arXiv:2601.03668v1 [math.NA] 7 Jan 2026

用于求解非光滑偏微分方程的间断伽辽金有限元算子网络*

KAPIL CHAWLA†，YOUNGJOON HONG‡，JAE YONG LEE§，和 SANGHYUN LEE†

摘要。我们提出了间断伽辽金有限元算子网络（DG– FEONet），这是一种无数据算子学习框架，结合了间断 Galerkin（DG）方法与神经网络的优势，用于求解具有不连续系数和非光滑解的参数化偏微分方程（PDE）。与 DeepONet 和傅里叶神经算子等传统算子学习模型不同，这些模型需要大量配对数据集，且通常难以处理尖锐特征；我们的方法采用对称内罚伽辽金（SIPG）格式，最小化基于 DG 的弱形式残差。DG-FEONet 通过神经网络预测逐单元解系数，从而无需预先计算输入—输出对即可进行无数据训练。我们通过收敛分析提供理论依据，并在一系列一维和二维偏微分方程问题上验证模型性能，结果表明，该模型能够准确恢复不连续性，在参数空间中具有很强的泛化能力，并展现出可靠的收敛率。我们的结果凸显了将局部离散化格式与机器学习相结合的潜力，从而在具有挑战性的偏微分方程问题中实现稳健且能感知奇异性的算子逼近。

关键词。间断伽辽金方法、有限元方法、物理信息驱动的算子学习、感知不连续性的学习

AMS 学科分类。65M60,65N30, 68T20, 68U07

1. 引言。偏微分方程（PDE）为复杂物理系统建模奠定了数学基础，应用范围涵盖流体流动、热传递、电磁学和弹性力学 [16]。虽然某些简单情形存在解析解，但现实中的偏微分方程问题通常涉及不规则区域、非线性或参数不确定性，因此数值方法对于实际计算至关重要 [5]。在这些方法中，有限元法（FEM）以其灵活性和准确性脱颖而出，尤其擅长处理复杂几何形状和边界条件。

尽管传统数值求解器可靠，但在反复求解参数化偏微分方程时，计算成本可能很高，例如设计优化、不确定性量化或控制应用。这推动了人们对机器学习方法的浓厚兴趣，这类方法能够近似算子本身。早期方法包括物理信息神经网络（PINNs） [39], ，它们通过将偏微分方程残差嵌入损失函数，实现无需数据的学习。然而，PINNs 往往难以处理陡峭梯度、边界层和不连续系数 [31, 34, 2]。此外，它们通常针对单个特定问题实例进行训练，因而泛化能力有限。

近年来，神经算子的概念受到关注，它能够直接学习参数化偏微分方程中随参数变化的解算子 [33, 26]。DeepONet [32]和傅里叶神经算子（FNO） [29] 等神经算子框架，旨在学习无限维函数空间之间的映射，

* **资助：** 本研究由美国能源部科学办公室能源地球计划资助，项目编号：DE-SC0024703。

†佛罗里达州立大学数学系，美国佛罗里达州塔拉哈西市学术路1017号，邮编32306（kc25bc@fsu.edu；slee17@fsu.edu）。

‡首尔国立大学数学科学系，大韩民国首尔市冠岳区冠岳路1号（hongyj@snu.ac.kr）。§中央大学人工智能系，大韩民国首尔市铜雀区黑石路84号（jaeyong@cau.ac.kr）。

从而无需针对每种新的参数设置重新求解偏微分方程，即可高效预测解。除 DeepONet 和 FNO 外，人们还提出了各种神经算子，包括基于图的 [4, 38], 基于小波的 [20],以及受 Transformer 启发的模型 [7, 48, 21], ，旨在处理不规则区域并提升跨参数空间的泛化能力。然而，这些方法通常依赖于使用大量输入-输出解对进行监督训练，而生成此类数据集可能成本高昂或难以实现，尤其是在高维或非线性问题中。

为克服这些挑战，近期研究探索了将受 PINN 方法启发的物理信息驱动损失引入算子学习框架 [19]。例如， [49] 和 [30] 分别将物理信息残差损失引入 DeepONet 和 FNO 架构，使网络能够在训练过程中利用控制偏微分方程的结构。在此研究方向的基础上，有限元算子网络（FEONet） [27, 22] 最近被提出，作为一种物理感知、无需数据的替代方案。FEONet 将有限元法的连续 Galerkin（CG）表述融入神经网络训练循环，通过最小化弱形式残差，而非解误差来实现训练。FEONet 通过预测有限元系数并利用基函数重构解，消除了对标注数据的需求，同时继承了有限元法的关键特性，例如网格灵活性和边界条件的精确施加。

然而，连续伽辽金公式要求解在单元边界处连续，且本身并不满足局部守恒性质。因此，它不适用于解或系数存在不连续性的情况，例如激波、裂隙介质或多相流。在这类问题中，连续 Galerkin 方法可能会引入伪振荡，且无法准确捕捉局部特征 [13]。本文提出 DG-FEONet，这是对 FEONet 的一种新型扩展，基于采用对称内罚 Galerkin（SIPG）公式的间断伽辽金（DG）方法。DG 框架放宽了单元间连续性要求，并通过数值通量和惩罚项保证一致性。因此，DG方法尤其适用于求解具有不连续性、非光滑解或非均质介质的偏微分方程，同时保持局部守恒律。

我们的方法使用神经网络逐单元预测局部 DG 系数，并基于 SIPG 残差定义损失函数。与原始 FEONet 一样，DG-FEONet 无需预计算解，并以无数据方式进行训练。然而，借助 DG 框架，它显著提升了连续伽辽金方法难以处理的问题中的准确性和鲁棒性。通过一系列数值实验，我们证明 DG-FEONet 能够准确捕捉参数化偏微分方程中的尖锐特征和不连续性，并为在复杂条件下进行算子学习提供一种可靠的无数据方法。

本文其余部分安排如下。第 2 节概述相关研究。第 3节介绍 DG-FEONet 的公式，包括简要概述 DG-FEM，以及基于对称内罚伽辽金（SIPG）格式构造弱形式损失。第 4 节对所提出的框架进行理论收敛性分析。第 5节通过一系列数值实验验证 DG-FEONet，这些实验涉及具有不连续系数的一维和二维参数化偏微分方程。我们还评估了其收敛行为，并展示了该方法在不同网格分辨率下的稳健性。最后，第6节总结全文，并指出未来的研究方向。

2. 相关工作。 有限元法长期以来一直是数值求解偏微分方程的基石 [53, 24]。连续伽辽金公式是经典的有限元法方法，通过全局连续的基函数对解进行近似。这种方法尤其适用于解平滑的问题。然而，强制要求单元间连续性，限制了其在具有陡峭梯度或不连续性问题中的适用性。间断伽辽金方法 [44, 40] 通过允许解在单元边界处不连续，克服了这些局限。间断伽辽金方法因其局部守恒性质、网格处理的灵活性以及与高阶近似的兼容性，已广泛应用于守恒律和多相流问题。在各种间断伽辽金方法变体中，对称内罚伽辽金（SIPG）方法 [35, 40] 因其稳定性和易于实现，在椭圆型和抛物型问题中广受欢迎。

深度学习的最新进展使神经网络成为求解奇异摄动偏微分方程的强大工具。针对这些难题，研究者也探索了物理信息机器学习 [2, 46, 8], ，但这类方法通常缺乏应用于更广泛场景所需的可扩展性。为更好地处理不连续性，近期研究提出了专门针对不连续解的基于物理信息神经网络的方法 [31, 50]。在神经算子领域，分量傅里叶神经算子 [28] 通过在其架构中引入渐近展开来处理奇异摄动，为提升基于神经网络的求解器的鲁棒性和精度提供了颇具前景的方向。

近期研究探索了将经典数值方法与神经网络相结合的混合方法，以发挥两种范式各自的优势。这些方法旨在提升偏微分方程求解器的效率、泛化能力和物理一致性。例如，研究者利用神经网络学习数据驱动的离散格式 [3], ，有效取代传统的有限差分算子。FEONet [27, 22] 将深度学习与连续伽辽金有限元法相结合，直接学习有限元法基下的解的展开系数，无需标注数据。类似地，谱方法包括谱系数算子网络学习（SCLON）[11] 和无监督勒让德伽辽金网络（ULGNet）[10] ，它们学习谱基或伽辽金基下的解系数，从而实现具备物理认知能力的无数据训练。近期，神经算子单元法（NOEM）[36] 应运而生，它将神经算子作为替代单元嵌入变分有限元框架，以降低多尺度问题中的网格剖分成本。最近， [9, 43, 17] 提出了基于间断伽辽金的神经网络框架，其中间断伽辽金公式与神经网络相结合，可用作试验/检验空间、混合求解器或增强基。

3. 模型描述。我们考虑使用结合算子学习的间断伽辽金有限元法来近似参数化偏微分方程的解算子。具体而言，我们考虑以下控制方程：

$$-\nabla \cdot (\varepsilon(x)\nabla u) + F(u) = f(x), \quad x \in \Omega, \tag{3.1}$$

在区域 $\Omega \subset \mathbb{R}^d (d = 1, 2, 3)$ 中，其中标量解 $u(x)$: $\Omega \to \mathbb{R}$ 在 $\partial\Omega$ 上满足适当的边界条件，且边界分解为狄利克雷部分和诺伊曼部分： $\partial\Omega = \partial\Omega_D \cup \partial\Omega_N$, 以及 $\partial\Omega_D \cap \partial\Omega_N = \emptyset$. 作为一个典型示例，我们考虑

$$F(u) = b(x)\nabla u(x) + c(x)u(x),$$

这通常称为对流–扩散–反应方程。其中， $\varepsilon(x)$、$b(x)$ 和 $c(x)$ 是空间变化系数， $f(x)$ 是给定的源函数，而这些量在 Ω 上可能不连续。我们的目标是学习其底层解算子

$$\mathcal{G} : (\varepsilon, b, c, f) \mapsto u, \tag{3.2}$$

它将一组空间变化系数和源项映射为偏微分方程（ $u(x)$（3.1）的相应解。以下任一场 $\varepsilon(x)$ 、 $b(x)$ 、 $c(x)$ 或 $f(x)$均可作为参数输入，具体取决于问题设定；DG-FEONet旨在针对所选输入配置逼近 $\mathcal{G}$ 。

3.1. 间断Galerkin有限元方法。与经典连续Galerkin有限元方法不同，间断伽辽金方法 [51, 41, 40, 12] 允许试探空间和检验空间在单元边界处不连续。设 $\mathcal{T}_h$ 为区域 Ω 的形状规则剖分，其中包含单元 K。 $\mathcal{T}_h$ 中的所有边构成的集合记为 $\varepsilon_h = \varepsilon_h^{\text{int}} \cup \varepsilon_h^D \cup \varepsilon_h^N$，该集合可分解为三个互不相交的子集。其中， $\varepsilon_h^{\text{int}}$为所有内部边的集合， ε_h^D 和 ε_h^N 分别为施加Dirichlet条件和Neumann条件的所有边界边的集合。为方便起见，我们还记 $\varepsilon_h^I := \varepsilon_h^{\text{int}} \cup \varepsilon_h^D$ 为内部边与狄利克雷边界边的并集。

对于每条边 $e \in \varepsilon_h$，令 K^+ 和 K^- 为满足 $e = \partial K^+ \cap \partial K^-$ 的相邻单元。令 $\mathbf{n}^+$ 和 $\mathbf{n}^-$ 分别为指向 ∂K^+ 和 ∂K^- 的单位外法向量。对于给定的向量函数 $\mathbf{q}$ 和标量函数 v，我们将加权平均算子 $\{\cdot\}$ 和跳跃算子 $[\![\cdot]\!]$ 定义为

$$\{\mathbf{q}\} = \frac{\mathbf{q}|_{K^+} + \mathbf{q}|_{K^-}}{2} \text{ on } e, \qquad [\![v]\!] = v|_{K^+}\mathbf{n}^+ + v|_{K^-}\mathbf{n}^- \text{ on } e.$$

如果 $e \in \partial\Omega$，则 e 仅属于一个单元 K，我们定义

$$[\![v]\!] = \{v\} = (v|_K)|_e.$$

对于 DG空间离散，我们定义离散解空间 $V_h \subset L^2(\Omega)$ 为破碎多项式空间：

$$V_h := \{v \in L^2(\Omega) \mid v|_K \in \mathbb{P}_r(K),\ \forall K \in \mathcal{T}_h\}, \tag{3.3}$$

其中， $\mathbb{P}_r(K)$ 表示单元 K 上次数不超过 r 的多项式空间。

离散解 $u_h \in V_h$ 由对称内罚 Galerkin（SIPG）形式 [45] 确定：求 $u_h \in V_h$ ，使得

$$a(u_h, v_h) = \ell(v_h) \quad \forall v_h \in V_h, \tag{3.4}$$

其中，双线性形式 $a(\cdot,\cdot)$ 定义为

$$\begin{aligned} a(u,v) = &\sum_{K\in\mathcal{T}_h} \int_K \varepsilon\nabla u\cdot\nabla v + (b\cdot\nabla u)v + cuv\,dx \\ &- \sum_{e\in\mathcal{E}_h^{\text{int}}} \int_e \{\!\{\varepsilon\nabla u\}\!\}\cdot[\![v]\!] + \{\!\{\varepsilon\nabla v\}\!\}\cdot[\![u]\!]\,ds \\ &+ \sum_{e\in\mathcal{E}_h^{\text{int}}} \int_e \frac{\sigma}{h_e}[\![u]\!]\cdot[\![v]\!]\,ds - \sum_{e\in\mathcal{E}_h^{\text{int}}} \int_e \widehat{bu}\cdot[\![v]\!]\,ds - \sum_{e\in\mathcal{E}_h^{\partial}} \int_e \widehat{bu}\cdot v\,ds \\ &- \int_{\partial\Omega_D} \varepsilon\nabla u\cdot\mathbf{n}\,v\,ds - \int_{\partial\Omega_D} \varepsilon\nabla v\cdot\mathbf{n}\,u\,ds + \int_{\partial\Omega_D} \frac{\sigma}{h}uv\,ds. \end{aligned} \tag{3.5}$$

这里， $[\![\cdot]\!]$ 表示内部面上的跳跃， $\{\!\{\cdot\}\!\}$ 表示平均值， h_e 表示局部网格尺寸， $\sigma > 0$ 表示惩罚参数。对流项的数值通量 $\widehat{bu}$ 定义为

$$\widehat{bu} = \frac{1}{2}(b^+ + b^-)\{u\} + \frac{1}{2}\left|b^+ + b^-\right|[\![u]\!],$$

第一项保证一致性，而第二项引入数值耗散以确保稳定性。

线性形式定义为：

$$\begin{aligned}\ell(v) = &\sum_{K\in\mathcal{T}_h}\int_K f(x)\,v(x)\,dx - \sum_{e\in\mathcal{E}_h^D}\int_e \big(\varepsilon\nabla v\cdot\mathbf{n}\big)\,u_D\,ds \\ &+ \sum_{e\in\mathcal{E}_h^D}\int_e \frac{\sigma}{h_e}\,u_D\,v\,ds + \sum_{e\in\mathcal{E}_h^N}\int_e u_N\,v\,ds.\end{aligned} \tag{3.6}$$

其中， u_D 在 $\partial\Omega_D$ 上给定狄利克雷边界条件，而 u_N 在 $\partial\Omega_N$上满足诺伊曼边界条件。

为将 DG 解表示为有限维形式，我们将 $u_h \in V_h$ 表示为局部支撑基函数 $\{\phi_k(x)\}_{k=1}^{N(h)}$ 的线性组合：

$$u_h(x) = \sum_{k=1}^{N(h)} \alpha_k(\omega)\phi_k(x),$$

其中， $\alpha_k(\omega) \in \mathbb{R}$ 是取决于问题数据 ω的展开系数，例如扩散系数 $\varepsilon(x)$、源项 $f(x)$或其他参数输入。这些系数通常通过求解由 DG 变分形式导出的线性系统来计算。

求解间断伽辽金离散产生的线性系统，计算开销很大，而且由于所得矩阵的自由度增加、具有块结构，求解过程往往并不简单 [14, 25]。FEONet框架提供了一种很有前景的替代方案：它直接学习从问题参数到解系数的映射，从而无需显式组装和求解线性系统。

3.2. 间断伽辽金有限元算子网络。使用传统 DG方法求解参数化偏微分方程时，每遇到一组新的输入参数样本（例如系数、源项），都需要组装并求解一个线性系统。DG-FEONet旨在学习从问题输入到解系数的直接映射，而不是反复求解该系统。

令 $\{\phi_k\}_{k=1}^{N(h)}$ 表示离散空间 V_h的局部间断伽辽金基。那么，近似解表示为：

$$u_h(x;\omega) = \sum_{k=1}^{N} \alpha_k(\omega)\phi_k(x),$$

其中， $\omega \in \Omega^{\text{param}}$ 表示一个参数输入样本。目标是利用神经网络 $\mathcal{N}_\theta$ 学习映射 $\omega \mapsto \alpha(\omega)$：

$$\alpha(\omega) \approx \mathcal{N}_\theta(\omega).$$

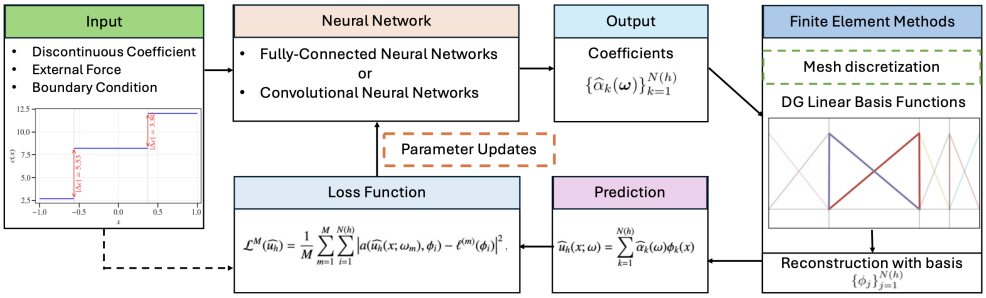


FIG. 1. *Schematic overview of the DG-FEONet framework. The PDE inputs, which we emphasize that these can be discontinuous, are passed to a neural network that predicts the coefficients for DG-FEM solutions. These coefficients are used to reconstruct the DG solution using basis functions defined over a triangulated domain. The network is trained in an data-free manner using the residual of the DG weak formulation as the loss function.*

这一思想受到物理信息神经网络（PINNs）的启发。PINNs 通过最小化偏微分方程残差，将物理定律融入学习过程。然而，与在配点处直接逼近解的经典 PINNs 不同，DG-FEONet 在有限维函数空间中运行。神经网络预测间断伽辽金基展开的系数，从而学习参数化解算子。这使 DG-FEONet 更接近 FEONet、DeepONet 和 FNO 等算子学习框架，但其独特优势在于采用间断伽辽金离散，更好地处理不连续性。

另一个关键区别是，DG-FEONet 以无数据方式训练，无需获取带标签的解。相反，我们基于 DG 变分公式的残差定义了一种物理信息损失：

$$\mathcal{L}(\theta) = \frac{1}{M}\sum_{m=1}^{M}\sum_{i=1}^{N}\left|a(u_h^{(m)},\phi_i) - \ell^{(m)}(\phi_i)\right|^2, \tag{3.7}$$

其中，$u_h^{(m)}(x) = \sum_{k=1}^{N}\mathcal{N}_\theta(\omega_m)_k\phi_k(x)$ 是第 m 个输入对应的预测解。该残差是相对于 DG 基中的试验函数 ϕ_i 进行评估的。

这种表述使得 能够直接基于偏微分方程的物理规律进行无数据算子学习，并利用 DG方法的结构。DG基函数的局部性和灵活性使该框架尤其适用于系数或解不连续的问题，而基于连续伽辽金的方法无法处理不连续解。图 1 示意展示了 DG-FEONet框架。

3.3. 损失表述与训练目标。 DG-FEONet框架通过最小化与对称内罚间断伽辽金离散相关的弱形式残差进行训练。给定如下形式的 DG 解神经网络预测

$$u_h(x;\omega) = \sum_{k=1}^{N}\hat{\alpha}_k(\omega)\phi_k(x),$$

其中，$\hat{\alpha}_k(\omega)$ 为网络预测的系数，$\{\phi_k\}_{k=1}^N$ 为局部间断伽辽金基，我们通过强制满足离散变分形式来定义训练目标：

$$a(u_h^{(m)},\phi_i) \approx \ell^{(m)}(\phi_i), \quad \forall i = 1,\dots,N,$$

算法 3.1DG-FEONet训练过程（SIPG 公式）

要求： 训练样本 $\{f(x;\,\omega_m)\}_{m=1}^{M}$
确保： 预测系数 $\widehat{\alpha}(\omega_m) = [\widehat{\alpha}_k(\omega_m)]_{k=1}^{N(h)}$ and 预测解 $\widehat{u}_h(x;\omega_m) = \sum_{k=1}^{N(h)} \widehat{\alpha}_k(\omega_m)\varphi_k(x)$

预计算：
组装 SIPG刚度矩阵 $A^{(s)}$ （罚项/跳跃项）。
组装对流矩阵 $A^{(c)}$ （数值通量和边界项）。
组装载荷向量 $F^{(m)}$，针对 $m = 1,\ \ldots\ldots,\ M$，使用 $f(x;\,\omega_m)$。
对于 $n = 1$、2、……、N_{epochs} **执行**
  $\mathcal{L}_{\text{DG}} \leftarrow 0$
  对 $m = 1$ **执行**$, 2, \ldots, M$ **do**
    $\widehat{\alpha}(\omega_m) \leftarrow$ DG-FEONet$(f(x;\,\omega_m))$
    $\mathcal{R}^{(m)} \leftarrow (A^{(s)} + A^{(c)})\widehat{\alpha}(\omega_m) - F^{(m)}$
    $\mathcal{L}_{\text{DG}} \leftarrow \mathcal{L}_{\text{DG}} + \|\mathcal{R}^{(m)}\|^2$
  结束循环
  更新 DG-FEONet 参数以最小化 $\mathcal{L}_{\text{DG}}$
结束循环

对于每个样本 $\omega_m \in \Omega_{\text{param}},\ m = 1,\ \ldots\ldots,\ M$。这里，$a(\cdot,\ \cdot)$ 是对称内罚间断伽辽金双线性形式，其中包括体积分、跨内部面的数值通量项和惩罚项，如方程（3.4）中所定义。

我们将第 m 个输入样本和第 i 个测试函数对应的残差定义为

$$\mathcal{R}_i^{(m)} := a(u_h^{(m)}, \phi_i) - \ell^{(m)}(\phi_i),$$

并将所有输入样本和测试函数的残差平方平均值定义为训练损失：

$$\mathcal{L}_{\text{DG}}(\theta) = \frac{1}{M}\sum_{m=1}^{M}\sum_{i=1}^{N}\left|\mathcal{R}_i^{(m)}\right|^2. \tag{3.8}$$

该损失函数以无数据方式衡量神经网络预测对每个样本的离散变分陈述的满足程度。无需带标签的解 $u(x;\omega)$，因此这种方法完全由物理信息驱动。通过最小化间断伽辽金测试空间上的残差，网络能够学习从输入函数（例如系数或源项）到解系数的映射，从而近似偏微分方程的解算子。

为确保计算效率，我们预先组装与刚度、对流和界面通量项对应的对称内罚间断伽辽金矩阵。训练期间，神经网络输出预测系数 $\hat{}\alpha(\omega_m)$，随后将其代入组装好的系统，以计算残差及其梯度，用于反向传播。

整体训练流程总结于算法 3.1，其中概述了 DG-FEONet 优化过程的关键步骤。

4. DG-FEONet 的收敛性分析。本节中，我们使用第 3 节引入的间断伽辽金设定。特别地，$\Omega \subset \mathbb{R}^d$ 表示物理区域，Ω_{param} 表示配备概率测度 $\mathbb{P}_{\Omega_{\text{param}}}$, V_h表示式（3.3）中定义的间断伽辽金分片空间，而 $a(\cdot,\cdot)$ 表示对称内罚间断伽辽金双线性形式。

DEFINITION 4.1（DG 能量范数）。对于 $v_h \in V_h$，定义

$$\|v_h\|_{\mathrm{DG}}^2 := \sum_{K\in\mathcal{T}_h} \left(\|\sqrt{\varepsilon}\nabla v_h\|_{L^2(K)}^2 + \|v_h\|_{L^2(K)}^2 \right) + \sum_{e\in\mathcal{E}_h^{\mathrm{int}}} \frac{\sigma}{h_e} \|[\![v_h]\!]\|_{L^2(e)}^2 + \sum_{e\subset\partial\Omega} \frac{\sigma}{h_e} \|v_h\|_{L^2(e)}^2. \tag{4.1}$$

假设4.2（强制性）。假设第3节中引入的对流-扩散-反应算子的系数满足标准的正性条件

$$c(x) - \tfrac{1}{2}\nabla\cdot b(x) \geq c_0 \geq 0 \qquad \text{a.e. in } \Omega, \tag{4.2}$$

并且 SIPG罚参数 $\sigma > 0$ 取值足够大。那么存在一个与 h无关的常数 $C_{\mathrm{coer}} > 0$，使得

$$a(v_h, v_h) \geq C_{\mathrm{coer}} \|v_h\|_{DG}^2 \qquad \forall v_h \in V_h. \tag{4.3}$$

当惩罚参数足够大时，SIPG扩散项的强制性已得到充分确立；例如，参见 [15]。对于对流-扩散-反应问题，在正性条件（4.2）下，DG 双线性型的稳定性可由标准DG理论推出；例如，参见 [42]。

4.1. 系数表示。设 $\{\phi_i\}_{i=0}^{N_h}$ 是 V_h的一组基。我们定义与 $a(\cdot,\ \cdot)$ 相关联的 DG 算子矩阵 $A \in \mathbb{R}^{(N_h+1)\times(N_h+1)}$ 为

$$A_{ij} := a(\phi_j, \phi_i). \tag{4.4}$$

在假设 4.2下， A 可逆；参见附录 A。

4.2. 逼近误差。

LEMMA 4.3. *[23]* 设 $A \in \mathbb{R}^{(N_h+1)\times(N_h+1)}$ 可逆（不一定对称或正定）。记其最小和最大奇异值分别为 $\sigma_{\min}(A)$和 $\sigma_{\max}(A)$。那么，对于任意 $\mathbf{x} \in \mathbb{R}^{N_h+1}$,

$$\sigma_{\min}(A)\,\|\mathbf{x}\|_2 \ \leq\ \|A\mathbf{x}\|_2 \ \leq\ \sigma_{\max}(A)\,\|\mathbf{x}\|_2.$$

等价地，令 $S = A^\top A$ （它是对称正定的），可得

$$\lambda_{\min}(S) = \sigma_{\min}(A)^2,\ \lambda_{\max}(S) = \sigma_{\max}(A)^2,$$

and

$$\lambda_{\min}(S)\,\|\mathbf{x}\|_2^2 \ \leq\ \|A\mathbf{x}\|_2^2 \ \leq\ \lambda_{\max}(S)\,\|\mathbf{x}\|_2^2.$$

在陈述逼近结果之前，我们先介绍系数空间中的神经网络逼近。令 $\mathcal{N}_n$ 表示容量参数为 n的神经网络类。令 $F \in \mathbb{R}^{N_h+1}$ 为 DG载荷向量，其分量为

$$F_i := \ell(\phi_i), \qquad i = 0, \ldots, N_h,$$

其中 $\{\phi_i\}_{i=0}^{N_h}$ 是 V_h 的基，且 $\ell(\cdot)$ 是式（3.6）中定义的线性泛函。我们定义损失泛函

$$\mathcal{L}(\alpha) := \|A\alpha - F\|_2^2,$$

以及总体最小化器

$$\widehat{\alpha}(n) := \arg\min_{\alpha\in\mathcal{N}_n} \mathcal{L}(\alpha). \tag{4.5}$$

有了这些记号后，我们现在证明系数空间中逼近误差的收敛性。

THEOREM 4.4. *[27]* 设 α^* 表示真实系数函数，设 $\widehat{\alpha}(n)$ 表示使用 n 与样本量无关的容量参数训练得到的 ***DG-FEONet***模型 M的逼近结果。则有如下收敛性：

$$\|\alpha^* - \widehat{\alpha}(n)\|_{L^2(\Omega_{param})} \to 0 \quad as\ n \to \infty. \tag{4.6}$$

4.3. 泛化误差。 我们采用 Rademacher 复杂度这一概念，它是衡量函数类拟合纯随机噪声能力的定量指标 [18, 47]。

DEFINITION 4.5. 设 $\{X_i\}_{i=1}^M$ 是一组独立同分布的随机变量。函数类 $\mathcal{G}$ 的 ***Rademacher*** 复杂度定义为

$$R_M(\mathcal{G}) \;=\; \mathbb{E}_{\{X_i,\varepsilon_i\}_{i=1}^M}\left[\sup_{f\in\mathcal{G}}\left|\frac{1}{M}\sum_{i=1}^{M}\varepsilon_i f(X_i)\right|\right],$$

其中 $\{\varepsilon_i\}_{i=1}^M$ 表示一列独立同分布的 ***Rademacher***随机变量，即$\mathbb{P}(\varepsilon_i = 1) = \mathbb{P}(\varepsilon_i = -1) = \frac{1}{2}$对每个 $i = 1,$ ....., M均成立。

THEOREM 4.6. 设 $\mathcal{G}$ 是一个 b-一致有界函数类。那么，对于任意$\delta > 0$，有

$$\sup_{f\in\mathcal{G}}\left|\frac{1}{M}\sum_{i=1}^{M} f(X_i) - \mathbb{E}[f(X)]\right| \;\le\; 2R_M(\mathcal{G}) + \delta,$$

以至少 $1 - \exp\Big(-\frac{M\delta^2}{2b^2}\Big)$. 的概率

证明。 证明直接由 [47] 中的定理 4.10 得出。 □

接下来，我们引入以下函数类：

$$\mathcal{G}_n := \{\, |A\alpha - F|^2 : \alpha \in \mathcal{N}_n \,\}. \tag{4.7}$$

由引理 4.3，存在一个常数 $C > 0$ 使得

$$\begin{aligned}\|A\alpha - F\|_{L^\infty(\Omega_{\text{param}})} &\le \|A\alpha\|_{L^\infty(\Omega_{\text{param}})} + \|F\|_{L^\infty(\Omega_{\text{param}})} \\ &\le C\Big(\|\alpha\|_{L^\infty(\Omega_{\text{param}})} + \|f\|_{C(\Omega_{\text{param}};L^1(\Omega))}\Big).\end{aligned} \tag{4.8}$$

在我们的设定下，下面的引理可由定理 4.6 直接得出。

LEMMA 4.7. 设 $\{\omega_m\}_{m=1}^M$ 是从分布 $\mathbb{P}_{\Omega_{param}}$ 中抽取的独立同分布样本。那么，对于任意 $\delta > 0$，以至少 $1 - 2\exp\big(-\frac{M\delta^2}{32\tilde{b}^2}\big)$ 的概率，有

$$\sup_{\alpha\in\mathcal{N}_n}\left|\mathcal{L}^M(\alpha) - \mathcal{L}(\alpha)\right| \;\le\; 2R_M(\mathcal{G}_n) + \tfrac{\delta}{2}. \tag{4.9}$$

现在利用引理 4.7建立泛化误差的收敛结果。全文中，我们假设 $\mathcal{G}_n$ 的 Rademacher 复杂度$M \to \infty$ 趋于零，这在多种情形下均成立 [18, 47]。

THEOREM 4.8. *[27]* 对于任意 $n \in \mathbb{N}$，假设 $\lim_{M\to\infty} R_M(\mathcal{G}_n) = 0$。那么，以概率 *1*，有

$$\lim_{n\to\infty} \lim_{M\to\infty} \|\widehat{\alpha}(n,M) - \widehat{\alpha}(n)\|_{L^2(\Omega_{param})} = 0.$$

LEMMA 4.9. *[23]* 设 $B \in \mathbb{R}^{(N_h+1)\times(N_h+1)}$ 为一个对称正定矩阵，并分别用 $\lambda_{\min}(B)$和 $\lambda_{\max}(B)$表示其最小和最大特征值。则对于每个向量 $x \in \mathbb{R}^{N_h+1}$,

$$\lambda_{\min}(B)\,\|x\|_2^2 \;\le\; x^\top B x \;\le\; \lambda_{\max}(B)\,\|x\|_2^2, \tag{4.10}$$

其中，$\|\cdot\|_2$ 表示 $\mathbb{R}^{N_h+1}$上的欧几里得范数。

4.4. DG-FEONet 的收敛性。 结合定理 4.4 与定理4.8，我们立即得到

$$\lim_{n\to\infty} \lim_{M\to\infty} \|\alpha^* - \widehat{\alpha}(n,M)\|_{L^2(\Omega_{\text{param}})} = 0. \tag{4.11}$$

THEOREM 4.10. 设 $n \in \mathbb{N}$，并令 $h > 0$、 $\varepsilon > 0$ 为固定值。假设对于每个n，其 *Rademacher* 复杂度满足

$$\widetilde{\mathcal{G}}_n := \{\, |A\alpha - F|^2 : \alpha \in \mathcal{N}_n \,\}$$

~满足 $R_M(\mathcal{G}_n) \to 0$ 当 $M \to \infty$时，从而定理 *4.8* 中的系数空间泛化结果成立。则以概率 *1*，

$$\lim_{n\to\infty} \lim_{M\to\infty} \left\|u_{\varepsilon,h}^{\text{DG}} - \widehat{u}_{\varepsilon,h,n,M}^{\text{DG}}\right\|_{L^2(\Omega_{param};L^2(\Omega))} = 0. \tag{4.12}$$

证明。令 $\{\phi_i\}_{i=0}^{N(h)}$ 为 V_h的一组基。对于每个参数 $\omega \in \Omega_{\text{param}}$，我们可以将 DG 解及其 DG-FEONet 近似写成系数形式：

$$u_{\varepsilon,h}^{\text{DG}}(\omega,x) \;=\; \sum_{i=0}^{N(h)} \alpha_i^*(\omega)\,\phi_i(x), \qquad \widehat{u}_{\varepsilon,h,n,M}^{\text{DG}}(\omega,x) \;=\; \sum_{i=0}^{N(h)} \widehat{\alpha}_i(n,M;\omega)\,\phi_i(x).$$

定义系数差

$$\Delta\alpha(\omega) \;:=\; \alpha^*(\omega) - \widehat{\alpha}(n,M;\omega) \in \mathbb{R}^{N(h)+1},$$

以及相应的 DG 函数差

$$\Delta u_{\varepsilon,h}^{\text{DG}}(\omega,x) \;:=\; u_{\varepsilon,h}^{\text{DG}}(\omega,x) - \widehat{u}_{\varepsilon,h,n,M}^{\text{DG}}(\omega,x) \;=\; \sum_{i=0}^{N(h)} \Delta\alpha_i(\omega)\,\phi_i(x).$$

对于每个固定的 $\omega \in \Omega_{\text{param}}$，我们有

$$\begin{aligned}
\|\Delta u_{\varepsilon,h}^{\text{DG}}(\omega,\cdot)\|_{L^2(\mathcal{T}_h)}^2 &:= \sum_{K\in\mathcal{T}_h} \int_K \Big|\sum_{i=0}^{N(h)} \Delta\alpha_i(\omega)\,\phi_i(x)\Big|^2 dx \\
&= \sum_{i,j=0}^{N(h)} \Delta\alpha_i(\omega)\,\Delta\alpha_j(\omega) \sum_{K\in\mathcal{T}_h} \int_K \phi_i(x)\,\phi_j(x)\,dx \\
&= \Delta\alpha(\omega)^\top M_h\,\Delta\alpha(\omega),
\end{aligned}$$

其中，（全局）DG质量矩阵 $M_h \in \mathbb{R}^{(N(h)+1)\times(N(h)+1)}$ 定义为

$$(M_h)_{ij} := \sum_{K\in\mathcal{T}_h} \int_K \phi_i(x)\,\phi_j(x)\,dx, \qquad 0 \le i,j \le N(h).$$

由于矩阵 M_h 是对称正定的，因此其所有特征值均严格为正。记 $\lambda_{\max}(M_h)$ 为其最大特征值，由引理 4.9可得

$$\Delta\alpha(\omega)^\top M_h\,\Delta\alpha(\omega) \;\le\; \lambda_{\max}(M_h)\,\|\Delta\alpha(\omega)\|_2^2.$$

因此，存在一个常数 $C_h > 0$，其仅取决于 h 和基，使得

$$\|\Delta u_{\varepsilon,h}^{\mathrm{DG}}(\omega,\cdot)\|^2_{L^2(\mathcal{T}_h)} \;\le\; C_h\,\|\Delta\alpha(\omega)\|_2^2 \qquad \forall\,\omega\in\Omega_{\mathrm{param}}. \tag{4.13}$$

接下来，我们在参数空间 Ω_{param} 上进行积分。

$$\begin{aligned}
\|u_{\varepsilon,h}^{\mathrm{DG}} - \widehat{u}_{\varepsilon,h,n,M}^{\mathrm{DG}}\|^2_{L^2(\Omega_{\mathrm{param}};L^2(\Omega))} &= \int_{\Omega_{\mathrm{param}}} \|\Delta u_{\varepsilon,h}^{\mathrm{DG}}(\omega,\cdot)\|^2_{L^2(\mathcal{T}_h)}\,d\mathbb{P}_{\Omega_{\mathrm{param}}}(\omega)\\
&\le \int_{\Omega_{\mathrm{param}}} C_h\,\|\Delta\alpha(\omega)\|_2^2\,d\mathbb{P}_{\Omega_{\mathrm{param}}}(\omega)\\
&= C_h\,\|\alpha^* - \widehat{\alpha}(n,M)\|^2_{L^2(\Omega_{\mathrm{param}})} \to 0.
\end{aligned}$$

证明完毕。 □

5. 数值实验。 本节通过若干数值实验评估所提出的 DG-FEONet框架的性能。我们重点考察其在求解具有不连续数据（例如系数和源项）的参数化偏微分方程时的准确性和泛化能力，此时解也可能是不连续的。我们将 DG-FEONet 的性能与基于经典CG的原始 FEONet架构（称为 CG-FEONet）进行比较，突出了 DG公式在处理不连续性方面的优势。

我们的实现结合使用 PyTorch [37] 进行神经网络建模，并使用 FEniCS [1] 进行有限元组装，遵循 [27] 中概述的方法。所有实验均使用 GPU 硬件（NVIDIA A6000）加速，以提高训练和评估效率。对于以下所有示例，一维情形下我们随机生成 2,000 个参数化输入样本，其中 1,000 个用作训练数据，另 1,000 个用作测试数据。对于二维情形，我们生成 1,000 个参数化输入样本，其中训练和测试各占 500 个。正如前文所述，模型以无数据方式进行训练，仅使用弱形式的残差，无需配对的输入输出数据。训练和测试所用的参考解均通过有限元求解器在足够精细的网格上计算，作为真实解。

我们通过计算一组 N_s 个测试样本上的平均相对离散 ℓ^2误差来评估 DG-FEONet 的性能。令 $u_i^{\mathrm{FEM}}{}_{,j}$ 和 $u_i^{\mathrm{nn}}{}_{,j}$ 分别表示第 i 个测试样本（$i = 1,\dots,N_s$）在第 j 个空间自由度处的有限元解和神经网络预测，其中 $j = 1,\dots,M$。那么，相对 ℓ^2 误差定义为：

$$\|u_{i,j}^{\mathrm{nn}} - u_{i,j}^{\mathrm{FEM}}\|_{l^2}^{\mathrm{Rel}} := \left(\frac{\sum_{j=1}^{M}\left(u_{i,j}^{\mathrm{nn}} - u_{i,j}^{\mathrm{FEM}}\right)^2}{\sum_{j=1}^{M}\left(u_{i,j}^{\mathrm{FEM}}\right)^2}\right)^{1/2}, \tag{5.1}$$

TABLE 1
Summary of numerical experiments for DG-FEONet.

实验	维度	输入	输出	不连续性来源
实验1（第5.1节）	1D	$f(x) = k$ 与 $k \in [0.1, 2]$	$u(x)$	$\varepsilon(x)$
实验2（第5.2节）	1D	$c(x)$ 在随机位置发生随机跳跃	$u(x)$	$\varepsilon(x)$ $c(x)$ 和 $f(x)$
Exp3（第5.3节）	2D	$f(x, y)$，具有随机正弦/余弦系数	$u(x, y)$	$\varepsilon(x, y)$
Exp4（第5.4节）	2D	$\varepsilon(x, y)$，具有随机半径和扩散值	$u(x, y)$	$\varepsilon(x, y)$
实验5（第5.5节）	2D	$\varepsilon(x, y)$，包含两个随机圆、半径和扩散值	$u(x, y)$	$\varepsilon(x, y)$

对每个第 i 个样本，相对 ℓ^2 误差的平均值定义为：

$$E_{\text{rel}} := \frac{1}{N_s} \sum_{i=1}^{N_s} \|u_{i,j}^{\text{nn}} - u_{i,j}^{\text{FEM}}\|_{l^2}^{\text{Rel}}. \tag{5.2}$$

设置。对于数值实验1–2，我们考虑如下方程的一维情形：

$$-\varepsilon(x)\, u''(x) + b(x)\, u'(x) + c(x)\, u(x) = f(x) \text{ in } \Omega, \tag{5.3a}$$
$$u(-1) = u(1) = 0, \tag{5.3b}$$

其中 $\Omega := (-1,1)$。对于这些实验，架构采用全连接神经网络（FCNN），包含 3 个隐藏层，每层有 32 个神经元。我们使用 SiLU（Sigmoid 加权线性单元）激活函数，并采用批量大小为 32 的 L-BFGS 优化器（有限记忆 Broyden-Fletcher-Goldfarb-Shanno）。

接下来，在数值实验3–5中，我们考虑如下方程的二维情形：

$$-\nabla \cdot (\varepsilon(x, y) \nabla u(x, y)) + \mathbf{v} \cdot \nabla u(x, y) = f(x, y) \text{ in } \Omega, \tag{5.4a}$$
$$u(x, y) = 0 \text{ on } \partial\Omega, \tag{5.4b}$$

其中 $\Omega := (-1,1)^2$。在这些实验中，我们采用二维卷积神经网络（CNN），包含 5 个隐藏层，每层有 32 个神经元，并使用 SiLU 激活函数。模型以无数据方式使用 L-BFGS 优化器进行训练。表 1详细总结了输入参数化以及每个数值实验中所考虑的不连续系数类型。

5.1. 数值实验1：不连续系数下的收敛性测试。我们首先考虑一个简单的一维情形，以建立基准性能并考察收敛行为。我们评估所提出的 DG-FEONet 在系数不连续时的性能，并研究解的收敛行为。这里，我们设 $c(x) = 0.001$ 和 $b(x) = 0.01$ 如式（5.3a）所示，但我们在扩散系数中引入一个不连续点：

$$\varepsilon(x) = \begin{cases} \varepsilon_{\min}, & x < 0, \\ \varepsilon_{\max} = m\, \varepsilon_{\min}, & x \geq 0, \end{cases} \qquad \varepsilon_{\min} = 0.01, \;\; m \in \{5, 10, 100\}.$$

这种不连续性会使解在界面 $x = 0$ 处出现梯度跳跃，这给采用神经网络和算子学习进行数值近似带来了挑战 [52]。在这种情况下，我们考虑将 $f(x) := k \in [0.1, 2]$ 作为输入，其中每个函数都取自该区间内随机抽取的一个常数值。为了评估

误差 E_{rel} 的收敛性，我们在均匀网格上获得 u^{nn}，网格包含 $N_e = 16$，32、64和128个单元，其中 N_e 表示区域内网格单元的数量。

图 2 展示了一个代表性输入对应的预测解以及收敛图。这些结果证明了 DG-FEONet 在学习具有不连续系数的偏微分方程解方面的稳健性，并实现了预期的近似次优收敛率 [6]。

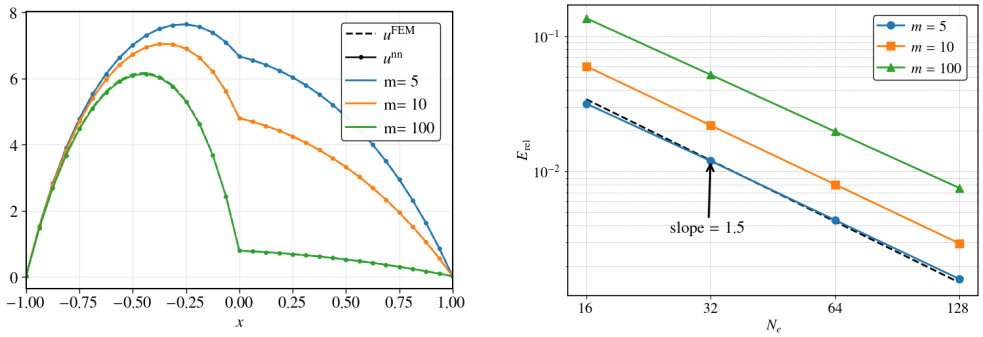


FIG. 2 *Performance of DG-FEONet for a discontinuous diffusion coefficient (Experiment 1).* 。、 *(a) Solution plots for* $m = 5$ 10 、 *and* 100. *(b) Convergence of the average relative* ℓ^2*-error* E_{rel} *for different values of* m.

此外，我们还将具有四个隐藏层的全连接架构 PINN 作为基线。该 PINN 在相同的偏微分方程（5.3a）上进行训练，其系数为b（x）$= 0.01$、c（x）$= 0.1$、f（x）$= 1.0$，以及

$$\varepsilon(x) = \begin{cases} 0.01, & x < 0, \\ 0.1, & x \geq 0. \end{cases}$$

图 3 表明，PINN 无法捕捉由不连续扩散系数引起的尖锐折角，而 DG-FEONet 与 DG 参考解相吻合。

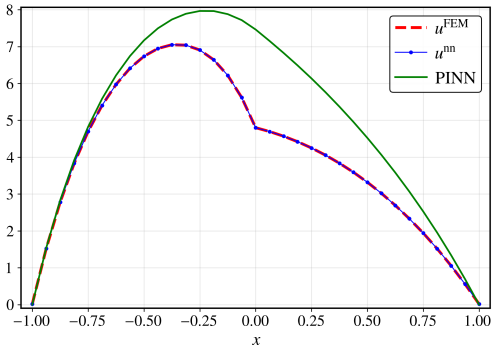


FIG. 3。 *Comparison of the DG reference solution, DG-FEONet prediction, and PINN approximation for Experiment 1 with a discontinuous diffusion coefficient* ε（x） *and jump factor* $m = 10$.

5.2. 数值实验2：随机位置间断与可变跳跃幅度。在本实验中，我们考虑存在多个

不连续性，其中跳跃幅度及其位置均随机变化。在此情况下，输入为反应系数 $c(x)$，其定义为具有两个不连续性的随机阶跃函数。其值分别从三个不同区间中随机采样：$[0,5)$、$[5,10)$ 和 $[10,15],$，分别对应左侧、中间和右侧子域。因此，

$$c(x) := \begin{cases} c_0 \in [0,5), & \text{if } x < x_0, \\ c_1 \in [5,10), & \text{if } x_0 \le x < x_1, \\ c_2 \in [10,15], & \text{if } x \ge x_1, \end{cases}$$

其中 x_0 和 x_1 也是在区域内随机选取的点（-1，1）。部分示例的设置见图 4。源项 $f(x)$ 定义为以下固定的分段常数函数：

$$f(x) = \begin{cases} 1.0, & \text{if } x < x_0, \\ -1.5, & \text{if } x_0 \le x < x_1, \\ 2.5, & \text{if } x \ge x_1. \end{cases}$$

扩散系数也不连续，定义如下：

$$\varepsilon(x,y) = \begin{cases} 0.01, & \text{if } x < x_0, \\ 0.02, & \text{if } x_0 \le x < x_1, \\ 0.03, & \text{if } x \ge x_1. \end{cases}$$

另一个系数设为 $b(x) = 0.01$。这种设置产生了多样的解剖面，对网络的泛化能力构成了严峻考验。

图4中，我们展示了输入函数 $c(x)$ 及预测解 $u^{\mathrm{nn}}(x)$ 的三个代表性样本。反应系数 $c(x)$ 呈现出明显的不连续性，相应的解 $u^{\mathrm{nn}}(x)$ 准确捕捉了这些不连续性引起的陡变。DG-FEONet 的预测与有限元解高度吻合，展示了其处理不连续系数的能力。作为比较，我们还使用相同的架构和训练数据训练了 CG-FEONet，但采用连续 Galerkin 离散。由于 CG空间是连续的，无法表示真正的跳跃，CG-FEONet 会将不连续性平滑掉，而 DG-FEONet 则保留了尖锐的跳跃，并与 DG参考解一致；见图 5。

5.3. 数值实验3：间断扩散系数下的收敛性测试。我们考虑给定偏微分方程（5.4a），研究间断扩散系数下解的收敛行为

$$\varepsilon(x,y) := \begin{cases} \varepsilon_{\min}, & \text{if } \sqrt{x^2+y^2} < 0.5, \\ \varepsilon_{\max} = m\,\varepsilon_{\min}, & \text{otherwise}, \end{cases} \qquad \varepsilon_{\min} = 0.1, \;\; m \in \{5,10,20\}.$$

不连续性降低了解的正则性，并引起与界面对齐的梯度跳跃。在此，我们设 $\mathbf{v} = (-1, 0)$，输入函数定义为

$$f(x,y) = m_0 \sin(n_0 x + n_1 y) + m_1 \cos(n_2 x + n_3 y),$$

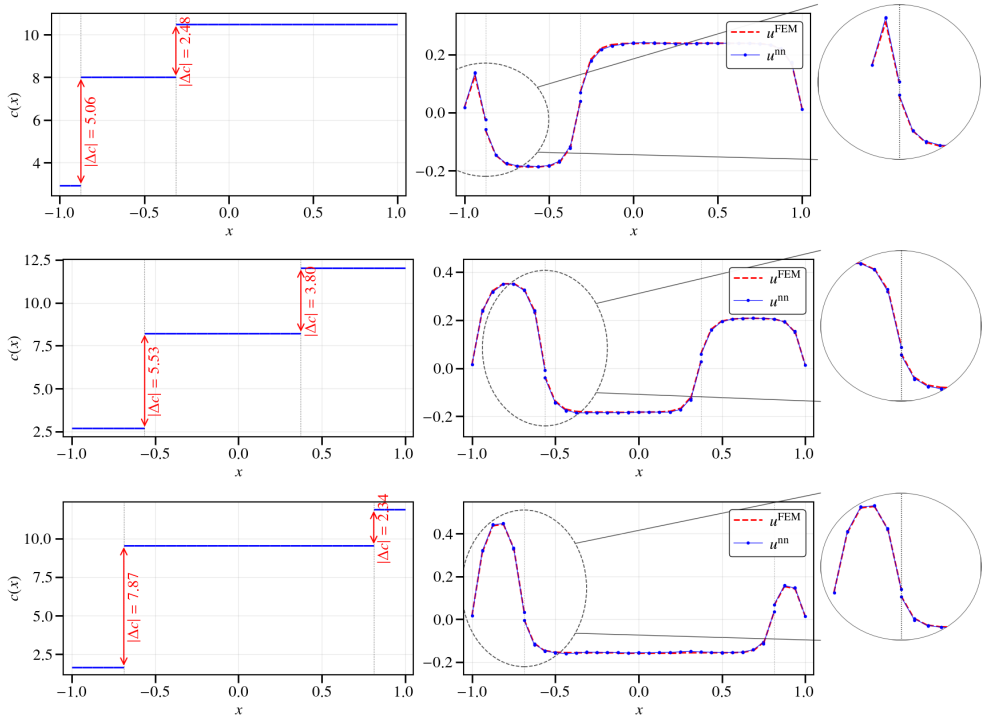


FIG. 4.

Three representative examples showing the input reaction coefficient $c(x)$ with discontinuities (left) and the DG-FEONet solution u^{nn} vs. FEM solution u^{FEM} (right) (Experiment 2). Annotated arrows indicate the magnitude of the jumps $|\Delta c|$ across elements. DG-FEONet successfully resolves both the discontinuous input and the resulting non-smooth output.

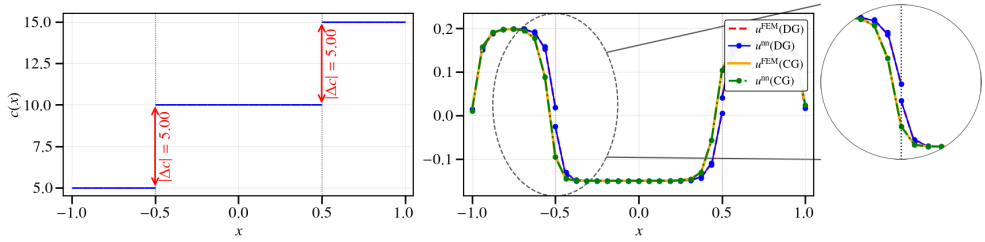
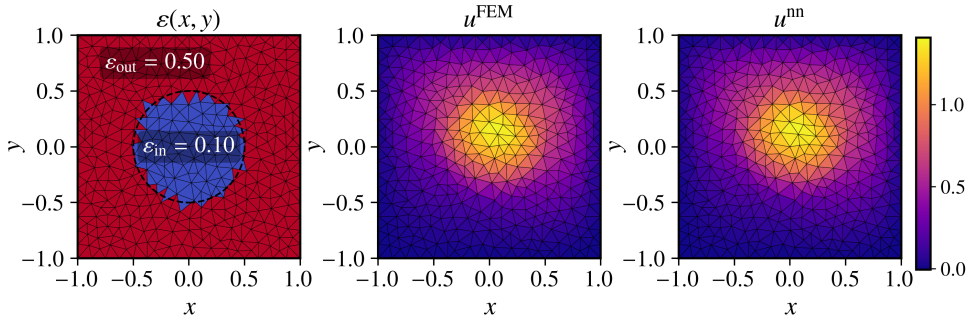


FIG. 5. *(Experiment 2). The CG-FEONet solution oversmooths the jumps, while DG-FEONet resolves them and agrees with the DG reference solution.*

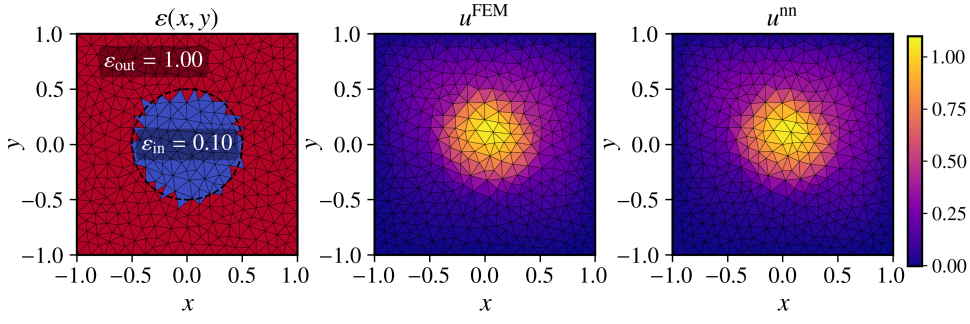
其中 m_0, $m_1 \in [1,2]$ 和 n_0, n_1, n_2, $n_3 \in [0,\pi]$ 均为随机采样。为评估平均相对 ℓ^2- 误差 E_{rel} （定义见 (5.2)）的收敛性，我们使用三角形网格对区域进行离散化，网格包含 $N_e = 32$, 128、512 和 2048 个单元（对应的自由度数分别为 96、384、1536 和 6144）。

图 6 展示了每个 $m \in \{5, 10, 20\}$对应的一个代表性样本的输出。第一幅图显示了间断扩散系数 $\varepsilon(x, y)$，并突出显示了圆形界面。第二幅和第三幅图分别展示了 FEM参考解

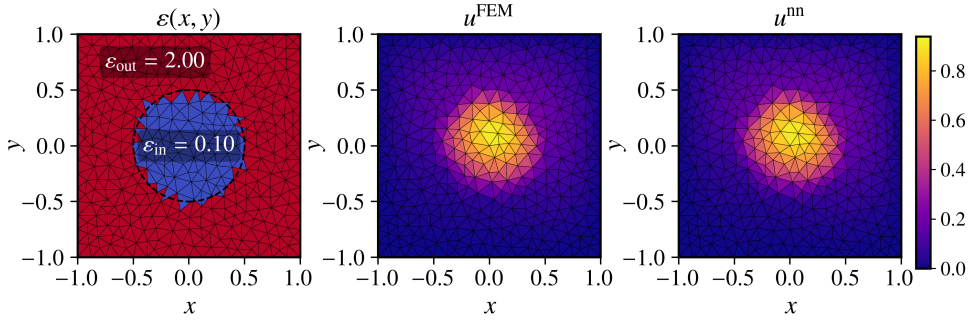
解 $u^{\mathrm{FEM}}(x,\ y)$ 和 DG-FEONet 预测 $u^{\mathrm{nn}}(x,\ y)$。这证实，DG-FEONet 能够稳定地逼近具有不连续系数的二维偏微分方程，并准确恢复不同跳跃幅度下由界面驱动的解结构。



（a）情况 1（$m = 5$）： $\varepsilon_{\mathrm{in}} = 0.10$、 $\varepsilon_{\mathrm{out}} = 0.50$。



（b）情况 2（$m = 10$）： $\varepsilon_{\mathrm{in}} = 0.10$、 $\varepsilon_{\mathrm{out}} = 1.00$。



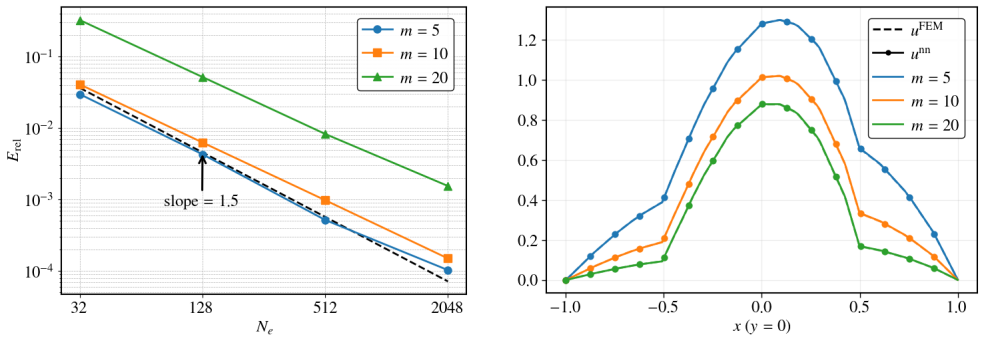
（c）情形 3（$m = 20$）： $\varepsilon_{\mathrm{in}} = 0.10$， $\varepsilon_{\mathrm{out}} = 2.00$。

FIG. 6 *Element-wise plots for three diffusion jumps* ε *x y (left) FEM reference solution* u^{FEM} 。 （， ） *(middle)*

， ， *and DG-FEONet prediction* u^{nn} *(right) in Experiment 3. Panels (a)–(c) correspond to* ， ， ， $m = 5$ $m = 10$ *and* $m = 20$ *respectively.*

图7 （a）展示了不连续系数条件下网格细化时的预期收敛行为。图中表明，该方法实现了稳定且接近最优的收敛率，与参考斜率高度吻合。图7 （b）展示了直线 $y = 0$上的解。我们观察到，对于这些问题，所提出的 DG-FEONet 表现出稳定且符合预期的收敛性能

具有二维连续/不连续系数。



(a) (b)

Fig. 7. *Convergence and solution profile for a discontinuous diffusion coefficient (Experiment 3). (a) Convergence of the average relative ℓ^2-error E_{rel} for different values of m. (b) Line-slice comparison of u^{nn} and u^{FEM} along $y = 0$ for $m = 5$ 10 and 20.*

5.4. 数值实验 4：半径和扩散值变化的随机圆形不连续界面。在本例中，我们通过引入半径和扩散值的变化，对前述设置进行扩展。输入扩散系数 $\varepsilon(x, y)$ 定义为

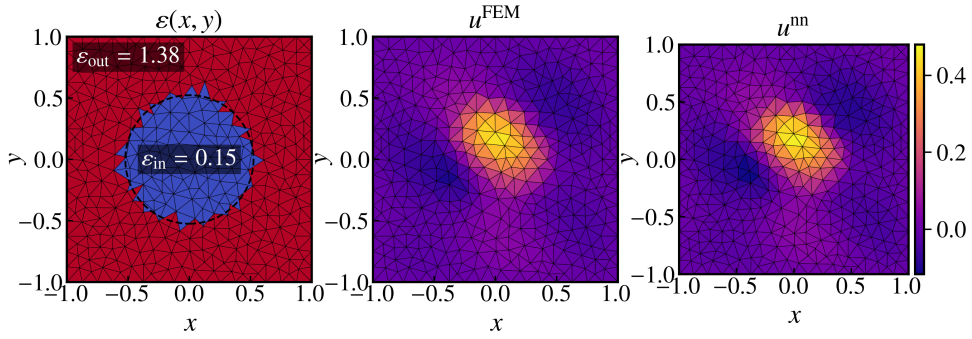
$$\varepsilon(x,y) = \begin{cases} \varepsilon_{\mathrm{in}}, & \text{if } \sqrt{x^2+y^2} < r, \\ \varepsilon_{\mathrm{out}}, & \text{otherwise}, \end{cases}$$

其中，$r \in (0, 1]$ 表示半径，$\varepsilon_{\mathrm{in}} \in [0.1, 0.5]$和 $\varepsilon_{\mathrm{out}} \in [1.0, 2.0]$ 在每次实现中独立采样。此设置生成了具有随机圆形界面和跳跃幅度的不连续扩散系数，为学习算子提供了具有挑战性的泛化测试。此处，源函数固定为$f(x, y) = 1.5\sin(\pi(x + 2y)) + 1.2\cos(\pi(1.5x + 0.5y))$。

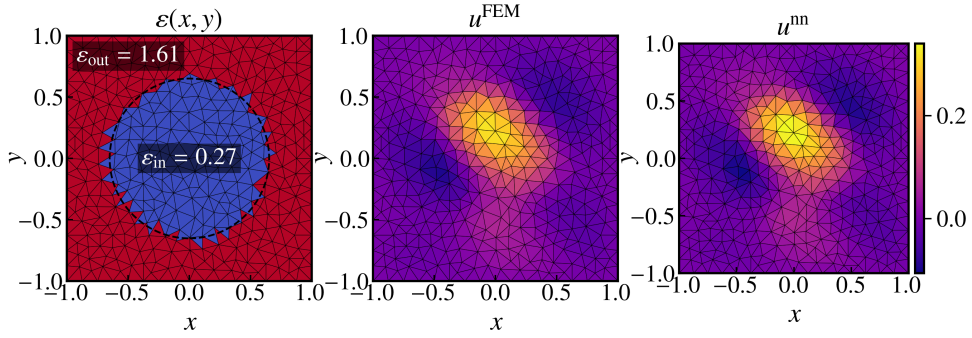
随机生成的扩散场示例如图8所示。图8 展示了两个具有代表性的测试用例的模型预测结果。我们展示了间断扩散系数 $\varepsilon(x, y)$、有限元解 $u^{\mathrm{FEM}}(x, y)$以及 DG-FEONet 预测 $u^{\mathrm{nn}}(x, y)$。结果凸显了模型捕捉界面诱导变化并在整个区域内保持较高预测准确率的能力。

为进一步研究半径和跳跃幅度对预测准确率的影响，我们进行了两项实验。首先，将半径固定为 $r = 0.5$，并通过随机采样 $\varepsilon_{\mathrm{in}} \in [0.1, 1.0]$来改变扩散对比度，同时固定 $\varepsilon_{\mathrm{out}} = 1.0$。这将扩散跳跃幅度对模型准确率的影响单独分离出来。图9 (a) 总结了实验结果，显示相对 ℓ^2 误差与跳跃幅度 $|\varepsilon_{\mathrm{out}} - \varepsilon_{\mathrm{in}}|$之间呈明显的上升趋势。扩散跳跃越大，预测误差越高，这符合预期，因为界面两侧的解不连续性更强。

接下来，我们固定 $\varepsilon_{\mathrm{in}} = 0.1$ 和 $\varepsilon_{\mathrm{out}} = 1.0$，改变夹杂半径 $r \in [0.1, 1.0]$，以考察几何形状（半径）如何影响预测准确率。图 9 （b）展示了相对 ℓ^2 误差与半径之间的关系。尽管回归线呈现出略微上升的趋势，但散点分布和较宽的置信带表明，相关性较弱且在统计上不显著。总体而言，在此范围内，DG-FEONet 对夹杂半径的变化基本不敏感。



(a) $r = 0.52$, $\varepsilon_{\mathrm{in}} = 0.15$, $\varepsilon_{\mathrm{out}} = 1.38$。



(b) $r = 0.65$, $\varepsilon_{\mathrm{in}} = 0.27$, $\varepsilon_{\mathrm{out}} = 1.61$。

FIG. 8. *Visualization of representative diffusion realizations with random circular interfaces (Experiment 4). Each row shows (left) the discontinuous diffusion coefficient ε x y (middle), the FEM reference solution u^{FEM}, and (right) the DG-FEONet prediction u^{nn} for the indicated $(r, \varepsilon_{\mathrm{in}}, \varepsilon_{\mathrm{out}})$ values.*

5.5. 数值实验 5：具有不同半径和扩散值的两个随机非重叠圆形不连续区域。 此示例引入两个互不相交的圆形夹杂，其中心、半径和扩散值均独立随机生成。中心 (x_1, y_1) 和 (x_2, y_2) 在区域 $\Omega = [-1,1]^2$ 内随机采样，并满足圆形区域互不重叠且完全位于区域内部的约束。随后，输入函数 $\varepsilon(x, y)$ 定义为

$$\varepsilon(x,y) = \begin{cases} \varepsilon_1, & \text{if } |(x,y)-(x_1,y_1)| < r_1, \\ \varepsilon_2, & \text{if } |(x,y)-(x_2,y_2)| < r_2 \text{ and not in circle 1}, \\ \varepsilon_{\mathrm{out}}, & \text{otherwise}, \end{cases} \tag{5.5}$$

其中，半径 r_1、$r_2 \in [0.2, 0.4]$、内部扩散值 ε_1、$\varepsilon_2 \in [0.1, 0.5]$ 以及背景扩散值 $\varepsilon_{\mathrm{out}} \in [1.0, 2.0]$ 均分别从各自区间中独立随机抽取。在此，我们设置 $\mathbf{v} = (-1, 0)$，并将源项固定为 $f(x,y) = 1.5\sin(\pi(x+2y)) + 1.2\cos(\pi(1.5x+0.5y))$。图 10 展示了一个具有代表性的测试样本，其中包括间断扩散系数 $\varepsilon(x,y)$、有限元参考解 $u^{\mathrm{FEM}}(x,y)$ 以及 DG-FEONet 预测 $u^{\mathrm{nn}}(x,y)$。所展示的设置包含两个半径分别为 $r_1 = 0.169$ 和 $r_2 = 0.397$ 的圆形夹杂，以及扩散系数 $\varepsilon_1 = 0.48$、$\varepsilon_2 = 0.30$ 和 $\varepsilon_{\mathrm{out}} = 1.09$。

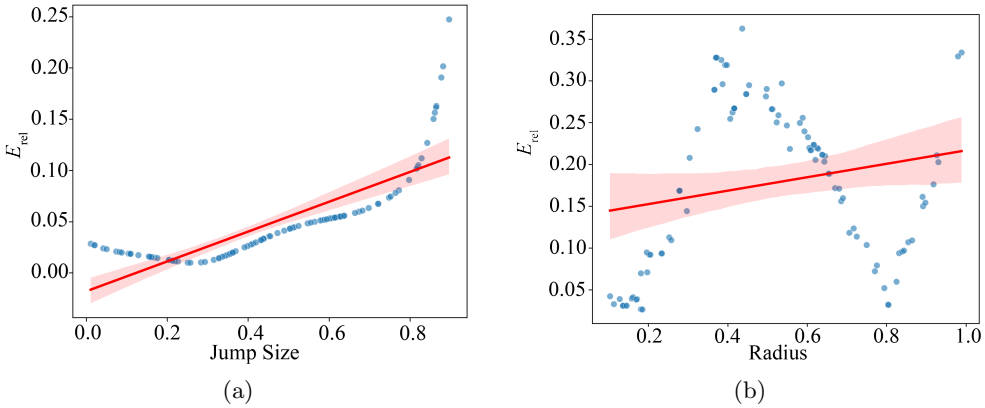


FIG. 9 *Sensitivity of DG-FEONet to diffusion jump size and inclusion radius (Experiment 4). . (a) Error versus diffusion jump size* $|\varepsilon_{\text{out}} - \varepsilon_{\text{in}}|$. *(b) Error versus inclusion radius* r. *The red solid line denotes the least-squares regression trend and the shaded band indicates the* 95 *confidence ,* % *interval for the fitted mean error.*

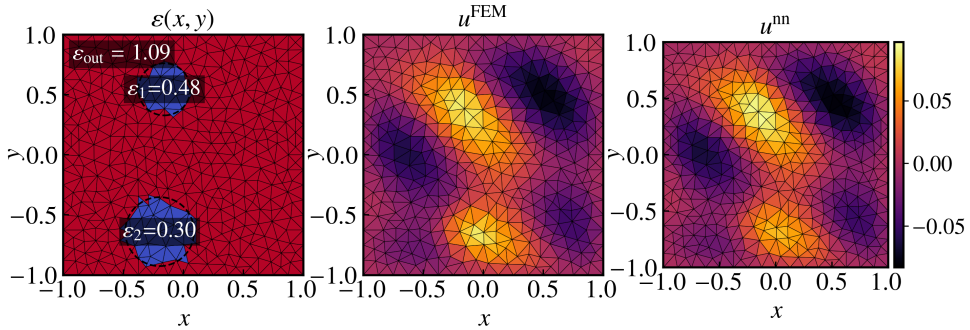


FIG. 10. *Representative test case with two circular inclusions of radii* $r_1 = 0.169$ *and* $r_2 = 0.397$, *and diffusion coefficients* $\varepsilon_1 = 0.48$ $\varepsilon_2 = 0.30$ *and* $\varepsilon_{\text{out}} = 1.09$ *(Experiment 5). From left to , , right: discontinuous diffusion field* ε x y *FEM solution* u^{FEM} x y *and DG-FEONet prediction* (,) (,) (, y) u^{nn} x .

图 11 展示了 u^{FEM}（x、y）与 u^{nn}（x、y）的三维曲面对比，表明二者在振幅和空间结构方面均高度一致；对于这一复杂设置下的 DG-FEONet 预测，仅观察到轻微的局部平滑。

6. 结论。本文提出了 DG-FEONet，这是一种基于间断伽辽金方法、用于求解参数化偏微分方程的物理信息神经算子框架。通过将间断伽辽金离散的弱形式嵌入损失函数，DG-FEONet 无需成对输入输出训练数据，即可直接从控制方程中进行无数据学习。

我们评估了 DG-FEONet 在一系列涉及光滑和不连续系数的一维与二维对流–扩散问题上的性能。数值结果表明，即使界面几何形状、位置和幅值存在显著变化，DG-FEONet 仍能准确捕捉由系数跳跃引起的解的尖锐特征。该模型表现出良好的收敛

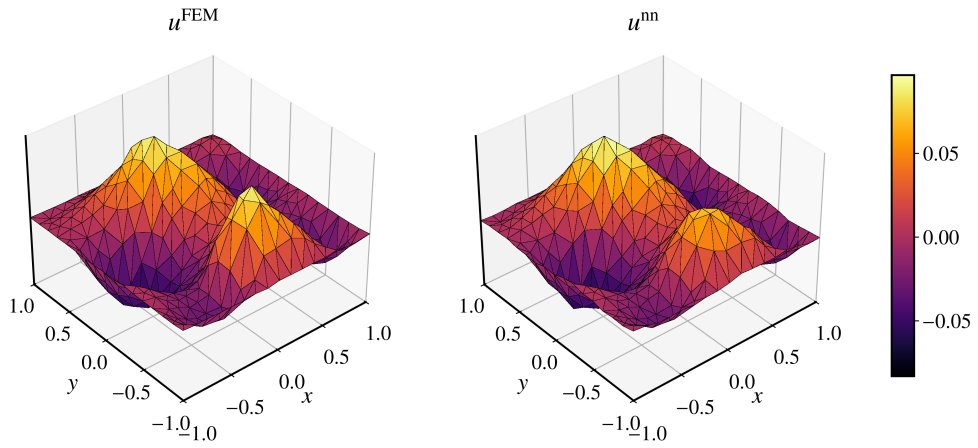


FIG. 11。 *Three-dimensional surface comparison for Experiment 5. Left: FEM reference solution* $u^{\mathrm{FEM}}(x,\ y)$. *Right: DG-FEONet prediction* $u^{\mathrm{nn}}(x,\ y)$。

行为和泛化能力，尤其是在具有多个不连续性的复杂情形中。

我们的实验进一步证实，基于 DG 的架构在处理非光滑问题时尤其有效，而传统神经求解器和基于配置点的物理信息驱动方法往往难以应对。DG-FEONet 利用间断伽辽金方法的局部性和灵活性，为在不规则且非均质环境中进行偏微分方程学习提供了一种可扩展且稳健的方法。

附录 A. DG 算子矩阵的可逆性。

LEMMAA.1（DG 算子矩阵的可逆性）。假设 *DG* 双线性型 $a(\cdot,\cdot)$ 在 *DG* 能量范数下是强制的，即存在 $C_{\mathrm{coer}} > 0$使得

$$a(v_h, v_h) \geq C_{\mathrm{coer}} \|v_h\|_{DG}^2 \quad \forall v_h \in V_h.$$

令 $A \in \mathbb{R}^{(N_h+1)\times(N_h+1)}$ 定义为

$$A_{ij} := a(\phi_j, \phi_i).$$

则 A 是非奇异的。

证明。 设 $\alpha = (\alpha_0, \ldots, \alpha_{N_h})^\top \in \mathbb{R}^{N_h+1}$ 满足

$$A\alpha = 0.$$

定义相应的 DG 函数 $u_h \in V_h$ 为

$$u_h := \sum_{j=0}^{N_h} \alpha_j\, \phi_j.$$

根据 4.4 对 A 的定义， i 的第 $A\alpha$ 个分量为

$$(A\alpha)_i = \sum_{j=0}^{N_h} A_{ij}\, \alpha_j = \sum_{j=0}^{N_h} a(\phi_j, \phi_i)\, \alpha_j = a\Big(\sum_{j=0}^{N_h} \alpha_j \phi_j,\ \phi_i\Big) = a(u_h, \phi_i),$$

对所有 $i = 0, \ldots, N_h$。因此 $A\alpha = 0$ 意味着

$$a(u_h, \phi_i) = 0 \qquad \forall\, i = 0, \ldots, N_h.$$

由于 $\{\phi_i\}_{i=0}^{N_h}$ 是 V_h 的一组基，任意 $v_h \in V_h$ 都可以写成

$$v_h = \sum_{i=0}^{N_h} \beta_i \phi_i.$$

由 $a(\cdot,\cdot)$ 的双线性性可知

$$a(u_h, v_h) = \sum_{i=0}^{N_h} \beta_i\, a(u_h, \phi_i) = 0 \qquad \forall\, v_h \in V_h.$$

特别地，取 $v_h = u_h$ 可得

$$a(u_h, u_h) = 0.$$

利用强制性假设（4.3），我们得到

$$0 = a(u_h, u_h) \;\geq\; C_{\mathrm{coer}}\, \|u_h\|_{DG}^2.$$

由于 $C_{\mathrm{coer}} > 0$，因此可得

$$\|u_h\|_{DG}^2 = 0,$$

因此，$u_h \equiv 0$ 属于 Ω，因为 $\|\cdot\|_{DG}$ 是 V_h上的一个范数。

最后，$u_h \equiv 0$ 以及展开式 $u_h = \sum_{j=0}^{N_h} \alpha_j \phi_j$，结合 $\{\phi_j\}_{j=0}^{N_h}$ 的线性无关性，可得

$$\alpha_j = 0 \qquad \forall\, j = 0, \ldots, N_h,$$

即，$\alpha = 0$。因此，A 的核中只有零向量，所以 A 是非奇异的，因而可逆。

致谢。 K. Chawla 和 S. Lee 的工作得到美国能源部科学办公室能源地球射击计划的资助，资助编号为 DE-SC 0024703。Y. Hong 的工作得到韩国政府（MSIT）资助的韩国国家研究基金会（NRF）基础科学研究计划（RS-2023-00219980）以及韩国政府（MSIT）资助的信息与通信技术规划评估院（IITP）项目的支持 [NO.RS-2021-II211343，人工智能研究生院项目（首尔国立大学）]。J.Y. Lee 的工作得到韩国政府（MSIT）资助的信息与通信技术规划评估院（IITP）项目（ RS-2021-II211341，人工智能研究生院项目（中央大学）） [以及韩国政府（ MSIT）资助的韩国国家研究基金会（NRF）项目（RS-2025-02303239）的支持] 。

参考文献。

[1] M. Alnæs, J. Blechta, J. Hake, A. Johansson, B. Kehlet, A. Logg, C. Richardson, J. Ring, M. E. Rognes, and G. N. Wells,*FEniCS*项目 *1.5* 版，《数值软件档案》，3（2015年），第 100 页。

[2] A. Arzani, K. W. Cassel, and R. M. D'Souza, 面向具有奇异摄动的边界层问题的理论引导型物理信息神经网络，《计算物理学杂志》，*473*（*2023*年），第 *111768* 页。[3] Y. Bar-Sinai, S. Hoyer, J. Hickey, and M. P. Brenner, 学习偏微分方程的数据驱动离散化，《美国国家科学院院刊》，*116*（*2019*年），第 *15344–15349* 页。[4] J. Brandstetter, D. E. Worrall, and M. Welling, 消息传递神经偏微分方程求解器，载于《国际学习表征会议》，*2022*年。
[5] S. C. Brenner and L. R. Scott,《有限元方法的数学理论》，施普林格，*2008*年。[6] Z. Cai, X. Yue, and S. Zhang, 用于界面问题的间断伽辽金有限元方法：先验与后验误差估计，《*SIAM* 数值分析杂志》，*49*（*2011*年），第 *1761–1787* 页。[7] S. Cao, 选择 *Transformer*：傅里叶还是伽辽金，载于《神经信息处理系统进展》，*2021*年。[8] C. Chen, Y. Yang, Y. Xiang, and W. Hao, 通过同伦动力学学习奇异摄动解，载于第四十二届国际机器学习会议论文集，*2025*年。
[9] G. Chen, S. Xu, D. Ni, and T. Zeng, *DGNN*：由间断伽辽金方法诱导的神经偏微分方程求解器，*2025*年。[10] J. Choi, N. Kim, and Y. Hong, 用于求解偏微分方程的无监督勒让德–伽辽金神经网络，《*IEEE Access*》，*11*（*2023*年），第 *23433–23446* 页。[11] J. Choi, T. Yun, N. Kim, and Y. Hong, 无需依赖数据的参数化偏微分方程谱算子学习，《计算机在应用力学与工程中的方法》，*420*（*2024*年），第 *116678* 页。[12] B. Cockburn, G. E. Karniadakis, and C.-W. Shu,《间断伽辽金方法：理论、计算与应用》，第 *11* 卷，施普林格科学与商业媒体，*2012*年。[13] B. Cockburn and C.-W. Shu, 守恒律的龙格–库塔间断伽辽金方法 *V*：多维系统，《计算物理学杂志》，*141*（*1998*年），第 *199–224* 页。[14] V. A. Dobrev, R. D. Lazarov, P. S. Vassilevski, and L. T. Zikatanov, 二阶椭圆方程间断伽辽金近似的两层预条件，《数值线性代数及其应用》，*13*（*2006*年），第 *753–770* 页。
[15] Y. Epshteyn and B. Rivière, 对称内罚伽辽金方法的惩罚参数估计，《计算与应用数学杂志》，*206*（*2007*年），第 *843–872* 页。[16] L. C. Evans,《偏微分方程》，第 *19* 卷，美国数学学会，*2022*年。[17] E. Franck, V. Michel-Dansac, and L. Navoret, 使用经物理信息神经网络增强的基函数构造近似良平衡间断伽辽金方法，《计算物理学杂志》，*512*（*2024*年），第 *113144* 页。[18] G. Gnecco and M. Sanguineti, 基于拉德马赫复杂度的逼近误差界，《应用数学科学》（俄罗斯），*2*（*2008*年），第 *153–176* 页。
[19] S. Goswami, A. Bora, Y. Yu, and G. E. Karniadakis,《物理信息深度神经算子网络》，施普林格国际出版，沙姆，*2023*年，第 *219–254* 页，*https://doi.org/10.1007/978-3-031-36644-4 6*。

[20] G. Gupta, X. Xiao, and P. Bogdan，基于多小波的微分方程算子学习，载于《神经信息处理系统进展》，*2021*年。[21] Z. Hao, C. Su, S. Liu, J. Berner, C. Ying, H. Su, A. Anandkumar,J. Song, and J. Zhu，*DPOT*：用于大规模偏微分方程预训练的自回归去噪算子变换器，载于《第*41*届国际机器学习会议论文集》，《机器学习研究论文集》第*235*卷，*PMLR*，*2024*年，第*17616–17635*页。[22] Y. Hong, S. Ko, and J. Lee，用于求解参数化二阶椭圆偏微分方程的有限元算子学习方法误差分析，*2024*年。[23] R. A. Horn and C. R. Johnson，《矩阵分析》，剑桥大学出版社，第*2*版，*2012*年。[24] T. J. R. Hughes，《有限元方法：线性静力与动力有限元分析》，多佛出版社，*2000*年。[25] O. Karakashian and C. Collins，双调和方程间断伽辽金近似的两层加性施瓦茨方法，《科学计算杂志》，*74*（*2018*年），第*573–604*页。[26] N. Kovachki, Z. Li, B. Liu, K. Azizzadenesheli, K. Bhattacharya, A. Stuart, and A. Anandkumar，神经算子：学习函数空间之间的映射，*2021*年。[27] J. Y. Lee, S. Ko, and Y. Hong，用于求解椭圆型参数化偏微分方程的有限元算子网络，《*SIAM*科学计算杂志》，*47*（*2025*年），第*C501–C528*页。[28] Y. Li, T. Du, Y. Pang, and Z. Huang，用于奇异摄动微分方程的分量傅里叶神经算子，载于《*AAAI*人工智能大会论文集》，第*38*卷，*2024*年，第*13691–13699*页。[29] Z. Li, N. B. Kovachki, K. Azizzadenesheli, B. liu, K. Bhattacharya,A. Stuart, and A. Anandkumar，用于参数化偏微分方程的傅里叶神经算子，载于《国际学习表征会议论文集》，*2021*年，*https://openreview.net/forum?id=c8P9NQVtmnO*。[30] Z. Li, H. Zheng, N. Kovachki, D. Jin, H. Chen, B. Liu, K. Azizzade-nesheli, and A. Anandkumar，用于学习偏微分方程的物理信息神经算子，《*ACM/IMS*数据科学杂志》，*1*（*2024*年）。[31] L. Liu, S. Liu, H. Xie, F. Xiong, T. Yu, M. Xiao, L. Liu, and H. Yong，使用物理信息神经网络进行间断计算，《科学计算杂志》，*98*（*2024*年），第*22*页。[32] L. Lu, P. Jin, G. Pang, Z. Zhang, and G. E. 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*（*2021*年），第*218–229*页。[33] L. Lu, X. Meng, S. Cai, Z. Mao, S. Goswami, Z. Zhang, and G. E.Karniadakis，基于*FAIR*数据对两种神经算子进行全面且公平的比较（含实际扩展），《计算机在应用力学与工程中的方法》，*393*（*2022*年），第*114778*页。[34] Z. Mao and X. Meng，采用基于残差/梯度的自适应采样方法求解具有尖锐解的偏微分方程的物理信息神经网络，《应用数学与力学》，*44*（*2023*年），第*1069–1084*页。[35] F. Müller, D. Schötzau, and C. Schwab，多边形椭圆问题的对称内罚间断 *Galerkin* 方法，《*SIAM* 数值分析杂志》

[36] W. Ouyang, Y. Shin, S.-W. Liu, and L. Lu, 神经算子单元法：借助可复用神经算子实现高效且可扩展的有限元法，*2025*年。[37] A. Paszke, *PyTorch*：一种命令式、高性能深度学习库，*2019*年。[38] T. Pfaff, M. Fortunato, A. Sanchez-Gonzalez, and P. Battaglia, 使用图网络学习基于网格的模拟，载于《国际学习表征会议论文集》，*2021*年。[39] M. Raissi, P. Perdikaris, and G. E. Karniadakis, 物理信息神经网络：用于求解涉及非线性偏微分方程的正问题与逆问题的深度学习框架，《计算物理学杂志》，*378*（*2019*年），第*686–707*页。[40] B. Rivière, 求解椭圆型与抛物型方程的间断伽辽金方法：理论与实现，工业与应用数学学会，*2008*年。[41] B. Rivière, M. F. Wheeler, and V. Girault, 椭圆问题的内部罚项、约束方法和间断伽辽金方法的改进能量估计。第一部分，《计算地球科学》，*3*（*1999*年），第*337–360*页。[42] B. Rivière, 求解椭圆型与抛物型方程的间断伽辽金方法，应用数学前沿丛书第 *35* 卷，工业与应用数学学会，费城，宾夕法尼亚州，*2008*年，*https://doi.org/10.1137/1.9780898717440*。[43] Y. Shen, J. Chen, and K. Wu, 一种用于求解双曲守恒律的混合间断伽辽金神经网络方法及其时间渐进式学习，*2025*年。[44] C.-W. Shu, 间断伽辽金方法：一般方法与稳定性，载于《偏微分方程的数值解》，*Birkh¨auser*，*2009*年，第*149–201*页。[45] S. Sun and M. F. Wheeler, 耦合流动与反应性输运问题的间断伽辽金方法，《应用数值数学》，*52*（*2005*年），第*273–298*页。[46] L. N. M. Tawfiq and K. M. M. Al-Abrahemee, 用于求解奇异摄动问题的神经网络设计，《应用与计算数学杂志》，*3*（*2014*年），第*1–5*页。[47] M. J. Wainwright, 高维统计学，剑桥统计与概率数学丛书第*48*卷，剑桥大学出版社，*2019*年。[48] S. Wang, J. H. Seidman, S. Sankaran, H. Wang, G. J. Pappas, andP. Perdikaris, *CViT*：用于算子学习的连续视觉变换器，载于《第十三届国际学习表征会议论文集》，*2025*年。[49] S. Wang, H. Wang, and P. Perdikaris, 使用物理信息 *DeepONet* 学习参数化偏微分方程的解算子，《科学进展》，*7*（*2021*年），第 *eabi8605* 页。[50] X. Wang, L. Yuan, and Y. Yu, 用于一维偏微分方程的间断混合神经网络，*2025*年。[51] M. F. Wheeler, 一种带内部罚项的椭圆配置有限元法，《*SIAM* 数值分析杂志》，*15*（*1978*年），第*152–161*页。[52] S. Wu and B. Lu, *INN*：作为一种易于使用的无网格方法求解界面偏微分方程问题的界面神经网络，《计算物理学杂志》，*470*（*2022*年），第*111588*页。[53] O. C. Zienkiewicz and R. L. Taylor, 有限元方法：基础与原理，*Butterworth–Heinemann*，*2000*年。