LSA-PINN：用于求解复杂几何形状上偏微分方程的线性边界连通性损失

Jian ChengWong*高性能计算研究院新加坡科技研究局新加坡
wongj@ihpc.a-star.edu.sg

Pao-Hsiung Chiu*高性能计算研究院新加坡科技研究局新加坡
chiuph@ihpc.a-star.edu.sg

Chinchun Ooi高性能计算研究院新加坡科技研究局新加坡
ooicc@ihpc.a-star.edu.sg

My HaDao高性能计算研究院新加坡科技研究局新加坡
daomh@ihpc.a-star.edu.sg

Yew-Soon Ong计算机科学与工程学院南洋理工大学新加坡
asysong@ntu.edu.sg

arXiv:2302.01518v2 [cs.LG] 2 Mar 2023

摘要—**我们提出了一种新颖的损失函数形式，利用物理信息神经网络（PINNs）从控制物理方程（通常由偏微分方程（PDEs）描述）中高效学习复杂动力学。在我们的实验中发现，现有版本的物理信息神经网络在许多问题上的学习效果较差，尤其是在复杂几何形状下，因为在近边界区域建立适当的采样策略变得越来越困难。如果局部梯度行为过于复杂，难以由物理信息神经网络充分建模，过于密集的采样可能会对训练收敛产生不利影响。另一方面，如果样本过于稀疏，现有的物理信息神经网络往往会对近边界区域过拟合，从而导致错误的解。为避免这些问题，我们提出了一种新的边界连通性（BCXN）损失函数，该函数为边界处的梯度行为提供线性局部结构近似（LSA）。我们的*BCXN*损失在训练过程中隐式施加局部结构，从而以稀疏一个数量级的训练样本，实现跨整个问题区域的快速物理信息学习。这种LSA-PINN方法在标准L2范数指标下的误差比现有方法小几个数量级，同时所需的训练样本和迭代次数大幅减少。我们提出的LSA-PINN不对神经网络的可微性提出任何要求，并且我们在当前PINN文献中常用的多层感知机和卷积神经网络版本上展示了其优势及易于实现的特点。**

索引词—**物理信息神经网络、局部结构近似、边界连通性、偏微分方程**

I. 引言

物理信息神经网络（PINN）已成为一种颇具前景的方法，可利用控制物理方程 [1] 学习动力系统的解。近年来，PINN 已应用于科学和工程领域的多种物理现象及应用研究，包括电磁学、流体动力学、热传递等 [2], [3]。PINN 的显著特点是将支配物理定律（通常以偏微分方程（PDE）的形式表示）作为学习目标。这种融入物理信息的学习——

*：贡献相同

约束可防止 PINN 在从问题域采样的所有训练点上违反底层物理规律。

现有的物理信息神经网络通过自动微分（AD）或数值微分（ND）类方法，在训练损失中评估偏微分方程约束。自动微分损失 已成为一种流行选择，因为它可以直接根据输入和网络权重，在任意采样位置精确计算导数项。两种方法各有优缺点，但数值微分损失 可以灵活地应用于多种不同的神经网络架构，包括多层感知机（MLP）和卷积神经网络（CNN），因为与自动微分不同，它不要求神经网络具备可微性。近期研究 [4]–[6]表明，数值微分方法，尤其是耦合自动微分-数值微分（*CAN*）损失 [6]，能够以更少的训练样本，更稳健、高效地生成精确解；而传统的自动微分损失 在训练过程中容易失效。这是因为数值微分方法利用相邻采样点的物理信息神经网络输出近似高阶导数，因此能够通过这些局部近似有效地将稀疏采样点连接成分段区域，从而以更稀疏的采样点促进整个区域内快速的物理信息学习。

处理不规则几何时，现有物理信息神经网络很难将区域内部的训练样本与边界完美连接起来。若无法做到这一点，物理信息神经网络在近边界区域开始过拟合时，可能会导致不理想的训练失败。由于许多具有实际意义的偏微分方程都是边值问题，因此理想情况下，物理信息神经网络应能模拟正确的边界行为。增加密集样本以更精细地划分边界附近的分段局部区域，或许能提高准确度，但采样密度需要增加到何种程度取决于经验，而且更密集的采样策略可能会妨碍训练收敛。此外，对于采用卷积神经网络架构的物理信息神经网络，可能无法在局部提高近边界区域的采样密度。

因此，本文提出了一种损失函数形式，有助于在边界处对局部

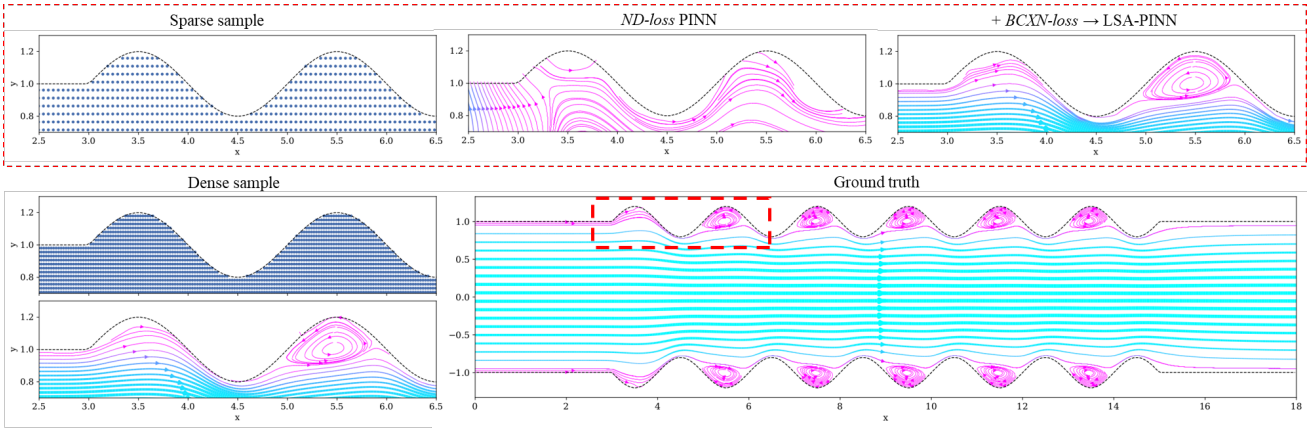


图 1. 物理信息神经网络在复杂几何形状中学习二维纳维–斯托克斯方程的解（波浪形通道流动问题，雷诺数 $=100$）。我们的 LSA-PINN 方法能够以更少的训练样本（减少 50%）和更快的速度（训练迭代次数减少 50%）学习到精确解。

梯度行为进行线性近似，从而恢复区域边界与边界附近内部样本之间的连通性。这种 局部结构近似 还可在物理信息神经网络训练过程中提供一定的结构偏置，以防止边界附近出现不准确的表示。这种新的边界连通性（*BCXN*）损失函数是新型 LSA-PINN方法的关键，使其能够利用更稀疏的训练样本更高效地学习偏微分方程的解，而不受区域几何形状的影响；示例见图 1。此外，该方法还可与其他物理信息神经网络优化方面的进展联合使用，例如损失平衡、区域分解和自适应采样。

在本文其余部分，我们将通过全面的实验展示：i）LSA-PINN 方法可灵活应用于多层感知机和卷积神经网络架构；ii）LSA-PINN 在学习多个复杂流体动力系统（包括用于元建模的系统）方面的有效性。与采用自动微分-和数值微分损失的传统物理信息神经网络相比，采用线性 *BCXN* 损失的 LSA-PINN 能够应对具有挑战性的偏微分方程问题，同时使用更稀疏的训练样本，从而进一步拓展物理信息神经网络学习复杂现实世界动力学的潜力。

II. 相关工作

物理信息神经网络中的高效采样。 先前研究 [7]，[8] 给出了物理信息损失学习在训练样本方面的理论极限。为提升物理信息神经网络在实际应用中的训练速度 [9]，，多项研究聚焦于高效采样策略，例如重要性采样、自适应采样和顺序采样，以减少物理信息神经网络训练所需的训练样本数量 [10]–[18]。此外，研究人员还探索了区域分解和并行化策略，以加速训练 [19]–[23]。与这些研究不同，我们通过一种新提出的*BCXN* 损失，使稀疏样本情形下的物理信息学习更加稳健。

用于物理信息神经网络的卷积神经网络架构和数值微分（ND）型损失。 基于卷积神经网络的公式化方法使我们能够设计和训练更大、更强大的神经网络，因此有望扩展到更复杂的大规模偏微分方程问题 [4]，[24]–[28]。然而，为了更好地处理不规则区域，可能需要进行繁琐的坐标变换。卷积神经网络架构 PINN 通常利用 ND 方法在其输入网格上计算偏微分方程损失函数，与基于自动微分的损失相比，计算成本低得多[4]。此外，也有研究提出将基于自动微分的损失与数值微分型损失相结合的混合框架，用于多层感知机架构的物理信息神经网络，以兼具两种方法的优势 [5], [6], [19], [29]。我们的方法进一步提升了数值微分型损失在不规则几何中的效率和适用性。

在物理信息神经网络损失函数中施加边界条件约束。 物理信息神经网络必须优先学习正确的边界行为 [30], [31]。通常通过罚函数法将边界条件作为软约束纳入物理信息神经网络损失函数。在此背景下，人们探索了各种策略，以便在物理信息神经网络训练期间动态调节偏微分方程约束与边界条件约束的相对重要性 [32]–[41]。此外，也有其他方法可以绕开损失平衡问题。例如，可以设计一个试探函数，使其在构造时就精确满足边界条件；或者将边界条件约束隐式地纳入偏微分方程损失函数 [42]–[48]，，这样便只需优化偏微分方程残差这一项损失。另一种方法是使用增广拉格朗日法，将边界条件约束纳入物理信息神经网络损失函数 [49]。我们的方法通过另一种策略将边界条件隐式地纳入偏微分方程损失函数。

III. 预备知识

A. 控制物理方程——不可压缩 *Navier–Stokes*（*N-S*）方程

从根本上说，研究已表明，物理信息神经网络适用于许多不同的物理系统。本研究着重于利用物理信息神经网络学习流体动力学。我们考虑的流体问题，其控制物理方程为由质量守恒和动量守恒推导出的稳态不可压缩纳维–斯托克斯方程：

$$\nabla \cdot \vec{u} = 0 \tag{1a}$$

$$(\vec{u} \cdot \nabla)\vec{u} = Re^{-1}\Delta\vec{u} - \nabla p \tag{1b}$$

在上述偏微分方程中，原始变量 $\vec{u}$ 和 p 分别表示速度向量和压力，而雷诺数（Re）表示惯性力与黏性力之比。$\vec{u}$ 在二维情况下由两个分量（u，v）组成。尽管本文使用笛卡尔坐标，但该公式也可推广到其他坐标系。

B. 多层感知机和卷积神经网络架构 *PINN*

我们使用全连接深度神经网络架构（例如，多层感知机、卷积神经网络）来表示动力过程$\boldsymbol{U}=[\vec{u},\ p]^T$的解。对于多层感知机，输入 $\vec{x}$是空间域中的一个点坐标。对于卷积神经网络，输入$\boldsymbol{X}$ 是一个具有固定形状的张量，包含整个（离散化的）空间域。给定输入 $\vec{x}$ 时，物理信息神经网络输出$\boldsymbol{U}(\vec{x};\boldsymbol{w})$的准确度由网络参数$\boldsymbol{w}$决定，这些参数在训练期间通过优化物理信息神经网络损失函数来确定。物理信息神经网络损失函数由偏微分方程损失分量（L_{PDE}）和边界条件损失分量（L_{BC}）构成:

$$L_{PINN} = \lambda_{PDE} L_{PDE} + \lambda_{BC} L_{BC} \tag{2a}$$

$$\begin{aligned} L_{PDE} = &\left\| \nabla \cdot \vec{u}(\vec{x};\boldsymbol{w}) \right\|_{\Omega}^2 + \left\| (\vec{u}(\vec{x};\boldsymbol{w}) \cdot \nabla)\, \vec{u}(\vec{x};\boldsymbol{w}) \right. \\ &\left. - Re^{-1} \Delta \vec{u}(\vec{x};\boldsymbol{w}) + \nabla p(\vec{x};\boldsymbol{w}) \right\|_{\Omega}^2 \end{aligned} \tag{2b}$$

$$L_{BC} = \left\| B[u(\vec{x};\boldsymbol{w})] - \vec{u}(\vec{x}) \right\|_{\partial\Omega}^2 \tag{2c}$$

偏微分方程损失会惩罚物理信息神经网络输出$\boldsymbol{U}(\vec{x};\boldsymbol{w})$在流体域$\vec{x} \in \Omega$内偏离纳维–斯托克斯控制方程的情况；边界条件损失则会惩罚其偏离预期狄利克雷边界条件 $\vec{u}_{BC}(\vec{x})$的情况，该条件约束区域边界 $\vec{x} \in \partial\Omega$处的速度场。式（2）a 中的相对权重 λ_{PDE} 和 λ_{BC} 控制不同损失分量之间的权衡。

偏微分方程和边界条件损失分量是在连续域上定义的，但出于实际考虑，我们在训练期间根据有限个n 样本$D = \{(\vec{x}_i)\}_{i=1}^n$ 来计算离散化的物理信息神经网络损失函数。卷积神经网络架构 PINN 的输入天然就是等距网格，所有样本均取自该网格。对于多层感知机架构 PINN，也可以方便地基于等距网格设计训练样本，然后在每次随机梯度下降（SGD）小批量训练迭代中，随机抽取$m(<n)$个样本来计算物理信息神经网络损失函数。

尽管让物理信息神经网络在非常密集的样本集上满足偏微分方程约束是理想的，即 $\Delta\vec{x} \to 0$ 其中 $\Delta\vec{x}$ 是相邻采样点之间的距离，但即使采用最先进的随机梯度下降变体，训练此类物理信息神经网络在计算上也可能不可行。对于更复杂的物理问题，在密集样本上进行训练相应地计算成本更高，并且由于某些具有非常复杂梯度行为的局部区域（例如异常高梯度 [40], [50]–[54]）的不利影响，训练也更有可能失败。对于同一个问题，在更稀疏的样本上训练物理信息神经网络意味着更少的限制（需要满足的偏微分方程约束），并且更有可能避开存在问题的梯度点，从而加快训练速度。通过降低样本密度来放宽偏微分方程约束，使学习具有物理意义的解成为可能，尽管这可能会给解引入一些近似误差。需要注意的是，如何最好地平衡样本密度与解的精度仍然是一个有待解决的问题。像 LSA-PINN 这样能够利用稀疏训练样本持续产生更准确解的方法，对于学习复杂的现实世界动力学具有极大优势。

C. 稀疏样本下的 *PINN* 学习挑战

当相对于局部复杂度而言样本较为稀疏时，高度灵活且过度参数化的物理信息神经网络可能会“过拟合”，即使所有采样点都满足偏微分方程约束（也就是说，收敛良好），仍会导致解不正确。对于自动微分损失 物理信息神经网络，这种现象可能出现在计算域内的任何位置。相反，数值微分损失 对样本密度更加稳健，因为数值微分方法利用相邻样本的输出来近似导数项，例如使用有限差分类模板，将稀疏样本连接成分段连续区域，从而促进整个问题域上的快速训练。对导数项采用较低阶的局部结构近似也为物理信息神经网络提供了一定的结构偏置，从而减少过拟合。例如，对异常高梯度进行正则化，使物理信息神经网络更容易学习解，而准确度则取决于所采用的近似方法。

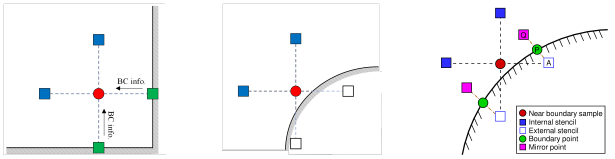
然而，如果近边界训练样本的模板未能与区域边界完美连通，数值微分损失仍可能出现训练失败（如图 2 所示）。在优化过程中，为满足近边界采样点处的偏微分方程约束，这些样本的输出实际上充当了自由参数；物理信息神经网络训练往往会“过拟合”近边界区域，以便更容易同时满足偏微分方程约束和边界条件约束，从而导致错误的解。对于不规则几何，这会成为数值微分损失的一大问题，因为几乎不可能设计出这样的采样位置，使其模板能够将区域边界与其他训练样本完美连通。

一种可能的解决方案是提高边界附近的采样密度，即 $\Delta\vec{x} \to 0$，以避免这种过拟合。但这会增加优化问题的复杂度，并可能对训练产生不利影响。对于采用卷积神经网络架构的物理信息神经网络而言，在局部细化采样密度也颇具挑战。为在不增加样本密度的情况下解决这一问题，我们提出 BCXN 损失，通过施加局部结构近似来约束近边界梯度行为，并恢复区域边界与近边界采样点之间的连通性。

IV. 方法

A. 在近边界采样点施加线性约束

本节概述所提出的 LSA-PINN方法。这里，我们将近边界样本的外部（区域外）模板称为 ES 点。如图 2c 所示，边界条件 $\vec{u}_{BC}$ 和场值 $\vec{u}_{MI}$ 分别定义在边界点 P 和区域内选定的镜像点 Q 处，而 ES 点 A 处的场值 $\vec{u}_{ES}$ 则是待确定的值。



(a) 模板 点 连接到边界　(b) 模板 点 落在区域之外　(c) 外推 沿法向

图 2。(a)-(b) 近边界训练样本（红色圆圈）及其模板点（方形）的示意图，这些点通常用于计算 数值微分损失中的偏微分方程约束。边界条件信息被传递至流体域中的模板（蓝色方形）：(a) 在规则区域中成功传递；(b) 在不规则区域中，由于区域外存在定义不足的模板（空心方形），传递失败。(c) 当前 LSA-PINN 框架下相关定义的示意图。*BCXN* 损失 在 3 个点之间施加约束：外部模板点 A、边界点 P，以及流体域内的镜像点 Q。

以局部坐标n （表面法向）为参照，沿AQ 对点 A 和 Q 相对于点 P 进行泰勒级数展开，可得：

$$\vec{u}_{ES} = \vec{u}_{BC} + \overline{AP}\frac{\partial \vec{u}_{BC}}{\partial n} + \frac{\overline{AP}^2}{2}\frac{\partial^2 \vec{u}_{BC}}{\partial n^2} + O(\overline{AP}^3) \tag{3a}$$

$$\vec{u}_{MI} = \vec{u}_{BC} - \overline{PQ}\frac{\partial \vec{u}_{BC}}{\partial n} + \frac{\overline{PQ}^2}{2}\frac{\partial^2 \vec{u}_{BC}}{\partial n^2} + O(\overline{PQ}^3) \tag{3b}$$

其中，$\overline{AP}$ 和$\overline{PQ}$分别是 A 与 P、P 与 Q 之间的距离，并且$AQ = AP + PQ$。因此，可以推导出 A 点处的场值 $\tilde{u}_{ES}$：

$$\vec{u}_{ES} = \vec{u}_{MI} + (\vec{u}_{BC} - \vec{u}_{MI}) \times \frac{\overline{AQ}}{\overline{PQ}} + O(\overline{AQ}^2) \tag{4}$$

从上述方程可以看出， $\tilde{u}_{ES}$ 受到 $\tilde{u}_{BC}$ 和 $\tilde{u}_{MI}$的线性约束。镜像点 Q 可以方便地选在模板中心，即近边界样本的采样位置；但我们选择沿边界点 P 的法向设置镜像点 Q（参见图 2c），因为这种做法更通用，且在我们的实验中表现更好。因此，$AP = PQ$，公式 (4) 变为：

$$\vec{u}_{ES} = 2\vec{u}_{BC} - \vec{u}_{MI} \tag{5}$$

对于纽曼型和罗宾型边界条件，可以采用类似的过程推导相应的线性约束。

B. 采用直接强制的边界连通性（*BCXN*）损失

我们可以直接应用式（*5*）计算场值 $\tilde{u}_{ES}$，以求得在使用 ND 型方案对近边界采样点施加偏微分方程约束时，任意外部模板点处的场值。这种直接强制方法与数值计算中的直接强制浸入边界法有关。接下来，我们可以使用以下*BCXN* 损失：

$$L_{LSA-PINN} = L_{PDE(BCXN)} \tag{6}$$

作为我们 LSA-PINN 的偏微分方程损失项，其中近边界采样点上的偏微分方程约束由局部线性约束进行调制。在这种实现方式中，边界条件被隐式融入训练损失中，因此无需使用边界条件损失项。

C. 镜像点处场值 $\tilde{u}_{MI}$ 的评估步骤

请注意，物理信息神经网络方法通常需要一个程序，既能从目标几何区域内部采样，也能从边界采样。与传统物理信息神经网络相比，LSA-PINN 还涉及额外步骤。我们首先介绍计算镜像点位置（沿法向）的步骤：

- **判断模板是否位于外部**。在本研究中，我们首先构造目标几何区域 ϕ [55],的水平集函数，由此可以找到模板点 A 与边界点 P 之间的最短距离AP 。根据水平集函数的符号，我们可以判断模板是否位于外部。

- **计算外部模板点的镜像点位置**。确定AP 后，我们计算流体域内镜像点 Q 的位置，其距离为$AP = PQ$。

— —

实际上，给定一组固定的训练样本后，所有外部模板点及其镜像位置都可以预先计算。另一方面，镜像点处的场值 $\tilde{u}_{MI}$ 取决于 LSA-PINN 的输出，因此会在训练迭代中计算：

- **多层感知机架构**。 $\tilde{u}_{MI}$ 值可通过在镜像点处直接计算 $\tilde{u}_{MI}(\tilde{x};\boldsymbol{w})$ 得到。
- **卷积神经网络架构**。当镜像点的位置与卷积神经网络网格不重合时，可使用以下反距离加权插值函数，求得镜像点 [56] 处的 $\tilde{u}_{MI}$ 值。

$$u(\tilde{x},\tilde{y}) = \Sigma u(x,y)\phi'\left(\frac{x-\tilde{x}}{h}\right)\phi'\left(\frac{y-\tilde{y}}{h}\right) \tag{7a}$$

$$\phi'\left(\frac{x-\tilde{x}}{h}\right) = \phi\left(\frac{x-\tilde{x}}{h}\right)/\Sigma\phi\left(\frac{x-\tilde{x}}{h}\right) \tag{7b}$$

$$\phi'\left(\frac{y-\tilde{y}}{h}\right) = \phi\left(\frac{y-\tilde{y}}{h}\right)/\Sigma\phi\left(\frac{y-\tilde{y}}{h}\right) \tag{7c}$$

$$\phi(r) = \begin{cases} \frac{1}{8}\left(3 - 2|r| + \sqrt{1+4|r|+4r^2}\right), & |r| \le 1 \\ \frac{1}{8}\left(5 - 2|r| - \sqrt{-7+12|r|-4r^2}\right), & 1 < |r| \le 2 \\ 0, & |r| > 2 \end{cases} \tag{7d}$$

($\tilde{x}$ $\tilde{y}$) 是待插值镜像点的位置，而 (x，y) 是流体域内训练样本的位置。在本研究中，h 取为 Δx，由此得到 a2Δ x 插值半径。一旦计算出 $\tilde{u}_{MI}$ 和 $\tilde{u}_{BC}$，我们便可在训练过程中根据式（5）计算 $\tilde{u}_{ES}$ 。

V. 结果

在本研究中，我们使用多层感知机和卷积神经网络架构，研究了所提出的 LSA-PINN 在学习多个复杂流体动力系统时的性能。我们进一步证明，LSA-PINN 广泛适用于不同的数值微分型训练损失，包括数值微分型（*BCXN-ND*）和 CAN 型（*BCXN-CAN*）方案。

基于各自物理信息神经网络解的均方误差（MSE），将 LSA-PINN（*BCXN-ND-*、*BCXN-CAN-loss*）与基线物理信息神经网络（*ND-*、*CAN-*、*AD-loss*）模型进行比较，

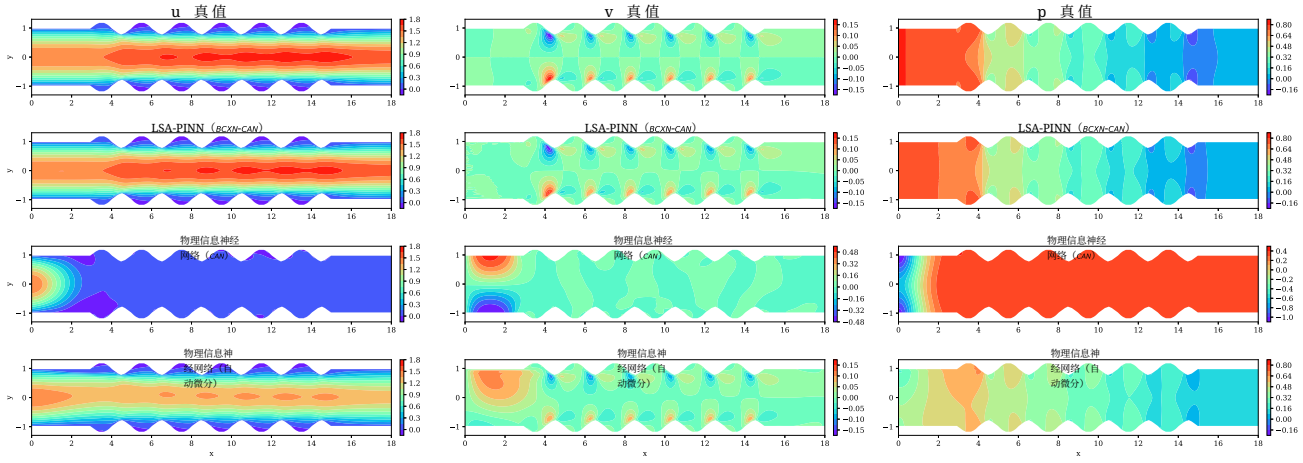


图 3. 波浪形通道流动问题中，真实值与物理信息神经网络解的u速度、v速度和压力p等值线对比（10 次运行的中位数均方误差），雷诺数 $=100$。

速度向量 $\vec{u}$ 相对于真实值的误差。为便于比较，所有模型采用相同的网络架构、初始化分布和训练设置。相关设置汇总于附录 A 的表 VI 和表 VII。每项实验均采用不同的初始化独立运行 10 次，并据此给出结果。真实值解由内部数值求解器通过改进的无散条件补偿（IDFC）方法 [57]求得。3 个流体动力学测试案例如下：

- **2D 半圆形顶盖驱动方腔流，**雷诺数 $=1000$。半圆形空腔内的流动由顶壁上的顶盖速度$u_{\text{顶盖}}=1$，$v_{\text{顶盖}}=0$驱动。其他壁面施加无滑移条件（$u=v=0$）。空腔中心附近有一个主涡。
- **2D 顶盖驱动方腔流，**雷诺数 $=1000$。1×1 单位正方形腔体内的流动由顶壁上的顶盖速度$u_{\text{顶盖}}=1$，$v_{\text{顶盖}}=0$ 驱动。其他壁面施加无滑移条件。由于包含复杂物理现象，顶盖驱动方腔流被广泛选作许多数值方法的基准算例。在雷诺数 $=1000$时，腔体中心附近有一个主涡，右下区域和左下区域则有次级涡和三级涡。
- **2D 波状通道流，**雷诺数 $=100$。本文研究流体流经长波状通道的流动。左侧边界处的入口速度分布定义为 $u(0,\ y)=-\frac{3}{2}y^2+\frac{3}{2}$，$v=0$。顶部和底部壁面采用无滑移条件，而右侧边界采用出口边界条件（$\frac{\partial u}{\partial x}=\frac{\partial v}{\partial x}=p=0$）。

A. 多层感知机架构 *LSA-PINN*

在多层感知机架构实验中，我们评估了 LSA-PINN 的性能，并将其与此前针对多层感知机架构发表的其他自动微分-、*CAN*-和数值微分损失基线方法进行比较。所有模型均使用 Adam算法训练，在相对稀疏的训练样本集上（半圆形顶盖驱动空腔、顶盖驱动空腔和波浪形通道流动问题分别使用 4,230、10,400 和 15,200 个点）进行 300,000 次迭代。多次独立运行的平均均方误差和相对 L2 误差汇总于表 I 中。

表 I 多层感知机架构物理信息神经网络的模型性能比较

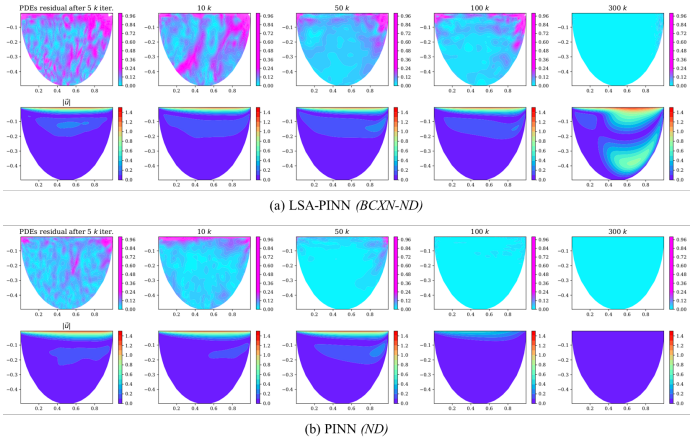
问题	模型	MSE	相对误差
半圆形顶盖驱动方腔流，雷诺数 $=1000$	*BCXN-CAN*	6.21×10^{-5}	3.55×10^{-2}
	BCXN-ND	1.61×10^{-4}	5.95×10^{-2}
	CAN	2.82×10^{-2}	7.62×10^{-1}
	ND	3.91×10^{-2}	8.56×10^{-1}
	AD（SIREN）	1.74×10^{-2}	5.95×10^{-1}
	自动微分（原版）	5.46×10^{-2}	9.83×10^{-1}
顶盖驱动方腔流，雷诺数 $=1000$	*BCXN-CAN*	5.10×10^{-5}	3.38×10^{-2}
	CAN	4.23×10^{-2}	9.77×10^{-1}
	AD（SIREN）	3.47×10^{-2}	8.87×10^{-1}
波状通道流，雷诺数 $=100$	*BCXN-CAN*	1.34×10^{-4}	7.28×10^{-2}
	CAN	5.96×10^{-1}	2.55×10^{0}
	AD（SIREN）	3.52×10^{-2}	7.19×10^{-1}

（I）基于先前工作 [6]实现 CAN-PINN 模型。（II）基于先前工作 [58]–[60]实现 SIREN-PINN 模型。（III）基于先前工作 [1]实现原版 PINN 模型。

结果表明，我们提出的 LSA-PINN 方法能够在涵盖不规则区域和复杂物理现象的多种问题中，成功学习到具有较低均方误差的良好解。所有 3 个测试案例均显示，采用 LSA-PINN 方法后解的精度有显著提升（>2 低了数个数量级 $\vec{u}$ 均方误差）。基线物理信息神经网络未能得出合理的解，也就是说，在当前采样密度和训练迭代次数下，它们无法正确学习流动情况，示例见图 3。

信息传播案例研究。从边界到内部区域的信息传播被认为是成功进行 PINN 建模的关键因素 [16], [31]。因此，我们研究了半圆形顶盖驱动空腔在采用和未采用*BCXN* 损失 的情况下，多次训练过程中偏微分方程残差的演变，以更好地了解*BCXN* 损失所带来的连通性增强和局部结构近似的影响。偏微分方程残差的示例等值线图以及

各训练迭代次数下的速度如图 4 所示。



(a) LSA-PINN *(BCXN-ND)*

(b) PINN *(ND)*

图 4. 半圆形顶盖驱动方腔流问题中，(a) LSA-PINN（*BCXN-ND*）和 (b) PINN（数值微分）的偏微分方程误差残差及速度等值线图，雷诺数 $= 1000$。

如前文所述，即使经过 300,000 次训练迭代，数值微分方法通常仍无法收敛到良好的模型，但在训练过程中加入*BCXN* 损失可以改善这一情况。速度等值线清楚地展示了收敛性的改善。此外，在训练早期阶段，偏微分方程残差在区域内的分布存在显著差异。对于典型的数值微分-PINN，偏微分方程残差会迅速降至零，尤其是在远离边界条件的区域。这也对应于数值微分损失PINN 的速度剖面收敛到错误解，表明边界条件中的信息在区域内部丢失了。另一方面，加入*BCXN* 损失后，区域内的偏微分方程残差会持续更长时间，这表明来自顶部边界的信息会持续传递至内部区域，直至 LSA-PINN 模型收敛。这也符合系统的物理动力学：正如“顶盖驱动方腔流”这一名称所示，整个区域内的流动很大程度上取决于顶盖的运动。

不同边界连通性采样点的案例研究。顶盖驱动方腔测试案例具有规则区域，但我们选择了单元中心型（即近边界点需要区域外的支撑点）采样位置，以展示所提出的方法对这类场景的适用性。为进一步说明区域外存在近边界模板点的影响，我们还研究了单元中心型采样点与节点中心型采样点（即近边界点使用与区域边界重合的支撑点）之间的差异。我们在雷诺数 $Re = 1000$ 的顶盖驱动腔流动中，针对*BCXN-CAN*、*CAN* 和自动微分 模型，使用不同的批量大小和点采样密度测试了这两组采样点，并分别训练了 300,000 次迭代。我们汇总了多次独立运行的均方误差和 L2相对误差

并列于表II中。对于该问题，定性上较好的解对应的均方误差为 1×10^{-3} 及以下。

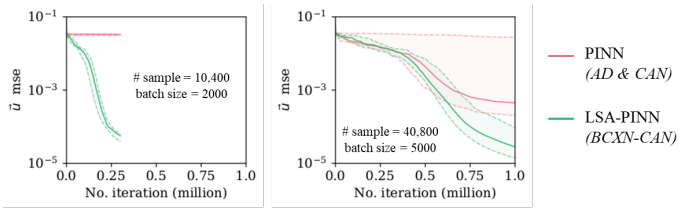
表II 不同采样密度下的模型性能。 [C] 和[N] 分别表示单元中心型和节点中心型采样布置。

批量大小	点数	模型	50×50		100×100	
			MSE	相对误差	MSE	相对误差
500	[C]	*BCXN-CAN*	2.63×10^{-4}	7.75×10^{-2}	3.06×10^{-3}	2.63×10^{-1}
		CAN 与自动微分	$> 1 \times 10^{-2}$	$> 5 \times 10^{-1}$	$> 1 \times 10^{-2}$	$> 5 \times 10^{-1}$
	[N]	*CAN*	$> 1 \times 10^{-2}$	$> 5 \times 10^{-1}$	$> 1 \times 10^{-2}$	$> 5 \times 10^{-1}$
		AD	$> 1 \times 10^{-2}$	$> 5 \times 10^{-1}$	$> 1 \times 10^{-2}$	$> 5 \times 10^{-1}$
1000	[C]	*BCXN-CAN*	2.23×10^{-4}	7.13×10^{-2}	3.97×10^{-4}	9.47×10^{-2}
		CAN 与自动微分	$> 1 \times 10^{-2}$	$> 5 \times 10^{-1}$	$> 1 \times 10^{-2}$	$> 5 \times 10^{-1}$
	[N]	*CAN*	$> 1 \times 10^{-2}$	$> 5 \times 10^{-1}$	$> 1 \times 10^{-2}$	$> 5 \times 10^{-1}$
		AD	$> 1 \times 10^{-2}$	$> 5 \times 10^{-1}$	$> 1 \times 10^{-2}$	$> 5 \times 10^{-1}$
2000	[C]	*BCXN-CAN*	2.36×10^{-4}	7.34×10^{-2}	7.14×10^{-5}	4.02×10^{-2}
		CAN 与自动微分	$> 1 \times 10^{-2}$	$> 5 \times 10^{-1}$	$> 1 \times 10^{-2}$	$> 5 \times 10^{-1}$
	[N]	*CAN*	7.41×10^{-4}	1.32×10^{-1}	$> 1 \times 10^{-2}$	$> 5 \times 10^{-1}$
		AD	$> 1 \times 10^{-2}$	$> 5 \times 10^{-1}$	$> 1 \times 10^{-2}$	$> 5 \times 10^{-1}$
4000	[C]	*BCXN-CAN*			2.38×10^{-5}	2.31×10^{-2}
		CAN 与自动微分			$> 1 \times 10^{-2}$	$> 5 \times 10^{-1}$
	[N]	*CAN*			2.94×10^{-4}	8.24×10^{-2}
		AD			$> 1 \times 10^{-2}$	$> 5 \times 10^{-1}$

通过这项实验，我们发现，对于较大的批量大小和更多的训练样本，节点中心型耦合自动数值微分损失 PINN 模型的表现优于单元中心版本。这进一步强调了我们的假设：准确地对边界条件进行建模是取得良好表现的重要因素。由于这是一个正方形区域，采用与区域边界完全连通的采样点分布，自然会带来更好的表现。

此外，在单元中心 LSA-PINN 中加入我们提出的 *BCXN* 损失 ，可以提升模型表现，使其误差低于经过适当对齐的基线 PINN 模型，尤其是在稀疏训练样本设置下。这表明，LSA-PINN 在采样策略的选择上具有灵活性，而传统采样策略可能需要采用与边界保持连通的采样方式。

不同采样密度下的案例研究。从上述实验中，我们注意到，采样密度增加 4 倍（从 50×50 到 100×100）可以提升模型性能。因此，我们进一步研究顶盖驱动方腔流问题，以了解精度收敛与采样密度之间的权衡。我们使用更密集（即 200×200）的样本集进行训练，直到*CAN* 和 自动微分 模型的准确度开始提升。请注意，更密集的采样需要更多训练迭代次数和更大的批量大小，而调整这些训练超参数可能非常耗时。尽管所有物理信息神经网络模型最终都能通过更密集的采样提升准确度，但图 5 表明，从密集样本集中学习要困难得多、速度也慢得多，这凸显了 LSA-PINN 的优势。值得鼓舞的是，在顶盖驱动方腔流问题中，LSA-PINN 所学习到的解更加准确（误差低 >1个数量级 $\bar{u}$ 均方误差），训练迭代次数更少（<50%），训练样本也更少（<50%），相比之下，基线 *CAN-* 和自动微分损失 物理信息神经网络模型的表现则较差。



(a) Sparse sample (b) Dense sample

图5. PINN分别在（a）稀疏和（b）密集训练场景下学习顶盖驱动方腔流的解。图中绘出了 $\vec{u}$ 均方误差的收敛趋势。粗线表示中位收敛路径，阴影区域表示从 $10^{th}-90^{th}$ 百分位数计算的百分位间距。

*B. CNN*架构的*LSA-PINN*

BCXN 损失 方法对神经网络的可微性没有要求，因此可以自由应用于任何神经网络架构，包括卷积神经网络。因此，我们将 LSA-PINN *(BCXN-ND)* 与此前针对卷积神经网络架构发表的基线物理信息神经网络*(ND)* 方法进行比较，并在表III中汇总了多次独立运行所得的平均均方误差和相对 L2 误差。表VII汇总了所用的 U-Net 架构和训练设置。本节中的所有物理信息神经网络模型均在单个离散化二维输入上训练，该输入由 3 个通道组成，分别为 x 坐标、y 坐标以及用于区分流体域和非流体区域的指示变量。模型使用 Adam算法进行了 200,000 次迭代训练，训练样本集相对稀疏；具体而言，半圆形和方形顶盖驱动方腔流问题分别使用属于流体域的 3,930 个和 10,000 个内部点。

表III 卷积神经网络架构 PINN 的模型性能比较

问题	模型	MSE	相对误差
半圆形顶盖驱动方腔流，雷诺数 $=1000$	*BCXN-ND*	1.14×10^{-4}	4.92×10^{-2}
	ND	3.96×10^{-2}	8.55×10^{-1}
顶盖驱动方腔流，雷诺数 $=1000$	*BCXN-ND*	2.52×10^{-5}	4.46×10^{-2}
	ND	2.38×10^{-2}	1.00×10^{0}
瞬态半圆形顶盖驱动腔流，雷诺数 $=500$	*BCXN-ND*	3.95×10^{-4}	1.29×10^{-1}
	ND	9.66×10^{-3}	6.81×10^{-1}

与多层感知机实验类似，面向卷积神经网络的 LSA-PINN 在解的精度方面较基线数值微分损失 物理信息神经网络有显著提升（均方误差降低了 2–3 个数量级 $\vec{u}$ ）。结果表明，LSA-PINN 能够有效学习采用 U-Net 架构的流体动力学问题的解；此外，*BCXN* 损失 还能通过解决离散化输入无法精确表示复杂几何形状所引发的训练问题，使卷积神经网络架构 PINN 受益。

瞬态二维半圆形顶盖驱动腔流，雷诺数 $=500$。除了稳态半圆形顶盖驱动腔流外，我们还进一步评估了将卷积神经网络架构 LSA-PINN 应用于此问题的瞬态版本建模所带来的改进。在本例中，问题的初始条件对应于静止流体域的初始状态（在 t = 0）。随后使用 PINN模型对半圆形顶盖驱动型空腔内的流动从 $t=0$ 到$t=2$, $\Delta t=0.1$的演化进行建模。模型设置见表VII。瞬态物理信息神经网络模型的输入中还额外包含t通道。

表II对比了LSA-PINN（*BCXN-ND*）与基线物理信息神经网络（数值微分）的性能。在$t=0$ 至$t=2.0$ 之间所有时间点的平均误差，在包含和不包含*BCXN* 损失 的情况下，均方误差分别为3.95e-4和9.66e-3。因此，在该瞬态问题中加入*BCXN* 损失 可使均方误差降低 >1 个数量级。图6给出了*BCXN-ND*模型和数值微分 模型的速度等值线图，以展示两者在收敛方面的差异。

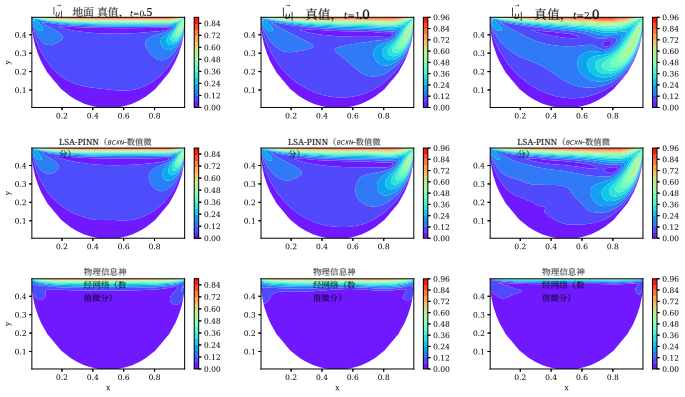


图6. 真实值和物理信息神经网络解在$t=0.5$、$t=1.0$ 和$t=2.0$ 时的瞬态模拟速度等值线。

问题复杂度增加时的性能。为更好地了解*BCXN* 损失所带来的准确度提升，我们还研究了LSA-PINN在更高效地建模复杂问题方面的作用。因此，我们针对一系列雷诺数，分别训练LSA-PINN和基线物理信息神经网络，以模拟半圆形顶盖驱动方腔流。随着雷诺数 增加，流动逐渐发展，其复杂性通常需要更先进的算法和更多训练样本，才能更好地捕捉非线性效应。

多次运行所得的平均均方误差和相对 L2 误差如表 IV 所示。我们清楚地看到，对于更复杂（雷诺数更高）的*Re* 问题，性能差异日益显著，这表明对于更复杂的问题，LSA-PINN 等方法越来越重要。

表 IV 二维半圆形顶盖驱动腔体问题的模型性能

指标	模型	雷诺数（*Re*）				
		50	100	200	400	1000
MSE	*BCXN-ND*	2.51×10^{-5}	2.76×10^{-5}	3.56×10^{-5}	4.58×10^{-5}	1.14×10^{-4}
	ND	2.98×10^{-4}	2.79×10^{-4}	3.74×10^{-4}	2.33×10^{-3}	3.96×10^{-2}
相对误差	*BCXN-ND*	0.0258	0.0269	0.0305	0.0335	0.0492
	ND	0.0823	0.0732	0.0780	0.211	0.855

元 PINN模型。 CNN 架构 PINN模型将整个空间域表示为单个输入。我们可以使用多个样本训练模型，每个样本代表不同的场景，例如几何形状或物理行为的变化。这样的元 PINN模型可用于预测新场景。因此，我们进一步研究 LSA-

PINN同时学习多个流动场景（跨越雷诺数 $=50-1000$）解的能力，场景为半圆形顶盖驱动腔体几何。模型架构和训练设置见附录 A 表VII。我们使用 LSA-PINN *(BCXN-ND)*模型预测训练期间未使用的雷诺数 $=300$，500，600，700，900场景。结果汇总于图 7。当前的 LSA-PINN方法成功学习到了良好的解，在大多数雷诺数场景下均方误差较低（即 <1 e-4），并取得了良好的预测准确度。

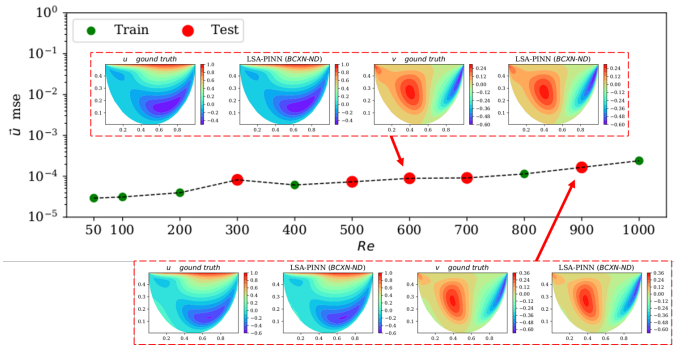


图 7. 使用卷积神经网络架构，LSA-PINN 并行学习多个二维半圆形顶盖驱动空腔流动场景的解。图中绘出了训练和测试雷诺数下真实值与 LSA-PINN 解之间的 $\bar{u}$ 均方误差。插图展示了雷诺数为 $=600$ 和 $=900$ 时真实值与 LSA-PINN 解的 u 速度等值线和 v 速度等值线。

C. *BCXN* 损失对训练时间的影响

我们进一步比较了在雷诺数为 $=1000$ 的半圆形顶盖驱动空腔问题中，使用和不使用 *BCXN* 损失 训练多层感知机和卷积神经网络模型所需的时间。10 次独立运行所得的平均均方误差和相对 L2 误差取平均值，并记录实际运行时间的最短值和最长值。所有实验均在一台配备 Intel Xeon Gold 6248 处理器和 Nvidia RTX 2080 Ti GPU 的 20 核工作站上，使用标准 TensorFlow 和 Keras 实现运行。

表 V 展示了多层感知机架构和卷积神经网络架构模型的定量比较。为了在不同架构类型内进行公平比较，所有多层感知机模型均采用相同的批量大小（1,000）和模型规模，进行相同次数的训练迭代（300,000）；所有卷积神经网络模型也采用相同的模型规模，进行相同次数的训练迭代（200,000）。

表 V 不同物理信息神经网络架构的模型性能

模型	方法	MSE	相对误差	时间（分钟）
MLP	*BCXN-CAN*	6.21×10^{-5}	3.55×10^{-2}	$55-57$
MLP	*BCXN-ND*	1.61×10^{-4}	5.95×10^{-2}	$23-24$
MLP	*CAN*	2.82×10^{-2}	7.62×10^{-1}	$40-46$
MLP	*ND*	3.91×10^{-2}	8.56×10^{-1}	$20-22$
MLP	*AD（SIREN）*	1.74×10^{-2}	5.95×10^{-1}	$58-61$
MLP	自动微分（原版）	5.46×10^{-2}	9.83×10^{-1}	$50-55$
CNN	*BCXN*	1.14×10^{-4}	4.92×10^{-2}	$126-136$
CNN	*ND*	3.96×10^{-2}	8.55×10^{-1}	$121-137$

从这项研究中我们发现，在所有案例中加入*BCXN* 损失只会略微增加模型训练所需的时间。对于 CNN 型架构和 ND-多层感知机型模型，增加幅度不到 5%。值得注意的是，尽管耗时略有增加，均方误差通常会降低一个数量级以上。在许多情况下，加入*BCXN* 损失后，模型训练时间与性能之间的这种权衡可能更为可取。

VI. 结论

我们提出了一种 LSA-PINN方法，能够利用稀疏训练样本，更高效地学习复杂几何形状下偏微分方程的解。这是通过在训练过程中隐式施加线性约束来实现的。由于推导中对泰勒级数展开进行了截断，本研究采用了线性约束；未来值得探究不同的高阶近似能否带来更好的性能。总体而言，我们全面的实验研究表明，LSA-PINN 在流体动力学的各类测试问题中具有实际优势：即使所需的训练迭代次数和训练样本少得多，解的准确度仍能提升数个数量级。

参考文献

[1] M. Raissi、P. Perdikaris 和 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和逆问题的深度学习框架》，*Journal of Computational physics*，第378卷，第686–707页，2019年。[2] G. E. Karniadakis、I. G. Kevrekidis、L. Lu、P. Perdikaris、S. Wang 和 L. Yang，《物理信息机器学习》，*Nature Reviews Physics*，第3卷第6期，第422–440页，2021年。[3] S. Cuomo、V. S. Di Cola、F. Giampaolo、G. Rozza、M. Raissi 和 F. Piccialli，《通过物理信息神经网络开展科学机器学习：现状与未来方向》，*arXiv*预印本 *arXiv:2201.05624*，2022年。[4] H. Gao、L. Sun 和 J.-X. Wang，《Phygeonet：用于求解不规则区域上参数化稳态偏微分方程的物理信息几何自适应卷积神经网络》，*Journal of Computational Physics*，第428卷，文章编号110079，2021年。[5] Z. Fang，《一种基于卷积神经网络的高效混合物理信息神经网络》，*IEEE Transactions on Neural Networks and Learning Systems*，2021年。[6] P.-H. Chiu、J. C. Wong、C. Ooi、M. H. Dao 和 Y.-S. Ong，《CAN-PINN：一种基于耦合自动微分与数值微分方法的快速物理信息神经网络》，*Computer Methods in Applied Me- chanics and Engineering*，第395卷，文章编号114909，2022年。[7] Y. Lu、H. Chen、J. Lu、L. Ying 和 J. Blanchet，《椭圆型偏微分方程的机器学习：快速收敛率泛化界、神经缩放定律与极小极大最优性》，*arXiv*预印本 *arXiv:2110.06897*，2021年。[8] S. Mishra 和 R. Molinaro，《物理信息神经网络近似一类偏微分方程反问题时的泛化误差估计》，*IMA Journal of Numerical Analysis*，第42卷第2期，第981–1022页，2022年。[9] S. Markidis，《新旧之争：物理信息深度学习能否取代传统线性求解器？》，*Frontiers in big Data*，第92页，2021年。[10] C. Anitescu、E. Atroshchenko、N. Alajlan 和 T. Rabczuk，《求解二阶边值问题的人工神经网络方法》，*Computers, Materials and Continua*，第59卷第1期，第345–359页，2019年。[11] L. McClenny 和 U. Braga-Neto，《使用软注意力机制的自适应物理信息神经网络》，*arXiv*预印本 *arXiv:2009.04544*，2020年。[12] C. L. Wight 和 J. Zhao，《使用自适应物理信息神经网络求解 Allen-Cahn 方程和 Cahn-Hilliard方程》，*arXiv*预印本 *arXiv:2007.04542*，2020年。

[13] M. A. Nabian、R. J. Gladstone 和 H. Meidani，《通过重要性采样高效训练物理信息神经网络》，Computer-Aided Civil and Infrastructure Engineering，第36卷，2021年，第977–962页，第8期。[14] L. Lu、X. Meng、Z. Mao 和 G. E. Karniadakis，《Deepxde：用于求解微分方程的深度学习库》，《SIAM评论》，第63卷第1期，第208–228页，2021年。[15] K. O. Lye、S. Mishra、D. Ray 和 P. Chandrashekar，《迭代代理模型优化（ISMO）：一种用于偏微分方程约束下深度神经网络优化的主动学习算法》，《Computer Methods in Applied Mechanics and Engineering》，第374卷，文章编号113575，2021年。[16] A. Daw、J. Bu、S. Wang、P. Perdikaris 和 A. Karpatne，《重新思考采样在物理信息神经网络中的重要性》，arXiv预印本 arXiv:2207.02338，2022年。[17] R. Mattey 和 S. Ghosh，《一种用于训练物理信息神经网络以求解 Allen-Cahn 方程和 Cahn-Hilliard方程的新型序贯方法》，《Computer Methods in Applied Mechanics and Engineering》，第390卷，文章编号114474，2022年。[18] C. Wu、M. Zhu、Q. Tan、Y. Kartha 和 L. Lu，《对物理信息神经网络的非自适应采样和基于残差的自适应采样进行全面研究》，《Computer Methods in Applied Mechanics and Engineering》，第403卷，文章编号115671，2023年。[19] A. D. Jagtap、E. Kharazmi 和 G. E. Karniadakis，《离散域上的守恒物理信息神经网络：守恒定律及其在正问题和逆问题中的应用》，《Computer Methods in Applied Mechanics and Engineering》，第365卷，文章编号113028，2020年。[20] A. D. Jagtap 和 G. E. Karniadakis，《扩展物理信息神经网络（XPINNs）：一种用于非线性偏微分方程的通用时空域分解深度学习框架》，载于 AAAI春季研讨会：MLPS，2021年。[21] K. Shukla、A. D. Jagtap 和 G. E. Karniadakis，《通过区域分解实现并行物理信息神经网络》，Journal of Computational Physics，第447卷，文章编号110683，2021年。[22] S. Dong 和 Z. Li，《用于求解线性和非线性偏微分方程的局部极限学习机与区域分解》，《Computer Methods in Applied Mechanics and Engineering》，第387卷，文章编号114129，2021年。[23] K. Li、K. Tang、T. Wu 和 Q. Liao，《D3m：一种用于偏微分方程的深度区域分解方法》，《IEEE Access》，第8卷，第5283–5294页，2019年。[24] N. Wandel、M. Weinmann 和 R. Klein，《从零开始学习不可压缩流体动力学——迈向快速、可微且具有泛化能力的流体模型》，arXiv预印本 arXiv:2006.08762，2020年。[25] ——，《在三维空间中将不可压缩纳维–斯托克斯方程教给快速神经代理模型》，《流体物理学》，第33卷第4期，文章编号047117，2021年。[26] R. Ranade、C. Hill 和 J. Pathak，《DiscretizationNet：一种使用有限体积离散求解纳维–斯托克斯方程的机器学习求解器》，《Computer Methods in Applied Mechanics and Engineering》，第378卷，文章编号113722，2021年。[27] N. Wandel、M. Weinmann、M. Neidlin 和 R. Klein，《Spline-PINN：利用快速、物理信息驱动的Hermite样条卷积神经网络实现无数据偏微分方程求解》，载于AAAI人工智能大会论文集，第36卷第8期，第8529–8538页，2022年。[28] P. Ren、C. Rao、Y. Liu、J.-X. Wang 和 H. Sun，《PhyCRNet：用于求解时空偏微分方程的物理信息卷积循环网络》，《Computer Methods in Applied Mechanics and Engineering》，第389卷，文章编号114399，2022年。[29] S. K. Mitusch、S. W. Funke 和 M. Kuchta，《混合有限元-神经网络模型：结合人工神经网络与有限元法》，Journal of Computational Physics，第446卷，文章编号110651，2021年。[30] Y. Shin、J. Darbon 和 G. E. Karniadakis，《论物理信息神经网络对线性二阶椭圆型和抛物型偏微分方程的收敛性》，arXiv预印本 arXiv:2004.01806，2020年。[31] S. Wang、S. Sankaran 和 P. Perdikaris，《训练物理信息神经网络，只需遵循因果性》，arXiv预印本 arXiv:2203.07404，2022年。[32] M. Elhamod、J. Bu、C. Singh、M. Redell、A. Ghosh、V. Podolskiy、W.-C. Lee 和 A. Karpatne，《CoPhy-PGNN：利用竞争损失函数学习物理引导神经网络以求解特征值问题》，《ACM智能系统与技术汇刊》（TIST），2020年。[33] R. Bischof 和 M. Kraus，《物理信息深度学习的多目标损失平衡》，arXiv预印本 arXiv:2110.09813，2021年。[34] X. Jin、S. Cai、H. Li 和 G. E. Karniadakis，“Nsfnets（navier-斯托克斯流网络）：用于不可压缩纳维–斯托克斯方程的物理信息神经网络，”《Journal of Computational Physics》，第 426 卷，文章编号 109951，2021年。[35] P. Thanasutives、M. Numao 和 K.-i. Fukui，“通过对抗式多任务学习增强物理信息神经网络以求解偏微分方程”，载于 2021年国际神经网络联合会议（IJCNN）。IEEE，2021年，第 1–9 页。[36] S. Maddu、D. Sturm、C. L. M¨uller 和 I. F. Sbalzarini，“逆狄利克雷加权使物理信息神经网络的可靠训练成为可能”，《机器学习：科学与技术》，第 3 卷，第 1 期，文章编号 015026，2022年。[37] T. de Wolff、H. C. Lincopi、L. Mart´ı 和 N. S anchez-Pi，“Mopinns：一种用于物理信息神经网络的进化式多目标方法”，载于遗传与进化计算会议附会论文集，2022年，第 228–231 页。[38] R. van der Meer、C. W. Oosterlee 和 A. Borovykh，“用于通过神经网络求解偏微分方程的最优加权损失函数”，《计算与应用数学杂志》，第 405 卷，文章编号 113887，2022年。[39] Z. Xiang、W. Peng、X. Liu 和 W. Yao，“自适应损失平衡物理信息神经网络”，《神经计算》，第 496 卷，第 11–34 页，2022年。[40] X. Huang、H. Liu、B. Shi、Z. Wang、K. Yang、Y. Li、M. Wang、H. Chu、J. Zhou、F. Yu 等，“一种用于求解含点源偏微分方程的通用 PINN 方法。”[41] S. Wang、Y. Teng 和 P. Perdikaris，“理解并缓解物理信息神经网络中的梯度流病态”，《SIAM 科学计算杂志》，第 43 卷，第 5 期，第 A3055–A3081 页，2021年。[42] I. E. Lagaris、A. Likas 和 D. I. Fotiadis，“用于求解常微分方程和偏微分方程的人工神经网络”，《IEEE 神经网络汇刊》，第 9 卷，第 5 期，第 987–1000 页，1998年。[43] I. E. Lagaris、A. C. Likas 和 D. G. Papageorgiou，“用于不规则边界边值问题的神经网络方法”，《IEEE 神经网络汇刊》，第 11 卷，第 5 期，第 1041–1049 页，2000年。[44] K. S. McFall 和 J. R. Mahan，“通过精确满足任意边界条件求解边值问题的人工神经网络方法”，《IEEE 神经网络汇刊》，第 20 卷，第 8 期，第 1221–1233 页，2009年。[45] J. Berg 和 K. Nystr¨o m，“一种用于复杂几何形状中偏微分方程的统一深度人工神经网络方法”，《神经计算》，第 317 卷，第 28–41 页，2018年。[46] M. A. Nabian 和 H. Meidani，“求解高维随机微分方程的深度学习方法”，《概率工程力学》，第 57 卷，第 14–25 页，2019年。[47] S. Karumuri、R. Tripathy、I. Bilionis 和 J. Panchal，“利用深度神经网络无模拟器求解高维随机椭圆偏微分方程”，《Journal of Computational Physics》，第 404 卷，文章编号 109120，2020年。[48] Z. Wang 和 Z. Zhang，“一种利用深度学习方法求解界面问题的无网格方法”，《Journal of Computational Physics》，第 400 卷，文章编号 108963，2020年。[49] L. Lu、R. Pestourie、W. Yao、Z. Wang、F. Verdugo 和 S. G. Johnson，“用于逆向设计且带有硬约束的物理信息神经网络”，《SIAM 科学计算杂志》，第 43 卷，第 6 期，第 B1105–B1132 页，2021年。[50] O. Fuks 和 H. A. Tchelepi，“物理信息机器学习在多孔介质非线性两相输运中的局限性”，《建模与计算机器学习杂志》，第 1 卷，第 1 期，2020年。[51] C. Michoski、M. Milosavljevi´c、T. Oliver 和 D. R. Hatch，“使用深度神经网络求解微分方程”，《神经计算》，第 399 卷，第 193–212 页，2020年。[52] A. A. Ramabathiran 和 P. Ramachandran，“Spinn：用于偏微分方程的稀疏、基于物理且部分可解释的神经网络”，《Journal of Computational Physics》，第 445 卷，文章编号 110600，2021年。[53] W. Ji、W. Qiu、Z. Shi、S. Pan 和 S. Deng，“Stiff-PINN：用于刚性化学动力学的物理信息神经网络”，《物理化学杂志 A》，第 125 卷，第 36 期，第 8098–8106 页，2021年。[54] D. Lucor、A. Agrawal 和 A. Sergent，“用于湍流自然对流代理建模的物理感知深度神经网络”，arXiv 预印本 arXiv:2103.03565，2021年。[55] S. Osher 和 N. Paragios，《成像、视觉与图形学中的几何水平集方法》。施普林格科学与商业媒体，2003年。[56] P.-H. Chiu 和 H. J. Poh，“改进的无散条件补偿耦合框架：求解具有

时变几何形状的流动问题”，《*International Journal for Numerical Methods in Fluids*》，第 93 卷，第 1 期，第 44–70 页，2021年。[57] P.-H. Chiu，“一种改进的无散条件补偿方法，用于在同位网格上求解不可压缩流”，*Com- puters & Fluids*，第 162 卷，第 39–54 页，2018年。[58] V. Sitzmann、J. Martel、A. Bergman、D. Lindell 和 G. Wetzstein，“具有周期性激活函数的隐式神经表示”，《*Advances in Neural Information Processing Systems*》，第 33 卷，第 7462–7473 页，2020年。[59] S. Wang、H. Wang 和 P. Perdikaris，“关于傅里叶特征网络的特征向量偏差：从回归到使用物理信息神经网络求解多尺度偏微分方程”，《*Computer Methods in Applied Mechanics and Engineering*》，第 384 卷，文章编号 113938，2021 年。[60] J. C. Wong 、 C. Ooi 、 A. Gupta 和 Y.-S. Ong，“利用物理信息神经网络在正弦空间中学习”，《*IEEE Transactions on Artificial Intelligence*》，2022年。

附录

A. 多层感知机和卷积神经网络的实验设置

表 VI 第 V-A 节实验所用物理信息神经网络（多层感知机）的架构和训练设置

问题类型	**正向**	**正向**	**正向**
测试案例	二维半圆顶盖驱动空腔 flow, 雷诺数 = 1000	二维顶盖驱动腔流动, 雷诺数 = 1000	二维波状通道 flow, 雷诺数 = 100
控制物理方程	二维 不可压缩纳维–斯托克斯方程	二维 不可压缩纳维–斯托克斯方程	二维 不可压缩纳维–斯托克斯方程
多层感知机架构 用于 *LSA-PINN*, *ND-PINN* & *AD-PINN*	$(x,\ y)$ – 64–30–30–30–[30–30–30–($\hat{u}$), 30–30–30–($\hat{v}$), 30–30–30–($\hat{p}$)]	(x,y) – 128–30–30–30–[30–30–30–($\hat{u}$), 30–30–30–($\hat{v}$), 30–30–30–($\hat{p}$)]	$(x,\ y)$ – 128–40–40–40–[40–40–40–($\hat{u}$), 40–40–40–($\hat{v}$), 40–40–40–($\hat{p}$)]
ND 方案用于 *LSA-PINN* & *ND-PINN*	CAN & 有限差分	CAN	CAN
训练样本（等间距间隔采样 @ 流体域& 边界）	$3,930+300$	$10,000+400$	$14,400+800$
批量大小	$1,000$	$2,000$	$1,000$
训练迭代次数	300,000	300,000	300，000

(I) 对于多层感知机架构，输入和输出之间的数字表示每个隐藏层中的节点数。例如，(x)–64–20–20–20–($\hat{u}$)表示一个输入x，后接4个隐藏层，每层分别有64、20、20和20个节点，以及一个输出$\hat{u}$。(II) 我们将正弦映射 [60] 融入物理信息神经网络的第一个隐藏层，并通过从正态分布$N(0,\sigma^2)$ $\sigma=1$中采样来初始化其权重。后续隐藏层使用“正弦”激活函数，但最终（输出）层使用“线性”激活函数；这些层的权重则通过He均匀分布初始化。(III) 批量大小：每次计算$L_{PINN}(=\lambda_{PDE}L_{PDE}+\lambda_{BC}L_{BC})$和$L^{PINN(df-BCXN)}(=\lambda_{PDE}L^{PDE(df-BCXN)})$所用的随机样本数。除非另有说明，我们使用默认的 $\lambda_{PDE}=1$ 和 $\lambda_{BC}=1$。(IV) 一次训练迭代：计算一次L_{PINN}或$L_{PINN(df-BCXN)}$，以反向传播权重梯度。我们将初始学习率设为1e-3，并在训练进入平台期时降低学习率，直至达到最低学习率5e-6。

表VII V-B节实验中使用的物理信息神经网络（卷积神经网络）架构和训练设置

问题类型	**正向**	**正向**	**正向**	**元建模**
测试案例	二维半圆形顶盖驱动方腔流，雷诺数 = 1000	二维顶盖驱动方腔流, 雷诺数 = 1000	二维非定常半圆形顶盖驱动方腔流, 雷诺数 = 500	二维半圆形顶盖驱动方腔流，雷诺数 = 1000
控制物理方程	二维 不可压缩纳维－斯托克斯方程	二维 不可压缩纳维–斯托克斯方程	二维 瞬态不可压缩纳维–斯托克斯方程	二维 不可压缩纳维–斯托克斯方程。
卷积神经网络（U-Net）架构 用于 *LSA-PINN*& *ND-PINN*	$(\boldsymbol{X}_{(56\times112\times3)})$–, 8–8–↓–16–16–↓–, 32–32–↑–64–64–↑–, 32–32–↑–16–16–↑–[, 8–8–8–($\hat{\boldsymbol{U}}_{u,56\times112}$), 8–8–8–($\hat{\boldsymbol{U}}_{v,56\times112}$), 8–8–8–($\hat{\boldsymbol{U}}_{p,\ 56\times112}$)	$(\boldsymbol{X}_{(104\times104\times3)})$–, 8–8–↓–16–16–↓–, 32–32–↑–64–64–↑–, 32–32–↑–16–16–↑–[, 8–8–8–($\hat{\boldsymbol{U}}_{u,104\times104}$), 8–8–8–($\hat{\boldsymbol{U}}_{v,104\times104}$), 8–8–8–($\hat{\boldsymbol{U}}_{p,\ 104\times104}$)	$(\boldsymbol{X}_{(56\times112\times4)})$–, 8–8–↓–16–16–↓–, 32–32–↑–64–64–↑–, 32–32–↑–16–16–↑–[, 8–8–8–($\hat{\boldsymbol{U}}_{u,56\times112}$), 8–8–8–($\hat{\boldsymbol{U}}_{v,\ 56\times112}$) , 8–8–8–($\hat{\boldsymbol{U}}_{p,\ 56\times112}$)	$(\boldsymbol{X}_{(56\times112\times4)})$–, 8–8–↓–16–16–↓–, 32–32–↑–64–64–↑–, 32–32–↑–16–16–↑–[, 8–8–8–($\hat{\boldsymbol{U}}_{u,56\times112}$), 8–8–8–($\hat{\boldsymbol{U}}_{v,56\times112}$), 8–8–8–($\hat{\boldsymbol{U}}_{p,\ 56\times112}$)
卷积核大小	4×4	6×6	4×4	6×6
ND 方案用于 *LSA-PINN*& *ND-PINN*	有限差分	有限差分	有限差分	有限差分
训练样本 / 批量大小	1 / 1	1 / 1	21 / 2	6 / 1
训练迭代次数	200，000	200,000	100，000	500，000

(I) 对于 U-net 架构，输入和输出之间的数字表示每个隐藏层中的滤波器数量。每个隐藏层由卷积操作、批归一化和非线性“正弦”激活组成。此外，↓表示 2×2 最大池化操作，↑表示 2×2 上采样操作。网络权重采用 He均匀分布初始化。(II) 批量大小：用于一次评估$L_{PINN}(=\lambda_{PDE}L_{PDE}+\lambda_{BC}L_{BC})$ 和 $L_{PINN(df-BCXN)}(=\lambda_{PDE}L_{PDE(df-BCXN)})$ 的样本数量。除非另有说明，我们使用默认的 $\lambda_{PDE}=1$ 和 $\lambda_{BC}=1$。(III) 一次训练迭代：评估L_{PINN}或$L_{PINN(df-BCXN)}$一次，以反向传播权重梯度。我们采用 1e-3 作为初始学习率，并在进入平台期时降低学习率，直至达到最低学习率 5e-6。(IV) 半圆形和规则方腔驱动流几何分别由 3,930 个和 10,000 个属于流体域的内部点构成。