用于热力耦合系统和非均质域的物理信息神经网络混合形式

Ali Harandi[1, *], Ahmad Moeineddin[2], Michael Kaliske[2], Stefanie Reese[1], Shahed Rezaei[3, *]

[1]应用力学研究所，亚琛工业大学，*Mies-van-der-Rohe-Str. 1*，*D-52074* 亚琛，德国

[2]结构分析研究所，德累斯顿工业大学，*Georg-Schumann-Str. 7*，*D-01187* 德累斯顿，德国

[3]*Access e.V*，*Intze-Str. 5*，*52072* 亚琛，德国

[*] 通讯作者：*ali.harandi@ifam.rwth-aachen.de*，*s.rezaei@access-technology.d* e

arXiv:2302.04954v2 [cs.CE] 6 Sep 2023

摘要

物理信息神经网络（PINNs）是一种求解边值问题的新工具，它根据控制方程、边界条件和初始条件定义神经网络的损失函数。近期研究表明，在为许多工程问题设计损失函数时，使用一阶导数并结合强形式和弱形式的方程，可以显著提高精度，尤其是在区域存在非均质性和变量跃变的情况下。这种新方法称为 PINNs 的混合形式，其思想源自混合有限元法。在该方法中，偏微分方程被重新表述为一个方程组，其中主要未知量是解的通量或梯度，次要未知量则是解本身。本文提出将混合形式应用于多物理场问题，具体而言，是求解一个稳态热力耦合方程组。此外，我们讨论序贯无监督训练和全耦合无监督训练，并比较二者的精度和计算成本。为提高网络的精度，我们引入硬边界约束，以确保预测结果有效。随后，我们研究不同优化器和网络架构对精度与效率的影响。最后，我们介绍一种类似于迁移学习的简单参数化学习方法。该方法结合数据与物理规律，以应对 PINNs 在计算成本方面的局限，并提升网络对未见案例的系统响应预测能力。本文的研究成果将有助于深度学习应用于许多其他耦合方程组的工程应用，从而实现快速、可靠的计算。

关键词：物理信息神经网络、热力耦合问题、非均质固体、硬约束、参数化学习

1. 引言

深度学习（DL）方法是机器学习（ML）的一个分支，在逼近任意函数和算子方面具有显著的灵活性 [1, 2]。它们通过从训练数据集中发现隐藏模式来完成这项任务（也称为监督学习）。此外，DL 方法仅通过考虑物理约束（无监督学习），就能够求解边值问题（BVP）。因此，

DL 模型已广泛应用于诸多工程领域，从材料设计到结构分析皆有涉及 [3]。与经典求解器相比，机器学习算法的一项优势是成功训练后能够大幅提速。这些模型的高效率使其在多尺度分析中尤为重要，因为此类分析的瓶颈在于上下尺度之间的信息变换过程（参见 [4, 5, 6] 及其中的参考文献）。一个重要问题是，网络必须针对足够多的观测数据进行训练，并涵盖所有相关场景。否则，就不能完全依赖网络的输出结果。在数据训练过程中加入已知的物理约束，有助于提升网络的预测能力（参见 [7] 及其中的参考文献）。要充分发挥将物理定律应用于经典数据驱动方法的优势，还需针对不同的测试案例和物理问题，对这一点进行充分研究。

在此，我们旨在探讨深度学习方法在计算工程领域的潜力。在该领域中，我们需要处理复杂的异质域以及由控制方程构成的耦合系统。下文将回顾一些重要研究成果，并介绍本研究的新贡献。

基于数值求解器生成的现有物理数据训练机器学习算法，可以加快解的预测过程。Park等 [8] 采用多尺度核神经网络，提高了对未见案例中应变场的预测能力。在传统的机器学习方法中，神经网络仅基于数据进行训练，问题的底层物理规律被编码在数据本身，而这些数据是由其他数值求解器生成的。Yang等 [9] 采用深度学习方法预测复合材料中的复杂应力场和应变场。Khorrami等 [10] 采用卷积神经网络，预测具有历史相关材料行为的应力场。另请参见 [11], ，作者基于低分辨率数据进行训练，采用傅里叶神经算子预测高分辨率应力场。神经网络代理模型是求解耦合热、化学和水文工程问题中复杂方程组的强大模型 [12]。数据驱动模型可根据驾驶工况估计荷电状态 [13] ，也可作为复杂系统（如质子交换膜燃料电池）的数字孪生 [14]。另请参见 [15, 16]，其中比较了适用于工业应用的不同神经算子的性能。

有趣的是，遵循物理信息神经网络（PINNs）的理念 [17]，并在最终损失函数中采用物理定律，我们便可以合理地信赖网络的输出。这是因为网络输出会自动遵循并满足重要的物理约束。然而，对于 PINN 方法的收敛性和泛化性质，目前仍缺乏恰当且深入的理论理解，相关研究仍在进行。因此，许多作者也进一步拓展并探索 PINNs 背后的理念，以提升网络在未见情形下的性能。相关示例见[18, 19, 20]。如果问题的底层物理规律已完全明确且完备，那么无需任何初始数据，神经网络便可仅依据给定的边值问题 [7]进行训练。后一点也称为无监督学习，因为事先不需要该问题的初始解。Hu等 [21] 率先探讨了基于域分解方法的扩展 PINNs 在多尺度和多物理场问题建模中的有效性。作者指出，这种方法在泛化能力方面存在权衡。一方面，将复杂偏微分方程的解分解为若干简单部分，可以降低问题的复杂度；另一方面，分解可能导致过拟合，使所得解的泛化能力下降。另见 Jagtap等 [19], 、D. Jagtap 和 Em Karniadakis[22], 、Henkes等 [23] 以及 Wang等 [24] 关于域分解理念的研究。Wang

等 [25] 研究了 PINNs 如何倾向于沿神经切线核的主导特征方向学习函数，并提出了能够构建稳健且准确的 PINNs 模型的架构。在 PINN 表述中，不同损失项的重要性可能有所不同。McClenny 和 Braga-Neto [26]提出为每个训练点采用自适应权重，使神经网络能够识别解中较难处理的区域，并被促使将注意力集中于这些区域。其基本思路是让权重随相应损失的增加而增大。Xiang等 [27]也指出，PINNs 的性能会受到损失权重的影响，并建立了高斯概率模型，通过为每个损失项设置自适应权重来定义自适应损失函数。另见 [28, 29]。另一种颇具前景的扩展思路是降低构造损失函数时所需的微分阶数。为此，研究人员结合偏微分方程变分形式中的不同思路，并将其应用于各种工程应用问题 [30]。这一思路还进一步拓展为所谓的混合形式，即同时采用强形式和弱形式 [7, 23, 31, 32]。

工程应用的一个重要方面是考虑多物理场特性。换言之，在许多实际问题中，必须同时考虑各种相互耦合且高度非线性的场。例如，在固体力学中，除了力学变形之外，通常还需要考虑热场 [33],化学浓度 [34], 电场和/或磁场 [35], 损伤场 [36, 37], 以及其他因素的影响。

多物理场背景下采用物理信息神经网络求解相关方程的思路，已由多位作者探讨。在这类问题中，我们需要处理代表不同物理过程的损失项，而这些损失项的数值大小可能截然不同。因此，通常有必要平衡不同的损失项。由此，经典物理信息神经网络模型仍需进一步改进。Amini 等人 [38] 研究了将物理信息神经网络用于求解涉及热-水-力过程的问题。作者提出采用控制方程的无量纲形式，并使用顺序训练策略，以提高多目标优化问题的求解精度。Raj 等人 [39] 利用物理信息神经网络研究了功能梯度材料的热-力问题及网络对位移的逼近能力，并指出，在逼近应力场时，网络产生的误差会高出一个数量级。Laubscher [40] 提出了用于物理信息神经网络的单一网络和分离网络，以预测干空气加湿问题中的动量、组分和温度分布，并研究了干空气加湿中的流体性质与热扩散。据报道，与单一网络相比，分离网络的损失更低。Mattey 和 Ghosh [41] 也指出，对于 Allen-Cahn 方程和 Cahn-Hilliard 方程等强非线性高阶偏微分方程，物理信息神经网络的精度会受到影响。作者提出使用单个神经网络，在连续的时间段上依序求解偏微分方程。另请参阅 Chen 等人[42], 、Nguyen 等人 [43] 以及 Bischof 和 Kraus [28]的研究。在采用弥散方法对固体中的裂纹扩展进行建模时，通常将损伤指定为一个场变量。在这种情况下，损伤场的演化方程通常与力学平衡方程相耦合[37]。Goswami 等人 [36] 提出了用于脆性断裂分析的 DeepONet 物理信息变分形式，另请参阅 [44, 45]的相关研究。

利用迁移学习算法 [46] 可以提高训练的计算效率并降低计算成本。迁移学习的核心思想是捕捉源域的共同特征和行为，并将其迁移到新的目标域。为此，可以迁移神经网络预训练的超参数，并将其初始化为新区域中神经网络的超参数

。后者可以降低训练的计算成本 [7, 47]。Tang 等人[48] 利用迁移学习物理信息神经网络求解涡激振动问题。Goswami 等人[49] 采用迁移学习的思路，对网络的超参数进行部分重新训练，以识别某一区域中的裂纹模式。Pellegrin 等人 [50] 利用迁移学习的思路，通过设置不同的配置（初始条件）来求解方程组，并表明，使用训练好的多头网络能够显著加快分支流动模拟，同时不损失精度。Desai 等人 [51] 利用物理信息神经网络，通过迁移学习求解偏微分方程线性方程组，例如泊松方程和薛定谔方程。Gao 等人 [52] 提出了一种基于奇异值的物理信息神经网络迁移学习算法，并展示了其在 Allen-Cahn方程中的性能。

根据所综述的文章可以清楚地看出，采用物理信息神经网络方法进行工程应用时，若要获得快速且可靠的结果，就必须谨慎行事。朴素版本的物理信息神经网络可能无法收敛到正确的解；即使能够收敛，与其他基于有限元法的数值求解器相比，其训练过程也可能耗时显著更长。本文探讨了如何解决上述问题。我们将重点讨论用于非均质固体稳态热力耦合系统的一阶物理信息神经网络表述，并比较耦合训练与序贯训练的性能（另见图 1）。第 2 节推导控制方程及其变分形式。第 3 节介绍新的物理信息神经网络混合形式的网络架构。第 4 节和第 5 节分别给出结果和结论。

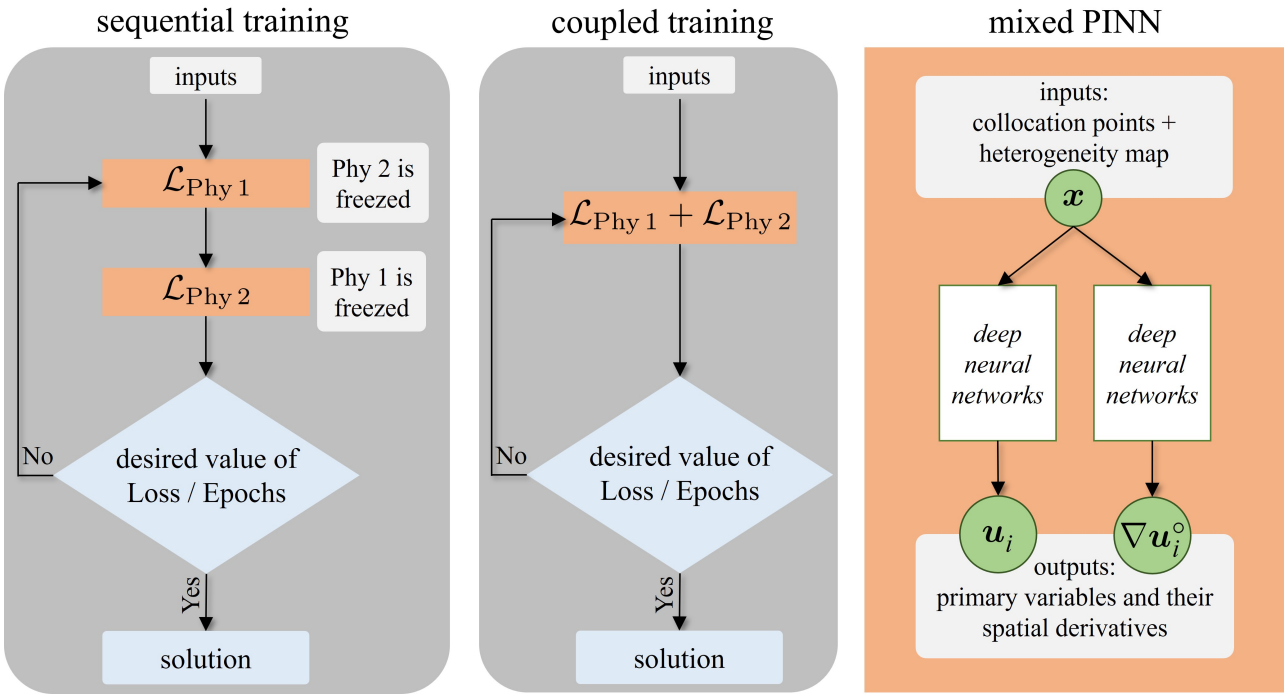


图 1：用于求解多物理场边值问题的序贯训练和耦合训练算法，以及混合物理信息神经网络架构。 $\mathcal{L}_{\text{Phy}}$ 表示相关的物理定律。变量 $\boldsymbol{x}$表示配点的位置。变量 $\boldsymbol{u}_i$ 表示第 i 个主变量（解），而 $\nabla \boldsymbol{u}_i^o$ 表示其空间梯度。随后利用相互独立的神经网络来预测主变量及其空间梯度。

2. 问题公式化

2.1. 非均质固体中的稳态热弹性

下面首先概述热-力问题的公式化。在这个问题中，我们处理的是由向量 $\boldsymbol{u}$ 表示的位移场，以及由标量参数 T 表示的温度场。通过不同的耦合项，这两个场会相互影响。从运动学开始，总应变 $\boldsymbol{\varepsilon}$ 可加性地分解为弹性部分和热部分，写作

$$\boldsymbol{\varepsilon} = \boldsymbol{\varepsilon}_e + \boldsymbol{\varepsilon}_t = \operatorname{sym}(\operatorname{grad}(\boldsymbol{u})) = \frac{1}{2}\left(\nabla \boldsymbol{u} + \nabla \boldsymbol{u}^T\right). \tag{1}$$

在式（1）中，$\boldsymbol{\varepsilon}_e$ 表示弹性应变张量，$\boldsymbol{\varepsilon}_t$ 表示由温度场引起的热应变。对于各向同性材料行为，热应变张量遵循线性膨胀定律，可写为

$$\boldsymbol{\varepsilon}_t = \alpha(x,y)\,(T - T_0)\,\boldsymbol{I}. \tag{2}$$

这里，α 为热膨胀系数；由于我们研究的是非均质微观结构，该系数可随空间变化。此外，T_0 为初始温度，$\boldsymbol{I}$ 为二阶单位张量。利用四阶弹性张量 $\mathbb{C}$，可将应力张量写为

$$\boldsymbol{\sigma} = \mathbb{C}(x,y)\,(\boldsymbol{\varepsilon} - \boldsymbol{\varepsilon}_t). \tag{3}$$

请注意，此处我们假设弹性性质与温度无关。考虑温度相关材料性质的问题将在未来工作中研究，而当前工作的研究内容可轻松扩展至此类问题。将式（3）写成Voigt记号形式为

$$\hat{\boldsymbol{\sigma}} = \hat{C}(x,y)\,(\hat{\boldsymbol{\varepsilon}} - \hat{\boldsymbol{\varepsilon}}_t). \tag{4}$$

在式（4）中，$\hat{\bullet}$ 表示Voigt记号下的张量（$\bullet$）。对于二维平面应变设置，位置相关弹性张量 $\mathcal{C}$ 写为

$$\hat{C}(x,y) = \frac{E(x,y)}{(1-2\nu(x,y))(1+\nu(x,y))}\begin{bmatrix} 1-\nu(x,y) & \nu(x,y) & 0 \\ \nu(x,y) & 1-\nu(x,y) & 0 \\ 0 & 0 & \dfrac{1-2\nu(x,y)}{2} \end{bmatrix}. \tag{5}$$

这里，杨氏模量 E 和泊松比 ν 表示材料的弹性常数。各材料点均满足各向同性假设。值得一提的是，异质域会导致材料参数依赖于配点的坐标。对于力学场，在不存在体力向量的情况下，线动量平衡方程以及狄利克雷和诺伊曼边界条件写作

$$\operatorname{div}(\boldsymbol{\sigma}) = \boldsymbol{0} \quad \text{in} \quad \Omega, \tag{6}$$

$$\boldsymbol{u} = \bar{\boldsymbol{u}} \quad \text{on} \quad \Gamma_D, \tag{7}$$

$$\boldsymbol{\sigma}\cdot\boldsymbol{n} = \boldsymbol{t} = \bar{\boldsymbol{t}} \quad \text{on} \quad \Gamma_N. \tag{8}$$

接下来，我们写出无热源情况下的能量平衡方程，以及稳态问题（[33]）

$$\operatorname{div}(\boldsymbol{q}) = 0 \quad \text{in } \Omega, \tag{9}$$
$$T = \bar{T} \quad \text{on } \bar{\Gamma}_D, \tag{10}$$
$$\boldsymbol{q} \cdot \boldsymbol{n} = q_n = \bar{q} \quad \text{on } \bar{\Gamma}_N. \tag{11}$$

在上述方程中， Ω和 Γ 分别表示力学场中物体内部和边界上的材料点。温度边界条件也在Γ上得到满足。在式（9）和（11）中， $\boldsymbol{q}$表示热通量，并根据傅里叶定律定义为

$$\boldsymbol{q} = -k(x, y)\, \nabla T. \tag{12}$$

这里， k（x， y）表示随位置变化的导热系数 .

为求得式（6）的伽辽金弱形式（WF），将其乘以试函数$\delta \boldsymbol{u}$ ，该试函数同样满足力学场的狄利克雷边界条件。经过分部积分并应用诺伊曼边界条件后，最终得到

$$\int_{\Omega} \delta \hat{\boldsymbol{\varepsilon}}_e^T\, \hat{C}(x, y)\, \hat{\boldsymbol{\varepsilon}}_e \; dV \; - \; \int_{\Gamma_N} \delta \boldsymbol{u}^T\, \bar{\boldsymbol{t}} \; dA = 0. \tag{13}$$

考虑力学内能 E_{int}^M 与力学外能 E_{ext}^M之间的平衡，我们将式（13）中的弱形式写为总能量的最小化问题，见[30], ：

$$\text{Minimize: } \underbrace{\int_{\Omega} \frac{1}{2} \hat{\boldsymbol{\varepsilon}}_e{}^T\, \hat{C}(x, y)\, \hat{\boldsymbol{\varepsilon}}_e \; dV}_{E_{\text{int}}^M} - \underbrace{\int_{\Gamma_N} \boldsymbol{u}^T\, \bar{\boldsymbol{t}} \; dA}_{E_{\text{ext}}^M}, \tag{14}$$

其满足式（7）和式（8）中的边界条件。将式（9）乘以试函数 δT ，对力学场采用类似的步骤，并应用式（10）和（11）中的边界条件，可得到热场的伽辽金弱形式:

$$\int_{\Omega} k(x, y)\, \nabla^T T\, \delta(\nabla T) \; dV \; + \; \int_{\bar{\Gamma}_N} \bar{q} \; \delta T \; dA \; = \; 0. \tag{15}$$

还可以将公式（15）中的弱形式重新表述为能量变分。后者启动了热内能与外能的平衡，由此得到

$$\text{Minimize: } \underbrace{\int_{\Omega} \frac{1}{2}\, k(x, y)\, \nabla^T T\, \nabla T \; dV}_{E_{\text{int}}^T} + \underbrace{\int_{\bar{\Gamma}_N} \bar{q}\, T \; dA}_{E_{\text{ext}}^T}. \tag{16}$$

在公式（10）和公式（11）中的边界条件约束下，对上述方程进行最小化。

在公式 (14) 中，耦合训练过程中针对位移场 $\boldsymbol{u}$ 和温度场 T 进行最小化。然而，通过采用序贯训练并固定一个场变量，相关损失函数则针对位移场 $\boldsymbol{u}$进行最小化。对于公式 (15)，由于公式化中不存在其他场，因此在两种训练过程中，最小化都是针对温度场 T 进行的。

与先前的研究类似，公式 (13) 和 (15) 中的表达式分别被添加到主变量 $\boldsymbol{u}$ 和 T对应的神经网络损失函数中。

根据图2，考虑两组不同的几何结构。第一种几何结构用于表示纤维基体结构。第二种几何结构代表一种先进的工程合金微观结构，其中包含矩形夹杂。相1（基体）的属性以下标”mat”表示，而相2（夹杂物）的材料属性以下标”inc”表示。接下来，我们定义比值 $E_{\mathrm{mat}}/E_{\mathrm{inc}} = k_{\mathrm{mat}}/k_{\mathrm{inc}} = R$。左侧和右侧几何结构中的值 R 分别等于0.3和2。针对这两种结构求解稳态热弹性问题。最后，边界条件汇总于图3。对于力学场，左右两部分被限制为沿 x 方向移动，上下边缘则沿 y方向固定。对于热场，左右边缘施加恒定温度，而上下边缘沿表面法向的热通量为零。

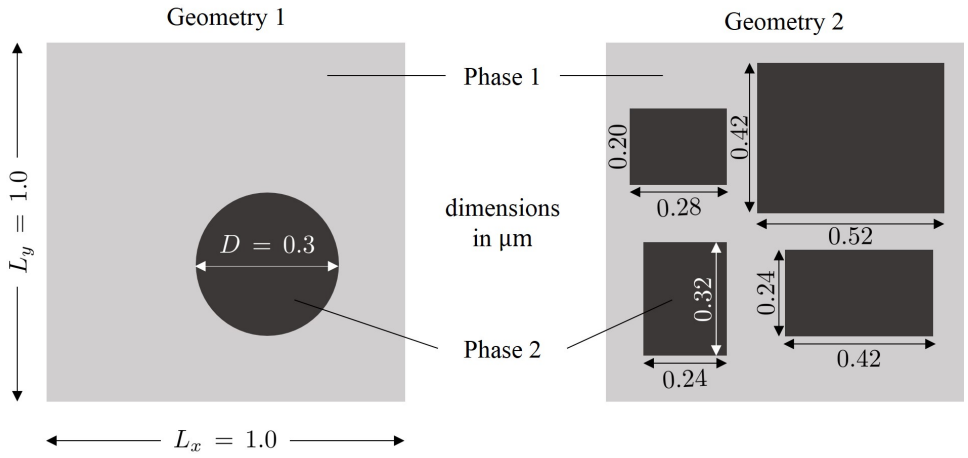


图 2：用于研究热弹性问题的选定几何结构。颜色变化表示杨氏模量和热导率的不同取值。

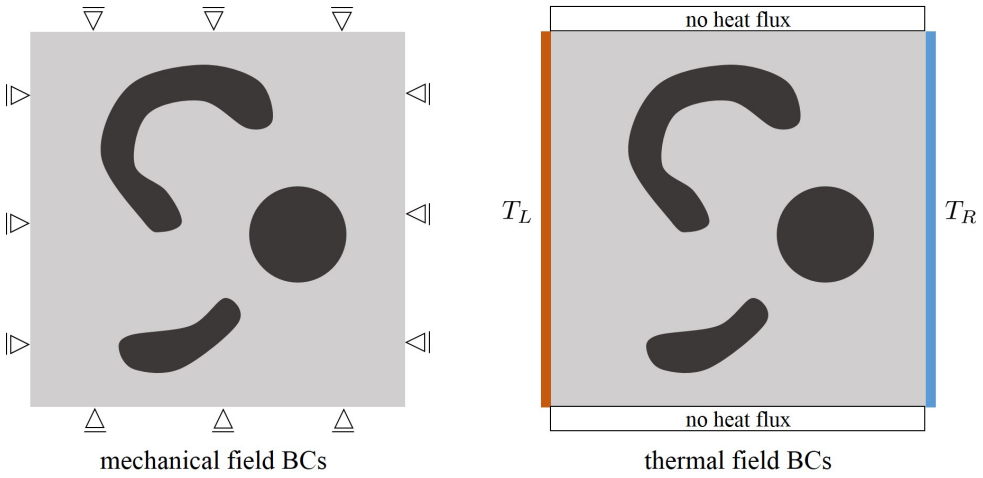


图 3：热力耦合问题的边界条件。左：力学场边界条件。右：热场边界条件。

3. 物理信息神经网络的网络架构

3.1. 热弹性问题

基于 Rezaei 等人提出的混合形式 [7], ，网络的输入为配点的位置（即二维情况下的坐标 x 和 y ）。随后，输出分为力学子问题和热学子问题中的场变量。对于力学场，输出为位移向量的分量以及应力张量（即 u_x、 u_y、 σ_{xx}、 σ_{xy}和 σ_{yy}）。对于热场，输出层包括标量温度值以及热通量的分量（即 T、 q_x和 q_y）。需要指出的是，我们计划针对每个单独的输出变量使用独立的全连接前馈神经网络（FFNN）（另见图4及 [7]中的说明）。每个神经网络的结构采用标准形式，可分为一个输入层、若干个隐藏层（可选）以及输出层。各层相互连接，以传递信息 [53]。每一层内部的神经元之间均无连接。因此，我们通过向量 $\boldsymbol{z}^l$表示从 $l-1$ 层到 l 的信息旁路。向量 $\boldsymbol{z}^l$ 的每个分量由

$$z_m^l = a(\sum_{n=1}^{N_l} w_{mn}^l z_n^{l-1} + b_m^l), \quad l = 1, \ldots, L. \tag{17}$$

在式（17）中， z_n^{l-1}是第 $l-1$层中的第 n个神经元。分量 w_{mn} 表示第 $l-1$ 层的第 n个神经元与第l层的第 m个神经元之间的连接权重。第 l个隐藏层中的每个神经元都有一个偏置变量 b_m^l。数值 N_l对应第 l个隐藏层中的神经元数量。隐藏层总数为 L。字母 a 表示每个神经元中的激活函数。激活函数 $a(\cdot)$通常是一个非线性函数。本研究采用双曲正切函数，其定义如下

$$\tanh(x) = \frac{\exp^x - \exp^{-x}}{\exp^x + \exp^{-x}}. \tag{18}$$

激活函数的适当选择取决于具体问题，应基于针对网络最终性能开展的超参数研究来确定。另请参阅 [7],[54]和 [55]中的讨论。

所述热-力问题的物理信息神经网络输入层由位置向量 x 和 y 坐标组成。对于输出层，力学场中包含 u_x、 u_y、 σ_{xx}、 σ_{xy}和 σ_{yy}，热场中包含 T、 q_x和 q_y。本文聚焦于二维情形，拓展到三维情形并不困难。接下来将讨论网络架构的其余部分。网络的可训练参数集用 $\boldsymbol{\theta} = \{\boldsymbol{W}, \boldsymbol{b}\}$表示，它们分别是神经网络的权重和偏置。将每种神经网络结构视为 $\mathcal{N}$时，网络在力学场和热场中的最终结果为

$$u_x = \mathcal{N}_{u_x}(\boldsymbol{X};\boldsymbol{\theta}),\ u_y = \mathcal{N}_{u_y}(\boldsymbol{X};\boldsymbol{\theta}),\ \sigma_{xx}^o = \mathcal{N}_{\sigma_{xx}}(\boldsymbol{X};\boldsymbol{\theta}),\ \sigma_{xy}^o = \mathcal{N}_{\sigma_{xy}}(\boldsymbol{X};\boldsymbol{\theta}),\ \sigma_{yy}^o = \mathcal{N}_{\sigma_{yy}}(\boldsymbol{X};\boldsymbol{\theta}), \tag{19}$$

$$T = \mathcal{N}_T(\boldsymbol{X};\boldsymbol{\theta}),\ q_x^o = \mathcal{N}_{q_x}(\boldsymbol{X};\boldsymbol{\theta}),\ q_y^o = \mathcal{N}_{q_x}(\boldsymbol{X};\boldsymbol{\theta}). \tag{20}$$

神经网络输出是可训练参数的函数，训练过程通过最小化不同的物理损失函数来完成。接下来，我们通过定义输入层和输出层来构建边值问题。为此，将偏微分方程及其相应的初始条件和/或边界条件表示为神经网络的损失函数。因此，我们需要求出输出层相对于输入层的导数（例如，应力相对于 x方向或 y方向的导数）。当前的深度学习软件包，如 Pytorch [56] 和 Tensorflow [57]，能够基于自动微分算法 [58]计算导数。本文开发的算法采用 SciANN 软件包 [59]实现，但该方法论很容易迁移到其他平台。

将力学相关损失项的总和记为 $\mathcal{L}_M$，其根据一组方程（式 (2)–(8)）定义如下：

$$\mathcal{L}_M = \underbrace{\mathcal{L}_M^{EF} + \mathcal{L}_M^{DBC}}_{\text{based on } \boldsymbol{u}} + \mathcal{L}_M^{cnc} + \underbrace{\mathcal{L}_M^{SF} + \mathcal{L}_M^{NBC}}_{\text{based on } \boldsymbol{\sigma}^o}. \tag{21}$$

在式 (21) 中，通过位移向量 $\boldsymbol{u}$ 的输出，最小化问题的能量形式 ($\mathcal{L}_M^{EF}$) 和给定的狄利克雷边界条件损失 ($\mathcal{L}_M^{DBC}$)。对于诺伊曼边界条件 ($\mathcal{L}_M^{NBC}$) 以及要求计算一阶位移导数的问题强形式 ($\mathcal{L}_M^{SF}$)，相关损失项应用于 $\boldsymbol{\sigma}^o$，

$$\mathcal{L}_M^{EF} = \text{MAE}_M^{EF}\left(\int_\Omega \frac{1}{2}\hat{\boldsymbol{\varepsilon}}_e^T\, \hat{C}(x,y)\, \hat{\boldsymbol{\varepsilon}}_e \; dV - \int_{\Gamma_N} \boldsymbol{u}^T\, \bar{\boldsymbol{t}}\; dA\right), \tag{22}$$

$$\mathcal{L}_M^{DBC} = \text{均方误差}_M^{DBC}(\boldsymbol{u} - \bar{\boldsymbol{u}}), \tag{23}$$

$$\mathcal{L}_M^{cnc} = \text{MSE}_M^{cnc}(\boldsymbol{\sigma}^o - \boldsymbol{\sigma}), \tag{24}$$

$$\mathcal{L}_M^{SF} = \text{MSE}_M^{SF}(\text{div}(\boldsymbol{\sigma}^o)), \tag{25}$$

$$\mathcal{L}_M^{NBC} = \text{MSE}_M^{NBC}\left(\boldsymbol{\sigma}^o \cdot \boldsymbol{n} - \bar{\boldsymbol{t}}\right). \tag{26}$$

损失函数　　与热场相关的项也根据式（9）—（16）计算。

与力学损失类似，热损失 $\mathcal{L}_T$ 由以下部分组成：$\mathcal{L}_T^{EF}$（热扩散问题的能量形式）、$\mathcal{L}_T^{DBC}$（狄利克雷边界条件）、$\mathcal{L}_T^{NBC}$（诺伊曼边界条件）、$\mathcal{L}_T^{SF}$（扩散问题的强形式）以及 $\mathcal{L}_T^{cnc}$（连接项），其表达式为

$$\mathcal{L}_T = \underbrace{\mathcal{L}_T^{EF} + \mathcal{L}_T^{DBC}}_{\text{based on } T} + \mathcal{L}_T^{cnc} + \underbrace{\mathcal{L}_T^{SF} + \mathcal{L}_T^{NBC}}_{\text{based on } \boldsymbol{q}^O}, \tag{27}$$

$$\mathcal{L}_T^{EF} = \text{MAE}_T^{EF}\left(\int_\Omega \frac{1}{2}\boldsymbol{q}^T \cdot \nabla T\, dV + \int_{\bar{\Gamma}} (\boldsymbol{q}\cdot\boldsymbol{n})\; T\, dA\right), \tag{28}$$

$$\mathcal{L}_T^{DBC} = \text{MSE}_T^{DBC}\left(T - \bar{T}\right), \tag{29}$$

$$\mathcal{L}_T^{cnc} = \text{MSE}_T^{cnc}(\boldsymbol{q}^o - \boldsymbol{q}), \tag{30}$$

$$\mathcal{L}_T^{SF} = \text{均方误差}_T^{SF}(\text{div}(\boldsymbol{q}^o)), \tag{31}$$

$$\mathcal{L}_T^{NBC} = \text{MSE}_T^{NBC}\left(\boldsymbol{q}^o \cdot \boldsymbol{n} - \bar{\boldsymbol{q}}\right). \tag{32}$$

为完整起见，均方误差和平均绝对误差定义如下：

$$\text{MSE}(\bullet)_{type} = \frac{1}{k_{type}}\sum_{i=1}^{k_{type}} (\bullet)^2, \quad \text{MAE}(\bullet)_{type} = \frac{1}{k_{type}}\sum_{i=1}^{k_{type}} |\bullet|. \tag{33}$$

数学优化问题写作:

$$\boldsymbol{\theta}^* = \arg\min_{\boldsymbol{\theta}} \mathcal{L}(\boldsymbol{X}; \boldsymbol{\theta}), \tag{34}$$

其中， $\boldsymbol{\theta}^*$ 是网络的最优可训练参数（权重和偏置），由该优化问题求得。

注 1 式（22）和式（28）中的能量形式是在全局层面计算的，所有配点共用一个损失项。其他损失项则在每个配点处分别最小化。正如 [7], 中所讨论的，能量损失采用平均绝对误差（MAE）有助于使其数值保持在与其他损失项相同的量级。

注 2 应当注意，网络的应力输出用于强形式损失和诺伊曼边界条件损失。因此，不会出现混合公式（例如 Hellinger-Reissner原理）中可能由鞍点约束引起的问题。

我们研究了两种不同的训练算法，以求解耦合方程组。一种方法采用所谓的序贯训练，另一种方法则采用图1所示的耦合训练。序贯训练算法通过最小化单个物理场的损失来实现，同时冻结其他物理参数。也欢迎读者参阅 Haghighat 等人 [60]，了解更多详情。序贯训练的流程与交错数值方案 [33, 37, 61]类似。对于序贯训练，我们定义以下损失项

$$\mathcal{L}_{Seq1} = \mathcal{L}_{\text{Phy1}} = \mathcal{L}_T, \tag{35}$$
$$\mathcal{L}_{Seq2} = \mathcal{L}_{\text{Phy2}} = \mathcal{L}_M. \tag{36}$$

在最小化式 (35) 中的损失函数时，与力学场相关的输出（即 u_x， u_y， σ^o_{xx}， σ^o_{xy}，以及 σ^o_{yy}）保持冻结。在最小化式 (36) 中的损失函数时，与热场相关的输出（即 T， q^o_x，以及 q^o_y）保持冻结。另一方面，我们采用耦合训练，同时最小化所有损失项。

对于耦合训练，总损失函数是所有损失项的总和，写作

$$\mathcal{L}_{Cpl} = \mathcal{L}_{\text{Phy1}} + \mathcal{L}_{\text{Phy2}} = \mathcal{L}_M + \mathcal{L}_T. \tag{37}$$

图 4 展示了用于求解热弹性问题的混合物理信息神经网络公式。除了位置向量（x, y）外，材料参数 C(x， y) 和 k(x， y) 也被添加到输入层。这样，网络便能够学习其他材料属性对应的解。

在所提出的网络结构中，主变量及其空间导数可直接求得。例如，对于力学问题，位移向量 $\boldsymbol{u}$ 和应力张量 $\boldsymbol{\sigma}^o$ 被定义为网络输出。此外，根据方程（1）和（2），利用预测的变形（u_x 和 u_y）以及预测温度 T，计算应变张量 $\boldsymbol{\varepsilon}$。将计算得到的应变张量（$\boldsymbol{\varepsilon}$）代入方程（4），即可计算耦合方程组中的应力张量 $\boldsymbol{\sigma}$ 。因此，由所求位移 $\boldsymbol{\sigma}$ 得到的应力张量通过方程（24）中的连接损失项 $\mathcal{L}_M^{cnc}$与网络的直接求值 $\boldsymbol{\sigma}^o$ 相连接。后者是一种软约束，用于将 $\boldsymbol{\sigma}^o$ 与 $\boldsymbol{\sigma}$ 值连接起来。

热场采用了类似的策略，其中根据方程（12）推导热流向量 $\boldsymbol{q}$；在该方程中，通过自动微分计算热场梯度。 $\boldsymbol{q}$ 通过方程（30）中所述的 $\mathcal{L}_T^{cnc}$ 与 $\boldsymbol{q}^o$ 相连接。

根据图4，序贯训练过程中的优化问题首先从第一个目标场损失函数开始，以实现对第一个物理场损失函数的最小化。

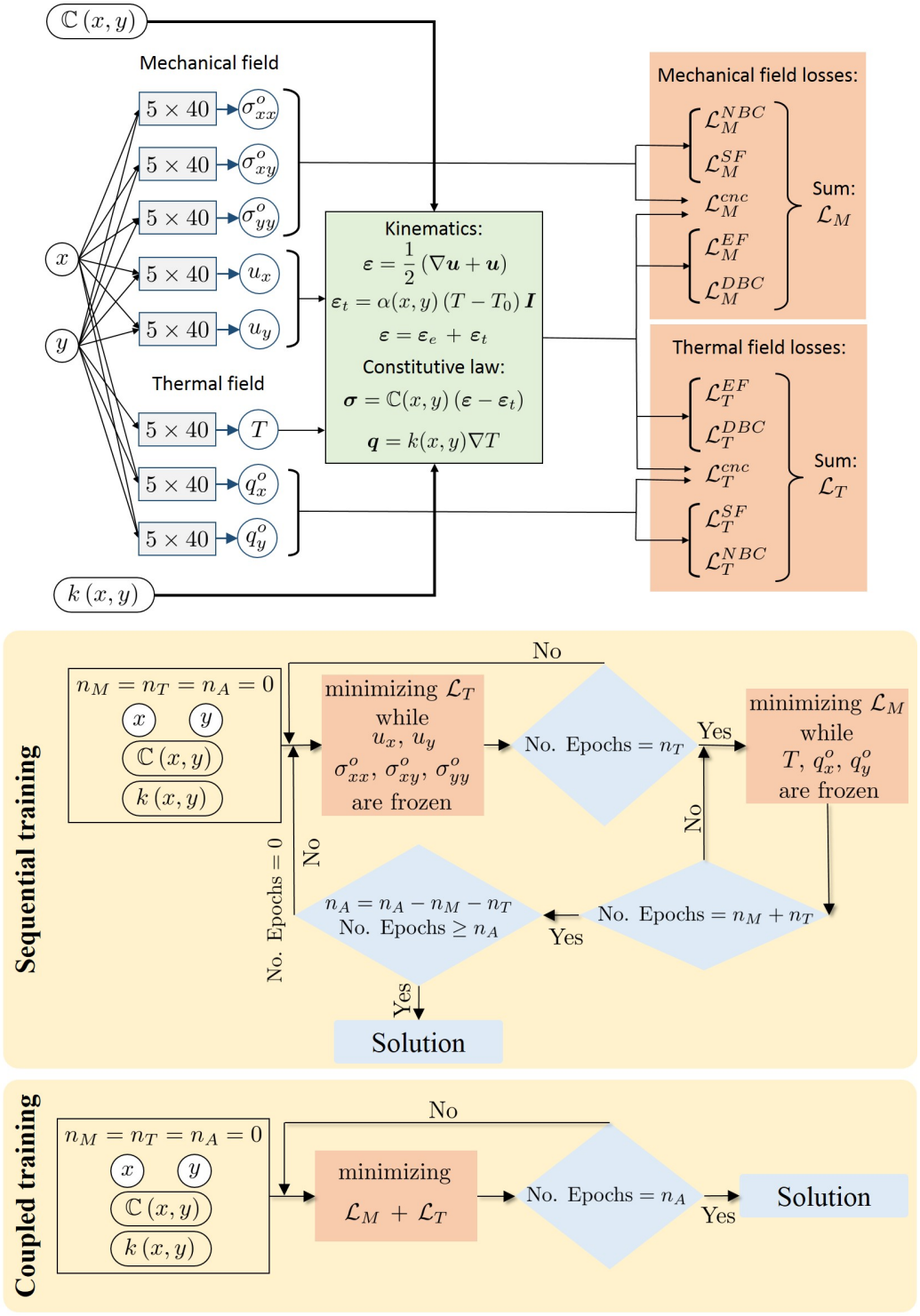


图 4：多物理场物理信息神经网络的网络架构和损失函数，以及热弹性问题的序贯训练和耦合训练流程图。

在热弹性问题中，训练首先通过寻找最优网络参数来满足热损失函数（$\mathcal{L}_T$）。这一选择基于以下事实：热诱导应变张量会影响力学场的结果。在我们提出的算法中，热场损失函数（第一个物理损失函数）的训练在达到一定训练轮数 n_T 后停止。随后，针对下一个物理损失函数（力学场损失函数 $\mathcal{L}_m$）进行训练，训练轮数可以相同，也可以不同（ n_M）。达到所需的总训练轮数 n_A 后，整个训练过程将停止。如果尚未达到所需的总训练轮数，则继续采用相同的过程，直至达到该值（最小化 $\mathcal{L}_T$ ，训练 n_T 轮；最小化 $\mathcal{L}_M$ ，训练 n_M 轮）。对于耦合训练过程，总损失函数通过将不同的物理场损失函数（$\mathcal{L}_M$ 和 $\mathcal{L}_T$）相加，合并为一个损失函数 $\mathcal{L}_{Cpl}$，见公式（37）和图4。

注释 3 研究表明，采用分离网络，并且在网络的损失函数中仅使用一阶导数，能够获得更好的性能 [7]。关于这一问题，也可参见 Chiu 等人 [62] 的研究。降低导数的阶数还有其他好处。例如，这样在选择激活函数时更加灵活（即可以避免高阶导数导致的梯度消失问题）。此外，与采用二阶求导的情况相比，计算成本也更低。

注释 4 在序贯训练中，停止每个训练过程的另一种方法是，使每个场的损失达到预期的特定值。实现这一点的方法是，在优化每个场的损失函数时，将训练轮数设为无限。只要达到预期损失函数的特定值，训练过程就会停止。

注释 5 当前的混合物理信息神经网络方法也可以解释为一个多目标优化问题。因此，不同损失项之间保持适当的平衡至关重要。在多物理场问题中，通常需要在开始最小化之前对各个量进行归一化处理 [23]。此外，损失项 L_T^{EF}和 L_M^{EF}根据绝对误差的均值定义，以使其与其他损失项处于相同量级。

本文采用 Adam 优化器，相关网络参数总结于表 1。请注意，对于不包含问题能量形式的简单 PINN 表述，可以使用多个批次来训练网络。

表 1：耦合和顺序训练过程的网络参数汇总 .

参数	值
输入、输出	$\boldsymbol{X}$, ($\boldsymbol{u}$, $\boldsymbol{\sigma}^o$, T, $\boldsymbol{q}^o$)
激活函数	tanh
层数和每层神经元数（L, N_l）	（5, 40）
批大小	混合物理信息神经网络公式的全批次
（学习率 α，训练轮数）	(10^{-3}, 10_5)
耦合训练	
总训练轮数（n_A）	10万训练轮次
序贯训练	
总迭代次数（n_A）	20万训练轮次
切换前热场的迭代次数（n_T）	2万训练轮次
切换前力学场的迭代次数（n_M）	2万训练轮次

根据图 3 中讨论的边界条件，下面针对给定的边值问题更具体地写出所有损失项。总应变张量和温度梯度向量的各个分量分别通过对位移场和温度场求导计算得到。根据式（19）和（20），我们定义以下网络

$$\mathcal{N}_{\varepsilon_x}(\boldsymbol{X};\boldsymbol{\theta}) = \partial_x \mathcal{N}_{u_x}, \tag{38}$$
$$\mathcal{N}_{\varepsilon_y}(\boldsymbol{X};\boldsymbol{\theta}) = \partial_y \mathcal{N}_{u_y}, \tag{39}$$
$$\mathcal{N}_{\varepsilon_{xy}}(\boldsymbol{X};\boldsymbol{\theta}) = \frac{1}{2}\left(\partial_y \mathcal{N}_{u_x} + \partial_x \mathcal{N}_{u_y}\right), \tag{40}$$
$$\mathcal{N}_{\nabla_x T}(\boldsymbol{X};\boldsymbol{\theta}) = \partial_x \mathcal{N}_T, \tag{41}$$
$$\mathcal{N}_{\nabla_y T}(\boldsymbol{X};\boldsymbol{\theta}) = \partial_y \mathcal{N}_T. \tag{42}$$

随后可推导出（各向同性）热应变张量的分量：

$$\mathcal{N}_{\varepsilon_{t_x}}(\boldsymbol{X};\boldsymbol{\theta}) = \mathcal{N}_{\varepsilon_{t_y}} = \alpha\left(\boldsymbol{X}\right)\left(\mathcal{N}_T - T_0\right). \tag{43}$$

考虑总应变张量和热致应变张量，弹性部分的应变张量可表示为：

$$\mathcal{N}_{\varepsilon_{e_x}}(\boldsymbol{X};\boldsymbol{\theta}) = \mathcal{N}_{\varepsilon_x} - \mathcal{N}_{\varepsilon_{t_x}}, \tag{44}$$
$$\mathcal{N}_{\varepsilon_{e_y}}(\boldsymbol{X};\boldsymbol{\theta}) = \mathcal{N}_{\varepsilon_y} - \mathcal{N}_{\varepsilon_{t_y}}, \tag{45}$$
$$\mathcal{N}_{\varepsilon_{e_{xy}}}(\boldsymbol{X};\boldsymbol{\theta}) = \mathcal{N}_{\varepsilon_{xy}}. \tag{46}$$

利用本构关系（胡克定律）以及傅里叶定律，可根据以下关系计算应力张量和热通量矢量的分量

$$\mathcal{N}_{\sigma_x}(\boldsymbol{X};\boldsymbol{\theta}) = \frac{E(\boldsymbol{X})}{(1+\nu(\boldsymbol{X}))(1-2\nu(\boldsymbol{X}))}\left((1-\nu(\boldsymbol{X}))\mathcal{N}_{\varepsilon_{e_x}} + \nu(\boldsymbol{X})\mathcal{N}_{\varepsilon_{e_y}}\right), \tag{47}$$
$$\mathcal{N}_{\sigma_y}(\boldsymbol{X};\boldsymbol{\theta}) = \frac{E(\boldsymbol{X})}{(1+\nu(\boldsymbol{X}))(1-2\nu(\boldsymbol{X}))}\left(\nu(\boldsymbol{X})\mathcal{N}_{\varepsilon_x} + (1-\nu(\boldsymbol{X}))\mathcal{N}_{\varepsilon_y}\right), \tag{48}$$
$$\mathcal{N}_{\sigma_{xy}}(\boldsymbol{X};\boldsymbol{\theta}) = \frac{E(\boldsymbol{X})}{2(1+\nu(\boldsymbol{X}))}\mathcal{N}_{\varepsilon_{e_{xy}}}, \tag{49}$$
$$\mathcal{N}_{q_x}(\boldsymbol{X};\,\boldsymbol{\theta}) = -k(\boldsymbol{X})\mathcal{N}_{\nabla_x T}, \tag{50}$$
$$\mathcal{N}_{q_y}(\boldsymbol{X};\,\boldsymbol{\theta}) = -k(\boldsymbol{X})\mathcal{N}_{\nabla_y T}. \tag{51}$$

为简化损失函数的构造，以下不显示神经网络输出 $\mathcal{N}(\boldsymbol{X};\boldsymbol{\theta})$ 对配点坐标 $\boldsymbol{X}$ 以及可训练参数$\boldsymbol{\theta}$ 的依赖关系。

最后，下面将展开用于计算力学场和热场损失函数的数学表达式。另请参见公式（22）至（32）以及图4,

$$\mathcal{L}_M^{EF} = \left| -\frac{1}{2N_b} \sum_{\boldsymbol{X}\in\Omega} \left(\mathcal{N}_{\sigma_x}\mathcal{N}_{\varepsilon_{ex}} + \mathcal{N}_{\sigma_y}\mathcal{N}_{\varepsilon_{ey}} + \mathcal{N}_{\sigma_{xy}}\mathcal{N}_{\varepsilon_{exy}} \right) \right|, \tag{52}$$

$$\mathcal{L}_M^{DBC} = \frac{1}{N_{DBC}^L} \sum_{\boldsymbol{X}\in\Gamma_D^L} \left(\mathcal{N}_{u_x} - 0\right)^2 + \frac{1}{N_{DBC}^R} \sum_{\boldsymbol{X}\in\Gamma_D^R} \left(\mathcal{N}_{u_x} - 0\right)^2, + \frac{1}{N_{DBC}^T} \sum_{\boldsymbol{X}\in\Gamma_D^T} \left(\mathcal{N}_{u_y} - 0\right)^2 + \frac{1}{N_{DBC}^B} \sum_{\boldsymbol{X}\in\Gamma_D^B} \left(\mathcal{N}_{u_y} - 0\right)^2, \tag{53}$$

$$\mathcal{L}_M^{cnc} = \frac{1}{N_b} \sum_{\boldsymbol{X}\in\Omega} \left(\left(\mathcal{N}_{\sigma_x^o} - \mathcal{N}_{\sigma_x}\right)^2 + \left(\mathcal{N}_{\sigma_y^o} - \mathcal{N}_{\sigma_y}\right)^2 + \left(\mathcal{N}_{\sigma_{xy}^o} - \mathcal{N}_{\sigma_{xy}}\right)^2 \right), \tag{54}$$

$$\mathcal{L}_M^{SF} = \frac{1}{N_b} \sum_{\boldsymbol{X}\in\Omega} \left(\left(\partial_x \mathcal{N}_{\sigma_x^o} + \partial_y \mathcal{N}_{\sigma_{xy}^o}\right)^2 + \left(\partial_y \mathcal{N}_{\sigma_y^o} + \partial_x \mathcal{N}_{\sigma_{xy}^o}\right)^2 \right), \tag{55}$$

$$\mathcal{L}_M^{NBC} = 0. \tag{56}$$

在上述方程中， Γ^L、 Γ^R、 Γ^T和 Γ^B分别表示左、右、上、下边界上各点的位置。DBC和NBC分别表示狄利克雷边界条件和诺伊曼边界条件。

请注意，根据所定义的边值问题，与诺伊曼边界条件相关的损失项（式（56））为零。我们没有对材料的微观结构施加任何外部牵引力。计算能量形式中的积分时，由于配点在区域内均匀分布，我们对配点进行等权求和。因此， N_b 表示区域（物体）内配点的数量。此外，N^L、 N^R、 N^T和 N^B分别表示左、右、上、下边界上的配点数量。另请参阅 [31] ，了解其他可能的方法。

与热场相关的损失函数如下:

$$\mathcal{L}_T^{EF} = \left| \frac{1}{2N_b} \sum_{\boldsymbol{X}\in\Omega} \left(\mathcal{N}_{q_x}\mathcal{N}_{\nabla_x T} + \mathcal{N}_{q_y}\mathcal{N}_{\nabla_y T} \right) + \frac{1}{N_{NBC}^L} \sum_{\boldsymbol{X}\in\Gamma_N^L} \left(\mathcal{N}_{q_x}\mathcal{N}_T \right) \right|, \tag{57}$$

$$\mathcal{L}_T^{DBC} = \frac{1}{N_{DBC}^L} \sum_{\boldsymbol{X}\in\Gamma_D^L} \left(\mathcal{N}_T - 1\right)^2 + \frac{1}{N_{DBC}^R} \sum_{\boldsymbol{X}\in\Gamma_D^R} \left(\mathcal{N}_T - 0\right)^2, \tag{58}$$

$$\mathcal{L}_T^{cnc} = \frac{1}{N_b} \sum_{\boldsymbol{X}\in\Omega} \left(\left(\mathcal{N}_{q_x^o} - \mathcal{N}_{q_x}\right)^2 + \left(\mathcal{N}_{q_y^o} - \mathcal{N}_{q_y}\right)^2 \right), \tag{59}$$

$$\mathcal{L}_T^{SF} = \frac{1}{N_b} \sum_{\boldsymbol{X}\in\Omega} \left(\partial_x \mathcal{N}_{q_x^o} + \partial_y \mathcal{N}_{q_y^o}\right)^2, \tag{60}$$

$$\mathcal{L}_T^{NBC} = \sum_{\boldsymbol{X}\in\Gamma_N^T} \frac{1}{N_{NBC}^T} \left(\mathcal{N}_{q_y^o} - 0\right)^2 + \sum_{\boldsymbol{X}\in\Gamma_N^B} \frac{1}{N_{NBC}^B} \left(\mathcal{N}_{q_y^o} - 0\right)^2. \tag{61}$$

4. 结果

4.1. 基体中的圆形非均质体（几何体*1*）

我们从图2所示的第一个几何结构开始。配点的设置如图5左侧所示。如 [7], 中所述，为提高结果的准确性，需要增加非均质区域、区域边界和基体中的配点数量。在x方向和 y方向上，范围均为0至 1 [毫米] 的矩形区域内约设置5000个配点。在这些配点处对损失函数（见第3.1节）进行最小化。当前边值问题的材料参数见表2。

	数值/单位
几何体1的热弹性问题	
基体杨氏模量和泊松比（E_{mat}, ν_{mat}）	（0.3 吉帕，0.3）
夹杂杨氏模量和泊松比（E_{inc}, ν_{inc}）	（1.0 吉帕，0.3）
基体热导率（k_{mat}）	0.3 瓦每米开尔文
夹杂物热导率（k_{inc}）	1.0 瓦每米开尔文

表2：线性热弹性问题中几何体1的输入参数

为评估网络响应，使用有限元程序 FEAP [63] 对相同区域的热弹性问题也进行了求解。有限元分析中的区域离散化采用这样的方式：单元与用于评估训练后网络响应的规则配点重合。此外，在图5中，颜色变化显示了 PINNs 配点处不同的杨氏模量和热导率值。这是因为有些点位于基体与夹杂的界面上。在有限元方法的组装过程中也采用了相同的处理方式。

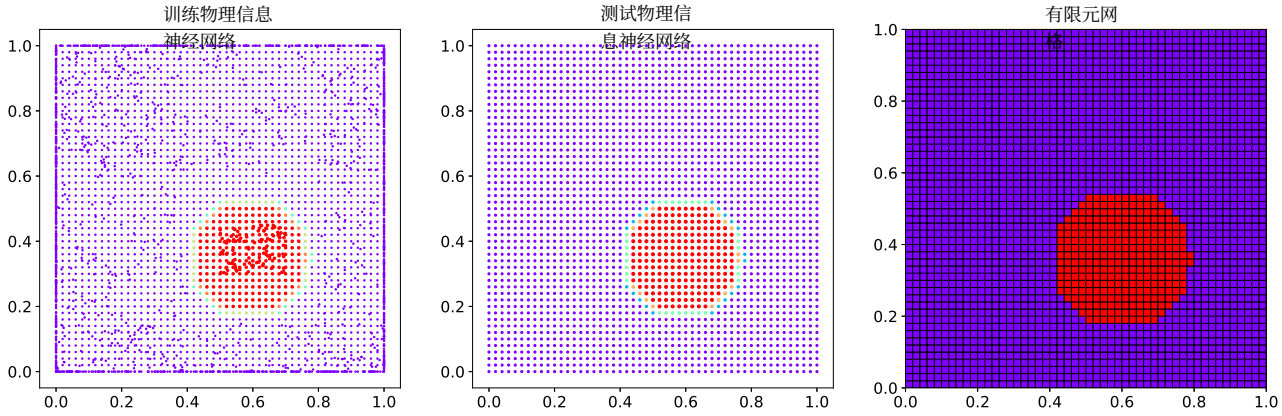


图5：左：用于最小化 PINNs 损失函数的配点。中：用于评估训练后网络响应的点。右：使用有限元法离散化的区域，用于求解相同的热弹性边值问题。颜色变化显示了杨氏模量的不同取值。

损失函数的最小化通过两种不同的方式进行：耦合训练和序贯训练（另见图 4）。两种情况下，训练均在配备 Platinum 8160 型号 Skylake CPU 的 Intel HNS2600BPB硬件节点类型上进行。对于耦合训练，每个损失项通过执行 100 k 训练轮次（迭代次数）进行最小化。为了公平地

比较序贯训练，我们总共执行 200k 训练轮次。在序贯训练中，我们对每个损失项进行 100 k 训练轮次的最小化。训练轮数根据最终损失值和结果的准确度来选择。这个数值可根据优化器类型以及对目标准确度的要求轻松调整。

4.1.1. 几何体*1*的耦合训练训练过程中主要损失项的衰减情况如图6所示。

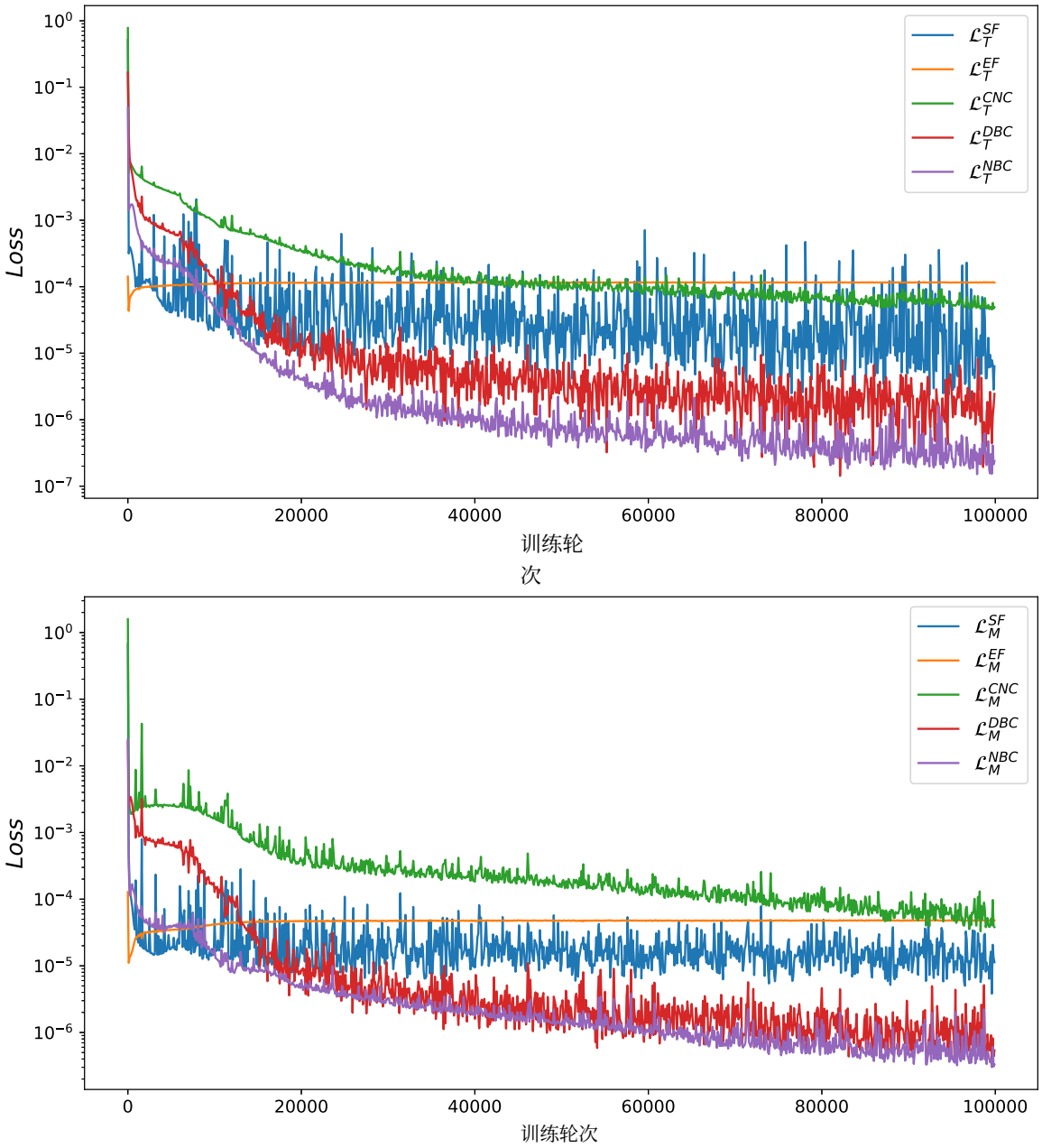


图6：采用耦合训练算法进行训练时，热场和力学场衰减对应的主要损失项。

在耦合训练的最小化过程中，我们试图找到最优的网络参数，以同时满足所有损失项。需要指出的是，为了便于理解并更好地

与序贯训练进行比较，图6分别绘出了与力学子问题和热学子问题相关的损失项。随着训练轮数增加，与控制方程和边界条件相关的所有损失项均逐渐减小。有趣的是，与能量相关的损失项在训练数千轮后基本保持不变。这一现象可能表明，该项对训练过程的贡献并不显著。然而，我们发现，省略这一项会导致结果不如有限元法的结果。相反，我们认为，系统的能量应当最小化，尽管它不一定要在平衡点处为零。这与其他损失项不同，后者是必须满足的直接方程。关于不同方法的更详细比较，读者也可参阅 [7] 。

图7绘出了 PINNs 和有限元法的变形构型。所有变形均由引入的温度梯度引起。为进行更详细的观察，图8对 PINNs 和有限元法结果中的位移及应力分量进行了比较。PINNs 结果列于第一列，而有限元法结果以及 PINNs 与有限元法之间的差异分别显示在第二列和第三列。

图8显示，对于热弹性问题，PINNs 获得的力学场结果与有限元分析结果吻合良好。 x 方向位移的平均相对差异为 0.59%， x方向应力的平均相对差异为 0.87%。尽管位移和应力分量的最大相对差异较高，但这些差异仅出现在区域内靠近夹杂边界的少数点处。在这些区域对配点进行局部加密可以进一步提高精度 [7]。

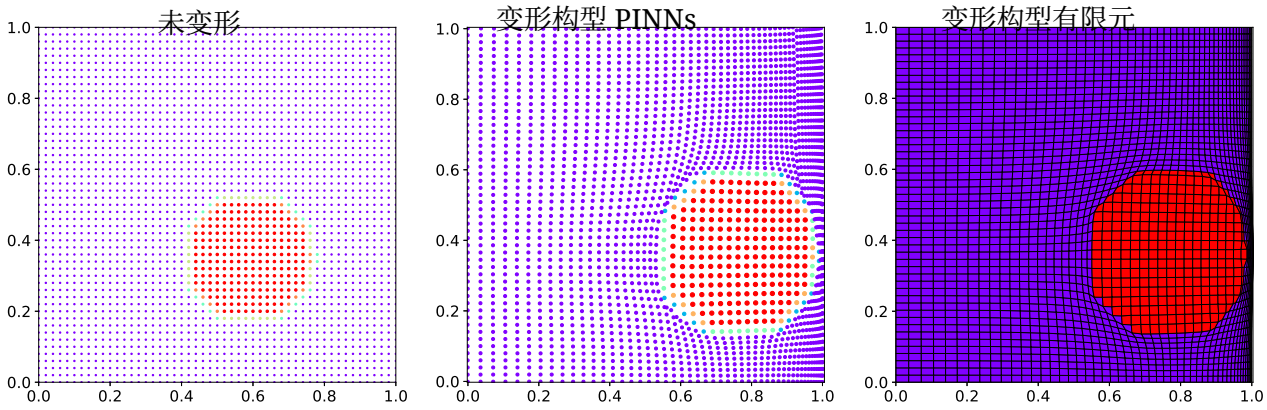


图7：左：物理信息神经网络中评估点的未变形构型。中：采用耦合训练的物理信息神经网络中评估点的变形构型。右：有限元分析中离散区域的变形构型。

最大差异出现在基体与非均质体的界面处。这是因为界面处材料属性的变化相当剧烈。解决这一问题的一种可能方法是在界面附近增加更多配置点，或采用区域分解方法，见 [23, 64, 65]。增加训练轮数也有所帮助。

图9还展示了耦合训练下热弹性问题中物理信息神经网络与有限元分析所得热场结果的一致性。 T的平均相对误差约为 0.3%， q_x 的平均相对误差约为 0.76%。与力学场类似，最大相对差异大多出现在界面上的少数点处。

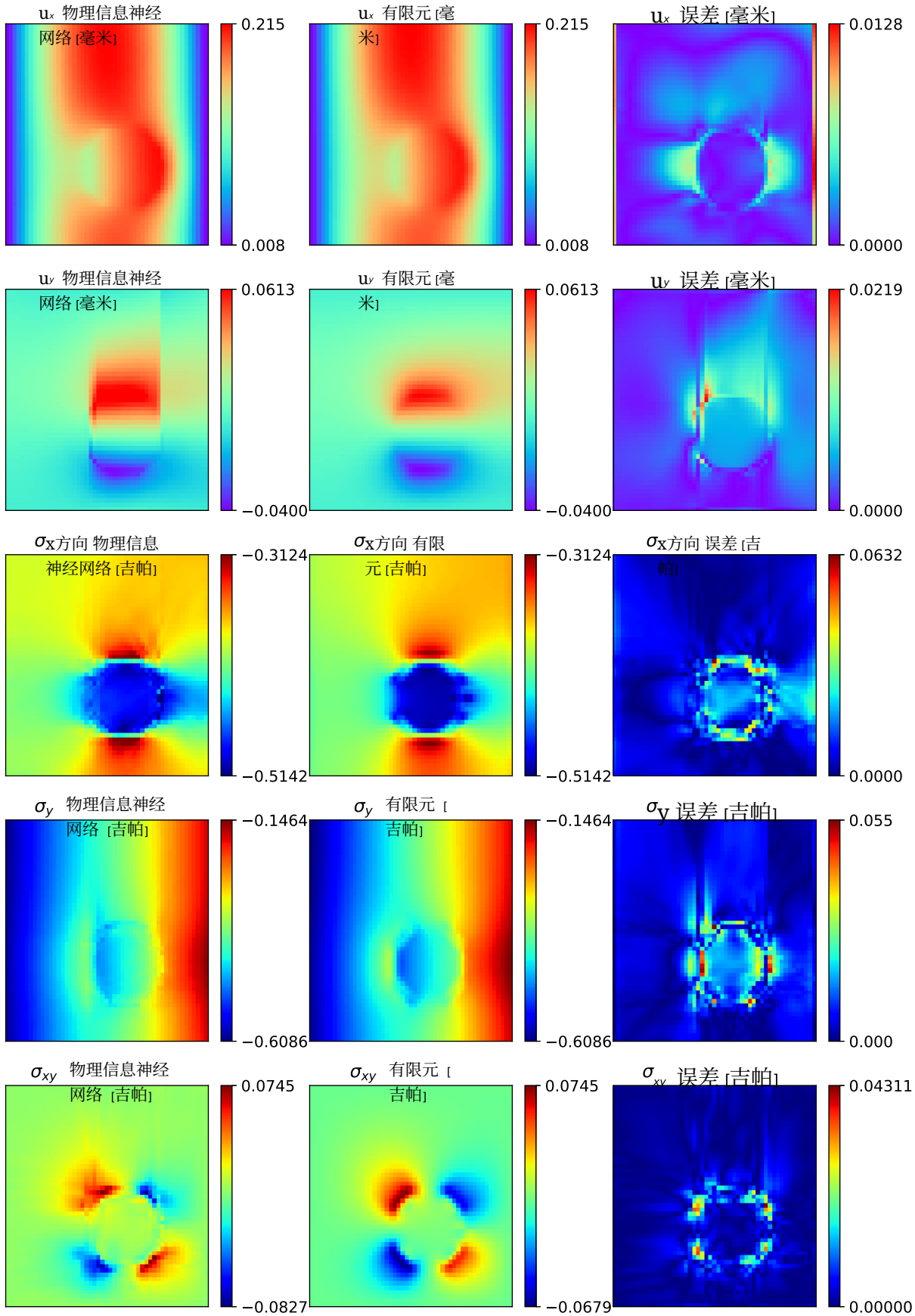


图 8：采用耦合训练过程求解热弹性问题得到的力学场结果。左：物理信息神经网络；中：有限元法；右：有限元法与物理信息神经网络的差值。

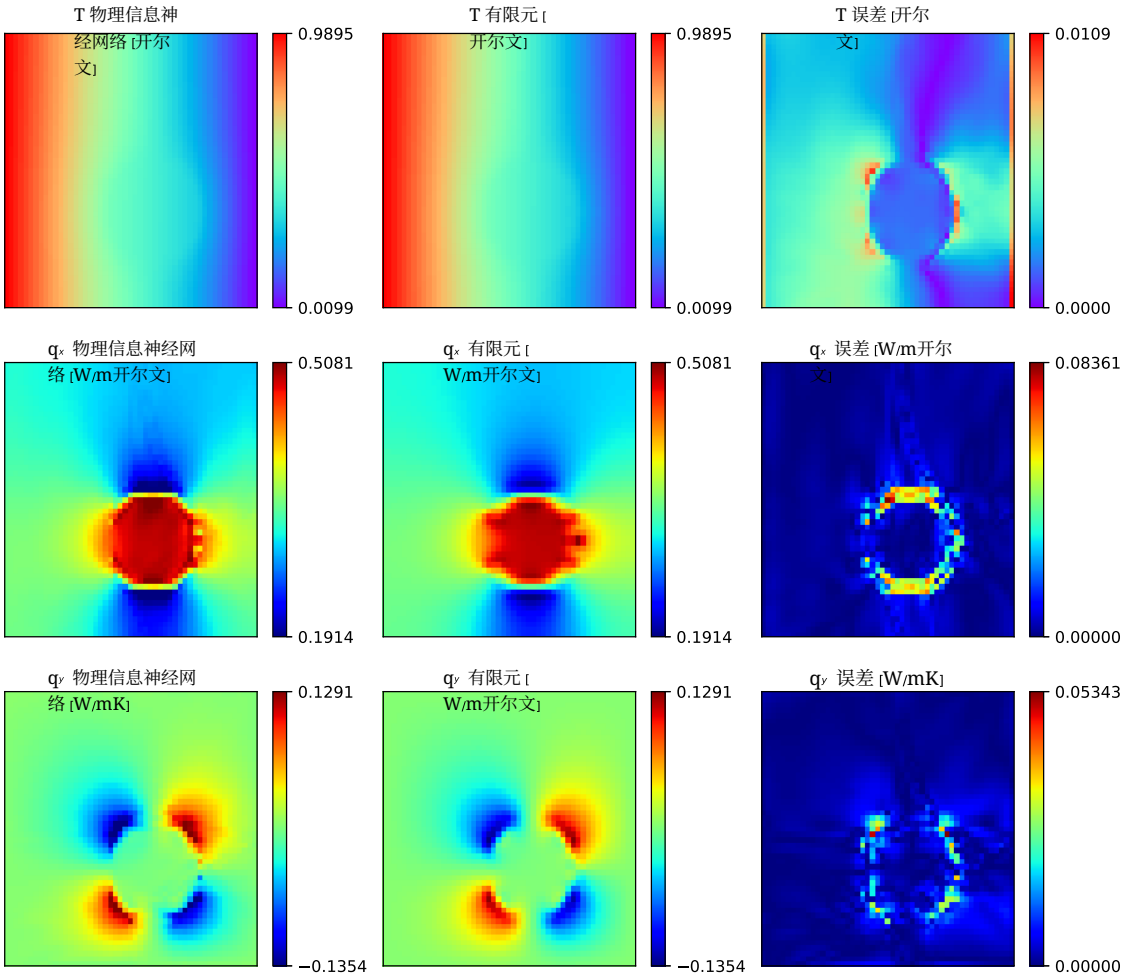


图9：采用耦合训练过程求解热弹性问题得到的热场结果。左：物理信息神经网络；中：有限元法；右：有限元法与物理信息神经网络之间的差异。

4.1.2. 顺序训练

采用顺序算法，使用图5中的相同配点来最小化热场和力学损失函数（$\mathcal{L}_T$ 和 $\mathcal{L}_M$）。如前所述，最小化过程总共进行200 k个训练轮次，因此每个场损失函数对应100 k个训练轮次。图10分别绘出了力学场和热场的主要损失随训练轮次的变化情况。

图10展示了所选 n_T和 n_M的影响。这两个数值决定了热场与力学损失函数最小化之间的切换时机。这两个参数的值设为 20 k。需要注意的是，在力学部分的训练期间不会计算热场的损失项，反之亦然。因此，为了展示完整的损失函数，会在各部分训练期间分别保存其损失，之后再将它们拼接起来。可以看出，对一个场的损失函数进行最小化，可能会导致其他场的损失增加。网络会尝试为当前活跃场寻找最优参数，而这些参数可能会改变先前为非活跃场找到的最优参数。对于每个损失函数而言，这一点在图10中竖线之后不太明显，尤其是热场损失。最后需要指出，在训练过程中交替训练各场若干迭代次数，对于找到系统的全局最小值至关重要。在本例中，n_M、n_T和 n_A^{Seq}的取值使得可以进行五次迭代。

还需要进一步研究，以确定各场之间的最优切换时间，以及最小化一个场损失函数时所需的训练轮次数。

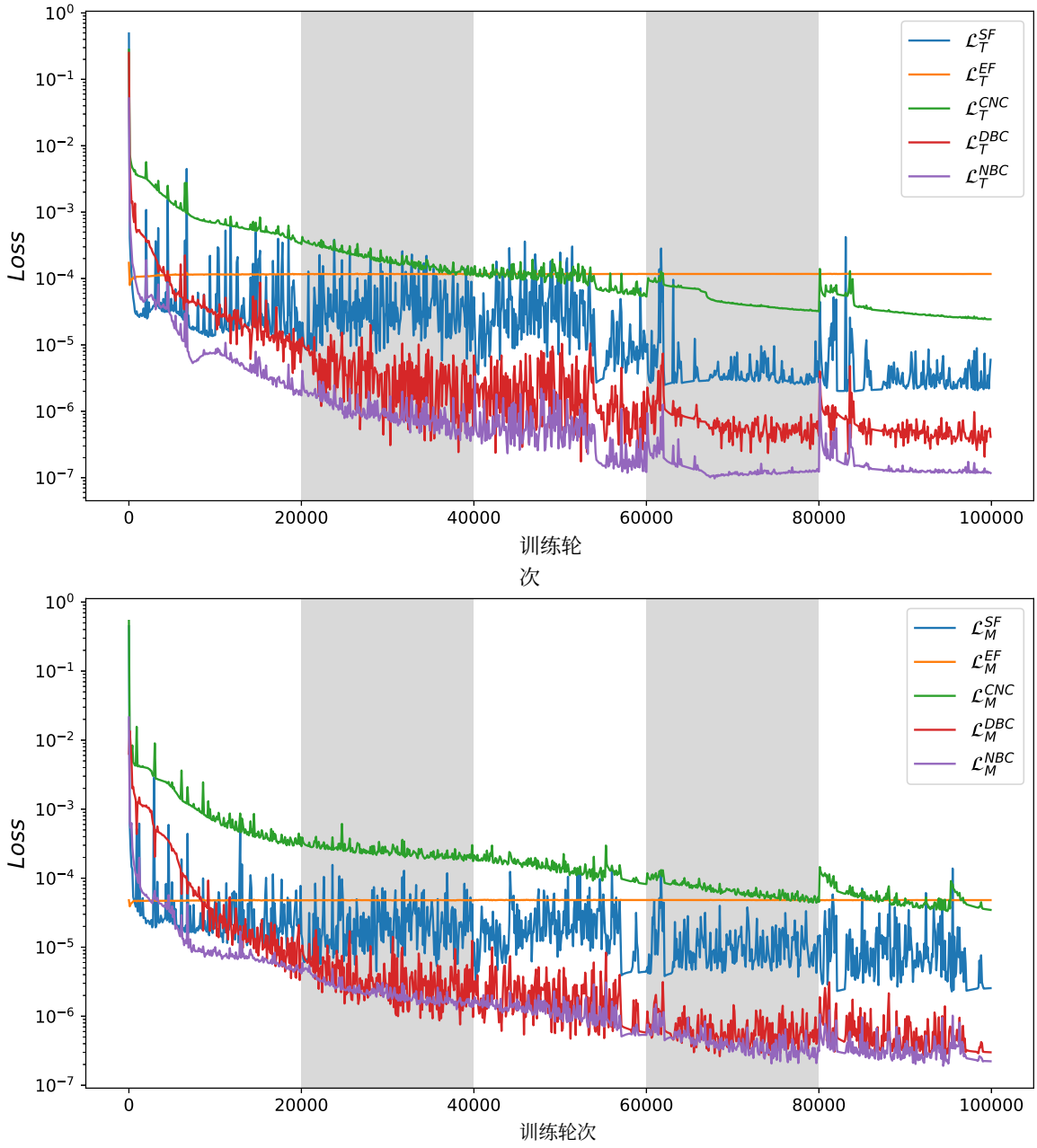


图10：上：使用序贯训练算法进行训练时，热场衰减的主要损失项。下：使用序贯训练算法进行训练时，力学场衰减的主要损失项。

采用序贯训练的混合物理信息神经网络，通过有限元分析得到的 x方向位移和应力分量的平均相对差异分别为 0.81 % 和 0.8 %。在热场中（见图14），温度场和 y方向的热通量 q_y 的平均相对误差分别为 0.26 % 和 0.66 %。

接下来，我们将所提出的混合物理信息神经网络公式与标准物理信息神经网络结构 [17], 以及深度能量法（深度能量法） [30], 进行比较，后者仅使用能量

项来训练神经网络。为进行公平比较，我们使用相同数量的训练轮次和相同的优化器，分别通过标准物理信息神经网络和深度能量法求解热弹性边值问题。值得注意的是，标准物理信息神经网络结构使用独立的神经网络来预测输出变量 $\boldsymbol{u}$ 和 T，因此公式中会出现二阶导数。深度能量法采用相同的网络架构；由于只使用能量项，因此仅需一阶导数。该网络包含五个隐藏层，每层有四十个神经元，并使用双曲正切 激活函数。我们分别采用耦合和序贯训练算法来训练标准物理信息神经网络、深度能量法和混合物理信息神经网络公式。

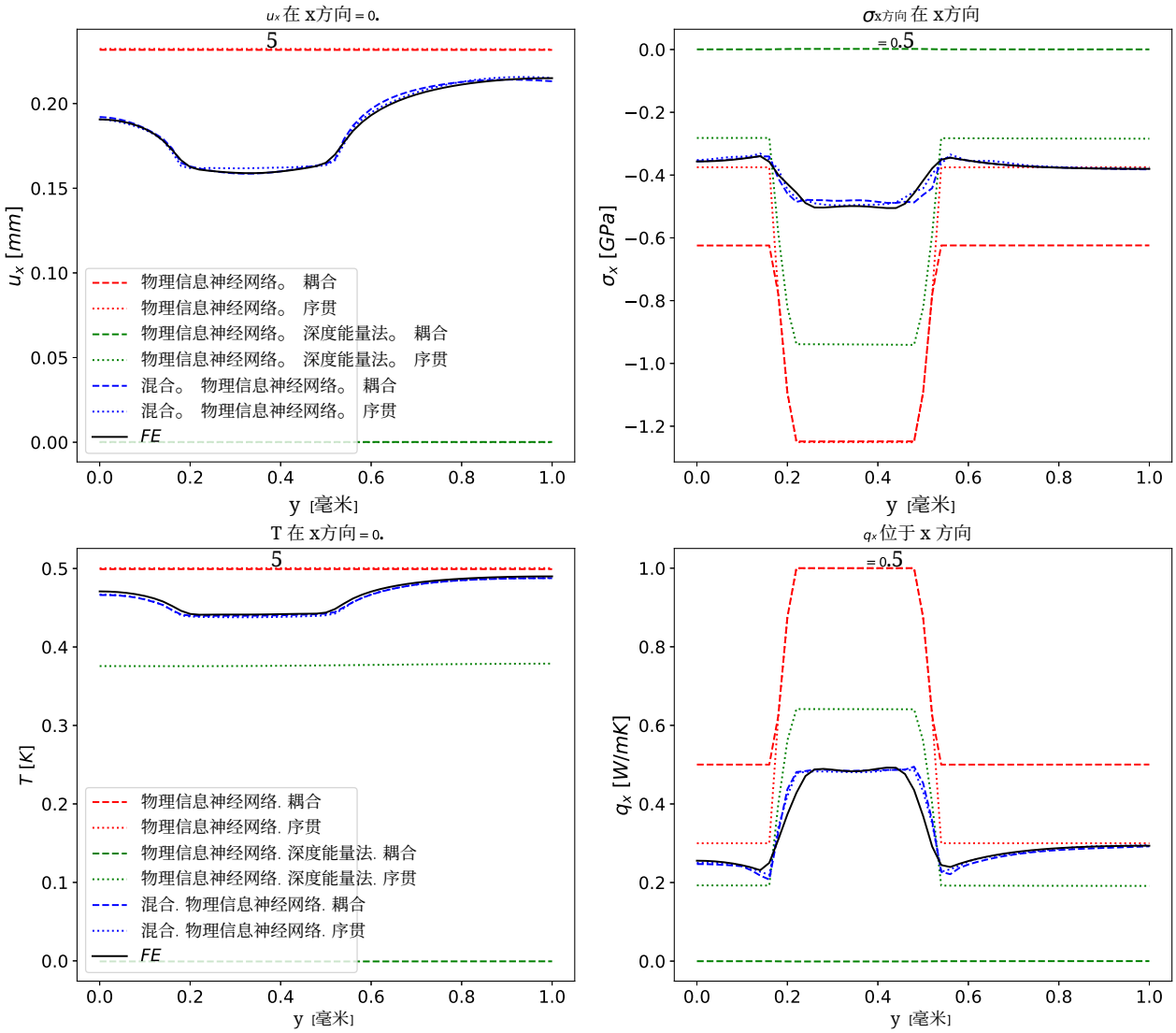


图 11：通过 $x = 0.5$毫米处的截面比较神经网络的主要输出。混合物理信息神经网络：表示本文提出的公式化方法，将 $\boldsymbol{\sigma}$ 和 $\boldsymbol{u}$ 作为输出，并在传统损失函数之外采用能量形式损失函数。物理信息神经网络：表示标准公式化方法，通常仅将位移场 $\boldsymbol{u}$ 和温度场 T 作为输出，并在损失函数中采用控制方程的强形式。相比之下，深度能量法（深度能量法）在损失函数中采用问题的能量形式，同样以 $\boldsymbol{u}$ 和 T 作为输出，但不一定强制满足方程的强形式。耦合：同时最小化所有场损失函数。序贯：采用所提出的序贯训练方法进行最小化。

图11展示了在 $x = 0.5\,\text{mm}$处截取的截面上，位移、应力和热通量的 x分量以及温度场。图中还添加了有限元分析结果作为参考解。

标准物理信息神经网络公式无法捕捉位移分量 u_x 和温度场 T的非均质性影响。此外，无论采用耦合训练还是序贯训练过程，其在预测 q_x 和 σ_x方面的表现都较差。然而，对于应力和热通量分量，序贯训练确实能取得略好的结果。对于深度能量法，也可以观察到类似现象：其结果优于标准物理信息神经网络，但仍不及我们使用有限元得到的参考解，这表明这些方法在异质域中的局限性。相比之下，混合物理信息神经网络公式表现出优异得多的性能，并且能够通过耦合训练算法和顺序训练算法捕捉非均质性的影响。进一步考察结果可以发现，耦合训练算法的性能略优于顺序训练算法。此外，无论采用何种方法来最小化所有损失项，序贯训练的性能都优于耦合训练。总体而言，我们的研究结果表明，在准确性和计算效率方面，混合物理信息神经网络公式优于标准物理信息神经网络和深度能量法，因此是求解复杂物理问题的一种有前景的方法。后者不仅通过仅基于一阶导数构建损失函数来实现，还通过同时采用问题的强形式和弱形式，并利用独立的神经网络预测输出而实现。

备注 6 耦合训练的每轮平均耗时约为 0.67 秒。对于序贯训练，该值降至 0.16 秒。然而，由于序贯训练至少需要进行两次迭代才能得到解，因此其计算时间节省约为 0.35 秒，与耦合训练方法相比，计算时间缩短了两倍以上。

4.2. 包含矩形夹杂物的复杂微观结构（几何体*2*）

为了进一步评估所提出方法求解非均质域中多物理场问题的性能，现考虑图12中的第二种几何结构。该
这种非均质性模式被定义为更具挑战性，具有尖角。这种模式也对应于某些特定的材料微观结构，例如镍合金体系，见 [66]。

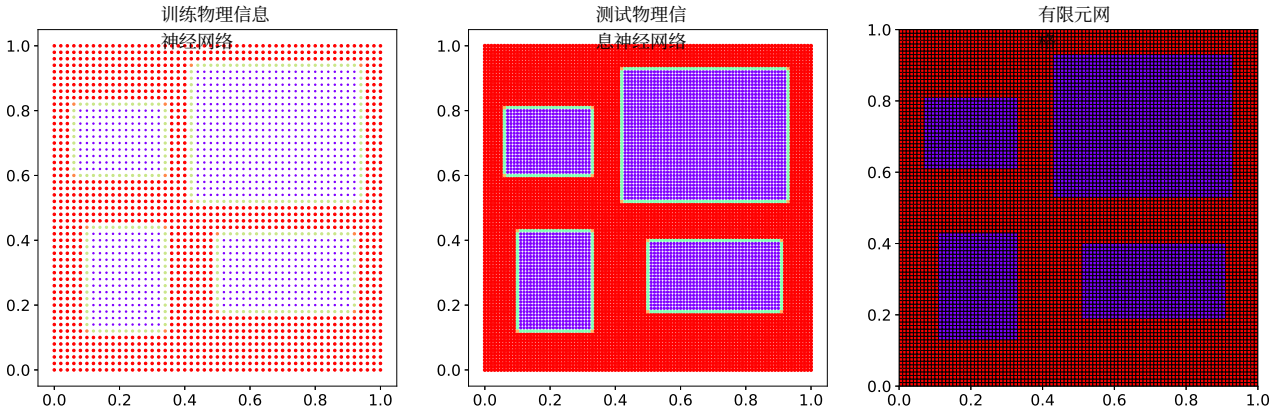


图 12：左：用于最小化第二种几何结构的损失函数的配点。中：用于评估训练后的网络响应的点。右：使用有限元法离散的区域，用于求解热弹性中的同一边值问题。颜色变化表示不同的杨氏模量值。

对于第二种几何结构，我们引入 2601 个配点来最小化损失函数（见图 12 左侧）。图 12 右侧展示了相应的网格，该网格对区域进行离散化，以便通过有限元法求解热弹性问题。为评估系统的响应，采用了 10201 个点。在此情况下，约 75 % 的点与配点的位置不同。后者旨在说明神经网络插值解的适用性。第二个边值问题中使用的材料参数列于表 3。

	数值/单位
几何体2的热弹性问题	
基体杨氏模量和泊松比（E_{mat}, ν_{mat}）	（1.0 吉帕，0.3）
夹杂杨氏模量和泊松比（E_{inc}, ν_{inc}）	（0.5 吉帕，0.3）
基体热导率（k_{mat}）	1.0 瓦每米开尔文
夹杂物热导率（k_{inc}）	0.5 瓦每米开尔文

表3：线性热弹性问题中几何体2的输入参数

4.2.1. 第二种几何结构的耦合训练

根据热弹性问题定义的损失函数通过针对第二种几何结构的耦合训练过程得以最小化。结果如图 13 所示，我们发现，尽管非均质区域的形状各异，且夹杂物的变形模式复杂，物理信息神经网络的混合形式仍能够预测这一多物理场问题的响应。然而，在变形构型的左侧，可以观察到左边界条件未得到完全满足。这是网络的力学场响应产生误差的主要来源。参见图 13 第一行，其中某些区域的左侧狄利克雷边界条件未得到完全满足。解决这一问题的一种方法是采用硬约束边界条件，或在边界处增加配点；本文将对此进行讨论。

与第一种几何结构的结果类似，差异主要集中在界面区域。 u_x 和 σ_x 的平均相对误差分别为 1.6 % 和 3.3 %。同样，我们观察到误差局部化在夹杂物周围极小的区域内。

对于温度预测，误差集中在左边界附近；而对于通量，右边界也至关重要，见图14。对于上述数值，平均相对误差分别为 1.3 % 和 1.8 %。

在图15中，我们通过在左边界增加配点，对同一个边值问题进行了进一步研究。我们在这条边上额外增加了约 1000 个点。这一过程类似于有限元法中的网格细化。因此，位移场和温度场的最大误差分别降低了 1.6 倍和 3.0 倍。此外，我们还观察到整个区域的总体平均误差有所降低，这表明满足边界条件具有重要意义。

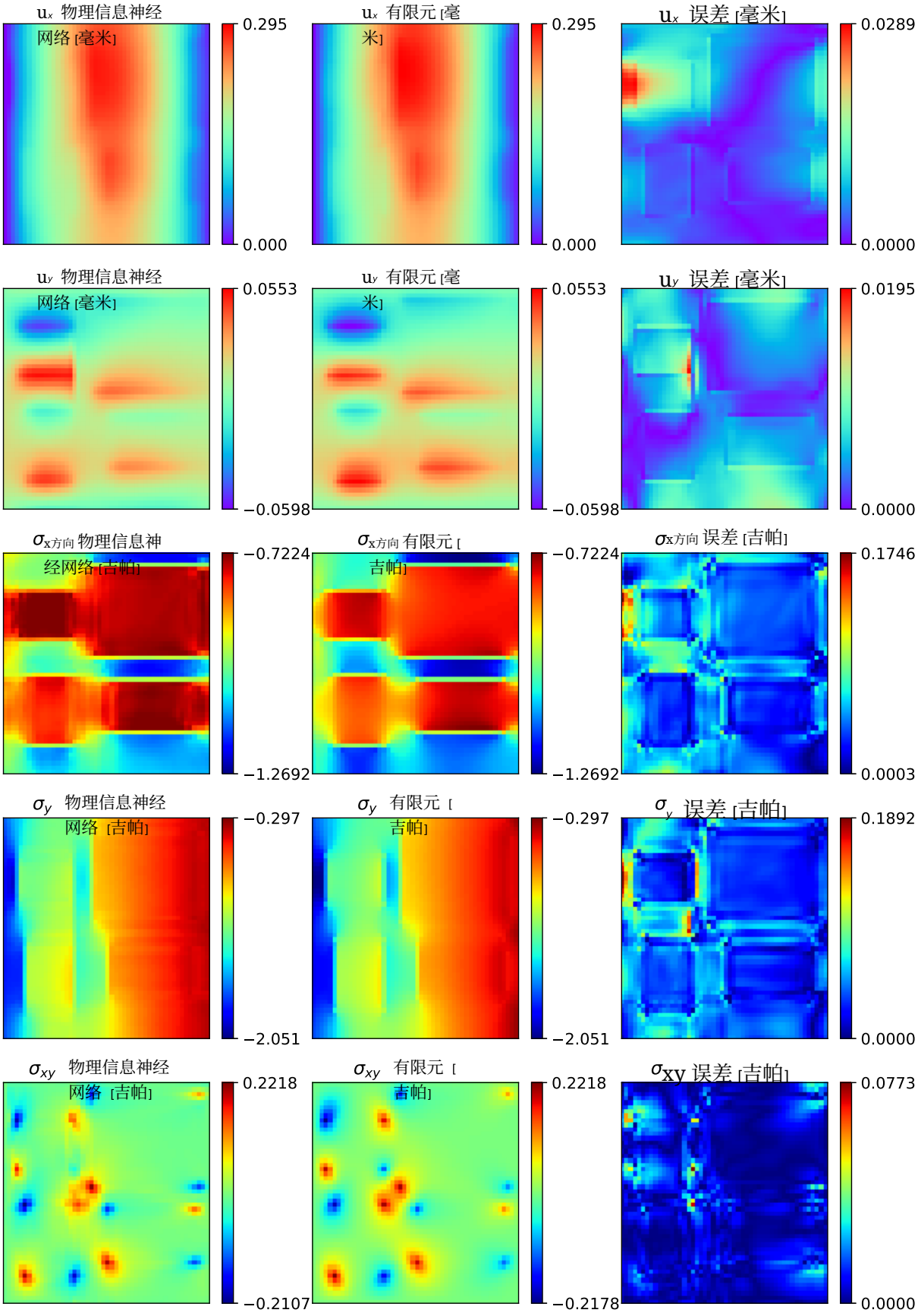


图13：采用耦合训练得到的几何体2热弹性力学场结果。左：物理信息神经网络；中：有限元法；右：有限元法与物理信息神经网络的差值。

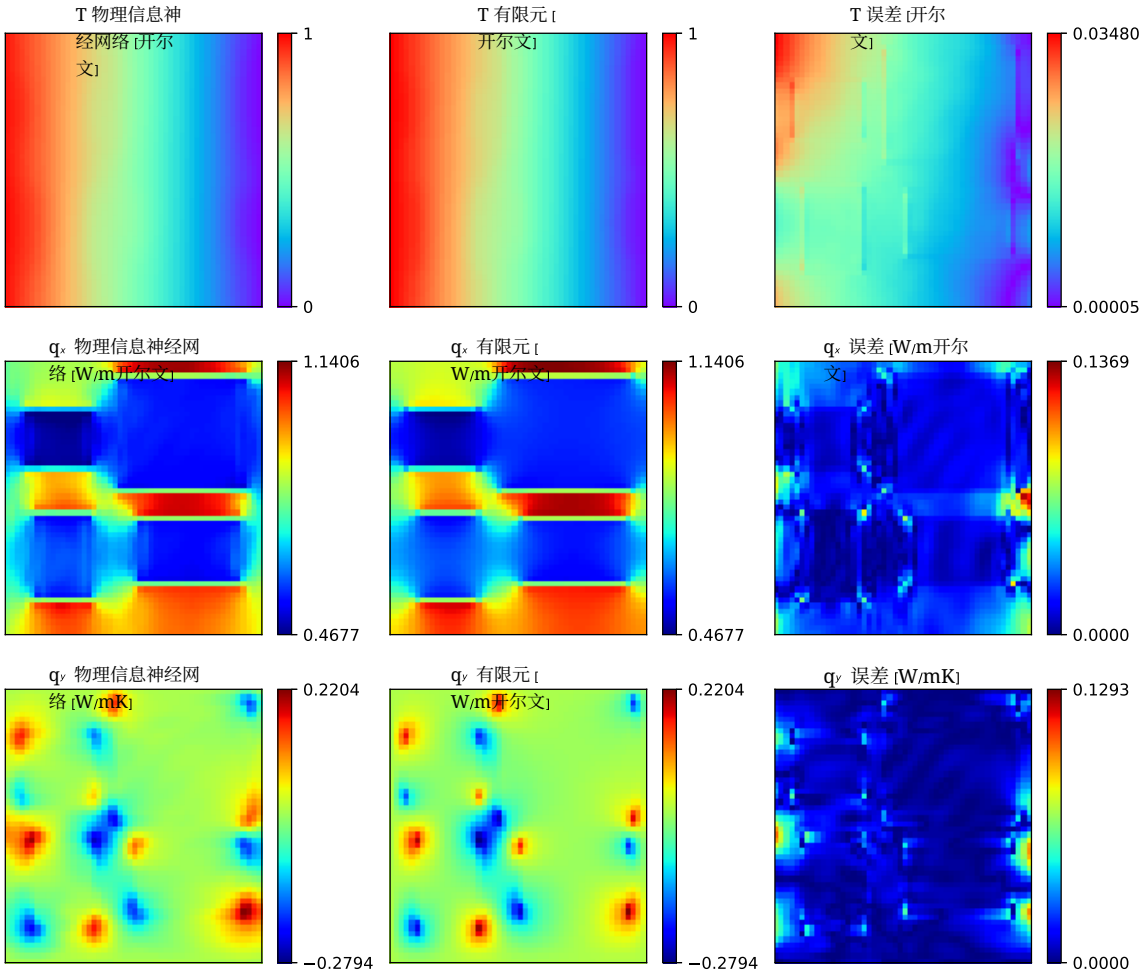


图14：采用耦合训练得到的第二种几何结构的热弹性热场结果。左：物理信息神经网络；中：有限元法；右：有限元法与物理信息神经网络的差异。

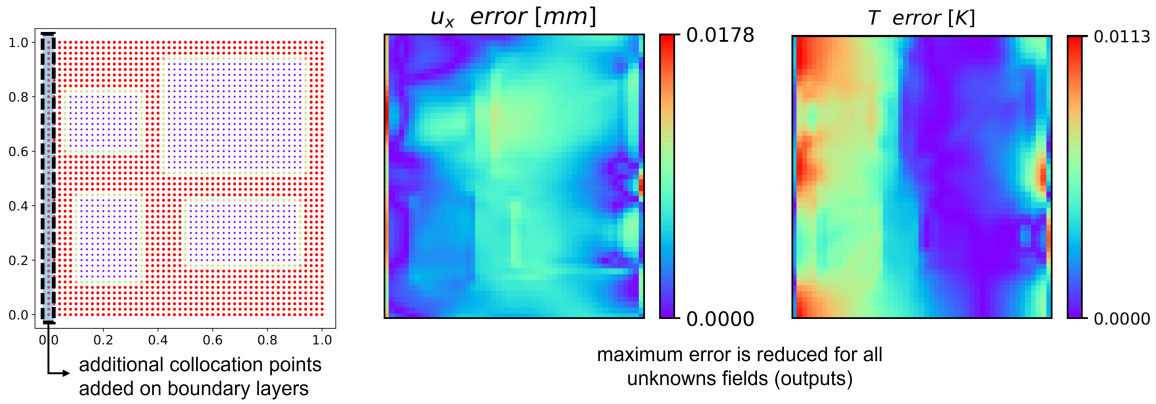


图15：在左边界细化配点后，误差得以降低。

4.2.2. 第二种几何结构的序贯训练

序贯训练方法中使用了与第 4.2.1 节相同数量的配点，以最小化损失函数。网络的相应结果部分绘于图 16 中。

我们还观察到，这种训练策略也出现了相同的问题，左侧的狄利克雷边界条件未得到完全满足。总体而言，结果与有限元法（FEM）的结果吻合良好，平均相对差异约为 0.49 % 和 2.12 %，分别对应 u_x 和 σ_x。对于热场，平均相对误差为 0.7 %。 y方向上的热通量相对误差为 1.82 %。

为更细致地检查采用耦合方法和序贯方法所得的结果，我们在 $x = 0.5$毫米处绘制截面，对比力学场输出分量、 x方向的热通量以及温度场。所述截面的选取方式确保其至少穿过一个夹杂部分。图16以有限元分析结果为参考解，对比了物理信息神经网络和耦合训练的结果。附录A中进行了网格收敛性分析，以确保通过有限元法获得的参考解准确。与有限元法结果相比，序贯训练的结果略优于耦合训练的结果。更具体地说，在夹杂附近，序贯训练的结果更接近参考解。相关的相对差异值也反映了顺序训练策略的优势。

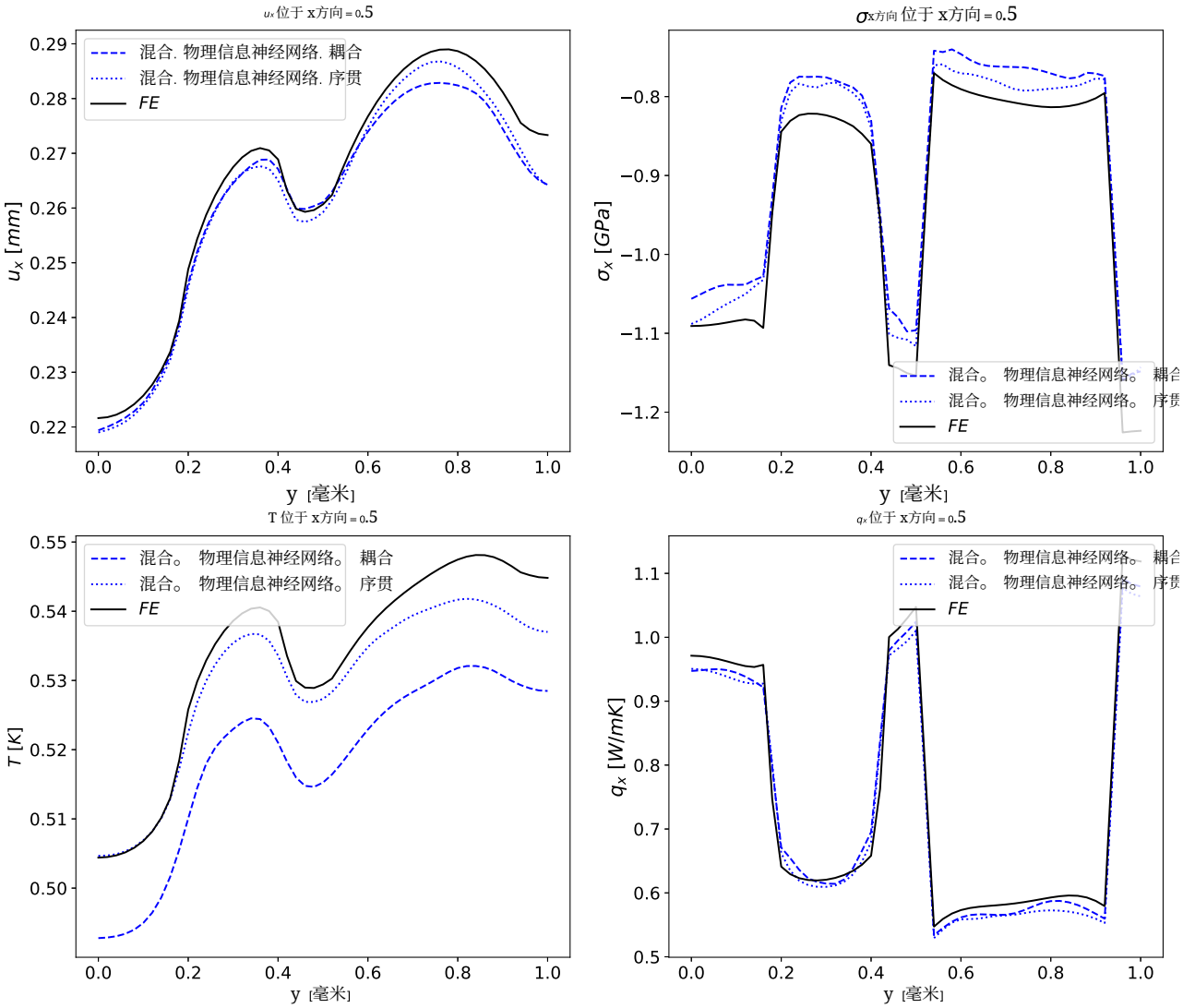


图16：比　　较几何体2在截面 $x = 0.5$毫米处网络输出的比较。

4.3. 提升模型性能的全面分析

本节讨论硬边界条件的概念，即调整网络的输出以满足指定的边界条件。正如 [67]和 [68] 所示，这一过程能够有效消除固定边界附近位移和温度场的预测误差。此外，本节还探讨了网络超参数和优化器的选择对训练后的网络采用所提出的混合物理信息神经网络公式求解热力学问题时的准确性和效率的影响。

为确保第二种几何结构满足狄利克雷边界条件，将网络的输出修改为

$$\hat{u}_x = u_x(x-1)(x), \tag{62}$$
$$\hat{u}_y = u_y(y-1)(y), \tag{63}$$
$$\hat{T}_y = T(x-1)(x) + 1 - x. \tag{64}$$

在式（62）中，符号 $(\hat{\bullet})$ 表示如何调整神经网络输出（$\bullet$），以确保训练后的网络满足狄利克雷边界条件。训练过程中通过耦合训练最小化其他损失函数。为比较该方法的效率，分别使用两种不同的优化器进行训练：Adam 优化器和基于拟牛顿方法的 L-BFGS。训练使用了 2601 个配点，如图 12 所示。

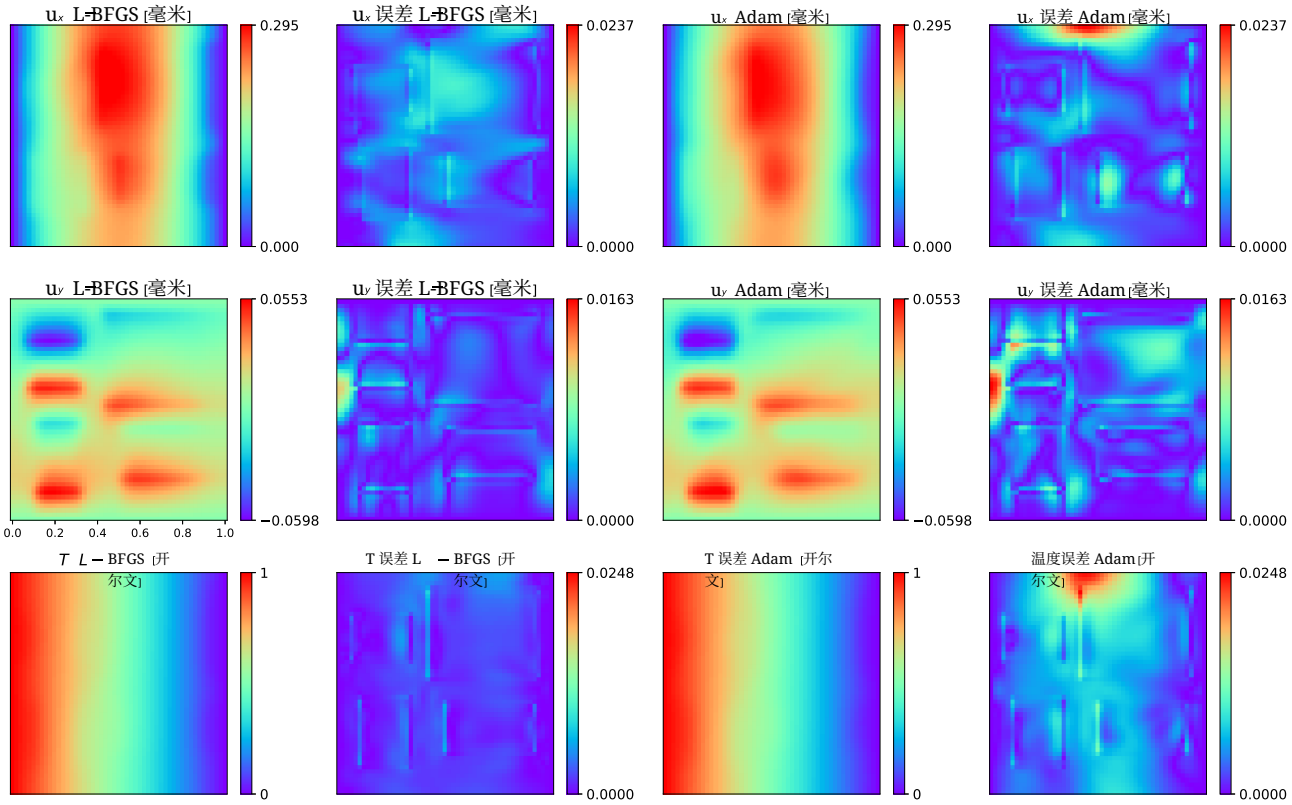


图 17：采用 L-BFGS 和 Adam 优化器时，施加硬狄利克雷边界条件对位移和温度场预测的影响。

图 17 展示了施加硬边界约束对力学场和温度场的影响。与采用软约束边界相比，施加硬边界约束可降低误差

条件。此外，使用拟牛顿算法优化器（L-BFGS）时，误差显著降低，几乎仅为使用 Adam 优化器进行耦合训练时的三分之一。图 18 显示，L-BFGS 优化器得到的总损失低于 Adam，表明其性能有所提升。

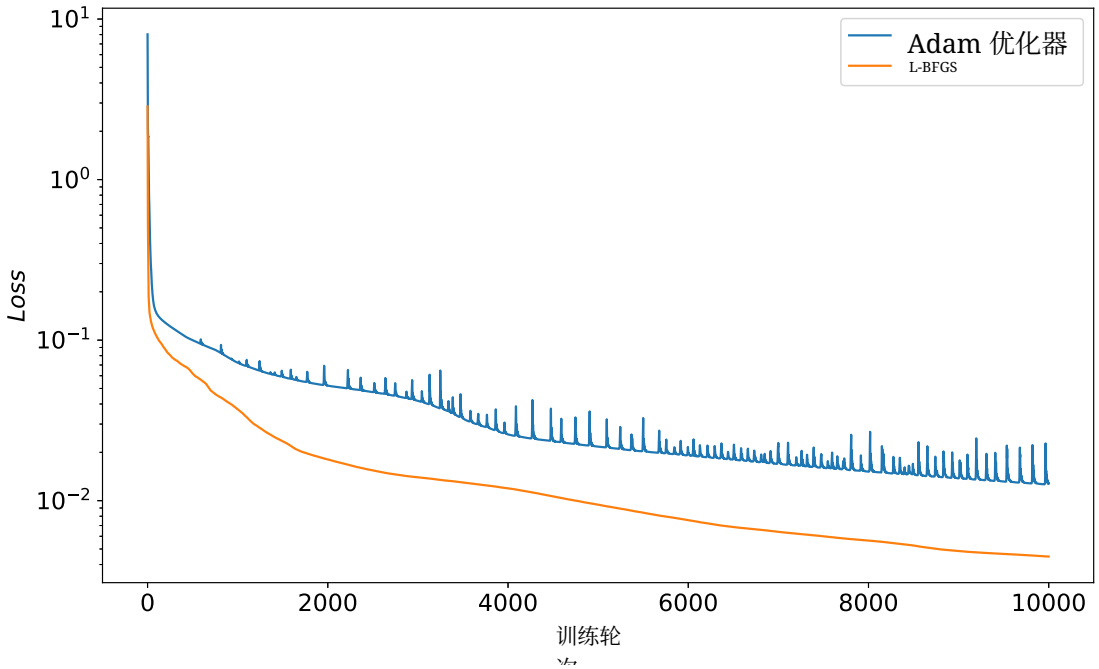


图 18：硬约束网络在 10000 个训练轮次（迭代次数）中，Adam 和 L-BFGS 优化器总损失的比较。

为考察网络超参数对网络架构有效性和效率的影响，我们研究了隐藏层数及其对网络性能的影响。我们采用了五种不同的网络架构，隐藏层数从一层到六层不等。所有网络的每个隐藏层均包含 40 个神经元，且各层均为全连接。图 20 所示结果表明，在各种量中，具有五层的网络始终优于其他网络，但 σ_x 除外；该处的误差高度局部化，见图 19。

确定网络的最佳层数后，我们进一步研究每层神经元数量的影响。沿用与前一项研究类似的方法，我们将隐藏层数固定为五层，并训练了每个隐藏层含五至六十个神经元的其他网络。结果见图 21。

此外，表4记录了每种网络架构每个训练轮次（迭代）的计算时间。报告的时间为一百个训练轮次（迭代）的平均耗时。

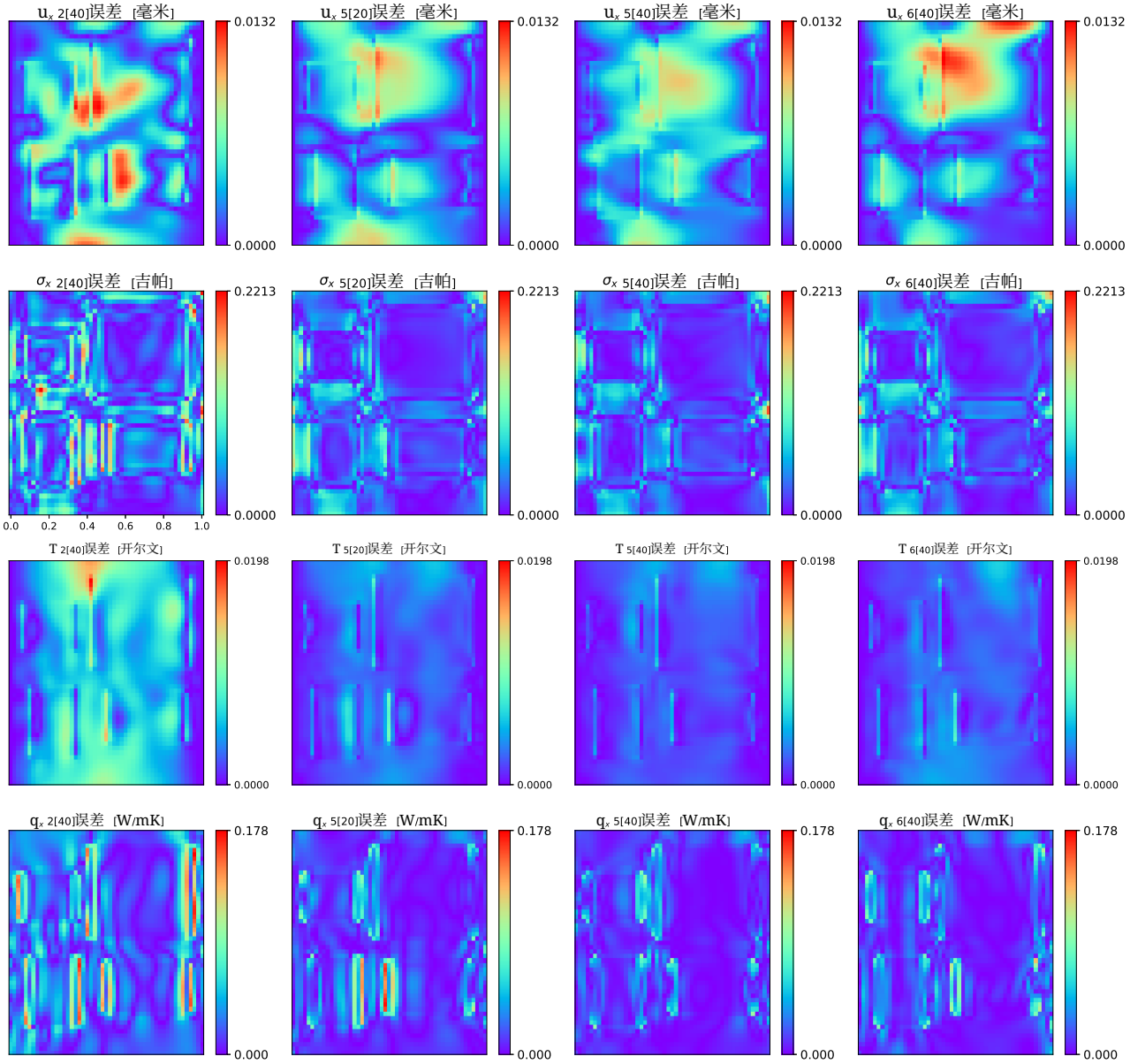


图19：关注量的不同网络架构比较，它们均使用相同的激活函数（双曲正切）和相同的迭代次数（10000）。

备注7 几何体二的边值问题使用 Adam 优化器训练一百万次迭代约需四小时。然而，切换到 L-BFGS优化器并施加硬约束后，仅需一万次迭代即可获得几乎相同的结果，在 Apple M2 Pro 平台上将训练时间缩短至约四十分钟。

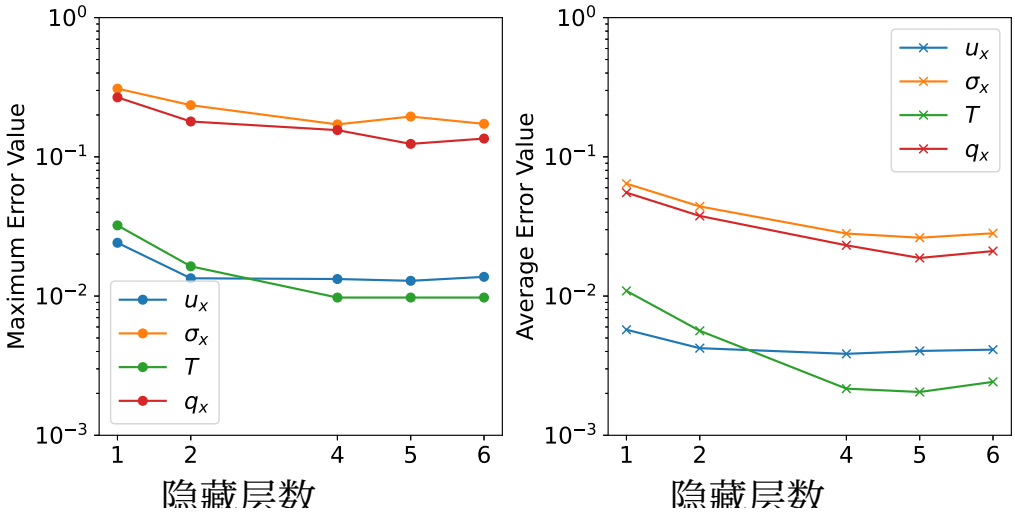


图20：比较每个隐藏层中神经元数量不同时关注量的最大误差和平均误差（共5个隐藏层，每层包含40个神经元）。

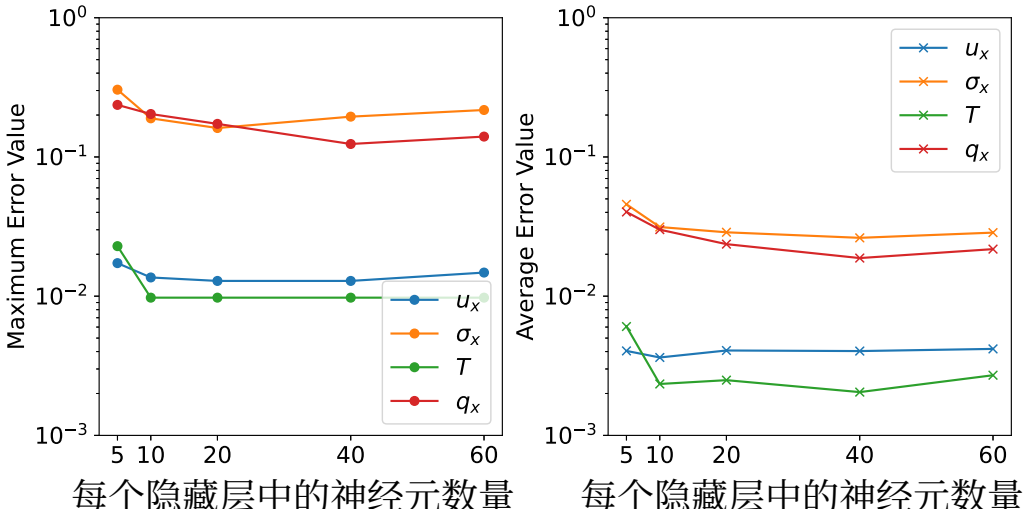


图21：比较每个隐藏层中神经元数量不同时关注量的最大误差和平均误差（隐藏层数固定为5）。

层 \ 神经元	5	10	20	40	60
1	0.2467	0.2492	0.2526	0.2712	0.2821
2	0.2595	0.26685	0.2706	0.2791	0.28845
4	0.2936	0.2948	0.3055	0.31865	0.3330
5	0.3141	0.3136	0.3294	0.3409	0.3600
6	0.3284	0.3346	0.3493	0.3718	0.3974

表4：使用 L-BFGS优化器在 Apple M2 pro 上，通过改变隐藏层数和神经元数量比较不同网络架构每个训练轮次的计算时间（单位：秒）。

4.4. 结合物理与数据的参数化学习

本节旨在通过结合数据、物理和参数化学习，获得未见案例的全场解。参数化学习使网络能够将其响应推广到不同的材料参数组。对于新的材料参数组，只需评估网络的响应，无需进一步训练，而训练仅需不到一秒 mS。后者比使用有限元等传统求解器求解系统快得多。

对于给定的边值问题，可以改变微观结构的拓扑结构、边界条件以及材料属性。为简化起见，我们仅关注所涉及的不同相的材料属性。因此，采用上一节中的几何体2，并为其设置不同的材料参数。为此，夹杂与基体之间的杨氏模量比以及它们的热导率比在1到10的范围内变化。对于每种情况，均通过有限元分析求解热弹性问题，随后将相应结果用于训练神经网络。有限元数据用于促进训练过程并降低计算成本。

训练网络时，需要将每个点的杨氏模量和热导率与配点坐标一同作为参数化学习的输入。所设计网络的架构如图 22 所示。除基于数据的损失外，还考虑了基于混合物理信息神经网络公式的损失， $\mathcal{L}_T$ 且 $\mathcal{L}_M$ 见公式 (35) 和 (36)。存在有限元分析数据时，总损失函数写为

$$\mathcal{L}_{total} = \underbrace{\mathcal{L}_{data}}_{\text{FE simulations}} + w \left(\underbrace{\mathcal{L}_T + \mathcal{L}_M}_{\text{Physics}} \right). \tag{65}$$

在式（65）中， w 表示物理在总损失函数中的权重。针对 w 参数从 0.0 到 1.0 变化的四种不同情形，对网络的预测进行了比较。在 $w = 0.0$情形下，研究的是数据驱动情形。考察一系列不同的权重值后，选取 $w = 0.1$ 的值，以获得最佳结果。训练采用迁移学习流程，针对不同相的杨氏模量 $E(x, y)$与热导率$k(x, y)$的特定比值，对式（65）中的 $\mathcal{L}_{total}$ 进行最小化。随后，经过 10 k 个训练轮次，求得第一组数值对应的网络最优参数。利用此前调优得到的参数，再针对新的比值进行 10 k 次迭代，以最小化损失。

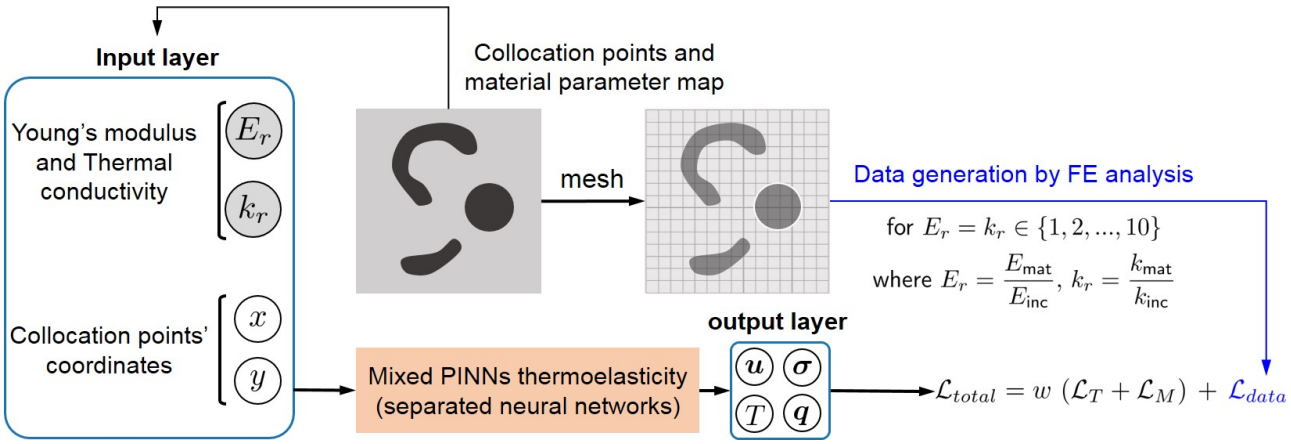


图 22：结合数据与物理信息并采用参数化算法时的网络输入和损失函数。

随后，调用基于数据与物理相结合的方法训练得到的网络，以及基于纯数据驱动方法训练得到的网络，预测 $E_r = K_r = 15$情形下的系统响应。该情形可视为这些比值的外推值，且不属于给定的训练集。

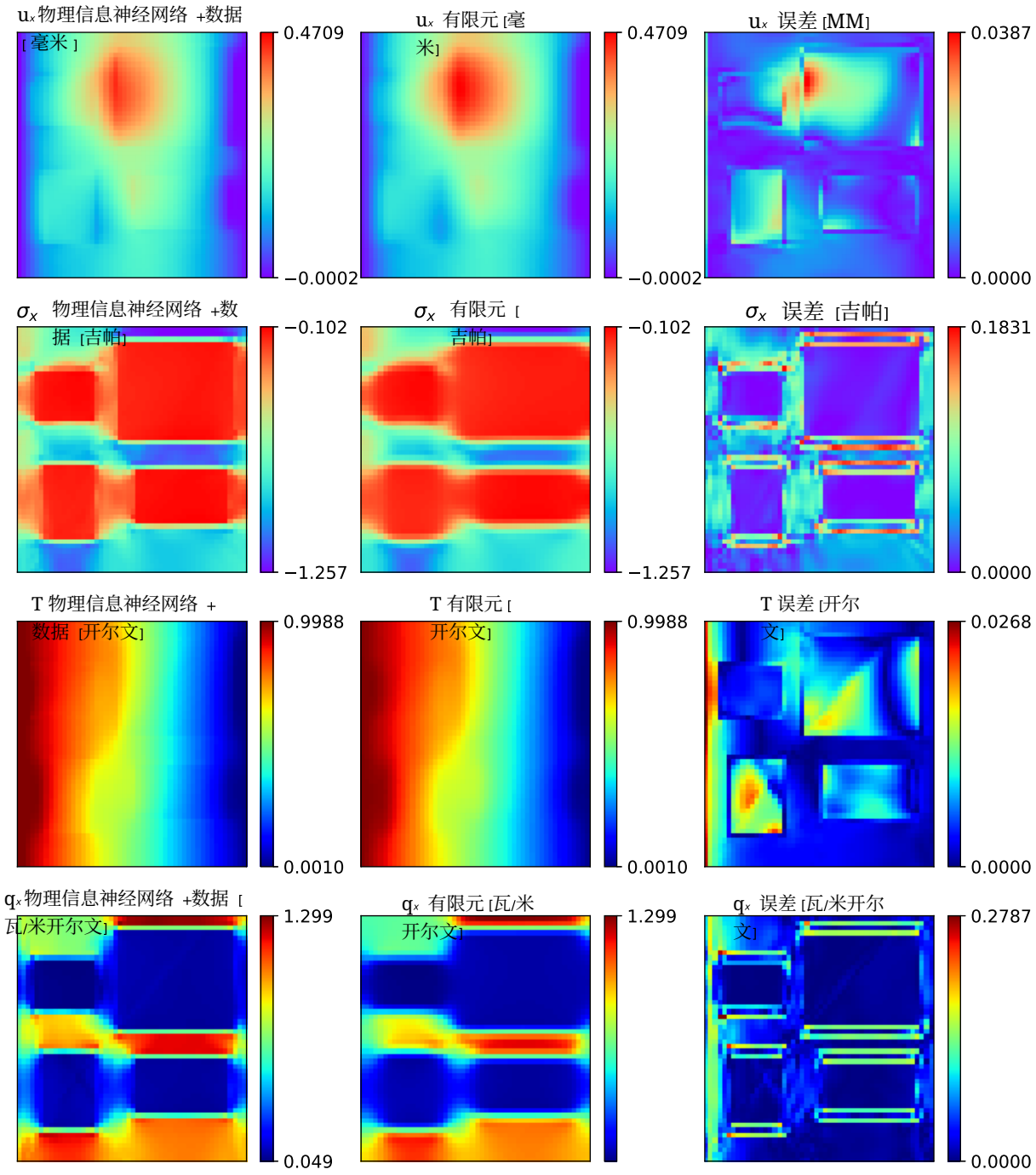


图23： $E_r = K_r = 15$情况下热弹性的力场和热场结果。左：物理信息神经网络 + 数据（$w = 0.1$），中：有限元法，右：有限元法与物理信息神经网络的差异。

图23和图24分别展示了结合数据与物理信息的情形以及纯数据驱动情形下，网络对主要场的预测。图中包含与

力学场相关的 T 和 q_x ，以及热场相关的 T 和 q_x 。采用迁移学习方案结合数据与物理信息，可更好地预测杨氏模量和热导率比均未见过的情形（外推）。

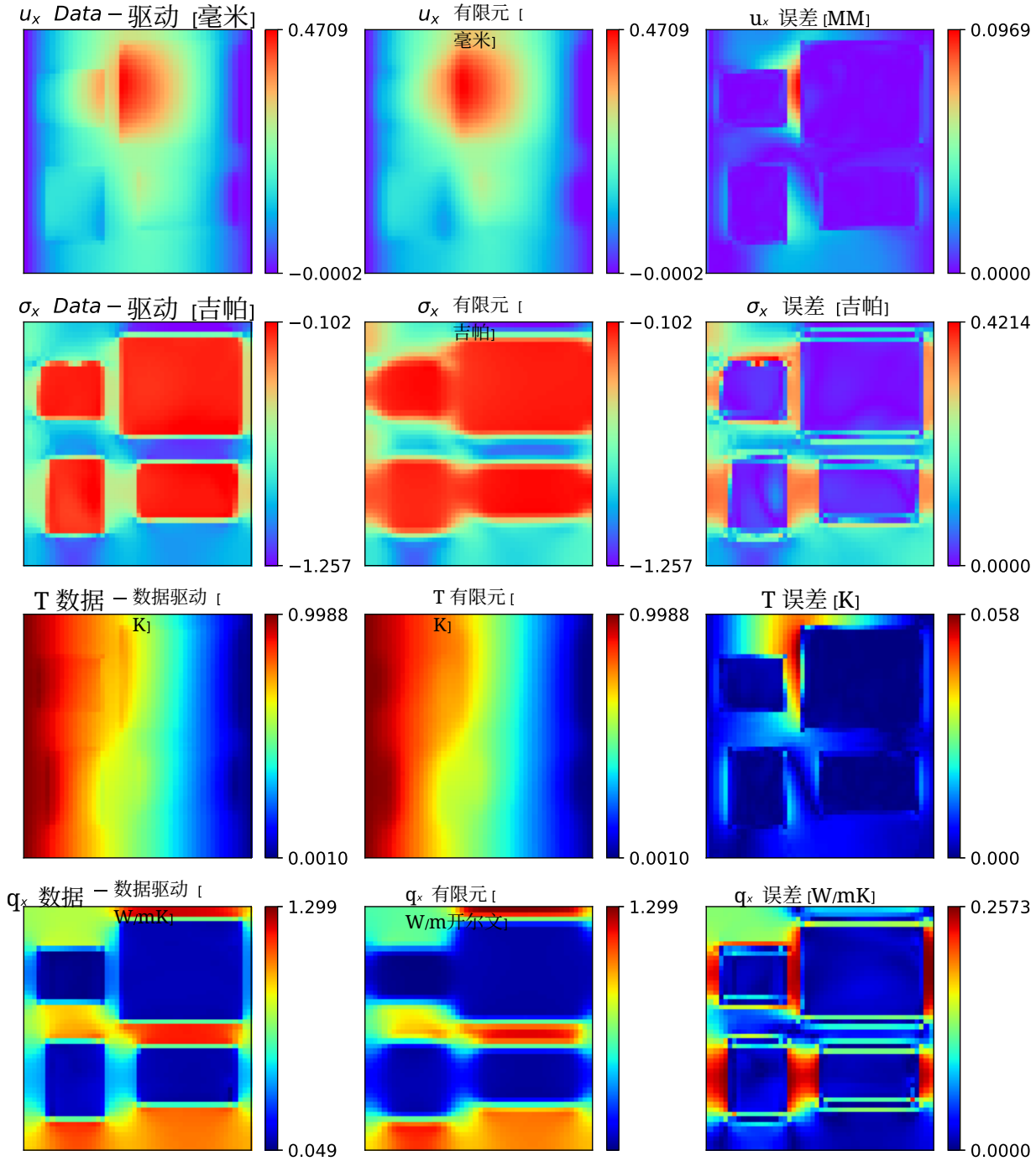


图24： $E_r = K_r = 15$情况下热弹性的力学场和热场结果。左：数据驱动（$w = 0.0$），中：有限元法，右：有限元法与数据驱动结果的差异。

总体而言，利用物理约束增强场预测，可将误差降低至纯数据驱动情形下误差的三分之一。采用物理约束时， x方向（q_x）热通量的最大误差略大，但这种

误差仅出现在区域内少数几个局部点上。此外，根据图 24，在数据驱动情形下，网络对夹杂物周围区域的预测准确度较低。

为了更好地进行比较，在 $x = 0.35$毫米处取一个截面。报告了数据与物理相结合的情形以及纯数据驱动情形的结果。该截面中关注的力学场输出为 u_x、 σ_x 。热场方面，选择 T和 q_x 进行比较。

总体而言，所述方法得到的相应结果与通过有限元分析计算的参考解吻合良好，见图 25。混合物理信息神经网络公式结合数据后，对边界值的近似效果更好。两种方法都能够捕捉夹杂物的影响；采用物理约束则提高了非均质区域内预测的准确性。

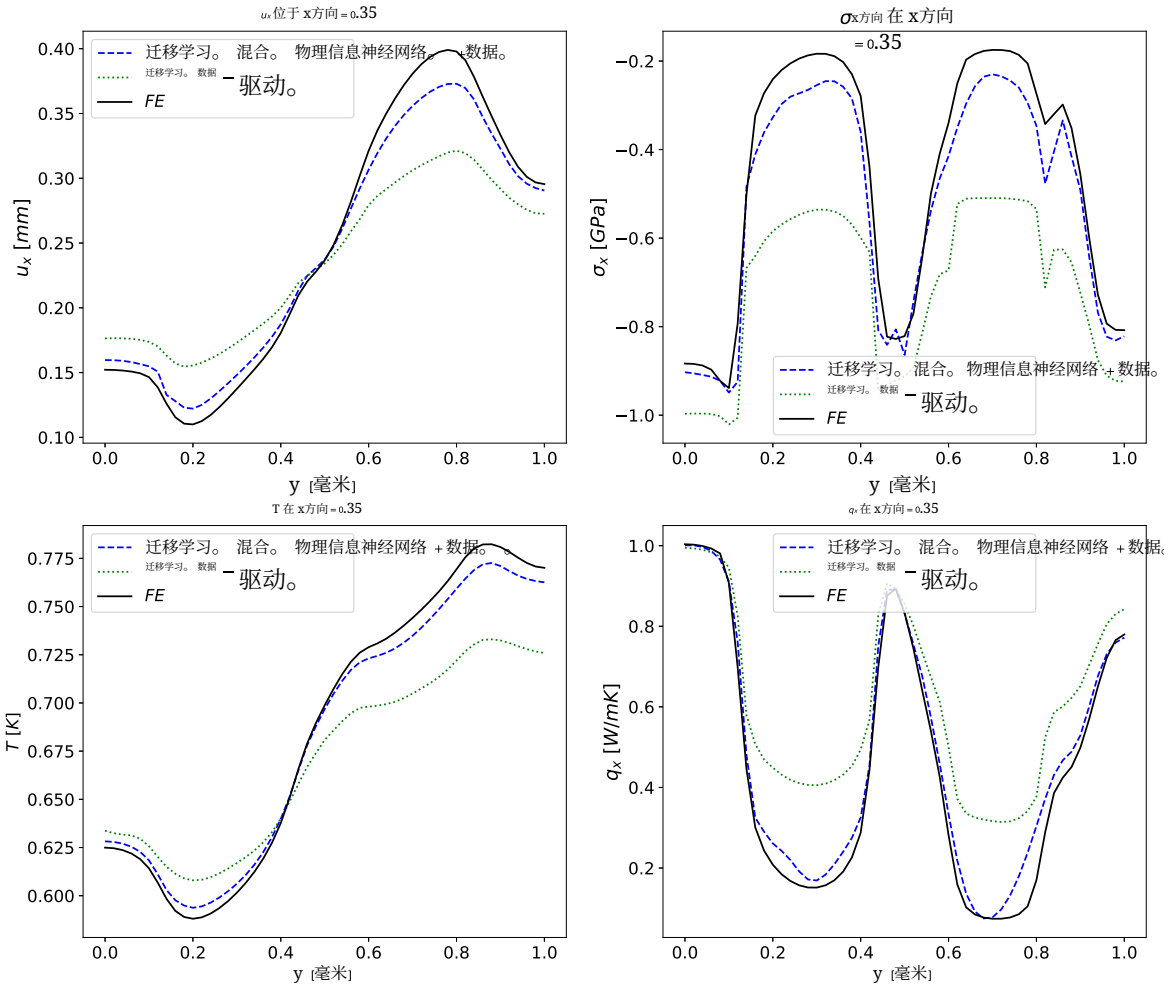


图25：采用混合物理信息神经网络公式结合数据，并利用迁移学习训练过程或纯数据驱动方案，对第二种几何结构在 $x = 0.5$毫米处截面的网络主要输出进行比较。

备注 8 采用迁移学习的数据驱动网络，其计算成本比通过结合数据与物理进行训练的网络低近 82%。构建物理约束时，需要计算输出参数相对于输入的导数，这是计算成本较高的主要原因。

5. 结论与展望

目前的研究表明，混合物理信息神经网络能够在没有任何真实数据的情况下，通过无监督学习有效求解非均质域中的多物理场问题。混合物理信息神经网络公式 [7, 23, 31], 仅采用一阶导数，并在损失函数中结合方程的强形式和弱形式，因此，与仅依赖强形式的传统场方法 [17]或仅采用弱形式的深度能量法 [30]相比，它能更准确地近似边值问题。此外，研究还探讨了序贯训练，即每次冻结一个场，并逐步进行损失函数最小化（另见 [38]）。结果表明，该方法在计算效率和网络精度方面均具有优势。为验证网络的能力，研究使用该网络求解一个准静态热弹性问题，并将结果与耦合训练所得结果进行比较；在耦合训练中，所有场的损失会同时最小化。

在后续阶段，将采用预设的网络架构并结合数据，预测此前未见案例的结果。目标是在显著降低计算成本、远低于有限元法等成熟方法的同时，即时预测不同材料属性下的系统响应。为此，训练过程将杨氏模量和热导率比作为额外输入参数。网络针对不同的比值进行训练，所得模型则用于预测（外推）此前未见比值对应的数值。训练后的网络几乎无需计算成本，即可预测给定边值问题的解。这一方法可能对数字孪生应用大有裨益，因为此类应用需要快速求得系统解。不过，要预测相同边界条件和初始条件下任意微观结构的响应，仍需开展更多研究。

未来研究中，至关重要的是将该方法论扩展至三维，并与其他高阶优化器（如 L-BFGS 方法，参见例如 [69]）比较性能。在许多应用中，将其扩展至弹塑性情形也十分必要 [70, 71]。此外，还可以着重求解其他耦合偏微分方程，例如相场损伤模型 [37, 49]。同时，研究更复杂的物理因素，以预测多物理场环境中的开裂 [33, 72] ，也已证明大有裨益。训练网络可用于结构优化，以及通过增材制造制备材料的优化设计。此外，本文讨论的迁移学习方法可用于推广网络，使其适用于不同的问题配置。将这些思路与算子学习方法 [2, 73]相结合并进行比较，也将具有重要价值。

致谢:
谨此感谢德国研究基金会（DFG）资助跨区域协作研究中心SFB/TRR 87子项目A6（项目编号138690629）以及SFB/TRR 339（项目编号453596084）。作者还感谢工业研究协会联合会“奥托·冯·格里克”注册协会（AiF）通过项目资助IGF 21348 N/3提供的经费支持。

作者声明:
Ali Harandi：方法论、软件、撰写、审阅与编辑。Ahmad Moidendin：方法论、软件、撰写、审阅与编辑。Michael Kaliske：指导、审阅与编辑。Stefanie Reese：经费获取、审阅与编辑。Shahed Rezaei：构思、方法论、指导、撰写、审阅与编辑。

附录A：有限元计算的收敛性研究

为验证第4.2节所用参考结果是否已收敛，通过采用三种不同的网格密度对区域进行离散化开展了研究。对含有多个夹杂物（几何体2）的材料系统进行网格划分，沿每条边分别设置50、100和200个单元。如图26所示，每个方向采用100个单元时，结果几乎已收敛。热场中的偏差可以忽略不计，且完全是由绘图比例尺造成的。

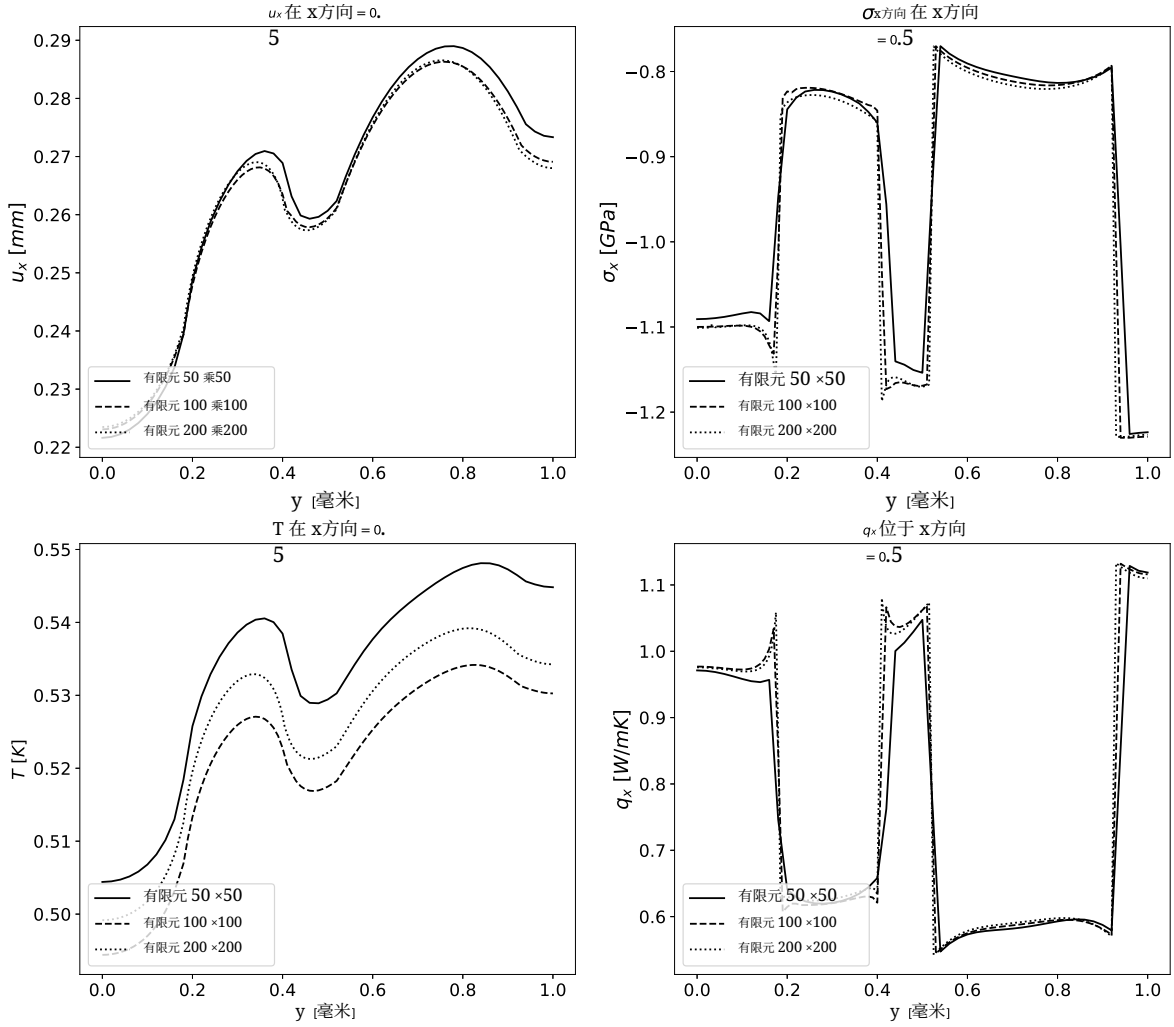


图26：对于包含多个夹杂物的第二种几何结构，有限元结果随单元数量的收敛情况。

附录B：利用狄利克雷硬边界条件和 L-BFGS优化器求得的全场解

在本附录中，　　我们展示了机械场和热场的完整场预测　　.

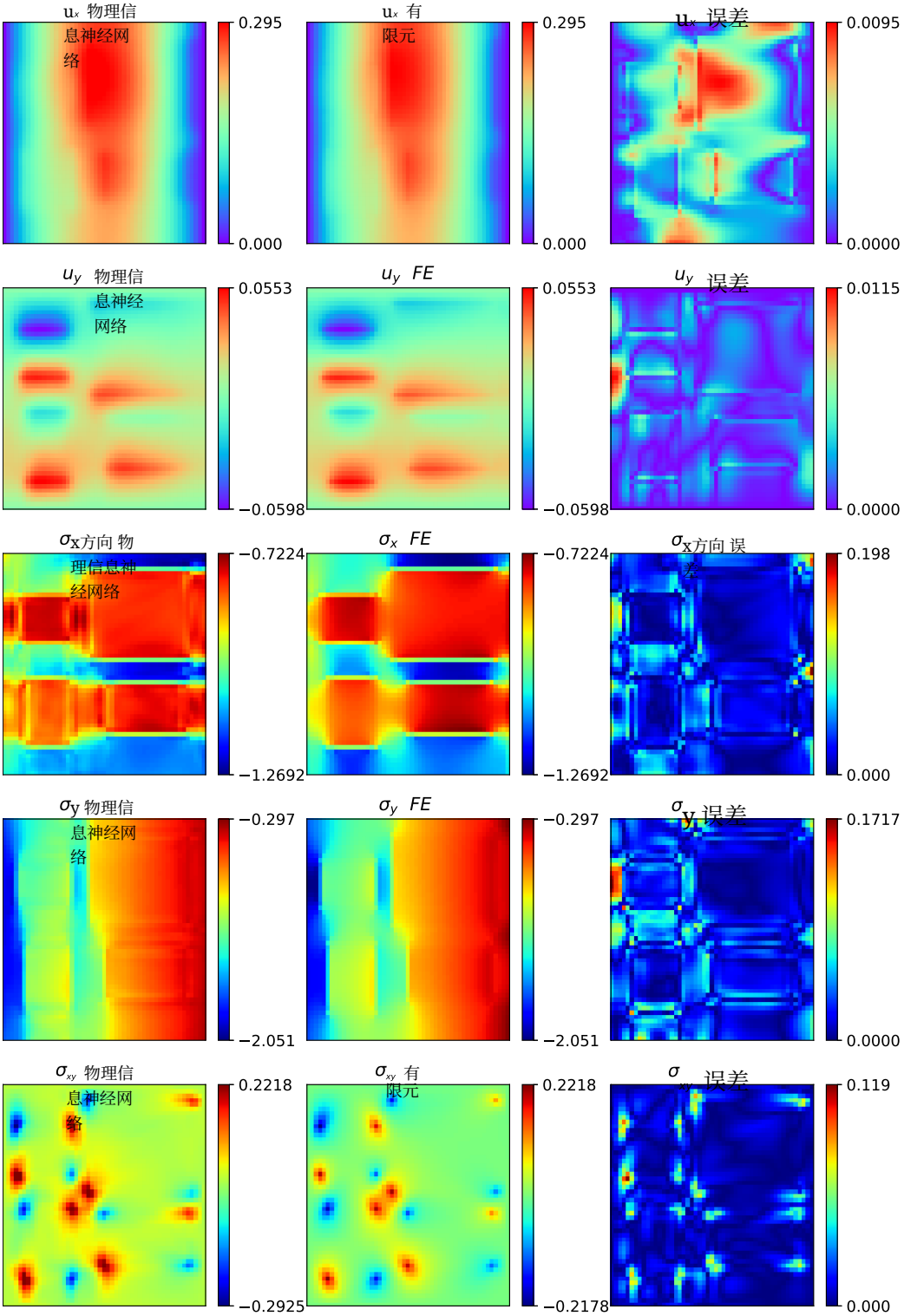


图27：采用耦合训练（L-BFGS优化器）和狄利克雷硬边界条件时，几何体2的热弹性力学场结果。左：物理信息神经网络；中：有限元法；右：有限元法与物理信息神经网络的差值。

为实现这些预测，我们采用了硬边界约束，并结合耦合训练过程使用 L-BFGS优化器。完成 10000 次迭代后，展示了结果。通过采用硬边界约束，我们确保预测准确，并符合狄利克雷边界条件。

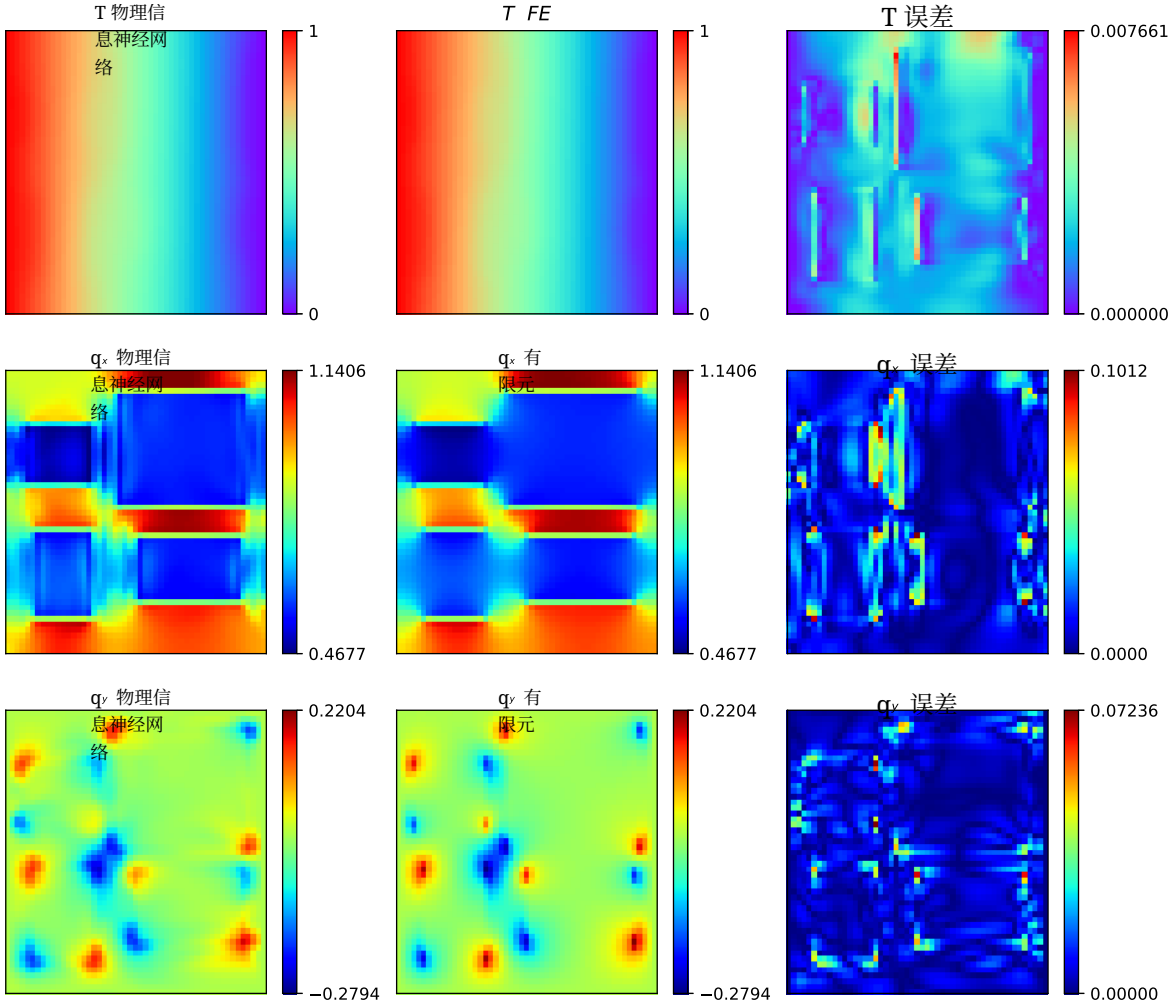


图28：采用耦合训练（L-BFGS优化器）和硬狄利克雷边界条件时，第二种几何结构的热弹性热场结果。左：物理信息神经网络；中：有限元法；右：有限元法与物理信息神经网络之差。

参考文献

[1] Zongyi Li、Nikola Kovachki、Kamyar Azizzadenesheli、Burigede Liu、Kaushik Bhattacharya、Andrew Stuart 和 Anima Anandkumar。用于参数化偏微分方程的傅里叶神经算子。*arXiv*预印本 *arXiv:2010.08895*，2020年。[2] Lu Lu、Pengzhan Jin、Guofei Pang、Zhongqiang Zhang 和 George Em Karniadakis。基于算子的通用逼近定理，通过 DeepONet 学习非线性算子。《自然·机器智能》，3：218–229，2021年。[3] Frederic E. Bock、Roland C. Aydin、Christian J. Cyron、Norbert Huber、Surya R. Kalidindi 和 Benjamin Klusemann。机器学习与数据挖掘方法在连续介质材料力学中的应用综述。《材料前沿》，6，2019年。[4] Mauricio Fern´andez、Shahed Rezaei、Jaber Rezaei Mianroodi、Felix Fritzen 和 Stefanie Reese。人工神经网络在界面力学预测中的应用：晶界本构行为研究。《工程科学高级建模与仿真》，7：1–27，2020年。[5] Grace C. Y. Peng、Mark Alber、Adrian Buganza Tepole、William R. Cannon、Suvranu De、Savador Dura-Bernal、Krishna Garikipati、George Karniadakis、William W. Lytton、Paris Perdikaris、Linda Petzold 和 Ellen Kuhl。多尺度建模与机器学习相遇：我们能学到什么？《工程计算方法档案》，28：1017–1037，2021年。[6] Jaber Rezaei Mianroodi、Shahed Rezaei、Nima H Siboni、Bai-Xiang Xu 和 Dierk Raabe。基于人工智能的无损多尺度本构弹性关系。*npj* 计算材料，8：1–12，2022年。[7] Shahed Rezaei、Ali Harandi、Ahmad Moeineddin、Bai-Xiang Xu 和 Stefanie Reese。面向非均质域中工程问题的物理信息神经网络混合形式：与有限元法的比较。*ComputerMethods in Applied Mechanics and Engineering*，401：115616，2022年。[8] Donggeun Park、Jiyoung Jung、Grace X. Gu 和 Seunghwa Ryu。可泛化且可解释的深度学习模型，用于提高网格复合材料应变场的预测精度。《材料与设计》，223：111192，2022年。[9] Zhenze Yang、Chi-Hua Yu 和 Markus J. Buehler。用于预测层级复合材料中复杂应力场和应变场的深度学习模型。《科学进展》，7：eabd7416，2021年。[10] Mohammad S Khorrami、Jaber R Mianroodi、Nima H Siboni、Pawan Goyal、Bob Svendsen、Peter Benner 和 Dierk Raabe。用于黏塑性多晶材料应力场代理建模的人工神经网络。*arXiv*预印本 *arXiv:2208.13490*，2022年。[11] Meer Mehran Rashid、Tanu Pittie、Souvik Chakraborty 和 N.M. Anoop Krishnan。利用傅里叶神经算子学习数字复合材料中的应力-应变场。《*iScience*》，25：105452，2022年。[12]Hannah Lu、Dinara Ermakova、Haruko Murakami Wainwright、Liange Zheng 和 Daniel M. Tartakovsky。用于多物理场仿真的数据驱动模拟器。《建模与计算机器学习期刊》，2：33–54，2021年。

[13] Marco Ragone、Vitaliy Yurkiv、Ajaykrishna Ramasubramanian、Babak Kashir 和 Farzad Mashayek。结合汽车仿真与多物理场建模的电动汽车电池荷电状态数据驱动估计。*Journal of Power Sources*，483：229108，2021年。[14] Bowen Wang、Guobin Zhang、Huizhi Wang、Jin Xuan 和 Kui Jiao。采用数据驱动代理模型的质子交换膜燃料电池多物理场解析数字孪生。*Energy and AI*，1：100004，2020年。[15] Lu Lu、Xuhui Meng、Shengze Cai、Zhiping Mao、Somdatta Goswami、Zhongqiang Zhang 和 George Em Karniadakis。基于公平数据对两种神经算子进行全面、公正的比较（含实用扩展）。*Computer Methods in Applied Mechanicsand Engineering*，393：114778，2022年。[16] Sarthak Kapoor、Jaber Rezaei Mianroodi、Mohammad Khorrami、Nima S Siboni 和 Bob Svendsen。针对材料非均质弹塑性固体应力代理建模而训练的两种人工神经网络的比较。*arXiv*预印本 *arXiv:2210.16994*，2022年。[17] 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问题与反问题的深度学习框架。*Journal of Computational Physics*，378：686–707，2019年。[18] Ameya D Jagtap 和 George E Karniadakis。扩展物理信息神经网络（xpinns）：一种基于广义时空域分解、用于非线性偏微分方程的深度学习框架。*Communications in Computational Physics*，28：2002–2041，2020年。[19] Ameya D. Jagtap、Ehsan Kharazmi 和 George Em Karniadakis。离散域上的守恒型物理信息神经网络：在守恒律正问题与反问题中的应用。*Computer Methods in Applied Mechanics and Engineering*，365：113028，2020年。[20] Ehsan Haghighat、Maziar Raissi、Adrian Moure、Hector Gomez 和 Ruben Juanes。用于固体力学反演与代理建模的物理信息深度学习框架。*Computer Methods in Applied Mechanics and Engineering*，379：113741，2021年。[21] Zheyuan Hu、Ameya D. Jagtap、George Em Karniadakis 和 Kenji Kawaguchi。扩展物理信息神经网络（xpinns）何时能提升泛化能力？*SIAMJournal on Scientific Computing*，44：A3158–A3182，2022年。[22] Ameya D. Jagtap 和 George Em Karniadakis。扩展物理信息神经网络（xpinns）：一种基于广义时空域分解、用于非线性偏微分方程的深度学习框架。*Communications in Computational Physics*，28：2002–2041，2020年。[23] Alexander Henkes、Henning Wessels 和 Rolf Mahnken。面向连续介质微观力学的物理信息神经网络。*Computer Methods in Applied Mechanics and Engineering*，393：114790，2022年。

[24] Yizheng Wang、Jia Sun、Wei Li、Zaiyuan Lu 和 Yinghua Liu。CENN：一种基于带子域神经网络的保守能量法，用于求解涉及非均质区域和复杂几何结构的变分问题。*Computer Methods in Applied Mechanicsand Engineering*，400：115491，2022年。[25] Sifan Wang、Hanwen Wang 和 Paris Perdikaris。关于傅里叶特征网络的特征向量偏差：从回归到使用物理信息神经网络求解多尺度偏微分方程。*Computer Methods in Applied Mechanics and Engineering*，384：113938，2021年。[26] Levi McClenny 和 Ulisses Braga-Neto。利用软注意力机制的自适应物理信息神经网络。*arXiv*预印本 *arXiv:2009.04544*，2020年。[27] Zixue Xiang、Wei Peng、Xu Liu 和 Wen Yao。自适应损失平衡物理信息神经网络。*Neurocomputing*，496：11–34，2022年。[28] Rafael Bischof 和 Michael Kraus。物理信息深度学习中的多目标损失平衡。*arXiv*预印本 *arXiv:2110.09813*，2021年。[29] Franz M Rohrhofer、Stefan Posch 和 Bernhard C Geiger。物理信息神经网络的帕累托前沿。*arXiv*预印本 *arXiv:2105.00862*，2021年。[30] E. Samaniego、C. Anitescu、S. Goswami、V.M. Nguyen-Thanh、H. Guo、K. Hamdia、X. Zhuang 和 T. Rabczuk。通过机器学习求解计算力学中的偏微分方程的能量方法：概念、实现与应用。*Computer Methods in Applied Mechanics and Engineering*，362：112790，2020年。[31] Jan N. Fuhg 和 Nikolaos Bouklas。用于解析有限应变超弹性中集中特征的混合深度能量法。*Journal of Computational Physics*，451：110839，2022年。[32] Diab W. Abueidda、Seid Koric、Erman Guleryuz 和 Nahil A. Sobh。用于超弹性的增强型物理信息神经网络。《国际工程数值方法杂志》，124(7)：1585–1601，2023年。[33] Hui Ruan、Shahed Rezaei、Yangyiwei Yang、Dietmar Gross 和 Bai-Xiang Xu。热力学相场断裂模型：在增材制造热裂纹模拟中的应用。*Journal of the Mechanics and Physics of Solids*，172：105169，2023年。[34] Shahed Rezaei、Armin Asheri 和 Bai-Xiang Xu。一种一致的化学-力学黏聚断裂框架及其在固态电池中的应用。*Journal of theMechanics and Physics of Solids*，157：104612，2021年。[35] Hongwei Guo、Xiaoying Zhuang、Pengwan Chen、Naif Alajlan 和 Timon Rabczuk。利用材料迁移学习和敏感性分析的物理信息深度配置法，分析非均质介质中的三维势问题。*Engineering with Computers*，2022年。[36] Somdatta Goswami、Minglang Yin、Yue Yu 和 George Em Karniadakis。一种物理信息变分 DeepONet，用于预测准脆性材料中的裂纹路径。*ComputerMethods in Applied Mechanics and Engineering*，391：114587，2022年。

[37] Shahed Rezaei、Ali Harandi、Tim Brepols 和 Stefanie Reese。各向异性黏聚断裂模型：长度尺度无关相场损伤模型的优势与局限。工程断裂力学，261：108177，2022年。[38] Danial Amini、Ehsan Haghighat 和 Ruben Juanes。多孔介质中热-水-力过程的物理信息神经网络求解。《工程力学杂志》，148(11)：04022070，2022年。[39]Mayank Raj、Pramod Yallappa Kumbhar 和 Ratna Kumar Annabattula。用于求解功能梯度材料热力学问题的物理信息神经网络。*arXiv*，abs/2111.10751，2021年。[40] R. Laubscher。利用物理信息神经网络模拟多组分流动与传热。《流体物理学》，33：087101，2021年。[41] Revanth Mattey 和 Susanta Ghosh。一种用于训练物理信息神经网络以求解 Allen–Cahn 方程和 Cahn-Hilliard方程的新型序贯方法。*Computer Methods in AppliedMechanics and Engineering*，390：114474，2022年。[42]Haotian Chen、Christopher Batchelor-McAuley、Enno K¨atelh¨on、Joseph Elliott 和 Richard G. Compton。对使用物理信息神经网络模拟伏安法的批判性评估：优势、劣势与最佳实践。电分析化学杂志，925：116918，2022年。[43] Long Nguyen、Maziar Raissi 和 Padmanabhan Seshaiyer。结合区域分解方法求解耦合多物理场问题的高效物理信息神经网络。载于《计算建模与仿真进展》，第41–53页。施普林格，2022年。[44] Bin Zheng、Tongchun Li、Huijun Qi、Lingang Gao、Xiaoqing Liu 和 Li Yuan。用于准脆性材料计算断裂的无标签数据物理信息机器学习模型。《国际力学科学杂志》，223：107282，2022年。[45] Yousef Ghaffari Motlagh、Peter K. Jimack 和 Ren´e de Borst。用于脆性断裂的深度学习相场模型。《国际工程数值方法杂志》，124：620–638，2023年。[46] Fuzhen Zhuang、Zhiyuan Qi、Keyu Duan、Dongbo Xi、Yongchun Zhu、Hengshu Zhu、Hui Xiong 和 Qing He。迁移学习综合综述。《*IEEE*会议录》，109：43–76，2021年。[47] Behnam Neyshabur、Hanie Sedghi 和 Chiyuan Zhang。迁移学习中迁移的究竟是什么？《神经信息处理系统进展》，33：512–523，2020年。[48] Hesheng Tang、Yangyang Liao、Hu Yang 和 Liyu Xie。一种用于涡激振动的迁移学习物理信息神经网络（TL-PINN）。《海洋工程》，266：113101，2022年。[49] Somdatta Goswami、Cosmin Anitescu、Souvik Chakraborty 和 Timon Rabczuk。通过迁移学习增强的物理信息神经网络，用于断裂相场建模。《理论与应用断裂力学》，106：102447，2020年。

[50] Rapha¨el Pellegrin、Blake Bullwinkel、Marios Mattheakis 和 Pavlos Protopapas。利用物理信息神经网络进行迁移学习，以高效模拟分支流，2022年。[51]Shaan Desai、Marios Mattheakis、Hayden Joy、Pavlos Protopapas 和 Stephen Roberts。物理信息神经网络的单次迁移学习。*arXiv*预印本 *arXiv:2110.11286*，2021年。[52] Yihang Gao、Ka Chun Cheung 和 Michael K. Ng。SVD-PINNs：通过奇异值分解对物理信息神经网络进行迁移学习，2022年。[53] J¨urgen Schmidhuber。神经网络中的深度学习：综述。神经网络，61：85–117，2015年。[54] Ameya D. Jagtap 和 George Em Karniadakis。激活函数在回归与分类中有多重要？综述、性能比较及未来方向，2022年。[55] J. Abbasi 和 P. Andersen。物理激活函数（PAFs）：一种将物理机制更高效地引入物理信息神经网络（PINNs）的方法。*arXiv*预印本 *arXiv:2205.14630*，2022年。[56] Adam Paszke、Sam Gross、Soumith Chintala、Gregory Chanan、Edward Yang、Zachary DeVito、Zeming Lin、Alban Desmaison、Luca Antiga 和 Adam Lerer。PyTorch中的自动微分。第*31*届神经信息处理系统大会，2017年。[57] Mart´ın Abadi、Paul Barham、Jianmin Chen、Zhifeng Chen、Andy Davis、Jeffrey Dean、Matthieu Devin、Sanjay Ghemawat、Geoffrey Irving、Michael Isard 等。 {TensorFlow}：一个用于 {大规模} 机器学习的系统。载于第*12*届*USENIX*操作系统设计与实现研讨会（*OSDI 16*），第265–283页，2016年。[58] Atilim Gunes Baydin、Barak A Pearlmutter、Alexey Andreyevich Radul 和 Jeffrey Mark Siskind。机器学习中的自动微分：综述。机器学习研究杂志，18：1–43，2018年。[59] Ehsan Haghighat 和 Ruben Juanes。SciANN：一个用于科学计算和基于人工神经网络的物理信息深度学习的Keras/TensorFlow封装器。*Computer Methods in Applied Mechanics and Engineering*，373：113552，2021年。[60] Ehsan Haghighat、Danial Amini 和 Ruben Juanes。采用应力分裂序贯训练的多相孔弹性物理信息神经网络模拟。*ComputerMethods in Applied Mechanics and Engineering*，397：115141，2022年。[61] Carlos A Felippa 和 Kwang-Chun Park。耦合机械系统的交错瞬态分析方法：公式化。*Computer Methods in Applied Mechanics and Engineering*，24：61–111，1980年。[62] Pao-Hsiung Chiu、Jian Cheng Wong、Chinchun Ooi、My Ha Dao 和 Yew-Soon Ong。CAN-PINN：一种基于耦合自动—数值微分法的快速物理信息神经网络。*Computer Methods in Applied Mechanics and Engineering*，395：114909，2022年。

[63] R . L. Taylor . FEAP - 有限元分析程序 2014
, .[64] Khemraj Shukla、Ameya D. Jagtap 和 George Em Karniadakis。通过区域分解实现并行物理信息神经网络。*Journal of Computational Physics*，447:110683，2021年。[65] Ehsan Kharazmi、Zhongqiang Zhang 和 George E.M. Karniadakis。hp-vpinns：结合区域分解的变分物理信息神经网络。*Computer Methodsin Applied Mechanics and Engineering*，374:113547，2021年。[66] J. Nguejio、F. Szmytka、S. Hallais、A. Tanguy、S. Nardone 和 M. Godino Martinez。增材制造与锻造镍合金625的微观结构特征和力学性能比较。材料科学与工程：*A*，764:138214，2019年。[67] Lu Lu、Raphael Pestourie、Wenjie Yao、Zhicheng Wang、Francesc Verdugo 和 Steven G Johnson。采用硬约束的物理信息神经网络进行逆向设计。*SIAMJournal on Scientific Computing*，43(6):B1105–B1132，2021年。[68]Shaikhah Alkhadhr 和 Mohamed Almekkawy。通过物理信息神经网络对波动方程进行建模：针对初始条件和边界条件的软约束与硬约束模型。传感器，23(5):2792，2023年3月。ISSN 1424-8220。[69] Diab W. Abueidda、Qiyue Lu 和 Seid Koric。用于三维固体力学的无网格物理信息深度学习方法。*International Journal for Numerical Meth-ods in Engineering*，122(23):7182–7201，2021年。[70] Sijun Niu、Enrui Zhang、Yuri Bazilevs 和 Vikas Srivastava。使用物理信息神经网络对有限应变塑性进行建模并评估网络性能。*Journal of the Mechanics and Physics of Solids*，172:105177，2023年。[71] Shahed Rezaei、Ahmad Moeineddin 和 Ali Harandi。利用物理信息神经网络学习非线性本构材料模型的解：Comm-pinn，2023年。[72] Shahed Rezaei、Jacob Niikoi Okoe-Amon、Cerun Alex Varkey、Armin Asheri、Hui Ruan 和 Bai-Xiang Xu。适用于化学-力学环境的内聚相场断裂模型：电池材料退化研究。《理论与应用断裂力学》，124:103758，2023年。[73] Sifan Wang、Hanwen Wang 和 Paris Perdikaris。利用物理信息DeepONet学习参数化偏微分方程的解算子。《科学进展》，7: eabi8605，2021年。