

基于 AGA-PINNs 方法求解具有间断解的方程^{*}

刘博宇, 江林峰, 杨凤莲

(河海大学 数学学院, 南京 210000)

摘要: 物理信息神经网络 (physics-informed neural networks, PINNs) 是求解偏微分方程的重要工具, 在偏微分方程的数值求解中, 具有间断解的方程是目前的研究难题, PINNs 及其现有的改进算法通常无法捕捉到间断的特性, 然而在流体力学领域需要考虑许多间断问题. 针对 PINNs 处理具有间断解的方程的不足, 本文提出了自适应梯度消灭的物理信息神经网络 (adaptive gradient-annihilated PINNs, AGA-PINNs) 方法来求解具有间断解的 Burgers 方程与 Allen-Cahn 方程. 该方法利用与梯度相关的权重函数来优化损失函数, 并在训练过程中根据当前的残差分布动态调整训练点, 以逐步强化模型对间断区域的学习能力. 算例结果表明, AGA-PINNs 方法比传统的 PINNs 方法、gPINNs 方法、GA-PINNs 方法、高阶 DG 方法等对于物理量的预测精度有显著提升, 在求解 Burgers 方程与 Allen-Cahn 方程时均方误差降低了大约一个数量级, 准确地再现了 Burgers 方程中冲击波的特征与 Allen-Cahn 方程中相分离现象.

关键词: AGA-PINNs; Burgers 方程; Allen-Cahn 方程; 自适应方法

中图分类号: O241 **文献标志码:** A **DOI:** 10.21656/1000-0887.460129

Solving Equations With Discontinuous Solutions Based on the AGA-PINNs Method

Liu Boyu, Jiang Linfeng, Yang Fenglian

(School of Mathematics, Hohai University, Nanjing 210000, P.R. China)

Abstract: The physical information neural networks (PINNs) are an important tool for solving partial differential equations. In the numerical solution of partial differential equations, equations with discontinuous solutions are currently a research challenge. The PINNs and their existing improved algorithms often cannot capture the characteristics of discontinuities. However, many discontinuous problems need to be considered in the field of fluid mechanics. In response to the shortcomings of the PINNs in handling equations with discontinuous solutions, an adaptive gradient-annihilated PINNs (AGA-PINNs) method was proposed to solve the Burgers equations and the Allen-Cahn equations with discontinuous solutions. The gradient related weight functions were utilized to optimize the loss function, and the training points were dynamically adjusted based on the current residual distribution during the training process to gradually enhance the model's learning ability in discontinuous regions. The experimental results show that, the AGA-PINNs method significantly improves the accuracy of predicted physical quantities compared to the traditional PINNs, the gradient-enhanced physics-informed neural

* 收稿日期: 2025-06-25; 修订日期: 2025-07-30

基金项目: 国家自然科学基金(12271140); 中央高校基本科研业务费(B220202081)

作者简介: 刘博宇(2001—), 男, 硕士生(E-mail: 231362010004@hhu.edu.cn);

江林峰(2000—), 男, 博士生(E-mail: mathlfjiang@hhu.edu.cn);

杨凤莲(1982—), 女, 副教授, 博士(通信作者. E-mail: yangfenglilian@hhu.edu.cn).

引用格式: 刘博宇, 江林峰, 杨凤莲. 基于 AGA-PINNs 方法求解具有间断解的方程[J]. 应用数学和力学, 2026, 47(7): 895-911.

networks (gPINNs), the gradient-annihilated physics-informed neural networks (GA-PINNs) method, and the high-order deep Galerkin (DG) method. In solving the Burgers equation and the Allen-Cahn equation, the mean square errors decrease by approximately 1 order of magnitude, accurately reproducing the characteristics of shock waves in the Burgers equation and the phase separation phenomena in the Allen-Cahn equation.

Key words: AGA-PINNs; Burgers equation; Allen-Cahn equation; adaptive method

0 引言

近年来,深度学习方法在各个领域取得巨大的成功,其强大的数据处理和学习能力使其在各类复杂任务中展现出卓越的性能,如计算机视觉^[1]、自然语言处理等^[2].同时,深度学习特别是深度神经网络也广泛应用于数学及力学领域^[3].物理信息神经网络(PINNs)是一种无网格方法^[4],可以用来解决各类偏微分方程(PDE)^[5],包括但不限于随机偏微分方程、分数阶偏微分方程^[6].PINNs的核心思想是将物理定律(通常以偏微分方程的形式出现)直接融入到神经网络的损失函数中^[7],使得网络的预测结果不仅依赖于数据,还遵循已知的物理规律^[8].Hornik等^[9]将神经网络作为一种函数逼近工具来求解常微分方程和部分偏微分方程,Raissi等^[10]提出使用PINNs方法求解非线性偏微分方程.伴随PINNs的迅速发展,部分学者拓展了使用PINNs求解偏微分方程的库的种类,例如DeepXDE^[11].偏微分方程残差的梯度信息隐藏着特定的物理规律,Yu等^[12]提出了梯度增强的PINNs(gPINNs)方法,使用一种新型的损失函数提高了PINNs的准确性.

现有PINNs及其改进算法通常只关注求解具有连续解的方程^[13],然而在流体力学、材料科学等领域有诸多间断问题待解决^[14],例如流体力学中具有间断解的Burgers方程、Allen-Cahn方程等.在间断点处,物理量的导数信息常常较难捕捉,导致使用一般的PINNs方法的训练效果不好^[15].Ferrer-Sánchez等^[16]提出了梯度消灭的PINNs(GA-PINNs)方法,利用与流体变量梯度相关的因子来检测物理时空中的不连续性,有效地捕捉到相对论流体动力学方程中的传播速度和流体变量不连续点的位置.然而GA-PINNs方法使用的权重函数极易对较大残差的区域过度抑制,并且无法动态调整训练点的分布,导致间断区域训练点的密度不足,整体精度较低.

本工作将GA-PINNs方法应用于经典的具有不连续解的方程,改进了GA-PINNs方法的权重函数,使用对数形式的权重函数避免在神经网络训练过程中对间断区域的过度抑制,并结合在高残差区域添加训练点、在相对平滑区域移除相同数目训练点的自适应采样点手段,从而提出了AGA-PINNs方法来求解具有间断解的Burgers方程与Allen-Cahn方程,模拟冲击波的传播和物质浓度的演化过程.算例结果表明,AGA-PINNs方法可以准确地捕捉到间断点及其周围点的信息,相比于PINNs、gPINNs、GA-PINNs、高阶deep Galerkin(DG)等方法,在解的精度和泛化能力上有显著提升.对于模型的泛化性,本文在理论上也给出了证明.

1 GA-PINNs 方法

本文考虑如下形式的偏微分方程:

$$f\left(\mathbf{x}, t; u; \frac{\partial u}{\partial x_1}, \frac{\partial u}{\partial x_2}, \dots; \frac{\partial^2 u}{\partial x_1^2}, \frac{\partial^2 u}{\partial x_1 \partial x_2}, \dots; \frac{\partial u}{\partial t}, \frac{\partial^2 u}{\partial t^2}, \dots; \lambda\right) = 0, \quad (1)$$

其中 $\mathbf{x} \in \Omega$ 为空间域, $t \in [0, T]$ 为时间域, λ 为偏微分方程的参数, $u(\mathbf{x}, t)$ 为偏微分方程的解,在间断点处, $|u_{\mathbf{x}}| \rightarrow +\infty$.方程的初始条件为

$$f_{\text{IC}}(\mathbf{x}, t) = 0, \quad (\mathbf{x}, t) \in \Omega \times \{0\}, \quad (2)$$

方程的边界条件为

$$f_{\text{BC}}(\mathbf{x}, t) = 0, \quad (\mathbf{x}, t) \in \partial\Omega \times (0, T]. \quad (3)$$

PINNs通常使用物理信息归纳残差来解决偏微分方程^[17],但PINNs方法倾向于解决具有光滑解的问题,对于间断问题,目前大多数方法通过适当延长训练时间来求解^[18],但这会导致解过于光滑^[19].

在Burgers方程与Allen-Cahn方程的间断点处,梯度在理论上将变为无穷大,不能用微分方程来描述.因此,在不连续点处常常会引入较大的误差,在所有点处取损失的平均值无法准确地处理有间断点的偏微分方

程. GA-PINNs 方法改进了 PINNs 方法的损失函数, 通过引入一个权重函数来动态调整不同区域的损失贡献, 为临界时空位置之前和之后的点提供有效的解来鲁棒地处理不连续性, 从而应对解的梯度较大或不连续的区域.

在使用 GA-PINNs 求解偏微分方程时, 首先需构造含有可训练参数 θ 的神经网络 $\hat{u}(\mathbf{x}, t; \theta)$ 来逼近偏微分方程的真实解 $u(\mathbf{x}, t)$, 然后根据方程的物理信息、初始条件、边界条件构造损失函数训练神经网络. 在 GA-PINNs 中, 神经网络通常为全连接网络, 各部分损失分别为^[20]

$$L_R = \frac{1}{N_R} \sum_{\Omega \times (0, T]} \left| f\left(\mathbf{x}, t; \hat{u}; \frac{\partial \hat{u}}{\partial x_1}, \frac{\partial \hat{u}}{\partial x_2}, \dots; \frac{\partial^2 \hat{u}}{\partial x_1^2}, \frac{\partial^2 \hat{u}}{\partial x_1 \partial x_2}, \dots; \frac{\partial \hat{u}}{\partial t}, \frac{\partial^2 \hat{u}}{\partial t^2}, \dots; \lambda\right) \right|^2, \quad (4)$$

$$L'_R = \kappa(\hat{u}, \partial_x \hat{u}) \cdot L_R, \quad (5)$$

$$L_{IC} = \frac{1}{N_{IC}} \sum_{\Omega \times 0} |f_{IC}(\mathbf{x}, t; \hat{u})|^2, \quad (6)$$

$$L_{BC} = \frac{1}{N_{BC}} \sum_{\partial \Omega \times T} |f_{BC}(\mathbf{x}, t; \hat{u})|^2, \quad (7)$$

训练总损失为

$$L_{GA-PINNs} = w_R L'_R + w_{IC} L_{IC} + w_{BC} L_{BC}, \quad (8)$$

其中 $\kappa(\hat{u}, \partial_x \hat{u})$ 为权重函数, $L_R, L'_R, L_{IC}, L_{BC}$ 分别为 PDE 损失、改进的 PDE 损失、初始条件损失、边界条件损失, N_R, N_{IC}, N_{BC} 分别为内部采样点、初始采样点、边界采样点的数量, w_R, w_{IC}, w_{BC} 为权重系数.

近年来, 多项研究表明, GA-PINNs 方法提升了物理信息的约束能力, 并在多个物理建模问题中表现出优于 PINNs 方法及其他改进 PINNs 方法的精度和收敛性, 但该方法应用于经典的流体力学方程时, 仍有显著不足:

① GA-PINNs 方法使用的权重函数随着 $|u_x|$ 的增加迅速下降, 使用其对 PDE 损失加权会造成 PDE 损失在空间不连续的区域停用, 从而分配更多的计算资源来解析平滑的区域, 而不是努力解决不连续性. 在间断区域神经网络的损失函数仅考虑初始条件的损失和边界条件的损失, 并没有考虑到方程的物理意义, 物理特性在训练过程中被丢失, 训练早期容易振荡, 对训练的稳定性造成影响, 极易在模型未收敛时, 预测解与真实解之间的误差出现极小值.

② GA-PINNs 方法无法动态调整训练点的分布, 当解存在剧烈振荡或间断区域训练点的密度不足, 模型未能充分训练该区域, 而在残差较小的区域有大量冗余点时, 会浪费计算资源, 降低整体的效率与精度. 根据误差驱动采样理论, 残差主导区域的点密度对训练结果影响显著, 单纯的固定采样方法对训练的限制极大.

③ 虽然模型在损失函数中显式考虑 PDE 残差项, 但在训练点分布不合理的情况下, 模型优化仍然主要关注平均误差最小化, 并倾向于“牺牲”难以拟合的区域以优化整体损失, 导致残差在间断处聚集而未能消除, 影响模型的泛化能力与局部预测精度.

2 AGA-PINNs 方法

针对 GA-PINNs 方法应用于经典流体力学方程时产生的各种问题, 本文在原 GA-PINNs 框架下提出一套改进方案, 并基于此提出 AGA-PINNs 方法. 该方法对原 GA-PINNs 方法的权重函数进行修改, 防止模型对间断区域的过度抑制, 并在训练过程中根据当前的残差分布动态调整训练点, 从而提升模型的泛化能力, 并在理论上给出了证明.

图 1 展示了 AGA-PINNs 的基本原理. 首先, 将偏微分方程的空间变量 x 、时间变量 t 作为输入, 其次通过全连接网络若干隐藏层的传递, 并在训练过程中不断使用自适应方法捕捉间断点的特征, 最终输出物理量的近似解.

2.1 权重函数

由于本文所考虑的方程在间断点处 $|u_x| \rightarrow +\infty$, 使用 GA-PINNs 原有的权重函数会使高残差区域的权重迅速接近于 0, 所以为了保持模型训练的可控性, 本文构造的梯度相关的权重函数为

$$\kappa(\hat{u}, \partial_x \hat{u}) = \frac{1}{1 + \ln(1 + \alpha_u |u_x|^{\beta_u})}, \quad (9)$$

其中 α_u, β_u 为两个额外的超参数, 在间断区域有

$$\lim_{|u_x| \rightarrow +\infty} \kappa(\hat{u}, \partial_x \hat{u}) = 0. \quad (10)$$

相反地, 在空间梯度较小、解较为平滑的区域, 特别是当 $|u_x| \rightarrow 0$ 时有

$$\lim_{|u_x| \rightarrow 0} \kappa(\hat{u}, \partial_x \hat{u}) = 1. \quad (11)$$

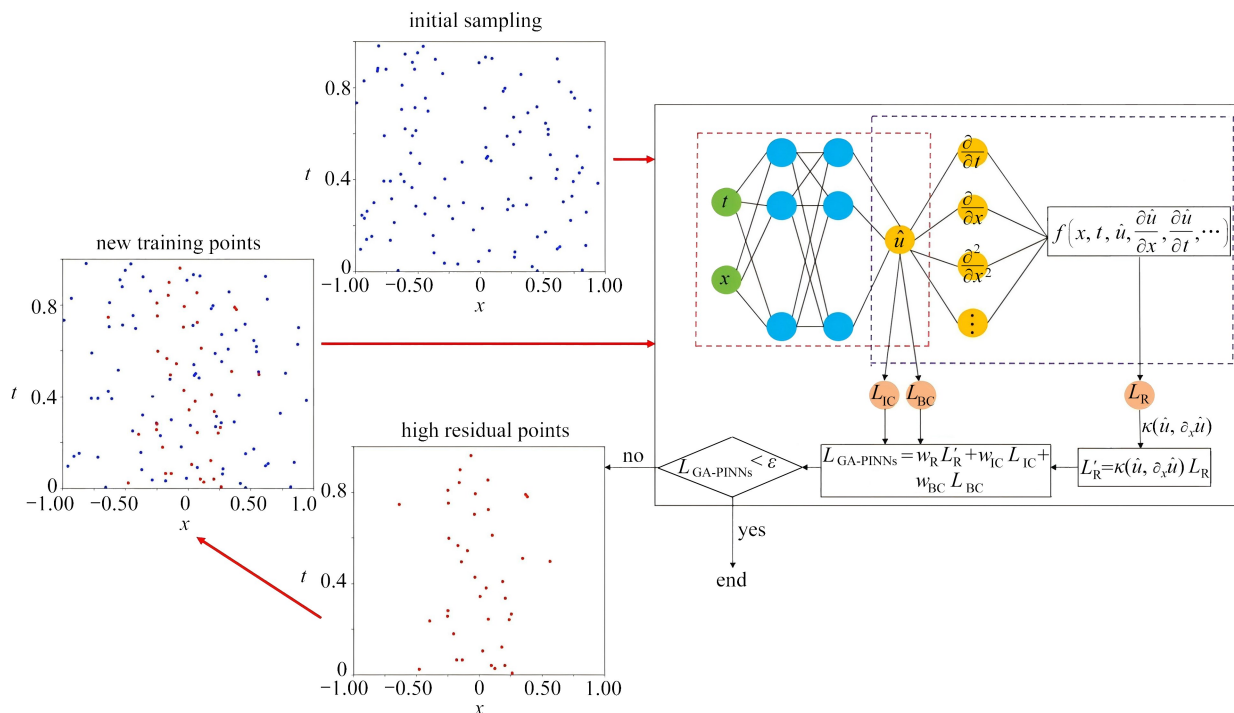


图1 AGA-PINNs 框架图

Fig. 1 The AGA-PINNs framework diagram

2.2 采样方法

在训练神经网络时, 目前较为普遍的采样方法是随机采样^[21]和均匀采样^[22], 但由于随机采样易发生点簇效应以及均匀采样点的特殊性, 对比各种统计方法优劣后, 本文选择 Sobol 采样方法^[23]. Sobol 序列具有高维均匀性, 在高维空间中, Sobol 序列可以更好地填充整个空间, 即使在每个维度上观察到不均匀, 但在整体高维空间中, Sobol 序列有更好的均匀分布, 可以有效地减少随机采样方法中常见的点簇效应. 在数学上, Sobol 采样的描述通常依赖于其 Van der Corput 序列, 对于整数 n , 其 Van der Corput 序列 $v(n)$ 的构造公式为

$$v(n) = \sum_{k=1}^n \frac{b_k}{2^k}, \quad (12)$$

其中 b_k 是 n 的二进制展开式中的第 k 位. 为了使得采样的均匀性更好, Sobol 采样通过引入方向数来调整各维度的采样, 并以准随机方式生成 2 的幂次方个采样点, 确保每个维度的采样序列具有低差异性, 从而减少物理空间中点的重叠, 并尽可能覆盖更多区域. 然而, 该方法仍然存在一定的随机性. 图 2 展示了使用随机采样、均匀采样和 Sobol 采样生成 256 个训练点的可视化结果.

2.3 自适应方法

在 PINNs 及 GA-PINNs 方法中, 残差点通常随机分布在时间及空间域中, 但使用神经网络训练时, 在间断区域需要大量的训练点才能保证在该区域训练充分. 为了捕捉到间断点的信息, 本文引入一个自适应过程^[24]来动态选择残差点, 不断更新训练过程中残差点的分布情况. 训练时, 自适应方法是在每次迭代中优先在残差较大的区域逐渐添加更多的训练点, 在解相对平滑的区域移除相同数量的训练点, 从而提升模型的泛

化能力.

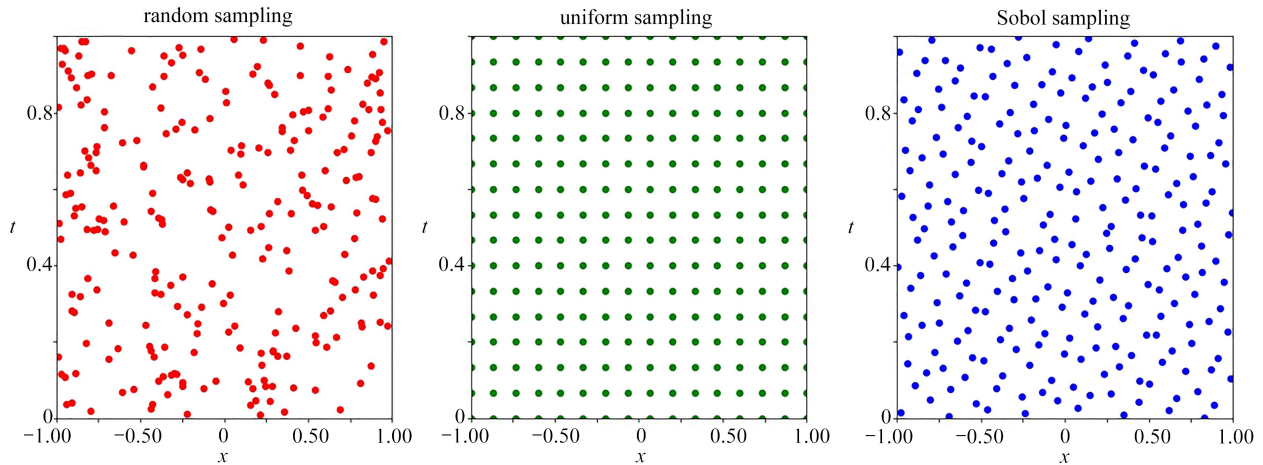


图 2 不同采样方法生成 256 个训练点的可视化结果

Fig. 2 Visualization results of 256 training points generated with different sampling methods

自适应过程主要分为三个步骤.首先生成初始训练集 $S = \{(x_1, t_1), \dots, (x_{N_R}, t_{N_R})\}$, 在时间和空间域上分别采样一定数量的点作为初始训练集.其次是选择高残差点,在每一轮训练完成后,使用 Sobol 采样生成 M 个新点 $\{(x'_1, t'_1), \dots, (x'_M, t'_M)\}$, 并使用 GA-PINNs 方法动态评估残差值,对各个点处的残差值进行排序 $R_1 > R_2 > \dots > R_M$, 筛选出残差较大的 m 个点 $\{(x''_1, t''_1), \dots, (x''_m, t''_m)\}$ 作为新的训练重点,这些高残差点通常位于模型预测精度较低的区域.最后是训练点集的更新,将筛选出的高残差点添加到原始训练点集 S 中,再将光滑部分 m 个低残差点舍去,保证每次更新时解的数量保持不变.在每一次自适应过程残差点更新后,基于当前的残差分布,模型可以集中学习难以拟合的区域,逐步提升全局解的质量.自适应过程如算法 1 所示.

算法 1 自适应方法

Step 1: 使用 GA-PINNs 方法在训练集 S 上对一定次数的迭代训练神经网络;

Step 2: 使用 Sobol 采样生成新点,并计算在这些点处的 PDE 残差;

Step 3: 找到残差最大的 m 个新点的索引,将这些点添加到训练集中,移除训练集中残差最小的 m 个点;

Step 4: Step 1—Step 3 重复 n 次.

下面给出自适应方法对模型泛化能力提升的证明.设时空域 $D = \Omega \times [0, T] \subset R^l, Y = L^p(D; R^m), Y^* \subset Y$ 为范数 $\|\cdot\|_{Y^*}$ 意义下的闭子空间,构造映射 $h: D \mapsto R^m$, 使得 $h \in Z^* \subset Y^*$, 考虑近似积分

$$h^* = \int_D h(y) dy, \quad (13)$$

其中 dy 为 $\tilde{l}$ 维 Lebesgue 测度, $\tilde{l} = l + 1$. 由于本文使用的采样方法为低差异序列,所以满足求积规则

$$\bar{h} = \sum_{i=1}^N w_i h(y_i), \quad (14)$$

其中 w_i 为权重, $y_i \in D$ 为求积点,并满足 Koksma-Hlawka 不等式:

$$|h^* - \bar{h}| \leq V_{HK}(h) D_N^*(\{y_i\}), \quad (15)$$

其中 $V_{HK}(h)$ 是 h 的 Hardy-Krause 变体, D_N^* 为采样点的星差异.由于 Sobol 采样生成的序列为低差异序列,所以对于初始采样序列, $V_{HK}(h)$ 是广义的变差度量,满足

$$V_{HK}(h) \leq C(\|h\|), \quad (16)$$

即 $V_{HK}(h)$ 是有界的,且采样点的星差异 $D_N^*(\{y_i\})$ 满足

$$D_N^*(\{y_i\}) \leq C' \frac{(\log N)^{\tilde{l}}}{N}, \quad (17)$$

其中 C' 为常数.在第 k 轮迭代中,我们替换 $m \ll N$ 个点,新采样的 m 个点也是通过 Sobol 采样生成,因此我们可以将新的点集视为扰动下的 Sobol 序列,根据 Sobol 序列在空间中容错的特性,点集的星差异度增长缓慢,

所以

$$D_N^*(\{y_i\}^k) \leq D_N^*(\{y_i\}^{k-1}) + \varepsilon_m = D_N^*(0) + \sum_{j=1}^k \varepsilon_m^{(j)} \leq D_N^*(0) + k \cdot \varepsilon_{\max}, \quad (18)$$

其中 $\varepsilon_{\max} \sim O\left(\frac{(\log N)^{\tilde{l}}}{N}\right)$, 因此有

$$|h^* - \tilde{h}| \leq C(\|h\|_{Z^*}, \tilde{l}) N^{-1} \quad (19)$$

满足正交误差是有界的。

设 $u(x, t)$ 为方程(1)的唯一真实解, 满足稳定性条件, $\hat{u}(x, t; \theta)$ 为神经网络近似解, 训练数据是基于求积规则式(14)获得的点集 S , 且设函数

$$J(\theta) = \sum_{n=1}^N w_n |R_\theta(y_n)|^p, \quad (20)$$

其中 w_n 为权重, $R_\theta(y_n)$ 为 PDE 残差. 在神经网络中, 需要寻找具有正则性的 $J(\theta)$ 达到全局最小或局部极小值时的参数

$$\theta^* = \arg \min (J(\theta) + \lambda_{\text{reg}} J_{\text{reg}}(\theta)), \quad (21)$$

其中 λ_{reg} 为平衡 $J(\theta)$ 与正则化项的因子, 此时神经网络的近似解为 $\hat{u}(x, t; \theta^*)$, 正交误差满足式(19), 训练误差 ε_T 为

$$\varepsilon_T = \left(\sum_{n=1}^N w'_n |R_{\theta^*}(y_n)|^p \right)^{1/p}, \quad (22)$$

其中 w'_n 为权重, 则可以得到泛化误差 ε_G 满足

$$\varepsilon_G \leq C_1 \varepsilon_T + C_1 C^{1/p} (\|h\|_{Z^*}, \tilde{l}) N^{-1/p}, \quad (23)$$

其中 C_1 为稳定性常数^[25]. 根据式(22), 训练误差与方程的残差成正比, 使用自适应方法, 整体及间断区域方程的残差均降低, 则神经网络的训练误差也降低, 而泛化误差的上界由训练误差控制, 所以使用自适应方法, 模型的泛化性能得到提升。

3 算 例

本节展示了使用 AGA-PINNs 方法求解具有间断解的 Burgers 方程与 Allen-Cahn 方程的算例结果, 并与 PINNs、gPINNs、GA-PINNs、高阶 DG 等方法进行了对比。

为了评估模型的准确性, 在算例中使用均方误差 (δ_{MSE}) 来评估模型的精度, 它表示模型预测得到的解与真实解之间的差异, δ_{MSE} 定义为

$$\delta_{\text{MSE}} = \frac{1}{N_x N_t} \sum_{i=1}^{N_x} \sum_{j=1}^{N_t} [u(x_i, t_j) - \hat{u}(x_i, t_j; \theta)]^2, \quad (24)$$

其中 N_x, N_t 分别为空间和时间域中测试点的个数。

本节中的所有算例均采用自动微分 (AD) 技术计算网络输出 $\hat{u}(x, t; \theta)$ 关于输入 x 和 t 的偏导数, 使其满足偏微分方程的约束, 来计算偏微分方程的残差. 使用 Adam 优化器来最小化损失函数, 从而不断更新网络的参数 θ , 使得预测解更加符合偏微分方程和相应的初始、边界条件。

3.1 超参数的调整

权重函数中超参数 α_u, β_u 的选取尤为重要, 这直接影响着 PDE 损失的大小, 进而影响模型的预测效果. 令式(9)中 $u_x = x$, 则 $\kappa(x) = 1/(1 + \ln(1 + \alpha |x|^\beta))$, $\kappa(x)$ 随超参数的增加呈下降趋势, 但 β 对于 $\kappa(x)$ 的影响更大, 这是因为 β 是 x 的指数. 所以在调整超参数 α_u 与 β_u 时, 首先初始化 $\alpha_u = \beta_u = 1$, 固定 α_u 的值, 并以 0.05 为步长逐渐增大 β_u 的值, 得到不同 β_u 值时 AGA-PINNs 预测物理量的均方误差如图 3 所示, 当 $\beta_u = 1.25$ 时, 模型预测物理量的均方误差都是最低的. 因此, 本文选取的权重函数的超参数为: $\alpha_u = 1, \beta_u = 1.25$ 。

3.2 Burgers 方程

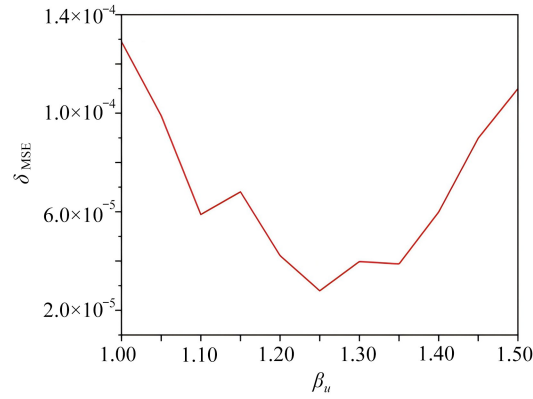
考虑如下 Burgers 方程:

$$\frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} = v \frac{\partial^2 u}{\partial x^2}, \quad x \in [-1, 1], t \in [0, 1], \quad (25)$$

其中 $v = \frac{0.01}{\pi}$, 方程的初始条件和边界条件为

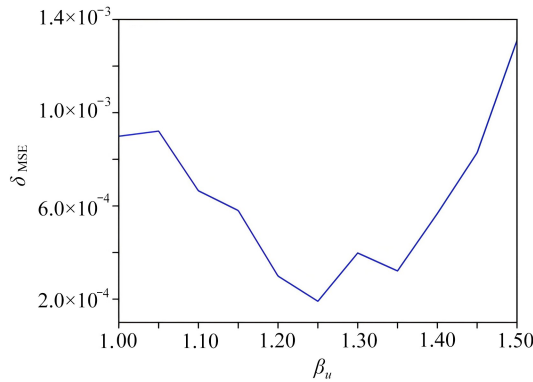
$$u(x, 0) = -\sin(\pi x), \quad (26)$$

$$u(-1, t) = u(1, t) = 0. \quad (27)$$



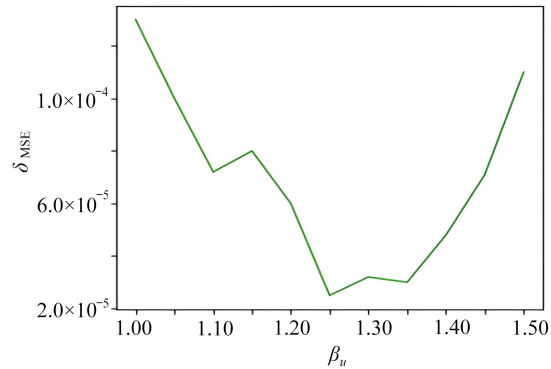
(a) Burgers 方程

(a) The Burgers equation



(b) Allen-Cahn 方程

(b) The Allen-Cahn equation



(c) Burgers 方程的 Riemann 问题

(c) The Riemann problem of the Burgers equation

图 3 不同 β_u 值时 AGA-PINNs 预测物理量的均方误差

Fig. 3 δ_{MSE} of predicting physical quantities using AGA-PINNs with different β_u values

Burgers 方程通常用于描述流体中的波动现象^[26].在 Burgers 方程的演化过程中,间断位置可能会形成冲击波,这是由于非线性项 uu_x 的作用导致的,并且对于具有不同初始条件和边界条件的 Burgers 方程,间断位置也可能会有所不同.本文考虑的 Burgers 方程的精确解如图 4 所示.

本文求解 Burgers 方程时,神经网络使用 4 层隐藏层,每层有 36 个神经元,激活函数使用 tanh 函数,并采用全连接网络,在该算例中使用 Adam 优化器,并固定学习率为 10^{-5} .

使用 GA-PINNs、PINNs、gPINNs、生成对抗网络(GAN)、基于注意力机制的 PINNs 方法(PINNsFormer)、高阶 DG 方法训练得到的损失及均方误差如表 1 所示.使用 PINNs、gPINNs 方法预测解的精度极低,PINNs 方法仅适用于预测较为平滑的区域,尽管 gPINNs 方法在损失函数中引入了物理量的梯度信息,但在 Burgers 方程的间断点处,速度的空间导数趋向于无穷,导致 gPINNs 损失函数的值也趋向于无穷,神经网络难以有效反向传播.考虑到 (x, t) 域内的采样,由于 Sobol 采样需要生成以 2 为基数的训练点,本文将采用 $N_R = 256 \times 256$ 个训练点.同样地,在对初始点与边界点进行采样时,也生成 $N_{\text{IC}} = N_{\text{BC}} = 256$ 个训练点.

该 Burgers 方程在 $x = 0$ 处十分陡峭,所以在该区域附近应该有更多的残差点.对于 AGA-PINNs 方法,首先使用 Sobol 采样方法生成初始采样点,然后在训练过程中逐渐添加残差点,并在残差较低的区域舍弃部分

残差点,共进行了5次实验,每次自适应过程分别添加10,20,30,40,50个残差点,即算法1中 m 分别取10,20,30,40,50,初始训练之后,在 $x=0$ 附近的区域具有最大的PDE残差,如图5所示.自适应方法自动在残差最大的区域附近添加更多的点,得到均方误差如图6所示,训练完成时的损失及 δ_{MSE} 如表2所示.模型在自适应添加30个点时,得到的均方误差最小,为 1.31×10^{-6} ,比普通的GA-PINNs方法降低了大约一个数量级,得到的预测解如图7所示.使用AGA-PINNs方法得到的预测解更加逼近真实解,并有效捕捉到了方程中间断点的位置.

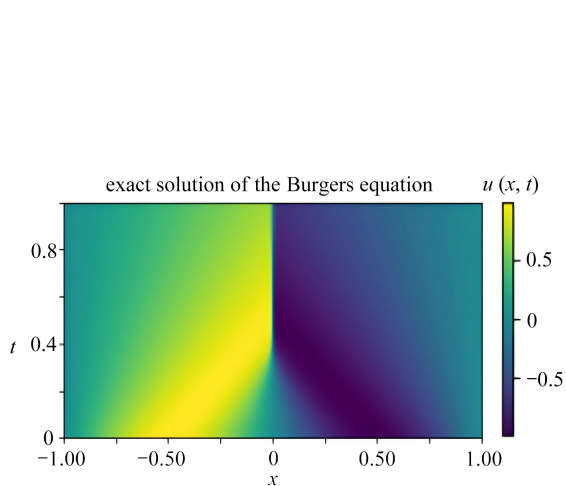


图4 Burgers方程精确解

Fig. 4 The exact solution of the Burgers equation

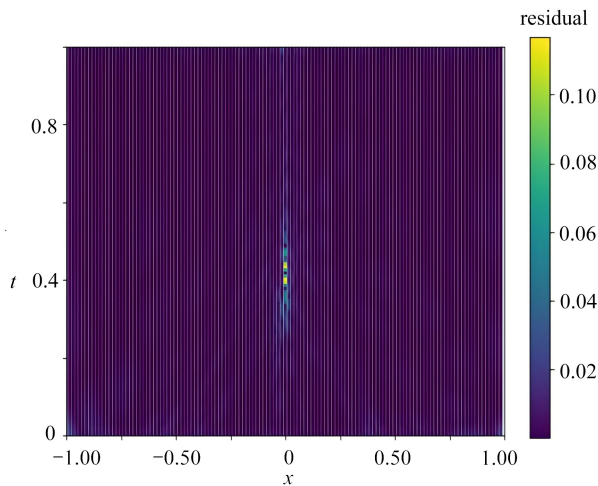


图5 求解 Burgers 方程初始训练后 PDE 残差

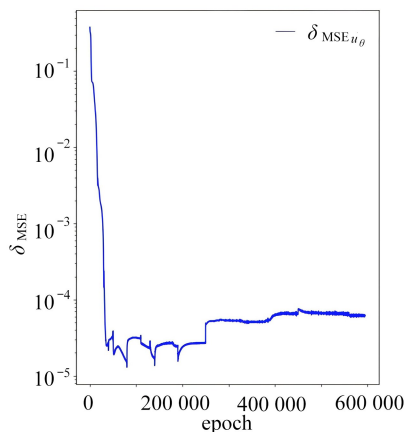
Fig. 5 PDE residuals after initial training in solving the Burgers equation

注 为了解释图中的颜色,读者可以参考本文的电子网页版本,后同。

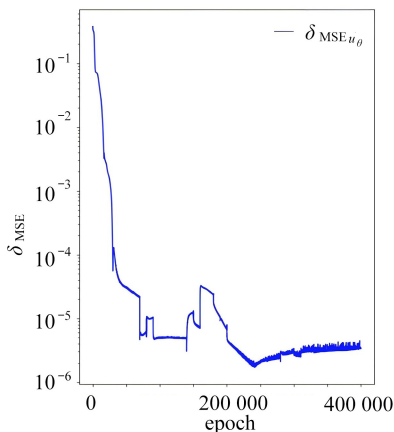
表1 不同方法求解 Burgers 方程损失及 δ_{MSE}

Table 1 Losses and δ_{MSE} in solving the Burgers equation with different methods

method	loss	δ_{MSE}
GA-PINNs	1.06×10^{-6}	6.23×10^{-5}
PINNs	1.49×10^{-5}	8.10×10^{-4}
gPINNs	7.48×10^{-5}	6.01×10^{-3}
GAN	4.82×10^{-6}	9.17×10^{-5}
PINNsFormer	9.58×10^{-6}	2.62×10^{-4}
high-order DG	—	7.99×10^{-5}



(a) $m = 10$



(b) $m = 20$

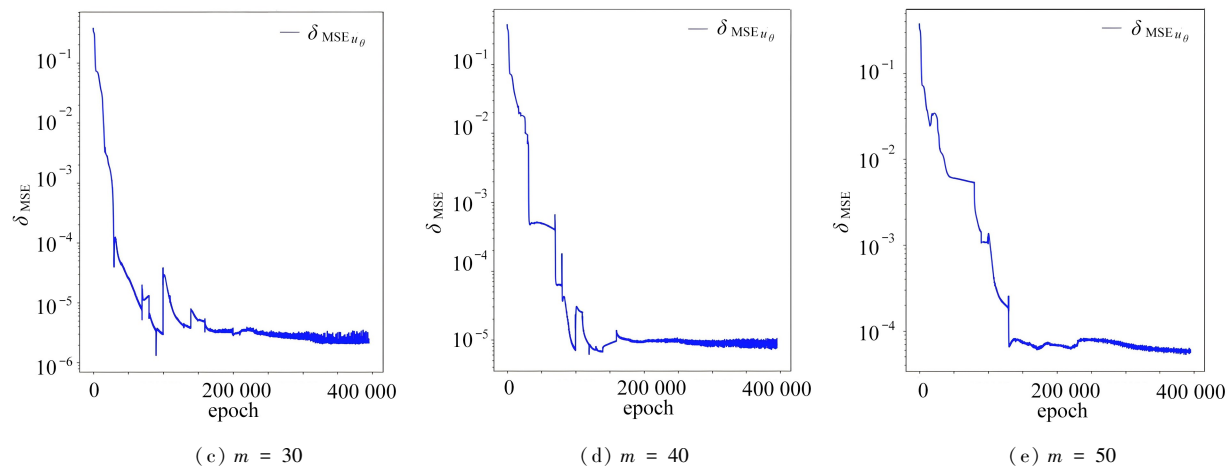


图 6 不同 m 值 AGA-PINNs 求解 Burgers 方程均方误差

Fig. 6 δ_{MSE} in solving the Burgers equation using AGA PINNs with different m values

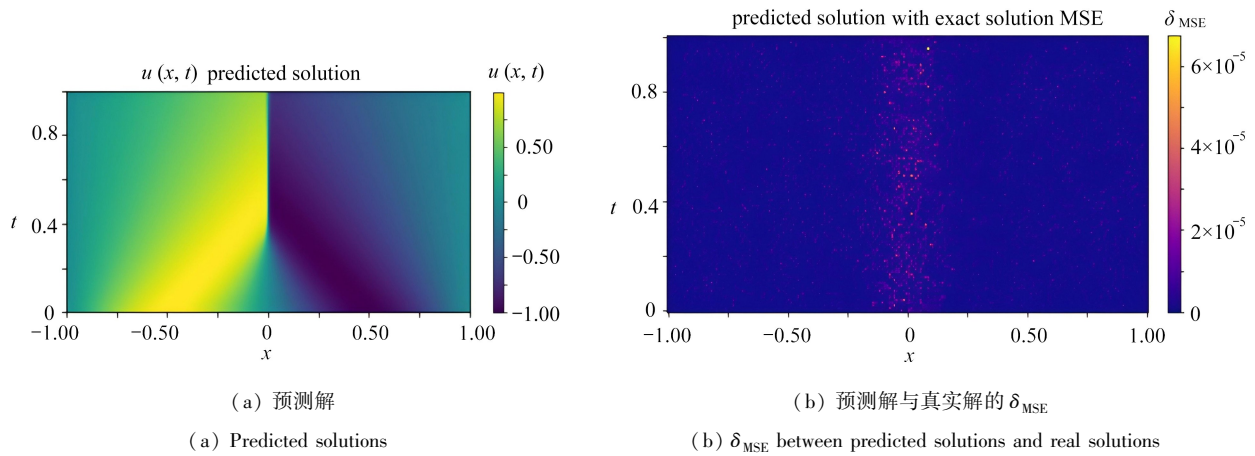


图 7 $m = 30$ 时 AGA-PINNs 求解 Burgers 方程预测结果

Fig. 7 Predicting results of AGA-PINNs with $m = 30$ for solving Burgers equation

表 2 不同 m 值 AGA-PINNs 求解 Burgers 方程损失及 δ_{MSE}

Table 2 Losses and δ_{MSE} in solving the Burgers equation with the AGA-PINNs for different m values

number of adaptive points	loss	δ_{MSE}
10	1.08×10^{-6}	7.13×10^{-5}
20	6.06×10^{-7}	3.19×10^{-6}
30	3.71×10^{-7}	1.31×10^{-6}
40	8.98×10^{-7}	8.02×10^{-6}
50	1.03×10^{-6}	5.55×10^{-5}

为验证 Sobol 采样的有效性,在算例中采用 $m = 30$ 的 AGA-PINNs 方法,分别利用均匀采样和随机采样生成相同数目的训练点进行训练,并与 Sobol 采样方法的实验结果进行对比.使用均匀采样的 AGA-PINNs 方法预测的结果如图 8(a) 所示,模型在 2.1×10^5 个周期达到收敛,均方误差为 9.01×10^{-4} ;使用随机采样的 AGA-PINNs 方法预测的结果如图 8(b) 所示,模型在 2.3×10^5 个周期达到收敛,均方误差为 2.17×10^{-3} .使用 Sobol 采样方法,模型预测得到的解与真实解之间的差异在整个空间和时间域上分布得更加均匀,并且能精确捕捉到间断点的信息,而使用均匀采样与随机采样,模型在间断点处表现较差,未能准确逼近方程的解.

3.3 Allen-Cahn 方程

考虑如下 Allen-Cahn 方程:

$$\frac{\partial u}{\partial t} = 0.001 \frac{\partial^2 u}{\partial x^2} + 5(u - u^3), \quad x \in [-1, 1], t \in [0, 1],$$

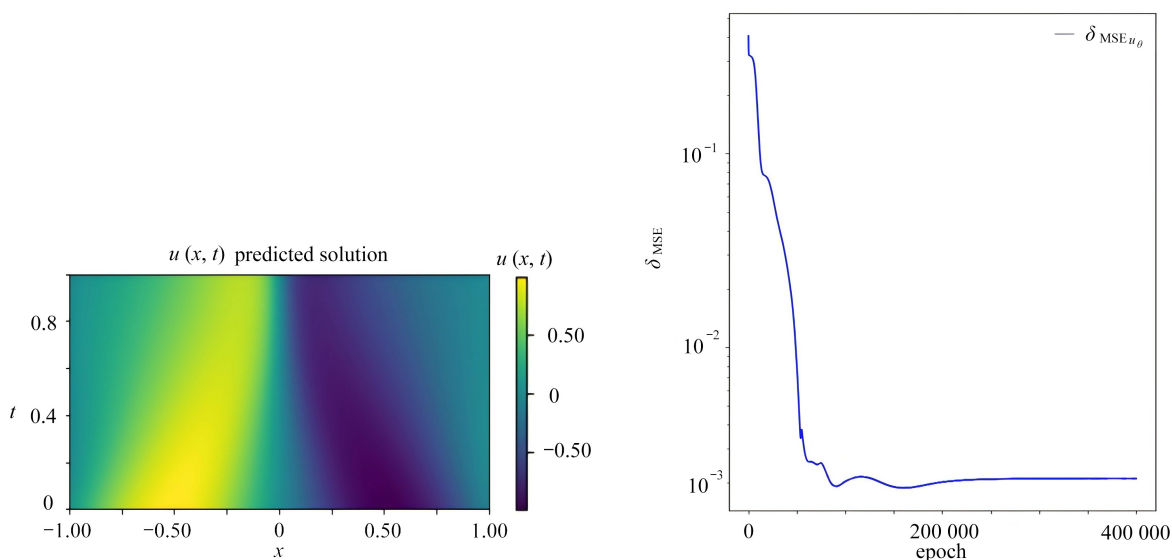
(28)

方程的初始条件和边界条件为

$$u(x, 0) = x^2 \cos(\pi x), \quad (29)$$

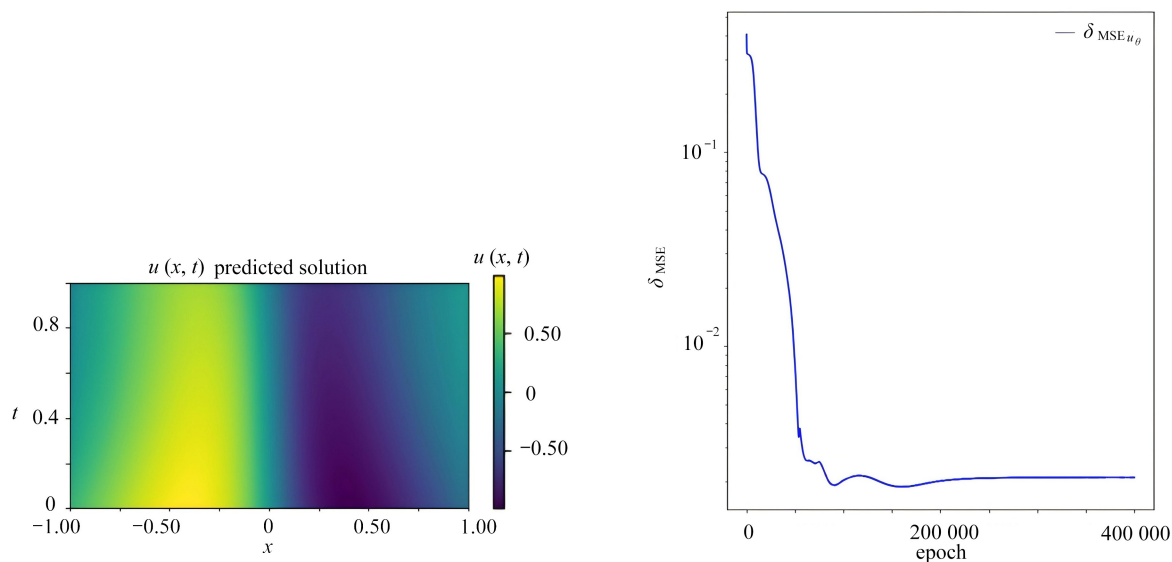
$$u(-1, t) = u(1, t), \quad u_x(-1, t) = u_x(1, t). \quad (30)$$

Allen-Cahn 方程通常用于描述相变界面中界面演化,例如液体到气体的相变、晶体生长等,其特点是在相变界面两侧值是截然不同的,在相变界面处方程有间断解^[27].本文考虑的 Allen-Cahn 方程的精确解如图 9 所示.



(a) 均匀采样 AGA-PINNs 方法

(a) The uniform sampling with the AGA-PINNs method



(b) 随机采样 AGA-PINNs 方法

(b) The random sampling with the AGA-PINNs method

图 8 不同采样方法 AGA-PINNs 求解 Burgers 方程预测结果

Fig. 8 Predicted results of the Burgers equation using the AGA-PINNs with different sampling methods

求解 Allen-Cahn 方程时,神经网络使用 4 层隐藏层,每层有 64 个神经元,激活函数使用 tanh 函数,并采用全连接网络,在该算例中使用 Adam 优化器,并固定学习率为 10^{-4} .

数值计算时在 (x, t) 域内采样 $N_R = 256 \times 256$ 个训练点,初始与边界采样 $N_{IC} = N_{BC} = 256$ 个训练点.使用 PINNs、gPINNs、GA-PINNs、GAN、PINNsFormer、高阶 DG 方法得到的损失及均方误差如表 3 所示,GA-PINNs

方法与高阶 DG 方法相比于其余几种方法预测方程的解时实现了更低的均方误差.对于 AGA-PINNs 方法,初始训练后,残差分布如图 10 所示,方程在 $x = -0.5$ 及 $x = 0.5$ 处存在不连续,应该添加更多的残差点,共进行了 5 次实验,每次自适应过程分别添加 10,20,30,40,50 个残差点,并在残差较小的区域分别舍弃相应个数的残差点,得到均方误差如图 11 所示,训练完成时的损失及 δ_{MSE} 如表 4 所示.模型在自适应添加 30 个点时,得到的均方误差最小,为 3.52×10^{-5} ,预测解如图 12 所示.

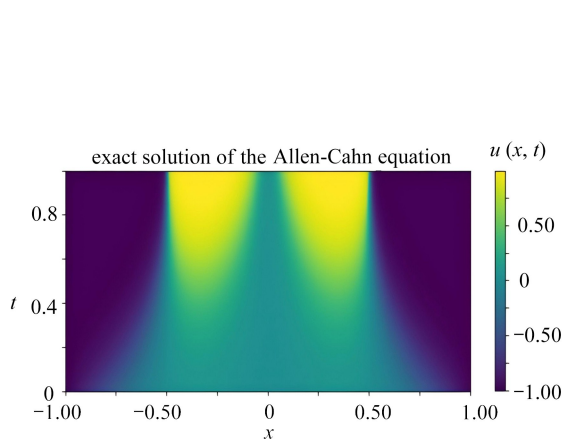


图 9 Allen-Cahn 方程精确解

Fig. 9 The exact solution of the Allen-Cahn equation

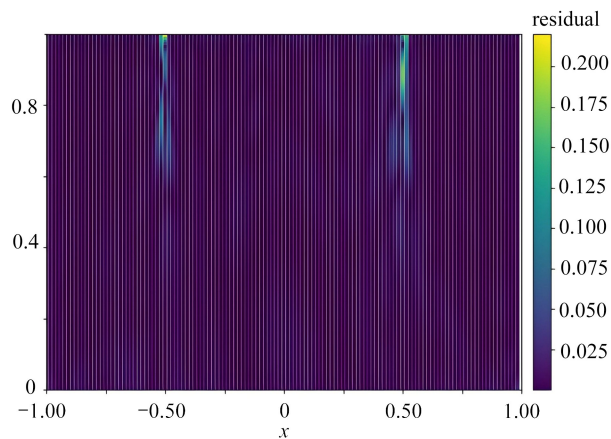


图 10 求解 Allen-Cahn 方程初始训练后 PDE 残差

Fig. 10 PDE residuals after initial training in solving the Allen-Cahn equation

表 3 不同方法求解 Allen-Cahn 方程损失及 δ_{MSE} Table 3 Losses and δ_{MSE} in solving the Allen-Cahn equation with different methods

method	loss	δ_{MSE}
GA-PINNs	9.18×10^{-6}	7.88×10^{-4}
PINNs	8.23×10^{-5}	2.37×10^{-3}
gPINNs	1.75×10^{-5}	8.05×10^{-4}
GAN	1.92×10^{-5}	9.39×10^{-4}
PINNsFormer	6.46×10^{-5}	1.69×10^{-3}
high-order DG	—	3.22×10^{-4}

3.4 Burgers 方程的 Riemann 问题

考虑如下无黏性 Burgers 方程^[28]

$$\frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} = 0, \quad x \in [-1, 2], t \in [0, 1], \quad (31)$$

其具有如下分段常数的初始条件

$$u(x, 0) = \begin{cases} u_L, & x \leq 0, \\ u_R, & x > 0, \end{cases} \quad (32)$$

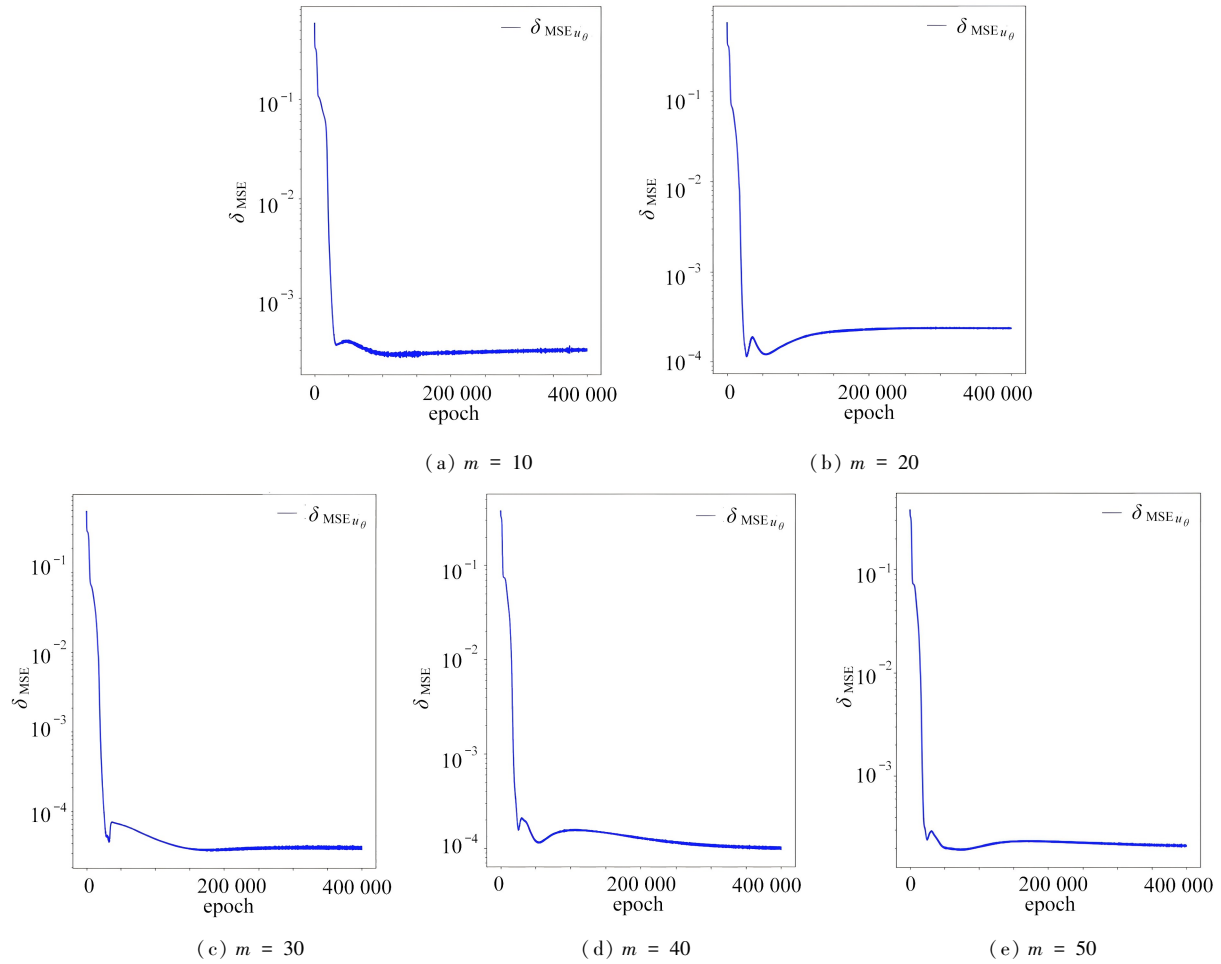
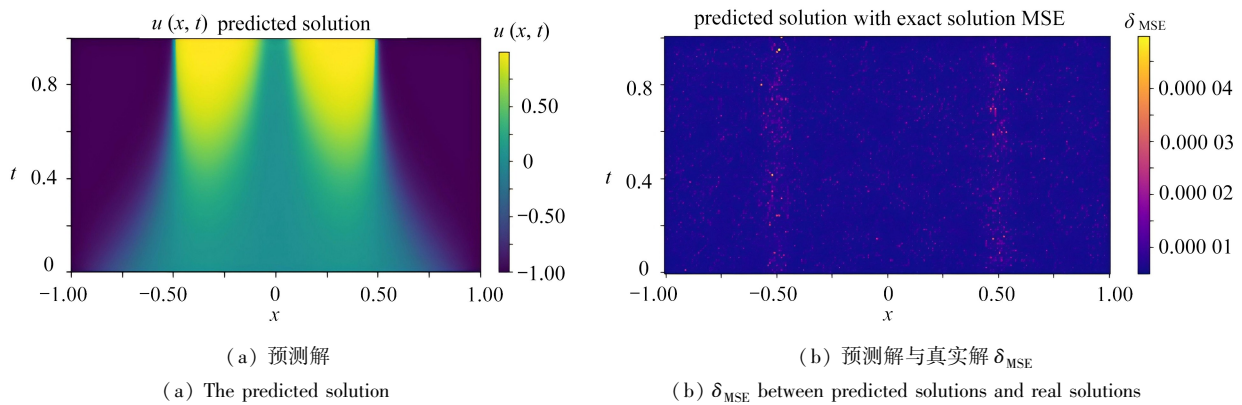
其中 u_L 和 u_R 是左右两侧的常数初始状态,本算例仅考虑 $u_L > u_R$ 的情况,且 $u_L = 2, u_R = 1$.在该情况下,两个波会相互撞击,形成激波,即解中存在一个不连续点,激波速度由 Rankine-Hugoniot 条件给出,为

$$s = \frac{1}{2}(u_L + u_R) = 1.5, \quad (33)$$

方程的精确解为

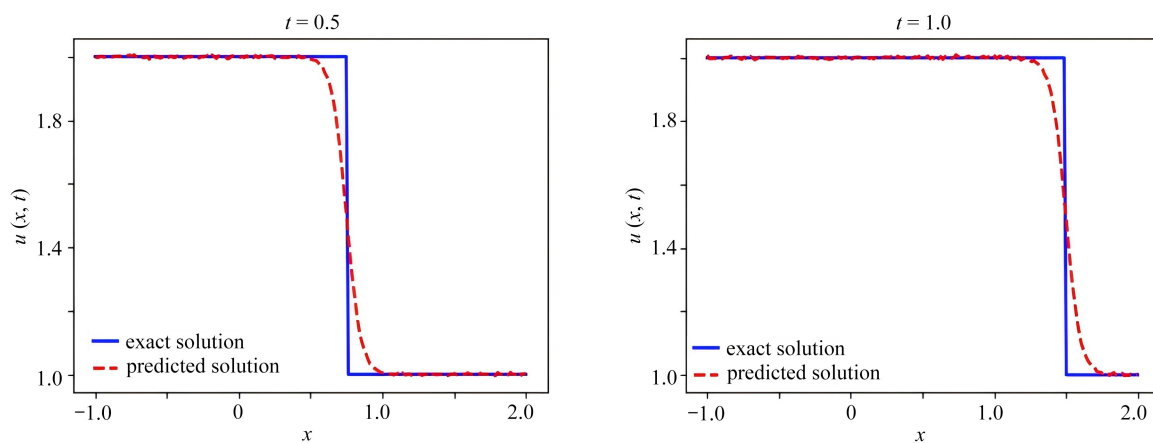
$$u(x, t) = \begin{cases} 2, & x \leq 1.5t, \\ 1, & x > 1.5t. \end{cases} \quad (34)$$

求解 Burgers 方程的 Riemann 问题时,神经网络使用 4 层隐藏层,每层有 20 个神经元,激活函数使用 tanh 函数,并采用全连接网络,在该算例中使用 Adam 优化器,并固定学习率为 10^{-4} .

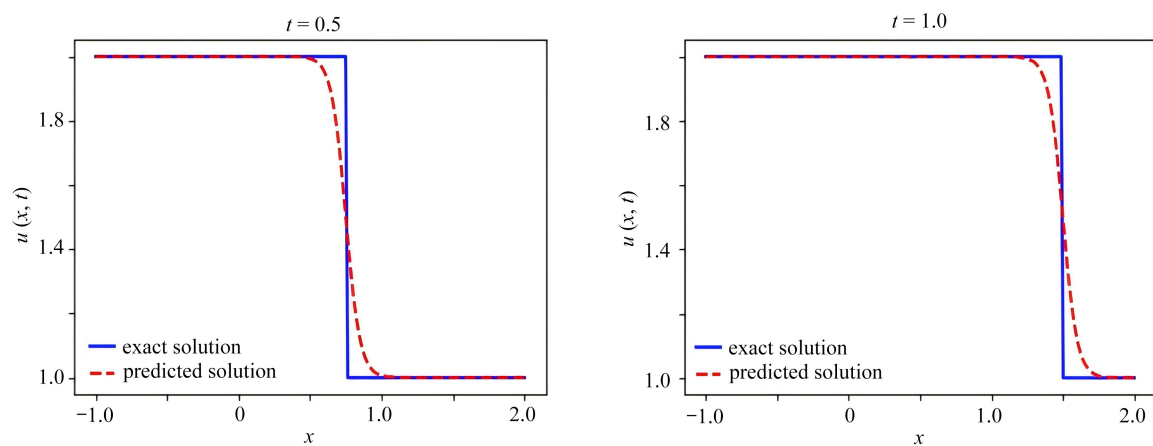
图 11 不同 m 值 AGA-PINNs 求解 Allen-Cahn 方程均方误差Fig. 11 δ_{MSE} in solving the Allen-Cahn equation using the AGA-PINNs with different m values图 12 $m = 30$ 时 AGA-PINNs 求解 Allen-Cahn 方程预测结果Fig. 12 Predicted results with the AGA-PINNs for $m = 30$ in solving the Allen-Cahn equation表 4 不同 m 值 AGA-PINNs 求解 Allen-Cahn 方程损失及 δ_{MSE} Table 4 Losses and δ_{MSE} in solving the Allen-Cahn equation with the AGA-PINNs for different m values

number of adaptive points	loss	δ_{MSE}
10	1.88×10^{-6}	3.19×10^{-4}
20	8.96×10^{-7}	2.57×10^{-4}
30	2.02×10^{-7}	3.52×10^{-5}
40	6.47×10^{-7}	1.01×10^{-4}
50	9.25×10^{-7}	2.41×10^{-4}

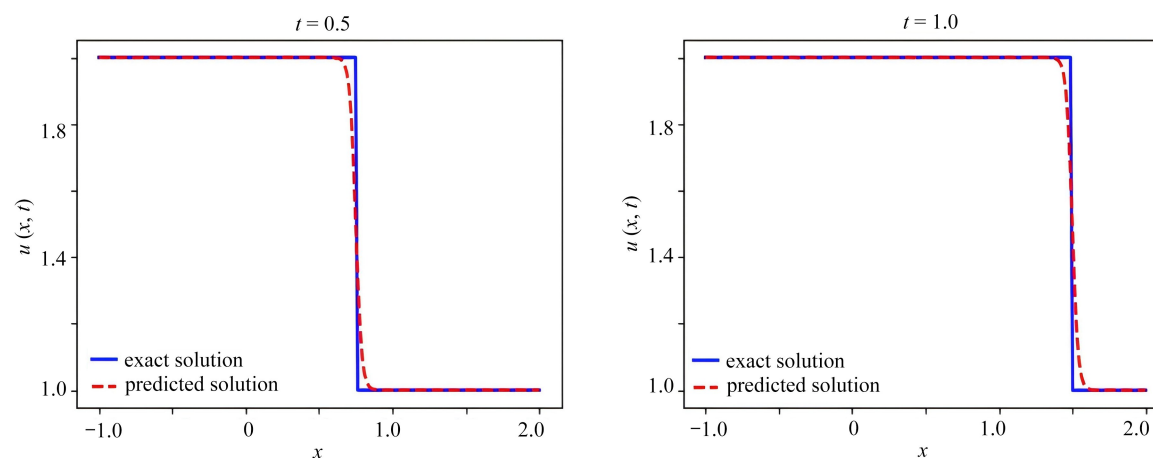
数值计算时在 (x, t) 域内采样 $N_R = 256 \times 256$ 个训练点, 初始采样 $N_{IC} = 256$ 个训练点. 使用 PINNs、gPINNs、GA-PINNs、GAN、PINNsFormer、高阶 DG 方法得到的 $t = 0.5$ 及 $t = 1.0$ 时的预测解如图 13 所示, 使用 PINNs、gPINNs、GA-PINNs、GAN、PINNsFormer、高阶 DG 方法得到均方误差分别为 7.65×10^{-2} , 2.77×10^{-2} , 3.89×10^{-3} , 3.25×10^{-3} , 8.49×10^{-4} , 1.61×10^{-3} , 且由图 13 观察到, PINNs 方法与 gPINNs 方法在间断点处预测解与真实解差异极大.



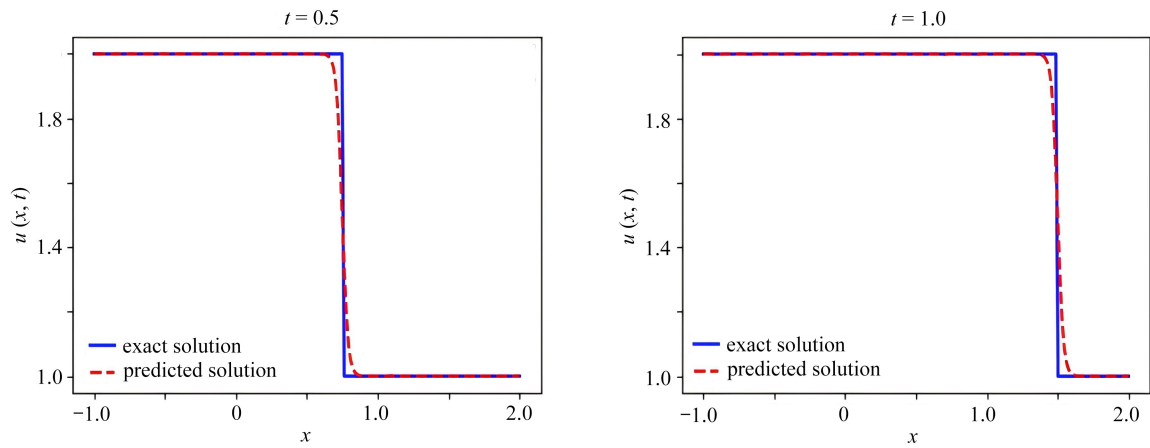
(a) PINNs



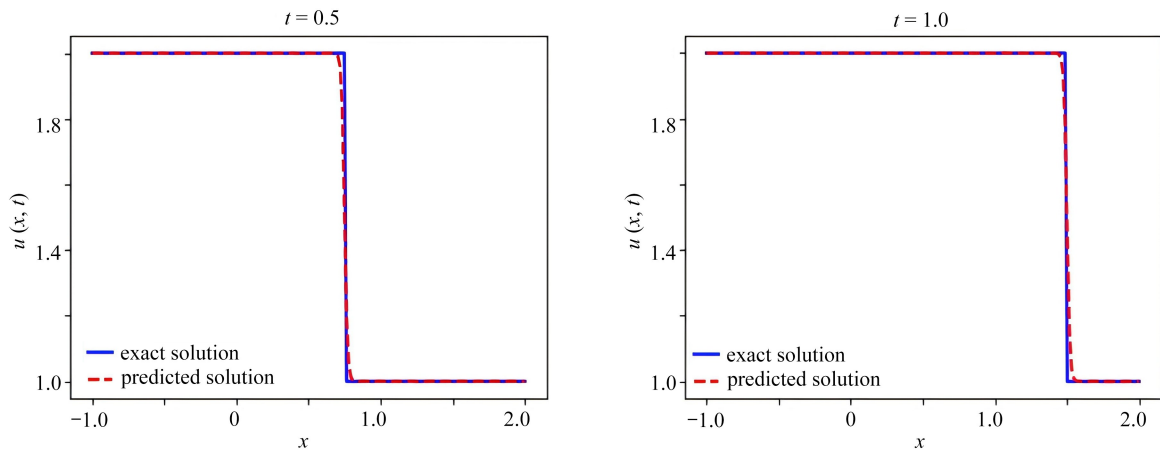
(b) gPINNs



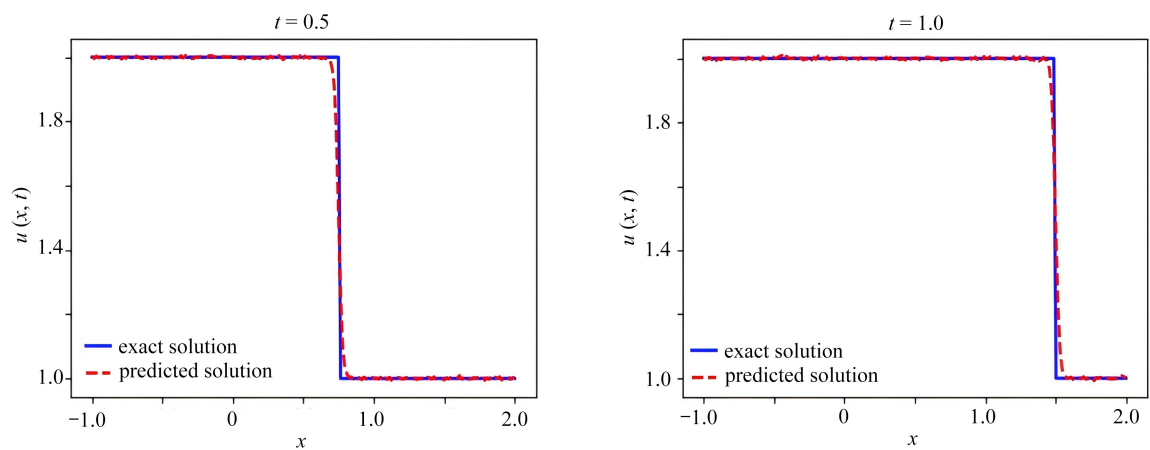
(c) GA-PINNs



(d) GAN



(e) PINNsFormer



(f) DG

图 13 不同方法求解 Burgers 方程的 Riemann 问题的预测结果

Fig. 13 Prediction results of the Burgers equation for the Riemann problem solved with different methods

对于 AGA-PINNs 方法,初始训练后,方程在 $x = 1.5$ 处存在不连续,应该添加更多的残差点,共进行了 5 次实验,每次自适应过程分别添加 10, 20, 30, 40, 50 个残差点,并在残差较小的区域分别舍弃相应个数的残差点,得到训练完成时的损失及 δ_{MSE} 如表 5 所示.模型在自适应添加 30 个点时,得到的均方误差最小,为 7.19×10^{-5} ,预测解如图 14 所示.

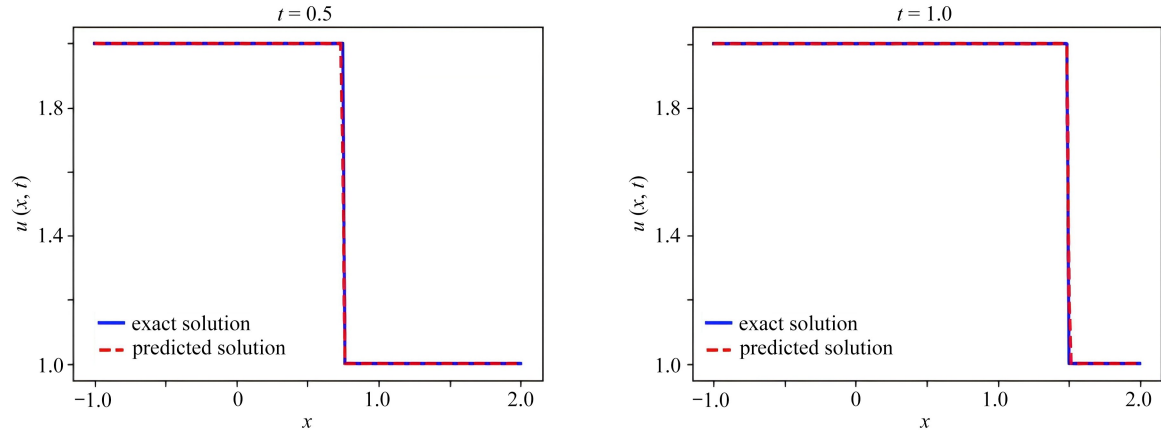


图 14 $m = 30$ 时 AGA-PINNs 求解 Burgers 方程的 Riemann 问题的预测结果

Fig. 14 Predicting results of AGA-PINNs for $m = 30$ in solving the Burgers equation of the Riemann problem

表 5 不同 m 值 AGA-PINNs 求解 Burgers 方程的 Riemann 问题的损失及 δ_{MSE}

Table 5 Losses and δ_{MSE} in solving the Burgers equation of the Riemann problem with the AGA-PINNs for different m values

number of adaptive points	loss	δ_{MSE}
10	7.49×10^{-6}	8.12×10^{-4}
20	1.18×10^{-6}	6.44×10^{-4}
30	4.26×10^{-7}	7.19×10^{-5}
40	6.58×10^{-7}	9.83×10^{-5}
50	9.32×10^{-7}	3.57×10^{-4}

4 结 束 语

本文提出一种 AGA-PINNs 方法来求解具有间断解的 Burgers 方程和 Allen-Cahn 方程,该方法不仅在损失函数中添加了与梯度相关的权重函数处理陡梯度,而且采用在方程残差较大的区域自适应调整采样点的手段提升模型的泛化性.数值算例结果表明,AGA-PINNs 方法在求解间断 Burgers 方程和 Allen-Cahn 方程时明显优于 PINNs、GA-PINNs、gPINNs、高阶 DG 等方法,实现了最低的均方误差,并捕捉到了间断点的跳跃,表明该方法在处理具有特征信息方程时是有效的.神经网络功能强大,可以用来逼近复杂的几何形状或高维度的偏微分方程,然而这些问题难以通过经典的数值计算方法进行模拟,未来可以将 AGA-PINNs 方法应用于处理高维间断方程^[29],特别需要考虑间断点的位置在多个方向上传播和相互作用.

在高维场景中,Sobol 采样作为低差异序列,较随机或均匀采样更能实现训练点的均匀覆盖,具有良好的数值稳定性.同时,残差排序过程的复杂度为 $O(n \log n)$,可借助并行计算进一步加速.为应对高维计算负担,我们可引入稀疏残差估计与局部点更新策略,避免全局冗余计算.此外,高维空间中梯度信息更为复杂,现有的权重函数可进一步优化,如采用多维梯度范数替代单一方向导数,以更有效识别高阶间断区域.在训练点的动态调整方面,可结合 KD-tree 或 Voronoi 分区将计算域划分为多个子区域^[30],并在高残差区域局部加密,通过区域内自适应迭代实现效率与精度的提升.

对于实际工程应用,考虑到计算资源受限、模型简化需求及复杂边界条件等因素,AGA-PINNs 可引入轻量级网络结构(如 ResNet、dropout、early stopping 等)与采样策略协同优化,有效避免冗余训练,提升训练效率.对于边界条件处理,未来亦可结合坐标映射方法与边界编码网络进一步增强其在复杂边界条件下的适应能力.

参考文献(References):

- [1] Zhang K, Zhao X G, Zhang L M, et al. Current status and prospect for the research and application of big data and intelligent optimization methods in oilfield development[J]. *Journal of China University of Petroleum (Edition of Natural Science)*, 2020, **44**(4): 28-38.
- [2] Wang Y, Cheung S W, Chung E T, et al. Deep multiscale model learning[J]. *Journal of Computational Physics*, 2020, **406**: 109071.
- [3] 邱天威, 魏光美, 宋禹欣, 等. 基于 PINN 方法的 KdV 类方程新孤子解的研究[J]. 应用数学和力学, 2025, **46**(1): 105-113. (Qiu Tianwei, Wei Guangmei, Song Yuxin, et al. Novel soliton solutions to KdV-type equations based on physics-informed neural networks[J]. *Applied Mathematics and Mechanics*, 2025, **46**(1): 105-113. (in Chinese))
- [4] Zha W, Li X, Li D, et al. Shale digital core image generation based on generative adversarial networks[J]. *Journal of Energy Resources Technology*, 2021, **143**(3): 033003.
- [5] 肖争光, 张春利, 陈伟球. 基于 PINNs 的压电半导体梁的非线性多场耦合力学分析[J]. 应用数学和力学, 2024, **45**(10): 1288-1299. (Xiao Zhengguang, Zhang Chunli, Chen Weiqiu. Analysis of nonlinear multi-field coupling mechanics of piezoelectric semiconductor beams via PINNs[J]. *Applied Mathematics and Mechanics*, 2024, **45**(10): 1288-1299. (in Chinese))
- [6] Pang G, Lu L, Karniadakis G E. fPINNs: fractional physics-informed neural networks[J]. *SIAM Journal on Scientific Computing*, 2019, **41**(4): A2603-A2626.
- [7] 闵建, 傅卓佳, 郭远. 课程-迁移学习物理信息神经网络用于曲面长时间对流扩散行为模拟[J]. 应用数学和力学, 2024, **45**(9): 1212-1223. (Min Jian, Fu Zhuojia, Guo Yuan. Curriculum-transfer-learning-based physics-informed neural networks for simulating long-term-evolution convection-diffusion behaviors on curved surfaces[J]. *Applied Mathematics and Mechanics*, 2024, **45**(9): 1212-1223. (in Chinese))
- [8] Li X, Chen H, Liu Z, et al. Identifying varying thermal diffusivity of inhomogeneous materials based on a hybrid physics-informed neural network[J]. *International Journal of Applied Mechanics*, 2022, **14**(7): 2250027.
- [9] Hornik K, Stinchcombe M, White H. Universal approximation of an unknown mapping and its derivatives using multilayer feedforward networks[J]. *Neural Networks*, 1990, **3**(5): 551-560.
- [10] Raissi M, Perdikaris P, Karniadakis G E. Physics-informed neural networks: a deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations[J]. *Journal of Computational Physics*, 2019, **378**: 686-707.
- [11] Lu L, Meng X, Mao Z, et al. DeepXDE: a deep learning library for solving differential equations[J]. *SIAM Review*, 2021, **63**(1): 208-228.
- [12] Yu J, Lu L, Meng X, et al. Gradient-enhanced physics-informed neural networks for forward and inverse PDE problems[J]. *Computer Methods in Applied Mechanics and Engineering*, 2022, **393**: 114823.
- [13] Guo H, Zhuang X, Fu X, et al. Physics-informed deep learning for three-dimensional transient heat transfer analysis of functionally graded materials[J]. *Computational Mechanics*, 2023, **72**(3): 513-524.
- [14] Raissi M, Yazdani A, Karniadakis G E. Hidden fluid mechanics: learning velocity and pressure fields from flow visualizations[J]. *Science*, 2020, **367**(6481): 1026-1030.
- [15] Bar-Kohany T, Jain A. Dissipation of boundary effects in multilayer heat conduction problems[J]. *International Journal of Heat and Mass Transfer*, 2024, **223**: 125207.
- [16] Ferrer-Sánchez A, Martín-Guerrero J D, de Austri-Bazan R R, et al. Gradient-annihilated PINNs for solving Riemann problems: application to relativistic hydrodynamics[J]. *Computer Methods in Applied Mechanics and Engineering*, 2024, **424**: 116906.

- [17] Mao Z, Jagtap A D, Karniadakis G E. Physics-informed neural networks for high-speed flows[J]. *Computer Methods in Applied Mechanics and Engineering*, 2020, **360**: 112789.
- [18] Xu H, Chang H, Zhang D. DL-PDE: deep-learning based data-driven discovery of partial differential equations from discrete and noisy data[J]. *Communications in Computational Physics*, 2025, **29**(3): 698-728.
- [19] Xu J, Wei H, Bao H. Physics-informed neural networks for studying heat transfer in porous media[J]. *International Journal of Heat and Mass Transfer*, 2023, **217**: 124671.
- [20] Lu L, Pestourie R, Yao W, et al. Physics-informed neural networks with hard constraints for inverse design [J]. *SIAM Journal on Scientific Computing*, 2021, **43**(6): B1105-B1132.
- [21] Mahmud M S, Huang J Z, Salloum S, et al. A survey of data partitioning and sampling methods to support big data analysis[J]. *Big Data Mining and Analytics*, 2020, **3**(2): 85-101.
- [22] Nabian M A, Gladstone R J, Meidani H. Efficient training of physics-informed neural networks via importance sampling[J]. *Computer-Aided Civil and Infrastructure Engineering*, 2021, **36**(8): 962-977.
- [23] Renardy M, Joslyn L R, Millar J A, et al. To Sobol or not to Sobol? The effects of sampling schemes in systems biology applications[J]. *Mathematical Biosciences*, 2021, **337**: 108593.
- [24] Wu C, Zhu M, Tan Q, et al. A comprehensive study of non-adaptive and residual-based adaptive sampling for physics-informed neural networks[J]. *Computer Methods in Applied Mechanics and Engineering*, 2023, **403**: 115671.
- [25] Mishra S, Molinaro R. Estimates on the generalization error of physics-informed neural networks for approximating PDEs[J]. *IMA Journal of Numerical Analysis*, 2023, **43**(1): 1-43.
- [26] de Lara F M, Ferrer E. Accelerating high order discontinuous Galerkin solvers using neural networks: 1D Burgers' equation[J]. *Computers & Fluids*, 2022, **235**: 105274.
- [27] Du Q, Yang J, Zhou Z. Time-fractional Allen-Cahn equations: analysis and numerical methods[J]. *Journal of Scientific Computing*, 2020, **85**(2): 42.
- [28] Toro E F. *Riemann Solvers and Numerical Methods for Fluid Dynamics: a Practical Introduction*[M]. Berlin: Springer, 2009.
- [29] Zeng S, Zhang Z, Zou Q. Adaptive deep neural networks methods for high-dimensional partial differential equations[J]. *Journal of Computational Physics*, 2022, **463**: 111232.
- [30] Zhang H, Xu Y, Liu Q, et al. Solving Fokker-Planck equations using deep KD-tree with a small amount of data [J]. *Nonlinear Dynamics*, 2022, **108**(4): 4029-4043.