arXiv:2602.07192v1 [cs.LG] 6 Feb 2026

深度材料网络在多尺度材料建模中的系统性能评估

何小龙[a,*]，魏浩然[a]，胡伟[a]，毛河南[b]，C. T. Wu[a]

[a]计算与多尺度力学组，*Synopsys Inc.*，美国加利福尼亚州利弗莫尔市，邮编 *94551*[b]*Synopsys Inc.*，美国华盛顿州雷德蒙德市，邮编 *98052*

摘要

深度材料网络（DMN）是一类保结构的机制型机器学习模型，其架构融入了微观力学原理，因此具备强大的外推能力，并有望显著加速复杂微观结构的多尺度建模。这类模型的一项关键优势是，只需使用线弹性数据进行训练，便可在在线预测阶段推广至非线性非弹性状态。尽管其应用日益广泛，但针对其完整离线—在线流程性能的系统评估仍然有限。本文从预测准确性、计算效率和训练鲁棒性等方面，对 DMN 进行了全面的比较评估。我们研究了离线训练选择的影响，包括初始化、批量大小、训练数据量和激活正则化对在线泛化性能及不确定性的影响。结果表明，随着训练数据量增加，预测误差和方差均会降低，而初始化和批量大小则会显著影响模型性能。此外，研究发现，激活正则化在控制网络复杂度、进而影响泛化性能方面发挥着关键作用。与原始 DMN 相比，无旋转的基于相互作用的材料网络（IMN）方案在离线训练中实现了 $3.4\times$- $4.7\times$ 倍加速，同时保持了相当的在线预测精度和计算效率。这些发现阐明了保结构材料网络在模型表达能力与效率之间的关键权衡，并为

*通讯作者电子邮箱：xiaolong.he@synopsys.com (Xiaolong He)

其在多尺度材料建模中的应用提供了实用指导。

关键词：深度材料网络，多尺度建模，降阶建模，均匀化，机理机器学习

---

1. 引言

经过精心设计微观结构的现代材料系统，为在广泛的工程应用中定制具有优异力学性能和功能的材料带来了新机遇。在复合材料、多晶材料、架构材料和多功能系统等各类材料中，多尺度建模已成为将微观尺度机制与宏观响应联系起来的核心范式，从而能够跨越不同长度和时间尺度进行预测性模拟与设计 [1–4]。对这些多尺度材料系统进行高效、准确的计算建模，不仅对于提升大规模结构分析的预测能力至关重要，也有助于加速先进材料的发现、优化与设计。

由于非均质材料具有多尺度特性，其宏观性能受到微观形貌、组分性能和相互作用的显著影响。尽管对大型非均质系统进行直接数值模拟（DNS）能够提供高保真预测，但其计算成本往往高得难以承受。因此，人们开发了多尺度建模方法，以衔接微观尺度行为与宏观响应。基于代表性体积单元（RVE）概念的均匀化 [5] 已成为多尺度材料建模的基石。可以采用有限元法（FEM） [2, 6] 和快速傅里叶变换方法 [7–9]等高保真DNS方法对RVE进行数值求解。数值RVE模型可与结构级有限元模型耦合，以对非均质结构进行建模 [2, 6, 10–15]。然而，基于RVE的DNS计算成本高昂，仍是主要瓶颈，限制了其在大规模工业问题中的应用 [16]。

机器学习（ML）和数据科学的最新进展为开发数据驱动材料建模和多尺度模拟方法带来了新的机遇 [4, 17–23]。无模型的数据驱动计算力学方法旨在绕过传统本构建模，直接在表征本构行为的材料数据库上求解物理约束优化问题，

以此实现上述目标 [24–33]。然而，这些方法在模拟非线性非弹性材料行为时面临挑战。与此同时，研究人员提出了基于机器学习的替代模型，以直接近似非线性本构定律 [22, 23, 34–40]。然而，由于缺乏明确的微观尺度物理，这类数据驱动代理通常需要大量训练数据，而且对未见过的材料系统和加载路径的外推能力往往有限。

为克服这些挑战并加速多尺度材料建模，研究人员开发了多种降阶建模技术，作为高保真模拟的替代方案，包括本征正交分解 [41–44]，自洽聚类分析 [45, 46]，以及一种称为深度材料网络（DMN）的机理性保结构机器学习模型 [47, 48]。保结构机器学习是指在模型架构、损失函数或训练过程中，显式嵌入已知数学、物理或几何结构（如守恒定律、对称性、约束、不变量和稳定性性质）的模型，而非单纯依赖对数据进行统计拟合。通过引入此类归纳偏置，保结构模型能够提高物理一致性、泛化能力和稳定性，同时确保所学习的替代模型遵循目标系统的底层结构，从而降低数据需求 [49–55]。从这个意义上说，DMN可视为一种保结构机器学习模型，因为其架构和组件显式编码了源自微观力学和经典均匀化理论的机理结构。具体而言，深度材料网络由一系列层级化的机理构建模块组成，这些模块具有解析均匀化解和物理上可解释的参数。因此，仅使用线弹性均匀化数据训练的深度材料网络，便能有效外推至非线性非弹性复合材料响应，同时保持热力学一致性，并继承其微观组分的应力-应变单调性 [47, 48, 56]。

受Liu等人原始工作的启发，[47, 48]，人们提出了DMN的紧凑变体，包括无旋转深度材料网络 [56] 和基于相互作用的材料网络（IMN） [57]。这些模型采用更高效的材料取向参数化方法，从而在保留基于微观力学且保持结构的均匀化理念的同时，减少独立参数的数量。在本研究中，我们将这些紧凑形式统称为IMN。DMN和IMN已成功应用于广泛的材料和结构，包括多相复合材料 [56, 58, 59]，机织结构 [60]，短-

纤维增强复合材料 [61–63], 多晶材料 [64], 微极复合材料本构模型 [65], 以及不可压缩非牛顿流体中的纤维悬浮液 [66]。与此同时，这些材料网络已得到扩展，可用于捕捉黏聚界面失效 [67], 应变局部化 [68],机织复合材料的热响应 [69], 耦合热-力行为 [70], 以及内部微观结构应力与应变分配 [71]。为高效建模具有不同微观结构（例如不同纤维取向或体积分数）的复合材料而无需重新训练，DMN结合基于物理信息描述符的迁移学习策略 [63, 72–74], ，并集成了基于图的 [75] 和基于Transformer的特征表示 [76], 以及基于微观力学的参数化网络架构 [77]。此外，人们还探索了初始化[78]和采样技术 [57, 59, 62, 79]等策略，以提高离线训练效率和在线预测精度。这些保持结构的材料网络作为材料微观结构的高效代理模型，已与有限元法结合，用于多尺度结构仿真，其速度相比基于直接数值模拟的方法提高了几个数量级[58, 61, 63, 73]。DMN已集成到多物理场仿真软件LS-DYNA中，用于非线性多尺度建模 [63]。

尽管取得了这些进展，但针对 DMN 和 IMN 在完整离线—在线流程中的系统性直接评估仍然有限，评估范围包括训练鲁棒性、计算成本和在线推理效率。本文通过从三个实用维度对 DMN 和 IMN 进行全面的性能评估，弥补了这一空白：(i) 对离线训练选择的敏感性，包括初始化、批量大小、训练数据量和激活正则化；(ii) 离线训练效率；以及 (iii) 在线预测精度，以及宏观材料点层面不同迭代求解器的计算成本。

本文其余部分安排如下。第 2 节回顾 DMN 和 IMN 的理论基础，包括其结构保持型机理构件、离线训练流程和在线预测算法。第 3 节考察初始化、批量大小、训练数据量和激活正则化对训练和泛化性能的影响。评估 IMN 的两种在线预测方案，包括不动点迭代方案和牛顿迭代；随后从离线训练成本和在线预测性能两方面对 DMN 和 IMN 进行系统比较。最后，第 4 节给出总结性评论与讨论。

2. 方法

本节回顾 DMN 和 IMN 的结构保持型机理构件，并总结离线训练和在线预测的公式及流程。

2.1. 深度材料网络（*DMN*）

深度材料网络（DMN） [47, 48] 是一种具有机理性且保持结构的机器学习模型，它嵌入了微观力学原理，用于数据驱动的多尺度建模。图 1 展示了采用二叉树网络结构构建的三层 DMN。一个 N 层的二叉树网络总共包含 $2^{N+1}-1$ 个节点，其中 2^N 个节点位于底层 N。这些底层节点称为基节点。

对于第 n 层中的第 i 个节点，定义了四个参数：一个节点权重w_i^n以及三个相互独立的欧拉角 $\{\alpha_i^n, \beta_i^n, \gamma_i^n\}$。其中，$0 \le i \le N$ 表示层索引，$1 \le n \le 2^i$ 表示层 i 内的节点索引，而$i = 0$ 对应于网络的输出层。基节点的权重通过修正线性单元（ReLU）激活，

$$w_N^n = ReLU(z^n) = max(z^n, 0), \quad 1 \le n \le 2^N, \tag{1}$$

其中 z^n 表示第 n个基节点的激活。父节点的节点权重w_i^n 计算为其两个子节点权重之和，

$$w_i^n = w_{i+1}^{2n-1} + w_{i+1}^{2n}, \quad 0 \le i \le N-1, \quad 1 \le n \le 2^i. \tag{2}$$

因此，DMN中的所有节点权重均可由基节点的激活 $\{z^n\}_{n=1}^{2^N}$计算得到。由此，DMN的独立参数由 2^N 个激活和 $3\times(2^{N+1}-1)$ 个欧拉角组成，共计 $7\times2^N-3$个参数。与纯数据驱动代理模型不同，这些参数具有明确的物理解释，其来源于保持均匀化物理特性的机理构件。

父节点的平均应力 $\bar{\boldsymbol{\sigma}}_i^n$ 可通过混合定则，利用其两个子节点的节点权重和应力得到，

$$\bar{\boldsymbol{\sigma}}_i^n = f_{i+1}^{2n-1}\boldsymbol{\sigma}_{i+1}^{2n-1} + f_{i+1}^{2n}\boldsymbol{\sigma}_{i+1}^{2n}, \tag{3}$$

其中体积分数定义为

$$f_{i+1}^{2n-1} = \frac{w_{i+1}^{2n-1}}{w_{i+1}^{2n-1} + w_{i+1}^{2n}}, \quad f_{i+1}^{2n} = \frac{w_{i+1}^{2n}}{w_{i+1}^{2n-1} + w_{i+1}^{2n}}. \tag{4}$$

¯因此，父节点的平均材料刚度 $\mathbf{C}_i^n$ 表示为其节点权重和两个子节点材料刚度的函数，

$$\bar{\mathbf{C}}_i^n = \mathbf{C}_{i+1}^{2n} - f_{i+1}^{2n-1}(\mathbf{C}_{i+1}^{2n} - \mathbf{C}_{i+1}^{2n-1})\boldsymbol{\mathcal{A}}, \tag{5}$$

其中$\boldsymbol{\mathcal{A}}$ 是 DMN构建模块中的应变集中矩阵，通过满足两相复合材料的界面平衡和运动学相容性推导得出 [48]。$\boldsymbol{\mathcal{A}}$ 的详细推导过程和表达式见 [48]。

为捕捉复杂微观结构的材料方向性行为，在网络的每个节点处定义欧拉角 $\{\alpha_i^n, \beta_i^n, \gamma_i^n\}$。对平均刚度和应力施加三维旋转（3D），即可得到旋转后的刚度和应力，

$$\mathbf{C}_i^n = \mathbf{R}^T(\alpha_i^n, \beta_i^n, \gamma_i^n)\bar{\mathbf{C}}_i^n\mathbf{R}(\alpha_i^n, \beta_i^n, \gamma_i^n), \tag{6}$$

$$\boldsymbol{\sigma}_i^n = \mathbf{R}(\alpha_i^n, \beta_i^n, \gamma_i^n)\bar{\boldsymbol{\sigma}}_i^n, \tag{7}$$

其中$\mathbf{R}(\alpha_i^n, \beta_i^n, \gamma_i^n)$是基于欧拉角的旋转矩阵。请注意，在将基节点刚度均匀化为父节点刚度之前，需要先对其进行旋转。

总之，如图 2(a) 所示，每个机理型 DMN构建模块都包含两种基本操作：平均和旋转，这两种操作保留了微观力学均匀化的基本结构。在整个网络中递归地对物理量（如应力、应变和刚度）应用这些操作，DMN便能有效地将基节点的微观材料行为均匀化，从而预测顶层的宏观复合材料响应。

2.1.1. 离线训练

令 $\boldsymbol{\psi} = \{\mathbf{z}\ \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}\}$ 表示 DMN 的独立可训练参数集合，其中 $\mathbf{z} = \{z^n\}_{n=1}^{2^N}$, $\boldsymbol{\alpha} = \{\alpha_i^n\}_{n=1}^{2^i}$, $\boldsymbol{\beta} = \{\beta_i^n\}_{n=1}^{2^i}$, $\boldsymbol{\gamma} = \{\gamma_i^n\}_{n=1}^{2^i}$,$0 \leq i \leq N$。两相复合材料在顶层节点（$i = 0$）处的均匀化材料刚度可表示为

$$\hat{\mathbf{C}} = \mathbf{C}_0^1 = \mathcal{H}(\tilde{\mathbf{C}}^{p1}, \tilde{\mathbf{C}}^{p2}, \boldsymbol{\psi}), \tag{8}$$

其中，$\tilde{\mathbf{C}}^{p1}$ 和 $\tilde{\mathbf{C}}^{p2}$ 表示两种基体材料的刚度矩阵，它们被指定给 DMN 的基节点，如图 1 所示。具体而言，奇数索引的基节点被指定为 $\tilde{\mathbf{C}}^{p1}$，而基节点

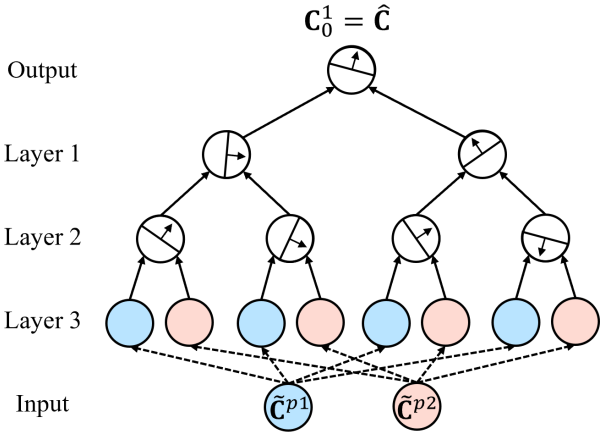


图1：三层深度材料网络的架构。

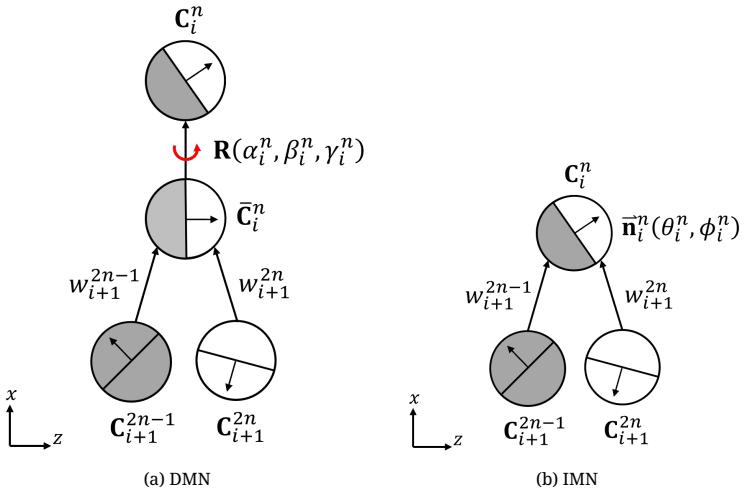


图2：(a) DMN；(b) IMN 的机理构建模块。

偶数索引对应的项被赋值为 $\tilde{\mathbf{C}}^{p2}$。算子 $\mathcal{H}$ 表示一种非线性保结构映射，该映射由从基节点到顶层节点、贯穿所有 DMN 层的一系列平均和旋转操作诱导得到。

通过最小化以下损失函数来确定 DMN 参数，

$$\mathcal{L}(\boldsymbol{\psi}) = \frac{1}{2N_s}\sum_{j=1}^{N_s}\frac{\left\|\mathbf{C}_j - \hat{\mathbf{C}}_j\right\|^2}{\left\|\mathbf{C}_j\right\|^2} + \eta\Big(\sum_{n=1}^{2^N} ReLU(z^n) - \xi\Big)^2, \tag{9}$$

其中，$\|\cdot\|$ 表示 Frobenius 范数，$\hat{\mathbf{C}}_j = \mathcal{H}(\tilde{\mathbf{C}}_j^{p1}, \tilde{\mathbf{C}}_j^{p2}, \boldsymbol{\psi})$ 是材料网络对第 j 个样本的预测。第一项衡量预测的均匀化刚度矩阵与参考均匀化刚度矩阵之间的相对误差，第二项则对基节点的激活进行正则化。该正则化控制有效的网络复杂度，进而影响在线预测精度和计算效率。第 3.2 节将考察正则化参数 η 和 ξ 对在线性能的影响。训练数据集由 $\{\tilde{\mathbf{C}}_j^{p1}, \tilde{\mathbf{C}}_j^{p2}, \mathbf{C}_j\}_j^{N_s}$ 构成，其中 $\tilde{\mathbf{C}}_j^{p1}$ 和 $\tilde{\mathbf{C}}_j^{p2}$ 分别表示第 j 个样本中两种基体材料的线弹性刚度矩阵，而$\mathbf{C}j$ 是对应代表性体积单元的均匀化刚度矩阵，通过直接数值模拟或实验测量获得。此处，N_s 表示材料样本总数。

2.1.2. 在线预测

尽管 DMN 是基于线弹性数据训练的，但它能够学习微观结构相互作用的本质，并可用于预测一般加载条件下复合材料的非线性非弹性行为。给定指定的宏观应变增量，DMN 会对材料信息进行正向均匀化和反向反均匀化，以预测多尺度材料响应。第 i 层中第 n 个节点的增量应力–应变关系定义为

$$\Delta\bar{\boldsymbol{\sigma}}_i^n = \bar{\mathbf{C}}_i^n \Delta\bar{\boldsymbol{\varepsilon}}_i^n + \delta\bar{\boldsymbol{\sigma}}_i^n, \tag{10}$$

其中，$\Delta\bar{\boldsymbol{\sigma}}_i^n$、$\Delta\bar{\boldsymbol{\varepsilon}}_i^n$ 和 $\mathbf{C}_i^n$分别表示第 n层中第 i个节点的平均增量应力、增量应变和刚度。变量$\delta\bar{\boldsymbol{\sigma}}_i^n$ 表示该节点的平均残余应力，它仅在

存在材料非线性时才会产生。与其他材料量类似，残余应力的正向均匀化可通过解析方法推导为

$$\delta\bar{\boldsymbol{\sigma}}_i^n = f_{i+1}^{2n-1}\delta\boldsymbol{\sigma}_{i+1}^{2n-1} + f_{i+1}^{2n}\delta\boldsymbol{\sigma}_{i+1}^{2n} + \boldsymbol{\mathcal{X}}, \tag{11}$$

其中，$\boldsymbol{\mathcal{X}}$ 是一个修正项，取决于两个子节点的刚度矩阵、节点权重和残余应力。$\boldsymbol{\mathcal{X}}$ 的详细推导和表达式见 [48]。子节点的旋转后残余应力 $\delta\boldsymbol{\sigma}_{i+1}^{2n-1}$ 和 $\delta\boldsymbol{\sigma}_{i+1}^{2n}$由下式得到

$$\delta\boldsymbol{\sigma}_i^n = \mathbf{R}(\alpha_i^n, \beta_i^n, \gamma_i^n)\delta\bar{\boldsymbol{\sigma}}_i^n. \tag{12}$$

DMN 在线迭代从底层开始，节点增量应变为 $\Delta\bar{\boldsymbol{\varepsilon}}_N^n$。在每个基节点处，通过评估微观基体材料的本构定律并结合式 (10)，得到相应的刚度 $\mathbf{C}_N^n$和残余应力 $\delta\bar{\sigma}_N^n$。随后，通过正向均匀化（式 (5) 和 (11)）和旋转（式 (6) 和 (12)），将刚度矩阵和残余应力向前传递至顶层，从而在顶层节点得到 $\mathbf{C}_0^1$ 和 $\delta\boldsymbol{\sigma}_0^1$。给定顶层节点处的宏观加载条件，使用式 (10) 计算 $\Delta\boldsymbol{\varepsilon}_0^1$ 和 $\Delta\boldsymbol{\sigma}_0^1$，然后依据各构建块内的应力-应变关系，通过逆旋转和反向去均匀化将其向后传递至底层，从而得到基节点处更新后的节点应变增量 $\Delta\bar{\boldsymbol{\varepsilon}}_N^n$。这一过程完成了一次由正向均匀化和反向去均匀化构成的 DMN 在线迭代，如图 3 所示。

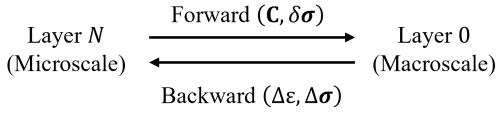


图 3：DMN 在线阶段的数据流。

深度材料网络在线迭代会持续进行，直至$\Delta\bar{\boldsymbol{\varepsilon}}_N^{n(k+1)}$ 与 $\Delta\bar{\boldsymbol{\varepsilon}}_N^{n(k)}$ 之间的相对差值足够小，其中 k 表示在线迭代计数器。收敛后，基节点处的微观尺度应力和内部状态变量（如等效塑性应变）将被更新，顶层节点处的应力增量 $\Delta\boldsymbol{\sigma}_0^1$ 则用于更新宏观尺度应力状态。需要指出的是，单次在线迭代中的操作总数与活跃基节点的数量成正比。算法 1

总结了 DMN 在线预测的流程。第 3.4 节将研究并讨论残余应力对 DMN 在线预测性能的影响。

算法 1 DMN 在线预测

输入：宏观应变增量 $\Delta\boldsymbol{\varepsilon}$
输出：宏观应力增量 $\Delta\boldsymbol{\sigma}$
初始化 $\Delta\bar{\boldsymbol{\varepsilon}}_N^{n(1)} = \mathbf{0}$
对 $k \leftarrow 1$ 到 N_{iter} **循环**
    计算刚度 $\bar{\mathbf{C}}_N^n$ 和残余应力 $\delta\bar{\boldsymbol{\sigma}}_N^n$，通过计算本构微观基材料的本构定律和公式 (10)
    正向均匀化 $\bar{\mathbf{C}}_N^n$ 和 $\delta\bar{\boldsymbol{\sigma}}_N^n$，以获得 $\mathbf{C}_0^1$ 和 $\delta\boldsymbol{\sigma}_0^1$
    通过 $\Delta\boldsymbol{\sigma}_0^1 =$计算应变增量 $\Delta\boldsymbol{\varepsilon}_0^1$ 和应力增量 $\Delta\boldsymbol{\sigma}_0^1$ $\mathbf{C}_0^1\Delta\boldsymbol{\varepsilon}_0^1 + \delta\boldsymbol{\sigma}^1{}_0$
    对应变和应力增量进行反向去均匀化，以获得 $\Delta\bar{\boldsymbol{\varepsilon}}_N^{n(k+1)}$
    检查是否满足 $\|\Delta\bar{\boldsymbol{\varepsilon}}_N^{n(k+1)} - \Delta\bar{\boldsymbol{\varepsilon}}_N^{n(k)}\|/\|\Delta\bar{\boldsymbol{\varepsilon}}_N^{n(k)}\| \leq tol.$ 如果不满足，则转到 3。
结束循环
宏观应力增量 $\Delta\boldsymbol{\sigma} = \Delta\boldsymbol{\sigma}^1{}_0$

2.2. 基于相互作用的材料网络（*IMN*）

基于相互作用的材料网络（IMN） [56, 57] 采用与 DMN 相同的二叉树网络结构。这两种模型的根本区别在于机理构件的构建方式。

在原始 DMN表述中， [47, 48], 每个构建块均基于两个子节点之间的相互作用界面定义。该界面的单位法向量最初与材料坐标系的z轴方向对齐，再通过由三个欧拉角参数化的三维旋转操作将其旋转至任意取向 $\{\alpha, \beta, \gamma\}$，如图2(a)所示。另一方面，IMN将每个构建块定义为具有相互作用界面的结构，该界面的单位法向量 $\vec{\mathbf{n}}$ 仅由两个角度参数化 $\{\theta, \phi\}$，$\vec{\mathbf{n}} = [\cos(2\pi\theta)\sin(\pi\phi), \sin(2\pi\theta)\sin(\pi\phi), \cos(\pi\phi)]$，，如图2(b)所示。因此，尽管DMN和IMN都能任意调整均匀化材料节点的取向，以捕捉材料的方向性行为，IMN所需的可训练参数更少，因为其取向参数化更为紧凑

。具体而言，在具有 N层的DMN中，每个节点对应三个独立的欧拉角，因此共有 $3\times(2^{N+1}-1)$个可训练取向参数；而具有 N层的IMN仅为每个父节点分配两个角度 $\{\theta,\phi\}$ ，因此共有 $2\times(2^{N}-1)$个可训练取向参数。加上底层的 2^N 个激活参数，具有 N层的IMN总计有 $3\times2^N-2$ 个独立可训练参数，而具有 N层的DMN则有 $7\times2^N-3$ 个。对于相同的网络深度，DMN中的可训练参数数量约为IMN的2.3倍。尽管IMN的参数效率更高，但这种模型复杂度降低对学习能力、离线训练效率和在线预测精度的影响尚不明确。本研究对DMN和IMN进行定量比较，系统评估两者之间的权衡。

下面简要回顾 IMN 的机理构建模块。为清晰起见，上标“h”表示与父节点相关的量，而上标“1”和“2”表示与其两个子节点相关的量。根据混合定则，父节点与其两个子节点之间的应变和应力关系为

$$\boldsymbol{\varepsilon}^h = f^1\boldsymbol{\varepsilon}^1 + f^2\boldsymbol{\varepsilon}^2. \tag{13}$$

$$\boldsymbol{\sigma}^h = f^1\boldsymbol{\sigma}^1 + f^2\boldsymbol{\sigma}^2, \tag{14}$$

其中 $f^1+f^2=1$。父节点与子节点之间满足能量守恒的 Hill-Mandel条件为

$$(\boldsymbol{\varepsilon}^h)^T\boldsymbol{\sigma}^h = (\boldsymbol{\varepsilon}^1)^T\boldsymbol{\sigma}^1 f^1 + (\boldsymbol{\varepsilon}^2)^T\boldsymbol{\sigma}^2 f^2. \tag{15}$$

将式 (13)–(14) 与式 (15) 结合可得

$$(\boldsymbol{\varepsilon}^1-\boldsymbol{\varepsilon}^2)^T(\boldsymbol{\sigma}^1-\boldsymbol{\sigma}^2)=0. \tag{16}$$

相互作用界面两侧的牵引连续性通过以下关系保证

$$\mathbf{H}^T(\boldsymbol{\sigma}^2-\boldsymbol{\sigma}^1)=\mathbf{0}, \tag{17}$$

其中，$\mathbf{H}$ 是与界面单位法向量$\vec{\mathbf{n}}$相关联的取向矩阵，定义为

$$\mathbf{H}(\vec{\mathbf{n}}) = \begin{bmatrix} \vec{n}_1 & 0 & 0 \\ 0 & \vec{n}_2 & 0 \\ 0 & 0 & \vec{n}_3 \\ 0 & \vec{n}_3 & \vec{n}_2 \\ \vec{n}_3 & 0 & \vec{n}_1 \\ \vec{n}_2 & \vec{n}_1 & 0 \end{bmatrix}. \tag{18}$$

结合式（16）和式（17），可得

$$(\boldsymbol{\varepsilon}^1 - \boldsymbol{\varepsilon}^2)^T = \frac{1}{f^1 f^2}(\mathbf{Hb})^T, \quad \forall \mathbf{b} \tag{19}$$

将式（13）—（14）与式（19）结合，子节点的应变可表示为均匀化应变的函数:

$$\boldsymbol{\varepsilon}^1 = \boldsymbol{\varepsilon}^h + \frac{1}{f^1}\mathbf{Hb}, \quad \boldsymbol{\varepsilon}^2 = \boldsymbol{\varepsilon}^h - \frac{1}{f^2}\mathbf{Hb} \tag{20}$$

式（13）—（20）构成了 IMN构件的力学基础。考虑父节点和子节点的线性本构关系， $\boldsymbol{\sigma}^h = \mathbf{C}^h\boldsymbol{\varepsilon}^h$, $\boldsymbol{\sigma}^1 = \mathbf{C}^1\boldsymbol{\varepsilon}^1$, $\boldsymbol{\sigma}^2 = \mathbf{C}^2\boldsymbol{\varepsilon}^2$，以及式（17）中的牵引连续条件和式（20）中的应变关系，可将$\mathbf{b}$ 向量表示为

$$\mathbf{b} = \mathbf{B}\boldsymbol{\varepsilon}^h, \quad \text{with } \mathbf{B} = f^1 f^2 \left[\mathbf{H}^T(f^2\mathbf{C}^1 + f^1\mathbf{C}^2)\mathbf{H}\right]^{-1}\mathbf{H}^T(\mathbf{C}^2 - \mathbf{C}^1), \tag{21}$$

父节点的均匀化刚度$\mathbf{C}^h$ 则可表示为

$$\mathbf{C}^h = f^1\mathbf{C}^1 + f^2\mathbf{C}^2 - (\mathbf{C}^2 - \mathbf{C}^1)\mathbf{HB}. \tag{22}$$

2.2.1. 离线训练

设 $\boldsymbol{\psi} = \{\mathbf{z}\ \boldsymbol{\theta}, \boldsymbol{\phi}\}$ 表示 IMN 的独立可训练参数集合，其中$\mathbf{z} = \{z^n\}_{n=1}^{2^N}$, $\boldsymbol{\theta} = \{\theta_i^n\}_{n=1}^{2^i}$, $\boldsymbol{\phi} = \{\phi_i^n\}_{n=1}^{2^i}$, $0 \le i \le N-1$。顶层节点处两相复合材料的均匀化材料刚度（$i = 0$）可表示为

$$\hat{\mathbf{C}} = \mathbf{C}_0^1 = \tilde{\mathcal{H}}(\tilde{\mathbf{C}}^{p1}, \tilde{\mathbf{C}}^{p2}, \tilde{\boldsymbol{\psi}}), \tag{23}$$

其中，算子 $\mathcal{H}$ 表示由从基节点到顶层节点、跨越所有IMN层应用均匀化操作所诱导的非线性保结构映射。这里， $\tilde{\mathbf{C}}^{p1}$ 和 $\tilde{\mathbf{C}}^{p2}$ 表示两种基体材料的线弹性刚度矩阵。采用ReLU激活获得底层中的节点权重，如公式（1）所示。通过最小化公式（9）中定义的损失函数来确定IMN参数，其中$\hat{\mathbf{C}}_j = \tilde{\mathcal{H}}(\tilde{\mathbf{C}}^{p1}, \tilde{\mathbf{C}}^{p2}, \tilde{\boldsymbol{\psi}})$。

2.2.2. 在线预测

不动点迭代。 类似于第2.1.2节中描述的DMN在线迭代，可以构建一种直接的IMN在线迭代方案，称为不动点迭代。该方案包括正向均匀化和反向去均匀化，如图4所示。与DMN在线算法的关键区别在于，IMN不动点迭代不显式涉及应力变量。

具体而言，IMN不动点在线迭代从底层的节点应变增量$\Delta\boldsymbol{\varepsilon}_N^n$开始。在每个基节点处，通过评估微观基材料的本构定律，得到相应的刚度矩阵$\mathbf{C}_N^n$。随后，通过正向均匀化（式（22））将刚度矩阵向前传播至顶层，从而得到顶层节点处的均匀化刚度$\mathbf{C}_0^1$。在这一正向传播过程中，还会同时确定式（21）中定义的所有父节点的矩阵$\mathbf{B}$。

给定顶层节点处规定的宏观加载条件 $\Delta\boldsymbol{\varepsilon}_0^1$，可通过式（20）和式（21）进行反均匀化，将其反向传播至底层，从而得到底层更新后的节点应变增量 $\Delta\boldsymbol{\varepsilon}_N^n$。这一过程完成一次IMN不动点在线迭代。

重复进行IMN不动点迭代，直至 $\Delta\boldsymbol{\varepsilon}_N^{n(k+1)}$ 与 $\Delta\boldsymbol{\varepsilon}_N^{n(k)}$ 之间的相对差值低于规定的容差，其中 k 表示迭代计数器。收敛后，根据各基节点的本构定律更新微观尺度应力，并使用顶层节点处得到的应力增量 $\Delta\boldsymbol{\sigma}_0^1$更新宏观尺度应力状态。算法2总结了使用不动点迭代进行IMN在线预测的完整流程。

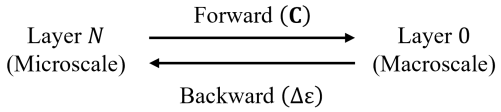


图4：IMN 定点在线阶段的数据流。

牛顿迭代。 根据式（20）中给出的父节点应变与其子节点应变之间的关系，底层的节点应变可表示为宏观应变的

算法2 采用不动点迭代的IMN在线预测

输入：宏观应变增量 $\Delta\boldsymbol{\varepsilon}$输出：宏观应力增量 $\Delta\boldsymbol{\sigma}$

1 : 初始化 $\Delta\boldsymbol{\varepsilon}_N^{n(1)}=\mathbf{0}$2: **对** $k\leftarrow 1$ 到 N_{iter} **执行**3: 通过评估微观基材料的本构定律，计算刚度$\mathbf{C}_N^n$ 4: 将刚度$\mathbf{C}_N^n$ 进行正向均匀化，以获得顶层节点的$\mathbf{C}_0^1$ 以及所有父节点的$\mathbf{B}_i^n$ 5: 对宏观应变增量 $\Delta\boldsymbol{\varepsilon}$ 进行反向去均匀化，以获得 $\Delta\boldsymbol{\varepsilon}_N^{n(k+1)}$ 6: 检查 $\|\Delta\boldsymbol{\varepsilon}_N^{n(k+1)}-\Delta\boldsymbol{\varepsilon}_N^{n(k)}\|/\|\Delta\boldsymbol{\varepsilon}_N^{n(k)}\|\leq tol.$。如果不是，则转到第 3 步。7: **结束循环**8: 通过评估微观基材料的本构定律，计算应力增量 $\Delta\boldsymbol{\sigma}_N^n$ 9: 计算宏观应力增量: $\Delta\boldsymbol{\sigma}=\sum_{n=1}^{2^N}w_N^n\Delta\boldsymbol{\sigma}_N^n$

顶层节点，$\boldsymbol{\varepsilon}_0^1$，作为

$$\boldsymbol{\varepsilon}_N^n=\boldsymbol{\varepsilon}_0^1+\sum_{i=1}^{N}\frac{\mathbf{H}\left(\vec{\mathbf{n}}_{i-1}^{1+r_{N-i+1}^n}\right)}{w_i^{1+r_{N-i}^n}}(-1)^{r_{N-i}^n}\left(w_{i-1}^{1+r_{N-i+1}^n}\mathbf{b}_{i-1}^{1+r_{N-i+1}^n}\right),\quad n=1,...,2^N, \tag{24}$$

其中 $r_i^n=\mathrm{mod}\,(n-1,2^i)$。定义向量 $\boldsymbol{\varepsilon}^{\mathrm{nodes}}=[\boldsymbol{\varepsilon}_N^1,\boldsymbol{\varepsilon}_N^2,...,\boldsymbol{\varepsilon}_N^2\,]^T$和 $\boldsymbol{\varepsilon}^{\mathrm{macro}}=[\boldsymbol{\varepsilon}_0^1,\boldsymbol{\varepsilon}_0^1,...,\boldsymbol{\varepsilon}_0^1]^T$，式 (24) 中的方程组可写为

$$\boldsymbol{\varepsilon}^{\mathrm{local}}=\boldsymbol{\varepsilon}^{\mathrm{nodes}}-\boldsymbol{\varepsilon}^{\mathrm{macro}}=\begin{bmatrix}\mathbf{A}_{1,1}&\mathbf{A}_{1,2}&\dots&\mathbf{A}_{1,2^N-1}\\\mathbf{A}_{2,1}&\mathbf{A}_{2,2}&\dots&\mathbf{A}_{2,2^N-1}\\\vdots&\vdots&\vdots&\vdots\\\mathbf{A}_{2^N,1}&\mathbf{A}_{2^N,2}&\dots&\mathbf{A}_{2^N,2^N-1}\end{bmatrix}\begin{bmatrix}\mathbf{a}_1\\\mathbf{a}_2\\\vdots\\\mathbf{a}_{2^N-1}\end{bmatrix}=\mathbf{A}\mathbf{a}, \tag{25}$$

其中 $i=1,...,N,\quad n=1,...,2^N$，以及

$$\begin{aligned}\mathbf{a}_{2^{i-1}+j-1} &= w_{i-1}^{j}\mathbf{b}_{i-1}^{j},\quad \text{for } j=1,...,2^{i-1},\\ \mathbf{A}_{n,2^{i-1}+j-1} &= \frac{\mathbf{H}\left(\vec{\mathbf{n}}_{i-1}^{1+r_{N-i+1}^{n}}\right)}{w_i^{1+r_{N-i}^{n}}}(-1)^{r_{N-i}^{n}},\quad \text{for } j=1+r_{N-i+1}^{n},\\ \mathbf{A}_{n,2^{i-1}+j-1} &= \mathbf{0}_{6\times 3},\quad \text{for } j\neq 1+r_{N-i+1}^{n}.\end{aligned}\tag{26}$$

向量**a** 称为跳跃向量，**A**表示代表 IMN架构的加权取向矩阵。对于给定的宏观应变 $\boldsymbol{\varepsilon}_0^1$，通过最小化残差向量来确定跳跃向量**a**，该残差向量用于确保 IMN 所有父节点处的界面牵引平衡。残差向量定义为

$$\mathbf{R}=\begin{bmatrix}\mathbf{R}_1\\ \mathbf{R}_2\\ \vdots\\ \mathbf{R}_{2^N-1}\end{bmatrix}=\begin{bmatrix}\mathbf{H}(\vec{\mathbf{n}}_0^1)^T(\boldsymbol{\sigma}_1^1-\boldsymbol{\sigma}_1^2)\\ \mathbf{H}(\vec{\mathbf{n}}_1^1)^T(\boldsymbol{\sigma}_2^1-\boldsymbol{\sigma}_2^2)\\ \vdots\\ \mathbf{H}(\vec{\mathbf{n}}_{N-1}^{2^{N-1}})^T(\boldsymbol{\sigma}_N^{2^N-1}-\boldsymbol{\sigma}_N^{2^N})\end{bmatrix}.\tag{27}$$

利用式 (14) 中的应力均匀化关系，残差向量**R**可表示为基节点处应力的函数:

$$\mathbf{R}=\mathbf{A}^T\mathbf{W}\begin{bmatrix}\boldsymbol{\sigma}_N^1\\ \boldsymbol{\sigma}_N^2\\ \vdots\\ \boldsymbol{\sigma}_N^{2^N}\end{bmatrix}=\mathbf{A}^T\mathbf{W}\boldsymbol{\sigma}^{\text{nodes}}=\mathbf{A}^T\mathbf{W}\mathbf{K}\boldsymbol{\varepsilon}^{\text{nodes}}\tag{28}$$

其中，分块对角矩阵**W** 和**K** 定义为

$$\mathbf{W}=\begin{bmatrix}\mathbf{W}^1 & \mathbf{0}_{6\times6} & \dots & \mathbf{0}_{6\times6}\\ \mathbf{0}_{6\times6} & \mathbf{W}^2 & \dots & \mathbf{0}_{6\times6}\\ \vdots & \vdots & \vdots & \vdots\\ \mathbf{0}_{6\times6} & \mathbf{0}_{6\times6} & \dots & \mathbf{W}^{2^N}\end{bmatrix},\mathbf{K}=\begin{bmatrix}\mathbf{C}^1 & \mathbf{0}_{6\times6} & \dots & \mathbf{0}_{6\times6}\\ \mathbf{0}_{6\times6} & \mathbf{C}^2 & \dots & \mathbf{0}_{6\times6}\\ \vdots & \vdots & \vdots & \vdots\\ \mathbf{0}_{6\times6} & \mathbf{0}_{6\times6} & \dots & \mathbf{C}^{2^N}\end{bmatrix}.\tag{29}$$

这里，$\mathbf{W}^n = w_N^n\mathbf{I}_{6\times6}$ 和$\mathbf{C}^n$ 分别表示第 n 个基节点的权重矩阵和切线刚度矩阵。

如果某个基节点的权重为零，则相应的残差未定义，必须将其从残差向量**R**中移除。相应地，相关的行

与这些基节点对应的行以及与其父节点对应的列都将从基体$\mathbf{A}$ 中移除，以确保约化系统的一致性。

结合式 (25) 和 (28) 得到

$$\mathbf{R} = \mathbf{A}^T\mathbf{W}\mathbf{K}(\mathbf{A}\mathbf{a} + \boldsymbol{\varepsilon}^{\mathrm{macro}}). \tag{30}$$

因此，残差相对于跳跃向量的雅可比矩阵为

$$\mathbf{J} = \frac{d\mathbf{R}}{d\mathbf{a}} = \mathbf{A}^T\mathbf{W}\mathbf{K}\mathbf{A}, \tag{31}$$

该矩阵用于牛顿型迭代方案中，以更新$\mathbf{a}$ ，直至残差范数 $\|\mathbf{R}\|$ 低于给定容差。基于牛顿迭代的完整 IMN 在线预测流程总结于算法 3 中。

算法 3 采用牛顿迭代的 IMN 在线预测

输入：宏观应变增量 $\Delta\boldsymbol{\varepsilon}$输出：宏观应力增量 $\Delta\boldsymbol{\sigma}$1：初始化 a $= \mathbf{0}$, R $= \mathbf{0}$2：构造 IMN 的 A 和 W 矩阵 3：当 $\|\mathrm{R}\| > tol.$ 时 4：根据式 (25) 计算局部应变增量： $\Delta\boldsymbol{\varepsilon}^{\mathrm{local}}$ =Aa 5：更新基节点处的应变增量： $\Delta\boldsymbol{\varepsilon}^{\mathrm{nodes}} = \Delta\boldsymbol{\varepsilon}^{\mathrm{macro}} + \Delta\boldsymbol{\varepsilon}^{\mathrm{local}}$6：根据微观基材料的本构定律，计算基节点处的应力增量和切线刚度： $\Delta\boldsymbol{\sigma}^{\mathrm{nodes}}$, K 7：根据式 (28) 计算残差：R $=$ $\mathrm{A}^T\mathrm{W}\Delta\boldsymbol{\sigma}^{\mathrm{nodes}}$8：计算雅可比矩阵：J $=\mathrm{A}^T\mathrm{WKA}$ 9：计算跳跃向量增量： δ a $= -\mathrm{J}^{-1}\mathrm{R}$ 10：更新跳跃向量：a $=$ a $+ \delta$a 11：结束循环 12：计算宏观应力增量： $\Delta\boldsymbol{\sigma} = \sum_{n=1}^{2^N} w_N^n \Delta\boldsymbol{\sigma}_N^n$

3. 实验

在本节中，我们研究初始化、批量大小、训练数据量和正则化如何影响模型训练与泛化性能。

我们评估了两种IMN在线预测算法，即第2.2.2节所述的不动点迭代方案和牛顿迭代方案。此外，还从离线训练成本和在线预测性能方面对DMN与IMN进行了比较。

为评估模型的泛化能力和鲁棒性，选取三种纤维体积分数为 60% 的单向纤维增强复合材料（如图 5 所示）进行测试。这些复合材料在相衬度和非弹性材料行为方面存在差异。基体和纤维相的材料参数汇总于表 1–2 中。在本研究中，所有直接数值模拟均使用 LS-DYNA，并通过专用的 RVE分析模块完成。使用 *RVE_ANALYSIS_FEM 关键字 [80], 可自动施加周期性边界条件并计算均匀化应力。六种独立加载工况下的均匀化材料响应如图 6 所示，包括沿三个正方向的单轴加载（图 6(a)–(c)）以及三种简单剪切加载工况（图 6(d)–(f)）。用于测试的三种复合材料定义如下。

- **复合材料 1**：基体：具有线性与指数组合硬化的各向同性弹塑性材料；纤维：正交各向异性线弹性材料。

- **复合材料 2**：基体：采用线性与指数组合硬化的各向同性弹塑性材料；纤维：各向同性线弹性材料。

- **复合材料 3**：基体：采用指数硬化的各向同性弹塑性材料；纤维：正交各向异性线弹性材料。

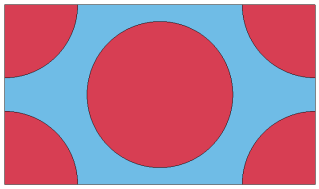


图 5：纤维体积分数为 60% 的单向纤维增强复合材料代表性体积单元。

在线预测精度通过六种加载工况下均质化应力的平均相对误差进行量化，定义如下：

$$e_{\boldsymbol{\sigma}} = \frac{1}{6}\sum_{i=1}^{6}\frac{\|\boldsymbol{\sigma}_i - \hat{\boldsymbol{\sigma}}_i\|_{L_2}}{\|\boldsymbol{\sigma}_i\|_{L_2}}, \tag{32}$$

其中， $\boldsymbol{\sigma}_i$ 表示第 i种加载条件下的参考均质化应力响应，$\hat{\boldsymbol{\sigma}}_i$ 为相应的模型预测。

Ta表 1：用于在线测试的三种复合材料的基体相材料参数。

	复合材料 1	复合材料 2	复合材料 3
杨氏模量（GPa）	3.8	2.1	3.8
泊松比	0.387	0.3	0.387
初始屈服强度（GPa）	0.01	0.029	0.03

表 2：在线测试中三种复合材料的纤维相材料参数 .

	复合材料 1	复合材料 2	复合材料 3
E_{11} （GPa）	19.8	72	19.8
E_{22} (GPa)	19.8	72	19.8
E_{33} (GPa)	245	72	245
G_{12} (GPa)	5.9	29.5	5.9
G_{13} (GPa)	29.2	29.5	29.2
G_{23} (GPa)	29.2	29.5	29.2
ν_{12}	0.67	0.22	0.67
ν_{13}	0.02	0.22	0.02
ν_{23}	0.02	0.22	0.02

DMN 和 IMN 模型均使用 PyTorch [81] 实现，并采用公式（9）中定义的相同损失函数进行训练。所有模型均使用 Adam优化器训练 10,000 训练轮次，初始学习率为 10^{-2}。如果验证损失连续 50 个训练轮次未改善，则采用学习率调度器将学习率降低至原来的 0.8 倍。所有离线训练均在 Intel(R) Xeon(R) W7-3465X CPU 的 20 个核心上进行，所有在线预测则均在第 11 代 Intel(R) i7-11850H CPU 上执行。

3.1. 初始化、批量大小和数据规模的影响

为研究初始化、批量大小和训练数据量对模型训练和泛化性能的影响，使用三个具有不同样本量和批量配置的数据集对深度材料网络进行训练：

- **数据集1**：训练样本数：400；验证样本数：100；批量大小：20、40、128

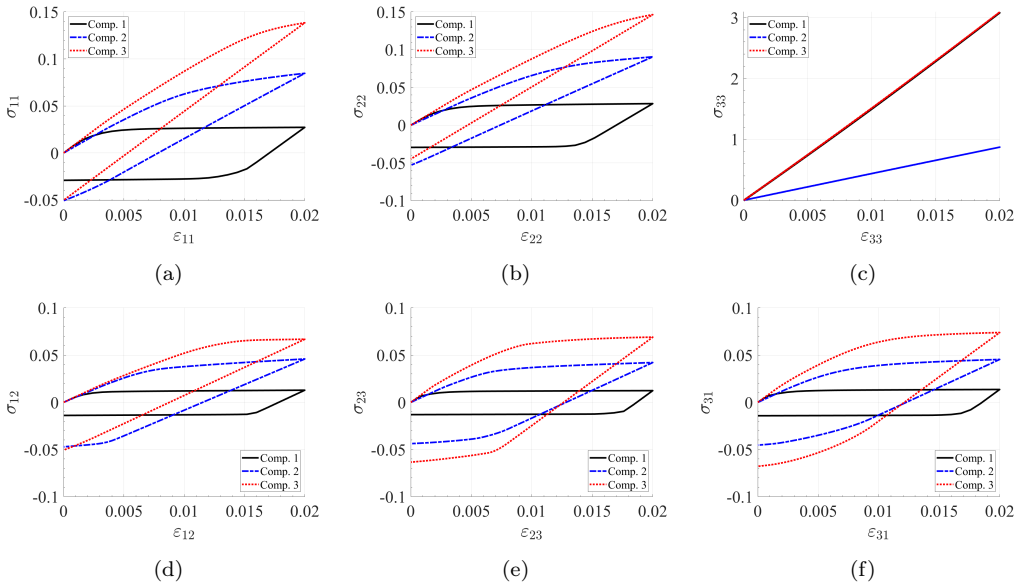


图6：在线测试中三种复合材料的材料响应：(a) $\varepsilon_{11} \sim \sigma_{11}$；(b)$\varepsilon_{22} \sim \sigma_{22}$；(c) $\varepsilon_{33} \sim \sigma_{33}$；(d) $\varepsilon_{12} \sim \sigma_{12}$；(e) $\varepsilon_{23} \sim \sigma_{23}$；(f) $\varepsilon_{31} \sim \sigma_{31}$。

- **数据集2**：训练样本数：1024；验证样本数：100；批量大小：40、64、128
- **数据集3**：2048 个样本用于训练；100 个样本用于验证；批量大小：128、256

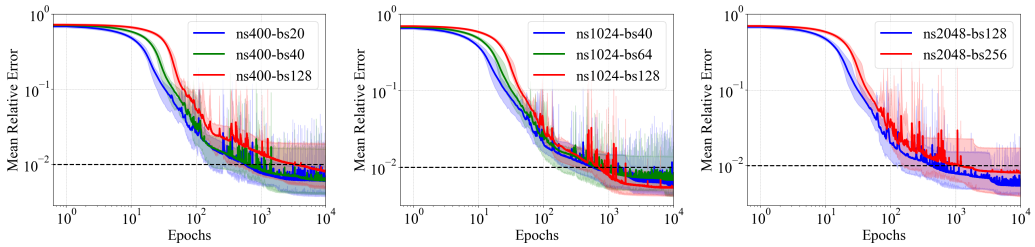
训练数据由 $\{\tilde{\mathbf{C}}_j^{p1}, \tilde{\mathbf{C}}_j^{p2}, \mathbf{C}_j\}_{j=1}^{N_s}$组成，其中 $\tilde{\mathbf{C}}_j^{p1}$和 $\tilde{\mathbf{C}}_j^{p2}$是第$j$个样本中两种组成基材料的线弹性刚度矩阵，而$\mathbf{C}j$是代表性体积单元（RVE）对应的均匀化刚度矩阵。本研究中，基体材料刚度矩阵 $\{\tilde{\mathbf{C}}^{p1}, \tilde{\mathbf{C}}^{p2}\}$ 按照 [48], 提出的基于正交各向异性弹性张量的方法进行采样。RVE 的均匀化刚度矩阵$\mathbf{C}$通过对 RVE 进行直接数值模拟（DNS）获得。

对于每种训练数据量与批量大小的组合，均采用 10 次随机初始化训练 DMN。总计训练了 80 个 DMN 模型，随后通过在线测试评估这些模型，测试对象为三种基材料属性未包含在训练集中的复合材料。结构保持深度材料网络的一项重要优势是，尽管它仅使用线弹性数据进行训练，却能准确预测由新基体材料构成的复合材料的非线性非弹性响应。

图 7 展示了式（33）中定义的平均相对误差的训练历史。实线表示 10 次随机初始化运行的均值，阴影区域表示相应的标准差，反映了由初始化带来的不确定性。

$$e_{\mathbf{C}} = \frac{1}{N_s} \sum_{j=1}^{N_s} \sqrt{\frac{||\mathbf{C}_j - \hat{\mathbf{C}}_j||^2}{||\mathbf{C}_j||^2}}, \tag{33}$$

由于验证误差与训练误差高度吻合，因此此处仅报告后者。结果表明，对于数据集1，批量大小为20或40时误差最低；对于数据集2和数据集3，批量大小为128时效果最佳。尽管未观察到数据规模与批量大小之间存在单调关系，但这些结果表明，较小的数据集往往受益于较小的批量大小，而较大的数据集可能更适合较大的批量大小。



(a) 数据集1　(b) 数据集2　(c) 数据集3

图7：采用以下设置训练的DMN的平均相对误差训练历史：(a) 使用400个样本，批量大小分别为20、40和128；(b) 使用1024个样本，批量大小分别为40、64和128；(c) 使用2048个样本，批量大小分别为128和256。实线表示采用随机初始化运行10次的均值，阴影区域表示相应的标准差。

图 8–10 将在 400 个样本上训练的深度材料网络的应力预测与三种测试复合材料的直接数值模拟结果进行了比较。结果表明，初始化和批量大小均对预测精度具有显著影响。阴影区域量化了由这些效应所引起的预测不确定性。在所测试的批量大小（20、40 和 128）中，批量大小为 40 时，在所有测试复合材料上均给出了最准确且不确定性最小的预测，尤其是在 11、22 和 12 方向上。该观察结果与图 7(a) 所示的平均相对误差相一致。使用 1024 和

附录 A 和附录 B 中提供了 2048 个样本，其预测精度和不确定性趋势也与图 7 所示的平均相对误差一致。

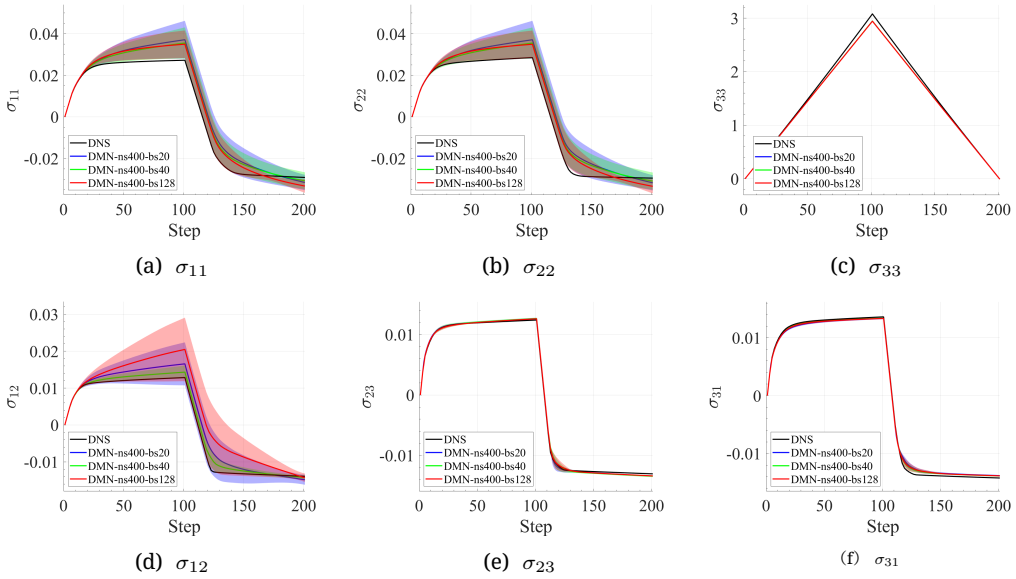


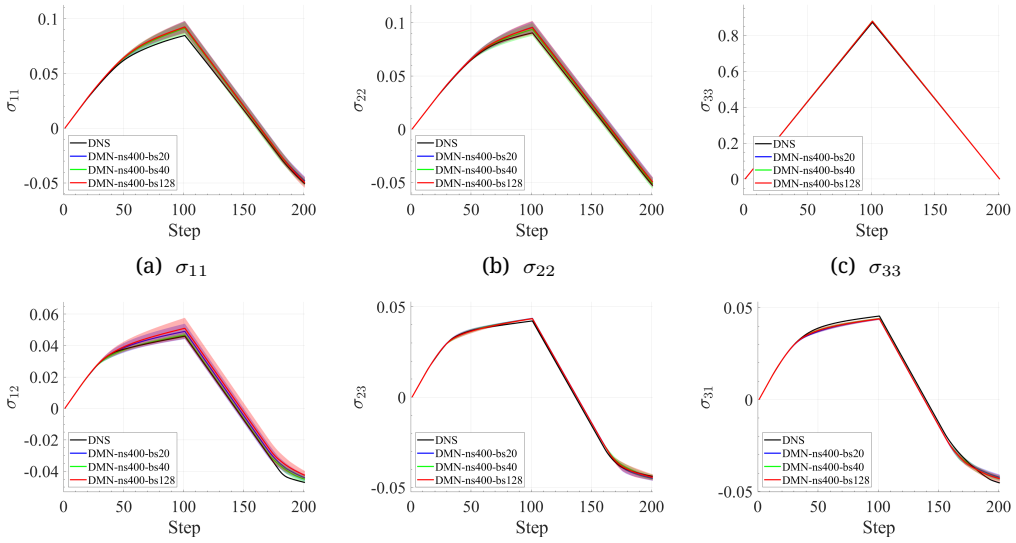
图 8：复合材料 1——使用 400 个样本训练的深度材料网络在批量大小为 20、40 和 128 时的应力预测：(a) σ_{11}；(b) σ_{22}；(c) σ_{33}；(d) σ_{12}；(e) σ_{23}；(f) σ_{31}；实线表示 10 次随机初始化运行结果的均值，阴影区域表示相应的标准差。

最后，图 11 汇总了全部 80 个深度材料网络模型的平均相对误差，涵盖所有训练数据规模、批量大小、初始化运行次数和测试复合材料。总体而言，结果表明，增大训练数据量可提高预测精度并降低不确定性。

3.2 正则化的影响

本节考察公式（9）中定义的损失函数里的正则化项如何影响在线预测性能。首先使用 $\xi = 1$ 训练隐式神经网络模型，同时将 η 从 10^{-2} 逐步调整至 10^{2}。为考虑初始化带来的差异，每种（ξ，η）配置均采用 5 次随机初始化进行训练。

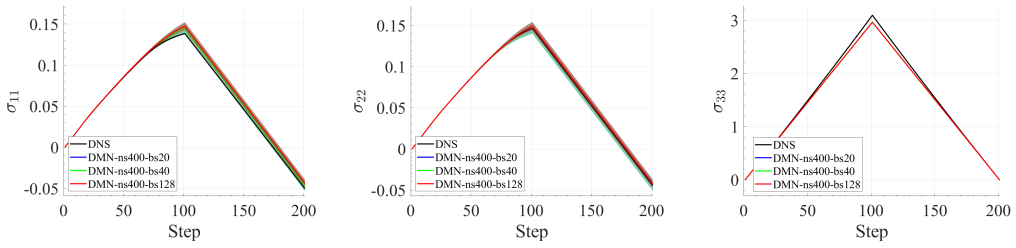
图 12(a) 显示，较小的 η 通常会带来更低的预测误差，该误差由公式（32）定义的度量指标衡量；最低预测误差出现在 $\eta = 1$。如图 12(c)-(d) 所示，两者



(a) σ_{11} (b) σ_{22} (c) σ_{33}

(d) σ_{12} (e) σ_{23} (f) σ_{31}

图 9：复合材料 2——使用 400 个样本训练的深度材料网络在批量大小为 20、40 和 128 时的应力预测：(a) σ_{11}；(b) σ_{22}；(c) σ_{33}；(d) σ_{12}；(e) σ_{23}；(f) σ_{31}；实线表示 10 次随机初始化运行结果的均值，阴影区域表示相应的标准差。



(a) σ_{11} (b) σ_{22} (c) σ_{33}

(d) σ_{12} (e) σ_{23} (f) σ_{31}

图10：复合材料 3——使用 400 个样本训练的 DMN 所得应力预测，批量大小分别为 20、40 和 128：(a) σ_{11}；(b) σ_{22}；(c) σ_{33}；(d) σ_{12}；(e) σ_{23}；(f) σ_{31}；实线表示 10 次随机初始化运行的均值，阴影区域表示相应的标准差。

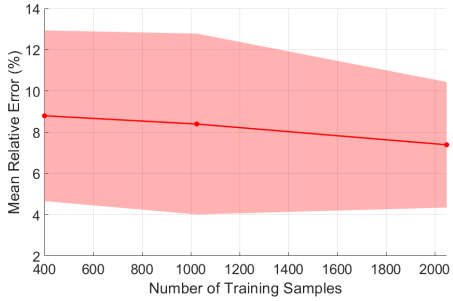


图11：全部 80 个 DMN 模型的平均相对误差，涵盖所有训练数据规模、批量大小、初始化运行次数和测试复合材料。实线表示均值，阴影区域表示相应的标准差。

随着 η 增加，收敛所需迭代次数和活跃基节点数量略有减少。因此，图 12(b) 所示的在线测试时间也随着 η增加而略有缩短。这些结果表明， η 对活跃基节点数量的影响相对有限，而活跃基节点数量反映了网络的复杂度，进而影响其在线计算效率。然而， η取值过大，会使正则化项对总损失的贡献失衡，从而导致模型性能下降（图 12(a)）。总体而言， $\eta = 1$在在线预测精度与效率之间实现了最佳权衡。

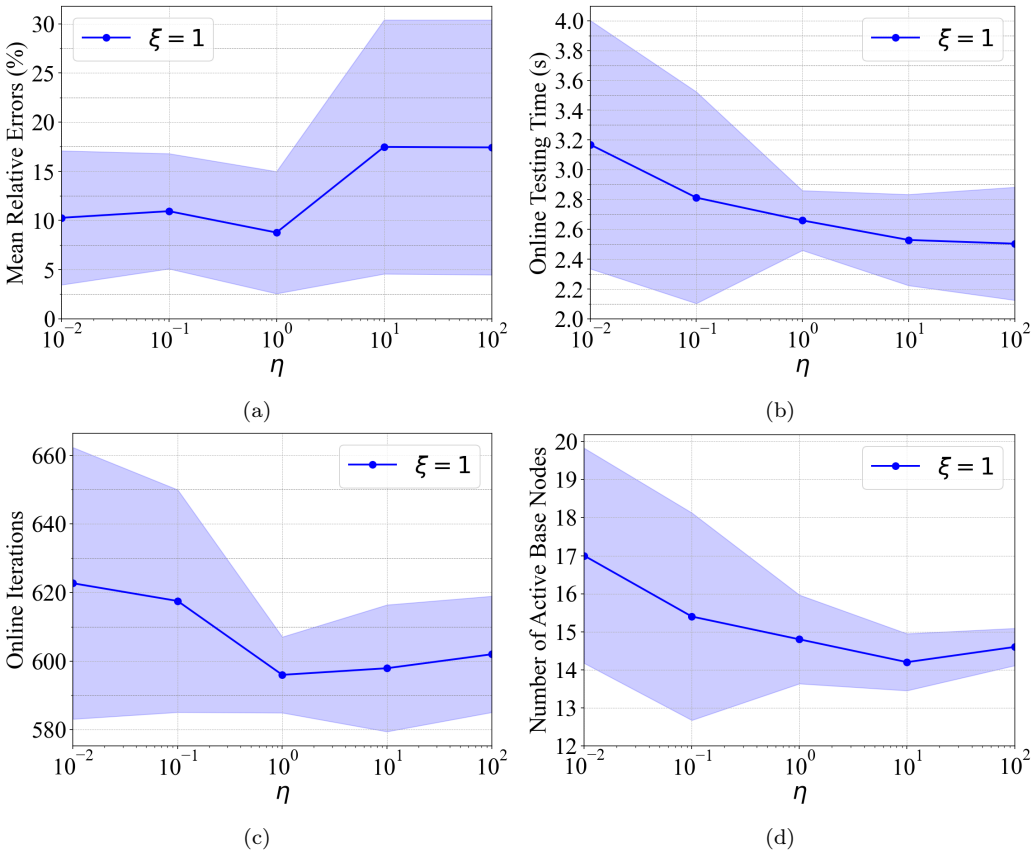


图12：损失函数（公式（9））中 η 对IMN在线预测性能的影响：（a）平均相对误差；（b）在线测试时间；（c）收敛所需迭代次数；（d）活跃基节点数量；实线表示5次随机初始化运行结果的均值，阴影区域表示相应的标准差。

接下来，我们通过训练采用 $\eta = 1$的隐式神经网络模型，考察 ξ 的影响，同时将 ξ 从 0.1 变化至 4。图 13(a) 显示，随着 ξ 增大，预测误差先降低后升高，并在$\xi = 1$处达到最小值。如图 13(b) 所示，在线测试时间随着 ξ增加而延长，这可归因于活跃基节点数量增加（图 13(d)），以及收敛所需的在线迭代次数增多（图 13(c)）。

这些观察结果表明，较大的 ξ 会促使更多基节点变为活跃状态，从而增加网络复杂度和在线计算成本。反之，当 ξ 过小时，网络中的活跃基节点数量不足，无法充分捕捉关键的微观结构相互作用和均匀化行为，导致预测精度下降。增大 ξ 可以提升网络表达能力，并在达到最佳值之前提高准确性。然而，超过这一点后，过高的网络复杂度可能会导致性能下降。基于这些参数研究，在后续所有 DMN 和 IMN 实验中，损失函数的 $\eta = 1$ 和 $\xi = 1$均采用选定值。

3.3. DMN 与 *IMN* 对比 - 离线训练

本节比较 DMN 和 IMN 的离线训练成本。两个模型均使用数据集1（400 个样本）进行训练，网络深度为 4 至 8 层，训练轮次为 10,000。为考虑初始化带来的差异，每种配置均采用 5 次随机初始化进行训练。图 14(a) 和 (b) 分别展示了 DMN 和 IMN 平均相对误差的训练历史。对于这两种模型，更深的网络通常会收敛到更低的训练误差。观察发现，当网络深度为 $N \geq 6$时，IMN 的误差低于 DMN。

图 14(c) 比较了 DMN 和 IMN 的离线训练时间。与 DMN 相比，IMN 的离线训练加速比达到 $3.4\times$ - $4.7\times$ 。这一加速的主要原因之一是，在给定网络深度的情况下，IMN 的可训练参数数量更少，如第 2.2 节所述。此外，IMN 前向传播的计算复杂度更低，这一点将在第 3.6 节中展示。离线训练效率的提升不仅有助于在实际应用中更快地基于预训练模型进行微调，还能使用更大、更多样化的训练数据集，从而增强鲁棒性和准确性，如第 3.1 节所示。

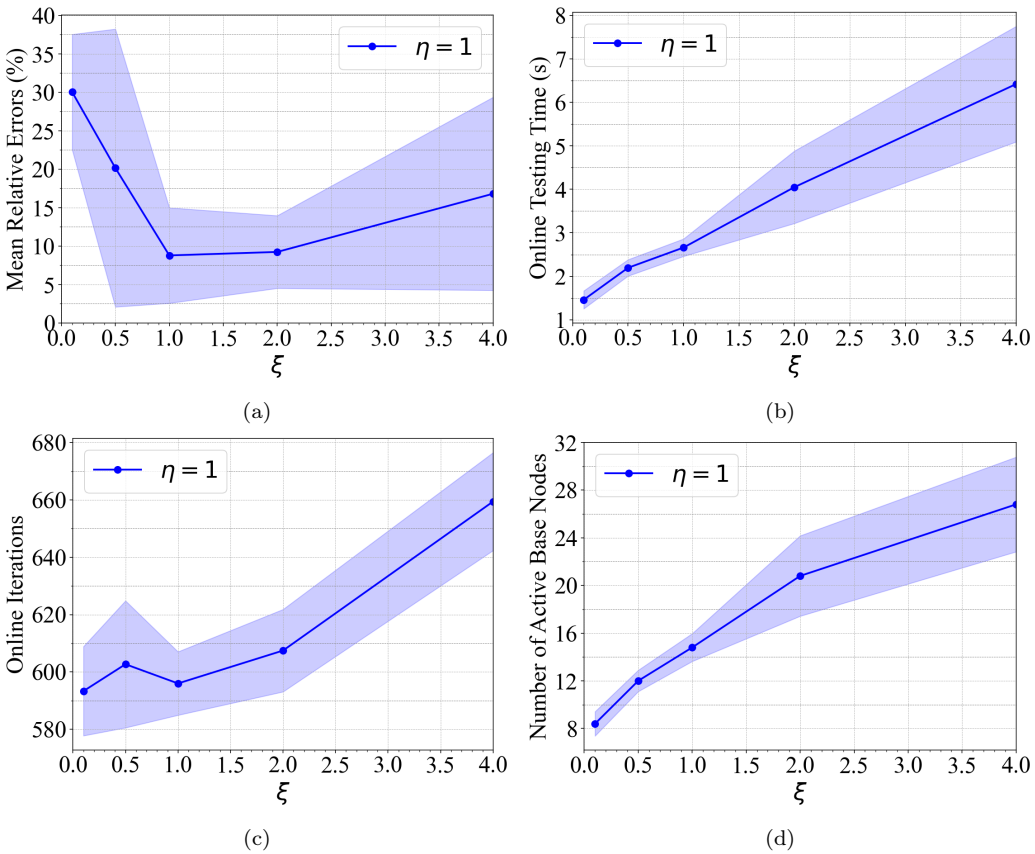
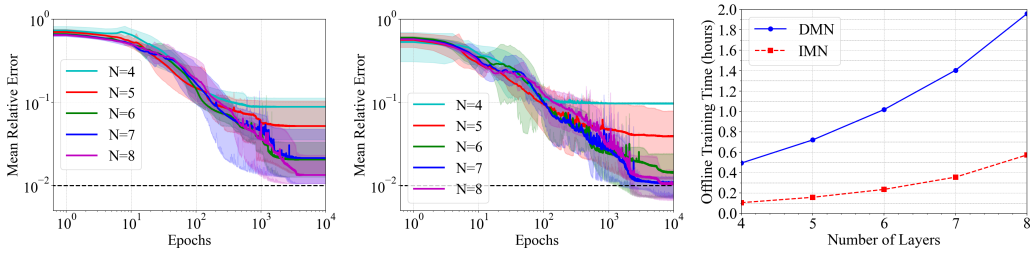


图13：损失函数（公式（9））中 ξ 对 IMN 在线预测性能的影响：(a) 平均相对误差；(b) 在线测试时间；(c) 收敛所需迭代次数；(d) 活跃基节点数量；实线表示采用随机初始化进行 5 次运行的均值，阴影区域表示相应的标准差。



(a) (b) (c)

图14：（a）网络深度为4至8的深度材料网络模型的平均相对误差训练历史；（b）不同层数的 IMN 的训练平均相对误差历史；（c）DMN 和 IMN 的离线训练时间；实线表示采用随机初始化运行5次的均值，阴影区域表示相应的标准差。

3.4. DMN 在线预测——残余应力的影响

本节研究深度材料网络在线迭代方案（公式（10）—（12））中的残余应力项对预测性能的影响，具体内容如第 2.1.2 节所述。使用包含残余应力和不包含残余应力的在线算法，对每个训练后的 DMN 进行评估。将训练后的深度材料网络模型用于预测复合材料 1 在六种加载工况下的材料响应。

图15（a）表明，加入残余应力项对预测精度没有显著影响，该精度采用公式（32）中定义的度量指标进行衡量。如图15（b）所示，包含残余应力的在线方案在在线测试期间所需的计算时间更少，平均加速比达到1.8× - 2.0×。此外，图15（c）表明，由于残余应力项的修正效应，该方案以显著更少的迭代次数收敛。

总体而言，在线计算代价取决于活跃基节点的数量、收敛所需迭代次数以及前向和反向传播的计算复杂度。为分离算法固有代价，在线测试时间同时按活跃基节点数量和迭代次数进行归一化，从而得到单次迭代单节点代价，如图15(d)所示。该度量指标反映了前向和反向传播的固有计算复杂度，并表明含残余应力的在线方案会产生更高的单次迭代单节点代价，约为不含残余应力方案的1.3× - 1.5×倍，这是由于与残余应力项相关的额外计算所致。

总体而言，这些结果表明，尽管单次迭代单节点计算成本更高，纳入残余应力的DMN在线方案相比不含残余应力的方案实现了更优的整体效率。

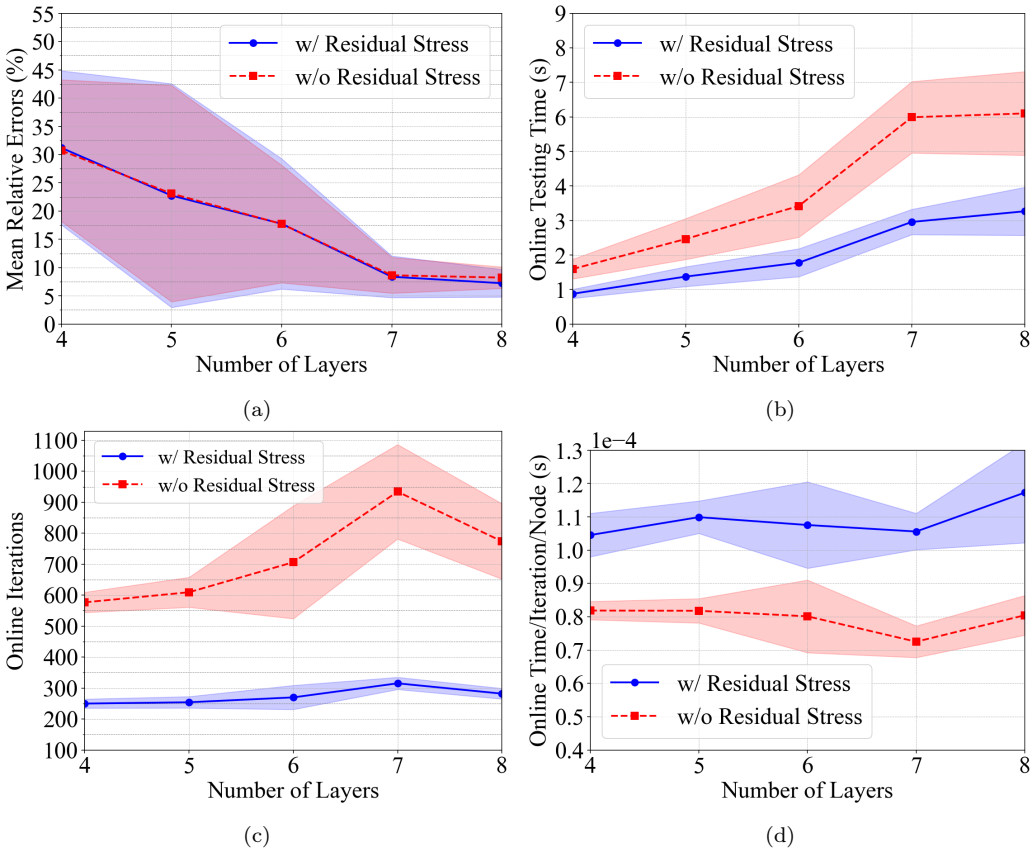


图15：残余应力对 IMN 在线预测性能的影响：(a) 平均相对误差；(b) 在线测试时间；(c) 收敛所需迭代次数；(d) 每个节点每次迭代的在线测试时间；实线表示 5 次随机初始化运行的均值，阴影区域表示相应的标准差。

3.5. IMN 在线预测——定点法与牛顿法

本节中，我们将按照第2.2.2节所述，采用不动点迭代方案和牛顿迭代方案，研究 IMN 的在线性能。每个训练好的 IMN 均使用这两种在线

算法进行评估。我们使用训练好的隐式神经网络模型，预测复合材料 1 在六种加载工况下的材料响应。

图 16(a) 显示，两个在线方案的预测准确性相当，衡量指标为公式（32）中定义的平均相对误差。如图 16(b) 所示，牛顿迭代方案在在线测试期间所需的计算时间更少，平均加速比达到 1.9× - 2.6×。此外，图 16(c) 显示，牛顿迭代方案收敛所需的迭代次数更少，这归因于其通过雅可比矩阵获得了更准确的搜索方向。

为单独考察算法固有代价，将在线测试时间分别除以活跃基节点数和迭代次数，得到单次迭代单节点代价，如图 16(d) 所示。该度量指标反映了两种在线算法前向和反向传播的固有计算复杂度，并表明牛顿迭代方案效率更高，平均加速比达到每次迭代每个节点 1.7× - 2.4× 。

总体而言，这些结果表明，牛顿迭代方案为IMN提供了更高的效率。在下一小节中，将采用牛顿迭代的IMN与DMN在在线性能方面进行比较。

3.6. DMN与IMN对比——在线预测

本节比较DMN与IMN的在线预测性能。训练后的模型用于预测复合材料1在六种加载工况下的材料响应。

图17(a)表明，当网络深度超过六层时，两种模型的预测准确性（公式（32））相当，而在较浅的网络中，DMN表现出更高的准确性。总体而言，预测不确定性随网络深度的增加而降低。图17(b)表明，两种模型的在线测试时间均随网络深度增加，且在所有配置下保持相当。

如图17(c)所示，DMN收敛所需的在线迭代显著更少（为IMN的0.4× - 0.5× ），这归因于其在线算法中的残余应力（见第2.1.2节和第3.4节）。DMN与IMN的迭代次数对网络深度相对不敏感，不过DMN表现出略大的变异性。图17(d)表明，两种模型的活跃基节点数量均随网络深度增加，且在各深度下数值相当。这表明，在相同的正则化设-

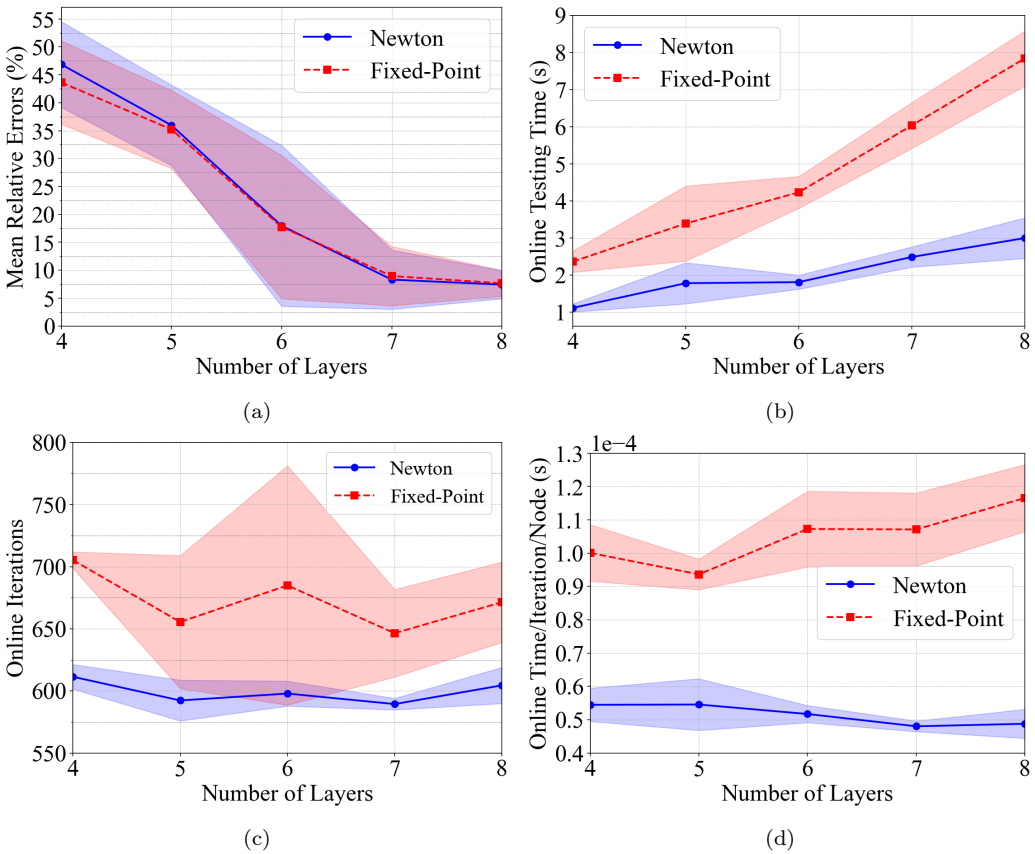


图16：IMN在线预测的不动点迭代与牛顿迭代对比：(a) 平均相对误差；(b) 在线测试时间；(c) 收敛所需迭代次数；(d) 每个节点每次迭代的在线测试时间；实线表示采用随机初始化进行5次运行的均值，阴影区域表示相应的标准差。

在损失函数设置中，DMN 和 IMN 的网络复杂度相近，因此可以公平地比较二者的性能。

为隔离两种模型的固有计算复杂度，将在线测试时间按活跃基节点数量和迭代次数进行归一化，得到图 17(e) 所示的单次迭代单节点代价。该度量指标反映了在线预测前向和反向传播的成本，不受迭代次数和节点数量影响。结果表明，IMN 的效率更高，每次迭代每个节点的加速比为 1.9× - 2.4× 。然而，IMN 收敛所需的迭代次数更多（图 17(c)），抵消了这一优势，因此其总体在线计算成本与 DMN 相当（图 17(b)）。总体而言，这些结果表明，DMN 和 IMN 的在线预测精度和效率相当。

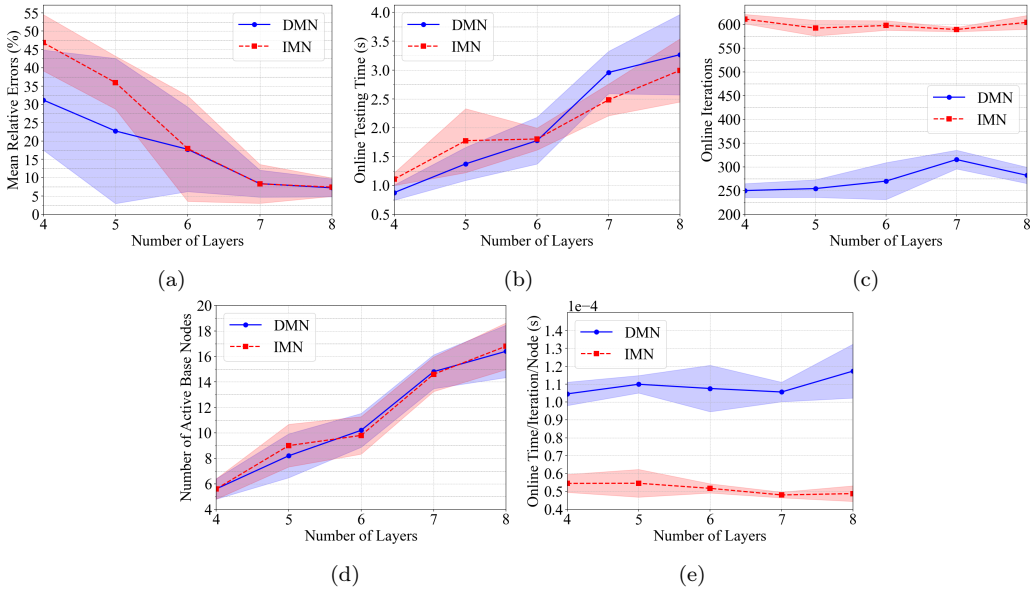


图17：DMN 与 IMN 的在线预测对比：(a) 平均相对误差；(b) 在线测试时间；(c) 收敛所需迭代次数；(d) 活跃基节点数量；(e) 每个节点每次迭代的在线测试时间；实线表示 5 次随机初始化运行的均值，阴影区域表示相应的标准差。

4. 结论

本研究对保结构材料网络进行了系统的性能评估，包括最初由 [47, 48] 提出的深度材料网络（DMN）及其变体——无旋转交互材料网络（IMN） [56, 57]。与纯数据驱动模型不同，DMN 将细观力学构建块嵌入具有物理可解释性参数的层级二叉树架构中。这种保结构设计赋予其强大的泛化能力：尽管网络仅使用线弹性数据进行训练，却能准确预测由此前未见过的基体材料组成的复合材料的非线性非弹性响应。这一特性凸显了它们在实际多尺度材料建模中的巨大潜力。

通过广泛的参数研究，我们考察了训练数据量、批量大小、初始化和损失函数正则化对模型准确性、鲁棒性及计算效率的影响。结果表明，尽管初始化和批量大小会影响模型训练及其性能，但增加训练数据量能够持续提高预测精度并降低不确定性。此外，基节点激活的正则化参数（ξ， η）会直接影响网络复杂度，进而影响在线预测精度和计算成本。具体而言，η 控制正则化强度，而 ξ 则会显著影响网络复杂度。基于参数分析，对于本研究所考察的示例，采用 $\eta = \xi = 1$ 可在准确性和效率之间实现平衡。

DMN 与 IMN 的全面比较揭示了二者各自的优势与权衡。离线训练结果表明，IMN 的训练效率更高，相较于 DMN 加速了 $3.4\times$–$4.7\times$，这主要归因于其可训练参数较少（第 2.2 节），且前向传播的计算复杂度较低（第 3.6 节）。在线性能评估表明，DMN 和 IMN 的预测精度及总体计算成本相近。DMN 的在线算法采用残余应力（第 3.4 节和第 3.6 节），因此收敛所需的迭代次数较少；而 IMN 的单次迭代、单个节点的计算成本则显著更低。这些因素相互抵消，使二者的总体在线效率相近。此外，研究表明，IMN 的基于牛顿法的在线迭代格式显著优于不动点迭代，在保持准确性相近的同时，最高可实现 $2.6\times$ 的加速比。

本研究结果为在多尺度模拟中部署这两种方法提供了实用指导。

总体而言，IMN中可训练参数数量的减少提高了离线训练效率，同时保持了与DMN相当的在线预测精度和效率。未来的工作将着重开发更高效、稳健的训练策略，以减轻模型对训练设置的敏感性，例如初始化 [78] 和激活正则化 3.2；同时，还将设计改进的采样策略 [62, 79] ，在最大限度提升学习性能的同时降低训练数据需求。

参考文献

[1] J. Fish, 《多尺度方法：连接科学与工程中的尺度》，牛津大学出版社，2010年。[2] M. G. Geers, V. G. Kouznetsova, W. Brekelmans，多尺度计算均匀化：趋势与挑战，《计算与应用数学杂志》234 (7) (2010) 2175–2182。[3] J. Fish, G. J. Wagner, S. Keten，材料的介观与多尺度建模，《自然·材料》20 (6) (2021) 774–786。[4] D. Bishara, Y. Xie, W. K. Liu, S. Li，基于机器学习的多尺度建模、仿真、均匀化与材料设计的最新综述，《计算工程方法档案》30 (1) (2023) 191–222。[5] R. Hill，增强固体的弹性性质：若干理论原理，《固体力学与物理学杂志》11 (5) (1963) 357–372。[6] F. Feyel, J.-L. Chaboche，用于模拟长纤维SiC/Ti复合材料弹黏塑性行为的有限元多尺度方法，《Computer Methods in Applied Mechanics and Engineering》183 (3-4) (2000) 309–330。[7] H. Moulinec, P. Suquet，一种计算具有复杂微观结构的非线性复合材料整体响应的数值方法，《Computer Methods in Applied Mechanics and Engineering》157 (1-2) (1998) 69–94。[8] T. De Geus, J. Vondřejc, J. Zeman, R. Peerlings, M. Geers，简明易用的有限应变FFT非线性求解器，《Computer Methods in Applied Mechanics and Engineering》318 (2017) 412–430。[9] C. Naili, I. Doghri, T. Kanit, M. Sukiman, A. Aissa-Berraies, A. Imad，短纤维增强复合材料：弹性力学中多种均匀化方法的无偏全场评估，《复合材料科学与技术》187 (2020) 107942。

[10] F. Feyel，一种利用广义连续体描述高度非线性结构响应的多层有限元方法（FE²），Computer Methods in Applied Mechanics and Engineering 192 (28-30) (2003) 3233–3244.[11] V. Kouznetsova、M. G. Geers、W. Brekelmans，多相材料的多尺度二阶计算均质化：一种嵌套有限元求解策略，Computer Methods in Applied Mechanics and Engineering 193 (48-51) (2004) 5525–5550.[12] J. Spahn、H. Andrä、M. Kabel、R. Müller，一种利用快速傅里叶变换对复合材料渐进损伤进行建模的多尺度方法，Computer Methods in Applied Mechanics and Engineering 268 (2014) 871–883.[13] J. Kochmann、S. Wulfinghoff、L. Ehle、J. Mayer、B. Svendsen、S. Reese，对弹黏塑性多晶体有效材料行为进行高效、精确的基于 FE-FFT 的双尺度预测，计算力学 61 (6) (2018年) 751–764.[14] V. B. C. Tan、K. Raju、H. P. Lee，用于非均质结构并发多层建模的直接 FE² ，Computer Methods in Applied Mechanics and Engineering 360 (2020) 112694.[15] X. He、K. Taneja、J.-S. Chen、C.-H. Lee、J. Hodgson、V. Malis、U. Sinha、S. Sinha，被动材料影响对骨骼肌变形与力输出的多尺度建模，《生物医学工程数值方法国际期刊》38 (4) (2022) e3571.[16]R. Xu、J. Yang、W. Yan、Q. Huang、G. Giunta、S. Belouettar、H. Zahrouni、T. B. Zineb、H. Hu，数据驱动多尺度有限元方法：从并发到分离，Computer Methods in Applied Mechanics and Engineering 363 (2020) 112893.[17] J. Ghaboussi、J. Garrett Jr、X. Wu，利用神经网络对材料行为进行基于知识的建模，《工程力学杂志》117 (1) (1991) 132–153.[18] Y. LeCun、Y. Bengio、G. Hinton，深度学习，《自然》521 (7553) (2015) 436–444.

[19] I. Goodfellow，深度学习，麻省理工学院出版社，2016年。

[20] K. Wang, W. Sun，通过递归均匀化与深度学习关联的多尺度多渗透率弹塑性模型，Computer Methods in Applied Mechanics and Engineering 334 (2018) 337–380。[21] F. Ghavamian, A. Simone，利用循环神经网络加速历史相关材料的多尺度有限元模拟，Computer Methods in Applied Mechanics and Engineering 357 (2019) 112594。[22] X. Liu, S. Tian, F. Tao, W. Yu，人工神经网络在复合材料本构建模中的综述，Composites Part B: Engineering 224 (2021) 109152。[23] J. N. Fuhg, G. A. Padmanabha, N. Bouklas, B. Bahmani, W. Sun, N. N. Vlassis, M. Flaschel, P. Carrara, L. De Lorenzis，固体数据驱动本构定律综述，《工程计算方法档案》32 (2025) 1841–1883。[24] T. Kirchdoerfer, M. Ortiz，数据驱动计算力学，Computer Methods in Applied Mechanics and Engineering 304 (2016) 81–101。[25] R. Ibanez, E. Abisset-Chavanne, J. V. Aguado, D. Gonzalez, E. Cueto, F. Chinesta，一种用于数据驱动计算弹性力学与非弹性力学的流形学习方法，《工程计算方法档案》25 (1) (2018) 47–57。[26] R. Eggersmann, T. Kirchdoerfer, S. Reese, L. Stainier, M. Ortiz，无模型数据驱动非弹性力学，Computer Methods in Applied Mechanics and Engineering 350 (2019) 81–99。[27] Q. He, J.-S. Chen，基于局部凸重构的物理约束数据驱动方法，用于含噪数据库，Computer Methods in Applied Mechanics and Engineering 363 (2020) 112791。[28] X. He, Q. He, J.-S. Chen, U. Sinha, S. Sinha，各向异性非线性弹性固体的物理约束局部凸性数据驱动建模，《数据中心工程》1 (2020) e19。

[29] X. He, Q. He, J.-S. Chen，用于物理约束数据驱动非线性材料建模的深度自编码器，Computer Methods in Applied Mechanics and Engineering 385 (2021) 114034。[30] Q. He, D. W. Laurence, C.-H. Lee, J.-S. Chen，基于流形学习的软生物组织数据驱动建模，生物力学杂志 117 (2021) 110124。[31] R. Eggersmann, L. Stainier, M. Ortiz, S. Reese，利用张量投票增强的无模型数据驱动计算力学，Computer Methods in Applied Mechanics and Engineering 373 (2021) 113499。[32] A. Gorgogianni, K. Karapiperis, L. Stainier, M. Ortiz, J. E. Andrade，数据驱动计算力学中的自适应目标导向数据采样，Computer Methods in Applied Mechanics and Engineering 409 (2023) 115949。[33] Y. Xu, J. Yang, Z. Kuang, Q. Huang, W. Huang, H. Hu，量子计算增强的距离最小化数据驱动计算力学，Computer Methods in Applied Mechanics and Engineering 419 (2024) 116675。[34] N. N. Vlassis, R. Ma, W. Sun，计算力学中的几何深度学习（一）：各向异性超弹性，Computer Methods in Applied Mechanics and Engineering 371 (2020) 113299。[35] F. Masi, I. Stefanou, P. Vannucci, V. Maffi-Berthier，基于热力学的人工神经网络用于本构建模，《固体力学与物理学杂志》147 (2021) 104277。[36] X. He, J.-S. Chen，用于路径相关材料数据驱动建模的热力学一致机器学习内变量方法，Computer Methods in Applied Mechanics and Engineering 402 (2022) 115348。[37] M. Maia, I. B. Rocha, P. Kerfriden, F. P. van der Meer，适用于路径相关非均质材料的物理循环神经网络：在数据驱动代理中嵌入本构模型，Computer Methods in Applied Mechanics and Engineering 407 (2023) 115934。

[38] Y. He、S. J. Semnani，用于路径相关材料行为建模的增量神经控制微分方程，Computer Methods in Applied Mechanics and Engineering 422 (2024) 116789.[39] R. E. Jones、A. L. Frankel、K. Johnson，一种借助内部状态变量对非弹性应力响应进行建模的神经常微分方程框架，《机器学习建模与计算杂志》3 (3) (2022).[40] B. Guo、Z. Lin、Q. He，历史感知神经算子：路径相关材料的稳健数据驱动本构建模，Computer Methods in Applied Mechanics and Engineering 447 (2025年) 118358.[41] J. Yvonnet 、 Q.-C. He，用于有限应变下超弹性介质非线性均质化的降阶模型多尺度方法（r3m），《计算物理学杂志》223 (1) (2007年) 341–368.[42] J. Oliver、M. Caicedo、A. E. Huespe、J. Hernández、E. Roubin，面向计算多尺度断裂的降阶建模策略，Computer Methods in Applied Mechanics and Engineering 313 (2017年) 560–595.[43] F. Fritzen、O. Kunc，采用降阶模型的非线性固体两阶段数据驱动均质化，《欧洲力学杂志A辑/固体》69 (2018年) 201–220.[44] S. Kaneko、H. Wei、Q. He、J.-S. Chen、S. Yoshimura，一种用于加速热循环诱发塑性变形建模的超降阶计算方法，《固体力学与物理学杂志》151 (2021) 104385.[45] Z. Liu、M. Bessa、W. K. Liu，自洽聚类分析：一种面向非弹性非均质材料的高效多尺度方案，Computer Methods in Applied Mechanics and Engineering 306 (2016年) 319–341.[46] Z. Liu、M. Fleming、W. K. Liu，用于弹塑性应变软化材料自洽聚类分析的微观结构材料数据库，Computer Methods in Applied Mechanics and Engineering 330 (2018年) 547–577.

[47] Z. Liu、C. T. Wu、M. Koishi，一种用于非均质材料多尺度拓扑学习与加速非线性建模的深度材料网络，Computer Methods in Applied Mechanics and Engineering 345 (2019年) 1138–1168.[48] Z. Liu、C. T. Wu，探索数据驱动的多尺度力学中深度材料网络的三维架构，《固体力学与物理学杂志》127 (2019年) 20–46.[49]Q. Hernández、A. Badías、D. González、F. Chinesta、E. Cueto，保结构神经网络，《计算物理学杂志》426 (2021) 109950.[50] E. Celledoni、M. J. Ehrhardt、C. Etmann、R. I. McLachlan、B. Owren、C.-B. Schonlieb、F. Sherry，保结构深度学习，《欧洲应用数学杂志》32 (5) (2021) 888–936.[51]Z. Zhang、Y. Shin、G. Em Karniadakis，Gfinns：面向确定性与随机动力系统的通用形式体系信息神经网络，《英国皇家学会哲学汇刊A辑》380 (2229) (2022) 20210207.[52] A. Gruber、M. Gunzburger、L. Ju、Z. Wang，度量辛系统的能量一致的模型降阶，Computer Methods in Applied Mechanics and Engineering 404 (2023) 115709.[53] A. Gruber、K. Lee、H. Lim、N. Park、N. Trask，高效参数化神经度量辛系统，arXiv预印本 arXiv:2405.16305 (2024).[54] J. S. R. Park、S. W. Cheung、Y. Choi、Y. Shin，tlasdi：热力学信息引导的潜空间动力学辨识，Computer Methods in Applied Mechanics and Engineering 429 (2024) 117144.[55] X. He、Y. Shin、A. Gruber、S. Jung、K. Lee、Y. Choi，参数化系统的热力学一致潜动力学辨识，arXiv预印本 arXiv:2506.08475 (2025年).[56] S. Gajek、M. Schneider、T. Böhlke，论深度材料网络的微观力学，《固体力学与物理学杂志》142 (2020) 103984.

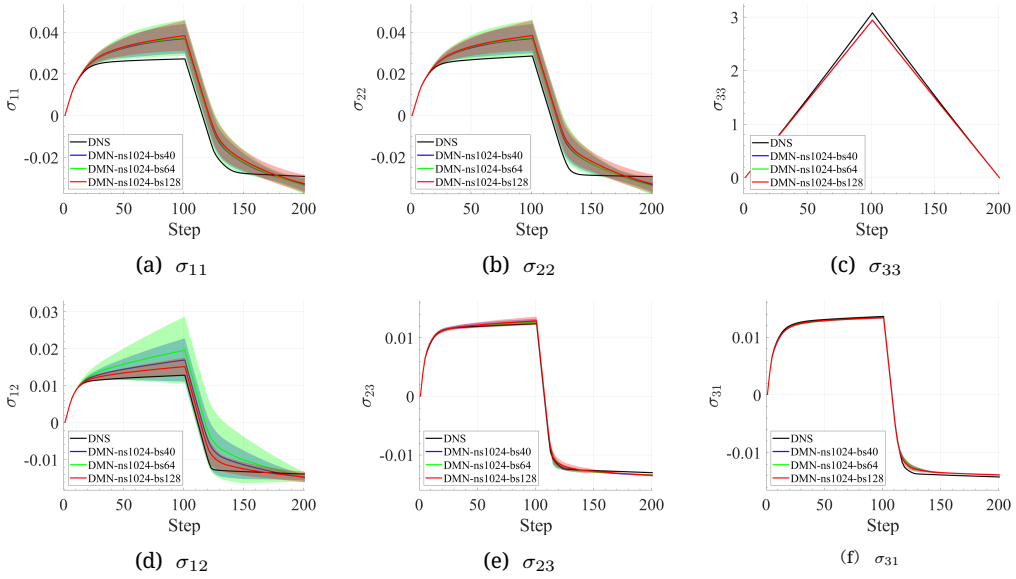
[57] V.-D. Nguyen, L. Noels，基于相互作用的材料网络：一种适用于（多孔）微结构材料的通用框架，Computer Methods in Applied Mechanics and Engineering 389 (2022) 114300.[58] S. Gajek, M. Schneider, T. Böhlke，一种用于短纤维增强塑料部件多尺度分析的有限元深度材料网络方法，Computer Methods in Applied Mechanics and Engineering 384 (2021) 113952.[59] V.-D. Nguyen, L. Noels，从相互作用视角重新审视基于微观力学的材料网络；面向多相复合材料的稳健高效实现，《欧洲力学杂志：A辑/固体》91 (2022) 104384.[60] L. Wu, L. Adam, L. Noels，用于机织复合材料多尺度分析的基于微观力学和数据驱动的降阶模型，Composite Structures 270 (2021) 114058.[61] S. Gajek, M. Schneider, T. Böhlke，一种用于热力耦合复合材料多尺度分析的有限元深度材料网络方法，计算力学 69 (5) (2022) 1087–1113.[62] A. P. Dey, F. Welschinger, M. Schneider, S. Gajek, T. Böhlke，采用基于非弹性信息的策略训练深度材料网络，以再现短纤维增强热塑性塑料的蠕变载荷，应用力学档案 92 (9) (2022) 2733–2755.[63] H. Wei, C. T. Wu, W. Hu, T.-H. Su, H. Oura, M. Nishi, T. Naito, S. Chung, L. Shen，基于 LS-DYNA 机器学习的多尺度方法：短纤维增强复合材料的非线性建模，《工程力学杂志》149 (3) (2023) 04023003.[64] T.-J. Wei, T.-H. Su, C.-S. Chen，用于多晶材料建模的具备方向感知的基于相互作用的深度材料网络，Computer Methods in Applied Mechanics and Engineering 441 (2025) 117977.[65] N. M. Francis, D. Shin, R. A. Lebensohn, F. Pourahmadian, R. Dingreville，微极深度材料网络，Computer Methods in Applied Mechanics and Engineering 446 (2025) 118329.

[66] B. Sterr, S. Gajek, A. Hrymak, M. Schneider, T. Böhlke，具有无限材料对比度的纤维悬浮液深度材料网络，International Journal for Numerical Methods in Engineering 126 (7) (2025年) e70014.[67] Z. Liu，带有黏聚层的深度材料网络：多阶段训练与界面失效分析，Computer Methods in Applied Mechanics and Engineering 363 (2020) 112913.[68] Z. Liu，应用于多尺度应变局部化建模的深度材料网络中的单元划分，Computer Methods in Applied Mechanics and Engineering 384 (2021) 113914.[69] D. Shin, P. J. Creveling, S. A. Roberts, R. Dingreville，用于热导率问题的深度材料网络：在机织复合材料中的应用，Computer Methods in Applied Mechanics and Engineering 431 (2024) 117279.[70] D. Shin, R. Alberdi, R. A. Lebensohn, R. Dingreville，一种用于预测复合材料热力学响应的深度材料网络方法，Composites Part B: Engineering 272 (2024) 111177.[71] D. Shin, R. A. Lebensohn, R. Dingreville，利用微观结构局部场信息训练 U-net 架构的深度材料网络，计算力学 (2026) 1–21.[72] Z. Liu, C. T. Wu, M. Koishi，深度材料网络的迁移学习：实现无缝结构–性能预测，计算力学 64 (2) (2019) 451–465.[73]Z. Liu, H. Wei, T. Huang, C. T. Wu，基于工艺引导的复合材料数据库的智能多尺度模拟，arXiv 预印本 arXiv:2003.09491 (2020).[74] T. Huang, Z. Liu, C. T. Wu, W. Chen，微观结构引导的深度材料网络：快速非线性材料建模与不确定性量化，Computer Methods in Applied Mechanics and Engineering 398 (2022) 115197.[75]J. G. Jean, T.-H. Su, S.-J. Huang, C. T. Wu, C.-S. Chen, 图增强深度材料网络：多尺度材料建模与

微观结构信息学，《计算力学》75 (1) (2025) 113–136。[76] T.-J. Wei、C.-S. Chen，复合材料微观结构基础模型：重构、刚度与非线性行为预测，《材料与设计》257 (2025) 114397。[77] T. Li，基于微观力学的参数化深度材料网络，用于预测具有变化形貌的非均质材料的物理行为，《Computer Methods in Applied Mechanics and Engineering》419 (2024) 116687。[78] D. Shin、R. Alberdi、R. A. Lebensohn、R. Dingreville，基于拼接策略的深度材料网络：通过可视化提高可解释性，并通过递归训练提升准确性，《npj计算材料》9 (1) (2023) 128。[79] P. B. K. Srinivas、M. Kabel、M. Schneider，通过基于位移的层合板公式和一种新颖的采样技术，为基于柔度的疲劳模型快速进行深度材料网络离线训练，《Computer Methods in Applied Mechanics and Engineering》449 (2026) 118517。[80] H. Wei、D. Lyu、W. Hu、C. T. Wu，使用 LS-DYNA 进行代表性体积单元分析，以实现高保真多尺度材料建模，arXiv 预印本 arXiv:2210.11761 (2022)。[81]A. Paszke、S. Gross、F. Massa、A. Lerer、J. Bradbury、G. Chanan、T. Killeen、Z. Lin、N. Gimelshein、L. Antiga 等，PyTorch：一种命令式风格的高性能深度学习库，《神经信息处理系统进展》32 (2019)。

附录 A. 在数据集2上训练的 DMN

本节展示了使用数据集2（1024 个样本）训练的 DMN 对三种复合材料进行的预测，批量大小分别为 40、64 和 128，如图 A.18–A.20 所示。在这三种配置中，批量大小为 128 时的预测更准确，且在所有测试的复合材料中不确定性更低，尤其是在 11、22 和 12 方向上；这与图 7(b) 所示相应的训练平均相对误差一致。



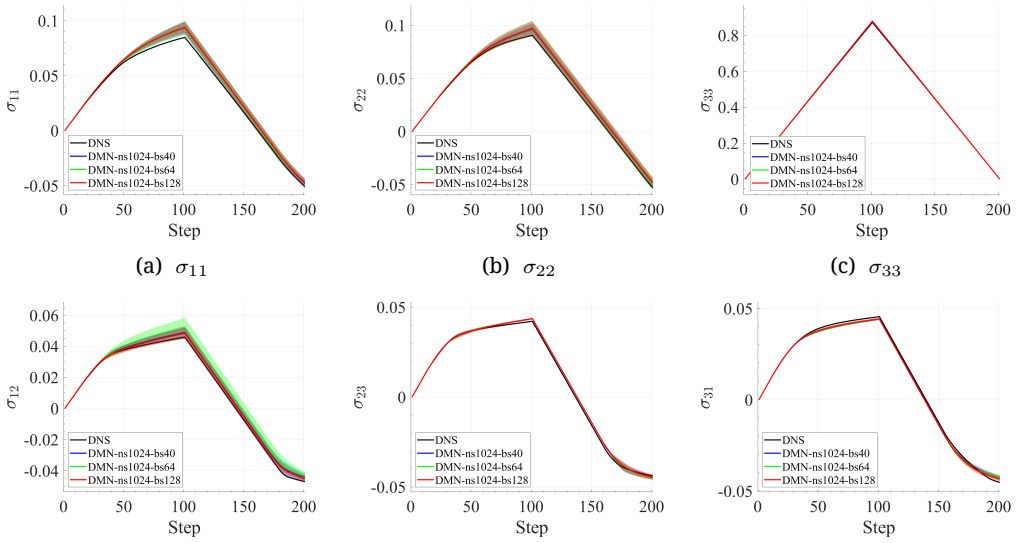
(a) σ_{11} (b) σ_{22} (c) σ_{33}

(d) σ_{12} (e) σ_{23} (f) σ_{31}

图 A.18：复合材料 1——使用 1024 个样本、批量大小分别为 40、64 和 128 训练的 DMN 的应力预测：(a) σ_{11}；(b) σ_{22}；(c) σ_{33}；(d) σ_{12}；(e) σ_{23}；(f) σ_{31}；实线表示 10 次随机初始化运行的均值，阴影区域表示相应的标准差。

附录 B。使用数据集3训练的 DMN

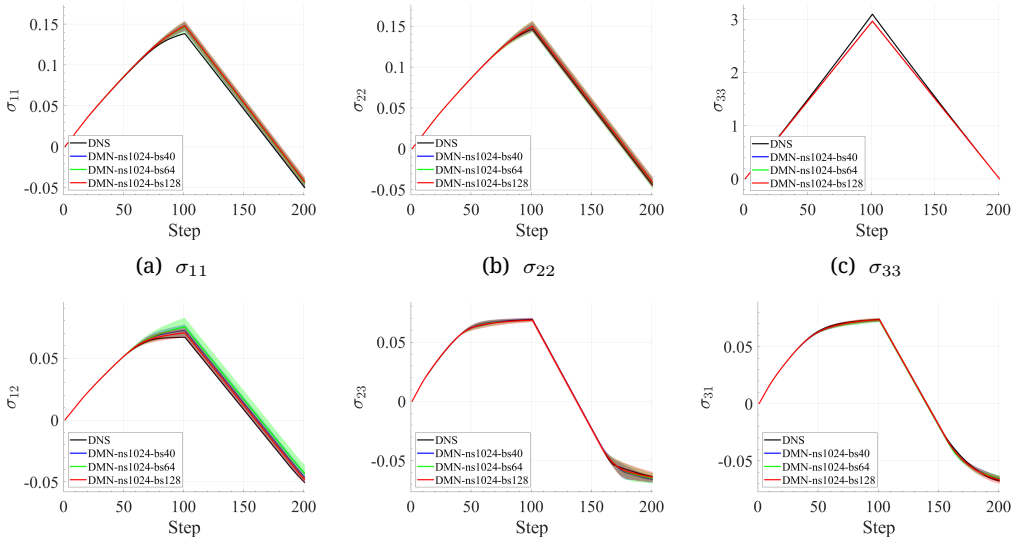
本节展示了使用数据集3（2048 个样本）以及批量大小 128 和 256 训练的 DMN 对三种复合材料的预测结果，如图 B.21—B.23 所示。在这些配置中，批量大小为 128 时的预测结果更为准确，并且在所有测试的复合材料中具有更低的不确定性，尤其是在 11、22 和 12 方向上；这与图 7(c) 所示的相应训练平均相对误差一致。



(a) σ_{11} (b) σ_{22} (c) σ_{33}

(d) σ_{12} (e) σ_{23} (f) σ_{31}

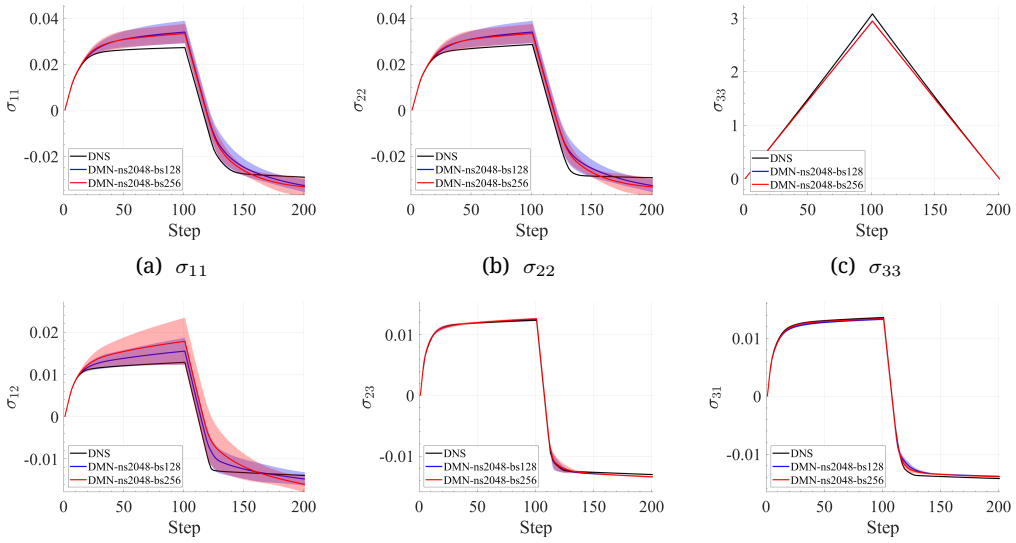
图 A.19：复合材料 2 - 使用 1024 个样本、批量大小分别为 40、64 和 128 训练的深度材料网络的应力预测：(a) σ_{11}；(b) σ_{22}；(c) σ_{33}；(d) σ_{12}；(e) σ_{23}；(f) σ_{31}；实线表示 10 次随机初始化运行的均值，阴影区域表示相应的标准差。



(a) σ_{11} (b) σ_{22} (c) σ_{33}

(d) σ_{12} (e) σ_{23} (f) σ_{31}

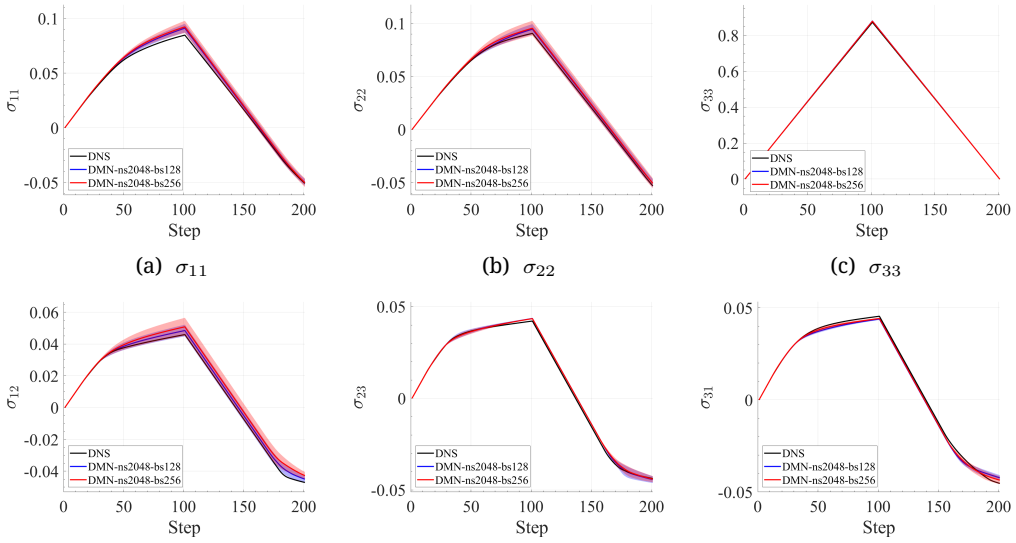
图 A.20：复合材料 3——使用批量大小为 40、64 和 128 的设置，以 1024 个样本训练的深度材料网络的应力预测：(a) σ_{11}；(b) σ_{22}；(c) σ_{33}；(d) σ_{12}；(e) σ_{23}；(f) σ_{31}；实线表示 10 次采用随机初始化运行结果的均值，阴影区域表示相应的标准差。



(a) σ_{11} (b) σ_{22} (c) σ_{33}

(d) σ_{12} (e) σ_{23} (f) σ_{31}

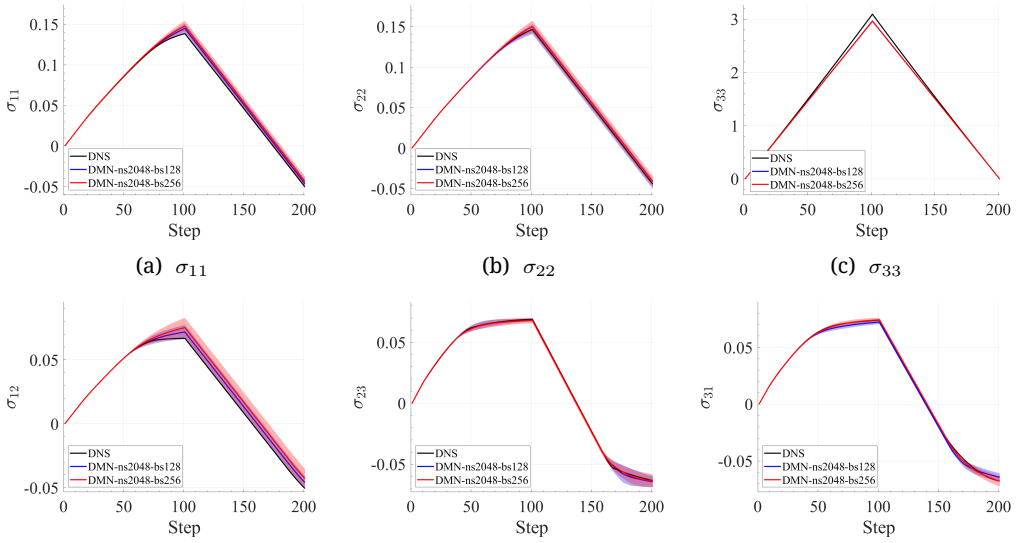
图 B.21：复合材料 1——使用 2048 个样本、批量大小为 128 和 256 训练的 DMN 的应力预测：（a）σ_{11}；（b）σ_{22}；（c）σ_{33}；（d）σ_{12}；（e）σ_{23}；（f）σ_{31}；实线表示 10 次随机初始化运行的均值，阴影区域表示相应的标准差。



(a) σ_{11} (b) σ_{22} (c) σ_{33}

(d) σ_{12} (e) σ_{23} (f) σ_{31}

图 B.22：复合材料 2 - 使用 2048 个样本训练的深度材料网络的应力预测，批量大小分别为 128 和 256：(a) σ_{11}；(b) σ_{22}；(c) σ_{33}；(d) σ_{12}；(e) σ_{23}；(f) σ_{31}；实线表示 10 次随机初始化运行结果的均值，阴影区域表示相应的标准差。



(a) σ_{11} (b) σ_{22} (c) σ_{33}

(d) σ_{12} (e) σ_{23} (f) σ_{31}

图 B.23：复合材料 3——使用 2048 个样本、批量大小分别为 128 和 256 训练的 DMN 的应力预测：(a) σ_{11}；(b) σ_{22}；(c) σ_{33}；(d) σ_{12}；(e) σ_{23}；(f) σ_{31}；实线表示采用随机初始化进行 10 次运行所得的均值，阴影区域表示相应的标准差。