基于频谱效率公平性的 XL-MIMO 系统预编码优化

李志立 1,2，傅友华 1,2，宋云超 1,2

(1.南京邮电大学电子与光学工程学院柔性电子(未来技术)学院，南京 210023；2.南京邮电大学射频集成与微组装技术国家地方联合工作实验室，南京 210023)

摘要：本文研究了在近场信道模型下基于频谱效率公平性的超大规模多输入多输出下行系统的预编码优化问题。考虑在该近场信道模型，即小区内同时存在视距(Line-of-Sight，LOS)和非视距(Non-Line-of-Sight，NLOS)的非平稳混合信道，其中 LOS 信道采用球面波模型，而 NLOS 信道则采用瑞利模型。文中以频谱效率的几何平均值作为优化目标，从而确保用户间的公平性并优化系统整体的频谱效率。为了处理复杂的优化目标函数，首先对其采用泰勒展开的一阶近似作为新的目标函数；接着，**使用拉格朗日对偶变换和二次变换将原始优化问题转化为更容易求解的等价问题**；最后，为了降低计算复杂度，采用了快速迭代收缩阈值算法与投影梯度下降算法结合的投影快速迭代收缩阈值**算法**(Projection Fast Iterative Shrinkage Threshold Algorithm，PFISTA)来解决等效优化问题。仿真结果显示，以几何平均值作为目标函数能够降低用户频谱效率之间的差异，实现用户频谱效率的均衡提升。此外，PFISTA 在获得与现有方法相当性能的同时，**具有较低的**计算复杂度。

关键词：非平稳；球面波；快速迭代收缩阈值算法；投影梯度下降算法；频谱效率的几何平均值；拉格朗日对偶变换；二次变换

中图分类号：TN929.5　　**文献标识码：**A

Precoding optimization of XL-MIMO system based on spectral efficiency fairness

LI Zhili[1,2] , FU Youhua[1,2]and SONG Yunchao[1,2]

(1.College of Electronic and Optical Engineering and College of Flexible Electronics (Future Technology) , Nanjing University of Posts and Telecommunications, Nanjing 210003, China；
2.National and Local Joint Engineering Laboratory of RF Integration and Micro-Assembly Technology , Nanjing University of Posts and Telecommunications, Nanjing 210003, China)

Abstract:This paper studies the precoding optimization problem for Extremely Large-Scale Multiple-Input-Multiple-Output downlink systems under a near-field channel model based on spectral efficiency fairness. The near-field channel model considers the coexistence of line-of-sight (LOS) and non-line-of-sight (NLOS) non-stationary mixed channels within the cell, where LOS channels are modeled using spherical wave models, while NLOS channels are modeled using Rayleigh models. The geometric mean of spectral efficiency is used as the optimization target to ensure fairness among users and optimize the overall spectral efficiency of the system. To handle the complex optimization objective function, a first-order Taylor expansion approximation is applied to create a simplified objective function; subsequently, Lagrangian dual transformation and quadratic transformations are used to transform the original optimization problem into an equivalent one that is easier to solve. Finally, to reduce computational complexity, the Projection Fast Iterative Shrinkage Threshold Algorithm (PFISTA), which combines fast iterative shrinkage thresholding algorithms with projected gradient descent, is employed to solve the equivalent optimization problem. Simulation results show that using the geometric mean as the objective function can reduce differences in spectral efficiencies among users, leading to a balanced improvement in user spectral efficiencies. Moreover, PFISTA achieves comparable performance to existing methods **while maintaining lower computational complexity.**

Keyword : non stationary； spherical wave； fast iterative shrinkage threshold algorithm； projected gradient descent；geometric mean of spectral efficiency；Lagrangian dual transformation；quadratic transformation

引言

在**第五代(5G)和**第六代(6G)移动通信之后，为了满足日益苛刻的互联网服务要求，开发了许多先进的技术[1,2]。这些关键技术之一是在超大规模多输入多输出(Extremely Large-Scale Multiple-Input-Multiple-Output，XL-MIMO)系统中部署超大规模天线阵列(Extremely Large Antenna Array，ELAA)来提高频谱效率(Spectral Efficiency，SE)[3,4]。随着天线数量的增加，XL-MIMO 系统面临着三大挑战：计算复杂度、可扩展性**和**非平稳性[5,6]。特别是，由于信道的空间非平稳特性，不同部分可能会经历不同的**传播条件**[7]。幸运的是已有**文献**通过 COST 2100 信道模型测量**了**可视区域(Visibility Regions，VRs)，并提出了可视区域的分布公式[8,9]。

基金项目：国家自然科学基金青年项目(62101282)。

此外，由于阵列孔径的增大，瑞利距离**也随之**增加，**导致**小区用户处于近场区域[10]。**因此，在 XL-MIMO 信道建模中，不仅需要考虑非平稳性，还需采用近场球面波模型。**

虽然用户与天线元件距离较近时，信道以视距(Line-of-Sight，LOS)传播为主。但对于小区内用户，由于环境中散射体和障碍物的存在，假设所有无线电链路都经历 LOS 传输是不准确的[11]。**因此**，如文献[12]**所述**，应合理地假设**部分**无线电链路**处于** LOS 机制下，而其余链路**处于**非视距(Non-Line-of-Sight，NLOS)机制下。

针对这些新的特点，已有不少文献从功率分配(Power Allocation，PA)、天线选择(Antenna Selection，AS)、波束赋形和预编码等不同的优化目标出发，对 XL-MIMO 系统进行了研究[13]。例如：文献[11]提出了支持最低服务质量(Quality of Service，QoS)的联合用户调度和 PA 技术，以提高服务用户的数量。尽管该文献考虑了小区内同时存在 LOS 和 NLOS 用户的情况，但并未考虑近场非平稳性和用户公平性，仅优化了系统的总频谱效率。文献[13]考虑了用户公平性和近场球面波建模，使用频谱效率的几何平均值(Geometric Mean of Spectral Efficiency，GM SE)进行 PA 优化，但忽略了信道的非平稳性和小区内同时存在 LOS 和 NLOS 信道的情况。文献[14]研究了分布式基站信号处理架构下 XL-MIMO 系统中非平稳信道下下行链路 AS 和 PA 的联合问题，但在信道模型中忽略了近场球面波、用户公平性和小区内同时存在 LOS 和 NLOS 信道的情况，仅优化了频谱效率的总和。文献[15]考虑了用户公平性，并使用 GM SE 作为目标函数对波束进行优化，但与上述文献不同，文献[15]研究的不是 XL-MIMO 系统而是可重构智能表面辅助通信系统的波束赋形，且未考虑小区内同时存在 LOS 和 NLOS 用户的情况，以及近场球面波和非平稳性。此外，除了文献[15]，其他文献主要关注功率分配优化，而未涉及预编码矩阵的优化。

本文针对多用户 XL-MIMO 通信考虑了一种基于球面波传播的非平稳多状态信道模型，该信道模型考虑到 LOS 和 NLOS 传输下的用户共存于同一通信小区[11]。**关注**到用户公平性，**采**用 GM SE 作为目标函数[15]**并提出复杂度较低的优化算法**。GM SE 优化问题最初使用泰勒展开一阶近似，随后使用拉格朗日对偶变换和二次变换，将其转化为易于求解的整式和问题[16]，最终为了降低**计算**复杂度使用投影梯度下降法(Projected Gradient Descent，PGD)和快速迭代收缩阈值算法(Fast Iterative Shrinkage Thresholding Algorithm，FISTA)相结合的投影快速迭代收缩阈值算法(Projection Fast Iterative Shrinkage Threshold Algorithm，PFISTA)[17,18]优化这一整式和问题。

1 系统模型与问题求解

1.1 信道模型

本文研究了单个小区内 XL-MIMO 系统的下行链路传输情况。其中该小区包含一个配备有 M 个天线的均匀线性阵列(Uniform Linear Array，ULA)的基站(Base Station，BS)，以及 K 个单天线用户。由于 XL-MIMO 信道模型是球面波模型，并且 LOS 和 NLOS 信道同处于一个小区内。所以定义了两个信道向量，一个用于 LOS 信道模型，另一个用于 NLOS 信道模型[11]。

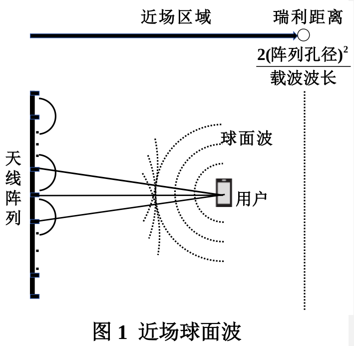


图 1 近场球面波

Fig. 1 Near field spherical wave

在 LOS 球面波信道模型中，基站天线和用户 k 之间的信道响应为向量 $\boldsymbol{a}_k^{\mathrm{LOS}} \in \mathrm{X}^{M\times 1}$，可表示为 $\boldsymbol{a}_k^{\mathrm{LOS}} = [a_{1,k},\ldots,a_{m,k},\ldots,a_{M,k}]^{\mathrm{T}}$，其元素[11]：

$$a_{m,k} = \sqrt{\beta_{m,k}^{\mathrm{LOS}}}\exp(-\mathrm{j}\frac{2\pi}{\lambda}r_{m,k}) \tag{1}$$

其中 $r_{m,k}$ 是天线 m 和用户 k 的距离，$\beta_{m,k}^{\mathrm{LOS}} = \Omega^{\mathrm{LOS}}(r_{m,k}/r_{\mathrm{h}})^{-v^{\mathrm{LOS}}}$ 是路径损耗，r_{h} 是参考距离，Ω^{LOS}是根据参考距离选取的路径损耗系数，v^{LOS}是路径损耗指数，$\lambda > 0$ 是载波波长[19]。

除此之外，NLOS 信道遵循独立同分布的瑞利衰落模型，即：

$$\boldsymbol{a}_k^{\mathrm{NLOS}} \sim \mathrm{XN}(0_M, \boldsymbol{\Sigma}_k) \tag{2}$$

其中 $\boldsymbol{\Sigma}_k = \mathrm{diag}([\beta_{1,k}^{\mathrm{NLOS}},\ldots,\beta_{m,k}^{\mathrm{NLOS}},\ldots\beta_{M,k}^{\mathrm{NLOS}}]^{\mathrm{T}})$ 是对角协方差矩阵，$\beta_{m,k}^{\mathrm{NLOS}} = \Omega^{\mathrm{NLOS}}(r_{m,k}/r_{\mathrm{h}})^{-v^{\mathrm{NLOS}}}$ 是路径损耗，Ω^{NLOS}是根据参考距离选取的路径损耗系数，v^{NLOS}是路径损耗指数。

设 $x_k \in \{0,1\}$ 是与用户 k 相关联的信道状态指示符，如果信道处于 LOS 状态，则 x_k 等于 1，或者如果信道处于 NLOS 状态，则 x_k 等于 0。为了简化分析，与文献[11]一致，假定信道状态指示符遵循参数 $0 \le \rho \le 1$ 的伯努利随机分布，即 ρ 为 LOS 概率，x_k 取值为 1 的概率为 ρ，取值为 0 的概率为 $1-\rho$。给定信道状态指示符的定义以及 LOS 和 NLOS 信道向量，多状态信道向量 $\boldsymbol{a}_k \in \mathrm{X}^{M\times 1}$ 可以被定义为：

$$\boldsymbol{a}_k = x_k \boldsymbol{a}_k^{\mathrm{LOS}} + (1-x_k)\boldsymbol{a}_k^{\mathrm{NLOS}} \tag{3}$$

除了考虑到 LOS 和 NLOS 路径同时存在的情况，还应该考虑到 XL-MIMO 信道的空间非平稳性，所以将第 k 个用户与基站之间的信道 $\boldsymbol{h}_k \in \mathrm{X}^{M\times 1}$ 建模为[20]:

$$\boldsymbol{h}_k = \boldsymbol{D}_k \boldsymbol{a}_k \tag{4}$$

其中 $\boldsymbol{D}_k \in \{0,1\}^{M\times M}$ 是通过 VRs 概念对信道非平稳特性进行建模的对角指示矩阵。采用文献[9]中描述的信道模型，其中每个用户的特征在于它们 VR 的两个属性：中心和长度。VR 的中心被建模为均匀的随机变量 $c_k \sim \mathrm{Y}(0,L)$，其中 L 表示 XL-MIMO 天线阵列的物理长度。而 VRs 的长度是 $l_k \sim \Lambda\mathrm{N}(\mu_l,\sigma_l)$。由于实际场景的多样性，LOS 和 NLOS 的 VRs 长度的参数可能有所不同。

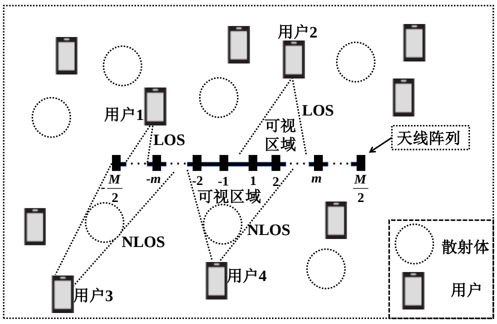


图 2 同处于一个小区内具有不同 VRs 的 LOS 和 NLOS 信道

Fig. 2 LOS and NLOS channels with different VRs within the same cell

1.2 优化问题

假设 $\boldsymbol{s} \in \mathrm{X}^{K\times 1} \sim \mathrm{XN}(0,\boldsymbol{I})$ 是 BS 发射给 K 个用户的数据信号，$\boldsymbol{p} = [p_1,\ldots,p_k,\ldots,p_K]^{\mathrm{T}}$ 是 K 个用户的数据信号功率。通过预编码矩阵 $\boldsymbol{V} \in \mathrm{X}^{M\times K}$ 进行预编码，BS 发射给用户的信号 $\boldsymbol{z}$ 为[13]:

$$\boldsymbol{z} = \boldsymbol{V}\,\mathrm{diag}^{1/2}\{\boldsymbol{p}\}\boldsymbol{s} \tag{5}$$

因此，第 k 个用户的接收信号可以表示为：

$$y_k = \boldsymbol{h}_k^{\mathrm{H}}\boldsymbol{w}_k s_k + \sum_{\substack{i=1\\ i\ne k}}^{K}\boldsymbol{h}_k^{\mathrm{H}}\boldsymbol{w}_i s_i + b_k \tag{6}$$

其中令 $\boldsymbol{W} = \boldsymbol{V}\,\mathrm{diag}^{1/2}\{\boldsymbol{p}\}$，它是满足总功率约束的预编码矩阵，并且 $\boldsymbol{W} = [\boldsymbol{w}_1,\ldots,\boldsymbol{w}_k,\ldots,\boldsymbol{w}_K]$，其列向量表示 BS 对每个用户发射信号的功率约束下的预编码向量，$b_k \sim \mathrm{XN}(0,\sigma^2)$ 是第 k 个用户的加性高斯白噪声。

根据给定的接收信号式(6)，第 k 个用户的信干噪比(Signal-to-Interference-plus-Noise Ratio，SINR)可以表示为：

$$\gamma_k(\boldsymbol{W}) = \frac{|\boldsymbol{h}_k^{\mathrm{H}}\boldsymbol{w}_k|^2}{\sum_{\substack{i=1\\ i\ne k}}^{K}|\boldsymbol{h}_k^{\mathrm{H}}\boldsymbol{w}_i|^2 + \sigma^2} \tag{7}$$

所以，第 k 个用户的 SE 可以表示为：

$$G_k(\boldsymbol{W}) = \log_2(1+\gamma_k(\boldsymbol{W})) \tag{8}$$

为了在 SE 和用户公平性之间取得平衡，选择用户 GM SE 作为主要优化目标。通过考虑预编码的总功率约束，优化问题可以表示为[15]:

$$\begin{aligned} &\max_{\boldsymbol{W}}\ R(\boldsymbol{W}) = [\prod_{k=1}^{K} G_k(\boldsymbol{W})]^{\frac{1}{K}} \\ &\mathrm{s.t.}\ \sum_{k=1}^{K}\|\boldsymbol{w}_k\|^2 \le P_{\max} \end{aligned} \tag{9}$$

其中 $P_{\max} = \|\boldsymbol{p}\|_1$。

2 优化算法

为了最大化用户 GM SE，可以将优化问题式(9)中最大化目标函数 $R(\boldsymbol{W})$ 变换为最小化 $g(\boldsymbol{W})=[\prod_{k=1}^{K}G_k(\boldsymbol{W})]^{\frac{-1}{K}}$ [13]。由 $g(\boldsymbol{W})$ 对 $G_k(\boldsymbol{W})$ 求二阶偏导，得 $\frac{\partial^2 g(\boldsymbol{W})}{\partial(G_k(\boldsymbol{W}))^2}=\frac{g(\boldsymbol{W})}{K^2(G_k(\boldsymbol{W}))^2}(1+K)$，可知 $g(\boldsymbol{W})$ 关于 $G_k(\boldsymbol{W})$ 是凸函数。然而，求得最优 $\boldsymbol{W}$ 存在较大难度。为了降低优化难度，参考文献[13]，考虑新目标函数 $g(\boldsymbol{W})$ 的泰勒展开近似：

$$
\begin{aligned}
g(\boldsymbol{W})&=g(\boldsymbol{W}^t)+\sum_{k=1}^{K}\frac{\partial g(\boldsymbol{W})}{\partial G_k(\boldsymbol{W})}|_{\boldsymbol{W}=\boldsymbol{W}^t}[G_k(\boldsymbol{W})-G_k(\boldsymbol{W}^t)]+o(G_k(\boldsymbol{W})-G_k(\boldsymbol{W}^t))\\
&\approx 2g(\boldsymbol{W}^t)-g(\boldsymbol{W}^t)\frac{1}{K}\sum_{k=1}^{K}\frac{G_k(\boldsymbol{W})}{G_k(\boldsymbol{W}^t)}
\end{aligned}
\tag{10}
$$

其中 $\boldsymbol{W}^t$ 是 $\boldsymbol{W}$ 经过 t 次迭代产生的值。求 $g(\boldsymbol{W})$ 对 $G_k(\boldsymbol{W})$ 的多阶偏导得:

$$
\frac{\partial^J g(\boldsymbol{W})}{\partial(G_k(\boldsymbol{W}))^J}=g(\boldsymbol{W})\prod_{j=1}^{J}\left(\frac{K(j-1)+1}{-KG_k(\boldsymbol{W})}\right)
\tag{11}
$$

其中 J 为偏导阶数。根据式(11)得到函数 $g(\boldsymbol{W})$ 在 $G_k(\boldsymbol{W}^t)$ 处泰勒展开的高阶项为: $g(\boldsymbol{W}^t)\left(\frac{G_k(\boldsymbol{W})-G_k(\boldsymbol{W}^t)}{G_k(\boldsymbol{W}^t)}\right)^J\prod_{j=1}^{J}\left(\frac{K(j-1)+1}{-Kj}\right)$。因为 $\left|\frac{G_k(\boldsymbol{W})-G_k(\boldsymbol{W}^t)}{G_k(\boldsymbol{W}^t)}\right|$ 和 $\left|\frac{K(j-1)+1}{-Kj}\right|$ 在迭代过程中是小于 1 的，故高阶项趋近于零，所以目标函数 $g(\boldsymbol{W})$ 可以使用一阶泰勒展开近似，即式(10)成立。

进一步，由于 $g(\boldsymbol{W}^t)>0$，最小化式(10)中的目标函数 $g(\boldsymbol{W})$ 可简化为最大化其一阶近似中的减数项，并代入第 k 个用户的 SE 表达式(8)，得到优化问题式(9)的等价优化问题为：

$$
\begin{aligned}
&\max_{\boldsymbol{W}}\ \ f_1(\boldsymbol{W})=\sum_{k=1}^{K}\frac{G_k(\boldsymbol{W})}{G_k(\boldsymbol{W}^t)}=\sum_{k=1}^{K}\alpha_k\log_2(1+\gamma_k(\boldsymbol{W}))\\
&\text{s.t.}\ \sum_{k=1}^{K}\|\boldsymbol{w}_k\|^2\le P_{\max}
\end{aligned}
\tag{12}
$$

其中，由于 $G_k(\boldsymbol{W}^t)$ 在每次迭代中相对不变，令 $\alpha_k=\frac{1}{G_k(\boldsymbol{W}^t)}$，引入了参数 $\boldsymbol{\alpha}=[\alpha_1,\dots,\alpha_k,\dots,\alpha_K]^{\mathrm{T}}\in \mathrm{P}^{K\times 1}$。

这样，考虑用户公平的 SE 优化问题就转换为式(12)这种类似于加权和速率的优化问题。在处理优化问题式(12)时，尽管文献[15]对类似形式的优化问题提出下降算法(Descent Algorithm，DA)作为解决方案，**但是**，该算法的计算复杂度相对较高。鉴于此，采用了一种新的优化算法，旨在降低计算复杂度。具体而言，首先，**针对**优化问题式(12)**中**目标函数的**的复杂形式**，即**包含**分式的对数求和，使用拉格朗日对偶变换[16,21]将其转化为**相对易于求解的**分式和问题；然后对转化后的分式和问题使用二次变换[16]，将其转化为易于求解的整式和问题；最后基于求解**该**整式和问题采用 PFISTA[17,18]降低计算复杂度。

2.1 优化问题的等效变换

对优化问题式(12)，首先引入辅助变量 δ_k，从目标函数 f_1 中提取与 $\boldsymbol{W}$ 相关的分式项 $\gamma_k(\boldsymbol{W})$，对优化问题式(12)应用拉格朗日对偶变换[16,21]，

$$
\begin{aligned}
&\max_{\boldsymbol{W}}\ \ f_2(\boldsymbol{W},\boldsymbol{\delta})=\sum_{k=1}^{K}\frac{1}{\ln 2}\frac{\alpha_k(1+\delta_k)\gamma_k(\boldsymbol{W})}{1+\gamma_k(\boldsymbol{W})}+\sum_{k=1}^{K}\alpha_k\log_2(1+\delta_k)-\frac{1}{\ln 2}\sum_{k=1}^{K}\alpha_k\delta_k\\
&\text{s.t.}\ \sum_{k=1}^{K}\|\boldsymbol{w}_k\|^2\le P_{\max}
\end{aligned}
\tag{13}
$$

其中 $\boldsymbol{\delta}=[\delta_1,\dots,\delta_k,\dots,\delta_K]^{\mathrm{T}}$，当优化问题式(12)的解 $\boldsymbol{W}$ 也是上式(13)的唯一解时，这两个优化问题等价，且它们的最优解也是相同的。可以验证：当 $\boldsymbol{W}$ 固定时，f_2 关于 $\boldsymbol{\delta}$ 是可微的凹函数，所以可以通过 $\frac{\partial f_2(\boldsymbol{W},\boldsymbol{\delta})}{\partial\delta_k}=0$ 确定最优 $\boldsymbol{\delta}$ 为[22]：

$$\delta_k^{\mathrm{opt}}=\gamma_k(\boldsymbol{W})=\frac{|\boldsymbol{h}_k^{\mathrm{H}}\boldsymbol{w}_k|^2}{\sum_{\substack{i\neq k\\i=1}}^{K}|\boldsymbol{h}_k^{\mathrm{H}}\boldsymbol{w}_i|^2+\sigma^2},k=1,\dots,K \tag{14}$$

将上式代入优化目标函数 f_2，此时 f_2 即等于原优化问题式(12)中的目标函数 f_1。于是原目标函数 f_1 中分式的对数和形式通过拉格朗日对偶变换转化为 f_2 的分式和形式。

那么考虑等价的优化问题式(13)，已知 $\boldsymbol{\delta}$ 最优解式(14)，固定 $\boldsymbol{\delta}$ 优化 $\boldsymbol{W}$。目标函数 f_2 中只有第一项与待求的 $\boldsymbol{W}$ 有关，于是代入 SINR 表达式(7)可得：

$$f_3(\boldsymbol{W})=\frac{1}{\ln 2}\sum_{k=1}^{K}\frac{\alpha_k(1+\delta_k)|\boldsymbol{h}_k^{\mathrm{H}}\boldsymbol{w}_k|^2}{\sum_{i=1}^{K}|\boldsymbol{h}_k^{\mathrm{H}}\boldsymbol{w}_i|^2+\sigma^2}+\mathrm{const}(\boldsymbol{\delta}) \tag{15}$$

其中 $\mathrm{const}(\boldsymbol{\delta})$ 是关于 $\boldsymbol{\delta}$ 的常数项。可以舍掉其常数项 $\mathrm{const}(\boldsymbol{\delta})$ 和系数 $\frac{1}{\ln 2}$ 优化 $\boldsymbol{W}$。对于目标函数 f_3 这种复杂分式和的优化问题，应考虑使用二次变换将其转化为易于求解的整式和问题[16]。引入辅助变量 ξ_k，对优化目标函数 f_3 应用二次变换，优化问题式(13)进一步等价为：

$$\begin{aligned}&\max_{\boldsymbol{W}}\ f_4(\boldsymbol{W},\boldsymbol{\xi})=\sum_{k=1}^{K}2\sqrt{\alpha_k(1+\delta_k)}\,\mathrm{Re}\{\xi_k^{\mathrm{H}}\boldsymbol{h}_k^{\mathrm{H}}\boldsymbol{w}_k\}-\sum_{k=1}^{K}\xi_k^{\mathrm{H}}\xi_k\left(\sum_{i=1}^{K}|\boldsymbol{h}_k^{\mathrm{H}}\boldsymbol{w}_i|^2\right)-\sigma^2\sum_{k=1}^{K}\xi_k^{\mathrm{H}}\xi_k\\&\mathrm{s.t.}\ \sum_{k=1}^{K}\|\boldsymbol{w}_k\|^2=\mathrm{Tr}(\boldsymbol{W}^{\mathrm{H}}\boldsymbol{W})\le P_{\max}\end{aligned} \tag{16}$$

其中 $\boldsymbol{\xi}=[\xi_1,\dots,\xi_k,\dots,\xi_K]^{\mathrm{T}}$，与上述拉格朗日对偶变换的情况类似，验证方式也相同。通过先固定 $\boldsymbol{W}$，然后令 $\frac{\partial f_4(\boldsymbol{W},\boldsymbol{\xi})}{\partial\xi_k}=0$ 确定最优 $\boldsymbol{\xi}$ 为：

$$\xi_k^{\mathrm{opt}}=\frac{\sqrt{\alpha_k(1+\delta_k)}\boldsymbol{h}_k^{\mathrm{H}}\boldsymbol{w}_k}{\sum_{i=1}^{k}|\boldsymbol{h}_k^{\mathrm{H}}\boldsymbol{w}_i|^2+\sigma^2},k=1,\dots,K \tag{17}$$

将上式代入优化目标函数 f_4，此时 f_4 即等于目标函数 f_3。于是目标函数 f_3 的分式和形式通过二次变换转化为 f_4 的整式和形式。

2.2 功率约束预编码 W 的优化

上一节对分式对数和形式的优化问题式(12)进行一些变换后得到简化的等价优化问题式(16)。现在求功率约束预编码 $\boldsymbol{W}$ 的最优解，首先固定 $\boldsymbol{\xi}$，然后通过 f_4 对 $\boldsymbol{w}_k$ 求二阶偏导，得 $\frac{\partial^2 f_4(\boldsymbol{W},\boldsymbol{\xi})}{\partial(\boldsymbol{w}_k)^2}=-2\left(\sum_{i=1}^{K}\xi_i^{\mathrm{H}}\xi_i\boldsymbol{h}_i\boldsymbol{h}_i^{\mathrm{H}}\right)$，发现其小于零，可知 f_4 关于 $\boldsymbol{w}_k$ 是凹函数。且可行域是凸集。所以根据卡罗需-库恩-塔克(Karush-Kuhn-Tucker，KKT)条件，得到 $\boldsymbol{w}_k$ 的最优解为：

$$\boldsymbol{w}_k^{\mathrm{opt}}=\sqrt{\alpha_k(1+\delta_k)}\xi_k\left(\sum_{i=1}^{K}\xi_i^{\mathrm{H}}\xi_i\boldsymbol{h}_i\boldsymbol{h}_i^{\mathrm{H}}+\mu\boldsymbol{I}_M\right)^{-1}\boldsymbol{h}_k \tag{18}$$

其中 μ 是引入的拉格朗日乘子，可以通过对分搜索确定[16]。到此已得到了优化问题的优化解，然而计算式(18)中矩阵求逆的复杂度为 $O(M^3)$，对于 M 较大的情况下，成本较高，并且，搜索 μ 以获得最优 $\boldsymbol{W}$ 也是耗时的[23]。所以为了降低复杂度，考虑使用 PGD[17]获得 $\boldsymbol{W}$ 的迭代公式为：$\widetilde{\boldsymbol{W}}^{t}=\boldsymbol{W}^{t-1}+\varphi\nabla f_4(\boldsymbol{W}^{t-1})$，其中 φ 为迭代步长，梯度 $\nabla f_4(\boldsymbol{W}^{t-1})=[\nabla f_4(\boldsymbol{w}_1^{t-1}),\dots,\nabla f_4(\boldsymbol{w}_k^{t-1}),\dots,\nabla f_4(\boldsymbol{w}_K^{t-1})]$。**由于 PGD 通常用于求解最小化问题，而式(16)是最大化问题，故需对目标函数取负，迭代式中的负号改为正号。**然而考虑到此迭代的收敛速率相对较低，通常仅为次线性的，具体表现为 $O(1/t)$，因此使用 FISTA[18]，使其收敛速率变为 $O(1/t^2)$。具体地，投影梯度为：

$$\widetilde{\boldsymbol{W}}^{t}=\boldsymbol{W}^{t-1}+\frac{d_{t-1}-1}{d_t}(\boldsymbol{W}^{t-1}-\boldsymbol{W}^{t-2})+\varphi\nabla f_4(\boldsymbol{W}^{t-1}+\frac{d_{t-1}-1}{d_t}(\boldsymbol{W}^{t-1}-\boldsymbol{W}^{t-2})) \tag{19a}$$

$$\nabla f_4(\boldsymbol{w}_k^{t-1})=\frac{\partial f_4(\boldsymbol{W})}{\partial \boldsymbol{w}_k}\Big|_{\boldsymbol{w}_k=\boldsymbol{w}_k^{t-1}}=2\sqrt{\alpha_k(1+\delta_k)}\boldsymbol{h}_k\xi_k-2\left(\sum_{i=1}^{K}\xi_i^{\mathrm{H}}\xi_i\boldsymbol{h}_i\boldsymbol{h}_i^{\mathrm{H}}\right)\boldsymbol{w}_k^{t-1} \tag{19b}$$

$$\boldsymbol{W}^t=\prod\nolimits_{\boldsymbol{C}}\left\{\tilde{\boldsymbol{W}}^t\right\}=\begin{cases}\tilde{\boldsymbol{W}}^t,\ \text{if}\ \ \mathrm{Tr}(\tilde{\boldsymbol{W}}^{t\,\mathrm{H}}\tilde{\boldsymbol{W}}^t)\le P_{\max}\\ \dfrac{\tilde{\boldsymbol{W}}^t}{\|\tilde{\boldsymbol{W}}^t\|_{\mathrm{F}}}\sqrt{P_{\max}},\ \text{otherelse}\end{cases} \tag{19c}$$

其中，式(19a)是 FISTA 迭代公式，式(19b)是 $\boldsymbol{w}_k=\boldsymbol{w}_k^{t-1}$ 时 $f_4(\boldsymbol{W})$ 相对于 $\boldsymbol{w}_k$ 的梯度公式，式(19c)是投影算法公式，t 是迭代次数，$\boldsymbol{W}$ 的可行域 $\boldsymbol{C}=\left\{\boldsymbol{W}\mid \mathrm{Tr}(\boldsymbol{W}^{\mathrm{H}}\boldsymbol{W})\le P_{\max}\right\}$，为了降低计算复杂度取迭代步长为固定步长 $\varphi=\dfrac{1}{2\|\sum_{i=1}^{K}\xi_i^{\mathrm{H}}\xi_i\boldsymbol{h}_i\boldsymbol{h}_i^{\mathrm{H}}\|_{\mathrm{F}}}$ [23,24]。式(19a)中，$d_0=1$，$d_t=\dfrac{1+\sqrt{1+4d_{t-1}^2}}{2},t=1,2,\dots$。而 $\nabla f_4(\boldsymbol{W}^{t-1}+\dfrac{d_{t-1}-1}{d_t}(\boldsymbol{W}^{t-1}-\boldsymbol{W}^{t-2}))$，可通过式(19b)将 $\boldsymbol{w}_k^{t-1}$ 换为 $\boldsymbol{w}_k^{t-1}+\dfrac{d_{t-1}-1}{d_t}(\boldsymbol{w}_k^{t-1}-\boldsymbol{w}_k^{t-2})$ 得到。

这样，将结合 PGD 和 FISTA 的算法称为 PFISTA[17,18]，即式(19)**的迭代求解过程**。

总之，基于泰勒展开一阶近似、拉格朗日对偶变换和二次变换的 PFISTA 已经为优化问题中需要优化的所有变量提供了闭式的解。该优化算法总结在算法 1 中。

算法 1 基于最大化 GM SE 的 PFISTA 迭代算法

输入：($\boldsymbol{H}$，$P_{\max}$，K，M，**循环最大值** $T_{\max}$，σ^2，**收敛精度** τ)
输出：$\boldsymbol{W}$
初始化：$\boldsymbol{W}_0$
1.循环，当 $t<T_{\max}$ 时；
2.根据式 $\alpha_k=G_k^{-1}(\boldsymbol{W}^t)$ 更新 $\boldsymbol{\alpha}$；
3.根据式(14)更新 $\boldsymbol{\delta}$；
4.根据式(17)更新 $\boldsymbol{\xi}$；
5.根据式(19)更新 $\boldsymbol{W}$；
6.根据式(9)判断 $|R(\boldsymbol{W}^t)-R(\boldsymbol{W}^{t-1})|\le\tau$ 是否成立，是结束循环，输出 $\boldsymbol{W}$，否转 1 且 t=t+1；

基于泰勒展开一阶近似、拉格朗日对偶变换和二次变换的 PFISTA 计算成本包括：变量 $\boldsymbol{\alpha}$、式(14)中的变量 $\boldsymbol{\delta}$、式(17)中的变量 $\boldsymbol{\xi}$ 和式(19)中的变量 $\boldsymbol{W}$ 的计算。其中 $\boldsymbol{\alpha}$、$\boldsymbol{\delta}$ 和 $\boldsymbol{\xi}$ 的计算复杂度都为 $O(K^2M)$。而 $\boldsymbol{W}$ 的计算复杂度主要取决于式(19b)的梯度计算，所以 $\boldsymbol{W}$ 的计算复杂度为 $O(KM^2)$。因此 PFISTA 的计算复杂度是 $O(I(3K^2M+KM^2))$，其中 I 为迭代次数。

3 仿真结果

假设用户均匀分布在范围为 $0.1\times0.1\mathrm{km}^2$ 的正方形小区内，其中用户到基站的最小距离设定为 30m。基站配备了一个超大规模天线阵列，其采用了 ULA 配置[25,26]。在该阵列中，各个天线元件之间的间距被精确设定为波长的一半。系统所使用的载波频率为 2.6 GHz[9]。此外，根据文献[9]、[11]、[19]、[25]和[26]的研究，在表 1 中汇总了仿真的参数配置。为了评估**所提**算法的性能，**文中**将如下算法进行比较：

带格式的：字体：(默认) Times New Roman，(中文) 宋体

表 1 仿真参数

Table 1 Simulation parameters

带格式表格

参数	参[illegible]	
M	基站天线数	**[72,128]**
K	用户数	**[4,8,12,16,20]**
r_{h}	参考距离	30m
ρ	LOS 概率	[0,0.25,0.5,0.75,1]
L	天线阵列长度	**[4.1510,7.3795]** m
σ^2	噪声方差	0dBm
c_k^{LOS}	LOS 的 VRs 中心	$\Upsilon(0,L)$
l_k^{LOS}	LOS 的 VRs 长度	$\Lambda\mathrm{N}\ (0.1L,0.1)$
c_k^{NLOS}	NLOS 的 VRs 中心	$\Upsilon(0,L)$
l_k^{NLOS}	NLOS 的 VRs 长度	$\Lambda\mathrm{N}\ (0.05L,0.3)$
Ω^{LOS}	LOS 路径损耗系数	0.214
v^{LOS}	LOS 路径损耗指数	2.20
Ω^{NLOS}	NLOS 路径损耗系数	0.220
v^{NLOS}	NLOS 路径损耗指数	3.80

其中算法(1)-(3)的目标函数为频谱效率的几何平均值(GM SE)，算法(4)的目标函数为总频谱效率(Sum Spectral Efficiency，Sum SE)

(1)PFISTA：结合泰勒展开一阶近似、拉格朗日对偶变换及二次变换的 PFISTA；

(2)PGD：此算法与(1)的处理流程相同但最后使用 PGD 进行优化；

(3)DA：文献**[15]中提出的一种基于泰勒展开一阶近似、不等式和拉格朗日乘子法的 DA 算法；**

(4)Sum-PGD[17,27]**：除优化过程不使用一阶泰勒展开外，其余部分与算法(2)相同。**

表 2 列出了基站天线数 M=128、用户数 K=16、视距(LOS)概率 $\rho=0.75$、信噪比(Signal-to-Noise Ratio，SNR)=20dB 和收敛精度 $\tau=10^{-3}$ 时所提 PFISTA 算法(1)与现有相关算法(2)-(4)的计算复杂度比较。其中 I_1 是外部迭代次数，I_2 为内部迭代次数。

图 3 展示了**在与表 2 相同条件下，不同算法的收敛曲线。结果显示，所有采用的优化算法均能实现收敛，但所需迭代次数存在差异。特别是在目标函数为频谱效

率的几何平均值(GM SE)的情况下，本文提出的 PFISTA 算法相较于 PGD 算法展现出较为显著的优势，仅需约 20 次迭代即可收敛，这一数值略高于 DA 算法。结合图 3 与表 2 的数据来看，尽管 DA 算法的外部迭代次数最少且性能几乎与 PFISTA 算法和 PGD 算法相当，但是它总计算复杂度最高。文献[15]中的 DA 算法复杂度较高，这可能与求取 ***W*** 时采用的拉格朗日乘子法有关。该方法涉及计算代价较高的矩阵求逆运算，并且由于 ***W*** 与拉格朗日乘子相互耦合，DA 算法采用了对分搜索法来迭代求解拉格朗日乘子，这进一步增加了算法的计算复杂度。此外，从图 3 还可以看出，与其他三种算法相比，第四种算法(Sum-PGD)在 GM SE 性能上明显处于劣势。

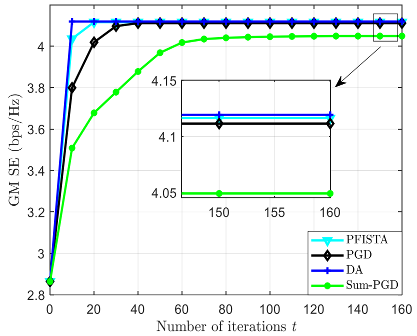


图 3 GM SE 与迭代次数 *t* 的关系

Fig. 3 The relationship between GM SE and the number of iterations *t*

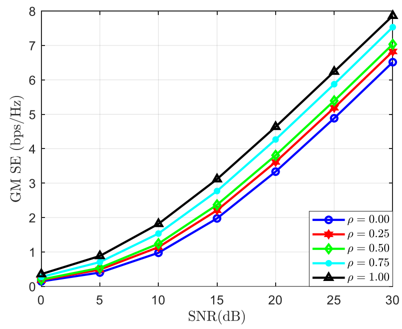


图 4 不同 LOS 概率下 SNR 和 GM SE 的关系

Fig. 4 The relationship between SNR and GM SE under different LOS probabilities

图 4 展示了**在**基站天线数 M=128、用户数 K=16 **和收敛精度** $\tau = 10^{-3}$ 的情况下，采用 **PFISTA 算法**在不同**视距(LOS)**概率下**信噪比(SNR)与频谱效率的几何平均值(GM SE)**的关系。从图中可以观察到，不论 LOS 概率如何变化，GM SE 都随着 SNR 的增加而**提高，且**随着 LOS 的概率增加，其 GM SE 也随之增加。

表 2 不同算法的计算复杂度比较

Table 2 Comparison of computational complexity of different algorithms

序号	算法	算法复杂度 $O(.)$	外部或内部迭代次数	总的复杂度
(1)	**PFISTA**	$O(I_1(3K^2M + KM^2))$	$I_1 = 20$	7.20896 E+6
(2)	**PGD**	$O(I_1(3K^2M + KM^2))$	$I_1 = \mathbf{40}$	**1.441792E+7**
(3)	**DA**	$O(I_1(K^2M + 2KM^2 + M^3 + I_2M^3))$	$I_1 = 10$，$I_2 = \mathbf{21}$	**4.66944E+8**
(4)	**Sum-PGD**	$O(I_1(2K^2M + KM^2))$	$I_1 = \mathbf{130}$	**4.25984E+7**

图 5 展示了在**信噪比 SNR=20dB**、**视距**(LOS)概率 $\rho = 0.75$、基站天线数 M=128、用户数 K=16 **和收敛精度** $\tau = 10^{-3}$ **的**情况下，使用迭代算法 PFISTA 在不同目标函数**下**每个用户的**频谱效率(SE)分布**。右图所采用的目标函数为考虑公平性的式(9)，而左图则为不考虑公平性的总频谱效率(Sum SE)，**即右图为算法 PFISTA，左图为算法 Sum-PGD 将 PGD 换为 PFISTA 的算法**。从图 5 中可以清楚地看出，相比不考虑公平性的情况，考虑公平性会使得每个用户之间的 SE 值差距减小。**同时，**考虑公平性时用户 SE 的**平**均值与不考虑公平性的用户 SE 的**平**均值几乎相当。

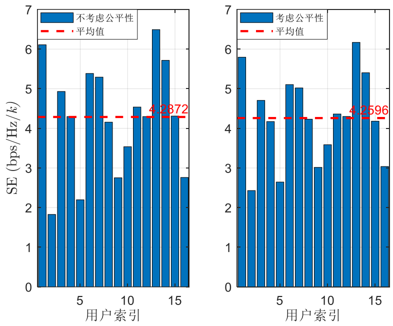


图 5 不同目标函数的每个用户的 SE

Fig. 5 SE for each user with different objective functions

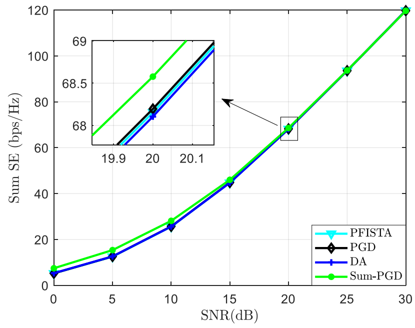


图 6 SNR 与 Sum SE 的关系

Fig. 6 The relationship between SNR and Sum SE

图 6 展示了在基站天线数 M=128、用户数 K=16、视距(LOS)概率 $\rho=0.75$ 和收敛精度 $\tau=10^{-3}$ 的条件下，信噪比(SNR)与总频谱效率(Sum SE)之间的关系。结果显示，随着 SNR 的增加，所有算法的 Sum SE 都呈现增加趋势。且当 $\text{SNR}\geq 20\,\text{dB}$ 时，以频谱效率的几何平均值(GM SE)作为目标函数的算法在 Sum SE 性能上与以 Sum SE 作为目标函数的算法差距较小。此外，不论 SNR 如何变化，以 GM SE 作为目标函数的三种算法的 Sum SE 性能都基本相同。

图 7 展示了在视距(LOS)概率 $\rho=0.75$、信噪比 SNR=20dB 和收敛精度 $\tau=10^{-3}$ 的条件下，用户数 K 与总频谱效率(Sum SE)之间的关系。结果显示，随着 K 的增加，所有算法的 Sum SE 都呈现上升趋势。特别地，当用户数 $K\leq 12$ 时，以频谱效率的几何平均值(GM SE)作为目标函数的算法相比于直接以 Sum SE 作为目标函数的算法，Sum SE 性能损失较小。此外，M=128 的场景性能优于 M=72 场景性能。进一步观察发现，无论 M 和 K 如何变化 PFISTA 算法与 PGD 算法和 DA 算法都表现出几乎相同的性能。

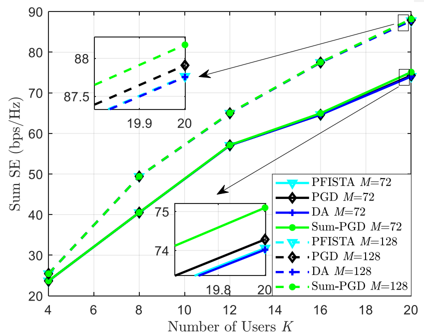


图 7 用户数 K 与 Sum SE 的关系

Fig. 7 The relationship between the number of users K and Sum SE

4 结束语

本文研究了近场模型下的基于频谱效率公平性的 XL-MIMO 系统预编码优化问题。在近场模型中，文中考虑了更符合实际情况的 LOS 和 NLOS 并存的非平稳信道。以频谱效率的几何平均值作为优化目标，通过对目标函数进行泰勒展开一阶近似、拉格朗日对偶变换和二次变换，将复杂的优化问题转化为更易于求解的形式。最后，使用 PFISTA 算法进行求解。仿真结果表明，所提算法不仅能够确保用户间的公平性，还能有效提升系统的整体频谱效率。并且，与现有方法相比，PFISTA 算法在提供几乎相同的性能增益的同时，具有较低的计算复杂度。

所提算法在 XL-MIMO 多用户通信系统中，能够改善用户体验和服务质量，以确保所有用户的公平接入和高效资源利用。未来的研究可以进一步探讨在非理想 CSI 条件下算法的鲁棒性和性能，以提高算法在实际环境中的适用性。

参考文献:

[1] Zhang Jiayi, Björnson E, Matthaiou M, et al. Prospective multiple antenna technologies for beyond 5G[J]. IEEE Journal on Selected Areas in Communications, 2020, 38(8): 1637-1660.

[2] Shi Enyu, Zhang Jiayi, Chen Shuaifei, et al. Wireless energy transfer in RIS-aided cell-free massive MIMO systems: opportunities and challenges[J]. IEEE Communications Magazine, 2022, 60(3): 26 -32.

带格式的: 缩进: 左侧: 0 厘米

[3] Iimori H, Takahashi T, Ishibashi K, et al. Joint activity and channel estimation for extra-large MIMO systems[J]. IEEE Transactions on Wireless Communications, 2022, 21 (9): 7253-7270.

[4] Cui Mingyao, Wu Zidong, Lu Yu, et al. Near-field MIMO communications for 6G: fundamentals, challenges, potentials, and future directions[J]. IEEE Communications Magazine, 2022, 61(1): 40-46.

[5] Wang Zhe, Zhang Jiayi, Du Hongyang, et al. Extremely large-scale MIMO: fundamentals, challenges,solutions, and future directions[J]. IEEE Wireless Communications, 2024, 31(3): 117-124.

[6] Lu Haiquan, Zeng Yong. Communicating with extremely large-scale array/surface: unified modeling and performance analysis[J]. IEEE Transactions on Wireless Communications, 2022, 21(6): 4039-4053.

[7] Filho J C M, Brante, G, Souza R D, et al. Exploring the non-overlapping visibility regions in XL-MIMO random access and scheduling[J]. IEEE Transactions on Wireless Communications, 2022, 21(8): 6597-6610.

[8] Gao Xiang, Tufvesson F, Edfors O. Massive MIMO channels—measurements and models[C]//2013 Asilomar Conference on Signals, Systems and Computers. 2013: 280-284.

[9] Amiri A, Manchón C N, Carvalho E D. A message passing based receiver for extra-large scale MIMO[C]//2019 IEEE 8th International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP). 2019: 564-568.

[10] Sherman J. Properties of focused apertures in the Fresnel region[J]. IRE Transactions on Antennas and Propagation, 1962, 10(4): 399-408.

[11] Souza J H I D, Filho J C M, Amiri A, et al. QoS-aware user scheduling in crowded XL-MIMO systems under non-stationary multi-state LoS/NLoS channels[J]. IEEE Transactions on Vehicular Technology, 2023, 72(6): 7639-7652.

[12] Liu Jiuyu, Ma Yi, Wang Jinfei, et al. A non-stationary channel model with correlated NLoS/LoS states for ELAA-mMIMO [C]//2021 IEEE Global Communications Conference (GLOBECOM). 2021: 1-6.

[13] Cui Yanling, Zhao Long, Gao Pengzun, et al. Optimal power allocation for XL-MIMO systems based on geometric mean of SE[C]//2023 8th IEEE International Conference on Network Intelligence and Digital Content (IC-NIDC). 2023: 227-231.

[14] Souza J H I D, Amiri A, Abrão T, et al. Quasi-distributed antenna selection for spectral efficiency maximization in subarray switching XL-MIMO systems[J]. IEEE Transactions on Vehicular Technology, 2021, 70(7): 6713-6725.

[15] Yu Hongwen, Tuan H D, Dutkiewicz E, et al. Maximizing the geometric mean of user-rates to improve rate-fairness: proper vs. improper Gaussian signaling[J]. IEEE Transactions on Wireless Communications, 2022, 21(1): 295-309.

[16] Shen Kaiming, Yu Wei. Fractional programming for communication systems—part I: power control and beamforming[J]. IEEE Transactions on Signal Processing, 2018, 66(10) : 2616-2630.

[17] Pellaco L, Bengtsson M, Jaldén J. Matrix-inverse-free deep unfolding of the weighted MMSE beamforming algorithm[J]. IEEE Open Journal of the Communications Society, 2022, 3: 65-81.

[18] Beck A, Teboulle M. A fast iterative shrinkage-thresholding algorithm for linear inverse

problems[J]. SIAM journal on imaging sciences, 2009, 2(1): 183-202.

[19] Ngo H Q, Larsson E G, Marzetta T L. Energy and spectral efficiency of very large multiuser MIMO systems[J]. IEEE Transactions on Communications, 2013, 61(4): 1436-1449.

[20] Amiri A, Rezaie S, Manchón C N, et al. Distributed receiver processing for extra-large MIMO arrays: A message passing approach[J]. IEEE Transactions on Wireless Communications, 2022, 21(4): 2654-2667.

[21] 朱小双,傅友华. STAR-RIS 辅助通信感知一体化系统联合波束成形设计[J].数据采集与处理,2024,(1):140-153.

ZHU Xiaoshuang, FU Youhua. Joint Beamforming Design for STAR-RIS Assisted Integrated Sensing and Communication System[J]. Journal of Data Acquisition and Processing,2024,(1):140-153.

[22] Shen Kaiming, Yu Wei. Fractional Programming for Communication Systems—Part II: Uplink Scheduling via Matching[J]. IEEE Transactions on Signal Processing, 2018, 66(10): 2631-2644.

[23] Wang Yajun, Fang Lili, Cai Shanjie, et al. Low-Complexity Algorithm for Maximizing the Weighted Sum-Rate of Intelligent Reflecting Surface Assisted Wireless Networks[J]. IEEE Internet of Things Journal, 2024, 11(6): 10490-10499.

[24] **Guo Huayan, Liang Yingchang , Chen Jie, et al. Weighted Sum-Rate Maximization for Reconfigurable Intelligent Surface Aided Wireless Networks[J]. IEEE Transactions on Wireless Communications, 2020, 19(5): 3064-3076.**

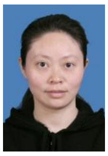
[25] Rodrigues V C, AmiriA, Abrão T, et al. Low-complexity distributed XL-MIMO for multiuser detection[C]//2020 IEEE International Conference on Communications Workshops (ICC Workshops). 2020: 1-6.

[26] Xu Bokai, Zhang Jiayi, Li Jiaxun, et al . Jac-PCG based low-complexity precoding for extremely large-scale MIMO systems[J]. IEEE Transactions on Vehicular Technology, 2023,72(12): 16811-16816.

[27] He Boxin, Zhang Aihua, Hao Wanming, et al. Multiple Beam Selection and Near-Optimal Digital Precoding for Multiuser Millimeter-Wave Massive MIMO Systems[J]. IEEE Transactions on Circuits and Systems II: Express Briefs , 2023,70 (2):811-855.

作者简介：

李志立(1999-)，男，硕士研究生，研究方向:研究方向：超大规模MIMO、预编码等，E-mail：1022020606@njupt.edu.cn。



傅友华(1978-),通信作者,女,博士，副教授，研究方向：MIMO 无线通信信号处理技术，E-mail：fuyh@njupt.edu.cn。

宋云超(1988-)，男，博士，副教授，研究方向：大规模 MIMO，智能反射面，强化学习。