通过最大化 ℓ^4 范数学习毫米波（mmWave）通信的稀疏化变换

Sueda Taner 和 Christoph Studer

arXiv:2601.04980v1 [cs.IT] 8 Jan 2026

摘要

毫米波（mmWave）载波频率下波传播的高度方向性，使用户设备与基站（BS）之间仅存在少量显著的传播路径。波传播的这种稀疏特性体现在波束空间域中；传统上，波束空间域是通过对基站（BS）均匀线性天线阵列进行空间离散傅里叶变换（DFT）得到的，其中每个 DFT 输出对应一个不同的波束。近年来，波束空间处理已成为一种颇具前景的技术，可降低工作于毫米波频率的大规模多用户（MU）多输入多输出（MIMO）全数字系统的基带复杂度和功耗。然而，对于有限维天线阵列而言，DFT 是否为最优稀疏化变换仍不明确。本文将 Zhai 等人通过最大化 ℓ^4 范数进行完备字典学习的框架扩展至复数情形，以学习新的稀疏化变换。我们为 ℓ^4 范数最大化提供理论基础，并提出两种适用的学习算法。随后，我们利用这些算法：（i）从理论上并通过仿真评估 DFT 在信道向量稀疏化方面的最优性；（ii）为真实世界和合成生成的信道向量学习性能更优的稀疏化变换。

索引词

波束空间、字典学习、毫米波通信、多天线通信、稀疏度。

I. 引言

MASSIVE 多用户（MU）多输入多输出（MIMO）和毫米波（mmWave）通信是第五代（5G）及未来无线系统的关键技术 [2], [3]。为抵消毫米波载波频率下较高的路径损耗，采用多天线基站（BS）架构进行波束成形至关重要。全数字基站架构充分发挥天线阵列的潜力：它们能够实现灵活的波束成形、简化信道估计并支持空间复用，同时产生与全模拟或混合架构相当的系统成本和射频（RF）功耗 [4]–[6]。然而，要在实践中实现全数字基站架构，需要用于信道估计、数据检测和多用户预编码的高效硬件与功耗的基带算法。

降低基带复杂度并简化硬件的一种有前景的方法是利用波束空间稀疏性。由于毫米波传播具有高度方向性，且因此产生较强的路径损耗，用户设备（UE）与基站天线阵列之间通常只有少数占主导地位的传输路径 [3], [7]。这一特性使毫米波信道向量能够采用稀疏表示。为获得这种稀疏表示，通常通过空间离散傅里叶变换（DFT）将信道向量从天线域变换到波束空间域，该变换作用于天线阵列（例如均匀线性阵列）。所得波束空间中的每个索引都对应一个特定的波束方向；若该方向存在传播路径，则该索引处的绝对值较大。

如 [8]–[22], 所示，利用波束空间稀疏性，能够为信道估计、数据检测和多用户预编码设计低复杂度基带算法与硬件架构。尽管对于极大规模天线阵列，使用离散傅里叶变换来实现信道向量稀疏化在理论上有其依据 [23], [24], ，但对于实际的有限维阵列，以及建模所得和真实世界（例如测量得到）的毫米波信道向量，离散傅里叶变换是否是最优的稀疏化变换仍不明确。

A. 贡献

本文在更广泛的复值向量酉稀疏化变换学习框架下探讨这一问题。我们的主要贡献如下：

- 我们将 ℓ^4范数最大化框架 [25] 从实值正交群扩展到复值酉群。
- 我们基于复值随机数据模型，构建了一个以 ℓ^4范数为基础的优化问题，从而能够学习适用于毫米波信道的酉稀疏化变换。

ST 就职于瑞典基斯塔的爱立信研究院，CS 就职于瑞士苏黎世联邦理工学院；电子邮件：sueda.taner@ericsson.com，studer@ethz.ch。本文部分内容已发表于 [1]。在此，我们通过补充定理和证明，提供 [1] 中省略的理论背景，以此阐释最大化 ℓ^4-范数对于学习完备酉字典的意义；我们还对离散傅里叶变换（DFT）在信道向量稀疏化中的最优性进行了更详细的分析，并针对特定实值正弦信号模型中的信号稀疏化，提出了离散余弦变换（DCT）具有次优性的新结果。

本研究部分获得瑞士国家科学基金会（NSF）资助，资助编号为 218704 和 207314；此外，项目 CHASER（CHIST-ERA-22-WAI-01）通过瑞士国家科学基金会（SNSF）资助编号 218704 获得 CHIST-ERA 资助。

- 我们将 [25] 提出的实值匹配、拉伸与投影（MSP）算法调整用于所提出的复值场景。我们刻画了该算法的不动点，但未确定它们是局部最优点还是鞍点。
- 我们提出了一种基于坐标上升（CA）的替代学习算法，该算法通过一系列吉文斯旋转和复相位移来实现。该算法便于分析其不动点的局部最优性。
- 我们证明，对于某种多径信道模型，离散傅里叶变换矩阵是MSP算法的不动点。此外，在自由空间视距信道模型下，我们证明离散傅里叶变换矩阵是CA算法的不动点，并且是稀疏性诱导目标函数的局部极大值点。
- 通过数值分析，我们证实，对于实值正弦信号模型，离散余弦变换（DCT）矩阵并非CA算法的不动点。这揭示了一个或许令人意外的事实：对于这一类信号，离散余弦变换并非最优稀疏化变换。
- 我们针对合成和真实世界的信道向量开展数值实验，比较使用离散傅里叶变换与学习得到的变换时，利用稀疏性的基带算法的性能。实验结果表明：（i）对于合成信道数据，离散傅里叶变换是出色的酉稀疏化变换，相比学习得到的变换仅有微小提升；（ii）对于来自真实世界大规模多用户多输入多输出系统的实测信道向量，学习得到的变换能显著提高稀疏度，进而大幅改善利用稀疏性的基带算法的误码率性能。

B. 相关现有技术

学习稀疏化变换与字典学习 [26], [27] 密切相关。字典学习旨在寻找一个（可能是过完备的）基或变换，称为字典，使一组输入信号能够在其中获得稀疏表示。在无线通信中，字典学习算法已用于稀疏信道估计 [28]–[37], 、降低信道状态信息反馈 [38], [39], 的开销，以及活跃用户检测 [40], [41]。这些文献通常研究以下正则化优化问题（或其变体），其中为便于求解，将 ℓ^0-伪范数替换为 ℓ^1-范数：

$$\underset{\mathbf{D}\in\mathbb{C}^{M\times N},\,\mathbf{X}\in\mathbb{C}^{N\times T}}{\text{minimize}} \ \|\bar{\mathbf{H}}-\mathbf{D}\mathbf{X}\|_{\mathrm{F}}^2+\lambda\|\mathbf{X}\|_0. \tag{1}$$

这里，矩阵 $\bar{\mathbf{H}}\in\mathbb{C}^{M\times T}$ 包含 T 个信道样本，每一列 $\bar{\mathbf{h}}_t\in\mathbb{C}^M$ 表示从单个发射天线到 M 根接收天线的信道向量；矩阵 $\mathbf{D}\in\mathbb{C}^{M\times N}$ 表示一个（可能是过完备的）字典，其包含 $M\le N$个元素（通常还会对$\mathbf{D}$ 的列施加单位长度约束）；矩阵 $\mathbf{X}=[\mathbf{x}_1\cdots\mathbf{x}_T]\in\mathbb{C}^{N\times T}$包含相应的稀疏系数向量 $\mathbf{x}_t$, $t=1,\dots,T$；而 $\lambda>0$ 是一个正则化参数，用于权衡稀疏表示误差与稀疏度。该字典学习问题的目标是用$\mathbf{D}\mathbf{X}$ 近似$\mathbf{H}$，同时尽可能减少$\mathbf{X}$中的非零元素数量。

针对式（1）中问题的另一种方法，是将其重新表述为一个约束优化问题，通常形式为

$$\underset{\mathbf{D}\in\mathbb{C}^{M\times N},\,\mathbf{X}\in\mathbb{C}^{N\times T}}{\text{minimize}} \ \|\bar{\mathbf{H}}-\mathbf{D}\mathbf{X}\|_{\mathrm{F}}^2 \ \text{ subject to } \ \|\mathbf{x}_t\|_0\le\eta,\ t=1,\dots,T. \tag{2}$$

例如，可以通过广泛使用的 K-SVD算法 [27] 或最优方向法（MOD） [42]对该问题进行近似求解。这两种方法都在稀疏编码 步骤和字典更新 步骤之间交替进行：稀疏编码步骤通过正交匹配追踪（OMP） [43] 或基追踪 [44]等算法，在固定字典下寻找新的最佳稀疏表示。字典更新步骤则在保持稀疏模式不变的同时，更新字典和稀疏系数向量。

通过此类方法获得的字典在满足单位列范数之外，是任意矩阵。因此，由这些字典诱导的稀疏化变换通常并非酉变换。consequently，将此类变换应用于基站天线阵列上的接收信号会改变噪声统计特性，这意味着，在变换后的（稀疏）域中执行的信号处理任务，在统计意义上不再等价于原始天线域中的相应任务。与这些结果不同，我们旨在通过建立在一种问题设置之上来学习酉 稀疏化变换；该问题设置采用了如 [25] 针对实值信号提出的、基于平滑 ℓ^4-范数并使用正交字典的稀疏度度量。

文献 [45] 也基于 ℓ^4范数的字典学习方法 [25], , 但其具体应用是利用 ℓ^3范数最大化进行盲数据检测。此外，文献 [46], 使用 ℓ^4范数定义一体化感知与通信的性能指标，以研究正交频分复用（OFDM）调制波形的最优性。与这两项研究不同，我们研究通过 ℓ^4范数最大化来学习信道向量稀疏化变换这一更一般的问题。

C. 符号说明

粗体小写字母和大写字母分别表示列向量和矩阵。对于向量 $\mathbf{a}$，第 k 个元素记作 $a_k=[\mathbf{a}]_k$；对于矩阵 $\mathbf{A}$，第 k 列记作 $\mathbf{a}_k$ ，第 (i,k) 个元素记作 $A_{i,k}=[\mathbf{A}]_{i,k}$。上标 $(\cdot)^*$、$(\cdot)^{\mathrm{T}}$ 和 $(\cdot)^{\mathrm{H}}$ 分别表示矩阵共轭、转置和厄米共轭。全零矩阵记作 $\mathbf{0}_{N\times M}$，单位矩阵记作 $\mathbf{I}_N$，酉DFT矩阵记作 $\mathbf{F}_N$；下标中的维度

在上下文明确时省略。 $N \times N$正交矩阵和酉矩阵的集合分别记作 $O(N; \mathbb{R})$ 和 $U(N; \mathbb{C})$。复值置换矩阵定义为一种酉矩阵，其中每一行和每一列都恰有一个非零元素；维度为 $N \times N$ 的复值置换矩阵集合记作 $\mathrm{CP}(N)$。逐元素乘法、克罗内克积、绝对值和 r 次幂分别记作 $\circ$、 $\otimes$、 $|\cdot|$和 $(\cdot)^{\circ r}$。按照 [25], , 对于 $p \geq 1$ ，矩阵 A 的逐元素 ℓ^p-范数记作 $\|\mathrm{A}\|_p$，因此 $\|\mathbb{A}\|_p^p \triangleq \sum_{i,\,k} |A_{i,\,k}|^p$。矩阵 A 的弗罗贝尼乌斯范数记作 $\|\mathbb{A}\|_{\mathrm{F}} \triangleq \sqrt{\sum_{i,\,k} |A_{i,\,k}|^2}$。单位球面上 N 维向量（即具有单位 ℓ^2-范数的向量）构成的集合记作 $\mathbb{S}^{N-1}$。矩阵 A 的奇异值分解记作 $\mathrm{SVD}(\mathrm{A})$，A 的第 i 大奇异值记作 $\sigma_i(\mathrm{A})$。整数集合 $\{1, \ldots, N\}$ 记作 $[N]$。实部和虚部分别用 $\Re\{\cdot\}$ 和 $\Im\{\cdot\}$ 表示。期望算子记作 $\mathbb{E}[\cdot]$。我们用 $\delta[n]$ 表示 $n \in \mathbb{Z}$ 的克罗内克 delta 函数，使得 $\delta[0] = 1$ ，且当 $n \neq 0$ 时 $\delta[n] = 0$ 。所有复值梯度均遵循 [47] 的定义。

D. 论文结构

本文其余部分安排如下。第二节介绍毫米波大规模多用户多输入多输出上行系统和信道模型，随后给出它们的波束空间表示。第三节阐述 ℓ^4-范数度量的动机，并详细介绍 ℓ^4-范数最大化问题设置。第四节提出两种算法，以近似求解 ℓ^4-范数最大化问题。第五节利用这两种算法考察离散傅里叶变换和离散余弦变换的最优性。第六节给出离散傅里叶变换用于信道向量稀疏化时的性能仿真结果。第七节作出总结。

II. System Model

A. 上行系统模型

我们考虑一个毫米波大规模多用户多输入多输出上行系统，其中数据由 U 个单天线用户设备传输至配备 B 根天线的基站。令 $\mathbf{H} \in \mathbb{C}^{B \times U}$ 表示基站与用户设备之间的上行信道矩阵，并令 $\mathbf{s} \in \mathcal{S}^U$表示取自星座图 $\mathcal{S}$ 的 U 数据符号。对于窄带传输，我们将天线域中的基带接收向量表示为

$$\mathbf{y} = \mathbf{H}\mathbf{s} + \mathbf{n}, \tag{3}$$

其中 $\mathbf{y} \in \mathbb{C}^B$ 为接收向量，$\mathbf{n}$ 表示独立同分布的圆对称高斯噪声 .

B. 信道模型

我们关注毫米波频率下的波传播 [7], , 并假设基站与用户设备（或散射体）之间的距离足够远。在理想硬件的假设下，对于配备均匀线性阵列（ULA）、天线间距为 $\lambda/2$（波长为 λ）的基站，表示特定用户设备与基站天线阵列之间信道的 MIMO 信道矩阵 $\mathbf{H}$ 的各列可建模如下 [48]:

$$\mathbf{h} = \sum_{\ell=1}^{L} c_\ell \mathbf{p}(\omega_\ell), \quad \mathbf{p}(\omega) = \left[1, e^{j\omega}, \ldots, e^{j(B-1)\omega}\right]^{\mathrm{T}}. \tag{4}$$

这里， $L \geq 1$ 表示到达天线阵列的传播路径总数（包括可能存在的视距路径）， $c_\ell \in \mathbb{C}$ 是与第 ℓ 条传播路径相关联的复数信道增益，$\mathbf{p}(\omega)$ 是阵列响应向量，而 ω_ℓ 是角频率，通常由关系式 $\omega_\ell = \pi \sin(\phi_\ell)$ 给出，其中 ϕ_ℓ 是第 ℓ 条传播路径相对于天线阵列的角度。

注 1。 如果基站天线阵列存在硬件非理想性，例如天线单元之间的衰减/相位不均匀、故障天线、天线耦合或射频串扰等，则式（*4*）中的信道模型将不再成立。如果已知硬件损伤且能够用可逆矩阵表示，则可以将失真矩阵的逆矩阵乘以接收向量，从而补偿这些损伤。经过这种补偿后，信道输入输出关系将简化为式（*3*），并再次采用式（*4*）中的信道模型，但噪声统计特性很可能会发生改变；这就要求在后续的基带处理操作中予以特别注意，例如波束成形或数据检测。

C. 波束空间表示

为揭示高载波频率下波传播的稀疏特性，可以通过对基站天线阵列中的接收信号应用离散傅里叶变换，获得波束空间（或角域）输入输出关系 [14], [15]。对于配备均匀线性阵列的基站，在理想硬件或已知且可补偿的硬件损伤假设下（如备注 1 所述），该波束空间表示可通过 $\bar{\mathbf{y}} = \mathbf{F}_B \mathbf{y}$ 获得。波束空间表示将$\mathbf{H}$的各列，即公式（4）中 L 个复正弦信号的叠加，变换到频域；当 L较小时，所得波束空间信道向量将具有稀疏性。然而，在存在未知或无法抑制的损伤时，信道矩阵的各列将不再能用公式（4）建模，从而可能对稀疏性产生不利影响。这一关键事实促使我们研究离散傅里叶变换在最大化接收向量稀疏度方面的最优性，并学习出相较于广泛使用的离散傅里叶变换具有更优稀疏化特性的替代酉变换。

III. 初级视觉皮层范数作为稀疏度度量

文献 [25] 针对一组给定的实值向量，在正交群 $O(N;\mathbb{R})$ 上构建了基于 ℓ^4 范数的字典学习框架。该框架背后的直觉是，在超球面上最大化矩阵的 ℓ^4 范数能够促进稀疏性；也就是说，在 ℓ^2 范数超球面上，最稀疏的点具有最小的 ℓ^1 范数和最大的 ℓ^4 范数 [25]。下文中，我们基于这一洞见，考虑带有随机数据模型的复值情形，旨在将该框架用于波束空间处理。

A. 问题设置

假设我们有遵循某种随机模型的数据样本，并希望找到一种酉变换，使这些数据样本的期望稀疏度最大化。具体而言，对于数据样本$\mathbf{A} \in U(N;\mathbb{C})$作用于$\mathbf{y} \in \mathbb{C}^N$后，我们使用基于 ℓ^4范数的目标函数$\hat{g}\colon U(N;\mathbb{C}) \to \mathbb{R}$来度量其期望稀疏度，该目标函数定义为

$$\hat{g}(\mathbf{A}) \triangleq \mathbb{E}_{\mathbf{y}}\left[\|\mathbf{A}\mathbf{y}\|_4^4\right]. \tag{5}$$

我们将目标表述为 fi　　寻找通过以下优化问题最大化稀疏度的酉变换　　m:

$$\underset{\mathbf{A}}{\text{maximize}}\ \hat{g}(\mathbf{A})\ \text{ subject to }\ \mathbf{A} \in U(N;\mathbb{C}). \tag{O1}$$

我们可以将对数据样本的概率分布取 ℓ^4范数的期望，理解为对无限多个数据样本的ℓ^4范数求平均。同样，对有限个观测样本的 ℓ^4范数求平均，等价于对仅在给定样本处非零的均匀概率质量函数（PMF）取期望。因此，对于观测样本有限的确定性情形，我们将式（5）中的期望替换为求和，并将稀疏度度量定义为 $g(\mathbf{A},\ \mathbf{Y})\colon U(N;\mathbb{C}) \times \mathbb{C}^{N\times M} \to \mathbb{R}$，如下所示：

$$g(\mathbf{A},\mathbf{Y}) \triangleq \sum_{m=1}^{M} \|\mathbf{A}\mathbf{y}_m\|_4^4 = \|\mathbf{A}\mathbf{Y}\|_4^4. \tag{6}$$

因此，相应的优化问题为

$$\underset{\mathbf{A}}{\text{maximize}}\ g(\mathbf{A},\mathbf{Y})\ \text{ subject to }\ \mathbf{A} \in U(N;\mathbb{C}), \tag{O2}$$

这对应于 [25] 的实值字典学习问题在复数域中的复值等价形式。因此，接下来我们将把 [25] 中的结果扩展到复数域。

B. 为什么选择 ℓ^4-范数?

我们首先针对 $U(N;\ \mathbb{C})$ 建立 ℓ^4-范数的关键性质，以更清楚地说明酉矩阵的稀疏度与其 ℓ^4-范数之间的关系；这些性质是对 [25] 中实值 $O(N;\ \mathbb{R})$ 情形的扩展。以下性质背后的核心思想是，最稀疏的酉矩阵是复置换矩阵。具体而言，对于任意$\mathbf{A} \in U(N;\ \mathbb{C})$，有 $N \le \|\mathbf{A}\|_0 \le N^2$，且当且仅当 $\mathbf{A} \in \mathrm{CP}(N)$ 时，下界能够取到。

下面的引理表明，酉群上 ℓ^4-范数的全局最大化点是复置换矩阵，而复置换矩阵是最稀疏的酉矩阵。证明见附录 A-A。

引理 1（酉群上 ℓ^4-范数的极值）。对于任意酉矩阵$\mathbf{A} \in U(N;\ \mathbb{C})$，有 $\|\mathbf{A}\|_4^4 \in [1,N]$。当且仅当$\mathbf{A} \in \mathrm{CP}(N)$时，上界可取到，即 $\|\mathbf{A}\|_4^4 = N$。

注 2。图 *1* 直观地说明了为什么酉球面上的 ℓ^4-范数是合适的稀疏度度量：将酉约束视为受限于单位 ℓ^2-球面。由此可见，在单位 ℓ^2-球面上，ℓ^4-范数在最稀疏的点处取得最大值，这些点以橙色星形标记表示；而典型的稀疏度度量 ℓ^1-范数在这些点处取得最小值。此例说明，最大化单位 ℓ^2-球面上的 ℓ^4-范数可以促进稀疏性。

以下引理表明，如果矩阵的 ℓ^4-范数“接近”其最大值，那么从弗罗贝尼乌斯范数的角度看，该矩阵“接近”一个复置换矩阵。证明见附录 A-B。

引理 2（ ℓ^4范数在酉群上的近似极大值）。 设 $\mathbf{W}$是一个酉矩阵，即 $\mathbf{W} \in U(N;\ \ \mathbb{C})$。若 $\frac{1}{N}\|\mathbf{W}\|_4^4 \ge 1-\epsilon$, $\forall \epsilon \in \left[0,\ \frac{1}{N}\left(1-\frac{1}{\sqrt{2}}\right)\right)$，则存在 $\mathbf{P} \in \mathrm{CP}(N)$，使得

$$\frac{1}{N}\|\mathbf{W}-\mathbf{P}\|_{\mathrm{F}}^2 \le 2\epsilon. \tag{7}$$

现在，我们给出酉群 $U(N;\ \mathbb{C})$ 上 ℓ^4-范数的一些基本性质，作为文献 [25] 中 $O(N;\ \mathbb{R})$ 情形的推广。我们通过令$\mathbf{A} \triangleq g(\mathbf{A},\ \mathbf{I}_N)$，定义 h，其中$\mathbf{Y} = \mathbf{I}_N$取自式（6），并将式（O2）中的优化问题改写如下：

$$\underset{\mathbf{A}}{\text{maximize}}\ h(\mathbf{A})\ \text{ subject to }\ \mathbf{A} \in U(N;\mathbb{C}). \tag{O3}$$

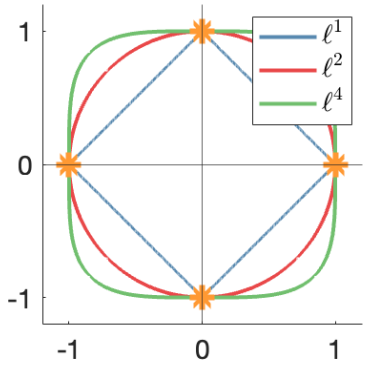

图 1. $\mathbb{R}^2$ 中的单位 ℓ^1-、 ℓ^2- 和 ℓ^4-球面。

我们在下面讨论的性质中使用 $h(\mathbf{A})$。

单位 群是复 Stiefel 流形的一种特殊情形，其切空间由 [49] 给出

$$T_{\mathbf{W}}U(N;\mathbb{C}) \triangleq \{\mathbf{Z} \in \mathbb{C}^{N\times N} : \mathbf{Z}^{\mathrm{H}}\mathbf{W} + \mathbf{W}^{\mathrm{H}}\mathbf{Z} = \mathbf{0}_{N\times N}\}, \tag{8}$$

其中 $\mathbf{W} \in \mathbb{C}^{N\times N}$。从 $U(N;\mathbb{C})$ 到切空间的投影运算 $\mathcal{P}_{T_{\mathbf{W}}U(N;\mathbb{C})}$: $\mathbb{C}^{N\times N} \to T_{\mathbf{W}}U(N;\mathbb{C})$ 定义为 [50]

$$\mathcal{P}_{T_{\mathbf{W}}U(N;\mathbb{C})}(\mathbf{Z}) \triangleq \mathbf{Z} - \frac{1}{2}\mathbf{W}(\mathbf{W}^{\mathrm{H}}\mathbf{Z} + \mathbf{Z}^{\mathrm{H}}\mathbf{W}) = \frac{1}{2}(\mathbf{Z} - \mathbf{W}\mathbf{Z}^{\mathrm{H}}\mathbf{W}). \tag{9}$$

我们用 $\nabla_{\mathbf{W}}h(\mathbf{W})$ 表示 $h(\mathbf{W})$ 关于$\mathbf{W}$ 的梯度，用 grad $h(\mathbf{W})$ 表示 $h(\mathbf{W})$ 在 $T_{\mathbf{W}}U(N;\mathbb{C})$ 上关于$\mathbf{W}$的黎曼梯度。因此，我们可以将 $\mathbf{W}$ 在 $T_{\mathbf{W}}U(N;\mathbb{C})$ 上的黎曼梯度表述如下:

$$\text{grad } h(\mathbf{W}) = \mathcal{P}_{T_{\mathbf{W}}U(N;\mathbb{C})}(\nabla_{\mathbf{W}}h(\mathbf{W})) = \frac{1}{2}(\nabla_{\mathbf{W}}h(\mathbf{W}) - \mathbf{W}(\nabla_{\mathbf{W}}h(\mathbf{W}))^{\mathrm{H}}\mathbf{W}). \tag{10}$$

We 利用式 (8)–(10) 分析下述两个引理中 $h(\mathbf{W})$ 的临界点 s.

引理 3。 $h(\mathbf{W})$ 在流形 $U(N;\mathbb{C})$ 上的临界点满足以下条件:

$$\mathbf{W}^{\mathrm{H}}(|\mathbf{W}|^{\circ 2} \circ \mathbf{W}) = (|\mathbf{W}|^{\circ 2} \circ \mathbf{W})^{\mathrm{H}}\mathbf{W}. \tag{11}$$

引理 3 的证明见附录 A-C。引理 3 表明，对于所有 $\mathbf{W} \in U(N;\mathbb{C})$， $U(N;\mathbb{C})$ 上 ℓ^4-范数临界点的条件由以下方程组给出:

$$\begin{cases} \mathbf{W}^{\mathrm{H}}(|\mathbf{W}|^{\circ 2} \circ \mathbf{W}) = (|\mathbf{W}|^{\circ 2} \circ \mathbf{W})^{\mathrm{H}}\mathbf{W} \\ \mathbf{W}^{\mathrm{H}}\mathbf{W} = \mathbf{I}. \end{cases} \tag{12}$$

我们利用引理 3 在引理 4 和第 IV-B 节中确定 $h(\mathbf{W})$的临界点。

引理 4。 酉群上 ℓ^4-范数的所有全局最大化点都是孤立临界点。

引理 4 的证明见附录 A-D。引理 4 表明，酉群上的 ℓ^4-范数目标函数的全局最大化点不会构成连续的解族；这意味着最大化点的微小扰动不会仍然是最大化点，这对于优化算法的理论保证和实际收敛都很有用。

IV. 学习稀疏化变换

现在，我们提出两种算法来近似求解 (O1)；只需像我们在 (6) 中所做的那样，将期望替换为求和，即可轻松扩展到近似求解 (O2)。介绍这些算法时，我们将保持说明的通用性，并在第 V 节和第 VI 节讨论信道向量稀疏化的具体应用。

A. MSP: 匹配、拉伸与投影

我们提出一种 MSP 算法的变体，它（i）将 [25, Alg. 2] 中的实值方法扩展到复值模型； （ii）支持使用随机模型，而非有限的观测样本集合。本质上，MSP算法在目标函数 (O1) 上执行步长为无穷大的投影梯度上升 (PGA)。在 MSP 中，每次迭代 t都会将估计值 与$\mathbf{A}_t$ 观测值 $\mathbf{y}$相匹配，将$\mathbf{A}_t\mathbf{y}$ 的所有元素代入梯度中的三次函数 $|\mathbf{A}_t\mathbf{y}|^{\circ 2} \circ (\mathbf{A}_t\mathbf{y})$,

并将其投影回酉群，如 [25] 中所述。我们使用以下引理将其投影到酉群上。

引理 5（投影到酉群）。 对于所有$\mathbf{A} \in \mathbb{C}^{N\times N}$，与$\mathbf{A}$的弗罗贝尼乌斯范数距离最小的酉矩阵如下：

$$\mathcal{P}_{U(N;\mathbb{C})}(\mathbf{A}) \triangleq \underset{\mathbf{M}\in U(N;\mathbb{C})}{\arg\min} \|\mathbf{M}-\mathbf{A}\|_{\mathrm{F}}^2 = \mathbf{U}\mathbf{V}^{\mathrm{H}}. \tag{13}$$

引理5的证明见附录 B-A。利用上述投影，求解优化问题 (O1) 的 MSP流程总结于算法 1.1 中。

算法 1.1（MSP）。 初始化$\mathbf{A}_0 \in U(N;\ \mathbb{C})$。对于每次迭代 $t=0,\ 1,\ \ldots,$ 直至收敛，计算目标函数关于$\mathbf{A}_t$的梯度：

$$\nabla_{\mathbf{A}_t}\hat{g}(\mathbf{A}_t) \triangleq \mathbb{E}_{\mathbf{y}}\left[2\big(|\mathbf{A}_t\mathbf{y}|^{\circ 2}\circ(\mathbf{A}_t\mathbf{y})\big)\mathbf{y}^{\mathrm{H}}\right], \tag{14}$$

并将梯度投影到酉群上：

$$\mathbf{A}_{t+1} = \mathcal{P}_{U(N;\mathbb{C})}(\nabla_{\mathbf{A}_t}\hat{g}(\mathbf{A}_t)). \tag{15}$$

*B. MSP*算法分析

现在，我们通过扩展 [25] 针对正交群的分析，在引理 6–8 中给出简化版 MSP算法在酉群上的收敛分析。此外，我们通过引理 9 和 10 给出算法 1.1 的不动点的一个充分条件。

我们的收敛分析基于（O3）中的优化问题，其目标函数为 $h(\mathbf{A}) = \|\mathbf{A}\|_4^4 = g(\mathbf{A}、\mathbf{I}_N)$。因此，我们采用算法 1.1 的一个略作修改的版本，分别用 $h(\mathbf{A}_t)$ 和 $\nabla_{\mathbf{A}_t}\hat{}g(\mathbf{A}_t) \triangleq 2\big(|\mathbf{A}_t|^{\circ 2}\circ\mathbf{A}_t\big)$ 替换 $\hat{}g(\mathbf{A}_t)$ 和 $\nabla_{\mathbf{A}_t}h(\mathbf{A}_t)$。为完整起见，我们在算法 1.2 中给出该算法的修改版本。

算法 1.2。 初始化 $\mathbf{A}_0 \in U(N;\ \mathbb{C})$。对于每次迭代 $t=0,\ 1、\ldots\ldots,$ 直至收敛，计算目标函数关于$\mathbf{A}_t$的梯度：

$$\nabla_{\mathbf{A}_t}h(\mathbf{A}_t) \triangleq 2\big(|\mathbf{A}_t|^{\circ 2}\circ\mathbf{A}_t\big), \tag{16}$$

并将梯度投影到酉群上：

$$\mathbf{A}_{t+1} = \mathcal{P}_{U(N;\mathbb{C})}(\nabla_{\mathbf{A}_t}h(\mathbf{A}_t)). \tag{17}$$

注意， $h(\mathbf{A})$ 关于$\mathbf{A}$上的酉旋转和复相位移具有不变性；因此，不失一般性，我们希望证明，对于（O3）中的优化问题，对于在 $U(N;\ \mathbb{C})$ 上随机初始化的任意 $\mathbf{A}_0$，算法 1.2 以概率 1 收敛到 CP(N) 中的某个复置换矩阵。

下面的引理阐明了初级视觉皮层目标函数 (O3) 的临界点的性质；引理 7 表明，采用任意固定步长 α（甚至 $\alpha=+\infty$）的 PGA 都能找到 (O3) 的临界点。接着，引理 8 表明，算法 1.2 在局部具有三次收敛速率，这比梯度下降方法快得多。

引理 6 (算法 1.2 的不动点)。 给定 $\mathbf{W} \in U(N;\ \mathbb{C})$，$\mathbf{W}$是算法 *1.2* 的不动点，当且仅当 $\mathbf{W}$是 $U(N;\ \mathbb{C})$ 上 ℓ^4-范数的临界点。

引理 6 的证明见附录 B-B。引理 6 意味着，算法 1.2 只会收敛到 $U(N;\ \mathbb{C})$ 上初级视觉皮层范数的临界点。

引理 7 (PGA 在任意步长下的收敛性)。 采用任意固定步长 $\alpha>0$ 的迭代投影梯度上升（*PGA*）算法（α 可为 $+\infty$，此时 *PGA* 等价于算法 *1.2*）能够找到该优化问题的一个鞍点：最大值$\underset{\mathbf{A}\in U(N;\mathbb{C})}{}\|\mathbf{A}\|_4^4$。

引理 7 的证明见附录 B-C。引理 7 表明，采用任意步长 $\alpha>0$（包括$\alpha=+\infty$）的 PGA 都能在 $\|\mathbf{A}\|_4^4$上找到 $U(N;\ \mathbb{C})$ 的临界点。

引理 *8*（全局极大值点附近的三次收敛率）。给定一个酉矩阵 $A \in U(N;\ \mathbb{C})$，令 A' 表示算法 *1.2* 经过一次迭代后的输出：$A' = UV^{\mathrm{H}}$，其中 $U\Sigma V^{\mathrm{H}} =$ 是 $A\big(A^{\circ 2}\circ A\big)$ 的奇异值分解。若 $\|A-C\|^2F = \epsilon$，对于 $\epsilon < 0.394$，则有 $\|A'-C'\|^2F < \|A-C\|^2F$，且 $\|A'-C'\|^2F < O(\epsilon^3)$，其中 C、$C' \in CP(N)$ 是对角复置换矩阵，使得 $c_{i,i} = e^{j\angle A_{i,i}}$ 且 $c'_{i,i} = e^{j\angle A'_{i,i}}$。

引理 8 的证明见附录 B-D。引理 8 表明，如果初始酉矩阵 $\mathbf{A}$ 足够接近一个复置换矩阵，那么算法 1.2 将以三次速率收敛到一个复置换矩阵。由于该算法具有排列不变性，我们不失一般性地假设，目标复排列是一个对角复置换矩阵 $\mathbf{C} \in \mathbb{C}^{N\times N}$。

到目前为止，我们已经对算法 1.2进行了收敛分析。接下来，我们给出一个引理，用于确定算法 1.1的不动点。

引理 *9.* 设 $A \in U(N; \mathbb{C})$ 是一个酉矩阵，其奇异值分解为 $SVD(A) = U\Sigma V^{\mathrm{H}}$，且 $D \in \mathbb{R}^{N\times N}$ 是一个对角矩阵。假设矩阵 $A_1 = DA$ 和 $A_2 = AD$ 的奇异值分解分别为 $SVD(A_1) = U_1\Sigma_1 V_1^{\mathrm{H}}$ 和 $SVD(A_2) = U_2\Sigma_2 V_2^{\mathrm{H}}$。那么，存在 A_1 和 A_2 的奇异值分解，使得 $\Sigma_1 = \Sigma_2 = D$，且 $A = UV^{\mathrm{H}} = U_1V_1^{\mathrm{H}} = U_2V_2^{\mathrm{H}}$。

引理 9 的证明见附录 B-E。现在，我们利用引理 9 为算法 1.1 的不动点给出一个充分条件。

引理 10 (MSP 的不动点)。 当存在一个对角矩阵$\mathbf{D} \in \mathbb{R}^{N\times N}$ 使得 $\nabla_{\mathbf{A}}\hat{} g(\mathbf{A}) = \mathbf{DA}$ 或 $\nabla_{\mathbf{A}}\hat{} g(\mathbf{A}) = \mathbf{AD}$时，酉矩阵$\mathbf{A} \in U(N; \mathbb{C})$ 是算法 *1.1* 的不动点。

引理 10 的证明可直接由引理 5 和引理 9 得出。在第 V-A 节中，我们将使用引理 10 来考察离散傅里叶变换在稀疏化信道向量方面的最优性。

注释 3. 要证明 *MSP*算法的不动点具有局部最优性颇具挑战：式 *(O1)* 中的优化问题是在酉群上进行的，而酉群是一个黎曼流形；要证明某一点是局部最大值 *[49]*，需要证明该点处的黎曼梯度为零，且黎曼 *Hessian* 为负半定。虽然针对 *[46]*，中的问题设置可以进行此类分析，但同样的方法并不适用于此处。通过接下来详述的步骤，我们可以更进一步地研究局部最优性。

C. CA：坐标上升

现在，我们提出一种替代方法，以近似求解式 (O1) 和 (O2) 中的优化问题。这里，我们利用这样一个事实：如 [51], [52] 所示，每个酉矩阵都可以表示为一系列实值 Givens旋转矩阵和复相位移的乘积。令 $\mathrm{G}(i,k,\alpha_{i,k}) \in \mathbb{C}^{N\times N}$表示具有如下形式的实值 Givens旋转矩阵：$G_{i,i} = G_{k,k} = \cos(\alpha_{i,k})$ $G_{i,k} = -G_{k,i} = \sin(\alpha_{i,k})$ $G_{\ell,\ell} = 1, \ell \neq i$；其他情况下为 $G_{\ell,m} = 0$。将矩阵从左侧乘以 $\mathrm{G}(i,k,\alpha_{i,k})$，相当于在 (i,k) 坐标平面内逆时针旋转 $\alpha_{i,k}$ 弧度。令 $\mathrm{R}(k,\beta_k) \in \mathbb{C}^{N\times N}$表示具有如下形式的对角复相移矩阵：$R_{k,k} = e^{j\beta_k}$，且当 $\ell \neq k$ 时为 $R_{\ell,\ell} = 1$。我们将式 (O1) 中的优化变量 A 改写为 $\mathrm{A} = \prod_{k=N}^{1}\prod_{i=k+1}^{N}\mathrm{G}(i,k,\alpha_{i,k})\prod_{\ell=k}^{N}\mathrm{R}(\ell,\beta_{i,\ell})$。这一分解得到以下与式 (O1) 等价的无约束优化问题：

$$\underset{\substack{\{\alpha_{i,k}\in\mathbb{R}:i,k\in[N],i>k\},\\ \{\beta_{k,\ell}\in\mathbb{R}:k,\ell\in[N],\ell\geq k\}}}{\text{maximize}} \ \mathbb{E}_{\mathbf{y}}\left[\left\|\left(\prod_{k=N}^{1}\prod_{i=k+1}^{N}\mathbf{G}(i,k,\alpha_{i,k})\prod_{\ell=k}^{N}\mathbf{R}(\ell,\beta_{k,\ell})\right)\mathbf{y}\right\|_4^4\right]. \tag{O4}$$

虽然（O4）中的优化问题与（O1）不同，不受约束，但由于目标函数关于优化变量是非凸的，直接求解仍然很困难。为此，我们提出一种基于坐标上升（CA）的算法（算法 2），通过迭代求得（O4）的近似解；该方法无需投影步骤，因为每次迭代更新都会保持酉性。

算法 2 （CA）。初始化 $\mathbf{A}_0 \in U$（N；$\mathbb{C}$）。对于每次迭代 $t = 01$、.....，直至收敛，选择一个索引对（i，k），使得 i、$k \in [N]$、$i > k$，并求出

$$\alpha_{i,k} \in \arg\max_{\tilde{\alpha}} \ \mathbb{E}_{\mathbf{y}}\left[\|\mathbf{G}(i,k,\tilde{\alpha})\mathbf{A}_t\mathbf{y}\|_4^4\right], \tag{18}$$

$$\{\beta_i,\beta_k\} \in \arg\max_{\tilde{\beta}_i,\tilde{\beta}_k} \ \mathbb{E}_{\mathbf{y}}\left[\left\|\mathbf{G}(i,k,\alpha_{i,k})\mathbf{R}(i,\tilde{\beta}_i)\mathbf{R}(k,\tilde{\beta}_k)\mathbf{A}_t\mathbf{y}\right\|_4^4\right], \tag{19}$$

并应用更新

$$\mathbf{A}_{t+1} = \mathbf{G}(i,k,\alpha_{i,k})\mathbf{R}(i,\beta_i)\mathbf{R}(k,\beta_k)\mathbf{A}_t. \tag{20}$$

该 CA算法的核心思想是：通过将吉文斯旋转矩阵和相移矩阵从右侧乘以给定的初始酉矩阵，依次更新其中一对索引，从而改进 ℓ^4-范数目标函数。请注意，由于 ℓ^4-范数对复相位移保持不变，因此在每次迭代中，我们先优化 $\alpha_{i,\ k}$，然后相应地优化 β_i 和 β_k。由于吉文斯旋转矩阵和相移矩阵与酉矩阵相乘后仍为酉矩阵，

该算法在每次迭代中都保持酉性，完全避免了向酉群进行投影。这一关键性质使我们能够在每次 CA迭代中，通过直接分析一阶导数和二阶导数来研究局部最优性。

注释 4. *CA*算法的复杂度通常高于*MSP*，因为我们必须至少遍历一次所有(i,k)索引对，并针对每个(i,k)，通过另一种迭代算法（例如梯度上升或线搜索）求解式 *(18)* 和式 *(19)* 中的优化问题。*CA*算法的优点在于，它能够建立局部最优性，因为在每次迭代中，我们求解的都是一个优化变量为标量（即 *Givens*旋转角）的无约束优化问题。相比之下，*MSP*算法求解的是一个约束问题，其中优化变量是一个 $N \times N$ 矩阵，因此最优性分析较为困难。下文将详细介绍 *CA*算法每次迭代的最优性准则。

D. 坐标上升分析

显然，算法 2 在每次迭代中的目标函数值都单调不减。由于目标函数是非凸的，我们无法断言其全局收敛；不过，我们可以研究该算法的不动点，以及每次迭代时式 (18) 中目标函数的局部极值。以下引理将分析这些性质。

引理 11。 设矩阵$\mathbf{G}$表示第 *IV-C* 节中定义的 *Givens*旋转矩阵，并设$\mathbf{x} \triangleq \mathbf{A}\mathbf{y}$为酉矩阵$\mathbf{A} \in U(N;\ \mathbb{C})$。矩阵$\mathbf{A}$是算法 *2* 的一个不动点，并且式 *(18)* 中的函数 $\|\mathbf{G}(i,k,\alpha_{i,\ k})\mathbf{A}\mathbf{y}\|_4^4$在 $\alpha_{i,\ k}=0$ 处取得局部最大值，当且仅当对于所有 i、$k \in [N]$、$i>k$，以下两个条件成立：

$$\text{(i)}\quad \mathbb{E}_{\mathbf{y}}\left[\frac{\partial}{\partial\alpha}\left(\|\mathbf{G}(i,k,\alpha)\mathbf{A}\mathbf{y}\|_4^4\right)\right]\Bigg|_{\alpha=0} = 4\,\mathbb{E}_{\mathbf{x}}\left[\Re\{x_i x_k^{\star}\}(|x_i|^2-|x_k|^2)\right]=0, \tag{21}$$

$$\text{(ii)}\quad \mathbb{E}_{\mathbf{y}}\left[\frac{\partial^2}{\partial\alpha^2}\left(\|\mathbf{G}(i,k,\alpha)\mathbf{A}\mathbf{y}\|_4^4\right)\right]\Bigg|_{\alpha=0} = 4\,\mathbb{E}_{\mathbf{x}}\left[2\Re\{x_k^2(x_i^*)^2\}+4|x_i|^2|x_k|^2-|x_k|^4-|x_i|^4\right]<0. \tag{22}$$

引理 11 的证明直接由一阶和二阶导数检验 [53], 式$\mathbf{A} = \mathbf{G}(i、\ k、0)\mathbf{A}$成立这一事实得出，并通过化简 $\alpha=0$ 的导数表达式完成。我们将在第 V-A 节和第 V-B 节中利用该引理，分别考察 DFT 在信道向量稀疏化方面的最优性，以及 DCT 在实值正弦波稀疏化方面的最优性。

V. Optimality of Known Transforms

现在，我们在第 IV 节中讨论算法的具体应用，以考察两种常见变换的最优性：(i) DFT 和 (ii) DCT。

A. 离散傅里叶变换

由于我们的首要目标是对第二节中提到的毫米波系统的信道向量进行稀疏化，因此我们采用随机数据模型$\mathbf{y} \in \mathbb{C}^B$，根据（4）可表示为：

$$y_b = \sum_{\ell=1}^{L} c_\ell e^{j\Omega_\ell(b-1)}. \tag{23}$$

其中，$\Omega_\ell \sim$服从均匀分布Unif$(0,2\pi)$，且对于 $\ell=1,..,L$，它们是独立同分布的。该数据模型对应于一种多径信道模型，其中与每条路径相关的角频率服从均匀分布。[1] 基于大量关于波束空间处理的文献（例如，参见 [8]–[22]），已知离散傅里叶变换是稀疏化信道向量的良好候选变换。因此，我们使用离散傅里叶变换初始化MSP和CA算法，以便（i）检验离散傅里叶变换的最优性，以及（ii）寻找可能更优的稀疏化变换。下面将详细分析离散傅里叶变换对于这两种算法的最优性。

*1) MSP*分析： 我们将证明，对于式（23）中的随机数据模型，离散傅里叶变换矩阵是MSP算法的不动点，具体结果如下。

定理 1. 对于式（*23*）中给出的随机数据模型$\mathbf{y}$，离散傅里叶变换矩阵 $\mathbf{F}_B$ 是算法 *1.1* 的一个不动点。

定理 1 的证明见附录C-A。定理 1 表明，算法 1.1 无法找到任何比离散傅里叶变换矩阵更能稀疏化式（23）中的多径信道模型的酉变换。需要注意的是，这一分析并未确定离散傅里叶变换矩阵对应的是目标函数（O1）的鞍点还是局部最大值。尽管如此，我们仍可使用所提出的坐标上升算法来分析局部最大值。

[1]另一种随机数据模型可以采用 $e^{j\pi\sin\Phi_\ell}$b，$\Phi_\ell \sim$Unif(0，2π)，该模型假设与每条路径相关联的入射角 [48] 服从均匀分布。然而，为便于进行理论分析，我们采用式 (23) 中的模型。

2）采用坐标上升进行分析：对于式 (23) 中的随机数据模型和自由空间视距信道模型（即 $L = 1$），以下定理证明，离散傅里叶变换矩阵是算法 2 的一个不动点，并且坐标上升可达到式 (18) 中目标函数的一个局部最大值。

定理 2。对于式 *(23)* 中给出的随机数据模型$\mathbf{y}$ 以及 $L = 1$，离散傅里叶变换矩阵 $\mathbf{F}_B$ 是算法 *2* 的不动点，并且式 *(18)* 中的目标函数 $\|\mathbf{G}(i,k,\alpha_{i,k})\mathbf{A}\mathbf{y}\|_4^4$ 对于所有 $i,k \in [B]$, $i > k$均取得局部最大值。

定理 2 的证明见附录 C-B。尽管我们无法形式化地证明函数 $\|\mathbf{G}(i,k,\alpha_{i,k})\mathbf{F}\mathbf{y}\|_4^4$（式 (18)）是否在 $\alpha_{i,k} = 0$ 处取得全局最大值，但对于式 (23) 中具有 $L = 1$ 的随机数据模型$\mathbf{y}$，在进行扰动和随机初始化后，我们没有观察到任何解能够在式 (5) 的目标函数中取得更高的值。我们将全局最优性的证明留待未来工作。

B. 离散余弦变换

上述关于离散傅里叶变换稀疏化复正弦信号的最优性结果可能看似显而易见。类似地，人们会预期，离散余弦变换作为离散傅里叶变换的实值对应形式，对于以下随机数据模型也同样是最优的$\mathbf{y} \in \mathbb{R}^B$，该模型中的信号为实值正弦信号，其形式为

$$y_b = \cos\left(\Omega(b-1) + \Phi\right),\ b = 1, \ldots, B, \tag{24}$$

其中 $\Omega \sim \text{Unif}(0, 2\pi)$，且 $\Phi \sim \text{Unif}(0, 2\pi)$。然而，正如下文将详细说明的那样，对于 B 的各种情况，DCT 的表现出人意料地次优——这表明，对于任意向量长度B，基于复值正弦波的数据模型中 DFT 的最优性并非显而易见。如果 DCT 矩阵既不是 MSP 算法的不动点，也不是坐标上升算法的不动点，那么该算法就能找到一个使目标函数增大的稀疏化变换，这意味着从 ℓ^4-范数的角度看，DCT 是次优的。

1) 坐标上升分析：本节中，我们用 $\mathbf{C}_B$ 表示 DCT-II 矩阵 $B \times B$。[2] 将 (21) 中的一阶导数在 $\mathbf{C}_B$ 代入 $\mathbf{A}$、并采用 (24) 中给出的随机数据模型 $\mathbf{y}$ 后进行化简，所得的闭式表达式见附录 C-C。虽然要证明这一闭式表达式对任意 *any* $B > 2$ 均非零并不容易，但我们的数值对该表达式在 $B = 3, \ldots, 200$ 处进行评估后发现，存在 (i,k)，使得 $\frac{\partial \int_0^{2\pi} \|\mathbf{G}(i,k,\alpha}{\partial \alpha} \omega \|_4^4 \omega \Big|_{\alpha=0} \neq 0$，这意味着 $\mathbf{C}_B$ 不是 CA 的不动点，也不是数据模型 (24) 的最优稀疏化变换。虽然我们尚未对 B 的更大取值进行数值计算，但预计结果相同。

注释 5. 鉴于 *DFT* 对复值正弦波是局部最优的，人们自然会认为 *DCT* 对实值正弦波是最优的，因此这一结果令人意外。不过，我们注意到 *DCT* 已接近最优，因为与 *DCT* 相比，*MSP* 或 *CA* 算法找到的新变换对 *(5)* 中目标函数的改进幅度很小。

VI. 数值结果

现在，我们利用算法 1.1，从给定的有限规模信道向量数据集中学习稀疏化变换，并开展两项实验：(i) 合成信道向量和 (ii) 真实世界（实测）信道向量。需要指出的是，在这些情况下，由于数据集中的信道向量分布不符合式 (23) 的数据模型（该模型假设与每条路径相关的角频率服从均匀分布），因此我们关于离散傅里叶变换最优性的理论结果不再有保证。具体而言，在 (i) 中，由于数据集规模有限，与每条路径相关的角频率并非均匀分布；在 (ii) 中，由于硬件损伤，式 (4) 中的信道模型甚至也不成立。

在所有实验中，我们将信道矩阵划分为训练集和测试集；使用训练集学习稀疏化变换；然后在测试集上测量初级视觉皮层范数，并针对不同的利用稀疏性的算法，模拟未编码比特误码率 (BER) 随信噪比 (SNR) 的变化。具体而言，我们使用 (i) [14] 中的波束空间信道估计 (BEACHES) 算法，以及 (ii) [15] 中的波束空间最大元素 (LE) 数据检测器；其密度系数为 0.125，仅处理稀疏数据中 12.5% 的元素，以降低复杂度。作为基线，我们还纳入了一种天线域线性最小二乘最小均方误差 (LMMSE) 数据检测器，它不依赖稀疏性，但仍使用 BEACHES 对信道向量进行去噪。

A. 合成信道向量

我们首先使用 QuaDRiGa mmMAGIC UMi模型 [55] 在 60吉赫兹载波频率下，采用具有 $\lambda/2$ 天线间距的 ULA，仿真视距（LoS）信道条件。尽管这代表了一种理想化场景，但该数据集中的信道向量并不直接遵循公式（23）中的理论模型，因为数据集包含的信道向量数量有限，并且信道向量中每条路径对应的角频率 Ω_ℓ 的分布不再是 Unif(0，2π)。

[2]尽管我们仅在附录 C-C 中给出了 DCT-II 的简化解析表达式，但我们通过数值验证确认，该结果同样适用于 DCT-I、DCT-III和DCT-IV。

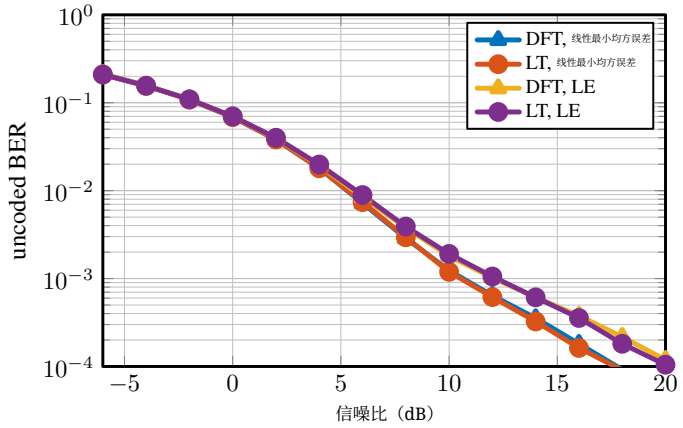

图 2. 使用合成 QuaDRiGa 信道、LMMSE 和利用稀疏性的 LE 检测器时，采用 DFT 或学习变换（LT）进行波束空间处理的未编码误码率，$B = 256$ 根基站天线和 $U = 16$ 个用户设备。

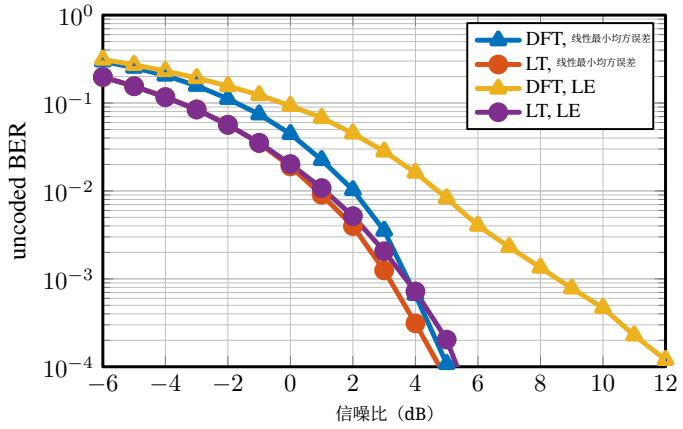

图 3. 使用实测信道向量 [54]、LMMSE 和 LE 检测器时，采用 DFT 或学习变换（LT）进行波束空间处理的未编码误码率，$B = 64$ 根基站天线（其中八根故障）和 $U = 1$ 个用户设备。

我们为一个毫米波大规模多输入多输出系统生成信道矩阵，该系统具有 $B = 256$ 基站天线和 $U = 16$ 单天线用户设备。用户设备随机分布在基站周围的 $120°$ 圆形扇区内，距离基站 10 米至 110米不等，并假设用户设备之间的最小间距为 $5°$。我们加入基站侧功率控制，使接收功率最高的用户设备最多比接收功率最低的用户设备高 6 dB。图 2 展示了使用离散傅里叶变换和学习变换（LT）对该信道向量集进行处理所得的误码率结果。我们观察到，与离散傅里叶变换相比，学习变换的误码率优势仅为边际优势。此外，使用学习变换时，测试集在波束空间域中的 ℓ^4-范数仅比使用离散傅里叶变换时高 18%，这可以解释为稀疏域中峰值幅度仅相差 2%。该结果表明，离散傅里叶变换（i）不再是最优的，但（ii）对于仿真毫米波视距信道而言，仍然是出色的稀疏化变换；这与我们针对公式（23）中所用近似模型的离散傅里叶变换最优性证明一致。因此，鉴于离散傅里叶变换可以通过快速傅里叶变换实现，而学习变换不一定具有快速算法，我们认为，为了获得微小改进而学习新的稀疏化变换并不值得。

我们还尝试了“半均匀阵列”。这一尝试的动机是太赫兹（THz）通信应用中大型天线阵列设计所面临的挑战；参见例如 [56]。在这种情况下，无法在子阵列块之间保持通常的 $\lambda/2$ 阵元间距，因此阵列由均匀子阵列组成，而子阵列之间的间距大于每个子阵列内部的阵元间距。例如，子阵列间距可以为 λ，而每个子阵列内部的阵元间距为 $\lambda/2$。对于这类阵列，我们以与上述类似的方式生成信道向量，并学习了一个稀疏化变换。结果表明，DFT 在这种半均匀阵列上的性能与均匀阵列情形接近。因此，学习变换的误差性能只能略优于 DFT；再一次，学习新的变换似乎并不值得。

B. 真实世界实测信道向量

现在，我们展示 IEEE Communications Theory Workshop Localization Competition [54] 提供的实测信道向量结果。由于采用了具有非理想硬件的矩形阵列，而非理想均匀线性阵列，该数据集中的信道测量结果与公式 (23) 中的理论模型存在显著偏差。这些信道测量基于

单用户设备向基站的传输，使用具有 8×8 方形天线阵列，阵元间距为 $\lambda/2$，载波频率为 1.27 吉赫兹。8 根基站天线发生故障，其输出已从数据集中排除；对于这些缺失项，我们采用了零填充。对于矩形阵列的波束空间处理，通常会按如下方式对这些数据应用二维离散傅里叶变换：将数据向量化为大小为 64 的向量，然后与$\mathbf{F}_8 \otimes \mathbf{F}_8$相乘。相反，我们使用这些长度为 64 的向量作为训练集，以学习一种稀疏化变换。

图3展示了使用离散傅里叶变换和学习变换（LT）处理该信道向量集时的误码率结果。我们观察到，与离散傅里叶变换搭配线性估计和线性最小均方误差检测器相比，LT 分别能以低约 5 dB 和 1 dB 的信噪比达到 0.1% 的目标误码率。此外，应用 LT 后测试集的 ℓ^4-范数最高比二维离散傅里叶变换高 $4\times$ 。这一结果表明，尽管在理想的 LoS 信道条件下，二维离散傅里叶变换非常适合波束空间处理，但学习新的稀疏化变换能够显著改善真实信道以及存在硬件损伤（例如天线故障）的通信系统的性能。

VII. 结论

我们通过将 [25] 中的实值字典学习问题推广到复数情形，构建了一个 ℓ^4-范数最大化问题，用于学习酉稀疏化变换。我们考虑了从随机模型和数据集两种途径进行学习，并针对这一优化问题提出了两种算法：（i）一种基于投影梯度上升的算法，该算法由 [25, 算法 2] 改编而来；（ii）一种无需投影到酉群上的新型坐标上升算法。我们使用这些算法研究了离散傅里叶变换在信道向量稀疏化方面的最优性。理论分析表明，对于多径信道模型，离散傅里叶变换矩阵是第一种算法的不动点。此外，对于自由空间视距信道模型，离散傅里叶变换矩阵是第二种算法的不动点，且每次迭代都能达到目标函数的局部最大值。为揭示这一结果并非显而易见，我们证明了对于实值正弦数据模型，离散余弦变换出人意料地是次优的。为证明离散傅里叶变换在理想毫米波信道模型下确实表现良好，我们给出了仿真结果。结果表明，对于采用非理想硬件的真实场景测量数据，学习得到的变换能够显著降低误码率。

尽管我们关注的是毫米波通信系统，但我们的算法也适用于更一般的稀疏化变换学习问题，可能包括其他随机数据模型或最优性论断，这为未来工作带来了许多研究方向。此外，学习具有比$O(N^2)$更低复杂度的快速算法的变换也将十分有趣，类似于快速傅里叶变换（FFT）。

APPENDIX A
PROOFS OF SECTION III

A. 引理 1 的证明

我们将 $\|\mathbf{A}\|_4^4 = \sum_{k=1}^N \sum_{\ell=1}^N |A_{k,\ell}|^4$ 的上界写为：

$$\|\mathbf{A}\|_4^4 \le \sum_{k=1}^N \sum_{\ell=1}^N |A_{k,\ell}|^4 + 2 \sum_{k=1}^N \sum_{1 \le \ell_1 < \ell_2 \le N} |A_{k,\ell_1}|^2 |A_{k,\ell_2}|^2 = \sum_{k=1}^N \left(\sum_{\ell=1}^N |A_{k,\ell}|^2 \right)^2 = N. \tag{25}$$

当且仅当 $A_{k,\ell_1} A_{k,\ell_2} = 0, \forall k, \ell_1 \neq \ell_2 \in [N]$ 时，上界才能达到，即 $\mathbf{A}$ 的每一行仅包含一个非零元素。结合 $\mathbf{A}$ 为酉矩阵这一条件，可推出 $\mathbf{A} \in \mathrm{CP}(N)$。

我们将 $\|\mathbf{A}\|_4^4$ 的下界写为以下形式，其中 $\tilde{\mathbf{a}} \triangleq \mathrm{vec}(\mathbf{A})$：

$$\|\mathbf{A}\|_4^4 = \|\tilde{\mathbf{a}}\|_4^4 = \sum_{k=1}^{N^2} |\tilde{a}_k|^4 \ge \frac{1}{N^2} \left(\sum_{k=1}^{N^2} |\tilde{a}_k|^2 \right)^2 = 1. \tag{26}$$

根据柯西–施瓦茨不等式 [57], , 当且仅当 $|\tilde{\mathbf{a}}|^{\circ 2}$ 和 $\mathbf{1}_{N^2}$ 线性相关时，才能达到下界。结合$\mathbf{A}$为酉矩阵这一条件，可得$|A_{i,\ k}| = \frac{1}{\sqrt{N}}$， $\forall i$， k。 —=

B. 引理 2 的证明

我们可以将 $\frac{1}{N}\|\mathbf{W}\|_4^4 \ge 1 - \epsilon$ 这一条件改写为

$$\frac{1}{N} \sum_{k=1}^N \|\mathbf{w}_k\|_4^4 \ge 1 - \epsilon, \tag{27}$$

其中每个列向量 $\mathbf{w}_k$ 满足 $\mathbf{w}_k \in \mathbb{S}^{N-1}$。根据引理 13，我们知道存在 $m_1, m_2, \ldots, m_N \in [N]$，使得以下结论对任意 $\epsilon \in [0,1]$均成立：

$$\frac{1}{N} \sum_{k=1}^N \left\| \mathbf{w}_k - \mathbf{e}_{m_k}^{\theta_k} \right\|_2^2 \le 2\epsilon. \tag{28}$$

这里，$\mathbf{e}_{m_k}^{\theta_k} \in \mathbb{C}^N$，其中 $[\mathbf{e}_{m_k}^{\theta_k}]_{m_k} = e^{j\theta_k}, \theta_k \in [0, 2\pi)$，且 $[\mathbf{e}_{m_k}^{\theta_k}]_\ell = 0, \forall \ell \neq m_k, \ell \in [N]$。事实上，根据引理 12，我们知道 (28) 由 $m_k \in \arg\max_{i\in[N]} |W_{i,k}|$ 和 $\theta_k = \angle W_{m_k,k}$具体满足。令 $\mathbf{P} \in \mathbb{C}^{N\times N}$，使得 $\mathbf{p}_k = \mathbf{e}_{m_k}^{\theta_k}$。到目前为止，我们只推导出存在一个 $\mathbf{P}$ ，其列仅有一个幅值为 1 的非零元素。我们尚未分析 $\mathbf{P}$ 是否为复置换矩阵；这要求 $m_k \neq m_j$, $\forall k \neq j$，即序列 $m_1, \ldots, m_N$中不能有重复项；这正是我们接下来要证明的。

通过 将总和 $\sum_{k=1}^N \|\mathbf{w}_k - \mathbf{e}_{m_k}^{\theta_k}\|_2^2 = \sum_{k=1}^N (2 - 2|W_{m_k,\ k}|)$ 简化后，我们将 (28) 重写如下：

$$\sum_{k=1}^{N} (1 - |W_{m_k,k}|) \leq N\epsilon. \tag{29}$$

令 $\mathcal{K}_i \triangleq \{k \in [N] : m_k = i\}$ 为 $\mathbf{P}$ 的列集合，其中非零项位于第 i 行，且 $c_i \triangleq |\mathcal{K}_i|$。令 $\mathcal{K}_i' \triangleq \{k \in \mathcal{K}_i : |W_{i,\ k}| \leq \frac{1}{\sqrt{2}}\}$。现在注意，若 $c_i \geq 2$，则必有 $|\mathcal{K}_i'| \geq c_i - 1$，因为根据 $\mathbf{W}$ 的酉性，$\sum_{k\in\mathcal{K}_i} |W_{i,\ k}|^2 \leq 1$。令 $N_r \triangleq \sum_{i\in\{m_k\}_{k=1}^N} (c_i - 1)$ 表示 m_1，...，m_N 中重复项的总数，其中 $0 \leq N_r \leq N - 1$。于是我们有

$$N_r\left(1 - \frac{1}{\sqrt{2}}\right) \leq \sum_{k\in\mathcal{K}'_{m_k}} (1 - |W_{m_k,k}|) \leq \sum_{k=1}^{N} (1 - |W_{m_k,k}|) \leq N\epsilon < \left(1 - \frac{1}{\sqrt{2}}\right), \tag{30}$$

由（30）可得 $N_r < 1$；由于 N_r 是非负整数，这意味着 $N_r = 0$，从而完成证明 f.

C. 引理 *3* 的证明

由于 $h(\mathbf{W}) = \|\mathbf{W}\|_4^4$，我们有 $\nabla_{\mathbf{W}} h(\mathbf{W}) = 2|\mathbf{W}|^{\circ 2} \circ \mathbf{W}$。将该梯度代入（10）可得：

$$\text{grad } h(\mathbf{W}) = \frac{1}{2}(2|\mathbf{W}|^{\circ 2} \circ \mathbf{W} - \mathbf{W}(2|\mathbf{W}|^{\circ 2} \circ \mathbf{W})^{\mathrm{H}}\mathbf{W}) = 0 \tag{31}$$

$$\implies |\mathbf{W}|^{\circ 2} \circ \mathbf{W} = \mathbf{W}(|\mathbf{W}|^{\circ 2} \circ \mathbf{W})^{\mathrm{H}}\mathbf{W}. \tag{32}$$

将左侧乘以 $\mathbf{W}^{\mathrm{H}}$ 后得到 $\mathbf{W}^{\mathrm{H}}(|\mathbf{W}|^{\circ 2} \circ \mathbf{W}) = (|\mathbf{W}|^{\circ 2} \circ \mathbf{W})^{\mathrm{H}}\mathbf{W}$。

D. 引理 *4* 的证明

不失一般性，我们证明对角复置换矩阵是一个孤立临界点，因为我们的目标函数在复置换群作用下保持不变。令 W_0 为一个对角矩阵，使得 $[\mathrm{W}_0]_{k,k} = e^{j\theta_k}$ for $k \in [N]$and $\theta_k \in [0, 2\pi)$，并且临界点条件，即 $\mathrm{W}_0^{\mathrm{H}}(|\mathrm{W}_0|^{\circ 2} \circ \mathrm{W}_0) = (|\mathrm{W}_0|^{\circ 2} \circ \mathrm{W}_0)^{\mathrm{H}}\mathrm{W}_0$ and $\mathrm{W}_0^{\mathrm{H}}\mathrm{W}_0 = \mathrm{I}_N$，成立。现在我们来考察，在 W_0, 的每个邻域中，是否都存在某个临界点 W；也就是说，是否存在一条路径 $\mathrm{W}(\cdot) : (-\epsilon, \epsilon) \to U(N; \mathbb{C})$，位于 W_0 附近，满足 $\mathrm{W}(0) = \mathrm{W}_0$，且满足以下条件：

$$\mathbf{W}(t)^{\mathrm{H}}(|\mathbf{W}(t)|^{\circ 2} \circ \mathbf{W}(t)) = (|\mathbf{W}(t)|^{\circ 2} \circ \mathbf{W}(t))^{\mathrm{H}}\mathbf{W}(t). \tag{33}$$

在 $t = 0$附近展开$\mathbf{W}(t)$，可得

$$\mathbf{W}(t) = \mathbf{W}_0 + t\mathbf{W}_1 + t^2\mathbf{W}_2 + \cdots. \tag{34}$$

将$\mathbf{W}(t)$代入式 (33) 的约束条件，可得

$$(\mathbf{W}_0 + t\mathbf{W}_1)^{\mathrm{H}} \left((\mathbf{W}_0 + t\mathbf{W}_1)^{\circ 2} \circ (\mathbf{W}_0 + t\mathbf{W}_1)^*\right) = \left((\mathbf{W}_0 + t\mathbf{W}_1)^{\circ 2} \circ (\mathbf{W}_0 + t\mathbf{W}_1)^*\right)^{\mathrm{H}} (\mathbf{W}_0 + t\mathbf{W}_1). \tag{35}$$

化简（35）的右侧可得

$$(\mathbf{W}_0 + t\mathbf{W}_1)^{\mathrm{H}} \left((\mathbf{W}_0 + t\mathbf{W}_1)^{\circ 2} \circ (\mathbf{W}_0 + t\mathbf{W}_1)^*\right) = \mathbf{I} + 2t\mathbf{W}_0^{\mathrm{H}} \circ \mathbf{W}_1 + t\mathbf{W}_0 \circ \mathbf{W}_1^* + t\mathbf{W}_1^{\mathrm{H}}\mathbf{W} \tag{36}$$

类似地，化简（35）的左侧可得

$$\left((\mathbf{W}_0 + t\mathbf{W}_1)^{\circ 2} \circ (\mathbf{W}_0 + t\mathbf{W}_1)^*\right)^{\mathrm{H}} (\mathbf{W}_0 + t\mathbf{W}_1) = \mathbf{I} + t\mathbf{W}_0^{\mathrm{H}}\mathbf{W}_1 + 2t\mathbf{W}_1^{\mathrm{H}} \circ \mathbf{W}_0 + t\mathbf{W}_0^{\mathrm{H}} \circ \mathbf{W}_1. \tag{37}$$

因此，将（35）化简为：

$$2t\mathbf{W}_0^{\mathrm{H}} \circ \mathbf{W}_1 + t\mathbf{W}_0 \circ \mathbf{W}_1^* + t\mathbf{W}_1^{\mathrm{H}}\mathbf{W}_0 = t\mathbf{W}_0^{\mathrm{H}}\mathbf{W}_1 + 2t\mathbf{W}_1^{\mathrm{H}} \circ \mathbf{W}_0 + t\mathbf{W}_0^{\mathrm{H}} \circ \mathbf{W}_1 \tag{38}$$

$$\implies 2\Im\{\mathbf{W}_0^{\mathrm{H}} \circ \mathbf{W}_1\} = \mathbf{W}_0^{\mathrm{H}}\mathbf{W}_1 - \mathbf{W}_1^{\mathrm{H}}\mathbf{W}_0. \tag{39}$$

为简化式 (39) 的左侧，我们使用以下结果：

$$\mathbf{W}(t)^{\mathrm{H}}\mathbf{W}(t) = (\mathbf{W}_0 + t\mathbf{W}_1)^{\mathrm{H}}(\mathbf{W}_0 + t\mathbf{W}_1) = \mathbf{I} + t\mathbf{W}_0^{\mathrm{H}}\mathbf{W}_1 + t\mathbf{W}_1^{\mathrm{H}}\mathbf{W}_0 \tag{40}$$

$$\implies \frac{d}{dt}(\mathbf{W}(t)^{\mathrm{H}}\mathbf{W}(t)) = \mathbf{W}_0^{\mathrm{H}}\mathbf{W}_1 + \mathbf{W}_1^{\mathrm{H}}\mathbf{W}_0. \tag{41}$$

$$\implies \frac{d}{dt}(\mathbf{W}(t)^{\mathrm{H}}\mathbf{W}(t))|_{t=0} = \frac{d}{dt}(\mathbf{I}_N)|_{t=0} = \mathbf{0}_{N\times N} \implies \mathbf{W}_0^{\mathrm{H}}\mathbf{W}_1 + \mathbf{W}_1^{\mathrm{H}}\mathbf{W}_0 = \mathbf{0}_{N\times N}. \tag{42}$$

因此，将 $\mathbf{W}_1^{\mathrm{H}}\mathbf{W}_0 = -\mathbf{W}_0^{\mathrm{H}}\mathbf{W}_1$ 代入式 (39)，可得

$$\Im\{\mathbf{W}_0^{\mathrm{H}} \circ \mathbf{W}_1\} = \mathbf{W}_0^{\mathrm{H}}\mathbf{W}_1. \tag{43}$$

由 (43) 可推得$\mathbf{W}_1 = \mathbf{0}_{N\times N}$ 必然成立。由于$\mathbf{W}(t)$ 是任意路径，这意味着由所有临界点构成的簇在$\mathbf{W}_0$附近不存在切空间。由此可知，$\mathbf{W}_0$ 是一个孤立临界点。

E. 引理 *13* 的辅助引理

引理 12. 设$\mathbf{q}$是一个 N 量，且位于单位球面上：$\mathbf{q} \in \mathbb{S}^{N-1}$。对于 $\epsilon \in [0,1]$，若 $\|\mathbf{q}\|_4^4 \geq 1-\epsilon$，则存在 $\exists k \in [N]$，使得

$$\left\|\mathbf{q} - \mathbf{e}_k^{\theta}\right\|_2^2 \leq 2\epsilon, \tag{44}$$

其中 $\theta = \angle q_k$，$\mathbf{e}_k^{\theta} \in \mathbb{C}^N$ 满足 $[\mathbf{e}_k^{\theta}]_k = e^{j\theta}$，且 $[\mathbf{e}_k^{\theta}]_\ell = 0, \forall \ell \neq k$。

令$\mathbf{q} = [q_1, q_1, \ldots, q_N]^{\mathrm{T}}$，不失一般性，假设

$$1 \geq |q_1|^2 \geq |q_2|^2 \cdots \geq |q_N|^2 \geq 0. \tag{45}$$

证明如下：我们将假设 $\|\mathbf{q}\|_4^4 \geq 1-\epsilon$ 重写为

$$|q_1|^4 + |q_2|^4 + \cdots + |q_N|^4 \geq 1-\epsilon, \tag{46}$$

结合 (45) 和 $\mathbf{q} \in \mathbb{S}^{N-1}$，我们得到

$$|q_1|^4 + |q_2|^2(|q_2|^2 + \cdots + |q_N|^2) = |q_1|^4 + |q_2|^2(1-|q_1|^2) \geq 1-\epsilon. \tag{47}$$

然后，由于 $q_1 \geq q_2$，我们有

$$|q_1|^4 + |q_1|^2(1-|q_1|^2) \geq |q_1|^4 + |q_2|^2(1-|q_1|^2) \geq 1-\epsilon \implies |q_1|^2 \geq 1-\epsilon. \tag{48}$$

由于 $\mathbf{q} \in \mathbb{S}^{N-1}$，因此我们知道

$$|q_2|^2 + \cdots + |q_N|^2 \leq \epsilon. \tag{49}$$

此外，由于 $|q_1| \leq 1$，(48) 还意味着

$$\epsilon \geq 1 - |q_1|^2 \geq 1 - |q_1| \implies (1-|q_1|)^2 \leq \epsilon^2. \tag{50}$$

结合 (49) 和 (50)，并利用 $\epsilon \leq 1$ 这一事实，我们得到

$$\left\||\mathbf{q}| - \mathbf{e}_1^0\right\|_2^2 = (|q_1|-1)^2 + |q_2|^2 + \cdots + |q_N|^2 \leq \epsilon^2 + \epsilon \leq 2\epsilon. \tag{51}$$

令 $\mathbf{e}_k^{\theta} = \mathbf{e}_1^{j\angle q_1}$。于是，我们得到$\left\|\mathbf{q} - \mathbf{e}_k^{\theta}\right\|_2^2 = \left\||\mathbf{q}| - \mathbf{e}_1^0\right\|_2^2 \leq 2\epsilon$，如所需。注意，这个引理是将 [25,引理 36] 推广到复值数的一个改编版本。

F. 引理 *2* 的辅助引理

引理 13. 设 $\mathbf{q}_0, \mathbf{q}_1, \ldots, \mathbf{q}_{K-1}$是单位球面上的 K 向量：$\mathbf{q}_k \in \mathbb{S}^{N-1}, \forall k \in [K]$。对于 $\epsilon \in [0,1]$，若

$$\frac{1}{K}\sum_{k=1}^{K}\|\mathbf{q}_k\|_4^4 \geq 1-\epsilon, \tag{52}$$

则存在 $\exists m_1, m_2, \ldots, m_K \in [N]$，使得

$$\frac{1}{K}\sum_{k=1}^{K}\left\|\mathbf{q}_k - \mathbf{e}_{m_k}^{\theta_k}\right\| \leq 2\epsilon, \tag{53}$$

其中 $\mathbf{e}_{m_k}^{\theta_k} \in \mathbb{C}^N$满足 $[\mathbf{e}_{m_k}^{\theta_k}]_{m_k} = e^{j\theta_k}, \theta_k \in \mathbb{R}$，且 $[\mathbf{e}_{m_k}^{\theta_k}]_\ell = 0, \forall \ell \neq m_k$。

证明过程如下：我们可以将 (52) 重写为

$$\sum_{k=1}^{K}\|\mathbf{q}_k\|_4^4 \geq K - K\epsilon. \tag{54}$$

由于 $\forall k \in [K]$, $0 \le \|\mathbf{q}_k\|_4^4 \le \|\mathbf{q}_k\|_2^4 = 1$，我们可以假设

$$\|\mathbf{q}_k\|_4^4 \le 1 - \alpha_k \epsilon, \forall k \in [K], \tag{55}$$

其中，我们引入满足 $\alpha_k \epsilon \in [0,1], \forall k \in [K]$ 且的辅助变量 α_k

$$\sum_{k=1}^{K} \alpha_k \le K \tag{56}$$

必须成立，才能同时满足 (54) 和 (55)。根据引理 12，存在 $\mathbf{e}_{m_k}^{\theta_k}$，使得

$$\left\|\mathbf{q}_k - \mathbf{e}_{m_k}^{\theta_k}\right\|_2^2 \le 2\alpha_k \epsilon, \forall k \in [K]. \tag{57}$$

结合 (56)，我们得到 $\frac{1}{K}\sum_{k=1}^{K}\left\|\mathbf{q}_k - \mathbf{e}_{m_k}^{\theta_k}\right\|_2^2 \le \frac{1}{K}\sum_{k=1}^{K} 2\alpha_k\epsilon = \frac{2\epsilon}{K}\sum_{k=1}^{K}\alpha_k \le 2\epsilon$，如所需。我们注意到，这个引理是对 [25, 引理 37] 在复值数情形下的推广。

APPENDIX B
PROOFS OF SECTION IV

A. 引理5的证明

注意到

$$\|\mathbf{M} - \mathbf{A}\|_{\mathrm{F}}^2 = \mathrm{tr}((\mathbf{M} - \mathbf{A})(\mathbf{M} - \mathbf{A})^{\mathrm{H}}) = \mathrm{tr}(\mathbf{I} - \mathbf{A}\mathbf{M}^{\mathrm{H}} - \mathbf{M}\mathbf{A}^{\mathrm{H}} + \mathbf{A}\mathbf{A}^{\mathrm{H}}) \tag{58}$$

$$= N - 2\Re\{\mathrm{tr}(\mathbf{A}\mathbf{M}^{\mathrm{H}})\} + \mathrm{tr}(\mathbf{A}\mathbf{A}^{\mathrm{H}}). \tag{59}$$

由于 $\mathbf{tr(AA^H)}$ 是常数，我们知道

$$\underset{\mathbf{M}\in U(N;\mathbb{C})}{\arg\min} \|\mathbf{M} - \mathbf{A}\|_{\mathrm{F}}^2 = \underset{\mathbf{M}\in U(N;\mathbb{C})}{\arg\max} \Re\{\mathrm{tr}(\mathbf{A}\mathbf{M}^{\mathrm{H}})\}. \tag{60}$$

令$\mathbf{U\Sigma V^H}$ =*SVD*(**A**)，则

$$\Re\{\mathrm{tr}(\mathbf{A}\mathbf{M}^{\mathrm{H}})\} = \Re\{\mathrm{tr}(\mathbf{U}\mathbf{\Sigma}\mathbf{V}^{\mathrm{H}}\mathbf{M}^{\mathrm{H}})\} = \Re\{\mathrm{tr}(\mathbf{\Sigma}\mathbf{V}^{\mathrm{H}}\mathbf{M}^{\mathrm{H}}\mathbf{U})\} \le \left|\mathrm{tr}(\mathbf{\Sigma}\mathbf{V}^{\mathrm{H}}\mathbf{M}^{\mathrm{H}}\mathbf{U})\right| \tag{61}$$

$$\le \sum_{i=1}^{N} \sigma_i(\mathbf{A})\sigma_i(\mathbf{V}^{\mathrm{H}}\mathbf{M}^{\mathrm{H}}\mathbf{U}) = \sum_{i=1}^{N} \sigma_i(\mathbf{A}), \tag{62}$$

其中，第二个不等式由冯·诺伊曼迹不等式得到；当且仅当$\mathbf{V^H M^H U}$是对角复置换矩阵时，等号成立。这意味着，对于某个$\mathbf{C} \in \mathrm{CP}(N)\, C_{i,k} = 0, \forall i \ne k$，有$\mathbf{V^H M^H U} = \mathbf{C} \implies \mathbf{M} = \mathbf{U C^H V^H}$。由于（58）不受**C**所引起的相位移影响，为简便起见，我们可以取 $\mathbf{C} = \mathbf{I}$。因此，$\mathbf{M} = \mathbf{U}\mathbf{V}^{\mathrm{H}}$。

B. 引理 6 的证明

令$\mathbf{U\Sigma V^H} = |\mathbf{W}|^{\circ 2} \circ \mathbf{W}$ 表示$SVD\big(|\mathbf{W}|^{\circ 2} \circ \mathbf{W}\big)$。

i) 若$\mathbf{W} \in U(N;\mathbb{C})$ 是算法 1.2 的不动点，则有 $\mathcal{P}_{U(N;\mathbb{C})}\big(|\mathbf{W}|^{\circ 2} \circ \mathbf{W}\big) = \mathbf{W}$；等价地，$\mathbf{W} = \mathbf{U}\mathbf{V}^{\mathrm{H}}$。利用这一等式，我们将（$|\mathbf{W}|^{\circ 2} \circ \mathbf{W}$）$^{\mathrm{H}}\mathbf{W}$简化如下：

$$(|\mathbf{W}|^{\circ 2} \circ \mathbf{W})^{\mathrm{H}}\mathbf{W} = (\mathbf{U}\mathbf{\Sigma}\mathbf{V}^{\mathrm{H}})^{\mathrm{H}}\mathbf{W} = \mathbf{V}\mathbf{\Sigma}\mathbf{U}^{\mathrm{H}}\mathbf{W} = \mathbf{V}\mathbf{\Sigma}\mathbf{U}^{\mathrm{H}}\mathbf{U}\mathbf{V}^{\mathrm{H}} = \mathbf{V}\mathbf{\Sigma}\mathbf{V}^{\mathrm{H}}, \tag{63}$$

这意味着 $(|\mathbf{W}|^{\circ 2} \circ \mathbf{W})^{\mathrm{H}}\mathbf{W}$ 是厄米对称的。因此，我们有

$$(|\mathbf{W}|^{\circ 2} \circ \mathbf{W})^{\mathrm{H}}\mathbf{W} = \mathbf{W}^{\mathrm{H}}(|\mathbf{W}|^{\circ 2} \circ \mathbf{W}), \tag{64}$$

根据引理 3，$\mathbf{W}$ 是酉群上 ℓ^4-范数的临界点。

ii）如果 $\mathbf{W}$ 是酉群上 ℓ^4-范数的临界点，那么我们有 $(|\mathbf{W}|^{\circ 2} \circ \mathbf{W})^{\mathrm{H}}\mathbf{W} = \mathbf{W}^{\mathrm{H}}(|\mathbf{W}|^{\circ 2} \circ \mathbf{W})$。因此，我们必有 奇异值分解$\big(|\mathbf{W}|^{\circ 2} \circ \mathbf{W})^{\mathrm{H}}\mathbf{W}\big)$ = 奇异值分解$\big(\mathbf{W}^{\mathrm{H}}(|\mathbf{W}|^{\circ 2} \circ \mathbf{W})\big)$。此外，根据奇异值分解的旋转不变性，我们有如下结果：

$$SVD\big((|\mathbf{W}|^{\circ 2} \circ \mathbf{W})^{\mathrm{H}}\mathbf{W}\big) = \mathbf{V}\mathbf{\Sigma}\mathbf{U}^{\mathrm{H}}\mathbf{W} = SVD\big(\mathbf{W}^{\mathrm{H}}(|\mathbf{W}|^{\circ 2} \circ \mathbf{W})\big) = \mathbf{W}^{\mathrm{H}}\mathbf{U}\mathbf{\Sigma}\mathbf{V}^{\mathrm{H}}. \tag{65}$$

最后，根据奇异值分解的唯一性，我们有 $\mathbf{V} = \mathbf{W}^{\mathrm{H}}\mathbf{U} \implies \mathbf{W} = \mathbf{U}\mathbf{V}^{\mathrm{H}}$。

C. 引理 7 的证明

~考虑以下目标函数 $h\colon\ U(N;\mathbb{C}) \to \mathbb{R}^+$:

$$\tilde{h}(\mathbf{A}) = \begin{cases} \frac{\alpha}{4}\|\mathbf{A}\|_4^4 + \frac{1}{2}\|\mathbf{A}\|_{\mathrm{F}}^2, & \text{when } \alpha < \infty \\ \|\mathbf{A}\|_4^4, & \text{when } \alpha = +\infty. \end{cases} \tag{66}$$

˜注意， $h(\mathbf{A})$ 在两种情况下都是凸函数。还要注意，Stiefel流形 $U(N;\mathbb{C})$ 是一个紧流形，因此根据 [58, 定理 1],，我们知道以下迭代更新，

$$\mathbf{A}_{k+1} \in \underset{\mathbf{W}\in U(N;\mathbb{C})}{\arg\max}\ \tilde{h}(\mathbf{A}_k) + \langle \partial\tilde{h}(\mathbf{A}_k), \mathbf{W}-\mathbf{A}_k\rangle \tag{67}$$

将找到一个鞍点 $\tilde{h}(\mathbf{A})$ 对于任意初始化 $\mathbf{A}_0 \in U(N;\ \mathbb{C})$，其中 $\langle\cdot,\ \cdot\rangle$: $U(N;\ \mathbb{C})\times U(N;\ \mathbb{C})\to\mathbb{C}$定义为

$$\langle \mathbf{W}_1, \mathbf{W}_2\rangle = \mathrm{tr}(\mathbf{W}_1^{\mathrm{H}}\mathbf{W}_2). \tag{68}$$

因此，将（68）代入（67）可得

$$\mathbf{A}_{k+1} = \begin{cases} \mathcal{P}_{U(N;\mathbb{C})}\big(\alpha|\mathbf{A}_k|^{\circ 2}\circ\mathbf{A}_k + \mathbf{A}_k\big) & \text{when } \alpha<\infty \\ \mathcal{P}_{U(N;\mathbb{C})}\big(|\mathbf{A}_k|^{\circ 2}\circ\mathbf{A}_k\big) & \text{when } \alpha=+\infty, \end{cases} \tag{69}$$

其中， $\mathcal{P}_{U(N;\mathbb{C})}$: $\mathbb{C}^{N\times N}\to U(N;\ \mathbb{C})$ 是到 $U(N;\ \mathbb{C})$ 上的投影

$$\mathcal{P}_{U(N;\mathbb{C})}(\mathbf{A}) = \mathbf{U}\mathbf{V}^{\mathrm{H}},\ \text{where } \mathbf{U}\mathbf{\Sigma}\mathbf{V}^{\mathrm{H}} = SVD(\mathbf{A})\,. \tag{70}$$

此外，注意到 $\|\mathbf{A}\|_{\mathrm{F}}^2 = N$对于酉 $\mathbf{A}$ 是一个常数，因此，在 $U(N;\ \mathbb{C})$ 上寻找 $\tilde{h}(\mathbf{A})$ 的临界点等价于寻找 $\|\mathbf{A}\|_4^4$ 的临界点。

D. 引理 *8* 的证明

令 $\mathbf{A} = \mathbf{D} + \mathbf{N}$，其中$\mathbf{D}$ 是$\mathbf{A}$ 的对角部分，而$\mathbf{N}$ 是非对角部分。因此，我们有 ve:

$$\big\||\mathbf{N}|^{\circ 2}\circ\mathbf{N}\big\|_{\mathrm{F}} \le \|\mathbf{N}\|_{\mathrm{F}}^2\,\|\mathbf{N}\|_{\mathrm{F}} \le \|\mathbf{A}-\mathbf{C}\|_{\mathrm{F}}^{3/2} = \epsilon^{3/2}, \tag{71}$$

其中，第一个不等式通过柯西–施瓦茨不等式得到，第二个不等式成立是因为 N 是 A –C 的非对角部分。我们可以将 |A|$^{\circ 2}\circ$A 视为 |D|$^{\circ 2}\circ$D 加上一个范数至多为$\epsilon^{3/2}$的小扰动 |N|$^{\circ 2}\circ$N。令 $\mathrm{U_A}\mathbf{\Sigma_A}\mathrm{V_A^H} = \mathrm{SVD}\big(|\mathrm{A}|^{\circ 2}\circ\mathrm{A}\big)$、$\mathrm{A}' = \mathrm{Q_A} = \mathrm{U_A}\,\mathrm{V_A^H}$，以及 $\mathrm{U_D}\mathbf{\Sigma_D}\mathrm{V_D^H} = \mathrm{SVD}\big(|\mathrm{D}|^{\circ 2}\circ\mathrm{D}\big)$、$\mathrm{Q_D} = \mathrm{U_D}\mathrm{V_D^H}$。根据 [59,定理 1], ，我们有

$$\|\mathbf{Q_A}-\mathbf{Q_D}\|_{\mathrm{F}} \le \frac{2\big\||\mathbf{A}|^{\circ 2}\circ\mathbf{A}-|\mathbf{D}|^{\circ 2}\circ\mathbf{D}\big\|_{\mathrm{F}}}{\sigma_N(|\mathbf{A}|^{\circ 2}\circ\mathbf{A})+\sigma_N(|\mathbf{D}|^{\circ 2}\circ\mathbf{D})} = \frac{2\big\||\mathbf{N}|^{\circ 2}\circ\mathbf{N}\big\|_{\mathrm{F}}}{\sigma_N(|\mathbf{A}|^{\circ 2}\circ\mathbf{A})+\sigma_N(|\mathbf{D}|^{\circ 2}\circ\mathbf{D})}. \tag{72}$$

由于 $|\mathbf{D}|^{\circ 2}\circ\mathbf{D}$是对角矩阵，$\mathbf{Q}_D = \mathbf{I}_N$，并且 $\sigma_N(|\mathbf{D}|^{\circ 2}\circ\mathbf{D}) = \min_k |A_{k,k}|^3$。此外 e,

$$\epsilon = \|\mathbf{A}-\mathbf{C}\|_{\mathrm{F}}^2 = \sum_{k=1}^{N}(|A_{k,k}|-1)^2 + \sum_{k=1}^{N}\sum_{\substack{\ell=1\\ \ell\neq k}}^{N}|A_{k,\ell}|^2 = \sum_{k=1}^{N}\sum_{\ell=1}^{N}|A_{k,\ell}|^2 - 2\sum_{k=1}^{N}|A_{k,k}| + N \tag{73}$$

$$= 2N - 2\sum_{k=1}^{N}|A_{k,k}| \iff \sum_{k=1}^{N}A_{k,k} = N - \frac{\epsilon}{2}. \tag{74}$$

不失一般性 性，我们可以假设 $1\ge|A_{0,\ 0}|\ge|A_{1,\ 1}|\ge\cdots\ge|A_{N-1,\ N-1}|>0$，因此有

$$|A_{N-1,N-1}| = N - \frac{\epsilon}{2} - \sum_{k=1}^{N-1}|A_{k,k}| \ge N - \frac{\epsilon}{2} - (N-1) = 1 - \frac{\epsilon}{2}, \tag{75}$$

这意味着

$$\sigma_N(|\mathbf{D}|^{\circ 2}\circ\mathbf{D}) = \min_k |A_{k,k}|^3 \ge \left(1-\frac{\epsilon}{2}\right)^3. \tag{76}$$

我们使用 [60, 定理 2.17] , 其表述见 [25, 引理 46] , 如下：设 $\mathrm{G}\in\mathbb{C}^{N\times N}$ 为一般的满秩矩阵，且 $\mathrm{G} = \mathrm{BM}$，其中 M 是一个对角矩阵，满足 $M_{k,k} = \|\mathrm{g}_k\|_2$，使得 B 的矩阵二范数为 1，即 $\|\mathrm{B}\|_2 = 1$。设 $\delta\mathrm{G} = \delta\mathrm{BM}$ 是 G 的一个扰动，满足 $\|\delta\mathrm{B}\|_2 \le \sigma_N(\mathrm{B})$。那么，$\frac{\|\sigma_i(\mathrm{G}) - \sigma_i(\mathrm{G}+\delta\mathrm{G})\|}{\sigma_i(\mathrm{G})} \le \|\delta\mathrm{B}\sigma_N\|_2(\mathrm{B})$。

我们将变量代换为 $\widetilde{\mathrm{G}} = |\mathrm{D}|^{\circ 2}\circ\mathrm{D} = \mathrm{BM}$ 和 $\delta\mathrm{G} = |\mathrm{A}|^{\circ 2}\circ\mathrm{A} - |\mathrm{D}|^{\circ 2}\circ\mathrm{D} = |\mathrm{N}|^{\circ 2}\circ\mathrm{N} = \delta\mathrm{BM}$。这里，我们知道 $\mathrm{B} = \mathrm{I}$，$\mathrm{M} = |\mathrm{D}|^{\circ 2}\circ\mathrm{D}$ $\delta\mathrm{B} = \delta\mathrm{G}\mathrm{M}^{-1} = (|\mathrm{N}|^{\circ 2}\circ\mathrm{N})(|\mathrm{D}|^{\circ 2}\circ\mathrm{D})^{-1}$，因此

$$\frac{|\sigma_N(|\mathbf{A}|^{\circ 2}\circ\mathbf{A}) - \sigma_N(|\mathbf{D}|^{\circ 2}\circ\mathbf{D})|}{\sigma_N(|\mathbf{D}|^{\circ 2}\circ\mathbf{D})} \le \frac{\|\delta\mathbf{B}\|_2}{\sigma_N(\mathbf{B})} = \|\delta\mathbf{B}\|_2\,, \tag{77}$$

这意味着

$$|\sigma_N(|\mathbf{A}|^{\circ 2}\circ\mathbf{A}) - \sigma_N(|\mathbf{D}|^{\circ 2}\circ\mathbf{D})| \le \big\|(|\mathbf{N}|^{\circ 2}\circ\mathbf{N})(|\mathbf{D}|^{\circ 2}\circ\mathbf{D})^{-1}\big\|_2\,\sigma_N(|\mathbf{D}|^{\circ 2}\circ\mathbf{D}) \tag{78}$$

$$\le \big\||\mathbf{N}|^{\circ 2}\circ\mathbf{N}\big\|_2\,\big\|(|\mathbf{D}|^{\circ 2}\circ\mathbf{D})^{-1}\big\|_2\,\sigma_N(|\mathbf{D}|^{\circ 2}\circ\mathbf{D}) \tag{79}$$

$$\leq \left\||\mathbf{N}|^{\circ 2}\circ\mathbf{N}\right\|_2\left(1-\frac{\epsilon}{2}\right)^{-3}\sigma_N(|\mathbf{D}|^{\circ 2}\circ\mathbf{D}) \tag{80}$$

$$\leq \epsilon^{3/2}\left(1-\frac{\epsilon}{2}\right)^{-3}\sigma_N(|\mathbf{D}|^{\circ 2}\circ\mathbf{D}), \tag{81}$$

其中，上述四个不等式分别由式（77）、柯西–施瓦茨不等式、式（76）和式（71）得出。因此，我们有

$$-\epsilon^{3/2}\left(1-\frac{\epsilon}{2}\right)^{-3}\sigma_N(|\mathbf{D}|^{\circ 2}\circ\mathbf{D})\leq\sigma_N(|\mathbf{A}|^{\circ 2}\circ\mathbf{A})-\sigma_N(|\mathbf{D}|^{\circ 2}\circ\mathbf{D}). \tag{82}$$

因此，由（82）和（76）分别可得以下两个不等式:

$$\sigma_N(|\mathbf{A}|^{\circ 2}\circ\mathbf{A})+\sigma_N(|\mathbf{D}|^{\circ 2}\circ\mathbf{D})\geq\left(2-\epsilon^{3/2}\left(1-\frac{\epsilon}{2}\right)^{-3}\right)\sigma_N(|\mathbf{D}|^{\circ 2}\circ\mathbf{D}) \tag{83}$$

$$\geq 2\left(1-\frac{\epsilon}{2}\right)^{3}-\epsilon^{3/2}. \tag{84}$$

将（71）和（84）代入（72），可得

$$\|\mathbf{Q_A}-\mathbf{Q_D}\|_{\mathrm{F}}\leq\frac{2\left\||\mathbf{N}|^{\circ 2}\circ\mathbf{N}\right\|_{\mathrm{F}}}{\sigma_N(|\mathbf{A}|^{\circ 2}\circ\mathbf{A})+\sigma_N(|\mathbf{D}|^{\circ 2}\circ\mathbf{D})}\leq\frac{2\epsilon^{3/2}}{2\left(1-\frac{\epsilon}{2}\right)^{3}-\epsilon^{3/2}}. \tag{85}$$

回顾$\mathbf{Q_A}=\mathbf{A}'$和$\mathbf{Q}_D=\mathbf{I}_N$，可得 $\|\mathbf{A}'-\mathbf{C}'\|_{\mathrm{F}}\leq\|\mathbf{A}'-\mathbf{I}\|_{\mathrm{F}}=\|\mathbf{Q_A}-\mathbf{Q_D}\|_{\mathrm{F}}\leq O(\epsilon^3)$。此外，当满足以下条件时，该运算为压缩映射

$$\frac{2\epsilon^{3/2}}{2\left(1-\frac{\epsilon}{2}\right)^{3}-\epsilon^{3/2}}\leq\|\mathbf{A}-\mathbf{D}\|_{\mathrm{F}}=\epsilon^{1/2} \tag{86}$$

$$\iff 2\epsilon-2\left(1-\frac{\epsilon}{2}\right)^{3}+\epsilon^{3/2}\leq 0\iff\epsilon<0.394, \tag{87}$$

至此，证明完毕。

E. 引理 *9* 的证明

由于满秩矩阵的左右奇异向量在相差复相位移的意义下唯一，因此 UV^{H} 唯一。由于 A 是酉矩阵，A 的一个奇异值分解为 $\mathrm{U}=\mathrm{I}_N,\boldsymbol{\Sigma}=\mathrm{I}_N,\mathrm{V}=\mathrm{A}$。于是 $\mathrm{A}_1=\mathrm{DA}=\mathrm{DV}=\mathrm{U}_1\boldsymbol{\Sigma}_1\mathrm{V}_1,$，因此存在 A_1的一个奇异值分解，使得 $\mathrm{U}_1=\mathrm{I}_N,\boldsymbol{\Sigma}_1=\mathrm{D}$，且 $\mathrm{V}_1=\mathrm{A}$。A 的另一个奇异值分解为 $\mathrm{U}=\mathrm{A}$，$\boldsymbol{\Sigma}=\mathrm{I}_N$，以及 $\mathrm{V}=\mathrm{I}_N$。于是，$\mathrm{A}_2=\mathrm{AD}=\mathrm{UD}=\mathrm{U}_2\boldsymbol{\Sigma}_2\mathrm{V}_2,$，因此存在 A_2的一个奇异值分解，使得 $\mathrm{U}_2=\mathrm{A}$ $\boldsymbol{\Sigma}_2=\mathrm{D}$，且 $\mathrm{V}_2=\mathrm{I}_N$。因此，我们有 $\mathrm{U}_1\mathrm{V}_1^{\mathrm{H}}=\mathrm{U}_2\mathrm{V}_2^{\mathrm{H}}=\mathrm{A}$。

F. 引理 *10* 的证明

设 $\mathbf{A}_t=\mathbf{A}$。若满足 $\nabla_{\mathbf{A}_t}\hat{g}(\mathbf{A}_t)=\mathbf{D}\mathbf{A}_t$或 $\nabla_{\mathbf{A}_t}\hat{g}(\mathbf{A}_t)=\mathbf{A}_t\mathbf{D}$，则根据引理 9，$\mathbf{A}_{t+1}=\mathcal{P}_{U(N;\mathbb{C})}(\nabla_{\mathbf{A}_t}\hat{g}(\mathbf{A}_t))=\mathbf{A}_t$成立。因此，矩阵 $\mathbf{A}$ 是算法 1.1 的一个不动点。

Appendix C
Proofs of Section V

A. 定理 *1* 的证明

为简洁起见，设 $\mathbf{F}=\mathbf{F}_B$。我们首先将式 (5) 中的目标函数展开如下:

$$\hat{g}(\mathbf{A})=\int_{\omega_1}\cdots\int_{\omega_L}\|\mathbf{A}\mathbf{y}\|_4^4\,f_{\Omega_1}(\omega_1)\cdots f_{\Omega_L}(\omega_L)\mathrm{d}\omega_1\cdots\mathrm{d}\omega_L. \tag{88}$$

令$\mathbf{A}=\mathbf{F}$，并将 $f_\Omega(\omega_\ell)=\frac{1}{2\pi}\mathbb{1}\{\omega\in(0,2\pi)\}$ 代入（88），可得到以下目标函数:

$$\hat{g}(\mathbf{F})=\frac{1}{(2\pi)^L}\int_{\omega_1=0}^{2\pi}\cdots\int_{\omega_L=0}^{2\pi}\|\mathbf{F}\mathbf{y}\|_4^4\,\mathrm{d}\omega_1\cdots\mathrm{d}\omega_L \tag{89}$$

$$=\frac{1}{(2\pi)^L}\int_{\omega_1=0}^{2\pi}\cdots\int_{\omega_L=0}^{2\pi}\sum_{i=1}^{B}\left|\sum_{n=1}^{B}F_{i,n}y_n\right|^4\mathrm{d}\omega_1\cdots\mathrm{d}\omega_L \tag{90}$$

$$=\sum_{i=1}^{B}\sum_{n=1}^{B}\sum_{p=1}^{B}\sum_{q=1}^{B}\sum_{r=1}^{B}F_{i,n}^*F_{i,p}F_{i,q}^*F_{i,r}C_L(p,q,r,n), \tag{91}$$

其中，我们定义 $C_L(p,q,r,n) \triangleq \frac{1}{(2\pi)^L}\int_{\omega_1=0}^{2\pi}\cdots\int_{\omega_L=0}^{2\pi} y_p y_q^* y_r y_n^* \mathrm{d}\omega_1\cdots\mathrm{d}\omega_L$。将（23）中的 y_n 展开式代入，可得

$$C_L(p,q,r,n)=\frac{1}{(2\pi)^L}\int_{\omega_1=0}^{2\pi}\cdots\int_{\omega_L=0}^{2\pi}\Big(\sum_{\ell=1}^{L}c_\ell e^{j\omega_\ell p}\Big)\Big(\sum_{\ell=1}^{L}c_\ell^* e^{-j\omega_\ell q}\Big)\Big(\sum_{\ell=1}^{L}c_\ell e^{j\omega_\ell r}\Big)\times\Big(\sum_{\ell=1}^{L}c_\ell^* e^{-j\omega_\ell n}\Big)\mathrm{d}\omega_1\ldots\mathrm{d}\omega_L. \quad (92)$$

为证明$\mathbf{F}$ 是不动点，我们利用引理 10 证明存在一个满秩对角矩阵$\mathbf{D}\in\mathbb{C}^{B\times B}$，使得 $\nabla_{\mathbf{F}}\hat{g}(\mathbf{F})=\mathbf{F}\mathbf{D}$，即:

$$\frac{\partial\hat{g}(\mathbf{F})}{\partial\mathbf{f}_n^*}=D_{n,n}\mathbf{f}_n,\ \forall n\in[B]. \quad (93)$$

等价地，对于所有索引对 (i,n)，我们需要证明

$$\frac{\partial\hat{g}(\mathbf{F})}{\partial F_{i,n}^*}F_{i,n}^*=D_{n,n}, \quad (94)$$

这意味着，对于所有索引对 (i,n)，$\frac{\partial\hat{g}(\mathbf{F})}{\partial F_{i,n}^*}F_{i,n}^*$ 仅取决于 n（而不取决于 i）。将 $F_{i,p}=\frac{1}{\sqrt{B}}e^{\frac{-j2\pi}{B}(i-1)(p-1)}$ 代入式 (91)，我们计算得到如下梯度:

$$\frac{\partial\hat{g}(\mathbf{F})}{\partial F_{i,n}^*}=\sum_{p=1}^{B}\sum_{q=1}^{B}\sum_{r=1}^{B}F_{i,p}F_{i,q}^*F_{i,r}C_L(p,q,r,n)=\frac{1}{B\sqrt{B}}\sum_{p=1}^{B}\sum_{q=1}^{B}\sum_{r=1}^{B}e^{\frac{-j2\pi}{B}(i-1)((p-1)-(q-1)+(r-1))}C_L(p,q,r,n). \quad (95)$$

通过改变下标来简化梯度表达式，可得

$$\frac{\partial\hat{g}(\mathbf{F})}{\partial F_{i+1,n+1}^*}=\frac{1}{B\sqrt{B}}\sum_{p=0}^{B-1}\sum_{q=0}^{B-1}\sum_{r=0}^{B-1}e^{\frac{-j2\pi i(p-q+r)}{B}}C_L(p,q,r,n). \quad (96)$$

将式 (96) 两边乘以 $F_{i+1,n+1}^*=\frac{1}{\sqrt{B}}e^{\frac{j2\pi in}{B}}$，可得

$$\frac{\partial\hat{g}(\mathbf{F})}{\partial F_{i+1,n+1}^*}F_{i+1,n+1}^*=\frac{1}{B^2}\sum_{p=0}^{B-1}\sum_{q=0}^{B-1}\sum_{r=0}^{B-1}e^{\frac{-j2\pi i(p-q+r-n)}{B}}C_L(p,q,r,n). \quad (97)$$

现在，我们考察 $C_L(p,q,r,n)$：对于 $L=1$，有

$$C_1(p,q,r,n)=\frac{|c_1|^4}{2\pi}\int_0^{2\pi}e^{j\omega(p-q+r-n)}\mathrm{d}\omega=|c_1|^4\delta[p-q+r-n]. \quad (98)$$

对于初级视觉皮层，我们有

$$\begin{aligned}C_2(p,q,r,n)&=\frac{1}{(2\pi)^2}\int_{\omega_1=0}^{2\pi}\int_{\omega_2=0}^{2\pi}\Big(\sum_{\ell=1}^{2}c_\ell e^{j\omega_\ell p}\Big)\Big(\sum_{\ell=1}^{2}c_\ell^* e^{-j\omega_\ell q}\Big)\Big(\sum_{\ell=1}^{2}c_\ell e^{j\omega_\ell r}\Big)\Big(\sum_{\ell=1}^{2}c_\ell^* e^{-j\omega_\ell n}\Big)\mathrm{d}\omega_1\mathrm{d}\omega_2 \quad (99)\\
&=(|c_1|^4+|c_2|^4)\delta[p-q+r-n]+|c_1|^2|c_2|^2\delta[p-q]\delta[r-n]+|c_1|^2c_1c_2^*\delta[p-q+r]\delta[n]\\
&+|c_1|^2c_1^*c_2\delta[p-q-n]\delta[r]+|c_1|^2|c_2|^2\delta[p-q]\delta[r-n]+c_1|c_2|^2c_2^*\delta[r]\delta[p-q-n]\\
&+c_1^*|c_2|^2c_2\delta[n]\delta[p-q+r]+|c_1|^2c_1c_2^*\delta[p+r-n]\delta[q]+c_1|c_2|^2c_2^*\delta[p]\delta[-q+r-n]\\
&+c_1^2(c_2^*)^2\delta[p+r]\delta[q+n]+|c_1|^2|c_2|^2\delta[p-n]\delta[-q+r]+|c_1|^2c_1^*c_2\delta[-q+r-n]\delta[p]\\
&+c_1^*|c_2|^2c_2\delta[q]\delta[p+r-n]+|c_1|^2|c_2|^2\delta[-q+r]\delta[p-n]+(c_1^*)^2c_2^2\delta[q+n]\delta[p+r]. \quad (100)\end{aligned}$$

我们观察到如下结果：对于初级视觉皮层，$C_L(p,q,r,n)$ 等于一个缩放后的 $\delta[p-q+r-n]$。对于 $L=2$，$C_L(p,q,r,n)$ 是以下两类项的线性组合：(i) $\delta[p-q+r-n]$ 项；(ii) 两个 $\delta[\cdot]$-函数相乘得到的项；在这两种情况下，每一项非零当且仅当 $p-q+r-n=0$。类似地，对于任意 L，$C_L(p,q,r,n)$ 是由若干项构成的线性组合，每一项是至多 L 个 $\delta[\cdot]$ 函数的乘积，且每一项非零当且仅当 $p-q+r-n=0$ 成立。因此，我们令 $p-q+r-n=0$，将 (97) 重写为

$$\frac{\partial\hat{g}(\mathbf{F})}{\partial F_{i+1,n+1}^*}F_{i+1,n+1}^*=\frac{1}{B^2}\sum_{p=0}^{B-1}\sum_{q=0}^{B-1}\sum_{r=0}^{B-1}C_L(p,q,r,n). \quad (101)$$

显然，$\frac{\partial\hat{g}(\mathbf{F})}{\partial F_{i,n}^*}F_{i,n}^*$ 与 i 无关，因此，对于任意 L，$\mathbf{F}$ 都是随机数据模型 $\mathbf{y}$（如式 (23) 所示）下算法 1.1 的不动点。

B. 定理 *2* 的证明

接下来，我们始终将 L 保留为变量，并仅在最后一步施加 $L=1$，以凸显这一约束在哪一步变得必要。

我们的目标是证明引理 11 中的一阶和二阶导数条件得到满足。通过将式 (21) 中关于 $\mathbf{y}$ 的一阶导数表达式按照式 (23) 重写，令 $\mathbf{A}=\mathbf{F}$，并代入 $f_\Omega(\omega_\ell)=\frac{1}{2\pi}\mathbb{1}\{\omega\in(0,2\pi)\}$，我们的目标是证明下式成立：

$$\begin{aligned}&\frac{\partial\int_{\omega_1}\cdots\int_{\omega_L}\|\mathbf{G}(i,k,\alpha)\mathbf{x}\|_4^4\,f_{\Omega_1}(\omega_1)\cdots f_{\Omega_L}(\omega_L)\mathrm{d}\omega_1\cdots\mathrm{d}\omega_L}{\partial\alpha}\Big|_{\alpha=0}\\&=\frac{4}{(2\pi)^L}\int_{\omega_1=0}^{2\pi}\cdots\int_{\omega_L=0}^{2\pi}\Re\{x_ix_k(x_i^*)^2-x_ix_k(x_k^*)^2\}\mathrm{d}\omega_1\cdots\mathrm{d}\omega_L=0,\end{aligned}\tag{102}$$

其中，$\mathbf{x}=\mathbf{F}\mathbf{y}$。为证明式 (102) 成立，我们首先化简以下积分：

$$\begin{aligned}&\frac{4}{(2\pi)^L}\int_{\omega_1=0}^{2\pi}\cdots\int_{\omega_L=0}^{2\pi}x_ix_k(x_i^*)^2\mathrm{d}\omega_1\ldots\mathrm{d}\omega_L\\&=\frac{4}{(2\pi)^L}\sum_{p=1}^{B}\sum_{q=1}^{B}\sum_{r=1}^{B}\sum_{n=1}^{B}F_{i,p}F_{k,r}F_{i,q}^*F_{i,n}^*\int_{\omega_1=0}^{2\pi}\cdots\int_{\omega_L=0}^{2\pi}y_py_q^*y_ry_n^*\mathrm{d}\omega_1\ldots\mathrm{d}\omega_L\qquad(103)\\&=\frac{4}{B^2}\sum_{p=0}^{B-1}\sum_{q=0}^{B-1}\sum_{r=0}^{B-1}\sum_{n=0}^{B-1}e^{-\frac{j2\pi(kr+i(p-q-n))}{B}}C_L(p,q,r,n).\qquad(104)\end{aligned}$$

这里，$C_L(p,q,r,n)$ 遵循式 (92) 中的定义。根据定理 1 的证明，我们知道，当且仅当 $p-q+r-n=0$ 成立时，$C_L(p,q,r,n)$ 非零。因此，

$$\frac{4}{(2\pi)^L}\int_{\omega_1=0}^{2\pi}\cdots\int_{\omega_L=0}^{2\pi}x_ix_k(x_i^*)^2\mathrm{d}\omega_1\ldots\mathrm{d}\omega_L=\frac{4}{B^2}\sum_{p=0}^{B-1}\sum_{q=0}^{B-1}\sum_{r=0}^{B-1}\sum_{n=0}^{B-1}e^{\frac{j2\pi(i-k)r}{B}}C_L(p,q,r,n).\tag{105}$$

观察式 (105) 左侧对 $(i-k)$ 的依赖关系，我们推得

$$\int_{\omega_1=0}^{2\pi}\cdots\int_{\omega_L=0}^{2\pi}x_ix_k(x_i^*)^2\mathrm{d}\omega_1\ldots\mathrm{d}\omega_L=\left(\int_{\omega_1=0}^{2\pi}\cdots\int_{\omega_L=0}^{2\pi}x_ix_k(x_k^*)^2\mathrm{d}\omega_1\ldots\mathrm{d}\omega_L\right)^*,\tag{106}$$

这意味着式 (102) 成立。因此，对于任意 L，函数 $\|\mathbf{G}(i,k,\alpha_{i,k})\mathbf{F}\mathbf{y}\|_4^4$（来自式 (18)）在 $\alpha_{i,k}=0$ 处，对于所有 $i,k\in[N]$，$i>k$，取得一个驻点；由此，我们证明了引理 11 中的第一个条件成立。

将式 (22) 中关于 $\mathbf{y}$ 的二阶导数表达式改写为式 (23) 中的形式，令 $\mathbf{A}=\mathbf{F}$ 并代入 $f_\Omega(\omega_\ell)=\frac{1}{2\pi}\mathbb{1}\{\omega\in(0,2\pi)\}$，定义

$$d_L(i,k)\triangleq\frac{4}{(2\pi)^L}\int_{\omega_1=0}^{2\pi}\cdots\int_{\omega_L=0}^{2\pi}\big(2\Re\{x_k^2(x_i^*)^2\}+4|x_i|^2|x_k|^2-|x_k|^4-|x_i|^4\big)\mathrm{d}\omega_1\cdots\mathrm{d}\omega_L.\tag{107}$$

我们的目标是证明 $d_L(i,k)<0$ 成立。在接下来的步骤中，我们再次利用以下事实：当且仅当 $p-q+r-n=0$ 成立时，$C_L(p,q,r,n)$ 非零。我们先考虑含有 $2\Re\{x_k^2(x_i^*)^2\}$ 的第一项的积分，首先简化 $x_k^2(x_i^*)^2$的积分：

$$\begin{aligned}&\frac{4}{(2\pi)^L}\int_{\omega_1=0}^{2\pi}\cdots\int_{\omega_L=0}^{2\pi}x_k^2(x_i^*)^2\mathrm{d}\omega_1\cdots\mathrm{d}\omega_L\\&=\frac{4}{(2\pi)^L}\sum_{p=1}^{B}\sum_{q=1}^{B}\sum_{r=1}^{B}\sum_{n=1}^{B}F_{k,p}F_{k,r}F_{i,q}^*F_{i,n}^*\int_{\omega_1=0}^{2\pi}\cdots\int_{\omega_L=0}^{2\pi}y_py_q^*y_ry_n^*\mathrm{d}\omega_1\ldots\mathrm{d}\omega_L\qquad(108)\\&=\frac{4}{B^2}\sum_{p=0}^{B-1}\sum_{q=0}^{B-1}\sum_{r=0}^{B-1}\sum_{n=0}^{B-1}e^{-\frac{j2\pi(k(p+r)+i(-q-n))}{B}}C_L(p,q,r,n)\qquad(109)\\&=\frac{4}{B^2}\sum_{p=0}^{B-1}\sum_{q=0}^{B-1}\sum_{r=0}^{B-1}\sum_{n=0}^{B-1}e^{\frac{j2\pi(i-k)(p+r)}{B}}C_L(p,q,r,n)\qquad(110)\end{aligned}$$

因此，我们得到

$$\frac{4}{(2\pi)^L}\int_{\omega_1=0}^{2\pi}\cdots\int_{\omega_L=0}^{2\pi}\big(2\Re\{x_k^2(x_i^*)^2\}\big)\mathrm{d}\omega_1\cdots\mathrm{d}\omega_L$$

$$= \frac{8}{B^2}\sum_{p=0}^{B-1}\sum_{q=0}^{B-1}\sum_{r=0}^{B-1}\sum_{n=0}^{B-1}\cos\left(\frac{2\pi}{B}(i-k)(p+r)\right)\Re\{C_L(p,q,r,n)\}. \tag{111}$$

现在，我们简化第二项的积分：

$$\frac{16}{(2\pi)^L}\int_{\omega_1=0}^{2\pi}\cdots\int_{\omega_L=0}^{2\pi} x_i x_i^* x_k x_k^* \mathrm{d}\omega_1\cdots\mathrm{d}\omega_L$$

$$= \frac{16}{(2\pi)^L}\sum_{p=1}^{B}\sum_{q=1}^{B}\sum_{r=1}^{B}\sum_{n=1}^{B} F_{i,p}F_{k,r}F_{i,q}^*F_{k,n}^*\int_{\omega_1=0}^{2\pi}\cdots\int_{\omega_L=0}^{2\pi} y_p y_q^* y_r y_n^* \mathrm{d}\omega_1\ldots\mathrm{d}\omega_L \tag{112}$$

$$= \frac{16}{B^2}\sum_{p=0}^{B-1}\sum_{q=0}^{B-1}\sum_{r=0}^{B-1}\sum_{n=0}^{B-1} e^{-\frac{j2\pi(k(r-n)+i(p-q))}{B}} C_L(p,q,r,n) \tag{113}$$

$$= \frac{16}{B^2}\sum_{p=0}^{B-1}\sum_{q=0}^{B-1}\sum_{r=0}^{B-1}\sum_{n=0}^{B-1} e^{\frac{j2\pi(i-k)(r-n)}{B}} C_L(p,q,r,n) \tag{114}$$

$$= \frac{16}{B^2}\sum_{p=0}^{B-1}\sum_{q=0}^{B-1}\sum_{r=0}^{B-1}\sum_{n=0}^{B-1}\cos\left(\frac{2\pi}{B}(i-k)(r-n)\right)\Re\{C_L(p,q,r,n)\}, \tag{115}$$

其中，最后一步是因为对绝对值平方进行积分，所得积分必为实数且非负。接下来，我们化简第三项中的积分：

$$\frac{4}{(2\pi)^L}\int_{\omega_1=0}^{2\pi}\cdots\int_{\omega_L=0}^{2\pi} x_i x_i^* x_i x_i^* \mathrm{d}\omega_1\cdots\mathrm{d}\omega_L \tag{116}$$

$$= \frac{4}{(2\pi)^L}\sum_{p=1}^{B}\sum_{q=1}^{B}\sum_{r=1}^{B}\sum_{n=1}^{B} F_{i,p}F_{i,r}F_{i,q}^*F_{i,n}^*\int_{\omega_1=0}^{2\pi}\cdots\int_{\omega_L=0}^{2\pi} y_p y_q^* y_r y_n^* \mathrm{d}\omega_1\ldots\mathrm{d}\omega_L \tag{117}$$

$$= \frac{4}{B^2}\sum_{p=0}^{B-1}\sum_{q=0}^{B-1}\sum_{r=0}^{B-1}\sum_{n=0}^{B-1} e^{-\frac{j2\pi i(p-q+r-n)}{B}} C_L(p,q,r,n) \tag{118}$$

$$= \frac{4}{B^2}\sum_{p=0}^{B-1}\sum_{q=0}^{B-1}\sum_{r=0}^{B-1}\sum_{n=0}^{B-1} C_L(p,q,r,n) = \frac{4}{B^2}\sum_{p=0}^{B-1}\sum_{q=0}^{B-1}\sum_{r=0}^{B-1}\sum_{n=0}^{B-1}\Re\{C_L(p,q,r,n)\}, \tag{119}$$

显然，这与 i无关；因此，我们可以推断，第四项中的积分等于第三项中的积分。最后，我们得到

$$d_L(i,k) = \frac{4}{B^2}\sum_{p=0}^{B-1}\sum_{q=0}^{B-1}\sum_{r=0}^{B-1}\sum_{n=0}^{B-1}\left(2\cos\left(\frac{2\pi}{B}(i-k)(p+r)\right)+4\cos\left(\frac{2\pi}{B}(i-k)(r-n)\right)-2\right)\Re\{C_L(p,q,r,n)\}. \tag{120}$$

遗憾的是，对于具有任意 L的多径模型，我们无法证明 $d_L(i,\ k) < 0$ 成立；因此，我们无法断定离散傅里叶变换是否会在任意 L下达到局部最大值。然而，对于 $L = 1$的单径模型，我们可以将$d_L(i,\ k)$化简为 $d_1(i,\ k) = \frac{8}{B^2}\left(3B\csc^2(\pi(i-k)/B) - \frac{2B^3+7B}{3}\right)$。现在，我们要证明以下结论：

$$d_1(i,k) < 0 \Leftrightarrow \csc^2(\pi(i-k)/B) < \frac{2B^2+7}{9} \Leftrightarrow \sin(\pi(i-k)/B) > \frac{3}{\sqrt{2B^2+7}}. \tag{121}$$

对于 $B \geq 2$，我们有

$$\sin(\pi(i-k)/B) \geq \sin(\pi/B) \geq \frac{\pi}{B} - \frac{1}{6}\left(\frac{\pi}{B}\right)^3 > \frac{3}{\sqrt{2B^2+7}}, \tag{122}$$

其中，第一个不等式成立是因为 $(i-k) \in \{1,\ldots,B-1\}$，第二个不等式根据 [61,定理 3.1], 中的下界得出，最后一个不等式在 $B \geq 1.75$ 时成立。因此，根据式 (122)，我们证明了 $d_1(i,\ k) < 0,\ \forall i,\ k \in [B],\ i > k$ 成立，这意味着引理 11 的第二个条件成立，并完成了证明。

C. 离散余弦变换的简化梯度

令$\mathbf{x} \triangleq \mathbf{C}_B\mathbf{y}$。我们有

$$x_{k+1} = \sqrt{\frac{2}{B}}\frac{1}{\sqrt{1+\delta[k]}}\sum_{n=0}^{B-1}\cos(\omega n+\phi)\cos\left(\frac{\pi}{B}(n+0.5)k\right). \tag{123}$$

由于$\mathbf{x} \in \mathbb{R}^B$，$f_\Omega(\omega_\ell) = \frac{1}{2\pi}\mathbb{1}\{\omega \in (0, 2\pi)\}$，且 $f_\Phi(\phi) = \frac{1}{2\pi}\mathbb{1}\{\phi \in (0, 2\pi)\}$，我们将式 (21) 改写为

$$\left.\frac{\partial \int_0^{2\pi}\int_0^{2\pi} \|\mathbf{G}(i,k,\alpha)\mathbf{x}\|_4^4 f_\Omega(\omega) f_\Phi(\phi) \mathrm{d}\omega \mathrm{d}\phi}{\partial \alpha}\right|_{\alpha=0} = \frac{4}{(2\pi)^2}\int_0^{2\pi} \left(x_i^3 x_k - x_k^3 x_i\right) \mathrm{d}\omega. \tag{124}$$

将 (123) 代入 (21)，再通过一系列繁琐的三角恒等式代数化简，可得

$$\begin{aligned}
&\left.\frac{\partial \int_0^{2\pi}\int_0^{2\pi} \|\mathbf{G}(i,k,\alpha)\mathbf{x}\|_4^4 f_\Omega(\omega) f_\Phi(\phi) \mathrm{d}\omega \mathrm{d}\phi}{\partial \alpha}\right|_{\alpha=0} \\
&= \frac{1}{B^2\sqrt{1+\delta[k]}} \sum_{p=0}^{B-1}\sum_{q=0}^{B-1}\sum_{r=0}^{B-1}\sum_{n=0}^{B-1} \cos\left(\frac{\pi}{B}\left(p+\frac{1}{2}\right)k\right)\cos\left(\frac{\pi}{B}\left(q+\frac{1}{2}\right)i\right) \\
&\quad \times \left(\cos\left(\frac{\pi}{B}\left(r+\frac{1}{2}\right)i\right)\cos\left(\frac{\pi}{B}\left(n+\frac{1}{2}\right)i\right) - \frac{1}{1+\delta[k]}\cos\left(\frac{\pi}{B}\left(r+\frac{1}{2}\right)k\right)\cos\left(\frac{\pi}{B}\left(n+\frac{1}{2}\right)k\right)\right) \\
&\quad \times \big(\delta[p-q+r-n] + \delta[p-q+r-n] + \delta[p-q-r+n]\big).
\end{aligned} \tag{125}$$

参考文献

[1] S. Taner 和 C. Studer，《离散傅里叶变换对波束空间大规模多用户多输入多输出通信的最优性》，发表于 IEEE 信息论国际研讨会（ISIT）论文集，2021年7月，第 2960–2965 页。[2] E. G. Larsson、F. Tufvesson、O. Edfors 和 T. L. Marzetta，《面向下一代无线系统的大规模多输入多输出》，IEEE 通信杂志，第 52 卷，第 2 期，第 186–195 页，2014年2月。[3] T. S. Rappaport、S. Sun、R. Mayzus、H. Zhao、Y. Azar、K. Wang、G. N. Wong、J. K. Schulz、M. Samimi 和 F. Gutierrez，《面向5G蜂窝通信的毫米波移动通信：这将成为现实！》，IEEE开放获取期刊，第 1 卷，第 335–349 页，2013年5月。[4] H. Yan、S. Ramesh、T. Gallagher、C. Ling 和 D. Cabric，《毫米波发射机波束成形架构中的性能、功耗和面积设计权衡》，IEEE 电路与系统杂志，第 19 卷，第 2 期，第 33–58 页，2019年5月。[5] K. Roth、H. Pirzadeh、A. L. Swindlehurst 和 J. A. Nossek，《面向多用户和不完美信道状态信息的低分辨率模数转换器：混合波束成形与数字波束成形比较》，IEEE 信号处理专题期刊，第 12 卷，第 3 期，第 484–498 页，2018年6月。[6] P. Skrimponis、S. Dutta、M. Mezzavilla、S. Rangan、S. H. Mirfarshbafan、C. Studer、J. Buckwalter 和 M. Rodwell，《移动毫米波和亚太赫兹接收机的功耗分析》，发表于第二届6G无线峰会论文集，2020年3月，第 1–5 页。[7] T. Rappaport、R. Heath、R. Daniels 和 J. Murdock，《毫米波无线通信》，普伦蒂斯·霍尔，2015年。[8] A. Sayeed 和 J. Brady，《面向毫米波频段高维多用户通信的波束空间多输入多输出》，发表于 IEEE 全球通信会议（GLOBECOM）论文集，2013年12月，第 3679–3684 页。[9] P. Schniter 和 A. Sayeed，《毫米波通信的信道估计与预编码器设计：稀疏方法》，发表于 Asilomar 信号、系统与计算会议论文集，2014年11月，第 273–277 页。[10] X. Gao、L. Dai、Z. Chen、Z. Wang 和 Z. Zhang，《波束空间毫米波大规模多输入多输出系统的近最优波束选择》，IEEE 通信快报，第 20 卷，第 5 期，第 1054–1057 页，2016年3月。[11] C. Chen、C. Tsai、Y. Liu、W. Hung 和 A. Wu，《压缩感知（CS）辅助的低复杂度毫米波多输入多输出系统波束空间混合预编码》，IEEE 信号处理汇刊，第 65 卷，第 6 期，第 1412–1424 页，2017年12月。[12] W. Ma 和 C. Qi，《采用混合预编码与合并的毫米波大规模多输入多输出系统波束空间信道估计》，IEEE 信号处理汇刊，第 66 卷，第 18 期，第 4839–4853 页，2018年9月。[13] M. Abdelghany、U. Madhow 和 A. T¨olli，《毫米波大规模多输入多输出的高效波束空间下行预编码》，发表于 Asilomar 信号、系统与计算会议，2019年11月，第 1459–1464 页。[14] S. H. Mirfarshbafan、A. Gallyas-Sanhueza、R. Ghods 和 C. Studer，《大规模多输入多输出毫米波系统的波束空间信道估计：算法与超大规模集成电路设计》，IEEE 电路与系统汇刊 I，第 1–14 页，2020年9月。[15] S. H. Mirfarshbafan 和 C. Studer，《大规模多用户多输入多输出毫米波系统的稀疏波束空间均衡》，发表于 IEEE 声学、语音与信号处理国际会议（ICASSP）论文集，2020年5月，第 1773–1777 页。[16] M. Mahdavi、O. Edfors、V. ¨Owall 和 L. Liu，《角域大规模多输入多输出检测：算法、实现与设计权衡》，IEEE 电路与系统汇刊，第 67 卷，第 6 期，第 1948–1961 页，2020年1月。[17] S. H. Mirfarshbafan、S. Taner 和 C. Studer，《SMUL-FFT：流式无乘法器快速傅里叶变换》，IEEE 电路与系统汇刊 II，第 68 卷，第 5 期，第 1715–1719 页，2021年3月。[18] E. G¨on¨ultas、S. Taner、A. Gallyas-Sanhueza、S. H. Mirfarshbafan 和 C. Studer，《面向全数字毫米波大规模多用户多输入多输出的硬件感知波束空间预编码》，IEEE 通信快报，2021年。[19] A. Osinsky、A. Ivanov 和 D. Yarotsky，《相干大规模多输入多输出中稀疏信道估计的空间去噪》，发表于 IEEE 车辆技术会议（VTC）论文集，2021年9月，第 1–5 页。[20] Y. Dai、M. E. Rasekh、S. H. Mirfarshbafan、H. Liew、A. Gallyas-Sanhueza、J. Dunn、U. Madhow、C. Studer 和 B. Nikoli´c，《分布式大规模多输入多输出基带处理系统的可适配、可扩展生成器》，《信号处理系统杂志》，第 94 卷，第 10 期，第 989–1003 页，2022年。[21] H. He、R. Wang、W. Jin、S. Jin、C.-K. Wen 和 G. Y. Li，《宽带毫米波多输入多输出的波束空间信道估计：一种模型驱动的无监督学习方法》，IEEE《无线通信汇刊》，第 22 卷，第 3 期，第 1808–1822 页，2023年3月。[22] D. Alimo、M. Hamamura 和 S. R. Sabuj，《采用速率分裂与正交预编码提升波束空间大规模多输入多输出系统的覆盖范围、吞吐量和能效》，IEEE开放获取期刊，第 12 卷，第 91 957–91 973 页，2024年。[23] A. M. Sayeed，《多天线衰落信道的解构》，IEEE 信号处理汇刊，第 50 卷，第 10 期，第 2563–2579 页，2002年10月。[24] W. U. Bajwa、J. Haupt、A. M. Sayeed 和 R. Nowak，《压缩信道感知：一种估计稀疏多径信道的新方法》，IEEE 会刊，第 98 卷，第 6 期，第 1058–1076 页，2010年。[25] Y. Zhai、Z. Yang、Z. Liao、J. Wright 和 Y. Ma，《通过在正交群上最大化 l4 范数进行完备字典学习》，《机器学习研究杂志》，第 21 卷，第 165 期，第 1–68 页，2020年8月。[26] B. Olshausen 和 D. Field，《采用超完备基组的稀疏编码：初级视觉皮层的一种策略 3325，1997 37，23，3311？》《视觉研究》，卷、期、页码：–。[27] M. Aharon、M. Elad 和 A. Bruckstein，《K-SVD：一种用于设计稀疏表示超完备字典的算法》，IEEE 信号处理汇刊，第 54 卷，第 11 期，第 4311–4322 页，2006年10月。

[28] Y. Ding 和 B. D. Rao，《基于字典学习的FDD大规模多输入多输出系统稀疏信道表示与估计》，IEEE《无线通信汇刊》，第 17 卷，第 8 期，第 5437–5451 页，2018年8月。[29] Y. Huang、Y. Liang 和 F. Gao，《基于块结构字典学习的FDD大规模多输入多输出系统信道估计》，发表于 IEEE 全球通信会议（GLOBECOM）论文集，2019年12月，第 1–6 页。[30] H. Xie 和 N. Gonz´alez-Prelcic，《混合频率选择性毫米波多输入多输出系统的信道估计字典学习》，IEEE《无线通信汇刊》，第 19 卷，第 11 期，第 7407–7422 页，2020年11月。[31] J. Yu、X. Jiang、Z. Zhong 和 X. Liu，《FDD多输入多输出系统的二维信道估计：基于字典学习的方法》，发表于信息、通信与信号处理国际会议（ICICSP），2021年9月，第 495–500 页。[32] M. Nazzal、M. Ali Ayg¨ul 和 H. Arslan，《MIMO-OFDM信道估计的多维稀疏性估计与利用》，发表于 IEEE 无线通信与网络会议（WCNC）论文集，2022年4月，第 980–985 页。[33] Z. Zhou、B. Cai、J. Chen 和 Y.-C. Liang，《RIS辅助MISO通信的基于字典学习的信道估计》，IEEE 无线通信快报，第 11 卷，第 10 期，第 2125–2129 页，2022年10月。[34] M. A. Ayg¨ul、M. Nazzal 和 H. Arslan，《面向毫米波大规模多输入多输出系统波束空间信道表示与估计的稀疏化字典学习》，IEEE开放获取期刊，第 11 卷，第 98 436–98 451 页，2023年9月。[35] H. Xie、J. Palacios 和 N. Gonz´alez-Prelcic，《存在硬件损伤和波束偏斜时的混合毫米波多输入多输出系统：信道模型与字典学习辅助配置》，IEEE《无线通信汇刊》，第 22 卷，第 10 期，第 6898–6913 页，2023年10月。[36] P. Maity、S. Khatri、S. Srivastava 和 A. K. Jagannatham，《硬件损伤下IRS辅助太赫兹MU-SIMO系统中基于字典学习（DL）的稀疏级联信道估计》，发表于 IEEE 统计信号处理研讨会（SSP）论文集，2023年7月，第 606–610 页。[37] M. Bayraktar、N. Gonz´alez-Prelcic、G. C. Alexandropoulos 和 H. Chen，《硬件损伤下毫米波RIS辅助的联合信道估计与定位：一种基于字典学习的方法》，IEEE《无线通信汇刊》，第 23 卷，第 12 期，第 19 696–19 712 页，2024年12月。[38] L. Lu、G. Y. Li、D. Qiao 和 W. Han，《用于大规模多输入多输出系统压缩感知信道反馈的增强稀疏性的基》，发表于 IEEE 全球通信会议（GLOBECOM）论文集，2015年12月，第 1–6 页。[39] P. K. Gadamsetty、K. V. S. Hari 和 L. Hanzo，《为FDD大规模MU-MIMO-OFDM系统CSI反馈学习公共字典》，IEEE 车辆技术开放期刊，第 4 卷，第 530–544 页，2023年8月。[40] Y. Bai、W. Chen、Y. Bai 和 B. Ai，《用于大规模MIMO大规模机器类通信的基于字典学习的信道估计与活动检测》，发表于 IEEE 国际通信会议（ICC）论文集，2022年5月，第 2272–2277 页。[41] S. Liang、W. Chen、N. Wang 和 B. Ai，《用于大规模MIMO大规模机器类通信的基于簇特定字典学习的活跃用户检测》，发表于 IEEE 全球通信会议（GLOBECOM）论文集，2023年12月，第 5823–5828 页。[42] K. Engan、S. Aase 和 J. Hakon Husoy，《框架设计的最优方向法》，发表于 IEEE 声学、语音与信号处理国际会议（ICASSP）论文集，第 5 卷，1999年3月，第 2443–2446 页。[43] Y. C. Pati、R. Rezaiifar 和 P. S. Krishnaprasad，《正交匹配追踪：递归函数近似及其在小波分解中的应用》，发表于 Asilomar 信号、系统与计算会议，1993年11月，第 40–44 页，第 1 卷。[44] S. S. Chen、D. L. Donoho 和 M. A. Saunders，《基追踪原子分解》，《SIAM科学计算杂志》，第 20 卷，第 1 期，第 33–61 页，1998年。[45] Y. Xue、Y. Shen、V. Lau、J. Zhang 和 K. B. Letaief，《通过在 Stiefel流形上最大化 ℓ_3范数实现大规模MIMO盲数据检测》，IEEE《无线通信汇刊》，第 1–1 页，2020年11月。[46] F. Liu、Y. Zhang、Y. Xiong、S. Li、W. Yuan、F. Gao、S. Jin 和 G. Caire，《在随机ISAC信令下，OFDM实现最低测距旁瓣》，10月。2024 [ ] 2407 06691。在线。可访问：https://arxiv.org/abs/。K. Kreutz-Delgado，《复梯度算子与CR微积分》，[47] 6月。2009 [ ] 0906 4835 1。在线。可[48] 访问：https://arxiv.org/abs/。v D. Tse 和 P. Viswanath，《无线通信基础》。美国：剑桥大学出版社，。[49] P.-A. Absil、R. Mahony 和 R. Sepulchre，2005《矩阵流形上的优化算法》。普林斯顿大学出版社，2009年。[50] A. Edelman、T. A. Arias 和 S. T. Smith，《带正交约束的算法几何》，《SIAM矩阵分析与应用杂志》，第 20 卷，第 2 期，第 303–353 页，1998年。[51] C. Studer、P. Luethi 和 W. Fichtner，《用于MIMO系统数据压缩导向矩阵反馈的超大规模集成电路架构》，发表于 IEEE 电路与系统国际研讨会（ISCAS）论文集，2008年5月，第 300–303 页。[52] B. Hassibi 和 B. Hochwald，《Cayley微分酉空时码》，IEEE 信息论汇刊，第 48 卷，第 6 期，第 1485–1503 页，2002年6月。[53] G. Strang，《微积分》，开放教材图书馆丛书。韦尔斯利—剑桥出版社，1991年，第 1 卷。[54] M. Gauger、M. Arnold 和 S. ten Brink，《住宅区中的大规模MIMO信道测量与可达速率》，发表于智能天线国际 ITG 研讨会（WSA）论文集，2020年2月，第 1–6 页。[55] S. Jaeckel、L. Raschkowski、K. B¨orner、L. Thiele、F. Burkhardt 和 E. Eberlein，《QuaDRiGa——准确定性无线信道生成器用户手册与文档》，弗劳恩霍夫海因里希·赫兹研究所，技术报告 v2.0.0，2017年。[56] A. A. Farid、A. S. H. Ahmed、A. Dhananjay 和 M. J. W. Rodwell，《用于无线通信的全封装135 GHz多用户MIMO发射机阵列模块》，IEEE 微波理论与技术汇刊，第 70 卷，第 7 期，第 3396–3405 页，2022年7月。[57] J. M. Steele，《柯西–施瓦茨大师课：数学不等式艺术导论》。剑桥大学出版社，2004年。[58] M. Journ´ee、Y. Nesterov、P. Richt´arik 和 R. Sepulchre，《稀疏主成分分析的广义幂法》，《机器学习研究杂志》，第 11 卷，第 517–553 页，2010年3月。[59] R.-C. Li，《酉极分解因子的新扰动界》，《SIAM矩阵分析与应用杂志》，第 16 卷，第 1 期，第 327–332 页，1995年1月。 [在线]。可访问：https://doi.org/ 10 1137 0895479893256359。/S J. Demmel 和 K. Veselic，《Jacobi方法比QR分解更精确》，《SIAM矩阵分析与应用杂志》，第 13 卷，[60] 第 4 期，第 1204–1245 页，1992年10月。 [在线]。可访问：https://doi.org/ 10 1137 0613074。/ R. Klen、[61] M. Visuri 和 M. Vuorinen，《双曲函数的Jordan型不等式》，《应用不等式杂志》，第 2010 卷，第 1–14 页，2010年。