



研究与开发

基于修正牛顿法的大规模 MIMO 低复杂度混合预编码算法

胡博, 王安定, 魏贵义

(浙江工商大学信息与电子工程学院, 浙江 杭州 310018)

摘要: 针对大规模多输入多输出 (multiple-input multiple-output, MIMO) 系统, 提出了一种基于修正牛顿 (modified Newton, MN) 法的相位跟踪算法, 有效地解决了传统高性能混合预编码方案中的高计算复杂度问题。该算法从子维度向量恢复的角度优化模拟预编码矩阵。在每个子维度优化中, 采用相位跟踪方法将模拟预编码向量的恢复转化为无约束的非线性优化问题, 并利用 MN 法进行求解。同时, 应用 Gerschgorin's Disk 定理和 Hermitian 矩阵分块求逆引理, 分别降低了 MN 法中计算修正因子以及 Hessian 矩阵求逆的计算复杂度。实验结果表明, 与仿真中几种传统的高性能混合预编码方案相比, 所提算法具有更高的频谱效率和更低的计算复杂度。

关键词: 大规模 MIMO; 混合预编码; 相位跟踪; 修正牛顿法

中图分类号: TN928

文献标志码: A

doi: 10.11959/j.issn.1000-0801.2023186

Low complexity hybrid precoding algorithm for massive MIMO based on modified Newton method

HU Bo, WANG Anding, WEI Guiyi

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

Abstract: A phase tracking algorithm was proposed for massive multiple-input multiple-output (MIMO) systems based on the modified Newton (MN) method, which effectively reduced the high computational complexity in traditional high-performance hybrid precoding schemes. The algorithm optimized the analog precoding matrix from the perspective of sub-dimensional vector recovery. In each sub-dimension optimization, the phase tracking method was used to transform the recovery of the analog precoding vectors into an unconstrained nonlinear optimization problem, which was then solved using the MN method. Concurrently, this strategy led to a marked reduction in the computational intricacy pertaining to both the computation of correction factors and the inversion of the Hessian matrix within the framework of the MN method. This was achieved through the insightful incorporation of Gerschgorin's Disk theorem and the Hermitian matrix block-inverse lemma. Simulation results show that the proposed algorithm has higher spectral efficiency and lower computational complexity than several conventional high-performance hybrid precoding schemes.

Key words: massive MIMO, hybrid precoding, phase tracking, modified Newton method

收稿日期: 2023-07-25; 修回日期: 2023-10-10

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

0 引言

随着无线通信业务的高速发展, 频谱资源日渐稀缺, 为了有效地扩展通信系统的带宽, 6G 将射频 (radio frequency, RF) 扩展至毫米波频段 (30~300 GHz)^[1]。然而, 高频的电磁波具有严重的路径损耗和较差的绕射能力^[2], 使得通信范围严重受限。为了解决这一问题, 学者们将毫米波与大规模 MIMO 技术结合, 不但有效地减小了天线阵列的尺寸, 而且补偿了毫米波的传输衰减, 为大规模天线阵列的部署提供了可操作性。除扩展使用频段外, 提高现有频段的频谱效率也是重点研究方向之一。而大规模 MIMO 系统可以借助预编码技术来提高频谱效率^[3]。

传统的纯数字预编码为每根天线配置专用的射频链路^[4-5], 因此, 纯数字预编码系统具有极高的硬件成本和功耗。为了解决该问题, 文献[6]提出了一种模拟/数字混合预编码结构, 它使用较少的射频链路将模拟预编码器和数字预编码器相连, 将混合预编码问题等效为稀疏信号重构问题, 并采用正交匹配追踪 (orthogonal matching pursuit, OMP) 算法进行求解, 从而降低了系统的硬件成本和功耗。为了降低 OMP 预编码的计算复杂度, 文献[6-10]提出了一些改进算法。其中, 文献[8]利用阶段确定的匹配追踪 (stage-determined matching pursuit, SdMP) 算法, 减少了 OMP 算法所需的迭代次数。文献[10]省略了 OMP 算法的迭代过程, 直接对阵列响应矩阵进行奇异值分解, 并选择与射频链路数量相同的列作为归一化后的模拟预编码矩阵, 从而显著降低了算法的复杂度, 但同时造成了较高的性能损失。文献[9]提出了正交性匹配追踪 (orthogonality-based matching pursuit, OBMP) 算法, 设计了二维正交码本, 使得 OMP 算法不再需要迭代。文献[7]利用天线阵列响应矩阵的相关性构建预编码码本候选集, 从而减小了 OMP 算法的匹配空间。文献[6-10]所提出的算法

均是基于码本的, 这使得它们在频谱效率方面有较大提升空间。

为了充分挖掘混合预编码算法在频谱效率上的潜力, 研究者将其表示为具有恒模约束的矩阵分解问题, 并采用交替最小化 (alternating minimization, AltMin) 准则设计数字/模拟混合预编码器。其中, 数字预编码矩阵通过最小二乘法计算, 而模拟预编码矩阵则分别利用共轭梯度 (conjugate gradient, CG)^[11]、梯度投影 (gradient projection, GP)^[12]、BB (barzilai-borwein) 梯度^[13]等算法得到。CG-AltMin 和 BB-AltMin 算法基于黎曼流形优化进行设计, 算法收敛后能获得较高的频谱效率。但二者只利用了目标函数的一阶导数信息, 这导致算法收敛速度较慢, 计算复杂度较高。GP-AltMin 算法利用梯度投影保证模拟预编码矩阵的恒模约束, 并通过适当放宽算法的收敛条件, 取得了较低的计算复杂度。然而, 与前两种算法相比, GP-AltMin 算法的性能略有下降。

上述研究主要针对单用户场景, 为解决多用户场景下的混合预编码问题, 文献[14]提出了一种简单有效的方法来充分利用大规模 MIMO 阵列的增益。该方法将基站与多个用户之间的信道纵向拼接, 构成组合信道矩阵, 并利用组合信道矩阵共轭转置的相位信息来更新模拟预编码矩阵。然而, 文献[14]只考虑了用户端配置单天线的情况。而文献[15-16]考虑了用户配置多根天线的情况。其中, 文献[15]提出了一种两阶段的混合预编码方案, 首先联合设计基站端和用户端的模拟预编码器和组合向量, 以最大化信道增益, 然后基于迫零 (zero-forcing, ZF) 数字预编码器抑制各用户之间的干扰。而文献[16]提出了一种混合最小均方误差 (hybrid minimum mean-squared error, hybrid MMSE) 预编码算法。在该算法中, 每个用户以最大化信道增益为目标来配置自己的模拟组合器, 并将等效信道反馈给基站端, 然后由基站端为所有用户设计混合预编码器。文献[16]考虑了噪



声对预编码的影响,并取得优于文献[15]的性能。然而,上述方法的和速率与纯数字预编码相比仍存在较大差距。为进一步提升和速率,文献[17]考虑了基站与各用户之间的通信环境,并引入注水法来优化用户功率的分配。另外,文献[11-13]所提出的基于交替优化的混合预编码算法也适用于多用户场景,但仍需要解决复杂度较高的问题。为降低复杂度,文献[18]重新考虑了混合预编码的设计问题,将预编码设计转化为等效信道特征值平方和的最大化问题,并提出了一种收发端联合设计的混合预编码方案。文献[19]从硬件层面入手,提出了一种基于电可控寄生单元阵列(electronically steerable parasitic array radiator, ESPAR)的混合预编码方案,该方案可以同时兼顾复杂度和性能,但 ESPAR 技术在频率限制和抗干扰方面仍存在较大的挑战^[20]。

近年来,基于深度学习的混合预编码技术^[21-22]受到广泛的关注,并取得了较优的性能表现。其中,文献[22]基于深度展开神经网络设计混合预编码方案,该方案将加权最小均方误差(weighted minimum mean-square error, WMMSE)算法改进为逐层(layer-wise)结构,并采用广义链式法则(generalized chain rule, GCR)计算可训练参数的梯度,提高了算法的收敛速度。文献[21]提出了一种基于复数反向传播(back propagation, BP)的混合预编码算法。该算法将模拟/数字预编码器的每个元素视为单个隐藏层神经网络的参数,并将混合预编码器的求解过程看作该神经网络的参数训练过程,最后利用 BP 算法,迭代求解出最优混合预编码器。

在上述文献中,尽管一些混合预编码方案^[11,13,17]具备较高的频谱效率,但它们通常需要在模拟预编码阶段进行大量高维度的矩阵运算,因而具有较高的计算复杂度,而多数优化算法^[12,18]在降低计算复杂度的同时却损失了一定的性能。与此同时,采用线下学习方式的深度学习方案^[22]在线下

训练阶段需要大量样本来精确调整网络参数,而训练完成的模型通常只适用于具体的系统参数,如基站天线数和用户数量。当系统参数变化时,需要重新进行训练,以确保网络模型的适用性^[23]。这将影响该类算法的实时性。

综上所述,非码本的大规模 MIMO 混合预编码算法在频谱效率和计算复杂度方面仍存在改进的空间。因此,本文提出了基于修正牛顿(modified Newton, MN)法的相位跟踪算法。该算法旨在通过 MN 法跟踪模拟预编码矩阵的相位参数,从而能够快速准确地估计模拟预编码矩阵。本文的主要贡献如下。

- 本文基于交替优化的混合预编码算法,提出了一种分离的求解模拟预编码矩阵的方案。该方案将整体优化问题划分为多个独立子维度的求解问题,可以并行地快速求解。由于待求解变量的维度较低,算法的计算复杂度显著降低。
- 本文将模拟预编码的设计视为 n 个单位模长复数 ($z \in \mathbb{C}^n : |z_1| = \dots = |z_n| = 1$) 约束下的最小二乘问题,针对该问题提出了一种无约束的非线性优化模型,并推导出该问题对应的目标函数的梯度向量和 Hessian 矩阵。这使得研究者可以利用多种经典的非线性优化算法求解模拟预编码矩阵。
- 本文研究了不同迭代停止条件对所提方案的频谱效率的影响,为进一步优化该类算法提供了新的思路。

符号说明: $\mathbb{C}$ 表示复数集合; $\mathbb{R}$ 表示实数集合; $(\cdot)^T$ 、 $(\cdot)^H$ 和 $(\cdot)^*$ 分别表示矩阵的转置、共轭转置和共轭; $[\cdot]_{ij}$ 表示矩阵第 i 行第 j 列的元素; $\| \cdot \|_F$ 表示矩阵的 Frobenius 范数; $| \cdot |$ 表示矩阵元素取模; $\odot$ 表示矩阵的 Hadamard 乘积; $\Re(\cdot)$ 表示取矩阵元素的实部; $\mathbf{A}$ 表示矩阵; $\mathbf{a}$ 表示向量; a 表示标量; $(\cdot)^{-1}$ 和 $(\cdot)^\dagger$ 分别表示矩阵求逆和矩阵求广义逆; $\mathbf{I}_n$ 表示 n 维单位矩阵。

1 系统描述

本节介绍了大规模 MIMO 混合预编码系统的传输模型及毫米波信道模型。

1.1 混合预编码系统的传输模型

单用户场景下混合预编码系统的传输模型如图 1 所示。发送端向接收端发送 N_s 条数据流，其中，发送端配置了 N_t 根天线，接收端拥有 N_r 根天线。为了实现多数据通信，需要在发送端和接收端配置 N_t^{RF} 和 N_r^{RF} 条射频链路，并且 $N_s \leq N_t^{\text{RF}} \leq N_t$ 、 $N_s \leq N_r^{\text{RF}} \leq N_r$ 。混合预编码器由模拟预编码矩阵 $\mathbf{F}_{\text{RF}} \in \mathbb{C}^{N_t \times N_t^{\text{RF}}}$ 和数字预编码矩阵 $\mathbf{F}_{\text{BB}} \in \mathbb{C}^{N_t^{\text{RF}} \times N_s}$ 相乘得到。离散时间的发射信号可以表示为 $\mathbf{x} = \mathbf{F}_{\text{RF}} \mathbf{F}_{\text{BB}} \mathbf{s}$ ，其中， $\mathbf{s} \in \mathbb{C}^{N_s \times 1}$ 为原始信号向量，并满足 $\mathbb{E}[\mathbf{s}\mathbf{s}^H] = \frac{1}{N_s} \mathbf{I}_{N_s}$ ， $\mathbf{I}_{N_s}$ 表示 N_s 维单位矩阵。此外，模拟预编码矩阵中的所有元素应具有相同的模值，即 $|\mathbf{F}_{\text{RF}}|_{ij}| = 1$ ，并且发送端归一化后的功率约束满足 $\|\mathbf{F}_{\text{RF}} \mathbf{F}_{\text{BB}}\|_F^2 = N_s$ 。

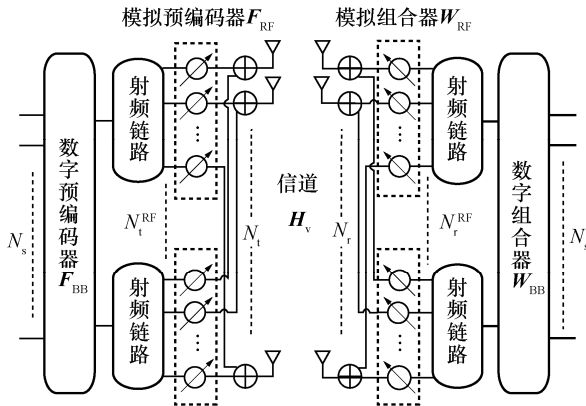


图 1 单用户场景下混合预编码系统的传输模型

一个窄带块衰落传播信道产生的接收信号为：

$$\mathbf{y} = \sqrt{\rho} \mathbf{H}_v \mathbf{F}_{\text{RF}} \mathbf{F}_{\text{BB}} \mathbf{s} + \mathbf{n} \quad (1)$$

其中， $\mathbf{H}_v$ 为 $N_r \times N_t$ 维复信道矩阵，满足 $\mathbb{E}[\|\mathbf{H}_v\|_F^2] = N_t N_r$ ； ρ 表示平均接收功率； $\mathbf{n} \sim \mathcal{CN}(0, \sigma_n^2 \mathbf{I}_{N_r})$ 为 $N_r \times 1$ 维高斯噪声向量， σ_n^2 为噪

声功率。

接收信号 $\mathbf{y}$ 经过接收端混合组合器 $\mathbf{W}_{\text{RF}} \mathbf{W}_{\text{BB}}$ 后可表示为：

$$\hat{\mathbf{s}} = \sqrt{\rho} \mathbf{W}_{\text{BB}}^H \mathbf{W}_{\text{RF}}^H \mathbf{H}_v \mathbf{F}_{\text{RF}} \mathbf{F}_{\text{BB}} \mathbf{s} + \mathbf{W}_{\text{BB}}^H \mathbf{W}_{\text{RF}}^H \mathbf{n} \quad (2)$$

其中， $\mathbf{W}_{\text{RF}} \in \mathbb{C}^{N_r \times N_r^{\text{RF}}}$ 为模拟组合矩阵， $\mathbf{W}_{\text{BB}} \in \mathbb{C}^{N_r^{\text{RF}} \times N_s}$ 为数字组合矩阵。此时，该毫米波通信系统的频谱效率可以表示为：

$$R = \text{lb} \left(\left(\mathbf{I}_{N_s} + \frac{\rho}{N_s} \mathbf{R}_n^{-1} \mathbf{W}_{\text{BB}}^H \mathbf{W}_{\text{RF}}^H \mathbf{H}_v \mathbf{F}_{\text{RF}} \mathbf{F}_{\text{BB}} \times \mathbf{F}_{\text{BB}}^H \mathbf{F}_{\text{RF}}^H \mathbf{H}_v^H \mathbf{W}_{\text{RF}} \mathbf{W}_{\text{BB}} \right) \right) \quad (3)$$

其中， $\mathbf{R}_n = \sigma_n^2 \mathbf{W}_{\text{BB}}^H \mathbf{W}_{\text{RF}}^H \mathbf{W}_{\text{RF}} \mathbf{W}_{\text{BB}}$ 为噪声协方差矩阵。本文以最大化频谱效率 R 为优化目标，展开对发送端和接收端的混合预编码器和组合器的研究。

1.2 毫米波信道模型

本文发送端和接收端均采用均匀平面阵列 (uniform planar array, UPA)。UPA 在 YOZ 坐标系的 y 轴部署 W 根天线、 z 轴部署 H 根天线，相邻天线单元之间具有相等的距离。UPA 的阵列响应向量 $\mathbf{a}_{\text{UPA}}(\phi, \theta)$ 可以表示为：

$$\mathbf{a}_{\text{UPA}}(\phi, \theta) = \frac{1}{\sqrt{N}} \left[1, \dots, e^{jkd(m \sin \phi \sin \theta + n \cos \theta)}, \dots, e^{jkd((W-1) \sin \phi \sin \theta + (H-1) \cos \theta)} \right]^T \quad (4)$$

其中， $k = \frac{2\pi}{\lambda}$ ， λ 为波长， d 表示天线单元之间的距离， ϕ 表示水平角， θ 表示俯仰角， m 和 n 表示天线单元在阵列平面中的索引，总天线数为 $N = WH$ 。

本文采用扩展的 saleh-valenzuela 窄带簇信道模型^[24]，假设信道由 N_{cl} 个散射簇组成，每个簇包含 N_{ray} 条传输路径。因此，离散时间的信道模型 $\mathbf{H}_v$ 可以表示为：

$$\mathbf{H}_v = \gamma \sum_{i=1}^{N_{\text{cl}}} \sum_{\ell=1}^{N_{\text{ray}}} \alpha_{i\ell} \mathbf{a}_r(\phi_{i\ell}^r, \theta_{i\ell}^r) \mathbf{a}_t(\phi_{i\ell}^t, \theta_{i\ell}^t)^H \quad (5)$$

其中， N_t 和 N_r 分别表示发送端和接收端的天线阵列



中的天线数量, $\gamma = \sqrt{N_t N_r / N_{cl} N_{ray}}$ 是归一化参数; $\alpha_{i\ell}$ 表示第 i 个散射簇中第 ℓ 条路径的复增益, 此外 $\phi_{i\ell}^r(\theta_{i\ell}^r)$ 和 $\phi_{i\ell}^t(\theta_{i\ell}^t)$ 分别表示接收端和发送端对应于第 i 个散射簇中的第 ℓ 条路径的水平 (俯仰) 角。向量 $\mathbf{a}_r(\phi_{i\ell}^r, \theta_{i\ell}^r) \in \mathbb{C}^{N_r \times 1}$ 和 $\mathbf{a}_t(\phi_{i\ell}^t, \theta_{i\ell}^t) \in \mathbb{C}^{N_t \times 1}$ 表示归一化后与角度 $\phi_{i\ell}^r(\theta_{i\ell}^r)$ 和 $\phi_{i\ell}^t(\theta_{i\ell}^t)$ 对应的接收端和发送端的阵列响应向量。此外, 复增益 $\alpha_{i\ell} \sim \mathcal{CN}(0, \sigma_{\alpha,i}^2)$, 其中 $\sigma_{\alpha,i}^2$ 表示第 i 个散射簇的功率。

2 基于修正牛顿法的相位跟踪算法

本节将在单用户场景下, 针对下行链路的混合预编码的设计进行研究。首先, 给出单用户混合预编码方案的设计模型, 并采用收发端分离的方式进行设计。随后, 详细分析了发送端混合预编码的设计模型, 并采用交替优化准则设计混合预编码器。在交替优化过程中, 将模拟预编码器的设计问题拆分为多个子维度的优化问题。通过相位跟踪方法, 将求解具有恒模约束的复数矩阵转化为求解复矩阵的相位矩阵, 并以最小化损失函数为优化目标, 提出了一种基于 MN 法的混合预编码算法。最后, 将所提算法拓展至接收端混合组合器的设计中。

2.1 单用户场景下混合预编码方案的设计模型

针对单用户背景, 下行链路的混合预编码的优化问题可表示为:

$$\begin{aligned} \max_{\mathbf{F}_{RF}, \mathbf{F}_{BB}, \mathbf{W}_{RF}, \mathbf{W}_{BB}} \quad & R \\ \text{s.t.} \quad & \|\mathbf{F}_{RF} \mathbf{F}_{BB}\|_F^2 = N_s \\ & |[\mathbf{W}_{RF}]_{ij}| = 1, \forall i, j \\ & |[\mathbf{F}_{RF}]_{ij}| = 1, \forall i, j \end{aligned} \quad (6)$$

显然, 式 (6) 是一个具有恒模约束的非凸优化问题, 因此难以直接求解。通过分开设计混合预编码器和组合器, 可将式 (6) 解耦成两个具有恒模约束的矩阵分解子问题, 并分别在第 2.2 节和第 2.3 节给出它们的求解方法。

2.2 发送端混合预编码的设计

参考文献[6], 混合预编码的优化问题可以表示为:

$$\begin{aligned} \min_{\mathbf{F}_{RF}, \mathbf{F}_{BB}} \quad & \|\mathbf{F}_{opt} - \mathbf{F}_{RF} \mathbf{F}_{BB}\|_F^2 \\ \text{s.t.} \quad & \|\mathbf{F}_{RF} \mathbf{F}_{BB}\|_F^2 = N_s \\ & |[\mathbf{F}_{RF}]_{ij}| = 1, \forall i, j \end{aligned} \quad (7)$$

式 (7) 实际上可归纳为求解混合预编码矩阵 $\mathbf{F}_{RF} \mathbf{F}_{BB}$ 与无约束纯数字预编码矩阵 $\mathbf{F}_{opt}$ 的最小欧氏距离问题。其中, $\mathbf{F}_{opt} \in \mathbb{C}^{N_t \times N_s}$ 由信道矩阵 $\mathbf{H}_v$ 奇异值分解后的右酉矩阵 $\mathbf{V}$ 的前 N_s 列构成。由于式 (7) 涉及两个变量 $\mathbf{F}_{RF}$ 和 $\mathbf{F}_{BB}$ 的求解, 可采用交替优化准则, 在每一轮迭代中, 固定其中一个变量并更新另一个变量, 此过程交替进行, 直至算法收敛。

首先, 算法通过固定 $\mathbf{F}_{RF}$ 来求解 $\mathbf{F}_{BB}$ 。在忽略发射功率约束的条件下, 优化问题可表示为:

$$\min_{\mathbf{F}_{BB}} \|\mathbf{F}_{opt} - \mathbf{F}_{RF} \mathbf{F}_{BB}\|_F^2 \quad (8)$$

基于最小二乘法, 得到该问题的最优解为:

$$\mathbf{F}_{BB} = \mathbf{F}_{RF}^{(k)\dagger} \mathbf{F}_{opt} \quad (9)$$

随后, 模拟预编码器 $\mathbf{F}_{RF}$ 的求解可采用相似的方法, 在固定 $\mathbf{F}_{BB}$ 的条件下求解。

$$\begin{aligned} \min_{\mathbf{F}_{RF}} \quad & \|\mathbf{F}_{opt} - \mathbf{F}_{RF} \mathbf{F}_{BB}\|_F^2 \\ \text{s.t.} \quad & |[\mathbf{F}_{RF}]_{ij}| = 1, \forall i, j \end{aligned} \quad (10)$$

将式 (10) 中的优化目标展开, 可得到:

$$\min \|\mathbf{F}_{opt} - \mathbf{F}_{RF} \mathbf{F}_{BB}\|_F^2 = \sum_{i=1}^{N_t} \min \left\| [\mathbf{F}_{opt}]_{i,:} - [\mathbf{F}_{RF}]_{i,:} \mathbf{F}_{BB} \right\|_F^2 \quad (11)$$

其中, $[\mathbf{F}_{opt}]_{i,:}$ 和 $[\mathbf{F}_{RF}]_{i,:}$ 分别表示矩阵 $\mathbf{F}_{opt}$ 和 $\mathbf{F}_{RF}$ 的第 i 行。因此, 式 (11) 可分解为 N_t 个独立的子优化问题。其中第 i 个子问题可以表示为:

$$\min_{[\mathbf{F}_{RF}]_{i,:}} \left\| [\mathbf{F}_{opt}]_{i,:} - [\mathbf{F}_{RF}]_{i,:} \mathbf{F}_{BB} \right\|_F^2, \text{s.t. } |[\mathbf{F}_{RF}]_{ij}| = 1, \forall j \quad (12)$$

在式 (12) 中, 基于恒模约束, 求解 $[\mathbf{F}_{RF}]_{i,:}$ 实际上等同于求解 $[\mathbf{F}_{RF}]_{i,:}$ 的相位 $[\mathbf{P}]_{i,:}$, 因此可以将

优化问题改写为:

$$\min_{\mathbf{F}_{\text{RF}}]_{i,:}} \left\| [\mathbf{F}_{\text{opt}}]_{i,:} - \mathbf{e}^{j[\mathbf{P}]_{i,:}} \mathbf{F}_{\text{BB}} \right\|_F^2 \quad (13)$$

其中, j 为虚数单位; $\mathbf{e}^{j[\mathbf{P}]_{i,:}}$ 表示以常数 e 为底、以矩阵 $[\mathbf{P}]_{i,:}$ 中各元素 $p_{i,:}$ 为指数的幂运算, 即 $\mathbf{e}^{j[\mathbf{P}]_{i,:}} = \mathbf{e}^{j p_{i,:}}$ 。为了便于表示, 记 $\mathbf{X} = \mathbf{P}^T$ 、 $\mathbf{O} = \mathbf{F}_{\text{opt}}^T$ 。

此外, $\mathbf{x}_i$ 和 $\mathbf{o}_i$ 分别表示 $\mathbf{X}$ 和 $\mathbf{O}$ 的第 i 列。设式(13)中的损失函数为 $f(\mathbf{x}_i)$, 可采用下降法求解此最小化问题 $\min f(\mathbf{x}_i)$, 基于迭代方程:

$$\mathbf{x}_i^{(k+1)} = \mathbf{x}_i^{(k)} + \mu_i^{(k)} \mathbf{d}_i^{(k)} \quad (14)$$

求解最优点 $\mathbf{x}_i^{\text{opt}}$, 其中 $k=1,2,\dots$ 表示迭代次数, $\mu_i^{(k)}$ 表示第 k 次迭代的步长, $\mathbf{d}_i^{(k)}$ 表示目标函数 $f(\mathbf{x}_i^{(k)})$ 在第 k 次迭代的搜索方向。目标函数在点 $\mathbf{x}_i^{(k)}$ 的二阶泰勒近似展开式为:

$$f(\mathbf{x}_i^{(k)} + \mathbf{d}_i^{(k)}) \approx f(\mathbf{x}_i^{(k)}) + (\nabla f(\mathbf{x}_i^{(k)}))^T \mathbf{d}_i^{(k)} + \frac{1}{2} (\mathbf{d}_i^{(k)})^T \nabla^2 f(\mathbf{x}_i^{(k)}) \mathbf{d}_i^{(k)} \quad (15)$$

其中,

$$\nabla f(\mathbf{x}_i^{(k)}) = \left. \frac{\partial f(\mathbf{x}_i)}{\partial \mathbf{x}_i} \right|_{\mathbf{x}_i = \mathbf{x}_i^{(k)}} \in \mathbb{R}^{N_{\text{RF}} \times 1} \quad (16)$$

$$\nabla^2 f(\mathbf{x}_i^{(k)}) = \left. \frac{\partial^2 f(\mathbf{x}_i)}{\partial \mathbf{x}_i^T \partial \mathbf{x}_i} \right|_{\mathbf{x}_i = \mathbf{x}_i^{(k)}} \in \mathbb{R}^{N_{\text{RF}} \times N_{\text{RF}}} \quad (17)$$

分别为 $f(\mathbf{x}_i^{(k)})$ 的梯度向量以及 Hessian 矩阵, 分别记作 $\mathbf{g}_i^{(k)}$ 和 $\mathbf{H}_i^{(k)}$, 两者的计算式分别为:

$$\mathbf{g}_i^{(k)} = 2\Re \left(\left(\mathbf{F}_{\text{BB}} \mathbf{F}_{\text{BB}}^H \mathbf{e}^{-j\mathbf{x}_i^{(k)}} - \mathbf{F}_{\text{BB}} \mathbf{o}_i^* \right) \odot j \mathbf{e}^{j\mathbf{x}_i^{(k)}} \right) \quad (18)$$

$$\mathbf{H}_i^{(k)} = \nabla_{\mathbf{x}_i^{(k)}} \left(\mathbf{g}_i^{(k)} \right)^T = \begin{bmatrix} \frac{\partial \mathbf{g}_{l_i}^{(k)}}{\partial \mathbf{x}_{l_i}^{(k)}} & \dots & \frac{\partial \mathbf{g}_{N_{\text{RF}}}^{(k)}}{\partial \mathbf{x}_{l_i}^{(k)}} \\ \vdots & \ddots & \vdots \\ \frac{\partial \mathbf{g}_{l_i}^{(k)}}{\partial \mathbf{x}_{N_{\text{RF}}}^{(k)}} & \dots & \frac{\partial \mathbf{g}_{N_{\text{RF}}}^{(k)}}{\partial \mathbf{x}_{N_{\text{RF}}}^{(k)}} \end{bmatrix} \quad (19)$$

其中, $\Re(\cdot)$ 表示取矩阵元素的实部, $\odot$ 表示矩阵的 Hadamard 乘积, $\mathbf{g}_{l_i}^{(k)}$ 和 $\mathbf{x}_{l_i}^{(k)}$ 分别表示 $\mathbf{g}_i^{(k)}$ 和 $\mathbf{x}_i^{(k)}$ 的第 l 行元素。令 $\mathbf{q}_i^{(k)} = \mathbf{F}_{\text{BB}} \mathbf{F}_{\text{BB}}^H \mathbf{e}^{-j\mathbf{x}_i^{(k)}} - \mathbf{F}_{\text{BB}} \mathbf{o}_i^*$, 此

时式(19)中 $\mathbf{H}_i^{(k)}$ 的元素为:

$$h_{ml}^{(k)} = \frac{\partial \mathbf{g}_{mi}^{(k)}}{\partial \mathbf{x}_{li}^{(k)}} = 2\Re \begin{cases} bb_{mm} - \mathbf{e}^{j\mathbf{x}_{mi}^{(k)}} \mathbf{q}_{mi}^{(k)}, m=l \\ bb_{ml} \mathbf{e}^{j\mathbf{x}_{mi}^{(k)}} \mathbf{e}^{-j\mathbf{x}_{li}^{(k)}}, m \neq l \end{cases} \quad (20)$$

其中, $\mathbf{q}_{mi}^{(k)}$ 为 $\mathbf{q}_i^{(k)}$ 的第 m 行元素, bb_{ml} 被定义为:

$$bb_{ml} = \left[\mathbf{F}_{\text{BB}} \mathbf{F}_{\text{BB}}^H \right]_{ml} = \sum_{k=1}^{N_i} b_{mk} b_{lk}^* \quad (21)$$

有关 $\mathbf{g}_i^{(k)}$ 和 $\mathbf{H}_i^{(k)}$ 的推导细节详见附录部分。

此时, 最优的下降方向 $\mathbf{d}_i^{(k)}$ 能够最小化式(22)所示的 $f(\mathbf{x}_i^{(k)})$ 的二阶泰勒逼近函数。

$$\min_{\mathbf{d}_i^{(k)}} f(\mathbf{x}_i^{(k)} + \mathbf{d}_i^{(k)}) = f(\mathbf{x}_i^{(k)}) + (\mathbf{g}_i^{(k)})^T \mathbf{d}_i^{(k)} + \frac{1}{2} (\mathbf{d}_i^{(k)})^T \mathbf{H}_i^{(k)} \mathbf{d}_i^{(k)} \quad (22)$$

在最优点时, 相对于参数向量 $\mathbf{d}_i^{(k)}$ 的梯度等于 0, 即:

$$\frac{\partial f(\mathbf{x}_i^{(k)} + \mathbf{d}_i^{(k)})}{\partial \mathbf{d}_i^{(k)}} = \mathbf{g}_i^{(k)} + \mathbf{H}_i^{(k)} \mathbf{d}_i^{(k)} = 0 \Leftrightarrow \mathbf{d}_i^{(k)\text{NT}} = -(\mathbf{H}_i^{(k)})^{-1} \mathbf{g}_i^{(k)} \quad (23)$$

其中, $\mathbf{d}_i^{(k)\text{NT}}$ 也被称为牛顿方向。若 Hessian 矩阵 $\mathbf{H}_i^{(k)}$ 正定, 则牛顿方向 $\mathbf{d}_i^{(k)\text{NT}}$ 是损失函数的下降方向, 这是因为 $-(\mathbf{g}_i^{(k)})^T \mathbf{H}_i^{(k)-1} \mathbf{g}_i^{(k)} < 0$ 。当 $\mathbf{x}_i^{(k)}$ 接近最优点 $\mathbf{x}_i^{\text{opt}}$ 时, 牛顿法具有二阶甚至更快的收敛速度。然而, 当 $\mathbf{x}_i^{(k)}$ 距离最优点较远时, $\mathbf{H}_i^{(k)}$ 可能会出现非正定或奇异的情况, 将导致算法无法继续迭代。

为了进一步优化算法, 本文采用 MN 法处理 $\mathbf{H}_i^{(k)}$ 非正定的情况, 即:

$$\mathbf{d}_i^{(k)} = -(\mathbf{H}_i^{(k)} + \delta_i^{(k)} \mathbf{I}_{N_{\text{RF}}})^{-1} \mathbf{g}_i^{(k)} \quad (24)$$

当修正因子 $\delta_i^{(k)} > -\lambda_{\min}$ 时, 可保证 $\mathbf{H}_i^{(k)} + \delta_i^{(k)} \mathbf{I}_{N_{\text{RF}}}$ 为对称正定矩阵, 其中 $\lambda_{\min}$ 为 $\mathbf{H}_i^{(k)}$ 的最小特征值。此外, 当 $\delta_i^{(k)}$ 较小时, 式(24)中



的搜索方向 $\mathbf{d}_i^{(k)}$ 偏向牛顿方向, 随着 $\delta_i^{(k)}$ 增大, $\mathbf{d}_i^{(k)}$ 会向负梯度方向偏移, 当 $\mathbf{H}_i^{(k)}$ 奇异时, 牛顿方向等于负梯度方向。

下面对 MN 法的特点、修正因子的选择以及低复杂度的矩阵求逆做以下说明。

说明 1 将 MN 法应用于优化问题需要计算损失函数的 Hessian 矩阵。在每一步迭代中, MN 法需要计算约 $\frac{n(n+1)}{2}$ 个二阶偏导数。此外, MN 法还需要对修正后的 Hessian 矩阵 (记为 $\mathbf{H} + \delta \mathbf{I}$, $\mathbf{I}$ 表示与 $\mathbf{H}$ 维度相同的单位矩阵) 进行求逆操作。MN 法的这些缺陷使得它在高维变量的计算中复杂度过高。对于式 (10) 中需要求解的变量 $\mathbf{F}_{\text{RF}}$, 其维度为 $N_t \times N_t^{\text{RF}}$ 。在大规模 MIMO 场景下, 基站端的天线阵元数量 N_t 通常可以达到 64、256, 甚至更多。由于计算复杂度问题, 不能直接使用 MN 法来求解 $\mathbf{F}_{\text{RF}}$ 。

为了解决这个问题, 本文利用式 (11) 将式 (10) 分解为 N_t 个如式 (12) 所示的子问题。每个子问题中需要求解的变量 $\mathbf{x}_i$ 的维度仅为 RF 链路数 N_t^{RF} 。值得注意的是, N_t^{RF} 满足 $N_s < N_t^{\text{RF}} < 2N_s \ll N_t$, 其中 N_s 一般设置为 2 或 3。因此, $\mathbf{x}_i$ 的维度相对较低, 这保证了 MN 法的可行性。

说明 2 为了使 MN 法计算出的搜索方向更接近牛顿方向, 修正因子应在满足 $\delta > -\lambda_{\min}$ 的同时尽可能小。通常情况下, 计算最小特征值 $\lambda_{\min}$ 需要将 Hessian 矩阵 $\mathbf{H}$ 的所有特征值求出, 然后选择其中最小的一个。但 MN 方法只需要找到最小特征值的下界来保证修正后的 Hessian 矩阵 $\mathbf{H} + \delta \mathbf{I}$ 的正定性。圆盘定理^[25]能够在较低的计算复杂度下提供合理的 Hessian 矩阵 $\mathbf{H}$ 特征值的下界估计, 从而确定修正因子 δ , 使得 $\mathbf{H} + \delta \mathbf{I}$ 保持正定, 同时也保证了 MN 法提供的搜索方向更接近牛顿方向。

圆盘定理规定, 对于一个 n 阶复矩阵或实矩

阵 $\mathbf{A}$, 其位于 (i,j) 位置上的元素为 a_{ij} , 则 $\mathbf{A}$ 的特征值应在式 (25) 表示的复平面中。

$$|z - a_{ii}| \leq r_i, i = 1, 2, \dots, n \quad (25)$$

其中, r_i 表示 $\mathbf{A}$ 的第 i 列或行元素去掉 a_{ii} 后的绝对值之和, 即:

$$r_i = \sum_{j=1, i \neq j}^n |a_{ij}| \quad (26)$$

结合式 (25) 和式 (26), 修正因子 δ 可取为:

$$\delta = \min \{ \tilde{\lambda}_1, \dots, \tilde{\lambda}_{N_t^{\text{RF}}} \} \quad (27)$$

其中, $\tilde{\lambda}_i = h_{ii} - \sum_{j=1, i \neq j}^{N_t^{\text{RF}}} |h_{ij}|$, h_{ij} 为 Hessian 矩阵 $\mathbf{H}$ 的元素。

说明 3 MN 法的每步迭代都需要计算修正后的 Hessian 矩阵 $\mathbf{H} + \delta \mathbf{I}$ 的逆矩阵。传统的高斯消元法是一种求逆速度较快的算法, 其计算复杂度为 $\mathcal{O}((N_t^{\text{RF}})^3)$ 。然而, 该算法未能利用 $\mathbf{H} + \delta \mathbf{I}$ 为 Hermitian 矩阵的特性。为了降低计算开销, 本文利用 Hermitian 矩阵分块求逆引理计算 $\mathbf{H} + \delta \mathbf{I}$ 的逆矩阵。下面对 Hermitian 矩阵分块求逆引理进行简要介绍。

Hermitian 矩阵的分块形式可以表示为:

$$\mathbf{H}_{m+1} = \begin{bmatrix} \mathbf{H}_m & \mathbf{r}_m \\ \mathbf{r}_m^H & \rho_m \end{bmatrix} \quad (28)$$

矩阵 $\mathbf{H}_{m+1}^{-1}$ 可由递推式给出:

$$\mathbf{H}_{m+1}^{-1} = \begin{bmatrix} \mathbf{H}_m & \mathbf{0}_m \\ \mathbf{0}_m^H & 0 \end{bmatrix} + \frac{1}{\rho_m} \begin{bmatrix} \mathbf{b}_m \mathbf{b}_m^H & \mathbf{b}_m \\ \mathbf{b}_m^H & 1 \end{bmatrix} \quad (29)$$

其中,

$$\mathbf{b}_m = -\mathbf{H}_m^{-1} \mathbf{r}_m \quad (30)$$

$$\beta_m = \rho_m + \mathbf{r}_m^H \mathbf{b}_m \quad (31)$$

此时, N_t^{RF} 阶修正的 Hessian 矩阵 $\mathbf{H} + \delta \mathbf{I}$ 求逆的计算复杂度为 $\mathcal{O}\left(\frac{(4N_t^{\text{RF}} - 1)(N_t^{\text{RF}} + 1)N_t^{\text{RF}}}{6}\right)$ 。

结合以上分析, 可以给出基于 MN 法的混合预编码方案的伪代码, 如算法 1 所示。

算法 1 MN-AltMin 混合预编码算法

输入 $\mathbf{F}_{\text{opt}}$
 输出 $\mathbf{F}_{\text{RF}}, \mathbf{F}_{\text{BB}}$
 1 利用随机相位构造 $\mathbf{X}^{(0)}$ 和 $\mathbf{F}_{\text{RF}}^{(0)}$ 并令 $k=0$;
 2 repeat
 3 固定 $\mathbf{F}_{\text{RF}}^{(k)}$, 求解 $\mathbf{F}_{\text{BB}}^{(k+1)} = \mathbf{F}_{\text{RF}}^{(k)\dagger} \mathbf{F}_{\text{opt}}$;
 4 固定 $\mathbf{F}_{\text{BB}}^{(k+1)}$, 利用算法 2 优化 $\mathbf{X}^{(k+1)}$ 和 $\mathbf{F}_{\text{RF}}^{(k+1)}$;
 5 $k \leftarrow k+1$;
 6 until 停止条件触发;
 7 $\mathbf{F}_{\text{BB}} = \sqrt{N_s} \frac{\mathbf{F}_{\text{BB}}}{\|\mathbf{F}_{\text{RF}} \mathbf{F}_{\text{BB}}\|_F}$;
 8 return $\mathbf{F}_{\text{RF}}, \mathbf{F}_{\text{BB}}$
 算法 1 中的迭代停止条件可设置为:

$$\left\| \mathbf{F}_{\text{opt}} - \mathbf{F}_{\text{RF}}^{(k+1)} \mathbf{F}_{\text{BB}}^{(k+1)} \right\|_F^2 - \left\| \mathbf{F}_{\text{opt}} - \mathbf{F}_{\text{RF}}^{(k)} \mathbf{F}_{\text{BB}}^{(k)} \right\|_F^2 < \varepsilon \quad (32)$$

或者迭代次数超过预定值 $T_{\max}^0$ 。其中, ε 为一个足够小的正值。算法 1 中第 7 步可以保证预编码算法满足发送功率约束。

算法 2 基于修正牛顿法的模拟预编码算法

输入 $\mathbf{F}_{\text{opt}}, \mathbf{F}_{\text{BB}}, \mathbf{F}_{\text{RF}}^{(0)}, \mathbf{X}^{(0)}$
 输出 $\mathbf{F}_{\text{RF}}, \mathbf{X}$
 1 parfor $i=1:N_t$
 2 令 $\mathbf{x}_i^0 = [\mathbf{X}^{(0)}]_{:,i}$, $\mathbf{o}_i = [\mathbf{F}_{\text{opt}}]_{i,:}^T$, $k=0$;
 3 repeat
 4 利用式 (24) 计算搜索方向 $\mathbf{d}_i^{(k)}$, 其中梯度向量 $\mathbf{g}_i^{(k)}$ 利用式 (18) 计算, Hessian 矩阵 $\mathbf{H}_i^{(k)}$ 利用式 (19) 计算, 修正因子 $\delta_i^{(k)}$ 和修正 Hessian 矩阵求逆分别利用式 (27) 和式 (29) 计算;
 5 利用 Armijo 线性回溯法计算步长 $\mu_i^{(k)}$;
 6 更新优化序列 $\mathbf{x}_i^{(k+1)} = \mathbf{x}_i^{(k)} + \mu_i^{(k)} \mathbf{d}_i^{(k)}$;

7 $k \leftarrow k+1$;
 8 until 停止条件触发;
 9 $[\mathbf{X}]_{:,i} = \mathbf{x}_i^{(k)}$;
 10 end
 11 $\mathbf{F}_{\text{RF}} = \mathbf{e}^{j\mathbf{X}^T}$;
 12 return $\mathbf{F}_{\text{RF}}, \mathbf{X}$

算法 2 的迭代停止条件一般设置为梯度的二范数 $\|\mathbf{g}_i^{(k)}\|_2 < \beta$ 或者迭代次数超过预定值 $T_{\max}^i$ 或者步长 $\mu_i^{(k)} < \rho$, 其中 β, ρ 均表示足够小的正值。在第 1 步中, 引入了 “parfor” 这一关键词, 以表示适宜并行执行的 “for” 循环。这一选择是基于各个循环之间无须进行数据通信和同步的考虑。因此, 以并行计算的方式迅速完成任务, 具备了充分的合理性。在第 5 步中提供了满足 Armijo 条件的步长 $\mu_i^{(k)}$ 以确保损失函数序列 $f(\mathbf{x}_i^{(k)})$ 单调递减, 从而保证算法的收敛性。

2.3 接收端混合组合器的设计

利用算法 1 计算出发送端混合预编码矩阵后, 将其与信道矩阵相乘可以得到等效信道 $\tilde{\mathbf{H}}_v = \mathbf{H}_v \mathbf{F}_{\text{RF}} \mathbf{F}_{\text{BB}}$, 进一步将其奇异值分解后得到的左酉矩阵 $\tilde{\mathbf{U}}$ 的前 N_s 列作为无约束全数字组合器 $\mathbf{W}_{\text{opt}} = [\tilde{\mathbf{U}}]_{:,1:N_s}$ 。参考发送端的设计思路, 可以写出接收端混合接收设计的优化问题。

$$\begin{aligned} \min_{\mathbf{W}_{\text{RF}}, \mathbf{W}_{\text{BB}}} & \left\| \mathbf{W}_{\text{opt}} - \mathbf{W}_{\text{RF}} \mathbf{W}_{\text{BB}} \right\|_F^2 \\ \text{s.t.} & \left| [\mathbf{W}_{\text{RF}}]_{ij} \right| = 1, \forall i, j \end{aligned} \quad (33)$$

接收端的混合接收器方案与发送端的混合预编码方案相似, 只需在执行算法 1 时, 不执行其第 7 步即可。

2.4 算法复杂度分析

本节旨在分析 MN-AltMin 算法中复数乘法和除法的次数, 并与典型的交替优化算法做比较。MN-AltMin 算法中每步操作所需的复数乘法和除法个数见表 1, MN-AltMin 算法与多种经典的交替优化算法的计算开销见表 2。



表 1 MN-AltMin 算法每步操作所需的复数乘法和除法个数

操作	复数乘法和除法个数
算法 1 (外部迭代)	$\mathbf{F}_{\text{RF}}^{(k)'} \mathbf{F}_{\text{opt}}$
算法 2 (内部子迭代)	$\mathbf{g}_i^{(k)}$
	$\mathbf{H}_i^{(k)}$
	$\mathbf{d}_i^{(k)}$

表 2 MN-AltMin 算法与多种经典的交替优化算法的计算开销

算法	计算开销
MN-AltMin	$N_{\text{iter}}^o (N_t (N_t^{\text{RF}})^2 + N_t N_t^{\text{RF}} N_s + f_{\text{inv}}(N_t^{\text{RF}})) + N_{\text{iter}}^i N_t \left(\frac{1}{6} (N_t^{\text{RF}})^3 + \left(N_s + \frac{15}{8} \right) (N_t^{\text{RF}})^2 + \left(N_s - \frac{1}{24} \right) N_t^{\text{RF}} + f_{\text{exp}}(N_t^{\text{RF}}) \right)$
CG-AltMin	$N_{\text{iter}}^o (N_t (N_t^{\text{RF}})^2 + N_t N_t^{\text{RF}} N_s + f_{\text{inv}}(N_t^{\text{RF}})) + N_{\text{iter}}^i N_t (4 N_s N_t N_t^{\text{RF}} + 11 N_t^{\text{RF}})$
BB-AltMin	$N_{\text{iter}}^o (N_t (N_t^{\text{RF}})^2 + N_t N_t^{\text{RF}} N_s + f_{\text{inv}}(N_t^{\text{RF}})) + N_{\text{iter}}^i N_t (2 N_s N_t N_t^{\text{RF}} + 8 N_t^{\text{RF}})$
GP-AltMin	$N_{\text{iter}}^o (N_t (N_t^{\text{RF}})^2 + N_t N_t^{\text{RF}} N_s + f_{\text{inv}}(N_t^{\text{RF}})) + N_{\text{iter}}^i N_t (2 N_s N_t N_t^{\text{RF}} + 2 N_t^{\text{RF}})$

对表 1 和表 2 中特殊的符号做以下说明:

- 在计算 $\mathbf{d}_i^{(k)}$ 所需的复数乘法次数中存在一个系数 $\alpha = 0.25$, 这是因为 $\mathbf{g}_i^{(k)}$ 和 $\mathbf{H}_i^{(k)}$ 都是实矩阵;

- 算法 1 中第 2~6 步为外部迭代过程, 算法所需外部迭代的总数记为 N_{iter}^o , 算法 2 中第 3~8 步为内部子迭代过程, N_{iter}^i 为所有外部迭代所需内部迭代个数的累加值;

f_{inv} 和 f_{exp} 分别为矩阵求逆运算和矩阵对应元素以 e 为底做幂运算, 其中 f_{inv} 采用高斯消元法求解, 所需的复数乘法个数为 n^3 ; f_{exp} 采用七阶麦克劳林公式计算, 所需的复数乘法个数为 $12\alpha n$ 。

3 多用户场景下 MN-AltMin 算法的应用

基于前文的分析, 本节将 MN-AltMin 算法从单用户场景扩展至多用户场景。多用户场景下混合预编码方案的系统框图如图 2 所示。设基站配置 N_{BS} 根天线, 并同时与 U 个用户进行通信。每个用户配备 N_{MT} 根天线和一条 RF 链路, 基站端配备 N_{RF} ($N_{\text{RF}} \geq N_s$) 条射频链路, 并传输 $N_s = U$ 路数据流。将 $U \times 1$ 维发送符号 $\mathbf{s} = [s_1, \dots, s_U]^T$ 送至混合预编码器 $\mathbf{F} = [\mathbf{f}_1, \dots, \mathbf{f}_U]$ 进行处理, 其中 s_u ($u=1, \dots, U$) 表示发送给第 u 个用户的信号, $\mathbf{f}_u = \mathbf{F}_{\text{RF}}[\mathbf{F}_{\text{BB}}]_{:,u}$ 表示与第 u 个用户对应的混合预编码向量。第 u 个用户经过模拟组合器 $\mathbf{w}_u \in \mathbb{C}^{N_{\text{MT}} \times 1}$ 处理后的信号可以表示为:

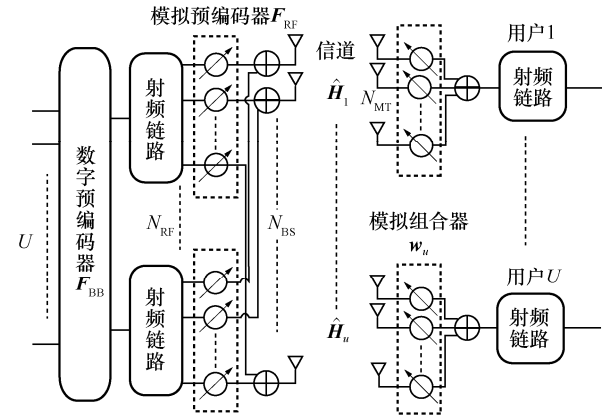


图 2 多用户场景下混合预编码方案的系统框图

$$y_u = \sqrt{\rho} \mathbf{w}_u^H \hat{\mathbf{H}}_u \sum_{i=1}^U \mathbf{f}_i s_i + \mathbf{w}_u^H \mathbf{n}_u \quad (34)$$

其中, ρ 为发射功率, $\hat{\mathbf{H}}_u \in \mathbb{C}^{N_{\text{MT}} \times N_{\text{BS}}}$ 为第 u 个用户对应的信道矩阵, 其表达式由式 (5) 给出; $\mathbf{n} \sim \mathcal{CN}(0; \sigma_n^2 \mathbf{I}_{N_{\text{MT}}})$ 为 $N_{\text{MT}} \times 1$ 维高斯噪声向量。

第 u 个用户可获得的数据速率为:

$$r_u = \text{lb} \left(1 + \frac{\frac{\rho}{U} |\mathbf{w}_u^H \hat{\mathbf{H}}_u \mathbf{f}_u|^2}{\frac{\rho}{U} \sum_{i \neq u} |\mathbf{w}_u^H \hat{\mathbf{H}}_u \mathbf{f}_i|^2 + \sigma_n^2} \right) \quad (35)$$

多用户混合预编码设计的优化目标为最大化

所有用户的和速率, 即:

$$\begin{aligned} & \max_{\mathbf{F}_{\text{RF}}, \mathbf{F}_{\text{BB}}, \{\mathbf{w}_u\}_{u=1}^U} \sum_{u=1}^U r_u \\ & \text{s.t. } [\mathbf{F}_{\text{RF}}]_{:,M} \in \mathcal{F}, \mathbf{w}_u \in \mathcal{W}, \|\mathbf{F}_{\text{RF}} \mathbf{F}_{\text{BB}}\|_{\text{F}}^2 = U \end{aligned} \quad (36)$$

其中, $\mathcal{F}$ 和 $\mathcal{W}$ 分别代表模拟预编码向量和模拟组合向量所有可能取值所组成的集合。直接求解式(36)需要搜索整个 $\mathcal{F}^U \times \mathcal{W}^U$ 空间中所有 $\mathbf{F}_{\text{RF}}$ 和 $\{\mathbf{w}_u\}_{u=1}^U$ 的组合, 因此很难实现。为此, 本文参照文献[16]的设计思路, 采用两阶段法将混合预编码器和模拟组合向量分开进行设计。

第一阶段, 每个用户设计各自的模拟组合向量以最大化信道增益, 即:

$$\max_{\mathbf{w}_u} \|\mathbf{w}_u^H \hat{\mathbf{H}}_u\|_2^2 \quad \text{s.t. } \mathbf{w}_u \in \mathcal{W} \quad (37)$$

第二阶段, 各用户计算自己的等效信道 $\mathbf{h}_u = \mathbf{w}_u^H \hat{\mathbf{H}}_u$ 并反馈给基站。接着, 基站计算出理想情况下的 ZF 纯数字预编码矩阵 $\mathbf{F}_{\text{ZF}} = \tilde{\mathbf{H}}^H (\tilde{\mathbf{H}} \tilde{\mathbf{H}}^H)^{-1}$, 其中 $\tilde{\mathbf{H}} = [\mathbf{h}_1, \dots, \mathbf{h}_U]^T$ 。最后, 基站可以参照单用户场景下的设计思路, 以 $\min \|\mathbf{F}_{\text{ZF}} - \mathbf{F}_{\text{RF}} \mathbf{F}_{\text{BB}}\|_{\text{F}}^2$ 为优化目标, 基于算法 1 计算 $\mathbf{F}_{\text{RF}}$ 和 $\mathbf{F}_{\text{BB}}$ 。

4 实验与分析

本节将基于 MATLAB R2021a 对算法性能进行仿真与评估, 仿真主机处理器型号为 Intel(R) Core(TM) i7-12700H, 主频为 2.7 GHz, RAM 为 32 GB。首先在单用户场景下设置仿真环境: 假设发射机和接收机都配置了 UPA, 其阵元数量分别为 64 个和 16 个; 发射机和接收机所配备的射频链路数统一用 N_{RF} 表示, 系统所传输的数据流数 $N_s = 2$ 。毫米波信道根据式(4)和式(5)生成, 散射簇数目 $N_{\text{cl}} = 8$, 每个簇的中心水平离开角 ϕ_i^l 和俯仰离开角 θ_i^l 分别在 $(-30^\circ, 30^\circ]$ 和 $(80^\circ, 100^\circ]$ 的范围内服从均匀分布, 中心到达角在 $(-180^\circ, 180^\circ]$ 的范围内服从均匀分布; 每个簇中有 $N_{\text{ray}} = 10$ 条路

径, 每条路径的到达角和离开角服从角度扩展为 7.5° 的拉普拉斯分布; 假设所有的散射簇具有相同的功率, 即 $\sigma_{\alpha,i}^2 = \sigma_{\alpha}^2 = 1$ 。将天线阵列阵元之间的距离 d 设为半波长, 即 $d = \frac{\lambda}{2}$, 噪声功率 σ_n^2 设置为 1。信噪比 (signal to noise ratio, SNR) 定义为 $\frac{\rho}{\sigma_n^2}$ 。

以下仿真将在随机生成 1 000 个信道实例的情况下, 比较本文提出的 MN-AltMin 算法与纯数字^[5]、OMP^[6]、CG-AltMin^[11]、GP-AltMin^[12]、BB-AltMin^[13]算法的性能。其中, 纯数字预编码算法用 OPT (optimum) 符号代指, 表示最优频谱效率。实验结果基于 1 000 次实验的平均值得出。为保证实验结果的准确性, 所有预编码器都配置相同的发射功率以及相同的迭代停止条件; 其中 ε 、 β 、 ρ 分别设置为 1.0×10^{-3} 、 1.0×10^{-6} 、 1.0×10^{-3} , T_{max}^i 和 T_{max}^o 分别设置为 1 000 和 100。

需要注意的是, GP-AltMin 算法每步迭代计算出梯度的二范数很难收敛到 1.0×10^{-6} 。因此, 该算法在更新模拟预编码时, 将超过最大迭代次数 (50 次) 作为迭代停止条件^[12]。此外, 由于 MN-AltMin 算法在更新模拟预编码时考虑的是分解后的子问题, 因此对其设置了更加严格的收敛条件, 即 $\beta = \frac{10^{-6}}{\sqrt{N_t}}$, 该条件可以保证当 N_t 个子问题迭代停止后, N_t 个梯度向量横向拼接构成的梯度矩阵的 Frobenius 范数满足 $\|\mathbf{g}_1, \dots, \mathbf{g}_{N_t}\|_{\text{F}} < 10^{-6}$ 。

射频链路数 $N_{\text{RF}} = 2$ 时不同算法的频谱效率对比如图 3 所示, 频谱效率的计算式为式(3)。随着 SNR 的增加, 所有算法的频谱效率均增加。由局部放大图可知, 在给定 SNR 的情况下, 本文所提出的 MN-AltMin 算法所实现的频谱效率均高于除 OPT 外的其他几种算法, 并且接近无约束 OPT 的性能。MN-AltMin 算法在收敛后, 能够保



证混合预编码矩阵 $\mathbf{F}_{\text{RF}}\mathbf{F}_{\text{BB}}$ 与纯数字预编码矩阵 $\mathbf{F}_{\text{opt}}$ 足够接近, 以及混合组合矩阵 $\mathbf{W}_{\text{RF}}\mathbf{W}_{\text{BB}}$ 与纯数字组合矩阵 $\mathbf{W}_{\text{opt}}$ 足够接近, 从而实现接近纯数字预编码的频谱效率。

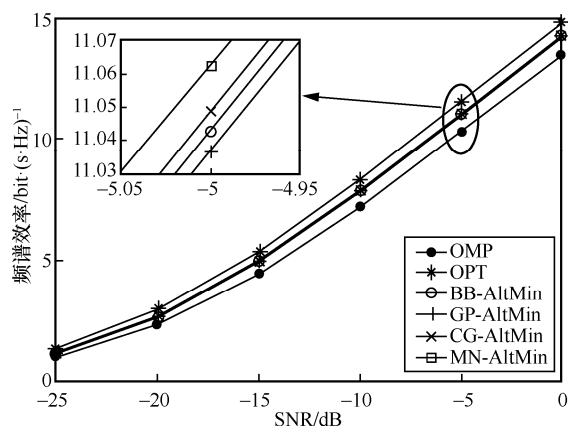


图3 $N_{\text{RF}} = 2$ 时不同算法的频谱效率对比

图3只考虑了 $N_{\text{RF}} = N_s$ 情况下多种预编码的性能。文献[26]的研究结果表明, 当 $N_{\text{RF}} \geq 2N_s$ 时, 混合预编码的设计问题存在闭式解, 这意味着混合预编码可以达到与纯数字预编码相同的性能。尽管本文主要关注 $N_s \leq N_{\text{RF}} < 2N_s$ 的情况, 但仍希望通过仿真实验, 验证所提出的算法能否在射频链路数充足的情况下接近纯数字预编码的性能表现。

SNR=0 时不同算法的频谱效率随着 N_{RF} 的变化情况如图4所示。随着 N_{RF} 的增加, MN-AltMin 算法与 OMP、CG-AltMin 和 BB-AltMin 算法的频谱效率逐渐提高。然而, GP-AltMin 算法在 $N_{\text{RF}} = 6$ 时的频谱效率出现下降, 这是因为该算法所设定

的迭代步长超过了 Lipschitz 常数^[12]的限制。在 $N_{\text{RF}} = 2$ 和 $N_{\text{RF}} = 3$ 的情况下, MN-AltMin 算法所实现的频谱效率均高于除 OPT 外的其他算法, 并且在 $N_{\text{RF}} = 2N_s = 4$ 的情况下, 达到接近纯数字预编码的性能水平。

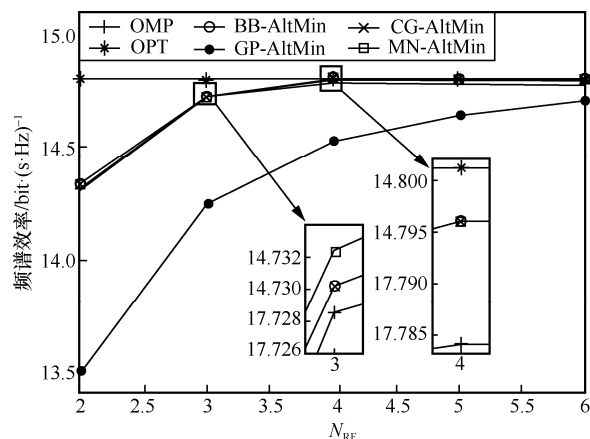


图4 SNR=0 时不同算法的频谱效率随着 N_{RF} 的变化情况

为了比较不同算法的复杂度, 表3列出了不同算法收敛时所需的平均迭代次数, 当 $N_{\text{RF}} = 2$ 和 $N_{\text{RF}} = 3$ 时本文提出的 MN-AltMin 算法具有远高于其他算法的收敛速度。

将表3的数据代入表2的计算式中, 可以得到不同算法的计算开销随着 N_{RF} 的变化情况, 如图5所示。图5表明本文提出的 MN-AltMin 算法的计算开销远低于其他算法。需要注意的是 BB-AltMin 算法的计算开销在 $N_{\text{RF}} \geq 4$ 时出现了激增, 这与文献[13]的仿真结果矛盾。出现这种情况的原因在于, BB 算法本质上是一种改进的最速下降法, 其在接近极小值点时的收敛速度会减慢,

表3 不同算法收敛时所需的平均迭代次数

N_{RF}	MN-AltMin		CG-AltMin		GP-AltMin		BB-AltMin	
	N_{iter}^i	N_{iter}^o	N_{iter}^i	N_{iter}^o	N_{iter}^i	N_{iter}^o	N_{iter}^i	N_{iter}^o
2	30.69	9.31	226.97	9.28	420.50	9.41	265.17	9.28
3	57.01	6.77	286.31	6.73	310.91	7.22	451.69	6.87
4	368.86	5.70	656.80	5.73	269.60	6.39	1 726.77	5.72
5	390.89	4.96	511.46	4.92	233.52	5.67	1 276.47	4.91
6	398.55	4.45	482.52	4.50	241.25	5.82	1 229.23	4.49

而文献[13]中设置了相对宽松的迭代停止条件,使得预编码算法的模拟域部分在未达到极小值点时就提前结束,从而得到了一个具有较低复杂度的 BB-AltMin 混合预编码算法。

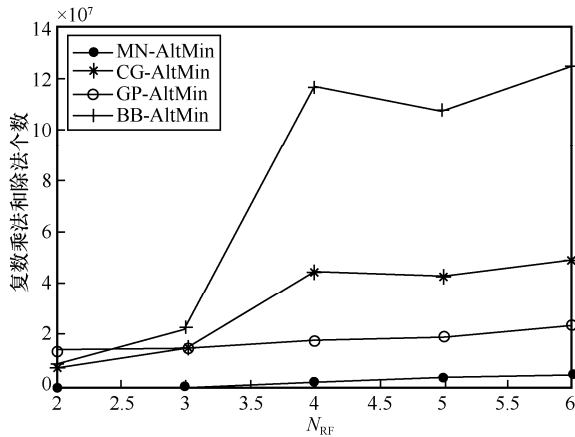


图5 不同算法的计算开销随着 N_{RF} 的变化情况

$N_{\text{RF}} = 3$ 、 $\text{SNR} = 0$ 时 MN-AltMin 算法在不同迭代停止条件下的频谱效率如图 6 所示。图 6 中, x 轴表示算法 2 中每个子问题求解时的最大可选迭代次数 T_{max}^i , y 轴表示损失函数在第 k 次迭代时的值与第 $k+1$ 次的值之间的残差的绝对值 ε , 而 z 轴表示算法的频谱效率。可以清楚地观察到 MN-AltMin 算法的频谱效率在 T_{max}^i 增加时改善较为有限; 然而, 随着 ε 的减小, 算法的频谱效率

显著提升。对比相同 ε 不同 T_{max}^i 的情况可以发现, 当 $T_{\text{max}}^i \leq 16$ 时算法的频谱效率保持稳步提升; 当 $T_{\text{max}}^i > 16$ 时, 算法的频谱效率在 $T_{\text{max}}^i = 16$ 时的频谱效率附近波动。此外, 通过表 2 中的计算式可以看出, MN-AltMin 算法每次内部迭代所需的计算开销远大于外部迭代。以 $N_s = 2$ 、 $N_t = 64$ 、 $N_{\text{t}}^{\text{RF}} = 3$ 为例, 每次内部迭代所需的复数乘法个数为 3 472, 而每次外部迭代所需的复数乘法个数为 987, 前者约为后者的 3.5 倍。

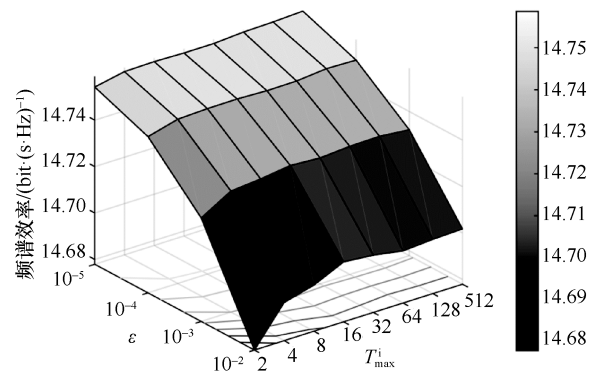


图6 $N_{\text{RF}} = 3$ 、 $\text{SNR} = 0$ 时 MN-AltMin 算法在不同迭代停止条件下的频谱效率

$N_{\text{RF}} = 3$ 时 MN-AltMin 算法在不同迭代停止条件下的平均迭代次数见表 4。表 4 进一步展示了图 6 所对应的不同迭代停止条件下的平均迭代次数。通过比较 $(T_{\text{max}}^i, \varepsilon)$ 为 $(512, 10^{-2})$ 与

表4 $N_{\text{RF}} = 3$ 时 MN-AltMin 算法在不同迭代停止条件下的平均迭代次数

ε		T_{max}^i							
		2	4	8	16	32	64	128	512
10^{-2}	N_{iter}^i	8.36	11.98	15.56	18.23	20.62	21.40	22.82	23.55
	N_{iter}^o	5.18	4.09	3.74	3.69	3.67	3.63	3.68	3.67
10^{-3}	N_{iter}^i	17.60	24.64	31.84	37.66	40.96	45.52	47.33	46.16
	N_{iter}^o	9.81	7.77	7.28	7.006	6.74	6.85	6.80	6.84
10^{-4}	N_{iter}^i	47.14	63.89	81.18	93.50	104.88	108.83	130.32	125.36
	N_{iter}^o	24.57	20.17	18.86	18.39	18.03	17.70	18.78	17.36
10^{-5}	N_{iter}^i	131.88	170.87	227.27	243.51	289.60	326.18	326.11	319.09
	N_{iter}^o	66.85	56.60	56.34	50.69	52.35	53.64	51.96	50.23



$(2,10^{-3})$ 、 $(512,10^{-3})$ 与 $(2,10^{-4})$ 、 $(512,10^{-4})$ 与 $(2,10^{-5})$ 的 3 组数据,可以观察到在每组内,仅 N_{iter}^o 有较大变化,这意味着两种迭代停止条件下的 MN-AltMin 算法的计算开销相近。同时,图 6 也表明在每组迭代停止条件中,具有较小 ε 值的 MN-AltMin 算法实现的频谱效率更高。这些结果表明,在保持计算开销相近的情况下,通过选择适当的迭代停止条件,可以提高算法的频谱效率。

下面将评估多用户场景下 MN-AltMin 算法的性能。假设基站端配备 64 根天线并同时与 U 个用户进行通信,每个用户配备 16 根天线。各用户和基站间的信道实例的产生方式与单用户场景下几乎相同,只需将簇中心角的到达角和离开角的范围改为 $(-180^\circ, 180^\circ]$ 。下面将在随机生成 1 000 个信道实例的情况下,将本文算法 MN-AltMin 与 Hybrid MMSE^[16]、CG-AltMin^[11]和 GP-AltMin^[12]进行对比与分析。此外,基于图 5 的分析,可将基于交替优化的 MN、CG 和 GP 算法的 T_{max}^i 和 ε 分别设置为 16 和 1.0×10^{-3} ,以同时兼顾算法的复杂度和性能。最后,上述算法中模拟组合向量均通过 OMP^[6] 计算,Hybrid MMSE 的码本由 2^6 组水平角和俯仰角参数对应的阵列响应向量构成,该 2^6 组参数则由水平和俯仰角度范围内均匀量化的 2^3 个水平角和俯仰角做笛卡儿积得到。

用户数量 $U=3$ 时不同算法的和速率随 SNR 的变化如图 7 所示,和速率的计算式由式 (36) 给出。从图 7 可以观察到,随着 SNR 的增加,所有算法的和速率都有所增加。对比 $N_{\text{RF}}=4$ 和 $N_{\text{RF}}=6$ 时 CG-AltMin、MN-AltMin 和 Hybrid MMSE 算法的表现,发现基于交替优化的 CG-AltMin 和 MN-AltMin 算法明显优于 Hybrid MMSE 算法。此外,当 N_{RF} 变大时,基于交替优化的 CG-AltMin 和 MN-AltMin 算法的性能相较于 Hybrid MMSE 算法提升明显。

SNR=0 时不同算法的和速率随用户数量 U 的变化如图 8 所示,考虑基站端 RF 链路数量高于用户数量的情况,即 $N_{\text{RF}}=1.5U$ ($\lfloor \cdot \rfloor$ 表示下

取整运算)。观察图 8 可以发现,随着用户数量 U 的增加,所有算法所实现的和速率都呈增长趋势。然而,用户数量的增加导致用户之间的干扰增加,因此增长趋势变得缓慢。在给定 U 的情况下, MN-AltMin 和 CG-AltMin 算法展现了相似的性能表现,两者的性能最接近纯数字预编码。此外,当 $U \leq 3$ 时,3 种基于交替优化算法 (MN-AltMin、CG-AltMin 和 GP-AltMin) 的性能几乎相同。然而,当 $U > 3$ 时, MN-AltMin 和 CG-AltMin 算法的性能开始明显优于 GP-AltMin 算法。

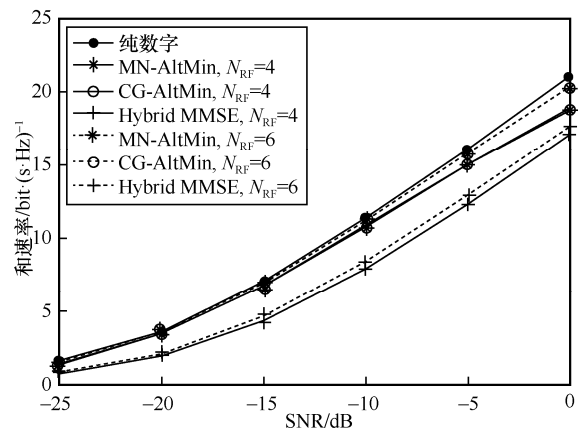


图 7 $U=3$ 时不同算法的和速率随 SNR 的变化

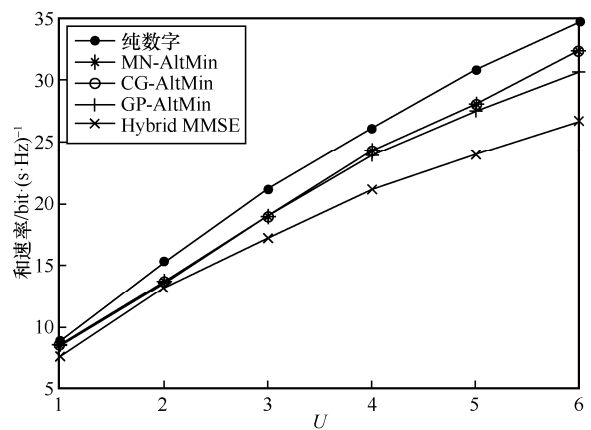


图 8 SNR=0 时不同算法的和速率随用户数量 U 的变化

综上所述,在单用户场景和多用户场景下,所提算法的频谱效率均高于参与比较的混合预编码算法,并接近纯数字预编码的性能。而在算法复杂度方面,与其他基于交替优化的算法相比,所提算法的复杂度明显降低。

5 结束语

本文针对大规模 MIMO 混合预编码系统, 提出一种基于交替优化的混合预编码方案和一种基于修正牛顿法的相位跟踪算法。该方案将模拟预编码矩阵的元素划分到多个独立的变量组, 并利用所提算法解决每个变量组所对应的子问题。本文还分析了该方案的计算复杂度和性能表现。在经典仿真场景下, 所提方案相较于其他类似的交替优化算法, 在计算复杂度明显降低的情况下仍具有一定程度的性能提升。未来将进一步研究不同射频链路和天线阵列连接情况下的混合预编码方案, 如子连接、部分连接和动态连接等, 并深入探究在超多用户接入时, 基站端和用户端天线数量之间的定量关系。

附录: $f(\mathbf{x}_i)$ 的梯度向量和 Hessian 矩阵表达式的推导

对 $f(\mathbf{P}) = \|\mathbf{F}_{\text{opt}} - \mathbf{e}^{j\mathbf{P}} \mathbf{F}_{\text{BB}}\|_{\text{F}}^2$ 微分可以得到:

$$\begin{aligned} d f(\mathbf{P}) &= d \left(\text{tr}((\mathbf{F}_{\text{opt}} - \mathbf{e}^{j\mathbf{P}} \mathbf{F}_{\text{BB}})^H (\mathbf{F}_{\text{opt}} - \mathbf{e}^{j\mathbf{P}} \mathbf{F}_{\text{BB}})) \right) = \\ &= d \left(\text{tr}(\mathbf{F}_{\text{BB}}^H \mathbf{e}^{-j\mathbf{P}^T} \mathbf{e}^{j\mathbf{P}} \mathbf{F}_{\text{BB}} - \mathbf{F}_{\text{BB}}^H \mathbf{e}^{-j\mathbf{P}^T} \mathbf{F}_{\text{opt}} - \mathbf{F}_{\text{opt}}^H \mathbf{e}^{j\mathbf{P}} \mathbf{F}_{\text{BB}}) \right) = \\ &= d \left(\text{tr}(\mathbf{F}_{\text{BB}}^H \mathbf{F}_{\text{RF}}^H \mathbf{F}_{\text{RF}} \mathbf{F}_{\text{BB}} - \mathbf{F}_{\text{BB}}^H \mathbf{F}_{\text{RF}}^H \mathbf{F}_{\text{opt}} - \mathbf{F}_{\text{opt}}^H \mathbf{F}_{\text{RF}} \mathbf{F}_{\text{BB}}) \right) = \\ &= \text{tr}(\mathbf{F}_{\text{BB}}^H d(\mathbf{F}_{\text{RF}}^H) \mathbf{F}_{\text{RF}} \mathbf{F}_{\text{BB}}) + \text{tr}(\mathbf{F}_{\text{BB}}^H \mathbf{F}_{\text{RF}}^H d(\mathbf{F}_{\text{RF}}) \mathbf{F}_{\text{BB}}) - \\ &= \text{tr}(\mathbf{F}_{\text{BB}}^H d(\mathbf{F}_{\text{RF}}^H) \mathbf{F}_{\text{opt}}) - \text{tr}(\mathbf{F}_{\text{opt}}^H d(\mathbf{F}_{\text{RF}}) \mathbf{F}_{\text{BB}}) = \\ &= \text{tr}(d(\mathbf{F}_{\text{RF}}^H) \mathbf{F}_{\text{RF}} \mathbf{F}_{\text{BB}} \mathbf{F}_{\text{BB}}^H) - \text{tr}(d(\mathbf{F}_{\text{RF}}^H) \mathbf{F}_{\text{opt}} \mathbf{F}_{\text{BB}}^H) + \\ &= \text{tr}(\mathbf{F}_{\text{BB}} \mathbf{F}_{\text{BB}}^H \mathbf{F}_{\text{RF}}^H d(\mathbf{F}_{\text{RF}})) - \text{tr}(\mathbf{F}_{\text{BB}} \mathbf{F}_{\text{opt}}^H d(\mathbf{F}_{\text{RF}})) = \\ &= \text{tr}((\mathbf{F}_{\text{RF}} \mathbf{F}_{\text{BB}} \mathbf{F}_{\text{BB}}^H - \mathbf{F}_{\text{opt}} \mathbf{F}_{\text{BB}}^H)^T d(\mathbf{F}_{\text{RF}})) + \\ &= \text{tr}((\mathbf{F}_{\text{RF}} \mathbf{F}_{\text{BB}} \mathbf{F}_{\text{BB}}^H - \mathbf{F}_{\text{opt}} \mathbf{F}_{\text{BB}}^H)^H d(\mathbf{F}_{\text{RF}})) \end{aligned} \quad (38)$$

定义矩阵 $\mathbf{Q} = (\mathbf{F}_{\text{RF}} \mathbf{F}_{\text{BB}} \mathbf{F}_{\text{BB}}^H - \mathbf{F}_{\text{opt}} \mathbf{F}_{\text{BB}}^H)^H$, 此时式 (38) 可

以被重新表述为:

$$\begin{aligned} d f(\mathbf{P}) &= \text{tr}(\mathbf{Q}^* d(\mathbf{e}^{-j\mathbf{P}})) + \text{tr}(\mathbf{Q} d(\mathbf{e}^{j\mathbf{P}})) = \\ &= \text{tr}(\mathbf{Q}^* (-j\mathbf{e}^{-j\mathbf{P}} \odot d\mathbf{P})) + \text{tr}(\mathbf{Q} (j\mathbf{e}^{j\mathbf{P}} \odot d\mathbf{P})) = \\ &= \text{tr}((\mathbf{Q}^* \odot -j\mathbf{e}^{-j\mathbf{P}^T}) d\mathbf{P}) + \text{tr}((\mathbf{Q} \odot j\mathbf{e}^{j\mathbf{P}^T}) d\mathbf{P}) = \\ &= \text{tr}(2\Re(\mathbf{Q} \odot j\mathbf{e}^{j\mathbf{P}^T}) d\mathbf{P}) \end{aligned} \quad (39)$$

由 $df(\mathbf{P}) = \text{tr}(\mathbf{A} d\mathbf{P}) \Leftrightarrow \nabla f(\mathbf{P}) = \mathbf{A}^T$ 可以直接给出 $f(\mathbf{P})$ 和 $f(\mathbf{X})$ 的梯度矩阵为:

$$(\nabla f(\mathbf{X}))^T = \nabla f(\mathbf{P}) = 2\Re(\mathbf{Q} \odot j\mathbf{e}^{j\mathbf{P}^T})^T \quad (40)$$

此时, 每个子问题中 $f(\mathbf{x}_i)$ 的梯度向量为:

$$\begin{aligned} \mathbf{g}_i &= (\nabla f(\mathbf{P}))_{:,i}^T = \nabla_{[\mathbf{P}]_{:,i}^T} f([\mathbf{P}]_{:,i}) = \nabla_{\mathbf{x}_i} f(\mathbf{x}_i^T) = \\ &= 2\Re([\mathbf{Q}]_{:,i} \odot j\mathbf{e}^{j\mathbf{x}_i}) = 2\Re((\mathbf{F}_{\text{BB}} \mathbf{F}_{\text{BB}}^H \mathbf{e}^{-j\mathbf{x}_i} - \mathbf{F}_{\text{BB}} \mathbf{o}_i^*) \odot j\mathbf{e}^{j\mathbf{x}_i}) \end{aligned} \quad (41)$$

结合式 (41) 可以写出 $\mathbf{g}_i$ 的元素表达式:

$$[\mathbf{g}_i]_m = g_{mi} = 2\Re\left(j\mathbf{e}^{j\mathbf{x}_{mi}} \left(\sum_{k=1}^{N_{\text{RF}}} b b_{mk} \mathbf{e}^{j\mathbf{x}_{ki}} - b_{mk} \mathbf{o}_{ki}^*\right)\right) \quad (42)$$

其中, $\mathbf{x}_{mi}$ 、 b_{mk} 和 $\mathbf{o}_{ki}$ 分别为 $\mathbf{X}$ 、 $\mathbf{F}_{\text{BB}}$ 和 $\mathbf{O}$ 的元素, 它们的下标表示其在矩阵中的位置。此外, $b b_{mk}$ 被定义为:

$$b b_{mk} = [\mathbf{F}_{\text{BB}} \mathbf{F}_{\text{BB}}^H]_{mk} = \sum_{l=1}^{N_s} b_{ml} b_{kl}^* \quad (43)$$

此时, 可以直接给出函数 $f(\mathbf{x}_i)$ 的 Hessian 矩阵:

$$\mathbf{H}_i = \nabla_{\mathbf{x}_i} \mathbf{g}_i^T = \begin{bmatrix} \frac{\partial g_{1i}}{\partial x_{1i}} & \dots & \frac{\partial g_{N_{\text{RF}}i}}{\partial x_{1i}} \\ \vdots & \ddots & \vdots \\ \frac{\partial g_{1i}}{\partial x_{N_{\text{RF}}i}} & \dots & \frac{\partial g_{N_{\text{RF}}i}}{\partial x_{N_{\text{RF}}i}} \end{bmatrix} \quad (44)$$

其中,

$$h_{ml} = \frac{\partial g_{mi}}{\partial x_{li}} = 2\Re \begin{cases} b b_{mm} - \mathbf{e}^{j\mathbf{x}_{mi}} [\mathbf{Q}]_{mi}, m=l \\ b b_{ml} \mathbf{e}^{j\mathbf{x}_{mi}} \mathbf{e}^{-j\mathbf{x}_{li}}, m \neq l \end{cases} \quad (45)$$

参考文献:

- [1] XIAO Z Y, ZHU L P, LIU Y M, et al. A survey on millimeter-wave beamforming enabled UAV communications and networking[J]. IEEE Communications Surveys & Tutorials, 2022, 24(1): 557-610.
- [2] RANGAN S, RAPPAPORT T S, ERKIP E. Millimeter-wave cellular wireless networks: potentials and challenges[J]. Proceedings of the IEEE, 2014, 102(3): 366-385.
- [3] ROH W, SEOL J Y, PARK J, et al. Millimeter-wave beamforming as an enabling technology for 5G cellular communications: theoretical feasibility and prototype results[J]. IEEE Communi-



- cations Magazine, 2014, 52(2): 106-113.
- [4] EREZ U, SHAMAI S, ZAMIR R. Capacity and lattice strategies for canceling known interference[J]. IEEE Transactions on Information Theory, 2005, 51(11): 3820-3833.
- [5] SPENCER Q H, SWINDLEHURST A L, HAARDT M. Zero-forcing methods for downlink spatial multiplexing in multi-user MIMO channels[J]. IEEE Transactions on Signal Processing, 2004, 52(2): 461-471.
- [6] 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.
- [7] 廖勇, 杨馨怡, 杜洁汝. 基于两阶段的毫米波大规模 MIMO 低复杂度混合预编码算法[J]. 电子学报, 2021, 49(7): 1298-1304.
- LIAO Y, YANG X Y, DU J R. A two stage based low complexity hybrid precoding algorithm for millimeter wave massive MIMO[J]. Acta Electronica Sinica, 2021, 49(7): 1298-1304.
- [8] UWAECHEA A N, MAHYUDDIN N M, AIN M F, et al. On the spectral-efficiency of low-complexity and resolution hybrid precoding and combining transceivers for mmWave MIMO systems[J]. IEEE Access, 2019(7): 109259-109277.
- [9] HUNG W L, CHEN C H, LIAO C C, et al. Low-complexity hybrid precoding algorithm based on orthogonal beamforming codebook[C]//Proceedings of 2015 IEEE Workshop on Signal Processing Systems (SIPS). Piscataway: IEEE Press, 2015: 1-5.
- [10] LIU X R, LI X M, CAO S, et al. Hybrid precoding for massive mmWave MIMO systems[J]. IEEE Access, 2019(7): 33577-33586.
- [11] 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.
- [12] CHEN J C. Gradient projection-based alternating minimization algorithm for designing hybrid beamforming in millimeter-wave MIMO systems[J]. IEEE Communications Letters, 2019, 23(1): 112-115.
- [13] MULLA M, ULUSOY A H, RIZANER A, et al. Barzilai-borwein gradient algorithm based alternating minimization for single user millimeter wave systems[J]. IEEE Wireless Communications Letters, 2020, 9(4): 508-512.
- [14] LIANG L, XU W, DONG X D. Low-complexity hybrid precoding in massive multiuser MIMO systems[J]. IEEE Wireless Communications Letters, 2014, 3(6): 653-656.
- [15] ALKHATEEB A, LEUS G, HEATH R W. Limited feedback hybrid precoding for multi-user millimeter wave systems[J]. IEEE Transactions on Wireless Communications, 2015, 14(11): 6481-6494.
- [16] NGUYEN D H N, LE L B, LE-NGOC T. Hybrid MMSE precoding for mmWave multiuser MIMO systems[C]//Proceedings of 2016 IEEE International Conference on Communications (ICC). Piscataway: IEEE Press, 2016: 1-6.
- [17] SU X F, JIANG Y. Optimal zero-forcing hybrid downlink precoding for sum-rate maximization[J]. IEEE Wireless Communications Letters, 2022, 11(3): 463-467.
- [18] WANG S G, LI Z T, HE M Y, et al. A joint hybrid precoding/combining scheme based on equivalent channel for massive MIMO systems[J]. IEEE Journal on Selected Areas in Communications, 2022, 40(10): 2882-2893.
- [19] ZHANG C, SHEN S P, HAN Z X, et al. Analog beamforming using ESPAR for single-RF precoding systems[J]. IEEE Transactions on Wireless Communications, 2023, 22(7): 4387-4400.
- [20] 李雪飞. 基于电可控寄生单元阵列天线的研究与应用[D]. 成都: 电子科技大学, 2021.
- LI X F. Research and application of array antenna based on electrically controllable parasitic element[D]. Chengdu: University of Electronic Science and Technology of China, 2021.
- [21] 张煜, 张治, 董晓岱. 毫米波 MIMO 系统中基于自适应梯度算法的混合预编码[J]. 通信学报, 2021, 42(10): 95-105.
- ZHANG Y, ZHANG Z, DONG X D. Adaptive gradient algorithm for hybrid precoding in mmWave MIMO system[J]. Journal on Communications, 2021, 42(10): 95-105.
- [22] HU Q Y, CAI Y L, SHI Q J, et al. Iterative algorithm induced deep-unfolding neural networks: precoding design for multiuser MIMO systems[J]. IEEE Transactions on Wireless Communications, 2021, 20(2): 1394-1410.
- [23] 章坚武, 王路鑫, 孙玲芬, 等. 人工智能在 5G 系统中的应用综述[J]. 电信科学, 2021, 37(5): 14-31.
- ZHANG J W, WANG L X, SUN L F, et al. An survey on application of artificial intelligence in 5G system[J]. Telecommunications Science, 2021, 37(5): 14-31.
- [24] SAYEED A M. Deconstructing multiantenna fading channels[J]. IEEE Transactions on Signal Processing, 2002, 50(10): 2563-2579.
- [25] WU H T, YANG J F, CHEN F K. Source number estimators using transformed Gerschgorin radii[J]. IEEE Transactions on Signal Processing, 1995, 43(6): 1325-1333.

- [26] ZHANG E, HUANG C. On achieving optimal rate of digital precoder by RF-baseband codesign for MIMO systems[C]//Proceedings of 2014 IEEE 80th Vehicular Technology Conference (VTC2014-Fall). Piscataway: IEEE Press, 2014: 1-5.

[作者简介]



胡博（1996—），男，浙江工商大学信息与电子工程学院硕士生，主要研究方向为大规模 MIMO、混合预编码。



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、空时信号处理以及大规模 MIMO 系统等下一代无线通信核心技术。



魏贵义（1973—），男，博士，浙江工商大学信息与电子工程学院院长、教授，主要研究方向为无线通信、移动计算、云计算、社交网络、博弈论。