ANALYSIS OF NONCONFORMING IFE METHODS AND A NEW SCHEME FOR ELLIPTIC INTERFACE PROBLEMS

HAIFENG JI1,*, FENG WANG2, JINRU CHEN2,3 AND ZHILIN LI4

Abstract. In this paper, an important discovery has been found for nonconforming immersed finite element (IFE) methods using the integral values on edges as degrees of freedom for solving elliptic interface problems. We show that those IFE methods without penalties are not guaranteed to converge optimally if the tangential derivative of the exact solution and the jump of the coefficient are not zero on the interface. A nontrivial counter example is also provided to support our theoretical analysis. To recover the optimal convergence rates, we develop a new nonconforming IFE method with additional terms locally on interface edges. The new method is parameter-free which removes the limitation of the conventional partially penalized IFE method. We show the IFE basis functions are unisolvent on arbitrary triangles which is not considered in the literature. Furthermore, different from multipoint Taylor expansions, we derive the optimal approximation capabilities of both the Crouzeix–Raviart and the rotated-$Q_1$ IFE spaces via a unified approach which can handle the case of variable coefficients easily. Finally, optimal error estimates in both $H^1$- and $L^2$-norms are proved and confirmed with numerical experiments.

Mathematics Subject Classification. 65N15, 65N30, 35R05.

Received December 20, 2022. Accepted May 16, 2023.

1. Introduction

Let $\Omega \subset \mathbb{R}^2$ be a convex polygonal domain and $\Gamma$ be a $C^2$-smooth interface immersed in $\Omega$. Without loss of generality, we assume that $\Gamma$ divides $\Omega$ into two disjoint sub-domains $\Omega^+$ and $\Omega^-$ such that $\Gamma = \partial \Omega^-$, see Figure 1 for an illustration. We consider the following second-order elliptic interface problem

\begin{align*}
-\nabla \cdot (\beta(x)\nabla u(x)) &= f(x) & \text{in } \Omega \setminus \Gamma, \\
[u]_\Gamma(x) &= 0 & \text{on } \Gamma,
\end{align*}

(1.1) (1.2)

Keywords and phrases. Interface problem, Nonconforming, Immersed finite element, Unfitted mesh.

1 School of Science, Nanjing University of Posts and Telecommunications, Nanjing 210023, P.R. China.
2 Key Laboratory of NSLSCS, Ministry of Education, Jiangsu International Joint Laboratory of BDMCA, School of Mathematical Sciences, Nanjing Normal University, Nanjing 210023, P.R. China.
3 School of Mathematical Sciences, Jiangsu Second Normal University, Nanjing 211200, P.R. China.
4 Department of Mathematics, North Carolina State University, Raleigh, NC 27695, USA.
*Corresponding author: hfji@njupt.edu.cn; hfji1988@foxmail.com

© The authors. Published by EDP Sciences, SMAI 2023

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.
\[ [\beta \nabla u \cdot n]_\Gamma (x) = 0 \quad \text{on } \Gamma, \]  
\[ u(x) = 0 \quad \text{on } \partial \Omega, \]  
where \( n(x) \) is the unit normal vector of the interface \( \Gamma \) at point \( x \in \Gamma \) pointing toward \( \Omega^+ \), and the notation \([v]_\Gamma\) is defined as  
\[ [v]_\Gamma := v^+|_\Gamma - v^-|_\Gamma \quad \text{with } v^\pm = v|_{\Omega^\pm} \]  
for a piecewise smooth function \( v \). The coefficient \( \beta(x) \) can be discontinuous across the interface \( \Gamma \) and is assumed to be piecewise smooth  
\[ \beta(x) = \beta^+(x) \text{ if } x \in \Omega^+ \quad \text{and } \beta(x) = \beta^-(x) \text{ if } x \in \Omega^-, \]  
with \( \beta^s(x) \in C^1(\overline{\Omega^s}) \), \( s = +, - \). We also assume that there exist two positive constants \( \beta_{\min} \) and \( \beta_{\max} \) such that \( \beta_{\min} \leq \beta^s(x) \leq \beta_{\max} \) for all \( x \in \overline{\Omega^s} \), \( s = +, - \).

It is well-known that traditional isoparametric finite element methods using an interface-fitted mesh can solve the interface problem with optimal convergence rates, see for example [5, 8, 31, 34]. For complicated interfaces or moving interfaces, unfitted meshes, which are not necessarily aligned with interfaces, have some advantages over interface-fitted meshes. However, traditional finite element methods using unfitted meshes only achieve suboptimal convergence rates (\( O(h^{1/2}) \) in the \( H^1 \) norm and \( O(h) \) in the \( L^2 \) norm) no matter how high degree of the polynomial is used, see [1,12].

The design and analysis of finite element methods on unfitted meshes with optimal convergence rates was started in [2,3]. Since then, many unfitted mesh finite element methods have been developed (see [7,9,18,19,27,33] for a few examples). Among these methods, immersed finite element (IFE) methods [14,17,20,21,25–27,29,30] are designed to recover the optimal convergence rates of traditional finite element methods on unfitted meshes while keeping the degrees of freedom and the structure unchanged. The basic idea of IFEs is to modify traditional shape functions on interface elements to satisfy interface conditions approximately. However, these modifications are done on each interface element independently, which may cause discontinuities of IFE basis functions across interface edges. Even for the \( P_1 \) conforming IFE method, these discontinuities are not negligible [21,28] and the optimal convergence rates cannot be achieved if the discontinuities are not treated appropriately. To overcome the difficulty, Lin et al. [28] proposed a partially penalized IFE method where extra penalty terms at interface edges were added to penalize the discontinuity. For nonconforming IFE (i.e., a modification to the traditional Crouzeix–Raviart element [10] or the rotated-\( Q_1 \) element [32]), we can choose midpoint values or integral-values on edges as degrees of freedom. If we choose the midpoint values of edges as degrees of freedom, the
discontinuities of IFE basis functions are also not negligible and the optimal convergence rates can be obtained by adding penalties (see [35]). In contrast to the case of midpoint values as degrees of freedom, if the integral-values on edges are used as the degrees of freedom, then the IFE basis functions have less severe discontinuity across interface edges since the basis functions maintain the integral-value continuous [17,35]. It seems that this choice of degrees of freedom might overcome the difficulty caused by the discontinuities without using penalties. Extensive numerical examples in the literature [17,25,29,35] support this opinion. However, the rigorous proof is missing and the current research tends to improve the analysis of the related algorithms, as quoted in [35, pp. 96]: “How to theoretically prove that the Galerkin IFE scheme with nonconforming rotated $Q_1$ IFE functions using integral-value degrees of freedom does converge optimally is an interesting future research topic.”

In this paper, we show that those nonconforming IFE methods using the integral-value degrees of freedom are not guaranteed to converge optimally without penalties unless the tangential derivative of the true solution (i.e., $\nabla u \cdot t$) or the jump of the coefficient $\beta$ is zero on the interface (see Thm. 4.4 and Rmk. 4.3). Furthermore, to validate our theoretical analysis, a nontrivial counter example with $\nabla u \cdot t \neq 0$ (see Example 6.1) is constructed to show that nonconforming IFE methods using the integral-values degrees without adding penalties may only achieve suboptimal convergence rates (i.e., $O(h^{1/2})$ in the $H^1$ norm and $O(h)$ in the $L^2$ norm). Note that it is relatively easy to construct an exact solution to satisfy the homogeneous interface conditions (1.2) and (1.3) when the exact solution is a constant along the interface (i.e., $\nabla u \cdot t = 0$). To the best of our knowledge, almost all existing numerical examples in the literature satisfy the condition $\nabla u \cdot t = 0$ on the interface and thus the optimal convergence rates are observed (see [17,25,29,35] for example) which is in agreement with our theoretical analysis (see Thm. 4.4 in Sect. 4) in this paper.

To achieve the optimal convergence rates, the natural way is to add penalties on interface edges [35]. However, the symmetric partially penalized IFE methods proposed in [35] need a manually chosen parameter which is assumed to be large enough. In [16], Guo et al. analyzed a partially penalized IFE methods for elasticity interface problems and derived a lower bound for the penalty parameter. In this paper, we develop a parameter-free nonconforming IFE method by using a lifting operator defined locally on interface edges. We consider both the Crouzeix–Raviart and the rotated-$Q_1$ element for solving interface problems with variable coefficients. The method is symmetric and the coercivity is ensured without requiring a sufficiently large parameter. To avoid integrating on curved regions, we also approximate the interface by line segments connecting the intersection points of the mesh and the interface. The optimal error estimates are derived rigorously and are verified by numerical experiments.

There are other contributions of this paper. First, we prove that Crouzeix–Raviart IFE basis functions are unisolvent on arbitrary triangles if the integral-values on edges are used as degrees of freedom. Our recent study in [24] shows that, for the IFEs using nodal values as degrees of freedom, the maximum angle condition, $\alpha_{\text{max}} \leq \pi/2$ on interface triangles, is necessary to ensure the unisolvence of the basis functions. The unisolvence of basis functions on arbitrary triangles is a significant advantage of the nonconforming IFEs using integral-value degrees of freedom over the IFEs using nodal values as degrees of freedom. Another contribution is a unified proof of the optimal interpolation error estimates for both the Crouzeix–Raviart and the rotated-$Q_1$ nonconforming IFE spaces for interface problems with piecewise smooth coefficients. Different from multipoint Taylor expansions [17,35], our proof is based on auxiliary functions constructed on interface elements and some useful inequalities developed by Li et al. in [31] and by Bramble and King in [4] for estimating errors in the region near the interface. The other contribution is a new theoretical result that the interpolation polynomial on one side of the interface can approximate the extensions of the exact solution optimally on the whole element $T$ no matter how small $T \cap \Omega^+$ or $T \cap \Omega^-$ might be (see Thm. 3.5 in Sect. 3.2). The result is useful for proving the optimal interpolation error estimates on interface edges (see (5.20) in the proof of Lemma 5.6).
2. Preliminaries

Throughout the paper we adopt the standard notation $W^k_p(\Lambda)$ for Sobolev spaces on a domain $\Lambda$ with the norm $\|\cdot\|_{W^k_p(\Lambda)}$ and the seminorm $|\cdot|_{W^k_p(\Lambda)}$. Specially, $W^2_2(\Lambda)$ is denoted by $H^k(\Lambda)$ with the norm $\|\cdot\|_{H^k(\Lambda)}$ and the seminorm $|\cdot|_{H^k(\Lambda)}$. As usual $H^0_0(\Lambda) = \{v \in H^1(\Lambda) : v = 0 \text{ on } \partial\Lambda\}$. For a domain $\Lambda$, we define $\Lambda^s := \Lambda \cap \Omega^s$, $s = +, -$ and a space

$$\tilde{H}^2(\Lambda) := \{v \in L^2(\Lambda) : v|\Lambda^s \in H^2(\Lambda^s), s = +, -, [v]_{\Gamma \cap \Lambda} = 0, [\beta \nabla v \cdot n]_{\Gamma \cap \Lambda} = 0\}$$

(2.1)

when $\Lambda^s \neq \emptyset$, $s = +, -$. Obviously, $\tilde{H}^2(\Lambda) \subset H^1(\Lambda)$. The space $\tilde{H}^2(\Lambda)$ is equipped with the norm $\|\cdot\|_{H^2(\Lambda^+ \cup \Lambda^-)}$ and the semi-norm $|\cdot|_{H^2(\Lambda^+ \cup \Lambda^-)}$ satisfying

$$\|\cdot\|_{H^2(\Lambda^+ \cup \Lambda^-)} = \|\cdot\|_{H^2(\Lambda^+)} + \|\cdot\|_{H^2(\Lambda^-)}, \quad |\cdot|_{H^2(\Lambda^+ \cup \Lambda^-)} = |\cdot|_{H^2(\Lambda^+)} + |\cdot|_{H^2(\Lambda^-)}.$$

By integration by parts, we can immediately derive the following weak formulation of (1.1)–(1.4): find $u \in H^1_0(\Omega)$ such that

$$a(u,v) := \int_{\Omega} \beta(x) \nabla u \cdot \nabla v \, dx = \int_{\Omega} f v \, dx \quad \forall v \in H^1_0(\Omega).$$

(2.2)

We have the following regularity theorem for the weak solution (see [23] for the case of piecewise smooth coefficients and [9,22] for the case of piecewise constant coefficients).

**Theorem 2.1.** If $f \in L^2(\Omega)$, then (1.1)–(1.4) has a unique solution $u \in \tilde{H}^2(\Omega)$ satisfying the following a priori estimate

$$\|u\|_{H^2(\Omega^+ \cup \Omega^-)} \leq C\|f\|_{L^2(\Omega)},$$

(2.3)

where $C$ is a positive constant depending only on $\Omega$, $\Gamma$ and $\beta$.

Let $\{T_h\}_{h>0}$ be a family of triangular or rectangular subdivisions of $\Omega$ such that no vertex of any element lies in the interior of an edge of another element. The diameter of $T \in T_h$ is denoted by $h_T$. We define $h = \max_{T \in T_h} h_T$ and assume that $T_h$ is shape regular, i.e., for every $T$, there exists a positive constant $q$ such that $h_T \leq q r_T$ where $r_T$ is the diameter of the largest circle inscribed in $T$. Denote $\mathcal{E}_h$ as the set of edges of the subdivision, and let $\mathcal{E}_h^s$ and $\mathcal{E}_h^b$ be the sets of interior edges and boundary edges. We adopt the convention that elements $T \in T_h$ and edges $e \in \mathcal{E}_h$ are open sets. Then, the sets of interface elements and interface edges are defined as

$$\mathcal{T}_h^\Gamma := \{T \in T_h : T \cap \Gamma \neq \emptyset\} \quad \text{and} \quad \mathcal{E}_h^\Gamma := \{e \in \mathcal{E}_h : e \cap \Gamma \neq \emptyset\},$$

and the sets of non-interface elements and non-interface edges are $\mathcal{T}_h^{non} = T_h \setminus \mathcal{T}_h^\Gamma$ and $\mathcal{E}_h^{non} = \mathcal{E}_h \setminus \mathcal{E}_h^\Gamma$. We can always refine the mesh near the interface to satisfy the following assumption.

**Assumption 2.2.** The interface $\Gamma$ does not intersect the boundary of any interface element at more than two points. The interface $\Gamma$ does not intersect the closure $\overline{e}$ for any $e \in \mathcal{E}_h$ at more than one point.
The interface $\Gamma$ is approximated by $\Gamma_h$ that is composed of all the line segments connecting the intersection points of the boundaries of interface elements and the interface. In addition, we assume that $\Gamma_h$ divides $\Omega$ into two disjoint sub-domains $\Omega^+_h$ and $\Omega^-_h$ such that $\Gamma_h = \partial \Omega^-_h$.

Given an interface element $T \in \mathcal{T}_h$, we denote the intersection points of $\Gamma$ and $\partial T$ by $D$ and $E$. The straight line $DE$ divides $T$ into $T^+_h = T \cap \Omega^+_h$ and $T^-_h = T \cap \Omega^-_h$, see Figure 2 for an illustration.

Let $n_h(x)$ be the unit normal vector of $\Gamma_h$ pointing toward $\Omega^+_h$. The tangent vector of $\Gamma_h$ can be defined as $t_h(x) = R_{\pi/2} n_h(x)$, where $R_\alpha$ is a rotation matrix $R_\alpha = \begin{bmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{bmatrix}$.

Denote $\text{dist}(x, \Gamma)$ as the distance between a point $x$ and the interface $\Gamma$, and $U(\Gamma, \delta) = \{ x \in \mathbb{R}^2 : \text{dist}(x, \Gamma) < \delta \}$ as the neighborhood of $\Gamma$ of thickness $\delta$. Define the meshsize of $\mathcal{T}_h$ by $h_h := \max_{T \in \mathcal{T}_h} h_T$.

It is obvious that $h_h \leq h$ and $\bigcup_{T \in \mathcal{T}_h} T \subset U(\Gamma, h_h)$.

We also define a signed distance function $\rho(x)$ with $\rho(x)|_{\Omega^+} = \text{dist}(x, \Gamma)$ and $\rho(x)|_{\Omega^-} = -\text{dist}(x, \Gamma)$. There exists a constant $\delta_0 > 0$ such that $\rho(x)$ is well-defined in $U(\Gamma, \delta_0)$ and $\rho(x) \in C^2(U(\Gamma, \delta_0))$ (see [11]).

**Assumption 2.3.** We assume that $h_h < \delta_0$ so that $T \subset U(\Gamma, \delta_0)$ for all $T \in \mathcal{T}_h$.

We extend the coefficients $\beta^s(x)$, $s = +, -$ smoothly to slightly larger domains $\Omega^s_c := \Omega^s \cup U(\Gamma, \delta_0)$, $s = +, -$ such that $\beta^s(x) \in C^1(\Omega^s_c)$ and $\beta^s_{\text{min}} \leq \beta^s(x) \leq \beta^s_{\text{max}}$, $s = +, -$, where the constants $\beta^s_{\text{min}}$ and $\beta^s_{\text{max}}$ depend on $\Gamma$, $\beta^\pm$ and $\delta_0$. Thus, there exists a constant $C_\beta$ such that

$$||\nabla \beta^s||_{L^\infty(\Omega^s_c)} \leq C_\beta, \quad s = +, -.$$  \hspace{1cm} (2.6)

By using the signed distance function $\rho$, we can evaluate the unit normal and tangent vectors of the interface as

$$n(x) = \nabla \rho, \quad t(x) = \left( -\frac{\partial \rho}{\partial x_2}, \frac{\partial \rho}{\partial x_1} \right)^T,$$  \hspace{1cm} (2.7)

---

**Figure 2.** Interface elements.
which are well-defined in the region $U(\Gamma, \delta_0)$. We note that the functions $n_h(x)$ and $t_h(x)$ are also viewed as piecewise constant vectors defined on interface elements. On each interface element $T \in T_h^\Gamma$, since $\Gamma$ is in $C^2$, by Rolle’s Theorem, there exists at least one point $x^* \in \Gamma \cap T$, see Figure 2, such that

$$n(x^*) = n_h(x^*) \quad \text{and} \quad t(x^*) = t_h(x^*).$$

(2.8)

Since $\rho(x) \in C^2(U(\Gamma, \delta_0))$, we have

$$n(x) \in (C^1(T))^2 \quad \text{and} \quad t(x) \in (C^1(T))^2 \quad \forall \ T \in T_h^\Gamma.$$

(2.9)

Using Taylor’s expansion at $x^*$, we further have

$$\|n - n_h\|_{L^\infty(T)} \leq C h_T \quad \text{and} \quad \|t - t_h\|_{L^\infty(T)} \leq C h_T \quad \forall \ T \in T_h^\Gamma.$$

(2.10)

The following lemma presents a $\delta$-strip argument that will be used for the error estimate in the region near the interface (see Lemma 2.1 in [31]).

**Lemma 2.4.** Let $\delta$ be sufficiently small. Then it holds for any $v \in H^1(\Omega)$ that

$$\|v\|_{L^2(U(\Gamma, \delta))} \leq C \sqrt{\delta} \|v\|_{H^1(\Omega)}.$$  

Furthermore, if $v|_{\Gamma} = 0$, then there holds

$$\|v\|_{L^2(U(\Gamma, \delta))} \leq C \delta \|\nabla v\|_{L^2(U(\Gamma, \delta))}.$$  

Recalling $T^s = T \cap \Omega^s$, $T^s_h = T \cap \Omega^s_h$, $s = +, -$ for all $T \in T_h^\Gamma$, we define

$$T^\Delta := (T^- \cap T^+_h) \cup (T^+ \cap T^-_h).$$

(2.11)

Since $\Gamma$ is in $C^2$, we have $|T^\Delta| \leq C h_T^3$. We shall need the following estimate on the region $T^\Delta$ (see Lemma 2 in [4]).

**Lemma 2.5.** Assume that $v \in H^1(T)$ and $T \in T_h^\Gamma$. Then there is a constant $C$, independent of $h$ and the interface location relative to the mesh, such that

$$\|v\|^2_{L^2(T^\Delta)} \leq C \left(h_T^2 \|v\|^2_{L^2(\Gamma \cap T)} + h_T^4 \|\nabla v\|^2_{L^2(T^\Delta)}\right).$$

3. Nonconforming IFE spaces and their properties

In this section, we describe nonconforming IFE spaces based on the Crouzeix–Raviart element or the rotated-$Q_1$ element and present their properties. To begin with, we define IFE shape function spaces. On a non-interface element $T \in T_h^{\text{non}}$, we use the traditional shape function space

$$V_h(T) = \begin{cases} \text{Span}\{1, x_1, x_2\}, & \text{for the Crouzeix–Raviart element (}T\text{ is a triangle)}, \\ \text{Span}\{1, x_1, x_2, x_1^2 - (\kappa_T x_2)^2\}, & \text{for the rotated-$Q_1$ element (}T\text{ is a rectangle)}, \end{cases}$$

where $\kappa_T = |e_1|/|e_2|$, $e_1$ and $e_2$ are edges of the rectangle and parallel to the $x_1$-axis and the $x_2$-axis, respectively. On an interface element $T \in T_h^\Gamma$, the IFE shape function space $S_h(T)$ is defined as the set of the following functions

$$\phi(x) = \begin{cases} \phi^+(x) \in V_h(T) & \text{if } x = (x_1, x_2)^T \in T_h^+, \\ \phi^-(x) \in V_h(T) & \text{if } x = (x_1, x_2)^T \in T_h^- \text{,} \end{cases}$$

(3.1)
Actually, we can choose \( x \) satisfying
\[
\beta^+_c (\nabla \phi^+ \cdot \mathbf{n}_h)(x_p) - \beta^-_c (\nabla \phi^- \cdot \mathbf{n}_h)(x_p) = 0,
\]
where \( x_p \) is an arbitrary point on \( \Gamma_h \cap T \) and the constants \( \beta^+_c \) and \( \beta^-_c \) are chosen such that
\[
\| \beta^s(x) - \beta^c \|_{L^\infty(T)} \leq C h_T, \quad s = +, -.
\]
Actually, we can choose \( \beta^s := \beta^s(x^*_p) \) with arbitrary \( x^*_p \in T, \ s = +, - \), to satisfy the condition (3.4) since we know that \( \beta^s(x) \in C^1(T), \ s = +, - \) from (2.5).

**Remark 3.1.** For the Crouzeix–Raviart element, the condition (3.2) is equivalent to
\[
\phi^+(D) = \phi^-(D), \quad \phi^+(E) = \phi^-(E),
\]
since \( \phi^s(x), \ s = +, - \), are linear functions. For the rotated-\( Q_1 \) element, we can write \( \phi^s(x) \) as
\[
\phi^s(x) = a^s + b^s x_1 + c^s x_2 + d^s (x_1^2 - (\kappa T x_2)^2), \quad x = (x_1, x_2)^T, \quad s = +, -,
\]
where \( a^s, b^s, c^s, d^s, s = +, - \), are constants. If we define a functional \( d : V_h(T) \to \mathbb{R} \) as
\[
d(\phi^s) = \frac{1}{2} \frac{\partial^2 \phi^s}{\partial x_1^2} = \frac{1}{\sqrt{(4\kappa_T^2 + 4)|T|}} |\phi^s|_{H^2(T)},
\]
then \( d^s = d(\phi^s) \). Similar to Lemma 2.1 in [20], the condition (3.2) is equivalent to
\[
\phi^+(D) = \phi^-(D), \quad \phi^+(E) = \phi^-(E), \quad d(\phi^+) = d(\phi^-).
\]

**Remark 3.2.** For the Crouzeix–Raviart element, \( \beta^s \nabla \phi^s \cdot \mathbf{n}_h, \ s = +, - \), are constants on the interface element. Thus, the condition (3.3) is equivalent to
\[
\beta^+_c (\nabla \phi^+ \cdot \mathbf{n}_h)(x) = \beta^-_c (\nabla \phi^- \cdot \mathbf{n}_h)(x) \quad \forall x \in \Gamma_h \cap T.
\]
However, for the rotated-\( Q_1 \) element, the relation (3.6) is no longer valid. In [29], the authors weakly enforce the continuity by using the following condition
\[
\int_{\Gamma_h \cap T} \beta^+_c (\nabla \phi^+ \cdot \mathbf{n}_h) - \beta^-_c (\nabla \phi^- \cdot \mathbf{n}_h) \, ds = 0
\]
which is a particular case of (3.3) since there exists a point \( x_p \in \Gamma_h \cap T \) such that
\[
\int_{\Gamma_h \cap T} (\beta^+_c \nabla \phi^+ - \beta^-_c \nabla \phi^-) \cdot \mathbf{n}_h \, ds = |\Gamma_h \cap T| (\beta^+_c \nabla \phi^+ - \beta^-_c \nabla \phi^-)(x_p) \cdot \mathbf{n}_h.
\]
Let \( \mathcal{I} = \{1, 2, 3\} \) for the Crouzeix–Raviart element and \( \mathcal{I} = \{1, 2, 3, 4\} \) for the rotated-\( Q_1 \) element. The degrees of freedom are defined as the mean values over edges
\[
N_i(\phi) := \frac{1}{|e_i|} \int_{e_i} \phi \, ds, \quad i \in \mathcal{I},
\]
where \( e_i, \ i \in \mathcal{I} \) are edges of the element \( T \), and \( |e_i| \) denotes the length of the edge \( e_i \). On an interface element \( T \in \mathcal{T}_h^T \), the immersed finite element is defined as \((T, S_h(T), \Sigma_T)\) with \( \Sigma_T = \{N_i, i \in \mathcal{I}\} \).
The nonconforming IFE space $V_h^{IFE}$ is defined as the set of all functions satisfying
\[
\left\{ \begin{array}{ll}
\phi|_T \in S_h(T) & \forall T \in T_h^T, \\
\phi|_T \in V_h(T) & \forall T \in T_h^{non}, \\
\int_{e} [\phi]_e \, ds = 0 & \forall e \in \mathcal{E}_h^e.
\end{array} \right.
\]

We also need a space for homogeneous boundary conditions
\[
V_{h,0}^{IFE} := \left\{ v \in V_h^{IFE} : \int_{e} v \, ds = 0 \quad \forall e \in \mathcal{E}_h^b \right\}.
\]

### 3.1. The unisolvence of IFE basis functions

It was proved in [17] that the function $\phi \in S_h(T)$ is uniquely determined by $N_i(\phi)$, $i = 1, 2, 3, 4$ for the rotated-$Q_1$ element, and $i = 1, 2, 3$ for the Crouzeix–Raviart element when the interface element is an isosceles right triangle. Now we prove that the result is also valid for arbitrary triangles in the following lemma. Note that for the IFEs using nodal values as degrees of freedom, the maximum angle condition, $\alpha_{\text{max}} \leq \pi/2$ on interface triangles, is necessary to ensure the unisolvence (see [24]). This property of the unisolvence of basis functions is one of advantages of nonconforming IFEs compared with the IFEs using nodal values as degrees of freedom.

**Lemma 3.3.** Let $T$ be an arbitrary interface triangle. For the Crouzeix–Raviart element, the function $\phi \in S_h(T)$ is uniquely determined by $N_i(\phi)$, $i = 1, 2, 3$.

**Proof.** We follow the argument proposed in [15, 17]. Consider a triangle $\triangle A_1A_2A_3$ with edges $e_1 = \overline{A_2A_3}$, $e_2 = \overline{A_1A_3}$ and $e_3 = \overline{A_1A_2}$. The interface $\Gamma$ cuts $e_1$ and $e_2$ at points $D$ and $E$, see Figure 2 for an illustration. Without loss of generality, we assume $T_h^T = \triangle EDA_3$ since the case $T_h^T = \triangle EDA_3$ can be treated by reversing $\beta^+_c$ and $\beta^-_c$. Let $\lambda_i(x)$, $i = 1, 2, 3$, be basis functions in $V_h(T)$ such that
\[
\frac{1}{|e_j|} \int_{e_j} \lambda_i(x) \, dx = \delta_{ij} \quad \forall j \in \{1, 2, 3\},
\]
where $\delta_{ij}$ is the Kronecker function. Using (3.2) and $|e_3|^{-1} \int_{e_3} \phi \, ds = N_3(\phi)$, we can write the IFE shape function $\phi(x)$ in (3.1) as
\[
\phi(x) = \begin{cases} 
\phi^+(x) = c_1 \lambda_1(x) + c_2 \lambda_2(x) + N_3(\phi) \lambda_3(x) & \text{if } x = (x_1, x_2)^T \in T_h^+, \\
\phi^-(x) = \phi^+(x) + c_0 n_h \cdot D \vec{x} & \text{if } x = (x_1, x_2)^T \in T_h^-,
\end{cases}
\]
(3.7)

where $c_0, c_1, c_2$ are unknowns. Applying the condition (3.3), the unknown $c_0$ can be expressed as
\[
c_0 = (\beta^+_c/\beta^-_c - 1) \nabla \phi^+ \cdot n_h = (\beta^+_c/\beta^-_c - 1) (c_1 \nabla \lambda_1 + c_2 \nabla \lambda_2 + N_3(\phi) \nabla \lambda_3) \cdot n_h.
\]
(3.8)

Substituting (3.8) into (3.7) and using $N_i(\phi) = |e_i|^{-1} \int_{e_i} \phi \, ds$, $i = 1, 2$, we obtain the following linear system of equations for other coefficients (see [15, 17] for details),
\[
(I + (\beta^+_c/\beta^-_c - 1) \delta \gamma^T) c = b,
\]
(3.9)

where
\[
\delta = \left( |e_1|^{-1} \int_{A_3D} L(x) \, ds, |e_2|^{-1} \int_{A_3E} L(x) \, ds \right)^T, \quad L(x) = n_h \cdot D \vec{x},
\]
\[
c = (c_1, c_2)^T, \quad \gamma = (\nabla \lambda_1 \cdot n_h, \nabla \lambda_2 \cdot n_h)^T,
\]
\[
b = \left( N_1(\phi) - \frac{(\beta^+_c/\beta^-_c - 1) N_3(\phi) \nabla \lambda_3 \cdot n_h}{|e_1|} \int_{A_3D} L(x) \, ds, 
\right.
\]
\[
N_2(\phi) - \frac{(\beta^+_c/\beta^-_c - 1) N_3(\phi) \nabla \lambda_3 \cdot n_h}{|e_2|} \int_{A_3E} L(x) \, ds \left)^T. \right.
\]
(3.10)
Set $k_1 = |A_3D|^{-1}$ and $k_2 = |A_3E|^{-1}$. Let $M_i$ be the midpoint of the edge $e_i$, $i = 1, 2, 3$, and $Q$ be the orthogonal projection of $M_2$ onto the line $A_2A_3$. We can find out $\gamma(1)$ and $\delta(1)$ as below

$$\gamma(1) = \nabla \lambda_1 \cdot n_h = |M_2Q|^{-1} \overrightarrow{M_2Q} |M_2Q|^{-1} \cdot t_h = |M_2Q|^{-1} R_{\pi/2} \left( \overrightarrow{M_2Q} |M_2Q|^{-1} \right) \cdot R_{\pi/2} (n_h)$$

$$= |M_2Q|^{-1} A_2A_3 A_2A_3^{-1} \cdot t_h = \left( \frac{1}{2} |e_2| \sin \angle A_3 \right)^{-1} |e_1|^{-1} A_2A_3 \cdot t_h,$$

and

$$\delta(1) = |e_1|^{-1} \int_{A_3D} n_h \cdot D \overrightarrow{x} \, ds = |e_1|^{-1} |A_3D| \cdot n_h \cdot D \overrightarrow{H} = -\frac{1}{2} k_1 |A_3A_{3,\perp}|,$$

where $\angle A_3 \in (0, \pi)$, $H$ is the midpoint of the line segment $A_3D$, and $A_{3,\perp}$ is the orthogonal projection of $A_3$ onto the line $DE$. Thus,

$$\gamma(1) \delta(1) = \nabla \lambda_1 \cdot n_h |e_1|^{-1} \int_{A_3D} D \overrightarrow{x} \, ds = -k_1 A_2A_3 \cdot t_h |A_3A_{3,\perp}| \angle A_3)^{-1}$$

$$= -\overrightarrow{DA}_3 \cdot t_h |A_3A_{3,\perp}| \angle A_3)^{-1}.$$

Analogously, we have

$$\gamma(2) \delta(2) = \nabla \lambda_2 \cdot n_h |e_2|^{-1} \int_{A_3E} D \overrightarrow{x} \, ds = -A_3E \cdot t_h |A_3A_{3,\perp}| \angle A_3)^{-1}.$$

Therefore,

$$\gamma^T \delta = \overrightarrow{ED} \cdot t_h |A_3A_{3,\perp}| \angle A_3)^{-1} = |DE| |A_3A_{3,\perp}| \angle A_3)^{-1}.$$

As long as $\angle A_3ED \in (0, \pi)$, it is true that $|A_3A_{3,\perp}| = k_2 |e_2| \sin \angle A_3ED$, which together with the relations $|DE| \sin \angle A_3 = k_1 |e_1| \sin \angle A_3ED$ yields

$$\gamma^T \delta = k_1 |e_1| (\sin \angle A_3ED)^{-1} k_2 |e_2| (\sin \angle A_3ED)(|e_1| |e_2|)^{-1} = k_1 k_2 \in [0, 1].$$

From the above inequality, we have

$$1 + \left( \frac{\beta^+}{\beta^-} - 1 \right) \gamma^T \delta \geq \min \left( 1, \frac{\beta^+}{\beta^-} \right) \geq \beta_{\min} / \beta_{\max} > 0. \quad (3.11)$$

Hence, by the well-known Sherman–Morrison formula, the linear system (3.9) has a unique solution

$$c = b - \frac{(\beta^+ / \beta^- - 1)(\gamma^T b) \delta}{1 + (\beta^+ / \beta^- - 1) \gamma^T \delta}, \quad (3.12)$$

which completes the proof of the lemma.

\[ \square \]

**Remark 3.4.** In [17], the authors also consider the case that the curved interface is not discretized and the interface condition (1.3) is enforced on a point on the exact interface, i.e., replacing $T_h^+$ and $T_h^-$ in (3.1) by $T^+$ and $T^-$, respectively, and replacing (3.3) by

$$\beta^+ \nabla \phi^+ \cdot n(F) - \beta^- \nabla \phi^- \cdot n(F) = 0,$$

where $F$ is a point on $\Gamma \cap T$. Thus, the result for $c_0$ in (3.8) involves $n(F) \cdot n_h$ (see [17, Eq. (4.7)]). In this paper, we do not consider this case. We rigorously analyze the error caused by discretizing the curved interface by line segments.
3.2. Optimal approximation capabilities of IFE spaces

On each element $T \in \mathcal{T}_h$, define a local interpolation operator $I_{h,T} : W(T) \to V_h(T)$ such that

$$N_i(I_{h,T} v) = N_i(v) \quad \forall i \in \mathcal{I},$$

where $W(T) = \{ v : N_i(v), i \in \mathcal{I} \text{ are well defined} \}$. Similarly, on each interface element $T \in \mathcal{T}_h^\Gamma$, define $I_{h,T}^{\text{IFE}} : W(T) \to S_h(T)$ such that

$$N_i(I_{h,T}^{\text{IFE}} v) = N_i(v) \quad \forall i \in \mathcal{I}.$$  \hfill (3.13)

The global IFE interpolation operator is defined by $I_{h}^{\text{IFE}} : H^1(\Omega) \to V_h^{\text{IFE}}$ such that

$$(I_{h}^{\text{IFE}} v)|_T = \begin{cases} I_{h,T}^{\text{IFE}} v & \text{if } T \in \mathcal{T}_h^\Gamma, \\ I_{h,T} v & \text{if } T \in \mathcal{T}_h^{\text{non}}. \end{cases}$$

For simplicity, define $v^s := v|_{\Omega^s}$, $s = +, -$ for all $v \in L^2(\Omega)$. With a small ambiguity of notation, given a function $v_h \in S_h(T)$, we define $v^s_h \in V_h(T)$, $s = +, -$ such that

$$v^s_h = v_h|_{\Omega^s}, \quad s = +, -.$$  \hfill (3.14)

To show that functions in $V_h^{\text{IFE}}$ can approximate a function $v$ in $\tilde{H}^2(\Omega)$ optimally, we need to interpolate extensions of $v^s$, $s = +, -$. It is well-known that (see [13]) for any $v \in \tilde{H}^2(\Omega)$ there exist extensions $v^s_E \in H^2(\Omega)$, $s = +, -$ such that

$$v^s_E|_{\Omega^s} = v^s \quad \text{and} \quad \|v^s_E\|_{H^2(\Omega)} \leq C\|v^s\|_{H^2(\Omega^s)}, \quad s = +, -.$$  \hfill (3.15)

The next two theorems state the optimal approximation properties of the immersed finite element spaces. The proofs are technical and thus are presented in the appendix.

**Theorem 3.5.** For any $v \in \tilde{H}^2(\Omega)$, there exists a constant $C$ independent of $h$ and the interface location relative to the mesh such that

$$\sum_{T \in \mathcal{T}_h^\Gamma} h_T^{2(m-1)} \|v^s_E - (I_{h}^{\text{IFE}} v)^s\|_{H^m(T)}^2 \leq C h_T^{2}\|v\|_{H^2(\Omega^{+} \cup \Omega^{-})}^2, \quad m = 0, 1, 2, \quad s = +, -.$$  \hfill (3.16)

**Proof.** See Appendix A.1. \hfill \Box

The above result shows that the interpolation polynomial on one side of the interface can approximate the extensions of the exact solution optimally on the whole element $T$ no matter how small $T \cap \Omega^+$ or $T \cap \Omega^-$ might be. This is the key in deriving the optimal error estimates on interface edges; see (5.20) in the proof of Lemma 5.6.

Taking into account the mismatch of $\Gamma$ and $\Gamma_h$, we can prove the optimal approximation capabilities of the nonconforming IFE spaces.

**Theorem 3.6.** For any $v \in \tilde{H}^2(\Omega)$, there exists a constant $C$ independent of $h$ and the interface location relative to the mesh such that

$$\sum_{T \in \mathcal{T}_h} \|v - I_{h}^{\text{IFE}} v\|_{H^m(T)}^2 \leq C h_T^{4-2m}\|v\|_{H^2(\Omega^{+} \cup \Omega^{-})}^2, \quad m = 0, 1.$$  \hfill (3.17)

**Proof.** See Appendix A.2. \hfill \Box
4. Analysis of the Nonconforming IFE Method Without Penalties

In this section, we analyze the nonconforming IFE method without penalties which is obtained from (2.2) by simply replacing the Sobolev space $H_0^1(\Omega)$ with the nonconforming IFE space $V_{h,0}^{\text{IFE}}$ (see [25, 29]). The method reads: find $u_h \in V_{h,0}^{\text{IFE}}$ such that

$$a_h(u_h, v_h) := \sum_{T \in T_h} \int_T \beta(x) \nabla u_h \cdot \nabla v_h \, dx = \int_{\Omega} f v_h \, dx \quad \forall \, v_h \in V_{h,0}^{\text{IFE}}. \quad (4.1)$$

We will show that the nonconforming IFE method without penalties is not guaranteed to converge optimally unless $[\beta]_T \nabla u \cdot \mathbf{t} = 0$ on $\Gamma$.

It is easy to see that $a_h(\cdot, \cdot) = a(\cdot, \cdot)$ on $H_0^1(\Omega)$ and it is positive-definite on $V_{h,0}^{\text{IFE}}$ because $a_h(v_h, v_h) = 0$, $v_h \in V_{h,0}^{\text{IFE}}$ implies $v_h = 0$. Thus, the discrete problem (4.1) has a unique solution. We define the energy norm

$$\|v\|_{a_h} := \sqrt{a_h(v, v)} \quad \forall \, v \in V_{h,0}^{\text{IFE}} + H_0^1(\Omega)$$

and quote the following well-known second Strang lemma (see Lemma 10.1.9 in [6]).

Lemma 4.1. Let $u$ and $u_h$ be the solutions of (2.2) and (4.1), respectively. Then

$$\|u - u_h\|_{a_h} \leq C \left\{ \inf_{v_h \in V_{h,0}^{\text{IFE}}} \|u - v_h\|_{a_h} + \sup_{w_h \in V_{h,0}^{\text{IFE}} \setminus \{0\}} \frac{|a_h(u - u_h, w_h)|}{\|w_h\|_{a_h}} \right\}. \quad (4.2)$$

Since $u \in \tilde{H}^2(\Omega)$, Theorem 3.6 implies

$$\inf_{v_h \in V_{h,0}^{\text{IFE}}} \|u - v_h\|_{a_h} \leq \|u - I_h^{\text{IFE}} u\|_{a_h} \leq C h \|u\|_{H^2(\Omega \cup \Omega^c)} \quad (4.3)$$

For the second term on the right-hand side of (4.2), we have

$$a_h(u - u_h, w_h) = \sum_{T \in T_h} \int_T \beta \nabla u \cdot \nabla w_h \, dx - \int_{\Omega} f w_h \, dx = \sum_{e \in \mathcal{E}_h} \int_e \beta \nabla u \cdot \mathbf{n}_e [w_h]_e \, ds,$$

where the jump $[w_h]_e \mathbf{n}_e$ across an edge $e$ is defined as follows. Let $e$ be an interior edge shared by two elements $T_1^e$ and $T_2^e$, and $\mathbf{n}_e$ the unit normal of $e$ pointing towards the outside of $T_1^e$. Define

$$[w_h]_e \mathbf{n}_e = (w_h|_{T_1^e} - w_h|_{T_2^e}) \mathbf{n}_e \quad \text{on} \ e.$$

If $e$ is an edge on the boundary of $\Omega$, then define $[w_h]_e \mathbf{n}_e = w_h \mathbf{n}_e$, where $\mathbf{n}_e$ is the unit normal of $e$ pointing towards the outside of $\Omega$. Given an edge $e$ and an element $T$, define the standard $L^2$ projection operators $P_0^e$ and $P_0^T$ as

$$P_0^e f = |e|^{-1} \int_e f \, ds, \quad P_0^T f = |T|^{-1} \int_T f \, dx.$$

It follows from the fact $\int_e [w_h]_e \, ds = 0$ that

$$|a_h(u - u_h, w_h)| = \sum_{e \in \mathcal{E}_h} \int_e \beta \nabla u \cdot \mathbf{n}_e [w_h]_e \, ds = \sum_{e \in \mathcal{E}_h} \int_e (\beta \nabla u \cdot \mathbf{n}_e - P_0^e(\beta \nabla u \cdot \mathbf{n}_e))[w_h]_e \, ds \leq \left( \sum_{e \in \mathcal{E}_h} \|\beta \nabla u \cdot \mathbf{n}_e - P_0^e(\beta \nabla u \cdot \mathbf{n}_e)\|_{L^2(e)}^2 \right)^{1/2} \left( \sum_{e \in \mathcal{E}_h} \|[w_h]_e\|_{L^2(e)}^2 \right)^{1/2}. \quad (4.4)$$
Let $\mathcal{T}_h^e = \{ T \in \mathcal{T}_h : e \subset \partial T \}$ for all $e \in \mathcal{E}_h$. For any $w_h \in V_h^{\text{IFR}}$, using the fact $w_h|_T \in H^1(T)$, we have (see [6])
\[
\| [w_h]_e \|_{L^2(e)}^2 \leq C |e| \sum_{T \in \mathcal{T}_h^e} \| w_h \|_{H^1(T)}^2 \quad \forall e \in \mathcal{E}_h.
\]
(4.5)

Thus, the following estimate holds true:
\[
\sum_{e \in \mathcal{E}_h} \| [w_h]_e \|_{L^2(e)}^2 \leq Ch \| w_h \|_{H^2}^2.
\]
(4.6)

The next step is to estimate $\| \beta \nabla u \cdot n_e - P_0^e (\beta \nabla u \cdot n_e) \|_{L^2(e)}$ in (4.4). Let $T$ be an element such that $e \subset \partial T$.
If $e \in \mathcal{E}_h^{\text{non}}$ and $T \in \mathcal{T}_h^{\text{non}}$, we have the standard estimate
\[
\| \beta \nabla u \cdot n_e - P_0^e (\beta \nabla u \cdot n_e) \|_{L^2(e)} \leq Ch^{1/2} |u|_{H^1(T)}.
\]
(4.7)

If $e \in \mathcal{E}_h^{\text{pass}}$ and $T \in \mathcal{T}_h^T$, the term can be estimated by using the fact that $e \in \Omega^s$, $s = +$ or $-$,
\[
\| \beta \nabla u \cdot n_e - P_0^e (\beta \nabla u \cdot n_e) \|_{L^2(e)} \leq Ch^{1/2} ||\beta^s u|_{T}^e - P_0^e (\beta^s v|_{T}^e) \|_{L^2(e)}
\]
\[
\leq Ch^{1/2} ||\beta^s u|_{T}^e - P_0^e (\beta^s v|_{T}^e) \|_{L^2(T)} + Ch^{1/2} |\beta \nabla u \cdot n_e|_{H^1(T)}
\]
\[
\leq Ch^{1/2} |\beta^s u|_{T}^e + |\beta \nabla u \cdot n_e|_{H^1(T)}.
\]
(4.8)

Hence, it follows from (4.7) to (4.8) and the extension result (3.15) that
\[
\sum_{e \in \mathcal{E}_h^{\text{pass}}} \| \beta \nabla u \cdot n_e - P_0^e (\beta \nabla u \cdot n_e) \|_{L^2(e)}^2 \leq Ch \sum_{i=+,-} |\beta^s u|_{T}^e_{H^2(\Omega)}^2 \leq Ch |u|_{H^2(\Omega \cup + \Omega)}^2.
\]
(4.9)

For interface edges $e \in \mathcal{E}_h^T$, we cannot conclude the optimal estimate since $(\beta \nabla u \cdot n_e)|_e$ may have a jump across $e \cap \Gamma$. Noticing that $[u]_T = 0$ implies $[\nabla u \cdot t]_T = 0$, the jump can be derived as
\[
[\beta \nabla u \cdot n_e]_T = [\beta \nabla u \cdot n_e]_T + [\beta \nabla u \cdot t]_T \cdot n_e = [\beta]_T (\nabla u \cdot t)(t \cdot n_e),
\]
(4.10)

where we have used
\[
[\beta \nabla u \cdot n]_T = 0 \quad \text{and} \quad [\beta \nabla u \cdot t]_T = \frac{1}{2} (\beta^+ + \beta^-) [\nabla u \cdot t]_T + [\beta]_T (\nabla u \cdot t) = [\beta]_T \nabla u \cdot t.
\]

The following lemma gives an estimate on interface edges.

\textbf{Lemma 4.2.} Let $u$ be the solution of (2.2). Assume the triangulation near the interface is quasi-uniform, i.e., there exists a constant $c$ such that $h_T \geq ch_T^0$ for all $T \in \mathcal{T}_h^T$. Then there holds
\[
\sum_{e \in \mathcal{E}_h^T} \| \beta \nabla u \cdot n_e - P_0^e (\beta \nabla u \cdot n_e) \|_{L^2(e)}^2 \leq Ch^T \sum_{T \in \mathcal{T}_h^T} |u|_{H^2(T \cup + )}^2 + C \| \beta \nabla u \cdot t \|_{H^1/2}^2.
\]
(4.11)

\textbf{Proof.} Define a function $z|_{\Omega^s} = z^s$, $s = +, -$, such that
\[
-\Delta z^s + z^s = 0 \quad \text{in} \ \Omega^s,
\]
\[
z^s = [\beta]_T \nabla u \cdot t \text{ on } \Gamma, \quad \frac{\partial z^s}{\partial n} = 0 \text{ on } \partial \Omega, \quad s = +, -,
\]
where $\nu$ is the outward unit normal vector to $\partial \Omega$. Since $[\beta]_T \nabla u \cdot t \in H^{1/2}(\Gamma)$, the function $z$ exists and satisfies
\[
z|_T = [\beta]_T \nabla u \cdot t \text{ and } |z|_{H^1(\Omega)} \leq C \| \beta \nabla u \cdot t \|_{H^{1/2}(\Gamma)},
\]
(4.12)
We use uniform triangulations as shown in Figure 3 and only consider these interface edges which are contained in Λ. The estimate (4.11) is sharp.

From (4.10) to (4.12), we have \([w - \hat{w}]_{Γ} = 0\). Thus, \(w - \hat{w} \in H^1(T)\). By the property of the \(L^2\) projection operator \(P^e_0\) and the standard trace inequality, we infer

\[
\|w - P^e_0(w)\|^2_{L^2(\hat{e})} \leq \|w - P^T_0(w - \hat{w})\|^2_{L^2(\hat{e})} = \|w - \hat{w} + \hat{w} - P^T_0(w - \hat{w})\|^2_{L^2(\hat{e})} \\
\leq 2\|w - \hat{w} - P^T_0(w - \hat{w})\|^2_{L^2(\hat{e})} + 2\|\hat{w}\|^2_{L^2(\hat{e})} \\
\leq C\left(h^-_T\|w - \hat{w} - P^T_0(w - \hat{w})\|^2_{L^2(T)} + h_T|w - \hat{w}|^2_{H^1(T)}\right) + 2\|z\|^2_{L^2(\hat{e})} \\
\leq C\left(h^-_T|w - \hat{w}|^2_{H^1(T)} + h_T^{-1}\|z\|^2_{L^2(T)} + h_T|z|^2_{H^1(T)}\right).
\]

Summing over all interface edges and using Lemma 2.4, we get

\[
\sum_{e \in \mathcal{E}^I_h} \|\beta \nabla u \cdot n_e - P^e_0(\beta \nabla u \cdot n_e)\|^2_{L^2(\hat{e})} \leq \sum_{T \in \mathcal{T}^h} C\left(h^-_T|w|_{H^1(T)}^2 + h_T^{-1}\|z\|^2_{L^2(U(Γ, h_T))} + h_T|z|_{H^1(Ω)}^2\right) \\
\leq Ch^T \sum_{T \in \mathcal{T}^h} |u|^2_{H^2(T_{∪Γ}T)} + Ch^-_T\|z\|^2_{L^2(U(Γ, h_T))} + Ch_T|z|_{H^1(Ω)}^2 \\
\leq Ch^T \sum_{T \in \mathcal{T}^h} |u|^2_{H^2(T_{∪Γ}T)} + C\|z\|^2_{H^1(Ω)},
\]

which together with (4.12) yields this lemma. □

**Remark 4.3.** The estimate (4.11) is sharp, i.e., we cannot find a better upper bound for the approximation error than \(O(1)\) when \([β]_Γ \nabla u \cdot t \neq 0\) on \(Γ\). We explain it by a concrete example as illustrated in Figure 3. The domain \(Ω\) contains a region \(Λ\) such that \(Λ = (0, 1)^2, Λ^+ = \{x = (x_1, x_2) \in Λ : x_1 > x_2\}\), \(Λ^- = \{x = (x_1, x_2) \in Λ : x_1 < x_2\}\). The interface contained in the region is \(Γ_∩Λ = \{x = (x_1, x_2) \in Λ : x_1 = x_2\}\)

We use uniform triangulations as shown in Figure 3 and only consider these interface edges which are contained in the region \(Λ\). Obviously, \(n_e = t = (\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})^T\) and \(n = (\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}})^T\). Let \(β^+ = 2, β^- = 1\) and the exact solution \(u(x_1, x_2)\) be a piecewise linear function on \(Λ^+, Λ^-\) such that

\[
β^+ \nabla u^+ \cdot n = β^- \nabla u^- \cdot n = 1 \quad \text{and} \quad u|_{Γ∩Λ} = \frac{1}{\sqrt{2}}(x_1 + x_2).
\]

Thus, \(\nabla u \cdot t = 1\) and \((\beta \nabla u \cdot n_e)|_{e^+} = 2, (β \nabla u \cdot n_e)|_{e^-} = 1\) for all \(e \subset Λ\) and \(e \in \mathcal{E}^I_h\). Therefore,

\[
\|β \nabla u \cdot n_e - P^e_0(β \nabla u \cdot n_e)\|^2_{L^2(\hat{e})} = \inf_{c_e \in \mathbb{R}} \|β \nabla u \cdot n_e - c_e\|^2_{L^2(\hat{e})} \\
= \inf_{c_e \in \mathbb{R}} \left(\frac{|c_e|}{2}(2 - c_e)^2 + \frac{|c_e|}{2}(1 - c_e)^2\right) \geq \frac{|c_e|^2}{4} \geq Ch.
\]

Using the fact that the number of interface edges contained in \(Λ\) is \(O(h^{-1})\), we observe that

\[
\sum_{e \in \mathcal{E}^I_h} \|β \nabla u \cdot n_e - P^e_0(β \nabla u \cdot n_e)\|^2_{L^2(\hat{e})} \geq \sum_{e \in \mathcal{E}^I_h, e \subset Λ} Ch \geq C.
\]
Theorem 4.4. Let $u$ and $u_h$ be the solutions of (2.2) and (4.1), respectively. Under the assumption of Lemma 4.2, the following discretization error estimate holds true:

$$
\|u - u_h\|_{a_h} \leq C h \|u\|_{H^2(\Omega^+ \cup \Omega^-)} + C h^{1/2} [\beta]_T \|\nabla u \cdot t\|_{H^{1/2}(\Gamma)},
$$

(4.13)

Proof. It follows from (4.9) and (4.11) that

$$
\sum_{e \in \mathcal{E}_h} \|\beta \nabla u \cdot n_e - P_0^e (\beta \nabla u \cdot n_e)\|_{L^2(e)}^2 \leq C h \|u\|^2_{H^2(\Omega^+ \cup \Omega^-)} + C \|\beta\|_T \|\nabla u \cdot t\|^2_{H^{1/2}(\Gamma)},
$$

(4.14)

which together with (4.4) and (4.6) yields

$$
|a_h(u - u_h, w_h)| \leq C h^{1/2} w_h \left(h^{1/2} \|u\|_{H^2(\Omega^+ \cup \Omega^-)} + \|\beta\|_T \|\nabla u \cdot t\|_{H^{1/2}(\Gamma)}\right).
$$

(4.15)

Combining (4.2), (4.3) and (4.15), we complete the proof. $\square$

Theorem 4.4 suggests that the solution of the nonconforming IFE method (4.1) only converges with a suboptimal convergence rate $O(h^{1/2})$ in the energy norm if $[\beta]_T \nabla u \cdot t \neq 0$ on $\Gamma$. Applying global trace inequalities on $\Omega^+$ and $\Omega^-$ to the second term on the right-hand side of (4.13), we get the following corollary.

Corollary 4.5. Let $u$ and $u_h$ be the solutions of (2.2) and (4.1), respectively. Suppose that $[\beta]_T \neq 0$ and $\nabla u \cdot t \neq 0$ on $\Gamma$. Under the assumption of Lemma 4.2, we have

$$
\|u - u_h\|_{a_h} \leq C h \|u\|_{H^2(\Omega^+ \cup \Omega^-)}.
$$

Remark 4.6. Using the standard duality argument (see [6, p. 284]), we can also derive the following suboptimal $L^2$ error estimate:

$$
\|u - u_h\|_{L^2(\Omega)} \leq C h \|u\|_{H^2(\Omega^+ \cup \Omega^-)}.
$$
5. A NEW NONCONFORMING IFE METHOD AND ERROR ESTIMATES

To recover the optimal convergence rates, different from the partially penalized IFE method in [35], we propose a parameter-free nonconforming IFE method. On each interface element $T \in \mathcal{T}_h^1$, define

$$W_h(T) := \{ w_h \in (L^2(T))^2 : w_h = \nabla v_h \quad \forall v_h \in S_h(T) \}. \quad (5.1)$$

We also define a space associated with an edge $e \in \mathcal{E}_h^1$ as

$$W_e := \{ w_h \in (L^2(\Omega))^2 : w_h|_{T_1^e} \in W_h(T_1^e), \quad w_h|_{T_2^e} \in W_h(T_2^e), \quad w_h|_{\Omega \setminus (T_1^e \cup T_2^e)} = 0 \}, \quad (5.2)$$

where $T_1^e$ and $T_2^e$ are elements sharing the common edge $e$. For simplicity of the implementation, the coefficient $\beta(x)$ is approximated by

$$\beta_h(x) = \begin{cases} 
\beta^+(x) & \text{if } x \in \Omega_h^+, \\
\beta^-(x) & \text{if } x \in \Omega_h^-.
\end{cases} \quad (5.3)$$

Define a local lifting operator $r_e : L^2(e) \rightarrow W_e$ such that

$$\int_{\Omega} \beta_h(x)r_e(\varphi) \cdot w_h \, dx = \int_e \{ \beta_h w_h \cdot n_e \} \varphi \, ds \quad \forall w_h \in W_e, \quad (5.4)$$

where $\{v\}_e = \frac{1}{2}(v|_{T_1^e} + v|_{T_2^e})$. Obviously, $r_e(\varphi)$ exists uniquely for any $\varphi \in L^2(e)$ and the support of $r_e(\varphi)$ is $T_1^e \cup T_2^e$. We emphasize that the cost in computing $r_e(\varphi)$ with given $\varphi$ is not significant in general in practical implementation because the dimension of the space $W_h(T)$ is 2 for the Crouzeix–Raviart element and 3 for the rotated-$Q_1$ element. We also note that for the Crouzeix–Raviart element we can express $r_e(\varphi)$ explicitly if we choose orthogonal basis functions of the space $W_h(T)$ as

$$\omega_1(x) = t_h, \quad \omega_2(x) = \begin{cases} 
\beta^- n_h & \text{if } x \in T_h^+ \\
\beta^+ n_h & \text{if } x \in T_h^-.
\end{cases} \quad (5.5)$$

The new nonconforming IFE method is to find $u_h \in V_{h,0}^{\text{IFE}}$ such that

$$A_h(u_h, v_h) := \tilde{a}_h(u_h, v_h) + b_h(u_h, v_h) + s_h(u_h, v_h) = \int_{\Omega} fv_h \, dx \quad \forall v_h \in V_{h,0}^{\text{IFE}}, \quad (5.6)$$

where

$$\tilde{a}_h(u_h, v_h) = \sum_{T \in \mathcal{T}_h} \int_T \beta_h(x) \nabla u_h \cdot \nabla v_h \, dx,$$

$$b_h(u_h, v_h) = -\sum_{e \in \mathcal{E}_h^1} \int_e \left( \{ \beta_h \nabla u_h \cdot n_e \}_e [v_h]_e + \{ \beta_h \nabla v_h \cdot n_e \}_e [u_h]_e \right) \, ds,$$  \quad (5.7)

$$s_h(u_h, v_h) = 4 \sum_{e \in \mathcal{E}_h^1} \int_{T_1^e \cup T_2^e} \beta_h(x) r_e([u_h]_e) \cdot r_e([v_h]_e) \, dx.$$  \quad (5.8)

Clearly, the new nonconforming IFE method is symmetric and parameter-free. In our recent work [24], we studied a similar parameter-free IFE method using nodal values as degrees of freedom. However, the method has a limitation, i.e., the maximum angle of the interface elements should be less than or equal to $\pi/2$. In contrast, the nonconforming IFE method proposed in this paper works without this restriction. The IFE basis functions
Lemma 5.1. We have with respect to the norm $\|v\|_h := \sqrt{a_h(v, v)}$

and

$$\|v\|_h := \left( \|v\|_{\bar{a}_h}^2 + \sum_{e \in \mathcal{E}_h^I} |e| \| \{ \beta_h \nabla v \}_e \|_{L^2(e)}^2 + \sum_{e \in \mathcal{E}_h^I} |e|^{-1} \| [v]_e \|_{L^2(e)}^2 + s_h(v, v) \right)^{1/2}.$$  \hspace{1cm} (5.6)

The continuity of the bilinear form $A_h(\cdot, \cdot)$ is verified directly from the Cauchy–Schwarz inequality

$$|A_h(w, v)| \leq \|w\|_h \|v\|_h \quad \forall \ w, v \in H^1_0(\Omega) \cap H^2(\Omega) + V_{h, 0}^{\text{IFE}}.$$  \hspace{1cm} (5.7)

The following lemma demonstrates the coercivity of the bilinear form $A_h(\cdot, \cdot)$ on the IFE space $V_{h, 0}^{\text{IFE}}$ with respect to the norm $\| \cdot \|_{\bar{a}_h}$.

**Lemma 5.1.** We have

$$A_h(v_h, v_h) \geq \frac{1}{2} \|v_h\|_{\bar{a}_h}^2 \quad \forall \ v_h \in V_{h, 0}^{\text{IFE}}.$$  \hspace{1cm} (5.8)

**Proof.** For all $v_h \in V_{h, 0}^{\text{IFE}}$, choosing $w_h|_{T^+_1 \cup T^+_2} = \nabla v_h$, $w_h|_{\Omega \setminus (T^+_1 \cup T^+_2)} = 0$ in (5.3) and using the fact that the support of $r_e(\varphi)$ is $T^+_1 \cup T^+_2$, we have

$$\int_{T^+_1 \cup T^+_2} \beta_h(x) r_e(\varphi) \cdot \nabla v_h \, dx = \int_{\Omega} \beta_h(x) r_e(\varphi) \cdot \nabla v_h \, dx = \int_{\Omega} \{ \beta_h \nabla v_h \cdot n_e \}_e \varphi \, ds.$$  \hspace{1cm} (5.9)

It follows from the Cauchy–Schwarz inequality that

$$|b_h(v_h, v_h)| = \left| 2 \sum_{e \in \mathcal{E}_h^I} \int_{\Omega} \{ \beta_h \nabla v_h \cdot n_e \}_e [v_h]_e \, ds \right| \leq \left( 4 \sum_{e \in \mathcal{E}_h^I} \int_{\Omega} \beta_h r_e([v_h]_e) \cdot [v_h]_e (\nabla v_h) \, dx \right)^{1/2} \left( \sum_{e \in \mathcal{E}_h^I} \int_{T^+_1 \cup T^+_2} \beta_h \nabla v_h \cdot \nabla v_h \, dx \right)^{1/2}.$$  \hspace{1cm} (5.10)

Since each interface element has at most two interface edges from Assumption 2.2, each element is calculated at most twice. Therefore,

$$\sum_{e \in \mathcal{E}_h^I} \int_{T^+_1 \cup T^+_2} \beta_h \nabla v_h \cdot \nabla v_h \, dx \leq 2 \sum_{T \in T_h} \int_T \beta_h(x) \nabla v_h \cdot \nabla v_h \, dx.$$  \hspace{1cm} (5.11)

Combining (5.5), (5.9) and (5.10), we have

$$|b_h(v_h, v_h)| \leq (s_h(v_h, v_h))^{1/2} \left( 2 \sum_{T \in T_h} \int_T \beta_h(x) \nabla v_h \cdot \nabla v_h \, dx \right)^{1/2} \leq s_h(v_h, v_h) + \frac{1}{2} \sum_{T \in T_h} \int_T \beta_h(x) \nabla v_h \cdot \nabla v_h \, dx,$$
which together with (5.5) yields the result

\[
A_h(v_h, v_h) = \tilde{a}_h(v_h, v_h) + b_h(v_h, v_h) + s_h(v_h, v_h) 
\geq \frac{1}{2} \sum_{T \in T_h} \int_T \beta_h(x) \nabla v_h \cdot \nabla v_h \, dx = \frac{1}{2} \| v_h \|_{\tilde{a}_h}^2.
\]

Next, we show the equivalence of the \( \| \cdot \|_{\tilde{a}_h} \)-norm and the \( \| \cdot \|_{a_h} \)-norm on the IFE space \( V_{h,0}^{IFE} \). To begin with, we need the following trace inequality for the IFE shape functions in \( S_h(T) \) which can be verified via straightforward calculations. We also refer readers to [29, Thm. 2.7].

**Lemma 5.2.** There exists a constant \( C \) independent of \( h \) and the interface location relative to the mesh such that

\[
\| \nabla v_h \|_{L^2(\partial T)} \leq Ch^{-1/2} \| \nabla v_h \|_{L^2(T)} \quad \forall \, v_h \in S_h(T) \quad \forall \, T \in T_h^\Gamma.
\]

(5.11)

We also need the following stability estimate for the local lifting operator \( r_e \).

**Lemma 5.3.** There exists a constant \( C \) independent of \( h \) and the interface location relative to the mesh such that

\[
\| r_e(\varphi) \|_{L^2(\Omega)} \leq C|e|^{-1/2} \| \varphi \|_{L^2(e)} \quad \forall \, \varphi \in L^2(e) \quad \forall \, e \in E_h^\Gamma.
\]

(5.12)

**Proof.** Since the support of \( r_e(\varphi) \) is \( T_e^+ \cup T_e^- \), choosing \( w_h = r_e(\varphi) \) in (5.3) yields

\[
\| r_e(\varphi) \|_{L^2(\Omega)}^2 \leq C \left\| \beta_h^{1/2} r_e(\varphi) \right\|_{L^2(T_e^+ \cup T_e^-)}^2 = C \int_{T_e} \left\{ \beta_h r_e(\varphi) \cdot n_e \right\} e \varphi \, ds
\leq C \| \{ \beta_h r_e(\varphi) \} e \|_{L^2(e)} \| \varphi \|_{L^2(e)} \leq C \| \varphi \|_{L^2(e)} \sum_{i=1,2} \| r_e(\varphi) \|_{T_e^\Gamma} \| \varphi \|_{L^2(e)}.
\]

Note \( r_e(\varphi)|_{T_e^\Gamma} \in W_h(T_e^\Gamma) \), we know from (5.1) that there exists a function \( v_h \in S_h(T_e^\Gamma) \) such that \( r_e(\varphi)|_{T_e^\Gamma} = \nabla v_h \). By the trace inequality (5.11) for the IFE shape functions, we have

\[
\| r_e(\varphi) \|_{T_e^\Gamma} \|_{L^2(e)} = \| \nabla v_h \|_{L^2(e)} \leq Ch^{-1/2} \| \nabla v_h \|_{L^2(T_e^\Gamma)} = Ch^{-1/2} \| r_e(\varphi) \|_{L^2(T_e^\Gamma)},
\]

which, together with (5.12) and a similar estimate on \( T_e^- \), completes the proof of this lemma. \( \square \)

**Remark 5.4.** Lemma 5.3 indicates that the parameter-free stabilization is weaker than the standard \( L^2 \) stabilization. Thus, our analysis is also valid for the nonconforming PPIFE method in [35], i.e., replacing \( s_h \) in (5.4) by \( s_h \) defined by

\[
s_h(u_h, v_h) = \sum_{e \in E_h^\Gamma} \eta_e |e| \int_e [u_h] e [v_h] e \, dx,
\]

where \( \eta_e > 0 \) should be sufficiently large.

We now prove the norm-equivalence in the following lemma.

**Lemma 5.5.** There exists a constant \( C \) independent of \( h \) and the interface location relative to the mesh such that

\[
\| v_h \|_{a_h} \leq \| v_h \|_{\tilde{a}_h} \leq C \| v_h \|_{a_h} \quad \forall \, v_h \in V_{h,0}^{IFE}.
\]

(5.13)
Proof. We just need to prove the second inequality since the first inequality is obvious. By the trace inequality (5.11) for the IFE shape functions, we can see that
\[
\sum_{e \in \mathcal{E}_h^I} |e| \|(\beta_h \nabla v_e)\|_{L^2(e)}^2 \leq C \sum_{e \in \mathcal{E}_h^I} \sum_{T \in T_h^e} \|\nabla v_h\|_{L^2(T)}^2 \leq C \|v_h\|_{a_h}^2.
\] (5.14)

From (4.5), we have
\[
\sum_{e \in \mathcal{E}_h^I} |e|^{-1} \|v_e\|_{L^2(e)}^2 \leq C \sum_{e \in \mathcal{E}_h^I} \sum_{T \in T_h^e} \|\nabla v_h\|_{L^2(T)}^2 \leq C \|v_h\|_{a_h}^2,
\] (5.15)
which, together with Lemma 5.3 for the local lifting operator, leads to
\[
s_h(v_h, v_h) \leq C \sum_{e \in \mathcal{E}_h^I} \|r_e(v_e)\|_{L^2(e)}^2 \leq C \sum_{e \in \mathcal{E}_h^I} |e|^{-1} \|v_e\|_{L^2(e)}^2 \leq C \|v_h\|_{a_h}^2.
\] (5.16)

Combining (5.6), (5.14)–(5.16), we get the second inequality in (5.13).

The following lemma provides an optimal estimate for the interpolation error in terms of the norm \(|| \cdot ||_h\).

**Lemma 5.6.** Suppose \(v \in \widetilde{H}^2(\Omega)\), then there exists a constant \(C\) independent of \(h\) and the interface location relative to the mesh such that
\[
|||v - I_h^{\text{IFE}}v|||_h \leq Ch ||v||_{H^2(\Omega^+ \cup \Omega^-)}.
\]

**Proof.** The first term in the norm \(||| \cdot ||_h\) can be bounded by Theorem 3.6,
\[
|||v - I_h^{\text{IFE}}v|||_a \leq Ch ||v||_{H^2(\Omega^+ \cup \Omega^-)}.
\] (5.17)

Since \((v - I_h^{\text{IFE}}v)|_T \in H^1(T)\) for all \(T \in T_h^I\), by the standard trace inequality and Theorem 3.5, we have
\[
\sum_{e \in \mathcal{E}_h^I} |e|^{-1} \|v - I_h^{\text{IFE}}v\|_{L^2(e)}^2 \leq C \sum_{e \in \mathcal{E}_h^I} \sum_{T \in T_h^e} \left(h_T^{-2} \|v - I_h^{\text{IFE}}v\|_{L^2(T)}^2 + \|v - I_h^{\text{IFE}}v\|_{H^1(T)}^2 \right)
\] (5.18)
which, together with Lemma 5.3, implies
\[
s_h(v - I_h^{\text{IFE}}v, v - I_h^{\text{IFE}}v) \leq C \sum_{e \in \mathcal{E}_h^I} |e|^{-1} \|v - I_h^{\text{IFE}}v\|_{L^2(e)}^2 \leq Ch^2 ||v||_{H^2(\Omega^+ \cup \Omega^-)}^2.
\] (5.19)

Let \(e^s = e \cap \Omega^s\), \(s = +, -\). Recalling the notations in (3.14) and (3.15), we have
\[
\|\{ \beta_h \nabla (v - I_h^{\text{IFE}}v) \}_e \|_{L^2(e)}^2 = \| \{ \beta_h \nabla (v - I_h^{\text{IFE}}v) \}_e \|_{L^2(e^+)}^2 + \| \{ \beta_h \nabla (v - I_h^{\text{IFE}}v) \}_e \|_{L^2(e^-)}^2 \leq C \| \{ \nabla (v_E^+ - (I_h^{\text{IFE}}v)^+) \}_e \|_{L^2(e^+)}^2 + C \| \{ \nabla (v_E^- - (I_h^{\text{IFE}}v)^-) \}_e \|_{L^2(e^-)}^2.
\]

Then using the standard trace inequality and Theorem 3.5, we infer
\[
\sum_{e \in \mathcal{E}_h^I} |e| \|\{ \beta_h \nabla (v - I_h^{\text{IFE}}v) \}_e \|_{L^2(e)}^2 \leq C \sum_{T \in T_h^s} \sum_{e = +, -} \left( \|v_E^s - (I_h^{\text{IFE}}v)^s\|_{H^1(T)}^2 + h_T^2 \|v_E^s - (I_h^{\text{IFE}}v)^s\|_{H^2(T)}^2 \right)
\] (5.20)
\[\leq Ch^2 ||v||_{H^2(\Omega^+ \cup \Omega^-)}^2.
\]
The lemma follows from (5.6), (5.17)–(5.20).

The following lemma concerns the errors caused by replacing \( \beta(x) \) by \( \beta_h(x) \).

**Lemma 5.7.** Let \( v \in \tilde{H}^2(\Omega) \) and \( w \in V_h^{\text{IFE}} + H^1(\Omega) \). Then there exists a constant \( C \) independent of \( h \) and the interface location relative to the mesh such that

\[
|a_h(v, w) - \tilde{a}_h(v, w)| \leq Ch_\Gamma \| v \|_{H^2(\Omega^+ \cup \Omega^-)} \left( \sum_{T \in T_h^\Gamma} \| \nabla w \|_{L^2(T^\Delta)}^2 \right)^{1/2}.
\]

(5.21)

Furthermore, if \( w \in \tilde{H}^2(\Omega) \), there holds

\[
|a_h(v, w) - \tilde{a}_h(v, w)| \leq Ch_\Gamma^2 \| v \|_{H^2(\Omega^+ \cup \Omega^-)} \| w \|_{H^2(\Omega^+ \cup \Omega^-)}.
\]

(5.22)

**Proof.** The Cauchy–Schwarz inequality gives

\[
|a_h(v, w) - \tilde{a}_h(v, w)| = \left| \sum_{T \in T_h^\Gamma} \int_{T^\Delta} (\beta - \beta_h) \nabla v \cdot \nabla w \, dx \right|
\]

\[
\leq C \left( \sum_{T \in T_h^\Gamma} \| \nabla v \|_{L^2(T^\Delta)}^2 \right)^{1/2} \left( \sum_{T \in T_h^\Gamma} \| \nabla w \|_{L^2(T^\Delta)}^2 \right)^{1/2}.
\]

(5.23)

Using Lemma 2.5 and the global trace inequalities on \( \Omega^+ \) and \( \Omega^- \), we can see that

\[
\sum_{T \in T_h^\Gamma} \| \nabla v \|_{L^2(T^\Delta)}^2 = \sum_{T \in T_h^\Gamma} \sum_{s=+,-} \| \nabla v^s \|_{L^2(T^\Delta \cap T^s)}^2 \leq \sum_{T \in T_h^\Gamma} \sum_{s=+,-} \| \nabla v^s \|_{L^2(T^\Delta)}^2
\]

\[
\leq C \sum_{T \in T_h^\Gamma} \sum_{s=+,-} \left( h_T^2 \| \nabla v^s \|_{L^2(T \cap \Gamma)}^2 + h_T^4 |v^s_E|_{H^2(T^\Delta)}^2 \right)
\]

\[
\leq Ch_\Gamma^2 \sum_{s=+,-} \| \nabla v^s \|_{L^2(\Gamma)}^2 + Ch_\Gamma^4 \sum_{s=+,-} |v^s_E|_{H^2(\Omega)}^2
\]

\[
\leq Ch_\Gamma^2 \sum_{s=+,-} \| v^s \|_{H^2(\Omega^s)}^2 = Ch_\Gamma^2 \| v \|_{H^2(\Omega^+ \cup \Omega^-)}^2.
\]

(5.24)

The estimate (5.21) follows from (5.23) and (5.24). If \( w \in \tilde{H}^2(\Omega) \), then similar to (5.24),

\[
\sum_{T \in T_h^\Gamma} \| \nabla w \|_{L^2(T^\Delta)}^2 \leq Ch_\Gamma^2 \| w \|_{H^2(\Omega^+ \cup \Omega^-)}^2.
\]

(5.25)

The estimate (5.22) then follows from (5.21) and (5.25).

With these preparations, we are ready to derive the \( H^1 \) error estimate for the new nonconforming IFE method.

**Theorem 5.8.** Let \( u \) and \( u_h \) be the solutions of (2.2) and (5.4), respectively. Then there exists a constant \( C \) independent of \( h \) and the interface location relative to the mesh such that

\[
\| u - u_h \|_h \leq Ch \| u \|_{H^2(\Omega^+ \cup \Omega^-)}.
\]

(5.26)
Proof. From Lemmas 5.1 and 5.5, we know that the bilinear form $A_h(\cdot, \cdot)$ is also coercive on $V_{h,0}^{\text{IFE}}$ with respect to the norm $\| \cdot \|_h$. Thus, the second Strang lemma implies

$$\| u - u_h \|_h \leq C \left\{ \inf_{v_h \in V_{h,0}^{\text{IFE}}} \| u - v_h \|_h + \sup_{w_h \in V_{h,0}^{\text{IFE}} \setminus \{0\}} \frac{|A_h(u - u_h, w_h)|}{\| w_h \|_h} \right\}. \quad (5.27)$$

The first term of the right-hand side of (5.27) can be bounded by Lemma 5.6

$$\inf_{v_h \in V_{h,0}^{\text{IFE}}} \| u - v_h \|_h \leq \| u - I_h^{\text{IFE}} u \|_h \leq C h \| u \|_{H^2(\Omega^+ \cup \Omega^-)} \quad (5.28)$$

Multiplying (1.1) by $w_h \in V_{h,0}^{\text{IFE}}$ and using integration by parts, we obtain

$$\int_{\Omega} f w_h \, dx = a_h(u, w_h) + s_h(u, w_h) - \sum_{e \in \mathcal{E}_h} \int_e \{ \beta \nabla u \cdot \mathbf{n}_e \}_e [w_h]_e + \{ \beta \nabla w_h \cdot \mathbf{n}_e \}_e [u]_e \, ds$$

$$= A_h(u, w_h) - \sum_{e \in \mathcal{E}_{h,\text{con}}} \int_e \beta \nabla u \cdot \mathbf{n}_e [w_h]_e \, ds + a_h(u, w_h) - \tilde{a}_h(u, w_h), \quad (5.29)$$

where $\beta_h(x) = \beta(x)$ on edges, $[u]_e = 0$, $[\beta \nabla u \cdot \mathbf{n}_e]_e = 0$ and $r_e([u]_e) = 0$ are used. It follows from (5.29) and (5.4) that

$$A_h(u - u_h, w_h) = \sum_{e \in \mathcal{E}_{h,\text{con}}} \int_e \beta \nabla u \cdot \mathbf{n}_e [w_h]_e \, ds + \tilde{a}_h(u, w_h) - a_h(u, w_h). \quad (5.30)$$

Hence, by (4.4), (4.6), (4.9), and Lemma 5.7, we have

$$|A_h(u - u_h, w_h)| \leq C h \| u \|_{H^2(\Omega^+ \cup \Omega^-)} \| w_h \|_h,$$

which, together with (5.27) and (5.28), completes the proof of the theorem. \qed

The optimal $L^2$ error estimate is also derived by using the standard duality argument below.

Theorem 5.9. Let $u$ and $u_h$ be the solutions of (2.2) and (5.4), respectively. Then there exists a constant $C$ independent of $h$ and the interface location relative to the mesh such that

$$\| u - u_h \|_{L^2(\Omega)} \leq C h \| u \|_{H^2(\Omega^+ \cup \Omega^-)}. \quad (5.31)$$

Proof. Let $z \in H^1_0(\Omega)$ be the solution of the following auxiliary problem

$$a(v, z) = \int_{\Omega} (u - u_h) v \, dx \quad \forall \, v \in H^1_0(\Omega). \quad (5.32)$$

Since $u - u_h \in L^2(\Omega)$, it follows from Theorem 2.1 that

$$z \in \tilde{H}^2(\Omega) \quad \text{and} \quad \| z \|_{H^2(\Omega^+ \cup \Omega^-)} \leq C \| u - u_h \|_{L^2(\Omega)}. \quad (5.33)$$

Let $z_h \in V_{h,0}^{\text{IFE}}$ be the solution of the new nonconforming IFE method applied to the auxiliary problem (5.32), i.e.,

$$A_h(v_h, z_h) = \int_{\Omega} (u - u_h) v_h \, dx \quad \forall \, v_h \in V_{h,0}^{\text{IFE}}. \quad (5.34)$$
Recalling that \( a_h(\cdot, \cdot) = a(\cdot, \cdot) \) on \( H^1_0(\Omega) \), and applying (5.32) and (5.34), we have
\[
\|u - u_h\|_{L^2(\Omega)}^2 = a(u, z) - A_h(u_h, z_h) = A_h(u, z) - A_h(u_h, z_h) - \tilde{a}_h(u, z) + a_h(u, z)
\]
\[
= A_h(u - u_h, z - z_h) + A_h(u - u_h, z_h) + A_h(u_h, z - z_h)
\]
\[
+ \left( a_h(u, z) - \tilde{a}_h(u, z) \right),
\]
where the relation \( b_h(u, z) = s_h(u, z) = 0 \) is used in the second identity since \([u]_e = [v]_e = 0\) for all edges. Lemma 5.7 provides the estimate for the last term
\[
a_h(u, z) - \tilde{a}_h(u, z) \leq C h^2 \|u\|_{H^2(\Omega + \cup_{\partial}\Omega)} \|z\|_{H^2(\Omega + \cup_{\partial}\Omega)}. \quad (5.36)
\]
The first terms on the right-hand side of (5.35) can be estimated using Theorem 5.8,
\[
A_h(u - u_h, z - z_h) \leq \|u - u_h\|_h \|z - z_h\|_h \leq C h^2 \|u\|_{H^2(\Omega + \cup_{\partial}\Omega)} \|z\|_{H^2(\Omega + \cup_{\partial}\Omega)}. \quad (5.37)
\]
We rewrite the second term on the right-hand side of (5.35) as
\[
A_h(u - u_h, z_h) = A_h(u - u_h, z_h - I_h^{IFE} z) + A_h(u - u_h, I_h^{IFE} z_h). \quad (5.38)
\]
It is easy to see that
\[
A_h(u - u_h, z_h - I_h^{IFE} z) \leq \|u - u_h\|_h \|z_h - I_h^{IFE} z\|_h \leq C h^2 \|u\|_{H^2(\Omega + \cup_{\partial}\Omega)} \|z\|_{H^2(\Omega + \cup_{\partial}\Omega)}. \quad (5.39)
\]
From (5.30), we have
\[
A_h(u - u_h, I_h^{IFE} z_h) = \sum_{e \in E_h^{non}} \int_{e} \beta \nabla u \cdot n_e \left[ I_h^{IFE} z_h \right]_{e} \, ds + \tilde{a}_h(u, I_h^{IFE} z_h) - a_h(u, I_h^{IFE} z_h). \quad (5.40)
\]
Since \([z]_e = 0\), the first term on the right-hand side can be estimated as
\[
\sum_{e \in E_h^{non}} \int_{e} \beta \nabla u \cdot n_e \left[ I_h^{IFE} z_h \right]_{e} \, ds = \sum_{e \in E_h^{non}} \int_{e} (\beta \nabla u \cdot n_e - P^e_0(\beta \nabla u \cdot n_e)) \left[ I_h^{IFE} z_h - z \right]_{e} \, ds \leq C h^2 \|u\|_{H^2(\Omega + \cup_{\partial}\Omega)} \|z\|_{H^2(\Omega + \cup_{\partial}\Omega)}, \quad (5.41)
\]
where the Cauchy–Schwarz inequality, (4.9), the standard trace inequality and Theorem 3.6 are used. Applying Lemma 5.7 and Theorem 3.6 again we obtain
\[
|\tilde{a}_h(u, I_h^{IFE} z_h) - a_h(u, I_h^{IFE} z_h)| \leq |\tilde{a}_h(u, z) - a_h(u, z)| + |\tilde{a}_h(u, I_h^{IFE} z_h - z) - a_h(u, I_h^{IFE} z_h - z)| \leq C h^2 \|u\|_{H^2(\Omega + \cup_{\partial}\Omega)} \|z\|_{H^2(\Omega + \cup_{\partial}\Omega)}. \quad (5.42)
\]
Combining (5.38)–(5.42), we find
\[
A_h(u - u_h, z_h) \leq C h^2 \|u\|_{H^2(\Omega + \cup_{\partial}\Omega)} \|z\|_{H^2(\Omega + \cup_{\partial}\Omega)}, \quad (5.43)
\]
and similarly,
\[
A_h(u_h, z - z_h) \leq C h^2 \|u\|_{H^2(\Omega + \cup_{\partial}\Omega)} \|z\|_{H^2(\Omega + \cup_{\partial}\Omega)}. \quad (5.44)
\]
Applying (5.35)–(5.37), (5.43)–(5.44), we arrive at the estimate
\[
\|u - u_h\|_{L^2(\Omega)}^2 \leq C h^2 \|u\|_{H^2(\Omega + \cup_{\partial}\Omega)} \|z\|_{H^2(\Omega + \cup_{\partial}\Omega)},
\]
which together with the regularity result (5.33) implies the estimate (5.31). □
6. Numerical examples

In this section, we present some numerical examples to validate the theoretical analysis. To avoid redundancy, we only report numerical results of IFE methods based on the Crouzeix–Raviart element since the results of IFE methods based on the rotated-$Q_1$ element are almost the same. We examine the convergence rate of IFE solutions using the following norms

$$|e_h|_{H^1} := \left( \sum_{T \in T_h} \left\| \beta_h \nabla (u - u_h) \right\|_{L^2(T)}^2 \right)^{1/2} \quad \text{and} \quad \|e_h\|_{L^2} := \|u - u_h\|_{L^2(\Omega)}.$$

For comparison, we replace $\beta(x)$ by $\beta_h(x)$ in the nonconforming IFE method (IFEM) without penalties (4.1) in our computation. Thus, the difference between the nonconforming IFEM without penalties and our new nonconforming IFEM (5.4) is the terms $b_h(\cdot, \cdot)$ and $s_h(\cdot, \cdot)$. In view of the analysis for our new method, the error resulting from replacing $\beta(x)$ by $\beta_h(x)$ does not affect the error estimates in Theorem 4.4 for the nonconforming IFE without penalties.

In all numerical examples, we set $\Omega = (-1,1) \times (-1,1)$ and use uniform meshes obtained as follows. We first partition the domain into $N \times N$ congruent rectangles, and then obtain the triangulation by cutting the rectangles along one of diagonals in the same direction (see Fig. 3). The interface $\Gamma$ and the subdomains $\Omega$ partition the domain into $\text{IFEM without penalties}$. For comparison, we replace $\beta(x)$ by $\beta_h(x)$ in the nonconforming IFE method (IFEM) without penalties (4.1). When the terms $b_h(\cdot, \cdot)$ and $s_h(\cdot, \cdot)$ are added to the scheme, i.e., the new nonconforming IFE (5.4), we observe the optimal convergence rates (see last two columns in Tab. 1 and 2).

6.1. A counter example with $\nabla u \cdot t \neq 0$ on $\Gamma$

We use this example to show that the nonconforming IFE method without penalties does not converge optimally, although the integral values on edges are used as degrees of freedom.

Example 6.1. We set $\varphi(x_1, x_2) = x_1^2 + x_2^2 - r_0^2$. Let $(r, \theta)$ be the polar coordinate of $x = (x_1, x_2)$. The exact solution is chosen as $u(x) = j(x)v(x)\omega(x)$, where $\omega(x) = \sin(\theta)$,

$$j(x) = \begin{cases} \exp\left(-\frac{1}{1-(r-r_0)^2/\eta^2}\right) & \text{if } |r-r_0| < \eta, \\ 0 & \text{if } |r-r_0| \geq \eta, \end{cases}$$

and

$$v(x) = \begin{cases} 1 + (r^2 - r_0^2)/\beta^+(x) & \text{if } x \in \Omega^+, \\ 1 + (r^2 - r_0^2)/\beta^-(x) & \text{if } x \in \Omega^-. \end{cases}$$

Let $r_0 = 0.5$, $\eta = 0.45$, $\beta^+(x)$ and $\beta^-(x)$ be positive constants. It is easy to verify that the jump condition (1.2) and (1.3) is satisfied and $\nabla u \cdot t \neq 0$ on $\Gamma$. We test two cases: $(\beta^+, \beta^-) = (10, 1000)$ and $(\beta^+, \beta^-) = (1000, 10)$. The exact solutions of these two cases are plotted in Figure 4.

We report numerical results in Tables 1 and 2 which clearly confirm our theoretical analysis. The second and third columns in Tables 1 and 2 indicate suboptimal convergence rates: $\|e_h\|_{L^2} \approx O(h)$, $|e_h|_{H^1} \approx O(h^{1/2})$ for the nonconforming IFE without penalties (4.1). When the terms $b_h(\cdot, \cdot)$ and $s_h(\cdot, \cdot)$ are added to the scheme, i.e., the new nonconforming IFE (5.4), we observe the optimal convergence rates (see last two columns in Tab. 1 and 2).

We also use this example to test the classic IFE method [21, 27] where the nodal values are used as degrees of freedom and no penalties are included. The suboptimal convergence rates are also observed (i.e., $O(h^{1/2})$ in the $H^1$ norm and $O(h)$ in the $L^2$ norm), which indicates that the error estimate in [21] is sharp. To avoid redundancy, we do not list the numerical results here.
Figure 4. Exact solutions of Example 6.1. Left: \((\beta^+, \beta^-) = (10, 1000)\); Right: \((\beta^+, \beta^-) = (1000, 10)\).

Table 1. Numerical results of Example 6.1 with \((\beta^+, \beta^-) = (10, 1000)\).

| \(N\) | \(\|e_h\|_{L^2}\) | Rate | \(\|e_h\|_{H^1}\) | Rate | \(\|e_h\|_{L^2}\) | Rate | \(\|e_h\|_{H^1}\) | Rate |
|---|---|---|---|---|---|---|---|---|
| 8  | 2.221E-01 |     | 2.140E+01 |     | 1.617E-01 |     | 1.781E+01 |     |
| 16 | 7.650E-02 | 1.54 | 1.037E+01 | 1.05 | 4.414E-02 | 1.87 | 6.889E+00 | 1.37 |
| 32 | 1.745E-02 | 2.13 | 5.970E+00 | 0.80 | 5.989E-03 | 2.88 | 3.851E+00 | 0.84 |
| 64 | 7.322E-03 | 1.25 | 3.597E+00 | 0.73 | 7.855E-04 | 2.93 | 1.784E+00 | 1.11 |
| 128| 3.204E-03 | 1.19 | 2.309E+00 | 0.64 | 1.935E-04 | 2.02 | 8.932E-01 | 1.00 |
| 256| 1.514E-03 | 1.08 | 1.548E+00 | 0.58 | 4.836E-05 | 2.00 | 4.461E-01 | 1.00 |
| 512| 7.276E-04 | 1.06 | 1.056E+00 | 0.55 | 1.197E-05 | 2.01 | 2.229E-01 | 1.00 |
| 1024| 3.603E-04 | 1.01 | 7.378E-01 | 0.52 | 2.992E-06 | 2.00 | 1.114E-01 | 1.00 |

Table 2. Numerical results of Example 6.1 with \((\beta^+, \beta^-) = (1000, 10)\).

| \(N\) | \(\|e_h\|_{L^2}\) | Rate | \(\|e_h\|_{H^1}\) | Rate | \(\|e_h\|_{L^2}\) | Rate | \(\|e_h\|_{H^1}\) | Rate |
|---|---|---|---|---|---|---|---|---|
| 8  | 1.736E-01 |     | 4.641E+01 |     | 1.540E-01 |     | 4.611E+01 |     |
| 16 | 7.679E-02 | 1.18 | 1.686E+01 | 1.46 | 6.402E-02 | 1.27 | 1.579E+01 | 1.55 |
| 32 | 1.186E-02 | 2.69 | 8.927E+00 | 0.92 | 5.868E-03 | 3.45 | 8.055E+00 | 0.97 |
| 64 | 5.188E-03 | 1.19 | 4.822E+00 | 0.89 | 8.575E-04 | 2.77 | 3.888E+00 | 1.05 |
| 128| 2.242E-03 | 1.21 | 2.802E+00 | 0.78 | 1.944E-04 | 2.14 | 1.949E+00 | 1.00 |
| 256| 1.061E-03 | 1.08 | 1.746E+00 | 0.68 | 4.841E-05 | 2.01 | 9.738E-01 | 1.00 |
| 512| 5.043E-04 | 1.07 | 1.133E+00 | 0.62 | 1.218E-05 | 1.99 | 4.868E-01 | 1.00 |
| 1024| 2.483E-04 | 1.02 | 7.654E-01 | 0.57 | 2.982E-06 | 2.03 | 2.434E-01 | 1.00 |
6.2. An example with variable coefficients and a non-convex interface

Example 6.2. We set

$$\varphi(x_1, x_2) = (3(x_1^2 + x_2^2) - x_1)^2 - x_1^2 - x_2^2 + 0.02.$$  

The exact solution is chosen as $u(x) = \varphi(x)/\beta(x)$, where

$$\beta(x_1, x_2) = \begin{cases} 
\beta^+(x_1, x_2) = 300(2 + \sin(6x_1 + 6x_2)) & \text{if } \varphi(x_1, x_2) > 0, \\
\beta^-(x_1, x_2) = 2 + \cos(6x_1 + 6x_2) & \text{if } \varphi(x_1, x_2) < 0. 
\end{cases}$$

It is easy to verify that the jump condition (1.2) and (1.3) is satisfied and $\nabla u \cdot t = 0$ on $\Gamma$. The exact solution and the interface are plotted in Figure 5.

To deal with variable coefficients, we choose $\beta^+_c = \beta^+(x_m)$, $\beta^-_c = \beta^-(x_m)$ on each interface element $T \in T_h^\Gamma$, where $x_m$ is the midpoint of $\Gamma_h \cap T$. Since $\nabla u \cdot t = 0$ on $\Gamma$, our theoretical analysis suggests the optimal convergence rates for both the nonconforming IFEM without penalties and our new nonconforming IFEM, which are confirmed by the results shown in Figure 6. We also test the parameter-free partially penalized immersed method (PPIFEM) using nodal values [24] and the nonconforming PPIFEM in [35] with $\eta_\epsilon = 10 \max(\beta^+(x_T), \beta^-(x_T))$, $x_T = e \cap \Gamma$ for each interface edge $e$ (see Rmk. 5.4). The numerical results in Figure 6 show that the convergence orders of all IFEMs are optimal.

6.3. An example with a straight interface and a piecewise linear solution

Inspired by Remark 4.3, we construct the following example with a straight interface and a piecewise linear solution.

Example 6.3. We choose

$$\varphi(x_1, x_2) = \frac{x_1 - x_2}{\sqrt{2}}, \quad u^\pm(x_1, x_2) = \frac{x_1 + x_2}{\sqrt{2}} + \frac{\varphi(x_1, x_2)}{\beta^\pm}.$$  

We set $\beta^+ = 2$ and $\beta^- = 1$. Note that the interface cut the boundary of the computational domain. Obviously, $\nabla u \cdot t \neq 0$ on $\Gamma$. Numerical results reported in Table 3 clearly show the suboptimal convergence of the nonconforming IFEM without penalties. However, the errors obtained by our new nonconforming IFEM are near machine precision.
6.4. An example with nonhomogeneous jump conditions and $\nabla u \cdot t \neq 0$ on $\Gamma$

Almost all the works in the literature only consider the examples satisfying the condition $\nabla u \cdot t = 0$ on $\Gamma$. The reason is that it is not easy to construct a function satisfying $\nabla u \cdot t \neq 0$ and the homogeneous conditions (1.2)–(1.3) on a curved interface simultaneously. For the problem with nonhomogeneous jump conditions, the relation $\nabla u \cdot t = 0$ can be easily violated. In the final example, we consider this case.

Example 6.4. We set

$$
\varphi(x_1, x_2) = x_1^2 + x_2^2 - 0.5^2,
\beta^+(x_1, x_2) = \sin(x_1 + x_2) + 2, \quad \beta^-(x_1, x_2) = \cos(x_1 + x_2) + 2,
$$

$$
u^+(x_1, x_2) = \ln(x_1^2 + x_2^2), \quad u^-(x_1, x_2) = \sin(x_1 + x_2).
$$

Clearly, $g_D := [u]_{\Gamma} \neq 0$ and $g_N := [\beta \nabla u \cdot n]_{\Gamma} \neq 0$. To deal with the nonhomogeneous jump conditions, we need a correction function $u_h^J$ defined as follows. We define $u_h^J | T = 0$ if $T \in \mathcal{T}_h^{\text{non}}$. On an interface element $T$, similarly to (3.2) and (3.3), we define $u_h^J | T^\pm = \phi^\pm$ with $\phi^\pm \in V_h(T)$ satisfying

$$
\phi^+(x) - \phi^-(x) = g_D(x) \quad \forall x \in \{D, E\},
$$

$$
\beta^+_c (\nabla \phi^+ \cdot n_h)(x_p) - \beta^-_c (\nabla \phi^- \cdot n_h)(x_p) = \frac{g_N(D) + g_N(E)}{2},
$$

$$
N_i(u_h^J | T) = 0 \quad \forall i \in \mathcal{I},
$$
Table 4. Numerical results of Example 6.4.

| $N$ | $\|e_h\|_{L^2}$ Rate | $|e_h|_{H^1}$ Rate | $\|e_h\|_{L^2}$ Rate | $|e_h|_{H^1}$ Rate |
|-----|----------------------|----------------------|----------------------|----------------------|
| 8   | 7.341E−02            | 1.194E+00            | 6.660E−02            | 1.052E+00            |
| 16  | 2.353E−02            | 1.64                 | 7.886E−01            | 0.60                 |
| 32  | 8.092E−03            | 1.54                 | 1.038E−03            | 0.97                 |
| 64  | 3.282E−03            | 1.30                 | 1.381E−01            | 0.99                 |
| 128 | 1.158E−03            | 1.50                 | 6.945E−02            | 0.99                 |
| 256 | 4.423E−04            | 1.39                 | 3.483E−02            | 1.00                 |
| 512 | 1.780E−04            | 1.31                 | 1.744E−02            | 1.00                 |
| 1024| 7.517E−05            | 1.24                 | 8.727E−03            | 1.00                 |

where $D$ and $E$ are endpoints of $\Gamma_h \cap T$. Now, for both methods (4.1) and (5.4), we replace the discrete trial space $V_{h,0}^{\text{IFE}}$ by $V_{h,g}^{\text{IFE}} + \{u'_h\}$, where $V_{h,g}^{\text{IFE}}$ is the nonconforming IFE space satisfying the nonhomogeneous Dirichlet boundary condition $g := u|_{\partial T}$ on boundary edges approximately. The numerical results reported in Table 4 show that the new nonconforming IFEM converges optimally in both $L^2$ and $H^1$ norms; however, the nonconforming IFEM without penalties only has the suboptimal convergence since $\nabla u \cdot t \neq 0$ on $\Gamma$.

We note that in (6.2) we use point values $g_N(D)$ and $g_N(E)$, which are not well defined if $g_N \in H^{1/2}(\Gamma)$. In this numerical example, $g_N(D)$ and $g_N(E)$ are meaningful since $g_N \in C(\Gamma)$. For the general case with $g_N \in H^{1/2}(\Gamma)$ and $g_D \in H^{3/2}(\Gamma)$, the point values should be replaced by averages in some sense. The corresponding IFE methods and theoretical analysis will be presented in a forthcoming paper.

7. Conclusions

In this paper, we have shown that the nonconforming IFE methods using the integral-value degrees of freedom on edges are not guaranteed to achieve optimal convergence rates without adding penalties although the continuity of IFE shape functions is weakly enforced through average values over edges. The suboptimal convergence rates have been confirmed by a counter numerical example where the tangential derivative of the exact solution is not zero on the interface. We think there is a similar issue for nonconforming IFE methods using integral-values on edges as degrees of freedom for solving elasticity and Stokes interface problems, which is an interesting topic in our future research.

To recover the optimal convergence rates, we have developed a new nonconforming IFE method with additional terms at interface edges. The new nonconforming IFE method is symmetric and the coercivity is ensured by a local lifting operator without a sufficiently large penalty parameter. We have also proved that IFE basis functions based on the Crouzeix–Raviart elements are unisolvent on arbitrary triangles which is one of advantages compared with the IFEs using nodal values as degrees of freedom. The optimal approximation capabilities of nonconforming IFE spaces based on the Crouzeix–Raviart and the rotated-$Q_1$ elements have been derived via a novel approach which can handle the case of variable coefficients easily. The optimal error estimates for the IFE solutions in the $H^1$ and $L^2$-norms have been derived and confirmed by some numerical examples.

Appendix A. Proof of optimal approximation capabilities of IFE spaces

Following the notations in the paper, on an interface element $T \in T_h^\Gamma$, we define IFE basis functions as

$$
\phi_i \in S_h(T), \quad N_j(\phi_i) = \delta_{ij} \quad \forall \ i, j \in I.
$$

(A.1)
Also define functions \( \phi_i^s \in V_h(T) \), \( s = +, - \), \( i \in \mathcal{I} \) such that \( \phi_i^s = \phi_i|_{T^s_i} \). Let \( \lambda_i \) be traditional basis functions, \( i.e. \)

\[
\lambda_i \in V_h(T), \quad N_j(\lambda_i) = \delta_{ij} \quad \forall \ i, j \in \mathcal{I}.
\]

(A.2)

Note that these functions depend on elements. We omit this dependence in our notation for simplicity. It is well-known that the traditional basis functions satisfy

\[
|\lambda_i|_{W^m(T)} \leq C h_T^{-m}, \quad i \in \mathcal{I}, \ m = 0, 1, 2,
\]

(A.3)

where the constant \( C \) only depends on the shape regularity parameter \( \rho \). The following lemma shows that the IFE basis functions also have similar estimates, which is one of essential ingredients for the success of IFE methods.

**Lemma A.1.** There exists a constant \( C \), depending only on \( \beta^+_c \), \( \beta^-_c \) and the shape regularity parameter \( \rho \), such that

\[
|\phi_i^+|_{W^m(T)} \leq C h_T^{-m}, \quad |\phi_i^-|_{W^m(T)} \leq C h_T^{-m}, \quad i \in \mathcal{I}, \ m = 0, 1, 2,
\]

(A.4)

\[
|\phi_i^+|_{W^m(T^+ \cup T^-)} \leq C h_T^{-m}, \quad i \in \mathcal{I}, \ m = 0, 1, 2.
\]

(A.5)

**Proof.** See Theorem 4.2 in [17] for the estimate (A.5) for the rotated-\( Q_1 \) element and the Crouzeix–Raviart element on right triangles. The estimate (A.4) can also be obtained easily from the proof of Theorem 5.6 in [15]. Next, we give a proof for general triangles without the constraint \( \alpha_{\max} = \pi/2 \). We just need to prove (A.4) since (A.5) is a direct consequence of (A.4). Using \( |N_j(\phi_i)| \leq 1, \|\nabla \lambda_i\|_{L^\infty(T)} \leq C h_T^{-1}, i, j \in \mathcal{I} \), we can estimate vectors in (3.10) as

\[
\|b\| \leq C, \quad |\gamma| \leq C h_T^{-1}, \quad |\delta| \leq C h_T,
\]

which, together with (3.11) and (3.12) lead to \( |c| \leq C \), where the constant \( C \) is independent of \( h_T \) and the interface location relative to the mesh. Thus, from (3.7), it follows

\[
|\phi_i^+|_{W^m(T)} \leq C h_T^{-m}, \quad i \in \mathcal{I}, \ m = 0, 1,
\]

where we have used the fact \( |\lambda_i|_{W^m(T)} \leq C h_T^{-m} \). It follows from (3.8) that \( |c_0| \leq C h_T^{-1} \) which together with (3.7) yields

\[
|\phi_i^-|_{W^m(T)} \leq C h_T^{-m}, \quad i \in \mathcal{I}, \ m = 0, 1.
\]

\[
\square
\]

Given two functions \( v^+ \in L^2(T) \) and \( v^- \in L^2(T) \), we define

\[
[v^\pm](x) := v^+(x) - v^-(x) \quad \forall \ x \in T.
\]

We next introduce auxiliary functions on each interface element \( T \in \mathcal{T}_h^\Gamma \). Recalling that \( D \) and \( E \) are intersection points of \( \Gamma \) and \( \partial T \), define auxiliary functions \( \Upsilon_1(x) \), \( \Upsilon_2(x) \) and \( \Psi(x) \) as

\[
\Upsilon_i(x) := \begin{cases} \Upsilon_i^+(x) \in V_h(T) & \text{if } x = (x_1, x_2) \in T^+_h, \\ \Upsilon_i^-(x) \in V_h(T) & \text{if } x = (x_1, x_2) \in T^-_h, \end{cases} \quad i = 1, 2,
\]

(A.6)

such that

\[
N_j(\Upsilon_i) = 0 \quad \forall \ j \in \mathcal{I}, \quad [\Upsilon_i^+](D) = 0, \quad i = 1, 2,
\]

\[
[\beta^+_c \nabla \Upsilon_1^+ \cdot n_h](x_p) = 1, \quad [\nabla \Upsilon_1^+ \cdot t_h](x_p) = 0, \quad [d(\Upsilon_1^+)] = 0, \quad i = 1, 2,
\]

(A.7)

\[
[\beta^-_c \nabla \Upsilon_2^+ \cdot n_h](x_p) = 0, \quad [\nabla \Upsilon_2^+ \cdot t_h](x_p) = 1, \quad [d(\Upsilon_2^+)] = 0,
\]
and
\[
\Psi(x) := \begin{cases} 
\Psi^+(x) \in V_h(T) & \text{if } x = (x_1, x_2) \in T^+_h, \\
\Psi^-(x) \in V_h(T) & \text{if } x = (x_1, x_2) \in T^-_h,
\end{cases}
\] (A.8)
such that
\[
N_j(\Psi) = 0 \ \forall \ j \in \mathcal{I}, \ \quad |\Psi^\pm|(D) = 1,
\]
\[
[\beta_c^\pm \nabla \Psi^\pm \cdot \mathbf{n}_h](x_p) = 0, \quad [\nabla \Psi^\pm \cdot t_h](x_p) = 0, \quad [d(\Psi^\pm)] = 0,
\] (A.9)
where \(d(\cdot)\) is defined in (3.5) and the point \(x_p \in \Gamma_h \cap T\) is the same as that in (3.3). For the rotated-\(Q_1\) element, we need another auxiliary function
\[
\Theta(x) := \begin{cases} 
\Theta^+(x) \in V_h(T) & \text{if } x = (x_1, x_2) \in T^+_h, \\
\Theta^-(x) \in V_h(T) & \text{if } x = (x_1, x_2) \in T^-_h,
\end{cases}
\] (A.10)
such that
\[
N_j(\Theta) = 0 \ \forall \ j \in \mathcal{I}, \ \quad |\Theta^\pm|(D) = 0,
\]
\[
[\beta_c^\pm \nabla \Theta^\pm \cdot \mathbf{n}_h](x_p) = 0, \quad [\nabla \Theta^\pm \cdot t_h](x_p) = 0, \quad [d(\Theta^\pm)] = 1.
\] (A.11)
In order to have a unified analysis, we also define \(\Theta = 0\) for the Crouzeix–Raviart element. Note that these auxiliary functions depend on the element \(T\). We omit the dependence for simplicity of notations.

**Lemma A.2.** On each interface element \(T \in \mathcal{T}_h^I\), these functions \(Y_1(x), Y_2(x), \Psi(x)\) and \(\Theta\) defined in (A.6)–(A.11) exist and satisfy
\[
|Y_i|^2_{H^m(T)} \leq Ch^{4-2m}_T, \quad m = 0, 1, 2, \ i = 1, 2, \ s = +, -,
\]
\[
|\Psi|^2_{H^m(T)} \leq Ch^{2-2m}_T, \quad |\Theta|^2_{H^m(T)} \leq Ch^{6-2m}_T, \quad m = 0, 1, 2, \ s = +, -,
\] (A.12)
where the constant \(C\) depends only on \(\beta^+_c, \beta^-_c\) and the shape regularity parameter \(q\).

**Proof.** We construct \(Y_i(x), i = 1, 2\) as follows,
\[
Y_i = z_i - I_{h,T}^{\text{IF}E} z_i, \quad z_i = \begin{cases} 
z^+_i & \text{in } T^+_h, \\
0 & \text{in } T^-_h,
\end{cases} \quad i = 1, 2,
\] (A.13)
where \(z^+_1\) and \(z^+_2\) are linear and satisfy
\[
z^+_i(D) = 0, \quad \beta^+_c \nabla z^+_i \cdot \mathbf{n}_h = 1, \quad \nabla z^+_i \cdot t_h = 0,
\]
\[
z^+_i(D) = 0, \quad \beta^-_c \nabla z^+_i \cdot \mathbf{n}_h = 0, \quad \nabla z^+_i \cdot t_h = 1.
\] (A.14)
It is easy to verify that \(z^+_1\) and \(z^+_2\) exist, and the constructed functions \(Y_i, i = 1, 2\) satisfy (A.6) and (A.7). From (A.14), we have
\[
||z^+_i||_{L^\infty(T)} \leq Ch_T, \ |\nabla z^+_i| \leq C, \ |z^+_i|_{W^2_2(T)} = 0, \ ||z_i||_{L^\infty(T)} \leq Ch_T, \quad i = 1, 2.
\] (A.15)
Since \(I_{h,T}^{\text{IF}E} z_i = \sum_j N_j(z_i) \phi_j\), it follows from (A.4) that, for \(m = 0, 1, 2, \ i = 1, 2, \ s = +, -\),
\[
\left|\left(I_{h,T}^{\text{IF}E} z_i\right)^s\right|_{W^m_\infty(T)} \leq \sum_{j \in \mathcal{I}} |N_j(z_i)| \ |\phi_j|^s_{W^m_\infty(T)} \leq Ch_T^{-m} \sum_{j \in \mathcal{I}} |N_j(z_i)|.
\]
Noticing
\[
|N_j(z_i)| \leq |e_j|^{-1} \int_{e_j} |z_i| \ ds \leq ||z_i||_{L^\infty(T)} \leq Ch_T,
\]
we have
\[ \left\| (I_{h,T}^{\text{IFE}} z)^s \right\|_{W^m_\infty(T)} \leq Ch^{1-m}_T, \quad m = 0, 1, 2, \ i = 1, 2, \ s = +, - , \]
which together with (A.13) and (A.15) implies
\[ |Y_i|_{W^m_\infty(T)} \leq Ch^{1-m}_T, \quad m = 0, 1, 2, \ i = 1, 2, \ s = +, - . \]
Finally, the first estimate in (A.12) is obtained by
\[ |\Psi^*|^2_{H^m(T)} \leq |\Psi^*|^2_{W^m_\infty(T)} \leq Ch^{4-2m}_T, \quad m = 0, 1, 2, \ i = 1, 2, \ s = +, - . \]

Other estimates in (A.12) can be proved similarly. We construct \( \Psi(x) \) as
\[ \Psi = z - I_{h,T}^{\text{IFE}} z, \quad z = \left\{ \begin{array}{ll} z^+ = 1 & \text{in } T^+_h, \\
0 & \text{in } T^-_h. \end{array} \right. \]
It is easy to verify that the constructed function \( \Psi \) satisfies (A.8) and (A.9). Since
\[ \|z^+\|_{L^\infty(T)} = 1, \quad |z^+|_{W^m_\infty(T)} = 0, \quad \|z\|_{L^\infty(T)} = 1, \quad m = 1, 2, \]
\[ |N_j(z)| \leq |e_j|^{-1} \int_{e_j} |z| \, ds \leq \|z\|_{L^\infty(T)} \leq 1, \quad j \in \mathcal{I}, \]
we get
\[ \left\| (I_{h,T}^{\text{IFE}} z)^s \right\|_{W^m_\infty(T)} \leq \sum_{j \in \mathcal{I}} |N_j(z)| \left| \phi_j^s \right|_{W^m_\infty(T)} \leq Ch^{-m}_T \sum_{j \in \mathcal{I}} |N_j(z)| \leq Ch^{-m}_T \]
and
\[ |\Psi^*|_{W^m_\infty(T)} \leq C h^{-m}_T, \quad s = +, - , \ m = 0, 1, 2, \]
which implies
\[ |\Psi^*|^2_{H^m(T)} \leq |\Psi^*|^2_{W^m_\infty(T)} \leq Ch^{2-2m}_T, \quad s = +, - , \ m = 0, 1, 2. \]

For the rotated-\( Q_1 \) element, we construct \( \Theta(x) \) as
\[ \Theta = z - I_{h,T}^{\text{IFE}} z, \quad z = \left\{ \begin{array}{ll} z^+ = (x_1 - m_1)^2 - k^2_T(x_2 - m_2)^2 & \text{in } T^+_h, \\
0 & \text{in } T^-_h, \end{array} \right. \]
where \( (m_1, m_2) \) is the center of the rectangle \( T \). Using (3.5), we easily verify that the constructed function \( \Theta \) satisfies (A.10) and (A.11). It follows that
\[ \|z^+\|_{L^\infty(T)} \leq Ch^2_T, \quad |z^+|_{W^2_2(T)} \leq Ch_T, \quad |z^+|_{W^2_\infty(T)} \leq C, \quad \|z\|_{L^\infty(T)} \leq Ch^2_T, \]
\[ |N_j(z)| \leq |e_j|^{-1} \int_{e_j} |z| \, ds \leq \|z\|_{L^\infty(T)} \leq Ch^2_T, \quad j \in \mathcal{I}, \]
which implies
\[ \left\| (I_{h,T}^{\text{IFE}} z)^s \right\|_{W^m_\infty(T)} \leq \sum_{j \in \mathcal{I}} |N_j(z)| \left| \phi_j^s \right|_{W^m_\infty(T)} \leq Ch^{-m}_T \sum_{j \in \mathcal{I}} |N_j(z)| \leq Ch^{2-m}_T \]
and
\[ |\Theta^*|^2_{H^m(T)} \leq |\Theta^*|^2_{W^m_\infty(T)} \leq |\Theta^*|^2_{W^m_2(T)} \leq C \left( |(z)^2_{W^2_\infty(T)} + |(I_{h,T}^{\text{IFE}} z)^2|_{W^2_\infty(T)} \right) |T| \]
\[ \leq Ch^{6-2m}_T, \quad s = +, - , \ m = 0, 1, 2. \]
For the Crouzeix–Raviart element, the above inequality (A.17) is also valid since we have defined \( \Theta = 0 \) if \( T \) is a triangle. \( \square \)
Lemma A.3. On each interface element $T \in T_h^I$, for any $v \in \tilde{H}^2(\Omega)$, the following identity holds:

$$
(I_h v_E^s)^{(s)}(x) - (I_h^{\text{IFE}} v)^{(s)}(x) = a\Psi^s(x) + \sum_{i=1,2} b_i \Phi_i(x) + \sum_{i \in I} g_i \phi_i^s + t\Theta^s(x) \quad \forall x \in T, \, s = +, -,
$$

(A.18)

with

$$
a = [I_h v_E^s](D), \quad b_i = [\beta_c^\pm \nabla(I_h v_E^s) \cdot \mathbf{n}_h](x_p), \quad b_2 = [\nabla(I_h v_E^s) \cdot \mathbf{t}_h](x_p),
$$

$$
g_i = |e_i|^{-1} \left( \int_{e_i^-} (I_h v_E^s - I_h v_E^-) \, ds + \int_{e_i^+} (I_h v_E^s - I_h v_E^+) \, ds \right), \quad i \in I,
$$

$$
t = [d(I_h v_E^s)],
$$

(A.19)

where $e_i = e_i \cap \Omega^s$, $s = +, -$, and $d(\cdot)$ is defined in (3.5). It is easy to see that $g_i = 0$ when $e_i^+ = \emptyset$ or $e_i^+ = e_i$.

Proof. For simplicity, define a function $w_h$ such that

$$
w_h |_{T^*_h} = w_h^s \quad \text{with} \quad w_h^s = I_h v_E^s - (I_h^{\text{IFE}} v)^s, \, s = +, -.
$$

It is obvious that $w_h^s \in V_h(T)$, $s = +, -$. Define another function

$$
v_h := [w_h^\pm](D)(\Psi(x) + [\beta_c^\pm \nabla w_h^\pm \cdot \mathbf{n}_h](x_p)\Phi_1(x)
$$

$$
+ [\nabla w_h^\pm \cdot \mathbf{t}_h](x_p)\Phi_2(x) + \sum_{i \in I} N_i(w_h)\phi_i + [d(w_h^\pm)]\Theta(x).
$$

(A.21)

Next, we prove $w_h = v_h$. From (A.1), (3.1)–(3.3) and Remark 3.1, the IFE basis functions $\phi_i$, $i \in I$ satisfy the following identities

$$
[\phi_i^\pm](D) = 0, \quad [\beta_c^\pm \nabla \phi_i^\pm \cdot \mathbf{n}_h](x_p) = 0, \quad [\nabla \phi_i^\pm \cdot \mathbf{t}_h](x_p) = 0,
$$

$$
[d(\phi_i^\pm)] = 0, \quad N_j(\phi_i) = \delta_{ij}, \quad i, j \in I.
$$

(A.22)

Combining (A.21), (A.22) and (A.6)–(A.11), we find

$$
[\beta_c^\pm \nabla w_h^\pm \cdot \mathbf{n}_h](x_p) = [\beta_c^\pm \nabla w_h^\pm \cdot \mathbf{n}_h](x_p), \quad N_i(v_h) = N_i(w_h), \quad i \in I
$$

(A.23)

and

$$
[w_h^\pm](D) = [w_h^\pm](D), \quad [\nabla w_h^\pm \cdot \mathbf{t}_h](x_p) = [\nabla w_h^\pm \cdot \mathbf{t}_h](x_p), \quad [d(w_h^\pm)] = [d(w_h^\pm)].
$$

(A.24)

It follows from (A.24) that $v_h - w_h$ is continuous across $\Gamma_h \cap T$, which together with (A.23) and (3.2)–(3.3), implies

$$
v_h - w_h \in S_h(T) \quad \text{and} \quad N_i(v_h - w_h) = 0 \quad \forall i \in I.
$$

Using the unisolvence of IFE shape functions (see Lemma 3.3), we know that the function $v_h - w_h$ is unique and $v_h - w_h = 0$ through a simple verification. Thus, from (A.21), we have

$$
w_h = v_h = a\Psi(x) + b_1 \Phi_1(x) + b_2 \Phi_2(x) + \sum_{i \in I} g_i \phi_i(x) + t\Theta(x),
$$

(A.25)

where

$$
a = [w_h^\pm](D), \quad b_1 = [\beta_c^\pm \nabla w_h^\pm \cdot \mathbf{n}_h](x_p), \quad b_2 = [\nabla w_h^\pm \cdot \mathbf{t}_h](x_p), \quad g_i = N_i(w_h), \quad t = [d(I_h v_E^\pm)].
$$

(A.20)

Using the following facts from (3.2) to (3.3)

$$
[(I_h^{\text{IFE}} v)^\pm](D) = 0, \quad [\beta_c^\pm \nabla(I_h^{\text{IFE}} v)^\pm \cdot \mathbf{n}_h](x_p) = 0, \quad [\nabla(I_h^{\text{IFE}} v)^\pm \cdot \mathbf{t}_h](x_p) = 0,
$$

(A.26)
we further have
\[ a = \|w_h^s\|_b(D) = \|I_h v_h^s - (I_h^{\text{IFEM}} v)_h\|_b(D) = \|I_h v_h^s\|_b(D) - \|(I_h^{\text{IFEM}} v)_h\|_b(D), \]
\[ b_1 = [\beta_c \nabla w_h^s \cdot n_h](x_p) = [\beta_c \nabla (I_h v_h^s - (I_h^{\text{IFEM}} v)_h) \cdot n_h](x_p) = [\beta_c \nabla (I_h v_h^s) \cdot n_h](x_p), \]
\[ b_2 = \|\nabla w_h^s \cdot t_h\|_b(x_p) = \|\nabla (I_h v_h^s - (I_h^{\text{IFEM}} v)_h) \cdot t_h\|_b(x_p) = \|\nabla (I_h v_h^s) \cdot t_h\|_b(x_p). \]

It remains to consider \( g_i, i \in I \). Define
\[ I_{h}^{BK} v := \begin{cases} I_h v_E^+ & \text{in } T_h^+, \\ I_h v_E^- & \text{in } T_h^- \end{cases} \]
then we know from (A.20) that \( w_h = I_h^{BK} v - I_h^{\text{IFEM}} v \). Using the fact that \( N_i(v - I_h^{\text{IFEM}} v) = 0 \) on interface elements from (3.13), we obtain
\[ g_i = N_i(w_h) = N_i(I_h^{BK} v - I_h^{\text{IFEM}} v) = N_i(I_h^{BK} v - v + I_h^{\text{IFEM}} v) = N_i(I_h^{BK} v - v) + N_i(v - I_h^{\text{IFEM}} v) = N_i(I_h^{BK} v - v) = |e_i|^{-1} \left( \int_{e_i^+} (I_h v_E^+ - v_E^-) \, ds + \int_{e_i^-} (I_h v_E^- - v_E^+) \, ds \right). \]

□

A.1. Proof of Theorem 3.5

Proof. On each interface element \( T \in \mathcal{T}_h^T \), by the triangle inequality, we have
\[ \left| v_E^s - (I_h^{\text{IFEM}} v)_h^s \right|_{H^m(T)} \leq \left| v_E^s - I_h v_E^s \right|_{H^m(T)} + \left| I_h v_E^s - (I_h^{\text{IFEM}} v)_h^s \right|_{H^m(T)}, \quad s = +, - . \tag{A.26} \]

The estimate of the first term is standard
\[ \left| v_E^s - I_h v_E^s \right|_{H^m(T)}^2 \leq C h_T^{2-2m} \left| v_E^s \right|_{H^2(T)}^2, \quad m = 0, 1, 2, \quad s = +, -. \tag{A.27} \]

For the second term on the right-hand side of (A.26), from Lemmas A.3, A.2 and (A.4), we have
\[ \left| I_h v_E^s - (I_h^{\text{IFEM}} v)_h^s \right|_{H^m(T)}^2 \leq \sum_{i=1,2} g_i^2 \left| \nabla^2 \Psi_i^s \right|_{H^m(T)} + \sum_{i \in I} g_i^2 \left| \nabla^2 \phi_i^s \right|_{H^m(T)} + \sum_{i \in I} g_i^2 \left| \nabla^2 \Theta_i^s \right|_{H^m(T)} \]
\[ \leq C h_T^{2-2m} \left( a^2 + \sum_{i \in I} g_i^2 \right) + C h_T^{4-2m} \sum_{i=1,2} b_i^2 + C h_T^{6-2m} t^2, \quad m = 0, 1, 2. \tag{A.28} \]

Here the constants \( t, a, b_1, b_2 \) and \( g_i, i \in I \) are defined in (A.19). We now estimate these constants one by one. Firstly, (3.5) implies
\[ t^2 = \left| \left[ d(I_h v_E^s) \right] \right| \leq \frac{1}{4k_T^4 + 4|T|} \left| I_h v_E^s \right|_{H^2(T)}^2 \leq C h_T^{-2} \sum_{s = +, -} \left| I_h v_E^s \right|_{H^2(T)}^2 \leq C h_T^{-2} \sum_{s = +, -} \left| v_E^s \right|_{H^2(T)}^2 \]
\[ \leq C h_T^{-2} \sum_{s = +, -} \left( \left| v_E^s \right|_{H^2(T)}^2 + \left| v_E^s - I_h v_E^s \right|_{H^2(T)}^2 \right) \leq C h_T^{-2} \sum_{s = +, -} \left| v_E^s \right|_{H^2(T)}^2. \tag{A.29} \]
where we used Sobolev’s inequality, scaling argument and the standard interpolation error estimate in the last inequality.

For the constant $b_1$, using the standard inverse inequality, we have

\[
b_1^2 = \| \beta_c \pm \nabla (I_h v_E) \cdot n_h \|^2_{L^2(\Omega)} \leq Ch_T^{-2} \| \nabla (I_h v_E) \cdot n_h \|^2_{L^2(\Omega)} \]

where

\[
b_1^2 \leq C \sum_{s=+,-} \| \nabla (I_h v_E) \cdot n_h - \nabla v_E \cdot n_h \|^2_{L^2(\Omega)} + \| \nabla v_E \cdot n_h \|^2_{L^2(\Omega)}.
\]

From (3.4), the second term can be estimate as

\[
h_T^{-2} \| (\beta_c^+ - \beta_c^-) \nabla (I_h v_E^+) \cdot n_h \|^2_{L^2(\Omega)} \leq \| \beta_c \pm \nabla (I_h v_E) \cdot n_h \|^2_{L^2(\Omega)} \leq C \sum_{s=+,-} \| \nabla (I_h v_E^+) \cdot n_h - \nabla v_E^+ \cdot n_h \|^2_{L^2(\Omega)} + \| \nabla v_E^+ \cdot n_h \|^2_{L^2(\Omega)}.
\]

For the first term on the right-hand side of (A.31), using (2.10), we get

\[
h_T^{-2} \| \nabla (I_h v_E^+) \cdot n_h \|^2_{L^2(\Omega)} = \| \nabla (I_h v_E^+) \cdot n_h + \beta_c^\pm \nabla v_E^+ \cdot (n_h - n + n) \|^2_{L^2(\Omega)} \]

\[
\leq Ch_T^{-2} \| \nabla (I_h v_E^+) \cdot n_h \|^2_{L^2(\Omega)}
\]

\[
+ \| n - n_h \|^2_{L^2(\Omega)} \| \beta_c^\pm \nabla v_E^+ \|^2_{L^2(\Omega)} + \| \beta_c^\pm \nabla v_E^+ \cdot n \|^2_{L^2(\Omega)} \]

\[
\leq C \sum_{s=+,-} \| v_E^+ \|^2_{H^1(\Omega)} + \| v_E^+ \|^2_{H^1(\Omega)} + Ch_T^{-2} \| \beta_c^\pm \nabla v_E^+ \cdot n \|^2_{L^2(\Omega)}.
\]

Combining above three inequalities yields

\[
b_2^2 \leq C \sum_{s=+,-} \| v_E^+ \|^2_{H^2(\Omega)} + \| v_E^+ \|^2_{H^1(\Omega)} + Ch_T^{-2} \| \beta_c^\pm \nabla v_E^+ \cdot n \|^2_{L^2(\Omega)}.
\]

Analogously,

\[
b_2^2 = \| \nabla (I_h v_E^+) \cdot t_h \|^2_{L^2(\Omega)} \leq \| \nabla (I_h v_E^+) \cdot t_h \|^2_{L^2(\Omega)} \leq Ch_T^{-2} \| \nabla (I_h v_E^+) \cdot t_h \|^2_{L^2(\Omega)} \]

\[
\leq Ch_T^{-2} \| \nabla (I_h v_E^+) - v_E^+ \cdot (t_h - t) \|^2_{L^2(\Omega)} \]

\[
\leq Ch_T^{-2} \left( \| \nabla (I_h v_E^+) - v_E^+ \cdot t_h \|^2_{L^2(\Omega)} + \| \nabla v_E^+ \cdot t \|^2_{L^2(\Omega)} \right) \]

\[
\leq C \sum_{s=+,-} \| v_E^+ \|^2_{H^2(\Omega)} + \| v_E^+ \|^2_{H^1(\Omega)} + Ch_T^{-2} \| \nabla v_E^+ \cdot t \|^2_{L^2(\Omega)}.
\]
Finally, the constants \( g_i, i \in \mathcal{I} \) in (A.19) can be estimated by the Cauchy–Schwarz inequality and the standard interpolation error estimate

\[
g_i^2 = |e_i|^2 \left[ \int_{e_i} (I_h^2 v_E^I - v_E^I) \, ds + \int_{e_i} (I_h^2 v_E - v_E) \, ds \right]^2 \leq Ch_i^2 \sum_{s=\pm,-} \|I_h^2 v_E^I - v_E^I\|_{L^2(e_i)}^2 \\
\leq C \sum_{s=\pm,-} \left( h_i^2 \|I_h^2 v_E^I - v_E^I\|_{L^2(T)}^2 + \|\nabla(I_h^2 v_E^I - v_E^I)\|_{L^2(T)}^2 \right) \leq Ch_i^2 \sum_{i=\pm,-} |v_e^I|^2_{H^2(T)}.
\]

We now combine (A.26)–(A.30), (A.32)–(A.34) to obtain the error estimate on interface elements

\[
h_T^{2(1-m)}|v_e^I - (I_h^I F E v)^I|^2_{H^m(T)} \leq Ch_T^2 \sum_{s=\pm} |v_e^I|^2_{H^2(T)} + C \left( \|\beta \pm \nabla v_E^I \cdot \mathbf{n}\|_{L^2(U(\Gamma, \mathcal{T}))}^2 + \|\nabla v_E^I \cdot \mathbf{t}\|_{L^2(T)}^2 \right).
\]

Summing up the estimate over all interface elements \( T \in T_h^\Gamma \) and using Assumption (2.3), we get

\[
\sum_{T \in T_h^\Gamma} h_T^{2(1-m)}|v_e^I - (I_h^I F E v)^I|^2_{W^2(T)} \leq Ch_T^2 \sum_{s=\pm} |v_e^I|^2_{H^2(T)} + C \left( \|\beta \pm \nabla v_E^I \cdot \mathbf{n}\|_{L^2(U(\Gamma, \mathcal{T}))}^2 + \|\nabla v_E^I \cdot \mathbf{t}\|_{L^2(T)}^2 \right).
\]

(A.35)

Since \( v \in \tilde{H}^2(\Omega) \), we know from the definition (2.1) that \( [\beta \pm \nabla v_E^I \cdot \mathbf{n}] (x) = 0 \) and \( [v_E^I] (x) = 0 \) for all \( x \in \Gamma \), which also implies \( [\nabla v_E^I \cdot \mathbf{t}] (x) = 0 \) on \( \Gamma \). Thus, by Lemma 2.4, and (2.5)–(2.6),

\[
\|\beta \pm \nabla v_E^I \cdot \mathbf{n}\|_{L^2(U(\Gamma, \mathcal{T}))} \leq Ch_T^2 \|\beta \pm \nabla v_E^I \cdot \mathbf{n}\|_{H^1(U(\Gamma, \mathcal{T}))} \leq Ch_T^2 \sum_{s=\pm} |v_e^I|^2_{H^2(\Omega)},
\]

\[
\|\nabla v_E^I \cdot \mathbf{t}\|_{L^2(T)} \leq Ch_T^2 \|\nabla v_E^I \cdot \mathbf{t}\|_{L^2(U(\Gamma, \mathcal{T}))} \leq Ch_T^2 \sum_{s=\pm} |v_e^I|^2_{H^2(\Omega)}.
\]

Substituting the above inequalities into (A.35) and using the extension result (3.15) we complete the proof. \( \square \)

**A.2. Proof of Theorem 3.6**

**Proof.** On each non-interface element \( T \in T_h^{\text{non}} \), the following estimate is standard

\[
|v - I_h^I F E v|^2_{H^m(T)} = |v - I_h v|^2_{H^m(T)} \leq Ch_T^{4-2m}|v|_{H^2(T)}^2, \quad m = 0, 1.
\]

(A.36)

On each interface element \( T \in T_h^\Gamma \), in view of the relations \( T = T^+ \cup T^- \) and \( T^* = (T^+ \cap T_h^+) \cup (T^- \cap T_h^-) \), \( s = +, - \), we have

\[
|v - I_h^I F E v|^2_{H^m(T)} = \sum_{s=+,-} |v^s - (I_h^I F E v)^s|^2_{H^m(T^\cap T_h^s)} + |v^+ - (I_h^I F E v)^+|^2_{H^m(T^\cap T_h^+)} + |v^- - (I_h^I F E v)^-|^2_{H^m(T^\cap T_h^-)}.
\]

(A.37)

By the triangle inequality,

\[
|v^s - (I_h^I F E v)^s|^2_{H^m(T^\cap T_h^s)} \leq 2|v^s - v_E^s|^2_{H^m(T^\cap T_h^s)} + 2|v_E^s - (I_h^I F E v)^s|^2_{H^m(T^\cap T_h^s)},
\]

\[
|v^+ - (I_h^I F E v)^+|^2_{H^m(T^\cap T_h^+)} \leq 2|v^+ - v_E^+|^2_{H^m(T^\cap T_h^+)} + 2|v_E^+ - (I_h^I F E v)^+|^2_{H^m(T^\cap T_h^+)},
\]

\[
|v^- - (I_h^I F E v)^-|^2_{H^m(T^\cap T_h^-)} \leq 2|v^- - v_E^-|^2_{H^m(T^\cap T_h^-)} + 2|v_E^- - (I_h^I F E v)^-|^2_{H^m(T^\cap T_h^-)}.
\]

(A.38)
Substituting (A.38) into (A.37) and using the definition (2.11), we get
\[ |v - I_h^{\text{FEM}}v|_{H^m(T)}^2 \leq C \sum_{s=+, -} |v^s - (I_h^{\text{FEM}}v)^s|_{H^m(T)}^2 + C\|v_E^+ - v_E^-\|_{H^m(T \setminus \Delta)}^2, \quad m = 0, 1. \] (A.39)

It follows from Lemma 2.5 and the fact \( \|v_E^+ \| = 0 \) on \( \Gamma \cap T \) that
\[
\|v_E^+ - v_E^-\|^2_{L^2(T \setminus \Delta)} \leq C h_T^4 \|v_E^+ \|^2_{H^1(T \setminus \Delta)} \leq C h_T^4 \sum_{s=+, -} |v^s|_{H^1(T)}^2,
\]
\[
\|\nabla(v_E^+ - v_E^-)\|^2_{L^2(T \setminus \Delta)} \leq C \left(h_T^2 \|\nabla v_E^+\|^2_{L^2(\Gamma \cap T)} + h_T^2 \|v_E^+\|^2_{H^2(T \setminus \Delta)}\right)
\leq C h_T^2 \sum_{s=+, -} \|\nabla v^s\|^2_{L^2(\Gamma \cap T)} + C h_T^4 \sum_{s=+, -} |v^s|_{H^2(T)}^2,
\]
which implies
\[ |v_E^+ - v_E^-|_{H^m(T \setminus \Delta)}^2 \leq C h_T^{4-2m} \sum_{s=+, -} \left(|v^s|_{H^1(T)}^2 + \|\nabla v^s\|^2_{L^2(\Gamma \cap T)}\right), \quad m = 0, 1. \] (A.40)

Combining (A.36), (A.39) and (A.40), we arrive at
\[
\sum_{T \in T_h} |v - I_h^{\text{FEM}}v|_{H^m(T)}^2 \leq C h_T^{4-2m} \sum_{T \in T_h} |v|_{H^2(T)}^2 + C \sum_{T \in T_h} \sum_{s=+, -} |v^s - (I_h^{\text{FEM}}v)^s|_{H^m(T)}^2
+ C h_T^{4-2m} \sum_{T \in T_h} \sum_{s=+, -} |v_E^s|_{H^1(T)}^2 + C h_T^{4-2m} \|\nabla v_E^+\|^2_{L^2(\Gamma)} m = 0, 1,
\]
which together with Theorem 3.5, the extension result (3.15) and the following global trace inequality
\[ \sum_{s=+, -} \|\nabla v_E^s\|^2_{L^2(\Gamma)} \leq C \left(\|v_E^+\|^2_{H^2(\Omega^+)} + \|v_E^-\|^2_{H^2(\Omega^-)}\right) \] (A.41)
yields the theorem.

**Acknowledgements.** The work of the first author was supported by the National Natural Science Foundation of China grants 11701291, 12101327, and the Natural Science Foundation of Jiangsu Province grant BK20200848. The work of the second and third authors was supported by the National Natural Science Foundation of China grants 12071227, 11871281, 11731007, the Natural Science Foundation of the Jiangsu Higher Education Institutions of China grant 20KJA110001, and the Ministry of Science and Technology of China Grant No. 2020YFA0713800. The work of the fourth author was supported by a Simons grant 633724.

**References**

[1] I. Babuška, The finite element method for elliptic equations with discontinuous coefficients. *Computing* 5 (1970) 207–213, 1970.

[2] J.W. Barrett and C.M. Elliott, A finite-element method for solving elliptic equations with Neumann data on a curved boundary using unfitted meshes. *IMA J. Numer. Anal.* 4 (1984) 309–325.

[3] J.W. Barrett and C.M. Elliott, Fitted and unfitted finite-element methods for elliptic equations with smooth interfaces. *IMA J. Numer. Anal.* 7 (1987) 283–300.

[4] J.H. Bramble and J.T. King, A robust finite element method for nonhomogeneous Dirichlet problems in domains with curved boundaries. *Math. Comput.* 63 (1994) 1–17.

[5] J.H. Bramble and J.T. King, A finite element method for interface problems in domains with smooth boundaries and interfaces. *Adv. Comput. Math.* 6 (1996) 109–138.
[6] S.C. Brenner and L.R. Scott, *The mathematical theory of finite element methods*, in Texts in Applied Mathematics. Vol. 15. Springer, Berlin (2008).
[7] E. Burman, J. Guzmán, M.A. Sánchez and M. Sarkis, Robust flux error estimation of an unfitted Nitsche method for high-contrast interface problems. *IMA J. Numer. Anal.* 38 (2018) 646–668.
[8] Z. Chen and J. Zou, Finite element methods and their convergence for elliptic and parabolic interface problems. *Numer. Math.* 79 (1998) 175–202.
[9] C.-C. Chu, I.G. Graham and T.-Y. Hou, A new multiscale finite element method for high-contrast elliptic interface problems. *Math. Comput.* 79 (2010) 1915–1955.
[10] M. Crouzeix and P.-A. Raviart, Conforming and nonconforming finite element methods for solving the stationary Stokes equations I. *ESAIM: Math. Model. Numer. Anal.* 7 (1973) 33–75.
[11] R.L. Foote, Regularity of the distance function. *Proc. Amer. Math. Soc.* 92 (1984) 153–155.
[12] S. Frei and T. Richter, A locally modified parametric finite element method for interface problems. *SIAM J. Numer. Anal.* 52 (2014) 2315–2334.
[13] D. Gilbarg and N.S. Trudinger, *Elliptic partial differential equations of second order*, in *Classics in Mathematics*. Springer-Verlag, Berlin (2001).
[14] 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 (2008) 472–495.
[15] R. Guo and T. Lin, A group of immersed finite-element spaces for elliptic interface problems. *IMA J. Numer. Anal.* 39 (2019) 482–511.
[16] R. Guo, T. Lin and Y. Lin, Error estimates for a partially penalized immersed finite element method for elasticity interface problems. *ESAIM: M2AN* 54 (2020) 1–24.
[17] R. Guo, T. Lin and X. Zhang, Nonconforming immersed finite element spaces for elliptic interface problems. *Comput. Math. Appl.* 75 (2018) 2002–2016.
[18] J. Guzmán, M.A. Sánchez and M. Sarkis, A finite element method for high-contrast interface problems with error estimates independent of contrast. *J. Sci. Comput.* 73 (2017) 330–365.
[19] A. Hansbo and P. Hansbo, An unfitted finite element method, based on Nitsche’s method, for elliptic interface problems. *Comput. Methods Appl. Mech. Eng.* 191 (2002) 5537–5552.
[20] X. He, T. Lin and Y. Lin, Approximation capability of a bilinear immersed finite element space. *Numer. Methods Partial Diff. Equ.* 24 (2008) 1265–1300.
[21] X. He, T. Lin, and Y. Lin, The convergence of the bilinear and linear immersed finite element solutions to interface problems. *Numer. Methods Partial Differ. Equ.* 28 (2012) 312–330.
[22] J. Huang and J. Zou, Some new a priori estimates for second-order elliptic and parabolic interface problems. *J. Differ. Equ.* 184 (2002) 570–586.
[23] J. Huang and J. Zou, Uniform a priori estimates for elliptic and static Maxwell interface problems. *Discrete and Continuous Dyn. Syst. Ser. B* 7 (2007) 145–170.
[24] H. Ji, F. Wang, J. Chen and Z. Li, A new parameter free partially penalized immersed finite element and the optimal convergence analysis. *Numer. Math.* 150 (2022) 1035–1086.
[25] D.Y. Kwak, K.T. Wee and K.S. Chang, An analysis of a broken $P_1$-nonconforming finite element method for interface problems. *SIAM J. Numer. Anal.* 48 (2010) 2117–2134.
[26] Z. Li, The immersed interface method using a finite element formulation. *Appl. Numer. Math.* 27 (1998) 253–267.
[27] Z. Li, T. Lin and X. Wu, New Cartesian grid methods for interface problems using the finite element formulation. *Numer. Math.* 96 (2003) 61–98.
[28] T. Lin, Y. Lin and X. Zhang, Partially penalized immersed finite element method for elliptic interface problems. *SIAM J. Numer. Anal.* 53 (2015) 1121–1144.
[29] T. Lin, D. Sheen and X. Zhang, A nonconforming immersed finite element method for elliptic interface problems. *J. Sci. Comput.* 79 (2019) 442–463.
[30] Z. Li, T. Lin, Y. Lin and R. Rogers, An immersed finite element space and its approximation capability. *Numer. Methods Partial Differ. Equ.* 20 (2004) 338–367.
[31] J. Li, J. Markus, B. Wohlmuth and J. Zou, Optimal a priori estimates for higher order finite elements for elliptic interface problems. *Appl. Numer. Math.* 60 (2010) 19–37.
[32] R. Rannacher and S. Turek, Simple nonconforming quadrilateral stokes element. *Numer. Methods Partial Differ. Equ.* 8 (1992) 97–111.
[33] H. Wu and Y. Xiao, An unfitted hp-interface penalty finite element method for elliptic interface problems. *J. Comput. Math.* 37 (2019) 316–339.
[34] J. Xu, Error estimates of the finite element method for the 2nd order elliptic equations with discontinuous coefficients. *J. Xiangtan University* 1 (1982) 1–5.
[35] X. Zhang, *Nonconforming immersed finite element methods for interface problems*. Ph.D. Thesis, Virginia Polytechnic Institute and State University (2013).