

基于扩展混合训练物理信息神经网络的 非线性薛定谔方程求解和参数发现*

王宇铎 陈嘉鑫 李彪†

(宁波大学数学与统计学院, 宁波 315211)

(2025 年 4 月 1 日收到; 2025 年 5 月 26 日收到修改稿)

提出了扩展混合训练物理信息神经网络 (X-MTPINNs), 该模型通过整合扩展物理信息神经网络 (X-PINNs) 的域分解技术与混合训练物理信息神经网络 (MTPINNs) 框架, 有效提升了非线性波动问题的求解能力. 相较于经典物理信息神经网络 (PINNs) 模型, 新模型具有双重优势: 1) 混合训练框架通过优化初边值条件的处理机制, 显著改善了模型收敛特性, 在提升非线性波解拟合精度的同时, 将计算时间降低约 40%; 2) X-PINNs 的域分解技术增强了模型对复杂动力学行为的表征能力. 基于非线性薛定谔方程 (NLSE) 的数值实验表明, X-MTPINNs 在亮双孤子解及三阶怪波求解以及参数反演等任务中均表现优异, 其预测精度较传统 PINNs 提升一至两个数量级. 对于逆问题, X-MTPINNs 算法在有噪声和无噪声条件下都能准确识别 NLSE 中的未知参数, 解决了经典 PINNs 在本研究条件下 NLSE 参数识别中完全失效的问题, 表现出很强的鲁棒性.

关键词: 物理信息神经网络, 非线性薛定谔方程, 扩展混合训练物理信息神经网络, 域分解, 参数发现

PACS: 02.30.Ik, 02.60.-x, 05.45.-a

DOI: 10.7498/aps.74.20250422

CSTR: 32037.14.aps.74.20250422

1 引言

20 世纪 90 年代初, Dissanayake 和 Thien^[1] 开创性地将神经网络用于求解偏微分方程 (PDEs). 到了 90 年代末, 人工神经网络 (ANN)^[2] 问世. 然而, 由于当时硬件能力的限制, 其发展受到了制约. PDEs 在科学研究中占有重要地位, 解决这些方程的传统数值方法包括有限体积法 (FVM)^[3]、有限差分法 (FDM)^[4]、有限元法 (FEM)^[5] 等. 尽管这些基于网格的方法被广泛使用, 但在实际应用中存在一定的局限性.

近年来, 随着 TensorFlow 和 PyTorch 等深度学习框架的出现, 深度学习在计算机视觉^[6]、生物医学^[7] 和自然语言处理^[8] 等领域取得了一定的进展, 使得人们对利用神经网络求解 PDEs 再次产生

了兴趣. 2019 年, Raissi 等^[9] 创新性地提出了物理信息神经网络 (PINNs). PINNs 背后的核心理念是通过偏微分方程组来设计损失函数, 将物理约束纳入神经网络, 从而使网络能够模拟物理规律.

PINNs 的问世激发了各个领域研究人员的灵感, 在能源^[10]、环境科学^[11]、工程^[12]、数学物理等领域都有应用. 尽管 PINNs 被广泛采用, 但它仍然容易受到训练失败、收敛缓慢等问题的影响. 为了应对这些挑战, 研究人员通过将 PINNs 与各种网络架构以及统计学方法相结合, 对 PINNs 进行了改进, 以满足不同场景的需求. 这导致了 CAN-PINNs^[13] 和 B-PINNs^[14] 等方法的发展. 采用拉格朗日参考系, LPINN 得到了发展^[15]. 此外, 域分解技术还催生了 XPINNs 等方法^[16], 它们在共享边界上应用适当的连续性条件, 将计算域分解为多个子域, 求解一系列独立的子问题, 从而重建全局解. 还有其

* 国家自然科学基金 (批准号: 12175111, 12235007, 12275144) 和宁波大学王宽诚幸福基金资助的课题.

† 通信作者. E-mail: libiao@nbu.edu.cn

他一些具有先进的深度学习算法的工作^[17-19].

在数学物理领域,可积系统中的许多方程都具有重要的研究价值, Li 等^[20]使用多尺度扰动法推导出了 5 阶 KdV 方程的渐进解. Chen 和 Wang^[21]使用微扰法给出了 KdV 方程的一种三孤子解. 同时,非线性 PDEs 的数值解也一直是领域内的热门话题. 2020 年, Li 和 Chen^[22,23]将 PINNs 框架应用于可积系统,研究了二阶和三阶 PDEs 的非线性演化. Song 等^[24]和 Wang 等^[25]利用 PINNs 研究了相关方程的数据驱动解和参数发现. 在此之后,人们利用可积系统的特性,为这类系统量身设计了新的深度神经网络模型. Qiu 等^[26]基于广义 Miura 变换的 PINNs 研究了 KdV 类方程的新孤子解. 李彪等^[27-29]提出了梯度下降、先验信息以及混合训练 PINNs 来求解非线性 PDEs 的怪波,在训练时间和精度方面取得了令人称道的成果. Zhou 等^[30]研究了与变系数 Hirota 方程相关的正演逆问题. Qiu 等^[31]将先验信息、自适应权重和激活函数相结合来优化预测. Liu 等^[32]利用 PINNs 对 PT 对称势下 NLSE 的多孤子解进行预测. Liu 等^[33]提出了混合采样 PINNs,通过预训练使得精度提高. 此外, Zhang 和 Chen^[34]将数学物理中的一些问题与方法融入并设计出了新的神经网络,利用 Riemann-Hilbert 问题和 PINN 方法研究了新型非局部 Sasa-Satsuma 方程的 N 孤子和数据驱动解,同时也将 Lax 对纳入神经网络,推导出 Korteweg-de Vries (KdV) 方程和 NLSE 等 PDEs 的解^[35].

PINNs 模型为解决 PDEs 的正问题和反问题提供了一种高效的数据驱动方法,尤其是在一些传统方法无法解决的情况下. 然而,对于一些 PDEs 较为复杂的数值解,经典 PINNs 模型面临了巨大挑战,在计算效率、准确性和鲁棒性方面都面临很大的困难,而且在参数识别任务中可能完全失败. 相比之下,我们提出的扩展混合训练 PINNs (X-MTPINNs) 模型巧妙地避免了将初边值条件直接加入损失函数,并降低了计算时间. 同时利用域分解技术,帮助模型对峰值更大的区域实现更好的拟合效果,与经典 PINNs 相比,能显著提高计算速度、预测精度和鲁棒性. 同时还能准确识别逆问题中的未知参数,提高泛化能力,克服经典 PINNs 在应对复杂动态挑战时的局限性.

本文第 2 节简要介绍了经典 PINNs 模型,并提出了扩展混合训练 PINNs 模型;第 3 节通过数

值实例,利用经典 PINNs 和 X-MTPINNs,研究了 NLSE 数据驱动解的动态行为,包括亮双孤子解以及三阶怪波,并以三阶怪波为基础,讨论了通过 X-MTPINNs 算法在有噪声和无噪声的 NLSE 中发现数据驱动的参数. 最后第 4 节对本文的工作进行了总结.

2 模型

2.1 经典 PINNs

首先,简要介绍经典的 PINNs 模型. 以下是使用 PINNs 求解 PDEs 时的基本形式,并给出了 Dirichlet 边界条件:

$$\begin{cases} f(x, t) = h_t + N_x[h] = 0, & (x, t) \in \Omega \times T, \\ h(x, t) = I(x), & x \in \Omega, t = t_0, \\ h(x, t) = B(x, t), & (x, t) \in \partial\Omega \times T, \end{cases}$$

其中, $N[\cdot]$ 是一个非线性微分算子; T 和 Ω 分别表示时间和空间区域, $\partial\Omega$ 是空间区域 Ω 的边界; $I(x)$ 和 $B(x, t)$ 是初始条件和边界条件算子.

PINNs 通过构建一个神经网络模型来逼近函数 $h(t, x)$. 网络输入通常包括时间 t 和空间变量 x , 而输出则是对应的函数值 $h(t, x)$. 网络的训练过程主要依赖于定义一个损失函数:

$$\text{MSE} = \text{MSE}_{\text{IC}} + \text{MSE}_{\text{BC}} + \text{MSE}_f,$$

该损失函数包含多个部分,以确保模型能够准确地拟合数据并遵循物理规律. 其中, MSE_{IC} 表示方程初值所产生的误差; MSE_{BC} 表示方程边值产生的误差; MSE_f 表示将训练得出的函数代入方程产生的残差.

2.2 X-MTPINNs

PINNs 出现后,已成为求解 PDEs 的强大工具. 虽然经典的 PINNs 已显示出巨大的前景,但面临着收敛速度慢、难以处理复杂的高维问题等挑战,尤其是在应用于具有局部特征或不连续性的大规模物理系统时. 为解决这些局限性,人们开发了 PINNs 的一个扩展版本,如扩展 PINNs (XPINNs),提高了解决不同领域 PDEs 的性能和鲁棒性. 这种方法降低了解决大规模问题的复杂性,加快了训练收敛速度,使流体动力学、电磁学和量子力学等领域的复杂系统研究成为可能. 混合训练 PINNs 将混合训练中随机采样点的初始条件和边界条件与

微分方程的残差结合起来作为神经网络的约束条件, 提高了模型的训练速度和精度. 近年来, 它在 NLSE 方程等怪波求解问题上取得了良好的效果.

本文提出 X-MTPINNs, 利用混合训练框架将方程的初边值点和内部点相混合, 作为同一神经网络的约束条件, 加入损失函数之中, 让模型能更专注于方程整体的内在性质和动力学行为, 也降低了计算时间, 加速了模型的训练. 同时, 使用 XPINNs 中的域分解技术, 对神经网络每个时间子域进行建模, 并在共享边界上实施适当的连续性条件. 这使该模型能更精准地捕捉孤子解、怪波解在特定区域上的动力学行为, 降低了整体域上拟合的复杂性. 因此, 该模型能显著提高预测精度与稳定性, 同时减少计算时间. 这一过程如图 1 所示. 与经典的 PINNs 类似, 进行域分解后, 在每个子域内训练一个神经网络. 最后, 对训练结果进行整合, 形成全局解. 调整后的 X-MTPINNs 见图 2.

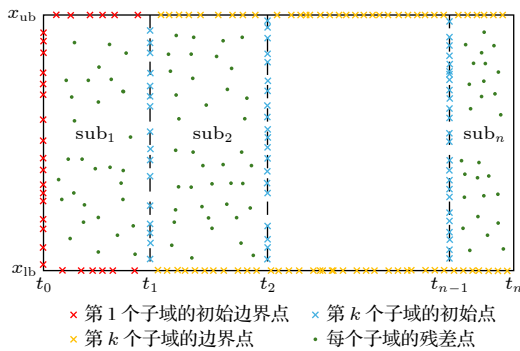


图 1 区域分解示意图

Fig. 1. Domain decomposition schematic diagram.

考虑上面同样的方程组. 首先将混合训练框架加入其中, 将初边值条件点和内部点混合到一起. 此时, 新的损失函数表达式为

$$\begin{aligned} \text{MSE} &= \text{MSE}_{\text{IB}} + \text{MSE}_f \\ &= \frac{1}{N_{\text{IB}}} \sum_{i=1}^{N_{\text{IB}}} |h(x_{\text{IB}}^i, t_{\text{IB}}^i) - h_{\text{IB}}^i|^2 \\ &\quad + \frac{1}{N_f} \sum_{i=1}^{N_f} |f(t_f^i, x_f^i)|^2, \end{aligned}$$

这里, $\{x_{\text{IB}}^i, t_{\text{IB}}^i, h_{\text{IB}}^i\}_{i=1}^{N_{\text{IB}}}$ 表示 $h(x, t)$ 在初边值条件下内部点混合集的训练数据, 而 $\{t_f^i, x_f^i\}_{i=1}^{N_f}$ 指定了 $f(x, t)$ 的配置点.

其次, 假设时间域为 $[T_0, T_1]$, 再将其划分为 n 个子域: $[T_0, T_1] = [t_0, t_n] = [t_0, t_1]^1 \cup [t_1, t_2]^2 \cup \dots \cup [t_{n-1}, t_n]^n$. 对于第 k 个子域 ($k > 0$), 考虑到划分后的都为闭集, 因此可以将新的边界作为初边值条件, 与内部点一起作为混合集, 并用拉丁超立方体采样来随机选择残差点. 此时, 第 k 个子域损失函数的表达式为

$$\begin{aligned} \text{MSE}_k &= \text{MSE}_{\text{IB}}^{(k)} + \text{MSE}_f^{(k)} \\ &= \frac{1}{N_{\text{IB}_k}} \sum_{i=1}^{N_{\text{IB}_k}} |h(x_{\text{IB}_k}^i, t_{\text{IB}_k}^i) - h_{\text{IB}_k}^i|^2 \\ &\quad + \frac{1}{N_{f_k}} \sum_{i=1}^{N_{f_k}} |f(t_{f_k}^i, x_{f_k}^i)|^2. \end{aligned}$$

最后, 通过减少每个循环中的 MSE_{IB} 和 MSE_f 三个惩罚项, 训练模型的学习能力.

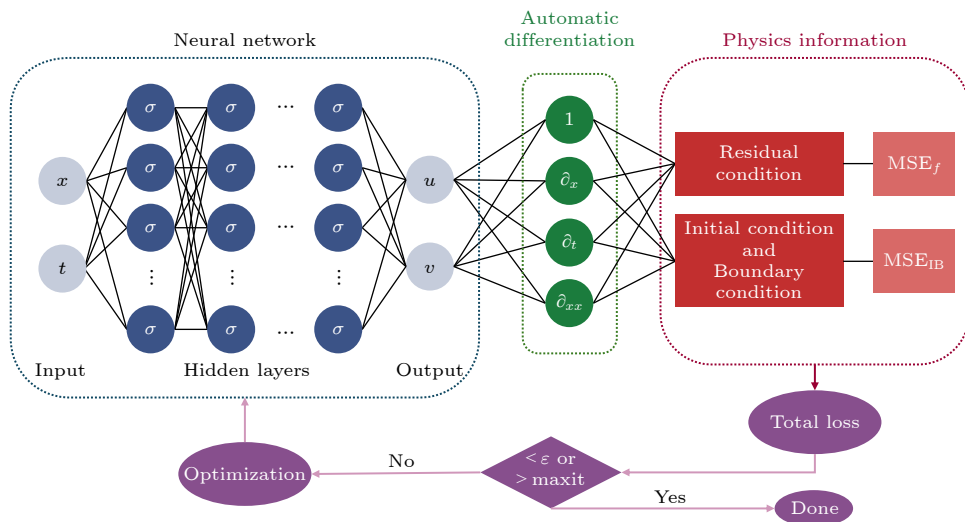


图 2 扩展混合训练物理信息神经网络 (X-MTPINNs)

Fig. 2. Extended mixed-training physics-informed neural networks (X-MTPINNs).

3 数值结果

本节以 NLSE 的亮双孤子解和三阶怪波的数值结果为例, 评估第 2 节所示的两个模型预测 NPDEs 解的性能. 在所有情况下, 在不失去通用性的前提下, 使用具有 4 个隐藏层、每个隐藏层中有 50 个神经元的深度全连接神经网络, 将双曲正切函数定义为模型的非线性激活函数. 使用默认的 Adam 和 L-BFGS 优化器分别训练了 50000 步. 所有模型均使用 Xavier 方法进行初始化, 并且没有使用其他额外的正则化技术.

非线性薛定谔方程 (nonlinear Schrödinger equation, NLSE) 的形式为

$$i\frac{\partial h}{\partial t} + \frac{1}{2}\frac{\partial^2 h}{\partial x^2} + |h|^2 h = 0, \quad (1)$$

其中, $h(x, t)$ 表示复幅度, x 和 t 分别表示空间和时间变量. 该方程通常用于建模光脉冲在非线性介质中的传播, 其中非线性项 $|h|^2 h$ 反映了介质的非线性效应.

3.1 数据驱动问题

3.1.1 亮双孤子解

本小节首先以 NLSE 的亮双孤子解作为数值实验. 亮双孤子解在描述各种物理现象中具有重要意义, 如光纤中的脉冲传播、海洋波浪、等离子体物理等. 这些孤子能够在不改变形状的情况下长时间传播, 反映了非线性波动的稳定性. 研究亮双孤子解有助于深入理解可积系统的行为, 通过数值模拟可以验证孤子解的特性, 帮助研究人员在其他非线性系统中寻找类似的解.

亮双孤子解^[36]的表达式为

$$h(x, t) = G/F,$$

其中,

$$G = a_1 \cosh(\theta_2) e^{i\phi_1} + a_2 \cosh(\theta_1) e^{i\phi_2} + ia_3 (\sinh(\theta_2) e^{i\phi_1} - \sinh(\theta_1) e^{i\phi_2}),$$

$$F = b_1 \cosh(\theta_1 + \theta_2) + b_2 \cosh(\theta_1 - \theta_2) + b_3 \cos(\phi_2 - \phi_1).$$

参数定义和取值为

$$\theta_k = \eta_k x - \eta_k \xi_k t - \theta_{k0} \quad (k = 1, 2),$$

$$\phi_k = \xi_k x + \frac{1}{2}(\eta_k^2 - \xi_k^2)t - \phi_{k0} \quad (k = 1, 2),$$

$$a_k = \frac{\eta_k}{2} [\eta_k^2 - \eta_{3-k}^2 + (\xi_1 - \xi_2)^2] \quad (k = 1, 2),$$

$$b_k = \frac{1}{4} \left\{ [\eta_1 + (-1)^k \eta_2]^2 + (\xi_1 - \xi_2)^2 \right\} \quad (k = 1, 2),$$

$$a_3 = \eta_1 \eta_2 (\xi_1 - \xi_2), \quad b_3 = -\eta_1 \eta_2.$$

$$\eta_1 = 0.8, \quad \eta_2 = 0.3, \quad \xi_1 = 0.5, \quad \xi_2 = -0.5, \quad \theta_{10} = 1.0, \quad \theta_{20} = -1.0, \quad \phi_{k0} = 0.$$

利用第 2 节介绍的两个模型以及第 3 节介绍的训练条件和优化方法, 得到了相应的数值结果, 如图 3 所示. 图 3 对第 2 节介绍的两个模型的预测性能进行了详细评估, 结果显示了两个局部波包. PINNs 模型捕捉到了较低波包的大致轮廓, 但在较高波包附近表现不佳, 这是深度学习模型中的一个常见问题. 高位波峰陡峭的梯度变化会导致显著的预测误差. 此外, PINNs 难以逼近亮双孤子解的整体结构, 最终导致预测效果不佳. 相比之下, X-MTPINNs 模型的表现符合预期, 仅在较高的峰值附近存在相对较高的绝对误差, 而在其他时空区域则更加平滑和稳定. 如表 1 所列, X-MTPINNs 的准确度比 PINNs 高出约一个数量级, 凸显了本文模型的优越性.

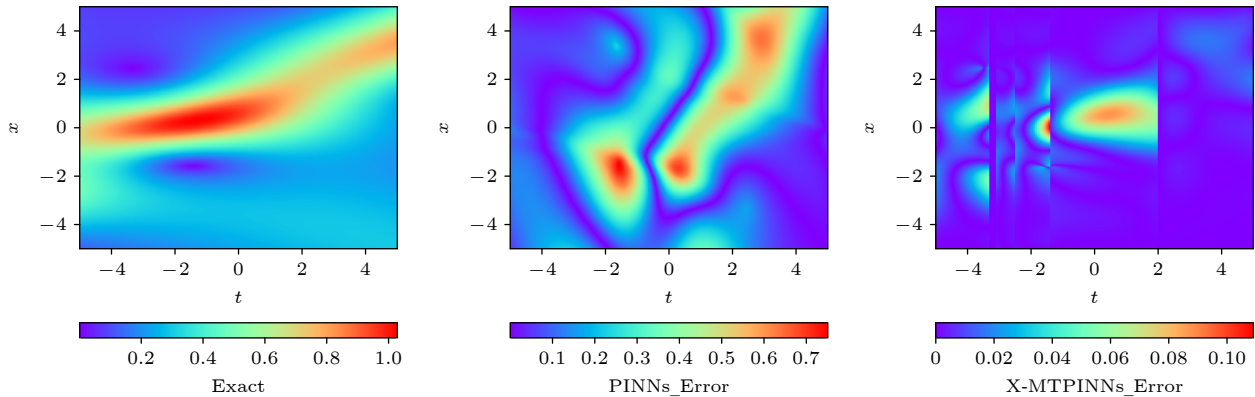


图 3 NLSE 亮双孤子解的真实值和不同模型的预测误差

Fig. 3. Exact values of two bright solitons in the NLSE and prediction errors from different models.

表 1 两种模型对 NLSE 亮双孤子解的数值预测结果

Table 1. Numerical prediction results of two models for two bright solitons in the NLSE.

亮双孤子	子域	隐藏层数	神经元数	L_2 误差	训练时间/s
X-MTPINNs	[-5.00, -3.30]	4	50	4.52×10^{-2}	2531.88
	[-3.30, -2.50]	4	50	2.63×10^{-2}	2459.23
	[-2.50, -1.40]	4	50	3.03×10^{-2}	2347.56
	[-1.40, 2.00]	4	50	5.73×10^{-2}	2489.87
	[2.00, 5.00]	4	50	1.70×10^{-2}	2489.87
PINNs	Global [-5, 5]	4	50	3.38×10^{-2}	2531.88
PINNs	[-5, 5]	4	50	2.51×10^{-1}	4176.69

3.1.2 三阶怪波

其次考虑三阶怪波. 三阶怪波作为怪波的高阶形式, 具有更复杂的结构和更高的峰值, 能在短时间和有限区域内释放出大量能量. 这种波动现象在多个领域具有重要的应用意义, 不仅为极端事件的数学建模提供了依据, 也有助于理解非线性系统中的异常行为.

三阶怪波^[36]的表达式为

$$h(x, t) = \left[-1 + \frac{G(x, t) + iK(x, t)}{D(x, t)} \right] e^{it},$$

其中 $G(x, t)$, $K(x, t)$ 和 $D(x, t)$ 的多项式如下:

$$\left\{ \begin{aligned} G(x, t) &= 24576x^{10} + (92160 + 368640t^2)x^8 + (322560 - 368640t^2 + 1228800t^4)x^6 + (-172800 + 691200t^2 \\ &\quad - 921600t^4 + 1720320t^6)x^4 + 1036800t^4 + 2580480t^6 + 1105920t^8 + 2703360t^{10} + 3041280t^8, \\ K(x, t) &= 49152t^{11} + (307200 + 245760x^2)t^9 + (-2672640 - 7372800x^2 + 4915200x^4)t^7 + (-7395840 \\ &\quad + 5253120x^2 - 2580480x^4 + 491520x^6)t^5 + (-4838400x^2 - 1382400x^4 - 1720320x^6 + 245760x^8)t^3 \\ &\quad + (-1843200x^8 - 460800x^6 - 1036800x^4 + 907200x^2 + 226800t), \\ D(x, t) &= 4096t^{12} + (129024 + 24576x^2)t^{10} + (956160 + 276480x^2 + 614400x^4)t^8 + (979200 \\ &\quad + 3456000x^2 + 614440x^4 + 81920x^6)t^6 + (5011200x^2 - 345600x^4 - 1843200x^6 + 614400x^8)t^4 \\ &\quad + (864000x^4 + 138240x^2 - 921600x^6 + 245760x^{10} + 4096x^{12}). \end{aligned} \right.$$

利用第 2 节介绍的两个模型以及第 3 节概述的训练条件和优化方法, 得到了相应的数值结果, 如图 4 所示. 图 4 进一步评估了两个模型的预测性能. 与一阶和二阶怪波相比, 三阶怪波的峰值更多, 形状和波动模式更复杂. 此外, 三阶怪波的峰高也高于一阶和二阶怪波. 因此, PINNs 显然无法进行

有效预测. 而 X-MTPINNs 则仍能较好地预测这种复杂的动力学行为, 只是在波峰的最高部分附近误差较大. 从表 2 可以看出, 与 PINNs 相比, X-MTPINNs 的精度大约高出两个数量级. 这也充分证明了本文模型的优越性.

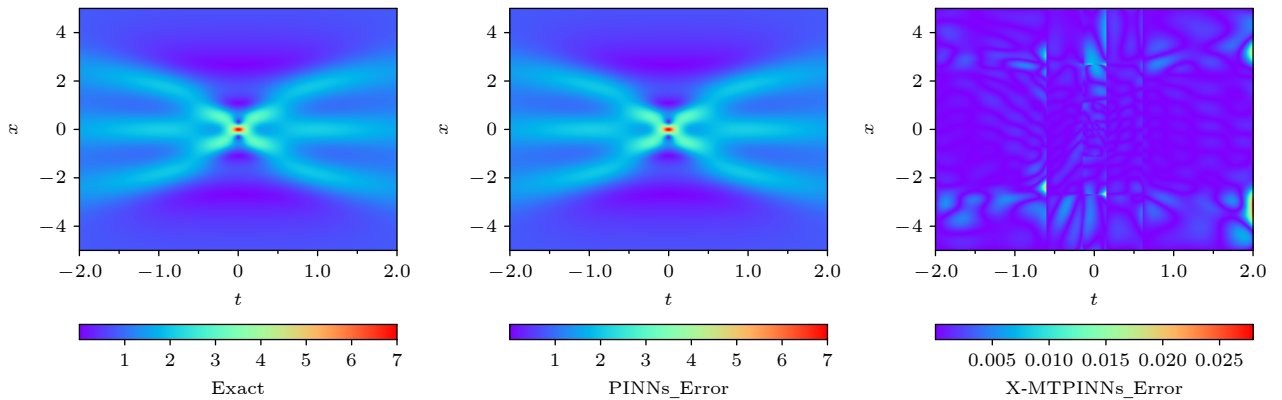


图 4 NLSE 三阶怪波的真实值和不同模型的预测误差

Fig. 4. Exact values of third-order rogue waves in the NLSE and prediction errors from different models.

表 2 两种模型对非线性薛定谔方程三阶怪波的数值预测结果

Table 2. Numerical prediction results of third-order rogue waves in the NLSE.

三阶怪波	子域	隐藏层数	神经元数	L_2 误差	训练时间/s
X-MTPINNs	[-2.00, -0.60]	4	50	1.25×10^{-3}	3065.07
	[-0.60, -0.15]	4	50	1.22×10^{-3}	2228.26
	[-0.15, 0.15]	4	50	3.89×10^{-3}	2018.29
	[0.15, 0.60]	4	50	1.31×10^{-3}	2438.84
	[0.60, 2.00]	4	50	1.84×10^{-3}	2339.57
	Global [-2, 2]	4	50	2.25×10^{-3}	3065.07
PINNs	[-2, 2]	4	50	4.94×10^{-1}	5538.17

3.2 方程参数发现

本节使用深度学习框架来研究 NLSE 的未知参数. 此时, NLSE 的具体形式如下:

$$i\hbar_t + \lambda_1 \hbar_{xx} + \lambda_2 |\hbar|^2 \hbar = 0, \quad (2)$$

$$(x, t) \in [-X, X] \times [-T, T],$$

可以定义 $f(t, x)$ 的形似如下:

$$f := i\hbar_t + \lambda_1 \hbar_{xx} + \lambda_2 |\hbar|^2 \hbar, \quad (3)$$

这里 λ_1 和 λ_2 都是可学习的参数. 同样地, 通过训练未知参数 λ_1 和 λ_2 , 未被发现的解 $\hat{h}(x, t)$ 可以通过最小化损失函数来进行近似.

接下来, 以三阶怪波求解为例进行数值实验. 神经网络初始化设置与前面所述一致.

实验结果如表 3 所列. 经典 PINNs 在本研究条件下 NLSE 的参数识别完全失效, 而即使添加 2% 和 5% 的噪声, X-MTPINNs 也能保持良好的预测能力, 这表明该模型具有很强的鲁棒性, 对复杂问题中的参数识别非常有效.

4 结论

本文结合 PINNs 算法和扩展 MTPINNs 算法, 系统探究了 NLSE 的数据驱动求解与参数反演问题. 具体而言, 将 XPINNs 中的域分解与 MTPINNs 结合起来, 在子域接口处添加伪初始点, 并行训练每个子域, 然后将结果整合形成全局解.

表 3 通过学习所得的 λ_1 和 λ_2 及其误差

Table 3. Learned λ_1 and λ_2 and their associated errors.

NPDEs	噪声	非线性演化方程	相对误差 $[\lambda_1, \lambda_2] / \%$
Correct	—	$i\hat{h}_t + 0.5\hat{h}_{xx} + \hat{h} ^2 \hat{h} = 0$	[0, 0]
[-2.00, -0.60]	0%	$\lambda_1 = 0.4999224, \lambda_2 = 0.9999491$	[0.01553, 0.00509]
	2%	$\lambda_1 = 0.5006779, \lambda_2 = 1.0010909$	[0.13558, 0.10909]
	5%	$\lambda_1 = 0.5017547, \lambda_2 = 1.0026274$	[0.35094, 0.26274]
[-0.60, -0.08]	0%	$\lambda_1 = 0.4995355, \lambda_2 = 0.9997470$	[0.09289, 0.02530]
	2%	$\lambda_1 = 0.4990919, \lambda_2 = 0.9979020$	[0.18162, 0.20980]
	5%	$\lambda_1 = 0.4983062, \lambda_2 = 0.9952767$	[0.33876, 0.47233]
[-0.08, 0.00]	0%	$\lambda_1 = 0.4907891, \lambda_2 = 0.9970491$	[1.84218, 0.29509]
	2%	$\lambda_1 = 0.4856022, \lambda_2 = 0.9931418$	[2.87955, 0.68582]
	5%	$\lambda_1 = 0.4838027, \lambda_2 = 0.9913622$	[3.23946, 0.86378]
[0.00, 0.08]	0%	$\lambda_1 = 0.4978686, \lambda_2 = 0.9941696$	[0.42627, 0.58304]
	2%	$\lambda_1 = 0.4953954, \lambda_2 = 0.9936743$	[0.92092, 0.63257]
	5%	$\lambda_1 = 0.4909737, \lambda_2 = 0.9887489$	[1.80527, 1.12511]
[0.08, 0.60]	0%	$\lambda_1 = 0.4999007, \lambda_2 = 0.9997627$	[0.01987, 0.02373]
	2%	$\lambda_1 = 0.4997132, \lambda_2 = 0.9992683$	[0.05737, 0.07317]
	5%	$\lambda_1 = 0.4974256, \lambda_2 = 0.9973150$	[0.51489, 0.26850]
[0.60, 2.00]	0%	$\lambda_1 = 0.5002598, \lambda_2 = 1.0001861$	[0.05195, 0.01861]
	2%	$\lambda_1 = 0.4997032, \lambda_2 = 0.9999705$	[0.05936, 0.00295]
	5%	$\lambda_1 = 0.4995049, \lambda_2 = 0.9998789$	[0.09903, 0.01211]
Global	0%	$\lambda_1 = 0.972926, \lambda_2 = 0.9987469$	[0.54148, 0.12531]
	2%	$\lambda_1 = 0.5035173, \lambda_2 = 0.9971460$	[0.70345, 0.28540]
	5%	$\lambda_1 = 4.9470665, \lambda_2 = 1.0050076$	[1.05867, 0.50076]

通过数值实验, 将这种扩展的 MTPINNs 方法应用于学习具有相对复杂结构的数据驱动解, 例如亮双孤子解和三阶怪波. 结果表明, 与经典 PINNs 相比, 预测精度提高了 1—2 个数量级, 同时计算时间减少了 40%. 对于逆问题, 使用扩展 MTPINNs 算法来学习无噪声、2% 和 5% 噪声条件下的 NLSE 未知参数. 训练结果表明, 即使增加了大量噪声, 扩展 MTPINNs 模型仍能准确恢复这些未知参数, 表现出很强的鲁棒性.

尽管本文取得了一些有意义的成果, 然而基于已有的对于可积系统的众多研究, 我们仍希望能寻找到更为巧妙的方法来处理偏微分方程的初边值条件, 研究其中方程的初边值条件与解的联系, 使得模型的效率以及预测精度都能得到更大的提升, 这也是我们未来工作的方向.

参考文献

- [1] Dissanayake M, Phan-Thien N 1994 *Commun. Numer. Methods Eng.* **10** 195
- [2] Lagaris I E, Likas A 1998 *IEEE Trans. Neural Networks* **9** 987
- [3] Eymard R, Gallouët T, Herbin R 2000 *Handb. Numer. Anal.* **7** 713
- [4] Zhang Y 2009 *Appl. Math. Comput.* **215** 524
- [5] Taylor C A, Hughes T J R, Zarinis C K 1998 *Comput. Methods Appl. Mech. Eng.* **158** 155
- [6] He K M, Zhang X Y, Ren S Q, Sun J 2016 *CPRV* 2016-December 770
- [7] Ding Y, Li J T, Zhang J X, Li P P, Bai H, Fang B, Fang H X, Huang K, Wang G Y, Nowell C J, Voelcker N H, Peng B, Li L, Huang W 2025 *Nat. Commun.* **16** 743
- [8] Liu J, Chen Y B, Liu K, Zhao J 2018 *32nd AAAI Conference on Artificial Intelligence* New Orleans, Louisiana, USA, February 2–7, 2018 p4865
- [9] Raissi M, Perdikaris P, Karniadakis G E 2019 *J. Comput. Phys.* **378** 686
- [10] Wang F J, Zhai Z, Zhao Z B, Di Y, Chen X F 2024 *Nat. Commun.* **15** 4332
- [11] James D, Alireza D, Soroush A 2019 *Sci. Total Environ.* **912** 168814
- [12] Xu S F, Sun Z X, Huang R F, Guo D L, Yang G W, Ju S J 2023 *AMS* **39** 322302
- [13] Chiu P H, Wong J C, Ooi C C, Dao H 2022 *Comput. Methods Appl. Mech. Eng.* **395** 114909
- [14] Yang L, Meng X H, Karniadakis G E 2021 *J. Comput. Phys.* **425** 109913
- [15] Mojgani R, Balajewicz M, Hassanzadeh P 2023 *Comput. Methods Appl. Mech. Eng.* **404** 115810
- [16] Jagtap A D, Karniadakis G E 2021 *CEUR Workshop Proc.* **296** 4
- [17] Che W J, Chen S C, Liu C, Zhao L C, Akhmediev N 2022 *Phys. Rev. A* **105** 043526
- [18] Liu C, Chen S C, Yao X K, Akhmediev N 2022 *Physica D* **433** 133192
- [19] Fang Y, Bo W B, Wang R R, Wang Y Y, Dai C Q 2022 *Chaos, Solitons Fractals* **165** 112908
- [20] Li W T, Chen Y P, Zhang Z, Yang X Y, Li B 2025 *Chin. Phys. Lett.* **42** 010202
- [21] Chen M, Wang Z 2023 *Chin. Phys. B* **32** 090504
- [22] Li J, Chen Y 2020 *Commun. Theor. Phys.* **72** 105005
- [23] Li J, Chen Y 2020 *Commun. Theor. Phys.* **72** 115003
- [24] Song J, Zhong M, Karniadakis G E, Yan Z Y 2024 *J. Comput. Phys.* **505** 112917
- [25] Wang X L, Wu Z K, Song J, Han W J, Yan Z Y 2024 *Chaos, Solitons Fractals* **180** 114509
- [26] Qiu T W, Wei G M, Song Y X, Wang Z 2025 *Appl. Math. Mech.* **46** 105
- [27] Tian S F, Li B 2023 *Acta Phys. Sin.* **72** 100202 (in Chinese) [田十方, 李彪 2023 物理学报 **72** 100202]
- [28] Li J H, Li B 2022 *Chaos, Solitons Fractals* **164** 112712
- [29] Tian S F, Li B, Zhang Z 2024 *Chin. Phys. Lett.* **41** 030201
- [30] Zhou H J, Pu J C, Chen Y 2023 *Nonlinear Dyn.* **111** 14667
- [31] Qiu W X, Si Z Z, Mou D S, Dai C Q, Li J T, Liu W 2025 *Nonlinear Dyn.* **113** 4063
- [32] Liu X M, Zhang Z Y, Liu W J 2023 *Chin. Phys. Lett.* **40** 070501
- [33] Liu H Y, Zhang Y B, Wang L 2024 *J. Syst. Sci.* **37** 494
- [34] Zhang W X, Chen Y 2025 *Nonlinear Dyn.* **113** 10211
- [35] Pu J C, Chen Y 2024 *J. Comput. Phys.* **510** 113090
- [36] Zhang J F, Dai C Q, Wang Y Y 2016 *Rogue Wave Theory and Its Applications Based on the Nonlinear Schrödinger Equation* (Beijing: Science Press) pp44–63 (in Chinese) [张解放, 戴朝卿, 王悦悦 2016 基于非线性薛定谔方程的畸形波理论及其应用 (北京: 科学出版社) 第 44—63 页]

Solving nonlinear Schrödinger equations and parameter discovery via extended mixed-training physics-informed neural networks^{*}

WANG Yuduo CHEN Jiaxin LI Biao[†]*(School of Mathematics and Statistics, Ningbo University, Ningbo 315211, China)*

(Received 1 April 2025; revised manuscript received 26 May 2025)

Abstract

In recent years, physics-informed neural networks (PINNs) have provided efficient data-driven methods for solving forward and inverse problems of partial differential equations (PDEs). However, when addressing complex PDEs, PINNs face significant challenges in computational efficiency and accuracy. In this study, we propose the extended mixed-training physics-informed neural networks (X-MTPINNs) as illustrated in the following figure, which effectively enhance the ability to solve nonlinear wave problems by integrating the domain decomposition technique of extended physics-informed neural networks (X-PINNs) in a mixed-training physics-informed neural networks (MTPINNs) framework. Compared with the classical PINNs model, the new model exhibits dual advantages: The first advantage is that the mixed-training framework significantly improves convergence properties by optimizing the handling mechanism of initial and boundary conditions, achieving higher fitting accuracy for nonlinear wave solutions while reducing the computation time by approximately 40%. And the second advantage is that the domain decomposition technique from X-PINNs strengthens the ability of the model to represent complex dynamical behaviors. Numerical experiments based on the nonlinear Schrödinger equation (NLSE) demonstrate that X-MTPINNs excel perform well in solving two bright solitons, third-order rogue waves, and parameter inversion tasks, with prediction accuracy improved by one to two orders of magnitude over traditional PINN. For inverse problems, the X-MTPINNs algorithm accurately identifies unknown parameters in the NLSE under noise-free, 2%, and 5% noisy conditions, solving the complete failure problem of NSLE parameter identification in classical PINNs in the studied scenario, thus demonstrating strong robustness.

Keywords: physics-informed neural networks, nonlinear Schrödinger equation, extended mixed-training physics-informed neural networks, domain decomposition, parameter identification

PACS: 02.30.Ik, 02.60.-x, 05.45.-a

DOI: [10.7498/aps.74.20250422](https://doi.org/10.7498/aps.74.20250422)

CSTR: [32037.14.aps.74.20250422](https://cstr.cn/32037.14.aps.74.20250422)

^{*} Project supported by the National Natural Science Foundation of China (Grant Nos. 12175111, 12235007, 12275144) and the K.C. Wong Magna Fund in Ningbo University, China.

[†] Corresponding author. E-mail: libiao@nbu.edu.cn

基于扩展混合训练物理信息神经网络的非线性薛定谔方程求解和参数发现

王宇铎 陈嘉鑫 李彪

Solving nonlinear Schrödinger equations and parameter discovery via extended mixed-training physics-informed neural networks

WANG Yuduo CHEN Jiaxin LI Biao

引用信息 Citation: *Acta Physica Sinica*, 74, 160201 (2025) DOI: 10.7498/aps.74.20250422

CSTR: 32037.14.aps.74.20250422

在线阅读 View online: <https://doi.org/10.7498/aps.74.20250422>

当期内容 View table of contents: <http://wulixb.iphy.ac.cn>

您可能感兴趣的其他文章

Articles you may be interested in

基于梯度优化物理信息神经网络求解复杂非线性问题

Solving complex nonlinear problems based on gradient-optimized physics-informed neural networks

物理学报. 2023, 72(10): 100202 <https://doi.org/10.7498/aps.72.20222381>

基于源项解耦的物理信息神经网络方法及其在放电等离子体模拟中的应用

Physics-informed neural networks based on source term decoupled and its application in discharge plasma simulation

物理学报. 2024, 73(14): 145201 <https://doi.org/10.7498/aps.73.20240343>

基于深度物理启发神经网络的微波波导器件逆设计方法

Inverse design of microwave waveguide devices based on deep physics-informed neural networks

物理学报. 2023, 72(8): 080201 <https://doi.org/10.7498/aps.72.20230031>

基于物理信息神经网络的绝热捷径动力学分析

Dynamic analysis of shortcut to adiabaticity based on physical information neural network

物理学报. 2025, 74(11): 114204 <https://doi.org/10.7498/aps.74.20250147>

基于物理信息神经网络的金属表面吸收率测量方法

Measurement method of metal surface absorptivity based on physics-informed neural network

物理学报. 2024, 73(9): 094301 <https://doi.org/10.7498/aps.73.20231453>

基于物理信息神经网络的光斑质心计算

Calculation of spot centroid based on physical informed neural networks

物理学报. 2022, 71(20): 200601 <https://doi.org/10.7498/aps.71.20220670>