arXiv:2302.02401v1 [cs.IT] 5 Feb 2023

迈向高效子阵列混合波束成形：基于注意力网络的FDD大规模多用户MIMO系统实用反馈

Zhilin Lu、Xudong Zhang、Rui Zeng 和 Jintao Wang，*IEEE* 高级会员

摘要—由于信道不互易，频分双工（FDD）多输入多输出（MIMO）系统需要进行信道状态信息（CSI）反馈。借助深度学习，已有许多研究成功重建了大规模 MIMO 的压缩理想 CSI。然而，仅进行 CSI 重建的实用性有限，因为这些方法没有考虑信道估计和目标波束成形设计。本文提出一种联合优化的网络，用于信道估计和反馈，从而学习频谱高效的波束成形器。此外，本文采用易于部署的子阵列混合波束成形架构，并专门设计了一种实用、轻量级的端到端网络。实验表明，与此前最先进的方法相比，所提出的网络在资源受限的用户设备上轻量十倍以上，而性能损失仅为轻微程度。

关键词—大规模MIMO、子阵列混合波束成形、CSI反馈、注意力网络、低复杂度

一、引言

MASSIVE 多输入多输出（MIMO）被广泛认为是未来无线通信系统中的一项重要技术。基站（BS）需要下行信道状态信息（CSI）来进行波束成形，从而实现更高的频谱效率 [1]。在频分双工（FDD）模式下，由于信道非互易性，需要在用户设备（UE）处估计下行链路信道状态信息（CSI），并将其反馈给基站。然而，由于大规模MIMO系统的天线规模庞大，直接反馈的开销大到难以接受。

自 CsiNet [2]展示出相较于传统压缩感知算法的显著优势以来，深度学习（DL）辅助的CSI压缩反馈便受到了广泛关注。然而，许多现有研究（如 CRNet [3] ）都假设信道估计是理想的，仅致力于改善CSI重构，因此无法保证波束成形获得更高的和速率。因此，在实际反馈中，应同时考虑信道估计和波束成形。

对于传统的基于块的通信系统，很难对相邻模块（如 CSI反馈和波束成形）进行联合优化。相反，将相邻模块的深度学习网络串联起来，并以端到端的方式联合优化整个流程，是一种自然的做法。例如，构建用于联合信道估计和反馈的自编码器，以提高 [4],中的信道重建精度；同时，在 [5]中联合优化反馈和波束成形，以获得更高的波束成形增益。

此外，一些研究尝试联合优化信道估计、反馈和波束成形（EFB），以系统性地提升波束成形和速率。文献 [6] 针对多用户 MIMO（MU-MIMO）系统实现了一种数字波束成形辅助的 EFB 流程。此外，文献 [7] 和 [8] 提出了适用于混合波束成形架构的 EFB 网络。然而，现有的 EFB 联合优化研究并不利于实际部署，因为它们仅考虑全连接波束成形结构，且所设计的网络规模较大。

本文着重于提升基于深度学习的 EFB 联合优化方法的实用性。主要贡献总结如下。

- 本文首次将子连接混合波束成形结构引入 EFB 联合优化任务。通过减少所需移相器（PSs）的数量，提高了能量效率。
- 为提升实用性，本文设计了一种低复杂度 EFB 网络。与以往最先进的方法相比，所提出的 EFBAttnNet 将资源受限用户设备处的编码器复杂度降低了 10 倍，而性能损失很小。

本文其余部分安排如下：第二节介绍系统模型，第三节阐述 EFB 问题的构建。第四节提出轻量级 EFBAttnNet，第五节分析实验结果。最后，第六节给出结论。

II. 系统模型

我们考虑一个包含一个基站和 K 个独立用户设备的频分双工大规模多用户 MIMO 系统。基站配备 N_t 根天线和 K 条射频（RF）链，而每个用户设备配备一根天线。第 k^{th} 个用户接收到的信号可表示为：

$$y_k = \mathbf{h}_k^H \mathbf{w}_k s_k + \sum_{i=1, i\neq k}^{K} \mathbf{h}_k^H \mathbf{w}_i s_i + z_k, \tag{1}$$

其中，$\mathbf{h}_k \in \mathbb{C}^{N_t \times 1}$ 和 $\mathbf{w}_k \in \mathbb{C}^{N_t \times 1}$ 分别表示该用户的下行链路信道向量和等效波束成形向量

作者来自清华大学电子工程系和北京信息科学与技术国家研究中心（BNRist），中国北京 100084。（电子邮箱：luzl18@mails.tsinghua.edu.cn、zxd22@mails.tsinghua.edu.cn、zengr21@mails.tsinghua.edu.cn、wangjintao@tsinghua.edu.cn）。主要研究结果可通过以下 GitHub 链接复现：https://github.com/Kylin9511/EFBAttnNet。

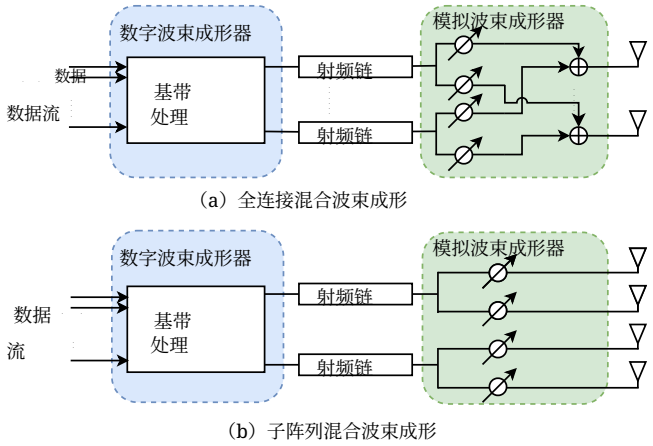


图 1. 全连接混合波束成形与子阵列混合波束成形的架构对比。

k^{th} 。 $s_k \in \mathbb{C}$ 表示发送给k^{th} 用户的符号，而 $z_k \sim \mathcal{CN}\left(0, \sigma^2\right)$ 表示接收端的加性白高斯噪声。

以往关于混合波束成形中 EFB 联合优化的研究主要关注图 1（a）所示的全连接架构。为了提高能量效率，我们首先将子阵列波束成形结构引入 EFB 流程。如图 1（b）所示， N_t 根天线均匀分配给 K 条射频链。每根天线只连接到一条射频链，而非所有射频链，因此所需的移相器数量大幅减少。因此，等效波束成形器$\mathbf{w}_k$ 遵循以下形式。

$$\mathbf{w}_k = \mathbf{ACd}_k, \tag{2}$$

其中，模拟波束成形器$\mathbf{A} \triangleq \operatorname{diag}\left(e^{j\boldsymbol{\theta}_\mathbf{A}}\right) \in \mathbb{C}^{N_t \times N_t}$ 是一个恒模对角矩阵， $\boldsymbol{\theta}_\mathbf{A} \in \mathbb{R}^{N_t \times 1}$ 是相位向量。同时，每个用户都有自己的数字波束成形器$\mathbf{d}_k \in \mathbb{C}^{K \times 1}$ ，我们将整体数字波束成形器表示为$\mathbf{D} \triangleq [\mathbf{d}_1, \ldots, \mathbf{d}_K] \in \mathbb{C}^{K \times K}$。天线连接器$\mathbf{C}$ 是一个块对角矩阵，它将每根天线连接到一条射频链，从而确保采用子阵列架构。

$$\mathbf{C} = \mathbf{I}_K \otimes \mathbf{1}_{N_m} \in \{0,1\}^{N_t \times K}, \tag{3}$$

其中$\mathbf{I}_K \in \mathbb{R}^{K \times K}$ 是单位矩阵， $\mathbf{1}_{N_m} \in \mathbb{R}^{N_m \times 1}$ 是全一向量，$N_m = N_t/K$ 表示连接到单个射频链的天线数量。

对于信道$\mathbf{h}_k$，本文采用影响深远的 Saleh-Valenzuela (SV) 聚类模型，如公式 (4) 所示。

$$\mathbf{h}_k = \frac{1}{L_{\text{path}}} \sum_{p=1}^{L_{\text{path}}} \alpha_{p,k} \mathbf{a}_t(\phi_{p,k}), \tag{4}$$

其中，$\alpha_{p,\ k}$ 和$\phi_{p,\ k}$ 分别是 k^{th} 用户在 p^{th} 路径上的复增益和出发角（AoD）。采用均匀线性阵列（ULA）模型，因此发射端响应向量$\mathbf{a}_t$ 设为：

$$\mathbf{a}_t(\phi) = \left[1, e^{j\frac{2\pi}{\lambda} d \sin(\phi)}, \ldots, e^{j\frac{2\pi}{\lambda} d(N_t-1)\sin(\phi)}\right]^T \tag{5}$$

III. 基于深度学习的EFB流程

基于深度学习的EFB流程从下行导频训练开始，在此过程中学习一个无偏置的全连接层。根据子阵列混合波束成形架构，全连接层的权重 $\boldsymbol{\theta}_{\tilde{\mathbf{X}}} \in \mathbb{R}^{N_t \times L}$ 用作 L 个不同导频的相位，并采用归一化幅度。

$$\tilde{\mathbf{X}} = \sqrt{P/N_t} e^{j\boldsymbol{\theta}_{\tilde{\mathbf{X}}}} = \sqrt{P/N_t}\left(\cos(\boldsymbol{\theta}_{\tilde{\mathbf{X}}}) + j\sin(\boldsymbol{\theta}_{\tilde{\mathbf{X}}})\right), \tag{6}$$

其中， P 是发射功率。如图 2 所示，所有˜用户共享相同的导频 $\mathbf{X}$。SubarrayPilotNet 实现的下行链路导频传输可描述为:

$$\tilde{\mathbf{y}}_k = \mathbf{h}_k^H \tilde{\mathbf{X}} + \tilde{\mathbf{z}}_k, \tag{7}$$

其中， ˜$\mathbf{z}_k \sim \mathcal{CN}(\mathbf{0}, \sigma^2\mathbf{I}_L)$ 是加性白高斯噪声。

对于用户设备端的编码器，信道估计和信道状态信息压缩被隐式地一并完成。如图 2 所示，反馈比特流 $\mathbf{q}_k \in \{\pm 1\}^{B \times 1}$ 直接从 ˜$\mathbf{y}_k$中学习得到。

$$\mathbf{q}_k = \mathcal{E}(\tilde{\mathbf{y}}_k, \Theta_\mathcal{E}); \ \ k \in \{1, \ldots, K\}, \tag{8}$$

其中， $\mathcal{E}(\cdot)$ 表示 EFB 编码器网络， $\Theta_\mathcal{E}$ 是编码器的可学习参数。

对于基站处的解码器，输入 $\mathbf{Q} = [\mathbf{q}_1, \ldots, \mathbf{q}_K]$汇集了用户发送的比特流。相位向量 $\boldsymbol{\theta}_\mathbf{A} \in \mathbb{R}^{N_t \times 1}$ 和原始数字波束成形器 $\hat{\mathbf{D}} \in \mathbb{C}^{K \times K}$ 被设为解码器的输出。

$$\left(\boldsymbol{\theta}_\mathbf{A}, \hat{\mathbf{D}}\right) = \mathcal{D}(\mathbf{Q}, \Theta_\mathcal{D}), \tag{9}$$

其中， $\mathcal{D}(\cdot)$ 表示参数为 $\Theta_\mathcal{D}$的 EFB 解码器网络。随后进行功率归一化，以满足功率约束 $\|\mathbf{ACD}\|_F^2 \le P$，如下所示。

$$\mathbf{D} = \frac{\sqrt{P}\hat{\mathbf{D}}}{\left\|\mathbf{AC}\hat{\mathbf{D}}\right\|_F} = \frac{\sqrt{P}\hat{\mathbf{D}}}{\left\|\operatorname{diag}\left(e^{j\boldsymbol{\theta}_\mathbf{A}}\right)\mathbf{C}\hat{\mathbf{D}}\right\|_F} \tag{10}$$

基于公式（1）和上述设计的波束成形器，系统的波束成形和速率可通过将所有用户的可达速率相加得到，具体如下。

$$R = \sum_{k=1}^{K} \log_2\left(1 + \frac{\left|\mathbf{h}_k^H \mathbf{ACd}_k\right|^2}{\sum_{i=1, i\neq k}^{K}\left|\mathbf{h}_k^H \mathbf{ACd}_i\right|^2 + \sigma^2}\right) \tag{11}$$

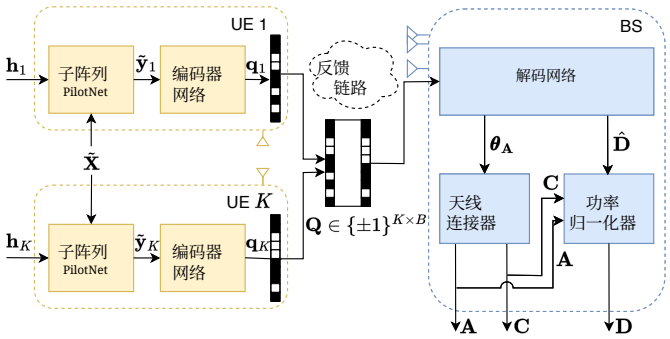
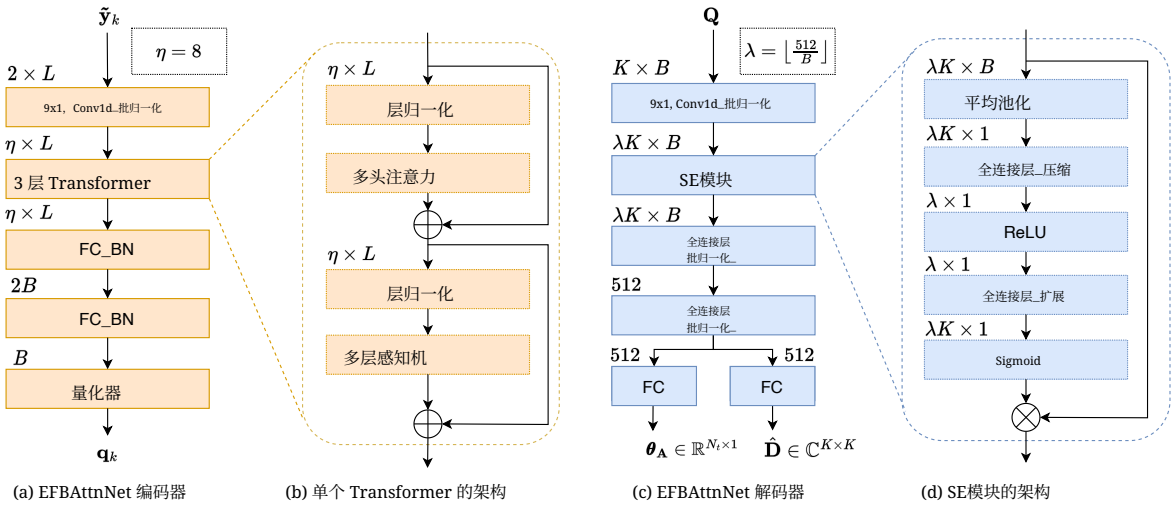


图 2. 通过网络联合优化对信道估计、反馈和波束成形进行协同设计的基于深度学习的EFB流程。



(a) EFBAttnNet 编码器　(b) 单个 Transformer 的架构　(c) EFBAttnNet 解码器　(d) SE模块的架构

图 3. 所提出的 EFBAttnNet 的设计。每个模块上方标注了输入维度，批量大小除外。

IV. 所提出的 EFBATTNNET

本节将详细介绍所提出的 EFBAttnNet。EFBAttnNet 编码器的架构如图 3（a）所示。 $\tilde{\mathbf{y}}_k$ 的实部和虚部拼接为两个通道，随后卷积层将通道数从 2 扩展到 8，以实现特征增强。

三个串联的 Transformer [9] 层构成编码器注意力模块，每个 Transformer 的架构如图 3 (b) 所示。自注意力结构擅长从具有复杂内部关系的输入中学习，例如根据导频信号估计信道 $\tilde{\mathbf{y}}_k$。

注意力模块之后，采用由批归一化（BN）构成的两层倒金字塔全连接层网络来调整输出维度。由于上行链路反馈需要比特流，因此使用最终量化器进行特征二值化。我们采用式 (12) 中的简单符号量化器，并在训练过程中使用 Sigmoid 调整的直通估计器，以便按照 [6] 对不可微符号函数进行梯度反向传播。

$$\mathcal{Q}(x) = \text{Sign}(\text{Sigmoid}(x) - 0.5) \in \{+1, -1\}, \qquad (12)$$

其中， $\mathcal{Q}(\cdot)$ 表示量化器。

值得注意的是，如图 2 所示，每个用户都有一个独立的 EFBAttnNet 编码器网络，而基站处仅有一个解码网络。正如图 3 (c) 所示，EFBAttnNet 解码器的输入也首先通过一个卷积核大小为 9×1 的一维卷积层进行扩展。扩展因子 λ 设置为 $\lfloor \frac{512}{B} \rfloor$，以在不同反馈比特数 B 下稳定解码器的复杂度。

在卷积扩展器之后添加一个压缩与激励（SE） [10] 模块，以实现通道注意力。如图 3 (d) 所示，网络学习一个注意力乘数，以重新校准通道权重，从而有利于后续两个全连接层的特征提取。最后，在末尾添加两个额外的全连接层，以生成维度正确的目标波束成形器。值得注意的是，每个批归一化层之后都添加了 ReLU 激活函数；为简洁起见，图 3 中省略了该部分。

EFBAttnNet 的设计主要基于注意力机制。借助 Transformer 和 SE模块，该网络能够在保持信道状态信息特征提取能力的同时，大幅降低复杂度。此外，专门设计了非平衡结构，使资源受限用户设备上的编码器复杂度远低于解码器。总体而言，所提出的 EFBAttnNet 旨在实现更实用的基于深度学习的反馈和波束成形。

V. 数值结果

A. 实验设置与基准方法

对于系统设置，SV信道模型中的路径数 L_{path} 设置为 2，而发射功率 P 设置为 1。所有实验均采用 10 dB 的信噪比（SNR）。对于深度学习相关设置，我们以无监督方式训练所提出的 EFBAttnNet，将损失函数设置为和速率的相反数 $-R$。按照 [3]，采用 Adam优化器和初始学习率为 10^{-3} 的余弦退火学习率（ LR ）调度器，并将批量大小设置为 1000。基于公式（4）中的信道模型，网络训练 300 个训练轮次，每个训练轮次包含 200 个批次。为进行公平比较，所有和速率性能均在 10000 个独立生成的信道上进行测试。

为展示所提出的 EFBAttnNet 的有效性，采用以下方法作为基准方法：

*1) HP-sub/*完整信道状态信息：假设信道估计和反馈均为理想情况，且基站拥有完整的信道状态信息。采用一种基于迫零的传统波束成形方法，称为 HP-sub [11]。

*2) HP-sub/OMP-CE/*无限反馈：用户设备使用正交匹配追踪（OMP）算法进行信道估计，以获取下行链路信道状态信息。随后，基站在理想反馈条件下采用 HP-sub 波束成形。

*3) HP-sub/OMP-CE/*有限反馈：用户设备使用 OMP 估计信道状态信息后，每个用户设备将信道参数 $\{\alpha_{ik}, \theta_{ik}\}_{i=1}^{L_{\text{path}}}$ 量化为 B 比特，以进行有限反馈。

4) EFBRefineNet：为进行公平的性能比较，我们将 [8] 中先前最先进的网络结构重新实现到我们的子阵列 EFB 流程中。

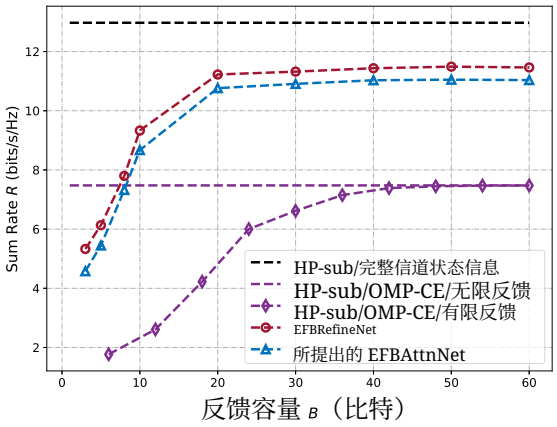


图4。在 $N_t = 64$、 $L = 8$条件下，双用户子阵列混合波束成形FDD MIMO系统中不同方法的和速率性能。

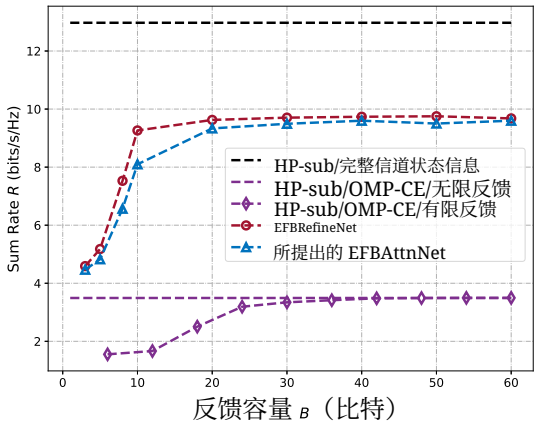


图5. 在具有 $N_t = 64$、 $L = 4$的双用户子阵列混合波束成形FDD MIMO系统中，不同方法的和速率性能。

表I 不同 B 下的复杂度比较，采用 $K = 2$， $L = 8$。

B	方法	参数量		浮点运算次数	
		总计[a]	UEs	总计	UEs
10	EFBRefineNet [8]	4376K	1103K	4453K	1117K
	EFBAttnNet	**849K**	**12K**	**1021K**	**100K**
40	EFBRefineNet [8]	4530K	1134K	4607K	1148K
	EFBAttnNet	**821K**	**26K**	**994K**	**114K**

[a] 解码器和两个编码器的总复杂度。

B. 波束成形性能与复杂度分析

本节将所提出的 EFBAttnNet 的和速率性能与上述基准方法进行比较。考虑一个具有 $N_t = 64$ 的双用户子阵列 EFB 系统，其中，基站的每条射频链都连接到 32 根天线，如式（2）和式（3）所述。

如图4所示，当 $L = 8$ 时，所提出的 EFBAttnNet 仅需 10 个反馈比特，性能便优于采用理想反馈的传统基准方法。此外，如图5所示，当导频数量进一步从 8 个减少到 4 个时，EFBAttnNet 的性能损失远小于传统基准方法。这证明，所提出的方法能够利用极其有限的导频（ $L \ll N_t$）提取信道特征并学习有效的子阵列混合波束成形器，而这对于传统的基于分块的 EFB 流程来说非常困难。

与此前最先进的 EFBRefineNet 相比，所提出的 EFBAttnNet 以少量性能损失换取了复杂度的大幅降低。如表 I 所示，参数规模和浮点运算次数（FLOPs）均大幅减少，尤其是资源受限用户设备中的编码器。此外，单个编码器和解码器的复杂度比从 16% 降至约 6%。这种非平衡结构也有利于实际部署。

VI. 结论

本文提出了一种用于FDD大规模多用户MIMO系统的新型信道估计、反馈和波束成形网络，名为EFBAttnNet。为提高部署的实用性，采用了子阵列混合波束成形架构，以提升能量效率。此外，利用注意力机制以较低成本提取更优的信道特征。实验结果证明，所提出的EFBAttnNet优于传统基准方法。与最先进的深度学习基准方法进一步比较发现，EFBAttnNet大幅降低了复杂度，尤其适用于资源受限用户设备，且仅造成和速率小幅损失。

参考文献

[1] 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*通信杂志，第55卷，第9期，第134–141页，2017年。[2] C.-K. Wen、W.-T. Shih和S. Jin，“大规模MIMO的深度学习CSI反馈”，*IEEE*无线通信快报，第7卷，第5期，第748–751页，2018年。[3] Z. Lu、J. Wang和J. Song，“大规模MIMO系统中基于深度学习的多分辨率CSI反馈”，载于*ICC 2020 - 2020*年*IEEE*国际通信会议（*ICC*），2020年，第1–6页。[4] X. Ma、Z. Gao、F. Gao和M. Di Renzo，“基于模型驱动深度学习的毫米波大规模混合MIMO系统信道估计与反馈”，*IEEE*通信领域精选期刊，第39卷，第8期，第2388–2406页，2021年。[5] J. Guo、C.-K. Wen和S. Jin，“单小区和多小区大规模MIMO系统中用于波束成形的基于深度学习的CSI反馈”，*IEEE*通信领域精选期刊，第39卷，第7期，第1872–1884页，2021年。[6] F. Sohrabi、K. M. Attiah和W. Yu，“频分双工大规模MIMO中的分布式信道反馈与多用户预编码深度学习”，*IEEE*无线通信汇刊，第20卷，第7期，第4044–4057页，2021年。[7] Z. Gao、M. Wu、C. Hu、F. Gao、G. Wen、D. Zheng和J. Zhang，“具有隐式CSI的空中大规模MIMO-OFDM系统的数据驱动深度学习混合波束成形”，*IEEE*通信领域精选期刊，第40卷，第10期，第2894–2913页，2022年。[8] M. Wu、Z. Gao、Z. Gao、D. Wu、Y. Yang和Y. Huang，“具有有限导频和反馈开销的频分双工大规模MIMO-OFDM系统深度学习混合预编码”，载于*2022*年*IEEE*国际通信会议研讨会（*ICC Workshops*），2022年，第318–323页。[9] A. Vaswani、N. Shazeer、N. Parmar、J. Uszkoreit、L. Jones、A. N. Gomez、L. Kaiser和I. Polosukhin，“注意力就是你所需要的一切”，载于第*31*届神经信息处理系统国际会议论文集，2017年，第6000–6010页。[10] J. Hu、L. Shen、S. Albanie、G. Sun和E. Wu，“挤压-激励网络”，*IEEE*模式分析与机器智能汇刊，第42卷，第8期，第2011–2023页，2020年。[11] Y. Guo、L. Li、X. Wen、W. Chen和Z. Han，“下行链路毫米波多用户大规模MIMO系统的基于子阵列的混合预编码设计”，载于*2017*年第九届无线通信与信号处理国际会议（*WCSP*），2017年，第1–4页。