物理信息神经网络中物理规律与数据一致性的作用

NicolásBecerra-Zuniga[a,*], Lucas Lacasa[b], Eusebio Valero[a,c], Gonzalo Rubio[a,c,*]

[a]*ETSIAE-UPM*-马德里理工大学航空学院，红衣主教西斯内罗斯广场*3*号，马德里，*28040*，马德里，西班牙[b]跨学科物理与复杂系统研究所（*IFISC*，*CSIC-UIB*），巴利阿里群岛大学校园，科学技术大楼，帕尔马，*07122*，巴利阿里群岛，西班牙[c]马德里理工大学计算模拟中心，蒙特甘塞多校区，博阿迪利亚德尔蒙特，*28660*，马德里，西班牙

arXiv:2602.10611v1 [cs.LG] 11 Feb 2026

摘要

物理信息神经网络（PINN）作为偏微分方程（PDE）的代理建模策略，尤其适用于标注数据稀缺且可利用物理约束来正则化学习过程的情形，因此受到广泛关注。然而在实际应用中，PINN 经常使用实验数据或数值数据进行训练，而这些数据会因测量噪声、离散误差或建模假设而无法与控制方程完全一致。此类数据与偏微分方程不一致性对 PINN 准确性和收敛性的影响仍未得到充分理解。本文系统分析了数据不一致性如何从根本上限制 PINN 所能达到的准确性。我们提出一致性障碍这一概念，将其定义为由数据保真度与偏微分方程残差的精确强制之间不匹配所导致的误差内在下界。为隔离并量化这一影响，我们研究了具有人工解析解的一维黏性伯格斯方程，从而能够全面控制数据保真度和残差误差。我们使用数值准确性逐步提高的数据集以及完全一致的解析数据训练 PINN。结果表明，尽管引入偏微分方程残差能够帮助 PINN 在一定程度上弥补低保真数据的不足，并恢复主要物理结构，但训练过程最终会饱和于一个由数据不一致性决定的误差水平。使用高保真度数值数据时，PINN 解与使用解析数据训练得到的解已无法区分，这表明一致性障碍实际上已被消除。这些发现阐明了 PINN 中数据质量与物理约束强制之间的相互作用，并为物理信息代理模型的构建与解读提供了实用指导。

*通讯作者。电子邮箱：n.becerra@upm.es (Nicolás Becerra-Zuniga), g.rubio@upm.es (Gonzalo Rubio)

关键词： 机器学习、物理信息神经网络（PINN）、数据-PDE 不一致性、伯格斯方程

1. 引言

随着机器学习技术近年来在物理科学中的兴起， [1],物理信息神经网络（PINN） [2, 3, 4] 已成为一种广受欢迎的监督学习框架 [5], ，这得益于其在数据稀缺的情况下构建偏微分方程（PDE）的数据驱动代理求解器的能力。该方法最初由 Lagaris 等人提出 [6] ，随后由 Karniadakis 等人将其扩展至深度神经网络 [5] [7], 。其核心思想是将控制规律——即神经网络给出的近似解满足偏微分方程的程度——直接作为（深度）神经网络损失函数中的一项。因此，PINN 参数的优化不仅由标准的数据学习原则驱动，也受到代理求解器试图近似的实际偏微分方程的引导。在损失中加入这样的“偏微分方程项”，相当于施加强归纳偏置，有效地将 PINN 的优化引导至其参数空间中可能带来更强泛化性能的区域：PINN 实际上被迫“学习物理规律”，因而不仅可能擅长插值，还可能外推至训练数据[8]中未包含的配置。因此，这种混合方法特别适用于训练数据稀缺、但控制方程已得到充分确立的科学与工程问题——例如，实验测量或数值模拟可能数量有限、获取成本高，或受到噪声污染。其应用涵盖流体动力学 [4], 、传热 [9]、电力系统 [10]、材料科学 [11], 、生理学 [12]和电磁学 [13], 等领域 [14]。

三个具有代表性的情景说明了物理信息神经网络在此类场景中的相关性。第一种情景是利用实验数据构建物理过程的代理求解器。实验数据可能对时间维度采样充分，但在空间和物理参数方面通常较为稀疏，而且往往受到测量噪声的影响。物理信息神经网络代理模型的目标是构建一个求解器，预测连续时间的时空场，使其既符合现有数据，也符合已知的底层物理规律。第二种情景出现在计算建模中，例如计算流体力学（CFD） [15], ，通常我们旨在针对某个固定几何形状和一组具体的物理参数（如雷诺数、马赫数等）求解纳维-斯托克斯方程（NS）。此时的任务是利用 CFD 仿真数据训练代理求解器。由于每次数值模拟本身都计算开销巨大，通常难以获取能密集覆盖各种几何形状和物理参数范围的 CFD 数据。在这种情况下，可以使用有限的高保真样本数据集训练物理信息神经网络，并通过强制要求网络中的物理变量遵循 NS 方程或某种近似模型 [16, 17] 来对训练过程进行正则化。成功训练后，物理信息神经网络将充当

快速代理求解器，使其能够在推理阶段针对广泛的条件近乎即时地进行预测。最后，第三种情景结合实验数据和数值模拟数据，构建一个（规模更大的）多保真度训练集，用于训练物理信息神经网络 [18]。

尽管物理信息神经网络前景广阔且已取得成功，但它们往往面临一个微妙却至关重要的挑战：确保物理规律（即其旨在求解的实际偏微分方程）与用于训练的数据来源之间的一致性。在大多数公式中，优化过程中会利用自动微分 [7]计算梯度，并据此更新物理信息神经网络的参数；计算梯度时涉及两类损失项（分别与数据残差和偏微分方程残差相关）。换言之，我们默认假设可用数据——包括用于训练的数据，以及用于测试物理信息神经网络在未见数据上性能的数据——严格遵循控制偏微分方程。然而，实验数据和数值模拟都无法准确反映潜在的物理现实。事实上，可用数据往往来自具有不确定性的异质来源，这些因素叠加后可能导致数据偏离上述假设。例如，实验数据原则上与偏微分方程一致，但会受到多种测量误差的影响。类似地，数值模拟数据虽然不存在测量误差，却可能受到算法问题的影响（从离散误差到特定偏微分方程积分算法选择所引入的偏差），从而导致数值解与偏微分方程的实际解不匹配 [19]。此外，数据与物理信息神经网络旨在求解的底层控制方程之间存在的这两类不一致性，在保真度或准确性方面也可能彼此不一致。总的来说，由此产生的差异可能以复杂的方式影响训练过程、降低收敛速度，并总体上削弱训练所得物理信息神经网络求解器的性能。

尽管已有多项研究展示了物理信息神经网络在数据—物理混合设置中的潜力， [7], 但对于数据对底层物理规律解析解公式的保真度如何影响物理信息神经网络可达到的精度，系统性分析仍然有限。特别是，训练数据的质量——无论是解析解、数值解还是实验数据——如何与精确施加的偏微分方程损失相互作用，以及这些不匹配在多大程度上会对模型的学习和后续性能造成内在限制，目前仍不清楚。在这项工作中，我们聚焦于这样一个事实：数据与物理规律之间的不一致性会导致基本的准确性界限，我们将其称为一致性障碍。

为弥补这一空白，本文对数据—偏微分方程一致性在物理信息神经网络训练中的作用开展了受控研究，并提出了一个统一框架，用于分析不同程度的数据保真度如何影响物理信息神经网络的收敛性和准确性。作为基准，我们考虑一维黏性伯格斯方程（一类具有可得解析解的典型非线性对流—扩散偏微分方程），并设计了一组训练实验，其中物理信息神经网络基于偏微分方程损失项和数据损失项进行训练，后者具有不同程度的保真度。这一设置使我们能够刻画一致性障碍，并

说明，即使在物理约束得到精确强制的情况下，物理信息神经网络所能达到的准确性仍受训练数据质量的限制。

本文其余部分安排如下。第 2 节概述 PINN 表述和误差分析所需的理论背景，为理解数据与物理不一致性如何影响 PINN 训练奠定基础。第 3 节介绍基准问题、数据生成和实验设置。第 4 节展示主要结果，分析数据与偏微分方程一致性对收敛和准确性的影响，并着重说明一致性障碍的出现。最后，第 5 节总结研究发现，并讨论其对实际物理信息神经网络应用的意义。

2. 预备知识

本节介绍理解限制物理信息神经网络（PINN）准确性的因素所需的理论背景。我们首先概述 PINN 的基本框架，包括符号约定、网络表示以及物理信息损失的表述。随后，我们说明现有训练数据中的误差如何通过损失函数中的残差传播，并引入偏差，进而影响网络可达到的准确性；由此自然引出一致性障碍这一概念，即由控制偏微分方程的保真度与训练所用数据之间的不匹配所导致的损失基本下限。

2.1. 物理信息神经网络基本框架

为便于说明，我们先考虑定义在空间域 $\Omega \subset \mathbb{R}$ 上的一维偏微分方程，

$$\frac{\partial u}{\partial t} + \mathcal{N}(u) = 0, \qquad x \in \Omega, \quad t \geq 0 \tag{1}$$

满足适当的边界条件，其中 $\mathcal{N}(\cdot)$ 表示一个（可能是非线性的）空间微分算子。在本研究中，我们聚焦于稳态解以简化分析，同时指出这一选择并不影响结果的普适性。在稳态下，$\frac{\partial u}{\partial t} = 0$，$u \equiv u(x)$ 表示（随空间变化的）场解。在物理信息神经网络框架中，真实解 $u(x)$由神经网络 $u_\theta(x)$近似，其中 θ 表示所有可训练参数（权重、偏差等）的集合。（深度）神经网络已被证明是通用函数逼近器，经过适当训练后，能够在整个定义域上对解进行连续表示。与标准神经网络不同，物理信息神经网络的训练过程通过在参数空间θ 上最小化复合损失函数 $\mathcal{L}(\theta)$来驱动，该函数由各项损失贡献（可能经过加权）求和构成：

$$\mathcal{L}(\theta) = \mathcal{L}_{\mathrm{PDE}}(\theta) + \mathcal{L}_{\mathrm{BC}}(\theta) + \mathcal{L}_{\mathrm{D}}(\theta), \tag{2}$$

其中， $\mathcal{L}_{\mathrm{PDE}}$ 包含微分算子作用于近似解 $u_\theta(x)$后在配点处得到的残差， $\mathcal{L}_{\mathrm{BC}}$ 表示 $u_\theta(x)$ 与边界条件处真实值之间的偏差，而 $\mathcal{L}_{\mathrm{D}}$ 则衡量模型预测与测量位置处训练数据之间的偏差。找到使 $\mathcal{L}(\theta)$最小化的参数 θ，即可得到解的近似；该近似不仅能较好地拟合测量位置处的解（并适用于插值场景），原则上也与控制方程相一致。这一表述为求解稳态偏微分方程提供了通用且灵活的框架，并可在此基础上针对具体问题构建相应的建模方案，并选择优化方法。

通常，边界条件和训练数据会归入 N_{data} 训练数据，并通过均方误差（MSE）来量化这些点处 $u_\theta(x)$与解 $u(x)$之间的不匹配程度，因此：

$$\mathcal{L}(\theta) = \frac{1}{N_{\mathrm{col}}} \sum_{i=1}^{N_{col}} [\mathcal{N}(u_\theta(x_i))]^2 + \frac{1}{N_{\mathrm{data}}} \sum_{j=1}^{N_{\mathrm{data}}} [u(x_j) - u_\theta(x_j)]^2, \tag{3}$$

其中， x_i 表示施加偏微分方程约束的 N_{col} 配点，而 x_j 表示有训练数据可用或施加了边界条件的 N_{data} 测量点。在任意迭代优化算法的每一步中，计算 $\mathcal{L}(\theta)$的第一项（尤其是算子 $\mathcal{N}(\cdot)$中出现的所有空间导数）都可借助神经网络中的自动微分来完成。

如前所述，我们讨论的是所谓的标准 物理信息神经网络，但请注意，我们也可以构建该问题的参数化 版本，在此版本中，微分算子显式依赖于一组额外参数 ν （例如物理参数）。这将得到一个更一般的偏微分方程，其稳态形式为

$$\mathcal{N}(u(x;\nu)) = 0, \qquad x \in \Omega, \nu \in \mathcal{P}, \tag{4}$$

其中， $\mathcal{P}$ 表示一个一般的参数空间。本文将研究标准物理信息神经网络和参数化物理信息神经网络。

2.2. 标签误差传播到损失函数中

如引言中所述，在 PINN 框架中使用的任意一组数据 $\{x, u(x;\ \nu)\}$ 都可能与偏微分方程给出的真实值存在不同程度的不一致性；对于数值数据，这是由离散化引起的偏差所致；对于实验数据，则是由测量误差所致。形式上，我们将 $u(x)$ x处偏微分方程的解析解记为 (true)，并将数据给出的解记为 $\tilde{u}(x)$，因此形式上

$$\varepsilon(x_k) := \tilde{u}(x_k) - u(x_k) \tag{5}$$

它量化了任意给定测量点 x_k处的数据标签残余不一致性。这些不一致性会传播到损失函数中，影响 PINN 在训练过程中达到的近似结果 $u_\theta(x)$。事实上，使用不一致数据 $\{x, \tilde{u}(x)\}$ 时得到的有效损失函数为

$$\mathcal{L}_{\text{eff}}(\theta) = \frac{1}{N_{\text{col}}} \sum_{i=1}^{N_{\text{col}}} [\mathcal{N}(u_\theta(x_i))]^2 + \frac{1}{N_{\text{data}}} \sum_{j=1}^{N_{\text{data}}} [\tilde{u}(x_j) - u_\theta(x_j)]^2. \tag{6}$$

利用方程 3 和 5，我们很容易得到

$$\mathcal{L}_{\text{eff}}(\theta) = \mathcal{L}(\theta) + \frac{1}{N_{\text{data}}} \sum_{j=1}^{N_{\text{data}}} 2\varepsilon(x_j)[u(x_j) - u_\theta(x_j)] + \frac{1}{N_{\text{data}}} \sum_{j=1}^{N_{\text{data}}} \varepsilon(x_j)^2, \tag{7}$$

其中，只有在所有数据不一致性都消除后，$\mathcal{L}_{\text{eff}}(\theta)$ 才等于 $\to \mathcal{L}(\theta)$，且 $\varepsilon(x_k)$ 等于 $\to 0\ \forall x_k$。$\mathcal{L}_{\text{eff}}(\theta)$ 中出现的附加项会影响 $u_\theta(x)$ 的优化，因此，若旨在最小化 $\mathcal{L}_{\text{eff}}(\theta)$ 而非 $\mathcal{L}(\theta)$，就会使学习过程偏离所有测量点处的真实目标 $u(x)$，进而影响全域解 $u_\theta(x)$ 的逼近。由于实践中难以量化和控制 $\varepsilon(x_k)$ 的表达式，由此产生的偏差可以用一致性障碍来理解。值得注意的是，当 ν 不受测量误差影响时（即使用数值数据时），同一个有效损失函数适用于参数化问题 $u(x;\nu)$；而使用实验数据时，ν 本身会受到不确定性的影响，并且有效损失中会出现附加项。

3. 方法

本节描述用于研究数据–偏微分方程一致性对物理信息神经网络（PINN）性能影响的实验框架。我们首先介绍本研究采用的基准问题，即具有制造解的黏性伯格斯方程，它为评估物理信息神经网络精度提供了平滑且可控的测试。随后，我们详细说明不同保真度水平数据集的生成，以及训练物理信息神经网络的四种相关一致性情景，以此作为受控方式，系统地调控偏微分方程残差与数据项之间的一致性，并量化数据保真度的影响。最后，我们给出标准配置与参数化配置所采用的物理信息神经网络架构、损失函数形式和训练设置。

3.1. 基准问题

我们将一维粘性Burgers方程作为基准问题，用于评估数据与偏微分方程一致性对物理信息神经网络性能的影响。该偏微分方程被广泛用作非线性输运和耗散过程的典型模型，

既是纳维–斯托克斯方程的简化极限，也是多个领域中的最简原型，包括湍流与能量级联研究、非线性声学、交通或网络流等。为了系统地解析控制其解和残差误差，我们采用制造解方法 [12], ，并在偏微分方程中加入强迫源项 $\mathcal{S}$ ，使其最终形式为

$$\frac{\partial u}{\partial t} + u\,\frac{\partial u}{\partial x} = \nu\,\frac{\partial^2 u}{\partial x^2} + \mathcal{S}(x,\nu), \qquad x \in [-1,1], \tag{8}$$

满足狄利克雷边界条件

$$u(-1;\nu) = u_{\mathrm{MMS}}(-1;\nu). \tag{9}$$

尽管黏性伯格斯方程在空间上是二阶的，但本文仅显式规定了流入边界条件，而在 $x = 1$ 处的出口边界条件保持自由，并由算子自然闭合，这有时被称为无操作边界条件 [20]。我们还将重点研究稳态解，即 $u \equiv u(x;\nu)$ 表示解场， ν 是黏度参数，而 $\mathcal{S}(x,\nu)$ 是强迫项，其定义满足人工解析解 $u(x;\nu) = u_{\mathrm{MMS}}(x;\ \nu)$，其中

$$u_{\mathrm{MMS}}(x;\nu) = \sin(2\pi x) + 0.5\,\log\left(\frac{1}{\nu}\right)\sin(6\pi x) + \log\left(\frac{1}{\nu}\right) + 2 \tag{10}$$

对于所有 ν 均严格满足控制偏微分方程（方程 8）。
这种基于制造解方法的构造具有三大主要优势。首先，对于任意黏度值，它都能产生一个光滑但非平凡的解，从而可以针对给定的 $u_\theta(x)$ 精确评估偏微分方程残差，同时避免与激波或陡峭梯度相关的复杂问题。其次，该解还表现出对黏度的非平凡（非线性）依赖，反映了现实中的参数敏感性，同时仍保持解析上的可处理性。最后，它使我们能够显式计算有限差分求解器和物理信息神经网络预测的离散误差。

如前所述，在本研究中，我们将分别分析标准物理信息神经网络和参数化物理信息神经网络。对于标准情形，我们将黏度参数固定为常数$\nu = 10^{-2}$。固定 ν 可对不同保真度的数据集进行受控比较。在这种情况下，物理信息神经网络唯一的输入特征是位置 x，唯一的输出特征则是速度 $u_\theta(x)$。相比之下，参数化物理信息神经网络（用于处理式 4）同时使用多个黏度值进行训练，这些黏度值在 $\nu \in [10^{-6}, 10^{-1}]$范围内按对数尺度采样。在这种情况下，黏度被视为额外的独立输入特征（即输入空间为二维），而学习到的解 $u_\theta(x;\nu)$使我们能够在空间和物理条件上进行插值，从而更深入地分析一致性效应。为了使各特征的数量级相近，实际输入物理信息神经网络的是经logMinMax缩放的黏度，而非实际黏度，其取值范围为 $[-1,1]$。

3.2. 所有一致性情景的数据生成

为评估数据保真度对物理信息神经网络性能的影响，我们使用有限差分求解器生成了三个数值数据集，其保真度各不相同（C1-C3）；此外，还构建了一个数据与式 10 的解析解完全一致的数据集（解析解情形，其数值误差可忽略不计，因此 $\varepsilon(x)$趋近于零）。这四个数据集定义了我们的一致性情景，后续将据此评估标准和参数化物理信息神经网络的性能。下文将详细说明这些数据集如何划分为训练数据和测试数据，并讨论生成PINN训练所需的配点和边界条件数据的方法。

数值数据生成– 对于数据集 C1-C3，通过对每个黏度值 ν 使用稳定的显式有限差分格式对方程 8 进行数值积分，得到相应的解 $\tilde{u}(x;\nu)$；黏度值在区间 $[10^{-6}, 10^{-1}]$ 内按对数取样。方程 8 中的扩散项采用二阶中心差分离散，而对流项则采用二阶迎风格式近似。由于伪时间推进仅用作迭代求解器，以求得稳态解 $\partial u/\partial t = 0$，且所有模拟均运行至离散残差收敛，因此最终的稳态数值解与具体选择的时间步长无关。因此，数据集 C1、C2 和 C3 中 $\tilde{u}(x;\nu)$ 的任何差异，都完全源于采用了不同的空间离散方法。

具体而言，数据集 C1 使用包含 $N^{\text{nodes}} = 81$ 个节点的粗空间网格生成，因此得到保真度较低的数值解，进而产生相对较大的离散误差 $\varepsilon(x)$。数据集 C2 使用更精细的空间网格（含 $N^{\text{nodes}} = 641$ 个节点），得到中等保真度的数值解；数据集 C3 则使用高度细化的空间网格（含 $N^{\text{nodes}} = 7121$ 个节点）生成，从而得到高保真度的数值解，并显著降低离散误差。

训练数据 – 对于标准物理信息神经网络（固定 ν），我们使用 $\nu = 10^{-2}$，并构建一个训练集，其中包含 4 个空间位置为 $\{x_j\}_{j=1}^{4} = \{-0.55, -0.3, 0.3, 0.8\}$ 的点（图1上图中的蓝点）。对于参数化物理信息神经网络，我们使用 $\nu \in \{10^{-1}, 10^{-3}, 10^{-6}\}$（经 logMinMax 重缩放），并针对每个黏度值定义 81 个空间位置，总计 243 个训练点（图1下图中的蓝点）。在 C1 中，这 81 个位置对应于数值积分所用的所有网格节点；而对于 C2 和 C3，我们选择在空间上最接近 C1 网格节点的网格节点（见图1）。这种受控的数据生成方法使我们能够量化与每个训练点相关的离散误差 $\varepsilon(x)$，并据此计算均方误差（类似于公式 7 中的最后一项），见表1。

测试数据与物理信息神经网络的双重评估 –对于三个数据集 C1-C3，测试数据由训练中未使用过的数值数据构成。对于标准物理信息神经网络（固定 ν），我们再次使用 $\nu = 10^{-2}$，并构建一个测试集，在以下位置设置 5 个点

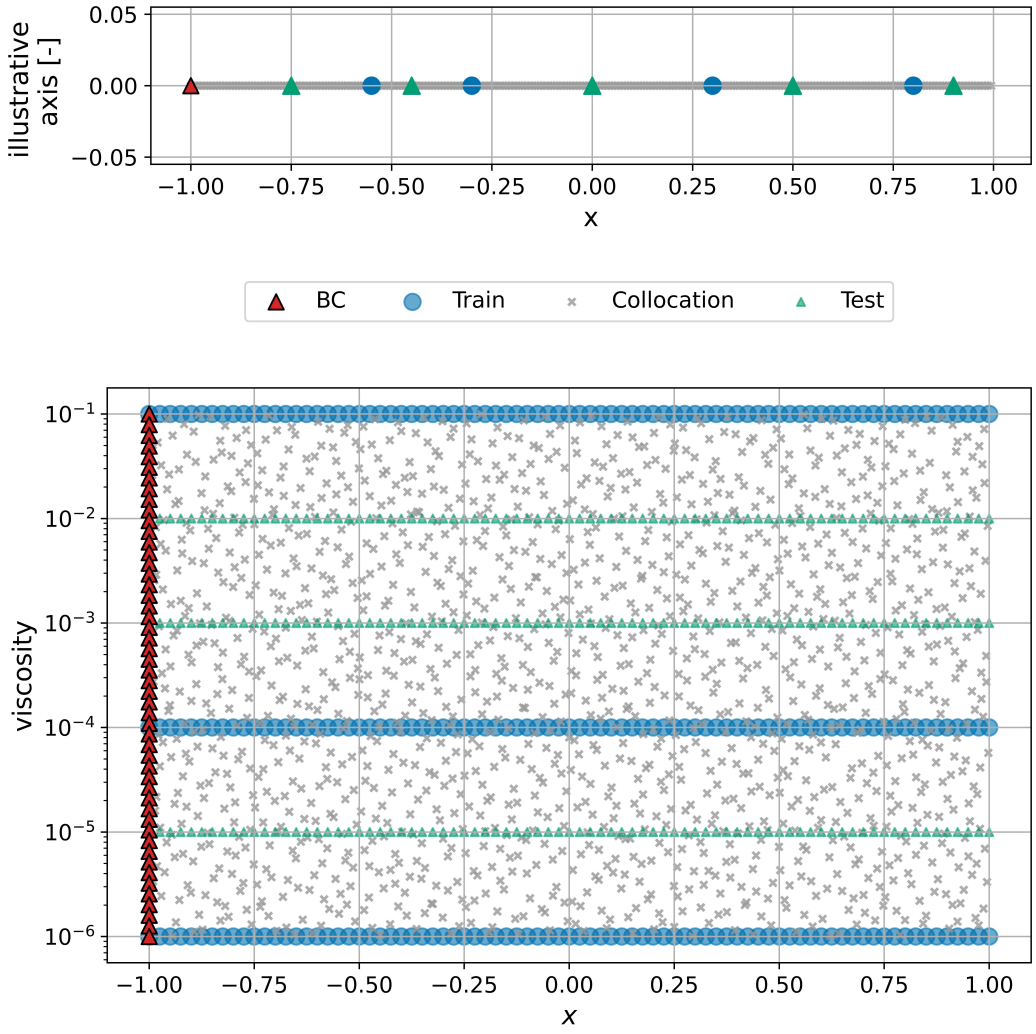


图1：训练和测试中使用的所有数据集点的分布。上图：标准物理信息神经网络。1000 个均匀分布的配点、4 个训练和测试数据点，以及 1 个边界条件点。下图：参数化物理信息神经网络，1500 个 Sobol 分布的配点、 3×81 个训练和测试数据点，以及 50 个均匀分布的边界条件点。

$\{x_j\}_{j=1}^4 = \{-0.75, -0.45, 0, 0.5, 0.9\}$（图1上图中的绿色三角形）。对于参数化物理信息神经网络，我们评估其在 ν 中进行插值的能力，因此使用 $\nu \in \{10^{-2}, 10^{-4}, 10^{-5}\}$（对数最小最大缩放）；与训练数据一样，每个黏度取 81 个空间位置，共计 243 个测试点（图1下图中的绿色三角形）。

那么，训练完成后，如何评估每个物理信息神经网络的性能？我们有两种选择：

将其预测结果与（数值）测试数据所提供的解（数据本身也不一致）进行比较，或与解析解（完全一致，但与测试数据不符）进行比较。每种评估方式都会揭示不同的见解；正如我们将在结果部分所说明的那样，本研究采用这两种方式进行双重评估，同时使用逐点全局指标（例如均方根误差）和更细致的指标。

配点 – 这是计算 $\mathcal{L}_{\mathrm{PDE}}$ 的空间位置。在标准物理信息神经网络配置中，使用 1000 个等距配点，并结合空间坐标（图1上图中的灰色叉号）。在参数化物理信息神经网络中，使用 Sobol序列在增广输入空间中抽取 1500 个配点，并对黏度维度进行对数缩放（图1下图中的灰色叉号）。

情景	数据来源	$\frac{1}{N_{\mathrm{training}}}\sum_j \varepsilon(x_j)^2$	一致性
C1	数值计算——粗网格	8.4×10^{-2}	Low
C2	数值计算——中网格	2.8×10^{-6}	中等
C3	数值计算——细网格	1.4×10^{-10}	High
解析解	解析解	0	精确

表1：标准物理信息神经网络情形下四种一致性情景（C1-C3 和解析解情景）的总结，以及数据一致性误差的量化结果。C1-C3 中的数据通过使用逐渐加密的计算网格对偏微分方程进行数值积分生成，从而得到逐渐减小的离散误差，并因此使数值解与真实解析解之间的不一致性逐渐减小，这一点通过均方意义进行了量化。

边界条件点 –这些点与 $x = -1$ 处的计算域端点重合。在标准物理信息神经网络中，这只有一个点。在参数化物理信息神经网络中，不同黏度下的边界条件始终位于 $x = -1$ 。我们在对数尺度上采样总计 50 个点，黏度范围为 10^{-6} 至 10^{-1}，从而在重缩放后的黏度范围 $[-1, 1]$ 内均匀采样出 50 个黏度值（图1中的红色三角形）。这些位置处的解根据式9直接由解析解计算得到，即在我们的研究中，边界条件不存在不一致性。

总体而言，这一设置能够在 C1-C3 情景下受控地降低数据保真度，从而提供一个系统性框架，用于研究离散化引起的数据与偏微分方程不一致性如何通过物理信息神经网络损失函数传播，并促成一致性障碍的出现。表1总结了数据来源、一致性情景及其相关的离散误差。

3.3. 物理信息神经网络架构与优化的详细说明

在本小节中，我们详细介绍标准型和参数化物理信息神经网络训练的各个方面，包括其架构、优化过程中使用的损失函数，以及

优化器。

3.3.1. 网络架构与损失函数

为简化研究，标准物理信息神经网络和参数化物理信息神经网络采用相同的网络架构和损失函数配置。我们使用一个全连接深度前馈神经网络，其中包含七个隐藏层，每层有 200 个神经元，并采用双曲正切作为激活函数。标准物理信息神经网络的输入为一维（x），参数化情形下的输入为二维（x，ν），从而使网络能够在不同黏度之间进行插值；输出始终为一维场 u_θ。

如前所述，在物理信息神经网络的优化过程中，可以为各个不同的损失项赋予权重，以便在训练过程中强调其重要性。这一点尤其重要，因为数据损失（边界条件损失 $\mathcal{L}_{\mathrm{BC}}$ 和其余训练数据损失 $\mathcal{L}_{\mathrm{D}}$）衡量的是解之间的简单不匹配（通常采用均方误差），而偏微分方程损失 $\mathcal{L}_{\mathrm{PDE}}$ 衡量的是解的行为和性质的残差，也就是说，它不仅考虑解本身，还考虑其导数。从这个意义上说，对于快速空间振荡的解，即使只有很小的不匹配，也可能导致较大的偏微分方程残差，因为场的空间导数所产生的影响不成比例地大于场本身的简单残差。简言之，$\mathcal{L}_{\mathrm{PDE}}$ 与 $\mathcal{L}_{\mathrm{BC}}$和 $\mathcal{L}_{\mathrm{D}}$本质上具有“不同的单位”，因此有必要对损失项进行适当加权，以实现平衡。一种初步方案是采用标准的固定权重损失组合，即

$$\mathcal{L}(\theta) = \alpha\,\mathcal{L}_{\mathrm{PDE}} + (1-\alpha)(\mathcal{L}_{\mathrm{D}} + \mathcal{L}_{\mathrm{BC}}), \tag{11}$$

其中，$\mathcal{L}_{\mathrm{PDE}}$ 等于公式 6 右端的第一项，而 $\mathcal{L}_{\mathrm{BC}}$ 和 $\mathcal{L}_{\mathrm{D}}$的形状均与公式 6 右端的第二项相同。可手动调整超参数 α，以探索物理规律遵循度与数据保真度之间的权衡。在结果部分——具体而言，在第 4.1.1 节——我们考察此类超参数 α 的不同取值如何影响（$\mathcal{L}_{\mathrm{PDE}}$，$\mathcal{L}_{\mathrm{D}}$）平面中的训练轨迹，从而估算帕累托前沿；这一分析遵循 Rohrhofer 等人 [21]的方法。通过这项探索性分析可以发现，α 的最优值很大程度上取决于具体问题，因此无法预先确定。因此，从此处开始以及后续分析中，我们采用一种称为 lbPINN 的动态加权策略。该方法受多任务学习 [22, 23], 启发，其中用于加权各损失项的前置系数集合 $\sigma = \{\sigma_{\mathrm{PDE}}, \sigma_{\mathrm{BC}}, \sigma_{\mathrm{D}}\}$ 本身是一组可训练参数，因此该集合也会在训练过程中动态更新。具体而言，我们使用以下全局损失函数，而不是公式 11

$$\mathcal{L}(\theta,\sigma) = \frac{1}{2\sigma_{\mathrm{PDE}}^2}\mathcal{L}_{\mathrm{PDE}} + \frac{1}{2\sigma_{\mathrm{BC}}^2}\mathcal{L}_{\mathrm{BC}} + \frac{1}{2\sigma_{\mathrm{D}}^2}\mathcal{L}_{\mathrm{D}} + \sum_{i\in\{\mathrm{PDE,BC,D}\}} \log\sigma_i. \tag{12}$$

这种自适应加权策略使网络能够在训练过程中动态调整各损失分量的贡献，从而促进不同保真度数据集的稳健收敛

，并确保没有任何单个项主导优化过程。再次需要指出，偏微分方程损失项衡量的并非解本身之间的差异，而是它们的属性（例如导数）之间的差异，因此其量级不一定与另外两项相同。在第 4.1.1 节中，我们比较了通过式 11 获得的训练轨迹与使用式 12 得到的训练轨迹，结论是后者确实能够达到帕累托前沿，而无需预先确定最佳权重组合。因此，后续分析将重点讨论通过式 12 训练物理信息神经网络。

3.3.2 优化算法

虽然深度监督学习中的多层感知机传统上使用（随机）梯度下降的某种扩展方法进行优化（例如 Adam 优化），但在物理信息神经网络的优化过程中，通常会在训练期间切换优化策略。本文采用这种混合策略。具体而言，首先使用 $AdamW$ 以实现稳定的全局收敛，随后使用 $L\text{-}BFGS$ 对解进行更平滑的调整。这种组合兼具一阶方法（$AdamW$）的稳定性和拟牛顿法（$L\text{-}BFGS$）的精确性。表 2 详细列出了每种优化器的训练轮次数。请注意，每个训练轮次内执行的迭代次数因优化器而异。对于 $AdamW$ 阶段，标准情形和参数化情形分别使用大小为 500 和 120 的小批量，因此每个训练轮次分别包含 2 次和 13 次内部迭代。相比之下，L-BFGS 是全批量优化器，每次迭代即为一个训练轮次。因此，我们分别将训练轮次数设为 3000 和 160000；此外，出于在收敛曲线（图 3 和图 5）中进行可视化的目的，对于 L-BFGS，我们分别每 60 次迭代（标准物理信息神经网络）和每 400 次迭代（参数化物理信息神经网络）绘制一次结果。因此，第 4.1.2 节和第 4.2.1 节中展示的收敛曲线以优化迭代次数而非训练轮次为单位进行报告。

此外，在启用这种混合优化之前，还会先对网络进行预热。具体来说，首先像训练基础多层感知机一样，对物理信息神经网络进行少量训练轮次的训练，也就是通过设置 $1/\sigma_{\mathrm{PDE}}^2 = 0$，有效关闭偏微分方程损失项；我们固定 $\sigma_{\mathrm{BC}} = \sigma_{\mathrm{D}} = 1$，只运行 $AdamW$ 几个训练轮次。对于标准物理信息神经网络，我们在这一初始预热过程中仅使用边界条件和训练数据。然而，对于参数化物理信息神经网络，我们首先需要足够数量的训练数据点，以便充分采样整个黏度范围。由于训练过程中不能使用测试数据，我们通过在$\nu \in \{10^{-2}, 10^{-4}, 10^{-5}\}$ 处加入虚假（错误）解来扩充训练数据，从而避免数据泄漏；这些解由 $\nu \in \{10^{-1}, 10^{-3}, 10^{-6}\}$处的解给出。

4. 结果

本节展示标准物理信息神经网络和参数化物理信息神经网络基准问题的数值结果。实验旨在系统评估数据保真度的变化如何

Case	预热	AdamW	L-BFGS
标准物理信息神经网络	50	2000	3000
参数化物理信息神经网络	500	20000	160000

表 2：标准物理信息神经网络和参数化物理信息神经网络各优化阶段的训练轮次数。AdamW 是一种小批量优化器，因此标准情形和参数化情形每个训练轮次分别进行 2 次和 13 次小批量更新。相比之下，L-BFGS 在全批量模式下运行，每个优化步骤都使用整个训练集进行计算。

影响物理信息神经网络解的准确性、稳定性和收敛性。我们结合第 2 节提出的一致性障碍来解读观察到的趋势。该障碍为数据保真度与控制偏微分方程不匹配时可达到的损失设定了下界。

4.1. 标准物理信息神经网络情形

我们从标准物理信息神经网络情形入手分析，其中黏度固定。作为初步探索，我们在第 4.1.1 节比较通过公式 11 和 12 进行的训练。将公式 11 的最小化解释为标量化的多目标优化后，我们构建其帕累托前沿，并表明通过公式 12 进行的训练无需预先找到最优权重，便可达到该前沿。因此，在第 4.1.2 节和第 4.1.3 节中，我们使用公式 12 中的动态加权损失，并评估该物理信息神经网络的收敛曲线和性能，考察不同数据一致性情景下的表现。

4.1.1. 损失函数中物理规律项与数据项的最优平衡：固定权重与 *lbPINN* 比较

在任何多目标优化（MOO）问题中，都存在一种权衡解——即所谓的帕累托最优解——在不使至少一个其他目标变差的前提下，无法改善其中任何一个目标。所谓的帕累托前沿，就是所有帕累托最优解的集合。

将物理信息神经网络优化视为一个多目标优化问题颇具启发性，其中一个目标是最小化偏微分方程残差 $\mathcal{L}_{\mathrm{PDE}}$，另一个目标是最小化数据相关损失$\mathcal{L}_{\mathrm{D}}$ 和 $\mathcal{L}_{\mathrm{BC}}$。通过全局损失函数（如式 11）合并这两个问题，会将原始的多目标优化问题转化为标量化问题。在这种情况下，对于固定的权衡权重α，我们执行优化（采用上文讨论的预热 $\rightarrow$ AdamW $\rightarrow$ L-BFGS 方法），并收集算法在所有优化步骤中得到的点对（$\mathcal{L}_{\mathrm{PDE}}$，$\mathcal{L}_{\mathrm{D}}$）。按照优化领域的术语，这构成了可行解集。对于该权衡权重 α，使式 11 中全局损失函数最小化的点（$\mathcal{L}_{\mathrm{PDE}}$，$\mathcal{L}_{\mathrm{D}}$）是标量化问题的帕累托最优解。因此，通过改变 α 找到的帕累托最优解集合构成帕累托前沿。

图 2 展示了在使用由以下公式定义的损失函数对物理信息神经网络进行优化的过程中，对于不同的 α （以颜色区分）所找到的所有可行解（$\mathcal{L}_{\mathrm{PDE}}$，$\mathcal{L}_{\mathrm{D}}$）

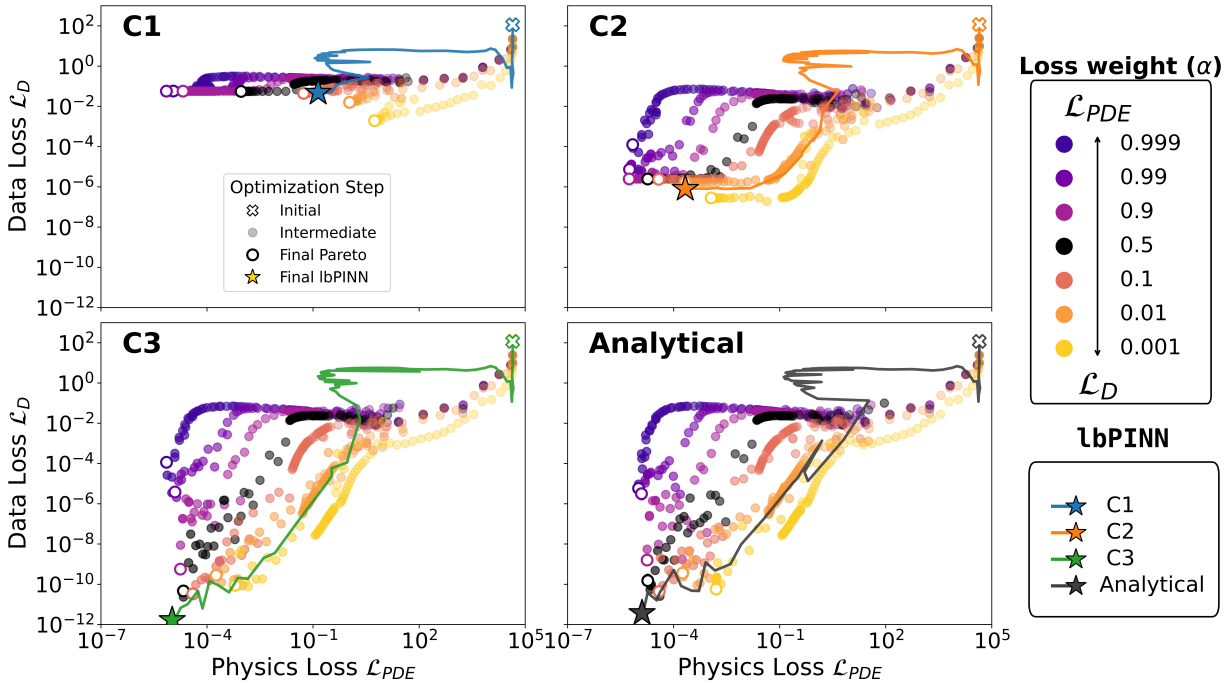


图 2：将标准物理信息神经网络通过最小化式 11 进行优化解释为一个（标量化的）多目标优化问题。我们在（$\mathcal{L}_{\mathrm{PDE}}$，$\mathcal{L}_{\mathrm{D}}$）平面中展示了优化过程中针对每个 α 值（以颜色区分）所达到的所有可行解。每次优化的最终值（以白色圆点突出显示）表示各权衡权重 α 对应的帕累托最优解，汇总起来近似构成帕累托前沿。每个面板展示了四种一致性情景（C1-C3 和解析解）下获得的这些信息。此外，每个面板还展示了通过 lbPINN 方法最小化式 12 得到的标准物理信息神经网络优化轨迹（实线）。该优化的最终解以相应颜色的星形标记表示。总体而言，lbPINN 得到的解与固定权重损失式 11 估计的帕累托前沿相容。

式 11。对于每个 α，帕累托最优解以白点突出显示，由此得到的曲线近似于帕累托前沿。该图的每个面板都针对前文定义的四种一致性情景分别重复了这一分析。当 α 减小时（因此偏微分方程损失项相对于数据损失项的重要性降低），通常优化会达到更小的 $\mathcal{L}_{\mathrm{D}}$ 可行值，这与预期一致。此外，训练数据的一致性水平会显著影响这一行为：保真度较高的数据集能够达到远小于低保真数据集的 $\mathcal{L}_{\mathrm{D}}$ 值。例如，注意到 $\mathcal{L}_{\mathrm{D}}$ 在 C3 中可达到约 10^{-11} 的量级，而在 C1 中仅能达到约 10^{-3} 的量级（且代价是 $\mathcal{L}_{\mathrm{PDE}}$ 显著增大）。直观而言，数据与偏微分方程的一致性越高，最小化 $\mathcal{L}_{\mathrm{PDE}}$ 与最小化 $\mathcal{L}_{\mathrm{D}}$ 之间的张力就越小（这两种最小化会更加一致）。

就帕累托前沿而言，可以看到在 C1 中，前沿大致在 $\mathcal{L}_D$ 轴上趋于饱和，

这清楚地表明数据是不一致的；由于偏微分方程与数据项之间存在明显张力，即使增大其相对权重，也无法降低数据拟合误差，这符合预期。换言之，对于 C1，偏微分方程的最小化与数据损失的最小化几乎不相容。C2、C3 和解析解的前沿呈现出经典的肘形——这表明偏微分方程与数据之间具有相容性。前沿拐角的位置给出了单位数据退化所对应的最大物理规律改进（反之亦然），其对应的 α 值则体现了一种权衡最优性的概念。还可以看到，C3 数据集（高分辨率数值数据）与解析参考解的帕累托前沿非常相似，这证明我们并不需要 $\varepsilon(x)$ 对所有 x 都精确为零，数据不一致性才会真正变得重要。

作为比较，我们还在同一张图中绘制了使用式（12）定义的 lbPINN 损失函数时得到的优化轨迹（实线）。尽管这些优化轨迹不同于任何固定 α 取值下对应的轨迹（这是预期的，因为 lbPINN 明确允许权衡权重在训练过程中动态变化），但 lbPINN 最终得到的解（图中以星号标出）位于式（11）估计的帕累托前沿上，甚至略优于该前沿。这一发现促使我们在本文其余部分采用 lbPINN 作为优化策略。

4.1.2. 收敛曲线

从现在起，我们使用第3.3.1节所述的动态加权损失来训练标准物理信息神经网络。图3以半对数图展示了测试均方根误差随训练迭代次数的变化情况。该指标衡量物理信息神经网络预测与测试集所提供的数值解之间的均方根误差

$$\texttt{testRMSE} = \sqrt{\frac{1}{N_{\text{test}}} \sum_{k=1}^{N_{\text{test}}} [\tilde{u}(x_k) - u_\theta(x_k)]^2}, \tag{13}$$

，涵盖所有四种一致性情景。图中标出了 AdamW 与 L-BFGS 优化器之间的切换。可以观察到，随着数据一致性的提高，测试均方根误差始终显著降低；C3 中的测试均方根误差基本与使用解析数据训练时达到的水平相同，这证实了物理信息神经网络正在学习，并且随着数据不一致性降低，这种学习会收敛至真实的物理规律。事实上，低保真数据集 C1–C2 的测试均方根误差会停滞在一定水平（请注意，该图为半对数图）；而对于 C3 和解析数据，切换至 L-BFGS 优化器则能进一步提升性能。

4.1.3. 将物理信息神经网络预测和数值数据与解析解进行比较

在上一节中，我们使用测试均方根误差来计算物理信息神经网络预测与测试数据中的（数值）解的匹配程度，即 $u_\theta(x)$ 与 $\tilde{u}(x)$ 的匹配程度，而后者

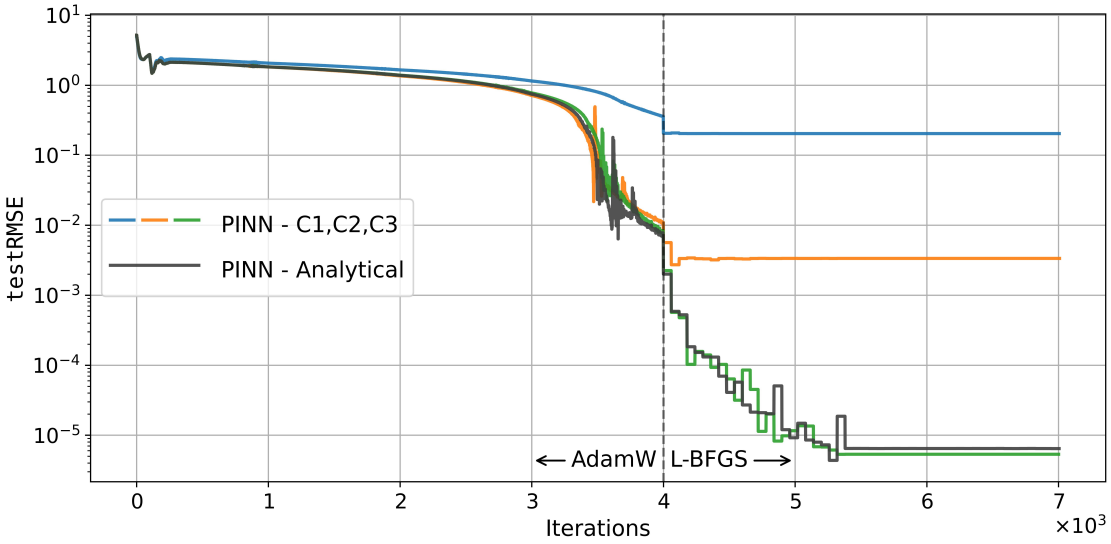


图3：对于在四种一致性情景下训练的标准物理信息神经网络，测试均方根误差（公式 13，用于衡量物理信息神经网络的预测与数值数据解之间的误差）随训练迭代次数的变化。图中突出显示了从 AdamW 优化器切换到 L-BFGS 优化器的过程。

本身会根据所分析的一致性情景呈现不同程度的不一致性。接下来，我们针对四种一致性情景，研究物理信息神经网络预测 $u_\theta(x)$ 与 $\tilde{u}(x)$ 的偏离程度，以及它最终与真实解析解 $u(x)$ 的匹配程度。为此，我们首先在图 4 中绘制预测解与解析解之间的空间分布绝对误差 $|u_\theta(x) - u(x)|$（实线）。图中采用半对数刻度。在高度一致的数据集（解析解和 C3）上训练的物理信息神经网络，在整个区域内都保持较低的绝对误差，仅出现轻微的高频局部振荡。相比之下，在低分辨率数据集（C1-C2）上训练的物理信息神经网络则表现出更大的偏离和更高的绝对误差。

作为补充，在同一张图中，我们还绘出了（虚线）数值解与真实解之间随空间分布的绝对误差，$|\tilde{u}(x_k) - u(x_k)|$；其中，位置 $x_k = x_1, x_2, \ldots, x_{N^{\text{nodes}}}$ 对应于第 3.2 节所定义的各一致性情景在积分方案中使用的计算网格节点，即$N^{\text{nodes}} = 81, 641, 7121$，分别对应 C1、C2 和 C3（对于解析解情景，我们也使用 C3 网格节点的位置）。比较实线和虚线可知，物理信息神经网络的预测结果系统性地比真实解 $u(x)$更接近基于其构建数据损失项的（不一致）数值数据 $\tilde{u}(x)$。这表明偏微分方程损失项使物理信息神经网络能够学习物理规律。同时，对于每种一致性情景，$u_\theta(x)$相较于 $\tilde{u}(x)$的改进并不显著。这凸显了一致性障碍：尽管

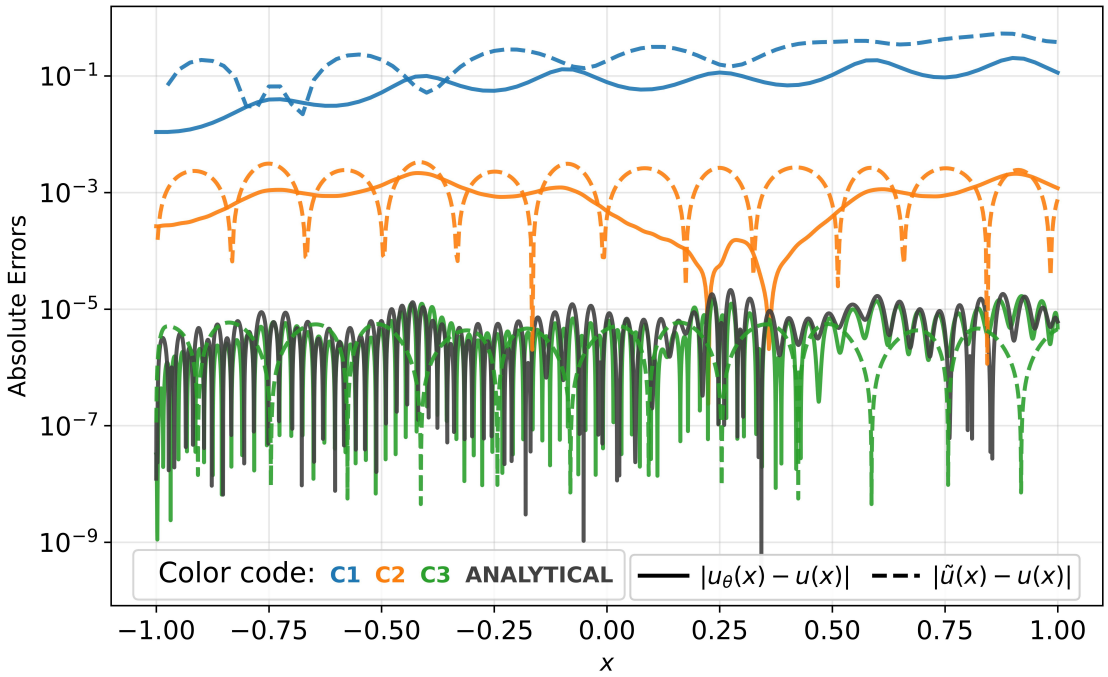


图4：对于在四种不同一致性情景下训练的标准物理信息神经网络，物理信息神经网络预测与解析解之间随空间变化的绝对误差 $|u_\theta(x) - u(x)|$ （实线）。作为比较，图中还以虚线表示数值数据随空间变化的绝对误差 $|\tilde{u}(x) - u(x)|$ 。

物理信息神经网络通过施加偏微分方程损失来帮助发现真实的 $u(x)$，但其能力最终受限于（不一致的）数据损失。

最后，我们构建了一个全局误差度量，通过计算 $u_\theta(x_k)$ 与 $u(x_k)$ 之间的均方根误差（RMSE），评估预测 $u_\theta(x)$ 与实际解析解 $u(x)$ 的差异；与前文相同，x_k 遍历用于积分求解数值解的计算网格节点位置（对于解析情景，我们使用 C3网格节点的位置）。为完整起见，我们还报告了相同网格节点处数值解 $\tilde{u}(x_k)$ 与解析解 $u(x_k)$ 之间的均方根误差。结果汇总于表3。

对于一致性最低的情景 C1 和一致性居中的情景 C2，物理信息神经网络的全局预测误差（相对于真实解）仅略小于数据因数值不一致性导致的实际误差。对于高分辨率数据（C3），物理信息神经网络的输出与解析参考解几乎无法区分。

总体而言，标准物理信息神经网络分析——包括帕累托前沿、收敛曲线和误差量化——证实，物理信息神经网络预测的准确性高度依赖训练

情景	$\sqrt{\frac{1}{N^{\text{nodes}}}\sum_{k=1}^{N^{\text{nodes}}}[\tilde{u}(x_k)-u(x_k)]^2}$	$\sqrt{\frac{1}{N^{\text{nodes}}}\sum_{k=1}^{N^{\text{nodes}}}[u_\theta(x_k)-u(x_k)]^2}$
C1	2.81e-01	9.65e-02
C2	1.89e-03	1.02e-03
C3	3.51e-06	5.34e-06
解析解	-	6.61e-06

表3：对于全部四种一致性情景，比较（i）数值（不一致）解与解析解之间的均方根误差（在 N^{nodes}集合上计算），以及（ii）物理信息神经网络预测与解析解之间的均方根误差。

数据的保真度，无论是与保真度相同的测试数据进行比较，还是与解析解进行比较，结果都是如此。高分辨率数据集（C3 和解析解）始终能获得最低误差和最稳定的预测，而低保真度数据则会导致更高的误差和更显著的空间变化。这些结果为数据分辨率的影响确立了明确的基准，并将指导对第 4.2 节后续参数化基准实验的解读。

4.2. 参数化物理信息神经网络案例

为将对一致性障碍的分析扩展到标准物理信息神经网络的设定之外，我们考虑一个具有两个输入特征（x 和 ν）的参数化物理信息神经网络。我们使用制造解，复现标准物理信息神经网络的实验框架，按照公式 12 训练物理信息神经网络，其中数据损失基于不同保真度（解析解、C1–C3）的数据集构建，同时保持物理残差精确。该参数化设置使我们能够系统评估数据质量如何影响模型在更复杂学习任务中的准确性；在此任务中，物理信息神经网络在物理条件（ν）上的插值能力尤其具有挑战性。

4.2.1. 收敛曲线

图 5 展示了参数化物理信息神经网络中，测试均方根误差（式 13）的变化情况。该误差衡量测试集所有点上物理信息神经网络预测 $u_\theta(x;\,\nu)$ 与（不一致的）测试数据解 $\tilde{u}(x;\,\nu)$ 之间的偏差，并以优化迭代次数为横轴。此处得到的收敛曲线在定性上与图 3 所示的标准物理信息神经网络情况相似：在存在不一致性的数值数据集（C1–C3）上训练的模型，其测试均方根误差依次趋于较低的平台值。请注意，C3 的测试均方根误差平台值略低于解析解情形；这是合理的，因为测试均方根误差衡量的是预测与测试集中的（不一致）数值解之间的偏差，因此，对于更接近解析解 $u(x;\nu)$ 的预测 $u_\theta(x;\nu)$，这种偏差预计会更大。

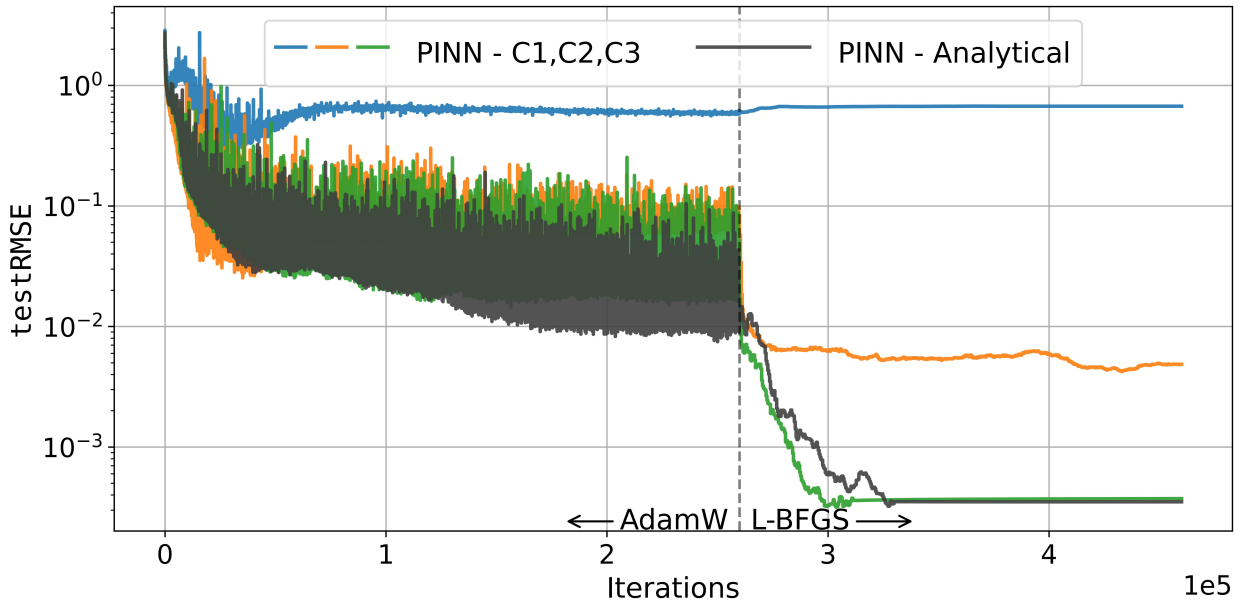


图 5：所有四种一致性情景下，参数化物理信息神经网络的测试均方根误差（公式 13）随训练迭代次数的变化。

优化器切换也很明显：初始阶段使用 AdamW，随后进入由 L-BFGS 驱动的平滑阶段。第二阶段中，C1、C3 和解析解情形均实现了完全收敛；但请注意，C2 的测试均方根误差曲线在经过 10^5 次迭代后仍有轻微波动。这种情况的一种可能原因是，对于 C2，数据损失项与偏微分方程损失项的量级完全相同，而最小化这两个损失项之间的矛盾尚未完全消除；不过，这一点可能还需要进一步分析。

4.2.2. 将物理信息神经网络预测和数值数据分别与解析解进行比较

与标准物理信息神经网络的情况类似，我们现在不再根据测试集中的数值解，而是根据解析解来评估参数化物理信息神经网络的性能。为了深入了解参数化物理信息神经网络在不同物理条件之间进行插值的能力，我们首先考察参数化物理信息神经网络在四种一致性情景下分别训练时，物理信息神经网络预测值 $u_\theta(x;\nu)$ 与解析解 $u(x;\nu)$ 之间的均方根误差（与标准物理信息神经网络的情况一样，对不同计算网格所确定的空间位置集合取平均）。图 6 以双对数尺度绘出了该均方根误差随 ν 的变化。此外，与图 4 中的做法类似，我们用空心圆点绘出 $\nu \in \{10^{-6}, 10^{-5}, 10^{-4}, 10^{-3}, 10^{-2}, 10^{-1}\}$ 对应的数值解 $\tilde{u}(x_k;\nu)$ 与真实解 $u(x_k;\nu)$ 之间的均方根误差，该误差是在相同位置上计算的

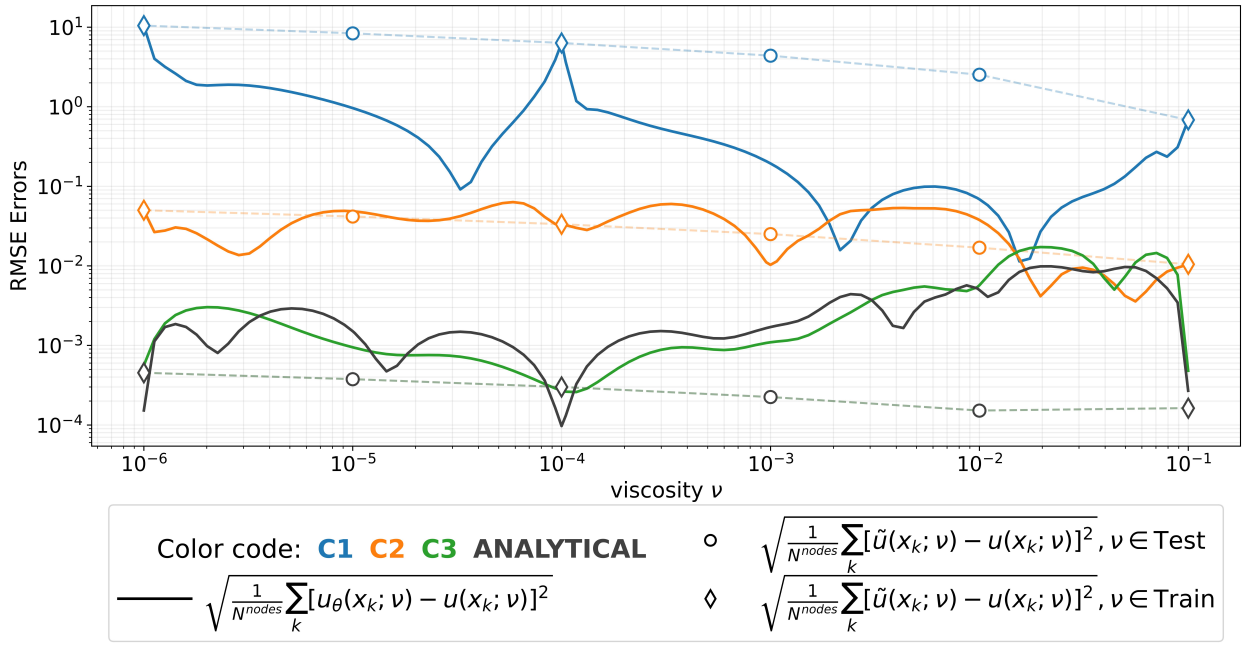


图 6：在不同黏度范围内，四种不同一致性情景（C1–C3、解析解）下，参数化物理信息神经网络预测与解析解之间的均方根误差（按空间位置计算；实线）。空心点表示六种黏度下数值解与解析解之间的均方根误差（在计算网格的所有空间位置上计算） $\nu \in \{10^{-6}, 10^{-5}, 10^{-4}, 10^{-3}, 10^{-2}, 10^{-1}\}$（虚线仅作视觉引导）。

$x_k = x_1,\ x_2,\ \ldots\ldots,\ x_{N^{\text{nodes}}}$ 比上一节更大（虚线为分段线性插值器，仅用于可视化）。

比较每种一致性情景下的实线和虚线，可以看出参数化物理信息神经网络插值性能如何受数据保真度影响。对于低一致性情景 C1，物理信息神经网络的性能明显优于数值数据；而在中等一致性情景 C2 下，预测结果与数值计算结果相当。高一致性情景（C3 和解析解）下，误差差异极小，与标准物理信息神经网络结果中观察到的饱和现象相似。

还需注意，在高黏度下，由于采用对数采样，配点数量减少，这限制了网络进一步降低误差的能力，导致其性能与情景 C2 相近。这些观察结果揭示了训练数据保真度、配置点分布与可达到的物理信息神经网络精度（相对于真实解析解）之间存在系统性关系。

图 7 还以箱线图的形式展示了黏度条件下的均方根误差分布，供参考（箱线图显示中位数、四分位间距、须线位于 ± 1.5 IQR 处，以及须线之外的离群值——

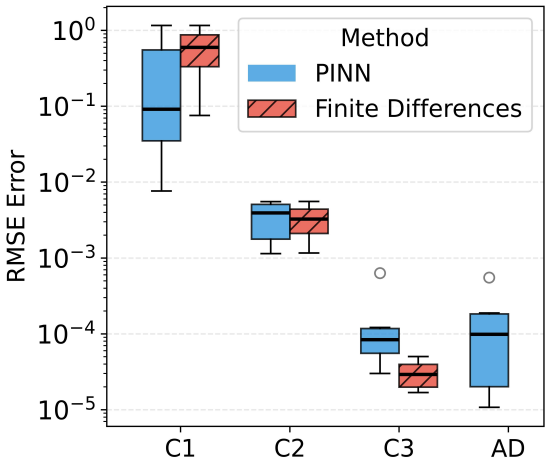


图7：（蓝色）对于每种一致性情景，参数化物理信息神经网络预测与解析解在计算网格节点处的黏度条件化均方根误差箱线图。（红色）对于每种一致性情景，数值解与解析解在计算网格节点处的黏度条件化均方根误差箱线图。

）。这很好地概括了前述观察结果：对于低一致性情景数据 C1，物理信息神经网络显著地学会了物理规律，因此偏离了有偏数值解；总体而言，与数值解相比，其相对于解析解的误差降低了约一个数量级。在 C2 中，物理信息神经网络的准确性与数值解相当（测试中位数 p-值 ≈ 1）。高保真情景（C3 和解析解）下，物理信息神经网络的准确性相近，但误差略大于数值解，这是因为在不同黏度间进行插值时，难以系统性地生成良好的预测。

总体而言，这些定量结果表明，物理信息神经网络能够有效改进基于低保真数据的预测，但其最终性能受到训练数据保真度和网络架构固有近似能力上限的共同制约。

5. 结论

物理信息神经网络被提出作为一种灵活的框架，通过将数据与嵌入其中的物理约束相结合，构建偏微分方程的替代解。本文表明，该方法的有效性从根本上取决于训练数据与控制方程之间的一致程度。具体而言，我们发现，当这种一致性不满足时，会出现一种一致性障碍，它构成了物理信息神经网络解所能达到的准确性下限。

以具有人工解析解的一维粘性Burgers方程

作为受控基准，我们系统地分析了数据保真度对PINN训练的影响。基于误差传播的理论分析解释了数据中的不一致性如何使数据–物理复合损失的优化产生偏差，令网络趋向于折中解而非精确的偏微分方程解，即使物理残差得到了精确约束也是如此。数值实验验证了这一观察：对于低保真度和中保真度数据集，物理信息神经网络表现出明显的误差饱和，其平台水平与数据中存在的离散误差密切相关。虽然基于物理的正则化能够部分修正低保真数据并恢复主要解结构，但所能实现的改进仍然有限。相较之下，使用高保真度数值数据或完全一致的解析数据进行训练时，得到的物理信息神经网络解会收敛至解析参考解，这表明一旦数据误差低于足够小的阈值，一致性障碍便可有效消除。标准和参数化物理信息神经网络形式均呈现类似趋势。对数据损失和物理损失进行帕累托前沿分析，进一步揭示了存在不一致数据时这两个目标之间固有的矛盾。随着数据保真度提高，这种矛盾逐渐减弱，帕累托前沿也向两种损失均可达到的更低值移动。

总体而言，这些发现阐明了数据质量在物理信息学习中的作用，并解释了在实际物理信息神经网络应用中常见的性能饱和现象。从实践角度来看，我们的结果强调，提高数据保真度或明确考虑数据误差，可能与网络架构或优化策略的进步同样重要。一致性障碍这一概念为解读物理信息神经网络的结果提供了有用的视角，并推动未来研究致力于减轻数据–偏微分方程不一致性的影响。未来工作应将这一分析扩展到更高维问题（例如纳维–斯托克斯方程）和含噪实验数据集，同时考虑物理参数和几何形状的一致性。另一个有前景的方向是开发自适应或不确定性感知型物理信息神经网络形式，将数据保真度明确纳入损失函数，以期减轻一致性障碍的影响。

致谢 – 作者感谢由 MCIN/AEI/10.13039/501100011033、西班牙资助的 TIFON 项目（PLEC2023- 010251）提供的资金支持。本研究获得了欧盟（ROSAS，项目编号 101138319）的资金支持。本研究获得了欧洲研究理事会执行局在 TRANSDIF- FUSE 项目（资助协议编号：101167322）下提供的资金支持。然而，文中表达的观点和意见仅代表作者本人，并不一定反映欧盟的观点和意见。欧盟及资助机构均不对这些观点和意见承担责任。LL 获得了 CSxAI 项目（PID2024-157526NB-I00）的部分支持，该项目由 MI- CIU/AEI/10.13039/501100011033/FEDER、UE 资助；同时还获得了由 MICIU/AEI/10.13039/501100011033 资助的玛丽亚·德·梅斯图卓越印章基金支持，以及

欧盟委员会芯片联合事业项目编号 101194363（NEHIL）的支持。

参考文献

[1] G. Carleo, I. Cirac, K. Cranmer, L. Daudet, M. Schuld, N. Tishby, L. Vogt-Maranto, L. Zdeborová, 《机器学习与物理科学》, 《现代物理学评论》91 (4) (2019) 045002。[2] G. E. Karniadakis, I. G. Kevrekidis, L. Lu, P. Perdikaris, S. Wang, L. Yang, 《物理信息机器学习》, 《自然评论：物理学》3 (6) (2021) 422–440。[3] S. Cai, Z. Mao, Z. Wang, M. Yin, G. E. Karniadakis, 《用于流体力学的物理信息神经网络（PINN）：综述》, 《力学学报》37 (12) (2021) 1727–1738。[4] C. Zhao, F. Zhang, W. Lou, X. Wang, J. Yang, 《物理信息神经网络及其在复杂流体动力学中应用的进展综述》, 《流体物理学》36 (10) (2024)。[5] I. J. Goodfellow, Y. Bengio, A. Courville, 《深度学习》，麻省理工学院出版社，美国马萨诸塞州剑桥，2016, http://www.deeplearningbook.org。[6] I. E. Lagaris, A. Likas, 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》9 (5) (1998) 987–1000。[7] M. Raissi, P. Perdikaris, 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反问题的深度学习框架》, 《计算物理学杂志》378 (2019) 686–707。[8] A. Bonfanti, R. Santana, M. Ellero, B. Gholami, 《物理信息神经网络在训练域之外的泛化及其影响超参数》, 《神经计算与应用》36 (36) (2024) 22677–22696。[9] S. Cai, Z. Wang, S. Wang, P. Perdikaris, G. E. Karniadakis, 《用于传热问题的物理信息神经网络》, 《传热学杂志》143 (6) (2021) 060801。[10] B. Huang, J. Wang, 《物理信息神经网络在电力系统中的应用：综述》, 《IEEE电力系统汇刊》38 (1) (2022) 572–588。[11] E. Zhang, M. Dao, G. E. Karniadakis, S. Suresh, 《利用物理信息神经网络分析材料的内部结构和缺陷》, 《科学进展》8 (7) (2022) eabk0644。

[12] F. Sahli Costabal, Y. Yang, P. Perdikaris, D. E. Hurtado, E. Kuhl, 《用于心脏激活映射的物理信息神经网络》, 《物理学前沿》8 (2020) 42。[13] A. Khan, D. A. Lowther, 《用于电磁分析的物理信息神经网络》, 《IEEE磁学汇刊》58 (9) (2022) 1–4。[14] S. Cuomo, V. S. Di Cola, F. Giampaolo, G. Rozza, M. Raissi, F. Piccialli, 《通过物理信息神经网络实现科学机器学习：现状与未来方向》, 《科学计算杂志》92 (3) (2022) 88。[15]J. H. Ferziger, M. Perić, R. L. Street, 《流体动力学计算方法》, 施普林格, 2019。[16] L. Sun, H. Gao, S. Pan, J.-X. Wang, 《基于不依赖仿真数据的物理约束深度学习的流体流动代理建模》, 《计算机方法在应用力学与工程》361 (2020) 112732。[17] H. Eivazi, M. Tahani, P. Schlatter, R. Vinuesa, 《用于求解雷诺平均纳维–斯托克斯方程的物理信息神经网络》, 《流体物理学》34 (7) (2022)。[18] M. Penwarden, S. Zhe, A. Narayan, R. M. Kirby, 《物理信息神经网络（PINN）的多保真度建模》, 《计算物理学杂志》451 (2022) 110844。[19] W. L. Oberkampf, C. J. Roy, 《科学计算中的验证与确认》, 剑桥大学出版社, 2010。[20] T. C. Papanastasiou, N. Malamataris, K. Ellwood, 《一种新的出口边界条件》, 《国际流体数值方法杂志》14 (5) (1992) 587–608。[21] F. M. Rohrhofer, S. Posch, C. Gößnitzer, B. C. Geiger, 《数据与物理：物理信息神经网络表面上的帕累托前沿》, 《IEEE开放期刊》11 (2023) 86252–86261。[22] A. Kendall, Y. Gal, R. Cipolla, 《利用不确定性对场景几何与语义的损失加权进行多任务学习》, 载于：IEEE计算机视觉与模式识别会议论文集, 2018, 第7482–7491页。[23] Z. Xiang, W. Peng, X. Liu, W. Yao, 《自适应损失平衡物理信息神经网络》, 《神经计算》496 (2022) 11–34。