

基于深度算子网络的偏微分方程求解方法

张茹¹, 娄久², 闫晓晨², 贺政萍¹

1. 哈尔滨商业大学 管理学院, 黑龙江 哈尔滨 150028

2. 哈尔滨工业大学 计算学部, 黑龙江 哈尔滨 150001

DOI: 10.61369/TACS.2026070007

摘 要 : 半导体制造流程中, 扩散工艺参数直接影响器件使用表现, 偏微分方程可完整描摹杂质扩散的动力学变化细节。有限差分法、有限元法求解高维、强非线性扩散方程时, 运算耗时久, 网格设置会带来诸多限制条件。研究人员搭建融合物理信息的深度算子网络 (PI-DeepONet) 用于偏微分方程求解, 依托分支—主干双网络架构, 学习源项函数与时空变量的非线性关联。扩散方程残差、初始条件、边界条件参与损失函数搭建, 使模型输出运行保持平稳。实验数据显示, 该模型测试集平均相对 L2 误差为 4.80×10^{-3} , 边界区域误差降至 2.95×10^{-3} 。复杂源项输入、长时间演化模拟场景中, 模型泛化表现突出, 运算速率理想。该方式弱化传统数值方法对网格的硬性限制, 为半导体扩散工艺建模与多物理场仿真研究提供了全新技术实现路径。

关 键 词 : 偏微分方程; 深度算子网络; 物理信息神经网络; 扩散工艺; 半导体制造

PDE Solving Method Based on Deep Operator Network

Zhang Ru¹, Lou Jiu², Yan Xiaochen², He Zhengping¹

1. School of Management, Harbin University of Commerce, Harbin, Heilongjiang 150028

2. Department of Computing, Harbin Institute of Technology, Harbin, Heilongjiang 150001

Abstract : In the semiconductor manufacturing process, diffusion process parameters directly affect device performance, and partial differential equations (PDEs) can fully describe the dynamic change details of impurity diffusion. When finite difference method (FDM) and finite element method (FEM) are used to solve high-dimensional and strongly nonlinear diffusion equations, they suffer from long computation time and numerous constraints imposed by mesh settings. Researchers have constructed a physics-informed deep operator network (PI-DeepONet) for PDE solving, which relies on a branch-trunk dual-network architecture to learn the nonlinear correlations between source term functions and spatiotemporal variables. The residual of the diffusion equation, initial conditions, and boundary conditions are incorporated into the loss function construction to ensure the stability of the model output. Experimental data show that the average relative L2 error of the model on the test set is 4.80×10^{-3} , and the error in the boundary region is reduced to 2.95×10^{-3} . In scenarios with complex source term inputs and long-time evolution simulations, the model exhibits excellent generalization performance and ideal computation speed. This method weakens the rigid constraints of traditional numerical methods on meshes, providing a new technical implementation path for semiconductor diffusion process modeling and multi-physics field simulation research.

Keywords : partial differential equation (PDE); deep operator network; physics-informed neural network (PINN); diffusion process; semiconductor manufacturing

引言

半导体制造的工艺尺度不断迈向更微观的层级, 器件生产过程中的扩散处理步骤, 会直接改变材料内部的结构与传输特性。想要摸清杂质在材料内部的变化过程, 就要用对应的数学方式展开完整推演, 这类推演工作也成为工艺优化的重要参考依据。

以往的方程计算方式, 需要借助网格划分来推进运算步骤。场景变得复杂时, 网格设置会给运算带来较多束缚, 运算过程繁琐, 得到的结果也和实际情况存在偏差。当下结合深度学习的两类计算模型, 在实际使用中都有各自的短板, 要么只能应对固定场景, 要么和实际物理规律的贴合度不足^[1]。针对这类应用情况, 研究人员将两类模型的长处结合, 构建出 PI-DeepONet 计算模型^[2]。经过多组测试验证, 该模型在扩散过程的运算中, 输出状态更平稳, 结果也更贴近实际工况, 为半导体领域的工艺仿真工作, 提供全新的实践方向^[3]。

项目信息: 本文系 2024 年度黑龙江省省属本科高校基本科研费项目资助 (项目名称: 基于 AIGC 的 "黑土优品+" 品牌价值提升策略研究)。

一、相关理论基础

(一) 半导体扩散工艺的偏微分方程模型

半导体杂质扩散遵循粒子守恒原理。扩散方程基于粒子守恒定律、费克定律与电场迁移项耦合推导，它能描述温度与电场共同作用下，杂质或缺陷浓度随时空的演化过程^[4]。

扩散方程通用形式如下：

$$\frac{\partial C_{ch}}{\partial t} = \nabla \cdot \left(D_{AX} \nabla C_A + \frac{Z_A C_A q E}{kT} \right) + S$$

式中， C_{ch} 为化学浓度（总杂质浓度）， D_{AX} 为扩散系数，

$\frac{Z_A C_A q E}{kT}$ 为漂移项， S 为源项。

研究人员会根据精度需求与计算效率，对扩散模型做简化，常用三类模型：

1. Fermi 模型：假设点缺陷处于热力学平衡，无需显式求解缺陷浓度，仅通过掺杂剂方程完成计算^[5]。

2. 二维模型：同时求解掺杂剂浓度与点缺陷浓度，适合中等复杂度场景。

3. 全耦合模型：同步求解掺杂扩散、载流子输运与电势分布，完整描述实际物理过程。

扩散问题的初始条件由离子注入剖面决定，边界条件常用 Dirichlet 定值边界、Neumann 通量边界、Robin 混合边界三类，分别适配恒定浓度、封闭边界、物质交换边界等不同工艺场景^[6]。

(二) 深度算子网络 (DeepONet)

DeepONet 是一种用于学习算子映射的深度学习架构，由分支网络与主干网络组成^[7]。分支网络提取输入函数的特征。主干网络提取查询点的位置特征。两个网络的输出做内积运算，得到目标函数在指定位置的预测值，其优势为一次训练、多场景预测，降低重复数值求解成本。

DeepONet 的数学表达式为：

$$G_\theta(u)(y) = \sum_{k=1}^p b_k(u) t_k(y)$$

其中， $b_k(u)$ 是分支网络的第 k 个特征分量， $t_k(y)$ 表示主干网络的第 k 个特征分量， p 为特征维度， θ 表示网络的全部可训练参数。

(三) 物理信息神经网络 (PINN)

物理信息神经网络 (PINN) 将控制方程、初边值条件转化为损失函数约束，在无网格框架下学习物理规律，降低标注数据依赖^[8]。但 PINN 仅针对单一问题训练，参数变化需重新训练，在长时间演化、强非线性问题中易出现收敛慢、误差累积等问题，仅适用于单场景求解^[9]。

二、PIDeepONet 模型构建

(一) 模型整体架构

PI-DeepONet 保留 DeepONet 的分支-主干双网络结构，搭配物理约束模块，构建完整的网络体系。分支网络提取输入函数高维特征，主干网络对时空坐标信息进行编码处理。残差模块通

过自动微分，完成扩散方程残差的计算，物理规律会伴随整个训练过程。模型学习算子映射关系的过程中，依照物理约束推进运算，运算走向匹配真实物理演化^[10]。

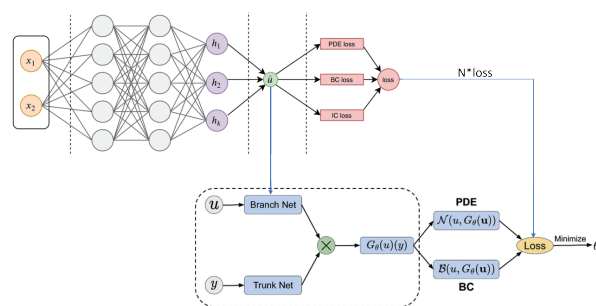


图1 PI-DeepONet 网络架构图

(二) 复合损失函数设计

PI-DeepONet 采用复合损失函数，它由边界条件损失与 PDE 残差损失两部分组成，数学表达式为：

$$L = \lambda_{bc} L_{bc} + \lambda_{res} L_{res}$$

式中， λ_{bc} 和 λ_{res} 为权重系数，初始值均设为 1。

边界条件损失 L_{bc} 衡量边界预测值与真实值的差异：

$$L_{bc} = \frac{1}{P} \sum_{i=1}^P (\hat{C}(x_i, t_i) - C(x_i, t_i))^2$$

式中， P 为边界采样点数量， $\hat{C}(x_i, t_i)$ 为模型预测值， C 为真实值。

PDE 残差损失 L_{res} 衡量域内采样点的方程残差：

$$L_{res} = \frac{1}{Q} \sum_{j=1}^Q r(x_j, t_j)^2$$

式中， Q 为域内残差采样点数量， r 为扩散方程残差。

复合损失函数让模型同时满足数据监督与物理约束，提升预测精度与物理一致性。

(三) 训练策略与采样方法

研究人员设计专用采样与训练策略，提升模型训练效率与稳定性：

1. 采样策略：边界区域加密采样，强化边界约束学习；时空域内随机选取配置点，计算 PDE 残差；训练样本加入高频振荡源项，提升复杂场景泛化能力。

2. 优化器与学习率：采用 Adam 优化器，初始学习率设为 10^{-3} ，每 2000 次迭代按指数衰减 10%。

3. 早停机制：验证损失连续 500 次迭代不下降时，提前终止训练，防止过拟合。

三、实验设计与实现

(一) 实验设置

实验以一维扩散方程为验证对象，扩散系数取常数 $D=0.01$ ，时空域为 $[0,1] \times [0,1]$ 。采用有限差分法生成数值解作为监督数据，通过高斯过程径向基函数生成 1000 组源项样本构建训练集，测试集采用密集时空网格保证评估精度。初始条件设为 $C(x,0)=0$ ，边界条件取齐次 Dirichlet 条件，基于 JAX 框架完成训练，利用

自动微分与 GPU 加速计算。

实验以相对 L_2 误差作为核心评价指标，公式为：

$$e_{rel} = \frac{\|\hat{u} - u\|_2}{\|u\|_2}$$

式中， $\hat{u}$ 为模型预测值， u 为参考数值解。

（二）实验结果分析

PIDeepONet 测试集平均相对 L_2 误差为 4.80×10^{-3} ，标准差为 2.01×10^{-3} 。和原始 DeepONet 相比，PIDeepONet 的整体误差水平显著降低，物理约束有效缓解了模型对样本的过度依赖。

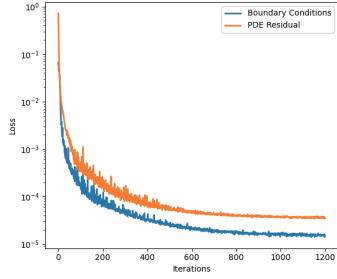


图2边界损失与PDE损失迭代图

训练过程中，边界损失与残差损失平稳下降，后期分别稳定在 10^{-5} 与 10^{-4} 量级，模型收敛状态良好。

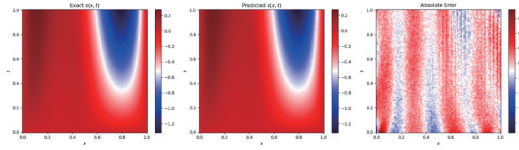


图3真实解与预测解对比图

模型预测解与数值解时空分布高度契合，无振荡、失真等问题，高梯度区域仍保持高精度拟合。

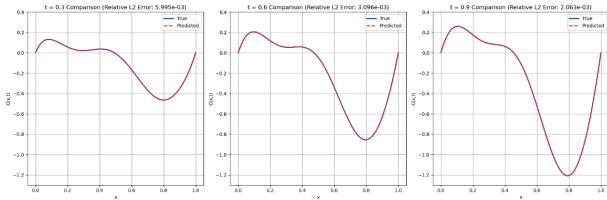


图4不同时刻真实值与预测值对比曲线

长时间演化测试中，PI-DeepONet 误差始终保持平稳， $t=0.9$ 时刻误差仅 2.063×10^{-3} ，远低于原始 DeepONet，长期稳定性优势显著。

与原始 DeepONet 相比，PI-DeepONet 整体精度提升约6倍，边界精度提升约12倍，可有效解决原始模型物理一致性不足的缺陷。

四、多方法对比实验

（一）实验设置

研究人员选取 DeepONet、PIDeepONet、DeepXDE、PINN 四种方法做对比。实验设置5组固定测试用例，空间网格数 $N_x=100$ ，时间网格数 $N_t=100$ ，扩散系数 $D=0.01$ 。所有方法

在统一环境下训练与测试，评价指标为平均相对 L_2 误差、训练时间、推理时间。

（二）性能对比结果

表1为四种方法的性能对比数据。

表1不同方法性能对比

方法	平均相对 L_2 误差	训练时间 (s)	适用范围
DeepONet	2.757e-02	26.3	同类多种条件
PI-DeepONet	1.651e-02	69.0	同类多种条件
DeepXDE	3.498e-02	549.7	同类多种条件
PINN	3.068e-01	287.2	单种条件

PIDeepONet 的平均相对 L_2 误差最低，预测精度最优。

PINN 误差最大，泛化能力受限。DeepONet 训练时间最短，PIDeepONet 以适中的训练时间换取精度大幅提升。DeepXDE 训练耗时最长，计算成本高。

（三）误差分布对比

对数坐标下，PINN 误差整体偏高，DeepXDE 与 DeepONet 在部分用例出现误差峰值。PIDeepONet 在所有测试用例中均保持低误差，曲线平稳。误差主要集中在梯度变化大的区域，PIDeepONet 在该区域的局部误差控制最优，最大误差仅0.014。结果证明，PIDeepONet 在精度、稳定性、鲁棒性上综合表现最佳。

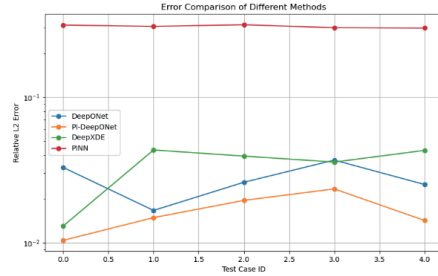


图5 5个测试用例 error 迭代图

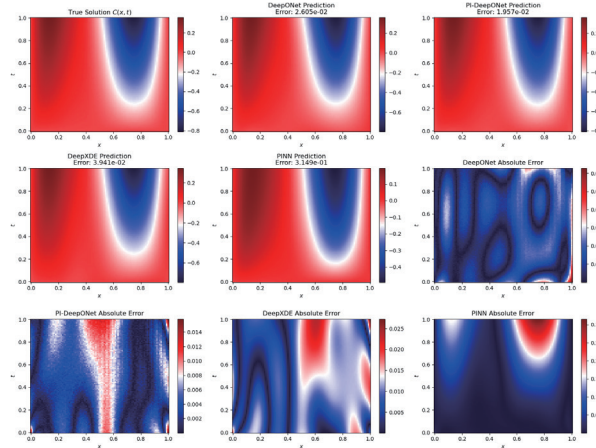


图6真实值与4个模型预测值对比图

五、结论与展望

（一）结论

本文提出的 PI-DeepONet 方法，将深度算子网络与物理规则类网络的长处加以整合，通过分支一主干双网络与复合损失函数

的设计,让模型训练全程顺着物理演化逻辑推进。实验测试所得数据能够表明,该模型输出结果的贴合度更高,边界区域的还原效果更理想,长时段模拟过程也能维持平稳状态,网格带来的运算束缚被大幅弱化,可用于半导体扩散相关模拟分析以及多物理场相关的计算作业。

(二) 展望

后续研究可从三个方向继续延伸。第一,拓宽模型的应用范

围,适配二维、三维空间内的扩散模拟场景,应对更多物理场相互作用的计算场景。第二,优化模型的采样方式与网络结构设计,减少训练过程的资源损耗,让模型在实际工程场景中更易落地运用。第三,把模型融入商用仿真工具之中,替换原有常规计算模块,让半导体制造领域的仿真技术得到进一步完善。

参考文献

[1] 孔宪熠,贾鹤萍.基于双通道物理信息神经网络的耦合方程逆问题求解[J/OL].山西大学学报(自然科学版),1-10.https://doi.org/10.13451/j.sxu.ns.2025063.

[2] 谢翔,江林峰,杨凤莲.基于自适应权重两阶段 PINNs 方法求解偏微分方程[J].广西师范大学学报(自然科学版),2026,44(02):175-189.

[3] 肖勇,周夏峰.基于物理信息神经网络方法求解中子扩散方程本征值问题[J].核动力工程,2025,46(S1):288-295.DOI:10.13832/j.jnpe.2025.S1.0288.

[4] 孙宇,朱海龙.局部自适应物理信息神经网络在求解 Burgers-Fisher 方程中的应用[J].山东理工大学学报(自然科学版),2025,39(04):47-53.

[5] 刘天源,张海湘,杨雪花,等.一类二维变系数反应扩散方程的交替方向隐式差分方法[J].湖南工业大学学报,2025,39(04):96-104.

[6] 王飞,党浩宁,尚勇,等.从传统数值方法到神经网络:偏微分方程数值解的演进与展望[J].现代应用物理,2025,16(01):41-66.

[7] 郭嘉,刘子源,侯臣平.求解含时偏微分方程的物理信息神经网络研究进展[J].现代应用物理,2025,16(01):81-93.

[8] 史雪婷,韩峰,王建国.偏微分方程智能求解算例分析及改进[J].现代应用物理,2025,16(01):256-267+290.

[9] 李芳,李昕,夏侯珍,等.带有 Dirichlet 边界条件的退化反应扩散方程解的收敛性(英文)[J].上海师范大学学报(自然科学版中英文),2024,53(06):711-718.DOI:10.20192/j.cnki.JSHNU(NS).2024.06.001.

[10] 陈煜谦.基于深度学习的偏微分方程求解[D].浙江大学,2024.