



XXXX

面向近场XL-MIMO通信的基于双分辨率表示的速率导向混合波束成形算法

高梅学健, 王安定

(浙江工商大学信息与电子工程学院, 浙江 杭州 310018)

摘 要: 针对近场超大规模多输入多输出 (extremely large-scale multiple-input multiple-output, XL-MIMO) 系统中球面波传播和空间非平稳性导致单原子混合波束成形表示能力受限的问题, 提出基于双分辨率表示的速率导向混合波束成形算法 (dual-resolution rate-oriented hybrid beamforming, DR^2 -HBF)。算法利用频谱效率梯度构造局部目标, 以双原子和受限复系数联合表征波束, 并通过两级前 K 名筛选、硬件投影和真实速率验证完成更新。仿真表明, 该算法在相同恒模约束下可提升频谱效率, 并能在不完美 CSI 下改善统计鲁棒性; 幅相可调架构的结果表明, 性能增益需结合硬件成本评价。

关键词: 近场通信; 超大规模多输入多输出; 混合波束成形; 双分辨率表示; 速率导向优化

中图分类号: TN929.5

文献标志码: A

doi: 10.11959/j.issn.1000-0801.

Rate-oriented hybrid beamforming algorithm based on dual-resolution representation for near-field XL-MIMO communications

Gao Meixuejian, Wang Anding

School of Information and Electronic Engineering, Zhejiang Gongshang University, Hangzhou 310018, China

Abstract: Conventional single-atom hybrid beamforming has limited representation capability in near-field extremely large-scale multiple-input multiple-output (XL-MIMO) systems because of spherical-wave propagation and spatial non-stationarity. A dual-resolution rate-oriented hybrid beamforming (DR^2 -HBF) algorithm is proposed. The algorithm constructs local objectives from the spectral-efficiency gradient, jointly represents each beam using two atoms and a constrained complex coefficient, and performs each update through two-stage Top-K candidate screening, hardware projection, and validation based on the original rate objective. Simulation results show that the proposed algorithm improves spectral efficiency under identical constant-modulus hardware constraints and enhances statistical ro-

收稿日期: 2026-08-17; 修回日期: 2026-09-08

通信作者: 王安定, anding_704@zjgsu.edu.cn

基金项目: 浙江省自然科学基金资助项目 (No. LY22F010013)

Foundation Items: The Natural Science Foundation of Zhejiang Province(No. LY22F010013)...



bustness under imperfect channel state information. Results obtained with amplitude-and-phase-adjustable architectures further indicate that performance gains should be evaluated together with hardware cost.

Key words: near-field communications, extremely large-scale multiple-input multiple-output, hybrid beamforming, dual-resolution representation, rate-oriented optimization

1 引言

随着载波频率提高和阵列孔径增大, 超大规模多输入多输出 (extremely large-scale multiple-input multiple-output, XL-MIMO) 链路在典型服务距离内可能进入菲涅耳 (Fresnel) 区域^[1-2], 传统远场平面波模型会产生角度—距离失配^[3-4], 因而近场信道需要同时描述球面波传播、角度—距离耦合^[5-7]及阵列空间非平稳性^[7-8]。这些传播特征一方面提供角度与距离联合聚焦自由度, 另一方面使波束字典、局部可见区域和硬件可实现约束相互耦合, 导致仅依赖单一离散波束原子的混合预编码难以同时兼顾表示精度、频谱效率与实现成本。

混合预编码通过少量射频 (radio frequency, RF) 链路与模拟网络共同完成波束成形, 是降低全数字 XL-MIMO 硬件成本和功耗的重要途径^[9-10]。现有架构在连接方式上可分为全连接与部分连接, 在模拟器件约束上又可分为恒模 (constant-modulus, CM) 相移器网络和幅相可调 (amplitude-and-phase adjustable, AP) 复权网络。CM 网络器件简单、功耗较低, 是 OMP 及多数经典混合预编码方法的常用假设^[11-13]; AP 网络可由矢量调制器或可变增益移相单元实现^[9-10], 能够提供更高的复权自由度, 但其器件数量、控制精度和功耗均更高。采用双移相器合成复权还可扩展模拟权值自由度, 但会增加器件数量^[12,14]。动态子阵、固定子阵及模块化阵列进一步在频谱效率、开关网络和超大规模可扩展性之间形成不同折中^[15-17]。因此, 算法性能必须在相同模拟硬件约束内比较, 跨架构结果只能用于说明性能—

器件—功耗权衡。

针对近场 XL-MIMO 混合波束成形, 已有研究基于极坐标域字典提出直接贪婪混合波束成形 (direct greedy hybrid beamforming, DG-HBF)、逐步独立贪婪混合波束成形 (successive independent greedy hybrid beamforming, SIG-HBF) 和逐步联合贪婪混合波束成形 (successive joint greedy hybrid beamforming, SJG-HBF) 等方法^[18-20], 动态混合架构与模块化阵列设计又分别扩展了 RF 连接和块间传播模型^[19-20]。本文关注的未解决问题是: 如何在明确的 AP 或 CM 硬件可行域内, 将离散原子对选择、连续复系数调节与原始 log-det 频谱效率统一起来, 并以有限候选规模完成可验证的局部更新; 同时, 需要明确 Top- K 截断只能获得受限候选集内的局部解, 不能据此宣称全局最优。

为解决上述问题, 本文提出基于双分辨率表示的速率导向混合波束成形算法 (dual-resolution rate-oriented hybrid beamforming, DR²-HBF)。这里的“双分辨率”不是两套物理阵列分辨率, 而是离散角度—距离原子索引与连续复系数共同形成的离散—连续联合表示。本文贡献概括如下: 1) 从 log-det 频谱效率的维廷格梯度出发, 构造 RF 链路—天线块的局部秩一目标, 并利用低维恒等式和线性方程求解避免显式形成大规模矩阵逆; 2) 在已有多原子线性表示思想上, 给出带 AP/CM 硬件投影的双原子天线块可行集, 明确单原子与双原子模型的包含关系、复系数物理含义及器件规模边界; 3) 采用两级 Top- K 离散筛选与受限连续系数估计控制候选数, 并分别给出完美 CSI 速率接受、非理想 CSI 鲁棒接受以及局部最

优诊断机制。本文只证明所选接受指标的单调有界性，不声称预编码变量收敛到全局最优解。

2 系统模型与问题建模

2.1 系统架构与硬件约束

考虑图1所示近场XL-MIMO下行系统^[1-2]。基站配备 N_t 根天线、 M_t 条RF链路和AP全连接模拟网络；用户端配备 N_r 根天线、 M_r 条RF链路。发射端采用均匀线阵（uniform linear array, ULA），阵元间距 $d=\lambda/2$ ， λ 为载波波长。阵列按索引分为 P 个连续等长天线块，每块 $N_{sub}=N_t/P$ 根天线；数据流数满足 $N_s \leq \min(M_t, M_r)$ 。

AP全连接网络含 $N_t M_t$ 个RF一天线支路，每支路通过矢量调制器或“可变增益器+移相器”输出一个复权^[9-10]。固定天线块仅划分 F_{RF} 的行索引以建模和更新，不切断物理连接。双原子在算法域合成一个目标复权后由同一支路实现；若采用双移相器及合路器，相移器数约为 $2N_t M_t$ ^[12,14]，其器件与功耗另行核算。器件功耗采用统一模型^[9-10]：

$$P_{\text{tot}} = P_{\text{BB}} + M_t P_{\text{RF}} + N_{\text{VM}} P_{\text{VM}} + N_{\text{PS}} P_{\text{PS}} + N_{\text{VGA}} P_{\text{VGA}} + N_{\text{C}} P_{\text{comb}} + N_{\text{SW}} P_{\text{SW}} + N_{\text{PA}} P_{\text{PA}} \quad (1)$$

其中， N_{VGA} 、 P_{VGA} 分别为可变增益放大器数量

和单位功耗， N_{PA} 、 P_{PA} 分别为功放数量和单位功耗。第4.7节表6按所列模块组合统一计数；其中AP分别计入VM和VGA项，具体集成实现需按实际配置核算。

AP、CM全连接分别取 $N_{\text{VM}}=N_t M_t$ 和 $N_{\text{PS}}=N_t M_t$ ；双移相器实现取 $N_{\text{PS}}=2N_t M_t$ 、 $N_{\text{C}}=N_t M_t$ ，等分部分连接通常仅需 N_t 个模拟支路。第4.7节报告器件数和归一化能效敏感性，不将归一化参数解释为实测绝对功耗。

模拟预编码矩阵 $F_{\text{RF}} \in \mathbb{C}^{N_t \times M_t}$ 与数字预编码矩阵 $F_{\text{BB}} \in \mathbb{C}^{M_t \times N_s}$ 构成 $F = F_{\text{RF}} F_{\text{BB}}$ ，接收信号 $y \in \mathbb{C}^{N_r \times 1}$ 为

$$y = HFs + n \quad (2)$$

其中， $H \in \mathbb{C}^{N_r \times N_t}$ 为近场信道，发送数据为 $s \in \mathbb{C}^{N_s \times 1}$ 满足 $\mathbb{E}[ss^H] = I_{N_s}$ ， $n \in \mathbb{C}^{N_r \times 1}$ 为加性高斯白噪声。接收合并等效并入信道或按理想合并处理。

2.2 信道模型

XL-MIMO信道可通过球面波阵列流型构建。与远场模型不同，近场路径的阵列响应由角度和距离共同决定^[2,7]。对于第 l 条路径，设其发射端角度和距离参数分别为 θ_l 和 r_l ，则发射端近场阵列流型可表示为

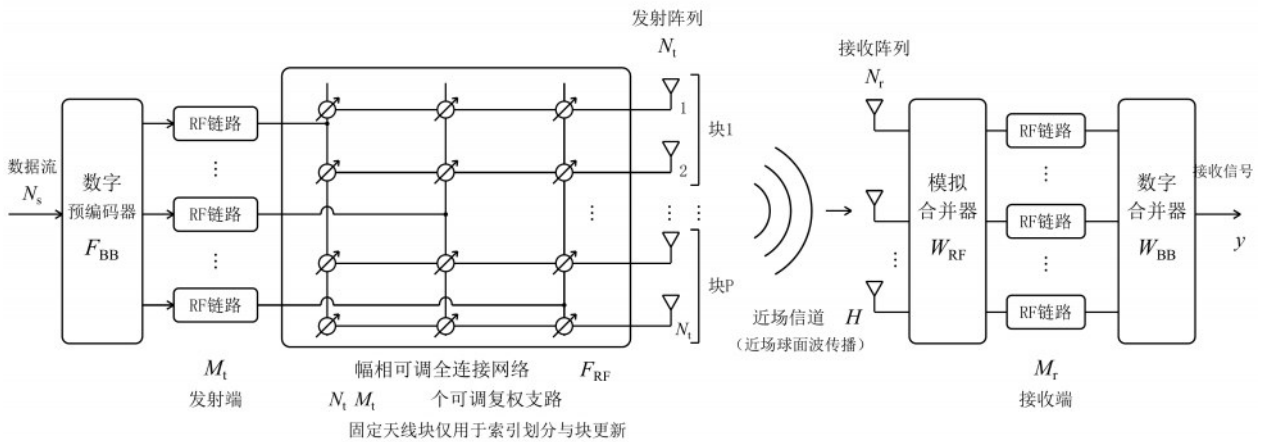


图1 近场XL-MIMO系统与幅相可调全连接结构



$$\mathbf{a}_l(\theta_l, r_l) = \frac{1}{\sqrt{N_l}} \left[e^{-j\frac{2\pi}{\lambda}(r_l^{(1)} - r_l)}, e^{-j\frac{2\pi}{\lambda}(r_l^{(2)} - r_l)}, \dots, e^{-j\frac{2\pi}{\lambda}(r_l^{(N_l)} - r_l)} \right]^T \quad \#(3)$$

其中, $r_l^{(n)}$ 表示第 n 个发射阵元到第 l 条路径对应散射体或用户的实际传播距离。接收端近场阵列流型可按相同方式定义, 其角度和距离参数分别为 ϕ_l 和 r_l' 。

在包含 L 条传播路径的几何簇状散射模型下, 将传统远场阵列响应替换为角度-距离联合的近场球面波阵列流型^[18,20], 近场 MIMO 信道可写为

$$\mathbf{H} = \sqrt{\frac{N_t N_r}{L}} \sum_{l=1}^L \beta_l \mathbf{a}_r(\phi_l, r_l') \mathbf{a}_t^H(\theta_l, r_l) \quad \#(4)$$

式中, β_l 为第 l 条路径的复增益, $\mathbf{a}_t(\cdot) \in \mathbb{C}^{N_t \times 1}$ 和 $\mathbf{a}_r(\cdot) \in \mathbb{C}^{N_r \times 1}$ 分别表示发射端和接收端的近场阵列流型。对于空间非平稳场景, 可进一步在发射端阵列流型中引入可见区域向量 $\mathbf{v}_l \in \mathbb{C}^{N_t \times 1}$, 用 $\mathbf{v}_l \odot \mathbf{a}_t(\theta_l, r_l)$ 描述第 l 条路径仅在部分子阵中可见的情况^[8,21], 其中 $\odot$ 表示哈达玛积 (Hadamard product)。

以首个阵元为参考点, 令 $\delta_n = n - 1$, r 和 θ 分别为相对该点的距离与角度, θ 从阵列法线计量。第 n 个阵元的精确传播距离为

$$r^{(n)} = \sqrt{r^2 + \delta_n^2 d^2 - 2r\delta_n d \sin \theta}$$

当阵元偏移相对于传播距离较小时, 对根式按 $\frac{\delta_n d}{r}$ 作二阶展开。采用文献^[6]的 Fresnel 近似, 并将其空间方向参数对应为 $\sin \theta$, 得到

$$r^{(n)} \approx r - \delta_n d \sin \theta + \frac{\delta_n^2 d^2 \cos^2 \theta}{2r} \quad \#(5)$$

式 (5) 适用于阵元偏移远小于传播距离且高阶相位误差可忽略的情况; 一次项对应平面波, 二次项描述波前弯曲。

对第 2.1 节的固定天线块, 块内采用平面波近似、块间保留球面波相位差^[19-20]; 在各块局部传播方向近似一致、共用块内响应时, 可进一步

写为

$$\mathbf{a}_l(\theta, r) \approx \mathbf{b}_l(\theta, r) \otimes \mathbf{u}_l(\theta) \quad \#(6)$$

其中, $\mathbf{b}_l(\theta, r)$ 和 $\mathbf{u}_l(\theta)$ 分别表示块间球面波与块内平面波响应, $\otimes$ 为克罗内克积 (Kronecker product)。

令 $D_{\text{sub}} = (N_{\text{sub}} - 1)d$ 为单个天线块孔径。当

$$\frac{D_{\text{sub}}^2}{\lambda r} \ll 1, \text{ 即 } r \gg \frac{2D_{\text{sub}}^2}{\lambda} \text{ 时, 块内二次相位项可忽略。}$$

默认参数下 $D_{\text{sub}} \approx 0.080\text{m}$ 、 $\frac{2D_{\text{sub}}^2}{\lambda} \approx 1.21\text{m}$,

字典距离为 $5 \sim 100\text{m}$, 满足该块内尺度条件。式 (6) 的共同块内响应还要求块间方向差足够小; 实际字典由式 (3) 的球面波流型截取并归一化, 不依赖这一共同方向近似。

2.3 频谱效率优化问题

在独立复高斯发送符号、等功率分配及等效

接收合并下, 频谱效率为^[9,11] $R = \log_2 \det \left(\mathbf{I}_{N_r} + \right.$

$$\left. \frac{\rho}{N_s} \mathbf{H} \mathbf{F} \mathbf{F}^H \mathbf{H}^H \right)$$
。利用 $\det(\mathbf{I}_{N_r} + \mathbf{A} \mathbf{B}) = \det(\mathbf{I}_{N_s} + \mathbf{B} \mathbf{A})$,

取 $\mathbf{A} = \sqrt{\eta} \mathbf{H} \mathbf{F}$, $\mathbf{B} = \sqrt{\eta} \mathbf{F}^H \mathbf{H}^H$, $\eta = \rho / N_s$, 得到等价低维形式:

$$R = \log_2 \det \left(\mathbf{I}_{N_s} + \frac{\rho}{N_s} (\mathbf{H} \mathbf{F})^H (\mathbf{H} \mathbf{F}) \right) \quad \#(7)$$

其中, ρ 为线性信噪比。全数字预编码逼近误差仅间接反映 log-det 速率^[11-12], 故本文直接优化式 (7)。

联合设计 $\mathbf{F}_{\text{RF}}$ 与 $\mathbf{F}_{\text{BB}}$, 得到

$$\max_{\mathbf{F}_{\text{RF}}, \mathbf{F}_{\text{BB}}} R(\mathbf{F}_{\text{RF}}, \mathbf{F}_{\text{BB}}), \quad \text{s.t. } \|\mathbf{F}_{\text{RF}} \mathbf{F}_{\text{BB}}\|_F^2 =$$

$$N_s, \quad \mathbf{f}_m^p \in D_{\text{p,H}}^{(2)} \quad \#(8)$$

其中, $\mathbf{f}_m^p$ 为第 m 条 RF 链路在第 p 个天线块上的模拟向量, $D_{\text{p,H}}^{(2)}$ 为第 3.2.1 节定义的经硬件投影后的双原子集合。

合成复权需经逐元素硬件投影 $\Pi_{\mathcal{H}}$: AP 幅度截断至 $[a_{\min}, a_{\max}]$, CM 幅度固定为 a_0 , 二者均以 $B\varphi$ 比特量化相位:

$$\Pi_{\text{AP}}(z) =$$

$$\text{clip}(|z|; a_{\min}, a_{\max}) \exp(jQ_{B\varphi}(\angle z)), \Pi_{\text{CM}}(z) = a_0 \exp(jQ_{B\varphi}(\angle z)) \# (9)$$

式 (8) 的块可行集取双原子组合经 $\Pi_{\mathcal{H}}$ 投影后的集合。连续相位取 $B\varphi \rightarrow \infty$; 归一化幅度范围 $a_{\min}$ 、 $a_{\max}$ 与同硬件基线共享。

式 (8) 是离散索引、连续系数与硬件投影耦合的非凸问题。下文采用梯度引导的顺序块搜索, 并以原始速率验证候选。

3 基于双分辨率表示的速率导向混合波束成形算法

3.1 速率导向的局部更新目标构造

固定信道 $\mathbf{H}$, 将整体预编码视为复矩阵变量。令 $\mathbf{F}_{\text{cur}} = \mathbf{F}_{\text{RF}} \mathbf{F}_{\text{BB}}$, $\eta = \rho/N_s$, $\mathbf{C} = (\mathbf{I}_{N_r} + \eta \mathbf{H} \mathbf{F}_{\text{cur}} \mathbf{F}_{\text{cur}}^H \mathbf{H}^H)^{-1}$ 。由 $d \ln \det X = \text{tr}(X^{-1} dX)$ 及 $d\mathbf{C}^{-1} = \eta \mathbf{H} (d\mathbf{F}_{\text{cur}} \mathbf{F}_{\text{cur}}^H + \mathbf{F}_{\text{cur}} d\mathbf{F}_{\text{cur}}^H) \mathbf{H}^H$, 令 $\mathbf{A} = \mathbf{H}^H \mathbf{C} \mathbf{H}$, 整理得到

$$dR = \frac{\eta}{\ln 2} \left[\text{tr} \left(\mathbf{F}_{\text{cur}}^H \mathbf{A} d\mathbf{F}_{\text{cur}} \right) + \text{tr} \left((\mathbf{A} \mathbf{F}_{\text{cur}})^T d\mathbf{F}_{\text{cur}}^* \right) \right]$$

$$\text{按 } dR = \text{tr} \left(\left(\frac{\partial R}{\partial \mathbf{F}_{\text{cur}}} \right)^T d\mathbf{F}_{\text{cur}} \right) + \text{tr} \left(\left(\frac{\partial R}{\partial \mathbf{F}_{\text{cur}}^*} \right)^T d\mathbf{F}_{\text{cur}}^* \right)$$

的维廷格微分约定, 比较 $d\mathbf{F}_{\text{cur}}^*$ 对应迹项的系数, 并代回 $\mathbf{A} = \mathbf{H}^H \mathbf{C} \mathbf{H}$, 得到共轭梯度

$$\frac{\partial R}{\partial \mathbf{F}_{\text{cur}}^*} = \frac{\eta}{\ln 2} \mathbf{H}^H \mathbf{C} \mathbf{H} \mathbf{F}_{\text{cur}} \# (10)$$

当 $\rho > 0$ 时, 省略正常数 $\eta/\ln 2$, 定义梯度代理 $\mathbf{G} = \mathbf{H}^H \mathbf{C} \mathbf{H} \mathbf{F}_{\text{cur}}$; $\rho = 0$ 时速率恒为 0。令 $\mathbf{Y} = \mathbf{H} \mathbf{F}_{\text{cur}}$, 由 $(\mathbf{I} + \eta \mathbf{Y} \mathbf{Y}^H) \mathbf{Y} = \mathbf{Y}(\mathbf{I} + \eta \mathbf{Y}^H \mathbf{Y})$ 且括号内矩阵均为厄米正定矩阵, 在两侧乘相应逆矩阵, 可得 $(\mathbf{I} + \eta \mathbf{Y} \mathbf{Y}^H)^{-1} \mathbf{Y} = \mathbf{Y}(\mathbf{I} + \eta \mathbf{Y}^H \mathbf{Y})^{-1}$, 故有

$$\mathbf{G} = \mathbf{H}^H \mathbf{Y} (\mathbf{I}_{N_s} + \eta \mathbf{Y}^H \mathbf{Y})^{-1}, \mathbf{Y} = \mathbf{H} \mathbf{F}_{\text{cur}} \# (11)$$

式 (11) 通过 N_s 维厄米正定线性方程求解, 采用 Cholesky 分解并在数值需要时加入对角加载; 无需显式形成 N_r 维矩阵逆。SVD 仅用于 $N_r \times M_t$ 等效信道的数字更新。

沿用固定分块, 令 $\mathcal{S}_p = \{(p-1)N_{\text{sub}} + 1, (p-1)N_{\text{sub}} + 2, \dots, pN_{\text{sub}}\}$, $p = 1, 2, \dots, P$, 模拟列按 $\mathbf{f}_m = [\mathbf{f}_m^{1T}, \mathbf{f}_m^{2T}, \dots, \mathbf{f}_m^{PT}]^T$, $\mathbf{f}_m^p \in \mathbb{C}^{N_{\text{sub}} \times 1}$ 分块, 局部向量是基本更新单元。

记 $\text{idx} = \mathcal{S}_p$, $\mathbf{G}_p = \mathbf{G}(\text{idx}, :)$ 及 $\mathbf{w} = \mathbf{F}_{\text{BB}}^H(:, m) = \mathbf{F}_{\text{BB}}(m, :)^H \in \mathbb{C}^{N_s \times 1}$ 。固定当前数字方向后, 局部模拟块对整体预编码的贡献为 $\mathbf{f}_m^p \mathbf{w}^H$, 故构造最小二乘目标

$$\min_g (\mathbf{G}_p - \mathbf{g} \mathbf{w}^H)^2_{\text{F}} \# (12)$$

固定梯度代理和数字方向, 对式 (12) 的目标关于 $\mathbf{g}^*$ 求导并令其为零, 得到正规方程 $(\mathbf{w}^H \mathbf{w}) \mathbf{g} - \mathbf{G}_p \mathbf{w} = 0$ 。当 $\mathbf{w}^H \mathbf{w} > 0$ 时, 目标为严格凸二次函数, 其唯一解为

$$\mathbf{g} = \frac{\mathbf{G}(\text{idx}, :)\mathbf{w}}{\mathbf{w}^H \mathbf{w}} \# (13)$$

$\mathbf{g} \in \mathbb{C}^{N_{\text{sub}} \times 1}$ 是固定 $\mathbf{w}$ 和 $\mathbf{G}$ 时的局部秩一投影; 若数字方向为零, 则跳过该块。数字更新后重算梯度和局部目标, 候选对原始速率的贡献仍须单独验证。

3.2 模拟预编码设计

将局部目标 $\mathbf{g}$ 投影到双原子近场字典, 并经硬件投影与速率接受形成模拟更新。

3.2.1 双原子天线块波束表示

依据第 2.2 节的角距联合模型^[5,18], 第 p 个天线块的 Q 个采样原子构成字典

$$\mathbf{A}_p = [\boldsymbol{\psi}_{p,1}, \boldsymbol{\psi}_{p,2}, \dots, \boldsymbol{\psi}_{p,Q}] \in \mathbb{C}^{N_{\text{sub}} \times Q} \# (14)$$

其中, $\boldsymbol{\psi}_{p,q}$ 由全阵流型 $\mathbf{a}_t(\theta_q, r_q)$ 在 $\mathcal{S}_p$ 上截取并归一化, 每列对应一个角距聚焦方向。



多原子线性组合属于已有稀疏表示思想^[11]。本文将其与速率梯度、受限系数及硬件可行域联

$$\mathcal{D}_p^{(2)} = \{\psi_{p,q_1} + \alpha\psi_{p,q_2} | q_1, q_2 \in \{1, \dots, Q\}, q_1 \neq q_2, \alpha \in \mathbb{C}, |\alpha| \leq \alpha_{\max}\} \# (15)$$

硬件投影前的局部合成候选为

$$f_m^p = \psi_{p,q_1} + \alpha_m^p \psi_{p,q_2} \# (16)$$

离散原子 ψ_{p,q_1} 、 ψ_{p,q_2} 选定角距方向， $\alpha_m^p = |\alpha_m^p| \exp(j(\angle \alpha_m^p))$ 连续调节辅助方向的相对幅相，以逼近两个局部传播分量或修正波前。合成后经 $\Pi_{\mathcal{H}}$ 投影，每个最终复权由一条既有支路实现，器件配置见第2.1节。

当 $\alpha_m^p = 0$ 时退化为单原子。在相同 $\Pi_{\mathcal{H}}$ 下有式 (17) 的包含关系，但硬件投影、Top-K 截断与速率接受使其不能直接推出真实速率严格提高。

$$\mathcal{D}_{p,\mathcal{H}}^{(1)} = \{\Pi_{\mathcal{H}}(\psi_{p,q})\} \subseteq \mathcal{D}_{p,\mathcal{H}}^{(2)} = \{\Pi_{\mathcal{H}}(\psi_{p,q_1} + \alpha\psi_{p,q_2})\} \# (17)$$

3.2.2 候选筛选与受限复系数求解

给定 $\mathbf{g}$ ，在 $q_1, q_2 \in \{1, 2, \dots, Q\}$ 且 $q_1 \neq q_2$ 下求解

$$\min_{q_1 \neq q_2, |\alpha_m^p| \leq \alpha_{\max}} \left\| \mathbf{g} - (\psi_{p,q_1} + \alpha_m^p \psi_{p,q_2}) \right\|_2^2 \# (18)$$

完整有序原子对有 $Q(Q-1)$ 组。两级 Top-K 先按相关性保留 K_1 个主原子，再按各主原子残差保留 K_2 个辅助原子，使高成本候选验证不超过 $K_1 K_2$ 次。截断可能遗漏较优原子，第4.5节以缩小字典局部枚举和初始化实验量化该风险。

对固定的非零原子对 $(\psi_{p,q_1}, \psi_{p,q_2})$ ，式 (18) 转为复系数的条件最小二乘问题。关于系数的共轭求导并令其为零，得到

$$-\psi_{p,q_2}^H (\mathbf{g} - \psi_{p,q_1} - \alpha_m^p \psi_{p,q_2}) = 0.$$

辅助原子非零使分母为正，因而无约束最小二乘解为

$$\alpha_m^p = \frac{\psi_{p,q_2}^H (\mathbf{g} - \psi_{p,q_1})}{\psi_{p,q_2}^H \psi_{p,q_2}} \# (19)$$

$\alpha_{\max} \geq 0$ 限制辅助原子的相对权重，不是器件

合，双原子集合定义为

功率上限。令 $\mathbf{z}$ 为待优化复系数，将目标配方后，与 $\mathbf{z}$ 有关的项为 $\|\psi_{p,q_2}\|_2^2 |\mathbf{z} - \alpha_m^p|^2$ ；其余项与 $\mathbf{z}$ 无关。因此，当式 (19) 的解在圆盘内时保持不变，否则径向投影至圆盘边界，得到

$$\alpha_c^* = \begin{cases} \alpha_m^p, & |\alpha_m^p| \leq \alpha_{\max} \\ \alpha_{\max} e^{j \arg(\alpha_m^p)}, & |\alpha_m^p| > \alpha_{\max} \end{cases} \# (20)$$

命题1 (双原子局部代理误差非增性) 对任意固定主原子 ψ_{p,q_1} 和任意非零辅助原子 ψ_{p,q_2} ，令 $\mathbf{r}_1 = \mathbf{g} - \psi_{p,q_1}$ ， $E_1 = \|\mathbf{r}_1\|_2^2$ ；在包含 $\alpha=0$ 的系数圆盘内采用式 (20) 的最优系数时，双原子最小代理误差 E_2 满足

$$E_2 = \min_{|\alpha| \leq \alpha_{\max}} \|\mathbf{r}_1 - \alpha \psi_{p,q_2}\|_2^2 \leq \|\mathbf{r}_1\|_2^2 = E_1 \# (21)$$

证明： $\alpha=0$ 属于系数可行域，故最小误差不大于单原子误差。无约束解在圆盘内且 $\psi_{p,q_2}^H \mathbf{r}_1 \neq 0$

时，误差减少 $\frac{|\psi_{p,q_2}^H \mathbf{r}_1|^2}{\|\psi_{p,q_2}\|_2^2}$ 。该结论限于硬件投影前

的局部代理；原始速率改善由后续验证决定。

3.2.3 模拟候选速率验证与矩阵组装

将候选块暂写入模拟矩阵，固定数字方向并按总功率重新归一化。以同一信道计算候选速率 R_{can} 与当前速率 R_{cur} ，按下式接受：

$$\mathbf{f}_{mp}^+ = \begin{cases} \mathbf{f}_{mp}^{\text{can}}, & R_{\text{can}} > R_{\text{cur}} \\ \mathbf{f}_{mp}^{\text{cur}}, & R_{\text{can}} \leq R_{\text{cur}} \end{cases} \# (22)$$

式 (22) 只保证固定设计信道上的速率不降。估计信道上的接受不能保证未知真实信道速率单调，不完美 CSI 采用下述样本规则。接受块按天线索引立即写回当前模拟列。

3.2.4 不完美 CSI 下的候选筛选与鲁棒接受

围绕估计信道 $\hat{H}$ 生成 S 个误差信道样本 $H(s)^{[6]}$ ，分别计算局部目标 $g(s)$ 。主原子按式

(23) 排序, 辅助原子以对应残差替代目标。\$S=1\$ 且 \$H(l)=\hat{H}\$ 时退化为确定性筛选。

$$s_q = \frac{1}{S} \sum_{s=1}^S |\psi_{p,q}^H \mathbf{g}^{(s)}|^2 \# (23)$$

固定一次设计的样本集, 令各样本速率为 \$R^{(s)}(F)\$, 以式 (24) 的保守效用 \$J\$ 接受候选。\$\kappa \geq 0\$ 控制标准差惩罚, \$\varepsilon\$ 为非负增益阈值。数字和模拟更新及最终分支选择均比较同一效用; 该规则不保证未知真实信道速率逐步提高, TT/ET/TE/EE 仅作离线诊断。

$$J(F) = \text{mean}_s [R^{(s)}(F)] - \kappa \cdot \text{std}_s [R^{(s)}(F)]$$

$$\text{accept iff } J(F^{\text{can}}) - J(F^{\text{cur}}) > \varepsilon \# (24)$$

3.3 数字预编码设计

固定 \$\mathbf{F}_{\text{RF}}\$, 令 \$\mathbf{H}_e = \mathbf{H}\mathbf{F}_{\text{RF}}\$ 并作奇异值分解 (singular value decomposition, SVD): \$\mathbf{H}_e = \mathbf{U}_e \mathbf{\Sigma}_e \mathbf{V}_e^H\$. 取 \$\mathbf{V}_e\$ 中最大 \$N_s\$ 个奇异值对应的右奇异向量 \$\mathbf{F}_{\text{BB}}^{(0)} = \mathbf{V}_e(:, 1:N_s)\$, 构造归一化候选

$$\mathbf{F}_{\text{BB}}^{\text{can}} = \frac{\sqrt{N_s} \mathbf{F}_{\text{BB}}^{(0)}}{\|\mathbf{F}_{\text{RF}} \mathbf{F}_{\text{BB}}^{(0)}\|_F} \# (25)$$

式 (25) 给出 SVD 数字候选^[12]。完美 CSI 按式 (26) 比较当前模拟矩阵下的候选与保留方向; 不完美 CSI 复用式 (24)。

$$\mathbf{F}_{\text{BB}}^{(t)} \in \underset{\mathbf{Z} \in \{\mathbf{F}_{\text{BB}}^{\text{can}}, \mathbf{F}_{\text{BB}}^{(t-1)}\}}{\text{argmax}} R(\mathbf{F}_{\text{RF}}^{(t-1)}, \mathbf{Z}) \# (26)$$

首轮采用式 (25), 后续及最终数字刷新均执行接受保护。拒绝仅保持数字方向, 模拟块仍继续搜索并按总功率缩放; 候选验证不重复 SVD。

3.4 DR²-HBF 总体流程与更新机制

算法 1 给出统一流程。完美与不完美 CSI 共享数字更新、梯度构造、顺序块搜索、硬件投影和归一化, 区别在于候选评分及接受判据。

3.5 单调性与收敛性分析

命题 2 (固定信道速率单调性) 固定设计信道及总功率约束, 以 \$l\$ 编号数字或模拟块更新。

算法 1 基于双分辨率表示的速率导向混合波束成形算法 (DR²-HBF)

输入 信道 \$H\$ 或样本集 \$\{H(s)\}\$, 局部近场字典 \$A_p\$, 硬件模式 \$\mathcal{H} \in \{AP, CM\}\$, \$K_1/K_2\$, \$I_{\text{max}}\$, \$\varepsilon\$, \$\alpha_{\text{max}}\$ 及鲁棒系数 \$\kappa\$。

初始化 按 \$\mathcal{H}\$ 初始化满足全连接固定分块参数化的 \$\mathbf{F}_{\text{RF}}\$; 初始化 \$\mathbf{F}_{\text{BB}}\$、当前速率或效用 \$J_{\text{best}}\$ 及 \$t=0\$。

步骤 1 \$t \leftarrow t+1\$。

步骤 2 固定 \$\mathbf{F}_{\text{RF}}^{(t-1)}\$, 按式 (25) 生成数字候选; 完美 CSI 按式 (26)、样本模式按式 (24) 接受, 并更新当前速率或效用 \$J_{\text{best}}\$。

步骤 3 令 \$Y = \mathbf{H}\mathbf{F}_{\text{RF}}^{\text{cur}}\$, 按式 (11) 通过 \$N_s\$ 维线性方程计算速率梯度 \$G\$; 样本模式分别计算 \$G(s)\$。

步骤 4 for \$m=1, \dots, M_t\$, \$p=1, \dots, P\$ do

步骤 5 根据式 (12) 建立局部逼近问题, 并按式 (13) 计算当前 RF 链路与天线块的局部更新目标 \$\mathbf{g}^{(l)}\$。

步骤 6 按单信道相关性或式 (23) 的样本平均相关能量选取 \$K_1\$ 个主原子候选。

步骤 7 对每个主原子计算单信道/样本残差, 并选取 \$K_2\$ 个辅助原子候选。

步骤 8 按式 (19) 和式 (20) 求 \$\alpha\$, 由式 (16) 合成双原子向量, 再按式 (9) 执行硬件投影 \$\Pi_{\mathcal{H}}\$。

步骤 9 暂时写入候选块并重新归一化 \$\mathbf{F}_{\text{BB}}\$; 完美 CSI 计算候选速率, 样本模式计算各样本速率及式 (24) 的效用 \$J\$。

步骤 10 完美 CSI 按式 (22) 接受; 不完美 CSI 按式 (24) 接受。拒绝则恢复, 接受则立即更新 \$\mathbf{F}_{\text{RF}}\$、\$\mathbf{F}_{\text{BB}}\$ 与 \$J_{\text{best}}\$。

步骤 11 end for (辅助原子候选循环)

步骤 12 end for (主原子候选循环)

步骤 13 end for (RF 链路与固定天线块循环)

步骤 14 记录本轮接受指标 \$J(t)\$。若相对增益低于 \$\varepsilon\$ 或 \$t=I_{\text{max}}\$ 则停止, 否则返回数字更新。

输出 基于最终 \$\mathbf{F}_{\text{RF}}\$ 生成数字候选并执行相同接受规则; 输出 \$\mathbf{F}_{\text{RF}}\$、\$\mathbf{F}_{\text{BB}}\$、\$F\$ 及候选/接受诊断信息。

若分别按式 (26) 和式 (22) 接受, 则保留解的速率满足

$$R^{(l+1)} \geq R^{(l)} \# (27)$$

证明: 数字更新保留两者中速率较高的方向; 模拟更新只接受严格增益, 否则恢复, 故有式 (27)。令 \$\lambda_i\$ 为 \$(\mathbf{H}\mathbf{F})^H(\mathbf{H}\mathbf{F})\$ 的特征值, 由 \$\|\mathbf{F}\|_F^2 = N_s\$ 及谱范数次乘性, 有 \$0 \leq \lambda_i \leq \|\mathbf{H}\|_2^2 \|\mathbf{F}\|_2^2 \leq N_s \|\mathbf{H}\|_2^2\$。将该界代入式 (7), 得到

$$0 \leq R^{(l)} \leq N_s \log_2 \left(1 + \rho \|\mathbf{H}\|_2^2 \right), \lim_{l \rightarrow \infty} R^{(l)} = R^* \# (28)$$



式 (28) 保证固定设计信道的速率值收敛。固定样本集和非负 κ 时, 式 (24) 使 J 单调, 其上界不超过各样本速率上界的均值, 故效用值亦收敛; 这不保证未知真实信道速率单调。离散索引、Top- K 截断和顺序块搜索使上述结论不能推出预编码变量收敛、一阶驻点或全局最优性。

所有候选被拒绝时, 仅为当前候选集上的块无改进点; 阈值停止仅表示近似停滞。第 4.5 节以 $Gap_K = (R_{full} - R_{TopK})/R_{full}$ 衡量缩小字典的局部枚举差距, 并报告多初始化敏感性。

3.6 计算复杂度分析

设字典规模为 Q 、外循环次数为 I_{out} 。由式 (11), 梯度代价为 $CG = O(N_r N_t N_s + N_r N_s^2 + N_s^3)$; 等效信道及 SVD 代价为 $CSVD = O(N_r N_t M_t + N_r M_t^2) (N_r \geq M_t)$ 。每个块的两级筛选代价为 $O((K_1 + 1)QN_{sub})$, 至多 $K_1 K_2$ 个候选需系数求解、硬件投影及速率验证。缓存 $Z = HF$ 并增量更新时, 单候选代价为 $CR = O(N_r N_{sub} N_s + N_r N_s^2 + N_s^3)$ 。总体为

$$\mathcal{C}_{DR2} = O\left(I_{out}\left[\mathcal{C}_{SVD} + \mathcal{C}_G + M_t P\left((K_1 + 1)QN_{sub} + K_1 K_2 \mathcal{C}_R\right)\right]\right) \quad (29)$$

Top- K 将高成本验证从 $Q(Q-1)$ 降至至多 $K_1 K_2$ 次, 但仍有 $K_1 Q$ 量级的辅助相关性筛选, 不能将总复杂度简写为 $O(K_1 K_2)$ 。大规模时等效信道乘法和字典相关性仍可能成为瓶颈; 实测时间及内存口径见第 4.6 节。

4 仿真结果与性能分析

4.1 仿真设置与比较原则

表 1 按硬件可行域分组: CM 组比较 CM-1Atom 与 CM-DR²-HBF[11, 18], 4 bit QPS-DR²-HBF 检验量化影响; AP 组比较 AP-AltMin 与 AP-DR²-HBF^[14]。FD-WF 仅作全数字上界; DG、SJG 和 SIG^[18-20]保留原生字典及连接结构, 仅作跨架构参考。

仿真采用 28 GHz 非平稳近场 XL-MIMO 信道^[7], 阵列与字典参数见表 2。基站采用半波长间距 ULA, 固定划分为 8 个连续等长天线块, 每块 16 根天线; 每条非视距路径随机遮挡 5 个块。

算法参数见表 2。AP 支路在 $[0, 1]$ 内连续调幅并连续调相; CM 采用恒幅 $1/\sqrt{128}$ 和连续相位; QPS 在 CM 约束上采用 4 bit 相位量化^[14, 17]。各算法均按相同总发射功率归一化。

主 SNR 及消融实验采用 200 个配对信道, 鲁棒性实验采用 50 个配对信道; 每次鲁棒设计使用 8 个训练不确定样本, 以 64 个独立样本验证均值和 5% 分位速率。Top- K 、系数上限及初始化诊断采用 20 组随机初始化, 复杂度实验每个设置重复 3 次。

4.2 不同硬件约束下的 SNR 性能

图 2 给出 -30~20 dB 的平均频谱效率: (a) 为 CM、QPS 及跨架构参考, (b) 为 AP 组, 并以 CM-DR²-HBF 参照硬件自由度。各曲线随 SNR 升高而增加, FD-WF 始终为上界。

表 1 不同硬件约束下的算法比较层级

方法	模拟硬件约束	连接/分块结构	局部表示	比较角色
FD-WF	无模拟网络	全数字	不受模拟约束	性能上界
CM-DG/OMP/OLS	恒模/相位量化	全连接/全阵	单原子	CM 组基线
CM-SJG/SIG	恒模或固定子阵	部分连接/固定子阵	单原子	跨架构参考
AP-single/AltMin	与本文相同幅相范围和量化	全连接/固定分块参数化	单原子或独立 AP 优化	AP 组公平基线
AP-DR ² -HBF	幅相可调	全连接/固定分块参数化	双原子+硬件投影	本文主方法
CM-DR ² -HBF	恒模/相位量化	全连接/固定分块参数化	双原子后恒模投影	受限变体

表2 主要仿真参数与统计设置

类别	参数	取值	说明
阵列与信道	f 、 N_t 、 N_r 、 P	28 GHz、128、8、8	ULA；半波长间距；固定连续分块
传输维度	M_t 、 N_s 、 L	4、4、4	RF链路数、数据流数、路径数
近场字典	$Q\theta$ 、 Qr 、 Q	64、20、1280	倒数距离 $0.01 \sim 0.20 \text{ m}^{-1}$
DR ² -HBF	I_{out} 、 K_1 、 K_2 、 α_{max}	5、48、24、4	阈值 10^{-4}
主实验	SNR、信道数	-30~20 dB、200	间隔5 dB；配对样本
鲁棒性	NMSE、训练/验证样本	0~0.20、8/64	50个信道；5%分位速率
扩展诊断	N_t 、初始化/重复数	128~2048、20/3	Top-K与复杂度分别统计

CM-DR²-HBF 在全部 SNR 采样点均优于 CM-1Atom。0 dB 时两者速率为 15.680 和 13.646 bit/s/Hz，配对增益 2.034 bit/s/Hz，95% 置信区间为 [1.876, 2.191] bit/s/Hz；20 dB 时为 39.432 和 37.307 bit/s/Hz，增益 2.125 bit/s/Hz，区间为 [2.026, 2.224] bit/s/Hz，支持相同恒模全连接约束下的算法有效性。

4 bit QPS 与连续相位 CM-DR²-HBF 曲线基本重合；20 dB 差值为 0.063 bit/s/Hz，95% 置信区间跨 0，本组参数下量化损失不显著。DG-OLS、SJG-OLS 和 SIG-OLS 的差异涉及字典、连接及硬件约束，仅反映系统级性能位置。

AP-DR²-HBF 在 -30~-15 dB 略优于 AP-AltMin，从 -10 dB 起低于后者。20 dB 时两者为 40.331 和 41.919 bit/s/Hz，前者低 1.588 bit/s/Hz，

95% 置信区间为 [-1.686, -1.491] bit/s/Hz，故不能声称其在连续幅相域内全面优于独立 AP 优化。相对 CM-DR²-HBF 的 0.898 bit/s/Hz 增益包含额外幅度自由度，属硬件一性能折中。

4.3 双原子、速率导向与保护机制消融

消融在同一批 200 个信道上比较单/双原子、代理/速率目标及 AP 保护版本。图 3 中，20 dB 时 AP-DR²-HBF 达到 40.331 bit/s/Hz，较 AP-1Atom-Rate、AP-2Atom-Proxy 和未保护的 AP-DR²-Native 分别提高 3.023、12.925 和 2.777 bit/s/Hz，相应 95% 置信区间均不含 0；0 dB 时保护增益仅 0.089 bit/s/Hz，表明其主要修正中高 SNR 的局部退化。

图 4 中，20 dB 时 CM-DR²-HBF 达到 39.432 bit/s/Hz，较 CM-1Atom-Rate 和 CM-2Atom-Proxy

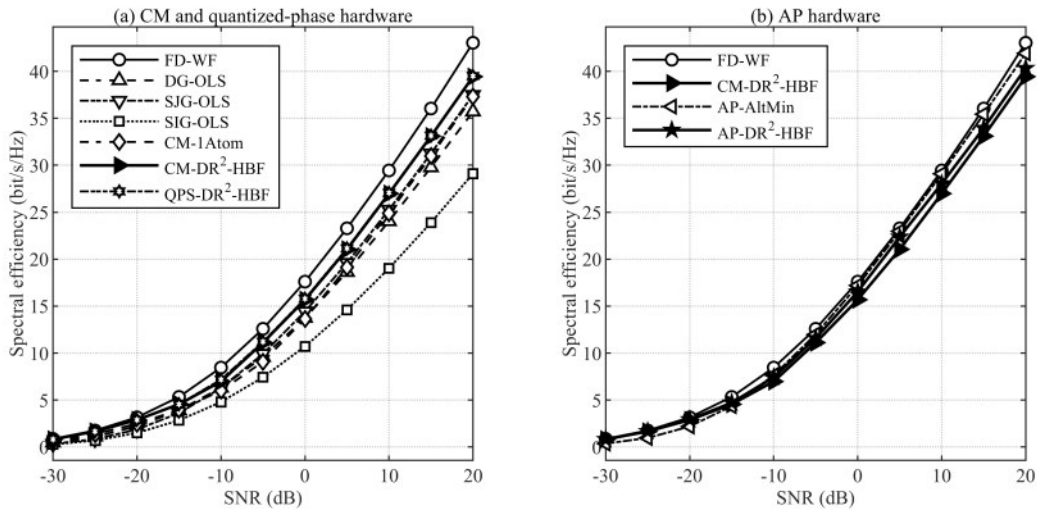
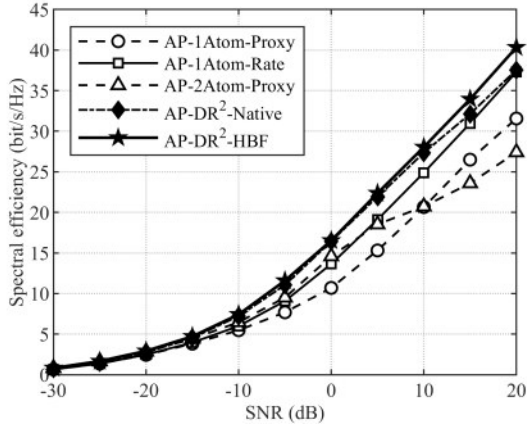
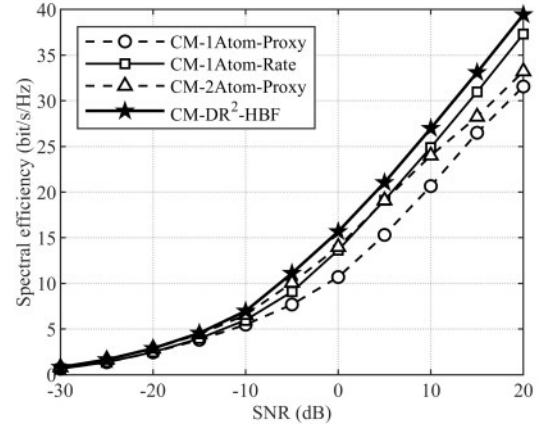


图2 不同硬件约束下的频谱效率随 SNR 变化曲线

图3 AP硬件下DR²-HBF关键模块消融结果图4 CM硬件下DR²-HBF关键模块消融结果

分别提高 2.125 和 6.214 bit/s/Hz; 0 dB 时增益为 2.034 和 1.697 bit/s/Hz。CM-2Atom-Proxy 在 15~20 dB 增益趋缓, 反映局部代理选出的新增原子经恒模投影后未必提高原始速率; 速率接受机制改善了这一现象。

两组消融表明, 双原子表示需结合硬件投影和原始 log-det 速率接受才能形成稳定增益; 此结论限于本文的候选集与顺序更新, 不构成全局最优性证明。

4.4 不完美 CSI 下的鲁棒性与误差机理

不完美 CSI 实验采用加性复高斯误差和联合发射/接收相位斜率、二次相位及复增益扰动的结构化误差^[6-7]。两类样本均经等能量变换, 使实际 NMSE 等于目标值。标准版本按估计信道筛选

和接受; 鲁棒版本使用 8 个训练样本, 取 $\kappa=0.5$ 并按第 3.2.4 节的保守效用规则接受更新, 并将标准解纳入候选。

图 5 给出 20 dB 的真实信道平均速率。NMSE=0.05 时, 鲁棒版本在加性/结构化误差下分别提高 0.364/0.176 bit/s/Hz, 95% 置信区间为 [0.153, 0.575]/[0.071, 0.280] bit/s/Hz。NMSE=0.20 时, 结构化误差下提高 0.260 bit/s/Hz; 加性误差下仅提高 0.061 bit/s/Hz 且区间含 0, 增益不显著。

图 6 基于 64 个独立验证样本计算 5% 分位速率。NMSE=0.05 和 0.20 时, 鲁棒版本在加性/结构化误差下的增益分别为 0.361/0.197 和 0.110/0.281 bit/s/Hz, 支持尾部速率改善。

图 7 给出已接受更新造成真实信道速率下降

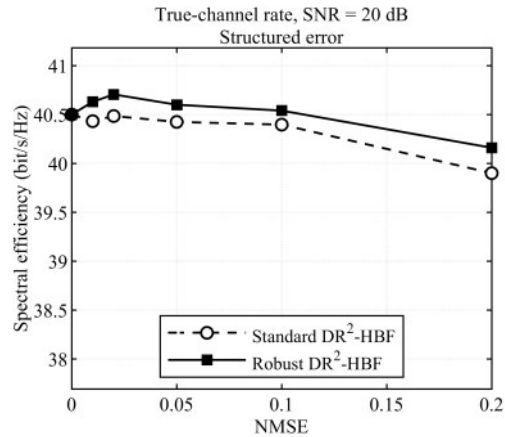
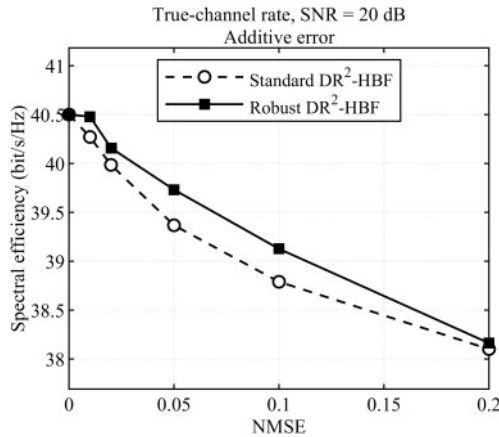


图5 不同信道误差下的真实信道平均频谱效率:(a)加性误差;(b)结构化误差

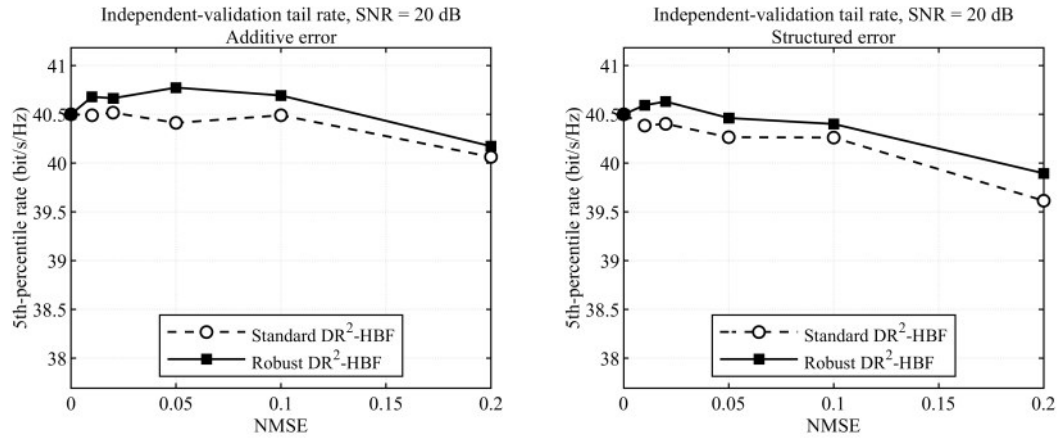


图6 独立验证样本的5%分位频谱效率

的比例。鲁棒版本在部分加性误差点降低该比例，但在NMSE=0.02、0.20时略有升高，结构化误差下也未一致降低。因此，统计鲁棒性改善不保证逐更新的真实信道速率单调。

表3在20 dB、NMSE=0.01的加性误差下进行TT/ET/TE/EE配对诊断；两个字母依次表示筛选和接受所用信道，T为真实信道，E为估计信道。

ET、TE相对TT的置信区间均含0，EE则显著损失0.419 bit/s/Hz；估计与真实信道的Top-K集合平均保持率为0.639。这支持候选排序与估计速率接受共同造成设计—评价失配的解释，不能将性能反转单独归因于双原子系数放大误差。鲁棒接受改善最终速率，但不能提供未知真实信道的逐更新保证。

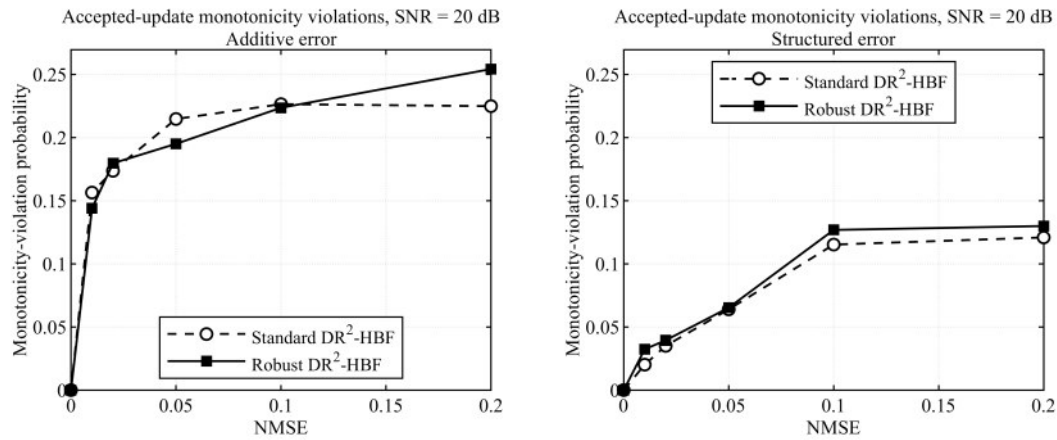


图7 已接受更新的真实信道单调性违背概率

表3 估计误差作用环节的TT ET TE EE配对诊断

组别	筛选信道	接受信道	平均速率/(bit·s ⁻¹ ·Hz ⁻¹)	相对TT差值及95%置信区间
TT	真实	真实	37.730	0
ET	估计	真实	37.749	0.018 [-0.252, 0.288]
TE	真实	估计	37.630	-0.100 [-0.315, 0.114]
EE	估计	估计	37.311	-0.419 [-0.750, -0.088]

共50个配对信道实现；T和E分别表示真实信道和估计信道。



4.5 Top-K截断与参数敏感性

为检验 Top-K 截断风险，在 $D=32$ 和 64 的缩小字典上枚举全部有序原子对，与不同 K_1/K_2 配对比较。表 4 的速率差为完整局部枚举减去 Top-K 结果，时间比为 Top-K 与局部枚举的运行时间之比。

默认 $K_1/K_2=48/24$ 时， $D=32$ 和 64 的平均速率差为 0.00069 和 0.00291 bit/s/Hz，时间降至局部枚举的 76.4% 和 29.0%。默认截断在该诊断中兼顾速率与时间，但局部枚举不能证明全局最优。

表 5 中， $\alpha_{max}=2$ 与 4 的统计量相同，默认 $\alpha_{max}=4$ 在当前设置中未形成约束； $\alpha_{max}=1$ 的均值仅高 0.040 bit/s/Hz，小于初始化波动。20 组初始化的标准差不超过 0.181 bit/s/Hz，说明对所用起点的敏感性有限，仍不能证明变量收敛至驻点或全局最优。

4.6 运行时间、候选规模与存储开销

复杂度实验固定 8 个天线块，将发射天线从 128 扩展至 2048，块长随之增加。每个规模执行 2 次外循环、重复 3 次；在线计时仅覆盖原生单分支核心搜索，不含离线字典生成。图 8 报告平均时间及 MATLAB 主变量内存占用之和。

天线数从 128 增至 2048 时，平均时间从 1.745 s 增至 7.338 s，内存代理从 0.0477 MiB 增至 0.5160 MiB。固定块数、 K_1/K_2 及外循环数使各规模均有 73728 个候选、73735 次原始速率计算；

增长主要来自信道—预编码乘法、字典相关性和更长块上的候选构造。该内存代理只是实现下界，并非进程峰值。

128 根发射天线， K_1/K_2 依次取 8/4、16/8、32/16 和 48/24 时，候选数为 2048、8192、32768、73728，平均时间为 0.0857、0.2445、0.7753、1.6128 s，体现高成本验证随 K_1/K_2 增长。增量实现与完整重算的最终速率差为 1.78×10^{-15} bit/s/Hz，但接收天线数为 8 时耗时 0.820 s，高于完整重算的 0.524 s，故数值一致性不意味着小接收维度下普遍加速。

AP 保护版本顺序执行原生 AP、CM 起点和 CM 热启动 AP 三个核心搜索，时间常数约为 3 次核心搜索，渐近复杂度阶不变；分支顺序复用内存保持同阶。图 8 只计单分支核心，应用时须计入上述保护成本。

4.7 硬件成本与归一化能效

表 6 统计 $N_t=128$ 、 $M_t=4$ 、 $P=8$ 时的器件成本。归一化模型中，基带、RF 链路、相移器、矢量调制器、可变增益放大器、合路器和功率放大器的单位功耗依次为 1、0.2、0.04、0.08、0.05、0.02 和 1，仅作同一假设下的相对比较。

AP 全连接功耗为 CM 的 2.14 倍；20 dB 时虽多 0.898 bit/s/Hz，归一化能效却为 0.4657，低于 CM 的 0.9732。QPS 与 CM 采用相同的单相移器功耗，能效接近；缺少分辨率相关器件功耗依据，

表 4 Top-K 截断与缩小字典局部枚举对比

D	K_1	K_2	平均速率差/(bit·s ⁻¹ ·Hz ⁻¹)	时间比	主/辅助原子保持率
32	8	4	0.15353	0.0402	0.711/0.450
32	16	8	0.05378	0.1366	0.914/0.727
32	32	16	0.00129	0.5150	0.997/0.858
32	48	24	0.00069	0.7642	1.000/0.895
64	8	4	0.50474	0.0099	0.458/0.280
64	16	8	0.08433	0.0344	0.786/0.573
64	32	16	-0.00035	0.1306	0.933/0.817
64	48	24	0.00291	0.2895	0.975/0.870

结果为 20 组随机初始化均值；局部枚举只覆盖缩小字典中的一次局部搜索，不是全局最优上界。 $D=64$ 、 $K_1/K_2=32/16$ 处的微小负差值源于贪婪路径依赖。

表5 复系数上限与随机初始化敏感性

$\alpha_{\max}$	平均速率/(bit·s ⁻¹ ·Hz ⁻¹)	标准差	最小值	最大值
1	23.658	0.181	23.261	23.910
2	23.618	0.168	23.239	23.886
4	23.618	0.168	23.239	23.886

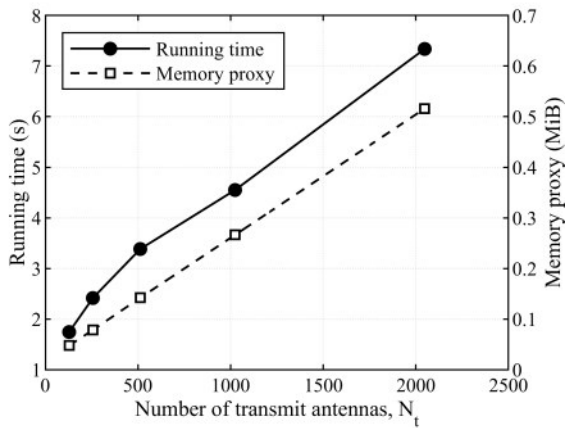


图8 发射天线数对单分支核心运行时间和内存代理的影响

不能声称4 bit器件的实际功耗更低。

部分连接的归一化功耗最低，双移相器全连接介于CM与AP之间；因缺少对应的同信道速率实验，两者仅比较器件数和功耗。本文直接算法证据来自CM同硬件组，AP结果反映幅相自由度、连续优化与实现成本的折中。RF链路数为4时，1024和2048根天线分别需4096和8192条AP支路，大规模应用须计入这一器件负担。

5 结束语

本文针对非平稳近场XL-MIMO提出DR²-HBF：由频谱效率梯度构造局部目标，以双原子

及受限复系数细化波束，通过两级Top-K筛选、硬件投影和原始速率验证完成更新。双原子局部代理包含单原子可行解；固定设计信道下，数字与模拟更新的速率保护使目标值单调有界，但不保证预编码变量收敛或全局最优。同硬件CM比较及消融支持该更新机制的有效性，AP版本并不普遍优于独立AP优化，其收益还需结合器件成本评价。

不完美CSI实验支持样本鲁棒机制对平均和尾部速率的改善，但部分误差点增益不显著，且未知真实信道上的逐更新单调性仍无保证。缩小字典诊断支持默认截断的速率—时间折中；大规模计时仅覆盖单分支核心，AP保护需另计多分支成本，归一化能效也不等同于实际能效。后续将研究连续AP域搜索、部分连接与双移相器实现、更强的鲁棒接受规则，以及基于实测器件功耗的能效评估和收敛条件。

表6 不同模拟架构的器件数、归一化功耗和能效

架构	NPS	NVM	NVGA	NComb	NRF/NPA	归一化功耗	20 dB 归一化能效
AP全连接	0	512	512	512	4/8	86.60	0.4657
CM全连接	512	0	0	512	4/8	40.52	0.9732
QPS全连接	512	0	0	512	4/8	40.52	0.9747
部分连接	128	0	0	0	4/8	14.92	—
双移相器全连接	1024	0	0	512	4/8	61.00	—

PS、VM、VGA和Comb分别表示相移器、矢量调制器、可变增益放大器和合路器；归一化能效为频谱效率与归一化功耗之比，不是实际bit/J。部分连接和双移相器架构没有对应的速率实验，故不计算能效。



参考文献:

- [1] CUI M Y, WU Z D, LU Y, et al. Near-field communications for 6G: fundamentals, challenges, potentials, and future directions[J]. IEEE Communications Magazine, 2023, 61(1): 40-46.
- [2] LIU Y W, WANG Z L, XU J Q, et al. Near-field communications: a tutorial review[J]. IEEE Open Journal of the Communications Society, 2023, 4: 1999-2049.
- [3] BOHAGEN F, ORTEN P, OIEN G E. Design of optimal high-rank line-of-sight MIMO channels[J]. IEEE Transactions on Wireless Communications, 2007, 6(4): 1420-1425.
- [4] DO H, LEE N, LOZANO A. Reconfigurable ULAs for line-of-sight MIMO transmission[J]. IEEE Transactions on Wireless Communications, 2021, 20(5): 2933-2947.
- [5] ZHANG H, SHLEZINGER N, GUIDI F, et al. Beam focusing for near-field multiuser MIMO communications[J]. IEEE Transactions on Wireless Communications, 2022, 21(9): 7476-7490.
- [6] CUI M Y, DAI L L. Channel estimation for extremely large-scale MIMO: far-field or near-field?[J]. IEEE Transactions on Communications, 2022, 70(4): 2663-2677.
- [7] LU Y, DAI L L. Near-field channel estimation in mixed LoS/NLoS environments for extremely large-scale MIMO systems[J]. IEEE Transactions on Communications, 2023, 71(6): 3694-3707.
- [8] DE CARVALHO E, ALI A, AMIRI A, et al. Non-stationarities in extra-large-scale massive MIMO[J]. IEEE Wireless Communications, 2020, 27(4): 74-80.
- [9] MOLISCH A F, RATNAM V V, HAN S Q, et al. Hybrid beamforming for massive MIMO: a survey[J]. IEEE Communications Magazine, 2017, 55(9): 134-141.
- [10] AHMED I, KHAMMARI H, SHAHID A, et al. A survey on hybrid beamforming techniques in 5G: architecture and system model perspectives[J]. IEEE Communications Surveys & Tutorials, 2018, 20(4): 3060-3097.
- [11] AYACH O E, RAJAGOPAL S, ABU-SURRA S, et al. Spatially sparse precoding in millimeter wave MIMO systems[J]. IEEE Transactions on Wireless Communications, 2014, 13(3): 1499-1513.
- [12] YU X H, SHEN J C, ZHANG J, et al. Alternating minimization algorithms for hybrid precoding in millimeter wave MIMO systems[J]. IEEE Journal of Selected Topics in Signal Processing, 2016, 10(3): 485-500.
- [13] 曹海燕, 杨晓慧, 胡文娟, 等. 毫米波大规模MIMO中基于部分连接的混合预编码方案[J]. 电信科学, 2020, 36(2): 35-42.
- CAO H Y, YANG X H, HU W J, et al. Hybrid precoding with sub-connected architecture for millimeter wave massive MIMO system[J]. Telecommunications Science, 2020, 36(2): 35-42.
- [14] SOHRABI F, YU W. Hybrid digital and analog beamforming design for large-scale antenna arrays[J]. IEEE Journal of Selected Topics in Signal Processing, 2016, 10(3): 501-513.
- [15] PARK S, ALKHATEEB A, HEATH R W. Dynamic subarrays for hybrid precoding in wideband mmWave MIMO systems[J]. IEEE Transactions on Wireless Communications, 2017, 16(5): 2907-2920.
- [16] 袁玥. 基于动态天线子阵列的毫米波混合预编码设计[D]. 西安: 西安邮电大学, 2019.
- YUAN Y. Hybrid precoding design for millimeter wave based on dynamic antenna subarray[D]. Xi'an: Xi'an University of Posts and Telecommunications, 2019.
- [17] HAN C, YAN L F, YUAN J H. Hybrid beamforming for terahertz wireless communications: challenges, architectures, and open problems[J]. IEEE Wireless Communications, 2021, 28(4): 198-204.
- [18] YANG S J, ELBIR A M, CHEN H, et al. Near-field hybrid beamforming for extremely large-scale (XL)-MIMO communications[J]. IEEE Transactions on Wireless Communications, 2025, 24(11): 9186-9200.
- [19] LIU M Z, LI M, LIU R, et al. Dynamic hybrid beamforming designs for ELAA near-field communications[J]. IEEE Journal on Selected Areas in Communications, 2025, 43(3): 644-658.
- [20] MENG C W, MA D Y, WANG Z L, et al. Near-field hybrid beamforming design for modular XL-MIMO ISAC systems[J]. IEEE Transactions on Communications, 2025, 73(11): 11840-11854.
- [21] ZHI K D, PAN C H, REN H, et al. Performance analysis and low-complexity design for XL-MIMO with near-field spatial non-stationarities[J]. IEEE Journal on Selected Areas in Communications, 2024, 42(6): 1656-1672.

王安定 (1975-), 男, 博士, 浙江工商大学信息与电子工程学院副教授, 主要研究方向为毫米波通信、空时信号处理以及 5G 通信中大规模 MIMO 系统的信道自适应跟踪、智能波束成形、信道容量分析及系统优化等无线核心技术。

高梅学健 (2001-), 男, 浙江工商大学信息与电子工程学院硕士生, 主要研究方向: 近场 MIMO 通信、混合波束成形、无线通信信号处理。

