用于重构流场的耦合扩散–编码器模型

AmirPouya Hemmasian1 1,2 1和 Amir Barati Farimani)美国宾夕法尼亚州匹兹堡，
卡内基梅隆大学机械工程系2)美国宾夕法尼亚州匹兹堡，卡内基梅隆大学机器学习系

（*电子邮箱：Barati@cmu.edu）

（*电子邮箱：Hemmas@cmu.edu）

（日期：2026年1月14日）

数据驱动的流场重建通常依赖自编码器架构，将高维状态压缩为低维潜在表示。然而，变分自编码器（VAE）等经典方法在面临强压缩时，往往难以保留流体流动的高阶统计结构。我们提出了 DiffCoder，这是一种耦合框架，将概率扩散模型与传统的卷积 ResNet 编码器相结合，并对两个组件进行端到端训练。编码器将流场压缩为潜在表示，而扩散模型则学习以压缩状态为条件的重建结果生成式先验。该设计使 DiffCoder 能够恢复分布特性和谱特性；这些特性并非最小化逐点重建损失所必需，却对于忠实表征流场的统计特性至关重要。我们在一个具有挑战性的 Kolmogorov流场数据集上，针对多种模型规模和压缩比，对 DiffCoder 与 VAE 基线进行了评估。在激进压缩条件下，DiffCoder 显著提升了谱精度，而 VAE 的性能则大幅下降。尽管两种方法的相对L2重建误差相当，DiffCoder 对流场底层分布结构的保留效果更好。在中等压缩水平下，规模足够大的 VAE 仍具有竞争力，这表明扩散模型先验在信息瓶颈严重时最能发挥优势。这些结果表明，基于扩散模型的生成式解码为复杂流场构建紧凑且统计一致的表示提供了一条很有前景的途径。

arXiv:2601.07946v1 [cs.LG] 12 Jan 2026

一、引言

复杂的流体流动受高维非线性动力学支配，模拟、存储和分析这些流动的成本都很高。在许多科学和工程场景中，尤其是在需要实时控制或嵌入式部署的情况下，人们越来越需要将这些流动压缩为紧凑的表示，同时保留其空间结构和统计结构。传统方法，如本征正交分解（POD）和动态模态分解（DMD），往往无法捕捉混沌状态下存在的非线性多尺度动力学[1]。

基于深度学习的自编码器在数据驱动的流动压缩与重建方面展现出巨大潜力，能够学习将高维状态映射到潜在表示的非线性映射。特别是，变分自编码器（VAE）支持采样和概率建模，但已知其在激进压缩下会产生模糊或平滑的重建结果[2–4]。这种模糊现象很大程度上源于高斯先验和训练过程中使用的变分界的局限性，它们可能无法捕捉流场数据分布的全部复杂性。

生成建模的最新进展，尤其是基于分数的扩散模型，为通过随机去噪过程学习复杂分布提供了一种强有力的替代方案[5,6]。这些模型在计算机视觉、物理模拟和设计等应用中展现出优异的性能。值得注意的是，近期的进展已将扩散模型应用于流场超分辨率[7]，、固体中的应力估计[8]，以及增材制造中的熔池建模[9]，，彰显了它们以高保真度建模复杂物理现象的能力。扩散模型的扩展还可用于逆向设计[10]，、不确定性量化[11]，和基于文本的模拟[12]，进一步展现了它们在结构化物理数据建模方面的多功能性。这些进展支持将扩散模型用作表达能力强的生成先验，以捕捉物理系统中各类量的底层分布。

在本研究中，我们提出**DiffCoder**，这是一种将卷积 ResNet 编码器与条件扩散模型相结合的混合框架。编码器将流场压缩为潜在编码（压缩场），扩散模型则以压缩场为条件，学习重建空间上的生成式先验。这种设置能够实现一种恢复机制，即使在压缩率很高的情况下，也能保留高阶统计量和空间细节。

近期有若干研究探索了传统 CAE 之外的流场建模方法，包括向量量化自编码器、基于 Transformer 的降阶模型以及物理信息损失设计[2,3,13]。DiffCoder 与应用于物理系统[5,7]，的条件扩散最为相关，但不同之处在于，它嵌入了由编码器网络学习得到的表征，并将该网络作为扩散训练过程的一部分进行联合训练。另有研究证明了卷积架构在学习时空生成先验[4,6] 以及将基于扩散的生成模型集成到物理建模流程[11,14]中的有效性。我们的方法建立在这些研究方向之上，并提出了一个端到端训练的统一系统。

本文其余部分安排如下：第二节回顾流场压缩和生成建模方面的先前工作。第三节简要介绍扩散模型的背景。第四节详细介绍 DiffCoder 的架构及其训练与推理过程。在第五节中，我们将展示 DiffCoder 与一个采用相同构建模块且参数量相同的变分自编码器

基线的对比结果，随后进行讨论并作出结论。

II. 相关工作

A. 降阶建模与流场压缩

经典降阶建模技术将高维流动状态投影到低维子空间。本征正交分解（POD）及其变体一直是流动分析与重建的基础[15,16],，而 Gappy-POD 等扩展方法则支持从不完整数据中进行重建[17]。稀疏降阶建模将基于传感器的动力学与全状态估计联系起来[18]。然而，线性方法难以捕捉复杂且混沌的流体流动中的非线性多尺度动力学[1]。

基于深度学习的自编码器通过学习高维状态与紧凑潜在表示之间的非线性映射，解决了这一局限。Erichson 等[19]证明，浅层神经网络能够根据有限的传感器测量值重建流场；Fukami 等[20] 则将 Voronoi 镶嵌与深度学习结合起来，利用稀疏传感器实现全局场重建。采用 β-变分自编码器（VAEs）[2] 及物理约束架构[3] 的扩展方法提升了泛化能力，但由于高斯似然假设的局限性，基于变分自编码器的解码器在高强度压缩下往往会生成过度平滑的重建结果。

B. 用于流场超分辨率的生成模型

生成对抗网络（GAN）已成功应用于流体流动的超分辨率。Güemes 等[21] 针对随机播种场开发了超分辨率生成对抗网络，在实验环境中获得了高质量重建结果。物理信息生成对抗网络已用于湍流增强[22],，无监督深度学习方法也展现出在湍流超分辨率重建方面的潜力[23]。然而，GAN 训练不稳定和模式坍塌仍是持续存在的挑战。在图像超分辨率领域，深度卷积网络[24] 和迭代细化方法[25] 已建立了强有力的基线，并被改编用于科学应用。

C. 物理系统的扩散模型

扩散模型已成为强大的生成先验，在图像合成质量方面超越了 GAN[26]。Ho 等[27]奠定了 DDPM 框架的基础，而 Song 等[28] 则提出了 DDIM，以加速确定性采样。改进的训练策略[29] 以及基于信噪比[30] 的高效损失加权方案，提升了样本质量和训练稳定性。

在流体力学领域，Shu 等[7] 开发了物理信息扩散模型，用于高保真流场重建，其谱精度优于确定性基线。Liu 和 Thuerey[31] 将扩散模型用作不确定性感知代理模型，以模拟翼型流动。Yang 和 Sommer[32] 探索了扩散模型在流体场预测中的应用；Li 等[33] 利用生成式扩散模型重建湍流旋转流；Abaidi 和 Adams[34] 则研究了扩散模型在可压缩流体场预测中的应用。

物理信息方法已融入扩散模型训练：Bastek 等[35] 开发了物理信息扩散模型，Shan 等[36] 提出了用于流场重建的物理信息残差扩散。Lu 和 Xu[37] 提出了利用物理引导扩散对偏微分方程求解器进行生成式降尺度的方法。

III. 背景：扩散模型

扩散模型是一类生成模型，通过学习逆转逐渐加噪的过程来生成数据。给定从目标分布中抽取的数据$\mathbf{x}_0$，前向过程会在T个时间步内，通过添加高斯噪声逐步破坏数据：

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

其中 $\{\beta_t\}_{t=1}^T$ 是控制每个步骤噪声大小的方差调度。常见的选择包括线性调度[27] 和余弦调度[29]。另一种参数化方式是通过对数信噪比定义调度 $\lambda_t = \log(\bar{\alpha}_t/(1-\bar{\alpha}_t))$，由此可得 $\bar{\alpha}_t = \sigma(\lambda_t)$，其中 σ表示 Sigmoid 函数。

定义 $\alpha_t = 1-\beta_t$和 $\bar{\alpha}_t = \prod_{s=1}^t \alpha_s$后，任意时间步$t$的边缘分布都可以用闭式形式表示：

$$q(\mathbf{x}_t|\mathbf{x}_0) = \mathcal{N}(\mathbf{x}_t; \sqrt{\bar{\alpha}_t}\,\mathbf{x}_0, (1-\bar{\alpha}_t)\mathbf{I}). \tag{2}$$

这使得可以通过重参数化$\mathbf{x}_t = \sqrt{\bar{\alpha}_t}\,\mathbf{x}_0 + \sqrt{1-\bar{\alpha}_t}\,\boldsymbol{\epsilon}$，直接从$\mathbf{x}_0$采样$\mathbf{x}_t$，其中 $\boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0},\mathbf{I})$。

生成过程通过学习反向过程$p_\theta(\mathbf{x}_{t-1}|\mathbf{x}_t)$来实现，该过程从高斯噪声$\mathbf{x}_T \sim \mathcal{N}(\mathbf{0},\mathbf{I})$开始，迭代地对样本去噪。反向过程被参数化为具有学习所得均值的高斯分布：

$$p_\theta(\mathbf{x}_{t-1}|\mathbf{x}_t) = \mathcal{N}(\mathbf{x}_{t-1}; \boldsymbol{\mu}_\theta(\mathbf{x}_t, t), \sigma_t^2\mathbf{I}). \tag{3}$$

训练通过最小化一个可化简为去噪目标的变分界来进行。网络可以训练为预测噪声 $\boldsymbol{\epsilon}$、干净数据$\mathbf{x}_0$, 或速度$\mathbf{v}_t = \sqrt{\bar{\alpha}_t}\,\boldsymbol{\epsilon} - \sqrt{1-\bar{\alpha}_t}\,\mathbf{x}_0$。Salimans 和 Ho[38], 提出的速度参数化在所有噪声水平下都具有数值稳定性，尤其适用于少步采样。

去噪扩散隐式模型（DDIM）[28] 将反向过程重新表述为确定性映射，使得采样所需的步数远少于原始 DDPM 表述[27]。给定时间步t处的预测估计值 $\hat{\mathbf{x}}_0$，DDIM 的计算如下：

$$\mathbf{x}_{t-1} = \sqrt{\bar{\alpha}_{t-1}}\,\hat{\mathbf{x}}_0 + \sqrt{1-\bar{\alpha}_{t-1}}\,\frac{\mathbf{x}_t - \sqrt{\bar{\alpha}_t}\,\hat{\mathbf{x}}_0}{\sqrt{1-\bar{\alpha}_t}}. \tag{4}$$

这种确定性表述允许跳过时间步，将采样步数从数百步减少到几十步，同时保持生成质量。

条件扩散模型通过将辅助信息$\mathbf{c}$ 纳入去噪网络来扩展这一框架，从而得到 $\boldsymbol{\epsilon}_\theta(\mathbf{x}_t,t,\mathbf{c})$，或等价地表示为$\mathbf{v}_\theta(\mathbf{x}_t, t,\mathbf{c})$。常见的条件机制包括沿通道维度拼接、交叉注意力或自适应归一化层。在我们的设定中，条件信号$\mathbf{c}$ =$\mathbf{z}$是由确定性编码器生成的压缩表示，使扩散模型能够据此重建高保真流场。

IV. 方法

基于扩散模型比基于高斯分布的解码器具有更丰富生成先验这一前提，**Diff-Coder** 通过两阶段设计将这一理念付诸实践：首先由确定性编码器压缩流场，再由条件扩散解码器通过迭代去噪重建流场。以下小节将介绍其架构、训练与推理，以及基线配置。

A. 架构概览

DiffCoder的编码器和条件去噪器由卷积与残差构件组成，使其能够为生成式重建过程（即反向扩散）学习最佳编码。图1展示了整体架构（上方）及其构件（下方）。

在变分自编码器基线中，编码器将输入场$\mathbf{x}$ 压缩为潜在表示$\mathbf{z}$，再由对称解码器直接重建$\mathbf{x}$ 。在 DiffCoder 中，同一编码器生成$\mathbf{z}$，但重建由 U-Net 在反向扩散过程中完成。U-Net 的输入包括噪声场$\mathbf{x}_t$、时间步t （通过平移缩放调制注入）以及条件信号$\mathbf{z}$ （通过拼接和可选的交叉注意力提供）。U-Net 不会直接输出重建结果，而是预测反向扩散过程中使用的速度$\mathbf{v}$ ，并在多个去噪步骤中迭代地细化输出。

基本构建单元是卷积块，由一个二维卷积层、均方根归一化和 SiLU 激活函数组成。当需要时间条件时（仅限 U-Net），该模块会通过线性投影，根据时间步嵌入引入自适应平移缩放调制。ResNet 块由两个卷积块和一个残差连接组成。每个层包含若干 ResNet 块，后接可选的自注意力和交叉注意力模块，再执行下采样或上采样操作；交叉注意力仅用于 U-Net，以压缩表示作为条件。

该**编码器**由一系列编码器层（E 层）组成，每层包含两个 ResNet 块，后接一个下采样操作（最后一层除外）。这一过程会逐步将输入场 $\mathbf{x} \in \mathbb{R}^{H\times W}$ 压缩为压缩表示 $\mathbf{z} \in \mathbb{R}^{h\times w}$，其中 $h < H$ 且 $w < W$。可以根据问题和数据集的特性进行调整，例如针对周期性边界条件使用循环填充。变分自编码器**解码器**以相反顺序镜像编码器：由一系列解码器层（D 层）组成，每层包含两个 ResNet 块和一个上采样操作。

U-Net采用标准的编码器-解码器结构，并带有跳跃连接[39]。收缩路径（E 层）通过连续的层进行下采样，中间层处理瓶颈特征，扩展路径（D 层）则通过来自对应编码器层的跳跃连接进行上采样。U-Net 使用时间步嵌入，通过移位和缩放操作调制卷积块，并将压缩场 $\mathbf{z}$与输入拼接后接收，也可选择使用交叉注意力。经验上，我们发现，在此应用中，不含注意力模块的模型在重建质量方面优于基于注意力的变体。这一发现与流场中以局部空间相关性为主的特征一致，因此卷积操作比全局注意力机制更适用。

B. 训练与推理

训练遵循第三节所述的离散时间扩散框架，包含 $T = 1000$ 个时间步。我们采用 Sigmoid 噪声调度[40,41], , 其将对数信噪比定义为线性插值：

$$\lambda_t = (1-t/T)\lambda_{\min} + (t/T)\lambda_{\max}, \tag{5}$$

其中 $\lambda_{\min}$ 和 $\lambda_{\max}$ 是控制噪声范围的超参数，而 $\bar{\alpha}_t = \sigma(\lambda_t)$。与线性替代方案相比，该调度在中间噪声水平上分配了更多容量。

对于每个训练样本，均匀抽取一个时间步 $t \in [1,T]$，网络根据编码器生成的潜在表示 $\mathbf{z}$，预测以其为条件的速度目标$\mathbf{v}_t$：

$$\hat{\mathbf{v}}_t = \mathbf{v}_\theta(\mathbf{x}_t, t, \mathbf{z}). \tag{6}$$

在三种常见的预测目标（$\boldsymbol{\epsilon}$、$\mathbf{x}_0$, 和 $\mathbf{v}$）中，速度目标函数在我们的实验中带来了最稳定的优化和最准确的重建结果[38]。训练损失为：

$$\mathscr{L} = \mathbb{E}_{\mathbf{x}_0,t,\boldsymbol{\epsilon}}\left[w_t \left\|\mathbf{v}_\theta(\mathbf{x}_t,t,\mathbf{z}) - \mathbf{v}_t\right\|_2^2\right], \tag{7}$$

其中 $w_t = \max(\mathrm{SNR}(t),1)$，且 $\mathrm{SNR}(t) = \bar{\alpha}_t/(1-\bar{\alpha}_t)$ 是截断信噪比权重[38]，用于强调高噪声时间步。编码器和 U-Net 使用 AdamW 结合 OneCycle 学习率调度进行联合优化。

在推理时，通过 DDIM 采样（式 4）进行重建。从高斯噪声$\mathbf{x}_T \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$开始，模型在每一步以压缩后的场$\mathbf{z}$ 为条件进行迭代去噪，将随机噪声转化为物理上合理的流场。所有实验均使用 20 个 DDIM 步数。

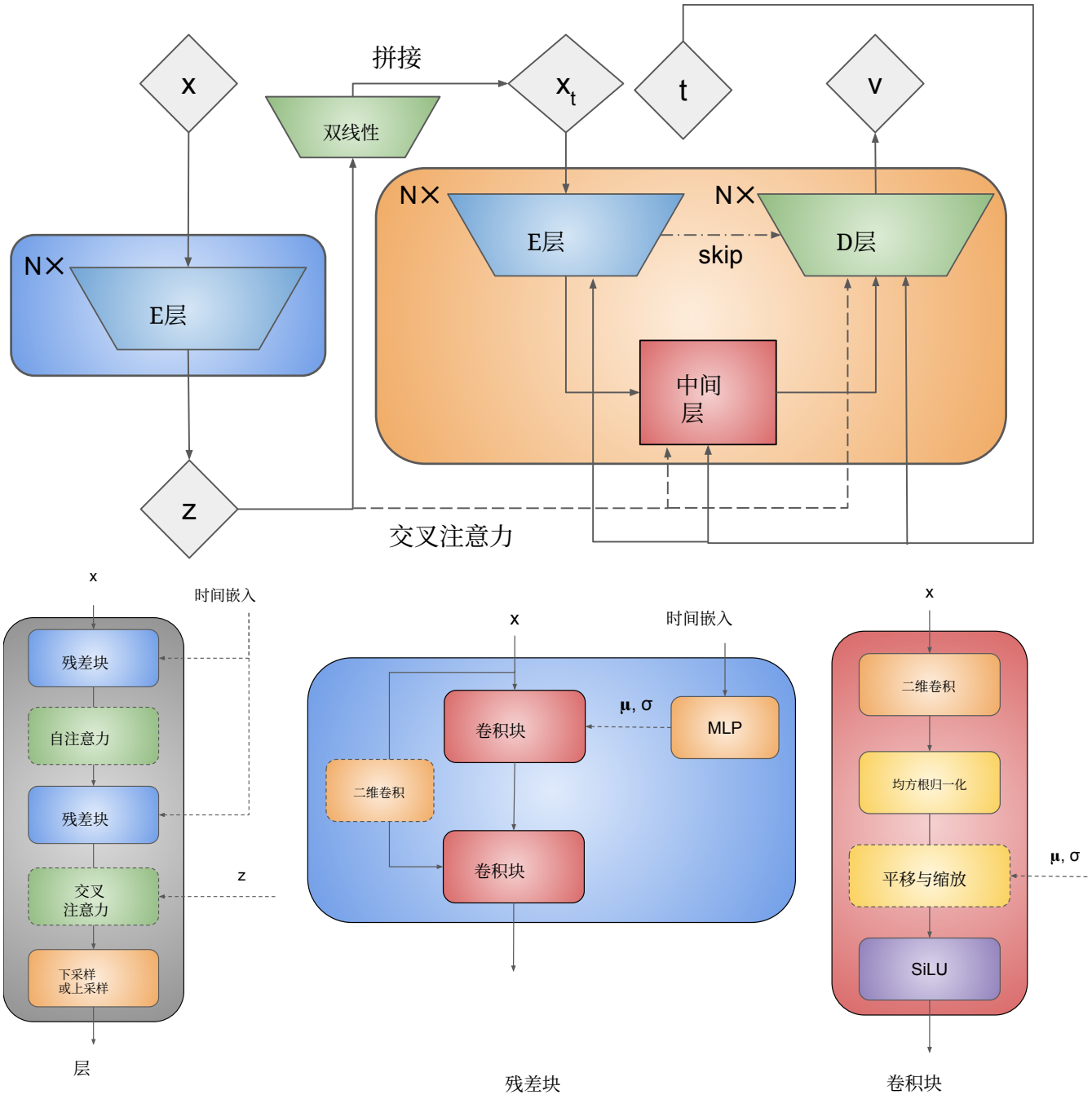


图 1. 上：DiffCoder 中与解码器对应的部分是一个条件 U-Net，它在反向扩散过程中执行迭代去噪。编码器生成 **z**，并通过与输入拼接以及可选的交叉注意力来对 U-Net 进行条件化。U-Net 通过平移缩放调制接收时间步 t，并预测反向扩散过程的目标（在本文中为速度 **v**）。下：DiffCoder 和变分自编码器基线的构建模块。

为单独评估基于扩散的解码器的贡献，我们将其与采用匹配编码器架构的变分自编码器基线进行比较。该变分自编码器使用相同的编码器和对称解码器（如上所述），并以标准证据下界（ELBO）目标进行训练，该目标结合了重建损失和 KL 散度正则化。此比较确保性能差异源自生成建模方法，而非架构容量。

V. 实验

我们在 Re = 1000 下的二维科尔莫哥洛夫流上评估 DiffCoder，并与变分自编码器基线进行比较，涵盖

不同的压缩比和模型容量。

A. 数据集与设置

我们使用 Li 等的工作中提供的 Kolmogorov 流数据集[42]，该数据集包含在雷诺数 $\mathrm{Re} = 1000$、具有周期性边界条件的 256×256 网格上模拟得到的 120 条柯尔莫哥洛夫流轨迹的涡量场快照。我们使用其中 100 条轨迹进行训练，20 条进行测试。

变分自编码器基线采用与 DiffCoder 相同的编码器架构，并配有对称的卷积解码器。两个模型均使用 AdamW 优化器训练，采用 OneCycle 学习率调度策略，最大学习率为 5×10^{-4}，批量大小为 16，共训练 10 个 epoch。推理时，DiffCoder 使用 20 个 DDIM 步数。

B. 评估指标

由于我们希望评估模型的谱精度，而不仅仅是逐点精度，因此定义了两个额外指标来评估结果，总计三个指标如下：

a. 相对 L_2 误差（涡量）。我们计算重建涡量场与真实涡量场之间的相对L_2 误差：

$$\varepsilon_\omega = \frac{\|\omega_{\mathrm{rec}} - \omega_{\mathrm{gt}}\|_2}{\|\omega_{\mathrm{gt}}\|_2}, \tag{8}$$

其中，为计算范数，将场展平为向量。

b. 全频谱误差。为评估能量谱的保真度，我们计算对数动能谱的相对L_2 误差：

$$\varepsilon_E = \frac{\|\log E_{\mathrm{rec}}(k) - \log E_{\mathrm{gt}}(k)\|_2}{\|\log E_{\mathrm{gt}}(k)\|_2}, \tag{9}$$

其中，$E(k)$ 表示波数k的动能谱。

c. 高波数谱误差。由于小尺度结构最难重建，我们还报告仅针对最高 25% 波数的谱误差，记作 $\varepsilon_E^{\mathrm{high}}$。

C. 结果

首先，我们使用 8M 参数规模、编码器深度为 3 的模型开展了一系列实验（压缩至 32×32 个场，意味着 $64\times$ 压缩），以确定注意力机制是否有利于 DiffCoder 或基线变分自编码器。我们评估了四种配置：编码器和解码器（或 U-Net）中均不使用注意力机制、仅在编码器中使用注意力机制、仅在解码器（或 U-Net）中使用注意力机制，以及在这两个组件中均使用注意力机制。对于 DiffCoder，完全移除注意力机制的表现最佳，这可能是因为相对较小的空间维度和流结构的局部特性无法从全局注意力机制中获益。我们还比较了线性、余弦和 Sigmoid 噪声调度。Sigmoid 噪声调度效果最佳，能在扩散过程中使信噪比平滑过渡。基于

在这些消融实验中，后续所有实验均使用不带注意力机制且采用 Sigmoid 噪声调度的 DiffCoder。实验结果显示，模型性能得到提升。图 2 总结了两种模型在四种注意力配置下的结果。

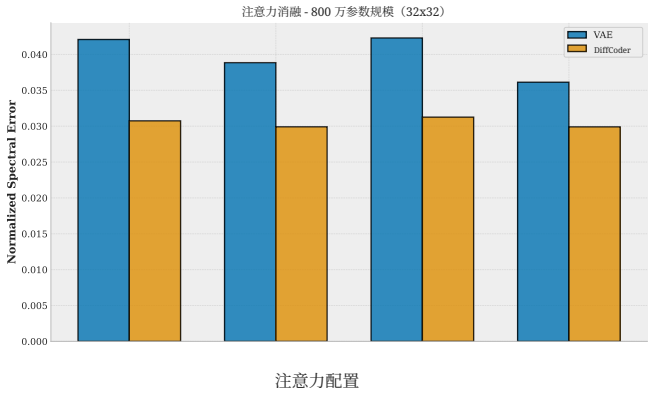


图 2。针对 800 万参数规模和编码器深度 3（64× 压缩）的注意力机制消融研究。比较四种注意力配置下的变分自编码器和 DiffCoder：无注意力、仅编码器、仅解码器/U-Net，以及编码器+解码器/U-Net。结果展示了不同注意力机制对归一化谱误差的影响（越低越好）。

接下来进行主要实验。我们在五种模型规模下训练 DiffCoder 和变分自编码器基线：1M、2M、4M、800 万参数和 16M，其中每个组件（编码器、解码器/U-Net）均按比例扩展。我们评估三种编码器深度设置：深度 2、深度 3 和深度 4，分别对应逐步加深的压缩。每增加一个深度级别，各轴的空间下采样因子都会增加 2 的幂次。图 3–5 展示了不同编码器深度和模型规模下的三项评估指标，表 I–III 则总结了所有配置中 DiffCoder 相对于变分自编码器基线的相对性能。接下来，我们将进一步考察每个压缩区间。

a. 轻度压缩（编码器深度 *2*）。当压缩程度很低时，在所有模型规模下，变分自编码器在涡量和谱指标方面均优于 DiffCoder。当潜在表示保留了大部分原始信息时，变分自编码器较为简单的解码器便已足够。高频谱误差（图 5）呈现出相似的趋势：两种模型都能较好地保留精细尺度内容，而变分自编码器的优势较为明显。

b. 中度压缩（编码器深度 *3*）。在这一中等压缩级别下，两种方法的性能相当；模型规模较小时，DiffCoder 在谱保真度方面更具优势。图 4 和图 5 显示了这一转折：随着模型规模增大，变分自编码器的曲线在谱误差方面逐渐接近或与 DiffCoder 的曲线相交，这表明只要参数足够多，变分自编码器同样可以出色地匹配频谱。

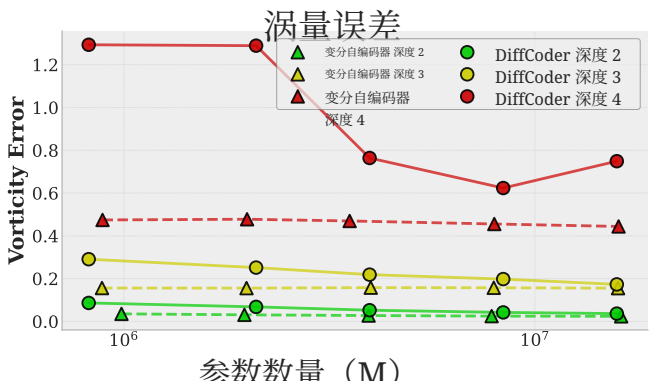


图 3. 不同模型规模和编码器深度下的涡量误差。随着容量增加，两种模型的表现均略有提升，其中变分自编码器始终表现更优。在更深层压缩下，两种模型的表现都未达到可接受水平：变分自编码器甚至未达到 40%，而 DiffCoder 的表现更差。

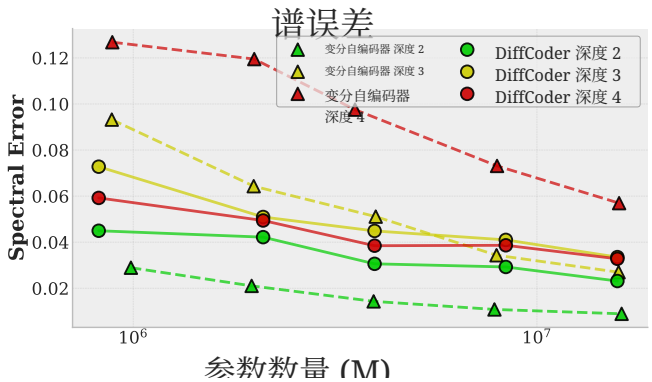


图 4. 不同模型规模和编码器深度下的谱误差。在编码器深度为 4 时，DiffCoder 展现出显著优势；此时谱内容损失最严重，也最难保留。在较浅的深度下，两种方法的表现相近；对于较小的模型规模，DiffCoder 仅在深度 3 时优于变分自编码器，而在其余实验中均逊于变分自编码器。

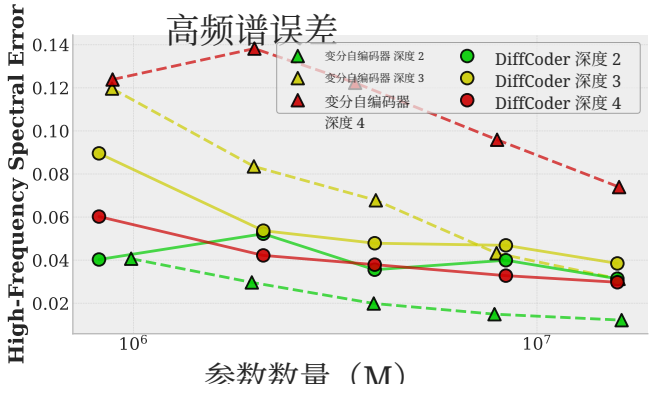


图 5. 不同模型规模和编码器深度下的高频谱误差。在深度 4 时，DiffCoder 的优势最为显著，此时高波数内容几乎完全丢失，甚至可能全部丢失。该指标用于衡量深度压缩下恢复小尺度结构的能力。

c. 深度压缩（编码器深度 *4*）。在激进压缩条件下，DiffCoder 在谱指标方面展现出明显优势。尽管两种模型的涡量重建误差均有所升高（图 3），反映出信息损失不可避免，但 DiffCoder 的谱误差显著低于变分自编码器（图 4）。在高波数谱误差 $\varepsilon^{\text{high}}{}_{E}$ 方面，这一优势最为突出（图 5），表明即使逐点重建颇具挑战，DiffCoder 仍能更好地恢复小尺度湍流结构。

这一趋势证实了我们的核心假设：当压缩瓶颈迫使解码器生成合理的高频内容，而非仅从潜在信息中重建时，基于扩散的解码所带来的益处最大。

D. 定性结果

我们在附录 A 中展示了 16M 模型在编码器深度 4 时，谱误差分别位于第 40 和第 60 百分位数的代表性测试样本重建结果。每个样本均以马赛克图呈现，其中包括真实值、模型（变分自编码器或 DiffCoder）的重建结果、插值基线（双线性和双三次插值）、相应的误差场、评估指标条形图以及动能谱。

变分自编码器的重建结果显得过于平滑，缺少真实值中存在的精细尺度结构。插值方法产生的结果更加模糊，无法恢复高频内容。DiffCoder 的重建结果虽然未能逐点匹配真实值，却呈现出逼真的小尺度特征，其视觉纹理与流场更为相符。

误差场显示，变分自编码器的误差在空间上与涡量梯度较大的区域相关，表明其存在系统性平滑。DiffCoder 的误差有时幅值更大，但在空间上分布更均匀，这说明即使模型无法精确重建，也能生成合理的结构。能量谱子图对此进行了量化：变分自编码器的谱在高波数处衰减，而 DiffCoder 在整个波数范围内都能保持更接近真实值的谱含量。

VI. 结论

我们提出了 DiffCoder，这是一种将卷积编码器与作为解码器的条件扩散模型相结合，用于压缩和重建流场的框架。DiffCoder 以迭代生成过程取代变分自编码器的简单解码器，从而重建仅凭高度压缩的表示无法推断出的、在物理上合理的高频内容。

我们对二维科尔莫哥洛夫流的实验揭示了基于扩散的解码具有明显优势的情形：在激进压缩（编码器深度 4）下，Diff-Coder 在保留动能谱方面显著优于匹配的变分自编码器基线，尤其是在高

表 I。相对涡量误差：DiffCoder 与变分自编码器的对比。数值显示为 $\varepsilon_{\omega}^{\mathrm{VAE}} \rightarrow \varepsilon_{\omega}^{\mathrm{Diff}}$，括号内为百分比变化。负值（绿色）表示 DiffCoder 的性能优于变分自编码器。

模型规模	深度 2	深度 3	深度 4
1M	0.0353 → 0.0863 (+144.8%)	0.1561 → 0.2908 (+86.3%)	0.4744 → 1.2933 (+172.6%)
2M	0.0309 → 0.0679 (+120.0%)	0.1557 → 0.2516 (+61.6%)	0.4777 → 1.2891 (+169.9%)
4M	0.0275 → 0.0524 (+90.8%)	0.1579 → 0.2187 (+38.4%)	0.4696 → 0.7641 (+62.7%)
8M	0.0248 → 0.0421 (+69.5%)	0.1573 → 0.1979 (+25.8%)	0.4555 → 0.6236 (+36.9%)
16M	0.0238 → 0.0365 (+53.8%)	0.1555 → 0.1731 (+11.3%)	0.4440 → 0.7491 (+68.7%)

表 II. 相对谱误差：DiffCoder 与变分自编码器的比较。数值显示为 $\varepsilon_{\mathrm{spec}}^{\mathrm{VAE}} \rightarrow \varepsilon_{\mathrm{spec}}^{\mathrm{Diff}}$，括号内为百分比变化。负值（绿色）表示 DiffCoder 的表现优于变分自编码器。

模型规模	深度 2	深度 3	深度 4
1M	0.0289 → 0.0450 (+55.5%)	0.0932 → 0.0728 (-21.9%)	0.1268 → 0.0592 (-53.3%)
2M	0。0210 → 0.0422（+100.5%）	0。0643 → 0.0509（-20.8%）	0。1195 → 0.0494（-58.6%）
4M	0.0143 → 0.0306 (+113.9%)	0.0511 → 0.0449 (-12.2%)	0.0975 → 0.0385 (-60.6%)
8M	0.0108 → 0.0293 (+170.6%)	0.0343 → 0.0410 (+19.4%)	0.0731 → 0.0387 (-47.1%)
16M	0。0089 → 0.0231（+159.6%）	0。0270 → 0.0335（+24.3%）	0。0570 → 0.0328（-42.4%）

表 III. 相对高频谱误差：DiffCoder 与变分自编码器的比较。数值显示 $\varepsilon_{\mathrm{HF}}^{\mathrm{VAE}} \rightarrow \varepsilon_{\mathrm{HF}}^{\mathrm{Diff}}$，百分比变化置于括号中。负值（绿色）表示 DiffCoder 优于变分自编码器。

模型规模	深度 2	深度 3	深度 4
1M	0.0407 → 0.0403 (-0.9%)	0.1197 → 0.0896 (-25.1%)	0.1239 → 0.0602 (-51.4%)
2M	0.0297 → 0.0523 (+76.0%)	0.0835 → 0.0537 (-35.7%)	0.1382 → 0.0422 (-69.5%)
4M	0.0199 → 0.0356 (+78.8%)	0.0678 → 0.0479 (-29.4%)	0.1223 → 0.0379 (-69.0%)
8M	0.0149 → 0.0400 (+167.4%)	0.0432 → 0.0469 (+8.4%)	0.0959 → 0.0328 (-65.8%)
16M	0.0122 → 0.0314 (+157.9%)	0.0313 → 0.0385 (+22.8%)	0.0739 → 0.0298 (-59.8%)

波数处，这些波数对应小尺度湍流结构。在轻度压缩（depth2）下，这一优势有所减弱，因为此时有足够的信息通过瓶颈得以保留，从而能够进行确定性重建。

这些发现对大规模流体流动模拟具有实际意义，在这类模拟中，存储限制使激进压缩成为必要。尽管在极高压缩比下，逐点重建误差可能难以避免，但保留统计特性和谱特性，或许能够对压缩数据进行更有意义的下游分析。

未来仍有若干值得探索的方向。将 Dif- fCoder 扩展到三维流动，可用于检验其在与实际应用最相关的高维场景中的可扩展性。学习得到的潜在空间也为潜在空间预测带来了可能：训练时序模型（例如 Transformer 或神经常微分方程），直接在压缩表示中预测动力学，再按需解码轨迹。这种方法有望在实现高效长时域预测的同时，利用扩散解码器在每个时间步重建具有统计保真度的流场。此外，纳入物理信息损失（例如无散约束或谱惩罚项）还可进一步提升统计保真度。

数据可用性声明

用于模型训练和评估的数据将在论文发表后公开。

附录 A：代表性样本重建 ions

本附录展示了 16M 模型在编码器深度 4 时（压缩场大小为 64×64）的选定测试样本在谱误差第 40 和第 60 百分位数处的代表性重建结果。每幅图均包含真实值、模型重建、双线性和双三次插值基线、误差场、重建误差条形图以及动能谱对比。

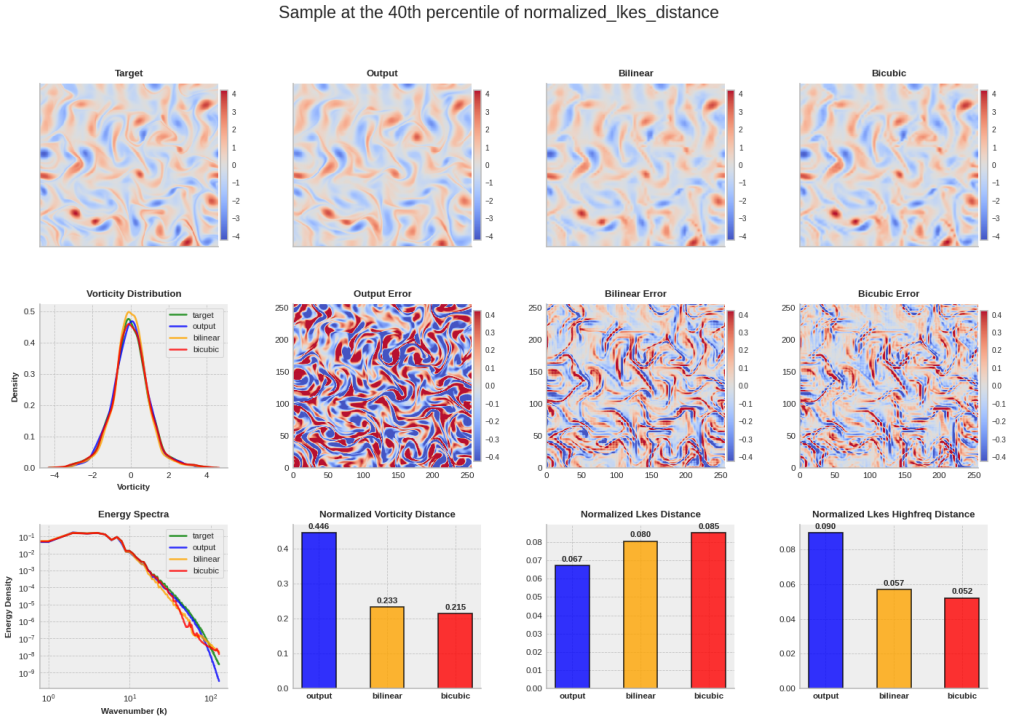


图 6. 谱误差第 40 百分位数处的变分自编码器重建结果（模型规模 16M，深度 4）。

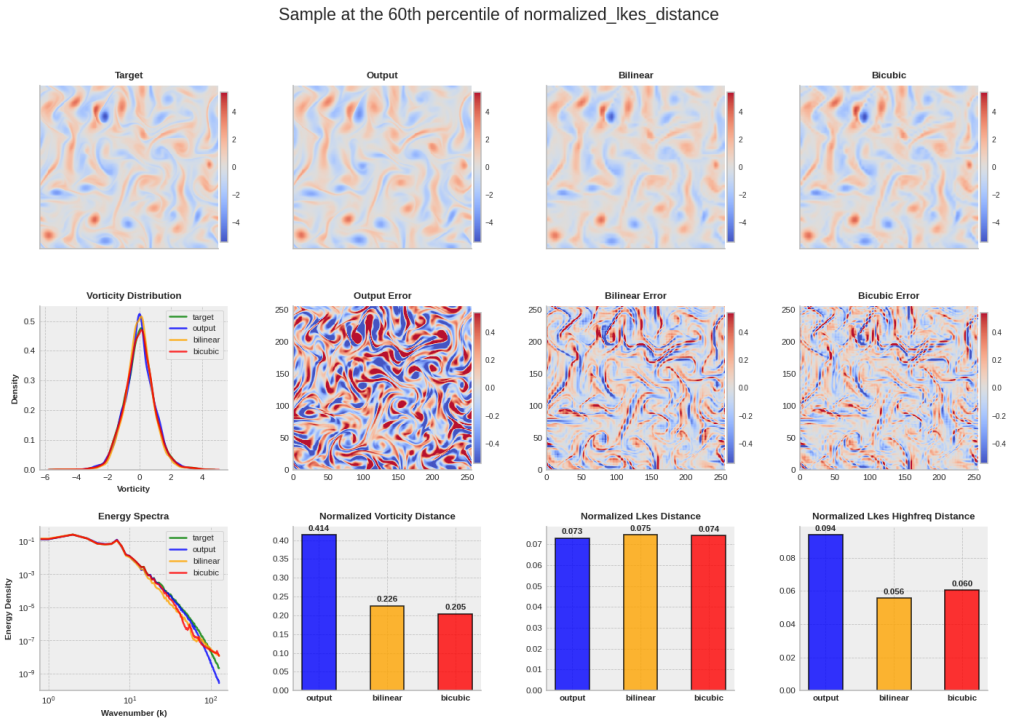


图 7. 谱误差第 60 百分位数处的变分自编码器重建结果（模型规模 16M，深度 4）。

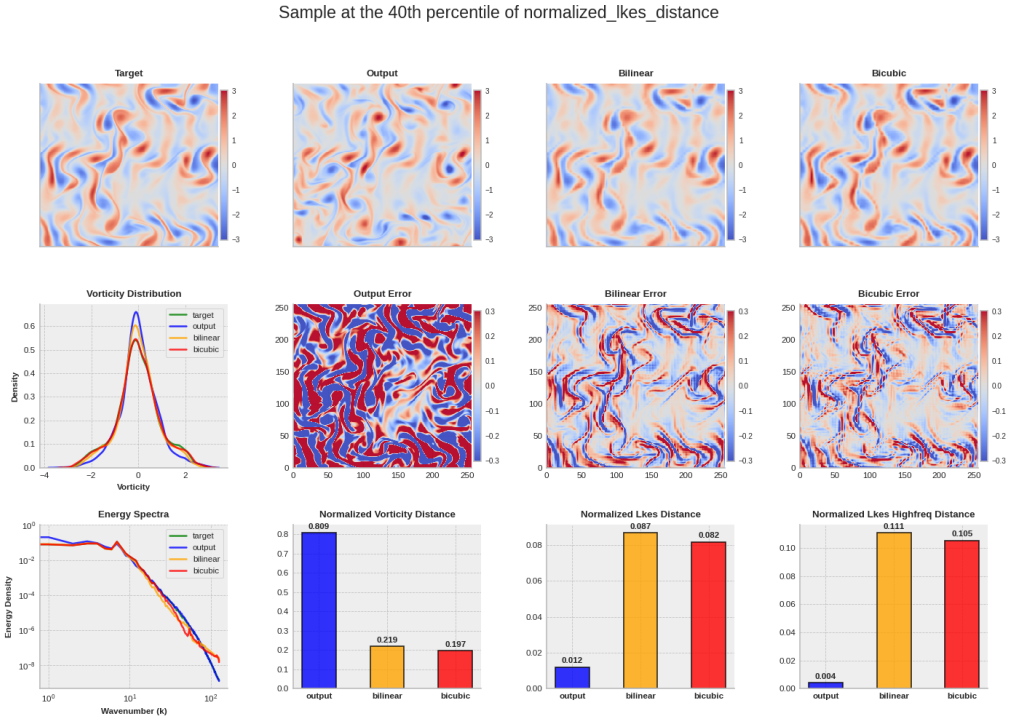


图 8. 谱误差第 40 百分位数处的 DiffCoder 重建（模型规模 16M，深度 4）。

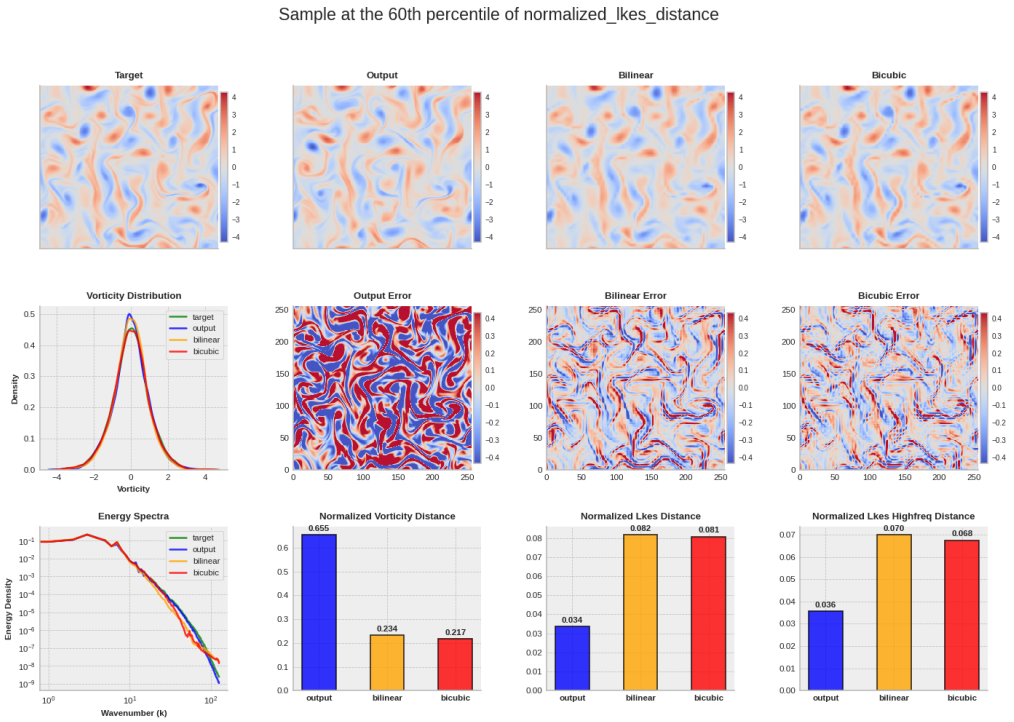


图 9. 谱误差第 60 百分位数处的 DiffCoder 重建（模型规模 16M，深度 4）。

[1]K. Fukami 等，《利用神经网络进行模型降阶：在层流和湍流中的应用》，SN 应用科学 **3**，655（2021）。[2]A. Solera-Rico 等，《用于流体流动降阶建模的β变分自编码器和 Transformer》，《自然-通讯》**15**，45578（2024）。[3]A. Glaws、P. Papas、M. Militzer 和 A. J. Christlieb，《利用深度学习对大型湍流模拟进行原位数据压缩》，《物理评论：流体》**5**，114602（2020）。[4]A. Racca 等，《利用卷积自编码器回声状态网络预测湍流动力学》，《流体力学杂志》（2023），10.1017/jfm.2023.xxx。[5]e. a. Anonymous，《用于多样化时空湍流生成的贝叶斯条件扩散模型》，arXiv 预印本 arXiv:2311.07896（2023），预印本。[6]A. Saydemir、M. Lienen 和 S. Günnemann，《展开时间：用于四维湍流流动的生成建模》，arXiv 预印本 arXiv:2406.11390（2024）。[7]D. Shu、Z. Li 和 A. B. 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）。[8]Y. Jadhav、J. Berthel、C. Hu、R. Panat、J. Beuth 和 A. B. Farimani，《Stressd：使用去噪扩散模型进行二维应力估计》，《计算机方法在应用力学与工程中的应用》**416**，116343（2023）。[9]F. Ogoke、Q. Liu、O. Ajenifujah、A. Myers、G. Quirarte、J. Malen、J. Beuth 和 A. B. Farimani，《利用生成式深度扩散实现增材制造中低成本、高保真的熔池模型》，《材料与设计》**245**，113181 2024( )。[10]Y. Jadhav、J. Berthel、C. Hu、R. Panat、J. Beuth 和 A. Barati Farimani，《利用三维扩散生成晶格单元以实现逆向设计：Glu3d》，《先进功能材料》**34**，2404165（2024）。[11]D. Shu 和 A. B. Farimani，《利用扩散概率模型进行零样本不确定性量化》，arXiv 预印本 arXiv:2408.04718（2024）。[12]A. Zhou、Z. Li、M. Schneier、J. R. Buchanan Jr 和 A. B. Farimani，《Text2pde：用于便捷物理模拟的潜在扩散模型，2025》，网址：https://arxiv. org/abs/2410.01153（2025）。[13]M. Momenifar 等，《用于湍流数据压缩的物理信息向量量化自编码器》，arXiv 预印本 arXiv:2201.03617（2022），预印本。[14]F. Ogoke、S. K. Suresh、J. Adamczyk、D. Bolintineanu、A. Garland、M. Hei-den 和 A. B. Farimani，《基于深度学习的光学图像超分辨率：利用生成式扩散模型对 LPBF 进行逐层原位监测》，《增材制造》，104790（2025）。[15]L. Sirovich，《湍流与相干结构的动力学。I. 相干结构》，《应用数学季刊》**45**，561–571（1987）。[16]T. Bui-Thanh、M. Damodaran 和 K. Willcox，《利用本征正交分解进行气动数据重建和逆向设计》，《AIAA杂志》**42**，1505–1516（2004）。[17]D. Venturi 和 G. E. Karniadakis，《圆柱绕流的缺失数据与重建方法》，《流体力学杂志》**519**，315–336（2004）。[18]J.-C. Loiseau、B. R. Noack 和 S. L. Brunton，《稀疏降阶建模：从基于传感器的动力学到全状态估计》，《流体力学杂志》**844**，459–490（2018）。[19]N. B. Erichson、L. Mathelin、Z. Yao、S. L. Brunton、M. W. Mahoney 和 J. N. Kutz，《利用浅层神经网络和有限传感器重建流体流动》，《英国皇家学会学报A辑》**476**，20200097（2020）。[20]K. Fukami、R. Maulik、N. Ramachandra、K. Fukagata 和 K. Taira，《利用 Voronoi 镶嵌辅助深度学习从稀疏传感器重建全局场》，《自然-机器智能》**3**，945–951（2021）。[21]A. Güemes、C. Sanmiguel Vila 和 S. Discetti，“随机播种场的超分辨率生成对抗网络”，《自然-机器智能》**4**，1165–1173（2022）。[22]A. Subramaniam、M. L. Wong、R. D. Borker、S. Nimmagadda 和 S. K. Lele，“ 利用物理信息生成对抗网络增强湍流 ”，（ 2020 ），arXiv:2003.01907 [physics.comp-ph]。[23]H. Kim、J. Kim、S. Won 和 C. Lee，“ 用于湍流超分辨率重建的无监督深度学习 ”，《流体力学杂志》**910**（2021）。[24]C. Dong、C. C. Loy、K. He 和 X. Tang，“使用深度卷积网络进行图像超分辨率”，IEEE模式分析与机器智能汇刊**38**，295–307（2015）。[25]C. Saharia、J. Ho、W. Chan、T. Salimans、D. J. Fleet 和 M. Norouzi，“通过迭代细化实现图像超分辨率”，IEEE模式分析与机器智能汇刊**45**，4713–4726（2022）。[26]P. Dhariwal 和 A. Nichol，“扩散模型在图像合成方面胜过 GAN”，神经信息处理系统进展**34**，8780–8794（2021）。[27]J. Ho、A. Jain 和 P. Abbeel，“去噪扩散概率模型”，神经信息处理系统进展**33**，6840–6851（2020）。[28]J. Song、C. Meng 和 S. Ermon，“去噪扩散隐式模型”，arXiv预印本 arXiv:2010.02502（2020）。[29]A. Q. Nichol 和 P. Dhariwal，“改进的去噪扩散概率模型”，载于国际机器学习会议 （PMLR，2021），第8162–8171页。[30]T. Hang、S. Gu、C. Li、J. Bao、D. Chen、H. Hu、X. Geng 和 B. Guo，“通过最小信噪比加权策略实现高效扩散训练”，载于*IEEE/CVF*国际计算机视觉会议论文集（2023），第7441–7451页。[31]Q. Liu 和 N. Thuerey，“利用去噪扩散概率模型进行翼型流动模拟的不确定性感知代理模型”，《AIAA杂志》，1–22（2024）。[32]G. Yang 和 S. Sommer，“用于流体场预测的去噪扩散模型”，arXiv预印本 arXiv:2301.11661（2023）。[33]T. Li、A. S. Lanotte、M. Buzzicotti、F. Bonaccorso 和 L. Biferale，“利用生成式扩散模型对湍流旋转流进行多尺度重建”，《大气》**15**，60（2023）。[34]R. Abaidi 和 N. A. Adams，“探索去噪扩散模型在可压缩流体场预测中的应用”，《计算机与流体》，106665（2025）。[35]J.-H. Bastek、W. Sun 和 D. M. Kochmann，“物理信息扩散模型”，arXiv预印本 arXiv:2403.14404（2024）。[36]S. Shan、P. Wang、S. Chen、J. Liu、C. Xu 和 S. Cai，“PIRD：用于流场重建的物理信息残差扩散”，arXiv预印本 arXiv:2404.08412（2024）。[37]Y. Lu 和 W. Xu，“利用物理引导扩散模型对偏微分方程求解器进行生成式降尺度”，arXiv预印本 arXiv:2404.05009（2024）。[38]T. Salimans 和 J. Ho，“用于扩散模型快速采样的渐进式蒸馏”，arXiv预印本 arXiv:2202.00512（2022）。[39]O. Ronneberger、P. Fischer 和 T. Brox，“U-Net：用于生物医学图像分割的卷积网络”，载于国际医学图像计算与计算机辅助干预会议 （施普林格，2015），第234–241页。[40]T. Chen，“论噪声调度对扩散模型的重要性”，arXiv预印本 arXiv:2301.10972（2023）。[41]Y. Song、J. Sohl-Dickstein、D. P. Kingma、A. Kumar、S. Ermon 和 B. 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）。[42]Z. Li、D. Shu 和 A. Barati Farimani，“用于偏微分方程代理建模的可扩展变换器”，神经信息处理系统进展**36**，28010 28039 2023–（）。