# An *hp*-Adaptive Sampling Algorithm on Dispersion Relation Reconstruction for 2D Photonic Crystals

Yueqi Wang<sup>\*</sup> and Guanglian Li<sup>†</sup>

### Abstract

Computing the dispersion relation for two-dimensional photonic crystals is a notoriously challenging task: It involves solving parameterized Helmholtz eigenvalue problems with high-contrast coefficients. To resolve the challenge, we propose a novel hp-adaptive sampling scheme that can detect singular points via adaptive mesh refinement in the parameter domain, and meanwhile, allow for adaptively enriching the local polynomial spaces on the elements that do not contain singular points. In this way, we obtain an element-wise interpolation on an adaptive mesh. We derive an exponential convergence rate when the number of singular points is finite, and a first-order convergence rate otherwise. Numerical tests are provided to illustrate its performance.

**Key words:** photonic crystals, band function reconstruction, adaptive mesh, element-wise interpolation

Mathematics Subject Classifications: 65N12, 65N50, 65N15, 65N30

# 1 Introduction

This work is concerned with propagation of electromagnetic waves within photonic crystals (PhCs). PhCs are dielectric materials with a period size comparable to the wavelength [16]. The properties of waves in PhCs depend heavily on their frequencies, and for certain frequencies, they may be unable to propagate through PhCs, leading to the band gap phenomenon. This intriguing feature has led to the development of important applications, e.g., optical transistors, photonic fibers, and low-loss optical mirrors [36, 26, 33, 23]. We focus on two-dimensional (2D) PhCs, which are homogeneous along the z axis and have high-contrast dielectric columns or holes in dielectric materials within the x-y plane.

The propagation of electromagnetic waves is governed by the Maxwell system. For 2D infinite crystals with perfect periodicity, Bloch theorem states that the problem is reduced to 2D parameterized Helmholtz eigenvalue problems defined over the unit cell with periodic boundary conditions. The associated parameter is the so-called wave vector  $\mathbf{k}$ , which varies in the irreducible Brillouin zone (IBZ) [22]. The *n*th band function  $\omega_n$  (or the square root of the *n*th largest eigenvalue up to a constant) is a function of the wave vector for all  $n \in \mathbb{N}_+$ . The band gap is the distance between two adjacent band functions. Thus, computing the *n*th band function  $\omega_n(\mathbf{k})$  involves solving infinite many Helmholtz eigenvalue problems defined on the unit cell. To obtain the band gap, it is necessary for the permittivity to take different values in the inclusion and background of the unit cell, and moreover, the ratio between these values, known as contrast, should be large. However, solving Helmholtz eigenvalue problems with high-contrast and piecewise constant coefficients poses significant challenges. Numerically, one

<sup>\*</sup>Department of Mathematics, The University of Hong Kong, Pokfulam Road, Hong Kong. Email: u3007895@connect.hku.hk YW acknowledges support from the Research Grants Council (RGC) of Hong Kong via the Hong Kong PhD Fellowship Scheme (HKPFS).

<sup>&</sup>lt;sup>†</sup>Department of Mathematics, The University of Hong Kong, Pokfulam Road, Hong Kong. Email: lotusli@maths.hku.hk GL acknowledges the support from Newton International Fellowships Alumni following-on funding awarded by The Royal Society, Young Scientists fund (Project number: 12101520) by NSFC and Early Career Scheme (Project number: 27301921), RGC, Hong Kong.

possible approach is to limit the parameter  $\mathbf{k}$  to a coarse mesh of the IBZ or even its edges only. However, it is inadequate for identifying small band gaps, e.g., avoided crossings, which are important in certain application areas of photonic materials. For example, it is important to avoid the smallest band gaps when manufacturing large-pitch photonic crystal fibers [9], whereas leakage channel fibers exploit the smallest band gaps [7]. While directly solving the eigenvalue problem over a very fine parametric discretization is failure-safe, it is highly inefficient. Band structure diagram paths based on crystallography categorize wave vectors by their symmetry and provide preferred band paths for band function reconstruction [14]. These important wave vectors are located at the face and edge centers, and vertices of Brillouin zone. However, there is no guaranteed accuracy.

The main goal of this paper is to provide a numerical method with guaranteed fast convergence rate. More specifically, let  $\mathcal{B}$  be the parameter domain, i.e., IBZ. Let  $\omega(\mathbf{k}) \in C(\mathcal{B})$  be a band function and  $P_n(\mathcal{B})$  the space of global polynomial functions of degree at most n. The best approximation to  $\omega(\mathbf{k})$  in  $P_n(\mathcal{B})$  with respect to the uniform norm  $\|\cdot\|_{\infty,\mathcal{B}}$  on  $\mathcal{B}$  can be measured by

$$E_n(\omega) := \inf_{p \in P_n(\mathcal{B})} \|\omega - p\|_{\infty, \mathcal{B}}.$$

If  $\omega \notin C^1(\mathcal{B})$ , then  $E_n(\omega)$  cannot decay faster than first order with respect to n, i.e.,  $E_n(\omega) = \mathcal{O}(n^{-1})$ [31, Chapter 7]. Furthermore, this decay rate may not be improved in a subdomain where the band function is analytic [18]. The decay rate is only  $\mathcal{O}(N^{-1/2})$  in terms of the number N of sampling points N. In [34], we have developed a global polynomial interpolation method (GPI) to reconstruct the band function, focusing on selecting sampling points to achieve this slow convergence rate. Thus, there is an imperative need to develop more effective computational techniques for band structure reconstruction.

To improve the convergence of GPI, it is critical to exploit piecewise analyticity of band functions and to utilize local polynomial interpolation. This work focuses on adaptively selecting the wave vectors. Inspired by the h, p, and hp versions of FEM [2], one promising approach is to discretize the parameter domain  $\mathcal{B}$  with a triangular mesh and approximate the band functions locally on each element. In the hversion of local interpolation, the mesh sizes tend to zero uniformly and the degree of the polynomials is fixed at a small number, and the convergence rate is still  $\mathcal{O}(N^{-1/2})$ . In the p version of local interpolation, the mesh is fixed and the degree of polynomials tends to infinity, and it can also achieve a convergence rate  $\mathcal{O}(N^{-1/2})$ . Our approach combines these two strategies and consists of two steps. First, we generate an adaptive mesh of  $\mathcal{B}$  by Algorithm 1 such that elements containing singular points are refined at every iteration. Second, given a slope parameter  $\mu$ , we assign an element-wise polynomial degree according to the number of refinement of each element, and develop conforming element-wise interpolation in Algorithm 2. To analyze the convergence of the method, we first provide in Theorem 5.4 the condition on the slope parameter  $\mu$  such that the element-wise interpolation error in the element with singular points dominates. Then we derive an exponential convergence rate in Theorem 5.5 when the number of singular points is finite and a first-order convergence rate in Theorem 5.6 otherwise.

Overall, the proposed *hp*-adaptive sampling algorithm significantly improves the computational efficiency and accuracy of band function approximation in PhCs. The adaptive refinement of the parameter domain and element-wise interpolation can effectively capture singularities and improve approximation accuracy, while the mechanism for selecting polynomial degrees can further enhance the convergence rate. Additionally, the computation of band functions at each selected wave vector is independent and embarrassingly parallel. This work provides a novel approach for photonic crystal band function computation, which holds potential for designing and optimizing photonic crystal structures.

The rest of the paper is organized as follows. In Section 2, we recap the derivation of band functions for 2D PhCs and regularity. We propose in Section 3 an hp-adaptive sampling algorithm, which adaptively refines elements containing singular points and allocates proper polynomial degrees for local Lagrange polynomial interpolation. We define the element-wise interpolation over the adaptive mesh and with a given element-wise degree in Section 4. In Section 5, we present a convergence analysis of the algorithm. We present extensive numerical experiments in Section 6 to complement the analysis. Finally, we give concluding remarks in Section 7. Throughout, the notation  $A \leq B$  stands for that A/B is bounded above by a constant independent of the number of sampling points N and the mesh size h.

# 2 Problem formulation

### 2.1 2D Band function computation

First we briefly derive the parameterized Helmholtz eigenvalue problem from the Maxwell equations, which was formulated in [34] to estimate band functions. For further details, we refer readers to [3, 34]. In the SI convention, the time harmonic Maxwell equations for linear, non-dispersive, and nonmagnetic media, with free charges and free currents, consist of a system of four equations [15],

$$\nabla \times \mathbf{E}(\mathbf{x}) - i\omega\mu_0 \mathbf{H}(\mathbf{x}) = 0, \qquad (2.1a)$$

$$\nabla \times \mathbf{H}(\mathbf{x}) + i\omega\epsilon_0\epsilon(\mathbf{x})\mathbf{E}(\mathbf{x}) = 0, \qquad (2.1b)$$

$$\nabla \cdot (\epsilon(\mathbf{x})\mathbf{E}(\mathbf{x})) = 0, \qquad (2.1c)$$

$$\nabla \cdot \mathbf{H}(\mathbf{x}) = 0. \tag{2.1d}$$

Here  $\mathbf{x} \in \mathbb{R}^3$ ,  $\mathbf{E}$  is the electric field,  $\mathbf{H}$  the magnetic field and  $\mathbf{D}$  the electric displacement field. The scalar  $\omega \geq 0$  is the frequency of the electromagnetic wave,  $\mu_0$  the vacuum permeability,  $\epsilon_0$  the vacuum permittivity, and  $\epsilon \in L^{\infty}(\mathbb{R}^3; \mathbb{R}^+)$  the relative permittivity.

Applying the curl operator to (2.1a) and using (2.1b), we obtain

$$\nabla \times (\nabla \times \mathbf{E}(\mathbf{x})) - (\omega c^{-1})^2 \epsilon(\mathbf{x}) \mathbf{E}(\mathbf{x}) = 0, \qquad (2.2)$$

with  $\epsilon_0 \mu_0 = c^{-2}$ , where c is the speed of light. Similarly, applying the curl operator to (2.1b) and using (2.1a), we obtain

$$\nabla \times \left( \epsilon(\mathbf{x})^{-1} \nabla \times \mathbf{H}(\mathbf{x}) \right) - \left( \omega c^{-1} \right)^2 \mathbf{H}(\mathbf{x}) = 0.$$
(2.3)



(a) PhCs with square lattice

(b) PhCs with hexagonal lattice

Figure 1: A schematic illustration of 2D PhCs. The materials have dielectric columns homogeneous along the z direction and periodic along x and y directions.

In this work, we focus on 2D PhCs. In Figure 1, we show 2D PhCs with a periodic arrangement of dielectric columns within a dielectric material. In practice, both have finite extensions in the z direction and finite periodicities in the x-y plane. Nevertheless, it is commonly assumed that the

material extends infinitely in the plane perpendicular to the columns. Thus, in the setting of 2D PhCs in Figure 1, the relative permittivity  $\epsilon(\mathbf{x})$  is assumed to be independent of the z direction. Then we can split the electromagnetic fields  $\mathbf{E} = (E_1, E_2, E_3)$  and  $\mathbf{H} = (H_1, H_2, H_3)$  in (2.2) and (2.3) into the transverse electric (TE) mode with  $H_1 = H_2 = E_3 = 0$  and transverse magnetic (TM) mode with  $E_1 = E_2 = H_3 = 0$ . Each mode is a scalar eigenvalue problem,

$$-\nabla \cdot (\epsilon(\mathbf{x})^{-1} \nabla H(\mathbf{x})) - (\omega c^{-1})^2 H(\mathbf{x}) = 0, \quad \mathbf{x} \in \mathbb{R}^2, \qquad (\mathbf{TE \ mode}), \qquad (2.4)$$
$$-\Delta E(\mathbf{x}) - (\omega c^{-1})^2 \epsilon(\mathbf{x}) E(\mathbf{x}) = 0, \quad \mathbf{x} \in \mathbb{R}^2, \qquad (\mathbf{TM \ mode}). \qquad (2.5)$$



Figure 2: Square lattice:  $\Omega$  (left) and the corresponding Brillouin zone (right). The IBZ  $\mathcal{B}$  is the shaded triangle with vertices  $\Gamma = (0,0), X = \frac{1}{a}(\pi,0)$  and  $M = \frac{1}{a}(\pi,\pi)$ .



Figure 3: Hexagonal lattice:  $\Omega$  (left) and the corresponding Brillouin zone (right). The IBZ  $\mathcal{B}$  is the shaded triangle with vertices  $\Gamma = (0,0), K = \frac{1}{a}(\frac{4}{3}\pi,0)$  and  $M = \frac{1}{a}(\pi,\frac{\sqrt{3}}{3}\pi)$ .

2D PhCs possess a discrete translational symmetry in the x-y plane (cf. Figure 1) [16], i.e., the relative permittivity  $\epsilon(\mathbf{x})$  satisfies

$$\epsilon(\mathbf{x} + c_1\mathbf{a}_1 + c_2\mathbf{a}_2) = \epsilon(\mathbf{x}), \quad \forall \mathbf{x} \in \mathbb{R}^2 \text{ and } c_1, c_2 \in \mathbb{Z}.$$

The primitive lattice vectors, denoted by  $\mathbf{a}_1$  and  $\mathbf{a}_2$ , are the shortest possible vectors that fulfill this condition and they span the fundamental periodicity domain  $\Omega$ , also known as unit cell, cf. Figures 2 and 3. In Figure 2, primitive lattice vectors of the square lattice are  $\mathbf{a}_i = ae_i$  for i = 1, 2, and in Figure 3, that of the hexagonal lattice  $\mathbf{a}_1 = \frac{a}{4}(e_1 + \sqrt{3}e_2)$  and  $\mathbf{a}_2 = \frac{a}{4}(e_1 - \sqrt{3}e_2)$ . Here,  $(e_i)_{i=1,2}$  is the canonical basis in  $\mathbb{R}^2$  and  $a \in \mathbb{R}^+$  lattice constant. The reciprocal lattice vectors are defined by

$$\mathbf{b}_i \cdot \mathbf{a}_j = 2\pi \delta_{ij}, \quad \text{for } i, j = 1, 2, \tag{2.6}$$

which generate the so-called reciprocal lattice. The elementary cell of the reciprocal lattice is the (first) Brillouin zone  $\mathcal{B}_F$ , i.e., the region closer to a certain lattice point than to any other lattice points in the reciprocal lattice.

Bloch's theorem [20] states that in periodic crystals, wave functions take the form of a plane wave modulated by a periodic function. Thus they can be written as  $\Psi(\mathbf{x}) = e^{i\mathbf{k}\cdot\mathbf{x}}u(\mathbf{x})$ , where  $\Psi$  is the wave function,  $u(\mathbf{x})$  is periodic sharing the periodicity of the crystal lattice and  $\mathbf{k}$  is the wave vector varying in the Brillouin zone  $\mathcal{B}_F$ . The periodic condition of  $u(\mathbf{x})$  implies that each wave function can be determined by its values within the unit cell  $\Omega$ . Thus, the solutions to (2.4) and (2.5) can be expressed as  $H(\mathbf{x}) = e^{i\mathbf{k}\cdot\mathbf{x}}u_1(\mathbf{x})$  and  $E(\mathbf{x}) = e^{i\mathbf{k}\cdot\mathbf{x}}u_2(\mathbf{x})$  for some periodic functions  $u_1(\mathbf{x})$  and  $u_2(\mathbf{x})$  in the unit cell  $\Omega$ , and the parameterized Helmholtz eigenvalue problems (2.4) and (2.5) reduce to

$$-(\nabla + i\mathbf{k}) \cdot \left(\epsilon(\mathbf{x})^{-1}(\nabla + i\mathbf{k})u_1(\mathbf{x})\right) - \left(\omega c^{-1}\right)^2 u_1(\mathbf{x}) = 0, \quad \mathbf{x} \in \Omega \qquad \text{(TE mode)}, \tag{2.7a}$$

$$-(\nabla + i\mathbf{k}) \cdot ((\nabla + i\mathbf{k})u_2(\mathbf{x})) - (\omega c^{-1})^2 \epsilon(\mathbf{x})u_2(\mathbf{x}) = 0, \quad \mathbf{x} \in \Omega \qquad (\mathbf{TM mode}),$$
(2.7b)

where  $\mathbf{x} \in \Omega \subset \mathbb{R}^2$ ,  $\mathbf{k}$  varies in the Brillouin zone (BZ), and  $u_i(\mathbf{x})$  satisfies the periodic boundary conditions  $u_i(\mathbf{x}) = u_i(\mathbf{x} + \mathbf{a_j})$  with  $\mathbf{a_j}$  being the primitive lattice vector for i, j = 1, 2. If the materials in the unit cell have additional symmetry, e.g. mirror symmetry, we can further restrict  $\mathbf{k}$  to the triangular irreducible Brillouin zone (IBZ), denoted by  $\mathcal{B}$ . The BZs and IBZs of the square lattice and hexagonal lattice are shown in Figures 2 and 3.

In sum, we can formulate both parameterized Helmholtz problems (2.7a) and (2.7b) by

$$-(\nabla + i\mathbf{k}) \cdot \alpha(\mathbf{x})(\nabla + i\mathbf{k})u(\mathbf{x}) - \lambda\beta(\mathbf{x})u(\mathbf{x}) = 0, \quad \mathbf{x} \in \Omega,$$
(2.8)

with  $\Omega \subset \mathbb{R}^2$ ,  $\mathbf{k} \in \mathcal{B}$ , and  $\lambda = (\omega c^{-1})^2$ . In the TE mode, u describes the magnetic field H in z-direction and the coefficients  $\alpha(\mathbf{x})$  and  $\beta(\mathbf{x})$  are  $\alpha(\mathbf{x}) := \epsilon(\mathbf{x})^{-1}$  and  $\beta(\mathbf{x}) := 1$ . Similarly, in the TM mode, udescribes the electric field E in the z-direction and the coefficients  $\alpha(\mathbf{x})$  and  $\beta(\mathbf{x})$  are  $\alpha(\mathbf{x}) := 1$  and  $\beta(\mathbf{x}) := \epsilon(\mathbf{x})$ . The variational formulation of (2.8) reads: find a non-trivial eigenpair  $(\lambda, u) \in (\mathbb{R}, H^1_{\pi}(\Omega))$ for  $\mathbf{k} \in \mathcal{B}$  such that

$$\begin{cases} \int_{\Omega} \alpha (\nabla + i\mathbf{k}) u \cdot (\nabla - i\mathbf{k}) \bar{v} - \lambda \beta u \bar{v} \mathrm{d}x = 0, \quad \forall v \in H^{1}_{\pi}(\Omega) \\ \|u\|_{L^{2}_{\beta}(\Omega)} = 1. \end{cases}$$
(2.9)

Here, we define  $L^2_{\beta}(\Omega)$  as the space of weighted square integrable functions equipped with the norm  $\|f\|_{L^2_{\beta}(\Omega)} := (\int_{\Omega} \beta(\mathbf{x}) |f(\mathbf{x})|^2 d\mathbf{x})^{\frac{1}{2}}$ . Let  $H^1(\Omega) \subset L^2_{\beta}(\Omega)$  be with square integrable gradient, equipped with the standard  $H^1(\Omega)$ -norm and  $H^1_{\pi}(\Omega) \subset H^1(\Omega)$  is composed of functions with periodic boundary conditions. Using the sesquilinear forms

$$a(u,v) := \int_{\Omega} \alpha(\nabla + i\mathbf{k})u \cdot (\nabla - i\mathbf{k})\bar{v} \,\mathrm{d}x \quad \text{and} \quad b(u,v) := \int_{\Omega} \beta u\bar{v} \,\mathrm{d}x,$$

problem (2.9) is equivalent to finding a non-trivial eigenpair  $(\lambda, u) \in (\mathbb{R}, H^1_{\pi}(\Omega))$  for  $\mathbf{k} \in \mathcal{B}$  such that

$$\begin{cases} a(u,v) = \lambda b(u,v), \text{ for all } v \in H^1_{\pi}(\Omega) \\ \|u\|_{L^2_{\beta}(\Omega)} = 1. \end{cases}$$

$$(2.10)$$

### 2.2 Regularity of band functions in the parameter domain

The regularity of band functions plays a crucial role in the convergence analysis. We have the following result taken from [34]. Here,  $Lip(\mathcal{B})$  denotes the space of Lipschitz continuous functions in the IBZ  $\mathcal{B}$  and  $\mathring{A}(\mathcal{B})$  the space of piecewise analytic functions with singular point set of zero Lebesgue measure.

**Theorem 2.1.** For 2D periodic PhCs,  $\omega_n(\mathbf{k}) \in Lip(\mathcal{B}) \cap \mathring{A}(\mathcal{B})$  for all  $n \in \mathbb{N}_+$ : each band function  $\omega_n(\mathbf{k})$  is analytic in  $\mathcal{B} \setminus \mathbf{S}$ , where  $\mathbf{S} \subset \mathcal{B}$  consists of the origin and points of degeneracy (i.e., the set of points at which certain band functions intersect, with zero Lebesgue measure).

**Remark 2.1.** Note that the singularity at  $\mathbf{k} = \mathbf{0}$  differs from that of others. The latter are known as branch points while the former arises from the composition  $g \circ \lambda$  with  $g(x) := \sqrt{x}$ . The derivative of the first band function  $\omega_1(\mathbf{k}) = c \cdot \sqrt{\lambda_1(\mathbf{k})}$ , which takes the value 0 at the origin, fail to be defined at  $\mathbf{k} = \mathbf{0}$ . To remove the effect of this singularity, we approximate  $\lambda_j(\mathbf{k})$  for  $j = 1, \dots, L-1$  with  $1 < L \in \mathbb{N}$ , instead.

# 3 hp-adaptive sampling algorithm in the parameter domain

Due to the presence of singularities, the convergence of numerical methods based on global or fixed-order local polynomial interpolation on uniform meshes may be slow. To optimally deal with singular points of band functions in PhCs, we propose a novel efficient numerical algorithm using an hp-adaptive interpolation method. It is inspired by the celebrated hp FEM [2, 1], and employs nonuniformly distributed polynomial degrees over graded, nonuniform meshes to discretize the problem. We shall design an hp interpolation strategy based on an element-wise quantity in order to approximate the first L-1 band functions for  $1 < L \in \mathbb{N}$ .

### 3.1 Adaptive mesh generation

We first recall the concept of a regular mesh without hanging nodes [27]. Let  $D \subset \mathbb{R}^2$  be a polygon with straight sides. A mesh  $\mathcal{T} = \{T_i\}_{i=1}^M$  is a partition of D into open disjoint triangles  $T_i$  with  $\overline{D} = \bigcup_{i=1}^M \overline{T_i}$ . For each  $T \in \mathcal{T}$ , we denote by  $h_T$  its diameter and  $\rho_T$  the diameter of the largest inscribed ball.

**Definition 3.1.** A mesh  $\mathcal{T} = \{T_i\}_{i=1}^M$  is said to be a regular mesh without hanging nodes if the following conditions hold. (i) For  $i \neq j$ ,  $\overline{T_i} \cap \overline{T_j}$  is either empty or it consists of a vertex or an entire edge of  $T_i$  and (ii) There exists a constant C > 0 such that  $h_T \leq C\rho_T$ .

We shall develop a strategy to generate a family of nested regular triangular meshes  $\mathcal{T}_n$ ,  $n = 1, 2, \cdots$ , based on an element-wise quantity (3.1) and a criterion for refinement (3.2). We need several notations for the mesh  $\mathcal{T}_n$ . The mesh size of  $\mathcal{T}_n$  is denoted by  $h_n := \max_{T \in \mathcal{T}_n} h_T$ . The collection of all faces and vertices over  $\mathcal{T}_n$  is denoted by  $\mathcal{F}_n$  and  $\mathcal{V}_n$ , respectively. For any  $T \in \mathcal{T}_n$ ,  $\mathcal{F}_T := \partial T \cap \mathcal{F}_n$  denotes its faces, and  $F_T^j$  for j = 1, 2, 3 the *j*th face,  $\mathcal{V}_T := \partial T \cap \mathcal{V}_n$  its vertices, and  $V_T^j$  for j = 1, 2, 3 the *j*th vertex. For any  $F \in \mathcal{F}_n$ ,  $\mathcal{T}_F := \bigcup \{T \in \mathcal{T}_n : F \in \mathcal{F}_T\}$  denotes the collection of elements having F as an edge. Finally,  $P_m(T)$  is the space of polynomial functions of degrees at most  $m \in \mathbb{N}_+$  on element T, and  $P_m(F)$  the space of polynomial functions of degrees at most  $m \in \mathbb{N}_+$  on each face  $F \in \mathcal{F}_n$ .

Since each band function is a piecewise analytic function of the wave vector, with singular points at branch points and the origin, the mesh design strategy should identify elements in which adjacent bands are in close proximity, and then refine them suitably to minimize the impact of singularities on the approximation. To this end, we employ an element-wise indicator that is based on the distance between two adjacent band functions

$$\eta(T) := \min_{1 \le q \le L-1} \|\omega_q - \omega_{q+1}\|_{\infty, \mathcal{V}_T} \text{ for } T \in \mathcal{T}_n.$$

$$(3.1)$$

That is, it uses the adjacent values on the vertices of each triangle element to ascertain the regularity of the band functions within each element. Furthermore, we introduce an element-wise tolerance

$$\operatorname{tol}_1(T) = \kappa h_T \max_{1 \le q \le L} \|\nabla \omega_q\|_{\infty, T \setminus \mathbf{S}},\tag{3.2}$$

where **S** is the set of singular points and  $\kappa \geq 2\sqrt{2}$  is a constant. The band function  $\omega_q(\mathbf{k})$  and the group velocity  $\nabla \omega_q(\mathbf{k})$  used in  $\eta(T)$  and  $\operatorname{tol}_1(T)$  are computed using the conforming Galerkin FEM, cf. (2.10) and (3.9).

Give a mesh  $\mathcal{T}_n$ , the collection of marked elements is

$$\mathcal{M}_n := \cup \left\{ T \in \mathcal{T}_n : \eta(T) \le \operatorname{tol}_1(T) \right\}.$$
(3.3)

We employ the iterated newest vertex bisection (NVB) [11] to refine the elements in  $\mathcal{M}_n$ . It involves repeatedly bisecting the longest edge of the marked element and inserting a new vertex at the midpoint until the desired level of refinement is achieved. Then the resulting mesh is sufficiently refined in areas where the singularities may impact the approximation accuracy.

Iterated NVB preserves the mesh regularity during refinement [32]. However, to achieve a set of regular meshes without any hanging nodes, it is often necessary to refine certain elements beyond the initial refinement. Thus, the number of elements to be refined, denoted as  $\#\mathcal{T}_{n+1} - \#\mathcal{T}_n$ , is typically larger than the number of marked elements  $\#\mathcal{M}_n$ . Nonetheless, the cumulative number of elements added in this manner does not inflate the total number of marked elements [4, Theorem 2.4].

**Theorem 3.1.** Let  $\mathcal{T}_1$  be a regular mesh in the sense of Definition 3.1. For  $n \geq 1$  let  $\{\mathcal{T}_n\}_{n\geq 1}$  be any sequence of the refinement of  $\mathcal{T}_1$ , where  $\mathcal{T}_{n+1}$  is a regular mesh generated from  $\mathcal{T}_n$  by the iterated NVB with a subset of the marked elements  $\mathcal{M}_n \subset \mathcal{T}_n$ . Then there exists a constant C solely depending on  $\mathcal{T}_1$ such that

$$#\mathcal{T}_n - #\mathcal{T}_1 \le C \sum_{i=1}^{n-1} #\mathcal{M}_i, \quad \forall n \ge 1.$$
(3.4)

We give the proposed mesh refinement procedure in Algorithm 1. The parameters  $tol_2$  and  $n_{max}$ denote the smallest allowed element size and the maximum number of iterations, respectively.

Algorithm 1 Adaptive mesh refinement

**Require:** initial mesh  $\mathcal{T}_1$ , tolerance tol<sub>2</sub>, and maximum loop number  $n_{\text{max}}$ Ensure:  $\mathcal{T}_n$  $n \leftarrow 1$ while  $n \leq n_{\max} \operatorname{do}$ Set  $\mathcal{M}_n = \emptyset$ for  $T \in \mathcal{T}_n$  do

if  $\eta(T) \leq \operatorname{tol}_1(T)$  and  $h_T \geq \operatorname{tol}_2$  then  $\mathcal{M}_n = \mathcal{M}_n \cup \{T\}$ 

$$\mathcal{M}_n = \mathcal{M}_n \cup \{$$

end if

### end for

Refine the collection of marked elements  $\mathcal{M}_n$  by iterated NVB and adjust the mesh until there are no hanging nodes.

```
n \leftarrow n+1
end while
```

#### Reliability of Algorithm 1 3.2

For the mesh  $\mathcal{T}_n$  at the *n*th iteration, let

$$\mathcal{T}_n^R := \cup \{T \in \mathcal{T}_n : T \cap \mathbf{S} = \emptyset\} \text{ and } \mathcal{T}_n^S := \mathcal{T}_n ackslash \mathcal{T}_n^R$$

be a partition of the mesh  $\mathcal{T}_n$  based on the singular point set **S**. The next result guarantees that Algorithm 1 refines all the elements in  $\mathcal{T}_n^S$  at the next iteration.

**Theorem 3.2**  $(\mathcal{T}_n^S \subset \mathcal{M}_n)$ . Algorithm 1 with the element-wise quantity (3.1) and the tolerance (3.2) quarantees that all elements containing branch points are marked for refinement.

*Proof.* The proof relies on piecewise analyticity and Lipschitz continuity of band functions. Let  $\mathcal{T}_n$  be a regular mesh generated by Algorithm 1 for  $n = 1, \dots$ . For any band function  $\omega(\mathbf{k})$  and any element  $T \in \mathcal{T}_n^R$ , by the mean value theorem, we obtain

$$\max_{\mathbf{k}\in T, \widehat{\mathbf{k}}\in\mathcal{V}_T} |\omega(\mathbf{k}) - \omega(\widehat{\mathbf{k}})| \le \sqrt{2}h_T \cdot \|\nabla\omega\|_{\infty,T}.$$
(3.5)

If  $T \in \mathcal{T}_n^S$ , then the first derivative of band functions may be discontinuous on singular points. The singular points on the line segment connecting  $\mathbf{k}$  and  $\hat{\mathbf{k}}$  divide the segment into several small intervals. By applying the mean value theorem on each subinterval, we obtain

$$\max_{\mathbf{k}\in T, \widehat{\mathbf{k}}\in\mathcal{V}_T} |\omega(\mathbf{k}) - \omega(\widehat{\mathbf{k}})| \le \sqrt{2h_T} \cdot \|\nabla \omega\|_{\infty, T\setminus \mathbf{S}}.$$
(3.6)

Note that  $T \notin \mathcal{M}_n$  implies any two adjacent band functions  $\omega_p(\mathbf{k})$  and  $\omega_{p+1}(\mathbf{k})$  for  $1 \leq p < L-1$  satisfy

$$\min_{\widehat{\mathbf{k}}\in\mathcal{V}_T} |\omega_p(\widehat{\mathbf{k}}) - \omega_{p+1}(\widehat{\mathbf{k}})| > \operatorname{tol}_1(T) = \kappa h_T \cdot \max_{1 \le q \le L} \|\nabla \omega_q\|_{\infty, T \setminus \mathbf{S}}$$

By the triangle inequality, this further implies for any  $\hat{\mathbf{k}} \in \mathcal{V}_T$  and  $\mathbf{k} \in T$ ,

$$\begin{aligned} \operatorname{tol}_1(T) &< |\omega_p(\mathbf{k}) - \omega_{p+1}(\mathbf{k})| \\ &\leq |\omega_p(\widehat{\mathbf{k}}) - \omega_p(\mathbf{k})| + |\omega_p(\mathbf{k}) - \omega_{p+1}(\mathbf{k})| + |\omega_{p+1}(\mathbf{k}) - \omega_{p+1}(\widehat{\mathbf{k}})|. \end{aligned}$$

This, together with (3.5) and (3.6), leads to

$$\min_{\mathbf{k}\in T} |\omega_p(\mathbf{k}) - \omega_{p+1}(\mathbf{k})| > (\kappa - 2\sqrt{2})h_T \cdot \max_{1 \le q \le L} \|\nabla \omega_q\|_{\infty, T \setminus \mathbf{S}} \ge 0.$$
(3.7)

That is, there is no branch points in T and  $T \notin \mathcal{T}_n^S$ . Hence, we have proved that  $T \notin \mathcal{T}_n^S$  if  $T \notin \mathcal{M}_n$ , or equivalently  $\mathcal{T}_n^S \subset \mathcal{M}_n$ .

Computing the element-wise tolerance  $\operatorname{tol}_1(T)$  (3.2) is impractical since it involves estimating  $\|\nabla \omega(\mathbf{k})\|_{\infty}$  for all  $\mathbf{k} \in T \setminus \mathbf{S}$ . However, numerical experiments have shown that the slope of the first several band functions (a.k.a. group velocity) does not exhibit high oscillations. Hermann et al [13] showed that in the first three bands, higher bands have low values of the group velocity and all of them are curves without significant fluctuations. Therefore, as a substitute for  $\operatorname{tol}_1(T)$ , we use a feasible tolerance

$$\widehat{\text{tol}}_1(T) := \kappa h_T \max_{1 \le q \le L} \|\nabla \omega_q\|_{\infty, \mathcal{V}_T \setminus \mathbf{S}}.$$
(3.8)

Computing  $\widehat{tol}_1(T)$  only requires the band function and its partial derivative over the vertices. Moreover, the group velocity can be derived directly from the following expression [34]

$$\frac{\partial\omega}{\partial k_i} = \frac{2k_i m_\alpha(u, u) + m_{\alpha i}(u, u)}{2\omega},\tag{3.9}$$

where the bilinear forms  $m_{\alpha i}(\cdot, \cdot)$  and  $m_{\alpha}(\cdot, \cdot)$  are defined as

$$m_{\alpha i}(u,v) = \int_{\Omega} i\alpha \left( u \frac{\partial \bar{v}}{\partial x_i} - \bar{v} \frac{\partial u}{\partial x_i} \right) \, \mathrm{d}\mathbf{x}, \quad i = 1, 2 \quad \text{and} \quad m_{\alpha}(u,v) = \int_{\Omega} \alpha u \bar{v} \, \mathrm{d}\mathbf{x}.$$

### 3.3 Conforming finite element space

Let  $\{\mathcal{T}_n\}_{n=1}^{\infty}$  be a sequence of nested regular triangular meshes generated by Algorithm 1. We shall construct an element-wise polynomial space for the mesh  $\mathcal{T}_n$ . We first define the layer of an element  $T \in \mathcal{T}_n$  given by Algorithm 1.

**Definition 3.2.** The layer  $\ell_T$  of an element  $T \in \mathcal{T}_n$  is  $\ell_T := n - r$ , where  $0 \le r \le n - 1$  is the number of refinements that have been performed on this element in Algorithm 1.

Note that the layer  $\ell_T$  of an element  $T \in \mathcal{T}_n$  implies the local regularity of the eigenfunction  $\lambda_j(\mathbf{k})$  for  $j = 1, \dots, L-1$ . Since the singular points fall into the marked element  $\mathcal{M}_n$ , we employ polynomials of lower degrees for the local interpolation in  $\mathcal{M}_n$ . Note also that the first band function vanishes at the origin, and is non-negative, sharing the symmetry of the lattice. This leads to the expression of  $(\omega_1(\mathbf{k}))^2$  similar to an elliptic cone. More precisely, since  $(\omega_1(\mathbf{k}))^2$  is analytic at  $\mathbf{k} = \mathbf{0}$  if  $\omega_1(\mathbf{0}) \neq \omega_2(\mathbf{0})$ , its Taylor expansion at the origin is

$$(\omega_1(\mathbf{k}))^2 = (\omega_1(k_1, k_2))^2 = \sum_{i=1}^{\infty} \sum_{j=0}^{i} a_{ij} (k_1)^{2j} (k_2)^{2i-2j},$$

with some  $a_{ij} \in \mathbb{R}$  for  $i \in \mathbb{N}$  and  $j = 0, \dots, i$ . Hence  $\lambda_1(\mathbf{k}) = (\omega_1(\mathbf{k})c^{-1})^2$  is quadratic near the origin, and we use polynomials of minimal degree 2 to interpolate eigenvalues over the marked elements  $\mathcal{M}_n$ . These observations motivate the following definition of the element-wise space of polynomials. The notation  $\lceil \cdot \rceil$  denotes the ceiling function.

**Definition 3.3.** For any element  $T \in \mathcal{T}_n$ , we take  $P_{n_T}(T)$  as the local interpolation space, with  $n_T := \max\{2, \lceil \mu \ell_T \rceil\}$  if  $T \notin \mathcal{M}_n$ , and  $n_T := 2$  otherwise, where the parameter  $\mu > 0$  is determined in (5.9) in Section 5 and  $\mathcal{M}_n$  is the collection of marked elements (3.3).

In this way, each element T is assigned a total degree  $n_T$ . However, neighboring elements may have different degrees on the common face. Suitable constraints are required to ensure the inter-element continuity, for which it is essential to associate with each face in the mesh a face degree.

**Definition 3.4** (Facewise space of polynomial functions). Let  $P_{m_F}(F)$  be the local interpolation space over F for  $F \in \mathcal{F}_n$  with  $m_F := \min_{T \in \mathcal{T}_F} \{n_T\}$ . This implies  $\max_{F \in \mathcal{F}_T} \{m_F\} \le n_T$ , for all  $T \in \mathcal{T}_n$ .

Finally, on a given mesh  $\mathcal{T}_n$  with element-wise and facewise degrees  $\{n_T\}_{T \in \mathcal{T}_n}$  and  $\{m_F\}_{F \in \mathcal{F}_n}$ , we define the conforming FEM space

$$V_n := \{ v \in C(\mathcal{B}) : v |_T \in P_{n_T}(T) \text{ and } v |_F \in P_{m_F}(F), \forall T \in \mathcal{T}_n \text{ and } F \in \mathcal{F}_n \},$$
(3.10)

which is a space of continuous piecewise polynomial functions on  $\mathcal{B}$ . Let N be its dimension. We aim at approximating the eigenvalues  $\lambda$ , for  $j = 1, \dots, L-1$ , over  $V_n$  by element-wise Lagrange interpolation.

### 4 Element-wise interpolation over the adaptive mesh

We develop Lagrange interpolation over each element  $T \in \mathcal{T}_n$ . For each element, we first map its closure  $\overline{T}$  into the reference equilateral triangle  $\hat{T} = \{\mathbf{z} := (x, y) : 0 \le y \le (1+x)\sqrt{3}, -1 \le x \le 0 \text{ or } 0 \le y \le (1-x)\sqrt{3}, 0 \le x \le 1\}$  and then derive Lagrange interpolation on  $\hat{T}$ . The vertices of  $\hat{T}$  are  $\hat{\mathbf{z}}_1 = (-1, 0), \hat{\mathbf{z}}_2 = (1, 0)$  and  $\hat{\mathbf{z}}_3 = (0, \sqrt{3})$ , with the face opposite to  $\hat{\mathbf{z}}_i$  denoted by  $\hat{F}^i$ . For any  $\{m, m_1, m_2, m_3\} \in \mathbb{N}^4_+$  with  $m \ge m_i$  for i = 1, 2, 3, (3.10) implies that the local space of polynomial functions over  $\hat{T}$  is

$$\hat{P} := \{ v \in P_m(\hat{T}) : v|_{\hat{F}^i} \in P_{m_i}(\hat{F}^i) \}.$$
(4.1)

Note that  $\hat{P} \subset P_m(\hat{T})$  and  $\hat{P} \subsetneq P_m(\hat{T})$  if  $m_i < m$  for some i = 1, 2, 3. Note also that the dimension of  $P_m(\hat{T})$  is  $\dim(P_m(\hat{T})) = (m+1)(m+2)/2$ , and the dimension of  $\hat{P}$  is  $\dim(\hat{P}) = \dim(P_m(\hat{T})) - \sum_{i=1}^3 (m-m_i) = 3 + ((m_1-1) + (m_2-1) + (m_3-1)) + (m-2)(m-1)/2$ .

# 4.1 Basis functions in the local polynomial space $\hat{P}$

We construct a set of basis functions for the local polynomial space  $\hat{P}$ , following [27]. We divide the basis functions into external and internal shape functions according to whether they vanish over the boundary  $\partial \hat{T}$ , and then divide the external shape functions further into nodal and side ones. The definition of nodal, side and internal shape functions is standard in the community of *p*-version and *hp*-adaptive FEMs, and it enables defining conforming FEM spaces with nonuniform degrees of polynomials over the mesh. We focus on developing an element-wise interpolation, and these shape functions play a role when formulating a special local polynomial space  $\hat{P}$ , which facilitates the definition of the element-wise interpolation in Section 4.2.

First, there are three nodal shape functions on  $\hat{T}$ :

$$N_1^1 = (1 - x - y/\sqrt{3})/2, \quad N_2^1 = (1 + x - y/\sqrt{3})/2 \text{ and } N_3^1 = y/\sqrt{3}.$$
 (4.2)

Note that for i = 1, 2, 3, we have  $N_i^1(\hat{\mathbf{z}}_i) = 1$  and  $N_i^1(\mathbf{z}) = 0$  for  $\mathbf{z} \in \hat{F}^i$ . Second, there are  $m_i - 1$  side shape functions over  $\hat{F}^i$ , i = 1, 2, 3, defined by

$$N_i^{m_i,j} = N_{i+1}^1 N_{i+2}^1 \Phi_{j-2} (N_{i+2}^1 - N_{i+1}^1),$$
(4.3)

with  $\Phi_{j-2}(x)$  being a polynomial of degree j-2 for  $j=2, \cdots, m_i$ , with the convention  $N_{j+3}^1 = N_j^1$  and  $\hat{\mathbf{z}}_{j+3} = \hat{\mathbf{z}}_j$  for j = 1, 2, 3. Note that  $\operatorname{supp}(N_i^{m_i,j})$  is the interior of  $\hat{F}^i$  and  $\hat{T}$ . Then the set of external shape functions is given by

$$\mathcal{E}_m^{(m_1,m_2,m_3)}(\hat{T}) := \left\{ N_j^1, N_j^{m_j,i} : 1 \le j \le 3 \text{ and } 2 \le i \le m_j \right\}.$$

By construction, the number of external basis functions is  $3 + ((m_1 - 1) + (m_2 - 1) + (m_3 - 1))$ .

Next, we construct a set of internal basis functions that vanish over  $\partial \hat{T}$ . Let  $b_{\hat{T}} := N_1^1 N_2^1 N_3^1$  be the basic bubble function on the reference element  $\hat{T}$  that vanishes over the edge  $\partial \hat{T}$ . Then the internal polynomial space  $\mathcal{I}_m(\hat{T})$  is defined by

$$\mathcal{I}_m(\hat{T}) := \operatorname{span}\{b_{\hat{T}}v|v \in P_{m-3}(\hat{T})\} \text{ for } m \ge 3.$$

$$(4.4)$$

The dimension of  $\mathcal{I}_m(\hat{T})$  is  $\hat{m} := (m-1)(m-2)/2$ . Let one set of basis functions be denoted as

$$\{N_i : i = 1, \cdots, \hat{m}\}.$$
 (4.5)

### 4.2 Local interpolation based on Lagrange polynomials

Now following [24], we construct a suitable interpolation operator  $\Pi : C(\hat{T}) \to \hat{P}$ , such that

$$\begin{cases} \Pi f \big|_{\hat{F}^i} \in P_{m_i}(\hat{F}^i) \text{ for } i = 1, 2, 3, \\ \Pi f \in \hat{P}, \end{cases} \quad \forall f \in C(\hat{T}). \end{cases}$$

It consists of the external interpolation E and internal interpolation I.

The external interpolation E is an interpolation operator over Fekete points along each face. Since the boundary points of Fekete points are the one-dimensional Gauss-Lobatto points [10], we get the Gauss-Lobatto interpolation along each face  $\hat{F}^i$  for i = 1, 2, 3. Next, we extend the polynomial Ef to the domain  $\Omega$ , and denote the extension by  $\mathbf{E}(Ef)$ . With the shape functions in Section 4.1, we can combine two procedures to obtain  $\mathbf{E}(Ef)$ . Specifically, we use the nodal and side shape functions to perform external interpolation along each face on Fekete points. For example, for  $\hat{F}^i$  with i = 1, 2, 3, the nodal shape functions are  $N_{i+1}^1$  and  $N_{i+2}^1$ , and we combine them with the  $(m_i - 1)$  side shape functions to obtain the interpolation on this face  $\hat{F}^i$ :

$$Ef|_{\hat{F}^{i}} = \left(\sum_{j=2}^{m_{i}} c_{i,j} N_{i}^{m_{i},j} + f(\hat{\mathbf{z}}_{i+1}) N_{i+1}^{1} + f(\hat{\mathbf{z}}_{i+2}) N_{i+2}^{1}\right) \bigg|_{\hat{F}^{i}},$$
(4.6)

where the coefficients  $\{c_{i,j}\}_{j=2}^{m_i}$  are to be determined such that  $Ef|_{\hat{F}^i}$  interpolates the  $m_i$ th order Fekete points on the face  $\hat{F}^i$ .

The extension  $\mathbf{E}(Ef)$  of Ef is then given by

$$\mathbf{E}(Ef) = \sum_{i=1}^{3} \sum_{j=2}^{m_i} c_{i,j} N_i^{m_i,j} + \sum_{i=1}^{3} f(\hat{\mathbf{z}}_i) N_i^1.$$
(4.7)

One can easily verify that  $\mathbf{E}(Ef)|_{\hat{F}^i} = Ef|_{\hat{F}^i}$  for i = 1, 2, 3.

Third, we define the internal interpolation operator  $I: C(\hat{T}) \to \mathcal{I}_m(\hat{T})$  that interpolates the Fekete points  $\{\mathbf{z}_i^*\}_{i=1}^{\hat{m}}$ , using the set of internal basis functions  $\{\hat{N}_i: i = 1, \cdots, \hat{m}\}$  (4.5):

$$\{\mathbf{z}_{j}^{*}\}_{j=1}^{\hat{m}} := \arg \max_{\{\mathbf{z}_{1}, \mathbf{z}_{2}, \cdots, \mathbf{z}_{\hat{m}}\} \subset \hat{T}} \det(V(\mathbf{z}_{1}, \mathbf{z}_{2}, \cdots, \mathbf{z}_{\hat{m}})).$$
(4.8)

Here,  $V(\mathbf{z}_1, \mathbf{z}_2, \cdots, \mathbf{z}_{\hat{m}}) \in \mathbb{R}^{\hat{m} \times \hat{m}}$  has entries  $V_{ij} = \mathring{N}_j(\mathbf{z}_i)$  for  $i, j = 1, \cdots, \hat{m}$ . Since  $\mathring{N}_i$  vanishes on  $\partial \hat{T}$  for  $i = 1, \cdots, \hat{m}, \{\mathbf{z}_i^*\}_{i=1}^{\hat{m}} \subset \hat{T} \setminus \partial \hat{T}$ . For any  $f \in C(\hat{T})$ , the internal interpolation If is then given by

$$If = \sum_{i=1}^{\hat{m}} a_i \mathring{N}_i \text{ s.t. } If(\mathbf{z}_j^*) = f(\mathbf{z}_j^*) \text{ for all } j = 1, \cdots, \hat{m}.$$
 (4.9)

Since the internal shape functions vanish along all three faces,  $If|_{\partial \hat{T}} = 0$  for all  $f \in C(\hat{T})$ .

Finally, we can define local interpolation operator for any function  $f \in C(T)$  by,

$$\Pi f := \mathbf{E}(Ef) + I(f - \mathbf{E}(Ef)). \tag{4.10}$$

By construction, we can derive the following properties of  $\Pi$ .

**Proposition 4.1.** The element-wise interpolation  $\Pi : C(\hat{T}) \to \hat{P}$  satisfies the following properties. (i)  $\Pi p = p$ , for all  $p \in \hat{P}$ ; (ii)  $\Pi f|_{\hat{F}^i} = Ef \in P_{m_i}(\hat{F}^i)$  is the Lagrange interpolation on Fekete points along the face  $\hat{F}^i$ , for i = 1, 2, 3.

### 4.3 Element-wise interpolation

For a given mesh  $\mathcal{T}_n$ , we define the operator  $\Pi_{\mathcal{T}_n} : C(\mathcal{B}) \to V_n$  by applying  $\Pi$  to each element  $T \in \mathcal{T}_n$ :

$$\Pi_{\mathcal{T}_n} f\big|_T := (\Pi(f \circ M_T)) \circ M_T^{-1}, \quad \forall T \in \mathcal{T}_n,$$
(4.11)

where the affine transformation  $M_T := A_T \mathbf{z} + \mathbf{b}$   $(A_T \in \mathbb{R}^{2 \times 2}$  is invertible and  $\mathbf{b} \in \mathbb{R}^2)$  maps the reference triangle  $\hat{T}$  to the "physical" element  $\bar{T}$ .

Now we can give the following algorithm for approximating band functions. By sequentially generating the adaptive mesh and the polynomial degree required for each element and each face, we can then use the interpolation formula (4.11) to adaptively interpolate  $\lambda_j$  and then generate the corresponding band function  $\omega_j$  for  $j = 1, 2, \dots, L-1$ . Algorithm 2 *hp*-adaptive sampling algorithm

**Require:**  $\mathcal{T}_n$ : an adaptive mesh from Algorithm 1;  $\{\ell_T\}_{T\in\mathcal{T}_n}$ : layer of each element defined in Definition 3.2;  $\mu$ : a positive parameter defined in Definition 3.3 satisfying condition (5.9) to be introduced; **Ensure:**  $\omega_j, j = 1, 2, \cdots, L-1$ Choose the element-wise degree  $\{n_T\}_{T\in\mathcal{T}_n}$  by Definition 3.3; Choose the facewise degree  $\{m_F\}_{F\in\mathcal{F}_n}$  by Definition 3.4; Compute the local interpolation  $\Pi\lambda_j$  by (4.10); Compute the global interpolation  $\Pi_{\mathcal{T}_n}\lambda_j$  by (4.11);  $\omega_j = c \cdot \sqrt{\lambda_j}, j = 1, 2, \cdots, L-1.$ 

Now we compare the computational complexity of Algorithm 2 with the global polynomial interpolation method [34] (ignoring their accuracy and computational cost of the adaptive mesh refinement in Algorithm 1). Let  $\mathcal{T}_n$  and  $V_n$  be the adaptive mesh generated by Algorithm 1 and conforming FEM space (3.10), and let N be the dimension of  $V_n$ , i.e., the number of sampling points. Suppose we use the same number N of sampling points for both methods, and the computational complexity in solving the Helmholtz eigenvalue problem is the same. The main difference in the complexity arises from the Lagrange interpolation. To compute N Lagrange polynomial functions using the classical Gauss elimination method involves  $\mathcal{O}(N^3)$  flops. In contrast, Algorithm 2 is based on element-wise interpolation, which has much lower computational complexity. Let the number of sampling points of each element  $T \in \mathcal{T}_n$  be  $N_T$ . Then the local Lagrange interpolation involves  $\mathcal{O}(N_T^3)$  flops. Since there is no communication among each element, local Lagrange interpolation over each element can be performed in parallel. Hence, the total computational complexity is  $\max_{T \in \mathcal{T}_n} \mathcal{O}(N_T^3)$  flops.

### 5 Convergence analysis

We establish in this section exponential and algebraic convergence of Algorithm 2 in Theorems 5.5 and 5.6 when the number of iterations  $n \to \infty$  in Algorithm 1, resting upon whether the number of crossings is finite or not. This indicates that Algorithm 2 outperforms the global polynomial interpolation methods [34] with a convergence rate that is at least twice as fast. To establish Theorems 5.5 and 5.6, we first derive the approximation property of the local polynomial interpolation operator II in Theorem 5.3. Then we formulate the element-wise approximation property of  $\Pi_{\mathcal{T}_n}$  in Theorem 5.4 and propose a proper assumption on the slope parameter  $\mu$  (5.9) such that the approximation error in the marked element patch  $\mathcal{M}_n$  dominates.

### 5.1 Local approximation property

Now we establish the approximation property of the local interpolation operator II. Recall that for any  $\{m, m_1, m_2, m_3\} \in \mathbb{N}^4_+$  with  $m \ge m_i$  for i = 1, 2, 3, the local space  $\hat{P}$  is defined by (4.1). First, we give the stability of the external interpolation and internal interpolation.

**Theorem 5.1.** Let E and I be the external interpolation and the internal interpolation defined by (4.6) and (4.9), respectively. Then the following stability estimates hold

$$\|Ef\|_{\infty,\hat{F}^{i}} \lesssim \log m_{i} \|f\|_{\infty,\hat{F}^{i}}, \qquad \forall f \in C(\hat{F}^{i}), \tag{5.1}$$

$$\|If\|_{\infty\hat{T}} \le \hat{m} \|f\|_{\infty\hat{T}}, \qquad \forall f \in C(\hat{T}).$$
(5.2)

*Proof.* Note that the Lebesgue constant of Gauss-Lobatto points in an interval (cf. [29]) is  $\mathcal{O}(\log m_i)$  and that of Fekete points in a triangle (cf. [30]) has an upper bound  $\hat{m}$ . Then the desired result follows.

Next, we derive the stability of the extension  $\mathbf{E}$ . To this end, we first introduce an inverse inequality for algebraic polynomials,

**Lemma 5.1** (Inverse estimate of Markov type [35]). For all  $m \in \mathbb{N}$  and  $p \in P_m(\hat{T})$ , there holds

$$\|\nabla p\|_{\infty,\hat{T}} \le 2m^2 \|p\|_{\infty,\hat{T}}$$

**Theorem 5.2.** For each  $p \in C(\partial \hat{T})$  with  $p|_{\hat{F}^i} \in P_{m_i}(\hat{F}^i)$  for all i = 1, 2, 3, the extension **E** defined in (4.7) satisfies

$$\|\mathbf{E}(p)\|_{\infty,\hat{T}} + m^{-2} \|\nabla \mathbf{E}(p)\|_{\infty,\hat{T}} \lesssim \|p\|_{\infty,\partial\hat{T}}.$$
(5.3)

Moreover, the extension map  $\mathbf{E}: p \mapsto \mathbf{E}(p)$  is a bounded linear operator.

Proof. Since Lemma 5.1 implies  $m^{-2} \|\nabla \mathbf{E}(p)\|_{\infty,\hat{T}} \leq 2 \|\mathbf{E}(p)\|_{\infty,\hat{T}}$ , it suffices to prove  $\|\mathbf{E}(p)\|_{\infty,\hat{T}} \lesssim \|p\|_{\infty,\partial\hat{T}}$ . We may assume p vanishes at all vertices and all sides but one, e.g., the third face  $\hat{F}^3 = \{(x,0): -1 \leq x \leq 1\}$ . Then  $\mathbf{E}(p)$  can be written in the form  $\mathbf{E}(p) = \sum_{i=2}^{m_3} c_{3,i} N_3^{m_3,i}$ , where the unknown coefficients  $\{c_{3,i}\}_{i=2}^{m_3}$  are to be determined such that

$$\mathbf{E}(p) = p \text{ on } \hat{F}^3. \tag{5.4}$$

Since  $N_2^1 - N_1^1 = x$ , after plugging into (4.3), we obtain

$$\mathbf{E}(p) = \sum_{i=2}^{m_3} c_{3,i} N_1^1 N_2^1 x^{i-2} = N_1^1 N_2^1 \sum_{i=2}^{m_3} c_{3,i} x^{i-2}.$$

Together with (4.2) and (5.4), this leads to

$$\begin{aligned} \max_{(x,y)\in\hat{T}} |\mathbf{E}(p)| &= \max_{(x,y)\in\hat{T}} \left| N_1^1 N_2^1 \sum_{i=2}^{m_3} c_{3,i} x^{i-2} \right| \\ &= \max_{\substack{0 \le x \le 1, 0 \le y \le \sqrt{3}(1-x) \\ -1 \le x \le 0, 0 \le y \le \sqrt{3}(1+x)}} \frac{1}{4} \left( 1 - x - y/\sqrt{3} \right) \left( 1 + x - y/\sqrt{3} \right) \left| \sum_{i=2}^{m_3} c_{3,i} x^{i-2} \right| \\ &\le \max_{-1 \le x \le 1} \frac{1}{4} (1 - x^2) \left| \sum_{i=2}^{m_3} c_{3,i} x^{i-2} \right| = \max_{(x,y)\in\hat{F}^3} |p|. \end{aligned}$$

This proves the desired result.

Finally, we can state the following quasi-optimal approximation property of  $\Pi$ .

**Theorem 5.3.** For every  $m \in \mathbb{N}_+$ , the interpolation operator  $\Pi : C(\hat{T}) \to \hat{P}$  satisfies

$$\|f - \Pi f\|_{\infty,\hat{T}} \lesssim \mathcal{C}_{\mathrm{stab}}(m) \inf_{p \in \hat{P}} \|f - p\|_{\infty,\hat{T}} \text{ with } \mathcal{C}_{\mathrm{stab}}(m) := m \log m.$$

*Proof.* The stability of E and  $\mathbf{E}$  in (5.1) and (5.3) gives the stability of  $\mathbf{E}E: C(\partial \hat{T}) \to C(\hat{T}):$ 

$$\|\mathbf{E}(Ef)\|_{\infty,\hat{T}} \lesssim \|Ef\|_{\infty,\partial\hat{T}} \lesssim \log m \|f\|_{\infty,\partial\hat{T}}.$$
(5.5)

Together with the stability estimate of I (5.2), it leads to

$$\begin{split} \|\Pi f\|_{\infty,\hat{T}} &\leq \|\mathbf{E}(Ef)\|_{\infty,\hat{T}} + \|I(f - \mathbf{E}(Ef))\|_{\infty,\hat{T}} \\ &\lesssim \log m \|f\|_{\infty,\partial\hat{T}} + m \|f - \mathbf{E}(Ef)\|_{\infty,\hat{T}} \\ &\leq \log m \|f\|_{\infty,\partial\hat{T}} + m (\|f\|_{\infty,\hat{T}} + \|\mathbf{E}(Ef)\|_{\infty,\hat{T}}). \end{split}$$

	_	

In view of (5.3) and (5.5), we obtain

$$\|\Pi f\|_{\infty,\hat{T}} \lesssim \log m \|f\|_{\infty,\partial\hat{T}} + m \|f\|_{\infty,\hat{T}} + m \log m \|f\|_{\infty,\partial\hat{T}} \lesssim m \log m \|f\|_{\infty,\hat{T}}$$

Note that  $\Pi p = p$  for all  $p \in \hat{P}$ , consequently, we arrive at

$$||f - \Pi f||_{\infty,\hat{T}} = ||(f - p) - \Pi (f - p)||_{\infty,\hat{T}} \lesssim m \log m ||f - p||_{\infty,\hat{T}}.$$

Taking the infimum over all  $p \in \hat{P}$  leads to the desired assertion.

Furthermore, we can derive the best approximation in  $\hat{P}$  in uniform norm.

**Lemma 5.2** (Best approximation in  $\hat{P}$ ). For all integer  $0 \leq j \leq \min(m_1, m_2, m_3)$  and all  $f \in W^{j+1,\infty}(\hat{T})$ , there exists a bounded constant  $C_{\text{bpa}}$  independent of j and  $h_{\hat{T}}$  such that

$$\inf_{p \in \hat{P}} \|f - p\|_{\infty, \hat{T}} \le C_{\text{bpa}} \left(\frac{h_{\hat{T}}}{j+1}\right)^{j+1} |f|_{W^{j+1,\infty}(\hat{T})},\tag{5.6}$$

where  $|\cdot|_{W^{k,\infty}(D)}$  is the  $W^{k,\infty}(D)$  semi-norm.

*Proof.* Recall the following well-known result [5, Equation (5.4.16)]: for all  $k \ge 0$ , there holds

$$\inf_{p \in P_k(\hat{T})} \|f - p\|_{\infty, \hat{T}} \le C_{\text{bpa}} \left(\frac{h_{\hat{T}}}{k+1}\right)^{k+1} |f|_{W^{k+1,\infty}(\hat{T})}, \quad \forall f \in W^{k+1,\infty}(\hat{T}).$$

By choosing  $k := \min(m_1, m_2, m_3)$  and noting  $P_k(\hat{T}) \subset \hat{P}$ , we establish the lemma.

**Remark 5.1** (Boundedness of the constant  $C_{bpa}$ ). Note that the best approximation error  $\inf_{p \in P_k(\hat{T})} ||f - p||_{\infty,\hat{T}}$  can be bounded by the product of the so-called Whitney constant and the modulus of smoothness of order k + 1 of a function  $f \in C(\hat{T})$ , and note also that the upper bound of the modulus of smoothness is given by  $\left(\frac{h_{\hat{T}}}{k+1}\right)^{k+1} |f|_{W^{k+1,\infty}(\hat{T})}$  for  $f \in W^{k+1,\infty}(\hat{T})$  [17, Equation (2.4)]. Furthermore, the Whitney constant is proved to be bounded [21, 28, 8, 12, 19]. Consequently, the constant  $C_{bpa}$  is bounded.

### 5.2 Global approximation error

Now we analyze the convergence of Algorithm 2. The following two lemmas are standard.

**Lemma 5.3** ([6, Theorem 3.1.2]). Suppose  $f \in W^{k,\infty}(T)$  for  $k \in \mathbb{N}$  and some element  $T \in \mathcal{T}_n$ . Then  $\hat{f} := f \circ M_T$  belongs to  $W^{k,\infty}(\hat{T})$ , and there exists a constant C(k) depending on k such that

$$\begin{aligned} |\hat{f}|_{W^{k,\infty}(\hat{T})} &\leq C(k) \|A_T\|^k |f|_{W^{k,\infty}(T)}, \\ |f|_{W^{k,\infty}(T)} &\leq C(k) \|A_T^{-1}\|^k |\hat{f}|_{W^{k,\infty}(\hat{T})}, \end{aligned}$$

where  $A_T$  is the matrix in the mapping  $M_T$ .

The next result gives bounds on  $||A_T||$  and  $||A_T^{-1}||$ .

**Lemma 5.4** ([6, Theorem 3.1.3]). There holds  $||A_T|| \le h_T \rho_{\hat{T}}^{-1}$  and  $||A_T^{-1}|| \le h_{\hat{T}} \rho_T^{-1}$ .

Next, we establish the element-wise approximation property of the global projection  $\Pi_{\mathcal{T}_n}$ .

**Theorem 5.4** (Element-wise approximation property of  $\Pi_{\mathcal{T}_n}$ ). Let  $n \in \mathbb{N}$  be sufficiently large. Let  $\mathcal{T}_n$  with element-wise and facewise degrees  $\{n_T\}_{T \in \mathcal{T}_n}$  and  $\{m_F\}_{F \in \mathcal{F}_n}$  be generated by Algorithm 1, Definitions 3.3 and 3.4 and let the hp approximation space be given by (3.10). Then for  $j = 1, \dots, L-1$ , there holds

$$\|\lambda_j - \Pi_{\mathcal{T}_n}\lambda_j\|_{\infty,T} \lesssim C_{\text{bpa}}h_T |\lambda_j|_{W^{1,\infty}(T)}, \qquad \qquad \text{for } T \in \mathcal{M}_n \tag{5.7}$$

$$\|\lambda_j - \Pi_{\mathcal{T}_n} \lambda_j\|_{\infty, T} \lesssim \mathcal{C}_{\mathrm{stab}}(n_T) \mathcal{C}_{\mathrm{bpa}} C(\tilde{n}) \left(\frac{\sqrt{3}h_T}{\tilde{n}}\right)^n |\lambda_j|_{W^{\tilde{n}, \infty}(T)}, \qquad \text{for } T \in \mathcal{T}_n \backslash \mathcal{M}_n, \tag{5.8}$$

where  $C_{\text{stab}}(n_T) := n_T \log n_T$ ,  $C(\tilde{n})$  is the constant in Lemma 5.3 and  $\tilde{n} = \min(m_{F_T^1}, m_{F_T^2}, m_{F_T^3}) + 1$ . Furthermore, if the slope parameter  $\mu$  satisfies the following condition

$$\ln\left(C_{\text{stab}}(\mu\ell_T)C(n_T+1)C_{\lambda_j,T}^{\lceil\mu\ell_T+1\rceil}\right) + \lceil\mu\ell_T+1\rceil\left(\ln\sqrt{3} + F(\ell_T) - \ln\lceil\mu\ell_T+1\rceil\right) \lesssim \ln\left(C_{\text{max}}\right) + F(1)$$
(5.9)

for all  $T \in \mathcal{T}_n \setminus \mathcal{M}_n$ , then

$$\max_{T \in \mathcal{T}_n \setminus \mathcal{M}_n} \|\lambda_j - \Pi_{\mathcal{T}_n} \lambda_j\|_{\infty, T} \lesssim \min_{T \in \mathcal{M}_n} \|\lambda_j - \Pi_{\mathcal{T}_n} \lambda_j\|_{\infty, T}.$$

Here,  $C_{\max} := \max_{T \in \mathcal{M}_n} |\lambda_j|_{W^{1,\infty}(T)}$ ,  $C_{\lambda_j,T}^{\tilde{n}} := |\lambda_j|_{W^{\tilde{n},\infty}(T)}$ , and  $F(\ell) := \ln h_1 - \frac{n-\ell}{2} \ln 2$  with  $h_1$  being the initial mesh size.

*Proof.* By the definition of the element-wise interpolation (4.11), we obtain

$$\|\lambda_j - \Pi_{\mathcal{T}_n} \lambda_j\|_{\infty,T} = \|(\lambda_j \circ M_T - \Pi(\lambda_j \circ M_T)) \circ M_T^{-1}\|_{\infty,T}, \quad \forall T \in \mathcal{T}_n.$$

Let  $\hat{\lambda}_j := \lambda_j \circ M_T$ . Then we derive

$$\|\lambda_j - \Pi_{\mathcal{T}_n} \lambda_j\|_{\infty,T} = \left\|\hat{\lambda}_j - \Pi \hat{\lambda}_j\right\|_{\infty,\hat{T}}$$

Together with Theorem 5.3, we derive

$$\|\lambda_j - \Pi_{\mathcal{T}_n} \lambda_j\|_{\infty,T} \lesssim C_{\text{stab}}(n_T) \inf_{p \in \hat{P}} \|\hat{\lambda}_j - p\|_{\infty,\hat{T}}.$$
(5.10)

Now the proof of the theorem consists of three steps. First, if  $T \in \mathcal{M}_n$ , then  $\lambda_j \in W^{1,\infty}(T)$  and  $\hat{\lambda}_j \in W^{1,\infty}(\hat{T})$ . Moreover, by Definition (3.3), we have  $n_T = 2$ . Thus,  $C_{\text{stab}}(n_T) \leq 3$ . By Lemma 5.2, we derive

$$\inf_{p \in \hat{P}} \|\hat{\lambda}_j - p\|_{\infty, \hat{T}} \le C_{\text{bpa}} h_{\hat{T}} |\hat{\lambda}_j|_{W^{1, \infty}(\hat{T})}$$

Then combining with Lemmas 5.3 and 5.4, this further leads to

$$\|\lambda_j - \Pi_{\mathcal{T}_n} \lambda_j\|_{\infty,T} \lesssim \mathcal{C}_{\mathrm{bpa}} h_{\hat{T}} |\lambda_j|_{W^{1,\infty}(\hat{T})} \lesssim \mathcal{C}_{\mathrm{bpa}} h_T |\lambda_j|_{W^{1,\infty}(T)}.$$

This proves the assertion (5.7). Second, if  $T \in \mathcal{T}_n \setminus \mathcal{M}_n$ , then  $\hat{\lambda}_j := \lambda_j \circ M_T \in W^{\tilde{n},\infty}(T)$ , and (5.10) and Lemma 5.2 yield

$$\inf_{p\in\hat{P}} \|\hat{\lambda}_j - p\|_{\infty,\hat{T}} \le C_{\text{bpa}} (h_{\hat{T}}/\tilde{n})^{\tilde{n}} |\hat{\lambda}_j|_{W^{\tilde{n},\infty}(\hat{T})}.$$

Together with Lemmas 5.3 and 5.4 and noting  $\frac{h_{\hat{T}}}{\rho_{\hat{T}}} = \sqrt{3}$ , we obtain

$$\begin{split} \|\lambda_j - \Pi_{\mathcal{T}_n} \lambda_j\|_{\infty,T} &\lesssim \mathcal{C}_{\mathrm{stab}}(n_T) \mathcal{C}_{\mathrm{bpa}}(h_{\hat{T}}/\tilde{n})^{\tilde{n}} |\hat{\lambda}_j|_{W^{\tilde{n},\infty}(\hat{T})} \\ &\leq \mathcal{C}_{\mathrm{stab}}(n_T) \mathcal{C}_{\mathrm{bpa}} C(\tilde{n}) (h_{\hat{T}}/\tilde{n})^{\tilde{n}} (h_T \rho_{\hat{T}}^{-1})^{\tilde{n}} |\lambda_j|_{W^{\tilde{n},\infty}(T)} \\ &= \mathcal{C}_{\mathrm{stab}}(n_T) \mathcal{C}_{\mathrm{bpa}} C(\tilde{n}) (\sqrt{3}h_T/\tilde{n})^{\tilde{n}} |\lambda_j|_{W^{\tilde{n},\infty}(T)}. \end{split}$$

This proves the assertion (5.8). Finally, we prove the last statement that if the slope parameter  $\mu$  satisfies (5.9), then for all elements  $T \in \mathcal{T}_n^R$  with  $n_T = \lceil \mu \ell_T \rceil$ , the right-hand side of (5.8) is no greater than the right-hand side of (5.7) for some element in  $\mathcal{M}_n$ . Definition 3.4 implies  $n_T \geq m_{F_T^i}$ , for i = 1, 2, 3, for all  $T \in \mathcal{T}_n \setminus \mathcal{M}_n$ . Moreover, if  $m_{F_T^i} \leq n_T$ , there exists  $\tilde{T} \in \mathcal{F}_F$  such that  $m_{F_T^i} = n_{\tilde{T}}$ . This motivates setting  $\tilde{n} = n_T + 1$  in (5.8) when analyzing the choice of  $\mu$ . Algorithm 1 and Definition 3.2 indicate the diameters of elements satisfy  $h_T \sim h_1 2^{-(n-\ell_T)/2}$ . Thus, ignoring all the absolute constants, the right hand side of (5.8) becomes

$$C_{\text{stab}}(n_T)C_{\text{bpa}}C(n_T+1)(\sqrt{3}h_12^{-(n-\ell_T)/2}/(n_T+1))^{n_T+1}|\lambda_j|_{W^{n_T+1,\infty}(T)},$$
(5.11)

and the right-hand side of (5.7) is

$$C_{bpa}h_1 2^{-(n-1)/2} |\lambda_j|_{W^{1,\infty}(T)}.$$
(5.12)

Then the condition (5.9) implies the maximum value of (5.12) for all  $T \in \mathcal{M}_n$  is bounded above by the minimum value of (5.12) for all  $T \in \mathcal{T}_n \setminus \mathcal{M}_n$ , and this completes the proof.

**Remark 5.2.** Note that the constant  $C_{bpa}$  is bounded, cf Remark 5.1. Assume that  $h_T$  and  $h_T^{\hat{n}}$  dominate the right-hand sides of (5.7) and (5.8) respectively, with a small enough initial mesh size,  $\mu = 1$  or  $\mu = 0.5$  is sufficient for (5.9). Then the point with the largest error lies in  $\mathcal{M}_n$ , as is numerically verified.

Next we discuss the approximation error under the assumption that the maximum error exists in some element belonging to  $\mathcal{M}_n$ . Since the mesh  $\mathcal{T}_n$  generated by the procedure is geometrically refined towards branch points with n layers, we derive a convergence rate according to the classification of branch points.

First, if the band functions we consider have a finite number of branch points, then the number of elements  $\#\mathcal{T}_n = \mathcal{O}(n)$  when n is sufficiently large, leading to the number of sampling points  $N = \mathcal{O}(n^3)$ . See Figure 4 for an illustration, in which the left two figures are the mesh refinement in case of a finite number of branch points in the IBZ of the square lattice. In this case, we derive an exponential convergence rate,

**Theorem 5.5** (Convergence rate for hp element-wise interpolation with finite branch points). Let  $n \in \mathbb{N}$  be sufficiently large. Let  $\mathcal{T}_n$  with element-wise and facewise degrees  $\{n_T\}_{T \in \mathcal{T}_n}$  and  $\{m_F\}_{F \in \mathcal{F}_n}$  be generated by Algorithm 1, Definitions 3.3 and 3.4, the hp approximation space be defined in (3.10), and the slope parameter  $\mu$  satisfy (5.9). If the band functions we consider have a finite number of branch points, then there holds

$$\|\lambda_j - \prod_{\mathcal{T}_n} \lambda_j\|_{\infty, \mathcal{B}} \lesssim \exp(-bN^{\frac{1}{3}}) |\lambda_j|_{W^{1,\infty}(\mathcal{B})} \text{ for } j = 1, \cdots, L-1.$$

Here, the constant b > 0 is independent of the number of sampling points N.

Proof. Theorem 3.2 indicates that the set of singular points **S** as a subset of the branch points is marked at the *n*th iteration, i.e.,  $\mathbf{S} \subset \mathcal{M}_n$ . Algorithm 1 implies that the diameter of each element  $T \in \mathcal{M}_n$ satisfies  $h_T \sim 2^{-n/2}$  for all  $T \in \mathcal{M}_n$ . The choice of  $\mu$  (i.e., the minimum value satisfying (5.9)) ensures that the element-wise interpolation error measured in the uniform norm attains its maximum over  $\mathcal{M}_n$ . This and Theorem 5.4 imply

$$\|\lambda_j - \Pi_{\mathcal{T}_n} \lambda_j\|_{\infty, \mathcal{B}} \lesssim 2^{-n/2} |\lambda_j|_{W^{1,\infty}(\mathcal{B})} \text{ for } j = 1, \cdots, L-1.$$
(5.13)

Since the number of sampling points  $N = \mathcal{O}(n^3)$  for sufficiently large n, the desired assertion follows.  $\Box$ 

Next we discuss the convergence rate when the number of branch points is infinite; see Figure 4 for an illustration of the mesh refinement. Asymptotically, we have the number of elements  $\#\mathcal{T}_n = \mathcal{O}(2^{\frac{n}{2}})$  and the number of sampling points  $N \sim 2^{\frac{n}{2}}$ . Similarly, we derive the following convergence rate for hp element-wise interpolation with infinite branch points.

**Theorem 5.6.** Under the conditions of Theorem 5.5, if the number of branch points is infinite, then

$$\|\lambda_j - \Pi_{\mathcal{T}_n}\lambda_j\|_{\infty,\mathcal{B}} \lesssim N^{-1}|\lambda_j|_{W^{1,\infty}(\mathcal{B})} \text{ for } j = 1, \cdots, L-1.$$

**Remark 5.3.** With  $n \in \mathbb{N}$  sufficiently large, the approximation error of the band function is given by

$$\begin{split} \|\omega_j - c \cdot \sqrt{\Pi_{\mathcal{T}_n} \lambda_j}\|_{\infty, \mathcal{B}} &= c^2 \cdot \left\| \frac{\lambda_j - \Pi_{\mathcal{T}_n} \lambda_j}{\omega_j + c \cdot \sqrt{\Pi_{\mathcal{T}_n} \lambda_j}} \right\|_{\infty, \mathcal{B}} \\ &\leq c \cdot \frac{\|\lambda_j - \Pi_{\mathcal{T}_n} \lambda_j\|_{\infty, \mathcal{B}}}{2\sqrt{\lambda_j(\tilde{\mathbf{k}})} - \sqrt{err}}, \end{split}$$

where  $\tilde{\mathbf{k}} = \arg \max_{\mathbf{k} \in \mathcal{B}} |\omega_j(\mathbf{k}) - c \cdot \sqrt{\Pi_{\mathcal{T}_n} \lambda_j(\mathbf{k})}|$  and  $err = \|\lambda_j - \Pi_{\mathcal{T}_n} \lambda_j\|_{\infty, \mathcal{B}}.$ 

Thus, the decay rates of  $\|\omega_j - c \cdot \sqrt{\prod_{\mathcal{T}_n} \lambda_j}\|_{\infty, \mathcal{B}}$  and  $\|\lambda_j - \prod_{\mathcal{T}_n} \lambda_j\|_{\infty, \mathcal{B}}$  are roughly of the same order.



(a) One singular point(b) Two singular points(c) One singular line  $k_2 = (d)$  Two singular lines  $k_2 = \frac{1}{a}(\frac{\pi}{3}, \frac{\pi}{3})$   $\frac{1}{a}(\frac{\pi}{3}, \frac{\pi}{3})$  and  $\frac{1}{a}(\frac{2\pi}{3}, \frac{\pi}{3})$   $k_1$  $k_1$  and  $k_2 + k_1 + \frac{\pi}{a} = 0$  with  $\frac{2}{a} \le k_1 \le \frac{3}{a}$ 

Figure 4: Algorithm 1 with finite (the first two) and infinite (the last two) singularities.

**Remark 5.4.** Theorems 5.5 and 5.6 hold asymptotically when the number of iterations n is sufficiently large. In practice, the largest allowed iteration  $n_{\max}$  of Algorithm 1 is not large. Then the number N of sampling points is  $N = \mathcal{O}(2^{\frac{n}{\theta}})$  with  $\theta > 2$ , which leads to a faster convergence rate. Indeed, we observe a second-order convergence rate in numerical experiments.

Last, we discuss the tolerance  $tol_2$  and stopping criterion  $n_{\text{max}}$  in Algorithm 1. By (5.13), we have

$$\|\lambda_j - \prod_{\mathcal{T}_n} \lambda_j\|_{\infty,\mathcal{B}} \lesssim 2^{-n/2} |\lambda_j|_{W^{1,\infty}(\mathcal{B})} \text{ for } j = 1, \cdots, L-1.$$

Thus,  $tol_2$  and  $n_{max}$  should satisfy  $2^{-n_{max}/2} \simeq tol_2 \simeq \epsilon$ , i.e., comparable with the target precision  $\epsilon$ .

# 6 Numerical experiments

To show the performance of Algorithm 2, we consider 2D PhCs with a square unit cell in Figure 2 and a hexagonal unit cell in Figure 3 as in [34]. Given  $\mathbf{k} \in \mathcal{B}$ , we solve the Helmholtz eigenvalue problem

(2.10) using the conforming Galerkin FEM. Due to the high contrast between the relative permittivity of the circular medium and the external one, we employ a fitted mesh generated by *distmesh2d* [25]; see Figures 5(a) and 5(b) for the mesh  $\mathcal{T}_h$  with a mesh size h = 0.025a for square lattice and h = 0.05afor hexagonal lattice. The associated conforming piecewise affine space is given by

$$V_h = \{ v_h \in C(\overline{\Omega}) : v_h |_K \in P_1(K) \ \forall K \in \mathcal{T}_h \}.$$

Given  $\mathbf{k} \in \mathcal{B}$ , the Galerkin FEM approximation to problem (2.10) is to find non-trivial eigenpair  $(\lambda_h, u_h) \in (\mathbb{R}, V_h)$  such that

$$\begin{cases} a(u_h, v_h) = \lambda_h b(u_h, v_h), \quad \forall v \in V_h, \\ \|u_h\|_{L^2_{\alpha}(\Omega)} = 1. \end{cases}$$

$$(6.1)$$





(a) Square lattice #DOFs=2107 (b) Hexagonal lattice #DOFs=1781

Figure 5: Discretization of the unit cell  $\Omega$  by distmesh2d.

In both cases, IBZ is a triangle. To obtain the reference solution with sufficient accuracy, we discretize  $\mathcal{B}$  with 20503 evenly distributed points (denoted by  $\hat{\mathcal{B}}$ ) to compute the interpolation error and the accuracy of the proposed method. The pointwise relative error is then computed on  $\hat{\mathcal{B}}$  by

$$e_i(\mathbf{k}) := \frac{|\omega_i(\mathbf{k}) - \hat{\omega}_i(\mathbf{k})|}{\omega_i(\mathbf{k})} \quad \text{for } \mathbf{k} \in \hat{\mathcal{B}} \text{ and } i = 1, \cdots, 5$$

Here,  $\omega_i(\mathbf{k})$  is the *i*th reference band function obtained directly by the Galerkin FEM over  $\hat{\mathcal{B}}$  using the same mesh on the unit cell  $\Omega$ , and  $\hat{\omega}_i(\mathbf{k})$  is by the element-wise polynomial interpolation. Since the first few low-frequency band functions are typically of practical interest [16], we show the performance of the numerical scheme using only the first five band functions. We measure the accuracy of the methods using maximum relative error

$$\operatorname{error}_{\infty} := \max_{i=1,\cdots,5} \max_{\mathbf{k} \in \hat{\mathcal{B}}} |e_i(\mathbf{k})|.$$

Figure 6 shows the location of singularities in the band functions with the distance between adjacent band functions. The areas where adjacent band functions are close or coincident are indicated in dark blue. These regions are targeted by Algorithm 1, allowing to detect the locations of singularities.

### 6.1 Numerical tests with square lattice in Figure 2

We first test Algorithm 2 on the square lattice with the TE mode and compare the performance of GPI with Fekete points and Chebyshev points of the second kind (Cheb2).

We present in Figure 7 the decay of  $\operatorname{error}_{\infty}$  against N on a log-log scale of Algorithm 2 with  $\kappa \in \{1, 2, 2\sqrt{2}, 3, 4\}$  and  $\mu \in \{1, 0.5\}$  as suggested in Remark 5.2. The performance of GPI with Fekete



Figure 6: Locations with distance between two adjacent band functions below 0.01 are marked blue.



Figure 7: Maximum relative error with respect to the number of sampling points.

points and Cheb2 are also provided to facilitate comparison. Since Fekete points for polynomial spaces with (total) degrees higher than 18 are still unknown [30], GPI with Fekete points is restricted to  $N \leq 190$ . We observe that Algorithm 2 outperforms GPI for all cases. Furthermore, an algebraic convergence rate between first order and second order is observed for any parameters of  $\kappa$  and  $\mu$ .

Figure 8 displays the adaptive meshes generated by Algorithm 1 with  $\kappa = 2\sqrt{2}$  and  $\kappa = 2$ , respectively. Comparing with Figure 6(a), one observes that Algorithm 1 adaptively refines the elements with crossings. Note that larger  $\kappa$  indicates more elements being marked and refined in each loop. The initial mesh has little effect on the resulting adaptive mesh since meshes overlap in the first several loops.

Figures 9 and 10 display the pointwise relative error  $e_i(\mathbf{k})$  of approximating the first five band functions under  $\kappa = 2\sqrt{2}$  and  $\kappa = 2$  and  $\mu = 1$  and  $\mu = 0.5$ . Due to a lack of information about the seventh band function, we are unable to determine the positions of all the singular points of the sixth band function. Therefore, the error analysis only takes into account the first five band functions, even though we interpolate the first six band functions. From Figures 9(a) and 10(a), one observes that with  $\mu = 1$  (i.e., the total polynomial degree on each element is consistent with its layer), the pointwise relative error  $e_i(\mathbf{k})$  dominates in the singular element patch  $\mathcal{T}_n^S$  with n = 8. This confirms Theorem 3.2, and (5.9). However, for a smaller slope parameter  $\mu = 0.5$ , the pointwise relative error  $e_i(\mathbf{k})$  in Figures 9(b) and 10(b) can take large values in the regular element patch  $\mathcal{T}_n^R$  with n = 8, indicating



Figure 8: Algorithm 1 with  $\kappa = 2\sqrt{2}$  (the top row) and  $\kappa = 2$  (the bottom row).



Figure 9:  $\kappa = 2\sqrt{2}$ : pointwise relative error  $e_i(\mathbf{k})$  under  $\mu = 1$  (left),  $\mu = 0.5$  (right).



that the condition (5.9) fails.

Figure 10:  $\kappa = 2$ : pointwise relative error  $e_i(\mathbf{k})$  under  $\mu = 1$  (left),  $\mu = 0.5$  (right).

### 6.2 Numerical tests with hexagonal lattice in Figure 3

Now we study 2D PhCs with infinite periodic hexagonal lattice, cf. Figure 3. The unit cell of the PhC is composed of six cylinders of dielectric material with dielectric constant  $\epsilon = 8.9$  embedded in the air and the lattice vectors  $\mathbf{a}_1 = \frac{a}{4}(e_1 + \sqrt{3}e_2)$  and  $\mathbf{a}_2 = \frac{a}{4}(e_1 - \sqrt{3}e_2)$  with lattice constant a = 3R, R is the length of hexagon edges and  $r = \frac{1}{3}R$  is the radius of cylinders. We observe similar results as before. Figure 11 shows the log-log plot of  $\operatorname{error}_{\infty}$  against N. The adaptive mesh generated by Algorithm 1 and pointwise relative errors  $\mathbf{e}_i(\mathbf{k})$  with  $\kappa = 2, 2\sqrt{2}, \ \mu = 0.5, 1$  are displayed in Figures 12, 13, 14, demonstrating the efficiency and reliability of Algorithm 2.



Figure 11: Maximum relative error with respect to the number of sampling points.



Figure 12: Algorithm 1 with  $\kappa = 2\sqrt{2}$  (the top row) and  $\kappa = 2$  (the bottom row).

# 7 Conclusion

In this work, we have presented an hp-adaptive sampling algorithm that can accurately approximate the band functions for two-dimensional photonic crystals. In the algorithm, we first generate an adaptive mesh that can refine elements containing singular points, and then assign proper polynomial degrees



Figure 13:  $\kappa = 2\sqrt{2}$ : pointwise relative error  $e_i(\mathbf{k})$  under  $\mu = 1$  (left),  $\mu = 0.5$  (right).



Figure 14:  $\kappa = 2$ : pointwise relative error  $e_i(\mathbf{k})$  under  $\mu = 1$  (left),  $\mu = 0.5$  (right).

to each element based on its level. Additionally, we assign proper polynomial degrees to each edge to ensure a continuous global polynomial space. Finally, we develop an element-wise interpolation scheme using the global polynomial space to approximate band functions. We proved a first-order convergence rate in the worst-case scenario and presented several numerical tests to verify its performance. One interesting future research question is about approximating three-dimensional band functions, which is highly nontrivial due to a lack of provable local polynomial interpolation method on a tetrahedral domain. We plan to research on designing an optimal 3D PhCs with a maximum band gap.

# References

- M. Ainsworth and T. Oden. A posteriori error estimation in finite element analysis. Comput. Methods Appl. Mech. Eng., 142(1-2):1–88, 1997.
- [2] I. Babuška and M. Suri. The p and h-p versions of the finite element method, basic principles and properties. SIAM Rev., 36(4):578–632, 1994.
- [3] G. Bao, L. Cowsar, and W. Masters, editors. *Mathematical modeling in optical science*. SIAM, Philadelphia, PA, 2001.
- [4] P. Binev, W. Dahmen, and R. DeVore. Adaptive finite element methods with convergence rates. *Numerische Mathematik*, 97:219–268, 2004.
- [5] C. Canuto, M. Y. Hussaini, A. Quarteroni, and T. A. Zang. Spectral methods. Springer-Verlag, Berlin, 2006. Fundamentals in single domains.

- [6] P. G. Ciarlet. The finite element method for elliptic problems. SIAM, Philadelphia, PA, 2002.
- [7] L. Dong, A. Mckay, A. Marcinkevicius, L. Fu, J. Li, K. Thomas, and E. Fermann. Extending effective area of fundamental mode in optical fibers. J. Lightwave Technol., 27(11):1565–1570, 2009.
- [8] V. K. Dzyadyk and I. A. Shevchuk. Theory of uniform approximation of functions by polynomials. Walter de Gruyter GmbH & Co. KG, Berlin, 2008.
- [9] T. Eidam, J. Rothhardt, F. Stutzki, F. Jansen, S. Hädrich, H. Carstens, C. Jauregui, J. Limpert, and A. Tünnermann. Fiber chirped-pulse amplification system emitting 3.8 GW peak power. *Optics Express*, 19(1):255–260, 2011.
- [10] L. Fejér. Bestimmung derjenigen Abszissen eines Intervalles, für welche die Quadratsumme der Grundfunktionen der Lagrangeschen Interpolation im Intervalle ein möglichst kleines Maximum besitzt. Annali della Scuola Normale Superiore di Pisa-Classe di Scienze, 1(3):263–276, 1932.
- [11] S. Funken, D. Praetorius, and P. Wissgott. Efficient implementation of adaptive P1-FEM in Matlab. Comput. Methods Appl. Math., 11(4):460–490, 2011.
- [12] J. Gilewicz, V. Kryakin, and A. Shevchuk. Boundedness by 3 of the Whitney interpolation constant. J. Approx. Theory, 119(2):271–290, 2002.
- [13] D. Hermann, M. Frank, K. Busch, and P. Wölfle. Photonic band structure computations. Optics Express, 8(3):167–172, 2001.
- [14] Y. Hinuma, G. Pizzi, Y. Kumagai, F. Oba, and I. Tanaka. Band structure diagram paths based on crystallography. *Comput. Mater. Sci.*, 128:140–184, 2017.
- [15] D. Jackson. *Classical electrodynamics*. Wiley, New York, 1999.
- [16] J. Joannopoulos, S. Johnson, J. Winn, and R. Meade. Molding the flow of light. Princeton Univ. Press, Princeton, NJ, 2008.
- [17] H. Johnen and K. Scherer. On the equivalence of the k-functional and moduli of continuity and some applications. In Constructive Theory of Functions of Several Variables: Proceedings of a Conference Held at Oberwolfach April 25–May 1, 1976, pages 119–140. Springer, 2006.
- [18] M. Kadec. On the distribution of points of maximum deviation in the approximation of continuous functions by polynomials. Amer. Math. Soc. Transl. (2), 26:231–234, 1963.
- [19] B. Karaivanov and P. Petrushev. Nonlinear piecewise polynomial approximation beyond Besov spaces. *Appl. Comput. Harm. Anal.*, 15(3):177–223, 2003.
- [20] C. Kittel. Introduction to solid state physics. John Wiley & Sons, USA, 2018.
- [21] V. Kryakin. On exact constants in the Whitney theorem. Math. Notes, 54(1):688–700, 1993.
- [22] P. Kuchment. Floquet Theory for Partial Differential Equations. Birkhäuser Verlag, Basel, 1993.
- [23] D. Labilloy, H. Benisty, C. Weisbuch, T. Krauss, V. Bardinal, and U. Oesterle. Demonstration of cavity mode between two-dimensional photonic-crystal mirrors. *Electronics Letters*, 33(23):1978– 1980, 1997.
- [24] J. M. Melenk. hp-finite element methods for singular perturbations. Springer-Verlag, Berlin, 2002.

- [25] P.-O. Persson and G. Strang. A simple mesh generator in MATLAB. SIAM Rev., 46(2):329–345, 2004.
- [26] P. Russell. Photonic crystal fibers. *Science*, 299(5605):358–362, 2003.
- [27] C. Schwab. p- and hp-finite element methods. The Clarendon Press, Oxford University Press, New York, 1998. Theory and applications in solid and fluid mechanics.
- [28] B. H. Sendov and M. D. Takev. A theorem of Whitney's type in R<sup>2</sup>. Pliska Stud. Math. Bulgar., 11(1):78–85, 1991.
- [29] B. Sündermann. Lebesgue constants in Lagrangian interpolation at the Fekete points. Mitt. Math. Ges. Hamburg, 11(2):204–211, 1983.
- [30] A. Taylor, A. Wingate, and E. Vincent. An algorithm for computing Fekete points in the triangle. SIAM J. Numer. Anal., 38(5):1707–1720, 2000.
- [31] A. F. Timan. Theory of approximation of functions of a real variable. Dover Publications, Inc., New York, 1994. Translated from the Russian by J. Berry, Translation edited and with a preface by J. Cossar, Reprint of the 1963 English translation.
- [32] C. T. Traxler. An algorithm for adaptive mesh refinement in n dimensions. *Computing*, 59:115–137, 1997.
- [33] S. Wang, S. Li, and Y.-S. Wu. An analytical solution of pressure and displacement induced by recovery of poroelastic reservoirs and its applications. *SPE Journal*, 28(03):1329–1348, 2023.
- [34] Y. Wang and G. Li. Dispersion relation reconstruction for 2D photonic crystals based on polynomial interpolation. To appear, J.Comput. Phys., 2023.
- [35] R. Wilhelmsen. A Markov inequality in several dimensions. J. Approx. Theory, 11(3):216–220, 1974.
- [36] F. Yanik, S. Fan, M. Soljačić, and D. Joannopoulos. All-optical transistor action with bistable switching in a photonic crystal cross-waveguide geometry. *Optics letters*, 28(24):2506–2508, 2003.