

The Discontinuous Galerkin Method by Patch Reconstruction for Helmholtz Problems

Di Li¹, Min Liu¹, Xiliang Lu¹ and Jerry Zhijian Yang^{1,2,*}

¹ School of Mathematics and Statistics, Wuhan University, Wuhan, Hubei 430072, China

² Computational Science Hubei Key Laboratory, Wuhan University, Wuhan, Hubei 430072, China

Received 10 January 2022; Accepted (in revised version) 24 March 2022

Abstract. This paper develops and analyzes interior penalty discontinuous Galerkin (IPDG) method by patch reconstruction technique for Helmholtz problems. The technique achieves high order approximation by locally solving a discrete least-squares over a neighboring element patch. We prove a priori error estimates in the L^2 norm and energy norm. For each fixed wave number k , the accuracy and efficiency of the method up to order five with high-order polynomials. Numerical examples are carried out to validate the theoretical results.

AMS subject classifications: 35J05, 65N30

Key words: Least-squares reconstruction, Helmholtz problems, patch reconstruction, discontinuous Galerkin, error estimates.

1 Introduction

The Helmholtz equation is a linear mathematical model that describes time-harmonic acoustic, elastic and electromagnetic steady state waves. One major problem in approximating this equation by classical finite element methods is the loss of ellipticity with an increasing excitation frequency.

For many years, the finite element method (and other type methods) has been widely used to discretize the Helmholtz equation with various types of boundary conditions, see [1,5–7,10,18,19,22,29] and the references therein. It is well known that, in every coordinate direction, one must put some minimal number of grid points in each wave length

*Corresponding author.

Emails: lidi.math@whu.edu.cn (D. Li), mliuf@whu.edu.cn (M. Liu), xllv.math@whu.edu.cn (X. Lu), zjyang.math@whu.edu.cn (J. Z. Yang)

$l = 2\pi/k$ in order to resolve the wave; that is, the mesh size h must satisfy the constraint $hk \lesssim 1$. In practice, 6-10 grid points are used in a wave length, which is often referred to as the "rule of thumb". However, this "rule of thumb" was probed rigorously not long ago by Ihlenburg [19] only in the one-dimensional case (called the preasymptotic error analysis). The main difficulty of the analysis is caused by the strong indefiniteness of the Helmholtz equation, which in turn makes it hard to establish stability estimates for the finite element solution under the "rule of thumb" mesh constraint. Standard finite element methods based on low-order polynomials do not perform well for the Helmholtz equation at high wavenumber. On the one hand, low-order polynomials do not well resolve the solution unless several grid points per wave length are used. On the other hand, such methods suffer from the so-called pollution effect: for a fixed number of grid points per wave length, the numerical error grows with the wavenumber [20, 21]. The detailed analysis of [15, 19] also shows that the pollution effect is inherent in the finite element method and is caused by the deterioration of stability of the Helmholtz operator as the wave number k becomes large. Indeed, it was suggested in [27, 28] that the pollution effect can be suppressed by using higher order polynomials for problems with higher wavenumber.

In the past fifteen years, DG methods have received a lot of attention and undergone intensive studies by many people. We refer the reader to [3, 4, 8, 12-14, 30, 32] and the references therein for a detailed account on DG methods for coercive elliptic and parabolic problems. We like to note that, in addition to the well known advantages of DG methods, the results of this paper also demonstrate the flexibility and effectiveness of DG methods for strongly indefinite problems, which was not well understood before. In this article, the discontinuous Galerkin method by patch reconstruction will be employed to study the Helmholtz problem. The method is an efficient numerical method for solving partial differential equations, was firstly introduced in [24] for the elliptic problems, and applied to many other model problems [23, 25, 26]. In [24], Li et al. proposed an arbitrary-order discontinuous Galerkin method for second-order elliptic problem on general polygonal mesh with only one degree of freedom per element by solving a local discrete least-squares over a neighboring element patch. In this work, our mesh-dependent sesquilinear forms penalize the jumps of the function values across the element edges/faces.

This paper has a small but vitally important idea that takes the penalty parameters as complex numbers of positive imaginary parts. This idea also contributes critically to the stability of the IPDG methods of this paper. The rest of the paper is organized as follows. In Section 2, we briefly describe the reconstruction finite element space, then state the basic properties of those spaces. In Section 3, we present the interior penalty discontinued Galerkin method for Helmholtz problem with the reconstructed approximation space and prove a priori error estimate. In Section 4, we perform several benchmark problems for Helmholtz problem to demonstrate the efficiency of the proposed method. Finally, in Section 5, we summarize the work and draw some conclusions.

2 The reconstructed finite element space

For a given polygonal domain Ω in \mathbb{R}^2 , its polygonal partition is \mathcal{T}_h , and the element is denoted by K and $\cup_{K \in \mathcal{T}_h} K = \Omega$. h_K denotes its diameter and $h := \max_{K \in \mathcal{T}_h} h_K$. Let \mathcal{E}_h denote the union of boundaries of element $K \in \mathcal{T}_h$. The partition \mathcal{T}_h must satisfy the same shape regularity conditions **A1**, **A2** as in [24]: There exist an integer number N independent of h , a real positive number σ independent of h , a compatible sub-decomposition $\tilde{\mathcal{T}}_h$, such that

- **A1.** Any element K admits a decomposition $\tilde{\mathcal{T}}_h|_K$ that consists of at most N triangles τ .
- **A2.** Any $\tau \in \tilde{\mathcal{T}}_h$ is shape-regular in the sense of Ciarlet-Raviart [11]: there exists σ such that $h_\tau/\rho_\tau \leq \sigma$ where h_τ is the diameter of τ and ρ_τ is the radius of the largest ball inscribed in τ .

Assumptions **A1** and **A2** impose quite weak constraints on the triangulation, which may contain elements with quite general shapes, for example, non-convex or degenerate elements are allowed.

The above shape regularity assumptions lead to some useful consequences, which will be extensively used in the later analysis.

- **M1.** For any $\tau \in \tilde{\mathcal{T}}_h$, there exists $\rho_1 \geq 1$ that depends on N and σ such that $h_K/h_\tau \leq \rho_1$.
- **M2 (Agmon inequality).** For all $v \in H^1(K)$, there exists C that depends on N and σ , but independent of h_K such that

$$\|v\|_{L^2(\partial K)}^2 \leq C \left(h_K^{-1} \|v\|_{L^2(K)}^2 + h_K \|\nabla v\|_{L^2(K)}^2 \right). \quad (2.1)$$

- **M3 (Approximation property).** For any positive integer m , there exists C that depends on N , m and σ , but independent of h_K such that for any $v \in H^{m+1}(K)$, there exists an approximation polynomial $\tilde{v} \in \mathbb{P}^m(K)$ such that

$$\|v - \tilde{v}\|_{L^2(K)} + h_K \|\nabla(v - \tilde{v})\|_{L^2(K)} \leq Ch_K^{m+1} |v|_{H^{m+1}(K)}. \quad (2.2)$$

- **M4 (Inverse inequality).** For any $v \in \mathbb{P}^m(K)$, there exists a constant C that depends only on N , m , σ and ρ_1 such that

$$\|\nabla v\|_{L^2(K)} \leq Ch_K^{-1} \|v\|_{L^2(K)}. \quad (2.3)$$

The proofs of above conclusions can be found in [24]. Moreover, we introduce some notations which will be used later. Let D denote a subdomain of Ω , it may be an element in \mathcal{T}_h or an aggregation of the elements belonging to \mathcal{T}_h . $H^m(D)$ and $\mathbb{P}^m(D)$ denote the

Sobolev space and the polynomial space with a total degree not greater than m on D , respectively.

Let U_h be the space consisting of piecewise constant functions related to \mathcal{T}_h :

$$U_h = \{v \in L^2(\Omega) \mid v|_K \in \mathbb{P}^0(K), \forall K \in \mathcal{T}_h\}.$$

Given the qualified partition \mathcal{T}_h , we need to define reconstruction spaces V_h and operators \mathcal{R} . The reconstruction operator \mathcal{R} is introduced in [24] which maps the piecewise constant space to discontinuous piecewise polynomial space V_h . Here we restate the basic process to construct the space and the operator to unify the notations.

The reconstruction space and the operator rely on two sets, namely, the element patch $S(K)$ and the sampling nodes $\mathcal{I}(K)$. The element patch $S(K)$ is a collection of the elements which contain K itself and neighboring elements around K . For each element K , a sampling node x_K is assigned in the interior of K . Let $\mathcal{I}(K)$ denote the set of sampling nodes belong to $S(K)$. For brevity, we omit the principles of how to choose the element patches and the sampling nodes here, the details can be found in [24].

For $\forall v \in U_h$ and $\forall K \in \mathcal{T}_h$, solving the following discrete least-square problem gives the local reconstruction $\mathcal{R}_K v$ on $S(K)$

$$\mathcal{R}_K v = \operatorname{argmin}_{p \in \mathbb{P}^m(S(K))} \sum_{x \in \mathcal{I}_K} |v(x) - p(x)|^2, \quad (2.4)$$

where the operator \mathcal{R}_K maps the piece-wise polynomial v to a m -th order polynomial p on $S(K)$.

Although $\mathcal{R}_K v$ gives the approximate polynomial on the element patch $S(K)$, we limit the approximate polynomial just on the element K . Thereafter, the global reconstruction operator \mathcal{R} is defined as follow,

$$\mathcal{R}|_K = \mathcal{R}_K.$$

As long as the reconstruction operator is given, the approximation space appears automatically. The operator \mathcal{R} embed the space U_h into an m -order discontinuous polynomial finite element space. The approximation space is defined by

$$V_h = \mathcal{R}U_h.$$

Before introducing the approximation property of \mathcal{R} , we discuss the size of element patch $S(K)$ first. Next, we compute the degree of freedom of the m -order polynomial, we have:

$$nop = \frac{(m+1)(m+2)}{2}.$$

Then the size of element patch denoted by $nppts$, should satisfy:

$$nppts \geq nop = \frac{(n+1)(n+2)}{2}.$$

Now we are ready to state the approximation property of the operator \mathcal{R} . The detailed proof can be found in [24]. Define

$$d_K := \text{diam}S(K) \quad \text{and} \quad d = \max_{K \in \mathcal{T}_h} d_K.$$

Moreover, we assume the sampling node set $\mathcal{I}(K)$ satisfy the following assumption.

Assumption 2.1. For every $K \in \mathcal{T}_h$, $p \in \mathbb{P}^m(S(K))$,

$$p|_{\mathcal{I}(K)} = 0 \quad \text{implies} \quad p|_{S(K)} \equiv 0. \quad (2.5)$$

Assumption 2.1 guarantees the uniqueness of solution to the least squares problem (2.4). It also implies the element patch $S(K)$ must be large enough, and the quantitative estimate is as follows,

$$\Lambda(m, \mathcal{I}(K)) < \infty$$

with

$$\Lambda(m, \mathcal{I}(K)) := \max_{p \in \mathbb{P}^m(S(K))} \frac{\|p\|_{L^\infty(S(K))}}{\|p|_{\mathcal{I}(K)}\|_{\ell_\infty}}. \quad (2.6)$$

The constant $\Lambda(m, \mathcal{I}(K))$ is similar to the Lebesgue constant in the approximation theory. We refer to [24] for the constraints of the uniform upper bound of $\Lambda(m, \mathcal{I}(K))$. The local reconstruction operator has the following approximation property.

Lemma 2.1 ([24, Lemma 3]). *If Assumption 2.1 holds, then there exists a unique solution to (2.4). Furthermore, $\mathcal{R}_K q$ satisfy*

$$\mathcal{R}_K q = q \quad \text{for all } q \in \mathbb{P}^m(S(K)). \quad (2.7)$$

For $q \in C^0(S(K))$, the stability property holds

$$\|\mathcal{R}_K q\|_{L^\infty(K)} \leq \Lambda(m, \mathcal{I}(K)) \sqrt{\#\mathcal{I}(K)} \|q|_{\mathcal{I}(K)}\|_{\ell_\infty}, \quad (2.8)$$

and the quasi-optimal approximation property is valid

$$\|q - \mathcal{R}_K q\|_{L^\infty(K)} \leq \Lambda_m \inf_{p \in \mathbb{P}^m(S(K))} \|q - p\|_{L^\infty(S(K))}, \quad (2.9)$$

where

$$\Lambda_m := \max_{K \in \mathcal{T}_h} \left\{ 1 + \Lambda(m, \mathcal{I}(K)) \sqrt{\#\mathcal{I}(K)} \right\}.$$

The optimal approximation property follows as

Lemma 2.2 ([24, Lemma 4]). *If Assumption 2.1 holds, for $q \in C^0(S(K)) \cap H^{m+1}(S(K))$, then there exists C such that*

$$\|q - \mathcal{R}_K q\|_{L^2(K)} \leq C \Lambda_m h_K d_K^m |q|_{H^{m+1}(S(K))}, \quad (2.10a)$$

$$\|\nabla(q - \mathcal{R}_K q)\|_{L^2(K)} \leq C (h_K^m + \Lambda_m d_K^m) |q|_{H^{m+1}(S(K))}. \quad (2.10b)$$

The global reconstruction operators \mathcal{R} and \mathcal{S} have the following approximation estimates.

Lemma 2.3. For $q \in H^{m+1}(\Omega)$, together with the Agmon inequality and the local approximation estimates (2.10a) and (2.10b), there exists a positive constant C , such that

$$\|q - \mathcal{R}q\|_{L^2(\Omega)} \leq C\Lambda_m h^{m+1} |q|_{H^{m+1}(\Omega)}, \tag{2.11a}$$

$$\|\nabla(q - \mathcal{R}q)\|_{L^2(\Omega)} \leq C\Lambda_m h^m |q|_{H^{m+1}(\Omega)}, \tag{2.11b}$$

where d is eliminated by h and C depends on the recursion depth t of the element patch.

3 Helmholtz problems

Now we consider the Helmholtz problems:

$$\begin{cases} -\Delta u - k^2 u = f & \text{in } \Omega, \\ \nabla u \cdot n + iku = g & \text{on } \partial\Omega, \end{cases} \tag{3.1}$$

where k is a given positive number and known as the wave number and i denotes the imaginary unit.

The discretized variational problem for Eqs. (3.1) reads: find $u_h \in U_h$, such that

$$a_h(\mathcal{R}u_h, \mathcal{R}v) - k^2(\mathcal{R}u_h, \mathcal{R}v) = (f, \mathcal{R}v)_h + \int_{\partial\Omega} g \overline{\mathcal{R}v} dx, \quad \forall v \in U_h. \tag{3.2}$$

The symmetric interior penalty method is employed to discretize the elliptic operator. For the second order elliptic operator, $a_h(\cdot, \cdot)$ is

$$\begin{aligned} a_h(v, w) := & \sum_{K \in \mathcal{T}_h} \int_K \nabla v \nabla \bar{w} dx - \sum_{e \in \mathcal{E}_h^i} \int_e ([[v]]) \{ \nabla \bar{w} \} + r [[\bar{w}]] \{ \nabla v \} ds \\ & + \sum_{e \in \mathcal{E}_h^i} i \int_e \eta_e h_e^{-1} [[v]] [[\bar{w}]] ds + \sum_{e \in \mathcal{E}_h^b} i \int_e k v \bar{w} ds, \end{aligned} \tag{3.3}$$

and

$$(f, \mathcal{R}v)_h := \sum_{K \in \mathcal{T}_h} \int_K f \overline{\mathcal{R}v} dx,$$

where η_e is a positive constant. More precisely, for $r = 1$, we obtain the classical (symmetric) interior penalty (SIP) method, for $r = -1$ the stabilized version of the Baumann-Oden method, usually referred to as non-symmetric interior penalty (NIP) method, and for $r = 0$ the incomplete interior penalty (IIP) method. We refer readers to [2, 9] for those interior penalty method. In this paper, we only consider the error estimation of the eigenvalue and eigenfunction with SIP method, but several benchmark problems for NIP and IIP method are carried out in Section 4.

Let \mathcal{E}_h^i denote the collection of the interior faces. The set of boundary faces is denoted as \mathcal{E}_h^b , and $\mathcal{E}_h = \mathcal{E}_h^i \cup \mathcal{E}_h^b$. Let e be an interior face shared by two neighboring elements K^+, K^- , and $\mathbf{n}^+, \mathbf{n}^-$ denote the corresponding outward unit normals. For the scalar-valued function q and the vector-valued function v , the average operator $\{\cdot\}$ and the jump operator $[[\cdot]]$ are defined as

$$\{v\} = \frac{1}{2}(v^+ + v^-), \quad [[v]] = \mathbf{n}^+ \cdot v^+ + \mathbf{n}^- \cdot v^-.$$

Here $v^+ = v|_{K^+}$ and $v^- = v|_{K^-}$. For $e \in \mathcal{E}_h^b$, we set

$$\{v\} = v|_K, \quad [[v]] = \mathbf{n} \cdot v|_K.$$

We define the energy norm $\|\cdot\|_h$ for any $v \in V_h = \mathcal{R}U_h$,

$$\|v\|_h^2 = \sum_{K \in \mathcal{T}_h} \|\nabla v\|_{L^2(K)}^2 + \sum_{e \in \mathcal{E}_h^i} h_e^{-1} \|[[v]]\|_{L^2(e)}^2 + \sum_{e \in \mathcal{E}_h^b} k \|v\|_{L^2(e)}^2 + \sum_{K \in \mathcal{T}_h} k^2 \|v\|_{L^2(K)}^2. \quad (3.4)$$

From the Agmon inequality (2.3) and Lemma 2.3, the following interpolation estimates are straightforward for the reconstruction operator in the energy norm.

For $g \in H^{m+1}(\Omega)$, there exists C that depends on N, σ, γ and m such that

$$\|g - \mathcal{R}g\|_h^2 \leq C \Lambda_m^2 h^{2m} (1 + kh + k^2 h^2) |g|_{H^{m+1}(\Omega)}^2. \quad (3.5)$$

By definition, we obtain the consistency of a_h in the sense that

$$a_h(u, \mathcal{R}v) - k^2(u, \mathcal{R}v) = (f, \mathcal{R}v) + \int_{\partial\Omega} g \overline{\mathcal{R}v} dx, \quad \forall v \in U_h. \quad (3.6)$$

Therefore, the following Galerkin orthogonality holds true.

$$a_h(u - \mathcal{R}u_h, \mathcal{R}v) - k^2(u - \mathcal{R}u_h, \mathcal{R}v) = 0, \quad \forall v \in U_h. \quad (3.7)$$

Next, we define an auxiliary bilinear form b_h and a mesh-dependent norm $\|\cdot\|$:

$$b_h(u, v) := a_h(u, v) + (k^2 u, v). \quad (3.8)$$

By the definition of energy norm and Cauchy's inequality, we have

$$|a_h(\mathcal{R}v, \mathcal{R}w) \pm k^2(\mathcal{R}v, \mathcal{R}w)| \leq 2 \| \mathcal{R}u \| \| \mathcal{R}w \|, \quad \forall v, w \in U_h. \quad (3.9)$$

Next, we need to prove the coercivity of b_h .

Lemma 3.1. *There is a constant $\lambda > 0$ independent of h such that for any $v \in U_h$, we have*

$$b_h(\mathcal{R}v, \mathcal{R}v) \geq \lambda \| \mathcal{R}v \|. \quad (3.10)$$

Proof. Let $v \in U_h$. By definition of b_h we have

$$\begin{aligned}
 b_h(\mathcal{R}v, \mathcal{R}v) &= \sum_{K \in \mathcal{T}_h} \|\nabla \mathcal{R}v\|_{L^2(K)}^2 - \sum_{e \in \mathcal{E}_h^i} 2Re \int_e [[\mathcal{R}v]] \{\nabla \overline{\mathcal{R}v}\} ds \\
 &\quad + \sum_{e \in \mathcal{E}_h^i} i\eta_e h_e^{-1} \|\llbracket \mathcal{R}v \rrbracket\|_{L^2(e)}^2 + \sum_{e \in \mathcal{E}_h^b} ik \|\mathcal{R}v\|_{L^2(e)}^2 + \sum_{K \in \mathcal{T}_h} k^2 \|\mathcal{R}v\|_{L^2(K)}^2. \quad (3.11)
 \end{aligned}$$

Using the Cauchy inequality and Young's inequality, for any $s > 0$, we have

$$\left| 2Re \int_e [[\mathcal{R}v]] \{\nabla \overline{\mathcal{R}v}\} ds \right| \leq \frac{s}{h} \|\llbracket \mathcal{R}v \rrbracket\|_{L^2(e)}^2 + \frac{h}{s} \|\{\nabla \overline{\mathcal{R}v}\}\|_{L^2(e)}^2. \quad (3.12)$$

Then we have

$$\begin{aligned}
 \sqrt{2}|b_h(\mathcal{R}v, \mathcal{R}v)| &\geq Re b_h(\mathcal{R}v, \mathcal{R}v) + Im b_h(\mathcal{R}v, \mathcal{R}v) \\
 &\geq \sum_{K \in \mathcal{T}_h} \|\nabla \mathcal{R}v\|_{L^2(K)}^2 - \sum_{e \in \mathcal{E}_h^i} \left(\frac{s}{h} \|\llbracket \mathcal{R}v \rrbracket\|_{L^2(e)}^2 + \frac{h}{s} \|\{\nabla \overline{\mathcal{R}v}\}\|_{L^2(e)}^2 \right) \\
 &\quad + \sum_{e \in \mathcal{E}_h^i} \eta_e h_e^{-1} \|\llbracket \mathcal{R}v \rrbracket\|_{L^2(e)}^2 + \sum_{e \in \mathcal{E}_h^b} k \|\mathcal{R}v\|_{L^2(e)}^2 + \sum_{K \in \mathcal{T}_h} k^2 \|\mathcal{R}v\|_{L^2(K)}^2.
 \end{aligned}$$

By Agmon inequality, we have

$$\|\{\nabla \overline{\mathcal{R}v}\}\|_{L^2(e)}^2 \leq C_{inv} h_K^{-1} \|\{\nabla \overline{\mathcal{R}v}\}\|_{L^2(K)}^2.$$

Then we have

$$\begin{aligned}
 \sqrt{2}|b_h(\mathcal{R}v, \mathcal{R}v)| &\geq \sum_{K \in \mathcal{T}_h} \|\nabla \mathcal{R}v\|_{L^2(K)}^2 - \sum_{K \in \mathcal{T}_h} t \|\nabla \mathcal{R}v\|_{L^2(K)}^2 + \sum_{e \in \mathcal{E}_h^i} \frac{th}{C_{inv}} \|\{\nabla \overline{\mathcal{R}v}\}\|_{L^2(e)}^2 \\
 &\quad - \sum_{e \in \mathcal{E}_h^i} \left(\frac{s}{h} \|\llbracket \mathcal{R}v \rrbracket\|_{L^2(e)}^2 + \frac{h}{s} \|\{\nabla \overline{\mathcal{R}v}\}\|_{L^2(e)}^2 \right) + \sum_{e \in \mathcal{E}_h^i} \eta_e h_e^{-1} \|\llbracket \mathcal{R}v \rrbracket\|_{L^2(e)}^2 \\
 &\quad + \sum_{e \in \mathcal{E}_h^b} k \|\mathcal{R}v\|_{L^2(e)}^2 + \sum_{K \in \mathcal{T}_h} k^2 \|\mathcal{R}v\|_{L^2(K)}^2 \\
 &\geq (1-t) \|\nabla \mathcal{R}v\|_{L^2(\Omega)}^2 + \sum_{e \in \mathcal{E}_h^i} \left(\frac{t}{C_{inv}} - \frac{1}{s} \right) h \|\{\nabla \overline{\mathcal{R}v}\}\|_{L^2(e)}^2 \\
 &\quad + \sum_{e \in \mathcal{E}_h^i} \left(\frac{\eta}{h} - \frac{s}{h} \right) \|\llbracket \mathcal{R}v \rrbracket\|_{L^2(e)}^2 + \sum_{e \in \mathcal{E}_h^b} k \|\mathcal{R}v\|_{L^2(e)}^2 + k^2 \|\nabla \mathcal{R}v\|_{L^2(\Omega)}^2.
 \end{aligned}$$

If we choose t and s , such that $\frac{t}{C_{inv}} - \frac{1}{s} > 0$ and $\eta - s > 0, 1 - t > 0$, i.e., $1 > t > \frac{C_{inv}}{s}, s < \eta$. Then for sufficiently large η_e , the result follows. \square

Theorem 3.1. Let $u \in H^2(\Omega)$ be the exact solution to the problem (3.1) and $\mathcal{R}u_h$ be the discrete solution of (3.2), if Assumption B holds, then

$$\| \|u - \mathcal{R}u_h\| \| \leq \left(1 + \frac{2}{\lambda}\right) \| \|u - \mathcal{R}u\| \| + \frac{2k}{\lambda} \| \|u - \mathcal{R}u\| \|_{L^2(\Omega)}. \quad (3.13)$$

Proof. Denote $v = \mathcal{R}u - \mathcal{R}u_h$, we obtain

$$\begin{aligned} \| \| \mathcal{R}u - \mathcal{R}u_h \| \| ^2 &\leq \frac{1}{\lambda} b_h(\mathcal{R}u - \mathcal{R}u_h, \mathcal{R}u - \mathcal{R}u_h) \\ &\leq \frac{1}{\lambda} |b_h(\mathcal{R}u - u, \mathcal{R}u - \mathcal{R}u_h)| + \frac{1}{\lambda} |b_h(u - \mathcal{R}u_h, \mathcal{R}u - \mathcal{R}u_h)|. \end{aligned} \quad (3.14)$$

By Galerkin orthogonality

$$a_h(u - \mathcal{R}u_h, \mathcal{R}v) = k^2(u - \mathcal{R}u_h, \mathcal{R}v), \quad \forall v \in U_h, \quad (3.15)$$

we have

$$\begin{aligned} |b_h(u - \mathcal{R}u_h, \mathcal{R}u - \mathcal{R}u_h)| &= |a_h(u - \mathcal{R}u_h, \mathcal{R}u - \mathcal{R}u_h) + k^2(u - \mathcal{R}u_h, \mathcal{R}u - \mathcal{R}u_h)| \\ &= 2|k^2(u - \mathcal{R}u_h, \mathcal{R}u - \mathcal{R}u_h)|. \end{aligned} \quad (3.16)$$

Then we will obtain

$$\| \| \mathcal{R}u - \mathcal{R}u_h \| \| ^2 \leq \frac{2}{\lambda} \| \|u - \mathcal{R}u\| \| \| \| \mathcal{R}u - \mathcal{R}u_h \| \| + \frac{2}{\lambda} |k^2(u - \mathcal{R}u_h, \mathcal{R}u - \mathcal{R}u_h)|. \quad (3.17)$$

By definition of the energy norm, we obtain

$$k \| \mathcal{R}u - \mathcal{R}u_h \|_{L^2(\Omega)} \leq \| \| \mathcal{R}u - \mathcal{R}u_h \| \| . \quad (3.18)$$

Then we have

$$\begin{aligned} |k^2(u - \mathcal{R}u_h, \mathcal{R}u - \mathcal{R}u_h)| &\leq k \| \|u - \mathcal{R}u_h\| \|_{L^2(\Omega)} \cdot k \| \| \mathcal{R}u - \mathcal{R}u_h \| \|_{L^2(\Omega)} \\ &\leq k \| \|u - \mathcal{R}u_h\| \|_{L^2(\Omega)} \| \| \mathcal{R}u - \mathcal{R}u_h \| \| . \end{aligned} \quad (3.19)$$

Then we obtain

$$\| \| \mathcal{R}u - \mathcal{R}u_h \| \| \leq \frac{2}{\lambda} \| \|u - \mathcal{R}u\| \| + \frac{2k}{\lambda} \| \|u - \mathcal{R}u_h\| \|_{L^2(\Omega)}. \quad (3.20)$$

From the above we can see

$$\begin{aligned} \| \|u - \mathcal{R}u_h\| \| &\leq \| \|u - \mathcal{R}u\| \| + \| \| \mathcal{R}u - \mathcal{R}u_h \| \| \\ &\leq \left(1 + \frac{2}{\lambda}\right) \| \|u - \mathcal{R}u\| \| + \frac{2k}{\lambda} \| \|u - \mathcal{R}u_h\| \|_{L^2(\Omega)}. \end{aligned} \quad (3.21)$$

Thus, we complete the proof. \square

Theorem 3.2. Let Ω be a bounded convex domain. Consider the following adjoint problem of (3.1):

$$\begin{cases} -\Delta\varphi - k^2\varphi = \omega & \text{in } \Omega, \\ \nabla\varphi \cdot n + ik\varphi = 0 & \text{on } \partial\Omega, \end{cases} \quad (3.22)$$

where $\omega \in L^2(\Omega)$ and the solution $\varphi \in H^2(\Omega)$. Then there is a constant $C_1 > 0$ only depending on Ω such that

$$\|\varphi\|_{H^2(\Omega)} \leq C_1(1+k)\|\omega\|_{L^2(\Omega)}. \quad (3.23)$$

Proof. From [31, Theorem 4.4], we obtain the result directly. \square

Theorem 3.3. Let $u \in H^{m+1}(\Omega)$ be the exact solution of (3.1). Let $u_h \in U_h$ be the discrete solution of the variational problem (3.2). Provided that the following threshold condition holds:

$$\frac{4k}{\lambda}CC_1\Lambda_m h(1+kh+k^2h^2)^{1/2}(1+k) < 1,$$

then there is a constant C only depends on $\Omega, N, \sigma, \gamma$ such that

$$\| \|u - \mathcal{R}u_h\| \| \leq Ch^m(1+kh+k^2h^2)^{1/2}|u|_{H^{m+1}(\Omega)}, \quad (3.24a)$$

$$\| \|u - \mathcal{R}u_h\|_{L^2(\Omega)} \| \leq Ch^{m+1}(1+kh+k^2h^2)(1+k)|u|_{H^{m+1}(\Omega)}. \quad (3.24b)$$

Proof. Using (3.5) and let $m=1, g=\varphi$, we obtain

$$\begin{aligned} \| \|\varphi - \mathcal{R}\varphi\| \| &\leq C\Lambda_m h(1+kh+k^2h^2)^{1/2}\|\varphi\|_{H^2(\Omega)} \\ &\leq CC_1\Lambda_m h(1+kh+k^2h^2)^{1/2}(1+k)\|\omega\|_{L^2(\Omega)}. \end{aligned} \quad (3.25)$$

If we take $\omega = u - \mathcal{R}u_h$, then

$$\| \|\varphi - \mathcal{R}\varphi\| \| \leq CC_1\Lambda_m h(1+kh+k^2h^2)^{1/2}(1+k)\|u - \mathcal{R}u_h\|_{L^2(\Omega)}. \quad (3.26)$$

By adjoint problem, we obtain

$$\| \|u - \mathcal{R}u_h\|_{L^2(\Omega)}^2 \| = a_h(u - \mathcal{R}u_h, \varphi) - k^2(u - \mathcal{R}u_h, \varphi). \quad (3.27)$$

Using Galerkin orthogonality (3.7), we have

$$\begin{aligned} \| \|u - \mathcal{R}u_h\|_{L^2(\Omega)}^2 \| &= a_h(u - \mathcal{R}u_h, \varphi - \mathcal{R}\varphi) - k^2(u - \mathcal{R}u_h, \varphi - \mathcal{R}\varphi) \\ &\leq 2\| \|u - \mathcal{R}u_h\| \| \| \|\varphi - \mathcal{R}\varphi\| \|. \end{aligned} \quad (3.28)$$

Combining with (3.25), we obtain

$$\| \|u - \mathcal{R}u\|_{L^2(\Omega)} \| \leq 2CC_1\Lambda_m h(1+kh+k^2h^2)^{1/2}(1+k)\| \|u - \mathcal{R}u_h\| \|. \quad (3.29)$$

By Theorem 3.1, we obtain

$$\| \|u - \mathcal{R}u_h\| \| \leq \left(1 + \frac{2}{\lambda}\right) \| \|u - \mathcal{R}u\| \| + \frac{4k}{\lambda} CC_1 \Lambda_m h (1 + kh + k^2 h^2)^{1/2} (1 + k) \| \|u - \mathcal{R}u_h\| \|. \quad (3.30)$$

If we assume that

$$\frac{4k}{\lambda} CC_1 \Lambda_m h (1 + kh + k^2 h^2)^{1/2} (1 + k) < 1,$$

we will obtain that there exists a constant C_{inf} , such that

$$\| \|u - \mathcal{R}u_h\| \| \leq C_{inf} \| \|u - \mathcal{R}u\| \|.$$

Combining with (3.5), we obtain

$$\| \|u - \mathcal{R}u_h\| \| \leq CC_{inf} \Lambda_m h^m (1 + kh + k^2 h^2)^{1/2} |u|_{H^{m+1}(\Omega)}, \quad (3.31a)$$

$$\| \|u - \mathcal{R}u_h\|_{L^2(\Omega)} \| \leq 2C^2 C_1 C_{inf}^2 \Lambda_m^2 h^{m+1} (1 + kh + k^2 h^2) (1 + k) |u|_{H^{m+1}(\Omega)}. \quad (3.31b)$$

So, we complete the proof. \square

4 Numerical experiments

In this section, we present some numerical results to show that our method is efficient for Helmholtz problems, and verify the theoretical estimates.

4.1 Numerical example

Here we present several benchmark problems for Helmholtz problems. As mentioned in (3.3), we take η_e as a positive constant. Before demonstrating the numerical results, we want to illustrate the purpose of the two examples. Example 4.1 is a smooth case which is served as the theoretical verification. Example 4.2 employs three different methods with the same uniform triangle mesh. The comparison wants to exhibit the proposed method possesses the high efficiency of utilizing the DOFs. The relative L2-error $\| \|u_h - u\|_{L^2}$ and the energy norm error $\| \|u_h - u\| \|$ are studied.

We consider the following 2D Helmholtz problem:

$$\begin{cases} -\Delta u - k^2 u = f := \frac{\sin(kr)}{r} & \text{in } \Omega, \\ \nabla u \cdot n + iku = g & \text{on } \partial\Omega. \end{cases} \quad (4.1)$$

Here g is chosen so that the exact solution is

$$u = \frac{\cos(kr)}{k} - \frac{\cos(k) + i\sin(k)}{k(J_0(k) + iJ_1(k))} J_0(kr)$$

in polar coordinates, where $J_\alpha(z)$ are Bessel functions of the first kind.

Example 4.1. In this example, we consider SIP method and the other numerical settings are as follows, the computational domain is square domain $\Omega = [0,1]^2$, the domain is partitioned into quasi-uniform triangular by *Gmsh* [16]. The size of the element patch is presented in Table 1, where m is the order of polynomials.

Table 1: The size of element patch $S(K)$ in Example 4.1 with different m -th order polynomial.

$m \backslash k$	$k=0.01$	$k=50$	$k=100$
$m=1$	4	4	5
$m=2$	8	10	8
$m=3$	17	17	15
$m=4$	19	19	20
$m=5$	28	28	28

Table 2: The relative L2-errors and convergence rates of SIP method with $k=0.01$ in Example 4.1.

Order	$h=1E-1$	$h=5E-2$	Rate	$h=2.5E-2$	Rate	$1.25E-2$	Rate
$m=1$	$3.10E-05$	$7.28E-06$	2.09	$1.63E-06$	2.16	$3.85E-07$	2.08
$m=2$	$1.16E-08$	$1.32E-09$	3.13	$1.37E-10$	3.27	$1.69E-11$	3.02
$m=3$	$8.43E-10$	$3.84E-11$	4.46	$2.12E-12$	4.18	$1.50E-13$	3.82

Table 3: The relative energy norm errors and convergence rates of SIP method with $k=0.01$ in Example 4.1.

Order	$h=1E-1$	$h=5E-2$	Rate	$h=2.5E-2$	Rate	$1.25E-2$	Rate
$m=1$	$1.51E-03$	$7.41E-04$	1.02	$3.69E-04$	1.00	$1.79E-04$	1.04
$m=2$	$5.46E-07$	$1.23E-07$	2.16	$2.63E-08$	2.22	$6.55E-09$	2.00
$m=3$	$3.97E-08$	$3.94E-09$	3.33	$4.51E-10$	3.13	$5.56E-11$	3.02

Table 4: The relative L2-errors and convergence rates of SIP method with $k=50$ in Example 4.1.

Order	$h=5E-2$	$h=2.5E-2$	Rate	$1.25E-2$	Rate	$6.25E-3$	Rate	$3.125E-3$	Rate
$m=1$	$7.41E-03$	$1.26E-03$	2.56	$4.10E-04$	1.62	$9.66E-05$	2.08	$2.38E-05$	2.02
$m=2$	$4.50E-03$	$3.89E-04$	3.53	$4.88E-05$	2.99	$6.19E-06$	2.98	$7.73E-07$	3.00
$m=3$	$3.06E-03$	$1.71E-04$	4.16	$1.09E-05$	3.98	$6.08E-07$	4.16	$3.84E-08$	3.99
$m=4$	$1.53E-03$	$5.72E-05$	4.74	$1.90E-06$	4.91	$5.00E-08$	5.25	$1.44E-09$	5.12
$m=5$	$2.90E-03$	$3.70E-05$	6.29	$5.03E-07$	6.20	$6.74E-09$	6.22	$9.37E-11$	6.17

Tables 2, 3 and Fig. 1 demonstrate the relative error and the convergence order with $k=0.01$. The convergence order can be observed which agree with the Theorem 3.3. The reason of showing until $m=3$ or $h=1.25E-2$ with $k=0.01$ is that the relative error achieves the machine precision with the highest approximation order and finest mesh.

Tables 4, 5 and Fig. 2 demonstrate the relative error and the convergence order with $k=50$. We can see that when h tends to zero, the convergence order conforms to the theorem.

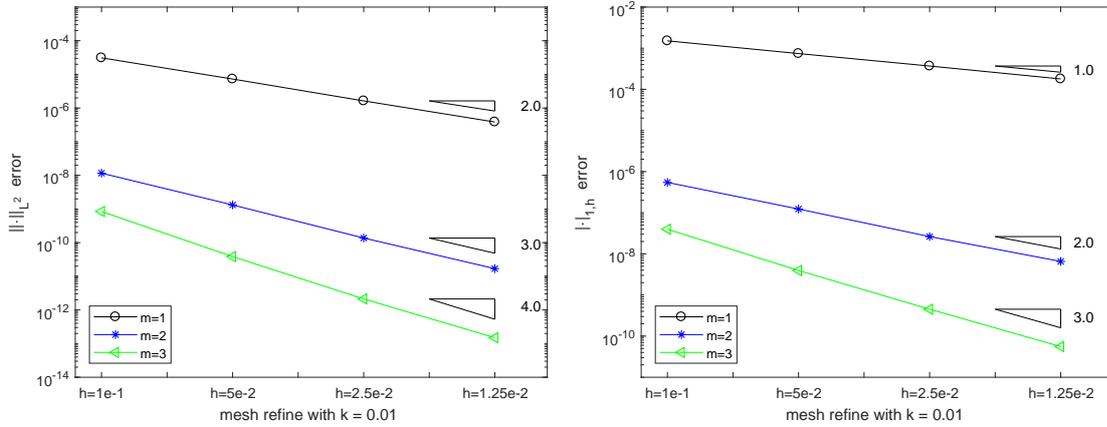


Figure 1: The convergence rates of relative L2-errors (left)/energy-norm errors (right) with SIP method and $k=0.01$ in Example 4.1.

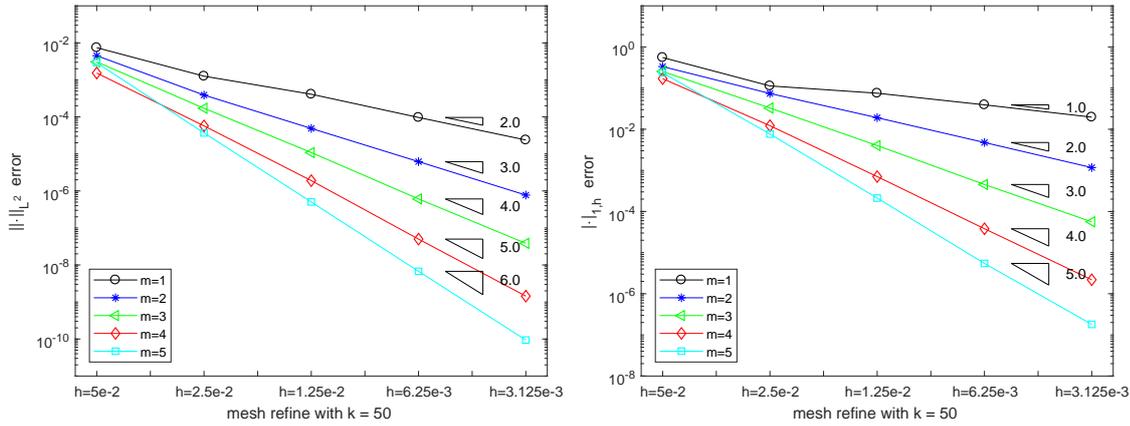


Figure 2: the convergence rates of relative L2-errors (left)/energy-norm errors (right) with SIP method and $k=50$ in Example 4.1.

Tables 6, 7 and Fig. 3 demonstrate the relative error and the convergence order with $k=100$. We can see that when h tends to zero, the convergence order conforms to the

Table 5: The relative energy norm errors and convergence rates of SIP method with $k=50$ in Example 4.1.

Order	$h=5E-2$	$h=2.5E-2$	Rate	$1.25E-2$	Rate	$6.25E-3$	Rate	$3.125E-3$	Rate
$m=1$	$5.54E-01$	$1.13E-01$	2.29	$7.52E-02$	0.59	$3.92E-02$	0.94	$1.97E-02$	0.99
$m=2$	$3.34E-01$	$7.40E-02$	2.18	$1.91E-02$	1.95	$4.76E-03$	2.01	$1.17E-03$	2.03
$m=3$	$2.54E-01$	$3.28E-02$	2.95	$4.01E-03$	3.03	$4.54E-04$	3.14	$5.62E-05$	3.01
$m=4$	$1.70E-01$	$1.22E-02$	3.80	$7.06E-04$	4.11	$3.81E-05$	4.21	$2.19E-06$	4.12
$m=5$	$2.35E-01$	$7.64E-03$	4.94	$2.12E-04$	5.17	$5.47E-06$	5.28	$1.80E-07$	4.93

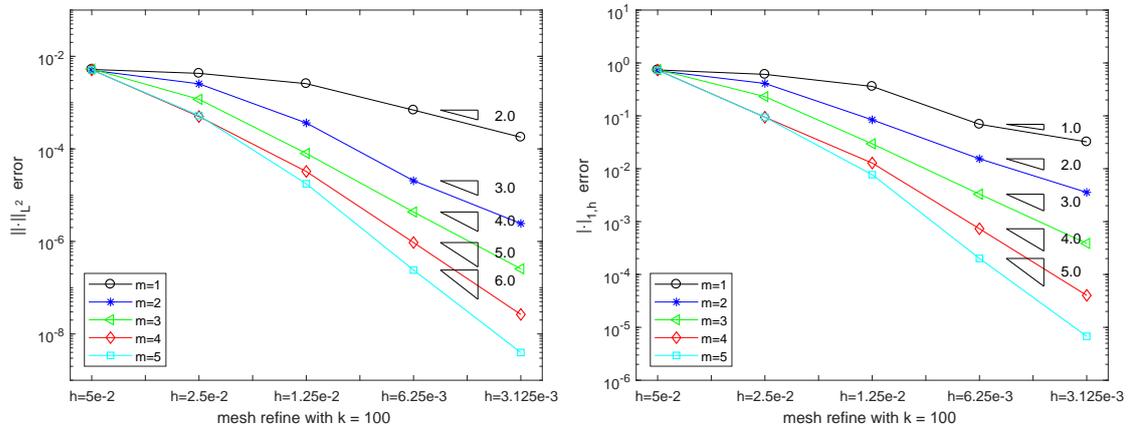


Figure 3: the convergence rates of relative L2-errors (left)/energy-norm errors (right) with SIP method and $k=100$ in Example 4.1.

Table 6: The relative L2-errors and convergence rates of SIP method with $k=100$ in Example 4.1.

Order	$h=5E-2$	$h=2.5E-2$	Rate	$1.25E-2$	Rate	$6.25E-3$	Rate	$3.125E-3$	Rate
$m=1$	$5.21E-03$	$4.29E-03$	0.28	$2.57E-03$	0.74	$6.89E-04$	1.90	$1.79E-04$	1.94
$m=2$	$5.00E-03$	$2.55E-03$	0.97	$3.64E-04$	2.81	$2.05E-05$	4.15	$2.44E-06$	3.07
$m=3$	$5.32E-03$	$1.18E-03$	2.17	$7.99E-05$	3.89	$4.33E-06$	4.21	$2.56E-07$	4.08
$m=4$	$5.13E-03$	$5.00E-04$	3.36	$3.25E-05$	3.94	$9.46E-07$	5.10	$2.64E-08$	5.17
$m=5$	$5.08E-03$	$5.26E-04$	3.27	$1.76E-05$	4.90	$2.41E-07$	6.19	$3.97E-09$	5.92

Table 7: The relative energy norm errors and convergence rates of SIP method with $k=100$ in Example 4.1.

Order	$h=5E-2$	$h=2.5E-2$	Rate	$1.25E-2$	Rate	$6.25E-3$	Rate	$3.125E-3$	Rate
$m=1$	$7.38E-01$	$6.09E-01$	0.28	$3.60E-01$	0.76	$6.92E-02$	2.38	$3.22E-02$	1.10
$m=2$	$7.14E-01$	$4.10E-01$	0.80	$8.41E-02$	2.29	$1.54E-02$	2.45	$3.57E-03$	2.11
$m=3$	$7.48E-01$	$2.31E-01$	1.70	$2.96E-02$	2.96	$3.39E-03$	3.16	$3.87E-04$	3.09
$m=4$	$7.34E-01$	$9.42E-02$	2.96	$1.28E-02$	2.88	$7.32E-04$	4.13	$4.03E-05$	4.18
$m=5$	$7.27E-01$	$9.45E-02$	2.94	$7.72E-03$	3.61	$2.01E-04$	5.28	$6.78E-06$	4.89

theorem.

Example 4.2. In this example, we consider NIP and IIP method and only consider $k=50$ and $k=100$. The other numerical settings are same with Example 4.1. Tables 8, 9, Tables 10, 11 demonstrate the relative error and the convergence order of NIP method with $k=50$ and $k=100$. Tables 12, 13, 14, and 15 demonstrate the relative error and the convergence order of IIP method with $k=50$ and $k=100$. We can observe that we obtain the ideal convergence order when h is small enough and the results are also coincide with the Theorem 3.3. It also can be seen that for k -th order polynomial degrees, if we consider relative energy norm error, the order of convergence is k . However if we consider relative L2-error and if k is even number, the order of convergence is k and if k is odd number, we

Table 8: The relative L2-errors and convergence rates of NIP method with $k=50$ in Example 4.2.

Order	$h=5E-2$	$h=2.5E-2$	Rate	$1.25E-2$	Rate	$6.25E-3$	Rate	$3.125E-3$	Rate
$m=1$	$1.01E-02$	$6.25E-03$	0.69	$2.12E-03$	1.56	$6.16E-04$	1.78	$1.68E-04$	1.87
$m=2$	$8.86E-03$	$7.88E-03$	0.17	$3.47E-03$	1.18	$7.12E-04$	2.29	$1.74E-04$	2.03
$m=3$	$8.10E-03$	$7.31E-04$	3.47	$4.79E-05$	3.93	$2.59E-06$	4.21	$1.65E-07$	3.97
$m=4$	$6.89E-03$	$8.76E-04$	2.97	$7.23E-05$	3.60	$3.82E-06$	4.24	$2.11E-07$	4.17
$m=5$	$3.80E-03$	$6.58E-05$	5.85	$3.03E-06$	4.44	$3.98E-08$	6.25	$6.55E-10$	5.93

Table 9: The relative energy norm errors and convergence rates of NIP method with $k=50$ in Example 4.2.

Order	$h=5E-2$	$h=2.5E-2$	Rate	$1.25E-2$	Rate	$6.25E-3$	Rate	$3.125E-3$	Rate
$m=1$	$4.95E-01$	$3.19E-01$	0.63	$1.21E-01$	1.40	$4.55E-02$	1.41	$1.90E-02$	1.26
$m=2$	$6.40E-01$	$5.66E-01$	0.18	$2.47E-01$	1.20	$5.19E-02$	2.25	$1.28E-02$	2.02
$m=3$	$5.79E-01$	$5.79E-02$	3.32	$3.37E-03$	4.10	$5.00E-04$	2.75	$6.11E-05$	3.03
$m=4$	$4.99E-01$	$6.20E-02$	3.01	$5.18E-03$	3.58	$2.74E-04$	4.24	$1.54E-05$	4.16
$m=5$	$2.84E-01$	$6.11E-03$	5.54	$2.00E-04$	4.93	$5.26E-06$	5.25	$1.54E-07$	5.09

Table 10: The relative L2-errors and convergence rates of NIP method with $k=100$ in Example 4.2.

Order	$h=5E-2$	$h=2.5E-2$	Rate	$1.25E-2$	Rate	$6.25E-3$	Rate	$3.125E-3$	Rate
$m=1$	$5.34E-03$	$4.95E-03$	0.11	$4.56E-03$	0.12	$3.44E-03$	0.41	$1.66E-03$	1.05
$m=2$	$5.18E-03$	$4.77E-03$	0.12	$4.17E-03$	0.19	$2.37E-03$	0.82	$6.26E-04$	1.92
$m=3$	$5.26E-03$	$4.55E-03$	0.21	$7.43E-04$	2.61	$5.69E-05$	3.71	$3.28E-06$	4.11
$m=4$	$5.18E-03$	$4.05E-03$	0.35	$1.21E-03$	1.75	$7.27E-05$	4.05	$4.51E-06$	4.01
$m=5$	$5.16E-03$	$2.67E-03$	0.95	$3.52E-05$	6.25	$5.37E-07$	6.04	$7.32E-09$	6.20

Table 11: The relative energy norm errors and convergence rates of NIP method with $k=100$ in Example 4.2.

Order	$h=5E-2$	$h=2.5E-2$	Rate	$1.25E-2$	Rate	$6.25E-3$	Rate	$3.125E-3$	Rate
$m=1$	$7.51E-01$	$7.02E-01$	0.10	$6.47E-01$	0.12	$4.88E-01$	0.41	$2.38E-01$	1.04
$m=2$	$7.33E-01$	$6.83E-01$	0.10	$5.94E-01$	0.20	$3.36E-01$	0.82	$8.84E-02$	1.93
$m=3$	$7.44E-01$	$6.50E-01$	0.20	$1.07E-01$	2.60	$7.30E-03$	3.87	$4.15E-04$	4.14
$m=4$	$7.35E-01$	$5.80E-01$	0.34	$1.71E-01$	1.76	$1.02E-02$	4.06	$6.36E-04$	4.01
$m=5$	$5.20E-01$	$2.78E-01$	0.91	$6.21E-03$	5.48	$1.64E-04$	5.25	$4.90E-06$	5.06

Table 12: The relative L2-errors and convergence rates of IIP method with $k=50$ in Example 4.2.

Order	$h=5E-2$	$h=2.5E-2$	Rate	$1.25E-2$	Rate	$6.25E-3$	Rate	$3.125E-3$	Rate
$m=1$	$1.00E-02$	$6.17E-03$	0.70	$2.19E-03$	1.49	$6.08E-04$	1.85	$1.58E-04$	1.95
$m=2$	$8.70E-03$	$5.14E-03$	0.76	$1.43E-03$	1.85	$3.04E-04$	2.23	$7.48E-05$	2.02
$m=3$	$8.02E-03$	$6.94E-04$	3.53	$3.95E-05$	4.13	$2.05E-06$	4.27	$1.34E-07$	3.94
$m=4$	$5.89E-03$	$4.20E-04$	3.81	$3.64E-05$	3.53	$1.91E-06$	4.25	$1.06E-07$	4.17
$m=5$	$4.43E-03$	$3.89E-05$	6.83	$1.47E-06$	4.72	$2.34E-08$	5.98	$3.89E-10$	5.91

observe the order of convergence is $k+1$.

Example 4.3. The square domain $\Omega = [0,1]^2$ is still considered, and isosceles right-angled

Table 13: The relative energy norm errors and convergence rates of IIP method with $k=50$ in Example 4.2.

Order	$h=5E-2$	$h=2.5E-2$	Rate	$1.25E-2$	Rate	$6.25E-3$	Rate	$3.125E-3$	Rate
$m=1$	$7.04E-01$	$4.42E-01$	0.67	$1.67E-01$	1.41	$5.45E-02$	1.61	$2.04E-02$	1.42
$m=2$	$6.31E-01$	$3.78E-01$	0.74	$1.04E-01$	1.87	$2.24E-02$	2.21	$5.54E-03$	2.01
$m=3$	$5.71E-01$	$5.73E-02$	3.32	$3.65E-03$	3.97	$4.56E-04$	3.00	$5.60E-05$	3.03
$m=4$	$4.33E-01$	$2.88E-02$	3.91	$2.57E-03$	3.49	$1.34E-04$	4.25	$7.65E-06$	4.14
$m=5$	$2.32E-01$	$6.03E-03$	5.27	$2.20E-04$	4.78	$5.74E-06$	5.26	$1.56E-07$	5.18

Table 14: The relative L2-errors and convergence rates of IIP method with $k=100$ in Example 4.2.

Order	$h=5E-2$	$h=2.5E-2$	Rate	$1.25E-2$	Rate	$6.25E-3$	Rate	$3.125E-3$	Rate
$m=1$	$5.34E-03$	$4.95E-03$	0.11	$4.56E-03$	0.12	$3.43E-03$	0.41	$1.66E-03$	1.05
$m=2$	$5.17E-03$	$4.70E-03$	0.14	$3.26E-03$	0.53	$1.27E-03$	1.36	$3.06E-04$	2.06
$m=3$	$5.26E-03$	$4.54E-03$	0.21	$5.65E-04$	3.01	$2.88E-05$	4.29	$1.70E-06$	4.09
$m=4$	$5.18E-03$	$3.49E-03$	0.57	$5.98E-04$	2.54	$3.63E-05$	4.04	$2.27E-06$	4.00
$m=5$	$5.15E-03$	$2.58E-03$	1.00	$4.45E-05$	5.86	$7.78E-07$	5.84	$1.40E-08$	5.80

Table 15: The relative energy norm errors and convergence rates of IIP method with $k=100$ in Example 4.2.

Order	$h=5E-2$	$h=2.5E-2$	Rate	$1.25E-2$	Rate	$6.25E-3$	Rate	$3.125E-3$	Rate
$m=1$	$7.52E-01$	$7.01E-01$	0.10	$6.46E-01$	0.12	$4.87E-01$	0.41	$2.37E-01$	1.04
$m=2$	$7.32E-01$	$6.75E-01$	0.12	$4.71E-01$	0.52	$1.81E-01$	1.38	$4.29E-02$	2.07
$m=3$	$7.44E-01$	$6.48E-01$	0.20	$8.49E-02$	2.93	$2.85E-03$	4.90	$3.81E-04$	2.90
$m=4$	$7.35E-01$	$5.05E-01$	0.54	$8.39E-02$	2.59	$5.04E-03$	4.06	$3.17E-04$	3.99
$m=5$	$7.32E-01$	$3.72E-01$	0.98	$6.19E-03$	5.91	$1.56E-04$	5.31	$5.01E-06$	4.96

triangles uniformly partition it. In this example, we give a numerical comparison among finite element method, discontinuous Galerkin method and the proposed method and we take SIP method in this example. The first two methods are preformed by FreeFem++ [17] with \mathbb{P}^k element pair ($k=2,3$), as \mathbb{P}^1 element pair has no convergence order when $k=100$. The results are consistent with those in [31, Table 2]. The meshes size list as follows: $10*10$, $20*20$, $40*40$, $80*80$, then we investigate the relative energy norm error with $k=100$. Fig. 4 shows the numerical performance of the different methods in which the horizontal ordinate is the number of the DOFs, and the vertical coordinate is also the relative energy norm error. The figure is the log-log scale plot that is capable of illustrating the convergence rate explicitly. Firstly, we can observe that all the methods achieve the optimal convergence rate as expected. The comparison of the three methods is conducted in the low-order approximation. The DG method is weak in the DOFs utilizing efficiency among the three methods. Meanwhile, the proposed method is comparable with the FEM. Finally, we focus on the performance of the proposed method. It employs the identical number of DOFs for arbitrary order approximations. The efficiency of utilizing the DOFs is improving while the approximation order is increasing.

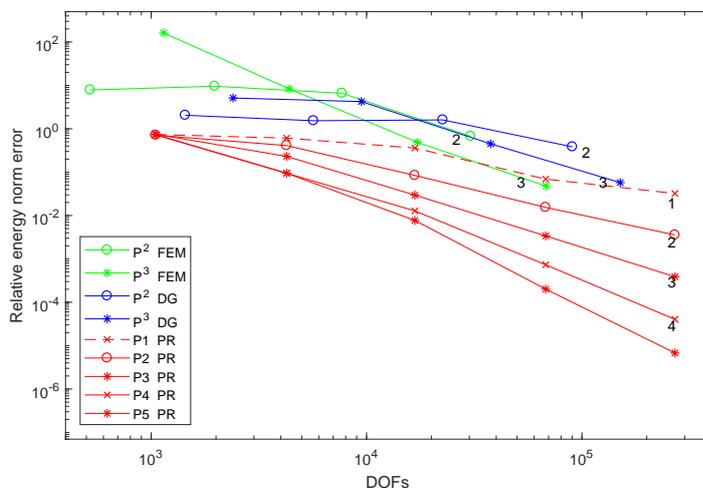


Figure 4: The relative energy norm error of Example 4.3.

5 Conclusions

This paper presents the discontinuous Galerkin method by patch reconstruction to solve the Helmholtz problems. We derive a priori estimates in the L^2 norm and energy norm for an IPDG method. For each fixed wave number k , the accuracy and efficiency of the method up to order five with high-order polynomials. Numerical experiments show that the convergence behavior of the proposed method is very similar to the theoretical results.

Acknowledgements

This work was supported by the National Key Research and Development Program of China (No. 2020YFA0714200), the Science and Technology Major Project of Hubei Province under Grant No. 2021AAA010, the National Science Foundation of China (Nos. 12125103 and 12071362), and by the Natural Science Foundation of Hubei Province (No. 2019CFA007). Numerical calculations are carried out on the supercomputing system of the Supercomputing Center of the Wuhan University.

References

- [1] G. B. ALVAREZ, A. LOULA, E. CARMO, AND F. A. ROCHINHA, *A discontinuous finite element formulation for helmholtz equation*, Comput. Methods Appl. Mech. Eng., 195(33/36) (2006), pp. 4018–4035.

- [2] P. F. ANTONIETTI, A. BUFFA, AND I. PERUGIA, *Discontinuous Galerkin approximation of the Laplace eigenproblem*, *Comput. Methods Appl. Mech. Eng.*, 195(25-28) (2006), pp. 3483–3503.
- [3] D. N. ARNOLD, *An interior penalty finite element method with discontinuous elements*, *SIAM J. Numer. Anal.*, 19(4) (1982), pp. 742–760.
- [4] D. N. ARNOLD, F. BREZZI, B. COCKBURN, AND L. D. MARINI, *Unified analysis of discontinuous Galerkin methods for elliptic problems*, *SIAM J. Numer. Anal.*, 39(5) (2002), pp. 1749–1779.
- [5] A. K. AZIZ AND R. B. KELLOGG, *A scattering problem for the helmholtz equation*, *Advances in Computer Methods for Partial Differential Equations–III*, 1979.
- [6] A. K. AZIZ AND A. WERSCHULZ, *On the numerical solutions of helmholtz's equation by the finite element method*, *SIAM J. Numer. Anal.*, 17(5) (2006), pp. 681–686.
- [7] I. I. BABUSKA, *Finite element solution of the helmholtz equation with high wave number part ii: The $h-p$ version of the fem*, *SIAM J. Numer. Anal.*, 34(1) (1997), pp. 315–358.
- [8] G. A. BAKER, *Finite element methods for elliptic equations using nonconforming elements*, *Math. Comput.*, 31 (1977), pp. 45–59.
- [9] A. BUFFA, P. HOUSTON, AND I. PERUGIA, *Discontinuous Galerkin computation of the Maxwell eigenvalues on simplicial meshes*, *J. Comput. Appl. Math.*, 204(2) (2007), pp. 317–333.
- [10] Z. CHEN AND H. WU, *An adaptive finite element method with perfectly matched absorbing layers for the wave scattering by periodic structures*, *SIAM J. Numer. Anal.*, (2003).
- [11] P. G. CIARLET, *The Finite Element Method for Elliptic Problems*, North-Holland Publishing Company, 1978.
- [12] B. COCKBURN, G. KARNIADAKIS, AND C.-W. SHU, *Discontinuous Galerkin Methods: Theory, Computation and Applications*, 2011.
- [13] B. COCKBURN AND C. W. SHU, *Local Discontinuous Galerkin Method for Time-Dependent Convection-Diffusionsystems*, 1997.
- [14] J. DOUGLAS AND T. DUPONT, *Interior penalty procedures for elliptic and parabolic galerkin methods*, *Lecture Notes Phys.*, 58 (1976), pp. 207–216.
- [15] J. DOUGLAS, J. E. SANTOS, D. SHEEN, AND L. BENNETHUM, *Frequency domain treatment of one-dimensional scalar waves*, *Math. Models Methods Appl. Sci.*, 03 (1993), pp. 171–194.
- [16] C. GEUZAIN AND J. F. REMACLE, *Gmsh: A 3-D finite element mesh generator with built-in pre- and post-processing facilities*, *Int. J. Numer. Methods Eng.*, 79(11) (2009), pp. 1309–1331.
- [17] F. HECHT, *New development in FreeFem++*, *J. Numer. Math.*, 20(3-4) (2012), pp. 251–265.
- [18] Y. HU, *Integral equation method for inverse scattering problem from the far-field data*, *Adv. Appl. Math. Mech.*, 13(6) (2021), pp. 1558–1574.
- [19] IHLENBURG, FRANK, BABUSKA, AND IVO, *Finite element solution of the helmholtz equation with high wave number part i1' the $h-p$ version of the fem*, *Siam J. Numer. Anal.*, (1997).
- [20] F. IHLENBURG, *Finite Element Analysis of Acoustic Scattering*, Springer Science & Business Media, 2006.
- [21] I. M. BABUSKA AND S. A. SAUTER, *Is the pollution effect of the fem avoidable for the helmholtz equation considering high wave numbers?*, *SIAM Rev.*, (2000).
- [22] K. ITO, Z. QIAO, AND J. TOIVANEN, *A domain decomposition solver for acoustic scattering by elastic objects in layered media*, *J. Comput. Phys.*, 227(19) (2008), pp. 8685–8698.
- [23] R. LI, P. B. MING, Z. Y. SUN, F. Y. YANG, AND Z. J. YANG, *A discontinuous Galerkin method by patch reconstruction for biharmonic problem*, *J. Comput. Math.*, 37(4) (2009), pp. 561–578.
- [24] R. LI, P. B. MING, Z. Y. SUN, AND Z. J. YANG, *An arbitrary-order discontinuous Galerkin method with one unknown per element*, *J. Sci. Comput.*, 80(1) (2019), pp 268–288.
- [25] R. LI, Z. SUN, AND F. YANG, *Solving eigenvalue problems in a discontinuous approximation space by patch reconstruction*, *SIAM J. Sci. Comput.*, 41(5) (2019), pp. A3381–A3400.

- [26] R. LI AND F. YANG, *A sequential least squares method for Poisson equation using a patch reconstructed space*, SIAM J. Numer. Anal., 58(1) (2020), pp. 353–374.
- [27] J. MELENK AND S. SAUTER, *Convergence analysis for finite element discretizations of the helmholtz equation with dirichlet-to-neumann boundary conditions*, Math. Comput., 79 (2010), pp. 1871–1914.
- [28] J. MELENK AND S. SAUTER, *Wavenumber explicit convergence analysis for galerkin discretizations of the helmholtz equation*, SIAM J. Numer. Anal., 49 (2011), pp. 1210–1243.
- [29] A. A. OBERAI AND P. M. PINSKY, *A multiscale finite element method for the helmholtz equation*, Comput. Methods Appl. Mech. Eng., 154(3–4) (1998), pp. 281–297.
- [30] B. RIVIÈRE, M. WHEELER, AND V. GIRAULT, *Improved energy estimates for interior penalty, constrained and discontinuous galerkin methods for elliptic problems. part I*, Comput. Geosci., 3 (1999), pp. 337–360.
- [31] C. W. SHU AND C. Y. LAM, *A phase-based interior penalty discontinuous galerkin method for the helmholtz equation with spatially varying wavenumber*, Comput. Methods Appl. Mech. Eng., 318 (2017), pp. 456–473.
- [32] M. WHEELER, *An elliptic collocation-finite element method with interior penalties*, SIAM J. Numer. Anal., 15 (1978), pp. 152–161.