arXiv:2603.04711v1 [math.NA] 5 Mar 2026

工业过程的物理信息深度学习：用于热传导的时间离散变分物理信息神经网络

Manuela Bastidas Olivares[a]，Josué David Acosta Castrillón[a]、Diego A. Muñoz[a]

[a]哥伦比亚国立大学，麦德林校区，理学院，麦德林，哥伦比亚

摘要

神经网络为求解偏微分方程（偏微分方程）提供了强大的工具。我们提出了一种专为抛物型问题设计的变分物理信息神经网络（VPINN）。我们的方法将经典时间离散与复合损失函数相结合，在每个时间步最小化残差对偶范数。我们通过模拟工业圆筒中的咖啡提取物冻结过程验证了该框架。该模拟考虑了温度相关性质和实验数据，并成功捕捉了这一过程的热动力学。

关键词：物理信息神经网络，抛物型偏微分方程，变分方法，工业冻结，残差最小化*2020*年数学学科分类：68T07, 65M15, 65M99

1. 引言

抛物型偏微分方程（偏微分方程）是科学与工程中扩散过程建模的基础，热方程是其主要原型。因此，为这类方程开发高效的数值方法仍然重要，而在物理信息神经网络已成为经典求解器的一种强大无网格替代方法的今天，这一点更为突出。通过将微分算子直接嵌入损失函数，经典物理信息神经网络无需复杂的网格生成即可求解偏微分方程，并具有计算优势，正如 [4, 8, 11]中所述。近期文献广泛记载了它们在抛物型问题中的应用 [1, 3, 9, 10, 14]。然而，标准物理信息神经网络通常

电子邮箱地址：mbastidaso@unal.edu.co (Manuela Bastidas Olivares)

针对方程的强形式，这要求解具有较高的正则性，但当解缺乏光滑性或呈现尖锐梯度时（这在实际应用中十分常见），该约束便成为局限[7, 18, 19]。

为克服这些理论限制，我们采用弱变分形式，与常见的有限元实现类似，将导数转移到（具有足够正则性的）试验函数上，从而允许正则性较低的解。我们提出一种稳健的变分物理信息神经网络（RVPINN）框架，将变分原理与时间离散相结合。我们的方法不使用均方误差 [2, 5, 6], ，而是在每个时间步最小化残差对偶范数。正如 [13, 16] 所示并在 [15, 17]中采用的那样，该范数可作为严格的误差估计器。

我们在工业冻结领域展示了该框架的实际价值。具体而言，我们对咖啡提取物中的传热过程进行建模，其热物性随温度变化。这里，我们使用参考数据集表征材料属性：跨越相变温度范围的密度、热导率和热容；实验记录则提供必要的边界条件。为确保物理保真度，我们使用从工业咖啡提取物冷冻过程中获得的实验数据，对主要数值实验进行验证。具体的实验装置、测量规程和原始数据采集见 [12]。基于这一实验框架，我们证实该方法能够准确捕捉真实场景中的复杂热行为。

本文结构如下。第2 节详细介绍数学框架，推导适用于抛物型问题的时间离散变分形式，并进行经典的适定性分析。在第3节，我们引入神经网络的概念，将其作为偏微分方程解的近似。在该节中，我们构造了一个损失函数，用于准确近似残差对偶范数，从而作为误差的估计量。第4节通过对基准热方程和工业咖啡提取物应用开展数值实验，对所提出的方法进行验证。最后，第5节总结研究发现并讨论未来的发展方向。

2. 数学框架

我们考虑定义在有界开域 $\Omega \subset \mathbb{R}^d$ 上、具有利普希茨边界 $\partial\Omega$ 的时间相关扩散问题。该物理过程由能量守恒支配，从而得到一个系数依赖于解的非线性抛物型偏微分方程。

设 $u(x,t)$ 表示所关注的标量场（例如温度），其中 $x \in \Omega$ 且 $t \in (0, t_{\text{end}}]$。控制方程如下：

$$\frac{\partial}{\partial t}\Big(\mathcal{C}(u)\, u\Big) \;-\; \nabla \cdot \Big(\mathcal{K}(u)\, \nabla u\Big) \;=\; f(x,t), \tag{1}$$

满足以下边界条件和初始条件：

$$\begin{aligned} u(x,t) &= u_b(t), && \text{on } \partial\Omega \times (0, t_{\text{end}}], \\ u(x,0) &= u_I(x), && \text{in } \Omega. \end{aligned}$$

在该能量平衡模型中，内能项表示为乘积 $\mathcal{C}(u)u$，其中系数 $\mathcal{C}(\cdot)$ 对应于体积热容，定义为温度相关的比热容与材料密度的乘积。平衡方程中的扩散项为 $\mathcal{K}\nabla u$，表示通过热传导进行的热传递。这一具体机制由傅里叶定律建模，其中 $\mathcal{K}(\cdot)$表示热导率或扩散系数。我们假设所有参数都是状态变量 u 的严格正有界函数。项 $f(x,\ t)$ 表示体积源项。此外，u_I是初始条件，u_b 是边界条件。

注 1。为简单起见，我们假设采用齐次狄利克雷边界条件，并考虑线性情形，其中模型参数仅取决于空间坐标，而不取决于 u。这些假设有助于更清晰地阐述数学框架。随后，在第 4节中，我们将该方法扩展至一般非线性情形，以展示其在更复杂场景中的鲁棒性。

2.1. 时间离散

我们将时间区间 $(0, t_{\text{end}}]$ 离散为 N 个大小为 Δt 的步长，将第 n 步的时间记为 $t^n = n\Delta t$，将解记为 $u^n(x) \approx u(x, t^n)$。

采用后向欧拉离散近似时间导数，可得半离散方程：

$$\mathcal{C} \cdot \frac{u^n \;-\; u^{n-1}}{\Delta t} \;-\; \nabla \cdot \Big(\mathcal{K}\, \nabla u^n\Big) \;=\; f(x, t^n).$$

重新整理各项后，我们得到一个线性椭圆问题，需要针对每个 n依次求解：

$$\mathcal{C}\, u^n \;-\; \Delta t\, \nabla \cdot \Big(\mathcal{K}\, \nabla u^n\Big) \;=\; \underbrace{\mathcal{C}\, u^{n-1} + \Delta t\, f(x, t^n)}_{:= f^n(x)}. \tag{2}$$

2.2. 变分公式

为推导弱形式，我们将半离散方程 (2) 乘以试验函数 $v \in H_0^1(\Omega)$ 并在区域 Ω 上积分（应用格林第一恒等式），得到：

$$-\int_\Omega \nabla \cdot \left(\mathcal{K}\,\nabla u^n\right) v\,dx \;=\; \int_\Omega \mathcal{K}\,\nabla u^n \cdot \nabla v\,dx - \int_{\partial\Omega} \mathcal{K}(\nabla u^n \cdot \mathbf{n})\,v\,ds,$$

其中$\mathbf{n}$ 是 $\partial\Omega$ 的外法向单位向量。由于 $v \in H_0^1(\Omega)$， v 在边界上的迹为零，因此该方程化简为：

$$\int_\Omega \mathcal{C}\,u^n\,v\,dx \;+\; \Delta t \int_\Omega \mathcal{K}\,\nabla u^n \cdot \nabla v\,dx \;=\; \int_\Omega f^n\,v\,dx.$$

因此，在时间步 n，变分问题是求 $u^n \in H_0^1(\Omega)$，使得：

$$a(u^n, v) \;=\; l^n(v), \quad \forall v \in H_0^1(\Omega), \tag{3}$$

其中，参数相关双线性形式 $a(\cdot,\cdot)$ 和线性泛函 $l(\cdot)$ 定义如下：

$$a(w, v) := \int_\Omega \mathcal{C}\,w\,v\,dx \;+\; \Delta t \int_\Omega \mathcal{K}\,\nabla w \cdot \nabla v\,dx,$$

$$l^n(v) := \int_\Omega f^n\,v\,dx.$$

2.3. 适定性分析

根据拉克斯-米尔格拉姆定理，一系列变分问题的可解性得以保证。为应用该定理，我们必须验证双线性型 $a(\cdot,\cdot)$ 在 $H_0^1(\Omega) \times H_0^1(\Omega)$ 上既连续又强制。

我们为 $H_0^1(\Omega)$ 配备梯度范数：

$$\|v\|_{H_0^1} := \|\nabla v\|_{L^2(\Omega)}.$$

根据庞加莱不等式，存在一个常数 $C_P > 0$ （仅取决于 Ω），使得 $\|v\|_{L^2} \le C_P \|v\|_{H_0^1}$ 对所有 $v \in H_0^1$（Ω）均成立。我们假设材料性质在物理上是一致的，即存在正的常数，使得：

$$0 < c_{\min} \le \mathcal{C}(\cdot) \le c_{\max}, \quad \text{and} \quad 0 < k_{\min} \le \mathcal{K}(\cdot) \le k_{\max}.$$

连续性： 利用柯西-施瓦茨不等式和系数的上界，我们有：

$$\begin{aligned}|a(w,v)| &\le \int_\Omega |\mathcal{C}\, w\, v|\, dx + \Delta t \int_\Omega |\mathcal{K}\, \nabla w \cdot \nabla v|\, dx \\ &\le c_{\max}\|w\|_{L^2}\|v\|_{L^2} + \Delta t\, k_{\max}\|\nabla w\|_{L^2}\|\nabla v\|_{L^2}.\end{aligned}$$

对 L^2 项应用庞加莱不等式，可得：

$$|a(w,v)| \le c_{\max} C_P^2 \|w\|_{H_0^1}\|v\|_{H_0^1} + \Delta t\, k_{\max}\|w\|_{H_0^1}\|v\|_{H_0^1}.$$

因此，该双线性型是有界的，其连续性常数为 $M = c_{\max}C_P^2 + \Delta t\, k_{\max}$。

强制性： 利用系数的下界，我们考察二次型：

$$\begin{aligned}a(v,v) &= \int_\Omega \mathcal{C}\, v^2\, dx + \Delta t \int_\Omega \mathcal{K}\, |\nabla v|^2\, dx \\ &\ge c_{\min}\|v\|_{L^2}^2 + \Delta t\, k_{\min}\|\nabla v\|_{L^2}^2.\end{aligned}$$

由于第一项 $c_{\min}\|v\|_{L^2}^2$ 非负，我们可以对该表达式给出如下下界：

$$a(v,v) \ge \Delta t\, k_{\min}\|\nabla v\|_{L^2}^2 = \Delta t\, k_{\min}\|v\|_{H_0^1}^2.$$

定义 $\gamma := \Delta t\, k_{\min}$，则有 $a(v,v) \ge \gamma\|v\|_{H_0^1}^2$。由于 $\Delta t > 0$ 且 $k_{\min} > 0$，γ 严格为正，因此该型是强制的。根据拉克斯-米尔格拉姆定理，对于每个时间步 n，都存在唯一的解 u^n。

因此，在每个时间步，所得问题都是适定的，为分析解的近似误差提供了一致的数学框架。

2.4. 稳定性估计与基于残差的误差界

对于此类偏微分方程的弱形式，解的稳定性由下确界-上确界 条件（或巴布什卡-布雷齐条件）决定。由于我们的试探空间与检验空间均为同一个希尔伯特空间$V = H_0^1(\Omega)$，上一节确立的强制性足以保证该条件成立。

令 γ 为上一节推导出的强制性常数。通过在上确界中选取试验函数 $v = w$，我们得到：

$$\sup_{v \ne 0} \frac{|a(w,v)|}{\|v\|_V} \;\ge\; \frac{a(w,w)}{\|w\|_V} \;\ge\; \gamma\|w\|_V.$$

因此，双线性型满足内插条件：

$$\inf_{w\neq 0}\sup_{v\neq 0}\frac{|a(w,v)|}{\|w\|_V\,\|v\|_V}\;\geq\;\gamma\;>\;0. \tag{4}$$

不等式（4）结合双线性型的连续性（常数为 M），为我们的残差最小化（神经网络训练）策略提供了理论基础。

设 $u^*\in V$ 为特定时间步 n、给定前一状态 u^{n-1}时变分问题（3）的精确解。根据定义，它满足：

$$a(u^*,v)=l^n(v),\quad \forall v\in V.$$

对于任意近似解 $u^n\in V$（例如神经网络的输出），我们将残差泛函 $\mathcal{R}(u^n)\in V^*$（对偶空间 $H^{-1}(\Omega)$）定义为：

$$\langle\mathcal{R}(u^n),v\rangle_{V^*\times V}:=a(u^n,v)-l^n(v). \tag{5}$$

由于 $a(\cdot,\cdot)$ 关于其第一个自变量具有线性性，我们可以用误差 $e=u^n-u^*$ 来重写残差：

$$\langle\mathcal{R}(u^n),v\rangle_{V^*\times V}\;=\;a(u^n,v)-a(u^*,v)\;=\;a(u^n-u^*,v).$$

利用连续性和内插条件，我们可以用残差的对偶范数来界定误差 $\|u^n-u^*\|_V$：

$$\frac{1}{M}\|\mathcal{R}(u^n)\|_{V^*}\;\leq\;\|u^n-u^*\|_V\;\leq\;\frac{1}{\gamma}\|\mathcal{R}(u^n)\|_{V^*}. \tag{6}$$

方程（6）表明，误差的能量范数与残差的对偶范数等价。因此，最小化 $\|\mathcal{R}(u)\|_{V^*}$ 在数学上等价于减小每个时间步的解的真实误差。

现在，本工作的主要目标是设计一种策略，以便在每个时间步准确计算这一对偶范数。

3. 神经网络离散化

与我们的时间步进公式一致，我们将区间 $(0,t_{\text{end}}]$ 划分为 N 个均匀时间步，并使用单个全连接前馈神经网络近似空间解序列 $\{u^n\}_{n=1}^N$。与将时间视为输入坐标的标准方法不同，

我们采用一种离散化策略，使网络将空间坐标 $x \in \Omega$ 映射为离散时间步上的解向量。

设 $u_\theta : \Omega \to \mathbb{R}^N$ 表示由权重和偏置集合 θ 参数化的神经网络。该架构由 L 层组成。第 j层的输出记为z_j，其递归定义为：

$$z_j = \sigma_j(W_j z_{j-1} + b_j), \quad \text{for } j = 1, \ldots, L,$$

其中 $z_0 = x$ 为空间输入。函数 σ_j 表示隐藏层的非线性激活函数（例如tanh），而输出激活函数σ_L 为恒等函数。最终输出向量对应于每个时间步的近似 $\left[\hat{u}^1(x) \ldots, \hat{u}^N(x)\right]^\top$。

为严格施加齐次狄利克雷边界条件，我们采用不可训练的截断函数 $\chi : \Omega \to \mathbb{R}$。该函数被构造为对所有 $x \in \partial\Omega$ 满足 $\chi(x) = 0$，并在内部满足 $\chi(x) > 0$。第 n 个时间步的最终试探解由下式给出：

$$u_\theta^n(x) := \chi(x)\, \hat{u}^n(x),$$

由此，它满足齐次边界条件。

我们通过最小化总损失来确定网络参数 θ，使用固定学习率的Adam优化器$\mathcal{L}(u_\theta)$。在此，梯度通过自动微分计算，并以迭代方式（通过反向传播）更新网络的权重和偏置。该架构和优化过程如图 1 所示

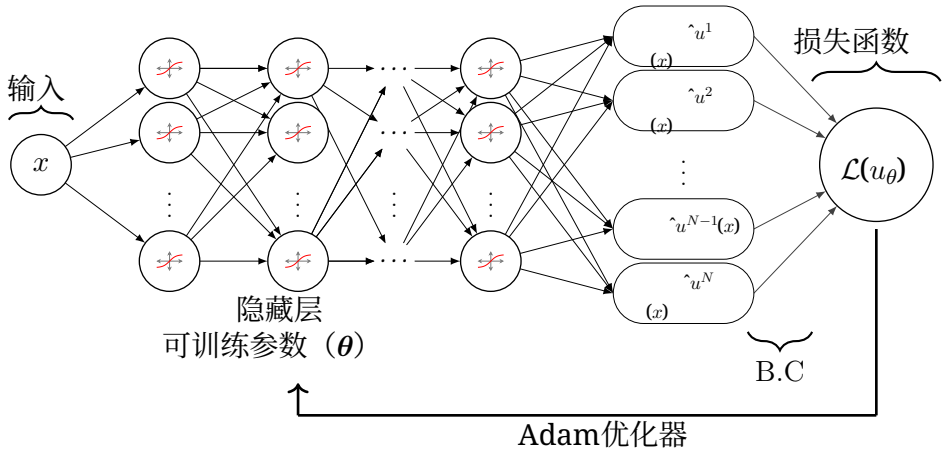

图 1：神经网络架构

3.1. 损失函数的计算

我们的目标是构造一个损失函数 $\mathcal{L}(u_\theta)$，使其最小化在数学上等价于最小化真实近似误差。

根据第 2 节的分析，在每个时间步，误差为 $\|u^n - u^*\|_{H_0^1}$，其中 u^* 是在时间 $t = n\Delta t$ 时的精确解。根据（6），该误差受残差 $\|\mathcal{R}（u^n）\|_{H^{-1}}$ 的对偶范数所界定。根据里斯表示定理，存在唯一的表示元 $\psi^n \in H_0^1$（Ω）使得：

$$\langle \mathcal{R}(u^n), v\rangle = (\psi^n, v)_{H_0^1}, \quad \text{and} \quad \|\mathcal{R}(u^n)\|_{H^{-1}} = \|\psi^n\|_{H_0^1}.$$

我们将表示元投影到测试空间 $H_0^1(\Omega)$ 的一组正交归一基 $\{\varphi_k\}_{k=1}^{N_{\text{test}}}$ 上，以此近似该范数。根据帕塞瓦尔恒等式，范数的平方等于各系数平方之和：

$$\|\mathcal{R}(u^n)\|_{H^{-1}}^2 = \|\psi^n\|_{H_0^1}^2 \approx \sum_{k=1}^{N_{\text{test}}} |\langle \mathcal{R}(u^n), \varphi_k\rangle|^2.$$

利用（5），我们便得到残差对偶范数的以下近似，从而能够稳健地估计每个时间步的误差

$$\begin{aligned} \|\mathcal{R}(u^n)\|_{H^{-1}}^2 &\approx \sum_{k=1}^{N_{\text{test}}} \Big| a(u^n, \varphi_k) - l(\varphi_k) \Big|^2 \\ &= \sum_{k=1}^{N_{\text{test}}} \left| \int_\Omega \mathcal{C}\, u^n\, \varphi_k \, dx + \Delta t \int_\Omega \mathcal{K}\, \nabla u^n \cdot \nabla \varphi_k \, dx - \int_\Omega f^n\, \varphi_k \, dx \right|^2 . \end{aligned} \tag{7}$$

因此，总损失函数定义为所有时间步上残差对偶范数的累积。代入（7），可得：

$$\mathcal{L}(u_\theta) = \Delta t \sum_{n=1}^{N_{\text{time}}} \sum_{k=1}^{N_{\text{test}}} \left| \int_\Omega \mathcal{C}\, u^n\, \varphi_k \, dx + \Delta t \int_\Omega \mathcal{K}\, \nabla u^n \cdot \nabla \varphi_k \, dx - \int_\Omega f^n\, \varphi_k \, dx \right|^2 \tag{8}$$

最小化(8) 可有效降低神经网络在整个时空域内近似解的误差。

注2. 例如，如果 $\Omega = (a, b)$，则关于内积 $(u, v)_{H_0^1} = \int_\Omega \nabla u \cdot \nabla v \, dx$ 的一种标准正交基选择为

$$\varphi_k(x) = \sqrt{\frac{2}{b-a}} \cdot \sin\left[k \left(\frac{\pi(x-a)}{b-a} \right) \right], \quad \textit{with} \quad k \in \mathbb{N}.$$

我们指出，也可以选择其他基函数；但是，这需要对与基函数表示相关的格拉姆矩阵求逆，而这一运算的计算开销较大，具体取决于所选函数 *[13]*。

4. 数值实验

在本节中，我们通过两项不同的数值实验评估所提出的变分物理信息神经网络（VPINN）框架的性能。

首先，我们考虑一个线性玩具问题：具有常系数的标准热方程。由于该问题具有已知的解析解，因此可将其用作基准，以验证我们方法的准确性。其次，我们将经过验证的框架应用于目标工业应用：咖啡提取物的非线性冻结。第二个案例涉及温度相关热物性，带来了更复杂的挑战，也展现了所提出的时间步进架构的稳健性。

4.1. 玩具问题：线性热方程

首先，我们考虑定义在区域$\Omega = (0, \pi)$上且满足$t_{\text{end}} = 1$的标准抛物型问题。我们寻求满足强形式（1）且$\mathcal{C}(\cdot) = \mathcal{K}(\cdot) \equiv 1$的弱解。此处，源项和初始条件的设置使得该问题的解析解为：

$$u^*(x,t) = e^{-t}\sin(x)\cos(x/2).$$

近似解由在*TensorFlow*中实现的全连接前馈神经网络进行参数化。该架构由 5 个隐藏层组成，每层包含 32 个神经元，并采用双曲正切（tanh）作为非线性激活函数。

为精确施加齐次狄利克雷边界条件，我们采用第 3 节所述的截断函数策略。原始网络输出与函数$\chi(x) = x(\pi - x)$相乘，从而确保试探解按构造在$x = 0$和$x = \pi$处为零。

网络参数在20^4次迭代中进行优化，使用学习率为10^{-2}并采用指数衰减的 Adam优化器。变分损失函数使用以下离散化方式计算：

- 时间：我们回顾一下，输出层的大小对应于时间离散中的时间步数，因此我们选择$N_{\text{时间}} = 128$步。

- 测试空间：我们采用前 $N_{\text{测试}}= 20$ 个傅里叶模态来近似测试空间 $H_0^1(\Omega)$。
- 数值求积：使用含有 $N_{\text{int}} = 128$ 个随机分布点的中点法来近似空间积分。

图 2比较了神经网络近似 u_θ 与四个不同时间步（$n = 1$、32、64、128）下的精确解。结果表明，该方法能够捕捉正弦波的扩散衰减，同时严格满足齐次边界条件。

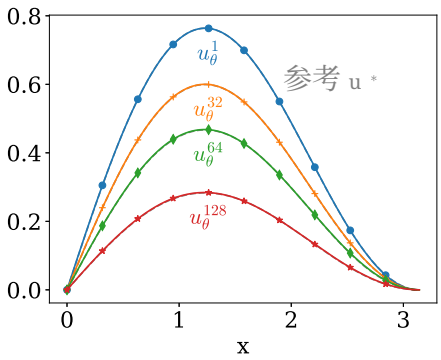

图 2：比较神经网络近似（u_θ^n）与四个不同时间步（$n = 1$、32、64、128）下的精确解。

图 3追踪了损失函数以及相对于精确解的相对误差的演变 u^*（x、t）。参考范数如下：

$$\|u^*\|_{L^2}^2 = \int_0^T \int_\Omega |u^*(x,t)|^2 \, dx \, dt \quad \text{and} \quad \|u^*\|_{H_0^1}^2 = \int_0^T \int_\Omega \left| \frac{\partial u^*}{\partial x}(x,t) \right|^2 dx \, dt.$$

这些参考值（$\|u^*\|_{L^2}$ 和 $\|u^*\|_{H_0^1}$）只需预先计算一次，并用作分母以对误差指标进行归一化。在图3中，我们观察到损失函数在收敛过程中出现振荡，这归因于变分公式的离散化限制。具体而言，测试空间使用了截断基，且积分约束采用数值近似，从而引入了积分噪声。尽管如此，这个简单示例表明，数值误差会随着训练损失同步降低，证实最小化（8）能够如预期那样有效降低误差（相差一个常数因子），这与（6）一致。

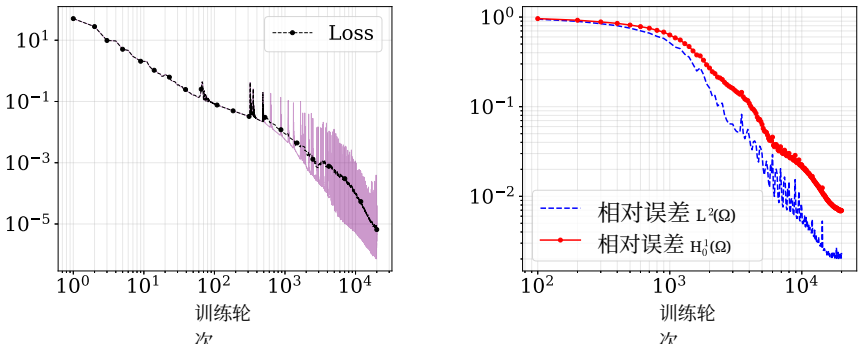

图3：训练过程中损失函数的变化（左）。相对误差（$L^2(\Omega)$ 和 $H_0^1(\Omega)$）（右）。

4.2. 咖啡提取物的非线性冻结

我们使用源自已发表实验装置的数据和参数来验证所提出的框架 [12]。在物理实验中，25千克咖啡提取物（总溶解固体（TDS）浓度为 31。1%）被装入一个圆柱形碳钢容器中。该容器内置分层塑料袋，并被置于温度维持在 -25℃ 的强制对流冷冻机中进行冷冻。我们直接从这项实验研究中建立的物理模型和组成模型中提取了高度依赖温度的材料性质：密度、热容和热导率。

为了评估我们的物理信息神经网络框架而设计数值实验，我们构建了该系统的一维简化。尽管实际的冻结过程发生在圆柱几何中，为了简化起见，我们假设存在一个平面一维笛卡尔域，其特征长度为 d，并假设该区域足够大，可以忽略横向效应。

热传递由热物性随温度变化的一维热传导方程控制：

$$\begin{aligned} \frac{\partial}{\partial \tau}\Big(\rho(T)c_p(T)\,T\Big) - \frac{\partial}{\partial s}\left(k(T)\frac{\partial T}{\partial s}\right) = 0, \quad s \in (0,d) \text{ and } \tau \geq 0, \\ T(0,\tau) = T(d,\tau) = T_b, \quad \tau \geq 0 \\ T(s,0) = T_I, \quad s \in [0,d]. \end{aligned} \tag{9}$$

此处， τ 表示时间 [s]， s 空间坐标 [m]，， T 以及温度 [℃]。参数 ρ、c_p和 k 分别表示密度 [kg/m^3]，、热容 [J/(kg $\cdot^\circ$ C)], 和热导率 [W/(m $\cdot^\circ$ C)]，（见图4）。

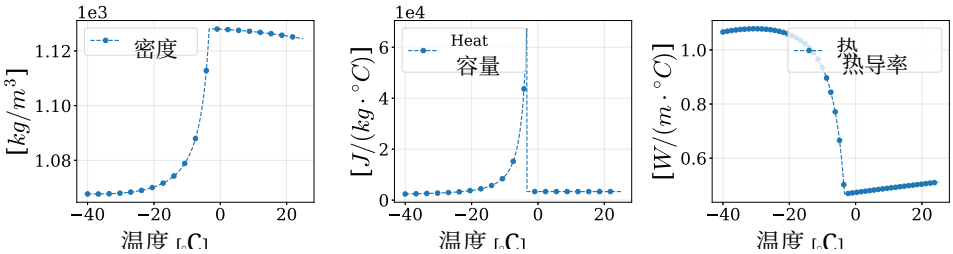

图4：咖啡提取物的热物性随温度的变化。

为提高数值稳定性和训练效率，我们引入以下无量纲变量：

$$u = \frac{T}{T_{\text{ref}}}, \quad x = \frac{s}{d}, \quad \text{and} \quad t = \frac{\tau}{t_0},$$

其中，T_{ref}为参考温度（取 $T_{\text{ref}} = T_I$），t_0 为特征扩散时间尺度，定义为 $t_0 = \frac{d^2 \rho_{\text{ref}} c_{p,\text{ref}}}{k_{\text{ref}}}$，其中 ρ_{ref}、$c_{p,\text{ref}}$和 k_{ref} 分别为参考密度、比热容和热导率。

将这些代入控制方程（9），可得无量纲形式：

$$\frac{\partial}{\partial t}\Big(\mathcal{C}(u)\, u\Big) - \frac{\partial}{\partial x}\left(\mathcal{K}(u)\frac{\partial u}{\partial x}\right) = 0, \quad x \in (0,1),\ t > 0,$$

其中，无量纲有效热容和热导率分别定义为 $\mathcal{C}(u) = \rho_{(T)} c_p(T)\, \rho_{\text{ref}} c_{p,\text{ref}}$和 $\mathcal{K}(u) = k_{(T)} k_{\text{ref}}$。

相应的无量纲边界条件和初始条件为：

$$u(0,t) = u(1,t) = \frac{T_b}{T_I}, \quad t \geq 0,$$
$$u(x,0) = 1, \qquad x \in [0,1].$$

为与 [12]，中的实验保持一致，我们将初始温度设为 $T_0 = 20°\text{C}$，并将圆柱体测得的边界温度作为数据驱动的边界条件。

我们使用全连接前馈神经网络对近似解进行参数化。该架构包含 5 个隐藏层，每层有 32 个神经元，并采用双曲正切（tanh）激活函数。

我们使用 Adam优化器对网络参数进行 10^5 次迭代优化，并采用初始值为 10^{-3} 的余弦衰减学习率。

我们在包含 128 个时间步的时间离散上计算变分损失，其大小与网络的输出层严格匹配。

需要指出的是，如前一项实验所述，我们使用截断函数将经验边界数据强制纳入神经网络。此外，在本实验中，我们在希尔伯特空间 $H^1(\Omega)$ 而非 $H_0^1(\Omega)$ 中定义正交归一基函数。我们采用 $N_{\text{test}} = 64$傅里叶模态，并通过随机中点数值求积规则，利用 256 个均匀采样的随机点计算弱形式积分。

图5 展示了两种情形下温度分布的时间演化：

- **非线性模型（u）：** 纳入实验测得的温度相关热物理参数，如图4所示。我们强调，神经网络通过求解直接逼近完整的非线性问题。这完全避免了任何对控制方程进行预先线性化，处理非线性项在损失形式中原生处理，类似于前向显式格式。
- **线性模型（u_0）：** 假设材料参数恒定（$\mathcal{C} = \mathcal{K} = 1$）.

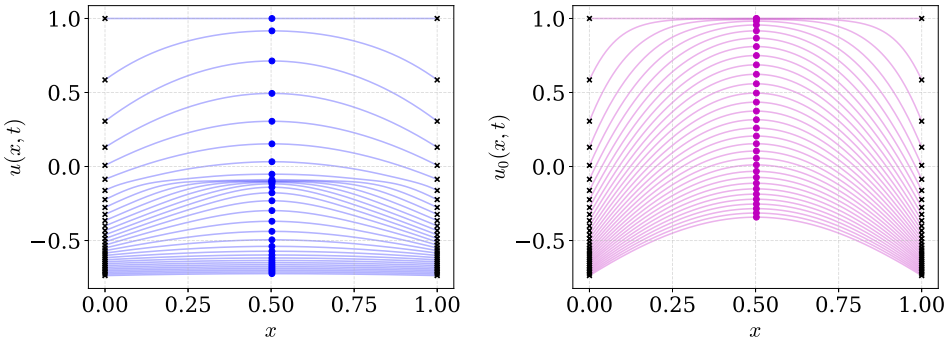

图5：温度分布的时间演化。非线性解 u （左）纳入了经验热物理参数，而线性对照解 u_0 （右）假设恒定属性。数据驱动的边界值以 ×标出，中点温度以 •突出显示。

比较这两个解，可以揭示非线性系数对物理过程的明确影响。线性模型 u_0 以恒定速率扩散，而非线性解 u 则表现出显著不均匀的扩散动力学，这是由有效热

容 $\mathcal{C}(u)$ 和导热系数 $\mathcal{K}(u)$的温度依赖性所驱动的。尤其是，当温度接近相变区域时，热物性的变化会改变局部扩散时间尺度，与线性情形相比，导致热渗透速度减慢。这一效应符合食品冻结过程预期的物理行为：潜热现象和材料结构变化会改变其表观热惯性。线性模型与非线性模型之间的差异不仅体现在定量方面，也体现在定性方面。在线性情形下，温度分布曲线始终保持对称且平滑扩散的特征。相比之下，非线性模型在中间时段的边界附近呈现更陡峭的梯度，反映了施加的实验边界数据与不断演化的材料系数之间的相互作用。这证实，忽略温度相关性质可能会导致对工业应用中瞬态热状态的系统性误表征。图6进一步凸显了这种差异，其中中点温度演化清晰地区分了这两种情形。与线性近似相比，非线性模型预测的冷却动力学有所延迟，捕捉到了由可变热容引起、且与实验结果一致的热缓冲效应。重要的是，VPINN框架保持稳定，并能准确解析这种非线性，无需预先进行线性化或算子分裂。网络直接在损失泛函中处理完全非线性的弱形式，表明残差对偶范数最小化为处理复杂的系数相互作用提供了足够的结构。

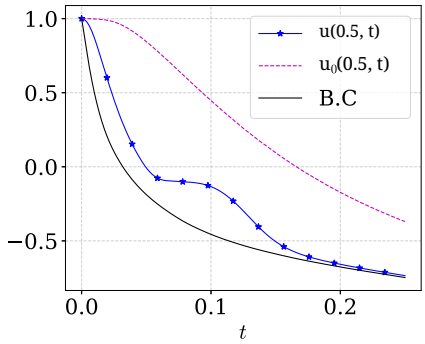

图6：两种模型的中点温度演化，并与实验边界条件一同绘出。

最后，图7 展示了训练过程中损失函数的收敛情况。我们在 10^5 次迭代后终止优化，因为进一步改进神经网络已无实际意义。

对神经网络的改进已无实际意义。总体而言，结果

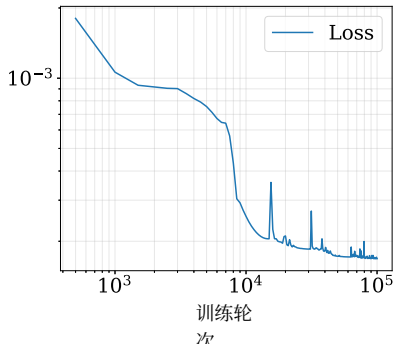

图7： 10^5 次训练迭代中损失函数的变化。

证实所提出的 RVPINN 框架能够在具有实际工业意义的场景中求解系数与温度相关的完全非线性抛物型问题，同时保持物理可解释性和数值稳健性。

5. 结论

我们提出了一种用于求解抛物型偏微分方程的时间离散变分物理信息神经网络（RVPINN）。该方法将经典时间离散与残差对偶范数最小化策略相结合，为标准均方误差物理信息神经网络提供了一种具有理论依据的替代方案。

数值实验证实了该方法的准确性和稳健性。尤其是工业冻结应用表明，该方法无需线性化即可成功处理温度相关系数，并捕捉线性与非线性扩散动力学之间具有物理意义的差异。

总体而言，所提出的 RVPINN 为瞬态扩散过程提供了一种数学上一致且计算稳定的经典物理信息神经网络替代方案。未来工作将研究更高维构型以及更复杂的多物理场耦合。

支持本研究的源代码和实验数据可在公共代码仓库中获取。其中包括基于 TensorFlow 的 RVPINN 实现，以及复现训练过程、图表和分析所需的全部笔记本。该仓库还提供了用于

实测边界条件和温度相关热物性参数的实测数据集。完整的代码仓库可免费访问，地址为
https://github.com/manubastidas/parabolicRVPINNs。

致谢

本研究由哥伦比亚国立大学麦德林校区通过项目 HERMES 63334 和 HERMES 63067 资助。

关于在写作过程中使用生成式人工智能和人工智能辅助技术的声明

在准备本研究期间，作者使用了 Gemini（谷歌）来协助写作、纠正英语用法，并改善文本的清晰度和组织结构。在使用该工具/服务后，作者根据需要对内容进行了审阅和编辑，并对出版物的内容承担全部责任。

参考文献

[1] C. Beck、鄂维南和 A. Jentzen，用于高维全非线性偏微分方程和二阶倒向随机微分方程的机器学习近似算法，《非线性科学杂志》，29（2019年），第 1563–1619 页。

[2] S. Berrone、C. Canuto 和 M. Pintore，通过变分物理信息神经网络求解偏微分方程：后验误差分析，《费拉拉大学年刊》，68（2022年），第 575–595 页。

[3] S. Cai、Z. Wang、S. Wang、P. Perdikaris 和 G. E. Karni- adakis，用于热传递问题的物理信息神经网络，《传热学杂志》，143（2021年），第 060801 页。

[4] S. Cuomo、V. S. Di Cola、F. Giampaolo、G. Rozza、M. Raissi 和 F. Piccialli，通过物理信息神经网络实现科学机器学习：现状与未来，《科学计算杂志》，92（2022年），第 88 页。

[5]E. Kharazmi、Z. Zhang 和 G. E. Karniadakis，用于求解偏微分方程的变分物理信息神经网络，arXiv 预印本 arXiv:1912.00873，（2019年）。

[6] ——，*hp*-变分物理信息神经网络：采用域分解的变分物理信息神经网络，《计算机方法：应用力学与工程》，374（2021年），第 113547 页。[7] A. Krishnapriyan、A. Gholami、S. Zhe、R. Kirby 和 M. W. Mahoney，物理信息神经网络中可能失效模式的特征分析，《神经信息处理系统进展》，34（2021年），第 26548–26560 页。[8] L. Lu、X. Meng、Z. Mao 和 G. E. Karniadakis，*DeepXDE*：用于求解微分方程的深度学习库，《SIAM评论》，63（2021年），第 208–228 页。[9]X. Meng、Z. Li、D. Zhang 和 G. E. Karniadakis，*PPINN*：用于随时间变化的偏微分方程的 *Parareal*物理信息神经网络，《计算机方法：应用力学与工程》，370（2020年），第 113250 页。[10] N. Nüsken 和 L. Richter，在倒向随机微分方程与物理信息神经网络之间插值：用于椭圆型和抛物型边值问题的深度学习，《机器学习杂志》，2（2023年），第 31–64 页。[11]M. Raissi、P. Perdikaris 和 G. E. Karniadakis，物理信息神经网络：用于求解涉及非线性偏微分方程的正向和逆向问题的深度学习框架，《计算物理学杂志》，378（2019年），第 686–707 页。[12] C. Restrepo-Díaz、N. Presiga-Posada、S. Velasquez、L. Gil 和 D. A. Muñoz，用于咖啡提取物高效工业冷冻过程的容器设计优化，《食品与生物产品加工》，（2025年）。[13]S. Rojas、P. Maczuga、J. Muñoz-Matute、D. Pardo 和 M. Paszyński，稳健变分物理信息神经网络，《计算机方法：应用力学与工程》，425（2024年），第 116904 页。[14] Y. Shin、J. Darbon 和 G. E. Karniadakis，关于物理信息神经网络求解线性二阶椭圆型和抛物型偏微分方程的收敛性，《计算物理学通讯》，28（2020年）。

[15] J. M. Taylor, M. Bastidas, D. Pardo, 和 I. Muga，用于求解时谐麦克斯韦方程的深度傅里叶残差方法，《计算物理学杂志》，523（2025年），第 113623 页。[16] J. M. Taylor、D. Pardo 和 I. Muga，利用神经网络求解偏微分方程的深度傅里叶残差方法，《计算机方法：应用力学与工程》，405（2023年），第 115850 页。[17] C. Uriarte、M. Bastidas、D. Pardo、J. M. Taylor 和 S. Ro- jas，利用最小二乘法优化变分物理信息神经网络，《计算机与应用数学》，185（2025年），第 76–93 页。[18] S. Wang、Y. Teng 和 P. Perdikaris，理解并缓解物理信息神经网络中的梯度流病态问题，《SIAM科学计算杂志》，43（2021年），第 A3055–A3081 页。[19] S. Wang、X. Yu 和 P. Perdikaris，物理信息神经网络何时以及为何训练失败：神经切线核视角，《计算物理学杂志》，449（2022年），第 110768 页。