arXiv:2602.10451v1 [cs.LG] 11 Feb 2026

一种具有分布级物理先验的多模态条件混合模型

Jinkyo Han
西北大学机械工程系
jinkyohan2030@u.northwestern.edu

Bahador Bahmani*
西北大学机械工程系
bahador.bahmani@northwestern.edu

2026年2月12日

摘要

许多科学与工程系统都表现出内在的多模态行为，这源于潜在的状态切换和物理机制的非唯一性。在这类情形中，如何以物理一致且可解释的方式学习可行结果的完整条件分布，仍然是一项挑战。尽管机器学习的近期进展已推动强大的多模态生成建模方法发展，但将其与物理约束科学建模相结合仍非易事，尤其是在必须保留物理结构或数据有限时。本研究提出一种基于混合密度表示的物理信息多模态条件建模框架。混合密度网络（MDN）为多模态条件分布提供了显式且可解释的参数化。通过分量特定正则化项嵌入物理知识，以惩罚违反控制方程或物理定律的情况。这种建模方式能够自然地处理非唯一性与随机性，同时保持计算效率，并适用于以情境输入为条件的建模。我们在一系列多模态性源于内在物理机制而非观测噪声的科学问题上评估了所提出的框架，包括非线性动力系统中的分岔现象、随机偏微分方程以及原子尺度冲击动力学。此外，我们还将所提出的方法与条件流匹配（CFM）模型这一具有代表性的最先进生成建模方法进行了比较，结果表明，混合密度网络能够取得有竞争力的性能，同时提供更简单、可解释性更强的建模方案。

关键词 混合密度网络、专家混合模型、物理信息神经网络、多模态数据

1 引言

许多科学与工程系统会表现出内在的多模态行为，这种行为源于材料响应的异质性、不完整观测、潜在机制切换，或输入与输出之间的非唯一映射。例如，分岔现象 [47], 随机偏微分方程 (SPDE) [11, 50], 冲击驱动的原子动力学 [8, 23]。在这些情况下，需要预测可行结果的完整条件分布；理想情况下，所得模型还应满足物理一致性，并具有有意义的可解释性。

计算科学中的传统建模方法通常依赖单峰不确定性表示 [46]。概率方法 [2–4, 12, 15, 20, 21, 25–27, 31, 41, 44, 45], 包括蒙特卡洛采样、多项式混沌展开、高斯过程和贝叶斯推断，为不确定性量化提供了严谨的工具，并已在实践中取得广泛成功。这些方法通常基于单峰表述构建，并明确定义随机输入；当多模态性源自隐式或尚未解析的物理机制时，这促使人们探索互补的方法。

机器学习的最新进展显著拓展了多峰生成模型建模的领域。生成对抗网络 (GANs) [16], 变分自编码器[22], 归一化流 [32], 扩散

*通讯作者。

模型[42], 基于分数的生成方法 [43], 和流匹配模型 [24]在计算机视觉和语言建模领域展现出卓越的成效。然而，将其应用于物理约束科学问题仍面临挑战， [1, 18, 38, 40], 因为纳入物理结构以及开发条件形式并非易事，且仍是活跃的研究领域 [7]。尽管这些模型功能强大、表达能力出色，但往往需要大量数据和审慎的训练流程；此外，以保留物理结构（如守恒定律、对称性或控制方程）的方式对其施加约束，也并非易事。此外，它们的潜在表示并不总是容易解释，也不易在既有的科学工作流程中进行操作。

本文重新审视混合密度表示 [5, 17, 28]，将其视为一种原理清晰但尚未得到充分探索的物理约束多模态建模替代方案。混合密度网络（MDN） [5], 最初作为条件概率模型提出，可将复杂的条件分布近似为简单参数化分量的有限混合。尽管相对简单，混合模型仍具备若干特点，使其特别适用于科学机器学习：对多模态性的显式表示、可解释的分量、易于处理的似然，以及对输入进行直接条件化。尽管基于混合模型的形式化方法历史悠久，但近年来，随着混合专家架构 [13, 39]在现代大语言模型中的应用，这类方法重新得到大规模采用。

本文提出一种基于混合密度表示的物理信息多模态条件建模框架，为数据驱动学习与基于物理的建模提供自然的接口。通过正则化项将物理先验和约束直接纳入混合模型形式中，这些正则化项用于惩罚对控制方程或物理定律的违反。此类约束施加于条件分布层面，而非单条预测轨迹或单个场实现。

该框架在一系列科学问题中进行了评估。在这些问题中，多模态性源于内在的物理机制，而非观测噪声，包括具有非唯一解的参数化偏微分方程、随机偏微分方程、原子尺度冲击动力学，以及非线性动力系统中的分岔现象。在这些情形下，基于混合模型的表述能够捕捉单峰模型无法表示的不同解分支、状态转变和结构化不确定性。

相关工作。 混合密度网络已在各种工程和应用科学领域中作为概率代理模型得到探索。其中，与本文工作最为接近的是 [9], , 该方法通过受物理因素启发的正则化项扩展混合密度网络。这些正则项主要具有统计性质，例如约束混合分量之间的方差，而非强制满足控制方程。此外，这些正则化项被均匀地应用于各个分量，没有考虑它们对条件分布的相对贡献；尽管混合密度网络具有较强的表达能力，所研究的数据却主要呈单峰分布。混合密度网络已被用作不确定性量化的随机代理模型 [35], , 包括用于热载荷下桥梁伸缩缝可靠性分析的循环概率预测 [49] , 以及地震响应建模 [36]。在设计和逆问题中，基于混合模型的方法已应用于微结构设计 [52] , 以及利用卷积混合密度网络对光子结构进行逆向设计 [48]。此外，研究者还探索了混合密度网络的集成形式，以进行概率风力预测 [29, 53]。此外，还有其他利用神经网络进行概率建模的方法，其中贝叶斯神经网络是最成熟的范式之一 [6, 30]。这类概率表述也与物理信息场景中的基于物理的正则化相结合，通过额外的损失项或先验纳入物理知识 [37, 51]。

总之，本工作的主要贡献如下：

1. 提出一种基于混合密度表示的物理信息建模框架，用于条件建模，并通过对均值函数进行分量特定的物理正则化来施加物理先验。2. 引入类别条件机制，在纳入离散上下文变量的同时，保留混合分量之间的共享结构和统计耦合，避免训练独立的类别专属模型可能造成的跨分量相关性损失。3. 将条件流匹配作为具有代表性的最先进生成建模方法进行比较。4. 展示该框架在多个科学领域中的通用性；在这些领域中，多模态性源于内在物理机制，而非观测噪声。

本文其余部分安排如下。在第 2节中，我们将建立用于建模条件多峰分布的混合密度网络核心公式，并将其扩展以纳入类别条件信息和基于物理的正则化。第 3节将展示一系列数值示例，说明所提出的

框架在具有内在多模态行为的物理系统上的应用。最后，第 4 节将总结本文，并概述主要成果与局限性。

2 公式构建

设 $(\Omega\ \mathcal{F}, \mathbb{P})$ 为一个概率空间。我们考虑一个条件随机场 $U\colon \Omega \times \mathcal{X} \to \mathbb{R}$，其中 $\mathcal{X}$表示一个指标集（例如，时空坐标）。集合 $\{U(x)\}_{x\in\mathcal{X}}$定义了 $\mathcal{X}$ 上的一个随机场。我们假设可以获得一个包含 N 个样本 $\mathcal{D} = \{(x_i, u_i)\}_{i=1}^{N}$ 的数据集，其中 $x_i \in \mathcal{X}$ 是评估指标，$u_i \in \mathbb{R}$ 是观测到的场值。我们的目标是学习场评估值的条件密度 $p(u|x)$，它可能呈现内在多模态性。为此，我们采用一种基于混合的概率表示，该表示兼具表达能力和解析可处理性。

有限混合密度模型通过将条件密度表示为分量密度的凸组合，为表示多模态概率分布提供了通用框架。在此处考虑的条件情形下，给定输入 x时的分布 u建模为

$$p(u|x) = \sum_{m=1}^{M} \pi_m(x)\, \phi_m(u|x), \tag{1}$$

其中 $\pi_m(x) \geq 0$ 满足 $\sum_{m=1}^{M} \pi_m(x) = 1$，而 $\phi_m(u|x)$ 表示分量条件密度。

我们重点关注这样一类条件混合模型：其中混合权重和分量密度的参数均由神经网络参数化。由此得到混合密度网络（MDN），其中所有混合参数均通过数据联合学习。将网络参数统记为 θ，则所得条件密度可写为

$$p(u|x;\theta) = \sum_{m=1}^{M} \pi_m(x;\theta)\, \phi_m(u|x;\theta). \tag{2}$$

这种参数化相对于输入变量诱导出一个连续条件随机场，这对于涉及基于梯度的约束的物理建模应用尤其有利。

虽然分量条件参数化密度 $\phi_m(u|x;\theta)$ 原则上可以选自广泛的分布类别，但本文始终仅考虑参数化的高斯分量。这一选择具有解析可处理性。

因此，每个分量密度均取为如下形式的高斯分布

$$\phi_m(u|x;\theta) = \frac{1}{\sigma_m(x;\theta)\sqrt{2\pi}} \exp\left(-\frac{\|u - \mu_m(x;\theta)\|^2}{2\,\sigma_m(x;\theta)^2}\right), \tag{3}$$

其中 $\mu_m(x;\ \theta)$ 和 $\sigma_m(x;\theta) > 0$ 分别表示分量均值和标准差。采用具有高斯条件分量的混合模型，其优点在于随机场的一阶和二阶条件统计矩可以写成闭式表达，这对于不确定性量化尤其有用。

$$\mathbb{E}[u|x] = \sum_{m=1}^{M} \pi_m(x)\, \mu_m(x), \tag{4}$$

$$\mathbb{E}[u^2|x] = \sum_{m=1}^{M} \pi_m(x)\, \left(\sigma_m(x)^2 + \mu_m(x)^2\right). \tag{5}$$

注 2.1。 对于标量值场，上述选择对应于各向同性高斯分布；将其扩展到具有结构化协方差的多变量输出也很简单。

模型参数通过最大似然估计进行学习。在总体层面上，这相当于最小化期望负对数似然

$$\mathcal{L}_{\mathrm{NLL}} = \mathbb{E}_{(x,u)}[-\log(\sum_{m=1}^{M} \pi_m(x)\, \phi_m(u|x))]. \tag{6}$$

这相当于对潜在混合成分分配进行边缘化。

在某些应用中，我们可以获得额外的类别信息，将每个观测与已知的物理状态或类别相关联。将此类信息作为额外的监督来源纳入模型，尤其

在数据稀缺场景中很有价值，还可以通过引导混合分量与物理状态之间的关联，帮助缓解混合密度模型中常见的可辨识性问题。令 $c \in \{1,\dots,C\}$ 表示与数据点 $(x,\ u)$ 相关联的类别标签，并令 g: $\{1,\dots,C\} \to \{1,\dots,M\}$ 表示从类别到混合分量的预设映射。这里并不将类别标签视为硬性约束，而是用它们在训练期间为混合成分分配提供弱引导。

具体而言，每个数据点的似然贡献仅限于与其类别标签相关联的混合分量，从而得到类别感知的负对数似然

$$\mathcal{L}_{\mathrm{NLL}}^{\mathrm{class}} = \mathbb{E}_{(x,u)}\left[-\log\left(\pi_{g(c)}(x)\,\phi_{g(c)}(u \mid x)\right)\right]. \tag{7}$$

该公式促使混合分量与已知物理状态保持一致，同时通过**在各分量之间保留共享参数化和统计耦合**，借助**共同的网络架构**实现这一点。在单个模型中进行联合训练，会通过共享表示隐式捕捉跨成分相关性；如果使用独立的网络分别对每种状态进行建模，这些相关性通常不会存在。

通过在似然目标函数中加入施加于成分均值函数 $\{\mu_m\}_{m=1}^{M}$ 的额外约束项，引入基于物理的先验信息。由于每个混合分量以概率 $\pi_m(x)$表示一个条件众数，因此相对于潜在成分分布的期望来施加物理一致性约束。令 $\mathcal{R}_{\mathrm{phys}}(\mu_m; x)$ 表示作用于第 m 个成分均值的基于物理的残差或约束泛函。物理信息正则化项定义为

$$\mathcal{L}_{\mathrm{phys}} = \mathbb{E}_{(x)}\left[\sum_{m=1}^{M} \pi_m(x)\,\mathcal{R}_{\mathrm{phys}}(\mu_m; x)\right]. \tag{8}$$

将混合权重 $\pi_m(x)$置于期望内部，体现了混合密度模型的概率语义。在混合密度网络中，每个分量对应一个潜在条件众数，并以概率 $\pi_m(x)$处于激活状态。因此，不应对所有分量均匀地施加物理一致性约束，而应根据每个分量在给定输入下对条件分布的似然贡献比例来施加约束。将基于物理的残差 $\mathcal{R}_{\mathrm{phys}}(\mu_m; x)$ 乘以 $\pi_m(x)$，可确保相对于潜在成分分布的期望来施加物理约束。在给定输入下，根据数据不太可能出现的分量对正则化目标函数的影响可以忽略，而占主导地位的分量则受到更强的约束。这避免了对未激活或缺乏充分数据支持的众数施加虚假的物理惩罚；否则，这些惩罚可能会使学习得到的条件分布产生偏差，或扭曲多模态结构。从概率角度看，公式8中的正则化项对应于输入与潜在混合成分分配的联合分布下的期望物理残差。该公式与将混合密度网络解释为潜变量模型的观点一致，并能够在完整条件分布层面实现物理信息学习，而不是仅针对单个成分实现。

总训练目标函数定义为

$$\mathcal{L}_{\mathrm{total}} = \mathcal{L}_{\mathrm{NLL}} + \lambda\,\mathcal{L}_{\mathrm{phys}}, \tag{9}$$

其中，$\lambda > 0$ 用于权衡数据保真度与物理一致性。

备注 2.2。对于成分参数固定且与样本无关的离散混合模型，最大似然估计通常采用期望最大化（EM）算法来执行，该算法交替进行潜在成分分配的估计与模型参数的更新 [28]。相比之下，在当前设定中，混合系数和成分参数被建模为由单个神经网络参数化的确定性输入相关函数。因此，无需引入显式的 EM 过程，即可直接使用基于梯度的方法优化似然。

2.1 网络架构

条件混合模型由一个神经网络参数化，该网络将输入变量（$z,\ x$) $\in \mathcal{Z} \times \mathcal{X}$映射为条件混合分布的参数。具体而言，该网络输出以下集合

$$\{\pi_m(z,x), \mu_m(z,x), \sigma_m(z,x)\}_{m=1}^{M},$$

对应的 t 分别为混合系数、分量均值和分量标准差 .

混合系数 $\{\pi_m(z,\ x)\}_{m=1}^{M}$ 通过应用 Softmax 激活函数获得，以确保其非负且经过归一化，使得 $\sum_{m=1}^{M}\pi_m(z,\ x) = 1$。分量标准差 $\{\sigma_m(z,\ x)\}_{m=1}^{M}$ 通过指数映射确保严格为正。分量均值 $\{\mu_m(z,\ x)\}_{m=1}^{M}$ 则不受额外约束，是模型中物理结构的主要载体。

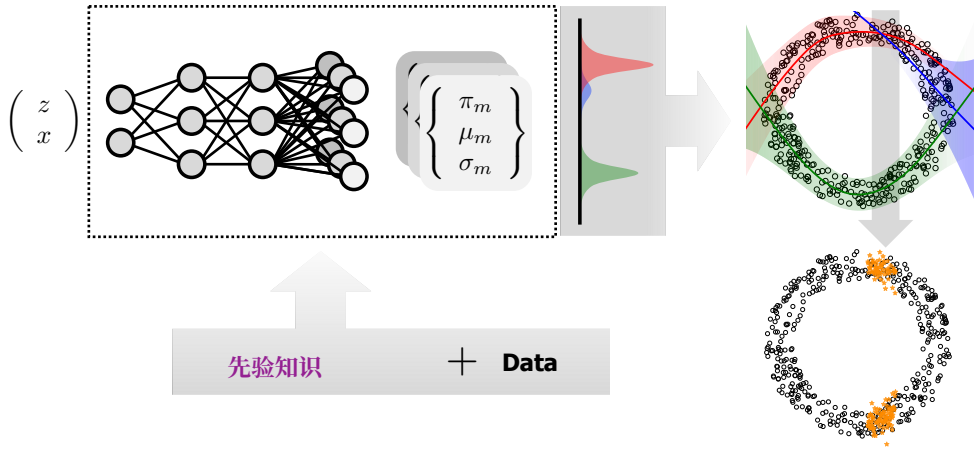


图1：混合密度网络架构和所提出的多模态条件建模框架示意图。

图 1 展示了混合密度网络（MDN）架构及所提出框架的示意图。单个网络将输入变量 $(z,\ x)$ 映射为条件混合分布的参数，包括混合系数、均值和方差，从而表示完整的条件密度。网络架构可由全连接层或其他适当的架构组成，具体取决于输入变量 $(z,\ x)$ 的结构。所有混合参数均通过基于梯度的优化联合学习。

3 数值示例

本节将通过四个数值示例展示所提出的框架。首先，我们考虑一个表现出分岔行为的动力系统，在模型容量和训练工作量相当的条件下比较混合密度网络和CFM的性能。随后，我们研究一个随机微分方程，以评估混合密度网络学习条件平衡分布的能力，并考察其外推行为。接下来，我们考虑从文献汇总得到的冲击雨贡纽数据，展示如何将单调性约束和类别条件结构作为先验知识纳入单个混合密度网络，以提高预测准确度和不确定性估计。最后，我们考察Chafee–Infante方程，以说明基于稳态偏微分方程残差的物理信息正则化的影响。实现细节和超参数设置见附录 A。

在所有示例中，均使用 `PyTorch` [34] 中的自动微分来计算基于物理的损失项所需的梯度。尽管混合分量的数量必须 先验地指定，但附录 B 提供了一项补充研究，表明所学习的生成模型在预测性能和不确定性估计两方面对于混合分量数量的敏感性均有限。

3.1 分岔图

第一个示例中，我们考虑一个呈现尖点型分岔的动力系统，

$$\dot{x} = -x^3 + \lambda x + \mu, \tag{10}$$

其中 x 表示系统状态，λ 是控制参数，而 μ 表示不完美参数。该系统曾被 [10] 用于基准测试，以在联合 $(x、\ \lambda、\ \mu)$ 空间中对分岔曲面和稳态进行生成建模。这里，我们将该系统视为有界缺陷 μ条件下的条件预测问题，并将其用作受控测试案例，以考察混合密度模型如何表示以控制参数 λ为条件的分岔型多峰稳态响应。我们还将所得的基于混合模型的表示与 CFM 生成模型进行比较。

当 $\mu = 0$ 时，系统发生对称叉形分岔；而非零的 μ会打破对称性并扰动分岔结构。我们关注稳态响应。

为构建数据集，我们对 $x \sim \mathcal{U}(-2、\ 2)$ 和 $\lambda \sim \mathcal{U}(-2.5,\ 2.5)$ 进行采样。仅选择满足 $|\mu| < 0.05$ 的样本，这些样本对应于较小但非零的缺陷。此过程生成的数据集中在理想分岔曲面附近，同时包含由缺陷引起的小幅扰动。

对于 λ的固定取值，由于底层分岔结构，多个稳态可能共存，而有界缺陷会使这些可接受解周围出现变异。因此，如图 2所示，给定 λ 时， x 的条件分布呈多峰分布，且具有非零离散程度。

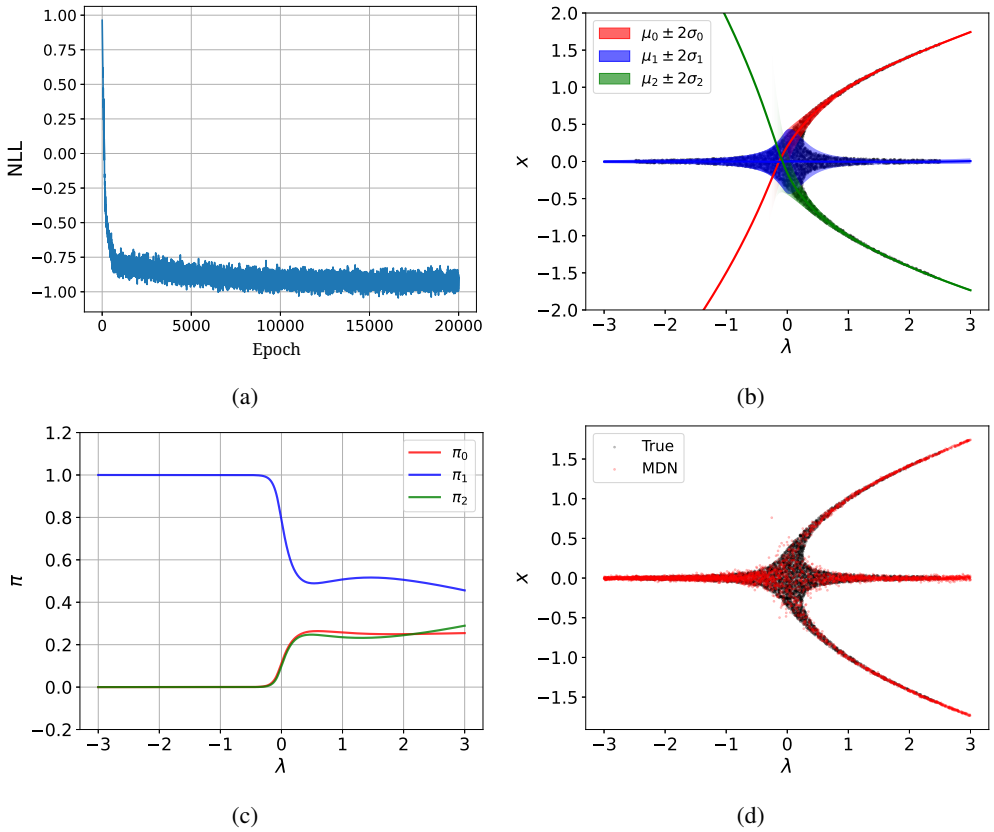


图2：用于对分岔响应分支进行概率建模的混合密度网络训练。（a）20,000次 ADAM 迭代中的损失历史。（b）预测的分量均值 μ_m 和标准差 σ_m 以及（c）预测的混合系数 π_m。（d）5,000 个从训练后的混合密度网络模型中抽取的样本。

混合密度网络的训练过程及其得到的概率表示如图2所示。图2a展示了优化过程中负对数似然目标函数的变化。图2b和2c 给出了模型学习到的各分量预测均值及不动点位置 μ_m、对应的标准差 σ_m和混合系数π_m 随条件参数 λ的变化。在区域 $\lambda < 0$ 中，仅存在一个稳态，因此无关分量的贡献会减小；而在分岔区域，多个分量均有贡献 $\lambda > 0$，体现了多模态行为的出现。图2d 展示了从训练后的混合密度网络模型中抽取的样本，说明该模型能够捕捉分岔响应结构以及有界缺陷所导致的离散程度。

图3 对比了模型容量相当（可训练参数数量相近）的混合密度网络和 CFM 模型。图3a展示了 CFM 在一组密集的控制参数值条件下生成的样本$\lambda \in [-3, 3]$，说明该模型能够捕捉整体分岔结构，但各分支上仍存在明显离散。图3b–d对比了真实数据、混合密度网络和 CFM 在 $\lambda_0 = -0.8$、0和0.8附近代表性条件值下的经验条件分布。

在 $\lambda_0 = -0.8$附近，响应为单峰，两种模型生成的分布都集中在单一稳态附近。在 $\lambda_0 = 0$附近，即分岔开始出现时，混合密度网络能够捕捉到变宽的条件分布，而 CFM 则倾向于将概率质量集中在中间值附近。在完全分岔区域 $\lambda_0 = 0.8$，混合密度网络能够分辨出三个不同的响应模态，而 CFM 生成的样本在各模态之间的中间区域存在明显的概率质量。

这些结果表明，相较于 CFM 等先进方法，混合密度网络能够提供一种具有竞争力的条件生成建模框架，同时显著简化训练和采样流程。我们

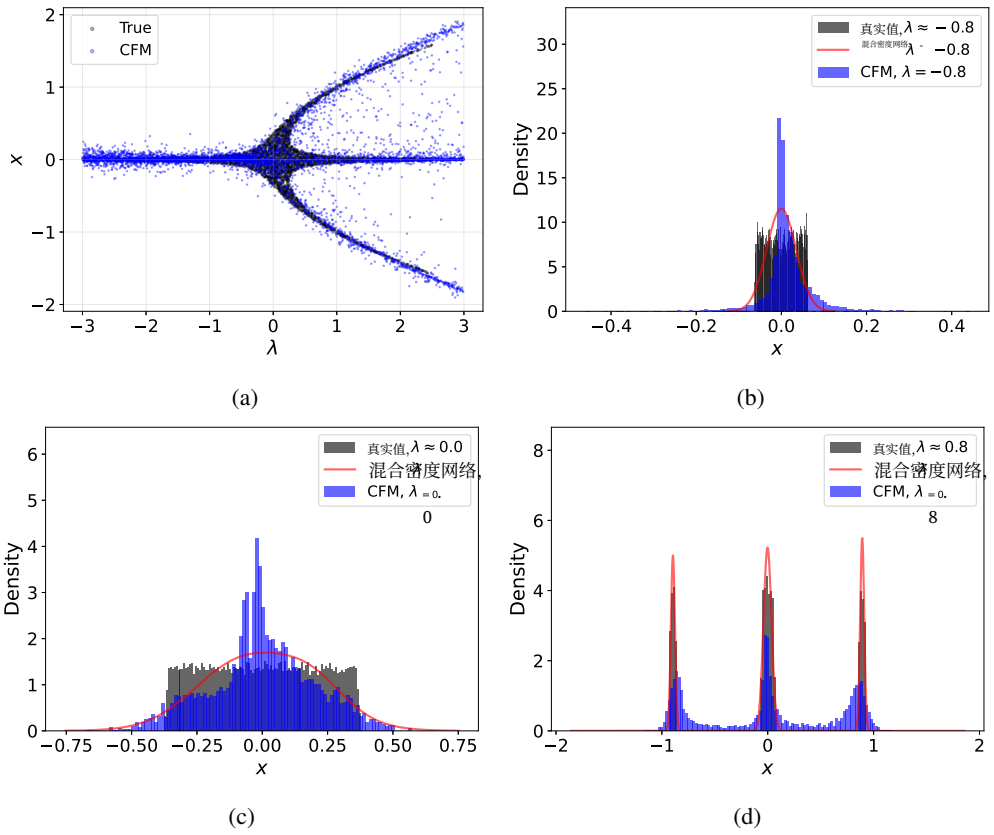


图3：MDN 和 CFM 在经过 20,000 次 ADAM 迭代后的比较。MDN 的模型参数数量为 457，而 CFM 的参数数量与之相近，为521 个。（a）在给定 5,000 个等距 λ 值（范围从 -3 到 3）时，CFM 模型输出的样本分布。真实分布、MDN 样本以及 5,000 个 CFM 样本的分布，分别以（b） $\lambda = -0.8$、（c） $\lambda = 0$和（d） $\lambda = 0.8$为条件。

强调，可能存在其他设置，使 CFM 的表现优于本研究中的结果；不过，在本研究中，两种方法都经过了相近程度的调优，且未进行大规模超参数优化。

3.2 多尺度随机微分方程

作为第二个例子，我们考虑一个二维多尺度随机微分方程，

$$\mathrm{d}u_1 = a_1\,\mathrm{d}t + a_2\,\mathrm{d}B_1, \tag{11}$$

$$\mathrm{d}u_2 = -\big(-1 + 0.2\,u_1 + 4\,u_2(-1+u_2^2)\big)\,\mathrm{d}t + a_3\,\mathrm{d}B_2, \tag{12}$$

该方程已在 [11] 中作为一种典型的慢快随机系统进行研究，其双稳态行为依赖于参数。在该系统中，慢变量 u_1 会调节快变量 u_2的有效势，从而使 u_2 的条件平衡分布在 u_1取固定值时通常呈多峰。

这类系统为研究多尺度动力学中的条件平衡测度提供了一个简洁而具有代表性的场景。在许多应用中，例如分子动力学和材料建模，人们往往关注如何采样与粗粒度变量或集体变量的给定取值相一致的细尺度构型。在此类场景中，生成式条件模型可以发挥补充基于物理的模拟的作用，通过提供有依据的初始化或提议分布，加速对目标条件测度的采样。

在此示例中，我们评估混合密度模型预测u_2 在给定 u_1时条件分布的性能。此外，我们还考察训练后的混合密度网络在训练期间观测到的条件值范围之外的外推行为。该问题存在平稳条件密度的解析表达式，其推导见附录 D。

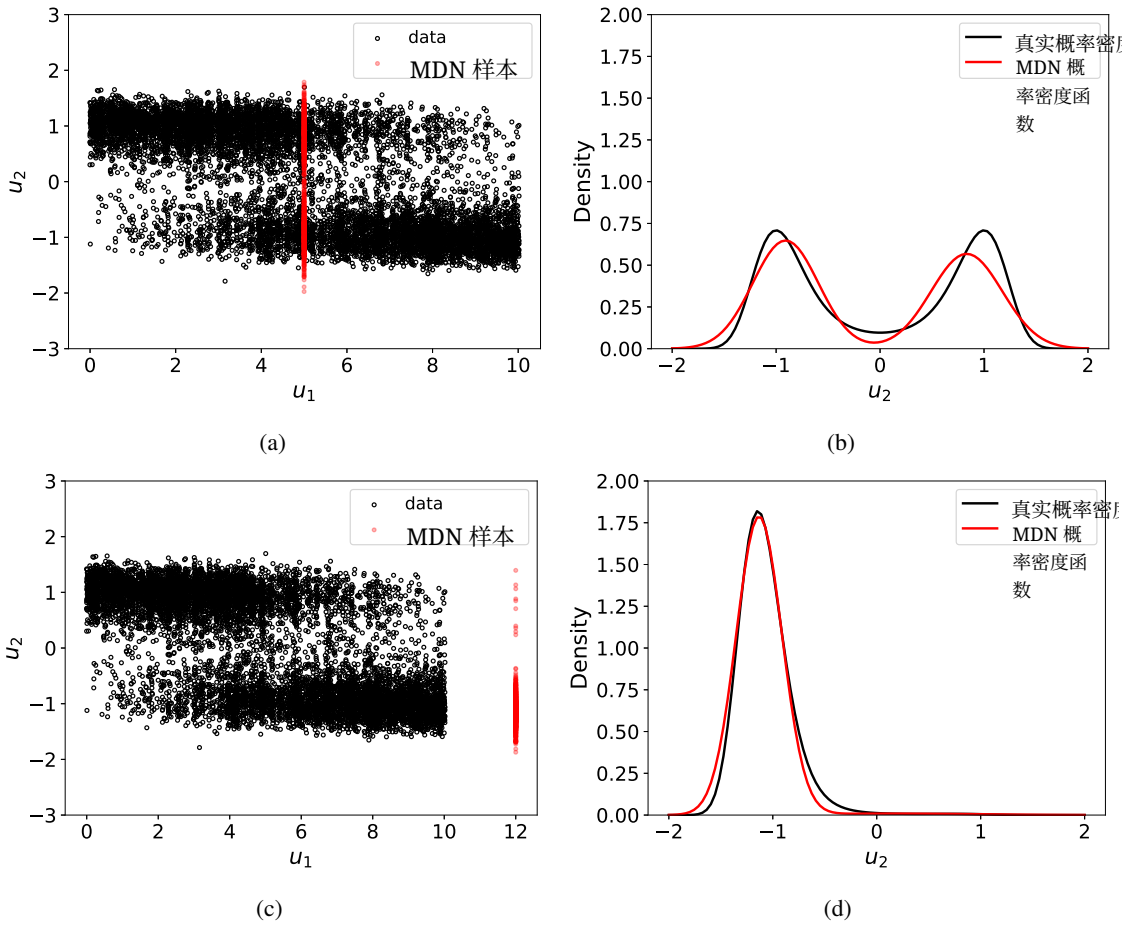


图 4：在 SPDE 数据上训练具有两个混合分量的混合密度网络的结果。为便于查看，从 10，000 个数据点中随机抽取 10，000，000 个 SPDE 数据点。（a）1，000 个由以 $u_1 = 5$ 为条件的混合密度网络模型生成的样本；（b）混合密度网络对应的概率密度函数与真实分布。（c）1，000 个由以 $u_1 = 12$为条件的混合密度网络模型生成的样本；（d）混合密度网络对应的概率密度函数与真实分布。

图 4 展示了混合密度网络在多尺度随机微分方程示例中的表现。图 4a 和 4b 展示了在 $u_1 = 5$ 处查询得到的混合密度网络输出。图4a 展示了混合密度网络在以该值为条件时生成的样本；图 4b 则将解析稳态条件密度与混合密度网络预测的密度进行了比较，结果表明，模型捕捉到了条件分布的双峰结构。图 4c和 4d 报告了$u_1 = 12$对应的相同量，其中该值位于训练范围之外。尽管处于这种外推情形，混合密度网络仍能生成与数据一致的样本，并恢复真实条件密度的整体结构。这些结果表明，混合密度网络能够准确捕捉快变量的多峰条件分布，并且在查询超出训练期间观测到的条件值范围时仍然有效。

3.3 冲击波物理

在第三个示例中，我们考虑单晶 3C–SiC 沿 [001] 取向的冲击雨贡纽曲线，并以冲击波速度–质点速度（U_s–U_p）关系表示。汇总后的数据集涵盖多个变形区间，包括弹性响应、塑性流动和冲击诱导相变，反映了冲击强度增加时波结构的演变。U_s–U_p 数据汇编自分子动力学模拟及 [8, 23, 33], 中报告的实验对比，如图5所示。我们通过曲线数字化，从已发表的图表中复现了数据。获取这些数据点均需耗费大量计算资源进行模拟（例如大规模分子动力学模拟）或开展实验，因此有必要使用机器学习模型作为代理。更重要的是，概率代理模型能够识别值得进一步探索的信息丰富区间，从而指导后续实验或模拟。

弹性区间、塑性区间和相变区间在 U_s–U_p 平面上分别形成不同分支，反映出不同的响应机制和多模态结构。与数据和热力学考量一致，冲击波速度 U_s 是质点速度 U_p 的非递减函数。我们将这一性质作为基于物理的先验，通过基于物理的正则化项施加单调性约束：

$$\mathcal{R}_{\text{phys}}(\mu_m; x) = \max\left(0, -\partial\mu_m\right). \tag{13}$$

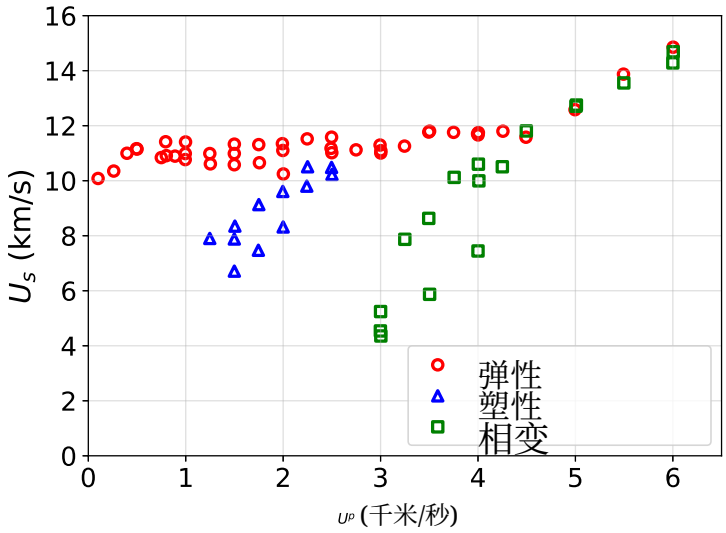


图 5：沿着 [001]取向的单晶 3C–SiC 数据的冲击雨贡纽曲线，数字化并转载自[8, 23, 33]。

图 6 展示了在混合密度网络训练中将单调性作为先验物理知识纳入模型的效果。左列显示了仅使用负对数似然损失公式6训练的三次独立试验结果，右列则显示了加入单调性正则化项公式 13后对应的试验结果。仅使用负对数似然损失进行训练时，各次试验学得的分量均值存在明显差异，且未能始终遵循 U_s 与 U_p之间预期的单调递增关系。具体而言，均值预测往往优先拟合整体数据分布，导致中间区域出现非单调行为。相比之下，加入单调性先验后，各次试验学得的分量均值在 U_p 方面更一致地遵循物理趋势。尽管这一弱物理约束未能完全区分弹性、塑性和相变区间，但它减轻了无约束设置下观察到的非物理行为，并使预测结果更能体现潜在的冲击物理。

在这个问题中，类别标签是已知的先验信息，因此可以将其作为额外监督来提升建模性能。在这种设定下，每个数据点都根据其所属的状态区域进行标记，混合分量也会被明确分配到相应的状态区域。图 7展示了将类别条件结构（公式 7 ）和单调性（公式 13）作为先验知识纳入混合密度网络训练的效果。与图 6中未利用类别信息的结果相比，利用类别信息的模型能够更清晰地区分响应分支，其分量均值也更稳定地与相应的状态区域对齐。这些结果突显了混合密度网络作为多模态物理系统灵活代理模型的价值：它能够将数据驱动学习与可解释的物理约束结构相结合。

即使纳入基于物理的知识，不同模型初始化之间的差异仍然存在。这种现象主要是由于问题的数据有限，大部分差异出现在缺乏数据支持的区域。此外，当前的表述并未显式建模认知不确定性。

3.4 具有随机场初始条件的微分方程

对于最后一个示例，我们考虑一个 Chafee–Infante 反应–扩散方程

$$u_t = u - u^3 + \nu\Delta u, \quad x \in [0, \pi],\ t \in [0, \infty), \tag{14}$$

$$u(0, t) = u(\pi, t) = 0. \tag{15}$$

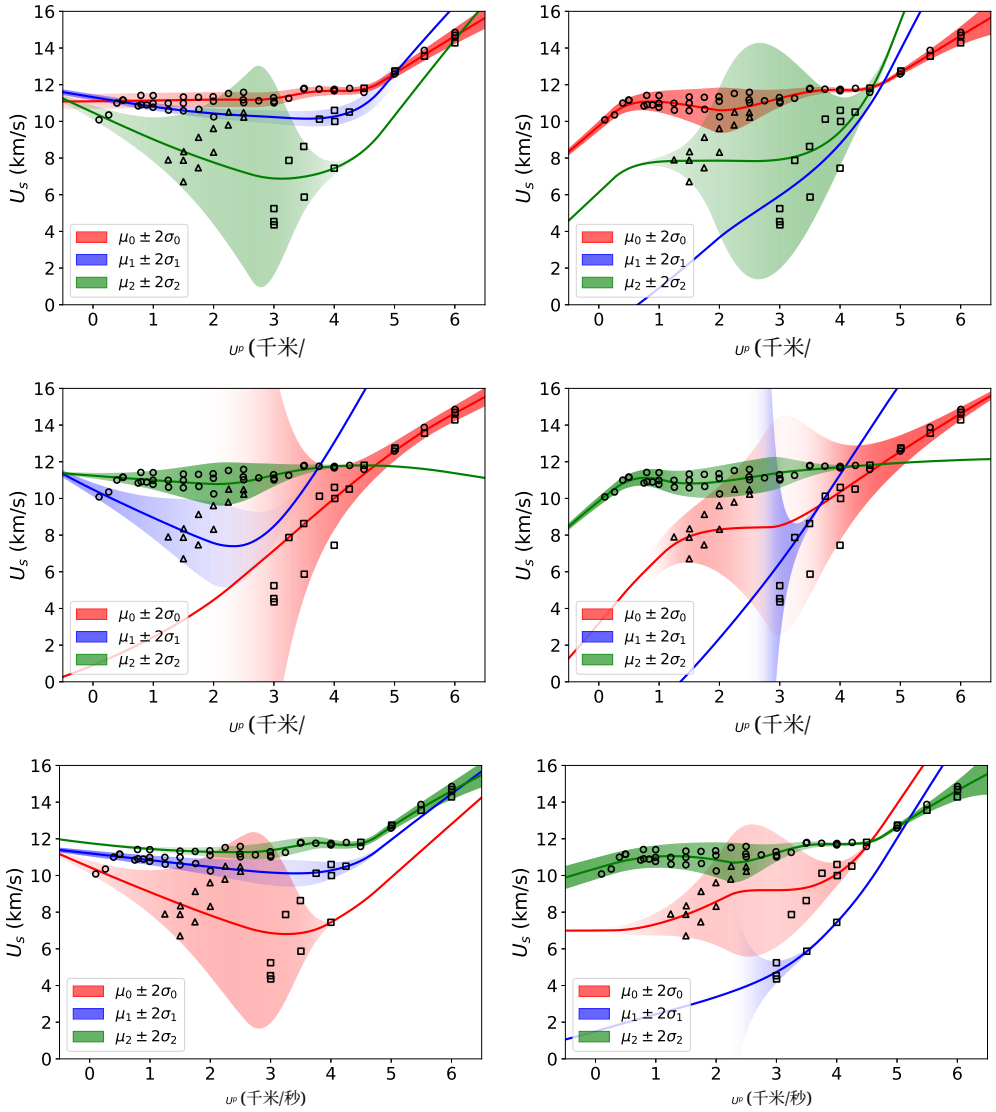


图 6：先验知识对混合密度网络模型训练的影响。经过 10,000 次 ADAM 迭代后，三次不同试验的结果：（左）仅使用负对数似然损失；（右）结合单调递增的先验知识。每个标记代表不同的状态：弹性状态用圆形表示，塑性状态用三角形表示，相变状态用方形表示。

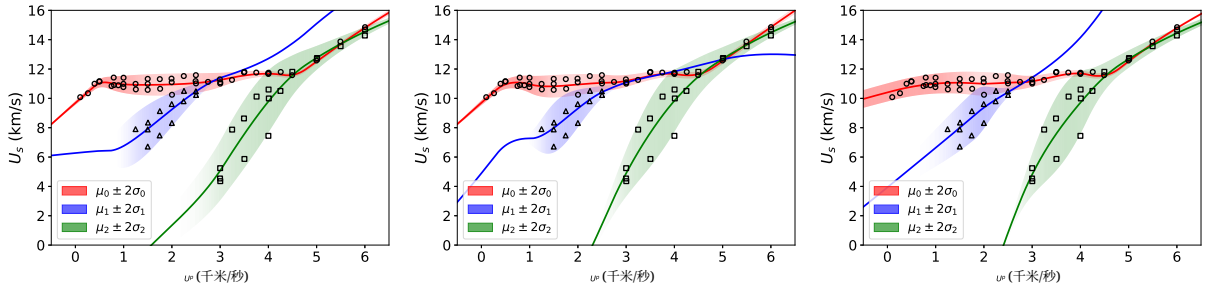


图 7：使用独热编码的类别条件先验知识和单调递增先验知识训练混合密度网络模型的三次不同试验，共进行 10,000 次 ADAM 迭代。

这一问题此前已被研究 [10]，并作为条件分数生成模型的基准 [43]。在此，我们选择 $\nu = 0.16$，其长时间动力学具有多个稳定解，并且已知会在二维惯性流形上演化 [14]。模拟以如下形式的随机初始条件进行初始化

$$u(x,0) = \sum_{n=1}^{3} a_n \sin(nx), \tag{16}$$

其中 $a_n \sim \mathcal{N}(0, 1)$。收集时间 $t = 4.5$（接近稳态）时的 100 个解剖面用于训练。

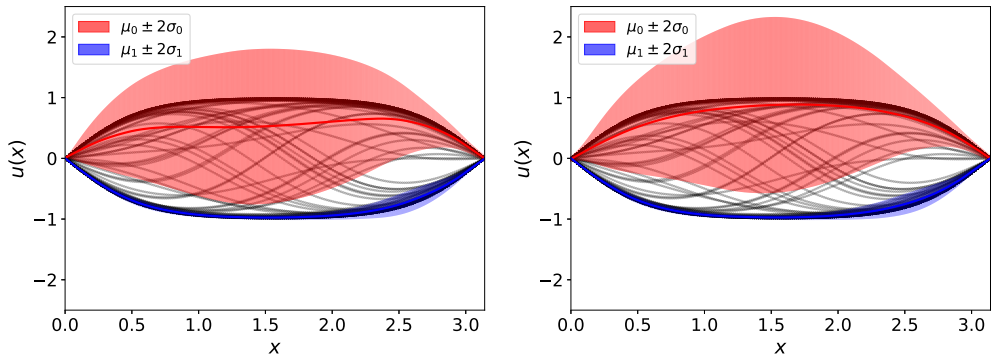
物理 本问题中的物理信息损失根据稳态偏微分方程残差定义如下 ，

$$\mathcal{R}_{\text{phys}}(\mu_m; x) = \left(\mu_m - \mu_m^3 + \nu \Delta \mu_m\right)^2. \tag{17}$$

图 8 比较了在不同训练目标函数下得到的混合密度网络预测结果。如图所示，数据集中在两个不同的解剖面附近。图 8a展示了仅使用负对数似然损失训练混合密度网络所得到的结果。从统计角度看，解系综主要集中在两个稳定模态（向上和向下）附近，模型能够对其进行粗略区分。然而，其中一个预测混合均值明显偏离了相应的稳态解。这种偏差是由于训练数据中存在尚未收敛至平衡状态的中间瞬态状态所导致的。我们有意将这类瞬态轨迹纳入数据集，以评估模型的性能。

图 8b 展示了将负对数似然损失与物理信息损失结合后得到的结果。一旦施加基于物理的先验，模型便会偏向于均值解。因此，与前一种情况相比，预测的混合均值更接近稳定稳态解。同时，预测的标准差增大。这种行为是预期之中的：由于混合均值受到约束而趋向物理上可接受的稳态，模型通过增大方差来解释由非平衡状态引起的剩余数据变异性。尽管这些状态所承载的概率质量低于稳定解，但其贡献仍不为零；增大的方差使模型能够以统计一致的方式表示这些状态。

本例中预测的对称高斯标准差可能显得不切实际，因为数据并未覆盖稳定解的一侧。这是混合密度网络中施加条件高斯假设的结果，
采用更灵活的条件分布进行扩展或许能够解决这一问题，但本文不予讨论。



(a) (b)

图 8：比较经过 50,000 次 ADAM 迭代后，采用（a）普通负对数似然损失和（b）负对数似然损失与基于物理的先验相结合的混合密度网络模型。

4 结论

在这项工作中，我们研究了使用混合密度网络作为灵活且可解释的概率框架，进行受物理约束的多峰条件建模。通过一系列具有代表性的示例——包括分岔现象、多尺度随机动力学、原子尺度冲击物理以及反应–扩散方程——我们展示了如何使用混合密度网络对由内在物理机制产生的多模态条件分布进行建模。

我们进一步表明，在所提出的框架中，可以将多种形式的先验知识——包括热力学原理、基于偏微分方程的约束以及类别条件结构——无缝集成到混合密度网络中。具体而言，我们引入了一种类别条件形式，并在训练过程中将其作为正则项施加，从而引导所学习的

混合分量趋向于对条件分布进行具有物理意义和语义意义的划分。引入此类先验知识，无需额外数据，也无需另行引入类别专属模型，即可增强各混合分量与已知物理行为的一致性，从而在单个概率模型中保留混合分量之间的共享结构和统计耦合。

此外，我们证明，在模型容量和训练工作量相近的情况下，混合密度网络（MDN）的性能可与 CFM（最先进的模型）相媲美，并且在某些情形下表现出比 CFM 更有结构的行为，同时保留了显式混合表示所带来的实际优势，即更简单的训练和采样流程。

重要的是，显式混合结构增强了可解释性：条件概率空间被分解为有限个分量，分别对应于先验定义的区域或状态。这种分解便于针对生成机制进行诊断和后处理，只需聚焦于各个分量，同时降低模型复杂度。此外，由于条件分布具有闭式表达，不确定性量化因而更加透明且便于解析处理——无需依赖蒙特卡洛采样，即可直接计算一阶和二阶条件统计矩。

局限性。虽然条件高斯假设适用于许多问题（包括本研究考察的若干问题），但在其他情形下可能并不现实，例如最后一个示例。尽管所提出的框架支持表达能力更强的条件分布，但本文未探讨这些扩展。此外，尽管该模型能够捕捉预测不确定性，却未明确考虑模型不确定性。

数据可用性

本研究使用的所有数据均已公开，或可根据所提供的信息复现；如有合理需求，可向通讯作者获取。

致谢

B.B. 感谢 Michael Shields 和 Yannis Kevrekidis 在 2025年春季进行的富有成效的讨论，尤其是关于冲击物理数据集和基于分数的生成建模的讨论。B.B. 还感谢西北大学提供的启动基金支持。

生成式人工智能使用声明

作者使用了 OpenAI 开发的大型语言模型 ChatGPT，协助进行英语语言编辑和语法润色。科学内容、技术解读和结论完全由作者负责。

参考文献

[1] Jan-Hendrik Bastek、WaiChing Sun 和 Dennis M Kochmann。物理信息扩散模型。*arXiv* 预印本 *arXiv:2403.14404*，2024年。[2] James L Beck 和 Lambros S Katafygiotis。模型及其不确定性的更新。I：贝叶斯统计框架。《工程力学杂志》，124(4):455–461，1998年。[3] James L Beck 和 Ka-Veng Yuen。利用响应测量进行模型选择：贝叶斯概率方法。《工程力学杂志》，130(2):192–203，2004年。[4] Ilias Bilionis 和 Nicholas Zabaras。多输出局部高斯过程回归：不确定性量化中的应用。《计算物理学杂志》，231(17):5718–5746，2012年。[5] Christopher M Bishop。混合密度网络。1994年。[6] Charles Blundell、Julien Cornebise、Koray Kavukcuoglu 和 Daan Wierstra。神经网络中的权重不确定性。载于国际机器学习会议，第1613–1622页。PMLR，2015年。[7] Sam Bond-Taylor、Adam Leach、Yang Long 和 Chris G Willcocks。深度生成建模：对变分自编码器、生成对抗网络、归一化流、基于能量的模型和自回归模型的比较综述。《*IEEE*模式分析与机器智能汇刊》，44(11):7327–7347，2021年。[8] Paulo S Branicio、Jingyun Zhang、José P Rino、Aiichiro Nakano、Rajiv K Kalia 和 Priya Vashishta。单晶与纳米晶碳化硅的平面冲击加载。《应用物理快报》，112(11)，2018年。

[9] Jie Chen、Yang Yu 和 Yongming Liu。用于不确定性量化的物理引导的混合密度网络。Reliability Engineering & System Safety，228:108823，2022年。[10] Ellis R Crabtree、Dimitris G Giovanis、Nikolaos Evangelou、Juan M Bello-Rivas 和 Ioannis G Kevrekidis。慢流形与分岔图的生成式学习。arXiv 预印本 arXiv:2504.20375，2025年。[11] ER Crabtree、JM Bello-Rivas 和 IG Kevrekidis。多尺度建模中的微观—宏观一致性：基于分数模型辅助的快慢动力系统采样。Chaos: An Interdisciplinary Journal of Nonlinear Science，34(5)，2024年。[12] Kathryn Farrell、J Tinsley Oden 和 Danial Faghihi。用于原子系统粗粒化模型自适应选择、校准与验证的贝叶斯框架。《计算物理学杂志》，295:189–208，2015年。[13] William Fedus、Barret Zoph 和 Noam Shazeer。Switch Transformer：通过简单高效的稀疏性扩展至万亿参数模型。Journal of Machine Learning Research，23(120):1–39，2022年。[14] Charles William Gear、IG Kevrekidis 和 BE Sonday。基于扩散映射参数化的慢流形积分。载于 Numerical Analysis and Applied Mathematics ICNAAM 2011：International Conference on Numerical Analysis and Applied Mathematics，第1389卷，第13–16页，2011年。[15] Roger G Ghanem 和 Pol D Spanos。随机有限元：谱方法。Courier Corporation，2003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。[17] Robert A Jacobs、Michael I Jordan、Steven J Nowlan 和 Geoffrey E Hinton。局部专家自适应混合。Neural Computation，3(1):79–87，1991年。[18] Christian Jacobsen、Yilin Zhuang 和 Karthik Duraisamy。Cocogen：用于正问题与逆问题的物理一致条件化基于分数生成模型。SIAM Journal on Scientific Computing，47(2):C399–C425，2025年。[19] Tero Karras、Miika Aittala、Timo Aila 和 Samuli Laine。阐明基于扩散的生成模型的设计空间。Advances in Neural Information Processing Systems，35:26565–26577，2022年。[20] Irfan Kaymaz。克里金法在结构可靠性问题中的应用。Structural Safety，27(2):133–151，2005年。[21] Marc C Kennedy 和 Anthony O'Hagan。计算机模型的贝叶斯校准。Journal of the Royal Statistical Society: Series B (Statistical Methodology)，63(3):425–464，2001年。[22] Diederik P Kingma 和 Max Welling。自编码变分贝叶斯。arXiv 预印本 arXiv:1312.6114，2013年。[23] WH Li、XH Yao、PS Branicio、XQ Zhang 和 NB Zhang。单晶与纳米晶碳化硅中的冲击诱发剥落。Acta Materialia，140:274–289，2017年。[24] Yaron Lipman、Ricky TQ Chen、Heli Ben-Hamu、Maximilian Nickel 和 Matt Le。用于生成建模的流匹配。arXiv 预印本 arXiv:2210.02747，2022年。[25] Wing Kam Liu、Ted Belytschko 和 A Mani。用于非线性结构动力学的概率有限元。Computer Methods in Applied Mechanics and Engineering，56(1):61–81，1986年。[26] James Martin、Lucas C Wilcox、Carsten Burstedde 和 Omar Ghattas。用于大规模统计逆问题的随机牛顿马尔可夫链蒙特卡罗方法及其在地震反演中的应用。SIAM Journal on Scientific Computing，34(3):A1460–A1487，2012年。[27] Youssef M Marzouk、Habib N Najm 和 Larry A Rahn。用于高效贝叶斯求解逆问题的随机谱方法。《计算物理学杂志》，224(2):560–586，2007年。[28] Geoffrey J McLachlan 和 Kaye E Basford。混合模型：推断与聚类应用。Statistics: Textbooks and Monographs，1988年。[29] Zhongxian Men、Eugene Yee、Fue-Sang Lien、Deyong Wen 和 Yongsheng Chen。利用混合密度神经网络集成进行短期风速与功率预测。Renewable Energy，87:203–211，2016年。[30] Radford M Neal。神经网络的贝叶斯学习，第118卷。Springer Science & Business Media，2012年。[31] Manolis Papadrakakis、Vissarion Papadopoulos 和 Nikos D Lagaros。利用神经网络和蒙特卡罗模拟对弹塑性结构进行结构可靠性分析。Computer Methods in Applied Mechanics and Engineering，136(1-2):145–163，1996年。

[32]George Papamakarios、Eric Nalisnick、Danilo Jimenez Rezende、Shakir Mohamed 和 Balaji Lakshminarayanan。《用于概率建模与推断的归一化流》。Journal of Machine Learning Research，22(57):1–64，2021年。[33] George D Pasparakis、Himanshu Sharma、Rushik Desai、Chunyu Li、Alejandro Strachan、Lori Graham-Brady 和 Michael D Shields。《用于预测冲击波雨贡纽曲线的物理约束高斯过程》。arXiv预印本 arXiv:2601.06655，2026年。[34] Adam Paszke、Sam Gross、Francisco Massa、Adam Lerer、James Bradbury、Gregory Chanan、Trevor Killeen、Zeming Lin、Natalia Gimelshein、Luca Antiga 等。《PyTorch：一种命令式风格、高性能的深度学习库》。Advances in Neural Information Processing Systems，32，2019年。[35] Han Peng 和 Jize Zhang。《随机模拟器的高效、可扩展仿真：一种基于混合密度网络的代理建模框架》。Reliability Engineering & System Safety，257:110806，2025 年。[36] Han Peng 、 Alexandros A Taflanidis 和 Jize Zhang。《利用可扩展混合密度网络随机代理模型加速地震响应分布估计》。《地震工程与结构动力学》，55(3):721–742，2026年。[37] Maziar Raissi、Paris Perdikaris 和 George E Karniadakis。《物理信息神经网络：用于求解涉及非线性偏微分方程的正问题与逆问题的深度学习框架》。《计算物理学杂志》，378:686–707，2019年。[38] Maximilian Rixner 和 Phaedon-Stelios Koutsourelakis。《一种用于半监督训练粗粒化代理模型并通过虚拟可观测量施加物理约束的概率生成模型》。《计算物理学杂志》，434:110218，2021年。[39] Noam Shazeer、Azalia Mirhoseini、Krzysztof Maziarz、Andy Davis、Quoc Le、Geoffrey Hinton 和 Jeff Dean。《规模惊人的大型神经网络：稀疏门控专家混合层》。arXiv 预印本 arXiv:1701.06538，2017年。[40] Dule Shu、Zijie Li 和 Amir Barati Farimani。《用于高保真流场重建的物理信息扩散模型》。《计算物理学杂志》，478:111972，2023年。[41] Timothy W Simpson、Timothy M Mauery、John J Korte 和 Farrokh Mistree。《用于基于仿真的多学科设计优化中全局近似的克里金模型》。《美国航空航天学会期刊》，39(12):2233–2241，2001年。[42] Jascha S ohl-Dickstein、Eric Weiss、Niru Maheswaranathan 和 Surya Ganguli。《利用非平衡热力学进行深度无监督学习》。载于《国际机器学习会议》，第2256–2265页。PMLR，2015年。[43] 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。[44] Pol D Spanos 和 Roger Ghanem。《随机介质的随机有限元展开》。《工程力学杂志》，115(5):1035–1053，1989年。[45] Bruno Sudret。《利用多项式混沌展开进行全局敏感性分析》。Reliability Engineering & System Safety，93(7):964–979，2008年。[46] C Raj Sundararajan。《概率结构力学手册：理论与工业应用》。Springer Science & Business Media，2012年。[47] John Michael Tutill Thompson 和 Giles William Hunt。《弹性稳定性一般理论》。（无标题），1973年。[48] Rohit Unni、Kan Yao 和 Yuebing Zheng。《用于层状光子结构逆向设计的深度卷积混合密度网络》。《ACS光子学》，7(10):2703–2712，2020年。[49] Yanjia Wang、Dong Yang 和 Francis TK Au。《一种利用循环混合密度网络对热载荷作用下桥梁伸缩缝进行可靠性分析的概率框架》。Reliability Engineering & System Safety，第111341页，2025年。[50] Yong Xu、Rencai Gu、Huiqing Zhang、Wei Xu 和 Jinqiao Duan。《有色噪声下双稳态 Duffing–van der Pol 振荡器中的随机分岔》。《物理评论E—统计、非线性与软物质物理》，83(5):056215，2011年。[51] Liu Yang、Xuhui Meng 和 George Em Karniadakis。《B-PINN：用于含噪数据下偏微分方程正问题与逆问题的贝叶斯物理信息神经网络》。《计算物理学杂志》，425:109913，2021年。[52] Zijiang Yang、Dipendra Jha、Arindam Paul、Wei-keng Liao、Alok Choudhary 和 Ankit Agrawal。《一种结合生成对抗网络与混合密度网络的微观结构材料设计逆向建模通用框架》。arXiv预印本 arXiv:2101.10553，2021年。

[53] Hao Zhang、Yongqian Liu、Jie Yan、Shuang Han、Li Li 和 Quan Long。《改进的深度混合密度网络，用于区域风电功率概率预测》。《*IEEE*电力系统汇刊》，35(4):2549–2560，2020年。

A 网络架构

各问题的超参数设置如下所示，顺序与第3 节和附录B中的介绍顺序一致。所有模型均使用学习率 10^{-3}进行训练，并采用两个隐藏层和ELU激活函数。对于使用混合密度网络的分岔示例，隐藏单元数设为16 ；对于使用CFM的分岔示例，设为20 ；对于多尺度随机微分方程、激波、Chafee–Infante和圆形示例，设为32 。

B 圆形数据

作为补充示例，我们考虑一个具有条件多模态性的合成二维圆形数据集。通过在 $[0, 2\pi)$ 内均匀采样角度，并在内半径为 $r_{\text{in}} = 0.35$ 、外半径为 $r_{\text{out}} = 0.5$、中心位于 (0.5,0.5) 的圆环内，按照与面积成正比的密度采样半径，从而生成数据点。共使用400个数据点。在相同的训练设置下，比较采用不同混合分量数量所得到的结果。

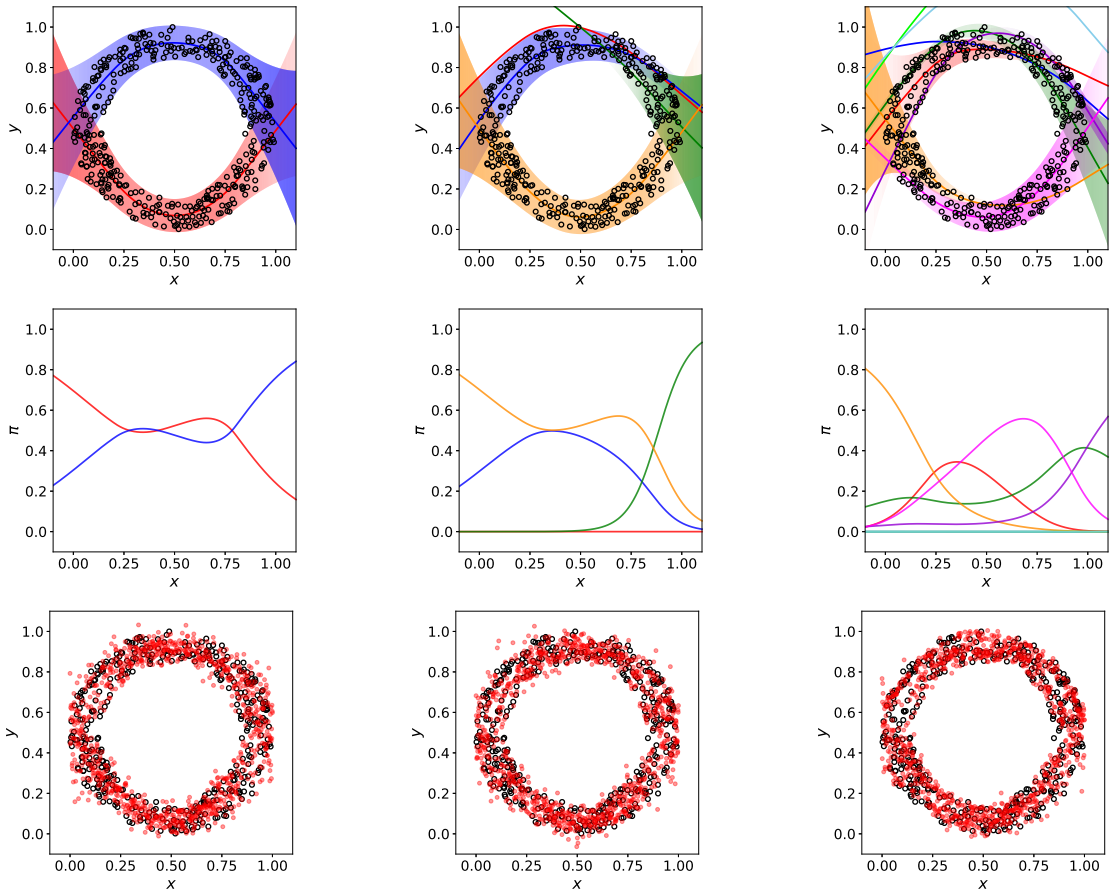


图9：圆形数据集在经过 5,000 次 ADAM 迭代后，采用不同数量混合分量的结果。顶部：学习得到的 $M = 2, 4, 8$ 的条件密度。中部：对应的混合权重 π_m。底部：从每个训练好的混合密度网络模型中抽取的 1,000 个样本。所有模型均在相同设置下训练。

图 9 比较了使用 $M = 2$、4和 8 个混合分量所得到的结果。顶行表明，在所有配置下，学习到的条件密度在视觉上相似，并且都恢复了数据相同的多模态结构。中间一行给出了相应的混合权重 π_m。当 M较大时，在整个输入域内，多个分量的权重都可以忽略不计：在 $M = 4$ 的情况下，一个分量的权重始终接近于零；而在 $M = 8$ 的情况下，只有五个分量具有非零权重，其余三个分量实际上被

抑制。底行展示了从每个训练好的模型中抽取的 1，000 个样本；对于 M 的所有取值，这些样本都与目标圆形分布高度重叠。总的来说，这些结果表明，对于混合分量数量的不同选择，可以得到相近的条件表示。

C 条件流匹配

本附录概述了 CFM 框架 [24]，该框架用于从数据中学习条件生成模型。CFM 旨在通过构建时变向量场，将易处理的基分布变换为目标数据分布，从而近似条件密度 $p_1(u|z,\ x)$。设 $p_0(u)$ 表示定义在 $\mathbb{R}^d$上的简单参考分布（例如，标准高斯分布），而 $p_1(u|z,\ x)$ 表示由数据诱导的条件分布。

为连接这些分布，CFM 引入了一个依赖上下文的时变速度场 $\mathbf{v}_\theta(u,\ t,\ z,\ x)$：$\mathbb{R}^d \times [0,1] \times \mathcal{Z} \times \mathcal{X} \to \mathbb{R}^d$，它通过常微分方程（ODE）定义确定性流：

$$\frac{du(t)}{dt} = \mathbf{v}_\theta(u(t), t, z, x), \qquad u(0) \sim p_0. \tag{18}$$

对于固定的上下文 $(z,\ x)$，从 18 到 $t = 0$ 的 $t = 1$ 积分方程定义了一个推前映射，将基分布 p_0 变换为对 $p_1(\cdot|z,\ x)$的目标近似。

训练采用无模拟目标函数，该目标函数基于噪声与数据样本之间预先设定的插值。给定一个数据样本 $u_1 \sim p_1(\cdot|z,\ x)$ 和一个独立抽样 $u_0 \sim p_0$，我们将条件概率路径定义为线性桥：

$$u_t = (1-t)u_0 + tu_1, \qquad t \in [0,1], \tag{19}$$

其路径速度恒定为 $\mathbf{w}_t := \frac{du_t}{dt} = u_1 - u_0$。通过将学习到的速度场 $\mathbf{v}_\theta$ 与目标速度 $\mathbf{w}_t$ 匹配，对参数 θ 进行优化。具体而言，从数据集中采样 $(z,\ x,\ u_1)$，$u_0 \sim p_0$，并令 $t \sim$ 服从均匀分布$(0,\ 1)$，通过最小化期望 L_2 回归损失来训练模型：

$$\mathcal{L}(\theta) = \mathbb{E}_{t,z,x,u_0,u_1}\left[\left\|\mathbf{v}_\theta(u_t,t,z,x) - (u_1-u_0)\right\|^2\right]. \tag{20}$$

训练完成后，抽取 $u(0) \sim p_0$，并使用数值积分器（例如欧拉格式）沿时间正向求解方程18，同时固定上下文 $(z,\ x)$，即可生成条件随机场的实现。

D 福克–普朗克方程的平稳解

考虑如下的一维伊藤随机微分方程（SDE）

$$du_2 = b(u_2; u_1)dt + a_3 dB(t), \tag{21}$$

其中，u_1 视为固定参数（例如慢变量），$B(t)$是标准布朗运动，$a_3 > 0$ 是常数噪声幅度（加性噪声）。

令 $p(u,t|u_1)$ 表示 $u_2(t)$的概率密度函数（概率密度函数）。相应的福克–普朗克方程如下 [19]:

$$\partial_t p = -\partial_u\big(b(u;u_1)\,p\big) + \frac{a_3^2}{2}\,\partial_{uu}p. \tag{22}$$

在平稳状态下，我们有:

$$-\partial_u\big(b(u;u_1)\,p_\infty(u|u_1)\big) + \frac{a_3^2}{2}\,\partial_{uu}p_\infty(u|u_1) = 0, \tag{23}$$

按如下方式引入概率通量 J

$$J(u) := b(u;u_1)\,p_\infty(u|u_1) - \frac{a_3^2}{2}\,\partial_u p_\infty(u|u_1), \tag{24}$$

方程 23 可写为守恒定律 $\partial_u J(u) = 0$，这意味着 $J(u) \equiv J_0$ 为常数。对于定义在 $\mathbb{R}$ 上且满足 $p_\infty(u|u_1) \to 0$ 当 $|u| \to \infty$ 时的可归一化平稳密度，稳态通量必须为零，即 $J_0 = 0$。因此，

$$\partial_u p_\infty(u|u_1) = \frac{2}{a_3^2}\,b(u;u_1)\,p_\infty(u|u_1), \tag{25}$$

由此得到

$$p_\infty(u_2|u_1) \propto \exp\left(\frac{2}{a_3^2}\int_{s=-\infty}^{u_2} b(s;u_1)\,ds\right). \tag{26}$$

比例 常数由归一化 $\int_{\mathbb{R}} p_\infty(u|u_1)\,du = 1$ 确定，前提是积分有限 ite.

在第3.2节的问题中，漂移函数为 $b(u_2;\, u_1) = -(-1 + 0.2\,u_1 + 4\,u_2(-1 + u_2^2))$。将此表达式代入方程26，并使用梯形法则数值计算相应的归一化因子：

$$p_\infty(u_2|u_1) \propto \exp\left(-\frac{2}{a_3^2}\left((u_2^2-1)^2 + (0.2u_1 - 1)u_2\right)\right). \tag{27}$$