ScienceDirect 上提供的目录《计算物理学杂志》期刊主页：
www.elsevier.com/locate/jcp



arXiv:2601.00199v1 [physics.optics] 1 Jan 2026

用于二维光子晶体正向设计与逆向设计的 POD–DeepONet 框架

Yueqi Wang[a]，Guanglian Li[c]，Guang Lin[a,b,*]

[a]普渡大学数学系，美国印第安纳州西拉斐特市 *Purdue Mall 610* 号，邮编 *47907*[b]普渡大学机械工程学院，美国印第安纳州西拉斐特市 *Purdue Mall 610* 号，邮编 *47907*[c]香港大学数学系，中国香港特别行政区薄扶林道

文章信息

文章历史：

关键词：POD–DeepONet、光子晶体、能带结构建模、布洛赫特征值问题、逆向设计。

摘　要

我们提出一种降阶算子学习框架，用于具有二值、基于像素的$p4m$对称晶胞的二维光子晶体的正向能带结构设计与逆向设计。我们将从高保真有限元能带快照中提取的POD主干网络，与预测降阶系数的神经分支网络相结合，构建沿标准高对称路径的离散能带映射的POD–DeepONet代理模型。该架构形成紧凑且可微的正向模型，并针对底层布洛赫特征值离散化进行了定制。我们证明了离散能带映射在松弛设计空间上的连续性，并证明POD–DeepONet代理模型具有一致逼近性质，因此总代理模型误差可自然分解为POD截断误差和网络逼近误差两部分。基于这一正向代理模型，我们提出两种端到端神经网络逆向设计流程，即从色散到结构的设计和带隙逆向设计；其训练目标结合了数据失配、二值化促进和监督正则化，以应对逆映射固有的非唯一性，并支持在松弛空间中进行稳定的基于梯度的优化。数值结果表明，所提出的框架能够实现准确的正向预测，并为实际的高对比度、基于像素的光子晶体布局生成有效的逆向设计。

© 2026 爱思唯尔公司。保留所有权利。

1. 引言

光子晶体（PhCs）是周期性结构的介电复合材料，能够调控电磁波的传播。自 Yablonovitch [1] 关于抑制自发辐射，以及 John [2]关于光子带隙材料中光局域的开创性研究以来，光子晶体已成为控制光的关键平台

*通讯作者：Guang Lin，电子邮箱：Guanglin@purdue.edu

在集成光子学中，光子晶体可用于实现具有定制色散特性的波导、腔、滤波器和激光器 [3, 4, 5, 6]。在频域设置下，其光谱特性由色散关系表征，即与麦克斯韦方程相关的参数化亥姆霍兹特征值问题所产生的能带函数族 $\{\bar{\omega}_n(\mathbf{k})\}_{n\geq 1}$。色散图中光子带隙的位置和宽度决定了慢光传输、限制和频率选择性等重要特性 [4, 5]。

从计算角度来看，色散关系是通过求解具有周期系数 [7]的参数化自伴复值亥姆霍兹特征值问题获得的。标准数值方法包括平面波展开法 [8], 、时域有限差分法（FDTD）格式 [9, 10], 以及有限元法（FEM） [11, 12]。因此，计算能带函数需要针对布里渊区内大量波矢求解这一亥姆霍兹特征值问题。为降低计算成本，大多数能带结构计算聚焦于经典高对称路径，即连接不可约布里渊区边界上高对称点的分段线性路径；这些路径已能揭示光子晶体应用中的带隙及其他显著谱特征 [4, 13, 14]。然而，即使在这一简化设定下，要获得高保真度的能带结构，仍需沿这些路径对 **k** 进行密集采样。因此，对于给定的光子晶体结构，仍须在细网格上求解大量广义亥姆霍兹特征值问题。在多查询场景中，这一计算成本尤为高昂：在高通量材料筛选和数据库构建过程中，必须针对数千种候选晶胞重复计算能带结构 [13, 15, 16, 17]。

在实际的光子晶体设计中，正向能带结构评估只是问题的一部分，而逆向设计往往是核心目标。根据应用需求，可以寻找一种周期微结构，使其能带结构尽可能接近给定的色散关系（色散到结构问题）；也可以优化几何结构，以最大化选定能带之间目标频率范围内的完全带隙（带隙设计）。经典方法将这些任务表述为偏微分方程约束优化问题，并通过拓扑优化求解，通常采用基于梯度的方法 [18, 19, 20] 或非梯度方案 [21, 22], , 以及用于最大化带隙的相关水平集公式 [23, 24]。这些算法从启发式初始猜测出发，反复求解大型特征值问题，在许多**k**点上评估目标函数和灵敏度，因此每次设计更新的计算成本都很高。由此，高维像素或水平集设计空间会带来巨大的计算成本，进而促使人们开发降阶和数据驱动的代理模型，尤其是基于深度学习的代理模型。此类模型能够以低得多的在线计算复杂度近似正向布洛赫映射，同时为逆向设计保留足够的精度。

近些年来，深度学习已成为建模超材料和光子晶体（PhCs）结构–性质映射的强大范式。当前方法包括使用全连接深度神经网络、ResNet 风格架构或卷积神经网络（CNNs），将几何形状或像素化的单胞映射到散射光谱或透射响应；随后，为实现逆向设计 [25, 26, 27, 28]，通过前向网络进行反向传播，或训练一个单独的网络，将目标响应映射为设计。自编码器经训练后可从单胞样本图像中提取拓扑特征，随后训练多层感知机（MLP），以建立带隙与拓扑特征之间的内在关系 [29]。后续研究采用变分自编码器或条件自编码器（VAE / cVAE）以及串联网络，根据带隙目标生成单胞 [30, 31, 32, 33, 34]。关于基于 CNN、MLP 和生成模型的超结构及光子晶体设计方法的最新综述，请参阅 [35, 36, 37]。大多数现有代理模型仍将能带结构视为高维黑箱向量：首先在布里渊区固定网格中选取的**k**点处计算拼接后的能带频率列表，然后由神经网络直接回归这些采样值。与此同时，算子学习已成为参数化偏微分方程数据驱动代理建模的一种强大框架。深度算子网络（DeepONets）并非近似有限维输入–输出映射，而是近似非线性算子；这类算子以系数、源项或边界条件等函数值输入为输入，并返回解场 [38, 39]。在适当假设下，DeepONets 满足算子的通用逼近定理 [38]。训练完成后，它们能够以极低的成本对新输入求值，同时保持对这些输入的可微性；这一特性对于基于梯度的逆问题尤为有吸引力。为提升高维情形下的效率和稳健性，多项研究将 DeepONets 与基于投影或多保真度的模型降阶方法相结合，包括 POD增强型深度算子网络框架 [40, 41, 42] 以及用于残差学习的多保真度深度算子网络架构 [43, 44]。这些方法利用了一个经验观察：在许多参数化偏微分方程中，解流形接近一个低维子空间。例如，可以使用本征正交分解（POD）得到的降阶基来表示输出场，同时由神经网络学习相应降阶系数如何

取决于输入。

据我们所知，降阶算子学习代理模型尚未应用于光子晶体中的布洛赫特征值问题或能带结构计算。其他参数化偏微分方程场景中的成功应用表明，这种方法既能加速能带结构评估，又能为模型施加透明的降阶结构。在本研究中，我们探讨一个正向问题和两个反问题：(i) 预测给定光子结构的能带函数；(ii) 以指定的能带函数为目标进行色散到结构设计；(iii) 以指定的带隙描述符为目标进行带隙设计。我们的正向问题可写为一个算子

$$\mathcal{G}:\ \epsilon \longmapsto (\widetilde{\omega}_1(\cdot;\epsilon),\dots,\widetilde{\omega}_{N_b}(\cdot;\epsilon)), \tag{1.1}$$

其中，ϵ 表示光子晶体的周期性介电常数分布，$\mathcal{G}$返回沿布里渊区高对称路径的前N_b个能带函数 $\widetilde{\omega}_n(\mathbf{k};\epsilon)$。第一个反问题，即色散到结构设计，旨在近似反演该算子，

$$\mathcal{I}_{\text{disp}}:\ (\widetilde{\omega}_1^*(\cdot),\dots,\widetilde{\omega}_{N_b}^*(\cdot)) \longmapsto \epsilon, \tag{1.2}$$

给定目标能带函数 $\widetilde{\omega}_n^*(\mathbf{k})$。第二个反问题考虑带隙描述符 $\mathbf{g} = (a,b,p)$，可表示为

$$\mathcal{I}_{\text{gap}}:\ \mathbf{g}^* \longmapsto \epsilon, \tag{1.3}$$

其中，$\mathbf{g}^* = (a^*,b^*,p^*)$表示目标带隙，即$a^*$ 和b^* 分别为带隙的下边缘和上边缘，而p^* 为带隙下方能带的索引。

我们引入晶胞的像素化表示，将介电分布编码为固定网格上的分片常数值。在这一离散设计空间上，我们对式 (1.1)–(1.3) 中的连续算子进行近似。具体而言，我们通过 POD–DeepONet代理模型 $\mathcal{G}$ 近似正向算子。每个能带函数$\widetilde{\omega}_n(\cdot;\epsilon)$均在采样波矢上展开为本征正交分解基。主干网络在选定的$\mathbf{k}$点处计算这一固定的本征正交分解基，而分支网络则将像素化微结构映射为相应的POD系数。随后，在训练得到的POD–DeepONet代理模型基础上，通过神经网络实现这两个反向算子。我们的主要贡献总结如下：

- **统一的算子学习 formulation**。 我们提出了一个用于二维光子晶体能带结构预测和两个逆向设计任务的统一算子学习框架。
- **误差分析**。 我们对所学习能带函数的总代理模型误差进行了明确分解，将其分为 POD 截断误差和网络逼近误差。
- **高效的端到端逆向设计，仅需极少的有限元调用**。 我们通过对可微代理模型进行反向传播，实现了快速的色散匹配和带隙目标设定，从而显著减少了重复有限元特征值求解的需求。

本文其余部分的组织如下。第 2 节对光子晶体的能带结构问题进行了表述。我们建立了有限元框架，并引入了晶胞的基于像素的参数化，从而对本文研究的正问题和反问题给出精确表述。第 3 节构建了用于正向能带映射的 POD–DeepONet 代理模型，即将像素设计向量映射为沿高对称路径采样的能带函数。我们介绍了主干网络的快照POD 构造以及分支网络的监督训练。随后，我们分析了逼近误差，以区分 POD 截断误差与神经网络误差的贡献。第 4 节通过将正向代理模型与用于色散到结构和带隙问题的逆向网络相结合，转向逆向设计。第 5 节给出了数值实验，展示了所提出方法的准确性和效率；第 6 节则以简要总结和展望作结。

2. 问题表述

本节简要回顾计算色散关系的数学框架，并构建正向与逆向频带结构问题，这些问题构成了本文的主要研究重点。

2.1. 色散关系计算

为明确符号约定，我们简要回顾从时谐麦克斯韦系统到参数化亥姆霍兹特征值问题的标准约化过程，该特征值问题描述二维（2D）光子晶体中的布洛赫模。推导过程沿用我们先前的工作 [45, 46]；背景介绍参见例如 [47, 48] 。

在国际单位制约定下，具有自由电荷和电流的线性、无色散、非磁性介质中的时谐麦克斯韦方程为

$$\nabla \times \mathbf{E}(\mathbf{x}) - i\omega\mu_0\mathbf{H}(\mathbf{x}) = 0, \tag{2.1a}$$
$$\nabla \times \mathbf{H}(\mathbf{x}) + i\omega\epsilon_0\epsilon(\mathbf{x})\mathbf{E}(\mathbf{x}) = 0, \tag{2.1b}$$
$$\nabla \cdot (\epsilon(\mathbf{x})\mathbf{E}(\mathbf{x})) = 0, \tag{2.1c}$$
$$\nabla \cdot \mathbf{H}(\mathbf{x}) = 0, \tag{2.1d}$$

其中，$\mathbf{x} \in \mathbb{R}^3$、$\mathbf{E}$ 和$\mathbf{H}$ 分别表示电场和磁场，$\omega \geq 0$ 为角频率，μ_0 和 ϵ_0 分别为真空磁导率和介电常数，$\epsilon \in L^\infty(\mathbb{R}^3;\mathbb{R}^+)$为相对介电常数。从（2.1a）–（2.1b）中消去$\mathbf{E}$ 或$\mathbf{H}$，可得到熟知的旋度–旋度形式

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

and

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

其中 $\epsilon_0\mu_0 = c^{-2}$，c为光速。

我们将研究范围限定为二维光子晶体，其沿z方向保持不变，并在x–y平面内具有周期性。因此，$\epsilon(\mathbf{x})$ 取为与z无关。在此假设下，场可分解为横电（TE）偏振，其中$H_1 = H_2 = E_3 = 0$，以及横磁（TM）偏振，其中$E_1 = E_2 = H_3 = 0$。将这些假设代入（2.2）–（2.3），可在 $\mathbb{R}^2$上得到标量亥姆霍兹特征值问题，

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

二维光子晶体在x–y 平面内具有离散平移对称性 [4], 因此，相对介电常数满足

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

其中，原始晶格矢量$\mathbf{a}_1$,$\mathbf{a}_2$ 张成一个基本周期域 Ω（晶胞）。倒格矢$\mathbf{b}_1$,$\mathbf{b}_2$ 定义为

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

并生成倒格子。其基本晶胞为（第一）布里渊区 $\mathcal{B}_F$，可将其刻画为倒易空间中距原点比距任何其他倒格子点都近的点集。在本文中，我们考虑原始矢量为$\mathbf{a}_i = a\mathbf{e}_i$（$i = 1 = 1, 2$）的正方晶格，其中（$\mathbf{e}_i$）$_{i=1,2}$ 是$\mathbb{R}^2$的标准基，$a > 0$为晶格常数。相应的倒格矢为$\mathbf{b}_i = \frac{2\pi}{a}\mathbf{e}_i$。

根据布洛赫定理 [49], , 式 (2.4)–(2.5) 的解可写为

$$\Psi(\mathbf{x}) = e^{i\mathbf{k}\cdot\mathbf{x}}u(\mathbf{x}),$$

其中，波矢$\mathbf{k}$ 位于 $\mathcal{B}_F$ 中，布洛赫因子u 关于晶格具有周期性。特别地，式 (2.4)–(2.5) 中的标量场可表示为

$$H(\mathbf{x}) = e^{i\mathbf{k}\cdot\mathbf{x}}u_1(\mathbf{x}), \qquad E(\mathbf{x}) = e^{i\mathbf{k}\cdot\mathbf{x}}u_2(\mathbf{x}),$$

其中，周期函数u_1,u_2 定义在晶胞 Ω 上。将其代入 (2.4)–(2.5)，可得定义在 Ω 上的参数化亥姆霍兹特征值问题，

$$-(\nabla + i\mathbf{k}) \cdot (\epsilon(\mathbf{x})^{-1}(\nabla + i\mathbf{k})u_1(\mathbf{x})) - (\omega c^{-1})^2 u_1(\mathbf{x}) = 0, \quad \mathbf{x} \in \Omega, \qquad \textbf{(TE 模式)}, \tag{2.7a}$$
$$-(\nabla + i\mathbf{k}) \cdot ((\nabla + i\mathbf{k})u_2(\mathbf{x})) - (\omega c^{-1})^2\epsilon(\mathbf{x})u_2(\mathbf{x}) = 0, \quad \mathbf{x} \in \Omega, \qquad \textbf{(TM 模式)}, \tag{2.7b}$$

其中$\mathbf{k}$ 在布里渊区内变化，布洛赫因子u_i 满足周期边界条件$u_i(\mathbf{x}) = u_i(\mathbf{x} + \mathbf{a}_j)$，其中$i$、$j = 1$取值为 2。若 ϵ 具有额外的点群对称性（例如镜面对称性），则波矢还可进一步限制在不可约布里渊区（IBZ）内，记为 $\mathcal{B} \subset \mathcal{B}_F$。正方晶格及其布里渊区的示例见图1。

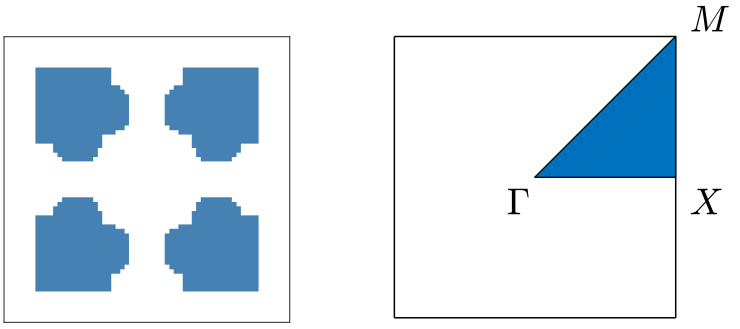


图 1. 方晶格晶胞 Ω （左）及其对应的第一布里渊区 $\mathcal{B}_F$ （右）示意图。在 Ω 中，蓝色表示介电常数为 8.9 的氧化铝，白色表示介电常数为 1 的空气；在 $\mathcal{B}_F$ 中，IBZ $\mathcal{B}$ 是顶点为 $\Gamma = (0,0)$、$X = (\pi/a, 0)$ 和 $M = (\pi/a, \pi/a)$ 的阴影三角形。

Both 参数化亥姆霍兹问题 (2.7a)–(2.7b) 可表示为统一形式

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

其中 $\Omega \subset \mathbb{R}^2$、$\mathbf{k} \in \mathcal{B}$ 和 $\lambda = (\omega c^{-1})^2$。在 TE 模式中，$u$ 表示 z 方向的磁场 H，系数为 $\alpha(\mathbf{x}) := \epsilon(\mathbf{x})^{-1}$ 和 $\beta(\mathbf{x}) := 1$；在 TM 模式中，u 表示 z 方向的电场 E，且$\alpha(\mathbf{x}) := 1$和 $\beta(\mathbf{x}) := \epsilon(\mathbf{x})$。

(2.8)的变分表述如下：对于每个$\mathbf{k} \in \mathcal{B}$，求一个非平凡特征对$(\lambda, u) \in \mathbb{R} \times H^1_\pi(\Omega)$，使得

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

其中，

$$L^2_\beta(\Omega) := \left\{ f \in L^2(\Omega;\mathbb{C}) \ : \ \|f\|^2_{L^2_\beta(\Omega)} := \int_\Omega \beta(\mathbf{x})|f(\mathbf{x})|^2\,\mathrm{d}\mathbf{x} < \infty \right\},$$

且周期索об列夫空间$H^1_\pi(\Omega) \subset H^1(\Omega;\mathbb{C})$定义为

$$H^1_\pi(\Omega) := \{v \in H^1(\Omega;\mathbb{C}) \ : \ v \text{ is periodic with respect to the lattice vectors of } \Omega\}.$$

这里，$H^1(\Omega;\mathbb{C})$表示由平方可积的复值函数及其平方可积弱梯度构成的索об列夫空间，并配备标准的H^1–范数。

以下结果是紧预解式二阶椭圆算子谱理论的一个标准推论 [50]。

定理2.1. 对于每个波矢$\mathbf{k} \in \mathcal{B}$，变分特征值问题(2.9) 在$H^1_\pi(\Omega)$ 上定义了一个自伴算子，其具有紧预解式。其频谱纯离散且非负，特征值可按非递减顺序排列（按有限重数重复），

$$0 \le \lambda_1(\mathbf{k}) \le \lambda_2(\mathbf{k}) \le \cdots \le \lambda_n(\mathbf{k}) \le \cdots \to +\infty.$$

此外，对于每个固定的 $n \in \mathbb{N}$，函数$\mathbf{k} \mapsto \lambda_n(\mathbf{k})$在 $\mathcal{B}$ 上连续，并且 $\lambda_n(\mathbf{k}) \to +\infty$ 在 $n \to \infty$ 时关于 $\mathcal{B}$ 一致。

布洛赫频率与特征值之间的关系为

$$\omega_n(\mathbf{k}) = c\,\sqrt{\lambda_n(\mathbf{k})},$$

并引入无量纲（归一化）能带函数

$$\widetilde{\omega}_n(\mathbf{k}) := \frac{a}{2\pi c}\,\omega_n(\mathbf{k}).$$

族 $\{\widetilde{\omega}_n(\cdot)\}_{n\geq 1}$ 称为光子晶体的色散关系。光子带隙是一个开区间 (ω_-, ω_+)，满足

$$\widetilde{\omega}_n(\mathbf{k}) \notin (\omega_-, \omega_+) \qquad \text{for all } \mathbf{k} \in \mathcal{B} \text{ and all } n \geq 1,$$

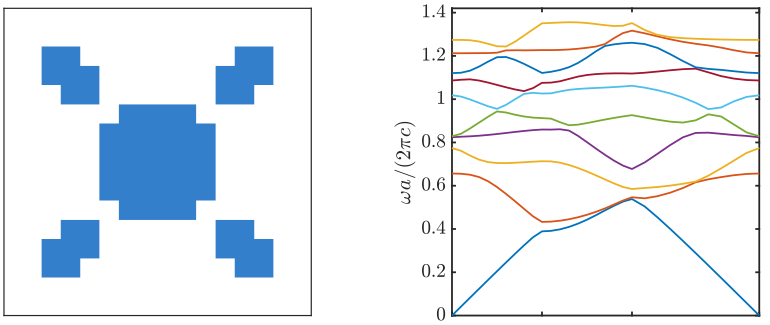
也就是说，在此频率范围内不存在布洛赫特征模 $(\mathbf{k}, \widetilde{\omega})$。因此，色散关系确定了所有允许的布洛赫模，尤其确定了光子带隙的位置和宽度 [51, 52]。

在大多数应用中，人们主要关注最低的几个能带函数，因为波导、腔和滤波器等光子器件通常在相对较低的频率下工作 [53, 6, 54], ，而最宽且最实用的带隙往往出现在前几个能带之间 [51, 52]。此外，实际的带隙工程通常依赖高对比度介电常数分布，因为较大的折射率对比度会增强布拉格散射，通常也会带来更强的场约束和更宽的带隙 [51, 52]。此外，利用晶格的点群对称性，将波矢 $\mathbf{k}$ 限制在不可约布里渊区内，并且仅沿连接其边界上高对称点的分段线性路径绘制 $\widetilde{\omega}_n(\mathbf{k})$，已成为标准做法（对于图1中的正方晶格，路径为 $\Gamma \to X \to M \to \Gamma$）。这样既大幅减少了采样的 $\mathbf{k}$–点数量，又能捕捉能带边缘、带隙及其他关键谱特征 [49, 7, 55, 13]。

令 $\mathcal{K}_{\mathrm{hs}} \subset \mathcal{B}$ 表示正方晶格不可约布里渊区边界上的这条标准高对称路径，定义如下：

$$\mathcal{K}_{\mathrm{hs}} := \overline{\Gamma X} \cup \overline{XM} \cup \overline{M\Gamma},$$

因此，我们关注沿高对称路径 $\mathcal{K}_{\mathrm{hs}} \subset \mathcal{B}$ 计算前 N_b 个能带函数 $\{\widetilde{\omega}_n(\mathbf{k})\}_{n=1}^{N_b}$。图 2 给出了一个示例，其中 2(a) 展示了一个晶胞，2(b) 展示了沿高对称路径 $\mathcal{K}_{\mathrm{hs}}$ 的相应归一化能带函数 $\widetilde{\omega}_n(\mathbf{k})$。



(a) 晶胞　　(b) 能带结构

图 2. 晶胞及其能带结构示例：(a) 16×16 **晶胞；(b) 沿高对称路径** $\mathcal{K}_{\mathrm{hs}}$ **的前** 10 **个 TE能带函数。**

2.2. 有限元离散

通过引入半双线性型

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

我们可以将问题 (2.9) 改写为紧凑形式：对于每个 $\mathbf{k} \in \mathcal{K}_{\mathrm{hs}}$，求一个非平凡特征对 $(\lambda, u) \in \mathbb{R}\times H^1_\pi(\Omega)$，使得

$$\begin{cases} a(u,v) = \lambda\, b(u,v), \quad \forall\, v \in H^1_\pi(\Omega), \\ \|u\|_{L^2_\beta(\Omega)} = 1. \end{cases} \tag{2.10}$$

设 $\mathcal{T}_h$ 为晶胞 Ω 的形状规则、协调且在相对面上周期的三角剖分。我们考虑协调、周期的 P_1 空间

$$V_h \;:=\; \{v_h \in H^1_\pi(\Omega) :\; v_h|_T \in \mathbb{P}_1(T) \;\; \forall\, T \in \mathcal{T}_h\},$$

并将其节点基记为 $\{\phi_i\}_{i=1}^{N_h} \subset V_h$。

离散特征值　问题表述如下：对每个 $\mathbf{k} \in \mathcal{K}_{\mathrm{hs}}$，求非平凡特征对 $(\lambda_h, u_h) \in (\mathbb{R}, V_h)$，使得

$$\begin{cases} a(u_h, v_h) = \lambda_h\, b(u_h, v_h), \quad \forall\, v_h \in V_h \\ b(u_h, u_h) = 1. \end{cases} \tag{2.11}$$

令 $u_h = \sum_{j=1}^{N_h} U_j \phi_j$，并用 $v_h = \phi_i$ 对 (2.11) 进行测试，可得广义厄米矩阵特征值问题

$$\mathbf{A}(\mathbf{k})\,\mathbf{U} \;=\; \lambda_h\, \mathbf{B}\,\mathbf{U}, \qquad \mathbf{U}^{*}\mathbf{B}\mathbf{U} = 1, \tag{2.12}$$

其中 $\mathbf{U} = (U_1, \ldots, U_{N_h})^{\top} \in \mathbb{C}^{N_h}$，且

$$\mathbf{A}_{ij}(\mathbf{k}) := \int_\Omega \alpha(\mathbf{x})\,(\nabla\phi_i + i\mathbf{k}\,\phi_i)\cdot(\nabla\phi_j - i\mathbf{k}\,\phi_j)\,\mathrm{d}\mathbf{x},$$
$$\mathbf{B}_{ij} := \int_\Omega \beta(\mathbf{x})\,\phi_i\,\phi_j\,\mathrm{d}\mathbf{x}.$$

给定波矢 $\mathbf{k}$，求解 (2.12) 可得特征对 $\{(\lambda_{h,n}(\mathbf{k}), \mathbf{U}_n(\mathbf{k}))\}_{n\ge 1}$，并按非递减顺序排列。我们将归一化能带函数定义为

$$\widetilde{\omega}_{h,n}(\mathbf{k}) := \frac{a}{2\pi c}\,\omega_{h,n}(\mathbf{k}) = \frac{a}{2\pi}\,\sqrt{\lambda_{h,n}(\mathbf{k})}.$$

The 然后，通过沿高对称路径 $\mathcal{K}$ 计算 $\widetilde{\omega}_{h,n}$，对能带结构进行近似 $_{\mathrm{hs}}$.

2.3. 晶胞介电常数的基于像素的参数化

接下来，我们引入具有 $p4m$ 平面对称性 [56] 的二值、分段常数晶胞介电常数。利用这种对称性，只需指定晶胞中一个经过对称性约化的子区域内的材料分布。我们将图 3 右上面板中的阶梯状三角形区域称为不可约对称楔形区域，以下简称楔形区域。

令 N_f 表示该楔形区域中的像素数量。我们引入一个二值设计向量

$$\boldsymbol{\rho} = (\rho_1, \ldots, \rho_{N_f})^{\top} \in \{0,1\}^{N_f},$$

其中，$\rho_j = 1$ 表示像素 j 填充了高介电常数材料，而 $\rho_j = 0$ 对应低介电常数背景。为便于后续使用，我们还定义了一个提升算子

$$E^{\mathrm{lift}} : [0,1]^{N_f} \longrightarrow [0,1]^{N_{\mathrm{pix}}}, \qquad \boldsymbol{\rho} \longmapsto E^{\mathrm{lift}}(\boldsymbol{\rho}) := (\tilde{\rho}_1, \ldots, \tilde{\rho}_{N_{\mathrm{pix}}})^{\top}, \tag{2.13}$$

该算子通过应用 $p4m$ 对称操作，将楔形设计向量 $\boldsymbol{\rho}$ 映射为完整晶胞上的像素向量。这里，N_{pix} 表示完整晶胞中的像素数量。

令 $\{P_j^{\mathrm{w}}\}_{j=1}^{N_f}$ 表示楔形区域中的像素子域。每个 P_j^{w} 都是图 3 所示的小方形像素之一。对这些楔形像素应用 $p4m$ 旋转和反射，可得到一组完整晶胞像素 $\{P_\ell\}_{\mathrm{l}=1}^{N_{\mathrm{pix}}}$，它们构成 Ω 的一个划分。相应的像素指示场为

$$\chi_{\boldsymbol{\rho}}(\mathbf{x}) := \sum_{\ell=1}^{N_{\mathrm{pix}}} \tilde{\rho}_\ell\, \mathbf{1}_{P_\ell}(\mathbf{x}), \qquad \mathbf{x} \in \Omega,$$

其中 $\mathbf{1}_{P_\ell}$ 是P_ℓ的特征函数。对于二值设计，有 $\tilde{\rho}_\ell \in \{0,1\}$，因此 $\chi_{\boldsymbol{\rho}}(\mathbf{x}) \in \{0,1\}$ 在 Ω 中几乎处处成立，其中 $\chi_{\boldsymbol{\rho}} = 1$ 位于高介电常数区域，$\chi_{\boldsymbol{\rho}} = 0$ 位于低介电常数背景中。

在所有数值示例中，我们考虑如下形式的二元复合材料

$$\epsilon(\mathbf{x}) \in \{\epsilon_{\text{air}}, \epsilon_{\text{alum}}\}, \qquad \mathbf{x} \in \Omega,$$

其中 $\epsilon_{\text{air}} = 1$ 和 $\epsilon_{\text{alum}} = 8。9$；见图1(a)。给定设计 $\boldsymbol{\rho}$，完整晶胞上的相应相对介电常数场为

$$\epsilon(\mathbf{x};\boldsymbol{\rho}) := \epsilon_{\text{air}} + (\epsilon_{\text{alum}} - \epsilon_{\text{air}})\chi_{\boldsymbol{\rho}}(\mathbf{x}), \qquad \mathbf{x} \in \Omega. \tag{2.14}$$

这定义了基于像素的可行集

$$\mathcal{E}_{\text{pix}} := \{\epsilon(\cdot;\boldsymbol{\rho}) : \boldsymbol{\rho} \in \{0,1\}^{N_f}\} \subset L^\infty(\Omega).$$

数值实验中的所有训练样本和验证样本均取自 $\mathcal{E}_{\text{pix}}$.

出于分析和基于梯度的优化的目的，我们还考虑设计变量的连续松弛。在这种松弛设定下，我们允许

$$\boldsymbol{\rho} \in [0,1]^{N_f},$$

同时保持定义 如上所述的$E^{\text{lift}}, \chi_{\boldsymbol{\rho}}$ 和 $\epsilon(\cdot;\boldsymbol{\rho})$。相应的松弛介电常数族 is

$$\mathcal{E}_{\text{rel}} := \{\epsilon(\cdot;\boldsymbol{\rho}) : \boldsymbol{\rho} \in [0,1]^{N_f}\}.$$

显然，我们有 $\mathcal{E}_{\text{pix}} \subset \mathcal{E}_{\text{rel}}$。在逆向设计公式中，我们在这个连续空间上进行优化，以便采用基于梯度的方法，然后将所得的松弛设计投影回二元构型。

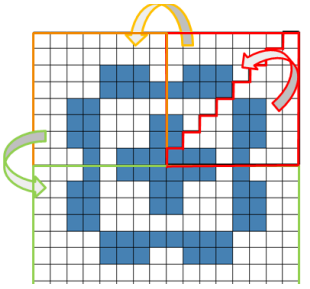


图3. 具有$p4m$平面对称性的晶胞的基于像素的参数化。蓝色像素表示介电常数为 ϵ_{alum} 的高介电常数材料；白色像素表示低介电常数背景 ϵ_{air}。红色三角形标示基本楔形区域。该区域与像素网格的交集包含 $N_f = 36$ 个楔形像素 $\{P_j^{\text{w}}\}_{j=1}^{N_f}$，每个像素由一个二元设计变量 $\rho_j \in \{0,1\}$ 控制。应用$p4m$旋转和反射（由彩色箭头标示），可将这些楔形像素映射到晶胞中的全部像素。

备注 2.2 (基于像素的参数化的一般性)。 晶胞中任何具有 $p4m$ 对称性的可测双相介电常数布局，都可以用一系列二值、分段常数的像素设计近似，因此，设计空间 $\mathcal{E}_{\text{pix}}$ 足以涵盖本文所研究的逆问题，并且仍与标准的基于像素的拓扑优化技术兼容。

2.4. 正向与逆向能带结构问题

现在，我们在离散情形下表述正向与逆向能带结构问题，并将其与算子映射 (1.1)–(1.3) 联系起来。

通过第 2.3 节中的基于像素的参数化，每个设计向量 $\boldsymbol{\rho} \in [0,1]^{N_f}$ 都会生成一个介电常数场 $\epsilon(\cdot;\boldsymbol{\rho}) \in \mathcal{E}_{\text{rel}}$。对于任意 $\boldsymbol{\rho} \in [0,1]^{N_f}$和任意波矢 $\mathbf{k} \in \mathcal{K}_{\text{hs}}$，离散特征值问题 (2.11) 得到一个非递减的特征值序列

$$0 < \lambda_{h,1}(\mathbf{k};\boldsymbol{\rho}) \le \lambda_{h,2}(\mathbf{k};\boldsymbol{\rho}) \le \cdots,$$

及其对应的特征向量 $\mathbf{U}_{h,n}(\mathbf{k};\boldsymbol{\rho})$。相应的归一化能带函数为

$$\widetilde{\omega}_{h,n}(\mathbf{k};\boldsymbol{\rho}) := \frac{a}{2\pi}\sqrt{\lambda_{h,n}(\mathbf{k};\boldsymbol{\rho})}, \qquad n \geq 1.$$

实际上，我们在有限组波矢上对能带结构进行采样

$$\{\mathbf{k}_\ell\}_{\ell=1}^{N_k} \subset \mathcal{K}_{\mathrm{hs}},$$

a并仅保留前N_b 条能带。随后，我们使用截断并离散化后的能带数据

$$\mathbf{W}_h(\boldsymbol{\rho}) := (\widetilde{\omega}_{h,n}(\mathbf{k}_\ell;\boldsymbol{\rho}))_{\substack{1\leq \ell\leq N_k\\ 1\leq n\leq N_b}} \in \mathbb{R}^{N_k\times N_b}.$$

这定义了设计空间上的离散能带映射

$$F_h^{\mathrm{pix}} : [0,1]^{N_f} \longrightarrow \mathbb{R}^{N_k\times N_b}, \qquad F_h^{\mathrm{pix}}(\boldsymbol{\rho}) := \mathbf{W}_h(\boldsymbol{\rho}), \tag{2.15}$$

它是式 (1.1) 中正向算子的基于像素的实现 $\mathcal{G}$。

正向问题

对于给定的设计向量 $\boldsymbol{\rho} \in [0,1]^{N_f}$，或等价地，给定周期性介电常数场 $\epsilon(\cdot;\boldsymbol{\rho})$，正向能带结构问题旨在计算沿高对称路径的能带函数。准确且高效地评估这一映射，对于表征候选单胞的色散关系，以及生成逆问题训练和目标函数所需的能带数据至关重要。在离散情形下，正向问题可归结为对映射 (2.15) 的评估。调用一次$F^{\mathrm{pix}}{}_h$，就需要求解N_k 个形如 (2.12) 的大规模矩阵特征值问题。

色散到结构的逆问题

在许多光子设计任务中，我们从目标色散关系出发，希望对其进行“反演”，也就是说，恢复一种周期微结构，使其能带图尽可能贴合目标。例如，基于期望的群速度或带隙位置等物理考量指定能带图，或在更高层次的建模中计算能带图时，就会出现这种情况；此时，人们希望找到一种可实现的晶胞来重现这种行为 [18, 5, 57, 17]。在连续层面，这对应于式 (1.2) 中的色散到结构算子 $\mathcal{I}_{\mathrm{disp}}$，该算子将一组目标能带函数映射为适当的介电常数分布。

设$\mathbf{W}^{\mathrm{targ}}{}_h \in \mathbb{R}^{N_k\times N^b}$ 为给定的目标能带结构，表示式 (1.2) 中所需能带函数 $(\widetilde{\omega}_1^*(\cdot),\ldots,\widetilde{\omega}_{N_b}^*(\cdot))$ 的离散采样值。我们寻求一种设计 $\boldsymbol{\rho} \in [0,1]^{N_f}$，使$\mathbf{W}_h(\boldsymbol{\rho})$ 接近$\mathbf{W}^{\mathrm{targ}}{}_h$。在离散层面，我们考虑以下优化问题

$$\min_{\boldsymbol{\rho}\in[0,1]^{N_f}} J_{\mathrm{disp}}(\boldsymbol{\rho}) := \frac{1}{N_k N_b}\left\|\mathbf{W}_h(\boldsymbol{\rho}) - \mathbf{W}_h^{\mathrm{targ}}\right\|_F^2, \tag{2.16}$$

并满足额外约束，例如强制材料分布为二值。

逆带隙设计问题

在许多应用中，例如设计波导和腔体 [1, 8, 18, 58, 19], ，主要目标并非重现整个色散关系，而是确保在指定频率范围内存在完整光子带隙。在这种情况下，设计目标是确定带隙的位置。

设 $1 \leq p < N_b$ 为指定的能带索引，并设$I_{\mathrm{gap}} = (a,b) \subset \mathbb{R}_+$ 为目标频率区间。我们寻求满足以下条件的设计 $\boldsymbol{\rho} \in [0,1]^{N_f}$，使第$p$条和第 $(p+1)$ 条能带满足

$$\widetilde{\omega}_{h,p}(\mathbf{k};\boldsymbol{\rho}) \leq a, \qquad \widetilde{\omega}_{h,p+1}(\mathbf{k};\boldsymbol{\rho}) \geq b, \qquad \forall\, \mathbf{k} \in \mathcal{K}_{\mathrm{hs}},$$

使得目标区间 I_{gap} 位于沿整个高对称性路径的能带 p 与$p+1$ 之间的完全带隙内。在离散情形下，这一要求变为

$$\widetilde{\omega}_{h,p}(\mathbf{k}_\ell;\boldsymbol{\rho}) \leq a, \qquad \widetilde{\omega}_{h,p+1}(\mathbf{k}_\ell;\boldsymbol{\rho}) \geq b, \qquad \ell = 1,\ldots,N_k. \tag{2.17}$$

就式 (1.3) 中引入的带隙描述符 $\mathbf{g} = (a,b,p)$ 而言，该反问题对应于算子$\mathcal{I}_{\mathrm{gap}}$，它将目标描述符 $\mathbf{g}$映射为一个可行设计。

我们通过最小化惩罚泛函 $J_{\text{gap}}(\boldsymbol{\rho})$ 来施加条件 (2.17)，该泛函用于衡量对这些条件的违反程度。其抽象形式为

$$\min_{\boldsymbol{\rho}\in[0,1]^{N_f}} \; J_{\text{gap}}(\boldsymbol{\rho}) := \frac{1}{N_k}\sum_{\ell=1}^{N_k} \Phi_{\text{gap}}(\widetilde{\omega}_{h,p}(\mathbf{k}_\ell;\boldsymbol{\rho}), \widetilde{\omega}_{h,p+1}(\mathbf{k}_\ell;\boldsymbol{\rho}), a, b), \tag{2.18}$$

其中 $\Phi_{\text{gap}}:\ \mathbb{R}^4 \to \mathbb{R}_+$ 是一个连续惩罚函数，使得

$$\Phi_{\text{gap}}(x,y,a,b) = 0 \quad \Longleftrightarrow \quad x \le a \text{ and } y \ge b.$$

我们还进一步采用正则化项，以抑制灰度中间值 .

直接通过反复进行有限元特征值求解来求解逆优化问题 (2.16) 和 (2.18)，代价将高得难以承受。因此，在本文余下部分，我们用 POD–DeepONet代理模型替代离散前向映射 $F^{\text{pix}}{}_h$，并将其与用于参数化可行设计的神经网络结合起来。由此得到的端到端可微模型使我们能够通过标准的基于梯度的训练来求解逆优化问题。

注 2.3（松弛优化与建模视角）。 逆问题 (2.16) 和 (2.18) 定义在松弛设计域 $[0,1]^{N_f}$ 上。采用这一选择可使目标泛函 J_{disp} 和 J_{gap} 对 $\boldsymbol{\rho}$ 可微，并允许使用基于梯度的算法。然而，物理上可实现的光子晶体仅采用两种材料相，其介电常数分别为 ϵ_{air} 和 ϵ_{alum}。因此，在求解松弛优化问题后，我们应用 0–1 投影，以获得二值设计 $\boldsymbol{\rho}^{\text{bin}} \in \{0,1\}^{N_f}$；详情见第 *4* 节。

3. POD–DeepONet 正向能带结构映射

本节中，我们为式 (2.15) 中定义的离散能带映射构建一个 POD–DeepONet 代理模型

$$F_h^{\text{pix}} : [0,1]^{N_f} \to \mathbb{R}^{N_k\times N_b}, \qquad \boldsymbol{\rho} \longmapsto \mathbf{W}_h(\boldsymbol{\rho}),$$

该代理模型遵循 DeepONet 的主干网络–分支结构。波矢$\mathbf{k}$ 的依赖关系由固定的低维 POD 基（即主干网络）表示，而设计向量 $\boldsymbol{\rho}$的依赖关系则由神经网络（即分支）学习。主干网络基由高保真频带数据计算得到，并在训练过程中保持固定。此设计利用了布洛赫算子的谱结构，在波矢变量上生成紧凑的低秩表示，并与通用的高维网络参数化相比，减少了有效可训练参数的数量。

3.1. 通过快照*POD*构建主干网络

我们首先通过快照本征正交分解（POD）构建固定的 POD 主干网络基。

设 $\{\boldsymbol{\rho}^{(i)}\}_{i=1}^{N_{\text{train}}} \subset \{0,1\}^{N_f}$ 为一组设计向量，$\epsilon^{(i)} := \epsilon(\cdot;\boldsymbol{\rho}^{(i)}) \in \mathcal{E}_{\text{pix}}$ 为式 (2.14) 中定义的对应分段常数介电常数。对于每个 $\boldsymbol{\rho}^{(i)}$，我们在所有波矢$\{\mathbf{k}_\ell\}_{\ell=1}^{N_k} \subset \mathcal{K}_{\text{hs}}$处求解离散特征值问题 (2.12)，然后保留前$N_b$ 个归一化能带，并组装截断能带数据

$$\mathbf{W}_h^{(i)} := \mathbf{W}_h(\boldsymbol{\rho}^{(i)}) = (\widetilde{\omega}_{h,n}(\mathbf{k}_\ell;\boldsymbol{\rho}^{(i)}))_{1\le\ell\le N_k,\ 1\le n\le N_b} \in \mathbb{R}^{N_k\times N_b}.$$

对于固定的i 和 $1 \le n \le N_b$，我们将该列

$$\widetilde{\omega}_{h,n}(\cdot;\boldsymbol{\rho}^{(i)}) := (\widetilde{\omega}_{h,n}(\mathbf{k}_\ell;\boldsymbol{\rho}^{(i)}))_{\ell=1}^{N_k} \in \mathbb{R}^{N_k}$$

视为能带函数的一个快照。汇集所有训练设计中的所有能带，可得到$N_s := N_{\text{train}}N_b$ 个快照，我们将其编索引为

$$s = (i-1)N_b + n, \qquad 1 \le i \le N_{\text{train}},\ 1 \le n \le N_b,$$

并组装成快照矩阵

$$X := (x_{s\ell})_{1\le s\le N_s,\ 1\le\ell\le N_k} \in \mathbb{R}^{N_s\times N_k}, \qquad x_{s\ell} := \widetilde{\omega}_{h,n}(\mathbf{k}_\ell;\boldsymbol{\rho}^{(i)}). \tag{3.1}$$

因此，X 的每一行都是在离散点 $\{\mathbf{k}_\ell\}_{\ell=1}^{N_k}$处采样的频带函数。

我们对转置后的快照矩阵进行奇异值分解 [59]

$$X^\top = U\Sigma V^\top, \tag{3.2}$$

其中，$U \in \mathbb{R}^{N_k \times N^k}$ 包含左奇异向量，而 $\Sigma = \mathrm{diag}(\sigma_1, \ldots, \sigma_{N_k})$，其中 $\sigma_1 \geq \cdots \geq \sigma_{N_k} \geq 0$。给定本征正交分解容差 $\tau_{\mathrm{POD}} > 0$，我们选择本征正交分解秩 $N_{\mathrm{POD}} \leq N_k$，使其为满足归一化尾部能量条件的最小整数:

$$\sum_{j>N_{\mathrm{POD}}} \sigma_j^2 / \sum_{j=1}^{N_k} \sigma_j^2 < \tau_{\mathrm{POD}}. \tag{3.3}$$

接下来，我们定义主干矩阵

$$\Phi_{\mathrm{tr}} := [\varphi_1, \ldots, \varphi_{N_{\mathrm{POD}}}] \in \mathbb{R}^{N_k \times N_{\mathrm{POD}}}, \tag{3.4}$$

其中，$\varphi_j \in \mathbb{R}^{N_k}$ 表示 j 是 U 的第几列。根据构造，列向量 $\{\varphi_j\}_{j=1}^{N_{\mathrm{POD}}}$ 构成本征正交分解子空间 $\mathcal{V}_{\mathrm{POD}} := \mathrm{span}\{\varphi_1, \ldots, \varphi_{N_{\mathrm{POD}}}\} \subset \mathbb{R}^{N_k}$ 的一组正交归一基。

奇异值分解的经典 Eckart–Young–Mirsky 定理 [60, 59] 给出了快照集投影到本征正交分解子空间上的投影误差的显式公式。对于每个快照能带矩阵 $\mathbf{W}^{(i)}{}_h \in \mathbb{R}^{N_k \times N_b} = 1, \ldots, N$，我们将其正交投影引入到该子空间上，定义为 $\mathcal{V}_{\mathrm{POD}}$ train

$$\mathbf{W}_{h,\mathrm{POD}}^{(i)} := \Phi_{\mathrm{tr}} C^{(i)}, \qquad C^{(i)} := \Phi_{\mathrm{tr}}^\top \mathbf{W}_h^{(i)} \in \mathbb{R}^{N_{\mathrm{POD}} \times N_b}.$$

命题 3.1 (快照集上的 POD 截断误差)。 设 $X \in \mathbb{R}^{N_s \times N^{\mathrm{train}}}$ 为 (3.1) 中定义的快照矩阵，$\{\sigma_j\}_{j=1}^{N_k}$ 为 $X^\top$ 按非递增顺序排列的奇异值，并设 Φ_{tr} 和 $\mathbf{W}^{(i)}{}_{h,\mathrm{POD}}$ 如上所定义。则数据集的平均投影误差满足

$$\frac{1}{N_{\mathrm{train}}} \sum_{i=1}^{N_{\mathrm{train}}} \left\| \mathbf{W}_h^{(i)} - \mathbf{W}_{h,\mathrm{POD}}^{(i)} \right\|_F^2 = \sum_{j>N_{\mathrm{POD}}} \sigma_j^2,$$

其中 $\|\cdot\|_F$ 表示弗罗贝尼乌斯范数。

对于一般设计 $\boldsymbol{\rho} \in [0,1]^{N_f}$，我们通过正交投影到 $\mathcal{V}_{\mathrm{POD}}$ 上来近似能带数据，并定义 POD 投影能带映射

$$F_h^{\mathrm{POD}} : [0,1]^{N_f} \to \mathbb{R}^{N_k \times N_b}, \qquad F_h^{\mathrm{POD}}(\boldsymbol{\rho}) := \mathbf{W}_h^{\mathrm{POD}}(\boldsymbol{\rho}), \tag{3.5}$$

其中，投影后的能带矩阵形式为

$$\mathbf{W}_h^{\mathrm{POD}}(\boldsymbol{\rho}) := \Phi_{\mathrm{tr}}\, C(\boldsymbol{\rho}), \qquad C(\boldsymbol{\rho}) := \Phi_{\mathrm{tr}}^\top \mathbf{W}_h(\boldsymbol{\rho}) \in \mathbb{R}^{N_{\mathrm{POD}} \times N_b}. \tag{3.6}$$

根据构造，$F^{\mathrm{POD}}{}_h$ 在快照设计上的限制会重现命题 3.1 中给出的最优平均投影误差。若快照 $\{\mathbf{W}^{(i)}{}_h\}_{i=1}^{N_{\mathrm{train}}}$ 充分采样 $\mathcal{E}_{\mathrm{rel}}$，且奇异值 $\{\sigma_j\}_{j=1}^{N_k}$ 快速衰减，则前 N_{POD} 个模态能够捕捉数据中所表示能带结构的大部分变化。在这种情况下，预计对于位于设计空间同一区域、但未包含在快照集中的设计，POD 截断误差也会保持较小。

在我们的 POD–DeepONet 构建中，主干矩阵 Φ_{tr} 在此离线阶段根据高保真度能带数据计算一次，随后在所有后续训练和逆向设计计算中固定用作主干网络。

3.2. POD–DeepONet 框架

在 POD–DeepONet 架构中，主干网络 由预计算矩阵 Φ_{tr} 给出，而分支网络 则近似 POD 系数对设计向量的非线性依赖关系。更具体地说，我们用一个全连接分支网络来近似式（3.6）中的系数映射 C

$$C_\theta : [0,1]^{N_f} \longrightarrow \mathbb{R}^{N_{\mathrm{POD}} \times N_b}, \qquad \boldsymbol{\rho} \longmapsto C_\theta(\boldsymbol{\rho}) := \left(c_{j,n}^\theta(\boldsymbol{\rho})\right)_{j=1,\ldots,N_{\mathrm{POD}}}^{n=1,\ldots,N_b},$$

其可训练参数为 θ。给定式（3.4）中的 Φ_{tr} 和分支网络的输出$C_\theta(\boldsymbol{\rho})$，POD–DeepONet 对采样$\mathbf{k}$点处离散能带数据的预测为

$$\mathbf{W}_h^{\text{POD-DO}}(\boldsymbol{\rho};\theta) := \Phi_{\mathrm{tr}}\, C_\theta(\boldsymbol{\rho}) = (\omega_{h,n}^{\text{POD-DO}}(\mathbf{k}_\ell;\boldsymbol{\rho},\theta))_{\ell=1,\dots,N_k}^{n=1,\dots,N_b}, \tag{3.7}$$

或者，按分量表示，

$$\omega_{h,n}^{\text{POD-DO}}(\mathbf{k}_\ell;\boldsymbol{\rho},\theta) = \sum_{j=1}^{N_{\text{POD}}} c_{j,n}^{\theta}(\boldsymbol{\rho})\,\varphi_j(\mathbf{k}_\ell), \qquad 1 \le \ell \le N_k,\ 1 \le n \le N_b,$$

其中，$c_{j,n}^{\theta}(\boldsymbol{\rho})$表示系数矩阵$C_\theta(\boldsymbol{\rho})$的第$j$行、第$n$列元素，而$\varphi_j$是第$j$个本征正交分解模态，即 Φ_{tr}的第j列。

为便于后文引用，我们将所得代理模型概括为如下算子

$$F_h^{\text{POD-DO}} : [0,1]^{N_f} \longrightarrow \mathbb{R}^{N_k\times N_b}, \qquad F_h^{\text{POD-DO}}(\boldsymbol{\rho};\theta) := \mathbf{W}_h^{\text{POD-DO}}(\boldsymbol{\rho};\theta) = \Phi_{\mathrm{tr}}\, C_\theta(\boldsymbol{\rho}), \tag{3.8}$$

该算子将在后续的正向和逆向计算中作为$F^{\text{pix}}{}_h$ 的高效近似。

式（3.8）定义了基于像素的能带映射$F^{\text{pix}}{}_h$ 的POD–DeepONet代理模型。预先计算得到的本征正交分解模态$\{\varphi_j\}_{j=1}^{N_{\text{POD}}}$在采样波矢上构成固定的物理信息主干基，而可训练分支网络C_θ 则学习本征正交分解系数如何依赖于高维设计向量 $\boldsymbol{\rho}$。这种变量分离方式结合了关于$\mathbf{k}$ 的降阶模型的高效性与可解释性，以及神经网络在设计空间中的灵活性。

*3.3. POD–DeepONet*前向映射的训练与评估

在数值实验中，我们以监督方式在训练集上训练$F^{\text{POD-DO}}{}_h$。该训练集包含从完整数据集 $\{(\boldsymbol{\rho}^{(i)}, \mathbf{W}^{(i)}{}_h)\}_{i=1}^{N_{\text{data}}}$ 中抽取的N_{train} 个带标签样本$\{(\boldsymbol{\rho}^{(i)},\mathbf{W}^{(i)}{}_h)\}_{i=1}^{N_{\text{train}}}$。

能带数据标准化

为稳定训练过程，我们对能带数据逐项进行仿射标准化（零均值和单位方差）。我们定义经验均值和方差

$$\mu := \frac{1}{N_{\text{train}}N_kN_b}\sum_{i=1}^{N_{\text{train}}}\sum_{\ell=1}^{N_k}\sum_{n=1}^{N_b}\widetilde{\omega}_{h,n}(\mathbf{k}_\ell;\boldsymbol{\rho}^{(i)}), \qquad \sigma^2 := \frac{1}{N_{\text{train}}N_kN_b}\sum_{i,\ell,n}(\widetilde{\omega}_{h,n}(\mathbf{k}_\ell;\boldsymbol{\rho}^{(i)})-\mu)^2.$$

标准化后的快照为

$$\mathbf{W}_h^{(i),\text{std}} := \frac{\mathbf{W}_h^{(i)} - \mu\mathbf{1}}{\sigma} \in \mathbb{R}^{N_k\times N_b}, \qquad i = 1,\dots,N_{\text{train}}, \tag{3.9}$$

其中减法和除法均按分量进行。以下损失函数（3.10）中使用这些标准化目标$\mathbf{W}^{(i),\text{std}}{}_h$。

监督训练目标

对于设计向量 $\boldsymbol{\rho} \in \{0,1\}^{N_f}$，标准化频带数据的 POD–DeepONet 预测由式 (3.7) 给出，即:

$$\mathbf{W}_h^{\text{POD-DO}}(\boldsymbol{\rho};\theta) := \Phi_{\mathrm{tr}}\, C_\theta(\boldsymbol{\rho}) \in \mathbb{R}^{N_k\times N_b}.$$

我们通过最小化经验均方误差来确定 θ

$$L(\theta) := \frac{1}{N_{\text{train}}N_kN_b}\sum_{i=1}^{N_{\text{train}}}\left\|\mathbf{W}_h^{\text{POD-DO}}(\boldsymbol{\rho}^{(i)};\theta) - \mathbf{W}_h^{(i),\text{std}}\right\|_F^2. \tag{3.10}$$

在实践中，我们通过随机梯度法（Adam）不断最小化 $L(\theta)$，直至收敛，并得到训练后的参数向量 θ^*。

在线评估

训练完成后，POD–DeepONet代理模型可对限定于像素参数化的离散前向映射进行快速、可微的近似。对于任意设计 $\boldsymbol{\rho} \in [0,1]^{N_f}$，我们通过逆标准化恢复能带数据：

$$\widetilde{\mathbf{W}}_h^{\text{POD-DO}}(\boldsymbol{\rho};\theta^*) := \sigma\,\mathbf{W}_h^{\text{POD-DO}}(\boldsymbol{\rho};\theta^*) + \mu\mathbf{1}. \tag{3.11}$$

POD–DeepONet代理模型的前向评估完整离线–在线流程总结于算法1中。

算法1： 使用POD–DeepONet代理模型进行前向评估

输入：数据集 $\{(\boldsymbol{\rho}^{(i)}, \mathrm{W}^{(i)}\mathrm{h})\}_{\mathrm{i}=1}^{N_{\text{data}}}$；POD 容差 τ_{POD}；分支网络 C_θ；训练集大小 $\mathrm{N}_{\text{train}}$；查询设计 $\boldsymbol{\rho}$。输出：预测的能带矩阵 $\widetilde{\mathrm{W}}^{\text{POD-DO}}\mathrm{h}(\boldsymbol{\rho};\theta^*)$。// 离线 1 随机选取训练集 $\{(\boldsymbol{\rho}^{(i)}, \mathrm{W}^{(i)}\mathrm{h})\}_{\mathrm{i}=1}^{N_{\text{train}}}$，从 $\{(\boldsymbol{\rho}^{(i)}, \mathrm{W}^{(i)}\mathrm{h})\}_{\mathrm{i}=1}^{N_{\text{data}}}$ 中选取；2 使用容差 τ_{POD} (3.4) 通过快照POD计算主干矩阵 Φ_{tr}；3 计算 μ、σ 和标准化数据 $\{\mathrm{W}^{(i),\text{std}}\mathrm{h}\}_{\mathrm{i}=1}^{N_{\text{train}}}$ (3.9)；// 训练 4 在 $\{(\boldsymbol{\rho}^{(i)}, \mathrm{W}^{(i),\text{std}}\mathrm{h})\}_{\mathrm{i}=1}^{N_{\text{train}}}$ 上使用损失函数 (3.10)，求得 $\theta^* := \arg\min_\theta \mathrm{L}(\theta)$；// 在线预测 5 计算POD–DeepONet对标准化数据的输出 $\mathrm{W}^{\text{POD-DO}}\mathrm{h}(\boldsymbol{\rho};\theta^*) = \mathrm{F}^{\text{POD-DO}}\mathrm{h}(\boldsymbol{\rho};\theta^*) := \Phi_{\text{tr}}\mathrm{C}_{\theta^*}(\boldsymbol{\rho})$；6 通过逆标准化获得预测的能带矩阵

$$\widetilde{\mathbf{W}}_h^{\text{POD-DO}}(\boldsymbol{\rho};\theta^*) := \sigma\,\mathbf{W}_h^{\text{POD-DO}}(\boldsymbol{\rho};\theta^*) + \mu\,\mathbf{1}.$$

备注 3.1（POD 主干网络中的固定**k**网格）。当前的 *POD–DeepONet*代理模型学习的是定义在预设高对称路径及固定**k**网格上的离散能带映射。*POD* 主干网络由在该网格上采样的能带快照提取而来，因此在相同路径和采样分辨率下提供最可靠的降阶表示。如果需要更密集的**k**采样，可以在细化后的网格上扩充快照集并重新构建*POD*基，同时保持分支网络架构和整体训练流程不变。同样，也可以通过重新生成相应的快照集合并构建相关的*POD* 主干网络，适配其他对称路径。开发连续依赖**k** 的主干网络（例如通过参数化路径坐标）是迈向离散不变评估的自然拓展。

*3.4. POD–DeepONet*代理模型的逼近性质

本小节分析 POD–DeepONet代理模型在放宽后的基于像素的设计空间中的逼近性质。我们首先证明离散能带映射的连续性。在此基础上，我们证明 POD–DeepONet 的通用逼近定理，并推导其总误差的分解。

我们首先给出松弛像素设计空间上离散能带映射的连续性结果。

命题 3.2（离散能带映射的连续性）。假设可容许介电常数满足

$$0 < \epsilon_{\min} \le \epsilon(\mathbf{x};\boldsymbol{\rho}) \le \epsilon_{\max} < \infty, \qquad \textit{for all } \mathbf{x} \in \Omega \textit{ and all } \boldsymbol{\rho} \in [0,1]^{N_f}.$$

对于 *TE* 偏振，我们设

$$\alpha(\mathbf{x};\boldsymbol{\rho}) := \epsilon(\mathbf{x};\boldsymbol{\rho})^{-1}, \qquad \beta(\mathbf{x};\boldsymbol{\rho}) := 1,$$

而对于 *TM* 偏振，我们设

$$\alpha(\mathbf{x};\boldsymbol{\rho}) := 1, \qquad \beta(\mathbf{x};\boldsymbol{\rho}) := \epsilon(\mathbf{x};\boldsymbol{\rho}).$$

设$\mathbf{A}(\mathbf{k};\boldsymbol{\rho})$ 和$\mathbf{B}(\boldsymbol{\rho})$ 为由这些系数组装得到的(2.12) 中的有限元矩阵。于是，对于每个波矢$\mathbf{k} \in \mathcal{K}_{\mathrm{hs}}$ 以及每个能带索引 $1 \le n \le N_b$，离散特征值

$$\lambda_{h,n}(\mathbf{k};\boldsymbol{\rho}) := \lambda_{h,n}(\mathbf{k};\epsilon(\cdot;\boldsymbol{\rho}))$$

连续依赖于 $\boldsymbol{\rho} \in [0,1]^{N_f}$。因此，离散能带映射

$$F_h^{\mathrm{pix}} : [0,1]^{N_f} \to \mathbb{R}^{N_k \times N_b}, \qquad \boldsymbol{\rho} \mapsto \mathbf{W}_h(\boldsymbol{\rho})$$

是一个连续映射。

证明。 根据式（2.14）中 $\epsilon(\mathbf{x};\boldsymbol{\rho})$ 的定义以及提升算子（2.13），可得

$$\epsilon(\mathbf{x};\boldsymbol{\rho}) = \epsilon_{\mathrm{air}} + \tilde{\rho}_\ell\,(\epsilon_{\mathrm{alum}} - \epsilon_{\mathrm{air}}), \qquad \mathbf{x} \in P_\ell,\ \ell = 1,\dots,N_{\mathrm{pix}}.$$

因此，在每个像素 P_ℓ 上， ϵ 的值连续依赖于标量 $\tilde{\rho}_\ell$。 $\alpha(\mathbf{x};\boldsymbol{\rho})$ 和$\beta(\mathbf{x};\boldsymbol{\rho})$ 也具有相同的性质，因此 α 和 β 是 $\boldsymbol{\rho} \in [0,1]^{N_f}$ 的连续函数。

接下来，在有限元离散化中，矩阵 $\mathbf{A}(\mathbf{k};\boldsymbol{\rho})$ 和 $\mathbf{B}(\boldsymbol{\rho})$ 由单元积分组装而成，其被积函数关于 α 和 β 呈线性（参见式（2.12））。因此，$\mathbf{A}(\mathbf{k};\boldsymbol{\rho})$ 和 $\mathbf{B}(\boldsymbol{\rho})$ 的每个元素都是 $\boldsymbol{\rho}$ 的连续函数，并且映射

$$\boldsymbol{\rho} \mapsto \mathbf{A}(\mathbf{k};\boldsymbol{\rho}), \qquad \boldsymbol{\rho} \mapsto \mathbf{B}(\boldsymbol{\rho})$$

相对于弗罗贝尼乌斯范数是连续的。

此外， $\epsilon(\mathbf{x};\boldsymbol{\rho})$ 的一致界意味着 $\beta(\mathbf{x};\boldsymbol{\rho})$ 一致地大于某个正数。因此，对于每个 $\boldsymbol{\rho} \in [0,1]^{N_f}$，质量矩阵 $\mathbf{B}(\boldsymbol{\rho})$ 都是厄米正定矩阵，并属于正定锥。由此我们可以定义厄米矩阵

$$\mathbf{H}(\mathbf{k};\boldsymbol{\rho}) := \mathbf{B}(\boldsymbol{\rho})^{-1/2}\mathbf{A}(\mathbf{k};\boldsymbol{\rho})\mathbf{B}(\boldsymbol{\rho})^{-1/2},$$

其中，$\mathbf{B}(\boldsymbol{\rho})^{-1/2}$ 表示主矩阵平方根。映射 $\mathbf{B} \mapsto \mathbf{B}^{-1/2}$ 在正定锥 [61] 上是解析的，因而是连续的。结合 $\mathbf{A}$ 和 $\mathbf{B}$ 的连续性可知， $\boldsymbol{\rho} \mapsto \mathbf{H}(\mathbf{k};\boldsymbol{\rho})$ 关于弗罗贝尼乌斯范数是连续的。

广义特征值问题 (2.12) 与 $\mathbf{H}(\mathbf{k};\boldsymbol{\rho})$ 的普通特征值问题相同，而 $\lambda_{h,n}(\mathbf{k};\boldsymbol{\rho})$ 是 $\mathbf{H}(\mathbf{k};\boldsymbol{\rho})$ 的第n个特征值，按非递减顺序排列。固定 $\mathbf{k} \in \mathcal{K}_{\mathrm{hs}}$ 和 $\boldsymbol{\rho},\boldsymbol{\rho}' \in [0,1]^{N_f}$，令 $\lambda_{h,n}(\mathbf{k};\boldsymbol{\rho})$ 和 $\lambda_{h,n}(\mathbf{k};\boldsymbol{\rho}'),n = 1,\dots,N_h$ 分别表示 $\mathbf{H}(\mathbf{k};\boldsymbol{\rho})$ 和 $\mathbf{H}(\mathbf{k};\boldsymbol{\rho}')$ 的有序特征值。厄米矩阵的霍夫曼 – 维兰特不等式 [61] 给出

$$\left|\lambda_{h,n}(\mathbf{k};\boldsymbol{\rho}) - \lambda_{h,n}(\mathbf{k};\boldsymbol{\rho}')\right| \le \left\|\mathbf{H}(\mathbf{k};\boldsymbol{\rho}) - \mathbf{H}(\mathbf{k};\boldsymbol{\rho}')\right\|_F, \tag{3.12}$$

对于每个 $1 \le n \le N_b$。因此，$\lambda_{h,\,n}(\mathbf{k};\,\cdot)$ 在 $[0,1]^{N_f}$ 上也是连续的。

最后，对于每个采样波矢 $\mathbf{k}_\ell$ 以及 $1 \le n \le N_b$，归一化能带值

$$\widetilde{\omega}_{h,n}(\mathbf{k}_\ell;\boldsymbol{\rho}) = \frac{a}{2\pi}\sqrt{\lambda_{h,n}(\mathbf{k}_\ell;\boldsymbol{\rho})}$$

是连续函数 $\lambda_{h,\,n}(\mathbf{k}_\ell;\,\cdot)$ 与 $[0,\infty)$ 上的平方根函数的复合，因此在 $[0,1]^{N_f}$ 上连续。将这些有限个标量函数汇集到 $\mathbf{W}_h(\boldsymbol{\rho})$ 中，可得离散能带映射 $F^{\mathrm{pix}}{}_h$ 的连续性，证毕。

□

现在我们来讨论 POD–DeepONet代理模型的逼近性质。

定理 3.2 (POD–DeepONet 的逼近性质)。 设 $\mathbf{W}^{\mathrm{POD}}{}_h$ 由 (3.6) 给出。固定 $1 \le N_{\mathrm{POD}} \le N_k$ 以及一个具有正交归一列的主干矩阵 $\Phi_{\mathrm{tr}} \in \mathbb{R}^{N_k \times N^{\mathrm{POD}}}$。对于任意 $\varepsilon > 0$，存在一个分支网络 $C_\theta\colon [0,1]^{N_f} \to \mathbb{R}^{N_{\mathrm{POD}} \times N^b}$ 及其对应的 *POD–DeepONet* 波段映射

$$\mathbf{W}_h^{\text{POD-DO}}(\boldsymbol{\rho}) := \Phi_{\mathrm{tr}} C_\theta(\boldsymbol{\rho}), \qquad \boldsymbol{\rho} \in [0,1]^{N_f},$$

使得

$$\sup_{\boldsymbol{\rho}\in[0,1]^{N_f}} \left\|\mathbf{W}_h^{\mathrm{POD}}(\boldsymbol{\rho}) - \mathbf{W}_h^{\text{POD-DO}}(\boldsymbol{\rho})\right\|_F \le \varepsilon.$$

证明。 根据命题 3.2，在 (2.15) 中定义的离散能带映射 $F^{\mathrm{pix}}{}_h$ 在 $[0,1]^{N_f}$ 上连续，因此相应的能带矩阵 $\mathbf{W}_h(\boldsymbol{\rho}) := F^{\mathrm{pix}}{}_h(\boldsymbol{\rho}) \in \mathbb{R}^{N_k \times N^b}$ 连续依赖于 $\boldsymbol{\rho}$。回顾本征正交分解系数映射

$$C : [0,1]^{N_f} \to \mathbb{R}^{N_{\mathrm{POD}} \times N_b}, \qquad C(\boldsymbol{\rho}) := \Phi_{\mathrm{tr}}^{\top} \mathbf{W}_h(\boldsymbol{\rho}),$$

在 (3.6) 中引入。由于 $\Phi_{\mathrm{tr}}^{\top}$ 是固定线性算子，$C(\cdot)$ 在 $[0,1]^{N_f}$ 上也是连续的。

根据前馈神经网络的通用逼近定理（例如，参见 [62, 63]），对于任意 $\varepsilon > 0$，存在一个分支网络 C_θ：$[0,1]^{N_f} \to \mathbb{R}^{N_{\mathrm{POD}} \times N^b}$，使得

$$\sup_{\boldsymbol{\rho} \in [0,1]^{N_f}} \left\| C(\boldsymbol{\rho}) - C_\theta(\boldsymbol{\rho}) \right\|_F \le \varepsilon.$$

利用 (3.6) 中 $\mathbf{W}^{\mathrm{POD}}{}_h$ 的定义以及上文中 $\mathbf{W}^{\text{POD-DO}}{}_h$ 的定义，可得对所有 $\boldsymbol{\rho} \in [0,1]^{N_f}$，

$$\left\| \mathbf{W}_h^{\mathrm{POD}}(\boldsymbol{\rho}) - \mathbf{W}_h^{\text{POD-DO}}(\boldsymbol{\rho}) \right\|_F = \left\| \Phi_{\mathrm{tr}}(C(\boldsymbol{\rho}) - C_\theta(\boldsymbol{\rho})) \right\|_F \le \left\| C(\boldsymbol{\rho}) - C_\theta(\boldsymbol{\rho}) \right\|_F \le \varepsilon,$$

其中用到了 Φ_{tr} 具有正交归一列，因此它在弗罗贝尼乌斯范数下是压缩映射。对 $\boldsymbol{\rho} \in [0,1]^{N_f}$ 取上确界，即得所需估计。 □

结合定理 3.2 与命题 3.1，可将代理模型的总逼近误差分解为两部分。对于任意 $\boldsymbol{\rho} \in [0,1]^{N_f}$，

$$\left\| \mathbf{W}_h(\boldsymbol{\rho}) - \mathbf{W}_h^{\text{POD-DO}}(\boldsymbol{\rho}) \right\|_F \le \left\| \mathbf{W}_h(\boldsymbol{\rho}) - \mathbf{W}_h^{\mathrm{POD}}(\boldsymbol{\rho}) \right\|_F + \left\| \mathbf{W}_h^{\mathrm{POD}}(\boldsymbol{\rho}) - \mathbf{W}_h^{\text{POD-DO}}(\boldsymbol{\rho}) \right\|_F.$$

右端第一项是*POD* 截断误差。它受用于选择 POD 秩的容差 τ_{POD} 控制，并且可以通过增大N_{POD} 或扩充快照集来减小；参见命题 3.1。第二项是网络逼近误差。对于固定的主干网络 Φ_{tr}，原则上，通过提高分支网络的表达能力，可以使该误差任意小，定理 3.2 保证了这一点。在实践中，我们首先选择N_{POD}，使 POD 截断误差低于规定的容差 τ_{POD}，然后选择网络架构和训练流程，使剩余偏差达到同阶或更小，从而使 POD 截断成为代理模型中的主要误差来源。

这些结果表明，对于固定的 POD 秩，POD–DeepONet代理模型可以在松弛设计空间 $[0,1]^{N_f}$上均匀逼近离散能带映射$F^{\mathrm{pix}}{}_h$。因此，下文逆问题中使用的基于代理模型的目标泛函，能够一致地逼近其离散有限元对应项。

注 3.3。 在本小节中，我们在第 *2.3* 节介绍的松弛设计空间 $[0,1]^{N_f}$ 上开展研究。实际应用中，离散能带映射和 *POD–DeepONet* 代理模型是在二值设计 $\boldsymbol{\rho} \in \{0,1\}^{N_f}$ 上评估的，而二值设计 $\boldsymbol{\rho} \in \{0,1\}^{N_f}$ 属于 $[0,1]^{N_f}$ 的子集，因此上述连续性和逼近结果尤其适用于具有实际意义的双材料胞元。采用松弛处理是为了简化分析，并便于使用基于梯度的逆向设计算法。在数值实验中，随后将松弛输出投影为二值像素，以获得两相胞元；结果表明，对于本文考虑的应用，这种松弛公式是适用的。

4. 基于 POD–DeepONet 的逆向能带结构设计

本节针对第 2.4 节介绍的两个逆能带结构问题，开发基于 POD–DeepONet 的算法。在这两种设定下，我们都旨在寻找设计向量 $\widehat{\boldsymbol{\rho}}$，使其对应的单胞能够再现指定的能带结构或实现指定的带隙。为进行基于梯度的优化，我们采用第 2.3 节介绍的松弛密度空间 $[0,1]^{N_f}$。我们构建逆神经网络，将目标规格映射为松弛设计 $\widehat{\boldsymbol{\rho}}$，并通过预训练的 POD–DeepONet代理模型评估这些网络；该代理模型能够快速进行带结构预测，并计算相对于$\widehat{\boldsymbol{\rho}}$的梯度。我们首先介绍色散到结构的逆向设计，随后介绍带隙逆问题。

4.1. 色散到结构的逆问题

色散到结构问题旨在寻找设计 $\widehat{\boldsymbol{\rho}}$，使其能带结构接近指定的目标能带矩阵$\mathbf{W}_h$ (2.16)。我们不针对每个目标分别求解 (2.16)，而是引入一个神经网络来表示逆映射，并以监督方式训练该网络。

我们重新使用正向数据集

$$\{(\boldsymbol{\rho}^{(i)}, \mathbf{W}_h^{(i)})\}_{i=1}^{N_{\text{data}}}$$

该数据集在第 3.3 节中介绍，并选取训练子集 $\{(\boldsymbol{\rho}^{(i)},\mathbf{W}^{(i)}{}_h)\}_{=1}^{N_{\text{train}}}{}_i$，其样本数为 $N_{\text{train}} \le N_{\text{data}}$。按照第 3.3 节中的相同标准化流程，我们得到标准化目标 $\mathbf{W}^{(i)}{}_{,\ \text{std}\,h} \in \mathbb{R}^{N_k \times N^b}$，其中 $i = 1, \cdots, N_{N_{\text{data}}}$。对于每个目标，我们将 $\mathbf{W}^{(i)}{}_{,}^{\text{std}}{}_h$ 展平为单列，从而构造特征向量 $\mathbf{y}^{(i)} \in \mathbb{R}^{N_\omega}$，$N_\omega := N_k N_b$。逆向网络

$$G_\phi^{\text{disp}} : \mathbb{R}^{N_\omega} \to \mathbb{R}^{N_f}, \qquad \mathbf{z} = G_\phi^{\text{disp}}(\mathbf{y}),$$

是一个参数为 ϕ 的全连接多层感知机。给定目标特征向量 $\mathbf{y}$，网络会生成一个不受约束的向量 $\mathbf{z} \in \mathbb{R}^{N_f}$，随后将其映射为松弛楔形区域，最后再按下文所述映射为二值楔形。

松弛参数化与二值投影

为了将无约束输出 $\mathbf{z} = G_\phi^{\text{disp}}(\mathbf{y})$ 转换为 $[0,1]^{N_f}$ 中的松弛楔形区域，我们首先应用逻辑映射

$$\sigma : \mathbb{R}^{N_f} \to (0,1)^{N_f}, \qquad \boldsymbol{\rho}^{\text{sig}} = \sigma(\mathbf{z}) := (\sigma(z_j))_{j=1}^{N_f}, \quad \sigma(t) := \frac{1}{1+e^{-t}},$$

该映射确保对于所有 $j = 1, \cdots, Nf$，均有 $0 < \rho^{\text{sig}}{}_j < 1$。为使松弛设计趋近于二值，我们随后逐分量应用平滑海维赛德投影。对于陡峭度参数 $\beta > 0$ 和阈值 $\eta \in (0,1)$，定义

$$\boldsymbol{\rho}(\mathbf{z}) := H_\beta(\boldsymbol{\rho}^{\text{sig}}) \in (0,1)^{N_f}, \qquad (H_\beta(\boldsymbol{\rho}^{\text{sig}}))_j := \frac{\tanh(\beta\eta) + \tanh(\beta(\rho_j^{\text{sig}} - \eta))}{\tanh(\beta\eta) + \tanh(\beta(1-\eta))}, \quad j = 1, \ldots, N_f. \tag{4.1}$$

训练期间，我们采用延续策略：将 β 从较小的初始值逐渐增大至较大的最终值，从而使 $\boldsymbol{\rho}(\mathbf{z})$ 逐渐趋近于 $\{0,1\}^{N_f}$，同时映射 $\mathbf{z} \mapsto \boldsymbol{\rho}(\mathbf{z})$ 保持平滑且可微。

将逆向网络与此参数化相结合，与目标特征向量$\mathbf{y}$ 对应的松弛楔形区域为

$$\boldsymbol{\rho}(\mathbf{y}; \phi, \beta) := H_\beta(\sigma(G_\phi^{\text{disp}}(\mathbf{y}))) \in [0,1]^{N_f}. \tag{4.2}$$

因此，网络输出 $\mathbf{z} = G_\phi^{\text{disp}}(\mathbf{y})$ 经由 sigmoid 和海维赛德变换，转换为连续设计空间 $[0,1]^{N_f}$ 中的松弛密度 $\boldsymbol{\rho}(\mathbf{y};\ \phi, \beta)$。

训练收敛后，我们通过硬阈值化得到一个二值楔形区域，

$$\rho_j^{\text{bin}}(\mathbf{y}; \phi, \beta) := \mathbf{1}_{\{\rho_j(\mathbf{y};\phi,\beta) > 1/2\}}, \qquad j = 1, \ldots, N_f, \tag{4.3}$$

从而得到一个离散双材料单胞。

训练目标

逆向网络应生成能带图与目标匹配、接近二值设计且靠近数据库中监督样本的单胞。令

$$\{(\mathbf{W}_h^{(i),\text{std}}, \mathbf{y}^{(i)}, \boldsymbol{\rho}^{(i)})\}_{i=1}^{N_{\text{train}}}$$

分别表示训练集 $\{0,1\}^{N_f}$ 中的标准化目标能带矩阵、其特征向量以及对应的楔形设计。对于每个 i，逆向网络和松弛参数化会生成

$$\widehat{\boldsymbol{\rho}}^{(i)}(\phi, \beta) := \boldsymbol{\rho}(\mathbf{y}^{(i)}; \phi, \beta) \in [0,1]^{N_f},$$

如式 (4.2) 所示。随后，参数冻结的 POD–DeepONet代理模型 θ^* 给出标准化能带预测 $\mathbf{W}^{\text{POD-DO}}{}_h\,(\widehat{\boldsymbol{\rho}}^{(i)}(\phi,\beta);\theta^*)$。

对于每个训练样本，我们定义三个贡献项:

$$J^{(i)}_{\text{MSE}}(\phi,\beta) := \frac{1}{N_k N_b}\left\|\mathbf{W}_h^{\text{POD-DO}}\widehat{\boldsymbol{\rho}}^{(i)}(\phi,\beta);\theta^*) - \mathbf{W}_h^{(i),\text{std}}\right\|_F^2, \tag{4.4}$$

$$R^{(i)}_{\text{bin}}(\phi,\beta) := \frac{1}{N_f}\sum_{j=1}^{N_f}\widehat{\rho}^{(i)}_j(\phi,\beta)(1-\widehat{\rho}^{(i)}_j(\phi,\beta)), \tag{4.5}$$

$$R^{(i)}_{\text{sup}}(\phi,\beta) := \frac{1}{N_f}\left\|\widehat{\boldsymbol{\rho}}^{(i)}(\phi,\beta) - \boldsymbol{\rho}^{(i)}\right\|_2^2. \tag{4.6}$$

其中，$J^{(i)}{}_{\text{MSE}}$ 衡量预测能带图与目标能带图之间的不匹配程度，$R^{(i)}{}_{\text{bin}}$ 惩罚灰度像素，并且仅在 $\{0,1\}$ 设计为二值时严格为零；而$R^{(i)}_{\text{sup}}$ 是一个弱邻近项，使松弛设计接近监督样本。

经验训练目标为

$$\mathcal{J}_{\text{disp}}(\phi,\beta) := \frac{1}{N_{\text{train}}}\sum_{i=1}^{N_{\text{train}}}\Big(J^{(i)}_{\text{MSE}}(\phi,\beta) + \gamma_{\text{bin}}R^{(i)}_{\text{bin}}(\phi,\beta) + \gamma_{\text{sup}}R^{(i)}_{\text{sup}}(\phi,\beta)\Big), \tag{4.7}$$

其中正则化权重为 γ_{bin}、$\gamma_{\text{sup}} > 0$。参数 ϕ 通过基于梯度的方法进行优化，同时保持θ^* 固定。三个贡献项的梯度均通过 POD–DeepONet代理模型以及映射 $\mathbf{y} \mapsto \boldsymbol{\rho}(\mathbf{y};\phi,\beta)$，利用自动微分进行传播。

在推理阶段，只需目标能带特征向量：对G^{disp}_ϕ 进行一次评估，再经过变换（4.2）和（4.3），即可得到一个能够近似复现给定能带图的二值楔形区域。

请注意，色散到结构映射和带隙映射都高度非单射，因为不同的微观结构可能产生本质上相同的能带图。因此，仅优化 $J^{(i)}{}_{\text{MSE}}$ 和 $R^{(i)}{}_{\text{bin}}$ 会留下一个庞大的可行设计等价类。引入监督项 $R^{(i)}_{\text{sup}}$ 作为一种数据驱动的正则化，通过引导逆向网络趋向数据库所表征的设计流形，从而在这些相互竞争的解中进行选择，并稳定训练过程。

注 4.1 (正向代理模型的松弛输入)。 在逆向训练过程中，*sigmoid*函数与*Heaviside*函数的结合使 $\boldsymbol{\rho}(\mathbf{y};\ \phi,\beta)$ 接近 $\{0,1\}^{N_f}$，因此使用松弛输入 $\boldsymbol{\rho} \in [0,1]^{N_f}$仅构成轻微外推。命题 *3.2* 给出了离散能带映射在 $[0,1]^{N_f}$ 上的连续性，而定理 *3.2* 保证 *POD–DeepONet* 能够在松弛域上逼近这一连续映射。尽管正向 *POD–DeepONet* 代理模型是在二值设计 $\boldsymbol{\rho} \in \{0,1\}^{N_f}$ 上训练的，并在反问题中通过松弛输入进行评估，数值结果表明，它在这些近二值输入上仍然准确。

4.2. 带隙反问题

对于带隙逆问题，我们采用类似的方法，其目标不是完整的能带图，而是指定能带a与b之间的带隙区间（a，b），其中p 和p + 1表示能带。我们用三维特征向量对这一设定进行编码

$$\mathbf{g} := (a,b,p)^\top \in \mathbb{R}^3,$$

其各分量在输入网络前均经过标准化。一个逆向网络 rk

$$G^{\text{gap}}_\phi : \mathbb{R}^3 \to \mathbb{R}^{N_f}, \qquad \mathbf{z} = G^{\text{gap}}_\phi(\mathbf{g}),$$

采用与G^{disp}_ϕ 相同的多层感知机架构，将$\mathbf{g}$ 映射为一个无约束参数向量$\mathbf{z}$。随后通过 Sigmoid 和 Heaviside投影得到松弛楔形区域,

$$\boldsymbol{\rho}(\mathbf{g};\phi,\beta) := H_\beta(\sigma(G^{\text{gap}}_\phi(\mathbf{g}))) \in [0,1]^{N_f},$$

并作为固定POD–DeepONet代理模型$\mathbf{W}^{\text{POD-DO}}{}_h\,(\boldsymbol{\rho};\theta^*)$的输入。

带隙逆网络应生成能够打开所需带隙的晶胞，同时尽可能接近二值设计，并接近数据库中的代表性示例。设

$$\{(\mathbf{g}^{(i)}, \boldsymbol{\rho}^{(i)})\}_{i=1}^{N_{\text{train}}}$$

表示标准化的带隙目标 $\mathbf{g}^{(i)} = (a^{(i)}, b^{(i)}, p^{(i)})^\top$ 以及用于训练的相应楔形设计 $\boldsymbol{\rho}^{(i)} \in \{0,1\}^{N_f}$。对于每个$i$，我们定义松弛楔形

$$\widehat{\boldsymbol{\rho}}^{(i)}(\phi,\beta) := \boldsymbol{\rho}(\mathbf{g}^{(i)};\phi,\beta) \in [0,1]^{N_f},$$

并评估具有冻结参数 θ^* 的 POD–DeepONet代理模型 $\mathbf{W}^{\text{POD-DO}}{}_h(\widehat{\boldsymbol{\rho}}^{(i)}(\phi,\beta);\ \theta^*)$。令 $\omega^{\text{POD-DO}}{}_{h,n}(\mathbf{k}_\ell;\ \boldsymbol{\rho},\theta^*)$ 表示 $\mathbf{W}^{\text{POD-DO}}{}_h(\boldsymbol{\rho};\ \theta^*)$ 的第 (l, n) 项。对于给定目标 $(a^{(i)}, b^{(i)}, p^{(i)})$，我们定义带隙约束项

$$\begin{aligned} J^{(i)}_{\text{gap}}(\phi,\beta) := \frac{1}{N_k N_b}\sum_{\ell=1}^{N_k}\sum_{n=1}^{N_b} \mathrm{ReLU}(\omega^{\text{POD-DO}}_{h,n}(\mathbf{k}_\ell;\widehat{\boldsymbol{\rho}}^{(i)}(\phi,\beta),\theta^*) - a^{(i)})\,\mathrm{ReLU}(b^{(i)} - \omega^{\text{POD-DO}}_{h,n}(\mathbf{k}_\ell;\widehat{\boldsymbol{\rho}}^{(i)}(\phi,\beta),\theta^*)) \\ + \frac{1}{N_k}\sum_{\ell=1}^{N_k}\Big(\mathrm{ReLU}(\omega^{\text{POD-DO}}_{h,p^{(i)}}(\mathbf{k}_\ell;\widehat{\boldsymbol{\rho}}^{(i)}(\phi,\beta),\theta^*) - a^{(i)})^2 + \mathrm{ReLU}(b^{(i)} - \omega^{\text{POD-DO}}_{h,p^{(i)}+1}(\mathbf{k}_\ell;\widehat{\boldsymbol{\rho}}^{(i)}(\phi,\beta),\theta^*))^2\Big), \end{aligned} \tag{4.8}$$

该项会惩罚位于 $(a^{(i)}, b^{(i)})$ 内的代理模型能带，并将第 $p^{(i)}$ 条能带压低至$a^{(i)}$ 以下，同时将第 $(p^{(i)}+1)$- 条能带抬高至$b^{(i)}$以上。此外，正如上文所述，我们还使用二值性惩罚项和较弱的邻近项，其定义见式 (4.5)–(4.6)。因此，经验训练目标为

$$\mathcal{J}_{\text{gap}}(\phi,\beta) := \frac{1}{N_{\text{train}}}\sum_{i=1}^{N_{\text{train}}}\Big(J^{(i)}_{\text{gap}}(\phi,\beta) + \gamma_{\text{bin}}R^{(i)}_{\text{bin}}(\phi,\beta) + \gamma_{\text{sup}}R^{(i)}_{\text{sup}}(\phi,\beta)\Big), \tag{4.9}$$

正则化权重为 $\gamma_{\text{bin}}, \gamma_{\text{sup}} > 0$。参数 ϕ 通过基于梯度的方法进行优化，而 θ^* 保持固定。梯度通过 POD–DeepONet代理模型以及映射 $\mathbf{g} \mapsto \boldsymbol{\rho}(\mathbf{g};\phi,\beta)$，利用自动微分进行传播。

在推理阶段，只需要带隙描述符 $\mathbf{g}$： 对G^{gap}_ϕ 进行一次求值，再经过变换 (4.2) 和 (4.3)，即可得到一个近似打开指定带隙的二值楔形区域。

注4.2。 带隙约束项 $J^{(i)}_{\text{gap}}$ 会惩罚落在目标区间 $(a^{(i)},b^{(i)})$ 内的代理能带值，并促使第 $p^{(i)}$ 条能带低于 $a^{(i)}$、第 $(p^{(i)}+1)$ 条能带高于 $b^{(i)}$。因此，该损失函数旨在确保指定区间包含在一个完全带隙内，而非强制带隙边缘精确匹配。不过在实践中，监督邻近项以及当前双相、受对称性和分辨率约束的参数化所能实现的有限带隙扩大，共同抑制了过度超调。因此，在我们的数值实验中，实际形成的带隙往往紧密包围目标区间，带隙边缘通常位于 $a^{(i)}$ 和 $b^{(i)}$ 附近。

算法2概述了用于色散到结构问题和带隙问题的基于梯度的逆向设计流程，其中所有标准化目标（能带图特征**y** 或带隙描述符**g**）均表示为特征向量**s** ，并通过统一流程进行处理。图4简要概述了整体工作流程，涵盖从快照POD和代理模型训练到正向查询以及两类逆向设计问题。

5. 数值实验

本节评估算法1中的POD–DeepONet 前向代理模型的性能，以及基于该模型的算法2中的逆向设计方案。

我们考虑具有方形晶胞、采用 TE 偏振的二维光子晶体，其由标量亥姆霍兹特征值问题 (2.7a) 描述。TM 情形也可用相同方式处理。所有示例均采用二值、$p4m$- 对称晶胞，像素网格为 16×16，其中 $\epsilon(\mathbf{x}) \in \{\epsilon_{\text{air}}, \epsilon_{\text{alum}}\}$，且 $\epsilon_{\text{air}} = 1$ 和 $\epsilon_{\text{alum}} = 8$。9。注意，每个$p4m$-对称晶胞均可由包含 $N_f = 36$ 个独立像素的楔形设计向量 $\boldsymbol{\rho} \in \{0,1\}^{N_f}$表示（见图5）。受 Jiang *et al.* [28]，的开放数据集启发，我们随机生成楔形设计向量，再将其扩展为完整的 16×16 晶胞。剔除高度碎片化或平凡的图案后，我们得到一个数据库，其中包含

算法2：采用 POD–DeepONet 的基于梯度的逆向设计 输入：问题4.1的训练数据 $\{(y^{(i)}, \boldsymbol{\rho}^{(i)})\}_{i=1}^{N_{\text{train}}}$，或问题4.2的训练数据 $\{(g^{(i)}, \boldsymbol{\rho}^{(i)})\}_{i=1}^{N_{\text{train}}}$；固定POD–DeepONet代理模型 $W^{\text{POD-DO}}h(\cdot;\theta^*)$；逆向网络 G_ϕ；Heaviside阈值 η；β的延续方案；正则化权重 $\gamma_{\text{bin}}, \gamma_{\text{sup}}$；查询目标 y 或 g。输出：二值单胞 $\epsilon^{\text{bin}}(x;\boldsymbol{\rho}^{\text{bin}})$。// 离线 1 对每个 $i = 1, \ldots, N_{\text{train}}$，将输入 $y^{(i)}$（色散目标）或 $g^{(i)}$（带隙目标）标准化，得到特征向量 $s^{(i)}$，并构建标准化训练集 $\{(s^{(i)}, \boldsymbol{\rho}^{(i)})\}_{i=1}^{N_{\text{train}}}$；// 训练 2 对每个 $i = 1, \ldots, N_{\text{train}}$，定义 $z^{(i)} := G_\phi(s^{(i)})$，$\widehat{\boldsymbol{\rho}}^{(i)}(\phi,\beta) := H_\beta(\sigma(z^{(i)}))$。使用 $\{(s^{(i)}, \boldsymbol{\rho}^{(i)}, \widehat{\boldsymbol{\rho}}^{(i)}(\phi,\beta))\}_{i=1}^{N_{\text{train}}}$ 构建式 (4.7) 中的经验目标函数 $\mathcal{J}_{\text{disp}}(\phi,\beta)$，或式 (4.9) 中的 $\mathcal{J}_{\text{gap}}(\phi,\beta)$。在训练集上求得 $\phi^* := \arg\min_\phi \mathcal{J}_{\text{disp}}(\phi,\beta)$ 或 $\phi^* := \arg\min_\phi \mathcal{J}_{\text{gap}}(\phi,\beta)$；// 在线预测 3 为查询目标 y 或 g 构建并标准化特征向量 s；4 $z \leftarrow G_{\phi^*}(s)$，$\boldsymbol{\rho}^* \leftarrow H_\beta(\sigma(z))$；5 $\rho_j^{\text{bin}} \leftarrow \mathbf{1}_{\{\rho_j^* > 1/2\}}$，$j = 1, \ldots, N_f$；6 通过像素参数化 (2.14) 构建二值单胞 $\epsilon^{\text{bin}}(x;\boldsymbol{\rho}^{\text{bin}})$。

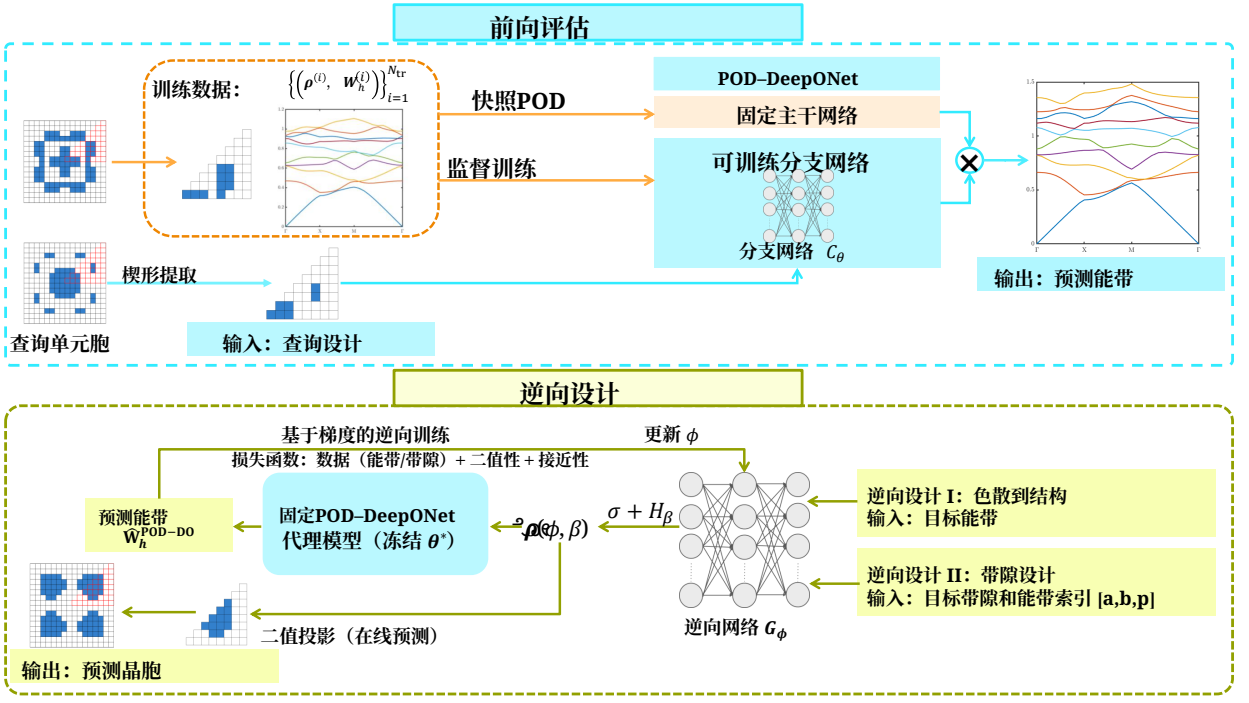


图 4. POD–DeepONet 框架的示意工作流程。上方概述了前向评估：快照POD构建固定主干网络，使用标准化数据上的能带均方误差损失训练分支网络（算法1），随后所得代理模型为查询设计预测能带结构。下方展示了两种基于 POD–DeepONet 的逆向设计流程，其中通过包含数据、二值性和接近性项的基于梯度的逆向训练，处理从色散到结构的设计以及带隙目标（算法2）。

$N_{\mathrm{data}} = 87{,}474$ 个不同的单胞。对于每个设计 $\boldsymbol{\rho}$，我们在晶胞的形状正则三角剖分上采用协调的 P_1 有限元离散化求解 TE 特征值问题（2.7a），得到矩阵特征值问题（2.12）。随后，我们计算沿高对称路径 $\mathcal{K}_{\mathrm{hs}}$ 均匀分布的 $N_b = 10$ 个波矢处的前 $N_k = 31$ 个归一化能带频率，并将其收集到

$$\mathbf{W}_h(\boldsymbol{\rho}) := (\widetilde{\omega}_{h,n}(\mathbf{k}_\ell;\boldsymbol{\rho}))_{\ell=1,\dots,N_k}^{n=1,\dots,N_b} \in \mathbb{R}^{N_k\times N_b}.$$

这些矩阵构成高保真度能带结构数据集。我们使用这些数据生成 POD 快照、训练 POD–DeepONet 前向代理模型和两个逆向设计模型。

完整的 $N_{\mathrm{data}} = 87{,}474$ 个晶胞数据集被随机划分为 $N_{\mathrm{train}} = 81{,}474$ 个训练样本、$N_{\mathrm{val}} = 5{,}000$ 个验证样本和 $N_{\mathrm{test}} = 1{,}000$ 个测试样本。训练前，所有能带数据 $\mathbf{W}_h(\boldsymbol{\rho})$ 均按第 3.3 节所述进行标准化，以下报告的所有误差均在逆标准化后测量。按照第 3.1 节，我们组装快照矩阵 $X \in \mathbb{R}^{N_s\times N^k}$（定义见式（3.1）），并对 $X^\top$ 进行奇异值分解。POD 秩 N_{POD} 根据容差 $\tau_{\mathrm{POD}} = 10^{-7}$ 按式（3.3）选取，由此得到 $N_{\mathrm{POD}} = 25$，相应的 POD 模态构成主干矩阵 $\Phi_{\mathrm{tr}} \in \mathbb{R}^{N_k\times 25}$，并在后续所有实验中保持固定。

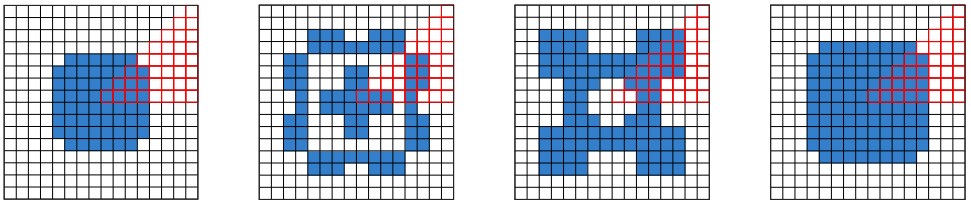


图5. 数据集中具有代表性的 16×16 基于像素的单胞。蓝色像素表示高介电常数材料 ϵ_{alum}，白色像素表示背景 ϵ_{air}。每个面板中的红色阶梯状区域标示设计楔形区域 $\boldsymbol{\rho}$，其中包含用于参数化单胞的$N_f = 36$ 个独立像素。

*5.1. POD–DeepONet*前向代理模型的准确性

我们首先评估POD–DeepONet代理模型在基于像素的设计空间上的前向映射$F^{\mathrm{pix}}{}_h$ 的准确性。

在分支网络C_θ的数值实现中，我们采用全连接架构，其中包含两个宽度为128的隐藏层，并使用 ReLU 激活函数。训练在PyTorch中进行，使用Adam 优化器，学习率为 10^{-3}，小批量大小为256，训练1000轮次。损失函数为均方误差，与式(3.10)相同。表1总结了本节使用的主要数据集规模和网络维度。

为凸显在固定主干网络基底中嵌入带隙信息的优势，我们还考虑了一个直接 MLP 基线，该基线学习从相同楔形特征到完整能带输出的映射，且不采用任何降阶结构。具体而言，该基线是一个全连接网络，包含一个宽度为 153、采用 ReLU 激活函数的隐藏层，以 36 维楔形向量为输入，直接输出 31×10 能带值。采用这一设置后，该基线约有 5.34×10^4 个可训练参数，与分支网络基本相当，从而确保比较公平。我们使用与 POD–DeepONet 相同的设置训练该基线。

表1. POD–DeepONet代理模型的数据集大小和网络维度汇总。

数量	符号	值
路径上的**k**点数量	N_k	31
能带数量	N_b	10
主干POD秩	N_{POD}	25
输入特征维度（分支）	N_f	36
训练样本	N_{train}	81,474
验证样本	N_{val}	5,000
测试样本	N_{test}	1,000

设 $\mathcal{I}_{\text{test}}$ 为N_{test} 个测试样本的索引集合。在测试集上，我们通过逐能带、逐样本和全局误差度量来量化预测准确度。对于每个测试索引$i \in \mathcal{I}_{\text{test}}$和每个波矢$\mathbf{k}_\ell$，其中$\mathrm{l} = 1, \cdots, N_k$，我们将 $\omega^{\text{POD-DO}}{}_{h,n}(\mathbf{k}_\ell;\, \boldsymbol{\rho}^{(i)}, \theta^*)$ 记为 POD–DeepONet 或 MLP 代理模型给出的预测能带值，将$\widetilde{\omega}_{h,n}(\mathbf{k}_\ell; \boldsymbol{\rho}^{(i)})$记为有限元求解器给出的对应高保真能带值。对于每个能带索引 $1 \le n \le N_b$，我们定义逐能带均方根误差

$$\mathrm{RMSE}_n := \left(\frac{1}{N_{\text{test}}\, N_k} \sum_{i \in \mathcal{I}_{\text{test}}} \sum_{\ell=1}^{N_k} \left(\omega_{h,n}^{\text{POD-DO}}(\mathbf{k}_\ell; \boldsymbol{\rho}^{(i)}, \theta^*) - \widetilde{\omega}_{h,n}(\mathbf{k}_\ell; \boldsymbol{\rho}^{(i)}) \right)^2 \right)^{1/2}, \tag{5.1}$$

相应的相对误差

$$\mathrm{relRMSE}_n := \frac{\mathrm{RMSE}_n}{\left(\frac{1}{N_{\text{test}}\, N_k} \sum_{i \in \mathcal{I}_{\text{test}}} \sum_{\ell=1}^{N_k} \widetilde{\omega}_{h,n}(\mathbf{k}_\ell; \boldsymbol{\rho}^{(i)})^2 \right)^{1/2}}, \tag{5.2}$$

以及逐能带平均相对误差

$$\mathrm{MRE}_n := \frac{100}{N_{\text{test}}\, N_k} \sum_{i \in \mathcal{I}_{\text{test}}} \sum_{\ell=1}^{N_k} \frac{\left| \omega_{h,n}^{\text{POD-DO}}(\mathbf{k}_\ell; \boldsymbol{\rho}^{(i)}, \theta^*) - \widetilde{\omega}_{h,n}(\mathbf{k}_\ell; \boldsymbol{\rho}^{(i)}) \right|}{\max\left(\left| \widetilde{\omega}_{h,n}(\mathbf{k}_\ell; \boldsymbol{\rho}^{(i)}) \right|, \tau \right)}. \tag{5.3}$$

为了衡量不同测试设计之间的离散程度，我们还针对每个 $i \in \mathcal{I}_{\text{test}}$ 定义样本误差

$$\mathrm{RMSE}_i := \left(\frac{1}{N_k N_b} \sum_{\ell=1}^{N_k} \sum_{n=1}^{N_b} \left(\omega_{h,n}^{\text{POD-DO}}(\mathbf{k}_\ell; \boldsymbol{\rho}^{(i)}, \theta^*) - \widetilde{\omega}_{h,n}(\mathbf{k}_\ell; \boldsymbol{\rho}^{(i)}) \right)^2 \right)^{1/2}, \tag{5.4}$$

$$\mathrm{relRMSE}_i := \frac{\mathrm{RMSE}_i}{\left(\frac{1}{N_k N_b} \sum_{\ell=1}^{N_k} \sum_{n=1}^{N_b} \widetilde{\omega}_{h,n}(\mathbf{k}_\ell; \boldsymbol{\rho}^{(i)})^2 \right)^{1/2}}, \tag{5.5}$$

以及样本平均相对误差

$$\mathrm{MRE}_i := \frac{100}{N_k N_b} \sum_{\ell=1}^{N_k} \sum_{n=1}^{N_b} \frac{\left| \omega_{h,n}^{\text{POD-DO}}(\mathbf{k}_\ell; \boldsymbol{\rho}^{(i)}, \theta^*) - \widetilde{\omega}_{h,n}(\mathbf{k}_\ell; \boldsymbol{\rho}^{(i)}) \right|}{\max\left(\left| \widetilde{\omega}_{h,n}(\mathbf{k}_\ell; \boldsymbol{\rho}^{(i)}) \right|, \tau \right)}, \tag{5.6}$$

这里，)式以避免除以过小的能带频率。测试集上的误差汇总于表2。对于全部十条能带，POD–DeepONet 的表现始终优于直接使用的多层感知机。尽管两种代理模型的误差均处于 10^{-3} 量级，但 POD–DeepONet 的优势具有一致性：逐能带平均相对误差从多层感知机的 0.914% 降至 POD–DeepONet 的 0.455%，且整个频谱范围内的均方根误差$_n$ 和相对均方根误差$_n$ 均更低。这表明，POD 引导的主干网络引入了有效的低秩归纳偏置，使预测受到具有物理意义的子空间约束，而非直接拟合完整的高维输出。为量化不同设计间的变异性，表3报告了逐样本汇总统计量，并列出按平均相对误差$_i$排序的代表性测试单元。POD–DeepONet 提高了平均精度，并使误差尾部更集中。具体而言，最大平均相对误差$_i$ 从多层感知机的 3.324% 降至 1.378%，最差情况下的均方根误差$_i$ 和相对均方根误差$_i$也相应更低。这种尾部表现的改善对逆向设计至关重要，因为优化过程可能会将黑箱代理模型推入约束不足的区域。图6进一步支持这一结论：对于最佳、中位数和最差测试单元，POD–DeepONet 均能很好地跟踪有限元参考解，最明显的偏差出现在交叉点附近的高阶能带以及快速变化的区段。总体而言，POD–DeepONet 在前十条 TE能带上具有更高的平均保真度和更好的可靠性，为后续逆向设计研究奠定了更坚实的基础。

5.2. 色散到结构的逆向设计

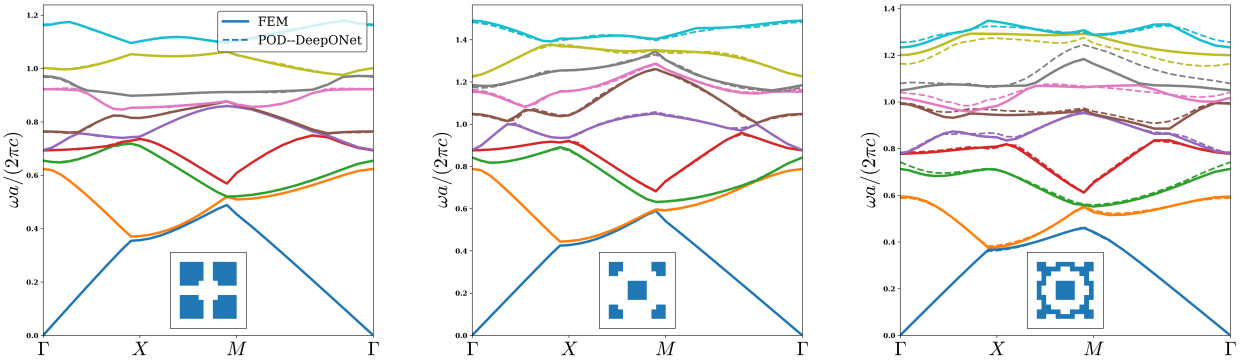
现在，我们评估逆向网络在色散到结构问题上的性能。我们沿用第 5.1 节中的同一数据库。逆向网络 G_ϕ^{disp} 采用全连接多层感知机实现，包含三个宽度为 256 的隐藏层，并使用 ReLU 激活函数。训练过程中，我们固定 POD–DeepONet 代理模型，仅通过最小化式 (4.7) 中权重为 $\gamma_{\text{bin}} = 10^{-3}$、$\gamma_{\text{sup}} = 10^{-2}$ 的经验损失函数来更新 ϕ。我们使用 Adam 优化器，

表2. POD–DeepONet 前向代理模型与多层感知机基线在测试集上的分带预测误差。

频带 n	POD–DeepONet			MLP		
	均方根误差$_n$	相对均方根误差$_n$	平均相对误差$_n$ (%)	均方根误差$_n$	相对均方根误差$_n$	平均相对误差$_n$ (%)
1	2.266×10^{-3}	6.000×10^{-3}	0.660	4.457×10^{-3}	9.837×10^{-3}	1.153
2	3.513×10^{-3}	6.613×10^{-3}	0.447	5.951×10^{-3}	1.088×10^{-2}	0.958
3	4.062×10^{-3}	6.190×10^{-3}	0.407	5.974×10^{-3}	9.141×10^{-3}	0.923
4	3.735×10^{-3}	4.866×10^{-3}	0.337	7.827×10^{-3}	8.542×10^{-3}	0.699
5	5.388×10^{-3}	5.611×10^{-3}	0.369	8.314×10^{-3}	9.267×10^{-3}	0.824
6	5.861×10^{-3}	6.069×10^{-3}	0.431	9.688×10^{-3}	1.002×10^{-2}	0.894
7	6.587×10^{-3}	6.340×10^{-3}	0.434	1.212×10^{-2}	9.953×10^{-3}	0.876
8	7.426×10^{-3}	6.696×10^{-3}	0.458	1.314×10^{-2}	1.205×10^{-2}	0.914
9	8.249×10^{-3}	7.002×10^{-3}	0.487	1.439×10^{-2}	1.253×10^{-2}	0.938
10	9.571×10^{-3}	7.596×10^{-3}	0.524	1.622×10^{-2}	1.328×10^{-2}	0.964
平均值	5.666×10^{-3}	6.298×10^{-3}	0.455	9.808×10^{-3}	1.055×10^{-2}	0.914

表3. POD–DeepONet 前向代理模型和多层感知机的逐样本汇总统计量及代表性测试单元。代表性样本按平均相对误差$_i$**排序。**

样本/统计量	POD–DeepONet			MLP		
	均方根误差$_i$	相对均方根误差$_i$	平均相对误差$_i$ (%)	均方根误差$_i$	相对均方根误差$_i$	平均相对误差$_i$ (%)
Mean	5.471×10^{-3}	5.938×10^{-3}	0.471	8.831×10^{-3}	9.554×10^{-3}	0.693
Max	1.998×10^{-2}	2.164×10^{-2}	1.378	4.544×10^{-2}	4.943×10^{-2}	3.324
Best	1.629×10^{-3}	2.001×10^{-3}	0.152	4.110×10^{-3}	4.287×10^{-3}	0.307
中位数	4.772×10^{-3}	4.592×10^{-3}	0.365	1.802×10^{-2}	8.832×10^{-3}	0.623
最差	1.736×10^{-2}	1.751×10^{-2}	1.378	2.037×10^{-2}	2.047×10^{-2}	3.324



（a）最佳测试样本　（b）中位数测试样本　（c）最差测试样本

图 6. 按 MSE_i **排序的最佳、中位数和最差测试样本。每个面板展示沿高对称路径的归一化 TE 能带结构，其中实线表示高保真度有限元法能带，虚线表示POD–DeepONet预测结果。相应的** 16×16 **晶胞以插图形式显示。**

学习率为 10^{-3}，批量大小为 512，并训练 $T = 150$ 个训练轮次。在第 $e = 1$、……、T 轮时，$H\beta$ 中的陡峭度参数设为

$$\beta(e) = \beta_0 + (\beta_{\max} - \beta_0)\Big(\frac{e}{T}\Big)^2, \qquad \beta_0 = 1,\ \beta_{\max} = 16,$$

使得 β 在训练过程中从 1 到 16 二次递增。

为进一步突出在本征正交分解风格主干网络中嵌入带隙信息的优势，我们还考虑了一个基于 MLP 的逆向设计基线，其正向代理模型采用与正向比较中相同的 MLP 架构，因此逆向性能差异主要反映主干网络信息引导表示的影响。

训练完成后，我们在测试索引集 $\mathcal{I}_{\text{test}}$ 上评估这一映射。对于每个 $i \in \mathcal{I}_{\text{test}}$，数据库提供一个参考楔形结构 $\boldsymbol{\rho}^{(i)}$ 及其对应的有限元法能带矩阵 $\mathbf{W}^{(i)}{}_h$。使用 $\mathbf{W}^{(i)}{}_h$，我们构造标准化特征向量 $\mathbf{y}^{(i)}$，并计算松弛逆向设计

$$\widehat{\boldsymbol{\rho}}^{(i)} := \boldsymbol{\rho}(\mathbf{y}^{(i)}; \phi^*, \beta) \in [0,1]^{N_f}.$$

随后，我们通过以 1/2 为阈值进行阈值化，强制实现严格二值性，

$$\boldsymbol{\rho}^{\text{bin},(i)} := \mathbf{1}_{\{\widehat{\boldsymbol{\rho}}^{(i)} > 1/2\}} \in \{0,1\}^{N_f}. \tag{5.7}$$

给定 $\boldsymbol{\rho}^{\text{bin},(i)}$，我们重建对应的基于像素的晶胞，使用高保真 FEM 求解器重新计算其 TE 能带结构，并获得归一化逆向能带矩阵 $\mathbf{W}^{\text{inv},(i)}{}_h$。根据第 5.1 节中的逐能带和逐样本误差度量，将 $\mathbf{W}^{\text{inv},(i)}{}_h$ 与 $\mathbf{W}^{(i)}{}_h$ 进行比较，从而量化逆向设计的准确性。

分频带逆向设计误差见表 5。在全部十个能带上，基于POD–DeepONet 的流程均表现出明显且稳定的优势。具体而言，平均值 RMSE_n 从基于 MLP 的方法的 3.246×10^{-2}降至基于POD–DeepONet 方法的 1.360×10^{-2}，平均值 relRMSE_n 也从 3.690×10^{-2} 降至 1.625×10^{-2}。逐能带平均相对误差的平均值也从 1.624% 降至 0.980%。这些系统性的下降表明，POD 引导的主干网络为色散目标提供了结构对齐的低秩归纳偏置，从而提升了整个频谱范围内的保真度。

为考察不同目标之间的差异，表 4 报告了逐样本汇总统计量，以及按 MRE_i排名的代表性逆向设计结果。基于POD–DeepONet 的逆向设计提高了平均精度（MRE_i从 1.615% 降至 0.989%），并收窄了误差尾部：最大值 MRE_i 从 16.232% 降至 10.578%，最差情况下的 RMSE_i 和 relRMSE_i也相应降低。这种尾部改善在实践中十分重要，因为逆向优化可能会将代理模型引向约束不足区域；POD 约束表示似乎能够稳定搜索过程。表 4 中的零误差最佳情况与在所采用的对称性和特征编码下精确恢复真实二值单胞的结果一致。

图 7 展示了这些代表性的逆向设计结果。在每个子图中，左列显示目标晶胞（上）和基于POD–DeepONet 的优化元胞（下），右侧面板则展示沿高对称路径的相应 TE 能带结构（实线：目标有限元法；虚线：逆向设计结果）。对于中位数和最差示例（图 7(b)–7(c)），优化设计与目标仅在少数像素上有所不同，但逆向设计元胞的有限元法能带仍与目标接近。随机选取的示例（图 7(d)）也呈现出类似现象。这与命题 3.2 所确立的离散能带映射连续性一致：介电常数场的微小逐像素扰动只会引起能带函数的微小变化。

总体而言，这些结果表明，尽管二元设计空间具有高度非凸性，基于 POD–DeepONet 的逆向网络仍能可靠地恢复 TE 色散图，误差仅为几个百分点。

5.3. 带隙逆向设计

基于带隙的逆向网络以与基于色散的情形类似的方式进行训练和评估。这里，输入是低维特征向量 $\mathbf{g} := (a, b, p)^\top \in \mathbb{R}^3$，用于描述能带 p 与$p + 1$之间的目标带隙区间（a，b）。一个 MLP G_ϕ^{gap}（其架构与G_ϕ^{disp} 相同）将$\mathbf{g}$ 映射为松弛楔形区域 $\boldsymbol{\rho}$；我们使用与第 5.2 节相同的 Heaviside 调度和二值化权重，最小化经验损失函数（4.9）。为分离 POD 引导的主干表示在正向约束中的贡献，我们使用直接 MLP 正向代理模型作为基线，重复执行相同的带隙反演过程。随后，将所得的二元楔形区域提升为完整元胞，并通过高保真 FEM 求解器重新计算其能带结构。

带隙目标取自第 5.1 节所使用的高保真 FEM 数据库。在 87,474 个单胞中，有 56,849 个至少包含前十条能带之间的一个带隙，总计得到 105,229 个带隙（部分晶胞贡献了多个带隙）。我们舍弃 FEM 带隙宽度 ≤ 0.01 的带隙，并针对每个剩余带隙记录能带索引 p、下边界和上边界 FEM 边缘 $g_{L,\,\mathrm{FEM}}$、$g_{U,\,\mathrm{FEM}}$以及宽度

$$w_{\mathrm{FEM}} := g_{U,\mathrm{FEM}} - g_{L,\mathrm{FEM}} > 0.01.$$

采用与第 5.1 节相同的晶胞数据库训练集/验证集/测试集划分，训练子集包含 35,858 个 $w_{\mathrm{FEM}} > 0.01$ 的带隙，验证子集包含 2,205 个此类带隙。这些带隙编码为描述符$\mathbf{g}^{(i)} = (g^{(i)}{}_{\mathrm{mid}}, w^{(i)}{}_{\text{有限元法}}, p^{(i)})$，并构成带隙到结构逆向网络的训练集和验证集。为进行定量评估，我们从测试集中随机选取$N^{\mathrm{gap}}{}_{\text{个测试集中的}} = 500$ 个 $w_{\mathrm{FEM}} > 0.01$ 带隙，并在下文的误差分析中将这些目标作为评估对象。

Ta表 4. 测试集上的逐样本逆向设计误差。代表性样本按平均相对误差排序

样本/统计量	基于POD–DeepONet			基于多层感知机		
	均方根误差	相对均方根误差	平均相对误差 (%)	均方根误差	相对均方根误差	平均相对误差 (%)
Mean	1.067×10^{-2}	1.112×10^{-2}	0.989	2.125×10^{-2}	2.473×10^{-2}	1.615
Max	1.214×10^{-1}	1.314×10^{-1}	10.578	2.073×10^{-1}	2.202×10^{-1}	16.232
Best	0	0	0	0	0	0
中位数	5.501×10^{-3}	6.172×10^{-3}	1.041	2.374×10^{-2}	3.614×10^{-2}	2.58
最差	2.292×10^{-2}	2.736×10^{-2}	10.578	1.673×10^{-1}	1.802×10^{-1}	16.232

表 5. 测试集上的分频带逆向设计误差。

频带 n	基于POD–DeepONet			基于多层感知机		
	均方根误差	相对均方根误差	平均相对误差 (%)	均方根误差	相对均方根误差	平均相对误差 (%)
1	6.442×10^{-3}	1.914×10^{-2}	0.933	1.306×10^{-2}	3.474×10^{-2}	1.774
2	1.067×10^{-2}	2.006×10^{-2}	0.988	2.112×10^{-2}	3.999×10^{-2}	1.866
3	1.151×10^{-2}	1.717×10^{-2}	0.941	2.438×10^{-2}	3.850×10^{-2}	1.791
4	1.010×10^{-2}	1.320×10^{-2}	0.855	2.401×10^{-2}	3.072×10^{-2}	1.654
5	1.337×10^{-2}	1.486×10^{-2}	0.979	3.322×10^{-2}	3.577×10^{-2}	1.188
6	1.483×10^{-2}	1.531×10^{-2}	1.034	3.487×10^{-2}	3.598×10^{-2}	1.134
7	1.487×10^{-2}	1.430×10^{-2}	1.004	3.772×10^{-2}	3.645×10^{-2}	1.179
8	1.682×10^{-2}	1.510×10^{-2}	0.983	4.019×10^{-2}	3.755×10^{-2}	1.836
9	1.861×10^{-2}	1.582×10^{-2}	1.030	4.519×10^{-2}	3.901×10^{-2}	1.901
10	1.874×10^{-2}	1.755×10^{-2}	1.051	5.080×10^{-2}	4.033×10^{-2}	1.914
平均值	1.360×10^{-2}	1.625×10^{-2}	0.980	3.246×10^{-2}	3.690×10^{-2}	1.624

训练完成后，对于测试集中的每个目标$\mathbf{g}^{(p_j)}$，我们都会获得一个优化后的楔形向量$\boldsymbol{\rho}^{(p_j)}{}_{\mathrm{inv}}$。随后，我们重建相应的晶胞，并使用高保真 FEM 求解器重新计算其能带结构。根据该能带图，我们提取频带之间的带隙

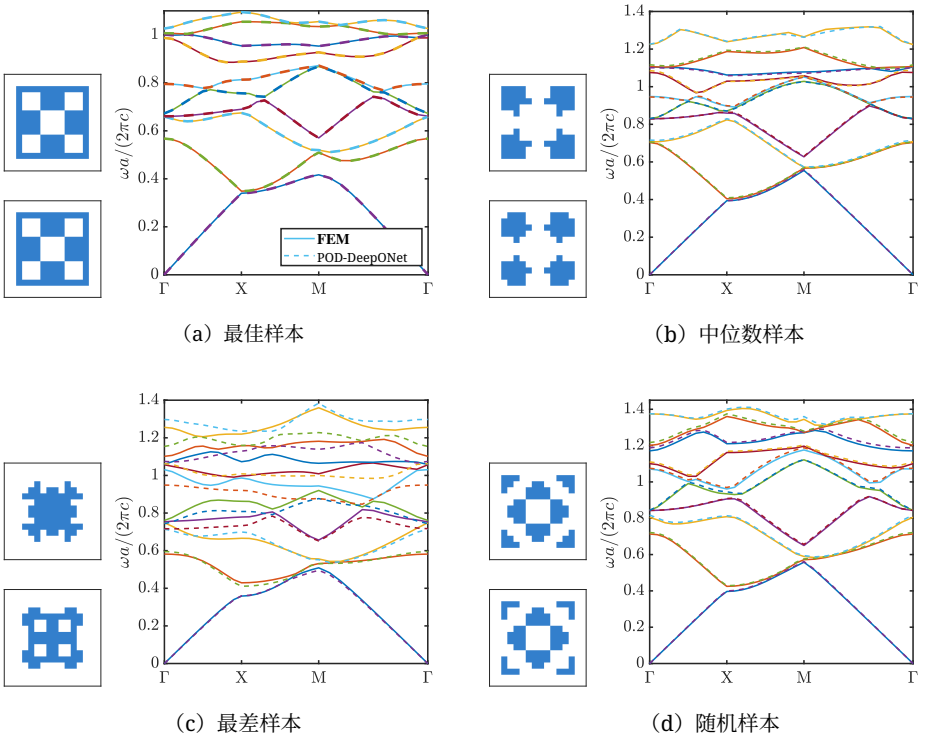
$$[g^{(p_j)}_{L,\mathrm{pred}}, g^{(p_j)}_{U,\mathrm{pred}}], \qquad w^{(p_j)}_{\mathrm{pred}} := g^{(p_j)}_{U,\mathrm{pred}} - g^{(p_j)}_{L,\mathrm{pred}},$$

p_j和$p_j + 1$。我们通过以下方式量化下边缘、上边缘及带宽的相对误差

$$e^{(p_j)}_{\mathrm{low}} = \left| \frac{g^{(p_j)}_{L,\mathrm{pred}} - g^{(p_j)}_{L,\mathrm{FEM}}}{g^{(p_j)}_{L,\mathrm{FEM}}} \right|, \quad e^{(p_j)}_{\mathrm{up}} = \left| \frac{g^{(p_j)}_{U,\mathrm{pred}} - g^{(p_j)}_{U,\mathrm{FEM}}}{g^{(p_j)}_{U,\mathrm{FEM}}} \right|, \quad e^{(p_j)}_{\mathrm{w}} = \left| \frac{w^{(p_j)}_{\mathrm{pred}} - w^{(p_j)}_{\mathrm{FEM}}}{w^{(p_j)}_{\mathrm{FEM}}} \right|, \tag{5.8}$$

并通过算术平均值对其进行总结

$$\mathrm{rMAE}_{\mathrm{low}} = \frac{1}{N^{\mathrm{gap}}_{\mathrm{test}}} \sum_{j=1}^{N^{\mathrm{gap}}_{\mathrm{test}}} e^{(p_j)}_{\mathrm{low}}, \quad \mathrm{rMAE}_{\mathrm{up}} = \frac{1}{N^{\mathrm{gap}}_{\mathrm{test}}} \sum_{j=1}^{N^{\mathrm{gap}}_{\mathrm{test}}} e^{(p_j)}_{\mathrm{up}}, \quad \mathrm{rMAE}_{\mathrm{w}} = \frac{1}{N^{\mathrm{gap}}_{\mathrm{test}}} \sum_{j=1}^{N^{\mathrm{gap}}_{\mathrm{test}}} e^{(p_j)}_{\mathrm{w}}. \tag{5.9}$$



（a）最佳样本 （b）中位数样本

（c）最差样本 （d）随机样本

图 7. 代表性逆向设计测试样本。在每个子图中，左列展示目标（上）和逆向设计（下）16×16 晶胞，右侧面板展示沿高对称路径的相应 TE 能带结构（实线：目标有限元法能带；虚线：逆向设计晶胞的能带）。

我们还考察预测的有限元法带隙与参考有限元法带隙的重合程度。对于每项实验，我们设定

$$I_{\mathrm{true}}^{(p_j)}=[g_{L,\mathrm{FEM}}^{(p_j)},g_{U,\mathrm{FEM}}^{(p_j)}],\qquad I_{\mathrm{pred}}^{(p_j)}=[g_{L,\mathrm{pred}}^{(p_j)},g_{U,\mathrm{pred}}^{(p_j)}],$$

并将它们交集和并集的长度定义为

$$\ell_{\mathrm{overlap}}^{(p_j)}=\max\Big\{0,\ \min(g_{U,\mathrm{FEM}}^{(p_j)},g_{U,\mathrm{pred}}^{(p_j)})-\max(g_{L,\mathrm{FEM}}^{(p_j)},g_{L,\mathrm{pred}}^{(p_j)})\Big\},\tag{5.10}$$

$$\ell_{\mathrm{union}}^{(p_j)}=\max(g_{U,\mathrm{FEM}}^{(p_j)},g_{U,\mathrm{pred}}^{(p_j)})-\min(g_{L,\mathrm{FEM}}^{(p_j)},g_{L,\mathrm{pred}}^{(p_j)}).\tag{5.11}$$

由此得到一个 Jaccard型重叠率和两个覆盖率，

$$R_{\mathrm{overlap}}^{(p_j)}=\frac{\ell_{\mathrm{overlap}}^{(p_j)}}{\ell_{\mathrm{union}}^{(p_j)}},\quad R_{\mathrm{cov,true}}^{(p_j)}=\frac{\ell_{\mathrm{overlap}}^{(p_j)}}{w_{\mathrm{FEM}}^{(p_j)}},\quad R_{\mathrm{cov,pred}}^{(p_j)}=\frac{\ell_{\mathrm{overlap}}^{(p_j)}}{w_{\mathrm{pred}}^{(p_j)}}.\tag{5.12}$$

我们通过其算术平均值 R_{overlap}、$R_{\mathrm{cov,\,true}}$ 和$R_{\mathrm{cov,\,pred}}$报告这些统计量。最后，如果预测带隙在较小相对容差范围内包含参考有限元法带隙，则将逆向设计判定为成功。对于给定的 $\tau_{\mathrm{gap}}>0$，我们定义

$$\chi_{\mathrm{succ}}^{(p_j)}=\mathbf{1}\Big\{g_{L,\mathrm{pred}}^{(p_j)}\le g_{L,\mathrm{FEM}}^{(p_j)}+\tau_{\mathrm{gap}}\,\big|g_{L,\mathrm{FEM}}^{(p_j)}\big|,\quad g_{U,\mathrm{pred}}^{(p_j)}\ge g_{U,\mathrm{FEM}}^{(p_j)}-\tau_{\mathrm{gap}}\,\big|g_{U,\mathrm{FEM}}^{(p_j)}\big|\Big\},\tag{5.13}$$

并在实验中取 $\tau_{\mathrm{gap}}=5\times10^{-3}$。总体有限元法级成功率为 $N_{\mathrm{succ}}/N^{\mathrm{gap}}$ 测试，其中 $N_{\mathrm{succ}}:=\sum^{N^{\mathrm{gap}}}$ 测试$_{j=1}$ $\chi^{(p_j}$）成功。

表6总结了所有N^{gap} 测试= 500 目标下逆带隙设计问题的性能。对于基于 POD–DeepONet 的流程，下带隙边缘和上带隙边缘的平均相对误差分别为 $\mathrm{rMAE}_{\mathrm{low}}=9.01\times10^{-3}$,

表6. 所有$N^{\mathrm{gap}}_{\text{测试}} = 500$目标的逆带隙设计问题汇总统计。

	基于POD–DeepONet	基于多层感知机
$\mathrm{rMAE}_{\mathrm{low}}$	9.01×10^{-3}	1.047×10^{-2}
$\mathrm{rMAE}_{\mathrm{up}}$	8.17×10^{-3}	9.288×10^{-3}
$\mathrm{rMAE}_{\mathrm{w}}$	5.78×10^{-2}	2.704×10^{-1}
$\overline{R}_{\text{重叠}}$	0.828	0.612
$\overline{R}_{\mathrm{cov,true}}$	0.874	0.675
$\overline{R}_{\mathrm{cov,\ pred}}$	0.902	0.674
$北_{\mathrm{succ}}/北^{\mathrm{gap}}_{\text{测试}}$	0.929	0.651

以及 $\mathrm{rMAE}_{\mathrm{up}} = 8.17 \times 10^{-3}$，均低于百分之一，表明逆向设计得到的带隙边缘与其有限元参考解非常接近。与基于多层感知机的对应方法相比，这些边缘误差始终较小，但更显著的优势体现在带隙宽度的恢复上：$\mathrm{rMAE}_{\mathrm{w}}$ 从 2.704×10^{-1}降至 5.78×10^{-2}，表明 POD约束逆映射能更有效地控制边缘错位的累积影响。这说明 POD 风格主干网络提供了与带隙信息在结构上相匹配的表示，相较于将完整能带数据视为通用输出，它能形成条件更优的逆向约束。平均而言，优化元胞产生的有限元带隙与目标区间和预测区间的并集重叠 82.8%（$\overline{R}_{\mathrm{overlap}} = 0.828$），覆盖目标宽度的 87.4%（$\overline{R}_{\mathrm{cov,\ true}} = 0.874$），且其自身宽度有 90.2% 位于目标区间内（$\overline{R}_{\mathrm{cov,\ pred}} = 0.902$）。这三项覆盖度指标均显著高于基于多层感知机的流程，表明 POD–DeepONet 约束更好地保留了目标频谱窗口的位置和范围。因此，大多数逆向设计满足式（5.13）中的成功判据，总体成功率达到$N_{\mathrm{succ}}/N^{\mathrm{gap}}_{\text{测试}} = 0.929$，显著高于基于多层感知机方法的 0.651 成功率。

图8对500个基于 POD–DeepONet 的逆向设计实验进行了补充可视化。每一对水平线段对应一个目标带隙：红色线段表示预设的有限元法带隙区间，黄色线段则表示由逆向设计流程生成的晶胞所对应的有限元法带隙。为便于阅读，500个样本分为两个面板：左侧面板显示样本1–250，右侧面板显示样本251–500。对于大多数目标，黄色线段与红色线段紧密重合，直观地证实了端点误差和带宽误差较小，也印证了表6所报告的高成功率。明显偏差相对少见，主要出现在窄带隙的高频区域。图9展示了逆带隙设计问题中两个随机选取的示例；尽管预测晶胞与相应的真实晶胞差异显著，其能带图仍几乎满足目标$\mathbf{g} = (a, b, p)$约束，说明逆映射并非唯一。

6. 结论

本文提出了一种 POD–DeepONet 框架，用于对具有二值、基于像素的$p4m$对称晶胞的二维光子晶体进行能带结构正向与反向分析。在正向问题中，我们从高保真能带快照中提取 POD 主干网络，并将其与分支网络结合，构建出紧凑且可微的代理模型；该模型在能带函数预测方面精度较高，平均误差约为 0.46%。基于此代理模型，我们提出了两种逆向设计策略。对于色散到结构任务，逆向网络的平均逐能带精度约为 1%，表明所提出的端到端可微公式能够可靠地恢复与目标色散数据相匹配的晶胞。此外，即使对于更具挑战性的带隙逆问题，数值结果仍令人鼓舞，表明学习得到的设计能够打开紧密覆盖指定目标区间的带隙。本研究聚焦于沿指定高对称路径的离散能带图设置，其中 POD 主干网络由固定$\mathbf{k}$网格上的快照构建。尽管这种选择能够提供紧凑、准确的降阶表示，从而实现高效的前向评估和端到端逆向优化，但这也意味着带隙识别和能带函数预测受所采用的路径和采样分辨率制约。此外，尽管我们的方法兼容训练中所用最低N_b条能带的离散排序，但当前实现并未明确处理交叉点或简并点附近的能带追踪稳定化。因此，未来工作将致力于扩展方法，以放宽这些限制。首先，我们将

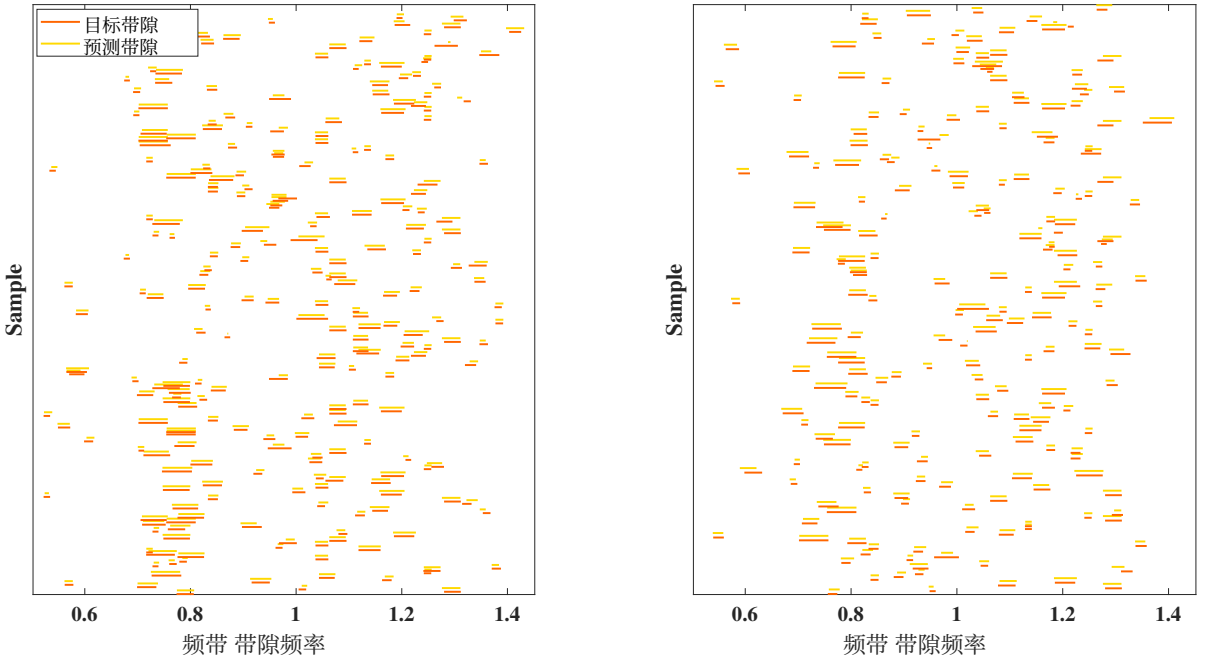


图 8. 500 个逆向设计目标的有限元法带隙区间。对于每个样本，红色线段表示指定的目标带隙，黄色线段表示逆向设计的晶胞所实现的有限元法带隙。为便于查看，样本分为两个面板：左侧为样本 1–250；右侧为样本 251–500。

开发连续**k** 或多分辨率主干网络构建方法，例如路径坐标参数化和自适应快照增广，以支持离散不变评估，并在加密或采用其他**k**采样时实现更可靠的带隙认证。其次，我们将研究能带追踪和交叉感知目标函数，以提升逆向优化中的标签一致性和梯度稳定性。第三，我们将把当前的二值、受对称性约束的参数化扩展至多材料和更高分辨率的设计空间。最后，我们将通过纳入代理模型误差指标和受制造工艺启发的扰动，评估不确定性感知逆向设计，并针对高保真求解器开展定向验证。

致谢

Guang Lin 谨向美国国家科学基金会（DMS-2533878、DMS-2053746、DMS-2134209、ECCS-2328241、CBET-2347401 和 OAC-2311848）、美国能源部科学办公室先进科学计算研究项目（DE-SC0023161）、SciDAC LEADS 研究所，以及美国能源部聚变能源科学项目（资助编号：DE-SC0024583）致谢。Guangliang Li 谨向香港研究资助局优配研究金（项目编号：17309925）致谢。

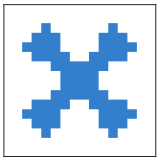
参考文献

[1] E. Yablonovitch, 《固体物理学与电子学中的自发辐射抑制》, 《物理评论快报》58 (1987) 2059。 [2] S. John, 《某些无序介电超晶格中的光子强局域化》, 《物理评论快报》58 (1987) 2486。 [3] S. G. Johnson, M. L. Povinelli, J. D. Joannopoulos, 《用于集成光学的新型光子晶体系统》, 载于：用于波分复用通信的有源与无源光学元件，第4532卷，国际光学工程学会，2001，第167–179页。 [4] J. D. Joannopoulos, S. G. Johnson, J. N. Winn, R. D. Meade, 《驾驭光的流动》, 普林斯顿大学出版社，美国新泽西州普林斯顿 [ua] (2008)。 [5] M. Notomi, 《利用强调制光子晶体操控光》, 《物理学进展报告》73 (2010) 096501。 [6] O. Painter, R. Lee, A. Scherer, A. Yariv, J. O'Brien, P. Dapkus, I. Kim, 《二维光子带隙缺陷模激光器》, 《科学》284 (1999) 1819–1821。 [7] P. A. Kuchment, 《偏微分方程的 Floquet 理论》, 第60卷，Springer Science & Business Media，1993。

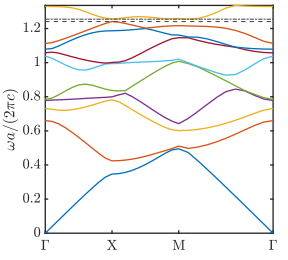
[8] K. Ho, C. T. Chan, C. M. Soukoulis, 《周期性介电结构中光子带隙的存在性》, 《物理评论快报》65 (1990) 3152。 [9] A. Taflove, S. C. Hagness, M. Piket-May, 《计算电磁学：有限差分时域法》, 《电气工程手册》3 (2005) 15。 [10] M. Qiu, S. He, 《计算含介电或金属夹杂的二维光子晶体中缺陷模的数值方法》, 《物理评论B》61 (2000) 12871。 [11] W. Axmann, P. Kuchment, 《计算光子与声学带隙材料谱的高效有限元法：I. 标量情形》, 计算物理学杂志150 (1999) 468–481。 [12] I. Andonegui, A. J. Garcia-Adeva, 《有限元法在二维光子晶体与谐振腔研究中的应用》, 《光学快报》21 (2013) 4072–4092。 [13] W. Setyawan, S. Curtarolo, 《高通量电子能带结构计算：挑战与工具》, 计算材料科学49 (2010) 299–312。 [14] E. Degir menci, P. Landais, 《太赫兹频率下二维金属光子晶体带隙与透射的有限元分析》, 应用光学52 (2013) 7367–7375。 [15] J. Li, J. Liu, S. A. Baronett, M. Liu, L. Wang, R. Li, Y. Chen, D. Li, Q. Zhu, X.-Q. Chen, 《拓扑声子材料的计算与数据驱动发现》, 《自然·通讯》12 (2021) 1204。 [16] R. K. Cersonsky, J. Antonaglia, B. D. Dice, S. C. Glotzer, 《三维光子晶体的多样性》, 《自然·通讯》12 (2021) 2543。 [17] Y. Wang, R. Craster, G. Li, 《用于三维光子晶体色散关系重构的 hp 自适应采样算法》, 计算物理学杂志521 (2025) 113572。 [18] O. Sigmund, J. Søndergaard Jensen, 《通过拓扑优化系统设计声子带隙材料与结构》, 《伦敦皇家学会哲学汇刊A辑：数学、物理与工程科学》361 (2003) 1001–1019。 [19] H. Men, K. Y. Lee, R. M. Freund, J. Peraire, S. G. Johnson, 《三维光子晶体带隙结构的鲁棒拓扑优化》, 《光学快报》22 (2014) 22632–22648。 [20] A. Dalklint, M. Wallin, K. Bertoldi, D. Tortorelli, 《通过拓扑优化设计可调声子带隙材料》, 《固体力学与物理学杂志》163 (2022) 104849。 [21] X. Zhang, J. Xing, P. Liu, Y. Luo, Z. Kang, 《通过非梯度拓扑优化实现声学超材料的完整带隙与定向带隙设计》, 《极端力学快报》42 (2021) 101126。 [22] Z. Jia, Y. Bao, Y. Luo, D. Wang, X. Zhang, Z. Kang, 《通过拓扑优化最大化声子晶体中的声学带隙》, 《国际机械科学杂志》270 (2024) 109107。 [23] C. Y. Kao, S. Osher, E. Yablonovitch, 《利用水平集方法最大化二维光子晶体中的带隙》, 《应用物理B辑》81 (2005) 235–244。 [24] X.-l. Cheng, J. Yang, 《正方晶格二维光子晶体中带隙的最大化》, 《美国光学学会杂志A辑》30 (2013) 2314–2319。 [25] J. Peurifoy, Y. Shen, L. Jing, Y. Yang, F. Cano-Renteria, B. G. DeLacy, J. D. Joannopoulos, M. Tegmark, M. Solja˘ci´c, 《利用人工神经网络进行纳米光子粒子模拟与逆向设计》, 《科学进展》4 (2018) eaar4206。 [26] M. H. Tahersima, K. Kojima, T. Koike-Akino, D. Jha, B. Wang, C. Lin, K. Parsons, 《集成光子功率分束器的深度神经网络逆向设计》, 《科学报告》9 (2019) 1368。 [27] C. Qiu, X. Wu, Z. Luo, H. Yang, G. He, B. Huang, 《基于不平衡数据集知识迁移的深度神经网络纳米光子逆向设计》, 《光学快报》29 (2021) 28406–28415。 [28] W. Jiang, Y. Zhu, G. Yin, H. Lu, L. Xie, M. Yin, 《基于深度学习的弹性超材料色散关系预测与结构逆向设计》, 《今日材料：物理》22 (2022) 100616。 [29] X. Li, S. Ning, Z. Liu, Z. Yan, C. Luo, Z. Zhuang, 《通过基于深度学习的数据驱动方法设计具有预期带隙的声子晶体》, 《计算方法在应用力学与工程中的应用》361 (2020) 112737。 [30] S. Han, Q. Han, C. Li, 《基于深度学习的声子晶体逆向设计，以实现预期的波衰减》, 《应用物理学杂志》132 (2022)。 [31] Y. Wang, R. Craster, G. Li, 《通过深度学习预测二维光子晶体的能带结构》, arXiv 预印本 arXiv:2411.06063 (2024)。 [32] B. Ma, R. Hao, H. Yan, H. Jiang, J. Chen, K. Tang, 《二维光子晶体完整光子带隙的深度学习逆向设计》, 《当代纳米科学》19 (2023) 423–431。 [33] X.-H. Wan, Y. Zhang, Q.-H. Guo, L.-Y. Zheng, 《基于深度学习的不规则声子晶体逆向设计》, 《国际机械科学杂志》(2025) 110335。 [34] T. V. Tran, S. Nanthakumar, X. Zhuang, 《用于按需逆向设计具有任意目标带隙的超材料的深度学习框架》, 《npj人工智能》1 (2025) 2。 [35] J. Song, J. Lee, N. Kim, K. Min, 《人工智能在创新超材料设计中的应用：综述》, 《国际精密工程与制造杂志》25 (2024) 225–244。 [36] R. Deng, W. Liu, L. Shi, 《光子晶体的逆向设计》, 《纳米光子学》13 (2024) 1219–1237。 [37] Z. Chen, A. Ogren, C. Daraio, L. C. Brinson, C. Rudin, 《如何利用可解释机器学习发现超材料中的隐藏模式》, 《极端力学快报》57 (2022) 101895。 [38] L. Lu, P. Jin, G. Pang, Z. Zhang, G. E. Karniadakis, 《基于算子通用逼近定理的 DeepONet 非线性算子学习》, 《自然·机器智能》3 (2021) 218–229。 [39] N. Kovachki, Z. Li, B. Liu, K. Azizzadenesheli, K. Bhattacharya, A. Stuart, A. Anandkumar, 《神经算子：学习函数空间之间的映射及其在偏微分方程中的应用》, 《机器学习研究杂志》24 (2023) 1–97。 [40] H. Eivazi, S. Wittek, A. Rausch, 《算子学习的非线性降阶》, arXiv 预印本 arXiv:2403.18735 (2024)。 [41] Q. Cheng, M. H. Sahadath, H. Yang, S. Pan, W. Ji, 《利用傅里叶基-深度算子网络及不确定性量化，对流动波动条件下的热传递进行代理建模》, 《核能进展》188 (2025) 105895。 [42] Y. Wang, G. Lin, 《针对边界数据和源数据独立变化的参数化偏微分方程的降阶基深度算子学习》, arXiv 预印本 arXiv:2511.18260 (2025)。 [43] N. Demo, M. Tezzele, G. Rozza, 《用于降阶建模中残差学习的 DeepONet 多保真方法》, 《先进建模》

以及《工程科学中的模拟》10 (2023) 12. [44] L. Lu, R. Pestourie, S. G. Johnson, G. Romano，用于高效学习偏微分方程的多保真深度神经算子及其在纳米尺度热传输快速逆向设计中的应用，《物理评论研究》4 (2022) 023210. [45] Y. Wang, G. Li，基于多项式插值的二维光子晶体色散关系重构，计算物理学杂志 498 (2024) 112659. [46] Y. Wang, G. Li，二维光子晶体色散关系重构的 hp自适应采样算法，arXiv 预印本 arXiv:2311.16454 (2023). [47] G. Bao, L. Cowsar, W. Masters，光学科学中的数学建模，SIAM，2001. [48] J. D. Jackson，经典电动力学，1999. [49] C. Kittel, P. McEuen，固体物理学导论，John Wiley & Sons，2018. [50] I. M. Glazman，奇异微分算子的定性谱分析直接方法，第2146卷，以色列科学翻译计划，1965.[51] J. D. Joannopoulos, P. R. Villeneuve, S. Fan，光子晶体，《固体通信》102 (1997) 165–173. [52] S. G. Johnson, J. D. Joannopoulos，光子晶体：从理论走向实践，Springer Science & Business Media，2001. [53] S. G. Johnson, S. Fan, P. R. Villeneuve, J. D. Joannopoulos, L. Kolodziejski，光子晶体板中的导模，《物理评论B》60 (1999) 5751. [54] Y. Akahane, T. Asano, B.-S. Song, S. Noda，二维光子晶体中的高Q值光子纳米腔，《自然》425 (2003) 944–947. [55] L. P. Bouckaert, R. Smoluchowski, E. Wigner，晶体中布里渊区理论与波函数的对称性，《物理评论》50 (1936) 58. [56] D. Schattschneider，平面晶体学群：识别与表示，《美国数学月刊》85 (1978) 439–450. [57] D. Castelló-Lurbe, V. Torres-Company, E. Silvestre，硅波导中的逆向色散工程，《美国光学学会杂志B辑》31 (2014) 1829–1835. [58] D. C. Dobson, S. J. Cox，最大化二维光子晶体中的带隙，《SIAM应用数学杂志》59 (1999) 2108–2120. [59] G. H. Golub, C. F. Van Loan，矩阵计算，约翰·霍普金斯大学出版社，2013. [60] C. Eckart, G. Young，用低秩矩阵逼近另一个矩阵，《心理计量学》1 (1936) 211–218. [61] R. Bhatia，矩阵分析，第169卷，Springer Science & Business Media，2013. [62] G. Cybenko，通过S形函数的叠加进行逼近，《控制、信号与系统数学》2 (1989) 303–314. [63] K. Hornik，多层前馈网络的逼近能力，《神经网络》4 (1991) 251–257.

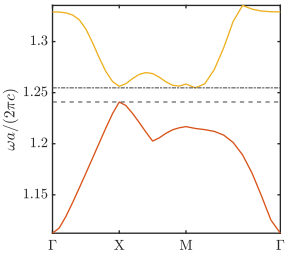
示例 A（目标 $\mathbf{g}$ $=(1.2409, 1.2547, 9)$）



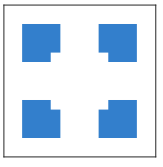
(a) 真实晶胞



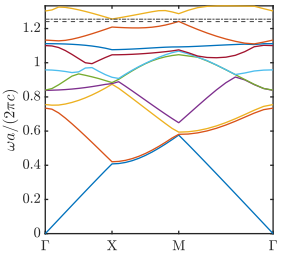
(b) 真实能带



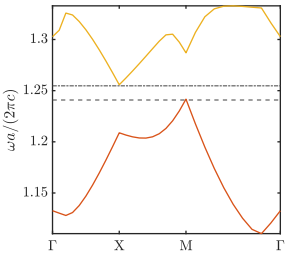
(c) 真实放大图



(d) 预测晶胞

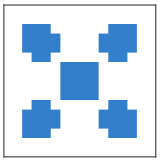


(e) 预测能带

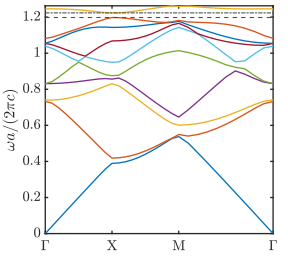


(f) 预测放大图

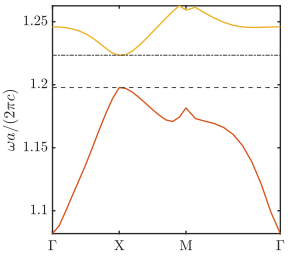
示例 B（目标 $\mathbf{g}$ $=(1.1978, 1.2234, 9)$）



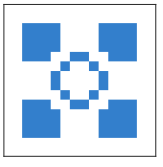
(g) 真实晶胞



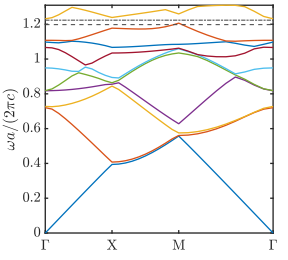
(h) 真实能带



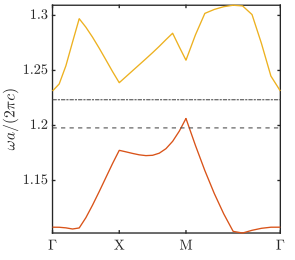
(i) 真实放大图



(j) 预测晶胞



(k) 预测能带



(l) 预测放大图

图9. 逆带隙设计问题的代表性结果。每个案例均比较真实与预测晶胞、沿 $\mathcal{K}_{\mathrm{hs}}$ 的相应能带图，以及目标带隙附近的放大图。