提升双层速率分割多址中的用户公平性：一种可移动天线方法

Ji Luo、Yaxuan Chen、Guangchi Zhang、Miao Cui、Hao Fu和Changsheng You

arXiv:2603.07127v1 [cs.IT] 7 Mar 2026

摘要——在双层速率分拆多址（RSMA）等先进多用户系统中提升用户公平性，是一项关键而又具有挑战性的任务。本文提出一种新颖的可移动天线（MA）方法来应对这一挑战。我们构建了一个最大最小公平问题，通过联合优化波束成形矩阵、用户聚类、公共速率分配以及天线位置向量（APV），最大化作为公平性关键指标的最小用户速率。为求解这一非凸问题，我们设计了一种高效的双环迭代算法。外环利用动态邻域剪枝粒子群优化方法寻找高质量的APV，而内循环则针对给定的APV优化其余变量。仿真结果验证了所提方法，表明该方案相较于各种基准方案可获得显著的公平性增益。

索引词——可移动天线、双层速率分割多址、用户公平性、双环迭代算法。

I. 引言

速率分拆（RATE-splitting multiple access，RSMA）已成为第六代（6G）网络中一种强大的非正交传输框架，这主要得益于其在干扰管理方面的灵活性 [1]。通过将消息拆分为公共流和私有流，该框架提供了一种统一结构，将空分多址（SDMA）和非正交多址（NOMA）作为其特例。这种适应性已被证明能够带来更优的增益 [2]。

传统单层速率分割多址的一个关键局限在于，其公共速率由信道最弱的用户决定，从而在大规模系统中形成公平性瓶颈。双层速率分割多址架构通过引入包含簇间和簇内公共流的分层结构，借助双层连续干扰消除（SIC）实现更高效的干扰管理，从而缓解了这一问题。尽管已有研究表明，这能够提高效率[3]，[4]，，但最终的系统性能，以及由此实现的用户公平性，仍然受到静态无线信道的根本限制。

近年来，移动天线（MA）技术被引入，作为一种主动设计无线

G. Zhang 和 M. Cui 的工作得到了广东省基础与应用基础研究基金（项目编号：2026A1515011208）的支持。C. You 的工作得到了国家自然科学基金（项目编号：62571227）的支持。（通信作者：Guangchi Zhang）

Ji Luo、Yaxuan Chen、Guangchi Zhang、Miao Cui 和 Hao Fu 均任职于广东工业大学信息工程学院，地址为中国广州 510006（电子邮箱：{luoji2、chenyaxuan}@mails.gdut.edu.cn；{gczhang、cuimiao、ecejasonhaofu}@gdut.edu.cn）。Changsheng You 任职于南方科技大学电子与电气工程系，地址为中国深圳 518055（电子邮箱：youcs@sustech.edu.cn）。

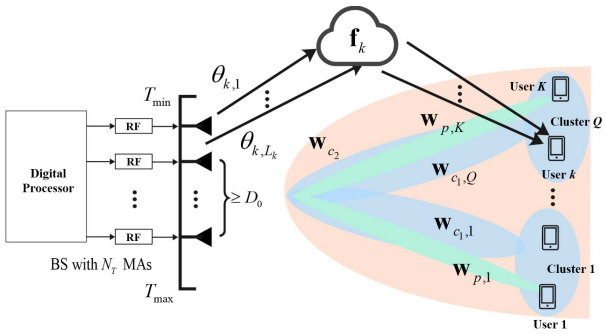


图1：移动天线辅助双层速率分裂多址系统。

信道 [5]。通过动态调整天线位置，移动天线能够创建更有利的信道条件，为通信设计提供额外的自由度（DoF）。与传统固定位置天线相比，移动天线在波束成形、空间复用和分集增益方面具有显著的性能提升 [5]–[7]。

移动天线与速率分拆多址的协同是一个颇具前景但尚处于起步阶段的研究领域。现有工作主要集中于移动天线辅助单层速率分割多址或和速率导向设计[8]，[9]。移动天线在双层RSMA框架中提升用户公平性的潜力仍未得到充分探索。为弥补这一空白，本文提出了一种专门用于提升双层速率分割多址系统中用户公平性的移动天线方案。本文的主要贡献有三点。首先，我们将移动天线引入双层速率分裂多址系统，缓解了传统单层速率分割多址系统中的公共速率瓶颈。其次，我们从用户公平性的角度，对波束成形矩阵、公共速率分配、用户聚类和天线位置向量（APV）进行联合优化建模。第三，针对这一非光滑、非凸问题，我们开发了一种高效的双环迭代算法。外环采用动态邻域剪枝粒子群优化（DNPPSO）算法优化APV，内循环则通过信道相似度和连续凸近似（SCA）技术确定用户聚类和资源分配。仿真结果表明，与各种基准方案相比，我们提出的方案实现了显著的公平性增益。

II. 系统模型与问题表述

如图 1 所示，我们考虑一个下行通信系统，其中配备 N_T 个移动天线（MAs）的基站（BS）通过双层速率分割多址方案为 K 个单天线用户提供服务

方案。移动天线和用户的索引集合分别表示为 $\mathcal{M}$、$\{1,2,\ldots,N_T\}$ 和 $\mathcal{K}$、$\{1,2,\ldots,K\}$。

根据 [7],，我们使用一维（1D）局部坐标系来表示移动天线位置。APV 定义为 $\mathbf{t}$ $[t_1,t_2,\ldots,t_{N_T}]$，其中坐标 t_m m个移动天线的位置受以下条件约束：

$$T_{\min} \le t_m \le T_{\max},\ \forall m \in \mathcal{M}, \tag{1}$$

$$|t_i - t_j| \ge D_0,\ \forall i,j \in \mathcal{M},\ i \ne j. \tag{2}$$

其中，$T_{\min}$ 和 $T_{\max}$ 定义了移动天线阵列线性区域的边界，D_0 是移动天线之间必须保持的最小间距，以减轻耦合效应。这意味着该区域最多可容纳 $N\max_T = \left\lfloor \frac{T_{\max}-T_{\min}}{D_0} \right\rfloor + 1$个移动天线，其中 $\lfloor\cdot\rfloor$ 表示向下取整算子。

为实施两层速率分割多址方案 [1],，将用户划分为 Q 个互不相交的簇。令 $\mathcal{Q}$,$\{1,2,\ldots,Q\}$ 为簇索引集合。每个簇 $q \in \mathcal{Q}$包含一个用户子集 $\mathcal{K}_q$，满足：

$$\bigcup_{q\in\mathcal{Q}} \mathcal{K}_q = \mathcal{K}, \quad \mathcal{K}_q \cap \mathcal{K}_{q'} = \emptyset, \forall q,q' \in \mathcal{Q}, q \ne q'. \tag{3}$$

传输信号由三种数据流组成：一个由所有用户解码的簇间公共流 s_{c_2}； Q个簇内公共流 $\{s_{c_1,q}\}_{q\in\mathcal{Q}}$，其中$s_{c_1,}$ $_q$ 由簇 q中的用户解码；以及 K 个私有流 $\{s_{p,k}\}_{k\in\mathcal{K}}$，其中$s_{p,}$ $_k$ 仅由用户 k解码。

每个数据流均使用专用的波束成形向量进行传输，分别记为 $\mathbf{w}_{c_2}$、$\mathbf{w}_{c_1,}$ $_q$和 $\mathbf{w}_{p,}$ $_k$。基站处的叠加信号为 $\mathbf{x}$

$$= {}_{c_2}s_{c_2} + \sum_{q\in\mathcal{Q}} {}_{c_1}{}_q s_{c_1}{}_q + \mathbf{w}\,\mathbf{w},\ , \sum_{k\in\mathcal{K}} {}_{p}{}_k s_{p}{}_k.$$

$\mathbf{w},\ ,$ 假设各数据流具有归一化的平均功率，且波束成形向量需满足基站处的总发射功率约束：

$$\|\mathbf{w}_{c_2}\|^2 + \sum_{q\in\mathcal{Q}} \|\mathbf{w}_{c_1,q}\|^2 + \sum_{k\in\mathcal{K}} \|\mathbf{w}_{p,k}\|^2 \le P_{\max}. \tag{4}$$

用户 k 接收到的信号为

$$y_k = \mathbf{h}_k^H(\mathbf{t})\mathbf{x} + n_k, \tag{5}$$

其中，$n_k \sim \mathcal{CN}(0,\sigma^2)$表示加性高斯白噪声，$\mathbf{h}_k(\mathbf{t})$为信道向量。基站到用户 k 的信道由 L_k 条路径构成 [6],，其中第 ℓ条路径的出发角（AoD）为 $\theta_{k,\ \ell}$。用户 k 的第 m个可移动天线（MA）的场响应向量（FRV）为

$$\mathbf{a}_k(t_m) = [e^{j\frac{2\pi}{\lambda}t_m\cos\theta_{k,1}}, \cdots, e^{j\frac{2\pi}{\lambda}t_m\cos\theta_{k,L_k}}]^T, \tag{6}$$

其中，λ 为波长。用户 k的综合场响应矩阵（FRM）为 $\mathbf{A}_k(\mathbf{t}) = [\mathbf{a}_k(t_1)\ldots,\ \mathbf{a}_k(t_{N_T})] \in \mathbb{C}^{L_k\times N_T}$。因此，信道向量为[1]

$$\mathbf{h}_k(\mathbf{t}) = \mathbf{A}_k^H(\mathbf{t})\mathbf{f}_k \in \mathbb{C}^{N_T\times 1}, \tag{7}$$

其中，$\mathbf{f}_k = [f_{k,1}, f_{k,2}, \ldots, f_{k,L_k}]^T \in \mathbb{C}^{L_k\times 1}$为路径响应向量（PRV），$f_{k,\ \ell}$ 表示用户 k的第 ℓ条路径的复增益。

[1]我们假设准静态信道模型以及完美信道状态信息（CSI），其可通过压缩感知[5], 获取，以建立理论性能基准。实际约束（例如移动时间、驱动能耗以及不完美信道状态信息）留待未来工作。

每个用户 $k \in \mathcal{K}_q$ 采用串行干扰消除来解码其目标数据流。解码顺序为：1）簇间公共流 s_{c_2}，2）簇内公共流$s_{c_1,q}$，以及3）私有数据流$s_{p,k}$。在每个阶段，成功解码的信号会在解码下一个数据流之前从接收信号中减去。用户 k 解码 s_{c_2}、$s_{c_1,q}$和$s_{p,k}$ 的信干噪比分别由下式给出

$$\gamma_{c_2,k} = \frac{|\mathbf{h}_k^H(\mathbf{t})\mathbf{w}_{c_2}|^2}{I_k}, \gamma_{c_1,q_k} = \frac{|\mathbf{h}_k^H(\mathbf{t})\mathbf{w}_{c_1,q}|^2}{I_k - |\mathbf{h}_k^H(\mathbf{t})\mathbf{w}_{c_1,q}|^2},$$

$$\gamma_{p,k} = \frac{|\mathbf{h}_k^H(\mathbf{t})\mathbf{w}_{p,k}|^2}{I_k - |\mathbf{h}_k^H(\mathbf{t})\mathbf{w}_{c_1,q}|^2 - |\mathbf{h}_k^H(\mathbf{t})\mathbf{w}_{p,k}|^2}, \tag{8}$$

其中，$I_k = \sum_{q\in\mathcal{Q}} |\ {}_k{}^H\ \ {}_{c_1}{}_q|^2 + \sum_{k\in\mathcal{K}} |\ {}_k{}^H\ \ {}_{p}{}_k|^2 + \mathbf{h}$ σ^2 $(\mathbf{t})\mathbf{w},\ \mathbf{h}\ (\mathbf{t})\mathbf{w},$ 表示在解码 s_{c_2}时，来自所有簇内数据流和私有流的干扰功率与噪声功率之和。相应的可达速率为 $R_{c_2,\ k} = \log_2(1+\gamma_{c_2,\ k})$、$R_{c_1,\ q_k} = \log_2(1+\gamma_{c_1,\ q_k})$和$R_{p,\ k} = \log_2(1+\gamma_{p,\ k})$。

为了使所有预期用户都能够解码公共流，其速率受信道条件最差的用户限制。也就是说，簇间数据流的速率为 $R_{c_2} = \min_{k\in\mathcal{K}} R_{c_2,\ k}$，而簇 q 的簇内数据流速率为 $R_{c_1,\ q} = \min_{k\in\mathcal{K}_q} R_{c_1,\ q_k}$。随后，这些公共速率将在相应用户之间进行分配。设$r_{c_2,\ k}$ 和$r_{c_1,\ q_k}$ 分别为分配给用户k的公共速率份额。必须满足以下约束：

$$\sum_{k\in\mathcal{K}} r_{c_2,k} \le R_{c_2}, \quad \sum_{k\in\mathcal{K}_q} r_{c_1,q_k} \le R_{c_1,q}, \forall q \in \mathcal{Q},$$

$$r_{c_2,k} \ge 0, \quad r_{c_1,q_k} \ge 0, \quad \forall k \in \mathcal{K}_q, \forall q \in \mathcal{Q}. \tag{9}$$

用户 k 的总速率是分配给它的公共速率与其私有速率之和，即 $R_{q_k} = r_{c_2,\ k} + r_{c_1,\ q_k} + R_{p,\ k}$。

最小用户速率取决于波束成形矩阵 $\mathbf{W} \triangleq [\mathbf{w}_{c_2}, \{\mathbf{w}_{c_1,\ q}\}, \{\mathbf{w}_{p,\ k}\}] \in \mathbb{C}^{N_T\times(1+Q+K)}$、用户聚类 $\{\mathcal{K}_q\}$、公共速率分配向量 $\mathbf{r} \triangleq [\{r_{c_2,k}\}, \{r_{c_1,q_k}\}] \in \mathbb{R}^{1\times 2K}$，以及 APV $\mathbf{t}$。为确保用户公平性，我们旨在通过联合优化这些变量来最大化最小用户速率。问题如下

$$\max_{\mathbf{W},\mathbf{t},\mathbf{r},\{\mathcal{K}_q\}} \min_{k\in\mathcal{K}} R_{q_k} \tag{10a}$$

$$\text{s.t.} \quad (1),(2),(3),(4),(9).$$

问题（10）由于最大最小目标函数是非凸的，且高维优化变量之间存在复杂耦合，因此难以求解。在下一节中，我们提出一种迭代算法，以求得高质量的次优解。

III. 针对问题 （10）

为求解问题（10），我们设计了一种将优化变量解耦的双环迭代算法。具体而言，外环使用DNPPSO算法优化 APV $\mathbf{t}$ 。对于外环生成的每个候选APV，内循环通过优化波束成形矩阵 $\mathbf{W}$、公共速率分配向量$\mathbf{r}$和用户聚类 $\{\mathcal{K}_q\}$，针对给定的 $\mathbf{t}$评估其适应度值。

A. 外循环优化

在外环中，我们通过求解以下问题来优化APV t:

$$\max_{\mathbf{t}} \min_{k\in\mathcal{K}} R_{q_k} \quad \text{s.t. } (1),(2). \tag{11a}$$

由于目标函数（11）不可微，我们采用无梯度方法DNPPSO算法。DNPPSO 利用粒子群进行搜索，每个粒子代表一个候选APV，以探索解空间。首先，生成一个由 P 个粒子组成的种群。对于第 p 个粒子，其位置向量和速度向量初始化为:

$$\mathbf{t}_p^{(0)} = [t_{p,1}^{(0)}, t_{p,2}^{(0)}, ..., t_{p,N_T}^{(0)}], \mathbf{v}_p^{(0)} = [v_{p,1}^{(0)}, v_{p,2}^{(0)}, ..., v_{p,N_T}^{(0)}], \tag{12}$$

其中，每个元素 $t_{p,\,m}^{(0)}$ 从$[T_{\min}, T_{\max}]$ 中随机选取，以满足（1）。为引导搜索，我们定义一个适应度函数，其中包含用于惩罚约束违反的惩罚项:

$$\mathcal{F}\left(\mathbf{t}_p^{(i)}\right) \triangleq R\left(\mathbf{t}_p^{(i)}\right) - \tau\mathcal{L}\left(\mathbf{t}_p^{(i)}\right), \tag{13}$$

其中，$R(\mathbf{t}_p^{(i)})$ 表示在迭代 i时，对给定 APV $\mathbf{t}_p^{(i)}$进行内循环优化所得到的最小用户速率。项 $\mathcal{L}(\mathbf{t}_p^{(i)})$ 统计违反约束（2）的移动天线数量，而 τ 是一个较大的正惩罚因子。该公式引导粒子朝可行域移动。

每个粒子根据其个体最优位置 $\mathbf{t}_{p,\,\text{pbest}}$，以及迄今找到的全局最优位置 $\mathbf{t}_{\text{gbest}}$ 来更新速度和位置。每次迭代的更新规则如下:

$$\mathbf{v}_p^{(i+1)} = \omega\mathbf{v}_p^{(i)} + c_1\tau_1\left(\mathbf{t}_{p,\text{pbest}} - \mathbf{t}_p^{(i)}\right) + c_2\tau_2\left(\mathbf{t}_{\text{gbest}} - \mathbf{t}_p^{(i)}\right), \tag{14}$$

$$\mathbf{t}_p^{(i+1)} = \mathcal{B}\left(\mathbf{t}_p^{(i)} + \mathbf{v}_p^{(i+1)}\right), \tag{15}$$

其中，c_1 和 c_2 分别为个体学习因子和全局学习因子，τ_1、$\tau_2 \sim U\in(0,1)$为随机参数，ω为惯性权重，在 I 次迭代中从 $\omega_{\max}$ 线性递减至 $\omega_{\min}$。函数 $\mathcal{B}(\cdot)$确保更新后的位置仍满足（1）。在每次迭代中对每个粒子执行内循环优化会带来显著的计算负担。为减轻这一负担，我们引入了DNPPSO [7]中的剪枝机制。其依据是，收敛至全局最优位置附近的粒子发现更优解的潜力较小。因此，我们定义了迭代 i 时的动态邻域半径 $d^{(i)}{}_{\text{R}}$，以及相应的邻域集合 $\mathcal{P}(d^{(i)}{}_{\text{R}})$:

$$\left\{d_{\text{R}}^{(i)} \mid \sum_{p=1}^{\tilde{P}^{(i)}} \xi\left(\left\|\mathbf{t}_p^{(i)} - \mathbf{t}_{\text{gbest}}\right\|_2 < d_{\text{R}}^{(i)}\right) \triangleq \tilde{P}^{(i)} - \tilde{P}^{(i+1)}\right\}, \tag{16}$$

$$\mathcal{P}\left(d_{\text{R}}^{(i)}\right) \triangleq \left\{\mathbf{t}_p^{(i)} \mid \|\mathbf{t}_p^{(i)} - \mathbf{t}_{\text{gbest}}\|_2 < d_{\text{R}}^{(i)}, 1 \le p \le \tilde{P}^{(i)}\right\}, \tag{17}$$

其中，$\xi(\cdot)$ 是指示函数，$\tilde{P}^{(i)}$ 是迭代 i时活跃粒子的数量。在每次迭代中，根据需要额外剪枝的粒子数量确定邻域半径 $d^{(i)}{}_{\text{R}}$，并将此范围内的粒子加入邻域集 $\mathcal{P}(d^{(i)}{}_{\text{R}})$，这样下一次迭代时便会跳过内循环。

算法1 所提出的用户聚类方案

1: 初始化簇 $\mathcal{K}_q = \emptyset$、 $\forall q$以及未聚类用户集合 $\mathcal{U} = \mathcal{K}$。
2: 计算矩阵C，其元素为$c_{a,\,b} = \frac{|\mathbf{h}_a^H \mathbf{h}_b|}{\|\mathbf{h}_a\|\|\mathbf{h}_b\|}$ 对所有 $a, b \in \mathcal{U}, a \neq b$。
3: 设置聚类索引 $q = 1$。
4: 重复
5: 求 $(a^*, b^*) = \arg\max_{a,\,b\in\mathcal{U},\,a\neq b} c_{a,\,b}$。
6: $\mathcal{K}_q \leftarrow \{a^*, b^*\}, \mathcal{U} \leftarrow \mathcal{U} \setminus \{a^*, b^*\}, q \leftarrow q + 1$。
7: 直到 $|\mathcal{U}| < 2$
8: $\mathcal{K}_Q \leftarrow \{u\}$，其中 u 是 $\mathcal{U}$ 中剩余的元素。
返回: $\{\mathcal{K}_q\}$。

B. 内循环优化

在内循环中，对于给定的 APV $\mathbf{t}_p^{(i)}$，我们求解用户聚类 $\{\mathcal{K}_q\}$、传输波束成形矩阵 $\mathbf{W}$以及公共速率分配向量 $\mathbf{r}$，以评估式（13）中的适应度值。该子问题表述为

$$\max_{\mathbf{W},\mathbf{r},\{\mathcal{K}_q\}} \min_{k\in\mathcal{K}} R_{q_k} \quad \text{s.t. } (3),(4),(9). \tag{18a}$$

问题（18）具有挑战性，这是由于离散聚类变量与连续波束成形变量和速率变量之间存在耦合。为使其具有可处理性，我们将其分解为两个依次求解的子问题：1）用户聚类；2）波束成形与速率分配的联合优化。

1）用户聚类优化：我们首先确定用户聚类 $\{\mathcal{K}_q\}$。直观地说，为最大化簇内公共速率，共享同一波束成形向量$\mathbf{w}_{c_1}$的 q 用户应具有高度相关的信道向量。因此，我们提出了一种基于信道相似度的贪心聚类算法。我们使用余弦相似度指标来量化任意两个用户 a 和 b 之间的相似度：$c_{a,\,b}$，$\frac{|\mathbf{h}_a^H \mathbf{h}_b|}{\|\mathbf{h}_a\|\|\mathbf{h}_b\|}$。该算法迭代地将信道相似度最高的两个未聚类用户进行配对。如果用户数为奇数，则将最后一个未配对用户分配到一个独立簇中。具体过程详见算法 1。[2]

2) 联合波束成形与速率分配优化：然后，我们通过求解来优化 W 和 r

$$\max_{\mathbf{W},\mathbf{r},} \min_{k\in\mathcal{K}} R_{q_k} \quad \text{s.t. } (4),(9). \tag{19a}$$

问题（19）是非凸的。为处理该问题，我们首先应用半正定松弛（半正定松弛）。我们定义 $\mathbf{W}_{c_2}$、$\mathbf{w}_{c_2}\mathbf{w}_{c_2}^H$、$\mathbf{W}_{c_1,q}$、$\mathbf{w}_{c_1,q}\mathbf{w}_{c_1,q}^H$，以及 $\mathbf{W}_{p,k}$、$\mathbf{w}_{p,k}\mathbf{w}_{p,k}^H$。令 $\widetilde{\mathbf{W}}\{\mathbf{W}_{c_2}\{\mathbf{W}_{c_1 q}\}\{\mathbf{W}_{p k}\}\}$ $\widetilde{\mathbf{W}}$、$\widetilde{\mathbf{W}}_i$、$\widetilde{\mathbf{W}}_i$为这些新矩阵 $\widetilde{\mathbf{W}}$ 变量的集合，其中 $\widetilde{\mathbf{W}}_i$ 表示 $\widetilde{\mathbf{W}}$ 中的第 i 个矩阵，即 $\widetilde{\mathbf{W}}_1 = \mathbf{W}_{c_2}$、$\widetilde{\mathbf{W}}_{1+q} = \mathbf{W}_{c_1,q}$ 和 $\widetilde{\mathbf{W}}_{1+Q+k} = \mathbf{W}_{p,k}$。

[2]为重点考察所提系统的性能潜力，我们采用固定规模聚类以实现低复杂度。可变规模聚类由于用户到簇的相似度设计、大规模簇中重新引入公共流瓶颈的风险，以及移动天线引起的动态特性而具有非平凡性；因此，将其留待未来研究。

这些新的矩阵变量之和定义为$\widetilde{\mathbf{W}}_s = \mathbf{W}_{c_2} + \sum_{q\in\mathcal{Q}} \mathbf{W}_{c_1,q} + \sum_{k\in\mathcal{K}} \mathbf{W}_{p,k}$。通过引入一个 z 最小速率的松弛变量，问题（19）可改写为：

$$\max_{\widetilde{\mathbf{W}},\mathbf{r},z} \quad z \tag{20a}$$

$$\text{s.t. } \mathrm{tr}\left(\widetilde{\mathbf{W}}_s\right) \le P_{\max}, \tag{20b}$$

$$r_{c_2,k} + r_{c_1,q_k} + \Gamma_{p,k} + \Lambda_{p,k} \ge z, \forall k \in \mathcal{K}, \tag{20c}$$

$$\Gamma_{c_2,k} + \Lambda_{c_2,k} \ge \sum_{k\in\mathcal{K}} r_{c_2,k}, \forall k \in \mathcal{K}, \tag{20d}$$

$$\Gamma_{c_1,q_k} + \Lambda_{c_1,q_k} \ge \sum_{k\in\mathcal{K}_q} r_{c_1,q_k}, \forall k \in \mathcal{K}_q, q \in \mathcal{Q}, \tag{20e}$$

$$\widetilde{\mathbf{W}}_i \succeq 0, \forall i \in \{1,\dots,1+Q+K\}, \tag{20f}$$

$$\mathrm{rank}\left(\widetilde{\mathbf{W}}_i\right) = 1, \forall i \in \{1,\dots,1+Q+K\}, \tag{20g}$$

$$r_{c_2,k} \ge 0, r_{c_1,q_k} \ge 0, \forall k \in \mathcal{K}_q, \forall q \in \mathcal{Q}. \tag{20h}$$

其中，(20c)-(20e)中的函数定义为 Γ_l , $\log_2(\lambda_l + \sigma^2)$, ($l \in \{\{c_2,k\},\{c_1,q_k\},\{p,k\}\}$), $\Lambda_{p,k}$, $-\log_2(\varphi_{p,k}+\sigma^2)$, $\Lambda_{c_2,k}$, $-\Gamma_{c_1,q_k}$, Λ_{c_1,q_k} , $-\Gamma_{p,k}$, 以及

$$\lambda_{c_2,k} = \mathbf{h}_k^H \widetilde{\mathbf{W}}_s \mathbf{h}_k, \ \lambda_{c_1,q_k} = \mathbf{h}_k^H \left(\widetilde{\mathbf{W}}_s - \widetilde{\mathbf{W}}_1\right) \mathbf{h}_k$$

$$\lambda_{p,k} = \mathbf{h}_k^H \left(\widetilde{\mathbf{W}}_s - \widetilde{\mathbf{W}}_1 - \widetilde{\mathbf{W}}_{1+q}\right) \mathbf{h}_k$$

$$\varphi_{p,k} = \mathbf{h}_k^H \left(\widetilde{\mathbf{W}}_s - \widetilde{\mathbf{W}}_1 - \widetilde{\mathbf{W}}_{1+q} - \widetilde{\mathbf{W}}_{1+Q+k}\right) \mathbf{h}_k. \tag{21}$$

由于秩一约束 (20g) 以及具有差分凸（D.C.）结构的速率约束，问题 (20) 仍然是非凸的。我们采用逐次凸近似（SCA）技术处理 D.C. 约束。具体而言，凸函数 $\Lambda_{c_2,k}$,Λ_{c_1,q_k} 和 $\Lambda_{p,k}$ 将由其在上一迭代所得解 [3]处的一阶泰勒下界替代，该解记为 $(\cdot)^{(j)}$。$\Lambda_{c_2,k}$ 的下界为

$$\Lambda_{c_2,k}^{(j,lb)} = -\log_2(\lambda_{c_1,q_k}^{(j)} + \sigma^2) - \frac{\lambda_{c_1,q_k} - \lambda_{c_1,q_k}^{(j)}}{(\lambda_{c_1,q_k}^{(j)} + \sigma^2)\ln 2}. \tag{22}$$

以类似方式处理 Λ_{c_1,q_k} 和 $\Lambda_{p,k}$ ，因此为简洁起见省略相关细节。代入这些下界后，第 $j+1$ 次迭代需要求解的问题为

$$\max_{\widetilde{\mathbf{W}},\mathbf{r},z} \quad z \tag{23a}$$

$$\text{s.t. } \Gamma_{c_2,k} + \Lambda_{c_2,k}^{(j,lb)} \ge \sum_{k\in\mathcal{K}} r_{c_2,k}, \forall k \in \mathcal{K}, \tag{23b}$$

$$\Gamma_{c_1,q_k} + \Lambda_{c_1,q_k}^{(j,lb)} \ge \sum_{k\in\mathcal{K}_q} r_{c_1,q_k}, \forall k \in \mathcal{K}_q, q \in \mathcal{Q}, \tag{23c}$$

$$r_{c_2,k} + r_{c_1,q_k} + \Gamma_{p,k} + \Lambda_{p,k}^{(j,lb)} \ge z, \forall k \in \mathcal{K}, \tag{23d}$$

(20b), (20f), (20g), (20h).

剩下唯一的非凸性是约束（20g）。此外，根据 [2], 中的定理 1，问题（23）在不含约束（20g）时的最优解总能满足秩一约束。因此，可以在不损失最优性的情况下去掉约束（20g），并通过内点法求解得到的半定规划（SDP）问题。

[3]在首次连续凸近似迭代中，公共流波束成形器初始化为零，而私有流波束成形器则在发射功率约束下采用最大比传输（MRT）进行初始化。

算法2 问题（10）的双层循环算法

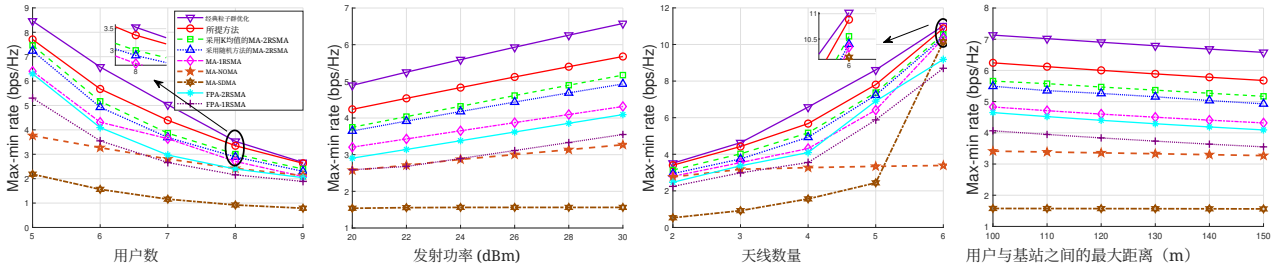
1: 初始化粒子速度 $\{\mathbf{v}_p^{(0)}\}_{p=1}^P$ 以及位置 $\{\mathbf{t}_p^{(0)}\}_{p=1}^P$.
2: 通过（$\mathcal{F}(\mathbf{t}_p^{(0)})$，计算每个粒子的初始适应度值13）。
3: 令 $\mathbf{t}_{p,\text{pbest}} \leftarrow \mathbf{t}_p^{(0)}, \forall p$，且 $\mathbf{t}_{\text{gbest}} \leftarrow \arg\max_p\{\mathcal{F}(\mathbf{t}_p^{(0)})\}$。
4: 对于 $i=1:\ I$，执行
5: 更新惯性权重 ω 和活跃粒子的数量 $\tilde{P}^{(i)}$。
6: 对于 $p=1:\ P$，执行
7: 如果粒子 $\mathbf{t}_p^{(i-1)}$ 未被剪枝，则
8: 更新 $\mathbf{v}_p^{(i)}$和 $\mathbf{t}_p^{(i)}$通过式（14）和式（15）。
9: 通过求解问题（$\mathcal{F}(\mathbf{t}_p^{(i)})$）来评估适应度 18）。
10: 更新 $\mathbf{t}_{p,\text{pbest}}$，若 $\mathcal{F}(\mathbf{t}_p^{(i)}) > \mathcal{F}(\mathbf{t}_{p,\text{pbest}})$。
11: 更新 $\mathbf{t}_{\text{gbest}}$，若 $\mathcal{F}(\mathbf{t}_p^{(i)}) > \mathcal{F}(\mathbf{t}_{\text{gbest}})$。
12: 结束 if
13: 结束 for
14: 确定邻域半径 $d_R^{(i)}$ 并设置 $\mathcal{P}(d^{(i)}{}_{\mathrm{R}})$ 通过 (16)、(17)。
15: 将 $\mathcal{P}(d^{(i)}{}_{\mathrm{R}})$ 内的粒子标记为在下一次迭代中跳过。
16: 结束 for
返回：最终 APV $\mathbf{t}_{\text{gbest}}$ 及对应的 W、r、$\{\mathcal{K}_q\}$

C. 总体算法

所提双环迭代算法总结于算法 2 中。在外环中，仅当粒子找到适应度值更高的位置时，才更新全局最优位置 $\mathbf{t}_{\text{gbest}}$ ，从而确保目标函数值在迭代过程中非递减。此外，由于功率约束（4），目标函数值存在上界。因此，所提算法具有单调且有界的特性。用户聚类的计算复杂度为 $\mathcal{O}(K^2)$。通过内点法求解问题（23）的计算复杂度为 $\mathcal{O}(\log(1/\epsilon)N_T^{4.5})$，其中 ϵ 为求解精度 [3]。因此，算法 2 的总计算复杂度为 $\mathcal{O}((K^2+\log(1/\epsilon)N_T^{4.5})\sum_{i=1}^I \tilde{P}^{(i)})$ [7]。

IV. 仿真结果

本节将评估所提算法的性能。对于所有用户，我们采用包含 $L_k = L = 6$ 条路径的几何信道模型 [5]。位置响应向量的元素$f_{k,\ell}$建模为独立同分布的圆对称复高斯（CSCG）随机变量，即$f_{k,\ell} \sim \mathcal{CN}\left(0 C_0 d_k^{-\alpha}/L_k\right)$，其中 $C_0 = -30$ dB为1 m参考距离处的信道功率增益，d_k 为基站到用户 k的距离，$\alpha = 2.8$ 为路径损耗指数。用户均匀分布在角度范围 $[0,\pi]$ 内，且 d_k 从 $[20,150]$ m中均匀选取。SCA逐次凸近似迭代的求解精度 ϵ 设为 10^{-3}。主要系统参数设置如下：$T_{\min} = 0$、$T_{\max} = 10\lambda$、$D_0 = \lambda/2$、$\lambda = 0.1$ m和 $\sigma^2 = -90$ dBm。DNPPSO的参数按照常用配置 [7]，进行设置，粒子数和迭代次数分别设为 $P = 50$ 和$I = 50$。采用线性剪枝策略，其中活跃粒子数随着 $\beta = 0.02$减少至 βP，并且 $c_1 = c_2 = 1.4$、$\tau = 20$、$\omega_{\min} = 0.4$、$\omega_{\max} = 0.9$。下文给出的结果是500次蒙特卡洛实现的平均性能。在每次实现中，用户位置和信道实现均随机生成。为减小随机初始化的影响，对于所有基准方法，外环优化均使用不同的随机种子执行多次，并选取性能最佳的解。



(a) (b) (c) (d)

图 2：最大最小速率与以下因素的关系：(a) 用户数，$N_T = 4, P_{\max} = 30$ dBm；(b) 发射功率，$N_T = 4, K = 6$；(c) 天线数量，$K = 6, P_{\max} = 30$ dBm；(d) 最大用户距离，$N_T = 4, K = 6, P_{\max} = 30$ dBm。所有子图共用同一图例。

我们将所提方案与以下几种基准方法进行比较：[4] **(1)** 经典粒子群优化：将DNPPSO方法替换为经典粒子群优化方法（即 $\mathcal{P}\left(d_R^{(i)}\right) = \emptyset$）。**(2)** 采用K均值的MA-2RSMA：使用K均值算法替换所提聚类方法 [3]。**(3)** 采用随机方法的MA-2RSMA：采用随机用户分簇。**(4) MA-1RSMA**：一种单层速率分割多址基准方案，其中省略了簇内公共流[9]。**(5) MA-NOMA**：一种传统的MA辅助非正交多址方案 [10]。**(6) MA-SDMA**：一种仅传输私有流的空分多址基准方案。**(7) FPA-2RSMA**：所提方案的固定位置天线对应方案。**(8) FPA-1RSMA**：一种基于固定位置天线的单层速率分割多址方案。

图 2(a) 展示了最大最小速率与用户数之间的关系。所提方法持续展现出竞争力。值得注意的是，移动天线为这些方案带来增益的根本机制有所不同。在空分多址中，移动天线用于使不同用户的信道向量正交；而在所提方法中，增益则源于提高部分用户之间的信道相关性，同时降低其他用户之间的信道相关性，并结合用户聚类。

图2(b)展示了最大最小速率随基站发射功率变化的关系。经典粒子群优化取得了最高性能，其次是所提方法，因为DNPPSO以略微的性能损失换取了显著降低的复杂度。尽管存在这种权衡，所提方法仍凭借以下三个因素显著优于其他基准方法：移动天线带来的额外空间自由度、双层速率分割多址更优越的干扰管理能力，以及有效的信道感知用户分簇。

图2(c)绘制了最大最小速率随发射天线数量变化的关系。增加天线数量能够增强阵列增益和波束赋形增益，从而提升所有方案的性能。所提方法始终优于基准方法，并以更少的天线达到目标速率，凸显了其降低硬件成本的潜力。

图 2(d) 展示了用户距离对最大最小速率的影响。所提出的双层速率分割多址方案与单层速率分割多址方案之间的性能差距尤为明显。在单层速率分割多址中，公共速率受信道条件最差的用户限制，但我们提出的双层方法利用簇内流，在较小的用户组内更有效地管理干扰。

[4]我们在所有仿真中遵循固定变量原则。对于所有需要聚类的方案，均采用相同数量的簇和相同的簇大小。对于所有基于SCA的优化，均采用相同的迭代设置。对于所有采用基于PSO的优化的方案，均采用相同的参数。

V. 结论

本文研究了一种移动天线辅助双层速率分裂多址系统，以提升用户公平性。我们通过联合设计波束成形矩阵、公共速率分配、用户聚类和APV，构建了一个以最大化最小用户速率为目标的优化问题。针对这一非凸优化问题，我们提出了一种高效的双环迭代算法。外环利用DNPPSO方法寻找高质量的APV，而内环则采用基于信道相似度的聚类方案和基于SCA的资源分配算法，对其余变量进行优化。仿真结果验证了所提框架的有效性，证实其在提升最小用户速率方面显著优于多种基准方案。

参考文献

[1] B. Clerckx等，“速率分拆多址导论：教程、误区与常见问题解答”，IEEE J. Sel. Areas Commun.，第41卷，第5期，第1265–1308页，2023年5月。[2] H. Fu等，“双用户下行链路MISO速率分拆系统的鲁棒安全波束成形设计”，IEEE Trans. Wireless Commun.，第19卷，第12期，第8351–8365页，2020年12月。[3] T. Qi等，“IRS辅助的分层速率分拆多址系统的发射功率最小化”，IEEE Wireless Commun. Lett.，第14卷，第5期，第1386–1390页，2025年5月。[4] Z. Li等，“面向无蜂窝大规模MIMO系统的具有均衡聚类设计的双层RSMA框架”，IEEE Internet Things J.，第12卷，第18期，第37339–37354页，2025年9月。[5] L. Zhu等，“可移动天线赋能无线通信的建模与性能分析”，IEEE Trans. Wireless Commun.，第23卷，第6期，第6234–6250页，2024年6月。[6] J. Ding等，“可移动天线通信系统的能效最大化”，IEEE Trans. Wireless Commun.，第25卷，第2624–2638页，2026年。[7] J. Ding等，“可移动天线辅助的近场多用户通信”，IEEE Wireless Commun. Lett.，第14卷，第1期，第138–142页，2025年1月。[8] C. Zhang等，“可移动天线辅助下行链路RSMA系统的和速率最大化”，载于IEEE国际通信会议（ICC）论文集，加拿大魁北克省蒙特利尔，2025年9月，第1049–1054页。[9] X. Li等，“面向超可靠低时延通信服务的可移动天线增强型RSMA短分组传输”，IEEE Internet Things J.，第12卷，第10期，第14771–14785页，2025年5月。[10] Z. Xiao等，“可移动天线辅助非正交多址：天线位置、预编码与解码设计的联合优化”，IEEE Trans. Wireless Commun.，第25卷，第4595–4612页，2026年。