capability. Now we start to apply bilinear IFE to different numerical methods and discuss the corresponding convergence analysis. In this chapter, we will first discuss the Galerkin method based on the bilinear IFE space $S_h(\Omega)$, which was originally introduced in [149]. Then we will analyze its convergence [110].

5.1 Galerkin method based on bilinear IFE

In this section we recall the Galerkin method based on the bilinear IFE space $S_h(\Omega)$ from [149]. Let $\mathcal{T}_h, h > 0$ be a family of rectangular meshes of the solution domain Ω that can be a union of rectangles. Let

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

We multiply the differential equation (1.1) by any $v \in H_0^1(\Omega)$ and integrate it over $\Omega^s (s = +, -)$ to have

$$-\int_{\Omega^s} \nabla \cdot (\beta^s \nabla u) v \; dxdy \;=\; \int_{\Omega^s} fv \; dxdy, \forall \; v \in H_0^1(\Omega).$$

Then a straightforward application of the Green's formula leads to

$$\int_{\Omega^s} \beta^s \nabla u \cdot \nabla v \; dxdy - \int_{\partial\Omega^s} \beta \frac{\partial u}{\partial \mathbf{n}} v \; ds \;=\; \int_{\Omega^s} fv \; dxdy, \; s = +, -, \forall \; v \in H_0^1(\Omega). \quad (5.1)$$

Summing (5.1) over s, we get the weak formulation

$$\int_{\Omega} \beta \nabla u \cdot \nabla v \; dxdy \;=\; \int_{\Omega} fv \; dxdy, \forall \; v \in H_0^1(\Omega). \quad (5.2)$$

Here we have used the flux jump condition (1.4) and $v \in H_0^1(\Omega)$.

Let $S_{h,0}(\Omega) \subset S_h(\Omega)$ consist of functions of $S_h(\Omega)$ vanishing on $\mathcal{N}_h \cap \partial\Omega$. The bilinear IFE Galerkin method can be described as follows: find $u_h \in S_h(\Omega)$ satisfying

$$\sum_{T\in\mathcal{T}_h} \int_T \beta \nabla u_h \cdot \nabla v_h \, dxdy = \int_\Omega f v_h \, dxdy, \forall \, v_h \in S_{h,0}(\Omega). \tag{5.3}$$

5.2 Numerical Examples

Since this bilinear IFE space has an $O(h^2)$ (in L^2-norm) and an $O(h)$ (in H^1-norm) approximation capability, we naturally expect the finite element method based on this IFE space to perform accordingly, which is confirmed numerically in [112, 149]. In this section, for the comparison among different numerical methods with bilinear IFE in Section 8.2.2, we will present some numerical results about the convergence of the IFE-Galerkin method for solving the following model interface problem

$$\begin{aligned} &-\nabla \cdot \big(\beta \nabla u\big) = f, \quad (x,y) \in \Omega, \\ &u|_{\partial\Omega} = g, \end{aligned}$$

together with the jump conditions on the interface Γ:

$$\begin{aligned} [u]\,|_\Gamma &= 0, \\ \left[\beta \frac{\partial u}{\partial n}\right]|_\Gamma &= 0. \end{aligned}$$

Here $\Omega = [-1,1] \times [-1,1]$. The interface curve Γ is a circle with radius $r_0 = \pi/6.28$ that separates Ω into two sub-domains Ω^- and Ω^+ with $\Omega^- = \{(x,y) \mid x^2 + y^2 \le r_0^2\}$. The coefficient function is

$$\beta(x,y) = \begin{cases} \beta^-, & (x,y) \in \Omega^-, \\ \beta^+, & (x,y) \in \Omega^+. \end{cases}$$

The boundary condition function $g(x,y)$ and the source term $f(x,y)$ are chosen such that the following function u is the exact solution.

$$u(x,y) = \begin{cases} \dfrac{r^\alpha}{\beta^-}, & \text{if } r \le r_0, \\ \dfrac{r^\alpha}{\beta^+} + \left(\dfrac{1}{\beta^-} - \dfrac{1}{\beta^+}\right) r_0^\alpha, & \text{otherwise}, \end{cases}$$

with $\alpha = 3, r = \sqrt{x^2+y^2}$. For simplicity, we only use the simple rectangular Cartesian meshes in our numerical experiments. In the following tables, $\|\cdot\|_0$ represents the usual L^2 norm, $|\cdot|_1$ is the usual semi-H^1 norm, and of course, they are computed numerically according

to the mesh used. The quantity $\|\cdot\|_\infty$ is the discrete infinity norm which is the maximum of the absolute values of the given function at all the nodes of a mesh.

Table 5.1 contains actual errors of the bilinear IFE solutions u_h with various partition size h for the interface problem with $\beta^- = 1, \beta^+ = 10$. We can easily see that the data in the second and third columns of this table satisfy

$$\|u_h - u\|_0 \approx \frac{1}{4} \|u_{\hat{h}} - u\|_0 , |u_h - u|_1 \approx \frac{1}{2} |u_{\hat{h}} - u|_1 ,$$

for $h = \hat{h}/2$. Using linear regression, we can also see that the data in this table obey

$$\|u_h - u\|_0 \approx 0.2789 \; h^{2.0204}, |u_h - u|_1 \approx 0.6855 \; h^{0.9525},$$

which indicates that the bilinear IFE solution u_h converges to the exact solution with convergence rates $O(h^2)$ and $O(h)$ in the L^2 norm and H^1 norm, respectively. This is in agreement with those error estimates for the bilinear IFE interpolation obtained in Chapter 4.

However, numerical experiments indicate that the bilinear IFE solution does not always have the second order convergence in the L^∞ norm because the data in the fourth column of Table 5.1 obey

$$|u_h - u|_\infty \approx 0.0179 \; h^{0.7643},$$

which clearly shows that the rate at which u_h converges to u is not $O(h^2)$. The question under what conditions the bilinear IFE solution can have a second order convergence in the L^∞ norm is still open.

The bilinear IFE method also works well for the case in which the coefficient function has a large jump, see Table 5.2, which contains actual errors of the bilinear IFE solutions u_h with various partition size h for the interface problem with $\beta^- = 1, \beta^+ = 10000$. The errors in this group of computations obey

$$\|u_h - u\|_0 \approx 0.2123 \; h^{1.9582}, |u_h - u|_1 \approx 0.8855 \; h^{1.0802},$$

which again are in agreement with those error estimates for the bilinear IFE interpolation. From Figure 5.1 for the linear regressions above, we can see that the data points match the linear regression lines very well.

5.3 The convergence of Galerkin method based on bilinear IFE

From the numerical experiments published up to now, we have observed that all of the IFE methods converge and can generate approximate solutions to interface problems with

h	$\|u_h - u\|_0$	$\|u_h - u\|_1$	$\|u_h - u\|_\infty$
1/8	4.2061×10^{-3}	9.6080×10^{-2}	-2.9687×10^{-3}
1/16	1.0652×10^{-3}	4.9346×10^{-2}	-2.7375×10^{-3}
1/32	2.4680×10^{-4}	2.4517×10^{-2}	-1.2725×10^{-3}
1/64	5.8112×10^{-5}	1.2633×10^{-2}	-8.0369×10^{-4}
1/128	1.6384×10^{-5}	6.9959×10^{-3}	-3.8749×10^{-4}

Table 5.1: Errors of the IFE solutions for the case when $\beta^- = 1, \beta^+ = 10$.

h	$\|u_h - u\|_0$	$\|u_h - u\|_1$	$\|u_h - u\|_\infty$
1/8	3.4231×10^{-3}	9.1187×10^{-2}	-2.4971×10^{-3}
1/16	9.5498×10^{-4}	4.5672×10^{-2}	-8.6241×10^{-4}
1/32	2.5688×10^{-4}	2.1478×10^{-2}	-4.6777×10^{-4}
1/64	6.1961×10^{-5}	9.6034×10^{-2}	-1.6250×10^{-4}
1/128	1.5168×10^{-5}	4.7067×10^{-3}	-6.6273×10^{-5}

Table 5.2: Errors of the IFE solutions for the case when $\beta^- = 1, \beta^+ = 10000$.

the same optimal convergence rates as the corresponding standard finite element spaces. For example, the Galerkin methods with the bilinear or linear IFE spaces have the $O(h^2)$ convergence rate for L^2 norm and $O(h)$ convergence rate for H^1 norm [112, 143, 144, 149]. However, to our knowledge, the convergence analysis for IFE methods has been carried out only for 1D problems [3, 150] where the IFE methods are conforming methods. The analysis for 2D and 3D IFE methods is more complicated because of the discontinuity in the functions of the involved IFE spaces. Since the 2D and 3D interface problems are much more important from the point of view of applications, the theoretical analysis on the convergence of the 2D and 3D IFE methods demands immediate attentions.

In this section, we will analyze the convergence of the bilinear Galerkin IFE solution for the model interface problem (1.1)-(1.4). The core effort of our analysis is to derive error bounds in which the constants C are independent of interface and mesh.

5.3.1 Some preliminaries and notations

In this section, we will introduce some preliminaries and notations which will be used for the convergence analysis. We will use the same definitions and notations defined at the beginning of Chapter 3 and Chapter 4. We will also use $\mathcal{T}_h$ to denote the collection of all elements in a mesh with size parameter h. We note that when h is small enough, most of elements in $\mathcal{T}_h$ are non-interface elements not intersecting with the interface Γ. Only those elements in the vicinity of Γ have the possibility to be cut through by Γ and become the

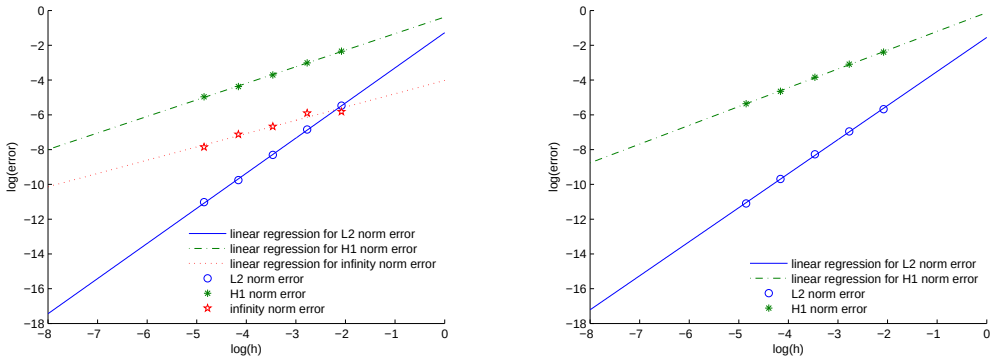

Figure 5.1: The plot on the left is for the linear regression of the data in Table 5.1 and the plot on the right is for the linear regression of the data in Table 5.2.

so-called interface elements. We will use $\mathcal{T}_{int}$ to denote the collection of all interface elements of $\mathcal{T}_h$ and let $\Omega_{int} = \cup_{T\in\mathcal{T}_{int}} T$.

In all the discussions from now on, we assume that the usual hypothesis (H_1)-(H_5) used in Chapter 3 and Chapter 4 are also hold. In addition, we also assume that

(H_6): The boundary of Ω and the interface Γ are such that the following Sobolev embedding inequality holds:

$$\|v\|_{0,p,\Omega} \leq Cp^{1/2} \|v\|_{1,\Omega}, \quad \forall p > 2, \quad \forall v \in PH^1_{int}(\Omega). \tag{5.4}$$

We would like to point out that when Ω has a smooth boundary, the lemma 2.1 in [180] states that (5.4) is true for $v \in H^1_0(\Omega)$. Using the extension technique, see Section 6.5 of [162] for example, we can see that this inequality is also true for $v \in H^1(\Omega)$.

For the convergence analysis, we will use the following notations to differentiate the regular bilinear finite elements and the bilinear IFE. On each element T, we first let $S_h^{non}(T)$ be spanned by the 4 regular bilinear nodal basis functions $\psi_i(x,y), i = 1,2,3,4$ on T. On each of the interface rectangular element T, we assume that the vertices of an interface element T are $A_i, i = 1,2,3,4$, with $A_i = (x_i, y_i)^t$. We define $\phi_i(X)$ to be the bilinear IFE function described by (3.4) such that

$$\phi_i(x_j, y_j) = \begin{cases} 1, & \text{if } i = j, \\ 0, & \text{if } i \neq j, \end{cases}$$

for $1 \leq i, j \leq 4$. We then let $S_h^{int}(T) = span\{\phi_i,\ i = 1,2,3,4\}$.

5.3.2 Bilinear interpolation of bilinear IFE functions

In this section, we present several lemmas about the bilinear interpolation of bilinear IFE functions. We will discuss both Type I and Type II interface elements configured as in Figure 3.1 for the local bilinear IFE space. Some notations introduced in Chapter 3 will be used. We use C to represent a generic constant whose value might be different from line to line. Unless otherwise specified, all the generic constants C in the presentation below are independent of interface and mesh.

We now consider the relationship between $S_h^{non}(T)$ and $S_h^{int}(T)$. Consider the following mappings:

$$\bar{I}_h : C(T) \longrightarrow S_h^{non}(T), \bar{I}_h\phi(X) = \sum_{i=1}^{4} \phi(A_i)\psi_i(X),$$

$$\tilde{I}_h : C(T) \longrightarrow S_h^{int}(T), \tilde{I}_h\phi(X) = \sum_{i=1}^{d} \phi(A_i)\phi_i(X).$$

By direct verification, we can easily see that

$$\bar{I}_h\tilde{I}_h\bar{u}_h = \bar{u}_h, \quad \forall \bar{u}_h \in S_h^{non}(T), \quad \tilde{I}_h\bar{I}_h\tilde{u}_h = \tilde{u}_h, \quad \forall \tilde{u}_h \in S_h^{int}(T),$$

and

$$\left\|\tilde{u}_h - \bar{I}_h\tilde{u}_h\right\|_{k,T} = \left\|\bar{u}_h - \tilde{I}_h\bar{u}_h\right\|_{k,T},$$

where

$$k = 0, 1, \ \bar{u}_h \in S_h^{non}(T), \tilde{u}_h \in S_h^{int}(T) \text{ such that } \bar{u}_h = \bar{I}_h\tilde{u}_h \text{ or } \tilde{u}_h = \tilde{I}_h\bar{u}_h.$$

In fact, our interpolation operator I_h for the bilinear IFE space can be locally described by

$$I_h = \begin{cases} \bar{I}_h, & \text{on non-interface elements,} \\ \\ \tilde{I}_h, & \text{on interface elements.} \end{cases}$$

Lemma 5.3.1 *There exists a constant C such that*

$$\left\|\bar{I}_h\tilde{u}_h\right\|_{0,T} \leq C\left\|\tilde{u}_h\right\|_{0,T}, \quad \forall \tilde{u}_h \in S_h^{int}(T). \tag{5.5}$$

Proof. Here we will use the notations and conclusions of Lemma 3.4.4 and Lemma 3.4.5. Let

$$\bar{u}_h = \bar{I}_h\tilde{u}_h = \sum_{i=1}^{4} \tilde{u}_h(A_i)\psi_i(X). \tag{5.6}$$

By the shape regular assumption, there exist two positive constants C_1 and C_2 independent of interface and mesh such that $C_1 h_T^2 \le |J_F| \le C_2 h_T^2$. Then using (3.17) and (5.6), we get

$$\begin{aligned} \|\bar{u}_h\|_{0,T}^2 &= \int_T \bar{u}_h^2(X)dX = \int_{\hat{T}} \widehat{\bar{u}_h}^2(\hat{X})\,|J_F|\,d\hat{X} \le Ch_T^2 \int_{\hat{T}} \widehat{\bar{u}_h}^2(\hat{X})d\hat{X} \\ &\le Ch_T^2 \|\vec{u}\|^2 \le Ch_T^2(\left\|\vec{u}^-\right\|^2 + \left\|\vec{u}^+\right\|^2) \\ &\le Ch_T^2 \|\vec{u}^s\|^2,\ s=+,-. \end{aligned} \tag{5.7}$$

Let s be such that $T_{1/t} \subset T^s$ for $t = 2$ or 4 depending on the type of T, then using representation (3.15) and (3.16) for $\tilde{u}_h$, we have

$$\begin{aligned} \|\tilde{u}_h\|_{0,T}^2 &= \left\|\tilde{u}_h^-\right\|_{0,T^-}^2 + \left\|\tilde{u}_h^+\right\|_{0,T^+}^2 \ge \|\tilde{u}_h^s\|_{0,T^s}^2 = \int_{T^s} [\tilde{u}_h^s(X)]^2\,dX \\ &= \int_{\hat{T}^s} \widehat{\tilde{u}_h^s}^2(\hat{X})\,|J_F|\,d\hat{X} \ge Ch_T^2 \int_{\hat{T}^s} \widehat{\tilde{u}_h^s}^2(\hat{X})d\hat{X} \\ &\ge Ch_T^2 \int_{\hat{T}_{1/t}} \widehat{\tilde{u}_h^s}^2(\hat{X})d\hat{X} = Ch_T^2 (\vec{u}^s)^t \hat{A}_{1/2}\vec{u}^s \\ &\ge Ch_T^2 \|\vec{u}^s\|^2 \end{aligned} \tag{5.8}$$

for some positive constant C where we have used the fact that matrix

$$\hat{A}_{1/2} = \left(\int_{\hat{T}_{1/t}} \hat{\psi}_i \hat{\psi}_j d\hat{X} \right)_{i,j=1}^4$$

is positive definite. Combining the two inequalities (5.7) and (5.8) leads to (5.5).

■

Lemma 5.3.2 *There exits a constant C such that*

$$\left\|\tilde{u}_h - \bar{I}_h\tilde{u}_h\right\|_{0,T} \le Ch\,\|\tilde{u}_h\|_{1,T},\quad \forall \tilde{u}_h \in S_h^{int}(T). \tag{5.9}$$

Proof. Let $\hat{\bar{I}}_h$ denote the interpolation $\bar{I}_h$ defined on $\hat{T}$. Applying the result in Lemma 5.3.1 on $\hat{T}$, we have

$$\begin{aligned} \left\|\widehat{\tilde{u}_h} - \widehat{\bar{I}_h\tilde{u}_h}\right\|_{0,\hat{T}} &= \left\|\widehat{\tilde{u}_h} - c - \left(\hat{\bar{I}}_h\widehat{\tilde{u}}_h - \hat{\bar{I}}_h c\right)\right\|_{0,\hat{T}} \\ &\le \left\|\widehat{\tilde{u}_h} - c\right\| + \left\|\hat{\bar{I}}_h\left(\widehat{\tilde{u}}_h - c\right)\right\|_{0,\hat{T}} \\ &\le C\left\|\widehat{\tilde{u}}_h - c\right\|_{0,\hat{T}} \\ &\le C\left\|\hat{\tilde{u}}_h - c\right\|_{1,\hat{T}}, \end{aligned}$$

for any constant c and this leads to

$$\begin{aligned}\left\|\widehat{\widetilde{u}_h}-\widehat{\bar{I}_h\widetilde{u}_h}\right\|_{0,\hat{T}} &\leq C\inf_c\left\|\widehat{\widetilde{u}}_h-c\right\|_{1,\hat{T}}\\ &\leq C\left|\widehat{\widetilde{u}}_h\right|_{1,\hat{T}}.\end{aligned}$$

Then (5.9) follows from the standard scaling argument.

■

For those interface elements as configured in Figures 3.1, we have the following lemma.

Lemma 5.3.3 *Assume f is a continuous piecewise bilinear function and its two pieces on $\overline{A_1A_2}$ is separated by $E\in\overline{A_1A_2}$. If $f(A_1)=f(A_2)=0$ and $\|A_1A_2\|=h$, then*

$$|f(E)| \leq \frac{1}{2}\left|[\nabla f(E)]\cdot\vec{v}_{\overline{A_1A_2}}\right|h. \tag{5.10}$$

Proof. For any continuous piecewise linear function f on $\overline{A_1A_2}$ with $f(A_1)=f(A_2)=0$, we can use the Taylor expansion to obtain the following:

$$\begin{aligned}0 &= f(A_1) = f(E^-)+\nabla f(E^-)\cdot(A_1-E) = f(E)+\nabla f(E^-)\cdot(A_1-E),\\ 0 &= f(A_2) = f(E^+)+\nabla f(E^+)\cdot(A_2-E) = f(E)+\nabla f(E^+)\cdot(A_2-E).\end{aligned}$$

Hence

$$f(E) = \frac{1}{2}[-\nabla f(E^-)\cdot(A_1-E)-\nabla f(E^+)\cdot(A_2-E)]$$

Let $\vec{v}_{\overline{A_1A_2}}$ denote the unit vector pointing from A_1 to A_2, then

$$f(E) = \frac{1}{2}[\nabla f(E^-)\cdot\vec{v}_{\overline{A_1A_2}}\|E-A_1\|-\nabla f(E^+)\cdot\vec{v}_{\overline{A_1A_2}}\|A_2-E\|]$$

Since $f(A_1)=f(A_2)=0$ and f is continuous, then

$$\begin{aligned}&\nabla f(E^-)\cdot\vec{v}_{\overline{A_1A_2}}>0,\ \nabla f(E^+)\cdot\vec{v}_{\overline{A_1A_2}}<0 \text{ if } f(E)>0,\\ &\nabla f(E^-)\cdot\vec{v}_{\overline{A_1A_2}}<0,\ \nabla f(E^+)\cdot\vec{v}_{\overline{A_1A_2}}>0 \text{ if } f(E)<0.\end{aligned}$$

Because $\|E-A_1\|\leq h$ and $\|A_2-E\|\leq h$, we get

$$\begin{aligned}f(E) &\leq \frac{1}{2}[\nabla f(E^-)\cdot\vec{v}_{\overline{A_1A_2}}h-\nabla f(E^+)\cdot\vec{v}_{\overline{A_1A_2}}h] \leq \frac{1}{2}\left|[\nabla f(E)]\cdot\vec{v}_{\overline{A_1A_2}}\right|h \text{ if } f(E)>0,\\ f(E) &\geq \frac{1}{2}[\nabla f(E^-)\cdot\vec{v}_{\overline{A_1A_2}}h-\nabla f(E^+)\cdot\vec{v}_{\overline{A_1A_2}}h] \geq -\frac{1}{2}\left|[\nabla f(E)]\cdot\vec{v}_{\overline{A_1A_2}}\right|h \text{ if } f(E)<0,\end{aligned}$$

which lead to (5.10).

■

We derive an estimate for the difference $\tilde{u}_h - \bar{I}_h\tilde{u}_h$ on ∂T in the following lemma.

Lemma 5.3.4 *There exits a constant C such that*

$$\left\|\tilde{u}_h - \bar{I}_h\tilde{u}_h\right\|_{0,\partial T} \leq Ch^{1/2}\left|\tilde{u}_h\right|_{1,T}, \quad \forall \tilde{u}_h \in S_h^{int}(T). \tag{5.11}$$

Proof. Without loss of generality, we consider those interface elements as configured in Figures 3.1. First we discuss interface elements of Type I. We note that $\tilde{u}_h$ is a C^0 piecewise bilinear function and $\bar{I}_h\tilde{u}_h$ is its bilinear interpolation. Hence

$$\left\|\tilde{u}_h - \bar{I}_h\tilde{u}_h\right\|_{0,\overline{A_2A_3}} = \left\|\tilde{u}_h - \bar{I}_h\tilde{u}_h\right\|_{0,\overline{A_3A_4}} = 0.$$

In addition, since $\tilde{u}_h(X) - \bar{I}_h\tilde{u}_h(X)$ is piecewise linear on $\overline{A_1A_2}$ and vanishes at A_1 and A_2, we can show that

$$\max_{X\in\overline{A_1A_2}} \left|\tilde{u}_h(X) - \bar{I}_h\tilde{u}_h(X)\right| \leq \left|\tilde{u}_h(E) - \bar{I}_h\tilde{u}_h(E)\right|. \tag{5.12}$$

Since $\bar{I}_h\tilde{u}_h$ is continuous on T, by applying (5.10) to $f(X) = \tilde{u}_h(X) - \bar{I}_h\tilde{u}_h(X)$, we have

$$\begin{aligned}
\left|\tilde{u}_h(E) - \bar{I}_h\tilde{u}_h(E)\right| &\leq \frac{1}{2}\left|\left[\nabla(\tilde{u}_h(E) - \bar{I}_h\tilde{u}_h(E))\right]\cdot\vec{v}_{\overline{A_1A_2}}\right| h_x \\
&= \frac{1}{2}\left|\left[\nabla\tilde{u}_h(E)\right]\cdot\vec{v}_{\overline{A_1A_2}}\right| h_x \\
&\leq C\left\|\nabla\tilde{u}_h^+(E)\right\| h_x \\
&\leq C\left|\tilde{u}_h\right|_{1,T},
\end{aligned} \tag{5.13}$$

where we have applied Lemma 4.1.1 and the inverse inequality (3.19) to obtain the last two inequalities above, respectively. Then, the inequalities (5.12) and (5.13) lead to

$$\begin{aligned}
\left\|\tilde{u}_h - \bar{I}_h\tilde{u}_h\right\|_{0,\overline{A_1A_2}} &= \left[\int_{\overline{A_1A_2}}\left[\tilde{u}_h(x) - \bar{I}_h\tilde{u}_h(x)\right]^2 dx\right]^{1/2} \\
&\leq \left[\left[\tilde{u}_h(E) - \bar{I}_h\tilde{u}_h(E)\right]^2\int_{\overline{A_1A_2}} dx\right]^{1/2} \\
&\leq h_x^{1/2}\left|\tilde{u}_h(E) - \bar{I}_h\tilde{u}_h(E)\right| \\
&\leq Ch^{1/2}\left|\tilde{u}_h\right|_{1,T}.
\end{aligned}$$

Similar estimate can be derived for $\left\|\tilde{u}_h - \bar{I}_h\tilde{u}_h\right\|_{0,\overline{A_4A_1}}$ and these estimates yield (5.11) for interface elements of Type I.

Similarly, for interface elements of Type II, we have

$$\begin{aligned}\left\|\tilde{u}_h-\bar{I}_h\tilde{u}_h\right\|_{0,\overline{A_2A_3}}&=\left\|\tilde{u}_h-\bar{I}_h\tilde{u}_h\right\|_{0,\overline{A_4A_1}}=0,\\ \left\|\tilde{u}_h-\bar{I}_h\tilde{u}_h\right\|_{0,\overline{A_1A_2}}&\leq Ch^{1/2}\left|\tilde{u}_h\right|_{1,T},\\ \left\|\tilde{u}_h-\bar{I}_h\tilde{u}_h\right\|_{0,\overline{A_3A_4}}&\leq Ch^{1/2}\left|\tilde{u}_h\right|_{1,T},\end{aligned}$$

which yield (5.11) for rectangular interface elements of Type II.

■

5.3.3 Error bounds for the bilinear IFE solution in H^1 norm

We now consider the error bound for the immersed finite element solution of the interface (1.1)-(1.4). Assuming the exact solution u of the interface problem (1.1)-(1.4) has the PH^2_{int} regularity, then the general Berger-Scott-Strang lemma [28] implies that the error in the IFE solution generated by (5.3) have the following error bound:

$$\left|u-u_h\right|_{1,h}\leq C\left(\inf_{v_h\in S_h(\Omega)}\left|u-v_h\right|_{1,h}+\sup_{v_h\in S_h(\Omega)}\frac{\left|a_h(u,v_h)-(f,v_h)\right|}{\left|v_h\right|_{1,h}}\right), \tag{5.14}$$

where

$$a_h(w,v)=\sum_{T\in\mathcal{T}_h}\int_T\beta\nabla w\nabla v dX,\ \ |v|^2_{1,h}=\sum_{T\in\mathcal{T}_h}|v|^2_{1,T}.$$

and $(w,v)_\Lambda$ denotes the L^2 inner product of w and v on a set Λ and we often omit its set symbol when $\Lambda=\Omega$.

Using Theorem 4.1.25, the first term on the right hand side of (5.14) has the following bound:

$$\inf_{v_h\in S_h(\Omega)}\left|u-v_h\right|_{1,h}\leq\left|u-I_hu\right|_{1,h}\leq Ch\left\|u\right\|_2. \tag{5.15}$$

Therefore, we will focus on the estimation of the consistency error term on the right hand of (5.14).

We first derive some preparation estimates. The following trace inequality is similar to the regular trace inequality on $H^2(T)$ [7], but we need to prove it on $PH^2_{int}(T)$. For each element $T=\square A_1A_2A_3A_4\in\mathcal{T}_h$, we let

$$E_1(\partial T)=\overline{A_1A_2},\ E_2(\partial T)=\overline{A_2A_3},\ E_3(\partial T)=\overline{A_3A_4},\ E_4(\partial T)=\overline{A_4A_1}.$$

Lemma 5.3.5 *We have the following trace inequality on $T\in\mathcal{T}_h$:*

$$\left\|\beta\frac{\partial v}{\partial n}\right\|^2_{0,E_i(\partial T)}\leq C\left(\frac{1}{h_T}\left|v\right|^2_{1,T}+h_T\left|v\right|^2_{2,T}\right),\quad\forall v\in PH^2_{int}(T),\quad 1\leq i\leq 4. \tag{5.16}$$

Proof. Without loss generality, we assume that $T \in \mathcal{T}_{int}$. For any $v \in PH^2_{int}(T)$, we let

$$\vec{q} = \beta \begin{bmatrix} v_x \\ v_y \end{bmatrix}, w = \beta(v_{xx} + v_{yy}).$$

Then for any $\phi \in C_0^\infty(T)$, we can easily see that

$$\int_T w\phi dX = -\int_T \vec{q} \cdot \nabla\phi dX.$$

This implies that $\vec{q} \in H(div, T)$ and $div(\vec{q}) = \beta(v_{xx} + v_{yy})$. On the reference element, we recall the following standard inequality for functions in $H(div, \hat{T})$ [76]:

$$\left\| \hat{\vec{q}} \cdot \mathbf{n} \right\|^2_{1/2, E_i(\partial\hat{T})} \leq C \left(\left\| \hat{\vec{q}} \right\|^2_{0,\hat{T}} + \left\| div(\hat{\vec{q}}) \right\|^2_{0,\hat{T}} \right).$$

Then, we can obtain (5.16) by using the above trace inequality on the reference element and the usual scaling procedure.

■

Lemma 5.3.6 *There exists a constant C such that for h small enough, we have*

$$\left(\sum_{T \in \mathcal{T}_{int}} \sum_{i=1}^{4} \left\| \beta \frac{\partial v}{\partial n} \right\|^2_{0, E_i(\partial T)} \right)^{1/2} \leq C \left| \log(h) \right|^{1/2} \|v\|_2, \forall v \in PH^2_{int}(\Omega). \tag{5.17}$$

Proof. Using Hölder's inequality, we have

$$\begin{aligned} \|f\|_{0,r,\Omega} &\leq \left[\int_\Omega f^r dxdy \right]^{\frac{1}{r}} \\ &\leq \left[\left[\int_\Omega (f^r)^{\frac{p}{r}} dxdy \right]^{\frac{r}{p}} \right]^{\frac{1}{r}} \left[\left[\int_\Omega 1^{\frac{r}{r-p}} dxdy \right]^{1-\frac{r}{p}} \right]^{\frac{1}{r}} \\ &= |\Omega|^{\frac{1}{r}-\frac{1}{p}} \|f\|_{0,p,\Omega}. \end{aligned}$$

Then letting $f = v_x,\ v_y$ and using the assumption $(\mathcal{H}_6)$ and the fact that $|\Omega_{int}| \leq Ch$, we get

$$|v|_{1,\Omega_{int}} \leq |\Omega_{int}|^{1/2-1/p} |v|_{1,p,\Omega_{int}} \leq Ch^{1/2-1/p} p^{1/2} \|v\|_2.$$

Then, using Lemma 5.3.5, we can obtain (5.17) as follows:

$$\begin{aligned}
\left(\sum_{T\in\mathcal{T}_{int}}\sum_{i=1}^{d}\left\|\beta\frac{\partial v}{\partial n}\right\|^2_{0,E_i(\partial T)}\right)^{1/2} &\leq C\left[\sum_{T\in\mathcal{T}_{int}}\left(\frac{1}{h_T}|v|^2_{1,T}+h_T|v|^2_{2,T}\right)\right]^{1/2}\\
&\leq C\left(\frac{1}{h}|v|^2_{1,\Omega_{int}}+h|v|^2_{2,\Omega_{int}}\right)^{1/2}\\
&\leq C\left(ph^{-2/p}\|v\|_2^2+h\|v\|_2^2\right)^{1/2}\\
&= C\left(ph^{-2/p}+h\right)^{1/2}\|v\|_2\\
&= Cp^{1/2}\left(h^{-2/p}+h/p\right)^{1/2}\|v\|_2\\
&\leq C|\log(h)|^{1/2}\|v\|_2.
\end{aligned}$$

In the last inequality, we let $p=|\log(h)|$ for h small enough and use $h^{1/log(h)}=e$.

■

Now, we are ready to derive an error bound in the discrete H^1 norm for the bilinear IFE solution.

Theorem 5.3.1 *There exists a constant C such that*

$$|u-u_h|_{1,h}\leq Ch^{1/2}|\log(h)|^{1/2}\|u\|_{2,\Omega}. \tag{5.18}$$

Proof. First, we multiply the differential equation (1.1) by any $v_h\in S_{h,0}$ and integrate it over $T\in\mathcal{T}_h$ to have

$$-\int_T\nabla\cdot(\beta\nabla u)v_h\,dxdy = \int_T fv_h\,dxdy, \forall\ v_h\in S_{h,0}.$$

Then a straightforward application of the Green's formula leads to

$$\int_T\beta\nabla u\cdot\nabla v_h\,dxdy-\int_{\partial T}\beta\frac{\partial u}{\partial\mathbf{n}}v_h\,ds = \int_T fv_h\,dxdy,\ \forall\ v_h\in S_{h,0}. \tag{5.19}$$

Summing (5.19) over $T\in\mathcal{T}_h$, we get

$$\sum_{T\in\mathcal{T}_h}\int_T\beta\nabla u\cdot\nabla v_h\,dxdy-\sum_{T\in\mathcal{T}_h}\int_{\partial T}\beta\frac{\partial u}{\partial\mathbf{n}}v_h\,ds = \sum_{T\in\mathcal{T}_h}\int_T fv_h\,dxdy,\ \forall\ v_h\in S_{h,0}.$$

Hence we have

$$a_h(u,v_h)-(f,v_h)=\sum_{T\in\mathcal{T}_h}\left(\beta\frac{\partial u}{\partial n},v_h\right)_{\partial T},\forall v_h\in S_{h,0}.$$

Since $\bar{I}_h v_h$ and $\beta\dfrac{\partial u}{\partial n}$ are continuous and the unit normal vectors of an edge in its two neighbor elements have opposite directions, then we can show

$$\left|\sum_{T\in\mathcal{T}_h}\sum_{i=1}^{4}\left(\beta\frac{\partial u}{\partial n},\bar{I}_h v_h\right)_{E_i(\partial T)}\right| \;=\; 0, \forall v_h\in S_{h,0}.$$

Hence

$$|a_h(u,v_h)-(f,v_h)| \;=\; \left|\sum_{T\in\mathcal{T}_h}\sum_{i=1}^{4}\left(\beta\frac{\partial u}{\partial n},v_h-\bar{I}_h v_h\right)_{E_i(\partial T)}\right|, \forall v_h\in S_{h,0}.$$

Then, by Lemma 5.3.4 and Lemma 5.3.6, we have

$$\begin{aligned}
|a_h(u,v_h)-(f,v_h)| &\le \sum_{T\in\mathcal{T}_{int}}\sum_{i=1}^{4}\left\|\beta\frac{\partial u}{\partial n}\right\|_{0,E_i(\partial T)}\left\|v_h-\bar{I}_h v_h\right\|_{0,E_i(\partial T)}\\
&\le Ch^{1/2}\sum_{T\in\mathcal{T}_{int}}\sum_{i=1}^{4}\left\|\beta\frac{\partial u}{\partial n}\right\|_{0,E_i(\partial T)}|v_h|_{1,T}\\
&\le Ch^{1/2}\left(\sum_{T\in\mathcal{T}_{int}}\sum_{i=1}^{4}\left\|\beta\frac{\partial u}{\partial n}\right\|^2_{0,E_i(\partial T)}\right)^{1/2}\left(\sum_{T\in\mathcal{T}_{int}}\sum_{i=1}^{4}|v_h|^2_{1,T}\right)^{1/2}\\
&\le Ch^{1/2}\left|\log(h)\right|^{1/2}|v_h|_{1,h}\,\|u\|_{2,\Omega}\,. \qquad (5.20)
\end{aligned}$$

Finally, the result of this theorem follows from applying the Berger-Scott-Strang inequality (5.14), the bound for the interpolation error (5.15) and (5.20).

■

The error bound obtained in Theorem 5.3.1 clearly indicates that the bilinear IFE solution converges to the exact solution of the interface problem when the mesh size h tends to zero. On the other hand, we note that the error bound in this theorem has a $O(h^{1/2})$ convergence rate which is sub-optimal from the point of view of the approximation capability of the bilinear IFE space. Recall that the H^1-norm error bound for the interpolation in the bilinear IFE space is $O(h)$, see Theorem 4.1.25 and [112]. In addition, all the numerical examples in [112, 149] indicate that the bilinear IFE Galerkin method can generate approximated solutions to the interface problem with the optimal convergence rate. The analysis to show that the bilinear IFE methods have the optimal convergence rate is still an elusive and interesting research topic.

Chapter 6

Bilinear immersed finite volume element method

Conservation law is important in physics since it governs energy, momentum, angular momentum, mass, electric charge and so on. It states that a particular measurable property of an isolated physical system does not change as the system evolves. Meanwhile, finite volume element (FVE) method, which is also called box method [19] or generalized difference method [139], plays an important role in the numerical methods for PDEs because it possesses the well known local conservation property. Therefore, it has been widely studied, extended and applied for numerous problems, such as flow problems [38, 83, 90, 94, 154, 160, 192, 193], conservation law [124, 153], parabolic and hyperbolic equations [81, 198], and elliptic problems [29, 82, 159], to name just a few. Recently some literatures also discussed the analysis of this method, see [5, 30, 31, 36, 37, 46, 53, 59, 79, 80, 85, 203] and related reference therein. We believe that the combination of the FVE's local conservation property and IFE's flexibility to handle interface jump conditions without using complicated meshes can generate competitive numerical methods for solving interface problems. Therefore, in this chapter we will follow the idea in [84] to apply the bilinear IFE space to solve the interface problem of the diffusion equation in the finite volume element formulation [113].

6.1 Implementation of finite volume element method with bilinear IFE

In this section, we will discuss the finite volume element method with the bilinear IFE space $S_h(\Omega)$ introduced in Chapter 3. The following set-up is well known for the finite volume element method [139], but we still repeat it here for the basic idea and the notations used in this chapter. To describe the method, for each mesh $\mathcal{T}_h$ of Ω, we introduce a dual mesh $\widehat{\mathcal{T}}_h$ by connecting the nearby centers of the elements in $\mathcal{T}_h$ in the vertical and horizontal directions,

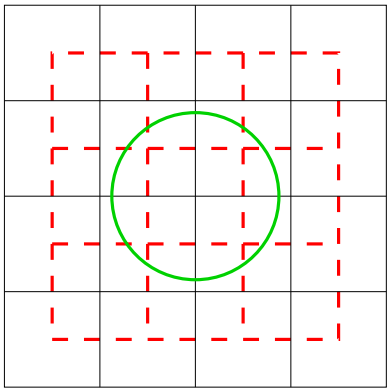

Figure 6.1: A mesh of Ω and the dual mesh for an interface problem. Elements in the mesh are solid rectangles and elements in the dual mesh are dash rectangles.

see the illustration in Figure 6.1 where the dual mesh $\widehat{\mathcal{T}}_h$ is sketched by the dash lines while $\mathcal{T}_h$ is sketched by solid lines.

First, we derive a weak form on each element of the dual mesh. Assume that the source term $f(X)$ is smooth enough so that the exact solution has the required smoothness in the discussion below. Let $\mathcal{N}_h$ be the set of all nodes of $\mathcal{T}_h$ and $\mathcal{N}_h^\circ$ be the set of all interior nodes of $\mathcal{T}_h$. Let $\widehat{K}_i$ be an element of $\widehat{\mathcal{T}}_h$ containing a node $X_i \in \mathcal{N}_h^\circ$. First, we integrate the differential equation (1.1) over $\widehat{K}_i$ to have

$$-\int_{\widehat{K}_i} \nabla \cdot (\beta \nabla u)\, dxdy \;=\; \int_{\widehat{K}_i} f\, dxdy.$$

If $\widehat{K}_i$ is not an interface element, then a straightforward application of the Green's formula leads to

$$-\int_{\partial \widehat{K}_i} \beta \frac{\partial u}{\partial \mathbf{n}}\, ds \;=\; \int_{\widehat{K}_i} f\, dxdy. \tag{6.1}$$

If $\widehat{K}_i$ is an interface element, then by applying the Green's formula piecewisely, we have

$$\begin{aligned} -\int_{\widehat{K}_i^-} \nabla \cdot (\beta \nabla u)\, dxdy - \int_{\widehat{K}_i^+} \nabla \cdot (\beta \nabla u)\, dxdy &= \int_{\widehat{K}_i} f\, dxdy, \\ -\int_{\partial \widehat{K}_i^-} \beta \frac{\partial u}{\partial \mathbf{n}}\, ds - \int_{\partial \widehat{K}_i^+} \beta \frac{\partial u}{\partial \mathbf{n}}\, ds &= \int_{\widehat{K}_i} f\, dxdy, \\ -\int_{\partial \widehat{K}_i} \beta \frac{\partial u}{\partial \mathbf{n}}\, ds - \int_{\partial \widehat{K} \cap \Gamma} \left[\beta \frac{\partial u}{\partial \mathbf{n}}\right]_\Gamma\, ds &= \int_{\widehat{K}_i} f\, dxdy, \end{aligned}$$

which leads to (6.1) again because of the flux jump condition (1.4). Hence, we conclude that the weak form (6.1) holds for any element $\widehat{K}_i \in \widehat{\mathcal{T}}_h$. This weak form enables us to introduce

the bilinear immersed finite volume element method as follows: find $u_h \in S_{h,E}(\Omega)$ such that

$$-\int_{\partial \widehat{K}_i} \beta \frac{\partial u_h}{\partial \mathbf{n}}\, ds \;=\; \int_{\widehat{K}_i} f\, dxdy, \quad \forall X_i \in \mathcal{N}_h^\circ. \tag{6.2}$$

Here, $S_{h,E}(\Omega) = \{v_h \in S_h(\Omega) : v_h(X) = g(X),\ \forall X \in \mathcal{N}_h \cap \partial\Omega\}$. We would like to point out that (6.2) indicates that the immersed FVE solution also have the local conservation property.

We now discuss some details in the implementation of the bilinear immersed FVE method. The key issue is the integrals used in this method. On each non-interface element $\widehat{K}_i$, standard Gaussian quadratures can be applied because we can assume that all the integrands involved are smooth enough. If $\widehat{K}_i$ is an interface element, both the line integral and the area integral in the bilinear immersed FVE method need to be treated carefully because of the discontinuity across the interface.

First, let us consider the area integral $\int_{\widehat{K}_i} f\, dxdy$ on the right hand side of (6.2). Under the assumption that $f(X)$ is piecewise smooth with respect to the interface Γ, we can approximate its integration over $\widehat{K}_i$ piecewisely by suitably partitioning $\widehat{K}_i$ into several sub-triangles. Assume that $\widehat{K}_i$ has vertices $\hat{X}_j,\ \ j = 1, 2, 3, 4$ and interface Γ intersects with the boundary of $\widehat{K}_i$ at D and E on two adjacent edges, see Figure 6.2. We can then use points D and E to partition $\widehat{K}_i$ into 4 triangles by adding 3 line segments: $\overline{DE}$, $\overline{D\hat{X}_3}$, $\overline{E\hat{X}_3}$. Note that the last two line segments are formed by connecting D and E to the vertex of $\widehat{K}_i$ not on the edges containing D and E. Hence,

$$\begin{aligned}\int_{\widehat{K}_i} f\, dxdy \;=\; &\int_{\triangle_{\hat{X}_1 ED}} f^-\, dxdy + \int_{\triangle_{E\hat{X}_2\hat{X}_3}} f^+\, dxdy \\ &+ \int_{\triangle_{D\hat{X}_3\hat{X}_4}} f^+\, dxdy + \int_{\triangle_{DE\hat{X}_3}} f\, dxdy.\end{aligned}$$

Gaussian quadratures with enough degree of precision can be applied straightforwardly to handel integrations on those sub-triangles within either Ω^- or Ω^+. A little extra care is need to handle the sub-triangles whose interiors intersect both Ω^- and Ω^+. For the case illustrated in Figure 6.2, when applying a Gaussian quadrature to compute $\int_{\triangle_{DE\hat{X}_3}} f\, dxdy$, we can replace the value of f at a quadrature node outside Ω^+ by the value of f at a point on Γ so long as this replacement has an $O(h^2)$ accuracy which can be achieved if Γ is smooth enough within $\widehat{K}_i$ [54]. Another way is to replace the value of f at a quadrature node outside Ω^+ by the value of f^+ at that quadrature node. That is, we use $\int_{\triangle_{DE\hat{X}_3}} f^+\, dxdy$ to replace $\int_{\triangle_{DE\hat{X}_3}} f\, dxdy$. A similar procedure can be developed for handling the case in which the interface Γ intersect with the boundary of $\widehat{K}_i$ at D and E on two opposite edges.

For an interface element $\widehat{K}_i$, the line integral on the left hand side of (6.2) also needs to be treated piecewisely to handle the discontinuity. Again, let us consider a dual element

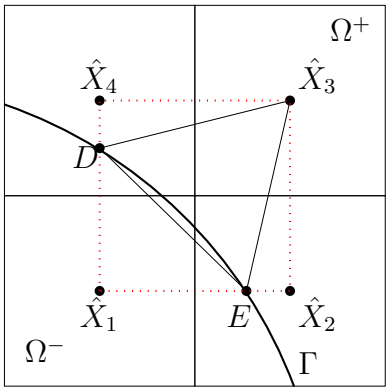

Figure 6.2: A dual element $\widehat{K} = \square\hat{X}_1\hat{X}_2\hat{X}_3\hat{X}_4 \in \widehat{\mathcal{T}}_h$ sketched by dash lines and 4 adjacent elements of $\mathcal{T}_h$. This element can be partitioned into 4 sub-triangles for the area integrals in the immerse FVE method.

$\widehat{K}_i = \square\hat{X}_1\hat{X}_2\hat{X}_3\hat{X}_4$, see Figure 6.3. Since $\widehat{K}_i$ has 4 edges, we have

$$\begin{aligned} -\int_{\partial\widehat{K}_i} \beta\frac{\partial u_h}{\partial \mathbf{n}}\, ds \;=\; & -\int_{\hat{X}_1\hat{X}_2} \beta\frac{\partial u_h}{\partial \mathbf{n}}\, ds - \int_{\hat{X}_2\hat{X}_3} \beta\frac{\partial u_h}{\partial \mathbf{n}}\, ds \\ & -\int_{\hat{X}_3\hat{X}_4} \beta\frac{\partial u_h}{\partial \mathbf{n}}\, ds - \int_{\hat{X}_4\hat{X}_1} \beta\frac{\partial u_h}{\partial \mathbf{n}}\, ds. \end{aligned}$$

Note that the flux $\beta\frac{\partial u_h}{\partial \mathbf{n}}$ on the boundary of $\widehat{K}_i$ is discontinuous at the points where $\partial\widehat{K}_i$ intersects either the edges of $\mathcal{T}_h$ or the interface Γ. Therefore, the line integrals on the right hand side above need to be computed on the small line segments between these discontinuous points. For the example demonstrated in Figure 6.3, we have

$$\begin{aligned} \int_{\hat{X}_1\hat{X}_2} \beta\frac{\partial u_h}{\partial \mathbf{n}}\, ds &= \int_{\hat{X}_1 A} \beta^-\frac{\partial u_h}{\partial \mathbf{n}}\, ds + \int_{\overline{AE}} \beta^-\frac{\partial u_h}{\partial \mathbf{n}}\, ds + \int_{E\hat{X}2} \beta^+\frac{\partial u_h}{\partial \mathbf{n}}\, ds, \\ \int_{\hat{X}_2\hat{X}_3} \beta\frac{\partial u_h}{\partial \mathbf{n}}\, ds &= \int_{\hat{X}_2 B} \beta^+\frac{\partial u_h}{\partial \mathbf{n}}\, ds + \int_{B\hat{X}_3} \beta^+\frac{\partial u_h}{\partial \mathbf{n}}\, ds, \\ \int_{\hat{X}_3\hat{X}_4} \beta\frac{\partial u_h}{\partial \mathbf{n}}\, ds &= \int_{\hat{X}_3 C} \beta^+\frac{\partial u_h}{\partial \mathbf{n}}\, ds + \int_{C\hat{X}_4} \beta^+\frac{\partial u_h}{\partial \mathbf{n}}\, ds, \\ \int_{\hat{X}_4\hat{X}_1} \beta\frac{\partial u_h}{\partial \mathbf{n}}\, ds &= \int_{\hat{X}_4 D} \beta^+\frac{\partial u_h}{\partial \mathbf{n}}\, ds + \int_{\overline{DF}} \beta^-\frac{\partial u_h}{\partial \mathbf{n}}\, ds + \int_{F\hat{X}_1} \beta^-\frac{\partial u_h}{\partial \mathbf{n}}\, ds. \end{aligned}$$

We note that all the integrands in the line integrals on the right hand sides above are polynomials; hence, a Gaussian quadrature with enough degree precision can be used to compute all of them precisely. As a consequence, this leads to another interesting fact that the matrix in the immersed FVE can be assembled exactly even if the interface Γ is a general curve. On the contrary, the matrices in the immersed finite element methods discussed in Section 5.1 and [112, 130, 143, 144, 149] cannot be formed precisely unless the interface Γ is trivial. In assembling the matrix in any of these immersed finite element methods over an

interface element $K \in \mathcal{T}_h$, assuming that the interface Γ intersects the edges of K at D and E, the error in the computation of the area integral over the region enclosed by $\overline{DE}$ and Γ is inevitable if Γ is a general curve.

Finally, we would like to point out that, for any given rectangular mesh $\mathcal{T}_h$ of Ω, the algebraic system of this bilinear immersed FVE method has the same structure as the algebraic system in the usual bilinear finite element method for the Dirichlet boundary value problem of the Poisson equation. The matrix in its algebraic system is guaranteed to be symmetric positive definite.

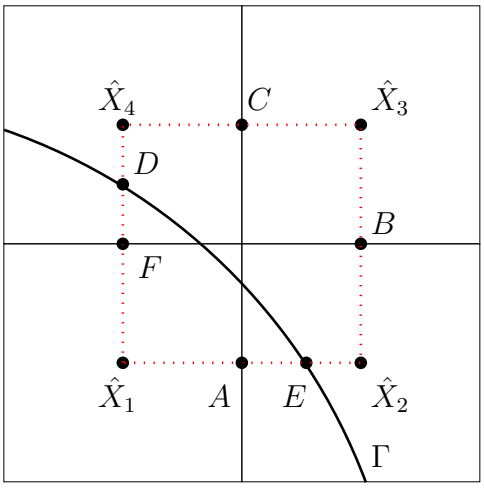

Figure 6.3: A dual element $\widehat{K}_i = \square\hat{X}_1\hat{X}_2\hat{X}_3\hat{X}_4 \in \widehat{\mathcal{T}}_h$ sketched by dash lines and 4 adjacent elements of $\mathcal{T}_h$. The edges of $\widehat{K}_i$ is partitioned by the discontinuous points of the flux for the line integrals in the immersed FVE method.

6.2 Numerical examples

In this section, we present numerical examples for the bilinear immersed finite volume element method to illustrate its features. We consider the same example as in Section 5.2.

Table 6.1 contains the errors of the bilinear immersed FVE solution u_h with various mesh size h and $\beta^- = 1, \ \beta^+ = 10$. Table 6.2 contains the errors of the bilinear immersed FVE solution u_h with $\beta^- = 1, \ \beta^+ = 10000$ representing a large jump. Table 6.3 contains the errors of the bilinear immersed FVE solution u_h with various mesh size h and $\beta^- = 10, \ \beta^+ = 1$. Table 6.4 contains the errors of the bilinear immersed FVE solution u_h with $\beta^- = 10000, \ \beta^+ = 1$. In these tables, $\|\cdot\|_0$ represents the usual L^2 norm, $|\cdot|_1$ is the usual semi-H^1 norm, and of course, they are computed numerically according to the mesh used. The quantity $\|\cdot\|_\infty$ is the discrete infinity norm which is the maximum of the absolute values of the given function at all the nodes of a mesh.

h	$\|u_h - u\|_0$	$\|u_h - u\|_1$	$\|u_h - u\|_\infty$
1/8	7.7394×10^{-3}	1.1705×10^{-1}	2.5110×10^{-3}
1/16	1.9658×10^{-3}	5.8644×10^{-2}	6.5026×10^{-4}
1/32	4.8127×10^{-4}	2.9255×10^{-2}	1.6598×10^{-4}
1/64	1.2173×10^{-4}	1.4550×10^{-2}	4.1413×10^{-5}
1/128	3.0115×10^{-5}	7.2699×10^{-3}	1.0611×10^{-5}
1/256	7.5436×10^{-6}	3.6362×10^{-3}	2.6485×10^{-6}

Table 6.1: Errors of the FV-IFE solution for the case with $\beta^- = 1, \quad \beta^+ = 10$

h	$\|u_h - u\|_0$	$\|u_h - u\|_1$	$\|u_h - u\|_\infty$
1/8	1.8420×10^{-3}	4.1025×10^{-2}	1.4562×10^{-3}
1/16	4.0555×10^{-4}	2.1051×10^{-2}	4.2813×10^{-4}
1/32	7.6016×10^{-5}	1.0193×10^{-2}	2.5606×10^{-4}
1/64	2.4890×10^{-5}	4.8512×10^{-3}	5.0649×10^{-5}
1/128	5.1332×10^{-6}	2.4100×10^{-3}	1.8048×10^{-5}
1/256	1.1050×10^{-6}	1.2110×10^{-3}	4.7363×10^{-6}

Table 6.2: Errors of the FV-IFE solution for the case with $\beta^- = 1, \quad \beta^+ = 10000.$

We can easily see that the data in the second and third columns of these tables satisfy

$$\|u_h - u\|_0 \approx \frac{1}{4} \|u_{\hat{h}} - u\|_0 , \quad |u_h - u|_1 \approx \frac{1}{2} |u_{\hat{h}} - u|_1 ,$$

for $h = \hat{h}/2$. Using linear regression, we can see that the data in Table 6.1 obey

$$\|u_h - u\|_0 \approx 0.5008 h^{2.0024}, \quad |u_h - u|_1 \approx 0.9427 \ h^{1.0025}, \quad \|u_h - u\|_\infty \approx 0.1559 \ h^{1.9788},$$

and the data in Table 6.2 obey

$$\|u_h - u\|_0 \approx 0.1422 \ h^{2.1154}, \quad |u_h - u|_1 \approx 0.3514 \ h^{1.0246}, \quad \|u_h - u\|_\infty \approx 0.0486 \ h^{1.6390},$$

h	$\|u_h - u\|_0$	$\|u_h - u\|_1$	$\|u_h - u\|_\infty$
1/8	7.6119×10^{-2}	1.0927×10^{0}	2.6593×10^{-2}
1/16	1.9110×10^{-2}	5.4809×10^{-1}	6.6274×10^{-3}
1/32	4.7894×10^{-3}	2.7425×10^{-1}	1.6796×10^{-3}
1/64	1.1967×10^{-3}	1.3715×10^{-1}	4.1590×10^{-4}
1/128	2.9946×10^{-4}	6.8576×10^{-2}	1.0489×10^{-4}
1/256	7.4846×10^{-5}	3.4288×10^{-2}	2.6144×10^{-5}

Table 6.3: Errors of the FV-IFE solution for the case with $\beta^- = 10, \quad \beta^+ = 1.$

h	$\|u_h - u\|_0$	$\|u_h - u\|_1$	$\|u_h - u\|_\infty$
1/8	7.6026×10^{-2}	1.0927×10^{0}	2.6270×10^{-2}
1/16	1.9119×10^{-2}	5.4813×10^{-1}	6.7172×10^{-3}
1/32	4.7613×10^{-3}	2.7425×10^{-1}	1.6608×10^{-3}
1/64	1.1930×10^{-3}	1.3714×10^{-1}	4.0496×10^{-4}
1/128	2.9813×10^{-4}	6.8575×10^{-2}	1.0940×10^{-4}
1/256	7.4494×10^{-5}	3.4288×10^{-2}	2.6902×10^{-5}

Table 6.4: Errors of the FV-IFE solution for the case with $\beta^- = 10000, \quad \beta^+ = 1$.

and the data in Table 6.3 obey

$$\|u_h - u\|_0 \approx 4.8643\ h^{1.9983}, \quad |u_h - u|_1 \approx 8.7375\ h^{0.9990}, \quad \|u_h - u\|_\infty \approx 1.6923\ h^{1.9974},$$

and the data in Table 6.4 obey

$$\|u_h - u\|_0 \approx 4.8715\ h^{1.9995}, \quad |u_h - u|_1 \approx 8.7379\ h^{0.9990}, \quad \|u_h - u\|_\infty \approx 1.6301\ h^{1.9861}.$$

For the linear regressions above, we obtain similar figures to Figure 5.1, which mean that the data points match the linear regression lines very well.

These results further indicate that the bilinear immersed FVE solution u_h converges to the exact solution with convergence rates $O(h^2)$ and $O(h)$ in the L^2 norm and H^1 norm, respectively. However, the actual computational results show that the solution does not always have the second order convergence in the L^∞ norm even though the mesh is fine enough. Similar phenomenon has been observed for IFE Galerkin method, see the numerical examples in Section 5.2. We guess this is mainly due to the non-conforming feature of the IFE space, and we plan to investigate this issue in our future research.

For a given rectangular mesh of Ω, we note that the linear system in this bilinear immerse FVE method has the same structure as that in the FVE method based on the standard bilinear finite elements for the Poisson's equation, especially from the point view of the number of non-zero entries and their locations in the matrix of the related linear system. This suggests that, on any given computer, the CPU time needed to solve the bilinear immerse FVE method should be comparable to that needed to solve the linear system in the standard bilinear FVE for simple Poisson's equation. Since it has become more and more difficult to obtain the precise CPU time usage of a computational procedure on a modern computer because of the complexity of the CUP unit (multi cores, cache, hardware parallelization, etc.) and the software (operating system, fire-wall, virus scan, etc.), we choose the number of iterations needed to make the preconditioned conjugate gradient (PCG) method to converge for a given error tolerance to illustrate the above observation, see Table 6.5. For the Ω specified at the beginning of this section, we use a rectangular mesh with $h = 1/128$, the incomplete Cholesky preconditioner, and the error tolerance $tol = 10^{-10}$ in all

the computations. From this table, we can see that, while the linear system in the bilinear FVE method for the Poisson's equations uses 221 PCG iterations, the the linear system in the bilinear immersed FVE method uses a 222 PCG iteration for the interface problem described in this section with $\beta^+ : \beta^- = 1 : 1.1$. We have also observed that the number of PCG iterations needed by the bilinear immersed FVE method gradually increases as the ratio $\beta^+ : \beta^-$ becomes larger. This increase is due to the fact that the interface problem is essentially more difficult than the simple boundary value problem of the Poisson's equation and will inevitably cost more time to solve by any method.

	bilinear FVE	$\beta^+ : \beta^- = 1 : 1.1$	$\beta^+ : \beta^- = 1 : 2$	$\beta^+ : \beta^- = 1 : 10$
# of iterations	221	222	279	299

Table 6.5: Comparison of the computational costs for solving linear systems in both the bilinear FVE method and the bilinear immersed FVE method.

Chapter 7

Immersed discontinuous Galerkin (IDG) method

The discontinuous Galerkin (DG) method was originally introduced for neutron transport equation [179] in 1973. Later on, the stability and convergence analysis was carried out [126, 127, 134, 176, 182]. Because of its high order accuracy, flexibility for mesh refinement, localizability, stability, parallelizability and less numerical diffusion/dispersion, the DG method has been widely extended and used to solve different kinds of partial differential equations, such as the total variation bounded(TVB) Runge-Kutta discontinuous Galerkin(RKDG) method [64, 65, 68, 69, 70, 72, 73, 177], the local discontinuous Galerkin(LDG) method [35, 44, 45, 47, 66, 67, 71, 104, 105, 131, 202], and many others, see [1, 20, 21, 88, 89, 92, 96, 99, 106, 107, 135, 140, 155, 172, 197] and reference therein.

In parallel, different discontinuous Galerkin finite element methods have also been developed for elliptic problems, including the interior penalty DG method [7, 10, 122, 128, 169, 183, 200] and the mixed DG method [48, 49, 50, 51]. Other related applications and analysis can be found in [43, 52, 78, 101, 102, 103, 132, 133, 190] and reference therein. For more details, we refer readers to the survey papers [8, 50].

The inclusion of penalty terms in the variational form defining a finite element method for elliptic equations is not new [9, 11, 18, 128, 200]. In 1979, M. Delves and C. A. Hall [77] developed a method called global element method, which is actually a DG method. The advantage of the method is that the linear system arising from the method is symmetric. However, it is not guaranteed to be positive semi-definite. In order to overcome the disadvantage, J. Douglas Jr. and T. Dupont [128], M. F. Wheeler [200] and D. N. Arnold [7] added some penalty terms which are some weighted L^2 inner products of the jumps in the function values across element edges. They also proved the continuity and coercivity of the penalty variational form and the optimal convergence rate. In 1998, J. T. Oden, I. Babuska and C. E. Baumann [10, 169] introduced a nonsymmetric DG method for the diffusion problem. After that, B. Riviere, M. F. Wheeler and V. Girault [183] added the same penalty terms

as above to the method. In 2000, T. J. R. Hughes, G. Engel, L. Mazzei and M. G. Larson [122] did many numerical experiments to analyze these interior penalty DG method. Other applications and analysis of the interior penalty DG method can be found in [100, 164] and reference therein.

Since the interior penalty DG method was successfully implemented to solve general elliptic problems, it is natural to extend it to elliptic interface problems. Because DG methods don't require the inter-element continuity of functions, they allow more flexible meshes than those permitted by conventional finite element methods. In a word, it is more efficient to implement local mesh refinement in DG methods. Additionally, because the IFE method allows the elements to be cut through by the interfaces, a structured mesh independent of interface, such as a Cartesian mesh, can be used for solving interface problems. Therefore, the combination of DG methods and IFE allows adaptive structured mesh to be used for solving interface problems. That is, a structured mesh can be refined wherever needed, such as around the interface and the singular source. In this chapter, we will discuss the immersed discontinuous Galerkin(IDG) method which combines the interior penalty DG method with IFE.

7.1 IDG method with bilinear IFE

In this section, we will introduce the IDG method with bilinear IFE. First we recall a well known discontinuous weak formulation from [51] with a slight modification in the definition of the involved space. Second, we follow [7, 10, 51, 77, 128, 169, 183, 200] to introduce the symmetric and nonsymmetric discontinuous weak formulations. Third, we introduce a bilinear IFE space and apply it to the discontinuous weak formulations. Finally, we compare the symmetric and nonsymmetric formulations. Even though the set-up in this section is well known [51], we will repeat it here in order to introduce the notations that will be used.

7.1.1 A well known discontinuous weak formulation

In this section, we will recall a well known discontinuous weak formulation [51]. Assume the solution u of the model interface problem (1.1)-(1.4) is in $PH^2_{int}(\Omega)$. For a given mesh $\mathcal{T}_h$ of Ω, a discontinuous weak formulation is formed on

$$\begin{aligned} PH^1(\mathcal{T}_h) \;=\; & \{v : \forall T \in \mathcal{T}_h, v|_T \in C(T); \\ & v|_T \in H^1(T) \text{ if } T \in \mathcal{T}_h \text{ is a non-interface element}; \\ & v|_{\tilde{T}^s} \in H^1(\tilde{T}^s) \text{ if } T \in \mathcal{T}_h \text{ is an interface element }, s = +, -\} \end{aligned}$$

as follows [51].

First, multiplying (1.1) by $v \in PH^1(\mathcal{T}_h)$ and integrating over each element $T \in \mathcal{T}_h$, we have

$$-(\nabla \cdot \big(\beta \nabla u\big), v)_T \quad = \quad (f, v)_T. \tag{7.1}$$

Applying Green's formula to the first term of (7.1), we have

$$-(\beta \nabla u \cdot \mathbf{n}, v)_{\partial T} + (\beta \nabla u, \nabla v)_T \quad = \quad (f, v)_T, \tag{7.2}$$

where $\mathbf{n}$ is the unit outer normal vector of ∂T. Then we sum (7.2) over $T \in \mathcal{T}_h$ to obtain

$$-\sum_{T\in\mathcal{T}_h} (\beta \nabla u \cdot \mathbf{n}, v)_{\partial T} + \sum_{T\in\mathcal{T}_h} (\beta \nabla u, \nabla v)_T \quad = \quad \sum_{T\in\mathcal{T}_h} (f, v)_T. \tag{7.3}$$

Now we analyze the first term in (7.3). First, let ε_h^D denote the edges of mesh $\mathcal{T}_h$ on the Dirichlet boundary of the problem, ε_h^o denote the interior edges of mesh $\mathcal{T}_h$ and $\varepsilon_h = \varepsilon_h^D \bigcup \varepsilon_h^o$. Second, $\forall e \in \varepsilon_h^o$, let T_1, T_2 be the two elements in $\mathcal{T}_h$ such that $e = \partial T_1 \cap \partial T_2$. For a vertical edge e, we define the element on its left side to be T_1 and for a horizontal edge e, we define the element below it to be T_1. Let ν be the unit normal vector of e exterior to T_2, see Figure 7.1. Finally, $\forall e \in \varepsilon_h^D$, let ν be the unit normal vector of e exterior to Ω. Then the first term in (7.3) can be rewritten as

$$\begin{aligned}
&\sum_{T\in\mathcal{T}_h} (\beta \nabla u \cdot \mathbf{n}, v)_{\partial T} \\
= \quad &\sum_{e\in\varepsilon_h^D} (\beta \nabla u \cdot \mathbf{n}, v)_e + \sum_{e\in\varepsilon_h^o} [(\beta \nabla u \cdot \mathbf{n}, v)_{\partial T_1 \cap e} + (\beta \nabla u \cdot \mathbf{n}, v)_{\partial T_2 \cap e}] \\
= \quad &\sum_{e\in\varepsilon_h^D} (\beta \nabla u \cdot \nu, v)_e + \sum_{e\in\varepsilon_h^o} [(\beta \nabla u \cdot \nu, v)_{\partial T_2 \cap e} - (\beta \nabla u \cdot \nu, v)_{\partial T_1 \cap e}].
\end{aligned} \tag{7.4}$$

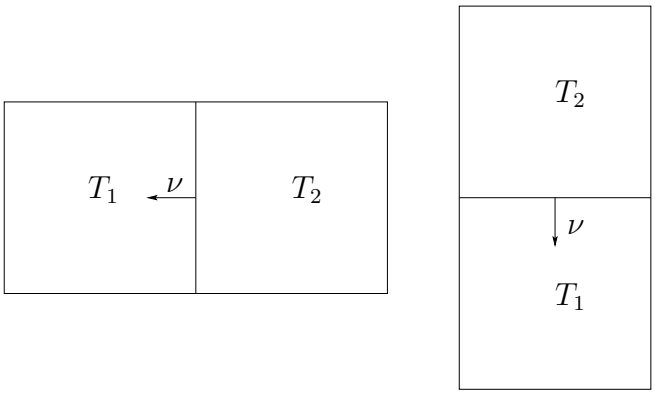

Figure 7.1: A sketch of T_1, T_2 and ν.

Now we use the jump and average terms to rewrite (7.4). First, for a function v, we define its jump and average on $e \in \varepsilon_h^o$ as follows.

$$\begin{aligned}
[v] \quad &= \quad (v|_{T_2})|_e - (v|_{T_1})|_e, \\
\{v\} \quad &= \quad \frac{1}{2}((v|_{T_2})|_e + (v|_{T_1})|_e).
\end{aligned}$$

$\forall e \in \varepsilon_h^D$, the jump and average functions are just the function itself. Second, by using the algebraic identity

$$ab - cd = \frac{1}{2}(a+c)(b-d) + \frac{1}{2}(a-c)(b+d),$$

$\forall e \in \varepsilon_h^o$, we can get

$$(\beta\nabla u \cdot \nu)|_{\partial T_2 \cap e} v|_{\partial T_2 \cap e} - (\beta\nabla u \cdot \nu)|_{\partial T_1 \cap e} v|_{\partial T_1 \cap e} = \{\beta\nabla u \cdot \nu\}[v] + [\beta\nabla u \cdot \nu]\{v\}.$$

Therefore, we have

$$\begin{aligned} & \sum_{e \in \varepsilon_h^o} [(\beta\nabla u \cdot \nu, v)_{\partial T_2 \cap e} - (\beta\nabla u \cdot \nu, v)_{\partial T_1 \cap e}] \\ = & \sum_{e \in \varepsilon_h^o} [(\{\beta\nabla u \cdot \nu\}, [v])_e + ([\beta\nabla u \cdot \nu], \{v\})_e]. \end{aligned} \tag{7.5}$$

In addition, if the fluxes $\beta\nabla u \cdot \nu$ are continuous almost everywhere in Ω, then we have

$$\sum_{e \in \varepsilon_h^o} ([\beta\nabla u \cdot \nu], \{v\})_e = 0. \tag{7.6}$$

Finally, using (7.3), (7.4), (7.5) and (7.6), we get

$$\sum_{T \in \mathcal{T}_h} (\beta\nabla u, \nabla v)_T - \sum_{e \in \varepsilon_h^o} (\{\beta\nabla u \cdot \nu\}, [v])_e - \sum_{e \in \varepsilon_h^D} (\beta\nabla u \cdot \nu, v)_e = \sum_{T \in \mathcal{T}_h} (f, v)_T.$$

Define

$$\begin{aligned} b(u,v) &= \sum_{T \in \mathcal{T}_h} (\beta\nabla u, \nabla v)_T, \\ J(u,v) &= \sum_{e \in \varepsilon_h^o} (\{\beta\nabla u \cdot \nu\}, [v])_e + \sum_{e \in \varepsilon_h^D} (\beta\nabla u \cdot \nu, v)_e, \\ a(u,v) &= b(u,v) - J(u,v), \\ L(v) &= \sum_{T \in \mathcal{T}_h} (f, v)_T. \end{aligned}$$

Then a discontinuous weak formulation of (1.1) is to find a $u \in PH^2_{int}(\Omega)$ such that

$$a(u,v) = L(v), \quad \forall v \in PH^1(\mathcal{T}_h). \tag{7.7}$$

7.1.2 The symmetric and nonsymmetric discontinuous weak formulations

Note that the bilinear form of the above weak formulation is nonsymmetric, hence M. Delves and C. A. Hall [77] introduced a symmetric discontinuous weak formulation as follows. Consider

$$J(v,u) = \sum_{e\in\varepsilon_h^o}(\{\beta\nabla v\cdot\nu\},[u])_e + \sum_{e\in\varepsilon_h^D}(\beta\nabla v\cdot\nu,u)_e.$$

If u is continuous almost everywhere in Ω, then $\sum_{e\in\varepsilon_h^o}(\{\beta\nabla v\cdot\nu\},[u])_e = 0$. Additionally, we have $u = g$ on ε_h^D, so we get

$$J(v,u) = \sum_{e\in\varepsilon_h^D}(\beta\nabla v\cdot\nu,g)_e.$$

Define

$$\begin{aligned} a^-(u,v) &= b(u,v) - J(u,v) - J(v,u), \\ L^-(v) &= L(v) - \sum_{e\in\varepsilon_h^D}(\beta\nabla v\cdot\nu,g)_e. \end{aligned}$$

Then, the weak formulation is to find a $u\in PH^2_{int}(\Omega)$ such that

$$a^-(u,v) = L^-(v), \quad \forall v\in PH^1(\mathcal{T}_h).$$

Here $a^-(\cdot,\cdot)$ is symmetric. However, the linear system of algebraic equations arising from this weak form is not guaranteed to be semi-definite. Hence J. Douglas Jr. and T. Dupont [128], M. F. Wheeler [200] and D. N. Arnold [7] added a penalty to the weak form as follows. Consider the following penalty term

$$J_\theta(u,v) = \sum_{e\in\varepsilon_h^o}\theta_e([u],[v])_e + \sum_{e\in\varepsilon_h^D}\theta_e(u,v)_e,$$

where θ_e is the penalty parameter. If u is continuous almost everywhere in Ω, then we can get

$$J_\theta(u,v) = \sum_{e\in\varepsilon_h^D}\theta_e(g,v)_e.$$

Define

$$\begin{aligned} a_\theta^-(u,v) &= b(u,v) - J(u,v) - J(v,u) + J_\theta(u,v), \\ L_\theta^-(v) &= L(v) - \sum_{e\in\varepsilon_h^D}(\beta\nabla v\cdot\nu,g)_e + \sum_{e\in\varepsilon_h^D}\theta_e(g,v)_e. \end{aligned}$$

Then, the symmetric discontinuous weak formulation is to find a $u \in PH^2_{int}(\Omega)$ such that

$$a^-_\theta(u,v) = L^-_\theta(v), \quad \forall v \in PH^1(\mathcal{T}_h). \tag{7.8}$$

Now we recall the work of J. T. Oden, I. Babuska, C. E. Baumann [10, 169] and B. Riviere, M. F. Wheeler, V. Girault [183] to generate a nonsymmetric discontinuous weak formulation. Define

$$\begin{aligned} a^+_\theta(u,v) &= b(u,v) - J(u,v) + J(v,u) + J_\theta(u,v), \\ L^+_\theta(v) &= L(v) + \sum_{e\in\varepsilon_h^D} (\beta\nabla v\cdot\nu, g)_e + \sum_{e\in\varepsilon_h^D} \theta_e(g,v)_e. \end{aligned}$$

Then, the nonsymmetric discontinuous weak formulation is to find a $u \in PH^2_{int}(\Omega)$ such that

$$a^+_\theta(u,v) = L^+_\theta(v), \quad \forall v \in PH^1(\mathcal{T}_h). \tag{7.9}$$

7.1.3 Bilinear immersed discontinuous Galerkin formulations

By using the local bilinear immersed finite element basis functions formed in Chapter 3, we construct a new bilinear immersed finite element(IFE) space $S_{h,D}(\Omega)$. We note that for each element T in a mesh $\mathcal{T}_h$, we have four nodal basis functions ϕ_i, $i = 1,2,3,4$. For a non-interface element T, they are the standard bilinear nodal basis functions; otherwise, they are the immersed bilinear basis functions. Then, we define a piecewise bilinear global nodal basis function $\psi_N(x,y)$ from each ϕ_i such that $\psi_N|_T = \phi_i$ and ψ_N is zero everywhere else. Finally, we define $S_{h,D}(\Omega)$ as the span of all these global nodal basis functions.

Now we use the discontinuous bilinear IFE space $S_{h,D}(\Omega) \subset PH^1(\mathcal{T}_h)$ and the weak formulations mentioned above to construct the IDG method as follows: find $u_h \in S_{h,D}(\Omega)$ such that

$$a^-_\theta(u_h,v_h) = L^-_\theta(v_h), \quad \forall v_h \in S_{h,D}(\Omega), \tag{7.10}$$

or find $u_h \in S_{h,D}(\Omega)$ such that

$$a^+_\theta(u_h,v_h) = L^+_\theta(v_h), \quad \forall v_h \in S_{h,D}(\Omega). \tag{7.11}$$

Remark 7.1.1 *The IDG method is actually a special case of the selective immersed DG method which will be discussed in Chapter 8, see Remark 8.1.2. Therefore, the convergence conclusion in Section 8.4 is also true for the IDG method.*

7.1.4 Comparison of the symmetric and nonsymmetric formulations

In this section, we will compare the symmetric and nonsymmetric formulations for their sensitivity to the penalty. According to [7, 51, 183], in order to obtain the stability and the optimal error estimate for the finite element solution of the interior penalty DG method, we need to choose the penalties as follows.

For the symmetric formulation, we should choose

$$\theta_e = \frac{C_*}{h}, \tag{7.12}$$

where C_* is a constant such that $C_* > C_0$ for some positive constant C_0. On the other hand, for the nonsymmetric formulation, we should choose

$$\theta_e = \frac{C_{**}}{h}, \tag{7.13}$$

where C_{**} is any positive constant.

We can see that the symmetric formulation leads to a symmetric system, but its penalty depends on a large enough positive number C_*. Therefore, the stability and convergence rates depend on the unknown penalty constant C_*, and it is not a trivial problem to choose a proper penalty parameter. On the other hand, the nonsymmetric formulation leads to a nonsymmetric system, but we can choose any penalty in the form (7.13) as long as $C_{**} > 0$. Therefore, the choice of penalty is straightforward. We emphasize that a small penalty may not work for symmetric formulation, but for nonsymmetric formulation. For the IDG method, we will observe the similar property.

7.2 Adaptive mesh

In this section, we will first point out a limitation of Galerkin finite element method for using adaptive Cartesian mesh. Then we will explain how the combination of the IDG method and bilinear IFE allows us to form an efficient numerical method with adaptive Cartesian mesh for solving interface problems.

In a Galerkin method based on Lagrange type finite elements, the involved finite element space can be described by the global nodal basis functions. These global basis functions usually are often required to have certain continuity at each node in the mesh. When a global basis function is restricted to an element, it is either the zero function or becomes one of the local basis function in that element. If we refine the mesh, new nodes usually need to be introduced in order to maintain this continuity. A new node becomes a so called hanging node if it is not a vertex of an element but it is on an edge of that element in the new mesh.

At a hanging node, the global basis function of the original type usually cannot be defined and the original Galerkin finite element method cannot be continued over the new mesh.

For example, let us consider a mesh formed by the four large rectangles in Figure 7.2. If we refine this mesh by cutting the lower left rectangular element into four small rectangles, then hanging nodes will be introduced at A and C. Assume that we want to use the bilinear finite elements. We note that node C belongs to three elements in the refined mesh. The global bilinear basis function corresponding to node C can be defined by local nodal basis functions in elements $\square ABCD$ and $\square DCEF$, but not in element $\square BGHE$. Hence, when a refined mesh is introduced, we need to make sure that there are no hanging nodes. However, this restriction usually decreases the efficiency because we have to use sufficient number of new nodes in the refined mesh so that none of new nodes becomes a hanging node and this will also increase the number of elements. For example, consider the mesh formed by the large rectangles in Figure 7.3. Assume that we want to refine element 1 because it contains the interface. To avoid hanging nodes, we have to refine elements 1, 2, 4, or element 1, 3, 5, or elements 1, 2, 3, 4, and 5 even though we only want to refine element 1.

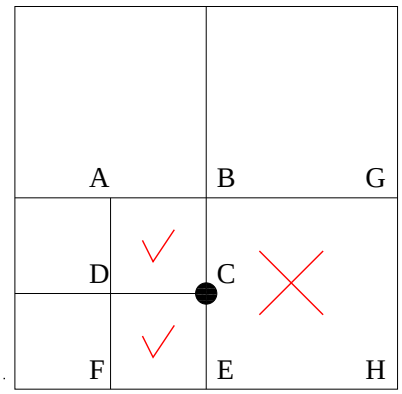

Figure 7.2: A sketch of a hanging node. The local basis functions are defined on the two elements on the left side, but not on the element on the right side. Therefore, the global basis function is not defined at the hanging node.

However, DG methods do not have this drawback because they don't require the inter-element continuity. According to the construction of the space $S_{h,D}(\Omega)$, we can see that each of the global basis function only consists of one local basis in one element and is zero zero everywhere else. Hence, for a hanging node, we can find the element in which this hanging node is a vertex and use the corresponding local nodal basis to introduce a global basis at this hanging node. For example, for the hanging node C in Figure 7.2, we only define a global basis from the local basis at C in the element $\square ABCD$ and another global basis from the local basis at C in the element $\square DCEF$. Because there is no local basis at the C in the element $\square BGHE$, no global basis is introduced at C with respect to this element. This is why a DG method does not have the drawback of hanging nodes and can allow us to employ the mesh refinement locally at the places needed.

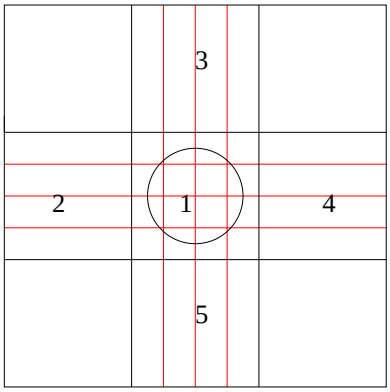

Figure 7.3: A sketch of an adaptive mesh allowed by standard Galerkin method.

Meanwhile, IFE allows us to use structured mesh for interface problems, so the IDG method allows us to use locally adaptive structured meshes for solving interface problems with non-trivial interfaces. That is, in a structured mesh, we can refine any region again and again while keeping the mesh of the rest region coarse. In particular, we can repeatedly refine only the interface elements along the interface, see Figure 7.4. This is very important for improving the efficiency of the algorithm. In addition, fast algebraic solvers, such as multigrid, can be easily applied because the meshes are structured.

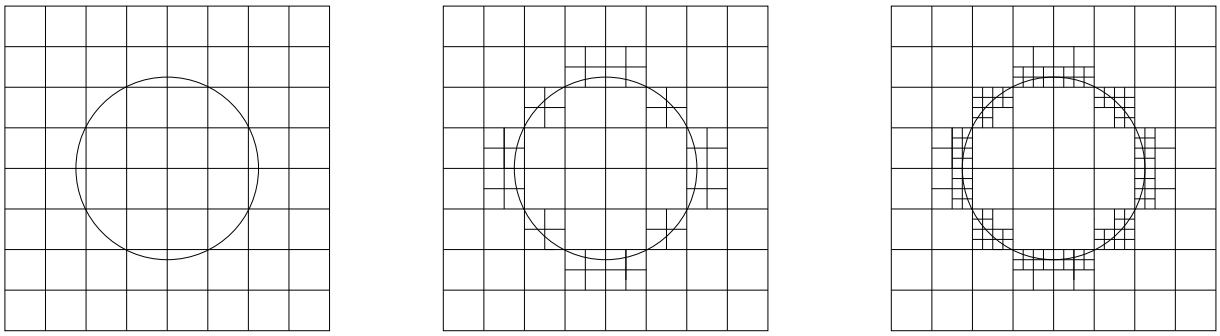

Figure 7.4: The left plot is the uniform mesh without refinement. In the middle plot, each interface element in the left mesh is refined into four congruent elements. In the right plot, each interface element in the middle graph is refined into four congruent elements.

7.3 Implementation of Gauss quadratures

In a DG method, we need to compute some integrals on the elements and the element edges, which involve the finite element basis functions and their derivatives. Usually we use Gauss quadratures, which require the continuity of the integrands, to compute these integrals. For non-interface elements, we can treat the corresponding integrals as usual. However, on the

interface elements, the derivative of the local bilinear IFE basis functions are discontinuous. In addition, on the edges of interface elements, the local bilinear IFE basis functions and their derivatives may also be discontinuous. Therefore, we need to treat the corresponding integrals carefully. In the following, we will discuss both area and line integrals separately.

(1) The integrals on interface elements:

Here we only discuss the Type I interface elements. The discussion for the Type II interface elements can be carried out similarly. Suppose we have a Type I interface element $\square ABCF$. We first connect the two end points D and E of the interface in the element by the segment $\overline{DE}$, see Figure 7.5. By using this segment and two other segments $\overline{DC}$, $\overline{EC}$, we can divide the interface element into four triangles $\triangle ADE$, $\triangle ECB$, $\triangle DFC$ and $\triangle CDE$. In each triangle, we specify one piece of the piecewise bilinear local IFE basis. Since each piece of the local bilinear IFE basis is continuous, the integrands are continuous on each triangle. Then we can apply Gauss quadratures to each triangle one by one. Finally, we add the four integrals on the four triangles together to get the integral on the whole interface element.

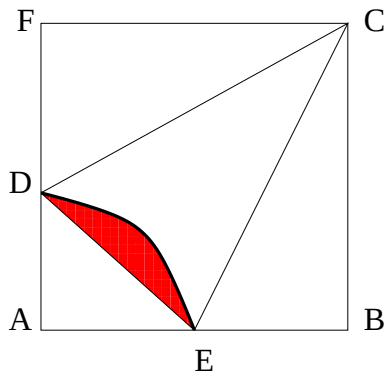

Figure 7.5: A sketch of the division of an interface element into 4 triangles.

(2) The integrals on the element edges which intersect the interface:

Suppose we have an element edge $\overline{AC}$ intersecting the interface at B. We first use the intersection point B to separate the edge $\overline{AC}$ into 2 segments $\overline{AB}$ and $\overline{BC}$, see Figure 7.6. On each segment, we can specify one piece of the local bilinear local IFE basis. Because each piece of the local bilinear IFE basis is continuous, the integrands are continuous on each of $\overline{AB}$ and $\overline{BC}$. Then we can apply the 1D Gauss quadratures them one by one. Finally, we add both of the two integrals on $\overline{AB}$ and $\overline{BC}$ together to get the integral on the whole element edge $\overline{AC}$.

Another implementation issue is about the line integrals on the interior edges for stiffness matrix. Consider an interior edge which has two neighboring elements. Each neighboring element has four local basis functions and two of them correspond to the two end points of this edge. Therefore, the two elements have eight local basis functions and four of them

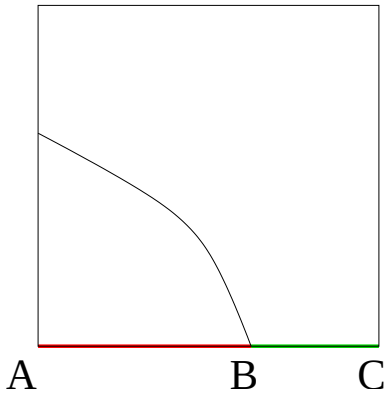

Figure 7.6: A sketch of the division of an interface edge into 2 segments AB and BC.

correspond to the two end points. Not only the four local basis functions corresponding to the two end points and their normal derivatives are nonzero on the edge, but also all the eight local basis functions and their normal derivatives may be nonzero on this edge. That is, if we arbitrarily choose one of the eight local basis functions as a trial function and one of the eight local basis functions as a test function, the line integral on this edge may be nonzero. Therefore, when we assemble the stiffness matrix, we need to compute the line integrals on this edge for all possible combinations of a trial function and a test function.

7.4 Numerical examples

In this section, we will present some numerical examples for the IDG method with the bilinear IFE space $S_{h,D}$. Since large error usually appears in interface elements, we will refine only the interface elements to construct the adaptive meshes as follows. First we construct a Cartesian mesh without mesh refinement, see the left plot of the Figure 7.4. Second we only refine the interface elements by using half of the step size of the original mesh, i.e., we refine each interface element in the left mesh into four congruent elements. Then we get the first-level refined mesh, see the middle plot of the Figure 7.4. Third, we refine the interface elements in the first-level refined mesh by using $\frac{1}{4}$ of the step size of the original mesh, i.e., we refine each interface element in the middle mesh into four congruent elements. Then we get the second-level refined mesh, see the right plot of Figure 7.4. Finally, if we continue this process, we can get n^{th}-level refined mesh.

We consider the same example as in Section 5.2 with $\beta^- = 1$ and $\beta^+ = 10$. In the following tables, $\|\cdot\|_0$ represents the usual L^2 norm, $|\cdot|_1$ is the usual semi-H^1 norm, and of course, they are computed numerically according to the mesh used. The quantity $\|\cdot\|_\infty$ is the discrete infinity norm which is the maximum of the absolute values of the given function at all the nodes of a mesh. We call the constant C_* in (8.2) and C_{**} in (7.13) as penalty constants.

7.4.1 Numerical results for the symmetric IDG method with bilinear IFE

In order to illustrate the convergence, we will show the numerical errors in L^2, H^1 and discrete infinity norm for the symmetric IDG method on the original rectangular mesh without mesh refinement, see the left plot of the Figure 7.4. Table 7.1 contains the errors of the solutions u_h with various partition sizes h and the penalty constant $C_* = 1000$.

h	$\|u_h - u\|_0$	$\|u_h - u\|_1$	$\|u_h - u\|_\infty$
1/16	9.2701×10^{-4}	4.6957×10^{-2}	1.9307×10^{-3}
1/32	2.2250×10^{-4}	2.2329×10^{-2}	5.0574×10^{-4}
1/64	5.8475×10^{-5}	1.0566×10^{-2}	1.3518×10^{-4}
1/128	1.4442×10^{-5}	5.2157×10^{-3}	3.4613×10^{-5}
1/256	3.6768×10^{-6}	2.6167×10^{-3}	8.6282×10^{-6}

Table 7.1: Errors of the symmetric IDG solutions on the original mesh with C_*=1000.

Using linear regression, we can also see that the data in Table 7.1 obey

$$\|u_h - u\|_0 \approx 0.2269\ h^{1.9901},\ \ |u_h - u|_1 \approx 0.8310\ h^{1.0429},\ \ \|u_h - u\|_\infty \approx 0.4342\ h^{1.9481},$$

which indicates that the IDG solution u_h converges to the exact solution with convergence rates $O(h^2)$ in the L^2 norm, $O(h)$ in the H^1 norm and $O(h^2)$ in the discrete infinity norm.

For the global effect of the local mesh refinement, we will compare the numerical errors in L^2, H^1 and discrete infinity norms on different meshes with the step sizes $h = \frac{1}{8}, \frac{1}{16}, \frac{1}{32}$ and $C_* = 1000000$. Table 7.2 to 7.4 contain numerical results generated on the original mesh and the refined meshes from the first-level to the fourth-level. Table 7.2 contains the L^2 norm errors. Table 7.3 contains the H^1 norm errors. Table 7.4 contains the discrete infinity norm errors. Note that the h is the the step size of the corresponding original mesh. From the three tables, we can see that local refinement of the first-level and second-level refined meshes dramatically reduces the global errors. Because the error in the non-interface area basically stays the same and becomes more and more dominant during the local refinement in interface elements, the third-level and fourth-level refined meshes don't reduce the global error much.

Finally, we would like to compare the discrete infinity norm errors on all interface elements in this section to see the local effect of the local mesh refinement. Table 7.5 contains the discrete infinity norm errors of the solutions u_h on different meshes with the step sizes $h = \frac{1}{8}, \frac{1}{16}, \frac{1}{32}$ and the penalty constant $C_* = 1000000$. When we refine the interface elements, the error around interface decreases quickly. Therefore, the combination of the DG method and IFE generates a numerical method that can efficiently control the error across the interface where interesting physics happen in many applications.

h	$\frac{1}{8}$	$\frac{1}{16}$	$\frac{1}{32}$
original mesh	2.1357×10^{-2}	4.7946×10^{-3}	7.2152×10^{-4}
first-level adaptive mesh	5.5170×10^{-3}	1.0046×10^{-3}	1.7156×10^{-4}
second-level adaptive mesh	2.2593×10^{-3}	5.8151×10^{-4}	1.5645×10^{-4}
third-level adaptive mesh	1.9584×10^{-3}	5.6831×10^{-4}	1.5215×10^{-4}
fourth-level adaptive mesh	1.9469×10^{-3}	5.6389×10^{-4}	1.5117×10^{-4}

Table 7.2: L^2 norm Errors of the symmetric IDG solutions on different meshes with C_*=1000000.

h	$\frac{1}{8}$	$\frac{1}{16}$	$\frac{1}{32}$
original mesh	1.7407×10^{-1}	7.2427×10^{-2}	2.7358×10^{-2}
first-level adaptive mesh	8.7645×10^{-2}	4.1428×10^{-2}	1.9983×10^{-2}
second-level adaptive mesh	6.7158×10^{-2}	3.7169×10^{-2}	1.9738×10^{-2}
third-level adaptive mesh	6.5053×10^{-2}	3.7000×10^{-2}	1.9663×10^{-2}
fourth-level adaptive mesh	6.4952×10^{-2}	3.6939×10^{-2}	1.9650×10^{-2}

Table 7.3: H^1 norm Errors of the symmetric IDG solutions on different meshes with C_*=1000000.

h	$\frac{1}{8}$	$\frac{1}{16}$	$\frac{1}{32}$
original mesh	3.0811×10^{-2}	8.0046×10^{-3}	3.2532×10^{-3}
first-level adaptive mesh	1.0102×10^{-2}	3.0655×10^{-3}	4.8067×10^{-4}
second-level adaptive mesh	5.2750×10^{-3}	1.4920×10^{-3}	4.2755×10^{-4}
third-level adaptive mesh	4.6491×10^{-3}	1.4413×10^{-3}	3.7831×10^{-4}
fourth-level adaptive mesh	4.6198×10^{-3}	1.4123×10^{-3}	3.7746×10^{-4}

Table 7.4: Discrete infinity norm Errors of the symmetric IDG solutions on different meshes with C_*=1000000.

h	$\frac{1}{8}$	$\frac{1}{16}$	$\frac{1}{32}$
uniform mesh	2.7755×10^{-2}	8.0046×10^{-3}	3.2532×10^{-3}
first-level adaptive mesh	7.3569×10^{-3}	3.0655×10^{-3}	3.2676×10^{-4}
second-level adaptive mesh	2.7734×10^{-3}	6.3936×10^{-4}	1.7824×10^{-4}
third-level adaptive mesh	3.9052×10^{-4}	2.0124×10^{-4}	8.3156×10^{-5}
fourth-level adaptive mesh	2.4619×10^{-4}	9.9134×10^{-5}	2.8141×10^{-5}

Table 7.5: Comparison of the discrete infinity norm errors of the symmetric IDG solutions on interface elements with C_*=1000000.

7.4.2 Numerical results for the nonsymmetric IDG method with bilinear IFE

Since the penalty constant C_{**} of the non-symmetric DG formulation doesn't depend on the problem, we would like to first show its convergence numerically in this section. We will compare the sensitivity of the symmetric and non-symmetric formulation in the next section. Table 7.6 contains the errors of the solutions u_h on the original rectangular mesh without mesh refinement with various partition sizes h and the penalty constant $C_{**} = 1000$.

h	$\|u_h - u\|_0$	$\|u_h - u\|_1$	$\|u_h - u\|_\infty$
1/16	9.0251×10^{-4}	4.6957×10^{-2}	1.9093×10^{-3}
1/32	2.1617×10^{-4}	2.2328×10^{-2}	5.0047×10^{-4}
1/64	5.6864×10^{-5}	1.0566×10^{-2}	1.3394×10^{-4}
1/128	1.4036×10^{-5}	5.2157×10^{-3}	3.4292×10^{-5}
1/256	3.5747×10^{-6}	2.6167×10^{-3}	8.5475×10^{-6}

Table 7.6: Errors of the nonsymmetric IDG solutions on the original mesh with C_{**}=1000.

Using linear regression, we can also see that the data in Table 7.6 obey

$$\|u_h - u\|_0 \approx 0.2210\ h^{1.9905},\ \ |u_h - u|_1 \approx 0.8309\ h^{1.0429},\ \ \|u_h - u\|_\infty \approx 0.4287\ h^{1.9474},$$

which indicates that the bilinear IDG solution u_h converges to the exact solution with convergence rates $O(h^2)$ in the L^2 norm, $O(h)$ in the H^1 norm and $O(h^2)$ in the discrete infinity norm.

Similar to the symmetric IDG method, we would like to compare the discrete infinity norm error on all interface elements to see how the local mesh refinement reduces the errors around the interface. Table 7.7 contains the discrete infinity norm errors of the solutions u_h with the step sizes $h = \frac{1}{8}, \frac{1}{16}, \frac{1}{32}$ and the penalty constant $C_{**} = 1000000$. From this table, we can see that the error decreases quickly during the local mesh refinement.

h	$\frac{1}{8}$	$\frac{1}{16}$	$\frac{1}{32}$
uniform mesh	2.7752×10^{-2}	8.0042×10^{-3}	3.2527×10^{-3}
first-level adaptive mesh	7.3559×10^{-3}	3.0647×10^{-3}	3.2675×10^{-4}
second-level adaptive mesh	2.7720×10^{-3}	6.3935×10^{-4}	1.7825×10^{-4}
third-level adaptive mesh	3.9047×10^{-4}	2.0124×10^{-4}	8.3093×10^{-5}
fourth-level adaptive mesh	2.4621×10^{-4}	9.9122×10^{-5}	2.8147×10^{-5}

Table 7.7: Comparison of the discrete infinity norm errors of the nonsymmetric IDG solutions on interface elements with C_{**}=1000000.

7.4.3 The effect of the penalty parameters on symmetric and nonsymmetric IDG formulations

From Section 7.1.4, we can see that the symmetric interior penalty DG method requires a large enough positive constant C_* for the penalty while the nonsymmetric interior penalty DG method only requires a positive constant C_{**} for the penalty. While a small penalty parameter works for the nonsymmetric method, it may cause the symmetric method to fail. Actually, we observe the similar performance for the IDG method. In this section, we will use two examples to show this phenomenon.

For the numerical results in Table 7.8 and Table 7.9, we use the original rectangular mesh without refinement. Table 7.8 contains the errors of the solutions u_h with various partition sizes h and the penalty constant $C_* = 1$ for the symmetric IDG method. Table 7.9 contains the errors of the solutions u_h with various partition sizes h and the penalty constant $C_{**} = 1$ for the nonsymmetric IDG method.

h	$\|u_h - u\|_0$	$\|u_h - u\|_1$	$\|u_h - u\|_\infty$
1/16	6.2958×10^{-2}	1.6052×10^{0}	1.1989×10^{-1}
1/32	4.3967×10^{-3}	2.8605×10^{-1}	1.1677×10^{-2}
1/64	7.3712×10^{-4}	1.1320×10^{-1}	3.2727×10^{-3}
1/128	1.3069×10^{-3}	3.9426×10^{-1}	5.6696×10^{-3}
1/256	1.5683×10^{-4}	9.6793×10^{-2}	5.1739×10^{-4}

Table 7.8: Errors of the nonsymmetric IDG solutions on the original mesh with C_*=1.

h	$\|u_h - u\|_0$	$\|u_h - u\|_1$	$\|u_h - u\|_\infty$
1/16	2.2230×10^{-3}	4.8928×10^{-2}	4.4368×10^{-3}
1/32	5.0232×10^{-4}	2.2835×10^{-2}	1.1160×10^{-3}
1/64	1.1751×10^{-4}	1.0707×10^{-2}	2.7990×10^{-4}
1/128	2.8279×10^{-5}	5.2514×10^{-3}	7.0089×10^{-5}
1/256	6.9225×10^{-6}	2.6257×10^{-3}	1.7537×10^{-5}

Table 7.9: Errors of the nonsymmetric IDG solutions on the original mesh with C_{**}=1.

The the data in Table 7.8 shows that the IDG solution u_h for symmetric formulation does not converge with the rate expected. However, using linear regression, we can also see that the data in Table 7.9 obey

$$\|u_h - u\|_0 \approx 0.6913\ h^{2.0805},\ \ |u_h - u|_1 \approx 0.8930\ h^{1.0560},\ \ |u_h - u|_\infty \approx 1.1252\ h^{1.9959},$$

which indicates that the nonsymmetric IDG solution u_h with bilinear IFE converges to the exact solution with convergence rates $O(h^2)$ in the L^2 norm, $O(h)$ in the H^1 norm and $O(h^2)$

in the discrete infinity norm. For all the linear regressions in this chapter, we obtain similar figures to Figure 5.1, which mean that the data points match the linear regression lines very well.

Chapter 8

Selective immersed discontinuous Galerkin (SIDG) method

In Chapter 7, we introduced the immersed DG method. One important advantage of this method is the local mesh refinement, which allows us to use finer meshes locally for some interesting parts of a problem domain. However, the immersed DG method is based on the interior penalty DG method [7, 10, 122, 128, 169, 183, 200], which needs much more finite element basis functions than Galerkin method on the same mesh. Hence much larger algebraic systems need to be solved by significantly more computational cost. This contradiction leads to a challenge about how to take advantage of the local refinement feature of the DG methods while keeping the computational cost as close to that of Galerkin method as possible. Coupling [60, 97, 173] or hybridization [61, 62, 63] of DG methods and conforming finite element methods is one efficient way to resolve this challenge. Following this idea, in this chapter, we will introduce a selective immersed discontinuous Galerkin(SIDG) method, which nicely resolves this challenge. Its basic idea is to use the discontinuous Galerkin formulation wherever local refinement is needed, such as around an interface or a singular source, but the regular Galerkin formulation everywhere else.

8.1 Formulations of the SIDG method

In this section, we will use our model interface problem to introduce the SIDG method. The same idea can be extended to other problems.

First we introduce some conventions. Assume $\mathcal{T}_h, h > 0$ is a mesh of the solution domain Ω. The mesh consists of interface elements whose interiors are cut through by the interfaces and the rest called non-interface elements. Let ε_S denote the collection of selected element edges where discontinuity will be allowed by the discontinuous Galerkin formulation, and let Ω_S the union of the elements in $\mathcal{T}_h$ having at least one of its edges in ε_S. The key

point of the SIDG method is to properly select these two sets based on the features of the problem to be solved. For our model interface problem, because of the discontinuity of the coefficient, the error around interface is usually much larger than the error on the rest of the domain. In order to apply local mesh refinement around the interface to improve the accuracy efficiently, one of our choice of Ω_S is the union of all interface elements and ε_S to be the set of all the edges of the elements in Ω_S. Let ε_h^D denote the element edges of $\mathcal{T}_h$ on the Dirichlet boundary of the problem and ε_h^o denote the interior element edges of $\mathcal{T}_h$. We first introduce the following space for the weak formulation:

$$\begin{aligned} PH_S^1(\mathcal{T}_h) &= \{v : \forall e \in \varepsilon_h^o/\varepsilon_S,\ [v]|_e = 0;\ \forall T \in \mathcal{T}_h, v|_T \in C(T); \\ &\quad v|_T \in H^1(T) \text{ if } T \in \mathcal{T}_h \text{ is a non-interface element}; \\ &\quad v|_{\tilde{T}^s} \in H^1(\tilde{T}^s) \text{ if } T \in \mathcal{T}_h \text{ is an interface element}, s = +, -\}. \end{aligned}$$

Now we recall the following definition. For a set $\Lambda \subset \Omega$ whose interior is cut through by Γ, we define

$$PH_{int}^2(\Lambda) = \left\{ u \in C(\Lambda), u|_{\Lambda^s} \in H^2(\Lambda^s), s = -, +, \left[\beta \frac{\partial u}{\partial \mathbf{n}_\Gamma}\right] = 0 \ \text{ on } \Gamma \cap \Lambda \right\}.$$

Similar to Section 7.1, we can formulate the selective discontinuous weak formulations for our model interface problem as follows.

$$\sum_{T\in\mathcal{T}_h} (\beta\nabla u, \nabla v)_T - \sum_{e\in\varepsilon_S} (\{\beta\nabla u \cdot \nu\}, [v])_e - \sum_{e\in\varepsilon_h^D} (\beta\nabla u \cdot \nu, v)_e = \sum_{T\in\mathcal{T}_h} (f, v)_T$$

Define

$$\begin{aligned} b(u,v) &= \sum_{T\in\mathcal{T}_h} (\beta\nabla u, \nabla v)_T \\ L(v) &= \sum_{T\in\mathcal{T}_h} (f, v)_T \\ J_S(u,v) &= \sum_{e\in\varepsilon_S} (\{\beta\nabla u \cdot \nu\}, [v])_e + \sum_{e\in\varepsilon_h^D} (\beta\nabla u \cdot \nu, v)_e \\ J_{S\theta}(u,v) &= \sum_{e\in\varepsilon_S} \theta_e([u], [v])_e + \sum_{e\in\varepsilon_h^D} \theta_e(u, v)_e \\ a_{S\theta}^-(u,v) &= b(u,v) - J_S(u,v) - J_S(v,u) + J_{S\theta}(u,v) \\ L^-\theta(v) &= L(v) - \sum_{e\in\varepsilon_h^D} (\beta\nabla v \cdot \nu, g)_e + \sum_{e\in\varepsilon_h^D} \theta_e(g, v)_e \end{aligned}$$

Then, following the idea in [7, 77, 128, 200], the symmetric selective discontinuous weak formulation is to find a $u \in PH_{int}^2(\Omega)$ such that

$$a_{S\theta}^-(u,v) = L_\theta^-(v), \quad \forall v \in PH_S^1(\mathcal{T}_h) \tag{8.1}$$

Based on the convergence analysis in [7, 51, 128, 200] and Section 8.4, we choose

$$\theta_e = \frac{C_*}{h}, \tag{8.2}$$

where C_* is a penalty constant such that $C_* > C_0$ for some positive constant C_0.

Even though the above discontinuous weak formulation is symmetric and positive definite, its penalty constant must be large enough, which means the choice of C_* depends on the problem. In order to remove this limitation, we follow the idea in [10, 169, 183] to generate a nonsymmetric selective discontinuous weak formulation as follows. Find a $u \in PH^2_{int}(\Omega)$ such that

$$a^+_{S\theta}(u,v) = L^+_\theta(v), \quad \forall v \in PH^1_S(\mathcal{T}_h) \tag{8.3}$$

where

$$\begin{aligned} a^+_{S\theta}(u,v) &= b(u,v) - J_S(u,v) + J_S(v,u) + J_{S\theta}(u,v) \\ L^+_\theta(v) &= L(v) + \sum_{e\in\varepsilon_h^D} (\beta\nabla v\cdot\nu, g)_e + \sum_{e\in\varepsilon_h^D} \theta_e (g,v)_e \end{aligned}$$

Based on [10, 169, 51, 183], we usually choose

$$\theta_e = \frac{C_{**}}{h}, \tag{8.4}$$

where C_{**} is a just positive penalty constant.

Now we use a selective immersed finite element space $S^S_h(\Omega) \subset PH^1_S(\mathcal{T}_h)$ and the weak formulations introduced above to construct the SIDG method as follows: find $u_h \in S^S_h(\Omega)$ such that

$$a^-_{S\theta}(u_h, v_h) = L^-_\theta(v_h), \quad \forall v_h \in S^S_h(\Omega), \tag{8.5}$$

or find $u_h \in S^S_h(\Omega)$ such that

$$a^+_{S\theta}(u_h, v_h) = L^+_\theta(v_h), \quad \forall v_h \in S^S_h(\Omega), \tag{8.6}$$

Remark 8.1.1 *The SIDG method depends on two key components: ε_S and $S^S_h(\Omega)$. We choose ε_S and/or Ω_S based on the features of the problem and our desire to use the DG formulation at the places we are interested in. The selective immersed finite element space $S^S_h(\Omega)$ can then be defined according to ε_S, Ω_S and our desire to include other features, such as the need to control the computational cost. As we can see later, a suitable choice of these components can generate a convergent method that can solve an interface problem on a rectangular mesh with the local mesh refinement capability at a reduced computational cost.*

Remark 8.1.2 *If we choose ε_S to be the set of all interior edges, then Ω_S is the union of all the elements in $\mathcal{T}_h$ and the SIDG method becomes the immersed DG method discussed in Chapter 7.*

8.2 SIDG method with bilinear IFE

In this section, we will first use the bilinear IFE functions introduced in [110, 111, 112, 149] to form a selective bilinear immersed finite element space $S_h^{Sb}(\Omega)$. Then we apply this space to the SIDG method and discuss its advantages and implementation.

8.2.1 Selective bilinear immersed finite element space

First, we briefly recall the piecewise bilinear immersed finite element function introduced in Section 3.1. Consider a typical rectangle element $T \in \mathcal{T}_h$. Assume that the four vertices of T are $A_i, i = 1, 2, 3, 4$, with $A_i = (x_i, y_i)^t$. If T is an interface element, then we use $D = (x_{_D}, y_{_D})^T$ and $E = (x_{_E}, y_{_E})^T$ to denote the interface points on its edges. Then we define the piecewise bilinear IFE functions as follows:

$$\phi(x,y) = \begin{cases} \phi^-(x,y) = a^-x + b^-y + c^- + d^-xy, & (x,y) \in \tilde{T}^-, \\ \phi^+(x,y) = a^+x + b^+y + c^+ + d^+xy, & (x,y) \in \tilde{T}^+, \\ \phi^-(D) = \phi^+(D), \quad \phi^-(E) = \phi^+(E), \quad d^- = d^+, & \\ \int_{\overline{DE}} \left(\beta^- \dfrac{\partial \phi^-}{\partial \mathbf{n}_{\overline{DE}}} - \beta^+ \dfrac{\partial \phi^+}{\partial \mathbf{n}_{\overline{DE}}} \right) ds = 0. & \end{cases} \tag{8.7}$$

We let $\phi_i(X)$ be the piecewise bilinear IFE function such that

$$\phi_i(x_j, y_j) = \begin{cases} 1, & \text{if } i = j, \\ 0, & \text{if } i \neq j \end{cases}$$

for $1 \leq i, j \leq 4$, and we call them the bilinear IFE nodal basis functions on an interface element T. For every element $T \in \mathcal{T}_h$, we let $S_h(T) = span\{\phi_i,\ i = 1, 2, 3, 4\}$, where ϕ_i, $i = 1, 2, 3, 4$ are the standard bilinear nodal basis functions for a non-interface element T; otherwise, ϕ_i, $i = 1, 2, 3, 4$ are the immersed bilinear basis functions defined above.

We need to introduce a few terminologies for describing the selective bilinear immersed finite element space. Assume that we start from a usual rectangular mesh $\mathcal{T}_h^0$ of Ω without any hanging nodes, and we call its elements the 0-th level elements. We refine $\mathcal{T}_h^0$ once by dividing each of a set of selected elements in $\mathcal{T}_h^0$ into 4 congruent sub-rectangles to generate a new mesh $\mathcal{T}_h^1$. We call those new smaller rectangles in in $\mathcal{T}_h^1$ the 1st level elements. Repeating this procedure to the n-th level, we can generate a refined mesh $\mathcal{T}_h$ which can contain elements from level 0 to level n.

For each node $(x_N, y_N)^t$ in a mesh $\mathcal{T}_h$, we define its associated elements to be those elements in which $(x_N, y_N)^t$ is a vertex. For example, the associated elements of the node K in Figure 8.1 are $\square NKIL$ and $\square JEKN$, but $\square EBFI$ is not an associated element of node K even though it is a neighboring element of K. The associated elements of node I in Figure 8.1 are $\square NKIL$, $\square EBFI$, $\square IFCG$ and $\square HIGD$. We then use $\mathcal{T}_{h,N} \subset \mathcal{T}_h$ to represent the collection of all the associated elements of a node $(x_N, y_N)^t$.

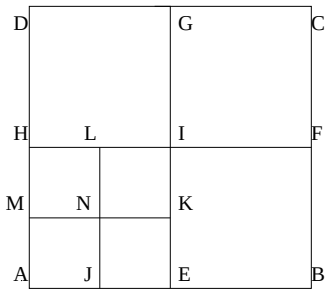

Figure 8.1: A sketch for the associated elements, coarsenable sets and hanging nodes.

At each node $(x_N, y_N)^t$ of $\mathcal{T}_h$, we also define a coarsenable set to be a subset $\mathcal{T}_{h,N,c} \subset \mathcal{T}_{h,N}$ such that

(1) $|\mathcal{T}_{h,N,c}| \geq 2$, i.e., $\mathcal{T}_{h,N,c}$ must contain at leat two associated elements of $(x_N, y_N)^t$.

(2) All the elements in $\mathcal{T}_{h,N,c}$ are in the same level.

(3) Each element in $\mathcal{T}_{h,N,c}$ shares at least one element edge with another element in $\mathcal{T}_{h,N,c}$.

For example, for the node I in Figure 8.1,

$$\{\square EBFI, \square IFCG, \square HIGD\},\ \{\square EBFI, \square IFCG\},\ \{\square IFCG, \square HIGD\}$$

are coarsenable sets, but $\{\square EBFI,\ \square HIGD\}$ is not a coarsenable set. For node I, there is no coarsenable set containing $\square NKIL$.

For each coarsenable set $\mathcal{T}_{h,N,c}$ at a node $(x_N, y_N)^t$, we can defined a piecewise bilinear function $\Psi_{N,c}$ as follows:

$$\begin{aligned}&\Psi_{N,c}|_{T\in\mathcal{T}_{h,N,c}} \in S_h(T),\quad \Psi_{N,c}(x_N, y_N) = 1 \text{ but } \Psi_{N,c} \text{ is zero at other nodes,}\\ &\Psi_{N,c}|_{T\notin\mathcal{T}_{h,N,c}} = 0,\ \Psi_{N,c}|_{\cup\mathcal{T}_{h,N,c}} \text{ is continuous at all nodes in } \mathcal{T}_{h,N,c}.\end{aligned}$$

Then, we define the selective global bilinear IFE basis functions associated with each node $(x_N, y_N)^t \in \mathcal{T}_h$ as follows:

Step 1: For each associated element $T_{N,i} \in \mathcal{T}_{h,N}$, we define a piecewise bilinear function $\Psi_{N,i}$ associated with $(x_N, y_N)^t$ by the zero extension of the local bilinear IFE or FE nodal basis function $\phi(x, y)$ on $T_{N,i}$ such that $\Psi_{N,i}(x_N, y_N) = 1$.

Step 2: We select a collection of coarsenable sets $\mathcal{T}_{h,N,c_j}, j = 1, 2, \cdots, m_N$ and introduce piecewise bilinear functions $\Psi_{N,c_j}, j = 1, 2, \cdots, m_N$.

Step 3: We discard every $\Psi_{N,i}$ defined above if $T_{N,i}$ is contained in one of the selected coarsenable sets $\mathcal{T}_{h,N,c_j}, j = 1, 2, \cdots, m_N$, and call all the remaining piecewise bilinear functions the selective global bilinear IFE basis functions associated with the $(x_N, y_N)^t$.

Finally, we define the selective bilinear immersed finite element (IFE) space $S_h^{Sb}(\Omega)$ to be the space spanned by the selective global bilinear IFE basis functions associated with all the nodes in $\mathcal{T}_h$.

Remark 8.2.1 *We note that in forming the global nodal basis functions associated with a node* $(x_N, y_N)^t$*, the* Step 3 *actually replaces those functions* $\Psi_{N,i}$ *with smaller supports by a function function* Ψ_{N,c_j} *with a larger support provided that* $T_{N,i} \in \mathcal{T}_{h,N,c_j}$*. This makes it possible to form* $S_h^{Sb}(\Omega)$ *of a lower dimension. Also, this step is very similar to the coarsen step in an adaptive mesh refinement procedure.*

The selection of coarsenable sets at each node in the *Step 2* is up to the user of this method. A user can use this opportunity to include desirable features in the method. One necessary requirement is to make sure that the global bilinear IFE basis functions associated with each node are linearly independent. We describe two specific rules for selecting coarsenable sets at each node so that this necessary requirement is fulfilled.

Selective rule 1: At each node $(x_N, y_N)^t$, we select only the largest coarsenable set that doesn't include any interface element.

For example, in Figure 8.1, suppose $\square AJNM$ is the only interface element around node N, then the coarsenable set to be used N is $\{\square JEKN, \square NKIL, \square MNLH\}$ according to **Selective rule** 1.

Selective rule 2: At each node $(x_N, y_N)^t$ of $\mathcal{T}_h$, if $(x_N, y_N)^t$ is also a node of $\mathcal{T}_h^0$, then we select only the largest coarsenable set which doesn't include any interface element. If $(x_N, y_N)^t$ is not a node of $\mathcal{T}_h^0$, we don't select any coarsenable set.

For example, suppose the $\mathcal{T}_h^0$ consists of elements $\square AEIH, \square EBFI, \square IFCG$ and $\square HIGD$ in Figure 8.1 and we refine $\square AEIH$ into four congruent rectangular elements to form the mesh $\mathcal{T}_h$. If $\square AJNM$ is the only interface element, then we don't select any coarsenable set for nodes J, K, L, M, N, but select the coarsenable set $\{\square EBFI, \square IFCG, \square HIGD\}$ for the node I.

Now we use the selective bilinear IFE space $S_h^{Sb}(\Omega) \subset PH_S^1$ and the selective discontinuous Galerkin formulations introduced in the previous section to construct the SIDG method with bilinear IFE as follows: find $u_h \in S_h^{Sb}(\Omega)$ such that

$$a_{S\theta}^-(u_h, v_h) = L_\theta^-(v_h), \quad \forall v_h \in S_h^{Sb}(\Omega), \tag{8.8}$$

or find $u_h \in S_h^{Sb}(\Omega)$ such that

$$a_{S\theta}^+(u_h, v_h) = L_\theta^+(v_h), \quad \forall v_h \in S_h^{Sb}(\Omega). \tag{8.9}$$

8.2.2 Advantages of the SIDG method

As discussed in Section 7.2, Galerkin method doesn't allow any hanging nodes in the meshes, which leads to some strict restrictions of mesh refinement. However, there is no such a restriction for the SIDG method. Meanwhile, IFE allows us to use structured meshes for interface problems, so the SIDG method with IFE spaces will allow us to use adaptive structured meshes with flexible local mesh refinement. That is, in a structured mesh, we can refine any region again and again while keeping the mesh in the rest region coarse. For example, we can refine only the interface elements along the interface. Of course, we can refine the interface elements as many times as needed, see Figure 7.4.

We would like to note that local refinement is a common advantage for all DG methods. Therefore, the immersed DG method discussed in Chapter 7 also has this advantage. However, as mentioned in the introduction section, the immersed DG method increases the computational cost by significant amount. In the following, we will explain how the selective feature in the SIDG method can help to keep the the computational cost low while retaining the local mesh refinement with structured mesh.

At each node outside of Ω_S, there is only one global nodal basis function, which is the same as that of regular Galerkin method. Therefore, the SIDG method only increases the number of global nodal basis functions at those nodes inside Ω_S. If the nodes inside Ω_S are much less than those outside of Ω_S, the SIDG method doesn't increase the number of global nodal basis functions much. Consider the model interface problem with Dirichlet boundary condition, suppose we have a uniform rectangular mesh $\mathcal{T}_h$ on a square with $N \times N$ nodes and L nodes in Ω_S. For Galerkin method, we use $(N-2)^2$ global nodal basis functions. For the immersed DG method, we need $4N^2 - 8N$ global nodal basis functions. For the SIDG method, the number of global nodal basis functions is at most $(N-2)^2 + 3L$. When L is much smaller than N^2, the SIDG method have much less global nodal basis functions than the immersed DG method, hence reduces the computationally cost dramatically. For example, in a mesh $\mathcal{T}_h$ fine enough for the numerical example in Section 8.3, most of its nodes are non-interface nodes, hence most of the global nodal basis functions of $S_h^S(\Omega)$ are just the usual bilinear global nodal basis functions except for few nodes in the vicinity of the interface Γ. Without loss of generality, we can assume $L = 2N$ for the numerical example in Section 8.3, then the number of global nodal basis functions for the SIDG method is at most $(N-2)^2 + 6N$, which is much closer to $(N-2)^2$ than $4N^2 - 8N$ when N is large.

Now we will use an example to compare the computation cost for Galerkin method, the immersed DG method and the SIDG method based on comparable accuracy. Tables 5.1, 7.1, 7.6, 8.2, and 8.10 present the IFE solution errors for Galerkin method, the symmetric immersed DG method, the nonsymmetric immersed DG method, the symmetric SIDG method and the nonsymmetric SIDG method for the same example as in Section 5.2 with $\beta^- = 1$ and $\beta^+ = 10$. From these tables, we can see that all of these methods achieve comparable accuracy on the same mesh.

Note that the number of global basis functions used in a finite element method determines the number of unknowns and the size of the algebraic system, hence reflects the computation cost. In Table 8.1, for different mesh sizes, we compare the number of global basis functions used by the numerical methods mentioned above. Note that a rectangular Cartesian mesh on $[-1,1]\times[-1,1]$ with a step size h has $(\frac{2}{h}+1)^2$ nodes. Let $\sharp 1$, $\sharp 2$, $\sharp 3$ denote the number of global basis functions used by Galerkin method, the immersed DG method and the SIDG method separately. Then we get the following table.

h	$\sharp 1$	$\sharp 2$	$\sharp 3$
1/16	961	4092	1357
1/32	3721	16380	4749
1/64	15625	65532	17681
1/128	64009	262140	68113
1/256	259081	1048572	267281

Table 8.1: Comparison of the number of global basis functions used by Galerkin method, the immersed DG method and the SIDG method.

Table 8.1 shows that $\frac{\sharp 1}{\sharp 3}$ is closer and closer to 1 when the mesh size h becomes smaller and smaller, but $\sharp 2$ stays around four times of $\sharp 1$. Therefore, the computation cost of the SIDG method is much less than that of the immersed DG method while keeping the local refinement feature. Even though Galerkin method uses less number of global basis functions than the SIDG method, it does have much more strict requirement on mesh refinement.

Additionally, the SIDG method only computes the jump terms for all the edges in ε_S but the regular interior penalty DG method needs to compute the jump terms for all the interior element edges. Hence the SIDG method dramatically reduces the computation cost for the jump terms.

8.2.3 Some Implementation issues

In this section, we will discuss some implementation issues for the SIDG method with bilinear IFE under Selective rule 2 and illustrate them by some examples.

As usual, finite element methods need three matrices T, P, and E to store the information of a mesh. The n^{th} column of matrix T stores the global nodal indices of the vertices of the n^{th} element. The n^{th} column of matrix P stores the coordinates of the n^{th} node. The n^{th} of matrix E stores the information of the n^{th} element edge. Because the SIDG method needs more information to be implemented, we introduce two new matrices HT and HE as follows. Basically we use HT to store the index of the global nodal basis functions corresponding to the nodes of all elements and the interface element index for IFE. We use HE to store the information of all element edges in ε_S.

First, we number all the global nodal basis functions in $S_h^S(\Omega)$. This index is different from that of all nodes since one node may be corresponding to more than one global basis functions, see Figure 8.2 for an example. The restriction of each global nodal basis function in $S_h^S(\Omega)$ on an element must be either 0 or a local nodal basis function. For the second case, we say that the global basis function corresponds to that local nodal basis function. Now we use the first four entries of the n^{th} column of matrix HT to store the indices of the global nodal basis functions corresponding to the four local nodal basis functions of the n^{th} element. These information will be used to assemble the global stiffness matrix from the local stiffness matrices. Note that the computation of the local stiffness matrices from $b(u,v)$ and the local load vectors from $L(v)$ on all elements is the same as that of Galerkin method.

Second, we number all the interface elements and call this index as interface element index. For those elements completely in Ω^-, we define their interface indices to be -1. For those elements completely in Ω^+, we define their interface indices to be 0. The left plot in Figure 8.2 gives an example of the interface element index. Then we use the fifth entry of the n^{th} column of matrix HT stores the interface element index of the n^{th} element. These interface information is for IFE.

Third, we number all the element edges in ε_S, see the right plot in Figure 8.3 for an example. Recall that in Section 7.1.1, for each interior edge e, we defined its two neighboring elements T_1, T_2 and the unit normal vector ν of e exterior to its T_2. For the n^{th} element edge $e \in \varepsilon_S$, we use the n^{th} column of matrix HE to store its information as follows. The first two entries store the global nodal indices of the two end points of this edge. The third and fourth entries store the two components of ν. The fifth and sixth entries store the indices of the elements T_1 and T_2 separately. These information are used for computing the line integrals on edges in ε_S and assemble them into the global stiffness matrix.

Now we use the following example to illustrate the above definitions. The left plot in Figure 8.2 is a mesh with indices of its elements and nodes. The number $(i),\ i = 1, \cdots, 16$ at the center of each element is the index of that element and the number $i,\ i = 1, \cdots, 25$ beside each node is the index of that node. The curve is the interface Γ. As usual, we can use the information in this graph to form the regular matrices T and P. The right plot in Figure 8.2 shows the indices of all elements and global basis functions. The number $i,\ i = 1, \cdots, 36$ beside each node is the index of one global nodal basis function corresponding to that node. The formation of those global nodal basis functions can be found in the previous section. The left plot Figure 8.3 shows the interface element index for all elements. The right plot Figure 8.3 shows the index of all element edges in ε_S.

For the partition introduce above, we get the matrices T, HT and HE as follows.

$$T = \begin{pmatrix} 1 & 2 & 3 & 4 & 6 & 7 & 8 & 9 & 11 & 12 & 13 & 14 & 16 & 17 & 18 & 19 \\ 6 & 7 & 8 & 9 & 11 & 12 & 13 & 14 & 16 & 17 & 18 & 19 & 21 & 22 & 23 & 24 \\ 7 & 8 & 9 & 10 & 12 & 13 & 14 & 15 & 17 & 18 & 19 & 20 & 22 & 23 & 24 & 25 \\ 2 & 3 & 4 & 5 & 7 & 8 & 9 & 10 & 12 & 13 & 14 & 15 & 17 & 18 & 19 & 20 \end{pmatrix},$$

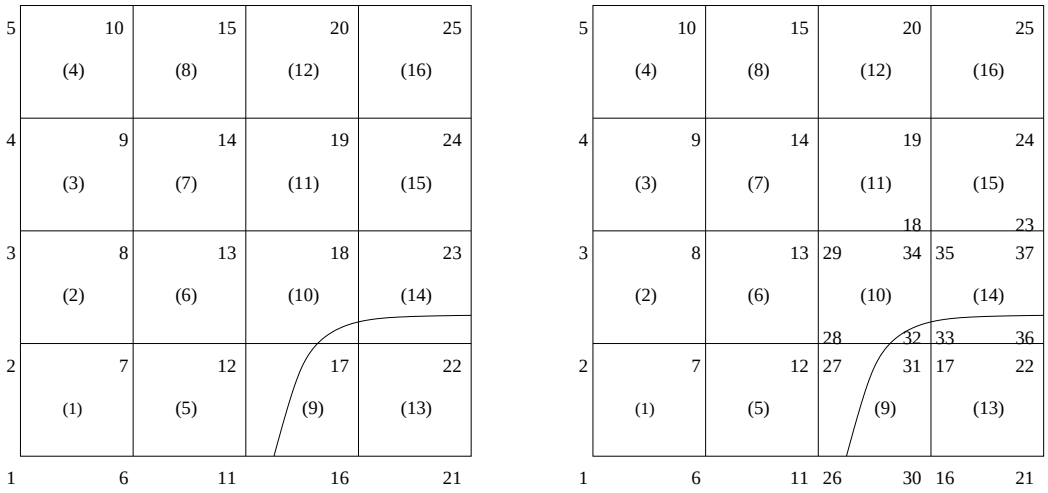

Figure 8.2: The left plot shows the indices of all elements and nodes. The right plot shows the indices of all elements and global nodal basis functions.

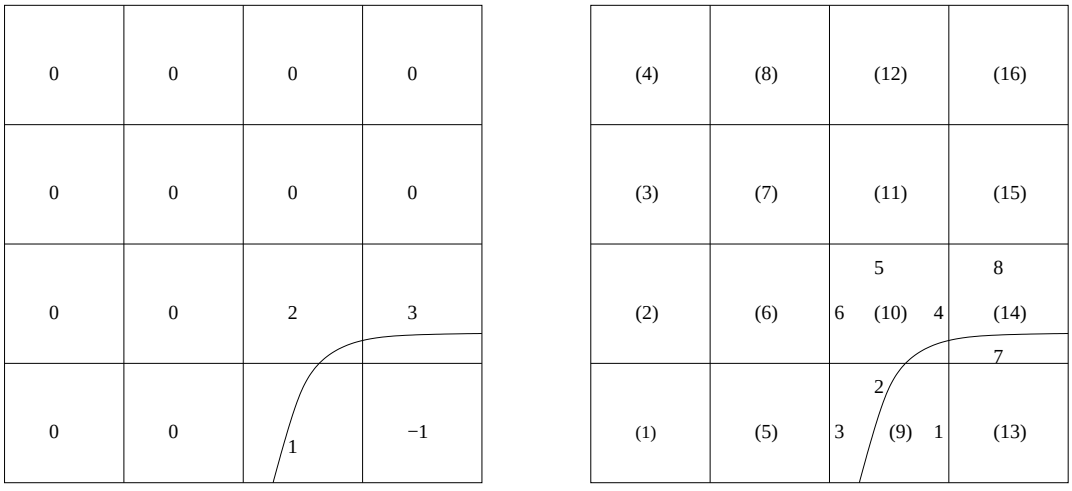

Figure 8.3: The left plot shows the interface element index. The right plot shows the index of all element edges in ε_S.

$$HT = \begin{pmatrix} 1 & 2 & 3 & 4 & 6 & 7 & 8 & 9 & 26 & 28 & 13 & 14 & 16 & 33 & 18 & 19 \\ 6 & 7 & 8 & 9 & 11 & 12 & 13 & 14 & 30 & 32 & 18 & 19 & 21 & 36 & 23 & 24 \\ 7 & 8 & 9 & 10 & 12 & 13 & 14 & 15 & 31 & 34 & 19 & 20 & 22 & 37 & 24 & 25 \\ 2 & 3 & 4 & 5 & 7 & 8 & 9 & 10 & 27 & 29 & 14 & 15 & 17 & 35 & 19 & 20 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 2 & 0 & 0 & -1 & 3 & 0 & 0 \end{pmatrix},$$

$$HE = \begin{pmatrix} 16 & 17 & 12 & 17 & 18 & 13 & 17 & 23 \\ 17 & 12 & 11 & 18 & 13 & 12 & 22 & 18 \\ -1 & 0 & -1 & -1 & 0 & -1 & 0 & 0 \\ 0 & -1 & 0 & 0 & -1 & 0 & -1 & -1 \\ 9 & 9 & 5 & 10 & 10 & 6 & 13 & 14 \\ 13 & 10 & 9 & 14 & 11 & 10 & 14 & 15 \end{pmatrix}.$$

Now we use two examples to show how we use the matrices above to generate a local stiffness matrix and assemble its entries to the global stiffness matrix. Suppose we are working on the 10^{th} element which is an interface element.

We first use $T(:, 10)$ to obtain the information of this element. Then we can use the information to compute the area integrals in (8.8) on this element, which form the local stiffness matrix of this element. This process is very similar to that of Galerkin method. However, we need to use matrix HT instead of T to assemble the entries in the local stiffness matrix into the global stiffness matrix. Assume $B = [b_{ij}]_{i,j=1}^{4}$ is the local stiffness matrix and A is the global stiffness matrix. Then we should assemble b_{ij} to $A(HT(i, 10), HT(j, 10))$.

As for the line integrals, suppose we are working on the 2^{nd} element edge in ε_S. $HE(:, 2)$ tells us all the information we need to compute the line integrals in (8.8) on this edge. Then we use $HE(5, 2) = 9$ and $HE(6, 2) = 10$ to get the indices of the two neighboring elements of this edge. Using them together with HT, we can assemble the results of these line integrals on the 2^{nd} element edge in ε_S into the global stiffness matrix A. For example, if the test function and trial function are the i^{th} and j^{th} local basis function of the 9^{th} element, then we assemble the corresponding line integral to $A(HT(i, 9), HT(j, 9))$. If the test function is the i^{th} local basis function of the 10^{th} element and the trial function is the j^{th} local basis function of the 9^{th}, then we assemble the corresponding line integral to $A(HT(i, 10), HT(j, 9))$.

Remark 8.2.2 *HT and HE introduced above provide basic information for the SIDG method. They are not necessarily most efficient. On the other side, we may add more information to HT and HE if necessary.*

Remark 8.2.3 *The immersed DG method and the interior penalty DG method with regular finite elements also need some matrices similar to the matrices HT and HE. However, since the SIDG method only use discontinuous formulation on part of the domain. The matrix HE for the SIDG method can be much smaller than that of the immersed DG method and the regular interior penalty DG method, depending how much we want to use the DG formulation, and this leads to reduction of memory and and computation time.*

8.3 Numerical examples

In this section, we will show some numerical examples for the SIDG method based on the selective bilinear IFE space. Consider the same example as in Section 5.2 with $\beta^- = 1$. The way to construct the refined meshes is the same as in Section 7.4 and we also call the original rectangular mesh without refinement in Section 7.4 0-level refined mesh. We will present numerical results for two cases with $\beta^+ = 10$ and $\beta^+ = 1000000$ separately.

Under the selective rule 2 in Section 8.2.1, at each node of the n^{th}-level ($n \geq 0$) refined mesh, if the node is still a node of the 0-level refined mesh, then we select only the largest coarsenable set which doesn't include any interface element. If the node is not a node of the 0-level refined mesh, then we don't select any coarsenable set. In this case Ω_S is actually the union of all interface elements of the 0-level refined mesh and ε_S is the set of all element edges of the n^{th}-level ($n \geq 0$) refined mesh in $\overline{\Omega_S}$.

8.3.1 Numerical results for the symmetric SIDG method with bilinear IFE

We will first numerically show the the convergence in L^2, H^1 and discrete infinity norms for the symmetric SIDG method with bilinear IFE on the original rectangular mesh without mesh refinement, see the left plot of Figure 7.4. Table 8.2 contains the errors of the solutions u_h with various partition sizes h, $\beta^+ = 10$ and $C_* = 1000$. Table 8.3, 8.4 and 8.5 contain the errors of the solutions u_h with various partition sizes h, $\beta^+ = 1000000$ and $C_* = 1000, 0.0001, 0$ separately.

h	$\|u_h - u\|_0$	$\|u_h - u\|_1$	$\|u_h - u\|_\infty$
1/16	9.3598×10^{-4}	4.6983×10^{-2}	1.9320×10^{-3}
1/32	2.2376×10^{-4}	2.2336×10^{-2}	5.0699×10^{-4}
1/64	5.8624×10^{-5}	1.0568×10^{-2}	1.3541×10^{-4}
1/128	1.4462×10^{-5}	5.2162×10^{-3}	3.4717×10^{-5}
1/256	3.6792×10^{-6}	2.6168×10^{-3}	8.6371×10^{-6}

Table 8.2: Errors of the symmetric SIDG method with bilinear IFE on the original mesh for $\beta^+ = 10$ and C_*=1000.

Using linear regression, we can also see that the data in Table 8.2 obey

$$\|u_h - u\|_0 \approx 0.2309\ h^{1.9933},\ \ |u_h - u|_1 \approx 0.8318\ h^{1.0431},\ \ \|u_h - u\|_\infty \approx 0.4347\ h^{1.9479},$$

and the data in Table 8.3 obey

$$\|u_h - u\|_0 \approx 0.2460\ h^{1.9996},\ \ |u_h - u|_1 \approx 0.8720\ h^{1.0726},\ \ \|u_h - u\|_\infty \approx 0.1488\ h^{1.5585}.$$

h	$\|u_h - u\|_0$	$\|u_h - u\|_1$	$\|u_h - u\|_\infty$
1/16	9.5450×10^{-4}	4.5462×10^{-2}	2.3301×10^{-3}
1/32	2.4804×10^{-4}	2.1547×10^{-2}	7.5253×10^{-4}
1/64	5.9188×10^{-5}	9.5477×10^{-3}	1.6324×10^{-4}
1/128	1.4644×10^{-5}	4.6827×10^{-3}	5.5396×10^{-5}
1/256	3.8419×10^{-6}	2.3697×10^{-3}	3.8733×10^{-5}

Table 8.3: Errors of the symmetric SIDG method with bilinear IFE on the original mesh for $\beta^+ = 1000000$ and C_*=1000.

h	$\|u_h - u\|_0$	$\|u_h - u\|_1$	$\|u_h - u\|_\infty$
1/16	7.0060×10^{-3}	2.2073×10^{0}	1.0925×10^{-1}
1/32	1.0443×10^{-2}	3.0078×10^{0}	5.7371×10^{-1}
1/64	3.0155×10^{-4}	3.6226×10^{-1}	7.8894×10^{-2}
1/128	2.3456×10^{-3}	6.1921×10^{-1}	8.1909×10^{-2}
1/256	5.4700×10^{-5}	2.8866×10^{-1}	2.8750×10^{-2}

Table 8.4: Errors of the symmetric SIDG solutions on the original mesh for $\beta^+ = 1000000$ and C_*=0.0001.

h	$\|u_h - u\|_0$	$\|u_h - u\|_1$	$\|u_h - u\|_\infty$
1/16	7.0081×10^{-3}	2.2085×10^{0}	1.0931×10^{-1}
1/32	1.3363×10^{-1}	2.8799×10^{0}	5.6452×10^{-4}
1/64	3.1109×10^{-4}	3.7472×10^{-1}	8.2323×10^{-2}
1/128	2.9876×10^{-3}	2.9423×10^{0}	4.2683×10^{-1}
1/256	6.4373×10^{-5}	3.3041×10^{-1}	3.5378×10^{-2}

Table 8.5: Errors of the symmetric SIDG solutions on the original mesh for $\beta^+ = 1000000$ and C_*=0.

From the linear regressions, we can see that the SIDG solution u_h converges to the exact solution with convergence rates $O(h^2)$ in the L^2 norm, $O(h)$ in the H^1 norm when the penalty constant C_* is large enough. However, Table 8.5 shows that it doesn't converge as expected with penalty constant $C^* = 0$. Therefore, the non-trivial penalty terms are necessary in order to guarantee the convergence. Our theoretic analysis also requires a large enough penalty constant C^* for convergence and the oscillating errors in Table 8.4 and 8.5 confirm this requirement numerically. Additionally, we can see that the solution u_h doesn't always converge to the exact solution with convergence rate $O(h^2)$ in the discrete infinity norm, which was also observed for Galerkin method and finite volume element method with IFEs before [112, 113, 144, 149]. The corresponding analysis leads to some interesting future research work.

In order to illustrate the global effect of the local mesh refinement, we will compare the numerical errors in L^2, H^1 and discrete infinity norms on different meshes with the step sizes $h = \frac{1}{8}, \frac{1}{16}, \frac{1}{32}$. Table 8.6 to 8.8 contain numerical errors in L^2, H^1 and discrete infinity norm on the original mesh and the refined meshes from the first-level to the fourth-level. Note that the h is the the step size of the corresponding original mesh. From the three tables, we can see that the effect of the first-level and second-level refined meshes are dramatic, but the effect of the third-level and fourth-level refined meshes is not much. That's because the error of the non-interface area is not reduced as much as that of the interface area.

h	$\frac{1}{8}$	$\frac{1}{16}$	$\frac{1}{32}$
original mesh	2.1366×10^{-2}	4.7971×10^{-3}	7.2254×10^{-4}
first-level adaptive mesh	4.3277×10^{-3}	8.2111×10^{-4}	1.7169×10^{-4}
second-level adaptive mesh	2.2492×10^{-3}	5.9412×10^{-4}	1.5710×10^{-4}
third-level adaptive mesh	2.0740×10^{-3}	5.8111×10^{-4}	1.5364×10^{-4}
fourth-level adaptive mesh	2.0634×10^{-3}	5.7744×10^{-4}	1.5281×10^{-4}

Table 8.6: L^2 norm errors of the symmetric SIDG solutions on different meshes.

h	$\frac{1}{8}$	$\frac{1}{16}$	$\frac{1}{32}$
original mesh	1.7413×10^{-1}	7.2458×10^{-2}	2.7376×10^{-2}
first-level adaptive mesh	7.8639×10^{-2}	3.9678×10^{-2}	1.9964×10^{-2}
second-level adaptive mesh	6.6543×10^{-2}	3.7186×10^{-2}	1.9708×10^{-2}
third-level adaptive mesh	6.5186×10^{-2}	3.7008×10^{-2}	1.9664×10^{-2}
fourth-level adaptive mesh	6.5084×10^{-2}	3.6968×10^{-2}	1.9656×10^{-2}

Table 8.7: H^1 norm errors of the symmetric SIDG solutions on different meshes.

To illustrate the local effect of the local mesh refinement around the interface for the symmetric SIDG method, we compare error deduction in the discrete infinity norm on the interface

h	$\frac{1}{8}$	$\frac{1}{16}$	$\frac{1}{32}$
original mesh	3.0810×10^{-2}	8.0050×10^{-3}	3.2539×10^{-3}
first-level adaptive mesh	8.3872×10^{-3}	1.9109×10^{-3}	4.6486×10^{-4}
second-level adaptive mesh	5.0352×10^{-3}	1.4862×10^{-3}	4.0208×10^{-4}
third-level adaptive mesh	4.6433×10^{-3}	1.4302×10^{-3}	3.7831×10^{-4}
fourth-level adaptive mesh	4.6150×10^{-3}	1.4105×10^{-3}	3.7724×10^{-4}

Table 8.8: Discrete infinity norm errors of the symmetric SIDG solutions on different meshes.

elements in Table 8.9. Note that the h is the the step size of the original mesh without refinement. From the data in this table, we can see that the local refinement reduce the error around interface efficiently.

h	$\frac{1}{8}$	$\frac{1}{16}$	$\frac{1}{32}$
uniform mesh	2.7754×10^{-2}	8.0050×10^{-3}	3.2539×10^{-3}
first-level adaptive mesh	5.4905×10^{-3}	1.7120×10^{-3}	3.1164×10^{-4}
second-level adaptive mesh	1.7126×10^{-3}	6.2751×10^{-4}	1.6193×10^{-4}
third-level adaptive mesh	3.7265×10^{-4}	1.9026×10^{-4}	8.3144×10^{-5}
fourth-level adaptive mesh	2.5030×10^{-4}	9.3606×10^{-5}	3.1081×10^{-5}

Table 8.9: Comparison of the discrete infinity norm errors of the symmetric SIDG solutions on interface elements.

8.3.2 Numerical results for the nonsymmetric SIDG method with bilinear IFE

Comparing the symmetric SIDG's dependence of the penalty constant on the problem, one important advantage of the non-symmetric SIDG method is that its penalty constant needs only to be positive. In this section, we will illustrate this feature by presenting the errors of the non-symmetric SIDG method with different penalty constants on the original rectangular mesh. Table 8.10 and 8.11 contain the errors of the solutions u_h with various partition sizes h, $\beta^+ = 10$ and the penalty constant $C_{**} = 1000, 0.0001, 0$ separately. Table 8.12 and 8.13 contain the errors of the solutions u_h with various partition sizes h, $\beta^+ = 1000000$ and the penalty constant $C_{**} = 1000, 0.0001$ separately.

Using linear regression, we can also see that the data in Table 8.10 obey

$$\|u_h - u\|_0 \approx 0.3056\ h^{1.9817},\ \ |u_h - u|_1 \approx 0.8244\ h^{1.0412},\ \ \|u_h - u\|_\infty \approx 0.4361\ h^{1.9260},$$

the data in Table 8.11 obey

$$\|u_h - u\|_0 \approx 0.3083\ h^{2.0443},\ \ |u_h - u|_1 \approx 0.8337\ h^{1.0430},\ \ \|u_h - u\|_\infty \approx 0.3937\ h^{1.9213},$$

h	$\|u_h - u\|_0$	$\|u_h - u\|_1$	$\|u_h - u\|_\infty$
1/8	4.8417×10^{-3}	9.4182×10^{-2}	7.4548×10^{-3}
1/16	1.2938×10^{-3}	4.6984×10^{-2}	2.1715×10^{-3}
1/32	3.1571×10^{-4}	2.2336×10^{-2}	5.6652×10^{-4}
1/64	8.1634×10^{-5}	1.0568×10^{-2}	1.5057×10^{-4}
1/128	2.0260×10^{-5}	5.2162×10^{-3}	3.8429×10^{-5}
1/256	5.1259×10^{-6}	2.6168×10^{-3}	9.5583×10^{-6}

Table 8.10: Errors of the nonsymmetric SIDG solutions on the original mesh for $\beta^+ = 10$ and C_{**}=1000.

h	$\|u_h - u\|_0$	$\|u_h - u\|_1$	$\|u_h - u\|_\infty$
1/8	5.5721×10^{-3}	9.4791×10^{-2}	8.4884×10^{-3}
1/16	9.2221×10^{-4}	4.7363×10^{-2}	1.7588×10^{-3}
1/32	2.1135×10^{-4}	2.2414×10^{-2}	4.4750×10^{-4}
1/64	6.0943×10^{-5}	1.0606×10^{-2}	1.1981×10^{-4}
1/128	1.5797×10^{-5}	5.2267×10^{-3}	3.9623×10^{-5}
1/256	4.0357×10^{-6}	2.6197×10^{-3}	9.6156×10^{-6}

Table 8.11: Errors of the nonsymmetric SIDG solutions on the original mesh for $\beta^+ = 10$ and C_{**}=0.0001.

h	$\|u_h - u\|_0$	$\|u_h - u\|_1$	$\|u_h - u\|_\infty$
1/8	3.3558×10^{-3}	9.0901×10^{-2}	7.0848×10^{-3}
1/16	9.4698×10^{-4}	4.5722×10^{-2}	2.4746×10^{-3}
1/32	2.4617×10^{-4}	2.1560×10^{-2}	7.4044×10^{-4}
1/64	5.8874×10^{-5}	9.5491×10^{-3}	1.6587×10^{-4}
1/128	1.4523×10^{-5}	4.6839×10^{-3}	5.4856×10^{-5}
1/256	3.7641×10^{-6}	2.3559×10^{-3}	2.0259×10^{-5}

Table 8.12: Errors of the nonsymmetric SIDG solutions on the original mesh for $\beta^+ = 1000000$ and C_{**}=1000.

h	$\|u_h - u\|_0$	$\|u_h - u\|_1$	$\|u_h - u\|_\infty$
1/8	4.1938×10^{-3}	8.9981×10^{-2}	7.5690×10^{-3}
1/16	5.6185×10^{-4}	4.5020×10^{-2}	1.6592×10^{-3}
1/32	1.1305×10^{-4}	2.0943×10^{-2}	4.2071×10^{-4}
1/64	5.4539×10^{-5}	9.5756×10^{-3}	1.3950×10^{-4}
1/128	1.2871×10^{-5}	4.6820×10^{-3}	3.8574×10^{-5}
1/256	3.0983×10^{-6}	2.3497×10^{-3}	1.0579×10^{-5}

Table 8.13: Errors of the nonsymmetric SIDG solutions on the original mesh for $\beta^+ = 1000000$ and C_{**}=0.0001.

h	$\|u_h - u\|_0$	$\|u_h - u\|_1$	$\|u_h - u\|_\infty$
1/8	4.1942×10^{-3}	8.9981×10^{-2}	7.5696×10^{-3}
1/16	5.6183×10^{-4}	4.5020×10^{-2}	1.6592×10^{-3}
1/32	1.1306×10^{-4}	2.0943×10^{-2}	4.2074×10^{-4}
1/64	5.4551×10^{-5}	9.5756×10^{-3}	1.3952×10^{-4}
1/128	1.2874×10^{-5}	4.6820×10^{-3}	3.8579×10^{-5}
1/256	3.0990×10^{-6}	2.3497×10^{-3}	1.0580×10^{-5}

Table 8.14: Errors of the nonsymmetric SIDG solutions on the original mesh for $\beta^+ = 1000000$ and C_{**}=0.

the data in Table 8.12 obey

$$\|u_h - u\|_0 \approx 0.2176\ h^{1.9756},\ \ |u_h - u|_1 \approx 0.8532\ h^{1.0682},\ \ \|u_h - u\|_\infty \approx 0.2778\ h^{1.7398},$$

and the data in Table 8.13 obey

$$\|u_h - u\|_0 \approx 0.1753\ h^{1.9831},\ \ |u_h - u|_1 \approx 0.8303\ h^{1.0634},\ \ \|u_h - u\|_\infty \approx 0.3172\ h^{1.8653},$$

and the data in Table 8.14 obey

$$\|u_h - u\|_0 \approx 0.1753\ h^{1.9830},\ \ |u_h - u|_1 \approx 0.8303\ h^{1.0634},\ \ \|u_h - u\|_\infty \approx 0.3172\ h^{1.8653}.$$

From the tables and linear regressions, we can see that the solution u_h converges to the exact solution with convergence rates $O(h^2)$ in the L^2 norm, $O(h)$ in the H^1 norm for penalty constant $C_{**} = 1000, 0.0001, 0$. This numerically confirms that the penalty constant C_{**} needs to be only a positive constant or even 0. The corresponding analysis leads to some future work. For all the linear regressions in this chapter, we obtain similar figures to Figure 5.1, which mean that the data points match the linear regression lines very well.

8.4 Convergence of the symmetric SIDG method with bilinear IFE

In this section, we will follow the framework in [7] to analyze the convergence of the symmetric SIDG method with bilinear IFE and prove the optimal convergence rate of the solutions in energy norm. As before, we use C to represent a generic constant whose value might be different from line to line. Unless otherwise specified, all the generic constants C in the presentation below are independent of interface and mesh.

Consider a mesh $\mathcal{T}_h$ with step size h. We first recall the following definitions. For any function $u \in PH^2_{int}(\Lambda)$, we define

$$\begin{aligned} \|u\|^2_{s,\Lambda} &= \|u\|^2_{s,\Lambda^+} + \|u\|^2_{s,\Lambda^-},\ s = 0,1,2, \\ |u|^2_{s,\Lambda} &= |u|^2_{s,\Lambda^+} + |u|^2_{s,\Lambda^-},\ s = 0,1,2. \end{aligned}$$

Consider an interface element $T \in \mathcal{T}_h$. For any function $w_h \in S_h(T)$ which will be recalled from [112, 149], we define

$$\begin{aligned} \|w_h\|^2_{s,T} &= \|u\|^2_{s,\tilde{T}^+} + \|u\|^2_{s,\tilde{T}^-},\ s = 0,1,2, \\ |w_h|^2_{s,T} &= |u|^2_{s,\tilde{T}^+} + |w_h|^2_{s,\tilde{T}^-},\ s = 0,1,2. \end{aligned}$$

For any function $w_h \in S_h(T)$ and $u \in PH^2_{int}(T)$, we define

$$\begin{aligned} \|w_h + u\|^2_{s,T} &= \|w_h + u\|^2_{s,\tilde{T}^+\cap T^+} + \|w_h + u\|^2_{s,\tilde{T}^+\cap T^-} + \|w_h + u\|^2_{s,\tilde{T}^-\cap T^+} + \|w_h + u\|^2_{s,\tilde{T}^-\cap T^-}, \\ |w_h + u|^2_{s,T} &= |w_h + u|^2_{s,\tilde{T}^+\cap T^+} + |w_h + u|^2_{s,\tilde{T}^+\cap T^-} + |w_h + u|^2_{s,\tilde{T}^-\cap T^+} + |w_h + u|^2_{s,\tilde{T}^-\cap T^-}, \\ & \quad s = 0,1,2. \end{aligned}$$

Here note that one of $\tilde{T}^+ \cap T^-$ and $\tilde{T}^- \cap T^+$ might be empty. In that case we can remove the norm and semi-norm on the empty set from the above definitions. For a set $\Lambda \subset \Omega$ whose interior is not cut through by Γ, we define $\|u\|_{s,\Lambda}$ and $|u|_{s,\Lambda}$ to be the usual H^s norm and H^s semi-norm on Λ $(s = 0, 1, 2)$ separately. Let $\varepsilon_{SD} = \varepsilon_h^D \bigcup \varepsilon_S$, then we define

$$\begin{aligned} \|u\|_{s,\mathcal{T}_h}^2 &= \sum_{T\in\mathcal{T}_h} \|u\|_{s,T}^2, \ s = 0, 1, 2, \\ |||u|||^2 &= \|u\|_{1,\mathcal{T}_h}^2 + \sum_{e\in\varepsilon_{SD}} \left(h^{-1}\|[u]\|_{0,e}^2 + h\|\{\beta\nabla u\cdot\nu\}\|_{0,e}^2\right). \end{aligned} \tag{8.10}$$

Applying the same argument in Lemma 5.3.5, we obtain a similar trace inequality on $S_h(T)$.

Lemma 8.4.1 *For each element $T = \square A_1A_2A_3A_4 \in \mathcal{T}_h$, define*

$$E_1(\partial T) = \overline{A_1A_2}, \ E_2(\partial T) = \overline{A_2A_3}, \ E_3(\partial T) = \overline{A_3A_4}, \ E_4(\partial T) = \overline{A_4A_1}.$$

Then we have the following trace inequality on every $T \in \mathcal{T}_h$:

$$\left\|\beta\frac{\partial w_h}{\partial n}\right\|_{0,E_i(\partial T)}^2 \le C\left(h^{-1}\left|w_h\right|_{1,T}^2 + h\left|w_h\right|_{2,T}^2\right), \quad \forall w_h \in S_h(T), \quad 1 \le i \le 4. \tag{8.11}$$

If $|\beta| \ge b_1 > 0$, then

$$\left\|\frac{\partial w_h}{\partial n}\right\|_{0,E_i(\partial T)}^2 \le C\left(h^{-1}\left|w_h\right|_{1,T}^2 + h\left|w_h\right|_{2,T}^2\right), \quad \forall w_h \in S_h(T), \quad 1 \le i \le 4. \tag{8.12}$$

The following two lemmas are similar to (2.2) and (2.8) in [7], but we need to show the constants C_1 and C_2 are independent of interface and mesh. Even though our space here is a little bit different from that in [7], the whole proof for Lemma 2.1 in [7] is still true because we have $w_h \in L^2(\Omega)$ for each $w_h \in S_h^{Sb}(\Omega)$.

Lemma 8.4.2 *There exists a constant C_1 independent of interface and mesh such that*

$$\|w_h\|_{0,\mathcal{T}_h}^2 \le C_1\left(\|\nabla w_h\|_{0,\mathcal{T}_h}^2 + \sum_{e\in\mathcal{E}_{SD}} h^{-1}\|[w_h]\|_{0,e}^2\right), \ \forall w_h \in S_h^{Sb}(\Omega)$$

Proof. Define $\psi \in H^2(\Omega)\bigcap H_0^1(\Omega)$ by $-\Delta\psi = w_h$. By the regularity of the elliptic problem, we have

$$\|\psi\|_{2,\mathcal{T}_h} \le C\|w_h\|_{0,\mathcal{T}_h}.$$

By trace theorem [7], we get

$$\|\nabla\psi\cdot\nu\|_{0,e}^2 \;\leq\; C(h^{-1}|\psi|_{1,T}^2+h|\psi|_{2,T}^2),$$

which leads to

$$\|\nabla\psi\|_{0,\mathcal{T}_h}^2+\sum_{e\in\varepsilon_{SD}} h\,\|\nabla\psi\cdot\nu\|_{0,e}^2 \;\leq\; C\,\|\psi\|_{2,\mathcal{T}_h}^2 \;\leq\; C\,\|w_h\|_{0,\mathcal{T}_h}^2\,.$$

Then

$$\begin{aligned}
\|w_h\|_{0,\mathcal{T}_h}^2 &= (w_h,w_h) \;=\; (w_h,-\Delta\psi)\\
&= \sum_{T\in\mathcal{T}_h}(\nabla w_h,\nabla\psi)_T-\sum_{e\in\varepsilon_{SD}}([w_h],\nabla\psi\cdot\nu)_e\\
&\leq C\left(\|\nabla w_h\|_{0,\mathcal{T}_h}^2+\sum_{e\in\varepsilon_{SD}} h^{-1}\,\|[w_h]\|_{0,e}^2\right)^{1/2}\left(\|\nabla\psi\|_{0,\mathcal{T}_h}^2+\sum_{e\in\varepsilon_{SD}} h\,\|\nabla\psi\cdot\nu\|_{0,e}^2\right)^{1/2}\\
&\leq C\left(\|\nabla w_h\|_{0,\mathcal{T}_h}^2+\sum_{e\in\varepsilon_{SD}} h^{-1}\,\|[w_h]\|_{0,e}^2\right)^{1/2}\|w_h\|_{0,\mathcal{T}_h},
\end{aligned}$$

which completes the proof.

■

Lemma 8.4.3 *There exists a constant C_2 such that*

$$\sum_{e\in\mathcal{E}_{SD}} h\|\{\nabla w_h\cdot\nu\}\|_{0,e}^2 \;\leq\; C_2\|\nabla w_h\|_{0,\mathcal{T}_h}^2,\ \forall w_h\in S_h^{Sb}(\Omega).$$

Proof. From (3.20), we have

$$|w_h|_{1,T}^2 \;\geq\; Ch^2|w_h|_{2,T}^2.$$

This inverse inequality is also obviously true for any non-interface element T since we use standard bilinear finite element functions on them. Combining this with the trace inequality (8.12), we get

$$\|\nabla w_h\cdot\nu\|_{0,e}^2 \;\leq\; C(h^{-1}|w_h|_{1,T}^2+h|w_h|_{2,T}^2) \;\leq\; Ch^{-1}|w_h|_{1,T}^2 \;\leq\; Ch^{-1}\,\|\nabla w_h\|_{0,T}^2\,,$$

which completes the proof.

■

Now we start to show the convergence for the symmetric SIDG method in the following theorem. Without loss of generality, we only prove for the case with homogeneous Dirichlet condition. In this case, we have

$$\begin{aligned} b(u,v) &= \sum_{T\in\mathcal{T}_h}(\beta\nabla u,\nabla v)_T, \\ J_S(u,v) &= \sum_{e\in\varepsilon_S}(\{\beta\nabla u\cdot\nu\},[v])_e, \\ J_{S\theta}(u,v) &= \sum_{e\in\varepsilon_S}\theta_e([u],[v])_e, \\ a^-_{S\theta}(u,v) &= b(u,v)-J_S(u,v)-J_S(v,u)+J_{S\theta}(u,v), \\ L^-_\theta(v) &= \sum_{T\in\mathcal{T}_h}(f,v)_T. \end{aligned}$$

It is easy to show that $a^-_{S\theta}(u,v)$ is bounded:

$$a^-_{S\theta}(u,v) \leq C|||u|||\,|||v|||. \tag{8.13}$$

The following lemma establishes the lower bound of $a^-_{S\theta}$.

Lemma 8.4.4 *If* $0<b_1\leq|\beta|\leq b_2<\infty$, $\theta_e=\frac{C_*}{h}$ *and* $C_*\geq\frac{4C_2b_2^2}{b_1}+\frac{b_1}{2}$, *then*

$$a^-_{S\theta}(w_h,w_h) \geq \alpha|||w_h|||^2+\frac{1}{2}J_{S\theta}(w_h,w_h),\ \forall w_h\in S_h^{Sb}(\Omega). \tag{8.14}$$

Proof. Since $\theta_e=\frac{C_*}{h}$ and $C_*\geq\frac{4C_2b_2^2}{b_1}+\frac{b_1}{2}$, then

$$\frac{1}{2}J_{S\theta}(w_h,w_h) \geq \left(\frac{2C_2b_2^2}{b_1}+\frac{b_1}{4}\right)\sum_{e\in\varepsilon_S}h^{-1}\|[w_h]\|^2_{0,e}.$$

Therefore, $\forall w_h \in S_h^{Sb}(\Omega)$, we have

$$\begin{aligned}
a_{S\theta}^{-}(w_h,w_h) &= \sum_{T\in\mathcal{T}_h}(\beta\nabla w_h,\nabla w_h)_T - 2\sum_{e\in\varepsilon_S}(\{\beta\nabla w_h\cdot\nu\},[w_h])_e + J_{S\theta}(w_h,w_h)\\
&\geq b_1\|\nabla w_h\|_{0,\mathcal{T}_h}^2 - 2b_2\sum_{e\in\varepsilon_S}\|\{\nabla w_h\cdot\nu\}\|_{0,e}\|[w_h]\|_{0,e} + J_{S\theta}(w_h,w_h)\\
&\geq b_1\|\nabla w_h\|_{0,\mathcal{T}_h}^2 - 2b_2\left(\sum_{e\in\varepsilon_S}h\|\{\nabla w_h\cdot\nu\}\|_{0,e}^2\right)^{\frac{1}{2}}\left(\sum_{e\in\varepsilon_S}h^{-1}\|[w_h]\|_{0,e}^2\right)^{\frac{1}{2}} + J_{S\theta}(w_h,w_h)\\
&= b_1\|\nabla w_h\|_{0,\mathcal{T}_h}^2 - 2\left(\frac{b_1}{2C_2}\sum_{e\in\varepsilon_S}h\|\{\nabla w_h\cdot\nu\}\|_{0,e}^2\right)^{\frac{1}{2}}\left(\frac{2C_2b_2^2}{b_1}\sum_{e\in\varepsilon_S}h^{-1}\|[w_h]\|_{0,e}^2\right)^{\frac{1}{2}}\\
&\quad + J_{S\theta}(w_h,w_h)\\
&\geq \frac{b_1}{2}\|\nabla w_h\|_{0,\mathcal{T}_h}^2 + \frac{b_1}{2}\|\nabla w_h\|_{0,\mathcal{T}_h}^2 - \frac{b_1}{2C_2}\sum_{e\in\varepsilon_S}h\|\{\nabla w_h\cdot\nu\}\|_{0,e}^2\\
&\quad - \frac{2C_2b_2^2}{b_1}\sum_{e\in\varepsilon_S}h^{-1}\|[w_h]\|_{0,e}^2 + J_{S\theta}(w_h,w_h). \qquad (8.15)
\end{aligned}$$

Applying Lemma 8.4.3, we get

$$\begin{aligned}
a_{S\theta}^{-}(w_h,w_h) &\geq \frac{b_1}{2}\|\nabla w_h\|_{0,\mathcal{T}_h}^2 - \frac{2C_2b_2^2}{b_1}\sum_{e\in\varepsilon_S}h^{-1}\|[w_h]\|_{0,e}^2 + J_{S\theta}(w_h,w_h)\\
&= \frac{b_1}{4}\|\nabla w_h\|_{0,\mathcal{T}_h}^2 + \frac{b_1}{4}\|\nabla w_h\|_{0,\mathcal{T}_h}^2 - \frac{2C_2b_2^2}{b_1}\sum_{e\in\varepsilon_S}h^{-1}\|[w_h]\|_{0,e}^2 + J_{S\theta}(w_h,w_h)\\
&\geq \frac{b_1}{4}\|\nabla w_h\|_{0,\mathcal{T}_h}^2 + \frac{b_1}{4C_2}\sum_{e\in\varepsilon_S}h\|\{\nabla w_h\cdot\nu\}\|_{0,e}^2 - \frac{2C_2b_2^2}{b_1}\sum_{e\in\varepsilon_S}h^{-1}\|[w_h]\|_{0,e}^2 + J_{S\theta}(w_h,w_h).
\end{aligned}$$

Then using Lemma 8.4.2, we have

$$\begin{aligned}
a_{S\theta}^{-}(w_h,w_h) &\geq \frac{b_1}{8}\|\nabla w_h\|_{0,\mathcal{T}_h}^2 + \frac{b_1}{8}\|\nabla w_h\|_{0,\mathcal{T}_h}^2 + \frac{b_1}{4C_2}\sum_{e\in\varepsilon_S}h\|\{\nabla w_h\cdot\nu\}\|_{0,e}^2\\
&\quad - \frac{2C_2b_2^2}{b_1}\sum_{e\in\varepsilon_S}h^{-1}\|[w_h]\|_{0,e}^2 + J_{S\theta}(w_h,w_h)\\
&\geq \frac{b_1}{8}\|\nabla w_h\|_{0,\mathcal{T}_h}^2 + \frac{b_1}{8C_1}\|w_h\|_{0,\mathcal{T}_h}^2 + \frac{b_1}{4C_2}\sum_{e\in\varepsilon_S}h\|\{\nabla w_h\cdot\nu\}\|_{0,e}^2\\
&\quad - \left(\frac{2C_2b_2^2}{b_1}+\frac{b_1}{8}\right)\sum_{e\in\varepsilon_S}h^{-1}\|[w_h]\|_{0,e}^2 + J_{S\theta}(w_h,w_h).
\end{aligned}$$

Finally, combining this with (8.15), we get

$$\begin{aligned}
a^{-}_{S\theta}(w_h, w_h) &\geq \frac{b_1}{8}\|\nabla w_h\|^2_{0,\mathcal{T}_h} + \frac{b_1}{8C_1}\|w_h\|^2_{0,\mathcal{T}_h} + \frac{b_1}{4C_2}\sum_{e\in\varepsilon_S} h\|\{\nabla w_h\cdot\nu\}\|^2_{0,e} \\
&\quad - \left(\frac{2C_2b_2^2}{b_1} + \frac{b_1}{8}\right)\sum_{e\in\varepsilon_S} h^{-1}\|[w_h]\|^2_{0,e} + \frac{1}{2}J_{S\theta}(w_h, w_h) + \frac{1}{2}J_{S\theta}(w_h, w_h) \\
&\geq \frac{b_1}{8}\|\nabla w_h\|^2_{0,\mathcal{T}_h} + \frac{b_1}{8C_1}\|w_h\|^2_{0,\mathcal{T}_h} + \frac{b_1}{4C_2}\sum_{e\in\varepsilon_S} h\|\{\nabla w_h\cdot\nu\}\|^2_{0,e} \\
&\quad - \left(\frac{2C_2b_2^2}{b_1} + \frac{b_1}{8}\right)\sum_{e\in\varepsilon_S} h^{-1}\|[w_h]\|^2_{0,e} + \left(\frac{2C_2b_2^2}{b_1} + \frac{b_1}{4}\right)\sum_{e\in\varepsilon_S} h^{-1}\|[w_h]\|^2_{0,e} + \frac{1}{2}J_{S\theta}(w_h, w_h) \\
&= \frac{b_1}{8}\|\nabla w_h\|^2_{0,\mathcal{T}_h} + \frac{b_1}{8C_1}\|w_h\|^2_{0,\mathcal{T}_h} + \frac{b_1}{4C_2}\sum_{e\in\varepsilon_S} h\|\{\nabla w_h\cdot\nu\}\|^2_{0,e} \\
&\quad + \frac{b_1}{8}\sum_{e\in\varepsilon_S} h^{-1}\|[w_h]\|^2_{0,e} + \frac{1}{2}J_{S\theta}(w_h, w_h) \\
&\geq \alpha|||w_h|||^2 + \frac{1}{2}J_{S\theta}(w_h, w_h).
\end{aligned}$$

In the last inequality we pick $\alpha = min\{\frac{b_1}{8}, \frac{b_1}{8C_1}, \frac{b_1}{4C_2}\}$.

■

Now we recall the definition of bilinear IFE interpolation. For a function $u \in PH^2_{int}(T)$, $T \in \mathcal{T}_h$, we let $I_{h,T}u \in S_h(T)$ be its interpolation such that $I_{h,T}u(X) = u(X)$ when X is a vertex of T. In general, For an element T with vertices A_1, A_2, A_3, A_4, we have

$$I_{h,T}u(X) = u(A_1)\phi_1(X) + u(A_2)\phi_2(X) + u(A_3)\phi_3(X) + u(A_4)\phi_4(X).$$

Accordingly, for a function $u \in PH^2_{int}(\Omega)$, we let I_hu be its interpolation such that $I_hu|_T = I_{h,T}(u|_T)$ for any $T \in \mathcal{T}_h$.

If we select the coarsenable sets such that the supports of the global nodal basis functions in $S^{Sb}_h(\Omega)$ are not overlapped, then it's easy to verify that $I_hu \in S^{Sb}_h(\Omega)$. This is a usual way to construct global basis functions in order to reduce the computational cost. Note that both the Selective Rule 1 and the Selective Rule 2 in Section 8.2.1 satisfy this requirement.

The following lemma provides an upper bound for the difference between the SIDG solution and the bilinear IFE interpolation of the analytic solution. For the convergence analysis here, we only need the boundedness, but this lemma may lead to some future work about superconvergence of the SIDG solution.

Lemma 8.4.5 *If $0 < b_1 \leq |\beta| \leq b_2 < \infty$, $\theta_e = \frac{C_*}{h}$ and $C_* \geq \frac{4C_2b_2^2}{b_1} + \frac{b_1}{2}$, then there exists a constant C such that*

$$|||I_hu - u_h|||^2 \leq C|||I_hu - u|||^2. \tag{8.16}$$

Proof. It's easy to verify that $S_h^{Sb}(\Omega) \subset PH_S^1(\mathcal{T}_h)$. Then (7.8) and (7.10) lead to

$$a_{S\theta}^-(u-u_h, v_h) \;=\; 0,\ \forall v_h \in S_h^{Sb}(\Omega). \tag{8.17}$$

Using (8.14), we get

$$\begin{aligned}\alpha|||I_h u - u_h|||^2 &\le a_{S\theta}^-(I_h u - u_h, I_h u - u_h)\\ &= a_{S\theta}^-(I_h u - u, I_h u - u_h) + a_{S\theta}^-(u - u_h, I_h u - u_h).\end{aligned}$$

Then by (8.17), we get

$$\alpha|||I_h u - u_h|||^2 \;\le\; a_{S\theta}^-(I_h u - u, I_h u - u_h).$$

Combining this with (8.13), we have

$$\begin{aligned}\alpha|||I_h u - u_h|||^2 &\le C|||I_h u - u|||\,|||I_h u - u_h|||\\ &\le \frac{C^2}{2\alpha}|||I_h u - u|||^2 + \frac{\alpha}{2}|||I_h u - u_h|||^2,\end{aligned}$$

which leads to (8.16).

■

Finally, the following theorem establishes the optimal convergence of the symmetric SIDG solution u_h in energy norm.

Theorem 8.4.1 *Assume the solution u of the model interface problem (1.1)-(1.4) is in $PH_{int}^2(\Omega)$. If $0 < b_1 \le |\beta| \le b_2 < \infty$, $\theta_e = \frac{C_*}{h}$ and $C_* \ge \frac{4C_2 b_2^2}{b_1} + \frac{b_1}{2}$, then there exists a constant C such that the symmetric immersed DG solution u_h has the following error bound*

$$|||u - u_h|||^2 \;\le\; Ch\|u\|_{2,\mathcal{T}_h} \tag{8.18}$$

Proof. Using (8.16), we have

$$\begin{aligned}|||u - u_h|||^2 &\le |||u - I_h u|||^2 + |||I_h u - u_h|||^2\\ &\le (1 + C)|||I_h u - u|||^2.\end{aligned} \tag{8.19}$$

Also, we have the following regular trace inequality [7]

$$\|v\|_{0,e}^2 \;\le\; C(h^{-1}\|v\|_{0,T}^2 + h|v|_{1,T}^2),\ \forall v \in H^1(T),\ e \subset \partial T, \tag{8.20}$$

and it is easy to verify that $S_h^{Sb}(\Omega) \subset H^1(T)$. Applying the same arguments in Lemma 5.3.5 to $I_h u - u$, we get

$$\|\nabla(I_h u - u)\cdot\nu\|_{0,e}^2 \;\le\; C\left(h^{-1}|I_h u - u|_{1,T}^2 + h|I_h u - u|_{2,T}^2\right). \tag{8.21}$$

Then using Theorem 4.1.13, Theorem 4.1.24, (8.10), (8.19), (8.20), and (8.21), we get

$$\begin{aligned}
& |||u-u_h|||^2 \\
\leq\ & (1+C)\left(\|I_hu-u\|^2_{1,\mathcal{T}_h}+\sum_{e\in\varepsilon_S}h^{-1}\|[I_hu-u]\|^2_{0,e}+\sum_{e\in\varepsilon_S}h\|\{\nabla(I_hu-u)\cdot\nu\}\|^2_{0,e}\right) \\
\leq\ & C\Bigg[\|I_hu-u\|^2_{1,\mathcal{T}_h}+\sum_{T\in\mathcal{T}_h}\left(h^{-2}\|I_hu-u\|^2_{0,T}+|I_hu-u|^2_{1,T}\right)+ \\
& \sum_{T\in\mathcal{T}_h}\left(|I_hu-u|^2_{1,T}+h^2|I_hu-u|^2_{2,T}\right)\Bigg] \\
\leq\ & Ch^2\|u\|^2_{2,\mathcal{T}_h}.
\end{aligned}$$

which completes the proof.

■

Chapter 9

Bilinear IFE for the non-homogeneous flux jump condition

Most of the previous articles about IFE are developed for solving the model interface problem with the homogeneous flux jump condition. Y. Gong, B. Li and Z. Li [98] present a homogenization using level set method to deal with the nonhomogeneous jump conditions. In this chapter, we will construct a new bilinear IFE space to deal with the following model interface problem with non-homogeneous flux jump condition.

$$-\nabla \cdot \big(\beta \nabla u\big) = f(x,y), \quad (x,y) \in \Omega, \tag{9.1}$$

$$u|_{\partial\Omega} = g(x,y) \tag{9.2}$$

together with the jump conditions on the interface Γ:

$$[u]\,|_{\Gamma} = 0, \tag{9.3}$$

$$\left[\beta \frac{\partial u}{\partial n}\right]|_{\Gamma} = Q(x,y). \tag{9.4}$$

This equation is critical to many applications, such as the charging problem for a conductor with induced charge. The basic idea is to add more basis functions, that can capture the non-homogeneous flux jump, to the original bilinear IFE space for homogeneous jump condition.

9.1 The bilinear IFE space for the non-homogeneous flux jump condition

In this section, we will construct a new IFE space over Ω which satisfies the non-homogeneous flux jump condition in a weak sense. We will use the notations introduced in Chapter 3.

Let $\mathcal{T}_h$ be a rectangular mesh of Ω. For a typical interface element $T \in \mathcal{T}_h$, we introduce the following new basis function:

$$\phi_J(x,y) = \begin{cases} \phi_J^-(x,y) = a^- x + b^- y + c^- + d^- xy, & (x,y) \in \tilde{T}^-, \\ \phi_J^+(x,y) = a^+ x + b^+ y + c^+ + d^+ xy, & (x,y) \in \tilde{T}^+, \\ \phi_J(x_j,y_j) = 0, \ j = 1,2,3,4, & \\ \phi_J^-(D) = \phi_J^+(D), \quad \phi_J^-(E) = \phi_J^+(E), \quad d^- = d^+, & \\ \int_{\overline{DE}} \left(\beta^- \dfrac{\partial \phi_J^-}{\partial \mathbf{n}_{\overline{DE}}} - \beta^+ \dfrac{\partial \phi_J^+}{\partial \mathbf{n}_{\overline{DE}}} \right) ds = 1, & \end{cases} \tag{9.5}$$

See Figure 9.1 for two typical bilinear IFE basis functions of Type I and Type II for nonhomogeneous flux jump condition. Their two pieces are separated by $\overline{DE}$.

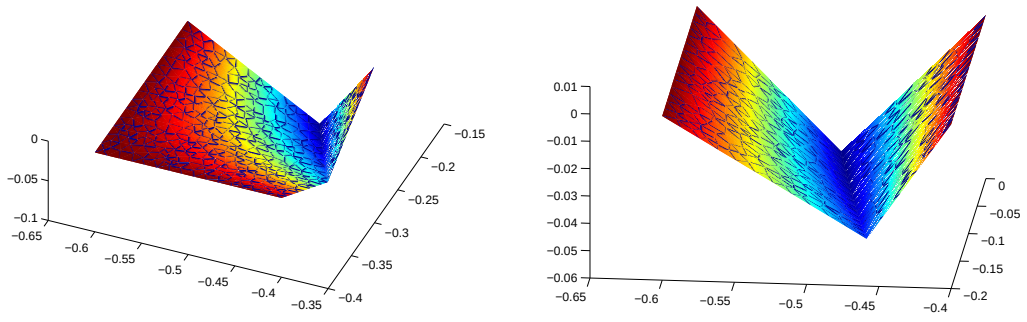

Figure 9.1: The plot on the left is a ϕ_J on a Type I interface element and the plot on the right is a ϕ_J on a Type II interface element. Both of them use $\overline{DE}$ to separate the two pieces.

Now we use this new basis function and the bilinear immersed finite element basis functions introduced in Section 3.1 to to define the new bilinear immersed finite element (IFE) space $S_h^J(\Omega)$. First, for every non-interface element $T \in \mathcal{T}_h$, we let $S_h(T) = span\{\phi_i, \ i = 1,2,3,4\}$, where ϕ_i, $i = 1,2,3,4$ are the standard bilinear nodal basis functions. For an interface element T, we let $S_h(T) = span\{\phi_J, \ \phi_i, \ i = 1,2,3,4\}$ where ϕ_i, $i = 1,2,3,4$ are the immersed bilinear basis functions defined by (3.4). Suppose there are N nodes and M interface elements in $\mathcal{T}_h$. Then, we define a continuous piecewise bilinear global nodal basis function $\psi_j(x,y)$ $(j = 1, \cdots, N)$ for each node $(x_j, y_j)^t$ of $\mathcal{T}_h$ such that $\psi_j(x_j, y_j) = 1$ but ψ_j is zero at other nodes, and $\psi_j|_T \in S_h(T)$ for any element $T \in \mathcal{T}_h$. Additionally, we define a piecewise bilinear global nodal basis function $\psi_{Jk}(x,y)$ $(k = 1, \cdots, M)$ for the k^{th} interface element T of $\mathcal{T}_h$ such that $\psi_{Jk}|_T = \phi_J$ and ψ_J is zero everywhere else. Finally, we define $S_h^J(\Omega) = span\{\psi_j, j = 1, \cdots, N, \psi_{Jk}, k = 1, \cdots, M\}$.

9.2 Finite element interpolation on $S_h^J(\Omega)$

In this section, we define the finite element interpolation on the new space $S_h^J(\Omega)$ and investigate its approximation capability numerically. For any subset $\Lambda \subset \Omega$ whose interior is cut through by the interface Γ, we define

$$PH^2_{int,J}(\Lambda) = \left\{ u \in C(\Lambda), u|_{\Lambda^s} \in H^2(\Lambda^s), s = -,+, \left[\beta \frac{\partial u}{\partial \mathbf{n_\Gamma}}\right] = Q \ \text{ on } \Gamma \cap \Lambda \right\}.$$

Note that this is not a linear space. For a function $u \in PH^2_{int,J}(T)$, we define

$$I_{h,T}u(X) = u(A_1)\phi_1(X) + u(A_2)\phi_2(X) + u(A_3)\phi_3(X) + u(A_4)\phi_4(X) + q\phi_J,$$

where

$$q = \int_{\overline{DE}} Q ds. \tag{9.6}$$

For a non-interface element T, we define

$$I_{h,T}u(X) = u(A_1)\phi_1(X) + u(A_2)\phi_2(X) + u(A_3)\phi_3(X) + u(A_4)\phi_4(X).$$

Accordingly, for a function $u \in PH^2_{int}(\Omega)$, we let $I_h^J u \in S_h^J(\Omega)$ be its interpolation such that $I_h^J u|_T = I_{h,T}(u|_T)$ for any $T \in \mathcal{T}_h$.

Based on the conclusions in Chapter 4, we naturally expect this new bilinear immersed finite element space has the optimal approximation capability as follows.

$$\begin{aligned} \|I_h u - u\|_{0,\Omega} &\leq Ch^2 \|u\|_{2,\Omega}, \\ |I_h u - u|_{1,\Omega} &\leq Ch \|u\|_{2,\Omega}. \end{aligned}$$

In the following, we use a numerical example to verify it numerically. For simplicity, we only present results obtained by using the bilinear IFE space based on uniformly rectangular Cartesian partitions in the rectangular domain $\Omega = (-1,1) \times (-1,1)$. The interface curve Γ is a circle with radius $r_0 = \pi/6.28$ which separates Ω into two sub-domains Ω^- and Ω^+ with $\Omega^- = \{(x,y) \ | \ x^2 + y^2 \leq r_0^2\}$. Here we choose

$$u(x,y) = \frac{(x^2+y^2)^{5/2}}{\beta^-},$$

which gives

$$Q(x,y) = 5(\beta^+ - \beta^-)\frac{(x^2+y^2)^{5/2}}{r_0}.$$

Table 9.1 contains actual errors of the IFE interpolation $I_h^J u$ with various partition sizes h for $\beta^- = 1, \beta^+ = 10$ and $\beta^- = 1, \beta^+ = 10000$ separately. Using linear regression, we can also see that the data in this table obey

$$\left\|I_h^J u - u\right\|_0 \approx 3.6279\ h^{1.9998}, \left|I_h^J u - u\right|_1 \approx 8.7742\ h^{0.9998},$$
$$\left\|I_h^J u - u\right\|_0 \approx 3.6288\ h^{1.9998}, \left|I_h^J u - u\right|_1 \approx 8.8116\ h^{1.0005},$$

which clearly indicate that the interpolation converges to u with convergence rates $O(h^2)$ and $O(h)$ in the L^2 norm and H^1 norm, respectively.

h	$\left\|I_h u - u\right\|_0$	$\left\|I_h u - u\right\|_1$
1/16	1.4172×10^{-2}	5.4838×10^{-1}
1/32	3.5460×10^{-3}	2.7443×10^{-1}
1/64	8.8666×10^{-4}	1.3724×10^{-1}
1/128	2.2167×10^{-4}	6.8620×10^{-2}
1/256	5.5418×10^{-5}	3.4310×10^{-2}
1/512	1.3855×10^{-5}	1.7155×10^{-2}
1/1024	3.4636×10^{-6}	8.5773×10^{-3}

Table 9.1: Errors in the interpolation $I_h^J u$ when $\beta^- = 1, \beta^+ = 10$

h	$\left\|I_h u - u\right\|_0$	$\left\|I_h u - u\right\|_1$
1/16	1.4173×10^{-2}	5.4848×10^{-1}
1/32	3.5468×10^{-3}	2.7578×10^{-1}
1/64	8.8677×10^{-4}	1.3756×10^{-1}
1/128	2.2169×10^{-4}	6.8659×10^{-2}
1/256	5.5420×10^{-5}	3.4316×10^{-2}
1/512	1.3855×10^{-5}	1.7156×10^{-2}
1/1024	3.4637×10^{-6}	8.5775×10^{-3}

Table 9.2: Errors in the interpolation $I_h^J u$ when $\beta^- = 1, \beta^+ = 10000$

Our numerical examples above indicate that the bilinear IFE space $S_h^J(\Omega)$ has the optimal approximation capability expected from bilinear polynomials. The interpolation error estimation leads to some interesting future work.

9.3 Galerkin method for solving the model interface problem with nonhomogeneous jump

In this section, we will introduce the Galerkin method for solving the model interface problem with non-homogeneous flux jump condition.

First, we multiply the differential equation (9.1) by any $v \in H_0^1(\Omega)$ and integrate it over $\Omega^s (s = +, -)$ to have

$$-\int_{\Omega^s} \nabla \cdot \big(\beta^s \nabla u\big) v \; dxdy \;\; = \;\; \int_{\Omega^s} fv \; dxdy, \forall \; v \in H_0^1(\Omega).$$

Then a straightforward application of the Green's formula leads to

$$\int_{\Omega^s} \beta^s \nabla u \cdot \nabla v \; dxdy - \int_{\partial\Omega^s} \beta \frac{\partial u}{\partial \mathbf{n}} v \; ds \;\; = \;\; \int_{\Omega^s} fv \; dxdy, \;\; s = +, -, \forall \; v \in H_0^1(\Omega). \quad (9.7)$$

Summing (9.7) over s, we get the weak formulation

$$\int_{\Omega} \beta \nabla u \cdot \nabla v \; dxdy \;\; = \;\; \int_{\Omega} fv \; dxdy - \int_{\Gamma} Qv ds, \forall \; v \in H_0^1(\Omega). \quad (9.8)$$

Here we use the flux jump condition (9.4) and $v \in H_0^1(\Omega)$.

Let $S_{h,0}^J(\Omega) \subset S_h^J(\Omega)$ consist of functions of $S_h^J(\Omega)$ vanishing on $\mathcal{N}_h \cap \partial\Omega$. The IFE Galerkin method can be described as follows: find $u_h \in S_h^J(\Omega)$ satisfying

$$\sum_{T \in \mathcal{T}_h} \int_T \beta \nabla u_h \cdot \nabla v_h \; dxdy \;\; = \;\; \int_{\Omega} fv_h \; dxdy - \int_{\Gamma} Qv_h ds, \forall \; v_h \in S_{h,0}^J(\Omega). \quad (9.9)$$

Suppose there are N nodes and M interface elements in $\mathcal{T}_h$. Let q_k denote the quantity q defined in (9.6) for the k^{th} interface element. Let $u_h = \sum_{j=1}^N u_j \psi_j + \sum_{k=1}^M q_k \psi_{Jk}$. As usual, u_j is the value of u_h at the j^{th} node. Finally, we get the linear system arising from the Galerkin IFE method as follows.

$$\begin{aligned} & \sum_{j=1}^N u_j \left(\sum_{T \in \mathcal{T}_h} \int_T \beta \nabla \psi_j \cdot \nabla \psi_i \; dxdy \right) \\ = \;\; & \int_{\Omega} f\psi_i \; dxdy - \int_{\Gamma} Q\psi_i ds - \sum_{k=1}^M q_k \left(\sum_{T \in \mathcal{T}_h} \int_T \beta \nabla \psi_{Jk} \cdot \nabla \psi_i \; dxdy \right), \\ & i = 1, \cdots, N. \end{aligned} \quad (9.10)$$

We note that the term on left side of this system and the first term on the right side are the same as the immersed Galerkin method for homogeneous jump condition discussed in

Section 5.1. Therefore, the only extra work here is to compute the second and third terms on the right side.

Since this bilinear IFE space has an $O(h^2)$ (in L^2-norm) and an $O(h)$ (in H^1-norm) approximation capability, we naturally expect the finite element method based on this IFE space to perform accordingly. To confirm this numerically, we consider the interface problem defined by (9.1)-(9.4) in which the boundary condition function $g(x, y)$ and the source term $f(x, y)$ are chosen such that the function $u(x, y) = \dfrac{(x^2 + y^2)^{5/2}}{\beta^-}$ is the exact solution. We use the same domain Ω and interface curve Γ as in the example of Section 9.2.

Table 9.3 contains actual errors of the bilinear IFE solutions u_h with various partition size h for the interface problem with the coefficient function $\beta(x, y)$ with $\beta^- = 1, \beta^+ = 10$.

Using linear regression, we can easily see that the data in this table obey

$$\|u_h - u\|_0 \approx 4.1440\ h^{1.9806}, |u_h - u|_1 \approx 8.5601\ h^{0.9906},$$

which indicate that the bilinear IFE solution u_h converges to the exact solution with convergence rates $O(h^2)$ and $O(h)$ in the L^2 norm and H^1 norm, respectively.

Table 9.4 contains actual errors of the bilinear IFE solutions u_h with various partition size h for the interface problem with the coefficient function $\beta(x, y)$ with $\beta^- = 1, \beta^+ = 10000$. The errors in this group of computations obey

$$\|u_h - u\|_0 \approx 28.5181\ h^{1.8124}, |u_h - u|_1 \approx 35.1895\ h^{0.8820}.$$

For all the linear regressions in this chapter, we obtain similar figures to Figure 5.1, which show that the data points match the linear regression lines very well.

h	$\|u_h - u\|_0$	$\|u_h - u\|_1$	$\|u_h - u\|_\infty$
1/16	1.8523×10^{-2}	5.5089×10^{-1}	1.2984×10^{-2}
1/32	3.9352×10^{-3}	2.7578×10^{-1}	3.2897×10^{-3}
1/64	1.0293×10^{-3}	1.3888×10^{-1}	2.4211×10^{-3}
1/128	3.0337×10^{-4}	6.9828×10^{-2}	1.0082×10^{-3}
1/256	6.9673×10^{-5}	3.5349×10^{-2}	4.9377×10^{-4}

Table 9.3: Errors of the IFE solutions for the case when $\beta^- = 1, \beta^+ = 10$.

h	$\|u_h - u\|_0$	$\|u_h - u\|_1$	$\|u_h - u\|_\infty$
1/16	3.1037×10^{-1}	3.1940×10^{0}	8.5700×10^{-1}
1/32	2.5783×10^{-2}	1.6276×10^{0}	1.3191×10^{-1}
1/64	1.6276×10^{-2}	8.3037×10^{-1}	1.2329×10^{-1}
1/128	4.4398×10^{-3}	4.9857×10^{-1}	4.1561×10^{-2}
1/256	1.3994×10^{-3}	2.7143×10^{-1}	-1.8584×10^{-2}

Table 9.4: Errors of the IFE solutions for the case when $\beta^- = 1, \beta^+ = 10000$.

Chapter 10

Conclusions, applications and future works

10.1 Conclusions

In this dissertation we carry out a systematic study of the bilinear immersed finite element(IFE) for interface problems. We have discussed all the three fundamental aspects for a new finite element method, including the development of the bilinear immersed finite element spaces, the implementation of numerical methods with these spaces, and the corresponding convergence analysis.

First, we construct a bilinear IFE space whose functions satisfy homogeneous jump conditions. Then we investigate basic properties for this space. Second, we use multi-variable Taylor expansion to carry out the error estimation for the bilinear IFE interpolation of a Sobolev function. The interpolation error estimates indicate that this space has the usual approximation capability expected from bilinear polynomials, which is $O(h^2)$ in L^2 norm and $O(h)$ in H^1 norm. Third, we implement several numerical methods with bilinear IFE, including Galerkin, finite volume element and discontinuous Galerkin (DG) methods. From all the numerical examples, we can see that all these methods with bilinear IFE have the same optimal convergence rates as those with standard bilinear finite element. We have also done some theoretical analysis for the convergence of these methods. For the symmetric immersed DG method and the selective immersed DG method with bilinear IFE, we prove its optimal convergence in energy norm. In addition, the convergence of Galerkin method with bilinear IFE is proved.

The selective immersed DG method only applies the DG formulation wherever necessary. The computational cost of this method can be maintained almost the same as that of the regular Galerkin method if the DG formulation is used only around the interface, but flexible local mesh refinement around the interface can still be carried out, which is not allowed by

regular Galerkin method. Finally, in order to deal with the nonhomogeneous jump condition, we add special basis functions to the bilinear IFE space to form a nonhomogeneous bilinear IFE space.

10.2 Future works

In fact, in this dissertation we haven't completely finished the research on bilinear IFE. Even though we have finished some convergence analysis, we still need to prove the optimal convergence for Galerkin method, the finite volume element method, and the nonsymmetric selective immersed DG method with bilinear IFE. Also, we need to carry out the analysis about the bilinear IFE for nonhomogeneous flux jump conditions.

Based on the immersed finite elements for the popular second order elliptic interface problem, we plan to extend the current immersed finite elements to more sophisticated problems, such as elasticity equation, Maxwell equation and Navier-Stokes equation with variable discontinuous coefficients and nonhomogeneous jump conditions. Additionally, trilinear immersed finite element is in our future plan since it is very useful for 3D simulations. As for higher order finite element space, we plan to study the bi-quadratic and tri-quadratic immersed finite elements.

The spitting extrapolation is one of the efficient techniques for accelerating the convergence of approximations [41, 42, 114, 115, 116, 121, 145, 146, 147, 148, 156, 157, 158, 168, 184, 185, 204]. This method can be naturally parallelized with a high degree of parallelism. It improves the accuracy with less computational complexity than Richardson extrapolation and requires only piecewise smoothness for the analytic solutions. The design of the independent parameters also gives us flexibility in choosing different kinds of meshes. These advantages of splitting extrapolation become more clear and powerful when the size of the problem is large and more independent meshes sizes are designed with domain decomposition. We plan to apply splitting extrapolation to IFE to accelerate its convergence and implement parallel algorithms. In addition, the extension to more intricate cases and the corresponding analysis for both finite volume and finite element methods are also in our future plan, especially for the finite volume method.

10.3 Application: simulation for charging in space

As mentioned in Chapter 1, there are many applications for the model interface problem and immersed finite elements, including electromagnetic problems, flow problems, topology/shape optimization, multiscale finite elements, bio-molecular problems, and the modeling of nonlinear phenomena, etc. What we have been working on is charging in space, which includes a series of important problems in aerospace engineering, such as spacecraft charging

under different types of solar winds, ion propulsion and lunar dust levitation, to name just a few. It involves the calculation of the electromagnetic environment, interaction between the environment and the spacecraft, effect of the solar winds on lunar dust and so on.

We present a simulation model on electrostatic levitation of lunar dust using the Particle-In-Cell(PIC) method [195, 196]. One time-consuming simulation step is to solve a 2D or 3D elliptic interface equation for the electric potential, so we apply the bilinear IFE or the 3D linear IFE in this step. For some simple cases, we also use a finite difference scheme to deal with the interface problems. Full particle PIC simulation are carried out to obtain plasma sheath and wake, surface charging, the transition point of surface electric field and the floating potential of lunar lander in the lunar terminator region. Test particle simulations and dust-in-plasma simulations are carried out to simulate the levitation of dust from lunar surface and dust transport around lunar lander. Results show that the dust levitation condition in the terminator region is sensitively influenced by ambient plasma and surface charging, and the levitation altitude varies significantly even for small changes of the sun elevation angle. In addition, the plasma sheath profiles and particle density distributions satisfy the theoretical analysis.

Based on these encouraging results, we are improving our simulation model and numerical methods to simulate more sophisticated cases, in which the lunar surface is concave up and down, the spacecraft and astronauts consists of different materials with complicated geometric structures, and some objects are moving.

We also plan to apply IFEs to other applications such as topology/shape optimization, multiscale finite element methods, the Navier-Stokes equation, and bio-molecular problems.

Bibliography

[1] S. Adjerid, K. D. Devine, J. E. Flaherty, and L. Krivodonova. A posteriori error estimation for discontinuous Galerkin solutions of hyperbolic problems. *Comput. Methods Appl. Mech. Engrg.*, 19(11-12):1097–1112, 2002.

[2] S. Adjerid and T. Lin. Higher-order immersed discontinuous Galerkin methods. *Int. J. Inf. Syst. Sci.*, 3(4):555–568, 2007.

[3] S. Adjerid and T. Lin. p-th degree immersed finite element for boundary value problems with discontinuous coefficients. *J. Appl. Numer. Math*, accepted.

[4] A. Almgre, J. B. Bell, P. Collela, and T. Marthaler. A Cartesian grid method for incompressible Euler equations in complex geometries. *SIAM J. Sci. Comput.*, 18:1289–1390, 1997.

[5] B. Amaziane and M. El Ossmani. Convergence analysis of an approximation to miscible fluid flows in porous media by combining mixed finite element and finite volume methods. *Numer. Meth. Part. Differ. Equ.*, 24(3):799–832, 2008.

[6] Y. Arakawa and M. Nakano. An efficient three-dimensional optics code for ion thruster research. In *AIAA-3196*, 1996.

[7] D. N. Arnold. An interior penalty finite element method with discontinuous elements. *SIAM J. Numer. Anal*, 19:742–760, 1982.

[8] D. N. Arnold, F. Brezzi, and B. Cockburn L. D. Marini. Unified analysis of discontinuous Galerkin methods for elliptic problems. *SIAM J. Numer. Anal.*, 39(5):1749–1779, 2001/02.

[9] I. Babuska. The finite element method with penalty. *Math. Comp.*, 27:221–228, 1973.

[10] I. Babuska, C. E. Baumann, and J. T. Oden. A discontinuous hp finite element method for diffusion problems:1D analysis. *Comput. & Math. Appl.*, 37:103–122, 1999.

[11] I. Babuska and M. Zlamal. Nonconforming elements in the finite element method with penalty. *SIAM J. Numer. Anal.*, 10:863–875, 1973.

[12] I. Babuška. The finite element method for elliptic equations with discontinuous coefficients. *Computing*, 5:207–213, 1970.

[13] I. Babuška, G. Caloz, and J. E. Osborn. Special finite element methods for a class of second order elliptic prolbems with rough coefficients. *SIAM J. Numer. Anal.*, 31:945–981, 1994.

[14] I. Babuška and J. Melenk. The partition of unity method. *Int. J. Numer. Meth. Eng.*, 40:727–758, 1997.

[15] I. Babuška and J. E. Osborn. Generalized finite element methods: their performance and relation to mixed methods. *SIAM J. Numer. Anal.*, 20(3):510–536, 1983.

[16] I. Babuška and J. E. Osborn. Finite element methods for the solution of problems with rough input data. In P. Grisvard, W. Wendland, and J.R. Whiteman, editors, *Singular and Constructive Methods for their Treatment, Lecture Notes in Mathematics, #1121*, pages 1–18, New York, 1985. Springer-Verlag.

[17] I. Babuška and J. E. Osborn. Can a finite element method perform arbitrarily badly? *Math. Comp.*, 69(230):443–462, 2000.

[18] G. A. Baker. Finite element methods for elliptic equations using nonconforming elements. *Math. Comp.*, 31(137):45–59, 1977.

[19] R. E. Bank and D. J. Rose. Some error estimates for the box method. *SIAM J. Numer. Anal.*, 24(4):777–787, 1987.

[20] F. Bassi and S. Rebay. A high-order accurate discontinuous finite element method for the numerical solution of the compressible Navier-Stokes equations. *J. Comput. Phys.*, 131(2):267–279, 1997.

[21] C. E. Baumann and J. T. Oden. An adaptive-order discontinuous Galerkin method for the solution of the Euler equations of gas dynamics. *Richard H. Gallagher Memorial Issue. Int. J. Numer. Meth. Engrg.*, 47(1-3):61–73, 2000.

[22] J. B. Bell, P. Colella, and H. M. Glaz. A second-order projection method for the incompressible Navier-Stokes equations. *J. Comput. Phys.*, 85:257–283, 1989.

[23] J. B. Bell and D. L. Marcus. A second-order projection method for variable-density flows. *J. Comput. Phys.*, 101:334–348, 1992.

[24] T. Belytschko, N. Moës, S. Usui, and C. Primi. Arbitrary discontinuities in finite elements. *Int. J. Numer. Meth. Eng.*, 50:993–1013, 2001.

[25] M. P. Bendsøe. *Optimization of structural topology, shape, and material.* Springer, 1995.

[26] M. P. Bendsøe and N. Kikuchi. Generating optimal topologies in optimal design using a homogenization method. *Comp. Meth. Appl. Mech. Engng.*, 71:197–224, 1998.

[27] M. P. Bendsøe and O. Sigmund. *Topology optimization*. Springer, 2nd edition, 2004.

[28] A. Berger, R. Scott, and G. Strang. Approximate boundary conditions in the finite element method. *Symposia Methematica*, 10:295–313, 1972.

[29] C. Bi and V. Ginting. Two-grid finite volume element method for linear and nonlinear elliptic problems. *Numer. Math.*, 108(2):177–198, 2007.

[30] C. Bi and H. Rui. Uniform convergence of finite volume element method with Crouzeix-Raviart element for non-self-adjoint and indefinite elliptic problems. *J. Comput. Appl. Math.*, 200(2):555–565, 2007.

[31] H. P. Bourgade and F. Filbet. Convergence of a finite volume scheme for coagulation-fragmentation equations. *Math. Comp.*, 77(262):851–882, 2008.

[32] I. Boyd, D. VanGilde, and X. Liu. Monte carlo simulation of neutral xenon flows in electric propulsion devices. *J. Propulsion and Power*, 14(6):1009–1015, 1998.

[33] D. Braess. *Finite elements*. Cambridge University Press, 1997.

[34] J. H. Bramble and J. T. King. A finite element method for interface problems in domains with smooth boundary and interfaces. *Adv. Comput. Math.*, 6:109–138, 1996.

[35] R. Bustinza, G. N. Gatica, and B. Cockburn. An a posteriori error estimate for the local discontinuous Galerkin method applied to linear and nonlinear diffusion problems. *J. Sci. Comput.*, 22/23:147–185, 2005.

[36] Z. Cai. On the finite volume element method. *Numer. Math.*, 58(7):713–735, 1991.

[37] Z. Cai and S. McCormick. On the accuracy of the finite volume element method for diffusion equations on composite grids. *SIAM J. Numer. Anal.*, 27(3):636–655, 1990.

[38] C. Calgaro, E. Creuse, and T. Goudon. An hybrid finite volume-finite element method for variable density incompressible flows. *J. Comput. Phys.*, 227(9):4671–4696, 2008.

[39] D. Calhoun. A Cartesian grid method for solving the two-dimensional Streamfunction-Vorticity equations in irregular regions. *J. Comput. Phys.*, 176:231–275, 2002.

[40] B. Camp, T. Lin, Y. Lin, and W. Sun. Quadratic immersed finite element spaces and their approximation capabilities. *Adv. Comput. Math.*, 24(1-4):81–112, 2006.

[41] Y. Cao, X.-M. He, and T. Lü. An algorithm using finite volume element method and its splitting extrapolation. *submitted*.

[42] Y. Cao, X.-M. He, and T. Lü. A splitting extrapolation for solving nonlinear elliptic equations with d-quadratic finite elements. *J. Comput. Phys.*, 228(1):109–122, 2009.

[43] P. Castillo. Performance of discontinuous Galerkin methods for elliptic PDEs. *SIAM J. Sci. Comput.*, 24(2):524–547, 2002.

[44] P. Castillo, B. Cockburn, I. Perugia, and D. Schötzau. An a priori error analysis of the local discontinuous Galerkin method for elliptic problems. *SIAM J. Numer. Anal.*, 38(5):1676–1706, 2000.

[45] P. Castillo, B. Cockburn, D. Schötzau, and C. Schwab. Optimal a priori error estimates for the *hp*-version of the local discontinuous Galerkin method for convection-diffusion problems. *Math. Comp.*, 71(238):455–478, 2002.

[46] P. Chatzipantelidis and R. D. Lazarov. Error estimates for a finite volume element method for elliptic PDEs in nonconvex polygonal domains. *SIAM J. Numer. Anal.*, 42(5):1932–1958, 2005.

[47] H. Chen. Pointwise error estimates of the local discontinuous Galerkin method for a second order elliptic problem. *Math. Comp.*, 74(251):1097–1116, 2005.

[48] H. Chen and Z. Chen. Stability and convergence of mixed discontinuous finite element methods for second-order differential problems. *J. Numer. Math.*, 11(4):253–287, 2003.

[49] H. Chen, Z. Chen, and B. Li. Numerical study of the *hp* version of mixed discontinuous finite element methods for reaction-diffusion problems: the 1D case. *Numer. Meth. Part. Differ. Equ.*, 19(4):525–553, 2003.

[50] Z. Chen. On the relationship of various discontinuous finite element methods for second-order elliptic equations. *East-West J. Numer. Math.*, 9(2):99–122, 2001.

[51] Z. Chen. *Finite element methods and their applications. Scientific Computation.* Springer-Verlag, Berlin, 2005.

[52] Z. Chen and H. Chen. Pointwise error estimates of discontinuous Galerkin methods with penalty for second-order elliptic problems. *SIAM J. Numer. Anal.*, 42(3):1146–1166, 2004.

[53] Z. Chen, R. Li, and A. Zhou. A note on the optimal l^2-estimate of the finite volume element method. *Adv. Comput. Math.*, 16(4):291–303, 2002.

[54] Z. Chen and J. Zou. Finite element methods and their convergence for elliptic and parabolic interface problems. *Numer. Math.*, 79:175–202, 1998.

[55] A. J. Chorin. A numerical method for solving incompressible viscous flow problems. *J. Comput. Phys*, 2:12, 1967.

[56] A. J. Chorin. Numerical solution of the Navier-Stokes equations. *Math. Comp.*, 22:745–762, 1968.

[57] A. J. Chorin. On the convergence of discrete approximations to the Navier-Stokes equations. *Math. Comp*, 23:341–353, 1969.

[58] A. J. Chorin. Numerical solutions of incompressible flow problems. *Stud. Numer. Anal.*, 2:64–71, 1970.

[59] S. H. Chou and X. Ye. Superconvergence of finite volume methods for the second order elliptic problem. *Comput. Methods Appl. Mech. Engrg.*, 196(37-40):3706–3712, 2007.

[60] B. Cockburn and C. Dawson. Approximation of the velocity by coupling discontinuous Galerkin and mixed finite element methods for flow problems. Locally conservative numerical methods for flow in porous media. *Comput. Geosci.*, 6(3-4):505–522, 2002.

[61] B. Cockburn and J. Gopalakrishnan. A characterization of hybridized mixed methods for second order elliptic problems. *SIAM J. Numer. Anal.*, 42(1):283–301, 2004.

[62] B. Cockburn and J. Gopalakrishnan. Error analysis of variable degree mixed methods for elliptic problems via hybridization. *Math. Comp.*, 74(252):1653–1677, 2005.

[63] B. Cockburn, J. Gopalakrishnan, and R. Lazarov. Unified hybridization of discontinuous Galerkin, mixed and continuous Galerkin methods for second order elliptic problems. *submitted to SIAM J. Numer. Anal.*

[64] B. Cockburn and J. Guzmán. Error estimates for the Runge-Kutta discontinuous Galerkin method for the transport equation with discontinuous initial data. *SIAM J. Numer. Anal.*, 46(3):1364–1398, 2008.

[65] B. Cockburn, S. C. Hou, and C.-W. Shu. The Runge-Kutta local projection discontinuous Galerkin finite element method for conservation laws. IV. The multidimensional case. *Math. Comp.*, 54(190):545–581, 1990.

[66] B. Cockburn, G. Kanschat, I. Perugia, and D. Schötzau. Superconvergence of the local discontinuous Galerkin method for elliptic problems on Cartesian grids. *SIAM J. Numer. Anal.*, 39(1):264–285, 2001.

[67] B. Cockburn, G. Kanschat, and D. Schötzau. A locally conservative LDG method for the incompressible Navier-Stokes equations. *Math. Comp.*, 74(251):1067–1095, 2005.

[68] B. Cockburn, S. Y. Lin, and C.-W. Shu. TVB Runge-Kutta local projection discontinuous Galerkin finite element method for conservation laws. III. one-dimensional systems. *J. Comput. Phys.*, 84(1):90–113, 1989.

[69] B. Cockburn and C.-W. Shu. TVB Runge-Kutta local projection discontinuous Galerkin finite element method for conservation laws. II. General framework. *Math. Comp.*, 52(186):411–435, 1989.

[70] B. Cockburn and C.-W. Shu. The Runge-Kutta local projection p_1-discontinuous-Galerkin finite element method for scalar conservation laws. *RAIRO Model. Math. Anal. Numer.*, 25(3):337–361, 1991.

[71] B. Cockburn and C.-W. Shu. The local discontinuous Galerkin method for time-dependent convection-diffusion systems. *SIAM J. Numer. Anal.*, 35(6):2440–2463, 1998.

[72] B. Cockburn and C.-W. Shu. The Runge-Kutta discontinuous Galerkin method for conservation laws. V. Multidimensional systems. *J. Comput. Phys.*, 141(2):199–224, 1998.

[73] B. Cockburn and C.-W. Shu. Runge-Kutta discontinuous Galerkin methods for convection-dominated problems. *J. Sci. Comput.*, 16(3):173–261, 2001.

[74] A. Dadone and B. Grossman. Progressive optimization of inverse fluid dynamic design problems. *Comput. Fluids*, 29:1–32, 2000.

[75] A. Dadone and B. Grossman. An immersed body methodology for inviscid flows on cartesian grids. In *AIAA*, 2002-1059.

[76] R. Dautray and J.-L. Lions. *Mathematical analysis and numerical methods for science and techonology*, volume 3. Springer-Verlag, New York, 1990.

[77] M. Delves and C. A. Hall. An implicit matching principle for global element calculations. *J. Inst. Math. Appl*, 23:223–234, 1979.

[78] V. Dolejší. Semi-implicit interior penalty discontinuous Galerkin methods for viscous compressible flows. *Commun. Comput. Phys.*, 4(2):231–274, 2008.

[79] O. Drblikova and K. Mikula. Convergence analysis of finite volume scheme for nonlinear tensor anisotropic diffusion in image processing. *SIAM J. Numer. Anal.*, 46(1):37–60, 2007/08.

[80] M. Dumbser, D. S. Balsara, E. F. Toro, and C. D. Munz. A unified framework for the construction of one-step finite volume and discontinuous Galerkin schemes on unstructured meshes. *J. Comput. Phys.*, 227(18):8209–8253, 2008.

[81] M. Dumbser, M. Käser, V. A. Titarev, and E. F. Toro. Quadrature-free non-oscillatory finite volume schemes on unstructured meshes for nonlinear hyperbolic systems. *J. Comput. Phys.*, 226(1):204–243, 2007.

[82] R. E. Ewing, R. Lazarov, T. Lin, and Y. Lin. Mortar finite volume element approximations of second order elliptic problems. *East-West J. Numer. Math.*, 8(2):93–110, 2000.

[83] R. E. Ewing, R. Lazarov, and Y. Lin. Finite volume element approximations of nonlocal in time one-dimensional flows in porous media. *Computing*, 64(2):157–182, 2000.

[84] R. E. Ewing, Z. Li, T. Lin, and Y. Lin. The immersed finite volume element methods for the elliptic interface problems. Modelling '98 (prague). *Math. Comput. Simulation*, 50(1-4):63–76, 1999.

[85] R. E. Ewing, T. Lin, and Y. Lin. On the accuracy of the finite volume element method based on piecewise linear polynomials. *SIAM J. Numer. Anal.*, 39(6):1865–1888, 2002.

[86] R. E. Ewing and H. Wang. A summary of numerical methods for time-dependent advection-dominated partial differential equations. *J. Comput. Appl. Math.*, 128:423–445, 2001.

[87] R. E. Ewing, H. Wang, and T.F. Russell. Eulerian-Lagrangian localized adjoint methods fo convection-diffusion equations and their convergence analysis. *IMA J. Numer. Anal.*, 15:405–495, 1995.

[88] R. S. Falk and G. R. Richter. Explicit finite element methods for symmetric hyperbolic equations. *SIAM J. Numer. Anal.*, 36(3):935–952, 1999.

[89] K. Fan, W. Cai, and X. Ji. A generalized discontinuous Galerkin (GDG) method for schrödinger equations with nonsmooth solutions. *J. Comput. Phys.*, 227(4):2387–2410, 2008.

[90] S. Faure, J. Laminie, and R. Temam. Colocated finite volume schemes for fluid flows. *Commun. Comput. Phys.*, 4(1):1–25, 2008.

[91] M. Feistauer and A. Ženišek. Finite element solution of nonlinear elliptic problems. *Numer. Math.*, 50(4):451–475, 1987.

[92] M. Feistauer and V. Kucera. On a robust discontinuous Galerkin technique for the solution of compressible flow. *J. Comput. Phys.*, 224(1):208–221, 2007.

[93] A. L. Fogelson and J. P. Keener. Immersed interface methods for neumann and related problems in two and three dimensions. *SIAM J. Sci. Comput.*, 22:1630–1654, 2001.

[94] F. Gao, Y. Yuan, and D. Yang. An upwind finite-volume element scheme and its maximum-principle-preserving property for nonlinear convection-diffusion problem. *Int. J. Numer. Meth. Fluids*, 56(12):2301–2320, 2008.

[95] A. Gersborg-Hansen, M. P. Bendsøe, and O. Sigmund. Topology optimization of heat conduction problems using the finite volume method. *Struct. Multidisc. Optim.*, 31:251–259, 2006.

[96] V. Girault, B. Riviere, and M. F. Wheeler. A discontinuous Galerkin method with nonoverlapping domain decomposition for the Stokes and Navier-Stokes problems. *Math. Comp.*, 74(249):53–84, 2005.

[97] V. Girault, S. Sun, M. F. Wheeler, and I. Yotov. Coupling discontinuous Galerkin and mixed finite element discretizations using mortar finite elements. *SIAM J. Numer. Anal.*, 46(2):949–979, 2008.

[98] Y. Gong, B. Li, and Z. Li. Immersed-interface finite-element methods for elliptic interface problems with non-homogeneous jump conditions. *SIAM J. Numer. Anal.*, 46:472–495, 2008.

[99] J. Gopalakrishnan and G. Kanschat. A multilevel discontinuous Galerkin method. *Numer. Math.*, 95(3):527–550, 2003.

[100] M. J. Grote, A. Schneebeli, and D. Schötzau. Interior penalty discontinuous Galerkin method for Maxwell 's equations: energy norm error estimates. *J. Comput. Appl. Math.*, 204(2):375–386, 2007.

[101] T. Gudi, N. Nataraj, and P. Pani. *hp*-discontinuous Galerkin methods for strongly nonlinear elliptic boundary value problems. *Numer. Math.*, 109(2):233–268, 2008.

[102] T. Gudi and A. K. Pani. Discontinuous Galerkin methods for quasi-linear elliptic problems of nonmonotone type. *SIAM J. Numer. Anal.*, 45(1):163–192, 2007.

[103] J. Guzmán. Pointwise error estimates for discontinuous Galerkin methods with lifting operators for elliptic problems. *Math. Comp.*, 75(255):1067–1085, 2006.

[104] J. Guzmán. Local and pointwise error estimates of the local discontinuous Galerkin method applied to the Stokes problem. *Math. Comp.*, 77(263):1293–1322, 2008.

[105] J. Haink and C. Rohde. Local discontinuous-Galerkin schemes for model problems in phase transition theory. *Commun. Comput. Phys.*, 4(4):860–893, 2008.

[106] P. Hansbo and M. G. Larson. Piecewise divergence-free discontinuous Galerkin methods for Stokes flow. *Comm. Numer. Meth. Engrg.*, 24(5):355–366, 2008.

[107] R. Hartmann and P. Houston. Adaptive discontinuous Galerkin finite element methods for the compressible Euler equations. *J. Comput. Phys.*, 183(2):508–532, 2002.

[108] J. Haslinger and R. A. E. Mäkinen. *Introduction to shape optimization: theory, approximation, and computation*, volume 7 of *Advances in Design and Control*. SIAM, Philadelphia, 2003.

[109] J. Haslinger and P. Neittaamäki. *Finite element approximation for optimal shape, material and topology design*. John Wiley & Sons, Ltd, Chichester, 1996.

[110] X.-M. He, T. Lin, and Y. Lin. The convergence of the bilinear and linear immersed finite element solutions to interface problems. *submitted*.

[111] X.-M. He, T. Lin, and Y. Lin. A selective immersed discontinuous Galerkin method for elliptic interface problems. *submitted*.

[112] X.-M. He, T. Lin, and Y. Lin. Approximation capability of a bilinear immersed finite element space. *Numer. Meth. Part. Differ. Equ.*, 24(5):1265–1300, 2008.

[113] X.-M. He, T. Lin, and Y. Lin. A bilinear immersed finite volume element method for the diffusion equation with discontinuous coefficients. *Commun. Comput. Phys.*, 6(1):185–202, 2009.

[114] X.-M. He and T. Lü. A splitting extrapolation method for second order hyperbolic equations. *submitted*.

[115] X.-M. He and T. Lü. Splitting extrapolation method for solving second order parabolic equations with curved boundaries by using domain decomposition and d-quadratic isoparametric finite elements. *Int. J. Comput. Math.*, 84(6):767–781, 2007.

[116] X.-M. He, T. Lü, and J. Wei. On a multi-variable asymptotic error expansion and the finite element splitting extrapolation for nonlinear parabolic and hyperbolic equations. *in preparation*.

[117] B. Heinrich. *Finite difference methods on irregular networks*, volume 82 of *International Series of Numerical Mathematics*. Birkhäuser, Boston, 1987.

[118] G. Hetzer and A. J. Meir. On an interface problem with a nonlinear jump condition, numerical approximation of solutions. *Int. J. Numer. Anal. Model.*, 4(3-4):519–530, 2007.

[119] D. W. Hewitt. The embedded curved boundary method for orthogonal simulation meshes. *J. Comput. Phys.*, 138:585–616, 1997.

[120] T. Y. Hou and B. T. R. Wetton. Second order convergence of a projection scheme for the incompressible Navier-Stokes equations with boundaries. *SIAM J. Numer. Anal.*, 30:609–629, 1993.

[121] J. Huang and T. Lü. Splitting extrapolations for solving boundary integral equations of linear elasticity dirichlet problems on polygons by mechanical quadrature methods. *J. Comput. Math.*, 24(1):9–18, 2006.

[122] T. J. R. Hughes, G. Engel, L. Mazzei, and M. G. Larson. A comparison of discontinuous and continuous Galerkin methods based on error estimates, conservation, robustness and efficiency, in Discontinuous Galerkin Methods, Theory, Computation and Applications. *Lecture Notes in Computational Science and Engineering*, 11:135–146, 2000.

[123] D. M. Ingram, D. M. Causon, and C. G. Mingham. Developments in Cartesian cut cell methods. *Math. Comput. Simulation*, 61(3-6):561–572, 2003.

[124] A. Jameson. The construction of discretely conservative finite volume schemes that also globally conserve energy or entropy. *J. Sci. Comput.*, 34(2):152–187, 2008.

[125] G.-W. Jang, S. Lee, Y. Y. Kim, and D. Sheen. Topology optimization using non-conforming finite elements: three-dismensional case. *Int. J. Numer. Meth. Engng*, 63(6):859–875, 2005.

[126] C. Johnson. *Numerical solution of partial differential equations by the finite element method*. Cambridge University Press, Cambridge, 1987.

[127] C. Johnson and J. Pitkäranta. An analysis of the discontinuous Galerkin method for a scalar hyperbolic equation. *Math. Comp.*, 46(173):1–26, 1986.

[128] J. Douglas Jr. and T. Dupont. Interior penalty procedures for elliptic and parabolic Galerkin methods. *Lecture Notes in Physics*, 58:207–216, 1976.

[129] R. Kafafy, T. Lin, Y. Lin, and J. Wang. Three-dimensional immersed finite element methods for electric field simulation in composite materials. *Int. J. Numer. Meth. Engrg.*, 64(7):940–972, 2005.

[130] R. Kafafy, J. Wang, and T. Lin. A hybrid-grid immersed-finite-element particle-in-cell simulation model of ion optics plasma dynamics. *Dyn. Contin. Discrete Impuls. Syst. Ser. B Appl. Algorithms*, 12:1–16, 2005.

[131] G. Kanschat. Block preconditioners for LDG discretizations of linear incompressible flow problems. *J. Sci. Comput.*, 22/23:371–384, 2005.

[132] O. A. Karakashian and F. Pascal. Convergence of adaptive discontinuous Galerkin approximations of second-order elliptic problems. *SIAM J. Numer. Anal.*, 45(2):641–665, 2007.

[133] M. G. Larson and A. J. Niklasson. Analysis of a family of discontinuous Galerkin methods for elliptic problems: the one dimensional case. *Numer. Math.*, 99(1):113–130, 2004.

[134] P. Lasaint and P. A. Raviart. On a finite element method for solving the neutron transport equation. *Mathematical aspects of finite elements in partial differential equations(Proc. Sympos., Math. Res. Center, Univ. Wisconsin, Madison, Wis., 1974), Academic Press, New York*, pages 89–123, 1974.

[135] C. Lasser and A. Toselli. An overlapping domain decomposition preconditioner for a class of discontinuous Galerkin approximations of advection-diffusion problems. *Math. Comp.*, 72(243):1215–1238, 2003.

[136] D. V. Le, B. C. Khoo, and J. Peraire. An immersed interface method for viscous incompressible flows involving rigid and flexible boundaries. *J. Comput. Phys.*, 220(6):109–138, 2006.

[137] R. J. LeVeque and Z. Li. The immersed interface method for elliptic equations with discontinuous coefficients and singular sources. *SIAM J. Numer. Anal.*, 34:1019–1044, 1994.

[138] J. Li, Y. Renardy, and M. Renardy. Numerical simulation of breakup of a viscous drop in simple shear flow through a volume-of-fluid method. *Phys. Fluids*, 12(2):269–282, 2000.

[139] R. Li, Z. Chen, and W. Wu. *Generalized difference methods for differential equations. Numerical analysis of finite volume methods. Monographs and Textbooks in Pure and Applied Mathematics.* 226. Marcel Dekker, 2000, New York.

[140] R. Li and T. Tang. Moving mesh discontinuous Galerkin method for hyperbolic conservation laws. *J. Sci. Comput.*, 27(1-3):347–363, 2006.

[141] Z. Li. The immersed interface method using a finite element formulation. *Appl. Numer. Math.*, 27(3):253–267, 1997.

[142] Z. Li and K. Ito. *The immersed interface method: Numerical solutions of PDEs involving interfaces and irregular domains. Frontiers in Applied Mathematics, 33.* SIAM, Philadelphia, PA, 2006.

[143] Z. Li, T. Lin, Y. Lin, and R. C. Rogers. An immersed finite element space and its approximation capability. *Numer. Meth. Part. Differ. Equ.*, 20(3):338–367, 2004.

[144] Z. Li, T. Lin, and X. Wu. New Cartesian grid methods for interface problems using the finite element formulation. *Numer. Math.*, 96(1):61–98, 2003.

[145] X. Liao and A. Zhou. A multi-parameter splitting extrapolation and a parallel algorithm for elliptic eigenvalue problem. *J. Comput. Math.*, 16(3):213–220, 1998.

[146] C. B. Liem, Tao Lü, and T. M. Shin. *The splitting extrapolation method. A new technique in numerical solution of multidimensional problems. With a preface by Zhongci Shi. Series on Applied Mathematics, 7.* World Scientific Publishing Co., Inc., River Edge, NJ, 1995.

[147] Q. Lin and T. Lü. The splitting extrapolation method for multidimensional problem. *J. Comput. Math.*, 1:45–51, 1983.

[148] Q. Lin and Q. D. Zhu. Undirectional extrapolations of finite difference and finite elements (in Chinese). *Engrg. Math.*, 1:1–12, 1984.

[149] T. Lin, Y. Lin, R. C. Rogers, and L. M. Ryan. A rectangular immersed finite element method for interface problems. In P. Minev and Y. Lin, editors, *Advances in Computation: Theory and Practice, Vol. 7*, pages 107–114. Nova Science Publishers, Inc., 2001.

[150] T. Lin, Y. Lin, and W. Sun. Error estimation of a class of quadratic immersed finite element methods for elliptic interface problems. *Discrete Contin. Dyn. Syst. Ser. B*, 7(4):807–823, 2007.

[151] T. Lin and J. Wang. An immersed finite element electric field solver for ion optics modeling. In *Proceedings of AIAA Joint Propulsion Conference, Indianapolis, IN, July, 2002*. AIAA, 2002-4263.

[152] T. Lin and J. Wang. The immersed finite element method for plasma particle simulation. In *Proceedings of AIAA Aerospace Sciences Meeting, Reno, NV, Jan., 2003*. AIAA, 2003-0842.

[153] Y. Liu, C.-W. Shu, E. Tadmor, and M. Zhang. Non-oscillatory hierarchical reconstruction for central and finite volume schemes. *Commun. Comput. Phys.*, 2(5):933–963, 2007.

[154] E. Lorin, A. Ali, and A. Soulaimani. A positivity preserving finite element-finite volume solver for the Spalart-Allmaras turbulence model. *Comput. Methods Appl. Mech. Engrg.*, 196(17-20):2097–2116, 2007.

[155] T. Lu and W. Cai. A fourier spectral-discontinuous Galerkin method for time-dependent 3-d Schrödinger-Poisson equations with discontinuous potentials. *J. Comput. Appl. Math.*, 220(1-2):588–614, 2008.

[156] T. Lü and Y. Feng. Splitting extrapolation based on domain decomposition for finite element approximations. *Sci. China Ser. E*, 40(2):144–155, 1997.

[157] T. Lü and J. Lu. Splitting extrapolation for solving second order elliptic systems with curved boundary in $\Re^d$ by using d-quadratic isoparametric finite element. *Appl. Numer. Math.*, 40(4):467–481, 2002.

[158] T. Lü, T. M. Shin, and C. B. Liem. *Splitting extrapolation and combination techniques(in Chinese)*. Scientific Press, Beijing, 1998.

[159] H. Man and Z. Shi. p_1-nonconforming quadrilateral finite volume element method and its cascadic multigrid algorithm for elliptic problems. *J. Comput. Math.*, 24(1):59–80, 2006.

[160] G. Manzini and A. Russo. A finite volume method for advection-diffusion problems in convection-dominated regimes. *Comput. Methods Appl. Mech. Engrg.*, 197(13-16):1242–1261, 2008.

[161] W. McLean. *Strongly elliptic systems and boundary integral equations*. Cambridge University Press, Cambridge, 2000.

[162] R.C. McOwen. *Partial differential equations: Methods and applications*. Prentice Hall, New Jersey, 2nd edition, 2003.

[163] N. Moës, J. Dolbow, and T. Belytschko. A finite element method for crack growth without remeshing. *Int. J. Numer. Meth. Eng.*, 46(1):131–150, 1999.

[164] I. Mozolevski, E. Süli, and P. R. Bösing. *hp*-version a priori error analysis of interior penalty discontinuous Galerkin finite element approximations to the biharmonic equation. *J. Sci. Comput.*, 30(3):465–491, 2007.

[165] Y. Muravlev and A. Shagayda. Numerical modelling of extraction systems in ion thrusters. In *IEPC-99-162*, 1999.

[166] M. Nakano and Y. Arakawa. Ion thruster lifetime estimation and modeling using computer simulation. In *IEPC-99-145*, 1999.

[167] Y. Nakayama and P. Wilbur. Numerical simulation of ion beam optics for many-grid systems. In *AIAA-2001-3782*, 2001.

[168] P. Neittaanmäki and Q. Lin. Acceleration of the convergence in finite difference methods by predictor corrector and splitting extrapolation methods. *J. Comput. Math.*, 5:181–190, 1987.

[169] J. T. Oden, I. Babuska, and C. E. Baumann. A discontinuous *hp* finite element method for diffusion problems. *J. Comput. Phys*, 146:491–519, 1998.

[170] Y. Okawa and H. Takegahara. Particle simulation on ion beam extraction phenomena in an ion thruster. In *IEPC-99-146*, 1999.

[171] X. Peng, W. Ruyten, V. Friedly, D. Keefer, and Q. Zhang. Particle simulation of ion optics and grid erosion for two-grid and three-grid systems. *Rev. Sci. Instrum.*, 65(5):1770, 1994.

[172] J. Peraire and P.-O. Persson. The compact discontinuous Galerkin (CDG) method for elliptic problems. *SIAM J. Sci. Comput.*, 30(4):1806–1824, 2008.

[173] I. Perugia and D. Schötzau. On the coupling of local discontinuous Galerkin and conforming finite element methods. *J. Sci. Comput.*, 16(2001)(4):411–433, 2002.

[174] C. S. Peskin. Flow patterns around heart valves. *J. Comput. Phys.*, 10:252–271, 1972.

[175] C. S. Peskin. Numerical analysis of blood flow in the heart. *J. Comput. Phys.*, 25:220–252, 1977.

[176] T. E. Peterson. A note on the convergence of the discontinuous Galerkin method for a scalar hyperbolic equation. *SIAM J. Numer. Anal.*, 28(1):133–140, 1991.

[177] J. Qiu, T. Liu, and B. C. Khoo. Simulations of compressible two-medium flow by Runge-Kutta discontinuous Galerkin methods with the ghost fluid method. *Commun. Comput. Phys.*, 3(2):479–504, 2008.

[178] A. Quarteroni and A. Valli. *Numerical approximation of partial differential equations*, volume 23 of *Computational Mathematics*. Springer-Verlag, New York, 2nd edition, 1997.

[179] W. H. Reed and T. R. Hill. Triangular mesh methods for the neutron transport equation. *Tech. Report No. LA-UR-73-479, Los Alamos Scientific Laboratory*, 1973.

[180] X. Ren and J. Wei. On a two-dimensional elliptic problem with large exponent in nonlinearity. *Trans. Amer. Math. Soc.*, 343(2):749–763, 1994.

[181] J. J. Rennilson and D. R. Criswell. Surveyor observations of lunar horizon-glow. *Earth, Moon, and Planets*, 10(2):121–142, 1974.

[182] G. R. Richter. An optimal-order error estimate for the discontinuous Galerkin method. *Math. Comp.*, 50(181):75–88, 1988.

[183] B. Rivière, M. F. Wheeler, and V. Girault. Improved energy estimates for interior penalty, constrained and discontinuous Galerkin methods for elliptic problems Part 1. *Comput. Geosci.*, 3:337–360, 1999.

[184] U. Rüde. Book review: The splitting extrapolation method (C. B. Liem, T. Lü and T. M. Shih). *SIAM Review*, 39:161–162, 1997.

[185] U. Rüde and A. Zhou. Multi-parameter extrapolation methods for boundary integral equations. Numerical treatment of boundary integral equations. *Adv. Comput. Math.*, 9(1-2):173–190, 1998.

[186] A. A. Samarskiĭ and V. B. Andreev. *Méthods aux différences pour équations elliptic-ques*. Mir, Moscow, 1978.

[187] S. A. Sauter and R. Warnke. Composite finite elements for elliptic boundary value problems with discontinuous coefficients. *Computing*, 77(1):29–55, 2006.

[188] O. Sigmund and J. Pertersson. Numerical instabilities in topology optimization: A survey on procedures dealing with checkerboards, mesh-dependencies and local minima. *Struct. Optim.*, 16:68–75, 1998.

[189] N. Sukemar, D. L. Chopp, N. Moës, and T. Belytschko. Modling holes and inclusions by level set in the extended finite-element method. *Comput. Meth. Appl. Mech. Eng.*, 190:6183–6200, 2001.

[190] T. Sun and D. Yang. Error estimates for a discontinuous Galerkin method with interior penalties applied to nonlinear Sobolev equations. *Numer. Meth. Part. Differ. Equ.*, 24(3):879–896, 2008.

[191] M. Tartz. Validation of a grid-erosion simulation by short-time erosion measurements. In *IEPC 99-147*, 1999.

[192] T. V. Voitovich and S. Vandewalle. Exact integration formulas for the finite volume element method on simplicial meshes. *Numer. Meth. Part. Differ. Equ.*, 23(5):1059–1082, 2007.

[193] C. Wang, H. Tang, and T. Liu. An adaptive ghost fluid finite volume method for compressible gas-water simulations. *J. Comput. Phys.*, 227(12):6385–6409, 2008.

[194] H. Wang. An improved WPE method for solving discontinuous Fokker-Planck equations. *Int. J. Numer. Anal. Model.*, 5(1):1–23, 2008.

[195] J. Wang, X.-M. He, and Y. Cao. Modeling spacecraft charging and charged dust particle interactions on lunar surface. *Proceedings of the 10th Spacecraft Charging Technology Conference, Biarritz, France*, 2007.

[196] J. Wang, X.-M. He, and Y. Cao. Modeling electrostatic levitation of dusts on lunar surface. *IEEE Trans. Plasma Sci.*, 36(5):2459–2466, 2008.

[197] K. Wang. A uniform optimal-order estimate for an Eulerian-Lagrangian discontinuous Galerkin method for transient advection-diffusion equations. *Numer. Meth. Part. Differ. Equ.*, 25(1):87–109, 2009.

[198] T. Wang. Alternating direction finite volume element methods for 2D parabolic partial differential equations. *Numer. Meth. Part. Differ. Equ.*, 24(1):24–40, 2008.

[199] T. S. Wang. A Hermite cubic immersed finite element space for beam design problems. Master's thesis, Virginia Polytechnic Institute and State University, 2005.

[200] M. F. Wheeler. An elliptic collocation-finite element method with interior penalties. *SIAM J. Numer. Anal*, 15:152–161, 1978.

[201] J. Wloka. *Partial differential equations*. Cambridge University Press, Cambridge, 1987.

[202] Y. Xu and C.-W. Shu. Local discontinuous Galerkin method for the Hunter-Saxton equation and its zero-viscosity and zero-dispersion limits. *SIAM J. Sci. Comput.*, 31(2):1249–1268, 2008/09.

[203] X. Ye. Analysis and convergence of finite volume method using discontinuous bilinear functions. *Numer. Meth. Part. Differ. Equ.*, 24(1):335–348, 2008.

[204] A. Zhou, C. B. Liem, T. M. Shih, and T. Lü. A multi-parameter splitting extrapolation and a parallel algorithm. *Syst. Sci. Math. Sci.*, 10(3):253–260, 1997.

[205] Y. C. Zhou and G. W. Wei. On the fictitious-domain and interpolation formulations of the matched interface and boundary (MIB) method. *J. Comput. Phys*, 219(1):228–246, 2006.

[206] Y. C. Zhou, S. Zhao, M. Feig, and G. W. Wei. High order matched interface and boundary method for elliptic equations with discontinuous coefficients and singular sources. *J. Comput. Phys*, 213(1):1–30, 2006.

Addendum for Ph.D. dissertation "Bilinear Immersed Finite Elements For Interface Problems"

Xiaoming He

Virginia Polytechnic Institute and State University

April 12, 2010

1. Change the Lemma 5.3.5 and its proof on pages 97-98 to be an assumption as follows:

   Now we assume the following trace inequality is true on $T \in \mathcal{T}_h$:

$$\left\| \beta \frac{\partial v}{\partial n} \right\|^2_{0,E_i(\partial T)} \leq C \left(\frac{1}{h_T} |v|^2_{1,T} + h_T |v|^2_{2,T} \right), \quad \forall v \in PH^2_{int}(T), \quad 1 \leq i \leq 4.$$

   This assumption is satisfied by $v \in H^2(T)$ according to the standard trace inequality. A function $v \in PH^2_{int}(T)$ can also satisfy this assumption, but the difficulty is to prove that the constant C is independent of the interface, which leads to some interesting future work.

2. Change the Lemma 8.4.1 and the line above it on page 143 to be an assumption as follows:

   For each element $T = \square A_1A_2A_3A_4 \in \mathcal{T}_h$, define

$$E_1(\partial T) = \overline{A_1A_2},\ E_2(\partial T) = \overline{A_2A_3},\ E_3(\partial T) = \overline{A_3A_4},\ E_4(\partial T) = \overline{A_4A_1}.$$

   Then we assume the following trace inequality is true on every $T \in \mathcal{T}_h$:

$$\left\| \beta \frac{\partial w_h}{\partial n} \right\|^2_{0,E_i(\partial T)} \leq C \left(h^{-1} |w_h|^2_{1,T} + h |w_h|^2_{2,T} \right), \quad \forall w_h \in S_h(T), \quad 1 \leq i \leq 4.$$

   If $|\beta| \geq b_1 > 0$, then

$$\left\| \frac{\partial w_h}{\partial n} \right\|^2_{0,E_i(\partial T)} \leq C \left(h^{-1} |w_h|^2_{1,T} + h |w_h|^2_{2,T} \right), \quad \forall w_h \in S_h(T), \quad 1 \leq i \leq 4.$$

   This assumption is satisfied by $w_h \in H^2(T)$ according to the standard trace inequality. A function $w_h \in S_h(T)$ can also satisfy this assumption, but the difficulty is to prove that the constant C is independent of the interface, which leads to some interesting future work.