arXiv:2601.13233v1 [cs.AI] 19 Jan 2026

RAG：一种基于随机森林的生成式设计框架，用于具有复杂功能响应要求的超材料不确定性感知设计

Bolin Chen[1], Dex Doksoo Lee[1],Wei “Wayne” Chen[2],和 Wei Chen[1*]

[1]西北大学机械工程系，美国伊利诺伊州埃文斯顿，邮编 60208 [2]德克萨斯农工大学 J. Mike Walker ’66 机械工程系，美国德克萨斯州大学城，邮编 77843[*]通讯作者：weichen@northwestern.edu

摘要

面向先进功能的超材料设计通常需要对非线性且依赖条件的响应（例如应力-应变关系和色散关系）进行逆向设计，这些响应由连续函数描述。现有的大多数设计方法聚焦于向量值响应（例如杨氏模量和带隙宽度），而功能响应的逆向设计仍面临挑战，原因在于其高维性、在逆向设计框架中纳入设计要求的复杂性，以及可行解可能不存在或不唯一。尽管生成式设计方法展现出良好前景，但它们往往需要大量数据，以启发式方式处理设计要求，且可能在未进行不确定性量化的情况下生成不可行设计。为应对这些挑战，我们提出一种基于随机森林的生成方法（RAG）。借助随机森林对小样本数据的适用性，并以离散化不变的方式重新表述正向映射，RAG能够以数据高效的方式预测高维功能响应。在逆向设计过程中，该框架估计以可灵活指定的设计要求为条件的解的似然。通过集成估计的似然能够量化生成设计的可信度，同时反映不同要求之间的相对难度。通过从条件似然中采样，该方法以单次生成的方式解决一对多映射问题。我们在以下案例中展示了RAG：1）具有指定部分通带/阻带的声学超材料；2）具有目标跳跃响应的机械超材料，分别使用500个和1057个样本。我们还在一个包含非线性应力-应变关系的公开机械超材料数据集上，将其数据效率与神经网络进行了基准比较。我们的框架为涉及功能响应、昂贵仿真和复杂设计要求的逆向设计提供了一种轻量且可信的途径，并可推广至超材料之外。

关键词 随机森林 ·生成式设计 · 功能响应 · 不确定性量化 n

1 引言

超材料是人工工程材料，其非凡特性源于几何结构而非材料成分 [1–3]。通过精心的几何设计，它们可实现广泛的应用，涵盖热学 [4], 机械 [5], 声学 [6], 和光学 [7] 领域，而这些应用是传统材料无法实现的。针对目标功能设计超材料，关键在于定制单元胞几何结构以获得所需响应。例如，通过镶嵌具有定制热导率张量的单胞，可以实现热斗篷 [8, 9]。通过空间分布具有局部目标刚度的微结构，可以实现展现目标变形行为的机械超材料

[10]。同样，通过排列具有所需带隙特性的单胞，可以调控声学超材料中的波传播 [11]。

单胞响应的丰富程度决定了可实现的系统级功能。现有超材料设计工作主要关注向量值响应，例如刚度张量 [10, 12–17]，带隙范围 [18–21], 热导率张量 [8]，或它们的组合 [22], ，这些都可以表示为有限维向量。尽管向量值响应能够有效表征其底层行为，但往往不足以对其进行全面描述。例如，刚度仅表征应力–应变关系的线弹性区间，却无法描述突跳等非线性力学行为。相比之下，功能响应以无限维函数 [23, 24]表示，能够描述具有复杂物理机制的非线性行为（例如非线性应力–应变关系），或随环境因素变化的特性（例如温度依赖型热导率）。功能响应承载着向量值响应所衍生自的底层行为。因此，基于功能响应构建的设计能够解锁更高级、更复杂的功能，而这些功能可能无法通过其他方式实现。例如，通过分布具有指定应力–应变关系的单胞，可以对软体执行器进行时空编程 [25] （图 1 (a)）。通过设计具有所需色散关系且满足特定通带/阻带要求的单胞，可以实现定向噪声滤波 [26] （图 1 (b)）。通过调整材料在不同温度下的热导率，可以实现温度可切换隐身和聚能 [27] （图 1 (c)）。文献中报道的其他实例还包括具有可调动态性能的鞋中底 [28], 、长曲棍球护胸和减振板 [29], ，以及能量吸收与耗散 [30–32]。

尽管前景广阔，但由于功能响应的高维性以及设计方案的不唯一性或不存在性，设计具有期望功能响应的超材料仍然面临挑战。为处理函数的无限维特性，通常的做法是将其离散化为高维向量 [25, 33]。这种高维性往往与潜在结构–性能关系的非线性相伴，不仅使功能响应的评估在计算上代价高昂，也使设计空间与响应空间之间的映射更加复杂。此外，对于功能响应给定的设计要求，如果相对容易满足，则可能存在许多可行设计；而如果其本质上十分严格，则可能一个也不存在 [23, 34]。具体而言，施加于功能响应的设计要求具有多种表达形式：它们既可以规定精确的响应值（例如图 1(a) 所示的完整应力–应变曲线），也可以仅约束某些特性（例如图 1(b) 所示色散关系的声学带隙特性）。需求规格说明的这种复杂性进一步增加了逆向设计的难度。

传统优化方法（例如拓扑优化） [35], 已广泛用于涉及功能响应的超材料设计 [30, 36–39]。在这些方法中，设计要求被表述为目标函数，例如最小化目标应力–应变关系与所得关系之间的差异 [30, 36], ，并通过迭代评估结构–性能关系来获得设计方案。基于梯度的拓扑优化已被证明可有效用于声学带隙的逆向设计 [37,38], 、拓扑态 [39],以及非线性应力—应变关系 [30, 36]。然而，解析灵敏度分析仅适用于解析可微的目标函数，这限制了设计要求规格说明的灵活性。为避免推导灵敏度，人们采用了遗传算法等非梯度优化方法 [31, 40, 41], ，以便更灵活地处理难以表述为可微目标函数的复杂设计要求。尽管如此，这些方法往往需要大量评估功能响应，而评估过程通常耗时较长。一般而言，迭代优化对初始猜测值较为敏感，这限制了其高效探索多样化设计方案的能力。

数据驱动设计 [42, 43] 为具有功能响应的超材料逆向设计提供了一种颇具前景的替代方案。该框架的核心是代理正向模型，它可以通过训练实现快速、即时预测，避免高昂的评估成本 [42], ，随后还能与下游优化相结合，加速搜索过程 [44–47], ，或与另一种机器学习模型结合，实现一次性、无迭代的逆向设计 [25, 28, 33, 48, 49]。近年来，条件深度生成模型，包括生成对抗网络 [50, 51], 、扩散模型 [52–54], 和变分自编码器 [55], ，已被用于生成以功能响应为条件的设计候选方案，例如应力–应变关系和声学色散关系。由于生成模型本质上具有随机性，相较于确定性模型，它们能够更有效地处理一对多映射 [56]。设计方案的多样性为考虑主要需求之外的其他设计目标提供了灵活性。

然而，利用机器学习处理功能响应的高维性通常需要大型数据集——样本量往往达到 10^3-10^5 个 [25, 28, 33]。当评估功能响应的计算成本较高时，获取如此大型的数据集便变得不切实际。此外，大多数生成模型直接以响应而非要求 [50–53, 55, 56], 为条件，因此只能支持与特定响应相关联的要求

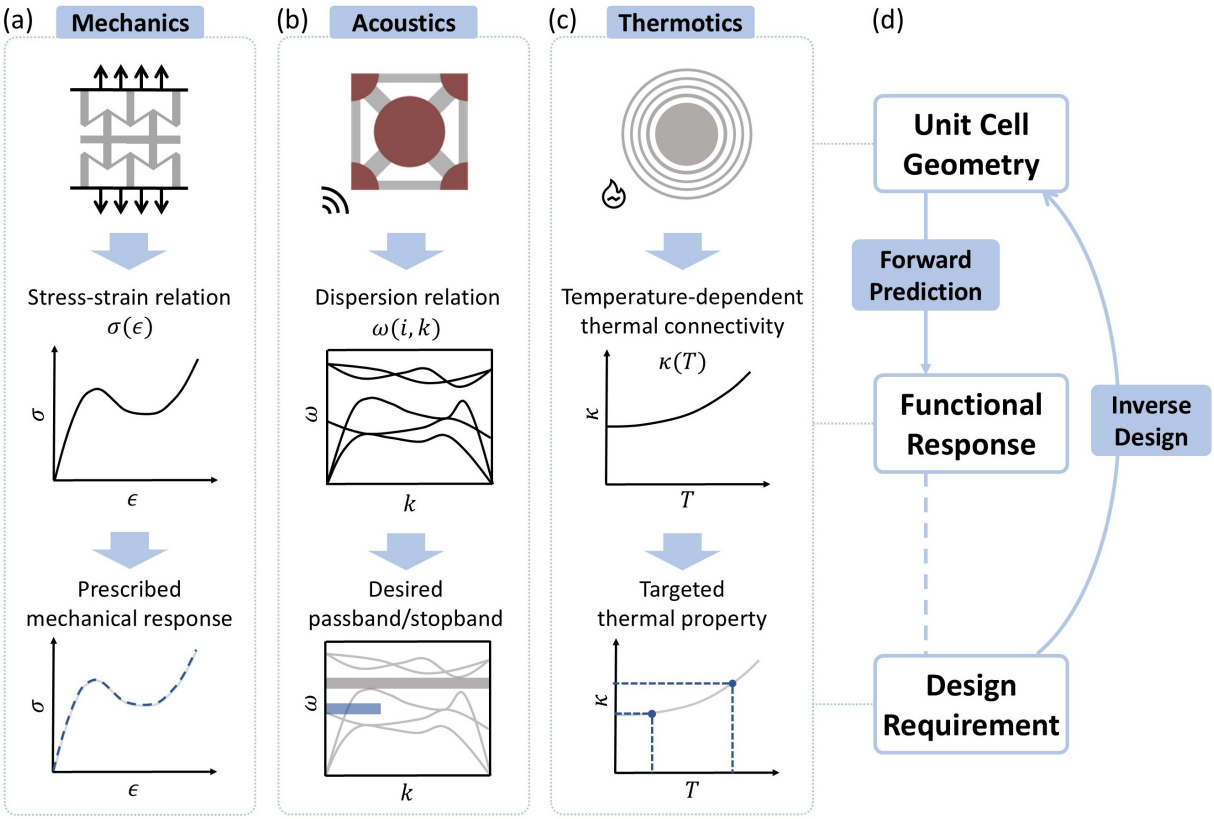

图1：不同类别超材料功能响应逆向设计示意图。每种情况下，中间面板展示功能响应，底部面板则展示相应的设计要求。(a) 机械超材料：功能响应为应力–应变关系，其中应力 σ 是应变 ϵ的函数。设计要求是预设的非线性力学行为，以蓝色虚线表示。(b) 声学超材料：功能响应为色散关系，其中频率 ω 是带索引 i 和波矢 k的函数。设计要求由蓝色和灰色阴影区域表示，分别代表所需的通带和阻带。(c) 热超材料：功能响应为温度依赖的热性质，其中热连通性 κ 是温度 T的函数。设计要求——高温时高导热、低温时低导热——以蓝色虚线表示。(d) 示意图展示了单元几何结构、功能响应与超材料设计要求之间的关系。设计要求的形式因应用而异，且通常难以明确规定。

（例如，在声学带隙逆向设计中，必须提供完整的色散关系作为条件响应 [33]）。此外，由于现有的生成式设计方法在条件推断中无法提供不确定性估计，我们无法评估生成的设计方案的可信度。鉴于可行解可能不存在或不唯一，这一点尤其成问题：如果不对这些解进行不确定性评估，该方法可能会生成实际上不可行的设计，尤其是在要求位于训练分布之外，或无法在预定义的设计空间内得到满足时 [23]。

迄今为止，大多数数据驱动的设计方法都采用神经网络（NN） [42,43]。尽管神经网络强大的表达能力有助于准确预测功能响应，但它们对数据需求量大，需要进行大量超参数调优，而且训练通常需要消耗大量资源。相比之下，随机森林是另一种应用广泛的机器学习方法，构建相对简单，并且对训练参数的变化更稳定 [57]。随机森林尤其适用于小数据场景，可降低超材料设计中的数据需求；在这类设计中，获取功能响应的仿真计算成本高昂。此外，其集成结构天然支持对模型预测进行不确定性量化 [58]。随机森林已被证明能够有效预测应力–应变关系 [59] 和光学透射光谱 [60], 等功能响应，但其固有的不确定性量化能力很少用于逆向设计。我们此前的工作 [24]

利用随机森林的集成结构，为生成的方案提供置信度估计。该方法的数据效率已通过声学超材料和光学超表面的光谱逆向设计得到验证，每个示例所需的样本均少于 250 个。不过，该框架仅适用于简化的定性功能响应，例如以二元形式表示声学和光学中的光谱行为，并未处理完全定量的响应。

在本研究中，我们提出了一种基于随机森林的生成式设计框架（RAG），可用于解决具有复杂需求规格说明的定量高维功能响应逆向设计问题，并兼容小数据集，同时量化生成设计的不确定性。我们训练随机森林进行正向映射，其中每棵决策树都根据给定的单元几何结构预测完整的功能响应。利用这些完整的功能响应信息，可以针对特定响应值或从响应中提取的特征来指定复杂设计要求。随后，每棵树都可以判断其预测的响应是否满足该要求。汇总随机森林中各棵树的投票结果，可得到以该要求为条件的似然估计，表明模型对于给定设计能实现所需功能响应的置信度。通过基于条件似然分布对设计空间进行采样，无需迭代即可一次性生成设计，从而解决一对多映射问题。

在 超材料功能响应逆向设计的背景下，RAG 的贡献如下：

1. **数据效率**。 RAG 使用小数据集即可准确预测高维功能响应，这在评估功能响应的计算成本较高时尤为有益。2. **处理复杂的需求规格说明**。 RAG 能够适应功能响应设计要求的复杂形式，为功能响应逆向设计提供更大的自由度。3. **量化生成方案的不确定性**。 RAG 在生成设计方案时能够提供预测不确定性。不确定性信息有助于避免接受对满足要求缺乏信心的设计，从而提高可信度，尤其是在要求位于分布之外时。

为验证所提出的框架，我们研究了两个功能响应逆向设计任务：（1）具有预设部分通带／阻带特性的声学超材料；（2）具有目标突跳响应的机械超材料 [25]。两项研究所使用的数据集均显著小于现有研究 [25, 28, 33, 44]——声学案例为500个样本，机械案例为1057个样本。这两个任务中的设计要求都难以表述为可微目标，因此难以使用拓扑优化求解。结果表明，RAG能够利用小规模数据高效处理高维功能响应和复杂设计要求。条件似然反映了模型的不确定性，不仅体现了针对给定要求生成设计的可信度，还揭示了不同要求之间的相对难度，从而有助于避免生成不可行设计。通过在公共数据集上与基于神经网络的设计框架进行比较，验证了数据效率 [25]。总体而言，这一轻量且不确定性感知的设计框架为小数据场景下的超材料逆向功能响应设计提供了一种快速且易于实现的方法。

本文其余部分安排如下。第2节介绍功能响应逆向设计问题的数学表述以及RAG的架构。第3节中，我们在声学超材料设计（第3.1节）和机械超材料设计（第3.2节）中验证RAG对小数据集的适用性及其不确定性量化能力。第4节和第5节分别阐述讨论与结论。

2 方法

本节首先对我们关注的功能响应逆向设计问题作出正式描述，并讨论相关的关键挑战，包括功能响应的高维性、复杂的设计要求，以及设计方案可能不存在或不唯一的情况（第2.1节）。为应对这些挑战，我们介绍RAG框架（第2.2节），该框架能够实现高维功能响应的数据高效正向预测（第2.2.1节），并在复杂需求规范下进行不确定性感知逆向设计（第2.2.2节）。

2.1 逆向功能响应设计问题的表述

图2（a）和（b）所示的从超材料单胞设计变量到其功能响应的正向映射可表示为

$$\mathcal{F}(\mathbf{x}) = f, \tag{1}$$

其中， $\mathcal{F}: \Omega_{\mathbf{x}} \to \{f: \Omega_{\mathbf{a}} \to \mathbb{R}\}$。$\mathbf{x}$ 是维度为 $d_{\mathbf{x}}$ 的设计变量向量，表示晶胞几何形状。相应的设计空间为 $\Omega_{\mathbf{x}} \subset \mathbb{R}^{d_{\mathbf{x}}}$。 f 表示关注的功能响应，其定义在查询点$\mathbf{a}$的定义域$\Omega_{\mathbf{a}} \subseteq \mathbb{R}^{d_{\mathbf{a}}}$ 上。例如， f 可表示应力–应变关系 $\sigma(\epsilon)$（其中查询点$\mathbf{a}$为应变 ϵ），或色散关系 $\omega(i,k)$（其中$\mathbf{a} = (i,k)$， i和 k 分别表示带索引和波矢）。$\Omega_{\mathbf{a}}$ 指定相应的关注范围。

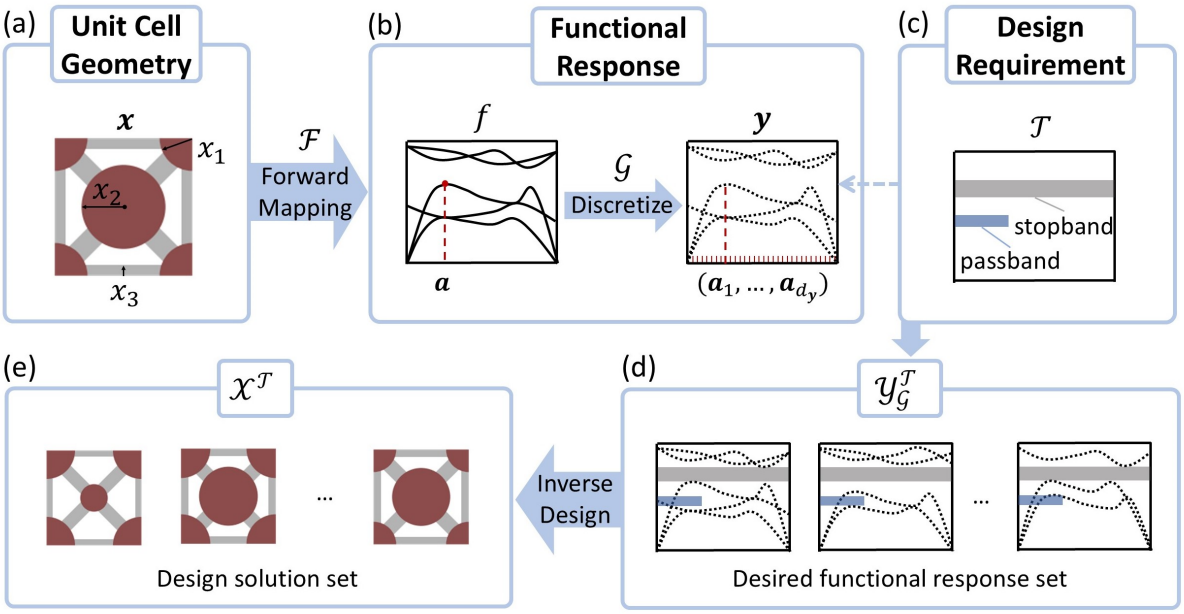

图2：正向映射与逆向设计的表述，展示了（a）晶胞几何形状$\mathbf{x}$、（b）相应的功能响应 f 及其离散形式$\mathbf{y}$、（c）设计要求 $\mathcal{T}$、（d）期望函数响应集 $\mathcal{Y}_{\mathcal{G}}^{\mathcal{T}}$和（e）设计解集 $\mathcal{X}^{\mathcal{T}}$之间的关系。

由于公式 1 中的功能响应 f 是无限维的，难以直接处理，因此我们在其定义域 $\Omega_{\mathbf{a}}$中的一系列查询点（$\mathbf{a}_1, \ldots, \mathbf{a}^{d_{\mathbf{y}}}$）上对其进行离散化，如图 2 （b）所示。在这些查询点处评估 f，即可定义一个离散化算子

$$\mathcal{G}(f) = \mathbf{y}, \tag{2}$$

其中 $\mathcal{G}: \{f: \Omega_{\mathbf{a}} \to \mathbb{R}\} \to \mathbb{R}^{d_{\mathbf{y}}}$。$\mathbf{y}$ 是离散化后的功能响应，其分量为 $y_q = f_{\mathbf{x}}(\mathbf{a}q) \quad q = 1, \ldots, d_{\mathbf{y}}$。离散化算子 $\mathcal{G}$ 的具体定义取决于 $\Omega_{\mathbf{a}}$上的采样策略，该策略必须确保$\mathbf{y}$ 保留原始无限维功能响应 f的必要信息。实际应用中，为了准确表示 f，其高度非线性的区域需要更高的查询点密度。因此，越复杂的功能响应通常需要越多查询点进行离散化，从而导致高维性。在不失一般性的前提下，为简便起见，本文全程采用均匀采样。具体而言，对定义域 $\Omega_{\mathbf{a}}$ 进行均匀采样，其中每个维度 a_j 均离散化为 n_j 个等距点 $(a_{j,1}, \ldots, a_{j,n_j}) \quad j = 1, \ldots, d_{\mathbf{a}}$。

因此，$\mathbf{y}$ 的总维度为 $d_{\mathbf{y}} = \prod_{j=1}^{d_{\mathbf{a}}} n_j$。这表明，$\mathbf{y}$ 的维度会随着$\mathbf{a}$的维数呈指数级增长，因为每增加一个维度，所需的查询点数量就会成倍增加。结合式 1 和式 2，可得到离散化的正向映射：

$$(\mathcal{G} \circ \mathcal{F})(\mathbf{x}) = \mathbf{y}, \tag{3}$$

其中 $\mathcal{G} \circ \mathcal{F}: \Omega_{\mathbf{x}} \to \mathbb{R}^{d_{\mathbf{y}}}$。

设计要求定义了哪些功能响应被视为理想响应。为便于判断某个响应是否符合期望，我们将要求施加于功能响应$\mathbf{y}$的离散形式。我们将期望响应集定义为 $\mathcal{Y}_{\mathcal{G}}^{\mathcal{T}}$所有离散响应$\mathbf{y}$的集合，这些响应是在 $\mathcal{G}$条件下构造的，并满足要求 $\mathcal{T}$（如图2(d)所示）。一种直接的 $\mathcal{T}$方式是规定目标功能响应$\mathbf{y}^*$，例如非线性应力–应变关系 [23,28,54,61,62],，此时 $\mathcal{Y}_{\mathcal{G}}^{\mathcal{T}} = \{\mathbf{y}^*\}$。如果允许目标存在一定偏差，则要求允许响应位于容差带 $\mathcal{Y}_{\mathcal{G}}^{\mathcal{T}} = \{\mathbf{y} \mid |y_i - y_i^*| \leq \delta_i, \quad \forall i = 1, \ldots\ldots, d_{\mathbf{y}}\}$内，其中$\delta_i$ 表示各分量的允许偏差。在其他情况下，要求关注的是功能响应的某些特征

，而非其完整曲线。典型示例包括规定机械超材料的 plateau stress 或吸收能量 [14, 25], ，或针对声学超材料设定特定的完整带隙，其下限为 ω_l，上限为 ω_u（如图2(c)和(d)中的灰色阴影区域所示），可表示为 $\mathcal{Y}_{\mathcal{G}}^{\mathcal{T}} = \{\mathbf{y} \mid y_i \notin [\omega_l, \omega_u], \quad \forall i = 1, \ldots\ldots, d_{\mathbf{y}}\}$。在此类与特征相关的要求下，完整的功能响应并非唯一确定，这意味着整体模式各异的响应仍可能是可行的。因此，$\mathcal{Y}_{\mathcal{G}}^{\mathcal{T}}$中的元素可能存在显著差异，如图2(d)所示。进而，这通常会产生多样化的可行设计解集合。

当指定需求 $\mathcal{T}$时，逆向设计旨在找到相应的设计解集 $\mathcal{X}^{\mathcal{T}}$，该集合包含设计空间 $\Omega_{\mathbf{x}}$ 中所有功能响应满足 $\mathcal{T}$ 的设计变量，如图2(e)所示。基于式3中的离散化正向映射，$\mathcal{X}^{\mathcal{T}}$定义为

$$\mathcal{X}^{\mathcal{T}} = \left\{\mathbf{x} \in \Omega_{\mathbf{x}} \mid (\mathcal{G} \circ \mathcal{F})(\mathbf{x}) \in \mathcal{Y}_{\mathcal{G}}^{\mathcal{T}}\right\}. \tag{4}$$

值得注意的是，如果 $\mathcal{T}$容易满足，式4中的解集 $\mathcal{X}^{\mathcal{T}}$可能包含多个可行设计；如果在给定设计空间内难以达到 $\mathcal{T}$，则该解集也可能为空。设计方案的这种非唯一性或可能不存在的情况，给逆向功能设计带来了重大挑战。

总之，功能响应的高维特性、设计要求的复杂性，以及设计方案的非唯一性或可能不存在的情况，使逆向设计过程极具挑战性。在第2.2节中，我们将介绍所提出的RAG框架，并展示它如何系统地应对这些挑战。

2.2 RAG框架

2.2.1 带不确定性量化的正向预测

根据公式 4，要确定解集 $\mathcal{X}^{\mathcal{T}}$，需要为正向映射$\mathbf{y} = (\mathcal{G} \circ \mathcal{F})(\mathbf{x})$构建代理模型。这里，$\mathcal{F}$ 表示无法更改的物理映射，而离散化算子$\mathcal{G}$ 则根据响应的非线性特征手动指定。构建代理模型最直接的方法是将$\mathbf{x}$ 作为输入，将$\mathbf{y}$ 作为输出。然而，学习到高维输出空间的映射会受到维数灾难的影响，因此需要大量数据才能实现有效的泛化。为缓解这一问题，现有方法通常采用：(i) 降维方法（例如主成分分析 [44, 64]），以获得$\mathbf{y}$的低维形式；(ii) 将$\mathbf{y}$ 手动参数化为紧凑形式 [28], ；或 (iii) 将$\mathbf{y}$ 划分为多个片段，并针对每个片段训练独立模型 [33]。尽管这些方法有效，但仍需要大型数据集和复杂的模型架构 [25, 28, 33]。此外，由于学习到的映射 $(\mathcal{G} \circ \mathcal{F})$ 与固定的离散化方式 $\mathcal{G}$绑定，因此无法在 $\Omega_{\mathbf{a}}$ 内任意查询点进行预测。

这两个问题可以通过将公式 1 中的原始映射重新表述为以下形式来解决：

$$\phi(\mathbf{x}, \mathbf{a}) = \mathcal{F}(\mathbf{x})(\mathbf{a}), \tag{5}$$

其中 ϕ： $\Omega_{\mathbf{x}} \times \Omega_{\mathbf{a}} \to \mathbb{R}$。如图 3 (a) 所示，将设计变量 $\mathbf{x}$ 和查询点 $\mathbf{a}$ 同时作为输入后，训练得到的代理模型 $\hat{}\,\phi$ 可以预测任意位置 $\mathbf{a} \in \Omega_{\mathbf{a}}$处的响应。这种离散化不变性使功能响应能够在不同离散化之间迁移。对于给定的离散化算子 $\mathcal{G}$，对应的离散响应为 $\mathbf{y} = (\mathcal{G} \circ \phi)(\mathbf{x})$，其各个分量为 $y_q = \phi(\mathbf{x}, \mathbf{a}_q)$。此外，由于输出维度为 1，代理模型可以用更少的数据进行训练，并更高效地学习正向映射 [65]。

在 RAG 中，使用随机森林来近似公式 5 中的映射。随机森林是一种集成学习方法，它针对同一任务训练多棵决策树，并汇总它们的预测结果以提高准确性。与神经网络相比，随机森林能够有效处理小数据集，所需的超参数调优较少，训练效率也更高。此外，随机森林能够自然地处理混合数据类型，因为 $\mathbf{x}$ 可能同时包含数值变量和分类变量。使用 N 棵树训练随机森林后，每棵树 $\hat{}\phi^{(n)}$ 都会预测其自身的离散功能响应 $\hat{}\,\mathbf{y}^{(n)} = (\mathcal{G} \circ \hat{}\phi^{(n)})(\mathbf{x})$ $n = 1, \ldots, N$。随后，将集成 $\hat{}\,\phi$ 定义为所有树 $\hat{}\phi(\mathbf{x}, \mathbf{a}) = \frac{1}{N}\sum_{n=1}^{N} \hat{}\phi^{(n)}(\mathbf{x}, \mathbf{a})$ 的平均值，其对应的离散预测为 $\hat{}\,\mathbf{y} = (\mathcal{G} \circ \hat{}\phi)(\mathbf{x})$。

除了通过平均来提高准确性之外，随机森林的集成结构还提供了预测不确定性，该不确定性源于各个树输出之间的差异 [58, 66], ，如图 3 (a) 所示。树预测的方差反映了树之间的分歧程度：方差较小表示预测一致且置信度较高，而方差较大则意味着不确定性更高。为简化计算，我们忽略树输出之间的相关性，并使用样本方差估计预测不确定性：$\hat{}\,\mathbf{s}^2(\mathbf{x}) = \frac{1}{N}\sum_{n=1}^{N}\left((\mathcal{G} \circ \hat{}\phi^{(n)})(\mathbf{x}) - (\mathcal{G} \circ \hat{}\phi)(\mathbf{x})\right)^{\circ 2}$，其中 $(\cdot)^{\circ 2}$ 表示逐元素平方。所得的 $\hat{}\,\mathbf{s}^2(\mathbf{x}) \in \mathbb{R}^{d_{\mathbf{y}}}$ 是

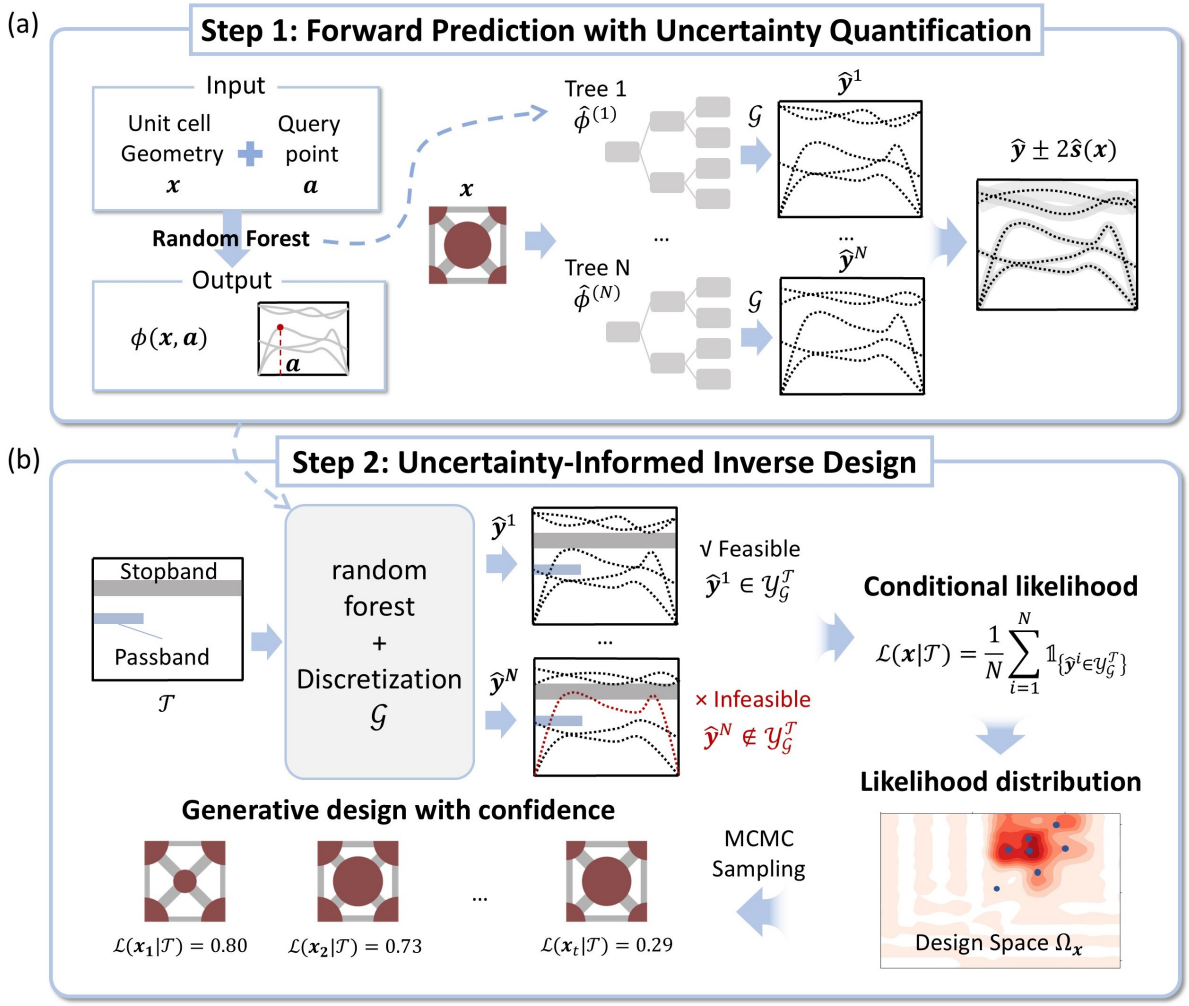

图3：RAG框架示意图。(a) 步骤1：对$\mathbf{y}$ 进行正向预测，基于$\mathbf{x}$ 并进行不确定性量化。(c) 步骤2：给定设计要求 $\mathcal{T}$，进行不确定性感知逆向设计。

一个向量值不确定性估计，其中每个条目 $\hat{s}_q^2(\mathbf{x})$ 用于量化相应查询点 $\mathbf{a}_q$处的预测方差。这提供了对模型置信度的估计，使我们能够评估逆向设计过程中所用预测的可信度。

2.2.2 不确定性感知逆向设计

如图2 (c)和(d)所示，指定设计要求 $\mathcal{T}$后，即可定义期望函数响应集$\mathcal{Y}_{\mathcal{G}}^{\mathcal{T}}$。对于给定的设计变量$\mathbf{x}$，随机森林中的每棵决策树$\hat{\phi}^{(n)}$ 都会判断$\mathbf{x}$ 是否满足要求（如图3 (b)所示）。根据将$\mathbf{x}$ 判定为可行的树所占比例，我们可以估计条件似然，以量化$\mathbf{x}$ 满足 $\mathcal{T}$的可能性：

$$\mathcal{L}(\mathbf{x} \mid \mathcal{T}) = \frac{1}{N} \sum_{n=1}^{N} \mathbb{1}_{\{(\mathcal{G} \circ \hat{\phi}^{(n)})(\mathbf{x}) \in \mathcal{Y}_{\mathcal{G}}^{\mathcal{T}}\}}, \tag{6}$$

其中，$\mathbb{1}$ 是指示函数：当树 n 给出的预测响应$\hat{\mathbf{y}}^{(n)} = (\mathcal{G} \circ \hat{\phi}^{(n)})(\mathbf{x})$属于期望响应集 $\mathcal{Y}_{\mathcal{G}}^{\mathcal{T}}$时返回1，否则返回0。公式6中的条件似然反映了模型预测的置信度。投票认为$\mathbf{x}$ 可行的树越多，模型对$\mathbf{x}$ 满足要求的置信度就越高。

我们指出，式 6 中计算 $\mathcal{L}(\mathbf{x} \mid \mathcal{T})$ 仅依赖正向预测，因此无需额外的机器学习模块来学习逆映射。此外，似然直接以 $\mathcal{T}$ 为条件。现有

生成模型通常学习条件分布 $p(\mathbf{x}|\mathbf{y})$，其条件直接取决于响应 $\mathbf{y}$，而非需求 $\mathcal{T}$ [50–53, 55, 56]。因此，这些方法通常需要手动选择一个代表性响应 $\mathbf{y} \in \mathcal{Y}_{\mathcal{G}}^{\mathcal{T}}$，并将其提供给生成模型以生成设计方案。当所需的功能响应无法明确定义时，这可能颇具挑战性。此外，这种启发式选择也不可避免地会引入偏差，从而限制方案的多样性；尤其是当 $\mathcal{Y}_{\mathcal{G}}^{\mathcal{T}}$包含高度多样化的响应模式时，如图2所示。相比之下，在我们的框架中，只要需求 $\mathcal{T}$ 能够表述为定义 $\mathcal{Y}_{\mathcal{G}}^{\mathcal{T}}$的约束，就可以直接估计相应的条件似然以进行设计生成。这使设计要求的指定方式具有很大的灵活性，正如第 3.1 节和第 3.2 节将进一步展示的那样。

基于似然函数，通过从设计空间上的似然分布 $\mathcal{L}(\mathbf{x} \mid \mathcal{T})$ 中采样，即可一次性获得满足需求的设计方案。由于 $\mathcal{L}(\mathbf{x} \mid \mathcal{T})$ 的形状可能会随 $\mathcal{T}$ 而呈现复杂变化，我们采用 Metropolis-Hastings 算法 [67]——一种广泛使用的马尔可夫链蒙特卡洛（MCMC）方法——从该似然分布中生成样本。在每次迭代中，算法根据指定的提议分布提出一个新状态，并依据 Metropolis-Hastings 接受概率决定是否接受该状态。该过程构造出一条马尔可夫链，其平稳分布与 $\mathcal{L}(\mathbf{x} \mid \mathcal{T})$ 一致。
在第 k 次迭代中，新提出的状态$\mathbf{x}_{k+1}$ 从以下分布中抽取：

$$\mathbf{x}_{k+1} \sim \mathcal{N}\big(\mathbf{x}_k,\, \mathrm{diag}(\boldsymbol{\sigma}^2)\big)\,, \tag{7}$$

其中，提议标准差定义为

$$\boldsymbol{\sigma} = c_0 d_{\mathbf{x}}^{-1/2} |\mathbf{u} - \mathbf{l}|, \tag{8}$$

其中，$\mathbf{l}$ 和$\mathbf{u}$ 分别表示设计变量的下界和上界。根据变量取值范围缩放提议标准差，可使提议步长自然适应设计空间。 c_0 用作控制整体步长的缩放参数。由于高斯提议具有对称性，接受概率可表示为

$$\alpha(\mathbf{x}_k, \mathbf{x}_{k+1}) = \min(1, \frac{\mathcal{L}(\mathbf{x}_{k+1} \mid \mathcal{T})}{L(\mathbf{x}_k \mid \mathcal{T})}). \tag{9}$$

似然值越高的设计变量被采样的概率越大，这意味着 RAG框架倾向于在代理模型置信度较高的区域生成设计方案。此外，每个生成方案对应的似然值还提供了其可信度的定量指标。如果需求 $\mathcal{T}$在设计空间 $\Omega_{\mathbf{x}}$内本质上难以满足，甚至不可能满足，那么整个域内的似然分布就会相对较低。因此，这种似然分布可作为代理指标，用于量化不同设计要求的任务难度；在采用深度生成建模进行逆向设计时，通常无法获得这一指标。

3 结果

本节评估 RAG 在声学超材料色散关系逆向设计（第 3.1 节）以及机械超材料应力–应变关系逆向设计（第 3.2 节）中的性能。对于这两项任务，随机森林均配置为包含 100 棵树，节点分裂所需的最小样本数设为 2，叶节点的最小样本数设为 1。节点分裂准则采用基尼不纯度 [68] 。为应对不同程度的非线性，我们针对每项任务调整最大树深度，以平衡预测能力与计算效率。

3.1 按需设计具有部分通带/阻带的声学超材料

3.1.1 问题陈述

利用在指定频率范围或方向上控制声波/弹性波传播的非凡能力，声学超材料已应用于振动控制 [69], 信号处理 [40], 和能量收集 [70]等众多领域。一般而言，要设计具有目标功能的声学超材料，关键在于逆向设计能够实现预期色散关系 [71, 72]的单胞。因此，公式 1 中的功能响应 f是频率函数 $\omega(i,\ \mathbf{k})$，其查询点为$\mathbf{a} = (i,\ \mathbf{k})$，其中 i表示带索引，$\mathbf{k}$ 表示波矢。尽管色散关系能够全面描述单胞晶格所支持的波动特性，但其高维性往往促使现有研究将其简化为少数基于频率的特征（例如完整带隙范围） [24, 73, 74]。这种简化限制了声学超材料实现更复杂波动操控的潜力。在本研究中，我们通过一个声学超材料案例研究证明，RAG 能够实现完整色散关系的逆向设计。

受 Sun 等人提出的微结构超材料设计启发 [75], , 该设计已通过实验验证具有足够丰富的动态特性空间，因此我们在本案例研究中考虑一种二维单胞设计

，其几何图案与之类似，如图4(a)所示。在支撑方形晶胞的中心和角部添加圆形微惯性，支杆半径为 r_{strut}。位于支撑结构中心的微惯性半径为 r_{center}，而位于方形晶胞角部的微惯性半径为 r_{corner}。设计空间 $\Omega_{\mathbf{x}}$ 由 $4 \leq r_{\text{strut}} \leq 6.41$、$\sqrt{2}r_{\text{strut}} \leq r_{\text{center}} \leq 20$ 和 $(\sqrt{2}+1)r_{\text{strut}} \leq r_{\text{corner}} \leq 20$ 定义（单位：mm）。r_{center} 和 r_{corner} 的下限根据 r_{strut} 确定，以确保支撑杆不会完全覆盖中心圆和角部圆。晶胞尺寸设为 $l = 60$ mm。杨氏模量 E、密度 ρ 和泊松比 ν 分别设为 70 GPa、2700 kg/m^3, 和 0.33。为展示需求规格说明的灵活性，我们考虑由多个部分通带和阻带组成的设计要求（即在特定波矢范围内定义的通带和阻带）。

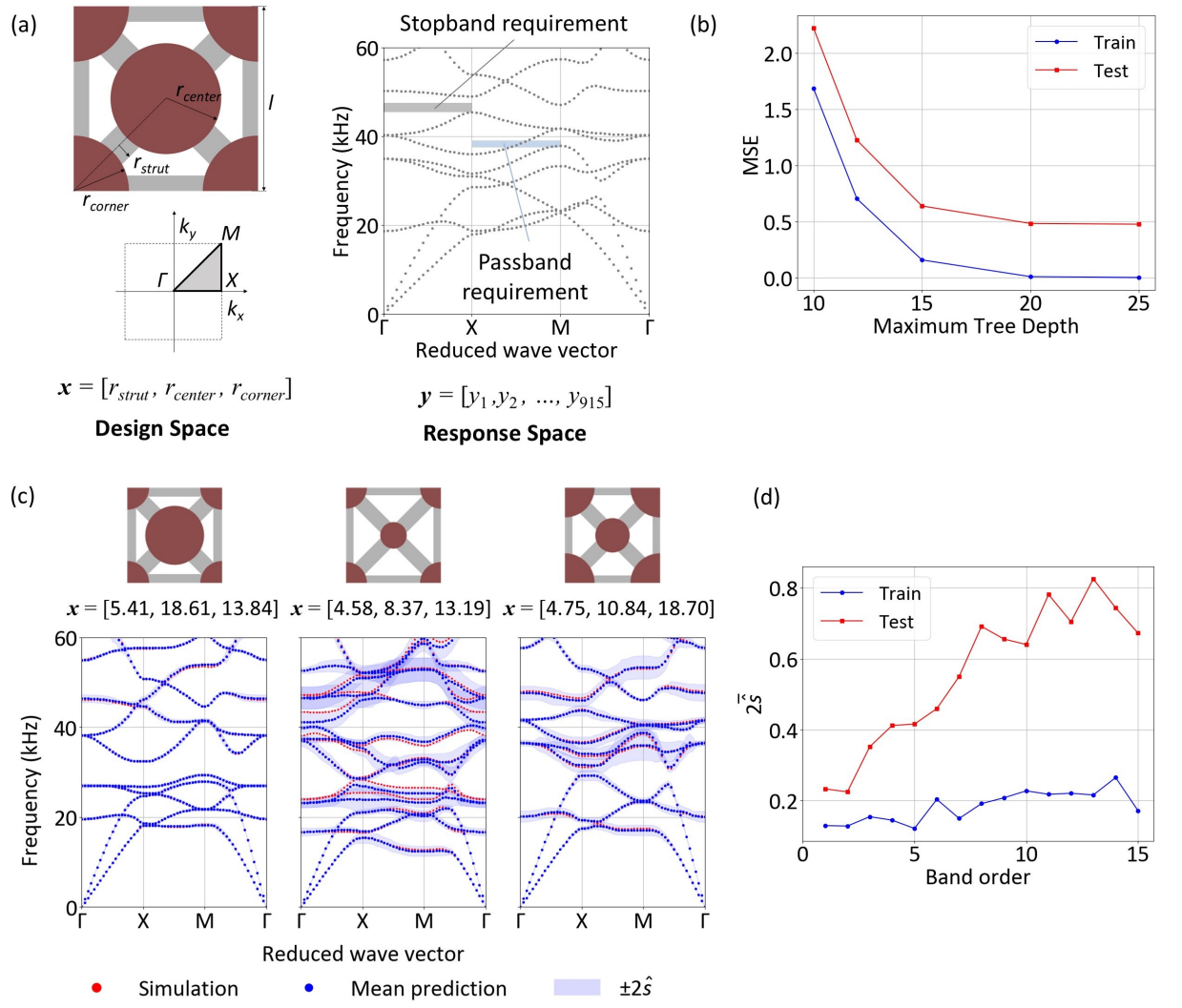

图 4：（a）案例研究中二维声学超材料的设计空间和响应空间。晶胞结构由三个几何参数 $\mathbf{x} = [r_{\text{strut}}, r_{\text{center}}, r_{\text{corner}}]$ 指定。（b）不同最大树深度下随机森林的正向预测精度，以均方误差（MSE）量化。（c）对三个测试样本进行带不确定性的正向预测。不确定性以 ± 2 标准差量化，记为 $\pm 2\hat{s}$。（d）不同带阶下色散关系的平均不确定性（以 $2\bar{\hat{s}}$ 量化）。

3.1.2 数据获取

弹性波传播的控制方程可表示为 [38, 76]:

$$\rho \ddot{\mathbf{u}} = \nabla(\lambda + \mu)\nabla \cdot \mathbf{u} + \nabla \cdot \mu \nabla \mathbf{u}, \tag{10}$$

其中，ρ 表示材料密度，$\mathbf{u} = \{u_x, u_y\}^\top$ 是位移向量，λ 和 μ 是拉梅系数。为对声学超材料单胞的无限平铺结构进行均匀化，采用周期性布洛赫-弗洛凯边界条件：

$$\mathbf{u}(\mathbf{r}, \mathbf{k}) = \tilde{\mathbf{u}}(\mathbf{r})\, e^{i(\mathbf{k}^{\mathrm{T}}\mathbf{r} + \omega t)}, \tag{11}$$

其中，$\mathbf{r} = \{x, y\}$ 表示位置向量，$\tilde{\mathbf{u}}$ 表示布洛赫位移向量，而$\mathbf{k} = (k_x, k_y)$表示波矢。将公式 11 代入公式 10，并基于有限元法进行离散化，可得到以下特征值问题：

$$\left(\mathbf{K}(\mathbf{k}) - \omega^2 \mathbf{M}(\mathbf{k})\right) \tilde{\mathbf{u}} = 0, \tag{12}$$

其中，$\mathbf{K}$ 和$\mathbf{M}$ 分别为布洛赫约化全局刚度矩阵和质量矩阵。由于$\mathbf{K}$ 和$\mathbf{M}$均依赖于波矢$\mathbf{k}$，因此得到的特征频率 ω 也都是$\mathbf{k}$的函数。

通过在设计空间内进行拉丁超立方抽样，生成了共 500 组几何参数集$\mathbf{x} = (r_{\text{strut}}, r_{\text{center}}, r_{\text{corner}})$。在 COMSOL Multiphysics 中进行布洛赫波分析，以求解公式 12 并获得色散关系。沿高对称路径$\Gamma \to X \to M \to \Gamma$对波矢进行采样来计算色散关系，该路径沿第一不可约布里渊区（IBZ）的边界延伸，如图4(a)所示。不可约布里渊区边界均匀离散为 61 个点。在每个采样波矢处，提取前 15 个特征频率，从而得到维度为 $d_{\mathbf{y}} = 915$ 的离散函数响应向量$\mathbf{y}$。

3.1.3 正向预测

数据获取后，随机抽取 400 个样本作为训练数据，用于训练随机森林以学习正向映射，其余 100 个样本则留作测试。根据公式 5 中重新表述的映射，我们使用设计变量 $\mathbf{x}$ 以及查询点 $\mathbf{a} = (i, k)$ 预测对应的频率，即分别预测响应 $\mathbf{y}$ 的各个分量。因此，随机森林的输入维度为 5，输出维度则降至 1。相应地，重新表述后的输入–输出对数量增加至原来的 $d_{\mathbf{y}}$ 倍，训练集和测试集分别包含 366,000 对和 91,500 对。通过调整最大树深度，在拟合能力与计算成本之间取得平衡。如图 4 (b) 所示，当树深度超过 20 时，训练集和测试集上的均方误差（MSE，定义为 $\frac{1}{d_{\mathbf{y}}}(\mathbf{y} - \hat{\mathbf{y}})^\top(\mathbf{y} - \hat{\mathbf{y}})$）不再下降。因此，将最大树深度设为 20，此时训练集和测试集上的 MSE 分别为 0.0108 和 0.4839。

测试集中三个样本的正向预测结果如图 4 (c) 所示。蓝色点表示随机森林的平均预测值，红色点表示通过 COMSOL 仿真获得的真实值。蓝色阴影区域表示 ± 2 标准差范围（记为 $\pm 2\hat{s}$），即各个决策树预测结果的范围，可用于衡量第 2.2.1 节中所述的预测不确定性。总体而言，真实值通常落在 ± 2 标准差界限内。由于高阶特征频率通常具有更大的变化，因此往往更难拟合，其预测精度低于低阶频率 [33]。如图 4 (d) 所示，量化的不确定性反映了这一点。蓝色和红色实线分别表示训练集$\bar{}$和测试集在不同带阶上的预测不确定性，量化指标为两个标准差的平均值 $2\hat{s}$。可以观察到，随着带阶增加，标准差也随之增大。图 4(c) 中的单个预测结果也呈现出类似趋势。不确定性与预测精度之间的一致性表明，方差可作为预测不确定性的合理度量。

3.1.4 逆向设计

训练正向模型后，我们使用同一模型进行逆向设计。为评估逆向设计的性能，我们设定了 10 项部分通带/阻带要求 $\mathcal{T}_k$，如表1所示。每项要求包含 2–3 个区段，每个区段指定一个波矢范围和一个频率范围。波矢区间的构造方式为：从 $\{\Gamma, X, M\}$ 中随机选择起点，从 $\{X, M, \Gamma\}$ 中随机选择终点，并要求起点沿不可约布里渊区边界的位置在终点之前。频率范围在 [20,50] 千赫兹内抽样，同时满足预设的带宽约束（1–5 千赫兹），并确保各区段的频率区间互不重叠。创建区段后，将每个区段随机指定为阻带或通带。对于每项要求，我们通过在设计空间上的相应似然分布中抽样，使用所提出的方法生成 30 个设计变量。需要注意的是，当模型认为某项要求无法实现时，设计空间内的似然可能处处为零。由于无法从此类似然分布中抽样生成设计变量，因此排除这些情况。

图5(a)和(b)分别展示了 RAG 针对要求 $\mathcal{T}_1$ 和 $\mathcal{T}_2$生成的设计变量。蓝色阴影区域表示目标通带，灰色阴影区域表示目标阻带。对于每项要求，我们展示似然最高的五个设计方案。对于 $\mathcal{T}_1$，由于施加的约束相对较少，可行设计很容易识别，

表1：逆向设计所用的设计要求。每个频带均由一个波矢范围和一个频率范围（单位：千赫兹）指定。

需求	阻带	通带
$\mathcal{T}_1$	$[\Gamma-X]$, [45.48, 47.52]	$[X-M]$, [37.49, 39.12]
$\mathcal{T}_2$	$[M-\Gamma]$, [26.20, 27.94] $[X-M-\Gamma]$, [47.85, 49.91]	$[M-\Gamma]$, [22.83, 25.49]
$\mathcal{T}_3$	$[X-M-\Gamma]$, [26.57, 28.22] $[\Gamma-X]$, [21.91, 24.34]	$[X-M]$, [45.27, 47.46]
$\mathcal{T}_4$	$[\Gamma-X]$, [47.04, 48.39]	$[\Gamma-X]$, [43.09, 45.96]
$\mathcal{T}_5$	$[M-\Gamma]$, [45.20, 46.21]	$[\Gamma-X-M-\Gamma]$, [28.42, 30.15]
$\mathcal{T}_6$	$[M-\Gamma]$, [44.96, 47.17]	$[\Gamma-X]$, [26.87, 28.72] $[X-M]$, [32.96, 34.03]
$\mathcal{T}_7$	$[M-\Gamma]$, [26.95, 28.12] $[\Gamma-X-M]$, [32.14, 33.41]	None
$\mathcal{T}_8$	$[M-\Gamma]$, [27.66, 28.82] $[\Gamma-X-M]$, [35.54, 38.05]	None
$\mathcal{T}_9$	$[\Gamma-X-M]$, [37.71, 40.34] $[X-M]$, [24.94, 26.90]	None
$\mathcal{T}_{10}$	$[\Gamma-X]$, [30.37, 33.22] $[\Gamma-X]$, [46.33, 48.07]	None

因此，RAG 生成的全部五个设计方案都具有较高的似然（均高于 0.90），且经仿真确认均满足要求。由于许多色散关系都能满足这些约束，相应的设计方案分布在设计空间的多个区域中。RAG 也很好地捕捉到了这一特征，因为生成的单元胞几何结构及其对应的色散关系明显不同。这表明，对于单一要求，RAG 能够通过生成多样化的方案，有效应对逆问题的一对多特性。相比之下，对于引入更多约束的 $\mathcal{T}_2$，生成设计的似然相对较低（均低于 0.4），且没有一个满足要求。总体而言，似然与可行性密切相关：似然较高的设计更有可能满足要求。利用这一信息，我们可以评估给定逆向设计要求的复杂程度，并通过排除低似然设计来降低风险。

图6(a)展示了全部10项预设要求的概况。灰色条形表示满足各项要求的训练样本数量。总体而言，数据集中可行样本越多，在设计空间内实现相应要求就越容易。因此，诸如 $\mathcal{T}_1$、 $\mathcal{T}_4$和 $\mathcal{T}_5$ 这样的要求相对容易实现，而 $\mathcal{T}_2$、 $\mathcal{T}_3$、 $\mathcal{T}_8$和 $\mathcal{T}_9$ 则更难实现，因为训练数据集中不存在可行样本。蓝色曲线所示的平均似然有效地反映了这种难度：较易实现的要求具有较高的平均似然，而缺少可行样本的要求则呈现较低的数值。这一趋势与图5一致，证实似然是衡量要求难度的有效指标。

为说明 RAG 能够生成满足需求的设计变量，我们计算满意率（定义为满足需求的生成设计所占的比例，以黄色曲线表示）以及平均重叠率，后者衡量生成响应与需求之间的平均频率范围重叠。满意率基于二元指标计算（设计变量要么完全满足需求，要么不满足），而重叠率则提供了一种更宽松的度量方式，可反映部分满足的情况。根据定义，平均重叠率的下界为满意率，这揭示了模型何时会生成虽不可行、但仍接近需求的设计变量。对于较容易满足的需求（$\mathcal{T}_1$、 $\mathcal{T}_4$、 $\mathcal{T}_5$），满意率和重叠率都较高；而对于较难满足的需求（$\mathcal{T}_2$、 $\mathcal{T}_3$、 $\mathcal{T}_8$、 $\mathcal{T}_9$），这两项指标都较低。重要的是，只有通过明确的可行性验证才能获得满意率和重叠率。相比之下，我们的框架利用似然信息，在验证之前标记出较不可能满足需求的设计变量。

图6(b)展示了以对应需求为条件，对全部300个生成设计的似然进行核密度估计（KDE）的结果。可以观察到，不可行设计通常具有较低的似然，而可行设计则对应较高的似然值。因此，通过应用似然度阈值筛除低似然设计，可以提高生成设计的整体满意率。图6(c)展示了选择率、满意率、重叠率与似然度阈值之间的关系。随着似然度阈值提高，低似然设计逐步被筛除，选择率随之下降，而保留的设计的满意率和重叠率则不断提高。这也表明，可根据不同任务可接受的置信水平灵活调整似然度阈值。

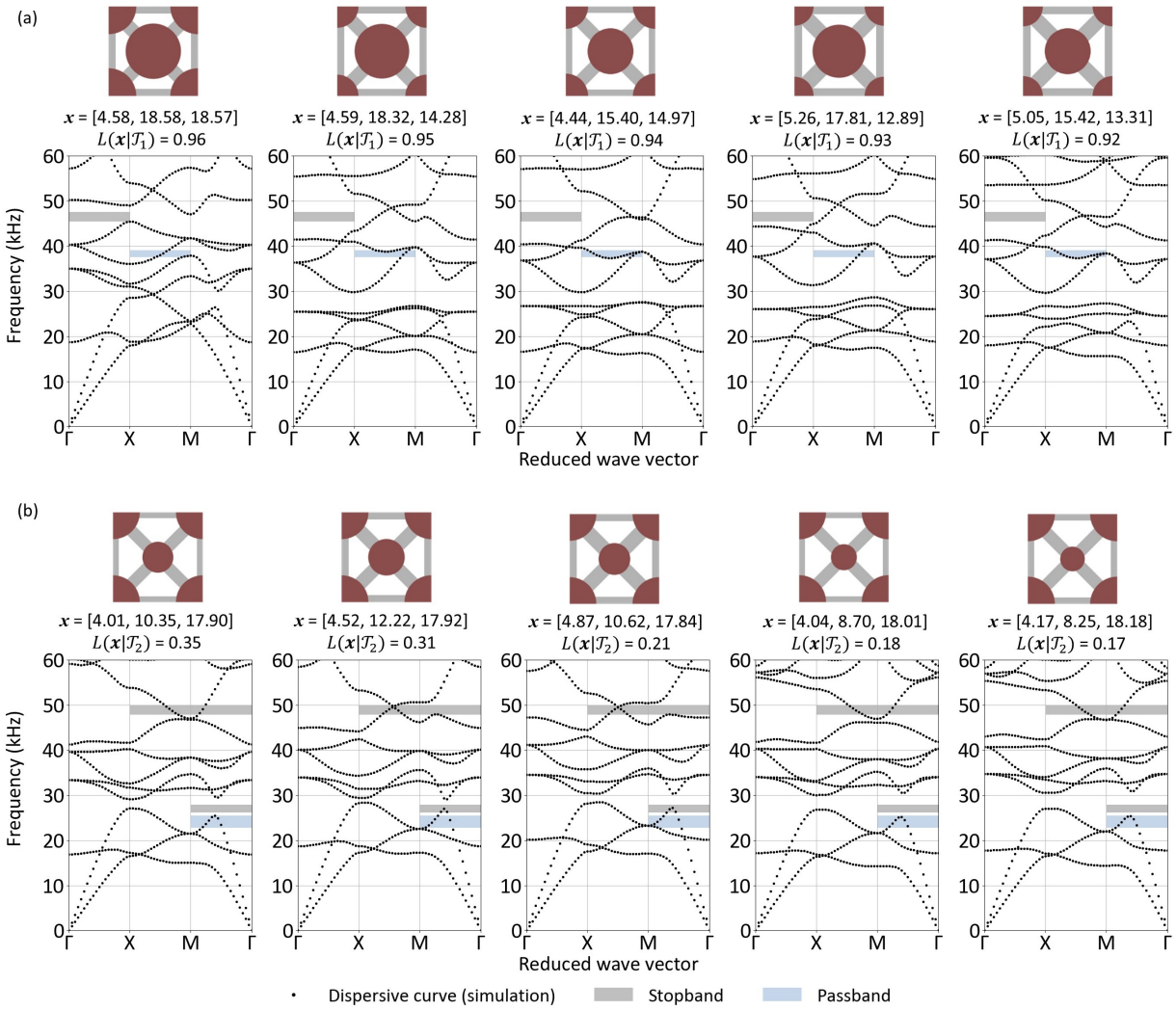

图5：针对两种不同要求 $\mathcal{T}_1$ （a）和 $\mathcal{T}_2$ （b）的声学超材料逆向设计结果。蓝色阴影区域表示部分通带，灰色阴影区域表示部分阻带。对于每项要求，图中展示了似然最高的前5种单元胞几何结构及其对应的色散关系。对于 $\mathcal{T}_1$，五种设计均满足该要求。对于 $\mathcal{T}_2$，没有一种设计满足该要求。全部十项要求的详细规格见表1。

3.2 设计具有预期跳跃响应的机械超材料

3.2.1 问题陈述

机械超材料因其单元胞结构所带来的前所未有的力学性能而受到广泛研究 [5]。这些性能通常通过其应力–应变关系来表征 [25, 28, 44, 59],，该关系可表示为函数 $\sigma(\epsilon)$，其中查询点$\mathbf{a}$：$=\epsilon$ 表示应变。为展示 RAG 在应力–应变关系逆向设计方面的能力及其数据效率优势，我们将其应用于 Chai等人提出的一个现有机械超材料系统 [25], ，仅使用原始训练数据集的40%。

在此任务中，设计空间由五个变量定义：倾斜梁的水平长度 l、悬臂梁与倾斜梁之间的角度 α、倾斜梁的宽度 w，以及列数 n_{col} 和行数 n_{row}，如图7 (a)所示。 n_{col} 和 n_{row} 只能取 1、2、3 或 4 的整数值。应力–应变关系均匀离散为 31 个点；由于第一个点的应力始终为零，我们只预测其余 30 个点，因此离散功能响应向量 $\mathbf{y}$ 的维度为 30。在本研究中，我们关注 [25], 中介绍的可重构软执行器逆向设计任务，在该任务中，通过简单的

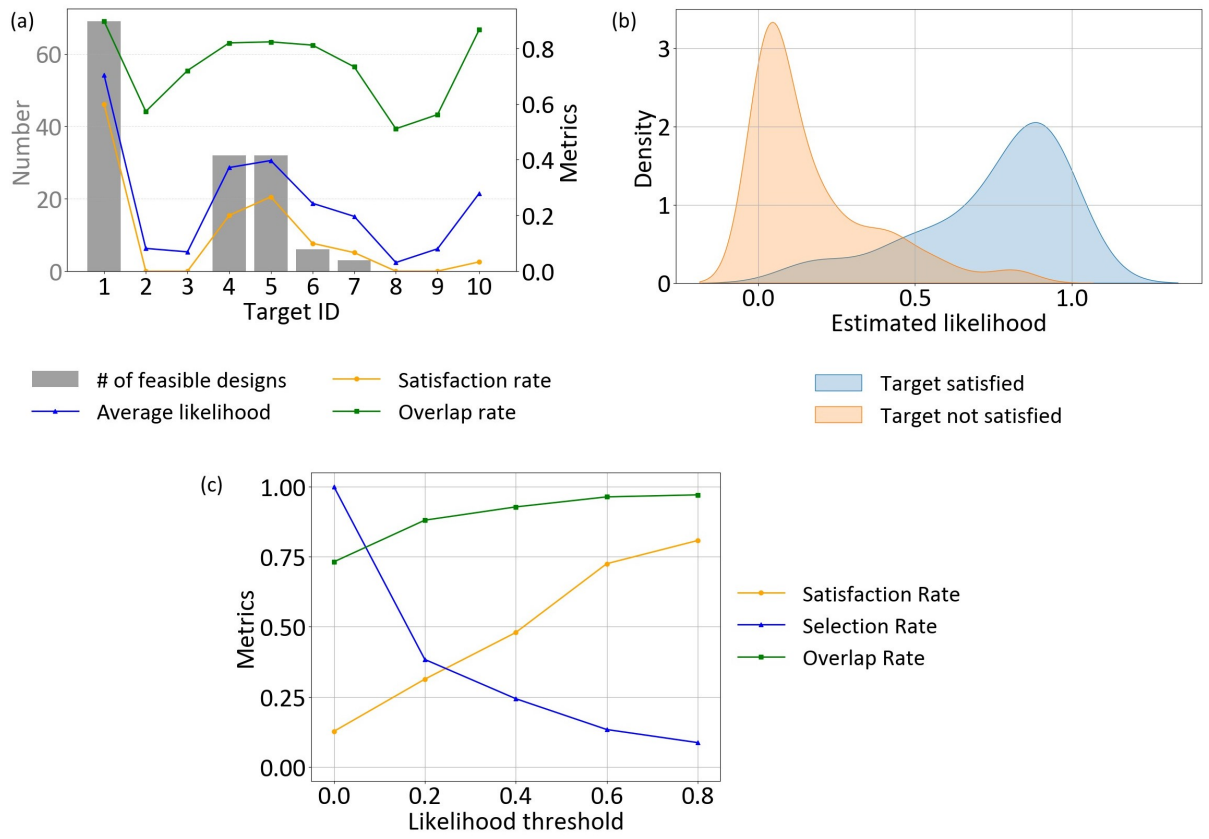

图6：根据表1中的10项规定的部分通带/阻带要求生成的全部300个设计变量的统计汇总。（a）10项要求的概览。灰色柱表示满足相应要求的训练样本数量。蓝色、黄色和绿色曲线分别表示平均似然、满意率和平均重叠率。（b）全部300个生成设计的估计似然的核密度估计（KDE）。（c）不同采样阈值下生成设计的满意率、重叠率和选择率。

气动输入即可实现不同的驱动模式。实现这一功能的关键是找到能够在特定力作用下发生突跳并提供所需行程的晶胞设计。因此，如图7(a)所示，对应力–应变关系施加的设计要求以目标突跳阈值力和行程的形式指定。

3.2.2 正向建模

与声学案例中的正向预测（第3.1.3节）类似，我们将设计变量$\mathbf{x}$和查询点$\mathbf{a}$——在本任务中对应于应变ϵ——共同作为输入，以预测相应的应力值。针对该任务，采用最大树深度为15的随机森林回归器。公开的基础数据集包含2065个训练样本和231个测试样本。为评估RAG的数据效率，我们在保持测试集不变的情况下，按照不同的下采样比例对原始训练集进行下采样。随后使用每个子集分别训练一个随机森林模型，其对应的精度如图7（b）所示。结果表明，在数据有限的情况下，随机森林仍然有效：当仅使用10%的训练数据时，模型在训练集和测试集上的平均相对误差仍分别达到2.70%和7.37%。相比之下，原研究采用神经网络（NN）时获得了更低的误差，训练集和测试集上的误差分别为1.08%和2.59%，但需要使用完整数据集。该结果表明，所提出的RAG框架仅需显著更少的训练数据，即可达到接近神经网络（NN）的预测精度，从而凸显其数据效率优势。在后续任务中，我们选用使用40%的训练数据（826个样本）训练的模型，该模型在训练集和测试集上的误差分别为2.09%和5.14%。

图7(c)分别展示了无突跳、包含一个突跳波以及包含两个突跳波的测试样本的正向预测及其不确定性量化结果。可以观察到，波动越大的曲线

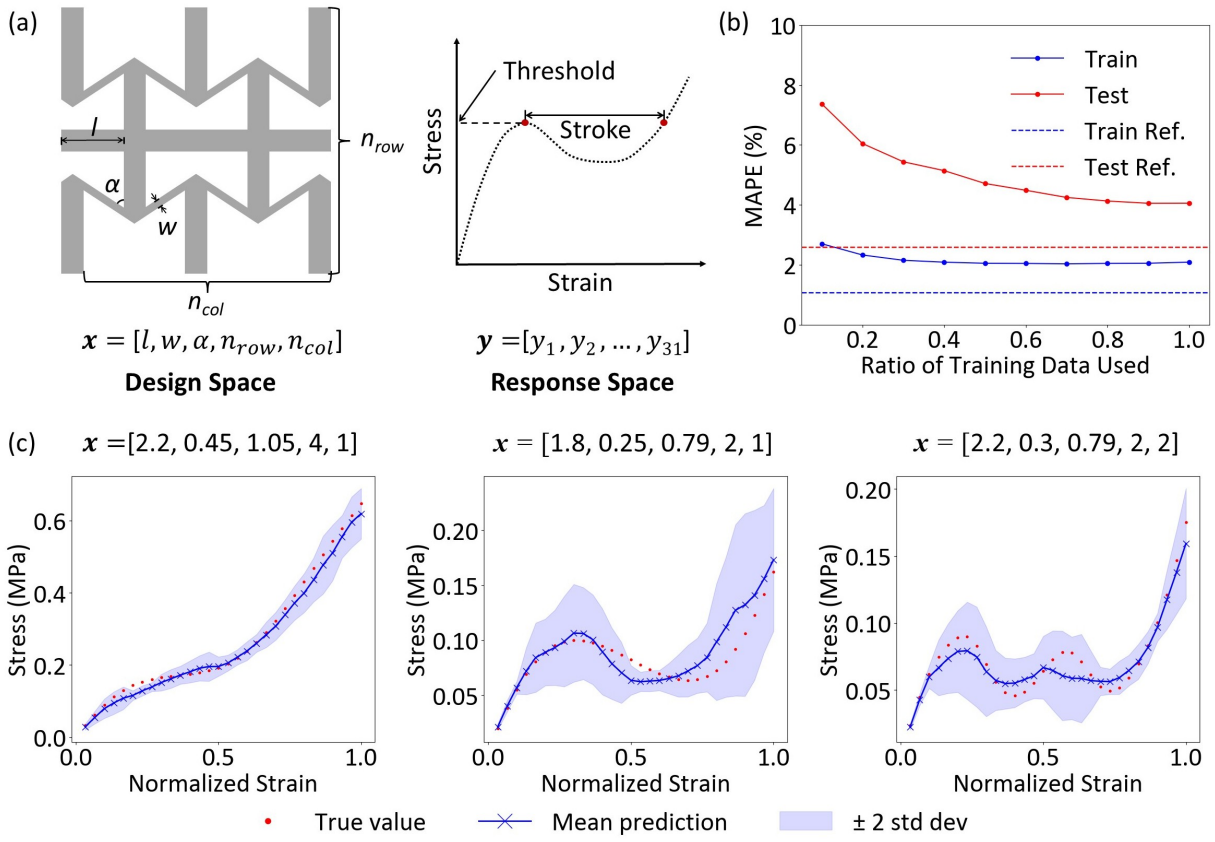

图 7：机械超材料的设计问题陈述与正向建模。(a) 机械超材料的设计空间和响应空间。(b) 以平均绝对百分比误差（平均绝对百分比误差）量化的预测精度与用于模型训练的数据比例之间的关系。蓝色和红色实线分别表示随机森林模型的训练精度和测试精度。蓝色和红色虚线分别表示文献 [25]中报告的神经网络训练精度和测试精度。(c) 分别对不发生突跳、出现一个突跳波和出现两个突跳波的测试样本进行正向预测和不确定性量化。

其不确定性范围越宽。真实值均处于2个标准差带内，表明预测准确度合理，且不确定性量化具有实际意义。

3.2.3 逆向推断

设计要求设定为目标阈值力 0.2 MPa 和目标行程 1.3 mm。我们通过为这些值引入容差来放宽要求——也就是说，生成设计的阈值力和行程值允许处于目标值规定百分比范围内。容差越小，要求越难满足。此处，我们考虑两种不同的容差水平，即 ±25% 和 ±15%，以研究要求的难度如何影响估计似然。

图 8(a) 和 (b) 展示了容差水平分别为 ±25% 和 ±15% 时设计空间上的估计似然分布。离散设计变量 n_{col} 和 n_{row} 将五维设计空间划分为十六个三维子空间。此处我们以对应于 $n_{col} = 4$ 和 $n_{row} = 2$ 的子空间为例进行可视化。为考察估计似然的准确性，我们还绘制了随机森林观测到的训练样本（即缩减后的训练数据集）。灰色点表示不满足要求的样本，蓝色点表示满足要求的样本。可以观察到，估计似然较高的区域位于可行样本附近，而较低的似然则出现在被不可行样本包围的区域。由于关于阈值和行程的要求并未对具体的应力—应变关系形状施加约束，不同类型的应力—应变关系均可能满足这些要求。因此，在整个设计空间中可以观察到多个高似然区域。此外，随着容差变得更严格，整体似然降低，但似然分布的形态保持相似，这是因为目标未发生变化。这些观察结果表明，似然

由 RAG 估计的分布能够有效反映可行设计解的分布，并且似然的大小也可作为衡量设计要求难度的可靠指标。

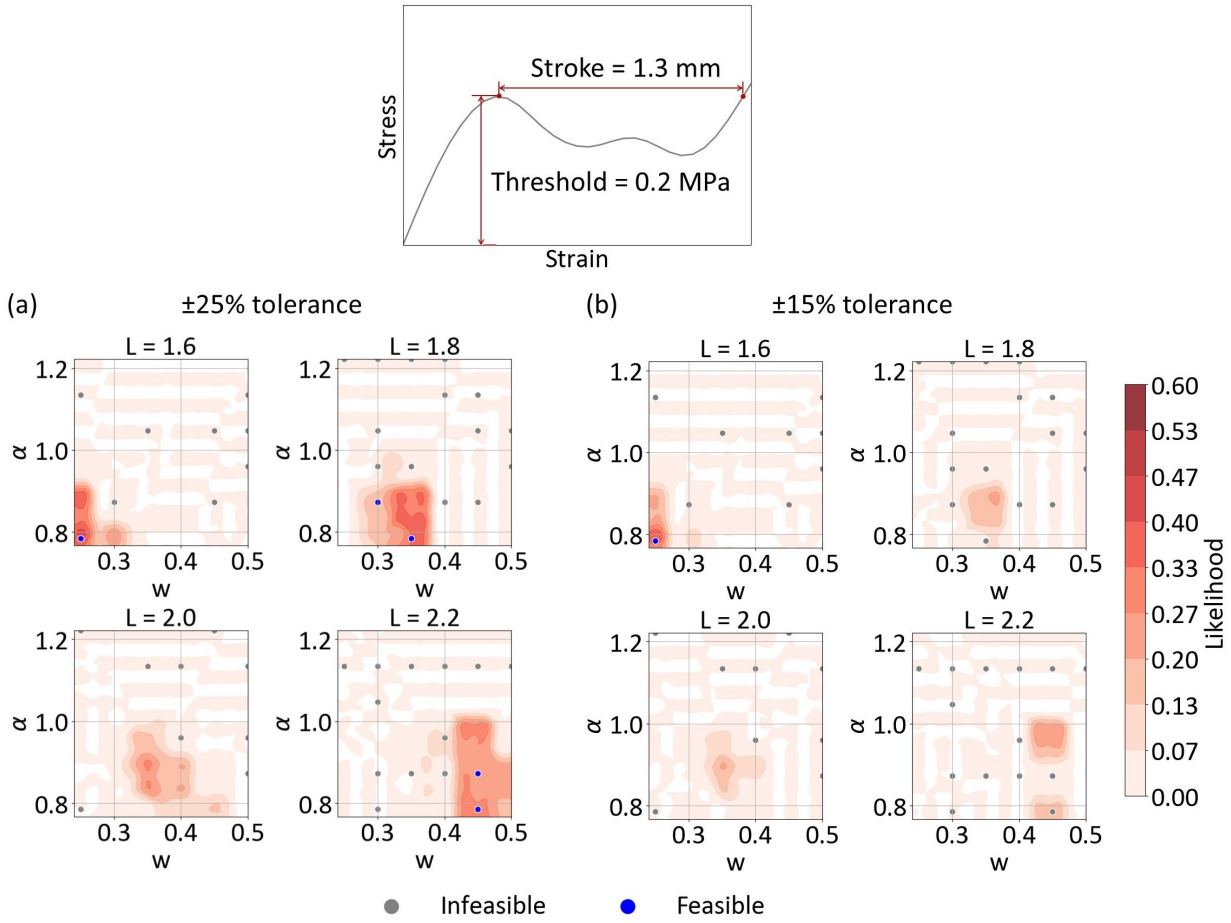

图8：在不同容差下，满足目标行程和阈值的设计空间中的似然分布。灰色/蓝色点分别表示训练数据集中的不可行/可行数据点。(a) 容差 = 25% (b) 容差 =15%。

4 讨论

本节将从需求规格说明的灵活性、数据效率和不确定性感知三个方面讨论 RAG 的优势。我们进一步分析其计算效率，以说明问题结构如何影响 RAG 的整体计算性能。

4.1 需求规格说明的灵活性

现有生成模型通常学习从功能响应 $\mathbf{y}$ 到设计变量 $\mathbf{x}$ 的逆映射。因此，必须手动指定满足设计要求的具有代表性的功能响应（例如，具有目标阈值和行程的应力–应变关系，或具有按需带隙属性的色散关系），作为输入以生成相应的设计方案。相比之下，RAG 通过估计以需求为条件的似然，直接学习从设计要求 $\mathcal{T}$ 到设计变量 $\mathbf{x}$ 的逆映射，即 $\mathcal{L}(\mathbf{x} \mid \mathcal{T})$。这消除了为设计生成手动指定具有代表性的功能响应的需求，并使需求规格说明更加灵活。

在本文考虑的两个设计变量任务——部分通带/阻带控制（第 3.1 节）和跳跃阈值/行程控制（第 3.2 节）——中，要求并未规定功能响应的确切形状，因此

允许功能响应呈现多种不同形状。在这种情况下，很难将要求表述为适用于拓扑优化的可微目标，而为生成模型确定具有代表性的色散关系或应力–应变关系同样并非易事。此外，使用某个特定的代表性功能响应来近似要求，不可避免地会引入偏差，并人为限制可行设计空间。相比之下，RAG 避免了这一约束来源，能够在设计空间内更全面地探索可行设计。

4.2 数据效率

与现有基于神经网络的方法完成类似任务所需的数据量相比，我们的方法只需少得多的数据，即可对完整的色散关系和应力–应变关系进行正向预测和逆向设计。RAG 的数据效率主要源于两个因素。首先，随机森林在小数据场景下天生比神经网络更稳健。其次，我们不将整个高维功能响应视为模型输出，而是将功能响应的查询点作为模型输入，并将模型的输出维度降至一维。这种重新表述大幅降低了学习复杂度，从而提高了数据效率。

4.3 不确定性感知

借助随机森林的集成结构，我们的框架能够在正向预测过程中量化响应空间中的不确定性，并在逆向设计过程中量化任意需求对应的设计空间中的不确定性。这一能力对于逆向设计问题尤为关键，因为这类问题可能是一对多的——需求相对容易满足，多个设计方案均可满足同一需求；也可能是一对无的——需求过于严苛，在预先定义的设计空间内不存在任何可行设计。现有大多数生成式逆向设计框架在生成设计方案时，都不会提供条件推断的任何度量。因此，即使需求难以满足甚至无法满足，这些方法也无法对此进行监测。相比之下，我们的方法能够提供模型不确定性信息，帮助判断在预先定义的设计空间内满足特定需求的难度。随后，可以通过选择置信度（似然）较高的设计方案，提高生成方案的可信度。

4.4 计算效率与可扩展性

随机森林的适用性受到其计算效率的制约。训练随机森林的时间复杂度为 $O(NMk\log(M))$，其中 N是树的数量， M是输入–输出对的数量， k是每次分裂时所考虑的特征数量。在数据获取过程中，我们通常会针对给定的设计变量 $\mathbf{x}$计算功能响应 $\mathbf{y}$。因此，训练设计样本的数量（记为 m）对应于不同 $\mathbf{x}$ –$\mathbf{y}$ 对的数量。根据公式 5 中的重构，引入查询点 $\mathbf{a}$ 作为额外输入，以预测相应的功能响应值。因此，每个原始的 $\mathbf{x}$ -$\mathbf{y}$ 对都被分解为 $d_{\mathbf{y}}$ 个输入–输出对。在此设定下，有 $M = m\sqrt{d_{\mathbf{y}}}$和 $k = \sqrt{d_{\mathbf{x}} + d_{\mathbf{a}}}$。因此，训练的时间复杂度可表示为 [77]

$$O(Nm\sqrt{d_{\mathbf{y}}(d_{\mathbf{x}} + d_{\mathbf{a}})}\log(m\sqrt{d_{\mathbf{y}}})) \tag{13}$$

该缩放关系说明，随机森林的计算效率取决于功能响应的维度和样本量。根据公式 13，可以评估由 $d_{\mathbf{y}}$ 确定的功能响应分辨率与训练效率之间的权衡。作为参考，本研究中每种情形的实际训练时间汇总于表 2，训练在配备 32 GB 内存的英特尔酷睿 i7-12700 中央处理器（2.10 GHz）上完成。可以观察到，两种情形的训练时间均在一分钟以内，明显快于基于神经网络的方法。声学超材料案例所需的训练时间更长，主要是因为其色散关系的维度远高于应力–应变关系。在设计生成阶段，效率

表 2：各任务中随机森林的训练时间。

Task	$d_{\mathbf{y}}$	m	$d_{\mathbf{x}}$	**训练时间（秒）**
声学超材料设计	915	400	3	58.88
机械超材料设计	30	826	5	3.01

马尔可夫链蒙特卡罗采样的时间主要受设计空间维度的影响 $d_{\mathbf{x}}$。高维设计空间会使马尔可夫链蒙特卡罗过程更难探索并找到高似然区域。当需求难以实现时，相应的高似然区域可能极小，甚至不存在，这会进一步增加采样时间。因此，为确保计算效率，所提出的方法更适用于维度相对较低的参数化设计空间。

5 结论

在这项工作中，我们提出了 RAG，这是一种基于随机森林的生成式逆向设计框架，能够利用小数据集处理定量的高维功能响应。RAG 只需训练一个正向随机森林模型，从而避免大量超参数调优，并实现快速、稳定的训练。我们不直接学习设计变量到功能响应的映射，而是将设计变量 $\mathbf{x}$ 与功能响应的查询点 $\mathbf{a}$ 作为输入，学习二者的联合映射。这种重新表述通过降低模型的输出维度，减轻了学习难度，并实现了离散化不变的预测。在逆向设计过程中，RAG 直接估计给定设计要求条件下的似然，为指定设计要求的方式提供了很大的灵活性。借助随机森林集成，该框架通过树输出之间的方差量化正向预测的不确定性，并通过估计似然量化逆向设计的不确定性。这些不确定性度量反映了模型推断结果的置信度。通过根据估计似然分布对设计空间进行采样，即可实现无需迭代、一次性完成的设计生成。

我们在两种超材料系统上验证了 RAG：具有指定部分通带/阻带的声学超材料，以及具有预期突跳行为的机械超材料；这些设计要求难以通过基于梯度的拓扑优化来满足。在这两种情况下，RAG 所需的数据量都远少于现有基于深度学习的方法，同时能够准确预测高维响应，并生成多样且可行的设计候选方案。似然估计有助于通过筛除低置信度解来确保生成设计的可信度，并指示设计要求的难度。

此外，我们指出了两个未来研究方向，可进一步发挥 RAG 的不确定性感知和易于构建的优势。首先，可利用 RAG 在正向预测中提供的不确定性信息进行自适应采样，从而提高数据效率。通过在高不确定性区域有选择地添加新的设计样本，可以大幅减少所需样本总数，同时保持预测精度。值得注意的是，RAG 学习的是功能响应的设计变量与查询点联合空间之间的映射。因此，可在设计空间 $\Omega_{\mathbf{x}}$ 和查询点空间 $\Omega_{\mathbf{a}}$ 中同时进行自适应采样。这意味着功能响应本身的离散化也可以优化。其次，得益于易于构建的特点，RAG 还可用于快速评估用户定义的设计空间是否足够丰富，能否满足给定的设计要求。在实践中，预定义的设计空间是否足以满足要求往往并不明确。RAG 为此类评估提供了一种高效工具。如果对于给定要求，RAG 估计的条件似然分布在整个设计空间内普遍较低，则表明预定义的设计空间不足，应予以扩展以满足该要求。

总体而言，RAG 为具有非线性和高维功能响应的超材料逆向设计提供了一种轻量、不确定性感知且数据高效的途径。其需求规格说明方式灵活、在小数据场景下稳健，并且兼容自适应学习策略，因此有望成为不确定性感知超材料设计流程的重要构件。除本研究展示的超材料设计外，RAG 还可用于更一般的逆向设计问题，这类问题涉及高维属性空间、昂贵的基于物理的仿真以及复杂的函数要求，只需进行少量简单修改即可应用。

致谢

Wei Chen、Doksoo Lee 和 Bolin Chen 感谢美国国家科学基金会推动工程领域变革性与公平性进步的研究构想（BRITE）研究员计划（CMMI-2227641）提供的支持。Wei Chen 和 Doksoo Lee 还感谢美国国家航空航天局通过太空制造中心（CISM-R2）：回收与月壤处理项目提供的少数族裔大学研究与教育项目机构研究机会（MIRO）部分资助（奖励编号 80NSSC24M0176）。Wei“Wayne”Chen 感谢美国国家科学基金会 EDSE 计划（CMMI 2434393）提供的支持。作者还感谢 Mohammad Charara 博士、Rachel Sun 和 Carlos M. Portela 教授就稿件提出了富有见地的讨论和建设性意见。

数据可用性声明

本文中声学超材料案例的源代码和数据可在以下地址获取：https://github.com/Sautoy/Random_Forest_Based_Generative_Design_Framework。

参考文献

[1] Xianglong Yu、Ji Zhou、Haiyi Liang、Zhengyi Jiang 和 Lingling Wu。与刚度、刚性和可压缩性相关的机械超材料：简要综述。Progress in Materials Science，94：114–173，2018。[2] Jun Wang、Gaole Dai 和 Jiping Huang。热超材料：基础、应用与展望。Iscience，23(10)，2020。[3] Steven A Cummer、Johan Christensen 和 Andrea Alù。利用声学超材料控制声音。Nature Reviews Materials，1(3)：1–13，2016。[4] Sophia R Sklan 和 Baowen Li。热超材料：功能与前景。National Science Review，5(2)：138–141，2018。[5] Pengcheng Jiao、Jochen Mueller、Jordan R Raney、Xiaoyu Zheng 和 Amir H Alavi。机械超材料及其他。Nature communications，14(1)：6004，2023。[6] Guangxin Liao、Congcong Luan、Zhenwei Wang、Jiapeng Liu、Xinhua Yao 和 Jianzhong Fu。声学超材料：理论、结构、制造方法与应用综述。Advanced Materials Technologies，6(5)：2000787，2021。[7] Costas M Soukoulis 和 Martin Wegener。三维光子超材料发展：既往成果与未来挑战。Nature photonics，5(9)：523–530，2011。[8] Yihui Wang、Wei Sha、Mi Xiao、Cheng-Wei Qiu 和 Liang Gao。深度学习赋能的热超材料智能设计。Advanced Materials，35(33)：2302387，2023。[9] Daicong Da 和 Wei Chen。用于热操控的双尺度数据驱动设计。International Journal of Heat and Mass Transfer，219：124823，2024。[10] Liwei Wang、Yu-Chin Chan、Faez Ahmed、Zhao Liu、Ping Zhu 和 Wei Chen。用于超材料系统基于机理的学习与设计的深度生成建模。Computer Methods in Applied Mechanics and Engineering，372：113377，2020。[11] Mourad Oudich、M Badreddine Assouar 和 Zhilin Hou。二维局域共振声子晶体板中的声波传播与波导。Applied Physics Letters，97(19)，2010。[12] Yifan Liu、Wei Huang、Zhiyong Wang、Jie Zhang 和 Jiayi Liu。基于机器学习的点阵结构力学性能预测与设计。International Journal of Mechanical Sciences，294：110230，2025。[13] 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。Materials & Design，211：110178，2021。[14] Kang Ang、Ji Qiu、Buyun Su、Zhiqiang Li、Xiaohu Yao、Zhihua Wang 和 Xuefeng Shu。基于深度学习的可编程无序超材料逆向设计。International Journal of Mechanical Sciences，第110712页，2025。[15] Zhenyang Gao、Hongze Wang、Nikita Letov、Yaoyao Fiona Zhao、Xiaolin Zhang、Yi Wu、Chu Lun Alex Leung 和 Haowei Wang。仿生复合超材料的数据驱动设计：实现极高可恢复性和超高比能量吸收。Composites Part B: Engineering，251：110468，2023。[16] Shengzhi Luan、Enze Chen、Joel John 和 Stavros Gaitanaros。有序与无序胞状超材料结构-性能关联的数据驱动框架。Science Advances，9(41)：eadi1453，2023。[17] Zihan Wang、Anindya Bhaduri、Hongyi Xu 和 Liping Wang。基于不确定性感知深度学习框架的超材料单元稳健设计优化。Structural and Multidisciplinary Optimization，68(3)：1–26，2025。[18] Jiapeng Sun、Yan Ma、Yulong He、Jianjiao Deng、Xi Chen、Dianlong Pan、Xin Li、Ming-Hui Lu 和 Yan- Feng Chen。低频振动衰减与基于集成学习的振动声学超材料逆向设计。Mechanical Systems and Signal Processing，239：113320，2025。[19] Than V Tran、SS Nanthakumar 和 Xiaoying Zhuang。用于按需逆向设计具有任意目标带隙的超材料的深度学习框架。npj Artificial Intelligence，1(1)：2，2025。[20] Hongyuan Liu、Yating Gao、Yongpeng Lei、Hui Wang 和 Qinxi Dong。穿孔负泊松比声子晶体弹性波操控的数据驱动逆向设计。Smart Materials and Structures，33(9)：095029，2024。[21] Zihan Wang、Weikang Xian、M Ridha Baccouche、Horst Lanzerath、Ying Li 和 Hongyi Xu。基于高斯混合β变分自编码器和迭代模型更新的声子带隙超材料设计。Journal of Mechanical Design，144(4)：041705，2022。

[22] Jeonghoon Park、Jaebum Noh、Jehyeon Shin、Grace X Gu 和 Junsuk Rho。利用解耦生成模型研究三维手性机械超材料的静态与动态行为。《Advanced Functional Materials》，35(2):2412901，2025。[23] Haoxuan Dylan Mu、Mingjian Tang、Wei Gao、Wei Chen 等。Guide：面向按需非线性功能响应的生成式与不确定性感知逆向设计。arXiv 预印本 arXiv:2509.05641，2025。[24] Wei Chen、Rachel Sun、Doksoo Lee、Carlos M Portela 和 Wei Chen。通过可解释学习对具有功能响应的超材料进行生成式逆向设计。《Advanced Intelligent Systems》，7(6):2400611，2025。[25] Zhiping Chai、Zisheng Zong、Haochen Yong、Xingxing Ke、Jiaqi Zhu、Han Ding、Chuan Fei Guo 和 Zhigang Wu。通过数据驱动逆向设计调整柔性弹跳机械超材料的应力–应变曲线，以实现按需机械响应。《Advanced Materials》，36(33):2404369，2024。[26] Sabiju Valiya Valappil、Alejandro M Aragón 和 Johannes FL Goosen。用于抑制夹装式超声流量计串扰的方向性带隙声子结构。《Mechanical Systems and Signal Processing》，224:112173，2025。[27] Pengfei Zhuang、Jun Wang、Shuai Yang 和 Jiping Huang。几何各向异性超材料中的非线性热响应。《Physical Review E》，106(4):044203，2022。[28] Chan Soo Ha、Desheng Yao、Zhenpeng Xu、Chenang Liu、Han Liu、Daniel Elkins、Matthew Kile、Vikram Deshpande、Zhenyu Kong、Mathieu Bauchy 等。通过机器学习，基于预设机械行为快速逆向设计超材料。《Nature Communications》，14(1):5765，2023。[29] Marco Maurizi、Derek Xu、Yu-Tong Wang、Desheng Yao、David Hahn、Mourad Oudich、Anish Satpati、Mathieu Bauchy、Wei Wang、Yizhou Sun 等。在图空间中设计具有可编程非线性响应和几何约束的超材料。《Nature Machine Intelligence》，7(7):1023–1036，2025。[30] Zhi Zhao、Rahul Dev Kundu、Ole Sigmund 和 Xiaojia Shelly Zhang。通过逆向设计利用层状材料实现极端非线性。《Science Advances》，11(20):eadr6925，2025。[31] Qingliang Zeng、Shengyu Duan、Zeang Zhao、Panding Wang 和 Hongshuai Lei。通过拓扑优化逆向设计吸能超材料。《Advanced Science》，10(4):2204977，2023。[32] Tark Raj Giri 和 Russell Mailen。多稳态机械超材料圆柱体中的可控弹跳序列与能量吸收。《International Journal of Mechanical Sciences》，204:106541，2021。[33] Kai Zhang、Yaoyao Guo、Xiangbing Liu、Fang Hong、Xiuhui Hou 和 Zichen Deng。基于深度学习的点阵超材料逆向设计以调控带隙。《Extreme Mechanics Letters》，69:102165，2024。[34] Jan-Hendrik Bastek、Siddhant Kumar、Bastian Telgen、Raphaël N Glaesener 和 Dennis M Kochmann。通过深度学习反转桁架超材料的结构–性能映射。《美国国家科学院院刊》，119(1):e2111505119，2022。[35] Ole Sigmund 和 Kurt Maute。拓扑优化方法：比较综述。《Structural and Multidisciplinary Optimization》，48(6):1031–1055，2013。[36] Jiashuo Xu、Mi Xiao 和 Liang Gao。通过拓扑优化逆向设计具有精确可编程非线性力学响应的结构。《Computer Methods in Applied Mechanics and Engineering》，446:118243，2025。[37] Vanessa Cool、Ole Sigmund 和 Niels Aage。通过拓扑优化设计具有振动声学带隙的超材料。《Computer Methods in Applied Mechanics and Engineering》，436:117744，2025。[38] Qiangbo Wu、Jingjie He、Wenjiong Chen、Quhao Li 和 Shutian Liu。具有预设带隙的声子晶体拓扑优化。《Computer Methods in Applied Mechanics and Engineering》，412:116071，2023。[39] Pegah Azizi、Rahul Dev Kundu、Weichen Li、Kai Sun、Xiaojia Shelly Zhang 和 Stefano Gonella。按需优化具有拓扑态的点阵材料。《美国国家科学院院刊》，122(32):e2506787122，2025。[40] Xiaopeng Zhang、Jian Xing、Pai Liu、Yangjun Luo 和 Zhan Kang。通过非梯度拓扑优化实现声学超材料的全向和方向性带隙设计。《Extreme Mechanics Letters》，42:101126，2021。[41] Xiaopeng Zhang、Yan Li、Yaguang Wang 和 Yangjun Luo。通过多材料拓扑优化设计声学超材料的超宽低频带隙。《Composite Structures》，306:116584，2023。[42] 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。

[43] Xiaoyang Zheng、Xubo Zhang、Ta-Te Chen 和 Ikumu Watanabe。机械超材料中的深度学习：从预测和生成到逆向设计。*Advanced Materials*，35(45):2302530，2023。[44] Bolei Deng、Ahmad Zareei、Xiaoxiao Ding、James C Weaver、Chris H Rycroft 和 Katia Bertoldi。通过神经加速进化策略，针对目标非线性响应逆向设计机械超材料。*AdvancedMaterials*，34(41):2206238，2022。[45] Yueyou Tang、Anfu Zhang、Qi Zhou、Mu He 和 Liang Xia。数据驱动设计拓扑优化负泊松比超材料，以定制应力–应变和泊松比–应变行为。*Materials & Design*，第114290页，2025。[46] Hamza Baali、Mahmoud Addouche、Abdesselam Bouzerdoum 和 Abdelkrim Khelif。利用数据驱动方法设计吸声超表面。《通讯材料》，4(1):40，2023。[47] Ertai Cao、Zhicheng Dong、Ben Jia 和 Heyuan Huang。通过数据驱动策略逆向设计各向同性负泊松比超材料。《材料前沿》，2025。[48] Zhibin Liang、Chongrui Liu、Yuze Liu、Yueyin Ma 和 Fuyin Ma。基于元启发式增强自编码器网络，面向需求的超薄宽带吸声超材料逆向设计与优化。*Composites Part B: Engineering*，第112643页，2025。[49] Weifeng Jiang、Yangyang Zhu、Guofu Yin、Houhong Lu、Luofeng Xie 和 Ming Yin。基于深度学习的弹性超材料色散关系预测与结构逆向设计。《今日材料：物理》，22:100616，2022。[50] Caglar Gurbuz、Felix Kronowetter、Christoph Dietz、Martin Eser、Jonas Schmid 和 Steffen Marburg。用于声学超材料设计的生成对抗网络。《美国声学学会杂志》，149(2):1162–1174，2021。[51] Jiahui Yan、Yingli Li、Guohui Yin、Song Yao 和 Yong Peng。利用深度学习逆向设计具有高自由度的声学超材料定制吸收。*MechanicalSystemsandSignalProcessing*，237:112989，2025。[52] Jan-Hendrik Bastek 和 Dennis M Kochmann。通过视频去噪扩散模型逆向设计非线性机械超材料。*NatureMachine Intelligence*，5(12):1466–1475，2023。[53] Nikolaos N Vlassis 和 WaiChing Sun。用于逆向设计微结构并精细调控非线性材料性质的去噪扩散算法。*ComputerMethods in Applied Mechanicsand Engineering*，413:116126，2023。[54] Qibang Liu、Seid Koric、Diab Abueidda、Hadi Meidani 和 Philippe Geubelle。迈向基于有符号距离函数的超材料设计：用于正向预测的神经算子变换器与用于逆向设计的扩散模型。*ComputerMethods in AppliedMechanics and Engineering*，446:118316，2025。[55] Che Wang、Yuhao Fu、Ke Deng 和 Chunlin Ji。用于超材料逆向设计的功能响应条件变分自编码器。载于***NeurIPS 2021*** 深度学习与逆问题研讨会，2021。[56] Wei Ma、Feng Cheng、Yihao Xu、Qinlong Wen 和 Yongmin Liu。基于深度生成模型与半监督学习策略的超材料概率表示和逆向设计。*AdvancedMaterials*，31(35):1901111，2019。[57] Victor Rodriguez-Galiano、Manuel Sanchez-Castillo、M Chica-Olmo 和 MJOGR Chica-Rivas。用于矿产远景预测的机器学习模型：对神经网络、随机森林、回归树和支持向量机的评估。《矿石地质学评论》，71:804–818，2015。[58] Hengrui Zhang、Wei Chen、Akshay Iyer、Daniel W Apley 和 Wei Chen。面向材料设计的不确定性感知混合变量机器学习。《科学报告》，12(1):19760，2022。[59] Jian He、Yaohui Wang、Hubocheng Tang、Guoquan Zhang、Ke Dong、Dong Wang、Liang Xia 和 Yi Xiong。可定制双稳态单元使软硬夹爪能够通过数据驱动设计处理多特征物体。《材料前沿》，12(12):4426–4433，2025。[60] Bhagwat Singh Chouhan、Ali Nawaz、Asit Das、Rohith KM、A mir Ahmad 和 Gagan Kumar。机器学习驱动的超宽带太赫兹多层超材料。《光波技术杂志》，2024。[61] 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。*Computer Methods in Applied Mechanics andEngineering*，432:117440，2024。[62] Xin-chun Zhang、Zhi-yi Song、Yi-nan Li、Li-jun Xiao、Zheng Xu、Li-xiang Rao、Tie-jun Ci 和 Xu-Long Hui。具有定制应力–应变响应的超材料生成式逆向设计。*International Journal ofMechanical Sciences*，第110875页，2025。

[63] Yuhao Bao、Zhiyuan Jia、Qiming Tian、Yangjun Luo、Xiaopeng Zhang 和 Zhan Kang。基于声子晶体的声学解复用器设计：通过带隙-通带拓扑优化实现。*Composite Structures*，351:118622，2025。[64] Sushan Nakarmi、Jeffery A Leiding、Kwan-Soo Lee 和 Nitin P Daphalapurkar。利用监督式自编码器预测介观结构多孔材料的非线性应力–应变响应。*ComputerMethods in AppliedMechanics and Engineering*，432:117372，2024。[65] Hong-yun Yang、Xiao-sun Wang、Lu Liu 和 Shi-jing Wu。基于深度学习的二维局域共振超材料低频带隙预测及弹性波传播特性研究。*European JournalofMechanics-A/Solids*，第105823页，2025。[66] Julia Ling、Maxwell Hutchinson、Erin Antono、Sean Paradiso 和 Bryce Meredig。利用数据驱动的实验设计和校准良好的不确定性估计，实现高维材料与工艺优化。*IntegratingMaterials and Manufacturing Innovation*，6(3):207–217，2017。[67] W Keith Hastings。使用马尔可夫链的蒙特卡洛抽样方法及其应用。1970。[68] Raisa Abedin Disha 和 Sajjad Waheed。基于基尼不纯度的加权随机森林（giwrf）特征选择技术在入侵检测系统中的机器学习模型性能分析。《*Cybersecurity*》，5(1):1，2022。[69] Xin Fang、Walter Lacarbonara 和 Li Cheng。非线性声学/弹性超材料与超结构的研究进展。《*Nonlinear Dynamics*》，第1–28页，2024。[70] Fahimeh Akbari-Farahani 和 Salman Ebrahimi-Nejad。从缺陷模态到拓扑超材料：用于能量收集的声子晶体与声学超材料前沿综述。《*Sensors and Actuators A: Physical*》，365:114871，2024。[71] Arash Kazemi、Kshiteej J Deshmukh、Fei Chen、Yunya Liu、Bolei Deng、Henry Chien Fu 和 Pai Wang。绘制色散曲线：通过非局域声子晶体定制能带结构。《*Physical Review Letters*》，131(17):176101，2023。[72] Alexander C Ogren、Berthy T Feng、Katherine L Bouman 和 Chiara Daraio。将高斯过程回归用作计算色散关系的代理模型。*ComputerMethods in AppliedMechanicsandEngineering*，420:116661，2024。[73] Ting-Wei Liu、Chun-Tat Chan 和 Rih-Teng Wu。基于深度学习的声学超材料设计：衰减听觉频段中的结构传声噪声。《*Materials*》，16(5):1879，2023。[74] Lige Chang、Xiaowen Li、Zengrong Guo、Yajun Cao、Yuyang Lu、Rinaldo Garziera 和 Hanqing Jiang。利用机器学习实现按需可调超材料设计以衰减噪声。*Materials& Design*，238:112685，2024。[75] Rachel Sun、Jet Lem、Yun Kai、Washington DeLima 和 Carlos M Portela。通过惯性设计实现微尺度超材料中的定制化超声传播。*ScienceAdvances*，10(45):eadq6425，2024。[76] Qiangbo Wu、Quhao Li、Renjing Gao 和 Shutian Liu。面向特定机械性能和带隙性能集成设计的声子晶体拓扑优化。*StructuralandMultidisciplinary Optimization*，68(9):1–20，2025。[77] Gilles Louppe。理解随机森林。*arXiv*预印本 *arXiv:1407.7502*，2014。