THE LEAST SQUARES FINITE ELEMENT METHOD FOR ELASTICITY INTERFACE PROBLEM ON UNFITTED MESH

FANYI YANG*

Abstract. In this paper, we propose and analyze the least squares finite element methods for the linear elasticity interface problem in the stress-displacement system on unfitted meshes. We consider the cases that the interface is C^2 or polygonal, and the exact solution $(\boldsymbol{\sigma}, \boldsymbol{u})$ belongs to $H^s(\text{div}; \Omega_0 \cup \Omega_1) \times H^{1+s}(\Omega_0 \cup \Omega_1)$ with $s > 1/2$. Two types of least squares functionals are defined to seek the numerical solutions. The first is defined by simply applying the L^2 norm least squares principle, and requires the condition $s \geq 1$. The second is defined with a discrete minus norm, which is related to the inner product in $H^{-1/2}(\Gamma)$. The use of this discrete minus norm results in a method of optimal convergence rates and allows the exact solution has the regularity of any $s > 1/2$. The stability near the interface for both methods is guaranteed by the ghost penalty bilinear forms and we can derive the robust condition number estimates. The convergence rates under L^2 norm and the energy norm are derived for both methods. We illustrate the accuracy and the robustness of the proposed methods by a series of numerical experiments for test problems in two and three dimensions.

Mathematics Subject Classification. 65N30.

1. INTRODUCTION

In this paper, we develop the least squares finite element methods (LSFEMs) for linear elasticity interface problems, which model the elasticity structure with different or even singular material properties, and have many applications in fields of materials science and continuum mechanics [2, 4, 25, 26, 39]. For such problems, the governing equations usually have discontinuous coefficients and involve the inhomogeneous jump conditions. Because of the discontinuity near the interface and the irregular geometry of the interface, it is still challenging to design efficient numerical methods for such equations.

The finite element method is an important numerical method for solving interface problems. In the last decades, various numerical schemes have been developed for the elliptic interface problem, and we refer to [6, 10, 11, 16, 33, 36, 40, 53] for some typical methods. The finite element methods can be roughly classified into fitted and unfitted methods based on types of grids. The body-fitted method requires the mesh to be aligned with the interface for representing the geometry of the interface accurately. For complex geometries, it is a challenging and time-consuming task to generate a high quality body-fitted mesh especially in high dimensions

College of Mathematics, Sichuan University, Chengdu, P.R. China.
*Corresponding author: yangfanyi@scu.edu.cn

[36]. In the unfitted method, the interface description is decoupled from the generation of the mesh, which provides a good flexibility when handling the problem with complex geometries. Examples of such methods are the cut finite element method [6, 10, 11, 30, 33, 58], the immersed finite element method [28, 40, 41] and the aggregated finite element method [3, 17, 37].

Recently, the unfitted finite element methods are also been applied to solve the linear elasticity interface problem. In [34], Hansbo and Hansbo proposed a linear finite element method and derived the optimal convergence rates in error measurements under the assumption that the exact solution $\boldsymbol{u}$ is piecewise H^2. In [4], Becker *et al.* developed a mixed finite element method with the linear accuracy for the displacement-pressure formulation. The inf-sup condition and the optimal error estimates are verified. Both methods are also called Nitsche extended finite element methods (Nitsche-XFEM), where the jump conditions are weakly imposed by the Nitsche penalty method [33]. In [58], Zhang followed the interface-penalty idea and presented a high-order unfitted method for the elasticity interface problem. Combining with the penalty method and the hybridizable discontinuous Galerkin approximation, Han *et al.* developed an X-HDG method for this problem [32]. Another type of unfitted finite element methods is the immersed finite element method [40], which mainly modifies the basis functions near the interface to capture the jump of the solution. In [45, 46], the authors presented the non-conforming immersed finite element methods for the linear elasticity interface problem. Other types of immersed finite element methods can be found in [28, 38]. We note that all above mentioned methods derive the error estimates under the assumption that the exact solution $\boldsymbol{u}$ has at least piecewise H^2 regularity. Many analysis techniques that are developed for piecewise H^2 solutions in the elliptic interface problem can be used in this case. But these techniques may be unavailable for the solution that has only piecewise $H^{1+s}(s < 1)$ regularity. In addition, for piecewise $H^1(\text{div})$ or piecewise $H^1(\text{curl})$ functions, applying these techniques to estimate the errors will result in a suboptimal convergence rate, see [43, 48] for unfitted methods in solving $H(\text{curl})$- and $H(\text{div})$-interface problems. To our best knowledge, there are few works on the interface problem of low regularity. In [29], the author presented an immersed finite element method for $H(\text{curl})$-interface problem with the optimal convergence rates. Only piecewise $H^1(\text{curl})$ regularity of the exact solution is required in the analysis.

In this paper, we develop least squares finite element methods on unfitted meshes for the linear elasticity interface problem, based on the stress-displacement formulation. For traditional linear elasticity problems, the LSFEMs have been investigated in [8, 12–14, 42, 52]. The LSFEM can offer the advantage of circumventing the inf-sup condition arising in mixed methods and ensure the resulting linear system is always symmetric positive definite. As the standard LSFEM, we define a least squares functional and seek the numerical solution by minimizing the functional over finite element spaces. Two types of least squares functionals are used in this paper. The first is defined by simply applying the L^2 norm least squares principle to the stress-displacement. The defined functional only involves the L^2 norms and the jump conditions are also enforced in the sense of L^2 norms. This method requires the exact solution $(\boldsymbol{\sigma}, \boldsymbol{u})$ has the regularity $H^s(\text{div}; \Omega_0 \cup \Omega_1) \times H^{1+s}(\Omega_0 \cup \Omega_1)$ with $s \geq 1$. The convergence rate in the $H(\text{div}) \times H^1$ norm is shown to be half order lower than the optimal rate. From the embedding theory, we know that any $\boldsymbol{\tau} \in H(\text{div}; \Omega_0 \cup \Omega_1)$ has the normal trace $\boldsymbol{n} \cdot \boldsymbol{\tau} \in H^{-1/2}(\Gamma)$, but the stronger L^2 norm is applied to handle the normal trace in this method. This is the reason that the condition $s \geq 1$ is required and the convergence rate is not optimal. To overcome this difficulty, we define another least squares functional with a discrete minus inner product. This method follows the ideas in [7,8], where the discrete minus norms corresponding to $H^{-1}(\Omega)$ are used. In this paper, we define a discrete minus inner product that is related to the inner product in the space $H^{-1/2}(\Gamma)$. The use of this inner product allows us to relax the regularity condition as $s > 1/2$, and gives the optimal convergence rate under the error measurement with respect to the required regularity. It is noticeable that the optimal convergence rate under the $H(\text{div})$ norm is achieved for functions in $H^s(\text{div}; \Omega_0 \cup \Omega_1)$. We also point out that in the unfitted methods, the H^1 trace estimate is usually the main tool to estimate the numerical error on the interface. For the low regularity case $s < 1$, this estimate is unavailable, and we use the embedding theory in the error estimation instead.

Another important issue for unfitted methods is the presence of small cuts near the interface. In our method, we employ the ghost penalty method [9] to cure the effects bringing by small cuts. The ghost penalty bilinear forms also correspond to L^2 norms, and they can be naturally added in the least squares functional. We can

prove a uniform upper bound of the condition number to the resulting linear system with proper penalty forms. We give a suitable penalty bilinear form with the polynomial local extension, and also the standard penalty forms given in [9, 30] can be used in our methods.

The rest of this article is organized as follows. In Section 2, we give the basic notation and introduce the stress-displacement formulation for the linear elasticity interface problem. The ghost penalty forms are also presented in this section. In Section 3, we define the associated least squares functional of the interface problem. Section 4 develops the numerical schemes. The least squares finite element methods with L^2 norms and with the discrete minus norm are established in Sections 4.1 and 4.2, respectively. The error estimations are also included. Finally, numerical results of test problems in two and three dimensions are presented in Section 5.

2. PROBLEM SETTING AND PRELIMINARIES

Let $\Omega \subset \mathbb{R}^d (d = 2, 3)$ be a convex polygonal (polyhedral) domain with the boundary $\partial\Omega$. Let $\Omega_0 \Subset \Omega$ be a polygonal (polyhedral) subdomain or a subdomain with a C^2-smooth boundary. We denote by $\Gamma := \partial\Omega_0$ the topological boundary, which can be regarded as an interface dividing Ω into two disjoint domains Ω_0 and Ω_1, where $\Omega_1 := \Omega\backslash\overline{\Omega}_0$, $\Omega_0 \cap \Omega_1 = \varnothing$ and $\overline{\Omega}_0 \cup \overline{\Omega}_1 = \overline{\Omega}$. The model problem considered in this paper is the linear elasticity interface problem defined on Ω, which reads: seek the stress $\boldsymbol{\sigma} = (\sigma_{ij})_{d\times d}$ and the displacement $\boldsymbol{u} = (u_j)_d$ such that

$$
\begin{aligned}
\mathcal{A}\boldsymbol{\sigma} - \boldsymbol{\varepsilon}(\boldsymbol{u}) &= \mathbf{0}, && \text{in } \Omega_0 \cup \Omega_1, \\
\nabla \cdot \boldsymbol{\sigma} + \boldsymbol{f} &= \mathbf{0}, && \text{in } \Omega_0 \cup \Omega_1, \\
\boldsymbol{u} &= \mathbf{0}, && \text{on } \partial\Omega, \\
[\![\boldsymbol{\sigma}]\!]_N = \boldsymbol{a}, \quad [\![\boldsymbol{u}]\!] &= \boldsymbol{b}, && \text{on } \Gamma,
\end{aligned}
\tag{1}
$$

where $\boldsymbol{f}$ is the source term, and $\boldsymbol{a}$, $\boldsymbol{b}$ are the jump conditions on the interface. The jump operators are defined as (6). The Lamé parameters λ, μ are assumed to be piecewise positive constant functions,

$$
(\lambda(\boldsymbol{x}), \mu(\boldsymbol{x})) := \begin{cases} (\lambda_0, \mu_0), & \text{in } \Omega_0, \\ (\lambda_1, \mu_1), & \text{in } \Omega_1, \end{cases} \quad \lambda_0, \lambda_1, \mu_0, \mu_1 > 0.
\tag{2}
$$

The constitutive law is expressed by the linear operator $\mathcal{A} : \mathbb{R}^{d\times d} \to \mathbb{R}^{d\times d}$:

$$
\mathcal{A}\boldsymbol{\tau} := \frac{1}{2\mu}\left(\boldsymbol{\tau} - \frac{\lambda}{d\lambda + 2\mu}\text{tr}(\boldsymbol{\tau})\mathbf{I}\right), \qquad \forall \boldsymbol{\tau} \in \mathbb{R}^{d\times d},
$$

where $\text{tr}(\cdot)$ denotes the trace operator, and $\mathbf{I} := (\delta_{ij})_{d\times d}$ is the identity tensor. The function $\boldsymbol{\varepsilon}(\boldsymbol{u})$ denotes the symmetric strain tensor:

$$
\boldsymbol{\varepsilon}(\boldsymbol{u}) = (\varepsilon_{i,j}(\boldsymbol{u}))_{d\times d}, \quad \varepsilon_{i,j}(\boldsymbol{u}) := \frac{1}{2}\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right), \qquad 1 \le i, j \le d.
$$

We assume that the interface problem (1) admits a unique solution $(\boldsymbol{\sigma}, \boldsymbol{u}) \in \boldsymbol{\Sigma}^s \times \mathbf{V}^{s+1}$ with $s > 1/2$, where

$$
\boldsymbol{\Sigma}^s := \mathbf{H}^s(\text{div}; \Omega_0 \cup \Omega_1), \quad \mathbf{V}^{s+1} := \left\{\boldsymbol{v} \in \mathbf{H}^{s+1}(\Omega_0 \cup \Omega_1) : \boldsymbol{v}|_{\partial\Omega} = \mathbf{0}\right\}.
\tag{3}
$$

For $i = 0, 1$, we let $\boldsymbol{\sigma}_i := \boldsymbol{\sigma}|_{\Omega_i}$, $\boldsymbol{u}_i := \boldsymbol{u}|_{\Omega_i}$. We further assume that $\boldsymbol{\sigma}_i$ and $\boldsymbol{u}_i$ can be extended to the whole domain Ω in the sense that there exist $\widetilde{\boldsymbol{\sigma}}_i \in \mathbf{H}^s(\text{div}; \Omega)$ and $\widetilde{\boldsymbol{u}}_i \in \mathbf{H}^{s+1}(\Omega)$ such that $\widetilde{\boldsymbol{\sigma}}_i|_{\Omega_i} = \boldsymbol{\sigma}_i$ with $\|\widetilde{\boldsymbol{\sigma}}_i\|_{H^s(\text{div};\Omega)} \le C\|\boldsymbol{\sigma}_i\|_{H^s(\text{div};\Omega_i)}$ and $\widetilde{\boldsymbol{u}}_i|_{\Omega_i} = \boldsymbol{u}_i$ with $\|\widetilde{\boldsymbol{u}}_i\|_{H^{s+1}(\Omega)} \le C\|\boldsymbol{u}_i\|_{H^{s+1}(\Omega_i)}$. Consequently, $\boldsymbol{\sigma}$ and $\boldsymbol{u}$ can be decomposed as

$$
\boldsymbol{\sigma} = \widetilde{\boldsymbol{\sigma}}_0 \cdot \chi_0 + \widetilde{\boldsymbol{\sigma}}_1 \cdot \chi_1, \quad \boldsymbol{u} = \widetilde{\boldsymbol{u}}_0 \cdot \chi_0 + \widetilde{\boldsymbol{u}}_1 \cdot \chi_1,
\tag{4}
$$

where χ_i is the characteristic function corresponding to the domain Ω_i. We refer to [1, 20, 35] for more details about the extension of Sobolev spaces. From (4), we formally introduce two projection operators $\pi_i(i = 0, 1)$ that $\pi_i\boldsymbol{\sigma} := \widetilde{\boldsymbol{\sigma}}_i \in \mathbf{H}^s(\text{div}; \Omega)$ and $\pi_i\boldsymbol{u} := \widetilde{\boldsymbol{u}}_i \in \mathbf{H}^{s+1}(\Omega)$.

Let us introduce the notation required in the definition to the scheme. We denote by $\mathcal{T}_h$ a quasi-uniform partition of Ω into triangle (tetrahedron) elements. The mesh $\mathcal{T}_h$ is unfitted that the element faces in $\mathcal{T}_h$ are not required to be aligned with the interface Γ. For any element $K \in \mathcal{T}_h$, we denote by h_K its diameter and by ρ_K the radius of the largest disk (ball) inscribed in K. Let $h := \max_{K \in \mathcal{T}_h} h_K$ be the mesh size, and let $\rho := \min_{K \in \mathcal{T}_h} \rho_K$. The mesh $\mathcal{T}_h$ is quasi-uniform in the sense that there exists a constant C_ν independent of h such that $h \leq C_\nu\rho$.

Next, we give the notation related to subdomains Ω_0 and Ω_1. For $i = 0, 1$, we define $\mathcal{T}_{h,i} := \{K \in \mathcal{T}_h \mid K \cap \Omega_i \neq \varnothing\}$ as the minimal subset of $\mathcal{T}_h$ that entirely covers Ω_i, and define $\mathcal{T}_{h,i}^\circ := \{K \in \mathcal{T}_{h,i} \mid K \subset \Omega_i\}$ as the set of interior elements inside the domain Ω_i. We define $\mathcal{T}_h^\Gamma := \{K \in \mathcal{T}_h \mid K \cap \Gamma \neq \varnothing\}$ as the collection of all cut elements. Their corresponding domains are defined as $\Omega_{h,i} := \text{Int}(\bigcup_{K \in \mathcal{T}_{h,i}} \overline{K})$, $\Omega_{h,i}^\circ := \text{Int}(\bigcup_{K \in \mathcal{T}_{h,i}^\circ} \overline{K})$, and $\Omega_h^\Gamma := \text{Int}(\bigcup_{K \in \mathcal{T}_h^\Gamma} \overline{K})$, respectively. Notice that there holds $\Omega_h^\Gamma = \Omega_{h,i} \backslash \overline{\Omega_{h,i}^\circ}$. For any element $K \in \mathcal{T}_h$, we let $K^i := K \cap \Omega_i$ and for any cut element $K \in \mathcal{T}_h^\Gamma$, we let $\Gamma_K := K \cap \Gamma$.

For any $K \in \mathcal{T}_h$, we define $\Delta(K) := \{K' \in \mathcal{T}_h \mid \overline{K'} \cap \overline{K} \neq \varnothing\}$ as the set of elements touching K. We denote by $B(\boldsymbol{z}, r)$ the disk (ball) centered at the point $\boldsymbol{z}$ with the radius r. Since $\mathcal{T}_h$ is quasi-uniform, there exists a generic constant C_Δ such that $\bigcup_{K' \in \Delta(K)} \overline{K'} \subset B(\boldsymbol{x}_K, C_\Delta h)$ for $\forall K \in \mathcal{T}_h$, where $\boldsymbol{x}_K$ is the barycenter of K. We set $B_{\Delta(K)} := B(\boldsymbol{x}_K, C_\Delta h)$ for $\forall K \in \mathcal{T}_h$.

Throughout this paper, C and C with subscripts are denoted to be generic positive constants that may vary in different lines, but are always independent of the mesh size h, how the interface Γ cuts the mesh $\mathcal{T}_h$, and the Lamé parameter λ defined in (2).

We make the following geometrical assumption on the mesh:

Assumption 1. *For any cut element $K \in \mathcal{T}_h^\Gamma$, the sets $\Delta(K) \cap \mathcal{T}_{h,0}^\circ$ and $\Delta(K) \cap \mathcal{T}_{h,1}^\circ$ are not empty.*

From Assumption 1, we can assign two interior elements $K_0^{\text{int}} \in \Delta(K) \cap \mathcal{T}_{h,0}^\circ$ and $K_1^{\text{int}} \in \Delta(K) \cap \mathcal{T}_{h,1}^\circ$ for any cut element $K \in \mathcal{T}_h^\Gamma$. In principle, $K_i^{\text{int}}(i = 0, 1)$ can be anyone in $\Delta(K) \cap \mathcal{T}_{h,i}^\circ$. In practice, one can select K_i^{int} to share a common face with K whenever possible. Consequently, Assumption 1 allows us to define two maps $M^i(\cdot)(i = 0, 1) : \mathcal{T}_h^\Gamma \to \mathcal{T}_{h,i}^\circ$ such that $M^i(K) = K_i^{\text{int}}$ for $\forall K \in \mathcal{T}_h^\Gamma$.

Remark 1. Assumption 1 can be relaxed as: for any cut element $K \in \mathcal{T}_h^\Gamma$, there exists a wide patch $S(K)$ of $\text{diam}(S(K)) \leq Ch_K$ such that $K \in S(K)$ and $S(K)$ contains two elements K_0^* and K_1^* satisfying

$$|K_i^* \cap \Omega_i| \geq C_p|K_i^*|, \quad i = 0, 1, \tag{5}$$

where $C_p \in (0, 1)$ is a user-defined constant. Then, we can let $M^i(K)$ be anyone in $S(K)$ satisfying (5) for $i = 0, 1$. Clearly, Assumption 1 coincides with the case that $S(K) = \Delta(K)$ and $C_p = 1$. By (5), there holds $\|v\|_{L^2(M^i(K))} \leq C\|v\|_{L^2(M^i(K) \cap \Omega_i)}$, and the properties **P1** and **P2** for the forms (12) with such $M^i(K)$ can be verified similarly. If we take $S(K) = \Delta(K)$ and select $C_p < 1$, then $M^i(K)$ can be a cut element, which has a large intersection to Ω_i. This idea is similar to the idea of merging small elements with neighbouring large elements to ensure the stability near the interface, see [17, 30, 37] for more details.

Let us introduce the notation of trace operators on the interface. Let $\boldsymbol{v}$ be a vector- or tensor-valued function, we define the jump operators $[\![\cdot]\!]$ and $[\![\cdot]\!]_N$ as

$$[\![\boldsymbol{v}]\!]|_\Gamma := \boldsymbol{v}^0|_\Gamma - \boldsymbol{v}^1|_\Gamma, \quad [\![\boldsymbol{v}]\!]_N|_\Gamma := \boldsymbol{n}_\Gamma \cdot (\boldsymbol{v}^0|_\Gamma - \boldsymbol{v}^1|_\Gamma), \tag{6}$$

where $\boldsymbol{v}^0 := \boldsymbol{v}|_{\Omega_0}, \boldsymbol{v}^1 := \boldsymbol{v}|_{\Omega_1}$, and $\boldsymbol{n}_\Gamma$ denotes the unit outward normal vector pointing to Ω_1 on Γ.

For an open bounded domain D, we let $H^r(D)$ denote the usual Sobolev spaces with the exponent $r \geq 0$, and we follow their corresponding inner products, seminorms and norms. We define $\mathbf{H}^r(D) := (H^r(D))^d$ and

$\mathbb{H}^r(D) := (H^r(D))^{d \times d}$ as the Sobolev spaces of vector and tensor fields, respectively. Let $L^2(D)$ coincide with $H^r(D)$ for $r = 0$. Further, we introduce $H^r(\text{div}; D) := \{\boldsymbol{v} \in \mathbf{H}^r(D) \mid \nabla \cdot \boldsymbol{v} \in H^r(D)\}$ with the norm $\|\boldsymbol{v}\|^2_{H^r(\text{div};D)} := \|\boldsymbol{v}\|^2_{H^r(D)} + \|\nabla \cdot \boldsymbol{v}\|^2_{H^r(D)}$, and let $\mathbf{H}^r(\text{div}; D) := (H^r(\text{div}; D))^d$ be the spaces of tensor fields. Each column of functions in $\mathbf{H}^r(\text{div}; D)$ belongs to $H^r(\text{div}; D)$. Let $H^{-1/2}(\partial D)$ be the dual space of $H^{1/2}(\partial D)$ with the norm

$$\|v\|_{H^{-1/2}(\partial D)} := \sup_{0 \neq \varphi \in H^{1/2}(\partial D)} \frac{(v, \varphi)_{L^2(\partial D)}}{\|\varphi\|_{H^{1/2}(\partial D)}}, \qquad \forall v \in H^{-1/2}(\partial D). \tag{7}$$

From the trace theory, we know that any function $\boldsymbol{v} \in H(\text{div}; D)$ has a normal trace $\boldsymbol{n} \cdot \boldsymbol{v} \in H^{-1/2}(\partial D)$ with $\|\boldsymbol{n} \cdot \boldsymbol{v}\|_{H^{-1/2}(\partial D)} \leq C\|\boldsymbol{v}\|_{H(\text{div};D)}$. For vector fields, we let $\mathbf{H}^{-1/2}(\partial D)$ be the dual space of $\mathbf{H}^{1/2}(\partial D)$, and the corresponding norm is analogous to (7).

To cure the effect bringing by small cuts near the interface, we follow the idea in the ghost penalty method [9, 30], which uses the data from the interior domain to ensure the stability near the interface. For this goal, we assume that we can construct two bilinear forms $s^r_{h,0} : L^2(\Omega_{h,0}) \times L^2(\Omega_{h,0}) \mapsto \mathbb{R}$ and $s^r_{h,1} : L^2(\Omega_{h,1}) \times L^2(\Omega_{h,1}) \mapsto \mathbb{R}$, with the induced seminorms $|v|^2_{s^r_{h,i}} := s^r_{h,i}(v,v)(i = 0, 1)$ for $\forall v \in L^2(\Omega_{h,i})$. In our method, we assume that the forms satisfy the following two properties:

P1: the L^2 norm extension property:

$$\|v\|_{L^2(\Omega_{h,i})} \leq C\Big(\|v\|_{L^2(\Omega_i)} + |v|_{s^r_{h,i}}\Big) \leq C\|v\|_{L^2(\Omega_{h,i})}, \qquad \forall v \in L^2(\Omega_{h,i}), \quad i = 0, 1. \tag{8}$$

P2: the weak consistency:

$$|v|_{s^r_{h,i}} \leq Ch^t\|v\|_{H^{s+1}(\Omega)}, \qquad \forall v \in H^{s+1}(\Omega), \quad i = 0, 1, \quad t = \min(s+1, r+1). \tag{9}$$

The suitable penalty forms $s^r_{h,i}(\cdot, \cdot)$ can be constructed by the face-based penalties and the projection-based penalties, see [10, 30]. In Section 2.7 of [30], the penalty forms constructed for elliptic interface problems satisfy **P1** and **P2**, which can be used in our method. We also note that extra properties of the penalty form in solving elliptic interface problems are required, as the H^1 seminorm extension property and the inverse estimate, see EP1–EP4 of [30]. In our method, **P1** and **P2** are enough to ensure the stability near the interface. Thus, the required ghost penalty may be more simple.

Here, we outline a method to construct the penalty bilinear forms by the local polynomial extension, and the implementation is easy. The idea of the local extension has also been widely used in unfitted methods, see [3, 11, 17, 36, 37, 57].

For any element $K \in \mathcal{T}_h$, we define a local extension operator $E^r_K(r \geq 0)$ that extends the function in $L^2(K)$ to the ball $B_{\Delta(K)}$ by

$$\begin{aligned} E^r_K : L^2(K) &\to \mathbb{P}_r\big(B_{\Delta(K)}\big), \\ v &\to E^r_K v, \end{aligned} \quad E^r_K v := (\mathcal{E}_K(\Pi^r_K v))|_{B_{\Delta(K)}}, \tag{10}$$

where $\mathcal{E}_K : \mathbb{P}_r(K) \to \mathbb{P}_r(\mathbb{R}^d)$ is the canonical extension of a polynomial to $\mathbb{R}^d$, and $\Pi^r_K : L^2(K) \to \mathbb{P}_r(K)$ is the L^2 projection operator on K. Since $K \in B_{\Delta(K)}$, for any $v \in L^2(K)$, $E^r_K v$ is the direct extension of $\Pi^r_K v$ from K to $B_{\Delta(K)}$. Particularly, for any polynomial $v \in \mathbb{P}_r(K)$, $E^r_K v$ is the direct polynomial extension of v to the ball $B_{\Delta(K)}$. From the definition (10), we can prove the following basic property of E^r_K,

$$\|E^r_K v\|_{L^2(B_{\Delta(K)})} \leq C\|\Pi^r_K v\|_{L^2(K)} \leq C\|v\|_{L^2(K)}, \qquad \forall v \in L^2(K), \quad \forall K \in \mathcal{T}_h. \tag{11}$$

The norm equivalence on the finite dimensional space gives us that $\|v\|_{L^2(B(\mathbf{0}, C_\Delta C_\nu))} \leq C\|v\|_{L^2(B(\mathbf{0},1))}$ for $\forall v \in \mathbb{P}_r(B(\mathbf{0}, C_\Delta C_\nu))$. Considering the affine mapping from $B(\mathbf{0}, 1)$ to $B(\boldsymbol{x}_K, \rho)$, we derive that

$$\|E^r_K v\|_{L^2(B_{\Delta(K)})} \leq C\|E^r_K v\|_{L^2(B(\boldsymbol{x}_K, \rho))} = C\|\Pi^r_K v\|_{L^2(B(\boldsymbol{x}_K, \rho))} \leq C\|\Pi^r_K v\|_{L^2(K)}, \qquad \forall v \in L^2(K),$$

which leads to the stability property (11). We refer to [17] for more discussion on the constants C appearing in (11).

For any $v \in L^2(\Omega_{h,i})(i = 0,1)$, there holds $v|_K \in L^2(K)$ and $E_K^r(v|_K) \in \mathbb{P}_r(B_{\Delta(K)})$ for $\forall K \in \mathcal{T}_{h,i}$. From (10), $E_K^r(v|_K)$ is the direct polynomial extension of $\Pi_K^r(v|_K)$ from K to the ball $B_{\Delta(K)}$. Hereafter, we simply write $E_K^r(v|_K)$ as $E_K^r v$ for $\forall v \in L^2(\Omega_{h,i})$. This notation is most frequently used for v is a piecewise polynomial function of degree r on the mesh $\mathcal{T}_{h,i}$. In this case, $E_K^r v \in \mathbb{P}_r(B_{\Delta(K)})$ is just the direct extension of $v|_K$ from K to $B_{\Delta(K)}$ for $\forall K \in \mathcal{T}_{h,i}$.

Based on E_K^r, two bilinear forms $s_{h,0}^r(\cdot, \cdot)$ and $s_{h,1}^r(\cdot, \cdot)$ that satisfy the properties **P1** and **P2** can be constructed as

$$s_{h,i}^r(v,w) := \sum_{K \in \mathcal{T}_h^\Gamma} \int_K \left(v - E_{M^i(K)}^r v \right)\left(w - E_{M^i(K)}^r w \right) \mathrm{d}\boldsymbol{x}, \qquad \forall v, w \in L^2(\Omega_{h,i}),\ i = 0,1. \tag{12}$$

Notice that any cut element $K \in \mathcal{T}_h^\Gamma$ has that $K \subset B_{\Delta(M^i(K))}$. Therefore, $E_{M^i(K)}^r v \in \mathbb{P}_r(B_{\Delta(M^i(K))})$ is well defined on K. Then, we verify the properties **P1** and **P2** for the forms (12). By (11), we know that

$$\left\| E_{M^i(K)}^r v \right\|_{L^2(K)} \le \left\| E_{M^i(K)}^r v \right\|_{L^2\left(B_{\Delta(M^i(K))}\right)} \le C\|v\|_{L^2(M^i(K))}, \qquad \forall v \in L^2(\Omega_{h,i}),\ \forall K \in \mathcal{T}_{h,i}.$$

Combining with the triangle inequality, one can find that for $\forall v \in L^2(\Omega_{h,i})$,

$$s_{h,i}^r(v,v) \le \sum_{K \in \mathcal{T}_h^\Gamma} \left(\|v\|_{L^2(K)}^2 + \left\| E_{M^i(K)}^r v \right\|_{L^2(K)}^2 \right) \le C \sum_{K \in \mathcal{T}_h^\Gamma} \left(\|v\|_{L^2(K)}^2 + \|v\|_{L^2(M^i(K))}^2 \right) \le C\|v\|_{L^2(\Omega_{h,i})}^2,$$

which gives the upper bound in (8). Again by (11) and the triangle inequality, there holds

$$\|v\|_{L^2(\Omega_h^\Gamma)}^2 = \sum_{K \in \mathcal{T}_h^\Gamma} \|v\|_{L^2(K)}^2 \le C \sum_{K \in \mathcal{T}_h^\Gamma} \left(\left\| v - E_{M^i(K)}^r v \right\|_{L^2(K)}^2 + \left\| E_{M^i(K)}^r v \right\|_{L^2(K)}^2 \right)$$

$$\le C s_{h,i}^r(v,v) + C \sum_{K \in \mathcal{T}_h^\Gamma} \|v\|_{L^2(M^i(K))}^2 \le C \left(|v|_{s_{h,i}^r}^2 + \|v\|_{L^2(\Omega_{h,i}^\circ)}^2 \right), \qquad \forall v \in L^2(\Omega_{h,i}).$$

The lower bound of (8) is reached and the property **P1** holds for (12). We turn to the weak consistency **P2**. Given any $v \in H^{s+1}(\Omega)$, and for any $K \in \mathcal{T}_h^\Gamma$, there exists $p_K \in \mathbb{P}_r(B_{\Delta(K)})$ such that $\|v - p_K\|_{L^2(B_{\Delta(K)})} \le Ch^t\|v\|_{H^{s+1}(B_{\Delta(K)})}$ [22]. From (11), we obtain that

$$\left\| v - E_{M^i(K)}^r v \right\|_{L^2(K)} \le \|v - p_K\|_{L^2(K)} + \left\| E_{M^i(K)}^r(p_K - v) \right\|_{L^2(K)}$$

$$\le \|v - p_K\|_{L^2(K)} + C\|p_K - v\|_{L^2(M^i(K))} \le C\|v - p_K\|_{L^2(B_{\Delta(K)})} \le Ch^t\|v\|_{H^{s+1}(B_{\Delta(K)})}.$$

Summation over all cut elements indicates the estimate (9), *i.e.* the property **P2** is reached.

Now the penalty forms $s_{h,i}^r(\cdot, \cdot)(i = 0,1)$ are defined for scalar-valued functions in $L^2(\Omega_{h,i})$, and it is natural to extend $s_{h,i}^r(\cdot, \cdot)$ for vector- and tensor-valued functions in a componentwise manner. Let $\boldsymbol{v}, \boldsymbol{w} \in \mathbf{L}^2(\Omega_{h,i})$ with $\boldsymbol{v} = (v_j)_d, \boldsymbol{w} = (w_j)_d$ be the vector-valued functions, and let $\boldsymbol{\tau}, \boldsymbol{\rho} \in \mathbb{L}^2(\Omega_{h,i})$ with $\boldsymbol{\tau} = (\tau_{jk})_{d \times d}, \boldsymbol{\rho} = (\rho_{jk})_{d \times d}$ be the tensor-valued functions, we define

$$s_{h,i}^r(\boldsymbol{v}, \boldsymbol{w}) := \sum_{j=1}^d s_{h,i}^r(v_j, w_j), \quad s_{h,i}^r(\boldsymbol{\tau}, \boldsymbol{\rho}) := \sum_{1 \le j,k \le d} s_{h,i}^r(\tau_{jk}, \rho_{jk}),$$

with the induced seminorms $|\boldsymbol{v}|_{s_{h,i}^r}^2 := s_{h,i}^r(\boldsymbol{v}, \boldsymbol{v})$ and $|\boldsymbol{\tau}|_{s_{h,i}^r}^2 := s_{h,i}^r(\boldsymbol{\tau}, \boldsymbol{\tau})$. The properties **P1** and **P2** can be extended for vector- and tensor-valued functions without any difficulty. In the computer implementation, the

bilinear forms are always with piecewise polynomial spaces. The operator E_K^r is equivalent to the direct extension operator, and there is no need to implement the L^2 projection in (10).

We close this section by giving the H^1 trace estimate [31, 33, 53] on the interface.

Lemma 1. *There exists a constant C such that*

$$\|w\|_{L^2(\Gamma_K)}^2 \leq C\Big(h_K\|w\|_{H^1(K)}^2 + h_K^{-1}\|w\|_{L^2(K)}^2\Big), \qquad \forall w \in H^1(K), \ \forall K \in \mathcal{T}_h^\Gamma. \tag{13}$$

We refer to [31] for the proof only assuming Γ is Lipschitz. The H^1 trace estimate is fundamental in the penalty-type unfitted finite element methods, as the main tool to handle the numerical error on the interface, such as [3, 10, 30, 33, 36, 37, 49, 53]. By (13), the error on the interface can be bounded by the estimates on elements. However, this trace estimate (13) requires the H^1 regularity. For the problem with the exact solution in $H^s(\mathrm{div}; \Omega_0 \cup \Omega_1)$ or $H^s(\mathrm{curl}; \Omega_0 \cup \Omega_1)$, applying (13) to estimate the numerical errors on the interface will lead to a suboptimal convergence rate. The optimal convergence needs a higher regularity assumption that the exact solution is piecewise H^{s+1}-smooth. see [43, 48] for unfitted methods on $H(\mathrm{div})$- and $H(\mathrm{curl})$-interface problems. In addition, the trace estimate (13) is not suitable for the solution of low regularity, such as the solution belongs to the space $H^s(\Omega_0 \cup \Omega_1)$ with $s < 1$.

3. Least squares functional for the interface problem

In this section, we introduce an associated least squares functional to the interface problem (1). Let $\boldsymbol{\Sigma} := \boldsymbol{\Sigma}^0$ and $\mathbf{V} := \mathbf{V}^1$ be the spaces coinciding with $s = 0$ in (3). We define the quadratic functional $\mathcal{J}(\cdot; \cdot)$ by

$$\mathcal{J}(\boldsymbol{\tau}, \boldsymbol{v}; \boldsymbol{f}, \boldsymbol{a}, \boldsymbol{b}) := J(\boldsymbol{\tau}, \boldsymbol{v}; \boldsymbol{f}) + B^{\boldsymbol{\sigma}}(\boldsymbol{\tau}; \boldsymbol{a}) + B^{\boldsymbol{u}}(\boldsymbol{v}; \boldsymbol{b}), \quad \forall(\boldsymbol{\tau}, \boldsymbol{v}) \in \boldsymbol{\Sigma} \times \mathbf{V}, \tag{14}$$

where

$$J(\boldsymbol{\tau}, \boldsymbol{v}; \boldsymbol{f}) := \|\mathcal{A}\boldsymbol{\tau} - \boldsymbol{\varepsilon}(\boldsymbol{v})\|_{L^2(\Omega_0 \cup \Omega_1)}^2 + \|\nabla \cdot \boldsymbol{\tau} + \boldsymbol{f}\|_{L^2(\Omega_0 \cup \Omega_1)}^2,$$
$$B^{\boldsymbol{\sigma}}(\boldsymbol{\tau}; \boldsymbol{a}) := \|[\![\boldsymbol{\tau}]\!]_N - \boldsymbol{a}\|_{H^{-1/2}(\Gamma)}^2, \quad B^{\boldsymbol{u}}(\boldsymbol{v}; \boldsymbol{b}) := \|[\![\boldsymbol{v}]\!] - \boldsymbol{b}\|_{H^{1/2}(\Gamma)}^2, \qquad \forall(\boldsymbol{\sigma}, \boldsymbol{v}) \in \boldsymbol{\Sigma} \times \mathbf{V}. \tag{15}$$

The trace terms $B^{\boldsymbol{\sigma}}(\cdot; \cdot)$ and $B^{\boldsymbol{u}}(\cdot; \cdot)$ are well-defined since $[\![\boldsymbol{\tau}]\!]_N|_\Gamma \in H^{-1/2}(\Gamma)$ and $[\![\boldsymbol{v}]\!]|_\Gamma \in H^{1/2}(\Gamma)$ for $\forall(\boldsymbol{\tau}, \boldsymbol{v}) \in \boldsymbol{\Sigma} \times \mathbf{V}$ from the trace theory. The exact solution $(\boldsymbol{\sigma}, \boldsymbol{u})$ clearly minimizes the functional $\mathcal{J}(\cdot; \cdot)$ since $\mathcal{J}(\boldsymbol{\sigma}, \boldsymbol{u}; \boldsymbol{f}, \boldsymbol{a}, \boldsymbol{b}) = 0$. In fact, we can prove that $(\boldsymbol{\sigma}, \boldsymbol{u})$ is also the unique solution to the minimization problem $\inf_{(\boldsymbol{\tau}, \boldsymbol{v}) \in \boldsymbol{\Sigma} \times \mathbf{V}} \mathcal{J}(\boldsymbol{\tau}, \boldsymbol{v}; \boldsymbol{f}, \boldsymbol{a}, \boldsymbol{b})$. For this purpose, we will give a norm equivalence property of $\mathcal{J}(\cdot, \cdot)$, and the norm equivalence is also crucial for the error analysis in the least squares finite element method [5]. The pair of spaces $\boldsymbol{\Sigma} \times \mathbf{V}$ can be naturally equipped with the norm

$$\|(\boldsymbol{\tau}, \boldsymbol{v})\|_{\mathbf{e}}^2 := \|\boldsymbol{\tau}\|_{H(\mathrm{div}; \Omega_0 \cup \Omega_1)}^2 + \|\boldsymbol{v}\|_{H^1(\Omega_0 \cup \Omega_1)}^2, \qquad \forall(\boldsymbol{\tau}, \boldsymbol{v}) \in \boldsymbol{\Sigma} \times \mathbf{V}.$$

The equivalence between $\mathcal{J}(\cdot; \cdot)$ and $\|\cdot\|_{\mathbf{e}}$ is given in the following lemma.

Lemma 2. *There exist constants C such that*

$$\|(\boldsymbol{\tau}, \boldsymbol{v})\|_{\mathbf{e}} \leq C\Big(\|\mathcal{A}\boldsymbol{\tau} - \boldsymbol{\varepsilon}(\boldsymbol{v})\|_{L^2(\Omega_0 \cup \Omega_1)} + \|\nabla \cdot \boldsymbol{\tau}\|_{L^2(\Omega_0 \cup \Omega_1)}$$
$$+ \|[\![\boldsymbol{\tau}]\!]_N\|_{H^{-1/2}(\Gamma)} + \|[\![\boldsymbol{v}]\!]\|_{H^{1/2}(\Gamma)}\Big) \leq C\|(\boldsymbol{\tau}, \boldsymbol{v})\|_{\mathbf{e}}, \qquad \forall(\boldsymbol{\tau}, \boldsymbol{v}) \in \boldsymbol{\Sigma} \times \mathbf{V}. \tag{16}$$

Proof. By the definition of $\mathcal{A}$, we have that

$$\|\mathcal{A}\boldsymbol{\tau}\|_{L^2(\Omega)}^2 = \frac{1}{(2\mu)^2}\left(\|\boldsymbol{\tau}\|_{L^2(\Omega)}^2 - \frac{\lambda(d\lambda + 4\mu)}{(d\lambda + 2\mu)^2}\|\mathrm{tr}(\boldsymbol{\tau})\|_{L^2(\Omega)}^2\right) \leq \frac{1}{(2\mu)^2}\|\boldsymbol{\tau}\|_{L^2(\Omega)}^2 \leq C\|\boldsymbol{\tau}\|_{L^2(\Omega)}^2. \tag{17}$$

The term $\|\mathcal{A}\boldsymbol{\tau} - \boldsymbol{\varepsilon}(\boldsymbol{v})\|_{L^2(\Omega_0 \cup \Omega_1)}$ can be bounded by $\|(\boldsymbol{\tau}, \boldsymbol{v})\|_{\mathbf{e}}$, following from the triangle inequality and (17). From the embedding theory, we know that

$$\|[\![\boldsymbol{\tau}]\!]_N\|_{H^{-1/2}(\Gamma)} \leq \|\boldsymbol{n} \cdot \boldsymbol{\tau}|_{\Omega_0}\|_{H^{-1/2}(\Gamma)} + \|\boldsymbol{n} \cdot \boldsymbol{\tau}|_{\Omega_1}\|_{H^{-1/2}(\Gamma)} \leq C\|\boldsymbol{\tau}\|_{H(\mathrm{div};\Omega_0 \cup \Omega_1)}.$$

Similarly, there holds $\|[\![\boldsymbol{v}]\!]\|_{H^{1/2}(\Gamma)} \leq C\|\boldsymbol{v}\|_{H^1(\Omega_0 \cup \Omega_1)}$. The upper bound in (16) is reached.

From Theorem 3.1 of [13], the lower bound in (16) holds for $\boldsymbol{\tau} \in \mathbf{H}(\mathrm{div}; \Omega)$ and $\boldsymbol{v} \in \mathbf{H}^1(\Omega)$, *i.e.* the lower bound holds for $\boldsymbol{\tau}$ and $\boldsymbol{v}$ satisfying $[\![\boldsymbol{\tau}]\!]_N|_\Gamma = [\![\boldsymbol{v}]\!]|_\Gamma = \mathbf{0}$. For the general case, we construct two auxiliary functions $\widetilde{\boldsymbol{\tau}}$ and $\widetilde{\boldsymbol{v}}$ to prove (16). Consider the elliptic problems

$$-\Delta \boldsymbol{w}_0 = \mathbf{0}, \quad \text{in } \Omega_0, \quad \partial_{\boldsymbol{n}} \boldsymbol{w}_0 = -[\![\boldsymbol{\tau}]\!]_N|_\Gamma, \quad \text{on } \Gamma,$$

and

$$-\Delta \boldsymbol{w}_1 = \mathbf{0}, \quad \text{in } \Omega_0, \quad \boldsymbol{w}_1 = -[\![\boldsymbol{v}]\!]|_\Gamma, \quad \text{on } \Gamma.$$

Since $[\![\boldsymbol{\tau}]\!]_N|_\Gamma \in \mathbf{H}^{-1/2}(\Gamma)$ and $[\![\boldsymbol{v}]\!]|_\Gamma \in H^{1/2}(\Gamma)$, we can know that both problems have solutions in $\mathbf{H}^1(\Omega_0)$ with $\|\boldsymbol{w}_0\|_{H^1(\Omega_0)} \leq C\|[\![\boldsymbol{\tau}]\!]_N\|_{H^{-1/2}(\Gamma)}$ and $\|\boldsymbol{w}_1\|_{H^1(\Omega_0)} \leq C\|[\![\boldsymbol{v}]\!]\|_{H^{1/2}(\Gamma)}$. We then extend $\boldsymbol{w}_0$ and $\boldsymbol{w}_1$ to the domain Ω by zero. Let $\widetilde{\boldsymbol{\tau}} := \boldsymbol{\tau} + \nabla \boldsymbol{w}_0$ and $\widetilde{\boldsymbol{v}} := \boldsymbol{v} + \boldsymbol{w}_1$, we have that $[\![\widetilde{\boldsymbol{\tau}}]\!]_N|_\Gamma = [\![\widetilde{\boldsymbol{v}}]\!]|_\Gamma = \mathbf{0}$, which allows us to derive that

$$\|(\boldsymbol{\tau}, \boldsymbol{v})\|_{\mathbf{e}} \leq \|(\widetilde{\boldsymbol{\tau}}, \widetilde{\boldsymbol{v}})\|_{\mathbf{e}} + \|(\boldsymbol{\tau} - \widetilde{\boldsymbol{\tau}}, \boldsymbol{v} - \widetilde{\boldsymbol{v}})\|_{\mathbf{e}} \leq \|(\widetilde{\boldsymbol{\tau}}, \widetilde{\boldsymbol{v}})\|_{\mathbf{e}} + C\Big(\|[\![\boldsymbol{\tau}]\!]_N\|_{H^{-1/2}(\Gamma)} + \|[\![\boldsymbol{v}]\!]\|_{H^{1/2}(\Gamma)}\Big),$$

and

$$\begin{aligned}
\|(\widetilde{\boldsymbol{\tau}}, \widetilde{\boldsymbol{v}})\|_{\mathbf{e}} &\leq C\Big(\|\mathcal{A}\widetilde{\boldsymbol{\tau}} - \boldsymbol{\varepsilon}(\widetilde{\boldsymbol{v}})\|_{L^2(\Omega_0 \cup \Omega_1)} + \|\nabla \cdot \widetilde{\boldsymbol{\tau}}\|_{L^2(\Omega_0 \cup \Omega_1)}\Big) \\
&\leq C\Big(\|\mathcal{A}\boldsymbol{\tau} - \boldsymbol{\varepsilon}(\boldsymbol{v})\|_{L^2(\Omega_0 \cup \Omega_1)} + \|\nabla \cdot \boldsymbol{\tau}\|_{L^2(\Omega_0 \cup \Omega_1)}\Big) + C\Big(\|\boldsymbol{\tau} - \widetilde{\boldsymbol{\tau}}\|_{H(\mathrm{div};\Omega_0 \cup \Omega_1)} + \|\boldsymbol{v} - \widetilde{\boldsymbol{v}}\|_{H^1(\Omega_0 \cup \Omega_1)}\Big) \\
&\leq C\Big(\|\mathcal{A}\boldsymbol{\tau} - \boldsymbol{\varepsilon}(\boldsymbol{v})\|_{L^2(\Omega_0 \cup \Omega_1)} + \|\nabla \cdot \boldsymbol{\tau}\|_{L^2(\Omega_0 \cup \Omega_1)} + \|[\![\boldsymbol{\tau}]\!]_N\|_{H^{-1/2}(\Gamma)} + \|[\![\boldsymbol{v}]\!]\|_{H^{1/2}(\Gamma)}\Big).
\end{aligned}$$

Combining the above two estimates leads to the lower bound in (16), which completes the proof. $\square$

It is noted that the generic constants in (16) are independent of λ, which ensure the proposed methods are robust when $\lambda \to \infty$.

From the definition to $\mathcal{J}(\cdot; \cdot)$, it can be observed that

$$\mathcal{J}(\boldsymbol{\tau}, \boldsymbol{v}; \mathbf{0}, \mathbf{0}, \mathbf{0}) = \|\mathcal{A}\boldsymbol{\tau} - \boldsymbol{\varepsilon}(\boldsymbol{v})\|_{L^2(\Omega_0 \cup \Omega_1)}^2 + \|\nabla \cdot \boldsymbol{\tau}\|_{L^2(\Omega_0 \cup \Omega_1)}^2 + \|[\![\boldsymbol{\tau}]\!]_N\|_{H^{-1/2}(\Gamma)}^2 + \|[\![\boldsymbol{v}]\!]\|_{H^{1/2}(\Gamma)}^2.$$

Therefore, $\mathcal{J}(\boldsymbol{\tau}, \boldsymbol{v}; \mathbf{0}, \mathbf{0}, \mathbf{0})$ is equivalent to $\|(\boldsymbol{\tau}, \boldsymbol{v})\|_{\mathbf{e}}^2$. Let $(\boldsymbol{\tau}, \boldsymbol{v}) \in \boldsymbol{\Sigma} \times \mathbf{V}$ be the solution to the problem (1) with $\boldsymbol{f} = \boldsymbol{a} = \boldsymbol{b} = \mathbf{0}$, which implies that $\mathcal{J}(\boldsymbol{\tau}, \boldsymbol{v}; \mathbf{0}, \mathbf{0}, \mathbf{0}) = 0$. From the equivalence (16), we immediately find that $\boldsymbol{\tau}$ and $\boldsymbol{v}$ are zero functions. Hence, the minimization problem $\inf_{(\boldsymbol{\tau}, \boldsymbol{v}) \in \boldsymbol{\Sigma} \times \mathbf{V}} \mathcal{J}(\boldsymbol{\tau}, \boldsymbol{v}; \boldsymbol{f}, \boldsymbol{a}, \boldsymbol{b})$ admits a unique solution.

4. LEAST SQUARES FINITE ELEMENT METHODS

In this section, we present the numerical schemes for solving the elasticity interface problem (1). We begin by introducing the approximation finite element spaces. For the mesh $\mathcal{T}_{h,i} (i = 0, 1)$, we define $\boldsymbol{\Sigma}_{h,i}^m \subset \mathbf{H}(\mathrm{div}; \Omega_{h,i})$ as the $H(\mathrm{div})$-conforming tensor-valued piecewise polynomial space of degree m. As the definition of $\mathbf{H}(\mathrm{div}; \Omega_{h,i})$, each column of functions in $\boldsymbol{\Sigma}_{h,i}^m$ belongs to $H(\mathrm{div}; \Omega_{h,i})$, *i.e.* the $\mathbf{BDM}_m$ space or the $\mathbf{RT}_m$ space. In this paper, the schemes are established on $\boldsymbol{\Sigma}_{h,i}^m$ is the $\mathbf{BDM}_m$ space. The methods and the analysis can be extended to the $\mathbf{RT}_m$ space without any difficulty. We define $\mathbf{V}_{h,0}^m \subset \mathbf{H}^1(\Omega_{h,0})$ as the vector-valued C^0 finite element space of degree m, and define $\mathbf{V}_{h,1}^m \subset \mathbf{H}^1(\Omega_{h,1})$ as the C^0 finite element space with zero trace on $\partial \Omega$, *i.e.* $\mathbf{V}_{h,1}^m|_{\partial \Omega} = \mathbf{0}$.

We define the extended approximation spaces $\boldsymbol{\Sigma}_h^m := \boldsymbol{\Sigma}_{h,0}^m \cdot \chi_0 + \boldsymbol{\Sigma}_{h,1}^m \cdot \chi_1$ and $\mathbf{V}_h^m := \mathbf{V}_{h,0}^m \cdot \chi_0 + \mathbf{V}_{h,1}^m \cdot \chi_1$ for the stress and the displacement, respectively, where the characteristic function χ_i is defined in (4). Then, any $\boldsymbol{\tau}_h \in \boldsymbol{\Sigma}_h^m$ and any $\boldsymbol{v}_h \in \mathbf{V}_h^m$ admit a unique decomposition that

$$\boldsymbol{\tau}_h = \boldsymbol{\tau}_{h,0} \cdot \chi_0 + \boldsymbol{\tau}_{h,1} \cdot \chi_1, \quad \boldsymbol{v}_h = \boldsymbol{v}_{h,0} \cdot \chi_0 + \boldsymbol{v}_{h,1} \cdot \chi_1, \tag{18}$$

where $\boldsymbol{\tau}_{h,i} \in \boldsymbol{\Sigma}_{h,i}^m$, $\boldsymbol{v}_{h,i} \in \mathbf{V}_{h,i}^m$. As the decomposition (4), we also formally let $\pi_i (i = 0, 1)$ be the projection operator such that $\pi_i \boldsymbol{\tau}_h := \boldsymbol{\tau}_{h,i} \in \boldsymbol{\Sigma}_{h,i}^m (\forall \boldsymbol{\tau}_h \in \boldsymbol{\Sigma}_h^m)$ and $\pi_i \boldsymbol{v}_h := \boldsymbol{v}_{h,i} \in \mathbf{V}_{h,i}^m (\forall \boldsymbol{v}_h \in \mathbf{V}_h^m)$.

The approximation spaces are conforming in the sense that $\boldsymbol{\Sigma}_h^m \subset \boldsymbol{\Sigma}$ and $\mathbf{V}_h^m \subset \mathbf{V}$. A natural idea to seek numerical solutions is minimizing the functional $\mathcal{J}(\cdot; \cdot)$ over the discrete conforming spaces, *i.e.* $\inf_{(\boldsymbol{\tau}_h, \boldsymbol{v}_h) \in \boldsymbol{\Sigma}_h^m \times \mathbf{V}_h^m} \mathcal{J}(\boldsymbol{\tau}_h, \boldsymbol{v}_h; \boldsymbol{f}, \boldsymbol{a}, \boldsymbol{b})$, which, however, is not an easy task. It is impossible to directly compute the trace terms $B^{\boldsymbol{\sigma}}(\cdot; \cdot)$ and $B^{\boldsymbol{u}}(\cdot; \cdot)$ in (15), and the minimization problem cannot be readily rewritten into a variational problem by the means of Euler–Lagrange equation because of the presence of $\| \cdot \|_{H^{1/2}(\Gamma)}$ and $\| \cdot \|_{H^{-1/2}(\Gamma)}$. The main idea of the proposed method is to apply some computational trace terms to replace $B^{\boldsymbol{\sigma}}(\cdot; \cdot)$ and $B^{\boldsymbol{u}}(\cdot; \cdot)$ in $\mathcal{J}(\cdot; \cdot)$ to define new functionals. The numerical approximations are then obtained by minimizing the new functionals.

For the convergence analysis, we define the spaces $\boldsymbol{\Sigma}_h := \{\boldsymbol{\sigma}\} \cup \boldsymbol{\Sigma}_h^m$ and $\mathbf{V}_h := \{\boldsymbol{u}\} \cup \mathbf{V}_h^m$, which contain the exact solution and all finite element functions, respectively. Here, we introduce the ghost penalty forms for the spaces $\boldsymbol{\Sigma}_h$ and $\mathbf{V}_h$ from the forms (12). By the definition of $\pi_i (i = 0, 1)$, there holds $\pi_i \boldsymbol{\tau}_h \in \mathbb{L}^2(\Omega_{h,i})(\forall \boldsymbol{\tau}_h \in \boldsymbol{\Sigma}_h)$ and $\pi_i \boldsymbol{v}_h \in \mathbf{L}^2(\Omega_{h,i})(\forall \boldsymbol{v}_h \in \mathbf{V}_h)$. From (12), we define the forms $\mathcal{S}_h^m(\cdot, \cdot)$ and $\mathcal{G}_h^m(\cdot, \cdot)$ for $\boldsymbol{\Sigma}_h$ and $\mathbf{V}_h$, respectively, as

$$\mathcal{S}_h^m(\boldsymbol{\tau}_h, \boldsymbol{\rho}_h) := s_{h,0}^m(\pi_0 \boldsymbol{\tau}_h, \pi_0 \boldsymbol{\rho}_h) + s_{h,1}^m(\pi_1 \boldsymbol{\tau}_h, \pi_1 \boldsymbol{\rho}_h), \qquad \forall \boldsymbol{\tau}_h, \boldsymbol{\rho}_h \in \boldsymbol{\Sigma}_h,$$

with the induced seminorm $|\boldsymbol{\tau}_h|_{\mathcal{S}_h^m}^2 := \mathcal{S}_h^m(\boldsymbol{\tau}_h, \boldsymbol{\tau}_h)(\forall \boldsymbol{\tau}_h \in \boldsymbol{\Sigma}_h)$, and

$$\mathcal{G}_h^m(\boldsymbol{v}_h, \boldsymbol{w}_h) := s_{h,0}^m(\pi_0 \boldsymbol{v}_h, \pi_0 \boldsymbol{w}_h) + s_{h,1}^m(\pi_1 \boldsymbol{v}_h, \pi_1 \boldsymbol{w}_h), \qquad \forall \boldsymbol{v}_h, \boldsymbol{w}_h \in \mathbf{V}_h,$$

with the induced seminorm $|\boldsymbol{v}_h|_{\mathcal{G}_h^m}^2 := \mathcal{G}_h^m(\boldsymbol{v}_h, \boldsymbol{v}_h)(\forall \boldsymbol{v}_h \in \mathbf{V}_h)$. From properties **P1** and **P2**, we have the following estimates,

$$\begin{aligned}
\|\pi_0 \boldsymbol{\tau}_h\|_{L^2(\Omega_{h,0})} + \|\pi_1 \boldsymbol{\tau}_h\|_{L^2(\Omega_{h,1})} &\leq C\Big(\|\boldsymbol{\tau}_h\|_{L^2(\Omega)} + |\boldsymbol{\tau}_h|_{\mathcal{S}_h^m}\Big) \\
&\leq C\Big(\|\pi_0 \boldsymbol{\tau}_h\|_{L^2(\Omega_{h,0})} + \|\pi_1 \boldsymbol{\tau}_h\|_{L^2(\Omega_{h,1})}\Big), \qquad \forall \boldsymbol{\tau}_h \in \boldsymbol{\Sigma}_h, \quad (19) \\
\|\pi_0 \boldsymbol{v}_h\|_{L^2(\Omega_{h,0})} + \|\pi_1 \boldsymbol{v}_h\|_{L^2(\Omega_{h,1})} &\leq C\Big(\|\boldsymbol{v}_h\|_{L^2(\Omega)} + |\boldsymbol{v}_h|_{\mathcal{G}_h^m}\Big) \\
&\leq C\Big(\|\pi_0 \boldsymbol{v}_h\|_{L^2(\Omega_{h,0})} + \|\pi_1 \boldsymbol{v}_h\|_{L^2(\Omega_{h,1})}\Big), \qquad \forall \boldsymbol{v}_h \in \mathbf{V}_h, \quad (20)
\end{aligned}$$

and

$$\begin{aligned}
|\boldsymbol{\tau}|_{\mathcal{S}_h^m} &\leq Ch^t \|\boldsymbol{\tau}\|_{H^s(\Omega)}, \qquad \forall \boldsymbol{\tau} \in \mathbb{H}^s(\Omega), \qquad t = \min(m, s), \\
|\boldsymbol{v}|_{\mathcal{G}_h^m} &\leq Ch^t \|\boldsymbol{v}\|_{H^{s+1}(\Omega)}, \qquad \forall \boldsymbol{v} \in \mathbf{H}^{s+1}(\Omega), \quad t = \min(m+1, s+1).
\end{aligned}$$

4.1. L^2 norm least squares finite element method for $s \geq 1$

We present the numerical method that defines the least squares functional using only L^2 norms. This method requires the exact solution has the regularity $s \geq 1$. As mentioned earlier, we will bound the trace terms $B^{\boldsymbol{\sigma}}(\cdot; \cdot)$ and $B^{\boldsymbol{u}}(\cdot; \cdot)$ by applying stronger and easily-computational norms. In this case, the space $\mathbb{H}^s(\Omega_i)$ continuously embeds into $\mathbb{L}^2(\Gamma)$, which indicates that $[\![\boldsymbol{\tau}]\!]_N|_\Gamma = \boldsymbol{n} \cdot ((\boldsymbol{\tau}|_{\Omega_0} - \boldsymbol{\tau}|_{\Omega_1}))|_\Gamma \in \mathbf{L}^2(\Gamma)$ for $\forall \boldsymbol{\tau} \in \boldsymbol{\Sigma}^s$. From the definition (7), one can find that

$$\|[\![\boldsymbol{\tau}]\!]_N\|_{H^{-1/2}(\Gamma)} \leq C\|[\![\boldsymbol{\tau}]\!]_N\|_{L^2(\Gamma)}, \qquad \forall \boldsymbol{\tau} \in \boldsymbol{\Sigma}^s. \tag{21}$$

For the displacement variable, we also give a stronger norm for $\|\cdot\|_{H^{1/2}(\Gamma)}$. We first consider the case that Γ is C^2-smooth. By the embedding relationship $\mathbf{H}^{s+1}(\Omega_i) \hookrightarrow \mathbf{H}^1(\Gamma)$, we know that $[\![v]\!]|_\Gamma \in \mathbf{H}^1(\Gamma)$ for $\forall v \in \mathbf{V}^{s+1}$. Here the norm $\|v\|^2_{H^1(\Gamma)} = \|v\|^2_{L^2(\Gamma)} + \|\nabla_\Gamma v\|^2_{L^2(\Gamma)}$, where ∇_Γ denotes the tangential gradient on the interface [23]. The tangential gradient $\nabla_\Gamma v(\boldsymbol{x})$ for $\forall v \in \mathbf{H}^1(\Gamma)$ only depends on values of v on $\Gamma \cap U$, where U is a neighbourhood of $\boldsymbol{x}$. For $i = 0, 1$, we note that any finite element function $v_{h,i} \in \mathbf{V}^m_{h,i}(i = 0, 1)$ is continuous and piecewise smooth on Γ, i.e. $v_{h,i}|_\Gamma \in C^0(\Gamma)$ and $v_{h,i}|_{\Gamma_K} \in C^2(\Gamma_K)(\forall K \in \mathcal{T}^\Gamma_h)$, which gives that $v_{h,i} \in \mathbf{H}^1(\Gamma)$. Thus, we have that $[\![v_h]\!]|_\Gamma = (\pi_0 v_h)|_\Gamma - (\pi_1 v_h)|_\Gamma \in \mathbf{H}^1(\Gamma)$ for $\forall v_h \in \mathbf{V}^m_h$. Consequently, we conclude that $[\![v_h]\!]|_\Gamma \in \mathbf{H}^1(\Gamma)$ for $\forall v_h \in \mathbf{V}_h$. By the Sobolev interpolation inequality ([47], Thm. 7.4) and the Cauchy–Schwarz inequality, we have that

$$\|[\![v]\!]\|^2_{H^{1/2}(\Gamma)} \le C\|[\![v]\!]\|_{L^2(\Gamma)}\|[\![v]\!]\|_{H^1(\Gamma)} \le C\Big(h^{-1}\|[\![v]\!]\|^2_{L^2(\Gamma)} + h\|[\![\nabla_\Gamma v]\!]\|^2_{L^2(\Gamma)}\Big), \qquad \forall v \in \mathbf{V}_h. \tag{22}$$

The right hand side of (22) only involves the L^2 norms, and can be easily computed.

For the case that Γ is polygonal (polyhedral), we let $\Gamma_j(1 \le j \le J)$ denote the sides of Γ, where every side Γ_j is a line segment (polygon). Since Γ has corners, the estimate (22) will be modified in a piecewise manner. From the embedding theory ([27], Thm. 1.5.2.1), we have that $\mathbf{H}^{s+1}(\Omega_i) \hookrightarrow \Pi^J_{j=1}\mathbf{H}^1(\Gamma_j)(i = 0, 1)$, which brings us that $[\![v]\!]|_{\Gamma_j} \in \mathbf{H}^1(\Gamma_j)(1 \le j \le J)$ for $\forall v \in \mathbf{V}^{s+1}$. Moreover, any $v^m_{h,i} \in \mathbf{V}^m_{h,i}(i = 0, 1)$ is continuous and piecewise smooth on each Γ_j. We can know that $(\pi_i v_h)|_{\Gamma_j} \in \mathbf{H}^1(\Gamma_j)$ and $[\![v_h]\!]|_{\Gamma_j} \in \mathbf{H}^1(\Gamma_j)$ for $\forall v_h \in \mathbf{V}^m_h$. Similar to (22), the norm $\|\cdot\|_{H^{1/2}(\Gamma)}$ can be bounded by

$$\|[\![v]\!]\|^2_{H^{1/2}(\Gamma)} \le C\sum_{j=1}^J\|[\![v]\!]\|^2_{H^{1/2}(\Gamma_j)} \le C\sum_{j=1}^J\|[\![v]\!]\|_{L^2(\Gamma_j)}\|[\![v]\!]\|_{H^1(\Gamma_j)} \le C\sum_{j=1}^J\Big(h^{-1}\|[\![v]\!]\|^2_{L^2(\Gamma_j)} + h\|[\![\nabla_\Gamma v]\!]\|^2_{L^2(\Gamma_j)}\Big)$$

$$\le Ch^{-1}\|[\![v]\!]\|^2_{L^2(\Gamma)} + Ch\sum_{j=1}^J\|[\![\nabla_\Gamma v]\!]\|^2_{L^2(\Gamma_j)}, \qquad \forall v \in \mathbf{V}_h. \tag{23}$$

The second inequality follows from the Sobolev interpolation inequality on each Γ_j. Compared with (22), the norm for the tangential gradient on Γ in (23) can be understood in a piecewise manner, i.e. we can formally let $\|\nabla_\Gamma \cdot\|^2_{L^2(\Gamma)} := \sum_{j=1}^J\|\nabla_\Gamma \cdot\|^2_{L^2(\Gamma_j)}$ for the case Γ is polygonal. Then, the upper bound of the estimate (23) will seem the same as the upper bound of (22). To simplify the representation, we use the notation consistent with the case that Γ is C^2.

Now, we define the discrete least squares functional $\mathcal{J}_h(\cdot; \cdot)$ that only involves the L^2 norms by

$$\mathcal{J}_h(\boldsymbol{\tau}_h, v_h; \boldsymbol{f}, \boldsymbol{a}, \boldsymbol{b}) := J(\boldsymbol{\tau}_h, v_h; \boldsymbol{f}) + B^{\boldsymbol{\sigma}}_h(\boldsymbol{\tau}_h; \boldsymbol{a}) + B^{\boldsymbol{u}}_h(v_h; \boldsymbol{b}) + |\boldsymbol{\tau}_h|^2_{\mathcal{S}^m_h} + |v_h|^2_{\mathcal{G}^m_h}, \quad \forall(\boldsymbol{\tau}_h, v_h) \in \boldsymbol{\Sigma}_h \times \mathbf{V}_h, \tag{24}$$

where $J(\cdot; \cdot)$ is defined in (15), and

$$B^{\boldsymbol{\sigma}}_h(\boldsymbol{\tau}_h; \boldsymbol{a}) := \|[\![\boldsymbol{\tau}_h]\!]_N - \boldsymbol{a}\|^2_{L^2(\Gamma)}, \quad B^{\boldsymbol{u}}_h(v_h; \boldsymbol{b}) := h^{-1}\|[\![v_h]\!] - \boldsymbol{b}\|^2_{L^2(\Gamma)} + h\|[\![\nabla_\Gamma v_h]\!] - \nabla_\Gamma \boldsymbol{b}\|^2_{L^2(\Gamma)}. \tag{25}$$

The last two seminorms in (24) are applied to guarantee the uniform upper bound of the condition number of the resulting linear system.

The numerical solutions are sought by solving the minimization problem

$$\inf_{(\boldsymbol{\tau}_h, v_h) \in \boldsymbol{\Sigma}^m_h \times \mathbf{V}^m_h} \mathcal{J}_h(\boldsymbol{\tau}_h, v_h; \boldsymbol{f}, \boldsymbol{a}, \boldsymbol{b}). \tag{26}$$

Since all terms in $\mathcal{J}_h(\cdot; \cdot)$ are defined with L^2 norms, the problem (26) can be solved by writing the corresponding Euler–Lagrange equation, which reads: seek $(\boldsymbol{\sigma}_h, \boldsymbol{u}_h) \in \boldsymbol{\Sigma}^m_h \times \mathbf{V}^m_h$ such that

$$a_{\mathcal{J}_h}(\boldsymbol{\sigma}_h, \boldsymbol{u}_h; \boldsymbol{\tau}_h, v_h) = l_{\mathcal{J}_h}(\boldsymbol{\tau}_h, v_h), \qquad \forall(\boldsymbol{\tau}_h, v_h) \in \boldsymbol{\Sigma}^m_h \times \mathbf{V}^m_h, \tag{27}$$

where the forms $a_{\mathcal{J}_h}(\cdot;\cdot)$ and $l_{\mathcal{J}_h}(\cdot)$ are defined as

$$
\begin{aligned}
a_{\mathcal{J}_h}(\boldsymbol{\tau}_h,\boldsymbol{v}_h;\boldsymbol{\rho}_h,\boldsymbol{w}_h) &:= (\mathcal{A}\boldsymbol{\tau}_h - \boldsymbol{\varepsilon}(\boldsymbol{v}_h), \mathcal{A}\boldsymbol{\rho}_h - \boldsymbol{\varepsilon}(\boldsymbol{w}_h))_{L^2(\Omega_0\cup\Omega_1)} + (\nabla\cdot\boldsymbol{\tau}_h, \nabla\cdot\boldsymbol{\rho}_h)_{L^2(\Omega_0\cup\Omega_1)} \\
&\quad + ([\![\boldsymbol{\tau}_h]\!]_N, [\![\boldsymbol{\rho}_h]\!]_N)_{L^2(\Gamma)} + h^{-1}([\![\boldsymbol{v}_h]\!], [\![\boldsymbol{w}_h]\!])_{L^2(\Gamma)} + h([\![\nabla_\Gamma\boldsymbol{v}_h]\!], [\![\nabla_\Gamma\boldsymbol{w}_h]\!])_{L^2(\Gamma)} \\
&\quad + \mathcal{S}_h^m(\boldsymbol{\tau}_h,\boldsymbol{\rho}_h) + \mathcal{G}_h^m(\boldsymbol{v}_h,\boldsymbol{w}_h),
\end{aligned}
$$

and

$$
l_{\mathcal{J}_h}(\boldsymbol{\tau}_h,\boldsymbol{v}_h) := -(\nabla\cdot\boldsymbol{\tau}_h, \boldsymbol{f})_{L^2(\Omega_0\cup\Omega_1)} + ([\![\boldsymbol{\tau}_h]\!]_N, \boldsymbol{a})_{L^2(\Gamma)} + h([\![\boldsymbol{v}_h]\!], \boldsymbol{b})_{L^2(\Gamma)} + h^{-1}([\![\nabla_\Gamma\boldsymbol{v}_h]\!], \nabla_\Gamma\boldsymbol{b})_{L^2(\Gamma)}.
$$

For the convergence analysis, we are aiming to show that $a_{\mathcal{J}_h}(\cdot;\cdot)$ is bounded and coercive, and satisfies a weakened Galerkin orthogonality property. Then, the error estimate can be reached in a standard argument.

We introduce the energy norm $\|\cdot\|_{\mathbf{e}_h}$ by

$$
\begin{aligned}
\|(\boldsymbol{\tau}_h,\boldsymbol{v}_h)\|_{\mathbf{e}_h}^2 &:= \|\boldsymbol{\tau}_h\|_{H(\mathrm{div};\Omega_0\cup\Omega_1)}^2 + \|\boldsymbol{v}_h\|_{H^1(\Omega_0\cup\Omega_1)}^2 + \|[\![\boldsymbol{\tau}_h]\!]_N\|_{L^2(\Gamma)}^2 \\
&\quad + h^{-1}\|[\![\boldsymbol{v}_h]\!]\|_{L^2(\Gamma)}^2 + h\|[\![\nabla_\Gamma\boldsymbol{v}_h]\!]\|_{L^2(\Gamma)}^2 + |\boldsymbol{\tau}_h|_{\mathcal{S}_h^m}^2 + |\boldsymbol{v}_h|_{\mathcal{G}_h^m}^2, \qquad \forall(\boldsymbol{\tau}_h,\boldsymbol{v}_h)\in\boldsymbol{\Sigma}_h\times\mathbf{V}_h.
\end{aligned}
$$

By (21)–(23), we have that

$$
B^{\boldsymbol{\sigma}}(\boldsymbol{\tau}_h;\mathbf{0}) = \|[\![\boldsymbol{\tau}_h]\!]_N\|_{H^{-1/2}(\Gamma)}^2 \le CB_h^{\boldsymbol{\sigma}}(\boldsymbol{\tau}_h;\mathbf{0}), \qquad \forall\boldsymbol{\tau}_h\in\boldsymbol{\Sigma}_h, \tag{28}
$$

and

$$
B^{\boldsymbol{u}}(\boldsymbol{v}_h;\mathbf{0}) = \|[\![\boldsymbol{v}_h]\!]\|_{H^{1/2}(\Gamma)}^2 \le CB_h^{\boldsymbol{u}}(\boldsymbol{v}_h;\mathbf{0}), \qquad \forall\boldsymbol{v}_h\in\mathbf{V}_h. \tag{29}
$$

Combining the definitions to $\mathcal{J}(\cdot;\cdot)$ and $\mathcal{J}_h(\cdot;\cdot)$ and the above two estimates, there holds

$$
\mathcal{J}(\boldsymbol{\tau}_h,\boldsymbol{v}_h;\mathbf{0},\mathbf{0},\mathbf{0}) \le C\mathcal{J}_h(\boldsymbol{\tau}_h,\boldsymbol{v}_h;\mathbf{0},\mathbf{0},\mathbf{0}) = a_{\mathcal{J}_h}(\boldsymbol{\tau}_h,\boldsymbol{v}_h;\boldsymbol{\tau}_h,\boldsymbol{v}_h), \qquad \forall(\boldsymbol{\tau}_h,\boldsymbol{v}_h)\in\boldsymbol{\Sigma}_h^m\times\mathbf{V}_h^m.
$$

From Lemma 2, we find that

$$
\|\boldsymbol{\tau}_h\|_{H(\mathrm{div};\Omega_0\cup\Omega_1)}^2 + \|\boldsymbol{v}_h\|_{H^1(\Omega_0\cup\Omega_1)}^2 \le Ca_{\mathcal{J}_h}(\boldsymbol{\tau}_h,\boldsymbol{v}_h;\boldsymbol{\tau}_h,\boldsymbol{v}_h), \qquad \forall(\boldsymbol{\tau}_h,\boldsymbol{v}_h)\in\boldsymbol{\Sigma}_h^m\times\mathbf{V}_h^m. \tag{30}
$$

The boundedness and the coercivity of the bilinear form $a_{\mathcal{J}_h}(\cdot;\cdot)$ directly follows from the Cauchy–Schwarz inequality, the estimate (17) and the estimate (30).

Lemma 3. *There exist constants C such that*

$$
\begin{aligned}
a_{\mathcal{J}_h}(\boldsymbol{\tau}_h,\boldsymbol{v}_h;\boldsymbol{\rho}_h,\boldsymbol{w}_h) &\le C\|(\boldsymbol{\tau}_h,\boldsymbol{v}_h)\|_{\mathbf{e}_h}\|(\boldsymbol{\rho}_h,\boldsymbol{w}_h)\|_{\mathbf{e}_h}, & \forall(\boldsymbol{\tau}_h,\boldsymbol{v}_h),(\boldsymbol{\rho}_h,\boldsymbol{w}_h)\in\boldsymbol{\Sigma}_h\times\mathbf{V}_h, & \tag{31} \\
a_{\mathcal{J}_h}(\boldsymbol{\tau}_h,\boldsymbol{v}_h;\boldsymbol{\tau}_h,\boldsymbol{v}_h) &\ge C\|(\boldsymbol{\tau}_h,\boldsymbol{v}_h)\|_{\mathbf{e}_h}^2, & \forall(\boldsymbol{\tau}_h,\boldsymbol{v}_h)\in\boldsymbol{\Sigma}_h^m\times\mathbf{V}_h^m. & \tag{32}
\end{aligned}
$$

Bringing the exact solution $(\boldsymbol{\sigma},\boldsymbol{u})$ into the discrete variational problem (27) yields that

$$
a_{\mathcal{J}_h}(\boldsymbol{\sigma}_h,\boldsymbol{u}_h;\boldsymbol{\tau}_h,\boldsymbol{v}_h) = l_{\mathcal{J}_h}(\boldsymbol{\tau}_h,\boldsymbol{v}_h) = a_{\mathcal{J}_h}(\boldsymbol{\sigma},\boldsymbol{u};\boldsymbol{\tau}_h,\boldsymbol{v}_h) - \mathcal{S}_h^m(\boldsymbol{\sigma},\boldsymbol{\tau}_h) - \mathcal{G}_h^m(\boldsymbol{u},\boldsymbol{v}_h),
$$

which leads to a weakened Galerkin orthogonality property of $a_{\mathcal{J}_h}(\cdot;\cdot)$.

Lemma 4. *Let $(\boldsymbol{\sigma},\boldsymbol{u})\in\boldsymbol{\Sigma}^s\times\mathbf{V}^{s+1}$ be the exact solution to (1), and let $(\boldsymbol{\sigma}_h,\boldsymbol{u}_h)\in\boldsymbol{\Sigma}_h^m\times\mathbf{V}_h^m$ be the numerical solution to (27), there holds*

$$
a_{\mathcal{J}_h}(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h,\boldsymbol{u}-\boldsymbol{u}_h;\boldsymbol{\tau}_h,\boldsymbol{v}_h) = \mathcal{S}_h^m(\boldsymbol{\sigma},\boldsymbol{\tau}_h) + \mathcal{G}_h^m(\boldsymbol{u},\boldsymbol{v}_h). \tag{33}
$$

To complete the error estimation, we present an approximation estimate under the norm $\|\cdot\|_{\mathbf{e}_h}$.

Lemma 5. *Let* $(\boldsymbol{\sigma}, \boldsymbol{u}) \in \boldsymbol{\Sigma}^s \times \mathbf{V}^{s+1}$ *be the exact solution to the problem* (1)*, there exists* $(\boldsymbol{\sigma}_I, \boldsymbol{u}_I) \in \boldsymbol{\Sigma}_h^m \times \mathbf{V}_h^m$ *such that*

$$\|(\boldsymbol{\sigma} - \boldsymbol{\sigma}_I, \boldsymbol{u} - \boldsymbol{u}_I)\|_{\mathbf{e}_h} \le Ch^{t-1/2}\Big(\|\boldsymbol{\sigma}\|_{H^s(\mathrm{div};\Omega_0\cup\Omega_1)} + \|\boldsymbol{u}\|_{H^{s+1}(\Omega_0\cup\Omega_1)}\Big), \qquad t = \min(m, s). \qquad (34)$$

Moreover, if $\boldsymbol{\sigma} \in \mathbb{H}^{s+1}(\Omega_0 \cup \Omega_1)$*, there holds*

$$\|(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h, \boldsymbol{u} - \boldsymbol{u}_I)\|_{\mathbf{e}_h} \le Ch^t\Big(\|\boldsymbol{\sigma}\|_{H^{s+1}(\Omega_0\cup\Omega_1)} + \|\boldsymbol{u}\|_{H^{s+1}(\Omega_0\cup\Omega_1)}\Big), \qquad t = \min(m, s). \qquad (35)$$

Proof. For $i = 0, 1$, we let $\boldsymbol{\sigma}_{I,i} \in \boldsymbol{\Sigma}_{h,i}^m$ be the interpolant of $\pi_i\boldsymbol{\sigma}$ into the space $\boldsymbol{\Sigma}_{h,i}^m$, and let $\boldsymbol{u}_{I,i} \in \mathbf{V}_{h,i}^m$ be the interpolant of $\pi_i\boldsymbol{u}$ into the space $\mathbf{V}_{h,i}^m$. From the approximation properties of finite element spaces [19,51], we have that

$$\|\pi_i\boldsymbol{\sigma} - \boldsymbol{\sigma}_{I,i}\|_{H(\mathrm{div};\Omega_{h,i})} \le Ch^t\|\pi_i\boldsymbol{\sigma}\|_{H^s(\mathrm{div};\Omega_{h,i})} \le Ch^t\|\boldsymbol{\sigma}\|_{H^s(\mathrm{div};\Omega_0\cup\Omega_1)},$$
$$\|\pi_i\boldsymbol{u} - \boldsymbol{u}_{I,i}\|_{H^1(\Omega_{h,i})} \le Ch^t\|\pi_i\boldsymbol{u}\|_{H^{s+1}(\Omega_{h,i})} \le Ch^t\|\boldsymbol{u}\|_{H^{s+1}(\Omega_0\cup\Omega_1)}. \qquad (36)$$

Let $\boldsymbol{\sigma}_I := \boldsymbol{\sigma}_{I,0}\cdot\chi_0 + \boldsymbol{\sigma}_{I,1}\cdot\chi_1$ and $\boldsymbol{u}_I := \boldsymbol{u}_{I,0}\cdot\chi_0 + \boldsymbol{u}_{I,1}\cdot\chi_1$. The errors $\|\boldsymbol{\sigma} - \boldsymbol{\sigma}_I\|_{H(\mathrm{div};\Omega_0\cup\Omega_1)}$ and $\|\boldsymbol{u} - \boldsymbol{u}_I\|_{H^1(\Omega_0\cup\Omega_1)}$ can be bounded directly by (36). The trace term $\|[\![\boldsymbol{\sigma} - \boldsymbol{\sigma}_I]\!]_N\|_{L^2(\Gamma)}$ is split into two parts by the triangle inequality that

$$\|[\![\boldsymbol{\sigma} - \boldsymbol{\sigma}_I]\!]_N\|_{L^2(\Gamma)} \le \|\pi_0\boldsymbol{\sigma} - \boldsymbol{\sigma}_{I,0}\|_{L^2(\Gamma)} + \|\pi_1\boldsymbol{\sigma} - \boldsymbol{\sigma}_{I,1}\|_{L^2(\Gamma)}.$$

Since $s \ge 1$, we are allowed to apply the H^1 trace estimate (13) and the approximation property (36) to find that

$$\|\pi_i\boldsymbol{\sigma} - \boldsymbol{\sigma}_{I,i}\|_{L^2(\Gamma)}^2 = \sum_{K\in\mathcal{T}_h^\Gamma} \|\pi_i\boldsymbol{\sigma} - \boldsymbol{\sigma}_{I,i}\|_{L^2(\Gamma_K)}^2 \le C\sum_{K\in\mathcal{T}_h^\Gamma}\Big(h^{-1}\|\pi_i\boldsymbol{\sigma} - \boldsymbol{\sigma}_{I,i}\|_{L^2(K)}^2 + h\|\pi_i\boldsymbol{\sigma} - \boldsymbol{\sigma}_{I,i}\|_{H^1(K)}^2\Big)$$
$$\le Ch^{2t-1}\|\boldsymbol{\sigma}\|_{H^s(\Omega_0\cup\Omega_1)}^2, \qquad i = 0, 1. \qquad (37)$$

Similarly, we apply the trace estimate (13) to obtain that

$$h^{-1}\|[\![\boldsymbol{u} - \boldsymbol{u}_I]\!]\|_{L^2(\Gamma)}^2 + h\|[\![\nabla_\Gamma(\boldsymbol{u} - \boldsymbol{u}_I)]\!]\|_{L^2(\Gamma)}^2 \le Ch^{2t}\|\boldsymbol{u}\|_{H^{s+1}(\Omega_0\cup\Omega_1)}^2. \qquad (38)$$

Finally, we apply the L^2 extension property (19) and (20) to deduce that

$$|\boldsymbol{\sigma} - \boldsymbol{\sigma}_I|_{\mathcal{S}_h^m} \le C\Big(\|\pi_0\boldsymbol{\sigma} - \boldsymbol{\sigma}_{I,0}\|_{L^2(\Omega_{h,0})} + \|\pi_1\boldsymbol{\sigma} - \boldsymbol{\sigma}_{I,1}\|_{L^2(\Omega_{h,1})}\Big) \le Ch^t\|\boldsymbol{\sigma}\|_{H^s(\Omega_0\cup\Omega_1)}$$
$$|\boldsymbol{u} - \boldsymbol{u}_I|_{\mathcal{G}_h^m} \le C\Big(\|\pi_0\boldsymbol{u} - \boldsymbol{u}_{I,0}\|_{L^2(\Omega_{h,0})} + \|\pi_1\boldsymbol{u} - \boldsymbol{u}_{I,1}\|_{L^2(\Omega_{h,1})}\Big) \le Ch^{t+1}\|\boldsymbol{u}\|_{H^{s+1}(\Omega_0\cup\Omega_1)}.$$

Combining all above estimates yields the desired estimate (34). In addition, if $\boldsymbol{\sigma}$ has a higher regularity that $\boldsymbol{\sigma} \in \mathbb{H}^{s+1}(\Omega_0\cup\Omega_1)$, the upper bound in (37) can be replaced by $Ch^{2t}\|\boldsymbol{\sigma}\|_{H^{s+1}(\Omega_0\cup\Omega_1)}^2$, which leads to the estimate (35). This completes the proof. $\qquad\square$

The error estimation to the numerical solution of (24) follows from Lemmas 3 to 5.

Theorem 1. *Let* $(\boldsymbol{\sigma}, \boldsymbol{u}) \in \boldsymbol{\Sigma}^s \times \mathbf{V}^{s+1}$ *be the exact solution to the problem* (1)*, and let* $(\boldsymbol{\sigma}_h, \boldsymbol{u}_h) \in \boldsymbol{\Sigma}_h^m \times \mathbf{V}_h^m$ *be the numerical solution to the problem* (24)*, there holds*

$$\|(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h, \boldsymbol{u} - \boldsymbol{u}_h)\|_{\mathbf{e}_h} \le Ch^{t-1/2}\Big(\|\boldsymbol{\sigma}\|_{H(\mathrm{div};\Omega_0\cup\Omega_1)} + \|\boldsymbol{u}\|_{H^{s+1}(\Omega_0\cup\Omega_1)}\Big), \qquad t = \min(m, s). \qquad (39)$$

Moreover, if $\boldsymbol{\sigma} \in \mathbb{H}^{s+1}(\Omega_0 \cup \Omega_1)$, there holds

$$\|(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h, \boldsymbol{u} - \boldsymbol{u}_h)\|_{\mathbf{e}_h} \leq Ch^t \Big(\|\boldsymbol{\sigma}\|_{H(\mathrm{div};\Omega_0\cup\Omega_1)} + \|\boldsymbol{u}\|_{H^{s+1}(\Omega_0\cup\Omega_1)}\Big), \qquad t = \min(m, s). \tag{40}$$

Proof. The proof is standard. For any $(\boldsymbol{\tau}_h, \boldsymbol{v}_h) \in \boldsymbol{\Sigma}_h^m \times \mathbf{V}_h^m$, we apply (31)–(33) to find that

$$\begin{aligned}
\|(\boldsymbol{\sigma}_h - \boldsymbol{\tau}_h, \boldsymbol{u}_h - \boldsymbol{v}_h)\|_{\mathbf{e}_h}^2 &\leq Ca_h(\boldsymbol{\sigma}_h - \boldsymbol{\tau}_h, \boldsymbol{u}_h - \boldsymbol{v}_h; \boldsymbol{\sigma}_h - \boldsymbol{\tau}_h, \boldsymbol{u}_h - \boldsymbol{v}_h) \\
&= C(a_h(\boldsymbol{\sigma} - \boldsymbol{\tau}_h, \boldsymbol{u} - \boldsymbol{v}_h; \boldsymbol{\sigma}_h - \boldsymbol{\tau}_h, \boldsymbol{u}_h - \boldsymbol{v}_h) - s_h(\boldsymbol{\sigma}, \boldsymbol{\sigma}_h - \boldsymbol{\tau}_h) - s_h(\boldsymbol{u}, \boldsymbol{u}_h - \boldsymbol{w}_h)) \\
&\leq C\big(\|(\boldsymbol{\sigma} - \boldsymbol{\tau}_h, \boldsymbol{u} - \boldsymbol{v}_h)\|_{\mathbf{e}_h} + |\boldsymbol{\sigma}|_{\mathcal{S}_h^m} + |\boldsymbol{u}|_{\mathcal{G}_h^m}\big)\|(\boldsymbol{\sigma}_h - \boldsymbol{\tau}_h, \boldsymbol{u}_h - \boldsymbol{v}_h)\|_{\mathbf{e}_h} \\
&\leq C\Big(\|(\boldsymbol{\sigma} - \boldsymbol{\tau}_h, \boldsymbol{u} - \boldsymbol{v}_h)\|_{\mathbf{e}_h} + h^s\|\boldsymbol{\sigma}\|_{H^s(\Omega_0\cup\Omega_1)} + h^{s+1}\|\boldsymbol{u}\|_{H^{s+1}(\Omega_0\cup\Omega_1)}\Big) \\
&\quad \times \|(\boldsymbol{\sigma}_h - \boldsymbol{\tau}_h, \boldsymbol{u}_h - \boldsymbol{v}_h)\|_{\mathbf{e}_h}.
\end{aligned}$$

Applying the triangle inequality, we find that

$$\|(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h, \boldsymbol{u} - \boldsymbol{u}_h)\|_{\mathbf{e}_h} \leq C\|(\boldsymbol{\sigma} - \boldsymbol{\tau}_h, \boldsymbol{u} - \boldsymbol{v}_h)\|_{\mathbf{e}_h} + Ch^s\Big(\|\boldsymbol{\sigma}\|_{H^s(\mathrm{div};\Omega_0\cup\Omega_1)} + \|\boldsymbol{u}\|_{H^{s+1}(\Omega_0\cup\Omega_1)}\Big).$$

Since $(\boldsymbol{\tau}_h, \boldsymbol{v}_h)$ is arbitrary, the error estimate under the norm $\|\cdot\|_{\mathbf{e}_h}$ follows from the approximation properties in Lemma 5. This completes the proof. $\qquad\square$

We estimate the condition number for the linear system (27), which is especially desired in the unfitted method.

Theorem 2. *There exists a constant C such that*

$$\kappa(A_{\mathcal{J}_h}) \leq Ch^{-2}, \tag{41}$$

where $A_{\mathcal{J}_h}$ is the linear system with respect to $a_{\mathcal{J}_h}(\cdot; \cdot)$.

Proof. Since the bilinear form $a_{\mathcal{J}_h}(\cdot; \cdot)$ is bounded and coercive with respect to the norm $\|\cdot\|_{\mathbf{e}_h}$. From Section 3.2 of [24], the main step is to show the relationship between the energy norm $\|\cdot\|_{\mathbf{e}_h}$ and the L^2 norm. Note that the finite element spaces $\boldsymbol{\Sigma}_h^m$ and $\mathbf{V}_h^m$ are the combinations of $\boldsymbol{\Sigma}_{h,0}^m$ and $\boldsymbol{\Sigma}_{h,1}^m$ and of $\mathbf{V}_{h,0}^m$ and $\mathbf{V}_{h,1}^m$, respectively, where the spaces $\boldsymbol{\Sigma}_{h,i}^m$ and $\mathbf{V}_{h,i}^m$ are defined on the domain $\Omega_{h,i}$. Hence, our goal is to show that

$$\begin{aligned}
\sum_{i=0,1} \Big(\|\pi_i \boldsymbol{\tau}_h\|_{L^2(\Omega_{h,i})}^2 &+ \|\pi_i \boldsymbol{v}_h\|_{L^2(\Omega_{h,i})}^2\Big) \leq C\|(\boldsymbol{\tau}_h, \boldsymbol{v}_h)\|_{\mathbf{e}_h}^2 \\
&\leq Ch^{-2} \sum_{i=0,1} \Big(\|\pi_i \boldsymbol{\tau}_h\|_{L^2(\Omega_{h,i})}^2 + \|\pi_i \boldsymbol{v}_h\|_{L^2(\Omega_{h,i})}^2\Big), \qquad \forall (\boldsymbol{\tau}_h, \boldsymbol{v}_h) \in \boldsymbol{\Sigma}_h^m \times \mathbf{V}_h^m. \tag{42}
\end{aligned}$$

Since $\|\cdot\|_{\mathbf{e}_h}$ is stronger than $\|\cdot\|_{\mathbf{e}}$, we have that

$$\|\boldsymbol{\tau}_h\|_{L^2(\Omega_0\cup\Omega_1)} + \|\boldsymbol{v}_h\|_{L^2(\Omega_0\cup\Omega_1)} \leq C\|(\boldsymbol{\tau}_h, \boldsymbol{v}_h)\|_{\mathbf{e}} \leq C\|(\boldsymbol{\tau}_h, \boldsymbol{v}_h)\|_{\mathbf{e}_h}.$$

The lower bound in (42) then follows from the estimates (19) and (20). For the upper bound, we apply the triangle inequality and the inverse estimate to find that

$$\|(\boldsymbol{\tau}_h, \boldsymbol{v}_h)\|_{\mathbf{e}_h}^2 \leq C \sum_{i=0,1} \Big(\|\pi_i \boldsymbol{\tau}_h\|_{H^1(\Omega_{h,i})}^2 + \|\pi_i \boldsymbol{v}_h\|_{H^1(\Omega_{h,i})}^2\Big) \leq Ch^{-2} \sum_{i=0,1} \Big(\|\pi_i \boldsymbol{\tau}_h\|_{L^2(\Omega_{h,i})}^2 + \|\pi_i \boldsymbol{v}_h\|_{L^2(\Omega_{h,i})}^2\Big).$$

From Corollary 3.4 of [24] and Section 2.6 of [30], the estimate (41) comes from (42), which completes the proof. $\qquad\square$

We have completed the error analysis for the discrete variational form (27). The scheme is robust in the sense that the constants appearing in the error bounds (39) and (40) are independent of λ, and also are independent of how the interface Γ cuts the mesh. From Theorem 1, the convergence rate of the numerical error under the energy norm is half order lower than the optimal rate without the higher regularity assumption. The major reason is that we use the stronger L^2 norm to replace the minus norm $\|\cdot\|_{H^{-1/2}(\Gamma)}$ to define the new quadratic functional, and the errors on the interface are essentially established by the H^1 trace estimate as (37) and (38). In addition, both estimates require the exact solution $(\boldsymbol{\sigma}, \boldsymbol{u})$ has at least $H^1(\mathrm{div}) \times H^2$ regularity. Consequently, this method and the analysis cannot be extended to the case of low regularity that $s < 1$.

4.2. Least squares finite element method with the discrete minus norm for $s > 1/2$

In this subsection, we give the numerical method that has the optimal convergence speed for $s > 1/2$. As stated before, the replacement of the L^2 norm cannot work and the H^1 trace estimate is also inappropriate for the case of low regularity. The natural choice is to involve the $\|\cdot\|_{H^{-1/2}(\Gamma)}$ in the discrete least squares functional, but the minus norm is not easy to compute. Here we follow the idea in [7, 8] to employ a discrete inner product, which is related to the minus norm $\|\cdot\|_{H^{-1/2}(\Gamma)}$. We first give an inner product in $\mathbf{H}^{-1/2}(\Gamma)$.

For any $\boldsymbol{v} \in \mathbf{H}^{-1/2}(\Gamma)$, consider the elliptic problem $\boldsymbol{w} - \Delta \boldsymbol{w} = \mathbf{0}$ in Ω_0 with $\partial_{\boldsymbol{n}} \boldsymbol{w} = \boldsymbol{v}$ on $\Gamma = \partial \Omega_0$. The corresponding weak form reads

$$(\nabla \boldsymbol{w}, \nabla \boldsymbol{t})_{L^2(\Omega_0)} + (\boldsymbol{w}, \boldsymbol{t})_{L^2(\Omega_0)} = (\boldsymbol{v}, \boldsymbol{t})_{L^2(\Gamma)}, \qquad \forall \boldsymbol{t} \in \mathbf{H}^1(\Omega_0). \tag{43}$$

The elliptic regularity theory indicates that the problem (43) admits a unique solution $\boldsymbol{w} \in \mathbf{H}^1(\Omega_0)$ with $\|\boldsymbol{w}\|_{H^1(\Omega_0)} \leq C\|\boldsymbol{v}\|_{H^{-1/2}(\Gamma)}$. This fact allows us to define an operator $T : \mathbf{H}^{-1/2}(\Gamma) \to \mathbf{H}^1(\Omega_0)$ such that $T\boldsymbol{v}$ is the solution to (43) for $\forall \boldsymbol{v} \in \mathbf{H}^{-1/2}(\Gamma)$. From T and (43), we define an inner product in $\mathbf{H}^{-1/2}(\Gamma)$ as

$$(\boldsymbol{v}, \boldsymbol{q})_{-1/2,\Gamma} := (\boldsymbol{v}, T\boldsymbol{q})_{L^2(\Gamma)}, \qquad \forall \boldsymbol{v}, \boldsymbol{q} \in H^{-1/2}(\Gamma), \tag{44}$$

with the induced norm $\|\boldsymbol{v}\|_{-1/2,\Gamma}^2 := (\boldsymbol{v}, \boldsymbol{v})_{-1/2,\Gamma}$. We then show that $\|\cdot\|_{-1/2,\Gamma}$ and $\|\cdot\|_{H^{-1/2}(\Gamma)}$ are equivalent. For $\forall \boldsymbol{v} \in \mathbf{H}^{-1/2}(\Gamma)$, we let $\boldsymbol{t} = T\boldsymbol{v}$ in (43), which directly gives that $\|T\boldsymbol{v}\|_{H^1(\Omega_0)} = \|\boldsymbol{v}\|_{-1/2,\Gamma}$. From the trace estimate $\|\boldsymbol{w}\|_{H^{1/2}(\Gamma)} \leq C\|\boldsymbol{w}\|_{H^1(\Omega_0)}(\forall \boldsymbol{w} \in \mathbf{H}^1(\Omega_0))$, we derive that

$$\|\boldsymbol{v}\|_{-1/2,\Gamma} = \frac{(\boldsymbol{v}, T\boldsymbol{v})_{L^2(\Gamma)}}{\|T\boldsymbol{v}\|_{H^1(\Omega_0)}} \leq C\frac{(\boldsymbol{v}, T\boldsymbol{v})_{L^2(\Gamma)}}{\|T\boldsymbol{v}\|_{H^{1/2}(\Gamma)}} \leq C\|\boldsymbol{v}\|_{H^{-1/2}(\Gamma)}, \qquad \forall \boldsymbol{v} \in \mathbf{H}^{-1/2}(\Gamma).$$

For any $\boldsymbol{z} \in \mathbf{H}^{1/2}(\Gamma)$, there exists $\boldsymbol{\varphi_z} \in \mathbf{H}^1(\Omega_0)$ such that $\boldsymbol{\varphi_z}|_\Gamma = \boldsymbol{z}$ and $\|\boldsymbol{\varphi_z}\|_{H^1(\Omega_0)} \leq C\|\boldsymbol{z}\|_{H^{1/2}(\Gamma)}$. Given $\boldsymbol{v} \in \mathbf{H}^{-1/2}(\Gamma)$, we find that

$$\begin{aligned}
(\boldsymbol{v}, \boldsymbol{z})_{L^2(\Gamma)} &= (\boldsymbol{v}, \boldsymbol{\varphi_z})_{L^2(\Gamma)} = (\nabla T\boldsymbol{v}, \nabla \boldsymbol{\varphi_z})_{L^2(\Omega_0)} + (T\boldsymbol{v}, \boldsymbol{\varphi_z})_{L^2(\Omega_0)} \\
&\leq \|T\boldsymbol{v}\|_{H^1(\Omega_0)}\|\boldsymbol{\varphi_z}\|_{H^1(\Omega_0)} = C\|\boldsymbol{v}\|_{-1/2,\Gamma}\|\boldsymbol{z}\|_{H^{1/2}(\Gamma)}, \qquad \forall \boldsymbol{z} \in \mathbf{H}^{1/2}(\Gamma),
\end{aligned}$$

which indicates that $\|\boldsymbol{v}\|_{H^{-1/2}(\Gamma)} \leq C\|\boldsymbol{v}\|_{-1/2,\Gamma}$. The equivalence between both norms is reached.

Roughly speaking, we have defined an equivalent norm and an inner product for the space $\mathbf{H}^{-1/2}(\Gamma)$, although the operator T and the inner product $(\cdot, \cdot)_{-1/2,\Gamma}$ still cannot be directly computed. In the numerical scheme, we introduce a discrete operator T_h to replace T to make the method computationally feasible. The discrete operator T_h is established with the space $\mathbf{V}_{h,0}^1$, which is the C^0 linear finite element space defined on the partition $\mathcal{T}_{h,0}$. Given any $\boldsymbol{v} \in \mathbf{H}^{-1/2}(\Gamma)$, we define the following discrete weak form: seek $\boldsymbol{w}_h \in \mathbf{V}_{h,0}^1$ such that

$$(\nabla \boldsymbol{w}_h, \nabla \boldsymbol{t}_h)_{L^2(\Omega_0)} + (\boldsymbol{w}_h, \boldsymbol{t}_h)_{L^2(\Omega_0)} + s_{h,0}^1(\boldsymbol{w}_h, \boldsymbol{t}_h) = (\boldsymbol{v}, \boldsymbol{t}_h)_{L^2(\Gamma)}, \qquad \forall \boldsymbol{t}_h \in \mathbf{V}_{h,0}^1. \tag{45}$$

Remark 2. The discrete variational form (45) can be regarded as an unfitted scheme for the elliptic system (43) based on the mesh $\mathcal{T}_{h,0}$. The ghost penalty form $s_{h,0}^1(\cdot, \cdot)$ ensures the uniform upper bound of the condition number to the corresponding linear system. From Theorem 2.16 of [30], the upper bound is $O(h^{-2})$ for (45).

It can be easily seen that the problem (45) admits a unique solution in $\mathbf{V}_{h,0}^1$. Similarly, this property allows us to define an operator $T_h : \mathbf{H}^{-1/2}(\Gamma) \to \mathbf{V}_{h,0}^1$ that $T_h \boldsymbol{v}$ is the solution to the discrete problem (45) for $\forall \boldsymbol{v} \in \mathbf{H}^{-1/2}(\Gamma)$. As (44), we define the discrete inner product $(\cdot, \cdot)_{-1/2,h}$ and the induced norm $\| \cdot \|_{-1/2,h}$ as

$$(\boldsymbol{v}, \boldsymbol{q})_{-1/2,h} := (\boldsymbol{v}, T_h \boldsymbol{q})_{L^2(\Gamma)}, \quad \|\boldsymbol{v}\|_{-1/2,h}^2 := (\boldsymbol{v}, \boldsymbol{v})_{-1/2,h}, \qquad \forall \boldsymbol{v}, \boldsymbol{q} \in \mathbf{H}^{-1/2}(\Gamma). \tag{46}$$

Let $\boldsymbol{t}_h = \boldsymbol{w}_h$ in (45), together with the property (8), it can be seen that

$$\|\boldsymbol{v}\|_{-1/2,h}^2 = (\boldsymbol{v}, T_h \boldsymbol{v})_{-1/2,h} = \|T_h \boldsymbol{v}\|_{H^1(\Omega_0)}^2 + |T_h \boldsymbol{v}|_{s_{h,0}^1}^2 \geq \|T_h \boldsymbol{v}\|_{H^1(\Omega_0)}^2, \tag{47}$$

and

$$\|\boldsymbol{v}\|_{-1/2,h}^2 = \|T_h \boldsymbol{v}\|_{H^1(\Omega_0)}^2 + |T_h \boldsymbol{v}|_{s_{h,0}^1}^2 \leq C \|T_h \boldsymbol{v}\|_{H^1(\Omega_{h,0})}^2. \tag{48}$$

Then, we give the relationship between T and T_h.

Lemma 6. *There exist constants C such that*

$$\|\boldsymbol{v}\|_{-1/2,h} \leq C \|\boldsymbol{v}\|_{-1/2,\Gamma}, \qquad \forall \boldsymbol{v} \in \mathbf{H}^{-1/2}(\Gamma), \tag{49}$$

and

$$\|\boldsymbol{v}\|_{-1/2,\Gamma} \leq C(\|\boldsymbol{v}\|_{-1/2,h} + h^{1/2} \|\boldsymbol{v}\|_{L^2(\Gamma)}), \qquad \forall \boldsymbol{v} \in \mathbf{L}^2(\Gamma). \tag{50}$$

Proof. For $\forall \boldsymbol{v} \in \mathbf{H}^{-1/2}(\Gamma)$, from (43) and (47), we have that

$$\|\boldsymbol{v}\|_{-1/2,h} = \frac{(\boldsymbol{v}, T_h \boldsymbol{v})_{L^2(\Gamma)}}{\|\boldsymbol{v}\|_{-1/2,h}} \leq \frac{(\boldsymbol{v}, T_h \boldsymbol{v})_{L^2(\Gamma)}}{\|T_h \boldsymbol{v}\|_{H^1(\Omega_0)}} \leq \sup_{\boldsymbol{t} \in \mathbf{H}^1(\Omega_0)} \frac{(\boldsymbol{v}, \boldsymbol{t})_{L^2(\Gamma)}}{\|\boldsymbol{t}\|_{H^1(\Omega_0)}} \leq \|T\boldsymbol{v}\|_{H^1(\Omega_0)} \leq \|\boldsymbol{v}\|_{-1/2,\Gamma},$$

which brings us the estimate (49).

From the Sobolev extension theory [1], there exists a linear extension operator $E_0 : \mathbf{H}^1(\Omega_0) \to \mathbf{H}^1(\Omega)$ such that $(E^0 \boldsymbol{w})|_{\Omega_0} = \boldsymbol{w}$, $\|E^0 \boldsymbol{w}\|_{H^1(\Omega)} \leq C \|\boldsymbol{w}\|_{H^1(\Omega_0)}$ for $\forall \boldsymbol{w} \in \mathbf{H}^1(\Omega_0)$. Let $I_{h,0}^1$ be the Scott-Zhang interpolation operator of the space $\mathbf{V}_{h,0}^1$, which satisfies the approximation estimates

$$\left\| \boldsymbol{w} - I_{h,0}^1 \boldsymbol{w} \right\|_{H^q(\Omega_{h,0})} \leq C h^{1-q} \|\boldsymbol{w}\|_{H^1(\Omega_{h,0})}, \quad q = 0, 1, \qquad \forall \boldsymbol{w} \in \mathbf{H}^1(\Omega_{h,0}).$$

From the trace estimate (13), it is quite standard to derive that $\|\boldsymbol{w} - I_{h,0}^1 \boldsymbol{w}\|_{L^2(\Gamma)} \leq C h^{1/2} \|\boldsymbol{w}\|_{H^1(\Omega_{h,0})}$ for $\forall \boldsymbol{w} \in H^1(\Omega_{h,0})$. Then, we deduce that

$$\begin{aligned}
\|\boldsymbol{v}\|_{-1/2,\Gamma}^2 = (\boldsymbol{v}, T\boldsymbol{v})_{L^2(\Gamma)} &= \left(\boldsymbol{v}, E^0(T\boldsymbol{v}) \right)_{L^2(\Gamma)} \\
&= \left(\boldsymbol{v}, I_{h,0}^1 E^0(T\boldsymbol{v}) \right)_{L^2(\Gamma)} + \left(\boldsymbol{v}, E^0(T\boldsymbol{v}) - I_{h,0}^1 \left(E^0(T\boldsymbol{v}) \right) \right)_{L^2(\Gamma)}.
\end{aligned}$$

The second term can be bounded by the approximation property, that is

$$\begin{aligned}
\left(\boldsymbol{v}, E^0(T\boldsymbol{v}) - I_{h,0} \left(E^0(T\boldsymbol{v}) \right) \right)_{L^2(\Gamma)} &\leq C h^{1/2} \|\boldsymbol{v}\|_{L^2(\Gamma)} \left\| E^0(T\boldsymbol{v}) \right\|_{H^1(\Omega_{h,0})} \leq C h^{1/2} \|\boldsymbol{v}\|_{L^2(\Gamma)} \|T\boldsymbol{v}\|_{H^1(\Omega_0)} \\
&\leq C h^{1/2} \|\boldsymbol{v}\|_{L^2(\Gamma)} \|\boldsymbol{v}\|_{-1/2,\Gamma}.
\end{aligned}$$

We let $\boldsymbol{t}_h = I_{h,0}^1(E^0(T\boldsymbol{v}))$ in (45) and apply (48) to bound the first term,

$$
\begin{aligned}
\left(\boldsymbol{v}, I_{h,0}^1\big(E^0(T\boldsymbol{v})\big)\right)_{L^2(\Gamma)} &\leq C\Big(\|T_h\boldsymbol{v}\|_{H^1(\Omega_0)} + |T_h\boldsymbol{v}|_{s_{h,0}^1}\Big)\Big(\big\|I_{h,0}^1\big(E^0(T\boldsymbol{v})\big)\big\|_{H^1(\Omega_0)} + |I_{h,0}^1\big(E^0(T\boldsymbol{v})\big)|_{s_{h,0}^1}\Big) \\
&\leq C\|\boldsymbol{v}\|_{-1/2,h}\big\|I_{h,0}^1\big(E^0(T\boldsymbol{v})\big)\big\|_{H^1(\Omega_{h,0})} \leq C\|\boldsymbol{v}\|_{-1/2,h}\big\|E^0(T\boldsymbol{v})\big\|_{H^1(\Omega_{h,0})} \\
&\leq C\|\boldsymbol{v}\|_{-1/2,h}\big\|E^0(T\boldsymbol{v})\big\|_{H^1(\Omega)} \leq C\|\boldsymbol{v}\|_{-1/2,h}\|T\boldsymbol{v}\|_{H^1(\Omega_0)} \leq C\|\boldsymbol{v}\|_{-1/2,h}\|\boldsymbol{v}\|_{-1/2,\Gamma}.
\end{aligned}
$$

Collecting all above estimates yields the second estimate (50), which completes the proof. $\qquad\square$

Now, let us define the least squares functional $\widetilde{\mathcal{J}}_h(\cdot;\cdot)$ by

$$
\widetilde{\mathcal{J}}_h(\boldsymbol{\tau}_h, \boldsymbol{v}_h; \boldsymbol{f}, \boldsymbol{a}, \boldsymbol{b}) := J(\boldsymbol{\tau}_h, \boldsymbol{v}_h; \boldsymbol{f}) + \widetilde{B}_h^{\boldsymbol{\sigma}}(\boldsymbol{\tau}_h; \boldsymbol{a}) + B_h^{\boldsymbol{u}}(\boldsymbol{v}_h; \boldsymbol{b}) + |\boldsymbol{\tau}_h|_{\mathcal{S}_h^m}^2 + |\boldsymbol{v}_h|_{\mathcal{G}_h^m}^2, \quad \forall(\boldsymbol{\tau}_h, \boldsymbol{v}_h) \in \boldsymbol{\Sigma}_h \times \mathbf{V}_h \quad (51)
$$

where $B_h^{\boldsymbol{u}}(\cdot;\cdot)$ is defined as in (25) and $\widetilde{B}_h^{\boldsymbol{\sigma}}(\cdot;\cdot)$ is defined as

$$
\widetilde{B}_h^{\boldsymbol{\sigma}}(\boldsymbol{\tau}_h; \boldsymbol{a}) := \|[\![\boldsymbol{\tau}_h]\!]_N - \boldsymbol{a}\|_{-1/2,h}^2 + h\|[\![\boldsymbol{\tau}_h]\!]_N - \boldsymbol{a}\|_{L^2(\Gamma)}^2, \qquad \forall \boldsymbol{\tau}_h \in \boldsymbol{\Sigma}_h. \quad (52)
$$

The numerical solutions are sought by minimizing the functional $\widetilde{\mathcal{J}}_h(\cdot;\cdot)$ over the finite element spaces $\boldsymbol{\Sigma}_h^m \times \mathbf{V}_h^m$. This minimization problem is equivalent to a variational problem by writing the Euler–Lagrange equation, which reads: seek $(\boldsymbol{\sigma}_h, \boldsymbol{u}_h) \in \boldsymbol{\Sigma}_h^m \times \mathbf{V}_h^m$ such that

$$
a_{\widetilde{\mathcal{J}}_h}(\boldsymbol{\sigma}_h, \boldsymbol{u}_h; \boldsymbol{\tau}_h, \boldsymbol{v}_h) = l_{\widetilde{\mathcal{J}}_h}(\boldsymbol{\tau}_h, \boldsymbol{v}_h), \qquad \forall(\boldsymbol{\tau}_h, \boldsymbol{v}_h) \in \boldsymbol{\Sigma}_h^m \times \mathbf{V}_h^m, \quad (53)
$$

where the forms $a_{\widetilde{\mathcal{J}}_h}(\cdot;\cdot)$ and $l_{\widetilde{\mathcal{J}}_h}(\cdot)$ are defined as

$$
\begin{aligned}
a_{\widetilde{\mathcal{J}}_h}(\boldsymbol{\tau}_h, \boldsymbol{v}_h; \boldsymbol{\rho}_h, \boldsymbol{w}_h) := &\ (\mathcal{A}\boldsymbol{\tau}_h - \boldsymbol{\varepsilon}(\boldsymbol{v}_h), \mathcal{A}\boldsymbol{\rho}_h - \boldsymbol{\varepsilon}(\boldsymbol{w}_h))_{L^2(\Omega_0 \cup \Omega_1)} + (\nabla \cdot \boldsymbol{\tau}_h, \nabla \cdot \boldsymbol{\rho}_h)_{L^2(\Omega_0 \cup \Omega_1)} \\
&+ ([\![\boldsymbol{\tau}_h]\!]_N, [\![\boldsymbol{\rho}_h]\!]_N)_{-1/2,h} + h([\![\boldsymbol{\tau}_h]\!]_N, [\![\boldsymbol{\rho}_h]\!]_N)_{L^2(\Gamma)} + h^{-1}([\![\boldsymbol{v}_h]\!], [\![\boldsymbol{w}_h]\!])_{L^2(\Gamma)} \\
&+ h([\![\nabla_\Gamma \boldsymbol{v}_h]\!], [\![\nabla_\Gamma \boldsymbol{w}_h]\!])_{L^2(\Gamma)} + \mathcal{S}_h^m(\boldsymbol{\tau}_h, \boldsymbol{\rho}_h) + \mathcal{G}_h^m(\boldsymbol{v}_h, \boldsymbol{w}_h),
\end{aligned}
$$

and

$$
\begin{aligned}
l_{\widetilde{\mathcal{J}}_h}(\boldsymbol{\tau}_h, \boldsymbol{v}_h) := &-(\nabla \cdot \boldsymbol{\tau}_h, \boldsymbol{f})_{L^2(\Omega_0 \cup \Omega_1)} + ([\![\boldsymbol{\tau}_h]\!]_N, \boldsymbol{a})_{-1/2,h} + h([\![\boldsymbol{\tau}_h]\!]_N, \boldsymbol{a})_{L^2(\Gamma)} \\
&+ h^{-1}([\![\boldsymbol{v}_h]\!], \boldsymbol{b})_{L^2(\Gamma)} + h([\![\nabla_\Gamma \boldsymbol{v}_h]\!], \nabla_\Gamma \boldsymbol{b})_{L^2(\Gamma)}.
\end{aligned}
$$

The convergence analysis of the solution in (53) is analogous to the discrete variational form (27). We introduce the energy norm $\|\cdot\|_{\widetilde{\mathbf{e}}_h}$ by

$$
\begin{aligned}
\|(\boldsymbol{\tau}_h, \boldsymbol{v}_h)\|_{\widetilde{\mathbf{e}}_h}^2 := &\ \|\boldsymbol{\tau}_h\|_{H(\mathrm{div};\Omega_0 \cup \Omega_1)}^2 + \|\boldsymbol{v}_h\|_{H^1(\Omega_0 \cup \Omega_1)}^2 + \|[\![\boldsymbol{\tau}_h]\!]\|_{-1/2,h}^2 + h\|[\![\boldsymbol{\tau}_h]\!]\|_{L^2(\Gamma)}^2 \\
&+ h^{-1}\|[\![\boldsymbol{v}_h]\!]\|_{L^2(\Gamma)}^2 + h\|[\![\nabla_\Gamma \boldsymbol{v}_h]\!]\|_{L^2(\Gamma)}^2 + |\boldsymbol{\tau}_h|_{\mathcal{S}_h^m}^2 + |\boldsymbol{v}_h|_{\mathcal{G}_h^m}^2, \qquad \forall(\boldsymbol{\tau}_h, \boldsymbol{v}_h) \in \boldsymbol{\Sigma}_h \times \mathbf{V}_h.
\end{aligned}
$$

From Lemma 6, there holds

$$
B^{\boldsymbol{\sigma}}(\boldsymbol{\tau}_h; \mathbf{0}) = \|[\![\boldsymbol{\tau}_h]\!]\|_{H^{-1/2}(\Gamma)}^2 \leq C\Big(\|[\![\boldsymbol{\tau}_h]\!]\|_{-1/2,h}^2 + h\|[\![\boldsymbol{\tau}_h]\!]\|_{L^2(\Gamma)}^2\Big) = \widetilde{B}_h^{\boldsymbol{\sigma}}(\boldsymbol{\tau}_h; \mathbf{0}), \qquad \forall \boldsymbol{\tau}_h \in \boldsymbol{\Sigma}_h.
$$

Together with the estimate (29), we can know that

$$
\mathcal{J}(\boldsymbol{\tau}_h, \boldsymbol{v}_h; \mathbf{0}, \mathbf{0}, \mathbf{0}) \leq C\widetilde{\mathcal{J}}_h(\boldsymbol{\tau}_h, \boldsymbol{v}_h; \mathbf{0}, \mathbf{0}, \mathbf{0}) = Ca_{\widetilde{\mathcal{J}}_h}(\boldsymbol{\tau}_h, \boldsymbol{v}_h; \boldsymbol{\tau}_h, \boldsymbol{v}_h), \qquad \forall(\boldsymbol{\tau}_h, \boldsymbol{v}_h) \in \boldsymbol{\Sigma}_h \times \mathbf{V}_h.
$$

As Lemma 3, it is similar to get that $a_{\widetilde{\mathcal{J}}_h}(\cdot;\cdot)$ is bounded and coercive under the energy norm $\|\cdot\|_{\widetilde{\mathbf{e}}_h}$, and satisfies a weakened Galerkin orthogonality property.

Lemma 7. *There exist constants C such that*

$$a_{\widetilde{\mathcal{J}}_h}(\boldsymbol{\tau}_h, \boldsymbol{v}_h; \boldsymbol{\rho}_h, \boldsymbol{w}_h) \leq C\|(\boldsymbol{\tau}_h, \boldsymbol{v}_h)\|_{\widetilde{\mathbf{e}}_h}\|(\boldsymbol{\rho}_h, \boldsymbol{w}_h)\|_{\widetilde{\mathbf{e}}_h}, \qquad \forall (\boldsymbol{\tau}_h, \boldsymbol{v}_h), (\boldsymbol{\rho}_h, \boldsymbol{w}_h) \in \boldsymbol{\Sigma}_h \times \mathbf{V}_h, \qquad (54)$$

$$a_{\widetilde{\mathcal{J}}_h}(\boldsymbol{\tau}_h, \boldsymbol{v}_h; \boldsymbol{\tau}_h, \boldsymbol{v}_h) \geq C\|(\boldsymbol{\tau}_h, \boldsymbol{v}_h)\|_{\widetilde{\mathbf{e}}_h}^2, \qquad \forall (\boldsymbol{\tau}_h, \boldsymbol{v}_h) \in \boldsymbol{\Sigma}_h^m \times \mathbf{V}_h^m. \qquad (55)$$

Lemma 8. *Let $(\boldsymbol{\sigma}, \boldsymbol{u}) \in \boldsymbol{\Sigma}^s \times \mathbf{V}^{s+1}$ be the exact solution to* (1)*, let $(\boldsymbol{\sigma}_h, \boldsymbol{u}_h) \in \boldsymbol{\Sigma}_h^m \times \mathbf{V}_h^m$ be the numerical solution to* (27)*, there holds*

$$a_{\widetilde{\mathcal{J}}_h}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h, \boldsymbol{u} - \boldsymbol{u}_h; \boldsymbol{\tau}_h, \boldsymbol{v}_h) = \mathcal{S}_h^m(\boldsymbol{\sigma}, \boldsymbol{\tau}_h) + \mathcal{G}_h^m(\boldsymbol{u}, \boldsymbol{v}_h). \qquad (56)$$

Proof. The proofs are similar to Lemmas 3 and 4. □

Moreover, we give the approximation estimate under the norm $\|\cdot\|_{\widetilde{\mathbf{e}}_h}$.

Lemma 9. *Let $(\boldsymbol{\sigma}, \boldsymbol{u}) \in \boldsymbol{\Sigma}^s \times \mathbf{V}^{s+1}$ be the exact solution to the interface problem* (1)*, there exists $(\boldsymbol{\sigma}_I, \boldsymbol{u}_I) \in \boldsymbol{\Sigma}_h^m \times \mathbf{V}_h^m$ such that*

$$\|(\boldsymbol{\sigma} - \boldsymbol{\sigma}_I, \boldsymbol{u} - \boldsymbol{u}_I)\|_{\widetilde{\mathbf{e}}_h} \leq Ch^t\Big(\|\boldsymbol{\sigma}\|_{H^s(\mathrm{div};\Omega_0 \cup \Omega_1)} + \|\boldsymbol{u}\|_{H^{s+1}(\Omega_0 \cup \Omega_1)}\Big), \qquad (57)$$

where $t = \min(m, s)$ for $s \geq 1$, and $t = \min(m, s) - \varepsilon$ with any $\varepsilon > 0$ for $1/2 < s < 1$.

Proof. We let $\boldsymbol{\sigma}_I := \boldsymbol{\sigma}_{I,0} \cdot \chi_0 + \boldsymbol{\sigma}_{I,1} \cdot \chi_1$, $\boldsymbol{u}_I := \boldsymbol{u}_{I,0} \cdot \chi_0 + \boldsymbol{u}_{I,0} \cdot \chi_1$ be the interpolants of $\boldsymbol{\sigma}$ and $\boldsymbol{u}$, which are defined in the proof in Lemma 5. The errors $\|\boldsymbol{\sigma} - \boldsymbol{\sigma}_I\|_{H(\mathrm{div};\Omega_0 \cup \Omega_1)}$, $\|\boldsymbol{u} - \boldsymbol{u}_I\|_{H^1(\Omega_0 \cup \Omega_1)}$, $|\boldsymbol{\sigma}_I|_{\mathcal{S}_h^m}$ and $|\boldsymbol{u}_I|_{\mathcal{G}_h^m}$ have been estimated in Lemma 5. We only bound the trace terms in $\|\cdot\|_{\widetilde{\mathbf{e}}_h}$. For the error $\|[\boldsymbol{\sigma} - \boldsymbol{\sigma}_I]_N\|_{-1/2,h}$, we apply the estimate (49), the approximation properties (36) and the trace theory of functions in $\mathbf{H}(\mathrm{div};\Omega_0 \cup \Omega_1)$ to find that

$$\|[\boldsymbol{\sigma} - \boldsymbol{\sigma}_I]_N\|_{-1/2,h} \leq C\|[\boldsymbol{\sigma} - \boldsymbol{\sigma}_I]_N\|_{-1/2,\Gamma} \leq C\|[\boldsymbol{\sigma} - \boldsymbol{\sigma}_I]_N\|_{H^{-1/2}(\Gamma)} \leq C \sum_{i=0,1} \|\boldsymbol{n} \cdot (\pi_i \boldsymbol{\sigma} - \pi_i \boldsymbol{\sigma}_I)\|_{H^{-1/2}(\Gamma)}$$

$$\leq C \sum_{i=0,1} \|\pi_i \boldsymbol{\sigma} - \pi_i \boldsymbol{\sigma}_I\|_{H(\mathrm{div};\Omega_i)} \leq Ch^{t_0}\|\boldsymbol{\sigma}\|_{H(\mathrm{div};\Omega_0 \cup \Omega_1)},$$

where $t_0 := \min(m, s)$. Since $\boldsymbol{u} \in \mathbf{H}^{1+s}(\Omega_0 \cup \Omega_1)$, the error $h^{-1}\|[\boldsymbol{u} - \boldsymbol{u}_I]\|_{L^2(\Gamma)}^2$ can be directly estimated by the H^1 trace estimate (13), *i.e.*

$$h^{-1}\|[\boldsymbol{u} - \boldsymbol{u}_I]\|_{L^2(\Gamma)}^2 \leq Ch^{2t_0}\|\boldsymbol{u}\|_{H^{s+1}(\Omega_0 \cup \Omega_1)}^2,$$

which is the same as in the proof to Lemma 5.

The rest is to bound the terms $h\|[\boldsymbol{\sigma} - \boldsymbol{\sigma}_I]_N\|_{L^2(\Gamma)}^2$ and $h\|[\nabla_\Gamma(\boldsymbol{u} - \boldsymbol{u}_I)]\|_{L^2(\Gamma)}^2$. For the case $s \geq 1$, both errors can be bounded by the H^1 trace estimate, and we derive that

$$h\|[\boldsymbol{\sigma} - \boldsymbol{\sigma}_I]_N\|_{L^2(\Gamma)}^2 \leq Ch \sum_{i=0,1} \|\pi_i \boldsymbol{\sigma} - \pi_i \boldsymbol{\sigma}_I\|_{L^2(\Gamma)}^2 \leq Ch \sum_{i=0,1} \sum_{K \in \mathcal{T}_h^\Gamma} \|\pi_i \boldsymbol{\sigma} - \pi_i \boldsymbol{\sigma}_I\|_{L^2(\Gamma_K)}^2$$

$$\leq C \sum_{i=0,1} \sum_{K \in \mathcal{T}_h^\Gamma} \Big(\|\pi_i \boldsymbol{\sigma} - \pi_i \boldsymbol{\sigma}_I\|_{L^2(K)}^2 + h^2\|\pi_i \boldsymbol{\sigma} - \pi_i \boldsymbol{\sigma}_I\|_{H^1(K)}^2\Big) \leq Ch^{2t_0}\|\boldsymbol{\sigma}\|_{H^s(\Omega_0 \cup \Omega_1)}, \quad (58)$$

and similarly, there holds

$$h\|[\nabla_\Gamma(\boldsymbol{u} - \boldsymbol{u}_I)]\|_{L^2(\Gamma)}^2 \leq Ch^{2t_0}\|\boldsymbol{u}\|_{H^{s+1}(\Omega_0 \cup \Omega_1)}.$$

For the case $1/2 < s < 1$, the error $h\|[\![\boldsymbol{\sigma} - \boldsymbol{\sigma}_I]\!]_N\|^2_{L^2(\Gamma)}$ cannot be bounded as (58) because the H^1 trace estimate is now unavailable. In this case, the embedding theory will be the main tool to estimate the errors. We let $\mathbb{V}^m_h \subset \mathbb{H}^1(\Omega)$ be the tensor-valued C^0 finite element space of degree m on the mesh $\mathcal{T}_h$. Since $\pi_i\boldsymbol{\sigma} \in \mathbb{H}^s(\Omega)$, we let $\pi^{\mathbf{s}}_i\boldsymbol{\sigma}$ be its Scott-Zhang interpolant into the space $\mathbb{V}^m_h$ [18], which satisfies that $\|\pi^{\mathbf{s}}_i\boldsymbol{\sigma} - \pi_i\boldsymbol{\sigma}\|_{L^2(\Omega)} \leq Ch^{t_0}\|\pi_i\boldsymbol{\sigma}\|_{H^s(\Omega)}$. For arbitrarily small $\varepsilon > 0$, the space $\mathbb{H}^{1/2+\varepsilon}(\Omega_i)$ continuously embeds into $\mathbb{L}^2(\Gamma)$. Notice that $\pi^{\mathbf{s}}_i\boldsymbol{\sigma}, \pi_i\boldsymbol{\sigma} \in \mathbb{H}^{1/2+\varepsilon}(\Omega_i)$ for small enough ε, and we apply the inverse estimate and the triangle inequality to derive that

$$
\begin{aligned}
h\|[\![\boldsymbol{\sigma} - \boldsymbol{\sigma}_I]\!]\|^2_{L^2(\Gamma)} &\leq Ch\sum_{i=0,1}\|\pi_i\boldsymbol{\sigma} - \pi_i\boldsymbol{\sigma}_I\|^2_{L^2(\Gamma)} \leq Ch\sum_{i=0,1}\|\pi_i\boldsymbol{\sigma} - \pi^{\mathbf{s}}_i\boldsymbol{\sigma}\|^2_{L^2(\Gamma)} + Ch\sum_{i=0,1}\|\pi^{\mathbf{s}}_i\boldsymbol{\sigma} - \pi_i\boldsymbol{\sigma}_I\|^2_{L^2(\Gamma)} \\
&\leq Ch\sum_{i=0,1}\|\pi_i\boldsymbol{\sigma} - \pi^{\mathbf{s}}_i\boldsymbol{\sigma}\|^2_{H^{1/2+\varepsilon}(\Omega)} + C\sum_{i=0,1}\sum_{K\in\mathcal{T}^\Gamma_h}\|\pi^{\mathbf{s}}_i\boldsymbol{\sigma} - \pi_i\boldsymbol{\sigma}_I\|^2_{L^2(K)} \\
&\leq Ch\sum_{i=0,1}\|\pi_i\boldsymbol{\sigma} - \pi^{\mathbf{s}}_i\boldsymbol{\sigma}\|^2_{H^{1/2+\varepsilon}(\Omega)} + C\sum_{i=0,1}\|\pi^{\mathbf{s}}_i\boldsymbol{\sigma} - \pi_i\boldsymbol{\sigma}_I\|^2_{L^2(\Omega_{h,i})} \leq Ch^{2t_1}\|\boldsymbol{\sigma}\|^2_{H^s(\Omega_0\cup\Omega_1)} \\
&\quad + C\sum_{i=0,1}\left(\|\pi_i\boldsymbol{\sigma} - \pi^{\mathbf{s}}_i\boldsymbol{\sigma}\|^2_{L^2(\Omega_{h,i})} + \|\pi_i\boldsymbol{\sigma} - \pi_i\boldsymbol{\sigma}_I\|^2_{L^2(\Omega_{h,i})}\right) \leq Ch^{2t_1}\|\boldsymbol{\sigma}\|^2_{H^s(\Omega_0\cup\Omega_1)},
\end{aligned}
$$

where $t_1 = \min(m, s) - \varepsilon$. From the embedding $\mathbf{H}^{3/2+\varepsilon}(\Omega_i) \hookrightarrow \mathbf{H}^1(\Gamma)$ [21], it is similar to get that

$$
h\|[\![\nabla_\Gamma(\boldsymbol{u} - \boldsymbol{u}_I)]\!]\|^2_{L^2(\Gamma)} \leq Ch^{2t_1}\|\boldsymbol{u}\|^2_{H^{s+1}(\Omega_0\cup\Omega_1)}.
$$

Collecting all above estimates leads to the approximation property (57), which completes the proof. $\qquad\square$

The error estimation for the numerical solution of (53) can be reached. The proof is the same as Theorem 1.

Theorem 3. *Let* $(\boldsymbol{\sigma}, \boldsymbol{u}) \in \boldsymbol{\Sigma}^s \times \mathbf{V}^{s+1}$ *be the exact solution to the interface problem* (1), *and let* $(\boldsymbol{\sigma}_h, \boldsymbol{u}_h) \in \boldsymbol{\Sigma}^m_h \times \mathbf{V}^m_h$ *be the numermical solution to the problem* (53), *there holds*

$$
\|(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h)\|_{\widetilde{\mathbf{e}}_h} \leq Ch^t\left(\|\boldsymbol{\sigma}\|_{H^s(\mathrm{div};\Omega_0\cup\Omega_1)} + \|\boldsymbol{u}\|_{H^{s+1}(\Omega_0\cup\Omega_1)}\right), \tag{59}
$$

where $t = \min(m, s)$ *for* $s \geq 1$, *and* $t = \min(m, s) - \varepsilon$ *with any* $\varepsilon > 0$ *for* $1/2 < s < 1$.

The condition number of the linear system of $a_{\widetilde{\mathcal{J}}_h}(\cdot; \cdot)$ also has a uniform upper bound. The proof is the same to Theorem 2.

Theorem 4. *There exists* C *such that*

$$
\kappa\left(A_{\widetilde{\mathcal{J}}_h}\right) \leq Ch^{-2}, \tag{60}
$$

where $A_{\widetilde{\mathcal{J}}_h}$ *is the linear system with respect to* $a_{\widetilde{\mathcal{J}}_h}(\cdot; \cdot)$.

From Theorem 3, it can be seen that the scheme is also robust since the constants in the error bounds (59) are independent of λ and the relative location between Γ and the mesh. Compared with the L^2 norm least squares finite element method, the convergence rate is nearly optimal without a higher regularity assumption, but the linear system is more complicated.

Ultimately, let us give some details about solving the linear system $A_{\widetilde{\mathcal{J}}_h}$. Since $a_{\widetilde{\mathcal{J}}_h}(\cdot; \cdot)$ involves the discrete minus inner product $(\cdot, \cdot)_{-1/2,h}$, we are required to solve the elliptic system (45) in the assembling of $A_{\widetilde{\mathcal{J}}_h}$. Let B be the sparse matrix of the bilinear form (45), which is invertible and satisfies that $\kappa(B) \leq Ch^{-2}$, see Remark 2. Let C be the sparse matrix corresponding to the inner product $([\![\boldsymbol{\tau}_h]\!], \boldsymbol{w}_h)_{L^2(\Gamma)}$ for $\forall(\boldsymbol{\tau}_h, \boldsymbol{v}_h) \in \boldsymbol{\Sigma}^m_h \times \mathbf{V}^m_h, \forall \boldsymbol{w}_h \in \mathbf{V}^1_{h,0}$. We note that this definition ensures the matrix C has the same column size as the matrix $A_{\widetilde{\mathcal{J}}_h}$. Then, $A_{\widetilde{\mathcal{J}}_h}$ has the form that

$$
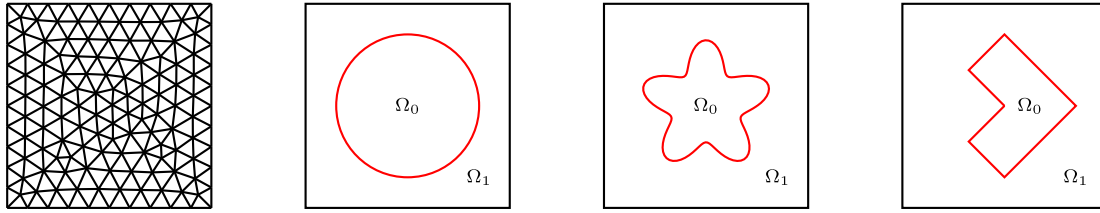A_{\widetilde{\mathcal{J}}_h} = A_{L^2} + C^T B^{-1} C, \tag{61}
$$

FIGURE 1. The unfitted mesh and the interfaces in two dimensions.

where A_{L^2} corresponds to all L^2 inner products in $a_{\widetilde{\mathcal{J}}_h}(\cdot, \cdot)$, *i.e.* A_{L^2} is the matrix of the bilinear form $a_{L^2}(\cdot; \cdot)$ that

$$
\begin{aligned}
a_{L^2}(\boldsymbol{\tau}_h, \boldsymbol{v}_h; \boldsymbol{\rho}_h, \boldsymbol{w}_h) := {}& (\mathcal{A}\boldsymbol{\tau}_h - \boldsymbol{\varepsilon}(\boldsymbol{v}_h), \mathcal{A}\boldsymbol{\rho}_h - \boldsymbol{\varepsilon}(\boldsymbol{w}_h))_{L^2(\Omega_0 \cup \Omega_1)} + (\nabla \cdot \boldsymbol{\tau}_h, \nabla \cdot \boldsymbol{\rho}_h)_{L^2(\Omega_0 \cup \Omega_1)} \\
& + h([\boldsymbol{\tau}_h]_N, [\boldsymbol{\rho}_h]_N)_{L^2(\Gamma)} + h^{-1}([\boldsymbol{v}_h], [\boldsymbol{w}_h])_{L^2(\Gamma)} + h([\nabla_\Gamma \boldsymbol{v}_h], [\nabla_\Gamma \boldsymbol{w}_h])_{L^2(\Gamma)} \\
& + \mathcal{S}_h^m(\boldsymbol{\tau}_h, \boldsymbol{\rho}_h) + \mathcal{G}_h^m(\boldsymbol{v}_h, \boldsymbol{w}_h).
\end{aligned}
\tag{62}
$$

Suppose that we have a fast algorithm to solve the linear system $B\mathbf{y} = \mathbf{z}$, then we can easily compute the matrix-vector products for the matrix $A_{\widetilde{\mathcal{J}}_h}$ by (61). Consequently, the Krylov iterative method (*e.g.* GMRES) can be used as the solver for the linear system. Although we have shown that the condition number of $A_{\widetilde{\mathcal{J}}_h}$ is $O(h^{-2})$, an effective preconditioner is still expected especially when h tends to zero. But $A_{\widetilde{\mathcal{J}}_h}$ involves the inverse matrix B^{-1}, it is costly to form it and it is also not convenient to even extract its diagonal. Traditional diagonal or block diagonal preconditioners are hard to use for $A_{\widetilde{\mathcal{J}}_h}$. One method is to use the matrix-free preconditioning technique to $A_{\widetilde{\mathcal{J}}_h}$. For example, one can use the hierarchically semiseparable (HSS) approximation for the given matrix to accelerate the convergence of iterative methods [15,54,55]. The construction of the HSS approximation may potentially only use matrix-vector products instead of the original matrix itself, and we refer to [44, 54] for fully matrix-free techniques to the construction of the HSS approximation. This idea has been used in the immersed finite element method solving the elliptic interface problem [56] . Once the HSS approximation H for $A_{\widetilde{\mathcal{J}}_h}$ is obtained in a structured form, it can be quickly factorized and the factors can be used as a preconditioner. Alternatively, we present a matrix-explicit preconditioning method for $A_{\widetilde{\mathcal{J}}_h}$. As (61), the matrix $A_{\widetilde{\mathcal{J}}_h}$ can be split into two parts. From (62), it can be easily seen that $a_{L^2}(\boldsymbol{\tau}_h, \boldsymbol{v}_h; \boldsymbol{\tau}_h, \boldsymbol{v}_h) = 0$ implies $\boldsymbol{\tau}_h = \mathbf{0}$, $\boldsymbol{v}_h = \mathbf{0}$ for $(\boldsymbol{\tau}_h, \boldsymbol{v}_h) \in \boldsymbol{\Sigma}_h^m \times \mathbf{V}_h^m$. Hence, A_{L^2} is symmetric positive definite. Any preconditioning technique can be applied to A_{L^2} because A_{L^2} is entirely explicit. As a numerical observation, the preconditioner from A_{L^2} can also significantly accelerate the iterative methods. We show the results in Example 1 of Section 5 by constructing the HSS approximation from A_{L^2} to obtain the preconditioner. For solving the linear system $B\mathbf{y} = \mathbf{z}$, we also use the HSS approximation as the preconditioner. The codes of the construction and the factorization of the HSS approximation are freely available in STRUMPACK [50]. A comprehensive analysis and a more appropriate method to solve the linear system are now left in the future study.

5. NUMERICAL RESULTS

In this section, a series of numerical tests are presented to show the numerical performance of the proposed method. For all tests, the source function $\boldsymbol{f}$ and the jump condition $\boldsymbol{a}, \boldsymbol{b}$ are taken from the exact solution accordingly. In Examples 1–5, we consider the elasticity interface problems in two dimensions on the squared domain $\Omega = (-1, 1)^2$. We adopt a family of triangular meshes with the mesh size $h = 1/5, 1/10, \ldots, 1/40$ to solve all tests. In Examples 1–3, the interface Γ is of class C^2 and is described by a level set function, see Figure 1. In Examples 4 and 5, the interface is taken to be the boundary of the L-shaped domain. In Example 6, we solve a three-dimensional interface problem to illustrate the numerical performance.

TABLE 1. The numerical results for Example 1 by the L^2 norm least squares finite element method (left)/the least squares finite element method with the discrete minus norm (right).

m	h	1/5	1/10	1/20	1/40	order	m	h	1/5	1/10	1/20	1/40	order
1	$\|\boldsymbol{u}-\boldsymbol{u}_h\|_{L^2(\Omega_0\cup\Omega_1)}$	1.86e−1	4.76e−2	1.19e−2	2.98e−3	2.00	1	$\|\boldsymbol{u}-\boldsymbol{u}_h\|_{L^2(\Omega_0\cup\Omega_1)}$	1.88e−1	4.83e−2	1.21e−2	3.03e−3	2.00
	$\|\boldsymbol{\sigma}-\boldsymbol{\sigma}_h\|_{L^2(\Omega_0\cup\Omega_1)}$	3.49e−0	1.71e−0	8.58e−1	4.32e−1	0.99		$\|\boldsymbol{\sigma}-\boldsymbol{\sigma}_h\|_{L^2(\Omega_0\cup\Omega_1)}$	3.50e−0	1.71e−0	8.58e−1	4.32e−1	0.99
	$\|(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h,\boldsymbol{u}-\boldsymbol{u}_h)\|_{\mathbf{e}}$	2.22e1	1.12e1	5.66e−0	2.85e−0	0.99		$\|(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h,\boldsymbol{u}-\boldsymbol{u}_h)\|_{\mathbf{e}}$	2.22e1	1.12e1	5.66e−0	2.84e−0	1.00
2	$\|\boldsymbol{u}-\boldsymbol{u}_h\|_{L^2(\Omega_0\cup\Omega_1)}$	4.01e−3	4.92e−4	6.15e−5	7.69e−6	2.99	2	$\|\boldsymbol{u}-\boldsymbol{u}_h\|_{L^2(\Omega_0\cup\Omega_1)}$	4.02e−3	4.93e−4	6.13e−5	7.68e−6	3.00
	$\|\boldsymbol{\sigma}-\boldsymbol{\sigma}_h\|_{L^2(\Omega_0\cup\Omega_1)}$	2.99e−1	7.48e−2	1.89e−2	4.77e−3	1.99		$\|\boldsymbol{\sigma}-\boldsymbol{\sigma}_h\|_{L^2(\Omega_0\cup\Omega_1)}$	3.00e−1	7.47e−2	1.90e−2	4.78e−3	2.00
	$\|(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h,\boldsymbol{u}-\boldsymbol{u}_h)\|_{\mathbf{e}}$	2.03e−0	5.13e−1	1.28e−1	3.23e−2	1.99		$\|(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h,\boldsymbol{u}-\boldsymbol{u}_h)\|_{\mathbf{e}}$	2.03e−0	5.12e−1	1.29e−1	3.22e−2	2.00
3	$\|\boldsymbol{u}-\boldsymbol{u}_h\|_{L^2(\Omega_0\cup\Omega_1)}$	1.65e−4	8.95e−6	5.42e−7	3.36e−8	4.01	3	$\|\boldsymbol{u}-\boldsymbol{u}_h\|_{L^2(\Omega_0\cup\Omega_1)}$	1.64e−4	8.93e−6	5.41e−7	3.38e−8	4.00
	$\|\boldsymbol{\sigma}-\boldsymbol{\sigma}_h\|_{L^2(\Omega_0\cup\Omega_1)}$	1.93e−2	2.31e−3	2.88e−4	3.59e−5	3.00		$\|\boldsymbol{\sigma}-\boldsymbol{\sigma}_h\|_{L^2(\Omega_0\cup\Omega_1)}$	1.92e−2	2.30e−3	2.87e−4	3.60e−5	3.01
	$\|(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h,\boldsymbol{u}-\boldsymbol{u}_h)\|_{\mathbf{e}}$	1.26e−1	1.60e−2	2.02e−3	2.53e−3	2.99		$\|(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h,\boldsymbol{u}-\boldsymbol{u}_h)\|_{\mathbf{e}}$	1.26e−1	1.61e−2	2.01e−3	2.51e−3	3.00

TABLE 2. Convergence steps for the linear system of the L^2 norm least squares finite element method in Example 1.

m	h	1/5	1/10	1/20	1/40
1	Preconditioner from HSS approximation	8	12	19	36
	Diagonal preconditioner	513	1003	2103	>3000
2	Preconditioner from HSS approximation	13	21	39	70
	Diagonal preconditioner	2273	>3000	>3000	>3000
3	Preconditioner from HSS approximation	21	32	56	91
	Diagonal preconditioner	>3000	>3000	>3000	>3000

Example 1. We first consider a linear elasticity interface problem with a circular interface centered at the origin with radius $r = 0.7$, see Figure 1. We solve this problem to study the convergence rates. The exact solution is taken as

$$\boldsymbol{u}(x,y) = \begin{bmatrix} \sin(2\pi y)(-1 + \cos(2\pi x)) + \frac{1}{1+\lambda}\sin(\pi x)\sin(\pi y) \\ \sin(2\pi x)(1 - \cos(2\pi y)) + \frac{1}{1+\lambda}\sin(\pi x)\sin(\pi y) \end{bmatrix} \quad \text{in } \Omega_0 \cup \Omega_1,$$

with the discontinuous parameters $\lambda|_{\Omega_0} = \lambda_0 = 5, \lambda|_{\Omega_1} = \lambda_1 = 1, \mu|_{\Omega_0} = \mu_0 = 2, \mu|_{\Omega_1} = \mu_1 = 1$. The numerical results are gathered in Table 1 by the L^2 norm least squares finite element method and the least squares finite element method with the discrete minus norm. For both methods, the energy norms $\|\cdot\|_{\mathbf{e}_h}$ and $\|\cdot\|_{\widetilde{\mathbf{e}}_h}$ are stronger than the norm $\|\cdot\|_{\mathbf{e}}$. We calculate $\|(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h, \boldsymbol{u}-\boldsymbol{u}_h)\|_{\mathbf{e}}$ to represent the errors under the energy norm. From Table 1, it can be seen that the errors under the energy norm approach zero at the optimal rates. For the L^2 errors, $\|\boldsymbol{u}-\boldsymbol{u}_h\|_{L^2(\Omega)}$ and $\|\boldsymbol{\sigma}-\boldsymbol{\sigma}_h\|_{L^2(\Omega)}$ have optimal/suboptimal convergence speeds. All numerically observed convergence orders are in agreement with the theoretical analysis in Theorems 1 and 3.

In this example, we demonstrate the numerical performance of the iterative method for solving the resulting linear system. For all tests, we use GMRES as the iterative solver and the iteration stops when the relative error $\|A\mathbf{x}^k - \mathbf{b}\|_{l^2}/\|\mathbf{b}\|_{l^2} < 10^{-8}$ at the stage k is smaller than the tolerance 10^{-8}. For the L^2 norm least squares finite element method, the resulting matrix is entirely explicit and any preconditioning technique can be used. Here, we try to use the HSS approximation and its diagonal to construct the preconditioners. The convergence steps are collected in Table 2. It can be seen that the convergence is apparently accelerated by the HSS preconditioning technique.

For the method with the discrete minus norm, we shall solve the linear system $A_{\widetilde{\mathcal{J}}_h}\mathbf{x} = \mathbf{b}$. As stated in the end of Section 4.2, we construct the HSS approximation H_{L^2} from A_{L^2} to obtain the preconditioner, and we also directly construct the HSS approximation H from A to obtain the preconditioner as a comparison. The

TABLE 3. Convergence steps for the linear system of the least squares finite element method with the discrete minus norm in Example 1.

| m	h	1/5	1/10	1/20	1/40
1	Preconditioner from HSS approximation H	9	13	21	38
	Preconditioner from HSS approximation H_{L^2}	10	16	23	42
	Identical preconditioner I	1738	>3000	>3000	>3000
2	Preconditioner from HSS approximation H	15	22	38	70
	Preconditioner from HSS approximation H_{L^2}	16	23	41	78
	Identical preconditioner I	>3000	>3000	>3000	>3000
3	Preconditioner from HSS approximation H	23	33	58	97
	Preconditioner from HSS approximation H_{L^2}	25	37	63	108
	Identical preconditioner I	>3000	>3000	>3000	>3000
B	Preconditioner from HSS approximation	4	7	13	21
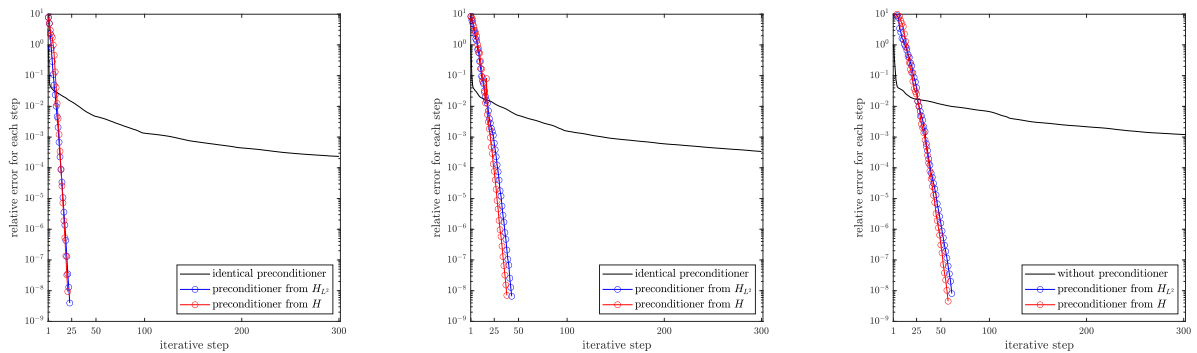
FIGURE 2. The convergence histories for iterative methods with the accuracy $m = 1$ (*left*)/$m = 2$ (*mid*)/$m = 3$ (*right*).

convergence steps are collected in Table 3. The convergence speeds of both preconditioned GMRES methods are significantly faster than the standard GMRES method. The convergence histories for the iterative methods on the mesh $h = 1/20$ are depicted in Figure 2. The numerical performances of both preconditioned methods are very close, and it is more convenient to construct the HSS approximation from A_{L^2} in a standard procedure. In every step of Krylov iteration, we are required to solve the linear system $B\mathbf{y} = \mathbf{z}$. We still use GMRES with the preconditioner from the HSS approximation as the solver for this system. The average iterative steps for different meshes are collected in Table 3. The iterative method also has a fast convergence speed. In the rest examples, we use the factors of H_{L^2} as the preconditioner to solve the linear system. Developing a more appropriate method to solve the linear system is now left in the future study.

Example 2. In this example, we consider the same interface problem as Example 1 but the Lamé parameters have a large jump across the interface. We solve this problem to demonstrate the robustness of the proposed method when $\lambda \to \infty$. The parameters are selected as $\lambda_0 = 10000/100, \lambda_1 = 1, \mu_0 = \mu_1 = 1$. The numerical results for both methods are displayed in Tables 4 and 5, respectively. We observe that the discrete solutions from both methods converge uniformly as $\lambda \to \infty$. The results illustrates the robustness when the parameter $\lambda \to \infty$.

TABLE 4. Numerical results for Example 2 by the L^2 norm least squares finite element method with $\lambda = 100, 10000$.

m	h / λ	1/5		1/10		1/20		1/40		Order
		100	10000	100	10000	100	10000	100	10000	
1	$\|\boldsymbol{u}-\boldsymbol{u}_h\|_{L^2(\Omega_0\cup\Omega_1)}$	1.840e−1	1.842e−1	4.709e−2	4.713e−2	1.180e−2	1.181e−2	2.971e−3	2.972e−3	2.00
	$\|\boldsymbol{\sigma}-\boldsymbol{\sigma}_h\|_{L^2(\Omega_0\cup\Omega_1)}$	3.565e−0	3.569e−0	1.751e−0	1.753e−0	8.800e−1	8.814e−1	4.369e−1	4.399e−1	1.00
	$\|(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h,\boldsymbol{u}-\boldsymbol{u}_h)\|_{\mathbf{e}}$	2.220e1	2.220e1	1.127e1	1.126e1	5.665e−0	5.665e−0	2.833e−0	2.835e−0	0.99
2	$\|\boldsymbol{u}-\boldsymbol{u}_h\|_{L^2(\Omega_0\cup\Omega_1)}$	4.016e−3	4.016e−3	4.924e−4	4.925e−4	6.149e−5	6.150e−5	7.692e−7	7.693e−7	3.00
	$\|\boldsymbol{\sigma}-\boldsymbol{\sigma}_h\|_{L^2(\Omega_0\cup\Omega_1)}$	3.055e−1	3.058e−1	7.626e−2	7.636e−2	1.932e−2	1.933e−2	4.797e−3	4.799e−3	2.00
	$\|(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h,\boldsymbol{u}-\boldsymbol{u}_h)\|_{\mathbf{e}}$	2.031e−0	2.030e−0	5.135e−1	5.135e−1	1.289e−1	1.289e−1	3.223e−2	3.222e−2	2.00
3	$\|\boldsymbol{u}-\boldsymbol{u}_h\|_{L^2(\Omega_0\cup\Omega_1)}$	1.651e−4	1.651e−4	8.925e−6	8.924e−6	5.410e−7	5.411e−7	3.379e−7	3.381e−7	4.01
	$\|\boldsymbol{\sigma}-\boldsymbol{\sigma}_h\|_{L^2(\Omega_0\cup\Omega_1)}$	2.222e−2	2.223e−2	2.419e−3	2.422e−3	2.940e−4	2.943e−4	3.631e−5	3.653e−5	3.02
	$\|(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h,\boldsymbol{u}-\boldsymbol{u}_h)\|_{\mathbf{e}}$	1.263e−1	1.263e−1	1.602e−2	1.603e−2	2.017e−3	2.017e−3	2.522e−4	2.523e−4	2.99

TABLE 5. Numerical results for Example 2 by the least squares finite element method of the discrete minus norm with $\lambda = 100, 10000$.

m	h / λ	1/5		1/10		1/20		1/40		Order
		100	10000	100	10000	100	10000	100	10000	
1	$\|\boldsymbol{u}-\boldsymbol{u}_h\|_{L^2(\Omega_0\cup\Omega_1)}$	1.859e−1	1.858e−1	4.773e−2	4.771e−2	1.199e−2	1.198e−2	3.018e−3	3.016e−3	1.99
	$\|\boldsymbol{\sigma}-\boldsymbol{\sigma}_h\|_{L^2(\Omega_0\cup\Omega_1)}$	3.572e−0	3.576e−0	1.753e−0	1.756e−0	8.806e−1	8.882e−1	4.447e−1	4.449e−1	0.99
	$\|(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h,\boldsymbol{u}-\boldsymbol{u}_h)\|_{\mathbf{e}}$	2.220e1	2.220e1	1.127e1	1.127e1	5.666e−0	5.665e−0	2.847e−0	2.847e−0	0.99
2	$\|\boldsymbol{u}-\boldsymbol{u}_h\|_{L^2(\Omega_0\cup\Omega_1)}$	4.017e−3	4.017e−3	4.925e−4	4.926e−4	6.150e−5	6.151e−5	7.693e−6	7.693e−6	3.00
	$\|\boldsymbol{\sigma}-\boldsymbol{\sigma}_h\|_{L^2(\Omega_0\cup\Omega_1)}$	3.053e−1	3.056e−1	7.627e−2	7.637e−2	1.932e−2	1.934e−2	4.893e−3	4.986e−3	1.98
	$\|(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h,\boldsymbol{u}-\boldsymbol{u}_h)\|_{\mathbf{e}}$	2.031e−0	2.030e−0	5.133e−1	5.133e−1	1.289e−1	1.289e−1	3.231e−2	3.233e−2	1.99
3	$\|\boldsymbol{u}-\boldsymbol{u}_h\|_{L^2(\Omega_0\cup\Omega_1)}$	1.649e−4	1.648e−4	8.920e−6	8.920e−6	5.409e−7	5.408e−7	3.325e−8	3.323e−8	4.03
	$\|\boldsymbol{\sigma}-\boldsymbol{\sigma}_h\|_{L^2(\Omega_0\cup\Omega_1)}$	2.217e−2	2.220e−2	2.418e−3	2.420e−3	2.939e−4	2.942e−4	3.633e−5	3.635e−5	3.02
	$\|(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h,\boldsymbol{u}-\boldsymbol{u}_h)\|_{\mathbf{e}}$	1.263e−1	1.264e−1	1.602e−2	1.602e−2	2.017e−3	2.018e−3	2.535e−4	2.536e−4	2.99

Example 3. In this example, we consider an elasticity interface problem with a star-shaped interface consisting of both concave and convex curve segments, see Figure 1. The interface Γ is governed by the polar angle θ that

$$r = \frac{1}{2} + \frac{\sin 5\theta}{7}.$$

The analytic solution $\boldsymbol{u}$ is taken in a piecewise manner as

$$\boldsymbol{u}(x,y) = \begin{cases} \boldsymbol{u}^0(x,y) & \text{in } \Omega_0, \\ \boldsymbol{u}^1(x,y), & \text{in } \Omega_1, \end{cases}$$

where

$$\boldsymbol{u}^0(x,y) = [\cos(\pi x)\cos(\pi y), \cos(\pi y)]^T, \quad \boldsymbol{u}^1(x,y) = [\sin(\pi x)\sin(\pi y), x(1-x)\sin(\pi y)]^T.$$

The parameters are chosen as $\lambda_0 = \lambda_1 = \mu_0 = \mu_1 = 1$. We present the numerical results in Table 6. The predicted convergence rates under the energy norms and the L^2 norms for both unknowns are verified from the convergence histories.

TABLE 6. Numerical results for Example 3 by the L^2 norm least squares finite element method (left)/the least squares finite element method with the discrete minus norm (right).

m	h	1/5	1/10	1/20	1/40	Order	m	h	1/5	1/10	1/20	1/40	Order
1	$\|\boldsymbol{u}-\boldsymbol{u}_h\|_{L^2(\Omega_0\cup\Omega_1)}$	9.131e−1	2.431e−2	6.244e−3	1.575e−3	1.98	1	$\|\boldsymbol{u}-\boldsymbol{u}_h\|_{L^2(\Omega_0\cup\Omega_1)}$	9.127e−2	2.430e−2	6.237e−3	1.573e−3	1.99
	$\|\boldsymbol{\sigma}-\boldsymbol{\sigma}_h\|_{L^2(\Omega_0\cup\Omega_1)}$	1.127e−0	5.245e−1	2.621e−1	1.326e−1	0.98		$\|\boldsymbol{\sigma}-\boldsymbol{\sigma}_h\|_{L^2(\Omega_0\cup\Omega_1)}$	1.127e−0	5.267e−1	2.627e−1	1.327e−1	0.99
	$\|(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h,\boldsymbol{u}-\boldsymbol{u}_h)\|_{\mathbf{e}}$	9.270e−0	4.722e−0	2.377e−0	1.193e−0	1.00		$\|(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h,\boldsymbol{u}-\boldsymbol{u}_h)\|_{\mathbf{e}}$	9.270e−0	4.722e−0	2.377e−0	1.193e−0	1.00
2	$\|\boldsymbol{u}-\boldsymbol{u}_h\|_{L^2(\Omega_0\cup\Omega_1)}$	3.047e−3	3.513e−4	4.412e−5	5.551e−6	2.99	2	$\|\boldsymbol{u}-\boldsymbol{u}_h\|_{L^2(\Omega_0\cup\Omega_1)}$	3.441e−3	3.513e−3	4.412e−5	5.551e−6	2.99
	$\|\boldsymbol{\sigma}-\boldsymbol{\sigma}_h\|_{L^2(\Omega_0\cup\Omega_1)}$	1.193e−1	2.432e−2	5.946e−3	1.503e−6	1.98		$\|\boldsymbol{\sigma}-\boldsymbol{\sigma}_h\|_{L^2(\Omega_0\cup\Omega_1)}$	1.182e−1	2.430e−2	5.953e−3	1.503e−3	1.98
	$\|(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h,\boldsymbol{u}-\boldsymbol{u}_h)\|_{\mathbf{e}}$	8.171e−1	2.079e−1	5.239e−2	1.315e−2	2.00		$\|(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h,\boldsymbol{u}-\boldsymbol{u}_h)\|_{\mathbf{e}}$	8.170e−1	2.079e−1	5.239e−2	1.315e−2	2.00
3	$\|\boldsymbol{u}-\boldsymbol{u}_h\|_{L^2(\Omega_0\cup\Omega_1)}$	1.595e−4	6.891e−6	3.799e−7	2.315e−8	4.03	3	$\|\boldsymbol{u}-\boldsymbol{u}_h\|_{L^2(\Omega_0\cup\Omega_1)}$	1.574e−4	6.863e−6	3.797e−7	2.315e−8	4.03
	$\|\boldsymbol{\sigma}-\boldsymbol{\sigma}_h\|_{L^2(\Omega_0\cup\Omega_1)}$	8.609e−3	8.586e−4	9.233e−5	1.135e−5	3.02		$\|\boldsymbol{\sigma}-\boldsymbol{\sigma}_h\|_{L^2(\Omega_0\cup\Omega_1)}$	7.772e−3	8.269e−4	1.003e−4	1.250e−5	3.00
	$\|(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h,\boldsymbol{u}-\boldsymbol{u}_h)\|_{\mathbf{e}}$	5.527e−2	6.517e−3	8.173e−4	1.027e−4	2.99		$\|(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h,\boldsymbol{u}-\boldsymbol{u}_h)\|_{\mathbf{e}}$	5.553e−2	6.517e−3	8.173e−4	1.026e−4	2.99

TABLE 7. Numerical results for Example 4 by the L^2 norm least squares finite element method (left)/the least squares finite element method with the discrete minus norm (right).

m	h	1/5	1/10	1/20	1/40	Order	m	h	1/5	1/10	1/20	1/40	Order
1	$\|\boldsymbol{u}-\boldsymbol{u}_h\|_{L^2(\Omega_0\cup\Omega_1)}$	8.082e−1	2.383e−1	6.356e−2	1.609e−2	1.98	1	$\|\boldsymbol{u}-\boldsymbol{u}_h\|_{L^2(\Omega_0\cup\Omega_1)}$	8.093e−1	2.416e−1	6.583e−2	1.651e−2	1.99
	$\|\boldsymbol{\sigma}-\boldsymbol{\sigma}_h\|_{L^2(\Omega_0\cup\Omega_1)}$	8.601e−0	3.583e−0	1.686e−0	8.412e−1	1.00		$\|\boldsymbol{\sigma}-\boldsymbol{\sigma}_h\|_{L^2(\Omega_0\cup\Omega_1)}$	8.471e−0	3.593e−0	1.699e−0	8.456e−1	1.00
	$\|(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h,\boldsymbol{u}-\boldsymbol{u}_h)\|_{\mathbf{e}}$	3.222e1	2.065e1	1.125e1	5.810e−0	0.96		$\|(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h,\boldsymbol{u}-\boldsymbol{u}_h)\|_{\mathbf{e}}$	3.221e1	2.065e1	1.125e1	5.811e−0	0.95
2	$\|\boldsymbol{u}-\boldsymbol{u}_h\|_{L^2(\Omega_0\cup\Omega_1)}$	8.849e−2	1.015e−2	1.205e−3	1.476e−4	3.03	2	$\|\boldsymbol{u}-\boldsymbol{u}_h\|_{L^2(\Omega_0\cup\Omega_1)}$	8.899e−2	1.013e−2	1.208e−3	1.477e−4	3.03
	$\|\boldsymbol{\sigma}-\boldsymbol{\sigma}_h\|_{L^2(\Omega_0\cup\Omega_1)}$	2.157e−0	4.519e−1	1.147e−1	2.914e−2	1.98		$\|\boldsymbol{\sigma}-\boldsymbol{\sigma}_h\|_{L^2(\Omega_0\cup\Omega_1)}$	2.127e−0	4.525e−1	1.153e−1	2.922e−2	1.98
	$\|(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h,\boldsymbol{u}-\boldsymbol{u}_h)\|_{\mathbf{e}}$	8.596e−0	2.652e−0	7.256e−1	1.875e−1	1.96		$\|(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h,\boldsymbol{u}-\boldsymbol{u}_h)\|_{\mathbf{e}}$	8.593e−0	2.658e−0	7.255e−1	1.875e−1	1.95
3	$\|\boldsymbol{u}-\boldsymbol{u}_h\|_{L^2(\Omega_0\cup\Omega_1)}$	1.453e−2	5.489e−4	3.156e−5	1.907e−6	4.05	3	$\|\boldsymbol{u}-\boldsymbol{u}_h\|_{L^2(\Omega_0\cup\Omega_1)}$	1.453e−2	5.486e−4	3.157e−5	1.908e−6	4.03
	$\|\boldsymbol{\sigma}-\boldsymbol{\sigma}_h\|_{L^2(\Omega_0\cup\Omega_1)}$	5.333e−1	4.399e−2	5.080e−3	6.361e−4	3.00		$\|\boldsymbol{\sigma}-\boldsymbol{\sigma}_h\|_{L^2(\Omega_0\cup\Omega_1)}$	5.272e−1	4.375e−2	5.087e−3	6.416e−4	2.99
	$\|(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h,\boldsymbol{u}-\boldsymbol{u}_h)\|_{\mathbf{e}}$	1.499e−0	2.609e−1	3.626e−2	4.699e−3	2.95		$\|(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h,\boldsymbol{u}-\boldsymbol{u}_h)\|_{\mathbf{e}}$	1.499e−0	2.609e−1	3.625e−2	4.699e−3	2.95

Example 4. In this test, we consider the interface problem with an L-shaped polygonal interface Γ, which is described by the following vertices, see Figure 1,

$$(0,0), \quad (-0.35, 0.35), \quad (0, 0.7), \quad (0.7, 0), \quad (0, -0.7), \quad (-0.35, 0.35).$$

The exact solution and the parameters are taken the same as Example 1. The numerical errors are displayed in Table 7. For the case of the polygonal interface, all numerically detected convergence speeds still agree with the theoretical results.

Example 5. In this test, we still consider the L-shaped interface and we investigate the performance of the method dealing with the problem of a singular solution. The exact solution is selected to be

$$\boldsymbol{u}(x,y) = \begin{cases} \boldsymbol{u}^0(x,y), & \text{in } \Omega_0, \\ [0,0]^T, & \text{in } \Omega_1, \end{cases}$$

with the parameters $\lambda_0 = \lambda_1 = \mu_0 = \mu_1 = 1$, where $\boldsymbol{u}^0$ is given as

$$u_r^0(r,\theta) = \frac{r^\alpha}{2\mu}(-(\alpha+1)\cos((\alpha+1)\theta) + (C_2 - (\alpha+1))C_1\cos((\alpha-1)\theta)),$$

$$u_\theta^0(r,\theta) = \frac{r^\alpha}{2\mu}((\alpha+1)\sin((\alpha+1)\theta) + (C_2 + \alpha - 1)C_1\sin((\alpha-1)\theta)),$$

in the polar coordinates (r,θ). We let $\alpha \approx 0.5444837$ be the solution of the following function

$$\alpha\sin(2w) + \sin(2w\alpha) = 0,$$

TABLE 8. Numerical errors for Example 5 by the least squares finite element method with discrete minus norm.

m	h	1/5	1/10	1/20	1/40	1/80	1/160	Order
	$\|\boldsymbol{u} - \boldsymbol{u}_h\|_{L^2(\Omega_0 \cup \Omega_1)}$	4.712e−2	2.100e−2	9.502e−3	3.369e−3	1.308e−3	5.158e−3	1.33
1	$\|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{L^2(\Omega_0 \cup \Omega_1)}$	7.779e−1	6.119e−1	4.950e−1	3.612e−1	2.506e−1	1.761e−1	0.50
	$\|(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h, \boldsymbol{u} - \boldsymbol{u}_h)\|_{\mathbf{e}}$	1.089e−0	7.980e−1	6.278e−1	4.428e−1	3.133e−1	2.170e−1	0.52
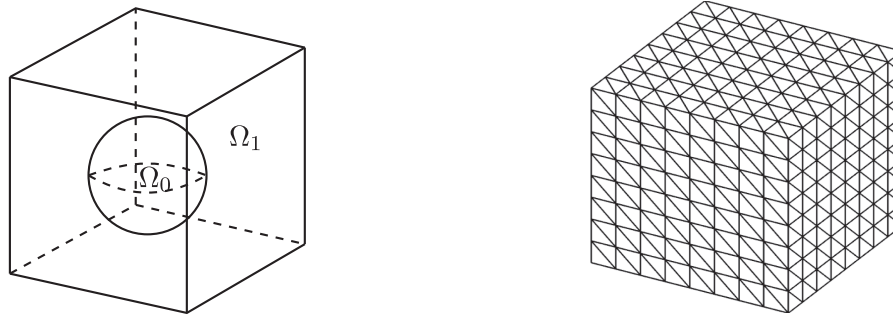
FIGURE 3. The spherical domain and the tetrahedral mesh of Example 3.

with $w = 3\pi/4$, and the constants C_1 and C_2 are

$$C_1 = -\frac{\cos((\alpha + 1)w)}{\cos((\alpha - 1)w)}, \quad C_2 = \frac{2(\lambda + 2\mu)}{\lambda + \mu}.$$

The source term $\boldsymbol{f} = 0$ on the domain $\Omega_0 \cup \Omega_1$. The exact solution has the regularity that $(\boldsymbol{\sigma}, \boldsymbol{u}) \in \mathbf{H}^{\alpha - \varepsilon}(\text{div}; \Omega_0 \cup \Omega_1) \times \mathbf{H}^{1 + \alpha - \varepsilon}(\Omega_0 \cup \Omega_1)$ for $\forall \varepsilon > 0$. Hence, we solve this problem by the least squares finite element with the discrete minus norm. We consider the linear accuracy $m = 1$ for this test, and the results are shown in Table 8. It can be observed that the convergence rate for the energy norm is $O(h^{0.52})$, which is consistent with the regularity of the exact solution and the theoretical analysis. The L^2 errors for both variables are less than the optimal convergence rates, i.e. $O(h^{1.3})$ and $O(h^{0.51})$ for $\boldsymbol{u}$ and $\boldsymbol{\sigma}$, respectively. The reason may be traced back to the singularity of the exact solution.

Example 6. In this test, we consider the interface problem in the cubic domain $\Omega = (0, 1)^3$. We take Γ as a spherical interface with the radius $r = 0.35$ centered at the point $(0.5, 0.5, 0.5)$. Let the exact displacement $\boldsymbol{u}$ be

$$\boldsymbol{u}(x, y, z) = \begin{bmatrix} 2^4 \\ 2^5 \\ 2^6 \end{bmatrix} x(1 - x)y(1 - y)z(1 - z), \quad \text{in } \Omega_0 \cup \Omega_1.$$

The parameters λ, μ are discontinuous across the interface with $\lambda_0 = 10, \lambda_1 = 1, \mu_0 = 10, \mu_1 = 1$. We adopt a series of tetrahedral meshes with the mesh $h = 1/4, 1/8, 1/16, 1/32$ to solve this problem, see Figure 3. The numerical errors under all error measurements are reported in Table 9 for both methods. The numerical results illustrate the accuracy of the methods in three dimensions.

TABLE 9. Numerical results for Example 6 by the L^2 norm least squares finite element method (left)/the least squares finite element method with the discrete minus norm (right).

m	h	1/4	1/8	1/16	1/32	Order	m	h	1/4	1/8	1/16	1/32	Order
1	$\|u - u_h\|_{L^2(\Omega_0\cup\Omega_1)}$	1.483e−1	4.786e−3	1.307e−2	3.335e−4	1.98	1	$\|u - u_h\|_{L^2(\Omega_0\cup\Omega_1)}$	1.149e−1	4.639e−2	1.291e−2	3.334e−3	1.95
	$\|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{L^2(\Omega_0\cup\Omega_1)}$	1.556e−1	7.528e−1	3.921e−1	1.948e−1	1.01		$\|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{L^2(\Omega_0\cup\Omega_1)}$	1.563e−0	7.551e−1	3.925e−1	1.960e−1	0.99
	$\|(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h, u - u_h)\|_{\mathbf{e}}$	5.089e−0	2.636e−0	1.353e−0	6.859e−1	0.98		$\|(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h, u - u_h)\|_{\mathbf{e}}$	5.089e−0	2.637e−0	1.353e−0	6.857e−1	0.98
2	$\|u - u_h\|_{L^2(\Omega_0\cup\Omega_1)}$	1.147e−2	9.065e−4	1.023e−4	1.233e−5	3.03	2	$\|u - u_h\|_{L^2(\Omega_0\cup\Omega_1)}$	1.153e−2	9.058e−4	1.022e−4	1.239e−5	3.04
	$\|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{L^2(\Omega_0\cup\Omega_1)}$	2.199e−1	5.603e−2	1.443e−2	3.685e−3	1.97		$\|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{L^2(\Omega_0\cup\Omega_1)}$	2.197e−1	5.693e−2	1.449e−2	3.691e−3	1.98
	$\|(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h, u - u_h)\|_{\mathbf{e}}$	7.433e−1	1.902e−1	4.888e−2	1.239e−2	1.98		$\|(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h, u - u_h)\|_{\mathbf{e}}$	7.433e−1	1.901e−1	4.887e−2	1.238e−2	1.98

REFERENCES

[1] R.A. Adams and J.J.F. Fournier, Sobolev Spaces. *Pure and Applied Mathematics (Amsterdam)*, 2nd edition. Vol. 140. Elsevier/Academic Press, Amsterdam (2003).

[2] A. Almqvist, C. Campañá, N. Prodanov and B.N.J. Persson, Interfacial separation between elastic solids with randomly rough surfaces: comparison between theory and numerical techniques. *J. Mech. Phys. Solids* **59** (2011) 2355–2369.

[3] S. Badia, F. Verdugo and A. Martín, The aggregated unfitted finite element method for elliptic problems. *Comput. Methods Appl. Mech. Eng.* **336** (2018) 533–553.

[4] R. Becker, E. Burman and P. Hansbo, A Nitsche extended finite element method for incompressible elasticity with discontinuous modulus of elasticity. *Comput. Methods Appl. Mech. Eng.* **198** (2009) 3352–3360.

[5] P.B. Bochev and M.D. Gunzburger, Finite element methods of least-squares type. *SIAM Rev.* **40** (1998) 789–837.

[6] S.P.A. Bordas, E. Burman, M.G. Larson and M.A. Olshanskii, editors, Geometrically Unfitted Finite Element Methods and Applications. *Lecture Notes in Computational Science and Engineering*, Vol. 121. Springer, Cham (2017), Held January 6–8, 2016.

[7] J.H. Bramble, R.D. Lazarov and J.E. Pasciak, A least-squares approach based on a discrete minus one inner product for first order systems. *Math. Comput.* **66** (1997) 935–955.

[8] J.H. Bramble, R.D. Lazarov and J.E. Pasciak, Least-squares methods for linear elasticity based on a discrete minus one inner product. *Comput. Methods Appl. Mech. Eng.* **191** (2001) 727–744.

[9] E. Burman, Ghost penalty. *C. R. Math. Acad. Sci. Paris* **348** (2010) 1217–1220.

[10] E. Burman, S. Claus, P. Hansbo, M.G. Larson and A. Massing, CutFEM: discretizing geometry and partial differential equations. *Int. J. Numer. Methods Eng.* **104** (2015) 472–501.

[11] E. Burman, M. Cicuttin, G. Delay and A. Ern, An unfitted hybrid high-order method with cell agglomeration for elliptic interface problems. *SIAM J. Sci. Comput.* **43** (2021) A859–A882.

[12] Z. Cai and G. Starke, First-order system least squares for the stress-displacement formulation: linear elasticity. *SIAM J. Numer. Anal.* **41** (2003) 715–730.

[13] Z. Cai and G. Starke, Least-squares methods for linear elasticity. *SIAM J. Numer. Anal.* **42** (2004) 826–842.

[14] Z. Cai, T.A. Manteuffel, S.F. McCormick and S.V. Parter, First-order system least squares (FOSLS) for planar linear elasticity: pure traction problem. *SIAM J. Numer. Anal.* **35** (1998) 320–335.

[15] S. Chandrasekaran, M. Gu and T. Pals, A fast ULV decomposition solver for hierarchically semiseparable representations. *SIAM J. Matrix Anal. Appl.* **28** (2006) 603–622.

[16] Z. Chen and J. Zou, Finite element methods and their convergence for elliptic and parabolic interface problems. *Numer. Math.* **79** (1998) 175–202.

[17] Z. Chen, K. Li and X. Xiang, An adaptive high-order unfitted finite element method for elliptic interface problems. *Numer. Math.* **149** (2021) 507–548.

[18] P. Ciarlet, Jr., Analysis of the Scott–Zhang interpolation in the fractional order Sobolev spaces. *J. Numer. Math.* **21** (2013) 173–180.

[19] L. Demkowicz and A. Buffa, H^1, $H(\mathrm{curl})$ and $H(\mathrm{div})$-conforming projection-based interpolation in three dimensions. Quasi-optimal p-interpolation estimates. *Comput. Methods Appl. Mech. Eng.* **194** (2005) 267–296.

[20] E. Di Nezza, G. Palatucci and E. Valdinoci, Hitchhiker's guide to the fractional Sobolev spaces. *Bull. Sci. Math.* **136** (2012) 521–573.

[21] Z. Ding, A proof of the trace theorem of Sobolev spaces on Lipschitz domains. *Proc. Amer. Math. Soc.* **124** (1996) 591–600.

[22] T. Dupont and L.R. Scott, Polynomial approximation of functions in Sobolev spaces. *Math. Comput.* **34** (1980) 441–463.

[23] G. Dziuk and C.M. Elliott, Finite element methods for surface PDEs. *Acta Numer.* **22** (2013) 289–396.

[24] A. Ern and J.-L. Guermond, Evaluation of the condition number in linear systems arising in finite element approximations. *ESAIM: M2AN* **40** (2006) 29–48.

[25] H. Gao, Y. Huang and F. Abraham, Continuum and atomisitic studies of intersonic crack propagation. *J. Mech. Phys. Solids* **49** (2001) 2113–2132.

[26] L.V. Gibiansky and O. Sigmund, Multiphase composites with extremal bulk modulus. *J. Mech. Phys. Solids* **48** (2000) 461–498.

[27] P. Grisvard, Elliptic Problems in Nonsmooth Domains. *Classics in Applied Mathematics*. Vol. 69. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA (2011).

[28] 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.

[29] R. Guo, Y. Lin and J. Zou, Solving two dimensional $\mathbf{H}(\mathbf{curl})$-elliptic interface systems with optimal convergence on unfitted meshes. Preprint arXiv:2011.11905 (2020).

[30] C. Gürkan and A. Massing, A stabilized cut discontinuous Galerkin framework for elliptic boundary value and interface problems. *Comput. Methods Appl. Mech. Eng.* **348** (2019) 466–499.

[31] J. Guzmán and M. Olshanskii, Inf-sup stability of geometrically unfitted Stokes finite elements. *Math. Comput.* **87** (2018) 2091–2112.

[32] Y. Han, X.-P. Wang and X. Xie, An interface/boundary-unfitted extended HDG method for linear elasticity problems. *J. Sci. Comput.* **94** (2023) 61.

[33] 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.

[34] A. Hansbo and P. Hansbo, A finite element method for the simulation of strong and weak discontinuities in solid mechanics. *Comput. Methods Appl. Mech. Eng.* **193** (2004) 3523–3540.

[35] R. Hiptmair, J. Li and J. Zou, Convergence analysis of finite element methods for $H(\mathrm{div}; \Omega)$-elliptic interface problems. *J. Numer. Math.* **18** (2010) 187–218.

[36] P. Huang, H. Wu and Y. Xiao, An unfitted interface penalty finite element method for elliptic interface problems. *Comput. Methods Appl. Mech. Eng.* **323** (2017) 439–460.

[37] A. Johansson and M.G. Larson, A high order discontinuous Galerkin Nitsche method for elliptic problems with fictitious boundary. *Numer. Math.* **123** (2013) 607–628.

[38] D.Y. Kwak, S. Jin and D. Kyeong, A stabilized P_1-nonconforming immersed finite element method for the interface elasticity problems. *ESAIM: M2AN* **51** (2017) 187–207.

[39] P.H. Leo, J.S. Lowengrub and Q. Nie, Microstructural evolution in orthotropic elastic media. *J. Comput. Phys.* **157** (2000) 44–88.

[40] Z. Li, The immersed interface method using a finite element formulation. *Appl. Numer. Math.* **27** (1998) 253–267.

[41] Z. Li and K. Ito, The Immersed Interface Method: Numerical Solutions of PDEs Involving Interfaces and Irregular Domains. *Frontiers in Applied Mathematics*. Vol. 33. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA (2006).

[42] R. Li and F. Yang, A least squares method for linear elasticity using a patch reconstructed space. *Comput. Methods Appl. Mech. Eng.* **363** (2020) 112902.

[43] R. Li, Q. Liu and F. Yang, A reconstructed discontinuous approximation on unfitted meshes to $H(\mathrm{curl})$ and $H(\mathrm{div})$ interface problems. *Comput. Methods Appl. Mech. Eng.* **403** (2023) 115723.

[44] L. Lin, J. Lu and L. Ying, Fast construction of hierarchical matrix representation from matrix-vector multiplication. *J. Comput. Phys.* **230** (2011) 4071–4087.

[45] T. Lin, D. Sheen and X. Zhang, A locking-free immersed finite element method for planar elasticity interface problems. *J. Comput. Phys.* **247** (2013) 228–247.

[46] T. Lin, D. Sheen and X. Zhang, A nonconforming immersed finite element method for elliptic interface problems. *J. Sci. Comput.* **79** (2019) 442–463.

[47] J.-L. Lions and E. Magenes, Non-Homogeneous Boundary Value Problems and Applications. Vol. I. Springer-Verlag, New York-Heidelberg (1972). Translated from the French by P. Kenneth, Die Grundlehren der mathematischen Wissenschaften, Band 181. MR 0350177.

[48] H. Liu, L. Zhang, X. Zhang and W. Zheng, Interface-penalty finite element methods for interface problems in H^1, $\mathbf{H}(\mathrm{curl})$, and $\mathbf{H}(\mathrm{div})$. *Comput. Methods Appl. Mech. Eng.* **367** (2020) 113137.

[49] R. Massjung, An unfitted discontinuous Galerkin method applied to elliptic interface problems. *SIAM J. Numer. Anal.* **50** (2012) 3134–3162.

[50] F.-H. Rouet, X.S. Li, P. Ghysels and A. Napov, A distributed-memory package for dense hierarchically semi-separable matrix computations using randomization. *ACM Trans. Math. Software* **42** (2016) 35.

[51] L.R. Scott and S. Zhang, Finite element interpolation of nonsmooth functions satisfying boundary conditions. *Math. Comput.* **54** (1990) 483–493.

[52] G. Starke, A. Schwarz and J. Schröder, Analysis of a modified first-order system least squares method for linear elasticity with improved momentum balance. *SIAM J. Numer. Anal.* **49** (2011) 1006–1022.

[53] H. Wu and Y. Xiao, An unfitted *hp*-interface penalty finite element method for elliptic interface problems. *J. Comput. Math.* **37** (2019) 316–339.

[54] Y. Xi, J. Xia and R. Chan, A fast randomized eigensolver with structured LDL factorization update. *SIAM J. Matrix Anal. Appl.* **35** (2014) 974–996.

[55] J. Xia, S. Chandrasekaran, M. Gu and X.S. Li, Fast algorithms for hierarchically semiseparable matrices. *Numer. Linear Algebra Appl.* **17** (2010) 953–976.

[56] J. Xia, Z. Li and X. Ye, Effective matrix-free preconditioning for the augmented immersed interface method. *J. Comput. Phys.* **303** (2015) 295–312.

[57] F. Yang and X. Xie, An unfitted finite element method by direct extension for elliptic problems on domains with curved boundaries and interfaces. *J. Sci. Comput.* **93** (2022) 75.

[58] X. Zhang, High order interface-penalty finite element methods for elasticity interface problems in 3D. *Comput. Math. Appl.* **114** (2022) 161–170.