

分类号： TN92

单位代码： 10335

密 级： 无

学 号： 21931108

浙江大学

硕士学位论文



中文论文题目： 5G和B5G大规模天线系统中
若干关键技术研究

英文论文题目： Research on Key Technologies for
5G/B5G Massive MIMO System

申请人姓名： 黄冠

指导教师： 刘安

专业名称： 信息与通信工程

研究方向： 无线通信

所在学院： 信息与电子工程学院

提交日期 2022年1月6日

5G和B5G大规模天线系统中
若干关键技术研究



论文作者签名: 黄冠

指导教师签名: 孙浩

论文评阅人1: 匿名

评阅人2: 匿名

评阅人3: 匿名

答辩委员会主席: 赵民建 教授 浙江大学信电学院

委员1: 郑史烈 教授 浙江大学信电学院

委员2: 余官定 教授 浙江大学信电学院

委员3: 林宏焘 研究员 浙江大学信电学院

委员4: 何先华 高工 诺基亚通讯有限公司

答辩日期: 2022年3月10日

浙江大学研究生学位论文独创性声明

本人声明所呈交的学位论文是本人在导师指导下进行的研究工作及取得的研究成果。除了文中特别加以标注和致谢的地方外，论文中不包含其他人已经发表或撰写过的研究成果，也不包含为获得 浙江大学 或其他教育机构的学位或证书而使用过的材料。与我一同工作的同志对本研究所做的任何贡献均已在论文中作了明确的说明并表示谢意。

学位论文作者签名：

签字日期： 2022 年 3 月 25 日

学位论文版权使用授权书

本学位论文作者完全了解 浙江大学 有权保留并向国家有关部门或机构送交本论文的复印件和磁盘，允许论文被查阅和借阅。本人授权浙江大学可以将学位论文的全部或部分内 容编入有关数据库进行检索和传播，可以采用影印、缩印或扫描等复制手段保存、汇编学位论文。

（保密的学位论文在解密后适用本授权书）

学位论文作者签名：

导师签名：

签字日期： 2022 年 3 月 25 日

签字日期： 2022 年 3 月 25 日

致 谢

“未觉池塘春草梦，阶前梧叶已秋声。”

三年的研究生生活转瞬即逝，而在求是园的日子竟也在不知不觉间步入了尾声。一晃七年，种种过往仿佛还如昨日般历历在目。自步入求是园融入信电学院，这一路走来，我曾畏惧、懈怠过，也曾风发、拼搏过。这期间，我遇到了太多可敬可亲的人们：恩师、益友、同窗，挚亲等等，是你们的包容、关怀和鼓励才支持我走到如今。想止于此，我提笔思忖，感怀难已……

首先，我要感谢我的导师刘安研究员。自从第一次在办公室与您见面，至今已三年有余。与您交流、讨论的次数越多，越能感受到老师的人格魅力。从本科时的毕业设计到如今研究生的毕业论文，期间一系列项目和课题工作都是在刘老师细致严谨的指导下完成的，包括文献调研、方案设计、展示讨论、实验分析和论文写作等等。刘老师在优化和通信领域渊博的学识，面对工作兢兢业业的态度和对待学生尽心尽责、循循善诱的培养方法……都给予我深深的感染和启发。我会始终记得老师“君子不器”的教诲，并以您为榜样在今后的学习工作中继续努力。谨以此表达对您衷心的感谢和谢意。

其次我要感谢赵民建教授。赵老师为人谦逊温和，对待教学、科研及实验室各项事务都细致入微、严谨认真。您在通信领域的高深造诣、一丝不苟的治学态度和宽宏仁厚的博大胸怀等等，称得上高山仰止。虽然少有机会得您在科研上的亲自指导，但您在课程及几次组会上传授的知识，都让我受益匪浅。

还要感谢李立言、雷鸣老师，在我遇到项目难题时曾给予过无私、恳切的指导和帮助。两位老师对硬件系统的深刻理解以及凡事亲历亲为的态度都令人折服。感谢蔡云龙教授、赵明敏和李旻老师，您们对学术研究饱满的热情也令我感染。感谢实验室的邵红霞、王婵和杨丽萍老师，谢谢你们为保持实验室正常运转和为我们营造轻松和谐的学习工作环境所付出的辛劳。

同样感谢在信电学院负责行政工作的蔡超、欧阳润清、陈超和周晓馨老师等。感谢您们在学生和党建工作中曾给予我的帮助和指导，让我感受到信电学院这个大家庭的温暖，祝你们事业顺利、身体健康。

感谢实验室已经毕业的陈晞涵博士师兄，您热情、严谨的学习工作态度让我非常敬佩，也是我今后努力的榜样。感谢即将毕业的刘冠颖师姐，经历过若干次项目合作后，您幽默风趣的谈吐、亲睦善良的性格值得我学习。感谢正在海外交流，追求博士学位的胡棋昱师兄，一同参加比赛的过往还历历在目，您笃行不倦的求学态度、温和谦逊的处事方法都让我印象深刻，祝您保重身体、学成归来。感谢北京大学的田畅师兄，您在学术课题上的细致指导让我收获颇多。感谢同届的黄哲、陈梓健、黄钰鹏、李绍标、窦为龙、章晓豪、程谦儒、叶远帆、吴怡芃和刘梦雪等同学曾经的帮助，是你们让我的研究生生活不再单调，祝你们都能顺利毕业并在今后的工作生活中一帆风顺。也感谢杨蕊、万雨博、胡智祥、徐文康、魏涵宇、程鹏远、方娇岚、盛思锴、张裔婷、张璐媛、王珂璇、武梦雨和林柏硕等师弟师妹的支持和鼓励，祝你们在求学途中竿头日上、不留遗憾。还有所遇的其他伙伴同学，一一罗列不得，感激不尽。

也要感谢我的研究生室友刘彦桢、陈嘉枫和吴承翰，及本科室友王润发、吴中伟、邓尧慷和刘一璟等同学为我原本单调枯燥的求学生活带来了诸多快乐。在漫长的相处中，我们经历过欢喜也分担过忧愁，这其中的酸甜苦辣将在我心中酿成一壶醇厚的酒，难以忘怀。

我还要把崇高的谢意和爱意献给我的父母：黄超先生和李玉华女士。感谢您们养育我长大并支持我完成学业；感谢您们努力为我营造的安心的成长环境；感谢您们在顺境时与我分享喜悦、在逆境时予我鼓励和安慰；感谢母亲教会我勤劳、踏实，父亲教会我乐观、勇敢……“谁言寸草心，报得三春晖”，养育之恩无以为报，惟愿流光轻弹以求慈乌反哺。当然也要感谢爷爷、大伯、姑姑和叔叔等亲人对我的关心和鼓励，让我体会到亲情的温暖，祝您们生活幸福、身体健康。

最后再次向以上的老师、同学、家人及其余未能一一提及的老师亲友献上衷心的感谢和祝福！“道由白云尽，春与青溪长。时有落花至，远随流水香……”，希望在我人生的下一个阶段也能继续与你们及更多人相遇。

黄冠

2022年1月于求是园

摘 要

移动通信技术的发展和普及为人类社会带来了巨大的影响，已经在潜移默化中彻底改变了人们的生产、生活方式和全世界的商业模式。受益于十年一代的通信技术的演进，当今的移动通信用户已达到数十亿，移动通信的峰值速率相较最初更是实现了近千倍的增长。在这其中，一种名为大规模阵列天线的技术因其在提高频谱和能效方面的卓越能力，已成为第五代移动网络（The fifth generation mobile networks, 5G）及超5G（Beyond 5G, B5G）通信系统的核心技术之一。然而，为了进一步挖掘大规模阵列天线的潜在优势并满足5G及B5G系统更多业务场景的需求，目前仍有许多关键问题需要解决，比如频分双工大规模阵列天线系统中的高精度低开销信道反馈问题、满足实际链路约束的更优多用户调度算法和面向5G/B5G不同业务场景的有效波束赋形和功率分配策略等。受此启发，本论文研究了5G/B5G大规模天线系统中的若干关键技术，包括信道反馈、用户调度、下行预编码和功率控制，其详细内容和主要贡献如下。

首先，我们研究了频分双工大规模阵列天线系统中的宽带信道状态信息（Channel state information, CSI）反馈问题，设计了一种称为TSAC的高精度低开销信道反馈方案。具体来说，我们首先从宽带信道中提取出需要反馈的二进制稀疏支撑矢量和信道系数信息。对二进制稀疏支撑矢量信息，本文通过利用大规模阵列天线信道在时延角度域上的动态突发稀疏特性和强时间相关性，提出了一种二级差分方案实现其无损反馈，且反馈开销可以根据信道变化速率自适应地调整。对信道系数信息，本文则设计了一种基于自适应字典的多原子矢量量化方案，实现了有效压缩和量化。相较现有通信协议标准中的NR Type-2反馈基线，我们的方案可在显著降低反馈开销的同时取得更优的传输误块率（Block error rate, BLER）表现。

其次，我们研究了多用户下行通信系统中以服务质量加权和速率最大化为目标的联合空频域多用户调度问题。为了使该问题的建模符合现有通信协议规范，我们考虑了5G标准中规定的连续频域资源分配约束。再考虑到大规模天线阵列和问题规模过大所导致的维度困难，这类问题是难以直接求解的。因此，本文通过构造一种增量容量函数将原问题转化为一系列子问题，并提出了一种JSFD-SMP算法来逐级求解。同时，为了降低算法复杂度，

本文还精心设计了一种KPS简化实施方法，可在不损失性能的前提下显著降低计算开销。最后，仿真结果证明了JSFD-SMP算法相较现有基线方案可以在保持优秀的用户体验速率公平性的同时显著提升小区吞吐量。

然后，我们研究了大规模阵列天线系统中低调制编码方案（Modulation and coding scheme, MCS）场景下以BLER为直接性能指标的下行预编码设计问题。由于BLER是一个与MCS、码率、传输块大小和信道编码方式等诸多实际链路参数有关的复杂函数，因此原编码系统优化问题是极难直接解决的。为了克服这个困难，本文首先引入链路分析和评估中常用的指数型等效信干噪比映射方法。之后，在考虑非理想CSI的场景下，我们基于EESM度量准则将原问题转化，重新构建了一个针对无编码系统中的优化问题并提出了一种E-S-WMMSE算法来有效求解。最后，基于大量的仿真实验和计算结果，我们证明了EESM度量准则在评估BLER性能上的准确性与合理性，同时也展示了所提算法的优越性。

最后，我们考虑一个包含动态队列的下行多用户多输入多输出系统，设计了一种基于正则化迫零预编码的功率控制策略，用以实现在满足各用户长期平均时延约束下的长期平均发送功率最小化。其涉及的问题是一个复杂的约束马尔可夫决策过程。本文设计了一种新颖的受约束深度强化学习（Constrained deep reinforcement learning, CDRL）算法，称为SCAPO，来有效求解。SCAPO算法通过将原问题中复杂的目标和约束函数用相应的凸代理函数替代，从而构造一系列针对转化后凸目标或可行集的凸优化问题。在每次更新中，SCAPO只需基于从在线样本中估计得到的目标和约束函数的一阶信息即可构造出凸替代函数，且其求解也是高效的。此外，SCAPO算法能够通过重用以往的旧样本数据来降低对在线样本的需求。数值仿真表明SCAPO相较现有的先进CDRL算法能实现更优的性能表现。

关键词：大规模阵列天线，信道状态信息反馈，频域资源分配，用户调度，指数型等效信干噪比映射，下行预编码，受约束深度强化学习，功率控制，超5G通信技术

Abstract

The development and popularization of mobile communication technology has brought a huge impact on human society, which has completely changed people's production, life style and business model all over the world. Benefit from decennial development of the communication technologies, the number of mobile communication users has reached billions, and the peak rate has increased by nearly a thousand times compared to the original. Among these, massive multiple-input multiple-output (MIMO) has become one of the core technologies in the fifth generation (5G) and beyond 5G (B5G) mobile communication systems due to its superior ability for improving spectrum and energy efficiency. However, to further unleash potential superiority of the massive MIMO and to meet more traffic requirements of the 5G/B5G, there are still many critical issues that need to be tackled, such as high-resolution and low-overhead channel state information (CSI) feedback scheme for FDD wideband massive MIMO system, more superior multi-user scheduling algorithm under practical link constraints, effective beamforming method and power control policy for distinctive 5G/B5G use cases and so on. Motivated by these, this thesis studies the key technologies for 5G/B5G massive MIMO system, involving channel feedback, user scheduling, downlink precoding and power control. The details and main contributions are listed as follows.

In the first part, we consider a high-resolution required wideband CSI feedback problem for FDD massive MIMO system, where a two-stage adaptive and compressed (TSAC) CSI feedback scheme is proposed to reduce the CSI feedback overhead and improve the feedback accuracy. Specifically, two binary sparse support vectors and the associated channel coefficients are first extracted from the massive MIMO channel. Then, by exploiting the dynamic burst-sparsity of the massive MIMO channel in the joint time-angular domain and strong time correlations, a two-stage differential lossless feedback scheme is proposed to compress and feed back the support vectors, where the feedback overhead is adaptively adjusted according to the channel changing rate. In addition, an adaptive-dictionary-based multi-atom vector quantization scheme is proposed for the channel coefficients. Compared with the 5G New Radio (NR) Type-2 feedback scheme in 3GPP

R15, the proposed scheme can significantly reduce the feedback overhead and achieve better block error rate (BLER) performance.

In the second part, we investigate a multi-user downlink transmission system, where the joint space-frequency domain user scheduler is optimized for maximizing quality-of-service (QoS) weighted sum-throughput. The formulated problem complies with standards specifications of the 5G and B5G, however, is hard to directly solve due to the intractable contiguous frequency-domain resource allocation (FDRA) constraint and curse of dimensionality. By tailoring an incremental weighted sum-throughput function, we transform the primary problem into several subproblems and propose a joint space-frequency domain successive message passing (JSFD-SMP) algorithm to iteratively solve them. Moreover, a method with no performance loss named key-point sampling (KPS) is carefully designed to reduce the computation overhead. Simulation results verify that proposed JSFD-SMP scheduling scheme can achieve significant throughput gain while maintaining good user experienced rate fairness over the state-of-the-art baseline schemes.

In the third part, we consider a beamforming design problem of the downlink massive MIMO system, in which case the optimization aim is to obtain good BLER performance under low modulation and coding scheme (MCS) value. Since BLER is a complicated function related to practical link parameters, i.e., MCS, code rate, transport block size (TBS) and channel coding scheme, the primary coded system optimization problem is non-trivial. To tackle this difficulty, we first introduce a performance metric called exponential effective signal-to-interference-plus-noise ratio mapping (EESM), which is an indispensable tool for link analysis and evaluation. Then, after taking imperfect CSI into consideration, we reformulate an uncoded system optimization problem based on EESM and propose an EESM-based successive weighted sum mean-square-error minimization (E-S-WMMSE) algorithm to efficiently solve it. Extensive simulation and computation results first verify reasonability and availability of EESM metric for evaluating BLER performance, and then show the superiority of proposed algorithm.

In the last part, we investigate the downlink transmission for multi-user MIMO (MU-MIMO) system with dynamic queues, in which the regularized zero-forcing precoder is adopted and the power allocation and regularization factor are optimized. Our aim is to find a power allocation and regularization factor control policy that can minimize the long-term average power consumption subject to long-term delay constraint for each user. The induced optimization problem is formulated as a constrained Markov decision process (CMDP), which is efficiently solved by the proposed constrained deep reinforcement learning (CDRL) algorithm, called successive convex

approximation policy optimization (SCAPO). The SCAPO is based on solving a sequence of convex objective/feasibility optimization problems obtained by replacing the objective and constraint functions in the original problems with convex surrogate functions. At each iteration, the SCAPO merely needs to estimate the first-order information and solve a convex surrogate problem that can be efficiently parallel tackled. Moreover, the SCAPO enables to reuse old experiences from previous updates, thereby significantly reducing the implementation cost. Numerical results show that the novel SCAPO can achieve the state-of-the-art performance over advanced baselines.

Keywords: massive MIMO, CSI feedback, FDRA, user scheduling, EESM, downlink precoding, CDRL, power control, beyond 5G

图目录

2.1	FDD massive MIMO系统模型.....	24
2.2	基于双域压缩的上行有限反馈系统框图.....	27
2.3	二进制支撑矢量的突发稀疏性和强时间相关性.....	28
2.4	四用户单流场景下的BLER性能	35
2.5	两用户双流场景下的BLER性能	35
2.6	各方案在不同传输场景下的吞吐量性能.....	36
3.1	单小区多用户下行系统模型.....	39
3.2	空分复用的的多用户频分调度.....	40
3.3	第 t 轮调度时的因子图	44
3.4	对数比矢量的排布图.....	48
3.5	10用户场景时的用户平均吞吐量性能.....	50
3.6	20用户场景时的用户平均吞吐量性能.....	51
3.7	30用户场景时的用户平均吞吐量性能.....	52
3.8	不同场景下各方案的JFI性能比较.....	52
4.1	ESM方法在通信链路中的应用	59
4.2	各次混合功率分配方案的BLER性能	67
4.3	双流MCS1传输时各方案的BLER性能	68
4.4	信道CEM5dB且双流MCS1传输时各方案的BLER性能.....	69
4.5	四流MCS1传输时各方案的BLER性能	69
4.6	双流MCS3传输时各方案的BLER性能	70
5.1	包含动态队列的下行MU-MIMO系统.....	72
5.2	功率控制策略.....	75
5.3	Q-value的估计方法	78
5.4	4用户8发送天线场景下的功率消耗学习曲线.....	81
5.5	4用户8发送天线场景下的队列时延学习曲线.....	82
5.6	8用户16发送天线场景下的功率消耗学习曲线.....	82

5.7 8用户16发送天线场景下的队列时延学习曲线..... 83

表目录

1.1	现有信道反馈方案总结.....	6
1.2	现有用户调度策略总结.....	13
1.3	现有下行预编码技术总结.....	19
2.1	指示器的详细注释.....	32
2.2	系统和信道参数设置.....	33
2.3	不同方案反馈开销比较.....	34
3.1	信道和业务模型参数设置.....	49
4.1	系统和信道仿真参数设置.....	66
4.2	各次混合功率分配方案的EESM目标函数值.....	67
5.1	系统和信道参数设置.....	80

缩写表

1G	The First Generation Mobile Networks	第一代移动网络
2G	The Second Generation Mobile Networks	第二代移动网络
3G	The Third Generation Mobile Networks	第三代移动网络
3GPP	The Third Generation Partnership Project	第三代合作伙伴计划
4G	The Fourth Generation Mobile Networks	第四代移动网络
5G	The Fifth Generation Mobile Networks	第五代移动网络
AMC	Adaptive Modulation and Coding	自适应调制编码
AWGN	Additive White Gaussian Noise	加性白高斯噪声
AoD	Angle-of-Departure	离开角
AoD-SS	AoD-Adaptive Subspace	自适应离开角子空间
B5G	Beyond 5G	超5G
BCD	Block Coordinate Descent	块坐标下降
BER	Bit Error Rate	误比特率
BET	Blind Equal Throughput	盲等吞吐量
BLER	Block Error Rate	误块率
BLI	Burst-Length Indicator	突发长度指示
BP	Belief Propagation	置信传播
BPI	Burst-Position Indicator	突发位置指示
BS	Base Station	基站
BWP	Bandwidth Part	部分带宽
CDI	Channel Direction Information	信道方向信息
CDL-B	Clustered Delay Line Type-B	簇延时线类型B
CDMA	Code Division Multiple Access	码分多址
CDRL	Constrained Deep Reinforcement Learning	受约束深度强化学习
CEI	Changed-Set Indicator	变化集合指示

CEM	Channel Error MSE	信道误差MSE
CESM	Capacity ESM	容量型ESM
CI	Channel Inversion	信道逆转
CIR	Channel Impulse Response	信道冲激响应
CMDP	Constrained Markov Decision Process	受约束马尔可夫决策过程
CNN	Convolutional Neural Network	卷积神经网络
CPO	Constrained Policy Optimization	受约束决策优化
CQI	Channel Quality Indicator	信道质量指示
CS	Compressed Sensing	压缩感知
CSI	Channel State Information	信道状态信息
CSI-RS	CSI Reference Signal	CSI参考信号
DCI	Downlink Control Information	下行控制信息
DFT	Discrete Fourier Transform	离散傅里叶变换
DM-RS	Demodulation Reference Signal	解调参考信号
DNN	Deep Neural Network	深度神经网络
DP	Dynamic Programming	动态规划
DPC	Dirty Paper Coding	脏纸编码
DS	Delay Spread	时延扩展
E-S-WMMSE	EESM-based Successive WMMSE	基于EESM的逐级WMMSE
EESM	Exponential Effective SINR Mapping	指数型等效信干噪比映射
EMI	Equal Mutual Information	等互信息
EPA	Equal Power Allocation	等功率分配
ESM	Effective SINR Mapping	等效信干噪比映射
FDD	Frequency Division Duplex	频分双工
FDMA	Frequency Division Multiple Access	频分多址
FDPS	Frequency Domain Packet Scheduling	频域包调度
FDRA	Frequency Domain Resource Allocation	频域资源分配
FHYP	Fast-Timescale Hybrid Precoding	快时间尺度混合预编码
FIFO	First-In First-Out	先入先出
GMD	Geometric Mean Decomposition	几何均值分解
GP	Geometric Programming	几何规划

GPF	Generalized Proportional Fair	广义比例公平
GPU	Graphics Processing Unit	图形处理单元
HARQ	Hybrid Automatic Repeat-Request	混合自动重传请求
HOL	Head-of-Line	队首
HPA	Hybrid Power Allocation	混合功率分配
HRFM	High Resolution Feedback Mode	高精度反馈模式
IoT	Internet of Things	物联网
JADE	Joint Allocation with Dual Ends	双端联合分配
JFI	Jain's Fairness Index	简氏公平性因子
JSFD	Joint Space-Frequency Domain	联合空频域
JSFD-SMP	JSFD Successive Message Passing	联合空频域逐级消息传播
KKT	Karush-Kuhn-Tucker	卡罗需-库恩-塔克
KPS	Key-Point Sampling	关键点采样
LAA	Lens Antenna Array	透镜天线阵
LBG	Linde-Buzo-Gray	林德-布佐-格雷算法
LEAP	Localized Expansion of Adjacent Positions	邻位局部扩展
LESM	Logarithmic ESM	对数型ESM
LRFM	Low Resolution Feedback Mode	低精度反馈模式
LTE	Long Term Evolution	长期演进
M-LWDF	Modified Largest Weighted Delay First	修正最大权重时延优先
MAC	Medium Access Control	媒体接入控制
MAP	Maximum-a-Posteriori	最大后验
MCS	Modulation and Coding Scheme	调制编码方案
MEC	Mobile Edge Computing	移动边缘计算
MF	Match Filter	匹配滤波
MIESM	Mutual Information ESM	互信息型ESM
MIMO	Multiple-Input Multiple-Output	多输入多输出
MISO	Multiple-Input Single-Output	多输入单输出
MMSE	Minimum Mean-Square Error	最小均方误差
MSE	Mean-Square Error	均方误差
MT	Maximum Throughput	最大吞吐量

MU-MIMO	Multi-User MIMO	多用户MIMO
MWF	Mercury Water-Filling	水银注水
NLOS	Non Line-of-Sight	非直视径
NP-hard	Non-Deterministic Polynomial Hard	非确定多项式困难
NR	New Radio	新空口
OFDM	Orthogonal Frequency Division Multiplexing	正交频分多址
OFDMA	Orthogonal Frequency Division Multiple Access	正交频分多址接入
OMP	Orthogonal Matching Pursuit	正交匹配追踪
ORA	Optimum Resource Allocation	最优资源分配
PAM	Pulse Amplitude Modulation	脉冲幅度调制
PDSCH	Physical Downlink Control Channel	物理下行控制信道
PF	Proportional Fair	比例公平
PHY	Physical Layer	物理层
PMI	Precoding Matrix Indicator	预编码矩阵指示
PPO	Proximal Policy Optimization	近端策略优化
PSK	Phase Shift Keying	相移键控
QAM	Quadrature Amplitude Modulation	正交幅度调制
QSI	Queue State Information	队列状态信息
QoS	Quality-of-Service	服务质量
RB	Resource Block	资源块
RBG	RB Group	资源块组
RR	Round Robin	轮询
RRM	Radio Resource Management	无线资源管理
RZF	Regularized Zero-Forcing	正则化迫零
SCAPO	Successive Convex Approximation Policy Optimization	逐级凸近似策略优化
SEI	Static-Set Indicator	不动集合指示
SINR	Signal-to-Interference-plus-Noise Ratio	信干噪比
SISO	Single-Input Single-Output	单输入单输出
SNR	Signal-to-Noise Ratio	信噪比
SRS	Sounding Reference Signal	探测参考信号
SU-MIMO	Single-User MIMO	单用户MIMO

SVD	Singular Value Decomposition	奇异值分解
TBS	Transport Block Size	传输块大小
TDD	Time Division Duplex	时分双工
TDMA	Time Division Multiple Access	时分多址
TDPS	Time Domain Packet Scheduler	时域包调度器
THP	Tomlinson-Hiroshima Precoding	汤姆林森-哈拉希玛预编码
THYP	Two-Timescale Hybrid Precoding	两级时间尺度混合预编码
TSAC	Two-Stage Adaptive and Compressed	二级自适应压缩
TTI	Transmission Time Interval	传输时延间隔
ULA	Uniform Linear Array	均匀线性天线阵列
UPA	Uniform Planar Array	均匀平面天线阵列
URLLC	Ultra-Reliable and Low-Latency Communications	超可靠低时延通信
VP	Vector Perturbation	矢量扰动
VQ	Vector Quantization	矢量量化
WF	Water-Filling	注水
WFQ	Weighted Fair Queuing	加权队列公平
WMMSE	Weighted Sum MSE Minimization	加权均方误差和最小化
ZF	Zero-Forcing	迫零
eMBB	Enhanced Mobile Broadband	增强移动宽带
mMTC	Massive Machine Type Communications	大规模机器类型通信

目 次

致谢	I
摘要	III
Abstract	V
图目录	VIII
表目录	X
缩写表	XI
目次	
1 绪论	1
1.1 研究背景与意义	1
1.2 关键技术研究进展及现状	4
1.3 论文主要内容和结构安排	19
2 大规模天线系统中基于双域压缩的高精度信道反馈技术	22
2.1 研究动机和主要成果	22
2.2 系统模型与方案概述	23
2.3 TSAC反馈方案设计	26
2.4 仿真结果与分析	33
2.5 本章小结	36
3 中重载场景下支持空域并发的多用户频分调度方案设计	37
3.1 研究动机和主要成果	37
3.2 系统模型与问题建模	38
3.3 JSFD-SMP用户调度算法	42
3.4 仿真结果与分析	49
3.5 本章小节	53
4 基于EESM准则的联合空频预编码及低MCS下增益空间探索	54
4.1 研究动机和主要成果	54
4.2 系统模型与问题建模	56
4.3 E-S-WMMSE算法	59
4.4 仿真结果与分析	65

4.5 本章小结	70
5 受约束深度强化学习框架下的时延感知功率控制策略研究	71
5.1 研究动机和主要成果	71
5.2 系统模型和问题建模	72
5.3 SCAPO算法.....	75
5.4 仿真结果与分析	80
5.5 本章小结	82
6 总结与展望	84
6.1 全文总结	84
6.2 未来工作展望	85
参考文献	87
附录	97
A 对引理1的证明	97
B 对定理2的证明	103
C 对引理3的证明	104
攻读硕士学位期间的研究成果	107

1 绪论

1.1 研究背景与意义

移动通信技术的发展和普及为人类社会带来了巨大的影响，已经在潜移默化中彻底改变了人们的生产、生活方式和全世界的商业模式。受益于第四代移动通信网络（The fourth generation mobile networks, 4G）中长期演进（Long term evolution, LTE）这一迄今为止最为成功的移动宽带技术，当今的移动通信用户已达到数十亿，移动通信的峰值速率相较最初更是实现了近千倍的增长。毫无疑问，移动宽带现在是、将来也必将是移动通信技术的重要组成部分。但随着科学技术和信息社会的进步，日益增长的无线数据量和蓬勃涌现的各类业务场景对移动通信网络的系统容量、频谱效率、用户覆盖密度、体验速率、传输时延和连接稳定性等诸多特性提出了更高的要求，这为第五代移动通信网络（The fifth generation mobile networks, 5G）乃至超5G（Beyond 5G, B5G）技术的发展提供了强大的驱动力。令人鼓舞的是，随着2018年6月的5G新空口（New radio, NR）无线接入技术为第三代合作伙伴计划（The third generation partnership project, 3GPP）组织所标准化并正式发布，遵循5G标准的商用网络已在全球范围内进行规模部署和商用。5G的产业链也日渐成熟，并逐步扩展于增强的移动宽带通信（Enhanced mobile broadband, eMBB）、超可靠低时延通信（Ultra-reliable and low-latency communications, URLLC）和大规模机器类型通信（Massive machine type communications, mMTC）这三类业务场景。然而，在移动通信技术蓬勃发展的现在，我们仍需要注意现有的无线频谱资源本身是稀缺的这一天然内禀属性。毫无疑问，这成为制约当今乃至未来无线通信技术有效发展的一项不容忽视的阻碍。因此对特定的通信场景来说，设计恰当的方案提高频谱利用效率、避免用户间的恶意干扰，同时尽可能减少无意义的资源消耗是至关重要的。虽然以往十年一代的通信技术，如第一代移动通信网络（The first generation mobile networks, 1G）的频分多址（Frequency division multiple access, FDMA）和蜂窝组网技术、第二代移动通信网络（The second generation mobile networks, 2G）的时分多址技术（Time division multiple access, TDMA）技术、第三代移动通信网络（The third generation mobile networks, 3G）的码分多址技术（Code division

multiple access, CDMA) 技术和4G的正交频分多址 (Orthogonal frequency division multiple access, OFDMA) 和多输入多输出 (Multiple-input multiple-output, MIMO) 技术等, 都有效解决了当时面临的瓶颈困难, 但并未根本地克服限制通信技术发展的关键难题。幸运的是, 在广大通信研究和技术人员的不懈努力下, 一种名为大规模多输入多输出 (Massive MIMO) 阵列天线的技术应运而生^[1]。由于其可以在不占用额外时、频传输资源的同时显著地提升系统容量和可靠性, 已经成为5G乃至未来通信系统的核心技术之一。基站 (Base station, BS) 通过部署大规模天线阵列, 充分利用波束赋形 (Beamforming)、混合预编码等技术, 可以显著提升频谱利用率且降低能量损耗。同时, 理论上可以证明当部署天线数量趋于无穷时, 各用户信道矢量会趋于正交并使得同一小区内的用户间干扰被消除。具体来说, massive MIMO技术的优势主要有以下几个方面:

- (1) **显著提升频谱和能量效率**: massive MIMO系统通过部署大量的天线, 能充分利用空间复用增益从而显著提升系统容量。在总的发射能量一定时, 通过使用波束赋形、预编码等技术, BS能够将发射能量集中到各目标移动终端上, 从而降低单根天线辐射功率和总的能量损耗, 同时也降低了对其他用户的干扰, 因此可在不增加额外带宽的情况下成倍地提高频谱利用率和能量效率。
- (2) **对单通道硬件非理想性不敏感**: massive MIMO系统降低了对单个天线的精度要求, 可以采用更价廉的低功耗硬件。当天线数目增加, massive MIMO系统中单根天线上的功率相较于传统 MIMO大大减少, 从而降低了系统对单个放大器或射频链路的线性度要求。因此可以使用廉价的功率放大器代替价格昂贵的高功率放大器, 这降低了设备的开销。
- (3) **系统可靠性提高**: 传统的MIMO系统对功率放大器以及射频链路的要求很高, 系统的某一模块出现故障都有可能使整个系统崩溃。而massive MIMO系统中天线单元众多, 部分模块出现问题往往不会影响整个系统的正常运行, 从而使系统的可靠和鲁棒性 (Robust) 得到提高。另一方面, 在使用massive MIMO技术的OFDMA系统中, 子载波之间的信道增益近似相同, 这简化了物理层 (Physical layer, PHY) 控制信号和媒体接入控制 (Medium access control, MAC) 层的设计难度。

因此, 得益于massive MIMO技术带来的空域设计自由度, 基站在获得各用户信道后可根据不同业务场景的需求更灵活地设计合适的MAC层用户调度及PHY下行预编码算法, 从而有效地避免用户间恶意竞争并显著提升传输性能。不过在上述过程中, 有以下几类关键问题仍需解决:

- (1) **高精度信道反馈**: 在设计用户调度和预编码算法前, 为了使发射波束精准传输到期望方向, 基站必须获得下行信道状态信息 (Channel state information, CSI)。

而CSI的获取在频分双工 (Frequency division duplex, FDD) 系统和时分双工 (Time division duplex, TDD) 系统中存在很大不同。TDD 系统中, 由于存在上下行信道的互易性 (Reciprocity), 基站可以根据用户发射的信号计算出上行CSI, 进而推导出完整或部分的下行CSI。FDD模式下CSI的获取相对复杂, 通常基于导频 (Pilot) 训练、信道估计和反馈的方式。首先由基站向各用户发送导频信号。当用户接收到导频信号后, 可以使用最小二乘或最小均方误差的方法估计出下行CSI。这些估计的CSI通过上行链路反馈到基站, 帮助实现用户调度、波束成形及功率控制等。但是随着天线数目的倍增和系统带宽的增加, 宽带信道的反馈开销也逐渐变大。特别是在有限反馈的massive MIMO系统中, 高精度低开销的信道反馈方法是亟需解决的一大问题。

- (2) **符合链路约束的用户调度**: 在实际无线链路系统中, 为了满足不同的业务需求, 基站需要为多个用户提供不同质量的连接服务。但受限于传输资源, 基站一般无法同时接入所有用户, 而是基于FDMA或TDMA等多址技术将不同用户的传输数据映射到不同时、频域资源上发送。这一过程一般由MAC层调度器负责完成并将结果通过下行控制信息 (Downlink control information, DCI) 和物理下行控制信道 (Physical downlink control channel, PDSCH) 发送至各用户。为了降低信令开销, 用户调度策略往往需要满足实际的链路约束, 如频域资源的连续性分配、各用户传输流和调制编码方案 (Modulation and coding scheme, MCS) 的一致性等等。因此, 我们虽然可以通过挖掘massive MIMO系统的空域自由度设计支持多用户空域并发的更优调度算法, 但倍增的天线规模和上述链路约束的存在使得这类问题难以求解。更重要的, 为了支持5G和B5G系统中URLLC类应用, 调度器的处理时间往往被限制在1ms (甚至更低) 内。这也对调度算法的复杂性也提出了更严苛的要求。
- (3) **不同业务场景下的预编码设计**: 基站在获得CSI并计算出用户调度策略后, 还需要设计合适的预编码算法来提升传输性能。在传统的大部分业务场景中, 系统容量往往是重要的性能指标。因此预编码器的优化目标大多是最大化小区用户的加权和速率。但随着5G和B5G系统中涌现出的更多业务场景类型, 这一优化目标也可能发生变化。比如在评估小区边缘用户传输可靠性的场景中, 设计以传输误块率 (Block error rate, BLER) 为指标的预编码算法是更符合实际的; 在考虑对传输时延敏感的业务中, 设计满足时延约束下发送功率小的控制策略则是更重要的。因此, 由丰富的业务类型催生出的不同优化目标对massive MIMO系统中预编码或功率控制算法的设计提出了更高的要求。

基于上述研究背景，本论文将针对现有5G及B5G大规模天线系统中PHY和MAC层涉及的一系列关键技术难题，即信道反馈、用户调度和下行预编码等，设计对应的关键算法和解决方案。具体研究内容将在本章第三节详细介绍，下一小节将调研以上三类关键技术的国内外研究现状。

1.2 关键技术研究进展及现状

本小节，我们将分三个子章节详细介绍现有通信系统中信道反馈、用户调度和下行预编码技术的研究进展及现状。

1.2.1 信道反馈研究现状

现有massive MIMO信道反馈方案可分为三大类：角度域（Angular-domain）稀疏信道反馈方案、直接反馈信道测量值方案和基于深度学习的压缩信道反馈方案。这三种方案都利用了massive MIMO信道的稀疏（Sparsity）特性来降低反馈信令开销。具体而言，由于massive MIMO系统中天线间距较小、有效传播路径有限等原因，信道之间往往有着较强的空间相关性。因此，许多研究者从空频域的角度来挖掘massive MIMO信道的稀疏性征，并利用该稀疏特征设计高效信道反馈方案。下面将详细阐述上述三种massive MIMO信道反馈方案的研究现状及目前存在问题。

A. 角度域稀疏信道反馈方案

在FDD模式下的无线通信系统中，由于信道不具有TDD模式的信道互易特性，只有通过下行导频训练和上行链路的信道反馈，基站才能获得信道状态信息。但是下行导频训练和上行信道反馈的开销与天线数目成正比增长，这对于massive MIMO系统的应用是一个巨大挑战。近年，一些研究和实验表明，由于基站周围的散射物有限，下行信道会在虚拟角度域上表现出稀疏特性^[2-4]。根据这一特性，学者提出了一种角度域稀疏信道反馈方案，可采用基于信道出射角（Angle-of-departure, AoD）的自适应码本设计降低信道反馈开销。现有角度域稀疏信道反馈方案又可分为快时间尺度和多时间尺度两种。

在快时间尺度反馈方案中，用户需要在每个信道相干时间内，同时反馈信道各径角度方向信息（Channel direction information, CDI）和信道质量指示（Channel quality indicator, CQI）以准确跟踪信道变化。比如文献^[3]提出了一种联合下行信道估计与上行信道反馈方案。用户在信道相关时间内接收到的导频信号后，首先采用压缩感知（Compressed sensing, CS）恢复算法进行信道估计。随后，再将估计信道的AoD和质量信息反馈给基站。文

献^[4]则设计了一种基于非均匀双向字典的信道反馈方案，提出在天线方向图中功率峰值附近进行更密集的量化，从而提高了角度分辨率并有效降低了量化误差。而在某些场景下，研究者利用基站周围散射物环境近似不变、角度方向信息比复增益信息变化更缓慢的特性，通过引入角度相关时间（Angle-coherence-time）的概念提出了多时间尺度反馈方案。相比于前者，用户只需在每个角度相关时间内反馈一次信道角度信息，有效降低了反馈开销。比如文献^[5]设计了一种AoD自适应的码本来减少量化比特。文献^[6]提出了均匀平面天线阵列情形下的自适应码本构造方法和另一种基于信道子空间的模拟反馈方法。文献^[7]则考虑了基于透镜天线阵（Lens antenna array, LAA）的毫米波massive MIMO系统，提出了一种降维子空间码本的设计方案来量化其等效信道。

综合来看，相对快时间尺度信道反馈方案，多时间尺度反馈虽然能进一步减少信道反馈开销，但它建立在信道AoD于角度相干时间内不发生变化的假设基础之上。在实际系统中，由于遮挡物的运动和环境的改变，信道AoD通常随时间是逐渐变化的，甚至有可能发生突变。当AoD变化积累到一定程度，多时间尺度信道反馈方案会有比较大的性能损失，因此在实际系统中的鲁棒性较差。另一方面，快时间尺度信道反馈方案虽然可以跟踪AoD的变化，但是反馈开销较大。所以，需要设计一种既可以跟踪AoD变化，同时又具有较小反馈开销的新型CSI压缩反馈方案。

B. 直接反馈信道测量值

最近，研究者们还提出了一种联合信道估计和反馈方案。用户直接反馈接收到的训练导频信号（信道测量值），并在基站端使用各种恢复算法来估计下行信道。这种方案的优势是避免了在用户端进行的信道估计和CSI投影计算，减轻了功率受限用户的计算负担。另一方面，基站也可以利用信道的时域相关性或者多用户信道的联合稀疏特征提高信道估计的性能，减少计算资源消耗。文献^[8,9]利用时变信道相邻时隙的信道冲激响应（Channel impulse response, CIR）的差分要比原CIR更稀疏的性质，提出了一种基于压缩感知的差分CSI估计方案。文献^[10,11]考虑到多用户系统接收端可能有相似的物理环境，如障碍物、反射物等，根据各用户信道的联合相关特性提出了一种分布式CSI估计方案。

直接反馈测量值的方案虽然在某些场景下能降低接收端的计算负担，但是并没考虑到反馈测量值经过上行信道的量化误差。同时与角度域稀疏信道方案相比，直接反馈信道测量值的开销通常更大。这是因为恢复信道所需的导频数，通常是信道稀疏度的2到3倍。

C. 基于深度学习的压缩反馈方案

由于人工智能、机器学习等方法在计算机视觉和自然语言处理等领域取得的成功，近年来，深度学习也被应用到无线通信领域以解决信道估计、压缩反馈等关键问题^[12]。这些深度学习网络模型大多是数据驱动的（Data-driven），即把系统架构视为一个需要训练数据确定参数的黑盒。这种数据驱动的深度学习方法除了需要充足的训练集，还需要大量的训练时间，而这些资源在实际无线通信系统中都是稀缺的。因此，除了纯粹的数据驱动方法，学者还提出了另一种基于实际物理模型和网络拓扑的模型驱动（Model-driven）方法^[13]。这种方法对训练数据量的要求相对较低，因而被更加广泛地应用于信道反馈等领域。文献^[14]提出了一种基于卷积神经网络（Convolutional neural network, CNN）的CsiNet来对信道矩阵进行自动编码和重构。这种网络结构实际是一组经过信道结构数据训练的编、解码器。接收端的编码器可以提取信道矩阵的特征并进行压缩表示。经过上行链路反馈后，基站端的解码器利用这些码字来恢复信道数据。文献^[15]则在CsiNet训练网络基础上进一步考虑信道时域相关性，提出了CsiNet-LSTM训练网络来解决时变信道的反馈问题。

综合来看，基于深度学习的信道反馈方法在某些场景下可以取得较好的反馈效果，但对于massive MIMO系统来说，可靠的信道训练数据获取在实际中是十分困难的。另一方面，在实际场景中，信道的统计特性往往随时间改变，因此训练网络的参数也要实时更新，否则会产生比较严重的估计误差。所以深度学习方法能否可靠地与实际通信系统中应用还有待进一步研究。

表 1.1 现有信道反馈方案总结

反馈方案	研究方向	优缺点
角度域稀疏信道反馈方案	1. 快时间尺度； 2. 多时间尺度。	1. 可以跟踪信道AoD变化，但反馈开销较大； 2. 反馈开销进一步减少，但无法处理信道突变情况，鲁棒性较差。
直接反馈信道测量值方案	1. 基于CIR的时间相关性； 2. 基于多用户间的联合稀疏性。	可以降低接收端的处理负担；基站可利用相关性降低计算开销，但直接反馈的开销更大，且上行有限反馈信道会影响恢复准确度。
基于深度学习的压缩反馈	1. 慢时间尺度反馈； 2. 利用信道时间相关性。	特定的训练网络可以有效降低反馈开销，提高恢复准确度，但需要大量的信道训练数据进行更新。计算开销大且无法处理信道突变等情况。

表1.1总结了现有各种信道反馈方案的优缺点。可以看到，这些方案相比传统的MIMO有限反馈方案，虽然可以有效降低信道反馈开销，但它们往往依赖于一些较强的假设或者特

定的场景，难以适应实际系统的鲁棒性要求。

1.2.2 用户调度研究现状

用户调度是目前5G及B5G系统中无线资源管理（Radio resource management, RRM）中的一项重要功能。受限于实际无线链路系统中发送天线数目和空域自由度，基站无法在同一时、频域传输资源上为所有用户提供通信服务，因此可按照一定准则选取部分用户进行通信，以获得更多的用户分集增益，提高系统整体的传输性能。

随着移动通信技术的发展和系统演进，不同时期调度方案和准则有所差别。在无线通信系统发展早期（3G及之前），基于TDMA或FDMA多址技术，基站将不同用户的传输数据映射到分离的时、频域资源上发送。因此，对每一个传输时隙或窄带资源，基站基于不同的用户调度算法和准则在系统容量和用户服务质量（Quality-of-service, QoS）性能间权衡。步入4G LTE后，系统结构更加精细化，时频域的最小调度资源为资源块（Resource block, RB）。频域包调度（Frequency domain packet scheduling, FDPS）问题作为传统用户调度问题的一般延申开始被广泛研究。相较之前用户调度准则，FDPS更加考虑实际LTE的系统特性，综合利用用户上报的CQI、用户传输队列长度、误块率-信噪比映射表等信息在每个传输时延间隔（Transmission time interval, TTI）上执行包调度准则以最大化系统吞吐量或服务性能。进一步的，5G NR系统通过广泛应用massive MIMO技术，极大地提升了系统容量和频谱利用率，同时也为调度方案的设计带来了更多的空域自由度。相较LTE的FDPS，5G系统的频域资源分配（Frequency domain resource allocation, FDRA）方案有了更精细的设计，即分为Type-1和Type-0两种，其区别在于调度的给某用户的RB集是否在部分带宽（Bandwidth part, BWP）内连续。而除此之外，实际系统的约束还包括不同用户的数据流阶数（Rank）、MCS和缓存器长度等，这无疑大大增加了调度方案设计的复杂度。为了降低设计难度，有不少研究者基于合理假设，设计忽略部分实际约束下的各种多用户调度和联合收发预编码方法以最大化系统的传输和服务性能。比如此时调度器会在各RB上独立的比较各用户的度量（Metric）值，并从中选择效用最大的用户提供服务，即

$$k_n^* = \arg \max_{k \in \mathcal{K}} \{m_{n,k}\}, \quad (1.1)$$

其中 $m_{n,k}$ 表示在第 n 个RB上调度用户 k 后可得到的度量值， \mathcal{K} 为系统用户集合。需要指出的是，上述准则只考虑每个RB只分配给一个用户的情况，称为单用户MIMO（Single-user MIMO, SU-MIMO）调度。随着发送天线数目增大，可以进一步利用空分复用（SDMA）在每个RB同时调度多个用户及发送多流以获得更高的调度增益，称为多用户MIMO（Multi-user MIMO, MU-MIMO）调度。根据不同的业务需求和优化目标，度量函数可由包括传输队列长度、

信道质量报告、历史调度结果、缓存器长度和服务需求质量在内的多种状态信息联合确定。我们需要重申，这种调度策略在每个RB上独立进行，虽然可以获得最大化的系统效用，但却忽略了上文提到的实际链路约束，这使得其实际实施的开销过大。因此在实际系统中，考虑实际链路约束的FDPS或FDRA方案可能是更易被采纳的。因此，通过类比文献^[16]，我们基于不同的假设将LTE及NR系统中的调度问题大致分为以下三类并详细介绍。

A. 未考虑链路约束且发端信道未知的调度策略

当用户信道条件未知时，调度策略一般基于时不变（Time-invariant）和无错传输介质等假设而设计。虽然这类场景在LTE或NR系统中几乎不会出现，但其设计思路对其它场景有一定的借鉴意义。具体来说，该场景下的调度策略分为以下几类：

- (1) 先入先出（First-in first-out, FIFO）准则：当信道条件未知时，最简单的调度策略可以是FIFO。此时度量值表示为

$$m_{n,k}^{\text{FIFO}} = t - T_k, \quad (1.2)$$

其中 t 和 T_k 分别表示当前TTI时刻和第 k 个用户调度请求到达的TTI时刻。因此，基于上述度量值，调度器完全根据各用户调度请求到达的先后顺序执行服务。这种方法虽然简单，但是没有考虑公平性和系统容量。

- (2) 轮询（Round robin, RR）准则：顾名思义，基于RR的调度算法策略是周期性地对连接用户进行轮流服务，以保证各用户的公平性。其度量值 $m_{n,k}^{\text{RR}}$ 为用户 k 距上一次调度经过的时间。这种调度策略虽然为服务用户提供了均等的接入机会，但它会造成系统资源的极大浪费，同时也无法确保为各接入用户提供相同的体验传输速率。

- (3) 加权队列公平（Weighted fair queuing, WFQ）准则：WFQ调度准则是RR调度策略的优化形式。其度量表示为

$$m_{n,k}^{\text{WFQ}} = w_k m_{n,k}^{\text{RR}}, \quad (1.3)$$

其中 w_k 表示用户 k 的轮询权重。调度器可以通过调整合适权重来提升系统容量性能，但是仍然无法完全保证用户的传输速率机会均等。

- (4) 盲等吞吐量（Blind equal throughput, BET）准则：BET调度可以保证接入用户的平均数据传输速率是均等的^[17]。调度器保存每个用户的历史平均吞吐量，并据此设计权重可变的度量值：

$$\begin{aligned} m_{n,k}^{\text{BET}} &= 1/\bar{R}_k(t-1), \\ \bar{R}_k(t) &= \alpha \bar{R}_k(t-1) + (1-\alpha)r_k(t), \end{aligned} \quad (1.4)$$

其中 $\bar{R}_k(t)$ 表示调度器保存的截止到 t 时刻关于用户 k 的历史平均吞吐量； $r_k(t)$ 表示 t 时刻用户 k 的实际传输速率。 $\alpha \in [0, 1]$ 为权衡因子，用以控制度量值更新的速度。可以看出，BET调度可以确保其服务用户的长期数据传输流量是均等的，这种特性在实际中往往非常重要。

B. 未考虑链路约束但发端信道已知的调度策略

对于实际LTE或NR系统，调度器往往可以获得各用户的下行CSI。对TDD系统，基站可利用上下行信道互易特性，基于上行探测参考信号（Sounding reference signal, SRS）获得下行信道的估值。对FDD系统，下行CSI的获取通过完整的CSI-report流程实现。这些信息可以辅助调度器设计性能更优的调度准则。该场景下调度策略有：

- (1) 最大吞吐量（Maximum throughput, MT）准则：MT调度准则下，调度器期望在每个TTI最大化下行吞吐量，度量值表示为

$$m_{n,k}^{\text{MT}} = \log(1 + \gamma_k^n(t)), \quad (1.5)$$

其中 γ_k^n 表示用户 k 在第 n 个RB上的接收信噪比（Signal-to-noise ratio, SNR）。MT虽然能够最大化小区吞吐量，但它无法为接入用户公平的分配频谱和传输资源。对处在蜂窝边缘的信道条件较差的用户，可能长时间无法被调度。因此实际调度器往往会在MT（使小区吞吐量最大化）和BET（保证吞吐量在用户之间公平分配）之间进行权衡，在尽可能利用各用户不同信道条件取得调度增益的同时，仍然满足一定程度的公平性。

- (2) 比例公平（Proportional fair, PF）准则：基于PF准则的调度算法期望在用户公平性和频谱效率间权衡，因此其度量值是MT和BET方法的组合，即

$$m_{n,k}^{\text{PF}} = m_{n,k}^{\text{MT}} \cdot m_{n,k}^{\text{BET}} = \log[1 + \gamma_k^n(t)] / \bar{R}_k(t - 1). \quad (1.6)$$

可以看出，PF准则将BET计算历史平均吞吐量的倒数作为权重因子，用于控制MT中各用户吞吐量度量值。总体来说，PF是应用较为广泛的调度准则，因此也有研究者在PF准则的基础上提出若干改进增强算法。比如文献^[18]提出了一种广义比例公平（Generalized proportional fair, GPF）准则，其度量值表示为

$$m_{n,k}^{\text{GPF}} = (m_{n,k}^{\text{MT}})^\xi \cdot (m_{n,k}^{\text{BET}})^\psi, \quad (1.7)$$

其中参数 ξ 和 ψ 各自作为MT和BET准则的幂指数系数，影响即时传输速率和历史平均吞吐量的重要性程度。MT、BET和PF等准则均可以视为GPF准则中参量取特定

值时的特殊场景。此外，文献^[19,20]也提出了类似方法，调度器可根据实际系统状况自适应的调整度量系数以满足不同的QoS需求。

- (3) 时频域联合调度：文献^[21]提出了一种时频域联合的两步调度策略。具体来说，对每个TTI，时域包调度器（Time domain packet scheduler, TDPS）首先从候选用户中挑选一组传输质量较好的用户子集。之后再由FDPS基于度量值进一步为每个RB确定调度用户。这种两步调度方案的优点是可以通过在每个TTI选取用户子集来缩小频域包调度的搜索优化空间，从而降低后续的计算开销。此外，时、频域调度器可使用包括MT, BET, PF等在内的不同的调度策略，实现更灵活的资源配置。但另一方面，由于FDPS每次只能在部分用户变量集合内调度，因此无法在更深层次上挖掘调度增益。
- (4) 时延敏感性调度：以上各调度准则大多只考虑在单个TTI内的对用户公平性和系统容量进行性能权衡，然而在实际传输系统中，包传输时延也是用户QoS体验的重要组成部分。基于此，各用户的包时延敏感性也作为一项重要的度量指标为一些调度器所考虑。比如文献^[22]提出一种跨层（Cross-layer）算法在MCS、BLER和发送功率等约束下解决平均时延最小化问题。文献^[23]则在MT度量基础上，将包传输时延作为权重因子，提出一种与PF准则相似的度量函数。该度量准则下，调度器可为包时延较大的用户设置更多的权重以为其分配更多的传输资源。文献^[24]提出的修正最大权重时延优先（Modified largest weighted delay first, M-LWDF）的调度准则也是类似的思路，即在PF度量函数的基础上补充与包传输时延相关的权重系数，改进的度量函数为

$$m_{n,k}^{\text{M-LWDF}} = \beta_k D_k \cdot m_{n,k}^{\text{PF}}, \quad (1.8)$$

其中 D_k 和 β_k 分别表示用户 k 的传输时延和重要性权重系数。

- (5) 缓存器容量受限调度：在实际通信系统中，数据发送和接收处理之前往往都会预保存在缓存器（Buffer）内。根据缓冲器大小和数据突发流量的强弱，调度模型可分为全缓冲（Full-buffer）和有限缓冲（Finite-buffer）两种^[25]。在全缓冲模型中，一般假设发送和接收端部署无限容量的缓冲器，上述几种调度准则均属此类。这种建模不必考虑数据包溢出或丢失等实际问题，因此算法设计比较简单。然而在大部分场景下，尤其是对某些流量突发（Bursty-traffic）业务来说，基站和用户的缓冲区中传输或接收的数据量是有限的，此时需要考虑发送数据量受限或接收包丢失等实际非理想约束。比如文献^[26]提出了一种几何规划（Geometric programming, GP）方法来解决在发端有限缓冲下的队列最小化问题。文献^[27]提出把有限缓冲的非理想效应建模为加权和速率的优化问题。文献^[28]则基于最大和（Max-

plus) 排队方法将包延迟中断概率约束转化为最小数据传输速率的约束进行求解。文献^[29]考虑接收端有限缓冲导致的包丢失问题, 提出一种期望同时满足中断概率约束和吞吐量公平性设计原则的调度策略。在该方案下, 调度器可基于用户上报的各自缓冲器剩余容量和其业务需求动态的调整资源分配优先级。

C. 考虑实际链路约束的调度策略

前两小节已介绍了LTE或NR系统中, 当调度器已知或未知各用户CSI时的各种调度策略。但需要指出的是, 在实际MU-MIMO调度场景中, 仍然存在许多耦合的非理想链路约束, 比如: 1) 为了降低系统复杂度, 各调度用户在其被分配的RB上需要保持相同的传输流数目和MCS阶数; 2) 考虑到实际BLER的影响, 用户的下行可达速率与MT准则中所建模的形如香农容量的公式不同, 需要基于MCS映射表分段查询获得^[30]; 3) NR系统引入了部分带宽, 对FDRA有了更精细的设计, 即可细分为Type-0和Type-1两类资源映射策略, 分别对应可非连续分配的资源块组(RB group, RBG)级调度和必须连续分配的RB级调度^[31,32]; 4) NR系统规定了更精细的帧结构, 每个TTI最多为1ms, 因此调度器需要谨慎考虑算法复杂度以满足实时性约束, 等等。基于此, 本小节将分点介绍考虑存在上述链路约束的调度方法, 具体包括:

- (1) SU/MU-MIMO的MCS和数据流约束: 根据各RB是否能被多用户复用, 调度问题可分为SU-MIMO和MU-MIMO场景。对SU-MIMO调度, 每个RB只能分配给一个用户使用, 因此目标优化形式相对简单。而在MU-MIMO问题中, 每个RB可被多用户复用以获得更多的空分复用增益, 但此时调度策略比较复杂。近年来有一系列研究者考虑了此场景下的调度问题。比如文献^[33]证明了在考虑多数据流分配时的优化是一个非确定多项式困难(Non-deterministic polynomial hard, NP-hard)问题, 并针对全反馈和部分反馈两种场景各提出了一种在PF准则下的基于贪婪迭代的近似算法。在此基础上, 文献^[34]针对SU-MIMO提出了一种广义的度量函数, 可以同时刻画包括PF在内的其它权重度量准则。在假设RB只能被多用户复用单流的前提下, 文献^[35]考虑了MU-MIMO场景的近似调度算法。文献^[36]考虑了包含MCS在内的LTE实际约束, 并同时研究了在积压流量模型和有限队列模型下的调度问题。文献^[37]提出了一种通用近似方法解决包含MCS优化的调度问题。但它们只涉及SU-MIMO场景。
- (2) 考虑BLER的调度策略: 在实际传输系统中, 当考虑传输BLER的影响时, 用户的

下行可达速率（吞吐量）与MT准则中形如香农容量的公式不同，而是

$$m_{n,k}^{\text{BLER}} = m_{n,k}^{\text{MT}} (1 - \text{BLER}_{n,k}^{\text{MCS}}), \quad (1.9)$$

其中 $\text{BLER}_{n,k}^{\text{MCS}}$ 为误块率，是一个与传输块大小（Transport block size, TBS）、信道编码方式及MCS均有关的复杂变量。由于其闭式表达式难以求得，因此不得不寻求简化方法。比如文献^[33,35]中采用查表法获得BLER的近似估计。调度器基于各用户关于CSI的上报信息，计算SNR并从SNR-BLER映射表中查询对应BLER。文献^[38]则将SNR与BLER直接建模为简单的二元映射关系，当SNR小于门限时，认为BLER为1（即实际吞吐量为0）。这类方法的优势是实现复杂度低，但缺点是需要基于大量链路级仿真提前获得各种信道、调制场景下的映射结果，鲁棒性较差。同时这种基于映射的方法大多建模为离散优化问题，难以直接求解，一般只能通过设计基于贪婪等的启发式方法近似求解。除此之外，文献^[39]提出使用基于深度学习的方法预测BLER，这类基于网络的方法在特定场景下可以获得较准确的预测结果，但网络训练开销较大，实际鲁棒性差。

- (3) 频域资源分配约束：根据用户频域资源分配的连续性，调度问题又可分为连续或非连续调度。以5G NR FDRA为例，系统规定了Type-0和Type-1两种资源分配模式^[31,32]。在Type-0模式下，系统可为每个用户分配非连续的频谱资源，此时最小的频域资源调度粒度是一个RBG。对Type-1模式，系统可在更精细的RB粒度上设计调度算法。但考虑到下行信令开销，系统限制为每个用户分配的RB资源必须连续。相较而言，由于连续性约束的存在，Type-1类型的调度问题更为复杂，这涉及到一个复杂的离散空间组合优化问题，因此大部分研究者通过设计各种启发式的近似算法来降低复杂度。比如文献^[40]提出了一种次优算法来解决在连续RB分配约束下的和速率最大化问题，并计算了所提方案与最优解的性能差距。文献^[41]针对加权和速率问题，提出一种基于邻位扩展的启发式调度方法。文献^[42]也是类似的思路，设计从可用带宽资源的两端逐步向中心扩展的资源分配算法以满足连续性约束。这类方法通过利用贪婪的思想虽然降低了计算复杂度，但也受限于简单的调度策略而无法挖掘更多的性能增益。
- (4) 计算复杂优化：在实际传输中，调度器需要在每个TTI（NR中最多为1ms）内确定调度策略以满足实时性约束。虽然前文所述的大部分调度策略都结合了启发式算法来降低计算复杂度，但当调度用户数或调度资源过多时，这些方法的计算开销可能仍然不满足实际需求。基于此，有研究者基于图形处理单元（Graphics processing unit, GPU）计算架构设计并行算法来降低处理时延。比如文献^[43]基于GPU并

行计算结构，将调度器部署在了Nvidia Quadro P6000平台上，测试可以满足100us的实时性调度需求。文献^[38]考虑了MU-MIMO场景中包含MCS和数据流约束的调度问题，基于商用GPU平台设计了一种mCore调度器以实现1ms的实时性调度需求。

表 1.2 现有用户调度策略总结

调度方案	研究方向	优缺点
未考虑实际链路约束的调度	1. 发端信道未知：FIFO、RR、WFQ和BET等； 2. 发端信道已知：MT、PF、GPF和M-LWDF等。	可以最大化地发掘调度增益且调度算法的复杂度低。但不考虑实际链路约束使得方案实施开销过大，几乎无法适用现有系统。
考虑实际链路约束的调度	1. SU-MIMO / MU-MIMO； 2. MCS和传输流约束； 3. BLER对吞吐量的影响； 4. 频域资源连续性分配约束； 5. 实时调度对计算复杂度的要求。	考虑链路约束的调度算法的实施开销低，且方案可行性高。但复杂约束使得调度算法难以设计。大部分启发式的调度算法无法挖掘所考虑场景的全部增益潜力。

表1.2总结了现有调度方案的优缺点。可以看到，在考虑实际链路约束的调度方案研究中，大部分算法仍然是启发式的。这虽然降低了实施复杂度，但是无法充分挖掘5G及B5G系统中潜在的调度增益。

1.2.3 下行预编码研究现状

基站在获得各用户下行CSI并确定当前时频资源上的服务用户后，可以通过预编码技术对各用户的发送信号进行预处理，从而有效抑制不同用户或天线间的干扰及噪声。基于预编码技术或波束赋形技术，基站也可以精细地调整波束方向以将能量集中到目标移动终端上，从而在不增加带宽和发送功率的情况下成倍地提高频谱利用率和能量效率。根据在发射端对信道矩阵的变换方法，预编码技术可以大致分为线性和非线性两类。作为一种经典非线性预编码方法，脏纸编码（Dirty paper coding, DPC）虽然可以实现近似最优的频谱效率，但是实现过程较为繁杂。尤其是对于massive MIMO系统，其计算复杂度更高。而线性预编码系统设计比较简单，易于实现。基于CSI在收发端是否已知，线性预编码技术又可以分为两类，一类是只考虑发射端处理的预均衡技术，如迫零（Zero-forcing, ZF）算法、匹配滤波（Match filter, MF）算法和最小均方误差（Minimum mean-square error, MMSE）预编码等。另外一类则考虑收发端联合处理的预编码算法，如奇异值分解（Singular value decomposition, SVD）预编码算法和几何均值分解（Geometric mean decomposition, GMD）

GMD预编码算法等。相较于非线性预编码方案，线性预编码因其较低的实现复杂度而成为massive MIMO系统的研究热点。此外，为了解决传统数字预编码方案中射频链路和天线数目的强对应关系，研究者还提出了一类混合预编码（Hybrid-precoding）的设计方案。通过将一个小高维模拟预编码器和一个小降维的数字预编码器相连，有效降低了基站的硬件成本和功耗。

然而需要注意，上述关于massive MIMO系统预编码方案的研究大多基于一些理想的假设和判定，比如：1) 认为输入的数据是理想高斯分布的；2) 假设接收端和发射端已知完美的CSI；3) 大多以最大化输入输出数据互信息（速率）为准则设计预编码方案等。然而在实际系统中，上述的一些假设可能无法得到满足。例如，输入信号往往需要通过正交幅度调制（Quadrature amplitude modulation, QAM）等，其取值并不满足高斯分布，而是来自有限字母（Finite-alphabet）星座集。同时，由于不可避免的信道反馈或估计的误差，收发端可能无法获得精确的CSI，从而对预编码系统的性能产生一定影响。另一方面，相较于最大化的发送速率，某些应用场景可能更重视误比特率（Bit error rate, BER）或BLER等指标。此时则需基于应用需求设定更合适的系统效用函数，这对预编码系统的优化设计提出了更高的挑战。基于此，我们在下面几小节中对目前的预编码技术和研究问题进行详细介绍。

A. 全数字线性或非线性预编码

根据信道矩阵的处理方式，目前数字预编码技术可以大致分为线性预编码和非线性预编码。前者只涉及对信道矩阵的简单变化，而后者由于需要一些求模等非线性处理，因此结构更为复杂。根据不同的设计准则以及收发端是否已知CSI，目前成熟的线性预编码可分为：完全在发射端实现的预均衡技术，如ZF、MF和MMSE等；以及收发端联合设计的预编码算法，如SVD和GMD。而对于非线性预编码，大部分是以DPC为基础的各类方法，如汤姆林森-哈拉希玛预编码（Tomlinson-Hiroshima precoding, THP）和矢量扰动（Vector perturbation, VP）预编码。虽然非线性预编码在信道容量、误码率特性等方面都优于线性预编码，但由于其较高的计算复杂度而无法广泛适用于massive MIMO系统。这些方案具体为：

- (1) ZF预编码：ZF预编码的核心思想是通过信道矩阵求逆处理进而计算得到预编码矩阵，因此也被称为信道逆转（Channel inversion, CI）预编码技术。经过ZF预编码处理后的信号，各用户间的干扰可被消除。ZF预编码在高SNR下可取得较好的传输性能；但在低SNR下，由于其未考虑噪声对系统的干扰，会因为放大噪声能

量而降低用户的可达速率^[44]。同时，ZF预编码由于涉及到矩阵求逆运算，计算复杂度较高，因此不适合问题规模较大的场景。

- (2) **MF预编码**：MF预编码的核心思想是只最大化的目标用户的接收SNR。该方案的最大优势是结构简单，计算复杂度极低。但由于其未考虑各用户间干扰，因此在并发用户信道相关性较高时，该方案的性能会急剧恶化^[45]。不过需要指出的是，在massive MIMO系统中，随着发送天线数目的增大，各用户接收天线对应的信道矢量将愈发地表现出彼此正交的趋势。因此在发端天线数较大时，MF可以展现出逼近ZF的性能优势。
- (3) **MMSE预编码**：MMSE算法的核心思想是通过最小化接收与发送信号之间的均方误差（Mean-square error, MSE）来兼顾考虑用户间干扰和高斯白噪声干扰的影响。通过调整用于衡量这两种干扰重要性大小的权重因子，MMSE可以在ZF和MF预编码间灵活权衡以充分结合二者的优势^[46]。因此，MMSE预编码也是massive MIMO系统中应用最广泛且性能可靠的重要预编码技术之一。
- (4) **SVD预编码**：若收发两端同时已知CSI，可以采取基于信道矩阵SVD分解的方式来进行联合设计。具体来说，收、发两端可分别取信道SVD分解后的左、由奇异向量酉矩阵作为各自的预编码矩阵。通过这种方式，SVD预编码将信道分解为一系列在空间上并行传输且互不干扰的子信道。此时的系统容量就是各子信道的容量和。已经证明，结合SVD预编码和基于注水（Water-filling, WF）的子信道功率分配方法是使得MIMO容量可达的最优设计方案^[47]。虽然SVD有显著的性能优势，但其计算涉及到矩阵SVD分解，因此复杂度也不低。
- (5) **GMD预编码**：将SVD分解预编码算法与注水算法结合虽然可以最大化MIMO系统容量，但也会可能导致系统子信道的能量相差较大，从而增加如调制、解调等系统链路的实现复杂度。GMD可有效弥补这一缺陷，它将系统信道分解成增益相同的子信道，使得其能量大致均衡。但作为代价，GMD在MIMO信道容量性能上相较SVD方案是有损失的。
- (6) **非线性预编码**：非线性预编码一般通过在发端对信号进行复杂、非线性的预运算操作，使得接收端在存在多用户干扰时也可基于逐级判决的结构将干扰消除。因此非线性预编码相较线性预编码在系统容量及BLER等传输性能上都有明显优势。但这类方法不仅要求收发两端已知完美CSI，还需要较大的运算量，系统实现复杂度也很高。因此其难以应用于massive MIMO系统，一般会用作分析系统容量性能上界的一种工具。非线性预编码包括涉及QR分解和Gosta编码的DPC预编码^[48]、涉及算数取模的THP预编码^[49]和基于最优扰动矢量选择的VP预编码^[50]，这里不再

详细介绍。

B. 模数混合预编码

传统的纯数字预编码方案通过为每根天线配置一条射频链路，实现对发送信号的全数字预编码处理。这类方法虽然具有较高的自由度，能实现接近理想的传输特性，但也带来了较大的实现复杂度。尤其是对massive MIMO和毫米波系统，基站的硬件成本和功耗难以忍受。另一方面，模拟波束赋形的结构能利用移相器对发送信号的相位进行调整，通过改变权重系数实现波束赋形，提高能量效率。它作为一种简单有效的方案，最初被应用于室内毫米波系统。不过，虽然模拟波束赋形可以实现相位调整，但却无法对信号的幅度进行控制。而且，由于天线阵列通过移相器网络与一条射频链路连接，传统的模拟波束赋形无法支持多流数据的传输。因此，结合纯数字预编码和模拟波束赋形两种结构的特点，混合预编码逐渐成为massive MIMO系统预编码中的重要研究方向之一。基于信道特性的变化速度，现有的混合预编码研究大致可以分为快时间尺度混合预编码（Fast-timescale hybrid precoding, FHYP）和两级时间尺度混合预编码（Two-timescale hybrid precoding, THYP）两类。

FHYP方案通常假设已知实时的CSI，并以此分别设计模拟和数字预编码器。文献^[51]针对massive MIMO系统，提出了一种低复杂度的相位迫零混合预编码方案。文献^[52]针对单用户的毫米波MIMO系统，提出了一种基于正交匹配追踪（Orthogonal matching pursuit, OMP）的稀疏预编码方案与合并算法。近些年来，深度学习也开始被应用到无线通信领域以解决massive MIMO系统的混合预编码问题^[53-55]。这些基于深度学习方法的核心思想，是利用深度神经网络（Deep neural network, DNN），将发送接收数据之间的处理过程视为黑盒，直接优化模型参数得到近似的混合预编码设计。这类方法的优势在于，由于神经网络的数据驱动特性，不需要关于系统模型的准确信息，深度神经网络就可以通过训练自动匹配未知的系统模型，因而对实际系统的非理想性并不敏感。但是，现有的深度学习方法尚还缺乏完整的理论支持，经验性质的调参方法使得网络泛化能力较弱。而且在实际系统中，为了保证性能，模型参数需要根据信道统计特性的变化而相应更新，这也会导致不低的计算开销。因此深度学习方法能否可靠地部署于实际通信系统还有待进一步证实。总结来说，FHYP方案虽然能灵活的调整预编码器的设计，但实际中获取实时CSI会导致大量的信令和反馈开销。另外，由于不同的子载波的实时CSI可能并不相同，FHYP需要为不同的子载波设计相应的模拟预编码器，这会大大提升设计的复杂度。

为了克服FHYP方案的缺点，文献^[56,57]提出了THYP的方法。在该方案中，模拟预编码

器的设计只与信道统计信息有关。数字预编码器根据有效信道（即实际信道与模拟预编码矩阵的乘积）来设计。基于此，文献^[58]提出了一种在线的算法框架，实现一般THYP问题的优化求解。文献^[59]基于两级随机优化方法，设计了一种用于下行链路的多小区大massive MIMO系统的THYP方法。可以发现，由于THYP算法根据信道统计特性设计模拟和数字预编码器，不需要实时的高维CSI，因此它可以显著降低用于CSI估计的信令开销。同时，由于不同子载波上的信道统计特性大致相同，因此THYP方案只需要设计一个模拟预编码器即可覆盖所有的子带（Subband），有效降低了硬件开销和系统复杂度。基于上述优势，THYP在实际通信系统中的应用要比FHYP更广泛。

C. 非理想CSI的线性预编码设计

在实际通信系统中，受限于有限的反馈信道、导频开销以及其他非理想的硬件条件，比如功率放大器的非线性效应、量化误差和正交两路信号失衡等，收发两端可能无法估计准确的CSI信息。另外，在massive MIMO系统中，若信道变化速率过快，基站也几乎不可能实时准确地获取完整的CSI。因此对于非理想CSI场景，应该综合考虑信道估计误差、反馈时延等带来的不确定影响，并以此设计更鲁棒的预编码方案。

目前克服非理想CSI的预编码设计方案大多基于鲁棒优化的方法。根据建模方法和优化目标的不同，鲁棒优化又可以分为最差优化（Worst-case-approach）和基于贝叶斯方法的随机优化（Stochastic-approach）。最差优化方法的目标是尽量改善最坏信道估计条件下的传输性能，一般通过假设信道估计误差在某预定义的不确定区域内来最优化其性能。文献^[60]针对单用户的场景，提出了一种用于波束成型的鲁棒设计方法。在考虑信道估计误差模值受限的条件下，通过优化最坏情况下的接收机最大SNR来设计预编码器。文献^[61]将这一方法扩展到了多用户场景。文献^[62]则在一定程度上改进了^[63]中提出的最差场景MMSE最优化的预编码设计方法。作者重新构造了MMSE的优化表达式，并将其应用到包括和速率最大化在内的其他鲁棒优化问题。

另一方面，贝叶斯方法和核心思想是通过建立以CSI估计值为条件的实际CSI随机分布来优化系统的平均性能。这类方法中，信道通常被直接建模为独立同分布的复高斯随机变量。比如文献^[64-66]分别考虑了以平均互信息，BER和MMSE为指标的优化方法。但也有研究者将因信道估计，信道老化（Channel-aging）导致的误差描述为信道不确定性，并将其建模为独立同分布的零均值单位方差复随机高斯变量。这种方法通常根据已有的先验信息和当前估计值，建立实际CSI的模型。比如，对于FDD模式下的massive MIMO系统，文献^[67]通过划分长期CSI和短期CSI，设计了一种双重结构的预编码器以最大化和速率。其中

长期CSI由于变化缓慢，可被视为已知的先验信息，而短期CSI则通过当前的估计值和不确定的复高斯随机变量的联合建模得到。上述的结构有效降低了信道反馈开销。文献^[68]则针对时变信道，将前一时帧的信道作为先验，联合不确定性的随机模型估计当前时帧的信道信息，并以此设计了一种鲁棒的屏蔽波束成形（Masked-beamforming）方法，实现了多用户的安全通信。

D. 有限星座集输入时的预编码设计

上述各类预编码方案，大多都假设其输入信号服从理想的高斯分布。尽管高斯输入信号可以实现信息论意义上最优性能，但在实际系统中，输入通常取自于离散有限的字母星座集合，例如相移键控（Phase shift keying, PSK）、幅度脉冲调制（Pulse amplitude modulation, PAM）和QAM等。而一系列工作表明^[69-73]，当输入分别是高斯信号和有限字母信号时，即使选择相同的预编码方案，其性能也会有较大差异。因此，对于这类基于Finite-alphabet输入的系统，需要重新评估并设计对应的预编码器。

近年来的大部分工作主要聚焦于如何在存在Finite-alphabet限制时，最大化输入输出之间的互信息。比如文献^[69]提出了一种开创性的水银注水（Mercury water-filling, MWF）算法，实现独立并行高斯信道上的最优功率分配。该工作的意义在于，它证明了当各子信道衰落相同时，具有更密集分布的星座输入需要分配更多的功率。而当各子信道输入星座相同时，需要把更多的功率分配给更好的子信道，这与传统注水方案的结论一致。这些结果表明，在输入信号为Finite-alphabet时，功率分配不仅与信道增益有关，还受星座集分布密集程度的影响。此外，在假设发射机已知完美CSI时，文献^[70]针对实值域的输入和信道数据设计一种迭代算法来求解预编码矩阵。但是需要指出，将输入信号、信道和预编码器参数限制在实数域只能得到一个次优的设计方案，求得的互信息也远远低于可能实现的全局最优值。文献^[71]则进一步在复数域考虑了上述问题。在证明互信息表达式是关于预编码矩阵的凹性函数后，作者设计了一种基于回溯线搜索的快速迭代算法求得近似最优的预编码矩阵。近年来，也有研究者开始针对massive MIMO系统提出低复杂度的预编码设计算法。比如文献^[72]考虑当输入为近似高斯分布时的并行高斯信道的广义功率分配问题，提出了一种鲁棒的注水功率分配方法。相较MWF，这种方法虽然损失了部分性能，但其不需要计算给定星座集下收发信号MMSE逆函数的精确闭式解，因此降低了实现复杂度。文献^[73]则考虑多用户的massive MIMO系统，着重研究低复杂度的ZF和MF线性预编码性能。在假设完美和不完美信道条件下，作者推导了在采取简单线性预编码时的可达互信息上界的表达式，并证明了当天线数目足够多时，采取ZF和MF预编码可以使有限星座输入达到与高斯输入相同的互信息性能。

综上，现有的针对Finite-alphabet输入的预编码研究大多只涉及简单的应用场景，比如通常以最大化输入输出互信息为准则，通过设计迭代算法求解近似最优的预编码矩阵。这类方法的计算复杂度通常比较高，虽然文献^[73]提出可以使用简单的线性ZF和MF预编码以降低计算开销，但其系统实际的吞吐量或BLER等性能会受到严重影响。

表 1.3 现有下行预编码技术总结

预编码方案	研究方向	优缺点
全数字的线性、非线性预编码	1. 线性预编码：ZF、MF、MMSE、SVD和GMD等； 2. 非线性预编码：DPC、THP和VP等。	线性预编码往往具有更低的实现复杂度，非线性预编码则在系统容量、BLER等性能上更有优势。在massive MIMO系统中，大部分都采用线性预编码方法。
模数混合预编码	1. 快时间尺度FHYP； 2. 双时间尺度THYP。	FHYP可根据瞬时CSI灵活调整预编码设计从而挖掘更大的性能增益，但需要极大的信令和反馈开销。THYP基于信道统计特性将模拟和数字预编码器分离设计、在不损失太多性能的同时有效降低了实现开销。
非理想CSI预编码技术	1. 最差场景优化方法； 2. 随机优化方法。	最差优化的目标是尽量改善最差估计信道下的传输性能；随机优化的目标是在给定实际CSI分布下优化系统的平均性能。
有限星座集输入时的预编码技术	1. 最优MWF功率分配； 2. 鲁棒近似算法等。	MFW虽然给出有限星座集下的最优功率分配，但涉及到计算关于MMSE逆函数精确且复杂的闭式解。近似算法虽然降低了计算复杂度，但性能损失不可忽略。

表1.3总结了现有下行预编码方案的研究技术。受限于篇幅，还有许多研究方向并未总结其中。但可以看到，现有大部分预编码大部分以系统容量为优化目标。同时考虑有限星座集输入、非理想CSI约束，并以编码系统传输性能（如BLER）为直接指标的预编码方案研究相对较少。

1.3 论文主要内容和结构安排

本论文主要研究了面向5G及B5G大规模天线系统中PHY和MAC层涉及的一系列待解决的关键技术，包括FDD系统高精度低开销信道反馈方案、支持空域并发且考虑链路约束的多用户频分调度策略、非理想CSI场景下以BLER为性能指标的下行预编码优化和时延敏感业务中长期功率控制策略的设计。为了解决这些关键技术难题，我们充分挖掘massive

MIMO信道的特性和问题本身的结构设计了一系列关键算法和完整解决方案，并将其转化为行之有效的实际应用成果。具体包括：1) 针对FDD系统中对宽带下行信道低开销且高精度的反馈需求等关键问题，本文提出了一种联合空频域的二级差分反馈压缩方案，设计了包含二级差分压缩、多原子矢量量化和离线-在线混合的字典更新等关键技术算法；2) 针对中重载场景下考虑频域资源分配约束同时支持多用户空域并发的用户调度关键难题，本文提出了一种非完全启发式的逐级消息传播频分调度算法，并精心设计了一种称为关键点采样的实施简化技术；3) 针对远距覆盖增强通信场景中以BLER为指标的编码系统下行预编码设计关键技术难题，本文通过引入链路评估方法中常用的度量函数提出了逐级迭代的下行预编码设计方法并进行了大量仿真实验，这为工业界评估此类问题的增益空间提供了详实的理论和实践指导；4) 针对时延敏感业务中常涉及的多用户长期平均时延约束下的长期功率控制难题，本文提出了一种基于逐级凸近似的新型受约束深度强化学习算法，降低了此类问题的求解复杂度并提高了数据利用效率。

需要指出，本文的研究成果不是某些算法在一些场景下的简单应用，而是结合当前通信系统中的关键技术难点并考虑实际系统链路约束后仔细设计的解决方案。比如本文所提出的信道反馈方案是在保证功能与适用场景与现有NR反馈技术几乎相同的前提下进行的优化设计；所提出的用户调度算法和简化技术也考虑了实际链路约束和实施开销；所设计的预编码算法及功率控制策略也都用于解决现有通信场景及业务中所面临的实际问题。具体而言，全文共分为六章，其研究内容和结构安排如下。

第一章为绪论，首先总体介绍5G和B5G通信系统中PHY和MAC层的研究背景和技术现状，之后针对几项关键技术，即信道反馈、用户调度和预编码等进行了详尽的研究调研和优缺点分析，最后总结本文的研究要点和各章节安排。

第二章针对FDD多用户massive MIMO系统的宽带下行CSI反馈问题设计了一种联合空频域的TSAC信道反馈方案。具体来说，通过利用massive MIMO信道的动态突发稀疏性和强时间、频域相关性，我们首先将宽带CSI投影至空、频域码本，提取出需要反馈的二进制稀疏支撑矢量和信道系数矢量信息。之后，我们精心设计了一种二级差分的方案实现了二进制稀疏矢量的低开销无损反馈，且反馈比特数可以根据信道变化速率自适应地调整。同时，为了保证量化精度的同时以较低的开销量化信道系数矢量，我们还提出了一种基于动态更新字典的多原子矢量量化方案，相较传统的标量量化方案显著降低了反馈开销。最后，我们基于协议标准搭建了完善的链路级仿真系统，证明了所提方案相较基线方案在反馈开销和量化性能上的巨大优势。

第三章研究了单小区中重载下行系统中支持空域并发的多用户频分调度问题，所提出的JSFD-SMP调度算法有效地处理了最大空域并发数和RB资源连续性分配这两个棘手的实

际链路约束。具体地，我们首先将原问题分解，通过在每轮迭代中构造一个只包含RB连续性分配约束的子问题将求解空间降低。对分解后的子问题，我们将用户和RB间仍然复杂的耦合关系建模成因子图，并设计了一种消息传播算法来有效求解。算法本身高度解耦的结构非常有利于在GPU等支持并行计算的多核平台上实现。另外，我们还提出了一种称为KPS的方法在不损失性能的同时进一步降低了实施开销。最后，我们基于大量的仿真实验比较了所提算法和现有几种基线调度方案在吞吐量和公平性等指标上的性能表现。实验结果证明了所提方案可以在保证用户体验速率公平性的同时显著提升小区吞吐量性能。

第四章研究了远距覆盖增强通信场景下的下行预编码设计问题。该类场景重点关心如何在低MCS下设计以BLER为性能指标的下行预编码方案。为了解决BLER函数难以刻画的困难，我们引入链路评估常用的准则将原本复杂的编码系统BLER优化问题转化为无编码系统中针对所选度量函数的优化问题。对转化后仍然非凸的问题，我们在非理想CSI场景下设计了一种基于逐级迭代的E-S-WMMSE算法来有效寻找目标函数的局部最优解。所提算法通过将转化问题泰勒展开，在每一次迭代中构造一个形如加权和速率最大化的子问题，并利用经典的WMMSE算法求解。为了验证方案的合理性，我们搭建了符合协议标准的链路级仿真平台，首先基于蒙特卡洛实验证明了度量准则对刻画此场景下BLER性能的准确性。之后，我们在理想CSI和非理想CSI两种场景下比较了所提方案和几种基线方案的BLER性能，并通过改变传输流数目和MCS阶数进一步探索了所考虑场景下的增益空间。仿真结果既证明了所提E-S-WMMSE算法的有效性，同时也为工业界评估此场景下可能的性能增益提供了详实的理论和数据指导。

第五章研究了5G时延敏感性业务场景下的长期功率控制问题。我们希望设计一种有效的功率控制策略，用以解决在满足各用户长期平均时延约束下的长期平均发送功率最小化的问题。为了克服传统受约束深度强化学习算法的更新计算复杂度高、对在线样本需求量大等缺陷，我们提出了一种新型的SCAPO算法框架。所提算法通过将原问题中复杂的目标和约束函数转化为其对应的凸代理函数，从而在每次迭代中只需求解一个凸优化问题，降低了更新计算复杂度。同时，在每次更新中，所提算法可以重用旧的样本数据，只需要少数的在线新样本即可完成训练，有效提高了数据利用效率。最后，我们将SCAPO算法与其他基线方案比较，仿真结果证明了所提算法的优越性。

第六章总结全文内容，并对后续工作和未来研究方向进行了展望。

2 大规模天线系统中基于双域压缩的高精度信道反馈技术

2.1 研究动机和主要成果

大规模阵列天线方案可以极大地提升系统容量和频谱利用率，已成为5G和B5G系统的核心技术之一。基站通过部署大规模天线阵列，充分利用波束赋形、混合预编码等技术，显著提升频谱利用率且降低能量损耗。理论上证明，当天线数量趋于无穷时，各用户信道矢量会趋于正交，使同一小区内的用户间干扰被消除。但是，massive MIMO技术在带来系统容量，频谱效率和鲁棒性等性能提升的同时也相应带来了诸多挑战。在波束赋形时，为了使发射波束精准传输到期望方向，BS需要获得下行CSI。而CSI的获取方式在FDD系统和TDD系统中存在很大不同。在TDD系统中，基站可以利用上下行信道之间的互易特性，根据用户在上行信道发送的信息，恢复下行信道状态信息。而FDD模式下CSI的获取较为复杂，通常基于导频训练，信道估计和信道反馈的方式。用户端利用最小二乘或最小均方误差等方法估计信道状态信息，随后将信道特征信息或预编码矢量通过上行链路反馈至基站，帮助实现资源调度和波束成形、功率控制。然而，随着天线数目和带宽资源倍增，导频训练和信道反馈的开销也逐渐增加。特别是在有限反馈的massive MIMO宽带系统中，如何降低信道反馈开销是亟需解决的一大问题。

从绪论中的调研结果可得知，现有关于信道反馈技术的研究路线可大致分为三类，即：1) 角度域的CSI反馈；2) 基于信道测量值的反馈方案；3) 基于深度学习的压缩反馈方案。在第二类方案中，用户不进行信道估计而直接将下行接收导频信号反馈至BS，由BS估计反馈信道。这种方法可以降低接收端的处理负担，但会导致极大的反馈开销。第三类反馈方案在特定场景下可以有效降低反馈开销，但数据驱动的网络需要大量的训练数据，计算开销大且鲁棒性差。因此目前FDD系统中基本采取的还是基于BS和用户端共享码本的CSI反馈方案。当用户端估计出信道状态信息后，只需反馈码本中与估计结果最相近的状态矢量下标。这种方法可以有效降低反馈开销，但是当天线数目过多时，高维码本的设计和更新，以较低时间复杂度遍历寻找最优状态矢量下标等都是比较困难的问题。在这类问题研究中，文献^[74]提出了一种基于CS的码本设计方法，但只适用于信道统计特性不变

的场景。文献^[6]利用信道在角度域上稀疏性，提出了一种称为AoD自适应的子空间码本设计方法。这种方案可以有效降低反馈开销，但只适用于信道AoD较长时间不变的场景，而这在实际中一般是无法实现的^[75]。文献^[76,77]研究了信道估计和反馈的联合设计方案。但這些隱式（Implicit）的反馈方法并没有提供一种直接反馈CSI矢量的方法，无法与现有的5G NR协议所兼容。另一方面，3GPP协议中也规定了一种称为Type-2的CSI反馈方法^[31]，可以跟踪信道AoD的变化并为BS提供高精度的信道反馈，但其代价是极大的反馈开销。

因此，为了克服上述现有方案的缺陷，本章节我们针对FDD多用户的massive MIMO宽带CSI反馈场景，通过利用massive MIMO信道的动态突发稀疏性（Burst-sparsity）、强时间相关性和频域相关性，设计一种联合空频域（Joint space-frequency domain, JSFD）的二级自适应压缩（Two-stage adaptive and compressed, TSAC）CSI反馈方案。所提方案相较现有方案的主要优势在于：

- (1) 不依赖如固定的AoD、信道统计信息等较严格的假设。所提方案可以逐帧跟踪信道AoD的变化并自适应地调整反馈比特。
- (2) 方案是一种直接反馈CSI矢量的方法，因此可与现有NR协议完美兼容。
- (3) 相较NR Type-2高精度基线方案，TSAC可以在实现更优的反馈性能的同时显著降低反馈开销。

本章剩余结构安排如下：第二子章节首先介绍系统模型和联合空频域的信道反馈方案框架；第三子章节详细介绍TSAC方案，具体分为对二进制支撑矢量的二级差分反馈方案、对稀疏系数的多原子矢量量化方案和基于离线-在线混合的量化字典更新方法；第四子章节将基于大量详尽的链路级仿真实验证明TSAC方案相较NR Type-2等几种基线在反馈开销和BLER传输性能上的优越表现；最后进行本章小结。

2.2 系统模型与方案概述

2.2.1 massive MIMO下行传输系统模型

如图(2.1)所示，本章节考虑一个下行的FDD massive MIMO系统，其中基站部署含有 N_t 根天线的均匀线性阵列（Uniform linear array, ULA），同时服务 K 个拥有 N_r 根天线的用户。一般 N_r 的典型取值为2或4，远小于 N_t 。系统可用的BWP由 N 个RB构成。记 $\mathbf{H}_{k,n} \in \mathbb{C}^{N_r \times N_t}$ 表示用户 k 在第 n 个RB上的下行信道。一般来说，在FDD下行链路中，BS无法像TDD系统一般直接基于上下行信道互易性获得下行CSI。因此其每个时间帧（Frame）将包含信道估计、信道反馈和下行传输三个阶段。具体来说，在下行信道估计阶段，BS向首先各用户发

送包括CSI参考信号（CSI reference signal, CSI-RS）在内的导频信号用于接收端的信道估计。在用户获得信道估值后，会将其量化压缩并反馈至BS。最后，BS将基于反馈的CSI执行调度策略或设计下行波束赋形矩阵以实现好的传输性能。

这里需要指出的是，根据BS处理业务或实现功能的不同，其要求用户上传的CSI形式也是有所区别的。比如当BS需要完整的下行CSI用于系统评估时，各用户将信道估值 $\hat{\mathbf{H}}_{k,n}$ 压缩量化后反馈至BS。由于这种方法的量化反馈开销非常大，因此大部分场景下各用户会上报数据量更小且对自己传输最有利的预编码矩阵指示（Precoding matrix indicator, PMI）信息^[31]。以基于信道SVD的预编码方案为例^[78]，此时CSI可以是 $\mathbf{H}_{k,n}^H \mathbf{H}_{k,n}$ 最大的 d 个特征值对应的特征向量构成的预编码矩阵 $\mathbf{V}_{k,n} \in \mathbb{C}^{N_t \times d}$ 。这两种上报CSI虽然在内容意义上有所区别，但就反馈形式而言，它们是一致的。因此我们不再过度区分二者，统称为 $\mathbf{H}_{k,n}$ 。

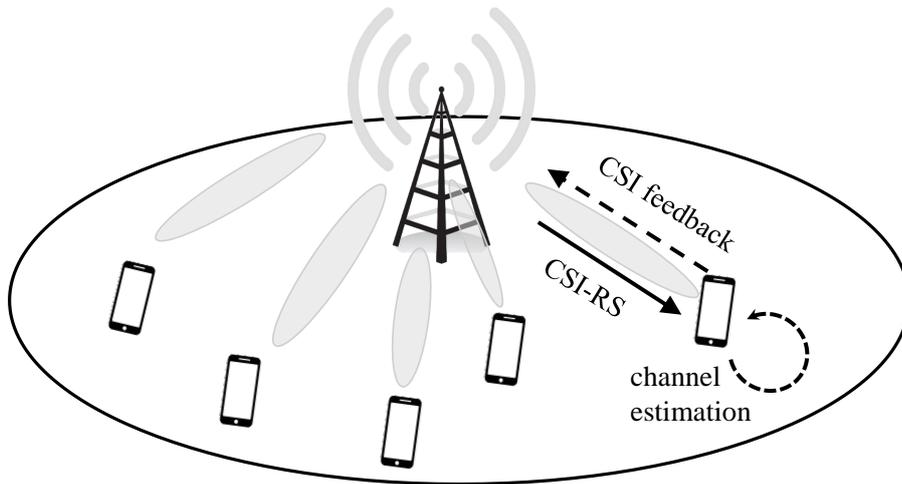


图 2.1 FDD massive MIMO系统模型

最后需要强调的是，本章只重点关注CSI反馈阶段的方案设计和性能比较，因此假设下行信道 $\mathbf{H}_{k,n}$ 是各用户完美已知的。学术和工业界均有许多关于massive MIMO信道估计的研究工作，这里不再赘述。

2.2.2 联合空频域的稀疏CSI的反馈模型

定义信道矢量 $\mathbf{h}_{k,n,i} \in \mathbb{C}^{N_t \times 1}$ 表示用户 k 在RB n 上的下行信道 $\mathbf{H}_{k,n}$ 的第 i 行或预编码 $\mathbf{V}_{k,n}$ 的第 i 列。考虑到各用户终端的反馈是独立进行的，我们假设各信道矢量 $\mathbf{h}_{k,n,i}, i = 1, \dots, N_r$ 也将独立进行量化反馈。为了书写清晰，我们将省略索引下标 k 和 i 。在实际中，为了降低反馈开销，一般不会直接反馈 \mathbf{h}_n ，而是利用 \mathbf{h}_n 的特性先进行信息压缩。比如，由于基站端周围的散射物有限，信道矢量在角度域上往往会表现出稀疏特性。在经典的基于射线的窄带

信道模型下^[79]，信道矢量可以表示为

$$\mathbf{h}_n = \sum_{j=1}^{P_n} g_{n,j} \mathbf{a}(\theta_{n,j}), \quad (2.1)$$

其中 P_n 表示可分辨的出射路径数目；对第 j 条子径， $g_{n,j}$ 和 $\theta_{n,j}$ 分别表示其复增益和出射角信息； $\mathbf{a}(\theta_{n,j}) \in \mathbb{C}^{N_t \times 1}$ 表示天线响应的指向向量，即

$$\mathbf{a}(\theta_{n,j}) = \left[1, e^{-j2\pi \frac{d}{\lambda} \sin(\theta_{n,j})}, \dots, e^{-j2\pi \frac{d}{\lambda} (N_t-1) \sin(\theta_{n,j})} \right]^T, \quad (2.2)$$

其中 d 为天线间距； λ 表示载波波长。实际中， P_n 一般远小于发送天线数目 N_t 。受此启发，用户可以首先求出 \mathbf{h}_n 在角度域码本 $\mathbf{A} \in \mathbb{C}^{N_t \times M_1}$ （ $M_1 \geq N_t$ 表示码本维度）上的稀疏表示，问题的数学表达式为

$$\min \|\mathbf{s}_n\|_0, \text{ subject to } \|\mathbf{h}_n - \mathbf{A}\mathbf{s}_n\| \leq \varepsilon, \quad (2.3)$$

其中 $\|\cdot\|_0$ 表示矩阵的0范数，即非零元素的个数。上式的物理意义可以描述为：在满足误差不超过 ε 的前提下，寻找信道矢量 \mathbf{h}_n 在 \mathbf{A} 上的最优稀疏投影矢量 \mathbf{s}_n 。这类问题可以通过诸如OMP等方法近似求解^[80,81]，这里不过多叙述。一般来说，空域码本 \mathbf{A} 可以直接选为离散傅里叶变换（Discrete Fourier transform, DFT）矩阵或通过离线字典学习（Dictionary-learning）等方法得到^[79]。因此在 \mathbf{A} 固定且收发双方都已知时，用户只需要上报 $\mathbf{s}_n \in \mathbb{C}^{M_1 \times 1}$ 的信息至BS即可。由于矢量 \mathbf{s}_n 是稀疏的，其中大部分元素都为零，因此通过设计合适的压缩量化方法，可使得其反馈开销有效降低。这类方案称为空域（或角度域、波束域）的压缩信道反馈方法。

同时，当信道的相干带宽较大，相邻资源块间频率相关性较强时，还可以通过对信道矢量在频域维度上压缩，在进一步减少反馈开销的同时，不会造成反馈性能的严重下降，称为双域（Double-domain），或联合空频域压缩反馈方案。此时问题(2.3)推广为形如下式双域稀疏表示问题，即

$$\min \|\mathbf{S}\|_0, \text{ subject to } \|\mathbf{W} - \mathbf{A}\mathbf{S}\mathbf{B}^H\| \leq \varepsilon, \quad (2.4)$$

其中 $\mathbf{W} = [\mathbf{h}_1, \mathbf{h}_2, \dots, \mathbf{h}_{N_3}] \in \mathbb{C}^{N_t \times N_3}$ 定义为由 N_3 个连续RB上的信道矢量组成的宽带待反馈CSI。需要指出，用于控制聚合等级的 N_3 实际上是权衡量化误差和反馈开销的参量，需要根据信道的实际频率选择性衰落特性而灵活选取。比如当信道频域相关性较大时，可以增大 N_3 以降低宽带反馈开销；但当信道频率选择性衰落明显时，会降低 N_3 以保证量化反馈性能。在本章中，我们取典型值13。另外， $\mathbf{A} \in \mathbb{C}^{N_t \times M_1}$ 和 $\mathbf{B} \in \mathbb{C}^{N_3 \times M_2}$ 分别表示空、频域码本，其各自的维度分别是 M_1 和 M_2 。上式的物理意义与式(2.3)类似，即在给定量化的

差下求解 \mathbf{W} 在双域稀疏基上的最稀疏表示 $\mathbf{S} \in \mathbb{C}^{M_1 \times M_2}$ 。这类称为双域稀疏表示的问题相较问题(2.3)更复杂些，但也有一系列工作提出了解决方法^[80,82,83]。比如一种简单方法是基于克罗内克积（Kronecker-product）首先将其转化为一维稀疏表示问题，再利用OMP或 l_1 -norm最小化等方法来求解^[80,81]。当然也有文献针对直接针对二维问题(2.4)设计了低复杂的二维迭代自适应方法^[82]。本章中，我们不赘述问题(2.3)和(2.4)的具体求解过程，而重点讨论在给定稀疏解 \mathbf{s}_n 和 \mathbf{S} 时的量化反馈方法。同时注意到 \mathbf{s}_n 可视为 \mathbf{S} 在 $M_2 = 1$ 时的特殊场景，因此下文中，我们仅介绍 \mathbf{S} 的反馈方法。

具体来说，通过利用 \mathbf{S} 在空、频域码本上的二维稀疏特性，我们可以将其进一步分解为两部分信息 $\{\mathbf{b}^s, \mathbf{b}^f\}$ 和 \mathbf{Z} 分别量化反馈。具体来说， $\mathbf{b}^s = [b_1^s, \dots, b_{M_1}^s] \in \{0, 1\}^{M_1}$ 和 $\mathbf{b}^f = [b_1^f, \dots, b_{M_2}^f] \in \{0, 1\}^{M_2}$ 都是一个二进制支撑（Support）向量，它们分别表示 \mathbf{S} 中非零元素的横、纵坐标位置，或称空、频域稀疏基的索引。换言之，当向量 \mathbf{S} 中第 i 行（或第 j 列）存在非零元素时，有 $b_i^s = 1$ （或 $b_j^f = 1$ ）；否则 $b_i^s = 0$ （或 $b_j^f = 0$ ）。 $\mathbf{Z} \in \mathbb{C}^{\|\mathbf{b}^s\|_0 \times \|\mathbf{b}^f\|_0}$ 则是依照 \mathbf{b}^s 和 \mathbf{b}^f 中非零元素所指示的行、列信息，从 \mathbf{S} 中取出的一个包含其所有非零元素的最小子矩阵，称为信道系数矩阵。因此，在获得空频域二进制支撑矢量 $\{\mathbf{b}^s, \mathbf{b}^f\}$ 和反馈的信道系数矩阵 $\hat{\mathbf{Z}}$ 后，BS可以恢复出 $\hat{\mathbf{S}}$ ，进而得到CSI矩阵 $\hat{\mathbf{W}}$ 。需要指出的是，由于 \mathbf{b}^s 和 \mathbf{b}^f 都是一个二进制矢量，实际可以做到无损反馈。但 \mathbf{Z} 会由于量化精度的原因而存在误差，这也是造成反馈误差的主要原因。

在下一章节，我们将利用二进制支撑矢量的突发稀疏性和相邻时帧相关性，设计一种称为二级差分反馈的压缩方案，以较低的开销实现逐帧的信道跟踪反馈。同时，对于 \mathbf{Z} 的量化，也将提出一种基于多字典原子的矢量量化（Vector quantization, VQ）方案。结合在线-离线混合的动态字典更新算法，这种量化方法在保证精度的同时有效地降低了反馈开销。

2.3 TSAC反馈方案设计

本章节将首先介绍所提出的TSAC方案的系统实现框图，之后在各个子章节详细介绍各模块的功能和实现算法。

如图(2.2)所示，用户端主要由空频域基求解器、二进制支撑矢量压缩反馈模块和多字典原子矢量量化器构成。其中空频域基求解器首先负责求解问题(2.4)获得稀疏解 \mathbf{S} 。对于从中提取的两部分信息 $\{\mathbf{b}^s, \mathbf{b}^f\}$ 和 \mathbf{Z} ，二进制支撑矢量压缩反馈模块和多字典原子矢量量化器分别将其压缩量化成 F_b 和 F_z 比特的信息并反馈至BS。BS在接收到上报信息后，首先可以利用 F_b 无误恢复出支撑矢量 $\{\mathbf{b}^s, \mathbf{b}^f\}$ 。之后，基于量化字典恢复包含量化误差的信道系数

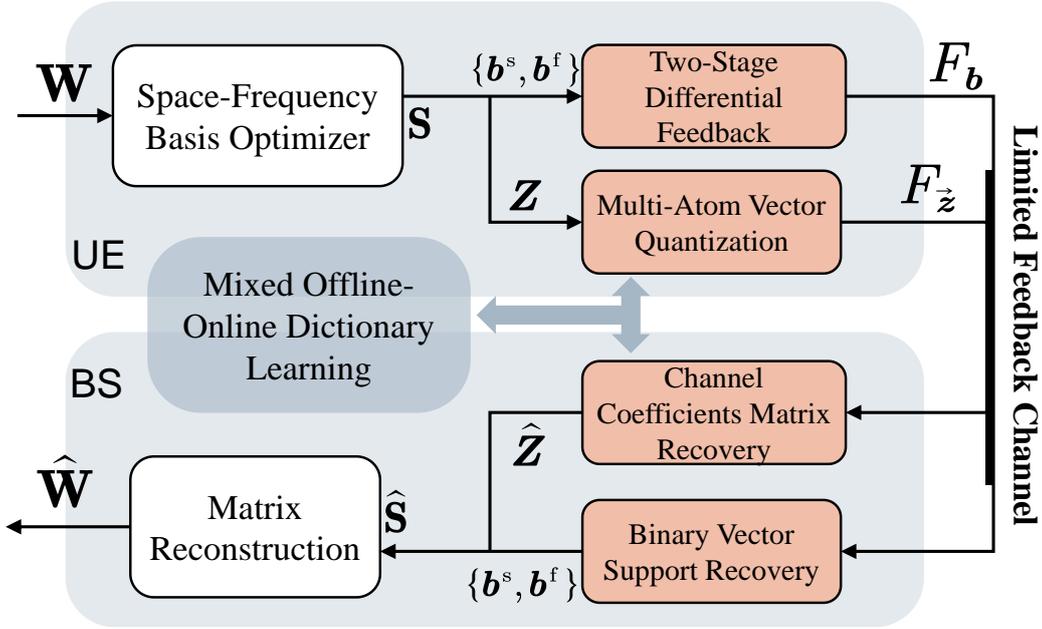


图 2.2 基于双域压缩的上行有限反馈系统框图

矩阵 $\hat{\mathbf{Z}}$ 。最后，通过查阅空、频码本 \mathbf{A} 和 \mathbf{B} 重构出信道矩阵 $\hat{\mathbf{W}}$ 。同时，为了保证信道系数矩阵的量化精度，混合离线-在线字典更新模块将同时部署于用户和 BS，负责维护并同步更新一个相同的量化字典。这种更新是只针对字典中单原子（Atom）的，因此几乎不会带来额外的计算量。

下面将细分三个子章节详细介绍除空频域稀疏基求解器之外的三个模块的具体工作过程和实现算法。为了书写简明，下文介绍算法时将不再区分 \mathbf{b}^s 和 \mathbf{b}^f ，而直接用 \mathbf{b} 来表示一个二进制稀疏支撑矢量。同时定义 $\mathbf{z} = \text{vec}(\mathbf{Z}) \in \mathbb{C}^{\|\mathbf{b}^s\|_0 \|\mathbf{b}^f\|_0 \times 1}$ 为系数矩阵 \mathbf{Z} 的列向量化形式，这不会丢失任何信息，只是方便下文叙述。

2.3.1 二进制支撑矢量的二级差分反馈方案

在本章中，我们假设 \mathbf{b} 中非零元素数目固定为 L 。这是一个合理的假设，因为在实际中，信道主径（Dominated-path）数通常变化的很慢，我们总可以取一个合适且足够大的 L （但 L 仍远小于码本维度）使得信道大部分能量都长时间集中在这些主径位置内。基于此，本小节将利用二进制支撑向量 \mathbf{b} 的突发稀疏性和相邻时帧相关性，设计一种低开销量化方案。

具体来说，有许多研究和实验证明，由于传播环境中散射体的有限和慢变特性，massive MIMO 信道通常会表现出突发稀疏性和很强的时间相关性^[84]。即二进制矢量 \mathbf{b} 的非零元素通常会集中在少数几个簇（Cluster）内，且不少非零元素位置在时相邻时帧内保持不变，

如图(2.3)所示。

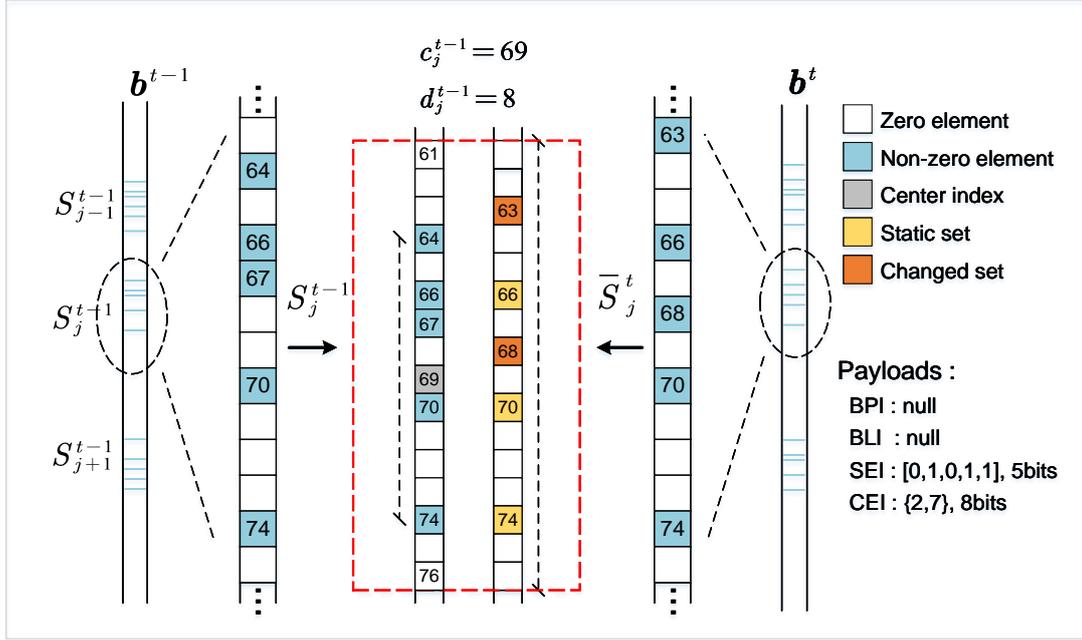


图 2.3 二进制支撑矢量的突发稀疏性和强时间相关性

基于这种结构化特性，下面将以第 t 帧二进制支撑矢量 \mathbf{b}^t 为例来详细介绍其反馈过程。具体来说，提出的二级差分反馈方案包含以下三个步骤：

(1) 步骤1: BS和用户对 \mathbf{b}^{t-1} 分段

为了利用图(2.3)所示的动态突发稀疏性，在每个时帧 t ，BS和用户会首先将上一时帧的二进制支撑矢量 \mathbf{b}^{t-1} 分割成 J^{t-1} 个片段 (Segment)，其中第 j 个片段中所有元素的索引集合记为 \mathcal{S}_j^{t-1} 。分段的准则可以描述为：在保证每个片段中非零元素所占比例接近1的同时使得总分段数 J^{t-1} 尽可能小。具体来说，该分段问题可建模为

$$\begin{aligned} \mathcal{P}_S : \quad & \min_{J^{t-1}, \{\mathcal{S}_1^{t-1}, \dots, \mathcal{S}_{J^{t-1}}^{t-1}\}} J^{t-1} \\ & \text{s.t. } \Omega^{t-1} \subseteq \cup_{j=1}^{J^{t-1}} \mathcal{S}_j^{t-1}, \\ & \frac{|\Omega_j^{t-1}|}{|\mathcal{S}_j^{t-1}|} \geq 1 - \epsilon, \forall j, \end{aligned} \quad (2.5)$$

其中 Ω^{t-1} 表示 \mathbf{b}^{t-1} 中所有非零元素的索引集合； Ω_j^{t-1} 则特别表示 \mathbf{b}^{t-1} 在分段 \mathcal{S}_j^{t-1} 中的非零元素索引集合； $\epsilon > 0$ 是一个控制非零元素占比的参量。上式中，第一个集合约束条件意味着划分的 J^{t-1} 个片段必须包含 \mathbf{b}^{t-1} 中所有的非零元素；第二个不等式约束说明每个分段 \mathcal{S}_j^{t-1} 的非零元素占比不得小于 $1 - \epsilon$ 。问题 \mathcal{P}_S 是一个困难的集合优化问题，基于穷搜的算法复杂度在实际中是难以接受的。因此，我们设计了一种迭代K-means的算法来寻找一个较好的分段实现，其计算过程总结于算

法1中。算法的基本思路是利用K-means聚类算法逐步增加分段数 J^{t-1} 直至第二个约束条件满足。需要强调的是，由于二进制支撑矢量是无损反馈的，BS和用户通过对 \mathbf{b}^{t-1} 执行算法1后可以获得一致的分段结果 $\mathcal{S}_j^{t-1}, \forall j$ 。这可作为一项重要的双端共享信息用以降低 \mathbf{b}^t 的反馈开销。

(2) 步骤2：用户端对 \mathbf{b}^t 的分段

基于前一时帧的共享信息 $\mathcal{S}_j^{t-1}, \forall j$ ，用户端会进一步将当前时帧的二进制支撑矢量 \mathbf{b}^t 分割成 \bar{J}^t 个片段，做法如下。首先，定义

$$\begin{aligned} c_j^{t-1} &= \lceil [\max(\Omega_j^{t-1}) + \min(\Omega_j^{t-1})] / 2 \rceil, \\ d_j^{t-1} &= 2^{\lceil \log_2 |\mathcal{S}_j^{t-1}| \rceil - 1}, \end{aligned} \quad (2.6)$$

分别表示 \mathbf{b}^{t-1} 第 j 个分段 \mathcal{S}_j^{t-1} 的中心（Center）位置和半径（Radius）长度，其中符号 $\lceil \cdot \rceil$ 表示上取整。注意这里将 d_j^{t-1} 设置成2的幂次是为了方便后续的量化和反馈。无需额外计算， \mathbf{b}^t 的第 j 个分段 $\bar{\mathcal{S}}_j^t$ 可以直接得到，即

$$\bar{\mathcal{S}}_j^t = \{i : c_j^{t-1} - d_j^{t-1} \leq i \leq c_j^{t-1} + d_j^{t-1} - 1\}, \quad (2.7)$$

其中 $j = 1, \dots, J^{t-1}$ 。在给定 $\bar{\mathcal{S}}_j^t, j = 1, \dots, J^{t-1}$ 时，我们移除 \mathbf{b}^t 矢量中在索引集合 $\cup_{j=1}^{J^{t-1}} \bar{\mathcal{S}}_j^t$ 中的所有非零元素，得到一个残余支撑矢量 $\bar{\mathbf{b}}^t$ 。之后对其应用算法1获得剩余的 $\bar{J}^t - J^{t-1}$ 个分段，即 $\bar{\mathcal{S}}_j^t, j = J^{t-1} + 1, \dots, \bar{J}^t$ ，再基于式(2.6)计算相应的中心位置信息 c_j^t 和半径 d_j^t 信息。至此，用户端得到 \mathbf{b}^t 的所有分段结果 $\bar{\mathcal{S}}_j^t, j = 1, \dots, \bar{J}^t$ 。我们需要再次强调， \mathbf{b}^t 的前 J^{t-1} 个分段不需要使用分段算法，而直接基于上一时帧的分段信息完成的。

(3) 步骤3：各分段的二级差分反馈

此步骤中，用户端将对每个分段 $\bar{\mathcal{S}}_j^t$ 执行二级差分压缩，并利用四个指示器（Indicators）承载压缩后的反馈信息。表(2.1)详细总结了这四个指示器的功能和量化开销。下面我们详细介绍二级差分反馈的具体实现过程。

在第一阶段中，指示器BPI和BLI被用于直接反馈 \mathbf{b}^t 中最后 $\bar{J}^t - J^{t-1}$ 个分段的 c_j^t 和 d_j^t ，其量化开销分别为每段 $\lceil \log_2 |\mathbf{b}^t| \rceil$ 和 $\lceil \log_2 \lceil \log_2 |\mathbf{b}^t| \rceil \rceil$ 比特。需要指出的是，这里BLI的开销之所以等于 $\lceil \log_2 \lceil \log_2 |\mathbf{b}^t| \rceil \rceil$ ，是因为 d_j^t 被设置为2的幂次，因而只能从 $\lceil \log_2 |\mathbf{b}^t| \rceil$ 个候选值中确定。在获取第一阶段的反馈信息后，BS可以无误重构出对 \mathbf{b}^t 的分段准则 $\bar{\mathcal{S}}_j^t, j = 1, \dots, \bar{J}^t$ 。

在第二阶段，各分段 $\bar{\mathcal{S}}_j^t$ 中的非零元素位置信息将会承载于SEI和CEI反馈至BS。具体地，用户首先把各分段 $\bar{\mathcal{S}}_j^t$ 中的非零元素划分为以下两个子集合，即不动集（Static-set） $\Omega_j^{t-1} \cap \Omega_j^t$ 和变化集（Changed-set） $\Omega_j^t - \Omega_j^{t-1} \cap \Omega_j^t$ 。之后，用户用一个长为 $|\Omega_j^{t-1}|$ 的

二进制矢量SEI记录不动集中的元素索引信息，换言之，SEI矢量中“1”即表示不动集中的元素。最后， $\bar{\mathcal{S}}_j^t$ 中变化集中的索引将直接通过CEI量化反馈，每个的开销为 $\lceil \log_2 |\bar{\mathcal{S}}_j^t| \rceil$ 比特。

基于上述的二级反馈方案，各用户可以用较低的开销将当前时帧的二进制支撑矢量 \mathbf{b}^t 无误的量化并反馈至BS。所提方案能极大降低开销的原因在于：1) 利用突发稀疏性将二进制支撑矢量分割成尺寸更小的子段来降低量化开销；2) 利用相邻时帧的强相关性提取共享信息（前 J^{t-1} 个分段信息、不动集信息）来降低反馈开销。此外，所提方案也可以根据信道变化的快慢而自适应地调整反馈开销。比如当信道变化速度快时，反馈开销会变大，因为此时分段结果 $\bar{J}^t - J^{t-1}$ 和变化集 $\Omega_j^t - \Omega_j^{t-1} \cap \Omega_j^t$ 都会发生较大变化。相反，当信道变化速度慢时，所需反馈开销也会相应减少。为了更清晰的描述上述反馈过程，图(2.3)展示一个分段的量化结果和所需开销。对于一般场景，上述方案的开销计算表达式则为

$$F_b = \Delta + \sum_{j=1}^{\bar{J}^t} \left[|\Omega_j^{t-1}| + |\Omega_j^t - \Omega_j^{t-1} \cap \Omega_j^t| \times \lceil \log_2 |\mathcal{S}_j^{t-1}| \rceil \right], \quad (2.8)$$

其中 $\Delta \triangleq (\bar{J}^t - J^{t-1}) \times (\lceil \log_2 |\mathbf{b}^t| \rceil + \lceil \log_2 \lceil \log_2 |\mathbf{b}^t| \rceil \rceil)$ 表示一阶段的反馈开销。

2.3.2 信道系数的多原子矢量量化

本章节介绍信道系数 \mathbf{z} 的量化反馈方法。具体来说，当给定一个包含 2^B 个原子（或称码矢）的合适的量化字典 $\mathbf{C} = [\mathbf{c}_1, \dots, \mathbf{c}_{2^B}] \in \mathbb{C}^{L \times 2^B}$ 后，我们提出了一种多原子的矢量量化方法，其分为以下两个步骤。

在第一步，用户首先会在字典 \mathbf{C} 中寻找关于 \mathbf{z} 的一个多原子表示，即 $\mathbf{z} = \sum_{i \in \mathcal{I}^*} p_i \mathbf{c}_i + \mathbf{e}$ 。其中 \mathbf{e} 记为表示误差， \mathcal{I}^* 表示字典各原子的索引，它可以通过求解形如下式的问题来获得，即

$$\mathcal{P}_Z : \min_{\mathcal{I} \subseteq \{1, \dots, 2^B\}} \left\| \mathbf{z} - \sum_{i \in \mathcal{I}} p_i \mathbf{c}_i \right\|^2, \text{ s.t. } |\mathcal{I}| = l, \quad (2.9)$$

其中 l 是用于表示 \mathbf{z} 的字典原子数目。问题 \mathcal{P}_Z 同样是一个集合优化问题，较难直接求解。因此，我们将其转化为如下的等价问题：

$$\mathcal{P}'_Z : \min_{\tilde{\mathbf{p}}} \|\mathbf{z} - \mathbf{C}\tilde{\mathbf{p}}\|^2, \text{ s.t. } \|\tilde{\mathbf{p}}\|_0 = l, \quad (2.10)$$

其中 $\tilde{\mathbf{p}} \in \mathbb{C}^{2^B \times 1}$ 记为信道系数 \mathbf{z} 的稀疏表示。 \mathcal{P}_Z 是一个标准的压缩感知恢复问题，有许多经典算法，如OMP^[81]等，可以低复杂度地有效求解，这里不再赘述。

算法 1 Iterative K-means算法

输入: 二进制支撑矢量 \mathbf{b}^t ; \mathbf{b}^{t-1} 中各分段 \mathcal{S}_j^{t-1} 的中心坐标 c_j^{t-1} ; 比例参数 ϵ

初始化: 分段数目 $N = 1$; 非零索引集合 $\Omega_j^t = \emptyset$; $c_j^0 = \lceil |\mathbf{b}^0| (2j - 1) / 2N \rceil, \forall j$

for $k = 1, 2, \dots$ **do**

 记 $c_j^t = c_j^{t-1}, j = 1, \dots, N$ 作为每个分段 \mathcal{S}_j^t 的中心。

repeat

 对于 \mathbf{b}^t 中每个非零元素, 根据其至 $c_j^t, \forall j$ 的最小距离划分至对应的索引集合 Ω_j^t 中。

 更新 $c_j^t = \lceil [\max(\Omega_j^t) + \min(\Omega_j^t)] / 2 \rceil$ 作为每个分段的新中心

until Ω_j^t 不发生改变

 取 $\mathcal{S}_j^t = \{i : \min(\Omega_j^t) \leq i \leq \max(\Omega_j^t)\}$, 并计算 $\sigma = \min_{j=1, \dots, N} |\Omega_j^t| / |\mathcal{S}_j^t|$ 。

if $\sigma + \epsilon \geq 1$ **then**

 停止并跳出**for**循环

else

$N = N + 1$

end if

end for

输出: $\mathcal{S}_j^t, \forall j$

在第二步, 用户会将原子集合 \mathcal{I}^* 与量化的系数 \mathbf{p} 反馈至BS。其中, \mathcal{I}^* 需要 $\lceil \log_2(C_{2B}^l) \rceil$ 比特的压缩开销; \mathbf{p} 中每个元素的幅值和相位分别用 m 、 n 比特量化。因此, \mathbf{z} 的量化开销为

$$F_z = \lceil \log_2(C_{2B}^l) \rceil + l \times (m + n), \quad (2.11)$$

其中字典维度 B 的典型取值一般为8到10。当然在实际中, 需要考虑信道条件并基于主径数 L 合理选取。另外需要说明的是, 当用户只需要反馈 \mathbf{z} 的矢量方向信息 $\mathbf{z}/\|\mathbf{z}\|$ 时(比如当反馈CSI是波束赋形矩阵时), F_z 可以进一步被压缩为

$$F_z = \lceil \log_2(C_{2B}^l) \rceil + \lceil \log_2 l \rceil + (l - 1) \times (m + n). \quad (2.12)$$

2.3.3 在线-离线混合的量化字典更新方法

在上一小节, 我们介绍了对 \mathbf{z} 的多原子矢量量化方法。需要指出的, 这种方法虽然相较于传统的标量量化方案显著降低了开销, 但是其量化性能与字典 \mathbf{C} 密切相关。不难想象, 若 \mathbf{C} 选取的不合理, 则会导致较大的表示误差 \mathbf{e} 。因此, 为了提升问题 \mathcal{P}_z^l 的求解精度, 我们

表 2.1 指示器的详细注释

名称	描述	每段的反馈开销 (比特)
BPI	突发位置指示 (Burst-position indicator) 被用于反馈最后 $\bar{J}^t - J^{t-1}$ 个分段的中心位置信息	$\lceil \log_2 \mathbf{b}^t \rceil$
BLI	突发长度指示 (Burst-length indicator) 用来量化每个分段的长度	$\lceil \log_2 \lceil \log_2 \mathbf{b}^t \rceil \rceil$
SEI	不动集合指示 (Static-set indicator) 是一个二进制矢量, 用于表示不动集中非零元素的索引	$ \Omega_j^{t-1} $
CEI	变化集合指示 (Changed-set indicator) 用来直接承载变化集中所有非零元素位置信息	$ \Omega_j^t - \Omega_j^{t-1} \cap \Omega_j^t \times \lceil \log_2 \mathcal{S}_j^{t-1} \rceil$

设计了一种离线-在线混合的字典学习算法并同时部署于BS和客户端。字典的离线训练和在线更新实际上是求解形如下式的问题, 即

$$\mathcal{P}_D : \min_{\mathcal{C}, \tilde{\mathbf{P}}} \|\mathcal{Z} - \mathcal{C}\tilde{\mathbf{P}}\|_F^2, \text{ s.t. } \|\tilde{\mathbf{p}}_i\|_0 \leq l, \forall i, \quad (2.13)$$

其中下标 F 表示Frobenius范数; $\mathcal{Z} = [\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_K]$ 是 K 个历史的信道系数训练数据; $\tilde{\mathbf{P}} = [\tilde{\mathbf{p}}_1, \tilde{\mathbf{p}}_2, \dots, \tilde{\mathbf{p}}_K]$ 是对应的稀疏表示训练数据。为了降低训练复杂度, 我们将 \mathcal{P}_D 转化为下面的聚类问题 \mathcal{P}'_D :

$$\begin{aligned} \mathcal{P}'_D : \min_{\mathcal{C}, \{\mathcal{Z}^i, i=1, \dots, 2^B\}} \sum_{i=1}^{2^B} \sum_{\mathbf{z}_j \in \mathcal{Z}^i} \|\mathbf{c}_i - \mathbf{z}_j\|^2 \\ \text{s.t. } \cup_{i=1}^{2^B} \mathcal{Z}^i = \mathcal{Z}, \\ \mathcal{Z}^p \cap \mathcal{Z}^q = \emptyset, \forall p, q \in \{1, \dots, 2^B\}, \end{aligned} \quad (2.14)$$

其物理意义是基于距离最近准则将训练数据集划分为以 \mathbf{c}_i 为中心的 2^B 个簇。这类问题也有许多经典的聚类算法可解, 如林德-布佐-格雷 (Linde-Buzo-Gray, LBG) 算法^[85]。因此在实际实施中, BS需要在系统运行前通过求解 \mathcal{P}'_D 获得一个较好的初始量化字典并广播至各接入用户。在系统运行中, 当用户和BS获得新的量化信道系数 $\hat{\mathbf{z}}^t$ 后, BS和用户的字典需要同步进行字典原子的更新以追踪信道特性的变化, 即

$$\mathbf{c}_i^{t+1} = (\tau_i^t \mathbf{c}_i^t + (1 - \tau_i^t) \hat{\mathbf{z}}^t) / \|\tau_i^t \mathbf{c}_i^t + (1 - \tau_i^t) \hat{\mathbf{z}}^t\|, \quad (2.15)$$

其中 \mathbf{c}_i 表示距离 $\hat{\mathbf{z}}^t$ 最近的字典原子; τ_i^t 是可控的步长参数。上式更新只涉及单个字典原子的更新, 在线计算量是极小的。同时需要指出的是, 由于BS和客户端使用相同的在线采样数据 $\hat{\mathbf{z}}^t$, 因此二者的量化字典会始终保持一致。综上, 基于这种离线-在线混合的同步更新方法, 量化字典会自适应地跟踪信道特性变化, 从而保证了信道系数的表示精度。

2.4 仿真结果与分析

本节中，我们将搭建符合LTE规范的链路级仿真系统来证明所提反馈方案相较现有的NR Type-2高精度反馈方案的显著优越性^[86,87]。信道模型采用现有3GPP协议规定的簇延时线类型B信道（Clustered delay line type-B, CDL-B）^[88]，这是一种非直视径（Non line-of-sight, NLOS）信道模型，其具体参数总结于表(2.2)中。同时，为了使得对比更详尽，我们还选择文献^[6]中提出的一种称为基于自适应离开角的子空间（AoD-adaptive subspace, AoD-SS）码本CSI反馈方案作为另一种基线。但需要指出，这种AoD-SS反馈方案在提出时并没有考虑联合空频域的CSI压缩反馈，同时也无法像所提出的TSAC反馈方案和NR-Type2基线一样可以逐帧跟踪二进制支撑矢量的变化。若要实现这两个功能，AoD-SS方案需要额外付出大量的量化开销。因此在本节中，我们只将AoD-SS作为传输性能仿真的一种参考基线，而不计算其所需的量化开销。

表 2.2 系统和信道参数设置

名称	设置
信道模型	CDL-B
用户数	4
用户移动速度	30km/h
发送天线数	32
接收天线数	4
天线阵列	ULA
天线是否极化	发送正交极化
载波频率	3GHz
最大时延扩展（Delay spread, DS）	1us
RB数目	52
子帧（Subframe）数目	500
子载波间隔	30kHz
MCS	8
各用户传输流数	最大为2

另外，在给定空、频域码本维度分别为 $M_1 = 64$ 和 $M_2 = 13$ 时，为了更详尽地比较分析，我们还通过改变信道系数矢量 \mathbf{z} 的量化精度，设置了两种不同精度的反馈模式，即：1) $\{l, m, n, L_s, L_f, B\} = \{6, 6, 6, 20, 10, 10\}$ 时的高精度信道反馈模式（High resolution feedback mode, HRFM）；2) $\{l, m, n, L_s, L_f, B\} = \{3, 4, 4, 10, 7, 8\}$ 时的低精度信道反馈模式（Low resolution feedback mode, LRFM）。其中 $\{l, m, n\}$ 在式(2.11)中介绍过，分别表示系数量化时所

用原子数和幅值、相位比特数； L_s 和 L_f 分别表示空、频域二进制支撑矢量的最大非零元素数； B 控制量化字典维度。下面各子章节将详细介绍计算和仿真结果，这里我们假设各用户反馈的是由各RB上双流SVD预编码矩阵组成的宽带CSI。

2.4.1 不同方案的反馈开销比较

表(2.3)计算并统计了所提方案和基线方案的反馈开销。表中数据是对各用户、各流及各个宽带信道样本平均后的结果。可以看出，所提方案的反馈开销显著少于NR Type-2基线。在高精度反馈模式下，二进制稀疏支撑矢量的反馈开销降低了约30%，信道系数矩阵的量化开销降低了40%多。低精度反馈模式下，所提方案的总开销仅为基线的25%。所提方案相较基线取得如此显著的反馈开销优势的原因主要在于：1) 我们所提的二级差分反馈方法充分地利用了二进制支撑矢量的动态突发稀疏性和强相关性来降低压缩开销，NR Type-2则是直接量化非零元素位置信息；2) 相较于基线方案直接对信道系数矩阵进行标量量化，所提的多原子矢量量化方法可以有效地对系数信息进行压缩。

表 2.3 不同方案反馈开销比较

项目	NR Type-2	TSAC-LRFM	TSAC-HRFM
空域二进制支撑矢量	18 bits	11.6 bits	15.7 bits
频域二进制支撑矢量	10 bits	3.9 bits	4.3 bits
信道系数矩阵	198 bits	40 bits	114 bits
总开销	226 bits	55.5 bits	134 bits

2.4.2 不同方案的BLER性能比较

图(2.4)和图(2.5)比较了不同方案的的用户平均BLER性能。我们分别考虑了四用户单流传输和两用户双流传输两种场景。可以看出，所提出的TSAC方案即使在低精度反馈模式下都可以取得相较NR Type-2和AoD-SS基线方案更优的性能。高精度反馈模式下则差距更大。同时可以注意，NR Type-2和AoD-SS基线方案的BLER性能并没有随着发送信噪比的增加一直下降，在我们仿真的区间内，二者甚至无法达到0.01的BLER性能，其原因是较差的量化反馈精度所导致的流间干扰过于强烈。这也因此证明了我们方案优秀的反馈精度性能。

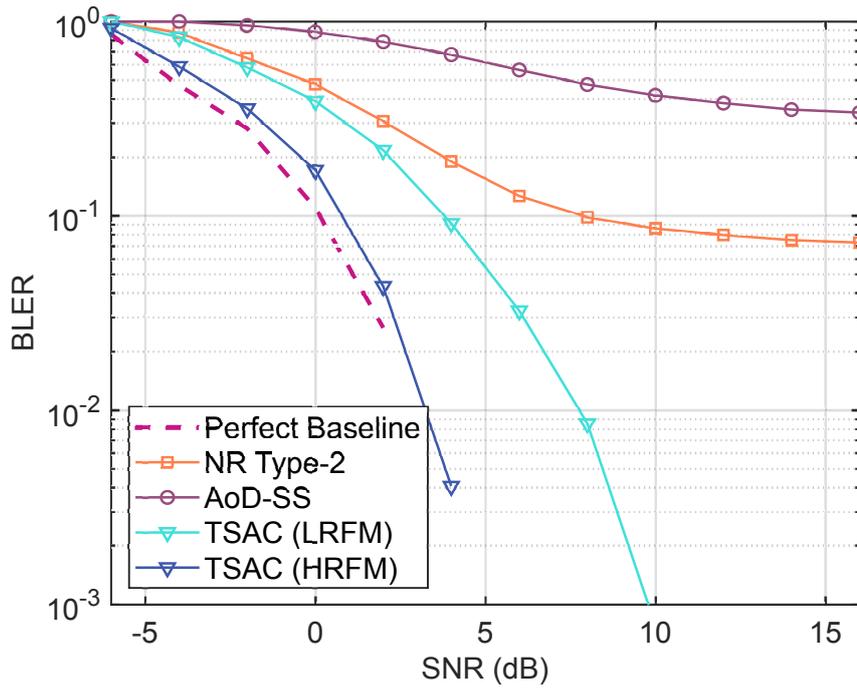


图 2.4 四用户单流场景下的BLER性能

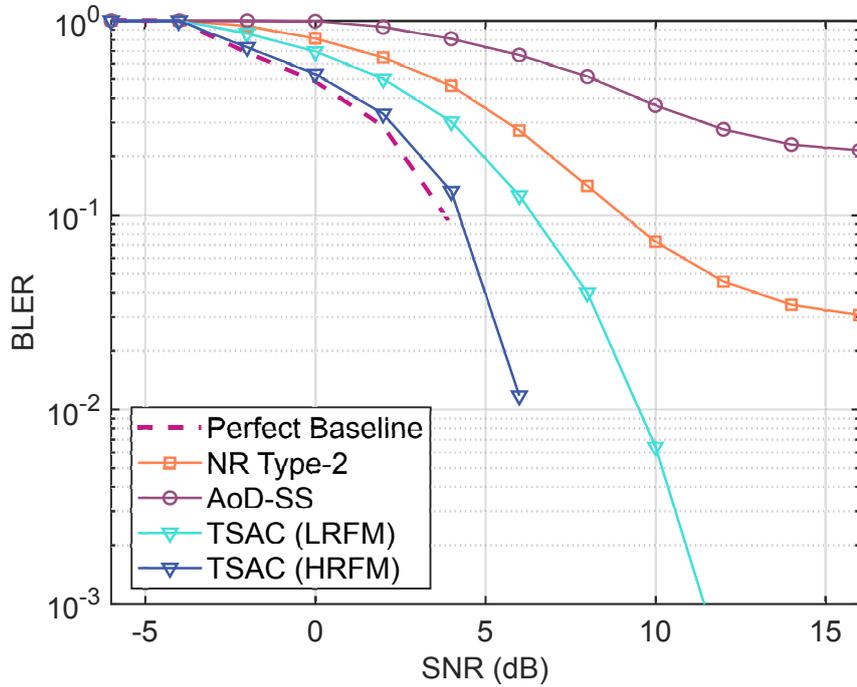


图 2.5 两用户双流场景下的BLER性能

2.4.3 不同方案的吞吐量性能比较

最后，我们比较各方案的吞吐量（Goodput）性能，如图(2.6)。这里吞吐量定义为 $R \times$

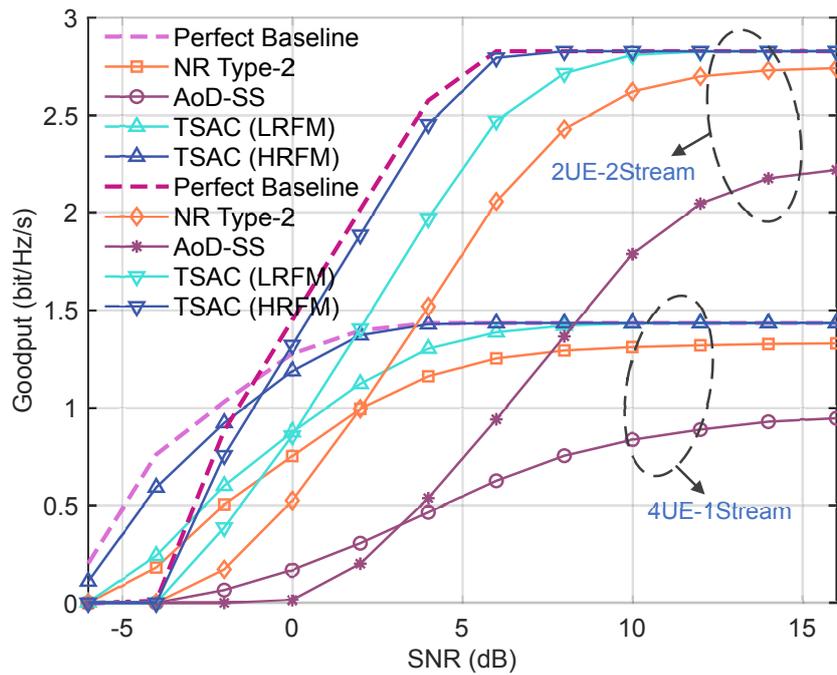


图 2.6 各方案在不同传输场景下的吞吐量性能

$(1 - \text{mean}(BLER))$), 其中 R 是各用户的物理传输速率 (由MCS确定)。图中各方案的性能差异趋势与BLER性能基本一致。这证明, 所提方案TSAC相较NR Type-2等基线也具有显著传输性能优势。这种优秀的量化性能得益于我们离线-在线混合的量化字典更新方法。

2.5 本章小结

在本章中, 我们针对FDD多用户的massive MIMO宽带CSI反馈场景提出了一种联合空频域的TSAC信道反馈方案。通过利用massive MIMO信道的动态突发稀疏性、强时间相关性和频域相关性, 我们首先将宽带CSI投影至空、频域码本, 提取出需要反馈的二进制稀疏支撑矢量和信道系数矢量信息。之后, 我们精心设计了一种二级差分的方案实现了二进制稀疏矢量的低开销无损反馈, 且反馈比特数可以根据信道变化速率自适应地调整。另外, 为了在保证量化精度的同时以较低的开销量化信道系数矢量, 我们提出了一种基于动态更新字典的多原子矢量量化方案, 相较传统的标量量化方案显著降低了反馈开销。最后, 我们基于协议标准搭建了完善的链路级仿真系统。详尽的计算和传输性能仿真结果证明了所提方案相较NR Type-2基线在反馈开销和量化性能上的巨大优势。

3 中重载场景下支持空域并发的多用户频分调度方案设计

3.1 研究动机和主要成果

用户调度是目前5G和B5G系统中无线资源管理的一项重要功能。受限于实际无线链路系统中发送天线数目和空域自由度，BS无法在同一时、频域传输资源上为所有用户提供通信服务，因此可按照一定准则选取部分用户进行通信，以获得更多的用户分集增益，提高系统整体的传输性能。此外，为了满足某些对特定用户业务的QoS要求，如吞吐量、用户公平性、延迟或可靠性等，BS需要解决的一个更关键的问题是如何将可用系统时间和频率资源调度给合适的终端用户。这类问题在4G LTE系统中一般被称为FDPS^[34,35]。在5G NR系统中，由于系统结构和时、频最小调度资源的精细化，这类问题又称为FDRA^[42]。在这些频域资源分配或用户调度问题中，根据实际链路约束的不同可大致分为如下几类研究方向。

一方面，根据每个RB能否被多个用户复用，调度问题可分为SU-MIMO和MU-MIMO两类。对SU-MIMO调度，系统约束每个RB只能分配给中至多一个用户使用，称为RB的独占性（Exclusivity）。此场景下只利用了单个用户的分集或复用增益，调度问题较为简单。而MU-MIMO调度则允许在同一RB的不同空间流上复用更多用户。这虽然提供了更大的空域灵活性，但也导致问题求解变得更为困难。近年来有一系列研究者考虑了此场景下的调度问题。比如文献^[33]证明了在考虑多数据流分配时的优化是一个NP-hard问题，并针对全反馈和部分反馈两种场景各提出了一种在PF准则下基于贪婪迭代的近似算法。在此基础上，文献^[34]针对SU-MIMO提出了一种广义的度量函数，可以同时刻画包括PF在内的其它权重度量准则。在假设RB只能被多用户复用单流的前提下，文献^[35]考虑了MU-MIMO场景下近似调度算法。文献^[36]考虑了包含MCS在内的LTE实际约束，并同时研究了在积压流量模型和有限队列模型下的调度问题。文献^[37]提出了一种通用近似方法解决包含MCS优化的调度问题。但它们只涉及SU-MIMO场景。

而另一方面，根据用户频域资源分配的连续性（Contiguity），调度问题又可分为连续或非连续调度。以5G NR FDRA为例，系统规定了Type-0和Type-1两种资源分配模式^[31,32]。

在Type-0模式下，系统可为每个用户分配非连续的频谱资源，此时最小的频域资源调度粒度是一个RBG。对Type-1模式，系统可在更精细的RB粒度上设计调度算法。但考虑到下行信令开销，系统限制为每个用户分配的RB资源必须连续。相较而言，由于连续性约束的存在，Type-1类型的调度问题更为复杂，这涉及到一个复杂的离散空间组合优化问题，因此大部分研究者通过设计各种启发式的近似算法来降低复杂度。比如文献^[40]提出了一种次优算法来解决在连续RB分配约束下的和率最大化问题，并计算了所提方案与最优解的性能差距。文献^[41]针对加权和速率问题，提出一种基于邻位扩展的启发式调度方法。文献^[42]也是类似的思路，设计从可用带宽资源的两端逐步向中心扩展的资源分配算法以满足连续性约束。

然而，在现有的调度算法中，很少有研究者同时考虑RB的多用户复用和存在频域连续约束下的调度方案设计。少部分研究也都采取基于可行分配图案（Feasible-allocation-pattern）的搜索解法，只适用于用户数少、带宽窄等问题规模较小的场景^[35,89]，无法满足中重载场景下的调度需求。而解决这类问题对进一步发掘5G/B5G系统的调度增益有着重要的意义。因此，本章节我们将着手于此，研究单小区中重载场景下支持空域并发的多用户频分调度问题。然而，求解上述问题并在实际系统中可靠实施还将面临以下三个关键技术难点：1) 考虑用户并发后，如何处理复杂的优化目标和实际的链路约束；2) 中重载场景下如何解决求解空间过大导致的维数灾难（Curse-of-dimensionality）；3) 如何控制算法的复杂度使其满足最小调度时间粒度，即TTI级的调度需求。

为了解决上述困难，我们将提出一种联合空频域的逐级消息传播（JSFD successive message passing, JSFD-SMP）用户调度算法和称为关键点采样（Key-point sampling, KPS）的实施简化技术。需要指出，虽然本章只考虑各用户单根天线，即多输入单输出（Multiple-input single-output, MISO）场景的调度算法设计，但所提算法可以容易地拓展至MIMO多流传输场景中。

3.2 系统模型与问题建模

3.2.1 系统模型

本章节中，我们将考虑一个单小区的多用户下行传输系统，如图(3.1)所示。部署 N_t 根天线的中心BS将为 K 个单天线用户提供传输服务，记用户集合为 \mathcal{K} 。系统可用的BWP由 N 个频域RB构成，其索引集合记为 \mathcal{N} 。其中各用户的流量业务类型可以是异构的（Heterogenous），具有不同的QoS权重。基于不同的调度准则，BS在每个TTI内从用户集 \mathcal{K} 中挑选出服务用户并为其分配若干RB资源用于传输业务数据。

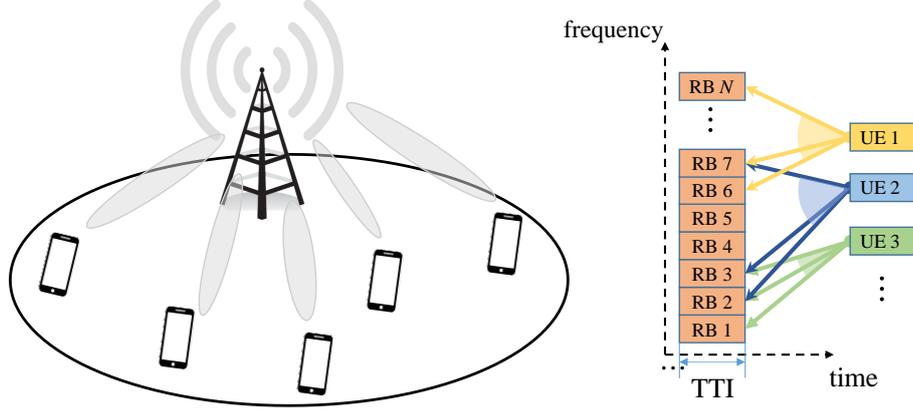


图 3.1 单小区多用户下行系统模型

为了方便说明，我们假设各用户的下行信道具有块衰落（Block-fading）特性，且在每个TTI内保持不变。更具体地，对特定的用户 k 来说，假设其在第 n 个RB上的信道 $\mathbf{h}_{k,n}^H \in \mathbb{C}^{1 \times N_t}$ 符合广泛使用的Saleh-Valenzuela多径信道模型^[90]，即

$$\mathbf{h}_{k,n}^T = \sqrt{\frac{N_t}{M_k}} \sum_{m=1}^{M_k} \beta_{m,k} e^{-j2\pi\tau_{m,k}f_n} \mathbf{a}_k(\varphi_{m,n,k}), \quad (3.1)$$

其中 M_k 是用户 k 信道中可分辨的径数； $\beta_{m,k}$ 和 $\tau_{m,k}$ 分别为第 m 根径的复增益和时延； $\varphi_{m,n,k}$ 定义为

$$\varphi_{m,n,k} = \frac{f_n}{c} d \sin \theta_m, \quad (3.2)$$

表示第 n 个RB上的空间指向。上式中， f_n 表示第 n 个RB的载波频率； d 定义为天线间隔，实际中一般为载波波长的一半，即 $d = c/2f_c$ （ c 表示光速）。本章中，我们假设系统BWP的数值远小于载波频率，这意味 f_n 与中心频率 f_c 近似相同，因此有 $\varphi_{m,n,k} \approx \frac{1}{2} \sin \vartheta_{m,n,k}$ 。最后， $\mathbf{a}_k(\varphi_{m,n,k})$ 是与 $\varphi_{m,n,k}$ 对应的阵列响应矢量。若假设天线阵列是ULA，则有 $\mathbf{a}_k(\varphi_{m,n,k}) = \frac{1}{\sqrt{N_t}} e^{-j2\pi\varphi_{m,n,k} \mathbf{p}_{a_k}}$ ，其中 $\mathbf{p}_{a_k} = [-\frac{N_t-1}{2}, -\frac{N_t+1}{2}, \dots, \frac{N_t-1}{2}]$ ^[91]。

在第 t 个TTI内，用户 k 的传输速率可以表示为

$$R_k(t) = \sum_{n \in \mathcal{B}_k(t)} r_{k,n}(t), \forall k \in \mathcal{K}, \quad (3.3)$$

其中 $\mathcal{B}_k(t)$ 表示此时系统为用户 k 分配的用于传输的RB索引集合（ $\mathcal{B}_k(t) = \emptyset$ 表示用户 k 在当前TTI内未被服务）； $r_{k,n}(t)$ 则表示用户 k 在第 n 个RB上的传输速率。在本章中，我们认为系统支持多个用户在同一RB的并发传输，此时 $r_{k,n}(t)$ 的数学表达式为（为了简洁，这里暂时省略了TTI索引）

$$r_{k,n} = \log_2 \left(1 + \frac{p_{k,n} |\mathbf{h}_{k,n}^H \mathbf{v}_{k,n}|^2}{\sum_{j \in \mathcal{S}_n \setminus k} p_{j,n} |\mathbf{h}_{k,n}^H \mathbf{v}_{j,n}|^2 + \sigma_{k,n}^2} \right), \quad (3.4)$$

这里 \mathcal{S}_n 表示第 n 个RB上并发复用的用户索引集合； $p_{k,n}$ 、 $\mathbf{v}_{k,n} \in \mathbb{C}^{N_t \times 1}$ 和 $\sigma_{k,n}^2$ 分别表示各自的发送功率、预编码矢量和噪声能量。可以看出，这种考虑空域并发的多用户频分调度方案固然可以帮助获取更高的调度增益，但也使用户调度的求解变得更为复杂，因为此时 $r_{k,n}$ 不仅与用户 k 本身的信道状态有关，还受该RB上复用的其余用户的干扰影响。

3.2.2 联合空频域的多用户调度问题

在本章中，我们希望设计一个TTI级的联合空频域用户调度算法，在给定各用户QoS权重和发送功率下实现优秀的小区吞吐量和用户感知速率性能。如图(3.2)所示，为了符合实际的链路约束，本章考虑的多用户调度需要满足3GPP 5G和B5G中的Type-1 FDRA规范，即调度给用户 k 的RB资源 \mathcal{B}_k 必须满足连续分配准则^[31,32]。同时考虑到实际系统中信道编码、调制器的实现开销，各RB上并发的最大用户数限制为 L 。上述约束条件使得本章所考虑的系统更符合实际，但也使得调度和资源分配的设计更加困难。

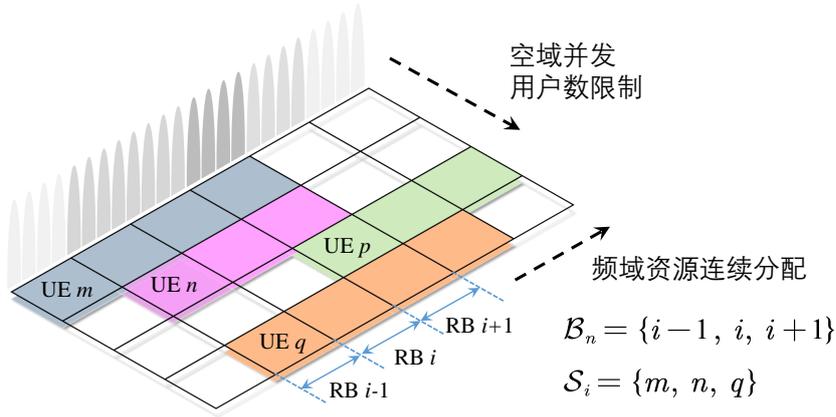


图 3.2 空分复用的多用户频分调度

具体来说，在每个TTI内，BS的MAC层调度器负责求解一个以加权和速率为目标，同时存在RB连续性分配约束和最大空域并发用户数限制下的多用户调度问题，其数学表达式可以表示为（注意这里都省略了表示TTI的索引）

$$\begin{aligned} \mathcal{P}_S : \max_{\mathcal{B}^* \text{ or } \mathcal{S}^*} \sum_{k \in \mathcal{K}} \sum_{n \in \mathcal{B}_k} w_k r_{k,n}(\mathcal{S}_n) \\ \text{s.t. } |\mathcal{B}_k| = \begin{cases} \max(\mathcal{B}_k) - \min(\mathcal{B}_k) + 1, & \text{if } \mathcal{B}_k \neq \emptyset \\ 0, & \text{if } \mathcal{B}_k = \emptyset \end{cases}, \forall k \in \mathcal{K}, \\ |\mathcal{S}_n| \leq L, \forall n \in \mathcal{N}, \end{aligned} \quad (3.5)$$

其中 \mathcal{B}_k 和 \mathcal{S}_n 已于前文介绍过，分别表示为用户 k 分配的RB索引集合和第 n 个RB上并发复用的用户索引集合。 $\mathcal{B}^* = \{\mathcal{B}_k, \forall k \in \mathcal{K}\}$ 和 $\mathcal{S}^* = \{\mathcal{S}_n, \forall n \in \mathcal{N}\}$ 的意义是等价的，即都表示一

个联合空频域的用户调度结果。式(3.5)的第一个约束条件表示各用户的RB分配必须满足连续性条件；第二个约束则限制了各RB上并发复用的用户数最大不超过 L 。同时，上式中的 w_k 表示用户 k 的QoS权重，在采用经典的M-LWDF的调度准则时^[24]，其具体数学表达式为

$$w_k = -(\log \delta_k / \tau_k) d_k / \bar{R}_k, \forall k, \quad (3.6)$$

其中 δ_k 、 τ_k 、 d_k 和 \bar{R}_k 分别表示用户 k 的可接受丢包率、可容许的最大时延、队首（Head-of-line, HOL）时延和历史平均吞吐量。

另外需要强调的是，对于 \mathcal{P}_S 中形如加权和速率最大化的目标问题，实际上只有通过将MAC层用户调度和PHY预编码一起联合优化才能找到全局最优解。但这种跨层联合优化问题是非常困难的。如何设计满足链路约束、复杂度低且能逼近全局最优的跨层算法至今都是悬而未决的一项难题。目前4G LTE及5G NR系统中还是采用MAC层和PHY功能分离的主流实现方案。所以在本章中，我们也遵循于此，只关注MAC层的用户调度 \mathcal{S}^* 方案的设计，而不考虑式(3.4)中关于预编码和功率分配权值的优化。

具体来说，我们假设调度器采用基于正则化迫零（Regularized zero-forcing, RZF）预编码和等功率权值分配的方案来估计 $r_{k,n}$ ^[92]。对第 n 个RB，给定用户调度结果 \mathcal{S}_n 后，记 $p_{k,n} = P / (N |\mathcal{S}_n|)$ 为用户 k 的分配功率（ P 是总发送功率）。同时定义 $\mathbf{V}_n \in \mathbb{C}^{N_t \times |\mathcal{S}_n|}$ 作为BS的归一化发送预编码矩阵，其数学表达式为

$$\mathbf{V}_n = [\mathbf{v}_{1,n}, \dots, \mathbf{v}_{|\mathcal{S}_n|,n}] = \tilde{\mathbf{H}}_n^H \left(\tilde{\mathbf{H}}_n \tilde{\mathbf{H}}_n^H + \alpha_n \mathbf{I} \right)^{-1} \boldsymbol{\Lambda}_n^{1/2}, \quad (3.7)$$

其中 α_n 是一个正则化因子； $\tilde{\mathbf{H}}_n = [\mathbf{h}_{1,n}, \dots, \mathbf{h}_{|\mathcal{S}_n|,n}]^H \in \mathbb{C}^{|\mathcal{S}_n| \times N_t}$ 表示该RB上所选调用户的信道矩阵； $\boldsymbol{\Lambda}_n^{1/2} = \text{diag} \left(\|\bar{\mathbf{v}}_{1,n}\|^{-1}, \dots, \|\bar{\mathbf{v}}_{|\mathcal{S}_n|,n}\|^{-1} \right)$ 是一个能量归一化对角矩阵， $\bar{\mathbf{v}}_{k,n}$ 表示矩阵 $\bar{\mathbf{V}}_n \triangleq \tilde{\mathbf{H}}_n^H \left(\tilde{\mathbf{H}}_n \tilde{\mathbf{H}}_n^H + \alpha_n \mathbf{I} \right)^{-1}$ 的第 k 列。

在给定上述预编码和功率分配方法下， $r_{k,n}$ 是只与 \mathcal{S}_n 有关的一个变量。但即使如此，问题 \mathcal{P}_S 也是一个比较困难的高维离散空间内的非凸组合优化问题。如采取类似可行分配图案的遍历解法，搜索空间的维度达到 $O(C_{N_2}^K)$ ，这是无法接受的^[93]。但如果采取启发式的低复杂的贪婪算法，则无法在更深层次上挖掘潜在的用户调度增益^[94]。因此，我们将在3.3章节提出一种JSFD-SMP算法来求解问题 \mathcal{P}_S 。所提出的算法的优势在于：1) 可在兼顾用户体验速率公平性的同时相较基线方案取得显著的吞吐量性能增益；2) 由于其结构的特殊性，算法非常适合采用并行计算等技术在硬件架构层面来降低处理时延，从而满足TTI级的实时调度需求。

3.3 JSFD-SMP用户调度算法

3.3.1 算法流程描述

本小节，我们将设计一种联合空频域的逐级消息传播算法JSFD-SMP来近似求解问题 \mathcal{P}_S 。为了避免任何可能的歧义，我们将忽略表示TTI的索引，本小节中出现的索引 t 均表示算法迭代的轮次，详述如下。

具体来说，由于原问题的高度复杂性和非凸性，我们首先将 \mathcal{P}_S 分解为 L 个子问题迭代求解。在第 t 轮迭代时，我们构建并求解形如下式的子问题 \mathcal{P}_S^t

$$\begin{aligned} \mathcal{P}_S^t : \quad & \max_{\{\mathcal{B}_0^t, \dots, \mathcal{B}_{|\mathcal{K}^t|-1}^t\}} \sum_{k \in \mathcal{K}^t} \sum_{n \in \mathcal{B}_k^t} A_{k,n}^t(\mathcal{S}_n^t) \\ \text{s.t.} \quad & |\mathcal{B}_k^t| = \begin{cases} \max(\mathcal{B}_k^t) - \min(\mathcal{B}_k^t) + 1, & \text{if } \mathcal{B}_k^t \neq \emptyset \\ 0, & \text{if } \mathcal{B}_k^t = \emptyset \end{cases}, \forall k \in \mathcal{K}^t, \\ & \mathcal{B}_k^t \cap \mathcal{B}_{k'}^t = \emptyset, \forall k \neq k', k, k' \in \mathcal{K}^t, \end{aligned} \quad (3.8)$$

其中， \mathcal{K}^t 和 \mathcal{S}_n^t 分别表示第 t 轮迭代时剩余的候选用户集和第 n 个RB上已分配的用户集； \mathcal{B}_k^t 表示此轮为用户 $k \in \mathcal{K}^t$ 分配的RB索引集合。 $A_{k,n}^t(\mathcal{S}_n^t)$ 定义为一个增量加权速率函数，具有如下数学表达式

$$A_{k,n}^t(\mathcal{S}_n^t) = \sum_{i \in \{\mathcal{S}_n^t \cup k\}} w_i r_{i,n}(\mathcal{S}_n^t \cup k) - \sum_{j \in \mathcal{S}_n^t} w_j r_{j,n}(\mathcal{S}_n^t), \quad (3.9)$$

其表述的物理意义是向第 n 个RB额外分配用户 k 后，可获得的加权和速率（目标函数）增量值。

可以看到，子问题 \mathcal{P}_S^t 实际上是 \mathcal{P}_S 在 $L = 1$ ，即不考虑空域并发场景下的简化版本。式(3.5)中关于RB并发数的约束变为此处对各用户RB分配的独占性约束。在每轮迭代中，我们通过限制每个RB最多只能分配给一个用户来降低所求的问题规模。

下面，我们设计一种消息传播方法来求解上述子问题。为了方便描述，我们定义一个二值优化变量矩阵 $\mathbf{X}^t \in \{0, 1\}^{|\mathcal{K}^t| \times N}$ 表示第 t 轮迭代的调度结果，其中 $x_{i,j} = 1$ 即表示将RB j 分配给在此轮候选用户集 \mathcal{K}^t 中的用户 i 。更简洁的，定义

$$\mathbf{x}^t = \text{vec}(\mathbf{X}^t) = [x_{1,1}, x_{1,2}, \dots, x_{i,j}, \dots, x_{\mathcal{K}^t,N}] \in \{0, 1\}^{|\mathcal{K}^t|N} \quad (3.10)$$

表示量化的二值优化向量，其中操作 $\text{vec}(\cdot)$ 表示将矩阵按列量化。基于此，对第 t 轮迭代中所求解的子问题 \mathcal{B}_k^t ，我们首先基于玻尔兹曼分布（Boltzmann-distribution）方程建立关于此时调度指示自变量 \mathbf{x}^t 的联合概率密度函数，其数学表达式可表示为

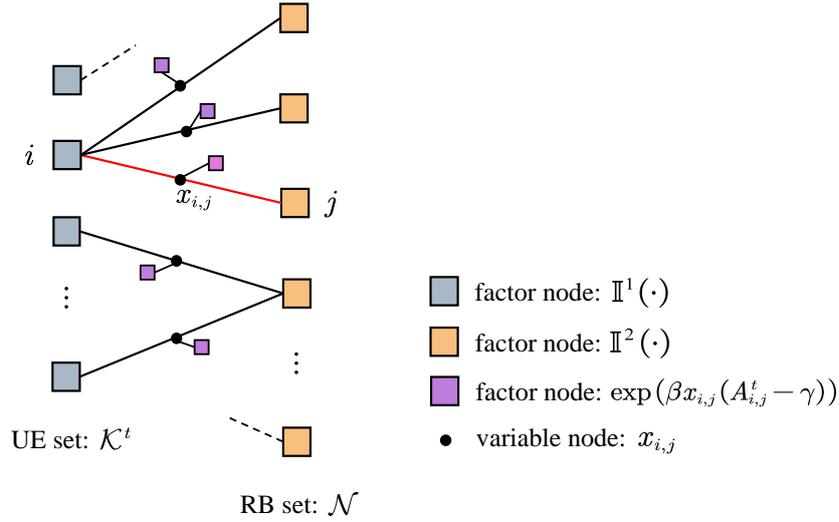
$$\Pr^t(\mathbf{x}^t) = \frac{1}{Z^t} \prod_{k \in \mathcal{K}^t} \mathbb{I}_k^1 \left(\sum_{n=1}^{N-1} \max(0, x_{k,n+1} - x_{k,n}) + \sum_{n=2}^N \max(0, x_{k,n-1} - x_{k,n}) \leq 2 \right) \cdot \prod_{n \in \mathcal{N}} \mathbb{I}_n^2 \left(\sum_{k=1}^{|\mathcal{K}^t|} x_{k,n} \leq 1 \right) \cdot \prod_{i \in \mathcal{K}^t, j \in \mathcal{N}} \exp(\beta x_{i,j} (A_{i,j}^t - \gamma)), \quad (3.11)$$

其中 Z^t 是一个用于确保概率密度和为1的归一化参量； $A_{i,j}^t$ 已于前文介绍过，它表示在给定第 n 个RB上已有的用户调度结果 \mathcal{S}_j^t 后，向其额外添加用户 i 后可获得的增量加权速率； $\beta > 0$ 和 $\gamma \geq 0$ 则是我们设置的两个经验补偿参数，一般情况下取 $\beta = 1$ ， $\gamma = 0$ 。更重要的，上式中的 $\mathbb{I}_k^1(\cdot)$, $\forall i \in \mathcal{K}^t$ 和 $\mathbb{I}_n^2(\cdot)$, $\forall n \in \mathcal{N}$ 表示两类指示函数，它们负责判断括号的条件是否满足并输出判决结果1或者0。其中， $\mathbb{I}_k^1(\cdot)$ 中不等式的物理意义是限制序列 $[x_{k,1}, \dots, x_{k,N}]$ 中元素跳变次数不超过2（这里跳变是指序列内存在相邻两个元素的数值不相同），即对应原问题中RB连续性分配约束； $\mathbb{I}_n^2(\cdot)$ 则表示限制第 n 个RB最多只能分配给一个用户，对应独占性约束。可以看出，只有当所有的指示函数值都为1，即各用户连续性约束和各RB的独占性约束全部满足时，概率密度函数才是一个非零值。这也意味着，只要我们最后找到的解使得 $P^t(\mathbf{x}^t)$ 非零，那么该组解也一定可以满足连续性和独占性约束条件。该问题可建模为形如下式的最大后验（Maximum-a-posteriori, MAP）分配问题，即

$$\mathbf{x}^{t*} = \arg \max_{x_{i,j} \in \{0,1\}^{|\mathcal{K}^t|N}} \Pr^t(\mathbf{x}^t). \quad (3.12)$$

需要指出，求解问题(3.12)得到的 \mathbf{x}^{t*} 就是子问题 \mathcal{P}_S^t 的最优解。但这是一个经典的NP-hard问题，难以直接处理。借助消息传播理论，我们设计了一种最大积（Max-product）置信传播（Belief propagation, BP）方法来近似求解上述问题。

具体来说，我们首先构建问题(3.12)的因子图（Factor-graph）连接模型，如图(3.3)所示。在图中，方型节点表示因子点（Factor-node），圆形节点表示变量点（Variable-node）。根据函数形式的不同，分别用青、粉、黄三种颜色来区分图中左、中、右侧的不同因子点。可以看出，所构建的因子图并不是树形图（Tree-graphical）模型，而存在许多内部环路。根据消息传播算法理论，对于一般的内部成环的因子图模型，BP算法实际上并不能保证一定在有限步内收敛。但同时也能注意到，图中因子点大致可分为左右两个互不相交的子集，且集合内的点没有直接边相连，这大体上类似于一种二分图（Bipartite-graph）结构。受此特殊的对称性结构启发，我们定义了两类四种不同的消息类型并设计了左右振荡式的更新传播方法。经过大量实验，这种更新策略可以使得算法在有限步内收敛。具体来说，BP算法按照以下四个步骤在因子图上迭代运行（为了书写简洁，这里暂时省略了表示轮次的上标 t ）：


 图 3.3 第 t 轮调度时的因子图

(1) 步骤1: 消息 $\nu_{ij \rightarrow i}(x_{i,j})$ 的更新

$\nu_{ij \rightarrow i}(x_{i,j})$ 表示变量点 $x_{i,j}$ 准备传向左因子点 i 的消息。根据消息传播规则，它直接将与其相连的另外两个因子点（粉色和黄色）传来的信息融合（相乘），其计算表达式为

$$\nu_{ij \rightarrow i}(x_{i,j}) \cong \mu_{j \rightarrow ij}(x_{i,j}) \cdot \exp(\beta x_{i,j}(A_{i,j}^t - \gamma)), \forall i, j. \quad (3.13)$$

在这里，我们不再与式(3.11)一样引入归一化常量 Z^t ，而是用更简明用符号 \cong 表示归一化赋值（Equal-up-to-a-normalization），即此时 $\nu_{ij \rightarrow i}(0) + \nu_{ij \rightarrow i}(1) = 1$ （以下同理）。

(2) 步骤2: 消息 $\mu_{i \rightarrow ij}(x_{i,j})$ 的更新

$\mu_{i \rightarrow ij}(x_{i,j})$ 表示因子点 i 准备传向变量点 $x_{i,j}$ 的消息。根据最大积消息传播算法，它收集来自除 $x_{i,j}$ 外的所有与其相连的变量点消息后，基于函数 $\mathbb{I}_i^1(\cdot)$ 对消息进行处理，并将计算的最大值传向 $x_{i,j}$ 。更新计算表达式为

$$\mu_{i \rightarrow ij}(x_{i,j}) \cong \max_{x_{i,n}, n \neq j} \mathbb{I}_i^1(x_{i,j}, [x_{i,n}, n \neq j]) \cdot \prod_{n \in \mathcal{N} \setminus \{j\}} \nu_{in \rightarrow i}(x_{i,n}), \forall i, j. \quad (3.14)$$

(3) 步骤3: 消息 $\nu_{ij \rightarrow j}(x_{i,j})$ 的更新

与步骤1类似， $\nu_{ij \rightarrow j}(x_{i,j})$ 表示变量点 $x_{i,j}$ 准备传向右因子点 j 的消息，它同样将来自左侧的两个因子点（青色和粉色）消息直接相乘融合，即

$$\nu_{ij \rightarrow j}(x_{i,j}) \cong \mu_{i \rightarrow ij}(x_{i,j}) \cdot \exp(\beta x_{i,j}(A_{i,j}^t - \gamma)), \forall i, j. \quad (3.15)$$

(4) 步骤4: 消息 $\mu_{j \rightarrow ij}(x_{i,j})$ 的更新

与步骤2类似， $\mu_{j \rightarrow ij}(x_{i,j})$ 表示右侧因子点 j 准备回传给变量点 $x_{i,j}$ 的消息。节点 j 收

集来自左侧的除 $x_{i,j}$ 外的所有变量点消息，用函数 $\mathbb{I}_j^2(\cdot)$ 处理后，将最大值传与 $x_{i,j}$ 。其计算表达式为

$$\mu_{j \rightarrow ij}(x_{i,j}) \cong \max_{x_{k,j}, k \neq i} \mathbb{I}_j^2(x_{i,j}, [x_{k,j}, k \neq i]) \cdot \prod_{k \in \mathcal{K}^t \setminus \{i\}} \nu_{kj \rightarrow j}(x_{k,j}), \forall i, j. \quad (3.16)$$

经过步骤1-4后，四种消息变量均各自更新了一次。重复这种左右振荡式的传播过程直至消息不再变化，此时BP算法收敛。之后将传向各变量点 $x_{i,j}$ 的信息融合，根据公式

$$b_{ij}(x_{i,j}) = \mu_{i \rightarrow ij}(x_{i,j}) \cdot \mu_{j \rightarrow ij}(x_{i,j}) \cdot \exp(\beta x_{i,j} (A_{i,j}^t - \gamma)), \forall i, j, \quad (3.17)$$

计算其边缘置信 (Marginal-beliefs) 概率 $b_{ij}(x_{i,j})$ ，并基于

$$x_{i,j} = \begin{cases} 1 & \text{if } b_{ij}(1) \geq b_{ij}(0) \\ 0 & \text{if } b_{ij}(1) < b_{ij}(0) \end{cases} \quad (3.18)$$

对其硬判决。

至此，我们完成对MAP分配问题(3.12)的近似求解，记为 $\mathbf{x}^{t*} = [x_{i,j}^*, i \in \mathcal{K}^t, j \in \mathcal{N}]$ 。仍需要再次指出的是，基于上述BP迭代算法所获得的解可能并不是最优的。一方面是因为所构建的因子图并非有界树形图，BP算法无法保证对任意初始值都能收敛至唯一解。另一方面，在上述迭代的步骤2和步骤4中，我们分别通过遍历除 $x_{i,j}$ 外其余变量 $[x_{i,n}]$ 和 $[x_{k,j}]$ 的取值来计算消息 $\mu_{i \rightarrow ij}(x_{i,j})$ 和 $\mu_{j \rightarrow ij}(x_{i,j})$ 的最大值，即相当于求 $x_{i,j}$ 的最大边缘 (Max-marginals) 概率值。但实际上，最后的解 \mathbf{x}^{t*} 可能无法保证该值的可达性，这也会造成一定的计算误差。但所幸，大量的关于消息传播算法的研究和应用显示了这种基于BP算法的解的性能总体是可观的^[95]。

之后，对各RB上的用户调度结果 \mathcal{S}_j^t 和候选用户集 \mathcal{K}^t 进行更新，即

$$\begin{aligned} \mathcal{S}_j^{t+1} &\leftarrow \mathcal{S}_j^t \cup i_j^*, \forall j \in \mathcal{N}, \\ \mathcal{K}^{t+1} &= \mathcal{K}^t \setminus \text{unique}([i_j^*, \forall j \in \mathcal{N}]), \end{aligned} \quad (3.19)$$

其中 $i_j^* = \arg \text{Find}(x_{i,j}^* \equiv 1)$ 表示在第 j 个RB上，从BP结果 $[x_{1,j}^*, x_{2,j}^*, \dots, x_{|\mathcal{K}^t|,j}^*]$ 中找到值等于1的用户下标；操作 $\text{unique}(\cdot)$ 表示通过删去向量中重复元素（只保留一个）来构造集合。

在这里我们需要指出，由于指示函数 $\mathbb{I}_n^2(\cdot)$ 的限制，这里最多只能找到符合条件的1个下标索引；同时由于指示函数 $\mathbb{I}_k^1(\cdot)$ 的限制，本轮各RB上所选的用户集 $[i_j^*, \forall j \in \mathcal{N}]$ 也将满足连续性分配的约束。换言之，不会存在满足 $j - j' > 1$ 的任意 (j, j') ，使得 $i_j^* = i_{j'}^*$ 成立（这里 $N \geq j > j' \geq 1$ ）。

最后基于本轮更新后的用户调度结果 \mathcal{S}_j^{t+1} ，重新计算增量容量 $A_{i,j}^{t+1}$ 、构建第 $t+1$ 轮的子问题并开始新一轮的BP计算……这样经过 L 轮计算后，算法最终获得联合空频域的用户调度结果。

为了更加清晰的表述上述算法的实施过程，我们将其总结于算法2中。

算法 2 JSFD-SMP算法

输入: 用户集 \mathcal{K} 和RB集 \mathcal{N} ；各用户信道 $\mathbf{H}_{k,n}$ 、QoS权重 $w_k, \forall k \in \mathcal{K}, n \in \mathcal{N}$ ；空域并发的最大用户个数 L

初始化: $\mathcal{S}_j^0 = \emptyset, \mathcal{K}^0 = \mathcal{K}$

for $t = 0, 1, \dots, L-1$ **do**

 给定当前RB上的用户调度结果 \mathcal{S}_j ，基于式(3.9)更新增量速率 $A_{i,j}^t, \forall i \in \mathcal{K}^t, j \in \mathcal{N}$

repeat

 基于式(3.13)更新 $\nu_{ij \rightarrow i}(x_{i,j}), \forall i, j$

 基于式(3.14)更新 $\mu_{i \rightarrow ij}(x_{i,j}), \forall i, j$

 基于式(3.15)更新 $\nu_{ij \rightarrow j}(x_{i,j}), \forall i, j$

 基于式(3.16)更新 $\mu_{j \rightarrow ij}(x_{i,j}), \forall i, j$

until BP算法收敛

 根据BP收敛结果计算 \mathbf{x}^{t*} 并按照式(3.19)更新 \mathcal{S}_j 和 \mathcal{K}^t

 如果 $|\mathcal{K}^t| = 0$ ，则提前结束循环

end for

输出: 用户调度结果 $\mathcal{B}^* = \{\mathcal{B}_k, \forall k \in \mathcal{K}\}$ 或 $\mathcal{S}^* = \{\mathcal{S}_n, \forall n \in \mathcal{N}\}$

3.3.2 算法复杂度分析和实施考量

本小节，我们将分析上述JSFD-SMP算法的计算复杂度和实施方案。综合来看，算法主要由内外两层迭代完成。外层迭代的次数由各RB上空域并发的用户数 L 确定；内层迭代次数与BP算法的收敛快慢有关，不妨记为 I 。在整个算法的运行过程中，所需计算量主要由增量加权和速率 $A_{i,j}^t$ 的更新、步骤2和步骤4中的最大积运算所主导。下面，我们分别考量这三者的浮点数计算复杂度。

首先考虑 $A_{i,j}^t$ 更新的计算复杂度。由式(3.4)可看出，其计算主要包括矩阵求逆和一系列矩阵相乘操作。在第 t 轮更新时，此时计算 $|\mathcal{K}^t|N$ 个取值，每个则需要约 $O(N_r^3 + N_r N_t t)$ 数量级的浮点运算操作。因此经过 L 轮迭代后，共需要数量级为 $\sum_{t=1}^L N |\mathcal{K}^t| \cdot O(N_r^3 + N_r N_t t)$ 的浮点数运算。需要注意，在每轮挑选中，剩余候选用户数 $|\mathcal{K}^t|$ 可能并非是稳定减少的，这

取决于用户信道状况和QoS权重的设置。有可能在少数几轮迭代后，所有用户均得到调度，此时算法提前结束；也有可能算法每轮只会确定一个调度用户，即 $|\mathcal{K}^{t+1}| = |\mathcal{K}^t| - 1$ 。考虑后者这种计算量最高的场景，此时需要约 $C_1 \triangleq O(NKLN_r^3 + NKL^2N_rN_t)$ 数量级的浮点数运算复杂度。

接着我们考量第 t 轮更新时的BP算法复杂度。此时在图(3.3)所示的因子图中，左侧用户因子点数量为 $|\mathcal{K}^t|$ ，右侧RB因子点数量为 N ，变量点数量为 $|\mathcal{K}^t|N$ 。在BP算法的每轮振荡更新中，步骤1和步骤3计算量相较步骤2和4可忽略。其中步骤4的最大积更新相对简单，只涉及一个 $|\mathcal{K}^t|$ 维数组的最大值查找操作，其复杂度为 $O(|\mathcal{K}^t|)$ 。但对步骤2，情况更为复杂。为了满足指示函数 $\mathbb{I}_i(\cdot)$ 的连续性约束，理论上需要从RB下标索引集合 $\{1, 2, \dots, N\}$ 中分别遍历起始坐标 n_{start} 和终止坐标 n_{end} 来生成满足连续性RB约束的 $1 + N(N+1)/2$ 个分配结果，并执行最大积运算以找出最大值。为了描述简单，定义 $S_{n_{\text{start}}, n_{\text{end}}}$ 来表示每个分配结果，即

$$S_{n_{\text{start}}, n_{\text{end}}} = [\dots, x_{i, n_{\text{start}}-1} = 0, x_{i, n_{\text{start}}} = 1, \dots, x_{i, n_{\text{end}}} = 1, x_{i, n_{\text{end}}+1} = 0, \dots], \quad (3.20)$$

特别的，当 $n_{\text{end}} < n_{\text{start}}$ 时， $S_{n_{\text{start}}, n_{\text{end}}}$ 的元素值全为零。可以看出，当 N 很大时，上述 $O(N^2)$ 数量级的计算开销可能是无法接受的。因此，为了降低此步的计算量，我们充分结合问题的特殊结构提出了一种KPS的方法，介绍如下。

KPS 方法：重新审视步骤2所表述最大积子问题。注意到输入消息 $\nu_{in \rightarrow i}(x_{in}), \forall n \in \mathcal{N} \setminus \{j\}$ 只有两个取值，即 $\nu_{in \rightarrow i}(0)$ 和 $\nu_{in \rightarrow i}(1)$ 。我们首先构造一个对数比向量

$$\vec{\nu}_i \triangleq \left[\log \left(\frac{\nu_{i0 \rightarrow i}(1)}{\nu_{i0 \rightarrow i}(0) + \epsilon} \right), \log \left(\frac{\nu_{i1 \rightarrow i}(1)}{\nu_{i1 \rightarrow i}(0) + \epsilon} \right), \dots, \log \left(\frac{\nu_{iN \rightarrow i}(1)}{\nu_{iN \rightarrow i}(0) + \epsilon} \right) \right] \in \mathbb{R}^N, \quad (3.21)$$

其中 $\epsilon > 0$ 是保证除法有效的一个极小修正项。图(3.4)展示了矢量 $\vec{\nu}_i$ 中一段序列的数值排布趋势。可以观察到图中有一些特殊的过零点，其下标索引记为

$$\mathcal{Q}_j = \{\dots, p-1, p, \dots, q-1, q, \dots, m-1, m, \dots, n-1, n, \dots\}. \quad (3.22)$$

将它和其他一些重要点，如边界点 $\{1, N\}$ 、 j 和其相邻点 $\{j-1, j, j+1\}$ 合并后，生成一个关键点集合，即

$$\mathbf{KP}_j = \mathcal{Q}_j \cup \{1, N, j-1, j, j+1\}. \quad (3.23)$$

由此，我们引出这样一个事实：满足步骤2中子问题(3.14)的最优分配结果 $S_{n_{\text{start}}^*, n_{\text{end}}^*}$ ，其起始点下标索引 n_{start}^* 和终止点下标索引 n_{end}^* 必然都属于集合 \mathbf{KP}_j 中。受限于篇幅，这里不给出具体的数学证明，但是可以举一些例子来佐证这个论点。比如仅考虑图(3.4)所示的一段

子序列, 对于 $x_{i,j} = 1$ 时的子问题(3.14), 候选的序列结果只可能有 S_{n_p, n_j} 、 $S_{n_j, n_{n-1}}$ 和 $S_{n_p, n_{n-1}}$; 当 $x_{i,j} = 0$ 时, 对应的最优候选序列只可能从 $S_{n_p, n_{q-1}}$ 和 $S_{n_m, n_{n-1}}$ 中取得。这一结果通过分析图示的排布趋势是容易获得的。

当RB数目 N 较大时, 这种KPS方法可以有效地降低遍历复杂度。因为在大部分场景下, 关键点个数一般是稀疏的。若记其稀疏度为 P , 则有 $|\text{KP}_j| \leq P \ll N, \forall j \in \mathcal{N}$ 。此时步骤2中的遍历复杂度降至约为 $O(P^2)$ 数量级。BP算法更新一次所需的浮点数计算数量级约为 $O(|\mathcal{K}^t|^2 N + |\mathcal{K}^t| NP^2)$ 。同样考虑每轮用户调度数缓慢减少的情况后, BP算法总体需要约 $C_2 \triangleq O(LIK^2 N + LIKNP^2)$ 数量级的运算。

综上, 采取KPS技术后, JSFD-SMP 算法运行所涉及的浮点数运算数量级是

$$C = C_1 + C_2 = O(NKLN_r^3 + NK L^2 N_r N_t + LIK^2 N + LIKNP^2). \quad (3.24)$$

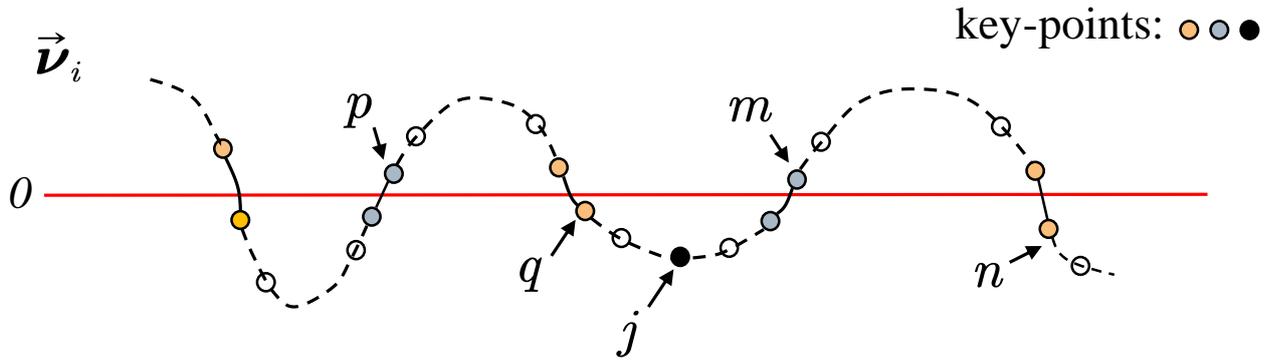


图 3.4 对数比矢量的排布图

在本节的末尾, 我们再简单介绍JSFD-SMP 算法在实际部署中的注意问题和一些加速计算方法。由上述分析结果可知, 算法的计算复杂度与 K^2 和 N 成正比。当用户和RB规模较大时, 所需的计算量可能仍然难以接受。一种合适的解决方法是在算法执行前, 通过用户分组、扩大RB聚合粒度等方法将问题规模缩减。这种方法可有效减少计算开销, 但代价就是由于调度精细度变差导致的性能损失。因此实际部署时需要根据场景需求灵活调整聚合等级以实现计算复杂度和性能之间的折中。同时, 注意到上述BP算法振荡更新过程(步骤1到4)中所涉及的计算是高度解耦的。基于这一特性, 我们可以利用包括GPU在内的多核硬件平台的并行计算架构进一步降低处理时延开销。这使得算法可以支持TTI级的实时调度需求。

3.4 仿真结果与分析

3.4.1 基线方案和性能指标

本子章节，我们将基于Saleh-Valenzuela信道模型搭建仿真系统，比较所提出的JSFD-SMP调度算法与其他几种基线方案在小区吞吐量和用户速率公平性指标上的性能表现。其中多用户的业务模型选为eMBB通信业务^[96]，且各TTI中用户 k 的数据包到达速率服从均值为 λ_k 的泊松到达分布（Poisson-arrival-distribution）。表(3.1)中详细的总结了系统参数。

表 3.1 信道和业务模型参数设置

名称	设置
信道模型	Saleh-Valenzuela
用户数	10, 20, 30
散射路径数	4
载波频率	3GHz
RBG数目	40
RB聚合粒度	1RBG=4RBs
带宽	58MHz
天线阵列	ULA
发送天线数	32
空域并发最大用户数	2, 4
路径增益	-10dB到10dB间均匀分布
业务模型	eMBB
QoS权重准则	M-LWDF
包到达平均速率	20Mbit/s到25Mbit/s间均匀分布
TTI持续时间	1ms
调度时间	100TTIs

在本章中，我们选取工业界中常用的邻位局部扩展（Localized expansion of adjacent positions, LEAP）^[41]和最近新提出的双端联合分配（Joint allocation with dual ends, JADE）^[42]作为比较的基线方案。需要指出的是，这两种基线都假设每个RB由一个用户独占，因而只考虑了在频域连续性分配约束下的用户调度问题。而其他同时考虑空域用户并发和频域连续约束的调度方案大多都采取基于可行分配图案的搜索解法，只适用于问题规模较小的场景^[35,89]。在本章所考虑的发送天线数、用户数和RB数较多的中重载场景下，这类方法由于搜索复杂度太高而基本上无法正常使用。因此，为了更详尽地比较，我们还参考

文献^[97]设置了一组多用户最优资源分配 (Optimum resource allocation, ORA) 基线方案。ORA方案忽略RB的连续性约束, 并行地在每个RB上从所有候选用户集中穷搜得到最优的并发用户集合。受限于计算复杂度, 我们设置该方案的最大用户并发数为2, 并以此作为性能上界 (Upper-bound) 参考。

对于性能指标的选取, 正如文献^[16]所提到的, 吞吐量和用户速率公平性之间的最优权衡一直是调度方案设计的重要目标之一。因此, 本章节将遵循于此, 选择小区用户平均吞吐量 (Throughput, THP) 和Jain's 公平性因子 (Jain's fairness index, JFI) 作为方案性能比较的两个重要指标, 其具体计算公式为

$$\begin{aligned} \text{THP} &= \frac{1}{KT} \sum_{t=1}^T \sum_{n \in \mathcal{N}} \sum_{k \in \mathcal{S}_n(t)} r_{k,n}(t), \\ \text{JFI} &= \frac{\left[\frac{1}{T} \sum_{t=1}^T \sum_{k \in \mathcal{K}} \sum_{n \in \mathcal{B}_k(t)} r_{k,n}(t) / \lambda_k \right]^2}{K \sum_{k \in \mathcal{K}} \left[\frac{1}{T} \sum_{t=1}^T \sum_{n \in \mathcal{B}_k(t)} r_{k,n}(t) / \lambda_k \right]^2}, \end{aligned} \quad (3.25)$$

其中JFI是的取值范围在 $[1/K, 1]$ 之间。当JFI为1时, 各用户的体验速率完全公平; 而JFI等于 $1/K$ 则表示系统只选择了一个用户服务, 此时用户调度最不公平。上式中其余各项符号参量均在前文介绍过, 这里不再赘述。

3.4.2 性能比较和分析

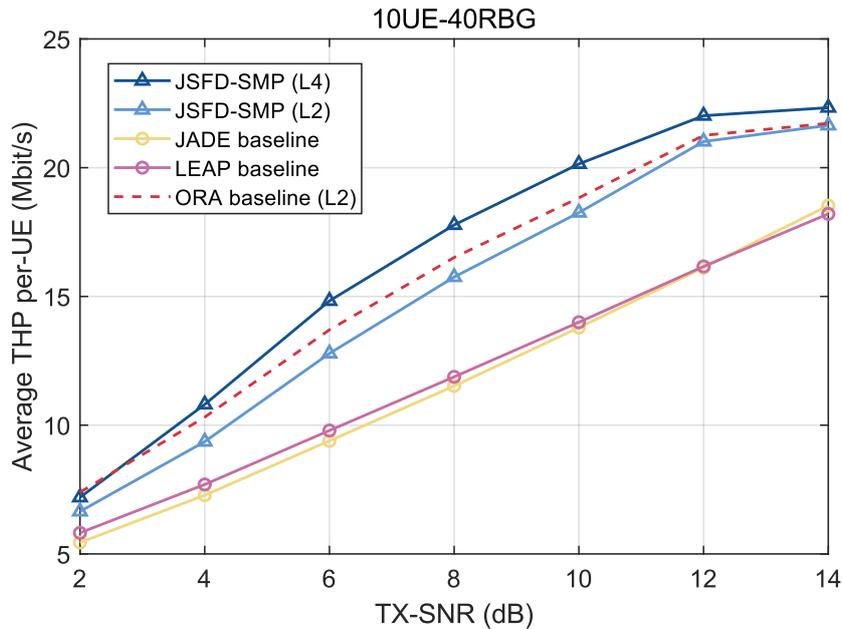


图 3.5 10用户场景时的用户平均吞吐量性能

如图(3.5)、(3.6)和(3.7)所示，我们分别仿真了在10、20和30用户场景下的小区用户平均吞吐量性能。需要指出的是，为了同时考虑Finite-buffer和Full-buffer两种场景下的吞吐量性能，我们在三种场景下固定了BS的发送功率范围。如(3.5)所示，在小区系统用户数较少时，当BS发送功率超过12dB后，可以看到用户平均吞吐量曲线逐渐饱和并逼近至各用户的平均包到达速率，即约23.5Mbit/s。这是显然的，因为用户的吞吐量不可能超过其来包数据量。这模拟了在有限队列时的各方案调度性能。而对于图(3.6)和图(3.7)所示的用户数规模较大的场景，由于功率受限，系统中各用户的发送速率将小于来包速率。这会造成各用户数据发送数据的逐渐累积，因此模拟的是Full-buffer下的调度性能。

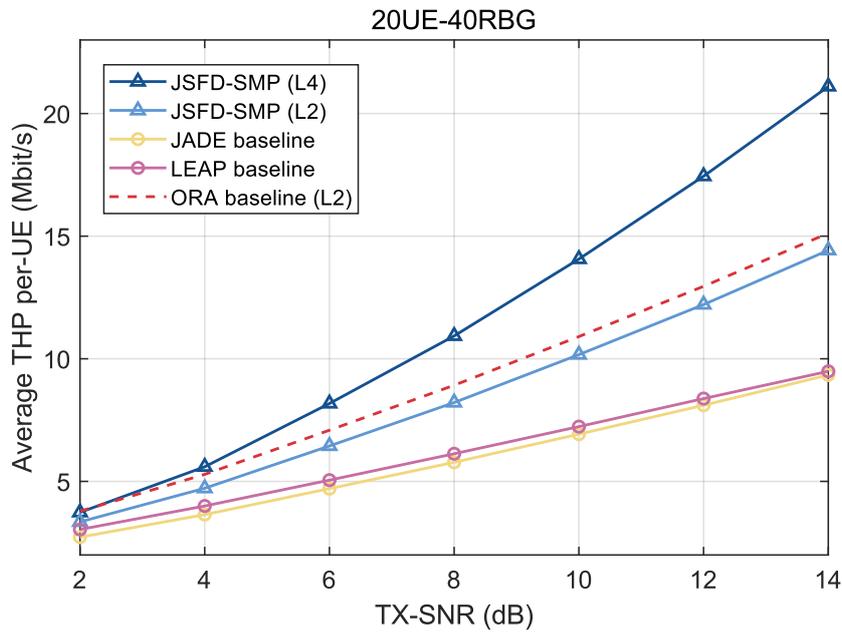


图 3.6 20用户场景时的用户平均吞吐量性能

可以看到，不论在何种场景下，所提出的JSFD-SMP调度方案相较现有的JADE和LEAP基线方案都有着显著的吞吐量性能提升。具体来说，JADE和LEAP两种基线在吞吐量性能上区别不大，都差于所提JSFD-SMP方案在空域并发数为2时的性能，其中等效发送功率增益约为4dB。在空域用户并发数为4时，所提方案则将性能差距进一步拉大，增益在6dB以上。同时，相较于空域并发数为2时的性能上界基线方案ORA，保持相同并发用户数的JSFD-SMP方案在三种场景下都可以达到与其非常接近的性能表现，这证明了所提方案的在吞吐量性能表现上的显著优越性。

图(3.8)比较了各方案在三种不同场景下的JFI性能。可以看到，各方案的JFI性能随着小区用户数的增长均有不同程度的恶化。具体来说，在三种基线方案中，LEAP和ORA的JFI性能表现优异且差异不大。JADE基线方案性能退化最为严重，其JFI指标在30用户数时降至0.9，

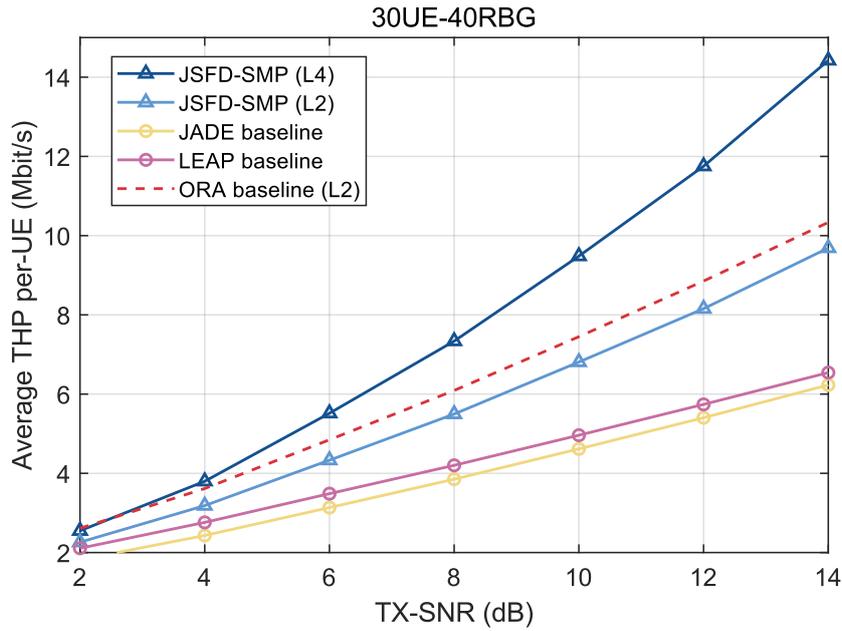


图 3.7 30用户场景时的用户平均吞吐量性能

在所有方案中最差。对于所提的JSF-SMP方案，在4用户并发时同样可以达到与LEAP和ORA基线方案一致最优的性能。在2用户并发时，虽然性能相较其余三者略有损失，但差距仅不到1%。这证明了所提方案在实现显著吞吐量性能提升的同时也可以优秀地兼顾用户体验速率的公平性。

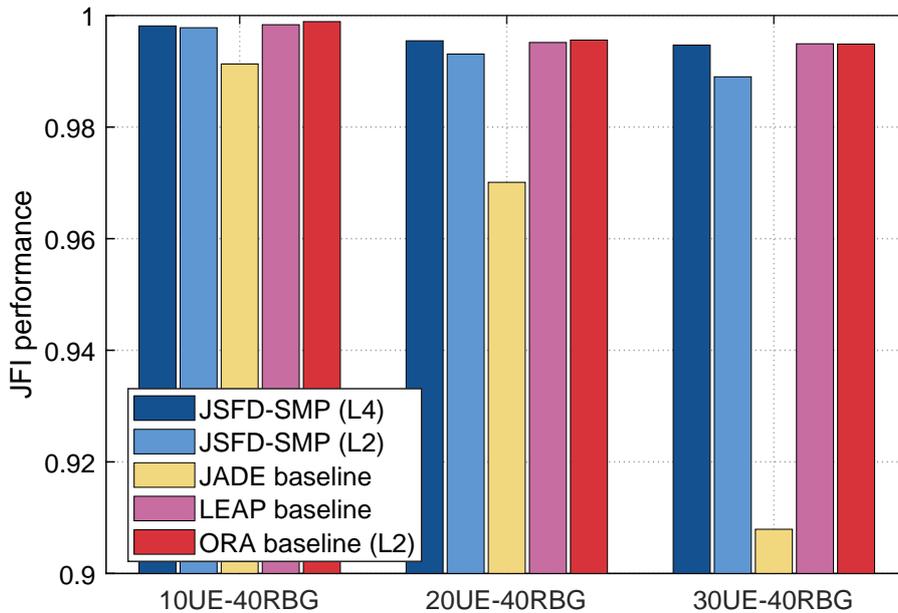


图 3.8 不同场景下各方案的JFI性能比较

3.5 本章小节

本章中，我们提出了一种JSFD-SMP算法来解决单小区下行系统中支持空域并发的多用户频分调度问题。受限于最大空域并发数和RB资源连续性分配这两个棘手的实际链路约束，该问题难以直接求解。为此，我们首先将原问题分解，通过在每轮迭代中构造一个只包含RB连续性分配约束的子问题将求解空间降低。对分解后的子问题，我们将用户和RB间仍然复杂的耦合关系建模成因子图，并设计了一种消息传播算法来有效求解。算法本身高度解耦的结构非常有利于在GPU等支持并行计算的多核平台上实现。另外，我们还提出了一种称为KPS的方法在不损失性能的同时进一步降低了算法复杂度。最后，我们基于大量的仿真实验比较了所提算法和现有几种基线调度方案在吞吐量和公平性等指标上的性能表现。实验结果证明了所提方案可以在保证小区用户体验速率公平性的同时显著提升小区吞吐量性能。

4 基于EESM准则的联合空频预编码及低MCS下增益空间探索

4.1 研究动机和主要成果

在massive MIMO系统中，BS通过利用预编码技术可将发射能量集中到目标移动终端上，在降低单根天线辐射功率的同时，减少对其他用户的干扰，从而可在不增加额外带宽的情况下成倍地提高频谱利用率和能量效率。

从绪论中的调研结果可以看出，当前学术界针对下行预编码设计方案的研究大多都基于相对理想的传输条件，比如假设瞬时CSI完美已知或者输入数据服从理想高斯分布。很少有工作直接优化传输的BLER性能，而是将优化目标选为基于香农容量公式的传输速率或系统吞吐量。在大部分场景下，该形式优化目标的选取是合理且有效的。因为现有的无线通信系统存在较完善的链路质量调整机制，称为自适应调制编码（Adaptive modulation and coding, AMC）。以下行传输为例，BS基于用户反馈的CQI或利用上下行信道的互易性推断得到下行信道的质量特征后，可以通过调整合适的MCS和码率（Code-rate）以确保链路的传输质量，将BLER指标控制在0.01以下。此场景下，BLER对系统容量的影响较小，因此如何选择合适的预编码方案保证系统传输速率最大化是至关重要的目标。但对于某些特殊的通信场景，比如在考虑如何为距离BS较远的用户尽可能提供稳定服务的问题中，由于用户信道质量相对较差，需要在较低的MCS和码率下的评估用户的链路性能。此时AMC可调整的范围较小，因此BLER性能指标往往成为系统的重要优化对象。在上述称为低MCS的远距覆盖增强问题中，设计以BLER性能为指标的预编码方案并分析此场景下的增益空间，对评估5G/B5G系统的BS覆盖范围和实际部署策略有重要的指导意义。

因此，基于上述研究现状，本章节将以实际的BLER传输性能为指标，设计一种同时考虑非理想CSI和Finite-alphabet输入的联合空频预编码方案，并探索所提方案在低MCS的远距覆盖增强通信场景的增益空间。然而，为了处理上述问题，我们需要解决以下三个关键技术难点：

- (1) 以BLER为优化目标的问题难以直接求解：BLER性能是一个极难刻画的目标函数，它不仅与信道和发送功率有关，也受MCS、码长、码率和信道编码方式等

诸多实际链路参数的影响。有研究者在致力于推导满足给定条件下BLER的性能界限或可达性的证明。比如文献^[98]考虑加性白高斯信道（Additive white Gaussian noise, AWGN）模型，在给定码长、信道容量和信道色散（Dispersion）信息下，推导了BLER的性能上界表达式；文献^[99]则基于块衰落信道和给定的码率，给出了BLER大于某一阈值的可达性证明。不过在我们的场景中，上述研究结果都存在明显的缺陷而无法适用：文献^[98]的推导没有考虑码字（Codeword）服从有限星座集分布的情况，无法准确的适用实际系统；文献^[99]提供的阈值估计太过松散，无法为预编码设计提供直接且相对准确的指导。据我们所知，几乎没有研究者直接基于编码系统（Coded-system）设计空频联合的预编码方案。

- (2) **低MCS和码率下增益空间探索比较困难**：在链路评估中，一般将BLER随SNR下降的曲线斜率称为分集阶数（Diversity-order），它刻画了BLER随着SNR增加时下降的速率^[100]。在低MCS和码率场景下，BLER-SNR曲线的斜率往往很大，呈现出一种“瀑布”效应。我们可能只需要将功率提升仅仅不到1dB，BLER就可以从1迅速降至0.01左右，而这会导致增益空间的探索变得困难。因此我们必须精细地设计功率分配或预编码方案并基于大量详实的仿真来排除偶然误差。
- (3) **非理想CSI下预编码方案设计**：同时考虑关键技术难点(1)和(2)后，非理想CSI对仿真结果的破坏性影响较其他高MCS通信场景也更大。这对空频联合的预编码方案设计提出了更大的挑战。

为了处理上述的各种技术难题，本章节我们将提出一种基于度量函数指示的逐级优化算法来解决此场景下针对BLER的联合空频预编码设计问题。具体来说，为了解决挑战(1)，我们首先引入一种称为指数型有效信干噪比映射（Exponential effective signal-to-interference-plus-noise ratio mapping, EESM）的合适的性能度量（Performance-metric）函数，将原本复杂的编码系统BLER优化问题可转化为无编码系统（Uncoded-system）中针对所选度量函数的优化问题。为了证明所选度量的合理性，我们进行了大量的蒙特卡洛仿真实验。实验结果说明所选的基于EESM的度量函数可以较好的刻画该场景下BLER性能。对于完美CSI的单用户场景，我们通过假设BS和用户采取基于信道SVD的预编码方案^[78]，构建一个只涉及功率权值分配的简易问题。然而，在考虑存在信道估计、信道反馈误差的非理想CSI场景时，转换后的函数仍然是非凸的。为此，我们提出了一种逐级迭代的方法，在每次迭代中建模一个考虑信道误差的加权和速率最大化的子问题，并利用Shi等人^[101]提出加权均方误差和最小化（Weighted sum mean-square-error minimization, WMMSE）算法求解。我们称所提的上述算法为基于EESM准则的逐级WMMSE（EESM-based successive WMMSE, E-S-WMMSE）优化方法。为了探索所提方案相较经典基线方案的增益空间，我们设置了大量

仿真实验，比较验证了不同MCS、传输数据流数目情况下的增益结果，证明了所提方案的有效性。实验结果可为工业界在处理类似覆盖增强场景下的性能增益空间评估问题提供一定的理论和实践指导。

本章节的剩余内容安排如下：第二子章节首先介绍系统模型和原始优化问题，随后提出基于EESM准则的问题转化方法；第三子章节针对转化后的问题，介绍了在考虑非理想CSI时的E-S-WMMSE算法框架，同时分析了算法复杂度；第四子章节基于大量详尽的链路级仿真实验，首先验证所选择的EESM度量函数有效性，之后在理想CSI和非理想CSI两种信道条件下比较所提方案和若干基线方案BLER性能差异，测试了在不同传输流和码率下的增益空间；最后进行本章小结。

4.2 系统模型与问题建模

4.2.1 系统模型

本章节，我们考虑一个宽带下行的massive MIMO系统。基站发送天线数和设备接收天线数分别即为 N_t 和 N_r ($N_t \gg N_r$)。在每个TTI，我们假设BS向用户传输的码字被编码承载于 N 个子载波和 L 个空间流上，其中 $L \leq N_r$ 。记 $\mathbf{V}_n = [\mathbf{v}_{n,1}, \mathbf{v}_{n,2}, \dots, \mathbf{v}_{n,L}] \in \mathbb{C}^{N_t \times L}$ 表示BS在第 n 个子载波上用于传输发送数据 $\mathbf{s}_n \in \mathbb{C}^{L \times 1}$ 的预编码矩阵，其中 $\mathbf{v}_{n,l} \in \mathbb{C}^{N_t \times 1}$ 表示用于承载第 l 个数据流的预编码矢量。因此，用户在第 n 个载波上的接收信号可以表示为

$$\mathbf{y}_n = \mathbf{H}_n \mathbf{V}_n \mathbf{s}_n + \mathbf{n}_n, \quad (4.1)$$

其中， $\mathbf{H}_n \in \mathbb{C}^{N_r \times N_t}$ 表示第 n 个载波上的下行信道矩阵； $\mathbf{n}_n \sim \mathcal{CN}(0, \sigma_n^2 \mathbf{I})$ 表示加性白高斯噪声。同时，我们假设发送数据符号矢量服从均值为0，方差为1的复高斯分布，即满足 $\mathbb{E}[\mathbf{s}_n \mathbf{s}_n^H] = \mathbf{I}_L$ （下标 L 表示单位阵的维度）。

用户在获得接收信号 \mathbf{y}_n 后，通过线性接收机均衡后得到恢复的估计信号

$$\hat{s}_{n,l} = \mathbf{u}_{n,l}^H \mathbf{y}_n, \quad (4.2)$$

其中 $\mathbf{U}_n = [\mathbf{u}_{n,1}, \mathbf{u}_{n,2}, \dots, \mathbf{u}_{n,L}] \in \mathbb{C}^{N_r \times L}$ 表示第 n 个子载波上的接收波束赋形矩阵。我们可将各载波上的不同数据流视为分离的虚拟收发端口，各自经历具有不同增益的单输入单输出（Single-input single-output, SISO）信道，对应的信干噪比（Signal-to-interference-plus-noise ratio, SINR）则具有如下的数学表达式

$$\gamma_{n,l} = \frac{|\mathbf{u}_{n,l}^H \mathbf{H}_n \mathbf{v}_{n,l}|^2}{\left| \mathbf{u}_{n,l}^H \sum_{j=1, j \neq l}^L \mathbf{H}_n \mathbf{v}_{n,j} \right|^2 + \sigma_n^2 \|\mathbf{u}_{n,l}\|^2}. \quad (4.3)$$

需要指出的是，在收发两端完美已知CSI的情况下，使得互信息最大的最优 $\{\mathbf{V}, \mathbf{U}\}$ 设计是将信道SVD后的左右奇异向量作为各流的收发波束赋形矢量，同时结合基于注水的功率分配方案^[47]。此时，公式(4.3)分母中描述流间干扰的第一项变为零，SINR只与流间的功率分配权值有关。但当完美CSI无法获得时，基于非理想CSI设计的 $\{\mathbf{V}\}$ 无法完全消除流间干扰，此时的问题求解将会变得复杂。

4.2.2 问题建模

本小节，我们希望求解在给定发送功率约束下的系统BLER性能最优的预编码矩阵设计问题，问题的数学表达式如下

$$\begin{aligned} \mathcal{P}_B : \min_{\{\mathbf{V}\}} & BLER(\{\mathbf{H}_n, \mathbf{V}_n, \forall n\}, mcs) \\ \text{s.t.} & \sum_{n=1}^N \text{Tr}(\mathbf{V}_n^H \mathbf{V}_n) \leq P, \end{aligned} \quad (4.4)$$

其中 P 为BS的最大发送功率；函数 $BLER(\cdot)$ 表示在给定当前TTI传输条件下所能取得的平均BLER性能。这里有两点值得说明，1) 由于系统瞬时CSI各不相同，我们无法在每个TTI重复发送数据包并统计经历多个噪声样本后的平均BLER性能。因此目标函数在实际系统中是无法验证的，但作为系统的优化对象，它的选取是合理的；2) BLER函数关于传输条件的表达式十分复杂，它不仅与式中所标注的瞬时CSI、波束赋形矩阵 $\{\mathbf{V}\}$ 、MCS有关，还受TBS、信道编码方式等一系列具体传输参数的影响。不过在实际系统中，这些传输参数往往都是确定的。因此在本章中，我们重点研究不同MCS下预编码方案的性能增益，而不关注后者参数对BLER数值的具体影响。不过即使如此，BLER目标函数仍具有强烈的非凸性，这导致原问题难以直接求解。幸运的是，3GPP协议的早期规范文档通过大量的链路仿真，给出了一种面向多载波正交频分多址（Orthogonal frequency division multiplexing, OFDM）系统中基于有效信干噪比（Effective SINR mapping, ESM）映射的BLER的近似表达式^[102]，即

$$BLER(\{\mathbf{H}_n, \mathbf{V}_n, \forall n\}, mcs) \stackrel{1}{\approx} BLER(\{\gamma_{n,l}, \forall n, \forall l\}, mcs) \stackrel{2}{\approx} BLER_{awgn}(\gamma_{\text{eff}}, mcs). \quad (4.5)$$

上式中， $\stackrel{1}{\approx}$ 表示原问题的BLER可以进一步写成与各载波各流上的等效信道增益有关的函数（ $\gamma_{n,l}$ 由公式(4.3)表述）； $\stackrel{2}{\approx}$ 是至关重要的，它进一步说明了这样一个事实：对一个承载于具有不同增益 $\gamma_{n,l}$ 的子信道上传输的码字，可以找到一个增益为 γ_{eff} 的AWGN平衰落信道，使得其两种传输下的BLER性能大致相同，且 γ_{eff} 可以完全由 $\gamma_{n,l}$ 映射得到（具体的映射表达式和物理意义将在下一小节详细介绍）。我们将看到，基于这种ESM的方法，原问

题中针对编码系统BLER优化的复杂问题可被转化为无编码系统中关于 γ_{eff} 的优化问题，这极大地降低了原问题的求解难度，同时为预编码方案的设计提供了可行的思路。

4.2.3 EESM准则下的BLER优化问题重构

ESM是4G LTE和5G NR无线系统中分析信道质量的重要方法。对于给定的码字，它基于其经历的多个子信道的不同增益信息 $\gamma_{n,l}$ ，计算出一个能保证相同码字错误率的等效平衰落信道增益 γ_{eff} ，其具体数学表达式为

$$\gamma_{\text{eff}} = \lambda f^{-1} \left(\frac{1}{NL} \sum_{n=1}^N \sum_{l=1}^L f \left(\frac{\gamma_{n,l}}{\lambda} \right) \right), \quad (4.6)$$

其中 $f(\cdot)$ 表示映射函数， $f^{-1}(\cdot)$ 是其逆函数；参数 $\lambda > 0$ 是与MCS有关的SINR缩放参数。根据映射函数不同，ESM又可细分为EESM、互信息型ESM（Mutual information ESM, MIESM）、容量型ESM（Capacity ESM, CESM）和对数型ESM（Logarithmic ESM, LESM）等。综合考虑应用场景和实现复杂度，EESM方法的应用最为广泛。它作为一种可将多个子信道SINR有效压缩的方法，被用于解决CQI反馈、链路自适应和频域资源调度等问题^[103]，如图(4.1)所示。例如，在实际系统中，各用户可基于EESM方法计算下行宽带信道的CQI并反馈至BS，以实现闭环的资源调度和链路速率适配^[104]。但另一方面，EESM方法作为一种能够准确指示码字错误率的技术^[105]，却很少被用于解决直接面向BLER性能提升的优化问题。其原因已在前文阐述过，即目前大部分通信场景更关注系统的容量性能，BLER可通过链路自适应和自动混合重传（Hybrid automatic repeat-request, HARQ）等技术被控制在较低的水平。但在我们所考虑的低MCS远距覆盖场景中，MCS调整的空间不大，设计面向BLER的预编码方法和增益空间评估理论是更有意义的。所以受此启发，我们将EESM方法引入所考虑BLER优化问题中。通过改写原问题 \mathcal{P}_B 中复杂的目标函数，将问题转化为针对等效信噪比 γ_{eff} 的无编码系统优化问题，即

$$\begin{aligned} \mathcal{P}'_B : \min_{\{\mathbf{V}\}} \exp \left(-\frac{\gamma_{\text{eff}}}{\lambda} \right) &= \frac{1}{NL} \sum_{n=1}^N \sum_{l=1}^L \exp \left(-\frac{\gamma_{n,l}}{\lambda} \right) \\ \text{s.t. } \sum_{n=1}^N \text{Tr} (\mathbf{V}_n^H \mathbf{V}_n) &\leq P. \end{aligned} \quad (4.7)$$

转化后的问题 \mathcal{P}'_B 具有更加清晰直观的数学形式，我们将在下一章节具体介绍的求解方法。不过这里需要指出的是，EESM关于BLER的估计性能非常依赖参数 λ 的选取。这意味着，为了获得不同的MCS下的最优SINR缩放参数，我们需要基于大量的链路数据样本 \mathcal{S} 求

解形如下式的关于BLER的最小均方误差问题：

$$\lambda = \arg \min_{\lambda} \left\{ \sum_{j \in \mathcal{S}} [BLER_j - BLER_{avgn}(\gamma_{\text{eff}}^j)]^2 \right\}. \quad (4.8)$$

不过幸运的是，已经有许多研究者针对此类问题进行了大量工作并总结绘制了不同MCS、码率下的参数选择表^[106]。在我们之后的仿真实验中，我们也依照此表确定 λ 的取值。

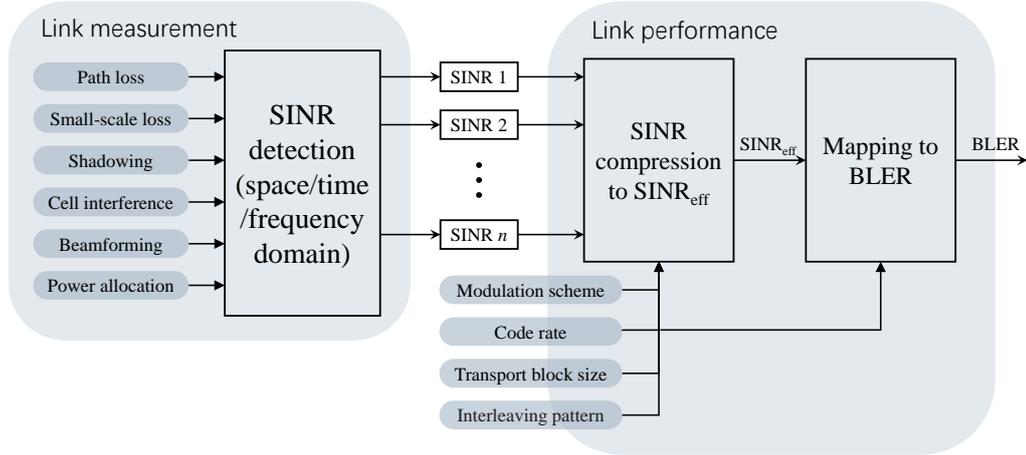


图 4.1 ESM方法在通信链路中的应用

4.3 E-S-WMMSE算法

本小节，我们将分别在完美CSI和非理想CSI两种场景下设计算法以求解转化后的问题 \mathcal{P}'_B 。在第一种场景，即收发两端已知完美的信道矩阵时，我们采取广为人知的基于信道SVD的预编码方法，将问题进一步转化为易于求解的功率权值分配问题。而对于存在信道估计或信道反馈误差的非理想CSI场景（假设误差协方差矩阵已知）， \mathcal{P}'_B 是一个棘手的非凸优化问题。为此，我们设计了一种基于逐级迭代的E-S-WMMSE算法来寻找目标函数的局部最优解。具体来说，通过将原问题泰勒展开，我们在每一次迭代中构造一个形如加权和速率最大化的子问题。对该子问题，我们利用它和加权均方误差和最小化问题的等价性，在Shi等人^[101]所提出WMMSE算法框架基础上设计了考虑信道误差的收发预编码迭代算法。算法的具体细节和复杂度分析将在以下各小节中详细介绍。

4.3.1 理想CSI场景下的算法设计

此场景下，假设BS和用户均可获得下行宽带CSI的完整信息，并采取基于信道SVD的波束赋形方案。需要事先指出的是，虽然Telatar^[107]已证明了采用SVD预编码和注水功率

分配方案是使得香农信道容量可达的最优方案。但对于我们所考虑的以BLER为优化目标的 \mathcal{P}'_B ，基于SVD的波束赋形不一定是最优的。因此本小节所介绍的转化问题实际上是对原问题的一种简化，可为低复杂度近似算法的设计提供一种思路。我们将在下一小节结合非理想CSI具体介绍关于问题 \mathcal{P}'_B 的完整求解框架。

在本小节考虑的场景下，对第 n 载波上的信道矩阵 \mathbf{H}_n ，其SVD可以表示 $\mathbf{H}_n = \bar{\mathbf{U}}_n \boldsymbol{\Sigma}_n \bar{\mathbf{V}}_n^H$ 。其中 $\bar{\mathbf{U}}_n \in \mathbb{C}^{N_r \times N_r}$ 和 $\bar{\mathbf{V}}_n \in \mathbb{C}^{N_t \times N_t}$ 均为酉矩阵； $\boldsymbol{\Sigma}_n \in \mathbb{C}^{N_r \times N_t}$ 是对角矩阵，其对角元素表示各SISO子信道的增益值。我们假设其已按降序排列且均为正数，即 $\sqrt{g_{n,1}} \geq \sqrt{g_{n,2}} \geq \dots \sqrt{g_{n,N_r}} > 0$ 。此时，SVD预编码下收发波束赋形矩阵 $\{\mathbf{V}_n, \mathbf{U}_n\}$ 分别由 $\bar{\mathbf{V}}_n$ 、 $\bar{\mathbf{U}}_n$ 的前 L 列构成，即

$$\begin{aligned} \mathbf{V}_n &= [\bar{\mathbf{v}}_{n,1}, \bar{\mathbf{v}}_{n,2}, \dots, \bar{\mathbf{v}}_{n,L}] \cdot \Lambda_n, \\ \mathbf{U}_n &= [\bar{\mathbf{u}}_{n,1}, \bar{\mathbf{u}}_{n,2}, \dots, \bar{\mathbf{u}}_{n,L}], \forall n, \end{aligned} \quad (4.9)$$

其中， $\Lambda_n = \text{diag}(\sqrt{p_{n,1}}, \sqrt{p_{n,2}}, \dots, \sqrt{p_{n,L}}) \in \mathbb{R}^{L \times L}$ 是功率分配对角阵。将上式代入式(4.7)后，问题 \mathcal{P}'_B 可进一步简化为

$$\begin{aligned} \min_{\mathbf{p}} \quad J &\triangleq \frac{1}{NL} \sum_{n=1}^N \sum_{l=1}^L \exp\left(-\frac{p_{n,l} g_{n,l}}{\lambda \sigma_n^2}\right) \\ \text{s.t.} \quad &\sum_{n=1}^N \sum_{l=1}^L p_{n,l} \leq P, \end{aligned} \quad (4.10)$$

其中 $\mathbf{p} = \{p_{n,l}, \forall n, l\} \in \mathbb{R}^{NL \times 1}$ 表示在各子信道上的功率分配结果。

容易看出，上述问题是只关于功率分配因子 $p_{n,l}$ 的凸优化问题，可以使用包括拉格朗日乘子法在内的一系列经典的凸优化方法求解，这里不再赘述。

4.3.2 非理想CSI场景下的算法设计

在实际的无线通信系统中，受限于信道估计、信道量化反馈的精度，用户和BS一般无法获得准确的下行信道信息，而只能得到其估计值，即

$$\hat{\mathbf{H}}_n = \mathbf{H}_n - \mathbf{H}_{n,e}, \quad (4.11)$$

其中， $\mathbf{H}_{n,e} \in \mathbb{C}^{N_r \times N_t}$ 定义为BS获得的第 n 个子载波上的信道误差。在这里，我们首先作出以下两个假设：1) 假设信道估计（或量化）是无偏的，即 $\mathbb{E}[\text{vec}(\mathbf{H}_{n,e})] = \mathbf{0}$ ；2) 假设信道误差的统计协方差矩阵 $\mathbf{R}_n \triangleq \mathbb{E}[\text{vec}(\mathbf{H}_{n,e}) \cdot \text{vec}(\mathbf{H}_{n,e})^H] \in \mathbb{C}^{N_t N_r \times N_t N_r}$ 是已知的。在上述两个假设中，假设二是合理的，因为BS和用户可根据大量的历史信道的样本统计信息计算得到 \mathbf{R}_n ，但假设一却并不总是成立。因为在现有的通信系统中，BS和用户是基于导频参考

信号，如SRS和解调参考信号（Demodulation reference signal, DM-RS）等，来获得上下行的估计信道。但这些参考信号并非每时每刻都发送，而是按照给定的导频图案（Pattern）和密度占据部分时频资源发送。对那些未被插入导频的子载波信道，其估值一般只能通过线性插值或外推等方法得到，这可能会导致有偏误差。不过在本章考虑的场景中，为了使得建模形式更加简洁，我们认为用户端可以获得下行CSI的完美估计，BS已知估计信道 $\hat{\mathbf{H}}_n, \forall n$ 和误差协方差矩阵 $\mathbf{R}_n, \forall n$ ，且此时的信道误差主要来源于信道反馈过程中的量化误差，因此假设一、二均得到满足。

基于上述假设，我们首先考察在非理想CSI时各流传输的平均可达速率，即

$$r_{n,l} = \mathbb{E}_{\mathbf{H}_{n,e}|\hat{\mathbf{H}}_n} \left[\log \left(1 + \mathbf{v}_{n,l}^H \mathbf{H}_n^H \left(\sum_{j=1, j \neq l}^L \mathbf{H}_n \mathbf{v}_{n,j} \mathbf{v}_{n,j}^H \mathbf{H}_n^H + \sigma_n^2 \mathbf{I}_{N_r} \right)^{-1} \mathbf{H}_n \mathbf{v}_{n,l} \right) \middle| \hat{\mathbf{H}}_n \right], \forall n, l, \quad (4.12)$$

其中， \mathbf{I}_{N_r} 表示维度为 N_r 的单位阵。需要指出，由于对数函数和矩阵求逆操作的存在，上式并没有一个简洁显式（Explicit）的期望表达式。因此，从实用的角度出发，我们一般会取如下近似^[108]：

$$r_{n,l} \approx \log \left(1 + \mathbf{v}_{n,l}^H \hat{\mathbf{H}}_n^H \left(\sum_{j=1, j \neq l}^L \hat{\mathbf{H}}_n \mathbf{v}_{n,j} \mathbf{v}_{n,j}^H \hat{\mathbf{H}}_n^H + \mathbf{J}_n + \sigma_n^2 \mathbf{I}_{N_r} \right)^{-1} \hat{\mathbf{H}}_n \mathbf{v}_{n,l} \right), \forall n, l, \quad (4.13)$$

其中 $\mathbf{J}_n \in \mathbb{C}^{N_r \times N_r}$ 表示由信道误差引起的额外干扰项，其具体表达式为

$$\begin{aligned} \mathbf{J}_n &= \mathbb{E}_{\mathbf{H}_{n,e}} \left[\sum_{i=1}^L \mathbf{H}_{n,e} \mathbf{v}_{n,i} \mathbf{v}_{n,i}^H \mathbf{H}_{n,e}^H \right] \\ &= \sum_{i=1}^L (\mathbf{v}_{n,i}^T \otimes \mathbf{I}_{N_r}) \mathbb{E} \left[\text{vec}(\mathbf{H}_{n,e}) \cdot \text{vec}(\mathbf{H}_{n,e})^H \right] (\mathbf{v}_{n,i}^T \otimes \mathbf{I}_{N_r})^H \\ &= \sum_{i=1}^L (\mathbf{v}_{n,i}^T \otimes \mathbf{I}_{N_r}) \mathbf{R}_n (\mathbf{v}_{n,i}^T \otimes \mathbf{I}_{N_r})^H, \end{aligned} \quad (4.14)$$

其中符号 \otimes 表示克罗内克乘积。

基于此，在非理想CSI场景下，我们利用香农容量公式 $r_{n,l} = \log(1 + \gamma_{n,l})$ 将问题 \mathcal{P}'_B 进一步改写并简化为

$$\begin{aligned} \mathcal{P}_E : \min_{\{\mathbf{V}\}} \psi(\mathbf{r}) &= \sum_{n=1}^N \sum_{l=1}^L \exp(-e^{r_{n,l}}/\lambda) \\ \text{s.t.} \quad \sum_{n=1}^N \text{Tr}(\mathbf{V}_n^H \mathbf{V}_n) &\leq P, \end{aligned} \quad (4.15)$$

其中 $\mathbf{r} = [r_{1,1}, \dots, r_{n,l}, \dots, r_{N,L}]^T \in \mathbb{R}^{NL \times 1}$ 表示各流信道的平均容量， $r_{n,l}$ 由式(4.13)和(4.14)给出。

上式仍然是一个复杂的非凸问题，难以直接求解。下面，我们通过将原问题目标函数逐步泰勒展开，设计一种迭代求解方法。具体来说，假设在第 t 次迭代时，我们得到原问题的一组解为 $\mathbf{r}^t = [r_{1,1}^t, \dots, r_{n,l}^t, \dots, r_{N,L}^t]^T$ ，将目标函数在该点泰勒展开后得到

$$\psi(\mathbf{r}) \approx \psi(\mathbf{r}^t) + \nabla_{\mathbf{r}} \psi(\mathbf{r}^t)^T \cdot (\mathbf{r} - \mathbf{r}^t), \quad (4.16)$$

其中， $\nabla_{\mathbf{r}} \psi(\mathbf{r}^t)^T = [-\rho_{1,1}, \dots, -\rho_{n,l}, \dots, -\rho_{N,L}] \in \mathbb{R}^{1 \times NL}$ 表示目标函数在解 \mathbf{r}^t 处的梯度值（为了书写简洁，我们在这里和下面推导中暂时省略上标 t ），且

$$\rho_{n,l} = (e^{r_{n,l}}/\lambda) \cdot \exp(-e^{r_{n,l}}/\lambda). \quad (4.17)$$

将上述展开代入式(4.15)后，我们得到在该步迭代中需要求解的子问题，可表示为

$$\begin{aligned} \max_{\{\mathbf{V}\}} \quad & \sum_{n=1}^N \sum_{l=1}^L \rho_{n,l} r_{n,l} \\ \text{s.t.} \quad & \sum_{n=1}^N \text{Tr}(\mathbf{V}_n^H \mathbf{V}_n) \leq P, \end{aligned} \quad (4.18)$$

这是一个形如加权和速率最大化的问题。由文献^[101]可知，它和另一类称为WMMSE的问题等价。基于此，我们终于得到所需求解的转化问题 $\mathcal{P}_{E,t}$ ，其数学表达式为

$$\begin{aligned} \mathcal{P}_{E,t} : \quad & \max_{\{\mathbf{V}, \mathbf{U}, w_{n,l}\}} \sum_{n=1}^N \sum_{l=1}^L \rho_{n,l} (w_{n,l} e_{n,l} - \log w_{n,l}) \\ \text{s.t.} \quad & \sum_{n=1}^N \text{Tr}(\mathbf{V}_n^H \mathbf{V}_n) \leq P, \end{aligned} \quad (4.19)$$

其中 $w_{n,l} > 0$ 表示一个正的权重矢量， $e_{n,l}$ 定义为 (n, l) 子信道上接收信号的MSE，即

$$\begin{aligned} e_{n,l} & \triangleq \mathbb{E}_{\mathbf{s}_n, \mathbf{n}, \mathbf{H}_n, e} [(\hat{s}_{n,l} - s_{n,l})(\hat{s}_{n,l} - s_{n,l})^H] \\ & = \mathbb{E}_{\mathbf{H}_n, e} \left[(1 - \mathbf{u}_{n,l}^H \mathbf{H}_n \mathbf{v}_{n,l}) (1 - \mathbf{u}_{n,l}^H \mathbf{H}_n \mathbf{v}_{n,l})^H + \sum_{j=1, j \neq l}^L \mathbf{u}_{n,l}^H \mathbf{H}_n \mathbf{v}_{n,j} \mathbf{v}_{n,j}^H \mathbf{H}_n^H \mathbf{u}_{n,l} + \sigma_n^2 \mathbf{u}_{n,l}^H \mathbf{u}_{n,l} \right] \\ & = (1 - \mathbf{u}_{n,l}^H \hat{\mathbf{H}}_n \mathbf{v}_{n,l}) (1 - \mathbf{u}_{n,l}^H \hat{\mathbf{H}}_n \mathbf{v}_{n,l})^H + \mathbf{u}_{n,l}^H \left(\sum_{j=1, j \neq l}^L \hat{\mathbf{H}}_n \mathbf{v}_{n,j} \mathbf{v}_{n,j}^H \hat{\mathbf{H}}_n^H + \mathbf{J}_n + \sigma_n^2 \mathbf{I}_{N_r} \right) \mathbf{u}_{n,l}. \end{aligned} \quad (4.20)$$

对上述问题，我们遵循文献^[101]所提出的基于块坐标下降（Block coordinate descent, BCD）的WMMSE迭代算法框架，逐步更新 $w_{n,l}$ 、 $\mathbf{v}_{n,l}$ 和 $\mathbf{u}_{n,l}$ 。具体来说，算法分为以下三个步骤：

(1) 步骤1: 更新 $w_{n,l}$

在给定 $\mathbf{v}_{n,l}$ 和 $\mathbf{u}_{n,l}$ 时, $e_{n,l}$ 也由式(4.20)确定。此时容易看出, $\mathcal{P}'_{B,t}$ 关于 $w_{n,l}$ 的子问题是一个无约束凸优化问题。直接应用一阶最优性条件, 可得:

$$w_{n,l}^{\text{opt}} = e_{n,l}^{-1}, \forall n, l. \quad (4.21)$$

(2) 步骤2: 更新 $\mathbf{v}_{n,l}$

在给定 $\mathbf{u}_{n,l}$ 和 $w_{n,l}$ 时, 关于 $\mathbf{v}_{n,l}$ 的优化子问题如下

$$\begin{aligned} \max_{\{\mathbf{V}\}} & \sum_{n=1}^N \sum_{l=1}^L \rho_{n,l} w_{n,l} (e_{n,l}) \\ \text{s.t.} & \sum_{n=1}^N \text{Tr}(\mathbf{V}_n^H \mathbf{V}_n) \leq P, \end{aligned} \quad (4.22)$$

其中 $e_{n,l}$ 由公式(4.20)给出。这是一个凸问题, 可使用经典的凸优化方法求解。根据一阶最优性条件, $\mathbf{v}_{n,l}$ 具有如下闭式解的形式

$$\mathbf{v}_{n,l}^{\text{opt}} = \left(\sum_{l=1}^L \rho_{n,l} w_{n,l} \cdot \hat{\mathbf{H}}_n^H \mathbf{u}_{n,l} \mathbf{u}_{n,l}^H \hat{\mathbf{H}}_n + \mathbf{K}_n + \mu \mathbf{I} \right)^{-1} \hat{\mathbf{H}}_n^H \mathbf{u}_{n,l} \rho_{n,l} w_{n,l}, \forall n, l, \quad (4.23)$$

其中 $\mu \geq 0$ 表示功率控制参数, 可以通过二分搜索法确定以使得功率约束条件满足; \mathbf{K}_n 与 \mathbf{J}_n 有类似的形式, 即

$$\mathbf{K}_n = \sum_{i=1}^L \rho_{n,l} w_{n,l} (\mathbf{u}_{n,i}^T \otimes \mathbf{I}_{N_t}) \tilde{\mathbf{R}}_n (\mathbf{u}_{n,i}^T \otimes \mathbf{I}_{N_t})^H, \quad (4.24)$$

注意这里 $\tilde{\mathbf{R}}_n \triangleq \mathbb{E}[\text{vec}(\mathbf{H}_{n,e}^H) \cdot \text{vec}(\mathbf{H}_{n,e})] \in \mathbb{C}^{N_t N_r \times N_t N_r}$ 与 \mathbf{R}_n 有细微区别, 但可由后者通过简单行列变换得到。

(3) 步骤3: 更新 $\mathbf{u}_{n,l}$

最后, 在给定 $\mathbf{v}_{n,l}$ 和 $w_{n,l}$ 时, 使得 $e_{n,l}$ 最小的即为经典的MMSE接收机, 即

$$\mathbf{u}_{n,l}^{\text{mmse}} = \left(\hat{\mathbf{H}}_n \mathbf{V}_n \mathbf{V}_n^H \hat{\mathbf{H}}_n^H + \mathbf{J}_n + \sigma_n^2 \mathbf{I} \right)^{-1} \hat{\mathbf{H}}_n \mathbf{v}_{n,l}, \forall n, l. \quad (4.25)$$

基于上述三个步骤循环迭代计算更新 $w_{n,l}$ 、 $\mathbf{v}_{n,l}$ 和 $\mathbf{u}_{n,l}$, 算法收敛后, 即可找到问题 $\mathcal{P}_{E,t}$ 的一组驻点解 (Stationary-solution) $\{w_{n,l}^\circ, \mathbf{v}_{n,l}^\circ, \mathbf{u}_{n,l}^\circ\}$ 。

之后根据文献^[101], 计算此时各流子信道的容量, 即 $\mathbf{r}^\circ = [r_{1,1}^\circ, \dots, r_{n,l}^\circ, \dots, r_{N,L}^\circ]^T$, 其中 $r_{n,l}^\circ = \log \left[\left(e_{n,l}^{\text{mmse}} \right)^{-1} \right] = \log w_{n,l}^\circ$, 并按照

$$\mathbf{r}^{t+1} = \mathbf{r}^t + \alpha_t (\mathbf{r}^\circ - \mathbf{r}^t) \quad (4.26)$$

更新 \mathbf{r}^{t+1} 。其中 $\{\alpha_t\}$ 表示一个大于零的更新步长，可基于经典的非精确线搜索方法（如Armijo-Goldstein准则^[109]等）确定。

最后，重新计算目标函数 $\psi(\mathbf{r})$ 在新速率点 \mathbf{r}^{t+1} 处泰勒展开表达式，并按照式(4.19)构建第 $t+1$ 次迭代的新的子问题 $\mathcal{P}_{E,t+1}$ ，之后重复步骤1至3……反复执行上述过程直到结果收敛，最终即可求得原问题 \mathcal{P}_E 的局部最优解。这里，我们不妨再回顾3.3.1小节中所提出的基于SVD的功率权值优化方案。虽然它只是针对问题 \mathcal{P}'_B 设计的一种近似解法，但可将其结果作为E-S-WMMSE算法的一组较好的初始点，从而降低后续算法的迭代次数。

至此，我们已经完整介绍了所提出的E-S-WMMSE算法，为了更加清晰的表述上述算法的实施过程，将其总结如下。

算法 3 E-S-WMMSE 算法

输入: 估计信道 $\hat{\mathbf{H}}_n$ ；信道误差协方差矩阵 \mathbf{R}_n ；学习步长 $\{\alpha_t\}$ ；

初始化: 通过求解式(4.10)所表述简单问题或基于其他先验信息获得 \mathbf{r}^0

for $t = 1, 2, \dots$ **do**

 将 \mathbf{r}^t 代入式(4.17)并基于式(4.19)构造WMMSE问题 $\mathcal{P}_{E,t}$

repeat

 基于式(4.21)更新 $w_{n,l}, \forall n, l$

 基于式(4.23)更新 $\mathbf{v}_{n,l}, \forall n, l$

 基于式(4.25)更新 $\mathbf{u}_{n,l}, \forall n, l$

until WMMSE算法收敛

 计算 \mathbf{r}° 、搜索步长 α_t 并根据式(4.26)更新 \mathbf{r}^{t+1}

end for

输出: 预编码 \mathbf{V}

4.3.3 算法收敛性分析

本小节将分析E-S-WMMSE算法的收敛性，并证明算法最终将收敛于问题 \mathcal{P}_E 的一个驻点解，即满足卡罗需-库恩-塔克（Karush-Kuhn-Tucker, KKT）条件。

定理 1. (E-S-WEESM算法的收敛性定则): E-S-WMMSE算法产生的任何一组极限点 (limiting-point) $\{\mathbf{r}^*, \mathbf{V}^*\}$ 都是原问题 \mathcal{P}_E 的一个驻点解。

上述定理是比较直观的，我们在此即给出证明。具体来说，为了证明上述定理，我们

需要说明两件事：1) 算法可以收敛到可行集内的一组极限点 $\{\mathbf{r}^*, \mathbf{V}^*\}$ ；2) 极限点 $\{\mathbf{r}^*, \mathbf{V}^*\}$ 满足问题 \mathcal{P}_E 的KKT条件。

一方面，根据 \mathbf{r} 的取值范围和式(4.17)中对原问题目标函数梯度的表述，不难发现 $\psi(\mathbf{r})$ 关于变量 \mathbf{r} 的偏导是李普希茨连续的 (Lipschitz-continuous-gradient)。因此，在采用基于Armijo-Goldstein准则的线搜索方法来确定步长 α_t 时，E-S-WMMSE算法总可以收敛至 \mathcal{P}_E 的一组极限点 $\{\mathbf{r}^*, \mathbf{V}^*\}$ ^[109]。

另一方面，对任意一组极限点 $\{\mathbf{r}^*, \mathbf{V}^*\}$ ，应用链式求导法则 (Chain-rule)，有

$$\frac{\partial \psi(\mathbf{r}^*)}{\partial \mathbf{v}_{n,l}} = \sum_{j=1}^L \frac{\partial \psi(\mathbf{r}^*)}{\partial r_{n,j}} \cdot \frac{\partial r_{n,j}(\mathbf{V}_n^*)}{\partial \mathbf{v}_{n,l}} = - \frac{\partial \left[\sum_{j=1}^L \rho_{n,j}^* r_{n,j}(\mathbf{V}_n^*) \right]}{\partial \mathbf{v}_{n,l}} \quad (4.27)$$

成立。而这说明了这样一个事实：问题 \mathcal{P}_E 在极限点 $\{\mathbf{r}^*, \mathbf{V}^*\}$ 处的KKT条件与权重为 $\rho_{n,l}^*, \forall n, l$ 的加权和速率问题(4.18)在极限点 $\{\mathbf{V}^*\}$ 处的KKT条件是等价的。文献^[101]已证明了WMMSE算法可以求得问题(4.18)的一组驻点解。因此在极限点 $\{\mathbf{r}^*, \mathbf{V}^*\}$ 处， \mathcal{P}_E 同样满足KKT条件，上述定理得证。

4.3.4 算法复杂度分析

本小节中，我们将分析E-S-WMMSE算法的计算复杂度。假设我们先通过求解式(4.10)所表述的简单问题获得一组初始解，之后开始逐级构建并求解WMMSE问题 $\mathcal{P}_{E,t}$ ，算法总体的计算复杂度则由这两部分构成。

对问题(4.10)，我们利用凸优化工具箱CVX的内点法求解。其复杂度是关于自变量数目的多项式函数，即需要约 $C_1 \triangleq O(N^{3.5}L^{3.5})$ 数量级的浮点数运算。对每次构建的WMMSE问题 $\mathcal{P}'_{B,t}$ ，其计算复杂度主要来源于 $w_{n,l}$ 、 $\mathbf{v}_{n,l}$ 和 $\mathbf{u}_{n,l}$ 的迭代更新计算（基于Armijo-Goldstein等准则搜索步长 α 的复杂度较之可以忽略不计）。这其中，式(4.23)和(4.25)中所涉及的矩阵求逆运算和包括矩阵克罗内克积在内的一系列矩阵乘法运算又占主导。具体来说，对 $\mathbf{v}_{n,l}, \forall n, l$ 的每次更新，需要 $C_2 \triangleq O(N \cdot N_t^3 + NL \cdot N_t^2 N_r)$ 量级的浮点数运算；对 $\mathbf{u}_{n,l}, \forall n, l$ 的每次更新，则需要 $C_3 \triangleq O(N \cdot N_r^3 + NL \cdot N_r^2 N_t)$ 量级。

综上分析，E-S-WMMSE算法所涉及的总体浮点数运算约为 $C \triangleq C_1 + I_1 I_2 (C_2 + C_3)$ 数量级，其中 I_1 和 I_2 分别表示算法3中两层循环的迭代次数。可以看出，在问题规模 N 、 L 、 N_t 和 N_r 既定时，选取好的迭代起始点可以有效地降低算法后续的计算复杂度。

4.4 仿真结果与分析

本章节，我们将搭建符合LTE规范的链路级仿真系统来比较所提出的联合空频域E-S-

WMMSE预编码方案和现有的几种功率分配基线方案在低MCS下的BLER性能。信道模型采用现有3GPP协议规定的CDL-B信道^[88]，部署为均匀平面天线阵列（Uniform planar array, UPA）。其他系统参数总结于表(4.1)中。

表 4.1 系统和信道仿真参数设置

名称	设置
信道模型	CDL-B
用户移动速度	3km/h
发送天线数	32
发送天线阵列	水平8-垂直4
接收天线数	4
天线是否极化	发送正交极化
载波频率	3GHz
DS	1us
RB数目	120
子帧数目	500
子载波间隔	30kHz
MCS	1, 3
用户传输流数	2, 4

对于基线预编码方案，我们假设它们都是采取4.3.1章节中提到的基于估计信道SVD分解和功率权值分配的预编码方案。具体地，本章采取最常使用的等功率分配（Equal power allocation, EPA）和使得互信息最大的注水（WF）法作为两种主要基线方案。同时为了更细致的比较，我们还采取了一种称为等互信息（Equal mutual information, EMI）的功率权值分配方案。EMI的思想与WF正好相反，它倾向于为质量较差（好）的子信道分配更多（少）的功率，从而保证式(4.3)中各子信道的接收信噪比完全相同，即 $\gamma_{n,l} = \gamma_{i,j}, \forall (n,l) \neq (i,j)$ 。

4.4.1 蒙特卡洛仿真测试

在正式比较所提方案与基线的BLER性能之前，我们先通过蒙特卡洛仿真实验验证本章节提出的基于EESM准则优化BLER性能的合理性。具体来说，我们首先随机产生若干组组合系数 $[\alpha_1, \alpha_2, \alpha_3]$ ，将上述提到的三种基线功率分配方案 \mathbf{P}_{EPA} 、 \mathbf{P}_{WF} 和 \mathbf{P}_{EMI} 线性组合得到的新的混合功率分配（Hybrid power allocation, HPA），即

$$\mathbf{P}_{HPA}^m = \alpha_1^m \mathbf{P}_{EPA} + \alpha_2^m \mathbf{P}_{WF} + \alpha_3^m \mathbf{P}_{EMI}, \quad (4.28)$$

其中上标 m 表示实验序号。对任意一次实验 m ，都有 $\alpha_i^m \geq 0, i = 1, 2, 3$ 和 $\alpha_1^m + \alpha_2^m + \alpha_3^m = 1$ 。

之后，我们计算每次HPA方案在式(4.10)所表述的EESM准则下的目标函数值 J ，并比较它和实际BLER仿真结果之间的对应关系。如表(4.2)和图(4.2)所示，它们分别记录了不同发送信噪比下10次典型的混合功率分配实验的EESM目标函数值和BLER结果。

表 4.2 各次混合功率分配方案的EESM目标函数值

混合功率分配实验	-3.6dB时的 J	-3.4dB时的 J	-3.2dB时的 J	-3.0dB时的 J
HPA1	0.7745	0.7652	0.7564	0.7471
HPA2	0.7978	0.7885	0.7792	0.7694
HPA3	0.8248	0.8155	0.8062	0.7964
HPA4	0.8490	0.8397	0.8304	0.8211
HPA5	0.8755	0.8671	0.8583	0.8490
HPA6	0.9015	0.8932	0.8848	0.8760
HPA7	0.9243	0.9164	0.9085	0.9006
HPA8	0.9588	0.9523	0.9457	0.9388
HPA9	0.9857	0.9802	0.9746	0.9685
HPA10	1.0164	1.0118	1.0071	1.0025

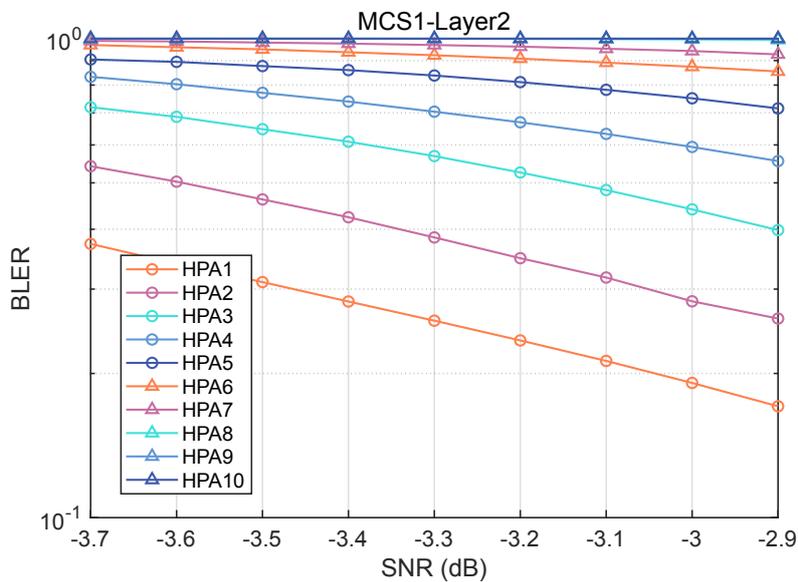


图 4.2 各次混合功率分配方案的BLER性能

可以看出，对不同混合功率分配方案，EESM目标值与BLER结果有很强相关性。具体表现为， J 值越低的混合功率分配方案，实际BLER性能越优。这说明EESM准则在此种信道仿真条件下是BLER指标的一个合适的度量准则。4.2章节提出的基于EESM目标函数最

小化优化得到的功率分配或预编码结果应该是该信道、传输条件下的近似优异解。

4.4.2 理想CSI场景下BLER性能比较

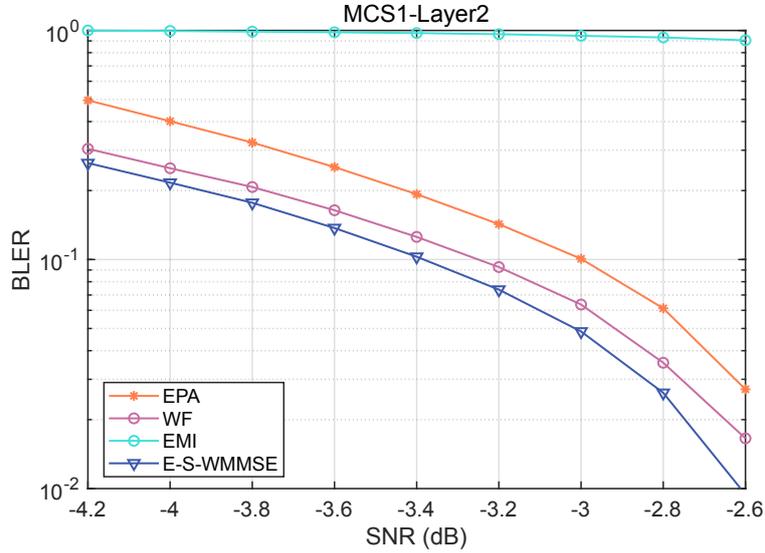


图 4.3 双流MCS1传输时各方案的BLER性能

图(4.3)比较了理想CSI场景下，在MCS为1、两流传输下各功率分配方案的BLER性能。此时可以看到，在三种基线方案中，WF性能最优秀，EPA和EMI表现较差。在BLER为0.1左右时，所提E-S-WMMSE方案相较最优基线WF领先发送功率增益约0.2dB。

同时，在发送功率为-3.4dB时，四种方案关于式(4.10)的目标函数值分别为 $J_{EMI}=0.8568$ 、 $J_{EPA}=0.7478$ 、 $J_{WF}=0.6875$ 和 $J_{E-S-WMMSE}=0.6827$ 。也可以发现其大小关系与BLER性能的对应相关性，即目标函数值 J 越大的功率分配方案对应BLER性能越差。

4.4.3 非理想CSI场景下BLER性能比较

图(4.4)比较了在BS端信道误差MSE (Channel error MSE, CEM) 为5dB时的各方案BLER性能。这里CEM定义为 $10 \log_{10} (\mathbb{E} (\|\mathbf{H}_n\|_F^2) / \mathbb{E} (\|\mathbf{H}_{n,e}\|_F^2))$ ，其中下标 F 表示矩阵Frobenius范数。可以看出，在MSE为5dB时，E-S-WMMSE方案相较最优WF基线方案的增益也大约为0.2dB，与理想CSI场景下几乎一致。这表明BS端信道的理想与否可能不是限制此场景下增益空间的主要原因。另外，同时比较图(4.4)和图(4.3)，可发现虽然各种方案在非理想CSI下的BLER性能均有一定的劣化，但所造成的等效发送功率损失并不大，仅有1dB左右。其原因即是前文所提到的低MCS和码率下BLER-SNR曲线展现的“瀑布效应”。

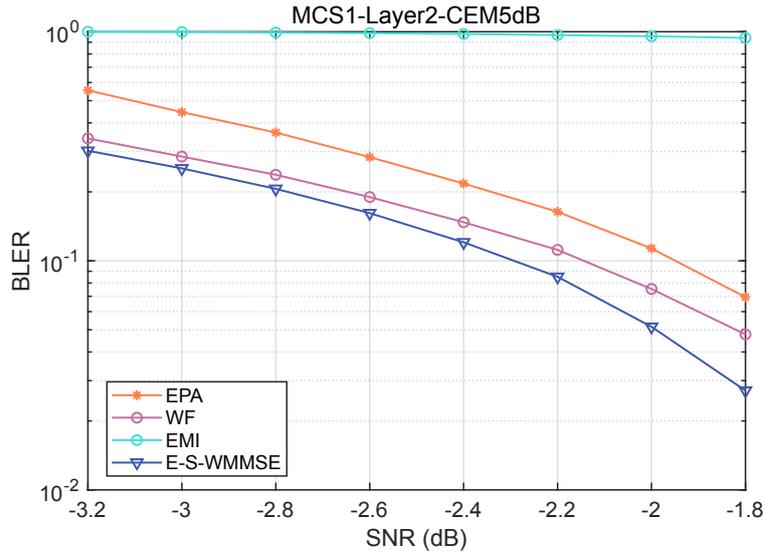


图 4.4 信道CEM5dB且双流MCS1传输时各方案的BLER性能

4.4.4 其他场景下增益空间的探索

为了进一步探索所提方案相较基线可能实现的BLER性能增益，我们在理想CSI场景下又分别增大发送数据流和MCS阶数后进行测试。BLER性能结果如图(4.5)和图(4.6)所示。

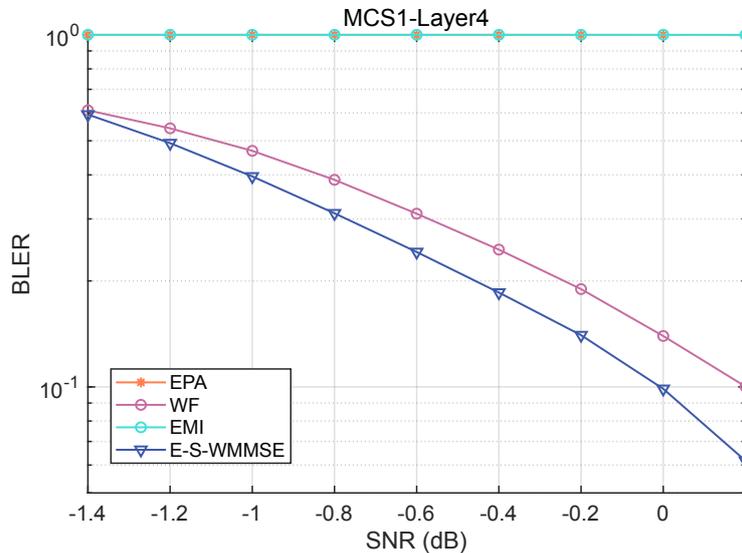


图 4.5 四流MCS1传输时各方案的BLER性能

可以看到，在增大到四流并发传输后，EPA和EMI基线方案的性能恶化非常严重，只有WF基线方案继续保持较好的性能。但此时所提方案相较WF的BLER性能增益几乎不变，仍只有0.2dB左右。但当MCS增大为3后，性能增益空间变大，达到了0.4dB。这间接说明了限制E-S-WMMSE算法取得增益的主要原因是较低的MCS。这也意味着，在本章考虑的信

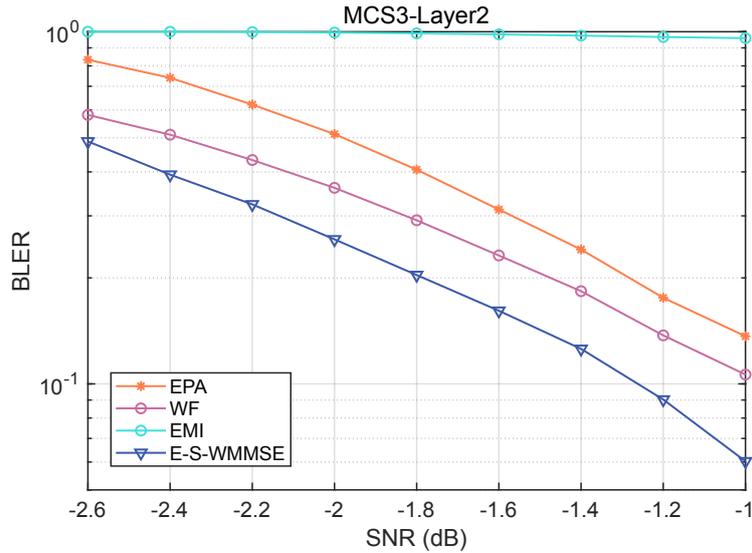


图 4.6 双流MCS3传输时各方案的BLER性能

道条件和低MCS传输场景下，相较WF基线可挖掘的BLER性能增益空间已经不大了，可能至多也只有0.2dB至0.4dB的发送功率增益提升。

4.5 本章小结

本章中，我们考虑了一种称为远距覆盖增强通信场景下的下行预编码设计问题。该类场景重点关心如何在低MCS下设计以BLER为性能指标的下行预编码方案。由于BLER是一个与MCS、码率和信道编码方式等诸多实际链路参数有关的复杂函数，因此原编码系统BLER优化问题是极难直接解决的。为了克服这个困难，我们引入链路分析和评估中常用的ESM方法，首先基于EESM准则将原本复杂的编码系统BLER优化问题转化为无编码系统中针对所选度量函数的优化问题。对转化后仍然非凸的问题，我们在非理想CSI场景下设计了一种基于逐级迭代的E-S-WMMSE算法来有效寻找目标函数的局部最优解。E-S-WMMSE算法通过将转化问题泰勒展开，在每一次迭代中构造一个形如加权和速率最大化的子问题，并利用经典的WMMSE算法求解。另外，在理想CSI下，我们也设计了只考虑功率权值优化的低复杂度SVD预编码方案。为了验证方案的合理性，我们搭建了符合协议标准的链路级仿真平台，首先基于蒙特卡洛实验证明了EESM度量准则对刻画此场景下BLER性能的准确性。之后，我们在理想CSI和非理想CSI两种场景下比较了所提方案和几种基线方案的BLER性能，并通过改变传输流数目和MCS阶数进一步探索了所考虑场景下的增益空间。仿真结果既证明了所提E-S-WMMSE算法的有效性，同时也为工业界评估此场景下可能的性能增益提供了详实的理论和数据指导。

5 受约束深度强化学习框架下的时延感知功率控制策略研究

5.1 研究动机和主要成果

对于MU-MIMO的下行传输系统，虽然下行预编码或功率分配的主要设计目标是为了获得更高的频谱效率以最大化小区容量。但在5G和B5G系统中，仍有一类特殊业务对传输时延和服务的可靠性非常敏感。这类业务场景称为URLLC，包括如交通安全、自动控制、物联网（Internet of things, IoT）和移动边缘计算（Mobile edge computing, MEC）等诸多服务实例^[110]。因此，为了支持这些时延感知（Delay-aware）敏感的应用服务，考虑时延约束下最优的功率控制技术是至关重要的。

然而，目前的大部分工作基本都只考虑了短期（Short-term）时延约束下的功率控制或资源管理策略，即假设系统会在一个信道相关时间（Channel-coherence-time）间隔内完成所有的数据传输^[111,112]。但在实际系统中，由于经常发生的信道深衰落（Deep-fading）效应和有限的计算资源，这种假设在实际中通常是很难满足的。因此，考虑一个系统的长期（Long-term）平均时延性能对于评估5G/B5G场景下的诸多低时延应用是至关重要的。对这类问题，文献^[113]采取了一种李雅普诺夫（Lyapunov）优化方案来解决系统的长期时延约束问题，但其只能保证传输队列的稳定性。文献^[114]通过将动态规划（Dynamic programming, DP）中的相对值迭代和一个虚拟连续时间系统相结合，提出了一种时延约束的传输策略。但这种方案高度依赖其考虑问题的特殊结构，难以推广至一般问题。

综合来看，这类长期时延约束下的功率控制策略设计问题困难的原因主要来源于以下两个方面：

- (1) **问题的马尔可夫性：**这类问题涉及到一个受约束的马尔可夫决策过程（Constrained Markov decision process, CMDP）^[115]。由于无法获得整个时间尺度上全部信道状态信息和数据到达速率信息，而且当前的决策会影响未来系统的状态，因此传统的基于DP的方法实际上是无法有效求解的^[116]。
- (2) **状态和求解空间维度大：**近些年，虽然受约束强化学习（Constrained deep reinforcement learning, CDRL）技术的巨大进步为解决CMDP问题提供了可行的指导方

法，但现有的基于DNN的主流CRL方法，如受约束决策优化（Constrained policy optimization, CPO）^[117]等，都几乎无法适用于实际系统。主要的原因是这些方法的在线更新往往需要大量的实时数据样本和较高的迭代计算开销，而这在实际通信系统中几乎是无法实现的。

因此在本章中，我们将考虑了一个包含动态队列的下行MU-MIMO系统，设计了一种基于RZF预编码方法的功率控制策略^[92]，用以解决在满足各用户长期平均时延约束下的长期平均发送功率最小化的问题。为了解决传统的CDRL算法更新计算复杂度高、对在线样本需求量大等缺陷，我们提出了一种称为逐级凸近似策略优化（Successive convex approximation policy optimization, SCAPO）的CDRL算法框架。所提算法通过将原问题中复杂的目标和约束函数转化为其对应的凸代理（Surrogate）函数，从而在每次迭代中只需求解一个凸优化问题，降低了更新计算复杂度。同时，在每次更新中，所提算法可以重用旧的样本数据，只需要少数的在线新样本即可完成训练，因而是数据有效（Data-efficiency）的。需要指出，我们在本章中通过假设RZF预编码从而只考虑优化功率权值的分配策略，这不仅是为了在性能和计算复杂度中进行权衡，更重要的是为了实现更快的网络收敛速度，这对实际通信系统是至关重要的。

本章的剩余结构安排如下：第二子章节将介绍系统模型和所考虑的CMDP问题；第三子章节会详细介绍所提出的SCAPO算法；第四子章节将基于仿真实验比较所提算法和几种基线CDRL方案在功率消耗、长期时延和收敛速度等性能上的优劣；最后总结本章内容。

5.2 系统模型和问题建模

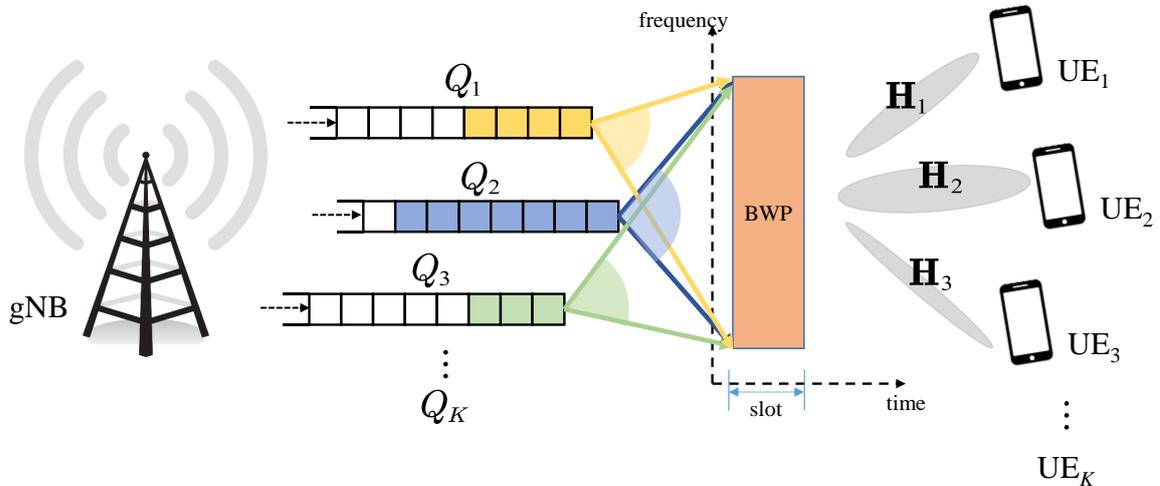


图 5.1 包含动态队列的下行MU-MIMO系统

5.2.1 MU-MIMO系统传输模型

如图(5.1)所示, 本章节我们将考虑一个MU-MIMO下行宽带传输系统。部署 N_t 根天线的中心BS将为 K 个单天线用户提供传输服务, 记用户集合为 \mathcal{K} 。系统可用的带宽记为 B 。为了方便说明, 我们假设信道具有块衰落特性, 且在每个TTI内保持不变。更具体的, 对用户 k 来说, 假设其在第 n 个载波上的信道 $\mathbf{h}_{k,n}^H \in \mathbb{C}^{1 \times N_t}$ 符合广泛使用的Saleh-Valenzuela多径信道模型^[90], 即

$$\mathbf{h}_{k,n}^T = \sqrt{\frac{N_t}{M_k}} \sum_{m=1}^{M_k} \beta_{m,k} e^{-j2\pi\tau_{m,k}f_n} \mathbf{a}_k(\varphi_{m,n,k}), \quad (5.1)$$

其中 M_k 是用户 k 信道中可分辨的径数; $\beta_{m,k}$ 和 $\tau_{m,k}$ 分别为第 m 根径的复增益和时延; $\varphi_{m,n,k}$ 定义为

$$\varphi_{m,n,k} = \frac{f_n}{c} d \sin \theta_m, \quad (5.2)$$

表示第 n 个载波的空间指向。上式中, θ_m 表示第 m 根径的离开角信息; f_n 表示第 n 个子载波的频率; d 定义为天线间隔, 实际中一般为载波波长的的一半, 即 $d = c/2f_c$ (c 表示光速)。在本章中, 我们假设系统是窄带且信道是平衰落的 (Flat-fading), 即 $B \ll f_c$, f_n 与中心频率 f_c 近似相同。因此我们省略表示子载波的下标 n , 有 $\varphi_{m,k} \approx \frac{1}{2} \sin \theta_{m,k}$ 。最后, $\mathbf{a}_k(\varphi_{m,k})$ 是与 $\varphi_{m,k}$ 对应的阵列响应矢量。若假设天线阵列是ULA的, 则有 $\mathbf{a}_k(\varphi_{m,k}) = \frac{1}{\sqrt{N_t}} e^{-j2\pi\varphi_{m,k}\mathbf{p}_{a_k}}$, 其中 $\mathbf{p}_{a_k} = [-\frac{N_t-1}{2}, -\frac{N_t+1}{2}, \dots, \frac{N_t-1}{2}]$ ^[91]。

因此记 $\mathbf{H} = [\mathbf{h}_1, \dots, \mathbf{h}_K]^H \in \mathbb{C}^{K \times N_t}$ 表示系统的全局CSI, 且由BS完美已知。在下行传输阶段, BS采取基于RZF的预编码方法^[92]。具体地, 定义 $\mathbf{V} = [\mathbf{v}_1, \dots, \mathbf{v}_K] \in \mathbb{C}^{N_t \times K}$ 表示发端预编码矩阵, 其数学表达式为

$$\mathbf{V} = \mathbf{H}^H (\mathbf{H}\mathbf{H}^H + \alpha_Z \mathbf{I})^{-1} \mathbf{\Lambda}^{1/2}, \quad (5.3)$$

其中 α_Z 是一个正则化因子; $\mathbf{\Lambda}^{1/2} = \text{diag}(\|\bar{\mathbf{v}}_1\|^{-1}, \dots, \|\bar{\mathbf{v}}_K\|^{-1})$ 是一个能量归一化对角矩阵, $\bar{\mathbf{v}}_k$ 表示矩阵 $\bar{\mathbf{V}} \triangleq \mathbf{H}^H (\mathbf{H}\mathbf{H}^H + \alpha_Z \mathbf{I})^{-1}$ 的第 k 列。因此, 用户 k 的下行传输速率可以表示为

$$R_k = B \log_2 \left(1 + \frac{p_k |\mathbf{h}_k^H \mathbf{v}_k|^2}{\sum_{j \neq k} p_j |\mathbf{h}_k^H \mathbf{v}_j|^2 + \sigma_k^2} \right), \forall k, \quad (5.4)$$

其中 p_k 、 $\mathbf{v}_k \in \mathbb{C}^{N_t \times 1}$ 和 σ_k^2 分别表示各用户的发送功率、预编码矢量和噪声能量。

5.2.2 动态队列模型

在本章考虑的系统, 定义系统的最小的时间粒度是持续时间为 τ 的1个时隙 (Slot)。在第 t 个slot内, BS为所有用户维持一个动态数据队列 $Q_k(t), \forall k \in \mathcal{K}$ 用以处理各种突发流量

业务。记 $\mathbf{Q}(t) = [Q_1(t), \dots, Q_K(t)]^T \in \mathbb{R}^K$ 为系统的全局队列状态信息 (Queue state information, QSI) 也由BS完美已知。

对用户 k 来说, 其数据队列的 $Q_k(t)$ 的动态变化模型为

$$Q_k(t) = [Q_k(t-1) + A_k(t)\tau - R_k(t)\tau]^+, \forall k \in \mathcal{K}, \quad (5.5)$$

其中 $[x]^+ \triangleq \max\{0, x\}$; $A_k(t)$ 表示第 t 个 slot 内用户 k 的数据包到达速率。本章中, 假设它服从平均到达速率为 λ_k 的经典泊松到达分布, 即 $\mathbb{E}[A_k] = \lambda_k$ 。根据 *Little's Law* 原理^[118], 此时用户 k 的传输时延可以表示为 $Q_k(t)/\lambda_k$ 。

5.2.3 功率控制策略和CMDP问题建模

在本章中, 我们的目标是获得一个控制策略, 用以解决在满足各用户长期平均时延约束下的长期平均发送功率最小化的问题, 如图(5.2)所示。控制策略的输入是每个 slot 系统的全局状态 (State) 信息 $s_t = \{\mathbf{Q}(t), \mathbf{H}(t)\}$, 输出的动作 (Action) 是包含功率分配和正则化因子的控制策略 $a_t = \{\mathbf{p}(t) \triangleq [p_1(t), \dots, p_K(t)]^T \in \mathbb{R}^K, \alpha_Z(t)\}$, 即

$$a_t = \pi_\theta(s_t), \quad (5.6)$$

这里下标 θ 表示 π 是由参数 θ (如DNN网络参数) 所控制的参数化策略。其中状态和动作空间维度分别为

$$\begin{aligned} \dim(s_t) &= \underbrace{K}_{\text{QSI}} + \underbrace{2KN_t}_{\text{CSI}}, \\ \dim(a_t) &= \underbrace{1}_{\alpha_Z} + \underbrace{K}_{p_k}. \end{aligned} \quad (5.7)$$

在第 t 个 slot, 定义 $C_k(s_t, a_t) = Q_k(t)/\lambda_k, k = 1, \dots, K$ 表示各用户的传输时延; $C_0(s_t, a_t) = \sum_{k \in \mathcal{K}} p_k(t)$ 表示系统的发送功率消耗; 因此, 我们的目标是通过求解形如下式CMDP问题来找到合适的控制参量 θ , 即

$$\begin{aligned} \mathcal{P}_C : \min_{\theta \in \Theta} J_0(\theta) &\triangleq \lim_{T \rightarrow \infty} \frac{1}{T} \mathbb{E} \left[\sum_{t=1}^T \sum_{k \in \mathcal{K}} p_k(t) \right] \\ \text{s.t. } J_k(\theta) &\triangleq \lim_{T \rightarrow \infty} \frac{1}{T} \mathbb{E} \left[\sum_{t=1}^T \frac{Q_k(t)}{\lambda_k} \right] - c_k \leq 0, \forall k \in \mathcal{K}, \end{aligned} \quad (5.8)$$

其中 $\Theta \subseteq \mathbb{R}^{n_\theta}$ 表示参数空间, n_θ 是其维度; c_1, \dots, c_K 表示各用户能容许的平均最大等待时延。

一般来说, 问题 \mathcal{P}_C 较难直接求解, 因为系统的状态分布会随着控制策略 π_θ 的更新而变化。这种非平稳的状态分布使得长期时延约束较难满足。在下一章节, 我们设计了一种称

为SCAPO的CDRL算法来求解CMDP问题 \mathcal{P}_C 。其优点是在线更新的复杂度低且只需要相对较少的在线样本，因此是数据有效的。

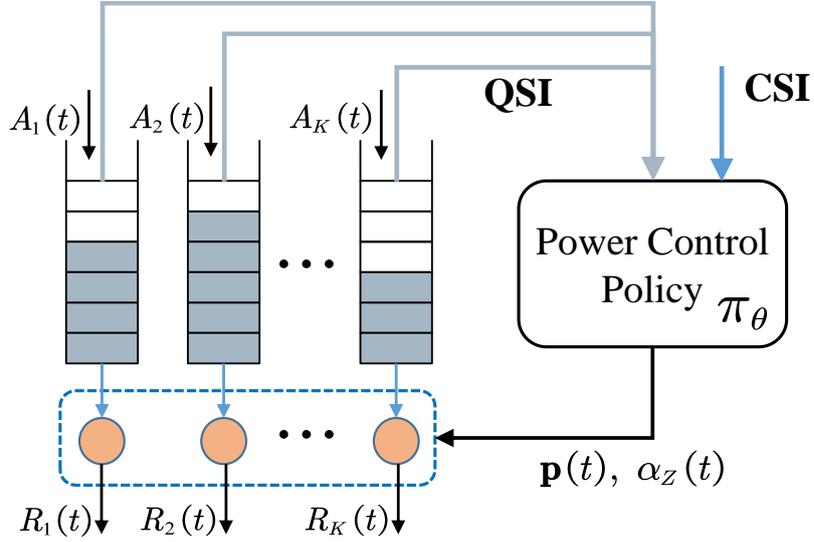


图 5.2 功率控制策略

5.3 SCAPO算法

本小节将针对问题 \mathcal{P}_C 设计一种计算复杂度低且数据有效的SCAPO算法。SCAPO的核心思想通过将原问题 \mathcal{P}_C 中复杂的目标和约束函数替换为对应的凸代理函数簇，进而在每次迭代中构建针对目标或可行解的优化问题来实现参数 θ 的更新。

5.3.1 算法总体框架

具体来说，在第 t 次迭代中，SCAPO首先将原问题的 \mathcal{P}_C 中的目标函数 $J_0(\theta)$ 和约束函数 $J_k(\theta), k = 1, \dots, K$ 替换为它们对应的代理函数 $\bar{J}_k^t(\theta), k = 0, \dots, K$ 。代理函数 $\bar{J}_k^t(\theta)$ 可以视为是利用 $J_k(\theta)$ 在第 t 次迭代参数值 θ^t 处的一阶信息而构造的一系列凸近似函数，其数学表达式为

$$\bar{J}_k^t(\theta) = \hat{J}_k^t + (\hat{g}_k^t)^T (\theta - \theta^t) + \varsigma_k \|\theta - \theta^t\|_2^2, k = 0, \dots, K, \quad (5.9)$$

其中 $\varsigma_k > 0$ 是用来保证函数凸性的正常数； $\hat{J}_k^t \in \mathbb{R}$ 和 $\hat{g}_k^t \in \mathbb{R}^{n_\theta}$ 分别是当前迭代下对函数值 $J_k(\theta)$ 和其梯度 $\nabla_\theta J_k(\theta)$ 的估值，其更新表达式为

$$\begin{aligned} \hat{J}_k^t &= (1 - \alpha_t) \hat{J}_k^{t-1} + \alpha_t \tilde{J}_k^t, \\ \hat{g}_k^t &= (1 - \alpha_t) \hat{g}_k^{t-1} + \alpha_t \tilde{g}_k^t, \end{aligned} \quad (5.10)$$

其中 $\{\alpha_t\}$ 是一个递减序列； \tilde{J}_k^t 和 \tilde{g}_k^t 是在第 t 次迭代中基于新的样本对函数和其梯度的估值，其具体形式将在下一小节中介绍。

基于每次迭代构建的代理函数 $\bar{J}_k^t(\theta)$, $k = 0, \dots, K$ ，算法首先求解关于 \mathcal{P}_O 的一个凸近似优化问题，即

$$\begin{aligned} \mathcal{P}_O : \bar{\theta}^t = \arg \min_{\theta \in \Theta} \bar{J}_0^t(\theta) \\ \text{s.t. } \bar{J}_k^t(\theta) \leq 0, k = 1, \dots, K. \end{aligned} \quad (5.11)$$

注意 \mathcal{P}_O 并不总是可解的。如其不可解，则通过求解下述问题来获得 $\bar{\theta}^t$ ：

$$\begin{aligned} \mathcal{P}_F : \bar{\theta}^t = \arg \min_{\theta \in \Theta, x} x \\ \text{s.t. } \bar{J}_k^t(\theta) \leq x, k = 1, \dots, K, \end{aligned} \quad (5.12)$$

其物理意义是最小化约束违反（Violation）值。

上述两个问题可以通过拉格朗日对偶（Lagrange-dual）等方法求解。在给定 $\bar{\theta}^t$ 结果后，参数 θ 的更新公式为

$$\theta^{t+1} = (1 - \beta_t) \theta^t + \beta_t \bar{\theta}^t, \quad (5.13)$$

这里 $\{\beta_t\}$ 也是一个递减序列。一般取 $\alpha_t = t^{-\kappa_1}$ 和 $\beta_t = t^{-\kappa_2}$ ，其中 $\kappa_1 \in (0.5, 1)$ ， $\kappa_2 \in (0.5, 1]$ 且 $\kappa_1 < \kappa_2$ 。上述SCAPO的运行过程总结于算法4中。

算法 4 SCAPO算法

输入：递减序列 α_t 和 β_t ；状态空间维度 n_s 和动作空间维度 n_a ；

初始化：策略参数 $\theta^0 \in \Theta$ ；基于状态预测动作的DNN网络；样本数据存储器 p_s

for $t = 0, 1, \dots$ **do**

 在策略 π_{θ^t} 下在线采样数据 ζ_t^n ，并更新存储器 p_s

 基于式(5.16)和(5.19)分别更新函数 \tilde{J}_k^t 和梯度 \tilde{g}_k^t , $k = 0, \dots, K$

 基于式(5.9)和(5.10)更新代理函数 $\bar{J}_k^t(\theta)$, $k = 0, \dots, K$

if 式(5.11)所表述的问题 \mathcal{P}_O 有可行解 **then**

 求解 \mathcal{P}_O 来获得 $\bar{\theta}^t$

else

 求解式(5.12)所表述的问题 \mathcal{P}_F 来获得 $\bar{\theta}^t$

end if

 根据式(5.13)更新 θ^{t+1}

end for

5.3.2 代理函数的构造方法

由公式(5.9)和(5.10)可知, 代理函数 $\tilde{J}_k^t(\theta)$ 每次更新需要基于新的函数 \tilde{J}_k^t 和其梯度信息估计值 \tilde{g}_k^t 。SCAPO算法提出了一种通过重用少量旧样本的数据有效的估计方法。具体来说, BS的代理 (Agent) 会保存最近经历的 $2T$ 个slot的样本轨迹 (Trajectory) p_s , 即

$$p_s = \{\zeta_{t-2T+1}, \dots, \zeta_{t-T+1}, \dots, \zeta_t\}, \quad (5.14)$$

其中 $\zeta_t \triangleq \{s_t, a_t, \{C'_k(s_t, a_t)\}_{k=0, \dots, K}\}$ 且

$$C'_k(s_t, a_t) = \begin{cases} \sum_{i \in \mathcal{K}} p_i(t), & \text{if } k = 0, \\ \frac{Q_k(t)}{\lambda_k} - c_k, & \text{if } k = 1, \dots, K. \end{cases} \quad (5.15)$$

基于上述的累计样本, 第 t 次迭代时 \tilde{J}_k^t 的更新将通过简单的样本平均方法得到, 即

$$\tilde{J}_k^t = \frac{1}{2T} \sum_{l=1}^{2T} C'_k(s_{t-2T+l}, a_{t-2T+l}), k = 0, \dots, K, \quad (5.16)$$

同时根据著名的策略梯度理论 (Policy-gradient-theorem) [119], 其梯度 $\nabla_{\theta} J_k(\theta)$ 由公式

$$\nabla_{\theta} J_k(\theta) = \mathbb{E}_{s \sim \mathbf{P}_{\pi_{\theta}}, a \sim \pi_{\theta}(\cdot|s)} [Q_k^{\pi_{\theta}}(s, a) \nabla_{\theta} \log \pi_{\theta}(a | s)] \quad (5.17)$$

所表征, 其中 $\mathbf{P}_{\pi_{\theta}}$ 是系统在策略 π_{θ} 下达到的稳定状态。同时定义

$$Q_k^{\pi_{\theta}}(s, a) = \mathbb{E}_{p_s \sim \pi_{\theta}} \left[\sum_{l=0}^{\infty} (C'_k(s_l, a_l) - J_k(\theta)) \mid S_0 = s, A_0 = a \right] \quad (5.18)$$

表示以 $S_0 = s$ 和 $A_0 = a$ 为起始, 在策略 π_{θ} 下的Q值 (Q-value) 函数。因此, 同样基于平均采样的方法, \tilde{g}_k^t 的计算方法为

$$\tilde{g}_k^t = \frac{1}{T} \sum_{l=1}^T \hat{Q}_k^{t-2T+l}(s_{t-2T+l}, a_{t-2T+l}) \nabla_{\theta} \log \pi_{\theta}(a_{t-2T+l} \mid s_{t-2T+l}), \quad (5.19)$$

其中

$$\hat{Q}_k^{t-2T+l}(s_{t-2T+l}, a_{t-2T+l}) = \sum_{l'=0}^{T-1} \left(C'_k(s_{t-2T+l+l'}, a_{t-2T+l+l'}) - \hat{J}_k^t \right) \quad (5.20)$$

表示从以 s_{t-2T+l} 和 a_{t-2T+l} 为起始, 延续 T 个slot的样本轨迹中得到的Q-value估值, 如图(5.3)所示。

可以看出, 式(5.16)和(5.19)所表述的更新过程只需要基于系统新采集的 $2T$ 个样本并通过简单的平均运算即可实现, 因此是一种低成本且数据有效的更新方法。

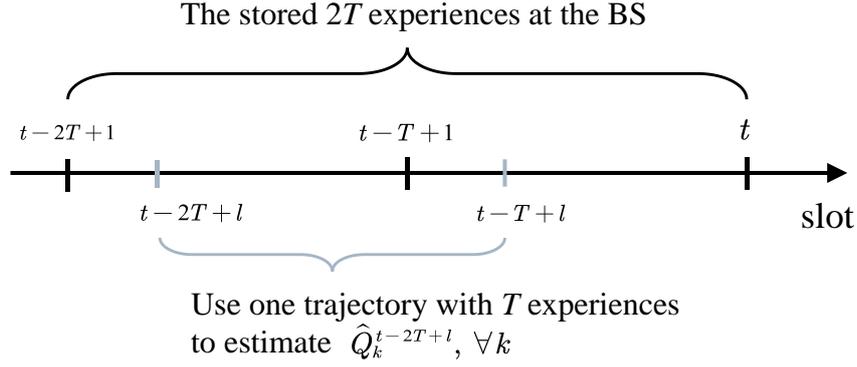


图 5.3 Q-value的估计方法

5.3.3 参数化的策略方法

本小节将介绍参数化的策略 π_θ 实现方法。具体来说，为了兼顾策略的探索（Exploration）和利用（Exploitation）功能，我们采取广泛应用的高斯策略（Gaussian-policy）方法^[120]，即

$$\pi_\theta(a | s) = (2\pi)^{-\frac{n_a}{2}} \prod_{i=1}^{n_a} \sigma'_i \exp\left(-\sum_{i=1}^{n_a} \frac{(a^i - \mu_{\theta'}^i(s))^2}{2\sigma_i'^2}\right), \quad (5.21)$$

其中 n_a 定义为动作空间的维度，由式(5.7)所给定； a^i 表示动作向量的第 i 个元素；策略的参数 θ 包括用于预测动作均值的DNN网络本身参数 θ' 和高斯策略下的标准差 $\sigma'_i, i = 1, \dots, n_a$ 。因此 σ'_i 作为参数的一部分控制策略探索的程度。特别的，当 $\sigma'_i \rightarrow 0, \forall i$ 时，高斯策略将退化成完全确定性策略^[121]。

5.3.4 算法的收敛性分析

本小节，我们将给出SCAPO算法的一般收敛性证明。我们将证明在满足一定假设下，对形如下式的一般CMDP问题 \mathcal{P}_G ，SCAPO算法都可以收敛至其KKT点。

$$\begin{aligned} \mathcal{P}_G : \min_{\theta \in \Theta} J_0(\theta) &\triangleq \lim_{T \rightarrow \infty} \frac{1}{T} \mathbb{E}_{p_s \sim \pi_\theta} \left[\sum_{t=0}^{T-1} C_0(s_t, a_t) \right] \\ \text{s.t. } J_i(\theta) &\triangleq \lim_{T \rightarrow \infty} \frac{1}{T} \mathbb{E}_{p_s \sim \pi_\theta} \left[\sum_{t=0}^{T-1} C_i(s_t, a_t) \right] - c_i \leq 0, i = 1, \dots, m. \end{aligned} \quad (5.22)$$

对问题 \mathcal{P}_G ，我们作出以下假设：

假设 1. （对问题 \mathcal{P}_G 的假设）

1. 参数空间 $\Theta \subseteq \mathbb{R}^{n_\theta}$ ($n_\theta > 0$) 是紧（Compact）且凸（Convex）的。
2. 状态空间 $\mathcal{S} \subseteq \mathbb{R}^{n_s}$ ($n_s > 0$) 和动作空间 $\mathcal{A} \subseteq \mathbb{R}^{n_a}$ ($n_a > 0$) 都是紧集。

3. 对任意的 $C_i, i \in 0, \dots, m$, 函数列 $J_i(\theta)$ 的一阶和二阶偏导都是一致有界 (Uniformly-bounded) 的; 且 π_θ 关于 θ 的函数满足李普希茨连续 (Lipschitz-continuous)。
4. 存在常数 $m_c > 0$ 和 $\rho \in (0, 1)$, 使得

$$\sup_{s \in \mathcal{S}} \|\mathbf{P}(S_t | S_0 = s) - \mathbf{P}_{\pi_\theta}\|_{\text{TV}} \leq m_c \rho^t, \forall t \geq 0 \quad (5.23)$$

成立。其中 \mathbf{P}_{π_θ} 是在策略 π_θ 下达到的稳定状态分布; 且

$$\|\mu - \nu\|_{\text{TV}} = \int_{s \in \mathcal{S}} |\mu(ds) - \nu(ds)| \quad (5.24)$$

定义为概率密度分布函数 μ 和 ν 之间的总变差 (Total-variation) 距离。

我们想强调的是, 假设1的第3点在DRL算法的收敛性证明中是经常用到的^[122]; 第4点是说明系统在策略 π_θ 下达到的平稳状态分布与起始状态是无关系的, 这也是经常用到的渐进假设^[120]。在本章考虑的问题中, 取 $C_0(s_t, a_t) \triangleq \sum_{k \in \mathcal{K}} p_k(t)$ 和 $C_i(s_t, a_t) \triangleq Q_i(t)/\lambda_i, i = 1, \dots, K$, 满足假设1。同时, 对式(5.10)和(5.13)中的更新步长 α_t 和 β_t , 我们也有如下假设:

假设 2. (对步长 α_t 和 β_t 的假设)

1. $\alpha_t \rightarrow 0$; 对 $\kappa \in (0, 1)$, 有 $\frac{1}{\alpha_t} \leq O(t^\kappa)$; $\sum_t \alpha_t t^{-1} < \infty$; $\sum_t \alpha_t \left(\log^2 t \sum_{t' = t - \log t}^t \beta_{t'} \right) < \infty$; $\sum_t (\alpha_t)^2 < \infty$ 。
2. $\beta_t \rightarrow 0$; $\sum_t \beta_t = \infty$; $\sum_t (\beta_t)^2 < \infty$ 。
3. $\lim_{t \rightarrow \infty} (\beta_t / \alpha_t) = 0$ 。

在本章中, 我们取 $\alpha_t = t^{-\kappa_1}$ 和 $\beta_t = t^{-\kappa_2}$ (其中 $\kappa_1 \in (0.5, 1)$, $\kappa_2 \in (0.5, 1]$ 且 $\kappa_1 < \kappa_2$), 亦满足假设2。基于此, 我们给出如下引理。

引理 1. (估值函数的渐进一致性): 对任意 $i \in \{0, \dots, m\}$, SCAPO算法产生的估值函数满足

$$\lim_{t \rightarrow \infty} \left| \hat{J}_i^t - J_i(\theta^t) \right| = 0, \quad (5.25a)$$

$$\lim_{t \rightarrow \infty} \left\| \hat{g}_i^t - \nabla J_i(\theta^t) \right\|_2 = 0. \quad (5.25b)$$

基于上述引理, 考虑一个收敛于极限点 θ^* 的子序列 $\{\theta^{t_j}\}_{j=1}^\infty$, 该更新过程获得的收敛代理函数 $\check{J}_i(\theta) \triangleq \lim_{j \rightarrow \infty} \check{J}_i^{t_j}(\theta), \forall i$ 自然满足

$$\begin{aligned} \left| \check{J}_i(\theta^*) - J_i(\theta^*) \right| &= 0, \\ \left\| \nabla \check{J}_i(\theta^*) - \nabla J_i(\theta^*) \right\|_2 &= 0. \end{aligned} \quad (5.26)$$

最后，我们作出关于收敛代理函数的斯莱特（Slater）条件假设：

假设 3.（Slater条件假设）问题 \mathcal{P}_G 的所有约束函数在极限点 θ^* 处满足slater条件，即存在一个绝对可行点 $\theta \in \text{relint}\Theta$ ，使得 $\check{J}_i(\theta) < 0, \forall i \in \{1, \dots, m\}$ 。

这也是KKT条件证明过程中的一项常用约束假设条件^[123]。基于引理1和假设3，我们给出SCAPO算法的收敛性定理。

定理 2.（SCAPO算法的收敛性定则）：在满足假设1和2时，对SCAPO算法由可行起始点 θ^0 出发并以足够小的初始步长 β_0 产生的迭代序列 $\{\theta^t\}_{t=1}^\infty$ ，满足假设3的每个极限点 θ^* 都是原问题 \mathcal{P}_G 的一个驻点解，即KKT条件成立。

引理1和定理2的证明详见附录A和B。

5.4 仿真结果与分析

本章节中，我们将基于Saleh-Valenzuela信道模型搭建仿真系统，比较所提SCAPO算法和其它几种CDRL算法的性能，包括发送功率消耗和时延约束的满足情况。具体的系统和信道参数详见表(5.1)。

表 5.1 系统和信道参数设置

名称	设置
信道模型	Saleh-Valenzuela
用户数	4, 8
散射路径数	4
载波频率	3GHz
带宽	10MHz
天线阵列	ULA
发送天线数	8, 16
路径增益	-10dB到10dB间均匀分布
噪声功率密度	-100dBm/Hz
包到达平均速率	10Mbit/s

在本章中，我们选择两个经典的CDRL算法，即拉格朗日近端策略优化（Proximal policy optimization, PPO）算法^[124]和CPO算法^[117]，作为两个主要的基线方案。另外，我们

也用等功率分配和渐进最优的正则化因子 $\alpha_Z = \sigma^2/p$ 作为RZF预编码的另一种基线方案，这里 σ^2 和 p 分别是各用户的噪声方差和发送功率。对所提出的SCAPO算法，我们假设BS在线样本数据存储器 p_s 的空间是 $2T=3000$ 个。在每次更新时，使用新加入的100个数据样本（Batch-size）和旧的2900个样本计算。同时为了更详尽地比较方案性能，我们还设置了两个对照方案，即SCAPO2和SCAPO3，它们不重用旧数据而分别基于新加入的1000和3000个新样本进行更新。需要指出，额外设置这两个对照方案是为了更好的对照性能，实际通信系统中往往无法为每次更新提供如此多的在线样本。

图(5.4)和图(5.5)展示了在用户数为4，发送天线数为8的场景下的学习曲线。可以看出所提出的SCAPO算法相较PPO-Lagrangian和CPO基线均有明显的性能优势，能以较低的传输功率满足平均时延约束。同时也能看出，SCAPO算法的收敛速度相较其余基线方案也更优。特别地，在与其他两种SCAPO对照基线（SCAPO2和SCAPO3）相比时，标准SCAPO方法（SCAPO1）可以实现基本与SCAPO3方案类似的性能表现和收敛速度。这也证明了我们的方法具有在线样本数据需求少、计算复杂度低的优势。

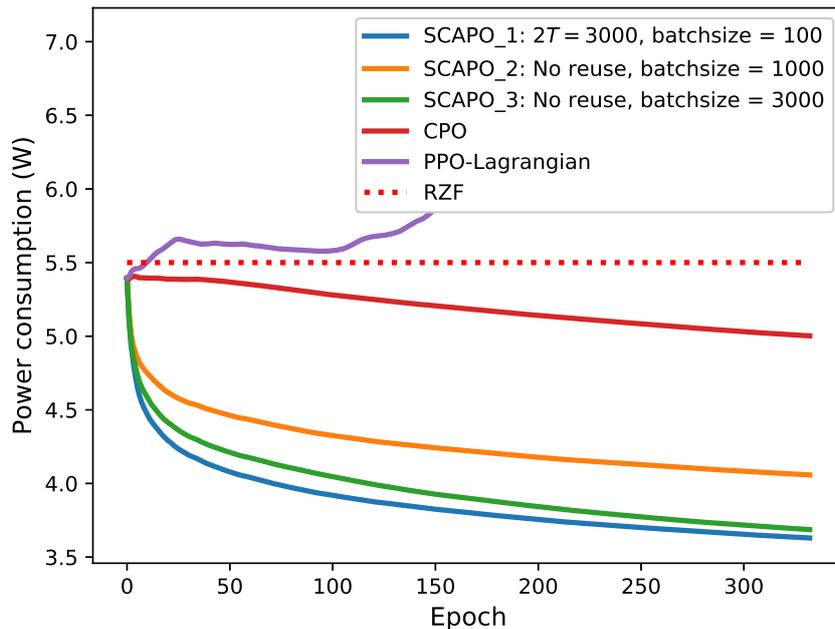


图 5.4 4用户8发送天线场景下的功率消耗学习曲线

图(5.6)和图(5.7)则展示了在用户数为8，发送天线数为16的场景下的学习曲线。在这种更复杂的传输场景下，几种方案都出现了不同的性能退化情况。但所提的SCAPO算法仍然能展现出显著的性能优势，可以在满足时延约束的同时以更低的计算复杂度实现更优的功率分配。

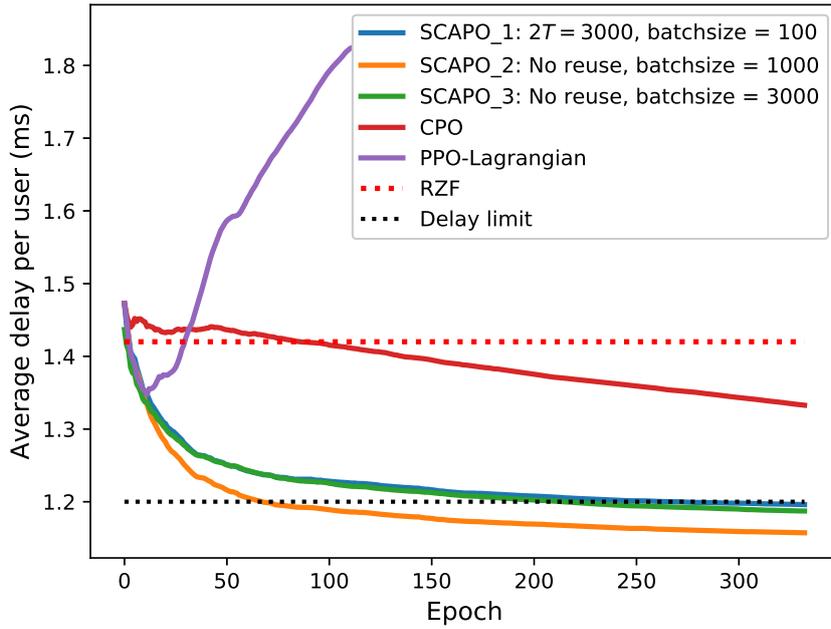


图 5.5 4用户8发送天线场景下的队列时延学习曲线

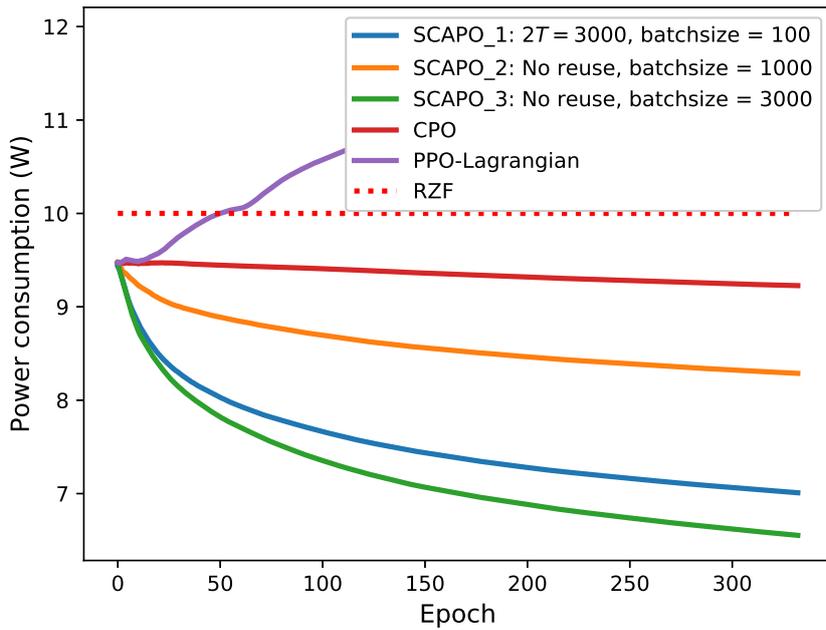


图 5.6 8用户16发送天线场景下的功率消耗学习曲线

5.5 本章小结

本章中，我们考虑一个包含动态队列的下行MU-MIMO系统，设计了一种基于RZF预编码的功率控制策略，用以解决在满足各用户长期平均时延约束下的长期平均发送功率最小化的问题。该问题是一个复杂的CMDP问题，传统的CDRL算法由于更新复杂度高、对在线样本需求量大等原因几乎无法适用于实际通信系统。为了解决这个困难，我们提出了

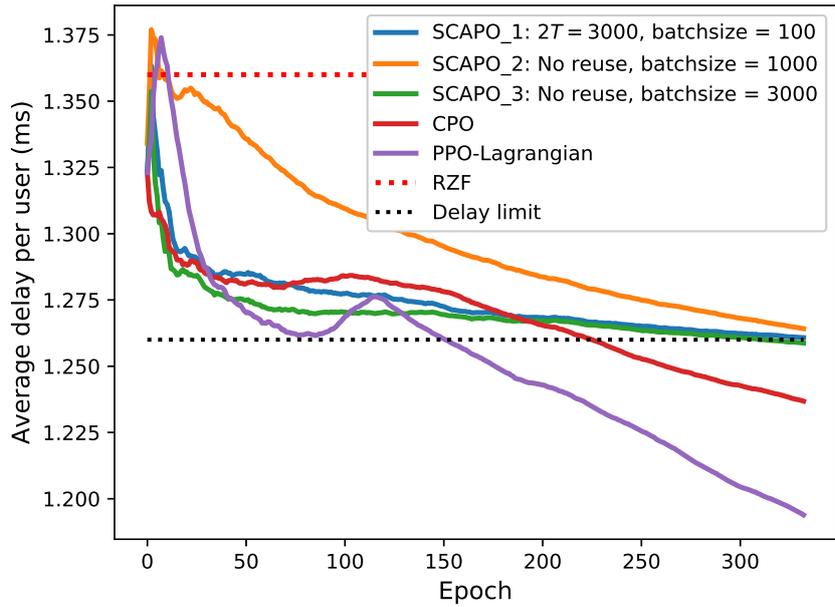


图 5.7 8用户16发送天线场景下的队列时延学习曲线

一种称为SCAPO的CDRL算法框架。所提算法通过将原问题中复杂的目标和约束函数转化为其对应的凸代理函数，从而在每次迭代中只需求解一个凸优化问题，降低了更新计算复杂度。同时，在每次更新中，所提算法可以重用旧的样本数据，只需要少数的在线新样本即可完成训练，因而是数据有效的。最后，我们将SCAPO算法与其他基线方案进行比较，仿真结果证明了所提算法的优越性。

6 总结与展望

6.1 全文总结

本文针对现有5G及B5G大规模天线系统中几项亟需解决的关键性技术难题，如FDD系统高精度信道反馈、空域并发的多用户频分调度、覆盖增强场景下的预编码设计和针对时延敏感应用的长时间尺度功率控制问题，提出了一系列行之有效的解决算法和实施技术。具体成果包括：

- (1) 研究了FDD多用户massive MIMO宽带系统中的高精度CSI反馈问题。我们通过挖掘massive MIMO信道的动态突发稀疏性、强时间相关性和频域相关性提出了一种联合空频域的TSAC信道反馈方案。方案包括三项核心设计，即针对二进制支撑矢量的二级差分压缩方案、针对稀疏系数的多原子矢量量化方法和离线-在线混合的字典更新技术。相较现有的NR Type-2基线反馈方法，所提方案可以同时取得反馈开销和量化性能上的显著优势。
- (2) 研究了单小区下行系统中支持空域并发的多用户频分调度问题。为了解决最大空域并发数和RB资源连续性分配这两个棘手的实际链路约束，我们提出了一种JSFD-SMP用户调度算法和KPS实施简化技术。相较现有基线方案，所提算法可以在保证小区用户体验速率公平性的同时显著提升小区吞吐量性能。
- (3) 研究了远距覆盖增强通信场景下以BLER为直接性能指标的下行预编码设计问题。为了解决BLER函数极难刻画的技术挑战，我们基于EESM准则构建了无编码系统中的转化优化问题，并设计了基于逐级迭代的E-S-WMMSE算法来有效寻找目标函数的局部最优解。为了探索所提方案相较经典基线方案的增益空间，我们设置了大量仿真实验，比较验证了不同MCS、传输数据流数目情况下的增益结果。实验结果不仅证明了所提方案的有效性，也为工业界在处理类似覆盖增强场景下的性能增益空间评估问题提供一定的理论和实践指导。
- (4) 研究了时延敏感应用中常涉及的多用户长期平均时延约束下的长期功率控制问题。对这类CMDP问题，为了解决传统CDRL算法计算复杂度高、数据利用效率差的缺

陷，我们提出了一种SCAPO的新型CDRL算法框架。通过将原问题中复杂的目标和约束函数转化为其对应的凸代理函数，从而在每次迭代中只需求解一个凸优化问题，降低了更新计算复杂度。同时，在每次更新中，所提算法可以重用旧的样本数据，只需要少数的在线新样本即可完成训练，因而是数据有效的。仿真实验证明了所提算法相较其他基线的性能优越性。

6.2 未来工作展望

本文是作者对研究生期间相关项目、课题的工作总结，针对现有及未来通信系统中诸如信道反馈、用户调度、下行预编码和功率控制等亟需解决关键技术进行了初步的研究。虽然取得了一些研究成果，但受限于时间精力和研究水平，本文工作的深度和价值也是有限的，上述课题还有可进一步发掘的探索研究方向。下面是作者对本文第二至第五章节课题中可进一步研究的方向或改进思路所提出的一些想法，包括：

- (1) 在本文第二章节考虑的高精度信道反馈问题中，所提的TSAC方法虽然相较现有NR Type-2方案具有显著的反馈开销和性能优势，但方案的反馈精度比较依赖于量化字典的选取。为了保证精度，我们提出了一种离线-在线混合的字典更新方法。然而，这种训练需要针对不同信道场景分别进行，灵活性受限。因此，如何基于机器学习或强化学习等方法设计一种能针对不同信道条件快速自适应学习的鲁棒字典更新算法，对进一步提高反馈精度及降低算法的实施开销具有重要的意义。
- (2) 在本文第三章节考虑的支持用户并发的多用户频分调度问题中，我们基于消息传播理论设计了一种JSFD-SMP算法来求解复杂的链路约束下的调度策略。虽然与传统的完全基于启发式的算法相比，JSFD-SMP性能更优。但需要指出，我们为了降低复杂度仍然采取了部分启发式的基于分轮迭代的求解方法，因此也只能找到原问题的近似解。同时，对于所构建的环形因子图模型，虽然我们通过结合其对称性结构设计了一种左右振荡更新算法，但大部分非对称环形结构因子图并无法在理论上保证算法的一致收敛性。不过幸运的是，消息传播算法本身显著的优点是其计算结构很适合利用并行计算平台加速实施。因此在未来如何有效地将此类方法应用于解决更多通信场景问题也是十分有价值的研究问题。
- (3) 在本文第四章节考虑的低MCS场景下针对BLER的下行预编码设计问题中，我们虽然基于现有的链路评估方法和大量仿真实验说明了EESM度量准则可以在一定范围内较好地刻画BLER的性能，但尚且缺乏坚实的理论分析。另外，这类方法的准确性也十分依赖于度量函数中参数的选取，而这需要通过大量的仿真实验才能

获得，复杂性高。因此，找到一种对刻画BLER性能更可靠且参数获取难度低的性能度量函数，对BLER增益空间探索乃至其它链路评估应用都是极有价值的。

- (4) 在本文第五章节考虑的长期时延约束下的功率控制问题中，所提出的SCAPO方法虽然相较现有CDRL算法是复杂度低且数据有效的。但不可否认的是，在问题规模较大（带宽大、用户数多）的场景中，算法将由于计算复杂度变大、在线收敛速度变慢而无法满足不同实时决策的需求。当然，这也是现有的深度强化学习算法几乎都无法适用于实际通信系统的主要原因。因此，如何降低神经网络的更新复杂度和在线收敛时间并将其可靠地应用于未来无线通信系统将会是一个极具理论和应用价值的研究方向。

参考文献

- [1] T. L. Marzetta. Noncooperative cellular wireless with unlimited numbers of base station antennas[J]. *IEEE Transactions on Wireless Communications*, 2010. 9(11):3590–3600. doi:10.1109/TWC.2010.092810.091092.
- [2] Z. Gao, L. Dai, Z. Wang, S. Chen. Spatially common sparsity based adaptive channel estimation and feedback for FDD massive MIMO[J]. *IEEE Transactions on Signal Processing*, 2015. 63(23):6169–6183. doi:10.1109/TSP.2015.2463260.
- [3] H. Almosa, S. Mosleh, E. Perrins, L. Liu. Downlink channel estimation with limited feedback for FDD multi-user massive MIMO with spatial channel correlation[C]//2018 IEEE International Conference on Communications (ICC). 2018:1–6. doi:10.1109/ICC.2018.8422121.
- [4] P. N. Alevizos, X. Fu, N. D. Sidiropoulos, Y. Yang, A. Bletsas. Limited feedback channel estimation in massive MIMO with non-uniform directional dictionaries[J]. *IEEE Transactions on Signal Processing*, 2018. 66(19):5127–5141. doi:10.1109/TSP.2018.2865412.
- [5] W. Shen, L. Dai, G. Gui, Z. Wang, R. W. Heath, F. Adachi. AoD-adaptive subspace codebook for channel feedback in FDD massive MIMO systems[C]//2017 IEEE International Conference on Communications (ICC). 2017:1–5. doi:10.1109/ICC.2017.7997168.
- [6] W. Shen, L. Dai, B. Shim, Z. Wang, R. W. Heath. Channel feedback based on AoD-adaptive subspace codebook in FDD massive MIMO systems[J]. *IEEE Transactions on Communications*, 2018. 66(11):5235–5248. doi:10.1109/TCOMM.2018.2849755.
- [7] W. Shen, L. Dai, Y. Li, Z. Wang, L. Hanzo. Channel feedback codebook design for millimeter-wave massive MIMO systems relying on lens antenna array[J]. *IEEE Wireless Communications Letters*, 2018. 7(5):736–739. doi:10.1109/LWC.2018.2818130.
- [8] W. Shen, L. Dai, Y. Shi, B. Shim, Z. Wang. Joint channel training and feedback for FDD massive MIMO systems[J]. *IEEE Transactions on Vehicular Technology*, 2016. 65(10):8762–8767. doi:10.1109/TVT.2015.2508033.
- [9] W. Shen, B. Wang, J. Feng, C. Gao, J. Ma. Differential CSIT acquisition based on compressive sensing for FDD massive MIMO systems[C]//2015 IEEE 81st Vehicular Technology Conference (VTC Spring). 2015:1–4. doi:10.1109/VTCSpring.2015.7145776.
- [10] X. Rao, V. K. N. Lau, X. Kong. CSIT estimation and feedback for FDD multi-user massive MIMO

- systems[C]//2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). 2014:3157–3161. doi:10.1109/ICASSP.2014.6854182.
- [11] X. Rao, V. K. N. Lau. Distributed compressive CSIT estimation and feedback for FDD multi-user massive MIMO systems[J]. *IEEE Transactions on Signal Processing*, 2014. 62(12):3261–3271. doi:10.1109/TSP.2014.2324991.
- [12] T. Wang, C. Wen, H. Wang, F. Gao, T. Jiang, S. Jin. Deep learning for wireless physical layer: Opportunities and challenges[J]. *China Communications*, 2017. 14(11):92–111. doi:10.1109/CC.2017.8233654.
- [13] Z. Xu, J. Sun. Model-driven deep-learning[J]. *National Science Review*, 2017. 5(1):22–24. ISSN 2095-5138. doi:10.1093/nsr/nwx099.
- [14] C. Wen, W. Shih, S. Jin. Deep learning for massive MIMO CSI feedback[J]. *IEEE Wireless Communications Letters*, 2018. 7(5):748–751. doi:10.1109/LWC.2018.2818160.
- [15] T. Wang, C. Wen, S. Jin, G. Y. Li. Deep learning-based CSI feedback approach for time-varying massive MIMO channels[J]. *IEEE Wireless Communications Letters*, 2019. 8(2):416–419. doi:10.1109/LWC.2018.2874264.
- [16] F. Capozzi, G. Piro, L. A. Grieco, G. Boggia, P. Camarda. Downlink packet scheduling in LTE cellular networks: Key design issues and a survey[J]. *IEEE Communications Surveys & Tutorials*, 2013. 15(2):678–700.
- [17] P. Kela, J. Puttonen, N. Kolehmainen, T. Ristaniemi, T. Henttonen, M. Moisio. Dynamic packet scheduling performance in UTRA long term evolution downlink[C]//2008 3rd International Symposium on Wireless Pervasive Computing. 2008:308–313. doi:10.1109/ISWPC.2008.4556220.
- [18] C. Wengerter, J. Ohlhorst, A. G. E. von Elbwart. Fairness and throughput analysis for generalized proportional fair frequency scheduling in OFDMA[C]//2005 IEEE 61st Vehicular Technology Conference. volume 3. 2005:1903–1907 Vol. 3. doi:10.1109/VETECS.2005.1543653.
- [19] M. Proebster, C. M. Mueller, H. Bakker. Adaptive fairness control for a proportional fair LTE scheduler[C]//21st Annual IEEE International Symposium on Personal, Indoor and Mobile Radio Communications. 2010:1504–1509. doi:10.1109/PIMRC.2010.5671970.
- [20] X. Li, B. Li, B. Lan, M. Huang, G. Yu. Adaptive pf scheduling algorithm in LTE cellular system[C]//2010 International Conference on Information and Communication Technology Convergence (ICTC). 2010:501–504. doi:10.1109/ICTC.2010.5674773.
- [21] A. Pokhariyal, K. I. Pedersen, G. Monghal, I. Z. Kovacs, C. Rosa, T. E. Kolding, P. E. Mogensen. HARQ aware frequency domain packet scheduler with different degrees of fairness for the UTRAN long term evolution[C]//2007 IEEE 65th Vehicular Technology Conference - VTC2007-Spring. 2007:2761–2765. doi:10.1109/VETECS.2007.567.
- [22] Fattah, Hossam, Alnuweiri, Hussein. A cross-layer design for dynamic resource block allocation in 3G long term evolution system[C]//2009 IEEE 6th International Conference on Mobile Adhoc and Sensor

- Systems. 2009:929–934. doi:10.1109/MOBHOC.2009.5337027.
- [23] P. Liu, R. Berry, M. L. Honig. Delay-sensitive packet scheduling in wireless networks[C]//2003 IEEE Wireless Communications and Networking, 2003. WCNC 2003. volume 3. 2003:1627–1632 vol.3. doi: 10.1109/WCNC.2003.1200630.
- [24] M. Andrews, K. Kumaran, K. Ramanan, A. Stolyar, P. Whiting, R. Vijayakumar. Providing quality of service over a shared wireless link[J]. IEEE Communications Magazine, 2001. 39(2):150–154. doi: 10.1109/35.900644.
- [25] P. Ameigeiras, Y. Wang, J. N. Ortiz, P. E. Mogensen. Traffic models impact on OFDMA scheduling design[J]. EURASIP Journal Wireless Communications and Networking, 2001. 39(2):150–154.
- [26] Kibeom Seong, R. Narasimhan, J. M. Cioffi. Queue proportional scheduling via geometric programming in fading broadcast channels[J]. IEEE Journal on Selected Areas in Communications, 2006. 24(8):1593–1602. doi:10.1109/JSAC.2006.879404.
- [27] G. Venkatraman, A. Tolli, J. Janhunen, M. Juntti. Low complexity multiuser MIMO scheduling for weighted sum rate maximization[C]//2014 22nd European Signal Processing Conference (EUSIPCO). 2014:820–824.
- [28] J. Mei, K. Zheng, L. Zhao, Y. Teng, X. Wang. A latency and reliability guaranteed resource allocation scheme for LTE V2V communication systems[J]. IEEE Transactions on Wireless Communications, 2018. 17(6):3850–3860. doi:10.1109/TWC.2018.2816942.
- [29] J. Huang, Z. Niu. Buffer-aware and traffic-dependent packet scheduling in wireless OFDM networks[C]//2007 IEEE Wireless Communications and Networking Conference. 2007:1554–1558. doi: 10.1109/WCNC.2007.293.
- [30] F. Brouwer, I. de Bruin, J. C. Silva, N. Souto, F. Cercas, A. Correia. Usage of link-level performance indicators for HSDPA network-level simulations in E-UMTS[C]//Eighth IEEE International Symposium on Spread Spectrum Techniques and Applications - Programme and Book of Abstracts (IEEE Cat. No.04TH8738). 2004:844–848. doi:10.1109/ISSSTA.2004.1371820.
- [31] 3GPP. NR: Physical layer procedures for data[J]. 3GPP, Technical Report 38.214, ver. 15.3.0, 2018.
- [32] 3GPP. NR: Radio resource control protocol specification[J]. 3GPP, Technical Report 38.331, ver. 16.2.0, 2020.
- [33] S. B. Lee, S. Choudhury, A. Khoshnevis, S. Xu, S. Lu. Downlink MIMO with frequency-domain packet scheduling for 3GPP LTE[C]//IEEE INFOCOM 2009. 2009:1269–1277. doi:10.1109/INFCOM.2009.5062041.
- [34] Y. Xu, H. Yang, F. Ren, C. Lin, X. Shen. Frequency domain packet scheduling with MIMO for 3GPP LTE downlink[J]. IEEE Transactions on Wireless Communications, 2013. 12(4):1752–1761. doi:10.1109/TWC.2013.022113.120678.

- [35] S. B. Lee, I. Pefkianakis, S. Choudhury, S. Xu, S. Lu. Exploiting spatial, frequency, and multiuser diversity in 3GPP LTE cellular networks[J]. *IEEE Transactions on Mobile Computing*, 2012. 11(11):1652–1665. doi:10.1109/TMC.2011.206.
- [36] H. Zhang, N. Prasad, S. Rangarajan. MIMO downlink scheduling in LTE systems[C]//2012 Proceedings IEEE INFOCOM. 2012:2936–2940. doi:10.1109/INFOCOM.2012.6195733.
- [37] A. Ragaleux, S. Baey, A Fladenmuller. An efficient and generic downlink resource allocation procedure for pre-5G networks[J]. *Wireless Communications and Mobile Computing*, 2016. 16:3089–3103.
- [38] Y. Chen, Y. Wu, Y. T. Hou, W. Lou. mCore: Achieving sub-millisecond scheduling for 5G MU-MIMO systems[C]//IEEE INFOCOM 2021 - IEEE Conference on Computer Communications. 2021:1–10. doi:10.1109/INFOCOM42981.2021.9488684.
- [39] Y. Huang, Y. T. Hou, W. Lou. A deep-learning-based link adaptation design for eMBB/URLLC multiplexing in 5G NR[C]//IEEE INFOCOM 2021 - IEEE Conference on Computer Communications. 2021:1–10. doi:10.1109/INFOCOM42981.2021.9488790.
- [40] S. Cical, V. Tralli. Adaptive resource allocation with proportional rate constraints for uplink SC-FDMA systems[J]. *IEEE Communications Letters*, 2014. 18(8):1419–1422. doi:10.1109/LCOMM.2014.2332997.
- [41] I. C. Wong, O. F. Oteri, J. W. McCoy. Resource allocation in multi-data stream communication link[M]. U.S. Patent 7911934B2, Mar. 2011.
- [42] S. Sun, S. Moon. Practical scheduling algorithms with contiguous resource allocation for next-generation wireless systems[J]. *IEEE Wireless Communications Letters*, 2021. 10(4):725–729. doi:10.1109/LWC.2020.3041836.
- [43] Y. Huang, S. Li, Y. T. Hou, W. Lou. GPF: A GPU-based design to achieve 100us scheduling for 5G NR[C]//Proceedings of the 24th Annual International Conference on Mobile Computing and Networking. New York, NY, USA, 2018:207–222. doi:10.1145/3241539.3241552.
- [44] T. K. Lyu. Capacity of multi-user MIMO systems with MMSE and ZF precoding[C]//2016 IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPs). 2016:1083–1084. doi:10.1109/INFOCOMW.2016.7562264.
- [45] C. Feng, Y. Jing, S. Jin. Interference and outage probability analysis for massive MIMO downlink with MF precoding[J]. *IEEE Signal Processing Letters*, 2016. 23(3):366–370. doi:10.1109/LSP.2015.2511630.
- [46] E. Mukubwa, O. A. Sokoya, D. S. Ilcev. Comparison and analysis of massive MIMO linear precoding schemes in the downlink[C]//2017 IEEE AFRICON. 2017:187–191. doi:10.1109/AFRCON.2017.8095479.
- [47] M. C. Thomas. *Elements of Information Theory*[M]. John Wiley & Sons, 1999.
- [48] M. Costa. Writing on dirty paper (Corresp.)[J]. *IEEE Transactions on Information Theory*, 1983.

- 29(3):439–441. doi:10.1109/TIT.1983.1056659.
- [49] M. Tomlinson. New automatic equalizer employing modulo arithmetic[J]. *Electronics Letters*, 1971. 7(5):138–139.
- [50] C. B. Peel, B. M. Hochwald, A. L. Swindlehurst. A vector-perturbation technique for near-capacity multi-antenna multiuser communication-part I: Channel inversion and regularization[J]. *IEEE Transactions on Communications*, 2005. 53(1):195–202. doi:10.1109/TCOMM.2004.840638.
- [51] L. Liang, W. Xu, X. Dong. Low-complexity hybrid precoding in massive multiuser MIMO systems[J]. *IEEE Wireless Communications Letters*, 2014. 3(6):653–656. doi:10.1109/LWC.2014.2363831.
- [52] O. E. Ayach, S. Rajagopal, S. Abu-Surra, Z. Pi, R. W. Heath. Spatially sparse precoding in millimeter wave MIMO systems[J]. *IEEE Transactions on Wireless Communications*, 2014. 13(3):1499–1513. doi:10.1109/TWC.2014.011714.130846.
- [53] H. Huang, Y. Song, J. Yang, G. Gui, F. Adachi. Deep-learning-based millimeter-wave massive MIMO for hybrid precoding[J]. *IEEE Transactions on Vehicular Technology*, 2019. 68(3):3027–3032. doi:10.1109/TVT.2019.2893928.
- [54] H. Huang, Y. Peng, J. Yang, W. Xia, G. Gui. Fast beamforming design via deep learning[J]. *IEEE Transactions on Vehicular Technology*, 2020. 69(1):1065–1069. doi:10.1109/TVT.2019.2949122.
- [55] A. M. Elbir, A. K. Papazafeiropoulos. Hybrid precoding for multiuser millimeter wave massive MIMO systems: A deep learning approach[J]. *IEEE Transactions on Vehicular Technology*, 2020. 69(1):552–563. doi:10.1109/TVT.2019.2951501.
- [56] A. Adhikary, J. Nam, J. Ahn, G. Caire. Joint spatial division and multiplexing - The large-scale array regime[J]. *IEEE Transactions on Information Theory*, 2013. 59(10):6441–6463. doi:10.1109/TIT.2013.2269476.
- [57] A. Liu, V. K. N. Lau. Phase only RF precoding for massive MIMO systems with limited RF chains[J]. *IEEE Transactions on Signal Processing*, 2014. 62(17):4505–4515. doi:10.1109/TSP.2014.2337840.
- [58] A. Liu, V. K. N. Lau, M. Zhao. Stochastic successive convex optimization for two-timescale hybrid precoding in massive MIMO[J]. *IEEE Journal of Selected Topics in Signal Processing*, 2018. 12(3):432–444. doi:10.1109/JSTSP.2018.2819084.
- [59] X. Chen, A. Liu, Y. Cai, V. K. N. Lau, M. Zhao. Randomized two-timescale hybrid precoding for downlink multicell massive MIMO systems[J]. *IEEE Transactions on Signal Processing*, 2019. 67(16):4152–4167. doi:10.1109/TSP.2019.2926017.
- [60] A. Abdel-Samad, A. B. Gershman, T. N. Davidson. Robust transmit beamforming based on imperfect channel feedback[C]//*IEEE 60th Vehicular Technology Conference*, 2004. VTC2004-Fall. 2004. volume 3. 2004:2049–2053 Vol. 3. doi:10.1109/VETEFC.2004.1400400.
- [61] B. Ottersten M. Bengtsson. Optimal downlink beamforming using semidefinite optimization[C]//*37th*

- Allerton Conference on Communication, Control, and Computing. 1999:977–996.
- [62] T. Endeshaw, B. Chalise, L. Vandendorpe. Robust sum rate optimization for the downlink multiuser MIMO systems: Worst-case design[C]//IEEE International Conference on Communications. 2010:3–7.
- [63] A. J. Tenenbaum, R. S. Adve. Improved sum-rate optimization in the multiuser MIMO downlink[C]//42nd Annual Conference on Information Sciences and Systems CISS. 2008:984–989.
- [64] S. A. Jafar, A. Goldsmith. On optimality of beamforming for multiple antenna systems with imperfect feedback[C]//Proceedings. 2001 IEEE International Symposium on Information Theory (IEEE Cat. No.01CH37252). 2001:321–. doi:10.1109/ISIT.2001.936184.
- [65] Wittneben. Optimal predictive TX combining diversity in correlated fading for microcellular mobile radio applications[C]//Proceedings of GLOBECOM 1995 Mini. 1995:48–54. doi:10.1109/CTMC.1995.502930.
- [66] A. Pascual Iserte, A. I. Perez-Neira, M. A. Lagunas Hernandez. Exploiting transmission spatial diversity in frequency selective systems with feedback channel[C]//2003 IEEE International Conference on Acoustics, Speech, and Signal Processing, 2003. Proceedings. (ICASSP '03). volume 4. 2003:IV–85. doi:10.1109/ICASSP.2003.1202547.
- [67] J. Park, C. Bruno. Multi-user linear precoding for multi-polarized massive MIMO system under imperfect CSIT[J]. IEEE Transactions on Wireless Communications, 2015. 14(5):2532–2547. doi:10.1109/TWC.2014.2388207.
- [68] M. Pei, J. Wei, K. Wong, X. Wang. Masked beamforming for multiuser MIMO wiretap channels with imperfect CSI[J]. IEEE Transactions on Wireless Communications, 2012. 11(2):544–549. doi:10.1109/TWC.2011.120511.110567.
- [69] A. Lozano, A. M. Tulino, S. Verdu. Optimum power allocation for parallel gaussian channels with arbitrary input distributions[J]. IEEE Transactions on Information Theory, 2006. 52(7):3033–3051. doi:10.1109/TIT.2006.876220.
- [70] Lamarca, Meritxell. Linear precoding for mutual information maximization in MIMO systems[C]//2009 6th International Symposium on Wireless Communication Systems. 2009:26–30. doi:10.1109/ISWCS.2009.5285230.
- [71] C. Xiao, Y. R. Zheng, Z. Ding. Globally optimal linear precoders for finite alphabet signals over complex vector gaussian channels[J]. IEEE Transactions on Signal Processing, 2011. 59(7):3301–3314. doi:10.1109/TSP.2011.2140112.
- [72] W. Cao, A. Dytso, M. Faub, G. Feng, H. V. Poor. Robust power allocation for parallel gaussian channels with approximately gaussian input distributions[J]. IEEE Transactions on Wireless Communications, 2020. 19(6):3685–3699. doi:10.1109/TWC.2020.2975784.
- [73] X. Zhu, Z. Meng, J. Xie, X. Tu, W. Zeng. Performance analysis of linear precoding in massive MIMO systems with finite-alphabet inputs[J]. IEEE Access, 2019. 7:85696–85704. doi:10.1109/ACCESS.2019.

- 2925939.
- [74] M. S. Sim, J. Park, C. Chae, R. W. Heath. Compressed channel feedback for correlated massive MIMO systems[J]. *Journal of Communications and Networks*, Feb. 2016. 18(1):95–104.
- [75] L. Wei, R. Q. Hu, Y. Qian, G. Wu. Key elements to enable millimeter wave communications for 5G wireless systems[J]. *IEEE Wireless Communications*, Dec. 2014. 21(6):136–143.
- [76] A. Alkhateeb, O. E. Ayach, G. Leusz, R. W. Heath. Channel estimation and hybrid precoding for millimeter wave cellular systems[J]. *IEEE Journal on Selected Areas in Communications*, Oct. 2014. 8(5):831–846.
- [77] D. Zhu, J. Choi, R. W. Heath. Two-dimensional AoD and AoA acquisition for wideband mmwave systems with cross-polarized MIMO[J]. *arXiv preprint arXiv:1702.02497*, Feb. 2017.
- [78] Ayachm O. E., S. Rajagopal, S. Abu-Surra, Z. Pi, R. W. Heath. Spatially sparse precoding in millimeter wave MIMO systems[J]. *IEEE Transactions on Wireless Communications*, Mar. 2014. 13(3):1499–1513.
- [79] Y. Ding, B. D. Rao. Dictionary learning-based sparse channel representation and estimation for FDD massive MIMO systems[J]. *IEEE Transactions on Wireless Communications*, Aug. 2018. 17(8):5437–5451.
- [80] Z. Zhang, Y. Xu, J. Yang, X. Li, D. Zhang. A survey of sparse representation: Algorithms and applications[J]. *IEEE Access*, 2015. 3:490–530.
- [81] T. Cai, L. Wang. Orthogonal matching pursuit for sparse signal recovery with noise[J]. *IEEE Transactions on Information Theory*, July 2011. 57(7):4680–4688.
- [82] M. J. Jahromi, M. H. Kahaei. Two-dimensional iterative adaptive approach for sparse matrix solution[J]. *IET Electronics Letters*, Jan. 2014. 50(1):45–47.
- [83] M. J. Jahromi, M. H. Kahaei. Complex two-dimensional TNIPM for l_1 norm-based sparse optimization to collocated MIMO radar[J]. *IEEE Transactions on Electrical Electronic Engineering*, 2016. 11:228–235.
- [84] A. Liu, V. K. N. Lau, W. Dai. Exploiting burst-sparsity in massive MIMO with partial channel support information[J]. *IEEE Transactions on Wireless Communications*, Nov. 2016. 15(11):7820–7830.
- [85] Y. Linde, A. Buzo, R. Gray. An algorithm for vector quantizer design[J]. *IEEE Transactions Communications*, Jan. 1980. 28(1):84–95.
- [86] R1-1709232. WF on type I and II CSI codebooks[J]. 3GPP TSG RAN WG1 (TSGR1) #89, May 2017.
- [87] R1-1905025. CSI enhancement for MU-MIMO support[J]. 3GPP TSG RAN WG1 Meeting #96, April 2019.
- [88] 3GPP. Study on channel model for frequencies from 0.5 to 100 GHz[J]. 3GPP, Technical Report 38.901, ver. 16.0.0, 2019.
- [89] Y. Sakata, T. Murakami, Y. Takatori, M. Mizoguchi, F. Maehara. Simple resource allocation scheme for heterogeneous traffic in MU-MIMO-OFDMA systems[C]//The 20th Asia-Pacific Conference on Commu-

- nication (APCC2014). 2014:189–193. doi:10.1109/APCC.2014.7091630.
- [90] R. W. Heath, N. Gonzalez-Prelcic, S. Rangan, W. Roh, A. M. Sayeed. An overview of signal processing techniques for millimeter wave MIMO systems[J]. *IEEE Journal of Selected Topics in Signal Processing*, 2016. 10(3):436–453. doi:10.1109/JSTSP.2016.2523924.
- [91] X. Gao, L. Dai, S. Zhou, A. M. Sayeed, L. Hanzo. Beamspace channel estimation for wideband millimeter-wave MIMO with lens antenna array[C]//2018 IEEE International Conference on Communications (ICC). 2018:1–6. doi:10.1109/ICC.2018.8422131.
- [92] H. Tataria, P. J. Smith, P. A. Dmochowski. On the general analysis of coordinated regularized zero-forcing precoding: An application to two-tier small-cell networks[J]. *IEEE Transactions on Communications*, 2017. 65(7):3133–3150. doi:10.1109/TCOMM.2017.2695199.
- [93] I. C. Wong, O. Oteri, W. Mccoy. Optimal resource allocation in uplink SC-FDMA systems[J]. *IEEE Transactions on Wireless Communications*, 2009. 8(5):2161–2165. doi:10.1109/TWC.2009.061038.
- [94] S. Sun, S. Moon. Practical scheduling algorithms with contiguous resource allocation for next-generation wireless systems[J]. *IEEE Wireless Communications Letters*, 2021. 10(4):725–729. doi:10.1109/LWC.2020.3041836.
- [95] M. Marc, M. Andrea. *Information, Physics, and Computation*[M]. 1st ed. Oxford University press, 2009.
- [96] P. Petar, T. Kasper Floe, O. Simeone, G. Durisi. 5G wireless network slicing for eMBB, URLLC, and mMTC: A communication-theoretic view[J]. *IEEE Access*, 2018. 6:55765–55779. doi:10.1109/ACCESS.2018.2872781.
- [97] C. S. Bontu, J. Ghimire, A. El-Keyi. Optimum resource allocation in MU-MIMO OFDMA wireless systems[C]//2020 IEEE 91st Vehicular Technology Conference (VTC2020-Spring). 2020:1–5. doi:10.1109/VTC2020-Spring48590.2020.9129372.
- [98] E. C. Song, G. Yue. Finite blocklength analysis for coded modulation with applications to link adaptation[C]//2019 IEEE Wireless Communications and Networking Conference (WCNC). 2019:1–7.
- [99] M. Zohdy, E. C. Song, G. Yue. Finite blocklength analysis of coded modulation for block fading channels with linear precoding[C]//2020 European Conference on Networks and Communications (EuCNC). 2020:311–315.
- [100] D. Duyck, J. J. Boutros, M. Moeneclaey. Precoding for outage probability minimization on block fading channels[J]. *IEEE Transactions on Information Theory*, 2013. 59(12):8250–8266.
- [101] Q. Shi, M. Razaviyayn, Z. Luo, C. He. An iteratively weighted MMSE approach to distributed sum-utility maximization for a MIMO interfering broadcast channel[J]. *IEEE Transactions on Signal Processing*, 2011. 59(9):4331–4340.
- [102] R1-031303. System-level evaluation of OFDM - further considerations[J]. 3GPP TSG RAN WG1 (TSGR1) #35, Nov. 2003.

- [103] Z. Hanzaz, H. D. Schotten. Analysis of effective SINR mapping models for MIMO OFDM in LTE system[C]//2013 9th International Wireless Communications and Mobile Computing Conference (IWCMC). 2013:1509–1515.
- [104] S. N. Donthi, N. B. Mehta. An accurate model for EESM and its application to analysis of CQI feedback schemes and scheduling in LTE[J]. *IEEE Transactions on Wireless Communications*, 2011. 10(10):3436–3448.
- [105] E. Westman. Calibration and evaluation of the exponential effective SINR mapping EESM in 802.16[D]. :Masters thesis, The Royal Institute of Technology (KTH), Sep. 2006.
- [106] J. Francis, N. B. Mehta. EESM-based link adaptation in OFDM: Modeling and analysis[C]//2013 IEEE Global Communications Conference (GLOBECOM). 2013:3703–3708.
- [107] E. Telatar. Capacity of multi-antenna gaussian channels[J]. *European Transactions on Telecommunications*, 1999. 10(6):585–595.
- [108] J. Pan, W. Ma, X. Xia, Y. Tian. Wmmse-based multiuser mimo beamforming: A practice-oriented design and lte system performance evaluation[C]//2015 IEEE 16th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC). 2015:435–439. doi:10.1109/SPAWC.2015.7227075.
- [109] J. Nocedal, S. Wright. *Numerical Optimization*[M]. 2nd ed. Springer, New York, NY, 2008.
- [110] M. Bennis, M. Debbah, H. V. Poor. Ultrareliable and low-latency wireless communication: Tail, risk, and scale[J]. *Proceedings of the IEEE*, 2018. 106(10):1834–1853. doi:10.1109/JPROC.2018.2867029.
- [111] Y. Cao, T. Lv, Z. Lin, W. Ni. Delay-constrained joint power control, user detection and passive beamforming in intelligent reflecting surface-assisted uplink mmWave system[J]. *IEEE Transactions on Cognitive Communications and Networking*, 2021. 7(2):482–495. doi:10.1109/TCCN.2021.3064973.
- [112] M. T. Kabir, C. Masouros. Delay-constrained beamforming and resource allocation in full duplex systems[J]. *IEEE Transactions on Vehicular Technology*, 2020. 69(3):3476–3480. doi:10.1109/TVT.2020.2967243.
- [113] T. X. Tran, A. Hajisami, D. Pompili. QuaRo: A queue-aware robust coordinated transmission strategy for downlink C-RANs[C]//2016 13th Annual IEEE International Conference on Sensing, Communication, and Networking (SECON). 2016:1–9. doi:10.1109/SAHCN.2016.7733010.
- [114] V. K. N. Lau, F. Zhang, Y. Cui. Low complexity delay-constrained beamforming for multi-user MIMO systems with imperfect CSIT[J]. *IEEE Transactions on Signal Processing*, 2013. 61(16):4090–4099. doi:10.1109/TSP.2013.2264058.
- [115] E. Altman. *Constrained markov decision processes*, RR-2574[R]. :INRIA, May 1995.
- [116] D. P. Bertsekas. *Dynamic Programming and Optimal Control*[M]. 3rd ed. Athena Scientific, 2005.
- [117] J. Achiam, D. Held, A. Tamar, P. Abbeel. *Constrained policy optimization*[J]. 2017. 70:22–31.
- [118] S. M. Ross. *Introduction to Probability Models*[M]. 6th ed. Academic Press, 1997.

- [119] R. S. Sutton, D. McAllester, S. Singh, Y. Mansour. Policy gradient methods for reinforcement learning with function approximation[C]//Advances in Neural Information Processing Systems. volume 12. 2000:1057–1063.
- [120] R. S. Sutton, A. G. Barto. Reinforcement Learning: An Introduction[M]. 2nd ed. The MIT Press, 2018.
- [121] D. Silver, G. Lever, N. Heess, T. Degris, D. Wierstra, M. A. Riedmiller. Deterministic policy gradient algorithms[C]//Proceedings of the 31th International Conference on Machine Learning. volume 32. 2014:387–395.
- [122] S. Zou, T. Xu, Y. Liang. Advances in neural information processing systems[C]//Proceedings of the 31th International Conference on Machine Learning. 2019:8668–8678.
- [123] M. Razaviyayn. Successive convex approximation: Analysis and applications[D]. :Ph.D. Dissertation, 2014.
- [124] J. Achiam A. Ray, D. Amodei. Benchmarking safe exploration in deep reinforcement learning[J]. 2019.
- [125] A. Ruszczyński. Feasible direction methods for stochastic programming problems[J]. Mathematical Programming, 1980. 19(1):220–229.
- [126] A. Y. Mitrophanov. Sensitivity and convergence of uniformly ergodic markov chains[J]. Journal of Applied Probability, 2005. 42(4).
- [127] A. Liu, V. K. N. Lau, B. Kananian. Stochastic successive convex approximation for non-convex constrained stochastic optimization[J]. IEEE Transactions on Signal Processing, 2019. 67(16):4189–4203.

附录

A 对引理1的证明

证明依赖于文献[125]中的Lemma1, 我们重写如下:

引理 2. 记 $(\Omega, \mathcal{F}, \mathbb{P})$ 是概率空间, $\{\mathcal{F}_t\}$ 是 \mathcal{F} 的 σ -域的递减序列。记 $\{\eta^t\}$ 和 $\{z^t\}$ 是 $\{\mathcal{F}_t\}$ 可测的随机向量序列, 并满足如下关系:

$$\begin{aligned} z^{t+1} &= \Pi_{\mathcal{Z}}(z^t + \alpha_t(\zeta^t - z^t)), z^0 \in \mathcal{Z}, \\ \mathbb{E}[\zeta^t | \mathcal{F}_t] &= \eta^t + b^t, \end{aligned} \tag{6.1}$$

其中 $\alpha_t \leq 0$; \mathcal{Z} 是一个凸闭集; $\Pi_{\mathcal{Z}}(\cdot)$ 表示在 \mathcal{Z} 上的投影。若

- (a). $\{\eta^t\}$ 所有的累积点都以概率1(w.p.1.)属于 \mathcal{Z} ;
- (b). 存在常数 C 使得 $\mathbb{E}[\|\zeta^t\|^2 | \mathcal{F}_t] \leq C$ 对所有的 $t \leq 0$ 成立;
- (c). $\sum_{t=0}^{\infty} \mathbb{E}[(\alpha_t)^2 + \alpha_t \|b^t\|] < \infty$;
- (d). $\sum_{t=0}^{\infty} \alpha_t = \infty$;
- (e). $\|\eta^{t+1} - \eta^t\| / \alpha_t \rightarrow 0$ w.p.1.;

成立, 则 $z^t - \eta^t \rightarrow 0$ w.p.1.。

基于上述引理2, 我们下面证明引理1。

A.1 证明式(5.25a)

根据假设2以及 $C_i(s, a)$ 的有界性, 可以看出引理2的条件(a)、(b)和(d)是满足的。对条件(c), 计算 t 时刻的随机偏差:

$$\begin{aligned} \|b_i^t\| &= \left| \mathbb{E}[\tilde{J}_i^t] - J_i(\theta_t) \right| \\ &\leq \left| \mathbb{E}[\tilde{J}_i^t] - \mathbb{E}[\tilde{J}_i^t] \right| + \left| \mathbb{E}[\tilde{J}_i^t] - J_i(\theta_t) \right|, \forall i, \end{aligned} \tag{6.2}$$

其中 \tilde{J}_i^t 是基于式(5.16)从另一个辅助轨迹 \tilde{p}_s 中计算得到的, 即

$$\begin{aligned} \tilde{p}_s = \{ & \tilde{s}_0, \tilde{a}_0, \tilde{s}_1, \dots, \tilde{s}_{t-2T-\tau+1}, \tilde{a}_{t-2T-\tau+1}, \tilde{s}_{t-2T-\tau+2}, \dots, \\ & \tilde{s}_{t-2T+1}, \tilde{a}_{t-2T+1}, \tilde{s}_{t-2T+2}, \dots, \tilde{s}_t, \tilde{a}_t \}. \end{aligned} \quad (6.3)$$

更具体地, 上述辅助轨迹是这么产生的, 在 $l = 0, 1, \dots, t - 2T - \tau$ 时, 按照 $a_l \sim \pi_{\theta^l}$ 采样; 在 $l = t - 2T - \tau + 1, \dots, t$ 时, 固定策略 $\pi_{\theta^{t-2T-\tau}}$, 并按照 $a_l \sim \pi_{\theta^{t-2T-\tau}}$ 采样. 基于此, 有

$$\mathbb{E} \left[\tilde{J}_i^t \right] = \frac{1}{2T} \sum_{l=1}^{2T} \int_{s \in \mathcal{S}} \mathbf{P} (S_{t-2T+l} = ds) \int_{a \in \mathcal{A}} \pi_{\theta^{t-2T+l}} (da | s) C'_i(s, a), \quad (6.4)$$

和

$$\mathbb{E} \left[\tilde{J}_i^t \right] = \frac{1}{2T} \sum_{l=1}^{2T} \int_{s \in \mathcal{S}} \mathbf{P} (\tilde{S}_{t-2T+l} = ds) \int_{a \in \mathcal{A}} \pi_{\theta^{t-2T-\tau}} (da | s) C'_i(s, a), \quad (6.5)$$

且

$$J_i(\theta_t) = \int_{s \in \mathcal{S}} \mathbf{P}_{\pi_{\theta^t}}(ds) \int_{a \in \mathcal{A}} \pi_{\theta^t}(da | s) C'_i(s, a). \quad (6.6)$$

因此, 有

$$\begin{aligned} & \left| \mathbb{E} \left[\tilde{J}_i^t \right] - \mathbb{E} \left[\tilde{J}_i^t \right] \right| \\ & \leq \frac{1}{2T} \sum_{l=1}^{2T} \left| \int_{s \in \mathcal{S}} \mathbf{P} (S_{t-2T+l} = ds) \int_{a \in \mathcal{A}} \pi_{\theta^{t-2T+l}} (da | s) C'_i(s, a) \right. \\ & \quad - \int_{s \in \mathcal{S}} \mathbf{P} (\tilde{S}_{t-2T+l} = ds) \int_{a \in \mathcal{A}} \pi_{\theta^{t-2T+l}} (da | s) C'_i(s, a) \\ & \quad + \int_{s \in \mathcal{S}} \mathbf{P} (\tilde{S}_{t-2T+l} = ds) \int_{a \in \mathcal{A}} \pi_{\theta^{t-2T+l}} (da | s) C'_i(s, a) \\ & \quad \left. - \int_{s \in \mathcal{S}} \mathbf{P} (\tilde{S}_{t-2T+l} = ds) \int_{a \in \mathcal{A}} \pi_{\theta^{t-2T-\tau}} (da | s) C'_i(s, a) \right| \\ & \stackrel{a}{\leq} \frac{1}{2T} \sum_{l=1}^{2T} \int_{s \in \mathcal{S}} \left| \mathbf{P} (S_{t-2T+l} = ds) - \mathbf{P} (\tilde{S}_{t-2T+l} = ds) \right| \\ & \quad \int_{a \in \mathcal{A}} \pi_{\theta^{t-2T+l}} (da | s) |C'_i(s, a)| \\ & \quad + \frac{1}{2T} \sum_{l=1}^{2T} \int_{s \in \mathcal{S}} \mathbf{P} (\tilde{S}_{t-2T+l} = ds) \\ & \quad \int_{a \in \mathcal{A}} |\pi_{\theta^{t-2T+l}} (da | s) - \pi_{\theta^{t-2T-\tau}} (da | s)| |C'_i(s, a)| \\ & = \frac{1}{2T} \sum_{l=1}^{2T} O \left(\left\| \mathbf{P} (S_{t-2T+l} \in \cdot) - \mathbf{P} (\tilde{S}_{t-2T+l} \in \cdot) \right\|_{\text{TV}} \right. \\ & \quad \left. + \left\| \theta^{t-2T+l} - \theta^{t-2T-\tau} \right\| \right), \end{aligned} \quad (6.7)$$

其中^a根据柯西-施瓦茨 (Cauchy-Schwarz) 不等式得到; 最后的等式来自假设1。进一步地, 由于参数空间 Θ 是紧集, 有

$$\|\theta^{t-2T+l} - \theta^{t-2T-\tau}\| = O\left(\sum_{t'=t-2T-\tau+1}^{t-2T+l} \beta_{t'}\right) \quad (6.8)$$

和

$$\begin{aligned} & \left\| \mathbf{P}(S_{t-2T+l} \in \cdot) - \mathbf{P}(\tilde{S}_{t-2T+l} \in \cdot) \right\|_{TV} \\ &= \int_{s' \in \mathcal{S}} \left| \mathbf{P}(S_{t-2T+l} = ds') - \mathbf{P}(\tilde{S}_{t-2T+l} = ds') \right| \\ &= \int_{s' \in \mathcal{S}} \left| \int_{s \in \mathcal{S}} \mathbf{P}(S_{t-2T+l-1} = ds) K_{\theta^{t-2T+l-1}}(s, ds') \right. \\ & \quad \left. - \int_{s \in \mathcal{S}} \mathbf{P}(\tilde{S}_{t-2T+l-1} = ds) K_{\theta^{t-2T-\tau}}(s, ds') \right| \\ &\stackrel{b}{\leq} \int_{s' \in \mathcal{S}} \int_{s \in \mathcal{S}} \left| \mathbf{P}(S_{t-2T+l-1} = ds) - \mathbf{P}(\tilde{S}_{t-2T+l-1} = ds) \right| \\ & \quad + \int_{s' \in \mathcal{S}} \int_{s \in \mathcal{S}} \mathbf{P}(\tilde{S}_{t-2T+l-1} = ds) \\ & \quad \left| K_{\theta^{t-2T+l-1}}(s, ds') - K_{\theta^{t-2T-\tau}}(s, ds') \right|, \\ &= \left\| \mathbf{P}(S_{t-2T+l-1} \in \cdot) - \mathbf{P}(\tilde{S}_{t-2T+l-1} \in \cdot) \right\|_{TV} \\ & \quad + O\left(\sum_{t'=t-2T-\tau+1}^{t-2T+l-1} \beta_{t'}\right), \end{aligned} \quad (6.9)$$

其中

$$K_{\theta}(s, ds') \triangleq \int_{a \in \mathcal{A}} P(ds' | s, a) \pi_{\theta}(da | s) \quad (6.10)$$

定义为转状态转移核 (Kernel) 函数; ^b推导与^a类似。因此, 基于上式的递推关系, 有

$$\begin{aligned} & \left\| \mathbf{P}(S_{t-2T+l} \in \cdot) - \mathbf{P}(\tilde{S}_{t-2T+l} \in \cdot) \right\|_{TV} \\ &= O\left(\sum_{j=t-2T-\tau+1}^{t-2T+l-1} \sum_{t'=t-2T-\tau+1}^j \beta_{t'}\right) = O\left((T-1+\tau) \sum_{t'=t-2T-\tau+1}^t \beta_{t'}\right). \end{aligned} \quad (6.11)$$

结合式(6.8)和(6.11), 有

$$\left| \mathbb{E}[\tilde{J}_i^t] - \mathbb{E}[J_i^t] \right| = O\left((T+\tau) \sum_{t'=t-2T-\tau+1}^t \beta_{t'}\right). \quad (6.12)$$

使用相似的技巧, 可以得到

$$\begin{aligned} & \left| \mathbb{E}[\tilde{J}_i^t] - J_i(\theta^t) \right| \\ &= \frac{1}{2T} \sum_{l=1}^{2T} O\left(\left\| \mathbf{P}(\tilde{S}_{t-2T+l} \in \cdot) - \mathbf{P}_{\pi_{\theta^t}} \right\|_{TV} + \|\theta^t - \theta^{t-2T-\tau}\|\right), \end{aligned} \quad (6.13)$$

其中

$$\begin{aligned}
& \left\| \mathbf{P} \left(\tilde{S}_{t-2T+l} \in \cdot \right) - \mathbf{P}_{\pi_{\theta^t}} \right\|_{\text{TV}} \\
& \leq \left\| \mathbf{P} \left(\tilde{S}_{t-2T+l} \in \cdot \right) - \mathbf{P}_{\pi_{\theta^{t-2T-\tau}}} \right\|_{\text{TV}} + \left\| \mathbf{P}_{\pi_{\theta^{t-2T-\tau}}} - \mathbf{P}_{\pi_{\theta^t}} \right\|_{\text{TV}} \\
& \stackrel{c}{\leq} m_c \rho^{\tau+l} + \left(\left\lceil \log_{\rho} m_c^{-1} \right\rceil + \frac{1}{1-\rho} \right) \|K_{\theta^t} - K_{\theta^{t-2T-\tau}}\|_q.
\end{aligned} \tag{6.14}$$

上式中， $\stackrel{c}{\leq}$ 的推导来自文献[126]的Theorem 3.1；操作符 $\|A\|_q \triangleq \sup_{\|q\|_{\text{TV}}=1} \|qA\|_{\text{TV}}$ 。因此有

$$\begin{aligned}
& \|K_{\theta^t} - K_{\theta^{t-2T-\tau}}\|_q \\
& = \sup_{\|q\|_{\text{TV}}=1} \left\| \int_{s \in \mathcal{S}} q(\text{d}s) (K_{\theta^t} - K_{\theta^{t-2T-\tau}})(s, \cdot) \right\|_{\text{TV}} \\
& = \sup_{\|q\|_{\text{TV}}=1} \int_{s' \in \mathcal{S}} \left| \int_{s \in \mathcal{S}} q(\text{d}s) (K_{\theta^t} - K_{\theta^{t-2T-\tau}})(s, \text{d}s') \right| \\
& \leq \sup_{\|q\|_{\text{TV}}=1} \int_{s' \in \mathcal{S}} \int_{s \in \mathcal{S}} |q(\text{d}s)| |K_{\theta^t}(s, \text{d}s') - K_{\theta^{t-2T-\tau}}(s, \text{d}s')| \\
& = O \left(\sum_{t'=t-2T-\tau+1}^t \beta_{t'} \right).
\end{aligned} \tag{6.15}$$

考虑到 $\rho^{\tau+l} \leq \rho^{\tau}$ ，有

$$\left| \mathbb{E} \left[\tilde{J}_i^t \right] - J_i(\theta^t) \right| = O \left(\rho^{\tau} + \sum_{t'=t-2T-\tau+1}^t \beta_{t'} \right), \tag{6.16}$$

和

$$\|b_i^t\| = O \left(\rho^{\tau} + (T + \tau) \sum_{t'=t-2T-\tau+1}^t \beta_{t'} \right). \tag{6.17}$$

使得 $\rho^{\tau} = O(\frac{1}{t})$ 后，可得引理2的条件(c)满足。对条件(e)，有

$$\begin{aligned}
|J(\theta_t) - J(\theta_{t-1})| & \stackrel{a}{=} O \left(\|\mathbf{P}_{\pi_{\theta_t}} - \mathbf{P}_{\pi_{\theta_{t-1}}}\|_{\text{TV}} + \|\theta^t - \theta^{t-1}\| \right) \\
& \stackrel{b}{=} O \left(\|\theta^t - \theta^{t-1}\| \right) = O(\beta_t),
\end{aligned} \tag{6.18}$$

其中 $\stackrel{a}{=}$ 的推导与6.7类似； $\stackrel{b}{=}$ 的推导来自文献[126]的Theorem 3.1。结合假设2对步长 α_t 和 β_t 的规定，可得引理2的条件(e)满足。因此引理2的所有条件满足，式(5.25a)成立。

A.2 证明式(5.25b)

构造辅助的梯度估计函数：

$$\hat{g}_i^t = (1 - \alpha_t) \hat{g}_i^{t-1} + \alpha_t \tilde{g}_i^t, \tag{6.19}$$

其中

$$\tilde{g}_i^t = \frac{1}{T} \sum_{l=1}^T \hat{Q}_i^{t-2T+l}(s_{t-2T+l}, a_{t-2T+l}) \nabla \log \pi_{\theta^t}(a_{t-2T+l} | s_{t-2T+l}), \quad (6.20)$$

且

$$\hat{Q}_i^{t-2T+l}(s_{t-2T+l}, a_{t-2T+l}) = \sum_{l'=0}^{T-1} (C'_i(s_{t-2T+l+l'}, a_{t-2T+l+l'}) - J_i(\theta^t)). \quad (6.21)$$

注意上式与式(5.20)的区别是用 $J_i(\theta^t)$ 代替 \hat{J}_k^t 。我们分如下两步证明式(5.25b):

$$\text{Step.1: } \lim_{t \rightarrow \infty} \|\hat{g}_i^t - \nabla J_i(\theta^t)\|_2 = 0, \quad (6.22)$$

$$\text{Step.2: } \lim_{t \rightarrow \infty} \|\hat{g}_i^t - \tilde{g}_i^t\|_2 = 0.$$

首先, 用前小节中相似的方法可以证明 \tilde{g}_i^t 同样满足引理2的条件(a)、(b)、(d)和(e)。同时, 结合式(5.17), 我们有

$$\mathbb{E}[\tilde{g}_i^t] = \frac{1}{T} \sum_{l=1}^T \int_{s \in \mathcal{S}} \mathbf{P}(S_{t-2T+l} = ds) \int_{a \in \mathcal{A}} \pi_{\theta^{t-2T+l}}(da | s) \mathbb{E}[\hat{Q}_i^{t-2T+l}(s, a)] \nabla \log \pi_{\theta^t}(a | s), \quad (6.23)$$

和

$$\begin{aligned} & \|\mathbb{E}[\tilde{g}_i^t] - \nabla J_i(\theta^t)\|_2 \\ &= \frac{1}{T} \sum_{l=1}^T O\left(\|\mathbf{P}(S_{t-2T+l} \in \cdot) - \mathbf{P}_{\pi_{\theta^t}}\|_{\text{TV}} + \|\theta^t - \theta^{t-2T+l}\|_2 + \left|\mathbb{E}\left[\hat{Q}_i^{t-2T+l}(s, a)\right] - Q_{\pi_{\theta^t}}(s, a)\right|\right). \end{aligned} \quad (6.24)$$

上式等号成立利用了 $C_i(s, a)$ 和 $\nabla \log \pi_{\theta}(a | s)$ 是有界的假设。与式(6.9)和(6.14)类似, 我们有

$$O\left(\|\mathbf{P}(S_{t-2T+l} \in \cdot) - \mathbf{P}_{\pi_{\theta^t}}\|_{\text{TV}}\right) \leq O\left(\rho^\tau + (T + \tau) \sum_{l'=t-2T-\tau+1}^t \beta_{l'}\right), \quad (6.25)$$

和

$$\begin{aligned} & \mathbb{E}\left[\hat{Q}_i^{t-2T+l}(s, a)\right] \\ &= C'_i(s, a) - J_i(\theta^t) + \int_{\mathcal{S}} P(ds' | s, a) \int_{\mathcal{A}} \pi_{\theta^{t-2T+l+1}}(da' | s') C'_i(s', a') - J_i(\theta^t) \\ & \quad + \sum_{l'=2}^{T-1} \left(\int_{\mathcal{S}} \mathbf{P}(S_{t-2T+l+l'} = ds') \int_{\mathcal{A}} \pi_{\theta^{t-2T+l+l'}}(da' | s') C'_i(s', a') - J_i(\theta^t) \right). \end{aligned} \quad (6.26)$$

结合式(5.18)的定义, 有

$$\begin{aligned} & \left| \mathbb{E}\left[\hat{Q}_i^{t-2T+l}(s, a)\right] - Q_i^{\pi_{\theta^t}}(s, a) \right| \\ &= O\left(\|\theta^t - \theta^{t-2T+l+1}\|_2 + \sum_{l'=2}^{T-1} \left(\|\mathbf{P}(S_{t-2T+l+l'} \in \cdot) - \mathbf{P}_t(S_{l'} \in \cdot)\|_{\text{TV}} + \|\theta^t - \theta^{t-2T+l+l'}\|_2\right) + |M_3|\right), \end{aligned} \quad (6.27)$$

其中

$$\begin{aligned}
 |M_3| &= \left| \sum_{l'=T}^{\infty} \left(\int_{\mathcal{S}} \mathbf{P}_t(S_{l'} = ds') \int_{\mathcal{A}} \pi_{\theta^t}(da' | s') C'_i(s', a') - \int_{\mathcal{S}} \mathbf{P}_{\pi_{\theta^t}}(ds') \int_{\mathcal{A}} \pi_{\theta^t}(da' | s') C'_i(s', a') \right) \right| \\
 &= \sum_{l'=T}^{\infty} O\left(\|\mathbf{P}_t(S_{l'} \in \cdot) - \mathbf{P}_{\pi_{\theta^t}}\|_{TV}\right) = O(\rho^T),
 \end{aligned} \tag{6.28}$$

上式的推导也与6.7类似。更进一步的，与式(6.11)类似，我们有

$$\|\mathbf{P}(S_{t-2T+l+l'} \in \cdot) - \mathbf{P}_t(S_{l'} \in \cdot)\|_{TV} = O\left(\sum_{j=t-2T+l+1}^{t-2T+l+l'} \sum_{l'=j}^t \beta_{l'}\right) \leq O\left(T \sum_{l'=t-2T-\tau}^t \beta_{l'}\right). \tag{6.29}$$

结合上述推导，对式(6.24)，有

$$\|b_i^t\| = \|\mathbb{E}[\tilde{g}_i^t] - \nabla J_i(\theta^t)\|_2 = O\left(\rho^\tau + (T + \tau) \sum_{l'=t-2T-\tau+1}^t \beta_{l'} + \rho^T + T^2 \sum_{l'=t-2T-\tau+1}^t \beta_{l'}\right). \tag{6.30}$$

同样让 $\rho^\tau = O(\frac{1}{t})$ 后，可得引理2的条件(c)满足。因此式(6.22)中Step.1成立。对Step.2的证明，我们首先根据式(5.10)和(6.19)得到

$$\begin{aligned}
 \hat{g}_i^t &= \sum_{l'=0}^t \prod_{j=l'+1}^t (1 - \alpha_j) \alpha_{l'} \tilde{g}_i^{l'}, \\
 \hat{\tilde{g}}_i^t &= \sum_{l'=0}^t \prod_{j=l'+1}^t (1 - \alpha_j) \alpha_{l'} \tilde{\tilde{g}}_i^{l'},
 \end{aligned} \tag{6.31}$$

因此有

$$\begin{aligned}
 \|\hat{g}_i^t - \hat{\tilde{g}}_i^t\|_2 &\leq \sum_{l'=0}^t \prod_{j=l'+1}^t (1 - \alpha_j) \alpha_{l'} e_{l'} \leq \sum_{l'=0}^t (1 - \alpha_t)^{t-l'} \alpha_{l'} e_{l'} \\
 &= \sum_{l'=0}^{n_t} (1 - \alpha_t)^{t-l'} \alpha_{l'} e_{l'} + \sum_{l'=n_t+1}^t (1 - \alpha_t)^{t-l'} \alpha_{l'} e_{l'} \\
 &\leq e_{t_a} \frac{(1 - \alpha_t)^{t-n_t}}{\alpha_t} + e_{t_b} \frac{\alpha_{n_t+1}}{\alpha_t},
 \end{aligned} \tag{6.32}$$

其中 $e_{l'} = \|\tilde{g}_i^{l'} - \tilde{\tilde{g}}_i^{l'}\|_2 = O\left(\left|\hat{J}_i^{l'} - J_i(\theta^{l'})\right|\right)$; $n_t = (1 - \kappa - \epsilon)t$, $\epsilon \in (0, 1 - \kappa)$; $e_{t_a} = \max_{l'=0, \dots, n_t} e_{l'}$ 且 $e_{t_b} = \max_{l'=n_t+1, \dots, t} e_{l'}$ 。基于假设2，可以推出

$$\lim_{t \rightarrow \infty} \|\hat{g}_i^t - \hat{\tilde{g}}_i^t\|_2 = O\left(\lim_{t_b \rightarrow \infty} \left|\hat{J}_i^{t_b} - J_i(\theta^{t_b})\right|\right) = 0. \tag{6.33}$$

由此，式(6.22)中Step.2也成立。因此引理1得证。

B 对定理2的证明

我们首先给出如下引理。

引理 3. 在假设1、2和 $\Theta_A^* \cap \bar{\Theta}_C^* = \emptyset$ 均成立时，若记 $\{\theta^t\}_{t=1}^\infty$ 表示SCAPO算法产生的迭代参数序列，则有

$$\begin{aligned} \limsup_{t \rightarrow \infty} J(\theta^t) &\leq 0, \text{ w.p.1.} \\ \lim_{t \rightarrow \infty} \|\bar{\theta}^t - \theta^t\| &= 0, \text{ w.p.1.} \end{aligned} \quad (6.34)$$

成立。其中 $J(\theta) \triangleq \max_{i \in \{1, \dots, m\}} J_i(\theta)$ 表示一个最大约束函数。

这里对引理3作进一步说明： Θ_A^* 表示由SCAPO算法产生的一系列极限点集合； $\bar{\Theta}_C^*$ 表示由一个非可行初始点出发，SCAPO算法运行并收敛至的非理想驻点解集，即

$$\bar{\Theta}_C^* \triangleq \{\theta : J(\theta) > 0, \theta \in \Theta_C^*\}, \quad (6.35)$$

其中 Θ_C^* 是下述问题的驻点解获得：

$$\min_{\theta \in \Theta} J(\theta) \triangleq \max_{i \in \{1, \dots, m\}} J_i(\theta). \quad (6.36)$$

需要强调，我们虽然定义了 $\bar{\Theta}_C^*$ ，但只要通过选择避开 $\bar{\Theta}_C^*$ 及其附近的合适的初始点（通过多次随机选择，这是容易做到的），SCAPO算法仍能收敛至问题 \mathcal{P}_G 的KKT解。因此上述引理的含义是：只要选择合适初始点保证 $\Theta_A^* \cap \bar{\Theta}_C^* = \emptyset$ 成立，SCAPO算法几乎一定（Almost-surely）收敛到可行集中，且 $\bar{\theta}^t$ 和 θ^t 的差距以概率1趋于0。上述引理的证明在附录C中。

下面，我们将分两步，首先证明当定理2的条件均满足时引理3中的条件 $\Theta_A^* \cap \bar{\Theta}_C^* = \emptyset$ 以概率1成立；再基于引理3的结论证明定理2。

首先， $\bar{\Theta}_C^* = \emptyset$ 时，显然有 $\Theta_A^* \cap \bar{\Theta}_C^* = \emptyset$ 。因此我们只考虑 $\bar{\Theta}_C^* \neq \emptyset$ 的情况，首先定义 $\mathcal{L}(x) \triangleq \{\theta : J(\theta) \leq x\}$ 表示 $J(\theta)$ 在 x 处的下水平集（Sublevel）；定义 $x_C = \min_{\theta \in \bar{\Theta}_C^*} J(\theta)$ 。根据 $\bar{\Theta}_C^*$ 的定义，有 $x_C > 0$ 。由于我们选择的初始点是可行的，即 $J(\theta^0) < 0$ ，因此一定有 $\theta^0 \in \mathcal{L}(0.5x_C)$ 。定义 Θ_S 是 $\mathcal{L}(0.5x_C)$ 的一个紧子集，且其所有的点都与 θ^0 相连。因此有 $\Theta_S \cap \bar{\Theta}_C^* = \emptyset$ 。记 $\hat{x} = 0.25x_C$ 和 $\mathcal{L}(\hat{x}, \Theta_S) = \mathcal{L}(\hat{x}) \cap \Theta_S$ 。由于 $J(\theta)$ 是李普希茨连续的（假设1），因此存在一个常数 $L > 0$ 使得

$$\min_{\theta \in \mathcal{L}(\hat{x}, \Theta_S)} \|\theta - \theta^0\| \geq L(0.25x_C - J(\theta^0)) \geq 0.25Lx_C \quad (6.37)$$

成立，其中 $\partial\mathcal{X}$ 表示集合 \mathcal{X} 的边界。

与附录C中类似，记 $\mathcal{T}_\epsilon = \{t: J(\theta^t) \geq \epsilon, \theta^t \in \Theta_S\}$ 和 $\mathcal{T}'_\epsilon = \mathcal{T}_\epsilon \cap \{t \geq t_\epsilon\}$ 后，由于 $\Theta_S \cap \bar{\Theta}_C^* = \emptyset$ ，式(6.45)和(6.46)同样成立。假设我们选择的初始步长满足 $\beta_0 < 0.25Lx_C / (R_\Theta t_\epsilon)$ ，其中 $R_\Theta \triangleq \max_{\theta, y \in \Theta} \|\theta - y\|$ 表示 Θ 空间的直径（Diameter），则根据式(5.13)和(6.37)，一定有 $\theta^t \in \mathcal{L}(\hat{x}, \Theta_S)$ 对任意 $t \leq t_\epsilon$ 成立。再根据式(6.45)，可知 $J(\theta^t)$ 几乎一定是递减的。因此，结合 $J(\theta)$ 的李普希茨连续性，有

$$\min_{\theta \in \mathcal{L}(\hat{x}, \Theta_S), \theta' \in \partial\Theta_S} \|\theta - \theta'\| \geq 0.25Lx_C. \quad (6.38)$$

因此，我们总可以取足够小的 ϵ 使得 $J(\theta) > 3\epsilon, \forall \theta \in \partial\Theta_S, \hat{x} > 2\epsilon$ 和 $0.25Lx_C > 2\epsilon$ 满足。根据式(6.38)，一旦 $\theta^t \in \mathcal{L}(\hat{x}, \Theta_S)$ 对 $t \geq t_\epsilon$ 成立， θ^{t+1} 必然属于 Θ_S （因为 $\|\theta^{t+1} - \theta^t\| \leq O(\beta_t) < \epsilon$ 对足够大的 t_ϵ 总成立），此时有两种情况：1) $J(\theta^t) \geq \epsilon$ ，此时根据式(6.45)有 $J(\theta^{t+1}) < J(\theta^t)$ ，因此有 $\theta^{t+1} \in \mathcal{L}(\hat{x}, \Theta_S)$ w.p.1.; 2) $J(\theta^t) < \epsilon$ ，此时根据式(6.46)有 $J(\theta^{t+1}) < 2\epsilon$ ，因此有 $\theta^{t+1} \in \mathcal{L}(\hat{x}, \Theta_S)$ w.p.1.。

综上，不论 $t \geq t_\epsilon$ 或 $t < t_\epsilon$ 时，总是有 $\theta^t \in \mathcal{L}(\hat{x}, \Theta_S) \subset \Theta_S$ 。由于 $\Theta_S \cap \bar{\Theta}_C^* = \emptyset$ ，因此 $\theta^t \notin \bar{\Theta}_C^*, \forall t$ w.p.1.，第一步证毕。

对第二步，记 $\{\theta^{t_j}\}_{j=1}^\infty$ 表示算法产生并收敛至极限点 θ^* 且满足Slater条件的子序列。根据第一步的结论和引理3，有

$$\lim_{j \rightarrow \infty} \|\bar{\theta}^{t_j} - \theta^{t_j}\| = 0, \text{ w.p.1.}, \quad (6.39)$$

和

$$\bar{\theta}^{t_j} = \operatorname{argmin}_{\theta \in \Theta} \bar{J}_0^{t_j}(\theta) \quad (6.40)$$

$$\text{s.t. } \bar{J}_i^{t_j}(\theta) \leq x^{t_j}, i = 1, \dots, m,$$

其中 $\lim_{j \rightarrow \infty} x^{t_j} = 0, \text{ w.p.1.}$ 。再结合引理1和 $\bar{J}_i^t(\theta), \hat{J}_i(\theta), \forall i$ 的强凸性，可得

$$\theta^* = \operatorname{argmin}_{\theta \in \Theta} \hat{J}_0(\theta) \quad (6.41)$$

$$\text{s.t. } \hat{J}_i(\theta) \leq 0, i = 1, \dots, m.$$

因此 θ^* 满足问题 \mathcal{P}_G 的KKT条件。

C 对引理3的证明

首先证明 $\limsup_{t \rightarrow \infty} J(\theta^t) \leq 0$ w.p.1.。

定义 $\mathcal{T}_\epsilon = \{t : J(\theta^t) \geq \epsilon\}$, $\forall \epsilon > 0$, 下面考量集合 \mathcal{T}_ϵ 是否有限。

可以看出 \mathcal{T}_ϵ 是无限集与 $\liminf_{t \in \mathcal{T}_\epsilon, t \rightarrow \infty} \|\bar{\theta}^t - \theta^t\| > 0$ 是相互矛盾的。用反证法, 假设其是无限集且 $\liminf_{t \in \mathcal{T}_\epsilon, t \rightarrow \infty} \|\bar{\theta}^t - \theta^t\| = 0$ 。此时存在一个子序列 $t^j \in \mathcal{T}_\epsilon$ 满足 $\lim_{j \rightarrow \infty} \|\bar{\theta}^{t^j} - \theta^{t^j}\| = 0$ 。定义 θ° 表示 $\{\theta^{t^j}\}$ 的极限点, 此时根据 SCAPO 算法的更新规则, 有两种情况: 1) 如果 θ° 是问题 $\mathcal{P}_\mathcal{O}$ 的最优解, 显然 $J(\theta^\circ) \leq 0$; 2) 如果 θ° 是问题 $\mathcal{P}_\mathcal{F}$ 的最优解, 则可以证明其满足问题 (6.36) 的 KKT 条件, 再根据 $\Theta_A^* \cap \bar{\Theta}_C^* = \emptyset$, 可得 $J(\theta^\circ) \leq 0$ 。

这两种情况都与 \mathcal{T}_ϵ 的定义矛盾, 因此 $\liminf_{t \in \mathcal{T}_\epsilon, t \rightarrow \infty} \|\bar{\theta}^t - \theta^t\| > 0$, \mathcal{T}_ϵ 是有限集。这意味着, 存在一个足够大的 t_ϵ 使得

$$\|\bar{\theta}^t - \theta^t\| \geq \epsilon', \forall t \in \mathcal{T}'_\epsilon \quad (6.42)$$

成立, 其中 $\epsilon' > 0$ 是一个常数且 $\mathcal{T}'_\epsilon = \mathcal{T}_\epsilon \cap \{t \geq t_\epsilon\}$ 。

定义 $\bar{J}^t(\theta) \triangleq \max_{i \in \{1, \dots, m\}} \bar{J}_i^t(\theta)$ 。根据其强凸性, 有

$$\nabla^T \bar{J}_i^t(\theta^t) d^t \leq -\eta \|d^t\|^2 + \bar{J}_i^t(\bar{\theta}^t) - \bar{J}_i^t(\theta^t), \quad (6.43)$$

其中 $d^t = \bar{\theta}^t - \theta^t$; $\eta > 0$ 是一个常数。根据假设 1, $J_i(\theta)$ 的梯度是李普希茨连续的, 因此存在一个 $L_J > 0$ 使得

$$\begin{aligned} J_i(\theta^{t+1}) &\leq J_i(\theta^t) + \beta_t \nabla^T J_i(\theta^t) d^t + L_J (\beta_t)^2 \|d^t\|_2^2 \\ &= J(\theta^t) + L_J (\beta_t)^2 \|d^t\|_2^2 + J_i(\theta^t) - J(\theta^t) + \beta_t (\nabla^T \bar{J}_i^t(\theta^t) + \nabla^T J_i(\theta^t) - \nabla^T \bar{J}_i^t(\theta^t)) d^t \\ &\stackrel{c}{\leq} J(\theta^t) + J_i(\theta^t) - J(\theta^t) - \eta \beta_t \|d^t\|_2^2 + \beta_t (\bar{J}_i^t(\bar{\theta}^t) - \bar{J}_i^t(\theta^t)) + o(\beta_t) \\ &\stackrel{d}{\leq} J(\theta^t) - \eta \beta_t \|d^t\|_2^2 + o(\beta_t), \forall i = 1, \dots, m \end{aligned} \quad (6.44)$$

成立。其中 $o(\beta_t)$ 表示 $\lim_{t \rightarrow \infty} o(\beta_t) / \beta_t = 0$ 。

上式中, 不等式 $\stackrel{c}{\leq}$ 成立的原因是利用 $\lim_{t \rightarrow \infty} \|\nabla J_i(\theta^t) - \nabla \bar{J}_i^t(\theta^t)\| = 0$; $\stackrel{d}{\leq}$ 成立是因为 $J_i(\theta^t) \leq J(\theta^t)$ 、 $\liminf_{t \rightarrow \infty} J(\theta^t) - \bar{J}_i^t(\bar{\theta}^t) \geq 0$ 和 $\lim_{t \rightarrow \infty} \|J_i(\theta^t) - \bar{J}_i^t(\theta^t)\| = 0$ 。因此, 通过选择一个足够大的 t_ϵ , 我们可以得到以下不等式

$$\begin{aligned} J(\theta^{t+1}) - J(\theta^t) &\leq -\beta_t \bar{\eta} \|d^t\|_2^2 \\ &\leq -\beta_t \bar{\eta} \epsilon', \forall t \in \mathcal{T}'_\epsilon, \end{aligned} \quad (6.45)$$

对某个 $\bar{\eta} > 0$ 成立。而且由于 $J(\theta)$ 是李普希茨连续的, 有

$$|J(\theta^{t+1}) - J(\theta^t)| \leq O(\|\theta^{t+1} - \theta^t\|_2) \leq O(\beta_t) < \epsilon, \forall t > t_\epsilon. \quad (6.46)$$

式(6.45)表示当 $J(\theta^t) \geq \epsilon$ 和 $t \geq t_\epsilon$ 时几乎一定是递减的。因此再结合式(6.46), 有

$$J(\theta^t) \leq 2\epsilon, \forall t \geq t_\epsilon. \quad (6.47)$$

由于上式对任意 $\epsilon > 0$ 成立, 因此 $\limsup_{t \rightarrow \infty} J(\theta^t) \leq 0$ 。

再根据 $\bar{J}_i^t(\theta)$ 的强凸性、 $J_i(\theta)$ 的有界性、 $\nabla J_i(\theta)$ 的李普希茨连续性和 $\sum_t \beta_t = \infty$, 我们可根据文献[127]中Appendix D-2的结论得到 $\lim_{t \rightarrow \infty} \|\bar{\theta}^t - \theta^t\| = 0$, w.p.1.。因此引理3得证。

攻读硕士学位期间的研究成果

已发表或录用的学术论文

1. **G. Huang**, A. Liu, and M. J. Zhao, “Two-stage adaptive and compressed CSI feedback for FDD massive MIMO,” *IEEE Transactions on Vehicular Technology*, vol. 70, no. 9, pp. 9602-9606, 2021. (SCI, 正文第二章)
2. C. Tian, **G. Huang**, A. Liu, and W. Luo, “Delay-aware power control for downlink multi-user MIMO via constrained deep reinforcement learning,” *2021 IEEE Global Communications Conference (GLOBECOM)*, 2021. (正文第五章)
3. C. Tian, A. Liu, **G. Huang**, and W. Luo, “Successive convex approximation based off-policy optimization for constrained reinforcement learning,” *IEEE Transactions on Signal Processing*, 2021. (SCI, 正文第五章)

已结题或在研的科研项目

1. 2019.09-2020.04, 华为公司合作项目-FDD高精度信道反馈, 项目主要参与成员。
2. 2020.04-2021.06, 华为公司合作项目-TDD下行预编码设计, 项目主要参与成员。
3. 2020.12-2021.06, 华为公司合作项目-TDD非天选信道反馈增强, 项目参与成员。
4. 2021.06-2021.12, 华为公司合作项目-5G多天线智能MU调度, 项目主要参与成员。
5. 2019.06-2019.12, 国防科技项目-非线性接收机敏感性测试方法, 项目参与成员。
6. 2020.04-2021.07, 西湖大学合作项目-红外微纳复眼结构设计, 项目主要参与成员。