



Research Paper

General enrichments of stable GFEM for interface problems: Theory and extreme learning machine construction [☆]

Dongmei Wang ^a, Hengguang Li ^b, Qinghui Zhang ^{a,*}

^a School of Science, Harbin Institute of Technology, Shenzhen, 518000, PR China

^b Department of Mathematics, Wayne State University, Detroit, MI 48202, USA

ARTICLE INFO

Keywords:

SGFEM
Interface
Complex geometry
Machine learning
Extreme learning machine

ABSTRACT

Generalized finite element methods (GFEMs), when applied to interface problems (IPs), need to be enriched with special functions to enhance approximation accuracy. These functions include distance functions, one-side distance functions, level set functions, and exponential forms of level set function. For the IP with geometrically complex interface curves, computation of the distance function or level set function could be challenging, and algorithms of computational geometry are usually involved. Moreover, theoretical analysis on optimal convergence of the GFEM enriched by these functions has not been fully investigated. In this study we propose a general enrichment scheme, based on which all the aforementioned enrichments can be viewed as special examples. We prove that a stable GFEM (SGFEM) with such a new enrichment scheme reaches the optimal convergence rate. Most importantly, the new scheme provides an instruction to construct machine learning (ML) based enrichments, which advances the ability of GFEM to handle geometrically complex interfaces. Two ML methods, deep neural network (DNN) and extreme learning machine (ELM), are studied. Among them, the ELM is highly suggested because it exhibits high accuracy for the interface curve with complex geometries. The learning dimension for the ML is one dimension less than that of the domain so that the proposed ML algorithm can be implemented efficiently. The numerical experiments demonstrate that the SGFEM with the ELM enrichment achieves the optimal convergence rates for the IP, as predicted theoretically.

1. Introduction

In interface modeling problems fundamental domains are divided into portions by various interface curves. Partial differential equations (PDEs) in these portions are equipped with different equation coefficients so that the solution of the PDE and its gradients may involve discontinuities across the interface. Such modeling problems exist in many engineering computations and physical phenomena, such as bi-materials, fluid-structure, multi-physics, and contact problems [1,2]. It is well-known that conventional numerical methods, e.g., finite element methods (FEM), finite volume methods, have a difficulty in mesh generations for the interface problem; the mesh has to be fitted to the interface to achieve reasonable accuracy [3–5]. As a result, the mesh generation consumes major com-

[☆] This research was partially supported by the Natural Science Foundation of China under grant 12471370 and Guangdong Provincial Natural Science Foundation of China under grant 2022A1515011187.

* Corresponding author.

E-mail addresses: li@wayne.edu (H. Li), zhangqh@hit.edu.cn, zhangqh6@mail.sysu.edu.cn (Q. Zhang).

<https://doi.org/10.1016/j.apnum.2025.03.009>

Received 7 March 2024; Received in revised form 21 October 2024; Accepted 31 March 2025

Available online 4 April 2025

0168-9274/© 2025 IMACS. Published by Elsevier B.V. All rights are reserved, including those for text and data mining, AI training, and similar technologies.

putational resources and may even cause failure of the numerical methods, especially in the interface problems with geometrically complex or evolving interface curves.

To reduce the burden of mesh generation, unfitted mesh has received extensive extension for the FEM to solve the interface problem in the last decades. A mesh is called unfitted if it is simple, fixed, and independent of interface positions [4,6,7]. The commonly-used unfitted FEMs include the immersed FEM [8–10], the cut FEM [11,12], the penalty FEMs [13,14,44,15], the generalized/extended FEM (GFEM/XFEM) [16–25], and many others [26–31]. The GFEM/XFEM augments the standard FEM with special functions that mimic local features of exact solutions to solve complicated non-smooth engineering problems [32–34]. These special functions are “pasted” by a partition of unity method [35,36]. For the interface problem, the merit of GFEM/XFEM consists in developing conforming methods free from penalty parameters or stabilization terms and constructing shape functions independent of the equation coefficients [6,19,23,20,37]. The stable GFEM (SGFEM) is a stable version of GFEM/XFEM. A GFEM/XFEM is stable if it maintains optimal convergence rates, and is well conditioned and robust in that the convergence and conditioning do not deteriorate when the interface approaches boundaries of meshes [38,39,19,40,24,41–43]. The applications of SGFEM to the IP are referred to [19,25,24,21,6].

The enriched local function is the major feature of the GFEM/XFEM and the SGFEM. For the interface problem the enrichments include distance function D [23,33,19], one-side distance function \tilde{D} [21,6,37] that is a cut distance function, absolute value of level set function [44,33], $|\varphi|$, and its modified version $|I_h\varphi| - I_h|I_h\varphi|$ [20,45,33], where I_h is the standard FE interpolation operator based on the unfitted mesh, exponential forms of D or φ [46,47]. There are also other enrichments used in the GFEM/XFEM, for instance, the Heaviside function [48,33]. However, the Heaviside function is discontinuous, and proper penalty schemes are necessary for stabilization. We do not relate the Heaviside enrichment in this study because we study conforming methods that are free from the penalty or stabilization techniques. On the one hand, except the distance function and one-side distance functions [19,6,21,45], there are no theories indicating whether the other enrichments can make the GFEM/XFEM achieve optimal convergence rates for the interface problem. A unified theoretical analysis on these enrichments is still missing. In other words, is it possible to have a general theory for these enrichments such that the optimal convergence rate can be proven? On the other hand, the evaluation of distance or level set functions involves certain iteration algorithms in implementation. This can significantly increase the computational cost, especially for the interface problem with complex geometries. This study is mainly focused on addressing these two issues.

We first propose a general condition for the enrichment \mathcal{G} in what follows:

$$\mathcal{G} \text{ is smooth in each portion cut by the interface } \Gamma, [\mathcal{G}]_\Gamma = 0, \left[\frac{\partial \mathcal{G}}{\partial \bar{n}_\Gamma} \right]_\Gamma \text{ has positive low bound,} \tag{1.1}$$

where $[v]_\Gamma$ represents the jump of v across Γ , and \bar{n}_Γ is a unit vector normal to Γ . We then prove that if an enrichment \mathcal{G} satisfies the condition (1.1), the SGFEM with \mathcal{G} shall achieve the optimal convergence rate, $O(h)$, in the energy norm. The aforementioned enrichments are covered by the condition (1.1), and thus their optimal convergence rates are proven in this paper in a unified way. More importantly, the condition (1.1) motivates us to develop machine learning (ML) enrichments [49]. A loss function is established using the condition (1.1), and the ML enrichments are derived by minimizing the loss function through certain training process. Merits of the ML enrichments are the following: (a) the evaluation of function is direct and does not need the iteration algorithm as used in the evaluation of distance function and level set function, and (b) the complex interface geometries can be resolved efficiently by means of nonlinear approximation ability of the ML functions. The ML techniques [49] have achieved great success in image recognition, speech recognition, and natural language processing [49,50]. In recent years the ML has been extensively applied to numerical solutions to PDEs, including the physics-informed neural network and DGM [51,52], energy methods [53,54], weak adversarial neural network [55], Nitsche based method [56], radial basis neural network [57], and many others [58–62]. The applications of ML to the interface problem are referred to [63–66]. This study enriches the SGFEM with the ML technique that produces the ML enrichment based on the general condition (1.1). We study two typical ML methods, deep neural network (DNN) [49] and the extreme learning machine (ELM) [67]. The ELM has been applied to the numerical solution to PDE, see [68–71] for instance. The ELM achieves high approximation accuracy in comparison with the DNN for the PDE with smooth solutions. In this study we find that the ELM can realize the condition (1.1) with very high accuracy even for the geometrically complex interfaces. Numerical experiments show that the SGFEM with the ELM enrichment achieves the optimal convergence rate, $O(h)$, as predicted by the theory. The training dimension of the ML enrichment is one dimension less than the space dimension, and the enrichments are only used at nodes of elements cut by the interface. Therefore, the employment of ML enrichment does not increase the overall computational complexity of the SGFEM.

The paper is organized as follows. The model problem is described in Section 2. The conventional FEM, GFEM and SGFEM are reviewed in Section 3, where we propose a general enrichment condition. The optimal convergence rate of the SGFEM with the general enrichment is proven in Section 4. Based on the general enrichment condition, the ML enrichments are constructed in Section 5. Two ML techniques, DNN and ELM, are considered and compared in Section 5. The numerical experiments and concluding remarks are presented in Sections 6 and 7, respectively.

2. Model problem

For a domain Δ in \mathbb{R}^2 , an integer m , and $1 \leq q \leq \infty$, we denote the usual Sobolev spaces by $W^{m,q}(\Delta)$ with norm $\|\cdot\|_{W^{m,q}(\Delta)}$ and semi-norm $|\cdot|_{W^{m,q}(\Delta)}$. The space $W^{m,q}(\Delta)$ will be represented by $H^m(\Delta)$ for $q = 2$ and $L^q(\Delta)$ when $m = 0$, respectively.

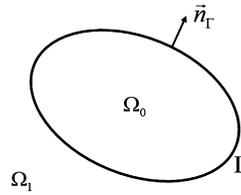


Fig. 1. The domains with a curved interface.

We consider a bounded and simply connected domain $\Omega \subset \mathbb{R}^2$ with a piecewise smooth boundary $\partial\Omega$. Let Γ be an interface that divides Ω into two sub-domains Ω_0 and Ω_1 such that $\overline{\Omega} = \overline{\Omega_0} \cup \overline{\Omega_1}$, $\Omega_0 \cap \Omega_1 = \emptyset$, and $\Gamma = \overline{\Omega_0} \cap \overline{\Omega_1}$. In this study, we consider the case that Γ is smooth and $\Gamma \cap \partial\Omega = \emptyset$, as shown in Fig. 1.

A point in the Cartesian coordinate system of \mathbb{R}^2 is denoted as $P = (x, y)$. Let a be a positive, piecewise-constant function given by

$$a(P) = \begin{cases} a_0, & P \in \Omega_0, \\ a_1, & P \in \Omega_1, \end{cases} \tag{2.1}$$

where a_0 and a_1 are positive constants, $0 < \zeta_0 \leq a_r \leq \zeta_1 < \infty$, $r = 0, 1$, and $\zeta_0, \zeta_1 \in \mathbb{R}$.

We are interested in the solution u of the interface problem:

$$\begin{aligned} -\nabla \cdot (a \nabla u) &= f, \text{ in } \Omega, \\ a \frac{\partial u}{\partial \vec{n}_b} &= g, \text{ on } \partial\overline{\Omega}, \end{aligned} \tag{2.2}$$

subject to jump conditions on the interface Γ

$$[u]_\Gamma = 0, \text{ on } \Gamma, \tag{2.3}$$

$$\left[a \frac{\partial u}{\partial \vec{n}_\Gamma} \right]_\Gamma = 0, \text{ on } \Gamma, \tag{2.4}$$

where \vec{n}_b and \vec{n}_Γ denote the unit outward normal to the boundary $\partial\Omega$ and to the interface Γ directed towards Ω_1 , respectively. The notation $[v]_\Gamma := v_0 - v_1$ defines the jump of a quantity v along the interface Γ , where $v_r := v|_{\Omega_r}$, $r = 0, 1$. The data f, g are given such that the solution $u \in \mathbb{M}_2$, where \mathbb{M}_2 is defined by

$$\mathbb{M}_2 := \{ u : u|_{\Omega_r} \in H^2(\Omega_r), r = 0, 1 \text{ and } \|\partial^\alpha u\|_{L^\infty(\Gamma)} < \infty, |\alpha| \leq 1 \} \tag{2.5}$$

with a norm

$$\|u\|_{\mathbb{M}_2} = \|u_0\|_{H^2(\Omega_0)} + \|u_1\|_{H^2(\Omega_1)} + \sum_{|\alpha| \leq 1} \|\partial^\alpha u\|_{L^\infty(\Gamma)}, \forall u \in \mathbb{M}_2.$$

Remark 2.1. We formulate the model problem using the homogenous interface conditions, (2.3) and (2.4). In this case, the solution is continuous, and its gradients are discontinuous. Such a discontinuity is referred to as weak discontinuity [11,12,23,25,9,19]. The method in this study can be extended to the nonhomogeneous interface condition, $[u]_\Gamma \neq 0$, $\left[a \frac{\partial u}{\partial \vec{n}_\Gamma} \right]_\Gamma \neq 0$ using a lifting technique like in [72,73]. The construction of shape functions is independent of the coefficient a so that the method in this paper can be applied to anisotropic interface coefficients [22] and vector-valued interface problems in a unified way. We present the model problem (2.2) with a Neumann boundary so that the solution u is unique up to a constant. Additional restrictive conditions are needed to attain a particular solution. For instance, when the solution value is given on a point in the domain, the solution is then unique. We use the Neumann boundary for simplicity of presentations. The proposed method can be used for the other boundary conditions directly, e.g., the Dirichlet or Robin conditions. \square

We define $\mathcal{E}(\Omega)$ to be the energy space with respect to the homogeneous IP given by

$$\mathcal{E}(\Omega) := \{ u \in H^1(\Omega) : \|u\|_{\mathcal{E}(\Omega)}^2 := \int_{\Omega} a \nabla u \cdot \nabla u \, dP < \infty \text{ and } [u]_\Gamma = 0 \text{ on } \Gamma \}. \tag{2.6}$$

The variational formulation associated with (2.2) based on a finite-dimensional subspace $\mathbb{S}_h \in \mathcal{E}(\Omega)$ is the following:

$$\text{Find } u_h \in \mathbb{S}_h \text{ such that } B(u_h, v_h) = L(v_h), \forall v_h \in \mathbb{S}_h, \tag{2.7}$$

where the bilinear form $B(\cdot, \cdot)$ and $L(\cdot)$ are defined by

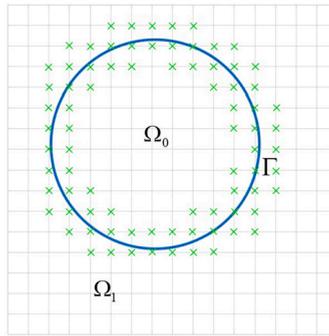


Fig. 2. The interface, mesh, and enriched nodes.

$$B(v, w) := \int_{\Omega} a \nabla v \cdot \nabla w \, dP, \quad L(v) := \int_{\Omega} f v \, dx + \int_{\partial\Omega} g v \, ds, \quad \forall v, w \in \mathcal{E}(\Omega). \tag{2.8}$$

According to the standard Céa’s Lemma we have

$$\|u - u_h\|_{\mathcal{E}(\Omega)} \leq \min_{v_h \in \mathcal{E}(\Omega)} \|u - v_h\|_{\mathcal{E}(\Omega)}, \quad \forall v_h \in \mathbb{S}_h. \tag{2.9}$$

We next describe GFEM and SGFEM subspaces \mathbb{S}_h , based on which the optimal convergence rate in the energy norm, $O(h)$, can be obtained from (2.9).

3. GFEM, SGFEM, and proposed general enrichment condition

Let $\mathcal{T}_h = \{e_s\}$ be a quasi-uniform finite element mesh with mesh size $0 < h < 1$; the finite elements e_s can be triangles or quadrilaterals. It is noted that the mesh \mathcal{T}_h is simple, fixed, and does not need to fit the interface Γ . The set FE nodes associated with the mesh \mathcal{T}_h is denoted by $\{P_i\}_{i \in I_h}$, where I_h is the index set. $\varphi_i, i \in I_h$ are the standard linear (bilinear for quadrilateral element) FE hat function φ_i . The closure of support of φ_i is denoted by ω_i . Since the mesh is quasi-uniform, we have

$$\|\varphi_i\|_{L^\infty(\Omega)} = \|\varphi_i\|_{L^\infty(\omega_i)} \leq 1, \quad \|\nabla \varphi_i\|_{L^\infty(\Omega)} = \|\nabla \varphi_i\|_{L^\infty(\omega_i)} \leq Ch^{-1}, \tag{3.1}$$

where the positive constant C is independent of h and i . It is well known that $\{\varphi_i\}_{i \in I_h}$ form a partition of unity (PU) [35,36], i.e.,

$$\sum_{i \in I_h} \varphi_i \equiv 1, \quad \text{in } \Omega.$$

The standard FEM subspace is given by

$$\mathbb{S}_h := \mathbb{S}_{FEM} = \text{span}\{\varphi_i(x) : i \in I_h\}. \tag{3.2}$$

It is well-known that the FEM (3.2) cannot yields highly accurate approximations because the mesh does not fit Γ [4,19].

The generalized or extended FEM (GFEM/XFEM) [32–34] is a typical technique to approximate the non-smooth problems by augmenting the FEM space \mathbb{S}_{FEM} (3.2) by non-polynomial enrichment space \mathbb{S}_{ENR} based on a partition of unity method (PUM) [35,36]. For the interface problem, the distance function to Γ (or the absolute of level set function $\phi, |\phi|$) [33,23,19,20] serves as the enriched function for the homogeneous IP, which is defined as follows:

$$D(P) = \text{dist}(P, \Gamma), \tag{3.3}$$

and the associated approximate subspace is

$$\mathbb{S}_h = \mathbb{S}_{FEM} \oplus \mathbb{S}_{ENR} \quad \text{and} \quad \mathbb{S}_{ENR} = \text{span}\{\varphi_i D : i \in I_{h, enr}^\Gamma\}, \tag{3.4}$$

where

$$I_{h, enr}^\Gamma = \{i \in I_h : P_i \in e_s \text{ where } e_s \cap \Gamma \neq \emptyset\}$$

is a set of enriched nodes, see Fig. 2. The GFEM (3.4) causes sub-optimal convergence rates due to so-called blending element errors [33], and the optimal rate can be restored by enlarging the enrichment set $I_{h, enr}^\Gamma$ or using corrected XFEM ideas [33].

There are various enrichments in the literature for the IP including a one-side distance function [21,6,37]

$$\tilde{D} = \begin{cases} D, & \in \Omega_0, \\ 0, & \in \Omega_1, \end{cases} \tag{3.5}$$

exponential form of ϕ [46,47],

$$\phi_{\text{exp}} = \begin{cases} 1 - e^{\mu\phi}, & \in \Omega_0, \\ -\phi, & \in \Omega_1, \end{cases} \tag{3.6}$$

where ϕ is a level set function [33,44] or a modified distance function, and a modified version of ϕ [20,45,33],

$$|\mathcal{I}_h\phi| - \mathcal{I}_h|\mathcal{I}_h\phi|, \tag{3.7}$$

where \mathcal{I}_h is the standard FE interpolation operator based on (3.2).

The stable GFEM (SGFEM) is a stable version of GFEM/XFEM, which improves conditioning of GFEM/XFEM. The approximate subspace of the SGFEM for the homogeneous IP [19,21,6] is given by

$$\mathbb{S}_h = \mathbb{S}_{FEM} \oplus \mathbb{S}_{ENR} \text{ and } \mathbb{S}_{ENR} = \text{span}\{\varphi_i(\xi - \mathcal{I}_h\xi) : i \in I_{h, \text{enr}}^\Gamma\}, \tag{3.8}$$

where ξ can be D or \tilde{D} . The convergence of SGFEM (3.8) was proven in [21,6], and we present it here.

Theorem 3.1. *Suppose that $u \in \mathbb{M}_2$ is the solution of the homogeneous IP ((2.2)-(2.4), and $u_{SG,h}$ is the SGFEM solution of (2.7) based on the finite-dimensional subspace \mathbb{S}_h (3.8), then there exists $C > 0$ independent of h such that*

$$\|u - u_{SG,h}\|_{\mathcal{E}(\Omega)} \leq Ch\|u\|_{\mathbb{M}_2}. \tag{3.9}$$

It was also shown in [21,6] that the SGFEM (3.8) is stable and robust in the sense that a scaled condition number (SCN) of stiffness matrices is $O(h^{-2})$ that is of same order as the FEM, and the convergence and SCN do not deteriorate as interface approaches boundaries of elements.

Remark 3.1. Except the enrichments (3.3) and (3.5) in SGFEM [6,21,45], the optimal convergence rates of other enrichments have not been investigated theoretically. Below, we propose a general enrichment for the SGFEM of the interface problem and prove the optimal convergence rate of this general enrichment. Then the above-mentioned enrichments can be viewed as the special instances so that their theoretical optimal convergence rates are obtained in a unified way. \square

A general enrichment condition By analyzing D and $|\phi|$ we find out that their key feature is that they are continuous across Γ and their derivatives normal to Γ have nonzero jumps across Γ , namely,

$$[D]_\Gamma = 0 \text{ and } \left[\frac{\partial D}{\partial \vec{n}_\Gamma} \right]_\Gamma \neq 0.$$

This feature also holds for the one-side distance function (3.5) and the exponential form of ϕ , (3.6). Such an observation motivates us to propose a general enrichment \mathcal{G} for the homogeneous IP as follows:

$$\mathcal{G} \in W^{2,\infty}(\bar{\Omega}_r), r = 0, 1, [\mathcal{G}]_\Gamma = 0, 0 < \kappa_1 \leq \left| \left[\frac{\partial \mathcal{G}}{\partial \vec{n}_\Gamma} \right]_\Gamma \right| \leq \kappa_2, \tag{3.10}$$

where $\kappa_2 > \kappa_1 > 0$ are constants independent of h and i . The approximation subspace of the SGFEM enriched by (3.10) is given by

$$\mathbb{S}_h = \mathbb{S}_{FEM} \oplus \mathbb{S}_{ENR} \text{ and } \mathbb{S}_{ENR} = \text{span}\{\phi_i(\mathcal{G} - \mathcal{I}_h\mathcal{G}) : i \in I_{h, \text{enr}}^\Gamma\}. \tag{3.11}$$

We prove in next section that the SGFEM (3.11) with the general enrichment \mathcal{G} can achieve the optimal convergence rate $O(h)$. As a consequence, the above-mentioned enrichments, D , $|\phi|$, (3.5), (3.6), satisfy the condition (3.10), and the SGFEM with these enrichments will also be proven to be optimally convergent. We mention that the enrichment (3.7) does not satisfy (3.10). However the optimal convergence of (3.7) can be also proven using the idea in this paper and the technique in [45].

Remark 3.2. Another important aspect of (3.10) is to instruct us to develop ML based enrichment for the IP. A loss function will be established based of (3.10) to train the DL functions to meet the condition (3.10). The ML enrichment possesses a great potential advantage for the IP with geometrically complex interface curves thanks to powerful nonlinear expression ability of the ML technique. We will execute this in the Section 5. \square

4. Theoretical analysis on the optimal convergence rate

We then present an approximation result of \mathcal{G} , which is important to estimate the convergence errors in situations of curved interfaces. Denote $\mathcal{G}_r = \mathcal{G}|_{\bar{\Omega}_r}, r = 0, 1$.

Lemma 4.1. For each $i \in I_{h, enr}^\Gamma$ and a point $Q_i \in \omega_i \cap \Gamma$, let

$$\tilde{\mathcal{G}}_i(P) = \mathcal{G}(Q_i) + \frac{\partial \mathcal{G}}{\partial \vec{n}_\Gamma^\perp}(Q_i) (P - Q_i) \cdot \vec{n}_\Gamma^\perp + \begin{cases} \frac{\partial \mathcal{G}_0}{\partial \vec{n}_\Gamma}(Q_i) (P - Q_i) \cdot \vec{n}_\Gamma, & P \in \bar{\Omega}_0 \cap \omega_i \\ \frac{\partial \mathcal{G}_1}{\partial \vec{n}_\Gamma}(Q_i) (P - Q_i) \cdot \vec{n}_\Gamma, & P \in \Omega_1 \cap \omega_i, \end{cases} \tag{4.1}$$

then there is a constant C independent of i and h , such that

$$\|\mathcal{G} - \tilde{\mathcal{G}}_i\|_{W^{l, \infty}(\omega_i)} \leq Ch^{2-l}, \quad l = 0, 1. \tag{4.2}$$

Proof. Let $T_i \mathcal{G}_r$ be the linear Taylor polynomial of \mathcal{G}_r at Q_i in $\bar{\Omega}_r \cap \omega_i$, $r = 0, 1$, respectively, written in the local orthogonal coordinate $(\vec{n}_\Gamma^\perp, \vec{n}_\Gamma)$, i.e.,

$$T_i \mathcal{G}_r(P) = \mathcal{G}_r(Q_i) + \frac{\partial \mathcal{G}_r}{\partial \vec{n}_\Gamma^\perp}(Q_i) (P - Q_i) \cdot \vec{n}_\Gamma^\perp + \frac{\partial \mathcal{G}_r}{\partial \vec{n}_\Gamma}(Q_i) (P - Q_i) \cdot \vec{n}_\Gamma, \quad P \in \bar{\Omega}_r \cap \omega_i, \quad r = 0, 1. \tag{4.3}$$

Using the Taylor theorem [74] and (3.10), we obtain

$$\|\mathcal{G}_r - T_i \mathcal{G}_r\|_{W^{l, \infty}(\bar{\Omega}_r \cap \omega_i)} \leq Ch^{2-l}, \quad l = 0, 1, \quad r = 0, 1.$$

Since $[\mathcal{G}]_\Gamma = 0$, we have

$$\mathcal{G}_0(Q_i) = \mathcal{G}_1(Q_i) = \mathcal{G}(Q_i), \quad \frac{\partial \mathcal{G}_0}{\partial \vec{n}_\Gamma^\perp}(Q_i) = \frac{\partial \mathcal{G}_1}{\partial \vec{n}_\Gamma^\perp}(Q_i) = \frac{\partial \mathcal{G}}{\partial \vec{n}_\Gamma^\perp}(Q_i). \tag{4.4}$$

According to (4.4), (4.1), and (4.3), we get $\tilde{\mathcal{G}}_i(P) = T_i \mathcal{G}_0$ for $P \in \bar{\Omega}_0 \cap \omega_i$, $\tilde{\mathcal{G}}_i(P) = T_i \mathcal{G}_1$ for $P \in \Omega_1 \cap \omega_i$, and thus

$$\begin{aligned} \|\mathcal{G} - \tilde{\mathcal{G}}_i\|_{W^{l, \infty}(\omega_i)} &= \max\{\|\mathcal{G} - \tilde{\mathcal{G}}_i\|_{W^{l, \infty}(\bar{\Omega}_0 \cap \omega_i)}, \|\mathcal{G} - \tilde{\mathcal{G}}_i\|_{W^{l, \infty}(\Omega_1 \cap \omega_i)}\} \\ &= \max\{\|\mathcal{G} - T_i \mathcal{G}_0\|_{W^{l, \infty}(\bar{\Omega}_0 \cap \omega_i)}, \|\mathcal{G} - T_i \mathcal{G}_1\|_{W^{l, \infty}(\Omega_1 \cap \omega_i)}\} \\ &\leq Ch^{2-l}, \quad l = 0, 1, \end{aligned}$$

which is the desired result (4.2). \square

We next present a technical result that will be used in proof of the main theorem. We continuously extend u_0 and u_1 to the whole domain Ω to get functions \tilde{u}_0 and \tilde{u}_1 in $H^2(\Omega)$ such that

$$\tilde{u}_r = u_r \text{ on } \bar{\Omega}_r \text{ and } \|\tilde{u}_r\|_{H^2(\Omega)} \leq C \|u_r\|_{H^2(\Omega_r)}, \quad r = 0, 1, \tag{4.5}$$

where C is a positive constant independent of h (see Theorem 1.4.5 in [74]).

Lemma 4.2. For each $i \in I_{h, enr}^\Gamma$, there is a linear polynomial z_i and a constant σ_i such that

$$\left. \begin{aligned} |u - z_i - \sigma_i \tilde{\mathcal{G}}_i|_{H^1(\omega_i)}^2 \\ |I_h(u - z_i - \sigma_i \tilde{\mathcal{G}}_i)|_{H^1(\omega_i)}^2 \end{aligned} \right\} \leq Ch^{4-2l} (\|\tilde{u}_0\|_{H^2(\omega_i)}^2 + \|\tilde{u}_1\|_{H^2(\omega_i)}^2) + Ch^{6-2l} \|\nabla u\|_{L^\infty(\Gamma)}^2, \quad l = 0, 1, \tag{4.6}$$

where ω_i is the patch associated with the node P_i .

Proof. For each $i \in I_{h, enr}^\Gamma$, we consider a point $Q_i \in \omega_i \cap \Gamma$. Let $T_i \tilde{u}_r$ be the linear Taylor polynomial of \tilde{u}_r at Q_i written in the local orthogonal coordinate $(\vec{n}_\Gamma^\perp, \vec{n}_\Gamma)$, namely,

$$T_i \tilde{u}_r(P) = \tilde{u}_r(Q_i) + \frac{\partial \tilde{u}_r}{\partial \vec{n}_\Gamma^\perp}(Q_i) (P - Q_i) \cdot \vec{n}_\Gamma^\perp + \frac{\partial \tilde{u}_r}{\partial \vec{n}_\Gamma}(Q_i) (P - Q_i) \cdot \vec{n}_\Gamma, \quad P \in \omega_i, \quad r = 0, 1.$$

Since u is continuous on Γ , we have $\tilde{u}_0(Q_i) = \tilde{u}_1(Q_i) = u(Q_i)$ and $\frac{\partial \tilde{u}_0}{\partial \vec{n}_\Gamma^\perp}(Q_i) = \frac{\partial \tilde{u}_1}{\partial \vec{n}_\Gamma^\perp}(Q_i) = \frac{\partial u}{\partial \vec{n}_\Gamma^\perp}(Q_i)$. Therefore,

$$T_i \tilde{u}_r(P) = u(Q_i) + \frac{\partial u}{\partial \vec{n}_\Gamma^\perp}(Q_i) (P - Q_i) \cdot \vec{n}_\Gamma^\perp + \frac{\partial \tilde{u}_r}{\partial \vec{n}_\Gamma}(Q_i) (P - Q_i) \cdot \vec{n}_\Gamma, \quad P \in \omega_i, \quad r = 0, 1. \tag{4.7}$$

Let

$$\sigma_i = \frac{\frac{\partial \tilde{u}_1}{\partial \vec{n}_\Gamma}(Q_i) - \frac{\partial \tilde{u}_0}{\partial \vec{n}_\Gamma}(Q_i)}{\frac{\partial \mathcal{G}_1}{\partial \vec{n}_\Gamma}(Q_i) - \frac{\partial \mathcal{G}_0}{\partial \vec{n}_\Gamma}(Q_i)}$$

and

$$z_i(P) = u(Q_i) + \frac{\partial u}{\partial \bar{n}_\Gamma^\perp}(Q_i) (P - Q_i) \cdot \bar{n}_\Gamma^\perp + \left(\frac{\partial \tilde{u}_1}{\partial \bar{n}_\Gamma}(Q_i) - \sigma_i \frac{\partial \mathcal{G}_1}{\partial \bar{n}_\Gamma}(Q_i) \right) (P - Q_i) \cdot \bar{n}_\Gamma - \sigma_i \left(\mathcal{G}(Q_i) + \frac{\partial \mathcal{G}}{\partial \bar{n}_\Gamma^\perp}(Q_i) (P - Q_i) \cdot \bar{n}_\Gamma^\perp \right), \quad P \in \omega_i.$$

Employing (4.1) and (4.7) we have

$$\begin{aligned} (u - z_i - \sigma_i \mathcal{G})(P) &= \sigma_i (\tilde{\mathcal{G}}_i - \mathcal{G})(P) + (u - z_i - \sigma_i \tilde{\mathcal{G}}_i)(P) \\ &= \sigma_i (\tilde{\mathcal{G}}_i - \mathcal{G})(P) + \begin{cases} (\tilde{u}_0 - T_i \tilde{u}_0)(P), & P \in \bar{\Omega}_0 \cup \omega_i, \\ (\tilde{u}_1 - T_i \tilde{u}_1)(P), & P \in \Omega_1 \cup \omega_i. \end{cases} \end{aligned} \tag{4.8}$$

Therefore,

$$\begin{aligned} |u - z_i - \sigma_i \mathcal{G}|_{H^l(\omega_i)}^2 &\leq 2|u - z_i - \sigma_i \tilde{\mathcal{G}}_i|_{H^l(\omega_i)}^2 + 2\sigma_i^2 |\mathcal{G} - \tilde{\mathcal{G}}_i|_{H^l(\omega_i)}^2 \\ &= 2|\tilde{u}_0 - z_i - \sigma_i \tilde{\mathcal{G}}_i|_{H^l(\Omega_0 \cap \omega_i)}^2 + 2|\tilde{u}_1 - z_i - \sigma_i \tilde{\mathcal{G}}_i|_{H^l(\Omega_1 \cap \omega_i)}^2 + 2\sigma_i^2 |\mathcal{G} - \tilde{\mathcal{G}}_i|_{H^l(\omega_i)}^2 \\ &= 2|\tilde{u}_0 - T_i \tilde{u}_0|_{H^l(\Omega_0 \cap \omega_i)}^2 + 2|\tilde{u}_1 - T_i \tilde{u}_1|_{H^l(\Omega_1 \cap \omega_i)}^2 + 2\sigma_i^2 |\mathcal{G} - \tilde{\mathcal{G}}_i|_{H^l(\omega_i)}^2 \\ &\leq 2|\tilde{u}_0 - T_i \tilde{u}_0|_{H^l(\omega_i)}^2 + 2|\tilde{u}_1 - T_i \tilde{u}_1|_{H^l(\omega_i)}^2 + 2\sigma_i^2 |\mathcal{G} - \tilde{\mathcal{G}}_i|_{W^{l,\infty}(\omega_i)}^2 |\omega_i| \\ &\leq 2Ch^{4-2l} \left(\|\tilde{u}_0\|_{H^2(\omega_i)}^2 + \|\tilde{u}_1\|_{H^2(\omega_i)}^2 \right) + 2C\sigma_i^2 h^{6-2l}, \quad l = 0, 1, \end{aligned} \tag{4.9}$$

where the last inequality uses the result (4.1). According to the condition (3.10), we have

$$|\sigma_i| \leq \frac{1}{\kappa_1} 2\|\nabla u\|_{L^\infty(\Gamma)} \tag{4.10}$$

Using (4.10) and (4.9), we get the first inequality of (4.6).

We next estimate the second term of (4.6). For $i \in I_{h, enr}^\Gamma$, we define $\omega_i^0 = \omega_i \cap \bar{\Omega}_0$ and $\omega_i^1 = \omega_i \setminus \omega_i^0$. Using (4.8), we have

$$\begin{aligned} \mathcal{I}_h(u - z_i - \sigma_i \mathcal{G})|_{\omega_i} &= \sum_{P_j \in \omega_i} (u - z_i - \sigma_i \mathcal{G})(P_j) \phi_j = \sum_{P_j \in \omega_i} \sigma_i (\tilde{\mathcal{G}}_i(P_j) - \mathcal{G}(P_j)) \phi_j \\ &\quad + \sum_{P_j \in \omega_i^1} (\tilde{u}_1 - T_i \tilde{u}_1)(P_j) \phi_j + \sum_{P_j \in \omega_i^0} (\tilde{u}_0 - T_i \tilde{u}_0)(P_j) \phi_j + \sum_{P_j \in \omega_i} \sigma_i (\tilde{\mathcal{G}}_i - \mathcal{G})(P_j) \phi_j. \end{aligned} \tag{4.11}$$

According to the Taylor Theorem [74] and the extension result (4.5), we have

$$|(\tilde{u}_r - T_i \tilde{u}_r)(P_j)| \leq \|\tilde{u}_r - T_i \tilde{u}_r\|_{L^\infty(\omega_i)} \leq Ch|\tilde{u}_r|_{H^2(\omega_i)}, \quad \text{for } P_j \in \omega_i, \quad r = 0, 1.$$

Therefore using (4.11), (3.1), (4.2), and the inequalities above, we get

$$\begin{aligned} |\mathcal{I}_h(u - z_i - \sigma_i \mathcal{G})|_{H^l(\omega_i)}^2 &\leq C \left[\sigma_i^2 \max_{P_j \in \omega_i} \{|\tilde{\mathcal{G}}_i(P_j) - \mathcal{G}(P_j)|^2\} + \max_{\substack{P_j \in \omega_i^r, \\ r=0,1}} \{|\tilde{u}_r - T_i \tilde{u}_r(P_j)|^2\} \right] \times \left(\sum_{P_j \in \omega_i} |\phi_j|_{H^l(\omega_i)}^2 \right) \\ &\leq Ch^{4-2l} \left(|\tilde{u}_0|_{H^2(\omega_i)}^2 + |\tilde{u}_1|_{H^2(\omega_i)}^2 \right) + C\sigma_i^2 h^{6-2l}, \quad l = 0, 1, \end{aligned} \tag{4.12}$$

where in the last inequality we use $|\phi_j|_{H^l(\omega_i)}^2 \leq Ch^{2-2l}$, $l = 0, 1$ according to (3.1). The second of (4.6) holds using the estimate of σ_i , (4.10). \square

Theorem 4.3. Suppose that $u \in \mathbb{M}_2$ is the solution of the homogeneous IP ((2.2)-(2.4)), and $\tilde{u}_{SG,h}$ is the SGFEM solution of (2.7) based on the finite-dimensional subspace \mathbb{S}_h (3.11) in which the distance function D is replaced by the general enrichment \mathcal{G} (3.10), then there exists $C > 0$ independent of h such that

$$\|u - \tilde{u}_{SG,h}\|_{\mathcal{E}(\Omega)} \leq Ch\|u\|_{\mathbb{M}_2}. \tag{4.13}$$

Proof. Denote

$$v := \mathcal{I}_h u + \sum_{i \in I_{h, enr}^\Gamma} \sigma_i \phi_i (\mathcal{G} - \mathcal{I}_h(\mathcal{G})),$$

where σ_i are the constants in Lemma 4.2. We see that v belongs to the space (3.11). Since $\{\phi_i\}_{i \in I_h}$ forms a PU and $z - \mathcal{I}_h z = 0$ on ω_i for any linear polynomial z , we have

$$\begin{aligned} u - v &= \sum_{i \in I_h} \phi_i(u - \mathcal{I}_h u) - \sum_{i \in I_{h, enr}^\Gamma} \sigma_i \phi_i(\mathcal{G} - \mathcal{I}_h(\mathcal{G})) \\ &= \sum_{i \in I_{h, NE}} \phi_i(u - \mathcal{I}_h u) + \sum_{i \in I_{h, enr}^\Gamma} \phi_i [I - \mathcal{I}_h] (u - z_i - \sigma_i \mathcal{G}) := T_1 + T_2, \end{aligned} \tag{4.14}$$

where $I_{h, NE} := I_h \setminus I_{h, enr}^\Gamma$, I is the identity operator, z_i is the linear polynomial in Lemma 4.2, and T_1, T_2 denote the two summation terms in the second equality above, respectively. Using (4.14), (4.6), and the similar argument to derive the estimates on T_1 and T_2 in Theorem 4.4 [21], we have

$$\|u - v\|_{\mathcal{E}(\Omega)}^2 \leq Ch^2 \|u\|_{\mathbb{M}_2}^2. \tag{4.15}$$

Now using the C ea’s Lemma (2.9) we have

$$\|u - u_{SG}\|_{\mathcal{E}(\Omega)} \leq \min_{v \in \mathcal{E}(\Omega)} \|u - v\|_{\mathcal{E}(\Omega)} \leq Ch \|u\|_{\mathbb{M}_2},$$

which is the desired result. \square

5. The ML enrichments of SGFEM for IP

The general enrichment (3.10) provides a practical instruction to construct the ML enrichment for the IP. The merit of ML enrichments consists in avoiding the computation of distance or level set functions, which requires certain algorithms of computational geometry. This is notable when the interface curve is geometrically complex; the ML techniques are powerful to deal with the complex geometry. Motivated from (3.10), we propose an easily-implemented form of (3.10) by the ML as follows: first construct a function \mathcal{G}_0 satisfying

$$\mathcal{G}_0 \in C^2(\bar{\Omega}), \quad \mathcal{G}_0|_\Gamma = 0, \quad \text{and} \quad \left| \frac{\partial \mathcal{G}_0}{\partial \bar{n}_\Gamma} \right|_\Gamma \geq \kappa_1, \tag{5.1}$$

and then the absolute of \mathcal{G}_0 ,

$$\mathcal{G}^{abs} = |\mathcal{G}_0| \tag{5.2}$$

or a one-side version

$$\mathcal{G}^{os} = \begin{cases} \mathcal{G}_0, & \text{in } \Omega_0, \\ 0, & \text{in } \Omega_1, \end{cases} \tag{5.3}$$

satisfy (3.11). This is because

$$\left| \frac{\partial \mathcal{G}^{abs}}{\partial \bar{n}_\Gamma} \right|_\Gamma = 2 \left| \frac{\partial \mathcal{G}_0}{\partial \bar{n}_\Gamma} \right|_\Gamma \geq 2\kappa_1, \quad \left| \frac{\partial \mathcal{G}^{os}}{\partial \bar{n}_\Gamma} \right|_\Gamma = \left| \frac{\partial \mathcal{G}_0}{\partial \bar{n}_\Gamma} \right|_\Gamma \geq \kappa_1.$$

We then approximate (5.1) using the ML technique in what follows: first construct a ML function \mathcal{G}_{ml} satisfying

$$\mathcal{G}_{ml} \in C^2(\bar{\Omega}), \quad \mathcal{G}_{ml}|_\Gamma \approx 0, \quad \text{and} \quad \left| \frac{\partial \mathcal{G}_{ml}}{\partial \bar{n}_\Gamma} \right|_\Gamma \geq \kappa_1, \tag{5.4}$$

and the associated \mathcal{G}_{ml}^{abs} and \mathcal{G}_{ml}^{os} defined like in (5.2) and (5.3) serve as the enrichments.

We establish the loss function of ML based on (5.4). To this end, we take N_Γ sampling points Q_i uniformly distributed on Γ , see Fig. 3, and denote $X = [Q_i]_{i=1}^{N_\Gamma}$ to be the input set for training of the ML. Let \mathbb{S}_{ML} be a function space generated by the ML with parameters θ (e.g., weights and bias of neural networks). The loss function we use to yield \mathcal{G}_{ml} in (5.1) is the following:

$$\text{Loss}(X; \theta) = \frac{1}{N_\Gamma} \sum_{i=1}^{N_\Gamma} \left(\left[\vartheta_\theta(Q_i) - 0 \right]^2 + \lambda \left[\frac{\partial \vartheta_\theta}{\partial \bar{n}_i}(Q_i) - \kappa_1 \right]^2 \right), \quad \forall \vartheta_\theta \in \mathbb{S}_{ML}, \tag{5.5}$$

where $\lambda > 0$ is a tradeoff parameter. Note that the direction $\bar{n}_i(Q_i)$ in (5.5) is an approximation of $\bar{n}_\Gamma(Q_i)$ in (5.4). Let P_i be the nearest point to Q_i in X , the unit vector $\bar{n}_i(Q_i)$ is defined to be orthogonal to $\overline{Q_i P_i}$ and towards Ω_1 as \bar{n}_Γ is towards, see Fig. 3. This reduces the computational complexity in the normal vector $\bar{n}_\Gamma(Q_i)$.

Minimizing the losses (5.5), we get the ML functions as follows:

$$\mathcal{G}_{ml} = \underset{\vartheta_\theta \in \mathbb{S}_{ML}}{\text{argmin}} \text{Loss}(X; \theta), \tag{5.6}$$

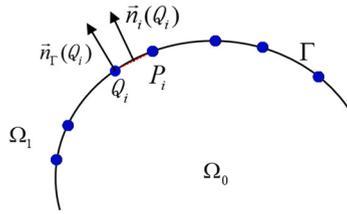


Fig. 3. The representation of Q_i , P_i , $\bar{n}_i(Q_i)$, and $\bar{n}_i(Q_i)$.

and then the associated \mathcal{G}_{ml}^{abs} and \mathcal{G}_{ml}^{os} defined like in (5.2) and (5.3) serve as the enrichments. The associated SGFEM subspace for the homogeneous IP is defined as

$$\mathbb{S}_h = \mathbb{S}_{FEM} \oplus \mathbb{S}_{ENR} \text{ and } \mathbb{S}_{ENR} = \text{span}\{\phi_i(\tilde{\mathcal{G}} - \mathcal{I}_h \tilde{\mathcal{G}}) : i \in I_{h, enr}^\Gamma\}, \tag{5.7}$$

where $\tilde{\mathcal{G}}$ is taken as \mathcal{G}_{ml}^{abs} or \mathcal{G}_{ml}^{os} . According to (4.13), the optimal convergence rate can be achieved if the ML functions \mathcal{G}_{ml} learn (5.4) accurately enough. This is verified in the numerical experiments below.

Remark 5.1. In the loss function (5.5) the ML function is only required to learn zero on Γ accurately, while $\frac{\partial \theta}{\partial \bar{n}_i}(Q_i)$ is not necessary to approximate κ_1 very accurately because $\left| \frac{\partial \mathcal{G}_{ml}}{\partial \bar{n}_i} \right|_\Gamma$ only needs a positive low bound, see (5.4). This provides us with a great flexibility to adjust the parameter λ in (5.5) to make (5.4) hold. We see in (5.5) that the dimension of learning is one dimension less than the problem dimension, two in \mathbb{R}^2 . Moreover, the assembling of stiffness matrices and the construction of ML enrichments can be implemented separately or in parallel. Therefore, yielding such a ML enrichment does not increase essential computational complexity in comparison with the SGFEM computation. \square

Remark 5.2. It is noted that there are other constructions of distance-like functions in the literature, such as the ML approach [75,76], generalized barycentric coordinates [77]. These distance-like functions are used globally in the domain, whose construction is also global. For instance, in [75,76] the values of distance function in the domain are used for the construction. This causes to compute more values of the distance function, which increases the computational complexity. By contrast, the ML function in this study only needs to satisfy the condition (5.4) on the interface Γ , and thus the computational dimension is one dimension less than that of domain computation, and there are no computations of distance function. Most importantly, according to the approximation theory (4.13), a positive low bound of $\left| \frac{\partial \mathcal{G}_{ml}}{\partial \bar{n}_i} \right|_\Gamma$ in (5.4) is needed to achieve the optimal convergence rate. This is not incorporated into the constructions in the literature. In our computation we find out that if $\left| \frac{\partial \mathcal{G}_{ml}}{\partial \bar{n}_i} \right|$ is close to zero on some locations on Γ (the positive low bound is lost), the optimal convergence rate of SGFEM will be damaged. \square

The description of deep neural network and extreme learning machine In this paper we consider two ML methods in (5.5) and (5.7), deep residual neural network (DSNN) and extreme learning machine (ELM), and in the associated enrichments, \mathcal{G}_{ml}^{abs} and \mathcal{G}_{ml}^{os} , “ml” will be taken as “dsnn” and “elm”, respectively. We describe them below.

Let $\mathbf{z}^k \in \mathbb{R}^{d_k}$ be vector of dimension d_k , and $\mathbf{z}^1 = (x, y)^T$. Let $\mathbf{W}^k \in \mathbb{R}^{d_{k+1} \times d_k}$ be $d_{k+1} \times d_k$ matrix and $\mathbf{b}^k \in \mathbb{R}^{d_{k+1}}$ be vector of dimension d_{k+1} . Denote the activation function by σ , and $\sigma(\mathbf{z}^k)$ is defined by $(\sigma(z_1^k), \dots, \sigma(z_{d_k}^k))^T$. A fully connected neural network function of L -layer is defined by

$$N_\theta := \mathbf{W}^L N_{L-1} \circ \dots \circ N_2 \circ N_1(\mathbf{z}^1) + \mathbf{b}^L, \quad N_k(\mathbf{z}^k) := \sigma(\mathbf{W}^k \mathbf{z}^k + \mathbf{b}^k), \quad k = 1, 2, \dots, L-1, \quad d_{L+1} = 1,$$

where $\theta = \{\mathbf{W}^k, \mathbf{b}^k\}_{k=1}^L$ is referred to as the parameter set, and $d_{L+1} = 1$ is because we consider a scalar problem. A fully connected neural network is exhibited in Fig. 4 Left. In this paper we use the ResNet [50], an improved version of fully connected network, which is generally used in the ML for PDE, see [53,60] for instance. The ResNet is obtained by stacking the residual blocks continuously. Each residual block consists of several fully connected layers, and its output is obtained by adding the output of the last layer and the input of the residual block. The merit of ResNet is the significant improvement of training speed and approximation error. A ResNet with four residual blocks is shown in Fig. 4 Middle. The parameters θ are derived by solving the minimization problem (5.6), which are updated using a stochastic gradient descent (SGD) method [78,49].

The second ML method is the extreme learning machine (ELM) [67]. The ELM is a shallow neural network with randomly selected wrights and bias. Let (x_c, y_c) be a relative center of Ω , and r_c be half the diameter of Ω . The functions of ELM are the following:

$$R_\theta = \sum_{j=1}^K w_j \sigma \left(m_j^1 \frac{x - x_c}{r_c} + m_j^2 \frac{y - y_c}{r_c} + b_j \right),$$

where $m_j^1, m_j^2, b_j, j = 1, 2, \dots, K$ are random numbers uniformly distributed on $[-1, 1]$, and $\theta = \{w_j\}_{j=1}^K$ is the parameter set of ELM. We note that the numbers m_j^1, m_j^2, b_j are generated in advance and fixed, which are not the trained parameters like in the DSNN. See

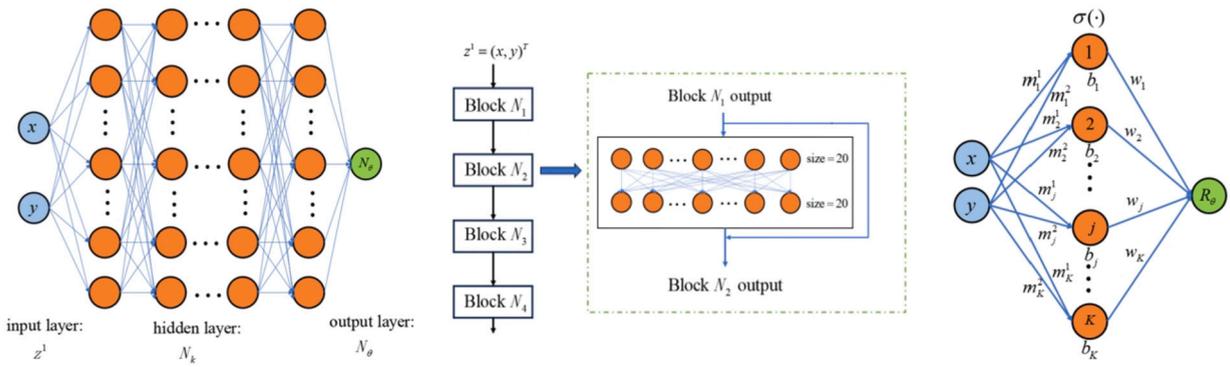


Fig. 4. Left: a fully connected neural network; Middle: a residual network incorporating four residual blocks, each of which contains two full connection layers and one residual item, and each layer contains 20 neurons; Right: an ELM.

Table 1
The approximation level of various ML enrichments \mathcal{G}_{ml} to the condition (5.4).

\mathcal{G}_{ml}	$\ \mathcal{G}_{ml} - 0\ _{L^\infty(\Gamma)}$	[min,max] of $\left \left[\frac{\partial \mathcal{G}_{ml}}{\partial \vec{n}_\Gamma} \right]_\Gamma \right $
\mathcal{G}_{dsnn}	6.425E-03	[0.989,1.009]
\mathcal{G}_{elm}	4.809E-13	[0.438,1.548]

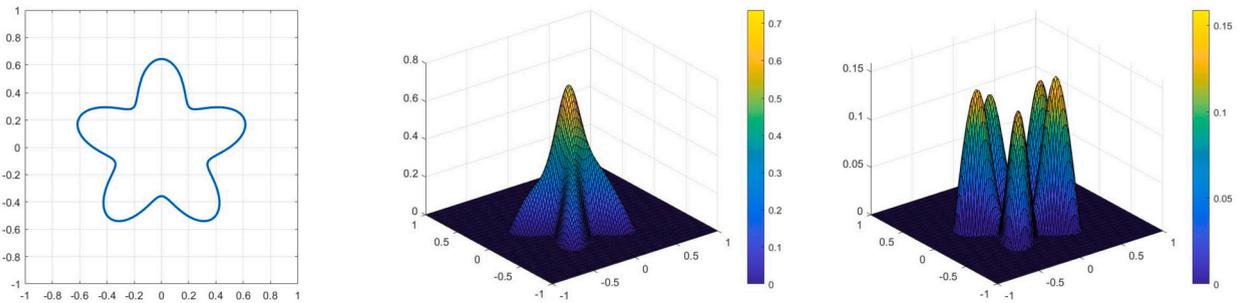


Fig. 5. Left: a pentagon interface. The plots in Ω_0 of ML methods: Middle: DSNN enrichment; Right: ELM enrichment. (For interpretation of the color(s) in the figure(s), the reader is referred to the web version of this article.)

Fig. 4 Right for an illustration of ELM. The optimization process of ELM is executed by solving a least square problem. The ELM has been applied to solve PEDs, see [68–71] for instance. We use the ELM in this paper to construct the enrichment (5.4).

In this paper we use the tanh activation function [49], and thus both the DSNN functions and ELM functions are C^∞ continuous. The DSNN and ELM based on the loss function (5.5) to produce the ML functions in (5.4) are denoted by \mathcal{G}_{dsnn} , \mathcal{G}_{elm} , respectively. In our computation we find out that the DSNN cannot produce the ML function with sufficient accuracy in (5.4). Therefore, the associated SGFEM generates the suboptimal convergence rate. On the contrary, the ELM meets (5.4) very accurately even for the interface is geometrically complex. Therefore, the ELM is suggested to yield the ML enrichment of the SGFEM for IP.

In the end of this section we carry out a numerical example to intuitively illustrate how effectively the proposed ML algorithm produces the ML enrichments in (5.4). We consider a pentagon interface with a parametric equation $r = 1/2 + 1/7 \sin(5\theta)$, where (r, θ) is the polar coordinate, see Fig. 5 Left. We uniformly sample N_Γ points Q_i on Γ as training points. In the DSNN, we set $N_\Gamma = 200$. The parameters of DSNN are the following: 4 residual blocks are used, each of which contains two full connection layers and one residual item, and each layer contains 20 neurons, iteration number is 10000, $\lambda = 0.1$, $\kappa_1 = 1$ in (5.5), the optimizer is “Adam”, the learning rate $\eta = 0.001$. In the ELM, we set $N_\Gamma = 340$, and $K = 300$ neurons are used, the parameters $m_j^1, m_j^2, b_j, j = 1, 2, \dots, K$ are randomly sampled based on the uniform distribution on $[-1, 1]$, $\lambda = 1 \times 10^{-12}$, $\kappa_1 = 1$ in (5.5), $x_c = 0, y_c = 0, r_c = 1/7$. The enrichments yielded by these ML methods based on the loss (5.5) are denoted by \mathcal{G}_{dsnn} and \mathcal{G}_{elm} , respectively. The plots of \mathcal{G}_{dsnn} and \mathcal{G}_{elm} in Ω_0 are shown in Fig. 5. To illustrate how these functions meet the condition (5.4) we compute the following indicators:

$$\|\mathcal{G}_{ml} - 0\|_{L^\infty(\Gamma)}, \quad [\min, \max] \text{ of } \left| \left[\frac{\partial \mathcal{G}_{ml}}{\partial \vec{n}_\Gamma} \right]_\Gamma \right|,$$

where $[\min, \max]$ are the smallest and largest values of $\left| \left[\frac{\partial \mathcal{G}_{ml}}{\partial \vec{n}_\Gamma} \right]_\Gamma \right|$ on Γ . These quantities of \mathcal{G}_{dsnn} , \mathcal{G}_{elm} are shown in Table 1.

From the Table 1 we see that the ELM enrichment based on the loss function (5.5) meets the condition (5.4) accurately, that is, $\|\mathcal{G}_{elm} - 0\|_{L^\infty(\Gamma)}$ is almost zero, and $\left| \left[\frac{\partial \mathcal{G}_{elm}}{\partial \bar{n}} \right]_{\Gamma} \right|$ has a positive low bound. The DSNN enrichment has relatively large error, $\|\mathcal{G}_{dsnn} - 0\|_{L^\infty(\Gamma)}$. The reason is that the approximation errors of DSNN (other DNN also) have a typical limitation, about 10^{-5} , which is difficult to improve by adjusting network parameters. The accuracy loss in the DSNN enrichment will damage the optimal convergence rate of SGFEM, as will be shown in the numerical experiments in the next section. In contrast, the ELM can achieve high approximation accuracy for complex interface curves. Therefore, we suggest the use of ELM enrichment in the SGFEM in this study.

6. Numerical results

We consider the model problem (2.2) in a domain $\Omega = (0, 1)^2$ with various curved interfaces Γ for the numerical experiments. The uniform $n \times n$ square FE mesh is used to discretize the domain $\Omega = (0, 1)^2$ with the mesh parameter $h = 1/n$. The nodes associated with the mesh are denoted by $\{P_i\}_{i \in I_h}$, where I_h is the index set. The meshes are not fitted to the interface. We will test different efficient materials, a_0 and a_1 . We set the manufactured exact solution u of (2.2) in the tests, and the loading functions f, g of (2.2) are calculated by using equation (2.2) and the manufactured exact solution u .

We will test the two ML enrichments with the learning schemes (5.5), DSNN and ELM, in the SGFEM (3.11), and identify the method with the best approximation result. We also test the existing enrichments in the literature for comparison and verification of Theorem 4.3. We list them in what follows: Specifically, the tested enrichments are the following:

- DF-the distance function (3.3), D ,
- ODF-the one-side distance function (3.5), \tilde{D} ,
- EDF-the exponential form of DF (3.6) with $\mu = 3$,
- ABDSNN-the absolute of proposed DSNN enrichment (5.2), \mathcal{G}_{dsnn}^{abs} , with the learning scheme (5.5),
- OSDSNN-the one-side DSNN function (5.2), \mathcal{G}_{dsnn}^{os} , with the learning scheme (5.5),
- ABELM-the absolute of proposed ELM enrichment (5.2), \mathcal{G}_{elm}^{abs} with the learning scheme (5.5),
- OSELM-the one-side ELM function (5.2), \mathcal{G}_{elm}^{os} with the learning scheme (5.5),

We compute and compare the relative error in the energy norm (EE), i.e.,

$$EE = \frac{\|u - u_h\|_{\mathcal{E}(\Omega)}}{\|u\|_{\mathcal{E}(\Omega)}}$$

for the approximation solution u_h obtained from these methods, and the scaled condition number (SCN) of associated stiffness matrices \mathbf{A} . The SCN of \mathbf{A} is defined by

$$\mathcal{K} := \kappa(\mathbf{DAD}), \tag{6.1}$$

where $\kappa(\cdot)$ is 1-condition number of a symmetric matrix, and \mathbf{D} be a diagonal matrix with

$$\mathbf{D}_{ii} = \mathbf{A}_{ii}^{-1/2}.$$

Setting for DSNN and ELM. In the numerical experiments below, we set the parameters in the DSNN and ELM as follows. The DSNN structure consists of 4 residual blocks, each of which contains two full connection layers and one residual item, where each layer contains 20 neurons, see Fig. 4 Middle. In the ELM, we use K neurons with the randomly generated weight and bias parameters in $[-1, 1]$, see Fig. 4 Right. The least square method [68–71] is used to solve the learning schemes (5.5). The number of sampling points for training the DSNN, the learning rate, and the iteration number of DSNN, and the number of neurons in the ELM will be given in the following subsections.

Integration for cut elements. We adopt a general numerical integration formula for the cut element. For elements cut by the interface, we connect the intersection points of the interface and the boundaries of an element by a straight line, and decompose the element into sub-triangles, on each of which the standard Gaussian rule for triangles is employed. We mention that a systematic study on the effect of numerical integration is not the objective of this work. We refer to for more details about the numerical integrations for the interface problems [23,33,79].

We now present our numerical results in the following sub-sections.

6.1. An interface problem with a circular interface

We first consider a circular interface Γ with the equation $(x - x_0)^2 + (y - y_0)^2 = r_0^2$, where $x_0 = \frac{1}{\sqrt{5}}, y_0 = \frac{1}{\sqrt{3}}, r_0 = \frac{1}{\sqrt{10}}$. In this case we set the manufactured solution of (2.2) as follows:

$$u = \begin{cases} \frac{2a_1 - a_1 r_0^2}{(a_1 - a_0)r_0^4} r^2 \cos(2\theta), & r < r_0 \ (\Omega_0), \\ \frac{a_1 + a_0 - a_0 r_0^2}{(a_1 - a_0)r_0^4} r^2 \cos(2\theta) + r^{-2} \cos(2\theta), & r \geq r_0 \ (\Omega_1), \end{cases}$$

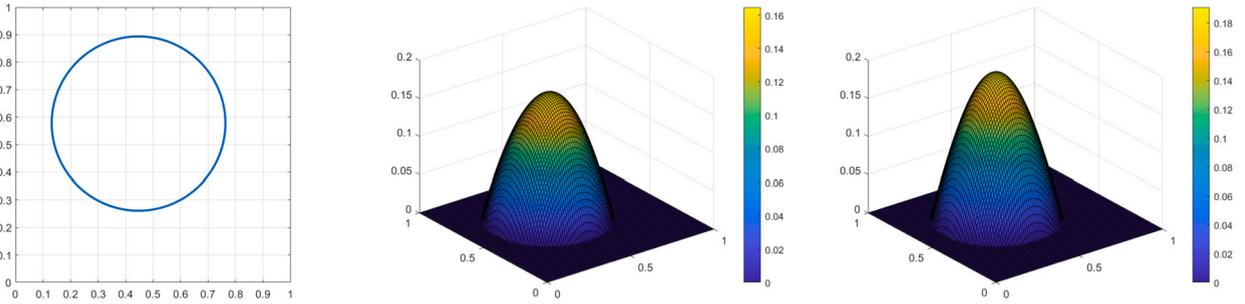


Fig. 6. Left: a circular interface. The plots in Ω_0 of ML methods: Middle: DSNN enrichment; Right: ELM enrichment.

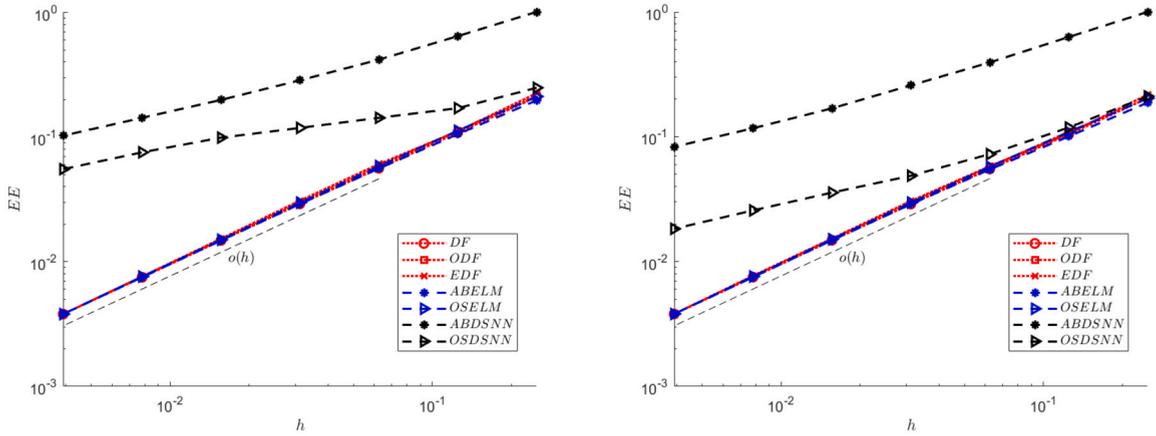


Fig. 7. EEs of SGFEM with various enrichments for a circular interface, the interface coefficients are left: $a_0 = 1, a_1 = 10$ ($c = 10$) and right: $a_0 = 1, a_1 = 100$ ($c = 100$).

Table 2
Approximation levels of various ML enrichments \mathcal{G}_{ml} to the condition (5.4) for the circular interface.

\mathcal{G}_{ml}	$\ \mathcal{G}_{ml} - 0\ _{L^\infty(\Gamma)}$	[min,max] of $\left \left[\frac{\partial \mathcal{G}_{ml}}{\partial \mathbf{m}_i} \right]_{\Gamma} \right $
\mathcal{G}_{dsnn}	2.015E-05	[0.998,1.003]
\mathcal{G}_{elm}	1.313E-14	[1.000,1.000]

where (r, θ) is the polar coordinate at the center (x_0, y_0) . It can be verified that u satisfies the interface condition (2.3) and (2.4). The mesh on the domain $[0, 1] \times [0, 1]$ is refined with $h^{-1} = n = 2^{j+1}$, $j = 1, 2, \dots, 7$. The interface Γ and a mesh with $n = 10$ are shown in Fig. 6 Left. We test two cases of coefficients $a(P)$: (i) $a_0 = 1$ and $a_1 = 10$, and (ii) $a_0 = 1$ and $a_1 = 100$. Their contrasts are $c = 10$ and $c = 100$, respectively.

The numbers of sampling points for training the DSNN and ELM are 340 in this situation. In the DSNN, we set the iteration number = 10000, $\lambda = 0.1$, $\kappa_1 = 1$, the optimizer = “Adam”, the learning rate $\eta = 0.001$. In the ELM, we set 300 neurons, $\lambda = 1$, $\kappa_1 = 1$, $x_c = 0, y_c = 0, r_c = 1$. The approximation levels of various ML enrichments \mathcal{G}_{ml} to the condition (5.4) are presented in Table 2, which are characterized by $\|\mathcal{G}_{ml} - 0\|_{L^\infty(\Gamma)}$ and $[\min, \max]$ of $\left| \left[\frac{\partial \mathcal{G}_{ml}}{\partial \mathbf{m}_i} \right]_{\Gamma} \right|$. In Table 2 we see that the ELM meet the condition (5.4) perfectly. Especially, $\|\mathcal{G}_{elm} - 0\|_{L^\infty(\Gamma)} \approx 10^{-14}$ is almost zero. However, the error $\|\mathcal{G}_{dsnn} - 0\|_{L^\infty(\Gamma)}$ is relatively larger so that the associated SGFEM cannot get the optimal convergence rate, see Fig. 7. The plots of DSNN and ELM enrichments in Ω_0 are shown in Fig. 6 Middle and Right.

The EEs and SCNs with respect to h of the SGFEM with various enrichments are presented for the different contrasts c (10 and 100) in Fig. 7 and Fig. 8, respectively. It is shown in these figures that the SGFEM with the enrichments DF, ODF, EDF, ABELM, OSELM can achieve the optimal convergence rate $O(h)$. This is predicted by the theoretical result, (4.13). Their conditioning is all of same order as that of FEM, and the merit of SGFEM is preserved. However, the rates of ABDSNN and OSDSNN are suboptimal, which is due to the accuracy loss of $\|\mathcal{G}_{dsnn} - 0\|_{L^\infty(\Gamma)}$, as remarked in the end of Section 5. Therefore, the ELM enrichment is suggested in this study. Our numerical results point out that there is not an essential difference between the ABELM and OSELM enrichments, and thus we will only test the ABELM in the experiment below.

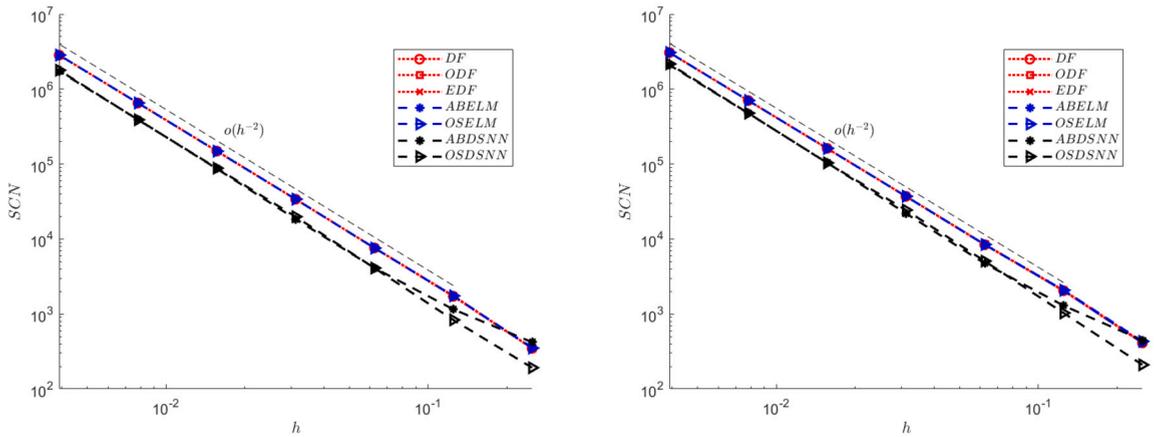


Fig. 8. SCNs of SGFEM with various enrichments for a circular interface, the interface coefficients are left: $a_0 = 1, a_1 = 10$ ($c = 10$) and right: $a_0 = 1, a_1 = 100$ ($c = 100$).

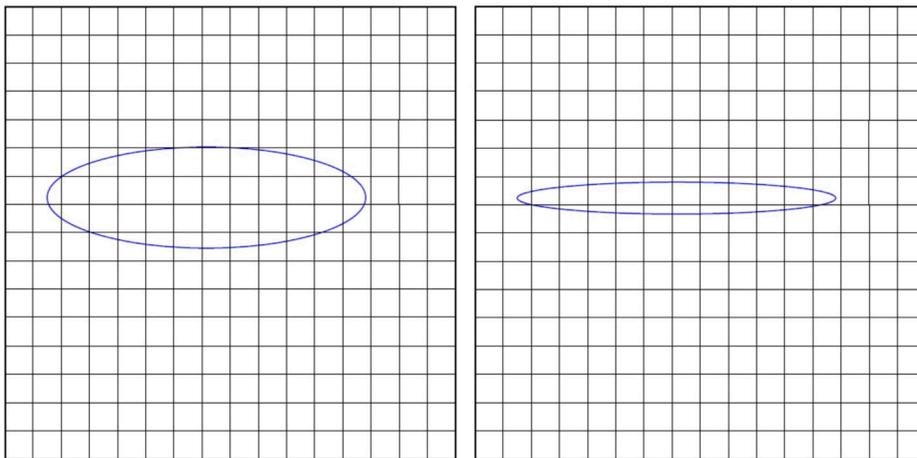


Fig. 9. The domain with elliptical interfaces. Left: $\alpha^2 = \frac{1}{8}, \beta^2 = \frac{1}{80}$; Right: $\alpha^2 = \frac{1}{8}, \beta^2 = \frac{1}{800}$.

6.2. Ellipsoid interfaces with differential ratios of major and minor axes

We next test another curved interface problem involving elliptical interfaces Γ with different ratios of major and minor axes as follows:

$$\frac{(x - x_0)^2}{\alpha^2} + \frac{(y - y_0)^2}{\beta^2} = 1,$$

where $x_0 = \frac{1}{\sqrt{5}}, y_0 = \frac{1}{\sqrt{3}}$. We test the cases (a) $\alpha^2 = \frac{1}{8}, \beta^2 = \frac{1}{80}$ and (b) $\alpha^2 = \frac{1}{8}, \beta^2 = \frac{1}{800}$ to verify applicability of the proposed general enrichments for varying curvatures. The mesh on the domain $[0, 1] \times [0, 1]$ is refined with $h^{-1} = n = 2^{j+1}, j = 1, 2, \dots, 7$. These interfaces and a mesh with $n = 16$ are shown in Fig. 9. The manufactured solution of (2.2) we consider is the following:

$$u = \begin{cases} a_1 \left(\frac{(x-x_0)^2}{\alpha^2} + \frac{(y-y_0)^2}{\beta^2} - 1 \right) e^{2x+y}, & \frac{(x-x_0)^2}{\alpha^2} + \frac{(y-y_0)^2}{\beta^2} < 1 \ (\Omega_0), \\ a_0 \left(\frac{(x-x_0)^2}{\alpha^2} + \frac{(y-y_0)^2}{\beta^2} - 1 \right) e^{2x+y}, & \frac{(x-x_0)^2}{\alpha^2} + \frac{(y-y_0)^2}{\beta^2} \geq 1 \ (\Omega_1). \end{cases}$$

It can be checked that u satisfies the interface condition (2.3) and (2.4). As illustrated in the situation of circular interface, in this case we only test the SGFEM with ABELM. The number of sampling points uniformly for training the ELM is 340, we set 300 neurons, $\kappa_1 = 1, x_c = \frac{1}{\sqrt{5}}, y_c = \frac{1}{\sqrt{3}}, r_c = \frac{1}{\sqrt{8}}$. For case (a), we set $\lambda = 1 \times 10^{-12}$; for case (b), we set $\lambda = 1 \times 10^{-11}$. The enrichments of case (a) and case (b) is drawn in Fig. 10, and the approximation levels of ELM enrichment to the condition (5.4) are presented in Table 3. For Table 3 we see that the condition (5.4) ELM is satisfied with very high accuracy for both elliptical interfaces.

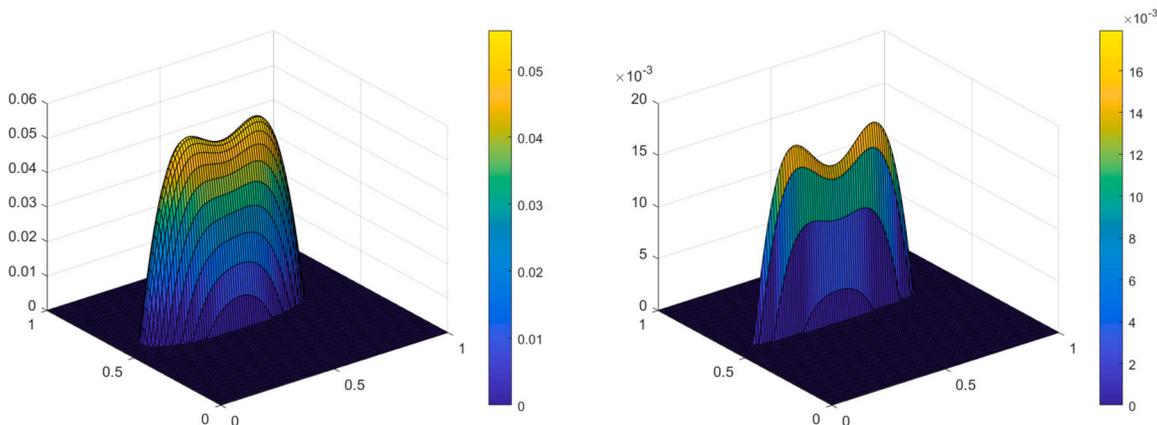


Fig. 10. The plots of ELM enrichments in Ω_0 . Left: $\alpha^2 = \frac{1}{8}, \beta^2 = \frac{1}{80}$; Right: $\alpha^2 = \frac{1}{8}, \beta^2 = \frac{1}{800}$.

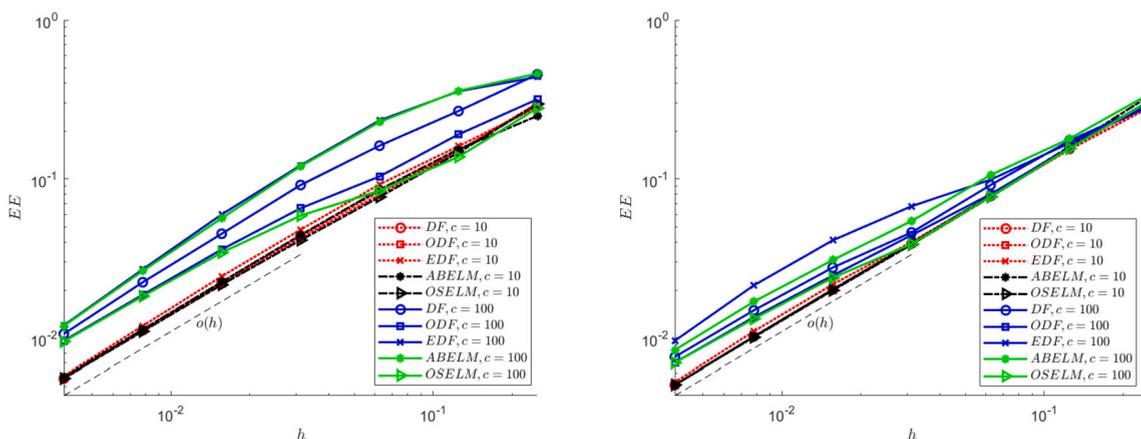


Fig. 11. EE of ellipse, left: $\alpha^2 = \frac{1}{8}, \beta^2 = \frac{1}{80}$ and right: $\alpha^2 = \frac{1}{8}, \beta^2 = \frac{1}{800}$.

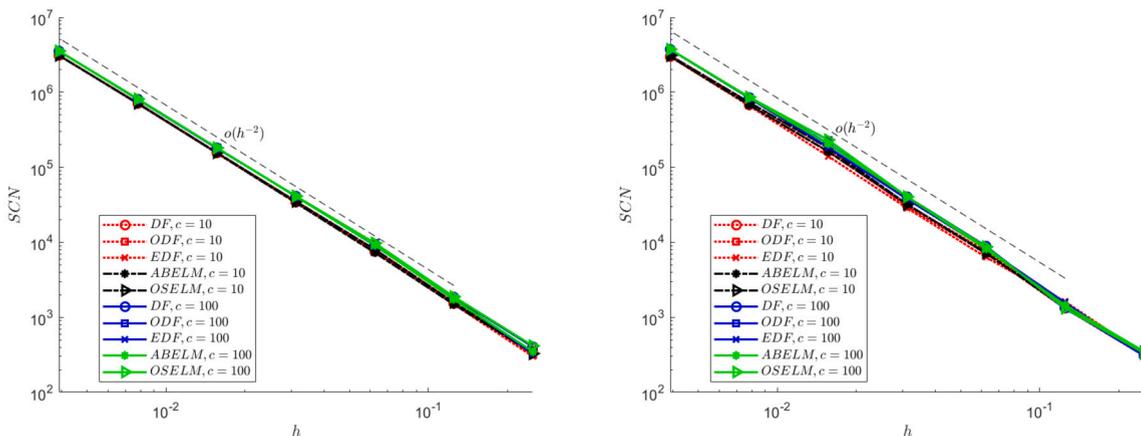


Fig. 12. SCN of ellipse, left: $\alpha^2 = \frac{1}{8}, \beta^2 = \frac{1}{80}$ and right: $\alpha^2 = \frac{1}{8}, \beta^2 = \frac{1}{800}$.

The EEs and SCNs with respect to h of the SGFEM with the ABELM are presented for the different ratios of major and minor axes (a) and (b) and different contrasts c (10 and 100) in Fig. 11 and Fig. 12, respectively. It is clearly shown in Fig. 11 that in the SGFEM with the ABELM along with other enrichments all converge with the optimal order $O(h)$, as predicted in the theoretical result, (4.13). The proposed ELM enrichment also applies to the curved interface problems with varying curvatures.

Table 3
Approximation levels of ELM enrichment to the condition (5.4) for the ellipsoid interfaces with different major and minor axes.

G_{elm}	$\ G_{elm} - 0\ _{L^\infty(\Gamma)}$	[min,max] of $\left \left[\frac{\partial G_{elm}}{\partial \bar{n}_c} \right]_\Gamma \right $
$\alpha^2 = \frac{1}{8}, \beta^2 = \frac{1}{80}$	1.041E-13	[0.545,1.219]
$\alpha^2 = \frac{1}{8}, \beta^2 = \frac{1}{800}$	6.593E-13	[0.209,1.276]

Table 4
Approximation levels of ABELM for the straight interfaces.

$\ G_{elm}^{abs} - 0\ _{L^\infty(\Gamma)}$	[min,max] of $\left \left[\frac{\partial G_{elm}^{abs}}{\partial \bar{n}_c} \right]_\Gamma \right $
6.423E-14	[0.893,1.042]

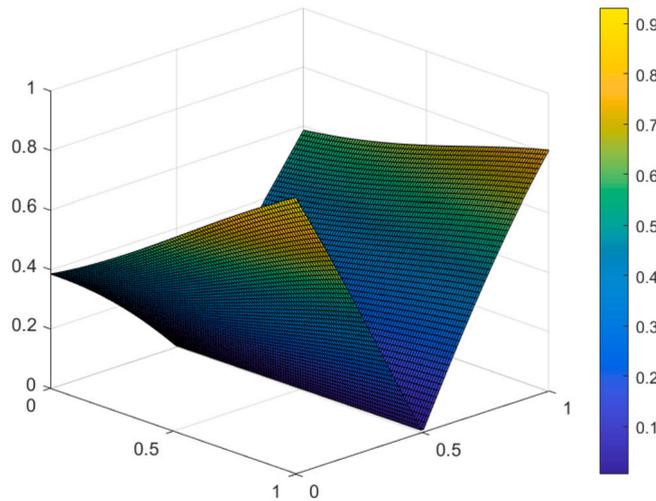


Fig. 13. The plot of ABELM for the straight interface.

6.3. Robustness test

We next test the robustness of SGFEM with the enrichment, ABELM. To this end, we fix $h = 1/16$ and consider a straight interface with equation $y = 0.5 + \delta$. We consider $\delta = 0.03 \times 2^{-j+1}$, $j = 1, 2, \dots, 20$ such that δ varies from 3×10^{-2} to 5.72×10^{-8} . In this case, the interface approaches a mesh-line $y = 0.5$ as δ decreases. The number of sampling points uniformly for training the ELM are 340, we set 300 neurons, $\kappa_1 = 1$, $x_c = 0$, $y_c = 0$, $r_c = 1$, $\lambda = 1 \times 10^{-12}$. The ABELM enrichments for the straight interface is drawn in Fig. 13, and the approximation levels of ELM to the condition (5.4) are presented in Table 4. Again the ELM enrichment meets the condition (5.4) perfectly.

The SCNs with respect to δ of SGFEM with the ABELM are presented in Fig. 14 for the different contrasts c (10 and 100). It is clear from Fig. 14 that the SCN of SGFEM with the ELM enrichment does not change as δ decreases (for fixed $h = 1/16$). This indicates that the proposed SGFEM with ABELM is robust.

7. Conclusions and comments

This paper proposed a general enrichment scheme for SGFEM of interface problem. The SGFEM enriched by such an enrichment was proven to converge with the optimal convergence rate. This theoretical analysis also holds for the commonly-used enrichments in the literature because these enrichments are special instances of the proposed general enrichment. Based on the general enrichment we developed the ML enrichment to replace the distance and level set functions. The ML enrichment possesses the merit in dealing with geometrically complex interfaces and reducing the complexity in evaluation of the distance and level set functions. In comparison with the DNN enrichment, the ELM enrichment achieved very high accuracy and was highly suggested in the SGFEM of interface problem. The extension of ML enrichments to three dimensional interface problems will be investigated in a forthcoming study.

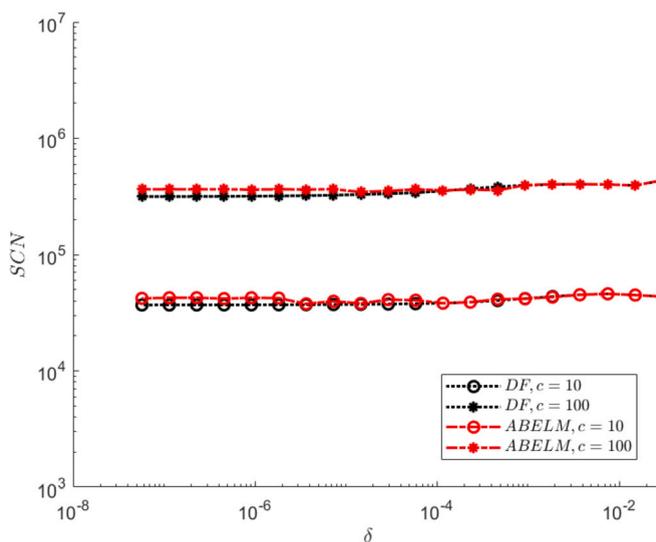


Fig. 14. The robustness test.

CRedit authorship contribution statement

Dongmei Wang: Data curation, Investigation, Writing – original draft. **Hengguang Li:** Conceptualization, Methodology, Validation, Writing – original draft, Writing – review & editing. **Qinghui Zhang:** Conceptualization, Funding acquisition, Investigation, Methodology, Project administration, Supervision, Validation, Writing – original draft, Writing – review & editing.

References

- [1] D.A. Edwards, H. Brenner, D.T. Wasan, *Interfacial Transport Process and Rheology*, Butterworths/Heinemann, London, 1991.
- [2] C.S. Peskin, Numerical analysis of blood flow in heart, *J. Comput. Phys.* 25 (1977) 220–252.
- [3] I. Babuška, The finite element method for elliptic equations with discontinuous coefficients, *Computing* 5 (1970) 207–213.
- [4] J.W. Barrett, C.M. Elliott, Fitted and unfitted finite-element methods for elliptic equations with smooth interfaces, *IMA J. Numer. Anal.* 7 (1987) 283–300.
- [5] Z. Chen, J. Zou, Finite element methods and their convergence for elliptic and parabolic interface problems, *J. Numer. Math.* 79 (1998) 175–202.
- [6] Q. Zhang, I. Babuška, A stable generalized finite element method (SGFEM) of degree two for interface problems, *Comput. Methods Appl. Mech. Eng.* 363 (2020) 112889.
- [7] Q. Zhang, Unfitted generalized finite element methods for Dirichlet problems without penalty or stabilization, *Numer. Methods Partial Differ. Equ.* 40 (2024) e23081.
- [8] Z. Li, T. Lin, Y. Lin, R.C. Rogers, An immersed finite element space and its approximation capability, *Numer. Methods Partial Differ. Equ.* 20 (2004) 338–367.
- [9] T. Lin, Y. Lin, X. Zhang, Partially penalized immersed finite element methods for elliptic interface problems, *SIAM J. Numer. Anal.* 53 (2015) 1121–1144.
- [10] S. Adjerid, I. Babuska, R. Guo, T. Lin, An enriched immersed finite element method for interface problems with nonhomogeneous jump conditions, *Comput. Methods Appl. Mech. Eng.* 404 (2023) 115770.
- [11] A. Hansbo, 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.
- [12] P. Hansbo, M.G. Larson, S. Zahedi, A cut finite element method for a Stokes interface problem, *Appl. Numer. Math.* 85 (2014) 90–114.
- [13] P. Huang, H. Wu, Y. Xiao, An unfitted interface penalty finite element method for elliptic interface problems, *Comput. Methods Appl. Mech. Eng.* 323 (2017) 439–460.
- [14] I. Harari, J. Dolbow, Analysis of an efficient finite element method for embedded interface problems, *Comput. Math.* 46 (2010) 205–211.
- [15] C. Lehrenfeld, A. Reusken, Analysis of a Nitsche XFEM-DG discretization for a class of two-phase mass transport problems, *SIAM J. Numer. Anal.* 51 (2013) 958–983.
- [16] J. Chessa, T. Belytschko, An extended finite element method for two-phase fluids, *J. Appl. Mech.* 70 (2003) 10–17.
- [17] M. Kirchhart, S. Gross, A. Reusken, Analysis of an XFEM discretization for Stokes interface problems, *SIAM J. Sci. Comput.* 38 (2016) A1019–A1043.
- [18] H. Sauerland, T.P. Fries, The extended finite element method for two-phase and free-surface flows: a systematic study, *J. Comput. Phys.* 230 (2011) 3369–3390.
- [19] K. Kergrene, I. Babuška, U. Banerjee, Stable generalized finite element method and associated iterative schemes: application to interface problems, *Comput. Methods Appl. Mech. Eng.* 305 (2016) 1–36.
- [20] N. Moës, M. Cloirec, P. Cartraud, J.F. Remacle, A computational approach to handle complex microstructure geometries, *Comput. Methods Appl. Mech. Eng.* 192 (2003) 3163–3177.
- [21] P. Zhu, Q. Zhang, T. Liu, Stable generalized finite element method (SGFEM) for parabolic interface problems, *J. Comput. Appl. Math.* 367 (2020) 112475.
- [22] Q. Zhang, C. Cui, U. Banerjee, I. Babuška, A condensed generalized finite element method (CGFEM) for interface problems, *Comput. Methods Appl. Mech. Eng.* 391 (2022) 114537.
- [23] K.W. Cheng, T.P. Fries, Higher-order XFEM for curved strong and weak discontinuities, *Int. J. Numer. Methods Eng.* 82 (2010) 564–590.
- [24] Q. Zhang, U. Banerjee, I. Babuška, Strongly stable generalized finite element method (SSGFEM) for a non-smooth interface problem, *Comput. Methods Appl. Mech. Eng.* 344 (2019) 538–568.
- [25] H. Sauerland, T.P. Fries, The stable XFEM for two-phase flows, *Comput. Fluids* 87 (2013) 41–49.
- [26] A. Patel, S.K. Acharya, A.K. Pani, Stabilized Lagrange multiplier method for elliptic and parabolic interface problems, *Appl. Numer. Math.* 120 (2017) 287–304.
- [27] P. Díez, R. Cottereau, S. Zlotnik, A stable extended FEM formulation for multi-phase problems enforcing the accuracy of the fluxes through Lagrange multipliers, *Int. J. Numer. Methods Eng.* 96 (2013) 303–322.

- [28] A.M. Aragón, B. Liang, H. Ahmadian, S. Soghrati, On the stability and interpolating properties of the hierarchical interface-enriched finite element method, *Comput. Methods Appl. Mech. Eng.* 362 (2020) 112671.
- [29] A. Zilian, A. Legay, The enriched space-time finite element method (EST) for simultaneous solution of fluid-structure interaction, *Int. J. Numer. Methods Eng.* 75 (2008) 305–334.
- [30] Z. Cai, X. Ye, S. Zhang, Discontinuous Galerkin finite element methods for interface problems: a priori and a posteriori error estimations, *SIAM J. Numer. Anal.* 49 (2011) 1761–1787.
- [31] R. Massjung, An unfitted discontinuous Galerkin method applied to elliptic interface problems, *SIAM J. Numer. Anal.* 50 (2012) 3134–3162.
- [32] I. Babuška, U. Banerjee, J. Osborn, Survey of meshless and generalized finite element methods: a unified approach, *Acta Numer.* 12 (2003) 1–125.
- [33] T.P. Fries, T. Belytschko, The extended/generalized finite element method: an overview of the method and its applications, *Int. J. Numer. Methods Eng.* 84 (2010) 253–304.
- [34] Y. Efendiev, T.Y. Hou, *Multiscale Finite Element Methods: Theory and Applications*, Springer, 2009.
- [35] I. Babuška, J.M. Melenk, The partition of unity finite element method, *Int. J. Numer. Methods Eng.* 40 (1997) 727–758.
- [36] C.A. Duarte, J.T. Oden, An h-p adaptive method using clouds, *Comput. Methods Appl. Mech. Eng.* 139 (1996) 237–262.
- [37] A.M. Aragón, C.A. Duarte, P.H. Geubelle, Generalized finite element enrichment functions for discontinuous gradient fields, *Int. J. Numer. Methods Eng.* 82 (2010) 242–268.
- [38] I. Babuška, U. Banerjee, Stable generalized finite element method, *Comput. Methods Appl. Mech. Eng.* 201–204 (2011) 91–111.
- [39] V. Gupta, C.A. Duarte, I. Babuška, U. Banerjee, A stable and optimally convergent generalized FEM (SGFEM) for linear elastic fracture mechanics, *Comput. Methods Appl. Mech. Eng.* 266 (2013) 23–39.
- [40] Q. Zhang, U. Banerjee, I. Babuska, High order stable generalized finite element methods, *Numer. Math.* 128 (2014) 1–29.
- [41] C. Cui, Q. Zhang, U. Banerjee, I. Babuška, Stable generalized finite element method (SGFEM) for three-dimensional crack problems, *Numer. Math.* 152 (2022) 475–509.
- [42] Q. Zhang, DOF-gathering stable generalized finite element methods (SGFEM) for crack problems, *Numer. Methods Partial Differ. Equ.* 36 (2020) 1209–1233.
- [43] C. Cui, Q. Zhang, Stable generalized finite element methods (SGFEM) for elasticity crack problems, *Int. J. Numer. Methods Eng.* 121 (2020) 3066–3082.
- [44] M. Duprez, A. Lozinski, ϕ -FEM: a finite element method on domains defined by level-sets, *SIAM J. Numer. Anal.* 58 (2020) 1008–1028.
- [45] W. Gong, H. Li, Q. Zhang, Improved enrichments and numerical integrations in SGFEM for interface problems, *J. Comput. Appl. Math.* 438 (2024) 115540.
- [46] B. Smith, B. Vaughan, D. Chopp, The extended finite element method for boundary layer problems in biofilm growth, *Commun. Appl. Math. Comput. Sci.* 2 (2007) 35–56.
- [47] A. Khoei, *Extended Finite Element Method: Theory and Applications*, Wiley Series in Computational Mechanics, John Wiley & Sons, Inc., Chichester, West Sussex, 2015.
- [48] C. Lang, D. Makhija, A. Doostan, K. Maute, A simple and efficient preconditioning scheme for Heaviside enriched XFEM, *Comput. Mech.* 54 (2014) 1357–1374.
- [49] I.J. Goodfellow, Y. Bengio, A. Courville, *Deep Learning*, MIT Press, 2016.
- [50] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, in: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 770–778.
- [51] M. Raissi, P. Perdikaris, G.E. Karniadakis, Physics-informed neural networks: a deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations, *J. Comput. Phys.* 378 (2019) 686–707.
- [52] J. Sirignano, K. Spiliopoulos, DGM: a deep learning algorithm for solving partial differential equations, *J. Comput. Phys.* 375 (2018) 1339–1364.
- [53] W. E, B. Yu, The deep Ritz method: a deep learning-based numerical algorithm for solving variational problems, *Commun. Math. Stat.* 6 (2018) 1–12.
- [54] E. Samaniego, C. Anitescu, et al., An energy approach to the solution of partial differential equations in computational mechanics via machine learning: concepts, implementation and applications, *Comput. Methods Appl. Mech. Eng.* 362 (2020) 112790.
- [55] Y. Zang, G. Bao, X. Ye, H. Zhou, Weak adversarial networks for high-dimensional partial differential equations, *J. Comput. Phys.* 411 (2020) 109409.
- [56] Y. Liao, P. Ming, Deep Nitsche method: deep Ritz method with essential boundary conditions, *Commun. Comput. Phys.* 29 (2021) 1365–1384.
- [57] J. Bai, G. Liu, A. Gupta, L. Alzubaidi, X. Feng, Y. Gu, Physics-informed radial basis network (PIRBN): a local approximating neural network for solving nonlinear partial differential equations, *Comput. Methods Appl. Mech. Eng.* 415 (2023) 116290.
- [58] L. Lu, X. Meng, Z. Mao, et al., DeepXDE: a deep learning library for solving differential equations, *SIAM Rev.* 63 (1) (2021) 208–228.
- [59] H. Sheng, C. Yang, Pfnm: a penalty-free neural network method for solving a class of second order boundary-value problems on complex geometries, *J. Comput. Phys.* 428 (2021) 110085.
- [60] S. Zeng, Z. Zhang, Q. Zou, Adaptive deep neural networks methods for high-dimensional partial differential equations, *J. Comput. Phys.* 463 (2022) 111232.
- [61] J.M. Taylor, P. David, I. Muga, A deep Fourier residual method for solving PDEs using neural networks, *Comput. Methods Appl. Mech. Eng.* 405 (2023) 115850.
- [62] X. Jin, S. Cai, H. Li, et al., NSFnets (Navier-Stokes flow nets): physics-informed neural networks for the incompressible Navier-Stokes equations, *J. Comput. Phys.* 426 (2021) 109951.
- [63] Z. Wang, Z. Zhang, A mesh-free method for interface problems using the deep learning approach, *J. Comput. Phys.* 400 (2020) 108963.
- [64] C. He, X. Hu, L. Mu, A mesh-free method using piecewise deep neural network for elliptic interface problems, *J. Comput. Appl. Math.* 412 (2022) 114358.
- [65] S. Wu, B. Lu, INN: interfaced neural networks as an accessible meshless approach for solving interface PDE problems, *J. Comput. Phys.* 470 (2022) 111588.
- [66] Y. Tseng, T. Lin, W. Hu, M. Lai, A cusp-capturing PINN for elliptic interface problems, *J. Comput. Phys.* 491 (2023) 112359.
- [67] G. Huang, Q. Zhu, C. Siew, Extreme learning machine: theory and applications, *Neurocomputing* 70 (1–3) (2006) 489–501.
- [68] V. Dwivedi, B. Srinivasan, Physics informed extreme learning machine (PIELM)—a rapid method for the numerical solution of partial differential equations, *Neurocomputing* 391 (2020) 96–118.
- [69] G. Fabiani, F. Calabrò, L. Russo, et al., Numerical solution and bifurcation analysis of nonlinear partial differential equations with extreme learning machines, *J. Sci. Comput.* 89 (2021) 44.
- [70] F. Calabrò, G. Fabiani, C. Siettos, Extreme learning machine collocation for the numerical solution of elliptic PDEs with sharp gradients, *Comput. Methods Appl. Mech. Eng.* 387 (1) (2021) 114–188.
- [71] S. Dong, J. Yang, Local extreme learning machines and domain decomposition for solving linear and nonlinear partial differential equations, *Comput. Methods Appl. Mech. Eng.* 387 (2021) 114129.
- [72] Z. Li, W. Wang, I.L. Chern, M. Lai, New formulations for interface problems in polar coordinates, *SIAM J. Sci. Comput.* 25 (2003) 224–245.
- [73] Y. Jiang, M. Nian, Q. Zhang, A stable generalized finite element method coupled with deep neural network for interface problems with discontinuities, *Axioms* 11 (2022) 384.
- [74] S.C. Brenner, L.R. Scott, *The Mathematical Theory of Finite Element Methods*, third edition, Springer, USA, 2008.
- [75] J. Berg, K. Nyström, A unified deep artificial neural network approach to partial differential equations in complex geometries, *Neurocomputing* 317 (2018) 28–41.
- [76] S. Manavi, E. Fattahi, T. Becker, A trial solution for imposing boundary conditions of partial differential equations in physics-informed neural networks, *Eng. Appl. Artif. Intell.* 127 (2024) 107236.
- [77] N. Sukumara, A. Srivastavab, Exact imposition of boundary conditions with distance functions in physics-informed deep neural networks, *Comput. Methods Appl. Mech. Eng.* 389 (2022) 114333.
- [78] D.P. Kingma, J.L. Ba, Adam: a method for stochastic optimization, arXiv:1412.6980, 2014.
- [79] T.P. Fries, S. Omerović, Higher-order accurate integration of implicit geometries, *Int. J. Numer. Methods Eng.* 106 (2016) 323–371.