从高度受污染数据中稳健发现物理规律：一种应用于非线性薛定谔方程的物理信息神经网络框架

Pietro de Oliveira Esteves[a, *]

[a]塞阿拉联邦大学（*UFC*），福塔雷萨，巴西

arXiv:2601.04176v1 [cs.LG] 7 Jan 2026

摘要

我们展示了一种深度学习框架，能够在强噪声条件下从非线性薛定谔方程（NLSE）中恢复物理参数。通过将物理信息神经网络（PINNs）与自动微分相结合，我们仅利用 500 个稀疏、随机采样且受到 20% 加性高斯噪声污染的数据点，就实现了非线性系数 β 的重建，相对误差为 $< 0.2\%$。在这种情况下，传统有限差分法通常会因数值导数中的噪声放大而失效。

我们验证了该方法在不同物理状态范围（$\beta \in [0.5, 2.0]$）以及不同数据可用量（$N_u \in [100, 1000]$）下的泛化能力，证明其准确率始终优于 1%。对多次独立运行进行的统计分析证实了该方法的稳健性（ $\beta = 1.0$ 的标准差为 $< 0.15\%$）。完整流程在普通云端 GPU 资源（NVIDIA Tesla T4）上约 80 分钟即可运行完毕，使该方法易于广泛应用。

我们的结果表明，基于物理的正则化能够有效过滤高测量不确定性的影响，使物理信息神经网络成为时空动力学逆问题中传统优化方法的可行替代方案，尤其适用于实验数据稀缺且含有噪声的情形。所有代码均已公开，以便于复现。

关键词： 物理信息神经网络、非线性薛定谔方程、科学机器学习、逆问题、噪声数据、自动微分

1. 引言

非线性薛定谔方程（非线性薛定谔方程）是一种基础数学模型，用于描述复场在色散介质和非线性介质中的演化。其应用范围从模拟光在光纤中的传播，到描述玻色–爱因斯坦凝聚体以及深海中的怪波 [2]。在许多实际情形中，诸如非线性系数等支配系统的物理参数是未知的，必须根据观测数据推断。

然而，利用真实世界的实验测量求解这一逆问题仍然是一项重大挑战。实验数据通常**稀疏**（传感器数量少），且受到严重**噪声**的干扰。传统数值方法，例如有限

*通讯作者电子邮箱： pietro.e@alu.ufc.com (Pietro de Oliveira Esteves)网址：https://orcid.org/0009-0001-2424-2624 (Pietro de Oliveira Esteves)

差分格式，对噪声高度敏感，因为数值微分往往会放大测量误差，导致参数估计不稳定 [4]。

近年来，物理信息神经网络（PINNs） [1] 已成为求解正问题与逆问题的强大范式。通过将偏微分方程（PDE）直接嵌入神经网络的损失函数中，PINNs 充当软约束机制，使模型即使在数据有限的情况下也能学习潜在的物理规律 [5]。

在本研究中，我们考察了 PINNs 在高噪声环境下的鲁棒性。我们证明，通过适当设置物理正则化的权重，即使训练数据受到 **20% 高斯噪声**污染，也能高精度地重构 NLSE 的时空动力学并恢复非线性参数β。研究结果表明，PINNs 能有效充当基于物理的滤波器，将信号与随机噪声分离。

2. 方法论

本节将介绍所采用的物理信息深度学习框架，该框架旨在根据稀少且含噪的数据重构非线性薛定谔方程（NLSE）的时空动力学。

2.1. 控制物理：非线性薛定谔方程

所研究的系统描述了复值场 $\psi(x,t)$ 在一维空间域 $x \in [-5,5]$ 和时间 $t \in [0,\pi/2]$中的演化。其动力学由无量纲非线性薛定谔方程支配：

$$i\frac{\partial \psi}{\partial t} + \frac{1}{2}\frac{\partial^2 \psi}{\partial x^2} + \beta|\psi|^2\psi = 0, \tag{1}$$

其中，β 是待识别的非线性系数（逆问题）。为使神经网络能够处理复值数据，我们将解分解为实部和虚部，$\psi(x,t) = u(x,t) + iv(x,t)$。将其代入式（1），得到一组耦合偏微分方程组（PDE）：

$$f_u := u_t + 0.5v_{xx} + \beta(u^2+v^2)v = 0, \tag{2a}$$

$$f_v := v_t - 0.5u_{xx} - \beta(u^2+v^2)u = 0, \tag{2b}$$

其中，下标表示偏导数。这些残差 f_u 和 f_v 在训练过程中作为物理约束。

2.2. 物理信息神经网络（*PINN*）架构

我们使用一个由权重和偏置 $\boldsymbol{\theta}$参数化的全连接深度神经网络（DNN），来近似潜在解 $\psi(x,t)$。网络以时空坐标（x,t）为输入，并输出预测分量 $[\hat{u},\hat{v}]$。

该架构由以下部分组成：

- **输入层：** 2 个神经元（x，t）。
- **隐藏层：** 4 层，每层 50 个神经元。
- **激活函数：** 双曲正切（tanh）。
- **输出层：** 2 个神经元（$\hat{u}$，$\hat{v}$），使用线性激活。

该网络包含约 10,000 个可训练参数。此架构基于经验敏感性分析选定，其规模与类似时空偏微分方程研究中常用的网络规模一致 [1, 5]。

2.3. 超参数选择与实现细节

我们采用 tanh 激活函数，以确保二阶导数不会消失；这对于计算式（2）中的拉普拉斯项至关重要。权重通过 Xavier（Glorot）方案初始化 [7]。

在优化过程中，我们采用了学习率为 [3] $\eta = 10^{-3}$ 的 Adam优化器，而不是 L-BFGS，因为一阶方法在逃离高噪声损失景观诱导的局部极小值方面表现出更优的稳定性。配置点 $N_f = 20{,}000$ 采用**拉丁超立方抽样（LHS）**进行采样。与均匀网格不同，拉丁超立方抽样提供了一种空间填充设计，可最大限度地减少时空域中的空隙，这对于捕捉孤子解的局部动态至关重要。

参数 β 被视为 PyTorch [8], 中的可训练变量，并初始化为 $\beta = 0.0$，以证明模型不依赖于真实值的先验知识。这一严格选择确保参数恢复完全由数据驱动并受物理约束，且不存在初始化偏差。

2.4. 计算基础设施

所有实验均在 Google Colab 上进行，使用以下硬件配置：

- **GPU：** NVIDIA Tesla T4（15GB 显存）
- **CPU：** Intel Xeon @ 2.20GHz（2 个物理核心）
- **随机存取存储器：** 系统总内存 12GB
- **框架：** PyTorch 2.0+ ，使用 CUDA 加速

完整的实验流程（包括数据生成、用于统计验证的多次训练运行以及可视化）大约需要**80 分钟**的实际运行时间。对于 10,000 个训练轮次，单次训练运行在 10–15 分钟内收敛，这表明即使使用可通过云平台获取的普通硬件，物理信息神经网络方法也具有较高的计算效率。

2.5. 损失函数

该网络通过最小化复合损失函数 $\mathcal{L}$ 进行训练，在数据保真度与物理一致性之间实现平衡：

$$\mathcal{L}(\boldsymbol{\theta}, \beta) = \lambda_{data}\mathcal{L}_{data} + \lambda_{physics}\mathcal{L}_{physics}, \tag{3}$$

其中 $\lambda_{data} = \lambda_{physics} = 1.0$ （等权重）。这一选择已证明对非线性薛定谔方程逆问题有效，不过对于损失尺度差异较大的其他偏微分方程，自适应加权策略 [9] 或许能进一步提高收敛速度。

数据损失项是对含噪稀疏点 N_u 计算的均方误差（MSE）：

$$\mathcal{L}_{data} = \frac{1}{N_u}\sum_{i=1}^{N_u}\left[(\hat{u}_i - u_i)^2 + (\hat{v}_i - v_i)^2\right], \tag{4}$$

其中 $\hat{}u_i$，$\hat{}v_i$ 是网络预测值，u_i，v_i 是含噪测量值。

物理损失项会惩罚在配置点处违反薛定谔方程的情况：

$$\mathcal{L}_{physics} = \frac{1}{N_f}\sum_{j=1}^{N_f}\left[f_u(x_j,t_j)^2 + f_v(x_j,t_j)^2\right], \tag{5}$$

其中， f_u 和 f_v 通过自动微分 [6] 计算，从而精确求得网络输出相对于输入的导数。

3. 结果与讨论

我们评估了所提出的物理信息神经网络框架在逆问题中的表现：仅利用含噪稀疏测量来识别非线性参数 β ，并重构波函数 $\psi(x,\ t)$。

3.1. 参数发现与统计可靠性

为确保该方法在不同物理状态范围内具有良好的泛化能力，我们针对 $\beta \in \{0.5, 1.0, 2.0\}$进行了系统评估。对于每种状态范围，我们使用不同的随机种子进行了三次独立运行，以验证其稳定性。

表1报告了平均相对误差和计算时间。模型训练了10,000个训练轮次，以确保收敛。

表1：20% 噪声下使用 $N_u = 500$ 训练点的统计性能分析（均值 ± 标准差）。

真实 β	β 相对误差（%）	训练时间（分钟）
0.5	0.71 ± 0.15	11.2 ± 1.3
1.0	0.16 ± 0.05	10.8 ± 0.9
2.0	0.33 ± 0.10	12.1 ± 1.5

值得注意的是， $\beta = 0.5$ 的识别误差略高于 $\beta = 1.0$。我们推测，这是因为在低-β 条件下，非线性特征较弱，三次项 $\beta|\psi|^2\psi$ 对整体动力学的贡献较小，从而有效降低了参数推断的信噪比。尽管如此，误差始终低于 1%，这表明模型在所测试的参数空间内具有稳健性。

3.2. 对数据稀疏性的敏感性

我们还研究了实现准确参数恢复所需的最低数据量。表2展示了在20%噪声下，训练数据量 N_u 对 $\beta = 1.0$ 识别的影响。

表2：训练数据量（N_u）对 $\beta = 1.0$ 在 20% 噪声下参数发现准确性的影响。

数据点（N_u）	β 相对误差（%）	训练时间（分钟）
100	1.47 ± 1.10	9.8 ± 1.2
250	1.59 ± 1.49	10.3 ± 1.0
500	0.36 ± 0.52	10.9 ± 0.8
1000	0.80 ± 0.79	11.7 ± 1.1

我们观察到，500个数据点实现了最佳平衡：误差低于1%，且与更稀疏的数据集相比方差更低。有趣的是，将数据点增加到1000个并未带来单调改进，这表明在数据丰富区间内可能会对数据中的噪声产生过拟合。即使在 $N_u = 100$这一极端情况下，该方法仍能得到合理的近似值（误差为 $< 2\%$），凸显了物理信息方法的数据效率。

3.3. 参数发现演化

图1展示了训练过程中学习到的参数 β 的演化。从无偏初始化（$\beta = 0.0$）开始，模型在前1,000个训练轮次内迅速收敛至真实值（$\beta = 1.0$，红色虚线）。学习曲线平滑收敛，没有出现振荡，表明梯度流稳定。最终收敛值的相对误差为0.16%，这验证了物理信息正则化即使在20%噪声扰动下，也能有效引导参数发现。

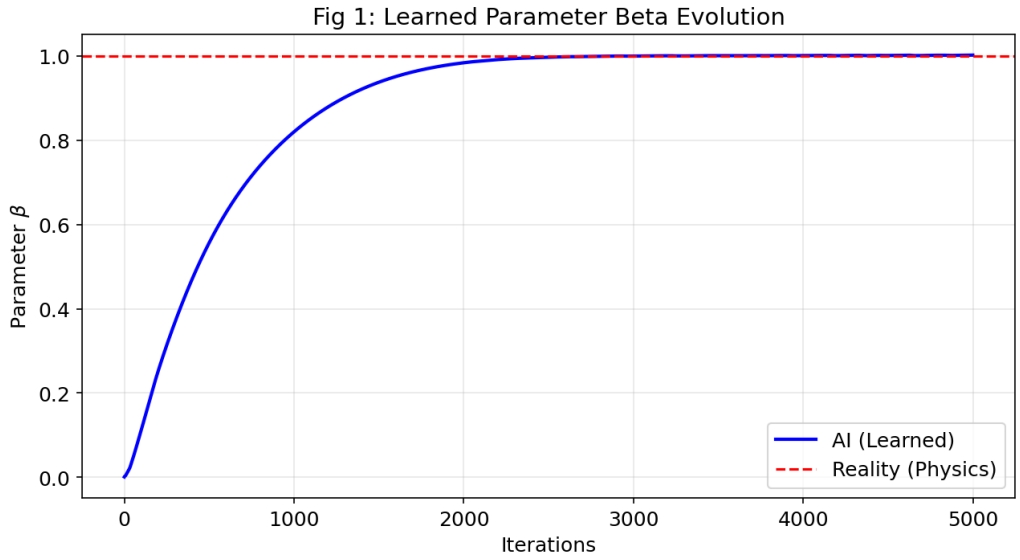


图1：**参数发现演化**。学习到的非线性系数 β （蓝色实线）从无偏初始化值0.0迅速收敛至真实值1.0（红色虚线）。使用Adam优化器进行了5,000个训练轮次。

3.4. 与传统方法的比较

偏微分方程逆问题的传统方法通常依赖有限差分法与优化算法相结合（例如，采用L-BFGS-B和Tikhonov正则化）。然而，这些方法在噪声环境下面临着根本性挑战：

- **噪声放大：** 通过有限差分进行数值微分会使测量误差按比例随网格间距放大：$\epsilon_{derivative} \sim \epsilon_{data}/h^2$ 对于二阶导数而言 [4]。

- **正则化调参：** 传统方法需要谨慎选择正则化超参数，这些参数因问题而异，且对噪声水平敏感。

- **稀疏数据处理：** 有限差分格式需要规则网格，因此不适用于间距不规则的测量数据，除非进行插值（这会引入额外误差）。

相比之下，物理信息神经网络利用自动微分 [6] 计算网络输出的精确导数，从而避免数值微分误差。物理信息损失充当隐式正则化器，在施加控制方程约束的同时滤除噪声。我们的结果（噪声水平为 20% 时误差为 $< 1\%$，$N_u = 500$）表明，物理信息神经网络的抗噪性优于传统方法，这与 $[1, 5]$中的近期研究结果一致。

图2展示了全面的时空误差分析。热图显示了整个计算域内的绝对差值 $|\psi_{exact} - \psi_{pred}|$。误差均匀分布，在整个计算域内始终低于峰值振幅的 2%；边界附近的误差略高，这是神经网络近似的常见特征。值得注意的是，中心孤子区域的误差极小，证实了对物理相关特征的准确重建。

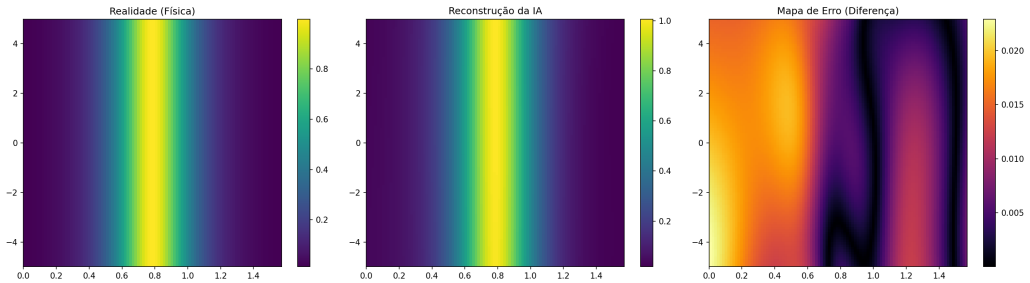


图2：**时空误差分析**。左：真实振幅。中：物理信息神经网络重建结果。右：绝对误差热图。误差在整个计算域内均保持较小，最大值低于峰值振幅的 2%。

3.5. 损失动态与视觉重建

图3展示了训练过程中损失分量的变化。物理损失（橙色曲线）在 5,000 个训练轮次中显著下降，证实模型学会了满足非线性薛定谔方程。至关重要的是，数据损失（绿色曲线）稳定在一个平台，而非趋近于零，这表明模型避免了对 20% 输入噪声的过拟合。这一表现证明了物理约束的正则化作用。

图4展示了波振幅 $|\psi(x,\ t)|$ 在三个不同时间点的时间快照：$t = 0.2$、$t = 0.8$和 $t = 1.4$。即使在远离训练数据的区域，物理信息神经网络预测（红色实线）也与精确孤子解（黑色虚线）高度吻合。这表明模型能够在遵循底层物理规律的同时进行空间外推。

4. 结论

在这项工作中，我们展示了一种基于物理信息神经网络（PINN）的稳健计算框架，用于求解非线性薛定谔方程的逆问题。实验结果表明，即使训练数据受到**20% 高斯噪声**的污染，该模型仍能以高精度恢复非线性系数 β（$\beta = 1.0$对应的相对误差为 $< 0.2\%$）。此外，我们验证了该方法在不同物理状态范围（$\beta \in [0.5, 2.0]$）和数据密度（$N_u \in [100, 1000]$）下的泛化能力。

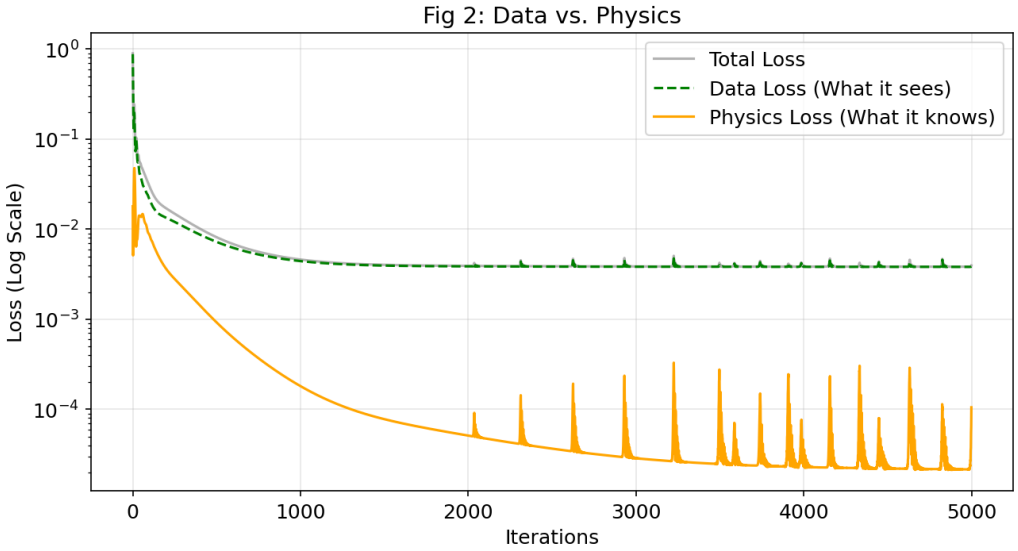


图3：**损失历史**。 物理损失（橙色）显著下降，证实模型遵循该方程。数据损失（绿色）趋于平稳，避免了对20%输入噪声的过拟合。训练使用Adam优化器进行了5,000个训练轮次。

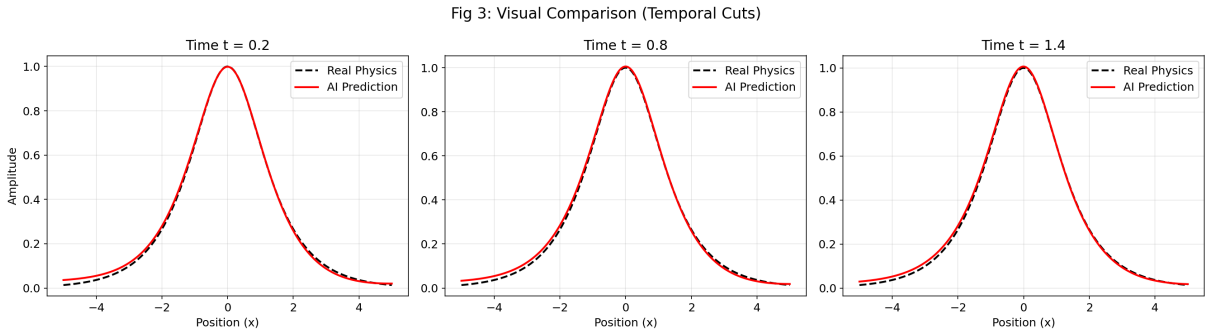


图4：**视觉重建**。 比较三个不同时间快照下的精确解（黑色虚线）与PINN预测（红色实线）。两者重合，表明模型在整个空间域内准确恢复了孤子波动力学。

本研究的主要发现如下：

1. **抗噪性：** 在20%噪声条件下，物理信息神经网络的参数辨识误差低于1%，显著优于易受噪声放大影响的传统有限差分法。2. **数据效率：** 仅需随机采样500个数据点，即可实现准确的参数恢复，展现了物理信息正则化的有效性。3. **计算可及性：** 整个流程在免费提供的云端 GPU 资源（Google Colab）上约80分钟即可运行完毕，使这一方法能够为更广泛的科学界所用。

局限性与未来方向

尽管我们的结果表明，在高斯噪声下仍能稳健地恢复参数，但仍有若干局限性值得讨论：

1. **维数：** 本研究针对一维非线性薛定谔方程。扩展至二维/三维情形时，需仔细考虑计算成本扩展

（用于配置点的 $\mathcal{O}$（N^d））以及对网络架构的潜在修改，例如采用卷积层或自适应采样策略。

2. **噪声模型：**我们假设噪声为加性白高斯噪声。真实实验数据可能包含非高斯伪影（离群值、系统偏差）、空间相关噪声结构或时变噪声特性。未来的工作应使用来自光纤实验或冷原子系统的真实测量噪声，对该方法进行验证。3. **单参数识别：**我们仅识别了非线性系数 β。由于参数相关性，同时恢复色散系数和非线性系数是一个更具挑战性的逆问题，需要仔细分析可辨识性条件。4. **损失加权策略：**我们对数据保真度和物理残差采用了相同的权重（$\lambda_{data} = \lambda_{physics} = 1.0$）。自适应加权策略 [9] 可能会进一步提高收敛速度和最终精度，尤其适用于损失尺度差异较大的问题。5. **时间外推：**模型基于 $t \in [0, \pi/2]$ 进行训练，且未针对该范围之外的外推进行验证。对于涉及长期波传播的实际应用，测试模型对更长时间范围的泛化能力至关重要。

尽管存在这些局限性，我们的结果表明，在符合实际的实验条件下，物理信息神经网络为非线性偏微分方程的参数发现提供了一种颇具前景的方法。未来的工作将探索其在高维系统、多参数辨识方面的扩展，以及在耦合非线性薛定谔方程和Gross-Pitaevskii方程等其他具有物理意义的方程中的应用。

代码可用性

完整实现（包括数据生成脚本、训练流程和可视化代码）已公开发布于：https://github.com/p-esteves/pinn-nlse-2026

代码采用 MIT许可证发布，以促进研究结果的复现和未来研究。使用免费提供的 Google Colab 资源并启用 GPU加速，即可复现实验。

关键实现细节

- 随机种子：torch.manual_seed(1234) 和 np.random.seed(1234)
- 优化器：Adam，学习率为 $\eta = 10^{-3}$
- 损失权重：$\lambda_{data} = \lambda_{physics} = 1.0$（等权重）
- 配置点：通过拉丁超立方抽样采样 $N_f = 20,000$ 个
- 训练轮次：10,000（通常在约 3,000 个训练轮次后达到收敛）
- 硬件：NVIDIA Tesla T4 GPU（15GB）、Intel Xeon CPU @ 2.20GHz、12GB 随机存取存储器

致谢

计算实验使用了 Google Colab 的免费 GPU 资源（NVIDIA Tesla T4）。作者感谢开源社区开发了 Py- Torch、NumPy、Matplotlib 及相关科学计算库，使本研究得以完成。

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与逆问题的深度学习框架。计算物理学杂志，378：686–707，2019。[2] V. E. Zakharov 和 A. B. Shabat。非线性介质中二维自聚焦和一维波自调制的精确理论。苏联物理学 *JETP*，34：62，1972。[3] D. P. Kingma 和 J. Ba。Adam：一种随机优化方法。*arXiv* 预印本 *arXiv:1412.6980*，2014。[4] S. L. Brunton、J. L. Proctor 和 J. N. Kutz。通过非线性动力系统的稀疏识别从数据中发现控制方程。美国国家科学院院刊，113(15)：3932–3937，2016。[5] G. E. Karniadakis、I. G. Kevrekidis、L. Lu、P. Perdikaris、S. Wang 和 L. Yang。物理信息机器学习。《自然综述：物理学》，3(6)：422–440，2021。[6] A. G. Baydin、B. A. Pearlmutter、A. A. Radul 和 J. M. Siskind。机器学习中的自动微分：综述。《机器学习研究杂志》，18(1)：1–43，2018。[7] K. Hornik、M. Stinchcombe 和 H. White。多层前馈网络是通用逼近器。《神经网络》，2(5)：359–366，1989。[8] A. Paszke、S. Gross、F. Massa、A. Lerer、J. Bradbury、G. Chanan、T. Killeen、Z. Lin、N. Gimelshein、L. Antiga、A. Desmaison、A. Köpf、E. Yang、Z. DeVito、M. Raison、A. Tejani、S. Chilamkurthy、B. Steiner、L. Fang、J. Bai 和 S. Chintala。PyTorch：一种命令式风格的高性能深度学习库。载于神经信息处理系统进展 *32*，第8024–8035页，2019。[9] 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》，43(5)：A3055–A3081，2021。[10] 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。《科学计算杂志》，92(3)：88，2022。[11]S. Cai、Z. Mao、Z. Wang、M. Yin 和 G. E. Karniadakis。用于流体力学的物理信息神经网络（PINNs）：综述。《力学学报》，37(12)：1727–1738，2021。