利用无条件扩散模型对具有定制力学性能的微结构进行风格约束逆向设计

Weipeng Xu[1], Ziyuan Xie[1], Haoju Lin[1],Xinyu Wang[1],Guangjin Mou[1],和 Tianju Xue *[1]

[1]香港科技大学土木与环境工程系，中国香港

arXiv:2601.06469v1 [cs.CE] 10 Jan 2026

摘要

深度生成模型，尤其是去噪扩散模型，已在具有所需性能和风格的架构化微结构高保真生成方面取得了显著成功。然而，这些近期方法通常依赖条件训练机制，并且需要投入大量计算资源来准备标注训练数据集；这使得它们缺乏灵活性，因为控制方程或边界条件的任何变化都需要进行完整的重新训练。在本研究中，我们提出了一种新的逆向设计框架，将无条件去噪扩散模型与可微编程技术相结合，用于生成架构化微结构。我们的方法无需为不同的问题设置准备昂贵的标注数据集，也无需重新训练。通过将扩散模型的噪声输入重新解释为可优化的设计变量，我们将设计任务表述为关于噪声输入的优化问题，从而能够控制反向去噪轨迹，引导生成的微结构趋向所需的力学性能，同时保留训练数据集中编码的风格约束。我们通过向量-雅可比积拼接建立了统一微分流程，以实现通过反向传播进行端到端梯度计算。若干数值算例展示了该框架的有效性及其在涉及多样性能和风格要求的先进设计任务中的潜力，这些算例涵盖了具有指定均匀化性能的微结构设计，以及具有目标超弹性和弹塑性行为的微结构设计。

关键词：扩散模型、逆向设计、可微编程、架构化微结构。

1 引言

架构化微结构已广泛应用于先进工程场景 [1, 2]，这得益于其优异且可调的力学性能。其复杂的几何构型使其能够突破基体材料固有的局限，从而实现一系列新颖功能，例如负泊松比 [3]，能量吸收 [4]，和应变隐身 [5]。架构化微结构的设计任务包含两个既不同又相互关联的方面：性能与风格。性能指微结构应表现出的期望力学行为或性能，例如特定的刚度、强度或变形特性。另一方面，风格涉及微结构的几何特征和美学特质，这些可能受到几何与制造约束 [6]，或视觉吸引力 [7]的影响。在实际应用中成功实现架构化微结构，关键在于使性能与风格达到和谐平衡。

*cetxue@ust.hk （通讯作者）

在设计方法方面，尽管基于启发式的直接设计 [8, 9] 和基于优化的逆向设计 [4, 10] 已被广泛用于定制微结构的性能，但要将风格约束纳入其中，往往还需付出额外努力 [11, 12]。近年来，数据驱动方法 [13, 14] 为先进材料与结构的计算分析和设计建立了新的范式。在这些方法中，深度生成模型，例如变分自编码器（VAE） [15], 生成对抗网络（GAN） [16], 基于流的模型 [17], 和去噪扩散模型 [18–20], 正逐渐成为强大的工具，可用于对具有定制力学性能的微结构进行风格约束逆向设计。早期研究 [21–25] 主要关注将 VAE 和 GAN 用于微结构设计。近年来，由于生成质量和多样性更出色，去噪扩散模型和基于流的模型也开始应用于这一任务。

近期研究通常采用条件扩散模型，在保留训练数据集中编码的风格特征的同时，引导生成过程朝向期望的性能。例如，Vlassis 和 Sun [26] 开发了一种条件去噪扩散算法，用于生成具有指定应力-应变响应和手写数字风格的超弹性微结构；生成过程由嵌入式特征向量引导，这些向量对目标力学行为进行编码。Bastek 和 Kochmann [27] 提出了一种视频扩散生成模型，以目标应力-应变曲线为条件，逆向设计周期性随机胞状结构，使其在大应变压缩下具有经过调控的非线性变形和应力响应。Wang 等人 [28] 提出了一种用于设计吸能微结构的条件去噪扩散框架 DiffMat，其中目标应力-应变响应通过交叉注意力机制进行编码。Li 等人 [29]利用条件去噪扩散模型生成具有目标超弹性性能的复合微结构。应变能值和反作用力被用作训练和生成过程的条件输入。Zhang 等人 [30] 开发了条件扩散模型 HyperDiff，用于逆向设计具有目标力-位移曲线的超弹性多孔微结构。对模拟得到的力-位移曲线进行平滑处理，并使用 B样条曲线对其参数化；归一化后的系数和节点则作为扩散模型的条件输入。

此外，近期一些研究探索了采用先进的模型架构和额外的引导技术，以提升生成质量和效率。Mirzaee 等人 [31] 提出了一种基于流的条件生成设计框架 PoreFlow，用于生成具有指定物理性能的多孔材料微结构。Yang 等人 [32] 提出了 GraphDGM，这是一种基于图的条件扩散模型，用于设计框架结构和晶格结构。Zhang 等人 [33]提出了条件扩散模型 LatticeOptDiff，用于设计具有定制弹性性能和体积分数的晶格结构；该模型通过性能回归网络引导生成过程，以提高设计精度。Zheng 等人 [34]提出了条件逆向设计框架 DiffuMeta，用于设计具有目标弹性性能、且在大变形下具有目标应力-应变响应的壳型结构。该方法引入基于代数语言的参数化表示，将复杂几何结构编码为数学词元序列，并将其集成到扩散Transformer架构中，以实现高质量生成。

上述条件生成设计方法必须依赖包含带标签的属性信息的大规模数据集进行训练。带标签的数据通常通过迭代数值模拟获得。控制方程或边界条件的任何变化，都需要重新生成完整的标注数据集并重新训练生成模型，这限制了它们在不同设计场景中的灵活性和适用性。

在本研究中，我们提出了一种新的逆向设计框架，将无条件 去噪扩散模型与先进的可微编程技术相结合 [35] ，用于设计具有定制力学性能并满足风格约束的架构化微结构。我们的方法无需针对不同问题设置进行昂贵的标注数据集准备和模型重新训练。该方法的核心思想是将扩散模型的噪声输入重新解释为

一种可优化的设计变量，用于编码目标微观结构的几何特征。通过将设计任务表述为针对噪声输入的优化问题，我们可以控制反向去噪轨迹，引导生成的微结构达到所需的力学性能。训练数据集中编码的风格约束在优化过程中得以自然保留，从而确保最终微结构保持所需的几何特征。我们通过向量-雅可比积（VJP）级联实现统一的求导流程，以支持通过反向传播进行端到端梯度计算 [36, 37], ，从而能够使用基于梯度的优化算法高效求解逆向设计问题。

本文其余部分的组织如下。在第 2 节中，我们介绍所提出的、针对具有定制力学性能的架构化微结构的受风格约束逆向设计框架。在第 3 节中，我们通过不同力学场景下的多个数值算例展示所提方法的有效性，包括具有指定均匀化性能、超弹性性能和弹塑性性能的微结构设计。此外，我们还讨论所提方法的优势、局限性及潜在的未来发展方向。最后，我们在第4节总结全文。

2 逆向　　设计框架：使用无条件扩散模型　　s

本节首先简要介绍去噪扩散概率模型（DDPM） [18] 和去噪扩散隐式模型（DDIM） [19]。随后，我们详细说明所提出的逆向设计算法，并介绍通过使用向量-雅可比积拼接的统一微分流程进行梯度计算的过程。最后，我们给出一个高斯混合模型的概念性示例，以阐释我们方法的核心机制。

2.1 去噪扩散模型

DDPM 的标准工作流程包含两个过程：正向扩散过程逐渐向数据中添加噪声，反向去噪过程则学习去除噪声并恢复原始数据分布。

在正向扩散过程中，给定一个从数据分布 $q(\boldsymbol{x}_0)$中抽取的样本 $\boldsymbol{x}_0$ ，通过在 T 个时间步中逐步添加高斯噪声，生成一系列潜变量 $\boldsymbol{x}_t$ （$t = 1, \ldots\ldots, T$）。该过程可定义为一个马尔可夫链

$$q(\boldsymbol{x}_{1:T} \mid \boldsymbol{x}_0) = \prod_{t=1}^{T} q(\boldsymbol{x}_t \mid \boldsymbol{x}_{t-1}), \tag{1}$$

其中，转移核 $q(\boldsymbol{x}_t \mid \boldsymbol{x}_{t-1})$ 定义为

$$q(\boldsymbol{x}_t \mid \boldsymbol{x}_{t-1}) = \mathcal{N}(\boldsymbol{x}_t; \sqrt{1-\beta_t}\boldsymbol{x}_{t-1}, \beta_t \mathbf{I}), \tag{2}$$

其中 $\beta_t \in (0, 1)$ 是一个方差调度，用于控制每个时间步添加的噪声量。随着 t 趋近于 T，该过程会逐步将数据分布转换为各向同性高斯分布 $\mathcal{N}(\mathbf{0}, \mathbf{I})$。正向过程还可以进一步表示为闭式形式：

$$q(\boldsymbol{x}_t \mid \boldsymbol{x}_0) = \mathcal{N}(\boldsymbol{x}_t; \sqrt{\bar{\alpha}_t}\boldsymbol{x}_0, (1-\bar{\alpha}_t)\mathbf{I}), \tag{3}$$

其中 $\alpha_t = 1 - \beta_t$ 和 $\bar{\alpha}_t = \prod_{s=1}^{t} \alpha_s$。然后，我们可以在任意时间步 t 直接从 $\boldsymbol{x}_0$和采样的噪声 $\boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ 中采样 $\boldsymbol{x}_t$ ，即

$$\boldsymbol{x}_t = \sqrt{\bar{\alpha}_t}\boldsymbol{x}_0 + \sqrt{1-\bar{\alpha}_t}\boldsymbol{\epsilon}. \tag{4}$$

反向去噪过程旨在学习参数化转移 $p_\theta(\boldsymbol{x}_{t-1} \mid \boldsymbol{x}_t)$，逐步从 $\boldsymbol{x}_T$ 中移除噪声，以恢复原始数据 $\boldsymbol{x}_0$。该过程可定义为

$$p_\theta(\boldsymbol{x}_{0:T}) = p(\boldsymbol{x}_T) \prod_{t=1}^{T} p_\theta(\boldsymbol{x}_{t-1} \mid \boldsymbol{x}_t), \tag{5}$$

其中 $p(\boldsymbol{x}_T) = \mathcal{N}(\boldsymbol{x}_T; \mathbf{0}, \mathbf{I})$ 是先验分布，作为反向过程的起点。转移核 $p_\theta(\boldsymbol{x}_{t-1} \mid \boldsymbol{x}_t)$ 参数化为

$$p_\theta(\boldsymbol{x}_{t-1} \mid \boldsymbol{x}_t) = \mathcal{N}(\boldsymbol{x}_{t-1}; \boldsymbol{\mu}_\theta(\boldsymbol{x}_t, t), \boldsymbol{\Sigma}_\theta(\boldsymbol{x}_t, t)), \tag{6}$$

其均值 $\boldsymbol{\mu}_\theta(\boldsymbol{x}_t,t)$ 为

$$\boldsymbol{\mu}_\theta(\boldsymbol{x}_t, t) = \frac{1}{\sqrt{\alpha_t}} \left(\boldsymbol{x}_t - \frac{\beta_t}{\sqrt{1-\bar{\alpha}_t}} \boldsymbol{\epsilon}_\theta(\boldsymbol{x}_t, t) \right), \tag{7}$$

其中 $\boldsymbol{\epsilon}_\theta$ 是一个去噪神经网络，其参数为 θ，用于预测每个时间步添加的噪声。方差 $\boldsymbol{\Sigma}_\theta$ 可以建模为固定参数或可学习参数。我们遵循原始去噪扩散概率模型（DDPM）论文 [18] 并将其设为 $\boldsymbol{\Sigma}_\theta(\boldsymbol{x}_t,t) = \tilde{\beta}_t \mathbf{I}$，其中 $\tilde{\beta}_t$ 定义为

$$\tilde{\beta}_t = \frac{1-\bar{\alpha}_{t-1}}{1-\bar{\alpha}_t} \beta_t. \tag{8}$$

去噪　oisi噪神经网络通过最小化以下目标函数进行训练：

$$\min_\theta \mathbb{E}_{t, \boldsymbol{x}_0, \boldsymbol{\epsilon}} \left[\left\| \boldsymbol{\epsilon} - \boldsymbol{\epsilon}_\theta(\sqrt{\bar{\alpha}_t}\boldsymbol{x}_0 + \sqrt{1-\bar{\alpha}_t}\boldsymbol{\epsilon}, t) \right\|_2^2 \right]. \tag{9}$$

训练好的模型可以从随机噪声 $\boldsymbol{x}_T \sim p(\boldsymbol{x}_T)$ 开始生成新样本，并通过迭代应用学习到的反向转移 $p_\theta(\boldsymbol{x}_{t-1} \mid \boldsymbol{x}_t)$ 得到 $\boldsymbol{x}_0$。对于每个时间步 t，采样过程可表示为

$$\boldsymbol{x}_{t-1} = \frac{1}{\sqrt{\alpha_t}} \left(\boldsymbol{x}_t - \frac{1-\alpha_t}{\sqrt{1-\bar{\alpha}_t}} \boldsymbol{\epsilon}_\theta(\boldsymbol{x}_t, t) \right) + \sigma_t \boldsymbol{z}, \tag{10}$$

其中 $\sigma_t = \sqrt{\tilde{\beta}_t}$和 $\boldsymbol{z} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ 是每个时间步添加的随机噪声。然而，使用马尔可夫过程（10）进行采样需要大量时间步才能获得高质量结果，这可能带来高昂的计算成本。

为了解决这一问题，DDIM [19] 提出了一种非马尔可夫反向过程，在保持样本质量的同时实现更快速的采样。DDIM 的前向过程与 DDPM 相同，而反向过程则被修改为

$$\boldsymbol{x}_{\tau_{i-1}} = \sqrt{\bar{\alpha}_{\tau_{i-1}}} \left(\frac{\boldsymbol{x}_{\tau_i} - \sqrt{1-\bar{\alpha}_{\tau_i}} \boldsymbol{\epsilon}_\theta(\boldsymbol{x}_{\tau_i}, \tau_i)}{\sqrt{\bar{\alpha}_{\tau_i}}} \right) + \sqrt{1-\bar{\alpha}_{\tau_{i-1}} - \sigma_{\tau_i}(\eta)^2} \boldsymbol{\epsilon}_\theta(\boldsymbol{x}_{\tau_i}, \tau_i) + \sigma_{\tau_i}(\eta) \boldsymbol{z}, \tag{11}$$

其中，时间步 $\{\tau_i\}$ 构成子序列 $S \subseteq \{1,.....T\}$，与 T 相比，其数量可以减少，从而加快采样。$\sigma_{\tau_i}(\eta)$ 是由超参数 $\eta \in [0,1]$控制的噪声尺度。通过设置 $\eta = 0$（即 $\sigma_{\tau_i} = 0$），采样过程会沿着固定轨迹进行，并变为确定性过程。初始噪声与生成样本之间的一一映射，加上采样效率，使 DDIM 特别适用于可追溯设计任务。因此，本研究采用 DDIM，并以 $\eta = 0$ 作为生成模型。

2.2 逆向设计方法

现有基于扩散模型的逆向设计方法通常会根据目标力学性能对反向去噪过程（5）进行条件约束。本研究提出了一种逆向设计框架，将无条件去噪扩散模型与可微编程技术相结合，以实现具有定制力学性能的微结构风格约束设计。该框架无需准备昂贵的标注数据集，也无需在考虑新的问题设定时重新训练模型。

图 1 展示了所提出的逆向设计框架的工作流程，该框架主要包含两个步骤：在图像数据集上训练无条件扩散模型，以及通过反向传播优化噪声输入以生成目标微结构。

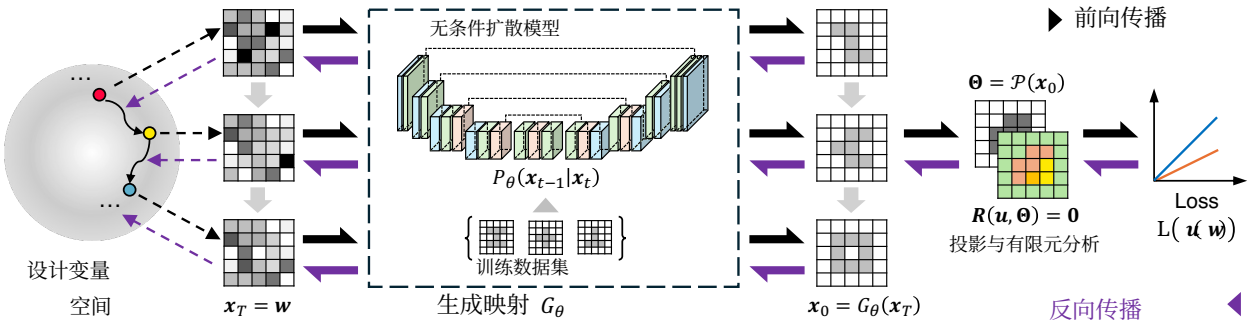

图1：利用无条件扩散模型对具有定制机械性能的微结构进行风格约束逆向设计框架的示意图。

第一步是在准备好的图像数据集上，按照第 θ 2.1节所述流程，使用神经网络参数训练一个无条件去噪扩散模型。训练好的扩散模型能够捕捉训练数据集中编码的几何特征和风格约束，并生成继承这些已学习特征的新样本。根据式（11）中的 $\eta = 0$DDIM 公式，我们得到一个确定性映射$\mathcal{G}_\theta$：$\mathbb{R}^n \to \mathbb{R}^n$ 该映射通过学习到的去噪过程，将噪声输入$\boldsymbol{x}_T \in \mathbb{R}^n$ 映射为微结构图像$\boldsymbol{x}_0 \in \mathbb{R}^n$，其中 n 表示图像维度。

接下来，我们将微结构设计任务表述为一个约束优化问题，并以噪声输入作为设计变量$\boldsymbol{w}$ $\in \mathbb{R}^n$。换言之，我们要寻找最优噪声输入$\boldsymbol{w}^*$，使生成的微结构呈现所需的力学性能。为了评估生成的微结构的力学响应，我们首先将图像$\boldsymbol{x}_0 = \mathcal{G}_\theta(\boldsymbol{w})$ 转换为有限元网格，并使用投影函数，根据像素强度值为每个单元赋予材料属性。然后，我们进行有限元分析（FEA），计算微结构在指定边界条件下的力学响应。接下来，通过损失函数 $\mathcal{L}$ 量化当前力学性能与目标力学性能之间的差异。设计目标是在满足机械系统平衡方程以及源自扩散模型生成的材料属性场的条件下，最小化 $\mathcal{L}$。由此得到以下优化问题：

$$\begin{aligned} &\min_{\boldsymbol{w}\in\mathbb{R}^n} \mathcal{L}\left(\boldsymbol{u}\left(\boldsymbol{w}\right)\right) \\ &\text{s.t.} \quad \boldsymbol{R}\left(\boldsymbol{u},\boldsymbol{\Theta}\right) = \boldsymbol{0}, \\ &\qquad \boldsymbol{\Theta} = \mathcal{P}\left(\boldsymbol{x}_0\right), \\ &\qquad \boldsymbol{x}_0 = \mathcal{G}_\theta(\boldsymbol{w}), \end{aligned} \tag{12}$$

其中 $\boldsymbol{u} \in \mathbb{R}^m$ 表示微结构的力学响应，$\boldsymbol{R}$： $\mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^m$ 表示

对机械系统控制方程进行离散化后得到的残差函数，$\boldsymbol{\Theta} \in \mathbb{R}^n$ 表示赋予微结构的材料属性场，而 $\mathcal{P}: \mathbb{R}^n \to \mathbb{R}^n$是一个投影函数，用于将生成图像 $\boldsymbol{x}_0 = \mathcal{G}_\theta(\boldsymbol{w})$ 映射至 $\boldsymbol{\Theta}$。该方法的核心在于重新诠释噪声输入：我们不再将其视为每次生成时可随意丢弃的初始状态，而是将其重新定义为一种可优化的设计变量，用于编码目标微观结构的几何特征。这一重新表述将典型的生成工作流程转变为由优化驱动的逆向设计过程。

为求解这一优化问题，梯度信息，即 $\mathrm{d}^{\mathcal{L}}\ \mathrm{d}w$ 至关重要。我们建立了统一微分流程，通过向量-雅可比积拼接实现，详见第2.3节，从而使力学响应 $\boldsymbol{u}$ 相对于材料属性场 $\boldsymbol{\Theta}$可微。整个框架由确定性扩散模型和有限元分析组成，其端到端可微性使我们能够通过自动微分计算梯度 [37]。具体而言，通过对整个框架进行反向传播 $\mathrm{d}^{\mathcal{L}}\ \mathrm{d}w$ 即可高效地获得，从而能够采用基于梯度的优化算法求解优化问题（12）。

2.3 Uni通过向量-雅可比积拼接实现的求导流程

本节将详细介绍如何通过统一微分流程计算损失函数 $\mathcal{L}$ 相对于设计变量 $\boldsymbol{w}$ 的梯度，即 $\mathrm{d}^{\mathcal{L}}\,\mathrm{d}w$。在自动微分 [37], 中，我们可以将前向计算过程分解为一系列初等运算，每个运算都表示为一个具有明确定义的输入和输出的函数。由此得到的计算图描述了这些运算之间的依赖关系，使我们能够通过链式法则系统地计算梯度。遵循这一原则，我们可以确定框架中的三个关键组成部分：生成映射 $\mathcal{G}_\theta$、投影函数 $\mathcal{P}$以及用于求解$\boldsymbol{u}$ 的力学求解器。应用链式法则后，损失函数 $\mathcal{L}$ 相对于设计变量 $\boldsymbol{w}$ 的总梯度可表示为

$$\frac{\mathrm{d}\mathcal{L}}{\mathrm{d}\boldsymbol{w}} = \frac{\partial \mathcal{L}}{\partial \boldsymbol{u}} \frac{\partial \boldsymbol{u}}{\partial \boldsymbol{\Theta}} \frac{\partial \boldsymbol{\Theta}}{\partial \boldsymbol{x}_0} \frac{\partial \boldsymbol{x}_0}{\partial \boldsymbol{w}}, \tag{13}$$

其中，每一项都不会显式计算为完整的雅可比矩阵，而是通过向量-雅可比积（VJP）高效地将梯度沿计算图反向传播。令$\boldsymbol{v}_{(*)} = \left[\frac{\partial \mathcal{L}}{\partial *}()\right]$ 表示中间变量$(*)$处传播的向量。然后，可以按顺序计算以下向量-雅可比积来执行反向传播：

$$\boldsymbol{v}_{\boldsymbol{u}}^{\top} = \frac{\partial \mathcal{L}}{\partial \boldsymbol{u}}, \tag{14a}$$

$$\boldsymbol{v}_{\boldsymbol{\Theta}}^{\top} = \boldsymbol{v}_{\boldsymbol{u}}^{\top} \frac{\partial \boldsymbol{u}}{\partial \boldsymbol{\Theta}}, \tag{14b}$$

$$\boldsymbol{v}_{\boldsymbol{x}_0}^{\top} = \boldsymbol{v}_{\boldsymbol{\Theta}}^{\top} \frac{\partial \boldsymbol{\Theta}}{\partial \boldsymbol{x}_0}, \tag{14c}$$

总梯度 $\mathrm{d}^{\mathcal{L}}\,\mathrm{d}w$ 可按如下方式求得：

$$\frac{\mathrm{d}\mathcal{L}}{\mathrm{d}\boldsymbol{w}} = \boldsymbol{v}_{\boldsymbol{x}_0}^{\top} \frac{\partial \boldsymbol{x}_0}{\partial \boldsymbol{w}}. \tag{15}$$

在这些向量-雅可比积组件中，(14a)、(14c) 和 (15) 可利用神经网络、投影函数和损失函数的可微性，通过标准反向传播算法轻松计算。关键在于计算式（$\boldsymbol{v}_{\boldsymbol{u}}^{\top}\,\partial_{\boldsymbol{u}}\,\partial\boldsymbol{\Theta}$ 14b）中的，它涉及力学响应 $\boldsymbol{u}$ 相对于材料属性场 $\boldsymbol{\Theta}$的灵敏度。这并不

简单，因为$\boldsymbol{u}$ 是由机械系统的平衡方程隐式定义的。

为求得 $\boldsymbol{v}_{\boldsymbol{u}}^{\top} \frac{\partial \boldsymbol{u}}{\partial \boldsymbol{\Theta}}$，我们采用隐式微分框架 [38]，通过伴随分析推导机械响应的基于伴随的向量-雅可比积规则$\boldsymbol{u}$，从而实现梯度通过机械求解器的高效传播 [39, 40]。我们首先对平衡方程$\boldsymbol{R}(\boldsymbol{u}, \boldsymbol{\Theta}) = \mathbf{0}$ 关于 $\boldsymbol{\Theta}$ 求导

$$\frac{\partial \boldsymbol{R}}{\partial \boldsymbol{u}} \frac{\partial \boldsymbol{u}}{\partial \boldsymbol{\Theta}} + \frac{\partial \boldsymbol{R}}{\partial \boldsymbol{\Theta}} = \mathbf{0}. \tag{16}$$

对上述方程重新整理可得

$$\frac{\partial \boldsymbol{u}}{\partial \boldsymbol{\Theta}} = -\left(\frac{\partial \boldsymbol{R}}{\partial \boldsymbol{u}}\right)^{-1} \frac{\partial \boldsymbol{R}}{\partial \boldsymbol{\Theta}}, \tag{17}$$

这意味着，隐式导数 $\frac{\partial \boldsymbol{u}}{\partial \boldsymbol{\Theta}}$ 可通过对右侧显式导数进行乘法运算来计算。然后，我们将 $\boldsymbol{v}_{\boldsymbol{u}}^{\top}$乘以式（17）两边，得到

$$\boldsymbol{v}_{\boldsymbol{u}}^{\top} \frac{\partial \boldsymbol{u}}{\partial \boldsymbol{\Theta}} = -\boldsymbol{v}_{\boldsymbol{u}}^{\top} \left(\frac{\partial \boldsymbol{R}}{\partial \boldsymbol{u}}\right)^{-1} \frac{\partial \boldsymbol{R}}{\partial \boldsymbol{\Theta}}. \tag{18}$$

为避免显式计算 $\frac{\partial \boldsymbol{R}}{\partial \boldsymbol{u}}$ 的逆矩阵，我们引入伴随变量 $\boldsymbol{\lambda}_{\boldsymbol{u}} \in \mathbb{R}^{m}$，通过求解以下伴随方程得到该变量：

$$\left(\frac{\partial \boldsymbol{R}}{\partial \boldsymbol{u}}\right)^{\top} \boldsymbol{\lambda}_{\boldsymbol{u}} = -\boldsymbol{v}_{\boldsymbol{u}}. \tag{19}$$

将 $\boldsymbol{\lambda}_{\boldsymbol{u}}$ 代入式（18）后，可将机械求解器的定制向量-雅可比积规则定义为

$$\boldsymbol{v}_{\boldsymbol{u}}^{\top} \frac{\partial \boldsymbol{u}}{\partial \boldsymbol{\Theta}} = \boldsymbol{\lambda}_{\boldsymbol{u}}^{\top} \frac{\partial \boldsymbol{R}}{\partial \boldsymbol{\Theta}}, \tag{20}$$

该规则仅涉及显式导数的计算 $\frac{\partial \boldsymbol{R}}{\partial \boldsymbol{\Theta}}$ 以及线性系统的求解（19）。定制向量-雅可比积规则（20）不依赖残差函数 $\boldsymbol{R}$的具体形式，因此适用于多种力学系统。将该定制向量-雅可比积规则与其他向量-雅可比积组件的标准微分规则串联起来，我们建立了统一微分流程，从而能够在整个逆向设计框架中高效计算梯度。算法基于 `JAX` [41] 生态系统实现，其中扩散模型使用 `Flax` [42]，有限元分析使用 `JAX-FEM` [43]。

2.4 概念示例

为说明所提出的方法，我们考虑一个涉及一维高斯混合模型（GMM）的概念示例。GMM 的概率密度函数定义如下：

$$p(x) = \sum_{i=1}^{n} \pi_i \mathcal{N}(x; \mu_i, \sigma_i^2), \tag{21}$$

其中，n 表示高斯分量的数量，π_i 表示第i 个分量的混合权重，满足 $\sum_{i=1}^{n} \pi_i = 1$；$\mathcal{N}(x; \mu_i, \sigma_i^2)$ 是均值为 μ_i、方差为 σ_i^2的高斯分布。在此示例中，我们构造了一个包含两个分量的混合分布，其中$n = 2$，参数为$\pi = [0.5、0.5]$、$\mu = [-2.5、2.5]$, 和 $\sigma = [0.5、0.5]$,，由此得到一个对称双峰分布。

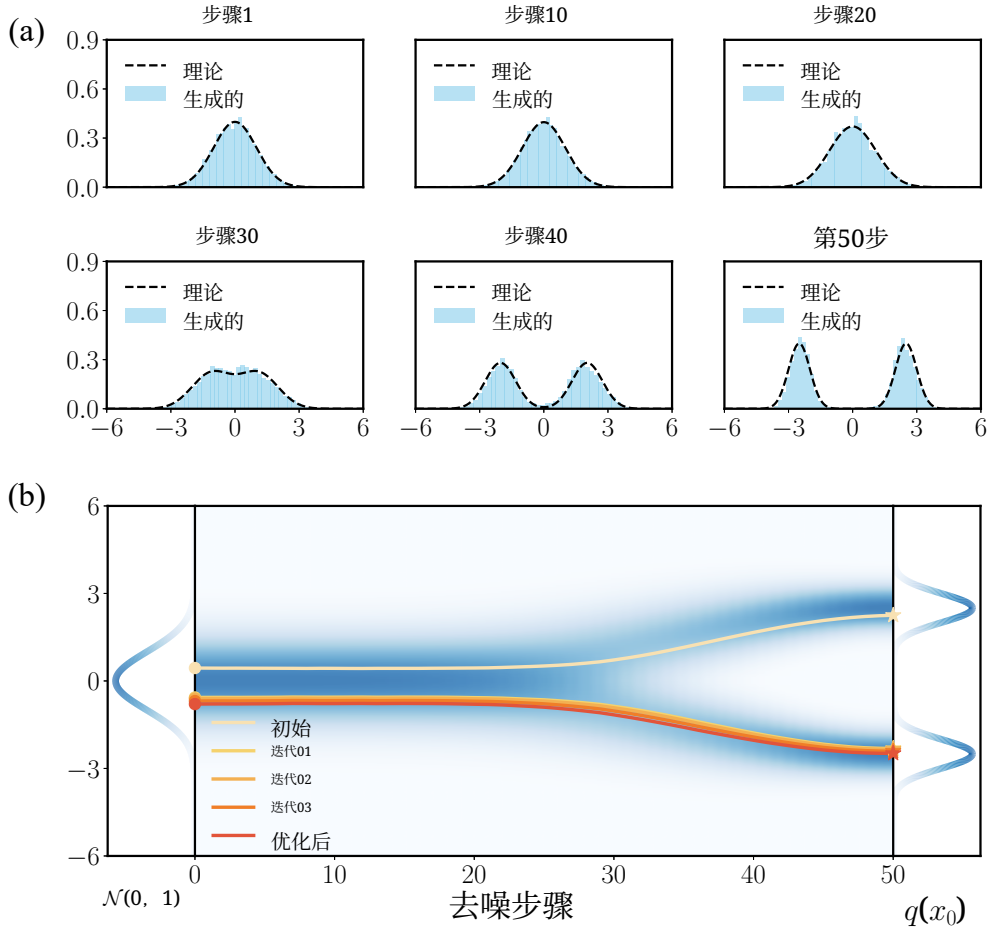

图2：概念性高斯混合示例的设计结果。(a) 训练后的扩散模型的反向去噪过程。(b) 优化过程中将样本从正模态转换到负模态时的采样轨迹。

首先，我们将此高斯混合模型用作目标数据分布$q(x_0)$，并抽取 20000 个样本来训练一个一维无条件去噪扩散模型。去噪神经网络由三个全连接层组成，隐藏维度为 64，并使用 Swish激活函数；时间步通过正弦位置嵌入进行编码，并通过加性连接整合。正向扩散过程采用$T = 1000$ 个时间步，方差调度为线性 β_t，范围从 10^{-4} 到 0.02。模型使用 Adam优化器训练 100 个训练轮数，学习率为 10^{-4}，批量大小为 128。反向去噪过程遵循公式（11），其中$\eta = 0$，并使用 50 个时间步。

图2 (a) 展示了训练后的扩散模型的反向采样过程，其中不同时间步的去噪样本 $\boldsymbol{x}_t$ 以浅蓝色直方图表示，而由前向过程推导出的理论分布则以黑色虚线表示。可以观察到，每个时间步的去噪样本都与理论分布高度吻合，样本分布逐渐从标准正态分布过渡到所需的双峰分布，验证了扩散模型的有效性。

接下来，我们考虑一个简单的设计任务：寻找最优噪声输入，以引导生成的样本趋向目标值 x^{obj}，其损失函数定义为

$$\mathcal{L}(\boldsymbol{w}) = \left[\boldsymbol{x}_0(\boldsymbol{w}) - x_{\text{obj}}\right]^2. \tag{22}$$

具体而言，我们使用基于梯度的 BFGS 算法优化噪声输入，使样本从正模态（接近 2.5）转移到负模态（接近 -20.5） [44]。图2(b) 展示了

优化迭代过程中的反向去噪过程，不同颜色的线条代表不同的迭代次数。左侧面板显示标准正态分布，右侧面板显示双峰高斯混合模型分布，中间面板展示了跨越所有时间步的去噪轨迹。背景色图表示每个去噪步骤的理论概率分布，颜色越深的蓝色表示密度越高的区域。

我们观察到，所有轨迹在每个去噪步骤中都会经过高概率区域，进一步验证了训练所得扩散模型的有效性。优化算法会在迭代过程中逐步修改噪声输入。到第一次迭代时，轨迹已经开始朝目标模态弯曲；到第四次迭代时，优化后的噪声输入成功引导样本精确达到负模态处的目标值 -2.5，且损失低于 10^{-3}。这一概念示例展示了我们方法的核心机制：通过优化噪声输入，我们能够控制反向去噪轨迹，引导生成的样本趋近所需目标。

3 结果与讨论

本节展示将所提出的逆向设计框架应用于各种微结构设计任务所得到的结果。我们首先介绍数据集和去噪神经网络，以及示例中使用的投影和优化设置。随后，我们展示所提出的方法在设计具有定制化均匀化性能、超弹性性能和弹塑性性能的微结构方面的有效性。最后，我们讨论所提出方法的优势与局限性，以及未来可能的研究方向。

3.1 准备

3.1.1 数据集与神经网络

本研究采用两种具有代表性的图像数据来训练扩散模型：MNIST 数据集中的手写数字灰度图像 [45], 以及二维正交各向异性超材料微结构的二值图像 [46]。MNIST 数据集包含 60000 张手写数字（0-9）的训练图像，分辨率为 28×28 像素，在机器学习领域被广泛用于图像生成和分类任务的基准测试。超材料数据集包含 90245 个正交各向异性微结构二值图像，分辨率为 50×50 像素。为便于优化过程，我们先向原始二值微结构添加均值为零、标准差为 0.01 的高斯噪声，再应用标准差为 1.2 的高斯滤波器对含噪图像进行平滑处理，从而将其转换为灰度图像。第 3.2 节将说明，引入中间灰度相既能避免优化过程中梯度信息的损失，又能保持微结构的整体几何特征。两个数据集均重新缩放至 $[-1, 1]$ 范围，并调整至 32×32像素，以用于扩散模型的训练。

去噪扩散神经网络采用 U-Net 架构 [47] ，具有跨越四个分辨率层级（32×32、 16×16、 8×8 和 4×4）的层次化编码器-解码器结构。编码器和解码器路径中均包含残差块 [48] 。在中间分辨率层级应用多头自注意力机制 [49], 以捕捉空间依赖关系和全局结构模式，使模型能够学习非局部相关性。时间步信息通过正弦位置嵌入进行编码，随后由具有 GELU 激活函数的两层前馈网络投影，并融入每个残差块中。正向扩散过程采用 $T = 1000$ 个时间步，使用线性方差调度 β_t , 范围从 10^{-4} 到 0.02。模型在上述数据集上训练 100 轮，批量大小为 128，使用 AdamW优化器 [50]和预热-余弦学习率调度

。反向去噪过程遵循公式（11），并采用 $\eta = 0$ 进行确定性生成，为提高效率使用 10 个时间步。

3.1.2 投影与优化

训练好的扩散模型可以接收随机噪声输入，并生成具有与训练数据集相似风格特征的灰度图像。每幅图像都被解释为单位正方形 $(0,1) \times (0,1)$上的一个 32×32 有限元网格，其中每个像素对应一个单元。随后，通过线性插值方案为每个单元赋予材料属性

$$\Theta(\tilde{\boldsymbol{x}}) = \Theta_0 + (\Theta_1 - \Theta_0)\tilde{\boldsymbol{x}}, \tag{23}$$

其中，Θ_0 和 Θ_1 分别是两种基体材料的属性。$\tilde{\boldsymbol{x}}$ 表示通过投影函数从生成图像 $\boldsymbol{x}_0$ 得到的材料密度场。具体而言，对于 MNIST 数据集，我们设定 $\tilde{\boldsymbol{x}} = 1 - \mathcal{P}(\boldsymbol{x}_0)$；对于超材料数据集，则设定 $\tilde{\boldsymbol{x}} = \mathcal{P}(\boldsymbol{x}_0)$。投影函数 $\mathcal{P}(\boldsymbol{x})$定义如下

$$\mathcal{P}(\boldsymbol{x}) = \frac{1}{2}\left[\tanh(\gamma \boldsymbol{x}) + 1\right], \tag{24}$$

其中，γ 是一个控制投影锐度的超参数。较小的 γ 会使材料相之间平滑过渡，而较大的 γ 则会产生更接近二元分布的结果。优化采用基于梯度的 BFGS 算法 [44]进行。所有计算均在配备 36 GB 内存的 Apple M3 Pro 平台上完成。

尽管中间材料相的存在有助于基于梯度的优化，但最终设计应为近似二值分布，以便实际制造。为解决这一问题，我们采用多阶段优化策略：优化从较小的 γ 开始，以实现平滑的材料过渡；随后，将最优设计作为下一阶段的初始猜测，并增大 γ。重复这一过程，直至达到足够大的 γ，从而得到近似二值的材料分布。为评估生成的微结构的二值质量，我们定义如下二值化指标：

$$\mathcal{B}(\tau) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{I}(\tilde{x}_i \le \tau \vee \tilde{x}_i \ge 1 - \tau), \tag{25}$$

其中，$\mathbb{I}(\cdot)$ 为示性函数，n 为像素总数，$\tilde{x}_i$ 为像素i的投影密度，τ 为一个较小的阈值，在本研究中取值为 10^{-3}。$\mathcal{B}$ 的值越高，表示设计的二值化程度越高。正如我们将在后续示例中展示的那样，这种多阶段策略能够有效引导优化，使其获得具有所需力学性能且可制造的设计。

3.2 定制化均匀化性能

在第一个示例中，我们将所提出的方法应用于具有定制均匀化性能的微结构设计，为先进复合材料提供基础构件。生成的微结构被建模为一个代表性体积单元（RVE）Y，并在宏观域 Ω上施加周期性边界条件，如图 3所示。通过向 RVE 施加一组预设的宏观应变载荷工况$\bar{\boldsymbol{\varepsilon}}$ 并求解以下边值问题，完成计算均匀化：

$$\begin{aligned}
-\nabla \cdot \boldsymbol{\sigma} &= \mathbf{0} && \text{in } Y, && (26a)\\
\boldsymbol{\sigma}(\boldsymbol{y}) &= \mathbb{C}(\boldsymbol{y}) : \boldsymbol{\varepsilon}(\boldsymbol{y}) && \text{for } \boldsymbol{y} \in Y, && (26b)\\
\boldsymbol{\varepsilon}(\boldsymbol{y}) &= \bar{\boldsymbol{\varepsilon}} + \nabla^s \tilde{\boldsymbol{u}} && \text{in } Y, && (26c)\\
\tilde{\boldsymbol{u}} &&& \text{is } Y\text{-periodic}, && (26d)\\
\boldsymbol{\sigma} \cdot \boldsymbol{n} &= \boldsymbol{t} && \text{is } Y\text{-antiperiodic on } \partial Y, && (26e)
\end{aligned}$$

其中，$\boldsymbol{y}$ 表示微观坐标，$\boldsymbol{\sigma}$ 表示微观柯西应力，$\boldsymbol{\varepsilon}$ 表示微观应变，$\mathbb{C}$ 表示根据生成的微结构赋予的局部弹性张量，$\tilde{\boldsymbol{u}}$ 表示待求解的周期性扰动位移场，$\boldsymbol{t}$ 表示 RVE 边界∂Y上的牵引向量，$\boldsymbol{n}$ 表示 ∂Y 上的外法向量。宏观应力 $\bar{\boldsymbol{\sigma}}$ 通过对 RVE 上的 $\boldsymbol{\sigma}$ 进行体积平均计算得到

$$\bar{\boldsymbol{\sigma}} = \frac{1}{|Y|}\int_{Y} \boldsymbol{\sigma}(\boldsymbol{y})\,\mathrm{d}\boldsymbol{y}, \tag{27}$$

其中，$|Y|$ 是代表性体积单元的体积。对代表性体积单元施加三个独立的宏观应变载荷工况，包括沿x和y方向的单轴拉伸以及纯剪切，以计算相应的宏观应力，并可根据宏观应力—应变关系推导出有效弹性张量 $\mathbb{C}^{\mathrm{hom}}$

$$\bar{\boldsymbol{\sigma}} = \mathbb{C}^{\mathrm{hom}} : \bar{\boldsymbol{\varepsilon}}. \tag{28}$$

有效弹性张量 $\mathbb{C}^{\mathrm{hom}}$ 表征了复合材料在宏观尺度上的均匀化弹性响应。因此，通过微结构设计实现目标 $\mathbb{C}^{\mathrm{obj}}$ 的能力对于先进材料应用非常重要。在此示例中，我们将生成的有效弹性张量与目标有效弹性张量各分量之间的平方误差设为损失函数

$$\mathcal{L}(\boldsymbol{w}) = \sum_{i,j=1}^{3}\left[\mathbb{C}_{ij}^{\mathrm{hom}}(\boldsymbol{w}) - \mathbb{C}_{ij}^{\mathrm{obj}}\right]^2. \tag{29}$$

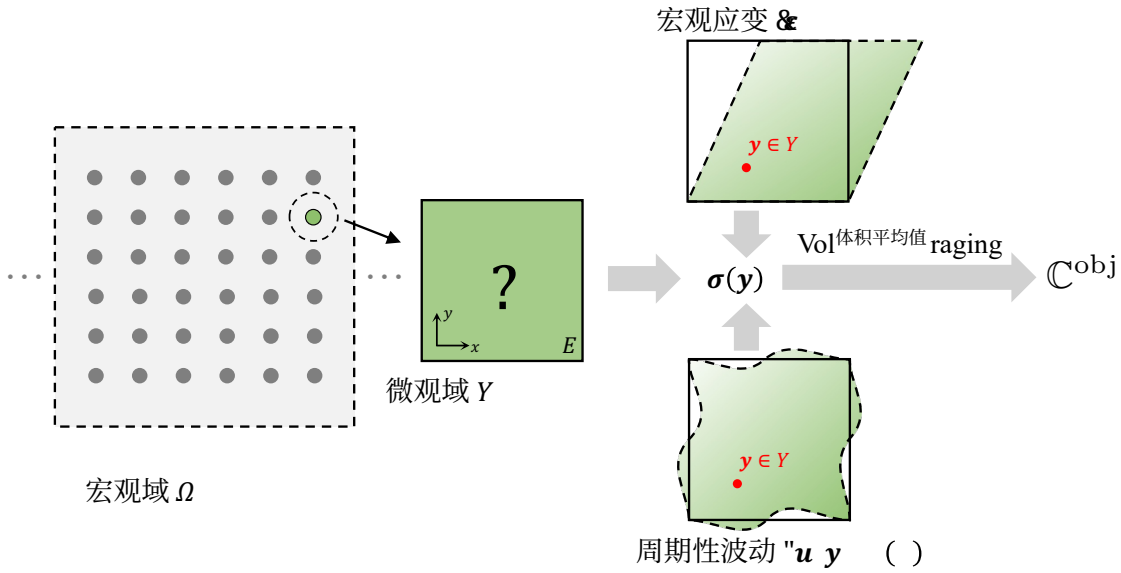

图3：具有定制均匀化性能的微结构设计示例的模型设置示意图。

在微结构设计中，我们考虑两种不同的基材：杨氏模量为 $E_0 = 1$ 的软材料，以及杨氏模量为 $E_1 = 100$ 的硬材料，两者的泊松比均为 $\nu = 0.3$。首先，我们使用基于 MNIST 的扩散模型执行微结构设计任务。我们考虑三组目标 $\mathbb{C}^{\mathrm{hom}}$，其对称性各不相同 [51]，包括双斜晶系、正交各向异性和四方晶系。每组的目标有效弹性张量

设定如下：

$$\mathbb{C}_{\text{bic}}^{\text{obj}} = \begin{bmatrix} 50 & 12 & 0 \\ 12 & 60 & -3 \\ 0 & -3 & 15 \end{bmatrix}, \quad \mathbb{C}_{\text{ort}}^{\text{obj}} = \begin{bmatrix} 50 & 12 & 0 \\ 12 & 60 & 0 \\ 0 & 0 & 15 \end{bmatrix}, \quad \mathbb{C}_{\text{tet}}^{\text{obj}} = \begin{bmatrix} 60 & 12 & 0 \\ 12 & 60 & 0 \\ 0 & 0 & 15 \end{bmatrix}. \tag{30}$$

每组以两种不同的噪声输入进行初始化，并采用多阶段策略进行优化，γ 取值为 5、10、20 和 80。当损失小于 10^{-2} 或达到最大迭代次数（100 次）时，即终止该阶段。

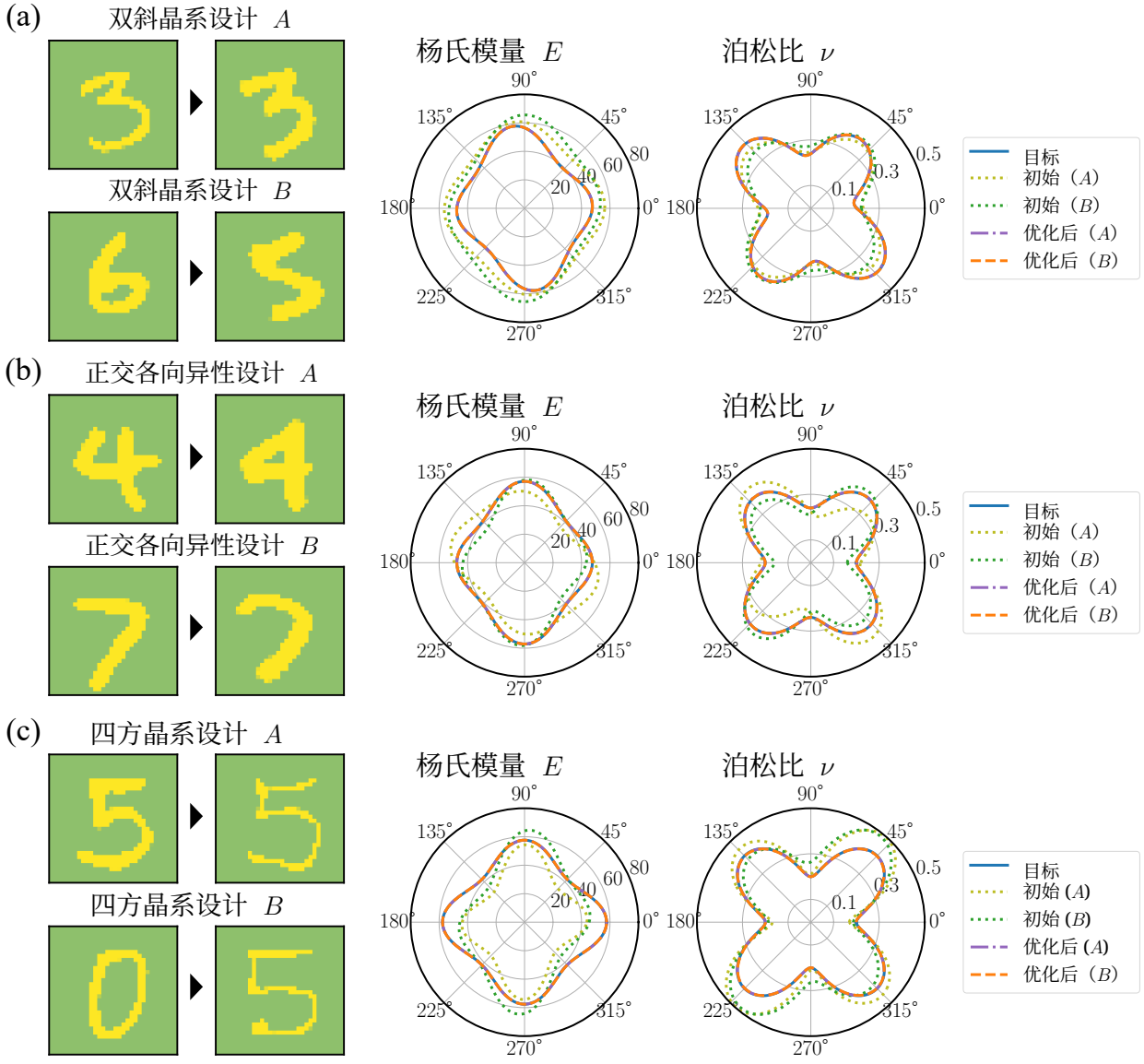

图 4：基于 MNIST 的微结构设计结果，用于实现具有不同对称性的目标有效弹性张量 [51]：(a) 双斜晶系，(b) 正交各向异性，(c) 四方晶系。

图 4展示了各组的设计结果，其中给出了初始和优化后的微结构，以及它们在不同方向上的杨氏模量和泊松比。对于大多数情况，优化后的微结构保留了初始猜测的整体形状，同时调整局部几何特征以满足目标属性。我们还观察到，与初始

猜测相比，一些优化后的微结构转变为不同的拓扑结构，但仍保持了与手写数字相似的几何特征。例如，在单斜 B 群中，优化后的微结构从类似数字 6 的形状转变为类似数字 5 的形状。尽管优化后的设计呈现出多样的几何特征，但它们都实现了所需的目标属性，证明了所提出方法能够在遵循所学习数据分布的同时，引导设计朝指定目标发展。所有优化后微结构的二值化指标 $\mathcal{B}$ 均约为 99%，表明多阶段优化策略能够有效生成可制造的设计。每个设计案例的运行时间如图 5所示，范围为 509 秒至 1402 秒。在优化的四个阶段中，第一阶段和最后阶段通常耗时更长，这是由于需要从初始猜测进行探索，以及在接近二值化且具有较大 γ时梯度信息的有效性较低。

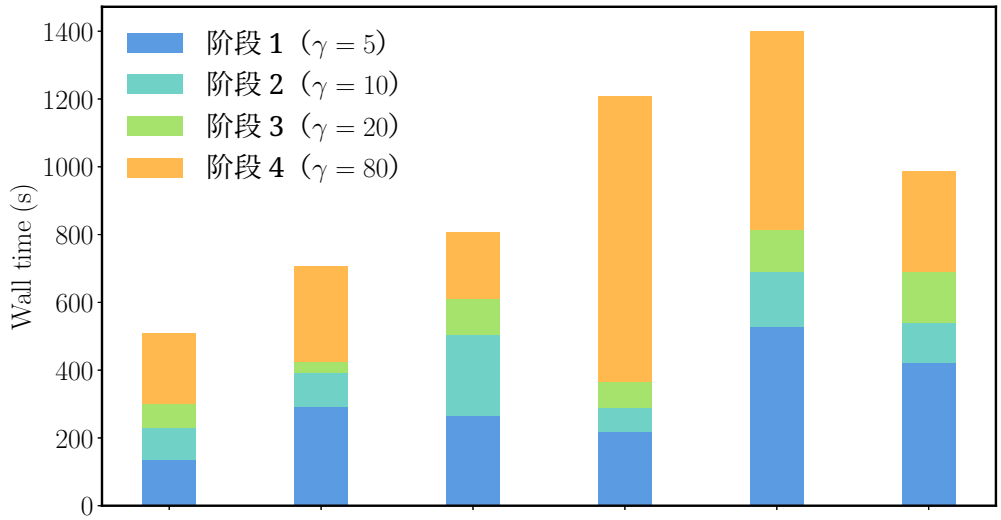

图 5：针对具有不同对称性的有效弹性张量，对基于 MNIST 的微结构进行逆向设计所需的运行时间。

在 MNIST 数据集设计示例之后，本文进一步使用包含二维正交各向异性超材料微结构的开放数据集 [46]，检验所提出的方法执行逆向设计任务的效果。基体材料和优化设置与基于 MNIST 的示例保持一致，其中包含四个优化阶段，γ 值分别为 5、10、20 和 100。目标有效弹性张量定义如下：

$$\mathbb{C}_{\text{ort}}^{\text{obj}} = \begin{bmatrix} 50 & 15 & 0 \\ 15 & 70 & 0 \\ 0 & 0 & 15 \end{bmatrix}, \qquad \mathbb{C}_{\text{tet}}^{\text{obj}} = \begin{bmatrix} 50 & 15 & 0 \\ 15 & 50 & 0 \\ 0 & 0 & 15 \end{bmatrix}. \tag{31}$$

每个对称群使用两个不同的噪声输入。

图6 展示了所考虑的正交各向异性和四方晶系对称性下的设计结果，包括初始微结构和优化后的微结构、5×5 优化后微结构的晶胞，以及对应的杨氏模量和泊松比极坐标图。所有案例的损失均低于 10^{-2}，二值化指标均高于 98%，表明所提出的方法能够有效生成具有所需均质化性能且可制造的微结构。此外，尽管训练图像具有严格的对称性，灰度预处理、扩散模型的本质特性以及优化过程仍可能引入轻微的不对称性

出现在最终设计中。这些微小偏差不会显著影响整体机械性能。对于正交各向异性和四方晶系对称性，可以选取生成图像的四分之一，应用对称操作以重构完整微结构，然后按照相同流程获得严格对称的设计。我们将在第3.5节进一步讨论这一问题。

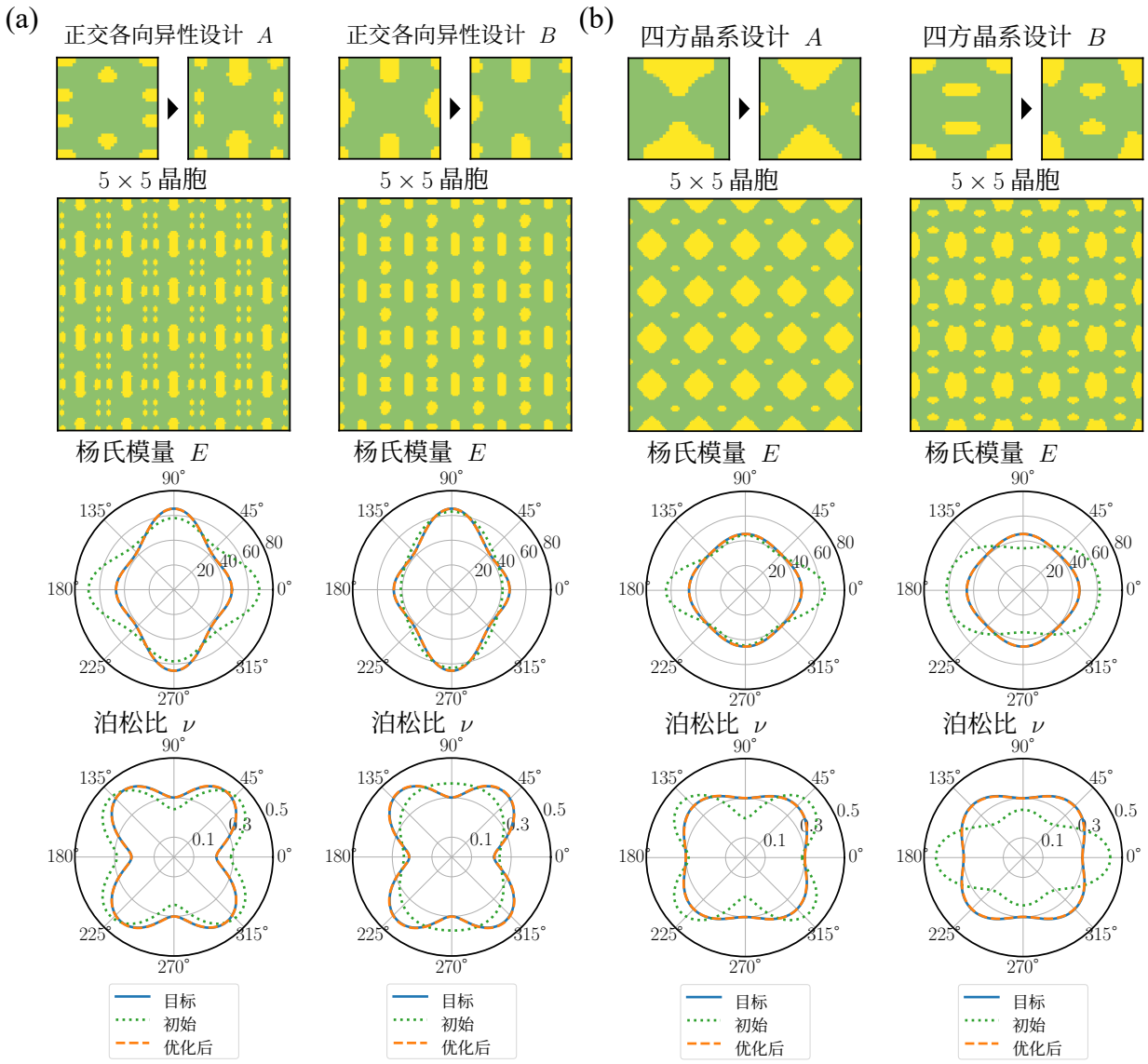

图 6：为实现具有不同对称性的目标有效弹性张量而设计的基于超材料的微结构结果：(a) 正交各向异性和 (b) 四方晶系。

如第3.1.1节所述，我们将原始二值微结构图像预处理为灰度图像，以便进行基于梯度的优化。为说明其中的原因，我们对基于灰度的扩散模型和基于二值的扩散模型进行了比较研究。具体而言，我们采用与第3.1.1 节相同的设置，使用原始二值数据训练另一个扩散模型。图7 展示了基于灰度的扩散模型和基于二值的扩散模型生成结果的比较，包括三个 γ 值（1、5 和 10）对应的投影密度场，以及分量C_{11}^{hom}相对于噪声输入的梯度。我们观察到，基于二值的扩散模型生成的密度场几乎为 0-1，且梯度幅值随着 γ 增大而逐渐减小。另一方面，基于灰度的模型生成的密度场相对平滑，包含中间值，并且仍能保持

显著的梯度幅值，即使 γ 较大也是如此。这是因为基于二值的扩散模型倾向于生成接近训练数据边界（-1 和 1）的像素值，当投影函数在 γ 较大时变得陡峭，便会导致梯度消失。相比之下，基于灰度的模型能够生成中间像素值，因此即使投影函数很陡峭，也能产生非零梯度。此外，生成的灰度微结构仍能捕捉原始二值设计的整体几何特征，并且通过多阶段策略，最终优化结果可以达到较高的二值化指标。因此，本研究采用基于灰度的扩散模型执行基于梯度的优化任务。

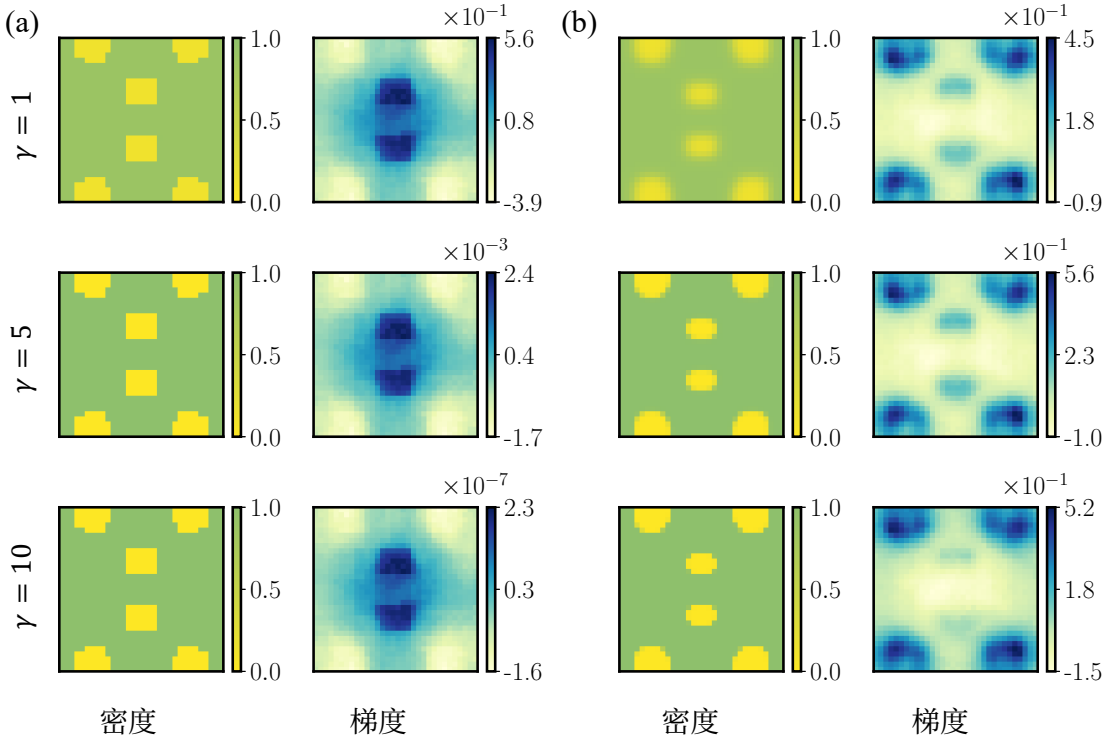

图7：不同斜率 γ 下的比较。对于（a）基于二值化的扩散模型和（b）基于灰度的扩散模型，投影密度场（左列）以及 $\mathbb{C}_{11}^{\text{hom}}$ 相对于噪声输入的梯度（右列）。

3.3 定制超弹性性能

除了实现特定的均匀化性能外，控制微结构的力学响应对于获得所需的功能性能也至关重要。具体而言，定制微结构的超弹性行为对于生物医学器件 [1]和软体机器人 [2]等应用至关重要。本节将讨论具有定制超弹性行为的微结构设计。超弹性体 Ω 变形的控制方程为

$$-\nabla \cdot \boldsymbol{P} = \boldsymbol{0} \quad \text{in } \Omega, \tag{32a}$$

$$\boldsymbol{u} = \bar{\boldsymbol{u}} \quad \text{on } \Gamma_D, \tag{32b}$$

$$\boldsymbol{P} \cdot \boldsymbol{N} = \bar{\boldsymbol{t}} \quad \text{on } \Gamma_N, \tag{32c}$$

其中，$\boldsymbol{u}$ 是待求解的位移场。$\boldsymbol{P} = {}^{\partial W}\!/_{\partial \boldsymbol{F}}$ 是第一类皮奥拉-基尔霍夫应力，W 是应变能密度函数，而$\boldsymbol{F} = \mathbf{I} + \nabla \boldsymbol{u}$ 是变形梯度。 $\bar{}\boldsymbol{u}$ 是规定位移

在狄利克雷边界 Γ_D 上，$\bar{\boldsymbol{t}}$ 是诺伊曼边界 Γ_N 上具有单位外法向$\boldsymbol{N}$的规定牵引力。基体材料采用 Neo-Hookean 超弹性模型，其W 定义为

$$W = \frac{G}{2}(J^{-2/3}I_1 - 3) + \frac{K}{2}(J-1)^2 \tag{33}$$

其中，$G = \frac{E}{2(1+\nu)}$ 为剪切模量，$K = \frac{E}{3(1-2\nu)}$ 为体积模量，E 为杨氏模量，ν 为泊松比，$I_1 = \mathrm{tr}(\boldsymbol{C})$ 为右柯西-格林变形张量$\boldsymbol{C} = \boldsymbol{F}^\top\boldsymbol{F}$的第一不变量，而$J = \det(\boldsymbol{F})$ 为$\boldsymbol{F}$的行列式。

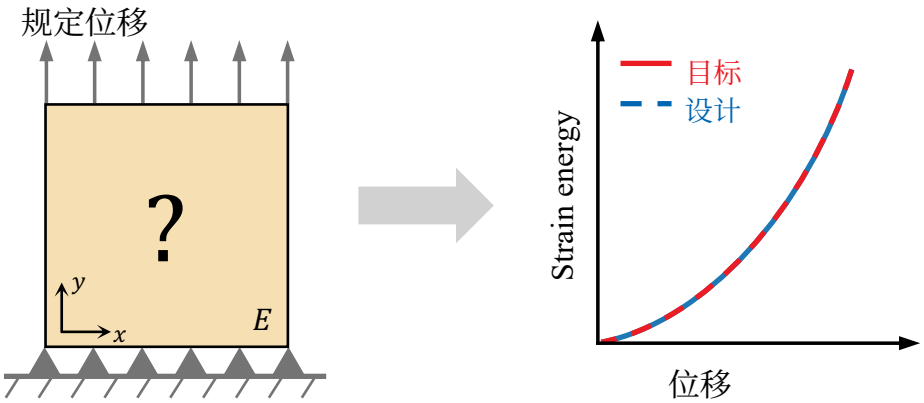

图 8：具有定制超弹性性能的微结构设计示例的模型设置示意图。

如图 8所示，我们考虑设计在单轴拉伸下具有定制应变能曲线的微结构。两种基材分别为杨氏模量为 $E_1 = 100$的硬材料和杨氏模量为 $E_0 = 1$的软材料，两者的泊松比均为 $\nu = 0.3$。计算域的上边界施加从 0.1 到 0.5 的均匀位移载荷，下边界固定。为表示目标超弹性行为，我们首先使用三次多项式近似由完整基材构成的微结构的应变能曲线函数：$P_1 = -19.880x^3 + 51.245x^2 + 0.472x$ 和$P_0 = -0.199x^3 + 0.512x^2 + 0.005x$，分别对应硬基材和软基材。随后，我们将目标应变能曲线定义为两种基材应变能曲线之间的线性插值

$$P^* = (1-\alpha)P_1 + \alpha P_0, \tag{34}$$

其中，$\alpha \in [0,1]$ 是控制所需超弹性行为的插值参数。损失函数定义为所有加载步中计算得到的应变能与目标应变能之间的均方误差。

$$\mathcal{L}(\boldsymbol{w}) = \frac{1}{N}\sum_{i=1}^{N}\left[P_i(\boldsymbol{w}) - P^*(u_i)\right]^2 \tag{35}$$

其中，N 是加载步总数，P_i 是生成的微结构在施加第i个位移u_i时的应变能，而 $P^*(u_i)$ 是对应的目标值。多阶段优化采用 γ 取值为2、5、10、20和100。每个阶段在损失小于 10^{-3} 或达到最大迭代次数（30次）时终止。

图 9 展示了具有相同应变能曲线 $\alpha = 0.3$ 的超弹性微结构设计结果，以及不同的初始设计，包括数字 3、4 和 5。初始微结构和优化后的微结构对应的应变能曲线与目标曲线的对比结果显示在

下方。尽管初始微结构表现出显著不同的超弹性行为，所提出的方法仍能有效地对其进行优化，使其达到所需的目标属性，同时保留手写数字风格。从数字 3 开始的优化后微结构保留了数字 3 的整体形状，但缩小了尺寸，以满足更高的刚度要求。相比之下，从数字 4 和 5 开始的优化分别收敛为类似数字 9 和 3 的不同形状，同时仍实现了所需的超弹性响应。所有优化后的微结构的二值化指标均超过 99% $\mathcal{B}$，表明多阶段优化策略成功生成了可制造的设计。

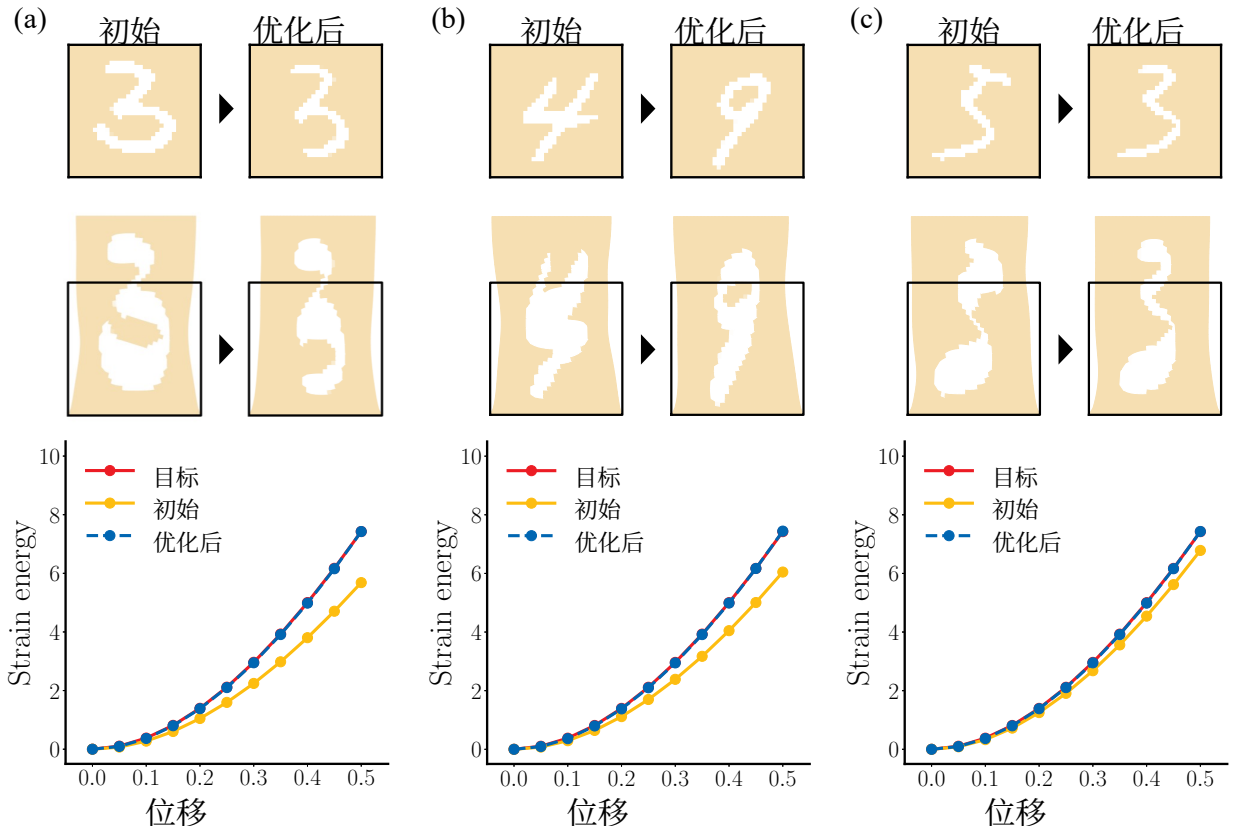

图9：具有相同应变能曲线、但初始设计不同的超弹性微结构设计结果：(a) 数字3；(b) 数字4；(c) 数字5。从上到下依次为：未变形构型、最大位移载荷作用下的变形构型，以及相应的应变能曲线。

接下来，我们研究在初始设计相同的情况下，不同目标超弹性行为对优化后微结构的影响。图 10 展示了初始设计和优化后的微结构，以及相应的应变能曲线，其中插值参数 α 的取值范围为 0.1至0.6。结果表明，所提出的方法能够生成与多种目标超弹性行为高度匹配的微结构。所有优化后的微结构的二值化指标均高于99%，这表明多阶段优化策略行之有效。

此外，我们可以观察到，优化后的微结构会逐渐过渡，以适应不断变化的超弹性响应。当 α 较小时，应变能曲线更接近硬基体材料的应变能曲线，因此会形成类似数字1和5的纤细形状。随着 α 增大，优化后的微结构会纳入更多软材料区域，以实现更低的刚度，并逐渐呈现类似数字9的形状，最终变为数字0。当 α 大于0.6时，优化后的微结构往往会趋于与 $\alpha = 0.6$相似的设计，但包含更多软材料区域。同时，优化后的噪声输入往往会偏离各向同性高斯分布。这些观察结果表明，在保留扩散模型所编码的清晰几何特征的同时，所选基材属性在实现极软超弹性性能方面存在局限。可以通过选择其他基材组合来解决这一问题，从而实现所需的性能和风格。这一现象将在第3.5节进一步讨论。

这些观察结果表明，在保留扩散模型所编码的清晰几何特征的同时，所选基材属性在实现极软超弹性性能方面存在局限。可以通过选择其他基材组合来解决这一问题，从而实现所需的性能和风格。这一现象将在第3.5节进一步讨论。

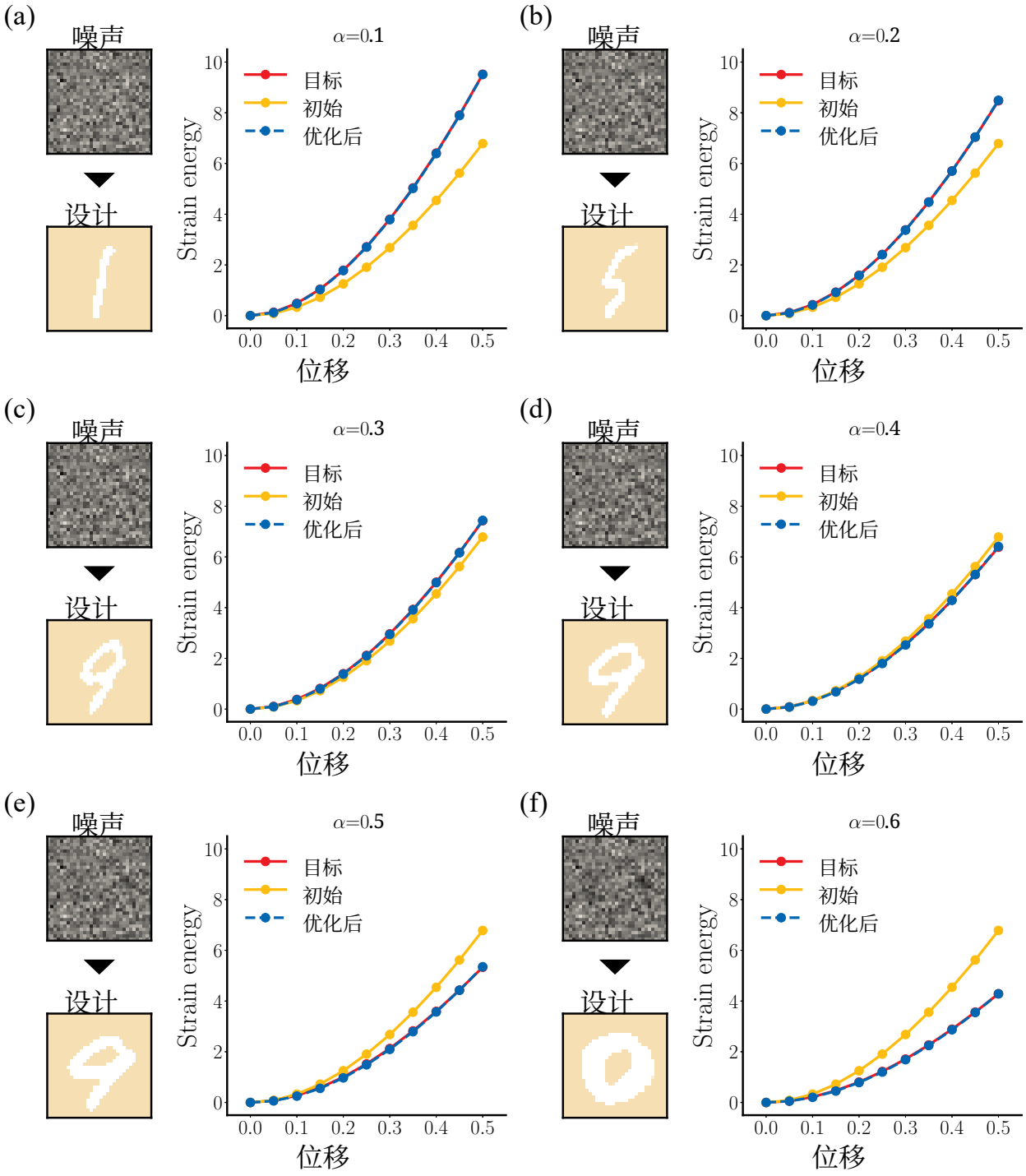

图10：具有相同初始设计、不同目标应变能曲线的超弹性微结构设计结果，这些曲线由 α 值表征。(a) 至(f) 分别对应 $\alpha = 0.1$、0.2、0.3、0.4、0.5 和 0.6。

3.4 定制弹塑性性能

除了超弹性响应外，弹塑性行为也是工程应用中常见的另一类重要非线性材料响应。本节探讨具有定制弹塑性性能的微结构逆向设计。材料弹塑性变形的控制方程为

$$-\nabla \cdot \boldsymbol{\sigma}^k = \mathbf{0} \quad \text{in } \Omega, \tag{36a}$$

$$\boldsymbol{u}^k = \bar{\boldsymbol{u}} \quad \text{on } \Gamma_D, \tag{36b}$$

$$\boldsymbol{\sigma}^k \cdot \boldsymbol{n} = \bar{\boldsymbol{t}} \quad \text{on } \Gamma_N, \tag{36c}$$

其中 $\boldsymbol{\sigma}^k = \boldsymbol{f}(\boldsymbol{u}^k, \boldsymbol{\varepsilon}^{k-1}, \boldsymbol{\sigma}^{k-1})$ 表示时间步 k处更新后的柯西应力，由应力更新函数 $\boldsymbol{f}$ 计算得出，该函数取决于当前位移 $\boldsymbol{u}^k$ 以及上一时间步的总应变 $\boldsymbol{\varepsilon}^{k-1}$ 和应力 $\boldsymbol{\sigma}^{k-1}$。边界条件相关符号与前文相同。此处我们采用小应变假设下的完美J2塑性模型 [52] 来描述基体材料的弹塑性行为。应力更新函数$\boldsymbol{f}$定义如下：

$$\boldsymbol{f}(\boldsymbol{u}^k, \boldsymbol{\varepsilon}^{k-1}, \boldsymbol{\sigma}^{k-1}) = \boldsymbol{\sigma}_{\text{trial}} - \frac{\boldsymbol{s}_{\text{trial}}}{s_{\text{trial}}} \langle f_{\text{yield}} \rangle_+, \tag{37a}$$

$$\boldsymbol{\sigma}_{\text{trial}} = \boldsymbol{\sigma}^{k-1} + \mathbb{C} : (\nabla^s \boldsymbol{u}^k - \boldsymbol{\varepsilon}^{k-1}), \tag{37b}$$

$$f_{\text{yield}} = s_{\text{trial}} - \sigma_0, \tag{37c}$$

其中 $\boldsymbol{\sigma}_{\text{trial}}$ 为试应力，$\mathbb{C}$ 为基体材料的弹性张量，$\boldsymbol{s}_{\text{trial}}$ 为 $\boldsymbol{\sigma}_{\text{trial}}$ 的偏量部分，$s_{\text{trial}} = \sqrt{\frac{3}{2}\boldsymbol{s}_{\text{trial}} : \boldsymbol{s}_{\text{trial}}}$为等效冯·米塞斯应力，$\sigma_0$为基体材料的屈服应力，$\langle \cdot \rangle_+$ 为斜坡函数。

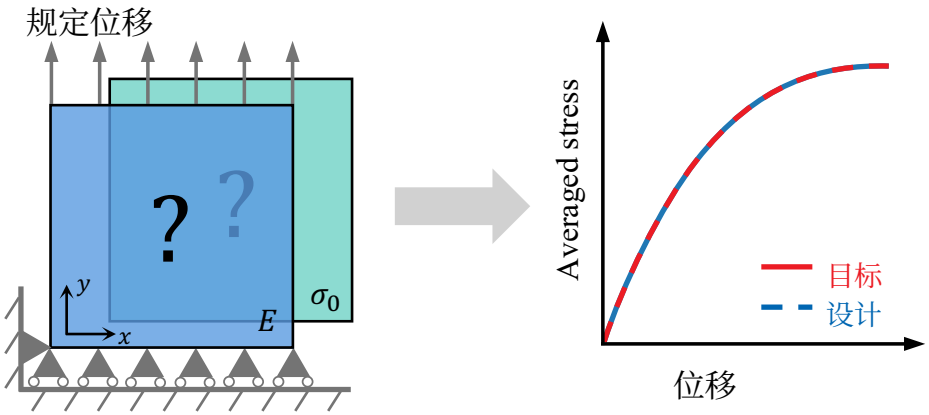

图11：针对定制弹塑性性能的微结构设计示例的模型设置示意图。

如图11所示，我们考虑在单轴拉伸条件下具有目标应力-位移曲线的微结构设计任务。考虑两种基材：杨氏模量为 $E_1 = 10^5$ 、屈服应力为 $\sigma_0^{(1)} = 300$的硬材料，以及杨氏模量为 $E_0 = 10^3$ 、屈服应力为 $\sigma_0^{(0)} = 30$的软材料。两种材料的泊松比均为 $\nu = 0.3$。计算域的上边界在y 方向上施加均匀位移载荷，范围从 5×10^{-4} 到 10^{-2}。下边界在 y 方向上固定，而

左下角也在x 方向上固定。设计目标函数如下：

$$\mathcal{L}(\boldsymbol{w}) = \frac{1}{N}\sum_{i=1}^{N}\left[\bar{\sigma}_i(\boldsymbol{w}) - \bar{\sigma}_i^{\text{obj}}\right]^2 \tag{38}$$

其中 N 是位移加载步的总数。 $\bar{\sigma}_i$ 表示生成的微结构在位移加载u_i下y 方向上的平均应力， $\bar{\sigma}^{\text{obj}}{}_i$ 表示相应的目标值。

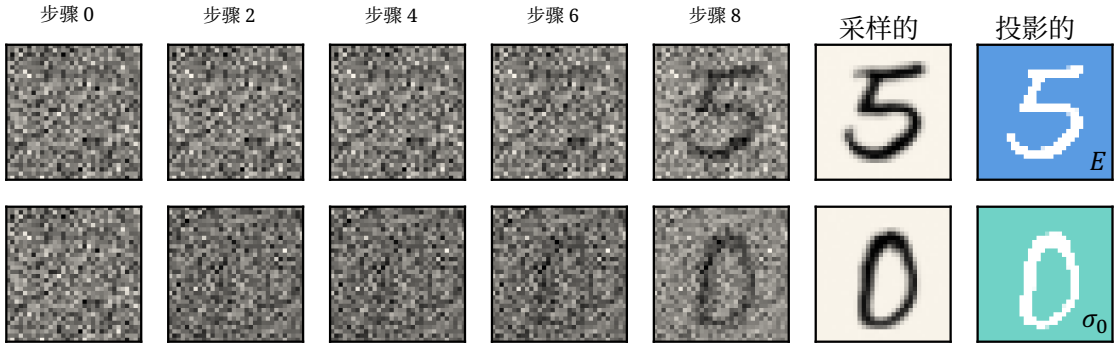

图 12：弹塑性设计示例中两个随机采样的噪声输入的反向去噪过程及投影微结构：杨氏模量场（上）；屈服应力场（下）。

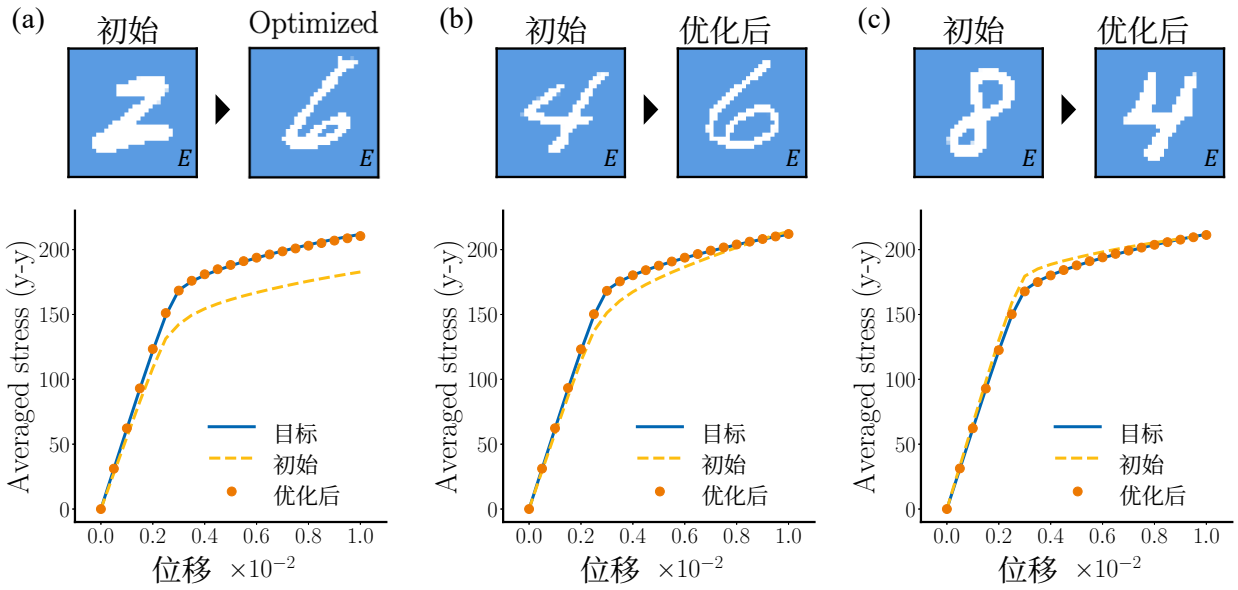

图 13：匹配相同目标应力-位移曲线的弹塑性微结构设计。杨氏模量场从不同初始设计开始进行优化：（a）数字 2；（b）数字 4；（c）数字 8。

我们首先研究单一材料属性场的设计，即杨氏模量 E 或屈服应力 σ_0，同时保持另一个属性场不变。如图 12所示，我们从训练好的基于MNIST的扩散模型中随机采样两个不同的数字，分别表示杨氏模量场（上）和屈服应力场（下）。最后一列展示了具有 $\gamma = 100$ 的投影微结构，如图 12所示，并用其计算应力-位移曲线：

作为目标。在设计杨氏模量场时，屈服应力场固定为图 12中的场，反之亦然。多阶段优化采用 γ 取值 5、10、20 和 100，损失容差为 0.2，每个阶段的最大迭代次数为 50。

图 13 展示了三种不同初始微结构（包括数字 2、4 和 8）的杨氏模量场设计结果。数字 2 和 4 的优化后微结构收敛为相似的设计，形似数字 6；数字 8 的优化后微结构则收敛为类似数字 4 的设计。如图 13底部所示，无论初始设计表面上有何不同，优化后的微结构均成功实现了目标弹塑性行为。

不同初始微结构的屈服应力场设计结果如图 14所示。研究了三种不同的初始设计，包括数字 3、5 和 7。数字 3 和 5 的优化后微结构分别收敛为形似数字 8 和 9 的相似设计，而数字 7 的优化后微结构则收敛为缩小版的数字 7。尽管优化后的屈服应力场呈现出不同的几何特征，但如图14底部所示，它们均成功实现了目标弹塑性行为。

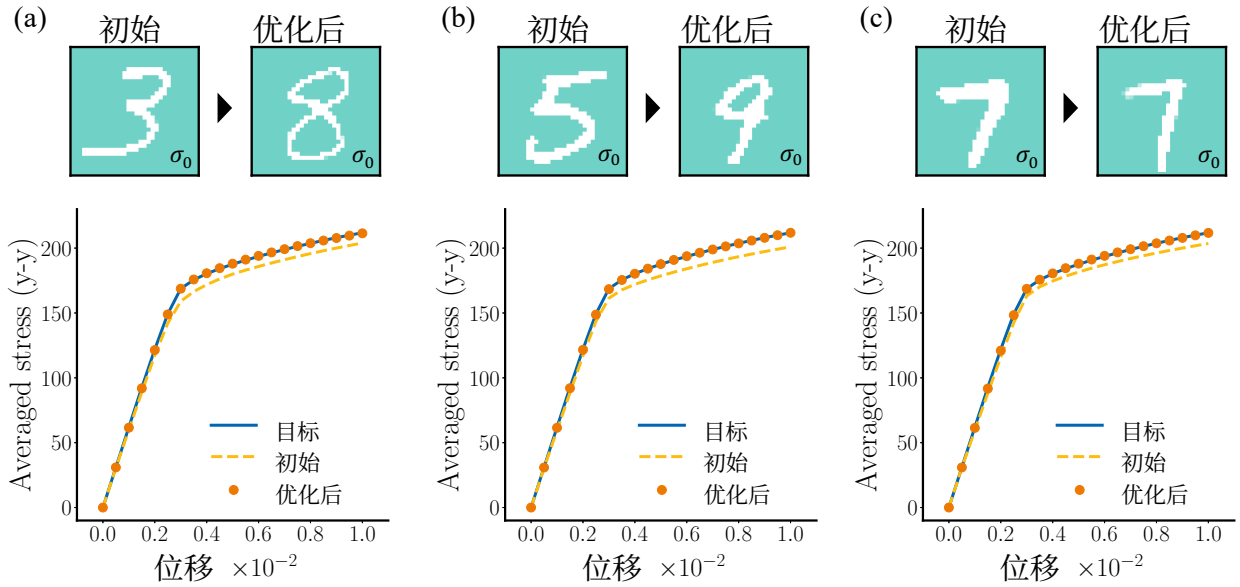

图 14：匹配同一目标应力-位移曲线的弹塑性微结构设计。从不同的初始设计出发优化屈服应力场：(a) 数字 3；(b) 数字 5；(c) 数字 7。

在展示所提出的方法在设计单个材料属性场方面的有效性后，我们考虑同步设计杨氏模量场和屈服应力场，以实现定制的弹塑性性能。为了用参数形式表示目标弹塑性行为，我们使用分段函数描述基于 MNIST 的微结构的应力-位移曲线

$$f(x) = \begin{cases} Ex, & x \le \varepsilon_y \\ \sigma_0 + \sum_{i=1}^{n} A_i \left(1 - e^{-\kappa_i (x - \varepsilon_y)}\right), & x > \varepsilon_y \end{cases} \tag{39}$$

其中，线弹性区域由杨氏模量E 表征，直至屈服应变ε_y，此时屈服应力为 σ_0；屈服后的塑性行为则采用一组指数函数之和建模，其振幅为A_i，速率为 κ_i。A_i 之和等于 $\sigma_\infty - \sigma_0$，其中 σ_∞ 为大应变下的渐近应力。由于 $\sigma_0 = E\varepsilon_y$，目标弹塑性行为可由参数 $\{E$、ε_y、σ_∞、A_i、$\kappa_i\}$完全定义。在此示例中，我们设定$n = 3$，并考虑三条具有不同特性的目标应力-位移曲线，相应参数如下

列于表1中。初始噪声输入与图12中的保持一致：杨氏模量场取数字5，屈服应力场取数字0。多阶段优化采用三个 γ 取值：10、20和100；损失容差设为0.5，每个阶段的最大迭代次数设为100。

表1：同步设计的目标弹塑性曲线参数。

组	E	ε_y	σ_∞	A_i	κ_i
1	4.8×10^4	2.7×10^{-3}	220	[4, 1, 85.4]	[1000, 88, 88]
2	6.8×10^4	2.8×10^{-3}	250	[12, 1, 46.6]	[1000, 87, 87]
3	8.2×10^4	2.8×10^{-3}	280	[18, 1, 31.4]	[1000, 63, 63]

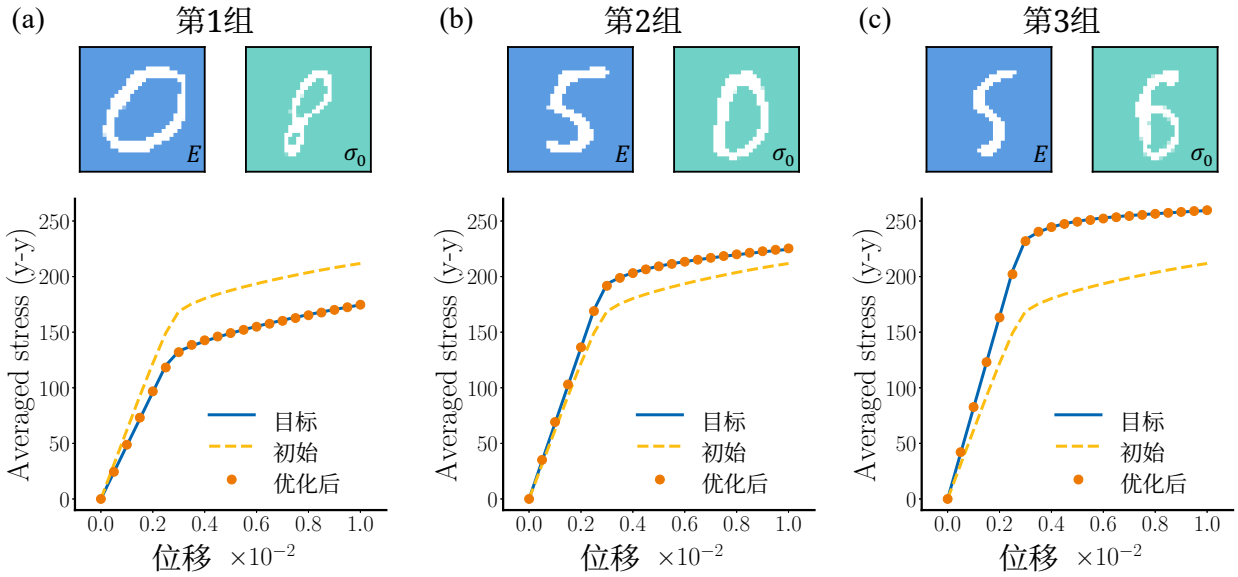

图15：采用相同初始设计，同时设计杨氏模量场和屈服应力场，以实现不同目标弹塑性行为的设计结果：(a) 第1组；(b) 第2组；(c) 第3组。

图15展示了三组杨氏模量场与屈服应力场的同步设计结果。优化后的微结构呈现出不同的几何特征，以实现不同的目标属性。在第1组中，优化后的杨氏模量场呈现出数字 0 的形状，并扩大了软材料区域，以实现较低的刚度；而屈服应力场则呈现出微小数字 8 的形状，以实现屈服后应力的持续增加。在第2组中，目标曲线与初始曲线较为接近，因此优化后的微结构与初始设计相似。在第3组中，优化后的杨氏模量场保留了初始数字 5 的整体形状，但尺寸有所缩小，以提高刚度；屈服应力场则呈现出尺寸相近的数字 6 的形状，以实现屈服后应力的缓慢增加。所有优化后的微结构的二值化指标均约为 99%。优化后微结构的所有应力-位移曲线均与目标曲线高度吻合，证明了所提出方法在同步设计多个材料属性场、实现定制弹塑性行为方面的有效性。

3.5 讨论

本节所示结果证明，所提出的逆向设计框架能够生成具有定制力学性能的微结构，同时保留训练数据中的风格特征。与现有基于条件生成模型的逆向设计方法相比，所提出的方法具有以下优势：

1. 设计范式的灵活性：现有的条件设计方法旨在通过依据目标属性对反向采样进行条件控制，同时生成具有所需力学性能和风格特征的微结构。这种范式通常需要准备大量微结构—属性对数据集，并设计复杂的条件嵌入机制。此外，控制方程或边界条件的任何变化通常都需要重新生成数据集，并重新训练条件生成模型，这既耗时又耗费计算资源。相比之下，本文提出的方法将定制机械性能与风格特征的实现解耦，重新将噪声输入解释为可优化的设计变量。这种新范式使扩散模型能够专注于捕捉训练数据的几何特征，同时由优化过程有效引导设计趋向所需的机械性能。这种灵活性使得预训练的无条件扩散模型能够直接应用于各种风格化任务。

2. 目标设定的普适性：得益于基于优化的设计过程，所提出的框架能够轻松适应各种形式的目标性能规格，包括特定的属性值和完整的响应曲线。这一特点使其能够更灵活地应对广泛的设计问题，无需针对不同目标设计专门的条件嵌入机制。

3. 实现效率：通过向量-雅可比积拼接实现的统一微分流程，可借助反向传播高效计算梯度。力学求解器的定制向量-雅可比积规则与具体问题无关，也易于扩展至其他力学系统。这一能力大幅减少了推导和实现所需的人工工作，从而加快设计流程。

尽管具有这些优势，当前研究仍存在一些局限，有待未来工作进一步探讨：

1. 噪声输入退化：正如超弹性设计示例所示，在以极端力学性能为目标时，优化后的噪声输入往往会偏离各向同性高斯分布。退化的噪声输入可能导致生成的微结构偏离所学习的数据分布，进而丧失风格特征并降低设计多样性。尽管可以通过选择不同的基材组合来缓解这些问题，但未来工作仍可探索在优化过程中更好地对噪声输入进行正则化的策略，并增强扩散模型的生成能力，以捕捉更广泛的设计变化。

2. 几何约束较为宽松：基于生成模型的设计方法能够自然地捕捉训练数据的风格特征，但未必能完全满足实际应用所要求的严格几何约束，正如均匀化设计示例所示。尽管训练数据集满足这些约束，生成样本仍可能因生成过程的性质以及我们方法中的优化流程而违反这些约束。除了采用多阶段投影方案来促进二值化，以及利用所讨论的对称操作来实现对称设计外，未来研究还可探讨将其他约束纳入优化

公式中，例如最小特征尺寸 [53] 和连通性要求 [54] ，这些约束在拓扑优化领域已得到充分发展。

4 结论

本文提出了一种新的逆向设计框架，将无条件去噪扩散模型与先进的可微编程技术相结合，实现了具有定制力学性能和风格特征的微结构设计。将噪声输入重新诠释为设计变量，可将传统生成过程转化为受约束优化问题的求解，从而有效控制反向去噪过程，并免去对条件嵌入的需求。通过拼接向量-雅可比积构建的统一微分流程，可借助反向传播高效、自动地进行灵敏度分析。本文通过多种微结构设计问题验证了所提方法的有效性，包括定制化的均质化性能、超弹性性能和弹塑性性能。数值结果表明，所提方法能够生成高质量的微结构，在保留训练数据风格特征的同时，使其力学性能与目标值高度匹配。本研究为将生成模型与基于物理的仿真相结合，推进先进材料与结构设计提供了一个有前景的方向。

5 致谢

作者感谢香港研究资助局通过 ECS项目（参考编号：26205024）提供资助。

6 数据可用性

我们的数据和代码可在以下网址获取：https://github.com/CMSL-HKUST/genop t.

参考文献

[1] Madihah Kazim、Aniket Pal 和 Debkalpa Goswami。用于生物工程的力学超材料：体外、可穿戴和可植入应用。*Advanced Engineering Materials*，27(7):2401806，2025年。https://doi.org/10.1002/adem.202401806。[2] Xiaoyang Zheng、Yuhao Jiang、Mustafa Mete、Jingjing Li、Ikumu Watanabe、Takayuki Yamada 和 Jamie Paik。超材料机器人学。*Science Robotics*，10(108):eadx1519，2025年。https://doi.org/10.1126/scirobotics.adx1519。[3] Lujie Zhang、Sen Yan、Wenlong Liu、Yao Liu、Wenjun Cai、Zidong Zhang 和 Ji Zhou。负泊松比力学超材料：综述。*Engineering Structures*，329:119838，2025年。https://doi.org/10.1016/j.engstruct.2025.119838。[4] Qingliang Zeng、Shengyu Duan、Zeang Zhao、Panding Wang 和 Hongshuai Lei。通过拓扑优化对吸能超材料进行逆向设计。*Advanced Science*，10(4):2204977，2023年。https://doi.org/10.1002/advs.202204977。[5] Liwei Wang、Jagannadh Boddapati、Ke Liu、Ping Zhu、Chiara Daraio 和 Wei Chen。通过数据驱动的非周期超材料设计实现力学隐身。美国国家科学院院刊，119(13):e2122185119，2022年。https://doi.org/10.1073/pnas.2122185119。

[6] Marco Maurizi、Derek Xu、Yu-Tong Wang、Desheng Yao、David Hahn、Mourad Oudich、Anish Satpati、Mathieu Bauchy、Wei Wang、Yizhou Sun 等。图空间中具有可编程非线性响应和几何约束的超材料设计。Nature Machine Intelligence，7(7):1023–1036，2025年。https://doi.org/10.1038/s42256-025-01067-x。[7] Lucas YM Sampaio、Pedro CM Cerˆantola 和 Leopoldo PR de Oliveira。作为隔音板美学解决方案的轻质装饰膜。Journal of Sound and Vibration，532:116971，2022年。https://doi.org/10.1016/j.jsv.2022.116971。[8] JTB Overvelde、S Shan 和 K Bertoldi。二维周期性柔性多孔结构中的屈曲致压实：孔形的影响。Advanced Materials，24(17):2337–2342，2012年。https://doi.org/10.1002/adma.201104395。[9] Sicong Shan、Sung H Kang、J ordan R Raney、Pai Wang、Lichen Fang、Francisco Candido、Jennifer A Lewis 和 Katia Bertoldi。用于捕获弹性应变能的多稳态架构材料。Adv. Mater，27(29):4296–4301，2015年。https://doi.org/10.1002/adma.201501708。[10] Wei Sha、Mi Xiao、Yihui Wang、Mingzhe Huang、Qishi Li 和 Liang Gao。热超材料的拓扑优化方法：综述。International Journal of Heat and Mass Transfer，227:125588，2024年。https://doi.org/10.1016/j.ijheatmasstransfer.2024.125588。[11] Tanguy Navez、Martin-Pierre Schmidt、Ole Sigmund 和 Claus BW Pedersen。由几何图案库引导的拓扑优化。Structural and Multidisciplinary Optimization，65(4):108，2022年。https://doi.org/10.1007/s00158-022-03197-x。[12]Weisheng Zhang、Yue Wang、Zongliang Du、Chang Liu、Sung-Kie Youn 和 Xu Guo。机器学习辅助的、具有艺术风格的建筑设计拓扑优化。Computer Methods in Applied Mechanics and Engineering，413:116041，2023年。https://doi.org/10.1016/j.cma.2023.116041。[13] Doksoo Lee、Wei Chen、Liwei Wang、Yu-Chin Chan 和 Wei Chen。超材料与多尺度系统的数据驱动设计：综述。Advanced Materials，36(8):2305254，2024年。https://doi.org/10.1002/adma.202305254。[14] Moritz Flaschel、Siddhant Kumar 和 Laura De Lorenzis。利用 euclid 自动发现广义标准材料模型。Computer Methods in Applied Mechanics and Engineering，405:115867，2023年。https://doi.org/10.1016/j.cma.2022.115867。[15] Diederik P Kingma 和 Max Welling。自编码变分贝叶斯。arXiv 预印本 arXiv:1312.6114，2013年。[16] Ian J Goodfellow、Jean Pouget-Abadie、Mehdi Mirza、Bing Xu、David Warde-Farley、Sherjil Ozair、Aaron Courville 和 Yoshua Bengio。生成对抗网络。Advances in Neural Information Processing Systems，27，2014年。https://doi.org/10.1145/3422622。[17] Laurent Dinh、Jascha Sohl-Dickstein 和 Samy Bengio。使用 Real NVP 进行密度估计。arXiv 预印本 arXiv:1605.08803，2016年。[18] Jonathan Ho、Ajay Jain 和 Pieter Abbeel。去噪扩散概率模型。Advances in Neural Information Processing Systems，33:6840–6851，2020年。[19] Jiaming Song、Chenlin Meng 和 Stefano Ermon。去噪扩散隐式模型。arXiv 预印本 arXiv:2010.02502，2020年。

[20] Yang Song、Jascha Sohl-Dickstein、Diederik P Kingma、Abhishek Kumar、Stefano Ermon 和 Ben Poole。通过随机微分方程进行基于分数的生成建模。arXiv 预印本 arXiv:2011.13456，2020年。[21]Liwei Wang、Yu-Chin Chan、Faez Ahmed、Zhao Liu、Ping Zhu 和 Wei Chen。用于超材料系统机理学习与设计的深度生成建模。《应用力学与工程中的计算方法》，372:113377，2020年。https://doi.org/10.1016/j.cma.2020.113377。[22] Qi Li、Jianwei Wang、Tao Lei、Tianyu Xiang、Chanchan Qin 和 Maoze Yang。基于变分自编码器的吸波器超材料设计。IEEE Access，12:92328–92336，2024年。https://doi.org/10.1109/ACCESS.2024.3422240。[23]Yunwei Mao、Qi He 和 Xuanhe Zhao。利用生成对抗网络设计复杂架构材料。《科学进展》，6(17):eaaz4169，2020年。https://doi.org/10.1126/sciadv.aaz4169。[24] Xiaoyang Zheng、Ta-Te Chen、Xiaofeng Guo、Sadaki Samitsu 和 Ikumu Watanabe。利用深度学习对负泊松比超材料进行可控逆向设计。《材料与设计》，211:110178，2021年。https://doi.org/10.1016/j.matdes.2021.110178。[25] Jun Wang、Wei Wayne Chen、Daicong Da、Mark Fuge 和 Rahul Rai。IH-GAN：一种用于基于隐式曲面的胞状结构逆向设计的条件生成模型。《应用力学与工程中的计算方法》，396:115060，2022年。https://doi.org/10.1016/j.cma.2022.115060。[26] Nikolaos N Vlassis 和 WaiChing Sun。用于微结构逆向设计的去噪扩散算法，可精细调控非线性材料属性。《应用力学与工程中的计算方法》，413:116126，2023年。https://doi.org/10.1016/j.cma.2023.116126。[27] Jan-Hendrik Bastek 和 Dennis M Kochmann。通过视频去噪扩散模型对非线性力学超材料进行逆向设计。Nature Machine Intelligence，5(12):1466–1475，2023年。https://doi.org/10.1038/s42256-023-00762-x。[28]Haoyu Wang、Zongliang Du、Fuyong Feng、Zhong Kang、Shan Tang 和 Xu Guo。Diffmat：利用扩散模型对吸能超材料进行数据驱动逆向设计。《应用力学与工程中的计算方法》，432:117440，2024年。https://doi.org/10.1016/j.cma.2024.117440。[29] Guangfa Li 和 Dehao Liu。基于条件扩散模型，对具有目标力学行为的复合材料进行逆向设计。Engineering Structures，345:121417，2025年。https://doi.org/10.1016/j.engstruct.2025.121417。[30] Yu Zhang、Lijie Liu、Shukai Li、KI Elkhodary 和 Shan Tang。HyperDiff：一种基于条件扩散模型的超弹性微结构逆向设计框架。2025年。[31]Hossein Mirzaee 和 Serveh Kamrava。使用条件连续归一化流对微结构进行逆向设计。《材料学报》，285:120704，2025年。https://doi.org/10.1016/j.actamat.2024.120704。[32]Zhenling Yang、Yilin Guo、Zhi Sun、Khalil I Elkhodary、Fuyong Feng、Zhong Kang、Shan Tang 和 Xu Guo。GraphDGM：一种用于框架和晶格结构的生成式数据驱动设计方法。《机械设计杂志》，147(10):101701，2025年。https://doi.org/10.1115/1.4068106。

[33] Jinlong Zhang、Shikun Chen、Robert J Martin、Baochang Liu、Ruixiong Zhang 和 Dengbao Xiao。用于晶格结构逆向设计的条件扩散模型。Structural and Multidisciplinary Optimization，68(3):1–23，2025年。https://doi.org/10.1007/s00158-025-03984-2。[34] Li Zheng、Siddhant Kumar 和 Dennis M Kochmann。DiffuMeta：利用扩散变换器进行超材料逆向设计的代数语言模型。arXiv 预印本 arXiv:2507.15753，2025年。[35] Mathieu Blondel 和 Vincent Roulet。可微编程要素。arXiv 预印本 arXiv:2403.14606，2024年。[36] David E Rumelhart、Geoffrey E Hinton 和 Ronald J Williams。通过误差反向传播学习表征。Nature，323(6088):533–536，1986年。https://doi.org/10.1038/323533a0。[37] Andreas Griewank 和 Andrea Walther。导数求值：算法微分的原理与技术。SIAM，2008年。https://doi.org/10.1137/1.9780898717761。[38] Mathieu Blondel、Quentin Berthet、Marco Cuturi、Roy Frostig、Stephan Hoyer、Felipe Llinares- L´opez、Fabian Pedregosa 和 Jean-Philippe Vert。高效且模块化的隐式微分。Advances in Neural Information Processing Systems，35:5230–5242，2022年。[39] Dan Givoli。逆问题伴随方法教程。《应用力学与工程中的计算方法》，380:113810，2021年。https://doi.org/10.1016/j.cma.2021.113810。[40]Yang Cao、Shengtai Li、Linda Petzold 和 Radu Serban。微分代数方程的伴随灵敏度分析：伴随微分代数方程组及其数值解。SIAM科学计算杂志，24(3):1076–1089，2003年。https://doi.org/10.1016/S0377-0427(02)00528-9。[41] James Bradbury、Roy Frostig、Peter Hawkins、Matthew James Johnson、Chris Leary、Dougal Maclaurin、George Necula、Adam Paszke、Jake VanderPlas、Skye Wanderman-Milne 和 Qiao Zhang。JAX：Python+NumPy 程序的可组合变换，2018年。[42] Jonathan Heek、Anselm Levskaya、Avital Oliver、Marvin Ritter、Bertrand Rondepierre、An- dreas Steiner 和 Marc van Zee。Flax：用于 JAX 的神经网络库与生态系统，2024年。[43] Tianju Xue、Shuheng Liao、Zhengtao Gan、Chanwook Park、Xiaoyu Xie、Wing Kam Liu 和 Jian Cao。JAX-FEM：一种用于自动逆向设计和机理数据科学的可微 GPU加速三维有限元求解器。《计算物理学通讯》，291:108802，2023年。https://doi.org/10.1016/j.cpc.2023.108802。[44]Jorge Nocedal 和 Stephen J Wright。数值优化。Springer，2006年。https://doi.org/10.1007/978-0-387-40065-5。[45] Yann LeCun、L´eon Bottou、Yoshua Bengio 和 Patrick Haffner。应用于文档识别的基于梯度的学习。IEEE会议录，86(11):2278–2324，2002年。https://doi.org/10.1109/5.726791。[46]Liwei Wang、J agannadh Boddapati、Ke Liu、Ping Zhu、Chiara Daraio 和 Wei Chen。二维正交各向异性超材料微结构数据集。2021年12月。https://doi.org/10.6084/m9.figshare.17141975.v1。[47] Olaf Ronneberger、Philipp Fischer 和 Thomas Brox。U-Net：用于生物医学图像分割的卷积网络。载于医学图像计算与计算机辅助干预国际会议，第234–241页。Springer，2015年。https://doi.org/10.1007/978-3- 319-24574-4 28。

[48] 何恺明、张祥雨、任少卿和孙剑。面向图像识别的深度残差学习。载于*IEEE*计算机视觉与模式识别会议录，第770–778页，2016年。https://doi.org/10.1109/CVPR.2016.90。[49] Ashish Vaswani、Noam Shazeer、Niki Parmar、Jakob Uszkoreit、Llion Jones、Aidan N Gomez、Łukasz Kaiser 和 Illia Polosukhin。注意力机制就是你所需要的一切。*Advances in Neural Information Processing Systems*，第30卷，2017年。[50] Ilya Loshchilov 和 Frank Hutter。解耦权重衰减正则化。*arXiv*预印本 *arXiv:1711.05101*，2017年。[51]Boris Desmorat 和 Nicolas Auffray。二维弹性材料的空间：一段几何之旅。连续介质力学与热力学，31(4):1205–1229，2019年。https://doi.org/10.1007/s00161-019-00794-y。[52]Juan C Simo 和 Thomas JR Hughes。计算非弹性力学。Springer，1998年。https://doi.org/10.1007/b98904。[53] Boyan S Lazarov、Fengwen Wang 和 Ole Sigmund。基于密度的拓扑优化中的长度尺度与可制造性。应用力学档案，86(1):189–218，2016年。https://doi.org/10.1007/s00419-015-1106-4。[54] Vanessa Cool、Niels Aage 和 Ole Sigmund。促进拓扑优化中连通性的一篇实用综述。*Structural and Multidisciplinary Optimization*，68(4):73，2025年。https://doi.org/10.1007/s00158-025-04004-z。