利用分段波导赋能的捏缩天线系统（SWAN）进行上行链路三混合波束成形

Hao Jiang、Chongjun Ouyang、Zhaolin Wang、Yuanwei Liu，IEEE会士、Arumugam Nallanathan，IEEE会士、Zhiguo Ding，IEEE会士，以及Robert Schober，IEEE会士

arXiv:2603.04684v1 [eess.SP] 4 Mar 2026

摘要——提出了一种基于分段波导赋能的捏缩天线系统（SWAN）的三混合波束成形架构，用于上行链路多用户MIMO通信，并联合优化数字、模拟和夹持波束成形。同时考虑了射频链与分段馈电点之间的全连接（FC）结构和部分连接（PC）结构。对于全连接架构，采用加权最小均方误差（WMMSE）和迫零（ZF）方法优化三重混合波束成形。具体而言，数字、模拟和夹持波束成形分量分别通过闭式解、黎曼流形优化和高斯-赛德尔搜索进行优化。对于部分连接架构，提出了一种面向SWAN接收机的交错拓扑，其中分配给每个射频链（子阵列）的分段与其他子阵列的分段相互交错。基于该结构，设计了一种基于WMMSE的三混合方案，其中用于全连接结构的黎曼流形更新被逐元素相位校准所取代，以利用模拟波束成形中的稀疏性。为深入理解所提系统的性能，针对全连接结构和部分连接结构推导了关于分段数的速率缩放规律。结果表明：i）对于全连接结构和部分连接结构，采用所提三重混合波束成形的SWAN始终优于传统混合波束成形以及采用夹持波束成形的传统挤压天线系统；ii）当分段数较大时，部分连接结构能够在和速率与能耗之间取得良好平衡；iii）可达速率并不一定随分段数的增加而提高。

索引词——捏合天线系统、三混合波束成形、和速率最大化、性能分析。

I. 引言

数十年来，多输入多输出（MIMO）技术在电信领域发挥着关键作用，为全球提供快速、无处不在的连接 [1], [2]。MIMO 技术的进步带来了更高的阵列增益、复用增益和分集增益，满足了日益增长的需求

Hao Jiang、Chongjun Ouyang 和 Arumugam Nallanathan 任职于伦敦玛丽女王大学电子工程与计算机科学学院，地址为英国伦敦 E1 4NS（电子邮箱： {hao.jiang, c.ouyang, a.nallanathan}@qmul.ac.uk）。Zhaolin Wang 和 Yuanwei Liu 任职于香港大学电气与计算机工程系，地址为香港（电子邮箱：{zhaolin.wang, yuanwei}@hku.hk）。Z. Ding 任职于新加坡南洋理工大学电气与电子工程学院（EEE），邮编 639798（电子邮箱：zhiguo.ding@ntu.edu.sg）。R. Schober 任职于德国埃尔朗根-纽伦堡弗里德里希-亚历山大大学（FAU）数字通信研究所（电子邮箱：robert.schober@fau.de）。

即更高的频谱效率、更低的时延和更强的可靠性 [3], [4]。

关于多输入多输出的研究可归纳为三大技术趋势：i）增大天线孔径；ii）提高天线布置密度；iii）采用可重构天线架构。具体而言，增大天线孔径是指集成数百或数千个天线单元，并使其间距为半个波长，从而形成非常大的物理孔径，例如大规模MIMO、超大规模天线阵列（ELAAs）和巨型MIMO [5], [6]。相比之下，提高天线布置密度则通过取消相邻单元之间的半波长约束来增加孔径密度，从而能够在固定孔径尺寸上形成全息表面或连续表面 [7]。

尽管前景广阔，前两种趋势仍依赖对无线信道的被动适应，导致硬件成本和实现复杂度增加。作为解决方案，能够操控无线信道的可重构天线架构近年来备受关注[8]。这类可重构天线能够针对特定任务调整无线传播，具体方式包括天线重定位或端口选择。将可重构天线集成到现有的混合波束成形架构中，可将其作为第三层加入，专门用于操控无线信道的电磁（EM）特性 [9]。由此实现了由数字波束成形、模拟波束成形和可重构电磁波束成形构成的三混合波束成形架构，为波束成形设计提供了额外的自由度（DoF）。利用这一新的自由度，可为下一代无线网络提供低成本、高能效的多输入多输出部署方案 [10]。

A. 先前工作

为实现三混合波束成形架构，可以采用不同类型的可重构天线，包括可移动天线（MAs） [11], 流体天线系统（FASs）[12], 和动态超表面（DMAs） [13]。然而，大多数可重构天线仅支持小尺度信道操控，因为其可重构性依赖于波长尺度操作，例如端口重新定位、极化调整和频率选择。为实现大尺度信道操控，NTT DOCOMO 提出了挤压天线系统（PASS） [14], ，该系统由用于远距离信号传输的波导和用于信号辐射的夹持

天线（PAs）组成。基于这些组件，PASS 通过波导内低衰减传播与灵活的自由空间传播相结合的方式，将预期信号传送给用户 [15]。与现有的可重构天线原型不同，PAs 可在数十米乃至数百米的范围内重新定位，从而实现大尺度信道可重构性。

除传统波束成形架构外，PASS 还支持天线位置优化，这种方法称为夹持波束成形，由此形成三混合波束成形架构。作为一项面向 6G 及未来通信的前沿技术，PASS 已受到全球工业界和学术界日益广泛的关注。具体而言，[16] 的作者提出了一种基于物理的波导内传播模型，为基于 PASS 的信道建模奠定了基础。在此模型基础上，研究人员针对下行链路三混合波束成形，考察了连续和离散 PA 位置设置下的功率最小化问题；这两种设置是实现夹持波束成形最常用的方法。此外，[17] 的作者研究了上行链路和下行链路中的三混合波束成形，以实现和速率最大化，并证明 PASS 优于传统多输入多输出，甚至优于大规模 MIMO。在上述研究的基础上，[18] 的作者设计了一种用于三混合波束成形的逐元素优化框架，实现了低复杂度的基带与夹持波束成形联合优化。此外，[19] 的作者改进了 PASS 的波导内信道模型，并指出，在相移器处辐射的比例功率可作为一种新的优化自由度。此外，[20] 的作者将波导视为一种新型资源块，并提出了多种用于夹持波束成形的传输结构。基于这一新开发的功率辐射模型，研究人员利用分支定界（BnB）方法确定了最优三混合波束成形方案，展示了 PASS 相较传统多输入多输出的真正优势。为克服夹持波束成形带来的计算复杂度，研究人员开发了基于机器学习（ML）的解决方案，例如图神经网络（GNNs）和 Transformer [21], ，将优化任务转移到离线训练阶段，从而实现低复杂度的在线实现。

此外，PASS赋能的三混合波束成形具有灵活性，可用于物理层安全（PLS）、通感一体化（ISAC）以及无线信息与能量同传（SWIPT）等特定应用，另请参见 [22]。

B. 研究动机与贡献

尽管早期研究取得了令人鼓舞的成果，但现有工作主要聚焦于三混合波束成形架构的全数字实现，因此引发了对其高昂硬件成本的担忧。

为缓解这一问题，可采用混合数字-模拟波束成形来减少射频链数量，因此它被视为实现大规模MIMO的实用方案 [23]。为充分发挥混合波束成形的优势，[24] 的作者针对下行链路中基于PASS的三重混合波束成形架构，研究了和速率最大化问题。此外，文献 [25]阐明了基于PASS的三重混合波束成形在分析方面的优势。然而，将这项工作扩展到上行链路可能会遇到问题，这是由于天线间辐射（IAR）所致：同一波导上的其他功率放大器会重新辐射某个功率放大器接收到的信号。此外，IAR效应可能会破坏上行链路与下行链路信道之间的对偶性，从而影响信道估计。分段波导赋能的捏缩天线系统（SWAN）可作为解决方案，通过以多个短分段波导取代传统的长距离波导来应对这些问题，每段波导均配有专用馈电点[26], [27]。缩短分段长度后，可通过在每段部署一个功率放大器来减轻IAR。此外，较短的分段波导能够减轻波导内损耗并便于维护，因此SWAN架构有望成为大规模部署PASS的理想方案。

为充分发挥基于 SWAN 的三混合波束成形架构的优势，我们提出了一种上行链路三混合波束成形设计的优化框架，并进行性能分析，以进一步刻画其性能极限。本文的主要贡献总结如下：

- 我们研究一种上行多用户MIMO系统，其接收机配备了基于 SWAN 的三混合波束成形架构。我们考虑射频链与分段馈电点之间的全连接（FC）和部分连接（PC）结构。
- 针对全连接结构，我们开发了基于加权最小均方误差（WMMSE）和迫零（ZF）的三重混合波束成形策略，以最大化上行链路和速率。我们采用块坐标下降（BCD）框架，统一实现：i）数字波束成形的闭式解；ii）模拟波束成形的黎曼流形优化；以及 iii）用于夹捏波束成形的高斯–赛德尔更新。
- 针对部分连接结构，我们提出了一种适用于 SWAN接收机的交错拓扑，其中分配给每个射频链（即子阵列）的线段与其他子阵列的线段交错排列。在此基础上，我们进一步开发了一种适用于部分连接结构的基于 WMMSE 的三重混合波束成形方法。具体而言，在同一块坐标下降框架中，我们以逐元素相位校准取代黎曼流形更新，从而利用稀疏模拟结构并降低复杂度。
- 我们进一步分析全连接和部分连接结构的速率缩放定律如何随分段数变化，以刻画其性能极限。结果表明，理论速率不一定会随着分段数的增加而提高。
- 大量仿真结果表明：i）基于 SWAN 的三重混合波束成形优于采用传统混合波束成形的大规模 MIMO 和采用钳缩波束成形的 PASS；ii）所提出的方法能够有效地为这两种结构设计三重混合波束成形。

C. 组织结构与符号说明

本文其余部分安排如下：第 II 节刻画上行信道并构建和速率最大化问题。在此基础上，第 III 节和第 IV 节分别针对全连接结构和部分连接结构提出三混合波束成形算法。第V节阐述这两种结构的性能极限。第 VI 节给出数值结果。最后，第 VII 节总结全文。

符号说明：标量、向量和矩阵分别用小写字母、黑体小写字母和黑体大写字母表示。 $\mathbb{C}^{M\times N}$ 和 $\mathbb{R}^{M\times N}$ 表示 $M\times N$ 复值矩阵和实值矩阵的空间。$(\cdot)^{\mathsf{T}}$、$(\cdot)^{\mathsf{H}}$、$(\cdot)^{*}$和$(\cdot)^{-1}$ 分别表示转置、共轭转置、共轭和求逆运算。$[\cdot]_{i,:}$、$[\cdot]_{:,j}$和$[\cdot]_{i,j}$ 分别表示提取矩阵的第 i行、第 j列和第(i,j)个元素的运算。 $\odot$和 ∇ 分别表示哈达玛积和梯度算子。 $\Re\{\cdot\}$ 和 $\Im\{\cdot\}$ 分别表示复值变量的实部和虚部。$\mathrm{j}=\sqrt{-1}$ 和 e 分别表示虚数单位和欧拉数。$\mathbb{E}\{\cdot\}$ 表示期望算子。 $\|\cdot\|$、$|\cdot|$和$(\cdot)^{|\cdot|}$分别表示 L_2-范数、绝对值和逐元素绝对值。算子 $\angle(\cdot)$表示提取复数相位的运算。

II. 系统模型

基于 SWAN 的三混合波束成形系统模型如图 1 所示，其中 K 个单天线用户向由 M 条线段组成的 SWAN接收机发送信号。SWAN 服务区域定义为 $D_x\times D_y$，其中 D_x和 D_y分别表示沿 x轴和 y轴方向的边长。每条线段包含一个功率放大器，其长度为 $L=D_x/M$。这些线段在高度 H 处依次连接，且平行于 x轴。第 m条线段上馈电点的 x坐标记为 x_m^{FD}，所有馈电点均位于线段的左端点。第 k个用户的几何位置为 $\mathbf{r}_k=[r_{x,k},r_{y,k},0]^{\mathsf{T}}$，而第 m个功率放大器的坐标为 $\mathbf{p}_m=[x_m,0,H]^{\mathsf{T}}$。由于所有功率放大器具有相同的 y和 z坐标，即对于所有 m，$y_m=0$ 和 $z_m=H$ 均成立，因此定义 x坐标向量 $\mathbf{x}=[x_1,...,x_M]^{\mathsf{T}}\in\mathbb{R}^{M\times 1}$，以简洁地表示功率放大器的位置。从波束成形架构的角度来看，馈电点与射频链之间考虑两种连接策略：全连接（FC）和部分连接（PC）。在全连接结构中，每条射频链通过移相器连接到所有馈电点；而在部分连接结构中，每条射频链仅连接到馈电点的一个子集。射频链数量和移相器数量分别记为 N_{RF} 和 N_{PS}。

A. 系统模型

在本小节中，我们对基于 SWAN 的三重混合波束成形架构的上行链路信道进行建模，其中包括自由空间传播和波导内传播。具体而言，

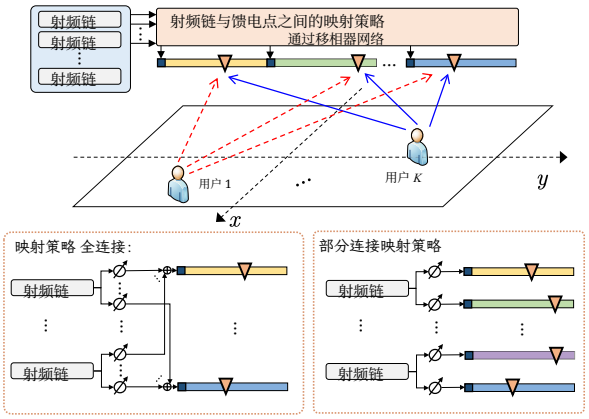


图 1：SWAN 提出的多用户上行链路三混合波束成形架构示意图。

从第 k 个用户到 SWAN 上 PAs 的自由空间信道向量为

$$\tilde{\mathbf{h}}_k(\mathbf{x})=\left[\frac{\sqrt{\eta}}{r_{1,k}}\mathrm{e}^{-\mathrm{j}\frac{2\pi}{\lambda_c}r_{1,k}},...,\frac{\sqrt{\eta}}{r_{M,k}}\mathrm{e}^{-\mathrm{j}\frac{2\pi}{\lambda_c}r_{M,k}}\right]^{\mathsf{T}},\tag{1}$$

其中，$\eta=\lambda_c^2/(16\pi^2)$ 表示自由空间路径增益常数，λ_c 为自由空间波长，$r_{m,k}=\|\mathbf{p}_m-\mathbf{r}_k\|$ 表示第 k 个用户与第 m 个 PA 之间的欧几里得距离。信号经 PAs 接收后，将在波导内传播至馈电点。令 Δ_m，$|x_m^{\mathrm{FD}}-x_m|$ 为第 m 个线段内的传播距离，则波导内信道可表示为

$$\mathbf{g}(\mathbf{x})=\left[10^{-\frac{\kappa\Delta_1}{20}}\mathrm{e}^{-\mathrm{j}\frac{2\pi}{\lambda_g}\Delta_1},...,10^{-\frac{\kappa\Delta_M}{20}}\mathrm{e}^{-\mathrm{j}\frac{2\pi}{\lambda_g}\Delta_M}\right]^{\mathsf{T}},\tag{2}$$

其中，$\lambda_g=\lambda_c/n_{\mathrm{eff}}$ 表示波导内波长，κ 表示以 dB/m 为单位测量的波导内损耗，n_{eff} 表示介质波导的有效折射率。因此，第 k 个用户的整体上行链路信道定义如下：

$$\mathbf{h}_k(\mathbf{x})=\tilde{\mathbf{h}}_k(\mathbf{x})\odot\mathbf{g}(\mathbf{x})\in\mathbb{C}^{M\times 1}.\tag{3}$$

通过在式 (3) 中设计 PA 位置，可实现信道重构。

基于所提出的信道模型，SWAN接收机处的接收信号可写为：

$$\mathbf{y}=\sum_{k=1}^{K}\mathbf{h}_k(\mathbf{x})s_k+\mathbf{n},\tag{4}$$

其中 s_k 表示第 k个用户发送的、满足 $\mathbb{E}\{|s_k|^2\}=P$的符号，P表示用户的上行发射功率，$\mathbf{n}\sim\mathcal{CN}(\mathbf{0}_M,\sigma^2\mathbf{I}_M)$表示功率为 σ^2的加性高斯噪声。更紧凑地，接收信号可写为：

$$\mathbf{y}=\mathbf{H}(\mathbf{x})\mathbf{s}+\mathbf{n},\tag{5}$$

其中，$\mathbf{H}(\mathbf{x})\triangleq[\mathbf{h}_1(\mathbf{x}),...,\mathbf{h}_K(\mathbf{x})]\in\mathbb{C}^{M\times K}$为上行信道矩阵，$\mathbf{s}=[s_1,s_2,...,s_K]^{\mathsf{T}}$为发送符号向量。式（5）中的接收信号经过模拟和数字波束成形处理。本文采用可达上行和速率作为衡量上行链路性能的指标 [17]，[28]。这里，我们将射频和数字波束成形矩阵分别定义为 $\mathbf{G}_{\mathrm{RF}}\in\mathbb{C}^{M\times N_{\mathrm{RF}}}$ 和

$\mathbf{G}_{\mathrm{BB}} \in \mathbb{C}^{N_{\mathrm{RF}} \times K}$。第 k 个用户的解调符号可表示如下:

$$\hat{s}_k = \mathbf{g}_k^{\mathsf{H}} \mathbf{G}_{\mathrm{RF}}^{\mathsf{H}} \mathbf{y} = \underbrace{\mathbf{g}_k^{\mathsf{H}}}_{\mathrm{D-BF}} \cdot \underbrace{\mathbf{G}_{\mathrm{RF}}^{\mathsf{H}}}_{\mathrm{A-BF}} \cdot \underbrace{\mathbf{H}(\mathbf{x})\mathbf{s}}_{\mathrm{P-BF}} + \mathbf{g}_k^{\mathsf{H}} \mathbf{G}_{\mathrm{RF}}^{\mathsf{H}} \mathbf{n}, \quad (6)$$

其中，$\mathbf{g}_k = [\mathbf{G}_{\mathrm{BB}}]_{:,\,k}$ 表示第 k 个用户的数字波束成形向量，“D-BF”、“A-BF”和“P-BF”分别代表数字波束成形、模拟波束成形和夹持波束成形。

B. 问题表述

根据式 (6) 中的信号模型，第 k 个用户的信干噪比（SINR）可表示如下:

$$\gamma_k(\mathbf{g}_k, \mathbf{G}_{\mathrm{RF}}, \mathbf{x}) = \frac{\left|\mathbf{g}_k^{\mathsf{H}} \mathbf{G}_{\mathrm{RF}}^{\mathsf{H}} \mathbf{h}_k(\mathbf{x})\right|^2}{\sum_{i=1, i\neq k}^{K} \left|\mathbf{g}_k^{\mathsf{H}} \mathbf{G}_{\mathrm{RF}}^{\mathsf{H}} \mathbf{h}_i(\mathbf{x})\right|^2 + \frac{\sigma^2}{P}\left\|\mathbf{g}_k^{\mathsf{H}} \mathbf{G}_{\mathrm{RF}}^{\mathsf{H}}\right\|^2}. \quad (7)$$

因此，综合考虑所有 K 个用户后，和速率可简洁地写为:

$$R(\mathbf{G}_{\mathrm{BB}}, \mathbf{G}_{\mathrm{RF}}, \mathbf{x}) = \sum_{k=1}^{K} \log_2\left(1 + \gamma_k(\mathbf{g}_k, \mathbf{G}_{\mathrm{RF}}, \mathbf{x})\right) \quad (8)$$

。本文考虑两种结构，即 PC 和 FC。在 FC 结构中，每个射频链的输出都通过移相器连接到各线段的所有馈电点，如图 1 左下方所示。在 PC 结构中，每个射频链的输出通过移相器连接到各线段馈电点的一个子集，如图 1 右下方所示。与 FC 结构相比，PC 结构的模拟波束成形矩阵是稀疏的，因此减少了优化变量的数量和所需移相器的数量。因此，PC 结构能够以较低的复杂度和较低的能耗实现混合波束成形[29]，，尽管由于模拟波束成形矩阵稀疏，其可用于优化的自由度较少。基于以上讨论，我们构建了基于 SWAN 的上行链路三混合波束成形问题。令 $\mathbf{W}_{\mathrm{BB}}$ 和 $\mathbf{F}_{\mathrm{BB}}$ 分别表示 FC 和 PC 结构的数字波束成形矩阵，并令 $\mathbf{W}_{\mathrm{RF}}$ 和 $\mathbf{F}_{\mathrm{RF}}$ 表示相应的模拟波束成形矩阵。因此，对于 FC 结构，我们定义 $(\mathbf{G}_{\mathrm{RF}}, \mathbf{G}_{\mathrm{BB}}) = (\mathbf{W}_{\mathrm{RF}}, \mathbf{W}_{\mathrm{BB}})$；对于 PC 结构，我们定义 $(\mathbf{G}_{\mathrm{RF}}, \mathbf{G}_{\mathrm{BB}}) = (\mathbf{F}_{\mathrm{RF}}, \mathbf{F}_{\mathrm{BB}})$。根据这些定义，上行和速率最大化问题表述如下:

$$\max_{\mathbf{x}, \mathbf{G}_{\mathrm{RF}}, \mathbf{G}_{\mathrm{BB}}} \quad R(\mathbf{G}_{\mathrm{RF}}, \mathbf{G}_{\mathrm{BB}}, \mathbf{x}) \quad \text{(P1a)}$$

$$\text{s.t.} \quad \left|[\mathbf{G}_{\mathrm{RF}}]_{i,j}\right| = 1, \ \forall i, j, \quad \text{(P1b)}$$

$$(\mathbf{G}_{\mathrm{RF}}, \mathbf{G}_{\mathrm{BB}}) = \begin{cases} (\mathbf{W}_{\mathrm{RF}}, \mathbf{W}_{\mathrm{BB}}), & \text{for FC}, \\ (\mathbf{F}_{\mathrm{RF}}, \mathbf{F}_{\mathrm{BB}}), & \text{for PC}, \end{cases} \quad \text{(P1c)}$$

$$\mathbf{x} \in \mathcal{F}. \quad \text{(P1d)}$$

约束 (P1b) 表示模拟波束成形器的单位模约束；约束 (P1c) 表示射频波束成形矩阵的具体实现；而约束 (P1d) 表示对每个功率放大器施加的位置约束，具体如下:

$$\mathcal{F} \triangleq \left\{ \mathbf{x} \;\middle|\; \begin{array}{c} (m-1)L \le [\mathbf{x}]_m \le mL, \forall m \in \mathcal{M} \\ |[\mathbf{x}]_m - [\mathbf{x}]_{m'}| \ge \Delta_{\min}, \forall m \neq m' \end{array} \right\}, \quad (9)$$

其中，$\mathcal{M}$ 定义为 $\mathcal{M}$、$\{1, 2, ..., M\}$，而 $\Delta_{\min}$ 表示两个 PAs 之间的最小间距，用于减轻互耦效应 [30]。

III. 全连接结构下基于 SWAN 的三混合波束成形方案

本节首先给出问题 (P1) 的加权最小均方误差（WMMSE）解。鉴于迫零（ZF）在实际多用户 MIMO 系统中的广泛应用，本文还给出了原问题的基于 ZF 的解，以提高所提出的基于 SWAN 的上行链路三混合波束成形架构的适用性。

A. 问题 (P1) 的基于WMMSE的解

在该方法中，我们采用块坐标下降法依次优化数字波束成形矩阵 $\mathbf{W}_{\mathrm{BB}}$、模拟波束成形矩阵 $\mathbf{W}_{\mathrm{RF}}$和 PA位置向量 $\mathbf{x}$。假设$\mathbb{E}\left\{|s_k|^2\right\} = P$、 $\mathbb{E}\{\mathbf{n}s^*\} = \mathbf{0}_M$以及 $\mathbb{E}\left\{\mathbf{n}\mathbf{n}^{\mathsf{H}}\right\} = \sigma^2\mathbf{I}_M$均成立，则第 k个用户的均方误差（MSE）可推导如下:

$$\begin{aligned} e_k &= \mathbb{E}\left\{\left|\hat{s}_k - s_k\right|^2\right\} \\ &= P\left|\mathbf{w}_k^{\mathsf{H}} \mathbf{W}_{\mathrm{RF}}^{\mathsf{H}} \mathbf{h}_k(\mathbf{x})\right|^2 + P\sum_{j=1, j\neq k}^{K}\left|\mathbf{w}_k^{\mathsf{H}} \mathbf{W}_{\mathrm{RF}}^{\mathsf{H}} \mathbf{h}_j(\mathbf{x})\right|^2 \\ &\quad - 2P\Re\left\{\mathbf{w}_k^{\mathsf{H}} \mathbf{W}_{\mathrm{RF}}^{\mathsf{H}} \mathbf{h}_k(\mathbf{x})\right\} + \sigma^2\left\|\mathbf{w}_k^{\mathsf{H}} \mathbf{W}_{\mathrm{RF}}^{\mathsf{H}}\right\|^2 + P. \end{aligned} \quad (10)$$

可以看出，式 (10) 中的均方误差具有二次型，因此比原始的和对数表达式更易于优化。此外，如[31] 和 [32], 所示，原始的和对数最大化问题等价于以下形式:

$$\min_{\mathcal{V} \cup \Omega} \quad \sum_{k=1}^{K}\left(\omega_k e_k - \log(\omega_k)\right) \quad \text{(P2a)}$$

$$\text{s.t.} \quad \text{(P1b) and (P1d)},$$

其中，$\mathcal{V} = \{\mathbf{W}_{\mathrm{BB}}、\mathbf{W}_{\mathrm{RF}}、\mathbf{x}\}$ 定义为优化变量集合，Ω $\{\omega_1, ..., \omega_K\}$ 表示所有 K 用户的权重集合。根据 [31], 中的证明，问题 (P2) 等价于 (P1)。基于问题 (P2) 的易处理形式，我们采用块坐标下降框架，将原始问题分解为关于各个优化变量的子问题。

1) 数字波束成形优化：根据块坐标下降的原理，我们固定优化变量 $\mathbf{x}$ 和 $\mathbf{W}_{\mathrm{RF}}$，并将模拟波束成形矩阵与信道向量相结合，形成等效信道向量，其由 $\check{\mathbf{h}}_k$ 、$\mathbf{W}_{\mathrm{RF}}^{\mathsf{H}}\mathbf{h}_k(\mathbf{x})$（对于 $\forall k$）给出。然后，在等效 $\check{}$信道 $\mathbf{h}_k$ 下，第 k 个用户的最优数字波束成形向量可由 (P2a) 的一阶最优性条件得到，其可写为如下形式:

$$\nabla_{\mathbf{w}_k^*} e_k = \sum_{j=1}^{K} P\check{\mathbf{h}}_j\check{\mathbf{h}}_j^{\mathsf{H}}\mathbf{w}_k - P\check{\mathbf{h}}_k + \sigma^2\mathbf{W}_{\mathrm{RF}}^{\mathsf{H}}\mathbf{W}_{\mathrm{RF}}\mathbf{w}_k = \mathbf{0}_{N_{\mathrm{RF}}},$$

由此可推导出最优波束成形向量如下：

$$\mathbf{w}_{k,\star} = \left(\sum_{j=1}^{K} \check{\mathbf{h}}_j \check{\mathbf{h}}_j^{\mathsf{H}} + (\sigma^2/P)\mathbf{W}_{\mathrm{RF}}^{\mathsf{H}}\mathbf{W}_{\mathrm{RF}}\right)^{-1} \check{\mathbf{h}}_k. \quad (11)$$

更紧凑地，最优射频域波束成形矩阵可写为 $\mathbf{W}_{\mathrm{BB},\star} = [\mathbf{w}_{1,\star},\ldots,\mathbf{w}_{K,\star}]$。

2）模拟波束成形优化：有了最优数字波束成形矩阵 $\mathbf{W}_{\mathrm{BB},\star}$ 后，现在考虑模拟波束成形矩阵的优化。具体而言，将 $\mathbf{W}_{\mathrm{BB},\star}$ 代回式（10）中的均方误差表达式，可得

$$e_{k,\mathbf{w}_{k,\star}} = P\mathbf{w}_{k,\star}^{\mathsf{H}}\mathbf{W}_{\mathrm{RF}}^{\mathsf{H}}\left(\sum_{j=1}^{K}\mathbf{h}_j\mathbf{h}_j^{\mathsf{H}} + (\sigma^2/P)\mathbf{I}_M\right) \times \mathbf{W}_{\mathrm{RF}}\mathbf{w}_{k,\star} - 2P\Re\left\{\mathbf{w}_{k,\star}^{\mathsf{H}}\mathbf{W}_{\mathrm{RF}}^{\mathsf{H}}\mathbf{h}_k\right\} + P. \quad (12)$$

因此，关于 $\mathbf{W}_{\mathrm{RF}}$ 的子问题可表述如下：

$$\min_{\mathbf{W}_{\mathrm{RF}}} \quad \sum_{k=1}^{K}\omega_k e_{k,\mathbf{w}_{k,\star}} \quad \text{(P3a)}$$
$$\text{s.t.} \quad \text{(P1b)}.$$

此外，定义 $\mathbf{R} \triangleq \sum_{j=1}^{K}\mathbf{h}_j\mathbf{h}_j^{\mathsf{H}} + (\sigma^2/P)\mathbf{I}_M$、$\mathbf{C} \triangleq P\sum_{k=1}^{K}\omega_k \mathbf{h}_{k,\star}\mathbf{h}_{k,\star}^{\mathsf{H}}$ 和 $\mathbf{B} \triangleq 2P\sum_{k=1}^{K}\omega_k \mathbf{h}_{k,\star}\mathbf{w}_{k,\star}^{\mathsf{H}}$，则第 k 个用户的均方误差表达式可改写为更紧凑的形式，即

$$\begin{aligned} f(\mathbf{W}_{\mathrm{RF}}) &\triangleq \sum_{k=1}^{K}\omega_k e_{k,\mathbf{w}_{k,\star}} \\ &= P\sum_{k=1}^{K}\omega_k \mathrm{tr}\left\{\mathbf{W}_{\mathrm{RF}}^{\mathsf{H}}\mathbf{R}\mathbf{W}_{\mathrm{RF}}\mathbf{w}_{k,\star}\mathbf{w}_{k,\star}^{\mathsf{H}}\right\} \\ &\quad - 2P\sum_{k=1}^{K}\Re\left\{\mathrm{tr}\left\{\mathbf{W}_{\mathrm{RF}}^{\mathsf{H}}\mathbf{h}_k\mathbf{w}_{k,\star}^{\mathsf{H}}\right\}\right\} + KP \\ &= P\mathrm{tr}\left\{\mathbf{W}_{\mathrm{RF}}^{\mathsf{H}}\mathbf{R}\mathbf{W}_{\mathrm{RF}}\sum_{k=1}^{K}\mathbf{w}_{k,\star}\mathbf{w}_{k,\star}^{\mathsf{H}}\right\} \\ &\quad - 2P\Re\left\{\mathrm{tr}\left\{\mathbf{W}_{\mathrm{RF}}^{\mathsf{H}}\sum_{k=1}^{K}\mathbf{h}_k\mathbf{w}_{k,\star}^{\mathsf{H}}\right\}\right\} + KP \\ &= \mathrm{tr}\left\{\mathbf{W}_{\mathrm{RF}}^{\mathsf{H}}\mathbf{R}\mathbf{W}_{\mathrm{RF}}\mathbf{C}\right\} - \Re\left\{\mathrm{tr}\left\{\mathbf{W}_{\mathrm{RF}}^{\mathsf{H}}\mathbf{B}\right\}\right\} + KP, \end{aligned} \quad (13)$$

这里用到了 $\mathrm{tr}\{\mathbf{AB}\} = \mathrm{tr}\{\mathbf{BA}\}$。

求解问题 (P3) 的关键在于单位模约束 (P1b)，该约束并非凸约束。为处理单位模约束，我们采用黎曼流形优化技术，将问题 (P3) 转换为无约束形式 [24], [29]。具体而言，单位模约束定义了 $M \times N$ 个半径为 1 的复圆，这是一类流形，其内积定义为 $\langle x_1, x_2\rangle = \Re\{x_1^* x_2\}$。根据这一定义，该流形由 $\mathcal{M}_{\mathrm{cc}} = \{|[\mathbf{W}_{\mathrm{RF}}]_{i,j}| = 1,\ \forall i,j\}$ 表征。因此，在该流形上执行梯度下降会自然地满足单位模约束，因为流形上的每个点都符合这一要求。流形上的梯度下降包括以下步骤：1）计算黎曼梯度；2）沿黎曼梯度的负方向更新；3）将终点重新投影回流形。按照这一顺序，第一步是计算由

算法1 黎曼梯度算法

初始化迭代索引 $t = 0$ 和优化变量 $\mathbf{W}_{\mathrm{BB},\star}$，$\mathbf{x}$，和 $\mathbf{W}^{(0)}$ 射频
计算下降方向 $\mathbf{D}^{(0)} = -\mathrm{grad} f(\mathbf{W}^{(0)}$ 射频$)$
重复
使用 Armijo 回溯线搜索选择步长 α_t
　通过回缩（16）确定下一个点 $\mathbf{W}^{(t+1)}$ 射频
　根据（15）计算黎曼梯度 $\mathbf{G}^{(t+1)}$
　计算 $\mathbf{G}_+^{(t)}$ 和 $\mathbf{D}_+^{(t)}$ 从 $\mathbf{G}^{(t)}$ 和 $\mathbf{D}^{(t)}$ 传输得到的结果
　根据（17）将 $\mathbf{W}^{(t)}$ 射频传输至$\mathbf{W}^{(t+1)}$ 射频
　计算 Polak–Ribi`ere 参数 by $\beta^{(t+1)} = \max\left\{0, \frac{\Re\left\{\text{平移}\left\{\left(\mathbf{G}_+^{(t)}\right)^{\mathsf{H}}\left(\mathbf{G}^{(t)} - \mathbf{G}^{(t)}\right)\right\}\right\}}{\Re\left\{\left(\mathbf{G}^{(t)}\right)^{\mathsf{H}}\mathbf{G}^{(t)}\right\}}\right\}$
　通过 $\mathbf{D}^{(t+1)} = -\mathbf{G}^{(t+1)} + \beta^{(t+1)}\mathbf{D}_+^{(t)}$ 计算共轭方向
　设置 $t = t + 1$
直到和速率的相对下降幅度低于一个 p重新定义的阈值
返回优化后的模拟波束成形矩阵$\mathbf{W}_{\mathrm{RF},\star}$，$\mathbf{W}^{(t)}$ 射频

欧几里得梯度诱导的黎曼梯度。忽略式 (36) 中的常数项，欧几里得梯度可按如下方式求得：

$$\nabla_{\mathbf{W}_{\mathrm{RF}}} f(\mathbf{W}_{\mathrm{RF}}) = 2\,\mathbf{R}\,\mathbf{W}_{\mathrm{RF}}\mathbf{C} - \mathbf{B}. \quad (14)$$

因此，令 t 为梯度下降过程的迭代索引，黎曼梯度可推导为

$$\begin{aligned} \mathbf{G}^{(t)} &= \mathrm{grad}\ f(\mathbf{W}_{\mathrm{RF}}^{(t)}) = \mathrm{proj}_{\mathbf{W}_{\mathrm{RF}}}\ \nabla_{\mathbf{W}_{\mathrm{RF}}^{(t)}} f\left(\mathbf{W}_{\mathrm{RF}^{(t)}}\right) \\ &= \nabla_{\mathbf{W}_{\mathrm{RF}}} f(\mathbf{W}_{\mathrm{RF}}^{(t)}) - \Re\{\nabla_{\mathbf{W}_{\mathrm{RF}}^{(t)}} f(\mathbf{W}_{\mathrm{RF}}^{(t)}) \odot (\mathbf{W}_{\mathrm{RF}}^{t})^*\} \odot \mathbf{W}_{\mathrm{R}}^{(t)\mathrm{F}}, \end{aligned} \quad (15)$$

这可以解释为从欧几里得梯度到流形 $\mathcal{M}$的切空间上的投影。因此，为确保更新后的$\mathbf{W}_{\mathrm{RF}}$ 仍位于该流形上，我们将回缩操作定义如下 [29]:

$$\mathbf{W}_{\mathrm{RF}}^{(t+1)} = \mathrm{retract}\left(\mathbf{W}_{\mathrm{RF}}^{(t)} + \alpha\,\mathbf{D}^{(t)}\right) = \frac{\mathbf{W}_{\mathrm{RF}}^{(t)} + \alpha\,\mathbf{D}^{(t)}}{(\mathbf{W}_{\mathrm{RF}}^{(t)} + \alpha\,\mathbf{D}^{(t)})^{|\cdot|}}, \quad (16)$$

其中，$\alpha > 0$ 为步长，$\mathbf{D}^{(t)}$ 表示点 $\mathbf{W}^{(t)}$ 射频处切空间中的矩阵。此外，为便于应用共轭梯度，我们定义了将 $\mathcal{M}_{\mathrm{cc}}$ 上两个不同点处的切空间相互映射的传输操作。根据 [29], , 该操作定义如下：

$$\begin{aligned} \mathbf{D}_+^{(t)} &= \mathrm{trans}(\mathbf{D}^{(t)}) \\ &= \mathbf{D}^{(t)} - \Re\left\{\mathbf{D}^{(t)} \odot \left(\mathbf{W}_{\mathrm{RF}}^{(t+1)}\right)^*\right\} \odot \mathbf{W}_{\mathrm{RF}}^{(t+1)}, \end{aligned} \quad (17)$$

其中，$\mathbf{D}^{(t)}$ 表示点 $\mathbf{W}^{(t)}$ 射频处切空间中的矩阵，而 $\mathbf{D}_+^{(t)}$ 表示点 $\mathbf{W}^{(t+1)}$ 射频 处切空间中的矩阵。整个梯度下降法总结于算法1中，该算法计算得到优化后的射频域波束成形矩阵，记为 $\mathbf{W}_{\mathrm{RF},\star}$。

3）夹持波束成形优化：对于固定的 $\mathbf{W}_{\mathrm{BB},\star}$和 $\mathbf{W}_{\mathrm{RF},\star}$，最后一步优化旨在找到最优 PA 位置。为解决这一问题，采用高斯-赛德尔方法来克服 PASS [16] 中的多模态优化目标。具体而言，每个线段的 x坐标区域被划分为一组离散的候选位置，

算法2 高斯-赛德尔算法

1: 初始化迭代索引 $t=0$ 和优化变量 $\mathbf{W}_{\mathrm{BB},\star}$, $\mathbf{W}_{\mathrm{RF},\star}$, $\mathbf{x}^{(0)}$
2: 重复
3: 对于 $m\in\{1,\ldots,M\}$ 执行
4: 通过降低均方误差 (P2a) 来更新 $x_m^{(t)}$，同时检查 $\mathbf{x}^{(t)}\in\mathcal{F}$的可行性
5: 结束循环
6: 设置 $t=t+1$
7: 直到均方误差的相对下降幅度低于预定义的阈值
8: 返回优化后的相控阵位置 $\mathbf{x}_\star$、$\mathbf{x}^{(t)}$

算法3: 基于加权最小均方误差的全连接三重混合波束成形

1: 初始化迭代索引 $t=0$ 和优化变量 $\mathbf{W}_{\mathrm{BB},\star}^{(0)}$, $\mathbf{W}_{\mathrm{RF},\star}^{(0)}$, $\mathbf{x}^{(0)}$
2: 重复
3: 根据式(11)求解$\mathbf{W}^{(t)}{}_{\mathrm{BB}\,\star}$
4: 通过算法1求解 $\mathbf{W}^{(t)}{}_{\mathrm{RF}\,\star}$
5: 通过算法2求解 $\mathbf{x}_\star^{(t)}$
6: 根据公式 (19) 求得$\Omega_\star^{(t)}$
7: 设置 $t=t+1$
8: 直到和速率的相对增幅低于预设阈值
9: 返回优化后的波束成形矩阵 $\mathbf{W}_{\mathrm{BB},\star}$、$\mathbf{W}^{(t)}{}_{\mathrm{BB}\,\star}$ 和 $\mathbf{W}_{\mathrm{RF},\star}\triangleq\mathbf{W}^{(t)}{}_{\text{射频},\star}$，以及优化后的 PA 位置 $\mathbf{x}_\star$、$\mathbf{x}_\star^{(t)}$

并使用一维搜索遍历这些候选位置。只有在满足可行性约束且均方误差降低时，才会更新 PA 位置。重复此过程，直至连续迭代之间的相对均方误差降至预设阈值以下。所得算法总结于算法2，其输出记为 $\mathbf{x}_\star$。

4）辅助变量集优化：将通过上述步骤获得的 $\mathbf{W}_{\mathrm{BB},\star}$、$\mathbf{W}_{\mathrm{RF},\star}$和 $\mathbf{x}_\star$ 代入式 (10) 中 e_k 的表达式 $\forall k$，可得 $e_{k,\star}$、$\forall k$。因此，通过求函数 (P2a) 关于 ω_k、$\forall k$的驻点，可以更新辅助变量 ω_k、 $\forall k$，得到以下方程：

$$\frac{\partial}{\partial\omega_k}\sum_{k=1}^{K}\left(\omega_k e_k-\log\left(\omega_k\right)\right)=e_k-1/\omega_k=0. \tag{18}$$

上述方程给出了 ω_k的更新规则，具体如下：

$$\omega_{k,\star}=1/e_{k,\star}. \tag{19}$$

因此，我们有 $\Omega_\star=\{\omega_{k,\star}\}_{k=1}^{K}$。

基于第 III-A、III-A2、III-A3 和 III-A4 节的讨论，整体块坐标下降算法总结于算法 3 中，可用于求解和速率最大化问题的一个局部最优解。整体计算复杂度推导为 $\mathcal{O}(I_{\mathrm{iter},\,1}(M^3+I_{\mathrm{iter}2}M+I_{\mathrm{iter}3}M+K)$，其中 $I_{\mathrm{iter}1}$、 $I_{\mathrm{iter}2}$ 和 $I_{\mathrm{iter}3}$ 分别表示块坐标下降、黎曼梯度下降和 Gauss-Seidel 搜索所需的迭代次数。

B. 基于零迫的解（P1）

在上行链路波束成形矩阵的设计中，迫零因其鲁棒性强且计算复杂度低而被广泛采用，因此它是评估和实现所提出的三混合波束成形架构的重要基准 [33], [34]。因此，本小节考虑基于迫零的混合波束成形。为确保迫零的可行性，我们假设 $K\le N_{\mathrm{RF}}$。通过施加迫零约束，可以消除用户间干扰。在这种情况下，优化问题可表示如下 [35]:

$$\max_{\mathcal{V}}\quad\sum\nolimits_{k=1}^{K}\log_2\left(1+\frac{P/\sigma^2}{\left\|\mathbf{W}_{\mathrm{RF}}\mathbf{w}_k\right\|^2}\right) \tag{P4a}$$

$$\text{s.t.}\quad\left|[\mathbf{W}_{\mathrm{RF}}]_{i,j}\right|=1,\ \forall i,j \tag{P4b}$$

$$\mathbf{W}_{\mathrm{BB}}^{\mathsf{H}}\mathbf{W}_{\mathrm{RF}}^{\mathsf{H}}\mathbf{H}\left(\mathbf{x}\right)=\mathbf{I}_K \tag{P4c}$$

$$\mathbf{x}\in\mathcal{F}. \tag{P4d}$$

需要注意的是，附加约束（P4c）强制用户的上行信号相互正交。与（P1）中复杂的分式形式相比，由于约束（P4c）的作用，（P4）中的目标函数得到了简化。因此，无需重新构造优化问题，即可直接应用块坐标下降框架。具体而言，我们依次优化数字波束成形矩阵 $\mathbf{W}_{\mathrm{BB}}$、模拟波束成形矩阵 $\mathbf{W}_{\mathrm{RF}}$ 以及 PA位置 $\mathbf{x}$。

1）数字波束成形优化：对于固定的 $\mathbf{W}_{\mathrm{RF}}$ 和 $\mathbf{x}$，第 k 个用户的数字波束成形向量的优化子问题可写为：

$$\min_{\mathbf{w}_k}\quad\left\|\mathbf{W}_{\mathrm{RF}}\mathbf{w}_k\right\|^2 \tag{P5a}$$

$$\text{s.t.}\quad\mathbf{w}_k^{\mathsf{H}}\mathbf{W}_{\mathrm{RF}}^{\mathsf{H}}\mathbf{H}\left(\mathbf{x}\right)=\mathbf{e}_k^{\mathsf{T}}, \tag{P5b}$$

其中，$\mathbf{e}_k=[\mathbf{I}_K]_{:,\,k}$ 表示从维数为 $K\times K$ 的单位矩阵中取出的第 k 列，即 $\mathbf{I}_K$。该子问题是凸的，因为它是一个带有仿射约束的二次最小化问题，因此可以得到最优解。根据 Karush-Kuhn-Tucker（KKT）条件，拉格朗日函数可写为：

$$\begin{aligned}\mathcal{L}\left(\mathbf{w}_k\right)=&\,\mathbf{w}_k^{\mathsf{H}}\mathbf{W}_{\mathrm{RF}}^{\mathsf{H}}\mathbf{W}_{\mathrm{RF}}\mathbf{w}_k\\&+\boldsymbol{\lambda}_k^{\mathsf{H}}\left(\check{\mathbf{H}}^{\mathsf{H}}\mathbf{w}_k-\mathbf{e}_k\right)+\left(\mathbf{w}_k^{\mathsf{H}}\check{\mathbf{H}}-\mathbf{e}_k\right)\boldsymbol{\lambda}_k,\end{aligned} \tag{20}$$

其中，等效信道矩阵定义为 $\check{\mathbf{H}}\triangleq\mathbf{W}_{\mathrm{RF}}^{\mathsf{H}}\mathbf{H}\left(\mathbf{x}\right)$，$\boldsymbol{\lambda}_k\in\mathbb{C}^{K\times1}$ 表示拉格朗日乘子。利用 $\mathcal{L}\left(\mathbf{w}_k\right)$ 的驻点条件，可得 $\nabla_{\mathbf{w}_k^*}\mathcal{L}\left(\mathbf{w}_k\right)=\mathbf{W}_{\mathrm{RF}}^{\mathsf{H}}\mathbf{W}_{\mathrm{RF}}\mathbf{w}_k+\check{\mathbf{H}}\boldsymbol{\lambda}_k=\mathbf{0}_{\mathrm{N}}$，其中 $\mathbf{0}_N\in\mathbb{C}^{N\times1}$ 表示全零向量。因此，第 k 个用户的波束成形向量最优解可表示为

$$\mathbf{w}_k=-\left(\mathbf{W}_{\mathrm{RF}}^{\mathsf{H}}\mathbf{W}_{\mathrm{RF}}\right)^{-1}\check{\mathbf{H}}\boldsymbol{\lambda}_k. \tag{21}$$

然后，利用约束 (P5b)，最优拉格朗日乘子$\boldsymbol{\lambda}_{k,\,\star}$ 可计算为

$$\boldsymbol{\lambda}_{k,\star}=-\left(\check{\mathbf{H}}^{\mathsf{H}}\left(\mathbf{W}_{\mathrm{RF}}^{\mathsf{H}}\mathbf{W}_{\mathrm{RF}}\right)^{-1}\check{\mathbf{H}}\right)^{-1}\mathbf{e}_k. \tag{22}$$

根据(21)和(22)，第 k个用户的数字波束成形向量的最优解，即$\mathbf{w}_{k,\,\star}$ 可以推导得到

如下：

$$\mathbf{w}_{k,\star}=\left(\mathbf{W}_{\mathrm{RF}}^{\mathsf{H}}\mathbf{W}_{\mathrm{RF}}\right)^{-1}\check{\mathbf{H}}\left(\check{\mathbf{H}}^{\mathsf{H}}\left(\mathbf{W}_{\mathrm{RF}}^{\mathsf{H}}\mathbf{W}_{\mathrm{RF}}\right)^{-1}\check{\mathbf{H}}\right)^{-1}\mathbf{e}_k.$$

因此，最优数字波束成形矩阵定义为$\mathbf{W}_{\mathrm{BB},\star}=\mathbf{D}^{-1}\check{\mathbf{H}}\left(\check{\mathbf{H}}^{\mathsf{H}}\mathbf{D}^{-1}\check{\mathbf{H}}\right)^{-1}\mathbf{I}_K$，其中$\mathbf{D}\triangleq\mathbf{W}_{\mathrm{RF}}^{\mathsf{H}}\mathbf{W}_{\mathrm{RF}}$。

2）模拟波束成形优化：给定最优数字波束成形矩阵$\mathbf{W}_{\mathrm{BB},\star}$，我们在单位模约束下优化$\mathbf{W}_{\mathrm{RF}}$。关于$\mathbf{W}_{\mathrm{RF}}$的优化问题可表述如下：

$$\max_{\mathbf{W}_{\mathrm{RF}}} f(\mathbf{W}_{\mathrm{RF}})\triangleq\sum_{k=1}^{K}\log_2\left(1+\frac{P/\sigma^2}{\mathrm{tr}\left\{\mathbf{E}_k\left(\check{\mathbf{H}}^{\mathsf{H}}\mathbf{D}^{-1}\check{\mathbf{H}}\right)^{-1}\right\}}\right) \tag{P6a}$$

$$\text{s.t. } |[\mathbf{W}_{\mathrm{RF}}]_{i,j}|=1,\ \forall i,j, \tag{P6b}$$

其中，$\mathbf{E}_k$，$\mathbf{e}_k\mathbf{e}_k^{\mathsf{T}}$表示元素选择矩阵。为处理单位模约束，我们再次采用黎曼流形方法。如第 III-A2 节所述，首先需要计算目标函数的欧几里得梯度，并将其转换到黎曼流形上。对于以下推导，下面这些定义很有用：

$$\mathbf{T}_1\triangleq\mathbf{D}^{-1},\quad \mathbf{T}_2=\check{\mathbf{H}}^{\mathsf{H}}\mathbf{T}_1\check{\mathbf{H}},\quad \mathbf{T}_3=\mathbf{T}_2^{-1} \tag{23a}$$

$$t_k\triangleq[\mathbf{T}_3]_{k,k},\quad c=P/\sigma^2. \tag{23b}$$

首先，$\partial f(\mathbf{W}_{\mathrm{RF}})/\partial r_k$表示为

$$\frac{\partial f(\mathbf{W}_{\mathrm{RF}})}{\partial t_k}=-\frac{c}{\ln(2)t_k(t_k+c)}. \tag{24}$$

因此，通过定义$\mathbf{T}_4\triangleq\mathrm{diag}\{[\partial f/\partial t_1,\ldots,\partial f/\partial t_K]\}$，$df(\mathbf{W}_{\mathrm{RF}})$可将$df(\mathbf{W}_{\mathrm{RF}})=\mathrm{tr}\{\mathbf{T}_4 d(\mathbf{T}_3)\}$紧凑地表示为。根据链式法则，我们进一步有$d\mathbf{T}_3=-\mathbf{T}_3(d\mathbf{T}_2)\mathbf{T}_3$。因此，总梯度可表示为$df=-\mathrm{tr}\{\mathbf{T}_3\mathbf{T}_4\mathbf{T}_3 d\mathbf{T}_2\}$，并可进一步展开如下：

$$df=-I_1-I_2-I_3, \tag{25}$$

其中我们使用如下定义I_1，$\mathrm{tr}\{\mathbf{T}_3\mathbf{T}_4\mathbf{T}_3\left(d\check{\mathbf{H}}^{\mathsf{H}}\right)\mathbf{T}_1\check{\mathbf{H}}\}$，$I_2$，$\mathrm{tr}\left\{\mathbf{T}_3\mathbf{T}_4\mathbf{T}_3\check{\mathbf{H}}^{\mathsf{H}}(d\mathbf{T}_1)\check{\mathbf{H}}\right\}$，以及$I_3$，$\mathrm{tr}\left\{\mathbf{T}_3\mathbf{T}_4\mathbf{T}_3\check{\mathbf{H}}^{\mathsf{H}}\mathbf{T}_1 d\check{\mathbf{H}}\right\}$。因此，$I_1$、$I_2$和$I_3$的精确形式得以推导得出。特别地，基于$\mathbf{H}$的$\check{}$定义，$I_1$和$I_3$可表示如下：

$$I_1=\mathrm{tr}\{\mathbf{T}_1\check{\mathbf{H}}\mathbf{T}_3\mathbf{T}_4\mathbf{T}_3\mathbf{H}^{\mathsf{H}}(d\mathbf{W}_{\mathrm{RF}})\}, \tag{26}$$

$$I_3=\mathrm{tr}\{\mathbf{H}\mathbf{T}_3\mathbf{T}_4\mathbf{T}_3\check{\mathbf{H}}^{\mathsf{H}}\mathbf{T}_1\left(d\mathbf{W}_{\mathrm{RF}}^{\mathsf{H}}\right)\}. \tag{27}$$

因此，下列等式成立：$I_1+I_3=2\Re\{\mathrm{tr}\{\mathbf{T}_1\check{\mathbf{H}}\mathbf{T}_3\mathbf{T}_4\mathbf{T}_3\mathbf{H}^{\mathsf{H}}(d\mathbf{W}_{\mathrm{RF}})\}\}$。随后，基于$d\mathbf{T}_1=-\mathbf{T}_1(d\mathbf{D})\mathbf{T}_1$以及$d\mathbf{D}=d\mathbf{W}_{\mathrm{RF}}^{\mathsf{H}}\mathbf{W}_{\mathrm{RF}}+\mathbf{W}_{\mathrm{RF}}^{\mathsf{H}}d\mathbf{W}_{\mathrm{RF}}$这一事实，项$I_2$可简化为如下形式：

$$\begin{aligned}I_2&=-\mathrm{tr}\left\{\mathbf{W}_{\mathrm{RF}}\mathbf{T}_1\check{\mathbf{H}}\mathbf{T}_3\mathbf{T}_4\mathbf{T}_3\check{\mathbf{H}}^{\mathsf{H}}\mathbf{T}_1 d\mathbf{W}_{\mathrm{RF}}^{\mathsf{H}}\right\}\\&\quad-\mathrm{tr}\left\{\mathbf{T}_1\check{\mathbf{H}}\mathbf{T}_3\mathbf{T}_4\mathbf{T}_3\check{\mathbf{H}}^{\mathsf{H}}\mathbf{T}_1\mathbf{W}_{\mathrm{RF}}^{\mathsf{H}}(d\mathbf{W}_{\mathrm{RF}})\right\}\\&=-2\Re\left\{\mathrm{tr}\left\{\mathbf{T}_1\check{\mathbf{H}}\mathbf{T}_3\mathbf{T}_4\mathbf{T}_3\check{\mathbf{H}}^{\mathsf{H}}\mathbf{T}_1\mathbf{W}_{\mathrm{RF}}^{\mathsf{H}}d\mathbf{W}_{\mathrm{RF}}\right\}\right\}.\end{aligned} \tag{28}$$

因此，欧几里得梯度可写为：

$$\nabla_{\mathbf{W}_{\mathrm{RF}}}f(\mathbf{W}_{\mathrm{RF}})=$$

$$\mathbf{W}_{\mathrm{RF}}\mathbf{T}_1\check{\mathbf{H}}\mathbf{T}_3\mathbf{T}_4\mathbf{T}_3\check{\mathbf{H}}^{\mathsf{H}}\mathbf{T}_1-\mathbf{H}\mathbf{T}_3\mathbf{T}_4\mathbf{T}_3\check{\mathbf{H}}^{\mathsf{H}}\mathbf{T}_1. \tag{29}$$

基于欧几里得梯度，可分别按照式 (15)、(16) 和 (17) 中的相同形式定义黎曼梯度、回缩操作和传输操作。利用这些操作，可以针对式 (P6a) 中的目标函数执行黎曼梯度上升。其过程总结于算法1中，其中共轭方向和目标函数中的负号已作相应替换。

3）夹持波束成形优化：基于已获得的$\mathbf{W}_{\mathrm{BB},\star}$和$\mathbf{W}_{\mathrm{RF},\star}$，关于功率放大器位置优化的剩余问题可通过高斯-赛德尔方法求解。具体步骤已汇总于算法2，为简洁起见，此处省略。

因此，整体算法可概括为算法3，但不再需要更新辅助变量集。由此得到的整体块坐标下降算法的计算复杂度为$\mathcal{O}(I_{\mathrm{iter},1}(M^3+I_{\mathrm{iter}2}M+I_{\mathrm{iter}3}M)$，其中$I_{\mathrm{iter}1}$、$I_{\mathrm{iter}2}$和$I_{\mathrm{iter}3}$分别表示块坐标下降、黎曼梯度下降和高斯-赛德尔搜索所需的迭代次数。

IV. 部分连接结构下基于 SWAN 的三混合波束成形求解方案

尽管全连接结构能够实现最佳性能，但由于模拟波束成形矩阵维度较高，仍面临计算复杂度高、能量效率降低的问题。部分连接结构通过使每条射频链经由移相器连接到一组天线，减少射频链数量，从而克服这些局限。因此，模拟波束成形的维度得以降低，移相器数量随之减少，能耗也相应降低。本节首先介绍一种适用于所提出 SWAN 架构的部分连接结构，然后介绍适用于部分连接结构的三混合波束成形架构。

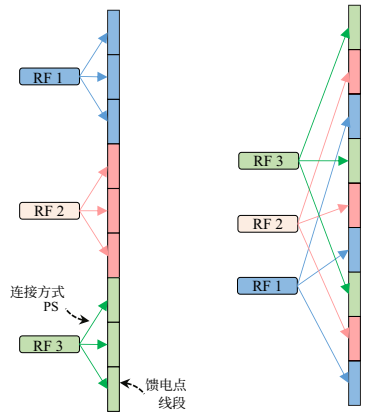
A. SWAN 的部分连接结构

在传统的基于固定位置阵列（FPA）的部分连接（PC）结构中，每条射频链以逐个子阵列的方式连接到馈电点，如图2a所示。如图所示，由一条射频链通过移相器连接多个馈电点构成的子阵列依次排列。因此，射频波束成形矩阵表示为

$$\mathbf{F}_{\mathrm{RF}}^{\mathrm{FPA}}=\begin{bmatrix}\mathbf{c}_1 & \mathbf{0}_{N_{\mathrm{par}}} & \cdots & \mathbf{0}_{N_{\mathrm{par}}}\\ \mathbf{0}_{N_{\mathrm{par}}} & \mathbf{c}_2 & & \\ \vdots & & \ddots & \vdots\\ \mathbf{0}_{N_{\mathrm{par}}} & \mathbf{0}_{N_{\mathrm{par}}} & \cdots & \mathbf{c}_{N_{\mathrm{RF}}}\end{bmatrix}, \tag{30}$$

其中，$\mathbf{c}_i=[\mathrm{e}^{\mathrm{j}\phi_{(i-1)N_{\mathrm{par}}+1}},\ldots,\mathrm{e}^{\mathrm{j}\phi_{iN_{\mathrm{par}}}}]^{\mathsf{T}}\in\mathbb{C}^{N_{\mathrm{par}}\times 1}$表示连接到第$i$条射频链的子阵列，$\phi_i\in(-\pi,\pi]$表示第$i$个移相器的相位。这里，一个子阵列中的天线数量，即$N_{\mathrm{par}}$，表示为$N_{\mathrm{par}}=M/N_{\mathrm{RF}}$。

这种顺序布局常用于传统的固定位置阵列（FPA）场景，因为不同子阵列所遭受的路径损耗几乎相同。相比之下，在



(a) 顺序连接 (b) 交错连接

图 2：部分连接结构示意图：顺序连接与交错连接。

SWAN架构中，每条线段通常长达数米。由于用户分布具有随机性，且SWAN架构长度较大，每个用户可根据其位置由一个或多个子阵列提供服务。因此，顺序布局不可避免地会降低用户间的公平性。为此，我们针对SWAN架构提出了一种新的部分连接（PC）结构，称为交错布局，如图2b所示。具体而言，将连接到不同射频链的线段交替排列，从而确保子阵列线段在整个SWAN架构中均匀分布。

为了从数学上对这种交错连接进行建模，我们定义一个置换矩阵，记为 $\mathbf{\Pi} \in \{0,1\}^{M\times M}$，其中 $M = N_{\mathrm{RF}}N_{\mathrm{par}}$ 表示移相器的总数。在此，我们将$\mathbf{F}_{\mathrm{RF}}^{\mathrm{FPA}}$的行顺序从顺序索引 $j=(j'-1)N_{\mathrm{par}}+i'$ 置换为交错索引 $i=(i'-1)N_{\mathrm{par}}+j'$，其中 $i'=1$、2、……、N_{par} 以及 $j'=1$、2、……、N_{RF}。基于上述定义，索引对集合可定义如下：

$$\mathcal{I}_{\mathrm{pair}} \triangleq \{\{(i'-1)N_{\mathrm{RF}}+j',(j'-1)N_{\mathrm{par}}+i'\} \mid i'=1,2,...,N_{\mathrm{par}},j'=1,2,...,N_{\mathrm{RF}}\} \tag{31}$$

$\mathbf{\Pi}$ 中等于1的第(i,j)个元素表示将$\mathbf{F}_{\mathrm{RF}}^{\mathrm{FPA}}$的第 j行置换到第 i 行的操作。基于这一解释，整个置换矩阵定义如下：

$$[\mathbf{\Pi}]_{i,j} = \begin{cases} 1, & \text{if } i \in \mathcal{I}_{\mathrm{pair}}, \\ 0, & \text{otherwise.} \end{cases} \tag{32}$$

因此，所提出的交错部分连接结构的射频波束成形矩阵，是通过对传统块对角矩阵的馈电点顺序进行置换而得到的，可表示为

$$\mathbf{F}_{\mathrm{RF}}^{\mathrm{SWAN}} = \mathbf{\Pi}\mathbf{F}_{\mathrm{RF}}^{\mathrm{FPA}}, \tag{33}$$

它是 $\mathbf{F}_{\mathrm{RF}}^{\mathrm{FPA}}$ 的一种行置换形式。下文中，为简化记号，我们用 $\mathbf{F}_{\mathrm{RF}}$ 表示 $\mathbf{F}_{\mathrm{RF}}^{\mathrm{SWAN}}$。由于 $\mathbf{F}_{\mathrm{RF}}$ 具有稀疏结构，ZF策略可能对 $\mathbf{F}_{\mathrm{RF}}$ 的条件数较为敏感，从而降低性能。因此，针对这种情况，我们提出一种 WMMSE策略，尽管按照第三节B中的相同步骤，也可以采用 ZF 方法。

B. 问题 (P1) 的基于WMMSE的解

对于交错式PC结构，优化问题可表述如下：

$$\max_{\mathcal{V}} \quad R(\mathbf{F}_{\mathrm{BB}},\mathbf{F}_{\mathrm{RF}},\mathbf{x}) \tag{P7a}$$
$$\text{s.t.} \quad |[\mathbf{F}_{\mathrm{RF}}]_{i,j}| = 1,\ \forall i,j \tag{P7b}$$
$$\mathbf{x} \in \mathcal{F}, \tag{P7c}$$

其中，$\mathbf{F}_{\mathrm{BB}}$ 表示数字波束成形矩阵。为求解该问题，我们采用第 III 节中介绍的基于WMMSE的解。为使本部分内容完整，下面简要介绍其核心步骤。首先，根据加权最小均方误差表达式，可得数字波束成形向量

$$\mathbf{f}_{k,\star} = \left(\sum_{k=1}^{K} P\bar{\mathbf{h}}_k\bar{\mathbf{h}}_k^{\mathsf{H}} + (\sigma^2/P)\mathbf{F}_{\mathrm{RF}}^{\mathsf{H}}\mathbf{F}_{\mathrm{RF}}\right)^{-1}\bar{\mathbf{h}}_k, \tag{34}$$

其中，$\bar{\mathbf{h}}_k$、$\mathbf{F}_{\mathrm{RF}}^{\mathsf{H}}\mathbf{h}_k$(x)。因此，数字波束成形矩阵构造为$\mathbf{F}_{\mathrm{BB},\star} = [\mathbf{f}_{1,\star},...,\mathbf{f}_{K,\star}]$。对于模拟波束成形矩阵的优化，$\mathbf{F}_{\mathrm{RF}}$ 的稀疏结构使得黎曼优化方法无法应用，因为流形的定义会受到破坏。因此，为了充分利用稀疏结构并降低计算复杂度，我们采用逐元素相位匹配方法来确定 $\mathbf{F}_{\mathrm{RF}}$中的移相器取值。具体而言，将 $\mathbf{F}_{\mathrm{BB},\star}$ 代入式 (10) 中的均方误差表达式，可得

$$e_{k,\mathbf{f}_{k,\star}} = P\mathbf{f}_{k,\star}^{\mathsf{H}}\mathbf{F}_{\mathrm{RF}}^{\mathsf{H}}\left(\sum_{j=1}^{K}\mathbf{h}_j\mathbf{h}_j^{\mathsf{H}} + \sigma^2\mathbf{I}_M\right)\mathbf{F}_{\mathrm{RF}}\mathbf{f}_{k,\star} - 2P\Re\left\{\mathbf{f}_{k,\star}^{\mathsf{H}}\mathbf{F}_{\mathrm{RF}}^{\mathsf{H}}\mathbf{h}_k\right\} + P. \tag{35}$$

因此，关于 $\mathbf{F}_{\mathrm{RF}}$ 的子问题可以写成以下形式：

$$\min_{\mathbf{F}_{\mathrm{RF}}} \quad \sum_{k=1}^{K}\omega_k e_{k,\mathbf{f}_{k,\star}} \quad \text{s.t.} \quad \text{(P1b).} \tag{P8a}$$

此外，定义 R $\sum_{j=1}^{K}\mathbf{h}_j\mathbf{h}_j^{\mathsf{H}} + (\sigma^2/P)\mathrm{I}_M$、C $P\sum_{k=1}^{K}\omega_k$ $\mathbf{h}_{k\star}^{\mathsf{H}}$ $2P\sum_{k=1}^{K}\omega_k$ $\mathbf{h}_{k\star}^{\mathsf{H}}$ f, f, 和 B h f, ，均方误差表达式可以改写为更紧凑的形式：

$$\begin{aligned}\sum_{k=1}^{K}\omega_k e_{k,\mathbf{f}_{k,\star}} &= P\sum_{k=1}^{K}\omega_k \mathrm{tr}\left\{\mathbf{F}_{\mathrm{RF}}^{\mathsf{H}}\mathbf{R}\mathbf{F}_{\mathrm{RF}}\mathbf{w}_{k,\star}\mathbf{w}_{k,\star}^{\mathsf{H}}\right\} \\ &\quad - 2P\sum_{k=1}^{K}\Re\left\{\mathrm{tr}\left\{\mathbf{F}_{\mathrm{RF}}^{\mathsf{H}}\mathbf{h}_k\mathbf{f}_{k,\star}^{\mathsf{H}}\right\}\right\} + KP \\ &= P\mathrm{tr}\left\{\mathbf{F}_{\mathrm{RF}}^{\mathsf{H}}\mathbf{R}\mathbf{F}_{\mathrm{RF}}\sum_{k=1}^{K}\mathbf{f}_{k,\star}\mathbf{f}_{k,\star}^{\mathsf{H}}\right\} \\ &\quad - 2P\Re\left\{\mathrm{tr}\left\{\mathbf{F}_{\mathrm{RF}}^{\mathsf{H}}\sum_{k=1}^{K}\mathbf{h}_k\mathbf{f}_{k,\star}^{\mathsf{H}}\right\}\right\} + KP \\ &= \mathrm{tr}\left\{\mathbf{F}_{\mathrm{RF}}^{\mathsf{H}}\mathbf{R}\mathbf{F}_{\mathrm{RF}}\mathbf{C}\right\} - \Re\left\{\mathrm{tr}\left\{\mathbf{F}_{\mathrm{RF}}^{\mathsf{H}}\mathbf{B}\right\}\right\} + KP. \end{aligned} \tag{36}$$

因此，为了在单位模约束下最小化上述方程，我们需要将 $\mathbf{F}_{\mathrm{RF}}$ 中的相位项与R、B和C对齐。为此，我们对PS相位采用逐元素优化策略。具体而言，我们依次更新$\mathbf{F}_{\mathrm{RF}}$ 的各个元素，即每次优化一个元素，同时保持其余元素固定。

为便于叙述，我们定义元素选择矩阵$\mathbf{E}_{m,n}$，将第(m, n)个元素设为1，其余所有元素设为0。因此，为更新$\mathbf{F}_{\mathrm{RF}}$的第(m,n)个元素，模拟波束成形矩阵可写成如下形式

$$\mathbf{F}_{\mathrm{RF}} = \breve{\mathbf{F}}_{m,n} + w\mathbf{E}_{m,n}, \tag{37}$$

其中，$\breve{\mathbf{F}}_{m,n}$ 是通过将 $\mathbf{F}_{\mathrm{RF}}$ 的第 (m,n) 个元素设为零，同时保持其余元素不变而得到的，而w 表示优化变量 $[\mathbf{F}_{\mathrm{RF}}]_{m,n}$。与全连接结构相比，部分连接结构所需的移相器数量更少。因此，（P9）中的逐元素相位校准只需应用于非零参数，从而由于稀疏结构而降低射频波束成形的复杂度。此外，与全连接结构相比，稀疏的部分连接结构能够实现更稳定的相位优化，因为每个相位仅与较少的一组其他相位相耦合。将（37）代入均方误差表达式后，均方误差表达式可改写为逐元素形式，即

$$\begin{aligned}\sum\nolimits_{k=1}^{K} \omega_k e_{k,\mathbf{f}_{k,\star}} &= |w|^2 \mathrm{tr}\left\{\mathbf{E}_{m,n}^{\mathsf{H}}\mathbf{R}\mathbf{E}_{m,n}\mathbf{C}\right\} \\ &+ \Re\left\{\mathrm{tr}\left\{w^*\left(2\mathbf{E}_{m,n}^{\mathsf{H}}\mathbf{R}\breve{\mathbf{F}}_{m,n}\mathbf{C} - \mathbf{E}_{m,n}^{\mathsf{H}}\mathbf{B}\right)\right\}\right\} + c_{1,m,n},\end{aligned}$$

其中，常数项定义为 $c_{1,m,n}\mathbf{tr}\{\breve{\mathbf{F}}_{m,n}^{\mathsf{H}}\mathbf{R}\breve{\mathbf{F}}_{m,n}\mathbf{C}\} - \Re\{\mathbf{tr}\{\breve{\mathbf{F}}_{m,n}^{\mathsf{H}}\mathbf{B}\}\} + KP$。此外，引入以下定义很方便：

$$c_{2,m,n} \triangleq \mathrm{tr}\left\{\mathbf{E}_{m,n}^{\mathsf{H}}\mathbf{R}\mathbf{E}_{m,n}\mathbf{C}\right\}. \tag{38}$$

$$c_{3,m,n} \triangleq \mathrm{tr}\left\{\mathbf{E}_{m,n}^{\mathsf{H}}\mathbf{B}\right\} - 2\mathrm{tr}\left\{\mathbf{E}_{m,n}^{\mathsf{H}}\mathbf{R}\breve{\mathbf{F}}_{m,n}\mathbf{C}\right\}. \tag{39}$$

基于这些定义，关于 $\mathbf{F}_{\mathrm{RF}}$ 的第 (m,n) 个元素的优化问题可表述如下：

$$\min_{w} \quad c_{2,m,n}|w|^2 - \Re\left\{\mathrm{tr}\left\{w^* c_{3,m,n}\right\}\right\} + c_{1,m,n} \tag{P9a}$$

$$\text{s.t.} \quad |w| = 1. \tag{P9b}$$

因此，问题 (P9) 的最优解可按如下方式获得：

$$w_{i,j,\star} = \mathrm{e}^{\mathrm{j}\angle c_{3,m,n}}. \tag{40}$$

将 (40) 应用于 $\mathbf{F}_{\mathrm{RF}}$ 中的所有元素，即可得到优化后的模拟波束成形矩阵 $\mathbf{F}_{\mathrm{RF},\star}$，记为$[\mathbf{F}_{\mathrm{RF},\star}]_{m,n} = w_{i,j,\star}$。随后，使用算法2中的高斯-赛德尔方法再次执行夹持波束成形。除上述步骤外，还根据 (19) 更新辅助变量集 Ω。因此，整体块坐标下降框架遵循算法3，唯一的区别是用所提出的逐元素相位优化替代黎曼梯度下降模块。计算复杂度可推导为 $\mathcal{O}(I_{\mathrm{iter},1}(M^3+N+I_{\mathrm{iter}3}M+K)$，其中$I_{\mathrm{iter}3}$ 归因于逐元素相位优化。

V. 基于SWAN的混合波束成形框架的性能分析

为刻画可达吞吐量随分段数的变化规律，我们考虑两种情形：1）一个射频链配备多个移相器，这是全连接结构的一种特殊情形；2）多个射频链配备一个功率放大器，这是部分连接结构的一种特殊情形。为简化分析，我们考虑单用户情形。

A. 全连接极限：单个射频链与多个移相器

在此情况下，优化变量为一个复值标量，即 $|w|^2 = 1$，以及一个模拟波束成形向量，即 $\mathbf{w}_{\mathrm{RF}} \in \mathbb{C}^{M\times 1}$。因此，优化问题可简化为以下形式：

$$\max_{\mathbf{w}_{\mathrm{RF}},\mathbf{x}} \quad \log_2\left(1 + P\left|\mathbf{w}_{\mathrm{RF}}^{\mathsf{H}}\mathbf{h}\left(\mathbf{x}\right)\right|^2/(M\sigma^2)\right) \tag{P10a}$$

$$\text{s.t.} \quad |[\mathbf{w}_{\mathrm{RF}}]_i| = 1,\ \forall i, \quad \mathbf{x} \in \mathcal{F}, \tag{P10b}$$

为简便起见，我们有意省略了用户索引。因此，在单位模约束下，$\mathbf{w}_{\mathrm{RF}}$ 的最优解可推导为 $\mathbf{w}_{\mathrm{RF},\star} = \mathrm{e}^{\mathrm{j}\angle\mathbf{h}(\mathbf{x})}$。因此，将 $\mathbf{w}_{\mathrm{RF},\star}$ 的表达式代入目标函数后，原优化问题可重写为 $\mathbf{x}_\star = \arg\max_{\mathbf{x}\in\mathcal{F}}(\sum_{m=1}^{M}|[\mathbf{h}(\mathbf{x})]_m|)^2$,，这是一个关于PA位置 $\mathbf{x}$ 的优化问题。此外，目标函数是关于信道增益的单调递增函数。基于上述结论以及式（3）中的信道向量表达式，信道增益为

$$\left(\sum\nolimits_{m=1}^{M}\left|[\mathbf{h}(\mathbf{x})]_m\right|\right)^2 \overset{(a)}{\simeq} \eta\left|\sum\nolimits_{m=1}^{M} 1/r_m\right|^2, \tag{41}$$

其中，(a) 是通过忽略每个线段中的波导内损耗得到的。值得注意的是，正如 [26]，中所证明的，对于适中的分段数，波导内损耗可以忽略不计。具体而言，当总波导长度为 100 米时，$M \geq 9$ 的信道增益损失可确保低于 10 %。因此，在我们的实验设置中，假设波导内损耗可以忽略不计是合理的。由此，最优 PA位置是将 PAs 尽可能靠近用户，以尽量减小每条链路的路径损耗。具体而言，第一步是选取包含用户 x 坐标的、索引为 $m_\star$ 的线段，并在该线段上将 PA 放置于用户的 x 坐标处，即 $x_{m_\star} = r_x$。随后，在满足 PAs 之间的间距约束以及每个线段的坐标范围的条件下，确定右侧的下一个 PA 位置，具体而言，即第 $(m_\star+1)$ 个线段上的位置，其确定方式为

$$x_{m_\star+1} = \max\{x_{m_\star} + \Delta_{\min}, x_{m_\star+1}^{\mathrm{FD}}\}. \tag{42}$$

依次将此方法应用于其余相移器，即可得到最优PA位置向量 $\mathbf{x}_\star$。因此，所得的最大信道增益如下：

$$|\mathbf{h}(\mathbf{x}_\star)|^2 = \eta\left|\sum\nolimits_{m=1}^{M}\left(1/\sqrt{\left(x_{m^\star} - r_x\right)^2 + r_y^2 + H^2}\right)\right|^2. \tag{43}$$

假设目标位于服务区域的中心，则最大信噪比可按如下方式求得：

$$\gamma_{\max}^{(1)} = \frac{P\eta}{M\sigma^2}\left(\frac{1}{\Delta_{yz}} + \sum\nolimits_{\tilde{m}=1}^{\frac{M-1}{2}}\frac{2}{\sqrt{L^2\left(\tilde{m}-0.5\right)^2 + \Delta_{yz}^2}}\right)^2, \tag{44}$$

其中 Δ_{yz}，$(r_y^2+H^2)^{1/2}$。根据欧拉–麦克劳林公式和 $(M-1)L \simeq ML = D_x$ [26]，当 M 较大时，上式可近似表示为：

$$\gamma_{\max}^{(1)} \simeq \frac{P\eta}{M\sigma^2}\left(\frac{1}{\Delta_{yz}}+\frac{2}{L}\sinh^{-1}\left(\frac{LM}{2\Delta_{yz}}\right)\right)^2. \quad (45)$$

因此，速率由下式给出

$$R_{\max}^{(1)} = \log_2(1+\gamma_{\max}^{(1)}). \quad (46)$$

注1.（关于分段数 M的速率缩放定律）基于(45)和(46)，对于给定的分段长度L，速率是关于分段数 M 的非单调函数。更具体地说，速率首先随 M 增大，直至 $M_\star = 2\Delta_{yz}^2/L$；此后，随着 M 进一步增大，速率下降。原因有两方面：1）随着分段数增加，聚合噪声功率（即 $M\sigma^2$）相应增加，从而降低信噪比；2）在总长度固定的情况下，新增分段逐渐远离用户，对波束增益增强的贡献递减。因此，增加的噪声功率抵消了波束成形增益，导致可达速率降低。当 M足够大时，我们有 $R_{\max}^{(1)} = \mathcal{O}((\log M)^2/M)$，并且当 $M\to\infty$时有 $R_{\max}^{(1)} \to 0$。

B. 部分连接极限：多个射频链配备单个移相器

本部分考虑单用户场景，其中射频链数量等于分段数，且每条链通过一个移相器连接到馈电点。在这种情况下，由于相移和幅度调节可通过数字波束成形向量 $\|\mathbf{w}\|=1$自由调整，因此我们可以将模拟波束成形向量 $\mathbf{w}_{\mathrm{RF}}$ 视为信道向量的固有部分，并省去对其进行优化。我们不失一般性地将移相器的相位设为零。因此，优化问题可表述如下：

$$\max_{\mathbf{w},\mathbf{x}} \quad \log_2(1+|\mathbf{w}^{\mathsf{H}}\mathbf{h}(\mathbf{x})|^2/(\sigma^2\|\mathbf{w}\|^2)) \quad \text{(P11a)}$$

$$\text{s.t.} \quad \mathbf{x}\in\mathcal{F}. \quad \text{(P11b)}$$

对于固定的功率放大器位置，最优波束成形向量遵循最大比合并（MRC）准则，即 $\mathbf{w}_\star = \frac{\mathbf{h}(\mathbf{x})}{\|\mathbf{h}(\mathbf{x})\|}$，其最优性可直接由柯西-施瓦茨不等式推出。因此，该问题可等价地写为 $\mathbf{x}_\star = \max_{\mathbf{x}\in\mathcal{F}}\|\mathbf{h}(\mathbf{x})\|^2$，其中信道增益可表示如下：

$$\|\mathbf{h}(\mathbf{x})\|^2 \overset{(a)}{\simeq} \eta\sum_{m=1}^{M} 1/r_m^2. \quad (47)$$

这里我们再次忽略波导内损耗。与第V节-A中讨论的前一种情形一样，最优PA位置，即 $\mathbf{x}_\star$，是通过将相移器尽可能靠近目标来确定的，因为这样可以使路径损耗最小。因此，最大化的信道增益为

$$\gamma_{\max}^{(2)} = \frac{P\eta}{\sigma^2}\left(\frac{1}{\Delta_{yz}^2}+\sum_{\tilde{m}=1}^{\frac{M-1}{2}}\frac{2}{L^2(\tilde{m}-0.5)^2+\Delta_{yz}^2}\right). \quad (48)$$

表 I：仿真参数。

参数（符号）	值
载波频率（f）	28 GHz
有效折射率（n_{eff}）	1.4
波导内损耗（κ）	0.08 dB/m
发射功率（P）	10 分贝毫瓦
噪声功率（σ^2）	−80 分贝毫瓦
服务区域尺寸 ($D_x \times D_y$)	80 m × 20 m
接收机高度 (H)	3 m
用户数 (K)	4（均匀分布）
射频链数量（三混合）(N_{RF})	25
分段数（三混合）（M）	50
夹持天线数量（N_{PA}）	50
高斯-赛德尔搜索分辨率	0.01 米
最小功率放大器夹持天线间距（$\Delta_{\min}$）	$\lambda_c/2$
块坐标下降停止准则（ϵ）	分数下降 $< 10^{-8}$

当奇数 M 足够大时，该方程可近似为

$$\gamma_{\max}^{(2)} \simeq \frac{P\eta}{\sigma^2}\left(\frac{1}{\Delta_{yz}^2}+\frac{2}{L\Delta_{yz}}\tan^{-1}\left(\frac{(M-1)L}{2\Delta_{yz}}\right)\right). \quad (49)$$

相应地，速率可表示为

$$R_{\max}^{(2)} = \log_2(1+\gamma_{\max}^{(2)}). \quad (50)$$

注 2.（关于分段数的速率缩放定律M）根据式 (49) 和式 (50)，当分段长度固定时，速率随 M 单调递增，但有上界L。速率随 M 增大而提升，并收敛至 $\gamma_{\max}^{(2,M\to\infty)} = \frac{P\eta}{\sigma^2}\left(\frac{1}{\Delta_{yz}^2}+\frac{\pi}{L\Delta_{yz}}\right)$。与注 1 中的前一种情况相比，在前一种情况下，可达速率会随 M 下降；而本情况下有界的速率缩放定律源于最大比合并波束成形对噪声进行非相干合并，因此有效噪声功率不会随分段数增加。因而，随着额外线段逐渐远离用户位置，这些线段带来的速率增益会因路径损耗增大而逐渐减小，使速率收敛到一个固定值。对于足够大的 M，有 $R_{\max}^{(2,M\to\infty)} = \lim_{M\to\infty} R_{\max}^{(2)} = \log_2\left(\frac{P\eta}{\sigma^2}\left(\frac{1}{\Delta_{yz}^2}+\frac{\pi}{L\Delta_{yz}}\right)\right)$。换言之，速率有界，即 $R_{\max}^{(2,M\to\infty)} \propto \mathcal{O}(1)$。

VI. 数值结果

除非另有说明，全文仿真均采用表 I 中的参数。此外，基于 WMMSE 的方法使用通过基于迫零的算法获得的波束成形器进行初始化。本节所 प्रस्तुत的数值结果基于 1000 次蒙特卡洛试验。为验证所提方法的有效性，考虑以下基准方法：

- 采用全连接结构和 WMMSE 的大规模MIMO（大规模 MIMO、全连接、WMMSE）：该基准采用传统的大规模MIMO接收机为多个用户提供服务。接收机配备了 $M = 50$ 根天线和$N_{\text{射频}} = 25$ 条射频链，阵列位于服务区域的中心。该基准的优化过程遵循第 III 节中基于 WMMSE 的方法。与所提出的三混合设计不同，该基准移除了夹持波束成形步骤，

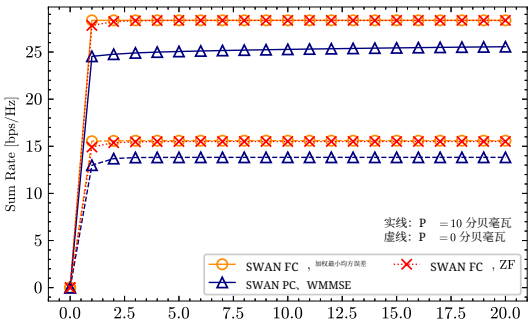


图 3： $P = 0$ 分贝毫瓦和 $P = 10$ 分贝毫瓦下的收敛行为示意图。

从而能够单独考察并突出三混合波束成形的优势。

- 无 IAR 的传统 PASS（无 IAR 的传统 PASS）：此基准方案采用传统 PASS 接收机为多个用户提供服务。与基于 SWAN 的架构不同，PASS 使用一根沿 x 方向延伸、覆盖整个服务区域的长波导。夹持天线的数量设为 $M = 50$，而受传统 PASS 限制，射频链数量限制为 $N_{\text{射频}} = 1$ 。更重要的是，我们重点研究不考虑 IAR 效应的理想情形。

与基准方案并列，“SWAN FC、WMMSE”和“SWAN PC、WMMSE”分别指针对全连接和部分连接结构提出的基于 SWAN 的三重混合 WMMSE 波束成形设计。“SWAN FC、ZF”则指针对全连接架构提出的基于 SWAN 的三重混合 ZF 波束成形设计。

A. 收敛行为与发射功率的影响

图 3 展示了所提算法在不同发射功率下的收敛行为。可以观察到，所提算法在 10 次迭代内迅速收敛，这表明所提方法的有效性。此外，由于部分连接结构带来的自由度损失，“SWAN PC、WMMSE”与“SWAN FC、WMMSE”结果之间存在性能差距。在所考虑的上行链路混合波束成形架构中，模拟合并器限制了数字合并器可用的有效自由度。因此，加权最小均方误差相对于迫零所提供的额外优化增益变得微不足道，从而产生几乎相同的性能。

在图 4 中，我们研究了和速率随发射功率的变化情况。随着发射功率增加，和速率也相应提高。由于缺少捏缩波束成形，即三混合波束成形架构的第三层，“大规模多输入多输出全连接结构，加权最小均方误差”基准的性能逊于所提出的方法。此外，由于射频链数量有限，“Conv. PASS”基准无法有效抑制多用户干扰（MUI），因此在所考虑的多用户场景中表现不佳。

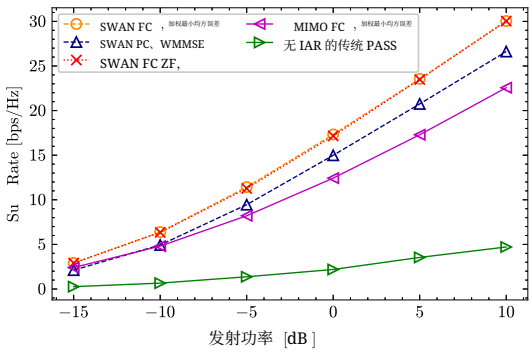


图 4：和速率随发射功率变化的示意图。

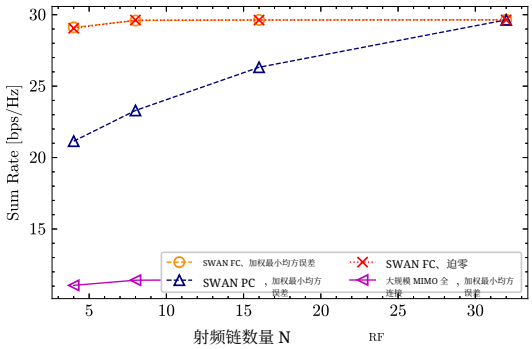


图 5：和速率与射频链数量的关系示意图。

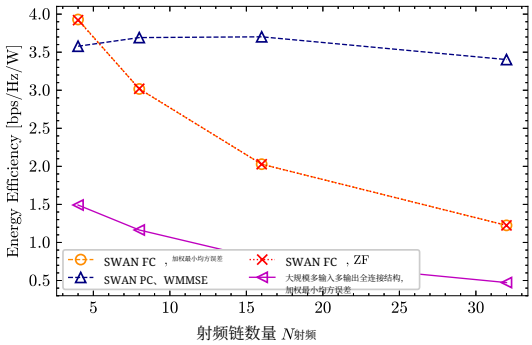
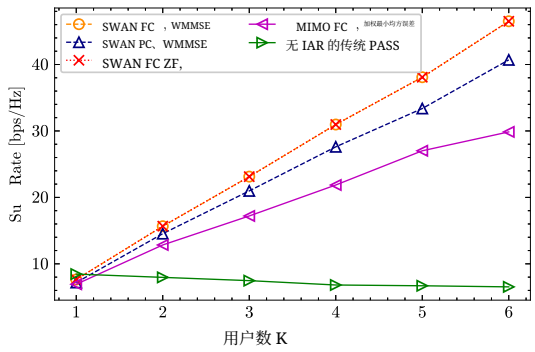


图 6：能量效率与射频链数量的关系示意图。

B. 射频链数量的影响

图 5 和图 6 展示了和速率与能量效率（能效）随射频链数量增加的变化情况。在本小节中，我们将分段数设为 $M = 32$，以确保部分连接结构中每个子阵列的分段数（即 N_{par}）为整数。此外，为了评估能效，我们采用了 [29], [36] 中的能耗模型。具体而言，设一个功率放大器、一个移相器和一条射频链的能耗分别为 $P_{\text{PA}} = 100$ mW、$P_{\text{PS}} = 10$ mW 和 $P_{\text{RF}} = 100$ mW。能量效率定义为 EE $= \frac{R}{P + N_{\text{RF}} P_{\text{RF}} + M P_{\text{PA}} + N_{\text{PS}} P_{\text{PS}}}$，其中 N_{PS} 表示移相器总数；对于全连接结构，该数量为 $N_{\text{PS}} = M \times N_{\text{RF}}$，对于部分连接结构，该数量为 $N_{\text{PS}} = M$。

如图 5 所示，随着射频链数量增加，干扰抑制能力增强，和速率也随之提高。此外，由于采用部分连接，



Fi图 7：和速率与用户数量的关系示意图 ers.

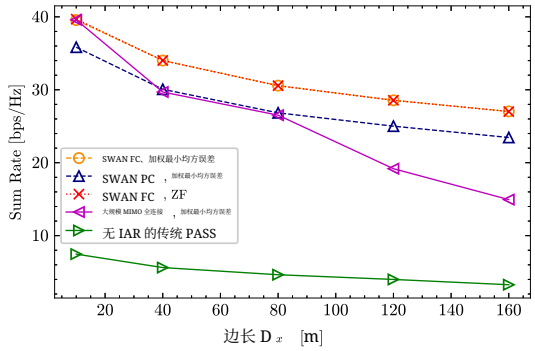


图8：和速率随边长的变化示意图。

所提出的“SWAN PC、WMMSE”与其他方法相比存在性能差距。此外，传统的大规模多输入多输出基线表现最差，这凸显了夹持波束成形的优势。从能效方面来看，如图 6 所示，由于分段数量较多且发射功率较低，随着 N_{RF} 增加，全连接结构的能效会下降。相比之下，部分连接通过采用稀疏模拟波束成形矩阵来降低硬件能耗，因此随着 N_{RF} 增加，其性能优于全连接结构。

C. 用户数量的影响

图7展示了和速率随用户数增加的变化情况。可以观察到，随着用户数增加，可实现的和速率也随之提高。然而，用户数越多，PC结构与FC结构之间的差距就越大。这是因为随着用户数增加，稀疏模拟波束成形矩阵自由度的减少会限制系统的整体性能。对于基准方案，由于缺乏信道可重构性，大规模多输入多输出框架的性能逊于所提方法。传统PASS的性能有限，这是因为它采用单射频链结构。

D. 边长的影响

图8展示了可达和速率随服务区域边长变化的情况。具体而言，我们增大 x方向的边长，同时保持 y方向的边长不变。此外，随着边长增大，在分段数 M固定的情况下，SWAN接收机的总长度也随之增加。随着边

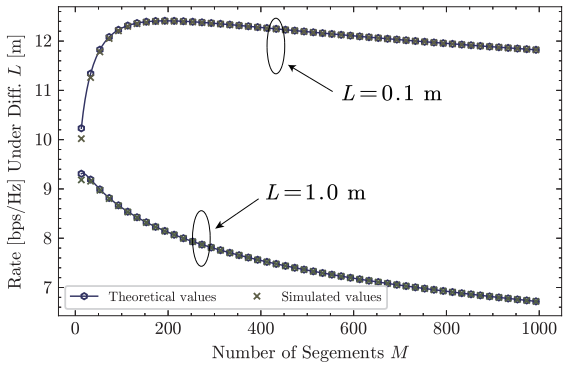


图9：单射频链和多个移相器下的FC极限示意图。

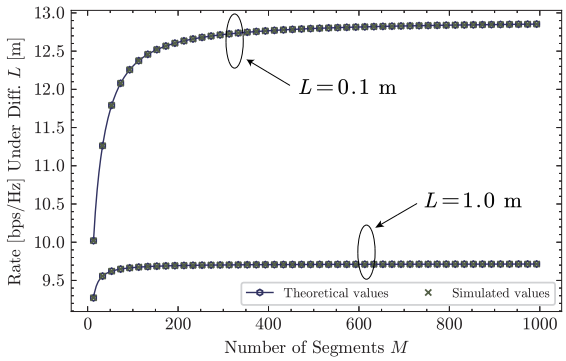


图10：多条射频链和单个移相器下的PC极限示意图。

长增加，可达和速率会降低。由于分段数保持不变，单位长度内的射频链数量减少，这表明信号处理能力有所下降。不过，与传统大规模多输入多输出基线相比，SWAN接收机的可达和速率降幅更为平缓。

E. 分段数量的速率缩放规律

本节中，我们验证全连接结构和部分连接结构的速率缩放规律。为此，如第V节所述，我们考虑单用户情形。具体而言，图9展示了采用单个射频链和多个相移器的全连接结构的速率缩放规律。如图所示，理论推导结果与仿真结果高度吻合。随着分段数增加，可达速率呈现先升后降的趋势，这与我们在第V节中的分析一致。对于给定的分段数，随着分段长度增加，可达速率会降低，因为SWAN两端的更多分段距离用户位置更远。此外，最优分段数会随着每段长度增加而减少，这表明距离用户较远的分段所产生的贡献会逐渐减小。

图10展示了具有多条射频链和单个移相器的部分连接结构的速率缩放定律。随着分段数增加，速率上升，但仍受各自渐近值的限制。此外，随着分段长度增加，速率下降，因为更远的分段对整体速率最大化的贡献较小。最后，

理论结果与仿真结果之间的一致性证实了我们推导的正确性。

VII. 结论

本文研究了在全连接（FC）和部分连接（PC）拓扑下，基于 SWAN 的三重混合波束成形，以实现上行多用户 MIMO 和速率最大化。针对全连接结构，本文提出了基于加权最小均方误差（WMMSE）和迫零（ZF）的三重混合波束成形方案，用于上行和速率最大化，并采用块坐标下降框架优化数字波束成形、采用黎曼流形优化进行模拟波束成形优化，以及采用基于高斯-赛德尔的方法重新定位功率放大器（PA）。针对部分连接结构，本文为 SWAN 提出了一种交错拓扑，并设计了一种基于 WMMSE 的三重混合方案。为充分利用部分连接结构的稀疏模拟架构，本文采用逐元素相位校准，以降低块坐标下降框架中的复杂度。此外，在简化的场景假设下，本文针对全连接和部分连接结构推导了速率缩放定律与分段数之间的关系。分析结果表明，增加分段数未必能提高可达速率。大量数值结果验证了分析结论，并表明对于全连接和部分连接结构，所提出的基于 SWAN 的三重混合波束成形始终优于采用混合波束成形的传统大规模 MIMO，以及采用仅收缩波束成形的传统 PASS。

参考文献

[1] L. Lu、G. Y. Li、A. L. Swindlehurst、A. Ashikhmin 和 R. Zhang，“大规模 MIMO 概述：优势与挑战”，IEEE Journal of Selected Topics in Signal Processing，第 8 卷，第 5 期，第 742–758 页，2014 年 10 月。[2] A. Alkhateeb 等人，“毫米波系统的 MIMO 预编码与合并方案”，IEEE Communications Magazine，第 52 卷，第 12 期，第 122–131 页，2014 年 12 月。[3] B. Ning 等人，“5G NR协议中的预编码矩阵指示器：3GPP波束成形码本教程”，IEEE Commun. Surveys Tuts.，第 1–1 页，2026 年，提前访问，doi:10.1109/COMST.2026.3653568。[4] M. Shafi 等人，“5G：标准、试验、挑战、部署与实践教程概述”，IEEE Journal on Selected Areas in Communications，第 35 卷，第 6 期，第 1201–1221 页，2017 年 6 月。[5] Y. Liu 等人，“近场通信：全面综述”，IEEE Commun. Surveys Tuts.，第 27 卷，第 3 期，第 1687–1728 页，2025 年 6 月。[6] E. Bj¨ornson 等人，“从天线丰裕到天线智能：面向 6G 超大规模 MIMO 系统”，arXiv 预印本 arXiv:2601.08326，2026 年。[7] T. Gong 等人，“全息MIMO通信：理论基础、赋能技术与未来方向”，IEEE Commun. Surveys Tuts.，第 26 卷，第 1 期，第 196–257 页，2024 年第 1 季度。[8] M. Liu、M. Li、R. Liu、Q. Liu 和 A. L. Swindlehurst，“可重构天线阵列：连接电磁学与信号处理”，arXiv 预印本 arXiv:2510.17113，2025 年。[9] J. Robert W. Heath、J. Carlson、N. V. Deshpande、M. R. Castellanos、M. Akrout 和 C.-B. Chae，“三混合MIMO架构”，arXiv 预印本 arXiv:2505.21971，2025 年。[10] M. R. Castellanos、S. Yang、C.-B. Chae 和 R. W. Heath，“在三混合MIMO架构中采用可重构天线，面向 6G 及未来发展”，IEEE Transactions on Communications，第 74 卷，第 381–401 页，2026 年，提前访问，doi:10.1109/TCOMM.2025.3621272。[11] L. Zhu、W. Ma 和 R. Zhang，“无线通信中的可移动天线：机遇与挑战”，IEEE Communications Magazine，第 62 卷，第 6 期，第 114–120 页，2024 年 6 月。[12] W. K. New、K.-K. Wong、H. Xu 等人，“6G网络流体天线系统教程：涵盖通信理论、优化方法与硬件设计”，IEEE Commun. Surveys Tuts.，第 27 卷，第 4 期，第 2325–2377 页，2025 年 8 月。[13] N. Shlezinger、G. C. Alexandropoulos、M. F. Imani、Y. C. Eldar 和 D. R. Smith，“面向 6G 极端大规模 MIMO 通信的动态超表面天线”，IEEE Wireless Communications，第 28 卷，第 2 期，第 106–113 页，2021 年 4 月。[14] A. Fukuda 、 H. Yamamoto 、 H. Okazaki 、 Y. Suzuki 和 K. Kawai，“捏缩天线：将介质波导用作天线”，NTT DOCOMO Technical Journal，第 23 卷，第 3 期，第 5–12 页，2022 年 1 月。[15] Z. Ding、R. Schober 和 H. Vincent Poor，“柔性天线系统：从捏缩天线视角出发”，IEEE Transactions on Communications，第 73 卷，第 10 期，第 9236–9253 页，2025 年 10 月。[16] Z. Wang 等人，“ 捏缩天线系统的建模与波束成形优化 ”，IEEE Transactions on Communications，第 73 卷，第 12 期，第 13 904–13 919 页，2025 年 10 月。[17] A. Bereyhi、C. Ouyang、S. Asaad、Z. Ding 和 H. V. Poor，“MIMO-PASS：通过 MIMO 捏缩天线系统进行上行与下行传输”，arXiv 预印本 arXiv:2503.03117，2025 年。[18] M. Sun、C. Ouyang、S. Wu 和 Y. Liu，“捏缩天线系统的多用户波束成形：逐元素优化框架”，IEEE Transactions on Wireless Communications，第 25 卷，第 6538–6552 页，2026 年，提前访问，doi:10.1109/TWC.2025.3625377。[19] X. Xu、X. Mu、Z. Wang、Y. Liu 和 A. Nallanathan，“捏缩天线系统（PASS）：功率辐射模型与最优波束成形设计”，arXiv 预印本 arXiv:2505.00218，2025 年。[20] J. Zhao 等人，“捏缩天线系统赋能的多用户通信：传输结构与波束成形优化”，IEEE Transactions on Communications，第 74 卷，第 2316–2330 页，2026 年。[21] J. Guo、Y. Liu 和 A. Nallanathan，“捏缩天线系统（PASS）中的三波束成形图变换器”，IEEE Journal on Selected Areas in Communications，第 1–1 页，2025 年，提前访问，doi:10.1109/JSAC.2025.3645715。[22] Y. Liu、H. Jiang、X. Xu 等人，“捏缩天线系统（PASS）：教程”，IEEE Transactions on Communications，第 74 卷，第 4881–4918 页，2026 年 1 月。[23] A. F. Molisch、V. V. Ratnam、S. Han、Z. Li、S. L. H. Nguyen、L. Li 和 K. Haneda，“大规模 MIMO 的混合波束成形：综述”，IEEE Communications Magazine，第 55 卷，第 9 期，第 134–141 页，2017 年。[24] C.-J. Zhao、Z. Wang、H. Shin 和 Y. Liu，“全连接捏缩天线系统的三重混合波束成形设计”，arXiv 预印本 arXiv:2511.14517，2025 年。[25] Z. Cheng、C. Ouyang 和 N. Marchetti，“采用捏缩天线的三重混合波束成形性能”，arXiv 预印本 arXiv:2511.01099，2025 年。[26] C. Ouyang 等人，“分段波导赋能的捏缩天线系统（SWAN）中的上行和下行通信”，第 74 卷，第 3688–3703 页，2026 年 1 月。[27] H. Jiang 等人，“用于通感一体化的分段波导赋能的捏缩天线系统（SWAN）”，arXiv 预印本 arXiv:2512.07649，2025 年。[28] T. Hou 等人，“上行捏缩天线系统（PASS）的性能”，IEEE Transactions on Communications，2025 年，提前访问，doi:10.1109/TCOMM.2025.3618726。[29] X. Yu 等人，“毫米波 MIMO 系统混合预编码的交替最小化算法”，IEEE Journal of Selected Topics in Signal Processing，第 10 卷，第 3 期，第 485–500 页，2016 年 4 月。[30] C. Ouyang、Z. Wang、Y. Liu 和 Z. Ding，“捏缩天线系统（PASS）的阵列增益”，IEEE Communications Letters，第 29 卷，第 6 期，第 1471–1475 页，2025 年 6 月。[31] Q. Shi 等人，“针对 MIMO 干扰广播信道分布式效用和最大化的迭代加权最小均方误差方法”，IEEE Transactions on Signal Processing，第 59 卷，第 9 期，第 4331–4340 页，2011 年 9 月。[32] T. Lin 等人，“采用最小均方误差准则的毫米波系统混合波束成形”，IEEE Transactions on Communications，第 67 卷，第 5 期，第 3693–3708 页，2019 年 5 月。[33] A. Wiesel、Y. C. Eldar 和 S. Shamai，“迫零预编码与广义逆”，IEEE Transactions on Signal Processing，第 56 卷，第 9 期，第 4409–4418 页，2008 年 9 月。[34] H. Q. Ngo、E. G. Larsson 和 T. L. Marzetta，“超大规模多用户 MIMO 系统的能效与频谱效率”，IEEE Transactions on Communications，第 61 卷，第 4 期，第 1436–1449 页，2013 年 4 月。[35] X. Su 和 Y. Jiang，“面向和速率最大化的最优迫零混合下行预编码”，IEEE Wireless Communications Letters，第 11 卷，第 3 期，第 463–467 页，2022 年 3 月。[36] H. Jiang 等人，“MIMO 流体天线系统的波束对准”，IEEE Journal on Selected Areas in Communications，2025 年，提前访问，doi:10.1109/JSAC.2025.3618212。