arXiv:2611.32499v1 [cs.CE] 24 Nov 2026

基于记忆学习非弹性均匀化：深度材料网络与离散伴随

BinhHuyNguyen[1] 和 Matti Schneider[2,3,4,*]

[1]材料成形中心（CEMEF），巴黎高科矿业学院[2]杜伊斯堡-埃森大学，工程数学研究所[3]杜伊斯堡-埃森纳米集成中心（CENIDE）[4]弗劳恩霍夫工业数学研究所ITWM，凯撒斯劳滕[*]通讯地址：matti.schneider@uni-due.de

2026年11月25日

摘要

深度材料网络（DMN）为计算均匀化提供了力学结构化代理模型。传统上，深度材料网络根据弹性数据进行识别，随后外推至非弹性响应。这一过程利用高效的线性求解器，并能够稳健地外推至不同的非弹性本构定律。然而，为给定的非弹性本构行为选择合适的线弹性训练数据并非易事。

在本研究中，我们直接根据宏观应力-应变轨迹识别 DMN。将自动微分应用于此类具有历史依赖的问题会带来巨大的内存需求，使大多数具有实际意义的问题无法处理。为缓解这一问题，我们推导出一种路径一致离散伴随，通过将完整加载历史中的平衡和内部变量演化视为耦合约束，避免传播逐参数状态灵敏度。

对于具有弹塑性基体的短纤维复合材料，伴随方法得到的灵敏度与直接微分以及独立自动微分实现所获得的结果一致；同时，该方法在每个时间步仅需对线性化状态方程进行一次反向求解，并避免逐参数传播状态灵敏度。非弹性训练还可提高预测准确度：在非弹性轨迹上训练的秩5 DMN，在不同加载路径下的表现优于在弹性均匀化数据上训练的更深层秩8 DMN。即使扰动组分非弹性材料参数，这一优势仍然保持。所得的非弹性训练 DMN 保持了在线部署所需的相同公式，并且速度仍比全场均匀化求解器快数个数量级。

关键词：深度材料网络；计算均匀化；非弹性；离散伴随 nt
方法；降阶模型；历史相关材料

1 引言

1.1 研究现状

计算均匀化提供基于物理的本构模型，可在无需解析宏观结构中各种非均质性的情况下考虑材料微观结构，而解析这些非均质性所需的计算成本高得难以承受 [1]。该方法在代表性体积单元（RVE）上求解微观平衡问题，并将所得有效响应赋予宏观实体的材料点 [2–4]。对于线性弹性材料，微观问题与宏观问题可以解耦：给定 RVE 的有效性质可预先计算并存储，随后用于均质结构分析。相比之下，宏观非弹性行为与微观组分演化密切相关，例如宏观塑性应变与晶体学滑移之间的关系 [5], ，因此必须采用耦合或嵌套的双尺度 formulation。在这种并行计算设置中，每个宏观材料点在每个载荷增量上都要通过求解非线性胞元问题来更新微观内部变量。成熟的数值技术，尤其是有限元法（FEM） [1,6,7] 和基于FFT的求解器 [8–10], （另见文献 [11–13] 及综述 [14], ），通常作为高保真微观求解器应用于 FE^2 [15,16]和 FE-FFT框架 [17,18]。尽管其成本低于完全解析的模拟，但反复求解非线性 RVE 问题仍然计算量很大，这限制了工业规模的并发多尺度模拟，也促使人们开发能够保留全场模型保真度的降阶微观均质化器，尤其适用于非弹性材料。Dvorak 提出的变换场分析（TFA） [19],将微观结构划分为若干子域，并假设各子域内的非弹性应变均匀。因此，为达到令人满意的准确度，可能需要进行精细划分 [19, 20]。为缓解这一局限，Michel 和 Suquet [21] 提出了非均匀变换场分析（NTFA），用空间非均匀、时间无关的模态以及随时间变化的活动系数来表示时空非弹性应变场。NTFA 被应用于非线性复合结构 [22, 23], ，并扩展至线性黏弹性复合材料 [24], ，还通过混合变分公式 [25] 和级数展开 [26]得到了改进。类似地，Liu 等 [27,28] 提出了自洽聚类分析，通过将 k-means 聚类与 Hill 思路下的自洽方案相结合，对 Lippmann–Schwinger方程进行近似 [29]。Schneider [30] 详细阐述了这些降阶均匀化方法之间的数学联系。正如 Michel 和 Suquet [26], 所指出的，这些方法为内部变量而非位移场或速度场构建降阶基；传统的基于投影的模型降阶则针对后者。后一类方法包括 Yvonnet 和 He 的本征正交分解模型 [31],，以及后续的超降阶技术 [32], ，例如经验求积法 [33, 34]和经验修正簇求积法 [35–37], 、Hashin–Shtrikman型降阶有限元公式 [38], ，以及基于规则网格的层合板降阶模型 [39]。

神经网络已成为表示非弹性材料行为的通用代理模型，早在20世纪90年代就已用于将应变序列映射为应力序列 [40–42]。近年来，循环神经网络（RNNs）及专为序列数据设计的变体（如长短期记忆网络）被用于表示路径依赖材料响应 [43–45], ，其中自洽性 [46]和热力学结构 [47] 被确定为重要要素。带有内变量的神经常微分方程 [48] 和历史感知神经算子 [49]也致力于实现相关目标，并以通用算子学习架构 [50, 51]为基础。在计算均匀化中，基于高维模型表示构建的神经网络也被用于取代小应变 [52] 和有限应变 [53, 54], 下的全场微观求解器，以加速机织复合材料的降阶模型 [55], ；傅里叶神经算子则被认为是通用线弹性均匀化器 [56]。

神经网络架构通常会根据具体问题的结构进行调整。例如，在图像处理中，卷积神经网络使用局部核，以反映相邻像素之间的相关性 [57]。同样，先验知识或归纳偏置也是设计旋转和平移不变网络 [58]及等变变换器 [59]的基础。这一原则也促使人们为微观力学均匀化设计物理信息和物理约束架构。Liu 等 [60, 61] 提出了深度材料网络（DMN），通过层合板层级结构来近似非均质 RVE 响应，这些层合板充当机制构件。对于每个构件，非弹性均匀化问题可以半解析地求解，因此计算成本较低；构件层级结构逐步生成有效响应。监督训练利用组分性质以及通过线性均匀化获得的有效刚度张量或柔度张量，来确定参数——各层合板的体积分数和旋转参考系。随后，所得网络可用于各相的非线性和非弹性材料行为。此类旋转 DMN 被扩展用于将迁移学习 [62], 应用于热力学问题 [63, 64] 和损伤问题 [65, 66], ，以及不确定性量化 [67], ；此外，人们还提出了其他网络构造，以提高其准确度 [68–70], ，例如图增强 [71] 和微极 [72]变体。旋转 DMN 还被用于工业环境中的短纤维增强复合材料 [73, 74]。文献 [75–77]提供了全面的概述和比较。

Gajek 等人 [78] 提供了理论见解和实践改进。他们证明，每个 DMN 都继承了其相材料的热力学一致性，且不受训练质量的影响，从而为每个经过训练的网络提供了物理保证。此外，他们还表明，对于具有消失记忆的材料，其有效响应在一阶近似下由线弹性均匀化函数决定。在实践方面，他们引入了直接 DMN，这类网络仅由体积分数和层合方向参数化。由于层合方向位于二维单位球面上，而一般旋转需要三个标量参数，因此直接 DMN 的训练参数少于 Liu 型 DMN。在离线训练中，Gajek 等人使用 Milton 公式 [79, 第 9] 章中的层合板构件，从而确保与深度学习框架的无缝集成。在在线阶段，他们引入了层级层合板的扁平化表示，该表示类似于传统离散化，并且相比层级实现可显著提高速度。随后，直接 DMN 被部署于弹塑性和热塑性问题的 FE^2 框架中 [80–82], ，还用于黏塑性和蠕变行为的逆向表征 [83]。其他应用还包括短纤维增强热塑性塑料的高度非线性载荷工况 [84], 、工艺诱发的不确定性的量化 [85], ，以及具有无限材料对比的纤维悬浮体 [86], ；此外，相关层合板构件的非退化性也在文献 [87] 中得到了研究。最近，Srinivas 等 [88] 引入了基于位移的公式，以加速离线训练。为进一步推广材料网络架构，Nguyen 和 Noels 引入了相互作用材料网络（IMN） [89, 90]。IMN 不再完全依赖层合板作为构件，而是采用更一般的相互作用机制，从而提供了额外的灵活性。Wei 等 [91]通过考虑晶体取向，将 IMN 扩展至多晶材料。

1.2 贡献

大多数先前研究都根据低成本的线性均匀化数据识别 DMN，随后将其与非线性组分本构律结合使用，并且这种向非弹性响应的迁移已被反复证明。因此，核心问题在于，弹性目标函数对历史相关均匀化算子的约束究竟有多强。Dey 等人 [92] 给出了一个尤其直接的示例：弹性验证误差非常相近的网络，可能表现出显著不同的蠕变误差。他们的非弹性信息驱动

策略根据非弹性验证误差，在存储的弹性训练迭代结果中进行选择，而无需计算非弹性训练梯度。Gajek 等人 [93] 随后表明，通过基于 Maxwell 材料本构结构的材料信息驱动采样和损失构造，可以显著改善弹性到黏弹性的迁移。Shin 等人 [94]则利用全场模拟的局部场信息，丰富了弹性训练。这些结果既展示了弹性识别的潜力，也揭示了其局限性：如果能够获得额外的本构信息，迁移效果可以得到改善，但由此形成的策略会受限于特定的材料类别。

Nguyen 和 Noels [90] 首次将从非线性轨迹直接辨识的方法引入基于相互作用的材料网络。他们的训练设置与本文所采用的设置较为接近：损失函数在由单调比例加载路径生成的均匀化响应轨迹上累积，并且可以将孔隙率等可选的均匀化描述符纳入目标函数。拟合参数的梯度通过对已收敛的网络平衡进行直接隐式微分获得。然而，对于其单调加载路径，先前已收敛的内部变量的敏感度不会在各增量之间传播。中间状态被视为单调路径上的子步进点。因此，相对于继承的本构状态，所得敏感度仅局限于当前已收敛步。

本文通过将完整加载轨迹上的平衡方程和本构更新方程视为耦合约束，消除了这一限制。已提交材料状态对可训练参数的依赖关系会从一个增量到下一个增量被显式保留。在直接形式下，该方法为完整的状态敏感度矩阵生成前向递推。在伴随形式下，同样的历史依赖性通过后续步之间的耦合向后传播。对于一般的历史依赖加载而言，这一区分至关重要，包括卸载、循环加载和非比例路径，因为在一般情况下，当前响应无法脱离先前已收敛状态而被独立微分。

当标量轨迹目标函数依赖于许多可训练参数时，伴随敏感性分析还具有额外优势，因为它无需针对每个参数传播一列状态敏感度。伴随方法广泛应用于敏感性分析和偏微分方程约束优化 [95–97]。在科学机器学习中，它还为神经常微分方程 [98] 和深度平衡模型 [99]提供了基础微分机制。此前已有研究使用非弹性轨迹数据训练材料网络 [90]。因此，本文工作的创新之处在于提出了路径一致的敏感度公式，并将其通过离散伴随实现应用于 DMN 识别。

本文有两项主要贡献。首先，将非弹性 DMN 训练表述为一个依赖历史的约束优化问题，并推导出完整的直接敏感度递推关系及相应的离散伴随。该公式仅需要宏观应力-应变轨迹：局部内部变量仍由本构演化方程控制，无需将其作为监督标签。其次，系统比较了不同网络秩和训练集规模下的弹性训练与非弹性训练，从而量化它们对预测精度、对未见加载历史的泛化能力、离线训练成本以及在线加速的影响。在跨网络秩比较中，我们区分了分布内精度与分布外泛化。

本文其余部分安排如下。第 2 节回顾 DMN 表示和弹性训练，并介绍非弹性训练及其直接灵敏度和伴随灵敏度，包括对这些灵敏度的数值验证。第 3 节针对短纤维复合材料，从准确度、计算成本、对未见加载路径的泛化能力以及向受扰动组分性质的迁移等方面比较弹性训练和非弹性训练。第 4 节总结经验教训。

2 深度材料网络与训练策略

2.1 DMN 表示

我们首先回顾 DMN 所近似的非线性均匀化问题，背景介绍参见 Suquet [100]。我们考虑一种微观结构 Y，其特征是由自由能密度 [101], 编码的非均质广义标准材料（GSM）

$$\Psi : Y \times \mathrm{Sym}(d) \times Z \to \mathbb{R}_{\geq 0}, \qquad (\boldsymbol{y}, \boldsymbol{\varepsilon}, \mathbf{z}) \mapsto \Psi(\boldsymbol{y}, \boldsymbol{\varepsilon}, \mathbf{z}), \tag{2.1}$$

以及耗散势

$$\Phi : Y \times Z \to \mathbb{R}_{\geq 0}, \qquad (\boldsymbol{y}, \dot{\mathbf{z}}) \mapsto \Phi(\boldsymbol{y}, \dot{\mathbf{z}}), \tag{2.2}$$

其中，对称二阶张量 $\boldsymbol{\varepsilon} \in$Sym($d$) 表示应变，而量 $\mathbf{z}$ 和 $\dot{\mathbf{z}}$分别表示巴拿赫空间 Z中的内变量及其变化率。给定时间区间 $[0, T]$上的宏观应变路径，

$$\bar{\boldsymbol{\varepsilon}} : [0, T] \to \mathrm{Sym}(d), \tag{2.3}$$

以及一个初始状态

$$\mathbf{z}_0 : Y \to Z, \tag{2.4}$$

均匀化将有效应力确定为体积平均值

$$\bar{\boldsymbol{\sigma}}(t) := \left\langle \frac{\partial \Psi}{\partial \boldsymbol{\varepsilon}}(\cdot, \bar{\varepsilon}(t) + \nabla^s \boldsymbol{u}(\cdot, t), \mathbf{z}(\cdot, t)) \right\rangle_Y \quad \text{for} \quad t \in [0, T]. \tag{2.5}$$

此处，位移涨落

$$\boldsymbol{u} : Y \times [0, T] \to \mathbb{R}^d \tag{2.6}$$

and 内变量 $\mathbf{z}$满足线动量平衡和Biot演化方程 ,

$$\mathrm{div} \left[\frac{\partial \Psi}{\partial \boldsymbol{\varepsilon}}(\cdot, \bar{\boldsymbol{\varepsilon}} + \nabla^s \boldsymbol{u}, \mathbf{z}) \right] = \mathbf{0}, \tag{2.7a}$$

$$\frac{\partial \Psi}{\partial \mathbf{z}}(\boldsymbol{y}, \bar{\varepsilon}(t) + \nabla^s \boldsymbol{u}(\boldsymbol{y}, t), \mathbf{z}(\boldsymbol{y}, t)) + \frac{\partial \Phi}{\partial \dot{\mathbf{z}}}(\boldsymbol{y}, \dot{\mathbf{z}}) = 0 \quad \text{for} \quad (\boldsymbol{y}, t) \in Y \times [0, T]. \tag{2.7b}$$

等价地，单胞定义了一个非线性、依赖历史的均匀化算子 $\mathcal{M}_Y$，该算子将截至当前时刻 t 的宏观应变历史以及初始微观状态映射为当前有效应力，参见例如文献 [78, 81], 。

$$\bar{\boldsymbol{\sigma}}(t) = \mathcal{M}_Y \left(\bar{\boldsymbol{\varepsilon}}|_{[0,t]}, \mathbf{z}_0 \right) \quad \text{for} \quad t \in [0, T]. \tag{2.8}$$

DMN为该算子提供了一个有限维且满足力学可容许性的代理模型。在直接DMN [78, 80, 81], 中，网络被组织为定向层合体的层级结构。对于本文所考虑的两相材料，秩为 K 的网络构成一个满二叉树，其中包含 $M = 2^K$ 个材料点和 $N = M - 1$ 个内部节点，如图1所示。叶节点被赋予两相的本构定律，并携带相应的材料点历史。特别是，它们不局限于线性弹性。内部节点通过其层合方向定义运动学相互作用，而材料点权重则决定叶节点的相对贡献。

不同于递归地计算层级结构，直接DMN可以展平成一种类似于传统离散化的形式 [78]。在全文中，根据 Mandel 记号，对称二阶张量 $\boldsymbol{\tau} \in$Sym(3) 均与空间 R^6 中的向量对应

$$\boldsymbol{\tau} \mapsto \left(\tau_{11}, \tau_{22}, \tau_{33}, \sqrt{2}\,\tau_{12}, \sqrt{2}\,\tau_{13}, \sqrt{2}\,\tau_{23} \right)^T, \tag{2.9}$$

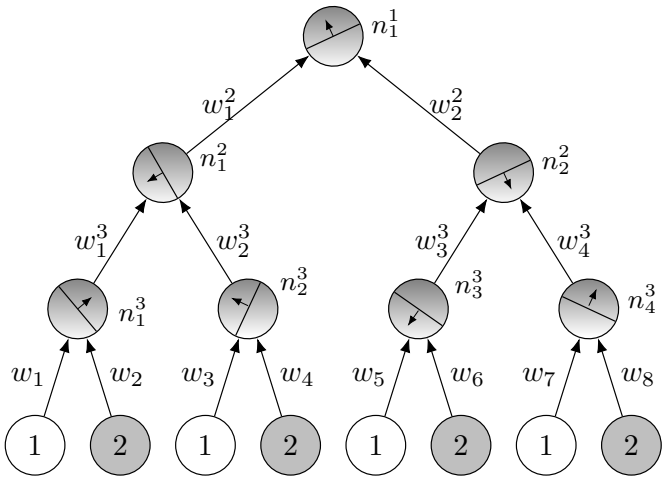


图1：二元复合材料的秩3直接DMN。 $M = 2^3$ 叶节点表示分配给相1或相2的材料点，并携带相应的本构响应，包括任何随时间演化的内部变量。 $N = M - 1$ 内部节点通过层合方向$\mathbf{n}$定义层次运动学，而边标签则表示相应的权重。因此，无论组分的响应是弹性还是非弹性，网络拓扑都保持不变。

从而张量的弗罗贝尼乌斯内积与相应向量的欧几里得内积相一致。本文中的所有应变和应力向量，包括给定的宏观应变和有效应力，均指这一表示形式。收集所有材料点处应变的向量记为

$$\vec{\boldsymbol{\varepsilon}} = A\bar{\boldsymbol{\varepsilon}} + D(\boldsymbol{w}, \boldsymbol{n})\boldsymbol{u}, \tag{2.10}$$

其中，矩阵 A 在材料点处复现给定的宏观应变，矩阵$D = D(\boldsymbol{w}, \boldsymbol{n})$表示相容对称梯度算子，向量 $\boldsymbol{u}$ 则汇集与内部节点相关的位移跳跃。由于每个局部波动应变都表示为位移跳跃与层合方向的对称秩一乘积，因此满足相容性。Gajek 等人 [78]给出了矩阵 D 的一般构造方法。为清楚起见，秩3网络的显式形式为

$$D = \begin{bmatrix} c_2^2 B(n_1^1) & c_2^3 B(n_1^2) & \mathbf{0} & c_2^4 B(n_1^3) & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ c_2^2 B(n_1^1) & c_2^3 B(n_1^2) & \mathbf{0} & -c_1^4 B(n_1^3) & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ c_2^2 B(n_1^1) & -c_1^3 B(n_1^2) & \mathbf{0} & \mathbf{0} & c_4^4 B(n_2^3) & \mathbf{0} & \mathbf{0} \\ c_2^2 B(n_1^1) & -c_1^3 B(n_1^2) & \mathbf{0} & \mathbf{0} & -c_3^4 B(n_2^3) & \mathbf{0} & \mathbf{0} \\ -c_1^2 B(n_1^1) & \mathbf{0} & c_4^3 B(n_2^2) & \mathbf{0} & \mathbf{0} & c_6^4 B(n_3^3) & \mathbf{0} \\ -c_1^2 B(n_1^1) & \mathbf{0} & c_4^3 B(n_2^2) & \mathbf{0} & \mathbf{0} & -c_5^4 B(n_3^3) & \mathbf{0} \\ -c_1^2 B(n_1^1) & \mathbf{0} & -c_3^3 B(n_2^2) & \mathbf{0} & \mathbf{0} & \mathbf{0} & c_8^4 B(n_4^3) \\ -c_1^2 B(n_1^1) & \mathbf{0} & -c_3^3 B(n_2^2) & \mathbf{0} & \mathbf{0} & \mathbf{0} & -c_7^4 B(n_4^3) \end{bmatrix}, \tag{2.11}$$

其中，标量 c_i^k 表示层级 k 中第 i 个节点相对于其父节点的体积分数，

$$c_i^k := \frac{w_i^k}{w_{2j-1}^k + w_{2j}^k}, \qquad j = \lceil i/2 \rceil, \tag{2.12}$$

而基体 B 表示对称化秩一映射

$$\mathbf{u} \mapsto \frac{1}{2}\left(\mathbf{n} \otimes \mathbf{u} + \mathbf{u} \otimes \mathbf{n}\right), \tag{2.13}$$

其形式为

$$B(\mathbf{n}) = \begin{bmatrix} n_1 & 0 & 0 \\ 0 & n_2 & 0 \\ 0 & 0 & n_3 \\ n_2/\sqrt{2} & n_1/\sqrt{2} & 0 \\ n_3/\sqrt{2} & 0 & n_1/\sqrt{2} \\ 0 & n_3/\sqrt{2} & n_2/\sqrt{2} \end{bmatrix} \tag{2.14}$$

采用 Mandel 记号（2.9）。在每个材料点处，局部本构律将应变历史和材料状态映射为应力。DMN 的总自由能是各材料点自由能的加权和。对位移跳跃施加平稳性条件，可得离散平衡方程

$$D^T W \vec{\boldsymbol{\sigma}}(A\bar{\boldsymbol{\varepsilon}} + D(\boldsymbol{w}, \boldsymbol{n})\boldsymbol{u}, z) = \mathbf{0}, \tag{2.15}$$

其中，分块对角矩阵 $W = W(\boldsymbol{w})$包含材料点权重，向量 z 收集材料点的内部变量，向量 $\vec{\boldsymbol{\sigma}}$ 汇集相应的应力。请注意，由 Gajek 等人 [78], 提出的方程（2.15）为 DMN 的非线性在线问题提供了一种全局组装、无旋转、类似有限元的表述；而较早的 DMN [60, 61, 65] 则沿层次网络通过递归均匀化和反均匀化来评估非线性响应。最终，方程（2.10）和（2.15）定义了用于弹性训练和非弹性训练的网络表示。具体而言，矩阵 D 和 W 描述学习得到的层次运动学和类求积权重，而分配给材料点的本构定律决定了所得算子是弹性的还是非弹性的。

2.2 网络参数化与弹性训练

材料网络表示由非负的材料点权重 $\boldsymbol{w}$ 和单位层合方向 $\boldsymbol{n}$ 决定。为进行优化，这些受约束量通过无约束的可训练坐标表示。材料点权重通过无约束的逻辑值 $\boldsymbol{\xi} \in \mathrm{R}^M$ 进行参数化，并采用当前实现中使用的带下限的 softmax,

$$w_i = w_{\min} + (1 - M w_{\min}) \frac{\exp(\xi_i)}{\sum_{j=1}^{M} \exp(\xi_j)}, \qquad i = 1, \ldots, M, \qquad w_{\min} = 10^{-9}. \tag{2.16}$$

因此，条件

$$w_i \geq w_{\min}, \qquad i = 1, \ldots, M, \qquad \text{and} \qquad \sum_{i=1}^{M} w_i = 1 \tag{2.17}$$

通过构造即可满足。对于每个内部节点，单位层合方向用球坐标 [78] 表示为

$$\boldsymbol{n}_j^i = \begin{bmatrix} \sin\alpha_j^i \cos\beta_j^i \\ \sin\alpha_j^i \sin\beta_j^i \\ \cos\alpha_j^i \end{bmatrix}, \tag{2.18}$$

其中，由于周期性和层合方向的符号不变性，这些角度可视为无约束的实坐标，其代表性范围为

$$\alpha_j^i \in [0, \pi), \tag{2.19a}$$
$$\beta_j^i \in [0, 2\pi). \tag{2.19b}$$

为简　　洁起见，我们将无约束的权重变量和所有角坐标汇集到向量　　or

$$\boldsymbol{\theta} := (\boldsymbol{\xi}, \boldsymbol{\alpha}, \boldsymbol{\beta}) \in \mathbb{R}^{N_\theta} \tag{2.20}$$

中，作为将在以下各节中通过弹性训练和非弹性训练策略确定的可训练参数。对于给定参数 $\boldsymbol{\theta}$，权重 $\boldsymbol{w}$ 和法向方向 $\boldsymbol{n}$ 通过算子 $D(\boldsymbol{\theta})$ 和 $W(\boldsymbol{\theta})$ 确定 DMN 表示。

弹性训练是针对同一 DMN 表示的一种识别策略。如果组成材料定律是线弹性的，则每个有向层合节点都具有解析均质化映射，

$$\mathbb{C}_j^i = \mathcal{G}\big(\mathbb{C}_{2j-1}^{i+1}, \mathbb{C}_{2j}^{i+1}, c_{2j-1}^{i+1}, c_{2j}^{i+1}, \boldsymbol{n}_j^i\big), \tag{2.21}$$

其中，四阶张量 C_{2j-1}^{i+1} 和 C_{2j}^{i+1} 表示两个子节点的均质化刚度，标量 c_{2j-1}^{i+1} 和 c_{2j}^{i+1} 表示它们的相对体积分数（2.12）。对于有向层合结构，父节点刚度 C_j^i 可根据 Milton 公式 [79, 第 9] 章进行计算

$$\Big[\big(\mathbb{C}_j^i - \mathbb{C}^0\big)^{-1} + \mathbb{\Gamma}^0\big(\boldsymbol{n}_j^i\big)\Big]^{-1} = c_{2j-1}^{i+1}\Big[\big(\mathbb{C}_{2j-1}^{i+1} - \mathbb{C}^0\big)^{-1} + \mathbb{\Gamma}^0\big(\boldsymbol{n}_j^i\big)\Big]^{-1} + c_{2j}^{i+1}\Big[\big(\mathbb{C}_{2j}^{i+1} - \mathbb{C}^0\big)^{-1} + \mathbb{\Gamma}^0\big(\boldsymbol{n}_j^i\big)\Big]^{-1}. \tag{2.22}$$

这里，四阶张量 C^0 表示参考刚度，四阶张量 $\mathbb{\Gamma}^0(\boldsymbol{n})$ 表示相应的层合算子。该算子作用于对称张量 $\boldsymbol{\tau}$，通过 $\mathbb{\Gamma}^0(\boldsymbol{n}) : \boldsymbol{\tau} := \operatorname{sym}\big(\boldsymbol{K}^0(\boldsymbol{n})^{-1}(\boldsymbol{\tau}\boldsymbol{n}) \otimes \boldsymbol{n}\big)$ 表示，其中声学张量为 $\boldsymbol{K}^0(\boldsymbol{n})\boldsymbol{a} := (\mathsf{C}^0 : (\boldsymbol{a} \otimes \boldsymbol{n}))\boldsymbol{n}$ [79, 第 9] 章。参考刚度 C^0 需要选择为使 Milton 公式（2.22）中出现的所有差值 $\mathsf{C} - \mathsf{C}^0$ 均为定号，例如可取为对称张量空间上单位张量的某个倍数，其模量大于所涉及全部刚度中的最大特征值；有关讨论请参见 Kabel等人 [102] 。这种递归是网络的线弹性特化形式，而不是其一般定义。它可以通过展平位移 formulation 递归求值，也可以通过其他等价的层合公式求值。Milton 公式的直接实现需要反复进行四阶张量求逆。作为替代方案，Srinivas 等 [88] 引入了基于位移的层合 formulation，以加速离线训练。

在传统的弹性训练中，可训练参数 $\boldsymbol{\theta}$ 通过线性响应加以识别，尤其是有效刚度张量。为评估拟合质量，可以使用不同的损失函数。在本文中，我们最小化有效刚度张量相对Frobenius误差的均方根，

$$\mathcal{L}(\boldsymbol{\theta}) := \sqrt{\frac{1}{N_s}\sum_{s=1}^{N_s}\left(\frac{\|\bar{\mathbb{C}}_s^{\mathrm{pred}}(\boldsymbol{\theta}) - \bar{\mathbb{C}}_s^{\mathrm{reference}}\|_F}{\|\bar{\mathbb{C}}_s^{\mathrm{reference}}\|_F}\right)^2} \longrightarrow \min_{\boldsymbol{\theta}}, \tag{2.23}$$

其中，整数 N_s 表示弹性样本的数量，四阶张量 $\bar{\mathsf{C}}_s^{\mathrm{pred}}$和$\bar{\mathsf{C}}_s^{\mathrm{reference}}$分别表示预测的和参考有效刚度张量。在本文中，$\|\cdot\|_F$表示张量、矩阵或完整轨迹数组的Frobenius范数，而 $\|\cdot\|_2$表示向量的欧几里得范数。由于弹性问题不包含随时间演化的材料历史，梯度可直接由链式法则得到，形式如下

$$\frac{\partial \mathcal{L}}{\partial \boldsymbol{\theta}} = \sum_{s=1}^{N_s} \frac{\partial \mathcal{L}}{\partial \bar{\mathbb{C}}_s^{\mathrm{pred}}} \frac{\partial \bar{\mathbb{C}}_s^{\mathrm{pred}}}{\partial \boldsymbol{\theta}}, \tag{2.24}$$

并且可以对代数层合板或基于位移的评估应用自动微分。这个较简单的求导问题并不意味着弹性辨识的成本低廉。生成一个参考有效刚度张量需要求解多个相互独立的全场线性均匀化问题；据报道，对于工业规模微观结构，采样足够广泛的材料属性空间的成本相当高 [103]。下文数值示例报告的数据生成耗时，为当前问题提供了直接的比较。

2.3 非弹性训练

2.3.1 训练公式

非弹性训练的目标是直接从非弹性响应轨迹中识别 DMN。线性均匀化产生单个有效刚度张量，而公式（2.8）中的非线性均匀化算子则将给定的应变历史映射为有效应力历史。因此，可训练网络参数通过最小化直接根据有效应力轨迹定义的损失函数获得，

$$\mathcal{L}(\boldsymbol{\theta}) := \frac{1}{N_s}\sum_{s=1}^{N_s}\left(\frac{1}{N_t}\sum_{n=1}^{N_t}\frac{\|\bar{\boldsymbol{\sigma}}_{s,n}^{\mathrm{pred}}(\boldsymbol{\theta}) - \bar{\boldsymbol{\sigma}}_{s,n}^{\mathrm{reference}}\|_2}{\|\bar{\boldsymbol{\sigma}}_{s,n}^{\mathrm{reference}}\|_2}\right) \longrightarrow \min_{\boldsymbol{\theta}}, \tag{2.25}$$

其中，整数 N_s 和 N_t 分别表示加载轨迹数和时间步数，而向量 $\bar{\boldsymbol{\sigma}}_{s\,n}^{\mathrm{pred}}$、和 $\bar{\boldsymbol{\sigma}}_{s\,n}^{\mathrm{reference}}$、分别表示网络预测的以及参考解得到的时间步 n 处的有效应力。求和从第一个加载步开始，因此参考应力不会为零；相对误差对所有加载步赋予相同权重，而与应力水平无关。实践中，欧几里得范数在残差为零处的不可微性并无影响。与第 3.1 节中引入的测试误差（3.6）不同，后者度量完整轨迹上的偏差，而逐步归一化避免了具有较大应力的后期加载步在训练中占据主导地位。形式上，梯度可由以下关系得到

$$\frac{\partial \mathcal{L}}{\partial \boldsymbol{\theta}} = \sum_{s=1}^{N_s}\sum_{n=1}^{N_t}\frac{\partial \mathcal{L}}{\partial \bar{\boldsymbol{\sigma}}_{s,n}^{\mathrm{pred}}}\frac{\mathrm{D}\bar{\boldsymbol{\sigma}}_{s,n}^{\mathrm{pred}}}{\mathrm{D}\boldsymbol{\theta}}. \tag{2.26}$$

然而，第二个因子包含了完整时间离散解对网络参数的隐式依赖。在时间步 n，位移跳跃向量 $\boldsymbol{u}_n \in \mathrm{R}^{3N}$ 和材料点变量向量$\boldsymbol{v}_n \in \mathrm{R}^{Mn_v}$ 取决于时间步 $n-1$ 的收敛解。这里，向量 $\boldsymbol{v}_n$ 汇集了所有材料点处的应力和内部变量，即每个材料点有 $n_v = 6+n_z$ 个标量，其中包括 n_z 个标量内部变量；基体 $\boldsymbol{S}_\sigma$ 用于提取应力分量。根据参考文献 [95], ，我们引入时间离散的全局残差和局部残差

$$\boldsymbol{R}_n(\boldsymbol{v}_n;\boldsymbol{\theta}) := D^T(\boldsymbol{\theta})W(\boldsymbol{\theta})\boldsymbol{S}_\sigma\boldsymbol{v}_n \tag{2.27}$$

and

$$\boldsymbol{H}_n(\boldsymbol{u}_n,\boldsymbol{v}_n,\boldsymbol{u}_{n-1},\boldsymbol{v}_{n-1};\boldsymbol{\theta}) := \boldsymbol{H}(\boldsymbol{v}_n,\boldsymbol{\varepsilon}_n,\boldsymbol{v}_{n-1},\boldsymbol{\varepsilon}_{n-1}) \quad \text{with} \quad \boldsymbol{\varepsilon}_n := A\bar{\boldsymbol{\varepsilon}}_n + D(\boldsymbol{\theta})\boldsymbol{u}_n, \tag{2.28}$$

它们分别取值于空间 R^{3N} 和 R^{Mn_v} 中，并分别对应线动量平衡（2.7a）和 Biot演化方程（2.7b）。因此，DMN 参数的辨识是一个历史相关约束优化问题

$$\underset{\boldsymbol{\theta}}{\text{minimize}}\ \mathcal{L}(\boldsymbol{\theta}) \qquad \text{subject to} \qquad \boldsymbol{R}_n = \boldsymbol{0}, \quad \boldsymbol{H}_n = \boldsymbol{0}, \quad n = 1,\ldots,N_t. \tag{2.29}$$

尽管损失函数（2.25）仅根据有效应力定义，但它对网络参数的依赖仍通过继承的本构状态的完整序列得以保留。因此，优化问题（2.29）中的耦合约束代表了完整的时间离散依赖链，并为路径一致的直接灵敏度和伴随灵敏度提供了基础。

2.3.2 隐式伴随微分

在推导伴随方程之前，我们先概述其基本策略。伴随方法对收敛后的增量方程求导，而不是对用于求解这些方程的牛顿迭代或返回映射迭代求导。与对展开后的前向求解器应用反向模式自动微分相比，该策略有若干优点。首先，内存需求取决于轨迹的收敛状态，与牛顿迭代和返回映射迭代的次数无关。其次，所计算的梯度不依赖于所采用的非线性求解器、初始猜测或迭代历史。第三，每个时间步的反向扫描只需对转置线性化状态方程求解一次，与可训练参数的数量无关；相比之下，第 2.3.3 节讨论的直接灵敏度则不同。缺点是需要存储收敛后的正向轨迹；除了常规的一致切线外，还需提供本构更新对前一状态的导数；并且所计算的梯度精度仅能达到前向求解器的容差。

为便于说明，我们考虑单条加载轨迹，因此，一个批次的贡献可通过对所有轨迹上得到的梯度求和获得。在每个时间步，我们将未知量汇集到状态向量中

$$\mathbf{q}_n := (\boldsymbol{u}_n, \boldsymbol{v}_n) \in \mathbb{R}^{3N} \times \mathbb{R}^{Mn_v} \tag{2.30}$$

并将两个残差合并为函数

$$\boldsymbol{F}_n(\mathbf{q}_n, \mathbf{q}_{n-1}; \boldsymbol{\theta}) := \begin{bmatrix} \boldsymbol{R}_n(\boldsymbol{v}_n; \boldsymbol{\theta}) \\ \boldsymbol{H}_n(\boldsymbol{u}_n, \boldsymbol{v}_n, \boldsymbol{u}_{n-1}, \boldsymbol{v}_{n-1}; \boldsymbol{\theta}) \end{bmatrix}, \tag{2.31}$$

因此，约束可写为 $\boldsymbol{F}_n = \mathbf{0}$ 对于 $n = 1, \ldots, N_t$。初始状态 $\mathbf{q}_0$ 已给定，因此与可训练参数无关。

W我们定义时间步 n处的损失贡献 as

$$L_n(\boldsymbol{\theta}) := \frac{1}{N_t} \frac{\|\bar{\boldsymbol{\sigma}}_n(\boldsymbol{\theta}) - \bar{\boldsymbol{\sigma}}_n^{\text{reference}}\|_2}{\|\bar{\boldsymbol{\sigma}}_n^{\text{reference}}\|_2}, \tag{2.32}$$

其形式取决于有效应力

$$\bar{\boldsymbol{\sigma}}_n(\boldsymbol{\theta}) := \bar{A}(\boldsymbol{\theta}) \boldsymbol{S}_\sigma \boldsymbol{v}_n \tag{2.33}$$

¯在时间步 n处的值，该值借助加权平均算子 $A(\boldsymbol{\theta})$计算得到。沿轨迹累积的损失定义了目标函数

$$\mathcal{J}(\boldsymbol{\theta}) := \sum_{n=1}^{N_t} L_n(\boldsymbol{\theta}). \tag{2.34}$$

对于约束优化问题（2.29），我们引入伴随拉格朗日函数

$$\mathcal{L}_{\text{adj}} := \mathcal{J}(\boldsymbol{\theta}) - \sum_{n=1}^{N_t} \boldsymbol{\lambda}_n^T \boldsymbol{F}_n = \sum_{n=1}^{N_t} L_n - \sum_{n=1}^{N_t} \boldsymbol{\gamma}_n^T \boldsymbol{R}_n - \sum_{n=1}^{N_t} \boldsymbol{\eta}_n^T \boldsymbol{H}_n, \tag{2.35}$$

其中，向量 $\boldsymbol{\lambda}_n = (\boldsymbol{\gamma}_n, \boldsymbol{\eta}_n) \in \mathrm{R}^{3N} \times \mathrm{R}^{Mn_v}$ 分别表示与全局残差 (2.27) 和局部残差 (2.28) 相关的伴随变量。具体而言，伴随变量属于

与相应未知量相同的空间，即伴随变量 $\boldsymbol{\gamma}_n$ 是伴随位移跳跃向量，而伴随变量 $\boldsymbol{\eta}_n$ 则与材料点处的应力和内部变量相关。对于收敛的正向解，约束 $\boldsymbol{F}_n = \mathbf{0}$ 得到满足，因此目标函数 $\mathcal{J}$ 和伴随拉格朗日函数 $\mathcal{L}_{\mathrm{adj}}$ 具有相同的敏感度。对伴随拉格朗日函数 (2.35) 关于参数向量 $\boldsymbol{\theta}$ 求导，并对乘以状态敏感度的各项进行归组，可得关系式

$$\begin{aligned}\frac{\mathrm{D}\mathcal{J}}{\mathrm{D}\boldsymbol{\theta}} = &\sum_{n=1}^{N_t}\left(\frac{\partial L_n}{\partial \boldsymbol{\theta}} - \boldsymbol{\lambda}_n^T\frac{\partial \boldsymbol{F}_n}{\partial \boldsymbol{\theta}}\right)\\ &+\sum_{n=1}^{N_t}\left[\frac{\partial L_n}{\partial \mathbf{q}_n} - \boldsymbol{\lambda}_n^T\frac{\partial \boldsymbol{F}_n}{\partial \mathbf{q}_n} - \boldsymbol{\lambda}_{n+1}^T\frac{\partial \boldsymbol{F}_{n+1}}{\partial \mathbf{q}_n}\right]\frac{\mathrm{D}\mathbf{q}_n}{\mathrm{D}\boldsymbol{\theta}},\end{aligned} \tag{2.36}$$

其中终端伴随变量设为 $\boldsymbol{\lambda}_{N_t+1} := \mathbf{0}$。在此及下文中，关于参数向量 $\boldsymbol{\theta}$ 的偏导数均在当前时间步 n 处计算，并将状态变量固定。若伴随变量满足以下方程，则关系式 (2.36) 中的状态敏感性项将消失

$$\left(\frac{\partial \boldsymbol{F}_n}{\partial \mathbf{q}_n}\right)^T\boldsymbol{\lambda}_n = \left(\frac{\partial L_n}{\partial \mathbf{q}_n}\right)^T - \left(\frac{\partial \boldsymbol{F}_{n+1}}{\partial \mathbf{q}_n}\right)^T\boldsymbol{\lambda}_{n+1}, \qquad n = N_t, \dots, 1. \tag{2.37}$$

这些伴随方程按时间倒序求解。涉及残差 $\boldsymbol{F}_{n+1}$ 的项将敏感性从后续时间步迁移到当前时间步，从而保留对本构历史的完整依赖。

根据定义 (2.31)，我们引入以下四个分块

$$\boldsymbol{H}_n^{\boldsymbol{v}} := \frac{\partial \boldsymbol{H}_n}{\partial \boldsymbol{v}_n}, \tag{2.38a}$$

$$\boldsymbol{H}_n^{\boldsymbol{\varepsilon}} := \frac{\partial \boldsymbol{H}_n}{\partial \boldsymbol{\varepsilon}_n}, \tag{2.38b}$$

$$\boldsymbol{H}_n^{\boldsymbol{v},-} := \frac{\partial \boldsymbol{H}_n}{\partial \boldsymbol{v}_{n-1}}, \tag{2.38c}$$

$$\boldsymbol{H}_n^{\boldsymbol{\varepsilon},-} := \frac{\partial \boldsymbol{H}_n}{\partial \boldsymbol{\varepsilon}_{n-1}} \tag{2.38d}$$

它们属于线性化本构残差。对于仅通过有效应力依赖状态的损失函数，伴随方程 (2.37) 可转化为耦合的全局-局部伴随系统

$$\begin{cases}\boldsymbol{\gamma}_n^T D^T W \boldsymbol{S}_\sigma + \boldsymbol{\eta}_n^T\boldsymbol{H}_n^{\boldsymbol{v}} + \boldsymbol{\eta}_{n+1}^T\boldsymbol{H}_{n+1}^{\boldsymbol{v},-} = \dfrac{\partial L_n}{\partial \bar{\boldsymbol{\sigma}}_n}\bar{A}\boldsymbol{S}_\sigma, \\ \boldsymbol{\eta}_n^T\boldsymbol{H}_n^{\boldsymbol{\varepsilon}} D + \boldsymbol{\eta}_{n+1}^T\boldsymbol{H}_{n+1}^{\boldsymbol{\varepsilon},-} D = \mathbf{0},\end{cases} \qquad n = N_t, \dots, 1, \tag{2.39}$$

采用约定 $\boldsymbol{\eta}_{N_t+1} := \mathbf{0}$表示局部伴随变量。伴随¯系统（2.39）中的所有网络算子，包括算子 D、 W 和 A，均在当前参数向量 $\boldsymbol{\theta}$ 处进行计算。伴随反向迭代在已收敛的正向轨迹上执行。在每个时间步，均在存储的已收敛状态下以解析方式重新计算本构雅可比矩阵块，并从 $n = N_t$ 到 $n = 1$反向求解伴随系统（2.39）。每个时间步需要进行一次耦合伴随求解，而此类求解的次数与可训练网络参数的数量无关。对于数值算例中采用的 J_2 塑性模型，这些本构雅可比矩阵块可根据弹性预测和塑性返回映射关系 [95] 以解析方式获得。将通过伴随系统（2.39）得到的伴随变量代入关系式（2.36），即可消除状态敏感度，仅保留网络算子的显式导数。对于任意参数扰动

$\delta\boldsymbol{\theta} \in \mathrm{R}^{N_\theta}$，相应的方向导数为

$$\begin{aligned}\mathrm{D}\mathcal{J}(\boldsymbol{\theta})[\delta\boldsymbol{\theta}] = \sum_{n=1}^{N_t} \Big\{ & \frac{\partial L_n}{\partial \bar{\boldsymbol{\sigma}}_n}\big[\delta\bar{A}\,\boldsymbol{S}_\sigma \boldsymbol{v}_n\big] \\ & - \boldsymbol{\gamma}_n^T\big[(\delta D)^T W \boldsymbol{S}_\sigma \boldsymbol{v}_n + D^T(\delta W)\boldsymbol{S}_\sigma \boldsymbol{v}_n\big] \\ & - \boldsymbol{\eta}_n^T\big[\boldsymbol{H}_n^{\boldsymbol{\varepsilon}}(\delta D)\boldsymbol{u}_n + \boldsymbol{H}_n^{\boldsymbol{\varepsilon},-}(\delta D)\boldsymbol{u}_{n-1}\big]\Big\},\end{aligned} \tag{2.40}$$

其中，矩阵 δD、δW 和 δA 分别表示算子 D、W 和 A 在参数向量 $\boldsymbol{\theta}$ 处沿方向 $\delta\boldsymbol{\theta}$ 的方向导数。方向导数 (2.40) 中的每一项都是相对于参数向量 $\boldsymbol{\theta}$ 的向量-雅可比积，因此无需计算逐参数敏感度 $\mathrm{D}\boldsymbol{u}_n/\mathrm{D}\boldsymbol{\theta}$ 和 $\mathrm{D}\boldsymbol{v}_n/\mathrm{D}\boldsymbol{\theta}$，即可求得损失梯度。自动微分仅用于算子 $D(\boldsymbol{\theta})$、$W(\boldsymbol{\theta})$ 和 $A(\boldsymbol{\theta})$ 的显式参数化，而本构模块则由材料更新提供。

2.3.3 隐式直接微分

为便于参考并验证隐式伴随公式，我们还考虑对收敛后的时间离散问题进行直接微分。对于单条轨迹，对目标函数 (2.34) 关于参数向量 $\boldsymbol{\theta}$ 求导，可得关系式

$$\frac{\mathrm{D}\mathcal{J}}{\mathrm{D}\boldsymbol{\theta}} = \sum_{n=1}^{N_t} \frac{\partial L_n}{\partial \bar{\boldsymbol{\sigma}}_n}\frac{\mathrm{D}\bar{\boldsymbol{\sigma}}_n}{\mathrm{D}\boldsymbol{\theta}}, \tag{2.41}$$

其中，有效应力 (2.33) 的直接敏感度为

$$\frac{\mathrm{D}\bar{\boldsymbol{\sigma}}_n}{\mathrm{D}\boldsymbol{\theta}} = \frac{\mathrm{D}\bar{A}}{\mathrm{D}\boldsymbol{\theta}}\boldsymbol{S}_\sigma \boldsymbol{v}_n + \bar{A}\boldsymbol{S}_\sigma \frac{\mathrm{D}\boldsymbol{v}_n}{\mathrm{D}\boldsymbol{\theta}}. \tag{2.42}$$

根据状态向量的定义 (2.30)，我们引入直接状态敏感度矩阵

$$\boldsymbol{Q}_n := \frac{\mathrm{D}\mathbf{q}_n}{\mathrm{D}\boldsymbol{\theta}} \in \mathbb{R}^{(3N+Mn_v)\times N_\theta}. \tag{2.43}$$

对收敛约束 $\boldsymbol{F}_n = \mathbf{0}$ 关于参数向量 $\boldsymbol{\theta}$ 求导，得到方程

$$\frac{\partial \boldsymbol{F}_n}{\partial \mathbf{q}_n}\boldsymbol{Q}_n + \frac{\partial \boldsymbol{F}_n}{\partial \mathbf{q}_{n-1}}\boldsymbol{Q}_{n-1} + \frac{\partial \boldsymbol{F}_n}{\partial \boldsymbol{\theta}} = \mathbf{0}, \tag{2.44}$$

因此，递推关系为

$$\boldsymbol{Q}_n = -\left(\frac{\partial \boldsymbol{F}_n}{\partial \mathbf{q}_n}\right)^{-1}\left[\frac{\partial \boldsymbol{F}_n}{\partial \boldsymbol{\theta}} + \frac{\partial \boldsymbol{F}_n}{\partial \mathbf{q}_{n-1}}\boldsymbol{Q}_{n-1}\right]. \tag{2.45}$$

递推式 (2.45) 按时间正向计算，并保留对完整材料历史的依赖。特别是，项 $(\partial \boldsymbol{F}_n/\partial \mathbf{q}_{n-1})\boldsymbol{Q}_{n-1}$ 传播了前一增量中已提交状态的敏感度。省略此项后，得到的是相对于继承状态的步局部敏感度，这正是 Nguyen 和 Noels [90]针对单调训练路径采用的简化方法。直接公式在数学上十分简单明了，但状态敏感度矩阵 $\boldsymbol{Q}_n$包含每个设计变量对应的一列敏感度，并且必须沿完整加载历史进行传播。由于递推式 (2.45) 中仅有矩阵 $\boldsymbol{Q}_{n-1}$ 参与计算，直接方法的内存需求为 $\mathcal{O}((3N + Mn_v)\ N_\theta)$ 量级。更重要的是，每个时间步都需要使用矩阵 $\partial \boldsymbol{F}_n/\partial \mathbf{q}_n$ 求解一个含有 N_θ 个右端项的线性系统，因此计算量会随可训练参数的数量线性增长。对于当前参数化，一个秩为 K 的二进制

DMN 有 $N_\theta = M + 2N$ 个可训练坐标，即秩为 6 时有 $N_\theta = 190$ 个，秩为 8 时有 $N_\theta = 766$ 个。作为比较，Nguyen 和 Noels 报道的最大基于相互作用的材料网络包含 415 个拟合参数 [90]。相比之下，伴随系统 (2.39) 每个时间步只需使用转置矩阵求解一次，与可训练参数的数量 N_θ 无关，但需要存储收敛后的正向轨迹，即内存需求为 $\mathcal{O}(N_t(3N+Mn_v))$ 量级。对于此处考虑的网络规模，两种方法的内存需求都适中，而伴随公式的优势在于每次梯度计算所需的线性求解次数较少。其余与参数相关的计算包括对网络算子的显式参数化进行向量-雅可比积运算，这些运算通过反向模式自动微分完成。

2.3.4 梯度验证

我们使用确定性单轨迹基准，通过中心有限差分对三种实现进行了验证，并且该验证独立于训练过程：第 2.3.3 节的直接微分、第 2.3.2 节的离散伴随，以及基于 JAX 自定义_vjp [104]、独立实现的隐式反向模式实现。后者为收敛的非线性 DMN 步分配了自定义向量-雅可比积：其反向规则求解转置的约化切线系统，并应用残差反向传播，而无需对牛顿迭代进行微分。因此，它实现了与伴随方程 (2.37) 相同的隐式离散伴随关系，同时为验证提供了独立的自动微分实现。为此，我们规定了比例宏观应变历史

$$\bar{\boldsymbol{\varepsilon}}_n := a_n \begin{bmatrix} 1 & -0.5 & -0.5 & 0.12 & 0 & 0 \end{bmatrix}^T, \tag{2.46}$$

在 Mandel 记号 (2.9) 中，其中载荷因子取值为

$$a_n \in \{0,\, 0.002,\, 0.004,\, 0.006,\, 0.008,\, 0.010,\, 0.006,\, 0.003,\, 0.007,\, 0.011\}. \tag{2.47}$$

因此，该轨迹包含加载、卸载和再加载，使得敏感性计算包括本构历史的传播。所采用的组分材料参数与计算示例部分中的相同。对于每个网络秩，均使用确定性参数初始化来计算目标有效应力轨迹 $\bar{\boldsymbol{\sigma}}_n^{\text{target}}$。为了在不进行优化的情况下获得非零目标函数和梯度，将参考轨迹构造为

$$\bar{\boldsymbol{\sigma}}_n^{\text{reference}}: = 0.97\,\bar{\boldsymbol{\sigma}}_n^{\text{target}} + 10^{-5} \begin{bmatrix} 1 & -0.5 & 0.25 & 0.1 & -0.05 & 0.075 \end{bmatrix}^T, \quad \text{for} \quad \begin{matrix} n = 1、\ldots、 \\ 9\text{</style}\}、 \end{matrix} \tag{2.48}$$

从 $n = 0$ 处的无应力初始状态开始。完整轨迹梯度通过对式（2.34）中的轨迹目标函数进行中心有限差分，按分量进行了近似，

$$g_i^{\text{FD}} := \frac{\mathcal{J}(\boldsymbol{\theta} + h_i \mathbf{e}_i) - \mathcal{J}(\boldsymbol{\theta} - h_i \mathbf{e}_i)}{2h_i}, \tag{2.49}$$

其中，向量$\mathbf{e}_i$ 表示与第 i 个可训练参数 θ_i 相关联的单位向量，标量h_i 表示步长

$$h_i := 10^{-6} \max(1, |\theta_i|). \tag{2.50}$$

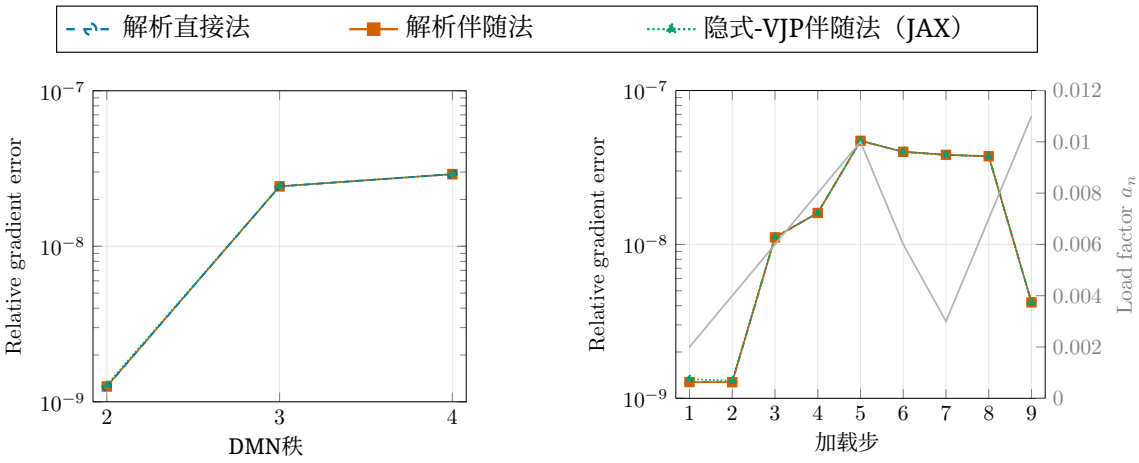
收集各分量 g_i^{FD}即可得到有限差分轨迹梯度$\mathbf{g}^{\text{FD}}$。我们还使用相对扰动 10^{-5} 重复进行了中心有限差分计算，作为对有限差分参考值的稳定性检查，但此处不展示结果。对于每个网络秩，我们考虑相对误差

$$e_{\text{grad}} := \frac{\|\mathbf{g} - \mathbf{g}^{\text{FD}}\|_2}{\|\mathbf{g}^{\text{FD}}\|_2}, \tag{2.51}$$

其中向量

$$\mathbf{g} := \frac{\mathrm{D}\mathcal{J}}{\mathrm{D}\boldsymbol{\theta}} \tag{2.52}$$

表示通过直接公式、解析伴随法或隐式-VJP伴随法获得的完整轨迹梯度。图 2(a) 展示了给定网络秩下完整参数梯度向量的相对欧几里得误差（2.51）。对于秩 2–4，三种实现均与中心有限差分参考值一致，误差约为 10^{-9}–10^{-8}。解析伴随法和隐式-VJP伴随法曲线几乎无法区分，这是预期结果，因为它们通过独立的导数实现实现了相同的离散伴随敏感性关系。它们与直接敏感度和有限差分参考值的一致性，证实了伴随方程（2.37）中对历史依赖性的处理是一致的。图 2(b) 展示了秩 4 时每个加载步 n 的损失贡献 L_n 梯度误差。沿完整轨迹，该误差保持在 10^{-9} 和 5×10^{-8} 之间。在初始加载过程中，误差增大至第 5 步；但在第 6 和第 7 步的卸载以及第 8 和第 9 步的再加载过程中，误差不再增长，并且三种实现均在每个加载步上保持一致。因此，继承的本构状态的敏感度能够从一个增量一致地传播到下一个增量。



（a）完整轨迹梯度
对于 DMN 秩 2–4

（b）秩 4 的每加载步梯度，
灰色表示加载因子 a_n

图 2：相对于中心有限差分参考值的梯度相对欧几里得误差（2.51）。三种实现彼此重合。

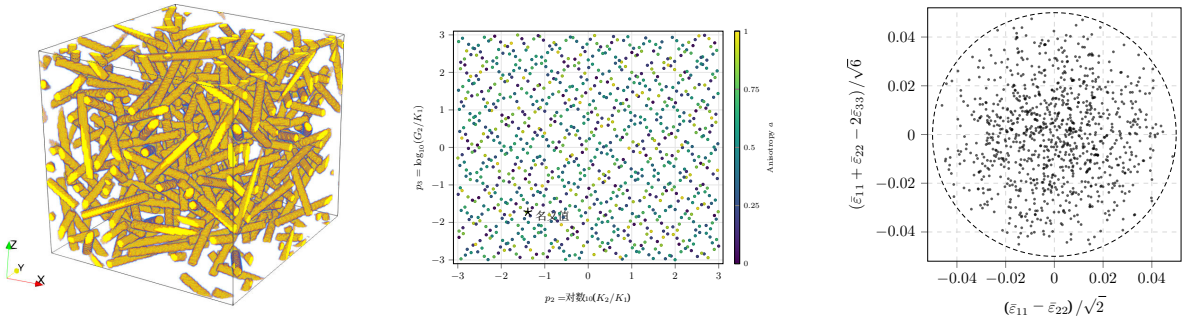
我们将数值比较限定于梯度的准确度。直接方法和伴随方法的计算复杂度将在第 2.3.3 节中讨论。

3 计算示例

3.1 问题设定

该数值研究比较了短纤维复合材料的弹性训练和非弹性训练。下面将说明其微观结构、参考数据、训练配置和误差度量。

微观结构与参考求解器。我们考虑一种两相复合材料，其代表性结构为聚合物基体，并由长度为 $100\,\mu$m、直径为 $10\,\mu$m 的短 E玻璃纤维增强。纤维以各向同性方式分散，并通过相应的纤维取向张量 [105], 编码，纤维体积分数



(a) 体素化短纤维复合材料
含有 128^3 个体素

（b）弹性数据：刚度对比

（c）非弹性数据：投影项终端应变，
$R^{\max} = 0.05$ （虚线）

图3：问题设定：微观结构和采样训练数据。

为 10 %。图 3(a) 所示的微观结构由顺序添加与迁移算法 [106], 生成，边长为 $256\,\mu$m，并以 $2\,\mu$m 的分辨率进行体素化，得到一个 128^3 体素网格。随后，将所得微观结构离散到旋转交错网格 [11] 上，并使用自主开发的基于快速傅里叶变换的均匀化求解器 [107] 求解。对于弹性数据集，采用线性共轭梯度法 [10]，求解；对于非弹性情形，则采用容差为 10^{-5} 的 Eyre–Milton 求解器 [108, 109] 。

弹性数据。 对于弹性辨识，我们采用 1000 对组分刚度张量，其采样遵循 Gajek 等人 [80]的正交各向异性采样。在 Mandel 记号下，两个组分张量为

$$\mathbf{C}_1 = 3K_1\mathbf{P}_{\mathrm{vol}} + 2G_1\mathbf{P}_{\mathrm{dev}}, \qquad \mathbf{C}_2 = 3K_2\mathbf{P}_{\mathrm{vol}} + 2G_2\big(\mathbf{P}_{\mathrm{dev}} - a\,\mathbf{N}' \otimes \mathbf{N}'\big), \tag{3.1}$$

其中，$\mathbf{P}_{\mathrm{vol}}$ 和$\mathbf{P}_{\mathrm{dev}}$ 分别为体积投影算子和偏应力投影算子，$\mathbf{N}'$ 为旋转后的单位范数偏张量。我们设定 $K_1 = 1$GPa 、 $G_1 = 10^{p_1}$ GPa 、 $K_2 = 10^{p_2}$ GPa 和 $G_2 = 10^{p_3}$ GPa。各向异性系数 $a \in [0, 1)$ 、指数 p_1、 p_2、 $p_3 \in [-3, 3)$ 以及四个角参数通过八维扰乱 Sobol 序列进行采样。这里，索引 1 和 2 分别表示纤维和基体。图3(b)展示了（p_2、 p_3）平面中采样得到的刚度参数，并按各向异性参数 a着色。非弹性数据中的各向同性组分以纤维的体积模量归一化后，对应于 $p_1 = -0.162$、 $p_2 = -1.389$、 $p_3 = -1.725$ 和 $a = 0$，见图3(b)中的标记，并且位于采样范围内。固定微观结构所对应的有效刚度张量构成弹性均匀化目标。样本被划分为 800 个训练样本、100 个验证样本和 100 个测试样本。该弹性辨识族不同于用于非弹性参考轨迹的固定各向同性组分性质。由于有效刚度关于组分刚度是一阶正齐次的，因此只有刚度对比起作用，归一化 $K_1 = 1$GPa 不构成限制。

非弹性数据。对于非弹性数据生成，聚合物基体遵循具有线性各向同性硬化的 von Mises 塑性模型，杨氏模量为 2.1 GPa，泊松比为 0.3，屈服应力为 29 MPa，硬化模量为 139 MPa。纤维相为线弹性材料，杨氏模量为 $E_f = 72$GPa，泊松比为 $\nu_f = 0.22$。终端宏观应变由 1000 个对称张量$\boldsymbol{E}_s$ 构成，这些张量的分量相互独立且服从标准正态分布；首先移除其体积部分，

$$\operatorname{dev}\boldsymbol{E}_s := \boldsymbol{E}_s - \frac{1}{3}\operatorname{tr}(\boldsymbol{E}_s)\,\boldsymbol{I}, \tag{3.2}$$

随后将所得的偏应变方向归一化为

$$\bar{\boldsymbol{\varepsilon}}_s^{\max} := R^{\max} \frac{\operatorname{dev} \boldsymbol{E}_s}{\|\operatorname{dev} \boldsymbol{E}_s\|_F}, \qquad R^{\max} = 0.05, \qquad s = 1, \ldots, 1000. \tag{3.3}$$

因此，所有终端应变均满足条件 $\operatorname{tr}(\bar{\boldsymbol{\varepsilon}}_s^{\max}) = 0$ and $\|\bar{\boldsymbol{\varepsilon}}_s^{\max}\|_F = R^{\max}$。对于每个终端应变，基于快速傅里叶变换的求解器都经历了单调递增的比例历史

$$\bar{\boldsymbol{\varepsilon}}_s(\kappa) := \kappa\, \bar{\boldsymbol{\varepsilon}}_s^{\max}, \qquad \kappa \in [0, 1], \tag{3.4}$$

按照加载因子 κ 表示，并离散为 $N_t = 10$ 个时间步，同时给出了相应的有效应力轨迹。1000 次 FFT 计算中有一次无效，因为最终有效应力的范数低于 10^{-6} GPa，因此被舍弃，最终获得了 999 条非弹性轨迹。

为直观展示采样结果，我们将终端偏应变投影到其正交归一偏应变 Mandel 基 [110] 上，并得到坐标

$$a_{1,s} := \frac{\bar{\varepsilon}_{11,s}^{\max} - \bar{\varepsilon}_{22,s}^{\max}}{\sqrt{2}}, \tag{3.5a}$$

$$a_{2,s} := \frac{\bar{\varepsilon}_{11,s}^{\max} + \bar{\varepsilon}_{22,s}^{\max} - 2\,\bar{\varepsilon}_{33,s}^{\max}}{\sqrt{6}}, \tag{3.5b}$$

$$a_{3,s} := \sqrt{2}\,\bar{\varepsilon}_{12,s}^{\max}, \tag{3.5c}$$

$$a_{4,s} := \sqrt{2}\,\bar{\varepsilon}_{13,s}^{\max}, \tag{3.5d}$$

$$a_{5,s} := \sqrt{2}\,\bar{\varepsilon}_{23,s}^{\max}. \tag{3.5e}$$

图 3(c) 展示了样本在（a_1，a_2）平面上的分布。下文中，上标 e 和 i 分别用于区分弹性训练和非弹性训练，而下标表示网络秩。例如，eDMN_5 和 iDMN_5 分别表示相应的秩为 5 的模型。附加在 iDMN 上的上标表示可用非弹性轨迹的数量。

训练配置。 深度材料网络（DMN）采用 JAX [104]实现，并使用 Adam [111]进行优化。学习率设为常数 10^{-3} 。eDMN 使用包含 64 个弹性刚度样本的小批次进行训练；iDMN 则使用包含 8 条完整应力-应变轨迹的小批次，并采用第 2.3.2 节推导的离散伴随梯度。每次 iDMN 优化都以一个样本的弹性预训练作为初始化，该预训练包含 1000 次 Adam 更新，学习率为 10^{-2}。

在主要优化过程中，每进行 25 次参数更新便评估一次验证损失，每进行 100 次参数更新便保存一次检查点。每当验证损失至少改善 10^{-6} 时，就保留当前最佳参数；若连续 20 次验证评估均未达到这一改善幅度，则终止训练，但参数更新次数最多为 10000 次。每种 eDMN 秩进行了五次独立运行；每种 iDMN 秩和训练集规模进行了三次独立运行。弹性数据集划分为 800 个训练样本、100 个验证样本和 100 个测试样本；999 条非弹性轨迹则划分为 899 条训练轨迹、49 条验证轨迹和 51 条测试轨迹。表 1 总结了弹性和非弹性 DMN 的训练配置。需要注意的是，iDMN^{128},iDMN^{256},iDMN^{512} 的数据集固定，并且是从 $\text{iDMN}^{\text{full}}$的数据集中嵌套划分得到的。此外，不同秩、训练集规模和初始化种子所用的验证轨迹与测试轨迹均保持不变。

误差度量。 对于测试样本 s 和独立训练的模型 r，轨迹相对有效应力误差定义为

$$e_s^{(r)} := \frac{\|\bar{\boldsymbol{\sigma}}_s^{\text{pred},(r)} - \bar{\boldsymbol{\sigma}}_s^{\text{FFT}}\|_F}{\|\bar{\boldsymbol{\sigma}}_s^{\text{FFT}}\|_F}. \tag{3.6}$$

Ta表 1：用于弹性 DMN 和非弹性 DMN 训练的数据集与优化配置 .

	eDMN	iDMN^{128}	iDMN^{256}	iDMN^{512}	iDMN^{full}
独立运行	5	3	3	3	3
训练样本	800	128	256	512	899
小批量大小	64	8	8	8	8
学习率	10^{-3}	10^{-3}	10^{-3}	10^{-3}	10^{-3}
验证间隔	25	25	25	25	25
检查点间隔	100	100	100	100	100
提前停止耐心值	20	20	20	20	20
最大更新次数	10000	10000	10000	10000	10000

对于每次训练运行，我们首先对 N_s 测试样本取该量的平均值，

$$\bar{e}^{(r)} := \frac{1}{N_s} \sum_{s=1}^{N_s} e_s^{(r)}. \tag{3.7}$$

由于每种网络配置都进行了多次训练，以下统计数据刻画了 N_r 次独立运行中平均测试误差的变化情况，

$$e_{\max} := \max_{1 \le r \le N_r} \bar{e}^{(r)}, \qquad e_{\text{mean}} := \frac{1}{N_r} \sum_{r=1}^{N_r} \bar{e}^{(r)}, \qquad e_{\text{std}} := \sqrt{\frac{1}{N_r - 1} \sum_{r=1}^{N_r} \left(\bar{e}^{(r)} - e_{\text{mean}} \right)^2}. \tag{3.8}$$

3.2 训练结果

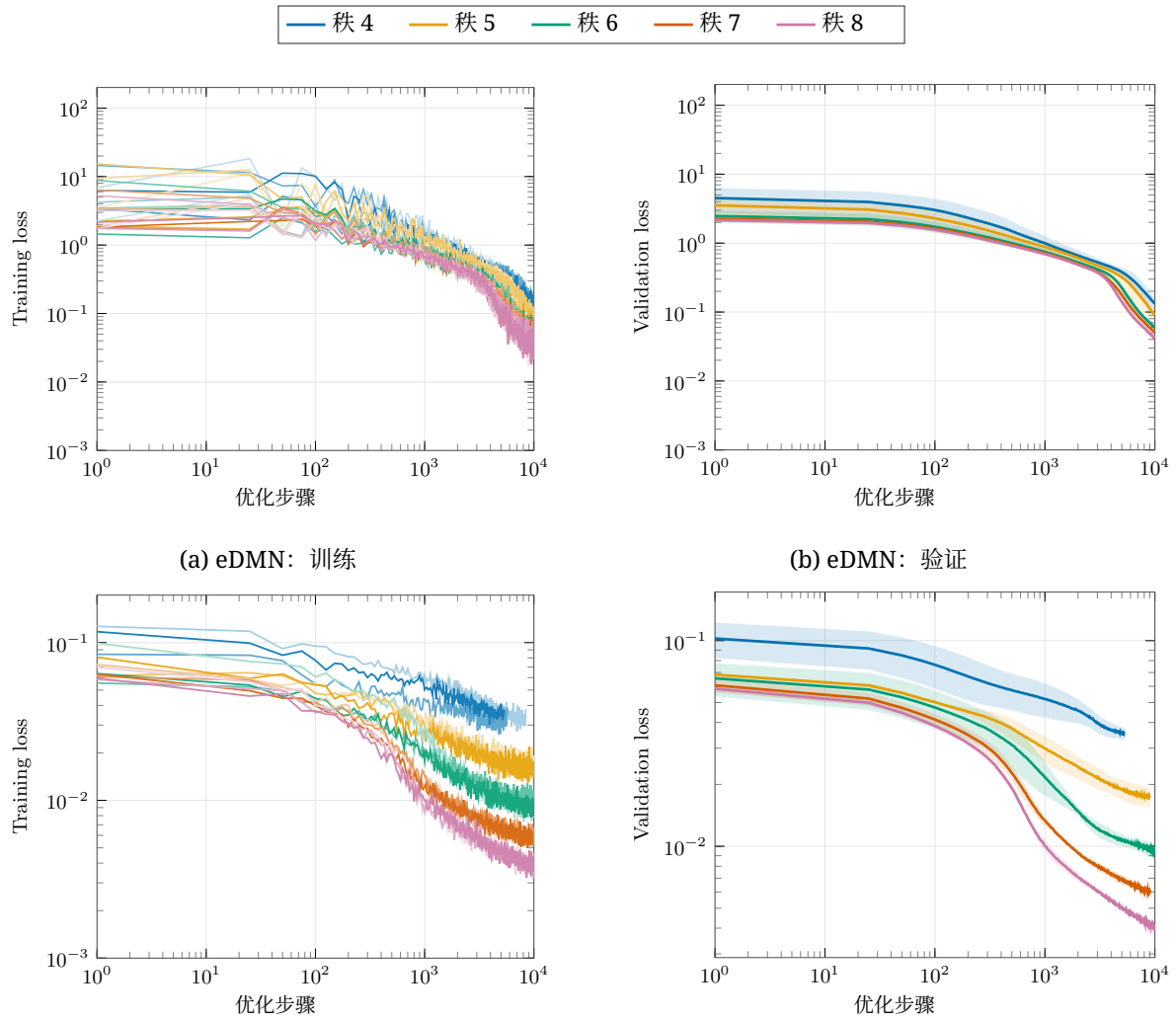
3.2.1 训练

图4展示了 eDMN 和 iDMN^{full} 的训练与验证历史。其他非弹性训练集规模也呈现出相似的行为。由于弹性训练和非弹性训练分别最小化方程（2.23）和（2.25）中的不同目标函数，因此它们的绝对损失值不能直接比较，只应在各自的训练策略内进行解读。在这两种情况下，随着网络秩增加，收敛后的损失通常会降低，这表明增加材料点和内部节点能够改善对微观均匀化器的表征。此处考察的最低秩模型出现较高的损失和较早的平台期，这与网络容量不足的情况相符。

非弹性优化采用上述弹性热启动进行初始化，该方法能够以较低成本提供一个具有力学依据的初始参数集。因此，如图4(c)所示，初始 iDMN损失通常低于 1。故下文将在共同的非弹性测试集上评估弹性训练与非弹性训练各自的优劣，而不比较它们不同的训练目标。

3.2.2 测试

训练后的模型在一个共同的非弹性测试集上进行了评估，该测试集包含未见过的单调递增比例应变路径。对于每个独立训练的模型，将公式 (3.6) 中的轨迹相对误差在测试样本上取平均。图5和表2报告了多次训练运行所得的统计结果。



(a) eDMN：训练

(b) eDMN：验证

(c) iDMN^{full}：训练

(d) iDMN^{full}：验证

图4：eDMN 和 iDMN^{full} 的训练与验证损失历史。颜色表示网络秩。对于训练历史，每种颜色的不同深浅区分了采用不同初始化的独立运行。对于验证历史，实线表示均值，阴影区域表示重复运行的标准差。

对于 eDMN，平均测试误差从秩4时的16.28%降至秩8时的5.12%。在较低秩数下，独立训练运行之间的离散程度较大，秩4和秩5时的标准差分别为4.33%和4.23%。在所研究的每个秩数下，非弹性训练都能得到显著更低的误差。在秩4时，iDMNs 的平均误差已经达到3.60%–3.75%，低于 eDMN 的5.12% $_8$，；而 iDMN$_8$ 达到0.50%–0.61%。因此，iDMN$_4^{128}$仅使用128条非弹性训练轨迹，便已优于 eDMN$_8$ 。

增加非弹性轨迹的数量会带来进一步改进，但存在边际收益递减。在秩为 8 时，iDMN^{128} 的平均误差仅从 0.61% 变为 iDMN^{512} 和 iDMN^{full} 的 0.50%。相应的标准差均低于 0.02%。iDMN^{256}, 、iDMN ^{512}，和 iDMN^{full} 模型在秩为 7–8 时已经几乎无法区分。这些结果区分了模型容量和训练集大小的影响。增加网络秩能够持续提高准确度，而当超过

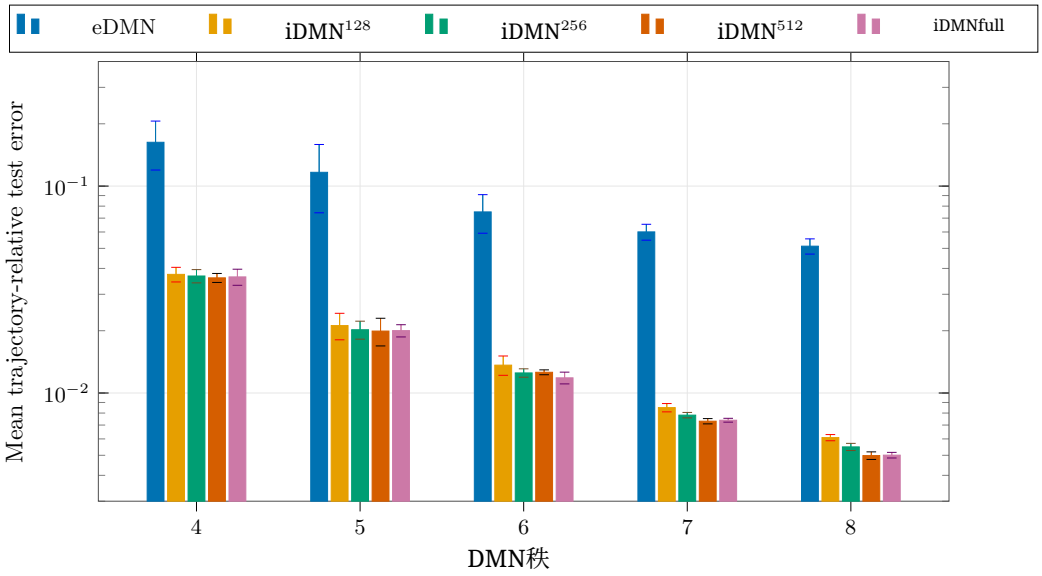


图5：eDMN 和四种 iDMN训练集规模下，非弹性测试集上的平均轨迹相对有效应力误差随网络秩的变化。

表2：非弹性测试集上的有效应力样本误差统计数据 e_s 。所有数值均以百分比表示。

Rank	eDMN			iDMN128			iDMN256			iDMN512			iDMN899		
	最大值	平均值	标准差	最大值	平均值	标准差	最大值	平均值	标准差	最大值	平均值	标准差	最大值	平均值	标准差
4	22.44	16.28	4.33	4.06	3.75	0.30	3.98	3.68	0.27	3.79	3.60	0.18	3.97	3.64	0.33
5	17.12	11.66	4.23	2.42	2.12	0.31	2.21	2.02	0.20	2.28	1.99	0.30	2.12	2.00	0.14
6	9.94	7.50	1.59	1.53	1.36	0.15	1.31	1.25	0.06	1.30	1.26	0.03	1.27	1.18	0.08
7	6.55	6.01	0.53	0.89	0.85	0.04	0.81	0.78	0.02	0.75	0.73	0.02	0.75	0.74	0.02
8	5.82	5.12	0.43	0.63	0.61	0.02	0.56	0.55	0.02	0.52	0.50	0.02	0.52	0.50	0.02

中等数量的非弹性轨迹后，对于当前加载分布，数据集规模的敏感性较弱。因此，对于此处考察的比例加载测试，一旦具备中等规模的非弹性训练集，在当前训练协议下，网络容量似乎就成为主要限制因素。

3.2.3 计算成本

本小节考察计算成本的三个组成部分：弹性和非弹性参考数据的生成、eDMN 和 iDMN 模型的训练，以及相较于全场基于FFT的求解器的在线加速。

数据生成 表 3 报告了用于构建训练目标的全场模拟的离线成本。所有全场计算均在配备 AMD EPYC 9354处理器的工作站上完成。每个弹性样本需要求解六个线性基于FFT的均匀化问题，每个问题在 12 个线程上耗时约 78.4 秒，即每个样本约需 1.57 CPU 小时，1000 个样本共需 1568.34 CPU 小时

。每条非弹性轨迹在六个线程上耗时 173.7 秒 $\pm$ 16.0秒墙钟时间，即约需 0.29 CPU 小时；999 条非弹性轨迹总计需要 289.27 CPU 小时，据此我们推算出生成 128、256 和 512 条轨迹所需的时间。因此，生成非弹性数据所消耗的 CPU 时间不到同等规模弹性数据集的五分之一。考虑到目标不同，解读这一比较时应有所区别：一个弹性样本包含完整的有效刚度张量，因此需要求解若干相互独立的线性均匀化问题；而一个非弹性样本包含的是对一条指定比例加载轨迹的响应。此外，基于 FFT 的求解器在非弹性情形下可利用外推：前序加载步的收敛解可为当前加载步提供准确的初始猜测 [14], ，因此每个加载步只需进行少量迭代。相比之下，弹性样本的六个加载工况彼此独立，无法进行这种外推。因此，较低的成本并不意味着非弹性本构积分本质上比线性均匀化更便宜。对于非弹性训练,

表 3：生成弹性和非弹性训练数据集所需的 CPU 时间。

数据集	样本	时间（CPU 小时）
弹性训练	1000	1568.34
非弹性训练	128	37.06
	256	74.13
	512	148.25
	999	289.27

参考数据仅生成一次，并在每个网络秩下重复使用，因此这部分离线成本取决于所选数据集规模，而不取决于 DMN 架构。

训练 eDMN 和 iDMN eDMN 和 iDMN 的训练时间列于表 4，并绘制于图 6 中。弹性训练的成本仍然很低，且仅弱依赖于秩：平均耗时从秩 4 时的 76.47 秒增加到秩 8 时的 102.83 秒。获取 eDMN 的成本几乎完全来自数据生成。另一方面，非弹性训练的成本要高得多，因为每次更新都要对完整的历史相关本构问题进行积分和微分。其平均耗时从秩 4 时约 439–540 秒增加到秩 8 时的 3.36×10^4–3.59×10^4 秒，对于最深的网络而言约需 9–10 小时。训练在 amplitUDE 超级计算机的一块 NVIDIA H100 PCIe GPU 上完成，而参考数据则在中央处理器上生成。由于这些资源无法直接比较，我们分别以 CPU 小时报告数据生成成本，并以 GPU 墙钟时间报告训练成本。以 iDMN^{128}, 为例，数据生成需要 37.06 CPU 小时，约为弹性数据集所需 1568.34 CPU 小时的四十分之一；而秩 6 时的训练耗时约为 0.8 GPU 小时，相比之下，eDMN 仅需约一分半钟。

在秩固定的情况下，iDMN 的训练时间几乎不受非弹性数据集规模的影响。这一现象是训练协议所导致的，该协议规定了小批量大小、验证间隔和参数更新的最大次数，但并不意味着处理一个训练轮次的时间与数据集规模无关。因此，成本的增加主要与非线性网络问题及其历史依赖伴随求解的规模有关，而不是与可用轨迹的数量成正比。这一结果补充了精度研究：将非弹性数据集规模增加到中等规模以上时，准确度的提升会逐渐减小；而提高秩则会迅速增加离线成本。因此，中等秩区域最为有利，在该区域，非弹性

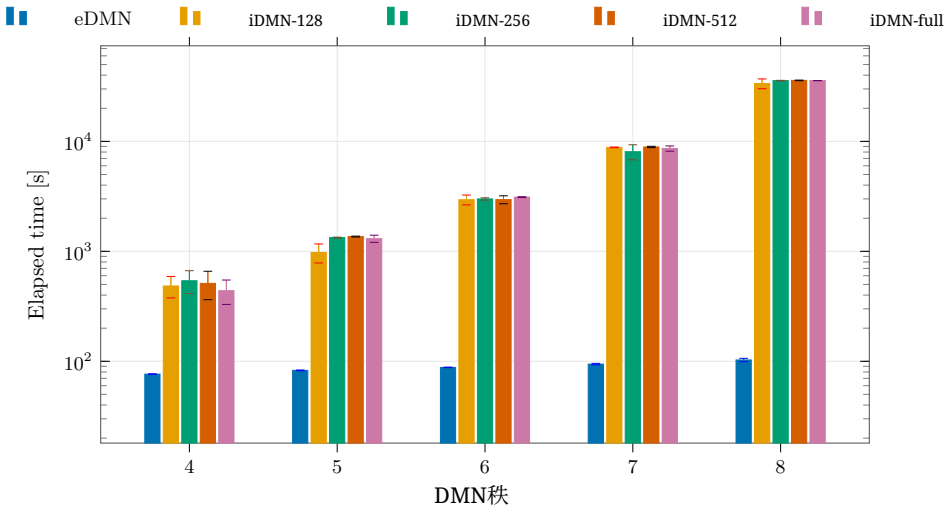


图 6：eDMN 和 iDMN 模型在不同网络秩和数据集大小下的训练时间。误差棒表示重复运行中的变异性。

表 4： t_{elapsed} eDMN 和 iDMN 训练任务的训练耗时统计。所有数值均以秒为单位。

Rank	eDMN			iDMN^{128}			iDMN^{256}			iDMN^{512}			iDMN^{899}		
	max	均值	标准差	max	mean	std	max	mean	std	max	mean	std	max	mean	std
4	77.00	76.47	0.38	576.82	484.18	107.16	635.15	540.15	126.52	673.47	511.24	147.74	564.77	438.78	109.70
5	83.78	82.67	0.62	1136.57	975.95	191.57	1348.38	1332.50	14.11	1372.98	1363.04	12.39	1369.14	1304.77	97.72
6	88.76	87.97	0.45	3140.60	2948.16	301.76	3043.29	2996.84	73.01	3136.23	2956.20	245.93	3132.54	3112.60	28.41
7	95.44	94.25	1.47	8888.53	8819.80	71.04	8817.01	8057.47	1263.36	8993.45	8901.72	123.21	8999.23	8617.79	470.48
8	106.73	102.83	3.63	35603.67	33575.98	3431.28	35926.38	35705.91	213.88	36092.48	35869.18	211.37	35780.80	35708.36	64.78

训练练可在无需付出秩-8 训练的高昂代价的情况下，获得大部分准确度提升 se.

与基于 FFT 的求解器的比较。对于在线评估，训练得到的参数以及相应的运动学算子 D 均保持固定。因此，Fortran 运行时不受这些材料网络参数是通过弹性训练还是非弹性训练获得的影响。每个时间步都会更新材料点本构状态，并使用 PETSc 和牛顿线搜索法 [112]求解非线性全局平衡方程（2.15）。全局残差容差设为 10^{-10}。一致约化切线采用矩阵自由方式应用，所得线性化系统则通过共轭梯度法求解。一个固定的弹性参考切线

$$K_{\text{el}} := \sum_i w_i D_i^T \mathbb{C}_i^{\text{el}} D_i, \tag{3.9}$$

根据弹性刚度 $\mathbb{C}_i^{\text{el}}$ 以及与材料点相关的算子 D 的分块 D_i 组装，并进行乔列斯基分解；随后在整个非线性轨迹中重复用作预条件子。在线预测在台式计算机（Intel Core Ultra 7 265）的一颗中央处理器上进行，使用 Intel ifx 2025.3.2 编译器和 PETSc 3.25.4 版。对于每种模型秩，我们进行 5 次模拟并计算平均耗时。

在线效率根据评估测试集中的一条指定宏观应变轨迹所需的墙钟时间进行评估。基于 FFT 的参考求解器耗时 182.968 s $\pm$ 11.159 s。表 5

报告了使用相同 Fortran 求解器评估的 eDMN 和 iDMN^{128} 模型所获得的加速比。随着网络秩的增加，在线加速幅度显著降低，因为数量

表 5：Fortran DMN 求解器相对于基于 FFT 的全阶求解器的加速比，根据相应的平均实际运行时间计算得出。

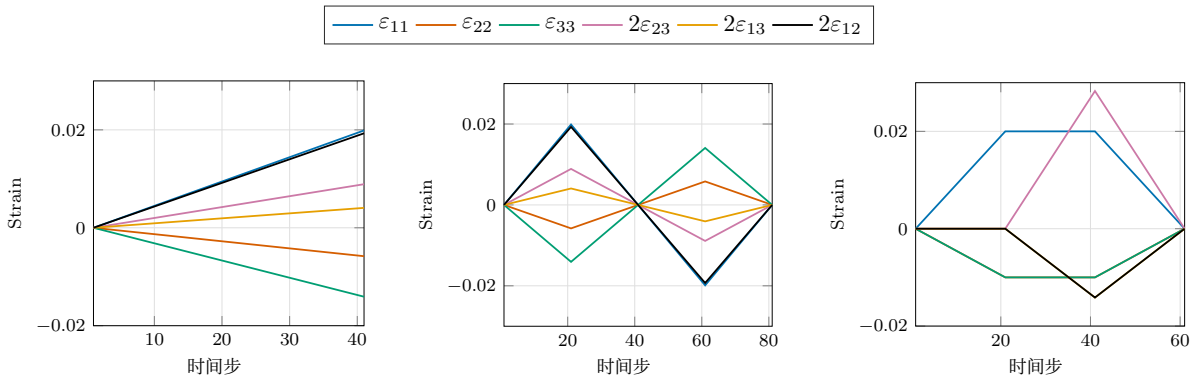
模型	秩 4	秩 5	秩 6	秩 7	秩 8
eDMN	1.43×10^5	5.14×10^4	1.14×10^4	1.74×10^3	2.57×10^2
iDMN^{128}	1.45×10^5	3.78×10^4	1.32×10^4	2.25×10^3	3.36×10^2

材料点的数量和平衡未知量的数量会随着网络深度增加。加速比从秩 4 时约为 1.5×10^5–2.2×10^5，下降到秩 8 时约为 3.3×10^2–3.4×10^2；秩 5–7 则提供了约为 10^3–10^4 的中间加速效果。因此，即使对于此处考虑的最大网络，降阶 DMN 的速度仍比基于 FFT 的全场求解器快两个数量级以上。在秩固定时，eDMN 和 iDMN^{128} 的在线计算成本相近。这是预期之中的，因为两者均采用相同的冻结材料网络公式进行评估，区别仅在于辨识出的网络参数。因此，训练策略主要影响离线辨识成本和预测精度，而非在线求解器的复杂度。

3.3 在线预测

本节评估经过训练的 eDMN 和 iDMN 模型对不同于训练数据的加载历史的预测能力。具体而言，我们使用三种未见加载历史对材料网络进行评估：采用更精细时间离散的单调比例递增路径、比例循环路径以及非比例路径。递增加载路径与训练数据属于同一类路径，可用于测试模型对时间分辨率变化的迁移能力；而循环路径和非比例路径则属于分布外（OOD）数据。这些测试逐步加大了与非弹性训练分布的偏离程度。递增加载情况保留了训练路径的单调和比例特征，同时引入了未见过的加载方向和更精细的时间离散。循环情况还引入了卸载、反向加载和再加载。最后，非比例情况则在整个轨迹中改变加载方向和加载历史。与 iDMN 训练数据集所采用的时间增量相比，这三条路径的离散化均采用了更小的时间增量，其时间步数分别为 $N_t = 40$、80 和 60，如图 7 所示。我们将经过训练的材料网络预测的有效应力轨迹与使用基于 FFT 的微观力学求解器得到的参考解进行比较。附录 B 展示了 eDMN$_8$ 和 iDMN$_8$ 的预测结果及其分量相对误差。主要误差度量为式（3.6）定义的轨迹相对范数，并将其应用于三条加载路径。图 8 和表 6 报告了这一全局轨迹误差随网络秩的变化。

从图 8 可以得出三点观察。首先，在三条未见加载路径上，iDMN$_8$ 的平均误差均低于 5% 的轨迹误差阈值，而 eDMN$_8$ 的平均误差在三条路径上均超过该阈值。非弹性训练的 DMN 将平均加载上升误差从 5.16% 降至 0.48%，将循环误差从 6.68% 降至 2.21%，并将非比例误差从 5.79% 降至 1.77%。从秩 5 起，iDMN^{full}在三条路径上的误差均低于 5% 的阈值，而且 iDMN$_5$ 在每条路径上的精度都高于 eDMN$_8$。此外，iDMN 在独立训练运行中的离散程度远小于 eDMN。由于两个模型的网络秩相同，因此这



（a）比例递增加载　　（b）循环比例加载　　（c）非比例加载

图 7：用于在线外推的应变路径。

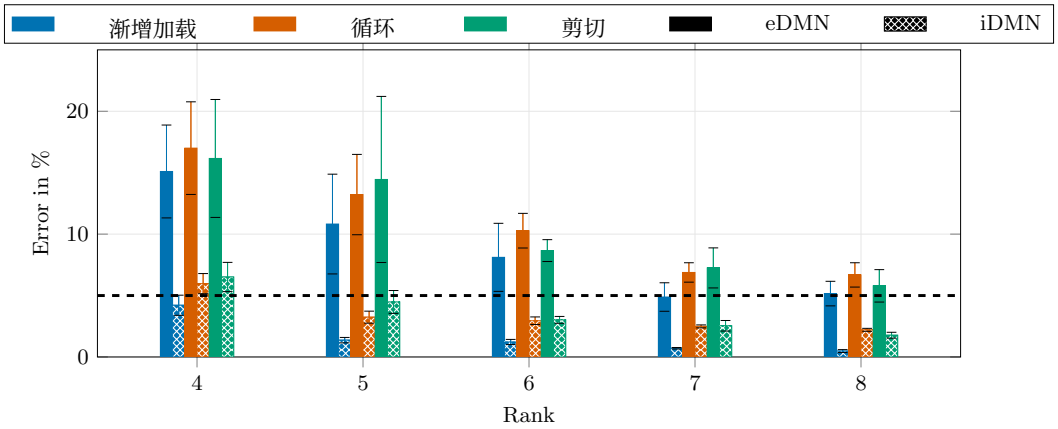


图 8：eDMN 和 iDMN^{full} 在三条未见加载路径上的轨迹相对有效应力误差。虚线表示 5%参考误差阈值。

表 6：Gajek 训练的 eDMN 和 iDMN^{full} 秩扫描在三条分布外加载路径上的全局轨迹相对 L_2 误差。数值为每个秩下五个独立 eDMN 训练随机种子和三个 iDMN 训练随机种子所得结果的平均值 ±样本标准差（%）。

Rank	渐增加载		循环		剪切	
	eDMN	iDMN^{full}	eDMN	iDMN^{full}	eDMN	iDMN^{full}
4	15.10 ± 3.78	4.21 ± 0.82	17.00 ± 3.77	5.97 ± 0.82	16.16 ± 4.80	6.51 ± 1.19
5	10.82 ± 4.06	1.37 ± 0.22	13.22 ± 3.27	3.24 ± 0.49	14.45 ± 6.76	4.48 ± 0.93
6	8.11 ± 2.77	1.22 ± 0.21	10.28 ± 1.41	2.95 ± 0.31	8.66 ± 0.89	3.02 ± 0.28
7	4.88 ± 1.16	0.71 ± 0.06	6.88 ± 0.79	2.49 ± 0.12	7.25 ± 1.63	2.54 ± 0.43
8	5.16 ± 1.00	0.48 ± 0.12	6.68 ± 0.99	2.21 ± 0.12	5.79 ± 1.32	1.77 ± 0.24

改进源于训练策略，而非网络容量的增加。其次，非弹性训练带来的准确度提升，在用于辨识的单调递增比例加载轨迹之外仍得以保留。渐增加载路径的误差最小；该路径保留了训练轨迹的比例特征，但改变了加载方向和时间

离散化。对于引入卸载和反向加载的循环加载，以及应变方向在整个轨迹中不断演化的非比例加载，这一优势仍然十分显著。因此，非弹性识别不仅改善了训练设置内的插值，还实现了向定性不同的加载历史的迁移。第三，与渐增加载情况相比，循环加载和非比例路径的误差更大，这表明经过训练的 DMN 仍然对用于辨识的加载历史分布敏感。当前的非弹性训练数据仅包含单调递增的比例加载轨迹。因此，提高对更一般加载历史的准确度，可能需要增加训练轨迹的多样性，而不仅仅是增加同一比例加载族的样本数量。

3.4 向受扰动组分性质的迁移

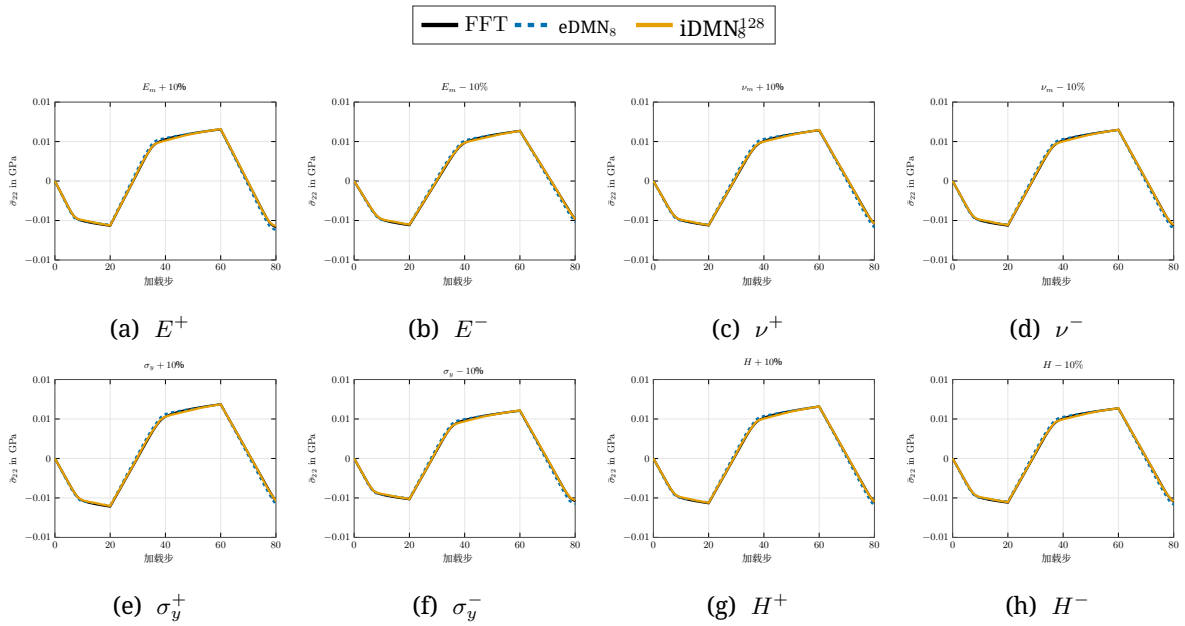
上述加载路径测试在保留用于非弹性识别的组分性质的同时，改变了宏观历史。接下来，我们考察 iDMN 在不同材料属性下的性能。保留了循环加载路径，因为它已涵盖卸载、反向加载和再加载，同时固定训练得到的网络参数。从名义基体参数出发，即杨氏模量为 $E_m = 2.1$GPa、泊松比为 $\nu_m = 0.30$、屈服应力为$\sigma_y = 29$MPa、硬化模量为 $H = 139\,$MPa，我们每次将一个参数扰动$\pm 10\%$。纤维仍为线弹性材料，其杨氏模量为 $E_f = 72$GPa，泊松比为$\nu_f = 0.22$。加上名义值情形，此过程共得到九种材料情景。我们再次使用基于 FFT 的求解器获得参考轨迹，并据此计算式 (3.6) 中的轨迹相对范数。

本实验还比较了弹性识别和非弹性识别预算。eDMN根据涵盖一系列组分性质的800个弹性刚度样本进行辨识，而iDMN^{128} 仅使用名义值材料参数下的128条非弹性轨迹。名义值基体以及全部八种扰动的刚度对比均处于弹性数据的采样范围内，参见第3.1节。因此，eDMN和iDMN^{128} 均应用于其训练数据附近。表7比较了所有秩和材料情景下的eDMN与iDMN^{128} 。完整的25模型扫描结果，包括iDMN^{256}, 、iDMN^{512}, 和iDMN^{full}，见附录C。

表7：循环加载下扰动矩阵性质的轨迹相对有效应力误差。数值为独立训练随机种子下的均值 $\pm$ 样本标准差（%）：eDMN为 $n = 5$ ，每种iDMN配置为 $n = 3$ 。

模型	Rank	Nom.	E_m^-	E_m^+	ν_m^-	ν_m^+	σ_y^-	σ_y^+	H^-	H^+
eDMN	4	17.00 ± 3.77	16.54 ± 3.29	17.54 ± 4.25	15.85 ± 2.97	18.50 ± 4.86	17.76 ± 4.14	16.39 ± 3.45	17.17 ± 3.94	16.86 ± 3.61
eDMN	5	13.22 ± 3.27	13.14 ± 2.99	13.35 ± 3.55	12.56 ± 2.71	14.07 ± 4.04	13.67 ± 3.51	12.87 ± 3.07	13.33 ± 3.30	13.11 ± 3.25
eDMN	6	10.28 ± 1.41	10.46 ± 1.37	10.16 ± 1.45	10.18 ± 1.45	10.39 ± 1.38	10.50 ± 1.36	10.13 ± 1.43	10.24 ± 1.44	10.32 ± 1.39
eDMN	7	6.88 ± 0.79	7.05 ± 0.73	6.79 ± 0.86	6.81 ± 0.71	6.99 ± 0.89	7.05 ± 0.81	6.78 ± 0.76	6.95 ± 0.83	6.83 ± 0.76
eDMN	8	6.68 ± 0.99	6.82 ± 0.95	6.61 ± 1.04	6.65 ± 1.00	6.71 ± 1.00	6.83 ± 1.12	6.59 ± 0.89	6.73 ± 0.97	6.63 ± 1.01
iDMN^{128}	4	5.99 ± 0.69	6.02 ± 0.49	5.90 ± 0.77	6.00 ± 0.65	6.04 ± 0.77	6.18 ± 0.85	5.77 ± 0.46	6.11 ± 0.67	5.88 ± 0.70
iDMN^{128}	5	3.45 ± 0.18	3.35 ± 0.14	3.60 ± 0.21	3.53 ± 0.21	3.40 ± 0.16	3.73 ± 0.21	3.25 ± 0.17	3.42 ± 0.17	3.48 ± 0.19
iDMN^{128}	6	3.10 ± 0.11	2.98 ± 0.09	3.24 ± 0.13	3.15 ± 0.07	3.06 ± 0.18	3.31 ± 0.09	2.95 ± 0.13	3.05 ± 0.14	3.16 ± 0.08
iDMN^{128}	7	2.39 ± 0.14	2.30 ± 0.14	2.49 ± 0.13	2.42 ± 0.09	2.36 ± 0.19	2.58 ± 0.13	2.24 ± 0.14	2.33 ± 0.10	2.44 ± 0.17
iDMN^{128}	8	2.27 ± 0.10	2.21 ± 0.09	2.34 ± 0.10	2.30 ± 0.12	2.24 ± 0.09	2.44 ± 0.10	2.14 ± 0.09	2.20 ± 0.10	2.34 ± 0.10

对于 iDMN^{128}, ，在表7中的45组同秩比较中，非弹性辨识模型的误差均低于 eDMN。在秩为8时，其名义值误差、平均扰动误差和最坏扰动误差分别为2.27%、2.28%和2.44%；相比之下，eDMN的对应数值为6.68%、6.70%和6.83%，其中



(a) E^+ (b) E^- (c) ν^+ (d) ν^-

(e) σ_y^+ (f) σ_y^- (g) H^+ (h) H^-

图9：基体材料参数扰动时预测的有效应力 $\bar{\sigma}_{22}$ ，每个模型展示一次训练运行的结果。

所有数值均为独立训练运行的平均值。对于这项循环迁移测试，即使是 iDMN$_4^{128}$，在所有九种材料工况下的准确度仍高于 eDMN$_8$ 。对于其他非弹性训练集规模，这些观察结果同样成立。

为进一步观察，图9展示了八组扰动材料属性下有效应力分量 $\bar{\sigma}_{22}$ 的轨迹。随着材料参数变化，iDMN^{128} 的轨迹始终接近FFT参考值，而eDMN则表现出更大的偏差，如表7中的误差所示。

因此，该实验表明，非弹性识别所获得的准确度并不局限于训练过程中使用的组分参数。在所研究的 $\pm$10%邻域内，即使弹性模量、泊松比、屈服应力和各向同性硬化模量发生变化，这种准确度仍得以保持。这是当前 J_2 本构律附近的局部材料迁移结果，并不能证明其可迁移到任意组分模型或参数范围。

4 结论

在这项工作中，我们直接根据非弹性应力-应变轨迹辨识了深度材料网络。为此，我们将训练表述为一个历史相关约束优化问题，并推导出路径一致离散伴随。该方法在每个时间步通过对线性化状态方程进行一次反向求解，即可获得轨迹损失的完整敏感度，且与可训练参数的数量无关。对于具有弹塑性基体的短纤维复合材料，经过非弹性训练的网络比经过弹性训练的对应网络更加准确，在独立训练运行中的离散程度更小，并且生成训练数据所需的CPU时间更少。下面，我们将讨论这些发现能够揭示的内容。

弹性训练是 DMN 的标准识别策略，这有充分的理由。它依赖高效的线性求解器，具有稳健性，而且每个经过训练的网络都从其组分继承了热力学一致性。然而，对于非弹性应用而言，其理论基础比实际成功所暗示的更为有限。Gajek 等人 [78] 表明，具有消失记忆的材料的有效响应在一阶近似下由线弹性均匀化函数决定。

因此，从弹性训练到非弹性行为的迁移，仅在渐近意义下，并且仅针对这一类材料得以确立。此外，即使在线弹性领域，DMN 的表达能力也仍未明确，因为尚不清楚由组分的任意微观结构可实现的有效刚度集合G 是否与由层级层合体可实现的有效刚度集合G_L 相一致 [79,113]。尽管非弹性训练无法解决这些问题，但它能够直接评估 DMN 对非弹性组分的近似质量。

在实践中，通常不清楚应如何训练 DMN 进行弹性预测，以及应使用哪些数据进行训练。Dey 等人 [92] 观察到，具有相近弹性误差的网络可能表现出显著不同的蠕变误差，而 Gajek 等人 [93] 通过采用针对组分 Maxwell 结构设计的采样方法，改进了弹性到黏弹性的迁移。我们的结果指向同一结论，并表明，对于当前问题和弹性采样而言，弹性目标函数对非弹性响应的约束并不充分。在秩 8 时，eDMN 的测试误差为 5.12%，比 iDMN 约 0.5%的误差大一个数量级。类似地，在未见加载路径上，eDMN 误差在独立训练运行之间的离散程度最高比 iDMN 大一个数量级。因此，只有当训练数据能够向网络提供有关依赖历史的应变分配信息时，增加网络容量才会有所助益。非弹性训练可以直接提供这一信息，而无需采用针对特定材料类别设计的采样策略。

每个网络层级精度的提升具有直接的实际意义，因为 DMN 的在线计算成本取决于材料点计算的数量 $M = 2^K$，而该数量会随着每增加一层而翻倍。因此，保持网络浅层结构对于高效的双尺度模拟至关重要。在我们的研究中，秩为 5 的 iDMN$_8$ 在测试集和所有考察的未见加载路径上都比 eDMN 更准确。同时，iDMN$_5^{128}$相较于全场求解器的在线加速比为 3.78×10^4，约为 eDMN$_8$ 加速比的 150 倍。

弹性训练的离线成本很容易被低估，因为数据生成的成本远高于辨识成本。每个弹性样本都需要求解六个相互独立的线性均匀化问题，而一条非弹性轨迹只需一次运行即可计算完成，并且可通过加载步之间的外推获益。对于此处研究的复合材料，非弹性数据集所需的 CPU 时间不到弹性数据集的五分之一；仅需 128 条轨迹，就足以达到 eDMN 在任何秩下都未能达到的准确度。具体而言，iDMN$_5^{128}$的数据生成所需 CPU 小时约为弹性数据集的四十分之一，而其训练耗时约为四分之一个 GPU 小时。从这一点来看，非弹性训练使 DMN 更接近经典模型降阶，因为无论如何都需要非弹性数据。对于基于投影的降阶模型，非弹性快照的适当正交分解构成训练 [31], ；而经验求积法 [33,34] 和经验修正簇求积法 [35–37] 等超降阶技术也同样基于非弹性数据来选择积分点或簇。与这些方法不同，DMN 不会用线性子空间来逼近解流形。因此，DMN 不受将解流形投影到固定线性子空间所带来的逼近障碍的限制，该障碍由 Kolmogorov n-宽度量化。

非弹性识别网络也能推广到训练设定之外。尽管训练数据仅包含一组材料参数下单调递增的比例路径，但在循环和非比例加载条件下，以及弹性和塑性基体参数发生 $\pm 10\%$ 扰动时，iDMN 仍比 eDMN 更准确。循环和非比例加载下相对较大的误差表明，未来的训练集应涵盖更多样的加载路径，而不是增加同一类路径的样本。

这些优势的代价是实现复杂得多。训练本身

只需要宏观应力轨迹，且内部变量不可观测。然而，要获得路径一致灵敏度，就需要将继承的本构状态的依赖关系从一个增量传播到下一个增量。为此，除了提供一致切线外，材料例程还必须提供本构更新相对于前一时间步状态的导数，这超出了标准用户材料例程的接口范围。此外，还必须沿存储的正向轨迹反向求解耦合的全局-局部伴随系统；由于本构响应会沿加载路径在弹性分支和塑性分支之间切换，因此所得梯度需要经过仔细验证。这些困难或许是 Nguyen 和 Noels [90] 针对单调加载路径采用步局部简化方法的原因之一。

在实际应用中，我们建议将弹性信息用作高效的初始化，并将高保真数据用于代理模型部署时所处的响应状态范围。当前工作的自然延伸包括使用循环和非比例加载路径的训练数据，研究其他类别的非弹性组分（如黏塑性材料或损伤材料），以及将经过非弹性训练的 DMN 部署于并发双尺度模拟中。

致谢

作者衷心感谢杜伊斯堡-埃森大学计算科学与模拟中心（CCSS）提供的计算时间。相关计算在信息与媒体服务中心（ZIM）的 amplitUDE 超级计算机上完成（德国研究基金会项目459398823；资助编号 INST 20876/423-1FUGG）。

资助

本研究由欧盟通过欧洲研究委员会在地平线欧洲计划框架下资助（ERC BeyondRVE启动资助，资助协议编号101040238）。然而，文中表达的观点和意见仅代表作者本人，不一定反映欧盟或欧洲研究委员会执行机构的立场。欧盟及资助机构均不对相关观点和意见承担责任。

CRediT 作者贡献声明

Binh Huy Nguyen： 方法论、软件、验证、形式分析、调查、数据整理、可视化、撰写——初稿、撰写——审阅与编辑。**Matti Schneider：** 概念化、方法论、形式分析、撰写——初稿、撰写——审阅与编辑、监督、经费获取。

利益冲突声明

作者声明，他们不存在任何可能被认为会影响本文所报告工作的已知经济利益冲突或个人关系。

数据可用性

数据可应要求提供。

写作过程中使用生成式人工智能和人工智能辅助技术的声明

在准备本项工作期间，作者使用了 OpenAI 的 ChatGPT 和 Anthropic 的 Claude 进行语言修订、文献检索和书目核查，使用 Anthropic 的 Claude Code 检查符号记法的一致性并准备图表，并使用 OpenAI 的 Codex 协助进行代码格式化。在使用这些工具后，作者根据需要对内容进行了审阅和编辑，独立核验了所有数学论证、参考文献和数值结果，并对本出版物的内容承担全部责任。

A 分析雅可比块

为便于理解，本节给出雅可比块（2.38）的闭式表达式。更多详情请参阅 Michaleris 等人 [95] 。对于本文研究的两相复合材料，我们分别计算弹性纤维和弹塑性基体的雅可比块。具体而言，对于属于弹性纤维相的每个材料点，状态向量 $\boldsymbol{v}_n^{\text{fiber}}$仅存储应力

$$\boldsymbol{v}_n^{\text{fiber}} = \boldsymbol{\sigma}_n \in \mathbb{R}^6, \tag{A.1}$$

而对于由 J_2 塑性控制的材料点，向量 $\boldsymbol{v}_n^{\text{matrix}}$由塑性应变 $\boldsymbol{\varepsilon}_n^p$、累积等效塑性应变 p_n、柯西应力 $\boldsymbol{\sigma}_n$ 以及当前步 n的累积塑性乘子λ_n 组成，

$$\boldsymbol{v}_n^{\text{matrix}} = \begin{bmatrix} \boldsymbol{\varepsilon}_n^p & p_n & \boldsymbol{\sigma}_n & \lambda_n \end{bmatrix}^T \in \mathbb{R}^{14}. \tag{A.2}$$

在每个增量步中，状态向量分别更新，随后拼接为全局向量

$$\boldsymbol{v}_n \in \mathbb{R}^{M(14+6)/2} \tag{A.3}$$

对于包含 M 个材料点的 DMN。此区别也会体现在局部残差及其偏导数中。我们用 $\mathbf{I}_{r\times s}$ 和 $\mathbf{0}_{r\times s}$ 分别表示维度为 $r \times s$的单位矩阵和零矩阵。弹性纤维相中的每个材料点对应的局部残差为

$$\boldsymbol{H}_n^{\text{fiber}} = \boldsymbol{\sigma}_n - \boldsymbol{\sigma}_{n-1} - \mathbb{C}_{\text{fiber}}(\boldsymbol{\varepsilon}_n - \boldsymbol{\varepsilon}_{n-1}), \tag{A.4}$$

使相应的雅可比块可表示为

$$\boldsymbol{H}_n^{\text{fiber},\boldsymbol{v}} = \mathbf{I}_{6\times 6}, \tag{A.5a}$$
$$\boldsymbol{H}_n^{\text{fiber},\boldsymbol{\varepsilon}} = -\mathbb{C}_{\text{fiber}}, \tag{A.5b}$$
$$\boldsymbol{H}_n^{\text{fiber},\boldsymbol{v},-} = -\mathbf{I}_{6\times 6}, \tag{A.5c}$$
$$\boldsymbol{H}_n^{\text{fiber},\boldsymbol{\varepsilon},-} = \mathbb{C}_{\text{fiber}}. \tag{A.5d}$$

弹塑性基体的材料点可能发生弹性变形或塑性变形，局部残差由下式给出

$$\boldsymbol{H}_n^{\text{matrix,elastic}} = \begin{bmatrix} \boldsymbol{\varepsilon}_n^p - \boldsymbol{\varepsilon}_{n-1}^p \\ p_n - p_{n-1} \\ \boldsymbol{\sigma}_n - \boldsymbol{\sigma}_{n-1} - \mathbb{C}_{\text{matrix}}\big(\boldsymbol{\varepsilon}_n - \boldsymbol{\varepsilon}_{n-1} - (\boldsymbol{\varepsilon}_n^p - \boldsymbol{\varepsilon}_{n-1}^p)\big) \\ \lambda_n - \lambda_{n-1} \end{bmatrix}, \tag{A.6}$$

and

$$\boldsymbol{H}_n^{\text{matrix,plastic}} = \begin{bmatrix} \boldsymbol{\varepsilon}_n^p - \boldsymbol{\varepsilon}_{n-1}^p - (\lambda_n - \lambda_{n-1})\mathbf{N}_n \\ p_n - p_{n-1} - (\lambda_n - \lambda_{n-1})\sqrt{\frac{2}{3}} \\ \boldsymbol{\sigma}_n - \boldsymbol{\sigma}_{n-1} - \mathbb{C}_{\text{matrix}}\big(\boldsymbol{\varepsilon}_n - \boldsymbol{\varepsilon}_{n-1} - (\boldsymbol{\varepsilon}_n^p - \boldsymbol{\varepsilon}_{n-1}^p)\big) \\ \|\boldsymbol{\sigma}_n^{\text{dev}}\| - \sqrt{\frac{2}{3}}(\sigma_y + Hp_n) \end{bmatrix}, \tag{A.7}$$

resp分别，其中屈服面的法向量 $\mathbf{N}$ 由偏应力 $\boldsymbol{\sigma}^{\text{dev}}$ vi 计算得到 a

$$\mathbf{N} = \frac{\boldsymbol{\sigma}^{\text{dev}}}{\|\boldsymbol{\sigma}^{\text{dev}}\|}. \tag{A.8}$$

对于基体，弹性分支上的雅可比块为

$$\boldsymbol{H}_n^{\text{matrix,elastic},\boldsymbol{v}} = \begin{bmatrix} \mathbf{I}_{6\times6} & \mathbf{0}_{6\times1} & \mathbf{0}_{6\times6} & \mathbf{0}_{6\times1} \\ \mathbf{0}_{1\times6} & 1 & \mathbf{0}_{1\times6} & 0 \\ \mathbb{C}_{\text{matrix}} & \mathbf{0}_{6\times1} & \mathbf{I}_{6\times6} & \mathbf{0}_{6\times1} \\ \mathbf{0}_{1\times6} & 0 & \mathbf{0}_{1\times6} & 1 \end{bmatrix}, \tag{A.9a}$$

$$\boldsymbol{H}_n^{\text{matrix,elastic},\boldsymbol{\varepsilon}} = \begin{bmatrix} \mathbf{0}_{6\times6} \\ \mathbf{0}_{1\times6} \\ -\mathbb{C}_{\text{matrix}} \\ \mathbf{0}_{1\times6} \end{bmatrix}, \tag{A.9b}$$

$$\boldsymbol{H}_n^{\text{matrix,elastic},\boldsymbol{v},-} = \begin{bmatrix} -\mathbf{I}_{6\times6} & \mathbf{0}_{6\times1} & \mathbf{0}_{6\times6} & \mathbf{0}_{6\times1} \\ \mathbf{0}_{1\times6} & -1 & \mathbf{0}_{1\times6} & 0 \\ -\mathbb{C}_{\text{matrix}} & \mathbf{0}_{6\times1} & -\mathbf{I}_{6\times6} & \mathbf{0}_{6\times1} \\ \mathbf{0}_{1\times6} & 0 & \mathbf{0}_{1\times6} & -1 \end{bmatrix}, \tag{A.9c}$$

$$\boldsymbol{H}_n^{\text{matrix,elastic},\boldsymbol{\varepsilon},-} = \begin{bmatrix} \mathbf{0}_{6\times6} \\ \mathbf{0}_{1\times6} \\ \mathbb{C}_{\text{matrix}} \\ \mathbf{0}_{1\times6} \end{bmatrix}. \tag{A.9d}$$

同时，在塑性分支上，雅可比块表示为

$$\boldsymbol{H}_n^{\text{matrix,plastic},\boldsymbol{v}} = \begin{bmatrix} \mathbf{I}_{6\times6} & \mathbf{0}_{6\times1} & -(\lambda_n - \lambda_{n-1})\frac{\mathbf{P}_{\text{dev}} - \mathbf{N}_n \otimes \mathbf{N}_n}{\|\boldsymbol{\sigma}_n^{\text{dev}}\|} & -\mathbf{N}_n \\ \mathbf{0}_{1\times6} & 1 & \mathbf{0}_{1\times6} & -\sqrt{\frac{2}{3}} \\ \mathbb{C}_{\text{matrix}} & \mathbf{0}_{6\times1} & \mathbf{I}_{6\times6} & \mathbf{0}_{6\times1} \\ \mathbf{0}_{1\times6} & -\sqrt{\frac{2}{3}}H & \mathbf{N}_n^T & 0 \end{bmatrix}, \tag{A.10a}$$

$$\boldsymbol{H}_n^{\text{matrix,plastic},\boldsymbol{\varepsilon}} = \begin{bmatrix} \mathbf{0}_{6\times6} \\ \mathbf{0}_{1\times6} \\ -\mathbb{C}_{\text{matrix}} \\ \mathbf{0}_{1\times6} \end{bmatrix}, \tag{A.10b}$$

$$\boldsymbol{H}_n^{\text{matrix,plastic},\boldsymbol{v},-} = \begin{bmatrix} -\mathbf{I}_{6\times6} & \mathbf{0}_{6\times1} & \mathbf{0}_{6\times6} & \mathbf{N}_n \\ \mathbf{0}_{1\times6} & -1 & \mathbf{0}_{1\times6} & \sqrt{\frac{2}{3}} \\ -\mathbb{C}_{\text{matrix}} & \mathbf{0}_{6\times1} & -\mathbf{I}_{6\times6} & \mathbf{0}_{6\times1} \\ \mathbf{0}_{1\times6} & 0 & \mathbf{0}_{1\times6} & 0 \end{bmatrix}, \tag{A.10c}$$

$$\boldsymbol{H}_n^{\text{matrix,plastic},\boldsymbol{\varepsilon},-} = \begin{bmatrix} \mathbf{0}_{6\times6} \\ \mathbf{0}_{1\times6} \\ \mathbb{C}_{\text{matrix}} \\ \mathbf{0}_{1\times6} \end{bmatrix}, \tag{A.10d}$$

其中，$\mathbf{P}_{\mathrm{dev}}$ 表示偏应力投影算子的矩阵表示。分支由标准返回映射算法 [114]选定。第 2.3 节所述的直接微分和伴随微分都会计算活动分支的导数；这些导数在弹性阶段或塑性阶段内均是光滑的。我们未遇到恰好位于弹塑性边界上的评估点，因为在该处导数未定义。

B 未见加载路径的八阶有效应力轨迹

以下各图将 eDMN$_8$ 和 iDMN$_8$ 的预测与 FFT参考值进行比较，涵盖全部六个独立有效应力分量。这些分量图补充了表 6 中报告的全局轨迹误差。对于应力分量 α 在时间步 n处，误差定义为

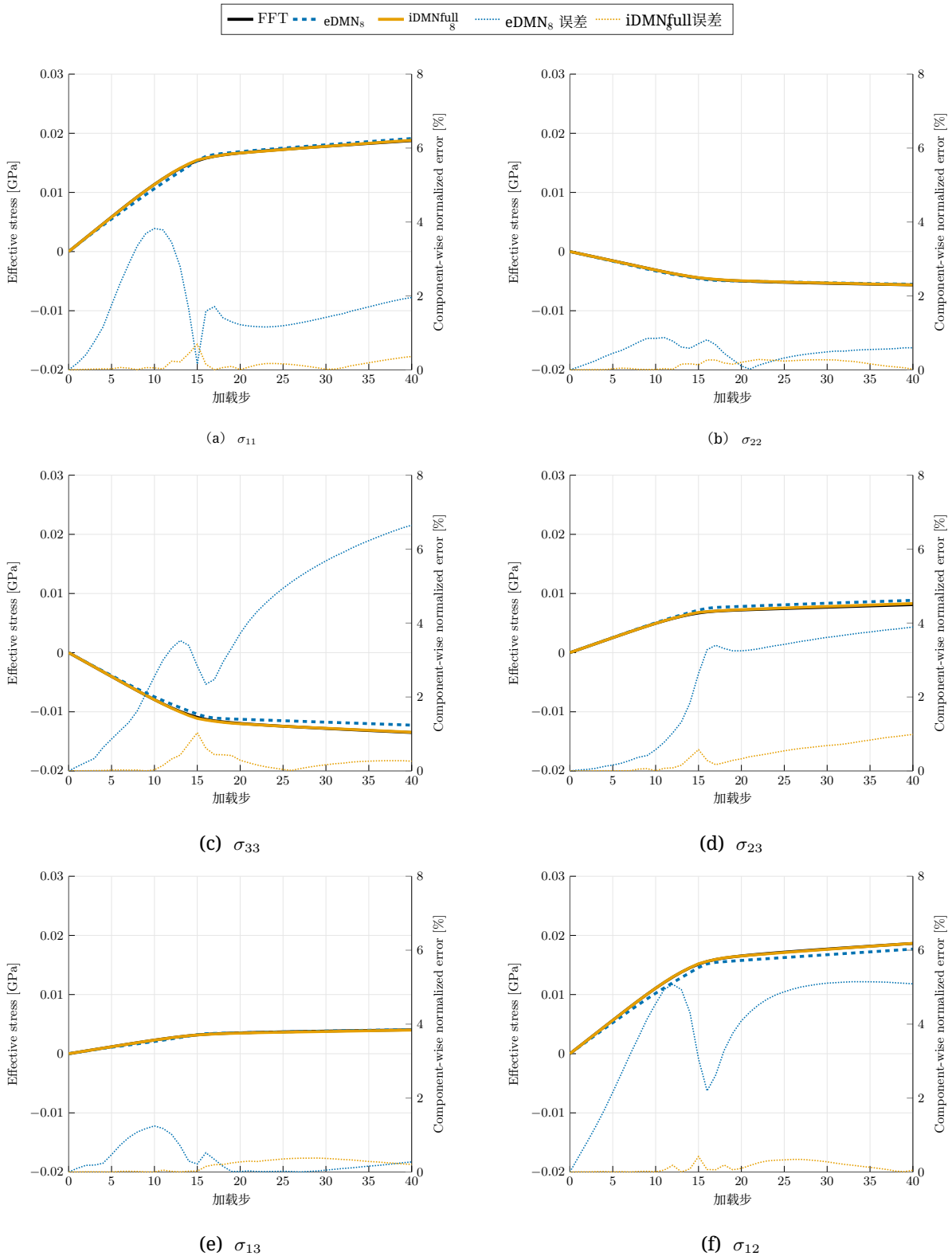
$$e_{\alpha,n} = \frac{\left|\bar{\sigma}^{\mathrm{DMN}}_{\alpha,n} - \bar{\sigma}^{\mathrm{FFT}}_{\alpha,n}\right|}{\max_{\substack{m=1,\dots,N_t \\ \beta=1,\dots,6}} \left|\bar{\sigma}^{\mathrm{FFT}}_{\beta,m}\right| + 10^{-12}}. \tag{B.1}$$

C 材料迁移轨迹误差

表 8 报告了第 3.4 节中使用的全部 225 个全局轨迹相对误差。

参考文献

[1] T. I. Zohdi 和 P. Wriggers，计算微观力学导论，应用与计算力学讲义第20卷。施普林格柏林海德堡，2005年。[2] V. V. Jikov、S. M. Kozlov 和 O. A. Oleinik，微分算子与积分泛函的均匀化。柏林，海德堡：施普林格，1994年。[3] D. Cioranescu 和 P. Donato，均匀化导论。牛津大学出版社，1999年。[4] G. A. Pavliotis 和 A. M. Stuart，多尺度方法：平均化与均匀化，应用数学丛书第53卷。纽约：施普林格，2008年。[5] J. R. Rice，“金属中时间相关塑性变形的应力-应变关系结构”，《应用力学学报》，第37卷，第3期，第728–737页，1970年。[6] J. M. Guedes 和 N. Kikuchi，“基于均匀化方法并结合自适应有限元方法的材料预处理与后处理”，《计算力学与工程方法》，第83卷，第2期，第143–198页，1990年。[7] M. G. Geers、V. G. Kouznetsova 和 W. A. M. Brekelmans，“多尺度计算均质化：趋势与挑战”，《计算与应用数学杂志》，第234卷，第7期，第2175–2182页，2010年。[8] H. Moulinec 和 P. Suquet，“计算复合材料线性与非线性力学性能的快速数值方法”，《法国科学院院刊。第二辑，力学、物理、化学、天文学》，第318卷，第11期，第1417–1423页，1994年。[9] H. Moulinec 和 P. Suquet，“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”，《计算力学与工程方法》，第157卷，第1-2期，第69–94页，1998年。[10] J. Zeman、J. Vondřejc、J. Novák 和 I. Marek，“用共轭梯度法加速用于周期介质数值均匀化的基于 FFT 的求解器”，《计算物理学杂志》，第229卷，第21期，第8065–8071页，2010年。[11] F. Willot，“用于计算具有精确局部场的复合材料力学响应的基于傅里叶的格式”，《法国科学院力学辑》，第343卷，第3期，第232–245页，2015年。[12] M. Kabel、T. Böhlke 和 M. Schneider，“用于大变形下弹性基于 FFT 的均匀化的高效不动点与 Newton–Krylov求解器”，《计算力学》，第54卷，第6期，第1497–1514页，2014年。[13] M. Schneider，“用于弹性和非弹性晶胞均质化问题的基于 FFT 的快速梯度法”，《计算力学与工程方法》，第315卷，第846–866页，2017年。

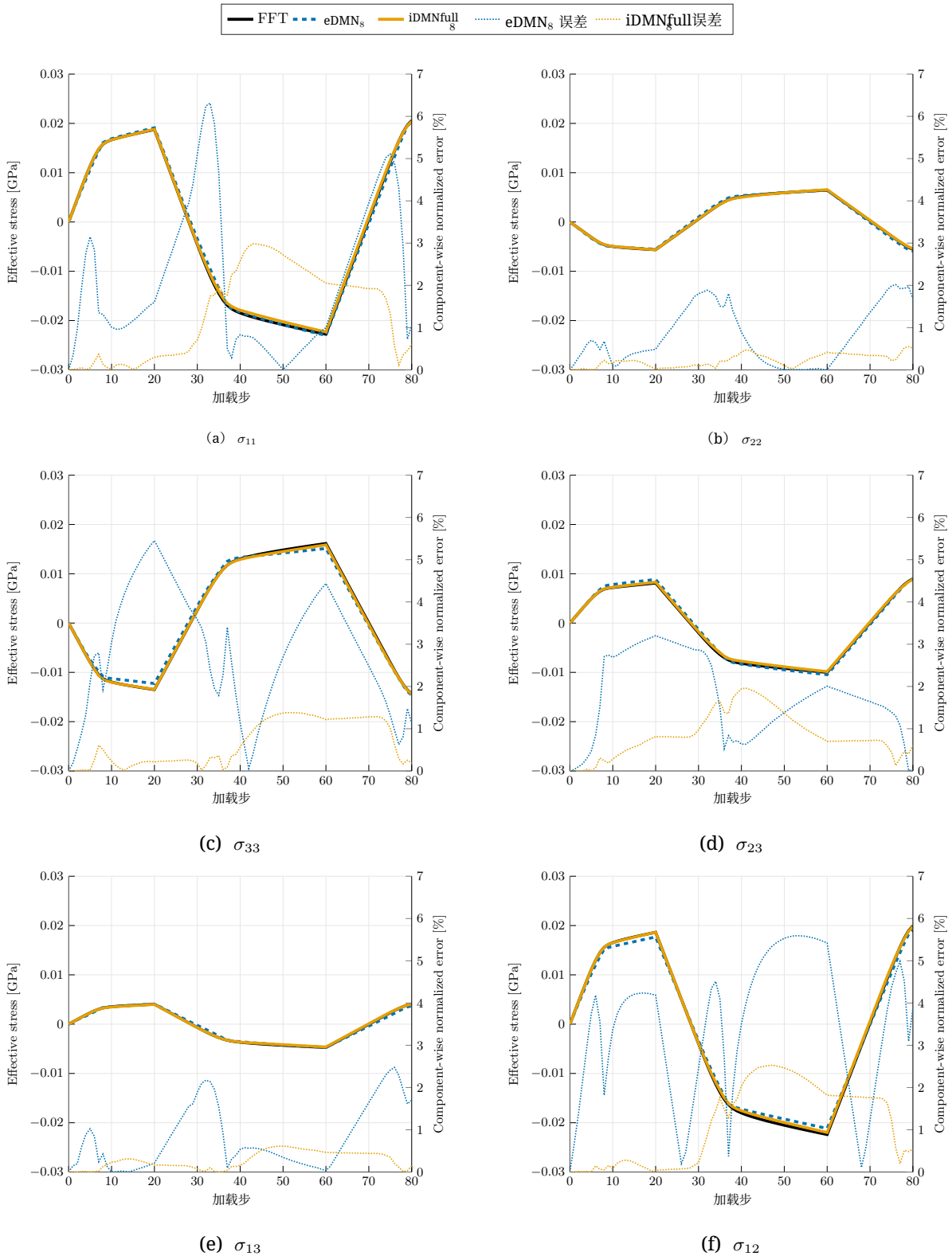


(a) σ_{11} (b) σ_{22}

(c) σ_{33} (d) σ_{23}

(e) σ_{13} (f) σ_{12}

图10：比例递增分布外路径的秩-8有效应力预测。

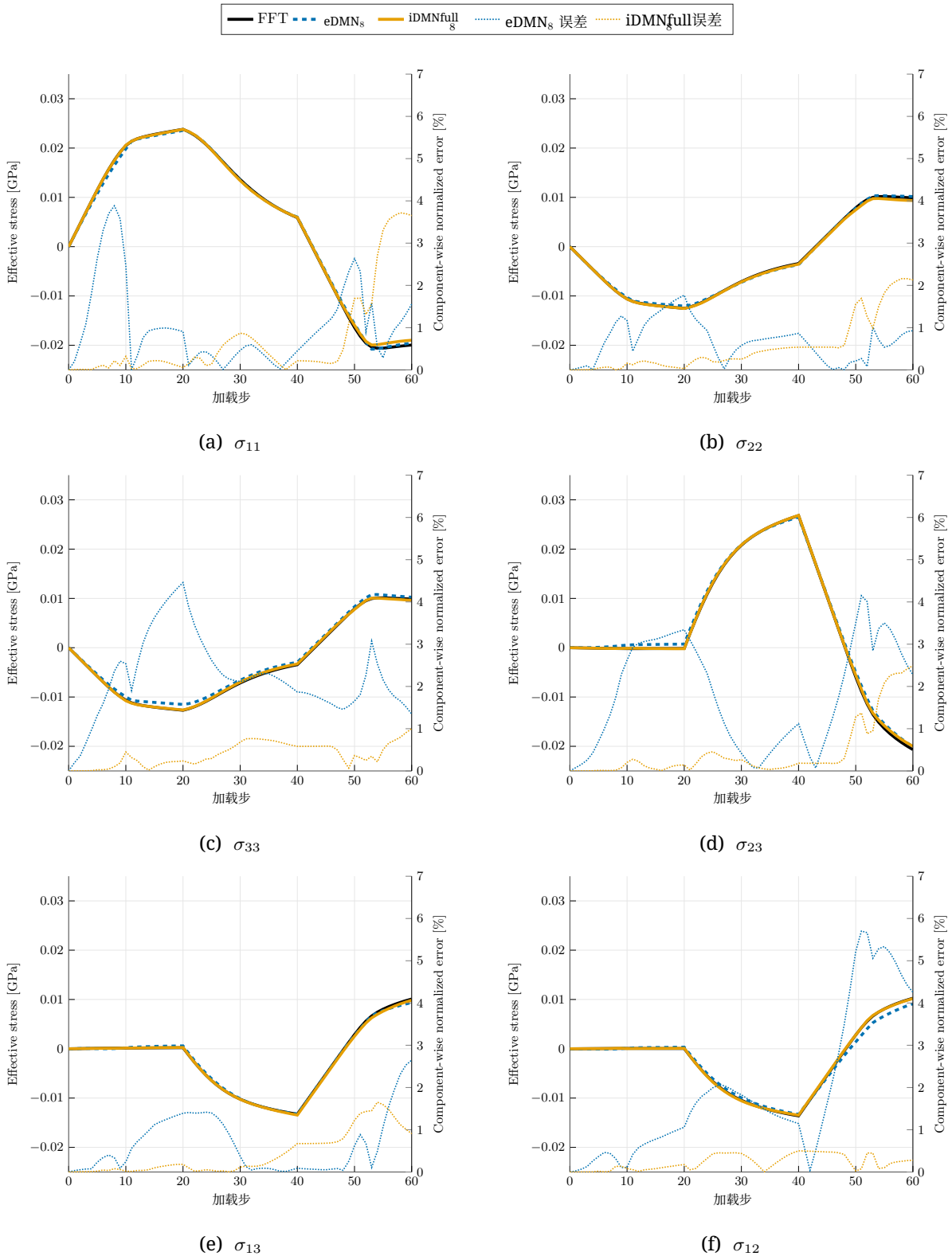


(a) σ_{11} (b) σ_{22}

(c) σ_{33} (d) σ_{23}

(e) σ_{13} (f) σ_{12}

图11：循环OOD路径的秩-8有效应力预测。



(a) σ_{11} (b) σ_{22}

(c) σ_{33} (d) σ_{23}

(e) σ_{13} (f) σ_{12}

图12：非比例OOD路径的秩-8有效应力预测。

表 8：循环加载下扰动矩阵性质的轨迹相对有效应力误差。数值为独立训练随机种子下的均值 ± 样本标准差（%）：eDMN 为 $n = 5$，每种 iDMN 配置为 $n = 3$。

模型	Rank	Nom.	E_m^-	E_m^+	ν_m^-	ν_m^+	σ_y^-	σ_y^+	H^-	H^+
eDMN	4	17.00 ± 3.77	16.54 ± 3.29	17.54 ± 4.25	15.85 ± 2.97	18.50 ± 4.86	17.76 ± 4.14	16.39 ± 3.45	17.17 ± 3.94	16.86 ± 3.61
eDMN	5	13.22 ± 3.27	13.14 ± 2.99	13.35 ± 3.55	12.56 ± 2.71	14.07 ± 4.04	13.67 ± 3.51	12.87 ± 3.07	13.33 ± 3.30	13.11 ± 3.25
eDMN	6	10.28 ± 1.41	10.46 ± 1.37	10.16 ± 1.45	10.18 ± 1.45	10.39 ± 1.38	10.50 ± 1.36	10.13 ± 1.43	10.24 ± 1.44	10.32 ± 1.39
eDMN	7	6.88 ± 0.79	7.05 ± 0.73	6.79 ± 0.86	6.81 ± 0.71	6.99 ± 0.89	7.05 ± 0.81	6.78 ± 0.76	6.95 ± 0.83	6.83 ± 0.76
eDMN	8	6.68 ± 0.99	6.82 ± 0.95	6.61 ± 1.04	6.65 ± 1.00	6.71 ± 1.00	6.83 ± 1.12	6.59 ± 0.89	6.73 ± 0.97	6.63 ± 1.01
iDMN^{128}	4	5.99 ± 0.69	6.02 ± 0.49	5.90 ± 0.77	6.00 ± 0.65	6.04 ± 0.77	6.18 ± 0.85	5.77 ± 0.46	6.11 ± 0.67	5.88 ± 0.70
iDMN^{128}	5	3.45 ± 0.18	3.35 ± 0.14	3.60 ± 0.21	3.53 ± 0.21	3.40 ± 0.16	3.73 ± 0.21	3.25 ± 0.17	3.42 ± 0.17	3.48 ± 0.19
iDMN^{128}	6	3.10 ± 0.11	2.98 ± 0.09	3.24 ± 0.13	3.15 ± 0.07	3.06 ± 0.18	3.31 ± 0.09	2.95 ± 0.13	3.05 ± 0.14	3.16 ± 0.08
iDMN^{128}	7	2.39 ± 0.14	2.30 ± 0.14	2.49 ± 0.13	2.42 ± 0.09	2.36 ± 0.19	2.58 ± 0.13	2.24 ± 0.14	2.33 ± 0.10	2.44 ± 0.17
iDMN^{128}	8	2.27 ± 0.10	2.21 ± 0.09	2.34 ± 0.10	2.30 ± 0.12	2.24 ± 0.09	2.44 ± 0.10	2.14 ± 0.09	2.20 ± 0.10	2.34 ± 0.10
iDMN^{256}	4	5.43 ± 0.85	5.42 ± 0.93	5.49 ± 0.77	5.44 ± 0.81	5.47 ± 0.89	5.69 ± 0.88	5.22 ± 0.83	5.60 ± 0.87	5.28 ± 0.84
iDMN^{256}	5	3.31 ± 0.21	3.23 ± 0.22	3.43 ± 0.20	3.45 ± 0.30	3.18 ± 0.14	3.51 ± 0.26	3.19 ± 0.17	3.28 ± 0.21	3.34 ± 0.21
iDMN^{256}	6	2.94 ± 0.32	2.80 ± 0.30	3.07 ± 0.34	2.95 ± 0.31	2.93 ± 0.33	3.14 ± 0.32	2.77 ± 0.32	2.87 ± 0.36	3.00 ± 0.29
iDMN^{256}	7	2.69 ± 0.04	2.57 ± 0.02	2.82 ± 0.05	2.71 ± 0.08	2.68 ± 0.08	2.92 ± 0.04	2.51 ± 0.04	2.64 ± 0.04	2.74 ± 0.04
iDMN^{256}	8	2.34 ± 0.24	2.29 ± 0.20	2.39 ± 0.28	2.40 ± 0.21	2.28 ± 0.28	2.52 ± 0.27	2.19 ± 0.22	2.27 ± 0.23	2.41 ± 0.25
iDMN^{512}	4	5.34 ± 1.52	5.33 ± 1.55	5.42 ± 1.48	5.42 ± 1.40	5.32 ± 1.63	5.64 ± 1.63	5.15 ± 1.40	5.42 ± 1.57	5.27 ± 1.47
iDMN^{512}	5	3.25 ± 0.57	3.15 ± 0.58	3.39 ± 0.58	3.38 ± 0.62	3.15 ± 0.53	3.49 ± 0.58	3.08 ± 0.59	3.24 ± 0.54	3.27 ± 0.61
iDMN^{512}	6	2.94 ± 0.07	2.84 ± 0.12	3.06 ± 0.04	2.98 ± 0.12	2.90 ± 0.03	3.17 ± 0.08	2.77 ± 0.08	2.86 ± 0.08	3.02 ± 0.08
iDMN^{512}	7	2.57 ± 0.09	2.44 ± 0.07	2.70 ± 0.10	2.58 ± 0.14	2.56 ± 0.05	2.79 ± 0.11	2.39 ± 0.06	2.51 ± 0.09	2.63 ± 0.09
iDMN^{512}	8	2.29 ± 0.09	2.25 ± 0.11	2.35 ± 0.07	2.34 ± 0.06	2.25 ± 0.12	2.48 ± 0.08	2.15 ± 0.09	2.23 ± 0.09	2.35 ± 0.09
iDMN^{full}	4	5.97 ± 0.82	6.10 ± 0.71	5.93 ± 0.93	6.00 ± 0.76	5.98 ± 0.90	6.10 ± 1.10	5.92 ± 0.59	6.10 ± 0.86	5.85 ± 0.78
iDMN^{full}	5	3.24 ± 0.49	3.18 ± 0.50	3.38 ± 0.49	3.35 ± 0.52	3.16 ± 0.47	3.50 ± 0.48	3.08 ± 0.52	3.25 ± 0.49	3.25 ± 0.50
iDMN^{full}	6	2.95 ± 0.31	2.87 ± 0.30	3.06 ± 0.35	3.02 ± 0.30	2.89 ± 0.34	3.18 ± 0.38	2.78 ± 0.27	2.90 ± 0.38	3.01 ± 0.26
iDMN^{full}	7	2.49 ± 0.12	2.37 ± 0.12	2.62 ± 0.12	2.50 ± 0.15	2.49 ± 0.10	2.71 ± 0.15	2.32 ± 0.09	2.44 ± 0.13	2.54 ± 0.12
iDMN^{full}	8	2.21 ± 0.12	2.16 ± 0.09	2.26 ± 0.15	2.25 ± 0.14	2.17 ± 0.13	2.38 ± 0.15	2.07 ± 0.10	2.15 ± 0.10	2.27 ± 0.14

[14] M. Schneider, 《非线性快速傅里叶变换计算均匀化方法综述》, 《力学学报》, 第232卷, 第6期, 第2051–2100页, 2021年。[15] R. Smit、W. Brekelmans 和 H. Meijer, 《通过多层次有限元建模预测非线性非均质系统的力学行为》, 《计算力学与工程方法》, 第155卷, 第181–192页, 1998年3月。[16] F. Feyel, 《复合结构的多尺度有限元弹黏塑性分析》, 《计算材料科学》, 第16卷, 第344–354页, 1999年12月。[17] J. Kochmann、S. Wulfinghoff、S. Reese、J. R. Mianroodi 和 B. Svendsen, 《基于双尺度有限元–快速傅里叶变换和相场的体材料微观结构演化与宏观材料行为计算建模》, 《计算力学与工程方法》, 第305卷, 第89–110页, 2016年。[18] J. Kochmann、L. Ehle、S. Wulfinghoff、J. Mayer、B. Svendsen 和 S. Reese, 《基于多尺度有限元–快速傅里叶变换的高效建模与仿真：具有非线性非均质微观结构的宏观变形过程》, 载于《非均质结构的多尺度建模》, 应用与计算力学讲义, 第129–146页, 施普林格国际出版公司, 2017年。[19] G. J. Dvorak, 《非弹性复合材料的变换场分析》, 伦敦皇家学会会刊：A辑：数学与物理科学, 第437卷, 第1900期, 第311–327页, 1992年。[20] J. Fish、K. Shek、M. Pandheeradi 和 M. S. Shephard, 《基于数学均匀化的复合结构计算塑性：理论与实践》, 《计算力学与工程方法》, 第148卷, 第53–73页, 1997年8月。[21] J.-C. Michel 和 P. Suquet, 《非均匀变换场分析》, 《国际固体与结构杂志》, 第40卷, 第25期, 第6937–6955页, 2003年。[22] J. Michel 和 P. Suquet, 《利用非均匀变换场分析对非线性复合结构进行计算分析》, 《计算力学与工程方法》, 第193卷, 第48–51期, 第5477–5502页, 2004年。计算塑性进展。[23] F. Fritzen 和 T. Böhlke, 《非均匀变换场分析的三维有限元实现》, 《国际工程数值方法杂志》, 第84卷, 第7期, 第803–829页, 2010年。[24] R. Largenton、J.-C. Michel 和 P. Suquet, 《将非均匀变换场分析扩展至存在老化和溶胀的线性黏弹性复合材料》, 《材料力学》, 第73卷, 第76–100页, 2014年。[25] F. Fritzen 和 M. Leuschner, 《基于混合增量形式的降阶基混合计算均匀化》, 《计算力学与工程方法》, 第260卷, 第143–154页, 2013年。[26] J.-C. Michel 和 P. Suquet, 《在保留本构关系变分结构的前提下，对材料微观力学进行模型降阶》, 《力学与固体物理学杂志》, 第90卷, 第254–285页, 2016年。[27] Z. Liu、M. Bessa 和 W. K. Liu, 《自洽聚类分析：一种用于非弹性非均质材料的高效多尺度方案》, 《计算力学与工程方法》, 第306卷, 第319–341页, 2016年。[28] Z. Liu、M. Fleming 和 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》, 《计算力学与工程方法》, 第330卷, 第547–577页, 2018年。[29] R. Hill, 《复合材料的自洽力学》, 《力学与固体物理学杂志》, 第13卷, 第4期, 第213–222页, 1965年。[30] M. Schneider, 《小应变下非线性材料自洽聚类分析的数学基础》, 《计算力学与工程方法》, 第354卷, 第783–801页, 2019年。[31] J. Yvonnet 和 Q.-C. He, 《有限应变下超弹性介质非线性均匀化的降阶模型多尺度方法（R3M）》, 《计算物理学杂志》, 第223卷, 第1期, 第341–368页, 2007年。[32] J. Hernández、J. Oliver、A. E. Huespe、M. Caicedo 和 J. Cante, 《计算多尺度均匀化中的高性能模型降阶技术》, 《计算力学与工程方法》, 第276卷, 第149–189页, 2014年。[33] J. A. Hernández、M. A. Caicedo 和 A. Ferrer, 《通过经验求积实现非线性有限元模型的维数超降阶》, 《计算力学与工程方法》, 第313卷, 第687–722页, 2017年。[34] N. Lange、G. Hütter 和 B. Kiefer, 《有限变形下采用聚类训练的整体式超降阶有限元方法》, 《计算力学与工程方法》, 第418卷, 文章编号116522, 2024年。[35]S. Wulfinghoff, 《利用经验校正簇求积（E3C）超降阶进行超弹性均匀化》, 《计算力学与工程方法》, 第444卷, 文章编号118137, 2025年。

[36] S. Wulfinghoff 和 J. Hauck，《非线性力学计算均匀化中的 E3C 超降阶》，《计算力学》，第76卷，第6期，第1563–1572页，2025年。[37] S. Wulfinghoff，《利用经验校正簇求积（E3C）超降阶进行计算晶体塑性均匀化》，《计算力学与工程方法》，第452卷，文章编号118681，2026年。[38] S. Wulfinghoff、F. Cavaliere 和 S. Reese，《利用 Hashin–Shtrikman 型有限元方法对非线性均匀化问题进行模型降阶》，《计算力学与工程方法》，第330卷，第149–179页，2018年。[39] M. Kabel、F. Ospald 和 M. Schneider，《一种用于有限应变计算均匀化的模型降阶方法：利用超弹性层合材料近似界面并应用于规则网格》，《计算力学与工程方法》，第309卷，第476–496页，2016年。[40] J. Ghaboussi、J. Garrett Jr 和 X. Wu，《利用神经网络进行基于知识的材料行为建模》，《工程力学杂志》，第117卷，第1期，第132–153页，1991年。[41] J. Ghaboussi、X. Wu 和 G. Kaklauskas，《神经网络材料建模》，Statyba，第5卷，第4期，第250–257页，1999年。[42] T. Furukawa 和 G. Yagawa，《利用神经网络进行黏塑性的隐式本构建模》，《国际工程数值方法杂志》，第43卷，第2期，第195–219页，1998年。[43] X. Li、C. C. Roth 和 D. Mohr，《基于机器学习的温度和速率相关塑性模型：在双相钢断裂实验分析中的应用》，《国际塑性学杂志》，第118卷，第320–344页，2019年。[44] M. B. Gorji、M. Mozaffar、J. N. Heidenreich、J. Cao 和 D. Mohr，《循环神经网络在路径相关塑性建模中的潜力》，《力学与固体物理学杂志》，第143卷，文章编号103972，2020年。[45] M. Mozaffar、R. Bostanabad、W. Chen、K. Ehmann、J. Cao 和 M. A. Bessa，《深度学习预测路径相关塑性》，《美国国家科学院院刊》，第116卷，第52期，第26414–26420页，2019年。[46] C. Bonatti 和 D. Mohr，《自洽性对于表征弹塑性固体的循环神经网络模型的重要性》，《力学与固体物理学杂志》，第158卷，文章编号104697，2022年。[47] F. Masi、I. Stefanou、P. Vannucci 和 V. Maffi-Berthier，《基于热力学的人工神经网络用于本构建模》，《力学与固体物理学杂志》，第147卷，文章编号104277，2021年。[48] R. E. Jones、A. L. Frankel 和 K. L. Johnson，《一种通过内变量对非弹性应力响应进行建模的神经常微分方程框架》，《建模与计算机器学习杂志》，第3卷，第3期，第1–35页，2022年。[49] B. Guo、Z. Lin 和 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》，《计算力学与工程方法》，第447卷，文章编号118358，2025年。[50] L. Lu、P. Jin、G. Pang、Z. Zhang 和 G. E. Karniadakis，《基于算子普适逼近定理的 DeepONet 非线性算子学习》，《自然·机器智能》，第3卷，第3期，第218–229页，2021年。[51] N. Kovachki、Z. Li、B. Liu、K. Azizzadenesheli、K. Bhattacharya、A. Stuart 和 A. Anandkumar，《神经算子：学习函数空间之间的映射及其在偏微分方程中的应用》，《机器学习研究杂志》，第24卷，第89期，第1–97页，2023年。[52] B. Le、J. Yvonnet 和 Q.-C. He，《利用神经网络对非线性弹性材料进行计算均匀化》，《国际工程数值方法杂志》，第104卷，第12期，第1061–1084页，2015年。[53] J. Yvonnet、D. Gonzalez 和 Q.-C. He，《用于非线性弹性非均质材料均匀化的数值显式势》，《计算力学与工程方法》，第198卷，第33–36期，第2723–2737页，2009年。[54] V. M. Nguyen-Thanh、L. T. K. Nguyen、T. Rabczuk 和 X. Zhuang，《利用基于高维模型表示的神经网络构建有限应变弹性静力学计算均匀化代理模型》，《国际工程数值方法杂志》，第121卷，第21期，第4811–4842页，2020年。[55] L. Wu、L. Adam 和 L. Noels，《机织复合材料多尺度分析的微观力学与数据驱动降阶模型》，《复合材料结构》，第270卷，文章编号114058，2021年。[56] B. H. Nguyen 和 M. Schneider，《用于线弹性周期均匀化问题的通用傅里叶神经算子》，《力学与固体物理学杂志》，第206卷，文章编号106418，2026年。[57] Y. LeCun、L. Bottou、Y. Bengio 和 P. Haffner，《将基于梯度的学习应用于文档识别》，《IEEE学报》，第86卷，第11期，第2278–2324页，1998年。

[58] J. Bruna 和 S. Mallat，“不变散射卷积网络”，IEEE《模式分析与机器智能汇刊》，第 35 卷，第 1872–1886 页，2013年8月。[59] K. S. Tai、P. Bailis 和 G. Valiant，“等变变换器网络”，载于《第36届国际机器学习会议论文集》，《机器学习研究论文集》第97卷，第6086–6095页，2019年。[60] Z. Liu、C. Wu 和 M. Koishi，“用于多尺度拓扑学习及非均质材料加速非线性建模的深度材料网络”，《计算力学与工程方法》，第345卷，第1138–1168页，2019年。[61] Z. Liu 和 C. 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”，《力学与固体物理学杂志》，第127卷，第20–46页，2019年。[62] Z. Liu、C. T. Wu 和 M. Koishi，“深度材料网络的迁移学习：实现结构—性能预测的无缝衔接”，《计算力学》，第64卷，第2期，第451–465页，2019年。[63] D. Shin、R. Alberdi、R. A. Lebensohn 和 R. Dingreville，“用于预测复合材料热力学响应的深度材料网络方法”，《复合材料B辑：工程》，第272卷，第111177页，2024年。[64] D. Shin、P. J. Creveling、S. A. Roberts 和 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”，《计算力学与工程方法》，第431卷，第117279页，2024年。[65] Z. Liu，“具有内聚层的深度材料网络：多阶段训练与界面失效分析”，《计算力学与工程方法》，第363卷，第112913页，2020年。[66] 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细分”，《计算力学与工程方法》，第384卷，第113914页，2021年。[67]A. E. Robertson、S. B. Inman、A. T. Lenau、R. A. Lebensohn、D. Shin、B. L. Boyce 和 R. Dingreville，“基于微观结构的变分深度材料网络：提升材料数字孪生中不确定性量化的稳健性”，npj《计算材料》，第12卷，第1期，第261页，2026年。[68] T. Huang、Z. Liu、C. Wu 和 W. Chen，“微观结构引导的深度材料网络：用于快速非线性材料建模和不确定性量化”，《计算力学与工程方法》，第398卷，第115197页，2022年。[69] D. Shin、R. Alberdi、R. A. Lebensohn 和 R. Dingreville，“通过拼接策略构建深度材料网络：用于可解释性分析的可视化，以及提升准确度的递归训练”，npj《计算材料》，第9卷，第1期，第128页，2023年。[70] T. Li，“微观力学信息驱动的参数化深度材料网络：用于预测形态变化的非均质材料的物理行为”，《计算力学与工程方法》，第419卷，第116687页，2024年。[71] J. G. Jean、T.-H. Su、S.-J. Huang、C.-T. Wu 和 C.-S. Chen，“图增强深度材料网络：利用微观结构信息学进行多尺度材料建模”，《计算力学》，第75卷，第1期，第113–136页，2025年。[72] N. M. Francis、D. Shin、R. A. Lebensohn、F. Pourahmadian 和 R. Dingreville，“微极深度材料网络”，《计算力学与工程方法》，第446卷，第118329页，2025年。[73] Z. Liu、H. Wei、T. Huang 和 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年。[74] H. Wei、C. T. Wu、W. Hu、T. H. Su、H. Oura、M. Nishi、T. Naito、S. Chung 和 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》，第149卷，第3期，04023003，2023年。[75] W.-N. Wan、T.-J. Wei、T.-H. Su 和 C.-S. Chen，“解码材料网络：探索深度材料网络和基于相互作用的材料网络的性能”，《力学杂志》，第40卷，第796–807页，2024年。[76] T.-J. Wei、W.-N. Wan 和 C.-S. Chen，“深度材料网络：概述、应用与当前方向”，《多尺度科学与工程》，第8卷，第1期，第1–19页，2026年。[77] X. He、H. Wei、W. Hu、H. Mao 和 C. T. Wu，“深度材料网络在多尺度材料建模中的系统性能评估。”arXiv预印本 arXiv:2602.07192，2026年。[78] S. Gajek、M. Schneider 和 T. Böhlke，“深度材料网络的微观力学研究”，《力学与固体物理学杂志》，第142卷，第103984页，2020年。[79] G. W. Milton，《复合材料理论》。工业与应用数学学会，2022年1月。

[80] S. Gajek、M. Schneider 和 T. Böhlke，“用于短纤维增强塑料构件多尺度分析的 FE–DMN 方法”，《计算力学与工程方法》，第384卷，第113952页，2021年。[81] S. Gajek、M. Schneider 和 T. Böhlke，“用于热机械复合材料多尺度分析的 FE-DMN 方法”，《计算力学》，第69卷，第5期，第1087–1113页，2022年。[82] S. Gajek、M. Schneider 和 T. Böhlke，“利用深度材料网络对微结构材料进行高效双尺度模拟”，PAMM，第21卷，第1期，e202100069，2021年。[83] A. P. Dey、F. Welschinger、M. Schneider、S. Gajek 和 T. Böhlke，“基于深度材料网络，快速逆向标定短纤维增强热塑性塑料黏塑性及蠕变行为的多尺度模型”，《国际塑性学杂志》，第160卷，第103484页，2023年。[84] A. P. Dey、F. Welschinger、M. Schneider、J. Köbler 和 T. Böhlke，“深度材料网络在高度非线性加载工况下短纤维增强热塑性塑料多尺度虚拟表征中的有效性”，《应用力学档案》，第94卷，第5期，第1177–1202页，2024年。[85] N. Meyer、S. Gajek、J. Görthofer、A. Hrymak、L. Kärger、F. Henning、M. Schneider 和 T. Böhlke，“用于量化片状模塑料工艺诱发不确定性的概率虚拟工艺链”，《复合材料B辑：工程》，第249卷，第110380页，2023年。[86] B. Sterr、S. Gajek、A. Hrymak、M. Schneider 和 T. Böhlke，“适用于无限材料对比纤维悬浮体的深度材料网络”，《国际工程数值方法杂志》，第126卷，第7期，e70014，2025年。[87] J. Lendvai 和 M. Schneider，“具有退化芯层的涂层层合板的非退化性”，《力学学报》，第237卷，第8期，第4717–4744页，2026年。[88] P. B. K. Srinivas、M. Kabel 和 M. Schneider，“通过基于位移的层合板公式以及针对基于柔度的疲劳模型的新型采样技术，对深度材料网络进行快速离线训练”，《计算力学与工程方法》，第449卷，第118517页，2026年。[89] V. D. Nguyen 和 L. Noels，“从相互作用视角重新审视基于细观力学的材料网络：针对多相复合材料的稳健高效实现”，《欧洲力学杂志——A辑：固体》，第91卷，第104384页，2022年。[90] V. D. Nguyen 和 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”，《计算力学与工程方法》，第389卷，第114300页，2022年。[91] T.-J. Wei、T.-H. Su 和 C.-S. Chen，“用于多晶材料建模的方向感知型基于相互作用的深度材料网络”，《计算力学与工程方法》，第441卷，第117977页，2025年。[92] A. P. Dey、F. Welschinger、M. Schneider、S. Gajek 和 T. Böhlke，“采用非弹性知情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加载”，《应用力学档案》，第92卷，第9期，第2733–2755页，2022年。[93] S. Gajek、M. Schneider 和 T. Böhlke，“基于材料信息的黏弹性深度材料网络训练”，PAMM，第22卷，第1期，e202200143，2023年。[94] D. Shin、R. A. Lebensohn 和 R. Dingreville，“结合微结构局部场信息的 U-net 架构深度材料网络训练”，《计算力学》，第78卷，第1期，第1–21页，2026年。[95] P. Michaleris、D. A. Tortorelli 和 C. A. Vidal，“瞬态非线性耦合问题的切线算子与设计灵敏度公式及其在弹塑性中的应用”，《国际工程数值方法杂志》，第37卷，第14期，第2471–2499页，1994年。[96] T. Schwedes、D. A. Ham、S. W. Funke 和 M. D. Piggott，“一项应用：潮汐涡轮机阵列布局优化”，载于《PDE约束优化中的网格依赖性：潮汐涡轮机阵列布局中的应用》，第79–107页，施普林格，2017年。[97] T. Bui-Thanh，“伴随变量及其在计算科学、工程学和数学中的作用”，2026年。[98] R. T. Chen、Y. Rubanova、J. Bettencourt 和 D. K. Duvenaud，“神经常微分方程”，载于《神经信息处理系统进展》，第31卷，2018年。[99] S. Bai、J. Z. Kolter 和 V. Koltun，“深度平衡模型”，载于《神经信息处理系统进展》，第32卷，2019年。[100] P. Suquet，“非弹性固体力学中的均匀化要素”，载于《复合介质均匀化技术》，《物理学讲义》，第193–278页，柏林、海德堡：施普林格，1987年。[101] B. Halphen 和 Q. S. Nguyen，“论广义标准材料”，《力学杂志》，第14卷，第39–63页，1975年。

[102] M. Kabel、D. Merkert 和 M. Schneider，《在基于FFT的均匀化中使用复合体素》，《计算力学与工程方法》，第 294 卷，页 168–188，2015年。[103] P. B. K. Srinivas、M. Krier、F. Welschinger、M. Kabel 和 M. Schneider，《深度材料网络与降阶模型：多尺度部件模拟的比较研究》，*PAMM*，第 26 卷，第 4 期，页 e70193，2026年。[104] J. Bradbury、R. Frostig、P. Hawkins、M. J. Johnson、C. Leary、D. Maclaurin、G. Necula、A. Paszke、J. VanderPlas、S. Wanderman-Milne 和 Q. Zhang，《JAX：Python+NumPy 程序的可组合变换》。http:// github.com/jax-ml/jax，2018年。[105] S. Montgomery-Smith、D. Jack 和 D. E. Smith，《含扩散的 Jeffery方程的快速精确闭合》，《非牛顿流体力学杂志》，第 166 卷，第 7-8 期，页 343–353，2011年。[106] M. Schneider，《用于生成短纤维增强塑料均匀化代表性体积元的顺序添加与迁移方法》，《计算力学》，第 59 卷，第 2 期，页 247–263，2017年。[107] M. Schneider，《关于基于FFT的计算均匀化中的 Barzilai–Borwein基本方案》，《国际工程数值方法杂志》，第 118 卷，第 8 期，页 482–494，2019年。[108] D. J. Eyre 和 G. W. Milton，《一种利用网格细化计算复合材料响应的快速数值方案》，《欧洲物理学杂志：应用物理》，第 6 卷，第 1 期，页 41–47，1999年。[109] M. Schneider、D. Wicht 和 T. Böhlke，《关于非弹性材料基于FFT的计算均匀化中的基于极化的方案》，《计算力学》，第 64 卷，第 4 期，页 1073–1095，2019年。[110] M. Krause 和 T. Böhlke，《任意阶张量调和基及其在弹性、弹黏塑性和基于织构的建模中的应用》，《固体数学与力学》，第 30 卷，第 5 期，页 1188–1212，2025年。[111] D. P. Kingma 和 J. Ba，《Adam：一种随机优化方法》，载于第三届国际学习表征会议（*ICLR*），（圣迭戈），2015年。[112] S. Balay、S. Abhyankar、M. F. Adams、S. Benson、J. Brown、P. Brune、K. Buschelman、E. M. Constantinescu、L. Dalcin、A. Dener、V. Eijkhout、J. Faibussowitsch、W. D. Gropp、V. Hapla、T. Isaac、P. Jolivet、D. Karpeev、D. Kaushik、M. G. Knepley、F. Kong、S. Kruger、D. A. May、L. C. McInnes、R. T. Mills、L. Mitchell、T. Munson、J. E. Roman、K. Rupp、P. Sanan、J. Sarich、B. F. Smith、S. Zampini、H. Zhang、H. Zhang 和 J. Zhang，《PETSc 网页》。https://petsc.org/，2026年。[113] G. W. Milton，《复合材料理论中的一些未解问题》，《英国皇家学会哲学汇刊*A*辑：数学、物理与工程科学》，第 379 卷，第 2201 期，页 20200115，2021年。[114] P. Wriggers，非线性有限元方法。施普林格科学与商业媒体，2008年。