arXiv:2603.03196v1 [math.NA] 3 Mar 2026

无限维生成式感知

PAOLO ANGELLA、VITO PAOLO PASTORE 和 MATTEO SANTACESARIA

Abstract. 深度生成模型已成为逆问题先验建模的标准方法，超越了传统的基于稀疏性的方法。然而，现有理论保证大多局限于有限维向量空间；当物理信号被建模为希尔伯特空间中的函数时，便形成了理论空缺。本文提出了一个适用于希尔伯特空间中生成式压缩感知的严谨框架。我们将局部相干性的概念推广到无限维设定，从而推导出最优且与分辨率无关的采样分布。借助对受限等距性质的推广，我们证明，当测量值的数量与先验的内在维数成正比（至多相差对数因子）时，可实现稳定恢复，且这一数量与环境维数无关。最后，针对达西流方程开展的数值实验验证了我们的理论结果，并表明在严重欠采样区间，使用低分辨率生成器可充当隐式正则化器，从而提高重建稳定性。

1. Introduction

从有限测量值中恢复高维信号，是信号处理、医学成像和科学计算领域的一项根本性挑战。经典压缩感知理论 [22, 28] 彻底改变了这一领域：它表明，若信号在某个基下具有稀疏表示，则只需数量与其稀疏性成正比、而非与其环境维数成正比的线性测量值（相差一个对数因子），即可恢复信号。

随着机器学习的兴起，深度生成模型的出现对稀疏性假设的僵化提出了挑战 [35, 32, 44]。大量研究表明，若生成先验将未知信号限制在低维生成器的值域内 $G\colon \mathbb{C}^k \to \mathbb{C}^N$，则其性能可以超过基于稀疏性的经典方法，而且通常所需的测量值显著更少 [18, 41, 15, 16, 6]。这与流形假设相吻合 [30]：信号集中于具有形变稳定性和各向异性正则性的低维几何结构上 [40, 33]。

在有限维情形中，这些方法的理论保证依赖于受限等距性质 [23, 14], ，这些性质确保测量算子以某种方式保持生成器的值域的几何结构。问题在于，许多具有实际意义的逆问题，例如源于偏微分方程、磁共振成像或计算机断层扫描的问题，本质上是无限维的 [29]。底层信号是希尔伯特空间中的函数，而非文献通常所假设的固定长度向量。

将这些问题先验地离散化会带来不同的理论与实际问题，例如“逆犯罪”，或对网格分辨率的依赖，而这种依赖会掩盖信号的真实复杂度 [12]。这使得将深度先验方法推广到压缩感知依赖于恰当的无限维表述。这正是本工作的目标。

我们将该逆问题建模为：根据酉算子 $\mathcal{F}$（例如傅里叶变换）的带噪加权求值，恢复未知函数 $x_0 \in \ell^2(\mathbb{N})$ 。我们假设 x_0 属于广义生成网络 G 的值域： $\mathbb{C}^k \to \ell^2(\mathbb{N})$。广义生成网络是一类深度 ReLU 网络

Key words and phrases. 逆问题，压缩感知，生成模型，深度学习，稀疏性。

在有限维空间之间进行映射，并具有一个以 $\ell^2(\mathbb{N})$ 为陪域的最终线性层。本文以 [16] 的有限维结果为出发点，将其框架推广到无限维设定。

本文的贡献如下：

- 我们引入了无限维局部相干性的概念，用于量化生成模型的值域与测量基之间的对齐程度。我们推导出一种基于相干性的最优采样分布 p^*，以最小化样本复杂度。
- 我们为无限维算子建立了 Gen-RIP。我们证明，为生成器的值域实现稳定嵌入所需的样本数量 m随内在维数 k 呈对数级增长，随局部相干性 $\mu_{\mathcal{U}}$呈多项式级增长，且与环境维数无关。
- 我们将分析扩展到考虑支撑集并非覆盖整个定义域的概率分布，并引入平衡性质，以控制这些实际情形下的近似误差。
- 通过对达西流进行数值实验，我们证明，基于相干性的采样显著优于均匀采样。此外，我们还发现一种违反直觉的现象：分辨率较低的生成器在欠采样区域中能获得更高的重建精度，这表明有限的分辨率可以发挥有益的隐式正则化器作用。

最后，我们指出，尽管我们的框架是在无限维环境空间中表述的，但此处考虑的生成先验本质上是有限维的，它将低维潜在编码映射到 $\ell^2(\mathbb{N})$的一个有限维子空间。

1.1. 相关工作。 *无限维压缩感知。* 从有限维到无限维压缩感知的转变，是由避免所谓逆犯罪以及经典表述中出现的离散化误差这一需要所推动的。广义采样理论表明，只要妥善处理离散化，便可以使用真正的无限维测量模型实现函数的稳定恢复 [7] （参见 [12] 关于非线性逆问题的推广）。在此基础上，发展了无限维压缩感知理论 [8, 9, 1, 5, 4]。此后，该框架已被推广到多种基和采样策略。特别地，已针对勒让德多项式基中的稀疏展开 [43], 以及呈现局部或结构化行为的采样方案导出了严格的恢复保证，并对后者建立了预言机型局部恢复保证 [3]。最近，已发展出面向逆问题的无限维压缩感知统一理论 [10, 11]。该框架能够在不对前向算子进行离散化的情况下，从局部测量（例如由拉东变换产生的测量）中恢复信号，尽管相关保证主要关注基于小波的稀疏性，而非学习得到的生成先验。

无限维生成模型。 将深度生成模型扩展至函数空间，已成为一个活跃的研究领域，其动机在于需要获得与分辨率无关的表示。早期方法主要着眼于扩展离散架构，而近期研究则直接在希尔伯特空间中构建生成建模。典型例子包括定义在函数空间上的无限维变分自编码器和基于分数的扩散模型，它们旨在学习无限维域上的概率测度 [31, 17]。与我们的设定尤为相关的是连续生成神经网络 [13], 和函数自编码器 [20], 这一类方法，它们专为学习有限维潜在空间与函数空间之间与分辨率无关的映射而设计（例如，$L^2(\mathbb{R}^d)$）。因此，它们自然适合作为我们理论框架中的生成器 G。

隐式生成先验。 近年来，基于分数的生成模型和扩散概率模型已成为解决逆问题的主流范式 [45,25]。不同于显式低维生成器 $G: \mathbb{C}^k \to \mathcal{H}$ 我们框架中采用的方法，这些方法通常利用由分数函数给出的隐式先验，并通过朗之万动力学等随机采样过程实现高质量重建（参见 [27] 近期综述）。尽管这些方法往往能达到最先进的感知质量，我们的工作则侧重于显式低维流形的确定性几何所带来的严格恢复保证。

算子学习与科学机器学习。 我们的工作也与快速发展的科学机器学习领域相关，尤其是算子学习 [19], ，即训练神经网络以逼近无限维函数空间之间的映射。傅里叶神经算子（FNO）[38]和 DeepONet [39]等架构在求解参数化偏微分方程方面展现出卓越的经验性能。误差分析和逼近理论方面的结果见 [37, 36] 。这些方法与我们的目标相同，均追求分辨率无关性，但通常依赖大型监督数据集来直接学习正向或逆算子。相比之下，我们的理论提供了一种互补视角，有望为训练神经算子或在数据稀缺情形下恢复输入提供最优数据采集策略。

1.2. **论文结构**。 本文其余部分安排如下。第 2节阐述无限维逆问题，并定义作为先验使用的广义生成网络类别。我们引入无限维局部相干性的概念，以推导最优采样策略；并建立平衡性质来处理非容许分布，同时给出主要的恢复保证。第3 节通过将函数自编码器 [20] 应用于达西流问题，对该框架进行数值验证。第4 节给出主要结果的证明，包括广义受限等距性质（Gen-RIP）的推导，以及恢复信号的误差界。最后，第 5 节总结本文的贡献，并讨论其局限性和有待探索的方向。

2. Setup and Main Results

本节将建立数学框架，利用生成先验求解无限维空间中的逆问题。假设有一个由无限维酉算子 F: $\ell^2(\mathbb{N}) \longmapsto \ell^2(\mathbb{N})$ 建模的物理过程。我们的目标是根据含噪测量向量 b 恢复未知信号 $x_0 \in \ell^2(\mathbb{N})$ ，使得:

$$b = SDFx_0 + \eta,$$

其中， S 是子采样算子， D 是加权算子（两者将在下文定义），而 η 表示测量噪声。

经典压缩感知依赖于固定基中的稀疏性。这里，我们不再依赖稀疏性，而是假设信号 x_0 位于深度生成模型值域附近。具体而言，我们将解限制在生成器的值域内:

$$G\colon \mathbb{C}^k \longmapsto \ell^2(\mathbb{N}).$$

为形式化这一先验，我们引入适用于我们设定的广义生成网络定义。

定义 2.1 ((k,d)-广义生成网络)。 固定整数 $2 \leqslant k: = k_0 \leqslant \ldots \leqslant k_d$。固定权重矩阵 $W^{(i)} \in \mathbb{C}^{k_i \times k_{i-1}}$，以及一个线性映射 $W: \mathbb{C}^{k_d} \longmapsto \ell^2(\mathbb{N})$。我们将函数 $G: \mathbb{C}^k \longmapsto \ell^2(\mathbb{N})$ 定义为一个 (k,d)-广义生成网络，其定义如下:

$$G(z) \coloneqq W\sigma\left(W^{(d)}\sigma\left(\cdots W^{(2)}\sigma\left(W^{(1)}z\right)\right)\right),$$

其中， σ 是 *ReLU* 激活函数，定义如下：

$$\sigma(x) = \max(0, x).$$

这是对经典生成网络的自然扩展，其陪域是有限维的；唯一的区别是，我们将一个线性映射 W 作为最后一层，将其映射到无限维空间。关键在于，我们将中间线性映射 $W^{(i)}$ 限制为具有有限维陪域。若允许它们映射到无限维空间，生成器的值域就会变成无限维，从而使后文给出的理论论证失效。在无限维情形下，无法实现均匀采样。因此，我们必须采用可变密度采样策略。我们引入定义在 $\mathbb{N}$上的概率分布 p，它告诉我们如何选择索引。为分析这一过程，我们定义相应的采样、加权和投影算子。

定义 2.2。 设 p 是定义在 $\mathbb{N}$上的概率分布。设 $\{e_j\}_{j\in\mathbb{N}}$ 表示 $\ell^2(\mathbb{N})$的典范基。我们将采样算子 $S: \ell^2(\mathbb{N}) \to \mathbb{C}^m$ 定义为由 m个独立同分布随机向量 s_1、.....、 s_m （取值于 $\ell^2(\mathbb{N})$）构成的线性映射，使得

$$\mathbb{P}(s_i = e_j) = p_j, \quad \forall j \in \mathbb{N}, \forall i \in [m],$$

其中我们记为 $[m] = \{1, \ldots, m\}$。 S的作用按分量通过内积定义：

$$(Sx)_i = \langle x, s_i \rangle, \quad \forall i \in [m].$$

为确保我们的估计量保持无偏，并对采样混合成分的能量进行归一化，我们使用对角加权算子 D。

定义 2.3。 给定 p 的一个概率分布，我们定义权重算子 D： $\ell^2(\mathbb{N}) \longmapsto \ell^2(\mathbb{N})$ ，使得

$$(Dx)_i = \begin{cases} \frac{x_i}{\sqrt{p_i}}, & for\ p_i > 0 \\ 0, & for\ p_i = 0 \end{cases}$$

ˆ 此外，我们通过算子 I确定采样分布的支撑集。该算子本质上是对采样概率非零的索引进行投影。

定义 2.4。 给定 p的一个概率分布，我们定义算子 ˆI： $\ell^2(\mathbb{N}) \longmapsto \ell^2(\mathbb{N})$

$$\hat{I} := \operatorname{diag}(\mathbb{1}_{\operatorname{supp} p}).$$

有了这些算子，我们的恢复策略可表述为一个优化问题。我们寻求一个潜在向量 ˆz ，使其最小化测量域中的数据一致性误差：

$$\hat{z} \in \arg\min_{z\in\mathbb{C}^k} \|b - SDFG(z)\|.$$

We 概述： D 这里仅需一个权重算子来平衡 p ，因为一般情况下它会 1
lea导致非均匀采样。恢复信号为：

$$\hat{x} = G(\hat{z}).$$

我们的最终分析目标是推导重建误差 $\|\hat{}x - x_0\|$的上界。总体设定与有限维情形相似，关键区别在于，我们能够处理数据的无限维特性，以及支撑集并非整个定义域的概率分布。尤其是，处理那些无法对某些相关坐标进行采样的概率分布，在有限维情形中也很重要，但以往文献中的研究尚未涉及这一问题。

2.1. **局部相干性与采样策略**。为确保更好的恢复率，采样分布 p 必须与算子 F 相对于信号模型的“几何结构”相匹配。若 F 将信号模型映射到 $\ell^2(\mathbb{N})$中的特定坐标，我们就必须以更高的概率对这些坐标进行采样。这种对齐关系由 局部相干性这一概念刻画，该概念源自经典压缩感知 [21]，并在 [15] 中扩展至生成模型。

设 $\mathcal{U} \subseteq \ell^2(\mathbb{N})$ 是 N 凸锥 $\bigcup_{i=1}^{N} \mathcal{C}_i$ 的并。我们定义 $\Delta(\mathcal{U}) := \bigcup_{i=1}^{N} \mathrm{span}(\mathcal{C}_i)$。直观地说，我们应将 $\mathcal{U}$ 视为生成器的值域

定义 2.5。令 $\{e_i\}_{i\in\mathbb{N}}$ 表示 $\ell^2(\mathbb{N})$的典范基。给定作用于 $\ell^2(\mathbb{N})$上的酉算子 F，我们将其第 i行定义为向量 $F_i := F^* e_i$。给定包含于有限维子空间中的凸锥有限并$\mathcal{U} \subseteq \ell^2(\mathbb{N})$，我们将 $\mathcal{U}$相对于 F 的局部相干性 $(\alpha_i)_{i\in\mathbb{N}}$ 定义为

$$\alpha_i = \sup_{x\in\Delta(\mathcal{U}),\|x\|=1} |\langle x, F_i\rangle| = \sup_{x\in\Delta(\mathcal{U}),\|x\|=1} |(Fx)_i|. \tag{1}$$

我们可以通过以下引理证明，在有限维子空间的框架下，这一定义是良定义的。

引理 2.1。设 $F\colon \ell^2(\mathbb{N}) \longmapsto \ell^2(\mathbb{N})$ 是一个酉算子，并设 $\mathcal{U} \subseteq \ell^2(\mathbb{N})$ 是一个包含于维数为 r 的子空间中的凸锥有限并，$\alpha = (\alpha_i)_{i\in\mathbb{N}}$ 是 $\mathcal{U}$ 按 (1)中的定义所得的局部相干性。则

$$\alpha \in \ell^2(\mathbb{N}) \quad and \quad \|\alpha\|^2 \leqslant r^2.$$

有限维情形下这显而易见，但在无限维情形下则不然。局部相干性α_i 量化了第 i 个测量基向量与模型集内任意归一化差向量之间的最大相关性。如果某个采样分布能够充分覆盖相干性较高的指标，就称其为“良好”。

定义 2.6. 给定 F 上的酉算子、$\ell^2(\mathbb{N})$ 中包含于某个有限维子空间的凸锥的有限并 $\mathcal{U} \subseteq \ell^2(\mathbb{N})$，以及 p 上的概率分布 $\mathbb{N}$。令 $(\alpha_i)_{i\in\mathbb{N}}$ 为 F 关于 $\mathcal{U}$ 的局部相干性。称 $\mu_{\mathcal{U}}(F、p)$ 为满足下式的最小常数：

$$\alpha_j \leqslant \mu_{\mathcal{U}}(F,p)\sqrt{p_j} \quad \forall j : p_j > 0$$

成立。若 $\mu_{\mathcal{U}}(F、p)$ 满足

$$\alpha_j \leqslant \mu_{\mathcal{U}}(F,p)\sqrt{p_j} \quad \forall j \in \mathbb{N},$$

我们称 p 为 可容许 的。

可容许分布可确保不会忽略相干性非零的指标。下面的引理给出了使所需样本复杂度最小化的最优采样分布。

引理 2.2. 给定 F 上的酉算子，以及包含于某个有限维子空间的凸锥的有限并 $\mathcal{U} \subseteq \ell^2(\mathbb{N})$，则在所有可容许的 p 中，$\mu_{\mathcal{U}}(F、p)$ 的最小值由 p^* 取得，其中

$$p_j^* = \frac{\alpha_j^2}{\|\alpha\|^2}, \tag{2}$$

在这种情况下，$\mu_{\mathcal{U}}(F、p^*)^2 = \|\alpha\|^2$。

我们注意到，最优分布 p^* 对应于按照归一化 Christoffel 函数（通常称为统计杠杆分数）进行采样。这一策略是最优最小二乘逼近的基础，相关分析见 [26]，近期综述见 [2]。

我们的分析通常需要假设 $\mu_{\mathcal{U}}(F,\ p) \geqslant 1$。如下文引理4.4所示，这一条件并不苛刻；它实际上排除了病态情形（例如相干性消失），并为数值验证提供了实用的判据。此外，引理4.3 证明，只要 p 是可容许的，这一条件便自动成立。

2.2. 恢复保证。 我们陈述主要的恢复结果。该定理为恢复信号 $\hat{x}$ 提供了一个误差界。它区分了一般情形（采样可能忽略某些活跃指标）和理想情形（概率分布可容许）。

我们采用集合运算的标准闵可夫斯基记号；具体而言，对于给定集合 A，我们将其差集记为 $A-A:=\{a-b \mid a,\ b\in A\}$。

定理 2.3。 设 F 是 $\ell^2(\mathbb{N})$ 上的酉算子，p 是 $\mathbb{N}$ 上的概率分布，其关联算子为 S、D 和 $\hat{I}$。设 G 是一个 (k,d)-广义生成网络，且$\mathcal{T}=R(G)-R(G)$ 是其差集，满足 $\mu_{\mathcal{T}}(F,p)\geqslant 1$。若测量值的数量$m\in\mathbb{N}$ 满足

$$m \geqslant 16\mu_{\mathcal{T}}(F,p)^2\left[2kd\log\left(\frac{2ek_d}{k}\right)+\log\left(\frac{4k}{\varepsilon}\right)\right],$$

那么，以至少 $1-\varepsilon$ 的概率，对于任意 $x_0\in\ell^2(\mathbb{N})$，$b:=SDFx_0+\eta$，$x^{\perp}:=x_0-\Pi_{R(G)}(x_0)$，以及满足 $\|SDF\hat{x}-b\|\leqslant\min_{x\in R(G)}\|SDFx-b\|+\hat{\varepsilon}$ 的估计量 $\hat{x}\in R(G)$，以下结论成立：

$$\|\hat{x}-x_0\|\leqslant\|x^{\perp}\|+\sqrt{\frac{2}{m}}\Big[2\|SDFx^{\perp}\|+2\|\eta\|+\hat{\varepsilon}\Big]+\left\|\hat{I}^{\perp}F(\hat{x}-\Pi_{R(G)}(x_0))\right\|. \tag{3}$$

此外，若 p 是可容许 分布，则最终的尾项消失，界简化为：

$$\|\hat{x}-x_0\|\leqslant\|x^{\perp}\|+\sqrt{\frac{2}{m}}\Big[2\|SDFx^{\perp}\|+2\|\eta\|+\hat{\varepsilon}\Big]. \tag{4}$$

我们注意到，采样概率分布的最佳选择是引理 2.2 中给出的分布，因此，若 α 是 F 相对于 $\mathcal{T}$ 的局部相干性，则界变为

$$m \geqslant 16\|\alpha\|^2\left[2kd\log\left(\frac{2ek_d}{k}\right)+\log\left(\frac{4k}{\epsilon}\right)\right].$$

当 p 可容许时，所得界与有限维情形下已知的界一致 [15, 16]。相比之下，不可容许的 p 情形给出了全新的界，构成了有限维和无限维情形下的新颖贡献。

注。 由于我们的测量模型 $b:=SDFx_0+\eta$ 中的加权算子 D，最终误差界中的噪声贡献按 $\frac{1}{\sqrt{m}}\|\eta\|$ 缩放。这带来了良好的渐近行为：对于有限能量噪声，误差趋于零；对于按 $\sqrt{m}$ 缩放的噪声模型，误差则始终严格有界。

2.2.1. 平衡性质。 在许多无限维情形中，如果相干性具有无限支撑集，或者我们受限于只能对指标的某个有限子集进行采样，那么可容许的分布（要求 $p_j>0$ 当 $\alpha_j>0$时成立）可能并不实用。在这种情况下，定理$\|\hat{I}^{\perp}F(\hat{x}-\Pi_{R(G)}(x_0))\|$ 中的项 2.3不会消失。为控制这一误差，我们引入平衡性质这一概念，它源自无限维压缩感知 [8]。

定义 2.7。 给定 F 中的酉算子 $\ell^2(\mathbb{N})$，以及 $\mathbb{N}$上的概率分布 p，其关联算子为$\hat{I}$。 G 一个 (k,d)-广义生成模型。如果对 $\theta\in(0,1)$满足平衡性质，则称

$$\|\hat{I}^{\perp}F\Pi_{R(G)}\|\leqslant\theta.$$

这一性质实质上要求模型值域中信号的能量集中在采样分布 p的支撑集上。如果这一条件成立，我们仍然可以推导出一个强恢复界，不过放大因子为 $(1-\theta)^{-1}$。

注。 有人可能会质疑平衡性质是否是一个现实的假设。我们注意到，对于任意给定的容差 $\theta > 0$、任意酉算子 F 以及任意有限维生成模型 G，都存在一个分布 p，使平衡性质成立。为此，考虑在前 ℓ 个指标 $p_\ell(n) = \frac{1}{\ell}\mathbb{1}_{[\ell]}$ 上的均匀分布，并令 $\hat{} I_\ell$ 为相应的算子，将向量投影到前 ℓ 个坐标上。其正交补 $\hat{} I_\ell^\perp$则投影到剩余的尾部指标上。

令 $\mathcal{V}$ 表示包含值域 $R(G)$ 的有限维子空间，并令 $R = F(\mathcal{V})$。我们有：

$$\|\hat{I}_\ell^\perp F \Pi_{\mathcal{V}}\| \leqslant \sup_{x\in R, \|x\|=1} \|\hat{I}_\ell^\perp x\|.$$

为了界定这部分尾部能量，令 $(u_1、\ldots、u_r)$ 为 R 的一组标准正交基。任意单位向量 $x \in R$ 都可表示为 $x = \sum_{i=1}^r c_i u_i$，其中 $\sum_{i=1}^r |c_i|^2 = 1$。根据柯西-施瓦茨不等式，第n个分量满足 $|x_n|^2 \leqslant \sum_{i=1}^r |(u_i)_n|^2$。对尾部求和可得：

$$\sup_{x\in R, \|x\|=1} \|\hat{I}_\ell^\perp x\|^2 = \sup_{x\in R, \|x\|=1} \sum_{n>\ell} |x_n|^2 \leqslant \sum_{n>\ell}\sum_{i=1}^r |(u_i)_n|^2.$$

由于每个基向量 u_i 都属于 $\ell^2(\mathbb{N})$，它们尾部能量的有限和必定会在 $\ell \to \infty$ 时严格趋于零。因此，$\| \hat{} I_\ell^\perp F \Pi_{\mathcal{V}}\| \to 0$。于是，通过选择足够大的支撑大小 ℓ，我们可以对任意 θ 满足平衡条件。

推论 2.4. 设 F 是 $\ell^2(\mathbb{N})$ 中的酉算子，p 是定义在 $\mathbb{N}$ 上的概率分布，满足 S，其关联算子为 D。令 G 为一个 (k,d)- 广义生成网络，并令 $\mathcal{T} = R(G) - R(G)$ 满足 $\mu_{\mathcal{T}}(F,p) \geqslant 1$。若平衡性质对 $\theta \in (0,1)$ 成立，且

$$m \geqslant 16\mu_{\mathcal{T}}(F,p)^2 \left[2kd\log\left(\frac{2ek_d}{k}\right) + \log\left(\frac{4k}{\epsilon}\right)\right],$$

则以下结论以至少 $1-\varepsilon$ 的概率成立：
对于任意 $x_0 \in \ell^2(\mathbb{N})$，令 $b := SDFx_0 + \eta$，其中 $\eta \in \mathbb{C}^m$ 且，$x^\perp := x_0 - \Pi_{R(G)}x_0$。令 $\hat{}$ $x \in R(G)$ 满足 $\|SDF\hat{}x - b\| \leqslant \min_{x\in R(G)}\|SDFx - b\| + \hat{}\varepsilon.$。则

$$\|\hat{x} - x_0\| \leqslant \frac{1}{1-\theta}\left(\|x^\perp\| + \sqrt{\frac{2}{m}}\Big[2\|SDFx^\perp\| + 2\|\eta\| + \hat{\varepsilon}\Big]\right)$$

换言之，我们保留了与可容许情形相同形式的稳定性保证，代价是常数更差；同时采用一种采样策略，使我们能够处理不可容许、因而也更贴近现实的概率分布，例如上文注中给出的分布。

3. Numerical Validation

前几章推导出的理论保证是在无限维希尔伯特空间 $\ell^2(\mathbb{N})$中表述的。尽管这种连续体表述给出了严格的界限，但在精确的连续极限下直接进行数值验证，计算上并不可行。因此，任何数值验证都依赖于有限维离散化。为了确保我们的观察并非特定网格选择所造成的假象，而是准确反映底层函数空间理论，我们必须采用一种能够在不同分辨率下保持一致运行的架构。为此，我们采用了函数自编码器（FAE）框架，该框架由 [20]提出。与标准离散自编码器不同，FAE 原生支持无网格学习，可学习函数空间之间的映射，而非固定的

欧几里得向量。采用这种架构，我们可以通过实证展示本文的核心主张：所提出的恢复保证和自适应采样策略对离散化维度具有稳健性。在接下来的章节中，我们将这一框架应用于达西流方程，这是科学机器学习中的一个经典应用。我们在不同的空间分辨率下进行实验，范围从 32×32 到 128×128，以验证重建误差和局部相干性估计所表现出的渐近行为是否符合理论上的无限维极限，且不受训练过程中所用具体网格粒度的影响。

3.1. **实验设置与方法。** 以下各节将详细介绍对我们理论框架的数值验证，具体针对与稳态达西流方程相关的逆问题 [24] [34]。该控制方程定义如下：

$$-\nabla \cdot (a(x)\nabla u(x)) = f(x), \quad x \in \Omega, \tag{5}$$

该方程描述了有界区域 $\Omega \in \mathbb{R}^2$ 上多孔介质中的流体流动，其中 $a(x)$ 表示介质的渗透率系数，$u(x)$ 表示由此产生的压力场（解），$f(x)$ 则为源项。

我们的目标是根据有限个含噪傅里叶测量值恢复压力场 $u \in L^2(\Omega)$。形式上，我们考虑以下逆问题

$$b = SD\mathcal{F}u + \eta,$$

其中，$\mathcal{F}$ 表示高保真分辨率下的离散傅里叶变换（在我们的实验中固定为 128×128），S 是一个采样算子，根据特定的概率分布（均匀分布或基于相干性的分布）选取部分 m频率，D 是用于确保无偏性的相应加权算子，η 则表示测量噪声。由于噪声依赖性并非本文的重点，所有实验均使用无噪声测量值。我们通过将解限制在 FAE 的值域内来处理这一不适定问题：$G : \mathbb{R}^k \to L^2(\Omega)$。该 FAE 经训练后可近似达西流的解流形。重建通过搜索使数据保真损失最小化的最优潜在编码 $\hat{z}$ 来完成：

$$\hat{z} = \underset{z \in \mathbb{R}^k}{\operatorname{argmin}} \|SD\mathcal{F}(U_r G_r(z)) - b\|_2^2,$$

其中，G_r 是以分辨率 r训练的生成器，U_r 是一个固定的上采样算子，可将生成器的输出映射至测量分辨率。这一表述使我们能够将生成分辨率与测量分辨率解耦，从而检验恢复保证的分辨率不变性。

3.1.1. 数据集与架构。 我们使用 [38], 中提供的达西流数据集验证我们的框架，具体实现细节见附录 A.3. [38, 2]。这些解呈现出复杂的多尺度特征，对标准线性重建方法构成了挑战。原始数据集包含 2048 张分辨率为 421 x 421 的图像；在我们的实验中，我们将其调整为 32 x 32、64 x 64 和 128 x 128。

我们采用了 [20]中介绍的 FAE。不过，我们对其架构进行了关键修改：将高斯误差线性单元（GeLU）激活函数替换为修正线性单元（ReLU）。此修改对于确保网络满足本文所建立的理论保证是必要的。除这一替换之外，网络的其余部分均与原始实现保持一致，并满足理论分析所需的假设。

3.1.2. 训练。 为展示 FAE 与分辨率无关的能力，训练在三种不同的空间分辨率下进行： 32×32、64×64和 128×128。FAE 的训练完全采用无监督方式进行。该网络由编码器$\mathcal{E}$ 和解码器 $\mathcal{D}$ 组成，经过优化后可将输入函数压缩为低维潜在表示 $z \in \mathbb{R}^k$ （在我们的实验中为 $k = 64$），随后再将其解压缩回原始函数域。其目标是在学习物理算子的离散化不变表示的同时，最小化重建误差。我们采用了 [20], 中所述的相同训练方案，具体而言，我们训练 FAE 50,000 个训练轮次，学习率设为 0.001，批次大小为 32。

3.1.3. 拟合潜在空间。 自编码器训练完成后，只要能够从潜变量分布中采样 $G(z)$ ，便可以独立使用解码器作为生成器 z 。然而，编码后的训练数据的聚合后验分布通常并不像变分自编码器 [35]那样符合标准高斯分布。因此，简单的高斯采样很可能会产生映射到分布外函数的潜在编码。为了准确捕捉所学习潜在流形的非高斯结构，我们对训练数据的潜在编码拟合一个包含 $K = 10$ 个混合成分的高斯混合模型（GMM）：

$$p(z) = \sum_{i=1}^{K} \pi_i \mathcal{N}(z|\mu_i, \Sigma_i). \tag{6}$$

该高斯混合模型有两个关键作用：它为采样提供有效的生成机制，并在恢复阶段充当结构化先验，使我们能够惩罚偏离训练分布支撑集的解。

3.1.4. 重建。 恢复问题被表述为寻找一个潜在空间元素z ，使其与观测到的欠采样傅里叶测量最匹配。我们实现中的一个关键方面，是如何处理多分辨率生成器与固定高分辨率真实值之间的关系；该真实值的分辨率固定为 128×128。令 u_{target} 表示该分辨率下的真实信号。对于在分辨率 r 下训练的生成器 G_r（其中 $r \in \{32, 64, 128\}$），我们求解以下优化问题：

$$\hat{z} = \arg\min_{z \in \mathbb{R}^k} \| S D \mathcal{F}(U_r^{128} G_r(z)) - S D \mathcal{F}(u_{\text{target}}) \|^2$$

其中 U_r^{128}表示一个双三次上采样算子，用于将生成器的输出映射到目标分辨率； $\mathcal{F}$、 S和 D 等价于第 2节中引入的算子，不同之处在于现在需要在有限维空间中运算，因此 $\mathcal{F}$ 是维度为128×128的快速傅里叶变换， S 是采样算子，根据选定的概率（在我们的实验中，概率由局部相干性诱导或设为均匀分布）对 m 傅里叶系数进行采样；最后， D 以相同方式对样本加权。该公式确保无论模型的内部分辨率如何，所有模型都针对同一个高保真物理真实值进行评估。我们还确保不同输出维度使用相同的采样算子，因为在测量之前，所有内容都会先升尺度，然后再进行傅里叶变换和采样。优化采用 Adam优化器进行，学习率为 0.01，迭代 500 次。我们采用一种实现方式，同时针对多个测试实例优化潜在编码。我们使用空间域上的均方误差（MSE）作为指标来评估性能。

3.2. **结果**。 现在我们展示实验结果。我们使用了一半的数据集，即 1024 张图像，在 3 种不同的条件下训练通用模型，并拟合高斯混合模型以适应潜在空间。另一半图像则用于计算其重建结果，依据的是

欠采样傅里叶变换。以下所有数值结果均为该测试集中所有图像的平均值。

3.2.1. 局部相干性与自适应采样。 要计算最优采样分布（2），需要评估局部相干性（定义2.5）。然而，该量无法显式计算，因此我们通过蒙特卡洛过程对其进行近似。

具体而言，我们从训练好的高斯混合模型中生成 $N = 1024$ 个样本，将其输入生成器，并将输出上采样至目标分辨率。随后，我们计算这些样本对应的傅里叶系数的最大幅值，得到

$$\alpha_j^r = \max_{i=1,\dots,N} \left| \left(\mathcal{F} U_r^{128}(G_r(z_i)) \right)_j \right|,$$

其中 $r \in \{32, 64, 128\}$ 和 j 表示傅里叶频率。然后，通过对 α_j^r 进行归一化得到采样概率 p_j。这种对 r 的依赖源于：在不同分辨率下训练的生成器呈现出不同的局部相干性结构。我们将这种自适应采样策略与均匀采样基线进行比较。

从理论角度看，局部相干性应取决于模型值域中信号的归一化差分，而非单个样本。受此观察启发，我们还测试了一种基于自差分的替代蒙特卡洛估计量。具体而言，给定一批生成的图像 $\{u_i\}_{i=1}^B$，我们计算所有成对差分 $u_i - u_j$，对每个差分进行归一化，逐通道应用二维傅里叶变换，然后在这批归一化差分中取模值的最大值。将所得相干性代理量平方并归一化，以得到采样概率。

虽然该估计量与局部相干性的理论定义更加吻合，但其表现始终不如上文所述的简单最大幅值估计量。经验上，原先基于最大化各个生成样本的傅里叶幅值的策略，得到了重建性能更好的采样分布。

图 1 将这两种方法与朴素的均匀采样方法进行了比较。尽管自差分估计量具有更充分的理论依据，但在我们的设定中，它并未带来更好的经验性能。

3.2.2. 自适应采样。 为评估重建性能，我们以对数尺度改变测量值数量 m，范围为傅里叶系数总数的 1% 至 32%。这使我们能够分析采样密度增加时恢复误差的渐近行为。如图 2 所示，在欠采样区域，自适应采样显著优于均匀采样基线；对于固定的误差阈值，其效率更高。随着采样率提高，两种策略的误差水平渐近趋于相当。此外，对于使用来自所考察的三种离散化的数据训练的模型，这一现象始终一致。图 3 对不同采样率下的重建结果进行了定性比较。

3.2.3. 重建误差的离散化不变性。 本实验的主要目的是表明，重建性能基本不受输出维度的影响。在图4中，我们报告了三种不同模型的恢复误差。每种模型均在所考虑的三个离散化级别之一上进行训练，并在生成三种不同分辨率（即 32×32、 64×64和 128×128）的输出时进行评估。

我们观察到，在可获得足够多测量值的渐近区间，恢复误差主要由模型的训练分辨率决定，基本不受目标输出维度影响。相比之下，在严重欠采样区间，模型分辨率对性能的影响较小，较低的输出维度会带来略好的重建结果。

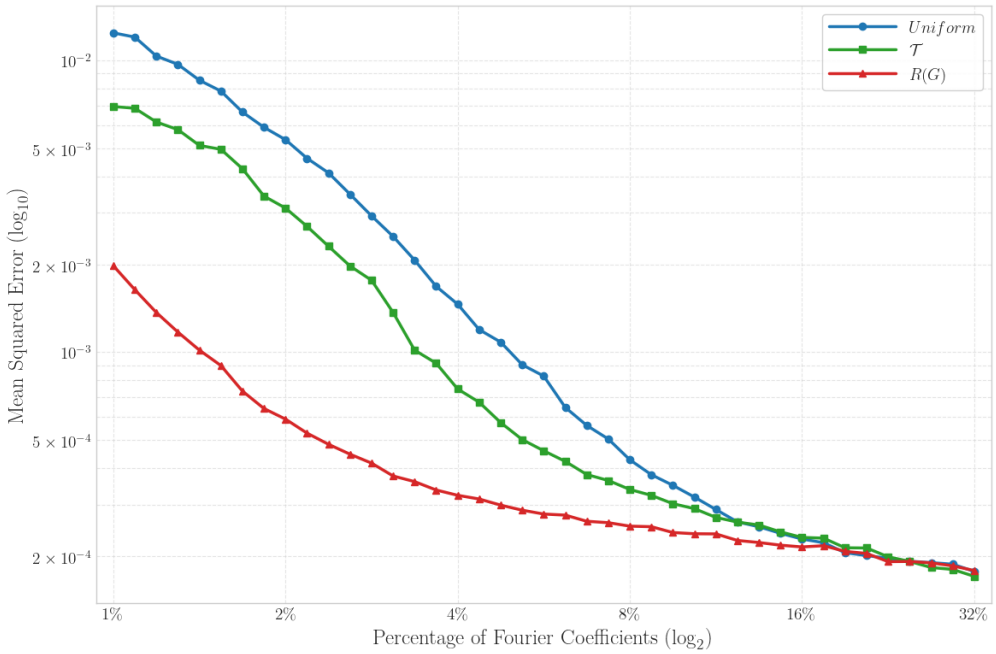


FIGURE 1. 原始的基于最大幅值的相干性估计器与基于自差的估计器之间的比较。尽管后者在理论上更符合局部相干性的定义，但前者的经验性能更优（模型经过训练，输出分辨率为 64 x 64）。

图 5定性展示了不同输出维度下的重建性能。我们观察到，在稀疏采样区域，恢复结果主要由低频特征决定，因此天然更有利于以较低分辨率生成的模型。此外，在样本充足的区域，以较低分辨率训练的模型表现更优。这种现象可以从神经网络的频谱偏差角度来理解[42]，：深度网络优先学习低频函数。在严重欠采样区间，可用测量值主要约束信号的低频分量。高分辨率生成器具有更大的容量，容易“臆造”出位于测量算子零空间中的高频伪影。相比之下，低分辨率生成器相当于一个显式的低通滤波器，可通过抑制这些虚假的高频分量，有效地对逆问题进行正则化，并确保恢复结果符合神经网络固有的频谱偏差。

3.2.4. 高斯混合模型作为正则化项。 在本实验中，我们研究将潜变量先验（6）作为显式正则化项纳入重建优化的效果。我们并非求解纯粹的数据保真问题，而是在目标函数中加入惩罚项，以鼓励潜变量停留在已拟合高斯混合模型下的高概率区域。这促使所得解与训练期间学习到的潜变量分布更加一致。与先前的实验不同，本研究并不关注分辨率不变性。

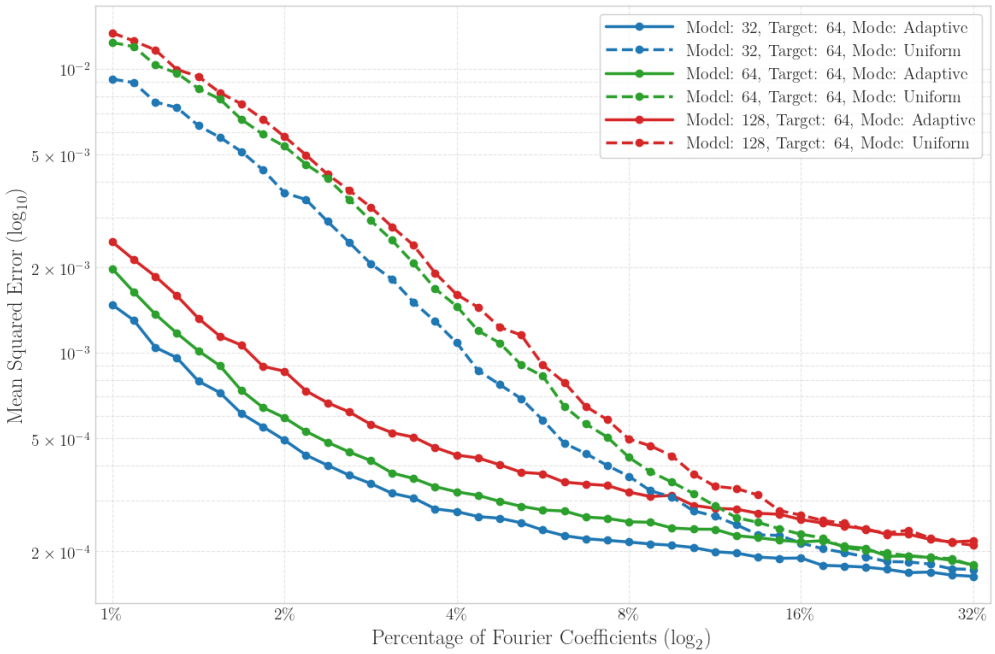


FIGURE 2. 均匀傅里叶采样和自适应傅里叶采样的重建误差与采样率的关系。

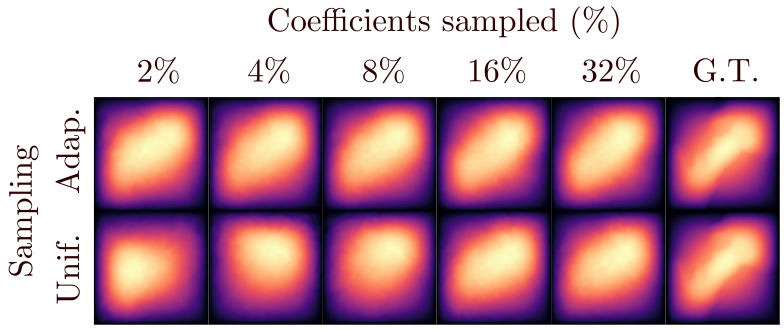


FIGURE 3. 不同采样率下均匀采样与自适应采样的重建示例（模型以 64 x 64 分辨率训练）。

由此得到的优化问题可以写为

$$\hat{z} = \arg\min_{z \in \mathbb{R}^k} \left\| SD\mathcal{F}\left(U_r^{128} G_r(z)\right) - SD\mathcal{F}(u_{\text{target}}) \right\|^2 + \lambda \left(-\log p(z)\right),$$

其中 $p(z)$ 是式（6）中定义的高斯混合模型先验，且 $\lambda \geqslant 0$ 控制正则化的强度。

我们没有预先固定 λ，而是使用数据集的 10% 作为验证集，通过参数调优过程确定最有效的正则化强度。具体而言，对于 λ的一组离散候选值，我们计算在所有采样级别 m上取平均的重建误差。随后，选择使这一汇总误差最小的 λ 作为最优正则化参数；在本例中，我们得到 $\lambda = 10^{-4}$。所得性能

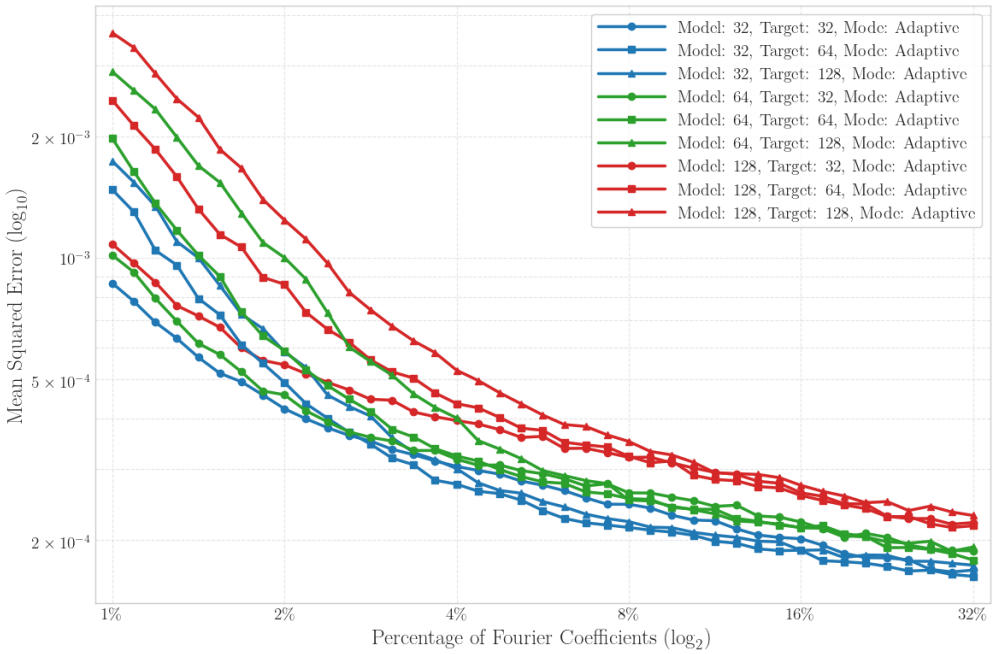


FIGURE 4. 不同训练离散化和输出离散化下的恢复误差。

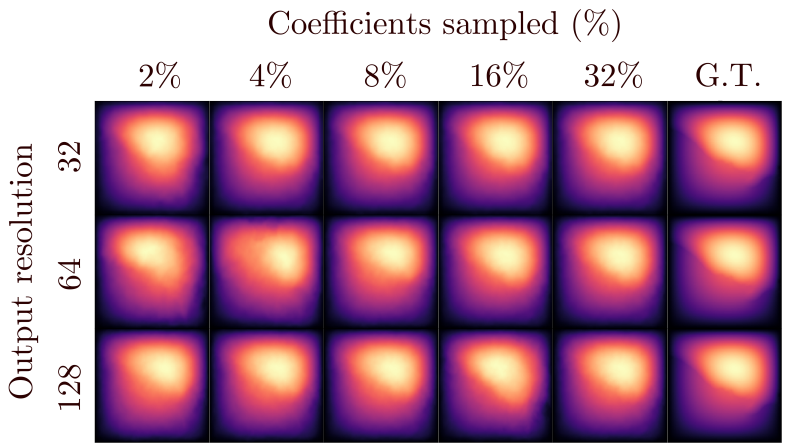


FIGURE 5.不同输出分辨率下的重建示例（模型在 64 x 64 分辨率下训练）

随后，我们将使用这一调优后的正则化项所获得的性能与未正则化情形（$\lambda = 0$）在其余数据点上的性能进行比较。

图 6展示了采用经过最优调节的正则化参数时，重建性能随测量值数量变化的情况。在严重欠采样区间，引入高斯混合模型先验相较于未正则化方法带来了明显的性能提升，因为该先验有效地将优化限制在潜在空间中与合理的生成器输出相对应的区域内。在此语境下，合理或逼真是指在以下分布下具有较高出现可能性的样本

该分布是生成器诱导的数据分布，而非基于任何外部物理真实性概念的判断。

随着测量值数量的增加，数据保真项提供的信息越来越多，正则化的优势也逐渐减弱。在高采样区域，非正则化形式优于正则化形式，这表明仅凭测量值就足以准确识别潜在编码，而不会引入先验带来的偏差。总体而言，这些结果凸显了测量算子与潜变量先验各自发挥的互补作用：当观测信息有限时，后者最为有益。

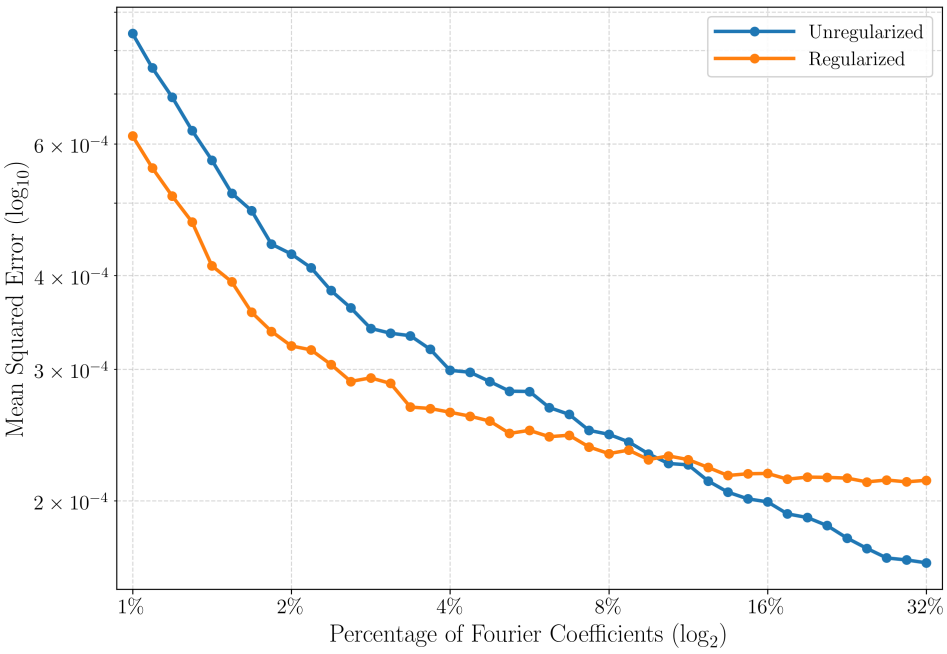


FIGURE 6. 我们提出的方法与经典压缩感知（CS）的比较。

4. PROOFS证明

现在，我们给出第 2 章结果所需的全部结果与证明

4.1. 设定与局部相干性。 首先，我们来刻画一个 (k, d)-广义生成网络的值域，特别是证明其值域包含于一个有限维空间中。

引理 4.1. 设 $2 \leqslant k := k_0 \leqslant \ldots \leqslant k_d$ 与 G 为一个 $(k、d)$-广义生成网络。则 $R(G)$ 是不超过 N 个、至多 k 维的多面锥的并，其中

$$N \leqslant \left(\frac{2ek_d}{k} \right)^{kd}.$$

证明。 本证明沿用 [15, 引理 S22]的计数论证。我们观察到，层宽度的几何平均数以最大宽度 k_d 为上界。因此，倒数第二层的值域 $\tilde{G}(z) = \sigma(W^{(d)} \dots \sigma(W^{(1)} z))$位于 $\mathbb{R}^{k_d}$中，并且是至多 $N \leqslant \binom{2ek_d}{k}^{kd}$ 个多面锥的并。我们的生成器的最后一层是线性算子$W:\ \mathbb{C}^{k_d} \to l^2(\mathbb{N})$。由于 W 是从有限维空间出发的线性映射，故其必然有界。有界线性算子将定义域中的多面锥映射为陪域中的多面锥（具体而言，该像是各生成器的图像的锥包）。因此，锥之并的几何结构在 $l^2(\mathbb{N})$ 中得以保持，而不会增加混合成分的数目 N。□

注。 由于 $R(G)$ 是至多由 N k维锥组成的并集，我们注意到闵可夫斯基和 $\mathcal{T} = R(G) - R(G)$ 是至多由 N^2 $2k$维锥组成的并集。

现在证明局部相干性是良定义的。

引理 *2.1* 的证明。 令 $(\varphi_i)_{i\in[r]}$ 为 $\Delta(\mathcal{U})$ 的一个基，且 $(F_j := F^* e_j)_{j\in\mathbb{N}}$ 其各行 F 构成 $\ell^2(\mathbb{N})$ 的标准正交基，因为 F 是酉矩阵。因此我们可以考虑

$$\alpha_j = \sup_{c\in\mathbb{C}^r:\ \|c\|=1} |\langle F_j, \sum_{i\in[r]} c_i \varphi_i\rangle| = |\langle F_j, \sum_{i\in[r]} \bar{c}_{i,j} \varphi_i\rangle|,$$

其中 $\bar{c}_{i,\ j}$ 是达到上确界的系数（由紧致性可知，该上确界可取到最大值）。于是我们可以得出

$$\begin{aligned}
\sum_{j\in\mathbb{N}} \alpha_j^2 &= \sum_{j\in\mathbb{N}} \langle F_j, \sum_{i\in[r]} \bar{c}_{i,j} \varphi_i\rangle^2 \\
&= \sum_{j\in\mathbb{N}} \left(\sum_{i\in[r]} \bar{c}_{i,j} \langle F_j, \varphi_i\rangle \right)^2 \\
&\leqslant r \sum_{j\in\mathbb{N}} \sum_{i\in[r]} \bar{c}_{i,j}^2 \langle F_j, \varphi_i\rangle^2 \\
&\leqslant r \sum_{j\in\mathbb{N}} \sum_{i\in[r]} \langle F_j, \varphi_i\rangle^2 \\
&= r \sum_{i\in[r]} \sum_{j\in\mathbb{N}} \langle F_j, \varphi_i\rangle^2 \\
&= r \sum_{i\in[r]} \|\varphi_i\|^2 = r^2 < +\infty.
\end{aligned}$$

□

以下两个引理给出了 $\mu_{\mathcal{U}}(F,\ p) \geqslant 1$成立的条件，这些条件对我们的主要结果至关重要。特别是，第二个引理保证只要 p是可容许的，该条件就成立。

引理 4.2。 设 F 是 $\ell^2(\mathbb{N})$上的酉算子， $\mathcal{U} \subseteq \ell^2(\mathbb{N})$ 是包含于有限维子空间中的凸锥的有限并，且 p 是定义在 $\mathbb{N}$ 上的概率分布。若存在单位向量 $\hat{}u \in \Delta(\mathcal{U})$ 使得 $\|IF\hat{}u\| = 1$，则 $\mu_{\mathcal{U}}(F, p) \geqslant 1$。

证明。 考虑 $\beta_j = (F\hat{}u)_j$。对于验证下式的最小 B，我们有 $\mu_{\mathcal{U}}(F,\ p) \geqslant B$

$$\beta_j \leqslant B\sqrt{p_j} \quad \forall j\colon p_j > 0,$$

我们可以对两边取范数平方，得到

$$\|\hat{I}F\hat{u}\| \leqslant B,$$

由于我们要寻找满足该不等式的最小常数，且 $\|\hat{}IF\hat{}u\| = 1$ 因此我们得出 $B = 1$ 和 $\mu_{\mathcal{U}}(F,$ $p) \geqslant B = 1$ □

引理 4.3。 设 F 是作用于 $\ell^2(\mathbb{N})$上的酉算子，且 $\mathcal{U} \subseteq \ell^2(\mathbb{N})$ 是包含于某个有限维子空间中的凸锥的有限并。若 p 是$\mathbb{N}$上的容许概率分布，则

$$\mu_{\mathcal{U}}(F, p) \geqslant 1.$$

证明。根据引理 4.2，我们只需证明存在 $\hat{}u \in \Delta(\mathcal{U})$ ，使得 $\|\hat{}u\| = 1$ ，从而$\|\hat{}IF\hat{}u\| = 1$，但这对所有此类对象都成立。若 $(F\hat{}u)_j \neq 0$，则根据定义有 $\alpha_j > 0$。 p 的可容许性保证对所有此类 j都有 $p_j > 0$ ，这意味着投影 I 在 $F\hat{}u$上起恒等作用。由于 F是酉算子，我们得出 $\|\hat{}IF\hat{}u\| = \|F\hat{}u\| = \|\hat{}u\| = 1$。□

直观地说，为了保证 $\mu_{\mathcal{U}}(F,\ p) \geqslant 1$，只需采样分布将其支撑集置于一组坐标上，使这些坐标的累积平方相干至少为 1。这确保测量算子为生成器的值域提供充分覆盖，正如下述引理所形式化的那样。

引理 4.4。 设 F 酉算子 $\ell^2(\mathbb{N}),\ \mathcal{U} \subseteq \ell^2(\mathbb{N})$ 为包含于$\hat{}$有限维子空间中的凸锥的有限并，且概率分布 p满足 I 在 p的支撑集上的投影。若有

$$\|\hat{I}\alpha\| \geqslant 1, \tag{7}$$

则 $\mu_{\mathcal{U}}(F、\ p) \geqslant 1$。

证明。 根据 $\mu_{\mathcal{U}}(F$的定义，$p)$，若记 Ω 为 p的支撑集，则有

$$\alpha_j^2 \leqslant \mu_{\mathcal{U}}(F, p)^2 p_j \quad \forall j \in \Omega.$$

对 Ω 中的所有指标将该不等式求和:

$$\sum_{j\in\Omega} \alpha_j^2 \leqslant \sum_{j\in\Omega} \mu_{\mathcal{U}}(F, p)^2 p_j = \mu_{\mathcal{U}}(F, p)^2 \sum_{j\in\Omega} p_j.$$

因此,

$$\sum_{j\in\Omega} \alpha_j^2 \leqslant \mu_{\mathcal{U}}(F, p)^2.$$

开平方意味着

$$\mu_{\mathcal{U}}(F, p) \geqslant \sqrt{\sum_{j\in\Omega} \alpha_j^2}.$$

因此，如果满足（7），则可直接推出 $\mu_{\mathcal{U}}(F,\ p) \geqslant 1$。 □

最后，我们证明如何利用相干性刻画用于信号恢复的最佳概率分布。

引理 *2.2* 的证明。 我们只需注意到 $\alpha_j^2 \leqslant \mu_{\mathcal{U}}(F,\ p)^2 p_j$对所有 $j \in \mathbb{N}$均成立，然后计算

$$\sum_{j\in\mathbb{N}} \alpha_j^2 \leqslant \sum_{j\in\mathbb{N}} \mu_{\mathcal{U}}(F, p)^2 p_j$$

这意味着 $\|\alpha\|^2 \leqslant \mu_{\mathcal{U}}(F,\ p)^2$。 □

4.2. **广义受限等距性质**。我们现在给出两个结果，证明我们的感知算子保留了信号集的几何结构。这一性质类似于经典压缩感知中的受限等距性质（RIP）；因此，我们将其称为广义受限等距性质（Gen-RIP）。

定义 4.1 (广义受限等距性质)。设 $\mathcal{K} \subseteq \ell^2(\mathbb{N})$ 是我们感兴趣的信号集，$\mathcal{F}$ 是酉算子，p 是定义在 $\mathbb{N}$ 上的概率分布，其对应的ˆ算子为 S、D 和 I。若 $\mathcal{A}$ 在 $\mathcal{K}$ 上满足常数为 $\delta \in (0、1)$ 的广义受限等距性质（*Gen-RIP*），则称

$$\sup_{x\in\mathcal{K},\|x\|=1}\left|\frac{1}{m}\|SDFx\|^2-\|\hat{I}Fx\|^2\right|\leqslant\delta$$

为了证明我们的感知算子具有这一性质，我们先给出一个辅助结果。

引理 4.5 (矩阵伯恩斯坦不等式 [46, 定理 5.4。1])。设 $X_1,\ldots,X_N$ 是相互独立、零均值的 $n\times n$ 对称随机矩阵，且 $\|X_i\|\leqslant K$ 几乎必然地，对所有 i均成立。则对每个 $t\geqslant 0$，有

$$\mathbb{P}\left\{\left\|\sum_{i=1}^N X_i\right\|\geqslant t\right\}\leqslant 2n\exp\left(-\frac{t^2/2}{\sigma^2+Kt/3}\right),$$

其中 $\sigma^2=\left\|\sum_{i=1}^N\mathbb{E}X_i^2\right\|$。

现在我们给出首批适用于一般集合 $\mathcal{U}$ 的结果。

引理 4.6. 设 F 是作用于 $\ell^2(\mathbb{N})$上的酉算子，且 $\mathcal{U}\subseteq\ell^2(\mathbb{N})$ 是一个 r维ˆ子空间。设 p 是 $\mathbb{N}$上的概率分布，其关联算子为 S、D和 I，且满足 $\mu_{\mathcal{U}}(F、p)\geqslant 1$。那么，对于 $\varepsilon\in(0、1)$ 以及 $\delta\in(0、1]$，有

$$\mathbb{P}\left[\sup_{x\in\mathcal{U},\|x\|=1}\left|\frac{1}{m}\|SDFx\|^2-\|\hat{I}Fx\|^2\right|\leqslant\delta\right]\geqslant 1-\varepsilon$$

给定

$$m\geqslant\frac{4\mu_{\mathcal{U}}(F,p)^2}{\delta^2}\log\left(\frac{2r}{\varepsilon}\right).$$

证明. 由于 $\mathcal{U}\subset\ell^2(\mathbb{N})$ 是一个 r维子空间，我们可以为其固定一组标准正交基 $(\psi_i)_{i\in[r]}$。接着定义算子 $P_{\mathcal{U}}$: $\ell^2(\mathbb{N})\to\mathbb{C}^r$ 如下:

$$(P_{\mathcal{U}}(x))_i=\langle x,\psi_i\rangle$$

我们观察到 $P_{\mathcal{U}}^*P_{\mathcal{U}}=\Pi_{\mathcal{U}}$，即对所有 $x\in\mathcal{U}$，$x=P_{\mathcal{U}}^*P_{\mathcal{U}}x$。现在来计算

$$\begin{aligned}
&\sup_{x\in\mathcal{U},\|x\|=1}\left|\frac{1}{m}\|SDFx\|^2-\|\hat{I}Fx\|^2\right|\\
&=\sup_{x\in\mathcal{U},\|x\|=1}\left|\frac{1}{m}\left\|SDFP_{\mathcal{U}}^*P_{\mathcal{U}}x\right\|^2-\left\|\hat{I}FP_{\mathcal{U}}^*P_{\mathcal{U}}x\right\|^2\right|\\
&=\sup_{u\in\mathbb{C}^r\cap\mathbb{S}^{r-1}}\left|\frac{1}{m}\left\|SDFP_{\mathcal{U}}^*u\right\|^2-\left\|\hat{I}FP_{\mathcal{U}}^*u\right\|^2\right|\\
&=\frac{1}{m}\sup_{u\in\mathbb{C}^r\cap\mathbb{S}^{r-1}}\left|u^*\left[\left(SDFP_{\mathcal{U}}^*\right)^*\left(SDFP_{\mathcal{U}}^*\right)-m\left(\hat{I}FP_{\mathcal{U}}^*\right)^*\left(\hat{I}FP_{\mathcal{U}}^*\right)\right]u\right|.
\end{aligned}$$

我们可以将其写为算子范数

$$\frac{1}{m}\left\|P_{\mathcal{U}}F^*DS^*SDFP_{\mathcal{U}}^* - mP_{\mathcal{U}}F^*\hat{I}FP_{\mathcal{U}}^*\right\| = \frac{1}{m}\left\|\sum_{i=1}^m\left[P_{\mathcal{U}}F^*\left[Ds_is_i^*D-\hat{I}\right]FP_{\mathcal{U}}^*\right]\right\|.$$

首先，我们注意到

$$\mathbb{E}[Ds_is_i^*D] = \hat{I}.$$

为简化记号，对每个 $i\in[m]$，我们将随机向量 $v_i\in\mathbb{C}^r$ 定义为

$$v_i := P_{\mathcal{U}}F^*Ds_i.$$

由于 D是自伴的，外积恰好为 $v_iv_i^* = P_{\mathcal{U}}F^*Ds_is_i^*DFP_{\mathcal{U}}^*$。因此，我们可以将算子和表示为

$$P_{\mathcal{U}}F^*DS^*SDFP_{\mathcal{U}}^* = \sum_{i=1}^m v_iv_i^*.$$

因此，我们可以将和中每一项的期望写为:

$$\mathbb{E}[v_iv_i^* - P_{\mathcal{U}}F^*\hat{I}FP_{\mathcal{U}}^*] = P_{\mathcal{U}}F^*\mathbb{E}[Ds_is_i^*D-\hat{I}]FP_{\mathcal{U}}^* = 0.$$

因此，构成该和的随机矩阵均为均值为零。接下来，我们几乎必然地对它们的算子范数进行界定。我们来计算 v_i的欧几里得范数。假设随机向量 s_i 取值为某个 $j\in\mathbb{N}$对应的标准基向量 e_j 。那么:

$$\begin{aligned}\|v_i\| &= \|P_{\mathcal{U}}F^*De_j\|\\ &= \frac{1}{\sqrt{p_j}}\|P_{\mathcal{U}}F^*e_j\|\\ &= \frac{1}{\sqrt{p_j}}\sup_{x\in\mathcal{U}\cap B}|\langle x, F^*e_j\rangle|\\ &\leqslant \mu_{\mathcal{U}}(F,p) =: \mu,\end{aligned}$$

其中，不等式直接由局部相干性 α_j和采样参数 $\mu_{\mathcal{U}}(F,p)$的定义得出。由于此界对任意实现 e_j均成立，因此我们有 $\|v_i\|\leqslant\mu$几乎处处成立。现在可以对每个均值为零矩阵的算子范数进行界定:

$$\|v_iv_i^* - P_{\mathcal{U}}F^*\hat{I}FP_{\mathcal{U}}^*\| \leqslant \|v_iv_i^*\| + \|P_{\mathcal{U}}F^*\hat{I}FP_{\mathcal{U}}^*\| \leqslant \mu^2+1\leqslant 2\mu^2,$$

其中，我们用到了秩一恒等式 $\|v_iv_i^*\| = \|v_i\|^2\leqslant\mu^2$、算子范数界 $\|P_{\mathcal{U}}F^*\hat{I}FP_{\mathcal{U}}^*\|\leqslant 1$，以及假设 $\mu\geqslant 1$。

接下来我们计算

$$\begin{aligned}\sigma^2 &= \left\|\sum_{i=1}^m\mathbb{E}\left[\left(v_iv_i^* - P_{\mathcal{U}}F^*\hat{I}FP_{\mathcal{U}}^*\right)^2\right]\right\|\\ &= \sup_{u\in\mathbb{C}^r\cap\mathbb{S}^{r-1}}\left\langle u, \sum_{i=1}^m\left(\mathbb{E}\left[v_iv_i^*v_iv_i^*\right] - \left(P_{\mathcal{U}}F^*\hat{I}FP_{\mathcal{U}}^*\right)^2\right)u\right\rangle\\ &\leqslant \sup_{u\in\mathbb{C}^r\cap\mathbb{S}^{r-1}}\left\langle u, \sum_{i=1}^m\|v_i\|^2\,\mathbb{E}\left[v_iv_i^*\right]u\right\rangle - \sum_{i=1}^m\left\|P_{\mathcal{U}}F^*\hat{I}FP_{\mathcal{U}}^*u\right\|^2\\ &\leqslant \sup_{u\in\mathbb{C}^r\cap\mathbb{S}^{r-1}}\sum_{i=1}^m\mu^2\langle u, \mathbb{E}\left[v_iv_i^*\right]u\rangle\\ &\leqslant \mu^2 m.\end{aligned}$$

第二个等式成立，是因为范数中的算子是半正定的，这意味着其算子范数等于其二次型在单位球面上的上确界。后续不等式是通过舍去非正的减项并应用ˆ界 $\|\mathbb{E}[v_i v_i^*]\| = \|P_{\mathcal{U}} F^* I F P_{\mathcal{U}}^*\| \leqslant 1$ 得出的。

由此，我们应用维度为 4.5、方差参数为 r、谱界为 $\sigma^2 \leqslant m\mu^2$ 的矩阵伯恩斯坦不等式（引理 $L \leqslant 2\mu^2$），得到：

$$\mathbb{P}\left\{\left\|\sum_{i=1}^{m}\left[P_{\mathcal{U}} F^*\left[D s_i s_i^* D - \hat{I}\right] F P_{\mathcal{U}}^*\right]\right\| \geqslant \gamma\right\} \leqslant 2r \exp\left(-\frac{\gamma^2/2}{m\mu^2 + \frac{2}{3}\mu^2\gamma}\right).$$

该界可转化为子空间上的算子范数。由于 $P_{\mathcal{U}}^*$ 是从 $\mathbb{C}^r$到 $\mathcal{U}$的等距映射，上述条件等价于：

$$\mathbb{P}\left\{\sup_{x\in\mathcal{U},\|x\|=1}\left|\frac{1}{m}\|SDFx\|^2 - \|\hat{I}Fx\|^2\right| \geqslant \frac{\gamma}{m}\right\} \leqslant 2r \exp\left(-\frac{\gamma^2/2}{m\mu^2 + \frac{2}{3}\mu^2\gamma}\right).$$

代入 $\delta = \frac{\gamma}{m}$后，指数变为：

$$-\frac{(m\delta)^2/2}{m\mu^2 + \frac{2}{3}\mu^2 m\delta} = -\frac{m\delta^2}{2\mu^2(1+\frac{2}{3}\delta)}.$$

假设 $\delta \leqslant 1$，则有 $(1+\frac{2}{3}\delta) \leqslant 5/3 < 2,$，这意味着 $\frac{1}{2(1+\frac{2}{3}\delta)} > \frac{1}{4}$。由此得到简化后的界：

$$\mathbb{P}\left\{\sup_{x\in\mathcal{U},\|x\|=1}\left|\frac{1}{m}\|SDFx\|^2 - \|\hat{I}Fx\|^2\right| \geqslant \delta\right\} \leqslant 2r \exp\left(-\frac{m\delta^2}{4\mu^2}\right).$$

最后，要求右侧不超过 ε，

$$2r \exp\left(-\frac{m\delta^2}{4\mu^2}\right) \leqslant \varepsilon,$$

我们得出结论：只要满足以下条件，该条件即成立

$$m \geqslant \frac{4\mu_{\mathcal{U}}(F,p)^2}{\delta^2} \log\left(\frac{2r}{\varepsilon}\right). \qquad \square$$

我们现在给出主要的广义受限等距性质（Gen-RIP）结果，该结果确立了受限等距对一个 (k, d)-广义生成网络值域内的信号成立。

定理 4.7（Gen-RIP）。设 F 为 $\ell^2(\mathbb{N})$ 上的酉算子，并设 p 为 $\mathbb{N}$ 上的概率分布，其关联算子为 S、 D以及ˆI。设 G 为一个 $(k$、 $d)$-广义生成网络，且 $\mathcal{T} = R(G) - R(G)$。对于 $\varepsilon > 0$、 $\delta \in (0$、 $1)$，假设测量值的数量满足

$$m \geqslant \frac{4\mu_{\mathcal{T}}(F,p)^2}{\delta^2}\left[2kd\log\left(\frac{2ek_d}{k}\right) + \log\left(\frac{4k}{\epsilon}\right)\right],$$

其中我们假设 $\mu_{\mathcal{T}}(F,p) \geqslant 1$。那么，以至少 $1-\varepsilon$ 的概率，以下结论成立：

$$\sup_{x\in\mathcal{T},\|x\|=1}\left|\frac{1}{m}\|SDF\boldsymbol{x}\|^2 - \|\hat{I}F\boldsymbol{x}\|^2\right| \leqslant \delta$$

证明。 根据引理 4.1， $\mathcal{T}$是一个至多包含 N^2 $2k$维子空间的集合。因此，我们可以写成 $\mathcal{T}=\bigcup_{j=1}^{N}\mathcal{L}_j$，并对每个 $\mathcal{L}_j$ 应用引理 4.6 ，令 $\varepsilon'>0$ 即可得到

$$\begin{aligned}&\mathbb{P}\left\{\sup_{x\in\mathcal{T},\|x\|=1}\left|\frac{1}{m}\|SDF\boldsymbol{x}\|^2-\|\hat{I}F\boldsymbol{x}\|^2\right|\geqslant\delta\right\}\\&\leqslant\sum_{j=1}^{N}\mathbb{P}\left\{\sup_{x\in\mathcal{L}_j,\|x\|=1}\left|\frac{1}{m}\|SDF\boldsymbol{x}\|^2-\|\hat{I}F\boldsymbol{x}\|^2\right|\geqslant\delta\right\}\\&\leqslant N\varepsilon'.\end{aligned}$$

为了求出 m所要求的速率，我们令 $\epsilon'=\epsilon/N^2$ 在引理 4.6中满足:

$$m\geqslant\frac{4\mu_{\mathcal{U}}(F,p)^2}{\delta^2}\log\left(\frac{2kN}{\epsilon}\right)=\frac{4\mu_{\mathcal{U}}(F,p)^2}{\delta^2}\left[\log N^2+\log\left(\frac{4k}{\epsilon}\right)\right].$$

将引理 $N\leqslant kd$ 4.1 中的界 $\log\binom{2ek_d}{k}$ log 代入 即可得到所需的样本复杂度。 □

4.3. **恢复界**。 现在，我们已具备证明主定理和推论所需的一切条件。

定理 *2.3* 的证明。 回顾 $x^{\perp}:=x_0-\Pi_{R(G)}(x_0)$。根据估计量的定义和三角不等式，我们有

$$\begin{aligned}\|SDF\hat{x}-b\|&\leqslant\min_{x\in R(G)}\|SDFx-b\|+\hat{\varepsilon}\\&\leqslant\left\|SDF\Pi_{R(G)}(x_0)-b\right\|+\hat{\varepsilon}\\&=\left\|SDFx^{\perp}+\eta\right\|+\hat{\varepsilon}\\&\leqslant\left\|SDFx^{\perp}\right\|+\|\eta\|+\hat{\varepsilon}.\end{aligned}$$

根据假设，测量值的数量 m 足够大，因此根据推论 4.7，测量算子 SDF 以至少 $\mathcal{T}=\Delta(R(G))$ 的概率，在差集 $\delta=1/2$, 上满足常数为 $1-\varepsilon$的受限等距性质。因此，由于差向量 $\hat{}x-\Pi_{R(G)}(x_0)\in\Delta(R(G))$，我们得到

$$\begin{aligned}\|SDF\hat{x}-b\|&=\left\|SDF\left(\hat{x}-\Pi_{R(G)}(x_0)\right)-SDFx^{\perp}-\eta\right\|\\&\geqslant\left\|SDF\left(\hat{x}-\Pi_{R(G)}(x_0)\right)\right\|-\left\|SDFx^{\perp}\right\|-\|\eta\|\\&\geqslant\sqrt{\frac{m}{2}}\left\|\hat{I}F\left(\hat{x}-\Pi_{R(G)}(x_0)\right)\right\|-\left\|SDFx^{\perp}\right\|-\|\eta\|.\end{aligned}$$

合并这两个不等式可得

$$\left\|\hat{I}F\left(\hat{x}-\Pi_{R(G)}(x_0)\right)\right\|\leqslant\sqrt{\frac{2}{m}}\left[2\left\|SDFx^{\perp}\right\|+2\|\eta\|+\hat{\varepsilon}\right].$$

应用三角不等式，并利用 F 的酉性，我们得到一般界:

$$\begin{aligned}\|\hat{}x-x_0\|&\leqslant\left\|x_0-\Pi_{R(G)}(x_0)\right\|+\left\|\hat{}x-\Pi_{R(G)}(x_0)\right\|\\&\leqslant\left\|x^{\perp}\right\|+\left\|\hat{I}F(\hat{x}-\Pi_{R(G)}(x_0))\right\|+\left\|\hat{I}^{\perp}F(\hat{x}-\Pi_{R(G)}(x_0))\right\|\\&\leqslant\left\|x^{\perp}\right\|+\frac{2\sqrt{2}}{\sqrt{m}}\left\|SDFx^{\perp}\right\|+\frac{2\sqrt{2}}{\sqrt{m}}\|\eta\|+\frac{\sqrt{2}}{\sqrt{m}}\hat{\varepsilon}+\left\|\hat{I}^{\perp}F(\hat{x}-\Pi_{R(G)}(x_0))\right\|.\end{aligned}$$

这证明了（3）。

对于第二部分，假设 p 是可容许的。我们观察到$\hat{I}^{\perp}Fu = 0$ 对所有 $u \in \Delta(R(G))$。这成立是因为，若局部相干性满足 $\alpha_j > 0$，则可容许性要求 $p_j >0$，从而意味着正交投影分量$\hat{I}^{\perp}_{j,j} = 0$。反之，若 $\alpha_j = 0$，则第 j个分量根据局部相干性的定义严格为零。由于$\hat{x} - \Pi_{R(G)}(x_0) \in \Delta(R(G))$，尾项 $\left\|\hat{I}^{\perp}F(\hat{x} - \Pi_{R(G)}(x_0))\right\|$完全消失，从而得到（4）。 □

推论的证明 *2.4*。我们只需注意到

$$\left\|\hat{I}^{\perp}F(\hat{x} - \Pi_{R(G)}(x_0))\right\| = \left\|\hat{I}^{\perp}F\Pi_{R(G)}(\hat{x} - x_0)\right\| \leqslant \theta\|\hat{x} - x_0\|,$$

只需代入（3），即可得到结论。 □

5. Conclusions

在这项工作中，我们为无限维生成式压缩感知奠定了理论基础。通过在 $\ell^2(\mathbb{N})$ 中恰当地表述问题并考虑离散化间隙，我们证明了生成先验能够有效用于函数重建。我们引入了无限维局部相干性的概念，并利用它推导出最优采样策略。我们在达西流问题上开展的数值实验验证了这些理论见解，表明基于相干性的采样能够显著减少实现精确重建所需的测量值数量。

此外，我们的实证结果揭示了模型分辨率与采样密度之间一种有益的相互作用。我们发现，在数据稀缺情形下，采用较低分辨率的生成器并进行显式升尺度，可以形成一种隐式正则化，从而稳定逆问题。我们的研究结果表明，最优生成先验未必是保真度最高的先验，而是最能匹配可用测量值频谱内容的先验。通过限制分辨率，我们利用网络 [42] 的频谱偏差，在数据稀缺情形下稳定反演。

尽管我们的理论框架具有稳健性，但它依赖于生成器采用分段线性激活函数（例如 ReLU）这一假设，以确保值域包含在有限维多面锥的并集中。将这一理论扩展到采用 GeLU 或 SiLU 等平滑激活函数的网络，仍是一项尚待解决的理论挑战。此外，尽管我们的理论给出了存在测量噪声时的误差界，但数值验证主要集中在无噪声情形。未来的工作可以探究该框架在不同噪声水平和分布下的实证稳健性。最后，将这些无限维 Gen-RIP 保证推广到其他实际测量算子（尤其是非酉算子），例如用于连续计算机断层扫描的拉东变换，是自然的下一步。

Acknowledgments

本材料基于美国空军科学研究办公室通过资助编号 FA8655-23-1-7083 支持的研究工作。该研究还部分得到授予热那亚大学数学系的 MIUR 卓越系项目支持，项目编号为 CUP D33C23001110001。该项目由区域 FSE+共同资助（欧盟条例 1060/2021 附件 IX 第 1.2 点）。MS 的研究得到大学与研究部资助项目 PRIN P2022XT498 的支持，该项目由欧盟委员会通过下一代欧盟计划资助。MS 是“意大利国家高等数学研究所”的成员。

Conflict of interest

作者声明，他们不存在任何利益冲突。

References

[1] Ben Adcock. 无限维压缩感知与函数插值。 *Foundations of Computational Mathematics* 18 3 661 701, 2018 , ( ): – . [2]

Ben Adcock. 最小二乘逼近的最优采样。 *Foundations of Computational Mathematics*, 25 6 1975 2034, 2025 ( ): – . [3] Ben Adcock、Claire Boyer 和 Simone Brugiapaglia. 关于压缩感知中预言机型局部恢复保证的研究。 *Information and Inference: A Journal of the IMA*, 10(1):1–49, 2021. [4] Ben Adcock、Simone Brugiapaglia、Nick Dexter 和 Sebastian Moraga. *On Efficient Algorithms for Computing Near-Best Polynomial Approximations to High-Dimensional Hilbert-Valued Functions from Limited Samples* , . 欧洲数学学会 - EMS - 出版社有限公司， . ISBN: 9783985470709 9783985475704 2747 9080, 2747 9099 ISSN: - - 2024 .[5] Ben Adcock、Simone Brugiapaglia 和 Clayton G Webster. *Sparse polynomial approximation of high-dimensional functions* 25 2022 , 第 . 卷。 SIAM， .[6] Ben Adcock、Juan M. Cardenas 和 Nick Dexter. 从任意线性样本中学习非线性模型类的统一框架。载于 *Forty-first International Conference on Machine Learning*, 2024. [7] Ben Adcock 和 Anders C Hansen. 任意基中稳定重建的广义采样定理。 *Journal of Fourier Analysis and Applications*, 18(4):685–716, 2012. [8] Ben Adcock 和 Anders C Hansen. 广义采样与无限维压缩感知。 *Foundations of Computational Mathematics* 16 5 1263 1329, 2016 , ( ): – . [9]

Ben Adcock、Anders C Hansen、Clarice Poon 和 Bogdan Roman. 打破相干性障碍：压缩感知新理论。载于 *Forum of mathematics, sigma*, 第 5 卷， 第 e4 页。剑桥大学出版社， 2017. [10] Giovanni S Alberti、Alessandro Felisi、Matteo Santacesaria 和 S Ivan Trapasso. 逆问题的压缩感知及稀疏拉东变换的样本复杂度。 *Journal of the European Mathematical Society* 2025 , .[11]

Giovanni S Alberti、Alessandro Felisi、Matteo Santacesaria 和 S Ivan Trapasso. 逆问题的压缩感知 II：反卷积、源恢复和磁共振成像中的应用。 *arXiv preprint arXiv:2501.01929*, 2025.[12] Giovanni S Alberti 和 Matteo Santacesaria. 有限测量下的无限维逆问题。*Archive for Rational Mechanics and Analysis*, 243(1):1–31, 2022. [13] Giovanni S. Alberti、Matteo Santacesaria 和 Silvia Sciutto. 连续生成神经网络：函数空间中的基于小波的架构。 *Numerical Functional Analysis and Optimization* 46 1 1 44, , ( ): – .[14]

Richard Baraniuk、Mark Davenport、Ronald DeVore 和 Michael Wakin. 随机矩阵受限等距性质的简单证明。 *Constructive approximation*, 28(3):253–263, 2008. [15] Aaron Berk、Simone Brugiapaglia、Babhru Joshi、Yaniv Plan、Matthew Scott 和 Ozgur Yilmaz. 一种刻画基于傅里叶测量的生成式压缩感知的相干性参数。 *IEEE Journal on Selected Areas in Information Theory* 3 3 502 512, 2022 , ( ): – .

[16] Aaron Berk、Simone Brugiapaglia、Yaniv Plan、Matthew Scott、Xia Sheng 和 Ozgur Yilmaz. 面向生成式压缩感知的模型自适应傅里叶采样。载于 *NeurIPS 2023 Workshop on Deep Learning andInverse Problems*, 2023.[17] Sam Bond-Taylor 和 Chris G. Willcocks. ∞-diff：采用子采样平滑状态的无限分辨率扩散。载于 *The Twelfth International Conference on Learning Representations*, 2024. [18] Ashish Bora、Ajil Jalal、Eric Price 和 Alexandros G. Dimakis. 使用生成模型进行压缩感知。载于 *Proceedings of the 34th International Conference on Machine Learning*, 第 537–546 页， 2017.[19] Nicolas Boulle´ *Handbook of Numerical Analysis* 25, 83 125 2024 e 和 Alex Townsend. 算子学习数学指南。载于 , 第 卷 页 – 。Elsevier， .[20] Justin Bunker、Mark Girolami、Hefin Lambley、Andrew M Stuart 和 TJ Sullivan. 函数空间中的自编码器。 *Journal of Machine Learning Research*, 26(165):1–54, 2025. [21] Emmanuel Candes 和 Justin Romberg. 压缩采样中的稀疏性与非相干性。 *Inverse problems*, 23 3( ):969, 2007. [22] Emmanuel J Cand`es、Justin Romberg 和 Terence Tao. 鲁棒不确定性原理：从高度不完整的频率信息中精确恢复信号。 *IEEE Transactions on information theory*, 52(2):489–509, 2006. [23] Emmanuel J Candes 和 Terence Tao. 通过线性规划进行解码。 *IEEE transactions on information theory* 51 12 4203 4215, 2005 , ( ): – .

[24] John A Cherry 和 R Allan Freeze. *Groundwater*, 第 370 卷。新泽西州恩格尔伍德克利夫斯：Prentice-Hall， 1979.

[25] Hyungjin Chung、Jeongsol Kim、Michael Thompson Mccann、Marc Louis Klasky 和 Jong Chul Ye。《面向一般噪声逆问题的扩散后验采样》。载于 *The Eleventh International Conference on Learning Representations* 2023，.

[26] Albert Cohen 和 Giovanni Migliorati。《最优加权最小二乘法》。*SMAI Journal of Computational Mathematics* 3 181 203, 2017，：–。[27] Giannis Daras、Hyungjin Chung、Chieh-Hsin Lai、Yuki Mitsufuji、Jong Chul Ye、Peyman Milanfar、Alexan- dros G Dimakis 和 Mauricio Delbracio。《用于逆问题的扩散模型综述》。*arXiv preprint arXiv:2410.00083* 2024，。[28] David L Donoho。《压缩感知》。*IEEE Transactions on information theory*，52(4):1289–1306，2006年。[29] Heinz Werner Engl、Martin Hanke 和 Andreas Neubauer。*Regularization of Inverse Problems*，*Mathematics and Its Applications*第375卷。施普林格科学与商业媒体，1996年。[30] Charles Fefferman、Sanjoy Mitter 和 Hariharan Narayanan。《检验流形假设》。*Journal of the American Mathematical Society* 29 4 983 1049, 2016，( ):–。[31] Giulio Franzese 和 Pietro Michiardi。《无限维中的生成扩散模型：综述》。*Philosophical Transactions of the Royal Society A* 383 2299 2025，( )，。[32] 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，,。[33] Philipp Grohs 和 Gitta Kutyniok。《抛物分子》。*Foundations of Computational Mathematics*，14 2 299( ):–337，2014年。[34] Joseph B Keller。《多孔介质流动的达西定律与双空间方法》。载于 *Nonlinear partial differential equations in engineering and applied science* 429 443 2017，第–页。劳特利奇，。[35] Diederik P Kingma 和 Max Welling。《自编码变分贝叶斯》。*International Conference on Learning Representations* 2013，。[36] Nikola Kovachki、Zongyi Li、Burigede Liu、Kamyar Azizzadenesheli、Kaushik Bhattacharya、Andrew Stuart 和 Anima Anandkumar。《神经算子：学习函数空间之间的映射及其在偏微分方程中的应用》。*Journal of Machine Learning Research*，24(89):1–97，2023年。[37] Samuel Lanthaler、Siddhartha Mishra 和 George E Karniadakis。《DeepONet 的误差估计：无限维中的深度学习框架》。*Transactions of Mathematics and Its Applications* 6 1 001, ，( ):tnac。[38] Zongyi Li、Nikola Borislavov Kovachki、Kamyar Azizzadenesheli、Burigede liu、Kaushik Bhattacharya、Andrew Stuart 和 Anima Anandkumar。《用于参数化偏微分方程的傅里叶神经算子》。载于 *International Conference on Learning Representations*，2021年。[39] Lu Lu、Pengzhan Jin、Guofei Pang、Zhongqiang Zhang 和 George Em Karniadakis。《基于算子通用逼近定理的 DeepONet 非线性算子学习》。*Nature Machine Intelli-gence*，3(3):218–229，2021年。[40] St´ephane Mallat。《群不变散射》。*Communications on Pure and Applied Mathematics*，65 10 1331( ):–1398，2012年。[41] Alireza Naderi 和 Yaniv Plan。《用于生成先验的无稀疏性压缩感知》。*IEEE Journal on Selected Areas in Information Theory* 3 3 493 501, 2022，( ):–。[42] Nasim Rahaman、Aristide Baratin、Devansh Arpit、Felix Draxler、Min Lin、Fred Hamprecht、Yoshua Bengio 和 Aaron Courville。《神经网络的频谱偏差》。载于 *International conference on machine learning* 5301 5310 2019，第–页。PMLR，。[43] Holger Rauhut 和 Rachel Ward。《通过 ℓ_1 -最小化实现稀疏勒让德展开》。*Journal of Approximation Theory* 164 5 517 533, 2012，( ):–。[44] Jascha Sohl-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 *International conference on machine learning*，第2256–2265页。PMLR，2015年。[45] 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》。载于 *International Conference on Learning Representations* 2021，。[46] Roman Vershynin。*High-dimensional probability: An introduction with applications in data science*，第47卷。剑桥大学出版社，2018年。

MaLGa Center, DIMA, University of Genoa, Via Dodecaneso 35, 16146, Genova, Italy
Email address: paolo.angella@edu.unige.itMaLGa Center, DIBRIS, University of Genoa, Via Dodecaneso 35, 16146, Genova, Italy*Email address*: vito.paolo.pastore@unige.it
MaLGa Center, DIMA, University of Genoa, Via Dodecaneso 35, 16146, Genova, Italy
Email address: matteo.santacesaria@unige.it