arXiv:2302.03577v2 [math.FA] 3 May 2024

逆问题的压缩感知与稀疏 Radon 变换的样本复杂度

GIOVANNI S. ALBERTI、ALESSANDRO FELISI、MATTEO SANTACESARIA 和 S. IVAN TRAPASSO

Abstract. 压缩感知能够通过少量测量恢复稀疏信号，而测量数量与未知信号的稀疏度成正比，至多相差对数因子。经典理论通常研究随机线性测量或子采样等距映射，并已应用于诸多领域，包括加速磁共振成像；其模型为子采样傅里叶变换。本文针对抽象逆问题（可能是不适定问题）发展了一套无限维压缩感知的一般理论，其中涉及任意前向算子。为此，我们考察了广义受限等距性质以及前向映射的拟对角化性质。

作为一项重要应用，我们首次针对稀疏 Radon 变换（即仅使用有限个角度 $\theta_1, \dots, \theta_m$）获得了严格的恢复估计。该变换在平行束和扇束设置下均可用于对计算机断层成像建模。当未知信号相对于由紧支撑小波构成的正交基是 s-稀疏时，我们证明，在满足以下条件的情况下可以实现稳定恢复

$$m \gtrsim s,$$

相差对数因子。

日期：2024年5月6日。2020年数学主题分类：42C40；44A12；60B20；94A20。关键词和短语：Radon 变换、无限维压缩感知、稀疏恢复、样本复杂度、小波、拟对角化、相干性、逆问题。本材料基于空军科学研究办公室资助的项目成果，资助编号为 FA8655-20-1-7027。我们感谢圣保罗公司基金会的支持。本项目由欧盟共同资助（欧洲研究委员会，SAMPDE，101041040）。然而，文中表达的观点和意见仅代表作者个人观点，并不一定反映欧盟或欧洲研究委员会的观点。欧盟和拨款机构均不对此承担责任。作者是“Gruppo Nazionale per l’Analisi Matematica, la Probabilit`a e le loro Applicazioni”以及“国家高等数学研究所”的成员。

Contents

1. 引言 2 1.1. 最先进水平 2 1.2. 主要贡献 3 1.3. 相关研究 8 1.4. 讨论与开放问题 10 1.5. 论文结构 11 2. 设置 12 2.1. 加权设置与记号 12 2.2. 字典、测量与最小化问题 13 3. 关于样本复杂度的一般结果 16 3.1. 主要结果 16 3.2. 关于 $\|G^{-1}\|$ 的定量界与拟对角化 18 3.3. 恢复误差分析 23 4. 稀疏 Radon 变换 24 5. 主要结果的证明 29 5.1. 预备结果 29 5.2. 稳健零空间性质 30 5.3. 广义受限等距性质 31 5.4. 通过子采样满足广义限制等距性质的充分条件 33 5.5. 定理 3.5 的证明 34 5.6. 关于 $\|G^{-1}\|$ 的定量界 38 5.7. 稀疏 Radon 变换 42 参考文献 45 附录 A. 定理 5.7 的证明 50 A.1. $f_\nu = 1$ 的情形 50 A.2. 一般情形 57

1. Introduction

1.1. 最先进水平。过去二十年来，压缩感知（CS） [38, 48, 51, 54, 52]一直是应用数学领域中十分活跃的研究方向，其特点是理论分析与技术进步之间的良性互动。压缩感知的重点是恢复稀疏信号，即那些能够表示为参考字典中少数向量线性组合的信号。利用稀疏性，可以在测量数量少于标准采样理论所需数量的情况下实现稳定恢复结果。

多年来，压缩感知理论已从有限维设定转向无限维设定 [8, 5, 1, 14, 9]。这一转变主要是由以下需求推动的：

涵盖越来越具挑战性的感知任务，而这些任务若在模拟世界中通过适当的算子和函数空间来表述，便能得到更忠实的刻画。事实上，研究如何根据间接测量恢复某个物理量的问题，是逆问题理论的主要关注点；这一领域在数学分析和数值分析中有着悠久的传统[73, 82, 92, 65]。大多数逆问题都由积分方程或微分方程描述，因此自然地建立在无限维设定中。

目前数学文献中关于逆问题的重建保证，都依赖于这样一个假设：除少数特殊情况外，观测者可以获取任意多的测量（参见 [13, 63, 15, 11] 及其中的参考文献；不过，这些文献并未利用任何稀疏性信息，而只利用了未知量的有限维性）。显然，在现实场景中，我们只能进行有限次测量；这些测量通常会受到噪声干扰，而且由于实验约束，对测量过程的控制也有限。因此，有必要发展适用于病态逆问题的样本复杂度理论，即研究实现稳定且准确的重建所需的最少（有限）测量次数。

直观而言，压缩感知的原理或许能在此发挥作用。事实上，如果未知量在某个字典中是稀疏的，那么我们就可以在测量算子的定义域中的某个有限维子空间内，对信号的位置进行近似定位。换言之，稀疏性条件天然伴随着有限维设定。不难理解，这种约化步骤相当微妙，最终可能会破坏有关真正无限维结构的重要信息。广义采样框架 [6, 8, 7, 5]正是为了解决这一问题：它通过确定合适的采样带宽和稀疏性截断带宽，实现稳定的有限维约化。

尽管压缩感知的思想已影响了逆问题的分析（例如，稀疏正则化 [17, 59, 58]），但最先进水平仍远未达到完全令人满意的程度，这主要是由于技术限制，以及旨在利用每个模型独特特征的逐案处理方法。尽管关于逆问题中稀疏性促进方法的数值分析已有大量文献（例如，参见[66, 70, 55]），但严格的理论保证仍然缺失（关于这一方向的一些结果，参见 [64, 14, 12] ）。

1.2. 主要贡献。

1.2.1. 摘要。本文旨在通过扩展当前压缩感知理论，为在现实的稀疏性约束下研究病态逆问题的样本复杂度迈出构建统一框架的第一步。主要目标是找出实现信号恢复所需的样本数量与该信号相对于合适字典的稀疏度之间的联系。

除了构建研究样本复杂度的一般框架外，我们还将其应用于稀疏角度层析成像问题。基本上，目标是根据通过 Radon 变换获得的有限数量的投影，获取物体内部结构的信息。断层成像技术应用于各种现实场景，包括无损检测、生物医学与地球物理成像，以及天文信号分析。断层成像逆问题的分析已是一个成熟的研究课题，同时仍是一个十分活跃的研究领域 [83, 84, 93, 92, 62]。

尽管稀疏 Radon 变换已从数值角度得到深入研究，且压缩感知的思想在其出现后不久便应用于断层成像，但迄今为止，压缩感知理论尚无法给出关于 Radon 变换样本复杂度的严格结果，也无法应对相关技术挑战。我们在这方面的主要贡献是证明：对于支撑在平面单位球内、且假定能在适当的小波字典中进行可压缩表示的信号，只要获得足够多的、对应于不同角度的 Radon 变换样本，就能以高概率重建——见下文定理 4.4、4.5 和 4.6。我们同时考虑平行束和扇束设置。简而言之，我们得到如下估计

$$\text{角度数} \;\gtrsim\; \text{稀疏度},$$

至多相差对数因子。

下面简要说明抽象样本复杂度框架及其在稀疏拉东反演中的应用所涉及的关键方面。

1.2.2. 稀疏病态逆问题的样本复杂度。设 $\mathcal{H}_1$、$\mathcal{H}_2$ 为可分复希尔伯特空间。我们假设测量由一族一致有界的线性映射 $(F_t)_{t\in\mathcal{D}}$ 表示：$\mathcal{H}_1 \to \mathcal{H}_2$。这些映射以测度空间 $(\mathcal{D}, \mu)$ 为索引，该测度空间赋有概率分布 ν，满足 $\mathrm{d}\nu = f_\nu \mathrm{d}\mu$，其中密度函数为 $f_\nu \in L^1(\mu)$，$\|f_\nu\|_{L^1} = 1$，并且对于某个 $0 < c_\nu \le 1$，满足 $f_\nu \ge c_\nu$。因此，该模型的前向映射 $F\colon \mathcal{H}_1 \to L^2_\mu(\mathcal{D};\, \mathcal{H}_2)$ 可自然地由关系式 $F_t u = Fu(t)$ 定义，该关系对几乎处处的 $t \in \mathcal{D}$ 成立。这一设定涵盖了两类测量：相对于某个字典通过内积给出的标量测量，以及逐点求值。此外，我们还考虑取值于希尔伯特空间的测量；这与压缩感知中的分块采样策略有关 [86, 20]。

我们固定一个正交归一基 $(\phi_i)_{i\in\Gamma}$，作为 $\mathcal{H}_1$ 的参考字典，并令 $\Phi\colon \mathcal{H}_1 \to \ell^2(\Gamma)$ 为相应的分析算子——这里，Γ 是一个至多可数的指标集，例如 $\Gamma = \mathbb{N}$ 或 $\Gamma \subset \mathbb{N}\times\mathbb{Z}$。我们所考虑的信号相对于 $(\phi_i)_i$ 是稀疏的。

假设字典、测量算子和概率密度由如下形式的相干性界相互关联：

$$\|F_t\phi_i\|_{\mathcal{H}_2} \le B\sqrt{f_\nu(t)}, \quad t \in \mathcal{D}, i \in \Gamma,$$

对于某个 $B \ge 1$。

现在考虑一个未知信号 $u^\dagger \in \mathcal{H}_1$。我们假设存在一个指标的有限子集 $\Lambda \subset \Gamma$，满足 $|\Lambda| = M$，使得 $\|P_\Lambda^\perp x^\dagger\|$ 很小，其中我们设定 $x^\dagger := \Phi u^\dagger$，并且 P_Λ 表示在张成空间 $\{e_i$ 上的正交投影：$i \in \Lambda\}$，$(e_i)_{i\in\Gamma}$ 为 $\ell^2(\Gamma)$ 的标准基。因此，应将 Λ 视为 $x^\dagger$ 的支撑的粗略估计。对于 $\|P_\Lambda^\perp x^\dagger\|_2$，还可以考虑更一般形式的尾部衰减率，这些衰减率通常与 $u^\dagger$ 和 Φ 的正则性相关。我们将在第 3 节和第 4 节中利用这一联系，讨论其在相关信号类中的应用。为简便起见，这里我们假设 $P_\Lambda^\perp x^\dagger = 0$。

给定从 ν 抽取的 m 独立同分布样本 t_1、.....、$t_m \in \mathcal{D}$，相应的测量值为

$$y_k = F_{t_k}u^\dagger + \varepsilon_k, \qquad k = 1, \dots, m,$$

其中，噪声向量 $\varepsilon \in \mathcal{H}_2^m$ 满足：对于某个 $\beta > 0$，对每个 k，都有 $\|\varepsilon_k\|_{\mathcal{H}_2} \le \beta$ ——换言之，以 t_k 为索引的测量值所带的噪声 ε_k 在所有样本上均一致有界。

我们尝试通过 ℓ^1-最小化问题重建 $x^\dagger$（并由此重建 $u^\dagger = \Phi^* x^\dagger$）

$$\min_{x\in\ell^2(\Lambda)} \|x\|_1 \quad : \quad \frac{1}{m}\sum_{k=1}^{m} \|F_{t_k}\Phi^* x - y_k\|_{\mathcal{H}_2}^2 \le \beta^2, \tag{1}$$

通过对模型的前向映射 F 进行稳定截断，实现向有限维设定的转换。更确切地说，我们考虑由下式给出的 $M \times M$ 矩阵

$$G = \sqrt{P_\Lambda \Phi F^* F \Phi^* P_\Lambda^*}.$$

稳定性最终归结为 G 的可逆性；若 F 满足一种称为 FBI 性质的受限单射性假设，则可确保这一点。许多值得关注的逆问题通常都满足该假设——例如，当算子 F 为单射时，详见下文定义 3.6。更一般地，若在最小化问题 (1) 中引入某种弹性网络正则化，则可以放宽这一条件，使得对于 $\alpha > 0$，可将 G 替换为 $G_\alpha = \sqrt{P_\Lambda \Phi F^* F \Phi^* P_\Lambda^* + \alpha^2 I_\Lambda}$，从而无需再考虑 G_α 的可逆性。为求简便，本文不讨论这一推广。

关于抽象逆问题样本复杂度的抽象结果，其大幅简化后的版本（定理 3.5）如下。

定理。在目前为止所述假设下，存在普适常数C_0、C_1、$C_2 > 0$，使得对于任意整数 $2 \le s \le |\Lambda|$，如果 $\widehat{x} \in \ell^2(\Lambda)$ 是最小化问题 (1) 的一个解，且样本数满足

$$m \ge C_0 B^2 \|G^{-1}\|^4 s,$$

在相差对数因子的范围内，以下恢复估计以极高概率成立：

$$\|x^\dagger - \widehat{x}\|_2 \le C_1 \frac{\sigma_s(x^\dagger)_1}{\sqrt{s}} + C_2 c_\nu^{-1/2} \|G^{-1}\| \beta,$$

项 $\sigma_s(x^\dagger)_1 = \inf\{\|x^\dagger - z\|_1 : \ z \in \ell^2(\Gamma) \ | \operatorname{supp}(z)| \le s\}$，表示对 $x^\dagger$ 进行最佳 s-稀疏逼近时产生的误差。

后续的大部分工作都致力于控制量 $\|G^{-1}\|$，因为它在前述结果中发挥着重要作用。在第 3.2 节中，我们讨论了几种为 $\|G^{-1}\|$ 给出显式界的策略，最终将其与 $F\Phi^* P_\Lambda^*$ 的奇异值联系起来。如果无法获取奇异值，我们会介绍一种替代方法，它依赖于一个在实践中通常成立的相容性条件：逆问题的测量算子在 $(\phi_i)_{i\in\Gamma}$ 上近似表现为对角算子，这与小波-模糊分解 [46] 类似。粗略地说，定义 3.7 中引入的拟对角化条件刻画了前向映射对信号稀疏度的影响，因此我们可以通过模拟奇异值分解来得到 $\|G^{-1}\|$ 的界。

此外，秉承统计逆向学习的思路 [21, 27, 29]（简要介绍见下文第 1.3 节），还可以进一步优化恢复估计，使其仅依赖于观察者能够调整的参数，即噪声水平 β 和样本数 m。更确切地说，在关于 $u^\dagger$ 相对于字典 $(\Phi_i)_i$ 的线性和非线性逼近误差满足适当条件时（若 $u^\dagger$ 属于某些信号类，例如卡通类图像，通常便能满足这些条件——更多详情见 §1.2.3 末尾），只要将 m 适当地选为噪声的函数，我们就能得到如下形式的估计

$$\|x^\dagger - \widehat{x}\|_2 \le c\beta^\alpha, \qquad \|x^\dagger - \widehat{x}\|_2 \le c\frac{1}{m^{\alpha'}},$$

结果相差对数因子，其中常数 $c > 0$取值适当。这里，α、$\alpha' > 0$ 显式依赖于线性和非线性逼近误差率，以及前向映射 F的病态程度。上述估计以高概率成立，不过也可以推导出期望误差的类似估计。

有关这些结果的更多细节和讨论，请参阅第3节；关于如何推广压缩感知的标准技术（尤其是受限等距性质及其

推论），以应对病态逆问题带来的困难，请参阅第5.2、5.3和5.4节。

1.2.3. 稀疏 Radon 变换。在第4节中，我们将讨论如何从有限样本中获得稀疏角度拉东变换求逆的恢复结果。具体而言，我们考虑从沿角度 $\theta \in [0, 2\pi$ 的 Radon 变换样本中，重建支撑于 $\mathbb{R}^2$ 中单位球内的信号。

回顾一下，对于平面上具有紧支撑的平方可积信号 u，沿角度 $\theta \in [0, 2\pi$ 的 Radon 变换定义为

$$\mathcal{R}_\theta u(s) \coloneqq \int_{e_\theta^\perp} u(y + se_\theta)\mathrm{d}y, \quad s \in \mathbb{R},$$

$\mathrm{d}y$ 是切片 $\theta^\perp$ 上的一维勒贝格测度。完整的 Radon 变换 $\mathcal{R}u \in L^2([0, 2\pi) \times \mathbb{R})$ 定义为 $\mathcal{R}u(s, \theta) = \mathcal{R}_\theta u(s)$。

下面描述相对于其而言，未知信号 $u^\dagger \in L^2(\mathcal{B}_1)$ 被假设为稀疏/可压缩的字典。我们考虑由具有紧支撑的适当二维小波 $(\phi_{j,n})_{(j,n)\in\Gamma}$ 构成的正交归一基。我们假设存在大小为 $|\Lambda| = M$ 的 $\Lambda \subset \Gamma$，使得 $\operatorname{supp} x^\dagger \subset \Lambda$（回顾 $x^\dagger = \Phi u^\dagger$）——因此 $\|P_\Lambda^\perp x^\dagger\|_2 = 0$。特别地，自然可以假设 $\Lambda = \Lambda_{j_0} = \{(j, n) \in \Gamma :$ 对于某个 $j_0 \in \mathbb{N}$，有 $j \le j_0\}$，即 $u^\dagger$ 属于与小波基相关的多尺度分解中具有有限最细尺度 j_0 的信号子空间。

我们的目标是为 $u^\dagger$ 提供恢复估计：给定 m 个对应 Radon 变换的样本，这些样本在不同角度 $\theta_1, \dots, \theta_m \in \mathbb{S}^1$ 下获得，角度从 $\mathbb{S}^1$ 中独立同分布地均匀随机抽取。这里我们仅考虑平行束模式，扇束采样模式将在第4节中讨论。

令 $W =$对角算子$(2^{j/2})_{j,n}$，u 的非线性恢复可通过求解以下最小化问题实现：

$$\min_u \|W^{-1}\Phi u\|_1 \quad : \quad \frac{1}{m}\sum_{k=1}^{m} \|\mathcal{R}_{\theta_k} u - y_k\|_{L^2(\mathbb{R})}^2 \le \beta^2, \tag{2}$$

其中，最小值是在张成空间 $\operatorname{span}(\phi_{j,n})_{(j,n)\in\Lambda_{j_0}}$ 上取得的，稀疏角度测量由 $y_k = \mathcal{R}_{\theta_k} u^\dagger + \varepsilon_k$ 给出，噪声 $\varepsilon \in (L^2(\mathbb{R}))^m$ 满足以下一致界$\|\varepsilon_k\|_{L^2(\mathbb{R})} \le \beta$，对每个 k 均成立。

我们关于稀疏 Radon 重建的主要结果（定理 4.5）的简化形式如下。

定理。在上述假设下，存在仅依赖于小波基的常数 C_0、C_1、$C_2 >$0，使得对于任意 $2 \le s \le |\Lambda_{j_0}|$，如果 $\widehat{u} \in$张成空间$(\phi_{j,\,n})_{(j,n)\in\Lambda_{j_0}}$ 是最小化问题 (2) 的一个解，且样本数量满足

$$m \ge C_0 s,$$

除对数因子外，则以下恢复估计以极高概率成立：

$$\|u^\dagger - \widehat{u}\|_{L^2} \leq C_1 2^{j_0/2} \frac{\sigma_s(W^{-1}\Phi u^\dagger)_1}{\sqrt{s}} + C_2 2^{j_0/2}\beta.$$

需要强调的是，使用权重可以使我们获得所需样本数更少的恢复估计，相较于将上述抽象结果应用于 Radon 设定时所需的样本数而言——见定理 4.4。

最后，我们简要阐述定理 4.6 中证明的结果。若将假设 $\|P_{\Lambda_{j_0}}^{\perp} x^\dagger\|_2 = 0$ 放宽为尾界 $\|P_{\Lambda_{j_0}}^{\perp} x^\dagger\|_2 \leq C2^{-aj_0}$（其中$a$、$C > 0$），且信号 $u^\dagger$ 满足可压缩性条件 $\sigma_s(\Phi u^\dagger)_1 \leq Cs^{1/2-p}$（其中 $p > 0$），则恢复估计可简化为

$$\|\widehat{u} - u^\dagger\|_{L^2} \leq c\beta^{\frac{2a}{2a+1}}, \qquad \|u^\dagger - \widehat{u}\|_{L^2} \leq c\left(\frac{1}{m}\right)^{\frac{ap}{a+p}},$$

精确到对数因子。这些估计成立的前提是噪声为确定性的。若假设噪声为随机变量，在自然假设下（例如，见 [21]），则有可能得到更精细的期望估计。

对于相对于多分辨率字典具有已知正则性的信号类，这一结果尤其有用。例如，已知小波空间中的此类稀疏性约束 $a = p = 1/2$ 对应于卡通类图像，即除沿曲线 [47, 75, 80]存在轻微不连续性外，由二维分段光滑函数构成的图像。因此，通过上述稀疏极小化方法，利用不同角度下 Radon 变换的 $m \asymp \beta^{-2}$ 含噪样本重建卡通类图像 $u^\dagger$是可行的，其误差为 $\|\widehat{u} - u^\dagger\|_{L^2} \leq C\beta^{1/2}$。

1.3. 相关研究。下面简要讨论文献中与本文工作具有一些共同特征的其他研究。

关于在无限维空间中进行采样和重建并获得恢复保证的方法，文献中已有若干研究。广义采样框架是在 [4, 10] 中发展起来的，旨在为可分希尔伯特空间提供保证。在这一框架下，给定任意框架中的元素与某个有限维子空间元素的有限个内积，可以恢复该有限维子空间中的元素。通过引入平衡性质 [8], ，解决了与重建稳定性相关的问题；当测量由酉算子建模时，这一性质尤其适用。已有多项研究在不同情形下针对各种非线性问题考察了这一问题——关于这一点，我们提及 [13, 63, 15, 11] 。

本文的一项主要结果（定理 3.5）在许多方面推广了此前已知的有限维 [32] 和无限维 [64, 8, 5, 14, 1]基于相干性的压缩感知结果。事实上，我们的理论能够处理线性不适定问题，这类问题通常涉及具有各向异性且可能取值为向量的紧算子

测量。由于我们的框架基于对采样/测量映射的一个非常宽泛的定义，能够同时涵盖标量积和逐点求值，因此这一结果也与逼近理论的研究范围相交，尤其涉及多项式稀疏逼近 [91] 和高维函数稀疏逼近 [3, 26, 45, 2]。

帮助我们处理前向映射的病态性和测量几何结构的核心概念是广义受限等距性质（g-RIP）（见第5.3节）。它是经典受限等距性质（RIP）的扩展。由于 RIP 在压缩感知中具有根本重要性，因此已得到广泛研究——例如参见 [39, 38, 33]。文献中还提出了许多相关变体。特别是，D-RIP [36, 74], 考虑了字典中可能存在的冗余；RIPL [16], 适用于具有分层稀疏性的信号；G-RIPL [1], 则适用于处理无限维信号。本文所考虑的 g-RIP 对应于不含分层、具有加权稀疏性的 G-RIPL。矩阵 G 在本文中发挥的作用略有不同，本文的主要目标是考虑逆问题的病态性。文献中还提出了压缩感知中受限等距性质的其他替代方法——例如参见 [19, 37]。文献 [59] 对受限等距性质与较弱的源条件概念进行了有趣的比较；不过，在实际情形中，验证是否满足源条件可能较为困难。

这里采用广义限制等距性质，以利用一种近似可对角化条件；在许多具有实际应用价值的情形中，这一条件都能满足——参见准可对角化的定义 3.7。事实上，文献中已经出现过类似的方法，其中一个典型例子是 [46] 引入并由 [35, 41] 进一步发展的、适用于更一般字典的小波-模糊分解（WVD）。还应指出， [64] 中的作者利用与前向映射的小波-模糊分解相关的对角化性质来获得恢复估计，尽管其设定和结果在若干方面与我们的方法有所不同。此外， [64] 所作的精确可对角化假设对于大多数具有实际应用价值的情形而言过于严格，尤其是对于本文讨论的稀疏 Radon 变换。

通过利用一种针对稀疏化字典的多尺度结构 $(\phi_{j,\ n})_{(j,n)\in\Gamma}$ 量身定制的跨尺度相干性概念，我们得以在定理 3.11 中推导出更精细的估计——见假设 3.10。文献中也有采用相干性多层信息的其他方法。例如，一些关于磁共振成像（MRI）的研究已经观察到非均匀的相干性模式 [76, 78, 79], ，并将其与多层采样方案结合起来 [96, 1, 5, 77]。我们采用了一种经过修改的相干性概念——见注 3.16——通过考虑前向映射的奇异值以某种方式衰减所量化的病态性，拓展了这些方法。

我们的抽象框架的主要应用——稀疏 Radon 变换的求逆——是 IP 社群众所周知的一项挑战。文献 [61], 中探讨了一种基于 Besov 范数惩罚、以小波为基础的促进稀疏性的方法，并给出了数值结果，显示其效果显著优于滤波反投影（FBP）和 Tikhonov 正则化等较经典的方法。文献中也使用了其他类小波字典（例如剪切波）来解决从稀疏数据中重建断层成像信号的问题——见 [30, 88, 28]。文献 [49] 中的重建是通过沿频域中的直线进行采样实现的，所采样的信号由狄拉克 δ 函数的线性组合给出。在适当的假设下，研究证明，精确重建所需的样本数与 δ 函数的数量成正比。文献 [70, 68], 开展了一项数值研究，以探究对于断层图像，准确重建所需的样本数与其稀疏度之间的关系；这里的稀疏度指非零像素的数量。数值实验表明，准确重建所需的标量采样值数量与信号的稀疏度成正比。

最后，我们提及一种替代方法，可为逆问题推导出与本文所述思路相似的恢复保证。这种方法由统计逆向学习理论提供，该领域位于逆问题与统计学习的交界处。简而言之，问题的表述如下。设 $\mathcal{H}_1$ 和$\mathcal{H}_2$ 为希尔伯特空间， $\mathcal{H}_2$ 由定义在某个输入空间 $(\mathcal{X}, \nu)$ 上的函数构成，其中 ν是 $\mathcal{X}$上的概率分布。设 $F\colon\ \mathcal{H}_1 \to \mathcal{H}_2$ 为线性算子。给定 $u \in \mathcal{H}_1$，若 X_1，…， $X_m \overset{\text{独立同分布}}{\sim} \nu$，则考虑

$$Y_k = Fu(X_k) + \varepsilon_k, \qquad k = 1, \dots, m,$$

其中 ε_k 是相互独立的噪声随机变量。目标是根据观测值 Y_1，…， Y_m恢复 u 。情形 $F = \mathrm{Id}$ 归结为经典的非参数回归问题。例如， [21, 89, 90, 57]中给出了该情形下的重建估计。

1.4. 讨论与开放问题。下面我们讨论该方法的一些局限性，以及未来研究的潜在方向。

- 本文提出的方法，以及大多数关于压缩感知的文献（关于特定模型的一些例外情况，见 [85, 22, 23, 56, 18]），都只适用于处理线性问题，因此仅适用于线性逆问题。非线性逆问题出现在许多领域，尤其是偏微分方程中的参数识别问题 [73, 65]。典型例子包括用于电阻抗断层成像的 Calder´on 问题 [24, 97] 以及逆散射问题 [42, 31]。发展非线性压缩感知理论以处理这些非线性问题，将会很有意义。

- 我们考虑固定的确定性噪声 $\varepsilon \in \mathcal{H}_2^m$，这代表最坏情况。统计反问题 [72, 95] 和统计逆向学习理论 [21, 57] 中常用的一种模型考虑统计噪声，即每个 ε_k 都是按照定义域 $\mathcal{H}_2$上的某个概率分布独立同分布地采样得到的。理解这种不同的设定会如何影响我们的方法和重构估计，是一个引人入胜的问题。特别是，由于我们的方法允许采用变密度采样策略，研究一类概率分布的最优性速率将会很有意义。

- 在这项工作中，我们假设 $(\psi_k)_{k\in\mathbb{N}}$ 和 $(\phi_i)_{i\in\Gamma}$ 是正交归一基。我们预计，整个方法也适用于较宽松的假设，例如，$(\phi_i)_{i\in\Gamma}$ 是一个框架和/或 Riesz基，如 [87, 14, 26] 中所示。此类推广将使我们能够考虑比小波正交归一基更一般的字典，例如曲波或剪切波；已知它们能为非线性逼近误差 [34, 60, 75]提供更好的收敛率。因此，这有望带来更好的样本复杂度估计。我们将这一推广留待未来研究。

- 在我们考虑的稀疏 Radon 变换模型中，角度数量有限 θ_k ，但对于每个固定角度，我们都会对每个平移 $s \in [-1, 1]$处的 Radon 变换进行采样。因此，这是一个测量的半离散模型，自然会让人想知道：若也对变量 s进行采样，是否仍能得到相同或类似的估计？文献 [49]包含了这方面的一些结果，但其采样是在频域进行的，且研究的是 Dirac δ脉冲情形。

- 我们选择仅展示逆问题抽象样本复杂度结果的一项应用（稀疏 Radon 变换）。不过，所提出的方法足够灵活，能够涵盖若干经典问题，包括从信号在傅里叶域中的样本恢复其小波系数，以及有限样本的反卷积问题。据我们所知，这是首次通过专门化同一抽象框架，一次性处理性质迥异的逆问题（即涉及“离散”和“连续”样本的逆问题）。我们决定将这一分析留待另一篇论文。

1.5. 论文结构。下面简要介绍本文的内容安排。第2节详细阐述了构建框架并陈述主要结果所需的预备概念。在此基础上，第3节给出了关于稀疏性约束下病态逆问题样本复杂度的一般性结果。本文的主要应用是稀疏角度层析成像问题，第4节对此进行了详尽讨论。主要结果和辅助结果的所有证明

均见第5节，但一项关键技术结果（定理 5.7）的证明除外，该证明见附录 A。

2. Setup

2.1. 加权设置与记号。令 $\mathbb{N}$ 表示正整数集合。对于 $n \in \mathbb{N}$，我们用 $[n]$ 表示不超过 n 的正整数集合，即 $[n] = \{1, \dots, n\}$。

令 Γ 为有限或可数指标集，例如 $\Gamma = \mathbb{N}$ 或 $\Gamma \subseteq \mathbb{N} \times \mathbb{Z}$。在我们的设定中，它是表示未知信号的字典的指标集。我们还将考虑指标的一个有限子集 $\Lambda \subseteq \Gamma$，它表示进行信号重建的指标集。我们用 P_Λ 表示 $\ell^2(\Gamma)$ 上的正交投影，其定义为

$$(P_\Lambda x)_i = \begin{cases} x_i & (i \in \Lambda) \\ 0 & (i \notin \Lambda) \end{cases}.$$

P_Λ 的像因此为 $\ell^2_\Lambda(\Gamma) := \operatorname{span}\{e_i : i \in \Lambda\}$, $(e_i)_{i\in\Gamma}$ 是 $\ell^2(\Gamma)$ 的标准基。为简化记号，我们常将 $\ell^2_\Lambda(\Gamma)$ 与 $\ell^2(\Lambda)$ 或 $\mathbb{C}^M$ 等同，其中 $M = |\Lambda|$。相应的伴随映射记为 ι_Λ，即典范嵌入 $\ell^2_\Lambda(\Gamma) \to \ell^2(\Gamma)$。我们还设 $P^\perp_\Lambda := I - P_\Lambda$，其中 I 为恒等算子。

如果对每个 $i \in \Gamma$ 都有 $\omega_i \geq 1$，我们就称序列 $\omega \in \mathbb{R}^\Gamma$ 为权重向量。集合 $S \subseteq \Gamma$ 的加权大小定义为

$$\omega(S) := \sum_{i\in S} \omega_i^2,$$

只要 S 有限即可。用 $|S|$ 表示集合 S 的基数，以下不等式可立即验证：

$$|S| \leq \omega(S) \leq |S| \sup_{i\in S} \omega_i^2.$$

对于 $0 < p \leq 2$，我们引入集合

$$\ell^p_\omega(\Gamma) := \{x \in \mathbb{C}^\Gamma : \|x\|_{p,\omega} < \infty\},$$

其中 ω-加权 ℓ^p 范数定义为

$$\|x\|_{p,\omega} := \left(\sum_{i\in\Gamma} |x_i|^p \omega_i^{2-p} \right)^{1/p}.$$

极限情形 $p \to 0$ 引出了加权稀疏性的概念。更确切地说，将 $x \in \mathbb{C}^\Gamma$ 的支撑记为 $\operatorname{supp}(x) = \{i \in \Gamma\colon\ x_i \neq 0\}$，若 x 是 s-ω-稀疏的，则

$$\|x\|_{0,\omega} := \omega(\operatorname{supp} x) \leq s.$$

注意，在非加权情形（即对所有 $i \in \Gamma$， $\omega_i = 1$ ）下，我们得到通常的 ℓ^p范数 $\|x\|_p$ 以及支撑集基数 $\|x\|_0 = |\operatorname{supp} x|$。

我们还引入了关于加权 ℓ^p 范数的 $x \in \ell^2(\Gamma)$ 的最佳 ω-加权 s-稀疏逼近误差：

$$\sigma_s(x)_{p,\omega} := \inf\{\|x - y\|_{p,\omega} : y \in \mathbb{C}^\Gamma, \|y\|_{0,\omega} \le s\}.$$

等价地，我们有

$$\sigma_s(x)_{p,\omega} = \inf\{\|x_{S^c}\|_{p,\omega} \colon S \subset \Gamma,\ \omega(S) \le s\},$$

其中 $x_S := P_S x$ 且 $S^c = \Gamma \setminus S$。若 $S \subset \Gamma$ 是一个满足 $\sigma_s(x)_{p,\,\omega} = \|x_{\tilde{S}^c}\|_{p,\,\omega}$ 且 $\omega(S) \le s$ 的集合，则称 $x_{\tilde{S}}$ 是关于 ℓ^p_ω 的 x 的最佳 s-ω-稀疏逼近。注意，通常 $x_{\tilde{S}}$ 并不唯一。

2.2. 字典、测量与最小化问题。下文出现的希尔伯特空间均假设为复可分空间。

希尔伯特空间。设 $\mathcal{H}_1$ 和 $\mathcal{H}_2$ 为希尔伯特空间。我们将 $\mathcal{H}_1$ 称为信号空间，将 $\mathcal{H}_2$ 称为测量空间。

字典。设 $(\phi_i)_{i\in\Gamma}$ 是 $\mathcal{H}_1$ 的一个正交归一基。令 $\Phi\colon \mathcal{H}_1 \to \ell^2(\Gamma)$ 为相应的分析算子，即 $\Phi u := (\langle u, \phi_i\rangle_{\mathcal{H}_1})_{i\in\Gamma}$。相应的合成算子 $\Phi^*\colon \ell^2(\Gamma) \to \mathcal{H}_1$ 由 $\Phi^* x = \sum_{i\in\Gamma} x_i \phi_i$ 给出。

测量空间。设 $(\mathcal{D}, \mu)$ 为一个测度空间。我们将 $\mathcal{D}$ 视为与测量相关的参数空间。实际上，我们相对于 $\mathcal{D}$ 上的概率测度 ν 进行随机采样，该概率测度相对于 μ 绝对连续——也就是说，存在一个正函数 $f_\nu \in L^1(\mu)$，使得 $\|f_\nu\|_{L^1} = 1$，且 $\mathrm{d}\nu := f_\nu \,\mathrm{d}\mu$。事实证明，为了获得最优的样本复杂度，概率分布的选择至关重要；然而，就本文的目的而言，它并不十分重要。我们将在未来的工作中进一步研究这一问题。

测量算子和前向映射。设 $(F_t)_{t\in\mathcal{D}}\colon \mathcal{H}_1 \to \mathcal{H}_2$ 是一族一致有界的线性映射，其中对每个 $t \in \mathcal{D}$，均有 $\|F_t\| \le C_F$，用于对我们问题中的测量算子建模。我们假设

$$\mathcal{D} \ni t \mapsto F_t u \in \mathcal{H}_2$$

对每个 $u \in \mathcal{H}_1$，都属于博赫纳空间 $L^2_\mu(\mathcal{D}; \mathcal{H}_2)$。我们定义前向映射

$$F\colon \mathcal{H}_1 \to L^2_\mu(\mathcal{D}; \mathcal{H}_2), \qquad (Fu)(t) = F_t u,$$

对于几乎所有的 $t \in \mathcal{D}$。我们假设 F 有界，且 $\|F\|_{\mathcal{H}_1 \to L^2_\mu(\mathcal{D};\ \mathcal{H}_2)} \le C_F$。

细心的读者会注意到，假设测量算子（非一致）有界，只会影响涉及截断误差的估计——参见命题 5.8。尽管如此，由于下文讨论的应用情境满足一致有界性假设，我们宁愿牺牲一些一般性来换取一致性。

噪声。令 $u^\dagger \in \mathcal{H}_1$ 表示未知信号，若 t_1、...、t_m 是来自 $\mathcal{D}$ 上分布 ν 的独立同分布样本，则含噪测量可写为

$$y_k := F_{t_k} u^\dagger + \varepsilon_k, \quad k = 1, \dots, m,$$

其中 $\varepsilon_k \in \mathcal{H}_2$。我们假设噪声满足以下一致有界条件：

$$\max_k \|\varepsilon_k\|_{\mathcal{H}_2} \leq \beta, \tag{3}$$

其中某个 $\beta \geq 0$。

截断误差。设 $\Lambda \subseteq \Gamma$ 为有限值。我们假设存在 $r \geq 0$，使得 $\|P_\Lambda^\perp x^\dagger\|_2 \leq r$，其中 $x^\dagger = \Phi u^\dagger$。截断误差 $\|P_\Lambda^\perp x^\dagger\|_2$ 通常称为线性逼近误差 [44]。下文将说明（见例 4.7），若 $u^\dagger$ 和 Φ 具有某种意义下的充分正则性（具体将在后文说明），则 $\|P_\Lambda^\perp x^\dagger\|_2$ 的衰减率可用 $M = |\Lambda|$ 明确表示。

最小化问题。我们考虑以下最小化问题：

$$\min_{x \in \ell^2(\Lambda)} \|x\|_{1,\omega} \quad : \quad \frac{1}{m} \sum_{k=1}^m \|F_{t_k} \Phi^* \iota_\Lambda x - y_k\|_{\mathcal{H}_2}^2 \leq \eta^2, \tag{4}$$

其中 $\eta \geq 0$ 为某个值；回顾一下，$\iota_\Lambda = (P_\Lambda)^*$。由于该问题是在二次约束下对凸泛函进行最小化，因此至少存在一个解——除非约束条件为空。

备注 2.1。文献中将 ℓ^1 最小化问题 (4) 称为综合形式。另一种方法称为分析形式

$$\min_u \|\Phi u\|_1 \quad : \quad \frac{1}{m} \sum_{k=1}^m \|F_{t_k} u - y_k\|_{\mathcal{H}_2}^2 \leq \eta^2,$$

其中，最小值是在 $\operatorname{span}(\phi_i)_{i \in \Lambda}$ 上取得的。

对于一般字典，这两种方法并不等价——例如参见[50]。不过，对于正交归一基，等价性成立，而本文讨论的正是这种情形。事实上，Φ 及其逆映射 Φ^* 是 $\mathcal{H}_1$ 与ℓ^2的等距同构（Γ）。本节定理的表述采用综合方法，因为这种表述在压缩感知和受限等距性质的背景下更为自然——参见第5节。相较之下，第4节采用分析方法，这更符合逆问题的理念。

两个例子。现在我们简要讨论两个符合我们框架的问题：从傅里叶测量中重建稀疏信号的小波系数，以及根据有限次逐点求值，从多项式基中恢复稀疏函数。尽管这些问题已得到深入研究——例如参见 [54, 第12节1]——我们的框架在许多方面推广了有界正交系统中的压缩感知，包括引入可能适定性不佳的前向映射，以及采用更一般的向量值测量算子。

示例 2.2（从傅里叶样本重建小波系数）。我们考虑这样一个问题：通过在适当的带宽$[N]_\pm$中采样信号的傅里叶系数，重建支撑于 $\mathbb{T} = [0,1]$ 中的信号的小波系数 $= \{-N, -N+1, \dots, N-1, N\}$；感兴趣的读者可参阅 [38, 79, 78]。

在此背景下，我们有 $\mathcal{H}_1 = L^2(\mathbb{T})$是信号空间，而 $\mathcal{H}_2 = \mathbb{C}$ 是标量测量空间。我们考虑一个正交归一基$(\phi_i)_{i\in\Gamma} = (\phi_{j,\ n})_{(j,n)\in\Gamma}$，由边界条件下的([40])或周期性的([81, 第3节。11])Daubechies小波构成。这里的指标集 $\Gamma \subset \mathbb{N}\times\mathbb{Z}$ 包含所有可能的尺度，其索引为 $j \in \mathbb{N}$，以及平移参数的所有可能取值，其索引为 $n \in \mathbb{Z}$。一般而言， n 的取值范围取决于尺度 j。

作为该模型的测量空间，我们考虑 $(\mathcal{D}, \mu) = ([N]_\pm, \mathfrak{c})$，其中 $\mathfrak{c}$ 是计数测度。前向映射 $F\colon L^2(\mathbb{T}) \to \mathbb{C}^{2N+1}$ 定义为 $F = P_N\mathcal{F}$，其中 $\mathcal{F}\colon L^2(\mathbb{T}) \to \ell^2(\mathbb{Z})$ 是傅里叶变换

$$\mathcal{F}u(t) = \int_{\mathbb{T}} u(x)e^{-2\pi i x t}\,\mathrm{d}x, \qquad t \in \mathbb{Z},$$

而 $P_N\colon \ell^2(\mathbb{Z}) \to \mathbb{C}^{2N+1}$ 是投影到由 $[N]_\pm$ 索引的分量上的投影。对于每个 $t \in [N]_\pm$，测量算子 $F_t\colon L^2(\mathbb{T}) \to \mathbb{C}$ 定义为 $F_t u = \mathcal{F}u(t) = Fu(t)$。我们选择 $f_\nu(0) := 1/C_\nu$ 和 $f_\nu(t) := 1/(|t|C_\nu)$（适用于 $t \neq 0$），其中 $C_\nu := 1 + \sum_{t=1}^{N} 2/t$ 是归一化常数；关于这一选择的进一步说明，参见注 3.2 中的 (i)。

截断带宽 $\Lambda \subset \Gamma$ 是指标 (j,n) 的子集，满足 $j \le j_0$。换言之，我们用一个具有有限最细尺度j_0的信号来近似原始信号。

例 2.3（用多项式近似）。我们考虑经典问题：给定有限数量的逐点样本 [91]，重构一个次数至多为 M 的多项式，该多项式相对于多项式的正交归一基是稀疏的。

设 $L^2_w(-1,1)$ 为关于权函数 $w\colon [-1,1] \to \mathbb{R}_+$ 的平方可积函数空间，其中 $w\mathrm{d}x$ 是概率测度。设 $(p_i)_{i\in\mathbb{N}}$ 为 $L^2_w(-1,1)$ 的多项式正交归一基。我们固定字典 $(\phi_i)_{i\in\Gamma} = (p_i)_{i\in\Gamma}, \Gamma = [M+1]$,

并将次数至多为 M 的多项式空间定义为 $\mathcal{P}_M := \operatorname{span}(p_i)_{i\in\Gamma} \subset L^2_w(-1,1)$。所考虑的模型由 $\mathcal{H}_1 = \mathcal{P}_M$ 和 $\mathcal{H}_2 = \mathbb{C}$ 表示。

测量空间为 $(\mathcal{D},\mu) = ([-1,1], w\mathrm{d}x)$。前向映射与恒等映射 $(F = I)$ 一致，并且对于每个 $t \in [-1,1]$，测量算子 $F_t\colon \mathcal{P}_M \to \mathbb{C}$ 由 $F_t u = u(t)$ 给出。我们选择 $\Lambda = \Gamma$ 和 $f_\nu \equiv 1$，使得 $\nu = \mu$。

3. A general result on sample complexity

3.1. 主要结果。为了陈述主要结果，我们需要引入一些假设。

假设 3.1。满足以下相干性界：

$$\|F_t\phi_i\|_{\mathcal{H}_2} \le B\sqrt{f_\nu(t)}\,\omega_i, \quad t \in \mathcal{D},\ i \in \Gamma,$$

对某个 $B \ge 1$，其中 f_ν 在第 2.2 节中定义，且 $\omega_i \ge 1$。

备注 3.2。我们来讨论这一界在前面讨论过的两个例子中的出现情况。

(i) 考虑例 2.2 中讨论的傅里叶-小波感知问题。根据 [67, 定理 2.1]，已知在对小波字典的尺度函数施加适当正则性假设的条件下，以下估计成立：

$$|\mathcal{F}\phi_{j,n}(0)| \le C, \quad (j,n) \in \Gamma,$$
$$|\mathcal{F}\phi_{j,n}(t)| \le \frac{C}{\sqrt{|t|}}, \quad t \in \mathbb{Z}\setminus\{0\},\ (j,n) \in \Gamma$$

对某个 $C \ge 1$。令 $C_\nu := 1 + \sum_{t=1}^{N} 2/t$，$f_\nu(0) = 1/C_\nu$ 且 $f_\nu(t) = 1/(|t|C_\nu)$，其中 $t \ne 0$ 且 $B := C\sqrt{C_\nu}$，则上述估计可改写为假设 3.1 的形式，其中 $\omega_i \equiv 1$。

(ii) 考虑例 2.3 中的问题。对于在 $[-1,1]$ 上使用勒让德多项式 $(p_i)_{i\in\mathbb{N}}$ 进行逼近的情形（参见 [91] 或 [98, 第 504] 页），我们有

$$\|p_i\|_\infty \le \sqrt{2i-1}, \quad i \in \mathbb{N}.$$

因此，只要权重的选取符合约束条件 $\omega_i \ge \sqrt{2i-1}$，假设 3.1 就能成立。

假设3.3. 对于某个 $c_\nu > 0$，满足以下估计：

$$c_\nu \le f_\nu \le 1.$$

注 3.4。这个不等式是一项相当技术性的假设；正如定理 3.5 的证明所示，它与这样一个事实有关：我们假设样本的噪声水平只有一个统一上界，见式（3）。通过提高对噪声的

要求，可以去掉这一假设，未来的工作将对此加以说明。本文不讨论这一扩展，因为它与本文研究的稀疏 Radon 问题无关。

若 A 是矩阵或算子，则用 $\|A\|$ 表示其算子范数。

定理 3.5。考虑第 2.2 节的设定。假设 3.1 和 3.3 均成立。令 $M = |\Lambda|$，并定义如下 $M \times M$ 矩阵：

$$G := \sqrt{P_\Lambda \Phi F^* F \Phi^* \iota_\Lambda}.$$

我们假设 G 可逆。

设 $x^\dagger \in \ell^2(\Gamma)$ 满足 $\|P_\Lambda^\perp x^\dagger\|_2 \le r$。考虑从概率分布 ν 中抽取的 m 个独立同分布样本 $t_1, \ldots\ldots, t_m \in \mathcal{D}$。令 $y_k := F_{t_k} \Phi^* x^\dagger + \varepsilon_k$，其中 $k = 1, \ldots\ldots, m$，且 $\max_k \|\varepsilon_k\|_{\mathcal{H}_2} \le \beta$。

存在绝对常数 C_0、C_1、$C_2 > 0$，使得以下结论成立。设 $\widehat{x} \in \ell^2(\Lambda)$ 是该最小化问题的一个解

$$\min_{x \in \ell^2(\Lambda)} \|x\|_{1,\omega} \quad : \quad \frac{1}{m} \sum_{k=1}^{m} \|F_{t_k} \Phi^* \iota_\Lambda x - y_k\|_{\mathcal{H}_2}^2 \le \left(\beta + C_3 \|G^{-1}\|^{-1} r\right)^2. \tag{5}$$

固定 $s \in [M]$ 取 $s \ge \|\omega\|_\infty^2 / 4$，并设 $\tau := B^2 \|G^{-1}\|^4 \|G\|^2 s$。对于任意 $\gamma \in (0, 1)$，若

$$m \ge C_0 \tau \max\{\log^3 \tau \log M, \log(1/\gamma)\},$$

则以超过 $1 - \gamma$ 的概率，以下恢复估计成立：

$$\|x^\dagger - \widehat{x}\|_{1,\omega} \le C_1 \sigma_s(P_\Lambda x^\dagger)_{1,\omega} + C_2 \sqrt{s} c_\nu^{-1/2} (\|G^{-1}\| \beta + C_3 r) + \|P_\Lambda^\perp x^\dagger\|_{1,\omega},$$

$$\|x^\dagger - \widehat{x}\|_2 \le C_1 \frac{\sigma_s(P_\Lambda x^\dagger)_{1,\omega}}{\sqrt{s}} + C_2 c_\nu^{-1/2} (\|G^{-1}\| \beta + C_3 r),$$

其中 $C_3 := c_\nu^{-1/2} C_4 (\|F \Phi^* \iota_{\Gamma \setminus \Lambda}\| \|G^{-1}\| + 1)$，且 C_4 仅取决于 C_F。

定理 3.5 给出了非一致类型的恢复保证，即对于固定的未知信号，以极高概率成立。尽管如此，证明中使用的所有工具本质上都适用于一致恢复，唯独命题 5.8 例外；该命题是获得截断误差精细估计所必需的。通过在估计中用可能更差的界 $\|G^{-1}\|(\beta + r)$ 替换项 $\|G^{-1}\| \beta + C_3 r$，可以得到一致恢复结果（即对一整类满足适当稀疏性条件的信号均成立）。特别地，若 $x^\dagger$ 的支撑包含于 Λ，从而 $r = 0$，则这两项相等，进而为这类信号得到一致恢复估计。

估计中出现的对数因子 $\log^3 \tau$ 已知并非最优——例如参见 [26]。优化对数的指数将大幅增加证明的篇幅，这超出了本文工作的范围。

为了将这一结果应用于具体的相关情形，必须明确计算若干量，其中最关键的是 $\|G^{-1}\|$。这将需要新的工具，并将成为 §3.2以下内容的重点；这里先介绍一些一般性观察。

对于 G 的可逆性，一个通常在若干值得关注的逆问题中都能满足的充分条件，是 F 具有如下所述的受限单射性。

定义 3.6。设 $\mathcal{H}$ 是一个复可分希尔伯特空间。若线性算子$V\colon \ell^2(\Gamma) \to \mathcal{H}$ 满足：对于每个有限指标族 $\Lambda \subset \Gamma$， V 在 $\ell^2(\Lambda)$ 上的限制是单射的，则称其具有有限基单射（FBI）性质。

给定希尔伯特空间 $\mathcal{H}_1$ 的一组正交归一基 $(\phi_i)_{i\in\Gamma}$，其分析算子为Φ。若算子 $F\colon \mathcal{H}_1 \to \mathcal{H}$ 关于 Φ 具有 FBI 性质（简称： Φ-FBI），则称其具有 FBI 性质。

假设 F 是 Φ-FBI。那么，对于每个有限$\Lambda \subset \Gamma$， $F\Phi^* \iota_\Lambda$ 都是单射，这意味着 $G \in \mathbb{C}^{M\times M}$ 是正定的，因而可逆。

正如 [25], 中已经指出的，在带有稀疏性约束的逆问题中，FBI 条件相当自然——事实上，它的限制比完全单射更弱，尽管后者在逆问题文献中更为常见。$\ell^2(\mathbb{N})$ 上一个非单射 FBI 算子的简单例子如下：

$$Ve_i = \begin{cases} \sum_j c_j e_j & (i=1) \\ e_{i-1} & (i \geq 2), \end{cases}$$

其中 $(e_i)_{i\in\mathbb{N}}$ 是 $\ell^2(\mathbb{N})$ 的标准基，而 $(c_j)_{j\in\mathbb{N}}$ 是任意复数序列，使得 $\sum_j |c_j|^2 < \infty$，并且对每个 j 都有 $c_j \neq 0$。

为了应用定理 3.5，还需要显式计算另一个量 $C_3 = c_\nu^{-1/2} C_4($ $\|F\Phi^* \iota_{\Gamma\setminus\Lambda}\| \|G^{-1}\| + 1)$。我们发现，在合理的条件下（见定理 3.11），量 C_3 可以由一个仅取决于 C_F、 c_ν 和准对角化界的常数来界定。

3.2. $\|G^{-1}\|$ 的定量界与拟对角化。我们已经证明，若 F 是 Φ-FBI，则矩阵 G 非退化。为了使定理 3.5 的估计具有可操作性，我们现在推导 $\|G^{-1}\|$的显式界。在无限维压缩感知的文献中，这类界是借助所谓的平衡性质 [8], 得到的；当测量由酉算子建模时，这种方法尤其适用。这里，我们在一定程度上利用了前向映射可能存在的病态性，以及下文将介绍的一种准对角化性质。许多我们感兴趣的前向映射（包括 Radon 变换）都满足这一性质。随后，我们推导出一些特定界，其特点与利用平衡性质得到的界不同。将平衡性质与准对角化性质结合起来的可能方法留待未来研究。

我们注意到，可以利用 $F\Phi^*$的奇异值对 $\|G^{-1}\|$ 进行显式估计。事实上，我们有

$$\|G^{-1}\| \le \sigma_M(F\Phi^*\iota_\Lambda)^{-1},$$

其中， $\sigma_M(F\Phi^*\iota_\Lambda)$ 是 $F\Phi^*\iota_\Lambda$ 的第 M 个奇异值，且 $M = |\Lambda|$。

拟对角化。前面的估计取决于 $F\Phi^*\iota_\Lambda$ 的奇异值。在实际情形中，估计 $\sigma_M(F\Phi^*\iota_\Lambda)$ 并非易事。因此，我们现在讨论一种估计 $\|G^{-1}\|$ 的替代方法，它是某个条件的推论，而该条件通常在多种情境下都成立——更准确地说，当能够找到一个字典，使其近似对角化模型的前向映射的作用时，该条件便成立。

现在我们考虑带有两个指标的（正交）多分辨率字典 $(\phi_{j,\ n})_{(j,n)\in\Gamma}$ （例如，正交小波）。参数 j 表示尺度，我们约定， j 的值越大，对应的尺度越精细。

定义 3.7. 如果存在 c、$C > 0$ 使得 F 满足关于 $(\Phi$、$b)$、$b \ge 0$ 的准对角化性质，则称 F 满足该性质。

$$c\sum_{(j,n)\in\Gamma} 2^{-2bj}|x_{j,n}|^2 \le \|F\Phi^*x\|^2_{L^2_\mu(\mathcal{D};\mathcal{H}_2)} \le C\sum_{(j,n)\in\Gamma} 2^{-2bj}|x_{j,n}|^2, \quad x \in \ell^2(\Gamma). \tag{6}$$

粗略地说，前述性质意味着，前向映射作用于字典 $(\phi_{j,\ n})_{(j,n)\in\Gamma}$ 中元素的结果，近似等同于一个对角算子的作用；该算子的二进系数仅依赖于参数 b 和“尺度”指标 j：一般而言， b 越大， F 的平滑性越强，因此相应逆问题的病态性也越严重。

现在我们给出一组确保准对角化性质成立的条件。回顾一下， $H^b(\mathbb{R}^d)$、$b \in \mathbb{R}$ 是广义函数 $u \in \mathcal{S}'(\mathbb{R}^d)$ 的 Sobole夫空间，其中 $(I-\Delta)^{b/2}u \in L^2(\mathbb{R}^d)$；而 $H^b(\mathbb{T}^d)$ 是广义函数 $u \in \mathcal{S}'(\mathbb{T}^d)$ 的 Sobole夫空间，其中 $(I-\Delta)^{b/2}u \in L^2(\mathbb{T}^d)$；这里 $(I-\Delta)^{b/2}$ 分别对应于 $\mathbb{R}^d$ 或 $\mathbb{Z}^d$ 上的傅里叶乘子 $(1+|\cdot|^2)^{b/2}$。

命题 3.8。设 $\mathcal{H}_1$ 为 $L^2(\mathbb{T}^d)$ 或 $L^2(\mathbb{R}^d)$。假设存在 $c, C > 0$, $b \ge 0$ ，使得：

- 前向映射 F 满足

$$c\|u\|^2_{H^{-b}} \le \|Fu\|^2_{L^2_\mu(\mathcal{D};\mathcal{H}_2)} \le C\|u\|^2_{H^{-b}}, \quad u \in \mathcal{H}_1. \tag{7}$$

- 字典 $(\phi_{j,\ n})_{(j,\ n)\in\Gamma}$ 是 $\mathcal{H}_1$ 的正交归一基，且

$$c\|u\|^2_{H^{-b}} \le \sum_{(j,n)\in\Gamma} 2^{-2bj}|\langle u, \phi_{j,n}\rangle_{L^2}|^2 \le C\|u\|^2_{H^{-b}}, \quad u \in \mathcal{H}_1, \tag{8}$$

我们将其称为字典的 Littlewood-Paley 性质。

那么， F 相对于（Φ、 b）满足准对角化性质（6）。

注 3.9. 若 $(\phi_{j,n})_{(j,n)\in\Gamma}$ 是一个 q-正则小波系统（依照 [81,第 2.2] 节的定义），那么对每个 $-q < b < q$，Littlewood-Paley 条件 (8) 都成立——例如参见 [81, 第 2.8 节中的定理 8 和第 3.11] 节中的 [80, 定理 9.2]。对于任意固定的 $q \in \mathbb{N}$，都存在紧支撑的 q-正则小波系统——参见 [81, 第 3.8] 节中的定理 3。

跨尺度相干性界。只要能获得针对字典 $(\phi_{j,n})_{(j,n)\in\Gamma}$ 多尺度结构的额外相干性界，定理 3.5 的结果在相对于 (Φ, b) 的拟对角化情形下便可得到改进。

特别地，我们考虑以下更严格版本的假设 3。 1.

假设 3.10。满足以下相干性界：

$$\|F_t\phi_{j,n}\|_{\mathcal{H}_2} \le B\frac{\sqrt{f_\nu(t)}}{d_{j,n}}\omega_{j,n}, \quad t \in \mathcal{D},\ (j,n) \in \Gamma.$$

对某些 $B \ge 1$, $1 \le d_{j,n} \le 2^{bj}$，其中 b 是准对角化 (6) 中出现的常数， f_ν 在第 2.2 节中定义，且 $\omega_{j,n} \ge 1$。

下文的注 3.16 将进一步阐明参数 $d_{j,n}$ 的作用。

我们可以利用这类更精细的相干性信息，大幅减少实现定理 3.5 中恢复所需的样本数量，同时保持相同的恢复速率。

对于 $j \in \mathbb{N}$，我们引入最细尺度为 j 的指标子集

$$\Lambda_j := \{(j', n) \in \Gamma :\ j' \le j\}. \tag{9}$$

此外，我们始终假设 $|\Lambda_j| < +\infty$。

定理 3.11。考虑第 2.2 节中的设定。设 $(\phi_{j,\ n})_{(j,n)\in\Gamma}$ 是满足 (6) 的多分辨率（正交归一）字典。假设 3.10 和 3.3 成立。固定 $j_0 \in \mathbb{N}$，并令 $M := |\Lambda_{j_0}|$。

设 $x^\dagger \in \ell^2(\Gamma)$ 满足 $\|P_{\Lambda_{j_0}}^{\perp} x^\dagger\| \le r$。考虑从概率分布 ν 中抽取的 m 个独立同分布样本 t_1, ..., $t_m \in \mathcal{D}$。令 $y_k := F_{t_k}\Phi^* x^\dagger + \varepsilon_k$，其中 $k = 1$, ..., m，且 $\max_k \|\varepsilon_k\|_{\mathcal{H}_2} \le \beta$。

存在常数 C_0、 C_1、 C_2、 $C_4 > 0$，它们仅依赖于 (6) 中的准对角化界以及 C_F，使得以下结论成立。令 $\zeta \in [0,1]$，并令 $W := \mathrm{diag}(2^{bj})_{(j,n)\in\Lambda_{j_0}}$。设 $\widehat{x}$ 是以下最小化问题的一个解：

$$\min_{x\in\ell^2(\Lambda_{j_0})} \|W^{-\zeta}x\|_{1,\omega} : \quad \frac{1}{m}\sum_{k=1}^{m} \|F_{t_k}\Phi^*\iota_{\Lambda_{j_0}}x - y_k\|_{\mathcal{H}_2}^2 \le \left(\beta + C_3 2^{-bj_0} r\right)^2, \tag{10}$$

其中 $C_3 = c_\nu^{-1/2} C_4$。

固定 $s \in [M]$，令 $s \geq \|\omega\|_\infty^2/4$，并设

$$\tau := B^2 \max_{(j,n)\in\Lambda_{j_0}} (d_{j,n}^{-2} 2^{2bj}) 2^{2(1-\zeta)bj_0} s.$$

If

$$m \geq C_0 \tau \max\{\log^3 \tau \log M, \log(1/\gamma)\}, \tag{11}$$

那么，以大于 $1-\gamma$ 的概率，以下恢复估计成立：

$$\|W^{-\zeta}x^\dagger - W^{-\zeta}\widehat{x}\|_2 \leq C_1 \frac{\sigma_s(P_{\Lambda_{j_0}} W^{-\zeta}x^\dagger)_{1,\omega}}{\sqrt{s}} + C_2 2^{-\zeta bj_0} c_\nu^{-1/2} (2^{bj_0}\beta + C_3 r).$$

特别地，我们有

$$\|x^\dagger - \widehat{x}\|_2 \leq C_1 2^{\zeta bj_0} \frac{\sigma_s(P_{\Lambda_{j_0}} W^{-\zeta}x^\dagger)_{1,\omega}}{\sqrt{s}} + C_2 c_\nu^{-1/2} \left(2^{bj_0}\beta + C_3 r\right).$$

注 3.12. 当 $(\phi_{j,n})_{j,n}$ 是类小波字典时，定理 3.11 中的参数 $M = |\Lambda_{j_0}|$ 通常满足 $M \asymp 2^{cj_0}$，其中 c 是某个正整数——例如，参见下文公式（15）。在这些情况下，样本复杂度中的量 $\log M$ 因此可以替换为 cj_0。

注 3.13. 得益于准对角化性质 (6)，定理 3.11 中与矩阵 G 相关的量具有显式形式。此外，得益于假设 3.10 给出的关于相干性的精细信息，以及参数 ζ 带来的灵活性，定理 3.11 中的样本复杂度相较于定理 3.5 得到了改进；事实上，在前一种情形中，样本复杂度为 $m = m(\tau)$，并且可以证明

$$\tau := B^2 \|G^{-1}\|^4 s \asymp B^2 2^{4bj_0} s,$$

回想起 $d_{j,\ n} \geq 1$ 和 $\zeta \in [0,1]$，它通常大于

$$\tau := B^2 \max_{(j,n)\in\Lambda_{j_0}} (d_{j,n}^{-2} 2^{2bj}) 2^{2(1-\zeta)bj_0} s.$$

在实践中，这将显著改善 Radon 变换的估计，正如第4节将要展示的那样。

定理 3.11 中的可调参数 ζ 表明，问题的样本复杂度与重建误差之间存在权衡：ζ的值越大，所需的测量数量越少，但重建误差可能越大。现在，我们讨论定理 3.11 中选择参数 ζ时两个特别值得关注的极端情形。首先，我们考虑 $\zeta = 0$的情形，此时测量数量最多，但重建误差的上界可能最好。

推论 3.14. 考虑定理 3.11 的设定，并假设满足相同的假设。令 $\widehat{x}$ 为以下最小化问题的一个解：

$$\min_{x\in\ell^2(\Lambda_{j_0})} \|x\|_{1,\omega} \quad : \quad \frac{1}{m}\sum_{k=1}^{m} \|F_{t_k}\Phi^*\iota_{\Lambda_{j_0}}x - y_k\|_{\mathcal{H}_2}^2 \le \left(\beta + C_3 2^{-bj_0} r\right)^2. \tag{12}$$

If

$$\tau \coloneqq B^2 \max_{(j,n)\in\Lambda_{j_0}} (d_{j,n}^{-2} 2^{2bj}) 2^{2bj_0} s.$$

and

$$m \ge C_0 \tau \max\{\log^3 \tau \log M, \log(1/\gamma)\},$$

那么，以高于 $1-\gamma$ 的概率，以下恢复估计成立：

$$\|x^\dagger - \widehat{x}\|_2 \le C_1 \frac{\sigma_s(P_{\Lambda_{j_0}} x^\dagger)_{1,\omega}}{\sqrt{s}} + C_2 c_\nu^{-1/2}(2^{bj_0}\beta + C_3 r).$$

接下来，我们考虑定理 3.11 中的情形 $\zeta = 1$。该情形所需的测量数量最少，但重建误差的界可能较大。

推论 3.15。考虑定理 3.11 的设定，并假设满足相同的假设。设 $\widehat{x}$ 是以下最小化问题的一个解：

$$\min_{x\in\ell^2(\Lambda_{j_0})} \|W^{-1}x\|_{1,\omega} : \quad \frac{1}{m}\sum_{k=1}^{m} \|F_{t_k}\Phi^*\iota_{\Lambda_{j_0}}x - y_k\|_{\mathcal{H}_2}^2 \le \left(\beta + C_3 2^{-bj_0} r\right)^2,$$

If

$$\tau \coloneqq B^2 \max_{(j,n)\in\Lambda_{j_0}} (d_{j,n}^{-2} 2^{2bj}) s \tag{13}$$

and

$$m \ge C_0 \tau \max\{\log^3 \tau \log M, \log(1/\gamma)\}, \tag{14}$$

那么，以高于 $1-\gamma$ 的概率，以下恢复估计成立：

$$\|F\Phi^* x^\dagger - F\Phi^*\widehat{x}\|_{L^2_\mu(\mathcal{D};\mathcal{H}_2)} \le C_1 \frac{\sigma_s(P_{\Lambda_{j_0}} W^{-1} x^\dagger)_{1,\omega}}{\sqrt{s}} + C_2 c_\nu^{-1/2}(\beta + 2^{-bj_0} C_3 r).$$

特别地，我们有

$$\|x^\dagger - \widehat{x}\|_2 \le C_1 2^{bj_0} \frac{\sigma_s(P_{\Lambda_{j_0}} W^{-1} x^\dagger)_{1,\omega}}{\sqrt{s}} + C_2 c_\nu^{-1/2}(2^{bj_0}\beta + C_3 r).$$

由推论 3.15 可直接得到以下精确恢复结果。若不存在截断误差（$x^\dagger \in \Lambda_{j_0}$，即 $r = 0$），若 $x^\dagger$ 是 s-稀疏的（$\sigma_s\left(P_{\Lambda_{j_0}} W^{-1} x^\dagger\right)_1 = \sigma_s\left(x^\dagger\right)_1 = 0$），且不存在噪声（$\beta = 0$），则

$$\widehat{x} = x^\dagger,$$

前提是 m 满足 (14)。

注 3.16. 我们来讨论一下 (13) 和 (14) 中出现的样本复杂度。为简单起见，考虑 $f_\nu = 1$, $\omega_{j,n} = 1$ 和 $\zeta = 1$ 的情形。在经典压缩感知中，$(F_t)_{t\in\mathcal{D}}$ 是测量算子，且 F 是等距算子，样本复杂度中出现的一个相关量是系统的相干性 $B > 0$，满足

$$\max_{(j,n)\in\Lambda_{j_0}} \|F_t \phi_{j,n}\|_{\mathcal{H}_2} \le B.$$

引入病态性的来源时（例如，若 F 是紧算子），必须用 $F\phi_{j,n}$ 的 L^2 范数对该量进行适当归一化。请注意，在这种设定下，假设 3.10 表示为 $\|F_t\phi_{j,n}\|_{\mathcal{H}_2} \le B d_{j\,n}^{-1}$ 。。另一方面，准对角化性质意味着 $\|F\phi_{j,n}\|_{L^2_\mu(\mathcal{D};\mathcal{H}_2)} \asymp 2^{-bj}$。因此，我们引入系统相对相干性的概念，即

$$\max_{(j,n)\in\Lambda_{j_0}} (B d_{j,n}^{-1} 2^{bj}).$$

从 (13) 和 (14) 可以明显看出，该系统的相对相干性在经典压缩感知中发挥着与通常的相干性相同的作用，其中对 $m \gtrsim B^2 s$ 给出了恢复保证，误差仅相差对数项。

3.3. 恢复误差分析。现在，我们将推导关于噪声 β 和测量次数 m 的显式估计。为实现这一目标，我们需要对假设 3.10 中参数 $d_{j,n}$ 的增长以及信号 $x^\dagger$ 的正则性施加额外假设。后者包括截断误差 $\|P^\perp_{\Lambda_{j_0}} x^\dagger\|_2$ 关于 $|\Lambda_{j_0}|$ 的某种衰减——即 $\|P^\perp_{\Lambda_{j_0}} x^\dagger\|_2 \le C 2^{-a j_0}$，其中 $|\Lambda_{j_0}| \asymp 2^{c j_0}$ ——以及对某个 $p \ge 1/2$ 的 p-可压缩性，这意味着$\sigma_s(x^\dagger)_1 \le C s^{1/2-p}$。

在许多问题中，这些估计都很自然，因为在这些问题里，假设信号 $x^\dagger$ 属于某个特定的信号类。例如，对于卡通类图像，我们有 $a = 1/2$，并且具有 $p = 1/2$ 意义下的近似 p-可压缩性，具体含义将在下文说明——参见例 4.7。

定理 3.17。考虑定理 3.14 的情形，并假设满足相同的假设。再假设下列估计成立：

$$|\Lambda_j| \le C 2^{cj}, \qquad d_{j,n} \ge 2^{dj},$$

$$\sigma_s(x^\dagger)_1 \le Cs^{1/2-p}, \qquad \|P_{\Lambda_j}^{\perp} x^\dagger\|_2 \le C2^{-aj},$$

其中 a、c、p、$C>0$ 和 $0 \le d \le b$ 为某些值。

固定 $\gamma \in (0,1)$，并令 $\beta \in (0,1)$ 足够小。令 $j_0 := \lfloor 1/((a+b)\log(1/\beta)) \rfloor$，并令 m 足够大。

存在仅取决于 a、b、c、p、C、假设 3.10 中出现的 B、(6) 中的准对角化界、C_F，以及对 γ 呈对数依赖的常数 C_0、$C_1 > 0$，使得以下结论成立。

令 $\widehat{x}$ 为最小化问题 (12) 的一个解。以超过 $1-\gamma$ 的概率，以下恢复估计成立：

$$\|x^\dagger - \widehat{x}\|_2 \le C_1 \left(\frac{\log^{4/p}(1/\beta)}{\beta^{2\frac{2b-d}{a+b}p} m^p} + \beta^{\frac{a}{a+b}} \right).$$

特别地，如果

$$m = \lfloor C_0 \beta^{-2\frac{(2b-d)}{a+b} - \frac{a}{p(a+b)}} \log^4(1/\beta) \rfloor,$$

then

$$\|x^\dagger - \widehat{x}\|_2 \le C_1 \beta^{\frac{a}{a+b}},$$

或者，等价地，

$$\|x^\dagger - \widehat{x}\|_2 \le C_1 \left(\frac{\log^4 m}{m} \right)^{\frac{ap}{a+4bp-2dp}}.$$

注 3.18. 遵循统计逆向学习文献中常用的方法——例如参见 [21], ——并为使表述简洁，我们在定理 3.17 中采用了渐近表述，其中 β 足够小且 m 足够大。分析该定理的证明后，可以将这一陈述定量地精确化。

4. The sparse Radon transform

现在，我们考虑通过沿角度 $\theta \in \mathbb{S}^1$ 对相应的 Radon 变换或扇束变换进行采样，来重建支撑在单位球 $\mathcal{B}_1 \subset \mathbb{R}^2$ 内的函数这一问题。

Radon 变换与扇束变换。下面回顾二维 Radon 变换的定义。

定义 4.1. 设 $K \subset \mathbb{R}^2$ 为一个紧集。令 $u \in L^2(K)$ 和 $\theta \in [0, 2\pi)$。沿角度 θ, $\mathcal{R}_\theta$: $L^2(K) \to L^2(\mathbb{R})$ 的 Radon 变换定义为

$$\mathcal{R}_\theta u(s) := \int_{e_\theta^\perp} u(y + se_\theta) \mathrm{d}y,$$

其中 $\mathrm{d}y$ 是 $e_\theta^\perp$ 上的一维勒贝格测度，且 $e_\theta := (\cos\theta, \sin\theta)$。

（完整）Radon 变换 $\mathcal{R}$: $L^2(K) \to L^2([0,2\pi) \times \mathbb{R})$ 定义为 $\mathcal{R}u(\theta, s) := \mathcal{R}_\theta u(s)$。

略微滥用记号，即使在 $\theta \in \mathbb{S}^1$ 的情况下，我们也将写作 $\mathcal{R}_\theta$，并采用 $[0,2\pi)$ 与 $\mathbb{S}^1$ 之间的自然等同。

更准确地说，$\mathcal{R}_\theta u$ 首先定义在 $L^2(K)$ 的一个稠密子空间中的函数上，使得积分表达式有意义，例如 $C^0(K)$。连续延拓到 $L^2(K)$，以及 $\mathcal{R}_\theta$ 和 $\mathcal{R}$ 的适定性，均由以下结果保证。

引理 4.2（[83, 定理 1.6]）。$\mathcal{R}_\theta$ 和 $\mathcal{R}$ 是定义良好的连续算子。此外，$\mathcal{R}$ 和 $\{\mathcal{R}_\theta\}_{\theta\in\mathbb{S}^1}$ 的范数均可由一个仅依赖于 $|K|$ 的常数一致地界定。

我们也考虑扇束变换，它在应用中更具实际意义。

定义 4.3. 令 $0 < d < \rho$。令 $u \in L^2(\mathcal{B}_d)$ 和 $\theta \in [0,2\pi)$。从角度 $\theta \in [0,2\pi)$、距离 ρ，$\mathcal{D}_\theta$ 处出发的扇束变换：$L^2(\mathcal{B}_d) \to L^2(-\frac{\pi}{2}, \frac{\pi}{2})$，定义为

$$\mathcal{D}_\theta u(\alpha) := \int_{\mathbb{R}} u(\rho e_\theta + t e_{\theta+\alpha})\mathrm{d}t.$$

（完整）扇束变换 $\mathcal{D}$: $L^2(\mathcal{B}_d) \to L^2([0,2\pi) \times (-\frac{\pi}{2}, \frac{\pi}{2}))$ 定义为$\mathcal{D}u(\theta,\alpha) := \mathcal{D}_\theta u(\alpha)$。

这不过是 Radon 变换的另一种参数化方式。我们考虑支撑包含在半径为 d 的球内的函数，并取中心位于半径为 ρ 的球面上的径向直线。

紧支撑小波。我们简要回顾紧支撑小波正交归一基 $(\phi_{j,\ n})_{j,\ n}$（属于）$L^2(\mathbb{R}^2)$ 的一种标准构造——参见 [43]，其中有全面的讨论。

我们首先在 $L^2(\mathbb{R})$ 中进行初步构造。对于 $j \in \mathbb{N}$ 和 $n \in \mathbb{Z}$，给定一个紧支撑尺度函数 $\chi \in L^2(\mathbb{R})$ 和一个母小波 $\psi \in L^2(\mathbb{R})$，我们定义

$$\chi_n := \tau_n \chi, \qquad \psi_{j,n} := D^1_{j-1}\tau_n \psi,$$

其中，$D^1_j f := 2^{j/2} f(2^j \cdot)$ 表示在一维中对 f进行二进伸缩，$\tau_n f := f(\cdot - n)$ 表示平移 n。

适当地选取 χ和 ψ后，我们有 $\{\chi_n\}_n \cup \{\psi_{j,n}\}_{j,n}$ 是 $L^2(\mathbb{R})$ 的正交归一基——参见 [43, 第 6] 章。此外，如 [81, 第3节中的定理 3.8] 所示，χ 和 ψ 可以取自 $C^2(\mathbb{R})$——更多细节请参见注 3.9。

现在，我们使用标准正交可分离小波的标准构造来建立一个正交归一基——例如参见 [80, 第7节.7]。回想一下，单变量函数 f、g 的张量积定义为 $(f \otimes g)(x,y) := f(x)g(y)$。对于 $j \in \mathbb{N}$ 和 $n_1,\ n_2 \in \mathbb{Z}$，我们设

$$\chi_{n_1,n_2} := \chi_{n_1} \otimes \chi_{n_2}, \qquad \psi^{(1)}_{0,n_1,n_2} := \psi_{0,n_1} \otimes \chi_{n_2},$$
$$\psi^{(2)}_{0,n_1,n_2} := \chi_{n_1} \otimes \psi_{0,n_2}, \qquad \psi^{(3)}_{0,n_1,n_2} := \psi_{0,n_1} \otimes \psi_{0,n_2}$$

以及 $\psi^{\varepsilon}_{j,\ n_1,\ n_2} := D_j^2 \psi^{\varepsilon}_{0,\ n_1,\ n_2}$，其中 $D_j^2 f := 2^j f(2^j \cdot)$ 是 f 在二维中的二进伸缩。可以证明

$$\{\chi_{n_1,n_2}\}_{n_1,n_2} \cup \{\psi^{(\varepsilon)}_{j,n_1,n_2}\}_{j,n_1,n_2,\varepsilon}, \qquad \varepsilon \in \{1,2,3\},$$

是 $L^2(\mathbb{R}^2)$ 的正交归一基。

最后，根据以下规则对上述小波重新排列，得到正交归一基 $(\phi_{j,\ n})_{j,\ n}$：对于 $j = 0,\ \ n = (n_1,\ n_2) \in \mathbb{Z}^2$，以及对于$j \geq 1,\ \ n = (n_1,\ n_2,\ \varepsilon) \in \mathbb{Z}^2 \times \{1,2,3\}$，令

$$\phi_{0,n} := \chi_{n_1,n_2}; \qquad \phi_{j,n} := \psi^{(\varepsilon)}_{j,n_1,n_2}.$$

令 Γ 为指标集合 $(j,\ n)$，满足 $\operatorname{supp} \phi_{j,\ n} \cap \mathcal{B}_1 \neq \emptyset$，即

$$\Gamma = \{(j,n) \in (\{0\} \times \mathbb{Z}^2) \cup \big(\mathbb{N} \times (\mathbb{Z}^2 \times \{1,2,3\})\big) : \operatorname{supp} \phi_{j,n} \cap \mathcal{B}_1 \neq \emptyset\}.$$

接下来，我们考虑空间 $\mathcal{H}_1 := \operatorname{span}(\phi_{j,\ n})_{(j,n)\in\Gamma}$。注意到$\overline{L^2(\mathcal{B}_1) \subseteq \mathcal{H}_1}$。容易看出，

$$|\{(j',n) \in \Gamma :\ j' = j\}| \asymp 2^{2j}, \qquad j \in \mathbb{N}. \tag{15}$$

事实上，根据构造，尺度 $j+1$ 处的小波数量约等于尺度 j 处对应数量乘以 2^2。结合式 (9)，这意味着 $|\Lambda_j| \asymp 2^{2j}$。

最后，我们选择 $0 < d < \rho$，使得对于每个 $u \in \mathcal{H}_1$，$\operatorname{supp} u \subseteq \mathcal{B}_d$，从而 $\mathcal{D}u$对每个 $u \in \mathcal{H}_1$都有定义。

主要结果。关于稀疏 Radon 变换的主要恢复结果如下。

定理 4.4。考虑如上定义的小波字典 $(\phi_{j,\ n})_{j,\ n\in\Gamma}$ 及相应的分析算子 $\Phi\colon\ \mathcal{H}_1 \to \ell^2(\Gamma)$。令 F 为 $\mathcal{R}$ 或 $\mathcal{D}$。存在常数 C_0、C_1、C_2、$C_3 > 0$；对于 Radon 变换，它们仅取决于所选的小波基；对于扇束变换，它们还取决于 d 和 ρ，使得以下结论成立。

固定 $j_0 \in \mathbb{N}$。设 $u^\dagger \in L^2(\mathcal{B}_1)$ 满足 $\|P^{\perp}_{\Lambda_{j_0}} \Phi u^\dagger\|_2 \leq r$。考虑从 $[0, 2\pi$ 上的均匀分布中抽取的 m 个独立同分布样本 $\theta_1, \ldots, \theta_m \in [0, 2\pi)$。

设 $y_k := F_{\theta_k} u^\dagger + \varepsilon_k$，其中 $k = 1, \dots, m$，且 $\max_k \|\varepsilon_k\|_{L^2} \le \beta$。设 $\widehat{u}$ 为以下最小化问题的一个解

$$\min_u \|\Phi u\|_1 \quad : \quad \frac{1}{m} \sum_{k=1}^{m} \|F_{\theta_k} u - y_k\|_{L^2}^2 \le \left(\beta + C_3 2^{-j_0/2} r\right)^2, \tag{16}$$

其中，最小值是在 $\mathrm{span}(\phi_{j,\ n})_{(j,n)\in\Lambda_{j_0}}$ 上取的。

令 $s \in [|\Lambda_{j_0}|]$，$s \ge 2$，并设 $\tau := 2^{j_0} s$。对于任意 $\gamma \in (0,1)$，若

$$m \ge C_0 \tau \max\{j_0 \log^3 \tau, \log(1/\gamma)\},$$

则以超过 $1-\gamma$ 的概率，以下恢复估计成立：

$$\|u^\dagger - \widehat{u}\|_{L^2} \le C_1 \frac{\sigma_s(P_{\Lambda_{j_0}} \Phi u^\dagger)_1}{\sqrt{s}} + C_2(2^{j_0/2}\beta + r). \tag{17}$$

上述结果对应于定理 3.11 中的情形 $\zeta = 0$（见推论 3.14），在这种情形下，增加测量次数可以得到更小的重建误差。在下一个定理中，我们考虑测量次数最少的情形（即推论 3.15 中讨论的 $\zeta = 1$ 情形），其测量次数与稀疏度 s 成正比，至多相差对数因子。这最适合恢复 $Fu^\dagger$，以及在没有噪声时精确恢复 $u^\dagger$，但通常会导致较差的误差界。

定理 4.5。考虑定理 4.4 的情形，并假设满足相同的假设。

存在常数 $C_0, C_1, C_2 > 0$，在 Radon 变换的情形下仅依赖于所选的小波基，而在扇束变换的情形下还依赖于 d 和 ρ，使得以下结论成立。令 $W := \mathrm{diag}(2^{j/2})_{j,\ n}$。令 $\widehat{u}$ 为最小化问题的一个解

$$\min_u \|W^{-1}\Phi u\|_1 \quad : \quad \frac{1}{m} \sum_{k=1}^{m} \|F_{\theta_k} u - y_k\|_{L^2}^2 \le \left(\beta + C_3 2^{-j_0/2} r\right)^2,$$

其中，最小值是在 $\mathrm{span}(\phi_{j,\ n})_{(j,n)\in\Lambda_{j_0}}$ 上取得的。

令 $s \in [|\Lambda_{j_0}|]$，$s \ge 2$。如果

$$m \ge C_0 s \max\{j_0 \log^3 s, \log(1/\gamma)\}, \tag{18}$$

th则以大于 $1-\gamma$ 的概率，以下恢复估计成立 :

$$\|Fu^\dagger - F\widehat{u}\|_{L^2(\mathbb{S}^1\times\mathbb{R})} \le C_1 \frac{\sigma_s\left(P_{\Lambda_{j_0}} W^{-1}\Phi u^\dagger\right)_1}{\sqrt{s}} + C_2(\beta + 2^{-j_0/2} r).$$

特别地，我们有

$$\|u^\dagger - \widehat{u}\|_{L^2} \le C_1 2^{j_0/2} \frac{\sigma_s\left(P_{\Lambda_{j_0}} W^{-1}\Phi u^\dagger\right)_1}{\sqrt{s}} + C_2(2^{j_0/2}\beta + r). \tag{19}$$

正如预期的那样，定理 4.5 所需的测量数量少于定理 4.4。另一方面，误差估计 (19) 可能比估计 (17) 更差，因为右端第一项乘有量 $2^{j_0/2}$ 。不过，值得注意的是，$\sigma_s\left(P_{\Lambda_{j_0}}W^{-1}\Phi u^\dagger\right)_1 \le \sigma_s\left(P_{\Lambda_{j_0}}\Phi u^\dagger\right)_1$，因此二者并不能直接比较。

作为定理 4.5 的直接推论，我们得到以下精确恢复结果。若没有截断误差（$u^\dagger \in \operatorname{span}(\phi_{j,n})^{(j,n)\in\Lambda_{j_0}}$，即 $r=0$），若 $\Phi u^\dagger$ 是s-稀疏的（$\sigma_s\left(P_{\Lambda_{j_0}}W^{-1}\Phi u^\dagger\right)_1 = \sigma_s\left(\Phi u^\dagger\right)_1 = 0$），且没有噪声（$\beta=0$），则

$$\widehat{u} = u^\dagger,$$

前提是 m 满足 (18)。

与第 3.3 节一样，我们现在推导关于噪声 β 和测量数量 m的显式估计。

定理 4.6。考虑定理 4.4 中的情形，并假设满足相同的假设。

假设以下估计成立：

$$\sigma_s(\Phi u^\dagger)_1 \le C s^{1/2-p}, \qquad \|P_{\Lambda_j}^{\perp}\Phi u^\dagger\|_2 \le C 2^{-aj},$$

对于某些 a、p、$C>0$。

固定 $\gamma \in (0,1)$，并令 $\beta \in (0,1)$ 足够小。令 $j_0 := \lfloor 2/(2a+1)\log(1/\beta)\rfloor$，并令 m 足够大。

存在常数 C_0、$C_1>0$，它们仅取决于所选的小波基以及 a、p、C，对 γ 呈对数依赖（对于扇束变换，还取决于 d 和 ρ），使得以下结论成立。

设 $\widehat{x}$ 是最小化问题 (16) 的一个解。那么，以超过 $1-\gamma$ 的概率，以下恢复估计成立：

$$\|u^\dagger - \widehat{u}\|_{L^2} \le C_1\left(\frac{\log^{4/p}(1/\beta)}{\beta^{\frac{2p}{2a+1}} m^p} + \beta^{\frac{2a}{2a+1}}\right).$$

特别是，如果

$$m = \lfloor C_0 \beta^{-\frac{2}{2a+1}-\frac{2a}{p(2a+1)}} \log^4(1/\beta)\rfloor.$$

then

$$\|u^\dagger - \widehat{u}\|_{L^2} \le C_1 \beta^{\frac{2a}{2a+1}},$$

或者，等价地，

$$\|u^\dagger - \widehat{u}\|_{L^2} \le C_1\left(\frac{\log^4 m}{m}\right)^{\frac{ap}{a+p}}.$$

在统计逆向学习 [27, 29] 的背景下，也得到了稀疏 Radon 变换的类似估计。这些估计与上文所述的结果在精神上相似，但由于两者的设定有很大差异，直接比较并不容易：前者未考虑稀疏性，而稀疏性在本文中起着至关重要的作用。

例 4.7（卡通类图像）。对于支撑在$\mathcal{B}_1$ 中的卡通类图像（见 [80, 第 9.2 节。4]——简而言之， C^2 除 C^2 边缘外均为平滑的信号），参照定理 4.6 的记号，我们有 $a = 1/2$ 和 p-可压缩性成立，且$p = 1/2$ 成立至对数项因子——参见第 5.7 节末尾。存在常数C_0、 $C_1 > 0$；对于 Radon 变换，它们仅取决于所选的小波基；对于扇束变换，还取决于 d 和 ρ 。以下 L^2 误差界成立（相差对数项因子）：

$$\|u^\dagger - \widehat{u}\|_{L^2} \le C_1 \left(\frac{1}{\beta^{1/2} m^{1/2}} + \beta^{1/2} \right). \tag{20}$$

取 $m = \lfloor C_0 \beta^{-2} \log^4(1/\beta) \rfloor$，可得（相差对数项因子）：

$$\|u^\dagger - \widehat{u}\|_{L^2} \le C_1 \beta^{1/2}, \qquad \|u^\dagger - \widehat{u}\|_{L^2} \le C_1 \left(\frac{1}{m} \right)^{1/4}.$$

关于（20）的证明，请参见第5.7节。

5. Proofs of the main results

在本节中，如果某个不等式在相差一个正的常数因子 $C > 0$ 的意义下成立，我们通常记作 $X \lesssim Y$，即 $X \le CY$。类似地，对于某个 $C > 0$，若 $X \ge CY$，我们记作 $X \gtrsim Y$。特别地， $X \asymp Y$ 表示 $X \lesssim Y$ 和 $X \gtrsim Y$ 均成立。一般而言，这些常数可能取决于推导出这些不等式的结果所涉及的其他量，具体依赖关系将在相应表述中明确说明。

5.1. 预备结果。对于第2.1节中引入的 ℓ^p_ω，求一个向量的最佳 s-ω-稀疏逼近，在实践中是一个颇具挑战的问题。 [91, 第 3] 节引入了准最佳稀疏逼近这一替代概念。事实上，后一种构造已足以满足下文证明的需要，因此我们在此简要回顾相关基本事实，以供读者参考。

给定 $x \in \ell^2(\Gamma)$、 $\omega \in [1, +\infty)^\Gamma$ 以及一个正整数 s，考虑与 $(|x_i|\omega_i^{-1})_{i\in\Gamma}$ 的非递增重排相关联的置换π： $\Gamma \to \Gamma$ 。令$i_0 \in \Gamma$ 为满足以下条件的最大索引

$$\omega(S_{i_0}) \le s, \quad S_{i_0} := \{\pi(1), \dots, \pi(i_0)\}.$$

因此，我们将 $x_{S_{i_0}}$ 定义为 x 的拟最佳 s-ω-稀疏逼近。注意，这一概念不依赖于 ℓ^p_ω 范数，而拟最佳 s-ω- 稀疏逼近 $x \in \ell^2(\Gamma)$ 关于 ℓ^p_ω 的相应误差定义为 $\tilde{\sigma}_s(x)_{p,\,\omega} := \|x_{S^c_{i_0}}\|_{p,\,\omega}$。根据定义，我们有

$$\sigma_s(x)_{p,\omega} \le \tilde{\sigma}_s(x)_{p,\omega}.$$

为便于后续使用，我们回顾以下 Stechkin 型不等式——证明见 [3, 引理 3.12]。若 $s > 0$ 且 $0 < p < q \le 2$，则

$$\sigma_s(x)_{q,\omega} \le \tilde{\sigma}_s(x)_{q,\omega} \le s^{1/q-1/p}\|x\|_{p,\omega}, \qquad x \in \ell^2(\Gamma). \tag{21}$$

为简化证明中的记号，下文还将引入以下与样本 $(t_1, \dots, t_m) \in \mathcal{D}^m$相关联的采样算子 $A\colon\ \ell^2(\Gamma) \to \mathcal{H}_2^m$：

$$Ax := \frac{1}{\sqrt{m}} \left(F_{t_1}\Phi^* x, \dots, F_{t_m}\Phi^* x\right).$$

略微滥用记号，我们也用 A 表示限制映射 $A\colon\ \ell^2(\Lambda) \to \mathcal{H}_2^m$到某个有限 $\Lambda \subset \Gamma$ 上的情形。

此外，针对问题 (4)，我们定义 $\varepsilon := \frac{1}{\sqrt{m}}(\varepsilon_1, \dots, \varepsilon_m)$ 和 $y := \frac{1}{\sqrt{m}}(y_1, \dots, y_m)$。注意，条件 $\max_k \|\varepsilon_k\|_{\mathcal{H}_2} \le \beta$ 意味着

$$\|\varepsilon\|^2_{\mathcal{H}_2^m} = \frac{1}{m}\sum_{k=1}^m \|\varepsilon_k\|^2_{\mathcal{H}_2} \le \beta^2.$$

此外，以下恒等式对于 $x \in \ell^2(\Lambda)$ 成立：

$$\|Ax - y\|^2_{\mathcal{H}_2^m} = \frac{1}{m}\sum_{k=1}^m \|F_{t_k}\Phi^* \iota_\Lambda x - y_k\|^2,$$

其中 $y_k = F_{t_k}\Phi^* x^\dagger + \varepsilon_k$，对于 $k = 1,\ \dots\dots,\ m$，如第 2.2 节所定义。

5.2. 稳健零空间性质。在基于 RIP 的情形下，获得恢复估计的典型策略是首先引入经过适当设计的稳健零空间性质（RNSP）版本；该性质对于确保稳定距离界，进而获得一致恢复结果，至关重要。我们提醒读者，除非另有说明，所有希尔伯特空间均假定为复数且可分。

定义 5.1. 设 $\mathcal{H}$ 为希尔伯特空间，Λ 为有限或可数集合。考虑权重向量 $\omega \in [1, +\infty)^\Lambda$、一个整数 $s \ge 1$以及实参数$0 < \rho < 1$, $\kappa > 0$。若算子 $A \in \mathcal{L}(\ell^2(\Lambda)\ \mathcal{H}^m)$ 满足相对于 $(\omega, \rho, \kappa, s)$ 的稳健零空间性质（RNSP），则

$$\|x_S\|_2 \le \frac{\rho}{\sqrt{s}}\|x_{S^c}\|_{1,\omega} + \kappa\|Ax\|_{\mathcal{H}^m}, \quad \forall\, x \in \ell^2(\Lambda),\ S \subset \Lambda : \omega(S) \le s.$$

注 5.2. 为使 A 满足相对于 $(\omega, \rho, \kappa, s)$ 的稳健零空间性质，一个充分条件可表述如下：存在 $0 < \rho' < 1/2$ 和 $\kappa' > 0$ 使得

$$\|x_S\|_2 \leq \frac{\rho'}{\sqrt{s}}\|x\|_{1,\omega} + \kappa'\|Ax\|_{\mathcal{H}^m}, \quad \forall\, x \in \ell^2(\Lambda),\ S \subset \Lambda : \omega(S) \leq s. \tag{22}$$

这确实是由估计直接得到的一个简单推论

$$\|x\|_{1,\omega} = \|x_S\|_{1,\omega} + \|x_{S^c}\|_{1,\omega} \leq \sqrt{s}\|x_S\|_2 + \|x_{S^c}\|_{1,\omega},$$

而该结果又由柯西-施瓦茨不等式推出。相关常数之间的关系为

$$\rho = \frac{\rho'}{1-\rho'}, \quad \kappa = \frac{\kappa'}{1-\rho'}.$$

现在我们证明，RNSP 蕴含适当的距离界。事实上，以下结果的证明很大程度上借鉴了压缩感知理论中的标准论证——例如，参见 [53, 定理 5]和 [91, 定理 4.2]。该证明与 [2, 引理 6.24]基本相同；将其推广到希尔伯特空间值测量的情形只需做一些小的修改。

定理 5.3（RNSP $\Rightarrow$ 距离界）。设 $\mathcal{H}$ 为希尔伯特空间，并假设 $A \in \mathcal{L}(\ell^2(\Lambda)\ \mathcal{H}^m)$ 满足定义 5.1 中关于 $(\omega, \rho, \kappa, s)$ 的 RNSP。则对所有 $x, z \in \ell^1_\omega(\Lambda)$,

$$\|x - z\|_{1,\omega} \leq c_1'\left(\|z\|_{1,\omega} - \|x\|_{1,\omega} + 2\sigma_s(x)_{1,\omega}\right) + c_2'\sqrt{s}\|A(x-z)\|_{\mathcal{H}^m},$$

其中我们设

$$c_1' = \frac{1+\rho}{1-\rho}, \quad c_2' = \frac{2\kappa}{1-\rho}.$$

此外

$$\|x - z\|_2 \leq \frac{c_1}{\sqrt{s}}\left(\|z\|_{1,\omega} - \|x\|_{1,\omega} + 2\sigma_s(x)_{1,\omega}\right) + c_2\|A(x-z)\|_{\mathcal{H}^m},$$

其中

$$c_1 = (\rho+1)c_1', \quad c_2 = (\rho+1)c_2' + \kappa.$$

5.3. 广义受限等距性质。现在我们定义标准受限等距性质（RIP）的一个广义版本，它对应于不分层且具有加权稀疏度的 G-RIPL，此概念已在 [1, 定义 3.5] 中引入。

定义5.4（广义受限等距性质）。设 $\mathcal{H}$ 是一个希尔伯特空间。给定 $G \in \mathcal{L}(\ell^2(\Lambda))$、$0 \leq \delta < 1$、$\lambda > 0$ 和一个权重向量 $\omega \in [1, +\infty)^\Lambda$，我们称 $A \in \mathcal{L}(\ell^2(\Lambda)\ \mathcal{H}^m)$、$m \in \mathbb{N}$ 关于 $(G, \omega, \delta, \lambda)$ 满足广义受限等距性质（g-RIP），如果

$$(1-\delta)\|Gx\|_2^2 \leq \|Ax\|_{\mathcal{H}^m}^2 \leq (1+\delta)\|Gx\|_2^2, \quad x \in \ell^2(\Lambda),\ \|x\|_{0,\omega} \leq \lambda.$$

注 5.5（g-RIP 的对角不变性）。考虑 Λ 为有限集的情形。设 $Z \coloneqq (z_i)_{i\in\Lambda}$ 是一个满足 $z_i \neq 0$ 的对角矩阵。由于 $\mathrm{supp}(x') = \mathrm{supp}(Zx')$，我们有 $\|x'\|_{0,\,\omega} = \|Zx'\|_{0,\,\omega}$。因此，经过变量替换 $x = Zx'$ 后，显然，A 关于 $(G,\omega,\delta,\lambda)$ 满足 g-RIP，当且仅当 AZ 关于 $(GZ,\omega,\delta,\lambda)$ 满足 g-RIP。

现在我们给出一些充分条件，在这些条件下，广义限制等距性质蕴含稳健零空间性质。另请参阅 [1, 定理 5.5]，其中给出了针对分层稀疏性调整后的这一结果版本。

定理 5.6（广义限制等距性质 $\Rightarrow$ 稳健零空间性质）。固定 $\rho' \in (0,1/2)$ 和 $\delta \in (0,1)$。假设$A \in \mathcal{L}(\ell^2(\Lambda)\ \mathcal{H}^m)$ 关于 $(G,\omega,\delta,\lambda)$ 满足广义限制等距性质，其中存在可逆的$G \in \mathcal{L}(L^2(\Lambda))$、权重 $\omega \in [1,+\infty)^\Lambda$ 以及 $\lambda \in \mathbb{R}^+$。对于任意 $s \geq \|\omega\|_\infty^2/4$，若

$$\lambda \geq \frac{5}{\rho'^2}\frac{1+\delta}{1-\delta}\|G^{-1}\|^2\|G\|^2 s,$$

则 A 关于 (ω,ρ,κ,s) 满足稳健零空间性质，其中

$$\rho = \frac{\rho'}{1-\rho'}, \quad \kappa = \frac{\|G^{-1}\|}{(1-\rho')\sqrt{1-\delta}}.$$

证明。根据注5.2，只需证明（22）成立。

固定 $x \in \ell^2(\Lambda)$以及满足 $\omega(S) \leq s$的 $S \subset \Lambda$。考虑 x_{S^c}。鉴于第5.1节中构造 x 的准最优 λ-ω-稀疏逼近的方法，我们现在按如下步骤进行。

令 $\pi\colon\ \mathbb{N} \to S^c$ 为产生$(|x_k|\omega_k^{-1})_{k\in S^c}$的非递增重排的枚举，使得$|x_{\pi(i)}|\omega_{\pi(i)}^{-1} \geq |x_{\pi(i+1)}|\omega_{\pi(i+1)}^{-1}$。令 $i_0 = 0$，并对 $j \geq 1$ 迭代定义 $i_j \in \mathbb{N}$ 为满足 $i_j > i_{j-1}$ 且.....的最大指标

$$\omega(\Omega_j) \leq \tilde{\lambda}, \quad \Omega_j \coloneqq \{\pi(i_{j-1}+1),\dots,\pi(i_j)\},$$

其中

$$\tilde{\lambda} = \frac{4}{\rho'^2}\frac{1+\delta}{1-\delta}\|G^{-1}\|^2\|G\|^2 s. \tag{23}$$

注意，由于 $\rho' \in (0,1/2)$，因此$\lambda \geq \tilde{\lambda} + \frac{1}{\rho'^2}\frac{1+\delta}{1-\delta}\|G^{-1}\|^2\|G\|^2 s \geq \tilde{\lambda} + s$。我们强调，每个 Ω_j 都非空，这要归功于假设条件 $\tilde{\lambda} \geq 4s \geq \|\omega\|_\infty^2$；并且由于 $\omega_i \geq 1$，它是有限的。

显然，$\cup_{j\geq 1}\Omega_j = S^c$。此外，对于 $j \geq 1$，根据定义可知，x_{Ω_j} 是 $x_{(\Omega_j\cup\Omega_{j+1})}$ 的准最佳 λ-ω-稀疏近似。Stechkin 型不等式 (21) 给出

$$\|x_{\Omega_j}\|_2 = \tilde{\sigma}_{\tilde{\lambda}}(x_{(\Omega_{j-1}\cup\Omega_j)})_{2,\omega} \leq \frac{1}{\sqrt{\tilde{\lambda}}}\|x_{(\Omega_{j-1}\cup\Omega_j)}\|_{1,\omega}, \quad j \geq 2. \tag{24}$$

由于 $x_S + x_{\Omega_1} = x - \sum_{j\geq 2} x_{\Omega_j}$ 是根据构造得到的， A 的广义限制等距性质给出

$$\begin{aligned}\|x_S\|_2 \leq \|x_S + x_{\Omega_1}\|_2 &\leq \|G^{-1}\|\|G(x_S + x_{\Omega_1})\|_2 \\ &\leq \frac{\|G^{-1}\|}{\sqrt{1-\delta}}\|A(x_S + x_{\Omega_1})\|_{\mathcal{H}^m} \\ &\leq \frac{\|G^{-1}\|}{\sqrt{1-\delta}}\left(\sum_{j\geq 2}\|Ax_{\Omega_j}\|_{\mathcal{H}^m} + \|Ax\|_{\mathcal{H}^m}\right).\end{aligned} \tag{25}$$

现在，由 (24) 可得

$$\|Ax_{\Omega_j}\|_{\mathcal{H}^m} \leq \sqrt{1+\delta}\|Gx_{\Omega_j}\|_2 \leq \sqrt{1+\delta}\|G\|\|x_{\Omega_j}\|_2 \leq \sqrt{1+\delta}\|G\|\frac{1}{\sqrt{\tilde{\lambda}}}\|x_{(\Omega_{j-1}\cup\Omega_j)}\|_{1,\omega}. \tag{26}$$

我们有

$$\sum_{j\geq 2}\|x_{(\Omega_{j-1}\cup\Omega_j)}\|_{1,\omega} \leq \sum_{j\geq 2}\|x_{\Omega_{j-1}}\|_{1,\omega} + \|x_{\Omega_j}\|_{1,\omega} \leq 2\|x\|_{1,\omega}. \tag{27}$$

结合 (25)、(26) 和 (27)，得到

$$\|x_S\|_2 \leq \frac{\sqrt{1+\delta}}{\sqrt{1-\delta}}\|G^{-1}\|\|G\|\frac{2\sqrt{s}}{\sqrt{\tilde{\lambda}}}\frac{1}{\sqrt{s}}\|x\|_{1,\omega} + \frac{\|G^{-1}\|}{\sqrt{1-\delta}}\|Ax\|_{\mathcal{H}^m}.$$

现在只需注意到 (23) 蕴含

$$\frac{\sqrt{1+\delta}}{\sqrt{1-\delta}}\|G^{-1}\|\|G\|\frac{2\sqrt{s}}{\sqrt{\tilde{\lambda}}} = \rho'.$$

(22) 的证明至此完成，该结论也随之得证。 □

5.4. 通过子采样保证广义限制等距性质的充分条件。现在，我们给出使模型的采样算子满足广义限制等距性质的充分条件。

定理 5.7（采样 $\Rightarrow$ 广义限制等距性质）。考虑第 2.2 节中的设定。假设假设 3.1 和 3.3 成立。定义

$$G \coloneqq \sqrt{P_\Lambda \Phi F^* F \Phi^* \iota_\Lambda}.$$

假设 G 可逆。

设 $(t_1, \ldots, t_m) \in \mathcal{D}^m$, $m \in \mathbb{N}$ 是从 ν 中独立同分布抽取的样本（可能出现的重复样本予以保留）。设 A 为与 $(t_1, \ldots, t_m)$ 相关联的采样算子，即

$$A \coloneqq \left(\frac{1}{\sqrt{m}}F_{t_k}\Phi^*\iota_\Lambda\right)_{k=1}^m,$$

并引入算子

$$Q \in \mathcal{L}(\mathcal{H}^m), \quad Q(h_1, \dots, h_m) = (f_\nu(t_1)^{-1/2} h_1, \dots, f_\nu(t_m)^{-1/2} h_m). \tag{28}$$

设 $M = |\Lambda|$ 和 $\lambda \in [M]$。假设 $\tau := B^2 \|G^{-1}\|^2 \lambda \geq 3$。存在一个普适常数 $C > 0$，使得对于任意 γ， $\delta \in (0, 1)$，若

$$m \geq C\delta^{-2}\tau \max\{\log^3 \tau \log M, \log(1/\gamma)\}, \tag{29}$$

则 QA 以超过 $1-\gamma$的概率满足关于 $(G, \omega, \delta, \lambda)$ 的广义限制等距性质。

这一结果的证明遵循压缩感知理论中一条成熟的思路，旨在为 δ 推导概率界——例如，可参见 [1, 定理 3.6], [74, 定理 3.1]，[77, 定理 3.2] 和 [91, 定理5.2]的证明中所采用的论证和技巧。关于定理 5.7 的证明，请参见附录 A。

5.5. 定理 3.5 的证明。要完成定理 3.5 的证明，还需要以下关于将无限维未知量 $x^\dagger$ 截断到有限维集合 Λ 所产生误差的估计。

命题 5.8（截断误差）。考虑定理 3.5 中的情形。假设 $\|P_\Lambda^\perp x^\dagger\|_2 \leq r$ 和

$$m \geq \|G_1^{-1}\|^2 \|G_2^{-1}\|^2 \|G_2\|^2 \log(1/\gamma) \tag{30}$$

对于某些可逆矩阵 G_1、G_2，且满足以下条件之一：

(1) $G_1 = G_2$，(2) 对某个常数 c，$\|G_1^{-1}\| \geq c$。

令 Q 按照 (28) 定义。则以大于 $1-\gamma$ 的概率，我们有

$$\|G_2^{-1}\| \|QAP_\Lambda^\perp x^\dagger\|_{\mathcal{H}_2^m} \leq C_1 c_\nu^{-1/2} \left(\|F\Phi^* \iota_{\Gamma\setminus\Lambda}\| \|G_2^{-1}\| + 1 \right) r,$$

其中 $C_1 > 0$ 是一个仅取决于 C_F 的常数（在第二种情况下还取决于 c）。

证明。设 $u_R := \Phi^* P_\Lambda^\perp x^\dagger$ 且 $c_F := \|F\Phi^* \iota_{\Gamma\setminus\Lambda}\|$。我们将引理 A.1 应用于随机变量 $Y_k(t) := \|QF_{t_k} u_R\|_{\mathcal{H}_2}^2 - \|Fu_R\|_{L^2_\mu(\mathcal{D};\mathcal{H}_2)}^2$，其中函数族 f_x 仅由恒等映射组成；因此，我们得到伯恩斯坦不等式的经典版本作为特例。注意，在这种情况下，

$$\|QAP_\Lambda^\perp x^\dagger\|_{\mathcal{H}_2^m}^2 = \frac{1}{m} \sum_{k=1}^m f_\nu(t_k)^{-1} \|F_{t_k} u_R\|_{\mathcal{H}_2}^2.$$

我们声称 $\mathbb{E}(f_\nu(t_k)^{-1}\|F_{t_k}u_R\|_{\mathcal{H}_2}^2 - \|Fu_R\|_{L^2_\mu(\mathcal{D};\mathcal{H}_2)}^2) = 0$。事实上，根据$F$的定义，我们有

$$\int_{\mathcal{D}} \|F_t u\|_{\mathcal{H}_2}^2 \mathrm{d}\mu(t) = \|Fu\|_{L^2_\mu(\mathcal{D};\mathcal{H}_2)}^2, \quad u \in \mathcal{H}_1,$$

这与该结论等价。此外，我们有

$$\begin{aligned}|f_\nu(t_k)^{-1}\|F_{t_k}u_R\|_{\mathcal{H}_2}^2 - \|Fu_R\|_{L^2_\mu(\mathcal{D};\mathcal{H}_2)}^2| &\le c_\nu^{-1}\|F_{t_k}u_R\|_{\mathcal{H}_2}^2 + \|Fu_R\|_{L^2_\mu(\mathcal{D};\mathcal{H}_2)}^2 \\ &\le 2C_F^2 c_\nu^{-1} r^2 =: K,\end{aligned}$$

其中我们使用了以下事实：$\|F_t\|$、$\|F\| \le C_F$。最后，我们有

$$\begin{aligned}&\mathbb{E}|f_\nu(t_k)^{-1}\|F_{t_k}u_R\|_{\mathcal{H}_2}^2 - \|Fu_R\|_{L^2_\mu(\mathcal{D};\mathcal{H}_2)}^2|^2 \\ &\quad= \mathbb{E}|f_\nu(t_k)^{-1/2}\|F_{t_k}u_R\|_{\mathcal{H}_2}|^4 - \left(\mathbb{E}\|Fu_R\|_{L^2_\mu(\mathcal{D};\mathcal{H}_2)}^2\right)^2 \\ &\quad\le C_F^2 c_\nu^{-1} r^2 \mathbb{E}|f_\nu(t_k)^{-1/2}\|F_{t_k}u_R\|_{\mathcal{H}_2}|^2 \\ &\quad= C_F^2 c_\nu^{-1} r^2 \|Fu_R\|_{L^2_\mu(\mathcal{D};\mathcal{H}_2)}^2 \le C_F^2 c_\nu^{-1} c_F^2 r^4 =: \Sigma^2.\end{aligned}$$

将 εm 代替 ε 应用于 Bernstein 不等式（引理 A.1），可得

$$\begin{aligned}&\mathbb{P}\left(\|QAu_R\|^2 \ge \|Fu_R\|_{L^2_\mu(\mathcal{D};\mathcal{H}_2)}^2 + \varepsilon\right) \\ &\quad= \mathbb{P}\left(\sum_{k=1}^m f_\nu(t_k)^{-1}\|F_{t_k}u_R\|_{\mathcal{H}_2}^2 \ge m\|Fu_R\|_{L^2_\mu(\mathcal{D};\mathcal{H}_2)}^2 + m\varepsilon\right) \\ &\quad\le \exp\left(-\frac{1}{C}\frac{m^2\varepsilon^2}{mC_F^2 c_\nu^{-1} c_F^2 r^4 + mC_F^2 c_\nu^{-1} r^4 c_F^2 + C_F^2 c_\nu^{-1} r^2 m\varepsilon}\right) \\ &\quad\le \exp\left(-\frac{1}{CC_F^2 c_\nu^{-1}}\frac{m\varepsilon^2}{c_F^2 r^4 + r^2\varepsilon}\right),\end{aligned}$$

其中，$C > 0$ 是一个绝对常数。

因此，以超过 $1-\gamma$ 的概率，以下估计成立：

$$\|QAu_R\|_{\mathcal{H}_2^m}^2 \le \|Fu_R\|_{L^2_\mu(\mathcal{D};\mathcal{H}_2)}^2 + \varepsilon,$$

前提是

$$m \ge \tilde{C}\frac{c_F^2 r^4 + r^2\varepsilon}{\varepsilon^2}\log(1/\gamma),$$

其中 $\tilde{C} := CC_F^2 c_\nu^{-1}$。

现在我们对 $m = m(\varepsilon)$ 与 ε 之间的关系求逆，得到 $\varepsilon = \varepsilon(m)$。然后，对于如式 (30) 中的 m 和 $\varepsilon = \varepsilon(m)$，上述不等式显然成立。我们有

$\varepsilon = \varepsilon(m)$ 可表示为

$$\varepsilon = \frac{\tilde{C}\log(1/\gamma)r^2 + \sqrt{\tilde{C}^2\log^2(1/\gamma)r^4 + 4\tilde{C}\log(1/\gamma)c_F^2 r^4 m}}{2m}.$$

因此，我们得到：其概率超过 $1-\gamma$，

$$\begin{aligned}\|QAu_R\|^2_{\mathcal{H}_2^m} &\le \|Fu_R\|^2_{L^2_\mu(\mathcal{D};\mathcal{H}_2)} + \varepsilon \\ &\le c_F^2 r^2 + \frac{\tilde{C}\log(1/\gamma)}{m} r^2 + \sqrt{\frac{\tilde{C}\log(1/\gamma)}{m}} c_F r^2 \\ &\le \left(c_F^2 + \frac{CC_F^2}{\|G_1^{-1}\|^2\|G_2^{-1}\|^2\|G_2\|^2} + \frac{\sqrt{CC_F^2}c_F}{\|G_1^{-1}\|\|G_2^{-1}\|\|G_2\|} \right) c_\nu^{-1} r^2.\end{aligned}$$

现在我们区分命题陈述中的两种情况。如果$G_1 = G_2 =: G$，那么

$$\begin{aligned}\|G^{-1}\|^2\|QAu_R\|^2_{\mathcal{H}_2^m} &\le \left(c_F^2\|G^{-1}\|^2 + \frac{CC_F^2}{\|G^{-1}\|^2\|G\|^2} + \frac{\sqrt{CC_F^2}c_F}{\|G\|} \right) c_\nu^{-1} r^2 \\ &\le \left(c_F^2\|G^{-1}\|^2 + CC_F^2 + c_F\|G^{-1}\|\sqrt{CC_F^2} \right) c_\nu^{-1} r^2 \\ &\le C_1^2 \left(c_F^2\|G^{-1}\|^2 + 1 \right) c_\nu^{-1} r^2,\end{aligned}$$

其中 C_1 仅取决于 C_F。

另一方面，如果 $\|G_1^{-1}\| \ge c$，则得到

$$\begin{aligned}\|G_2^{-1}\|^2\|QAu_R\|^2_{\mathcal{H}_2^m} &\le \left(c_F^2\|G_2^{-1}\|^2 + \frac{CC_F^2}{\|G_1^{-1}\|^2\|G_2\|^2} + \frac{\sqrt{CC_F^2}c_F}{\|G_1^{-1}\|\|G_2\|}\|G_2^{-1}\| \right) c_\nu^{-1} r^2 \\ &\le \left(c_F^2\|G_2^{-1}\|^2 + CC_F^2 c^{-2}\|G_2^{-1}\|^2 + c_F\sqrt{CC_F^2}c^{-1}\|G_2^{-1}\|^2 \right) c_\nu^{-1} r^2 \\ &\le C_1^2 \left(c_F^2\|G_2^{-1}\|^2 + 1 \right) c_\nu^{-1} r^2,\end{aligned}$$

其中 C_1 仅取决于 C_F 和 c。 □

定理 3.5 的证明可通过串联本节中已证明的结果得出。

定理 3.5 的证明。我们有 m 满足定理 5.7 中的 (29)，其中

$$\lambda = C_0\|G^{-1}\|^2\|G\|s.$$

由定理 5.7 可知，QA（其中 Q 的定义见式 (28)）关于 $(G, \omega, \delta, \lambda)$ 满足广义限制等距性质，且 $\delta = 1/2$。利用定理 5.6（取 $\delta = 1/2$ 和 $\rho' = 1/3$），可知 QA 满足关于

(ω,ρ,κ,s) 的稳健零空间性质，且 $\rho = 1/2$ 和 $\kappa = \frac{3}{2\sqrt{1-\delta_*}}\|G^{-1}\| = \frac{3}{\sqrt{2}}\|G^{-1}\|$。我们有 $C_0 = \frac{5}{\rho'^2}(1+\delta)/(1-\delta) = 135$。

注意，$x^\dagger \in \ell^2(\Gamma)$，而 QA 在有限维子空间 $\ell^2(\Lambda) \subset \ell^2(\Gamma)$ 上满足 g-RIP性质。我们可以将误差项拆分如下：

$$\|x^\dagger - \widehat{x}\|_2 \le \|P_\Lambda x^\dagger - \widehat{x}\|_2 + \|P_\Lambda^\perp x^\dagger\|_2 \le \|P_\Lambda x^\dagger - \widehat{x}\|_2 + r.$$

现在我们来估计第一项。参照定理5.3中应用于 $P_\Lambda x^\dagger$ （代替 $x^\dagger$ ）的 ℓ^2 界，可得以下估计：

$$\|P_\Lambda x^\dagger - \widehat{x}\|_2 \le \frac{c_1}{\sqrt{s}}\left(\|\widehat{x}\|_{1,\omega} - \|P_\Lambda x^\dagger\|_{1,\omega} + 2\sigma_s(P_\Lambda x^\dagger)_{1,\omega}\right) + c_2\|G^{-1}\|\|QA(P_\Lambda x^\dagger - \widehat{x})\|_{\mathcal{H}_2^m}$$

其中 c_1、$c_2 > 0$ 为绝对常数。注意，$P_\Lambda x^\dagger$ 满足问题 (5) 的约束；事实上，我们可以利用假设3.3以及在$G_1 = G_2 = G$ 情形下应用的命题 5.8，得出以下界以超过 $1-\gamma$的概率成立：

$$\begin{aligned}\left(\frac{1}{m}\sum_{k=1}^m \|F_{t_k}\Phi^*\iota_\Lambda x - y_k\|_{\mathcal{H}_2}^2\right)^{\frac{1}{2}} &= \|AP_\Lambda x^\dagger - y\|_{\mathcal{H}_2^m} \\ &\le \|Ax^\dagger - y\|_{\mathcal{H}_2^m} + \|AP_\Lambda^\perp x^\dagger\|_{\mathcal{H}_2^m} \\ &\le \beta + \|Q^{-1}\|\|QAP_\Lambda^\perp x^\dagger\|_{\mathcal{H}_2^m} \\ &\le \beta + C_3\|G^{-1}\|^{-1}r.\end{aligned}$$

利用 $\widehat{x}$ 是 (5) 的极小值点这一事实，可得

$$\|P_\Lambda x^\dagger - \widehat{x}\|_2 \le \frac{2c_1}{\sqrt{s}}\sigma_s(P_\Lambda x^\dagger)_{1,\omega} + c_2\|G^{-1}\|\|QA(P_\Lambda x^\dagger - \widehat{x})\|_{\mathcal{H}_2^m}. \tag{31}$$

我们可以将 (31) 右端的最后一项拆分如下：

$$\|QA(P_\Lambda x^\dagger - \widehat{x})\|_{\mathcal{H}_2^m} \le \|QAx^\dagger - Qy\|_{\mathcal{H}_2^m} + \|QA\widehat{x} - Qy\|_{\mathcal{H}_2^m} + \|QAP_\Lambda^\perp x^\dagger\|_{\mathcal{H}_2^m}.$$

第一项满足

$$\|QAx^\dagger - Qy\|_{\mathcal{H}_2^m} \le c_\nu^{-1/2}\beta$$

根据关于噪声 水平的假设以及假设3.3。第二项满足

$$\|QA\widehat{x} - Qy\|_{\mathcal{H}_2^m} \le c_\nu^{-1/2}(\beta + C_3\|G^{-1}\|^{-1}r)$$

因为 $\widehat{x}$ 满足问题 (5) 的约束。最后，第三项满足 s

$$\|QAP_\Lambda^\perp x^\dagger\|_{\mathcal{H}_2^m} \le C_3\|G^{-1}\|^{-1}r \le c_\nu^{-1/2}C_3\|G^{-1}\|^{-1}r$$

根据命题 5.8 在情形 $G_1 = G_2 = G$ 下的应用，其概率大于 $1-\gamma$。

这些界共同推出了右端最后一项的以下界：
(31)：

$$\|G^{-1}\|\|QA(P_\Lambda x^\dagger - \widehat{x})\|_{\mathcal{H}_2^m} \le 2c_\nu^{-1/2}(\|G^{-1}\|\beta + C_3 r).$$

命题 5.8 给出的界成立，同时 QA 以大于 $1-2\gamma$ 的概率满足广义限制等距性质。这就完成了 ℓ^2 情形的证明。 ℓ^1_ω情形的处理方法类似。□

5.6. 关于 $\|G^{-1}\|$ 的定量界。我们证明第 3.2 节中陈述的结果。

命题 3.8 的证明。假设意味着

$$\|Fu\|^2_{L^2_\mu(\mathcal{D};\mathcal{H}_2)} \asymp \|u\|^2_{H^{-b}} \asymp \sum_{(j,n)\in\Gamma} 2^{-2bj}|\langle u, \phi_{j,n}\rangle|^2.$$

现在考虑 $x \in \ell^2(\Gamma)$，并令 $u := \Phi^* x$。根据假设，$(\phi_{j,\ n})_{(j,n)\in\Gamma}$ 是一个正交归一基，因此我们有 $x_{j,\ n} = \langle u, \phi_{j,\ n}\rangle_{L^2}$。于是

$$\|F\Phi^* x\|^2_{L^2_\mu(\mathcal{D};\mathcal{H}_2)} = \|Fu\|^2_{L^2_\mu(\mathcal{D};\mathcal{H}_2)} \asymp \sum_{(j,n)\in\Gamma} 2^{-2bj}|\langle u, \phi_{j,n}\rangle|^2 = \sum_{(j,n)\in\Gamma} 2^{-2bj}|x_{j,n}|^2,$$

因此， F 满足所要求的准对角化性质。 □

命题 5.9. 假设算子 F 满足关于 (Φ, b) $b \ge 0$ 的拟对角化性质 (6)。令 $G := \sqrt{P_\Lambda \Phi F^* F \Phi^* \iota_\Lambda}$，且 $Z := \mathrm{diag}(z_{j,n})_{(j,n)\in\Lambda}$ 为对角矩阵，其中对每个 $(j,n) \in \Lambda$，都有 $z_{j,n} > 0$ 。那么有以下估计：

$$\|(GZ)^{-1}\|^2 \asymp \max_{(j,n)\in\Lambda} (z_{j,n}^{-2} 2^{2bj}),$$

其中隐含常数取决于 (6) 中的准对角化常数。特别地，若 $z_{j,n} \equiv 1$ 且 $\Lambda = \Lambda_{j_0}$，则

$$\|G^{-1}\|^2 \le C 2^{2bj_0}.$$

证明. 令 $W := \mathrm{diag}(2^{bj})_{(j,n)\in\Lambda}$。由 (6) 可得

$$\|GZx\|_2^2 = \|F\Phi^* \iota_\Lambda Zx\|^2_{L^2_\mu(\mathcal{D};\mathcal{H}_2)} \ge c\|W^{-1}Zx\|_2^2 \ge c \min_{(j,n)\in\Lambda} (z_{j,n}^2 2^{-2bj})\|x\|_2^2.$$

结合这两个估计可得

$$\|x\|_2^2 \le \frac{1}{c} \max_{(j,n)\in\Lambda} (z_{j,n}^{-2} 2^{2bj})\|GZx\|_2^2.$$

这就证明了 $\|(GZ)^{-1}\|^2$ 的上界。

另一方面，令 $(j', n') \in \Lambda$ 满足

$$z^2_{j',n'} 2^{-2bj'} = \min_{(j,n)\in\Lambda} (z_{j,n}^2 2^{-2bj}).$$

则有

$$\|GZe_{j',n'}\|_2^2 \le C\|W^{-1}Ze_{j',n'}\|_2^2 = C\min_{(j,n)\in\Lambda}(z_{j,n}^2 2^{-2bj}).$$

令 $f_{j',n'} := \frac{GZe_{j',n'}}{\sqrt{C}}\,{}^{n'}\min_{(j,n)\in\Lambda}(z_{j,n}2^{-bj})$。前一个界表明 $\|f_{j',}$**{v23>**。因此，我们有

$$\|(GZ)^{-1}\|^2 = \sup_{\|y\|_2\le 1}\|(GZ)^{-1}y\|^2 \ge \|(GZ)^{-1}f_{j',n'}\|_2^2 = \frac{1}{C\min_{(j,n)\in\Lambda}(z_{j,n}^2 2^{-2bj})}.$$

这证明了 $\|(GZ)^{-1}\|^2$的下界。 □

~定理 3.11 的证明。令 $D := \mathrm{diag}(d_{j,n})_{(j,n)\in\Lambda_{j_0}}$，并令 D 为作用于 $\mathrm{span}(\phi_{j,n})_{(j,n)\in\Gamma}$ 上的算子，其定义为 $D\phi_{j,n} := d_{j,n}\phi_{j,n}$，当 $(j,n) \in \Lambda_{j_0}$, $D\phi_{j,n} = \phi_{j,n}$ 时；否则取 $\Phi^* D = \tilde{D}\Phi^*$。此外，我们将 W，因而也将 W^ζ，扩展为作用于$\ell^2(\Gamma)$上的算子，令 $W = \mathrm{diag}(w_{j,n})_{j,n}$，其中

$$w_{j,n} = \begin{cases} 2^{bj} & \text{if } (j,n)\in\Lambda_{j_0}, \\ 1 & \text{otherwise.} \end{cases}$$

同理，令 $\tilde{W}^\zeta = \Phi^* W^\zeta \Phi$。

~考虑由 $(F_t D)$ 给出的测量算子。相应的前向映射~为 FD，而截断矩阵（即起到 G 作用的矩阵）为 $\tilde{G} := \sqrt{D^* G^* G D}$，相应的采样矩阵为 AD。此外，$\|(\tilde{G})^{-1}\| = \|(GD)^{-1}\|$。

根据命题 5.9 和准对角化性质，可得以下界：

$$\|(GD)^{-1}\|^2 \asymp \max_{(j,n)\in\Lambda_{j_0}}(d_{j,n}^{-2}2^{2bj}), \quad \|(GW^\zeta)^{-1}\|^2 \asymp 2^{2(1-\zeta)bj_0}, \quad \|GW^\zeta\| \lesssim 1, \tag{32}$$

其中，隐含常数取决于准对角化常数。令 λ 定义为

$$\lambda := \tilde{C}_0\|GW^\zeta\|^2\|(GW^\zeta)^{-1}\|^2 s,$$

~其中，$C_0 = 135$ 是定理 5.6 中 λ 表达式里出现的常数，其中$\delta = 1/2$ 且 $\rho' = 1/3$。

利用这些界，我们可以推得，若 C_0 足够大，则 (11) 中的样本复杂度m满足 (29)，即

$$m \ge \frac{C}{4}\tau\max\{\log^3\tau\log M, \log(1/\gamma)\}$$

其中 $\tau = B^2\|(GD)^{-1}\|^2\lambda \ge 3$。由定理 5.7 可得，$QAD$ 以极高概率满足关于 $(GD, \omega, 1/2, \lambda)$ 的广义限制等距性质。对角

注 5.5 中证明的广义限制等距性质的对角不变性表明，后一个结果意味着 QA 满足关于 $(G, \omega, 1/2, \lambda)$ 的广义限制等距性质。

根据注 5.5 中所示的广义限制等距性质的对角不变性， QAW^{ζ}等价地满足关于 $(GW^{\zeta}, \omega, 1/2, \lambda)$ 的广义限制等距性质。定理 5.6 现可推出， QAW^{ζ} 满足关于 $(\omega, \rho, \kappa, s)$ 的稳健零空间性质，其中 $\rho = 1/2$ 且$\kappa = \frac{3}{\sqrt{2}}\|(GW^{\zeta})^{-1}\| \lesssim 2^{(1-\zeta)bj_0}$。

下面的恒等式成立：

$$y = AW^{\zeta}z^{\dagger} + \varepsilon,$$

其中 $z^{\dagger} := W^{-\zeta}x^{\dagger}$。注意， $\|P^{\perp}_{\Lambda_{j_0}}z^{\dagger}\|_2 = \|W^{-\zeta}P^{\perp}_{\Lambda_{j_0}}x^{\dagger}\|_2 \leq 2^{-\zeta bj_0}r =: r'$。设 $\widehat{z}$是该问题的一个解

$$\min_{z\in\ell^2(\Lambda_{j_0})} \|z\|_{1,\omega}: \quad \|AW^{\zeta}z - y\|^2_{\mathcal{H}^m_2} \leq \left(\beta + C_3 2^{(\zeta-1)bj_0}r'\right)^2 = \left(\beta + C_3 2^{-bj_0}r\right)^2,$$

其中 C_3是稍后指定的常数。我们还引入记号 $\widehat{x} := W^{\zeta}\widehat{z}$。变量替换 $x = W^{\zeta}z$ 表明，这个问题等价于 (10)。

现在，我们按照定理 3.5 的证明继续推导。由于

$$\|W^{-\zeta}x^{\dagger} - W^{-\zeta}\widehat{x}\|_2 = \|z^{\dagger} - \widehat{z}\|_2 \leq \|P_{\Lambda_{j_0}}z^{\dagger} - \widehat{z}\|_2 + r',$$

只需得到 $\|P_{\Lambda_{j_0}}z^{\dagger} - \widehat{z}\|_2$的一个上界即可。

与测量算子 $(F_t\tilde{W}^{\zeta})$ 相关联的伴随前向映射为 $F\tilde{W}^{\zeta}$。利用准对角化性质 (6)，我们推得$\|F\tilde{W}^{\zeta}\| \leq C$，其中 C 是准对角化上界中的常数；因此，与该前向映射相关联的一个上界为 C。使用命题 5.8，并取情形 $G_1 = GD$ 和 $G_2 = GW^{\zeta}$ （注意，根据 (32)， $\|(GD)^{-1}\|$ 有一个下界，该下界由准对角化界确定，因为 $d_{j,n} \leq 2^{bj}$） ，可得以大于$1-\gamma$的概率，以下界成立：

$$\begin{aligned}
\|QAW^{\zeta}P^{\perp}_{\Lambda_{j_0}}z^{\dagger}\|_{\mathcal{H}^m_2} &\leq Cc_{\nu}^{-1/2}\|(GW^{\zeta})^{-1}\|^{-1}\left(\|F\Phi^*W^{\zeta}\iota_{\Gamma\setminus\Lambda_{j_0}}\|\|(GW^{\zeta})^{-1}\| + 1\right)r' \\
&= Cc_{\nu}^{-1/2}\left(\|F\Phi^*W^{\zeta}\iota_{\Gamma\setminus\Lambda_{j_0}}\| + \|(GW^{\zeta})^{-1}\|^{-1}\right)r' \\
&\leq Cc_{\nu}^{-1/2}2^{(\zeta-1)bj_0}r' = Cc_{\nu}^{-1/2}2^{-bj_0}r,
\end{aligned}$$

其中， $C > 0$ 仅取决于准对角化常数；我们还利用（32）对 $\|(GW^{\zeta})^{-1}\|^{-1} \lesssim 2^{(\zeta-1)bj_0}$ 进行了界定，并利用准对角化性质对 $\|F\Phi^*W^{\zeta}\iota_{\Gamma\setminus\Lambda_{j_0}}\| \lesssim 2^{(\zeta-1)bj_0}$ 进行了界定。令 $C_3 := Cc_{\nu}^{-1/2}$。按照定理 3.5 的证明，我们可以推断 $z^{\dagger}$ 满足最小化问题的约束；利用定理 5.3 中的距离界，即得结论。□

推论 3.14 的证明。只需在 $\zeta = 0$ 的情形下应用定理 3.11。 □

推论 3.15 的证明。只需在 $\zeta = 1$ 的情形下应用定理 3.11，并注意到，根据准对角化性质 (6)，

$$\|F\Phi^* x^\dagger - F\Phi^* \widehat{x}\|_{L^2_\mu(\mathcal{D};\mathcal{H}_2)} \asymp \|W^{-1}x^\dagger - W^{-1}\widehat{x}\|_2. \qquad \square$$

定理 3.17 的证明。我们先引入便于使用的记号

$$\tilde{m} := m/(\tilde{C}_0 B \log^3 \tau \log M) \gtrsim m/(\tilde{C}_0 B c j_0 \log^3 \tau),$$

其中 $C_0 = C_0 \log(1/\gamma)$，且 C_0, B 是定理 3.14 中出现的常数，而 $\tau := 2^{bj_0}s$ 对某个 $s \geq 3$ 成立。注意，根据 j_0 的选取，我们有 $j_0 \asymp \log(1/\beta)$，并且

$$\log \tau \leq 2bj_0 + \log s \leq bj_0 + \log M \asymp j_0 \asymp \log(1/\beta).$$

由此可得

$$\tilde{m} \gtrsim m/\log^4(1/\beta). \tag{33}$$

设 β 足够小，使得 $\log^3 \tau \log M \geq \log(1/\gamma)$——回想一下，$j_0 = j_0(\beta)$。根据定理 3.14，如果

$$\tilde{m} \geq 2^{2(b-d)j_0} 2^{2bj_0} s \geq \max_{(j,n)\in\Lambda_{j_0}} (d_{j,n}^{-2} 2^{2bj}) 2^{2bj_0} s$$

那么，以大于 $1-\gamma$ 的概率，以下误差估计成立 s:

$$\|x^\dagger - \widehat{x}\|_2 \leq C_1 \left(\frac{\sigma_s(x^\dagger)_1}{\sqrt{s}} + 2^{bj_0}\beta + r \right).$$

假设对于某个 $\zeta > 0$，有 $2^{-bj_0} \asymp \beta^\zeta$。因此 $\tilde{m} = \beta^{-2}{}_{(2b-d)b}$ ~~ζs~~，并且

$$\|x^\dagger - \widehat{x}\|_2 \leq C_1 \left(s^{-p} + \beta^{1-\zeta} + \beta^{\frac{a}{b}\zeta} \right).$$

施加 $1 - \zeta = a\zeta/b$，可得 $\zeta = b/(a+b)$，从而得到 $2^{-bj_0} \asymp \beta^{\frac{b}{a+b}}$。

结合 s 和 ζ 的表达式以及这些估计，可得

$$\|x^\dagger - \widehat{x}\|_2 \leq C_1 \left(\frac{\beta^{-2\frac{(2b-d)}{a+b}p}}{\tilde{m}^p} + \beta^{\frac{a}{a+b}} \right).$$

利用 (33)，即可得到命题中的第一个估计。

此外，若 $\tilde{m} = \beta^{-2\frac{(2b-d)}{a+b} - p\ \frac{a}{(a+b)}}$，则 $\tilde{m}^p = \beta^{-2\frac{(2b-d)}{a+b}p - \frac{a}{a+b}}$，因此

$$\frac{\beta^{-2\frac{(2b-d)}{a+b}p}}{\tilde{m}^p} = \beta^{\frac{a}{a+b}}.$$

第二个估计随即得出。

最后一个估计是通过对关系 $\tilde{m} = \tilde{m}(\beta)$ 求逆得到的，再次使用 (33)，并注意到 $\log\ m \asymp \log 1/\beta$。$\square$

5.7. 稀疏 Radon 变换。我们回顾第4节中的一些记号。

这两个问题的信号空间均由 $\mathcal{H}_1 = \overline{\text{span}(\phi_{j,n})_{(j,n)\in\Gamma}}$ 给出，而 $\mathcal{H}_2 = L^2(\mathbb{R})$ 是固定角度 $\theta \in [0, 2\pi)$ 下 Radon 变换的陪域，$\mathcal{H}_2 = L^2(-\pi/2, \pi/2)$ 则是角度 θ 下扇束变换的陪域。令 K 为如下定义的紧集：

$$K := \overline{\bigcup_{(j,n)\in\Gamma} \operatorname{supp} \phi_{j,n}},$$

使得 $\mathcal{H}_1 \subset L^2(K)$。

该模型的一个合适测量空间是 $(\mathcal{D}, \mu) = ([0, 2\pi), \mathrm{d}\theta)$，其中 $\mathrm{d}\theta$ 是 $[0, 2\pi$ 上的均匀概率测度。对于每个 $\theta \in \mathbb{S}^1$, $F_\theta : \mathcal{H}_1 \subset L^2(K) \to L^2(\mathbb{R})$ 要么由固定角度 $F_\theta = \mathcal{R}_\theta$ 处的 Radon 变换给出，要么由固定角度 $F_\theta = \mathcal{D}_\theta$发出的扇束变换给出。前向映射 F: $\mathcal{H}_1 \subset L^2(K) \to L^2([0, 2\pi) \times \mathbb{R})$ 表示为 Radon 变换 $F = \mathcal{R}$ 或扇束变换 $F = \mathcal{D}$。注意，Radon 变换的陪域$L^2([0, 2\pi)\times\mathbb{R})$ 可通过规范识别 $L^2([0, 2\pi) \times\mathbb{R}) \cong L^2([0, 2\pi) L^2(\mathbb{R}))$ 纳入我们的框架，扇束变换同样如此。

引理 5.10. 对某些依赖于 K 的常数 c_1、$C_1 > 0$，以下不等式成立：

$$c_1\|u\|_{H^{-1/2}} \le \|\mathcal{R}u\|_{L^2} \le C_1\|u\|_{H^{-1/2}}, \quad u \in L^2(K).$$

证明. 证明与 $[83,$ 定理 5.1$]$, 类似，其中针对 $u \in L^2(\mathcal{B}_1)$ 证明了这些不等式。□

拟对角化. 定理 5.10 意味着 Radon 变换满足条件 (7)，其中 $b = 1/2$。另一方面，如果小波具有注 3.9 中所述的足够正则性，则 Littlewood-Paley 性质 (8) 对 $b = 1/2$ 成立。于是，命题 3.8 得到以下准对角化性质：

$$c_1 \sum_{(j,n)\in\Gamma} 2^{-j}|x_{j,n}|^2 \le \|\mathcal{R}\Phi^* x\|_{L^2}^2 \le C_1 \sum_{(j,n)\in\Gamma} 2^{-j}|x_{j,n}|^2, \quad x \in \ell^2(\Gamma). \tag{34}$$

我们现在利用以下恒等式证明扇束变换的准对角化性质（例如，参见 $[83,$ 第 III.3$]$ 节或 $[92,$ 第 16.3$]$ 节）：

$$\mathcal{D}u(\theta, \alpha) = \mathcal{R}u(\theta + \alpha - \pi/2, \rho \sin\alpha).$$

我们有

$$\begin{aligned}\|\mathcal{D}u\|_{L^2}^2 &= \int_{\mathbb{S}^1}\int_{-\pi/2}^{\pi/2} |\mathcal{R}u(\theta + \alpha - \pi/2, \rho\sin\alpha)|^2 \mathrm{d}\alpha\mathrm{d}\theta \\ &= \int_{\mathbb{S}^1}\int_{-\pi/2}^{\pi/2} |\mathcal{R}u(\theta, \rho\sin\alpha)|^2 \mathrm{d}\alpha\mathrm{d}\theta\end{aligned}$$

$$= \int_{\mathbb{S}^1} \int_{-r}^{r} |\mathcal{R}(\theta, s)|^2 \frac{1}{\sqrt{\rho^2 - s^2}} \mathrm{d}s\mathrm{d}\theta,$$

这意味着

$$\rho^{-1/2} \|\mathcal{R}u\|_{L^2} \leq \|\mathcal{D}u\|_{L^2} \leq (\rho^2 - d^2)^{-1/4} \|\mathcal{R}u\|_{L^2}. \tag{35}$$

结合引理 5.10 和 (34)，可知 $\mathcal{D}$ 满足拟对角化性质，其常数为 $b = 1/2$，并且取决于所选的小波基以及 d 和 ρ。

相干性界。我们注意到，根据构造，对于每条直线 $\ell \subset \mathbb{R}^2$，$\operatorname{supp} \phi_{j,\ n} \cap \ell$ 都包含在一个长度为 $C2^{-j}$ 的线段内，其中 C 仅取决于所选的小波基。于是有

$$\begin{aligned}
\|\mathcal{R}_\theta \phi_{j,n}\|_{L^2(\mathbb{R})}^2 &= \int_{\mathbb{R}} \left| \int_{\theta^\perp} \phi_{j,n}(y + se_\theta) \mathrm{d}y \right|^2 \mathrm{d}s \\
&\leq C2^{-j} \int_{\mathbb{R}} \int_{\theta^\perp} |\phi_{j,n}(y + se_\theta)|^2 \mathrm{d}y\mathrm{d}s \\
&= C2^{-j} \|\phi_{j,n}\|_{L^2(\mathbb{R}^2)}^2 = C2^{-j}.
\end{aligned}$$

同理，有

$$\begin{aligned}
\|\mathcal{D}_\theta \phi_{j,n}\|_{L^2}^2 &= \int_{-\pi/2}^{\pi/2} \left| \int_{\mathbb{R}} \phi_{j,n}(\rho e_\theta + te_{\theta+\alpha}) \mathrm{d}t \right|^2 \mathrm{d}\alpha \\
&\leq C2^{-j} \int_{-\pi/2}^{\pi/2} \int_{\mathbb{R}} |\phi_{j,n}(\rho e_\theta + te_{\theta+\alpha})|^2 \, \mathrm{d}t\mathrm{d}\alpha \\
&= C2^{-j} \int_{-\pi/2}^{\pi/2} \int_{\mathbb{R}} |\phi_{j,n}(\rho e_\theta + te_{\theta+\alpha})|^2 \, t \, \frac{1}{t} \, \mathrm{d}t\mathrm{d}\alpha \\
&\leq \frac{C2^{-j}}{\rho - d} \int_{-\pi/2}^{\pi/2} \int_{\mathbb{R}} |\phi_{j,n}(\rho e_\theta + te_{\theta+\alpha})|^2 \, t \, \mathrm{d}t\mathrm{d}\alpha = \frac{C2^{-j}}{\rho - d}.
\end{aligned}$$

因此，在两种情况下，$d_{j,\ n} = 2^{j/2}$ 和一个常数 B 都满足假设 3.10；对于扇束变换，该常数取决于 d、ρ 以及所选的小波基：

$$\|F_\theta \phi_{j,n}\|_{L^2} \leq \frac{B}{2^{j/2}}, \quad \theta \in [0, 2\pi), \ (j, n) \in \Gamma. \tag{36}$$

在这两种情况下，准对角化性质与命题 5.9 一起推出以下估计：

$$\|(GD)^{-1}\|^2 \leq C \max_{(j,n) \in \Lambda_{j_0}} (2^j 2^{-j}) = C,$$

and

$$\|G^{-1}\|^2 \leq C2^{j_0}.$$

此外，根据 (15)，我们有 $\log M \asymp 2j_0$。

定理 4.4 的证明。只需将推论 3.14 应用于 Radon 场景，并利用准对角化性质 (34)、(35) 以及相干性界 (36)。□

定理 4.5 的证明。只需将推论 3.15 应用于 Radon 场景，并利用准对角化性质 (34)、(35) 以及相干性界 (36)。□

定理 4.6 的证明。只需将定理 3.17 应用于 Radon 场景，并利用准对角化性质 (34)、(35) 以及相干性界 (36)。□

(20) 的证明。我们主要沿用 [80, 第 9.3 节1]中的论证。根据紧支撑小波基的构造，可得 $\|\phi_{j,n}\|_{L^2} = 1$以及 $|\text{supp}(\phi_{j,n})| = C2^{-2j}$。因此，我们得到

$$|\langle u^\dagger, \phi_{j,n}\rangle| \leq \|u^\dagger\|_{L^\infty}\|\phi_{j,n}\|_{L^1} \leq C2^{-j}.$$

另一方面，根据注 3.9，结合 $b = -2$，可知 $|\langle u^\dagger,\phi_{j,n}\rangle| \lesssim 2^{-2j}$如果 $\text{supp}(\phi_{j,n})$ 与 $u^\dagger$ 的不连续性不相交。

令 Σ_j^1 为尺度 j 上支撑与 $u^\dagger$ 的不连续性相交的小波指标集合，并令 Σ_j^2 为尺度 j 上支撑与 $u^\dagger$ 的不连续性不相交的小波指标集合。参照第 9.3 节中关于分段正则图像的非线性逼近部分进行论证，可以看出 $|\Sigma_j^1| \asymp 2^j$ 和 $|\Sigma_j^2| \asymp 2^{2j}$。因此，我们有

$$\begin{aligned}
\|P_{\Lambda_{j_0}}^{\perp}\Phi u^\dagger\|_2^2 &= \sum_{j>j_0}\sum_n |\langle u^\dagger, \phi_{j,n}\rangle|^2 \\
&= \sum_{j>j_0}\left(\sum_{(j,n)\in\Sigma_j^1} |\langle u^\dagger, \phi_{j,n}\rangle|^2 + \sum_{(j,n)\in\Sigma_j^2} |\langle u^\dagger, \phi_{j,n}\rangle|^2\right) \\
&\lesssim \sum_{j>j_0}(2^j 2^{-2j} + 2^{2j}2^{-4j}) \lesssim 2^{-j_0}.
\end{aligned}$$

这意味着 $\|P_{\Lambda_j}^{\perp}\Phi u^\dagger\|_2 \leq C2^{-aj}$ 在 $a = 1/2$ 下成立。

根据类似的论证（参见 [80] 中的公式 (9.63)），可得 $|(\Phi u^\dagger)^{(i)}| \leq i^{-1}$，其中 $|(\Phi u^\dagger)^{(\cdot)}|$ 是 $|\Phi u^\dagger|$ 的非递增重排。无法推导出 p-可压缩性估计，因为 $\Phi u^\dagger$ 甚至可能不可求和，因此我们无法采用定理 4.6。相反，我们将利用这样一个事实：稀疏误差

定理 4.4 的恢复估计中的稀疏误差取决于截断信号 $P_{\Lambda_{j_0}}\Phi u^\dagger$。更确切地说，对于某个固定的稀疏度 $s \geq 1$，有：

$$\sigma_s(P_{\Lambda_{j_0}}\Phi u^\dagger)_1 \leq \sum_{i=s}^{M} |(\Phi u^\dagger)^{(i)}| \leq \sum_{i=s}^{M} \frac{1}{i} \lesssim \log(M/s) \leq \log(M),$$

其中 $M = |\Lambda_{j_0}|$。回想起 $M \asymp 2^{2j_0}$，并按照定理 4.6 将 j_0 选为 β 的函数，我们得到 $\log(M) \,.\, \log(1/\beta)$。因此我们得出

$$\sigma_s(P_{\Lambda_{j_0}}\Phi u^\dagger)_1 \lesssim \log(1/\beta) s^{1/2-p}$$

其中 $p = 1/2$。其余估计的推导方式与定理 4.6 的证明完全相同，由此得到 (20)。□

References

[1] B. Adcock、V. Antun 和 A. C. Hansen。《无限维压缩感知中的一致恢复及其在结构化二进制采样中的应用》。Appl. Comput. Harmon. Anal., 55:1–40, 2021年。[2] B. Adcock、S. Brugiapaglia、N. Dexter 和 S. Moraga。《利用有限样本计算高维希尔伯特空间值函数近乎最佳多项式近似的高效算法》。arXiv 预印本 arXiv:2203.13908, 2022年。[3] B. Adcock、S. Brugiapaglia 和 C. G. Webster。《高维函数的稀疏多项式近似》。工业与应用数学学会，费城，PA，2022年。[4] B. Adcock、A. Hansen 和 C. Poon。《超越一致重建：广义采样的最优性与尖锐界，以及在均匀重采样问题中的应用》。SIAM 数学分析杂志，45，01 2013年。[5] B. Adcock、A. Hansen、C. Poon 和 B. Roman。《打破相干性屏障：压缩感知的新理论》。Forum Math. Sigma, 5, 02 2017年。[6] B. Adcock、A. Hansen、B. Roman 和 G. Teschke。《广义采样：连续域上的稳定重建、逆问题与压缩感知》。成像与电子物理进展第182卷，第187–279页。爱思唯尔，2014年。[7] B. Adcock 和 A. C. Hansen。《任意基下稳定重建的广义采样定理》。傅里叶分析与应用杂志，18(4):685–716, 2012年。[8] B. Adcock 和 A. C. Hansen。《广义采样与无限维压缩感知》。计算数学基础，16(5):1263–1323, 2016年。[9] B. Adcock 和 A. C. Hansen。《压缩成像：结构、采样与学习》。剑桥大学出版社，剑桥，2021年。Vegard Antun 参与撰写。[10] B. Adcock、A. C. Hansen 和 C. Poon。《基于傅里叶样本的最优小波重建：稳定采样率的线性与普适性》。Appl. Comput. Harmon. Anal., 36(3):387–415, 2014年。[11] G. S. Alberti、A. Arroyo 和 M. Santacesaria。《低维流形上的逆问题》。《非线性》，36(1):734–808, 2023年。[12] G. S. Alberti、P. Campodonico 和 M. Santacesaria。《压缩感知光声断层成像可归结为欠采样傅里叶测量的压缩感知》。SIAM J. Imaging Sci., 14(3):1039–1077, 2021年。

[13] G. S. Alberti 和 M. Santacesaria。《有限次测量下的 Calder´on 逆问题》。Forum Math. Sigma, 7:e35, 2019年。[14] G. S. Alberti 和 M. Santacesaria。《基于各向异性测量的无限维压缩感知及其在偏微分方程逆问题中的应用》。Appl. Comput. Harmon. Anal., 50:105–146, 2021年。[15] G. S. Alberti 和 M. Santacesaria。《有限次测量下的无限维逆问题》。Arch. Ration. Mech. Anal., 243(1):1–31, 2022年。[16] A. Bastounis 和 A. C. Hansen。《压缩感知的诸多现实应用中不存在一致恢复，以及受限等距性质和分层零空间性质》。SIAM J. Imaging Sci., 10(1):335–371, 2017年。[17] M. Benning 和 M. Burger。《逆问题的现代正则化方法》。Acta Numer., 27:1–111, 2018年。.[18] A. Berk、S. Brugiapaglia、B. Joshi、Y. Plan、M. Scott 和 O. Yilmaz。《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年。[19] P. J. Bickel、Y. Ritov 和 A. B. Tsybakov。《Lasso 与 Dantzig 选择器的同步分析》。Ann. Statist., 37(4):1705–1732, 2009年。[20] J. Bigot、C. Boyer 和 P. Weiss。《压缩感知中的分块采样策略分析》。IEEE Transactions on Information Theory, 62(4):2125–2139, 2016年。[21] G. Blanchard 和 N. M¨ucke。《统计逆学习问题正则化的最优速率》。计算数学基础，18(4):971–1013, 2018年。[22] T. Blumensath。《非线性观测下的压缩感知及相关非线性优化问题》。IEEE Trans. Inform. Theory, 59(6):3466–3474, 2013年。[23] A. Bora、A. Jalal、E. Price 和 A. G. Dimakis。《使用生成模型的压缩感知》。载于 D. Precup 和 Y. W. Teh 编，《第34届国际机器学习会议论文集》，Proceedings of Machine Learning Research 第70卷，第537–546页，澳大利亚悉尼国际会议中心，2017年8月6–11日。PMLR。[24] L. Borcea。《电阻抗断层成像》。Inverse Problems, 18(6):R99–R136, 2002年。[25] K. Bredies 和 D. A. Lorenz。《迭代软阈值法的线性收敛》。傅里叶分析与应用杂志，14(5-6):813–837, 2008年。[26] S. Brugiapaglia、S. Dirksen、H. C. Jung 和 H. Rauhut。《有界 Riesz 系统中的稀疏恢复及其在偏微分方程数值方法中的应用》。Appl. Comput. Harmon. Anal., 53:231–269, 2021年。[27] T. A. Bubba、M. Burger、T. Helin 和 L. Ratti。《统计逆学习问题中的凸正则化》。Inverse Probl. Imaging, 17(6):1193–1225, 2023年。[28] T. A. Bubba、T. Heikkil¨a、H. Help、S. Huotari、Y. Salmon 和 S. Siltanen。《稀疏动态断层成像：一种用于植物茎碘灌注成像的基于剪切波的方案》。Inverse Problems, 36(9):094002, 2020年9月。[29] T. A. Bubba 和 L. Ratti。《统计逆学习中的基于剪切波的正则化及其在X射线断层成像中的应用》。Inverse Problems, 38(5):论文编号054001，43页，2022年。[30] T. A. Bubba、S. Siltanen、M. Lassas、M. M¨arz 和 Z. Purisha。《稀疏动态断层成像中的基于剪切波的正则化》。31页，2017年8月。[31] F. Cakoni、D. Colton 和 H. Haddar。《逆散射理论与透射特征值》，CBMS-NSF 应用数学区域会议系列第88卷。工业与应用数学学会（SIAM），费城，PA，2016年。

[32] E. Cand`es 和 J. Romberg。《压缩采样中的稀疏度与不相干性》。Inverse problems，23(3):969，2007年。[33] E. J. Cand`es。《受限等距性质及其对压缩感知的意义》。巴黎科学院数学报告，346(9):589–592，2008年。[34] E. J. Cand`es 和 D. Donoho。《曲波的新紧框架及具有 $c^2$ 奇异性的对象的最优表示》。《纯粹与应用数学通讯》，57:219 – 266，2004年02月。[35] E. J. Cand`es 和 D. L. Donoho。《恢复不适定逆问题中的边缘：曲波框架的最优性》。统计年鉴，30(3):784–842，2002年。[36] E. J. Cand`es、Y. C. Eldar、D. Needell 和 P. Randall。《利用相干且冗余的字典进行压缩感知》。Appl. Comput. Harmon. Anal.，31(1):59–73，2011年。[37] E. J. Cand`es 和 Y. Plan。《压缩感知的概率理论与无受限等距性质理论》。IEEE Transactions on Information Theory，57(11):7235–7254，2011年。[38] E. J. Cand`es、J. Romberg 和 T. Tao。《稳健不确定性原理：从高度不完整的频率信息中精确重建信号》。IEEE Trans. Inform. Theory，52(2):489–509，2006年。[39] E. J. Cand`es 和 T. Tao。《通过线性规划解码》。IEEE Trans. Inform. Theory，51(12):4203–4215，2005年。[40] A. Cohen、I. Daubechies 和 P. Vial。《区间上的小波与快速小波变换》。Appl. Comput. Harmon. Anal.，1(1):54–81，1993年。[41] F. Colonna、G. Easley、K. Guo 和 D. Labate。《利用剪切波表示反演 Radon 变换》。Appl. Comput. Harmon. Anal.，29:232–250，2010年09月。[42] D. Colton 和 R. Kress。《反声学与电磁散射理论》，应用数学科学第93卷。施普林格，纽约，第三版，2013年。[43] I. Daubechies。《小波十讲》。工业与应用数学学会，1992年。[44] R. A. DeVore。《非线性逼近》。载于《数值学报》1998年第7卷，Acta Numer.，第51–150页。剑桥大学出版社，剑桥，1998年。[45] N. Dexter、H. Tran 和 C. Webster。《参数化偏微分方程稀疏同时逼近的一种混合 ℓ_1 正则化方法》。ESAIM Math. Model. Numer. Anal.，53(6):2025–2045，2019年。[46] D. L. Donoho。《通过小波–模糊分解非线性求解线性逆问题》。Appl. Comput. Harmon. Anal.，2(2):101–126，1995年。[47] D. L. Donoho。《图像的稀疏分量与最优原子分解》。Constr. Approx.，17(3):353–382，2001年。[48] D. L. Donoho。《压缩感知》。IEEE Transactions on Information Theory，52(4):1289–1306，2006年。[49] C. Dossal、V. Duval 和 C. Poon。《沿径向直线采样傅里叶变换》。SIAM J. Numer. Anal.，55(6):2540–2564，2017年。[50] M. Elad、P. Milanfar 和 R. Rubinstein。《信号先验中的分析与合成》。载于《2006年第14届欧洲信号处理会议》，第1–5页，2006年。[51] Y. C. Eldar 和 G. Kutyniok（编）。《压缩感知》。剑桥大学出版社，剑桥，2012年。理论与应用。[52] M. Fornasier 和 H. Rauhut。《压缩感知》。载于《成像数学方法手册》第1、2、3卷，第205–256页。施普林格，纽约，2015年。[53] S. Foucart。《威布尔矩阵与冗余字典下 ℓ^1-最小化的稳定性与稳健性》。Linear Algebra and its Applications，441:4–21，2014年。

[54] S. Foucart 和 H. Rauhut。《压缩感知数学导论》，应用与数值调和分析丛书。Birkh¨auser/施普林格，纽约，2013年。[55] H. Garde 和 K. Knudsen。《部分数据下电阻抗断层成像的稀疏性先验》。Inverse Probl. Sci. Eng.，24(3):524–541，2016年。[56] A. C. Gilbert、H. W. Levinson 和 J. C. Schotland。《用于逆散射的非线性迭代硬阈值法》。《SIAM成像科学期刊》，13(1):108–140，2020年1月。[57] M. Giordano 和 R. Nickl。《椭圆型反问题中采用高斯过程先验的贝叶斯推断的一致性》。Inverse Problems，36(8):085001，35，2020年。[58] M. Grasmair、M. Haltmeier 和 O. Scherzer。《带有 ℓ^q 惩罚项的稀疏正则化》。Inverse Problems，24(5):055020，13，2008年。[59] M. Grasmair、M. Haltmeier 和 O. Scherzer。《 ℓ^1-正则化线性收敛的充要条件》。Comm. Pure Appl. Math.，64(2):161–182，2011年。[60] K. Guo 和 D. Labate。《利用剪切波实现最优稀疏多维表示》。SIAM 数学分析杂志，39(1):298–318，2007年。[61] K. H¨am¨al¨ainen、A. Kallonen、V. Kolehmainen、M. Lassas、K. Niinim¨aki 和 S. Siltanen。《稀疏断层成像》。SIAM J. Sci. Comput.，35(3):B644–B665，2013年。[62] P. C. Hansen、J. Jørgensen 和 W. R. B. Lionheart。《计算机断层成像：算法、洞见与恰到好处的理论》。工业与应用数学学会，费城，PA，2021年。[63] B. Harrach。《有限个电极条件下电阻抗断层成像的唯一性与利普希茨稳定性》。Inverse Problems，35(2):024005，2019年。[64] E. Herrholz 和 G. Teschke。《用于逆问题与不适定问题的压缩感知原理和迭代稀疏恢复》。Inverse Problems，26(12):125012，24，2010年。[65] V. Isakov。《偏微分方程的反问题》，应用数学科学第127卷。施普林格，尚，第三版，2017年。[66] B. Jin 和 P. Maass。《参数识别问题的稀疏性正则化》。Inverse Problems，28(12):123001，70，2012年。[67] A. D. Jones、B. Adcock 和 A. C. Hansen。《通过相干性分析多维压缩感知问题的结构》。CoRR，abs/1610.07497，2016年。[68] J. Jørgensen、S. Coban、W. Lionheart、S. McDonald 和 P. Withers。《SparseBeads数据：稀疏正则化计算机断层成像的基准测试》。Meas. Sci. Technol.，28，2017年09月。[69] J. S. Jørgensen、S. B. Coban、W. R. B. Lionheart、S. A. McDonald 和 P. J. Withers。《SparseBeads数据：稀疏正则化计算机断层成像的基准测试》。Meas. Sci. Technol.，28(12):124005，2017年11月。[70] J. S. Jørgensen 和 E. Y. Sidky。《需要多少数据才够？稀疏正则化X射线计算机断层成像的相图分析》。Phil. Trans. R. Soc. A.，373(2043):20140387，2015年。[71] J. S. Jørgensen、E. Y. Sidky、P. C. Hansen 和 X. Pan。《计算机断层成像中稀疏图像欠采样可恢复性的定量研究》。2012年。[72] J. Kaipio 和 E. Somersalo。《统计反问题：离散化、模型降阶与反问题犯罪》。J. Comput. Appl. Math.，198(2):493–504，2007年。[73] A. Kirsch。《反问题数学理论导论》，应用数学科学第120卷。施普林格，纽约，第二版，2011年。[74] F. Krahmer、D. Needell 和 R. Ward。《利用冗余字典和结构化测量进行压缩感知》。载于《2015年国际采样理论与应用会议（SampTA）》，第25–29页，2015年。[75] G. Kutyniok 和 W.-Q. Lim。《紧支撑剪切波具有最优稀疏性》。J. Approx. Theory，163(11):1564–1589，2011年。

[76] P. E. Z. Larson、S. Hu、M. Lustig、A. B. Kerr、S. J. Nelson、J. Kurhanewicz、J. M. Pauly 和 V. D. B. 面向超极化碳-13研究的快速动态三维MR波谱成像：采用压缩感知和多频带激发脉冲。磁共振医学，65(3):610–9，2011年。[77] C. Li 和 B. Adcock。具有局部结构的压缩感知：分层稀疏类的一致恢复保证。Appl. Comput. Harmon. Anal.，46(3):453–477，2019年。[78] M. Lustig、D. Donoho 和 J. M. Pauly。稀疏MRI：压缩感知在快速MR成像中的应用。磁共振医学，58(6):1182–1195，2007年。[79] M. Lustig、D. L. Donoho、J. M. Santos 和 J. M. Pauly。压缩感知MRI。IEEE信号处理杂志，25(2):72–82，2008年。[80] S. Mallat。《信号处理小波之旅：稀疏之道》。爱思唯尔/学术出版社，阿姆斯特丹，第3版，2009年。[81] Y. Meyer。《小波与算子》。剑桥大学出版社，剑桥，1992年。[82] J. L. Mueller 和 S. Siltanen。《线性与非线性反问题及其实际应用》，计算科学与工程第10卷。工业与应用数学学会（SIAM），费城，宾夕法尼亚州，2012年。[83] F. Natterer。《计算机断层成像数学》。工业与应用数学学会，2001年。[84] F. Natterer。X射线断层成像。载于《逆问题与成像》，数学讲义第1943卷，第17–34页。施普林格，柏林，2008年。[85] H. Ohlsson、A. Yang、R. Dong 和 S. Sastry。CPRL——将压缩感知扩展至相位恢复问题。载于 F. Pereira、C. J. C. Burges、L. Bottou 和 K. Q. Weinberger（编），《神经信息处理系统进展25》，第1367–1375页。Curran Associates公司，2012年。[86] A. C. Polak、M. F. Duarte 和 D. L. Goeckel。压缩感知中分组非相干测量的性能界。IEEE信号处理汇刊，63(11):2877–2887，2015年。[87] C. Poon。压缩感知中的结构依赖采样：紧框架的理论保证。Appl. Comput. Harmon. Anal.，42(3):402–451，2017年。[88] Z. Purisha、S. Karhula、J. Ketola、J. Rimpel¨ainen、M. Nieminen、S. Saarakkala、H. Kr¨oger 和 S. Siltanen。自动正则化方法：稀疏数据在三维X射线微型CT重建中的应用。IEEE医学成像汇刊，38:1–1，2018年08月。[89] A. Rastogi。用于非线性统计反问题的过度平滑惩罚Tikhonov 正则化。纯粹与应用分析通讯，19(8):4111–4126，2020年。[90] A. Rastogi、G. Blanchard 和 P. Math´e。非线性统计反问题的Tikhonov 正则化收敛性分析。电子统计学杂志，14(2):2798–2841，2020年。[91] H. Rauhut 和 W. Rachel。通过加权 ℓ^1 最小化进行插值。Appl. Comput. Harmon. Anal.，40(2):321–351，2016年。[92] O. Scherzer（编）。《成像数学方法手册》，第1、2、3卷。施普林格，纽约，第2版，2015年。[93] O. Scherzer、M. Grasmair、H. Grossauer、M. Haltmeier 和 F. Lenzen。《成像中的变分方法》，应用数学科学第167卷。施普林格，纽约，2009年。[94] S. Siltanen、V. Kolehmainen、S. J¨arvenp¨a¨a、J. P. Kaipio、P. Koistinen、M. Lassas、J. Pirttil¨a 和 E. Somersalo。少量射线照片下医学X射线断层成像的统计反演：I. 一般理论。物理医学与生物学，48(10):1437，2003年。[95] A. M. Stuart。逆问题：贝叶斯视角。Acta Numer.，19:451–559，2010年。[96] Y. Tsaig 和 D. L. Donoho。压缩感知扩展。信号处理，86(3):549–571，2006年。《信号与图像处理中的稀疏逼近》。

[97] G. Uhlmann。电阻抗断层成像与 Calder´on 问题。《逆问题》，25(12):123011,39，2009年。
[98] V. A. Zorich。《数学分析 II》。Universitext，施普林格，柏林、海德堡，2016年。

AppendixA. Proof of Theorem 5.7

我们首先回顾 Bernstein-Talagrand 不等式的一个变体，稍后将会用到。

引理 A.1（Bernstein 不等式 [54, 定理 8.42]）。设 $f_x\colon \mathbb{C}^M \to \mathbb{R}$ 是以 $x \in \mathcal{B}_*$ 为指标的可数函数族。设 $Y_1, \dots, Y_m$ 是相互独立的随机变量，使得对每个 $k = 1, \dots, m$ 以及 $x \in \mathcal{B}_*$，以下条件均成立：

$$\mathbb{E} f_x(Y_k) = 0; \quad |f_x(Y_k)| \le K \ \ a.s.; \quad \mathbb{E}|f_x(Y_k)|^2 \le \Sigma^2.$$

考虑随机变量

$$Z \coloneqq \sup_{x \in \mathcal{B}_*} \sum_{k=1}^{m} f_x(Y_k).$$

那么，对每个 $\varepsilon > 0$,

$$\mathbb{P}(Z \ge \mathbb{E}Z + \varepsilon) \le \exp\left(-\frac{\varepsilon^2}{2(m\Sigma^2 + 2K\mathbb{E}Z) + 2K\varepsilon/3}\right).$$

A.1. 情形 $f_\nu = 1$。我们先考虑 $f_\nu = 1$ 的情形，此时 $\mu = \nu$ 是一个概率测度。因此，假设 3.1 变为

$$\|F_t \phi_i\|_{\mathcal{H}_2} \le B\omega_i, \quad t \in \mathcal{D}, \ i \in \Gamma. \tag{37}$$

此外，我们有 $Q = I$。

考虑定义在 $\mathbb{C}^{M \times M}$ 上的以下半范数：

$$\|B\|_{G,\omega,\lambda} \coloneqq \sup_{x \in \mathcal{B}} |\langle Bx, x\rangle|,$$

其中，我们令

$$\mathcal{B} \coloneqq \{x \in \mathbb{C}^M \colon \ \|Gx\|_2 \le 1, \ \|x\|_{0,\omega} \le \lambda\}.$$

通过标准论证，可以将满足关于 $(G, \omega, \delta, \lambda)$ 的广义限制等距性质的 A 所对应的最小 $\delta > 0$ 表征为

$$\delta_* \coloneqq \|A^*A - G^*G\|_{G,\omega,\lambda}. \tag{38}$$

此外，令

$$X_t \coloneqq F_t \Phi^* \iota_\Lambda, \quad t \in \mathcal{D}.$$

初等计算表明

$$A^*A = \frac{1}{m}\sum_{k=1}^{m} X_{t_k}^* X_{t_k}, \tag{39}$$

$$G^*G = \mathbb{E}X_t^* X_t. \tag{40}$$

现在的目标是给出 $\mathbb{E}\delta_* = \mathbb{E}\|A^*A - G^*G\|_{G,\,\omega,\,\lambda}$ 的一个界。为此，我们引入一个 Rademacher 向量 $(\varepsilon_1, \ldots, \varepsilon_m)$，它与样本 $(t_1, \ldots, t_m)$ 相互独立，并考虑以下随机变量，其中 $u \in \mathcal{B}$:

$$Z_u := \sum_{k=1}^{m} \varepsilon_k \|X_{t_k} u\|_{\mathcal{H}_2}^2.$$

不难验证，$\{Z_u\colon\ u \in \mathcal{B}\}$ 是一个中心对称的次高斯随机过程 [54, 第 8 节6]，其度量为以下伪度量:

$$d(u,v) := \left(\sum_{k=1}^{m} (\|X_{t_k} u\|_{\mathcal{H}_2}^2 - \|X_{t_k} v\|_{\mathcal{H}_2}^2)^2\right)^{1/2}.$$

因此，我们可以使用 Dudley 不等式。

命题 A.2. 以下估计成立:

$$\mathbb{E}\delta_* \le \frac{16\sqrt{2}}{m}\mathbb{E}\int_0^\infty \sqrt{\log \mathcal{N}(\mathcal{B}, d, y)}\,\mathrm{d}y.$$

证明. 根据 (38)、(39) 和 (40)，我们有

$$\mathbb{E}\delta_* = \mathbb{E}\left\|\frac{1}{m}\sum_{k=1}^{m}(X_{t_k}^* X_{t_k} - \mathbb{E}X_{t_k}^* X_{t_k})\right\|_{G,\omega,\lambda}.$$

使用标准的对称化技巧（例如，参见 [54, 引理 8.4]），我们得到

$$\mathbb{E}\delta_* \le \frac{2}{m}\mathbb{E}\left\|\sum_{k=1}^{m}\varepsilon_k X_{t_k}^* X_{t_k}\right\|_{G,\omega,\lambda} = \frac{2}{m}\mathbb{E}\sup_{u\in\mathcal{B}}\left|\sum_{k=1}^{m}\varepsilon_k\|X_{t_k}u\|_{\mathcal{H}_2}^2\right|.$$

所需结论可直接由 Dudley 不等式推出。 □

为了获得可处理的覆盖数估计，我们首先对底层度量进行略微修改。在此之前，我们需要证明一个涉及该常数的技术性估计

$$\tau := B^2\|G^{-1}\|^2\lambda.$$

引理 A.3. 假设（37）成立。则

$$\|X_t x\|_{\mathcal{H}_2} \le \sqrt{\tau}, \quad t \in \mathcal{D},\ x \in \mathcal{B}.$$

证明。设 $t \in \mathcal{D}$, $x \in \mathcal{B}$。设 $S :=\text{supp}(x)$。利用 (37)，可得

$$\begin{aligned}\|F_t \Phi^* x\|_{\mathcal{H}_2} &\le \sum_{i\in S} |x_i| \|F_t \Phi^* e_i\|_{\mathcal{H}_2} \le \sum_{i\in S} |x_i| B\omega_i \\ &\le \|x\|_2 B\sqrt{\omega(S)} \le \|G^{-1}\| \|Gx\|_2 B\sqrt{\lambda} \le \sqrt{\tau}.\end{aligned}$$

□

对于 $1 < p < \infty$，我们定义如下半范数：

$$\|u\|_{X,p} \coloneqq \left(\sum_{k=1}^m \|X_{t_k} u\|_{\mathcal{H}_2}^{2p}\right)^{1/2p}, \quad u \in \mathbb{C}^M.$$

我们将相应的伪距离记为 $d_{X,\ p}$。

引理 A.4。设 $1 < r < \infty$ 和 r'，满足 $\frac{1}{r} + \frac{1}{r'} = 1$。则有

$$d(u,v) \le 2\tau^{(r-1)/2r} m^{1/2r} \|A^*A\|_{G,\omega,\lambda}^{1/2r} d_{X,r'}(u,v), \qquad u, v \in \mathcal{B}.$$

证明。根据 H¨older 不等式，可得

$$\begin{aligned}d(u,v) &= \left(\sum_{k=1}^m \left(\|X_{t_k} u\|_{\mathcal{H}_2}^2 - \|X_{t_k} v\|_{\mathcal{H}_2}^2\right)^2\right)^{1/2} \\ &= \left(\sum_{k=1}^m \left(\|X_{t_k} u\|_{\mathcal{H}_2} + \|X_{t_k} v\|_{\mathcal{H}_2}\right)^2 \left(\|X_{t_k} u\|_{\mathcal{H}_2} - \|X_{t_k} v\|_{\mathcal{H}_2}\right)^2\right)^{1/2} \\ &\le \left(\sum_{k=1}^m \left(\|X_{t_k} u\|_{\mathcal{H}_2} + \|X_{t_k} v\|_{\mathcal{H}_2}\right)^{2r}\right)^{1/2r} \left(\sum_{k=1}^m \left(\|X_{t_k} u\|_{\mathcal{H}_2} - \|X_{t_k} v\|_{\mathcal{H}_2}\right)^{2r'}\right)^{1/2r'} \\ &\le 2 \sup_{w\in\mathcal{B}} \left(\sum_{k=1}^m \|X_{t_k} w\|_{\mathcal{H}_2}^{2r}\right)^{1/2r} \left(\sum_{k=1}^m \left(\|X_{t_k}(u-v)\|_{\mathcal{H}_2}\right)^{2r'}\right)^{1/2r'} \\ &= 2 \sup_{w\in\mathcal{B}} \left(\sum_{k=1}^m \|X_{t_k} w\|_{\mathcal{H}_2}^{2r}\right)^{1/2r} d_{X,r'}(u,v).\end{aligned}$$

由引理 A.3 和式（39），我们得到

$$\sup_{w\in\mathcal{B}} \left(\sum_{k=1}^m \|X_{t_k} w\|_{\mathcal{H}_2}^{2r}\right)^{1/2r} = \sup_{w\in\mathcal{B}} \left(\sum_{k=1}^m \|X_{t_k} w\|_{\mathcal{H}_2}^2 \|X_{t_k} w\|_{\mathcal{H}_2}^{2(r-1)}\right)^{1/2r}$$

$$\leq \tau^{(r-1)/2r} \sup_{w\in\mathcal{B}} \left(\sum_{k=1}^{m} \|X_{t_k} w\|_{\mathcal{H}_2}^2 \right)^{1/2r}$$
$$= \tau^{(r-1)/2r} m^{1/2r} \|A^* A\|_{G,\omega,\lambda}^{1/2r}. \qquad \square$$

刚刚证明的 $d(u,v)$ 与 $d_{X,\ r}(u,v)$ 之间的关系，反映为相应覆盖数之间的一个不等式。

命题 A.5. 对任意 $r > 1$，我们有

$$\int_0^\infty \sqrt{\log \mathcal{N}(\mathcal{B}, d, y)}\, \mathrm{d}y \leq C(r) \int_0^\infty \sqrt{\log \mathcal{N}(\mathcal{B}, d_{X,r'}, y)}\, \mathrm{d}y,$$

其中 $C(r) := 2\tau^{(r-1)/2r} m^{1/2r} \|A^* A\|_{G\,\omega\ ,,\lambda}^{1/2r}$。

证明. 引理 A.4 蕴含

$$d(u,v) \leq C(r) d_{X,r'}(u,v), \quad u, v \in \mathcal{B}.$$

St覆盖数的标准性质（例如，参见 [54]）意味着

$$\mathcal{N}(\mathcal{B}, d, y) \leq \mathcal{N}(\mathcal{B}, C(r) d_{X,r'}, y) = \mathcal{N}(\mathcal{B}, d_{X,r'}, y/C(r)),$$

因此

$$\int_0^\infty \sqrt{\log \mathcal{N}(\mathcal{B}, d, y)} \mathrm{d}y \leq \int_0^\infty \sqrt{\log \mathcal{N}(\mathcal{B}, d_{X,r'}, y/C(r))} \mathrm{d}y.$$

代入 $y' = y/C(r)$ 最终证明了这一结论。 $\square$

由于度量发生了修改，我们相应地改变了底层空间。考虑集合

$$T_\omega^{\lambda,M} := \{x \in \mathbb{C}^M : \ \|x\|_2 \leq 1, \ \|x\|_{0,\omega} \leq \lambda\}.$$

容易证明

$$\mathcal{B} \subseteq \|G^{-1}\| T_\omega^{\lambda,M}.$$

命题 A.6. 以下估计成立：

$$\int_0^\infty \sqrt{\log \mathcal{N}(\mathcal{B}, d_{X,r'}, y)} \mathrm{d}y \leq \|G^{-1}\| \int_0^\infty \sqrt{\log \mathcal{N}(T_\omega^{\lambda,M}, d_{X,r'}, y)} \mathrm{d}y.$$

证明。只需将恒等式 $\mathcal{N}(\mathcal{B}, Cd, y) = \mathcal{N}(\mathcal{B}, d, y/C)$ 应用于 $\mathcal{B} \subseteq \|G^{-1}\| T^{\lambda}_{\ ,\ \omega\ M}$，相对于 $d_{X,r'}$，然后考虑代换 $y' = y/\|G^{-1}\|$。 $\square$

迄今为止，为改变覆盖数的底层结构所付出的努力，因将问题化简为更容易处理的情形而得到了回报，进而使我们能够得到一个方便的上界。事实上，以下结果的证明与 [91, 定理5.2],中的证明基本相同，只需做少量调整，因此略去。

命题 A.7. 以下估计成立：

$$\int_0^\infty \sqrt{\log \mathcal{N}(T_\omega^{\lambda,M}, d_{X,r'}, y)}\mathrm{d}y \le C_1 \sqrt{r' m^{1/r'} \lambda \log^2 \lambda \log M},$$

其中 $C_1 > 0$ 是一个普适常数。

至此，我们可以证明关于 $\mathbb{E}\delta_*$的所需界。

命题 A.8。以下估计成立：

$$\mathbb{E}\delta_* \le C_2 \sqrt{\frac{\tau \log^3 \tau \log M}{m}} \sqrt{\mathbb{E}\delta_* + 1},$$

其中 $C_2 > 0$ 是一个普适常数。

证明。结合命题 A.2 与迄今为止证明的所有估计，我们最终得到

$$\mathbb{E}\delta_* \le C_2' m^{-1+1/2r+1/2r'} \tau^{(r-1)/2r} \|G^{-1}\| (r' \lambda \log^2 \lambda \log M)^{1/2} \mathbb{E}\|A^*A\|_{G,\omega,\lambda}^{1/2r},$$

其中，$C_2' > 0$ 是一个普适常数。注意到 $m^{-1+1/2r+1/2r'} = m^{-1/2}$。鉴于 r的选择具有任意性，不妨设

$$r = 1 + (\log \tau)^{-1} \quad \Rightarrow \quad r' = 1 + \log \tau.$$

因此，我们得到

$$\tau^{(r-1)/2r} = \tau^{1/2r'} = \tau^{1/2(1+\log \tau)} \le e^{1/2},$$

并利用$\log\ \tau \ge \log 3 \ge 1$,

$$\sqrt{r'} = \sqrt{1 + \log \tau} \le \sqrt{2} \sqrt{\log \tau}.$$

最后，注意到

$$\|A^*A\|_{G,\omega,\lambda} \le \|A^*A - G^*G\|_{G,\omega,\lambda} + \|G^*G\|_{G,\omega,\lambda}.$$

特别地，对于 $x \in \mathcal{B}$，我们有

$$|\langle G^*Gx, x\rangle| = \|Gx\|_2^2 \le 1,$$

因此$\|G^*G\|_{G,\ \omega,\ \lambda} \le 1$。因此，

$$\begin{aligned} \|A^*A\|_{G,\omega,\lambda}^{1/2r} &\le \left(\|A^*A - G^*G\|_{G,\omega,\lambda} + 1\right)^{1/2r} \\ &\le \left(\|A^*A - G^*G\|_{G,\omega,\lambda} + 1\right)^{1/2}. \end{aligned}$$

根据詹森不等式和（38），我们得出

$$\mathbb{E}\|A^*A\|_{G,\omega,\lambda}^{1/2r} \le \mathbb{E}\left(\|A^*A - G^*G\|_{G,\omega,\lambda} + 1\right)^{1/2} \le \sqrt{\mathbb{E}\delta_* + 1}. \qquad \square$$

根据刚刚证明的界，很容易推导出使 $\mathbb{E}\delta_*$ 足够小所需的样本数 m 的下界。

命题 A.9。存在一个普适常数 $C > 0$，使得对于任意 $\delta \in (0, 1)$，

$$m \geq C\delta^{-2}\tau \log^3 \tau \log M \quad \Rightarrow \quad \mathbb{E}\delta_* \leq \delta/2.$$

证明。显然，对于任意 $C \geq 16C_2^2$，这些假设意味着

$$C_2\sqrt{\frac{\tau \log^3 \tau \log M}{m}} \leq \frac{\delta}{4}.$$

由命题 A.8 可知，这意味着

$$\mathbb{E}\delta_* \leq \frac{\delta}{4}\sqrt{\mathbb{E}\delta_* + 1}. \tag{41}$$

直接计算可得，若 $a \in (0, 1)$ 且 $x > 0$，则

$$x \leq a\sqrt{x+1} \quad \Rightarrow \quad x \leq 2a.$$

因此，由 (41) 可推得

$$\mathbb{E}\delta_* \leq \delta/2,$$

这就是所要证明的结论。 □

一旦建立了δ_* 偏差的界限，就可以最终证明定理 5.7 的结论，其思路与引理 A.1 中给出的集中不等式一致。为此，首先注意到映射

$$\mathbb{C}^M \ni x \mapsto |\langle (A^*A - G^*G)x, x\rangle| \in \mathbb{R}$$

是连续的。因此，用 $\mathcal{B}'$ 表示 $\mathcal{B}$ 的一个可数稠密子集，我们得到

$$\delta_* = \sup_{x \in \mathcal{B}'} |\langle (A^*A - G^*G)x, x\rangle|.$$

此外，令 $\mathcal{P} := \{\pm 1\}$，即可写出

$$\delta_* = \sup_{(x,p)\in \mathcal{B}'\times\mathcal{P}} p\langle (A^*A - G^*G)x, x\rangle,$$

而且，由 (39) 可进一步得到

$$\delta_* = \frac{1}{m}\sup_{(x,p)\in \mathcal{B}'\times\mathcal{P}} \sum_{k=1}^{m} f_{x,p}(X_{t_k}), \tag{42}$$

其中，我们令 $f_{x,p}(B) := p\langle (B^*B - G^*G)x, x\rangle$。

现在我们可以为 δ_* 推导一个尾界。事实上，下面将说明，伯恩斯坦型不等式在引理 A.1 中的假设均已满足。

引理 A.10. 对于每个 $k = 1, \ldots, m$，以及对于 $(x, p) \in \mathcal{B}' \times \mathcal{P}$，

$$\mathbb{E}f_{x,p}(X_{t_k}) = 0, \qquad |f_{x,p}(X_{t_k})| \leq \tau, \qquad \mathbb{E}|f_{x,p}(X_{t_k})|^2 \leq \tau.$$

证明. $f_{x,p}(X_{t_k})$ 的零均值性质由 (40) 得出。

关于 $|\mathbb{E}f_{x,p}(X_{t_k})|$ 的界，引理 A.3 表明

$$\|X_{t_k}x\|_{\mathcal{H}_2}^2 \le \tau, \quad x \in \mathcal{B}.$$

另一方面，我们有

$$\|Gx\|_2^2 \le 1 \le \tau, \quad x \in \mathcal{B}.$$

因此，对于 $(x,p) \in \mathcal{B}' \times \mathcal{P}$，我们得到

$$\begin{aligned} |f_{x,p}(X_{t_k})| &\le \max\left(\|X_{t_k}x\|_{\mathcal{H}_2}^2 - \|Gx\|_2^2, \|Gx\|_2^2 - \|X_{t_k}x\|_{\mathcal{H}_2}^2\right) \\ &\le \max\left(\|X_{t_k}x\|_{\mathcal{H}_2}^2, \|Gx\|_2^2\right) \le \tau. \end{aligned}$$

最后，我们使用已证明的 $x \in \mathcal{B}$ 的不等式，即

$$\|X_{t_k}x\|_{\mathcal{H}_2}^2 \le \tau; \quad \mathbb{E}\|X_{t_k}x\|_{\mathcal{H}_2}^2 = \|Gx\|_2^2 \le 1.$$

因此，对于 $(x,p) \in \mathcal{B}' \times \mathcal{P}$，我们得到

$$\begin{aligned} \mathbb{E}|f_{x,p}(X_{t_k})|^2 &= \mathbb{E}|\|X_{t_k}x\|_{\mathcal{H}_2}^2 - \|Gx\|_2^2|^2 \\ &= \mathbb{E}\|X_{t_k}x\|_{\mathcal{H}_2}^4 - \left(\mathbb{E}\|X_{t_k}x\|_{\mathcal{H}_2}^2\right)^2 \\ &\le \tau\mathbb{E}\|X_{t_k}x\|_{\mathcal{H}_2}^2 - \|Gx\|_2^4 \le \tau. \end{aligned}$$

□

利用引理 A.1 可得所需的集中界。

命题 A.11. 存在一个普适常数 $C > 0$，使得对于任意δ，$\gamma \in (0,1)$，若

$$m \ge C\delta^{-2}\tau\log(1/\gamma) \quad and \quad \mathbb{E}\delta_* \le \delta/2,$$

则 $\mathbb{P}(\delta_* \ge \delta) \le \gamma$。

证明. 由于根据假设有 $\mathbb{E}\delta_* \le \delta/2$，我们得到

$$\mathbb{P}(\delta_* \ge \delta) \le \mathbb{P}(\delta_* \ge \mathbb{E}\delta_* + \delta/2) = \mathbb{P}(m\delta_* \ge \mathbb{E}(m\delta_*) + \varepsilon),$$

其中我们令 $\varepsilon = m\delta/2$。回顾 (42)，结合引理 A.1（对于 $m\delta_*$）与引理 A.10 中的估计，可得

$$\begin{aligned} \mathbb{P}(m\delta_* \ge \mathbb{E}(m\delta_*) + \varepsilon) &\le \exp\left(-\frac{\delta^2m^2/4}{2(m\tau + m\tau\delta) + m\tau\delta/3}\right) \\ &= \exp\left(-\frac{1}{4}\frac{1}{2(1+\delta)+\delta/3}\frac{\delta^2 m}{\tau}\right) \\ &\le \exp\left(-\frac{1}{C}\frac{\delta^2 m}{\tau}\right), \end{aligned}$$

其中 $C \ge 18$。因此，该结论由关于 m 的假设得出。 □

A.2. 一般情形。我们现在可以在完全一般的情形下完成定理 5.7 的证明。到目前为止，我们已经证明，若 $f_\nu \equiv 1$，也就是说，若 $\mu = \nu$ 是概率测度，那么根据命题 A.9 和命题 A.11，我们得到 $\delta_* \le \delta$，其概率超过 $1-\gamma$。特别地， A 以极高概率满足关于 $(G, \omega, \delta, \lambda)$ 的广义限制等距性质。

现在考虑 $f_\nu \not\equiv 1$ 的情形。我们引入归一化测量˜算子 F_t，定义如下：

$$\tilde{F}_t \coloneqq f_\nu(t)^{-1/2} F_t.$$

假设 3.1 现在表述如下：

$$\|F_t \phi_i\|_{\mathcal{H}_2} \le B\sqrt{f_\nu(t)}\omega_i, \quad t \in \mathcal{D},\ i \in \Gamma,$$

后者等价于

$$\|\tilde{F}_t \phi_i\|_{\mathcal{H}_2} \le B\omega_i, \quad t \in \mathcal{D},\ i \in \Gamma.$$

˜此外，注意到对于每个 $u \in \mathcal{H}_1$， F_t 关于 $\mathrm{d}\nu = f_\nu \mathrm{d}\mu$ 满足以下恒等式：

$$\int_{\mathcal{D}} \|\tilde{F}_t u\|^2_{\mathcal{H}_2} \mathrm{d}\nu(t) = \int_{\mathcal{D}} f_\nu(t)^{-1} \|F_t u\|^2_{\mathcal{H}_2} f_\nu(t) \mathrm{d}\mu(t) = \|Fu\|^2_{L^2_\mu(\mathcal{D};\mathcal{H}_2)}.$$

因此，证明再次归结为测量算子˜F_t 及其伴随前向映射 F 满足 $f_\nu \equiv 1$ 的情形。特别地，以超过 $1-\gamma$ 的概率，˜我们得到 A 关于 $(G, \omega, \delta, \lambda)$ 满足广义限制等距性质，其中

$$\tilde{A} \coloneqq \left(\frac{1}{\sqrt{m}} \tilde{F}_{t_k} \Phi^* \iota_\Lambda\right)_{k=1}^m = \left(\frac{1}{\sqrt{m}} f_\nu(t_k)^{-1/2} F_{t_k} \Phi^* \iota_\Lambda\right)_{k=1}^m = QA.$$

MaLGa Center, Department of Mathematics, University of Genoa, Via Dodecaneso 35, 16146 Genova, Italy.
电子邮箱地址：giovanni.alberti@unige.it

MaLGa Center, Department of Mathematics, University of Genoa, Via Dodecaneso 35, 16146 Genova, Italy.
电子邮箱地址：alessandro.felisi@edu.unige.it

MaLGa Center, Department of Mathematics, University of Genoa, Via Dodecaneso 35, 16146 Genova, Italy.
电子邮箱地址：matteo.santacesaria@unige.it

Department of Mathematical Sciences “G. L. Lagrange”, Politecnico di Torino, Corso Duca degli Abruzzi, 24, 10129 Torino, Italy电子邮箱地址：salvatore.tra
p asso @ p olito.it