Available online at www.sciencedirect.com
ScienceDirect
Comput. Methods Appl. Mech. Engrg. 404 (2023) 115744
Computer methods in applied mechanics and engineering
www.elsevier.com/locate/cma

# Error analysis of Petrov-Galerkin immersed finite element methods

Cuiyu He[a,*], Shun Zhang[b,1], Xu Zhang[a,2]

[a] *Department of Mathematics, Oklahoma State University, Stillwater, OK, 74078, United States of America*
[b] *Department of Mathematics, City University of Hong Kong, Kowloon Tong, Hong Kong, China*

Received 19 April 2022; received in revised form 23 October 2022; accepted 26 October 2022
Available online xxxx

**Abstract**

This paper designs and analyzes a new and stable Petrov–Galerkin (PG) immersed finite element method (IFEM) for the second-order elliptic interface problems by introducing stabilization terms based on the classical PG-IFEM, which lacks the local positivity. The Petrov–Galerkin immersed finite element method uses the immersed finite element functions for the trial space and the standard finite element functions for the test space. Both the *a priori* and *a posteriori* error estimates of the method are analyzed in this paper. We prove the continuity and inf-sup condition and the *a priori* error estimate of the energy norm. The proposed *a posteriori* error estimator is proved to be both reliable and efficient, with both reliability and efficiency constants independent of the location of the interface. Extensive numerical results confirm the numerical scheme's optimal convergence and indicate the robustness with respect to the interface-mesh intersection and the coefficient contrast, despite the robustness of the inf-sup constant with respect to the interface-mesh intersection has yet been theoretically proved.


*MSC:* 35R05; 65N15; 65N30

*Keywords:* Petrov–Galerkin; Immersed finite element; Interface problems; A priori error estimates; A posteriori error estimates

## 1. Introduction

In the past decades, there has been a growing interest in numerical simulations for real-world multi-physics problems, such as multi-phase flow, fluid–structure interaction, biology membranes, and protein simulations, to name a few. The numerical simulations are often carried out over domains consisting of multiple materials separated by interfacial surfaces. The governing models for the numerical simulations are the commonly used partial differential equations (PDE) of the interface type, widely known as interface problems.

There are generally two classes of numerical methods for interface problems: *fitted-mesh methods* and *unfitted-mesh methods*. The fitted-mesh methods, such as the classical finite element method (FEM), require the computational mesh to align with the interface; otherwise, the convergence and accuracy might be compromised. Such requirement often hinders applications with complex interfacial geometry or time-dependent problems with evolving

* Corresponding author.
*E-mail addresses:* cuiyu.he@okstate.edu (C. He), shun.zhang@cityu.edu.hk (S. Zhang), xzhang@okstate.edu (X. Zhang).

[1] The research of this author was partially supported by the Research Grants Council of the Hong Kong SAR, China under the GRF Grant Projects No. CityU 11302519 and CityU 11305319.

[2] The research of this author was partially supported by the National Science Foundation grant DMS-2110833 and the 2021 ORAU Ralph E. Powe Junior Faculty Enhancement Award.

https://doi.org/10.1016/j.cma.2022.115744
0045-7825/

interfaces. Because generating a high-quality three-dimensional mesh of a complicated geometry is extremely time-consuming. In contrast, unfitted-mesh methods, such as Extended FEM [1–4], Cut-FEM [5–8], unfitted Hybrid High-Order (HHO) method [9,10], Multiscale FEM [11–14] and Immersed FEM (IFEM) [15–18], etc., significantly alleviate the body-fitting restriction on the mesh generation, therefore, have gained enormous popularity recently.

The main idea of IFEM is to incorporate interface jump conditions in designing IFE shape functions around interfaces. The development of IFEM can be traced back to [15] for one-dimensional (1D) elliptic interface problems using piecewise linear polynomial approximations. Since then, the IFEM has been extended to high-order polynomial approximations [19,20], and multi-dimensional PDE problems [16–18,21–25]. Besides, the immersed idea has been applied in other computational frameworks such as nonconforming FEM [26,27], discontinuous Galerkin method [28,29], finite volume method [30,31], and weak Galerkin method [32].

For multi-dimensional interface problems, using the classical IFEM [16,17,23] can only guarantee a sub-optimal convergence order. This sub-optimality is caused by its lack of consistency since the IFE basis functions are not in the $H^1$-conforming subspace. One remedy is to use the partially-penalized IFEM [18], in which the consistency and stability terms are added on interface edges. The optimal convergence in the energy norm has been rigorously proved [18,33].

An alternative strategy to ensure consistency is to use standard finite element functions in the test space. We note that the approximation property is only required for the trial space. Therefore, we have some flexibility for the test space as long as the resulting linear system is stable, i.e., the resulting variational formulation satisfies the continuity and the inf-sup condition. The resulting numerical scheme lies in the Petrov–Galerkin framework due to the difference in test and trial functional spaces. Some applications of the Petrov–Galerkin method in multiscale finite element methods can be found in [34,35]. The $P_1$ Petrov–Galerkin type immersed finite element method has been developed and widely applied for 2D elliptic interface problems, e.g., see [36–39], and for 3D problems [40,41]. In [42], the authors developed a high-degree discontinuous PG-IFEM. The theoretical analysis of the inf-sup condition for the one-dimension case has been proved in [43]. Optimal convergence rates are observed for numerous numerical results in various dimensions. Theoretically, however, there is minimal analysis concerning the well-posedness and optimal convergence rates for the error to our knowledge.

This paper analyzes the stability of several PG-IFEM schemes for two-dimensional interface problems. The main difficulty is proving the inf-sup stability because the local stiffness matrix is not always positive due to arbitrary interface-mesh intersection. Note that when the jump of coefficients vanishes, the PG-IFEM evolves into the classical FEM, which is stable. This indicates that one can possibly obtain stability when the contrast is relatively small. Such stability results in interface problems with mild coefficient contrast ($< 16$) are proved in [42]. In this paper, without any restriction on the coefficient's contrast, we also prove the inf-sup stability for the classical PG-IFEM scheme under certain geometrical conditions of the interface.

Note that the positivity of local stiffness matrices is only a sufficient condition typically used to guarantee the inf-sup stability. Therefore, lacking this property does not necessarily indicate that the global system is not inf-sup stable. Indeed, though theoretically, the inf-sup stability has not been proved, extensive numerical results for the classical PG-IFEM have shown its effectiveness in solving elliptic interface problems.

Since we cannot manage to prove the unconditional stability of the classical PG-IFEM scheme, we develop a novel enhanced PG-IFEM with additional penalty terms for which we can theoretically establish the stability. Note that one can no longer directly apply the classical penalty based on the solution jump along the interface edges since the test function is conforming, nullifying the penalty term. The intuition is to penalize the distance between the test and trial spaces. To do so, we first project the test function in the classical FE space into the IFEM space by nodal interpolation and then penalize its solution jump along interface edges. We further add the penalization of the normal flux jump along the interface edges. This enhanced formulation is consistent, continuous, and inf-sup stable without any assumption on the mesh-interface intersection or the coefficient contrast. A priori error estimates are established in the energy norm. We note that the inf-sup constant in the a priori error estimation has yet been proved independent of the interface-mesh intersection in this paper. Nevertheless, extensive numerical results confirm the numerical scheme's optimal convergence and indicate robustness with respect to the interface-mesh intersection and the coefficient contrast. We also derive the residual-based a posteriori error estimator. Furthermore, the a posteriori error estimate is proved to be both reliable and efficient, with constants independent of the interface-mesh intersection.

Although the theoretical analysis is based on the enhanced PG-IFEM, extensive numerical results indicate that the classical PG-IFEM without any penalty performs as well as the enhanced PG-IFEM. The numerical performance

will not be compromised, plus the condition number is usually smaller. We believe our proposed PG-IFEM is of great interest to mechanical and aerospace engineering applications which seek a simple, accurate, and robust numerical scheme for their applications of interface problems [44–48]. The proposed enhanced PG-IFEM can be naturally applied to problems in three dimensions. Several numerical experiments of 3D interface problems are provided in Section 6.

The rest of the paper is organized as follows. In Section 2, we describe the interface problems and develop the enhanced Petrov–Galerkin IFEM. In Section 3, we analyze the stability of the classical and enhanced Petrov–Galerkin IFEMs. We analyze the *a priori* and *a posteriori* error estimates of the enhanced PG-IFEM in Sections 4 and 5, respectively. In Section 6, we report some numerical examples to demonstrate the features of these PG-IFEMs.

## 2. Interface problems and Petrov–Galerkin IFEM

Let $\Omega \subset \mathbb{R}^2$ be a polygonal domain with a Lipschitz boundary $\partial\Omega = \overline{\Gamma}_D \cup \overline{\Gamma}_N$, where $\Gamma_D \cap \Gamma_N = \emptyset$. Assume that $\text{meas}(\Gamma_D) > 0$. We consider the elliptic interface problem:

$$-\nabla \cdot (\beta \nabla u) = f \quad \text{in } \Omega^+ \cup \Omega^- \tag{2.1}$$

with boundary conditions

$$u = 0 \ \text{ on } \Gamma_D \quad \text{and} \quad -\beta \nabla u \cdot \mathbf{n} = g_N \ \text{ on } \Gamma_N.$$

Here, $f \in L^2(\Omega)$, $g_N \in L^2(\Gamma_N)$, and $\mathbf{n}$ is the unit vector outward normal to $\partial\Omega$. The zero Dirichlet boundary condition is assumed for simplicity. The method can be readily extended to non-homogeneous boundary conditions. The notations $\nabla$ and $\nabla\cdot$ are the gradient and divergence operators, respectively. Furthermore, assume that $\Omega$ is separated by a closed smooth interface curve $\Gamma$ into $\Omega^+$ and $\Omega^-$ such that $\overline{\Omega} = \overline{\Omega^+ \cup \Gamma \cup \Omega^-}$. The diffusion coefficient $\beta$ is assumed to be a positive piecewise constant function as follows

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

Denote by $r = \dfrac{\beta^+}{\beta^-}$ the ratio of the coefficient jump. The solution is assumed to satisfy the following interface jump conditions:

$$[\![u]\!]_\Gamma = 0 \quad \text{and} \quad [\![\beta \nabla u \cdot \mathbf{n}]\!]_\Gamma = 0, \tag{2.2}$$

where the jump of a function $v$ across the interface $\Gamma$ is defined by

$$[\![v]\!]_\Gamma = v^+|_\Gamma - v^-|_\Gamma.$$

We use the standard notations for the Sobolev spaces. Let

$$H^1_D(\Omega) = \{v \in H^1(\Omega) : v = 0 \text{ on } \Gamma_D\}.$$

Then the variational problem for (2.1) is to find $u \in H^1_D(\Omega)$ such that

$$a(u, v) := (\beta \nabla u, \nabla v) = (f, v) - \langle g_N, v \rangle_{\Gamma_N}, \quad \forall\, v \in H^1_D(\Omega), \tag{2.3}$$

where $(\cdot, \cdot)_\omega$ or $\langle \cdot, \cdot \rangle_\omega$ is the $L^2$-inner product on $\omega \subset \mathbb{R}^d$ or $\mathbb{R}^{d-1}$. The subscript $\omega$ is omitted when $\omega = \Omega$.

### 2.1. Triangulation and notations

In this paper, we only consider the triangular meshes in two dimensions. Let $\mathcal{T}_h$ be a triangulation of $\Omega$ that is regular and interface-independent. Thus, the mesh may not be interface-fitted. Denote the set of all vertices of the triangulation $\mathcal{T}_h$ by

$$\mathcal{N} := \mathcal{N}_I \cup \mathcal{N}_D \cup \mathcal{N}_N$$

where $\mathcal{N}_I$ is the set of all interior vertices, $\mathcal{N}_D$ and $\mathcal{N}_N$ are the sets of vertices on $\bar{\Gamma}_D$ and $\Gamma_N$, respectively. Denote the set of all edges of the triangulation $\mathcal{T}_h$ by

$$\mathcal{E} := \mathcal{E}_I \cup \mathcal{E}_D \cup \mathcal{E}_N$$

where $\mathcal{E}_I$ is the set of all interior edges, $\mathcal{E}_D$ and $\mathcal{E}_N$ are the sets of boundary edges on $\Gamma_D$ and $\Gamma_N$, respectively. For each element $K \in \mathcal{T}_h$, denote by $h_K$ the diameter of $K$, and by $\mathcal{N}_K$ and $\mathcal{E}_K$ the sets of all vertices and edges on $K$, respectively. Without loss of generality, we assume the interface $\Gamma$ and the background mesh $\mathcal{T}_h$ satisfy the following conditions

**(I)** The interface $\Gamma$ cannot intersect an edge of any element at more than two points unless the edge is part of $\Gamma$.

**(II)** If $\Gamma$ intersects the boundary of an element at two points, these intersection points must be on different edges of this element.

**(III)** The interface $\Gamma$ is a piecewise $C^2$-continuous function, and for any $K \in \mathcal{T}_h$, $\Gamma_K := \Gamma \cap K$ is a straight line segment if $\Gamma_K \neq \emptyset$.

We assume $\Gamma_K$ is a straight line for simplicity. In the general case, the smooth interface is approximated by a line segment for linear elements, and the corresponding approximation error will not be greater than $h_K^2$ in the $L^2$ norm, which preserves the optimal convergence for the numerical method. Based on the above assumptions, elements in $\mathcal{T}_h$ can be categorized into two classes: *non-interface elements* that either has no intersection with $\Gamma$ or $\Gamma \cap K \subset \partial K$, and *interface elements* whose interior is cut through by $\Gamma$. Denote the set of all interface elements by $\mathcal{T}_h^i$, and the set of non-interface elements by $\mathcal{T}_h^n = \mathcal{T}_h / \mathcal{T}_h^i$. For each $K \in \mathcal{T}_h^i$, we also let $K^+ = K \cap \Omega^+$ and $K^- = K \cap \Omega^-$.

For an edge $F \in \mathcal{E}$, if $F$ is cut through by $\Gamma$, i.e., $F \cap \Gamma \neq \emptyset$ and $F \not\subset \Gamma$, then $F$ is called an interface edge. We denote by $\mathcal{E}^i$ the set of all such interface edges. For each $F \in \mathcal{E}$, denote by $h_F$ the length of $F$. Denote by $\mathbf{n}_F = (n_1, n_2)$ and $\mathbf{t}_F = (-n_2, n_1)$ fixed unit vectors normal and tangential to $F$, respectively. Let $K_{F,1}$ and $K_{F,2}$ be the two elements sharing the common edge $F \in \mathcal{E}_I$ such that the unit vector out normal to $K_{F,1}$ coincides with $\mathbf{n}_F$. When $F \subset \partial\Omega$, $\mathbf{n}_F$ is the unit outward vector normal to $\partial\Omega$, and denote by $K_{F,1}$ the boundary element having the edge $F$. For a function $v$ that is defined on $K_{F,1} \cup K_{F,2}$, denote its traces on $F$ by $v_F^1$ and $v_F^2$ restricted on $K_{F,1}$ and $K_{F,2}$, respectively. Define the jump of a function $v$ on the edge $F$ by

$$[\![v]\!]_F = \begin{cases} v_F^1 - v_F^2, & \text{for } F \in \mathcal{E}_I, \\ v_F^1, & \text{for } F \in \mathcal{E}_D \cup \mathcal{E}_N \end{cases}$$

and the average of a function $v$ on the edge $F$ by

$$\{v\}_F = \begin{cases} \left(v_F^1 + v_F^2\right)/2, & \text{for } F \in \mathcal{E}_I, \\ v_F^1, & \text{for } F \in \mathcal{E}_D \cup \mathcal{E}_N. \end{cases}$$

It is easy to verify that

$$[\![vw]\!]_F = [\![v]\!]_F \{w\}_F + \{v\}_F [\![w]\!]_F, \quad \forall\, F \in \mathcal{E}. \tag{2.4}$$

For simplicity, we may drop the subscript $F$ in the notations $[\![\cdot]\!]_F$ and $\{\cdot\}_F$ if there is no confusion on where the jump and average are defined.

### 2.2. Petrov Galerkin IFEM

For simplicity, we assume that the interface does not intersect with the boundary, which indicates that $\mathcal{E}^i \subset \mathcal{E}_I$. For each interface element $K \in \mathcal{T}_h^i$, define the local IFE space by

$$\tilde{P}_1(K) = \left\{ v \in H^1(K) : \beta \nabla v \in H(\mathrm{div}, K),\, v|_{\bar{K}^\pm} \in P_1(\bar{K}^\pm) \right\}$$

where $P_1(w)$ is the space of all polynomial functions in the domain $w$ of degree no more than 1. We refer readers to [16–18] for more details about the construction of the linear IFE space $\tilde{P}_1(K)$. The global IFE space $\mathcal{S}(\mathcal{T}_h)$ is then defined to include all functions such that

(1) $v|_K \in \tilde{P}_1(K)$ for all $K \in \mathcal{T}_h^i$, $v|_K \in P_1(K)$ for all $K \in \mathcal{T}_h^n$, and
(2) $v$ is continuous at every vertex $z \in \mathcal{N}$.

Note that for each $z \in \mathcal{N}$, there exists a unique IFE nodal basis function [16,17], denoted by $\tilde{\lambda}_z \in \mathcal{S}(\mathcal{T}_h)$, such that

$$\tilde{\lambda}_z(z') = \delta_{zz'}, \quad \forall\, z' \in \mathcal{N}$$

where $\delta$ is the Kronecker delta function. We also define the classical $H^1$ conforming linear finite element space by $\mathcal{V}(\mathcal{T}_h)$, i.e.,

$$\mathcal{V}(\mathcal{T}_h) = \{v \in H^1(\Omega) : v|_K \in P_1(K)\ \forall K \in \mathcal{T}_h\}.$$

For each $v \in \mathcal{S}(\mathcal{T}_h)$, we define the discrete gradient operator $\nabla_h$ by

$$(\nabla_h v)|_K = \nabla(v|_K), \quad \forall\, K \in \mathcal{T}_h.$$

We now define two interpolation operators to the spaces $\mathcal{S}(\mathcal{T}_h)$ and $\mathcal{V}(\mathcal{T}_h)$,

$$\Pi(v) := \sum_{z\in\mathcal{N}} v(z)\tilde{\lambda}_z, \quad \mathcal{I}(v) := \sum_{z\in\mathcal{N}} v(z)\lambda_z,$$

respectively, where $\lambda_z$ is the classical nodal basis function of $z \in \mathcal{N}$.

The enhanced Petrov–Galerkin IFEM for the interface problem is to find $u_h \in \mathcal{S}_D(\mathcal{T}_h)$ such that

$$\tilde{a}_h(u_h, v) = (f, v) - \left\langle g_N, v\right\rangle_{\Gamma_N}, \forall v \in \mathcal{V}_D(\mathcal{T}_h) \tag{2.5}$$

where

$$\mathcal{S}_D(\mathcal{T}_h) := \left\{v \in \mathcal{S}(\mathcal{T}_h) \,:\, v|_{\Gamma_D} = 0\right\}, \quad \mathcal{V}_D(\mathcal{T}_h) = \{v \in \mathcal{V}(\mathcal{T}_h), v|_{\Gamma_D} = 0\},$$

and

$$\tilde{a}_h(u_h, v) := a_h(u_h, v) + \sum_{F\in\mathcal{E}^i} \int_F \left(\gamma_1 h_F^{-1}\beta[\![u_h]\!][\![\Pi v]\!] + \gamma_2 h_F \beta[\![\nabla u_h \cdot \mathbf{n}_F]\!][\![\nabla \Pi v \cdot \mathbf{n}_F]\!]\right) ds, \tag{2.6}$$

with

$$a_h(w, v) := \sum_{K\in\mathcal{T}_h} \int_K \beta \nabla w \cdot \nabla v \, dx. \tag{2.7}$$

It is easy to verify that the following error equation holds

$$\tilde{a}_h(u - u_h, v) = 0, \quad \forall v \in V_D(\mathcal{T}_h). \tag{2.8}$$

**Remark 2.1.** The classical PG-IFEM is a special case for the enhanced PG-IFEM, i.e., $\gamma_1 = \gamma_2 = 0$. The main difficulty of the stability analysis for the classical PG-IFEM is proving the inf-sup stability because the local stiffness matrix is not always positive definite due to arbitrary interface-mesh intersection [38,39,42]. Though this does not necessarily lead to the instability of the classical PG-IFEM, it is highly challenging, if not impossible, to analyze its stability theoretically. On the other hand, adding the penalty terms makes the analysis much more manageable.

## 3. Stability

In this section, we analyze the stability of the PG-IFEM. First, we will prove the conditional stability for the classical PG-IFEM without penalties under some geometric constraints. Next, we will prove the unconditional stability of the enhanced PG-IFEM.

### *3.1. Stability for the classical PG-IFEM*

In the following lemma, we prove a special case for which the classical PG-IFEM is stable. Recall that for the classical PG-IFEM, $\gamma_1 = \gamma_2 = 0$ in (2.6).

**Assumption 3.1.** Assume that for each $K \in \mathcal{T}_h^i$, there exists an edge $F \in \mathcal{E}_K$ such that $F$ is parallel to $\Gamma_K$.

**Lemma 3.1.** Assume that $\mathcal{T}_h$ satisfies *Assumption* 3.1. Then the bilinear form $a_h(\cdot,\cdot)$ defined in (2.7) satisfies the following inf-sup condition

$$a_h(v, \mathcal{I}v) \geq C\|\sqrt{\beta}\nabla v\|_{\Omega}^2 \quad \forall v \in \mathcal{S}(\mathcal{T}_h), \tag{3.1}$$

where the constant $C$ is independent of the jump of the coefficient and the interface-mesh intersection.

**Proof.** We prove (3.1) on the reference element. Let $K$ be the reference triangle and denote by $A_i$, $i = 1, 2, 3$ the vertices of $A$ (counterclockwise) with $A_1 = (0, 0)$ being the vertex on the right angle. For the first case, we consider $K$ as an interface element, and $D$, $E$ are the intersection points on the $x$ and $y$ axes, respectively. Assume that $D = (d, 0)$ and $E = (0, e)$. Assume that $A_1DE$ is the triangle in $\Omega^+$. The normal of the interface is chosen to be $\mathbf{n}_\Gamma|_K = (e, d)/\sqrt{d^2 + e^2}$.

Based on the fact that $\nabla v$ is piecewise constant, a simple calculation leads to the following,

$$\begin{aligned}&(\beta\nabla v, \nabla\mathcal{I}v)_K = (\beta\nabla v, \nabla v)_K + (\beta\nabla v, \nabla\mathcal{I}v - \nabla v)_K\\ =&0.5\beta^+ de\|\nabla v^+\|^2 + 0.5\beta^-(1 - de)\|\nabla v^-\|^2 + (\beta\nabla v, \nabla\mathcal{I}v - \nabla v)_K,\end{aligned} \tag{3.2}$$

where $\|\mathbf{v}\|$ without a subscript denotes the Euclidian 2-norm of the vector $\mathbf{v}$. It is also straightforward to obtain the following identities:

$$\partial_x\mathcal{I}v = d\partial_x v^+ + (1 - d)\partial_x v^- \quad \text{and} \quad \partial_y\mathcal{I}v = e\partial_y v^+ + (1 - e)\partial_y v^-. \tag{3.3}$$

Define $[\![\partial_x v]\!] = (\partial_x v^+ - \partial_x v^-)$ and $[\![\partial_y v]\!] = (\partial_y v^+ - \partial_y v^-)$. Then by direct calculations we have

$$\begin{aligned}&(\beta\nabla v, \nabla\mathcal{I}v - \nabla v)_K = (\beta^+\nabla v^+, (\nabla\mathcal{I}v - \nabla v^+))_{K^+} + (\beta^-\nabla v^-, (\nabla\mathcal{I}v - \nabla v^-))_{K^-}\\ =&-0.5de\beta^+\left((1 - d)\partial_x v^+[\![\partial_x v]\!] + (1 - e)\partial_y v^+[\![\partial_y v]\!]\right)\\ &+ 0.5(1 - de)\beta^-\left(d\partial_x v^-[\![\partial_x v]\!] + e\partial_y v^-[\![\partial_y v]\!]\right).\end{aligned} \tag{3.4}$$

Thanks to the fact $[\![\beta\nabla v \cdot \mathbf{n}_\Gamma]\!] = 0$ and $[\![v]\!]|_\Gamma = 0$ there also holds

$$\beta^+\nabla v^+ \cdot \mathbf{n}_\Gamma = \beta^-\nabla v^- \cdot \mathbf{n}_\Gamma, \quad \nabla v^+ \cdot \mathbf{n}_\Gamma^\perp = \nabla v^- \cdot \mathbf{n}_\Gamma^\perp, \tag{3.5}$$

where $\mathbf{n}_\Gamma^\perp$ is the clockwise perpendicular vector to $\mathbf{n}_\Gamma$. In this case, we have $\mathbf{n}_\Gamma^\perp = (d, -e)/\sqrt{d^2 + e^2}$. Recall that $r := \beta^+/\beta^-$, then

$$\begin{aligned}\nabla v^+ &= (\nabla v^+ \cdot \mathbf{n}_\Gamma)\mathbf{n}_\Gamma + (\nabla v^+ \cdot \mathbf{n}_\Gamma^\perp)\mathbf{n}_\Gamma^\perp = \frac{1}{r}(\nabla v^- \cdot \mathbf{n}_\Gamma)\mathbf{n}_\Gamma + (\nabla v^- \cdot \mathbf{n}_\Gamma^\perp)\mathbf{n}_\Gamma^\perp,\\ \nabla v^- &= (\nabla v^- \cdot \mathbf{n}_\Gamma)\mathbf{n}_\Gamma + (\nabla v^- \cdot \mathbf{n}_\Gamma^\perp)\mathbf{n}_\Gamma^\perp = r(\nabla v^+ \cdot \mathbf{n}_\Gamma)\mathbf{n}_\Gamma + (\nabla v^+ \cdot \mathbf{n}_\Gamma^\perp)\mathbf{n}_\Gamma^\perp,\end{aligned} \tag{3.6}$$

and hence,

$$\begin{aligned}\nabla v^+ - \nabla v^- &= (1 - r)\,(\nabla v^+ \cdot \mathbf{n}_\Gamma)\mathbf{n}_\Gamma,\\ [\![\partial_x v]\!] &= (1 - r)\,(\nabla v^+ \cdot \mathbf{n}_\Gamma)\sin(\theta),\\ [\![\partial_y v]\!] &= (1 - r)\,(\nabla v^+ \cdot \mathbf{n}_\Gamma)\cos(\theta),\end{aligned} \tag{3.7}$$

where $\theta$ is the angle formed by $A_1D$ and $DE$. Note that $\mathbf{n}_\Gamma = (\sin(\theta), \cos(\theta))$.

From (3.7), (3.4) can be rewritten as

$$\begin{aligned}(\beta\nabla v, \nabla\mathcal{I}v - \nabla v)_K =\ &0.5\beta^+ de(r - 1)(1 - d)(\nabla v^+ \cdot \mathbf{n}_\Gamma)\partial_x v^+\sin(\theta)\\ &+ 0.5\beta^+ de(r - 1)(1 - e)(\nabla v^+ \cdot \mathbf{n}_\Gamma)\partial_y v^+\cos(\theta)\\ &+ 0.5\beta^-(1 - de)d(1 - r)(\nabla v^+ \cdot \mathbf{n}_\Gamma)\partial_x v^-\sin(\theta)\\ &+ 0.5\beta^-(1 - de)e(1 - r)(\nabla v^+ \cdot \mathbf{n}_\Gamma)\partial_y v^-\cos(\theta).\end{aligned} \tag{3.8}$$

When $d = e$, i.e, $\Gamma_K = DE$ is parallel to the hypotenuse of the reference triangle, (3.8) can be simplified as follows:

$$\begin{aligned}&(\beta\nabla v, \nabla\mathcal{I}v - \nabla v)_K\\ =&0.5\beta^+ e^2(1 - e)(r - 1)(\nabla v^+ \cdot \mathbf{n}_\Gamma)^2 + 0.5\beta^-(1 - e^2)e\,(1 - r)\,(\nabla v^+ \cdot \mathbf{n}_\Gamma)(\nabla v^- \cdot \mathbf{n}_\Gamma)\\ =&0.5\beta^+ e^2(1 - e)(r - 1)(\nabla v^+ \cdot \mathbf{n}_\Gamma)^2 + 0.5\beta^+(1 - e^2)e\,(1 - r)\,(\nabla v^+ \cdot \mathbf{n}_\Gamma)^2\\ =&0.5\beta^+ e(1 - e)(1 - r)(\nabla v^+ \cdot \mathbf{n}_\Gamma)^2.\end{aligned} \tag{3.9}$$

If $r \leq 1$, we immediately have that $(\beta\nabla v, \nabla\mathcal{I}v - \nabla v)_K \geq 0$ and therefore,

$$(\beta\nabla v, \nabla\mathcal{I}v)_K \geq \|\sqrt{\beta}\nabla v\|_K^2.$$

If $r \geq 1$, we could instead prove by using an alternative format. From (3.5) and (3.9), we have

$$(\beta\nabla v, \nabla\mathcal{I}v - \nabla v)_K = 0.5\beta^- e(1-e)\left(\frac{1}{r} - 1\right)(\nabla v^- \cdot \mathbf{n}_\Gamma)^2, \tag{3.10}$$

which, combined with (3.2), gives

$$\begin{aligned}
(\beta\nabla v, \nabla\mathcal{I}v)_K =& 0.5\beta^+ e^2\|\nabla v^+\|^2 + 0.5\beta^-(1-e^2)\|\nabla v^-\|^2 \\
&+ 0.5\beta^-(e-e^2)\left(\frac{1}{r} - 1\right)(\nabla v^- \cdot \mathbf{n}_\Gamma)^2 \\
=& 0.5\beta^+ e^2\|\nabla v^+\|^2 + 0.5\beta^-\left((1-e) + \frac{1}{r}(e-e^2)\right)(\nabla v^- \cdot \mathbf{n}_\Gamma)^2 \\
&+ 0.5\beta^-(1-e^2)(\nabla v^- \cdot \mathbf{n}_\Gamma^\perp)^2 \\
\geq& \|\sqrt{\beta}\nabla v\|_{K^+}^2 + \min\left\{\frac{(1-e) + \frac{1}{r}(e-e^2)}{1-e^2}, 1\right\}\|\sqrt{\beta}\nabla v\|_{K^-}^2 \\
\geq& \frac{1}{1+e}\|\sqrt{\beta}\nabla v\|_K^2 \geq \frac{1}{2}\|\sqrt{\beta}\nabla v\|_K^2.
\end{aligned} \tag{3.11}$$

For the other cases when $\Gamma_K$ is parallel to the adjacent sides of the right angle, we can prove (3.1) in a similar way. This completes the proof of the lemma.

**Remark 3.1.** Since the eigenvalues of the local stiffness matrix continuously depend on the variables $d$ and $e$, the local matrix should remain to be positive for a range of $d$ and $e$ values. Therefore, the unconditional stability result will continue to hold for interfaces that are "almost" parallel to mesh edges.

We also note that even though currently we are not able to prove the unconditional stability (inf-sup condition), numerical results by us and other works, e.g., see [36–38,40,49], have shown that the classical PG-IFEM always yields the optimal convergence rates for various test problems.

### *3.2. The unconditional stable scheme with penalty*

We now prove the unconditional well-posedness, i.e., continuity and inf-sup condition, for the stabilized PG-IFEM. Define the following energy norms: for $w \in \mathcal{S}(\mathcal{T}_h)$,

$$|||w||| := \left(\|\sqrt{\beta}\nabla_h w\|_\Omega^2 + \sum_{F\in\mathcal{E}^i} h_F^{-1}\|\sqrt{\beta}[\![w]\!]\|_F^2\right)^{1/2};$$

for $w \in \mathcal{S}(\mathcal{T}_h) \oplus H^1(\Omega)$,

$$|||w|||_* := \left(\|\sqrt{\beta}\nabla_h w\|_\Omega^2 + \sum_{F\in\mathcal{E}^i} h_F^{-1}\|\sqrt{\beta}[\![w]\!]\|_F^2 + h_F\|\sqrt{\beta}[\![\nabla w \cdot \mathbf{n}_F]\!]\|_F^2\right)^{1/2},$$

and for $v \in \mathcal{V}(\mathcal{T}_h)$,

$$|||v||| := \left(\|\sqrt{\beta}\nabla_h v\|_\Omega^2 + \sum_{F\in\mathcal{E}^i} h_F^{-1}\|\sqrt{\beta}[\![\Pi v]\!]\|_F^2\right)^{1/2}.$$

Note that we used the same notation $|||\cdot|||$ for $w \in \mathcal{S}(\mathcal{T}_h)$ and $v \in \mathcal{V}(\mathcal{T}_h)$ because indeed they can be defined in the same format since $\Pi w = w$ for any $w \in \mathcal{S}(\mathcal{T}_h)$.

We will use the following stability results for the interpolation operator $\mathcal{I}$ (Lemma 4.6 in [50]):

$$c\|\nabla v\|_K \le \|\nabla \mathcal{I} v\|_K \le C\|\nabla v\|_K \quad \forall K \in \mathcal{T}_h^i,\ \forall v \in \mathcal{S}(\mathcal{T}_h), \tag{3.12}$$

where the constants $c, C$ are independent of the mesh-interface intersection.

For simplicity, from now on we will use $\lesssim$ to represent $\le C$ where the generic constant $C$ is independent of the mesh size, mesh-interface intersection, and coefficient contrast.

**Lemma 3.2** (*Continuity*)**.** For any $w \in H^1(\Omega) \oplus \mathcal{S}(\mathcal{T}_h)$ and $v \in \mathcal{V}(\mathcal{T}_h)$ we have the following continuity result,

$$\tilde{a}_h(w, v) \lesssim |\!|\!| w |\!|\!|_* |\!|\!| v |\!|\!|, \tag{3.13}$$

where $\tilde{a}_h(\cdot,\cdot)$ is defined in (2.6).

**Proof.** By definition, it is sufficient to prove that

$$\sum_{F\in\mathcal{E}^i} h_F \|\sqrt{\beta}[\![\nabla \Pi v \cdot \mathbf{n}_F]\!]\|_F^2 \lesssim |\!|\!| v |\!|\!|^2.$$

Applying the trace inequality and the stability result in (3.12) give

$$h_F^{1/2}\|\sqrt{\beta}[\![\nabla \Pi v \cdot \mathbf{n}_F]\!]\|_F \lesssim \|\sqrt{\beta}\nabla \Pi v\|_{K_F^1\cup K_F^2} \lesssim \|\sqrt{\beta}\nabla v\|_{K_F^1\cup K_F^2}, \tag{3.14}$$

where $K_F^1$ and $K_F^2$ are the two elements sharing the interior edge $F$. □

**Lemma 3.3.** There holds the following inf-sup type condition: given $\gamma_1$ and $\gamma_2$ large enough in (2.6), there exists a constant $C$ independent of mesh size such that

$$\tilde{a}(v, \mathcal{I} v) \ge C|\!|\!| v |\!|\!|^2, \quad |\!|\!| \mathcal{I} v |\!|\!| \lesssim |\!|\!| v |\!|\!|, \quad \forall v \in \mathcal{S}(\mathcal{T}_h). \tag{3.15}$$

**Proof.** Note that if $v \in \mathcal{S}(\mathcal{T}_h)$, then $\Pi \circ \mathcal{I}$ is the identity operator. The second part of (3.15) is then a direct consequence of (3.12).

By a direct calculation, we have

$$\tilde{a}(v, \mathcal{I} v) = \sum_{K\in\mathcal{T}_h} (\beta\nabla v, \nabla \mathcal{I} v)_K + \sum_{F\in\mathcal{E}^i} \gamma_1 h_F^{-1}\|\sqrt{\beta}[\![v]\!]\|_F^2 + \sum_{F\in\mathcal{E}^i} \gamma_2 h_F \|\sqrt{\beta}[\![\nabla v \cdot \mathbf{n}_F]\!]\|_F^2. \tag{3.16}$$

To estimate the first term in (3.16), we first apply the add-and-subtract technique:

$$(\beta\nabla v, \nabla \mathcal{I} v)_K = (\beta\nabla v, \nabla v)_K + (\beta\nabla v, \nabla \mathcal{I} v - \nabla v)_K. \tag{3.17}$$

Using integration by parts, Cauchy–Schwartz inequality, trace and inverse inequalities, we have

$$\begin{aligned}
&\sum_{K\in\mathcal{T}_h} (\beta\nabla v, \nabla \mathcal{I} v - \nabla v)_K = \sum_{F\in\mathcal{E}^i} \langle [\![\beta\nabla v \cdot \mathbf{n}_F]\!], \{\mathcal{I} v - v\}\rangle_F - \sum_{F\in\mathcal{E}^i} \langle \{\beta\nabla v \cdot \mathbf{n}_F\}, [\![v]\!]\rangle_F \\
&\le \sum_{F\in\mathcal{E}^i} \sum_{K\in\{K_F^1, K_F^2\}} \left( \|\beta^{1/2}[\![\nabla v \cdot \mathbf{n}_F]\!]\|_F \|\beta^{1/2}(\mathcal{I} v - v)\|_F + \|\beta\nabla v \cdot \mathbf{n}_F\|_{-1/2,F} \|[\![v]\!]\|_{1/2,F} \right) \\
&\lesssim \sum_{F\in\mathcal{E}^i} \sum_{K\in\{K_F^1, K_F^2\}} \left( \|\beta^{1/2}[\![\nabla v \cdot \mathbf{n}_F]\!]\|_F \, \|\beta^{1/2}(\mathcal{I} v - v)\|_F + \|\beta\nabla v\|_K \, h_F^{-1/2} \|[\![v]\!]\|_F \right).
\end{aligned} \tag{3.18}$$

In the last inequality, we used the trace inequality, and therefore the constant only depends on the shape regularity of the triangle.

We now proceed to bound the term $\|\beta^{1/2}(\mathcal{I}v - v)\|_F$. By direct computations and the Poincaré inequality, thanks to the fact that $\mathcal{I}v - v = 0$ at the endpoints of $F$, we have

$$\begin{aligned}
&\|\beta^{1/2}(\mathcal{I}v - v)_K\|_F^2 = \beta^+\|\mathcal{I}v - v\|_{F^+}^2 + \beta^-\|\mathcal{I}v - v\|_{F^-}^2\\
\leq& C_p\left(\beta^+|F^+|^2\|\nabla(\mathcal{I}v - v)\cdot\mathbf{t}_F\|_{F^+}^2 + \beta^-|F^-|^2\|\nabla(\mathcal{I}v - v)\cdot\mathbf{t}_F\|_{F^-}^2\right)\\
\leq& C_p\left(\beta^+|F^+|^2\|\nabla(\mathcal{I}v - v)\|_{F^+}^2 + \beta^-|F^-|^2\|\nabla(\mathcal{I}v - v)\|_{F^-}^2\right)\\
=& C_p\left(\beta^+|F^+|^2\frac{|F^+|}{|K^+|}\|\nabla(\mathcal{I}v - v)\|_{K^+}^2 + \beta^-|F^-|^2\frac{|F^-|}{|K^-|}\|\nabla(\mathcal{I}v - v)\|_{K^-}^2\right)\\
\leq& C_p h_F\left(\beta^+\frac{|F^+|^2}{|K^+|}\|\nabla(\mathcal{I}v - v)\|_{K^+}^2 + \beta^-\frac{|F^-|^2}{|K^-|}\|\nabla(\mathcal{I}v - v)\|_{K^-}^2\right)\\
\leq& C_p\max\left\{\frac{|F^+|^2}{|K^+|}, \frac{|F^-|^2}{|K^-|}\right\}h_F\|\sqrt{\beta}\nabla(\mathcal{I}v - v)\|_K^2\\
\leq& C_1 h_F\|\sqrt{\beta}\nabla v\|_K^2.
\end{aligned} \tag{3.19}$$

where $C_p$ is the Poincaré constant. Indeed, by a direct calculation, $C_p$ can be replaced by $1/3$. In the last inequality, we used (3.12). Unfortunately, the constant $\max\left\{\frac{|F^+|^2}{|K^+|}, \frac{|F^-|^2}{|K^-|}\right\}$, and hence, $C_1$, depends on the interface-mesh intersection.

Combining (3.18) and (3.19) yields

$$\begin{aligned}
&\sum_{K\in\mathcal{T}_h}(\beta\nabla v, \nabla\mathcal{I}v - \nabla v)_K\\
\lesssim& C_1\sum_{F\in\mathcal{E}^i}\sum_{K\in\{K_F^1, K_F^2\}}\|\sqrt{\beta}\nabla v\|_K h_F^{1/2}\|\sqrt{\beta}[\![\nabla v\cdot\mathbf{n}_F]\!]\|_F\\
&+\sqrt{\frac{\beta^+}{\beta^-}}\sum_{F\in\mathcal{E}^i}\sum_{K\in\{K_F^1, K_F^2\}}\|\sqrt{\beta}\nabla v\|_K\, h_F^{-1/2}\|\sqrt{\beta}[\![v]\!]\|_F\\
\lesssim& C_1\sum_{F\in\mathcal{E}^i}\frac{h_F}{2\epsilon_1}\|\sqrt{\beta}[\![\nabla v\cdot\mathbf{n}_F]\!]\|_F^2 + \sqrt{r}\frac{h_F^{-1}}{2\epsilon_1}\|\sqrt{\beta}[\![v]\!]\|_F^2\\
&+\frac{C_1\epsilon_1 + \sqrt{r}\epsilon_2}{2}\sum_{K\in\mathcal{T}_h^i}\|\sqrt{\beta}\nabla v\|_K^2.
\end{aligned} \tag{3.20}$$

Finally, choosing $\epsilon_i, i = 1, 2$ small enough and $\gamma_i, i = 1, 2$ large enough proves the first part of (3.15). This completes the proof of the lemma. □

## 4. *A priori* error estimation

**Lemma 4.1.** We have the following best approximation result:

$$|\!|\!|u - u_h|\!|\!|_* \leq C\min_{v\in\mathcal{S}_D(\mathcal{T}_h)}|\!|\!|u - v|\!|\!|_*, \tag{4.1}$$

where the constant $C$ depends on the constants in (3.12) and (3.15).

**Proof.** Thanks to (2.8), (3.13) and (3.15), we have for any $v\in\mathcal{S}_D(\mathcal{T}_h)$

$$\begin{aligned}
&|\!|\!|u_h - v|\!|\!|^2 \leq C\tilde{a}_h(u_h - v, \mathcal{I}(u_h - v)) = C\tilde{a}_h(u - v, \mathcal{I}(u_h - v))\\
\lesssim& C|\!|\!|u - v|\!|\!|_*|\!|\!|\mathcal{I}(u_h - v)|\!|\!| = C|\!|\!|u - v|\!|\!|_*|\!|\!|u_h - v|\!|\!|,
\end{aligned} \tag{4.2}$$

which, combining with the triangle inequality gives (4.1). □

**Lemma 4.2.** We have the following optimal convergence result for the energy norm:

$$\min_{v\in\mathcal{S}_{\mathcal{D}}(\mathcal{T}_h)} |||u-v|||_* \le Ch\|u\|_{PH^2(\Omega)}, \tag{4.3}$$

where $\|u\|^2_{PH^2(\Omega)} := \|u\|^2_{2,\Omega^+} + \|u\|^2_{2,\Omega^-}$.

**Proof.** Choose $v = \Pi u$. From Theorem 4.1 in [33], we have

$$\sum_{F\in\mathcal{E}^i} h_F\|\beta^{1/2}[\![\nabla(u-\Pi u)\cdot\mathbf{n}_F]\!]\|^2_F \le Ch^2\|u\|^2_{PH^2(\Omega)}, \tag{4.4}$$

where the constant $C$ is independent of the mesh-interface intersection.

By the triangle, Young's, and the trace inequalities, we have the following bounds,

$$\begin{aligned}\sum_{F\in\mathcal{E}^i} h_F^{-1}\|\sqrt{\beta}[\![u-v]\!]\|^2_F &\lesssim \sum_{F\in\mathcal{E}^i} h_F^{-1}\|\sqrt{\beta}(u-v)|_{K_F^1}\|^2_F + h_F^{-1}\|\sqrt{\beta}(u-v)|_{K_F^2}\|^2_F\\ &\le C\sum_{F\in\mathcal{E}^i}\left(h_F^{-2}\|u-v\|^2_{K_F^1\cup K_F^2} + \|\nabla(u-v)\|^2_{K_F^1\cup K_F^2}\right).\end{aligned} \tag{4.5}$$

Note that the constants in (4.4) and (4.5) depend on the coefficient $\beta$. Finally, combining (4.4) and (4.5) with the following optimal interpolation result in [17]

$$\sum_{K\in\mathcal{T}_h} h_K^{-2}\|u-\Pi u\|^2_K + \|\nabla(u-\Pi u)\|^2_K \lesssim h^2\|u\|^2_{PH^2(\Omega)},$$

yields (4.3). This completes the proof of the lemma. □

## 5. *A posteriori* error estimation

In this section, we derive the residual-based *a posteriori* error analysis based on the formulation (2.5). The PG-IFEM algorithm without the stabilization terms shares the same error estimator.

### 5.1. Residual-based a posteriori error estimator and indicator

For every $K\in\mathcal{T}_h$ we define the local error indicator $\eta_K$ by

$$\begin{aligned}\eta_K^2 = &\sum_{F\in\mathcal{E}_K\cap\mathcal{E}^i}\left(\frac{h_F}{2}\|\beta^{1/2}[\![\nabla u_h\cdot\mathbf{n}_F]\!]\|^2_F + \frac{h_F}{2}\|\beta^{1/2}[\![\nabla u_h\cdot\mathbf{t}_F]\!]\|^2_F\right)\\ &+\sum_{F\in\mathcal{E}_K\cap\mathcal{E}_I\setminus\mathcal{E}^i}\frac{h_F}{2}\|\beta_F^{-1/2}[\![\beta\nabla u_h\cdot\mathbf{n}_F]\!]\|^2_F + \sum_{F\in\mathcal{E}_K\cap\mathcal{E}_N} h_F\|\beta_F^{-1/2}[\![\beta\nabla u_h\cdot\mathbf{n}_F]\!]\|^2_F\end{aligned} \tag{5.1}$$

where $\beta_F = \max\left(\beta|_{K_F^1}, \beta|_{K_F^2}\right)$. The global error estimator $\eta$ is then defined by

$$\eta = \left(\sum_{K\in\mathcal{T}_h}\eta_K^2\right)^{1/2}. \tag{5.2}$$

Since the error estimator $\eta_K$ is the same as in the penalized immersed finite element method introduced in [51], we refer [51] for the efficiency proof of the local error indicator $\eta_K$. It remains to prove the global reliability.

### 5.2. Global reliability

In this subsection, we establish the reliability bound of the global estimator $\eta$ given in (5.2). The analysis here is similar to that as in [51]. In this paper, for simplicity, we assume that the computational interface is exact and therefore carries no geometric approximation on the interface. For the general geometry case, we refer to [51] in which the curved interface is considered. For the curved interface, an additional geometry approximation error

should also be added to the a posteriori error estimation. However, numerical experiments have shown reliability without adding such a term for interfaces of various regularity.

Define the following Sobolev space

$$H_N^1(\Omega) = \left\{ v \in H^1(\Omega) : \int_\Omega v\, dx = 0 \quad \text{and} \quad \frac{\partial v}{\partial t} = 0 \text{ on } \Gamma_N \right\}.$$

Here, $\frac{\partial v}{\partial t} = \nabla u \cdot \mathbf{t}$ denotes the tangential derivative of $u$. For $\phi \in H^1(\Omega)$, define the adjoint curl operator by $\nabla^\perp \phi = \left( -\frac{\partial \phi}{\partial y}, \frac{\partial \phi}{\partial x} \right)$.

**Lemma 5.1** (*Helmholtz Decomposition*)**.** Let $u$ and $u_h$ be the solutions of (2.3) and (2.5), respectively. Then there exist uniquely $\phi \in H_D^1(\Omega)$ and $\psi \in H_N^1(\Omega)$ such that

$$\beta \nabla u - \beta \nabla_h u_h = \beta \nabla \phi + \nabla^\perp \psi. \tag{5.3}$$

Moreover,

$$(\nabla \phi, \nabla^\perp \psi) = 0 \quad (\beta \nabla_h e, \nabla_h e) = (\beta \nabla \phi, \nabla \phi) + (\beta^{-1} \nabla^\perp \psi, \nabla^\perp \psi), \tag{5.4}$$

where $e = u - u_h$ and $\nabla^\perp = \left( \frac{\partial}{\partial y}, -\frac{\partial}{\partial x} \right)$.

A proof of the lemma can be found in Lemma 4.1 [51].

We now define a Clément-type interpolation operator $I_c : H_D^1(\Omega) \to \mathcal{V}(\mathcal{T}_h)$ by

$$I_c(v) = \sum_{z \in \mathcal{N}} (\pi_z v) \lambda_z(x) \tag{5.5}$$

where $\pi_z$ is defined by

$$\pi_z(v) = \begin{cases} \dfrac{\int_{\omega_z} \lambda_z v\, dx}{\int_{\omega_z} \lambda_z\, dx}, & \forall\, z \in \mathcal{N} \setminus \mathcal{N}_D, \\ 0, & \forall\, z \in \mathcal{N}_D, \end{cases} \tag{5.6}$$

where $\lambda_z$ is the classical barycentric hat function of $\mathcal{V}(\mathcal{T}_h)$ associated to $z \in \mathcal{N}$ and $\omega_z$ is the union of elements sharing $z$ as a common vertex. Note that

$$(v - \pi_z v, \lambda_z)_{\omega_z} = 0 \quad \forall\, z \in \mathcal{N} \setminus \mathcal{N}_D. \tag{5.7}$$

By Lemma 6.1 in [52] there holds for all $v \in H_D^1(\Omega)$

$$\|v - \pi_z v\|_{\omega_z} \le C \operatorname{diam}(\omega_z) \|\nabla v\|_{\omega_z}, \quad \forall\, z \in \mathcal{N}. \tag{5.8}$$

Note that $\Pi(I_c(v))$ is the corresponding modified Clément type interpolation of $v$ into the space of $\mathcal{S}(\mathcal{T}_h)$ defined in [51]. We recall the approximation and stability results for the above two types of Clément-type interpolation operators.

**Lemma 5.2** (*Clément-type Interpolation*)**.** Let $v \in H_D^1(\Omega)$, and $I_c v \in \mathcal{V}(\mathcal{T}_h)$ be the interpolation of $v$ defined in (5.5). Then there exists a constant $C > 0$ that is independent of the mesh size and the location of the interface such that

$$\begin{aligned} &\left\| v - I_c v \right\|_K + \left\| v - \Pi(I_c(v)) \right\|_K \le C h_K \left\| \nabla v \right\|_{\omega_K}, \forall\, K \in \mathcal{T}_h, \\ &\left\| \nabla(v - I_c v) \right\|_K + \left\| \nabla(v - \Pi(I_c(v))) \right\|_K \le C \left\| \nabla v \right\|_{\omega_K}, \ \forall\, K \in \mathcal{T}_h, \\ &\left\| (v - I_c v)|_K \right\|_F + \left\| (v - \Pi(I_c(v)))|_K \right\|_F \le C h_F^{1/2} \left\| \nabla v \right\|_{\omega_K}, \forall\, F \in \mathcal{E}_K, \end{aligned} \tag{5.9}$$

where $\omega_K$ is the union of all elements sharing at least one vertex with $K$.

**Lemma 5.3.** Let $\phi$ and $\psi$ be given in (5.3). Then we have the following error representations in the weighted semi-$H^1$ norm:

$$\begin{aligned}\|\sqrt{\beta}\nabla\phi\|_{\Omega}^2 &= (f,\phi-v) - \sum_{F\in\mathcal{E}_I\cup\mathcal{E}_N}\int_F [\![\beta\nabla u_h\cdot\mathbf{n}_F]\!](\phi-v)ds \\ &\quad + \sum_{F\in\mathcal{E}^i}\int_F \left(\gamma_1 h_F^{-1}\beta[\![u_h]\!][\![\Pi v]\!] + \gamma_2 h_F\beta[\![\nabla u_h\cdot\mathbf{n}_F]\!][\![\nabla\Pi v\cdot\mathbf{n}_F]\!]\right) ds.\end{aligned} \tag{5.10}$$

for any $v\in V_D(\mathcal{T}_h)$ and

$$\|\beta^{-1/2}\nabla^{\perp}\psi\|_{\Omega}^2 = -\sum_{F\in\mathcal{E}^i}\int_F [\![u_h]\!]\big(\nabla^{\perp}\psi\cdot\mathbf{n}_F\big)ds. \tag{5.11}$$

**Proof.** Let $v\in\mathcal{V}(\mathcal{T}_h)$ be arbitrary. Applying (5.4), (5.3), and integration by parts gives

$$\begin{aligned}&(\beta\nabla\phi,\nabla\phi) = (\beta\nabla_h e,\nabla\phi) = \Big(\beta\nabla_h e,\nabla(\phi-v)\Big) + \big(\beta\nabla_h e,\nabla v\big)\\ &= \sum_{K\in\mathcal{T}_h}\left(\int_K \big(f,\phi-v\big)\,dx + \int_{\partial K}\big(\beta\nabla e\cdot\mathbf{n}\big)\big(\phi-v\big)\,ds\right) + \big(\beta\nabla_h e,\nabla v\big)\\ &= (f,\phi-v) - \sum_{F\in\mathcal{E}_I\cap\mathcal{E}_N}\int_F [\![\beta\nabla u_h\cdot\mathbf{n}_F]\!](\phi-v)\,ds + \big(\beta\nabla_h e,\nabla v\big).\end{aligned} \tag{5.12}$$

Applying (2.8) and the facts that $[\![u]\!] = [\![\phi]\!] = [\![\beta\nabla u\cdot\mathbf{n}_F]\!] = 0$ for all $F\in\mathcal{E}_I$,

$$\begin{aligned}&\big(\beta\nabla_h e,\nabla v\big) = \big(\beta\nabla_h e,\nabla v\big) - \tilde{a}_h(e,v)\\ &= \sum_{F\in\mathcal{E}^i}\int_F \left(\gamma_1 h_F^{-1}\beta[\![u_h]\!][\![\Pi v]\!] + \gamma_2 h_F\beta[\![\nabla u_h\cdot\mathbf{n}_F]\!][\![\nabla\Pi v\cdot\mathbf{n}_F]\!]\right) ds.\end{aligned} \tag{5.13}$$

(5.10) is a direct consequence of (5.12) and (5.13).

To prove (5.11), by (5.4), (5.3), integration by parts, and the facts that

$$[\![e]\!]_F = -[\![u_h]\!]_F \quad\text{and}\quad [\![\nabla^{\perp}\psi\cdot\mathbf{n}_F]\!]_F = 0, \quad \forall\, F\in\mathcal{E}_I,$$

we have

$$\begin{aligned}&(\beta^{-1}\nabla^{\perp}\psi,\nabla^{\perp}\psi) = (\nabla e,\nabla^{\perp}\psi) = \sum_{K\in\mathcal{T}_h}\int_{\partial K} e(\nabla^{\perp}\psi\cdot\mathbf{n})\,ds\\ &= -\sum_{F\in\mathcal{E}^i}\int_F [\![u_h]\!]\big(\nabla^{\perp}\psi\cdot\mathbf{n}_F\big)\,ds.\end{aligned}$$

This completes the proof of the lemma. □

Define the data oscillation term,

$$H_f(\mathcal{T}_h) = \left(\sum_{z\in\mathcal{N}\setminus\mathcal{N}_D} h_z^2\|\beta^{-1/2}(f-\pi_z f)\|_{\omega_z}^2 + \sum_{z\in\mathcal{N}_D} h_z^2\|\beta^{-1/2}f\|_{\omega_z}^2\right)^{1/2},$$

where $\pi_z$ is defined in (5.6) and $h_z$ is the diameter of $\omega_z$.

**Theorem 5.4** (*Global Reliability*)**.** There exists a constant $C_r>0$ that is independent of the location of the interface and the mesh size, such that

$$\|\sqrt{\beta}(\nabla u-\nabla_h u_h)\|_{\Omega}\le C_r(\eta + H_f(\mathcal{T}_h)). \tag{5.14}$$

**Proof.** Recall from (5.10),

$$\|\sqrt{\beta}\nabla\phi\|_{\Omega}^2 = (f,\phi-v)_{\Omega} - \sum_{F\in\mathcal{E}_I\cup\mathcal{E}_N}\int_F [\![\beta\nabla u_h\cdot\mathbf{n}_F]\!](\phi-v)ds + \sum_{F\in\mathcal{E}^i}\int_F \left(\gamma_1 h_F^{-1}\beta[\![u_h]\!][\![\Pi v-\phi]\!] + \gamma_2 h_F\beta[\![\nabla u_h\cdot\mathbf{n}_F]\!][\![\nabla\Pi v\cdot\mathbf{n}_F]\!]\right) ds.$$

Let $v = I_c\phi$ in (5.10). Applying the definition of $I_c\phi$, partition of unity, (5.7), (5.9) and the Poincaré inequality yields

$$\begin{aligned}
&(f,\phi-v)_{\Omega} = \left(f,\left(\sum_{z\in\mathcal{N}}\lambda_z\right)\phi - \sum_{z\in\mathcal{N}}\pi_z\phi\lambda_z\right)\\
=&\sum_{z\in\mathcal{N}}(f,(\phi-\pi_z\phi)\lambda_z) = \sum_{z\in\mathcal{N}\setminus\mathcal{N}_D}(f-\pi_z f,(\phi-\pi_z\phi)\lambda_z) + \sum_{z\in\mathcal{N}_D}(f,\phi\lambda_z)\\
\le&\sum_{z\in\mathcal{N}\setminus\mathcal{N}_D}\|f-\pi_z f\|_{\omega_z}\|\phi-\pi_z\phi\|_{\omega_z} + \sum_{z\in\mathcal{N}_D}\|f\|_{\omega_z}\|\phi\|_{\omega_z}\\
\le&C\left(\sum_{z\in\mathcal{N}\setminus\mathcal{N}_D}h_z^2\|\beta^{-1/2}(f-\pi_z f)\|_{\omega_z}^2 + \sum_{z\in\mathcal{N}_D}h_z^2\|\beta^{-1/2}f\|_{\omega_z}^2\right)^{1/2}\|\sqrt{\beta}\nabla\phi\|_{\Omega}.
\end{aligned}\tag{5.15}$$

By the triangle and trace inequalities and (5.9), we have for all $F\in\mathcal{E}^i$

$$h_F^{-1/2}\|[\![\phi-\Pi(I_c\phi)]\!]\|_F \le \sum_{K\in\{K_F^1,K_F^2\}}h_F^{-1/2}\|(\phi-\Pi(I_c\phi))|_K\|_F \le C\sum_{K\in\{K_F^1,K_F^2\}}\|\nabla\phi\|_{\omega_K}.\tag{5.16}$$

Similarly, we have

$$h_F^{-1/2}\|(\phi-I_c\phi)|_K\|_F \le C\|\nabla\phi\|_{\omega_K},\quad \forall F\in\mathcal{E}_I.\tag{5.17}$$

Combining the triangle inequality and (5.9) also yields

$$\begin{aligned}
&h_F^{1/2}\|\sqrt{\beta}[\![\nabla\Pi(I_c\phi)\cdot\mathbf{n}_F]\!]\|_F \lesssim \sum_{K\in\{K_F^1,K_F^2\}}h_F^{1/2}\|\sqrt{\beta}\nabla(\Pi(I_c\phi))\cdot\mathbf{n}_F|_K\|_F\\
&\le C\sum_{K\in\{K_F^1,K_F^2\}}\|\sqrt{\beta}\nabla\phi\|_{\omega_K}.
\end{aligned}\tag{5.18}$$

From Lemma 4.4 in [51], we also have that

$$\|[\![u_h]\!]\|_F \le Ch_F\|[\![\nabla u_h\cdot\mathbf{t}_F]\!]\|_F,\quad \|[\![u_h]\!]\|_{1/2,F} \le h_F^{1/2}\|[\![\nabla u_h\cdot\mathbf{t}_F]\!]\|_F\quad \forall F\in\mathcal{E}^i.\tag{5.19}$$

Combining the Cauchy–Schwartz inequality and (5.15)–(5.19) gives that

$$\|\sqrt{\beta}\nabla\phi\|_{\Omega} \le C_r(\eta + H_f(\mathcal{T})).\tag{5.20}$$

Finally, by applying (5.11), the duality, (5.19), and trace inequality, we have

$$\begin{aligned}
&\|\beta^{-1/2}\nabla^{\perp}\psi\|_{\Omega}^2 = -\sum_{F\in\mathcal{E}^i}\int_F [\![u_h]\!](\nabla^{\perp}\psi\cdot\mathbf{n}_F)ds\\
\le&\sum_{F\in\mathcal{E}^i}\|[\![u_h]\!]\|_{1/2,F}\|\nabla^{\perp}\psi\cdot\mathbf{n}_F\|_{-1/2,F} \lesssim \sum_{F\in\mathcal{E}^i}h_F^{1/2}\|[\![\nabla u_h\cdot\mathbf{t}_F]\!]\|_F\|\nabla^{\perp}\psi\cdot\mathbf{n}_F\|_{-1/2,F}\\
\lesssim&\left(\sum_{F\in\mathcal{E}^i}h_F\|[\![\nabla u_h\cdot\mathbf{t}_F]\!]\|_F^2\right)\|\nabla^{\perp}\psi\|_{\Omega}.
\end{aligned}\tag{5.21}$$

Finally, (5.14) is a direct result of (5.4), (5.20) and (5.21). Note that here we use the fact that $[\![u_h]\!]|_F\in H_{00}^{1/2}(F)$ since $[\![u_h]\!]|_F$ takes zero value at the end points of $F$. This completes the proof of the theorem. □

## 6. Numerical results

In this section, we report some numerical results to demonstrate the performance of Petrov–Galerkin IFEM. In particular, we will compare three different PG-IFEMs as follows

$$a_h^{(i)}(u_h, v) = (f, v) - \langle g_N, v\rangle_{\Gamma_N} \tag{6.1}$$

where the bilinear forms take the form:

$$\text{Algorithm 1}: \ a_h^{(1)}(u, v) = \sum_{K\in\mathcal{T}_h} \int_K \beta \nabla u \cdot \nabla v \, dx. \tag{6.2}$$

$$\text{Algorithm 2}: \ a_h^{(2)}(u, v) = a_h^{(1)}(u, v) + \sum_{F\in\mathcal{E}^i} \gamma_1 h_F^{-1} \int_F \beta [\![u]\!][\![\Pi v]\!] \, ds. \tag{6.3}$$

$$\text{Algorithm 3}: \ a_h^{(3)}(u, v) = a_h^{(2)}(u, v) + \sum_{F\in\mathcal{E}^i} \gamma_2 h_F \int_F \beta [\![\nabla u \cdot \mathbf{n}_F]\!][\![\nabla \Pi v \cdot \mathbf{n}_F]\!] \, ds. \tag{6.4}$$

Here, $a^{(1)}(\cdot,\cdot)$ is the classical Petrov–Galerkin IFEM without any penalty, $a^{(2)}(\cdot,\cdot)$ takes into account the solution jump, and $a^{(3)}(\cdot,\cdot)$ includes additional normal flux jump.

We use unfitted Cartesian triangular meshes for all numerical experiments. The adaptive mesh refinement follows the standard procedure:

$$\text{Solve} \longrightarrow \text{Estimate} \longrightarrow \text{Mark} \longrightarrow \text{Refine}.$$

The residual-based error indicator $\eta_K$ on each element are computed in (5.1). We adopt the equilibration marking strategy, i.e., construct a minimal subset $\hat{\mathcal{T}}_h$ of $\mathcal{T}_h$ such that

$$\sum_{K\in\hat{\mathcal{T}}} \eta_K^2 \geq \theta^2 \eta^2,$$

where the threshold $\theta = 0.5$. Finally, we refine the marked triangles by the newest vertex bisection [53]. We use log–log plots for all error reporting figures. In our numerical experiments, we report errors of numerical solutions in the $L^2$ norm and semi-$H^1$ norm. We note that the $H^1$ semi-norm should have the similar convergence behavior to the weighted semi-$H^1$ norm, i.e., $\|\sqrt{\beta}\nabla(u-u_h)\|_\Omega$, for reasonable large jumps of $\beta$ used in our numerical tests. In addition, we report the error in $L^\infty$ norm, which measures the largest discrepancy among all the mesh points.

**Example 6.1** (*Smooth Interface with Moderate Jump*)**.** In this example, we consider a diffusion problem with a smooth circular interface which has been reported in [18]. Let $\Omega = (-1, 1)^2$, and the interface $\Gamma$ is a circle centered at $(x_0, y_0) = (0, 0)$ with radius $r_0 = \pi/6.28$. The interface separates $\Omega$ into two sub-domains, denoted by $\Omega^-$ and $\Omega^+$ such that

$$\Omega^- = \{(x, y) \in \Omega : r(x, y) < r_0\} \quad \text{and} \quad \Omega^+ = \{(x, y) \in \Omega : r(x, y) > r_0\},$$

where

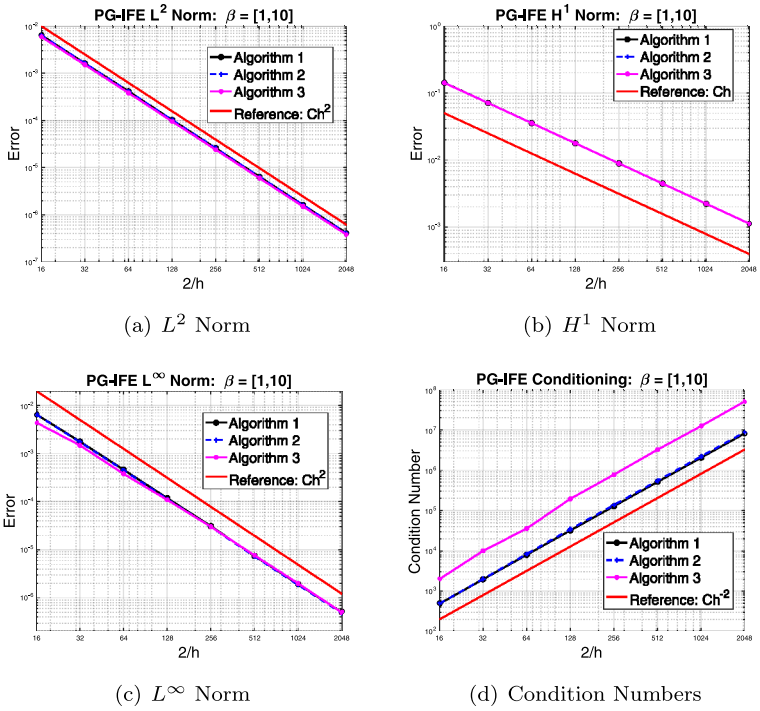
$$r(x, y) = (x - x_0)^2 + (y - y_0)^2.$$

The exact solution to this interface problem is

$$u(x, y) = \begin{cases} \frac{1}{\beta^-} r^p, & \text{if } (x, y) \in \Omega^-, \\ \frac{1}{\beta^+} r^p + \left(\frac{1}{\beta^-} - \frac{1}{\beta^+}\right) r_0^p, & \text{if } (x, y) \in \Omega^+. \end{cases} \tag{6.5}$$

Here $\beta^\pm > 0$ are the diffusion coefficients, and $p = 5$ is the regularity parameter. We let $(\beta^-, \beta^+) = (1, 10)$ which represents a moderate jump of the diffusion coefficients.

We use uniform meshes to test the convergence rates. We start from an $8 \times 8$ Cartesian triangular mesh and perform uniform mesh refinement until $1024 \times 1024$. All three algorithms (6.2)–(6.4) are examined. Errors in $L^2$ and $H^1$ norms are reported Fig. 6.1(a)–(b), respectively. It can be observed that errors in $L^2$, $H^1$ norms decay in optimal orders, which confirms our theoretical results (4.1). The error in the $L^\infty$ norm which measures the maximum error




(a) $L^2$ Norm (b) $H^1$ Norm

(c) $L^\infty$ Norm (d) Condition Numbers

**Fig. 6.1.** Uniform mesh convergence results of Example 6.1 with $(\beta^-, \beta^+) = (1, 10)$.

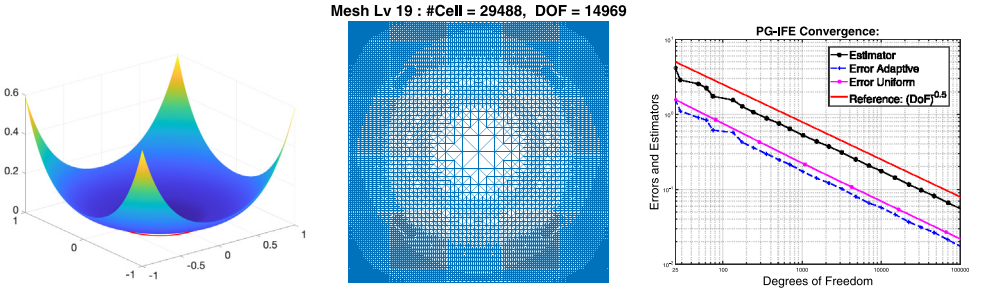
among mesh points is reported in Fig. 6.1(c). We also observe that the error in $L^\infty$ norm is second-order. Condition numbers of these algorithms are reported in Fig. 6.1(d). All three algorithms demonstrate a similar $\mathcal{O}(h^{-2})$ growth rate, while the Algorithm 3 (6.4) has a relatively larger condition number due to additional penalties.

For the adaptive mesh refinement test, we use the Algorithm 3 (6.4) to carry out all computation. However, we note that the classical PG-IFEM, i.e., Algorithm 1 (6.2), has similar numerical performance and is computationally much simpler than Algorithm 3 (6.4). The numerical solution is plotted in Fig. 6.2(a). The mesh depicted in Fig. 6.2(b) shows the adaptive refinement. We observe relatively dense mesh refinements around the four corners where the solution changes more dramatically. There is no extensive mesh refinement around the interface, because PG-IFEM itself can resolve the interface accurately for a moderate coefficient jump. This expected feature was also observed for partially penalized IFEM [51]. In Fig. 6.2(c), we report the convergence of PG-IFEM and the error estimators under adaptive mesh refinements. The decay rates are both close to $(\text{DoF})^{-1/2}$, which indicates the optimal order of errors with respect to the number of degrees of freedom (DoF). In comparison, errors of PG-IFEM under uniform refinement are very close to the errors of adaptive refinements.

**Example 6.2** (*Smooth Interface with Large Jump*)**.** In this example, we use the same problem in Example 6.1, but with a larger jump, i.e., $\beta^- = 1$, $\beta^+ = 1000$.

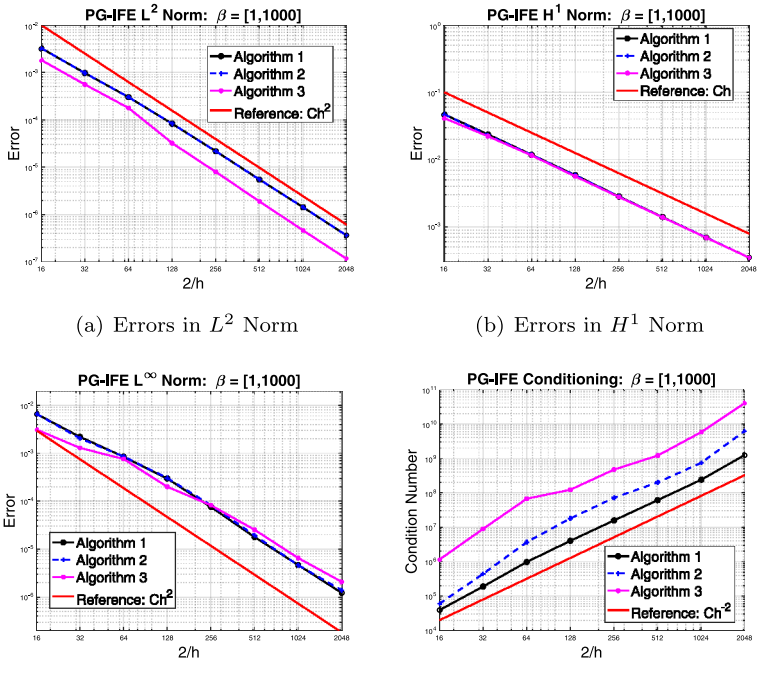
The errors in $L^2$, $H^1$, and $L^\infty$ norms decay in optimal orders, as reported in Fig. 6.3(a)–(c). In this example, we see a slightly better performance in Algorithm 3 (6.4) in the $L^2$ norm. On the other hand, (6.4) also gives the largest condition number among all three algorithms as shown in Fig. 6.3(d). The error surfaces on the $256 \times 256$ mesh are depicted in Fig. 6.4 for these algorithms. We can observe that similar magnitudes of errors for all three algorithms, while the behavior of the first two algorithms (6.2) and (6.3) are more similar.

A numerical solution is depicted in Fig. 6.5(a). For adaptive mesh refinement tests, we can see in Fig. 6.5(b) that the mesh is refined around the interface where the changes of solutions are more dramatic due to the large coefficient variation. In Fig. 6.5(c), the errors and the estimators decay optimally with respect to the degrees of freedom. Moreover, the errors are slightly smaller for adaptive mesh refinement than for uniform refinement, although the latter also decays at an optimal rate.




(a) PG-IFE solution (b) Adaptive Mesh (level 19) (c) Convergence Result

**Fig. 6.2.** Adaptive mesh results of Example 6.1 with $(\beta^-, \beta^+) = (1, 10)$.




(a) Errors in $L^2$ Norm (b) Errors in $H^1$ Norm

(c) Errors in $L^\infty$ Norm (d) Condition Numbers

**Fig. 6.3.** Uniform mesh convergence results of Example 6.2 with $(\beta^-, \beta^+) = (1, 1000)$.

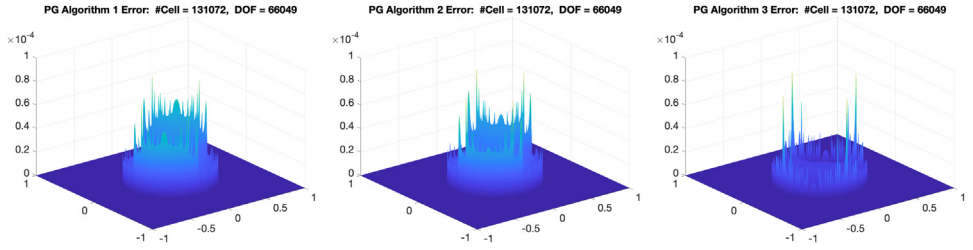
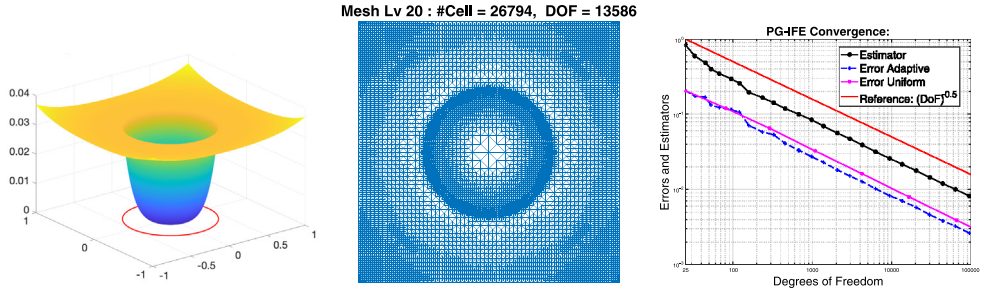



**Fig. 6.4.** From left: Error surfaces of Algorithms 1–3 on a 256 × 256 mesh of Example 6.2.

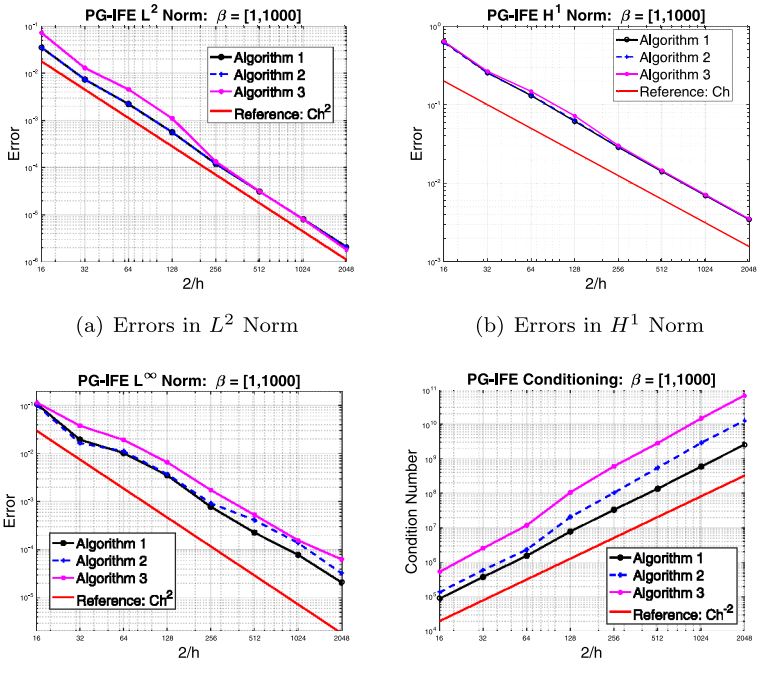
**Example 6.3** (*Complicated Interfacial Shape*)**.** In this example, we consider an interface problem with a more complicated interfacial shape. The exact solution has the following form

$$u(x, y) = \begin{cases} \frac{1}{\beta^-}\phi(x, y), & \text{if } \phi(x, y) < 0, \\ \frac{1}{\beta^+}\phi(x, y), & \text{if } \phi(x, y) \geq 0, \end{cases} \quad \text{in } \Omega = (-1, 1)^2 \tag{6.6}$$




(a) PG-IFE solution (b) Adaptive Mesh (level 19) (c) Convergence Result

**Fig. 6.5.** Adaptive mesh convergence results of Example 6.2 with $(\beta^-, \beta^+) = (1, 1000)$.




(a) Errors in $L^2$ Norm (b) Errors in $H^1$ Norm

(c) Errors in $L^\infty$ Norm (d) Condition Numbers

**Fig. 6.6.** Uniform mesh convergence results of Example 6.3 with $(\beta^-, \beta^+) = (1, 1000)$.

where $\phi$ is a level set function of the petal-shaped interface as follows

$$\phi(x, y) = (x^2 + y^2)^2 \left(1 + 0.5 \sin\left(12 \tan^{-1}\left(\frac{y}{x}\right)\right)\right) - 0.3.$$

This example was reported in [51]. We choose a large jump ratio $\beta^- = 1$ and $\beta^+ = 1000$.

The errors in $L^2$, $H^1$, and $L^\infty$ norms decay in optimal orders, as reported in Fig. 6.6(a)–(c). The performances are similar to all three algorithms. The algorithm 3 (6.4) gives the largest condition number among all three as shown in Fig. 6.6(d). The error surfaces are depicted in Fig. 6.7.

For adaptive mesh refinement tests, we start with a finer $32 \times 32$ mesh due to the complex shape of the interface. A numerical solution is plotted in Fig. 6.8(a). Fig. 6.8(b) demonstrates that the mesh is refined around the interface where the changes of solutions are more dramatic. In Fig. 6.8(c), the error and the estimators decay optimally with respect to the degrees of freedom. Compared to Example 6.2, the adaptive mesh refinement are more beneficial in this example due to the combined difficulty in high jump ratio and the complex interface shape. Errors in adaptive mesh refinement are smaller than those from the uniform refinement, although the latter is also an optimal rate.

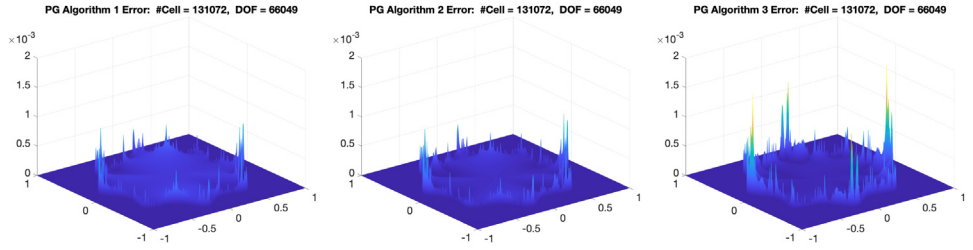
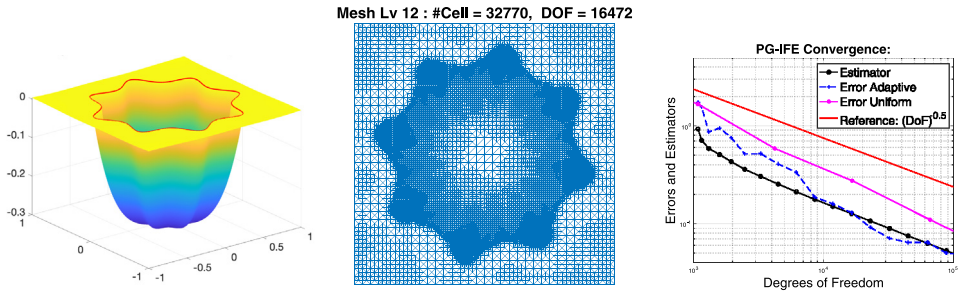



**Fig. 6.7.** From left: Error surfaces of Algorithms 1–3 on a 256 × 256 mesh of Example 6.3.




(a) PG-IFE solution (b) Adaptive Mesh (level 12) (c) Convergence Result

**Fig. 6.8.** Adaptive mesh convergence results of Example 6.3 with $(\beta^-, \beta^+) = (1, 1000)$.

**Example 6.4** (*Complicated Interface Shape with Flipped Coefficient*)**.** In this example, we consider the same problem as Example 6.3 with a flipped coefficient, i.e., $\beta^- = 1000$ and $\beta^+ = 1$.

The errors in $L^2$, $H^1$, and $L^\infty$ norms decay in optimal orders and behave similarly for all three algorithms, as shown in Fig. 6.9(a)–(c). The Algorithm 3 (6.4) gives the largest condition number among all three as shown in Fig. 6.9(d). The error surfaces of three algorithms are depicted in Fig. 6.10.

We start again from the 32 × 32 mesh for adaptive mesh refinement tests. A numerical solution has the shape Fig. 6.11(a). Fig. 6.11(b) demonstrate that mesh are refined around the interface where the changes of solutions are more dramatic. Not much refinement are performed inside the interface due to the solution is flat in this region. In Fig. 6.11(c), the error and the estimators decay optimally with respect to the degrees of freedom.

**Example 6.5** (*Sharp-Corner Interface*)**.** In this example, we consider the case when the interface has a sharp corner, as used in [27]. Let $\Omega = (-1, 1)^2$. The interface is defined by the level-set function:

$$\Gamma(x, y) = -y^2 + ((x - 1)\tan(\theta))^2 x. \tag{6.7}$$

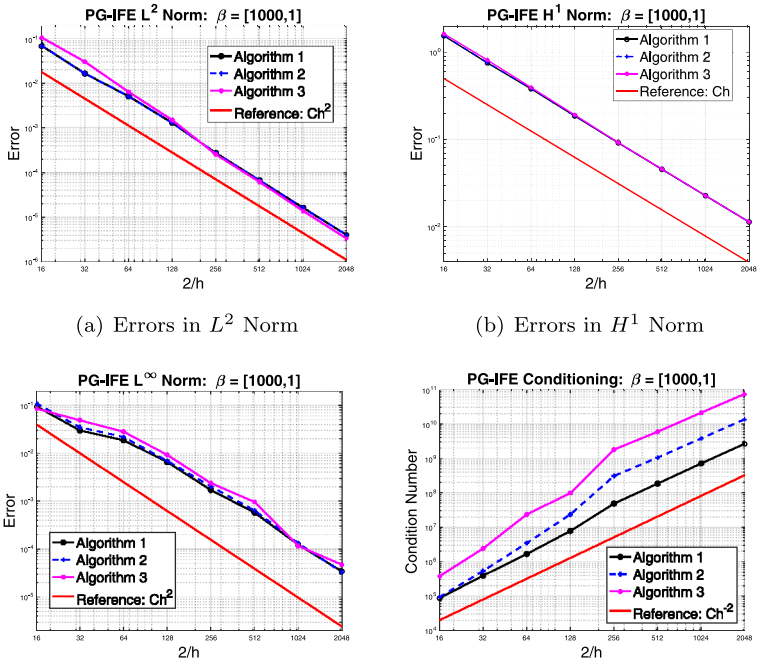
The subdomains are defined as $\Omega^+ = \{(x, y) \in \Omega : \Gamma(x, y) > 0\}$, and $\Omega^- = \{(x, y) \in \Omega : \Gamma(x, y) < 0\}$. The exact solution is chosen as:

$$u(x, y) = \begin{cases} \dfrac{1}{\beta^-}\Gamma(x, y), & (x, y) \in \Omega^-, \\ \dfrac{1}{\beta^+}\Gamma(x, y), & (x, y) \in \Omega^+. \end{cases} \tag{6.8}$$

We let $\beta^- = 1$ and $\beta^+ = 1000$.

The errors in $L^2$, $H^1$, and $L^\infty$ norms decay in optimal orders and behave similarly for all three algorithms, as shown in Fig. 6.12(a)–(c). The Algorithm 3 (6.4) gives the largest condition number among all three as shown in Fig. 6.12(d). The error surfaces are depicted in Fig. 6.13.

We start again from the 8 × 8 mesh for adaptive mesh refinement tests. A numerical solution has the shape plotted in Fig. 6.13(a). Fig. 6.14(b) demonstrates that mesh are refined around and inside the interface where the




(a) Errors in $L^2$ Norm (b) Errors in $H^1$ Norm

(c) Errors in $L^\infty$ Norm (d) Condition Numbers

**Fig. 6.9.** Uniform mesh convergence results of Example 6.4 with $(\beta^-, \beta^+) = (1000, 1)$.

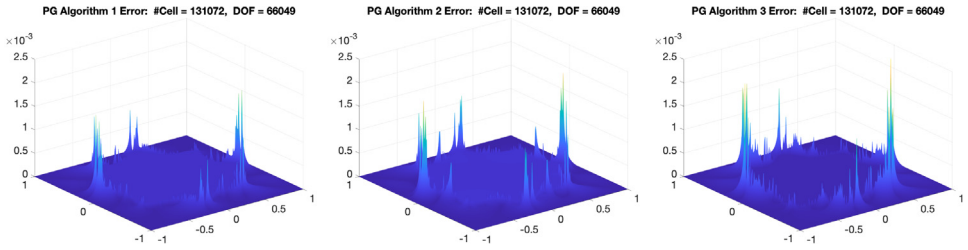
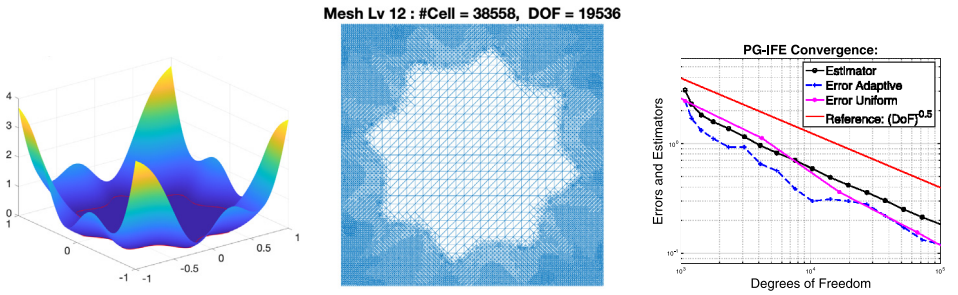
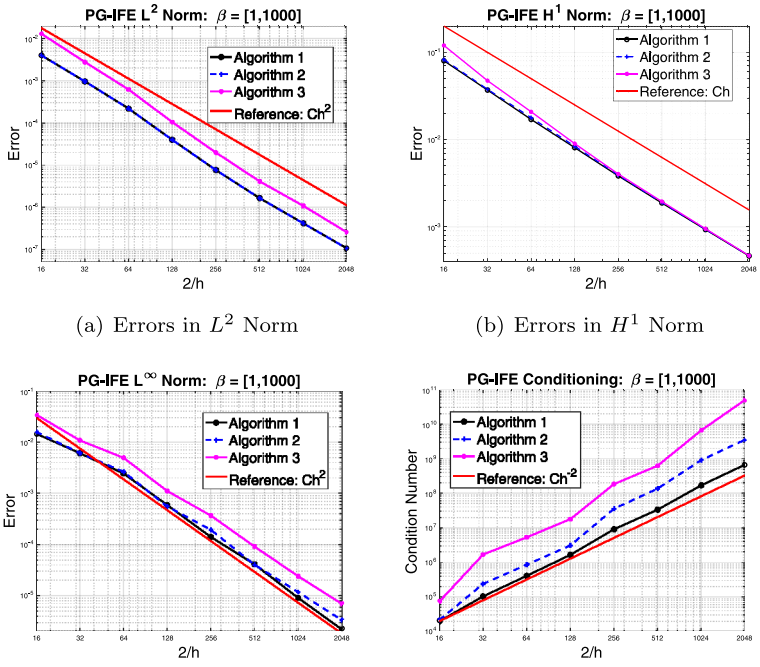



**Fig. 6.10.** From left: Error surfaces of Algorithms 1–3 on a 256 × 256 mesh of Example 6.4.




(a) PG-IFE solution (b) Adaptive Mesh (level 12) (c) Convergence Result

**Fig. 6.11.** Adaptive mesh convergence results of Example 6.4 with $(\beta^-, \beta^+) = (1000, 1)$.




(a) Errors in $L^2$ Norm (b) Errors in $H^1$ Norm

(c) Errors in $L^\infty$ Norm (d) Condition Numbers

**Fig. 6.12.** Uniform mesh convergence results of Example 6.5 with $(\beta^-, \beta^+) = (1, 1000)$.

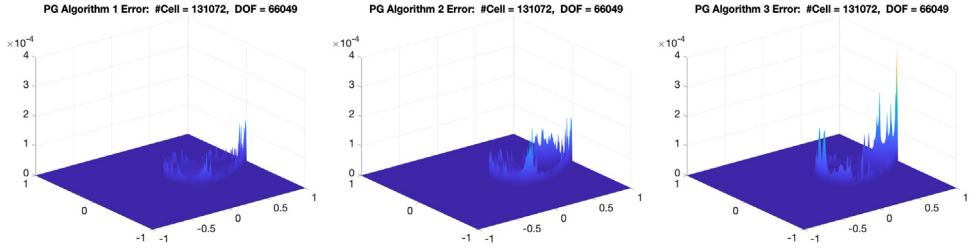



**Fig. 6.13.** From left: Error surfaces of Algorithms 1–3 on a 256 × 256 mesh of Example 6.5.

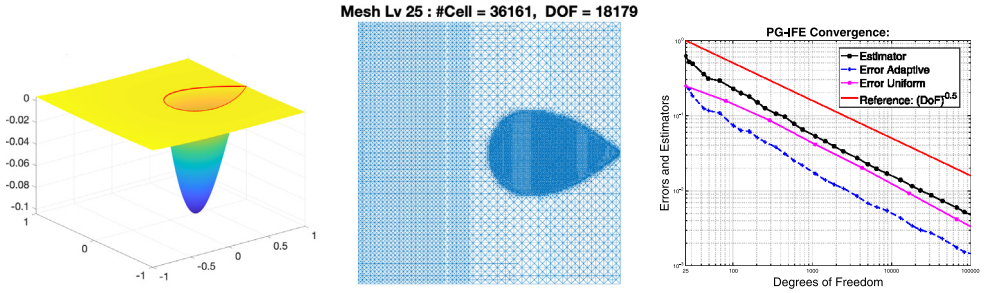
changes of solutions are more dramatic. In Fig. 6.14(c), the error and the estimators decay optimally with respect to the degrees of freedom. Again, we see some superiority in adaptive-mesh solutions over uniform-mesh solutions.

**Example 6.6** (*Multiple Interfaces*)**.** In this example, a multi-domain interface problem with multiple interfaces is considered. The interface consists of four circular interfaces defined as follows:

$$\begin{aligned}
\Gamma_1(x, y) &= (x + 0.5)^2 + (y + 0.5)^2 - (\pi/10)^2,\\
\Gamma_2(x, y) &= (x + 0.5)^2 + (y - 0.5)^2 - (\pi/9)^2,\\
\Gamma_3(x, y) &= (x - 0.5)^2 + (y + 0.5)^2 - (\pi/8)^2,\\
\Gamma_4(x, y) &= (x - 0.5)^2 + (y - 0.5)^2 - (\pi/7)^2.
\end{aligned}$$

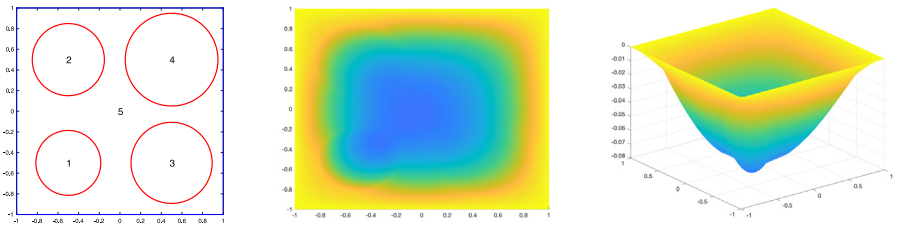
These interfaces separate the domain $\Omega = (-1, 1)^2$ into five subdomains $\Omega_i = \{(x, y) \in \Omega : \Gamma_i(x, y) < 0\}$ for $i = 1, 2, 3, 4$, and $\Omega_5 = \Omega/\overline{\Omega_1 \cup \Omega_2 \cup \Omega_3 \cup \Omega_4}$. Assume the boundary condition $u = 0$ on $\partial\Omega$, and the source function $f(x, y) = 1$. We consider three coefficient configurations as follows

- **Case 1**: $\beta_1 = 1$, $\beta_2 = 2$, $\beta_3 = 3$, $\beta_4 = 4$, and $\beta_5 = 5$.




(a) PG-IFE solution (b) Adaptive Mesh (level 25) (c) Convergence Result

**Fig. 6.14.** Adaptive mesh convergence results of Example 6.5 with $(\beta^-, \beta^+) = (1, 1000)$.




(a) coefficient (Case 1) (b) IFE solution (top view) (c) IFE solution (3D view)

**Fig. 6.15.** Coefficient configuration and numerical solutions of Case 1 of Example 6.6.

- **Case 2**: $\beta_1 = 1$, $\beta_2 = 2$, $\beta_3 = 3$, $\beta_4 = 4$, and $\beta_5 = 100$.
- **Case 3**: $\beta_1 = 1000$, $\beta_2 = 2$, $\beta_3 = 1$, $\beta_4 = 2000$, and $\beta_5 = 50$.

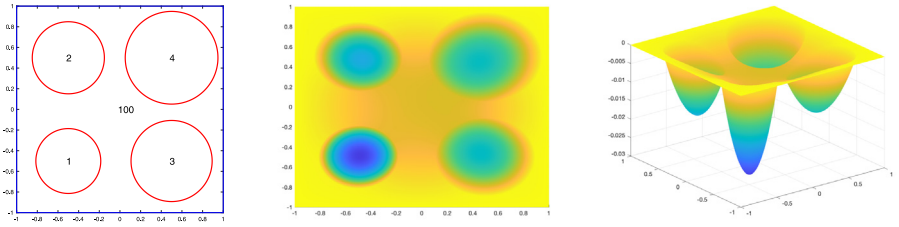
In all these cases, the analytical form of the exact solution is unknown.

We compute the numerical solutions using the PG-IFE Algorithm 1 defined in (6.2). The coefficient configuration, and numerical solution in 2D view and 3D view are depicted in Fig. 6.15, Fig. 6.16, and Fig. 6.17 for Case 1, Case 2, and Case 3, respectively. For adaptive mesh refinement tests, the adaptive meshes are plotted in Fig. 6.18 for three cases. In Case 1, all coefficients in five sub-domains have small jumps, thus the mesh refinement is quasi-uniform. Not much in any sub-domain nor the interface. In Case 2, due to the large jump in $\Omega_5$ over other subdomains, the mesh refinements concentrate around the interface. Also, note that mesh is slightly finer inside $\Omega_1$ due to the solutions being steeper inside this region among all subdomains. In Case 3, the mesh refinements are centered around interfaces and $\Omega_2$ and $\Omega_3$, but not so much in $\Omega_1$ and $\Omega_4$. This is due to a relatively large jump compared to Case 1, and the relative flat solution inside subregion $\Omega_1$ and $\Omega_4$, as depicted in Fig. 6.17. In Fig. 6.19, we report the convergence of the error estimators in all cases. We can see that the convergence is optimal with respect to the degrees of freedom. Since the exact solution is unknown, we cannot report the decay of true error. However, we believe the error will decay optimally based on the effectiveness of Example 6.1–6.5.

**Example 6.7** (*3D Sphere Interface Problem*)**.** In this example, we apply our method to a 3D interface problem. The construction of 3D IFE spaces are reported in [23–25]. Let $\Omega = (-1, 1)^3$ and let $\Gamma = \{(x, y, z) : \gamma(x, y, z) = 0\}$ be a spherical interface where $\gamma(x, y, z) = x^2 + y^2 + z^2 - r_0^2$. The exact solution is

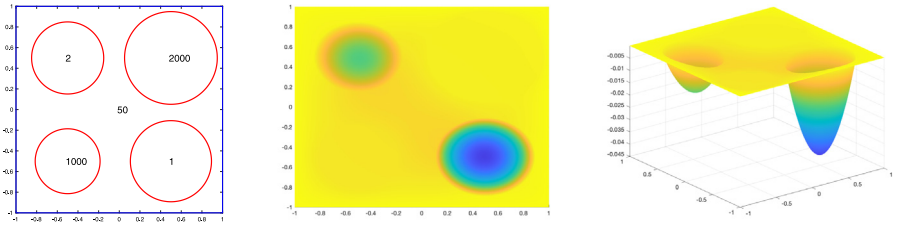
$$u(x, y, z) = \begin{cases} -\cos\left(\dfrac{\pi(x^2 + y^2 + z^2)}{2r_0^2}\right) & \text{in } \Omega^- := \{(x, y, z) \in \Omega : \gamma(x, y, z) < 0\}, \\ x^2 + y^2 + z^2 - r_0^2 & \text{in } \Omega^+ := \{(x, y, z) \in \Omega : \gamma(x, y, z) > 0\}. \end{cases} \tag{6.9}$$

The parameters are chosen to be $r_0 = \pi/4$ and $\beta^- = 1$, and $\beta^+ = \frac{\pi}{2r^2} \approx 2.5465$. This example has been used in [25].



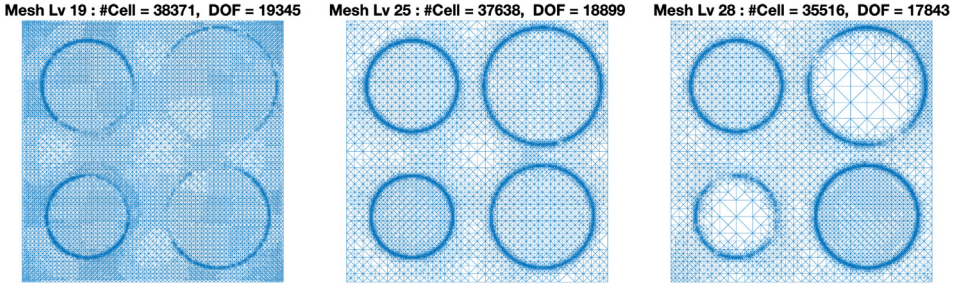

(a) coefficient (Case 2) (b) IFE solution (top view) (c) IFE solution (3D view)

**Fig. 6.16.** Coefficient configuration and numerical solutions of Case 2 of Example 6.6.




(a) coefficient (Case 3) (b) IFE solution (top view) (c) IFE solution (3D view)

**Fig. 6.17.** Coefficient configuration and numerical solutions of Case 3 of Example 6.6.




(a) Case 1 (b) Case 2 (c) Case 3

**Fig. 6.18.** Adaptive meshes of all three cases in Example 6.6.

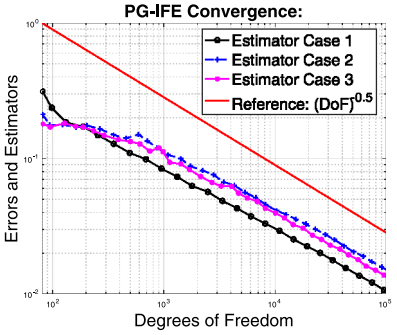



**Fig. 6.19.** Adaptive mesh convergence results of Example 6.6.

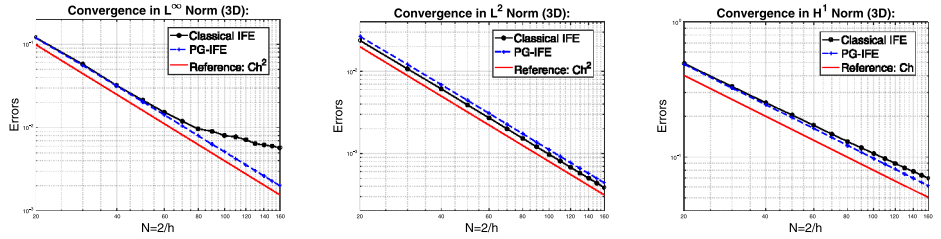
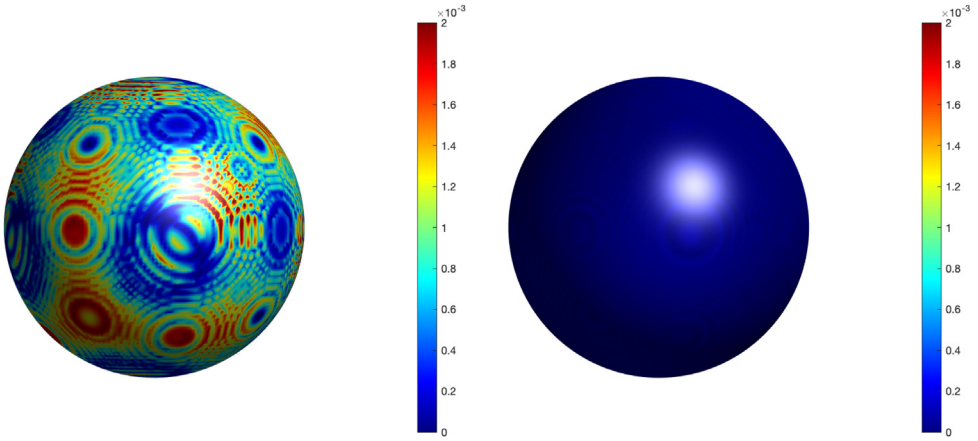



**Fig. 6.20.** Uniform mesh convergence results of Example 6.7.




(a) Errors Surface of IFE Solution (b) Errors Surface of PG-IFE Solution

**Fig. 6.21.** Error surface of interface on mesh $N = 100$ of Example 6.7.

Our computation is carried out on a family of uniform tetrahedral meshes obtained by first partitioning the domain into $N^3$ cuboids and then cutting each cuboid into six congruent tetrahedrons. We start from a coarse mesh with $N = 20$ and stretch to a very fine mesh with $N = 160$ by incrementing 10 more vertices in each direction for each finer mesh. In Fig. 6.20, we report errors and convergence rates in $L^\infty$, $L^2$, and $H^1$-norms for both PGIFE and classical IFE methods. In $L^\infty$ norm, the PG-IFE solution is superior to the classical IFE solution when the mesh size is small, and the PGIFE solution maintains the optimal second-order convergence. The accuracy of PG-IFE approximation is also illustrated in the comparison of errors on the interface surface in Fig. 6.21. The convergence rates in $L^2$, and $H^1$-norms are both optimal.

**Example 6.8** (*3D Orthocircle Interface Problem*)**.** In this example [25], we consider an interface problem with more complicated shape and a larger jump. Let $\Omega = (-1.2, 1.2)^3$, and let the interface be $\Gamma = \{(x, y, z) \in \Omega : \gamma(x, y, z) = 0\}$ where

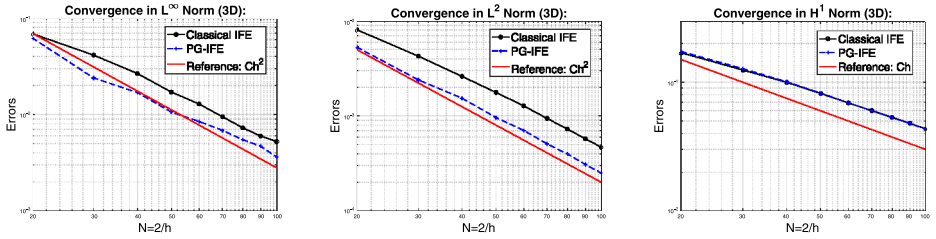
$$\gamma = \left[(x^2 + y^2 - 1)^2 + z^2\right]\left[(x^2 + z^2 - 1)^2 + y^2\right]\left[(y^2 + z^2 - 1)^2 + x^2\right] - 0.075^2\left[1 + 3(x^2 + y^2 + z^2)\right].$$

Let the exact solution be

$$u(x, y, z) = \begin{cases} \dfrac{1}{\beta^-}\gamma(x, y, z) & \text{in } \Omega^- := \{(x, y, z) \in \Omega : \gamma(x, y, z) < 0\}, \\ \dfrac{1}{\beta^+}\gamma(x, y, z) & \text{in } \Omega^+ := \{(x, y, z) \in \Omega : \gamma(x, y, z) > 0\}. \end{cases} \tag{6.10}$$

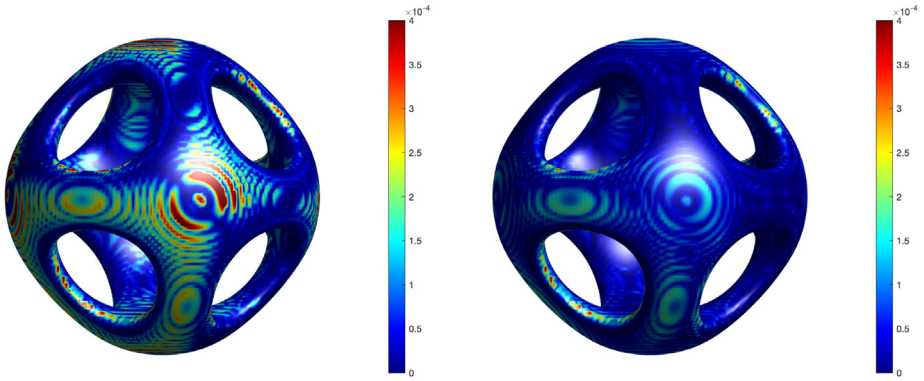
The coefficients are chosen to have a contrast as $\beta^- = 1$ and $\beta^+ = 100$.

In Fig. 6.22, we report errors and convergence rates in the $L^\infty$, $L^2$, and $H^1$-norms for both PGIFE and classical IFE methods. The convergence rates are optimal for this complex-shape interface problem in all three norms. In




(a) Errors in $L^\infty$ Norm (b) Errors in $L^2$ Norm (c) Errors in $H^1$ Norm

**Fig. 6.22.** Uniform mesh convergence results of Example 6.8.




(a) IFE Errors (b) PG-IFE Errors

**Fig. 6.23.** Error surface of interface on mesh $N = 100$ of Example 6.8.

$L^\infty$, $L^2$ norms, the PGIFE method outperforms the classical IFE method, which can be also seen from the error comparison on the interface surface in Fig. 6.23.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Data availability

No data was used for the research described in the article.

## References

[1] I. Babuška, J.E. Osborn, Generalized finite element methods: Their performance and their relation to mixed methods, SIAM J. Numer. Anal. 20 (3) (1983) 510–536.
[2] I. Babuška, U. Banerjee, Stable generalized finite element method (SGFEM), Comput. Methods Appl. Mech. Engrg. 201/204 (2012) 91–111.
[3] J. Dolbow, N. Moës, T. Belytschko, An extended finite element method for modeling crack growth with frictional contact, Comput. Methods Appl. Mech. Engrg. 190 (51–52) (2001) 6825–6846.
[4] B.L. Vaughan Jr., B.G. Smith, D.L. Chopp, A comparison of the extended finite element method with the immersed interface method for elliptic equations with discontinuous coefficients and singular sources, Commun. Appl. Math. Comput. Sci. 1 (2006) 207–228 (electronic).
[5] A. Hansbo, P. Hansbo, An unfitted finite element method, based on Nitsche's method, for elliptic interface problems, Comput. Methods Appl. Mech. Engrg. 191 (47–48) (2002) 5537–5552.
[6] E. Burman, S. Claus, P. Hansbo, M.G. Larson, A. Massing, CutFEM: Discretizing geometry and partial differential equations, Internat. J. Numer. Methods Engrg. 104 (7) (2015) 472–501.

[7] E. Burman, D. Elfverson, P. Hansbo, M.G. Larson, K. Larsson, Shape optimization using the cut finite element method, Comput. Methods Appl. Mech. Engrg. 328 (2018) 242–261.
[8] E. Burman, D. Elfverson, P. Hansbo, M.G. Larson, K. Larsson, A cut finite element method for the Bernoulli free boundary value problem, Comput. Methods Appl. Mech. Engrg. 317 (2017) 598–618.
[9] E. Burman, A. Ern, An unfitted hybrid high-order method for elliptic interface problems, SIAM J. Numer. Anal. 56 (3) (2018) 1525–1546.
[10] E. Burman, M. Cicuttin, G. Delay, A. Ern, An unfitted hybrid high-order method with cell agglomeration for elliptic interface problems, SIAM J. Sci. Comput. 43 (2) (2021) A859–A882.
[11] T.Y. Hou, X.-H. Wu, A multiscale finite element method for elliptic problems in composite materials and porous media, J. Comput. Phys. 134 (1) (1997) 169–189.
[12] T.Y. Hou, X.-H. Wu, Z. Cai, Convergence of a multiscale finite element method for elliptic problems with rapidly oscillating coefficients, Math. Comp. 68 (227) (1999) 913–943.
[13] Y. Efendiev, T.Y. Hou, Multiscale Finite Element Methods, in: Surveys and Tutorials in the Applied Mathematical Sciences, vol. 4, Springer, New York, 2009, p. xii+234, Theory and applications.
[14] C.-C. Chu, I.G. Graham, T.-Y. Hou, A new multiscale finite element method for high-contrast elliptic interface problems, Math. Comp. 79 (272) (2010) 1915–1955.
[15] Z. Li, The immersed interface method using a finite element formulation, Appl. Numer. Math. 27 (3) (1998) 253–267.
[16] Z. Li, T. Lin, X. Wu, New cartesian grid methods for interface problems using the finite element formulation, Numer. Math. 96 (1) (2003) 61–98.
[17] Z. Li, T. Lin, Y. Lin, R.C. Rogers, An immersed finite element space and its approximation capability, Numer. Methods Partial Differential Equations 20 (3) (2004) 338–367.
[18] T. Lin, Y. Lin, X. Zhang, Partially penalized immersed finite element methods for elliptic interface problems, SIAM J. Numer. Anal. 53 (2) (2015) 1121–1144.
[19] S. Adjerid, T. Lin, A $p$-th degree immersed finite element for boundary value problems with discontinuous coefficients, Appl. Numer. Math. 59 (6) (2009) 1303–1321.
[20] W. Cao, X. Zhang, Z. Zhang, Superconvergence of immersed finite element methods for interface problems, Adv. Comput. Math. 43 (4) (2017) 795–821.
[21] S.-H. Chou, D.Y. Kwak, K.T. Wee, Optimal convergence analysis of an immersed interface finite element method, Adv. Comput. Math. 33 (2) (2010) 149–168.
[22] J. Guzmán, M.A. Sánchez, M. Sarkis, On the accuracy of finite element approximations to a class of interface problems, Math. Comp. 85 (301) (2016) 2071–2098.
[23] R. Kafafy, T. Lin, Y. Lin, J. Wang, Three-dimensional immersed finite element methods for electric field simulation in composite materials, Internat. J. Numer. Methods Engrg. 64 (7) (2005) 940–972.
[24] R. Guo, T. Lin, An immersed finite element method for elliptic interface problems in three dimensions, J. Comput. Phys. 414 (1) (2020) 109478.
[25] R. Guo, X. Zhang, Solving three-dimensional interface problems with immersed finite elements: A-priori error analysis, J. Comput. Phys. 441 (2021) 110445.
[26] D.Y. Kwak, K.T. Wee, K.S. Chang, An analysis of a broken $P_1$-nonconforming finite element method for interface problems, SIAM J. Numer. Anal. 48 (6) (2010) 2117–2134.
[27] T. Lin, D. Sheen, X. Zhang, A nonconforming immersed finite element method for elliptic interface problems, J. Sci. Comput. 79 (1) (2019) 442–463.
[28] X. He, T. Lin, Y. Lin, Interior penalty bilinear IFE discontinuous Galerkin methods for elliptic equations with discontinuous coefficient, J. Syst. Sci. Complex. 23 (3) (2010) 467–483.
[29] T. Lin, Q. Yang, X. Zhang, A *priori* error estimates for some discontinuous Galerkin immersed finite element methods, J. Sci. Comput. 65 (3) (2015) 875–894.
[30] K. Liu, Q. Zou, Analysis of a special immersed finite volume method for elliptic interface problems, Int. J. Numer. Anal. Model. 16 (6) (2019) 964–984.
[31] W. Cao, X. Zhang, Z. Zhang, Q. Zou, Superconvergence of immersed fnite volume methods for one-dimensional interface problems, J. Sci. Comput. 73 (2–3) (2017) 543–565.
[32] L. Mu, X. Zhang, An immersed weak Galerkin method for elliptic interface problems, J. Comput. Appl. Math. 362 (2019) 471–483.
[33] R. Guo, T. Lin, Q. Zhuang, Improved error estimation for the partially penalized immersed finite element methods for elliptic interface problems, Int. J. Numer. Anal. Model. 16 (4) (2019) 575–589.
[34] T.Y. Hou, X.-H. Wu, Y. Zhang, Removing the cell resonance error in the multiscale finite element method via a Petrov–Galerkin formulation, Commun. Math. Sci. 2 (2) (2004) 185–205.
[35] J.S. Hesthaven, S. Zhang, X. Zhu, High-order multiscale finite element method for elliptic problems, Multiscale Model. Simul. 12 (2) (2014) 650–666.
[36] S. Hou, X.-D. Liu, A numerical method for solving variable coefficient elliptic equation with interfaces, J. Comput. Phys. 202 (2) (2005) 411–445.
[37] S. Hou, W. Wang, L. Wang, Numerical method for solving matrix coefficient elliptic equation with sharp-edged interfaces, J. Comput. Phys. 229 (19) (2010) 7162–7179.
[38] L. Wang, H. Zheng, X. Lu, L. Shi, A Petrov–Galerkin finite element interface method for interface problems with Bloch-periodic boundary conditions and its application in phononic crystals, J. Comput. Phys. 393 (2019) 117–138.

[39] L. Wang, S. Hou, L. Shi, P. Zhang, A bilinear Petrov–Galerkin finite element method for solving elliptic equation with discontinuous coefficients, Adv. Appl. Math. Mech. 11 (1) (2019) 216–240.
[40] S. Hou, P. Song, L. Wang, H. Zhao, A weak formulation for solving elliptic interface problems without body fitted grid, J. Comput. Phys. 249 (2013) 80–95.
[41] L. Wang, S. Hou, L. Shi, A numerical method for solving three-dimensional elliptic interface problems with triple junction points, Adv. Comput. Math. 44 (1) (2018) 175–193.
[42] Q. Zhuang, R. Guo, High degree discontinuous Petrov–Galerkin immersed finite element methods using fictitious elements for elliptic interface problems, J. Comput. Appl. Math. 362 (2019) 560–573.
[43] H. Ji, Q. Zhang, B. Zhang, Inf-sup stability of Petrov–Galerkin immersed finite element methods for one-dimensional elliptic interface problems, Numer. Methods Partial Differential Equations 34 (6) (2018) 1917–1932.
[44] D. Han, P. Wang, X. He, T. Lin, J. Wang, A 3D immersed finite element method with non-homogeneous interface flux jump for applications in particle-in-cell simulations of plasma-lunar surface interactions, J. Comput. Phys. 321 (2016) 965–980.
[45] D. Han, X. He, D. Lund, X. Zhang, PIFE-PIC: Parallel immersed finite element particle-in-cell for 3-D kinetic simulations of plasma-material interactions, SIAM J. Sci. Comput. 43 (3) (2021) C235–C257.
[46] S. Adjerid, N. Chaabane, T. Lin, An immersed discontinuous finite element method for Stokes interface problems, Comput. Methods Appl. Mech. Engrg. 293 (2015) 170–190.
[47] S. Adjerid, N. Chaabane, T. Lin, P. Yue, An immersed discontinuous finite element method for the Stokes problem with a moving interface, J. Comput. Appl. Math. 362 (2019) 540–559.
[48] D. Han, J. Wang, X. He, A nonhomogeneous immersed-finite-element particle-in-cell method for modeling dielectric surface charging in plasmas, IEEE Trans. Plasma Sci. 44 (8) (2016) 1326–1332.
[49] R. Guo, Y. Lin, J. Zou, Solving two dimensional H(curl)-elliptic interface systems with optimal convergence on unfitted meshes, 2020, arXiv preprint arXiv:2011.11905.
[50] R. Guo, Solving parabolic moving interface problems with dynamical immersed spaces on unfitted meshes: Fully discrete analysis, SIAM J. Numer. Anal. 59 (2) (2021) 797–828.
[51] C. He, X. Zhang, Residual-based a posteriori error estimation for immersed finite element methods, J. Sci. Comput. 81 (2019) 2051–2079.
[52] C. Carstensen, R. Verfürth, Edge residuals dominate a posteriori error estimates for low order finite element methods, SIAM J. Numer. Anal. 36 (5) (1999) 1571–1587 (electronic).
[53] J.M. Maubach, Local bisection refinement for $n$-simplicial grids generated by reflection, SIAM J. Sci. Comput. 16 (1) (1995) 210–227.