瞬态对流占优问题中稳定化有限元解的物理信息后处理

Süleyman Cengizci[a,*], Ömür Uğur[b], Srinivasan Natesan[c]

[a]安塔利亚科学大学计算机技术系，安塔利亚 *07190*，土耳其[b]中东技术大学应用数学研究所，安卡拉 *06800*，土耳其[c]印度理工学院古瓦哈蒂分校数学系，古瓦哈蒂 *781039*，印度

arXiv:2603.03259v1 [math.NA] 3 Mar 2026

摘要

由于时空区域中存在陡峭梯度和传播前沿，对流占优瞬态输运现象的数值模拟面临着重大的计算挑战。经典数值离散化方法往往会产生伪振荡，因此需要采用先进的稳定化策略。此外，即使稳定化抑制了此类数值不稳定性，仍可能需要额外的正则化来解析局部陡峭梯度。另一方面，单独使用物理信息神经网络（PINN）方法难以捕捉对流主导问题中出现的尖锐层，且通常需要大量训练轮次才能从头学习解。为此，本文提出一种混合计算框架，将最初针对稳态问题开发的PASSC（物理信息神经网络增强的SUPG与激波捕捉）方法扩展至非稳态情形，并结合稳定化有限元方法（FEM）与PINN求解瞬态对流-扩散-反应（CDR）方程。本文采用半离散稳定化有限元方法，通过流线上风/Petrov–Galerkin（SUPG）形式并辅以YZβ 激波捕捉算子实现稳定化，以解析陡峭梯度和激波类特征。PINN后处理校正并非在整个时空区域内进行，而是有选择地应用于终止时刻附近；此时，神经网络通过吸收来自最后 K_s 个快照的空间数据来增强有限元解，同时施加由控制方程和边界条件导出的残差约束。网络架构采用带有随机傅里叶特征的残差块，以高效捕捉多尺度解结构；渐进训练策略结合自适应损失加权，平衡数据保真度与物理一致性。针对五个基准问题开展的全面数值实验——涵盖边界层和内部层、行波以及非线性Burgers动力学——表明，与单独使用稳定化有限元解相比，所提出的框架显著提高了终止时刻的解精度。

关键词：物理信息神经网络、有限元、对流主导、SUPG、激波捕捉、深度学习 g

*通讯作者电子邮箱： suleyman.cengizci@antalya.edu.tr (Süleyman Cengizci), ougur@metu.edu.tr (Ömür Uğur), natesan@iitg.ac.in (Srinivasan Natesan)

1. 引言

科学与工程中对物理现象的数学描述，在很大程度上依赖于（非线性）偏微分方程（偏微分方程）。对流-扩散-反应方程是尤为重要的一类偏微分方程，因为它们描述了质量、热量和动量等量的输运与转化。由于这类方程的解析解——尤其是在涉及非线性或耦合系统时——很少能写成闭式表达式，因此计算技术变得不可或缺。为此，人们通常采用数值离散格式、半解析技术、机器学习方法，或将这些方法结合起来，以可接受的精度和计算成本近似求解。

对流–扩散–反应型偏微分方程的高保真数值模拟要求对空间离散和时间离散均进行细致处理。经典的数值空间离散方法，包括有限差分、有限元和有限体积格式，在对流主导扩散时往往会产生（节点到节点的）伪振荡，从而降低计算结果的物理可靠性。这一长期存在的困难推动了过去数十年中众多稳定化公式的发展。在有限元领域，流线上风/Petrov-Galerkin（SUPG） [1–5], Galerkin/最小二乘法（GLS） [6], 以及变分多尺度方法（VMS） [7]作为有效的稳定化策略已获得广泛认可。尽管取得了成功，这些技术在陡峭梯度、激波或薄边界层和内部层附近仍可能产生局部非物理振荡。因此，激波捕捉（SC）（亦称间断捕捉）算子常被附加进来，以进一步抑制解快速变化区域中的此类振荡。然而，一个持续存在的挑战在于稳定化参数的恰当确定，这些参数通常逐单元计算。尽管存在众多基于经验的参数定义，普遍适用的形式仍难以获得，且参数调节可能带来显著的计算开销，尤其是对于瞬态和高维问题。关于稳定化有限元技术及其应用的全面讨论可见于 [8–23] 及其所引文献。

深度学习技术的快速发展，激发了人们将其应用于偏微分方程求解器的日益浓厚的兴趣。物理信息神经网络（PINNs）是一项尤为前景广阔的进展：通过微分算子编码的物理定律被直接嵌入训练损失。通过在选定配点处最小化控制方程的残差，物理信息神经网络无需依赖标注数据集即可逼近解。尽管如此，物理信息神经网络与既有数值方法——包括标准Galerkin有限元法（GFEM）——在很大程度上沿各自独立的路径发展。然而，近期研究表明，Machine Learning工具，尤其是物理信息神经网络，在增强经典求解器方面具有相当大的潜力，可为数值稳定性所需的人工耗散提供数据驱动估计。尽管已探索了物理信息神经网络与有限元公式之间的多种耦合策略，但将激波捕捉机制系统性地纳入基于连续有限元法的物理信息神经网络框架，仍受到极为有限的关注。填补这一空白有望在精度与效率两方面带来改进，并将稳定化方法的适用范围扩展到要求更高的应用。

许多研究致力于利用人工神经网络（人工神经网络）提升传统数值格式的稳定性和保真度。例如，Ray 和 Hesthaven [24, 25], 利用

多层感知器检测高阶有限元计算中的振荡单元，并在局部应用斜率限制器以减轻非物理行为。Veiga 等人 [26] 为双曲守恒律开发了一种无参数稳定化策略，通过训练人工神经网络定位需要人工耗散的区域；他们在标量对流和欧拉系统问题上，基于间断伽辽金有限元法（间断伽辽金有限元法）验证了该方法。Discacciati 等人 [27] 将人工神经网络与龙格–库塔间断伽辽金有限元求解器相结合，构建了一个通用人工黏性框架，既能为光滑解保持最优精度，又能可靠地捕捉不连续性。Schwander 等人 [28] 训练网络，在傅里叶谱背景下评估局部正则性，并将其方法与熵黏性方法进行基准比较，测试对象包括一维和二维守恒律以及欧拉方程。

Raissi 等人 [29] 关于物理信息神经网络的奠基性研究表明，无需带标签训练数据，即可通过以无监督或半监督的方式最小化配点处的残差来求解微分方程。Karniadakis 等人 [30] 随后将这一范式扩展至高维反问题、涉及含噪观测的情形以及潜在物理规律的发现，凸显了物理信息神经网络在正向建模和反向建模中的多功能性。Mitusch 等人 [31] 介绍了一种基于开源库 FEniCS [32, 33] 和 dolfin-adjoint [34],构建的混合有限元–物理信息神经网络实现，并将其应用于泊松问题、对流–扩散–反应问题以及心脏电生理学。Huang 等人 [35] 对用于对流–扩散–反应型方程的深度神经网络方法进行了分类，并评估了这些方法在科学与工程领域中的适用性。Jin 等人 [36]提出了 NSFnets，这是一种专为不可压缩Navier–Stokes方程设计的物理信息神经网络架构，适用于层流和湍流状态。

Yu 和 Hesthaven [37] 为间断伽辽金有限元法构建了一种数据驱动人工黏性模型，在光滑度检测方面显著优于经典模态衰减指标。Yadav 和 Ganesan [38] 提出了 SPDE-ConvNet，这是一种结合 SUPG 稳定化的卷积神经网络，用于奇异摄动对流–扩散–反应方程，并展示了其在参数优化和振荡抑制方面的有效性。Huang [39] 提出了一种面向对流主导问题的变分一致的节点积分无网格方法，以解决节点积分固有的伽辽金精确性损失问题。Le-Duc 等人 [40] 开发了用于参数化结构力学模拟的有限元信息神经网络（FEI-NN）。在后续研究中，Yadav 和 Ganesan [41] 提出了 SPDE-NetII，这是一种人工神经网络增强的流线迎风佩特罗夫–伽辽金格式，在边界层和内部层的分辨率方面有所提升，并在基准比较中优于标准物理信息神经网络和变分神经网络（VarNet） [42] 。Luong 等人 [43]通过一个两阶段过程解决了复杂几何形状上的精确边界条件施加问题：首先构造满足齐次本质边界条件的基于深度神经网络的试函数，然后求解由此得到的边值问题。

Grossmann等人 [44] 针对线性偏微分方程和非线性偏微分方程（包括泊松方程、Allen–Cahn方程和半线性薛定谔方程），对物理信息神经网络与有限元法进行了严格比较，结论是有限元法通常能获得更高的精度，而物理信息神经网络在某些情形下可能具有计算优势。Yamazaki等人 [45] 将弱伽辽金格式与隐式欧拉时间步进相结合，提出了有限算子学习（有限算子学习），在非均质介质中的含时热传导问题上取得了高精度。加速训练的努力包括由Buzaev等所探索的傅里叶神经算子和库普曼神经网络

人 [46], 以及Wang等人 [47], 的课程学习策略，该策略优先处理较容易的子域，然后再处理高损失区域。Chang等人 [48, 49] 提出了奇异层物理信息神经网络（sl-PINNs），将边界层结构直接嵌入网络架构，从而改善奇异摄动问题的求解性能。Frerichs-Mihov等人 [50] 为对流主导流设计了专门的损失函数，而Hsieh和Huang [51] 提出了一种多尺度稳定化物理信息神经网络公式，结合弱边界约束、索博列夫训练和迁移学习，以增强边界层附近的稳定性。Kütük和Yücel [52] 将Allen–Cahn方程的能量耗散结构嵌入物理信息神经网络损失函数，并证明其能够在一维到三维范围内准确模拟相分离和亚稳态。最近，Raina等人 [53] 研究了流体力学中的双参数奇异摄动问题，指出了标准物理信息神经网络在分辨边界层方面的局限性，并提出了受有限元方法启发的有限基物理信息神经网络（FB-PINNs）。

在本研究中，我们将当前研究的作者在 [54] 提出的混合框架——称为 PASSC（**P**INN-**A**ugmented **S**UPG with **S**hock-**C**apturing，PINN增强型SUPG与激波捕捉）——从最初针对平流主导的稳态对流-扩散-反应型偏微分方程的框架，扩展至非稳态情形。本研究的主要贡献概括如下。首先，我们提出一种混合训练范式，利用两种精密的有限元法（FEM）要素——SUPG稳定化和 YZβ 激波捕捉机制 [55, 56]——为具有边界层和/或内部层的对流主导问题生成高保真训练数据。尽管 SUPG稳定化在许多此类情形下表现良好，但要充分解析陡峭梯度和内部间断，通常仍需辅助激波捕捉 [57, 58]。因此，我们采用 SUPG-YZβ 解作为训练数据，同时仅在稳定的内部区域有选择地施加偏微分方程残差约束。其次，所提出的架构与维数无关，采用了具有傅里叶特征嵌入的深度残差网络、层归一化以及审慎选择的超参数。第三，该方法具有良好的可扩展性，可处理一维、二维和三维空间问题。

2. 数学框架

令 $n_{\mathrm{sd}} \geq 1$ 表示空间维数，令 $\Omega \subset \mathbb{R}^{n_{\mathrm{sd}}}$ 表示一个有界利普希茨区域，该区域代表流动所占据的物理区域，即问题定义所在的空间域。为建立本文讨论的数学框架，我们考虑以下时变对流-扩散-反应型偏微分方程作为一个代表性模型问题：

$$\frac{\partial u(t,\mathbf{x})}{\partial t} - \varepsilon \Delta u(t,\mathbf{x}) + \mathbf{b}(t,\mathbf{x}) \cdot \nabla u(t,\mathbf{x}) + c(\mathbf{x}) u(t,\mathbf{x}) = f(t,\mathbf{x}) \quad \text{in } \Omega \times (t_0, t_{\mathrm{f}}], \tag{1}$$

$$u(t,\mathbf{x}) = u_D(t,\mathbf{x}) \text{ on } \partial\Omega \times (t_0, t_{\mathrm{f}}], \tag{2}$$

$$u(t_0,\mathbf{x}) = u_0(\mathbf{x}) \quad \text{in } \Omega. \tag{3}$$

这里，u 表示未知标量场，其中 $t \in (t_0, t_{\mathrm{f}}]$ 为时间变量，$\mathbf{x} = (x_1, x_2, \ldots, x_{n_{\mathrm{sd}}}) \in \Omega$为空间位置向量。正常数 ε 表示扩散系数，并假定其在对流主导区域中满足 $0 < \varepsilon \ll 1$。向量场$\mathbf{b} \in L^\infty([t_0, t_{\mathrm{f}}]; \mathcal{W}^{1,\infty}(\Omega)^{n_{\mathrm{sd}}})$表示对流速度，$c \in L^\infty(\Omega)$为反应系数，$f \in L^2([t_0, t_{\mathrm{f}}]; L^2(\Omega))$为源项。

函数 $u_0 \in L^2(\Omega)$规定初始条件，而 $u_D \in C([t_0, t_\mathrm{f}]; \mathcal{H}^{1/2}(\partial\Omega))$指定与时间相关的狄利克雷边界数据，其中 $\mathcal{H}^{1/2}(\partial\Omega)$表示 $\mathcal{H}^1(\Omega)$在 $\partial\Omega$上的迹空间，使得对每个 $t \in [t_0, t_\mathrm{f}]$，函数 $u_D(t, \cdot)$均可延拓至 $\mathcal{H}^1(\Omega)$。在本文中，我们采用标准索伯列夫空间记号： $L^2(\Omega)$与 $L^\infty(\Omega)$分别表示平方可积函数空间与本质有界函数空间； $\mathcal{W}^{1,\infty}(\Omega)$是一阶弱导数本质有界的函数空间； $\mathcal{H}^1(\Omega)$表示一阶导数平方可积的函数空间。对于巴拿赫空间 X，记号 $C([t_0, t_\mathrm{f}]; X)$与 $L^2([t_0, t_\mathrm{f}]; X)$分别指连续的与平方可积的$X$值函数的标准空间。

注1。 尽管控制方程(1)–(3)配备了狄利克雷边界条件，所提出的框架同样能够方便地容纳诺伊曼型条件。为使阐述简洁清晰，在后续推导中我们仅关注狄利克雷情形。

注2。 为便于第2.1–2.3节有限元公式中的记号，当场量的含义可由上下文清楚看出时，省略其显式的空间依赖性。因此，对流速度$\mathbf{b}(t,\ \mathbf{x})$、反应系数 $c(\mathbf{x})$和源项 $f(t,\ \mathbf{x})$此后分别记为$\mathbf{b}$、 c和f。

2.1. 标准*Galerkin*有限元法形式

与式(1)所定义的模型问题相对应的标准Galerkin有限元形式寻求解 $u^h \in \mathcal{S}_u^h$， 使得对所有测试函数 $w^h \in \mathcal{V}_u^h$: –(3)

$$\int_\Omega w^h \frac{\partial u^h}{\partial t}\, d\Omega + \int_\Omega \left(\varepsilon \nabla w^h \cdot \nabla u^h + w^h\, \mathbf{b} \cdot \nabla u^h + c\, w^h u^h \right) d\Omega = \int_\Omega w^h f\, d\Omega. \tag{4}$$

在该形式中，上标 h 表示属于有限维函数空间。离散试探函数空间与测试函数空间定义如下:

$$\mathcal{S}_u^h = \left\{ u^h \in \mathcal{H}^{1h}(\Omega) : u^h \doteq u_D^h \text{ on } \partial\Omega \right\}, \tag{5}$$

$$\mathcal{V}_u^h = \left\{ w^h \in \mathcal{H}^{1h}(\Omega) : w^h \doteq 0 \text{ on } \partial\Omega \right\}, \tag{6}$$

其中， u_D^h 表示边界数据 u_D 在有限元空间上的离散表示，符号 $\doteq$表示边界值是按照有限元插值的意义赋予的。空间 $\mathcal{H}^{1h}(\Omega)$ 表示索伯列夫空间 $\mathcal{H}^1(\Omega)$ 的一个有限维子空间，其定义如下:

$$\mathcal{H}^{1h}(\Omega) = \left\{ \phi^h \in \mathcal{C}^0(\bar{\Omega}) : \phi^h\big|_{\Omega^e} \in \mathbb{P}_1(\Omega^e),\ \forall \Omega^e \in \mathcal{T}^h \right\}. \tag{7}$$

这里， $\mathcal{C}^0(\bar{\Omega})$ 表示定义在计算域 Ω 的闭包上的连续函数空间。集合 $\mathbb{P}_1(\Omega^e)$ 表示定义在每个有限元 Ω^e 上的分片线性多项式函数空间，而$\mathcal{T}^h$ 表示将区域剖分为有限元的网格剖分。

这种标准 Galerkin 公式为对流–扩散–反应型偏微分方程的数值处理提供了基准。然而，在对流主导区域中，众所周知，这种方法通常会遇到数值不稳定性和非物理振荡，因此需要采用后续章节将讨论的稳定化技术，例如 SUPG 和激波捕捉方法。

*2.2. SUPG*稳定化有限元公式

与方程（1）–（3）中的模型问题相对应的 SUPG稳定化有限元公式旨在求得离散解 $u^h \in \mathcal{S}_u^h$，使其满足以下条件：对于所有测试函数 $w^h \in \mathcal{V}_u^h$:

$$\int_\Omega w^h \frac{\partial u^h}{\partial t}\, d\Omega + \int_\Omega \left(\varepsilon \nabla w^h \cdot \nabla u^h + w^h\, \mathbf{b} \cdot \nabla u^h + c\, w^h u^h\right) d\Omega - \int_\Omega w^h f\, d\Omega + \sum_{e=1}^{n_{\mathrm{el}}} \int_{\Omega^e} \tau_{\mathrm{SUPG}} \left(\mathbf{b} \cdot \nabla w^h\right) \underbrace{\left(\frac{\partial u^h}{\partial t} - \varepsilon \Delta u^h + \mathbf{b} \cdot \nabla u^h + c u^h - f\right)}_{\mathcal{R}(u^h)} d\Omega^e = 0. \quad (8)$$

公式 (8) 中的第一个积分对应标准 GFEM 公式，第二行则表示 SUPG贡献项。解空间和试验函数空间 $\mathcal{S}_u^h$ 和 $\mathcal{V}_u^h$ 与 GFEM 公式中定义的相同。索引 n_{el} 表示网格离散中的有限元数量，而 τ_{SUPG} 是与 SUPG 公式相关的稳定化参数。

备注 3. 在本研究中，所有有限元计算均采用连续分片线性空间 $\mathcal{H}^{1h}(\Omega)$。尽管高阶多项式基（例如 $\mathbb{P}_2$ 或 $\mathbb{P}_3$）可能在光滑区域获得更高的精度，但其计算成本较高，并且在对流占主导地位时容易出现数值不稳定性，尤其是在存在尖锐内层或陡峭梯度的情况下。此外，在缺乏充分稳定化时，较高的多项式次数往往会加剧此类情形下的伪振荡。相比之下，线性单元的组装和求解时间较短，且能够与 SUPG稳定化和激波捕捉技术无缝结合。因此，我们有意采用稳定化的 $\mathbb{P}_1$ 单元，以在计算效率和稳健性之间实现有效折中，从而支持对广泛的对流主导问题进行准确模拟。

在形式(8)中，参数 τ_{SUPG} 决定了为稳定化目的而添加到离散系统中的人工扩散水平。该参数的选取需要精细平衡：过大的取值会产生过度扩散的解，从而损害精度；而过小的取值则会使振荡伪影持续存在。因此，必须根据具体问题对 τ_{SUPG} 进行细致调节。尽管文献中提供了多种稳健的定义，但尚未出现单一的普遍最优公式。在本研究中，我们倾向于采用Shakib等人提出的稳定化参数 [59], ，更多讨论见 [9, 60]:

$$\tau_{\mathrm{SUPG}} = \left(\left(\frac{2}{\Delta t}\right)^2 + \left(\frac{2\|\mathbf{b}\|}{h_K}\right)^2 + \left(\frac{4\varepsilon}{h_K^2}\right)^2\right)^{-1/2}. \quad (9)$$

此处， $\|\mathbf{b}\|$ 表示对流速度的大小（其中 $\|\cdot\|$ 表示本文中始终一致使用的欧几里得范数）， Δt 表示时间步长， h_K 是与单元 $K = \Omega^e$相关联的特征网格尺寸，定义为

$$h_K = \mathrm{diam}(K), \quad \forall K \in \mathcal{T}^h. \quad (10)$$

注 4. 式 (8) 中 SUPG 公式的一个关键性质是，稳定化项将控制偏微分方程 (1) 的残差 $\mathcal{R}(u^h)$ 作为加权因子纳入其中。这一构造保证了相容性：每当将精确解代入该形式时，稳定化贡献恒为零。

大量数值证据 [14, 16, 17, 23, 57, 58, 61–65] 表明，SUPG 方法虽然在许多情形下有效，但在抑制陡峭梯度或类激波特征附近的节点间振荡方面可能不足。为弥补这一不足，通常需要补充的激波捕捉机制。在本工作中，我们采用 YZβ 激波捕捉技术，该技术最初由 Tezduyar 及其合作者 [55, 56, 66–68], 提出，用以提高呈现陡峭空间变化区域的解的质量。

2.3. SUPG-YZβ形式

加入 SUPG 稳定化和 YZβ 激波捕捉后，GFEM 有限元公式为：求 $u^h \in \mathcal{S}_u^h$，使得对所有 $w^h \in \mathcal{V}_u^h$,

$$\begin{aligned}\int_\Omega w^h \frac{\partial u^h}{\partial t}\, d\Omega + \int_\Omega \Big(\varepsilon \nabla w^h \cdot \nabla u^h + \mathbf{b}\cdot\nabla u^h\, w^h + c\, u^h w^h - f w^h\Big)\, d\Omega \\ + \sum_{e=1}^{n_{\rm el}} \int_{\Omega^e} \tau_{\rm SUPG}(\mathbf{b}\cdot\nabla w^h)\Big(\frac{\partial u^h}{\partial t} - \varepsilon\Delta u^h + \mathbf{b}\cdot\nabla u^h + c\, u^h - f\Big)\, d\Omega \\ + \sum_{e=1}^{n_{\rm el}} \int_{\Omega^e} \nu_{\rm SHOC}(\nabla w^h \cdot \nabla u^h)\, d\Omega = 0. \quad (11)\end{aligned}$$

式 (11) 的第一行表示标准 Galerkin 公式，第二行引入 SUPG 稳定化项，第三行则包含激波捕捉项。函数空间 $\mathcal{S}_u^h$ 和 $\mathcal{V}_u^h$ 的定义同前。YZβ 激波捕捉参数 $\nu_{\rm SHOC}$可按如下方式定义 [55, 56, 66–68]:

$$\nu_{\rm SHOC} = \left|\mathrm{Y}^{-1}\mathrm{Z}\right| \left(\sum_{i=1}^{n_{\rm sd}} \left|\mathrm{Y}^{-1}\frac{\partial u^h}{\partial x_i}\right|^2\right)^{\frac{\beta}{2}-1} \left(\frac{h_{\rm SHOC}}{2}\right)^{\beta}, \quad (12)$$

其中，Y 是一个标量，表示解 u的参考值。

注释 5。 参数 Y 是一个非零缩放常数，用于体现解的参考尺度。一种典型的定义为

$$\mathrm{Y} = \frac{u^{\max} + u^{\min}}{2}, \quad (13)$$

表示预期解范围的均值；更一般地，也可以在 $u^{\min}$ 和$u^{\max}$ 之间选取任意合适的值。对于瞬态问题，初始条件通常可为确定 Y 提供合适的依据。在耦合方程组的情形下，Y 替换为对角矩阵$\mathbf{Y}$ ，其各元素编码了每个控制方程的参考尺度；进一步讨论见 [61, 68] 。

式（12）中的量 Z 表示时变控制方程的逐点残差，表示为 [69]

$$\mathrm{Z} = \mathcal{R}(u^h) = \frac{\partial u^h}{\partial t} - \varepsilon\Delta u^h + \mathbf{b}\cdot\nabla u^h + c u^h - f. \quad (14)$$

备注 6。 该残差定义扩展了 Tezduyar [55, 56, 66–68] 的原始形式，将控制方程中的所有项都纳入其中。因此，当达到精确解时，激波捕捉贡献恒为零，从而避免产生虚假的数值扩散，并确保该形式的相容性。对于耦合方程组，标量 Z 可推广为残差向量 $\mathbf{Z}$；根据备注 5，绝对值 $\left|\mathrm{Y}^{-1}\mathrm{Z}\right|$ 则替换为范数 $\|\mathbf{Y}^{-1}\mathbf{Z}\|$。

备注 7。 方程（12）中的指数 β 控制激波捕捉算子对解梯度的敏感度。具体而言， $\beta = 1$ 适用于较为陡峭的过渡，而 $\beta = 2$ 则能为极陡的解梯度提供更强的正则化。如 [68]中所述，也可以采用中间值。鉴于本文针对的是具有明显解前沿的强对流主导问题，我们在全文中采用 $\beta = 2$ 。

激波捕捉长度尺度 h_{SHOC} 定义如下：

$$h_{\mathrm{SHOC}} = \left[2 \left(\sum_{a=1}^{n_{\mathrm{en}}} |\mathbf{j} \cdot \nabla N_a| \right) \right]^{-1}, \quad \text{with} \quad \mathbf{j} = \frac{\nabla u^h}{\|\nabla u^h\|}, \tag{15}$$

其中 n_{en} i 是单元节点的数量，而 N_a 是与节点 a 对应的形函数 .

注 8。 许多激波捕捉公式会采用沿局部对流场方向加权的度量来定义长度尺度 h_{SHOC} 。尽管这类方法有望提高精度，但通常需要额外的网格相关预处理，也会增加实现的复杂性。本文采用一种更简单的策略，将 $h_{\mathrm{SHOC}} = h_K$ 设为（见式（10）），以避免方向依赖并降低计算成本。虽然对于强各向异性单元配置，这种近似的精度可能较低，但本文给出的数值结果表明，该方法在所有考察的测试用例中均能提供稳定且足够准确的预测。

注 9。 根据注 7 和注 8 中的讨论，式中的激波捕捉黏度 ν_{SHOC} 可简化为 (12)

$$\nu_{\mathrm{SHOC}} = \frac{\left|\mathrm{Y}^{-1}\mathrm{Z}\right| h_K^2}{4}. \tag{16}$$

与 SUPG 稳定化参数一样，为 ν_{SHOC} 选取合适的值仍是一项具有挑战性且本质上依赖具体问题的任务。实际应用中，通常通过经验标定或启发式调节来确定该系数。这一难题凸显了 PINN 等数据驱动方法的优势：它们可以直接依据物理约束和可用数据学习并调整稳定化参数。除了参数估计之外，PINN 还可充当后处理校正器，通过最小化由控制方程和边界/初始条件导出的残差，对稳定化有限元解进行细化。下一节将详细探讨这一能力。

注记 10. 所提出混合方法的有限元组成部分并不限于本文所采用的特定稳定化与激波捕捉方案。其他稳定化公式，包括GLS或VMS，均可无缝纳入该框架。同样，其他激波捕捉算子或层间虚假振荡衰减（SOLD）方法也可替代本文所采用的YZβ 技术。选择将SUPG稳定化与YZβ 激波捕捉相结合，主要是基于作者此前在对流主导输运问题中使用这些方法的经验。

3. 物理信息神经网络与SUPG-YZβ 形式的集成

沿用针对稳态问题在 [54], 中提出的框架，本节介绍一种混合计算策略，将稳定化有限元近似与物理信息神经

网络相结合。网络架构及多阶段训练过程如图1和图2所示。

注 11. 尽管所提出的框架适用于一维、二维或三维空间维度中的对流-扩散-反应型问题，除非另有说明，所有形式与算法描述均针对二维情形（$n_{\mathrm{sd}} = 2$）给出。采用这一约定是为了记号的简洁与可读性；向其他维度的推广是直接的。

3.1 混合*PINN-FEM*集成的动机

将物理信息神经网络与SUPG-YZβ 形式耦合，可缓解任一方法单独应用时所出现的关键缺陷。SUPG-YZβ 方法尽管在数值上稳健，却依赖于用户指定参数，其最优值通常通过经验规则或逐案标定来确定，这可能限制其在不同流动情形下的性能。相反，传统物理信息神经网络尽管概念上精巧，却常常难以分辨薄边界层，并且在强对流主导情形下可能面临训练时间过长的问题。所提出的混合策略通过若干创新克服了这些障碍。数值稳定的SUPG-YZβ 解提供了信息丰富的先验，使神经网络能够专注于分辨细尺度特征，而无需从一开始就学习基本的偏微分方程行为。与此同时，该框架利用SUPG-YZβ 离散化的稳定性来加速收敛，同时保留物理信息神经网络对复杂物理现象的表征灵活性。这构成了本研究相对于现有有限元法-物理信息神经网络耦合策略的主要贡献（参见例如 [31, 40, 41]）：通过将网络训练锚定于稳定化有限元解，该方法继承了经典数值格式的稳健性，同时受益于数据驱动学习的适应性。

3.2 网络架构与傅里叶特征映射

物理信息神经网络架构采用随机傅里叶特征映射，以增强网络捕捉边界层和内部激波特有的高频变化的能力。给定时空输入坐标$\mathbf{z} = (t,\mathbf{x}) \in \mathbb{R}^{n_{\mathrm{sd}}+1}$，傅里叶特征嵌入定义如下：

$$\boldsymbol{\phi}(\mathbf{z}) = \begin{bmatrix} \mathbf{z} \\ \sin(\mathbf{B}^\top \mathbf{z}) \\ \cos(\mathbf{B}^\top \mathbf{z}) \end{bmatrix} \in \mathbb{R}^{n_{\mathrm{sd}}+1+2n_{\mathrm{F}}}, \tag{17}$$

其中，$\mathbf{B} \in \mathbb{R}^{(n_{\mathrm{sd}}+1)\times n_{\mathrm{F}}}$ 是一个随机频率矩阵，其元素独立采样自正态分布 $\mathcal{N}(0, \sigma^2)$，并在整个优化过程中保持固定（即不参与训练）。这里，n_{F} 表示傅里叶特征的数量，σ 控制频率带宽。三角函数逐元素作用于投影$\mathbf{B}^\top \mathbf{z} \in \mathbb{R}^{n_{\mathrm{F}}}$，所得特征向量 $\boldsymbol{\phi}(\mathbf{z})$将原始坐标与正弦、余弦投影拼接起来。

傅里　　er增强特征首先通过输入层映射到隐藏维度 n_h　　r:

$$\mathbf{h}_0 = \mathtt{SiLU}\Big(\mathbf{W}^{(0)}\boldsymbol{\phi}(\mathbf{z}) + \tilde{\mathbf{b}}^{(0)}\Big), \tag{18}$$

其中$\mathbf{W}^{(0)} \in \mathbb{R}^{n_h \times (n_{\mathrm{sd}}+1+2n_{\mathrm{F}})}$ 和 $\tilde{\mathbf{b}}^{(0)} \in \mathbb{R}^{n_h}$。网络随后通过 n_r个残差块处理该表示，每个残差块按如下方式变换隐藏状态：

$$\mathbf{p}_i = \mathtt{SiLU}(\mathtt{Linear}_1(\mathbf{h}_i)), \tag{19}$$

$$\mathbf{q}_i = \mathtt{Linear}_2(\mathbf{p}_i), \tag{20}$$

$$\mathbf{h}_{i+1} = \mathtt{SiLU}(\mathtt{LayerNorm}(\mathbf{q}_i + \mathbf{h}_i)), \tag{21}$$

其中，$\mathbf{h}_i \in \mathbb{R}^{n_h}$表示 $i = 0, \mathbf{1}, \ldots, n_r - 1$中第 i 个模块的隐藏状态。这里，Linear_1 和 Linear_2表示如下形式的标准全连接（仿射）变换：

$$\mathtt{Linear}(\mathbf{v}) = \mathbf{W}\mathbf{v} + \tilde{\mathbf{b}}, \tag{22}$$

其中，可训练权重为 $\mathbf{W} \in \mathbb{R}^{n_h \times n_h}$，偏置为 $\mathbf{b} \in \mathbb{R}^{n_h}$；每个残差模块使用两个此类层，且参数集各不相同。跳跃连接 $\mathbf{q}_i + \mathbf{h}_i$ 后接层归一化，可缓解深度架构中的梯度退化。第5节报告了各项数值实验所采用的具体超参数值（n_h、 n_r、 n_{F}、 σ）。

SiLU（sigmoid加权线性单元，Swish）激活函数定义如下：

$$\mathtt{SiLU}(x) = x \cdot \mathtt{sigmoid}(x) = \frac{x}{1 + e^{-x}}. \tag{23}$$

与基于 ReLU（修正线性单元）的替代方案相比，这种平滑的非单调激活函数有助于改善梯度流，并已在物理信息学习任务中展现出更优性能，尤其适用于解变化平滑的问题。LayerNorm 通过归一化提高训练稳定性：

$$\mathtt{LayerNorm}(\mathbf{x}) = \frac{\mathbf{x} - \mu_{\mathrm{LN}}}{\sqrt{\sigma_{\mathrm{LN}}^2 + \epsilon_{\mathrm{LN}}}} \odot \boldsymbol{\gamma}_{\mathrm{LN}} + \boldsymbol{\beta}_{\mathrm{LN}}, \tag{24}$$

其中， μ_{LN} 和 σ_{LN}^2 分别为在特征维度上计算的均值和方差， $\boldsymbol{\gamma}_{\mathrm{LN}}$ 和 $\boldsymbol{\beta}_{\mathrm{LN}}$ 分别为可学习的缩放与偏移参数， ϵ_{LN} 为用于保证数值稳定性的小常数， $\odot$ 表示逐元素（Hadamard）乘积。

3.3. 数据驱动训练方法

我们使用通过采用 SUPG-YZβ 形式获得的数值稳定有限元近似，并使用均方误差数据损失函数训练物理信息神经网络模型：

$$\mathcal{L}_{\mathrm{data}}(\theta) = \frac{1}{K_s} \sum_{k=1}^{K_s} \left[\frac{1}{N_h} \sum_{i=1}^{N_h} \left| u_{\mathrm{NN}}(t_k, \mathbf{x}_i; \theta) - u^h_{\text{SUPG-YZ}\beta}(t_k, \mathbf{x}_i) \right|^2 \right], \tag{25}$$

其中， K_s 表示在终止时刻附近选取的时间快照数量， N_h 表示有限元网格 $\mathcal{T}^h$中的节点顶点数量，从而得到总计 $N = K_s \times N_h$ 个时空训练对。对于每个快照 t_k，内层求和计算所有网格节点 $\{\mathbf{x}_i\}_{i=1}^{N_h}$上的空间均方误差，外层求和则对 K_s 个选定离散时间层上的这些逐快照误差取平均。时空坐标（t_k,$\mathbf{x}_i$）对应于在离散时间层 t_k 和网格节点 $\mathbf{x}_i$处计算得到的有限元法解。该项

$u_{\mathrm{NN}}(t_k, \mathbf{x}_i; \theta)$表示神经网络预测，$u^h_{\mathrm{SUPG\text{-}YZ}\beta}(t_k, \mathbf{x}_i)$表示相应的稳定化有限元解，而 θ 表示可训练网络参数。该损失函数促使神经网络学习由稳定化有限元方法所捕捉的时空演化和边界层结构。数据驱动方法具有多项优势：(i) 利用成熟有限元求解器中蕴含的计算经验；(ii) 避免纯物理信息训练中平衡多个损失项所带来的挑战；(iii) 确保所学习的解继承底层 SUPG-YZβ 形式的稳定性性质。

注释 12。 第3.5节详述的选择性施加强制策略仅适用于偏微分方程残差损失$\mathcal{L}_{\mathrm{pde}}(\theta)$，该损失仅在内部配置点处计算。数据损失 $\mathcal{L}_{\mathrm{data}}(\theta)$和边界损失$\mathcal{L}_{\mathrm{bc}}(\theta)$均在各自的完整点集上计算，不进行任何基于距离的筛选或选择性筛选：$\mathcal{L}_{\mathrm{data}}$ 使用所有 $N = K_s \times N_h$ 时空点，以保持全局解保真度；而 $\mathcal{L}_{\mathrm{bc}}$ 则使用 $\partial\Omega$上的所有边界节点。

3.4. 采用选择性施加强制策略的混合物理信息训练

为在保持数据保真度的同时纳入物理约束，我们提出了一种采用选择性物理约束的多阶段混合训练策略。该方法认识到，计算域中的不同区域并非都同样适合基于物理的学习。混合损失函数结合了加权分量，例如：

$$\mathcal{L}_{\mathrm{hybrid}}(\theta) = w_{\mathrm{data}}\mathcal{L}_{\mathrm{data}}(\theta) + w_{\mathrm{pde}}\mathcal{L}_{\mathrm{pde}}(\theta) + w_{\mathrm{bc}}\mathcal{L}_{\mathrm{bc}}(\theta), \tag{26}$$

其中，各分量将在下文定义，非负标量权重 w_{data}、w_{pde}、w_{bc} 分别控制数据一致性、偏微分方程残差最小化和边界条件施加的相对重要性。

注释 13。这些权重用作平衡因子，用于调整各项对整体训练目标的贡献。这些权重在训练过程中的动态变化实现了一种课程学习方法：网络先学习逼近参考解，再逐步纳入物理约束。这样可以避免在建立基本解近似之前物理项占主导而导致的训练不稳定性。

偏微分方程残差损失在随机采样的内部配置点上强制满足控制型CDR方程：

$$\mathcal{L}_{\mathrm{pde}}(\theta) = \frac{1}{N_{\mathrm{int}}} \sum_{j=1}^{N_{\mathrm{int}}} \left| \mathcal{R}^{\mathrm{NN}}(t_j, \mathbf{x}_j; \theta) \right|^2, \tag{27}$$

其中，内部配点集定义为

$$\mathcal{X}_{\mathrm{int}} = \{(t_j, \mathbf{x}_j) \in [t_{\mathrm{start}}, t_{\mathrm{end}}] \times \Omega \ : \ d_{\partial\Omega}(\mathbf{x}_j) > d_{\min}\}. \tag{28}$$

此处，时空配置点 $(t_j,\ \mathbf{x}j)$ 在训练时间窗口 $[t_{\mathrm{start}}, t_{\mathrm{end}}]$ 和空间域 Ω 内均匀随机抽取，$d_{\partial\Omega}(\mathbf{x})$ 表示点 $\mathbf{x}$ 到最近边界 $\partial\Omega$ 的欧几里得距离。阈值参数 $d_{\min}$ 控制边界附近的排除区域，在该区域内不施加偏微分方程

残差；其取值取决于具体问题，应根据预期的边界层厚度和网格分辨率进行选择（详情见第3.5.1节）。术语 N_{int} 表示满足距离准则的内部配置点数量。物理信息神经网络的强形式偏微分方程残差 $\mathcal{R}^{\mathrm{NN}}$ 定义如下：

$$\mathcal{R}^{\mathrm{NN}}(t,\mathbf{x};\theta)=\frac{\partial u_{\mathrm{NN}}}{\partial t}-\varepsilon\,\Delta u_{\mathrm{NN}}+\mathbf{b}(t,\mathbf{x})\cdot\nabla u_{\mathrm{NN}}+c(\mathbf{x})\,u_{\mathrm{NN}}-f(t,\mathbf{x}),\tag{29}$$

其中，$u_{\mathrm{NN}}=u_{\mathrm{NN}}(t,\ \mathbf{x};\theta)$，所有偏导数均通过自动微分相对于时空输入坐标计算。时间导数 $\partial u_{\mathrm{NN}}/\partial t$ 完全通过自动微分获得，无需使用有限差分近似，也无需访问相邻时间步的有限元解。

边界条件损失通过惩罚神经网络预测与给定边界值之间的偏差，来强制满足 Ω 上的狄利克雷条件 $u=u_D$ ∂：

$$\mathcal{L}_{\mathrm{bc}}(\theta)=\frac{1}{N_{\mathrm{bc}}}\sum_{k=1}^{N_{\mathrm{bc}}}\left|u_{\mathrm{NN}}(t_k,\mathbf{x}_k;\theta)-u_D(t_k,\mathbf{x}_k)\right|^2,\tag{30}$$

其中，$\mathbf{x}_k\in\partial\Omega$ 是沿区域边界的每条边均匀分布的配点，$t_k\in\{t_{\mathrm{start}},t_{\mathrm{end}}]$ 在训练窗口内的离散时间层采样，u_D $(t_k,\ \mathbf{x}_k)$ 表示给定的狄利克雷数据，通常可能同时依赖于空间和时间。边界配点总数为 $N_{\mathrm{bc}}=N_{\mathrm{sides}}\times N_{\mathrm{pts}}\times N_{\mathrm{times}}$，其中 N_{sides} 为边界边数，N_{pts} 为每条边上的点数，N_{times} 为采样时间层数。

注 14。 用于边界条件损失的时间层数 N_{times} 不必与数据损失 $\mathcal{L}_{\mathrm{data}}(\theta)$ 中使用的快照数 K_s 一致。在数据损失（式 (25)）中，K_s 表示在终止时刻附近选取的有限元解快照数，每个快照贡献 N_h 个空间训练点。在边界损失（式 (30)）中，N_{times} 控制边界上的时间采样密度，可独立于 K_s 选择，以平衡边界保真度与计算成本。特别是，由于边界数据通常随时间平滑变化或呈分段常数，较小的 N_{times} 值往往就足够了。

注 15。 当网络架构中包含提升函数 u_{lift} 和距离函数 d（$\mathbf{x}$）$=x_1(1-x_1)\,x_2(1-x_2)$，使得输出具有如下形式时

$$u_{\mathrm{NN}}(t,\mathbf{x};\theta)=u_{\mathrm{lift}}(t,\mathbf{x})+d(\mathbf{x})\,\mathcal{N}(t,\mathbf{x};\theta),\tag{31}$$

狄利克雷条件在构造时便得到精确满足，因为 $d(\mathbf{x})=0$ 在 $\partial\Omega$ 上意味着边界上满足 $u_{\mathrm{NN}}=u_{\mathrm{lift}}=u_D$。在这种情况下，$\mathcal{L}_{\mathrm{bc}}\equiv 0$，因此可以从训练目标中省去边界损失项。当边界数据 u_D 足够光滑时，可以采用这种基于提升函数的强制施加方法；对于边界条件不连续或分段定义的问题，则改用显式惩罚形式 (30)。备注 16。 本文框架通过 $\mathcal{L}_{\mathrm{bc}}(\theta)$ 软性强制施加边界条件，而不是采用硬强制施加技术，例如 Luong 等人提出的统一物理信息神经网络方法 [43], ，该方法通过试函数对网络输出施加代数约束，从而精确满足本质边界条件。硬强制施加虽然无需调节 w_{bc}，并能保证精确满足狄利克雷条件，但此处采用软强制施加

有以下几个原因：(i) 在多阶段训练转换过程中，它能够灵活平衡相互竞争的目标；(ii) 选择性物理约束策略在计算偏微分方程残差时排除了边界邻近区域，而单独的边界损失则可以显式控制这些区域的边界条件保真度；(iii) 由于 SUPG-YZβ 解在构造时已满足边界条件，因此边界数据会通过 $\mathcal{L}_{\text{data}}(\theta)$ 隐式学习，使 $\mathcal{L}_{\text{bc}}(\theta)$ 主要在物理主导训练阶段发挥正则化机制的作用。

3.5. 增强型选择性物理约束策略

其关键创新在于选择性地应用物理约束，这解决了混合PINN训练中的一个根本难题——确定在何处以及何时应用基于物理的学习。我们不再在整个区域内均匀施加偏微分方程残差，而是采用多准则选择策略，将物理约束限制在最能发挥作用的区域。整体流程如图1所示。

3.5.1. 基于距离的选择准则

我们引入一个距离函数，用于衡量点 $\mathbf{x} \in \Omega$ 到矩形区域 $\Omega = [a, b]^2$的边界 $\partial\Omega$ 的距离：

$$d_{\text{boundary}}(\mathbf{x}) = \min_i \left\{ \frac{x_i - a}{b - a}, \frac{b - x_i}{b - a} \right\}, \tag{32}$$

其中，最小值是在所有空间坐标方向上取的。该函数计算到任一边界面的最小归一化距离，其中每一项表示沿第 i 个坐标方向到最近边界的距离，并按区域宽度 $(b - a)$ 进行缩放。我们采用以下距离准则：

$$d_{\text{boundary}}(\mathbf{x}) > d_{\min}, \tag{33}$$

其中，$d_{\min}$ 的取值应使物理约束与边界层保持达到最佳平衡。这种选择性约束确保只有满足距离准则的内部点才会对偏微分方程残差损失 $\mathcal{L}_{\text{pde}}(\theta)$ 作出贡献，从而增强数值稳定性并保持边界完整性。

注 17. 阈值参数 $d_{\min}$ 定义为相对于区域特征长度尺度 L 归一化的量。对于一般的有界区域 $\Omega \subset \mathbb{R}^n$，令

$$d(\mathbf{x}, \partial\Omega) = \inf_{\mathbf{y} \in \partial\Omega} \|\mathbf{x} - \mathbf{y}\|_2 \tag{34}$$

表示点 $\mathbf{x} \in \Omega$ 到边界 $\partial\Omega$ 的欧几里得距离。归一化距离定义为

$$\tilde{d}(\mathbf{x}) = \frac{d(\mathbf{x}, \partial\Omega)}{L}. \tag{35}$$

对于矩形区域 $\Omega = [0, L_1] \times [0, L_2] \times \cdots \times [0, L_n]$，特征长度为 $L = \max_{1 \le i \le n} L_i$，距离可简化为

$$d(\mathbf{x}, \partial\Omega) = \min_{1 \le i \le n} \min(x_i, L_i - x_i). \tag{36}$$

˜选择性强制条件 $d(\mathbf{x}) > d_{\min}$ 确保了具有不同区域尺寸的问题均具有尺度不变行为。对于数值算例中采用的单位区域（$L = 1$），归一化距离与绝对距离相同。

3.5.2. 实现细节与计算考量

选择性物理约束通过基于实际代码实现的高效掩码操作来实现：

$$\mathcal{S}_{\text{physics}}(\mathbf{x}) = \mathbb{I}_{d_{\text{boundary}}(\mathbf{x})>d_{\min}}, \quad \text{provided that } N_{\text{int}} \geq N_{\text{int}}^{\min}, \tag{37}$$

其中，指示函数 $\mathbb{I}_{d_{\text{boundary}}(\mathbf{x})>d_{\min}}$ 用于选取满足距离准则的内部点，而全局条件 $N_{\text{int}} \geq N_{\text{int}}^{\min}$则确保内部点的数量不少于最低要求，以便稳定地计算偏微分方程损失。

边界点的识别通过向量化距离计算来完成。对于单位正方形域 $[0,1]^2$：

$$d_{\text{left}} = x_1, \quad d_{\text{right}} = 1 - x_1, \tag{38}$$

$$d_{\text{bottom}} = x_2, \quad d_{\text{top}} = 1 - x_2, \tag{39}$$

$$d_{\text{boundary}} = \min(d_{\text{left}}, d_{\text{right}}, d_{\text{bottom}}, d_{\text{top}}). \tag{40}$$

注释 18。 方程（38）–（40）适用于数值算例中使用的单位正方形域 $\Omega = [0,1]^2$ 。对于一般矩形域 $\Omega = [0, L_1] \times [0, L_2]$，这些表达式可推广为 $d_{\text{left}} = x_1$,$d_{\text{right}} = L_1 - x_1$, $d_{\text{bottom}} = x_2$, $d_{\text{top}} = L_2 - x_2$，归一化距离为 $d(\mathbf{x}) = d_{\text{boundary}}/L$ ，其中$L = \max(L_1, L_2)$，详见注释17。

PDE配点的数量 N_{pde}是一个问题相关超参数，其取值需在计算成本与足够的物理采样密度之间取得平衡。一般而言， N_{pde} 应选取得足够大，以确保内部区域得到充分覆盖——尤其是在具有薄层或陡峭梯度的问题中——同时又应足够小，以避免二阶自动微分带来过高的计算开销。各测试用例所采用的具体取值与相应的数值算例一并在第5节中给出。类似地，边界排除阈值 $d_{\min}$ 和最小内部点数 $N_{\text{int}}^{\min}$（见式(37)）也是问题相关参数；各测试用例的取值在第5节中给出。

3.6. 多阶段权重演化

为在整个训练过程中平衡数据保真度与物理约束，我们采用一种自适应加权策略，使各项损失项的相对重要性随训练进度而演化。分配给数据损失 $\mathcal{L}_{\text{data}}$、偏微分方程残差损失 $\mathcal{L}_{\text{pde}}$和边界损失 $\mathcal{L}_{\text{bc}}$ 的相对权重会根据训练进度在不同阶段动态演化，例如：

$$\text{Phase I (Data-dominant):} \quad w_{\text{data}} = w_{\text{data}}^{\text{I}}, \quad w_{\text{pde}} = w_{\text{pde}}^{\text{I}}, \quad w_{\text{bc}} = w_{\text{bc}}^{\text{I}}, \tag{41}$$

$$\text{Phase II (Transition):} \quad w_{\text{data}} = w_{\text{data}}^{\text{II}}, \quad w_{\text{pde}} = w_{\text{pde}}^{\text{II}}, \quad w_{\text{bc}} = w_{\text{bc}}^{\text{II}}, \tag{42}$$

$$\text{Phase III (Physics-dominant):} \quad w_{\text{data}} = w_{\text{data}}^{\text{III}}, \quad w_{\text{pde}} = w_{\text{pde}}^{\text{III}}, \quad w_{\text{bc}} = w_{\text{bc}}^{\text{III}}, \tag{43}$$

其中，权重的上标表示其所属的阶段。

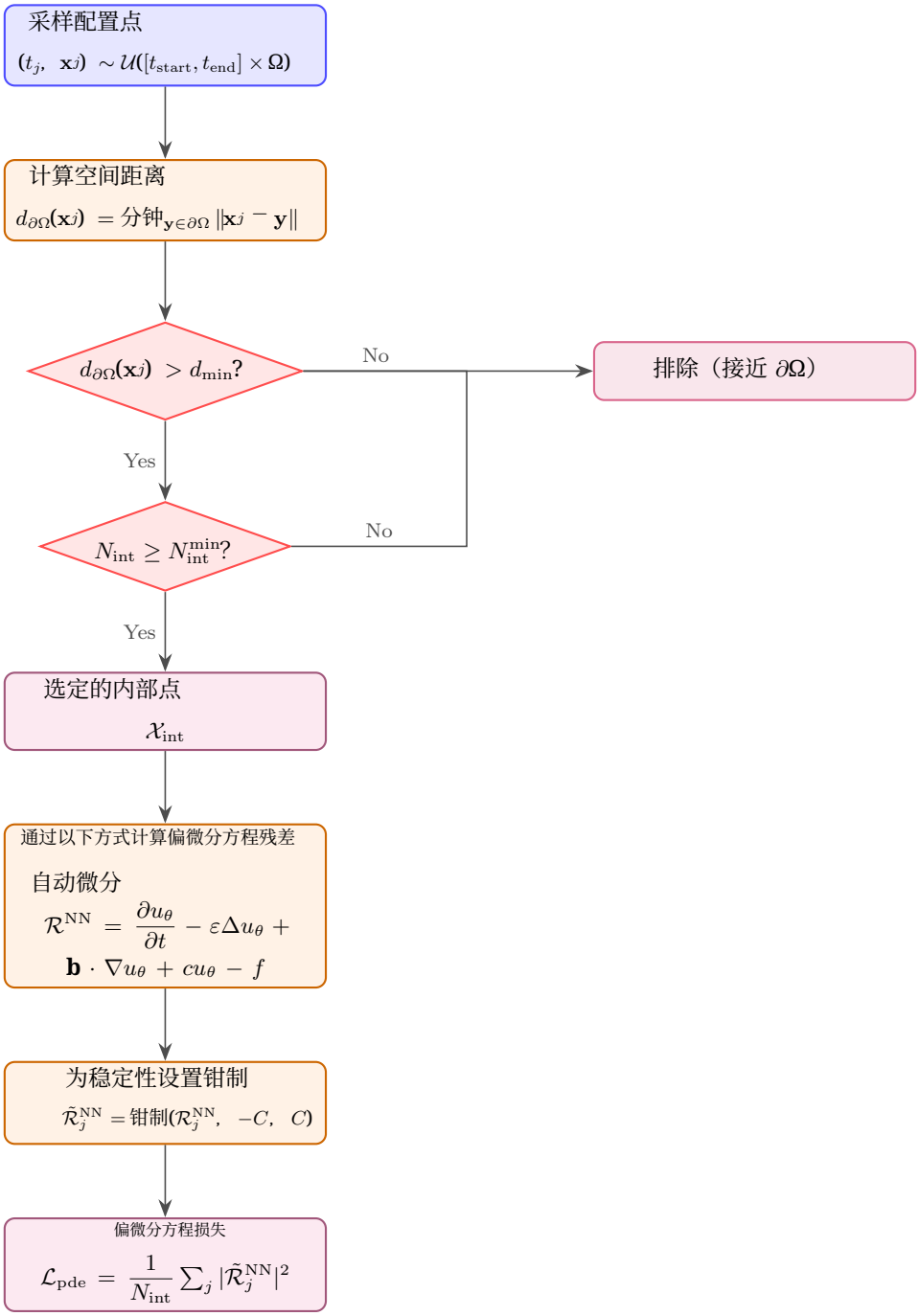


图1：选择性物理约束策略示意流程图。在每次训练迭代中，均在训练时间窗口和空间域内随机均匀采样时空配置点（t_j，$\mathbf{x}_j$）。仅保留空间距离满足 $d_{\partial\Omega}(\mathbf{x}_j) > d_{\min}$ 且符合最小计数准则 $N_{\text{int}} \geq N_{\text{int}}^{\min}$的点，用于偏微分方程残差计算。为保持陡峭梯度附近的数值稳定性，排除边界邻近点。强形式残差（包括通过自动微分获得的时间导数 $\partial u_\theta / \partial t$ ）会被截断至 $[-C, C]$ ，然后再计算损失，以防止梯度爆炸。

所采用的阶段边界和损失权重选择（见式（41）–（43））是根据数值实验确定的。例如，对于一维对流-扩散-反应问题，约 $1000-2000$ 个总训练轮次即可实现稳定且准确的收敛，而二维问题通常需要约 $2500-5000$ 个训练轮次。因此，这些数值应视为基于经验的指导，而非普遍最优的超参数；对于其他偏微分方程或几何形状，可能需要进行调整。尽管近年来，PINN 领域的相关研究已探索了更具原则性的动态加权策略，例如梯度范数平衡或基于不确定性的缩放，但我们经验证调优的三阶段方法在所考虑的所有基准问题上均表现稳健。将此类自适应机制融入混合有限元–物理信息神经网络框架，是未来颇具前景的研究方向，有望进一步提升该方法的泛化能力。

采用这种渐进式策略有三方面的理由：（i）以数据主导学习作为初始化，可避免在形成合理的解近似之前物理项占主导而导致的早期训练不稳定性；（ii）逐步过渡使网络能够平稳适应不断变化的优化景观；（iii）最后阶段的物理主导精细化通过直接施加控制偏微分方程，使混合方法有望进一步改进有限元参考解。各阶段的具体权重取决于问题，将在数值算例中报告。

*3.7. FEM-PINN*数据传输与计算接口

FEniCS与PyTorch之间的无缝集成通过精心设计的数据传输协议实现，该协议在保持数值精度的同时支持高效的GPU加速训练。空间坐标$\{\mathbf{x}_i\}_{i=1}^{N}$通过FEniCS函数空间映射，借助V.tabulate_dof_coordinates()从有限元网格中提取，并与快照时刻 $\{t_k\}_{k=1}^{K_s}$配对，以构造时空训练元组 $\{(t_k,\mathbf{x}_i,u^k_{\text{SUPG-YZ}\beta}(\mathbf{x}_i))\}$。这确保了有限元节点与跨多个时间层用于损失计算的点之间的精确对应。利用FEniCS内置的点求值功能，对每个快照处的 $u^k_{\text{SUPG-YZ}\beta}(\mathbf{x})$进行高精度求值，从而保持有限元逼近的精度。采用空间容差 $\epsilon_{\text{boundary}}=10^{-8}$ 和几何邻近检验对边界点进行高效检测。

该混合框架利用PyTorch的自动微分能力来计算时空偏微分方程残差。网络接受三维输入(t,x_1,x_2)，并产生标量输出$u_{\text{NN}}(t,x_1,x_2)$。时间导数 $\partial u_{\text{NN}}/\partial t$ 与空间梯度 ∇u_{NN} 使用torch.autograd.grad并设置create_graph=True同时计算，从而能够通过同一计算图进行高阶微分。拉普拉斯算子 Δu_{NN} 通过对空间一阶导数的逐次微分获得。强形式偏微分方程残差 $\mathcal{R}^{\text{NN}}$在训练时间窗口 $[t_{\text{start}},t_{\text{end}}]$内的随机采样内部配置点处求值，并通过距离掩码$(d>0.02)$排除靠近边界的点，以避免与显式边界条件损失相互干扰。对多个配点上所有导数的高效批量计算利用了GPU并行化。

3.8. 优化策略与训练配置

训练采用 AdamW 随机优化算法 [70], ，因其相较于标准 Adam [71]具有更强的泛化能力和解耦的权重衰减机制。优化器

配置了动量参数 $\beta_1 = 0.9$、 $\beta_2 = 0.999$，以及 $\lambda = 10^{-7}$ 的权重衰减系数，作为隐式 L^2 正则化，以防止过拟合并促进更平滑的解剖面——这对内部层问题尤其有益，因为解的正则性至关重要。

初始　　学习率根据批量大小，按照平方根缩放规则进行调整　　：

$$\alpha = 8 \times 10^{-5} \times \sqrt{\frac{B}{256}}, \tag{44}$$

其中， B 表示小批量的大小。这一缩放启发式方法能够在不同批量配置下保持梯度方差稳定，并防止不同损失分量的梯度方向相互冲突时出现振荡；在时变设置中，时间导数会引入额外的梯度竞争，使这一现象更加显著。

ReduceLROnPlateau 调度器会监控轮次平均总损失，并在停滞 150 个训练轮次后，将学习率乘以衰减因子 $\gamma = 0.9$ ，同时将最小学习率下限设为 $\alpha_{\min} = 10^{-6}$。较长的耐心周期能够适应物理信息损失景观固有的变化性；在网络学会平衡多个目标的过程中，暂时的平台期往往预示着显著的改进。乘法衰减 $\alpha \leftarrow \gamma\alpha$ 能够实现渐进式退火，在保持训练动量的同时，促进后续训练轮次中的精细收敛。

3.8.1. 梯度稳定化与数值保护措施

为确保梯度稳定性并抑制训练过程中的数值发散，我们从梯度、损失和残差三个层面实施全面的数值保护措施。

采用最大范数约束 $\|\nabla_\theta \mathcal{L}\|_2 \leq 1.0$进行梯度裁剪，以防止在解梯度陡峭、二阶导数可能变大的区域发生灾难性的参数更新。向量化梯度验证机制会在每个优化步骤检查拼接后的参数梯度张量：

$$\text{if } \neg\,\texttt{isfinite}\left(\bigoplus_{i=1}^{L} \nabla_{\theta_i}\mathcal{L}\right) \text{ then discard update and increment NaN counter}, \tag{45}$$

其中， $\bigoplus$ 表示对所有 L 参数组进行拼接。如果累计 NaN 计数超过 10，则终止训练，这表明数值不稳定性持续存在。在损失层面，也会对各个分量进行类似的保护：

$$\text{if } \texttt{isnan}(\mathcal{L}_{\text{hybrid}}) \text{ or } \texttt{isinf}(\mathcal{L}_{\text{hybrid}}) \text{ then } \mathcal{L}_{\text{hybrid}} \leftarrow \mathcal{L}_{\text{data}}, \tag{46}$$

确保训练在偏微分方程残差出现瞬时尖峰时，能够仅依靠数据保真损失恢复。

此外，在计算的多个阶段中，采用 clamp(x, a, b) 函数进行数值裁剪，以截断极端值：

$$\text{Residual clipping:} \quad \mathcal{R}^{\text{NN}}(\mathbf{x}) \leftarrow \texttt{clamp}\big(\mathcal{R}^{\text{NN}}(\mathbf{x}),\, -10.0,\, 10.0\big), \tag{47}$$

$$\text{Loss clipping:} \quad \mathcal{L}_{\text{hybrid}} \leftarrow \texttt{clamp}\big(\mathcal{L}_{\text{hybrid}},\, 0.0,\, 100.0\big), \tag{48}$$

其中，clamp 函数定义如下：

$$\texttt{clamp}(x,a,b)=\begin{cases}a, & \text{if } x<a,\\ x, & \text{if } a\le x\le b,\\ b, & \text{if } x>b.\end{cases} \tag{49}$$

± 10.0 的残差裁剪界限经过选择，旨在抑制尖锐层附近配点的异常值贡献，同时保留具有物理意义的梯度信息。

3.8.2. 批处理与混合精度训练

所有训练数据——由 $K_s \times N_h$ 个时空元组构成——均在初始化时预先完整加载到图形处理器上，以消除主机到设备传输开销。轮次级打乱通过在设备上使用 torch.randperm 进行原地随机排列来实现，每轮产生 $\lceil N_{\mathrm{data}}/B \rceil$ 个小批次，每个小批次都会执行一次完整的优化器更新。每轮训练期间会定期评估偏微分方程残差，而非在每个小批量上都进行评估，从而将计算开销巨大的二阶自动微分分摊到多个数据批次上。在默认配置中，每轮都会在第一个小批量上计算残差；不过，该频率被视为可调参数，可根据具体问题进行调整——例如，在提高物理采样率有益或需要进一步降低每批次计算成本时，每隔 k 个小批次评估一次残差。各测试用例采用的具体评估频率见第5节。在所有情况下，每次评估时都会在训练时间窗口 $[t_{\mathrm{start}}, t_{\mathrm{end}}] \times \Omega$ 内随机采样 N_{pde} 个配点。

在支持 CUDA 的设备上，有选择地通过自动混合精度（AMP）进行混合精度训练：数据保真度和边界条件的前向计算采用 float16 运算，以提高计算吞吐量；而涉及高阶自动微分的偏微分方程残差计算则保持完整的 float32 精度，以确保二阶导数计算的数值精度。动态 GradScaler 会调整损失缩放因子，以防止低精度下出现梯度下溢。根据轮次平均总损失最小值确定的最佳模型状态保存在 CPU内存中，并每 100 个训练轮次保存一次检查点到磁盘，以确保容错能力。在本文考虑的所有测试用例中，训练均进行 $N_{\mathrm{epochs}} = 5000$ 个训练轮次。

3.9. 超参数选择

网络架构采用 n_r 个残差块，隐藏维度为 n_h ，并使用尺度为 σ 的 n_{F} 个随机傅里叶特征；各测试问题所采用的具体取值见第5节。在所有示例中，架构容量的选择旨在平衡表征能力与过拟合风险：参数过少会难以解析尖锐界面，而网络规模显著增大则可能只记住有限元数据，却无法学到其底层物理规律。傅里叶特征嵌入提供了表示高频层结构所需的频谱偏差校正 [72]。

训练使用 SUPG-YZβ 解的最后 K_s 个时间快照，覆盖终止时刻附近的一个时间窗口$[t_{\mathrm{start}}, t_{\mathrm{end}}] \subset [0, t_f]$。这种窗口化方法将物理信息神经网络的校正能力集中于

近终端动力学，此时平流锋面已形成最具挑战性的特征；同时，它也为网络提供了充足的时间背景，使其能够通过自动微分学习有意义的 $\partial u/\partial t$ 行为。训练数据集总计包含 $K_s \times N_h$ 个时空点； K_s 的取值和批量大小 B 取决于具体问题，并在各数值示例中一并说明。

对于具有齐次狄利克雷数据的问题，备注15中描述的基于提升的距离函数强制使边界损失变得不再必要（$w_{\mathrm{bc}} = 0$）。当边界数据分段不连续或非齐次时，则改为使用显式边界损失，每条边设置 n_{bc} 个配点。类似地，PDE配点的数量 N_{pde} 和边界排除阈值 $d_{\min}$ 会针对每个问题分别选取，以平衡物理采样密度与计算成本；第5节给出了有代表性的取值。

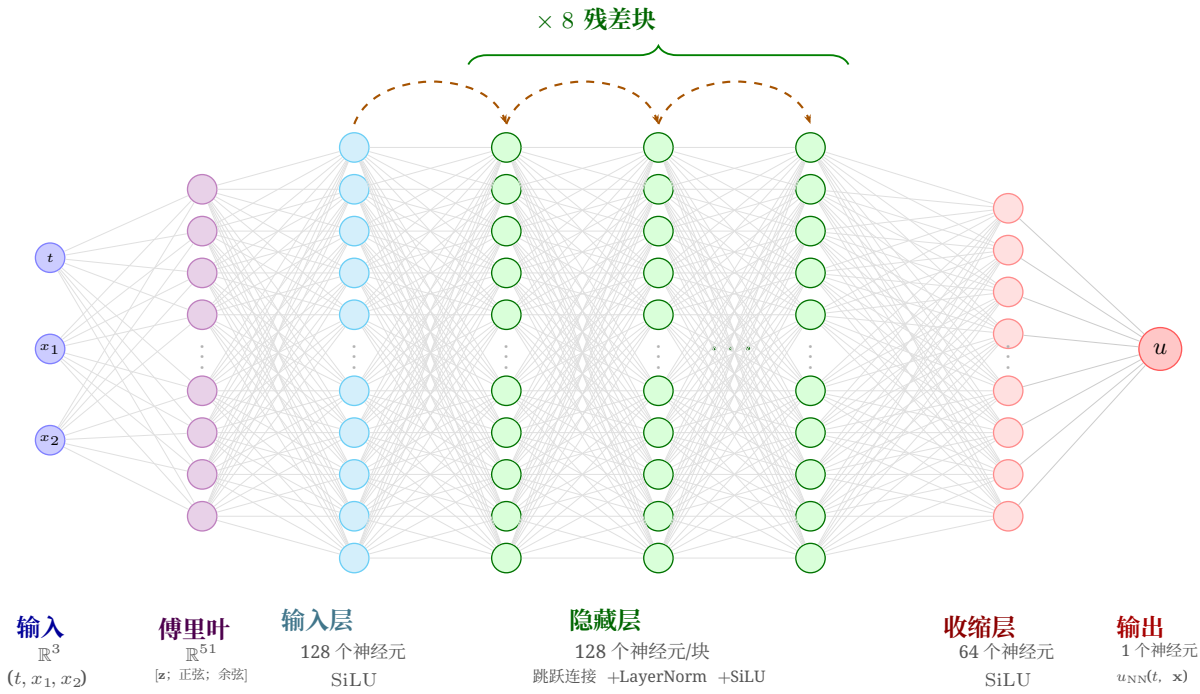


图 2：PASSC 神经网络架构示意图（展示 $n_h = 128$, $n_r = 8$, $n_{\mathrm{F}} = 24$, $n_{\mathrm{sd}} = 2$）。时空输入 $\mathbf{z} = (t,\ x_1,\ x_2) \in \mathbb{R}^3$ 经随机傅里叶特征嵌入映射为 $\boldsymbol{\varphi}(\mathbf{z}) \in \mathbb{R}^{51}$，通过带有 SiLU 激活函数的输入层投影至 128 维隐藏维度，随后由 8 个残差块进行处理——每个残差块均包含两个全连接层、一个跳跃（残差）连接和层归一化。经过收缩层 ($128 \to 64$) 后，单个输出神经元生成预测解 $u_{\mathrm{NN}}(t,\ \mathbf{x})$。

4. 更多计算细节

本节介绍为实现和验证所提出的混合方法而开发的计算框架，涵盖科学计算环境、空间离散格式、网格生成策略、所得线性与非线性代数方程组的求解技术，以及 PyTorch 环境和 CUDA 技术。

4.1. 科学计算环境：*FEniCS*

FEniCS [32, 33, 73] 是一个开源科学计算平台，专为偏微分方程（PDE）的数值求解而设计。它提供了稳健且灵活的基础设施，用于开发基于有限元法的求解器，因此被选为本研究的主要计算环境。凭借其多功能性，FEniCS 为实现稳定化有限元方法和物理信息神经网络（PINNs）提供了理想框架。其 C++ 和/或基于 Python 的接口可简洁地表达复杂数学模型，并能与众多机器学习库无缝集成——这是在统一环境中优化稳定化和激波捕捉参数的一项关键能力。此外，FEniCS 兼容高性能计算（HPC）系统，支持并行计算，这对于求解大规模、计算密集型问题至关重要。在处理时变问题和三维偏微分方程时，这种可扩展性尤具优势，有助于增强本研究中数值研究的稳健性和灵活性。本研究采用 FEniCS（2019.2.0.64.dev0）作为核心计算平台，用于实现和验证标准及稳定化有限元公式，并将其与物理信息神经网络相结合。SUPG稳定化和 YZβ 激波捕捉技术在 FEniCS 环境中开发，并与机器学习策略相结合。

FEniCS 还有一个更新的版本，称为 FEniCSx [74], ，它通过现代 C++ 架构提供更高的性能、改进的可扩展性，以及与当代科学计算生态系统更好的集成。有关 FEniCS项目旧版与新版的更多详情，感兴趣的读者可参阅项目网页 https://fenicsproject.org/。

4.2. PyTorch 与 *CUDA*

本研究将 PyTorch [75] 用作实现物理信息神经网络的核心深度学习框架。其动态计算图、友好的用户界面，以及通过 CUDA（统一计算设备架构）对 GPU 加速的有力支持，使其成为在复杂的物理驱动场景中训练神经网络的合适选择。PyTorch 能够对损失函数进行自动微分，这些损失函数不仅包含标准偏微分方程残差，还包含第 3 节所介绍的 SUPG稳定化与 YZβ 激波捕捉项的贡献。这一能力对于在反向传播过程中高效计算梯度并以稳定方式更新模型参数至关重要。为提升计算性能，尤其是针对高分辨率时空问题与深层网络架构，所有基于 PyTorch 的物理信息神经网络模型均使用 CUDA 在英伟达图形处理器上进行训练。CUDA 后端可显著加速前向传播与梯度计算，从而大幅缩短总训练时间。

Python 生态系统促进了 PyTorch 与 FEniCS 之间的集成，使有限元求解器与神经网络之间能够无缝传输数据。具体而言，在 FEniCS 中计算得到的单元级参数、几何特征以及偏微分方程残差等量被作为输入传递给 PyTorch 网络，从而在训练期间实现局部自适应。

有关 PyTorch 的更多详情，请参阅该项目的官方网站：https://pytorch. org/。

4.3. 有限元网格构建

本研究考虑了一维和二维几何形状，并生成有限元网格以支持数值模拟。研究采用 FEniCS 内置的 mshr 模块，在 Python 环境中直接生成有限元网格。这使得网格生成更加快捷，并能无缝集成到整体模拟工作流程中。

4.4. 时间积分

采用后向欧拉法对稳定化有限元公式所得到的空间半离散系统进行时间推进。已知时间层 t^n 的离散解 $u^{h,n} \in \mathcal{S}_u^h$，通过求解以下全离散变分问题，可得到时刻 $t^{n+1} = t^n + \Delta t$ 的解 $u^{h,n+1}$：求 $u^{h,n+1} \in \mathcal{S}_u^h$，使得

$$\begin{aligned}&\int_\Omega \frac{u^{h,n+1} - u^{h,n}}{\Delta t}\, w^h \, d\Omega + \int_\Omega \left(\varepsilon \, \nabla u^{h,n+1} \cdot \nabla w^h + (\mathbf{b}^{n+1} \cdot \nabla u^{h,n+1})\, w^h + c\, u^{h,n+1}\, w^h - f^{n+1}\, w^h \right) d\Omega \\ &+ \sum_{e=1}^{n_{\mathrm{el}}} \int_{\Omega^e} \tau_{\mathrm{SUPG}} \, (\mathbf{b}^{n+1} \cdot \nabla w^h)\, \mathcal{R}^h(u^{h,n+1})\, d\Omega + \sum_{e=1}^{n_{\mathrm{el}}} \int_{\Omega^e} \nu_{\mathrm{SHOC}}\, \nabla w^h \cdot \nabla u^{h,n+1}\, d\Omega = 0,\end{aligned} \tag{50}$$

对所有 $w^h \in \mathcal{V}_u^h$ 均成立，其中，$\mathbf{b}^{n+1} = \mathbf{b}(t^{n+1}, \mathbf{x})$ 和 $f^{n+1} = f(t^{n+1}, \mathbf{x})$ 分别表示在隐式时间层取值的对流速度和源项。对于对流速度依赖于解本身的非线性问题（例如，Burgers 方程中 $\mathbf{b} = (u, u)^\top$），其隐式依赖关系 $\mathbf{b}^{n+1} = \mathbf{b}(u^{h,n+1})$ 通过下一节所述的牛顿 – 拉夫森线性化处理。时间层 $n+1$ 的强形式残差定义为

$$\mathcal{R}^h(u^{h,n+1}) = \frac{u^{h,n+1} - u^{h,n}}{\Delta t} - \varepsilon \, \Delta u^{h,n+1} + \mathbf{b}^{n+1} \cdot \nabla u^{h,n+1} + c\, u^{h,n+1} - f^{n+1}. \tag{51}$$

4.5. 线性与非线性代数方程组的数值求解

如式 (11) 所述，控制方程的稳定化形式导出一组非线性方程。这些非线性方程组采用牛顿–拉夫森（N–R）法求解，该方法为非线性偏微分方程的线性化与迭代求解提供了稳健框架。在每一次非线性迭代中，所得线性方程组根据问题规模与计算设置采用不同策略处理。对于小规模问题，由于效率高且易于实现，采用基于 LU（下三角-上三角）分解的直接求解器。在串行计算中，利用 FEniCS 中内置的 PETSc（可移植、可扩展科学计算工具包） [76] 支持来访问多种线性代数例程。当由于内存限制或问题规模过大而使直接求解器不可行时，采用 GMRES（广义最小残量法） [77] 等迭代求解器。GMRES 辅以 ILU（不完全下三角-上三角）预条件，以加速稀疏且条件不良系统的收敛。

5. 测试计算

本节通过一组全面的数值算例，评估所提出混合框架的性能与稳健性。测试问题经过精心挑选，涵盖范围广泛的瞬态

情景，旨在评估该方法的准确性、稳定性和自适应能力，尤其是在存在陡峭梯度、不连续性和激波类特征的情况下。在适用时，本文还与经典方法进行比较，以突出所开发方法的优势。所有计算均在一台配备 Intel Core Ultra 9 275HX 处理器（24 核）和 NVIDIA GeForce RTX 5070 GPU（8 GB 显存）、运行 Ubuntu 24.04.3 LTS 和 CUDA Toolkit 13.0 的系统上完成。

5.1. 示例 *1*。

作为第一个 测试示例，考虑以下一维抛物型初边值问题 [7 8]:

$$\begin{cases} u_t - \varepsilon u_{xx} + (1 + x(1-x))u_x = f(t,x), & (t,x) \in (0,1] \times (0,1), \\ u(0,x) = u_0(x), & 0 < x < 1, \\ u(t,0) = 0, \quad u(t,1) = 0, & 0 \le t \le 1, \end{cases} \tag{52}$$

其中 t 初始数据 $u_0(x)$ 和源函数 $f(t,\ x)$ 选取为使其符合精确解

$$u(t,x) = \exp(-t)\left(C_1 + C_2 x - \exp\left(\frac{x-1}{\varepsilon}\right)\right), \tag{53}$$

其中 $C_1 = \exp(-1/\varepsilon)$ 和 $C_2 = 1 - \exp(-1/\varepsilon)$。

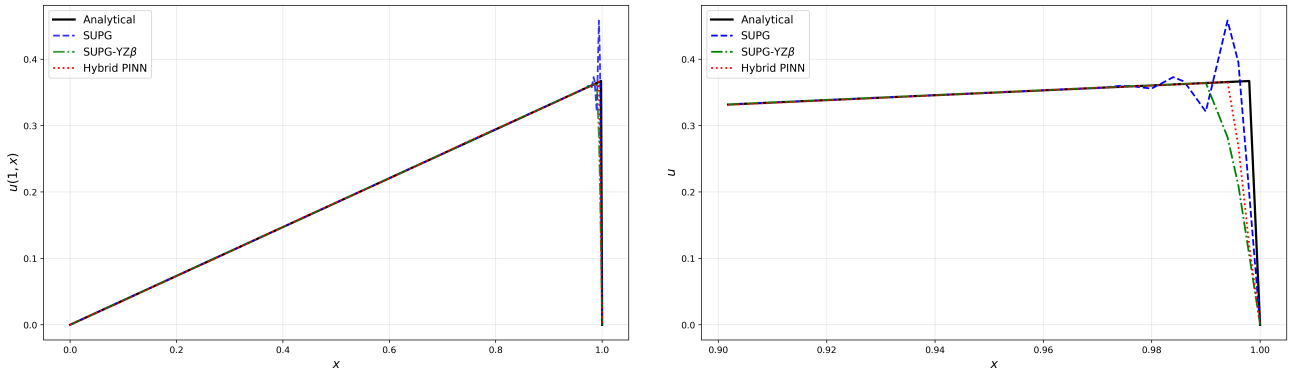
注 19。 在示例 *1*中，由于问题是一维的，我们通过使用 x 代替 x_1 来简化表示空间坐标的记号。

对于有限元设置，我们选择 $\varepsilon = 10^{-4}$、 $n_{\mathrm{el}} = 200$、 $\Delta t = 0.0025$ （$N_t = 400$）和 Y $= 0.1$。物理信息神经网络架构采用 $n_h = 128$、 $n_r = 8$、 $n_{\mathrm{F}} = 24$ 和 $\sigma = 4.0$。训练使用终止时刻 $t_{\mathrm{f}} = 1.0$ 附近的最后 $K_s = 10$ 个时间快照，批量大小为 256，初始学习率为 $\alpha = 10^{-4}$，梯度裁剪阈值为 $g_{\max} = 0.3$。每三个小批量，在 $N_{\mathrm{pde}} = 512$ 个随机采样的内部配置点处评估偏微分方程残差，以降低计算成本。阶段边界和自适应权重见表 1。

表 1：求解示例 *1*时使用的阶段边界和自适应权重。

阶段	训练轮次	w_{data}	w_{pde}	w_{bc}
I	0–1499	1.0	0.5	0.1
II	1500–2999	0.8	0.95	0.1
III	3000–4999	0.35	5.0	0.1

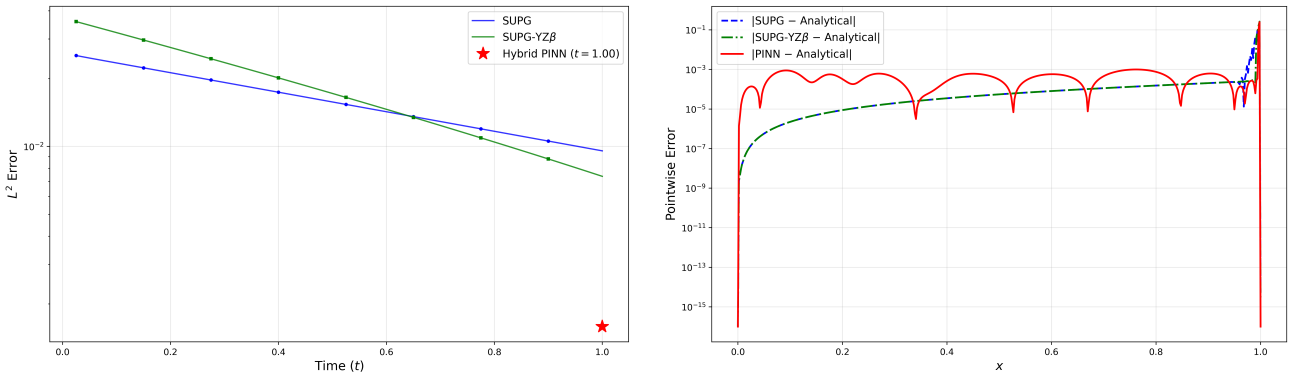
图3展示了终止时刻 $t_{\mathrm{f}} = 1$ 时 $\varepsilon = 10^{-4}$的数值解。在全局尺度上（图3a），三种数值近似——SUPG、SUPG-YZβ和混合 PINN——在区域内部的视觉表现都与解析解无法区分，此处解变化平滑且近乎线性。关键差异出现在 $x = 1$附近的边界层区域，如图3b所示的放大图。在那里，SUPG解呈现出明显的节点间振荡，并向



(a) (b)

图3：SUPG、SUPG-YZβ和混合PINN所得数值解与解析解在示例 ***1*** 且 $\varepsilon = 10^{-4}$下的比较：(a) 整个区域；(b) $x = 1$附近边界层区域的放大图。

区域内部延伸，这是指数陡峭梯度宽度为 $\mathcal{O}(\varepsilon)$ 时横风耗散不足所导致的众所周知的伪影。加入 YZβ 激波捕捉算子可抑制振荡过冲，得到单调剖面；然而，SUPG-YZβ 解会使该层略微弥散，因为各向同性人工扩散将过渡区域拓宽至超出其解析范围。混合 PINN 校正则恢复了更尖锐、更准确的边界层表示，与解析剖面高度吻合。这一改进得益于物理信息精化：网络从 SUPG-YZβ 参考数据中学习全局解结构后，通过自动微分施加控制偏微分方程残差，从而发挥正则化项的作用，使该层更加尖锐，同时不会重新引入伪振荡。

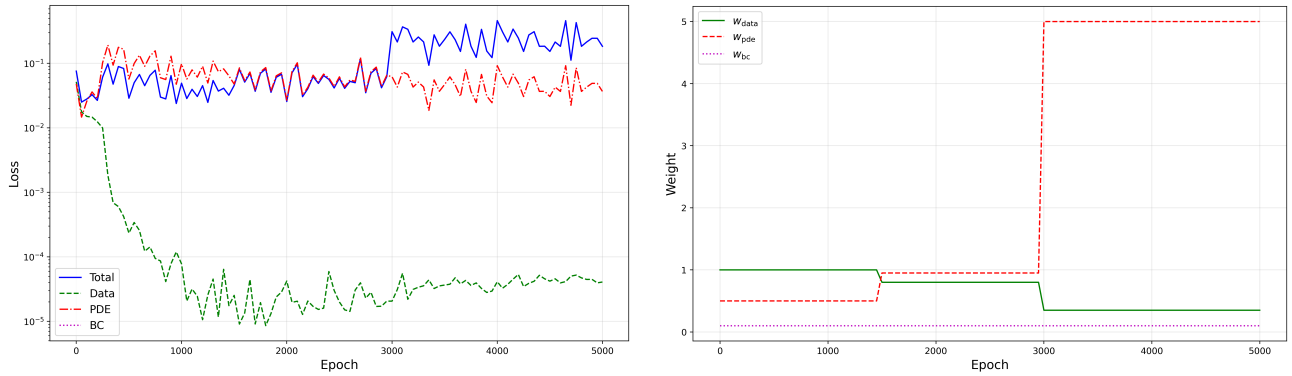


(a) (b)

图4：示例 ***1*** 在 $\varepsilon = 10^{-4}$下的误差分析：(a) SUPG和SUPG-YZβ在时间区间 $[0, 1]$ 上 L^2 误差的演化，终止时刻 $t_{\mathrm{f}} = 1$ 的混合 PINN 误差以星形标记标出；(b) $t_{\mathrm{f}} = 1$ 处空间域上的逐点绝对误差。

图4中的误差分析对这三种方法进行了定量评估。图4a展示了 L^2 误差范数随时间的演变。在整个模拟区间内，两种有限元求解器的误差均单调递减，其中 SUPG-YZβ 形式始终优于仅采用 SUPG 的离散化

，这是因为额外的激波捕捉耗散。值得注意的是，终止时刻$t_{\mathrm{f}}=1$ 的混合 PINN 校正（红色星号）使 L^2 误差比 SUPG-YZβ 解低约两个数量级，展现了物理信息后处理步骤带来的显著精度提升。图 4b 所示的 $t_{\mathrm{f}}=1$ 处点误差空间分布，进一步揭示了各方法之间的结构差异。SUPG解在平滑内部区域的误差维持在 $\mathcal{O}(10^{-5})$，但在 $x=1$附近出现尖峰，此处的边界层振荡导致误差超过 $\mathcal{O}(10^{-1})$。SUPG-YZβ 解在整个区域内将误差更均匀地分布在 $\mathcal{O}(10^{-5})$，有效消除了边界层尖峰；代价是额外的人工扩散使平滑区域的误差略有升高。相比之下，混合 PINN 误差呈现出截然不同的特征：点误差在 $\mathcal{O}(10^{-7})$和 $\mathcal{O}(10^{-3})$之间振荡，并在若干局部位置降至接近机器精度。这种非均匀误差分布反映了傅里叶特征嵌入的谱逼近特性：它将表示能力分配到多个频率尺度，而非使误差均匀分布。



(a) (b)

图 5：使用 $\varepsilon=10^{-4}$ 的示例 *1* 训练诊断：(a) 在对数尺度下，5000 个训练轮次中各个损失分量（总损失、数据损失、偏微分方程损失和边界条件损失）的变化；(b) 控制三阶段训练策略的自适应权重调度 w_{data}、w_{pde} 和 w_{bc}。

示例 *1* 的混合PINN训练动态如图 5 所示。图 5a 中的损失演化揭示了三个训练阶段中各个分量的不同表现。在第一阶段（训练轮次 0–1499），数据保真度占主导地位，权重为 $w_{\mathrm{data}}=1.0$，数据损失从 $\mathcal{O}(10^{-2})$ 快速下降至 $\mathcal{O}(10^{-4})$，表明网络成功地从 SUPG-YZβ参考快照中学习了全局解结构。此阶段的偏微分方程损失保持在 $\mathcal{O}(10^{-1})$，因为物理约束仅获得了适中的权重（$w_{\mathrm{pde}}=0.5$）。过渡到第二阶段（训练轮次 1500–2999）后，偏微分方程的影响略微增至 $w_{\mathrm{pde}}=0.95$，同时数据权重降至 $w_{\mathrm{data}}=0.8$，使学习过程逐渐转向物理驱动学习。在第三阶段（训练轮次 3000–4999），偏微分方程权重急剧升至 $w_{\mathrm{pde}}=5.0$，而数据权重降至 $w_{\mathrm{data}}=0.35$，如图 5b 的权重调度所示。总损失随之增加，反映出偏微分方程残差的贡献此时已占主导；然而，数据损失仍持续稳步下降，证实物理信息精化并未损害已学习到的解保真度。整个训练过程中，边界条件损失始终为零，因为基于提升的距离函数 $d(x)=x(1-x)$ 施加

通过构造精确满足齐次狄利克雷条件（见备注15），因此权重 $w_{\mathrm{bc}} = 0.1$ 无关紧要。

5.2. 例*2*。

考虑以下二维对流-扩散-反应方程（隆起高度变化） [17, 79]:

$$\begin{cases} \dfrac{\partial u}{\partial t} - \varepsilon \Delta u + \mathbf{b} \cdot \nabla u + cu = f, & (x_1, x_2) \in \Omega, \quad t \in (0, t_{\mathrm{f}}], \\ u(0, x_1, x_2) = u_0(x_1, x_2), & (x_1, x_2) \in \Omega, \\ u = 0, & (x_1, x_2) \in \partial\Omega, \quad t \in [0, t_{\mathrm{f}}], \end{cases} \tag{54}$$

其中 $\Omega = (0,1)^2$, $t_{\mathrm{f}} = 0.5$, $\varepsilon = 10^{-6}$, $\mathbf{b} = (2,3)^{\top}$，且 $c = 1$。强迫项 f、初始条件 u_0和边界条件的选取使得精确解

$$u(t, x_1, x_2) = 16 \sin(\pi t)\, x_1(1 - x_1)\, x_2(1 - x_2) \left[\frac{1}{2} + \frac{\arctan\left(2\varepsilon^{-1/2}\left(0.25^2 - (x_1 - 0.5)^2 - (x_2 - 0.5)^2\right)\right)}{\pi} \right] \tag{55}$$

成立。

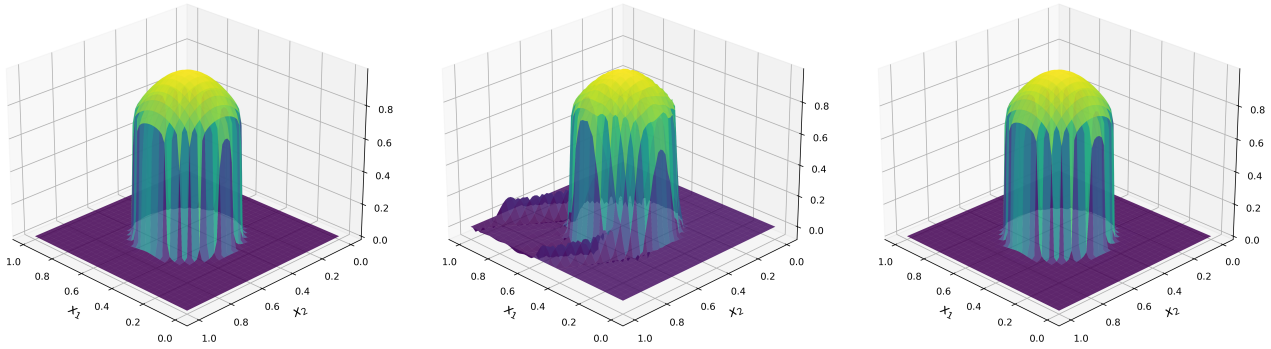
该问题也在 [17, 79] 中进行了研究。在 [17], 中，作者比较了几种激波捕捉机制，并指出它们在这一特定问题中均无法奏效，都会在隆起周围产生伪条纹。

对于有限元设置，我们选择 $\varepsilon = 10^{-6}$, $n_{\mathrm{el}} = 8{,}192$（ 64×64 网格的结构化三角剖分），$\Delta t = 0.0025$ ($N_t = 200$)，以及 $\Upsilon = 0.7$。物理信息神经网络架构采用 $n_h = 128$, $n_r = 8$, $n_{\mathrm{F}} = 24$ 和 $\sigma = 4.0$。训练使用靠近 $t_{\mathrm{f}} = 0.5$ 的最近 $K_s = 5$ 个时间快照，批量大小为 16000，初始学习率为 $\alpha = 8 \times 10^{-5}\sqrt{N_{\mathrm{batch}}/256} \approx 6.3 \times 10^{-4}$ （平方根缩放），梯度裁剪阈值为 $g_{\max} = 1.0$。每个小批量在 $N_{\mathrm{pde}} = 256$ 个随机采样的内部配置点上评估偏微分方程残差。阶段边界和自适应权重见表2。

表2：求解例2所用的阶段边界和自适应权重。

阶段	训练轮次	w_{data}	w_{pde}	w_{bc}
I	0–999	1.0	0.05	0.01
II	1000–1499	0.5	0.1	0.05
III	1500–4999	0.1	0.5	0.1

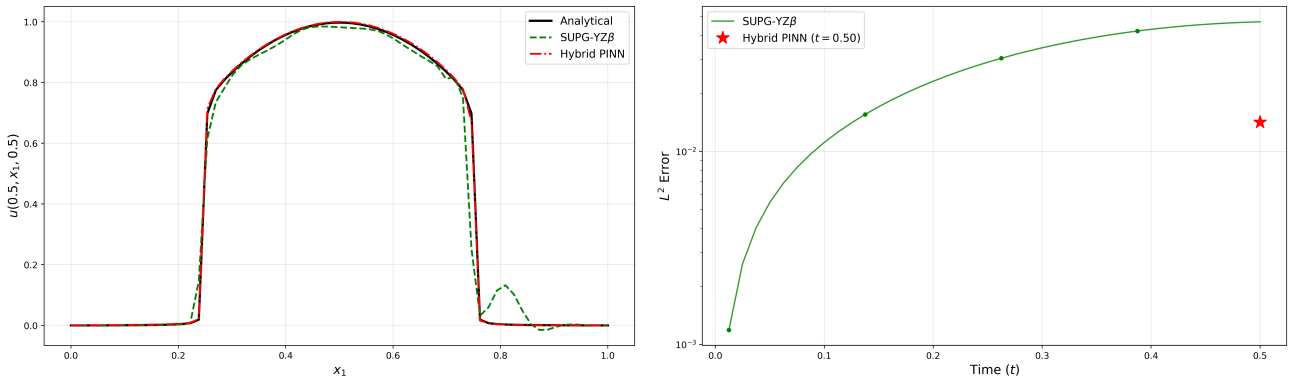
图6展示了例2在 $t = 0.5$ 时的三维解剖面。解析解（图6a）呈现一个以 $(x_1,\ x_2) = (0.5,\ 0.5)$ 为中心、具有紧支撑的平滑隆起，其振幅受因子 $16 \sin(\pi t)$ 调制，并在所选时间层达到峰值。该隆起具有极其陡峭的侧翼，其形状由宽度为 $\mathcal{O}(\sqrt{\varepsilon}) = \mathcal{O}(10^{-3})$ 的反正切层控制，因此这是一个极具挑战性的测试用例。如 [17], 所述，所有经典激波捕捉机制在此类问题上均告失效，图6b中的 SUPG-YZβ 近似证实了这一点：隆起底部及侧翼周围出现明显的伪条纹，同时在解本应恒为零的平坦区域内出现局部振荡



(a) (b) (c)

图6：例*2* 在 $t = 0.5$ 且 $\varepsilon = 10^{-6}$时解的三维曲面图：(a) 解析解；(b) SUPG-YZβ 有限元逼近；(c) 混合 PINN 校正。

。这些伪影源于各向同性激波捕捉扩散与内层的圆对称性发生不利相互作用，从而产生沿对流方向$\mathbf{b} = (2,\ 3)^{\top}$传播的横风畸变。混合PINN校正（图6c）大幅消除了这些伪影，恢复出的表面轮廓与解析解非常接近。隆起形状、峰值振幅和陡峭侧翼均得到了忠实再现，同时有效抑制了有限元近似中可见的振荡条纹。该结果展示了物理信息后处理能够校正即使在应用稳定化和激波捕捉之后仍然存在的结构性伪影。

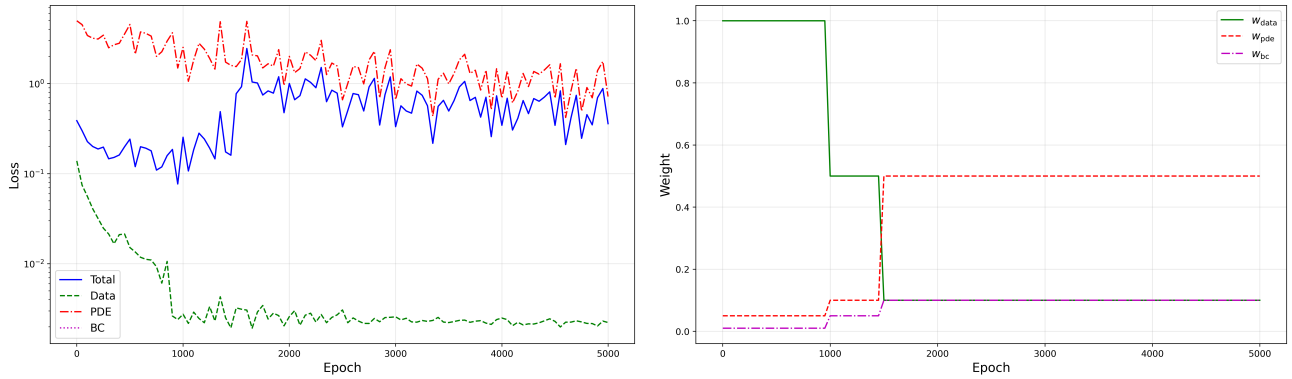


(a) (b)

图7： $\varepsilon = 10^{-6}$条件下例*2* 的定量比较：(a) $t = 0.5$时沿 $x_2 = 0.5$ 的横截面剖面，对比解析解、SUPG-YZβ和混合 PINN；(b) SUPG-YZβ 在时间区间 $[0, 0.5]$ 内的 L^2 误差演变，终止时刻的混合 PINN 误差以星形标记表示。

图7提供了更定量的评估。沿 $x_2 = 0.5$ 在 $t = 0.5$处的横截面剖面（图7a）揭示了内层的详细结构。解析解呈现出尖锐、对称的隆起，在接近单位值处形成平坦平台，陡峭侧翼则在距离 $\mathcal{O}$（$\sqrt{\varepsilon}$）内过渡至零。SUPG-YZβ 近似解捕捉到了整体形状，但存在两个典型缺陷：陡峭

侧翼相较于精确剖面更为平滑，且在 $x_1 \approx 0.8$附近出现了幅值为 ~ 0.13 的伪次级隆起，这与图6b三维视图中可见的条纹伪影相对应。混合PINN校正彻底消除了这一伪特征，并恢复了更陡峭的侧翼，使其更贴合解析剖面，但在下游过渡附近$x_1 \approx 0.75$仍观察到轻微欠冲。图7b中 L^2 误差范数随时间的变化显示，SUPG-YZβ 解的误差单调递增：早期从 $\mathcal{O}(10^{-3})$ 升至终止时刻 $t = 0.5$的 $\mathcal{O}(10^{-2})$。这一增长反映了随着隆起幅值按 $\sin(\pi t)$ 增大，条纹伪影逐渐累积。终止时刻的混合 PINN 误差（红色星号）低于 SUPG-YZβ 误差曲线，证实尽管物理信息校正仅作用于最终时间快照，仍能显著提升整体精度。



(a) (b)

图 8：使用 $\varepsilon = 10^{-6}$ 的例*2* 训练诊断：(a) 对数尺度下，5000 个训练轮次中各个损失分量（总损失、数据损失、偏微分方程损失和边界条件损失）的变化；(b) 控制三阶段训练策略的自适应权重调度 w_{data}、w_{pde} 和 w_{bc}。

例*2* 的训练诊断如图 8 所示。与示例 *1*的一维情形（见图 5）相比，损失景观呈现出截然不同的特征，反映了二维隆起问题的复杂性有所增加。如图 8a 所示，在第一阶段（训练轮次 0–999）中，数据损失从 $\mathcal{O}(10^{-1})$ 稳步下降至 $\mathcal{O}(10^{-2})$，此时数据权重以 $w_{\text{data}} = 1.0$ 占主导。然而，在整个训练过程中，偏微分方程损失始终居高不下，维持在 $\mathcal{O}(1)$–$\mathcal{O}(10)$，远高于示例 *1*。这种较高的残差是意料之中的：宽度为 $\mathcal{O}(\sqrt{\varepsilon}) = \mathcal{O}(10^{-3})$ 的陡峭内层会产生较大的点态偏微分方程残差，难以将其最小化，尤其是在隆起侧翼周围的过渡区域。

图 8b 中的权重调度揭示了针对该问题所采用的审慎物理约束策略。在第 1000 个训练轮次从第一阶段过渡到第二阶段时，数据权重从 1.0 降至 0.5，而偏微分方程权重则从 0.05 小幅升至 0.1。在第三阶段（训练轮次 1500–4999），权重稳定在 $w_{\text{data}} = 0.1$ 和 $w_{\text{pde}} = 0.5$——明显比示例 *1* ($w_{\text{pde}} = 5.0$)中采用的激进物理主导配置更为温和。这种克制是有意为之：对隆起问题施加过大的偏微分方程权重，可能会重新引入 PINN 校正旨在消除的伪条纹伪影。整个训练过程中，边界条件损失始终处于较低水平，这与通过以下距离函数施加的基于提升函数的强制施加相符

$d(\mathbf{x}) = x_1(1-x_1)\,x_2(1-x_2)$。

5.3. 例*3*。

考虑以下二维对流-扩散-反应方程（行波） [79]:

$$\begin{cases} \dfrac{\partial u}{\partial t} - \varepsilon\Delta u + \mathbf{b}\cdot\nabla u + cu = f, & (x_1,x_2)\in\Omega, \quad t\in(0,t_{\mathrm{f}}], \\ u = 0, & (x_1,x_2)\in\partial\Omega, \quad t\in[0,t_{\mathrm{f}}], \\ u(0,x_1,x_2) = u_0(x_1,x_2), & (x_1,x_2)\in\Omega, \end{cases} \tag{56}$$

其中 $\Omega = (0,1)^2$, $t_{\mathrm{f}} = 1$, $\varepsilon = 10^{-8}$, $\mathbf{b} = (\cos(\pi/3), \sin(\pi/3))^{\top}$, 且 $c = 1$。强迫项 f、初始条件 u_0 和边界条件的选取使得精确解

$$u(t,x_1,x_2) = 0.5\sin(\pi x_1)\sin(\pi x_2)\left[\tanh\left(\frac{x_1+x_2-t-0.5}{\sqrt{\varepsilon}}\right)+1\right] \tag{57}$$

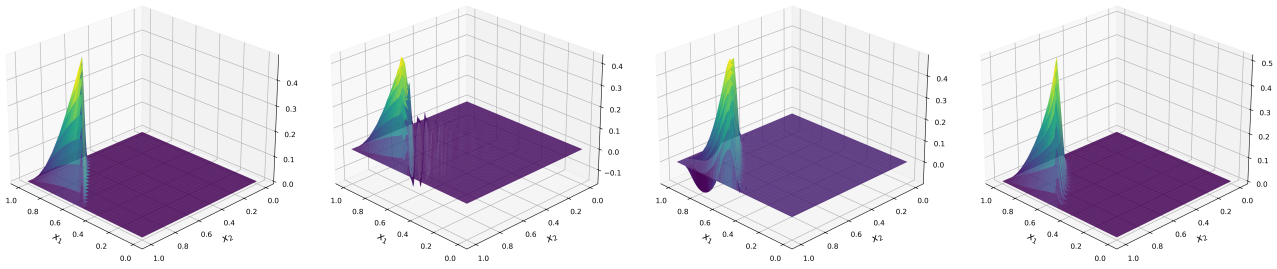
满足该方程。该解具有宽度为 $\mathcal{O}(\sqrt{\varepsilon})$ 的移动内层。

对于有限元设置，我们取 $n_{\mathrm{el}} = 8192$ （ 64×64 网格的结构化三角剖分）、 $\Delta t = 0.001 (N_t = 1000)$ 和 Y= 0.25。宽度为 $\mathcal{O}(\sqrt{\varepsilon})$ 的移动内层 $= \mathcal{O}(10^{-4})$，因此需要比例*2* 更高的时间分辨率，才能准确追踪传播前沿。物理信息神经网络架构采用较小的网络，包含 $n_h = 96$、 $n_r = 6$、 $n_{\mathrm{F}} = 16$ 和 $\sigma = 4.0$，这反映了与隆起问题相比，行波解具有更低维的结构。训练使用靠近 $t_{\mathrm{f}} = 1.0$ 的最后 $K_s = 5$ 个时间快照，批量大小为 16000，初始学习率为 $\alpha \approx 6.3\times 10^{-4}$ （平方根缩放），梯度裁剪阈值为 $g_{\max} = 1.0$。每个小批量在 $N_{\mathrm{pde}} = 384$ 个随机采样的内部配置点上评估偏微分方程残差。由于网络架构融入了基于提升的距离函数$d(\mathbf{x}) = x_1(1-x_1)\,x_2(1-x_2)$，因此在所有训练阶段均完全省略边界损失（$w_{\mathrm{bc}} = 0$），因为通过构造可精确满足狄利克雷条件（见备注15）。阶段边界和自适应权重见表3。

表3：用于求解例*3*的阶段边界和自适应权重。

阶段	训练轮次	w_{data}	w_{pde}	w_{bc}
I	0–999	1.0	0.05	0.0
II	1000–1499	0.5	0.1	0.0
III	1500–4999	0.1	0.5	0.0

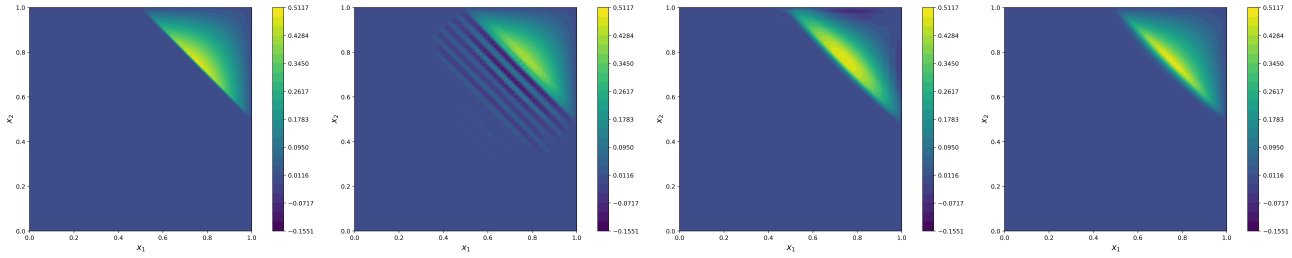
图9展示了例*3*在终止时刻 $t_{\mathrm{f}} = 1$ 的三维解剖面。解析解（图9a）呈现出一个局部行波峰，这是正弦包络 $0.5\sin(\pi x_1)\sin(\pi x_2)$ 与由宽度为 $\mathcal{O}(\sqrt{\varepsilon}) = \mathcal{O}(10^{-4})$ 的双曲正切剖面所控制的内层相互作用的结果。这个极薄的移动前沿对所有数值方法都构成了严峻挑战。SUPG解（图9b）表现极差：由于缺乏激波捕捉，未稳定的横风



(a) (b) (c) (d)

图9：例*3* 在 $t_{\mathrm{f}}=1$ 且 $\varepsilon=10^{-8}$条件下解的三维曲面图：(a) 解析解；(b) SUPG有限元近似；(c) SUPG-YZβ 有限元近似；(d) 混合PINN校正。

扩散产生了显著的下冲，幅度约为 -0.1，同时还出现了遍布整个区域的振荡伪影。加入 YZβ 激波捕捉（图9c）后，负向下冲被消除，全局振荡受到抑制，得到非负解；然而，峰值振幅明显减弱，且各向同性人工扩散使层变宽，形成弥散剖面，低估了真解的幅值。混合PINN校正（图9d）恢复了清晰的层结构和正确的峰值振幅，所得表面轮廓与解析解高度吻合。峰值定位准确，周围的平坦区域没有伪振荡，并且从高值区域到零平台的过渡，比任一种有限元近似都更加精确地得到捕捉。本例凸显了混合框架在涉及极薄移动前沿的问题中的独特优势：对于这类问题，即使经过稳定化和激波捕捉的有限元解，也会受到过度数值扩散的影响。

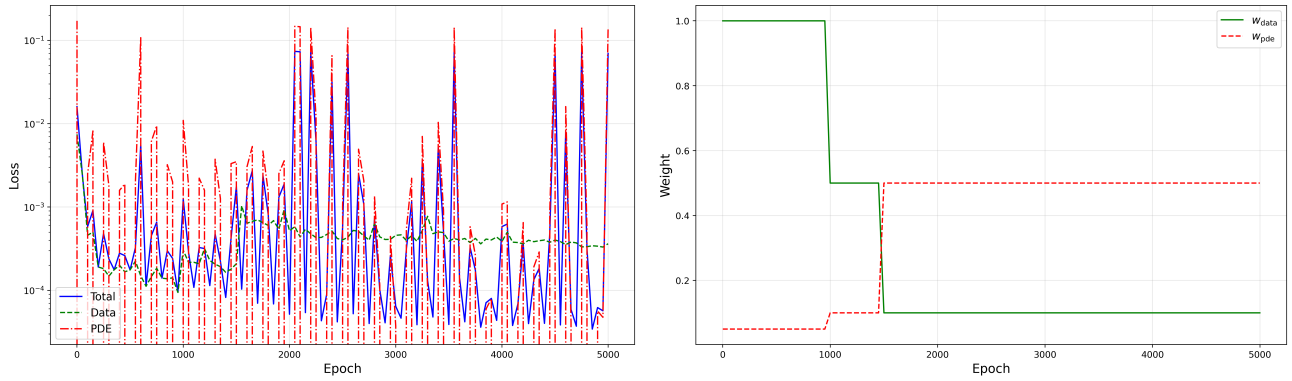


(a) (b) (c) (d)

图10：例*3* 在 $t_{\mathrm{f}}=1$ 且 $\varepsilon=10^{-8}$条件下解的等值线图：(a) 解析解；(b) SUPG有限元近似；(c) SUPG-YZβ 有限元近似；(d) 混合PINN校正。所有子图使用相同的色标。

图10中的等值线图以俯视视角展示了相同的解，更清晰地揭示了内层及其数值近似的空间结构。解析解（图10a）呈现出沿对流方向$\mathbf{b}=(\cos(\pi/3),\sin(\pi/3))^{\top}$取向的尖锐对角锋面，将右上角附近的升高区域与零值外部区域分隔开来。SUPG等值线图（图10b）将横风振荡显示为一系列平行条纹，这些条纹垂直于对流方向排列，并深入延伸至区域内部。这些伪影表现为交替的对角带

正负误差，是结构化网格上对流主导问题中横风扩散不足的典型特征。SUPG-YZβ 解（图10c）完全消除了条纹，产生了平滑的等值线场；然而，与解析参考解相比，锋面明显变宽，且右上角的峰值有所衰减。混合PINN等值线（图10d）紧密再现了解析模式：锋面锐度得以恢复，对角线取向被正确捕捉，升高区域的解幅值与参考解一致，且没有SUPG-YZβ 近似中观察到的振幅损失。



(a) (b)

图11：使用 $\varepsilon = 10^{-8}$进行例*3* 的训练诊断：(a) 在 5000 个训练轮次内，各损失分量（总损失、数据损失和偏微分方程损失）在对数尺度上的变化；(b) 控制三阶段训练策略的自适应权重调度 w_{data} 和 w_{pde} 。由于通过距离函数精确施加了狄利克雷条件，因此省略边界条件权重（$w_{\mathrm{bc}} = 0$）。

示例 *3* 的训练诊断结果如图 11 所示。图 11a 中的损失演化呈现出与前述示例截然不同的特征。偏微分方程损失出现明显的间歇性尖峰，幅度跨越近三个数量级，从 $\mathcal{O}(10^{-4})$ 到 $\mathcal{O}(10^{-1})$，并贯穿所有训练阶段。这些突增源于内层极端陡峭：当随机采样的配点恰好落在宽度为 $\mathcal{O}(\sqrt{\varepsilon}) = \mathcal{O}(10^{-4})$ 的过渡区域内时，陡峭梯度会导致偏微分方程残差幅值急剧增大，从而产生孤立的高损失批次。尽管偏微分方程损失如此波动，数据损失仍从 $\mathcal{O}(10^{-3})$ 稳步下降至 $\mathcal{O}(10^{-4})$，这表明网络成功学到了全局解结构，且未受偶发的偏微分方程残差尖峰影响而发生不稳定。尖峰出现时，总损失随偏微分方程损失项变化；而在尖峰之间，则回落至数据驱动基线。图 11b 中的权重调度与例*2*采用相同的三阶段结构，其中 w_{data} 从 1.0 降至 0.1，w_{pde} 则在三个阶段中从 0.05 升至 0.5。值得注意的是，整个训练过程中边界条件权重设为 $w_{\mathrm{bc}} = 0$，因为基于提升的距离函数 $d(\mathbf{x}) = x_1(1 - x_1)\, x_2(1 - x_2)$ 在构造时就精确满足齐次狄利克雷条件，因此无需单独设置边界条件损失项。第三阶段保守的偏微分方程权重上限为 $w_{\mathrm{pde}} = 0.5$，与例*2* 相同，且远低于示例 *1*中使用的 $w_{\mathrm{pde}} = 5.0$；这体现了对极薄层问题所需的审慎克制，因为过度加强物理约束可能会放大随机残差尖峰，并使训练过程不稳定。

5.4. 示例 *4*。

考虑以下二维非耦合 Burgers 方程 [62]:

$$\begin{cases} \dfrac{\partial u}{\partial t} + u\dfrac{\partial u}{\partial x_1} + u\dfrac{\partial u}{\partial x_2} = \dfrac{1}{\mathrm{Re}}\left(\dfrac{\partial^2 u}{\partial x_1^2} + \dfrac{\partial^2 u}{\partial x_2^2}\right), & t \in (0, t_{\mathrm{f}}], \quad (x_1, x_2) \in \Omega, \\ u(t_0, x_1, x_2) = u_0(x_1, x_2), & (x_1, x_2) \in \Omega, \\ u(t, x_1, x_2) = u_D(t, x_1, x_2), & (x_1, x_2) \in \partial\Omega, \quad t \in (0, t_{\mathrm{f}}], \end{cases} \tag{58}$$

其中，t 初始条件和狄利克雷边界条件的设定使得精确解

$$u(t, x_1, x_2) = \frac{1}{1 + \exp\left(\frac{\mathrm{Re}(x_1 + x_2 - t)}{2}\right)} \tag{59}$$

满足。

对于有限元设置，我们设定 $t_0 = 0$、 $t_{\mathrm{f}} = 1.0$、$\mathrm{Re} = 10^4$ （$\varepsilon = 1/\mathrm{Re} = 10^{-4}$）、 $n_{\mathrm{el}} = 4608$ （ 48×48 网格的结构化三角剖分）、 $\Delta t = 0.01$ （$N_t = 100$）以及 Y= 0.2。物理信息神经网络架构采用 $n_h = 96$、$n_r = 6$、 $n_{\mathrm{F}} = 16$ 和 $\sigma = 4.0$。训练使用靠近 $t_{\mathrm{f}} = 1.0$ 的最后 $K_s = 5$ 个时间快照，批量大小为 5000，初始学习率为 $\alpha \approx 3.5 \times 10^{-4}$ （平方根缩放），梯度裁剪阈值为 $g_{\max} = 1.0$。每个小批量在 $N_{\mathrm{pde}} = 256$ 个随机采样的内部配置点处计算偏微分方程残差。

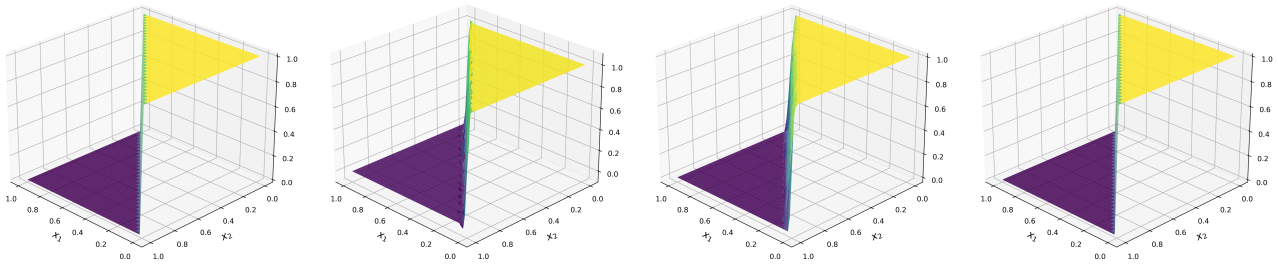
该示例训练策略的一个显著特点是，在所有训练阶段中，数据权重始终固定为$w_{\mathrm{data}} = 1.0$ ，这与示例 *1*–例*3*中采用的数据权重递减调度形成对比。作出这一设计选择是出于 Burgers 方程的非线性特性：由于仅凭偏微分方程残差无法唯一确定解（无黏极限下存在无穷多个弱解），因此必须在所有阶段维持数据保真度，使网络锚定在正确的物理解支上。相反，偏微分方程权重在四个阶段中逐步增加，从第一阶段的 $w_{\mathrm{pde}} = 0.01$ 增加到阶段 IV 的$w_{\mathrm{pde}} = 5.0$ ，在保持数据锚定的同时逐步强化物理约束。由于网络架构采用了形式为 $u_{\mathrm{lift}}(t,\ \mathbf{x}) = \left(1 + \exp(\mathrm{Re}\,(x_1 + x_2 - t)/2)\right)^{-1}$ 的提升函数，能够精确满足边界条件，因此边界条件权重在整个过程中设为 $w_{\mathrm{bc}} = 0$ 。阶段边界和自适应权重见表4。

表4：用于求解例*4*的阶段边界和自适应权重。

阶段	训练轮次	w_{data}	w_{pde}	w_{bc}
I	0–999	1.0	0.01	0.0
II	1000–1999	1.0	0.1	0.0
III	2000–3499	1.0	1.0	0.0
IV	3500–4999	1.0	5.0	0.0

备注20。对于具有非齐次边界条件的问题，提升函数 u_{lift} 必须满足给定边界数据，但不必与精确解一致。在本例中，二维Burgers方程的解析

解可作为方便的提升；不过在实践中，u_{lift} 仅凭边界数据即可构造，例如通过调和延拓或将边界值双线性插值到内部区域。或者，SUPG-YZβ 有限元解本身也可用作提升函数，此时网络修正 $d(\mathbf{x})\mathcal{N}_\theta(t,\ \mathbf{x})$ 学习有限元近似与真解之间的逐点偏差。

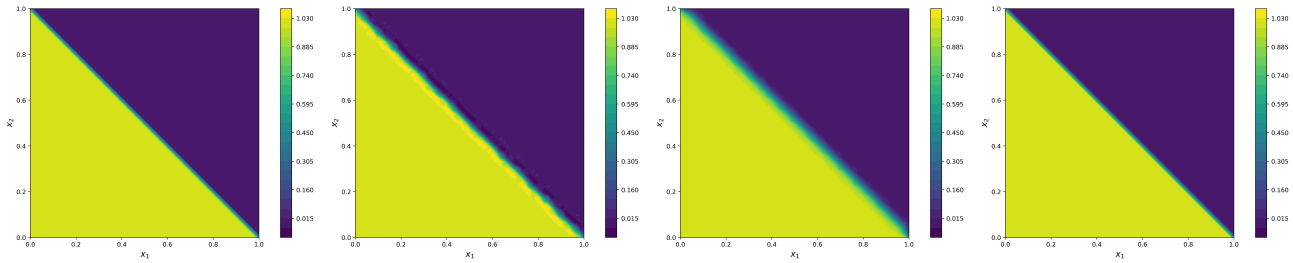


(a) (b) (c) (d)

图12：在 $t_{\text{f}} = 1$ 且 Re 为 $= 10^4$ 时，例*4* （二维Burgers方程）解的三维曲面图：(a) 解析解；(b) SUPG有限元近似；(c) SUPG-YZβ 有限元近似；(d) 混合PINN校正。

图12展示了 $t_{\text{f}} = 1$ 时例*4*的三维解剖面。解析解（图12a）沿对角线 $x_1 + x_2 = t$呈现 sigmoid 型内层，从区域 $x_1 + x_2 < t$ 中的$u \approx 1$ 到其补集中的 $u \approx 0$ 发生陡峭过渡；内层宽度受雷诺数控制，如 $\mathcal{O}(1/\text{Re}) = \mathcal{O}(10^{-4})$。这个问题尤其具有挑战性，因为对流速度依赖于解，导致前沿位置和陡峭度通过非线性自平流机制相互耦合。SUPG解（图12b）捕捉到了整体的 sigmoid 结构和前沿位置，但过渡区域出现了明显弥散：原本尖锐的解析前沿被跨越数个网格宽度的平缓斜坡所取代。SUPG-YZβ 近似（图12c）生成了相近的剖面；然而，前沿附近的上平台处可见轻微过冲，表明对于这个非线性问题，激波捕捉算子未能完全抑制类Gibbs振荡。混合PINN校正（图12d）恢复了更陡峭的前沿，与解析解更加吻合，其过渡区域也明显比任一有限元近似更加尖锐。这种改进在对角线方向上最为显著，因为 Burgers 方程的自锐化机制会将解梯度集中到越来越窄的带状区域中，有限元网格无法充分解析这一变化，而神经网络在连续时空域中运行，能够以更高的保真度近似该解。

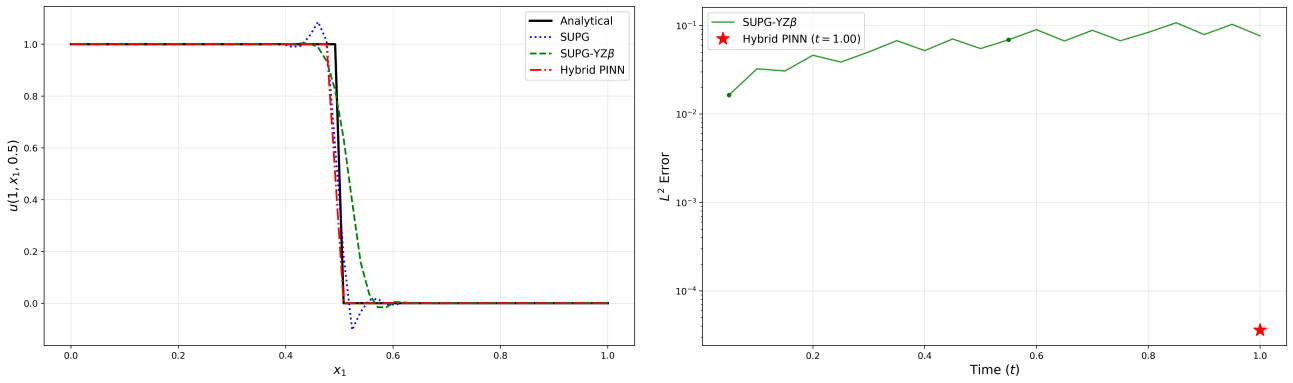
图13中的等值线图提供了俯视图，更清楚地展现了内层结构以及四种解之间的差异。解析解（图13a）显示，沿着 $x_1 + x_2 = 1$ 存在一条清晰的对角前沿，$u \approx 1$ （黄色，左上方）向 $u \approx 0$ （紫色，右下方）的过渡发生在 $\mathcal{O}(1/\text{Re}) = \mathcal{O}(10^{-4})$ 的距离范围内。等值线几乎笔直且彼此平行，反映出精确sigmoid解沿特征方向的平面对称性。SUPG等值线（图13b）



(a) (b) (c) (d)

图13：例*4*（二维Burgers方程）在 $t_{\mathrm{f}}=1$ 且 $\mathrm{Re}=10^{4}$ 时解的等值线图：(a) 解析解；(b) SUPG有限元近似；(c) SUPG-YZβ 有限元近似；(d) 混合PINN校正。所有面板共用同一色标。

再现了正确的前沿位置和方向，但过渡带明显更宽：绿色的中间等值线跨越了数个网格宽度，表明数值扩散使层的展宽超过了解析解中的范围。SUPG-YZβ 解（图13c）的前沿展宽程度相近；不过，左上角附近的等值线间距呈现出细微的不对称，这表明各向同性激波捕捉扩散与非线性自对流相互作用，形成了略微不均匀的层剖面。混合PINN等值线（图13d）恢复了更清晰的前沿，过渡带也更窄，与解析参考解更为吻合；同时，等值线保持了精确解的平行、平面结构。绿色中间区域的宽度尤其能体现这一改进：与两种有限元近似相比，PINN结果中的该区域明显更窄。

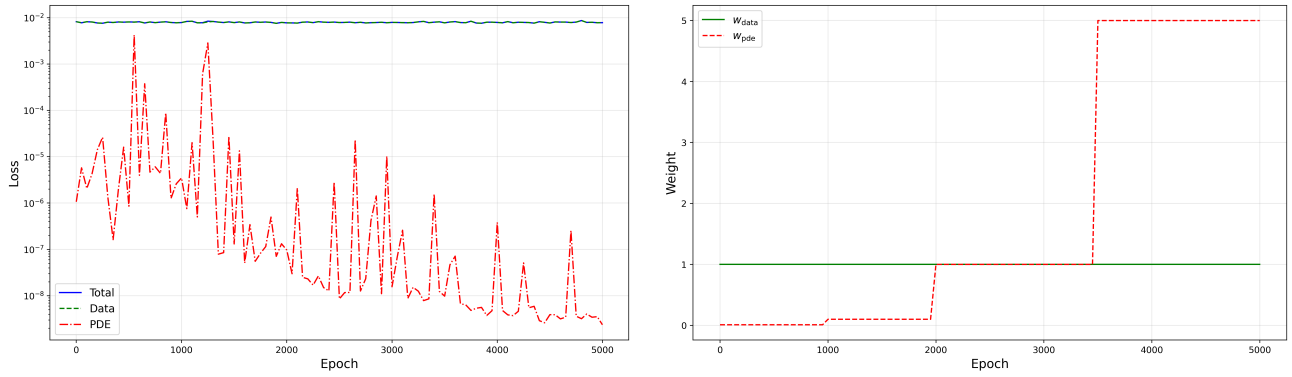


(a) (b)

图14：例*4*（二维Burgers方程）在 $\mathrm{Re}=10^{4}$ 时的定量比较：(a) $t_{\mathrm{f}}=1$时沿 $x_2=0.5$ 的横截面剖面，对比解析解、SUPG、SUPG-YZβ 和混合PINN；(b) SUPG-YZβ在时间区间 $[0,1]$ 内的 L^2 误差演化，星形标记表示终止时刻的混合 PINN 误差。

图14a所示的在 $t_{\mathrm{f}}=1$时沿 $x_2=0.5$ 的横截面剖面，提供了穿过 $x_1+x_2=1$处对角内层的一维切片，即前沿在 $x_1=0.5$处与切片相交。解析解呈现出从 $u=1$ 到 $u=0$ 的陡峭S形过渡，过渡距离为 $\mathcal{O}(1/\mathrm{Re})=\mathcal{O}(10^{-4})$。SUPG解捕捉到了前沿位置，但在 $x_1\approx 0.55$附近出现了明显的零下欠冲，这是.....的典型特征

对于非线性对流而言，这是横风稳定化不足导致的类吉布斯伪影。SUPG-YZβ 近似消除了欠冲，但引入了显著弥散：过渡在 x_1方向上延展约0.1个单位，使层宽度比解析宽度扩大了约两个数量级。混合 PINN 校正恢复了明显更陡的前沿，与解析剖面高度吻合，既没有SUPG的欠冲，也没有SUPG-YZβ的过度弥散。图14b中 L^2 误差的时间演化显示，SUPG-YZβ 解的误差呈增长趋势，从早期的 $\mathcal{O}$（10^{-3}）增至$\mathcal{O}$（10^{-1}）（ $t_{\mathrm{f}} = 1$）。误差单调增长反映了尖锐前沿在区域内传播时数值扩散的逐步累积：在每个时间步，人工激波捕捉耗散都会进一步展宽该层，这种扩散误差并在整个模拟区间内不断累积。终止时刻的混合 PINN 误差（红色星号）约为 $\mathcal{O}$（10^{-4}），比SUPG-YZβ 误差低近三个数量级，展现了本研究所有示例中最显著的改进。这一显著提升归因于基于提升函数的架构与四阶段训练策略的共同作用：前者将网络锚定在正确的边界行为上，后者则持续保持数据保真度（全程为$w_{\mathrm{data}} = 1.0$），从而避免非线性偏微分方程残差驱使解趋向伪分支。



(a) (b)

图15：例*4*（二维Burgers方程）在 $\mathrm{Re} = 10^4$ 下的训练诊断：(a) 在对数尺度下，5000 个训练轮次中各损失分量（总损失、数据损失和偏微分方程损失）的变化；(b) 控制四阶段训练策略的自适应权重调度 w_{data} 和w_{pde} 。边界条件权重未显示（$w_{\mathrm{bc}} = 0$），因为通过提升函数精确施加了狄利克雷条件。

示例 *4* 的训练诊断结果如图15所示，与此前所有示例相比，呈现出截然不同的模式。如图15a所示，训练过程中总损失与数据损失几乎无法区分，两者均保持在约 $\mathcal{O}(10^{-2})$的平台水平。这是基于提升函数的架构直接导致的结果：由于提升函数 u_{lift} 已精确满足边界条件，网络修正 $d(\mathbf{x})\mathcal{N}_\theta$ 只需学习相对较小的扰动，而数据损失——用于衡量网络输出与 SUPG-YZβ 参考解之间的差异——则稳定在有限元离散误差的固有水平。偏微分方程损失的量级则截然不同，其范围从 $\mathcal{O}(10^{-8})$ 到 $\mathcal{O}(10^{-3})$，并伴有间歇性尖峰；随着训练进行，尖峰的频率和幅度均有所下降。这一极低的

偏微分方程残差——比示例 *1*–例*3*中观察到的数值低几个数量级——反映了修正场的平滑性：网络只需对已经足够准确的提升函数进行小幅调整，因此产生的偏微分方程残差也相应较小。当随机采样的配点与 $x_1 + x_2 = t$ 处的薄内层相交时，就会出现尖峰；此处解梯度的量级为 $\mathcal{O}(\mathrm{Re})$。图15b中的权重调度展示了此示例特有的四阶段策略。与示例 *1*–例*3*中采用的递减调度不同，数据权重在所有阶段始终固定为 $w_{\mathrm{data}} = 1.0$。这一设计选择是考虑到 Burgers 方程弱解的不唯一性，确保网络始终锚定在物理上正确的解分支上。偏微分方程权重逐步增加，从第一阶段的 $w_{\mathrm{pde}} = 0.01$ 增至阶段 II 和 III 的 $w_{\mathrm{pde}} = 0.1$ 和 1.0，最终在阶段 IV（训练轮次 3500–4999）达到 $w_{\mathrm{pde}} = 5.0$。尽管最后阶段的偏微分方程权重较高，总损失仍由数据分量主导，因为偏微分方程残差本身较小；这证实物理信息训练能够优化解，同时不会与数据保真目标相竞争。

5.5. 例*5*。

考虑以下 二维时变对流–扩散方程，其具有一个 L 形内部层 [69] ：

$$\begin{cases} \dfrac{\partial u}{\partial t} + \mathbf{b} \cdot \nabla u - \nabla \cdot (\varepsilon \nabla u) = 0, & (x_1, x_2) \in \Omega, \quad t \in (0, t_{\mathrm{f}}], \\ u(0, x_1, x_2) = u_0(x_1, x_2), & (x_1, x_2) \in \Omega, \\ u = g(t, x_1, x_2), & (x_1, x_2) \in \Gamma_D, \quad t \in [0, t_{\mathrm{f}}], \end{cases} \tag{60}$$

其中 $\Omega = (0,\ 1)^2$ 为单位正方形， $t_{\mathrm{f}} = 0.25$，扩散率为 $\varepsilon = 10^{-8}$。对流速度为

$$\mathbf{b} = |\mathbf{b}|(\cos\theta, \sin\theta)^{\top} = \left(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2} \right)^{\top}, \tag{61}$$

其大小为 $\|\mathbf{b}\| = 1$ ，相对于网格的对流角为 $\theta = 45^{\circ}$。初始条件由一个 L 形块组成：

$$u_0(x_1, x_2) = \begin{cases} 1, & \text{if } (x_1, x_2) \in \Omega_L, \\ 0, & \text{otherwise}, \end{cases} \tag{62}$$

其中 L 形区域 Ω_{L} 定义为

$$\Omega_L = \left\{ (x_1, x_2) : 0 \le x_1 \le \tfrac{1}{2}L,\ 0 \le x_2 \le \tfrac{1}{4}L \right\} \cup \left\{ (x_1, x_2) : 0 \le x_1 \le \tfrac{1}{4}L,\ 0 \le x_2 \le \tfrac{1}{2}L \right\}, \tag{63}$$

其中 $L = 1$。狄利克雷边界条件为

$$g(t, x_1, x_2) = \begin{cases} 1, & \text{if } x_2 = 0 \text{ and } 0 \le x_1 \le \tfrac{1}{2}L, \\ 1, & \text{if } x_1 = 0 \text{ and } 0 \le x_2 \le \tfrac{1}{2}L, \\ 0, & \text{otherwise on } \partial\Omega. \end{cases} \tag{64}$$

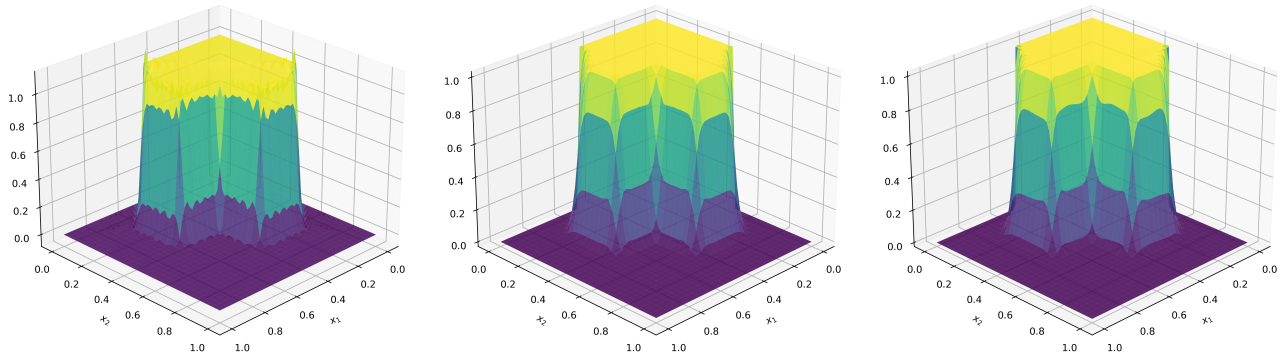
对于有限元设置，我们设定 $n_{\mathrm{el}} = 8192$ （64×64 网格的结构化三角剖分）、 $\Delta t = 0.001$ （$N_t = 250$）、$t_{\mathrm{f}} = 0.25$ 和 Y$= 0.25$。由于该问题没有解析解，因此通过对解剖面的定性比较以及不同方法间 L^2 范数的定量比较来进行评估。物理信息神经网络架构采用 $n_h = 128$、 $n_r = 8$、 $n_{\mathrm{F}} = 24$ 和 $\sigma = 4.0$。训练使用靠近 $t_{\mathrm{f}} = 0.25$ 的最后 $K_s = 10$ 个时间快照，批量大小为 2048，初始学习率为 $\alpha = 3 \times 10^{-4}$，梯度裁剪阈值为 $g_{\max} = 1.0$。每个 epoch 在 $N_{\mathrm{pde}} = 4096$ 个随机采样的内部配点处评估偏微分方程残差——这是所有示例中最大的取值，旨在充分解析沿流出边界形成的两个内层之间复杂的相互作用。

与前述示例不同，本问题的训练策略采用渐进式四阶段调度，其中三个权重同时演化。这一设计源于缺乏解析解：SUPG-YZβ 参考数据是唯一可用的真实值，权重调度必须仔细平衡数据保真度与物理约束，以避免将网络导向满足偏微分方程残差、却偏离有限元参考解且无法独立验证的解。数据权重从第一阶段的 $w_{\mathrm{data}} = 0.1$ 增加到阶段 IV 的 $w_{\mathrm{data}} = 0.8$ ，反映出随着物理约束加强，锚定参考解的重要性日益增加。偏微分方程权重从 $w_{\mathrm{pde}} = 0.5$增加到 $w_{\mathrm{pde}} = 10.0$，逐步强化物理约束。边界条件权重从 $w_{\mathrm{bc}} = 0.1$ 增加到 $w_{\mathrm{bc}} = 1.0$，以确保无法通过简单提升函数纳入的非齐次分段狄利克雷数据在偏微分方程损失占主导时仍得到良好满足。阶段边界和自适应权重见表5。

表5：求解例5所用的阶段边界和自适应权重。

阶段	训练轮次	w_{data}	w_{pde}	w_{bc}
I	0–999	0.1	0.5	0.1
II	1000–1999	0.35	0.8	0.3
III	2000–3499	0.5	5.0	0.8
IV	3500–4999	0.8	10.0	1.0

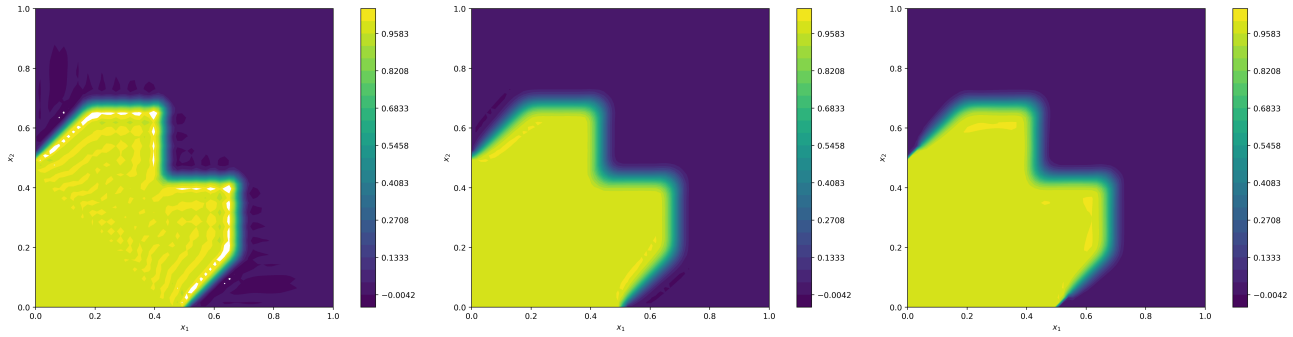
图16展示了示例 5在终止时刻 $t_{\mathrm{f}} = 0.25$ 的三维解剖面。由于该问题没有解析解，因此评估依赖于对三种数值近似的定性比较。SUPG解（图16a）捕捉到了整体的 L 形块结构及其向右上角的对角对流；然而，前沿的前沿边缘可见明显的振荡伪影，尤其是在 L 形水平臂与垂直臂交汇的拐角附近。这些横风振荡是基于流线的稳定化在存在陡峭横向梯度时的典型缺陷，它们污染了本应始终保持为 1 的平坦平台区域，也影响了周围的零值外部区域。加入 YZβ 激波捕捉（图16b）后，这些振荡特征得到显著抑制，剖面也更加平滑、单调。前沿边缘的分辨效果更好，平台区域更加平坦，外部区域的伪振荡也



(a) (b) (c)

图16：例*5* 在 $t_{\mathrm{f}} = 0.25$ 且 $\varepsilon = 10^{-8}$时解的三维曲面图：(a) SUPG有限元近似；(b) SUPG-YZβ 有限元近似；(c) 混合PINN。

基本消除了。不过，尖锐过渡处仍可观察到轻微弥散，这与激波捕捉算子引入的各向同性人工扩散相符。混合PINN校正（图16c）生成的表面轮廓在整体结构上与 SUPG-YZβ 近似十分相似，同时还有两项显著改进：解保持在物理允许范围 $[0, 1]$ 内，没有出现有限元近似中的过冲或欠冲伪影；并且，沿对流输运的 L 形块前沿边缘，前沿过渡看起来略微更加尖锐。其有界性改善归因于两方面的共同作用：基于表现良好的 SUPG-YZβ 参考解进行数据驱动训练，以及偏微分方程残差正则化；二者共同引导网络远离非物理极值。

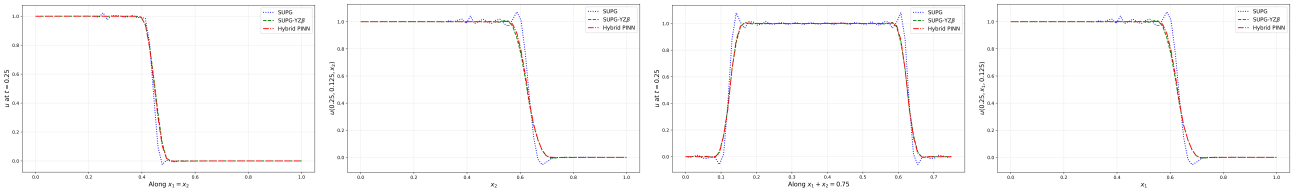


(a) (b) (c)

图17：例*5* 在 $t_{\mathrm{f}} = 0.25$且 $\varepsilon = 10^{-8}$时解的等值线图：(a) SUPG有限元近似；(b) SUPG-YZβ 有限元近似；(c) 混合PINN。

图17中的等值线图提供了俯视图，更清晰地展现了对流输运的L形前沿的空间结构，以及三种数值近似之间的差异。SUPG解（图17a）显示出明显的横风振荡，表现为沿前沿边缘呈现的一系列同心波纹状等值线畸变。这些伪影深入平台区域和零值外部区域，是在结构化网格上对流占优地输运不连续初始数据时横风扩散不足的典型表现

。这种振荡模式在L形的再入角附近尤为严重，因为水平前沿与垂直前沿相互作用，产生了彼此竞争的梯度方向，而基于流线的SUPG稳定化无法充分解析这些方向。SUPG-YZβ 解（图17b）完全消除了这些振荡特征，生成了平滑的等值线场，且前沿边界清晰。然而，与初始尖锐不连续相比，前沿过渡明显变宽，这反映了激波捕捉算子引入的各向同性人工扩散。前沿后方的平台区域保持了接近1的近乎均匀值，证实了YZβ 机制能够有效保留陡峭梯度以外区域的解结构。混合PINN等值线（图17c）与SUPG-YZβ 模式十分接近，但前沿边缘处略有差异：过渡区似乎稍窄，表明偏微分方程残差约束部分抵消了激波捕捉算子引入的数值扩散。此外，混合PINN解在整个区域内始终保持在 $[0, 1]$ 范围内，避免了有限元近似色标中可见的轻微欠冲 -0.0042——这是因为数据保真度损失与偏微分方程残差损失在训练过程中共同惩罚了非物理解值。

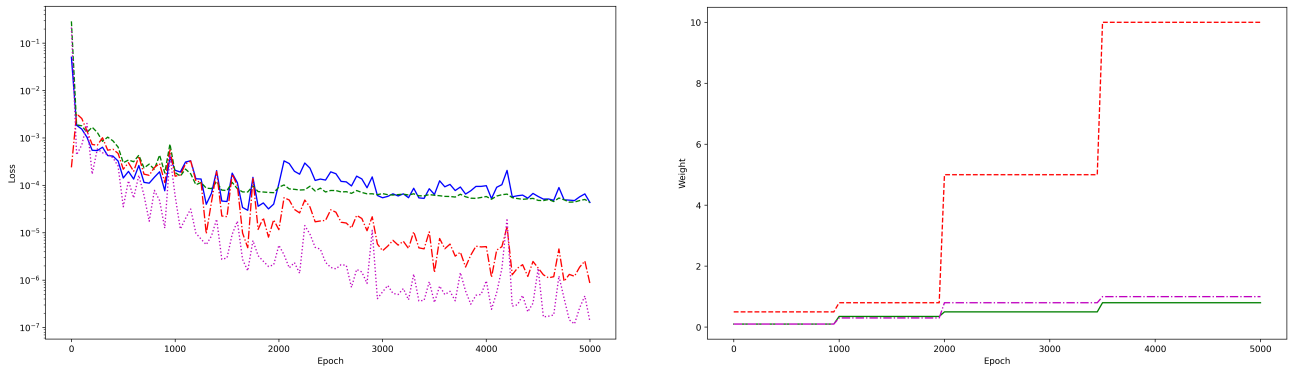


(a) (b) (c) (d)

图18： $t_{\mathrm{f}} = 0.25$ 处 例*5* 的横截面剖面（ $\varepsilon = 10^{-8}$），比较 SUPG、SUPG-YZβ 和混合PINN：(a) 沿对角线 $x_1 = x_2$；(b) 沿 $x_1 = 0.125$ （穿过垂直臂）；(c) 沿 $x_1 + x_2 = 0.75$ （垂直于对流方向）；(d) 沿 $x_2 = 0.125$ （穿过水平臂）。

图18展示了在终止时刻 $t_{\mathrm{f}} = 0.25$ 沿四条具有代表性的直线提取的一维横截面剖面，对三种方法进行了定量比较。沿$x_1 = x_2$ 的对角线切片（图18a）沿对流方向穿过对流输运的L形前沿 $\mathbf{b} = (\sqrt{2}/2,\ \sqrt{2}/2)^{\top}$。三种方法都捕捉到了 $u \approx 1$ 到 $u \approx 0$ 在 $s \approx 0.45$附近的过渡；然而，SUPG解在前沿附近出现了明显的振荡过冲，而 SUPG-YZβ 和混合 PINN 剖面则几乎单调。PINN 剖面与 SUPG-YZβ 解高度吻合，且在前沿处的过渡略微更陡。位于 $x_1 = 0.125$ 的垂直切片（图18b）穿过L形的垂直臂。平台区在 $u \approx 1$ 延伸至 $x_2 \approx 0.6$，这与初始块从 $x_2 \in [0, 0.5]$对流输运而来的情况一致。SUPG解在前沿后沿附近出现了明显的零下欠冲，而 SUPG-YZβ 和混合 PINN解均保持非负剖面。与 SUPG-YZβ 近似相比，混合PINN在前沿处的下降略为陡峭。沿 $x_1 + x_2 = 0.75$ 的切片（图18c）垂直于对流方向，是最能揭示差异的比较。该横截面横向穿过前沿，此处横风效应最为显著。SUPG解沿整个剖面呈现显著的振荡伪影，反映了众所周知的横风

扩散不足。SUPG-YZβ 剖面平滑得多，但略有展宽。混合PINN沿袭了 SUPG-YZβ 剖面，同时在 $[0,1]$, 内保持有界，这与物理信息训练的正则化效应一致。位于 $x_2 = 0.125$ 的水平切片（图18d）穿过L形的水平臂。$x_1 \approx 0.65$ 附近的前沿表现相似：SUPG解出现振荡，SUPG-YZβ 解呈现平滑但扩散的过渡，而混合PINN则紧密贴合稳定化有限元法参考解。在全部四个横截面上，PINN 校正始终保留了 SUPG-YZβ 近似的整体解结构，同时在经验上满足最大值原理——这是两种有限元公式都无法通过构造保证的性质。尽管架构中并未强制施加这一性质，但它自然源于在有界参考数据上训练并在内部进行偏微分方程残差最小化的结合。



(a) (b)

图19：例*5* 在 $\varepsilon = 10^{-8}$条件下的训练诊断：(a) 各损失分量（总损失、数据、偏微分方程和边界条件）在5000个训练轮次中的演化（对数尺度）；(b) 支配四阶段训练策略的自适应权重调度 w_{data}、 w_{pde}和 w_{bc} 。

例5的训练诊断结果如图19所示。图19a中的损失演化揭示了四个训练阶段中各损失分量的不同表现。在第一阶段（训练轮次 0–999）中，偏微分方程权重为中等水平 $w_{\text{pde}} = 0.5$ ，数据权重较低，为 $w_{\text{data}} = 0.1$；此时所有损失分量都从 $\mathcal{O}(10^{-1})$ 迅速降至 $\mathcal{O}(10^{-4})$ ，这表明网络成功地从 SUPG-YZβ 参考快照中学习了全局解结构，同时开始满足偏微分方程约束。第一阶段结束时，数据损失稳定在 $\mathcal{O}(10^{-4})$ 附近，并在后续各阶段持续缓慢下降，这证实了逐步增大的数据权重（w_{data} ：$0.1 \to 0.8$）有效地将网络锚定于参考解。偏微分方程损失呈现间歇性尖峰，幅度跨越约两个数量级，从 $\mathcal{O}(10^{-6})$ 到 $\mathcal{O}(10^{-4})$ ，并在整个训练过程中持续出现。当随机采样的配点与尖锐的 L形锋面相交时，便会产生这些突增；在此处，解梯度陡峭，偏微分方程残差的幅值也相应增大。尽管表现出这种波动，偏微分方程损失总体上仍随训练推进而稳步下降。在所有阶段，边界条件损失始终保持在较低水平 $\mathcal{O}(10^{-5})$ –$\mathcal{O}(10^{-7})$ ，这反映出针对分段不连续狄利克雷数据显式施加边界条件的有效性。图19b中的权重调度展示了此问题所采用的四阶段策略。偏微分方程权重

经历了最显著的变化：从第一阶段的 $w_{\mathrm{pde}} = 0.5$ 增至第二阶段的 $w_{\mathrm{pde}} = 0.8$ ，再升至第三阶段的 $w_{\mathrm{pde}} = 5.0$ ，最终在阶段 IV（训练轮次 3500–4999）达到 $w_{\mathrm{pde}} = 10.0$ 。如此激进地提高最终的偏微分方程权重，是推动物理信息精化超越单纯数据插值所必需的。与此同时，数据权重从 $w_{\mathrm{data}} = 0.1$ 增至 $w_{\mathrm{data}} = 0.8$，确保随着偏微分方程约束的加强，网络仍锚定于 SUPG-YZβ 参考解；这一设计考虑到缺乏解析解可用于独立验证。边界条件权重从 $w_{\mathrm{bc}} = 0.1$ 逐步增至 $w_{\mathrm{bc}} = 1.0$，反映出随着训练过渡到物理主导精细化，维持非齐次分段狄利克雷数据的边界保真度日益重要。

6. 结论与未来工作

本文提出了一种混合计算框架，将稳定化有限元方法与物理信息神经网络相结合，用于数值求解瞬态对流主导型 CDR 偏微分方程。该方法采用结合 YZβ 激波捕捉算子的 SUPG 公式生成高保真参考解，随后将其用作配备残差块和随机傅里叶特征嵌入的深度神经网络的训练数据。PINN 校正并非针对整个时空区域进行训练，而是聚焦于终止时刻附近最后 K_s 个时间快照；在此过程中，网络吸收有限元数据，同时通过自动微分以强形式施加控制偏微分方程。

所提出方法的一个核心要素是选择性物理约束策略：该策略将偏微分方程残差损失限制在满足基于距离的排除准则的内部配置点上。这一设计避免了物理信息训练干扰稳定化有限元离散已能充分解析的尖锐边界层与内部层结构。多阶段自适应权重演化——从数据主导训练过渡到过渡阶段，再到物理主导训练——对于平衡数据保真度与物理一致性至关重要，也有助于避免训练初期的不稳定性；当物理项在建立合理的解近似之前就占据主导时，通常会出现这类不稳定性。

该框架在涵盖一维和二维情形的五个基准问题上得到了验证：带边界层的线性抛物型方程（示例 *1*）、隆起高度变化的时变 CDR 方程（例*2* ）、传播内部层问题（例*3*）、含移动激波前沿的黏性 Burgers 方程（例*4*），以及含不连续 L 形内部层的对流–扩散方程（例*5*）。在所有测试用例中，混合PINN校正始终优于独立的 SUPG-YZβ 有限元解，尤其是在存在陡峭梯度、移动前沿和局部振荡伪影的区域。结果表明，所提出的方法继承了经典稳定化公式的鲁棒性和稳定性，同时利用神经网络的表征灵活性提升了解的质量。

本研究自然引出了未来研究的若干方向。首先，扩展到三维问题是迈向实际工程应用的关键一步，包括复杂几何中的计算流体力学和传热。尽管网络架构和训练策略

在设计上与维数无关——时空输入 $\mathbf{z} = (t, \mathbf{x}) \in \mathbb{R}^{n_{\mathrm{sd}}+1}$ 以及傅里叶特征嵌入能够适应任意空间维数——但生成三维 SUPG-YZβ 参考数据的计算成本，以及偏微分方程残差评估所需增加的配点数量，都带来了额外挑战，值得进行系统性研究。

其次，将该方法应用于耦合偏微分方程组，例如不可压缩Navier–Stokes方程、磁流体动力学流动和多组分反应输运，是一项极具价值的拓展。在这类情形中，方程间耦合会引入额外的非线性，而稳定化参数和激波捕捉参数必须考虑各控制方程不同的特征尺度。备注5和6中讨论的对角缩放矩阵 $\mathbf{Y}$ 和残差向量 $\mathbf{Z}$ 为将 YZβ 机制推广到向量值问题提供了自然的起点；同时，PINN组件可扩展为多输出架构，从而同时近似所有解场。

第三，采用时空有限元格式（例如，参见 [80–84]）是一项自然且在方法论上意义重大的拓展。在当前框架中，空间与时间离散是分别处理的：空间域采用稳定化有限元进行离散，而时间则通过向后欧拉格式进行积分。相比之下，时空格式将时间维度与空间坐标置于同等地位，在整个时空区域 $\mathcal{Q} = (t_0,\ t_f] \times \Omega$ 上构造有限元网格，并以统一方式求解由此得到的变分问题。这种方法为所提出的混合框架带来若干潜在优势：(i) 能够同时在空间和时间上进行局部加密，这对于存在移动前沿和瞬态层结构的问题尤其有益；(ii) SUPG稳定化和 YZβ 激波捕捉可自然拓展至时空情形，因为稳定化参数 τ_{SUPG} 已通过式（9）同时纳入空间与时间尺度；(iii) PINN组件本身已基于时空输入坐标 $\mathbf{z} = (t,\ \mathbf{x})$ 运行，并通过自动微分计算时间导数，因此无需修改架构即可与时空离散兼容。研究时空 SUPG-YZβ 参考解与选择性物理约束策略之间的相互作用——特别是是否应将基于距离的排除准则推广为时空度量——是一个重要的开放问题。

第四，也可能是最重要的一点是，通过物理信息神经网络对稳定化参数和激波捕捉参数进行数据驱动优化，为消除目前支撑 τ_{SUPG} 和 ν_{SC} 选择的经验标定提供了一条很有前景的途径。与其依赖解析表达式，不如训练辅助神经网络，通过最小化同时惩罚偏微分方程残差和解振荡的损失泛函，来预测逐单元的最优稳定化参数。这种方法将把参数选择从启发式公式转向基于学习的范式，有望在无需人工干预的情况下，在更广泛的问题类别中实现更优性能。

其他探索方向包括：引入由基于PINN的误差指示器引导的自适应网格细化策略；探索更有原则的动态损失加权机制，例如梯度范数平衡或基于不确定性的缩放，以取代当前依赖经验调优的多阶段调度；以及研究迁移学习技术，利用针对一种问题配置训练的模型来

ac加速具有不同参数值或几何形状的相关场景的收敛 s.

作者声明

CRediT 作者贡献声明

SC:概念构思；正式分析；经费获取；调查研究；方法学；软件；验证；可视化；撰写初稿；审阅与编辑。**ÖU:** 概念构思；正式分析；调查研究；审阅与编辑。**SN:** 概念构思；正式分析；调查研究；审阅与编辑。

利益冲突

作者声明不存在任何利益冲突。

致谢

本研究的第一作者得到了土耳其科学技术研究委员会（TUBITAK）225M468号资助的支持。作者感谢TUBITAK的支持。

参考文献

[1] T. J. R. Hughes, A. N. Brooks, 一种无横风扩散的多维迎风格式, 载于: T. J. R. Hughes（编）, 对流主导流动的有限元方法, AMD第34卷, ASME, 纽约, 1979, 第19–35页.

[2] A. N. Brooks, T.J. R.Hughes, 针对对流主导流动的流线迎风/Petrov-Galerkin公式，特别侧重于不可压缩纳维-斯托克斯方程, Computer Methods in Applied Mechanics and Engineering 32 (1982) 199–259. doi:10.1016/0045-7825(82)90071-8.

[3] T. E. Tezduyar，T. J. R. Hughes，《针对一阶双曲系统的时间精确有限元技术的开发，特别关注可压缩欧拉方程》，NASA技术报告NASA-CR-204772，NASA，1982年。

[4] T. E. Tezduyar，T. J. R. Hughes，《对流主导流动的有限元公式，特别关注可压缩欧拉方程》，载于：AIAA第21届航空航天科学会议论文集，AIAA论文 83-0125，里诺，内华达，1983年。doi:10.2514/6.1983-125。

[5] T. J. R. Hughes，T. E. Tezduyar，《一阶双曲系统的有限元方法，特别关注可压缩欧拉方程》，Computer Methods in Applied Mechanics and Engineering 45（1984）217–284。doi:10.1016/ 0045-7825(84)90157-9。

[6] T. J. R. Hughes, L. P. Franca, G. M. Hulbert，《一种新的计算流体力学有限元形式：VIII. 对流-扩散方程的 Galerkin/最小二乘法》, Computer Methods in Applied Mechanics and Engineering 73 (1989) 173–189. doi:10.1016/0045-7825(89)90111-4.

[7] T. J. R. Hughes，多尺度现象：格林函数、狄利克雷-诺伊曼公式、亚格子尺度模型、气泡与稳定化方法的起源，Computer Methods in Applied Mechanics and Engineering 127 (1995) 387–401. doi:10.1016/0045-7825(95)00844-9.

[8] T. E. Tezduyar、S. Mittal、R. Shih，采用四边形速度-压力单元进行时间精确不可压缩流动计算，Computer Methods in Applied Mechanics and Engineering 87（1991）363–384。doi:10.1016/0045-7825(91)90014-W。

[9] L. P. Franca, T. J. R. Hughes, 对流-扩散模型的应用，《Computer Methods in Applied Mechanics and Engineering》95 (1992年) 253–276。doi:10.1016/0045-7825(92)90143-8.[10] T. E. Tezduyar, Y. Osawa，基于单元矩阵和向量计算的有限元稳定化参数，《Computer Methods in Applied Mechanics and Engineering》190 (2000) 411–430。doi:10.1016/S0045-7825(00)00211-5.[11] L. P. Franca, F. Valentin，关于对流–反应–扩散方程的一种改进型非常规稳定有限元方法，《Computer Methods in Applied Mechanics and Engineering》190 (2000) 1785–1800。doi:10.1016/s0045-7825(00)00190-0.[12] T. J. R. Hughes, G. Scovazzi, L. P. Franca，多尺度方法与稳定化方法，载于：E. Stein, R. de Borst, T. J. R. Hughes（编），计算力学百科全书，2004年。doi:10.1002/9781119176817.ecm2051.[13] E. Burman, P. Hansbo，Galerkin逼近对流–扩散–反应问题的边稳定化，《Computer Methods in Applied Mechanics and Engineering》193 (2004) 1437–1453。doi:10.1016/j.cma.2003.12.032.[14] V. John, P. Knobloch，关于对流-扩散方程的层间虚假振荡衰减（SOLD）方法：第一部分——综述，《Computer Methods in Applied Mechanics and Engineering》196 (2007年) 2197–2215。doi:10.1016/j.cma.2006.11.013.[15] L. P. Franca, G. Hauke, A. Masud，重新审视对流–扩散方程的稳定化有限元方法，《Computer Methods in Applied Mechanics and Engineering》195 (2006年) 1560–1572。doi:10.1016/j.cma.2005.05.028.[16] V. John, P. Knobloch，关于对流-扩散方程的层间虚假振荡衰减（SOLD）方法：第二部分——针对 P_1 和 Q_1 有限元的分析，《Computer Methods in Applied Mechanics and Engineering》197 (2008) 1997–2014。doi:10.1016/j.cma.2007.12.019.[17] V. John, E. Schmeyer，具有小扩散项的时变对流–扩散–反应方程有限元方法，《Computer Methods in Applied Mechanics and Engineering》198 (2008) 475–494。doi:10.1016/j.cma.2008.08.016.[18] R. Codina，使用正交子尺度分析Oseen方程的稳定有限元逼近，《Applied Numerical Mathematics》58 (2008) 264–283。doi:10.1016/j.apnum.2006.11.011.[19] C. Jiang, Z. Zhang, X. Han, G. Liu, G. Gao, T. Lin，一种用于不可压缩层流黏性流动的准隐式特征线罚有限元法，《International Journal for Numerical Methods in Engineering》114 (2018) 147–171。doi:10.1002/nme.5738.[20] M. Liu, G. Gao, H. Zhu, C. Jiang，一种采用隐式SUPG/SPGP/分步法稳定化的基于单元平滑有限元法，用于不可压缩流动，《Engineering Analysis with Boundary Elements》124 (2021年) 194–210。doi:10.1016/j. enganabound.2020.12.018.[21] C. Jiang, C. Hong, T. Wang, G. Zhou，一种用于不可压缩流动传热问题的N侧基于单元平滑有限元法，《Engineering Analysis with Boundary Elements》146 (2023年) 749–766。doi:10.1016/j.enganabound.2022.11.025.[22] M. Liu, C. Jiang, B. C. Khoo, H. Zhu, G. Gao，使用可实现k-ε 模型和混合网格分析湍流的基于单元平滑有限元模型，《Journal of Computational Physics》501 (2024年) 112783。doi:10.1016/j.jcp.2024.112783.[23] S. Cengizci, H. F. Öztop, S. Natesan，一种用于模拟触基性驱动癌症侵袭的间断捕捉SUPG有限元框架，《Mathematics and Computers in Simulation》241 (2026) 271–292。doi:10.1016/j.matcom.2025.10.011.[24] D. Ray, J. S. Hesthaven，将人工神经网络用作问题单元指示器，《Journal of Computational Physics》367 (2018年) 166–191。doi:10.1016/j.jcp.2018.04.029.[25] D. Ray, J. S. Hesthaven，利用神经网络检测二维非结构网格上的问题单元，《Journal of Computational Physics》397 (2019年) 108845。doi:10.1016/j.jcp.2019.07.043。

[26] M. H. Veiga, R. Abgrall，利用多层感知机神经网络为守恒律构建通用稳定化方法：一维标量方程与方程组，载于：欧洲计算力学会议暨第七届欧洲计算流体动力学会议（ECCM-ECFD 2018），英国格拉斯哥，2018年，第2525–2550页。2018年6月11–15日。[27] N. Discacciati, J. S. Hesthaven, D. Ray，利用神经网络调节的人工黏性控制高阶间断伽辽金格式中的振荡，《Journal of Computational Physics》409 (2020年) 109304。doi:10.1016/j.jcp.2020.109304.[28] L. Schwander, D. Ray, J. S. Hesthaven，利用神经网络控制的局部人工黏性抑制谱方法中的振荡，《Journal of Computational Physics》431 (2021年) 110144。doi:10.1016/j.jcp.2021.110144.[29] 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，《Journal of Computational Physics》378 (2019年) 686–707。doi:10.1016/j.jcp.2018.10.045.[30] G. E. Karniadakis, I. G. Kevrekidis, L. Lu, P. Perdikaris, S. Wang, L. Yang，物理信息机器学习，《自然综述：物理》3 (2021年) 422–440。doi:10.1038/s42254-021-00314-5.[31] S. K. Mitusch, S. W. Funke, M. Kuchta，混合有限元-神经网络模型：结合人工神经网络与有限元法，《Journal of Computational Physics》446 (2021年) 110651。doi:10.1016/j.jcp.2021.110651.[32] A. Logg, K.-A. Mardal, G. Wells，利用有限元法自动求解微分方程：FEniCS书，第84卷，计算科学与工程讲义，施普林格出版社（柏林·海德堡），2012年。doi:10.1007/978-3-642-23099-8.[33] M. Alnæs, J. Blechta, J. Hake, A. Johansson, B. Kehlet, A. Logg, C. Richardson, J. Ring, M. E. Rognes, G. N. Wells，FEniCS项目1.5版，《数值软件档案》3 (2015年) 9–23。doi:10.11588/ans.2015.100.20553.[34] S. Mitusch, S. W. Funke, J. S. Dokken，dolfin-adjoint 2018.1：适用于FEniCS和Firedrake的自动伴随算子，《开放源代码软件杂志》4 (2019年) 1292。doi:10.21105/joss.01292.[35] S. Huang, W. Feng, C. Tang, Z. He, C. Yu, J. Lv，偏微分方程与深度神经网络的结合：综述，《IEEE神经网络与学习系统汇刊》36 (2025年) 13649–13669。doi:10.1109/tnnls.2025.3545967.[36] X. Jin, S. Cai, H. Li, G. E. Karniadakis，NSFnets（纳维-斯托克斯流网络）：用于不可压缩Navier-Stokes方程的物理信息神经网络，《Journal of Computational Physics》426 (2021年) 109951。doi:10.1016/j.jcp.2020.109951.[37] J. Yu, J. S. Hesthaven，一种用于间断伽辽金方法的数据驱动激波捕捉方法，《计算机与流体》245 (2022年) 105592。doi:10.1016/j.compfluid.2022.105592.[38] S. Yadav, S. Ganesan, SPDE-ConvNet：利用深度学习求解奇异摄动偏微分方程，载于：第八届欧洲应用科学与工程计算方法大会（ECCOMAS Congress 2022），挪威奥斯陆，2022年。2022年6月5–9日。[39] T.-H. Huang，针对对流主导问题的稳定化且变分一致的积分无网格公式，《Computer Methods in Applied Mechanics and Engineering》403 (2023年) 115698。doi:10.1016/j.cma.2022.115698.[40] T. Le-Duc, H. Nguyen-Xuan, J. Lee，一种用于结构力学参数化模拟的有限元信息神经网络，《分析与设计中的有限元》217 (2023年) 103904。doi:10.1016/j.finel.2022.103904.[41] S. Yadav, S. Ganesan，人工神经网络增强稳定有限元法，《Journal of Computational Physics》499 (2024年) 112702。doi:10.1016/j.jcp.2023.112702.[42] R. Khodayi-Mehr, M. M. Zavlanos，VarNet：用于求解偏微分方程的变分神经网络，2019年。doi:10.48550/arxiv.1912.07443。

[43] K. A. Luong、T. Le-Duc、J. Lee，通过统一物理信息神经网络自动施加边值问题的边界条件，Engineering with Computers 40（2024年）1717–1739。doi:10.1007/s00366-023-01871-2。[44] T. G. Grossmann、U. J. Komorowska、J. Latz、C.-B. Schönlieb，物理信息神经网络能否胜过有限元法?，IMA Journal of Applied Mathematics 89（2024年）143–174。doi:10.1093/imamat/hxae011。[45] Y. Yamazaki、A. Harandi、M. Muramatsu、A. Viardin、M. Apel、T. Brepols、S. Reese、S. Rezaei，面向任意区域上时空偏微分方程的基于有限元的物理信息算子学习框架，Engineering with Computers 41（2024年）1–29。doi:10.1007/s00366-024-02033-8。[46] F. Buzaev、J. Gao、I. Chuprov、E. Kazakov，物理信息神经网络的混合加速技术：比较分析，Machine Learning 113（2023年）3675–3692。doi:10.1007/s10994-023-06442-6。[47] W. Yufeng、X. Cong、Y. Min、Z. Jin，减少对困难区域的强调：针对奇异摄动对流-扩散-反应问题的物理信息神经网络课程学习，East Asian Journal on Applied Mathematics 14（2024年）104–123。doi:10.4208/eajam. 2023-062.170523。[48] T. Y. Chang、G. M. Gie、Y. Hong、C. Y. Jung，Burgers 方程的奇异层物理信息神经网络方法，arXiv预印本 arXiv:2410.09723（2024年）。[49] T.-Y. Chang、G.-M. Gie、Y. Hong、C.-Y. Jung，平面平行流的奇异层物理信息神经网络方法，Computers & Mathematics with Applications 166（2024年）91–105。doi:10.1016/j.camwa.2024.04.025。[50] D. Frerichs-Mihov、L. Henning、V. John，关于稳态对流主导对流-扩散问题的物理信息神经网络损失泛函，Communications on Applied Mathematics and Computation（2024年）。doi:10.1007/ s42967-024-00433-7。[51] T.-Y. Hsieh、T.-H. Huang，一种用于模拟对流主导流动的、带弱施加边界条件的多尺度稳定化物理信息神经网络迁移学习方法，Engineering with Computers 40（2024年）3353–3387。doi:10. 1007/s00366-024-01981-5。[52] M. Kütük、H. Yücel，用于Allen–Cahn方程的能量耗散保持型物理信息神经网络，Journal of Computational Science 87（2025年）102577。doi:10.1016/j.jocs.2025.102577。[53] A. Raina、S. Badireddi、S. Natesan，应用物理信息神经网络求解流体动力学中出现的边界层问题，Mathematical Foundations of Computing 0（2025年）1–20。doi:10.3934/mfc.2025024。[54] S. Cengizci、O. Uğur、S. Natesan，一种用于计算稳态对流主导流动的、带激波捕捉的物理信息神经网络增强SUPG稳定化混合有限元框架，Advances in Engineering Software 216（2026年）104135。doi:10.1016/j.advengsoft.2026.104135。[55] T. E. Tezduyar，移动边界与界面的计算及稳定化参数，International Journal for Numerical Methods in Fluids 43（2003年）555–575。doi:10.1002/fld.505。[56] T. E. Tezduyar、M. Senga，可压缩流SUPG 公式中的稳定化与激波捕捉参数，Computer Methods in Applied Mechanics and Engineering 195（2006年）1621–1632。doi:10.1016/j.cma.2005.05.032。[57] S. Cengizci、O. Uğur、S. Natesan，化学反应对流主导非线性模型的SUPG-YZβ 计算，International Journal of Computer Mathematics 100（2022年）283–303。doi:10.1080/00207160.2022.2114794。[58] S. Cengizci、O. Uğur、S. Natesan，采用激波捕捉的对流主导三维椭圆偏微分方程的基于SUPG的稳定化有限元计算，Journal of Computational and Applied Mathematics 451（2024年）116022。doi:10.1016/j.cam.2024.116022。[59] F. Shakib，可压缩Euler方程与Navier-Stokes方程的有限元分析，博士论文，机械工程系，斯坦福大学，1988年。

[60] T. E. Tezduyar，用于不可压缩流计算的稳定化有限元公式，Advances in Applied Mechanics 28（1992年）1–44。doi:10.1016/S0065-2156(08)70153-4。[61] S. Cengizci，一种用于模拟由反应-对流-扩散方程耦合系统支配的自然现象的增强型SUPG稳定化有限元公式，Mathematical Modelling and Numerical Simulation with Applications 3（2023年）297–317。doi:10.53391/mmnsa.1387125。[62] S. Cengizci、O. Uğur，一种带间断捕捉的稳定化有限元公式，用于求解高雷诺数下的Burgers型方程，Applied Mathematics and Computation 442（2023年）127705。doi:10.1016/j.amc.2022.127705。[63] S. Cengizci、O. Uğur，用于计算浅水方程的、以YZβ 激波捕捉增强的SUPG 公式，ZAMM——应用数学与力学杂志 / Zeitschrift für Angewandte Mathematik und Mechanik 103（2023年）。doi:10.1002/zamm.202200232。[64] S. Cengizci、O. Uğur，采用带激波捕捉的稳定化有限元公式模拟高哈特曼数磁流体力学管道流的计算研究，Journal of Computational Science 81（2024年）102381。doi:10.1016/j.jocs.2024.102381。[65] S. Cengizci、O. Uğur，Heston随机波动率模型下欧式与美式期权定价的计算研究：SUPG-YZβ 形式的应用，Computational Economics（2024年）。doi:10.1007/s10614-024-10704-3。[66] T. E. Tezduyar，具有移动边界与界面的流体动力学有限元方法，载于：E. Stein、R. D. Borst、T. J. R. Hughes（编），计算力学百科全书，第3卷：流体，Wiley，2004年。doi:10.1002/0470091355. ecm069。[67] T. E. Tezduyar、M. Senga，采用YZβ 激波捕捉的无黏超声速流SUPG有限元计算，《计算机与流体》36（2007年）147–159。doi:10.1016/j.compfluid.2005.07.009。[68] T. E. Tezduyar、M. Senga、D. Vicker，采用SUPG 公式与YZβ 激波捕捉计算圆柱与球体周围的无黏超声速流，Computational Mechanics 38（2006年）469–481。doi:10.1007/s00466-005-0025-6。[69] Y. Bazilevs、V. M. Calo、T. E. Tezduyar、T. J. R. Hughes， $YZ\beta$ 对流主导过程的间断捕捉及其在动脉药物输送中的应用，International Journal for Numerical Methods in Fluids 54（2007年）593–608。doi:10.1002/ fld.1484。[70] I. Loshchilov、F. Hutter，解耦权重衰减正则化，2017年。doi:10.48550/arXiv.1711.05101。arXiv:1711.05101。[71] D. P. Kingma、J. L. Ba，Adam：一种随机优化方法，载于：ICLR 2015：国际学习表征会议，ArXiv，纽约州伊萨卡，2015年，第13页。[72] M. Tancik、P. P. Srinivasan、B. Mildenhall、S. Fridovich-Keil、N. Raghavan、U. Singhal、R. Ramamoorthi、J. T. Barron、R. Ng，傅里叶特征使网络能够学习低维域中的高频函数，载于：神经信息处理系统进展（NeurIPS 2020），加拿大温哥华，2020年。[73] H. P. Langtangen、A. Logg，用Python求解偏微分方程：FEniCS教程 I，*Simula*计算*Springer*简报第3卷，Springer International Publishing，2016年。[74] I. A. Baratta、J. P. Dean、J. S. Dokken、M. Habera、J. S. Hale、C. N. Richardson、M. E. Rognes、M. W. Scroggs、N. Sime、G. N. Wells，DOLFINx：下一代FEniCS问题求解环境，2023年。doi:10.5281/zenodo.10447666。[75] A. Paszke、S. Gross、F. Massa、A. Lerer、J. Bradbury、G. Chanan、T. Killeen、Z. Lin、N. Gimelshein、L. Antiga、A. Desmaison、A. Kopf、E. Yang、Z. DeVito、M. Raison、A. Tejani、S. Chilamkurthy、B. Steiner、L. Fang、J. Bai、S. Chintala，PyTorch：一种命令式风格的高性能深度学习库，载于：第32届神经信息处理系统会议论文集，第32卷，Curran Associates, Inc.，加拿大温哥华，2019年。

[76] S. Balay, S. Abhyankar, M. F. Adams, S. Benson, J. Brown, P. Brune, K. Buschelman, E. Constantinescu, L. Dalcin, A. Dener, V. Eijkhout, J. Faibussowitsch, W. D. Gropp, V. Hapla, T. Isaac, P. Jolivet, D. Karpeev, D. Kaushik, M. G. Knepley, F. Kong, S. Kruger, D. A. May, L. C. McInnes, R. T. Mills, L. Mitchell, T. Munson, J. E. Roman, K. Rupp, P. Sanan, J. Sarich, B. F. Smith, H. Suh, S. Zampini, H. Zhang, H. Zhang, J. Zhang, PETSc/TAO 用户手册，技术报告 ANL-21/39 - 修订版 3.23，阿贡国家实验室，2025年。doi:10.2172/2565610.[77] Y. Saad, M. Schultz，GMRES：求解非对称线性方程组的广义最小残差算法，《SIAM科学与统计计算杂志》7 (1986年) 856–869。doi:10.1137/0907058.[78] S. Gowrisankar, S. Natesan，等分布网格上奇异摄动对流–扩散抛物型初始–边界值问题的稳健数值格式，《计算机物理通讯》185 (2014年) 2008–2019。doi:10.1016/j.cpc.2014. 04.004.[79] S. Giere, T. Iliescu, V. John, D. Wells，对流主导的对流–扩散–反应方程的SUPG降阶模型，Computer Methods in Applied Mechanics and Engineering 289 (2015年) 454–474。doi:10.1016/j.cma.2015.01.020.[80] T. J. R. Hughes, G. M. Hulbert，弹性动力学的时空有限元方法：形式与误差估计，Computer Methods in Applied Mechanics and Engineering 66 (1988年) 339–363。doi:10.1016/0045-7825(88)90006-0.[81] S. Mittal, T. E. Tezduyar，关于非定常不可压缩流稳定化时空有限元公式的注记，《计算机物理通讯》73 (1992年) 93–112。doi:10.1016/0010-4655(92)90031-s.[82] T. E. Tezduyar, S. Sathe，采用时空有限元对流固相互作用进行建模：求解技术，International Journal for Numerical Methods in Fluids 54 (2007年) 855–900。doi:10.1002/fld.1430.[83] T. E. Tezduyar, S. Sathe, R. Keedy, K. Stein，用于流固相互作用计算的时空有限元技术，Computer Methods in Applied Mechanics and Engineering 195 (2006年) 2002–2027。doi:10.1016/j.cma.2004.09.014.[84] O. Steinbach，抛物型问题的时空有限元方法，Computational Methods in Applied Mathematics 15 (2015年) 551566。doi:10.1515/cmam-2015-0026。