arXiv:2602.01737v1 [physics.flu-dyn] 2 Feb 2026

用于参数化偏微分方程的物理信息切比雪夫多项式神经算子

Biao Chen、Jing Wang*、Hairun Xie、Qineng Wang、Shuai Zhang、Yifan Xia、Jifa Zhang*

2026年2月3日

摘要

神经算子已成为强大的深度学习框架，可用于近似参数化偏微分方程（PDE）的解算子。然而，当前方法主要依赖多层感知机（MLP）将输入映射为解；由于固有的谱偏差和固定的激活函数，这会削弱其在物理信息设置下的训练稳健性。为克服架构上的局限，我们提出物理信息切比雪夫多项式神经算子（CPNO）这一新型无网格框架。该框架利用基变换，以数值稳定的切比雪夫谱基取代不稳定的单项式展开。通过将参数依赖调制机制整合到主网络中，CPNO 能够在近乎最优的函数空间中构造偏微分方程解，使模型摆脱对多层感知机的特定依赖，并增强多尺度表示能力。理论分析表明，切比雪夫基具有近极小极大一致逼近性质和优越的条件性，且勒贝格常数随阶数呈对数增长，从而减轻谱偏差，并确保优化过程中的梯度流稳定。针对基准参数化偏微分方程开展的数值实验表明，CPNO 具有更高的精度、更快的收敛速度，并对超参数表现出更强的稳健性。跨声速翼型流动实验则展示了 CPNO 表征复杂几何问题的能力。

关键词： 偏微分方程；神经算子；切比雪夫多项式；调制；物理信息神经网络

1 引言

偏微分方程（PDE）是描述众多科学与工程学科中复杂现象的基础数学工具 [1, 2, 3, 4, 5, 6]„ 利用神经网络求解偏微分方程已广泛应用于流体力学 [7], 、固体力学 [8], 、材料科学 [9], 、地球科学 [10], 、生物医学工程 [11], 和金融数学 [12]。参数化偏微分方程描述了系统行为如何响应物理参数、几何构型、边界条件或初始状态的变化，这在气候建模 [13] 和航空航天工程 [14]等应用中尤为重要。然而，这些方程的解通常呈现多尺度特征和复杂模式，例如空间局域现象与参数依赖分岔。因此，高效且准确地求解偏微分方程对于理解、预测和优化这些复杂系统至关重要 [15, 16]。

深度学习的发展催生了多种基于神经网络的偏微分方程求解策略。这些偏微分方程求解方法可沿着两个基本维度自然分类：训练是否依赖大量预计算解数据集，以及目标是求解单个偏微分方程还是一族参数化偏微分方程 [17, 18, 19, 20]。需要大型监督数据集的方法包括针对单个问题的经典物理信息神经网络，以及用于参数化问题族的数据驱动神经算子，例如 DeepONet [21] 和傅里叶神经算子（FNO） [22]。相比之下，纯物理信息方法完全不使用解数据，而是将控制方程、边界条件和初始条件直接嵌入损失函数 [23], ；这一分支涵盖了针对单个问题的多种改进型 PINN [24, 25]。

此外，物理信息神经算子，例如 PINO [26], 物理信息 DeepONet [27], 元自编码器 [28], 能够在没有任何标注解的情况下学习解算子 [29]。图1以象限图的形式展示了这一领域，并将代表性方法置于相应位置。本研究聚焦于最具挑战性、但在科学上也最有价值的象限：仅凭物理损失学习参数化解算子。

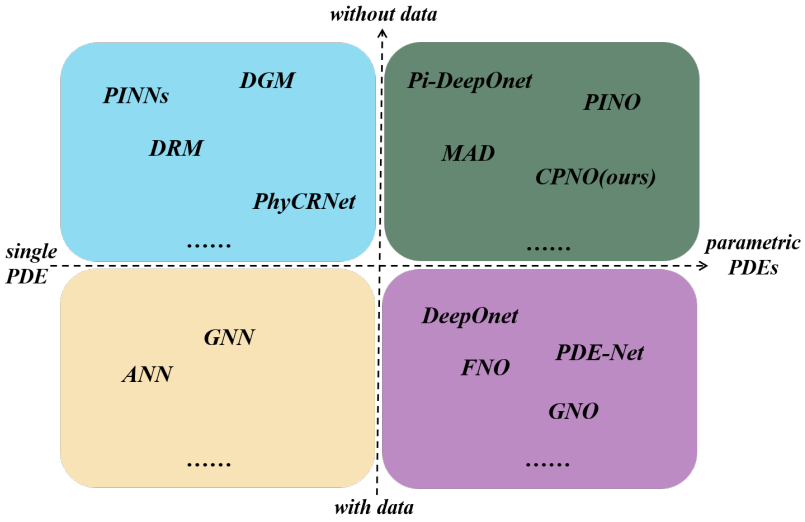

图1：根据处理单一偏微分方程与参数化偏微分方程的能力以及数据需求，对偏微分方程求解方法进行分类。

在实践中，参数化偏微分方程的物理信息算子学习仍以难度极高而著称。近期对训练动态的研究表明，在纯物理监督下，损失景观通常会呈现尖锐、高曲率的沟谷，并伴随偏微分方程残差中持续存在的高频噪声 [30, 31, 32]。这种病态行为根本上源于人们广泛依赖采用固定逐点非线性的多层感知机。这类架构存在明显的谱偏差，倾向于捕捉低频分量，却难以表示许多偏微分方程解所具有的陡峭梯度和多尺度特征 [33, 34]。大量研究尝试通过在损失设计或优化层面采取补救措施来缓解这些困难。典型例子包括因果加权损失和索伯列夫范数损失 [35, 25], 、基于残差的自适应细化和重要性采样 [36, 37], 、序列到序列训练计划 [38], ，以及对抗式或变分形式 [39]。尽管这些技术在特定基准测试上能够带来显著提升，但它们本质上仍是事后修正，并未触及底层架构的局限：传统多层感知机所张成的函数空间天生容易出现严重的梯度失衡和优化刚性。一种替代且更直接的策略是彻底摒弃逐点多层感知机，转而显式地在多项式基上参数化解算子 [40]。虽然这种方法规避了多层感知机的谱偏差，但采用单项式基的朴素实现会重新引入高次多项式插值中众所周知的数值病态，其特征是 Gram矩阵的条件数随次数呈指数增长，且对系数扰动极为敏感 [41, 42, 43, 44, 45]。

为了实现本质上稳定且高精度的物理信息算子学习，需要一种架构先验，它在保留多项式通用逼近能力的同时，从一开始就对数值条件性施加严格控制——从而完全避开对复杂损失重加权或自适应采样策略的需求。逼近理论以第一类切比雪夫多项式的形式，为这一挑战提供了决定性的解决方案。在参考区间 $[-1, 1]$上，该基是唯一的极小极大多项式族：对于任意固定的多项式次数p，它在该次数的所有多项式中给出相对于最佳逼近的最小可能最大偏差 [46, 47, 48, 49]。至关重要的是，其余弦分布根使节点在定义域边界附近聚集，从而有效消除困扰等距性或基于单项式的插值的龙格现象。作为直接结果，相关联的条件数在离-

散系统中仅表现为 $O(p^2)$，并随多项式次数p呈现极为温和的增长，相较指数级 $O(e^p)$爆炸，后者由单项式基 [50, 51, 52, 53]所呈现。通过直接基于切比雪夫谱基构建神经算子，并辅以轻量的参数依赖调制机制，所提出的切比雪夫多项式神经算子（CPNO）在可证明条件良好且谱均衡的函数空间中作为先验运行。这一基础性架构决策从根源上消除了训练病态的主要来源，从而在纯粹基于物理的监督下实现稳健收敛，并在可用时从容纳入稀疏观测数据，正如第2.5节所严格证明的那样。

基于这些见解，我们提出了一种新颖的物理信息切比雪夫多项式神经算子（CPNO）。本文的主要贡献总结如下：

- 我们成功构建了一个稳健且高效的切比雪夫多项式神经算子框架。该框架系统地在近乎最优且数值稳定的函数基中构造解的表示。这为设计范式开辟了新思路，突破了传统固定激活函数的局限。

- 我们对切比雪夫基的一致逼近性质和数值稳定性进行了严谨的理论分析。该分析阐明了这些性质如何直接应对并解决现有神经算子中普遍存在的谱偏差和训练不稳定性问题。

- 通过在多个基准参数化偏微分方程上进行大量数值实验，我们证明，CPNO在解的精度、收敛速度和稳健性方面显著优于当前的物理信息无网格神经算子方法。

作为一种物理信息、无网格方法，CPNO旨在为具有挑战性的参数化系统提供高效、准确且稳健的求解途径。本文其余部分安排如下：第2节介绍参数化偏微分方程的算子学习，并详细说明模型架构。第3节介绍切比雪夫多项式的理论优势。第4节展示并分析数值实验结果，包括与基线方法的比较及消融研究。第5节总结全文并讨论未来的研究方向。

2 方法

本节详细介绍物理信息切比雪夫多项式神经算子（CPNO），这是一种采用切比雪夫谱特征来逼近参数化偏微分方程解的无网格架构。文中描述了模型架构的核心组件以及物理信息训练范式。

2.1 问题定义

本小节定义了该框架所处理的参数化偏微分方程的一般类别，并将待逼近的解算子加以形式化。考虑一类定义在时间域 $[0,T]$ 和空间域 $\Omega \subset \mathbb{R}^d$ 上的一般参数化偏微分方程（PDE）。目标是确定一个解 u：$[0,T]\times\Omega \to \mathbb{R}$，其位于合适的巴拿赫空间 V中。该系统由取自参数空间 Θ的参数 θ 支配。对于给定的 $\theta \in \Theta$，解$u(\cdot;\theta) \in V$满足：

$$\begin{cases} \mathcal{L}(\theta)u(x;\theta) = f(x;\theta), & x \in (0,T]\times\Omega, \\ \mathcal{B}(\theta)u(x;\theta) = g(x;\theta), & x \in (0,T]\times\partial\Omega, \\ \mathcal{I}(\theta)u(x;\theta) = u_0(x;\theta), & x \in \{0\}\times\Omega, \end{cases} \tag{1}$$

其中， $\mathcal{L}(\theta)$、 $\mathcal{B}(\theta)$ 和 $\mathcal{I}(\theta)$ 分别表示微分算子、边界算子和初始算子，它们可能是非线性的，并且依赖于参数。参数空间 Θ 被构造为一个积空间，以统一各种变化来源： $\theta = (a, f, g, u_0)$，其中， $a \in \mathcal{A}$ 对 $\mathcal{L}$的系数进行参数化， $\mathcal{L}$为 Ω 上的函数空间； $\mathcal{A}$为 Ω 上的函数空间； $f \in \mathcal{F}(L^2((0, T) \times\Omega)$ 表示源项， $g \in \mathcal{G}(L^2((0, T) \times\partial\Omega))$ 表示狄利克雷边界； $u_0 \in \mathcal{U}_0$为 Ω 上的函数空间。假设空间 $\Theta = \mathcal{A}\times\mathcal{F}\times\mathcal{G}\times\mathcal{U}_0$ 具有

紧致性和度量结构，以确保解映射的稳定性和连续性。算子学习的核心任务是逼近解算子 $\mathcal{S}$，该算子将每个参数实例 $\theta \in \Theta$ 映射到其唯一的弱解 $u(\cdot;\theta) \in V$：

$$\mathcal{S} : \Theta \to V, \quad \theta \mapsto \mathcal{S}(\theta) = u(\cdot;\theta). \tag{2}$$

在偏微分方程适定的条件下，$\mathcal{S}$ 连续且定义良好。神经算子构造参数化近似 $\mathcal{S}_\phi \approx \mathcal{S}$，其中 ϕ 表示可训练的网络参数。

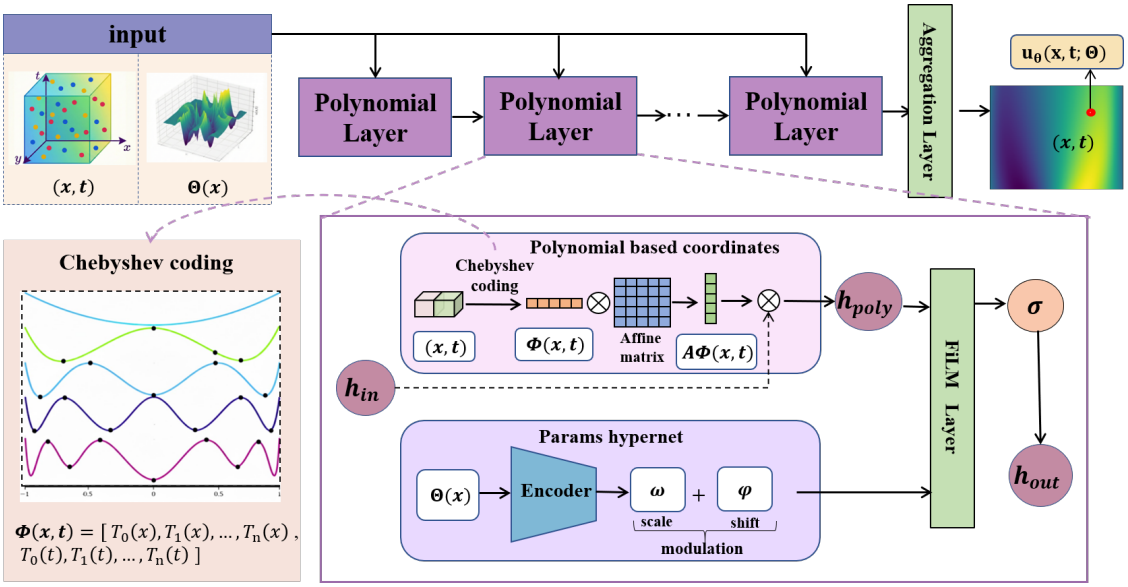

图 2: 物理信息切比雪夫多项式神经算子（CPNO）的主要框架)

2.2 多项式堆叠神经算子框架

我们提出一种基于多项式神经算子的多层堆叠框架，以高效求解参数化偏微分方程。所提出的 CPNO 架构如图 2所示。该框架以时空坐标（x, t）$\in \Omega \times [0, T]$ 和参数 $\theta \in \Theta$ 作为输入，通过逐层变换迭代构建解 $u(x, t; \theta)$。该方法以多项式逼近理论为基础，利用高阶多项式空间的非线性映射，继承了魏尔斯特拉斯定理的稠密性，并适用于捕捉偏微分方程的多尺度特征。其模块化设计支持灵活调整多项式形式，从理论上能够逼近任意连续解，并具有稳健的泛化能力。

所提出的框架是一种多层神经算子架构，由 q 变换映射 $\{\mathcal{F}_j\}_{j=1}^{q}$构成，旨在逼近参数化偏微分方程（PDE）的解。该过程从 $\mathbf{h}^{(0)} = \Phi(x, t)$ 开始，其中 $\Phi(x, t)$ 作为初始隐藏状态，通过切比雪夫谱特征对时空坐标（x, t）进行编码。各层的演化由以下递归关系控制：

$$\mathbf{h}^{(j)} = \mathcal{F}_j(\mathbf{h}^{(j-1)}; \theta), \quad j = 1, 2, \ldots, q, \tag{3}$$

其中，θ 表示影响算子行为的全局偏微分方程参数。最终解通过线性投影获得：

$$u(x, t; \theta) = \mathbf{W}_q \mathbf{h}^{(q)} + b_q, \tag{4}$$

其中，$\mathbf{W}_q \in \mathbb{R}^{d_u \times d_h}$ 和 $b_q \in \mathbb{R}^{d_u}$ 分别表示权重矩阵和偏置向量，用于将终端隐藏状态 $\mathbf{h}^{(q)} \in \mathbb{R}^{d_h}$ 映射到输出空间 $\mathbb{R}^{d_u}$。

每个变换映射 $\mathcal{F}_j$ 定义为参数依赖的非线性映射：

$$\mathcal{F}_j(\mathbf{h}^{(j-1)}; \theta) = \sigma\left(\mathbf{f}_m^{\theta}\left(\mathbf{h}^{(j-1)}\right)\right), \tag{5}$$

其中 $\mathbf{f}_m^\theta$ 表示依赖于 θ的调制函数，而 σ 是非线性激活函数，有助于在各层之间自适应地调整隐藏状态。依次应用这些算子可得到复合变换：

$$\mathbf{h}^{(q)} = \mathcal{F}_q \circ \mathcal{F}_{q-1} \circ \cdots \circ \mathcal{F}_1(\Phi(x,t);\theta), \tag{6}$$

与线性投影结合后，即可得到解：

$$u(x,t;\theta) = \mathbf{W}_q \left(\mathcal{F}_q \circ \mathcal{F}_{q-1} \circ \cdots \circ \mathcal{F}_1(\Phi(x,t);\theta)\right) + b_q. \tag{7}$$

这种分层组合有效地构建了多项式结构，其显式表示将递归展开如下：

$$u(x,t;\theta) = \mathbf{W}_q \left(\sigma\left(\mathbf{f}_m^\theta\left[\sigma\left(\cdots\sigma\left(\mathbf{f}_m^\theta\left(\Phi(x,t)\right)\right)\cdots\right)\right]\right)\right), \tag{8}$$

其中，嵌套应用 $\sigma \circ \mathbf{f}_m^\theta$ 共进行 q 次，从切比雪夫编码后的输入 Φ（x，t）开始。根据魏尔斯特拉斯逼近定理，切比雪夫多项式在 $[-1,1]$ 上具有稠密性。由此得到的多项式展开使该框架能够通过控制深度 q 和参数优化，以可控精度逼近一大类偏微分方程解。

2.3 多项式层：切比雪夫基与调制变换

在所提出的框架中，第 j 层变换算子 $\mathcal{F}_j$ 是整体多层堆叠结构中的关键组成部分，它根据 $\mathbf{h}^{(j)} = \mathcal{F}_j$（$\mathbf{h}^{(j-1)}$；$\theta$），将前一层的输出 $\mathbf{h}^{(j-1)}$ 映射为$\mathbf{h}^{(j)}$。该层采用切比雪夫谱编码与调制变换协同设计，以捕捉参数化偏微分方程（PDE）解的多尺度特征，从而提升学习效率和数值稳定性。

切比雪夫谱基该子层将低维时空坐标 $(x,\ t)$ 与 $x = (x_1, \ldots\ldots, x_d) \in \Omega$ 转换为高维谱特征表示，以解析多尺度特性。每个空间坐标分量和时间坐标都经过仿射变换，归一化到切比雪夫多项式的标准区间 $[-1,1]$：

$$x_i' = 2\frac{x_i - x_{i,\min}}{x_{i,\max} - x_{i,\min}} - 1, \quad t' = 2\frac{t - t_{\min}}{t_{\max} - t_{\min}} - 1 \tag{9}$$

其中，$x_{i,\ \min}$ 和 $x_{i,\ \max}$ 分别表示 Ω 在第 i 个维度上的边界，$t_{\min}$ 和 $t_{\max}$ 表示 $[0,T]$的边界。编码采用第一类切比雪夫多项式 $T_k(z)$ 的三项递推关系：

$$T_0(x) = 1, \quad T_1(x) = x, \quad T_{k+1}(x) = 2xT_k(x) - T_{k-1}(x), \quad k \geq 1, \tag{10}$$

对每个归一化坐标 x_i' 和 t'计算至 P 阶的值。随后，将各分量对应的序列拼接起来，形成谱特征向量 $\Phi(x,\ t)$：

$$\Phi(x,t) = [T_0(x_1'), \ldots, T_P(x_1'), \ldots, T_0(x_d'), \ldots, T_P(x_d'), T_0(t'), \ldots, T_P(t')]^T \in \mathbb{R}^{(d+1)(P+1)}. \tag{11}$$

切比雪夫多项式的正交性 $(\int_{-1}^{1} T_j(x)T_k(x)(1-x^2)^{-1/2}dx = \delta_{jk}/\pi)$ 可确保数值稳定性，而递归计算减轻了网络从原始坐标中学习基函数的负担，正如先前研究所指出的。特征向量 $\Phi(\mathbf{x}\ t)$ 作为后续调制子层的输入，可显著增强多尺度表示。

调制变换 该子层通过递归组合上一层的隐藏状态$\mathbf{h}^{(l-1)}$与输入的切比雪夫谱特征 $\Phi(\mathbf{x}\ t)$，系统地构造高阶多项式项。首先，通过可学习线性变换$\mathbf{A}_l$将 $\Phi(\mathbf{x}\ t)$映射到与隐藏状态维度相同的空间。随后，将结果与上一层的隐藏状态进行哈达玛积运算，以提高多项式次数：

$$\mathbf{h}^{(j)}(\mathbf{x},t) = (\mathbf{A}\Phi(\mathbf{x},t)) \odot \mathbf{h}^{(j-1)}(\mathbf{x},t), \tag{12}$$

算法1 CPNO 的前向与反向传播

需要： 时空坐标 $\mathbf{x} \in \mathbb{R}^d$，偏微分方程参数 θ，映射网络 ϕ，合成网络参数 $\psi = \{A_j, \omega_j, \phi_j\}_{j=1}^{q}$，$W_q$，物理信息损失 $\mathcal{L}$

确保： 预测解 $u(\mathbf{x};\theta)$，更新后的参数 ϕ，ψ

前向传播：

$x' \leftarrow$ 归一化(x)，$t' \leftarrow$ 归一化(t) ▷ 映射到 $[-1,1]^d$和 $[-1,1]$

$\Phi(x,t) \leftarrow \bigoplus_{i=1}^{d}[T_0(x_i')\ldots,T_P(x_i')] \oplus [T_0(t')\ldots,T_P(t')]$ ▷ 切比雪夫特征

$\mathbf{h}^{(0)} \leftarrow \Phi(x,t) \in \mathbb{R}^{d_h}$ ▷ 初始化隐藏状态

对 j = 1至 q 执行

$\mathbf{h}^{(j)}_{\text{多项式}} \leftarrow (A_j\Phi(x,t)) \odot \mathbf{h}^{(j-1)}$ ▷多项式项

$m_j \leftarrow M_{\varphi,j}(\theta)$ ▷ 调制向量

$\omega_j \leftarrow \omega_{,j}(m_j)$; $\phi_j \leftarrow \phi_{,j}(m_j)$ ▷ 缩放与平移

$\mathbf{h}^{(j)} \leftarrow \sigma(\omega_j \odot \mathbf{h}^{(j)}_{\text{多项式}} + \phi_j)$ ▷ 调制激活

结束循环

$u(x,t;\theta) \leftarrow W_q\mathbf{h}^{(q)}$ ▷ 投影到解空间

其中$\mathbf{h}^{(0)}$ 通常初始化为常数向量。然后，通过参数 θ对这一递归生成、包含丰富高阶交叉项的特征$\mathbf{h}^{(j)}$进行动态调制。

调制使用振幅向量 $\boldsymbol{\omega}_l(\theta) \in \mathbb{R}^{d_h}$ 和相位向量 $\boldsymbol{\varphi}_l(\theta) \in \mathbb{R}^{d_h}$，它们由参数映射网络生成，随后应用非线性激活函数 σ:

$$\mathbf{h}^{(j)}(\mathbf{x},t;\theta) = \sigma\left(\boldsymbol{\omega}_l(\theta) \odot \mathbf{h}^{(j)}(\mathbf{x},t) + \boldsymbol{\varphi}_j(\theta)\right) \tag{13}$$

这里，我们采用高斯误差线性单元（GELU）作为激活函数 σ。这一设计将多项式次数的扩展与参数调制解耦，使网络能够在结构化且逐步丰富的特征空间中，自适应地调整其表示，以适配特定的偏微分方程解。该设计具有灵活性，其中 $\Phi(\mathbf{x}\ t)$ 分量可以替换为勒让德基等其他谱编码，而使用 $\boldsymbol{\omega}_l(\theta)$ 和$\boldsymbol{\varphi}_l(\theta)$ 进行调制则保留了其核心的参数自适应特性。

基于上述讨论，我们将前向算法总结于算法1。

2.4 优化目标和损失函数

本小节概述了用于训练 CPNO 参数的优化目标和损失函数，以确保满足物理约束并纳入可用数据。CPNO 的可训练参数 ϕ 通过最小化复合损失函数 $\mathcal{L}(\phi)$进行优化。该函数确保符合偏微分方程系统的物理定律，并在有观测数据时将其纳入其中。总损失表示为以下四个分量的加权和:

$$\mathcal{L}(\phi) = \lambda_{\text{pde}}\mathcal{L}_{\text{pde}}(\phi) + \lambda_{\text{ic}}\mathcal{L}_{\text{ic}}(\phi) + \lambda_{\text{bc}}\mathcal{L}_{\text{bc}}(\phi) + \lambda_{\text{data}}\mathcal{L}_{\text{data}}(\phi), \tag{14}$$

其中，λ_{pde}、λ_{ic}、λ_{bc} 和 λ_{data} 表示用户定义的超参数，用于平衡各项的贡献。偏微分方程残差损失 $\mathcal{L}_{\text{pde}}$ 确保预测解 $u_\phi(x,t;\theta)$满足控制方程。该损失通过计算偏微分方程残差的均方误差得到，评估范围为从时空域和参数空间中采样的配点。 $\mathcal{L}(\theta)$中的微分算子通过自动微分获得:

$$\mathcal{L}_{\text{pde}}(\phi) = \mathbb{E}_{(x,t)\sim\mathcal{U}(\Omega\times(0,T]),\theta\sim\mathcal{P}(\Theta)}\left[\|\mathcal{L}(\theta)u_\phi(x,t;\theta) - f(x,t;\theta)\|_2^2\right], \tag{15}$$

其中，$\mathcal{U}$ 表示定义域上的均匀分布，$\mathcal{P}(\Theta)$表示参数空间上的分布。初始条件损失 $\mathcal{L}_{\text{ic}}$ 和边界条件损失 $\mathcal{L}_{\text{bc}}$ 分别在各自的定义域上施加约束。两者均定义为预测值与规定值之间的均方误差，以促使预测与初始状态及边界行为相一致:

$$\mathcal{L}_{\text{ic}}(\phi) = \mathbb{E}_{x\sim\mathcal{U}(\Omega),\theta\sim\mathcal{P}(\Theta)}\left[\|\mathcal{I}(\theta)u_\phi(0,x;\theta) - u_0(x;\theta)\|_2^2\right], \tag{16}$$

$$\mathcal{L}_{\mathrm{bc}}(\phi) = \mathbb{E}_{(x,t)\sim\mathcal{U}(\partial\Omega\times(0,T]),\theta\sim\mathcal{P}(\Theta)}\left[\|\mathcal{B}(\theta)u_\phi(x,t;\theta) - g(x,t;\theta)\|_2^2\right] \tag{17}$$

当存在稀疏解测量时，数据保真损失 $\mathcal{L}_{\mathrm{data}}$ 使预测与真实观测值相一致。该项针对数据集 $\mathcal{D}$中的已知输入输出对（x， t， θ， u_{ref}）最小化差异:

$$\mathcal{L}_{\mathrm{data}}(\phi) = \mathbb{E}_{(x,t,\theta,u_{\mathrm{ref}})\sim\mathcal{D}}\left[\|u_\phi(x,t;\theta) - u_{\mathrm{ref}}\|_2^2\right]. \tag{18}$$

训练目标用于确定使总损失最小化的最优参数 ϕ^*：

$$\phi^* = \arg\min_\phi \mathcal{L}(\phi). \tag{19}$$

此优化过程采用随机梯度法。使用 Adam 优化器，利用其自适应学习率，并有效处理深度网络训练。在每一步中，基于来自相应分布的小批量配点和参数样本，计算损失以及相对于 ϕ 的梯度。随后按照 Adam 更新规则迭代更新参数。指数衰减等学习率退火计划有助于在训练后期进行微调，从而促进收敛至精确且符合物理规律的解算子。

2.5 T切比雪夫多项式神经算子的理论分析

本节将进行严格的理论分析，证明 CPNO 框架能够同时实现近乎最优的表达能力与可训练性。从经典逼近理论的角度来看，切比雪夫谱编码所产生的逼近误差，其收敛速率可任意接近最佳多项式逼近所能达到的速率。与此同时，从优化动力学和数值线性代数的角度来看，切比雪夫基具有近乎最优的条件性，其勒贝格常数仅随多项式次数对数增长，因此能够有效抑制谱偏差，确保前向传播和反向传播中的数值稳定性，并在整个训练过程中保证梯度流稳健且条件良好——而这些性质已被证明是声名狼藉、条件极差的单项式基无法实现的。这些通过分析推导出的谱收敛和动力学稳定性特性，共同支撑了该框架在实证中表现出的优异精度、泛化性能和训练可靠性。

2.5.1 切比雪夫谱近似与误差分解

为了严格量化 CPNO 框架的学习目标，我们从谱投影的角度分析近似误差。考虑目标解算子 $\mathcal{S}$：$\Theta \to V$，它将参数实例 θ 映射为解 $u(x;\theta)$。神经算子构造近似 $u_\phi(x;\theta)$。我们不假设 u_ϕ 严格位于某个多项式空间内，而是利用三角不等式分解总误差，以分离函数基的贡献与网络的优化能力:

$$\|u(x;\theta) - u_\phi(x;\theta)\|_{L^\infty} \le \|u(x;\theta) - \mathcal{P}_N u(x;\theta)\|_{L^\infty} + \|\mathcal{P}_N u(x;\theta) - u_\phi(x;\theta)\|_{L^\infty} \tag{20}$$

其中，第一项为截断误差，第二项为估计误差。 $\mathcal{P}_N$ 表示投影到次数不超过 N的切比雪夫多项式空间上的最优谱投影算子:

$$\mathcal{P}_N u(x;\theta) = \sum_{k=0}^{N} c_k^*(\theta)T_k(x), \quad c_k^*(\theta) = \frac{\langle u(\cdot;\theta), T_k\rangle_w}{\|T_k\|_w^2} \tag{21}$$

此处， $\langle\cdot,\ \cdot\rangle_w$表示权函数为 $w(x) = (1-x^2)^{-1/2}$ 的加权内积。第一项，$\mathcal{E}_{tru} = \|u(x;\ \theta) - \mathcal{P}_N u(x;\ \theta)\|_{L^\infty}$，表示误差的理论下界，该下界仅由基函数的选择和解的正则性决定。

定理 1 (截断误差的指数衰减**)**。 设 $u(x;\theta)$ 在伯恩斯坦椭圆 $\mathcal{E}_\rho$ 内解析，该椭圆的焦点为 ± 1 ，半长轴与半短轴之和等于 $\rho > 1$。切比雪夫级数的截断误差满足：

$$\mathcal{E}_{app} = \left\| u(x;\theta) - \sum_{k=0}^{N} c_k^*(\theta) T_k(x) \right\|_{L^\infty} \leq \frac{C\rho^{-N}}{\rho - 1} \tag{22}$$

其中 C 是依赖于 u在 $\mathcal{E}_\rho$内最大模的常数。

定理 1 确立了 CPNO 主干网络的谱收敛性质。误差随 N呈指数衰减，与基于网格的方法的代数衰减形成对比。这确保了适度的 N 足以捕捉复杂的解流形。证明见附录 B。

2.5.2 估计误差界与数值稳定性

方程的第二项20 是估计误差 $\mathcal{E}_{est}$，量化了最优谱投影与网络输出之间的差异：

$$\mathcal{E}_{est} = \left\| \sum_{k=0}^{N} c_k^*(\theta) T_k(x) - u_\phi(x;\theta) \right\|_{L^\infty} \tag{23}$$

最小化 $\mathcal{E}_{est}$ 要求神经网络准确逼近映射 $\theta \mapsto \{c_k^*(\theta)\}_{k=0}^N$。这将偏微分方程求解任务转化为系数回归问题。CPNO 框架的一个显著优势在于其固有的数值稳定性，可通过目标函数 $u(x;\theta)$及其导数的界进行严格控制。

定理 2 （函数振幅界）。 利用切比雪夫多项式的性质，$|T_k(x)| \leq 1$ 对所有 $x \in [-1,1]$成立，目标函数的振幅满足

$$|u_s(x;\boldsymbol{\theta})| = \left| \sum_{k=0}^{N} c_k(\boldsymbol{\theta}) T_k(x) \right| \leq \sum_{k=0}^{N} |c_k(\boldsymbol{\theta})| \cdot |T_k(x)| \leq \sum_{k=0}^{N} |c_k(\boldsymbol{\theta})|. \tag{24}$$

取 L^∞ 范数可得到更紧凑的界：

$$\|u_s(x;\boldsymbol{\theta})\|_{L^\infty([-1,1])} \leq \|\mathbf{c}(\boldsymbol{\theta})\|_1, \tag{25}$$

其中 $\mathbf{c}(\theta) = [c_0(\theta) \ldots, c_N(\theta)]^T$，而 $\|\cdot\|_1$ 表示向量 ℓ^1 范数。因此， u_s 的全局振幅由其谱系数向量的 L^1 范数决定。

通过对网络 $\mathcal{M}_c$ 的输出施加约束，例如采用 $GELU$ 激活函数或范数裁剪，可强制满足条件 $\|\mathbf{c}(\boldsymbol{\theta})\|_1 \leq C$ ，从而保证 u_s 有界。该机制可确保前向传播的数值稳定性，并防止解发散。

在物理信息设置中，解的导数稳定性至关重要。将 m 阶微分算子定义为 $\mathcal{D}^m = \frac{d^m}{dx^m}$，则有：

$$\mathcal{D}^m u_s(x;\boldsymbol{\theta}) = \sum_{k=0}^{N} c_k(\boldsymbol{\theta}) \mathcal{D}^m T_k(x). \tag{26}$$

定理 3 (导数振幅界**)**。 切比雪夫多项式的 m 阶导数满足界 $\|\mathcal{D}^m T_k(x)\|_{L^\infty} \leq C_m k^{2m}$，其中 C_m 是仅依赖于 m 的常数。因此，解的导数大小有如下界：

$$|\mathcal{D}^m u_s(x;\boldsymbol{\theta})| \leq \sum_{k=0}^{N} |c_k(\boldsymbol{\theta})| \cdot |\mathcal{D}^m T_k(x)| \leq C_m \sum_{k=0}^{N} |c_k(\boldsymbol{\theta})| k^{2m}. \tag{27}$$

解的导数大小由谱系数的加权 L^1 范数决定。为确保高阶导数的稳定性，网络会受到隐式引导，学习谱衰减系数，使得 $|c_k(\boldsymbol{\theta})|$ 能够足够快速地衰减，以抵消 k^{2m} 的多项式增长。这一特性与现实物理系统中解的谱衰减行为一致，并能有效防止大系数高阶项引发的数值崩溃，从而增强前向传播的数值鲁棒性。

2.5.3 反向传播中数值稳定性的理论分析

我们方法的优势在训练过程的数值稳定性方面尤为显著。首先考虑偏微分方程的线性微分算子 $\mathcal{L}$，其定义为 $\mathcal{L}[u](x) = f(x;\boldsymbol{\theta})$。残差为：

$$\mathcal{R}(x;\boldsymbol{\theta}) = \mathcal{L}[u_s](x;\boldsymbol{\theta}) - f(x;\boldsymbol{\theta}) = \sum_{k=0}^{N} c_k(\boldsymbol{\theta})\mathcal{L}[T_k](x) - f(x;\boldsymbol{\theta}). \tag{28}$$

损失函数定义为残差的平方 L^2 范数：$Loss = \int_{-1}^{1}(\mathcal{R}(x;\boldsymbol{\theta}))^2 dx$。关于谱系数 $c_j(\boldsymbol{\theta})$的梯度为：

$$\frac{\partial Loss}{\partial c_j} = 2\int_{-1}^{1} \mathcal{R}(x;\boldsymbol{\theta})\frac{\partial \mathcal{R}}{\partial c_j} dx. \tag{29}$$

由于算子 $\mathcal{L}$具有线性性质，因此有 $\frac{\partial \mathcal{R}}{\partial c_j} = \mathcal{L}[T_j](x)$，这是一个固定的基函数。令$\phi_j(x) = \mathcal{L}[T_j](x)$，则梯度为：

$$\frac{\partial Loss}{\partial c_j} = 2\int_{-1}^{1} \left(\sum_{k=0}^{N} c_k(\boldsymbol{\theta})\phi_k(x) - f(x;\boldsymbol{\theta})\right)\phi_j(x) dx. \tag{30}$$

梯度向量 $\nabla_{\mathbf{c}} Loss$ 可写成矩阵形式：

$$\nabla_{\mathbf{c}} Loss = 2(\mathbf{G}\mathbf{c} - \mathbf{b}), \tag{31}$$

其中，$\mathbf{G}$ 是 Gram矩阵，其元素为；$\mathbf{b}$ 是一个向量，其元素为

$$G_{jk} = \int_{-1}^{1} \phi_j(x)\phi_k(x) dx = \int_{-1}^{1} \mathcal{L}[T_j](x)\mathcal{L}[T_k](x) dx$$

$$b_j = \int_{-1}^{1} f(x)\phi_j(x) dx$$

训练稳定性的核心取决于 Gram矩阵$\mathbf{G}$的谱性质，尤其是其条件数 $\kappa(\mathbf{G})$。该矩阵的性质由函数基的选择和线性算子 $\mathcal{L}$决定。为了从根本上揭示基选择的影响，我们详细比较由单项式基和切比雪夫基生成的 Gram矩阵。Gram矩阵的条件数 $\kappa(\mathbf{G})$直接刻画了二次损失景观的几何形状。因此，分析由切比雪夫基和标准单项式基生成的 Gram矩阵，能够揭示它们在数值稳定性方面的内在差异。为阐明这些内在性质，我们考虑最简单但最具代表性的情形：恒等算子，即 $\mathcal{L} = I$。在这种情形下，直接采用谱展开来拟合目标函数。

我们首先分析标准单项式基 $\{\psi_k(x) = x^k\}_{k=0}^{N}$。相应格拉姆矩阵 $\mathbf{G}_M$ 的元素为：

$$G_{jk}^{(M)} = \langle x^j, x^k\rangle_{L^2} = \int_{-1}^{1} x^j x^k dx = \int_{-1}^{1} x^{j+k} dx. \tag{32}$$

该积分可直接计算：

$$G_{jk}^{(M)} = \left[\frac{x^{j+k+1}}{j+k+1}\right]_{-1}^{1} = \frac{1-(-1)^{j+k+1}}{j+k+1}, \tag{33}$$

可进一步简化为：

$$G_{jk}^{(M)} = \begin{cases} \frac{2}{j+k+1} & \text{if } j+k \text{ is even}, \\ 0 & \text{if } j+k \text{ is odd}. \end{cases} \tag{34}$$

该矩阵具有一种结构，其数值性质与臭名昭著的**希尔伯特矩阵**（其元素为 $H_{ij} = \frac{1}{i+j-1}$）极为相似。其条件数 $\kappa(\mathbf{G}_M)$ 随矩阵大小 N 呈指数增长，近似为 $\kappa(\mathbf{G}_M) \sim O(e^{3.5N})$。这表明，由单项式基导出的格拉姆矩阵具有极差的条件性，意味着任何基于高阶单项式的学习框架在本质上都存在数值不稳定性。

接下来，我们分析切比雪夫基 $\{\psi_k(x) = T_k(x)\}_{k=0}^{N}$。其格拉姆矩阵$\mathbf{G}_T$的元素为：

$$G_{jk}^{(T)} = \langle T_j, T_k \rangle_{L^2} = \int_{-1}^{1} T_j(x) T_k(x) dx. \tag{35}$$

必须指出，该内积是标准的无权 L^2 内积，而非加权内积 $\langle \cdot, \cdot \rangle_w$（权函数为$(1-x^2)^{-1/2}$），切比雪夫多项式在后者下完全正交。因此，所得矩阵并非严格对角。该积分具有已知的解析解：

$$G_{jk}^{(T)} = \begin{cases} 0 & \text{if } j+k \text{ is odd}, \\ \frac{-2}{(j+k-1)(j+k+1)} & \text{if } j \neq k \text{ and } j+k \text{ is even}, \\ \pi & \text{if } j = k = 0, \\ \frac{\pi}{2} & \text{if } j = k > 0. \end{cases} \tag{36}$$

对该矩阵结构的分析表明，对角元素 $G_{kk}^{(T)}$（取值为 π 或 $\pi/2$）是矩阵中最大的元素。非对角元素 $G_{jk}^{(T)}$ 随其距对角线距离的增大而迅速减小，衰减率为 $O((j+k)^{-2})$。因此，在标准 L^2 内积下由切比雪夫基导出的格拉姆矩阵是对角占优的。此类矩阵已知是良态的，其条件数 $\kappa(\mathbf{G}_T)$随 N增长非常缓慢，通常呈低阶多项式速率。这为我们算子框架的稳定性、收敛速度和最终精度提供了坚实的数学基础。

对于非线性微分算子 $\mathcal{N}$，虽然损失函数 L 不再是简单的二次型，且学习问题变成了非凸优化，但切比雪夫基仍能保证稳定性。首先，解的有界性（$|T_k(x)| \leq 1$）及其导数的有界性（$|T_k'(x)| \leq k^2$）确保计算复杂的非线性残差时，不会因基函数本身数值爆炸而失败。其次，其近似正交性确保优化过程中遇到的局部子问题具有相对良好的条件性，从而避免单项式基可能引发的病态行为。

3 实验结果与讨论

为严格评估我们的方法，我们针对参数化偏微分方程开展了实验，包括 Burgers 方程、扩散-反应系统和艾伦-卡恩方程。所有问题的参数空间均使用高斯随机场（GRF）生成，其平均值为零，并采用指数二次核 $k_l(x_1, x_2) = \exp\left(-\frac{|x_1-x_2|^2}{2l^2}\right)$。我们将所提出的方法与旨在查询参数化偏微分方程解的物理信息无网格方法进行比较。基线方法包括 PI-DeepONet、未经微调的 MAD 和 Hyper-PINNs。为确保公平比较，所有方法均采用具有 4 个隐藏层、每层宽度为 64 的全连接神经网络实现。我们使用 GELU 激活函数和 Adam 优化器。所有实验均在单台 NVIDIA V100 GPU 上完成。性能主要通过相对误差 L_2 进行评估：

$$\text{Error} = \frac{\|u_{\text{pred}} - u_{\text{true}}\|_{L_2}}{\|u_{\text{true}}\|_{L_2}} \tag{37}$$

其中 u_{true} 表示参考解，而 u_{pred} 表示神经网络预测。该指标提供了一种归一化的解精度度量，不受解的幅值影响，并广泛应用于偏微分方程相关任务。更多详情请参阅 https://github.com/biao023/CPNO.git。

我们的实验首先将 CPNO 的基本测试结果与基线方法进行比较分析。随后，我们研究该方法的收敛特性，以及其逼近多尺度误差流场的有效性。最后，我们对切比雪夫编码阶数对解精度的影响进行了消融研究。

3.1 比较数值结果

为初步评估所提出的 CPNO 在求解参数化偏微分方程（PDE）方面的性能，本节将其与该领域常见的几种代表性物理信息无网格方法进行直接比较。这些基线方法包括物理信息 DeepONet（PI-DeepONet）、元学习辅助 DeepONet（MAD）和基于超网络的物理信息神经网络（HyperPINNs）。对于 CPNO，我们评估了两种不同的训练范式：第一种是纯物理信息零样本学习方法（CPNO-zero），不依赖任何真实解的观测数据；第二种是少样本学习方法（CPNO-few），通过加入少量观测数据来增强物理信息训练，具体而言，这些数据对应三个参数实例的解。基准偏微分方程包括伯格斯方程、艾伦-卡恩方程、扩散-反应系统和纳维-斯托克斯方程，其详细介绍见附录 A。

表1：CPNO 与基线方法在参数化偏微分方程中的相对 $L2$ 误差

方法	Burgers	Allen-Cahn	扩散-反应	Navier-Stokes
PI-DeepoNet	0.0809	0.0237	0.0318	0.0342
MAD	0.1063	0.2262	0.0750	0.1918
HyperPINNs	0.0455	0.0476	0.0314	0.0406
CPNO-zero	0.0306	0.0169	0.0296	0.0272
CPNO-few	0.0290	0.0109	0.0299	0.0166

表 1总结了 CPNO 和基线方法在三个基准问题上取得的相对 L2 误差，表明所提出的 CPNO 方法在测试的参数化偏微分方程问题中表现出竞争力。在纯物理信息零样本设置（CPNO-zero）中，CPNO 始终优于所有竞争方法；在纳维–斯托克斯基准上，相较于表现最强的既有方法（HyperPINN），误差最多降低了 62%，并在整套基准测试中创下新的最佳结果。训练过程中仅纳入少量（$\leq$ 5%）稀疏采样的解观测时，少样本变体（CPNO-few）在四个问题中的三个问题上进一步提高了精度，并在 Allen–Cahn 方程（较 CPNO-zero 提高 35%）、扩散–反应方程和纳维–斯托克斯方程（提高 39%）上取得迄今报告的最低误差。在 Burgers方程上，CPNO-few 的表现与其零样本版本相差不到 5%，同时仍显著优于所有基线方法。这些结果凸显了 CPNO 出色的数据效率、架构表达能力，以及在稳健的物理信息基础上无缝利用有限监督信号的能力。这些初步发现表明，无论采用零样本模式还是少样本模式，CPNO 框架都能为参数化偏微分方程提供高精度的无网格解。

3.2 收敛速度与精度研究

为全面评估我们所提出的CPNO的性能，并严格验证其核心组件——切比雪夫谱编码——的有效性，我们针对一个关键消融模型开展了比较分析。该基线记为w/o-CPNO，具有相同的架构，但省略了谱编码模块，直接处理原始时空坐标。本节详细比较了CPNO方法与w/o-CPNO在全部三个基准参数化偏微分方程问题上的收敛率，以及达到预定义精度阈值所需的训练迭代次数。本节采用的所有训练方法均为纯物理驱动。

图 3 展示了应用于各个参数化偏微分方程的不同方法的收敛曲线。横坐标表示训练轮次，纵坐标表示

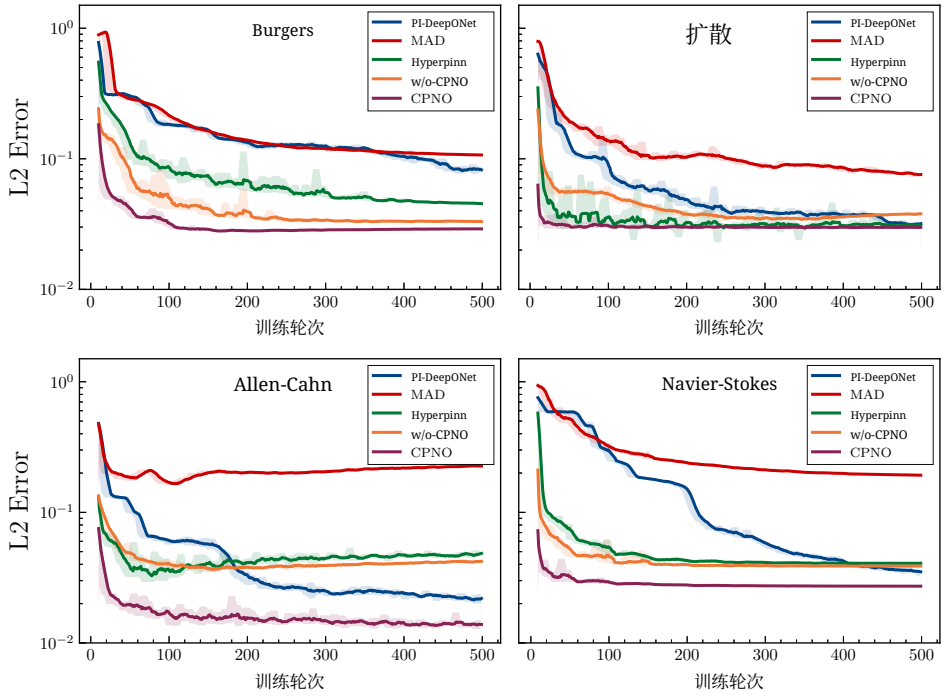

图3：不同方法在测试用例上的收敛曲线对比。

相应的相对 L_2 测试误差。在所有三个基准问题中，我们都观察到一致的结果：所提出的 CPNO 收敛最快，最终精度也最高。在基线方法中，HyperPINN 的整体表现最为出色，但其参数数量显著多于其他基线方法。相比之下，PI-DeepONet 和 MAD 收敛较慢，最终精度也较低。值得注意的是，不采用切比雪夫编码增强的 w/o-CPNO 在 Burgers方程上的表现优于 HyperPINN，但在 Allen-Cahn 和扩散-反应系统上的表现略逊于 HyperPINN。相比之下，得益于增强的输入表示和高效的参数调制，CPNO 的误差曲线始终呈现快速的初始下降趋势，并在较少的训练轮次内稳定在最低误差水平。这凸显了它在学习效率和模型表达能力方面的综合优势。

图4 通过可视化最终解精度与计算效率之间的权衡，全面评估了模型性能。在该图中，每种方法收敛后的相对 L_2 测试误差与相应的训练成本进行对比，训练成本以收敛所需的训练轮次数量衡量。在这一空间中，最佳性能对应于误差低且计算成本低的区域，即图的左下角。在所有三个基准问题中，所提出的 CPNO 始终位于这一最佳区域，从而构成所评估方法的帕累托前沿。尽管某些基线方法能够达到相当的精度，但其训练成本却高得多。相反，其他收敛速度更快的方法，其最终误差则明显更高。这些结果表明，CPNO 框架在精度和效率之间实现了更优的平衡，能够以显著降低的计算开销，为参数化偏微分方程生成高保真解。这种效率对于需要快速迭代模型或探索广阔参数空间的实际应用尤其重要。

图 5 以可视化方式对比了 Burgers 方程某个选定参数实例的结果，展示：(i) 各方法的预测解；(ii) 预测解与真实解之间逐点绝对误差的等高线图 $|u_{pred} - u_{true}|$；以及 (iii) 展示该逐点误差场空间傅里叶谱能量分布的等高线图。直接观察预测解，可以初步评估各方法解析激波等关键特征的能力。误差等高线图进一步凸显了 CPNO 在整体误差控制方面的优势；其误差等高线通常颜色更浅，表明误差幅值更低。值得注意的是，CPNO

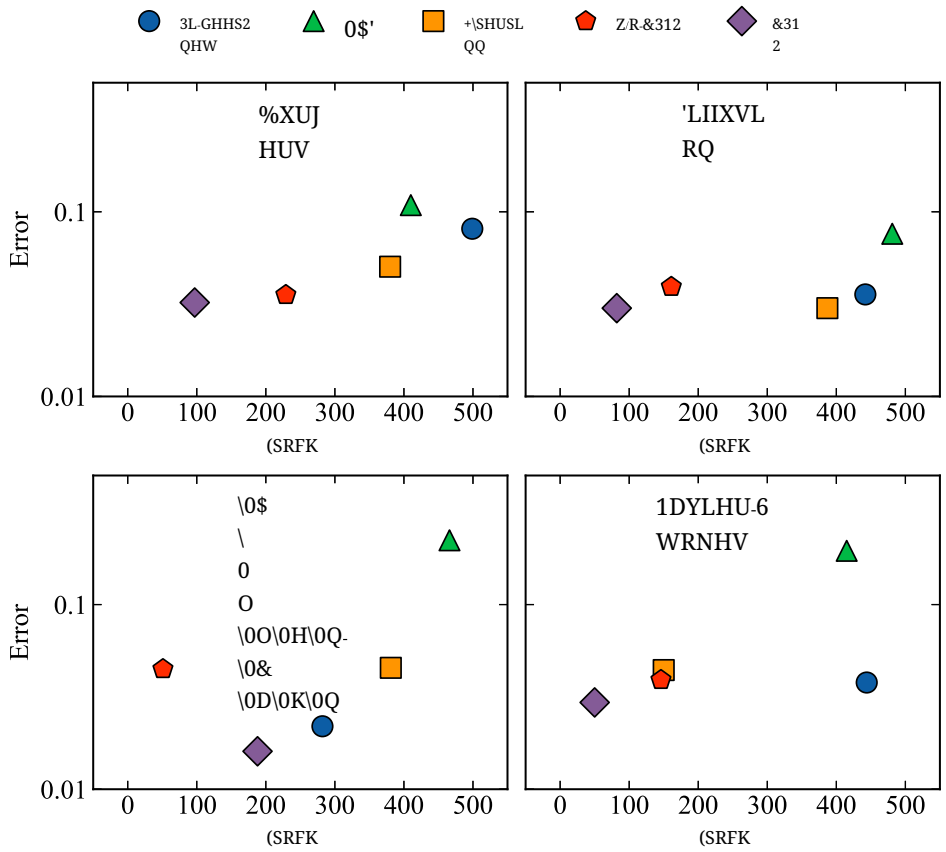

图4：不同方法在测试用例上达到特定测试准确率所需的最少训练轮数比较。

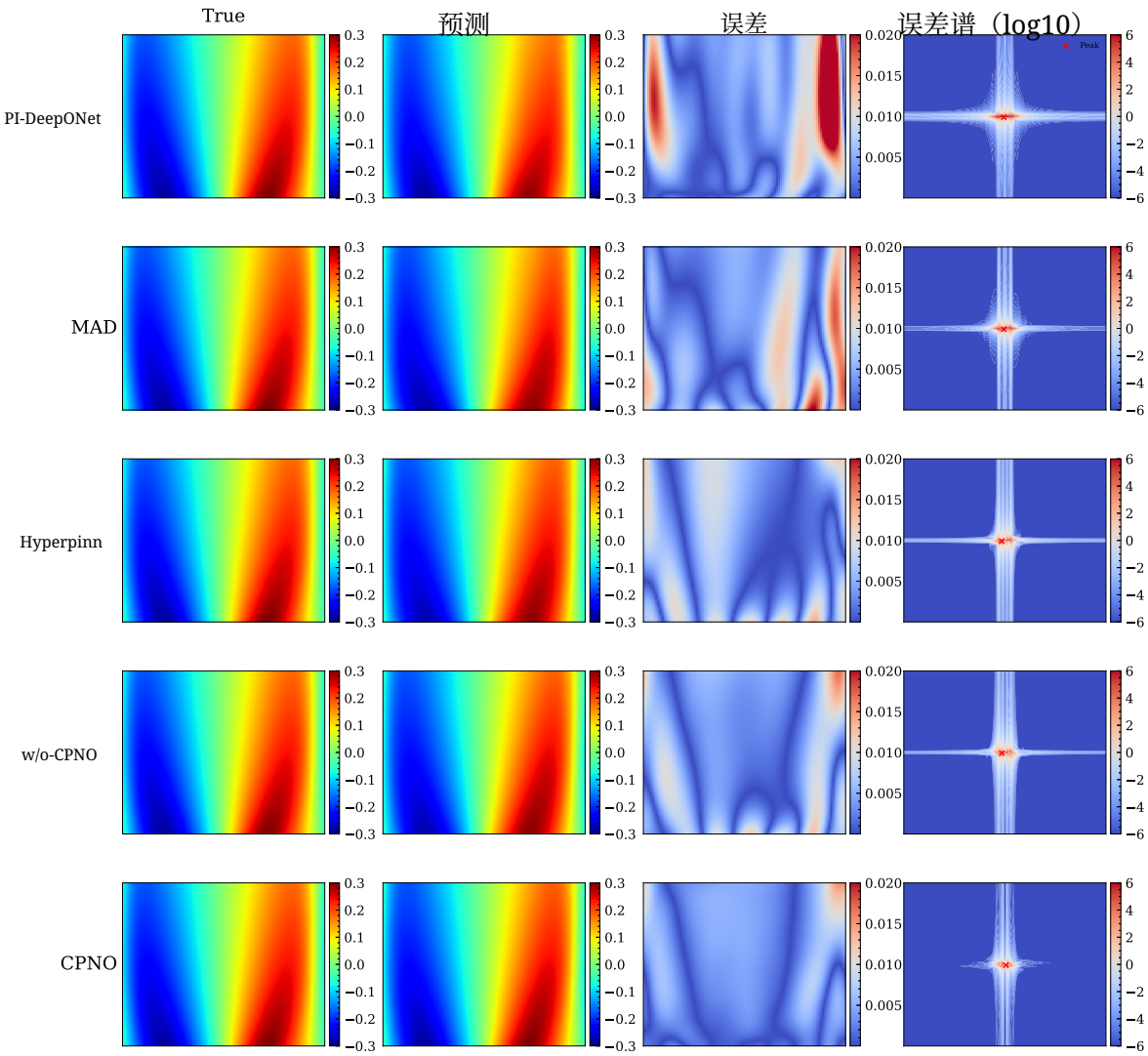

图 5：Burgers 的预测解、误差等值线和误差谱等值线的可视化比较。

即使在解梯度较大的区域（如激波前沿附近），其预测误差仍然较小；而一些基线方法在这些关键区域可能会出现明显的误差集中。误差谱等高线图深入展示了空间域中不同频率模态下的误差分布。与其他方法相比，CPNO不仅显著抑制了高频空间分量中的误差——这表明它在捕捉细节和高波数解内容方面能力更强——而且在低频区域的误差水平通常也更低。与基线方法相比，CPNO在整个频谱范围内都能更忠实地逼近Burgers方程的解，这一全面的频谱误差降低充分证明了其能力。

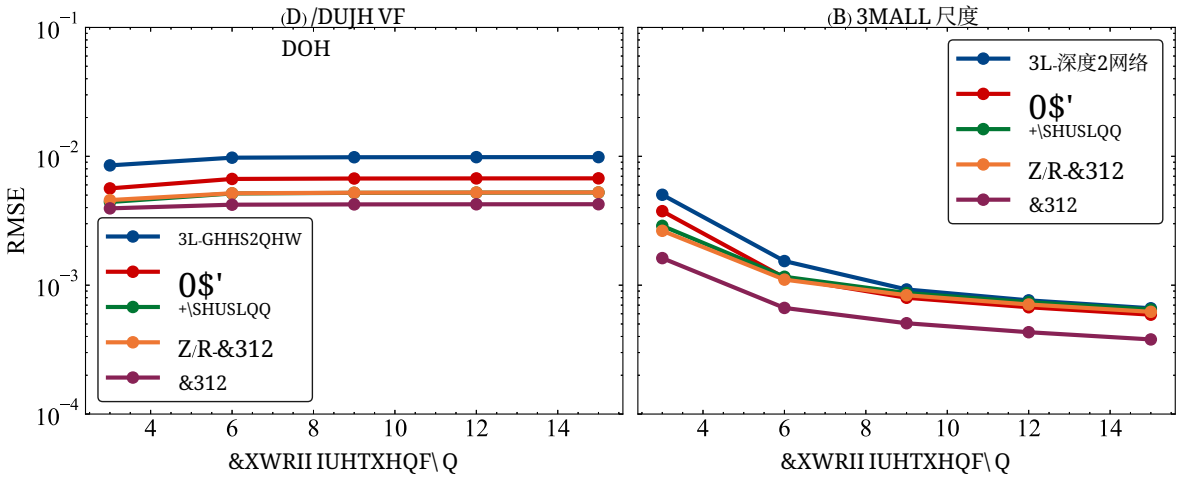

图6：Burgers方程多尺度均方根误差比较。(a) 大尺度分量的误差随截止波数的变化；(b) 小尺度分量的误差随截止波数的变化

为进一步考察不同方法在 Burgers方程中捕捉不同尺度解结构的能力，我们对预测结果进行了多尺度误差分解。通过在傅里叶域应用具有不同截止波数的低通和高通滤波器，将真实解以及各模型的预测解分解为大尺度（低频）和小尺度（高频）分量。随后，分别计算这两个尺度分量的均方根误差（RMSE），以量化各模型在不同尺度上的预测精度。图 6 展示了 Burgers 的一个测试案例中这些多尺度误差的比较分析，对比了 CPNO 与基线方法。图 6a 显示大尺度分量的均方根误差，图 6b 显示小尺度分量的均方根误差，两者均以截止波数为变量。从结果中可以清楚地看出，在所有考察的截止波数下，CPNO 在大尺度和小尺度分量上的预测误差始终最低。这表明，CPNO 不仅能够准确捕捉解的全局宏观结构（大尺度特征），还展现出更强的能力，可解析精细细节和高频振荡（小尺度特征）。这一结果有力地证明了我们提出的方法在解决多尺度问题方面具有全面的性能优势和广泛的适用性。

3.3 超参数敏感性分析

为全面评估我们提出的 CPNO 的性能，并严格验证其核心组件的有效性，我们将其与消融模型 w/o-CPNO 进行比较分析。两种架构均在前文介绍的零样本和少样本训练范式下进行评估，得到四种用于比较的模型变体：CPNO-zero、CPNO-few、w/o-CPNO-zero 和 w/o-CPNO-few。本节将在一系列超参数设置下，对这四种变体进行详细分析，以评估其性能和稳健性。我们研究这些模型对关键超参数的敏感性，包括批量大小和学习率。批量大小从集合 1000、2000、4000 中选取，学习率从 0.005、0.001、0.0005 中选取，共得到 9 种不同的超参数配置。对于每种配置，所有模型均在 Burgers 上独立进行训练和评估，并记录相对 {v1}测试误差。

共得到 9 种不同的超参数配置。对于每种配置，所有模型均在 burgers 上独立进行训练和评估，并记录相对 L_2 测试误差。

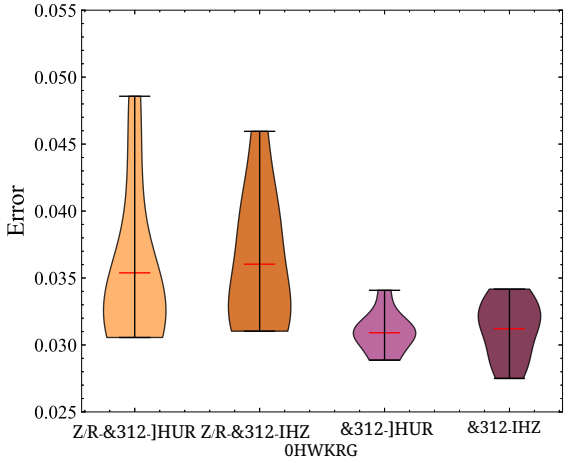

图7：不同超参数配置下 CPNO 和 m-PINO 的平均相对误差及标准差比较

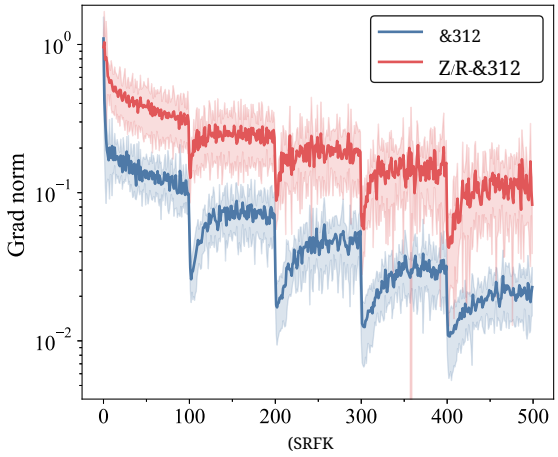

图8：不同超参数配置下，训练过程中 CPNO 和 w/o-CPNO 的逐轮次平均梯度范数比较。

图7 通过展示四种模型变体在九种不同超参数配置下计算所得的相对误差小提琴图，详细呈现了超参数鲁棒性。 L_2 每个小提琴图的形状都展示了误差的完整概率密度分布，其宽度对应于特定误差水平下结果的密度。与单独使用汇总统计量相比，这种可视化能更全面地展现性能的一致性。显而易见，CPNO-zero 和 CPNO-few 对应的小提琴图明显更紧凑，且集中在较低的误差值范围内。这不仅表明其中位数精度更高，也说明模型在测试的超参数范围内具有很高的性能一致性。相比之下，w/o-CPNO 变体的小提琴图明显更细长、更宽，表明它们对超参数选择更为敏感。这些图中较长的上尾意味着某些超参数选择会导致性能大幅下降，而这种风险无法通过简单的标准差指标充分体现。这些结果有力地证明，引入切比雪夫谱编码有助于形成更稳定、更鲁棒的优化景观，并显著降低模型对超参数变化的敏感性。

为进一步阐明切比雪夫编码对模型训练动态和稳定性的影响，我们比较了 CPNO 与 w/o-CPNO 在此前详述的 9 种超参数配置下，逐训练轮次平均 L_2 梯度范数的变化，并记录了每个训练轮次的平均 L_2 梯度范数。图8 展示了 CPNO 和 w/o-CPNO 在全部九种超参数配置下逐训练轮次平均 L_2梯度范数的均值曲线及其对应的标准差（以阴影区域表示）。横坐标表示训练轮次，纵坐标表示对数尺度下的平均 L_2 梯度范数。从图中可以清楚地看出，在整个训练过程中，CPNO 的平均梯度范数整体上不仅幅度较低，而且与 w/o-CPNO 方法相比，标准差也显著更小。具体而言，w/o-CPNO 的梯度范数曲线波动较大，在不同超参数设置下保持较高的均值和较大的方差，这可能表明其优化景观较为敏感。相比之下，CPNO 的梯度范数曲线平滑得多，均值更低，标准差也明显更小。这表明其训练过程更加稳定，且较少受到超参数选择变化的影响。更小且更稳定的梯度范数通常与更易优化的损失景观和更可靠的优化轨迹相关，这与图8中报告的 CPNO 更高的精度和超参数鲁棒性相一致。

3.4 鲁棒性分析：对输入参数噪声的稳定性

理想的神经算子不仅必须针对精确参数得出准确的解，还应对输入中的微小扰动具有鲁棒性。为系统评估我们提出的 CPNO 在这方面的性能，我们以 Burgers方程为基准设计了噪声鲁棒性实验。在该实验中，我们向定义初始条件的参数中加入不同程度的高斯白噪声，标准差 σ 设为 $\{0.005、0.01、0.03\}$，并将结果与无噪声情况进行比较。在这些含噪输入条件下，我们评估了 CPNO 模型及其消融变体 w/o-CPNO 的性能。所有模型均在纯物理信息（零样本）范式下训练。

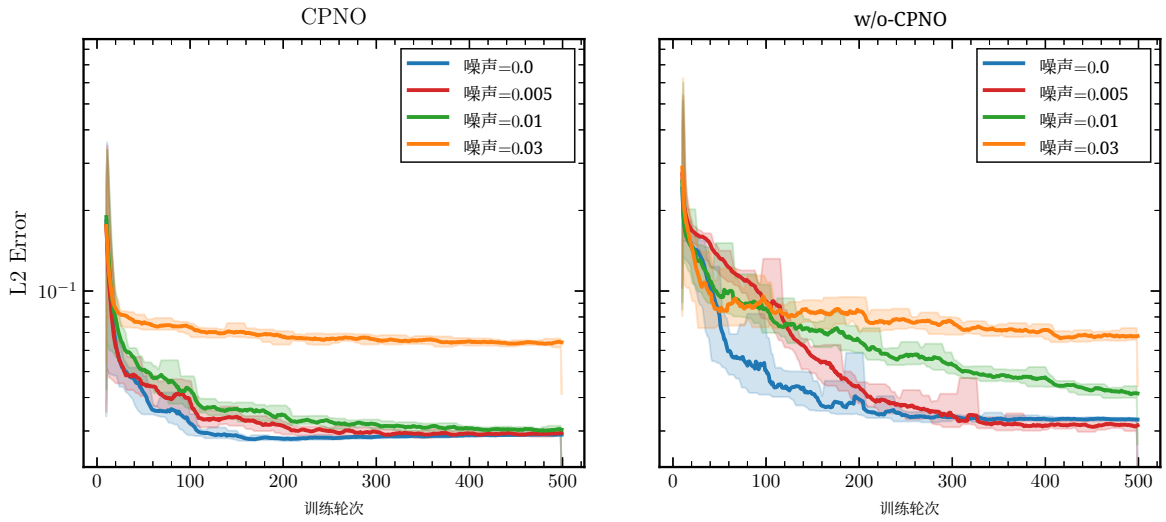

图 9：在不同参数噪声水平下，CPNO 和 w/o-CPNO 对 Burgers方程的训练收敛情况。

图9 展示了两个模型在不同噪声水平下的训练收敛过程。可以清楚地观察到，低噪声水平下 CPNO 的收敛曲线几乎无法区分，并且快速、稳定地收敛至极低的误差，这表明低水平噪声对其训练动态和最终精度的影响可以忽略不计。虽然较高的噪声水平会导致最终误差增大，但收敛依然迅速。相比之下，w/o-CPNO 的收敛曲线对所有噪声水平都表现出显著的敏感性。随着噪声水平升高，其收敛过程变得更加振荡，最终误差也大幅增加，表明参数扰动严重影响了其训练过程。

为更直观地理解这种性能差异，图10 和11 针对一个测试案例，展示了不同噪声水平下预测解和逐点误差场的定性比较。如图10所示，CPNO 的预测结果在视觉上几乎与无噪声预测完全相同，并且其误差场在所有噪声水平下都保持在极低的量级。这表明 CPNO 已成功学习到稳定的算子映射。然而，图11揭示了 w/o-CPNO 模型的脆弱性。随着噪声水平升高，其预测解的质量急剧下降，出现显著的非物理振荡，尤其是在激波前沿附近。相应的误差场中存在大面积高误差区域，证实学习到的算子映射并不稳定。这些结果表明，我们的算子框架对低水平噪声具有鲁棒性，这主要归功于切比雪夫多项式提供的数值稳定特征空间。这使算子能够学习更平滑的参数到解流形，从而不受输入扰动的影响。

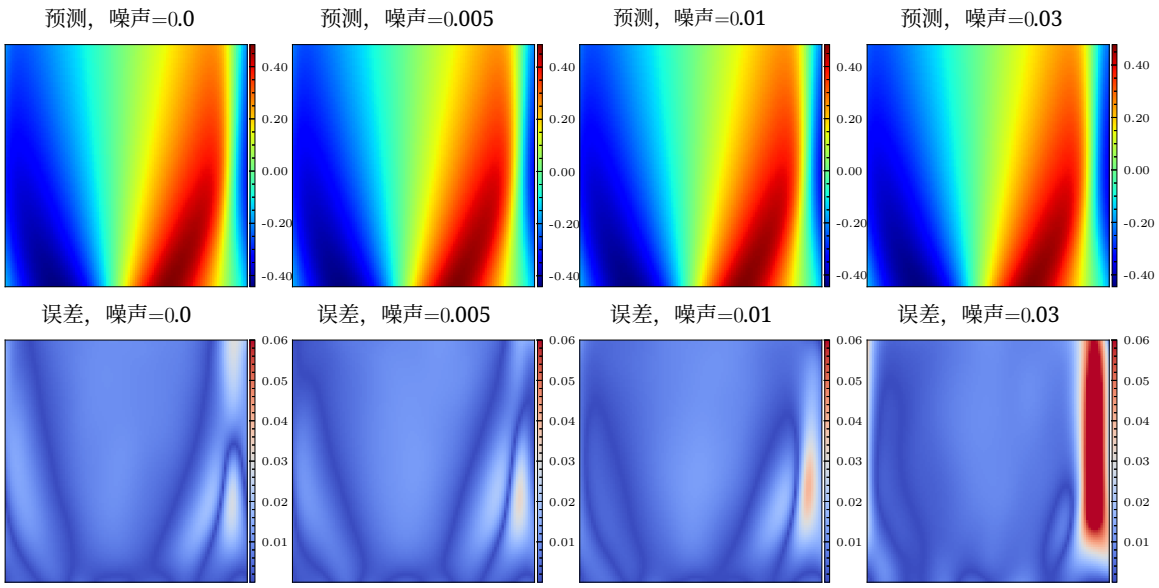

图10：CPNO 在不同输入噪声水平下的预测解（上排）和逐点误差场（下排）。

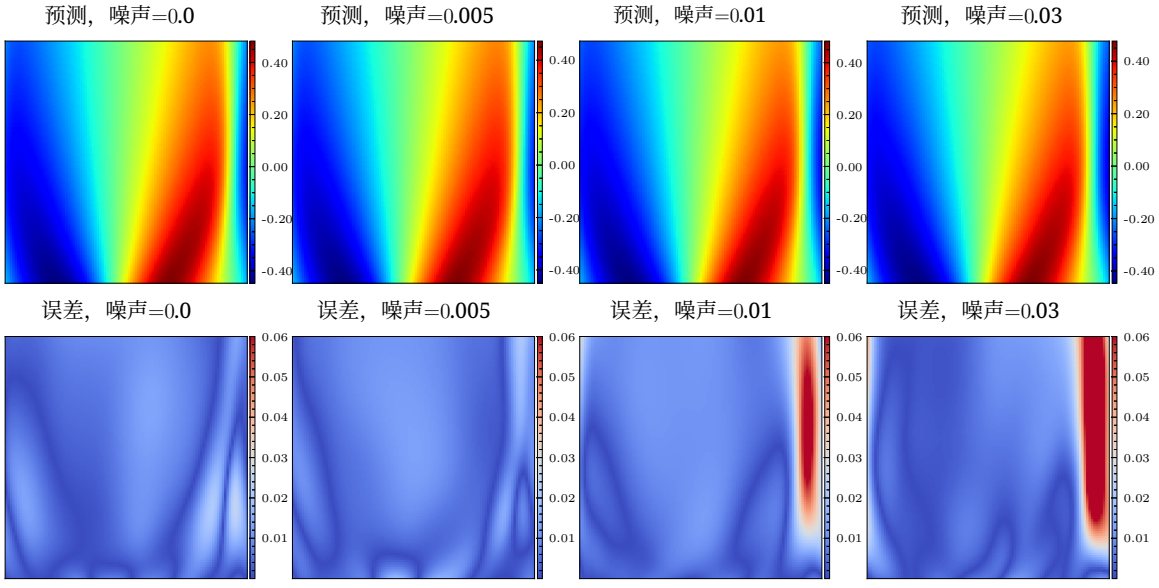

图11：w/o-CPNO 在不同输入噪声水平下的预测解（上排）和逐点误差场（下排）。

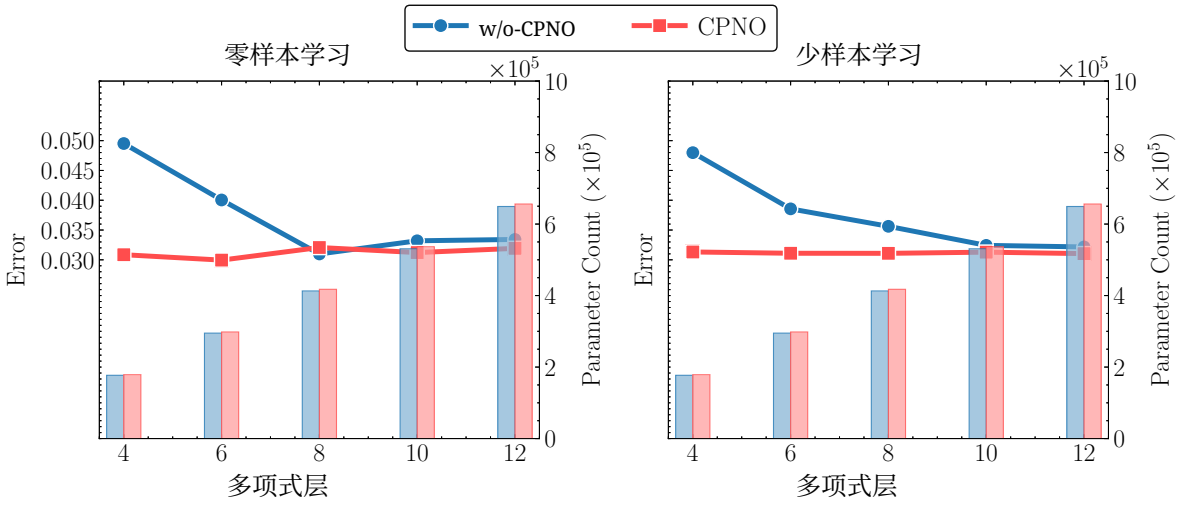

图12：Co　伯格斯方程案例中不同多项式层深度的最终测试误差比较

3.5 消融研究

3.5.1 多项式层深度与特征表示

为探究合成网络深度对模型性能的影响，并进一步阐明切比雪夫谱编码与网络架构之间的协同关系，我们进行了消融研究。我们在不同网络深度下评估了四种模型变体（w/o-CPNO-zero、w/o-CPNO-few、CPNO-zero 和 CPNO-few）的性能，层数分别设为 4、6、8、10 和 12。所有实验均基于 Burgers方程进行。

图 12分别通过两个子图展示了模型准确率和网络参数数量随深度变化的趋势，对应零样本和少样本范式。在每个子图中，横轴表示合成网络的层数，左、右纵轴分别表示最终相对 L_2 测试误差和网络参数总数。我们观察到几个关键现象。对于直接处理原始坐标的 w/o-CPNO 模型，较浅的网络（例如 4 层）误差相对较高，这表明浅层网络难以从原始输入中构造足够复杂的特征。随着网络加深，其误差逐渐降低，但这需要以参数数量线性增加为代价。相比之下，采用切比雪夫编码的 CPNO 模型表现出截然不同的特点。即使网络较浅（例如 4 层或 6 层），CPNO 也能达到很低的误差。进一步加深网络仅能带来微小的准确率提升，其性能曲线在较宽的深度范围内始终保持平缓，且误差水平较低。

这一比较为切比雪夫谱编码的一项核心优势提供了有力证据：设计良好、信息丰富的输入特征空间可显著降低模型对网络深度的依赖。如图X(a)所示，6层CPNO-zero模型所达到的精度远优于12层w/o-CPNO-zero模型，而参数量仅为后者的一半。这清楚地表明，通过在输入阶段进行高效的谱特征变换，我们可以用更紧凑、更高效的架构达到甚至超越更深、更复杂网络的性能。因此，切比雪夫编码不仅直接提升了模型的最终精度，更重要的是提高了特征表示的质量，从而使得构建更轻量、参数效率更高的神经算子成为可能。

3.5.2 用于解映射的切比雪夫阶数

我们提出的方法中的一个关键超参数是谱编码模块所用切比雪夫多项式的最高阶数。为研究模型性能对这一参数的敏感性，我们针对伯格斯方程开展了消融研究。在此实验中，网络深度

固定为 8 层合成网络，并在系统地改变切比雪夫阶数 $P \in 3$、4、5、6、7、8 的同时，评估 CPNO-zero 模型的性能。图 13 展示了最终相对 L_2 测试误差随编码阶数的变化。可以观察到一个明显趋势：当 P 从 3 增至 5 时，误差显著降低。这表明，低阶谱特征不足以充分捕捉 Burgers解流形的复杂性，而纳入更高阶的谱信息对于实现高精度至关重要。值得注意的是，对于 $P > 5$，误差曲线趋于平坦，表明模型性能已收敛。

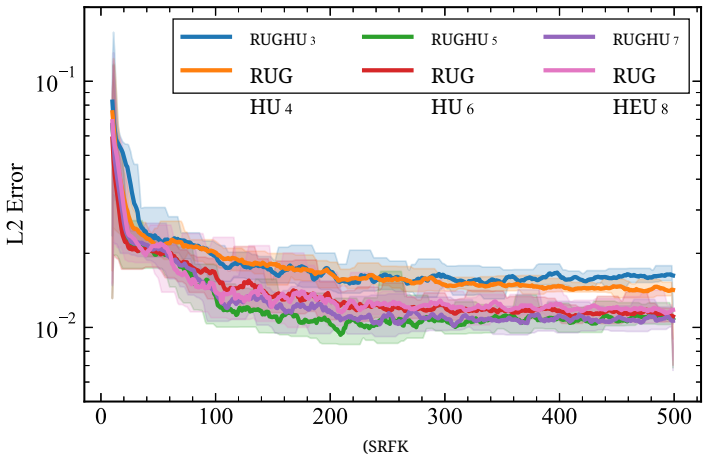

图13：Co　　不同切比雪夫多项式阶数在艾伦-卡恩方程算例上的最终测试误差比较

这一结果表明，针对该问题的切比雪夫编码存在一个有效的收敛阶。这体现了我们方法的稳健性，因为其性能对P的选择并非无限敏感，并且在超过某一足够阶数后趋于稳定。有限且适中阶数的切比雪夫多项式足以准确表示复杂的解流形，从而在表达能力与计算成本之间取得平衡。

3.6 向复杂几何的扩展：跨声速翼型流动

尽管此前的实验已展示 CPNO 对基准参数化偏微分方程的稳健零样本求解能力，但这些问题主要局限于规则域（例如矩形域或具有周期边界的域）。为进一步验证 CPNO 在不规则工程几何中的有效性，本节采用数据驱动方法评估其在跨声速翼型流动上的表现——这是一项航空航天工程中的典型挑战，其特点是边界复杂，并存在激波和湍流边界层等多尺度现象。我们利用高保真 CFD仿真数据训练 CPNO，并重点评估其在复杂几何形状中的精度和泛化能力。跨声速翼型流动由雷诺平均 Navier-Stokes（RANS）方程描述，其边界条件包括远场自由来流条件和翼型表面的无滑移条件。该数据集包含约 6400 个样本，每个样本的输入参数包括翼型几何描述符和攻角，输出流场则涵盖压力和速度分布。CPNO 的多项式阶数设为 $P = 10$ ，切比雪夫编码度设为

$$\mathcal{L} = \frac{1}{N}\sum_{i=1}^{N}\|u_{\phi_i}(\theta_i) - u_{\mathrm{ref}_i}(\theta_i)\|_2^2 \quad 16 \quad 1$$

训练采用数据驱动范式，通过最小化均方误差损失（x;) (x;) 进行，并依据 L 绝对误差指标进行评估。

表2：CPNO用于跨音速翼型流动预测的L1绝对误差

L1误差	u		v		p	
	mean	max	mean	max	mean	max
训练集	0.00175	0.07744	0.00094	0.08072	0.00126	0.05493
测试集	0.00216	0.08727	0.00103	0.08209	0.00121	0.05463

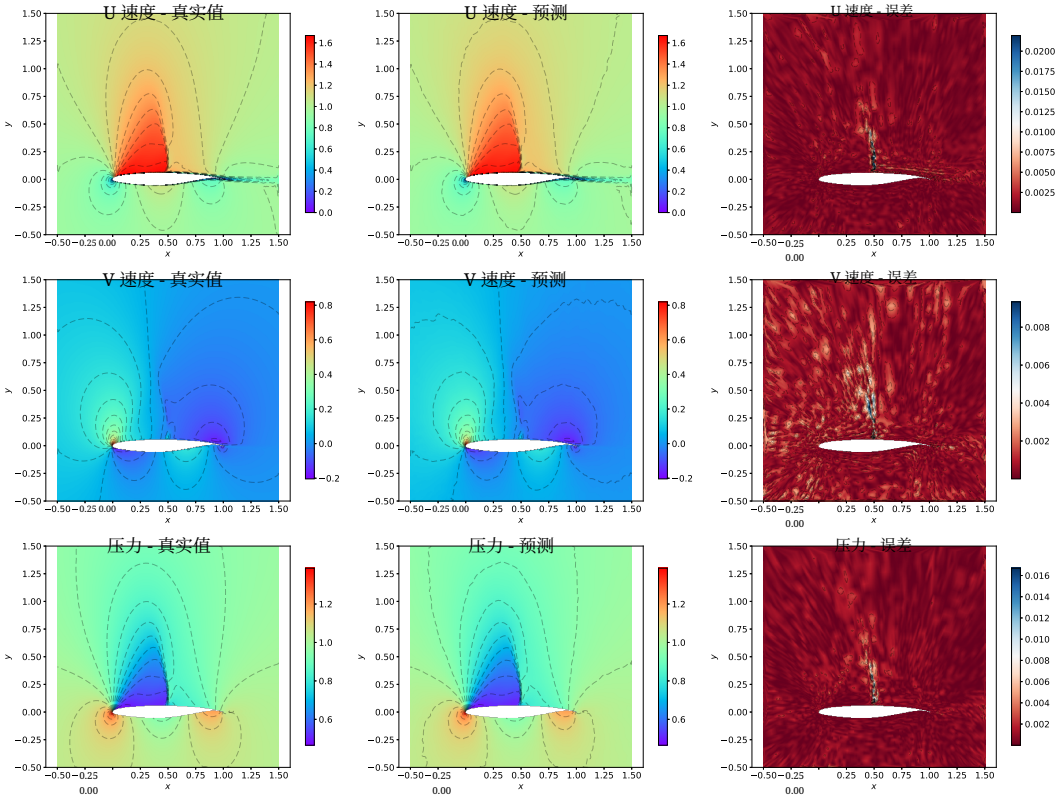

图14：速度分量（u、v）与压力（p）场的预测与真实全局分布对比

训练集和测试集的评估结果见表 2，平均误差始终处于 10^{-3} 的量级，最大误差则控制在 10^{-1} 的量级。训练误差与测试误差高度一致，凸显了该算法对几何形状和来流条件变化的强泛化能力。图 14 展示了模型在测试用例上的预测结果，可见模型能够有效捕捉激波结构，并展现出 CPNO 的高保真预测能力。这些发现表明，CPNO 能够有效拓展至复杂工程几何，为飞机设计优化提供有价值的参考。

4 结论

本文提出一种新颖的物理信息切比雪夫多项式神经算子（CPNO），旨在准确而稳健地求解参数化偏微分方程（PDE）。我们构建了堆叠算子框架，其中非线性由多项式展开表示，从而无需预先定义非线性分量。我们在每一层中采用切比雪夫多项式作为基础基函数，以同时增强网络的表示能力与训练稳定性。通过对基准参数化偏微分方程问题进行系统的数值实验，我们对CPNO的性能进行了严格评估。结果表明，与PI-DeepONet、MAD和HyperPINNs等代表性基线方法相比，CPNO不仅在最终解的精度上具有显著优势，而且展现出更快的收敛率。多尺度误差分析与误差谱分析进一步证实，CPNO能够出色地准确捕捉宏观解结构以及细粒度的高频细节。此外，对超参数敏感性和训练过程梯度范数的评估表明，CPNO具有更强的稳健性与训练稳定性。

尽管CPNO在多个方面展现出优越性能，但当前框架仍存在一定局限性。在处理无结构的可变参数时，目前仍依赖插值方法，因为其坐标表示和网络架构更适

用于规则域或可通过简单方式参数化的域。未来研究面临的一项重大挑战，是如何将CPNO及其相关的编码和调制策略有效扩展至任意复杂几何。未来有几个颇具前景的研究方向。首先，通过利用大量高保真数据整合数据驱动策略，可以引导网络学习复杂几何域上的行为，或用于校准和提升底层物理模型的预测准确性。其次，对于涉及更高维参数空间及时空域的复杂问题，有必要进一步优化CPNO架构和训练算法，以提升可扩展性和计算效率，从而更好地缓解维数灾难。最后，扩展CPNO的适用范围以求解呈现混沌行为的动力系统，也是一个颇具吸引力的研究方向。这类问题对长期预测稳定性以及准确捕捉对初始条件和参数的敏感性提出了严格要求，可能需要引入先进的时间序列建模技术或不确定性量化方法。

参考文献

[1] Pingki Datta、Nikhil Pawar 和 Salah Faroughi。利用物理信息神经网络对干燥颗粒流动进行建模。载于*AGU*秋季会议摘要，第 2022 卷，页 H52N–0633，2022年12月。[2]Micka¨el Delcey、Yoann Cheny、Jean-Baptiste Keck、Adrien Gans 和 S´ebastien Kiesgen de Richter。利用物理信息神经网络识别含泥沙流的沉降速度。《计算力学与工程中的计算方法》，432:117389，2024年。[3] S.A. Faroughi、C. Fernandes、J. Miguel N´obrega 和 G.H. McKinley。黏弹性流体中平移球体的阻力系数闭合模型。《非牛顿流体力学杂志》，277:104218，2020年。[4] Salah A. Faroughi、Ramin Soltanmohammadi、Pingki Datta、Seyed Kourosh Mahjour 和 Shirko Faroughi。采用周期激活函数的物理信息神经网络在非均质多孔介质溶质运移中的应用。《数学》，12(1)，2024年。[5] Xia Yan、Jingqi Lin、Sheng Wang、Zhao Zhang、Piyang Liu、Shuyu Sun、Jun Yao 和 Kai Zhang。利用物理信息神经网络模拟非均质裂隙多孔介质中的两相流。《水资源进展》，189:104731，2024年。[6] Bicheng Yan、Manojkumar Gudala、Hussein Hoteit、Shuyu Sun、Wendong Wang 和 Liangliang Jiang。用于地热能回收非迭代优化的物理信息机器学习。《应用能源》，365:123179，2024年。[7] Carlos Str¨ofer、Michel´en、Jin-Long Wu、Heng Xiao 和 Eric Paterson。利用卷积神经网络从流场中提取数据驱动、基于物理的特征。《计算物理通讯》，25(3):625–650，2019年。[8] Haoteng Hu、Lehua Qi 和 Xujiang Chao。用于计算固体力学的物理信息神经网络（pinn）：数值框架与应用。《薄壁结构》，205:112495，2024年。[9] Wei Zhang、Ao Yu、Haiyan Mao、Guangxia Feng、Cheng Li、Guanzhi Wang、Jinfa Chang、David Halat、Zhao Li、Weilai Yu、Yaping Shi、Shengwen Liu、David W. Fox、Hao Zhuang、Angela Cai、Bing Wu、Fnu Joshua、John R. Martinez、Lei Zhai、M. Danny Gu、Xiaonan Shan、Jeffrey A. Reimer、Yi Cui 和 Yang Yang。三相界面纳米多孔电催化剂上的动态气泡平衡了主动二氧化碳捕集与还原。146(31):21335–21347。出版方：美国化学会。[10] Lei Xu、Nengcheng Chen、Zeqiang Chen、Chong Zhang 和 Hongchu Yu。地球系统科学中的时空预测：方法、不确定性、可预测性及未来方向。

地球 - *Science Reviews*, 222:103828, 2021年.

[11] Kaan Sel、Amirmohammad Mohammadi、Roderic I. Pettigrew和Roozbeh Jafari。用于无袖带血压估计的生理时间序列建模的物理信息神经网络。6(1):110.[12] Federico Gatta、Vincenzo Schiano Di Cola、Fabio Giampaolo、Francesco Piccialli和Salvatore Cuomo。通过物理信息神经网络进行美式期权定价的无网格方法。《工程分析与边界元》, 151:68–82, 2023年.[13] Fr´ed´eric Hourdin、Daniel B. Williamson、Catherine Rio、Fleur Couvreux、Romain Roehrig、Najda Villefranque、Ionela Musat、Fatoumata Bint Diallo、Laurent Fairhead和Victoria Volodina。利用机器学习的基于过程的气候模型开发：II. 从单柱到全球的模型校准。《地球系统建模进展》, 13, 2020年.[14]Tianyixing Han、Shuyao Hu、Chongwen Jiang和Chun-Hian Lee。泰勒–麦科尔方程的基本解。《流体力学杂志》, 999:A48, 2024年.[15] Nanzhe Wang、Haibin Chang和Dongxiao Zhang。用于代理模型构建与反演建模的理论引导自编码器。《计算力学与工程中的计算方法》, 385:114037, 2021年.[16] Suihong Song、Dongxiao Zhang、Tapan Mukerji和Nanzhe Wang。GANSIM-surrogate：一种用于随机条件地质建模的集成框架。《水文学杂志》, 620:129493, 2023年.[17] Lu Lu、Pengzhan Jin和George Em Karniadakis。DeepONet：基于算子通用逼近定理学习非线性算子以识别微分方程。*CoRR*, abs/1910.03193, 2019年.[18] Zongyi Li、Nikola B. Kovachki、Kamyar Azizzadenesheli、Burigede Liu、Kaushik Bhattacharya、Andrew M. Stuart和Anima Anandkumar。神经算子：用于偏微分方程的图核网络。*CoRR*, abs/2003.03485, 2020年.[19] Kaushik Bhattacharya、Bamdad Hosseini、Nikola B. Kovachki和Andrew M. Stuart。用于参数化偏微分方程的模型降阶与神经网络，2021年.[20] Lu Lu、Pengzhan Jin、Guofei Pang、Zhongqiang Zhang和George Em Karniadakis。基于算子通用逼近定理通过DeepONet学习非线性算子。《自然·机器智能》, 3(3):218–229, 2021年.[21] Xiaoyi Lu和Yi Wang。使用物理信息深度神经算子网络进行辐射传热代理建模。《燃烧学会论文集》, 40(1):105282, 2024年.[22] Zongyi Li、Nikola Kovachki、Kamyar Azizzadenesheli、Burigede Liu、Kaushik Bhattacharya、Andrew Stuart和Anima Anandkumar。用于参数化偏微分方程的傅里叶神经算子。2021年.[23] M. Raissi、P. Perdikaris和G.E. Karniadakis。物理信息神经网络：一种用于求解涉及非线性偏微分方程的正问题与反问题的深度学习框架。《计算物理学杂志》, 378:686–707, 2019年.[24] Chenxi Wu、Min Zhu、Qinyang Tan、Yadhu Kartha和Lu Lu。物理信息神经网络的非自适应采样与基于残差的自适应采样综合研究。《计算力学与工程中的计算方法》, 403:115671, 2023 年 .[25]Jeremy Yu 、 Lu Lu 、 Xuhui Meng 和 George Em Karniadakis。用于偏微分方程正问题与反问题的梯度增强物理信息神经网络。《计算力学与工程中的计算方法》, 393:114823, 2022年.[26] Zongyi Li、Hongkai Zheng、Nikola B. Kovachki、David Jin、Haoxuan Chen、Burigede Liu、Kamyar Azizzadenesheli和Anima Anandkumar。用于学习偏微分方程的物理信息神经算子。*CoRR*, abs/2111.03794, 2021年.

[27] Sifan Wang、Hanwen Wang 和 Paris Perdikaris。利用物理信息 DeepONets 学习参数化偏微分方程的解算子。[28] Xiang Huang、Zhanhong Ye、Hongsheng Liu、Beiji Shi、Zidong Wang、Kang Yang、Yang Li、Bingya Weng、Min Wang、Haotian Chu、Fan Yu、Bei Hua、Lei Chen 和 Bin Dong。用于求解参数化偏微分方程的元自动解码器。*arXiv* 电子预印本，第 arXiv:2111.08823 页，2021年11月。[29] Filipe de Avila Belbute-Peres、Yi-Fan Chen 和 Fei Sha。Hyperpinn：利用物理信息超网络学习参数化微分方程。*CoRR*，abs/2111.01008，2021年。[30] Sifan Wang、Yujun Teng 和 Paris Perdikaris。理解并缓解物理信息神经网络中的梯度病态，2020年。[31] Aditi S. Krishnapriyan、Amir Gholami、Shandian Zhe、Robert M. Kirby 和 Michael W. Mahoney。刻画物理信息神经网络中可能的失效模式，2021年。[32] Pancheng Niu、Jun Guo、Yongming Chen、Yuqian Zhou、Minfu Feng 和 Yanchao Shi。改进物理信息神经网络以缓解梯度相关失效。《神经计算》，638:130167，2025年。[33] Nasim Rahaman、Aristide Baratin、Devansh Arpit、Felix Draxler、Min Lin、Fred A. Hamprecht、Yoshua Bengio 和 Aaron Courville。论神经网络的谱偏差，2019年。[34] Sifan Wang、Xinling Yu 和 Paris Perdikaris。物理信息神经网络何时以及为何训练失败：神经切线核视角，2020年。[35] Sifan Wang、Shyam Sankaran 和 Paris Perdikaris。训练物理信息神经网络时遵循因果性。《计算力学与工程中的计算方法》，421:116813，2024年。[36] Jian Cheng Wong、Abhishek Gupta、Chin Chun Ooi、Pao-Hsiung Chiu、Jiao Liu 和 Yew-Soon Ong。物理信息神经网络的进化优化：进化物理信息神经网络的前沿与机遇，2025年。[37] Shuji Chang、Piyush Agarwal、Chris McCready、Luis Ricardez-Sandoval 和 Hector Budman。用于物理信息神经网络的基于灵敏度的自适应采样。《IFAC-PapersOnLine》，58(14):325–330，2024年。第十二届 IFAC 化工过程先进控制研讨会 AD-CHEM 2024。[38] Ying-Yi Hong 和 Daryll John Medina。用于台风强度和轨迹预测的基于相位的物理信息序列到序列神经网络。《机器学习与应用》，22:100774，2025年。[39] Shengzhu Shi、Yao Li、Zhichang Guo、Boying Wu 和 Yang Zhao。通过白盒对抗攻击进行物理信息神经网络失效区域定位与细化，2025年。[40] Rajhans Singh、Ankita Shukla 和 Pavan Turaga。面向大型多样化数据集的多项式隐式神经表示，2023年。[41]Chunyu Guo、Lucheng Sun、Shilong Li、Zelong Yuan 和 Chao Wang。用于流体力学的基于切比雪夫多项式的物理信息 Kolmogorov-Arnold 网络，2025年。[42] Tianchi Yu、Yiming Qi、Ivan Oseledets 和 Shiyi Chen。谱信息神经网络：一种高效且低内存占用的物理信息神经网络。《计算与应用数学杂志》，477:117178，2026年5月。[43] Pengsong Yin、Shuo Ling 和 Wenjun Ying。用于求解偏微分方程的切比雪夫谱神经网络，2024年。[44] Zhang Jianan、Wang Yiyi、Duan Hongyi 和 Li Qingyang。基于多种算法的 Lunger 现象案例研究，2023年。

[45] Pablo D. Brubeck、Yuji Nakatsukasa 和 Lloyd N. Trefethen。结合 Arnoldi 方法的范德蒙德矩阵。《SIAM Review》，63(2):405–415，2021年。[46] Liujun Meng、Xuelin Zhang 和 Hanquan Wang。用于椭圆方程边值问题的切比雪夫多项式神经网络求解器。《East Asian Journal on Applied Mathematics》，13(2):230–245，2023年。[47] ZeinabHajimohammadi、Fatemeh Baharifard、Ali Ghodsi 和 Kourosh Parand。用于求解微分模型的分数阶切比雪夫深度神经网络（fcdnn）。《Chaos, Solitons & Fractals》，153:111530，2021年。[48] Pengsong Yin、Shuo Ling 和 Wenjun Ying。用于求解偏微分方程的切比雪夫谱神经网络，2024年。[49] S.M. Sivalingam、Pushpendra Kumar 和 V. Govindaraj。用于求解分布阶分数阶微分方程的基于切比雪夫神经网络的数值格式。《Computers & Mathematics with Applications》，164:150–165，2024年。[50] Lloyd N. Trefethen。《逼近理论与逼近实践》。SIAM，费城，2013年。[51] Samah M. Dardery 和 Mohamed M. Allan。用于求解一类奇异积分方程的切比雪夫多项式。《Applied Mathematics》，5(4):753–764，2014年。[52] J.C. Mason 和 David C. Handscomb。《切比雪夫多项式》。Chapman and Hall/CRC，第1版，2002年。[53] H. Nyengeri、R. Nizigiyimana、J.-P. Mutankana、H. Bayaga 和 F. Bayubahe。幂级数与切比雪夫级数变换公式及其在通过 Frobenius 方法和 Taylor 方法求解常微分方程中的应用。《Open Access Library Journal》，8:1–19，2021年。

附录

A 参数化偏微分方程的表述

本附录详细介绍了第4.1节中评估的参数化偏微分方程（PDE），包括伯格斯方程、艾伦-卡恩方程和非线性扩散-反应系统。这些方程描述了多种物理现象，其参数变化通过高斯随机场生成，用于测试所提出的 CPNO 框架的性能。

Burgers 方程

伯格斯方程描述了没有外力作用时不可压缩黏性流体的动力学，对于理解湍流的发展和激波的形成至关重要。我们考虑具有不同初始条件和周期性边界条件的以下伯格斯方程：

$$\begin{cases} \frac{ds}{dt} + s\frac{ds}{dx} - \nu\frac{d^2s}{dx^2} = 0, \quad (x,t) \in [0,1]^2 \\ s(x,0) = u(x), \\ s(0,t) = s(1,t), \quad \frac{ds}{dx}(0,t) = \frac{ds}{dx}(1,t) \end{cases} \tag{38}$$

我们的目标是将初始条件映射为方程的解。初始条件 $u(x)$ 通过随机生成，得到 900 个训练样本和 10 个测试样本。

扩散-反应系统

非线性扩散-反应系统刻画了反应物与产物如何因扩散和反应过程而随时间和空间演化。我们考虑如下具有零初始条件和边界条件的非线性扩散-反应系统：

$$\begin{cases} \frac{\partial s}{\partial t} = D\frac{\partial^2 s}{\partial x^2} + ks^2 + u(x), \quad (x,t) \in (0,1]^2 \\ s(0,t) = 0, s(x,0) = 0 \end{cases} \tag{39}$$

其中 $D = 0.01$ 为扩散系数， $k = 0.01$ 为反应速率。我们的目标是将源项 $u(x)$映射到偏微分方程解 $s(x)$。我们生成1000个源项 $u(x)$案例用于训练，以及10个案例用于测试。

艾伦-卡恩方程

艾伦-卡恩方程是一种基础的反应扩散方程，广泛用于描述材料科学中的相分离现象，例如二元合金的粗化过程，并描述标量序参量随时间的演化。我们考虑具有不同初始条件和周期边界条件的以下艾伦-卡恩方程：

$$\begin{cases} \partial_t u = d_1 \Delta u + d_2 u(1-u^2), \\ u(0, \boldsymbol{x}) = f_0(\boldsymbol{x}), \end{cases} \tag{40}$$

其中， ϵ 是控制界面宽度的参数。我们的目标是学习从不同初始条件 $u(x)$到对应解 $s(x,\ t)$的映射。初始条件 $u(x)$是随机生成的，其中900个样本用于训练，10个样本用于测试。

纳维-斯托克斯方程 纳维-斯托克斯方程的涡量形式描述了标量涡量场及其对应的无散速度在黏性流动中的时空演化。我们考虑具有周期边界条件的以下涡量-速度形式的二维不可压缩纳维-斯托克斯方程：

$$\begin{cases} \partial_t w(\boldsymbol{x}, t) + u(\boldsymbol{x}, t) \cdot \nabla w(\boldsymbol{x}, t) = \nu \Delta w(\boldsymbol{x}, t) + f(\boldsymbol{x}), \\ w(\boldsymbol{x}, 0) = w_0(\boldsymbol{x}) \end{cases} \tag{41}$$

源项为 $f(\boldsymbol{x}) = 0.1(\sin(2\pi(x+y)) + \cos(2\pi(x+y)))$，且 $\nu = 0.1$。我们的目标是学习映射$w(x,t)$ 到 $w(x,t)$，其中 $\Delta t = 0.1s$。我们生成了500个与傅里叶神经算子相似的案例[22]，分辨率为 $32 * 32$。

B CPNO 截断误差的证明

与切比雪夫基相关的截断误差，以及神经网络引入的系数估计误差：

$$E_{tru} = \left\| u(x;\theta) - \sum_{k=0}^{N} c_k^*(\theta) T_k(x) \right\|_{L^\infty} \tag{42}$$

其中， e_k^* 是通过加权内积得到的真实切比雪夫系数：

$$c_k^* = \frac{\langle u(x;\theta), T_k \rangle}{\langle T_k, T_k \rangle}, \quad \langle f, g \rangle = \int_{-1}^{1} \frac{f(x)g(x)}{\sqrt{1-x^2}}\, dx \tag{43}$$

且切比雪夫多项式的正交性满足 $\langle T_m,\ T_n \rangle = \frac{\pi}{2}\delta_{mn}$（其中 $\langle T_0,\ T_0 \rangle = \pi$，当 $m,\ n \neq 0$时）。

截断误差衡量切比雪夫基使用有限的 N 项级数逼近 $u(x;\theta)$ 的内在能力。真实系数 e_k^* 通过正交投影计算：

$$c_k^* = \frac{\int_{-1}^{1} u(x;\theta) T_k(x) w(x)\, dx}{\int_{-1}^{1} T_k(x)^2 w(x)\, dx} \tag{44}$$

其中，权函数 $w(x) = 1/\sqrt{1-x^2}$ 确保正交性，分母 $\langle T_k,\ T_k \rangle = \pi/2$ 适用于 $k \geq 1$（其中 $\langle T_0,\ T_0 \rangle = \pi$）。因此， $u(x;\theta)$ 可表示为无穷级数表示$u(x;\theta) = \sum_{k=0}^{\infty} e_k^* T_k(x)$，而截断误差对应于尾项 $\sum_{k=N+1}^{\infty} e_k^* T_k(x)$。

鉴于 $u(x;\theta)$ 是一个光滑函数，我们假设它在一个伯恩斯坦椭圆内解析，该椭圆的焦点位于 ± 1，半长轴为.....，因此 $u(x;\theta)$ 可以扩展为该椭圆内的全纯函数，这意味着切比雪夫系数 e_k^* 随 k 指数衰减，即对于某些 $C > 0$ 和 $\rho > 1$，有 $|c_k^*| \leq C\rho^{-k}$，其中 ρ 取决于解析半径。通过计算级数尾项的 L^∞ 范数，可以得到截断误差的上界。利用切比雪夫多项式的一致有界性（对于 $x \in [-1, 1]$，有 $|T_k(x)| \leq 1$），可得：

$$\left| \sum_{k=N+1}^{\infty} c_k^* T_k(x) \right| \leq \sum_{k=N+1}^{\infty} |c_k^*| \cdot |T_k(x)| \leq \sum_{k=N+1}^{\infty} C\rho^{-k}. \tag{45}$$

这是一个几何级数，由于 $\rho > 1$，它收敛。其和为：

$$\sum_{k=N+1}^{\infty} \rho^{-k} = \rho^{-(N+1)} \cdot \frac{1}{1-\rho^{-1}} = \frac{\rho^{-(N+1)}}{\rho - 1}. \tag{46}$$

代入系数界后，得到：

$$\sum_{k=N+1}^{\infty} C\rho^{-k} = C \cdot \frac{\rho^{-(N+1)}}{\rho - 1}. \tag{47}$$

注意到 $\rho^{-(N+1)} = \rho^{-N} \cdot \rho^{-1}$，并调整常数 C 以计入初始偏移以及 u 在椭圆域上的上确界，该界可简化为：

$$\left\| u(x;\theta) - \sum_{k=0}^{N} c_k^* T_k(x) \right\|_{L^\infty} \leq \sup_{x\in[-1,1]} \left| \sum_{k=N+1}^{\infty} c_k^* T_k(x) \right| \leq \frac{C\rho^{-N}}{\rho - 1}, \tag{48}$$

其中，将 C 重新缩放，以反映解析区域内 u的最大值。该结果证实了切比雪夫基的谱收敛性，并将其与传统多项式基的代数收敛区分开来。

C 非线性微分算子 Gram矩阵的分析

为通过实证验证前述理论分析，并揭示两种基在更实际场景下数值稳定性的差异，我们设计了一项关键数值实验。我们采用阻尼拉普拉斯算子作为测试案例：

$$\mathcal{L}[u] = \alpha \frac{d^2u}{dx^2} + u, \quad \text{with } \alpha = 0.01. \tag{49}$$

这一算子的选择很有代表性。与最简单的恒等算子不同，加入二阶导数项后，算子能够作用于不同尺度，从而更全面地测试基在微分运算下的稳定性。同时，通过选择较小的α，我们确保微分项的影响足以区分两种基的性能，同时又不至于过大，以免单项式基因数值溢出而立即失效。我们为单项式基 $\{\psi_k(x) = x^k\}$ 和切比雪夫基 $\{\psi_k(x) = T_k(x)\}$构造格拉姆矩阵$\mathbf{G}$ ，其元素由积分 $G_{jk} = \langle \mathcal{L}[\psi_j], \mathcal{L}[\psi_k] \rangle_{L^2}$定义。数值积分采用高精度高斯-勒让德求积，并分析条件数 $\kappa(\mathbf{G})$ 随最大多项式次数 N 从 2 变化至 14 时的变化。

图15 清楚地揭示了两种基在数值稳定性方面的显著差距：单项式基的格拉姆矩阵条件数， $\kappa(\mathbf{G}_M)$，随着次数 N的增加呈指数式爆炸增长。在 $N = 14$时，其条件数高达 $10^{11},$，由于舍入误差被灾难性地放大，任何基于该矩阵的数值计算都完全不可靠。相比之下，切比雪夫基的条件数， $\kappa(\mathbf{G}_T)$，尽管也在增加，但增长速度慢得多，呈现低阶多项式趋势。

值得注意的是，与恒等算子的情况相比，切比雪夫基的条件数确实有所增加。这是因为引入微分算子 $\alpha u''$ 扰乱了该基在标准 L^2 内积下的近正交结构。然而，最关键的观察在于增长速率存在数量级上的差异。切比雪夫基固有的稳定性确保了这种微分扰动造成的条件性恶化受到控制，且呈多项式增长。相比之下，单项式基本身就存在固有的病态性，而微分算子只会加剧其固有的指数不稳定性。在 $N = 14$处，两种条件数之比约为 $10^5,$，这清楚地表明了切比雪夫基的绝对

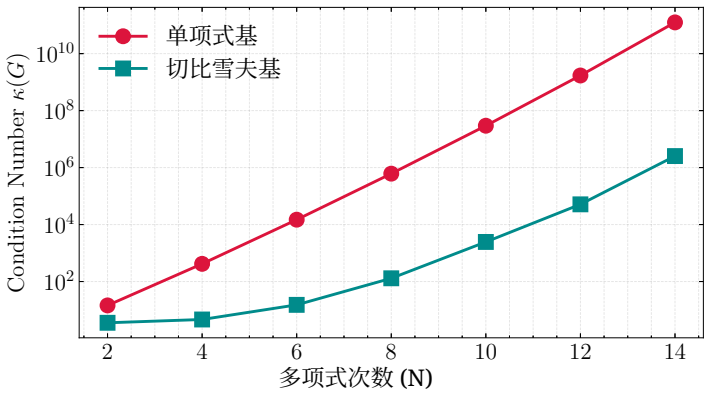

图15　　：条件数 $k(G)$ 随多项式阶数 N 变化的比较　　.

优势。这次数值实验为我们的核心理论提供了有力的实证依据：凭借优越的内在数学性质，切比雪夫基在构建涉及微分算子的学习框架时，能够使优化问题显著更加稳定。这进一步确立了切比雪夫基作为高阶、高精度算子学习基础选择的地位。