Joint Design of Wireless Fronthaul and Access Links in Massive MIMO CRANs

YINGJIA HUANG

Joint Design of Wireless Fronthaul and Access Links in Massive MIMO CRANs

Yingjia Huang

This thesis is submitted for the degree of
Doctor of Philosophy

Department of Engineering
Faculty of Science
University of Durham
England

April 2023

Joint Design of Wireless Fronthaul and Access Links in Massive MIMO CRANs

Yingjia Huang

Submitted for the degree of Doctor of Philosophy
April 2023

Abstract

Cloud radio access network (CRAN) has emerged as a promising mobile network architecture for the current 5th generation (5G) and beyond networks. This thesis focuses on novel architectures and optimization approaches for CRAN systems with massive multiple-input multiple-output (MIMO) enabled in the wireless fronthaul link. In particular, we propose a joint design of wireless fronthaul and access links for CRANs and aim to maximize the network spectral efficiency (SE) and energy efficiency (EE).

Regarding downlink transmission in massive MIMO CRANs, the precoding designs of the access link are optimized by accounting for both perfect instantaneous channel state information (CSI) and stochastic CSI of the access link separately. The system design adopts a decompress-and-forward (DCF) scheme at the remote radio heads (RRHs), with optimization of the multivariate compression covariance noise. Constrained by the maximum power budgets set for the central unit (CU) and RRHs, we aim to maximize the network sum-rate and minimize the total transmit power for all user equipments (UEs). Moreover, we present a separate optimization design and compare its performance, feasibility, and computational efficiency with the proposed joint design. Considering the uplink transmission, we utilize a compress-and-forward (CF) scheme at the RRHs. Assuming that perfect CSI is available at the CU, our objective is to optimize the precoding matrix of the access link while adopting conventional precoding methods for the fronthaul link. This thesis also proposes an unmanned aerial vehicle (UAV)-enabled CRAN architecture with

a massive MIMO CU as a supplement system to the terrestrial communication networks. The locations of UAVs are optimized along with compression noise, precoding matrices, and transmit power. To tackle the non-convex optimization problems described above, we employ efficient iterative algorithms and conduct a thorough exploration of practical simulations, yielding promising results that outperform benchmark schemes.

In summary, this thesis explores future wireless CRAN architectures, leveraging promising technologies including massive MIMO and UAV-enabled communications. Furthermore, this work presents comprehensive optimization designs aimed at further enhancing the network efficiency.

Declaration

The work in this thesis is based on research carried out at the Department of Engineering, Durham University, England. No part of this thesis has previously been submitted for a degree in this or any other institution.

Acknowledgements

This thesis would not be finished without the contribution of many people.

First and foremost, I would like to express sincere gratitude to my supervisor, Dr Aissa Ikhlef, who has continuously provided me with invaluable guidance and support throughout my entire PhD study. His deep expertise in the field and rigorous academic attitude have laid a solid foundation for my research and will be a treasure for my future career. I want to extend a special thanks to my reviewer, Prof. Hongjian Sun, who has offered me valuable advice and supportive encouragement in every stage of my PhD. I would also like to thank Durham University for providing me with this PhD position and doctoral studentship to support my research. Dr Dick Maryopi and Hamilton HPC service team have also contributed a lot to my research conscientiously and responsibly.

My deepest appreciation is to my parents, who have always fully supported and respected me as an independent adult, no matter how small I was. Love would not exist without strength, honesty, self-awareness, and endurance. Focusing on the process not the result is one of many things I learned from their practice to become who I am today. If I could ever achieve or have achieved anything in my life, all credit belongs to my mum and dad.

It has been tough for me to continue PhD study during the pandemic. I would like to thank all my friends who have been with me and supported me to overcome the anxiety and difficulty of an unpredictable future. Their presence has made my academic journey much more colorful and fulfilling.

Today is ChúXī, the eve of YúanDàn (Chinese New Year). This is the day we call “out with the old, in with the new”. It has been a long journey and now it has finally come to an end. I could not have imagined how bittersweet this would taste at the beginning, but it has become the past and also the most invaluable treasure for my future life.

21/January/2023

Contents

List of Figures

List of Tables

List of Algorithms

Nomenclature

Mathematical Notations

$(.)^*$	Complex conjugate
$(.)^T$	Transpose
$(.)^H$	Hermitian transpose
i	Imaginary unit i $= \sqrt{-1}$
$\mathbf{0}$	Zero matrix
$\mathbf{I}$	Identity matrix
$\mathbb{E}[.]$	Expectation operator
$\mathbb{C}^{m\times n}$	Complex coordinate space of dimension $m \times n$
$\mathbb{R}^{m\times n}$	Real coordinate space of dimension $m \times n$
$\mathcal{CN}(m, n)$	Complex normal distribution
$\| \cdot \|$	Norm
$\mathrm{tr}(.)$	Trace
$\mathrm{diag}\,(.)$	Multiplication by an identity matrix
$\log\det(.)$	Log-determinant
$I(\mathbf{x};\mathbf{y})$	Mutual information
$I(\mathbf{x};\mathbf{y}\|\mathbf{z})$	Conditional mutual information
$h(\mathbf{x})$	Differential entropy
$h(\mathbf{x}\|\mathbf{y})$	Conditional differential entropy

Superscripts

$(.)^{(r)}, (.)^{(l)}$	Iteration indices

$(.)^{\star}$	Optimi solution of the variable
$(.)^{\text{lb}}$, $(.)^{\text{ub}}$	Lower bound, Upper bound

Subscripts

$(.)_i$	Index of RRH
$(.)_j$	Traversal index for all UEs (UAV-UE pairs in Chap. 6), especially when k is used to indicate specific UE
$(.)_k$	Index of UE
$(.)_m$	Index of $\mathcal{N}_{\text{R}}$ subsets
$(.)_{\text{C}}$	For the CU
$(.)_{\text{R}}$	For the RRHs (UAVs)
$(.)_{\text{U}}$	For the UEs

Math Mode Accents

$\bar{\mathbf{X}}$ (channel matrices)	LoS components
$\bar{\mathbf{X}}$ (precoding matrices)	Multiplication of matrix $\mathbf{X}$ and its conjugate transpose $\mathbf{X}^H$
$\hat{\mathbf{x}}$	Quantized version of signal $\mathbf{x}$
$\widetilde{\mathbf{X}}$	Given value of variable $\mathbf{X}$

Greek Symbols

$\boldsymbol{\omega}$	Gaussian quantization noise for encoded signal $\bar{\mathbf{x}}_{\text{R},i}$
$\boldsymbol{\Omega}$	Covariance matrix of quantization noise
$\alpha_{k,i}$	Large-scale fading coefficient of the access channel
β_i^{LoS}	Large-scale fading coefficient in the fronthaul link
δ	Antenna spacing
$\Gamma_{\mathcal{S}_m}$	Stacking matrix composed of $\mathcal{S}_m$ identity matrices
λ	Carrier Wavelength
$\boldsymbol{\Theta}_i$	Substitute function used inside of function φ_m^{ub}
$\boldsymbol{\Xi}_k$	Substitute function used inside of function $C_{\text{ac},k}^{\text{lb}}$
ϕ	Inclined angle from the array of CU to the x-z plane
ϕ_i	Azimuth angle from the CU to the i-th RRH
$\sigma_{\text{U}}^2,\ \sigma_{\text{R}}^2,\ \sigma_{\text{C}}^2$	Noise power level of signals received by UEs, RRHs, the CU
$\varphi_m,\ \varphi_i$	Quantization functions

ξ	Pass loss at the reference distance $d_0 = 1m$

Latin Alphabet

$\mathbf{B}$	Precoding matrix for the fronthaul link
$\mathcal{C}$	Set of rates for all UEs
$C_{\text{ac},k}$	Rate function of message M_k
C_i	Rate of message V_i
$C_{\text{fr},i}$	Rate function of message V_i
d_{min}	Minimum safe distance between any two UAVs
d_{ik}	Distance between UAV i and UE k
$f_{ij}, \mathbf{f}_i, \bar{\mathbf{f}}_i$	Substitute functions used inside of function $C_{\text{ac},k}^{\text{lb}}$
$g_{ik}, \boldsymbol{g}_k, \mathbf{G}$	Access channel model
$H_{\text{C}}, H_{\text{R}}$	Height of the CU, the UAVs
$h_{mi}, \mathbf{h}_i, \mathbf{H}$	Fronthaul channel model
K	Number of receive antennas equipped at each UE
l_{mi}	Distance between the m-th antenna at the CU and the i-th UAV
M	Total number of antennas equipped at the CU
M_k	Message intended for the k-th UE
$M_x,\ M_y$	Number of antenna columns, rows equipped at the CU
$\mathcal{N}_{\text{U}}, \mathcal{N}_{\text{R}}, \mathcal{N}_{\mathcal{S}}$	Set of UEs, RRHs, all $\mathcal{N}_{\text{U}}$ subsets
N	Number of transmit antennas equipped at each RRH
$N_{\text{U}}, N_{\text{R}}$	Number of UEs, RRHs
$n_{\text{C},i}, \mathbf{n}_{\text{C}}, \mathbf{n}_{\text{U},k},\ \mathbf{n}_{\text{R},i}$	Received transmit noise
n_{B}	Number of coherence blocks
n_{L}	Block length
n_{R}	Number of data streams for UEs
$P^{\max}$	Maximum transmit power budget
$\mathbf{p}_{\text{C}}$	CU transmit power for all signals intended for RRHs (UAVs)
$p_{\text{R},i},\ \mathbf{p}_{\text{R}}$	CU transmit power for signals intended for RRHs (UAVs)
R_k	Rate of message M_k
$\mathcal{S}_m$	The m-th subset of $\mathcal{N}_{\mathcal{S}}$

$s_{\mathrm{U},k},\ \mathbf{s}_{\mathrm{U},k}$	Encoded signal from message M_k
$s_{\mathrm{R},i}$	Baseband signal encoded from V_i (Signals transmitted from the i-th RRH in Chapter 5)
$S_{ik},\ \mathbf{S}$	Slack Variables
$\bar{\mathcal{U}}$	Set of precoding matrices for the access link in downlink transmission
$\mathbf{U}$	Precoding matrix for the access link
$\bar{\mathcal{V}}$	Set of precoding matrices for the access link in uplink transmission
$\mathbf{v}_{\mathrm{U},k},\ \mathbf{v}_{\mathrm{C},m},\ \mathbf{q}_i$	Location of the k-th UAV, the m-th antenna at the CU, the i-th UAV
V_i	Index set of the compressed $\hat{\mathbf{x}}_{R,i}$ (Index set of the compressed $\hat{\mathbf{y}}_{\mathrm{R},i}$ in Chapter 5)
$\mathbf{w}_i$	Normalized beamforming vector for transmit signal $s_{\mathrm{R},k}/\mathbf{s}_{\mathrm{R},k}$
$x_{\mathrm{R},i},\ \mathbf{x}_{\mathrm{R},i},\ \mathbf{x}_{\mathrm{U},k},\ \mathbf{x}_{\mathrm{C}}$	Transmit signals
(x, y, z)	3-D Coordinate System
$\mathbf{y}_{\mathrm{U},k},\ \mathbf{y}_{\mathrm{R},i}$	Received signals

Acronyms / Abbreviations

2G	**2**nd-**G**eneration
3-D	**3**-**D**imentional
3GPP	**3G P**artnership **P**roject
3G	**3**rd-**G**eneration
4G	**4**th-**G**eneration
5G	**5**th-**G**eneration
AF	**A**mplify-and-**F**orward
BBU	**B**ase**b**and **U**nit
BCD	**B**lock **C**oordinate **D**escent
BLER	**Bl**ock **E**rror **R**ate
bpcu	**b**its **p**er **c**hannel **u**se
BSC	**B**ase **S**tation **C**ontroller
BS	**B**ase **S**tation
BTS	**B**ase **T**ransceiver **S**tation
CDMA	**C**ode-**D**ivision **M**ultiple **A**ccess
CF	**C**ompress-and-**F**orward
CoMP	**Co**ordinated **M**ulti-**P**oint transmission and reception
CPU	**C**entral **P**rocessing **U**nit
CRAN	**C**loud/**C**entrali **R**adio **A**ccess **N**etwork
CSI	**C**hannel **S**tate **I**nformation
CU	**C**entral **U**nit
DCF	**Dec**ompress-and-**F**orward
DC	**D**ifference of **C**onvex
DF	**D**ecode-and-**F**orward
E-UTRAN	**E**volved **UTRAN**
EE	**E**nergy **E**fficiency
eMBB	**e**nhanced **M**obile **B**road**b**and
eNodeB	**e**volved **NodeB**
FDD	**F**requency **D**ivision **D**uplex

FDMA	**F**requency **D**ivision **M**ultiple **A**ccess
GERAN	**G**SM **E**volution **RAN**
GPS	**G**lobal **P**ositioning **S**ystem
GSM	**G**lobal **S**ystem for **M**obile Communications
H-CRAN	**H**eterogeneous-**CRAN**
HetNet	**Het**erogeneous **Net**work
HPC	**H**igh **P**erformance **C**omputing
IoT	**I**nternet **o**f **T**hings
IRS	**I**ntelligent **R**eflecting **S**urface
JCAS	**J**oint **C**ommunication **a**nd **S**ensing
JDD	**J**oint **D**ecompression and **D**ecoding
KKT	**K**arush-**K**uhn-**T**ucker
LoS	**L**ine-**o**f-**S**ight
LP	**L**inear **P**rogramming
LTE	**L**ong-**T**erm **E**volution
MEC	**M**ulti-access **E**dge **C**omputing
MF	**M**atched **F**ilter
MIMO	**M**ultiple-**I**nput and **M**ultiple-**O**utput
mIoT	**m**assive **IOT**
MMSE	**M**inimum **M**ean **S**quare **E**rror
mMTC	**m**assive **M**achine **T**ype **C**ommunications
MM	**M**ajorization-**M**inimization
MR	**M**aximum-**R**atio
MU-MIMO	**M**ulti-**U**ser **MIMO**
NLoS	**N**one-**L**ine-**o**f-**S**ight
NR	**N**ew **R**adio
OFDMA	**O**rthogonal **FDMA**
QCQP	**Q**uadratically **C**onstrained **Q**uadratic **P**rogramming
QoS	**Q**uality **o**f **S**ervice
QP	**Q**uadratic **P**rogramming

RAN	**R**adio **A**ccess **N**etwork
RF	**R**adio **F**requency
RNC	**R**adio **N**etwork **C**ontroller
RoF	**R**adio **o**ver **F**iber
RRH	**R**emote **R**adio **H**ead
RZF	**R**egularized **Z**ero-**f**orcing
SAA	**S**ample **A**verage **A**pproximation
SC-FDMA	**S**ingle-**C**arrier **FDMA**
SCA	**S**uccessive **C**onvex **A**pproximation
SDP	**S**emi**d**efinite **P**rogramming
SDR	**S**emi**d**efinite **R**elaxation
SE	**S**pectral **E**fficiency
SINR	**S**ignal-to-**I**nterference-plus-**N**oise **R**atio
SSUM	**S**ucce**s**sive **U**pper bound **M**inimization
TDD	**T**ime **D**ivision **D**uplex
TDMA	**T**ime-**D**ivision **M**ultiple **A**ccess
UAV	**U**nmanned/**U**ncrewed **A**erial **V**ehicle
UE	**U**ser **E**quipment
ULA	**U**niform **L**inear **A**rray
UMi	**U**rban **Mi**crocell
UMTS	**U**niversal **M**obile **T**elecommunications **S**ystem
UPA	**U**niform **P**lanar **A**rray
URLLC	**U**ltra-**R**eliable and **L**ow **L**atency **C**ommunications
UTRAN	**U**MTS **T**errestrial **RAN**
W-CDMA	**W**ideband **C**ode-**D**ivision **M**ultiple **A**ccess
ZF	**Z**ero-**F**orcing

Chapter 1

Introduction

By the end of 2022, over 95% of the global population has been covered by *3rd-generation* (3G) and advanced mobile broadband networks, making mobile internet connectivity accessible to over 55% of the worldwide population, according to the latest reports [1, 2]. The current *5th-generation* (5G) network is designed to support up to 10^6 devices per km^2 and deliver peak data rates of up to 20 gigabits/s for downlink, along with ultra-low latency ranging from 1 and 10 milliseconds [3]. The target by 2030 is to provide mobile network access to 100% of the population, with over 54% of connections being supported by 5G [1, 4]. Beyond the enhancements in data throughput and bandwidth offered by 5G networks over their *4th-generation* (4G) predecessors, 5G and beyond networks also represent a paradigm shift in networking technology, unlocking novel applications and opportunities in industry 4.0 [5], smart cities, healthcare, retail and so on. A wide range of usage scenarios have been explored and proposed for 5G and future networks [6, 7], which can be broadly classified into three categories [8, 9]:

- **Enhanced mobile broadband (eMBB):** eMBB offers wireless high-speed broadband services in densely populated areas. Its key features comprise enhanced seamless connectivity and robust support for user mobility.
- **Massive machine type communications (mMTC):** mMTC, also known as *massive Internet of Things* (mIoT), aims to stably support ultra-dense and massive numbers

of devices with low rates and energy.

- **Ultra-reliable and low latency communications (URLLC):** URLLC prioritizes mission-critical applications that are highly sensitive to latency and system reliability. It is vital to optimize every step of the uplink and downlink transmission process, as well as strategies aiming at shortening data processing response time.

The impact of future communication networks is far-reaching and encompasses every aspect of our daily lives. It is not just about delivering faster or superior services, but rather utilizing technologies as a catalyst to enable a range of services that will be integral to our lives. This thesis aims to explore some potential scenarios for future applications in an efficient and cost-effective manner.

1.1 Motivation

With the accelerated development of manufacturing in the recent several decades, hardware components could be massively produced with higher accuracy and lower costs. It enables the *central processing unit* (CPU) to achieve higher processing speed and become smaller and lighter. Portable electronic devices, as a result, are capable of being equipped with complicated modules to meet users' daily requirements. Especially since the beginning of the 21st century, there has been an explosive growth in demand for internet access from mobile phones. Wireless *radio access network* (RAN) technologies, which enable wireless communication between user devices and the core network, are thus widely studied both commercially and academically.

The *2nd-generation* (2G) cellular network is the first one that implemented entirely digital transmission since the early 1990s [10]. *Global system for mobile communications* (GSM) is a standard employed in 2G networks that utilizes *time-division multiple access* (TDMA) technologies. The enhanced data rates for *GSM evolution RAN* (GERAN) based on GSM represents one of the 2G RANs, which comprises *base station controllers* (BSCs) and *base transceiver stations* (BTSs). It was superseded by 3G *universal mobile telecommunications service* (UMTS), which utilizes *wideband code-division multiple*

access (W-CDMA) as its primary technology [11]. This was accomplished through the implementation of *radio network controllers* (RNCs) and NodeBs in *UMTS terrestrial RANs* (UTRANs). Subsequently, the 4G *long-term evolution* (LTE) technology was introduced, which involves the use of *evolved NodeBs* (eNodeBs). The widely-used method in LTE *evolved UTRANs* (EUTRANs) is *frequency division multiple access* (FDMA), with *orthogonal FDMA* (OFDMA) employed in the downlink and *single-carrier FDMA* (SC-FDMA) in the uplink. Each evolution has resulted in a significant increase in transmission rate and a decrease in latency. Since *3rd Generation Partnership Project* (3GPP) release 15 [12], the first phase of 5G specifications started in 2017 and the era of 5G *new radio* (NR) has come in. Until the submission date of this thesis, the standardization of 5G Advanced is in progress and the corresponding release 18 is expected to be finished in 2024 [13].

In future wireless networks, delivering a high data rate will remain the main driver along with other features such as high reliability and low latency while supporting a large number of users in a scalable fashion. In the physical layer, some criteria are used to evaluate the quality of wireless communication networks. In this thesis, we mainly take the following two aspects into account:

- **Spectral Efficiency (SE):** Over a limited bandwidth at the given frequency spectrum, SE is a measurement of how efficiently the information can be transmitted. The unit is written as bits/s/Hz, which measures the average number of bits that can be reliably transmitted over the channel per second per hertz. Channel capacity determines the upper bound of SE that can be achieved for a specific range of bandwidth, given the transmit power and noise variance.
- **Energy Efficiency (EE):** There are plenty of definitions of EE across various scenarios [14] to evaluate the communication network, but its fundamental objective remains consistent, which is the minimization of the total energy consumption while maintaining a certain *quality of service* (QoS) requirement, such as high throughput and low latency.

To meet the exponential growth of *internet of things* (IoT) demand, RAN becomes a

vital part of wireless telecommunication systems which realize cooperation among all units on a large scale. With the rapid evolution of telecommunication technologies, potential next-generation RAN is under research to further improve SE and EE. *Cloud/Centralized RAN* (CRAN) is known to be one of the most competitive structures for future networks [15].

Inspired by centralized signal processing technology, the CRAN architecture was first proposed in 2010 [16] with the objective of providing higher SE and EE. "C" represents centralized processing, Cooperative radio, Cloud and Clean (green). The primary idea of CRAN is to pool the *baseband units* (BBUs) of several *base stations* (BSs) altogether in a *central unit* (CU) rather than placing them separately in each BS [17]. In general, CRAN physically decouples the function of traditional BSs into two main parts: the major signal processing is centralized and merged into a BBU pool, and the distributed *remote radio heads* (RRHs) replace the conventional BSs and operate as relays. One of the major advantages of CRAN systems is to alleviate the burden of classical BSs by reducing their computational and hardware complexities. The CU undertakes most of the computational burden from the conventional distributed BSs to provide centralized high-performance computations, large-area cooperation and interference management. Meanwhile, RRHs with low complexity and power consumption can be deployed cost-effectively to realize interference mitigation and cooperation. To exploit this, the RRHs should be distributed remotely and cover a large service area. The network densification can thus be achieved to further improve SE. Figure 1.1 shows a general structure of CRANs. There are two links involved in the CRAN architecture. The transmission link for RRHs to serve *user equipments* (UE) in their local cells is referred to as the access link. Each RRH is connected to the BBU and the transmission link between them is referred to as the fronthaul link. The combined structure of one centralized BBU and multiple RRHs can manage a highly cooperative communication network and provide a wide range of seamless coverage.

In this thesis, we propose a joint design of wireless fronthaul and access links in CRANs and optimize both network SE and EE. Before conducting the research, there are a few key questions that need to be answered:

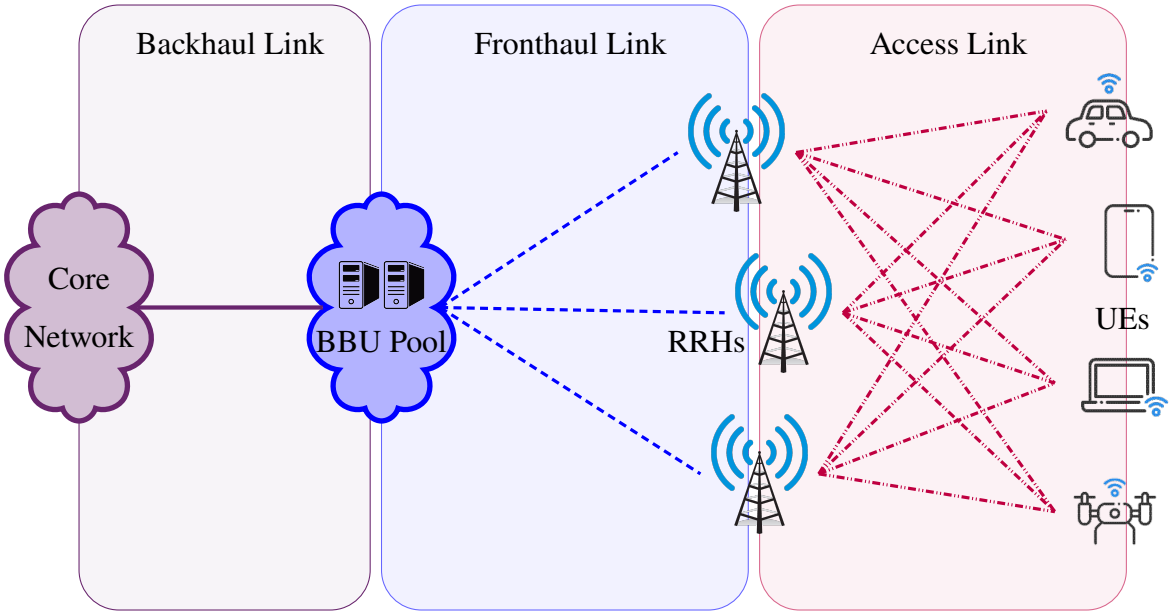


Figure 1.1: General CRAN architecture.

- **Why consider CRAN?**

  The rise of IoT has stimulated the exploration of new RANs, such as *heterogeneous network* (HetNet) [18] as a cellular network structure which integrates diverse cell sizes, such as macro/microcells, picocells, and femtocells. Achieving high network densification and cell cooperation to mitigate intra-cell interference demands the flexible deployment and centralized processing support. Additionally, cell-free systems leveraging massive *multiple-input and multiple-output* (MIMO) technologies [19] offer another promising direction for future RANs, where a large number of access points provides wide coverage while serving relatively few UEs. CRAN is capable of providing high-performance computing and cooperation for both architectures. Moreover, emerging technologies such as *intelligent reflecting surface* (IRS) [20, 21], *joint communication and sensing* (JCAS) [22], and *multi-access edge computing* (MEC) [23] have garnered significant attention and can be seamlessly integrated and supported by the CRAN architecture.

- **Why consider wireless fronthaul links?**

  Most of the existing works have mostly regarded the fronthaul link as a finite-

capacity link [24, 25] or a wired link (such as *radio of fiber* (RoF) technique [26]), constraining network designs by a fixed fronthaul capacity. However, wired networks can be prohibitively expensive, particularly for ultra-dense deployment in future RANs. Wireless communication, on the other hand, can provide mobile and temporary connectivity and support multiple access with a single transmitter. Moreover, in certain deployment scenarios such as difficult terrains and rural areas, using the wireless medium for fronthaul links could be more advantageous due to its lower deployment cost and greater flexibility in placement and adjustment. Furthermore, wireless collaboration or integration with other telecommunication networks, such as aerial or satellite communication, can be enabled by employing antennas at the CU and leveraging the CU's capabilities.

- **Why consider fully centralized structure?**

  Since some parts of the responsibilities can be centralized and managed by the CU, various protocols to split the function between the central and distributed units can be employed based on specific applications. As more functions become centralized in the cloud, the fronthaul link bears a greater burden [27]. Our goal is to improve the whole network performance at the physical layer by increasing the maximum achievable throughput and reducing the cost of transmit power, where the fully centralized structure presents the greatest challenge in achieving this goal. Moreover, partial centralization is inadequate in supporting some advanced functionalities such as spatial cooperation among distributed MIMO and *coordinated multiple point transmission and reception* (CoMP) for the access link [27]. To this end, we adopt a fully centralized structure and aim to jointly optimize both fronthaul and access links to improve SE and EE in this thesis.

1.2 Research Challenges

However, challenges always coexist when guiding new visions into reality. Therefore, this thesis aims to address the following limitations and difficulties of the CRAN while

exploring and extending practical scenarios.

- **Limited Wireless Fronthaul Capacity:**

  In the CRAN architecture, the fronthaul link plays a crucial role as it constrains the maximum throughput that can be communicated between the CU and RRHs. Although a wired fronthaul link can provide high throughput, the lack of flexibility and high cost of deployment and maintenance is non-negligible [17]. Wireless fronthaul link further increases the burden on the whole network. Improving the SE and EE of fronthaul links is key to enhancing the overall performance of the network.

- **Fast and Low-cost Deployment and Mobility:**

  To achieve densely deployed networks, it is vital to minimize the size, cost, and weight of the traditional BSs. Full centralization can alleviate the burden on the RRHs, making them the lightest and most cost-effective units. These units can be swiftly mobilized and conveniently relocated (especially wirelessly connected), while also demanding minimal maintenance in the future. However, this approach leads to a heavy burden on the fronthaul link, thereby emphasizing the need for effective compression schemes.

- **Imperfect *Channel State Information* (CSI):**

  Most of the work on CRAN assumes that the CU perfectly knows the instantaneous CSI of the wireless access link, on which the beamforming matrices are designed based. However, obtaining perfect CSI is costly even in *time division duplex* (TDD) mode. It is more practical to have only stochastic CSI available but the absence of accurate CSI can result in a reduction in the achievable capacity.

- **Support for Aerial Communication:**

  Given the CU's ability to operate high-performance computations, integrating aerial communications into the CRAN is a promising architecture. This integration enables the clustering and centralization of most of the signal processing and network control from conventional flying platforms to the CU, resulting in a lighter weight and longer hovering time for drones. Seamless extensive cooperation with both aerial

and terrestrial entities can be achieved and enhanced in a cost-effective manner. However, it presents challenges to wireless access of flying platforms, where both SE and EE are significantly limited, especially for the fronthaul link. Furthermore, the positions of aerial units and compression schemes need to be designed properly.

1.3 Contributions

To address the challenges mentioned above and further improve SE and EE, we propose novel CRAN architectures and evaluate their performances in this thesis. Our contributions can be summarized as follows:

- **Joint design of wireless fronthaul and access link in Massive MIMO CRANs:** Unlike most of the existing works in CRAN, where the wired fronthaul is assumed as an ideal bit-pipe with a fixed capacity constraint, we propose a novel architecture of CRAN with massive MIMO enabled in the fronthaul link and focus on optimizing varying capacity constraints in the wireless fronthaul link. A joint cooperative design of fronthaul and access links is proposed while considering the impact of multivariate compression noise and interference that naturally exist in wireless channels. In particular, the precoding matrices for the access link, the quantization covariance matrices, and transmit powers are jointly optimized to maximize the network sum-rate and minimize the total power consumption. Additionally, we present designs with lower computational complexity, including the independent point-to-point compression strategy and separate optimization designs [28, 29] for wireless fronthaul and access links, which are compared to the proposed joint design in terms of achieved performance, feasibility, and computational time.

  - In the downlink transmission, the *decompress-and-forward* (DCF) scheme is utilized at the RRHs. Also, different levels of CSI knowledge, namely instantaneous CSI and stochastic CSI, are both considered in the optimization problems aiming to improve network SE and EE. While the framework for

incorporating instantaneous CSI into the system design is relatively simple, the lack of CSI knowledge presents a significant challenge. To address this issue, we present an ergodic sum-rate maximization problem that assumes the availability of only the stochastic CSI.

– In the uplink transmission, perfect knowledge of CSI is assumed to be available at the CU. Two conventional precoding methods, namely *maximum-ratio* (MR) and *zero-forcing* (ZF) combining, are adopted for the access link. The *compress-and-forward* (CF) scheme is utilized at the RRHs aiming to improve the sum-rate of the proposed designs.

All the aforementioned problems are non-convex and thus hard to solve. We utilize iterative algorithms based on *successive convex approximation* (SCA) to tackle the corresponding problems by transforming them into relaxed convex *semidefinite relaxation* (SDR) problems [30]. We also provide proof that the optimized solution from the proposed algorithms is feasible and optimal for the original problems. Furthermore, we provide comprehensive numerical results and compare our proposed designs to benchmark schemes.

- **Multi-UAV-enabled CRANs with Massive MIMO:**

  We propose a *unmanned/uncrewed aerial vehicle* (UAV)-enabled CRAN with a massive MIMO BBU as a supplement system to the terrestrial communication networks. UAVs are regarded as flying RRHs that are integrated into CRAN[1] which aims to reduce the circuit and signal processing complexity as well as power consumption in UAVs. Unlike the existing works, we adopt a more realistic system model with DCF relaying at the flying RRHs and jointly optimize both wireless fronthaul and access links including the placements of the UAVs. The idea of UAV-UE paired is described and the corresponding optimization problem is presented. We propose an iterative algorithm by decomposing and reformulating the original problems into two convex SDR subproblems, the solution of which is proven to be

[1] As a result, the proposed architecture will inherit all the benefits of the conventional CRAN architecture.

optimal for the proposed model. Numerical simulations are conducted to present the trajectory of UAVs' optimized placements. In addition, the network performance is evaluated by comparing it to the benchmark schemes.

1.4 Thesis Structure

The remainder of this thesis is organized as follows:

- **Chapter 2:** In this chapter, we introduce the essential background of relevant technologies and concepts used in this thesis, followed by a comprehensive literature review of state-of-the-art research.
- **Chapter 3:** The downlink of a CRAN architecture with wireless fronthaul and access links is investigated in this chapter, where a CU equipped with a very large antenna array serves multiple multi-antenna RRHs that in turn provide service to a number of UEs. The assumption of perfect knowledge of instantaneous CSI available at the CU is made for designing the precoding matrices of the access link. Constrained by the transmit power budgets, a joint design of both fronthaul and access links is proposed. The corresponding optimization problems aiming to maximize the sum-rate and minimize the total power consumption are formulated. Separate designs of fronthaul and access links are also investigated with low complexity. Iterative algorithms are presented for solving the resulting problems. The proposed algorithms are evaluated and compared with benchmarks.
- **Chapter 4:** Adopting the same system model proposed in Chapter 3, only the knowledge of stochastic CSI of the access link is available at the CU for system designing in this chapter. We propose both joint and separate designs targeting to maximize the ergodic sum-rate over multiple coherence blocks and minimize the transmit power consumption. The resulting optimization problems are hard to solve directly. Iterative algorithms were presented and compared to benchmark schemes

to demonstrate their superiority in achieving higher SE and EE.

- **Chapter 5:** The uplink transmission in CRAN with massive MIMO utilized in the fronthaul link is considered in this chapter. Maximizing the system SE and EE involves jointly optimizing the precoders at the UEs, the quantization noise covariance matrices, and transmit powers at the multi-antenna RRHs and UEs. To solve the resulting non-convex problems, iterative algorithms based on the *majorization-minimization* (MM) approach method are proposed. Two precoding schemes at the central unit are considered, namely MR and ZF combining.

- **Chapter 6:** A novel architecture of multi-UAV-enabled CRAN is proposed in this chapter. In particular, we propose to deploy the UAVs as flying RRHs to serve ground UEs. The CU is equipped with a large-scale antenna array to serve the flying RRHs and affords all the baseband signal processing. To optimize the proposed architecture, the problem of maximizing the minimum rate of UEs is considered by jointly optimizing UAVs placement, quantization noise variance, and power control. The corresponding optimization problem is not convex and to solve it an efficient iterative algorithm is derived by combining the block coordinate descent and *successive convex approximation* (SCA) methods. Numerical results demonstrate the superior performance of the proposed algorithm compared to the two benchmark schemes.

- **Chapter 7:** In this chapter, conclusions are drawn about the primary findings and results from our work. Potential future work is also discussed.

1.5 List of Publications

The publications related to the main contributions of this thesis are listed as follows:

- **Chapter 3:**
  - Y. Huang and A. Ikhlef, "Joint optimization of wireless fronthaul and access links in CRAN with a massive MIMO central unit," *ICC 2022 - IEEE International Conference on Communications*, Seoul, Korea, Republic of, 2022, pp. 1906-1911, doi: 10.1109/ICC45855.2022.9839160.
- **Chapter 4:**
  - Y. Huang and A. Ikhlef, "Joint optimization of wireless fronthaul and access links in massive MIMO CRAN with stochastic CSI," submitted to *2023 IEEE Global Communications Conference*.
- **Chapter 5:**
  - D. Maryopi, Y. Huang and A. Ikhlef, "Sum-rate maximization in uplink CRAN with a massive MIMO fronthaul," *2021 IEEE Globecom Workshops (GC Wkshps)*, Madrid, Spain, 2021, pp. 1-6, doi: 10.1109/GCWkshps52748.2021.9681970.
- **Chapter 6:**
  - Y. Huang and A. Ikhlef, "Joint Design of Fronthaul and Access Links in Massive MIMO Multi-UAV-Enabled CRANs," in *IEEE Wireless Communications Letters*, vol. 10, no. 11, pp. 2355-2359, Nov. 2021, doi: 10.1109/LWC.2021.3100320.

Chapter 2

Background and Literature Review

This chapter provides a concise overview of the background and state-of-the-art research on several relevant technologies employed in this thesis. We first provide an overview of the concept of CRAN architecture in Section 2.1. In Section 2.2, we introduce the fundamental concept of MIMO and relevant technologies. Section 2.3 introduces UAV-aided communication with advantages and challenges. Section 2.4 gives a brief introduction to convex optimization problems which are widely used for solving our proposed models. Finally, the latest research on the aforementioned technologies, as well as their combined design is presented in Section 2.5.

2.1 CRAN Architecture

Ultra-dense deployment and flexible centralized processing are the key enablers for future networks [31]. CRAN is a promising technology to achieve both functions and meet the requirements of future wireless communication networks [15]. The primary idea of CRAN is to migrate the baseband processing from the BSs to the CU [17]. Therefore, the low-complexity and low-cost BSs, also known as the RRHs, can be managed to cooperate with each other and effectively suppress interference, resulting in higher SE and EE. The CU

communicates with the RRHs through the fronthaul links while the RRHs communicate with the UEs via the access links.

2.1.1 Function Splits

As in conventional RANs, the protocol layers operated by BSs can be generally categorized as the L3 network layer, L2 data link layer and L1 physical layer. Since some features can be centralized into the CU, different split options in these three layers can be applied between the central and distributed units [32]. According to the standardization of 3GPP [33], there are 8 separate options depending on the applications, as shown in Figure 2.1. Several works have discussed and compared different function splits [27, 31, 34, 35].

Full centralization is the most promising CRAN architecture where the functional split happens between the *radio frequency* (RF) and *physical* (PHY) layer. In another word, all three layers are executed in the CU whereas only RF functionality is performed at the RRHs. This structure is clear and simple with significant benefits in terms of flexible deployment, low cost and simple maintenance [36]. However, large amounts of data need to be transmitted uplink/downlink through the fronthaul link which thus requires high capacity and low latency in a fully centralized structure.

Centralized processing allows distributed RRHs to operate as relays and forward messages to UEs. There are various strategies for processing at a relay. *Decode-and-forward* (DF) is one of the most common techniques in which relay nodes decode the received signal before forwarding the decoded signals to the destination. However, in the downlink transmission, DF requires the local CSI at each RRH to encode signals for UEs within the coverage. Additionally, it suffers from intra-cell interference among RRHs. An alternative approach to overcome these issues is to generate the precoding design of the access link at the CU to achieve interference management and transmit its compressed information to the RRHs prior to the data transmission. In the uplink, RRHs require the knowledge of codebooks used by UEs in order to decode the signals correctly, as well

Figure 2.1: Flexible functional split between the CU and RRHs.

as the knowledge of fronthaul CSI to encode the signals transmitted to CU. However, acquiring this information leads to a heavy burden on the fronthual link and is not feasible in practice. Also, both transmission requires RRHs to undertake part of baseband signal processing. Therefore, signal quantization and compression schemes are preferable for a full centralization structure of CRANs with wireless fronthaul link [27].

2.1.2 Quantization and Compression

In the recent decade, there have been many studies on compression and decompression schemes for wireless fronthaul links, aiming to increase network capacity. From the perspective of information theory, the compression process can be viewed as a Gaussian test channel [37], in which the input consists of uncompressed signals and the output comprises compressed signals.

For the DCF scheme in the CRAN downlink transmission, denote $\mathbf{x}^n$ as an input sequence of n independent and identically distributed (i.i.d.) samples in a point-to-point compression system. For a given test channel with a rate R, $\mathbf{x}^n$ is compressed into an index M within a codebook $\mathcal{C} = \{\hat{\mathbf{x}}^n(1), \cdots, \hat{\mathbf{x}}^n(2^{nR})\}$ for transmission. At the receiver side, M is decompressed to obtain the corresponding sequence $\hat{\mathbf{x}}^n$. In order to recover the signals successfully,

$$R \geq I(\mathbf{x}^n; \hat{\mathbf{x}}^n) \tag{2.1}$$

needs to be satisfied. For multivariate compression systems [38–40], assuming that there are N_{R} RRHs and $\mathcal{N}_{\mathrm{R}} \triangleq \{1, \ldots, N_{\mathrm{R}}\}$ is the RRHs set. Denote M chosen codebooks $\mathcal{C}_1, \cdots, \mathcal{C}_{N_{\mathrm{R}}}$ known at both transmitters and receivers, where each codebook $\mathcal{C}_i = \{\hat{\mathbf{x}}_i^n(1), \cdots, \hat{\mathbf{x}}_i^n(2^{nR_i})\}$ has rate R_i. The sequence $\mathbf{x}^n$ is compressed randomly and conditionally independently into N_{R} indices, which indicates the corresponding codewords $\hat{\mathbf{x}}_i^n, \ldots, \hat{\mathbf{x}}_{N_{\mathrm{R}}}^n$. The signals can be reliably decompressed only if

$$\sum_{i \in \mathcal{S}} R_{\mathcal{S}} \geq \sum_{i \in \mathcal{S}} h(\hat{\mathbf{x}}_i^n) - h(\hat{\mathbf{x}}_{\mathcal{S}}^n | \mathbf{x}^n), \quad \forall \mathcal{S} \subseteq \{1, \ldots, N_{\mathrm{R}}\}. \tag{2.2}$$

This thesis utilizes and compares both independent and multivariate compression schemes for the proposed models.

For the CF scheme in the CRAN uplink transmission, denote $\mathbf{y}^n$ as the signals received at the RRH intended to be compressed as $\hat{\mathbf{y}}_i^n$ and forwarded to the CU. For the conventional single-user compression scheme,

$$R \geq I(\mathbf{y}^n; \hat{\mathbf{y}}^n) \tag{2.3}$$

has to be satisfied in order to recover $\mathbf{y}^n$ successfully. For Wyner–Ziv distributed coding [40, 41], define a compression order $\{\pi(1), \ldots, \pi(N_{\mathrm{R}})\}$. Therefore, the CU can leverage the decompressed signals $\{\hat{\mathbf{y}}_{\pi(1)}^n, \ldots, \hat{\mathbf{y}}_{\pi(i-1)}^n\}$ as the side information before decompressing the signal $\hat{\mathbf{y}}_{\pi(i)}^n$. All signals can be decompressed successfully only if the conditions

$$R_{\pi(i)} \geq I(\mathbf{y}_{\pi(i)}^n; \hat{\mathbf{y}}_{\pi(i)}^n | \hat{\mathbf{y}}_{\pi(1)}^n, \ldots, \hat{\mathbf{y}}_{\pi(i-1)}^n), \ \forall i \in \mathcal{N}_{\mathrm{R}}, \tag{2.4}$$

are satisfied.

2.2 MIMO

MIMO is a wireless communication technology that enables spatial multiplexing by employing multiple antennas at both the transmitter and receiver to increase the data transmission capacity [42, 43]. Let M and K denote the number of transmit and receive antennas, respectively, the channel model $\mathbf{H} \in \mathbb{C}^{K \times M}$ is given as

$$\mathbf{H} \triangleq \begin{bmatrix} \mathbf{h}_1 \\ \mathbf{h}_2 \\ \vdots \\ \mathbf{h}_K \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & \cdots & h_{1M} \\ h_{21} & h_{22} & \cdots & h_{2M} \\ \vdots & & & \vdots \\ h_{K1} & h_{K2} & \cdots & h_{KM} \end{bmatrix}, \tag{2.5}$$

where the (i, j)-entry of $\mathbf{H}$ represents the channel response from the j-th transmit antenna to the i-th receive antenna. For a Rician distributed channel, h_{ij} is given by

$$h_{ij} = \sqrt{\beta_i} \left(\sqrt{\frac{\kappa}{\kappa + 1}} \bar{h}_{ij} + \sqrt{\frac{1}{\kappa + 1}} \tilde{h}_{ij} \right), \tag{2.6}$$

where β_i represents the large-scale fading. $\kappa \geq 0$ corresponds to the Rician factor defined as the power ratio between the *line-of-sight* (LoS) and *non-line-of-sight* (NLoS) components. When $\kappa = 0$, the channels exhibit Rayleigh fading with complex small scale fading $\tilde{h}_{ij}$ following a normal distribution

$$\tilde{h}_{ij} \sim \mathcal{CN}(0, 1). \tag{2.7}$$

When $\kappa \to \infty$, in contrast, the channels behave as deterministic LoS propagation with phase shift, i.e.,

$$\bar{h}_{ij} = e^{-\mathrm{i}\frac{2\pi}{\lambda}D_{ij}}, \tag{2.8}$$

where λ is the carrier wavelength and D_{ij} is the distance between the i-th antenna at the transmitter and the j-th antenna at the receiver. i denotes the imaginary unit.

Considering a horizontal ULA used for transmission, Figure 2.2 shows the path of a transmitted waveform from each antenna in the ULA. In reality, when the distance between the receiver and transmitter is large enough, the rays transmitted from antennas can be approximated as parallel. Therefore, the LoS channel from the ULA to the i-th receive antenna at the receiver side is given as

$$\mathbf{h}_i = [h_{11}h_{12}\cdots h_{1M}] = \sqrt{\beta_i}\left[1\ e^{-\mathrm{i}\frac{2\pi}{\lambda}d\cos(\phi_i)}\ \ldots\ e^{-\mathrm{i}\frac{2\pi}{\lambda}(M-1)d\cos(\phi_i)}\right], \tag{2.9}$$

where $d \triangleq \frac{\lambda}{2}$ denotes the antenna spacing and $\phi_i \in [0, 2\pi)$ is the azimuth angle of departure. Provided an accurate azimuth angle ϕ_k between the transmitter and a specific

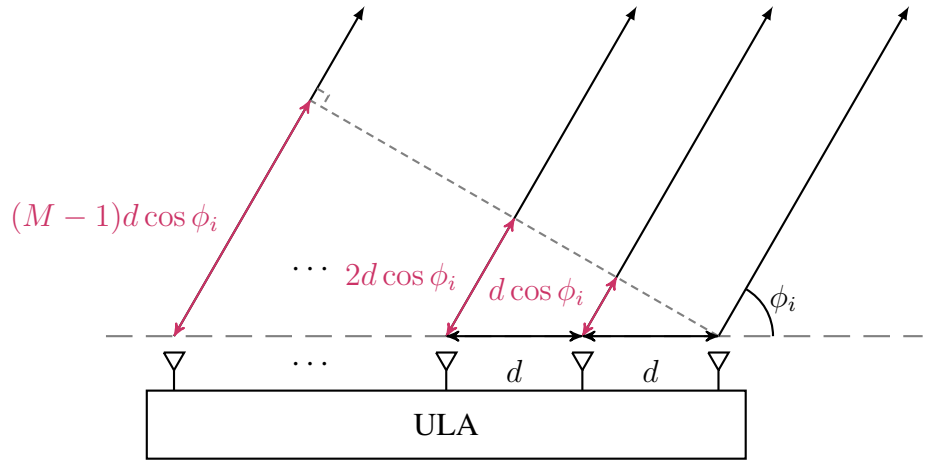


Figure 2.2: LoS propagation line at the transmitter side with a ULA.

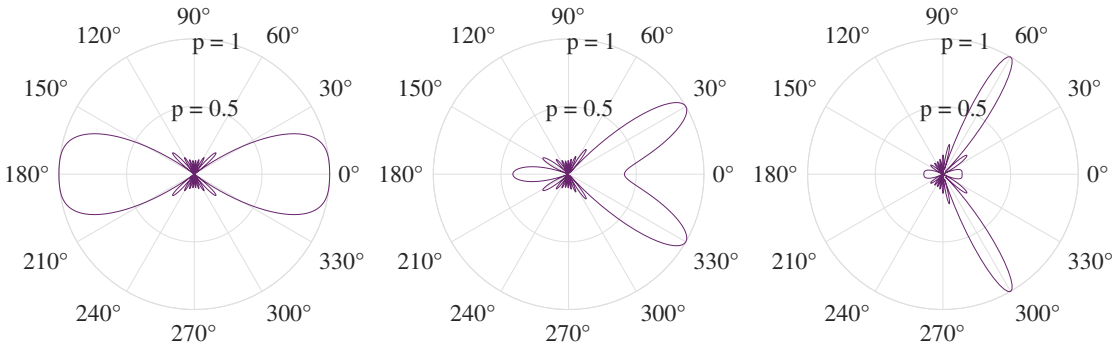
k-th receiver, applying the normalized digital beamforming, represented by

$$\frac{\mathbf{h}_k^H}{\|\mathbf{h}_k^H\|} = \frac{1}{M}[\bar{h}_{11}^* \bar{h}_{12}^* \cdots \bar{h}_{1M}^*] = \frac{1}{M}\left[1\ e^{\mathrm{i}\frac{2\pi}{\lambda}d\cos(\phi_k)}\ \ldots\ e^{\mathrm{i}\frac{2\pi}{\lambda}(M-1)d\cos(\phi_k)}\right], \tag{2.10}$$

as the steering vector, can effectively direct and concentrate the beams towards the intended target. Figure 2.3 depicts the received gain for various azimuth angles.

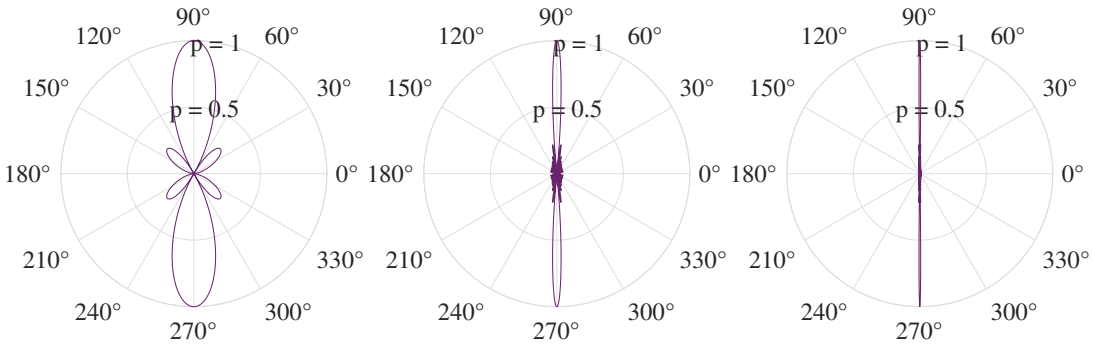
2.2.1 Massive MIMO

Massive MIMO is an extension of MIMO that describes systems with hundreds or even thousands of antennas for transmission and/or reception. In both LoS propagation and Rayleigh fading channels, narrower beams can be achieved by using more antennas for transmission, which allows for more precise targeting of signals towards their intended destinations. Assuming that all transmit antennas are omnidirectional, Figure 2.4 illustrates the beam shape generated by different numbers of transmit antennas M, all with the same total transmit power. As the number of transmit antennas increases significantly, a phenomenon known as channel hardening occurs [42]. This means that the channels directed towards the target tend to be deterministic, with favorable propagation characteristics that reduce the impact of small-scale fading and frequency dependence as the number of antennas increases. The more antennas applied at the transmitter side, the more



(a) $\phi_i = 0°$ (b) $\phi_i = 30°$ (c) $\phi_i = 60°$

Figure 2.3: Power gain of beams transmitted from ULA versus different azimuth angles ϕ_i for 2GHz wavelength, $M = 10$, $d = \frac{\lambda}{2}$, and p represents normalized power.



(a) $M = 4$ (b) $M = 20$ (c) $M = 100$

Figure 2.4: Power gain of beams transmitted from ULA with different numbers of array elements M for 2GHz wavelength, $\phi_k = 90°$, $d = \frac{\lambda}{2}$, and p represents normalized power.

accurately the beam power is directed at the receivers. Specifically, these channels are asymptotically orthogonal, meaning that they behave as

$$\frac{\|\mathbf{h}_i\|^2}{\mathbb{E}\{\|\mathbf{h}_i\|^2\}} \xrightarrow{M\to\infty} 1, \qquad \forall i = 1, 2, \ldots, K, \tag{2.11a}$$

$$\frac{\mathbf{h}_i^H \mathbf{h}_{i'}}{\sqrt{\mathbb{E}\{\|\mathbf{h}_i\|^2\}\mathbb{E}\{\|\mathbf{h}_{i'}\|^2\}}} \xrightarrow{M\to\infty} 0, \qquad \forall i, i' = 1, 2, \ldots, K, \ \ i \neq i'. \tag{2.11b}$$

The conclusion in [44] showed that asymptotic orthogonality would exist in the LoS scenario if a ULA or *uniform planar array* (UPA) is applied, but not *uniform circular array* (UCA). It appears that the asymptotic orthogonality of channels leads to a higher *signal-to-interference-plus-noise ratio* (SINR), which lets the effective throughput approach capacity bounds as the number of antennas M increases.

2.2.2 CSI Acquisition

Point-to-point MIMO, *multi-user MIMO* (MU-MIMO), and massive MIMO in TDD and *frequency division duplex* (FDD) modes require varying levels of CSI [42]. Assuming that the number of transmit antennas is greater than or equal to the number of receive antennas, i.e., $M \geq K$, in the downlink transmission, the channel training burden is strongly affected by the number of antennas M at the BS. Due to channel reciprocity in TDD mode, the CSI

estimated from the uplink channel training can be utilized in downlink data transmission in the same coherence time [45]. The BS can learn the CSI from the uplink pilot transmission with a minimum length of K and then easily estimate the downlink channels and precode signals for downlink transmission. Especially for massive MIMO, downlink training is unnecessary when $M \to \infty$. In FDD mode, on the other hand, downlink training needs M pilots, whereas uplink training requires K pilots and extra resources for sending quantized CSI back to the UE. It is non-practical for channel training in FDD mode with a large number of BS antennas due to high complexity. It is clear that TDD operation is more suitable for massive MIMO with $M \gg K$.

2.3 UAV-enabled Communication

In recent years, wireless communication systems which employ UAVs have attracted extensive attention and have been widely studied for both military and civilian applications [46, 47]. According to different functionalities, sizes, control methods, and application scenarios, UAVs can be classified into two main categories based on wing configuration [47]. The first category includes fixed-wing UAVs that can have higher flying velocity and better endurance, while the second category is the rotary-wing UAVs that can hover in the sky and takeoff/land vertically like a tiny counterpart of the helicopter. UAVs are regarded as a good candidate for coverage extension due to their low cost, fast deployment, dynamic adjustment, and adaptability in severe environments and non-reachable areas [48–50]. Particularly, UAV-aided communications are indispensable in some situations where conventional terrestrial communication systems are overloaded or even nonexistent, e.g., temporary network coverage of major sports events or emergency network coverage in disasters [51].

Due to the varying altitudes of UAVs typically ranging from 1.5 to 300 meters, the wireless channels associated with UAV communication systems are often more dominated by the LoS channel component [47, 52] than those associated with terrestrial communica-

tions. As a result, Rician fading [53, 54] and free space fading [55, 56] channel models are commonly utilized in UAV wireless communications. This is beneficial for UAV-enabled communication to achieve stable connection and QoS requirements. The diagram presented in Figure 2.5 depicts a scenario wherein multiple UAVs are employed in a CRAN architecture as flying RRHs.

2.4 Convex Optimization

Generally, optimization problems are NP-hard according to computational complexity theory, which cannot be solved within polynomial time. However, convex optimization is one type of mathematical problems which has effective solutions, e.g., interior-point methods [57]. As convex functions and convex optimization problems are widely used in this thesis, a brief introduction is given in this section.

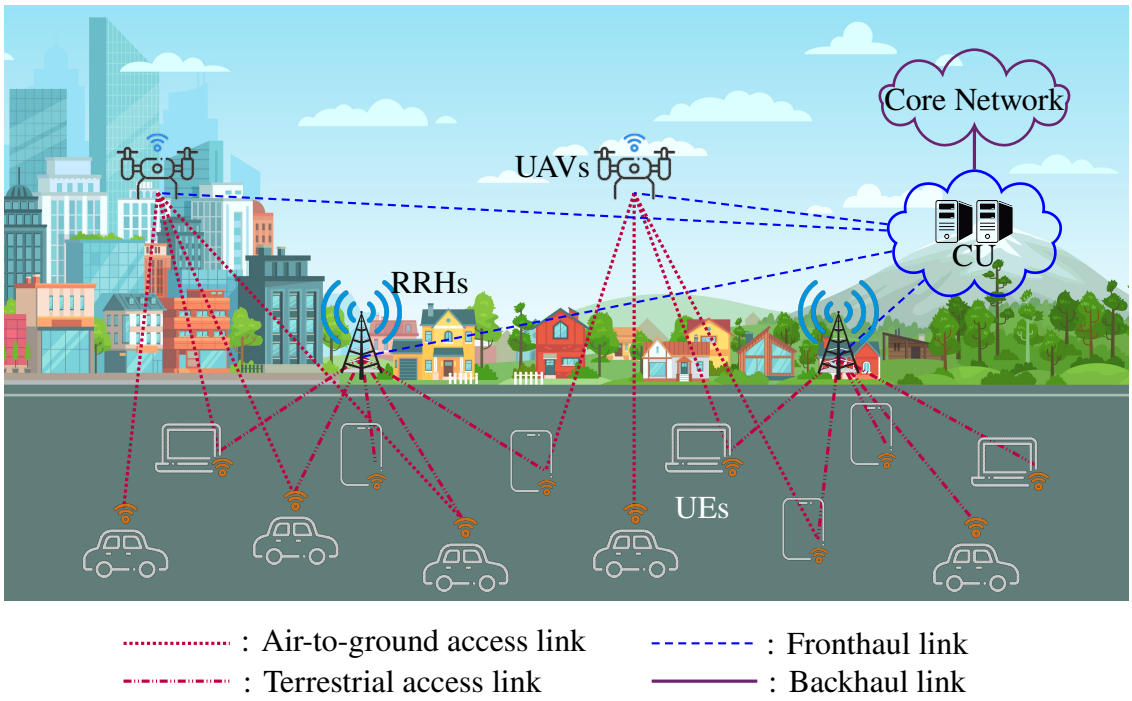


Figure 2.5: UAV-enabled CRAN architectures.

Assuming that the domain of function $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ is a convex set, which is equivalent to [57]

$$\theta x_1 + (1-\theta)x_2 \in \mathbf{dom} f,\ \forall x_1, x_2 \in \mathbf{dom} f \text{ and } \forall \theta \in [0,1]\,, \tag{2.12}$$

function $f(x)$ can be defined as convex if and only if

$$f(\theta x_1 + (1-\theta)x_2) \leq \theta f(x_1) + (1-\theta) f(x_2),\ \forall x_1, x_2 \in \mathbf{dom} f \text{ and } \forall \theta \in [0,1]\,, \tag{2.13}$$

is satisfied. Especially, function f is affine if the equality in (2.13) holds, whereas function f is strictly convex if only the inequality holds.

An optimization problem can be generally formed as

$$\min_{x \in \mathbb{R}^n} \quad f_0(x), \tag{2.14a}$$

$$\text{s.t.} \quad f_i(x) \leq 0,\ i = 1, \ldots, m, \tag{2.14b}$$

$$h_i(x) = 0,\ i = 1, \ldots, p, \tag{2.14c}$$

where x is optimization variable, $f_0(x)$ in (2.14a) is the objective function aimed to be minimized. (2.14b) and (2.14c) are the constraints need to be satisfied. The problem (2.14) is convex if and only if inequality constraints $f_0, \ldots, f_m$ are convex and equality constraints $h_1, \ldots, h_p$ are affine.

2.4.1 Common Convex Problems

It is clear that if functions $f_0, \ldots, f_m$ are all affine, as a special case of convex problems, the problem is called LP, which could be written in a standard form as

$$\min_{\mathbf{x} \in \mathbb{R}^n} \quad \mathbf{b}_0^T \mathbf{x}, \tag{2.15a}$$

$$\text{s.t.} \quad \mathbf{b}_i^T \mathbf{x} + c_i \leq 0,\ i = 1, \ldots, m, \tag{2.15b}$$

$$h(\mathbf{x}) = 0, \tag{2.15c}$$

where $\mathbf{b}_0, \ldots, \mathbf{b}_m$ are vectors with constant elements and $c_1, \ldots, c_m$ are scalars. With convex quadratic functions included in both objective function and constraints, it leads to QCQP, which has the form as

$$\min_{\mathbf{x}\in\mathbb{R}^n} \quad \mathbf{x}^T\mathbf{A}_0\mathbf{x} + \mathbf{b}_0^T\mathbf{x}, \tag{2.16a}$$

$$\text{s.t.} \quad \mathbf{x}^T\mathbf{A}_i\mathbf{x} + \mathbf{b}_i^T\mathbf{x} + c_i \leq 0,\ i = 1, \ldots, m, \tag{2.16b}$$

$$h(\mathbf{x}) = 0. \tag{2.16c}$$

where $\mathbf{A}_i \in \mathbb{R}^{n\times n}$ satisfies $\mathbf{A}_i = \mathbf{A}_i^H \succeq \mathbf{0}, \forall i$. LP is a special case of the QCQP problem with $\mathbf{A}_i = \mathbf{0}, \forall i$.

In fact, all the aforementioned convex optimization problems can be collectively seen as a more general problem called SDR. The standard model of SDR is given as [30]

$$\min_{\mathbf{x}\in\mathbb{R}^n} \quad \text{tr}(\mathbf{C}^T\mathbf{X}), \tag{2.17}$$

$$\text{s.t.} \quad \text{tr}(\mathbf{A}_i^T\mathbf{X}) = \mathbf{b}_i, \tag{2.18}$$

$$\mathbf{X} \succeq \mathbf{0}. \tag{2.19}$$

Convex SDR can be solved efficiently by interior-point methods and is broadly used in this thesis. However, almost all the proposed problems in this thesis are non-convex problems and thus difficult to solve. By transforming these non-convex problems into convex SDR problems while preserving the same constraints, we can obtain solutions and attain convergence. The conversion and relaxation of these problems, as well as the associated algorithms, are of practical significance and will be elaborated upon in subsequent chapters

2.4.2 Convexity Preservation

Certain operations and compositions of functions are known to preserve convexity. Based on the frequent usage in this thesis, we enumerate some examples here.

- **Operations:** Denote non-negative scalar x and positive semi-definite matrix $\mathbf{X} \succeq \mathbf{0}$

as variables, we have [57]

Function	Curvature	Monotonicity
$\sum_i a_i x_i,\ \forall a_i \geq 0$	affine	non-decreasing
$1/x$	concave	non-increasing
$\lvert x\rvert,\ \lVert\mathbf{X}\rVert_2$	convex	non-monotonic
$\log(x),\ \log\det(\mathbf{X})$	concave	non-decreasing

- **Function composition:** Consider a composition function $f = h \circ g$ defined as [57]

$$f(x) = h(g(x)), \quad \textbf{dom} f = \{x \in \textbf{dom} g | g(x) \in \textbf{dom} h\}, \tag{2.20}$$

with the existence of second derivative with respect to variable x, which is given as

$$f''(x) = h''(g(x))g'(x)^2 + h'(g(x))g''(x). \tag{2.21}$$

f is convex only if one of the following is satisfied

  - h is convex and non-decreasing and g is convex;
  - h is convex and non-increasing and g is concave.

f is concave only if one of the following is satisfied

  - h is concave and non-decreasing and g is concave;
  - h is concave and non-increasing and g is convex.

2.5 State-of-the-art Research

In this section, the latest research on architectures and technologies related to our proposed models are introduced.

2.5.1 CRAN

Since the proposal of the CRAN architecture in 2011, there have been numerous surveys conducted on this topic. [32, 35] presented a list of promising research directions

of CRAN research at a very early stage. [27] focused on the research on the constrained fronthaul as well as signal compression and quantization for both uplink and downlink transmission. [58, 59] presented the evolution of RANs and a vision of the current 5G network with its expected QoS. Also, comprehensive research on CRAN and advanced *heterogeneous-CRAN* (H-CRAN) was provided from both academic and commercial perspectives, including mobile operators and vendors. [36] summarized recent studies on CRAN architecture, SE and EE enhancement, interference management, delay, and security, along with their respective solution approaches. Besides throughput enhancement, resource allocation in CRAN is an important topic to improve EE, which has been extensively covered in several review papers. [60] surveyed the EE improvement using large-scale cooperative processing and cloud-based CoMP in H-CRAN. [14] provided a comprehensive survey of EE optimization problems in CRAN in terms of different adjectives, constraints, corresponding algorithms, and applications. Furthermore, various structures of CRAN are currently under review and widely discussed. [34] presented an extensive overview of the different function splits on the protocol stack between the CU and DU. [15, 61] have compared the full and partial centralized CRAN.

The CRAN architecture involves three types of links, namely backhaul, fronthaul, and access links, and their joint design is of practical interest to improve network SE and EE. The backhaul link connects CRANs to the core network, and reviews on backhaul solutions in the 5G perspective were provided by [62]. Some works have simply assumed ideal fronthaul links without constraints while focusing on MU-MIMO in the access link. Different from traditional cellular networks, spatial multiplexing can be achieved by employing a large number of distributed RRHs serving multiple UEs simultaneously. CoMP was identified as a key feature in the previous LTE-A structure [63], but it was mainly achieved by neighboring cells, where the number of multi-points was normally less than five. However, with centralized processing in CRAN, large-scale cloud-based CoMP can be achieved, allowing UEs to receive signals from a large number of distributed low-cost RRHs. [64, 65] considered each RRH and UE equipped with only a single antenna in the downlink transmission. [64] used MIMO in the access link by enabling RRH selection

to serve UEs. Power allocation was optimized aiming to maximize the average weighted sum-rate of the CRAN network. However, this work utilized *regularized zero-forcing* (RZF) as the precoding method and only the regularization factor was optimized for the problems. [65] applied massive MIMO in the access link where the number of RRHs is ten times that of UEs. EE was aimed to be maximized and constrained by the target network sum-rate. However, the CSI was assumed perfectly known at the transmitters. Further works proposed to use multiple antennas in both RRHs and UEs. [66] considered stochastic CSI in the sum-rate maximization problem to mitigate the interference cancellation.

However, large-scale cooperative processing among distributed RRHs requires high-performance fronthaul links with large capacity and low latency. Since large amounts of data need to be collected and processed by the CU and also be transmitted from the CU via the fronthaul link, it can become a bottleneck even if an optical-fiber cable is deployed [40]. [26] provided an in-depth analysis of the modulation and resource allocation schemes in CRANs with RoF fronthaul links. The signal compression and quantization scheme is the key to improving the system performance such as SE and EE [17, 27]. As studied in [67], there are two prominent schemes in the CRAN system according to RRH operation type, namely DF and DCF-based schemes. Both schemes are proposed to maximize the weighted sum-rate and compared with benchmarks. However, it is worth noting that in this work, only the independent point-to-point DCF scheme was considered. [68] introduced a spatial DCF scheme. In the downlink transmission, the multivariate compression studied in [38, 39] is one of the promising approaches in CRAN systems. The work in [39] proved that the effect of quantization noises can be suppressed by designing its correlation properly. In [38], a joint design of precoding scheme and multivariate fronthaul compression was proposed for CRAN architecture. In this work, the system design was considered for both perfect CSI and imperfect effective CSI scenarios. However, this work assumed the fronthaul link to have finite capacity and considered only hybrid analog-digital beamforming. It is worth noting that fully digital precoding can achieve better performance. For uplink transmission, CF schemes such as conventional single-user compression and Wyner-Ziv coding can be employed

at the RRHs to quantize the received signals from the UEs, and then forward them to the CU [41, 69–76]. [69, 75, 76] focused on the *joint decompression and decoding* (JDD) in the uplink transmission with finite fronthaul capacity. [72, 73] further considered the successive decoding method utilized at the CU, along with the Wyner-Ziv coding to compress the signals received at the RRHs. However, the above-mentioned works only considered the fronthaul link with a given limited capacity and did not take its optimization into account. [74] aimed to jointly optimize both the fronthaul and access links in the uplink transmission and compare the performance of DF and CF. However, in this work, only conventional single-user compression was considered, and a very limited number of receive antennas were utilized at the CU. Furthermore, [77] jointly maximized the sum-rate of uplink and downlink transmission in full-duplex CRANs by optimizing beamforming and compression covariance matrices. However, this work only considered the independent point-to-point compression scheme for the downlink and the conventional single-user compression for the uplink.

Meanwhile, some works have focused on the joint design under the assumption that both links are wireless, particularly in the case of the full centralization structure, which places a greater load on the fronthaul links. [78] studied several beamforming designs for both wireless fronthaul and access links and provided theoretical performance bounds for the proposed system. However, the beamforming matrices of the two links were optimized separately and only *amplify-and-forward* (AF) relaying mechanism was considered at the RRHs. [28] managed to mitigate the inter-cluster interference by jointly designing the wireless fronthaul and access links, which were compared to the separate designs. The SE was maximized by optimizing the beamforming schemes and compression strategies. However, this paper only considered the independent point-to-point compression scheme. Also, a very limited number of antennas were used at each unit for transmission or reception, e.g., 4 antennas at the CU and a single antenna at both the RRHs and UEs. [79] studied the joint beamforming and multivariate compression scheme in the downlink transmission, while [74] considered joint precoding design and conventional single-user compression scheme in the uplink. However, both of only considered a limited number

of antennas at the CU and did not exploit the potential benefits of massive MIMO for the fronthaul link.

2.5.2 UAV-enabled Communication

Wireless communication is a crucial enabling technology for UAVs, which are anticipated to present promising business prospects in the forthcoming decade [47, 51, 80]. Based on the Rician fading model widely used in air-to-ground communications, the trade-off between minimum data rate and UAV altitude for maximum coverage was studied in [53]. [54] presented a UAV-enabled cooperation scheme for cancelling the interference in ground cellular networks with the Rician fading channel model. Both [53] and [54] did not take the fronthaul links into account. [80] considered CoMP in the CRAN by applying UAVs as RRHs and dynamically optimized the UAV placement for max-min rate fairness. For access links, the authors derived lower and upper bound expressions for the achievable rates in both LoS and Rayleigh fading channels. However, the fronthaul links in [80] were regarded as ideal channels without any limitations and interference. [55] and [56] both studied the UAV trajectory while [55] focused on optimizing the EE of the system and [56] concentrated on optimizing the system throughput. Both of them utilized the free space channel model which is not realistic.

The use of multi-UAV-enabled wireless communication systems with centralized signal processing was considered in [80–83]. [80] studied CoMP in CRAN by using UAVs as RRHs and dynamically optimized the UAV placement for max-min rate fairness. [81] considered multi-UAV trajectory control to maximize the minimum rate. In both [80] and [81] the fronthaul links were assumed to have unlimited/very high capacity. [82] maximized the sum-rate in the uplink of CRAN with LoS fronthaul and access links by optimizing the UE association, UAV placement, and UEs' and UAVs' transmit powers. However, both the fronthaul and access links were assumed to be interference-free by using orthogonal frequency bands for different users. [56] studied the joint design of UAV trajectory and power control. The authors in [80] proposed a CoMP architecture with the

optimization of UAV placements. However, the fronthaul link in these papers was only considered as an ideal capacity-limited transmission link. In our work, we propose a more practical architecture by studying the transmission strategy and the design of power control in both fronthaul and access links. In order to mitigate the interference in the fronthaul links, we consider massive MIMO because the channel coefficients become asymptotically orthogonal as the number of CU antennas goes to infinity [44, 84].

The constrained flying time is a severe defect due to the limited battery capacity in UAVs [55]. Apart from the power used for hovering, energy usage for other purposes such as signal processing and transmission needs to be reduced in order to achieve longer service time. The CRAN framework is a promising solution for this challenge. CRAN can be regarded as a special relay-enabled communication architecture where the selection and placement of RRHs and UEs have a significant impact. [85] introduced a two-step optimization problem designing the precoding scheme and RRH selection jointly. [86] studied the imperfect CSI scenario with the optimization of EE. However, the RRHs in both papers had fixed locations and thus the authors aimed to select appropriate UEs to be served for higher data throughput. Integrating massive MIMO technology at the CU to support drones shows significant potential to further improve the network throughput. In [87], a comprehensive *3-dimentional* (3-D) LoS channel model was proposed under massive MIMO scenarios with UAVs. The results showed that the low complexity encoding, e.g. MF and ZF, is capable of achieving excellent data rates.

In our work, unlike the existing works, we adopt a more practical system model with a massive MIMO CU and a DCF relaying protocol at the flying RRHs and jointly optimize both the fronthaul and access links including the placements of the UAVs. To the best of the authors' knowledge, this work is the first to use massive MIMO and compression at the CU and DCF at the flying RRHs to improve the capacity of the wireless fronthaul links in multi-UAV-enabled CRANs, where usually the fronthaul link is the bottleneck.

Chapter 3

Design of Wireless Downlink Transmission in Massive MIMO CRANs with Perfect Instantaneous CSI

3.1 Introduction

In this chapter, we leverage massive MIMO to enhance the fronthaul link's capacity in a CRAN architecture. This improvement in fronthaul capacity has a positive impact on the entire system's overall capacity. We aim to maximize the sum-rate of all UEs and minimize power consumption by jointly designing the wireless fronthaul and access links. To achieve this, we formulate the optimizing problems that take both links into account. The proposed optimization problems adopt the DCF scheme at the RRHs and account for the quantization noise. We also optimize the precoding matrices of the access link and transmit power of both links. However, solving these problems can be challenging. To address this issue, we first linearize relevant terms within non-convex constraints and then reformulate the problem as an SDR problem. The proposed algorithm is proven to be convergent and provides optimal solutions for the original problem. We also consider

the problem where the fronthaul and access links are separately designed [28, 29] and the corresponding algorithm is provided.

The system model of downlink transmission using wireless fronthaul and access links is presented in Section 3.2. In Section 3.3, we formulate the optimization problems based on the assumption that perfect knowledge of instantaneous CSI is available at the CU. Additionally, we provide a detailed description of the corresponding iterative algorithms. In Section 3.4, the superiority of our proposed approaches is demonstrated by comparing their abilities to increase the sum-rate and reduce total power consumption with two benchmark schemes. This chapter is concluded in Section 3.5.

3.2 System Model

We consider the downlink transmission in a CRAN consists of a CU equipped with a massive ULA of M omnidirectional antennas for transmission. The CU provides wireless service to N_{U} multi-antenna UEs through N_{R} RRHs that are serving as relays. In particular, $M \gg N_{\mathrm{R}}$. Each RRH is equipped with a single antenna for reception and N antennas for transmission while each UE has K receive antennas. The distance between antennas in each node is equal to half of the carrier wavelength. For notational convenience, we let $\mathcal{N}_{\mathrm{R}} \triangleq [1, \ldots, N_{\mathrm{R}}]$ and $\mathcal{N}_{\mathrm{U}} \triangleq [1, \ldots, N_{\mathrm{U}}]$ denote the sets of RRHs and UEs, respectively.

The downlink transmission involves two wireless links, namely the fronthaul and access links, which are assumed to operate in different time/frequency domains to avoid interference between them. Specifically, let $\mathbf{h}_i \in \mathbb{C}^{1\times M}$ denote the channel vector between the CU and the i-th RRH, $i \in \mathcal{N}_{\mathrm{R}}$, and thus the fronthaul link can be defined as $\mathbf{H} = [\mathbf{h}_1^T, \mathbf{h}_2^T, \ldots \mathbf{h}_{N_{\mathrm{R}}}^T]^T$. By appropriately placing the CU and RRHs, a clear LoS path can be achieved and thus the channels are modelled as

$$\mathbf{h}_i = \sqrt{\beta_i^{\mathrm{LoS}}} \left[1 \; e^{\mathrm{i}\frac{2\pi}{\lambda} d\cos(\phi_i)} \; \ldots \; e^{\mathrm{i}\frac{2\pi}{\lambda}(M-1)d\cos(\phi_i)}\right], \tag{3.1}$$

where β_i^{LoS} represents the large-scale fading coefficient, d is the antenna spacing, and

$\phi_i \in [0, 2\pi)$ is the azimuth angle of arrival to RRH i from the CU. β_i^{LoS} is assumed to be constant over frames and known at the CU as both the CU and RRHs are fixed. i denotes the imaginary unit. We assume $\phi_i \neq \phi_j$ for $i \neq j$ such that $\lim_{M\to\infty} |\bar{\mathbf{h}}_i^H \bar{\mathbf{h}}_j| \to 0$ if $i \neq j$ [43]. Similarly, the access link is expressed as $\mathbf{G}_k = [\mathbf{G}_{k,1}, \ldots, \mathbf{G}_{k,N_{\mathrm{R}}}]$, where $\mathbf{G}_{k,i} \in \mathbb{C}^{K\times N}$ represents the channel matrix between the k-th UE and the i-th RRH, $k \in \mathcal{N}_{\mathrm{U}},\ i \in \mathcal{N}_{\mathrm{R}}$. We adopt a block fading channel model, where the fading coefficients are constant within each time block and statistically independent from one block to another. The access link is modelled as Rayleigh fading channels, which can be expressed as

$$\mathbf{G}_{k,i} = \sqrt{\alpha_{k,i}}\, \tilde{\mathbf{G}}_{k,i}, \tag{3.2}$$

where $\alpha_{k,i}$ denotes the large-scale fading coefficient and $\tilde{\mathbf{G}}_{k,i} \in \mathbb{C}^{K\times N}$ is a matrix with elements modelled as $\mathcal{CN}(\mathbf{0}, \mathbf{I})$. The schematic diagram of the proposed system is shown in Figure 3.1.

In terms of downlink data transmission, we adopt the DCF relaying strategy [39] where the RRHs first decompress the received signal from the CU and then forward the decompressed signals to the UEs. Let $M_k \in \{1, \ldots, 2^{n_{\mathrm{L}} R_k}\},\ k \in \mathcal{N}_{\mathrm{U}}$, denote the message to be transmitted from the CU to the k-th UE, where n_{L} is the block length and R_k is the

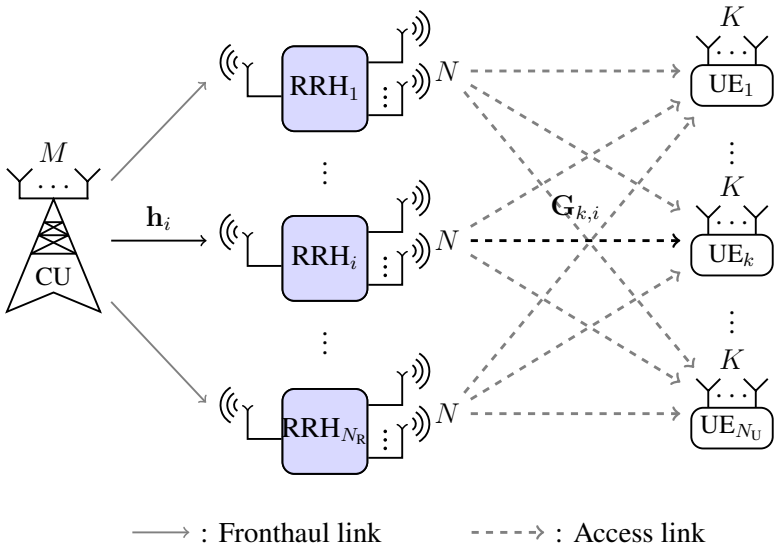


Figure 3.1: System model of downlink transmission in the proposed massive MIMO CRAN architecture.

rate of message M_k. First, the CU encodes the message M_k as $\mathbf{s}_{\mathrm{U},k} \in \mathbb{C}^{n_{\mathrm{R}}\times 1} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I})$, where $n_{\mathrm{R}} \leq \min(K, NN_{\mathrm{R}})$ denotes the number of data streams, and then precodes the encoded signal as

$$\mathbf{x}_{\mathrm{R},i} = \sum_{k\in\mathcal{N}_{\mathrm{U}}} \mathbf{U}_{k,i}\mathbf{s}_{\mathrm{U},k}, \tag{3.3}$$

where $\mathbf{U}_{k,i} \in \mathbb{C}^{N\times n_{\mathrm{R}}}$ denotes the precoding matrix for signal $\mathbf{s}_{\mathrm{U},k}$. To account for the limited capacity of the wireless fronthaul link, the CU quantizes and compresses the precoded signals $\mathbf{x}_{\mathrm{R},i} \in \mathbb{C}^{N\times 1}$. We adopt the Gaussian quantization test channel to model the quantization process [39], and hence, the resulting quantized signal, $\hat{\mathbf{x}}_{\mathrm{R},i}$, can be written as

$$\hat{\mathbf{x}}_{\mathrm{R},i} = \mathbf{x}_{\mathrm{R},i} + \boldsymbol{\omega}_i, \tag{3.4}$$

where $\boldsymbol{\omega}_i \in \mathbb{C}^{N\times 1} \sim \mathcal{CN}(\mathbf{0}, \boldsymbol{\Omega}_{ii})$ represents the Gaussian quantization noise, which is independent of $\mathbf{x}_{\mathrm{R},i}$. $\boldsymbol{\Omega}_{ii}$ denotes the covariance of the compression noise. Intuitively, the smaller the quantization noise level the more accurate the signal, and vice versa. The CU then compresses the quantized signal $\hat{\mathbf{x}}_{\mathrm{R},i}$ to generate the compression index $V_i \in \{1, 2, \ldots, 2^{n_{\mathrm{L}}C_i}\}$, where C_i is the rate of message V_i that can be interpreted as the capacity of the fronthaul link between the CU and RRH i. Note that the signals $\hat{\mathbf{x}}_{\mathrm{R},i}$, $i \in \mathcal{N}_{\mathrm{R}}$, could be either compressed independently or jointly [39]. In particular, if the signals are independently compressed the quantization noises for different RRHs will be uncorrelated, i.e., $\mathbb{E}[\boldsymbol{\omega}_i\boldsymbol{\omega}_j^H] = \mathbf{0}$, for $i \neq j$. However, if the signals are jointly compressed the quantization noises for different RRHs will be correlated, i.e., $\mathbb{E}[\boldsymbol{\omega}_i\boldsymbol{\omega}_j^H] \neq \mathbf{0}$, for $\forall i, j \in \mathcal{N}_{\mathrm{R}}$. Using joint compression allows to improve the achievable rates compared to independent compression. Since the independent compression can be seen as a special case of joint compression, we adopt joint compression and the two compression schemes will be compared in the numerical results. For notational convenience, denote $\boldsymbol{\Omega} \triangleq \mathbb{E}[\boldsymbol{\omega}\boldsymbol{\omega}^H]$, where $\boldsymbol{\omega} \triangleq [\boldsymbol{\omega}_1^T, \boldsymbol{\omega}_2^T, \ldots, \boldsymbol{\omega}_{N_{\mathrm{R}}}^T]^T$. For independent compression, $\boldsymbol{\Omega}$ is a block diagonal matrix, i.e., $\boldsymbol{\Omega} = \mathrm{diag}\{\boldsymbol{\Omega}_{11}, \ldots, \boldsymbol{\Omega}_{N_{\mathrm{R}}N_{\mathrm{R}}}\}$. The compressed signal V_i is then encoded as a baseband signal $s_{\mathrm{R},i} \sim \mathcal{CN}(0, 1)$. The signal $s_{\mathrm{R},i}$ is next beamformed towards RRH i,

using the beamforming vector $\mathbf{w}_i \in \mathbb{C}^{M\times 1}$, to produce the following transmit signal

$$\mathbf{x}_C = \sum_{i\in\mathcal{N}_\mathrm{R}} \sqrt{p_i}\mathbf{w}_i s_{\mathrm{R},i}, \tag{3.5}$$

where p_i is the CU transmit power corresponding to signal $s_{\mathrm{R},i}$ intended for RRH i. In this chapter, without loss of generality, we consider the ZF precoding method[1], which is a conventional low-complexity linear precoder [43]. Hence, $\mathbf{w}_i$, $i \in \mathcal{N}_\mathrm{R}$, is expressed as

$$\mathbf{w}_i = \frac{\mathbf{b}_i}{\|\mathbf{b}_i\|}, \tag{3.6}$$

where $\mathbf{b}_i$ is the i-th column of

$$\mathbf{B} \triangleq \mathbf{H}^H(\mathbf{H}\mathbf{H}^H)^{-1}. \tag{3.7}$$

Assuming that the CU has a maximum average transmit power $P_\mathrm{C}^{\max}$, from (3.5) and (3.6), the average transmit power at the CU is constrained as $\mathbb{E}[\|\mathbf{x}_C\|^2] = \sum_{i\in\mathcal{N}_\mathrm{R}} p_i \le P_\mathrm{C}^{\max}$.

Through the fronthaul link, the received signal at the i-th RRH from the CU is given by

$$\mathbf{y}_{\mathrm{R},i} = \mathbf{h}_i\mathbf{x}_C + \mathbf{n}_{\mathrm{R},i}, \tag{3.8}$$

where $\mathbf{n}_{\mathrm{R},i} \in \mathbb{C}^{M\times 1} \sim \mathcal{CN}(0, \sigma_\mathrm{R}^2\mathbf{I})$ is the noise received at the i-th RRH. From (3.5) and (3.8), the received signal at RRH i can be recast as

$$\mathbf{y}_{\mathrm{R},i} = \sqrt{p_i}\mathbf{h}_i\mathbf{w}_i s_{\mathrm{R},i} + \sum_{j\in\mathcal{N}_\mathrm{R}\setminus\{i\}} \sqrt{p_j}\mathbf{h}_i\mathbf{w}_j s_{\mathrm{R},j} + \mathbf{n}_{\mathrm{R},i}. \tag{3.9}$$

Next, RRH i decodes $s_{\mathrm{R},i}$ based on $\mathbf{y}_{\mathrm{R},i}$ and consequently recovers the message V_i. The achievable rate C_i of message V_i, is given as

$$C_i \le C_{\mathrm{fr},i}(\mathbf{p}) \triangleq \log \frac{\sum_{j\in\mathcal{N}_\mathrm{R}} p_j\mathbf{h}_i\mathbf{w}_j\mathbf{w}_j^H\mathbf{h}_i^H + \sigma_\mathrm{R}^2}{\sum_{j\in\mathcal{N}_\mathrm{R}\setminus\{i\}} p_j\mathbf{h}_i\mathbf{w}_j\mathbf{w}_j^H\mathbf{h}_i^H + \sigma_\mathrm{R}^2}, \tag{3.10}$$

where $\mathbf{p} \triangleq [p_1, p_2, \ldots, p_{N_\mathrm{R}}]^T$. Based on the decoded message V_i, RRH i can determine the quantized signal $\hat{\mathbf{x}}_{\mathrm{R},i}$. With joint compression at the CU, in order for RRH i to recover

[1]Other methods such as *matched filter* (MF), RZF, and *minimum mean square error* (MMSE) could easily be considered.

$\hat{\mathbf{x}}_{\mathrm{R},i}$, according to the rate-distortion theory, the condition

$$\varphi_m(\{\mathbf{U}_j\}_{j\in\mathcal{N}_\mathrm{U}},\boldsymbol{\Omega}) \triangleq \sum_{i\in\mathcal{S}_m} \log\det\left(\boldsymbol{\Gamma}_i^H\left(\sum_{j\in\mathcal{N}_\mathrm{U}}\mathbf{U}_j\mathbf{U}_j^H\right)\boldsymbol{\Gamma}_i+\boldsymbol{\Omega}_{ii}\right) - \log\det\left(\boldsymbol{\Gamma}_{\mathcal{S}_m}^H\boldsymbol{\Omega}\boldsymbol{\Gamma}_{\mathcal{S}_m}\right) \leq \sum_{i\in\mathcal{S}_m} C_i \tag{3.11}$$

where $\mathbf{U}_j \triangleq [\mathbf{U}_{j,1}^T,\mathbf{U}_{j,2}^T,\ldots,\mathbf{U}_{j,N_\mathrm{R}}^T]^T$, $j\in\mathcal{N}_\mathrm{U}$, must be satisfied for all subsets $\mathcal{S}_m\subseteq\mathcal{N}_\mathrm{R}$, $m\in\mathcal{N}_\mathcal{S}=\{1,\ldots,2^{N_\mathrm{R}}-1\}$ [39]. Matrix $\boldsymbol{\Gamma}_{\mathcal{S}_m}$ denotes a matrix that is obtained by stacking matrices $\boldsymbol{\Gamma}_i$, $i\in\mathcal{S}_m$, horizontally, e.g., if $\mathcal{S}_m=\{1,2,4\}$ then $\boldsymbol{\Gamma}_{\mathcal{S}_m}=[\boldsymbol{\Gamma}_1\boldsymbol{\Gamma}_2\boldsymbol{\Gamma}_4]$. $\boldsymbol{\Gamma}_i\in\mathbb{C}^{NN_\mathrm{R}\times N}$ is an all zero matrix except the submatrix from row $(i-1)N+1$ to iN, which is an identity matrix of size N. Note that in order to decompress the received signal, each RRH needs to be informed by the CU about the used compression codebooks. The decoded signal $\hat{\mathbf{x}}_{\mathrm{R},i}$ is then transmitted by RRH i to all UEs through the access link. Assuming that RRH i has a maximum average transmit power $P_{\mathrm{R},i}^{\max}$, from (3.4), the average transmit power at RRH i must be constrained as

$$\mathbb{E}[\|\hat{\mathbf{x}}_{\mathrm{R},i}\|^2] = \sum_{k\in\mathcal{N}_\mathrm{U}} \mathrm{tr}\left(\boldsymbol{\Gamma}_i^H\mathbf{U}_k\mathbf{U}_k^H\boldsymbol{\Gamma}_i\right)+\mathrm{tr}\left(\boldsymbol{\Omega}_{ii}\right)\leq P_{\mathrm{R},i}^{\max}, \tag{3.12}$$

where the partial entries $\boldsymbol{\Omega}_{ii}$ can be expressed as $\mathrm{diag}\left(\boldsymbol{\Omega}_{ii}\right)\triangleq\boldsymbol{\Gamma}_i^H\boldsymbol{\Omega}\boldsymbol{\Gamma}_i,\ i\in\mathcal{S}_m$.

For the access link, the signals received by the k-th UE is given by

$$\mathbf{y}_{\mathrm{U},k} = \sum_{i\in\mathcal{N}_\mathrm{R}}\mathbf{G}_{k,i}\hat{\mathbf{x}}_{\mathrm{R},i}+\mathbf{n}_{\mathrm{U},k}, \tag{3.13}$$

where $\mathbf{n}_{\mathrm{U},k}\in\mathbb{C}^{K\times 1}\sim\mathcal{CN}(0,\sigma_\mathrm{U}^2\mathbf{I})$ is the Gaussian noise at the k-th UE. By substituting (3.4) into (3.13), the received signal at UE k can be rewritten as

$$\mathbf{y}_{\mathrm{U},k} = \mathbf{G}_k\mathbf{U}_k\mathbf{s}_{\mathrm{U},k}+\sum_{j\in\mathcal{N}_\mathrm{U}\setminus\{k\}}\mathbf{G}_k\mathbf{U}_j\mathbf{s}_{\mathrm{U},j}+\sum_{i\in\mathcal{N}_\mathrm{R}}\mathbf{G}_{k,i}\boldsymbol{\omega}_i+\mathbf{n}_{\mathrm{U},k}. \tag{3.14}$$

The first, second, and third terms in (3.14) represent the desired signal to be decoded, the interference from other UEs, and the quantization noise contribution, respectively. Based on the received signal in (3.14), UE k decodes the message M_k and its achievable average rate is given by

$$C_{\mathrm{ac},k}\left(\{\mathbf{U}_j\}_{j\in\mathcal{N}_\mathrm{U}},\boldsymbol{\Omega}\right)\triangleq\log\det\left(\mathbf{I}+\frac{\mathbf{G}_k\mathbf{U}_k\mathbf{U}_k^H\mathbf{G}_k^H}{\sum\limits_{j\in\mathcal{N}_\mathrm{U}\setminus\{k\}}\mathbf{G}_k(\mathbf{U}_j\mathbf{U}_j^H+\boldsymbol{\Omega})\mathbf{G}_k^H+\sigma_\mathrm{U}^2\mathbf{I}}\right). \tag{3.15}$$

3.3 Problem Formulation

In this chapter, we assume that the CU has perfect instantaneous CSI knowledge, which is used to design the precoding and quantization noise covariance matrices. We propose two optimization problems: the sum-rate maximization problem and the total power minimization problem. However, solving these problems is challenging due to their non-convexity. To address this, we transform them into convex problems, which can easily be solved using interior point methods. We provide the algorithms to obtain the final optimized solutions and prove that the solutions obtained by the proposed algorithms are also feasible for the original problems.

3.3.1 Sum-rate Maximization

Given the ever-increasing demand for mobile data throughput, we start by focusing on the sum-rate maximization problem. Our objective is to maximize the sum-rate among all UEs, subject to the constraints imposed by the wireless fronthaul capacity, quantization noise, and power budgets of both the CU and RRHs. In this section, both joint and separate designs are proposed to maximize the sum-rate, along with corresponding algorithms to solve these problems.

3.3.1.a Joint Design of Fronthaul and Access Links

Based on the system model presented above, we propose a joint design to optimize both wireless fronthaul and access links. First, the corresponding optimization problem for network sum-rate maximization can be formulated as

$$\underset{\mathbf{p}\geq 0,\ \{\mathbf{U}_j\}_{j\in\mathcal{N}_\mathrm{U}}\succeq\mathbf{0},\ \mathbf{\Omega}\succeq\mathbf{0}}{\text{maximize}} \quad \sum_{k\in\mathcal{N}_\mathrm{U}} C_{\mathrm{ac},k}\left(\{\mathbf{U}_j\}_{j\in\mathcal{N}_\mathrm{U}}, \mathbf{\Omega}\right) \tag{3.16a}$$

$$\text{s.t.} \quad \varphi_m(\{\mathbf{U}_j\}_{j\in\mathcal{N}_\mathrm{U}}, \mathbf{\Omega}) \leq \sum_{i\in\mathcal{S}_m} C_{\mathrm{fr},i}(\mathbf{p}),\ \forall m \in \mathcal{N}_\mathcal{S}, \tag{3.16b}$$

$$\sum_{i\in\mathcal{N}_\mathrm{R}} p_i \leq P_\mathrm{C}^{\max}, \tag{3.16c}$$

$$\sum_{k\in\mathcal{N}_\mathrm{U}} \mathrm{tr}\left(\mathbf{\Gamma}_i^H \mathbf{U}_k \mathbf{U}_k^H \mathbf{\Gamma}_i\right) + \mathrm{tr}\left(\mathbf{\Omega}_{ii}\right) \leq P_{\mathrm{R},i}^{\max},\ \forall i \in \mathcal{N}_\mathrm{R}. \tag{3.16d}$$

Constraint (3.16b) ensures that the quantized signals are recoverable at the RRHs, while constraints (3.16c) and (3.16d) are the power constraints for the CU and RRHs, respectively.

By inspecting the objective function (3.16a), the fronthaul constraint (3.16b), and the power constraints the fronthaul constraint (3.16c) and (3.16d), we can identify that this problem belongs to the class of non-convex problems. Since it is difficult to solve, we reformulate the problem into a class of DC problems and solve it using the MM method [88, 89].

Adopting the approach given in [39], the first step is to make a change of variable

$$\bar{\mathbf{U}}_j \triangleq \mathbf{U}_j \mathbf{U}_j^H \succeq \mathbf{0}. \tag{3.17}$$

For simplicity, denote $\bar{\mathcal{U}} \triangleq \{\bar{\mathbf{U}}_j : \forall j \in \mathcal{N}_{\mathrm{U}}\}$ as the set of precoding matrices in the access link, such that we have a function $C_{\mathrm{ac},k}(\bar{\mathcal{U}}, \mathbf{\Omega}_i)$ in the objective function (3.16a) and $\varphi_i(\bar{\mathcal{U}}, \mathbf{\Omega}_i)$ in the first constraint function (3.16b). However, the change of variable $\mathbf{U}_j$ in (3.17) does not guarantee that the optimal solution of $\mathbf{U}_j$ can be obtained from the optimized value of $\bar{\mathbf{U}}_j$. To this end, we introduce a rank constraint into the problem (3.16) to ensure the solution is feasible for the original problem, resulting in the following reformulation

$$\underset{\mathbf{p}\succeq 0,\, \bar{\mathcal{U}}\succeq \mathbf{0},\, \mathbf{\Omega}\succeq \mathbf{0}}{\text{maximize}} \quad \sum_{k\in\mathcal{N}_{\mathrm{U}}} C_{\mathrm{ac},k}\left(\bar{\mathcal{U}}, \mathbf{\Omega}\right) \tag{3.18a}$$

$$\text{s.t.} \quad \varphi_m(\bar{\mathcal{U}}, \mathbf{\Omega}) \leq \sum_{i\in\mathcal{S}_m} C_{\mathrm{fr},i}(\mathbf{p}),\ \forall m \in \mathcal{N}_{\mathcal{S}}, \tag{3.18b}$$

$$\sum_{i\in\mathcal{N}_{\mathrm{R}}} p_i \leq P_{\mathrm{C}}^{\max}, \tag{3.18c}$$

$$\sum_{k\in\mathcal{N}_{\mathrm{U}}} \mathrm{tr}\left(\mathbf{\Gamma}_i^H \bar{\mathbf{U}}_k \mathbf{\Gamma}_i\right) + \mathrm{tr}\left(\mathbf{\Omega}_{ii}\right) \leq P_{\mathrm{R},i}^{\max},\ \forall i \in \mathcal{N}_{\mathrm{R}}, \tag{3.18d}$$

$$\mathrm{rank}\left(\bar{\mathbf{U}}_k\right) \leq n_{\mathrm{R}},\ \forall k \in \mathcal{N}_{\mathrm{U}}, \tag{3.18e}$$

where rank constraint (3.18e) guarantees the recovery of $\mathbf{U}_j$ from a given $\bar{\mathbf{U}}_j$. Solving this problem is still challenging due to the rank constraint (3.18e), and the non-convex or non-concave functions $C_{\mathrm{ac},k}\left(\bar{\mathcal{U}}, \mathbf{\Omega}\right)$, $\varphi_m(\bar{\mathcal{U}}, \mathbf{\Omega})$ and $C_{\mathrm{fr},i}(\mathbf{p})$. As a result, finding an optimal solution can be computationally complex and time-consuming. To address the challenges

presented by the non-convex constraints, we apply the MM method. Specifically, we use this method to reformulate the non-convex functions as the difference of two concave functions. This reformulation enables us to linearize the terms responsible for the non-convexity of the problem, thereby making it possible to solve the problem efficiently. Based on (3.15), $C_{\text{ac},k}(\bar{\mathcal{U}}, \mathbf{\Omega})$ can be rewritten as

$$\begin{aligned} C_{\text{ac},k}(\bar{\mathcal{U}}, \mathbf{\Omega}) = &\log\det\left(\sum_{j\in\mathcal{N}_{\text{U}}} \mathbf{G}_k(\bar{\mathbf{U}}_j + \mathbf{\Omega})\mathbf{G}_k^H + \sigma_{\text{U}}^2\mathbf{I}\right) \\ &- \log\det\left(\sum_{j\in\mathcal{N}_{\text{U}}\setminus\{k\}} \mathbf{G}_k(\bar{\mathbf{U}}_j + \mathbf{\Omega})\mathbf{G}_k^H + \sigma_{\text{U}}^2\mathbf{I}\right), \end{aligned} \tag{3.19}$$

which is the difference of two concave functions with respect to the optimization variables $\bar{\mathcal{U}}$ and $\mathbf{\Omega}$. To transform the objective function (3.18a) into a concave function, it is necessary to linearize the second term of (3.19). Since the second term is a concave function, it is upper bounded by its first order Taylor expansion at the given point $(\tilde{\bar{\mathcal{U}}}, \widetilde{\mathbf{\Omega}})$. Therefore, $C_{\text{ac},k}(\bar{\mathcal{U}}, \mathbf{\Omega})$ is lower bounded by its convex expression as

$$\begin{aligned} C_{\text{ac},k}(\bar{\mathcal{U}}, \mathbf{\Omega}) \geq\ & C_{\text{ac},k}^{\text{lb}}\left(\bar{\mathcal{U}}, \mathbf{\Omega} | \tilde{\bar{\mathcal{U}}}, \widetilde{\mathbf{\Omega}}\right) \\ \triangleq\ & \log\det\left(\mathbf{G}_k\left(\sum_{j\in\mathcal{N}_{\text{U}}} \bar{\mathbf{U}}_j + \mathbf{\Omega}\right)\mathbf{G}_k^H + \sigma_{\text{U}}^2\mathbf{I}\right) - \log\det\left(\mathbf{\Xi}_k\right) \\ & - \frac{1}{\ln(2)}\text{tr}\left(\mathbf{\Xi}_k^{-1} \times \left(\mathbf{G}_k\left(\sum_{j\in\mathcal{N}_{\text{U}}\setminus\{k\}} \bar{\mathbf{U}}_j + \mathbf{\Omega}\right)\mathbf{G}_k^H - \mathbf{\Xi}_k\right)\right), \end{aligned} \tag{3.20}$$

where $\mathbf{\Xi}_k \triangleq \mathbf{G}_k\left(\sum_{j\in\mathcal{N}_{\text{U}}\setminus\{k\}} \tilde{\bar{\mathbf{U}}}_j + \widetilde{\mathbf{\Omega}}\right)\mathbf{G}_k^H + \sigma_{\text{U}}^2\mathbf{I}$.

To convexify constraint (3.18b), it is necessary to transform $\varphi_m(\bar{\mathcal{U}}, \mathbf{\Omega})$ and $C_{\text{fr},i}(\mathbf{p})$ into convex and concave functions, respectively. Based on (3.11), the left-side function $\varphi_m(\bar{\mathcal{U}}, \mathbf{\Omega})$ is the difference of two concave terms. To make this function convex, the first Taylor expansion can be applied to linearize the first term, which is the non-convex part.

Therefore, for the given point at $(\tilde{\bar{\mathcal{U}}}, \tilde{\bar{\mathbf{\Omega}}})$, $\varphi_m(\bar{\mathcal{U}}, \mathbf{\Omega})$ can be expressed by its upper bound as

$$\begin{aligned}\varphi_m\left(\bar{\mathcal{U}}, \mathbf{\Omega}\right) \leq\ & \varphi_m^{\text{ub}}\left(\bar{\mathcal{U}}, \mathbf{\Omega} | \tilde{\bar{\mathcal{U}}}, \tilde{\bar{\mathbf{\Omega}}}\right) \\ \triangleq\ & -\log\det\left(\mathbf{\Gamma}_{\mathcal{S}_m}^H \mathbf{\Omega} \mathbf{\Gamma}_{\mathcal{S}_m}\right) + \sum_{i\in\mathcal{S}_m} \log\det \mathbf{\Theta}_i, \\ & + \frac{1}{\ln(2)} \sum_{i\in\mathcal{S}_m} \text{tr}\left(\mathbf{\Theta}_i^{-1} \times \left(\mathbf{\Gamma}_i^H \sum_{j\in\mathcal{N}_{\text{U}}} \bar{\mathbf{U}}_j \mathbf{\Gamma}_i + \mathbf{\Omega}_{ii} - \mathbf{\Theta}_i\right)\right), \end{aligned} \tag{3.21}$$

where $\mathbf{\Theta}_i \triangleq \mathbf{\Gamma}_i^H \sum_{j\in\mathcal{N}_{\text{U}}} \tilde{\bar{\mathbf{U}}}_j \mathbf{\Gamma}_i + \tilde{\bar{\mathbf{\Omega}}}_{ii}$. Similarly, the right-hand side function $C_{\text{fr},i}(\mathbf{p})$ of constraint (3.18b) needs to be converted to a concave function. In particular, (3.10) can be rewritten as

$$C_{\text{fr},i}\left(\mathbf{p}\right) = \log\left(\mathbf{f}_i \mathbf{p} + \sigma_{\text{R}}^2\right) - \log\left(\bar{\mathbf{f}}_i \mathbf{p} + \sigma_{\text{R}}^2\right), \tag{3.22}$$

where $\mathbf{f}_i \triangleq [f_{i,1}, f_{i,2}, \ldots, f_{i,N_{\text{R}}}]$ with its j-th element $f_{i,j} = \mathbf{h}_i \mathbf{w}_j \mathbf{w}_j^H \mathbf{h}_i^H$, and $\bar{\mathbf{f}}_i \triangleq [f_{i,1}, \ldots, f_{i,i-1}, 0, f_{i,i+1}, ..., f_{i,N_{\text{R}}}]$. In order to convexify $C_{\text{fr},i}(\mathbf{p})$, the second concave term $\log_2\left(\bar{\mathbf{f}}_i \mathbf{p} + \sigma_{\text{R}}^2\right)$ needs to be linearized by using the first order Taylor expansion. Hence, $C_{\text{fr},i}(\mathbf{p})$ can be replaced by its lower bound at a given point $\tilde{\mathbf{p}}$ as

$$\begin{aligned} C_{\text{fr},i}\left(\mathbf{p}\right) \geq\ & C_{\text{fr},i}^{\text{lb}}\left(\mathbf{p} | \tilde{\mathbf{p}}\right) \\ \triangleq\ & \log\left(\mathbf{f}_i \mathbf{p} + \sigma_{\text{R}}^2\right) - \log\left(\bar{\mathbf{f}}_i \tilde{\mathbf{p}} + \sigma_{\text{R}}^2\right) - \frac{1}{\ln(2)} \frac{\bar{\mathbf{f}}_i}{\bar{\mathbf{f}}_i \tilde{\mathbf{p}} + \sigma_{\text{R}}^2} (\mathbf{p} - \tilde{\mathbf{p}}). \end{aligned} \tag{3.23}$$

Due to the convexness of $\varphi_m^{\text{ub}}\left(\bar{\mathcal{U}}, \mathbf{\Omega}\right)$ and concaveness of $C_{\text{fr},i}^{\text{lb}}(\mathbf{p})$, constraint (3.18b) can be replaced by a convex inequality as

$$\varphi_m^{\text{ub}}\left(\bar{\mathcal{U}}, \mathbf{\Omega} | \tilde{\bar{\mathcal{U}}}, \tilde{\bar{\mathbf{\Omega}}}\right) \leq \sum_{i\in\mathcal{S}_m} C_{\text{fr},i}^{\text{lb}}\left(\mathbf{p} | \tilde{\mathbf{p}}\right), \tag{3.24}$$

which is a tighter constraint than (3.18b) and thus all solutions satisfying (3.24) are feasible to (3.18b).

According to (3.19)-(3.24), problem (3.18) can be reformulated as

$$\begin{aligned} \underset{\mathbf{p}\geq 0,\ \bar{\mathcal{U}}\succeq\mathbf{0},\ \mathbf{\Omega}\succeq\mathbf{0}}{\text{maximize}} \quad & \sum_{k\in\mathcal{N}_{\text{U}}} C_{\text{ac},k}^{\text{lb}}\left(\bar{\mathcal{U}}, \mathbf{\Omega} | \tilde{\bar{\mathcal{U}}}, \tilde{\bar{\mathbf{\Omega}}}\right) && \text{(3.25a)} \\ \text{s.t.} \quad & \text{(3.24), (3.18c) - (3.18e).} \end{aligned}$$

Furthermore, we relax the rank constraint (3.18e) [90] and the resulting problem becomes an SDR problem. As the constraint on the rank of $\bar{\mathcal{U}}$ is relaxed, there is no guarantee that the resulting solution $\bar{\mathcal{U}}^{\star}$ satisfies the rank constraint. However, we can prove that the rank of the optimal solution $\bar{\mathcal{U}}^{\star}$ obtained from the relaxed problem always satisfies the rank constraint (3.18e) and hence $\bar{\mathcal{U}}^{\star}$ is also an optimal solution to problem (3.25). We have the following theorem.

Theorem 3.3.1 The solution to problem (3.25), $\bar{\mathbf{U}}_k^{\star}, \forall k \in \mathcal{N}_{\mathrm{U}}$, always satisfies $\mathrm{rank}(\bar{\mathbf{U}}_k^{\star}) \leq \min(K, NN_{\mathrm{R}})$.

Proof 1 *Please see the Appendix.*

As a result, optimization problem (3.25) without rank constraint (3.18e) is a convex problem and can be solved easily using interior point methods such as the CVX optimization toolbox [91]. In addition, constraint (3.24) is locally tighter than the original constraint (3.18b) in problem (3.18) at the given point. Furthermore, objective function (3.25a) to be maximized is the lower bound of (3.18a). Therefore, the optimized variables derived from problem (3.25) are within the feasible domain of the original problem (3.18). The optimal precoding matrix $\mathbf{U}_j^{\star}, \forall j \in \mathcal{N}_{\mathrm{U}}$, can be obtained from $\bar{\mathbf{U}}_j^{\star}$ by using the well-known eigenvalue decomposition technique. Specifically, we have $\bar{\mathbf{U}}_j^{\star} = \bar{\mathbf{V}}\bar{\mathbf{D}}\bar{\mathbf{V}}^H$, where $\mathbf{D}$ is the diagonal matrix containing the eigenvalues and $\mathbf{V}$ is the matrix containing the corresponding eigenvectors. Therefore, $\mathbf{U}_j^{\star}$ is given by $\mathbf{U}_j^{\star} = \mathbf{V}\mathbf{D}^{\frac{1}{2}}$, where $\mathbf{D}$ is a diagonal matrix whose diagonal elements are the non-zero diagonal elements of $\bar{\mathbf{D}}$ and the columns of $\mathbf{V}$ are the corresponding eigenvectors.

Algorithm 1 outlines the steps for maximizing the sum-rate of all UEs. Due to the convexity of problem (3.25), the optimized achievable sum-rate is non-decreasing and the convergence is guaranteed to be achieved when the iteration number $r \to \infty$ [92].

3.3.1.b Separate Design of Fronthaul and Access Links

The proposed Algorithm 1 has offered a solution for the joint design of both the fronthaul and access links. However, despite the convexity of problem (3.25), the computational

Algorithm 1. Joint design based sum-rate maximization for perfect instantaneous CSI in downlink transmission.

1: **Input:** Essential system parameters, including fronthaul channels $\mathbf{H}$, access channels $\mathbf{G}$;

2: **Initialization:** Set $r \coloneqq 0$ and a feasible point $\{\mathbf{p}^{(0)} \geq 0, \bar{\mathcal{U}}^{(0)} \succeq \mathbf{0}, \mathbf{\Omega}^{(0)} \succeq \mathbf{0}\}$;

3: **repeat**

4: Update $r \coloneqq r + 1$;

5: Find the optimal solution $\{\mathbf{p}^{(r)}, \bar{\mathcal{U}}^{(r)}, \mathbf{\Omega}^{(r)}\}$ by solving convex problem (3.25) based on the given point $\{\mathbf{p}^{(r-1)}, \bar{\mathcal{U}}^{(r-1)}, \mathbf{\Omega}^{(r-1)}\}$:

$$\underset{\mathbf{p}^{(r)} \geq 0,\, \bar{\mathcal{U}}^{(r)} \succeq \mathbf{0},\, \mathbf{\Omega}^{(r)} \succeq \mathbf{0}}{\text{maximize}} \quad \sum_{k \in \mathcal{N}_{\mathrm{U}}} C_{\mathrm{ac},k}^{\mathrm{lb}} \left(\bar{\mathcal{U}}^{(r)}, \mathbf{\Omega}^{(r)} | \bar{\mathcal{U}}^{(r-1)}, \mathbf{\Omega}^{(r-1)} \right)$$

$$\text{s.t.} \quad \varphi_m^{\mathrm{ub}} \left(\bar{\mathcal{U}}^{(r)}, \mathbf{\Omega}^{(r)} | \bar{\mathcal{U}}^{(r-1)}, \mathbf{\Omega}^{(r-1)} \right) \leq \sum_{i \in \mathcal{S}_m} C_{\mathrm{fr},i}^{\mathrm{lb}} \left(\mathbf{p}^{(r)} | \mathbf{p}^{(r-1)} \right), \quad \forall m \in \mathcal{N}_{\mathcal{S}}, \tag{3.26a}$$

$$\sum_{i \in \mathcal{N}_{\mathrm{R}}} p_i^{(r)} \leq P_{\mathrm{C}}^{\max}, \tag{3.26b}$$

$$\sum_{k \in \mathcal{N}_{\mathrm{U}}} \mathrm{tr} \left(\mathbf{\Gamma}_i^H \bar{\mathbf{U}}_k^{(r)} \mathbf{\Gamma}_i \right) + \mathrm{tr} \left(\mathbf{\Omega}_{ii}^{(r)} \right) \leq P_{\mathrm{R},i}^{\max}, \ \forall i \in \mathcal{N}_{\mathrm{R}}, \tag{3.26c}$$

6: **until** convergence;

7: **Output:** $\{\mathbf{p}^{\star}, \bar{\mathcal{U}}^{\star}, \mathbf{\Omega}^{\star}\} \coloneqq \{\mathbf{p}^{(r)}, \bar{\mathcal{U}}^{(r)}, \mathbf{\Omega}^{(r)}\}$.

complexity is still significant, and the solution requires a large amount of time and memory, which will be further discussed in Subsection 3.4.1. For this reason, to reduce the complexity of the proposed problems, we propose to optimize the fronthaul and access links separately [28,29]. In particular, the fronthaul link is optimized at first and its corresponding optimization problem aims to maximize the fronthaul data rate $C_{\text{fr},i}(\mathbf{p}),\ \forall i \in N_{\text{R}}$, with respect to the power allocation variable $\mathbf{p}$. There are two optimization strategies to maximize the data rate for the system, namely max-min and sum-rate. The max-min approach aims to ensure that all users receive a minimum guaranteed data rate. In other words, the goal is to allocate resources in such a way that the data rate for the worst-performing user is maximized. On the other hand, the sum-rate approach aims to maximize the total data rate across all users, regardless of the individual channel qualities for each user.

For the **individual-rate** problem, we aim to maximize the individual rate of all $C_{\text{fr},i}(\mathbf{p}),\ \forall i \in \mathcal{N}_{\text{R}}$, that is

$$\underset{\mathbf{p}\geq 0}{\text{maximize}} \quad \min_{i\in\mathcal{N}_{\text{R}}} \quad C_{\text{fr},i}(\mathbf{p}), \tag{3.27a}$$

$$\text{s.t.} \quad \sum_{i\in\mathcal{N}_{\text{R}}} p_i \leq P_{\text{C}}^{\max}. \tag{3.27b}$$

For the **sum-rate** problem, we aim to maximize the sum of all fronthaul rates $\sum_{i\in\mathcal{N}_{\text{R}}} C_{\text{fr},i}(\mathbf{p})$ and the problem can be formulated as

$$\underset{\mathbf{p}\geq 0}{\text{maximize}} \quad \sum_{i\in\mathcal{N}_{\text{R}}} C_{\text{fr},i}(\mathbf{p}), \tag{3.28a}$$

$$\text{s.t.} \quad \sum_{i\in\mathcal{N}_{\text{R}}} p_i \leq P_{\text{C}}^{\max}. \tag{3.28b}$$

It is clear that the only difference between problems (3.27) and (3.28) is the objective function. The main reason for introducing two problems is that the entire system's design is not based solely on the fronthaul links. Simply ensuring that all RRHs have fair and ensured fronthaul rates may not be sufficient to guarantee that all UEs achieve the highest data rates via the access link. The comparison of two problems (3.27) and (3.28) will be discussed in Subsection 3.4.2.

By convexifying the objective function $C_{\text{fr},i}(\mathbf{p})$ as shown in (3.23), problem (3.27) can

easily be transformed into a convex problem as

$$\underset{\mathbf{p}\geq 0}{\text{maximize}} \quad \min_{i\in\mathcal{N}_{\text{R}}} \quad C^{\text{lb}}_{\text{fr},i}\left(\mathbf{p}|\tilde{\mathbf{p}}\right) \tag{3.29a}$$

$$\text{s.t.} \quad \text{(3.27b)}.$$

Similarly, the problem (3.28) can be rewritten as

$$\underset{\mathbf{p}\geq 0}{\text{maximize}} \quad \sum_{i\in\mathcal{N}_{\text{R}}} C^{\text{lb}}_{\text{fr},i}\left(\mathbf{p}|\tilde{\mathbf{p}}\right) \tag{3.30a}$$

$$\text{s.t.} \quad \text{(3.27b)}.$$

Hence, both convex problems (3.29) and (3.30) can be solved using the CVX optimization toolbox.

Next, by defining $\{C^{\star}\} \coloneqq \{C^{\star}_i | C^{\star}_i = C_{\text{fr},i}(\mathbf{p}^{\star}), i \in \mathcal{N}_{\text{R}}\}$ as the achievable fronthaul capacity, where $\mathbf{p}^{\star}$ is the optimal solution obtained from problem (3.29) or (3.30), we can formulate the corresponding sum-rate maximization problem for the access link as

$$\underset{\bar{\mathcal{U}}\succeq\mathbf{0},\ \mathbf{\Omega}\succeq\mathbf{0}}{\text{maximize}} \quad \sum_{k\in\mathcal{N}_{\text{U}}} C_{\text{ac},k}\left(\bar{\mathcal{U}},\mathbf{\Omega}\right) \tag{3.31a}$$

$$\text{s.t.} \quad \varphi_m(\bar{\mathcal{U}},\mathbf{\Omega}) \leq \sum_{i\in\mathcal{S}_m} C^{\star}_i,\ \forall m \in \mathcal{N}_{\mathcal{S}}, \tag{3.31b}$$

$$\text{(3.18d)},\ \text{(3.18e)}.$$

Similar to the transformation in (3.20) and (3.21) and relaxation of rank constraint (3.18e), non-convex problem (3.31) can be reformulated as

$$\underset{\bar{\mathcal{U}}\succeq\mathbf{0},\ \mathbf{\Omega}\succeq\mathbf{0}}{\text{maximize}} \quad \sum_{k\in\mathcal{N}_{\text{U}}} C^{\text{lb}}_{\text{ac},k}\left(\bar{\mathcal{U}},\mathbf{\Omega}|\tilde{\bar{\mathcal{U}}},\tilde{\mathbf{\Omega}}\right) \tag{3.32a}$$

$$\text{s.t.} \quad \varphi^{\text{ub}}_m\left(\bar{\mathcal{U}},\mathbf{\Omega}|\tilde{\bar{\mathcal{U}}},\tilde{\mathbf{\Omega}}\right) \leq \sum_{i\in\mathcal{S}_m} C^{\star}_i,\ \forall m \in \mathcal{N}_{\mathcal{S}}, \tag{3.32b}$$

$$\text{(3.18d)},$$

which is a convex problem and hence can be solved using the CVX optimization toolbox. Additionally, similar to the problem (3.25) in the previous Subsection 3.3.1.a, all the optimized values from problem (3.32) are also in the domain of the original problem (3.31).

In summary, the problem has been divided into two subproblems. The fronthaul link is optimized firstly with respect to transmit power $\mathbf{p}$ by solving problem (3.29) or (3.30). Using the optimized capacity of the fronthual channel, problem (3.32) optimizes the access link with respect to precoding matrices $\bar{\mathcal{U}}$ and quantization covariance matrices $\mathbf{\Omega}$. The detailed steps to solve problems (3.29) and (3.32) are given in Algorithm 2. To study the different optimization methods on the fronthaul capacity, both individual-rate and sum-rate maximization problems are included in Algorithm 2 based on problems (3.29) and (3.30), respectively.

Compared to Algorithm 1, the proposed separate design in Algorithm 2 decomposes the problem into two simpler optimization problems and thus has lower complexity than the proposed joint design. We will compare and discuss the achievable sum-rates and the actual optimization time of the proposed algorithms in Subsection 3.4.2.

3.3.2 Total Transmit Power Minimization

In addition to data rate, power consumption is a critical factor when designing communication networks. This section focuses on the design of a communication system that aims to minimize the total transmit power while ensuring the fulfillment of QoS requirements of achieving guaranteed minimum data rates for UEs. To achieve this goal, we propose a joint optimization of both the fronthaul and access links. Our objective is to minimize the total transmit power at both the CU and RRHs, taking into account the required data rates of the UEs, as well as the power budgets of the CU and RRHs. Depending on the specific situation, there are two approaches to ensuring that UEs achieve their target data rates. One way is to consider each UE individually, where each UE k has a desired rate γ_k, $\forall k \in \mathcal{N}_{\mathrm{U}}$. We have

$$C_{\mathrm{ac},k}(\bar{\mathcal{U}}, \mathbf{\Omega}) \geq \gamma_k, \ \forall k \in \mathcal{N}_{\mathrm{U}}. \tag{3.34}$$

Algorithm 2. Separate design based sum-rate maximization for perfect instantaneous CSI in downlink transmission.

Input: Essential system parameters, including fronthaul channels $\mathbf{H}$, access channels $\mathbf{G}$;

Initialization (first loop): Set $l := 0$ and a feasible point $\{\mathbf{p}^{(0)} \geq 0\}$;

repeat

Update $l := l + 1$;

Find the optimal solution $\mathbf{p}^{(l)}$ and corresponding set $\{C_{\text{fr},i}^{\text{lb}}(\mathbf{p}), \forall i \in \mathcal{N}_{\text{R}}\}$ based on the known $\mathbf{p}^{(l-1)}$ by solving one of the following problems:

- **Individual-rate** problem (3.29):
$$\begin{aligned} \underset{\mathbf{p}^{(l)} \geq 0}{\text{maximize}} \quad & \min_{i \in \mathcal{N}_{\text{R}}} \quad C_{\text{fr},i}^{\text{lb}}(\mathbf{p}^{(l)} | \mathbf{p}^{(l-1)}) \\ \text{s.t.} \quad & \sum_{i \in \mathcal{N}_{\text{R}}} p_i^{(l)} \leq P_{\text{C}}^{\max}; \end{aligned}$$

- **Sum-rate** problem (3.30):
$$\begin{aligned} \underset{\mathbf{p}^{(l)} \geq 0}{\text{maximize}} \quad & \sum_{i \in \mathcal{N}_{\text{R}}} C_{\text{fr},i}^{\text{lb}}(\mathbf{p}^{(l)} | \mathbf{p}^{(l-1)}) \\ \text{s.t.} \quad & \sum_{i \in \mathcal{N}_{\text{R}}} p_i^{(l)} \leq P_{\text{C}}^{\max}; \end{aligned}$$

until convergence;

update $\mathbf{p}^{\star} := \mathbf{p}^{(l)}$ and $\{C^{\star}\} := \{C_i^{\star} | C_i^{\star} = C_{\text{fr},i}(\mathbf{p}^{\star}), i \in \mathcal{N}_{\text{R}}\}$;

Initialization (second loop): Set $r := 0$ and a feasible point $\{\bar{\mathcal{U}}^{(0)} \succeq \mathbf{0}, \mathbf{\Omega}^{(0)} \succeq \mathbf{0}\}$;

repeat

Update $r := r + 1$;

Find the optimal solution $\{\bar{\mathcal{U}}^{(r)}, \mathbf{\Omega}^{(r)}\}$ by solving problem (3.32) based on the known $\{\mathbf{p}^{\star}, \bar{\mathcal{U}}^{(r-1)}, \mathbf{\Omega}^{(r-1)}\}$:

$$\begin{aligned} \underset{\bar{\mathcal{U}}^{(r)} \succeq \mathbf{0},\ \mathbf{\Omega}^{(r)} \succeq \mathbf{0}}{\text{maximize}} \quad & \sum_{k \in \mathcal{N}_{\text{U}}} C_{\text{ac},k}^{\text{lb}} \left(\bar{\mathcal{U}}^{(r)}, \mathbf{\Omega}^{(r)} | \bar{\mathcal{U}}^{(r-1)}, \mathbf{\Omega}^{(r-1)} \right) \\ \text{s.t.} \quad & \varphi_m^{\text{ub}} \left(\bar{\mathcal{U}}^{(r)}, \mathbf{\Omega}^{(r)} | \bar{\mathcal{U}}^{(r-1)}, \mathbf{\Omega}^{(r-1)} \right) \leq \sum_{i \in \mathcal{S}_m} C_i^{\star}, \ \forall m \in \mathcal{N}_{\mathcal{S}}, \\ & \sum_{k \in \mathcal{N}_{\text{U}}} \text{tr} \left(\mathbf{\Gamma}_i^H \bar{\mathbf{U}}_k^{(r)} \mathbf{\Gamma}_i \right) + \text{tr} \left(\mathbf{\Omega}_{ii}^{(r)} \right) \leq P_{\text{R},i}^{\max}, \ \forall i \in \mathcal{N}_{\text{R}}, \end{aligned}$$

until convergence;

Output: $\{\mathbf{p}^{\star}, \bar{\mathcal{U}}^{\star}, \mathbf{\Omega}^{\star}\} := \{\mathbf{p}^{(l)}, \bar{\mathcal{U}}^{(r)}, \mathbf{\Omega}^{(r)}\}$.

Another way is to consider the sum-rate of all the UEs, which means

$$\sum_{k\in\mathcal{N}_{\mathrm{U}}} C_{\mathrm{ac},k}(\bar{\mathcal{U}},\boldsymbol{\Omega}) \geq \gamma, \tag{3.35}$$

where γ is a given required sum-rate of the access link.

Similar to the sum-rate problem in the previous section, the total power minimization problem can thus be formulated as

$$\underset{\mathbf{p}\geq 0,\, \bar{\mathcal{U}}\succeq\mathbf{0},\, \boldsymbol{\Omega}\succeq\mathbf{0}}{\text{minimize}} \quad \sum_{k\in\mathcal{N}_{\mathrm{U}}} \mathrm{tr}\left(\bar{\mathbf{U}}_k\right) + \mathrm{tr}\left(\boldsymbol{\Omega}\right) + \sum_{i\in\mathcal{N}_{\mathrm{R}}} p_i \tag{3.36a}$$

$$\text{s.t.} \quad (3.34) \text{ or } (3.35), \tag{3.36b}$$

$$(3.18b) \text{ - } (3.18e).$$

Following the derivation steps (3.20) - (3.24) and the relaxation of rank constraint (3.18e), problem (3.36) can be transformed into the following SDR problems

- **Individual-rate:**

$$\underset{\mathbf{p}\geq 0,\, \bar{\mathcal{U}}\succeq\mathbf{0},\, \boldsymbol{\Omega}\succeq\mathbf{0}}{\text{minimize}} \quad \sum_{k\in\mathcal{N}_{\mathrm{U}}} \mathrm{tr}\left(\bar{\mathbf{U}}_k\right) + \mathrm{tr}\left(\boldsymbol{\Omega}\right) + \sum_{i\in\mathcal{N}_{\mathrm{R}}} p_i$$

$$\text{s.t.} \quad C^{\mathrm{lb}}_{\mathrm{ac},k}\left(\bar{\mathcal{U}},\boldsymbol{\Omega}|\tilde{\bar{\mathcal{U}}},\tilde{\boldsymbol{\Omega}}\right) \geq \gamma_k,\ \forall k\in\mathcal{N}_{\mathrm{U}}, \tag{3.37a}$$

$$(3.24), (3.18c), (3.18d),$$

- **Sum-rate:**

$$\underset{\mathbf{p}\geq 0,\, \bar{\mathcal{U}}\succeq\mathbf{0},\, \boldsymbol{\Omega}\succeq\mathbf{0}}{\text{minimize}} \quad \sum_{k\in\mathcal{N}_{\mathrm{U}}} \mathrm{tr}\left(\bar{\mathbf{U}}_k\right) + \mathrm{tr}\left(\boldsymbol{\Omega}\right) + \sum_{i\in\mathcal{N}_{\mathrm{R}}} p_i$$

$$\text{s.t.} \quad \sum_{k\in\mathcal{N}_{\mathrm{U}}} C^{\mathrm{lb}}_{\mathrm{ac},k}\left(\bar{\mathcal{U}},\boldsymbol{\Omega}|\tilde{\bar{\mathcal{U}}},\tilde{\boldsymbol{\Omega}}\right) \geq \gamma, \tag{3.37b}$$

$$(3.24), (3.18c), (3.18d),$$

which can be optimally solved using interior point algorithms such as the CVX optimization toolbox [91]. However, it is worth noting that setting a higher value for γ may render the problems infeasible. Similar to Algorithms 1 and 2 in Subsection 3.3.1, the optimized results obtained from the reformulated convex problem (3.37) is feasible to the original problem (3.36). Starting from a feasible point $\{\mathbf{p}^{(0)}\geq 0, \bar{\mathcal{U}}^{(0)}\succeq\mathbf{0}, \boldsymbol{\Omega}^{(0)}\succeq\mathbf{0}\}$, a local optimized result can be achieved by solving problem (3.37) until convergence. The algorithm to solve the transmit power minimization problem is detailed in Algorithm 3.

Algorithm 3. Joint design based transmit power minimization for perfect instantaneous CSI in downlink transmission.

1: **Input:** Essential system parameters, including fronthaul channels $\mathbf{H}$, access channels $\mathbf{G}$, and

- **Average Initialization:** an average distributed feasible point $\{\mathbf{p}^{(0)} \geq 0, \bar{\mathcal{U}}^{(0)} \succeq \mathbf{0}, \mathbf{\Omega}^{(0)} \succeq \mathbf{0}\}$;
- **Sum-rate Based Initialization:** an optimized feasible point $\{\mathbf{p}^{(0)} \geq 0, \bar{\mathcal{U}}^{(0)} \succeq \mathbf{0}, \mathbf{\Omega}^{(0)} \succeq \mathbf{0}\}$ obtained from the sum-rate maximization problem solved in Algorithm 1;

2: **Initialization:** Set $r := 0$;

3: **repeat**

4: Update $r := r + 1$;

5: Find the optimal solution $\{\mathbf{p}^{(r)}, \bar{\mathcal{U}}^{(r)}, \mathbf{\Omega}^{(r)}\}$ by solving problem (3.37) based on the given point $\{\mathbf{p}^{(r-1)}, \bar{\mathcal{U}}^{(r-1)}, \mathbf{\Omega}^{(r-1)}\}$:

- **Individual-rate:**

$$\begin{aligned} \underset{\mathbf{p}^{(r)} \geq 0,\, \bar{\mathcal{U}}^{(r)} \succeq \mathbf{0},\, \mathbf{\Omega}^{(r)} \succeq \mathbf{0}}{\text{minimize}} \quad & \sum_{k \in \mathcal{N}_{\mathrm{U}}} \mathrm{tr}\left(\bar{\mathbf{U}}_k^{(r)}\right) + \mathrm{tr}\left(\mathbf{\Omega}^{(r)}\right) + \sum_{i \in \mathcal{N}_{\mathrm{R}}} p_i^{(r)}, \\ \text{s.t.} \quad & C_{\mathrm{ac},k}^{\mathrm{lb}}\left(\bar{\mathcal{U}}^{(r)}, \mathbf{\Omega}^{(r)} | \bar{\mathcal{U}}^{(r-1)}, \mathbf{\Omega}^{(r-1)}\right) \geq \gamma_k, \ \forall k \in \mathcal{N}_{\mathrm{U}}, \\ & \text{(3.26a), (3.26b), (3.26c);} \end{aligned}$$

- **Sum-rate:**

$$\begin{aligned} \underset{\mathbf{p}^{(r)} \geq 0,\, \bar{\mathcal{U}}^{(r)} \succeq \mathbf{0},\, \mathbf{\Omega}^{(r)} \succeq \mathbf{0}}{\text{minimize}} \quad & \sum_{k \in \mathcal{N}_{\mathrm{U}}} \mathrm{tr}\left(\bar{\mathbf{U}}_k^{(r)}\right) + \mathrm{tr}\left(\mathbf{\Omega}^{(r)}\right) + \sum_{i \in \mathcal{N}_{\mathrm{R}}} p_i^{(r)}, \\ \text{s.t.} \quad & \sum_{k \in \mathcal{N}_{\mathrm{U}}} C_{\mathrm{ac},k}^{\mathrm{lb}}\left(\bar{\mathcal{U}}^{(r)}, \mathbf{\Omega}^{(r)} | \bar{\mathcal{U}}^{(r-1)}, \mathbf{\Omega}^{(r-1)}\right) \geq \gamma, \\ & \text{(3.26a), (3.26b), (3.26c);} \end{aligned}$$

6: **until** convergence;

7: **Output:** $\{\mathbf{p}^{\star}, \bar{\mathcal{U}}^{\star}, \mathbf{\Omega}^{\star}\} := \{\mathbf{p}^{(r)}, \bar{\mathcal{U}}^{(r)}, \mathbf{\Omega}^{(r)}\}$.

However, it is worth noting that by choosing the initial point based on the sum-rate maximization problem, which has been fully discussed in Subsection 3.3.1, a better locally optimized result may be achieved. The main concern, in this case, is to what extent the choice of the initialization point affects the final optimized results. Previously, a uniformly distributed initialization point (denoted as **average initial** method) was chosen among all the involved units in the system, such as allocating the total transmit power of CU equally and maximally to all antennas at CU. This initialization method is intended to ensure that the optimization process starts from a fair and balanced starting point, with each unit having an equal opportunity to contribute to the optimization results. To the best of our knowledge, there is no initialization method that has been proven to be superior to others in terms of low complexity, high feasibility and practicality. Except for the averaged initialization which has been used for the sum-rate maximization problem, in this case, another initialization point (denoted as **sum-rate based initial** method) can be utilized for the power minimization problem by taking the optimal solutions obtained from Algorithm 2. Apparently, this method incurs longer computational time and memory than the one that utilizes average initialization because the sum-rate maximization problem must be solved first. The impact of the initialization point on the final optimized results will be shown and compared in Subsection 3.4.3, in order to determine the effectiveness of using an optimized initialization point for the power minimization problem.

3.4 Numerical Results

In this section, we evaluate the performance of the proposed algorithms by simulating a 2 GHz downlink communication channel. The system parameters in our simulation, unless otherwise stated, are set as follows: $N_{\mathrm{R}} = 5$, $N_{\mathrm{U}} = 10$, $M = 200$, $N = 2$, $K = 1$, $P_{\mathrm{C}}^{\max} = 10$ W and $P_{\mathrm{R}}^{\max} \triangleq P_{\mathrm{R},i}^{\max} = 10$ W, $\forall i \in \mathcal{N}_{\mathrm{R}}$. We assume that the UEs are uniformly distributed [2] [19] inside a square area of 1 km^2. The heights of CU, RRHs, and

[2]Note that non-uniform UE distributions [93] can also be employed for simulations, without affecting the behavior of the proposed models and algorithms.

UEs are assumed as 25 m, 10 m, and 1.5 m from the ground, respectively. The minimum distance among all RRHs and between CU and RRHs is set as 100 m. The large-scale fading of the fronthaul link β^{LoS} and the access link α^{NLoS} are in line with the urban macrocell and urban microcell Street Canyon model in Table 7.4.1-1 of [94], respectively. The noise level σ_{R}^2, σ_{U}^2 are modelled as thermal noise with the room temperature 290 K and 5 dB noise figure over 20 MHz bandwidth channel.

To solve the problems, we use the CVX optimization toolbox (version 2.2) in MATLAB [91]. Although all of our problems have been transformed into SDR problems and they are theoretically guaranteed to be solvable, no solver is perfect, and there is no "best" solver that can outperform the others on every model [91] in terms of stability and speed. In our simulations, we employ both SDPT3 [95, 96] and SeDuMi [97] solvers comparing their feasibility and computational time in this section. The precision of the CVX solver is set to "best" with an iteration stopping precision of 10^{-2}.

While CVX is a powerful and well-developed tool for convex problems, it is important to note that "it will sometimes fail to converge even for problems known to have solutions" [91]. CVX applies the successive approximation method, particularly with functions like $\log\det(.)$ that we frequently use in our problems. This method is generally effective in solving most of our problems with concrete realizations and parameters, but it may still fail occasionally, even when confirmed solutions exist. Some potential solutions for solving $\log\det(.)$ problems, such as Mosek solver, are unable to solve most of our problems. The primary reason for this difficulty is that our optimization problems typically involve a large number of variables and constraints, which can require significant memory and computational resources to obtain results. Also, the high complexity of the proposed problems can also contribute to a lack of stability. Also, it seems CVX is not reliable with MATLAB parallel computing toolbox, at least not formally supported, which caused problems to allocate massive parallel realizations into multi-core HPC and evaluate the average performance over a large number of scenarios. Given the "imperfection" of CVX, thus, we will also evaluate the feasibility of our proposed problems using different solvers and algorithms in the actual simulations throughout this thesis.

To demonstrate the effectiveness of the proposed algorithms, we propose two benchmark schemes:

- **Benchmark 1: Independent block quantization [39].** In this benchmark scheme, we consider the independent block quantization method. Specifically, the signal intended for the i-th RRH is compressed independently, which means

$$\mathbb{E}[\boldsymbol{\omega}_i \boldsymbol{\omega}_j^H] = \mathbf{0}, \ \forall i \neq j \in \mathcal{N}_{\mathrm{R}}, \tag{3.38}$$

  and thus $\boldsymbol{\Omega}$ is a block-diagonal matrix in the optimization problem (3.18). The steps to solve the corresponding problem are the same as those for problem (3.18).

- **Benchmark 2: Independent diagonal quantization.** In this benchmark scheme, each element of the signal $\mathbf{x}_{\mathrm{R},i}$ is compressed and quantized independently. Therefore, there is no correlation among all the quantization noise elements, which means

$$\boldsymbol{\Omega}_{ii} = \mathbb{E}[\boldsymbol{\omega}_i \boldsymbol{\omega}_i^H] = \mathrm{diag}\left(\omega_{i,1}, \cdots, \omega_{i,N}\right), \ \forall i \in \mathcal{N}_{\mathrm{R}}, \tag{3.39}$$

  where $\omega_{i,1}, \cdots, \omega_{i,N}$ denote the quantization power for each antenna. Therefore, $\boldsymbol{\Omega}$ is a diagonal real-valued matrix. The steps to solve the corresponding problem are the same as those used for problem (3.18).

It is worth noting that for Benchmarks 1 and 2, the term $\log\det\left(\boldsymbol{\Gamma}_{\mathcal{S}_m}^H \boldsymbol{\Omega} \boldsymbol{\Gamma}_{\mathcal{S}_m}\right)$ in the quantization condition (3.11) could be expressed as $\sum_{i\in\mathcal{S}_m} \log\det\left(\boldsymbol{\Gamma}_i^H \boldsymbol{\Omega} \boldsymbol{\Gamma}_i\right)$. Therefore, the constraint (3.18b) can be collapsed into a simple version

$$\varphi_i(\bar{\mathcal{U}}, \boldsymbol{\Omega}) \leq C_{\mathrm{fr},i}(\mathbf{p}), \ i \in \mathcal{N}_{\mathrm{R}}. \tag{3.40}$$

Previously, the number of constraint (3.18b) is $2^{N_{\mathrm{R}}} - 1$ and the size of optimized variable $\boldsymbol{\Omega}$ is $N^2 N_{\mathrm{R}}^2$ for Algorithm 1 and 2. However, the number of constraint (3.40) is only N_{R} and $\boldsymbol{\Omega}$ has $N^2 N_{\mathrm{R}}$ and $N N_{\mathrm{R}}$ elements to optimize in the Benchmarks 1 and 2, respectively. The computational complexity is related to the number of constraints and the size of optimized variables [91].

3.4.1 Convergence and Complexity Comparison

To compare different algorithms and solvers in terms of computational time, feasibility, convergence, and achievable optimized results such as sum-rates and power consumption, we conduct a simulation using 100 randomly generated sets of RRH and UE positions. Fronthaul channel models are generated based on the 3-D distances between CU and RRH antenna elements, as described in equation (3.1). For each set of locations and fronthaul model, we generate 5 random access channel models to account for extreme scenarios. Therefore, there are 500 realizations in total, of which the following results are based on the average performance. The metric we adopt in this thesis to measure the complexity of the solutions is the actual simulation running time. The simulations are performed using MATLAB 2022a, with CVX (version 2.2) as the accompanying software platform, while the hardware setup consists of 2 CPU cores (2 x AMD EPYC 7702) and 256GB RAM.

Figures 3.2 and 3.3 compare the computational running time of both solvers for different algorithms. Figure 3.2 shows the actual running time for optimization problem within each iteration while Figure 3.3 displays the overall time required for the proposed algorithms to converge. Although the y-axes have different scales, both figures exhibit similar trends. This suggests that two figures have a similar proportionality constant, which represents the number of interactions required by the proposed algorithms to converge. Due to the low computational complexity of Benchmarks 1 and 2, it is reasonable to observe significantly

Table 3.1: Feasibility of joint design in Algorithm 1, separate design in Algorithm 2, and two benchmarks using solvers SDPT3 and SeDuMi.

Solver / Algorithm	SDPT3	SeDuMi
Joint design	100.0%	100.0%
Separate design - Ind-rate	100.0%	100.0%
Separate design - Sum-rate	94.8%	99.8%

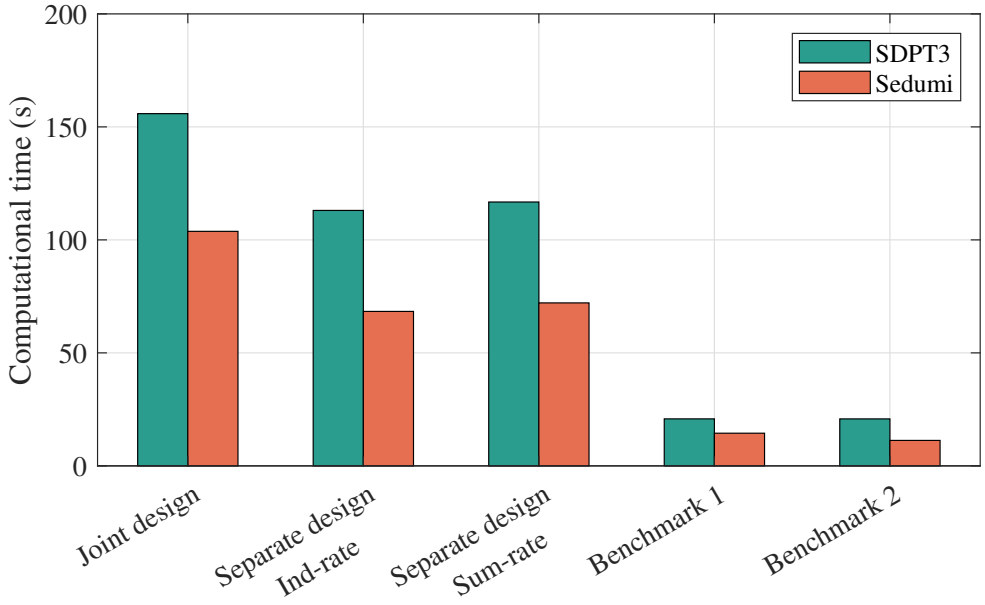


Figure 3.2: Actual average running time within each iteration of the proposed algorithms for both SDPT3 and SeDuMi solvers. The computational time is based on MATLAB running on 2 CPU cores (2 x AMD EPYC 7702) with 256GB RAM.

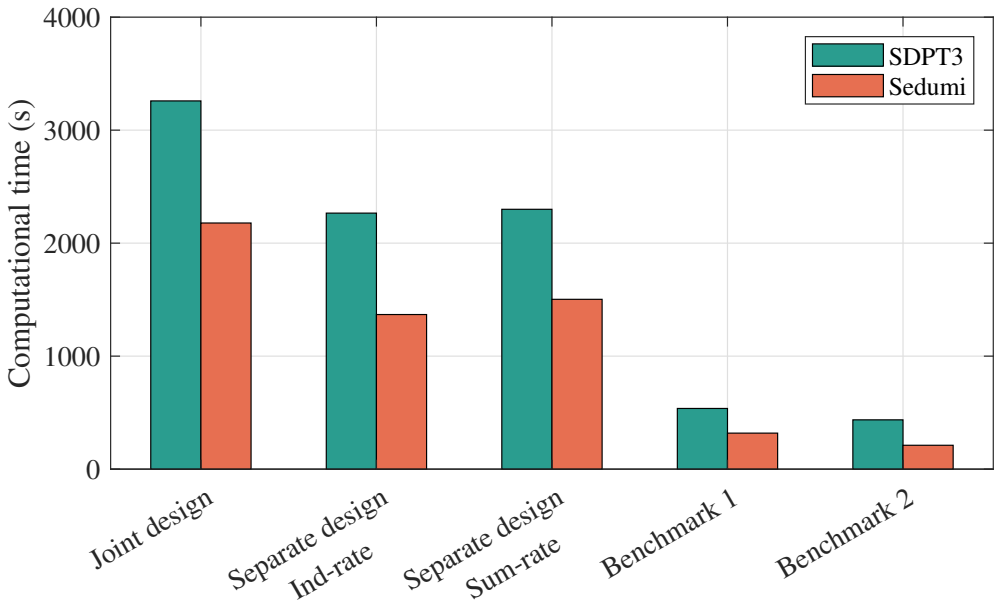


Figure 3.3: Total running time of the proposed algorithms for both SDPT3 and SeDuMi solvers. The computational time is based on MATLAB running on 2 CPU cores (2 x AMD EPYC 7702) with 256GB RAM.

lower running time for these approaches in comparison to the proposed algorithms.

Figure 3.4 illustrates the optimized results for different approaches, where both solvers perform similarly. Joint design and separate design with sum-rate fronthaul design stand out from the other methods, achieving a significantly higher sum-rate. It is noteworthy that these two designs also require the longest computational time, as shown in Figures 3.2 and 3.3. The final obtained sum-rate is similar for both solvers. As shown in Table 3.1, it is clear that both solvers demonstrate high feasibility.

Figure 3.5 displays the convergence speed of various methods. We only show the results obtained using the SDPT3 solver here as SeDuMi exhibits a similar trend. It is apparent that all methods exhibit nearly identical non-decreasing growth rates and converge within 20 to 30 iterations.

3.4.2 Sum-rate Maximization Problem

Now, we evaluate the impact of several key parameters on the achievable sum-rate using joint design, detailed in Algorithm 1, and separated design, presented in Algorithm 2. Specifically, we investigate how some parameters affect the performance, including the number of RRHs and UEs operating in the system, the number of antennas equipped on these units, and the power budgets. To evaluate the performance of different scenarios, we generate hundreds of scenarios with randomly located RRHs and UEs, along with hundreds of random channel realizations. The final results are obtained as the average of all realizations. In this section, 100 random locations of RRHs and UEs with their corresponding fronthaul and access links are generated for simulations. The following results are based on the average performance of these 100 models.

Figure 3.6 shows the impact of the number of antennas M at the CU on the achievable sum-rate. It can be observed that the achievable sum-rate keeps increasing with the increasing number of antennas at the CU. It is clear that the joint design in Algorithm 1 outperforms separate designs and benchmarks, as demonstrated not only in Figure 3.6 but

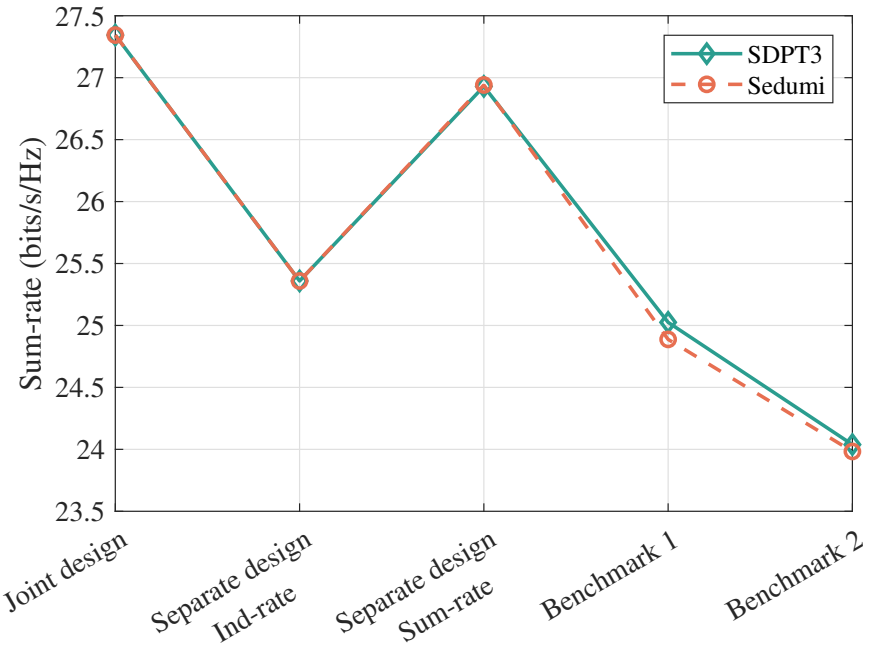


Figure 3.4: Achievable sum-rate of the proposed algorithms and benchmarks using the SDPT3 and SeDuMi solvers.

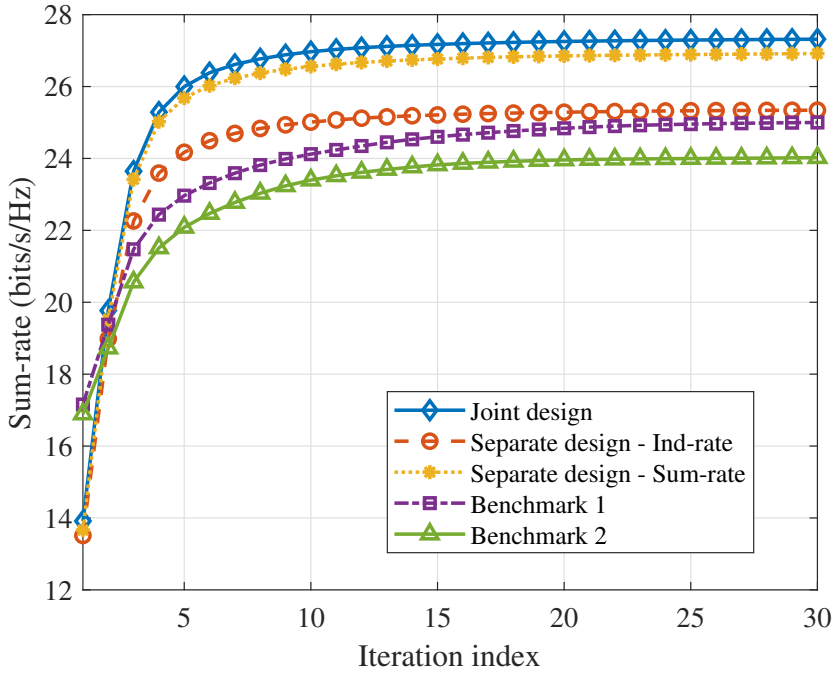


Figure 3.5: Convergence speed of the proposed algorithms and benchmarks.

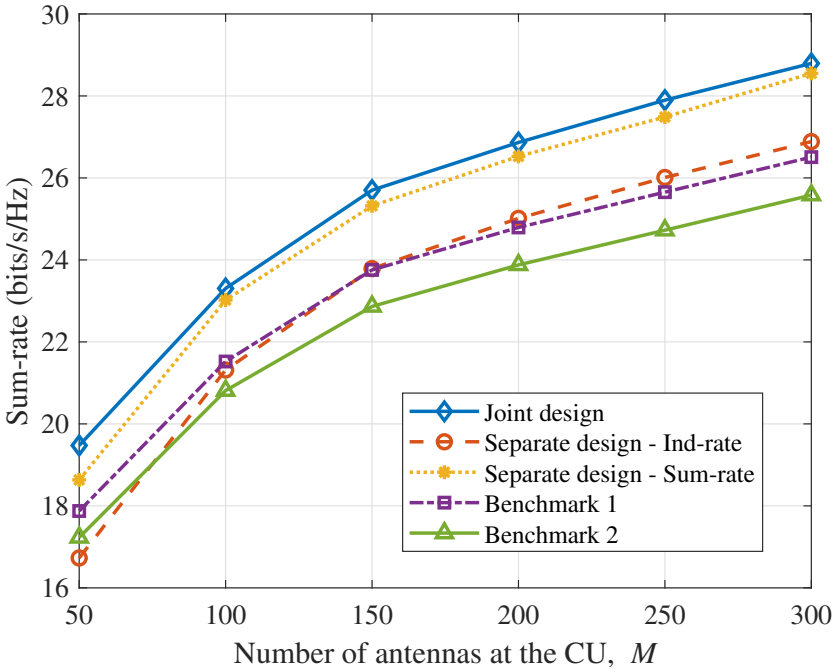


Figure 3.6: Achievable sum-rate versus number of antennas M at the CU.

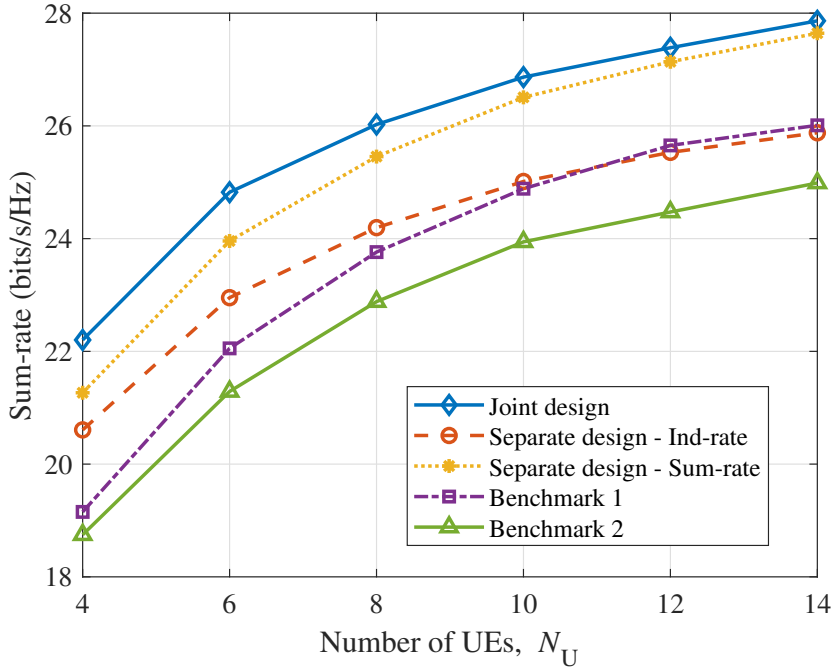


Figure 3.7: Achievable sum-rate versus number of UEs N_{U}.

also across all the following results. This figure indicates that the separate design with optimized fronthaul sum-rate achieves a network sum-rate close to that of the joint design, and much better than the separate design with individually optimized fronthaul rates. This difference is attributed to the variable $\{C^\star\}$ that connects two separate problems in Algorithm 2. The separate fronthaul design with individual-rate aims to maximize $\min_i C_i^\star$ by taking the worst channel into account, while the one with sum-rate maximizes $\sum_i C_i^\star$ by taking the summation, which is the one consistent with the objective of the main sum-rate maximization problem (3.31). Additionally, when the number of antennas $M \leq 100$, the separate design exhibits a relatively low sum-rate. This observation indicates that the separate design is more sensitive to the limitations of the fronthaul capacity.

Figure 3.7 depicts the achievable sum-rate as a function of the number of UEs. The sum-rate increases as the number of UEs increases, but the rate of growth decreases with more UEs. The achievable rates are not proportional to the number of UEs, indicating that the average resource allocated to each UE decreases as more UEs are added and may be saturated with the larger number of UEs.

Figures 3.8 and 3.9 investigate the impact of the transmit power on the achievable rates, where Figures 3.8 focuses on $P_{\mathrm{C}}^{\max}$ and Figure 3.9 on $P_{\mathrm{R},i}^{\max}$. It is observed from Figure 3.8 that $P_{\mathrm{C}}^{\max}$ has little effect on the achievable sum-rates, even if it is relatively small, etc., $P_{\mathrm{C}}^{\max} \leq 1\mathrm{W}$. On the other hand, we observe that the optimized sum-rate increases rapidly when $P_{\mathrm{R},i}^{\max} \leq 5\mathrm{W}$ and continues to increase with a higher power allocation, as shown in Figure 3.9. The reason behind this is that the fronthaul link uses massive MIMO, where channel hardening can be achieved with larger numbers of antennas M even with very little transmit power. It proves that massive MIMO is an effective technique for consuming less power while ensuring a high transmission rate. Referring back to Figure 3.6, adding more antennas at the CU could potentially further improve the sum-rate of the access link without requiring a large amount of energy. Figure 3.9 shows that the transmit power budgets on the RRHs strongly affect the achievable rate, especially when $P_{\mathrm{R},i}^{\max} \in (0, 5)\mathrm{W}$. One potential solution to increase the network sum-rate could be adding more transmit antennas for the access link.

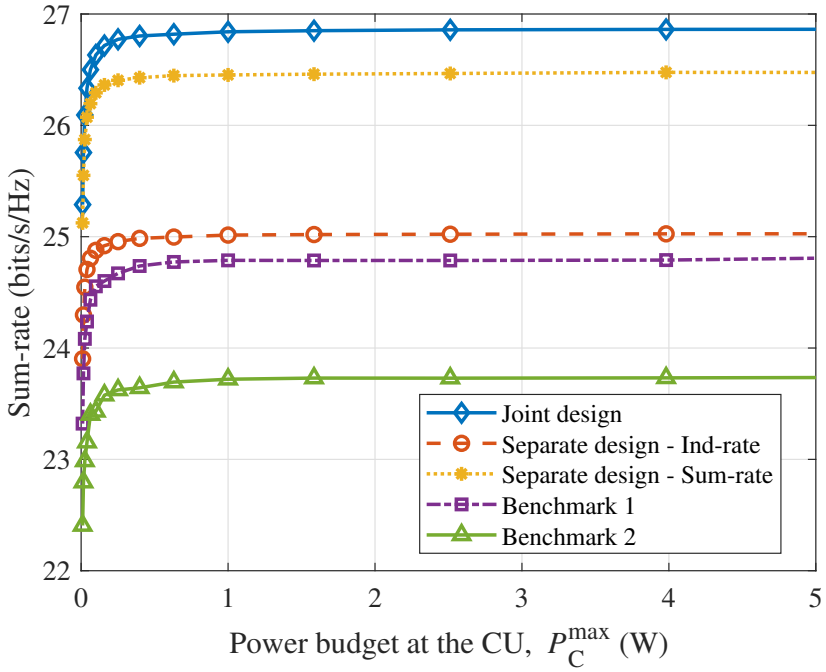


Figure 3.8: Achievable sum-rate versus transmit power budget P_C^{max} at the CU.

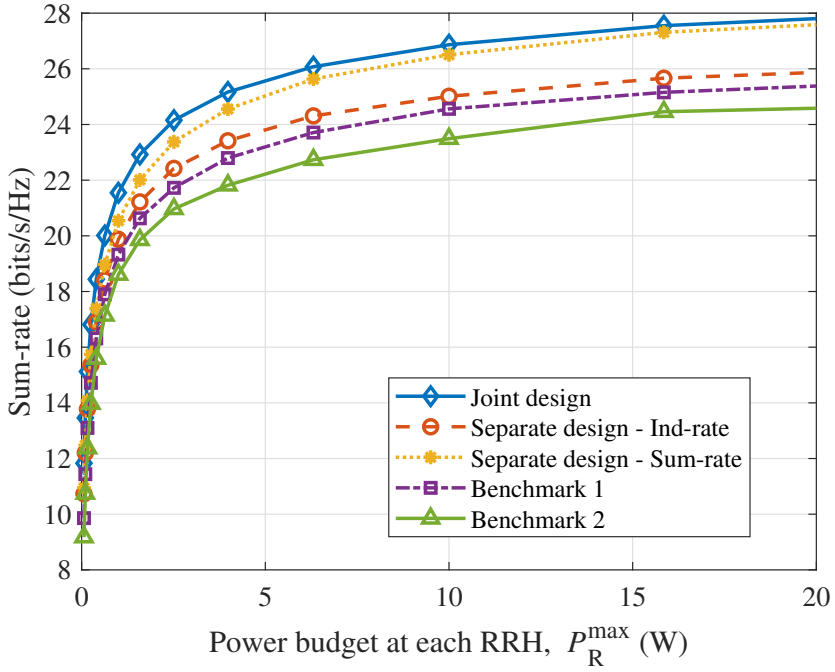


Figure 3.9: Achievable sum-rate versus transmit power budget P_R^{max} at each RRH.

3.4.3 Total Power Minimization Problem

Algorithm 3 aims to minimize the total transmit power while ensuring a given sum-rate for the proposed network. However, it is possible that no solutions can satisfy a high sum-rate under the given power budget. To address this, a feasibility check of the proposed problem can be performed to confirm a proper value of the achievable sum-rate before minimizing the total power consumption. The feasibility problem is based on problem (3.37) and can be formulated as

$$\begin{aligned} \underset{\mathbf{p}\succeq 0,\, \bar{\mathcal{U}}\succeq \mathbf{0},\, \mathbf{\Omega}\succeq \mathbf{0}}{\text{minimize}} \quad & 0 \\ \text{s.t.} \quad & \text{(3.37a) or (3.37b)}, \\ & \text{(3.24), (3.18c), (3.18d)}. \end{aligned} \tag{3.41}$$

If the problem (3.41) can be solved, it means that the given sum-rate is achievable and the power consumption might be further reduced by applying Algorithm 3 and vice versa.

We first study the impact of constraint (3.37a), which ensures minimum individual-rate for each UE, solved in Algorithm 3. Figure 3.10 plots the feasibility of problem (3.41) against the minimum required rate per user for various values of $P_{\mathrm{C}}^{\max}$ and $P_{\mathrm{R}}^{\max}$ with different compression schemes, namely the multivariate and independent point-to-point schemes. The initialization method is based on average initialization. It can be observed that the feasibility increases with increasing power budgets $P_{\mathrm{C}}^{\max}$ and $P_{\mathrm{R}}^{\max}$. However, independent compression can achieve higher feasibility compared to multivariate compression under all power constraints. The lower feasibility of multivariate compression used in Algorithm 3 is due to the large number of inequalities in (3.24), which is given as $2^{N_{\mathrm{R}}} - 1 = 31$, while Benchmark 1 only has five inequalities for this constraint.

Figure 3.11 investigates the total power consumption versus the number of antennas M under different minimum required rates. The optimized power consumption achieved by the proposed algorithm notably decreases with increasing M whereas the gain is relatively flat when $M \geq 200$. Furthermore, it is shown that a higher requirement of the minimum achievable rate results in the need for more transmit power, and the proposed joint design

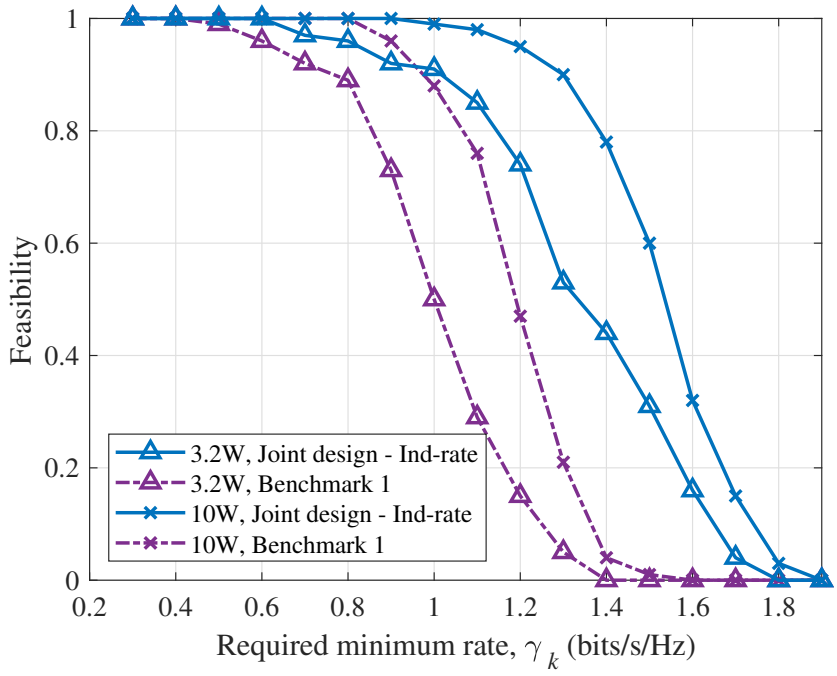


Figure 3.10: Feasibility of the proposed scheme versus minimum required rate per UE, and for different power constraints $P_{\mathrm{C}}^{\max} = P_{\mathrm{R}}^{\max}, \forall i \in \mathcal{N}_{\mathrm{R}}$.

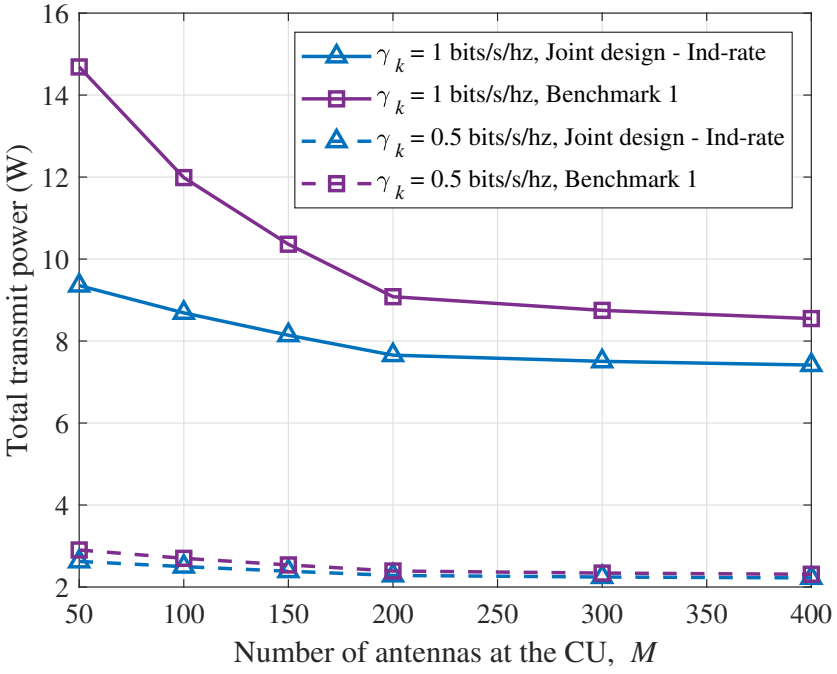


Figure 3.11: Total power consumption of the proposed algorithm versus the number of antennas M at the CU with $\gamma_k = 1$ bits/s/Hz and $\gamma_k = 0.5$ bits/s/Hz.

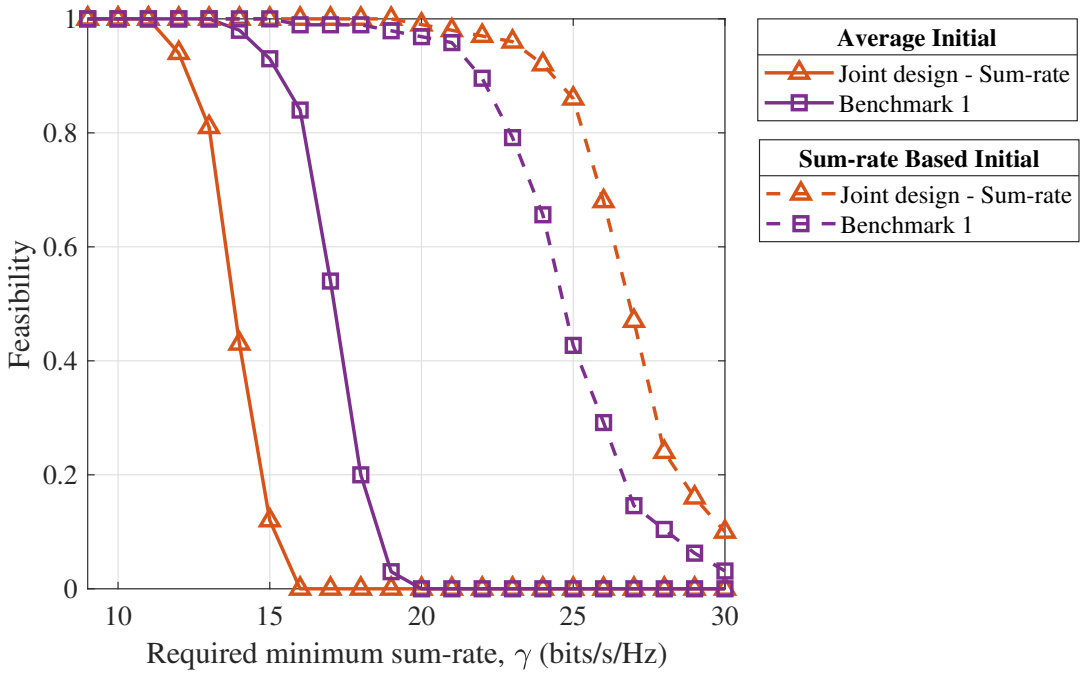


Figure 3.12: Feasibility of the proposed scheme versus required sum-rate with different initialization methods.

solved in Algorithm 3 with multivariate compression saves more power than independent compression.

Next, we focus on the impact of constraint (3.37b), which ensures the achievable sum-rate of all UEs. The effects of utilizing different initialization methods, based on both the average allocation and the optimal variables obtained from Algorithm 1, are also compared. We begin by evaluating the feasibility of the proposed joint design by simulating problem (3.41) with sum-rate constraint (3.37b), as shown in Figure 3.12. Similar to Figure 3.10, the independent compression yields higher feasibility than the multivariate compression when utilizing the average initialization. However, by applying the optimized solutions as the initial point to minimize the total power, the feasibility of multivariate compression surpasses that of independent compression. In addition, both methods demonstrate higher feasibility in comparison to the usage of average initialization. It is important to note that, in this method, the sum-rate maximization problem given in Algorithm 1 must be solved prior to the total power minimization problem.

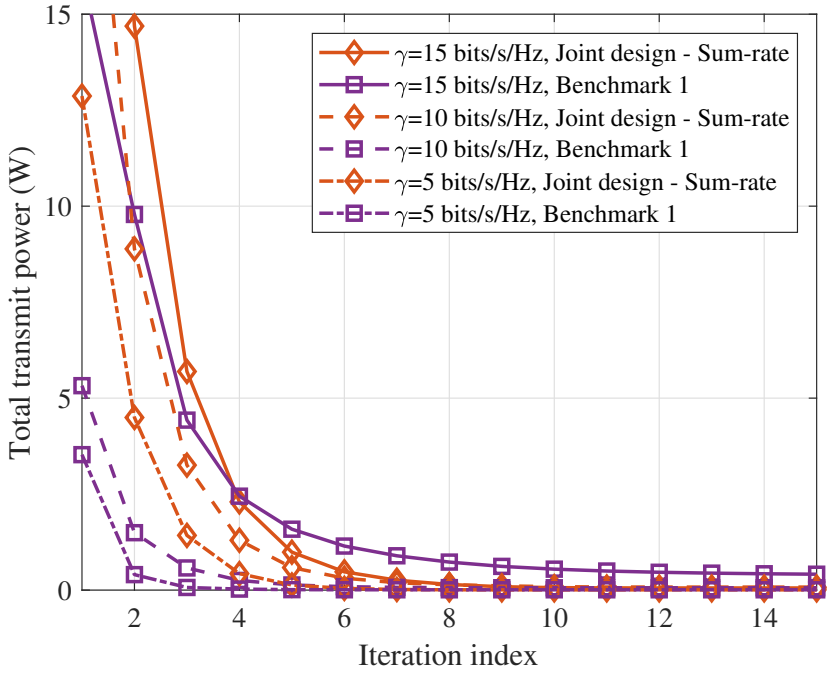


Figure 3.13: Convergence speed of the joint design proposed in Algorithm 3 against the number of iterations with average initialization.

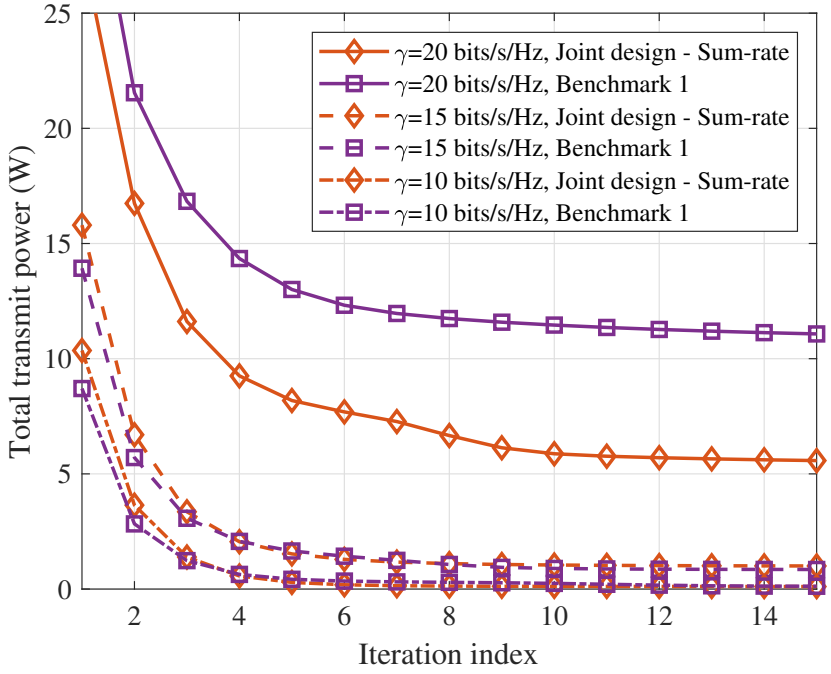


Figure 3.14: Convergence speed of the joint design proposed in Algorithm 3 against the number of iterations with sum-rate based initialization.

Figures 3.13 and 3.14 investigate the transmit power with different requirements of minimum achieved sum-rate γ. Figure 3.13 considers both the joint design and Benchmark 1 initialized by the uniformly allocated power and precoding matrices. It can be observed that a higher sum-rate γ requires larger power consumption. For all values of γ, the joint design outperforms Benchmark 1, which indicates that the total power can be further reduced by applying the joint design. The same pattern is also reflected in Figure 3.14, wherein the optimal solution obtained from the sum-rate maximization problem given in Algorithm 1 is utilized as the starting point for solving the total power minimization problems. However, although high transmit power is required, it is feasible to solve the power minimization problem with $\gamma = 20$ bits/s/Hz in Figure 3.14, whereas in Figure 3.13, it is not possible due to the low feasibility of the average initialization method.

3.5 Conclusion

In this chapter, both joint and separate designs of the wireless fronthaul and access links were proposed in a CRAN architecture featuring a massive MIMO CU. The proposed designs focused on the downlink transmission, and assumed that the CU has access to the perfect instantaneous CSI, on which the precoding designs of the access link were based. The DCF compression scheme was adopted at the RRHs and the multivariate compression noise was considered for the system optimizations. In particular, the optimization problems were formulated aiming to maximize the sum-rate and minimize the total transmit power. The problems were constrained by the power budgets of the CU and RRHs, as well as the need to satisfy compression constraints to recover compressed signals successfully.

The resulting optimization problems were non-convex, making them difficult to solve. However, by replacing the optimized variables, we were able to convexify the non-convex functions and relax the rank constraint. This led to the formulation of convex SDR problems, which can be optimally solved using interior point algorithms such as the CVX optimization toolbox. We have provided a proof that the optimal results obtained

from the reformulated problems are valid for the original optimization problems, as the variable substitution does not alter the feasibility of the original problems. To solve the reformulated problems, we proposed iterative algorithms with guaranteed convergence, resulting in optimal solutions for the original problem.

Comprehensive numerical results have been provided in this chapter. We first discussed the feasibility and computational complexity of the proposed algorithms and benchmarks using two of the most common solvers, namely SDPT3 and SeDuMi, while also presenting the convergence speed. We also performed numerous simulations to evaluate the impact of system parameters, including the number of antennas M at the CU, the number of UEs N_{U}, and power budgets $P_{\mathrm{C}}^{\max}$ and $P_{\mathrm{R}}^{\max}$, on the proposed algorithms and benchmarks. The numerical results clearly demonstrated that the proposed joint design outperforms the separate designs and both benchmarks in terms of the achieved sum-rate. However, it should be noted that the separate designs and both benchmarks have significantly lower complexity, and therefore require less computational time and memory. Furthermore, the implementation of massive MIMO in the fronthaul link was observed to significantly enhance the achievable sum-rate, suggesting the potential to also increase the sum-rate for the access link for future models. Additionally, a superior decrease in power consumption can be achieved by the proposed joint design compared to the separate designs and two benchmark schemes.

Chapter 4

Design of Wireless Downlink Transmission in Massive MIMO CRANs with Stochastic CSI

4.1 Introduction

Obtaining perfect instantaneous CSI for the access link at the CU can improve system design and increase data throughput. However, the acquisition of this information requires a significant amount of resources, such as transmitting training sequences and forwarding received pilots from the RRHs to the CU. Therefore, in this chapter, we consider a more realistic scenario where only stochastic CSI for the access link is available at the CU [98]. To design the system, we formulate two non-convex problems to maximize the ergodic sum-rate and minimize the total transmit power, respectively. Similar to Chapter 3, both joint and separate designs are proposed for both problems. However, the presence of expectation operations within the objective expressions makes the optimization problems challenging to solve. Therefore, we employ the *sample average approximation* (SAA) method to approximate the ergodic sum-rate. Then, we transform the problems into

convex SDR problems using the SCA method, which can be solved using interior-point methods. The iterative algorithms are presented to solve the reformulated problems and compare their performance with two proposed benchmark schemes.

In Section 4.2, we propose the joint and separate designs, along with their respective optimization problems. These designs have the primary objectives of maximizing the ergodic sum-rate and minimizing total power consumption. Also, the iterative algorithms are presented and compared with benchmarks. The numerical results in Section 4.3 demonstrate the superiority of the proposed algorithms. Finally, we conclude the chapter in Section 4.4.

4.2 Problem Formulation

While the fronthaul link can achieve a deterministic LoS channel by appropriately placing the CU and RRHs, the presence of rich scatters resulting from the mobility and large number of UEs creates challenges for the acquisition of CSI in the access channels. Instead, the system design based on the stochastic CSI is of practical interest.

Since perfect instantaneous CSI of the access link is unavailable, the channel precoding matrices $\bar{\mathcal{U}}$ and compression noise $\mathbf{\Omega}$ are designed and optimized for all coherence blocks. In the following, optimization problems with stochastic CSI are formulated and the corresponding algorithms to solve them are outlined and discussed.

4.2.1 Ergodic Sum-rate Maximization

Similar to the sum-rate maximization problem (3.18) for perfect instantaneous CSI, we first aim to maximize the sum-rate of the system by optimizing transmit power $\mathbf{p}$, precoders $\bar{\mathcal{U}}$ and quantization noise covariance matrix $\mathbf{\Omega}$, which is subjected to the quantization constraint and power limits at CU and each RRH.

4.2.1.a Joint Design of Fronthaul and Access Links

We follow the same steps used in Subsection 3.3.1 for the joint design of both fronthaul and access links. The optimization problem can thus be formulated as

$$\underset{\mathbf{p}\geq 0,\, \bar{\mathcal{U}}\succeq \mathbf{0},\, \mathbf{\Omega}\succeq \mathbf{0}}{\text{maximize}} \sum_{k\in\mathcal{N}_{\mathrm{U}}} R_k \tag{4.1a}$$

$$\text{s.t.} \quad \varphi_m(\bar{\mathcal{U}}, \mathbf{\Omega}) \leq \sum_{i\in\mathcal{S}_m} C_{\mathrm{fr},i}(\mathbf{p}),\ \forall m \in \mathcal{N}_{\mathcal{S}}, \tag{4.1b}$$

$$\sum_{i\in\mathcal{N}_{\mathrm{R}}} p_i \leq P_{\mathrm{C}}^{\max}, \tag{4.1c}$$

$$\sum_{k\in\mathcal{N}_{\mathrm{U}}} \mathrm{tr}\left(\mathbf{\Gamma}_i^H \bar{\mathbf{U}}_k \mathbf{\Gamma}_i\right) + \mathrm{tr}\left(\mathbf{\Omega}_{ii}\right) \leq P_{\mathrm{R},i}^{\max},\ \forall i \in \mathcal{N}_{\mathrm{R}}, \tag{4.1d}$$

$$\mathrm{rank}\left(\bar{\mathbf{U}}_k\right) \leq n_{\mathrm{R}},\ \forall k \in \mathcal{N}_{\mathrm{U}}, \tag{4.1e}$$

where R_k represents the ergodic achievable rate for the k-th UE and is defined as [77, 98]

$$R_k \triangleq \mathbb{E}\left[C_{\mathrm{ac},k}\left(\bar{\mathcal{U}}, \mathbf{\Omega}\right)\right]. \tag{4.2}$$

Constraint (4.1b) is the condition in order to dequantize the received signals successfully. Constraints (4.1c) and (4.1d) are the transmit power constraints and (4.1e) represents the rank constraint.

It is obvious that this problem is hard to solve due to the objective function (4.1a) as the actual channel realization $\mathbf{G}$ is unknown. To tackle this problem, we use the SAA method [38, 99] to obtain an approximate of the ergodic sum-rate. Precisely, we assume that there are n_{B} finite coherence blocks within the transmission time and thus R_k could be rewritten as

$$R_k = \frac{1}{n_{\mathrm{B}}} \sum_{n=1}^{n_{\mathrm{B}}} C_{\mathrm{ac},k}^{(n)}\left(\bar{\mathcal{U}}, \mathbf{\Omega}\right),\ \text{when } n \to \infty, \tag{4.3}$$

where $C_{\mathrm{ac},k}^{(n)}\left(\bar{\mathcal{U}}, \mathbf{\Omega}\right)$ represents the capacity corresponding to the l-th channel $\mathbf{G}_k^{(l)}$. Therefore, problem (4.1) can be re-expressed as

$$\underset{\mathbf{p}\geq 0,\, \bar{\mathcal{U}}\succeq \mathbf{0},\, \mathbf{\Omega}\succeq \mathbf{0}}{\text{maximize}} \quad \frac{1}{n_{\mathrm{B}}} \sum_{n=1}^{n_{\mathrm{B}}} \sum_{k\in\mathcal{N}_{\mathrm{U}}} C_{\mathrm{ac},k}^{(n)}\left(\bar{\mathcal{U}}, \mathbf{\Omega}\right) \tag{4.4a}$$

$$\text{s.t.} \quad \text{(4.1b) - (4.1e).}$$

Apparently, the reformulated optimization problem (4.4) is still hard to solve due to the non-convex function $C_{\mathrm{ac},k}^{(n)}\left(\bar{\mathcal{U}},\mathbf{\Omega}\right)$ in the objective function (4.4a) and $\varphi_m(\bar{\mathcal{U}},\mathbf{\Omega})$ and $C_{\mathrm{fr},i}(\mathbf{p})$ in constraint (4.1b). Following the derivation (3.20) - (3.24) and the relaxation of rank constraint (4.1e), problem (4.4) can be re-expressed as a convex problem

$$\underset{\mathbf{p}\geq 0,\,\bar{\mathcal{U}}\succeq\mathbf{0},\,\mathbf{\Omega}\succeq\mathbf{0}}{\text{maximize}} \quad \frac{1}{n_{\mathrm{B}}}\sum_{n=1}^{n_{\mathrm{B}}}\sum_{k\in\mathcal{N}_{\mathrm{U}}} C_{\mathrm{ac},k}^{\mathrm{lb},(n)}\left(\bar{\mathcal{U}},\mathbf{\Omega}|\tilde{\bar{\mathcal{U}}},\tilde{\mathbf{\Omega}}\right) \tag{4.5a}$$

$$\text{s.t.} \quad \varphi_m^{\mathrm{ub}}\left(\bar{\mathcal{U}},\mathbf{\Omega}|\tilde{\bar{\mathcal{U}}},\tilde{\mathbf{\Omega}}\right) \leq \sum_{i\in\mathcal{S}_m} C_{\mathrm{fr},i}^{\mathrm{lb}}\left(\mathbf{p}|\tilde{\mathbf{p}}\right),\ \forall m\in\mathcal{N}_{\mathcal{S}}, \tag{4.5b}$$

$$(4.1c),\ (4.1d),$$

where constraint (4.5b) is locally tighter than the original constraint (4.1b) and the objective function (4.4a) is locally lower bounded by (4.5a). Therefore, all solutions satisfying (4.5) are feasible to original problem (4.4).

Algorithm 4 outlines the steps for maximizing the average sum-rate over the given n_{B} channels used to approximate the ergodic sum-rate. Due to the convexity of problem (4.5), the optimized ergodic achievable sum-rate is non-decreasing and the convergence is guaranteed to be achieved when the iteration number $r\to\infty$ [92].

4.2.1.b Double-loop Joint Design of Fronthaul and Access Links

Although the complexity of Algorithm 4 is reasonably low due to the convexity of problem (4.5), it is vital to know the impact of channel numbers n_{B} on the average sum-rate. For this reason, we adopt *successive upper bound minimization* (SSUM) method [77, 98], which includes two nested loops, to solve the original problem (4.1). Algorithm 5 shows the steps to optimize the sum-rate by adding new coherence channels iteratively. The outer loop starts from one random single channel realization and one random set of feasible variables $\{\mathbf{p}^{(0)}\geq 0,\bar{\mathcal{U}}^{(0)}\succeq\mathbf{0},\mathbf{\Omega}^{(0)}\succeq\mathbf{0}\}$. The optimized solution is updated by solving Algorithm 4 in the inner loop. Then, one more random channel realization is generated and added to the problem in the next outer loop. Precisely, in the n_{B}-th outer loop, there are n_{B} coherence channels involved and the optimized solution $\{\mathbf{p}^{(n)}\geq 0,\bar{\mathcal{U}}^{(n)}\succeq\mathbf{0},\mathbf{\Omega}^{(n)}\succeq\mathbf{0}\}$

is obtained from the inner loop by averaging over n_{B} channel realizations.

Due to the randomness of the new generated channel, the solutions obtained from the previous loop may achieve a better or worse sum-rate for the new channel and thus can result in an increase or decrease of the average sum-rate, respectively. Accordingly, the trend of the optimized results from the outer loop is not guaranteed to be non-decreasing or non-increasing. However, the optimized sum-rate becomes stabilized when $r \to \infty$. To explain it further, simulation results of convergence will be discussed in Subsection 4.3.1.

Algorithm 4. Joint design based ergodic sum-rate maximization for stochastic CSI in downlink transmission.

1: **Input:** Total number of channel realizations $n_{\mathrm{B}} > 0$, with n_{B} channel realizations $\{\mathbf{G}^{(1)}, \ldots, \mathbf{G}^{(n_{\mathrm{B}})}\}$ generated based on the same stochastic CSI;

2: **Initialization:** Set $r \coloneqq 0$ and a feasible point $\{\mathbf{p}^{(0)} \geq 0, \bar{\mathcal{U}}^{(0)} \succeq \mathbf{0}, \mathbf{\Omega}^{(0)} \succeq \mathbf{0}\}$;

3: **repeat**

4: Update $r \coloneqq r + 1$;

5: Find the optimal solution $\{\mathbf{p}^{(r)}, \bar{\mathcal{U}}^{(r)}, \mathbf{\Omega}^{(r)}\}$ by solving problem (4.5) based on the given point $\{\mathbf{p}^{(r-1)}, \bar{\mathcal{U}}^{(r-1)}, \mathbf{\Omega}^{(r-1)}\}$:

$$
\begin{aligned}
\underset{\mathbf{p}\geq 0,\, \bar{\mathcal{U}}\succeq \mathbf{0},\, \mathbf{\Omega}\succeq \mathbf{0}}{\text{maximize}} \quad & \frac{1}{n_{\mathrm{B}}} \sum_{n=1}^{n_{\mathrm{B}}} \sum_{k\in\mathcal{N}_{\mathrm{U}}} C_{\mathrm{ac},k}^{\mathrm{lb},(n)} \left(\bar{\mathcal{U}}^{(r)}, \mathbf{\Omega}^{(r)} | \bar{\mathcal{U}}^{(r-1)}, \mathbf{\Omega}^{(r-1)}\right) \\
\text{s.t.} \quad & \varphi_m^{\mathrm{ub}} \left(\bar{\mathcal{U}}^{(r)}, \mathbf{\Omega}^{(r)} | \bar{\mathcal{U}}^{(r-1)}, \mathbf{\Omega}^{(r-1)}\right) \leq \sum_{i\in\mathcal{S}_m} C_{\mathrm{fr},i}^{\mathrm{lb}} \left(\mathbf{p}^{(r)} | \mathbf{p}^{(r-1)}\right), \ \forall m \in \mathcal{N}_{\mathcal{S}}, \\
& \sum_{i\in\mathcal{N}_{\mathrm{R}}} p_i^{(r)} \leq P_{\mathrm{C}}^{\max}, \\
& \sum_{k\in\mathcal{N}_{\mathrm{U}}} \mathrm{tr}\left(\mathbf{\Gamma}_i^H \bar{\mathbf{U}}_k^{(r)} \mathbf{\Gamma}_i\right) + \mathrm{tr}\left(\mathbf{\Omega}_{ii}^{(r)}\right) \leq P_{\mathrm{R},i}^{\max}, \ \forall i \in \mathcal{N}_{\mathrm{R}};
\end{aligned}
$$

6: **until** convergence;

7: **Output:** $\{\mathbf{p}^{\star}, \bar{\mathcal{U}}^{\star}, \mathbf{\Omega}^{\star}\} = \{\mathbf{p}^{(r)}, \bar{\mathcal{U}}^{(r)}, \mathbf{\Omega}^{(r)}\}$.

Algorithm 5. Double-loop joint design based ergodic sum-rate maximization for stochastic CSI in downlink transmission.

1: **Initialization (outer loop):** Set $l \coloneqq 0$, $R^{(0)} \coloneqq 0$, maximum number of iterations $L_{\max}$, and a feasible point $\{\mathbf{p}^{(0)} \geq 0, \bar{\mathcal{U}}^{(0)} \succeq \mathbf{0}, \mathbf{\Omega}^{(0)} \succeq \mathbf{0}\}$;

2: **repeat**

3: Update $l \coloneqq l + 1$;

4: Generate a random channel matrix $\mathbf{G}^{(l)}$ based the given stochastic CSI;

5: **Initialization (inner loop):** Set $r \coloneqq 0$ and $\{\mathbf{p}^{(l,0)}, \bar{\mathcal{U}}^{(l,0)}, \mathbf{\Omega}^{(l,0)}\} \coloneqq \{\mathbf{p}^{(l-1)}, \bar{\mathcal{U}}^{(l-1)}, \mathbf{\Omega}^{(l-1)}\}$;

6: **repeat**

7: Update $r \coloneqq r + 1$;

8: Find the optimal solution $\{\mathbf{p}^{(l,r)}, \bar{\mathcal{U}}^{(l,r)}, \mathbf{\Omega}^{(l,r)}\}$ by solving problem (4.5) based on the given point $\{\mathbf{p}^{(l,r-1)}, \bar{\mathcal{U}}^{(l,r-1)}, \mathbf{\Omega}^{(l,r-1)}\}$:

$$\begin{aligned}
\underset{\mathbf{p}\geq 0,\, \bar{\mathcal{U}}\succeq\mathbf{0},\, \mathbf{\Omega}\succeq\mathbf{0}}{\text{maximize}} \quad & \frac{1}{l}\sum_{n=1}^{l}\sum_{k\in\mathcal{N}_{\mathrm{U}}} C_{\mathrm{ac},k}^{\mathrm{lb},(n)}\left(\bar{\mathcal{U}}^{(l,r)}, \mathbf{\Omega}^{(l,r)} | \bar{\mathcal{U}}^{(l,r-1)}, \mathbf{\Omega}^{(l,r-1)}\right) \\
\text{s.t.} \quad & \varphi_m^{\mathrm{ub}}\left(\bar{\mathcal{U}}^{(l,r)}, \mathbf{\Omega}^{(l,r)} | \bar{\mathcal{U}}^{(l,r-1)}, \mathbf{\Omega}^{(l,r-1)}\right) \leq \sum_{i\in\mathcal{S}_m} C_{\mathrm{fr},i}^{\mathrm{lb}}\left(\mathbf{p}^{(l,r)} | \mathbf{p}^{(l,r-1)}\right), \\
& \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \forall m \in \mathcal{N}_{\mathcal{S}}, \\
& \sum_{i\in\mathcal{N}_{\mathrm{R}}} p_i^{(l,r)} \leq P_{\mathrm{C}}^{\max}, \\
& \sum_{k\in\mathcal{N}_{\mathrm{U}}} \mathrm{tr}\left(\mathbf{\Gamma}_i^H \bar{\mathbf{U}}_k^{(l,r)} \mathbf{\Gamma}_i\right) + \mathrm{tr}\left(\mathbf{\Omega}_{ii}^{(l,r)}\right) \leq P_{\mathrm{R},i}^{\max},\ \forall i \in \mathcal{N}_{\mathrm{R}};
\end{aligned}$$

9: **until** convergence;

10: Update $\{\mathbf{p}^{(l)}, \bar{\mathcal{U}}^{(l)}, \mathbf{\Omega}^{(l)}\} \coloneqq \{\mathbf{p}^{(l,r)}, \bar{\mathcal{U}}^{(l,r)}, \mathbf{\Omega}^{(l,r)}\}$, $R^{(l)} \coloneqq \frac{1}{l}\sum_{n=1}^{l}\sum_{k\in\mathcal{N}_{\mathrm{U}}} C_{\mathrm{ac},k}^{\mathrm{lb},(n)}\left(\bar{\mathcal{U}}^{(l,r)}, \mathbf{\Omega}^{(l,r)}\right)$;

11: **until** $|R^{(l)} - R^{(l-1)}| < 10^{-3}$ or $l \geq L_{\max}$;

12: **Output:** $\{\mathbf{p}^{\star}, \bar{\mathcal{U}}^{\star}, \mathbf{\Omega}^{\star}\} = \{\mathbf{p}^{(l)}, \bar{\mathcal{U}}^{(l)}, \mathbf{\Omega}^{(l)}\}$.

4.2.1.c Separate Design of Fronthaul and Access Links

Similar to Subsection 3.3.1.b, we propose a separate design in order to reduce computational complexity. According to our system model described in Section 4.2, the UE side has no prior acknowledgment of the access channels and thus we aimed to maximize the ergodic sum-rate for the access link in the previous section. However, the fronthaul link is assumed to be deterministic and thus perfect instantaneous CSI can be obtained. The separate optimization of the fronthaul link is similar to the one used in Algorithm 2. The details of the proposed separate designs are provided in Algorithm 6.

4.2.2 Total Transmit Power Minimization

Similar to Subsection 3.3.2, we aim to minimize the total transmit power while ensuring that the UEs can achieve the required individual-rate or sum-rate over multiple time blocks during the transmission. In order to achieve ergodic rates of UEs, similar to problem (3.36), it can be expressed as

- **Individual-rate:**

$$\begin{aligned} \underset{\mathbf{p}\succeq 0,\, \bar{\mathcal{U}}\succeq \mathbf{0},\, \mathbf{\Omega}\succeq \mathbf{0}}{\text{minimize}} \quad & \sum_{k\in\mathcal{N}_{\mathrm{U}}} \mathrm{tr}\left(\bar{\mathbf{U}}_k\right) + \mathrm{tr}\left(\mathbf{\Omega}\right) + \sum_{i\in\mathcal{N}_{\mathrm{R}}} p_i \\ \text{s.t.} \quad & R_k \geq \gamma_k,\ \forall k \in \mathcal{N}_{\mathrm{U}}, \qquad (4.9a) \\ & (3.18b) \text{ - } (3.18e), \end{aligned}$$

- **Sum-rate:**

$$\begin{aligned} \underset{\mathbf{p}\succeq 0,\, \bar{\mathcal{U}}\succeq \mathbf{0},\, \mathbf{\Omega}\succeq \mathbf{0}}{\text{minimize}} \quad & \sum_{k\in\mathcal{N}_{\mathrm{U}}} \mathrm{tr}\left(\bar{\mathbf{U}}_k\right) + \mathrm{tr}\left(\mathbf{\Omega}\right) + \sum_{i\in\mathcal{N}_{\mathrm{R}}} p_i \\ \text{s.t.} \quad & \sum_{k\in\mathcal{N}_{\mathrm{U}}} R_k \geq \gamma, \qquad (4.9b) \\ & (3.18b) \text{ - } (3.18e). \end{aligned}$$

The objective function and other constraints for ergodic rates are the same as those in (3.36). Obviously, problem (4.9) is difficult to solve due to non-convex constraints (4.9a), (4.9b), and (3.18b), as well as rank constraint (3.18e). By following the steps presented

Algorithm 6. Separate design based ergodic sum-rate maximization algorithm for stochastic CSI.

Input: Total number of channel realizations $n_{\mathrm{B}} > 0$, with n_{B} channel realizations $\{\mathbf{G}^{(1)}, \ldots, \mathbf{G}^{(n_{\mathrm{B}})}\}$ generated based on the same stochastic CSI;

Initialization (first loop): Set $l \coloneqq 0$ and a feasible point $\{\mathbf{p}^{(0)} \geq 0\}$;

repeat

Update $l \coloneqq l + 1$;

Find the optimal solution $\mathbf{p}^{(l)}$ and corresponding set $\{C_{\mathrm{fr},i}^{\mathrm{lb}}(\mathbf{p}), \forall i \in \mathcal{N}_{\mathrm{R}}\}$ based on the known $\mathbf{p}^{(l-1)}$ by solving one of the following problems:

- **Individual-rate** problem (3.29):
$$\begin{aligned} \underset{\mathbf{p}^{(l)} \geq 0}{\text{maximize}} \quad & \min_{i \in \mathcal{N}_{\mathrm{R}}} \; C_{\mathrm{fr},i}^{\mathrm{lb}}(\mathbf{p}^{(l)} | \mathbf{p}^{(l-1)}) \\ \text{s.t.} \quad & \sum_{i \in \mathcal{N}_{\mathrm{R}}} p_i^{(l)} \leq P_{\mathrm{C}}^{\max}; \end{aligned}$$

- **Sum-rate** problem (3.30):
$$\begin{aligned} \underset{\mathbf{p}^{(l)} \geq 0}{\text{maximize}} \quad & \sum_{i \in \mathcal{N}_{\mathrm{R}}} C_{\mathrm{fr},i}^{\mathrm{lb}}(\mathbf{p}^{(l)} | \mathbf{p}^{(l-1)}) \\ \text{s.t.} \quad & \sum_{i \in \mathcal{N}_{\mathrm{R}}} p_i^{(l)} \leq P_{\mathrm{C}}^{\max}; \end{aligned}$$

until convergence;

Update $\mathbf{p}^{\star} \coloneqq \mathbf{p}^{(l)}$ and $\{C^{\star}\} \coloneqq \{C_i^{\star} | C_i^{\star} = C_{\mathrm{fr},i}(\mathbf{p}^{\star}), i \in \mathcal{N}_{\mathrm{R}}\}$;

Initialization (second loop): Set $r \coloneqq 0$ and a feasible point $\{\bar{\mathcal{U}}^{(0)} \succeq \mathbf{0}, \mathbf{\Omega}^{(0)} \succeq \mathbf{0}\}$;

repeat

Update $r \coloneqq r + 1$;

Find the optimal solution $\{\bar{\mathcal{U}}^{(r)}, \mathbf{\Omega}^{(r)}\}$ based on the known $\{\mathbf{p}^{\star}, \bar{\mathcal{U}}^{(r-1)}, \mathbf{\Omega}^{(r-1)}\}$ by solving the following problem:

$$\begin{aligned} \underset{\bar{\mathcal{U}}^{(r)} \succeq \mathbf{0},\, \mathbf{\Omega}^{(r)} \succeq \mathbf{0}}{\text{maximize}} \quad & \frac{1}{n_{\mathrm{B}}} \sum_{n=1}^{n_{\mathrm{B}}} \sum_{k \in \mathcal{N}_{\mathrm{U}}} C_{\mathrm{ac},k}^{\mathrm{lb},(n)} \left(\bar{\mathcal{U}}^{(r)}, \mathbf{\Omega}^{(r)} | \bar{\mathcal{U}}^{(r-1)}, \mathbf{\Omega}^{(r-1)} \right) \\ \text{s.t.} \quad & \varphi_m^{\mathrm{ub}} \left(\bar{\mathcal{U}}^{(r)}, \mathbf{\Omega}^{(r)} | \bar{\mathcal{U}}^{(r-1)}, \mathbf{\Omega}^{(r-1)} \right) \leq \sum_{i \in \mathcal{S}_m} C_i^{\star}, \; \forall m \in \mathcal{N}_{\mathcal{S}}, \\ & \sum_{k \in \mathcal{N}_{\mathrm{U}}} \mathrm{tr} \left(\mathbf{\Gamma}_i^H \bar{\mathbf{U}}_k^{(r)} \mathbf{\Gamma}_i \right) + \mathrm{tr} \left(\mathbf{\Omega}_{ii}^{(r)} \right) \leq P_{\mathrm{R},i}^{\max}, \; \forall i \in \mathcal{N}_{\mathrm{R}}, \end{aligned}$$

until convergence;

Output: $\{\mathbf{p}^{\star}, \bar{\mathcal{U}}^{\star}, \mathbf{\Omega}^{\star}\} \coloneqq \{\mathbf{p}^{(l)}, \bar{\mathcal{U}}^{(r)}, \mathbf{\Omega}^{(r)}\}$.

in Subsection (3.3.2), we employ a similar method presented in Algorithm 3 to solve the problem (4.9). The steps are detailed in Algorithm 7.

4.3 Numerical Results of Stochastic CSI

This section presents simulations of the proposed model described in Section 4.2. We study and compare the performance of the proposed Algorithms 3a-d and 7 are studied and compared, especially in terms of the feasibility of the problem, optimized objective value, convergence speed, and computational time. Unless otherwise stated, the simulation parameters used in this section are the same as those used in Section 3.4.

4.3.1 Convergence and Complexity Comparison

Since all algorithms simulate a limited number of access channel models n_{B} to approximate the ergodic sum-rate, it is vital to determine the appropriate number of channels to use in the simulations. A small n_{B} can lead to a loss of generality, while a large n_{B} can result in increased problem complexity. Using an inappropriate number of n_{B} in the simulations can be problematic. Therefore, we first start by investigating the impact of n_{B}.

Figure 4.1 compares the single-loop joint design presented in Algorithm 4 and double-loop joint design proposed in Algorithm 5 in terms of the average sum-rate versus the number of channels involved n_{B}. The solver utilized to optimize the problems is SDPT3. It can be seen that the achievable sum-rate has a decreasing trend when more channel realizations are averaged for both algorithms. This trend is expected because one set of optimized variables, that can achieve a high rate for one channel realization, may perform badly for the others. It is worth noting that neither Algorithm 4 nor the outer-loop of Algorithm 5 is a convex problem with respect to n_{B}. Therefore, there is no guarantee that the ergodic sum-rate converges to a critical point but will become stabilized if n_{B} is large enough.

Algorithm 7. Joint design based transmit power minimization for stochastic CSI.

1: **Input:** Essential system parameters, including fronthaul channels $\mathbf{H}$, access channels $\mathbf{G}$, and

- **Average Initialization:** a random feasible point $\{\mathbf{p}^{(0)} \geq 0, \bar{\mathcal{U}}^{(0)} \succeq \mathbf{0}, \mathbf{\Omega}^{(0)} \succeq \mathbf{0}\}$;
- **Sum-rate Based Initialization:** an optimized feasible point $\{\mathbf{p}^{(0)} \geq 0, \bar{\mathcal{U}}^{(0)} \succeq \mathbf{0}, \mathbf{\Omega}^{(0)} \succeq \mathbf{0}\}$ obtained from the sum-rate maximization problem solved in Algorithm 4;

2: **Initialization:** Set $r \coloneqq 0$;

3: **repeat**

4: Update $r \coloneqq r + 1$;

5: Find the optimal solution $\{\mathbf{p}^{(r)}, \bar{\mathcal{U}}^{(r)}, \mathbf{\Omega}^{(r)}\}$ of problem (4.9) based on the given point $\{\mathbf{p}^{(r-1)}, \bar{\mathcal{U}}^{(r-1)}, \mathbf{\Omega}^{(r-1)}\}$ by solving one of the following problems:

- **Individual-rate:**

$$\underset{\mathbf{p}^{(r)} \geq 0,\, \bar{\mathcal{U}}^{(r)} \succeq \mathbf{0},\, \mathbf{\Omega}^{(r)} \succeq \mathbf{0}}{\text{minimize}} \quad \sum_{k \in \mathcal{N}_{\mathrm{U}}} \mathrm{tr}\left(\bar{\mathbf{U}}_k^{(r)}\right) + \mathrm{tr}\left(\mathbf{\Omega}^{(r)}\right) + \sum_{i \in \mathcal{N}_{\mathrm{R}}} p_i^{(r)}$$

$$\text{s.t.} \quad \frac{1}{n_{\mathrm{B}}} \sum_{n=1}^{n_{\mathrm{B}}} C_{\mathrm{ac},k}^{\mathrm{lb},(n)} \left(\bar{\mathcal{U}}^{(r)}, \mathbf{\Omega}^{(r)} | \bar{\mathcal{U}}^{(r-1)}, \mathbf{\Omega}^{(r-1)}\right) \geq \gamma_k, \ \forall k \in \mathcal{N}_{\mathrm{U}},$$

(3.24), (3.18c), (3.18d);

- **Sum-rate:**

$$\underset{\mathbf{p}^{(r)} \geq 0,\, \bar{\mathcal{U}}^{(r)} \succeq \mathbf{0},\, \mathbf{\Omega}^{(r)} \succeq \mathbf{0}}{\text{minimize}} \quad \sum_{k \in \mathcal{N}_{\mathrm{U}}} \mathrm{tr}\left(\bar{\mathbf{U}}_k^{(r)}\right) + \mathrm{tr}\left(\mathbf{\Omega}^{(r)}\right) + \sum_{i \in \mathcal{N}_{\mathrm{R}}} p_i^{(r)}$$

$$\text{s.t.} \quad \frac{1}{n_{\mathrm{B}}} \sum_{n=1}^{n_{\mathrm{B}}} \sum_{k \in \mathcal{N}_{\mathrm{U}}} C_{\mathrm{ac},k}^{\mathrm{lb},(n)} \left(\bar{\mathcal{U}}^{(r)}, \mathbf{\Omega}^{(r)} | \bar{\mathcal{U}}^{(r-1)}, \mathbf{\Omega}^{(r-1)}\right) \geq \gamma,$$

(3.24), (3.18c), (3.18d);

6: **until** convergence;

7: **Output:** $\{\mathbf{p}^\star, \bar{\mathcal{U}}^\star, \mathbf{\Omega}^\star\} \coloneqq \{\mathbf{p}^{(r)}, \bar{\mathcal{U}}^{(r)}, \mathbf{\Omega}^{(r)}\}$.

Figure 4.2 shows the average total running time for the single-loop and double-loop joint design with the different number of channels n_{B}. It is worth noting that for a given $n_{\mathrm{B}} \geq 4$, the double-loop joint design requires only a very short time to obtain the converged optimal solution. According to Figure 4.1, it seems that the double-loop joint design is not able to achieve an optimal sum-rate as high as the single-loop joint design. One possibility for this observation is that the double-loop design might quickly converge to a local minimum based on the first three involved channels. Although the double-loop design has a very short computational time for each specific number of channels n_{B}, the accumulated time becomes longer than the single-loop design when $n_{\mathrm{B}} \geq 3$. Therefore, the single-loop joint design actually outperforms the double-loop design in terms of a high achievable sum-rate and small computational time.

Since the fluctuation of both algorithms is very little when $n_{\mathrm{B}} \geq 10$, a balance between performance (such as improvement on achievable maximized sum-rate) and consumed resources (such as computational memory and time) could be achieved when n_{B} is large.

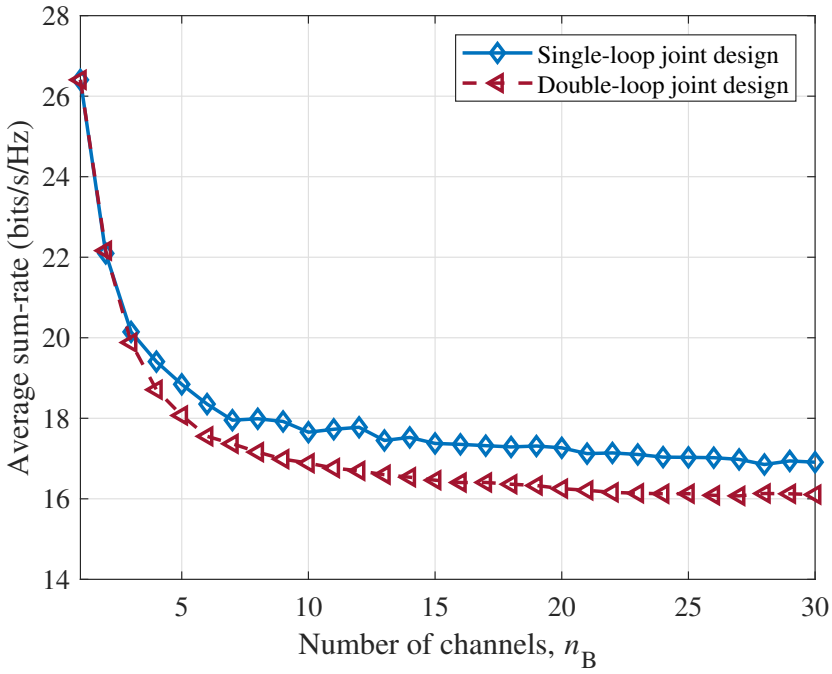


Figure 4.1: Average sum-rate versus the number of channels n_{B} for the single-loop joint design presented in Algorithm 4 and double-loop joint design proposed in Algorithm 5.

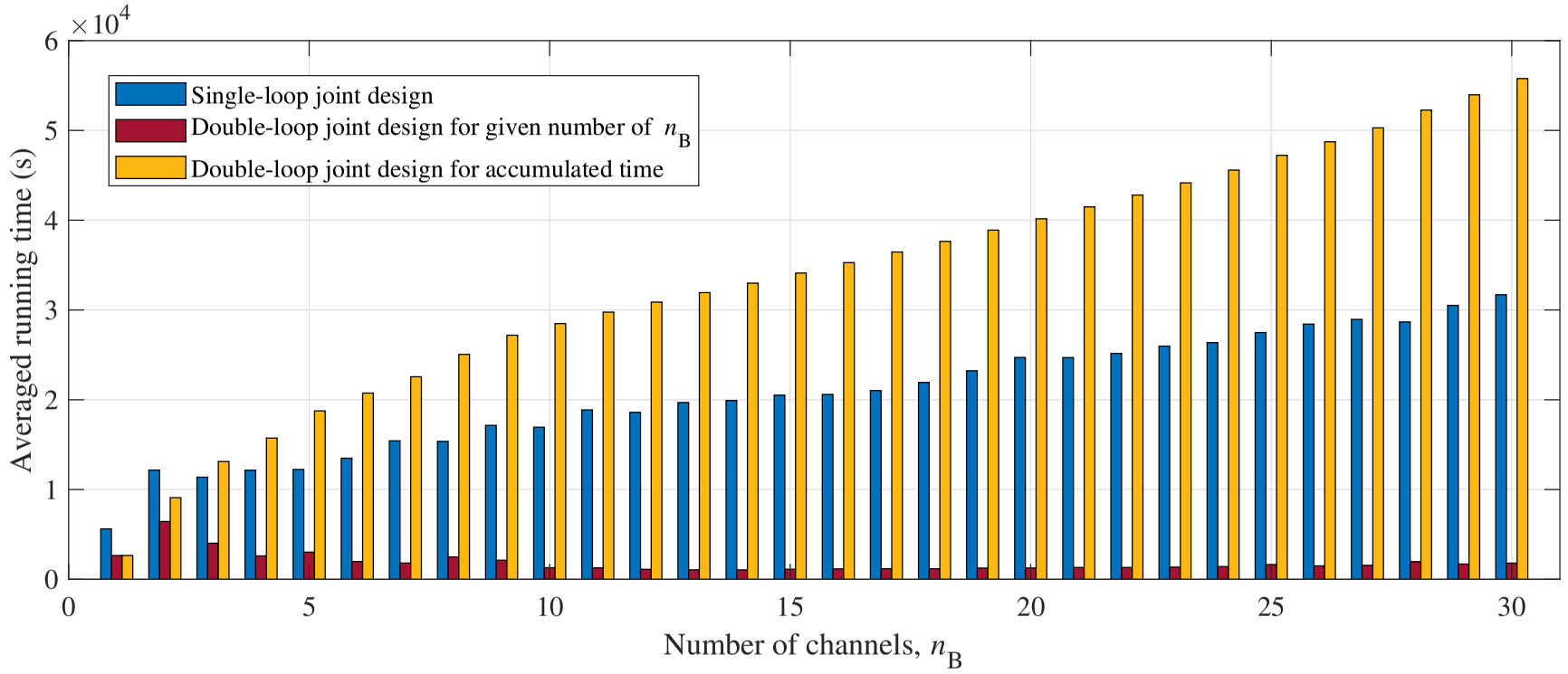


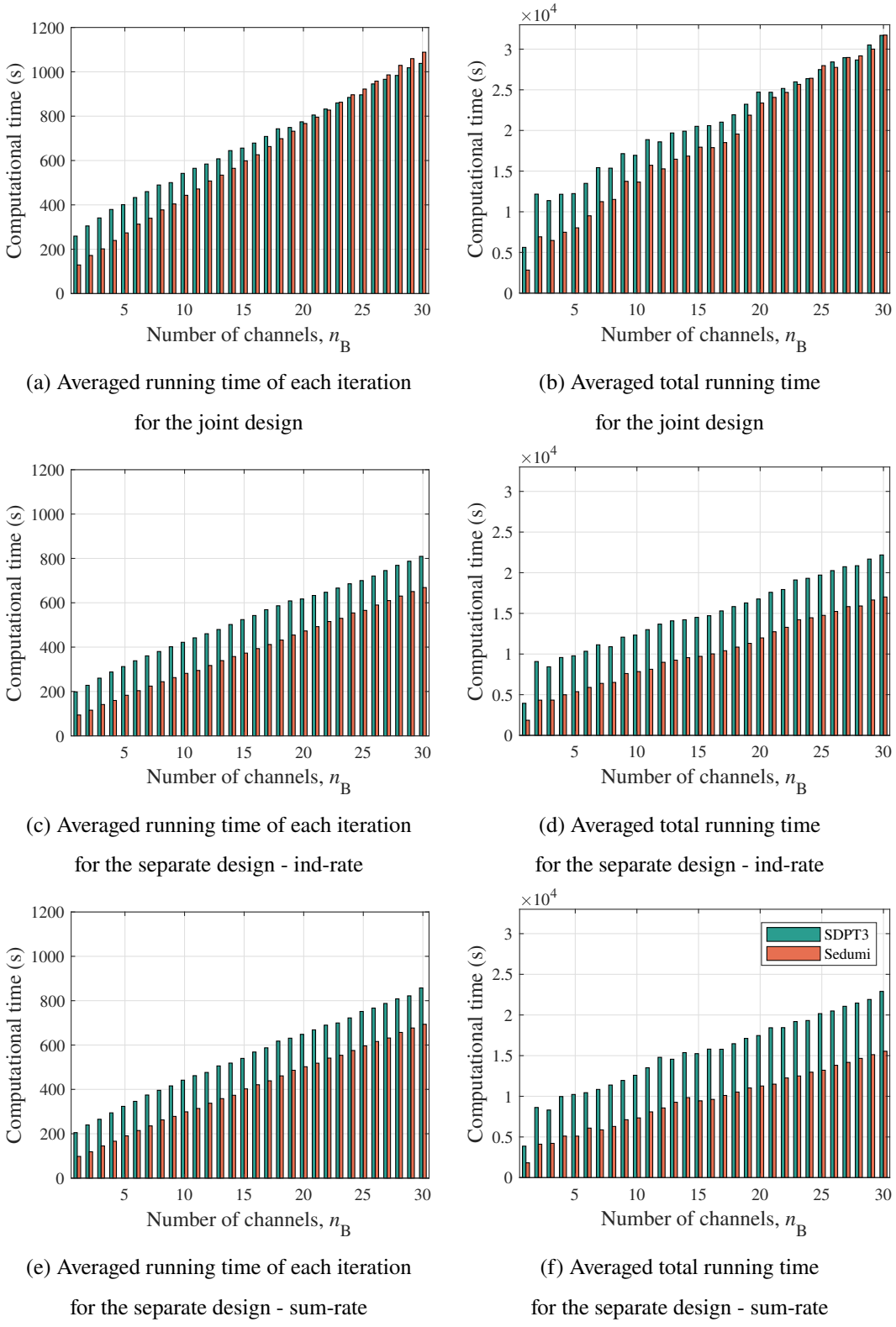
Figure 4.2: Actual averaged total computational time with different number of channels n_B for the single-loop joint design presented in Algorithm 4 and double-loop joint design in Algorithm 5. The computational time is based on MATLAB running on 2 CPU cores (2 x AMD EPYC 7702) with 256GB RAM.

Table 4.1: Feasibility of the proposed joint and separate designs, as well as two benchmarks using different solvers SDPT3 and SeDuMi.

Algorithm \ Solver	SDPT3	SeDuMi
Single-loop joint design	100/100	18/20
Separate design - Ind-rate	100/100	18/20
Separate design - Sum-rate	94/100	15/20

In order to save some time without loss of generality, we choose $n = 20$ channel realizations for the rest of the simulations and the double-loop joint design prescribed in Algorithm 5 will not be considered in the following Sections 4.3.2 and 4.3.3.

Figure 4.3 compares the running time among the joint design presented in Algorithm 4, the separate designs proposed in Algorithm 6, and two benchmarks with two solvers, while Table 4.1 lists their feasibility. Please note that the computational time shown in Figures 4.2 and 4.3 are calculated based on a single CPU core (2 x AMD EPYC 7702). In real scenarios, the calculation will be processed by the CU and thus faster computational speed can be achieved by utilizing multiple more powerful CPU cores. The convergence time of the proposed algorithms is shown in Figure 4.3 is just a reference to time consumption. The exact values are not of much interest to us but the difference among algorithms and the trend with the increased number of involved channels n_{B} is vital here, which can be seen as an approximation of reality. Additionally, the computational time for the first loop of separate designs in Algorithm 6 is very short and negligible compared to the second loop. Therefore, the running time in Subfigures 4.3c - 4.3f is counted only based on the second loop. Figure 4.3 illustrates that Sedumi typically requires less time to solve a single convex problem and converges in most cases, in contrast to SDPT3, which requires longer computation time but exhibits higher feasibility, as shown in Table 4.1. Both solvers exhibit distinct advantages: Sedumi demonstrates higher computational efficiency but less stability, whereas SDPT3 shows the opposite - greater stability but longer computation time.



(a) Averaged running time of each iteration for the joint design

(b) Averaged total running time for the joint design

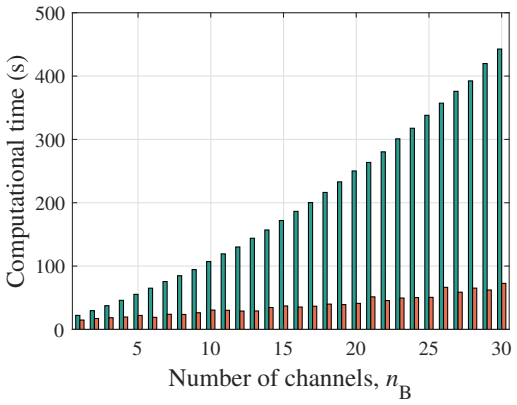
(c) Averaged running time of each iteration for the separate design - ind-rate

(d) Averaged total running time for the separate design - ind-rate

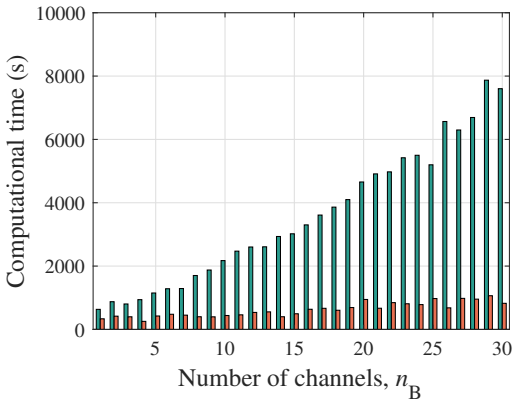
(e) Averaged running time of each iteration for the separate design - sum-rate

(f) Averaged total running time for the separate design - sum-rate

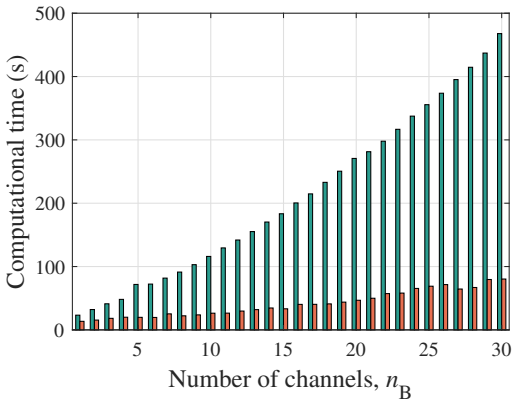
Figure 4.3: Actual running time against the number of iterations with the proposed algorithms and benchmarks using different solvers SDPT3 and SeDuMi. The computational time is based on MATLAB running on 2 CPU cores (2 x AMD EPYC 7702) with 256GB RAM.



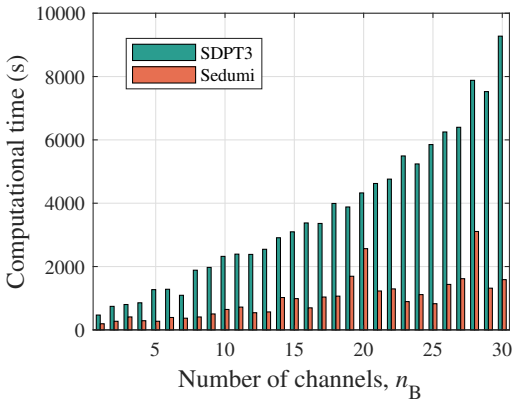
(g) Averaged running time of each iteration for Benchmark 1



(h) Averaged total running time for Benchmark 1



(i) Averaged running time of each iteration for Benchmark 2



(j) Averaged total running time for Benchmark 2

Figure 4.3: (Continued) Actual running time against the number of iterations with the proposed algorithms and benchmarks using different solvers SDPT3 and SeDuMi. The computational time is based on MATLAB running on 2 CPU cores (2 x AMD EPYC 7702) with 256GB RAM.

Figure 4.4 presents average convergence behaviors of the proposed algorithms and Benchmarks 1 and 2. We observe that the average achievable sum-rates of all the schemes are non-decreasing and the convergence is achieved after approximately 20 iterations. In this realization, the joint design scheme proposed in Algorithm 4 has shown better performance than all other schemes, while the separate design prescribed in Algorithm 6 with optimized individual-rate has performed the worst. This can be attributed to the fact that the separate designs of the fronthaul and access links, as implemented in Algorithm 6, leads to a slight decrease in the achievable sum-rate compared to the joint design. The separate design scheme converges more slowly compared to the other schemes. This is due to the fact that the optimization of the fronthaul links has to be completed before the optimization of the access links can begin, which requires more iterations and hence slower convergence. Although the independent point-to-point compression scheme can still achieve higher rates than the separate designs, its feasibility is quite low, as shown in Table 4.1. Benchmarks 1 and 2 show similar performance in the initial iterations, but Benchmark 2 converges faster while Benchmark 1 achieves better performance towards

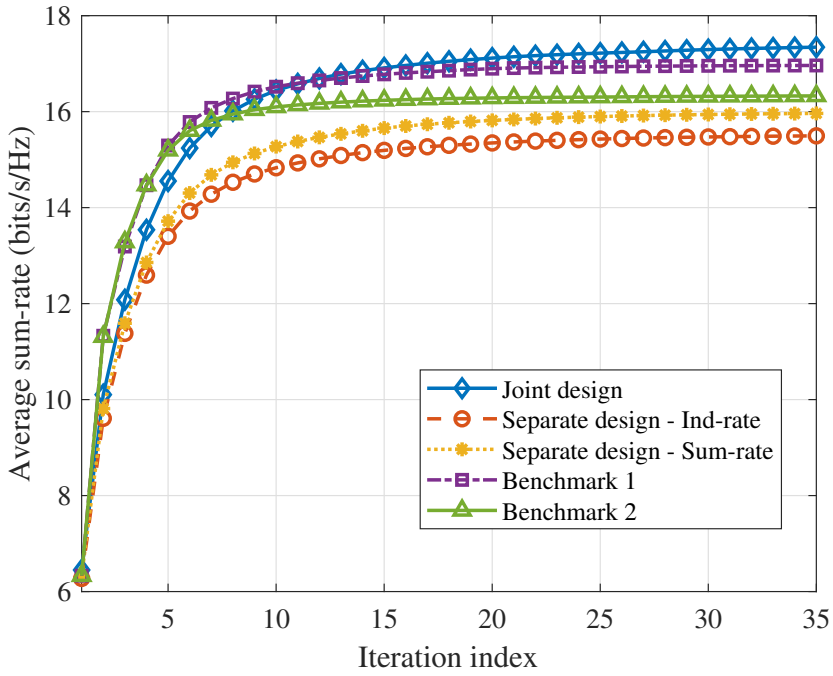


Figure 4.4: Convergence speed of the proposed algorithms versus the number of iterations with the proposed algorithms.

the end. This is because Benchmark 1 involves more elements of the quantization noise covariance matrix $\mathbf{\Omega}$ in optimization than Benchmark 2, which leads to better performance but a longer convergence time.

It is worth noting that the Benchmarks outperform the separate designs, unlike the numerical results presented in previous Chapter 3, where the achievable sum-rate of the separate designs is very close to that of the joint design. This indicates that although the compression scheme is not jointly optimized among all RRHs, the ergodic network sum-rate benefits a lot from the joint design of both fronthaul and access links. In the case of a single channel with the knowledge of perfect instantaneous CSI, the compression scheme takes on greater importance. Conversely, for multiple coherence channels, the joint design is more vital for improving the ergodic sum-rate.

4.3.2 Sum-rate Maximization Problem

Figure 4.5 depicts the average sum-rate versus the number of antennas M at the CU. It can be observed that the performance increases with increasing M. The proposed joint design in Algorithm 4 performs better than Benchmarks 1 and 2 while the proposed separate design performs the worst. Recall that in Benchmarks 1 and 2 the fronthaul and access links are jointly designed. Therefore, the joint design of both fronthaul and access links helps the fronthaul link achieve the best results, as expected.

In Figure 4.6, the average sum-rate versus the number of UEs N_{U} is shown. It is clear that the average sum-rate increases almost linearly with the number of UEs N_{U}. The average sum-rate against the power budget is evaluated in Figures 4.7 and 4.8. Figure 4.7 focuses on the impact of the available power at the CU with $P_{\mathrm{R}}^{\max} = 10\mathrm{W}$. For all the schemes, we can observe that the average sum-rate increases with increasing power budget at the CU but quickly saturates when $P_{\mathrm{C}}^{\max} \geq 0.5$ W. It is observable that the growth of the average sum-rate is not significantly changing, especially when $P_{\mathrm{C}}^{\max} \geq 0.4$ W. On the contrary, the average sum-rate keeps increasing with increasing $P_{\mathrm{R},i}^{\max}, \forall i \in \mathcal{N}_{\mathrm{R}}$, as shown

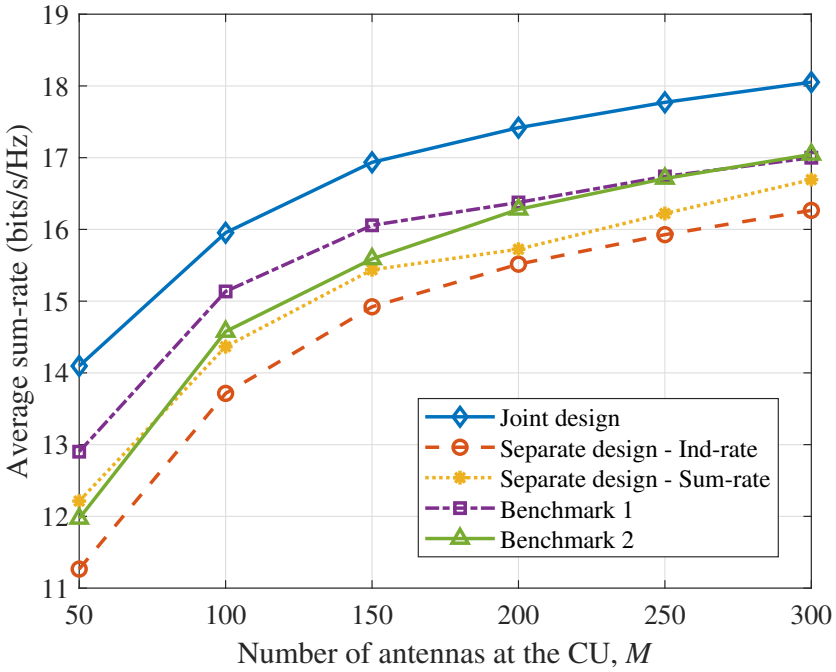


Figure 4.5: Average sum-rate versus the number of antennas M at the CU for the proposed algorithms.

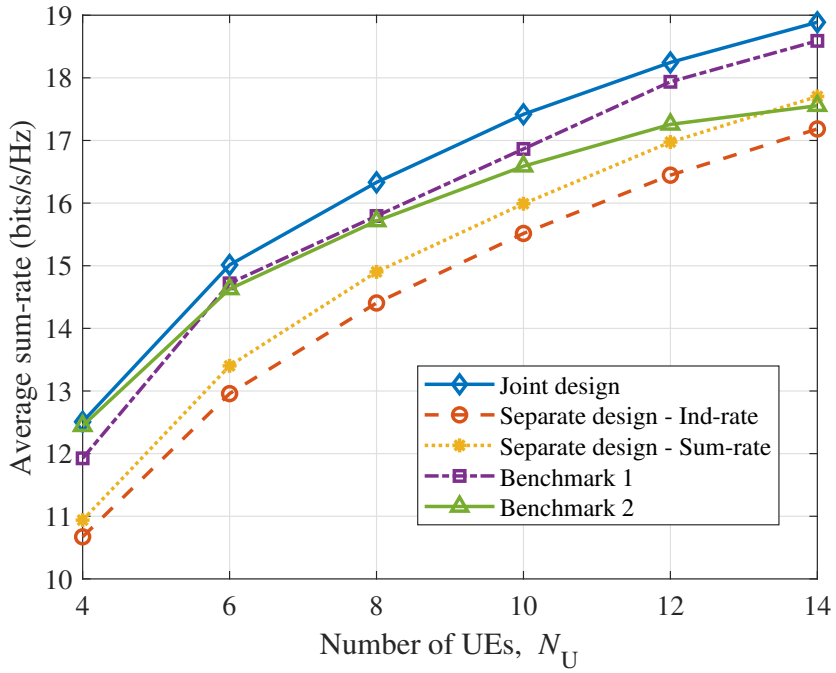


Figure 4.6: Average sum-rate versus the number of UEs N_{U} for the proposed algorithms.

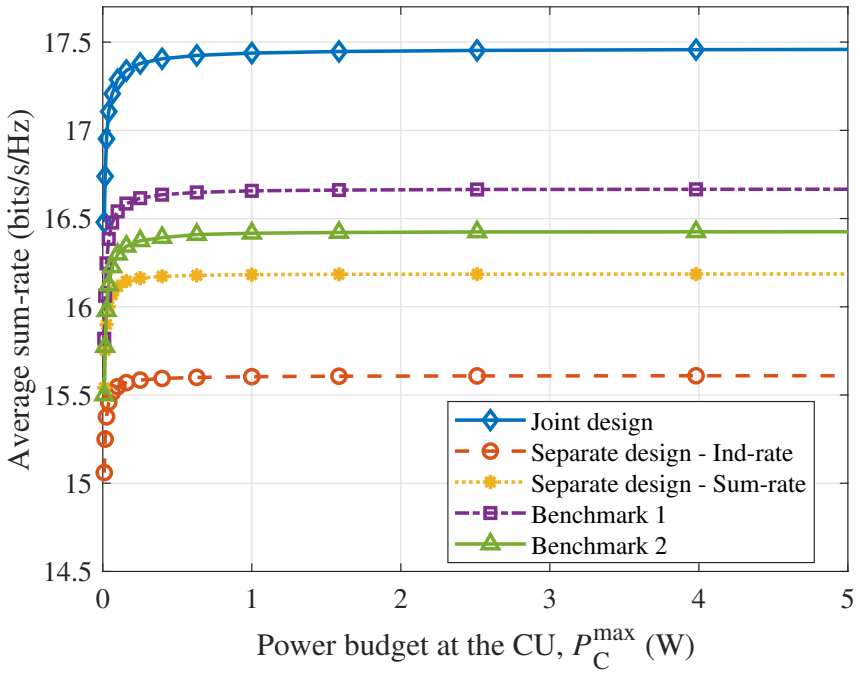


Figure 4.7: Average sum-rate versus the power budget $P_{\mathrm{C}}^{\max}$ at the CU for the proposed algorithms.

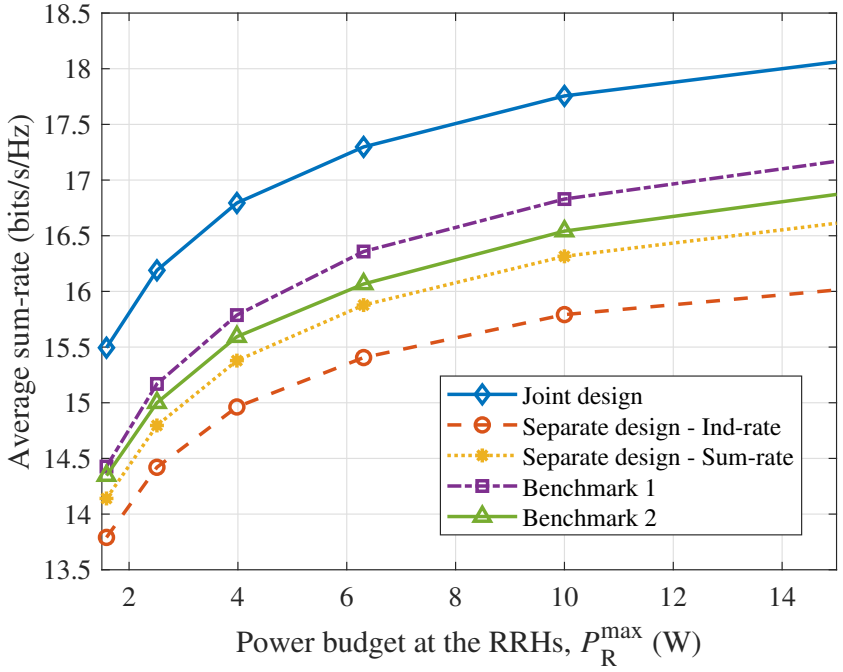


Figure 4.8: The average sum-rate versus the power budget $P_{\mathrm{R}}^{\max}$ at each RRH.

in Figure 4.8. All the schemes benefit from the increased power budget at RRHs and have further potential improvements. According to Figure 4.7, the bottleneck of the proposed design is caused by the access link. To further improve the sum-rate performance of the proposed system, future research may focus on optimizing the access link. Potential solutions include employing more RRHs, increasing the number of transmit antennas equipped at the RRHs, and allocating a higher power budget allocated for the access link.

4.3.3 Total Power Minimization Problem

Figure 4.9 shows the minimized total transmit power for different sum-rate requirements γ with the joint design and two benchmarks. The joint design is solved in Algorithm 7 with ensured sum-rate constraint, and initialized with the solution obtained from the sum-rate maximization problem solved in Algorithm 4. Similar to Figures 3.12 and 3.13 presented in Subsection 3.4.3, both methods are non-increasing and converge rapidly within 10 iterations. Larger sum-rate requirement γ results in higher power consumption. Although the difference is almost negligible, the joint design proposed in Algorithm 7 saves more energy than the two benchmarks, where each element of the signal $\bar{\mathbf{x}}_{\mathrm{R},i}$ is compressed and quantized independently.

4.4 Conclusion

Unlike the scenario of perfect instantaneous CSI, which we have thoroughly discussed in Chapter 3, we have considered the case where only stochastic CSI of the access link is available at the CU in this chapter. In the fronthaul link employing massive MIMO, the use of a large number of transmit antennas results in channel hardening, which increases the capacity of the fronthaul links in our designs. In contrast, the limited information available for the access channel poses challenges for designing the precoding matrices. We proposed joint and separate designs for optimizing both the fronthaul and access links,

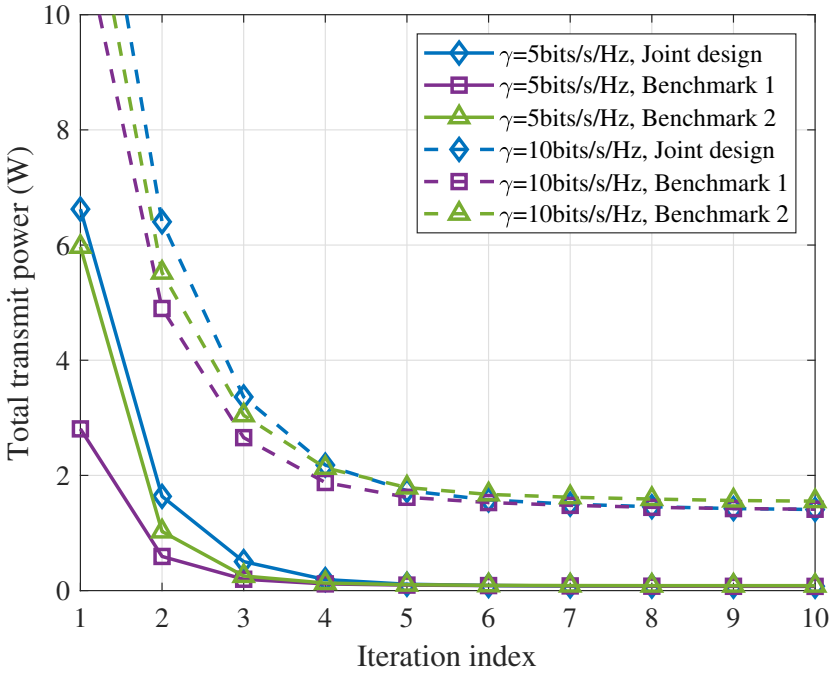


Figure 4.9: Convergence of the joint design in Algorithm 7 against the number of iterations with optimized initialization based on the sum-rate maximization problem solved in Algorithm 4.

with the goal of maximizing the ergodic sum-rate subject to a certain power budget, as well as minimizing the total transmit power while satisfying required rate constraints. As the resulting optimization problems were non-convex, they were difficult to solve directly. To address this issue, we transformed them into convex SDR problems and described iterative algorithms to solve the reformulated problems. As the optimized variables were designed to be employed for all coherence blocks, we also provided nested double-loop algorithms to examine how the number of coherence blocks considered in the optimization problems affects the results.

Numerical results have been provided to examine the proposed designs and algorithms in this chapter. The computational time, feasibility, and convergence speed of all proposed algorithms and benchmarks were presented and compared.

Chapter 5

Design of Wireless Uplink Transmission in Massive MIMO CRANs

5.1 Introduction

In this chapter, we consider the uplink transmission for the same CRAN architecture proposed in Chapters 3 and 4, focusing on the joint design of wireless fronthaul and access links with massive MIMO enabled for the fronthaul link. Specifically, we propose two optimization problems. The first problem aims to maximize the total network sum-rate, while the second problem aims to minimize the power consumption for data transmission with an ensured sum-rate. Assuming that the CSI for both links is perfectly known to the CU, we consider two conventional precoding methods for the fronthaul link, namely MR and ZF combining. At the RRHs, the CF strategy is adopted and the point-to-point independent compression scheme is considered. The compression covariance noise constrains the capacity in both links, which poses a challenging problem for the system design.

The system model is introduced in Section 5.2. In Section 5.2.1, we formulate the

optimization problems for sum-rate maximization and total power minimization. Similar to the problems introduced in Chapter 3, these problems are challenging to solve. To address this issue, we change the optimized variables and transform them into convex SDR problems by linearizing non-convex constraints. Iterative algorithms are proposed to obtain optimal results for both optimization problems. In Section 5.4, numerical results are presented and compared to the benchmarks. Finally, the chapter is concluded in Section 5.5.

5.2 System Model

We consider the uplink of a CRAN system, where N_{U} multi-antenna UEs wish to wirelessly send their messages to the CU in two hops through N_{R} multi-antenna RRHs, as depicted in Figure 5.1. The CU is equipped with a massive ULA of size M. For $k \in \mathcal{N}_{\mathrm{U}} = \{1, \ldots, N_{\mathrm{U}}\}$, the k-th user is equipped with K antennas, which brings on the total number of antennas at all UEs to $\sum_{\mathcal{N}_{\mathrm{U}}} K$. For $i \in \mathcal{N}_{\mathrm{R}} = \{1, \ldots, N_{\mathrm{R}}\}$, the i-th RRH utilizes N antennas for reception, with the total number of antennas at all RRHs for reception is $\sum_{\mathcal{N}_{\mathrm{R}}} N$, and one antenna for transmission. This setting is favorable because in the fronthaul links we have a massive MIMO channel and hence the assumption of using single antenna RRHs for transmission in the fronthaul is of practical interest. We assume that the access links (UEs-RRHs links) and fronthaul links (RRHs-CU links) are separated in the time domain to avoid interference between them.

Let $M_k \in \{1, \cdots, 2^{n_{\mathrm{L}} R_k}\}$ denote the message to be transmitted by UE k to the CU, where n_{L} is the block length and R_k is the information rate in *bits per channel use* (bpcu). Then, the k-th UE encodes its message using a Gaussian codebook into a data stream $\mathbf{s}_{\mathrm{U},k} \in \mathbb{C}^{n_{\mathrm{R}} \times 1}$ of unit variance and transmits it using a linear precoder $\mathbf{V}_k \in \mathbb{C}^{K \times n_{\mathrm{R}}}$. The resulting signal $\mathbf{x}_{\mathrm{U},k} = \mathbf{V}_k \mathbf{s}_{\mathrm{U},k}$, with power constraint $\mathbb{E}\{\|\mathbf{x}_{\mathrm{U},k}\|^2\} \leq P_{\mathrm{U},k}^{\max}$, is then transmitted through the channel $\mathbf{G}_{k,i}^T \in \mathbb{C}^{N \times K}$ to the i-th RRH. A Rayleigh flat-fading channel model is used for $\mathbf{G}_{k,i}$, i.e., $\mathbf{G}_{k,i} = \sqrt{\alpha_{k,i}} \tilde{\mathbf{G}}_{k,i}$, where $\alpha_{k,i}$ is the large-scale fading

coefficient of the channel between UE k and RRH i, and $\tilde{\mathbf{G}}_{k,i}$ is a matrix of independent Rayleigh coefficients with entries modeled as $\mathcal{CN}(0,1)$. The received signal at the i-th RRH from all UEs is then given by

$$\mathbf{y}_{\mathrm{R},i} = \sum_{k \in \mathcal{N}_{\mathrm{U}}} \mathbf{G}_{k,i}^{T} \mathbf{x}_{\mathrm{U},k} + \mathbf{n}_{\mathrm{R},i}, \tag{5.1}$$

where $\mathbf{n}_{\mathrm{R},i} \sim \mathcal{CN}(\mathbf{0}, \sigma_{\mathrm{R}}^2 \mathbf{I})$ is the additive noise at the i-th RRH. Next, the received signal $\mathbf{y}_{\mathrm{R},i}$ is transferred to the CU via a wireless fronthaul link.

Similar to [41], we adopt the CF relaying strategy at the RRHs. So, before forwarding to the CU via the fronthaul links, the signal $\mathbf{y}_{\mathrm{R},i}$ at the i-th RRH is quantized and compressed. We adopt the Gaussian quantization test channel to model the quantization process [39], and hence the resulting quantized signal, $\hat{\mathbf{y}}_{\mathrm{R},i}$, can be written as

$$\hat{\mathbf{y}}_{\mathrm{R},i} = \mathbf{y}_{\mathrm{R},i} + \boldsymbol{\omega}_i, \tag{5.2}$$

where $\boldsymbol{\omega}_i \sim \mathcal{CN}(\mathbf{0}, \boldsymbol{\Omega}_i)$ is the quantization noise, which is independent of $\mathbf{y}_{\mathrm{R},i}$. Next, $\hat{\mathbf{y}}_{\mathrm{R},i}$ is compressed to generate the compression index $V_i \in \{1, \ldots, 2^{n_{\mathrm{L}} C_i}\}$, where C_i is the rate. Here, we assume point-to-point compression. The index V_i is then mapped into a complex scalar symbol $x_{\mathrm{R},i}$ to be further transmitted by the i-th RRH, using a single antenna, to the CU.

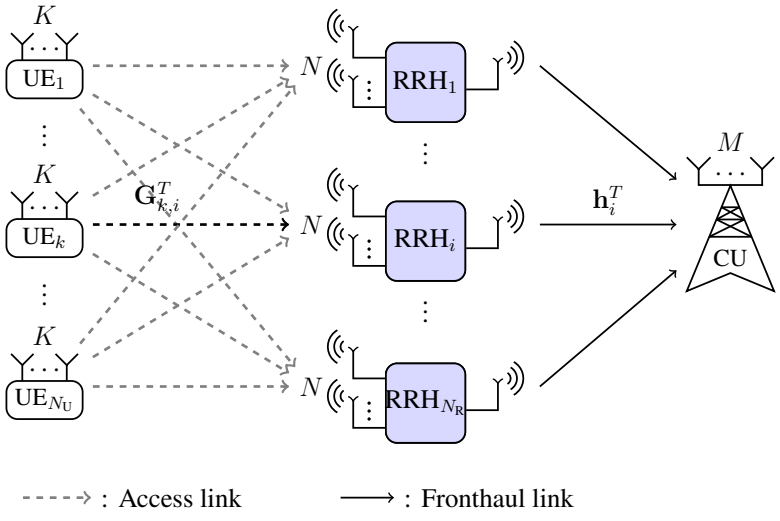


Figure 5.1: System model of downlink transmission in the proposed massive MIMO CRAN architecture.

The received signal at the CU can be written as

$$\mathbf{y}_{\mathrm{C}} = \sum_{i \in \mathcal{N}_{\mathrm{R}}} \mathbf{h}_i^T x_{\mathrm{R},i} + \mathbf{n}_{\mathrm{C}}, \tag{5.3}$$

where $\mathbf{h}_i^T \in \mathbb{C}^{M \times 1}$ is the channel between the i-th RRH and the CU, $\mathbf{n}_{\mathrm{C}} \sim \mathcal{CN}(0, \sigma_{\mathrm{C}}^2 \mathbf{I})$ is the additive noise at the CU and $\mathbb{E}\{|x_{\mathrm{R},i}|^2\} = p_{\mathrm{R},i} \leq P_{\mathrm{R},i}^{\max}$ is the transmit power constraint per RRH. Considering a scenario with LoS propagation, the channel vector $\mathbf{h}_i^T$ can be given by [43, 84]

$$\mathbf{h}_i = \sqrt{\beta_i^{\mathrm{LoS}}}\, \bar{\mathbf{h}}_i, \tag{5.4}$$

where β_i^{LoS} is the large scale fading from the i-th RRH and

$$\bar{\mathbf{h}}_i = \begin{bmatrix} 1 & e^{j2\pi \frac{\delta}{\lambda} \cos(\phi_i)} & \dots & e^{j2\pi(M-1) \frac{\delta}{\lambda} \cos(\phi_i)} \end{bmatrix}. \tag{5.5}$$

Here, δ is the antenna spacing, λ is the carrier wavelength and ϕ_i is the angle of arrival from the i-th RRH distributed uniformly in $[0, 2\pi)$. At the CU, an estimate of the received symbol $\hat{x}_{\mathrm{R},i}$ is obtained from (5.3) using beamformer $\mathbf{w}_i \in \mathbb{C}^M$, which is designed based on the perfectly known channel information available at the CU.

We further assume that the user codebooks and the RRH codebooks are available at the CU. Consequently, a lossless decoding from the detected symbol $\hat{y}_{C,l}$ to an index V_i can be realized.

5.2.1 Joint Design

Under the model given in the previous section, we now discuss our proposed design, where we use an optimization framework. The main question is how to jointly optimize the access and fronthaul links to maximize the system sum-rate. To that end, before we formulate the optimization problem, we first need to specify the fronthaul receive combining scheme and the closed-form expression of the achievable rate per user. Here, we assume that the perfect CSI for the access and fronthaul links is available at the CU.

At the CU, the first step is to detect the signals $x_{\mathrm{R},i}$, for all $i \in \mathcal{N}_{\mathrm{R}}$, sent by the RRHs from the received signal $\mathbf{y}_{\mathrm{C}}$ in (5.3). In particular, to detect the signal sent by the i-th RRH, $x_{\mathrm{R},i}$, a combining vector $\mathbf{w}_i$ is applied to the received signal at the CU and the detected signal is given by

$$\hat{x}_{\mathrm{R},i} = \mathbf{w}_i \mathbf{h}_i^T x_{\mathrm{R},i} + \sum_{j \in \mathcal{N}_{\mathrm{R}} \setminus \{i\}} \mathbf{w}_i \mathbf{h}_j^T x_{\mathrm{R},j} + \mathbf{w}_i \mathbf{n}_{\mathrm{C}}. \tag{5.6}$$

Thus, we can compute the SINR of the i-th RRH as

$$\mathrm{SINR}_i = \frac{|\mathbf{w}_i \mathbf{h}_i^T|^2 p_{\mathrm{R},i}}{\sigma_{\mathrm{C}}^2 \|\mathbf{w}_i\|^2 + \sum\limits_{j \in \mathcal{N}_{\mathrm{R}} \setminus \{i\}} |\mathbf{w}_i \mathbf{h}_j^T|^2 p_{\mathrm{R},j}}. \tag{5.7}$$

Therefore, the rate C_i between RRH i and the CU is achievable if the condition

$$\begin{aligned} C_i \leq C_{\mathrm{fr},i}(\{p_{\mathrm{R},i}\}) &\triangleq I(x_{\mathrm{R},i}; \hat{x}_{\mathrm{R},i}) \\ &= \log_2 \left(1 + \frac{|\mathbf{w}_i \mathbf{h}_i^T|^2 p_{\mathrm{R},i}}{\sigma_{\mathrm{C}}^2 \|\mathbf{w}_i\|^2 + \sum\limits_{j \in \mathcal{N}_{\mathrm{R}} \setminus \{i\}} |\mathbf{w}_i \mathbf{h}_j^T|^2 p_{\mathrm{R},j}} \right) \end{aligned} \tag{5.8}$$

is satisfied.

Later in Section 5.4, we will provide a showcase for two popular and simple combining techniques which are MR and ZF processing. The combining matrix, $\mathbf{W} \triangleq \left[\mathbf{w}_1^T, \ldots, \mathbf{w}_{N_{\mathrm{R}}}^T\right]^T$, is given by

$$\mathbf{W} = \begin{cases} \bar{\mathbf{H}}^*, & \text{for MR,} \\ (\bar{\mathbf{H}}^* \bar{\mathbf{H}}^T)^{-1} \bar{\mathbf{H}}^*, & \text{for ZF,} \end{cases} \tag{5.9}$$

with $\bar{\mathbf{H}} \triangleq \left[\bar{\mathbf{h}}_1^T, \ldots, \bar{\mathbf{h}}_{N_{\mathrm{R}}}^T\right]^T$. Plugging (5.9) into (5.7), the fronthaul SINR from the i-th RRH under the MR and ZF schemes can be computed, respectively, as

$$\mathrm{SINR}_i^{\mathrm{MR}} = \frac{\beta_i^{\mathrm{LoS}} \|\bar{\mathbf{h}}_i\|^4 p_{\mathrm{R},i}}{\sigma_{\mathrm{C}}^2 \|\bar{\mathbf{h}}_i\|^2 + \sum_{j \in \mathcal{N}_{\mathrm{R}} \setminus \{i\}} \beta_j |\bar{\mathbf{h}}_i \bar{\mathbf{h}}_j^T|^2 p_{\mathrm{R},j}}, \tag{5.10}$$

and

$$\mathrm{SINR}_i^{\mathrm{ZF}} = \frac{\beta_i^{\mathrm{LoS}} p_{\mathrm{R},i}}{[(\bar{\mathbf{H}}^* \bar{\mathbf{H}}^T)^{-1}]_{i,i}}. \tag{5.11}$$

Based on the decoded signal $x_{\mathrm{R},i}$, we get the corresponding compression index T_i. Given T_i, the CU can determine the quantized signal $\hat{\mathbf{y}}_{\mathrm{R},i}$. For independent point-to-point compression at each RRH, the condition for successful decompression represents a special case of (5.12), which is given as

$$\begin{aligned}\varphi_i(\{\mathbf{V}_k\},\mathbf{\Omega}_{ii}) &\triangleq I(\mathbf{y}_{\mathrm{R},i};\hat{\mathbf{y}}_{\mathrm{R},i})\\ &= \log\det\left(\sum_{k\in\mathcal{N}_{\mathrm{U}}}\mathbf{G}_{k,i}^T\mathbf{V}_k\mathbf{V}_k^H\mathbf{G}_{k,i}^* + \sigma_{\mathrm{C}}^2\mathbf{I} + \mathbf{\Omega}_{ii}\right) - \log\det\left(\mathbf{\Omega}_{ii}\right)\\ &\leq C_i, \quad \forall i \in \mathcal{N}_{\mathrm{R}}.\end{aligned} \tag{5.12}$$

5.2.2 Achievable Rate

We now determine the achievable rate for each UE. First, define $\hat{\mathbf{y}}_{\mathrm{R}} \triangleq \left[\hat{\mathbf{y}}_{\mathrm{R},1}^T,\ldots,\hat{\mathbf{y}}_{\mathrm{R},N_{\mathrm{R}}}^T\right]^T$, $\mathbf{G}_k \triangleq [\mathbf{G}_{k,1},\ldots,\mathbf{G}_{k,N_{\mathrm{R}}}]$, $\mathbf{n}_{\mathrm{R}} \triangleq \left[\mathbf{n}_{\mathrm{R},1}^T,\ldots,\mathbf{n}_{\mathrm{R},N_{\mathrm{R}}}^T\right]^T$ and $\boldsymbol{\omega} \triangleq \left[\boldsymbol{\omega}_1^T,\ldots,\boldsymbol{\omega}_{N_{\mathrm{R}}}^T\right]^T \sim \mathcal{CN}(0,\mathbf{\Omega})$ with $\mathbf{\Omega} = \mathrm{diag}(\{\mathbf{\Omega}_i\}_{i\in\mathcal{N}_{\mathrm{R}}})$. Accordingly, we can write the received signal after quantization from all RRHs by plugging (5.1) into (5.2) such that we obtain

$$\begin{aligned}\hat{\mathbf{y}}_{\mathrm{R}} &= \sum_{k\in\mathcal{N}_{\mathrm{U}}}\mathbf{G}_k^T\mathbf{V}_k\mathbf{s}_{\mathrm{U},k} + \mathbf{n}_{\mathrm{R}} + \boldsymbol{\omega}\\ &= \mathbf{G}_k^T\mathbf{V}_k\mathbf{s}_{\mathrm{U},k} + \sum_{j\in\mathcal{N}_{\mathrm{U}}\setminus\{k\}}\mathbf{G}_j^T\mathbf{V}_j\mathbf{s}_j + \mathbf{n}_{\mathrm{R}} + \boldsymbol{\omega}.\end{aligned} \tag{5.13}$$

By treating the interference as noise, the rate achievable by the k-th UE is given by

$$\begin{aligned}R_k \leq& I(\mathbf{s}_{\mathrm{U},k};\hat{\mathbf{y}}_{\mathrm{R}})\\ =& \log\det\left(\mathbf{I} + \frac{\mathbf{G}_k^T\mathbf{V}_k\mathbf{V}_k^H\mathbf{G}_k^*}{\sum_{j\in\mathcal{N}_{\mathrm{U}}\setminus\{k\}}\mathbf{G}_j^T\mathbf{V}_j\mathbf{V}_j^H\mathbf{G}_j^* + \sigma_{\mathrm{R}}^2\mathbf{I} + \mathbf{\Omega}}\right)\\ =& \log\det\left(\sum_{k\in\mathcal{N}_{\mathrm{U}}}\mathbf{G}_k^T\mathbf{V}_k\mathbf{V}_k^H\mathbf{G}_k^* + \sigma_{\mathrm{R}}^2\mathbf{I} + \mathbf{\Omega}\right)\\ &- \log\det\left(\sum_{j\in\mathcal{N}_{\mathrm{U}}\setminus\{k\}}\mathbf{G}_j^T\mathbf{V}_j\mathbf{V}_j^H\mathbf{G}_j^* + \sigma_{\mathrm{R}}^2\mathbf{I} + \mathbf{\Omega}\right)\\ \triangleq& C_{\mathrm{ac},k}(\{\mathbf{V}_k\},\mathbf{\Omega}).\end{aligned} \tag{5.14}$$

5.3 Problem Formulation

In this section, our objective is to enhance the network SE and EE in the uplink transmission through joint optimization of wireless fronthaul and access links. Assuming that perfect instantaneous CSI for both links is available to all the units involved in the system, we aim to maximize the total sum-rate for UEs and also minimize the power consumption with the ensured rate.

5.3.1 Sum-rate Maximization

Since the achievable rates for both links have been determined in Subsection 5.2.2, we now aim at maximizing the weighted sum-rate, with a weighting coefficient w_k, by optimizing the quantization noise covariance matrices $\mathbf{\Omega}_i$, the precoding matrices $\mathbf{V}_k$ and the transmit power allocation $p_{\mathrm{R},i}$ at each RRH. The weighting coefficient w_k can be selected to give a priority to a specific user. It is important to note that the rate in (5.14) is achievable if conditions (5.12) and (5.8) are satisfied.

Similar to the problems proposed in the previous Chapter 3, denote $\bar{\mathcal{V}} \triangleq \{\bar{\mathbf{V}}_j : \forall j \in \mathcal{N}_{\mathrm{U}}\}$ as the set of precoding matrices in the access link, for simplicity, where $\bar{\mathbf{V}}_j \triangleq \mathbf{V}_j \mathbf{V}_j^H, \ \forall j$. Accordingly, the optimization problem can thus be formulated as follows

$$\underset{\bar{\mathcal{V}} \succeq \mathbf{0},\ \mathbf{\Omega}_i \succeq \mathbf{0},\ \mathbf{p}_{\mathrm{R}} \geq 0}{\text{maximize}} \quad \sum_{k \in \mathcal{N}_{\mathrm{U}}} w_k C_{\mathrm{ac},k}(\bar{\mathcal{V}}, \mathbf{\Omega}) \tag{5.15a}$$

$$\text{s.t.} \quad \varphi_i(\bar{\mathcal{V}}, \mathbf{\Omega}) \leq \sum_{i \in \mathcal{S}_m} C_{\mathrm{fr},i}(\mathbf{p}_{\mathrm{R}}), \ \forall m \in \mathcal{N}_{\mathcal{S}}, \tag{5.15b}$$

$$p_{\mathrm{R},i} \leq P_{\mathrm{R},i}^{\max}, \forall i \in \mathcal{N}_{\mathrm{R}}, \tag{5.15c}$$

$$\mathrm{tr}\left(\bar{\mathbf{V}}_k\right) \leq P_{\mathrm{U},k}^{\max}, \forall k \in \mathcal{N}_{\mathrm{U}}. \tag{5.15d}$$

Constraint (5.15c) follows the assumption that all RRHs have their own power budget of $P_{\mathrm{R},i}^{\max}$. This should be allocated by each RRH for transmission and quantization. In this

case, a lower quantization noise can be achieved by allocating more power for quantization with the consequence of a lower power remaining for transmission. Constraint (5.15d) limits the maximum transmit power for each UE with their own budgets.

Algorithm 8. Joint design based sum-rate maximization for perfect instantaneous CSI in uplink transmission.

1: **Input:** Essential system parameters, including fronthaul channels $\mathbf{H}$, access channels $\mathbf{G}$;

2: **Initialization:** Set $r \coloneqq 0$ and a feasible point $\{\bar{\mathcal{V}}^{(0)} \succeq \mathbf{0}, \mathbf{\Omega}^{(0)} \succeq \mathbf{0}, \mathbf{p}_{\mathrm{R}}^{(0)} \geq 0\}$;

3: **repeat**

4: Update $r \coloneqq r + 1$;

5: Find the optimal solution $\{\bar{\mathcal{V}}^{(r)}, \mathbf{\Omega}^{(r)}, \mathbf{p}_{\mathrm{R}}^{(r)}\}$ by solving convex optimization problem (5.15) based on the known $\{\bar{\mathcal{V}}^{(r-1)}, \mathbf{\Omega}^{(r-1)}, \mathbf{p}_{\mathrm{R}}^{(r-1)}\}$

$$
\begin{aligned}
\underset{\bar{\mathcal{V}}^{(r)} \succeq \mathbf{0},\, \mathbf{\Omega}^{(r)} \succeq \mathbf{0},\, \mathbf{p}_{\mathrm{R}}^{(r)} \geq 0}{\text{maximize}} \quad & \sum_{k \in \mathcal{N}_{\mathrm{U}}} w_k C_{\mathrm{ac},k}^{\mathrm{lb}}(\bar{\mathcal{V}}^{(r)}, \mathbf{\Omega}^{(r)} | \bar{\mathcal{V}}^{(r-1)}, \mathbf{\Omega}_i^{(r-1)}), \\
\text{s.t.} \quad & \varphi_i^{\mathrm{ub}}(\bar{\mathcal{V}}^{(r)}, \mathbf{\Omega}^{(r)} | \bar{\mathcal{V}}^{(r-1)}, \mathbf{\Omega}^{(r-1)}) \leq C_i, \forall i \in \mathcal{N}_{\mathrm{R}}, \\
& C_{\mathrm{fr},i}^{\mathrm{lb}}(\mathbf{p}_{\mathrm{R}}^{(r)} | \mathbf{p}_{\mathrm{R}}^{(r-1)}) \geq C_i, \forall i \in \mathcal{N}_{\mathrm{R}}, \\
& p_{\mathrm{R},i}^{(r)} \leq P_{\mathrm{R},i}^{\max}, \forall i \in \mathcal{N}_{\mathrm{R}}, \\
& \mathrm{tr}\left(\bar{\mathbf{V}}_k^{(r)}\right) \leq P_{\mathrm{U},k}^{\max}, \forall k \in \mathcal{N}_{\mathrm{U}},
\end{aligned}
$$

6: **until** convergence;

7: **Output:** $\{\bar{\mathcal{V}}^{\star}, \mathbf{\Omega}^{\star}, \mathbf{p}_{\mathrm{R}}^{\star}\} \coloneqq \{\bar{\mathcal{V}}^{(r)}, \mathbf{\Omega}^{(r)}, \mathbf{p}_{\mathrm{R}}^{(r)}\}$.

Due to the non-convex objective function (5.15a) and fronthaul constraint (5.15b), problem (5.15) is difficult to solve. Similar to the solutions we provided in Chapter 3, we approximate those functions by a sequence of concave/convex functions following the MM method and solve the problem iteratively. Specifically, we approximate the nonconvex/nonconcave terms, $C_{\mathrm{ac},k}\left(\bar{\mathcal{V}}, \mathbf{\Omega}\right)$ in the objective function (5.15a), as well as $\varphi_m(\bar{\mathcal{V}}, \mathbf{\Omega})$ and $C_{\mathrm{fr},i}(\mathbf{p}_{\mathrm{R}})$ in constraint (5.15b), by linearizing them using a Taylor series $\Psi(\mathbf{A}, \mathbf{B})$ around the optimal point $\mathbf{B}$ obtained from an given point $\{\tilde{\bar{\mathcal{V}}}, \tilde{\mathbf{\Omega}}, \tilde{\mathbf{p}}_{\mathrm{R}}\}$, which is

defined by

$$\Psi(\mathbf{A},\mathbf{B}) \triangleq \log\det(\mathbf{B}) + \frac{1}{\ln 2}\mathrm{tr}(\mathbf{B}^{-1}(\mathbf{A}-\mathbf{B})). \tag{5.17}$$

Therefore, $C_{\mathrm{ac},k}\left(\bar{\mathcal{V}},\boldsymbol{\Omega}\right)$ in the objective function (5.15a) can be approximated by its convex lower bound as

$$\begin{aligned} C_{\mathrm{ac},k}(\bar{\mathcal{V}},\boldsymbol{\Omega}_i) \geq\ & C_{\mathrm{ac},k}^{\mathrm{lb}}(\bar{\mathcal{V}},\boldsymbol{\Omega}_i|\tilde{\bar{\mathcal{V}}},\tilde{\boldsymbol{\Omega}}_i) \\ \triangleq\ & \log\det\left(\sum_{k\in\mathcal{N}_{\mathrm{U}}}\mathbf{G}_k^T\bar{\mathbf{V}}_k\mathbf{G}_k^* + \sigma_{\mathrm{R}}^2\mathbf{I} + \boldsymbol{\Omega}\right) \\ & - \Psi\left(\sum_{j\in\mathcal{N}_{\mathrm{U}}\setminus\{k\}}\mathbf{G}_j^T\bar{\mathbf{V}}_j\mathbf{G}_j^* + \sigma_{\mathrm{R}}^2\mathbf{I} + \boldsymbol{\Omega}, \sum_{j\in\mathcal{N}_{\mathrm{U}}\setminus\{k\}}\mathbf{G}_j^T\tilde{\bar{\mathbf{V}}}_j\mathbf{G}_j^* + \sigma_{\mathrm{R}}^2\mathbf{I} + \tilde{\boldsymbol{\Omega}}\right). \end{aligned} \tag{5.18}$$

In addition, for constraint (5.15b), $\varphi_m(\bar{\mathcal{V}},\boldsymbol{\Omega})$ is upper bounded by a concave function as

$$\begin{aligned} \varphi_i(\bar{\mathcal{V}},\boldsymbol{\Omega}_i) \leq\ & \varphi_i^{\mathrm{ub}}(\bar{\mathcal{V}},\boldsymbol{\Omega}_i|\tilde{\bar{\mathcal{V}}},\tilde{\boldsymbol{\Omega}}_i) \\ \triangleq\ & \Psi\left(\boldsymbol{\Gamma}_i^H\left(\sum_{k\in\mathcal{N}_{\mathrm{U}}}\mathbf{G}_k^T\bar{\mathbf{V}}_k\mathbf{G}_k^*\right)\boldsymbol{\Gamma}_i + \sigma_{\mathrm{C}}^2\mathbf{I} + \boldsymbol{\Omega}_{ii},\right. \\ & \qquad \left.\boldsymbol{\Gamma}_i^H\left(\sum_{k\in\mathcal{N}_{\mathrm{U}}}\mathbf{G}_k^T\tilde{\bar{\mathbf{V}}}_k\mathbf{G}_k^*\right)\boldsymbol{\Gamma}_i + \sigma_{\mathrm{C}}^2\mathbf{I} + \tilde{\boldsymbol{\Omega}}_{ii}\right) \\ & - \log\det\left(\boldsymbol{\Gamma}_{\mathcal{S}_m}^H\boldsymbol{\Omega}\boldsymbol{\Gamma}_{\mathcal{S}_m}\right), \end{aligned} \tag{5.19}$$

whereas $C_{\mathrm{fr},i}(\mathbf{p}_{\mathrm{R}})$ is lower bounded by its convex approximation given as

$$\begin{aligned} C_{\mathrm{fr},i}(\mathbf{p}_{\mathrm{R}}) \geq\ & C_{\mathrm{fr},i}^{\mathrm{lb}}(\mathbf{p}_{\mathrm{R}}|\tilde{\mathbf{p}}_{\mathrm{R}}) \\ \triangleq\ & \log_2\left(\sigma_{\mathrm{C}}^2\|\mathbf{w}_i\|^2 + \sum_{i\in\mathcal{N}_{\mathrm{R}}}|\mathbf{w}_i\mathbf{h}_i^T|^2 p_{\mathrm{R},i}\right) \\ & - \Psi\left(\sigma_{\mathrm{C}}^2\|\mathbf{w}_i\|^2 + \sum_{j\in\mathcal{N}_{\mathrm{R}}\setminus\{i\}}|\mathbf{w}_i\mathbf{h}_j^T|^2 p_{\mathrm{R},j},\ \sigma_{\mathrm{C}}^2\|\mathbf{w}_i\|^2 + \sum_{j\in\mathcal{N}_{\mathrm{R}}\setminus\{i\}}|\mathbf{w}_i\mathbf{h}_j^T|^2\tilde{p}_{\mathrm{R},j}\right). \end{aligned} \tag{5.20}$$

In Algorithm 8, we summarize the steps to iteratively solve the resulting convex optimization problem. The optimized objective value is guaranteed to be non-decreasing and reach convergence [92].

5.3.2 Total Power Minimization

In this section, we further consider increasing EE for a guaranteed achievable sum-rate of the whole network. Similar to Subsections 3.3.2 and 4.2.2, the power minimization problems for the uplink transmission can be formulated as

$$\underset{\mathbf{p}\succeq 0,\ \bar{\mathcal{V}}\succeq \mathbf{0},\ \mathbf{\Omega}\succeq \mathbf{0}}{\text{minimize}} \quad \sum_{k\in\mathcal{N}_{\mathrm{U}}} \mathrm{tr}\left(\bar{\mathbf{V}}_k\right) + \sum_{i\in\mathcal{N}_{\mathrm{R}}} p_{\mathrm{R},i}, \tag{5.21a}$$

$$\text{s.t.} \quad \sum_{k\in\mathcal{N}_{\mathrm{U}}} C_{\mathrm{ac},k}^{\mathrm{lb}}\left(\bar{\mathcal{V}}, \mathbf{\Omega}\right) \geq \gamma, \tag{5.21b}$$

$$\varphi_i(\bar{\mathcal{V}}, \mathbf{\Omega}) \leq \sum_{i\in\mathcal{S}_m} C_{\mathrm{fr},i}(\mathbf{p}_{\mathrm{R}}),\ \forall m \in \mathcal{N}_{\mathcal{S}}, \tag{5.21c}$$

$$p_{\mathrm{R},i} \leq P_{\mathrm{R},i}^{\max}, \forall i \in \mathcal{N}_{\mathrm{R}}, \tag{5.21d}$$

$$\mathrm{tr}\left(\bar{\mathbf{V}}_k\right) \leq P_{\mathrm{U},k}^{\max}, \forall k \in \mathcal{N}_{\mathrm{U}}, \tag{5.21e}$$

where γ denotes the minimum guaranteed sum-rate for all UEs. The resulting problem (5.21) is hard to solve due to non-convex constraints (5.21b) and (5.21c). The MM method can be adopted to solve the problem iteratively and the steps are detailed in Algorithm 9.

5.4 Numerical Results

To evaluate the performance of our proposed system, we provide in this section some numerical simulations. We consider a square area of size 1 km^2, in which a CU is placed in the center. Around the CU, there are $N_{\mathrm{R}} = 4$ uniformly distributed RRHs and $N_{\mathrm{U}} = 10$ uniformly distributed UEs, each of which is equipped with $K = 2$ and $N = 3$ antennas, respectively. We make sure that the distance from RHHs to the CU is no closer than 10 m, and the distance between RRHs is no closer than 100 m. We consider the setup where the transmission between the fronthaul links and access links are separated in the time domain and operated at the same carrier frequency 1.9 GHz with the system bandwidth $B = 20$ MHz. As a result, we assume there is no self-interference between the transmit and receive antennas at the RRHs. For the fronthaul link, we consider LoS channel

Algorithm 9. Joint design based transmit power minimization for perfect instantaneous CSI in uplink transmission.

1: **Input:** Essential system parameters, including fronthaul channels $\mathbf{H}$, access channels $\mathbf{G}$, and

- **Average Initialization:** a random feasible point $\{\mathbf{p}^{(0)} \geq 0, \bar{\mathcal{U}}^{(0)} \succeq \mathbf{0}, \mathbf{\Omega}^{(0)} \succeq \mathbf{0}\}$;
- **Sum-rate Based Initialization:** an optimized feasible point $\{\mathbf{p}^{(0)} \geq 0, \bar{\mathcal{U}}^{(0)} \succeq \mathbf{0}, \mathbf{\Omega}^{(0)} \succeq \mathbf{0}\}$ obtained from the sum-rate maximization problem solved in Algorithm 1;

2: **Initialization:** Set $r \coloneqq 0$;

3: **repeat**

4: Update $r \coloneqq r + 1$;

5: Find the optimal solution $\{\mathbf{p}^{(r)}, \bar{\mathcal{U}}^{(r)}, \mathbf{\Omega}^{(r)}\}$ by solving problem (5.21) based on the given point $\{\mathbf{p}^{(r-1)}, \bar{\mathcal{U}}^{(r-1)}, \mathbf{\Omega}^{(r-1)}\}$:

$$
\begin{aligned}
\underset{\mathbf{p}^{(r)} \geq 0,\, \bar{\mathcal{V}}^{(r)} \succeq \mathbf{0},\, \mathbf{\Omega}^{(r)} \succeq \mathbf{0}}{\text{minimize}} \quad & \sum_{k \in \mathcal{N}_{\mathrm{U}}} \mathrm{tr}\left(\bar{\mathbf{V}}_k^{(r)}\right) + \sum_{i \in \mathcal{N}_{\mathrm{R}}} p_{\mathrm{R},i}^{(r)}, \\
\text{s.t.} \quad & \sum_{k \in \mathcal{N}_{\mathrm{U}}} C_{\mathrm{ac},k}^{\mathrm{lb}}\left(\bar{\mathcal{V}}^{(r)}, \mathbf{\Omega}^{(r)} | \bar{\mathcal{V}}^{(r-1)}, \mathbf{\Omega}^{(r-1)}\right) \geq \gamma, \\
& \varphi_i(\bar{\mathcal{V}}^{(r)}, \mathbf{\Omega}^{(r)} | \bar{\mathcal{V}}^{(r-1)}, \mathbf{\Omega}^{(r-1)}) \leq \sum_{i \in \mathcal{S}_m} C_{\mathrm{fr},i}(\mathbf{p}_{\mathrm{R}}^{(r)} | \mathbf{p}_{\mathrm{R}}^{(r-1)}), \ \forall m \in \mathcal{N}_{\mathcal{S}}, \\
& p_{\mathrm{R},i}^{(r)} \leq P_{\mathrm{R},i}^{\max}, \ \forall i \in \mathcal{N}_{\mathrm{R}}, \\
& \mathrm{tr}\left(\bar{\mathbf{V}}_k^{(r)}\right) \leq P_{\mathrm{U},k}^{\max}, \ \forall k \in \mathcal{N}_{\mathrm{U}};
\end{aligned}
$$

6: **until** convergence;

7: **Output:** $\{\mathbf{p}^{\star}, \bar{\mathcal{U}}^{\star}, \mathbf{\Omega}^{\star}\} \coloneqq \{\mathbf{p}^{(r)}, \bar{\mathcal{U}}^{(r)}, \mathbf{\Omega}^{(r)}\}$.

in (5.4) with a free-space path loss model. For the access link, we consider Rayleigh fading channel with a rather realistic path loss model specified by 3GPP in [100], where we use the scenario of *urban microcell* (UMi) with NLoS, and set the height of RRHs $h_{\mathrm{RHH}} = 22.5$ m and the height of UEs $h_{\mathrm{UE}} = 1.5$ m. Unless otherwise stated, the power budgets for the RRHs and the UEs are set as $P_{\mathrm{R}}^{\max} \triangleq P_{\mathrm{R},i}^{\max} = 30$ dBm, $\forall i \in \mathcal{N}_{\mathrm{R}}$, and $P_{\mathrm{U}}^{\max} \triangleq P_{\mathrm{U},k}^{\max} = 20$ dBm, $\forall k \in \mathcal{N}_{\mathrm{U}}$, respectively. Throughout this section, we consider that all UEs have the same weighting coefficients, i.e., $w_k = 1, \forall k \in \mathcal{N}_{\mathrm{U}}$.

We first show in Figure 5.2 the convergence of the proposed algorithm for the MR and ZF combining schemes with different numbers of antennas at the CU, M. We can see that for both MR and ZF, the convergence is achieved within 40 iterations. As expected, the ZF schemes outperforms MR for all cases. Furthermore, increasing the number of antennas at the CU always leads to a higher network sum-rate.

In Figure 5.3, we investigate the average sum-rate against the UEs transmit power $P_{\mathrm{U}}^{\max}$, where the average sum-rate is computed using Algorithm 8 with $P_{\mathrm{R}}^{\max} = 30$ dBM over several channel realizations. Figure 5.3 shows that the average sum-rate increases linearly with $P_{\mathrm{U}}^{\max}$. As we increase the number of antennas at the CU, M, the slope of the curve increases for both MR and ZF combining schemes. Moreover, we can observe that the ZF scheme generally outperforms the MR combining scheme. To achieve the same average sum-rate of 30 bits/s/Hz, the MR combining scheme requires 150 more CU antennas with slightly more UE transmit power than the ZF scheme. This is due to the well-known fact that ZF can suppress the interference between RRHs more effectively.

To further investigate the effect of the fronthaul links on the performance, we evaluate the average sum-rate against $P_{\mathrm{R}}^{\max}$ for $P_{\mathrm{U}}^{\max} = 20$ dBm in Figure 5.4. The average sum-rate initially increases as $P_{\mathrm{R}}^{\max}$ increases, but then appears to be bounded at high $P_{\mathrm{R}}^{\max}$ for both MR and ZF combining schemes. Note that we only specify the maximum available transmit power $P_{\mathrm{R}}^{\max}$, which is the same for all RRHs. However, different RRHs can have different transmit power $p_{\mathrm{R},i}$ as a variable of the optimization problem. Since Algorithm 8 should jointly give the optimal $p_{\mathrm{R},i}$ and $\mathbf{\Omega}_i$, increasing $P_{\mathrm{R}}^{\max}$ is not expected to increase

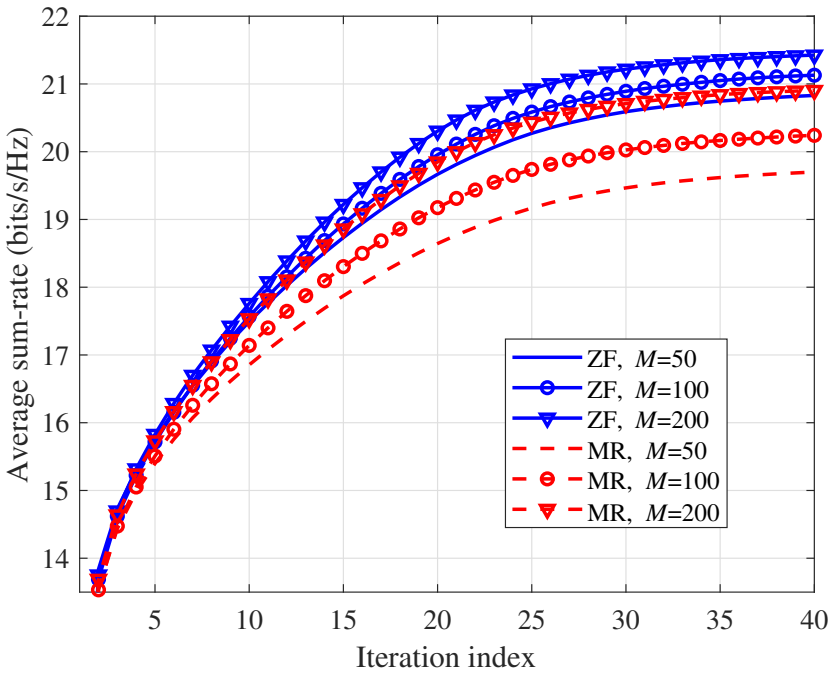


Figure 5.2: Convergence speed in terms of average sum-rate against iteration index of the proposed algorithm for MR and ZF fronthaul beamforming with different number of antennas at the CU, M.

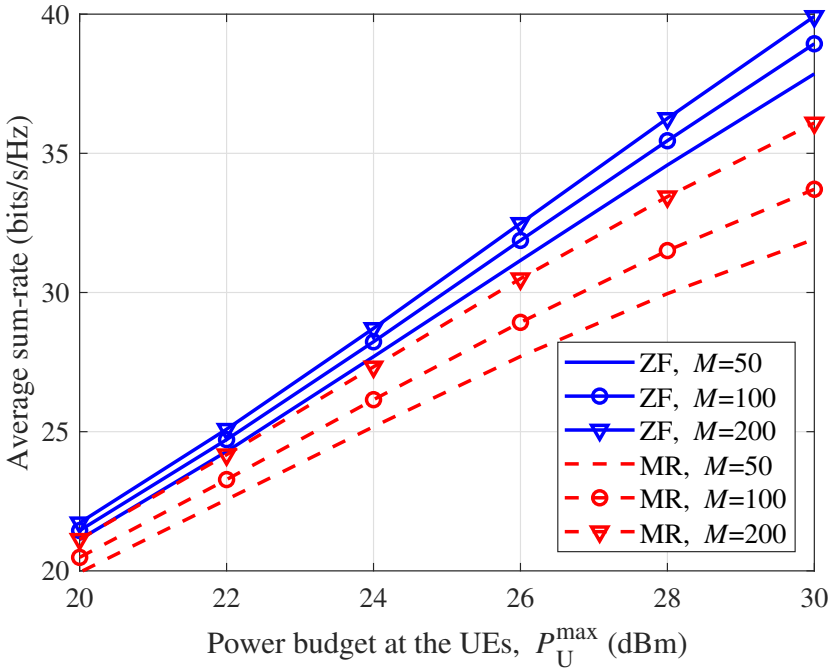


Figure 5.3: Average sum-rate with respect to the UE transmit power $P_{\mathrm{U}}^{\max}$ for MR and ZF combining schemes, and for different number of antennas at the CU, M.

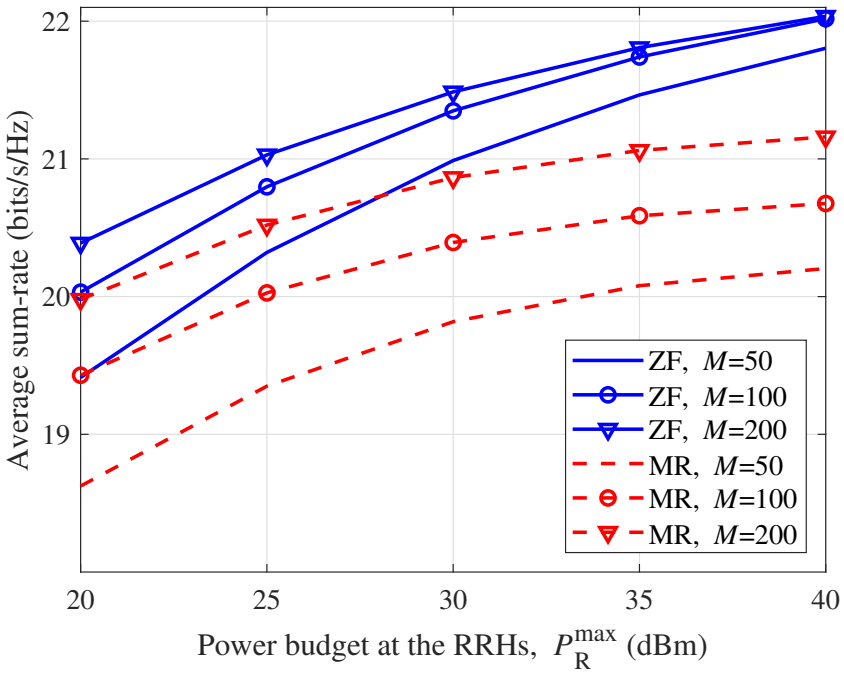


Figure 5.4: Average sum-rate with respect to the maximum available transmit power at RRHs $P_{\mathrm{R}}^{\max}$ for MR and ZF combining schemes, and for different numbers of antennas at the CU, M.

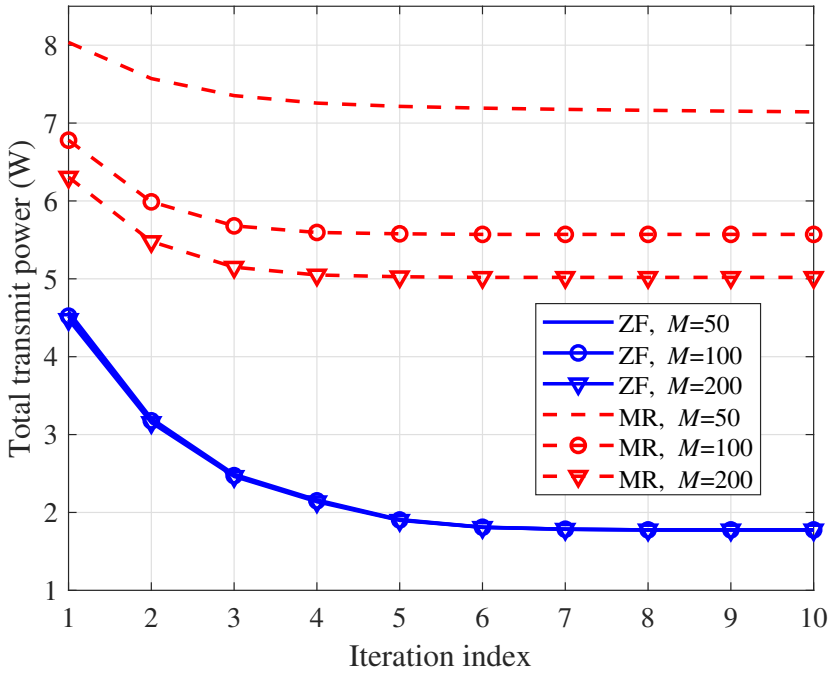


Figure 5.5: Convergence speed of the power minimization problem in Algorithm 9 with sum-rate based initialization and minimum required sum-rate $\gamma = 30$ bits/s/Hz for MR and ZF combining schemes, and for different numbers of antennas at the CU, M.

the sum-rate beyond a certain point to satisfy constraint (5.15b). This might explain the asymptotic behavior observed in Figure 5.4. In this regard, increasing the number of antennas M and using a better fronthaul combining scheme can enlarge the feasibility set, as can be seen from the curve that shifts upwards. As we increase M, the gap between the curves appears to get smaller. It is shown more clearly for ZF, which indicates a bound in M. We leave the further investigation of this bound as our future work.

In Figure 5.5, the total power minimization problem is simulated by utilizing the average initialization as the starting point. Both ZF and MR combining schemes are considered, along with different numbers of antennas M equipped at the CU. It is obvious that the convergence is achieved rapidly within 10 iterations. For the MR scheme, a larger number of antennas at the CU results in an obvious reduction in total consumption. However, this trend is almost negligible for the ZF scheme. Despite the small impact of increasing M, the ZF scheme still outperforms MR and achieves a significant reduction in power consumption through the joint design presented in Algorithm 9.

Figures 5.6 and 5.7 compare the different initialization methods with various required sum-rate γ. Figure 5.6 utilizes the average initialization method and the feasibility of such problem is very small with large γ, especially when $\gamma \geq 10$ bits/s/Hz. Conversely, Figure 5.7 employs the optimal solutions obtained from the sum-rate maximization problem given in Algorithm 8, which is able to minimize the power consumption under a much higher ensured sum-rate. Figure 5.6 illustrates that at the start of the optimization, the ZF scheme requires slightly more total power than the MR scheme, whereas both schemes substantially reduce power consumption to a low level by the end of the optimization process. Conversely, Figure 5.7 demonstrates a distinct gap between the ZF and MR schemes, with the ZF scheme requiring significantly less power, particularly when ensuring a high sum-rate. For both scenarios, the proposed joint power minimization problem can significantly reduce the total transmit power for both ZF and MR schemes.

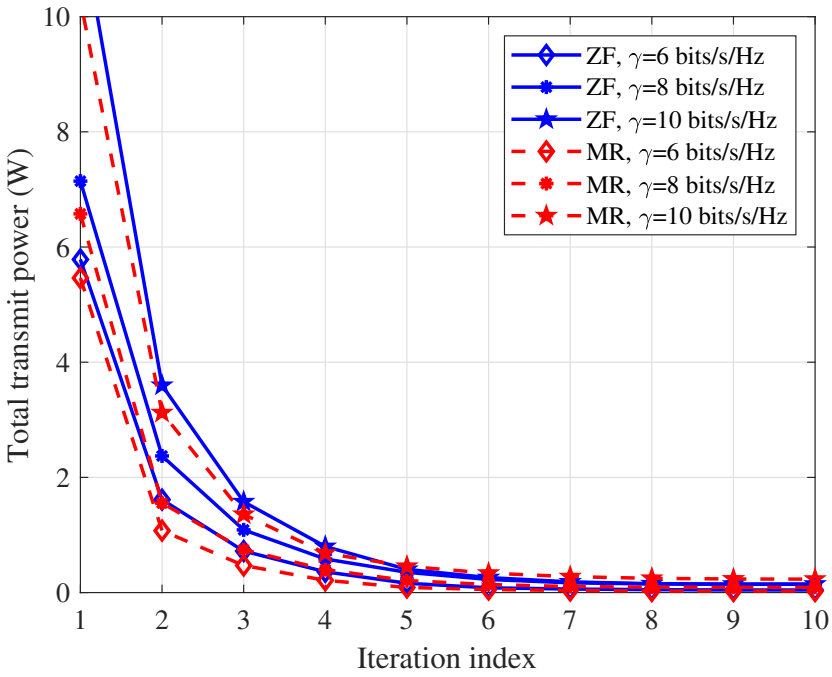


Figure 5.6: Convergence speed of the power minimization problem in Algorithm 9 with average initialization method, $M = 200$, for different required sum-rate γ.

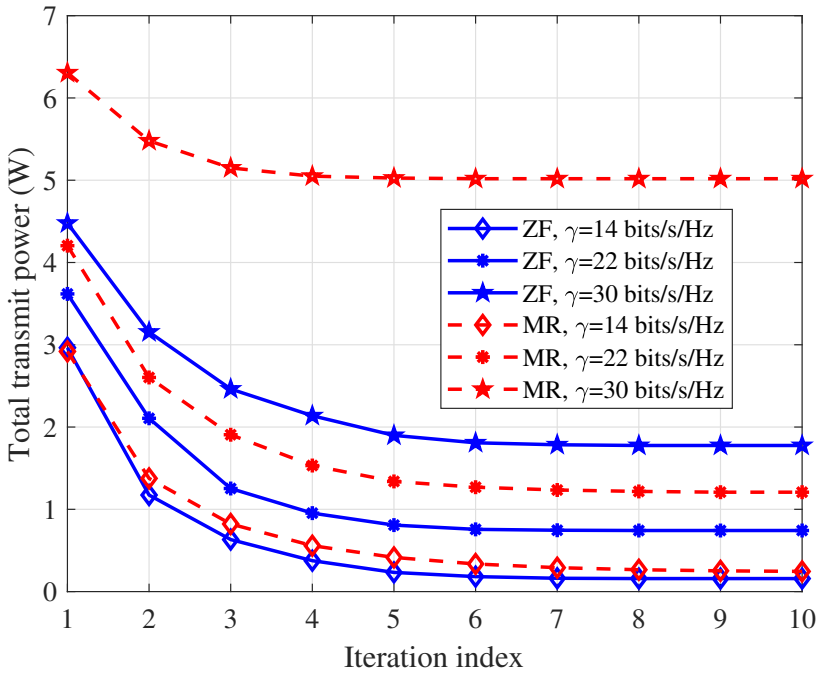


Figure 5.7: Convergence speed of the power minimization problem in Algorithm 9 with sum-rate based initialization method, $M = 200$, for different required sum-rate γ.

5.5 Conclusion

In this chapter, we studied the utilization of massive MIMO fronthaul in the uplink transmission of the CRAN system. To maximize the sum-rate and minimize the total power consumption, we jointly optimized the precoding matrices at the UEs, the quantization covariance matrices, and the transmit powers at the RRHs. The conventional precoding methods, namely MR and ZF combining, were employed for the fronthaul transmission. Furthermore, the CF method was used at the RRHs. Iterative algorithms have also been presented to solve non-convex optimization problems based on the MM approach.

Numerical simulations were conducted to evaluate the performance of our proposed design with respect to the MR and ZF combining schemes. As expected, the ZF scheme outperformed the MR combining scheme. For both schemes, we observed an asymptotic behavior of the sum-rate with respect to the maximum available transmit power at the RRHs. This insight suggests that relaxing the fronthaul bottleneck by increasing the available power at RRHs may be less effective. Alternatively, increasing the number of antennas at the central unit may result in greater improvements.

Chapter 6

Multi-UAV-enabled CRANs with Massive MIMO

6.1 Introduction

In this chapter, we propose a novel architecture for multi-UAV-enabled CRAN. In particular, UAVs are proposed to be deployed as flying RRHs to serve ground UEs. Due to many advantages such as fast configuration, high mobility, and low cost, using UAVs as flying RRHs in CRAN is very promising in certain applications where conventional CRAN might not be suitable. During unexpected or occasional situations, such as urgent accidents or provisional events, the demand for data throughput to support monitoring and information exchange can rapidly increase. Also, in case of disasters where terrestrial communications might stop working (e.g., BSs are damaged), UAVs can be deployed quickly and timely as flying BSs to ensure reliable coverage of the impacted area. The combination of UAV-enabled communication and CRAN structure can realize temporary and cost-effective communication requirements with fast deployment and wide-range coverage. As the fronthaul links in the proposed CRAN architecture, unlike the conventional one, have to be wireless, we propose to use a large-scale antenna array equipped at CU to

enhance the capacity of the fronthaul link.

The system model of the proposed architecture is presented in Section 6.2. To validate the proposed model and aim to maximize the whole network, a max-min fairness optimization problem and one of the potential solutions are proposed in Section 6.3. Section 6.4 provides numerical results to verify the effectiveness of the proposed system. Finally, the conclusion of this chapter is drawn in Section 6.5.

6.2 System Model

Consider the downlink transmission in a UAV-enabled CRAN system consisting of $N_{\mathrm{U}} > 1$ single-antenna UE and single-antenna UAV pairs [1]. In such case, $N_{\mathrm{R}} = N_{\mathrm{U}}$ represents the number of UAVs. The CU is assumed to be equipped with a large-scale UPA of $M \gg N_{\mathrm{U}}$ antennas. TDD transmission mode is adopted here to coordinate the transmissions among the fronthaul and access links [2]. For the sake of implementation simplicity, this chapter assumes that direct links between the CU and the UEs either do not exist or are too weak to be exploited.

Figure 6.1 shows the geometric model of the system. A 3-D coordinate system (x, y, z) is established where the ground is set as the x-y plane. One corner antenna of UPA is located at $(0, 0, H_{\mathrm{C}})$, where H_{C} is the CU height. We assume that the UPA is equipped with M_y rows of antennas where all the rows are parallel with the x-axis. Each row has M_x antennas with an antenna spacing $\delta = \lambda/2$, where λ is the wavelength. Hence, the total number of antennas at the CU is $M = M_x M_y$. For notational convenience, we define

[1] We note that a UAV can serve multiple UEs via using multiple access techniques such as FDMA, TDMA, *code-division multiple access* (CDMA), OFDMA, etc.

[2] Note that the investigation of the signalling overheads among the CU, UAVs, and UEs as well as their costs and timeliness is beyond the scope of this chapter.

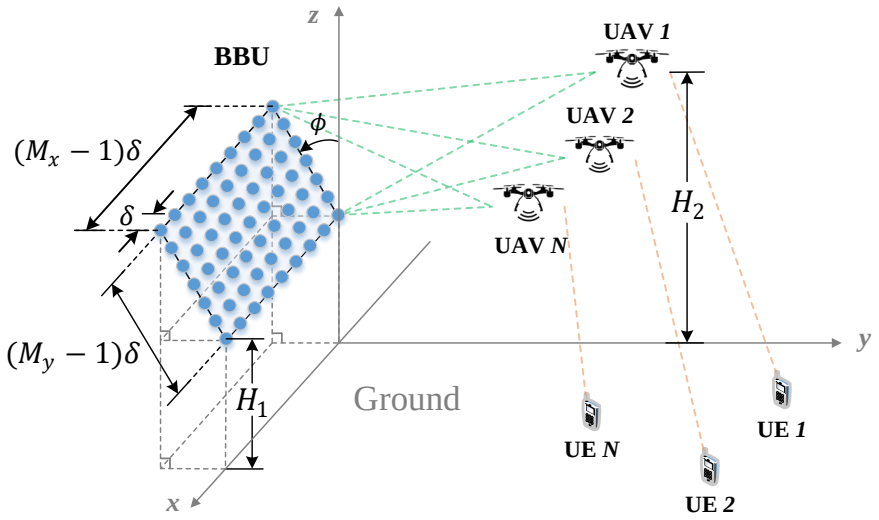


Figure 6.1: System model of the proposed UAV-enabled CRAN.

the antenna element in the i-th column and the j-th row as the m-th antenna where

$$m = i + (j-1)M_x, \quad \forall i \in \{1, \ldots, M_x\}, j \in \{1, \ldots, M_y\}. \tag{6.1}$$

Also, we define $\mathcal{N}_{\mathrm{U}} = \{1, \ldots, N_{\mathrm{U}}\}, \mathcal{N}_{\mathrm{R}} = \{1, \ldots, N_{\mathrm{U}}\}$ and $\mathcal{N}_{\mathrm{C}} = \{1, \ldots, M\}$ as the sets of UEs, UAVs, and the antennas at the CU, respectively.

Based on this coordinate system, the location of the m-th antenna is defined as

$$\mathbf{v}_{\mathrm{C},m} = ((i-1)\delta, \ -(j-1)\delta \sin\phi, \ (j-1)\delta\cos\phi + H_{\mathrm{C}}), \tag{6.2}$$

where ϕ is the inclined angle between the array and the x-z plane as shown in Figure 6.1. We assume that the UAVs fly at a constant altitude H_{R}. The locations of the i-th UAV and the k-th UE are defined as

$$\boldsymbol{q}_i = (x_i, \ y_i, \ H_{\mathrm{R}}), \ \forall i \in \mathcal{N}_{\mathrm{R}}, \tag{6.3}$$

and

$$\mathbf{v}_{\mathrm{U},k} = (x_{\mathrm{U},k}, \ y_{\mathrm{U},k}, \ 0), \ \forall k \in \mathcal{N}_{\mathrm{U}}, \tag{6.4}$$

respectively. Assuming that the locations of the UEs are constant and perfectly known. Therefore, distance l_{mi} between the m-th CU antenna element and the i-th UAV is

$$l_{mi} = \|\boldsymbol{q}_i - \mathbf{v}_{\mathrm{C},m}\|. \tag{6.5}$$

and distance d_{ik} between UAV i and UE k is

$$d_{ik} = ||\boldsymbol{q}_i - \mathbf{v}_{\mathrm{U},k}|| = \sqrt{(x_i - x_{\mathrm{U},k})^2 + (y_i - y_{\mathrm{U},k})^2 + H_{\mathrm{R}}^2}, \tag{6.6}$$

It is worth noting that some essential information, such as the locations of CU and all UEs and CSI for both links, is assumed to be well-known to all the units and keep constant in the same frame due to the change in a very short time being negligible.

Let $\mathbf{h}_i \triangleq [h_{1i}, \ldots, h_{Mi}]^T$ and $\boldsymbol{g}_k \triangleq [g_{1k}, \ldots, g_{Nk}]^T$ denote the downlink channel coefficients from all antennas at the CU to the i-th UAV and from all UAVs to the k-th UE, respectively. Since the ground-to-air and air-to-ground channels are generally dominated by the LoS link in practice, for simplicity, we assume that both the fronthaul and access links are modelled as LoS channels [56, 80]. As a result, the channels between the m-th CU antenna and the i-th UAV and between the i-th UAV and the k-th UE are defined as

$$h_{mi} = \sqrt{\frac{\xi}{l_{mi}^2}}\bar{h}_{mi} \quad \text{and} \quad g_{ik} = \sqrt{\frac{\xi}{d_{ik}^2}}\bar{g}_{ik}, \tag{6.7}$$

respectively, where

$$\bar{h}_{mi} = e^{-\mathrm{i}\frac{2\pi}{\lambda}l_{mi}} \quad \text{and} \quad \bar{g}_{ik} = e^{-\mathrm{i}\frac{2\pi}{\lambda}d_{ik}} \tag{6.8}$$

denote the phase shifts of the fronthaul and access links, respectively, and ξ represents the path loss at the reference distance $d_0 = 1$m in both links. For future use, we define $\bar{\mathbf{h}}_i \triangleq [\bar{h}_{1i}, \ldots, \bar{h}_{Mi}]^T$.

Denote the message that the CU intends to transmit to the k-th UE is encoded as $s_{\mathrm{U},k} \sim \mathcal{CN}(0, 1)$. Thus the precoded signal sent from the paired k-th UAV is given as

$$\bar{x}_{\mathrm{R},k} = \sqrt{p_{\mathrm{R},k}}\, s_{\mathrm{U},k}, \tag{6.9}$$

where $p_{\mathrm{R},k}$ is the power of $\bar{x}_{\mathrm{R},k}$. The CU first quantizes and compresses $\bar{x}_{\mathrm{R},k}$. We adopt the decompress-and-forward relaying strategy and use the Gaussian quantization test channel to model the quantization process [39]. Hence, the resulting quantized signal, $x_{\mathrm{R},k}$, can be expressed as

$$x_{\mathrm{R},k} = \bar{x}_{\mathrm{R},k} + \omega_k, \tag{6.10}$$

where $\omega_k \sim \mathcal{CN}(0, \Omega_k)$ is the compression noise. By assuming that the signals $\bar{x}_{\mathrm{R},k}$, $k \in \mathcal{N}_{\mathrm{U}}$, are compressed independently, their corresponding quantization noises are hence uncorrelated, i.e., $\mathbb{E}[\omega_i \omega_j^*] = 0$ for $i \neq j$ [39]. Let $\mathbf{\Omega} = [\Omega_1, \ldots, \Omega_N]^T$. The CU then compresses the quantized signal $x_{\mathrm{R},i}$ to generate the compression index $V_i \in \left\{1, 2, \ldots, 2^{n_{\mathrm{L}} C_i}\right\}$, where n_{L} denotes the coding block length and C_i denotes the rate of message V_i. Subsequently, the CU encodes each message V_i to obtain the encoded baseband signal $s_{\mathrm{R},i} \sim \mathcal{CN}(0, 1)$ and generates the normalized precoding vector $\mathbf{w}_i \in \mathbb{C}^{M \times 1}$ for $s_{\mathrm{R},i}$. $\mathbf{w}_i$ denotes the i-th column of the precoding matrix $\mathbf{W} \in \mathbb{C}^{M \times N}$. Here, we consider two conventional low-complexity linear precoders, namely the MF and ZF precoding methods. Thus, $\mathbf{W}$ can be expressed as

$$\mathbf{W} \triangleq \begin{cases} \bar{\mathbf{H}}^*/M, & \text{for MF,} \\ \bar{\mathbf{H}}^* \left(\bar{\mathbf{H}}^T \bar{\mathbf{H}}^*\right)^{-1}, & \text{for ZF,} \end{cases} \tag{6.11}$$

where $\bar{\mathbf{H}} \triangleq [\bar{\mathbf{h}}_1, \ldots, \bar{\mathbf{h}}_{N_{\mathrm{R}}}]$. It is worth noting that (6.11) is slightly different from ZF method used in (3.7) as l_{mi} is no longer fixed and needs to be taken into account in this chapter. Next, the CU sends the precoded signal $\sum_{i \in \mathcal{N}_{\mathrm{R}}} \mathbf{w}_i \sqrt{p_{\mathrm{C},i}} s_{\mathrm{R},i}$ to the UAVs, where $p_{\mathrm{C},i}$ is a real-valued coefficient represents transmit power. The signal received by the i-th UAV is given by

$$y_{\mathrm{R},i} = \mathbf{h}_i^T \mathbf{w}_i \sqrt{p_{\mathrm{C},i}} s_{\mathrm{R},i} + \mathbf{h}_i^T \sum_{j \in \mathcal{N}_{\mathrm{R}} \setminus \{i\}} \mathbf{w}_j \sqrt{p_{\mathrm{C},j}} s_{\mathrm{R},j} + n_{\mathrm{R},i}, \tag{6.12}$$

where $n_{\mathrm{R},i} \sim \mathcal{CN}(0, \sigma_{\mathrm{R}}^2)$ denotes the noise at the i-th UAV, $i \in \mathcal{N}_{\mathrm{R}}$. Therefore, the rate C_i through the fronthaul link should be constrained as

$$C_i \leq C_{\mathrm{fr},i}(\mathbf{Q}, \mathbf{p}_{\mathrm{C}}) \triangleq \log_2 \left(1 + \frac{p_{\mathrm{C},i} \left|\mathbf{h}_i^T \mathbf{w}_i\right|^2}{\sum\limits_{j \in \mathcal{N}_{\mathrm{R}} \setminus \{i\}} p_{\mathrm{C},j} \left|\mathbf{h}_i^T \mathbf{w}_j\right|^2 + \sigma_{\mathrm{R}}^2}\right), \tag{6.13}$$

where $\mathbf{Q} = \left[\boldsymbol{q}_1^T, \ldots, \boldsymbol{q}_N^T\right]^T$ and $\mathbf{p}_{\mathrm{C}} = [p_{\mathrm{C},1}, \ldots, p_{\mathrm{C},N}]^T$. Then, each UAV i decompresses the signal received from the CU to recover the message $x_{\mathrm{R},i}$. In order to decompress the signals successfully at the UAVs, the message rate C_i should be bounded as [67]

$$C_i \geq \varphi_i(p_{\mathrm{R},i}, \Omega_i) \triangleq \log_2 \left(1 + \frac{p_{\mathrm{R},i}}{\Omega_i}\right), \; i \in \mathcal{N}_{\mathrm{R}}. \tag{6.14}$$

The message recovered by UAV i, $x_{\mathrm{R},i}$, is then forwarded to its paired UE i. The signal received by the k-th UE is given by

$$y_{\mathrm{U},k} = g_{kk}\sqrt{p_{\mathrm{R},k}}s_{\mathrm{U},k} + \sum_{j\in\mathcal{N}_{\mathrm{R}}\setminus\{k\}} g_{jk}\sqrt{p_{\mathrm{R},j}}s_{\mathrm{U},j} + \sum_{j\in\mathcal{N}_{\mathrm{R}}} g_{jk}\omega_j + n_{\mathrm{U},k}, \tag{6.15}$$

where $n_{\mathrm{U},k} \sim \mathcal{CN}(0,\sigma_{\mathrm{U}}^2)$ denotes the noise at UE k. Hence, the achievable rate for UE k can be obtained as

$$R_k \triangleq C_{\mathrm{ac},k}\left(\mathbf{Q},\mathbf{p}_{\mathrm{R}},\mathbf{\Omega}\right). \tag{6.16}$$

In order to simplify the steps to solve the problem, it is easy to decompose $C_{\mathrm{ac},k}\left(\mathbf{Q},\mathbf{p}_{\mathrm{R}},\mathbf{\Omega}\right)$ into two parts

$$R_k = R_{k,1}\left(\mathbf{Q},\mathbf{p}_{\mathrm{R}},\mathbf{\Omega}\right) - R_{k,2}\left(\mathbf{Q},\mathbf{p}_{\mathrm{R}},\mathbf{\Omega}\right). \tag{6.17}$$

where

$$R_{k,1}\left(\mathbf{Q},\mathbf{p}_{\mathrm{R}},\mathbf{\Omega}\right) = \log_2\Big(\sum_{j\in\mathcal{N}_{\mathrm{R}}}\left(p_{\mathrm{R},j}+\Omega_j\right)\left|g_{jk}\right|^2+\sigma_{\mathrm{U}}^2\Big), \tag{6.18a}$$

$$R_{k,2}\left(\mathbf{Q},\mathbf{p}_{\mathrm{R}},\mathbf{\Omega}\right) = \log_2\Big(\sum_{j\in\mathcal{N}_{\mathrm{R}}\setminus\{k\}} p_{\mathrm{R},j}\left|g_{jk}\right|^2+\sum_{j\in\mathcal{N}_{\mathrm{R}}}\Omega_j\left|g_{jk}\right|^2+\sigma_{\mathrm{U}}^2\Big). \tag{6.18b}$$

In the next section, we jointly optimize the power control coefficients, the quantization noise variance, and the UAV placement according to the max-min fairness criterion[3].

6.3 Max-min Fairness Optimization Problem

In this section, our goal is to maximize the minimum rate among all UEs via the optimization of UAVs placement $\mathbf{Q}$, power control $\mathbf{p}_{\mathrm{C}},\mathbf{p}_{\mathrm{R}}$, and quantization noise variance $\mathbf{\Omega}$. According to (6.13), (6.14), and (6.17), the corresponding optimization problem can

[3]Note that it is straightforward to extend this work to other criteria such as the sum-rate criterion.

be formulated as

$$\max_{\mathbf{Q},\mathbf{p}_{\mathrm{C}},\mathbf{p}_{\mathrm{R}},\mathbf{\Omega},\mathcal{C}} \ \min_{k} \ R_k\left(\mathbf{Q},\mathbf{p}_{\mathrm{R}},\mathbf{\Omega}\right),\ \forall k \in \mathcal{N}_{\mathrm{U}}, \tag{6.19a}$$

$$\text{s.t.} \quad C_i \le C_{\mathrm{fr},i}(\mathbf{Q},\mathbf{p}_{\mathrm{C}}),\ \forall i \in \mathcal{N}_{\mathrm{R}}, \tag{6.19b}$$

$$C_i \ge \varphi_i(p_{\mathrm{R},i},\Omega_i),\ \forall i \in \mathcal{N}_{\mathrm{R}}, \tag{6.19c}$$

$$p_{\mathrm{R},i} + \Omega_i \le P_{\mathrm{R},i}^{\max},\ \forall i \in \mathcal{N}_{\mathrm{R}}, \tag{6.19d}$$

$$\mathrm{tr}\left\{\mathrm{diag}\left(p_{\mathrm{C},1},\ldots,p_{\mathrm{C},N_{\mathrm{R}}}\right)\mathbf{W}^H\mathbf{W}\right\} \le P_{\mathrm{C}}^{\max}, \tag{6.19e}$$

$$\|\boldsymbol{q}_i - \boldsymbol{q}_j\|^2 \ge d_{\min}^2,\ \forall i,j \in \mathcal{N}_{\mathrm{R}},\ i \ne j, \tag{6.19f}$$

$$p_{\mathrm{C},i} \ge 0,\ p_{\mathrm{R},i} \ge 0,\ \Omega_i \ge 0,\ \forall i \in \mathcal{N}_{\mathrm{R}}, \tag{6.19g}$$

where $\mathcal{C} \triangleq \{C_i : \forall i \in \mathcal{N}_{\mathrm{R}}\}$ and $d_{\min}$ is the minimum safety distance between any two UAVs to avoid collisions. (6.19d) and (6.19e) refer to the power constraints at each UAV and the CU, respectively. $P_{\mathrm{R},i}^{\max}$ and $P_{\mathrm{C}}^{\max}$ denote the maximum transmit powers at RRH i and CU, respectively.

It is clear that problem (6.19) is very hard to solve because it is non-convex. To solve it, we propose to decompose it into two convex sub-problems and solve them iteratively by applying the BCD and SCA method [56].

6.3.1 Sub-Problem 1: UAV placement optimization

Firstly, for any given $\mathbf{p}_{\mathrm{C}},\mathbf{p}_{\mathrm{R}},\mathbf{\Omega}$, we aim to optimize the UAV placement $\mathbf{Q}$. Therefore, sub-problem 1 is derived from problem (6.19) as

$$\max_{\mathbf{Q},\mathcal{C}} \ \min_{k} \ R_{k,1}\left(\mathbf{Q}\right) - R_{k,2}\left(\mathbf{Q}\right),\ \forall k \in \mathcal{N}_{\mathrm{U}}, \tag{6.20a}$$

$$\text{s.t.} \quad C_i \le C_{\mathrm{fr},i}(\mathbf{Q}),\ \forall i \in \mathcal{N}_{\mathrm{R}}, \tag{6.20b}$$

$$C_i \ge \varphi_i,\ \forall i \in \mathcal{N}_{\mathrm{R}}, \tag{6.20c}$$

$$\|\boldsymbol{q}_i - \boldsymbol{q}_k\|^2 \ge d_{\min}^2,\ \forall i,k \in \mathcal{N}_{\mathrm{R}},\ i \ne k. \tag{6.20d}$$

It is observed that the objective function (6.20a) and constraints (6.20b) and (6.20c) are all non-convex. Define slack variables $\mathbf{S} \triangleq \{S_{ik} :\ \forall i \in \mathcal{N}_{\mathrm{R}}, k \in \mathcal{N}_{\mathrm{U}}\}$. Then, $R_{k,2}$ in (6.18a)

can equivalently be rewritten as

$$R_{k,2}(\mathbf{S}) = \log_2\left(\xi \sum_{i\in\mathcal{N}_\mathrm{R}\setminus\{k\}} \frac{p_{\mathrm{R},i}+\Omega_i}{S_{ik}} + \xi\frac{\Omega_j}{S_{kk}} + \sigma_\mathrm{R}^2\right). \tag{6.21}$$

Thus, problem (6.20) can be recast as

$$\max_{\mathbf{Q},\mathbf{S},\mathcal{C}} \ \min_{k} \ R_{k,1}(\mathbf{Q}) - R_{k,2}(\mathbf{S}), \ \forall k \in \mathcal{N}_\mathrm{U}, \tag{6.22a}$$

$$\text{s.t.} \ \ S_{ik} \le \|\boldsymbol{q}_i - \mathbf{v}_{\mathrm{U},k}\|^2, \ \forall i \in \mathcal{N}_\mathrm{R}, k \in \mathcal{N}_\mathrm{U}, \tag{6.22b}$$

$$(6.20\text{b}), \ (6.20\text{c}), \ (6.20\text{d}).$$

Note that $R_{k,1}(\mathbf{Q})$ is neither convex nor concave with respect to $\boldsymbol{q}_i$, but it is convex with respect to $\|\boldsymbol{q}_i - \mathbf{v}_{\mathrm{U},k}\|^2$. Define $\boldsymbol{q}_i^{(r)} = \left\{\boldsymbol{q}_i^{(r)}, \forall i \in \mathcal{N}_\mathrm{R}\right\}$ as the placement solution from the $(r-1)$-th iteration. By taking the first-order Taylor expansion at the point $\|\boldsymbol{q}_i^{(r)} - \mathbf{v}_{\mathrm{U},k}\|^2$, we can get the lower bound for $R_{k,1}$ with respect to $\|\boldsymbol{q}_i - \mathbf{v}_{\mathrm{U},k}\|^2$ as

$$\begin{aligned} R_{k,1} &= \log_2\left(\xi \sum_{i\in\mathcal{N}_\mathrm{R}} \frac{p_{\mathrm{R},i}+\Omega_i}{\|\boldsymbol{q}_i - \mathbf{v}_{\mathrm{C},k}\|^2} + \sigma_\mathrm{U}^2\right) \\ &\ge \sum_{i\in\mathcal{N}_\mathrm{R}} -A_{k,i}^{(r)}\left(\|\boldsymbol{q}_i - \mathbf{v}_{\mathrm{C},k}\|^2 - \|\boldsymbol{q}_i^{(r)} - \mathbf{v}_{\mathrm{C},k}\|^2\right) + B_k^{(r)} \\ &\triangleq R_{k,1}^{\mathrm{lb}}, \ \forall k \in \mathcal{N}_\mathrm{U}, \end{aligned} \tag{6.23}$$

where

$$A_{k,i}^{(r)} = \frac{\xi(p_{\mathrm{R},i}+\Omega_i)\log_2 e}{\|\boldsymbol{q}_i^{(r)} - \mathbf{v}_{\mathrm{C},k}\|^4\left(\sum\limits_{k\in\mathcal{N}_\mathrm{R}} \frac{\xi(p_{\mathrm{R},k}+\Omega_k)}{\|\boldsymbol{q}_k^{(r)} - \mathbf{v}_{\mathrm{C},k}\|^2} + \sigma_\mathrm{U}^2\right)} \tag{6.24}$$

and

$$B_k^{(r)} = \log_2\left(\sum_{k\in\mathcal{N}_\mathrm{R}} \frac{\xi(p_{\mathrm{R},k}+\Omega_k)}{\|\boldsymbol{q}_k^{(r)} - \mathbf{v}_{\mathrm{C},k}\|^2} + \sigma_\mathrm{U}^2\right) \tag{6.25}$$

are constants.

In the fronthaul links, since the antenna spacing at the CU is negligible compared with the distance between the CU and UAVs, it is reasonable to assume that all the links to the same UAV are identical and hence $l_{mi} \approx l_i, \forall m \in \mathcal{N}_\mathrm{C}$. Therefore, the fronthaul channel

model in (6.7) can be re-expressed as $h_{mi} = \sqrt{\xi/l_i^2}\,\bar{h}_{mi},\ \forall i \in \mathcal{K}_{\mathrm{R}}$ and thus the capacity $C_{\mathrm{fr},i}(\mathbf{Q}, \mathbf{p}_{\mathrm{C}})$ in (6.13) can be simplified as

$$C_{\mathrm{fr},i}(\mathbf{Q}, \mathbf{p}_{\mathrm{C}}) = \log_2\left(1 + \frac{p_{\mathrm{C},i}\left|\sqrt{\frac{\xi}{l_i^2}}\,\bar{\mathbf{h}}_i^T\mathbf{w}_i\right|^2}{\sum\limits_{j\in\mathcal{N}_{\mathrm{R}}\setminus\{i\}} p_{\mathrm{C},j}\left|\sqrt{\frac{\xi}{l_i^2}}\,\bar{\mathbf{h}}_i^T\mathbf{w}_j\right|^2 + \sigma_{\mathrm{R}}^2}\right), \tag{6.26}$$

As indicated in [44], asymptotic orthogonality would exist in LoS scenario if a UPA is applied. The first term of the numerator in (6.26) tends to zero when M tends to infinity. Therefore, we ignore this part for simplifying the optimization problem and the constraint (6.20b) is replaced as

$$\begin{aligned} C_i \le C_{\mathrm{fr},i}(\mathbf{Q}) &\approx \log_2\left(1 + \frac{\xi p_{\mathrm{C},i}}{\sigma_{\mathrm{R}}^2\left|\left|\boldsymbol{q}_i - \mathbf{v}_{\mathrm{U},1}\right|\right|^2}\right) \\ &\triangleq \widetilde{C}_{\mathrm{fr},i}(\boldsymbol{q}_i),\ \forall i \in \mathcal{K}_{\mathrm{R}}. \end{aligned} \tag{6.27}$$

$\widetilde{C}_{\mathrm{fr},i}(\boldsymbol{q}_i)$ is convex with respect to $\|\boldsymbol{q}_i - \mathbf{v}_{\mathrm{U},1}\|^2$, but constraint (6.27) is still non-convex with respect to variable $\boldsymbol{q}_i$. Therefore, we derive the lower bound at the point $\|\boldsymbol{q}_i^{(r)} - \mathbf{v}_{\mathrm{U},1}\|^2,\ \forall i \in \mathcal{N}_{\mathrm{R}}$, which is given by

$$\begin{aligned} \widetilde{C}_{\mathrm{fr},i}(\boldsymbol{q}_i) &\ge -E_i^{(r)}\left(\|\boldsymbol{q}_i - \mathbf{v}_{\mathrm{U},1}\|^2 - \|\boldsymbol{q}_i^{(r)} - \mathbf{v}_{\mathrm{U},1}\|^2\right) + F_i^{(r)} \\ &\triangleq \widetilde{C}_{\mathrm{fr},i}^{\mathrm{lb},1}(\boldsymbol{q}_i), \end{aligned} \tag{6.28}$$

where

$$E_i^{(r)} = \frac{p_{\mathrm{C},i}\log_2 e}{\frac{\sigma_{\mathrm{R}}^2}{\xi}\|\boldsymbol{q}_i^{(r)} - \mathbf{v}_{\mathrm{U},1}\|^4 + p_{\mathrm{C},i}\|\boldsymbol{q}_i^{(r)} - \mathbf{v}_{\mathrm{U},1}\|^2} \tag{6.29}$$

and

$$F_i^{(r)} = \log_2\left(1 + \frac{\xi p_{\mathrm{C},i}}{\sigma_{\mathrm{R}}^2\|\boldsymbol{q}_i^{(r)} - \mathbf{v}_{\mathrm{U},1}\|^2}\right) \tag{6.30}$$

are constants. Similarly, $\|\boldsymbol{q}_i - \mathbf{v}_{\mathrm{U},k}\|^2$ is convex with respect to $\boldsymbol{q}_i$, hence constraint (6.22b) is non-convex. By applying the first-order Taylor expansion, the lower bound of $\|\boldsymbol{q}_i - \mathbf{v}_{\mathrm{U},k}\|^2$ at $\boldsymbol{q}_i^{(r)}$ is given by

$$\|\boldsymbol{q}_i - \mathbf{v}_{\mathrm{U},k}\|^2 \ge \|\boldsymbol{q}_i^{(r)} - \mathbf{v}_{\mathrm{U},k}\|^2 + 2(\boldsymbol{q}_i^{(r)} - \mathbf{v}_{\mathrm{U},k})^T(\boldsymbol{q}_i - \boldsymbol{q}_i^{(r)}). \tag{6.31}$$

Also, we can get the lower bound for the $\|\boldsymbol{q}_i - \boldsymbol{q}_k\|^2$ in constraint (6.20d) as

$$\|\boldsymbol{q}_i - \boldsymbol{q}_k\|^2 \geq -\|\boldsymbol{q}_i^{(r)} - \boldsymbol{q}_k^{(r)}\|^2 + 2(\boldsymbol{q}_i^{(r)} - \boldsymbol{q}_k^{(r)})^T(\boldsymbol{q}_i - \boldsymbol{q}_k). \tag{6.32}$$

Now, using (6.23)-(6.32), problem (6.20) can be reformulated as

$$\max_{\mathbf{Q},\mathbf{S},\mathcal{C}} \ \min_{k} \ R_{k,1}^{\text{lb}}(\mathbf{Q}) - R_{k,2}(\mathbf{S}) \tag{6.33a}$$

$$\text{s.t.} \ \ S_{ik} \leq \|\boldsymbol{q}_i^{(r)} - \mathbf{v}_{\text{U},k}\|^2 + 2\big(\boldsymbol{q}_i^{(r)} - \mathbf{v}_{\text{U},k}\big)^T \times \big(\boldsymbol{q}_i - \boldsymbol{q}_i^{(r)}\big), \ \forall i \in \mathcal{N}_\text{R}, k \in \mathcal{N}_\text{U}, \tag{6.33b}$$

$$C_i \leq \widetilde{C}_{\text{fr},i}^{\text{lb},1}(\boldsymbol{q}_i), \ \forall i \in \mathcal{N}_\text{R}, \tag{6.33c}$$

$$d_{\min}^2 \leq -\|\boldsymbol{q}_j^{(r)} - \boldsymbol{q}_k^{(r)}\|^2 + 2\big(\boldsymbol{q}_j^{(r)} - \boldsymbol{q}_k^{(r)}\big)^T \times \big(\boldsymbol{q}_j - \boldsymbol{q}_k\big), \ \forall j,k \in \mathcal{N}_\text{R}, \ j \neq k, \tag{6.33d}$$

(6.20c),

which is a convex problem and can thus be solved iteratively using a standard convex optimization toolbox such as CVX [91].

6.3.2 Sub-Problem 2: Power control and quantization noise variance optimization

Now, for any given UAV placement $\mathbf{Q}$, we aim to optimize the power coefficients $\mathbf{p}_\text{C}, \mathbf{p}_\text{R}$ and quantization noise $\boldsymbol{\Omega}$. From problem (6.19), we can obtain sub-problem 2 as follows

$$\max_{\mathbf{p}_\text{C},\mathbf{p}_\text{R},\boldsymbol{\Omega},\mathcal{C}} \ \min_{k} \ R_{k,1}(\mathbf{p}_\text{R},\boldsymbol{\Omega}) - R_{k,2}(\mathbf{p}_\text{R},\boldsymbol{\Omega}), \ \forall k \in \mathcal{N}_\text{U}, \tag{6.34a}$$

$$\text{s.t.} \ \ C_i \leq C_{\text{fr},i}(\mathbf{p}_\text{C}), \ \forall i \in \mathcal{N}_\text{R}, \tag{6.34b}$$

$$C_i \geq \varphi_i(p_{\text{R},i}, \Omega_i), \ \forall i \in \mathcal{N}_\text{R}, \tag{6.34c}$$

$$p_{\text{R},i} + \Omega_i \leq P_{\text{R},i}^{\max}, \ \forall i \in \mathcal{N}_\text{R}, \tag{6.34d}$$

$$\text{tr}\left\{\text{diag}\left(p_{\text{C},1}, \ldots, p_{\text{C},N_\text{R}}\right)\mathbf{W}\mathbf{W}^H\right\} \leq P_\text{C}^{\max}, \tag{6.34e}$$

$$p_{\text{C},i} \geq 0, \ p_{\text{R},i} \geq 0, \ \Omega_i \geq 0, \ \forall i \in \mathcal{N}_\text{R}. \tag{6.34f}$$

Problem (6.34) is hard to solve due to the non-convexity of the objective function (6.34a) and constraints (6.34b) and (6.34c). Let us define $\mathbf{p}_{\mathrm{C}}^{(r)}$ and $\mathbf{p}_{\mathrm{R}}^{(r)}$ as the given power coefficients from the $(r-1)$-th iteration where $\mathbf{p}_{\mathrm{C}}^{(r)} = \left[p_{\mathrm{C},1}^{(r)}, \ldots, p_{\mathrm{C},N}^{(r)}\right]^T$, $\mathbf{\Omega}^{(r)} = \left[\Omega_1^{(r)}, \ldots, \Omega_N^{(r)}\right]$, and $\mathbf{p}_{\mathrm{R}}^{(r)} = \left[p_{\mathrm{R},1}^{(r)}, \ldots, p_{\mathrm{R},N}^{(r)}\right]^T$. For the objective function (6.34a), we can derive an upper bound for the concave function $R_{k,2}\left(\mathbf{p}_{\mathrm{R}}, \mathbf{\Omega}\right)$ at a given point $(\mathbf{p}_{\mathrm{R}}^{(r)}, \mathbf{\Omega}^{(r)})$, which is given by

$$\begin{aligned} R_{k,2}(\mathbf{p}_{\mathrm{R}}, \mathbf{\Omega}) &\leq \sum_{i \in \mathcal{N}_{\mathrm{R}} \setminus \{k\}} I_{k,i}^{(r)}(p_{\mathrm{R},i} - p_{\mathrm{R},i}^{(r)}) + \sum_{i \in \mathcal{N}_{\mathrm{R}}} I_{k,i}^{(r)}(\Omega_i - \Omega_i^{(r)}) + J_k^{(r)} \\ &\triangleq R_{k,2}^{\mathrm{ub}}\left(\mathbf{p}_{\mathrm{R}}, \mathbf{\Omega}\right), \ \forall k \in \mathcal{N}_{\mathrm{U}}, \end{aligned} \tag{6.35}$$

where

$$I_{k,i}^{(r)} = \frac{|g_{ik}|^2 \log_2 e}{\sum\limits_{j \in \mathcal{N}_{\mathrm{R}} \setminus \{k\}} p_{\mathrm{R},j}^{(r)} |g_{jk}|^2 + \sum\limits_{j \in \mathcal{N}_{\mathrm{R}}} \Omega_j^{(r)} |g_{jk}|^2 + \sigma_{\mathrm{U}}^2} \tag{6.36}$$

and

$$J_k^{(r)} = \log_2 \Big(\sum_{j \in \mathcal{N}_{\mathrm{R}} \setminus \{k\}} p_{\mathrm{R},j}^{(r)} |g_{jk}|^2 + \sum_{j \in \mathcal{N}_{\mathrm{R}}} \Omega_j^{(r)} |g_{jk}|^2 + \sigma_{\mathrm{U}}^2 \Big) \tag{6.37}$$

are constants. Similarly, the right-hand-side of constraint (6.34c) has the following upper bound at the point $(\mathbf{p}_{\mathrm{R}}^{(r)}, \mathbf{\Omega}^{(r)})$

$$\begin{aligned} \varphi_i(p_{\mathrm{R},i}, \Omega_i) &\leq \frac{\log_2 e}{p_{\mathrm{R},i}^{(r)} + \Omega_i^{(r)}}(p_{\mathrm{R},i} + \Omega_i) + \log_2 \left(p_{\mathrm{R},i}^{(r)} + \Omega_i^{(r)}\right) - \log_2 e - \log_2 \left(\Omega_i\right) \\ &\triangleq \varphi_i^{\mathrm{ub}}(p_{\mathrm{R},i}). \end{aligned} \tag{6.38}$$

For constraint (6.34b), the lower bound for $C_{\mathrm{fr},i}(\mathbf{p}_{\mathrm{C}})$ with respect to $\mathbf{p}_{\mathrm{C}}$ using the first-order Taylor expansion at $\mathbf{p}_{\mathrm{C}}^{(r)}$ is

$$\begin{aligned} C_{\mathrm{fr},i}(\mathbf{p}_{\mathrm{C}}) &\geq \log_2 \Big(\sum_{j \in \mathcal{N}_{\mathrm{R}}} \left|\mathbf{h}_i^T \mathbf{w}_j\right|^2 p_{\mathrm{C},j} + \sigma_{\mathrm{R}}^2 \Big) - \sum_{k \in \mathcal{N}_{\mathrm{R}} \setminus \{i\}} X_{i,k}^{(r)}(p_{\mathrm{C},k} - p_{\mathrm{C},k}^{(r)}) - Y_i^{(r)} \\ &\triangleq C_{\mathrm{fr},i}^{\mathrm{lb},2}(\mathbf{p}_{\mathrm{C}}), \end{aligned} \tag{6.39}$$

where

$$X_{i,k}^{(r)} = \frac{\left|\mathbf{h}_i^T \mathbf{w}_k\right|^2 \log_2 e}{\sum_{j \in \mathcal{N}_{\mathrm{R}} \setminus \{k\}} \left|\mathbf{h}_i^T \mathbf{w}_j\right|^2 p_{\mathrm{C},j}^{(r)} + \sigma_{\mathrm{R}}^2} \tag{6.40}$$

and

$$Y_i^{(r)} = \log_2 \left(\sum_{j \in \mathcal{N}_\mathrm{R} \setminus \{i\}} \left| \mathbf{h}_i^T \mathbf{w}_j \right|^2 p_{\mathrm{C},j}^{(r)} + \sigma_\mathrm{R}^2 \right). \tag{6.41}$$

Therefore, using (6.35)-(6.39), problem (6.34) can be recast as

$$\max_{\mathbf{p}_\mathrm{C}, \mathbf{p}_\mathrm{R}, \mathbf{\Omega}, \mathcal{C}} \quad \min_k \; R_{k,1}\left(\mathbf{p}_\mathrm{R}, \mathbf{\Omega}\right) - R_{k,2}^{\mathrm{ub}}\left(\mathbf{p}_\mathrm{R}, \mathbf{\Omega}\right), \; \forall k \in \mathcal{N}_\mathrm{U}, \tag{6.42a}$$

$$\text{s.t.} \quad C_i \leq C_{\mathrm{fr},i}^{\mathrm{lb},2}(\mathbf{p}_\mathrm{C}), \; \forall i \in \mathcal{N}_\mathrm{R}, \tag{6.42b}$$

$$C_i \geq \varphi_i^{\mathrm{ub}}(p_{\mathrm{R},i}, \Omega_i), \; \forall i \in \mathcal{N}_\mathrm{R}, \tag{6.42c}$$

$$(6.34\mathrm{d}), \; (6.34\mathrm{e}), \; (6.34\mathrm{f}),$$

which is a convex problem and hence can be solved iteratively using a standard convex optimization toolbox such as CVX.

6.3.3 Iterative Algorithm

Similar to [56], we adopt the BCD algorithm and the overall solution for problem (6.19) is given in Algorithm 10. It should be noted that the convergence of Algorithm 10 is guaranteed [56]. Note that in sub-problem 1, we approximate (6.13) by (6.27) to simplify constraint (6.19b), and as a result, the obtained optimal solution for sub-problem 1 may not meet all the conditions in the original problem (6.19). However, a tighter function (6.39) for the same constraint is used in sub-problem 2 and thus making sure any ultimate solutions are all feasible for problem (6.19).

6.4 Numerical Results

In this section, we assess the performance of the proposed algorithm. We assume that the UAVs are deployed in a rural environment and the UEs are randomly and uniformly distributed within a square area of 1 km^2 where the distance between its center and the

Algorithm 10. Joint design based iterative sum-rate maximization for downlink transmission in UAV-enabled Massive MIMO CRANs.

1: **Input:** Essential system parameters and a feasible point $\{\mathbf{Q}^{(0)}, \mathbf{\Omega}^{(0)}, \mathbf{p}_{\mathrm{C}}^{(0)}, \mathbf{p}_{\mathrm{R}}^{(0)}\}$;
2: **Initialization:** Set $r \coloneqq 0$;
3: **repeat**
4: Update $r \coloneqq r + 1$;
5: Solve problem (6.33) for given $\{\mathbf{Q}^{(r-1)}, \mathbf{p}_{\mathrm{C}}^{(r-1)}, \mathbf{p}_{\mathrm{R}}^{(r-1)}, \mathbf{\Omega}^{(r-1)}\}$, and denote the optimal solution as $\{\mathbf{Q}^{(r)}\}$;
6: Solve problem (6.42) for given $\{\mathbf{Q}^{(r)}, \mathbf{\Omega}^{(r-1)}, \mathbf{p}_{\mathrm{C}}^{(r-1)}, \mathbf{p}_{\mathrm{R}}^{(r-1)}\}$, and denote the optimal solution as $\{\mathbf{\Omega}^{(r)}, \mathbf{p}_{\mathrm{C}}^{(r)}, \mathbf{p}_{\mathrm{R}}^{(r)}\}$;
7: **until** convergence;
8: **Output:** $\{\mathbf{Q}^{\star}, \mathbf{\Omega}^{\star}, \mathbf{p}_{\mathrm{C}}^{\star}, \mathbf{p}_{\mathrm{R}}^{\star}\} \coloneqq \{\mathbf{Q}^{(r)}, \mathbf{\Omega}^{(r)}, \mathbf{p}_{\mathrm{C}}^{(r)}, \mathbf{p}_{\mathrm{R}}^{(r)}\}$.

Table 6.1: Parameter settings for the simulation.

Parameters	Description	Value
H_{C}	Altitude of CU	30 m
H_{R}	Altitude of UAVs	60 m
ϕ	Tilted angle of CU antenna arrays	0°
λ	Wavelength	0.5 m
δ	Antenna spacing at CU	0.25 m
σ_{R}^2, σ_{U}^2	Noise level in both Links	-100 dBm
ξ	Channel gain at 1m for both links	-40 dB
$P_{\mathrm{C}}^{\max}$	Transmit power for CU	1 W
$P_{\mathrm{R},i}^{\max}$, $\forall i \in \mathcal{N}_{\mathrm{R}}$	Transmit power for each UAV	0.1 W
$d_{\min}$	Minimum safe distance between UAVs	10 m

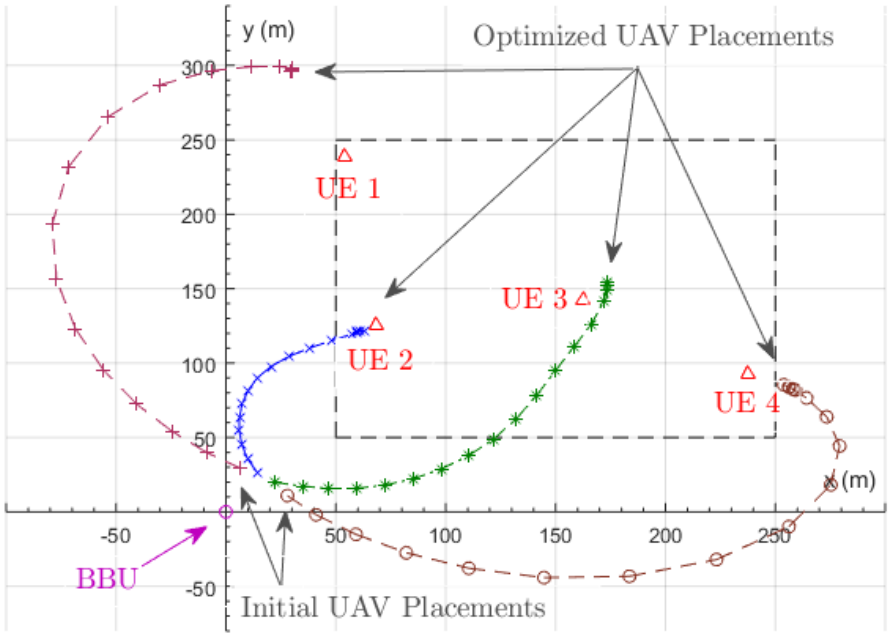


Figure 6.2: Path of UAV optimized positions in the proposed algorithms.

CU is denoted by D. Unless otherwise specified, we consider the following simulation parameters are set as shown in Table 6.1.

Figure 6.2 shows the trajectory of the UAV placement optimized with iterations. It is observed that the steps are getting smaller with each iteration and eventually become stationary where the best placement is achieved.

Figure 6.3 investigates the convergence of the proposed algorithm and illustrates the minimum rate against the iteration number for difference values of D. We consider the scenario of $N = 5$ UAV-UE pairs and ZF precoding at the CU. For comparison purposes, we consider two benchmark schemes. In the first scheme, referred to as Scheme I, only the placement of UAVs is optimized; and in the second scheme, referred to as Scheme II, only the power control at the CU and RRH is optimized. We can clearly see that the proposed algorithm significantly outperforms the two benchmark schemes. Also, we notice that as the distance between the CU and the served area, D, increases from 2000m to 4000m the convergence speeds and achieved minimum rates of the algorithms decrease. We also observe that the proposed algorithm and Scheme I have approximately the same

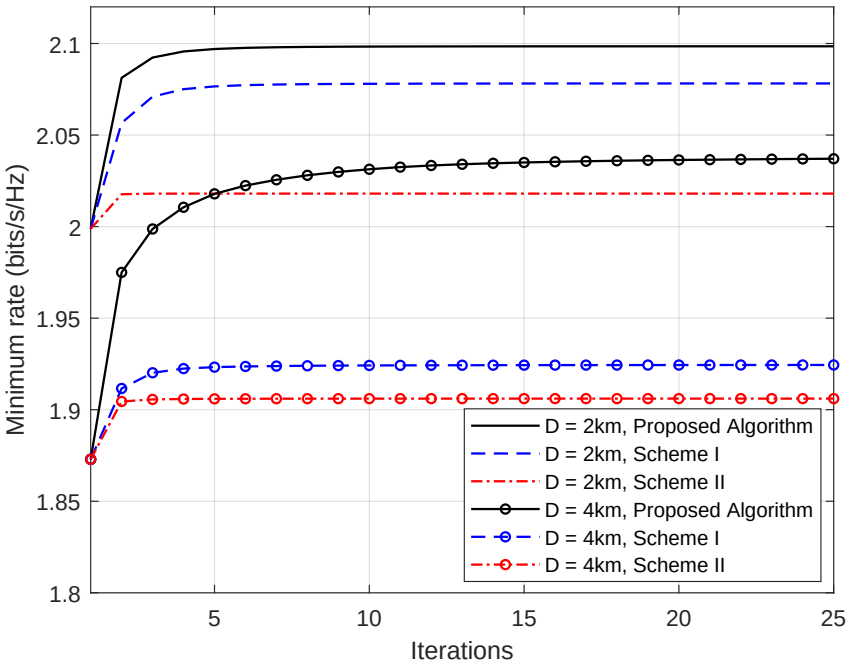


Figure 6.3: Convergence speed of the proposed algorithm for $N = 5$ UAV-UE pairs with ZF in the fronthaul link.

convergence speed, which is slower than that of Scheme II.

Figure 6.4 demonstrates the impact of the transmit power of UAVs on the minimum rate achieved by each UE for different values of CU tranmit power. We consider $K = 2$ and ZF precoding at the CU. We can see clearly that As the UAV transmit power increases the minimum rate increases. Also, increasing the CU transmit power results in an improved minimum rate. It is important to notice that as the CU transmit power increases the rate of improvement of the minimum rate decreases and this can be explained by the fact that beyond a certain value of the CU tranmit power the access link becomes the bottleneck link.

Figure 6.5 compares the performance of the ZF and MF precoding methods in terms of minimum rate versus the number of UEs, N, for different values of D. It is obvious that ZF consistently performs better than MF in all cases. However, we can see that the minimum rate drops rapidly with the increased number of UEs for both precoding methods. This is because the rate is limited by the interference among UEs in the access link. One solution

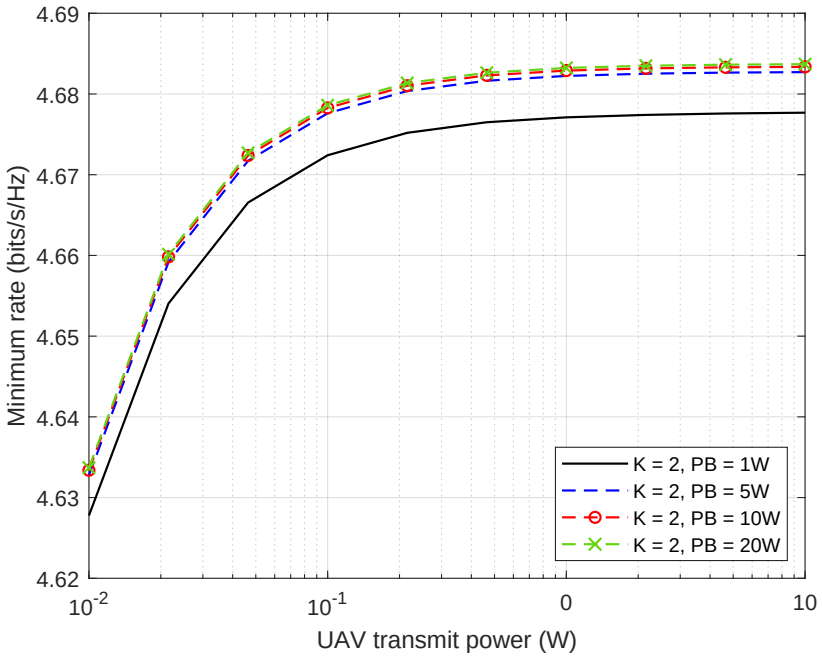


Figure 6.4: Minimum rate versus UAVs transmit power for different values of CU transmit power.

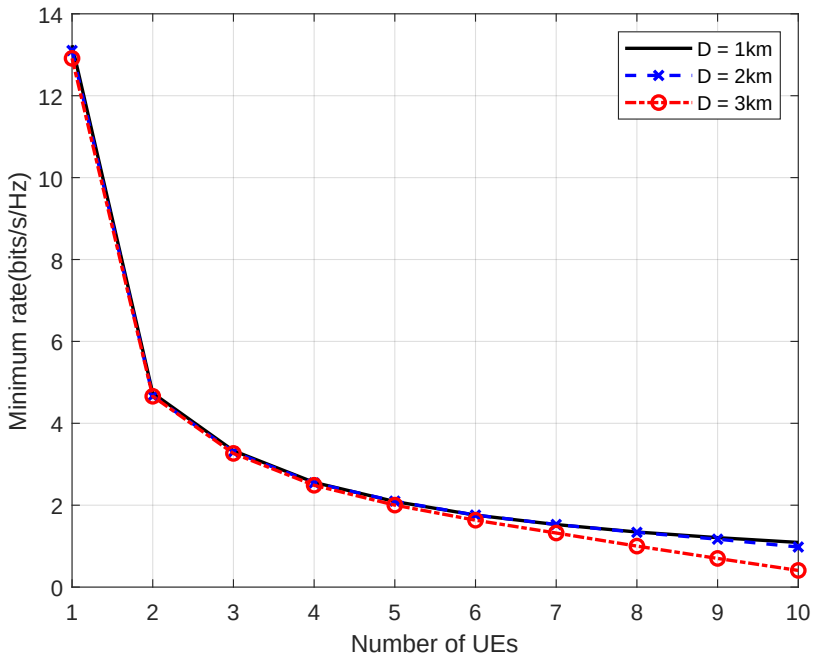


Figure 6.5: Minimum rate versus the number of UEs for ZF and MF for different values of D.

to improve the performance is to combine the use of multiple antennas and precoding in the access link, which is beyond the scope of the current work and is left as an interesting future work.

6.5 Conclusion

In this chapter, we have introduced a novel architecture of multi-UAV-enabled CRANs, which features a massive MIMO fronthaul connection. The proposed architecture employs UAVs as flying RRHs to establish connectivity with the terrestrial UEs. Unlike previous designs presented in Chapters 3 and 4, the difficulties of this system model centered on the placement of UAVs. While positioning UAVs closer to the UEs can result in better access link quality, it can also lead to higher interference levels among different UEs and a reduced capacity of the fronthaul link. Also, the precoding matrices of the access link and quantization noise were considered in the proposed joint designs.

The optimization problem formulated to maximize the minimum rate among all UEs posed significant challenges due to its high complexity. To address this issue, the problem was divided into two sub-problems: one focused on optimizing the UAV placements, while the other aimed to optimize the precoding and compression noise covariance matrices. By transforming both optimization problems into convex SDR problems, an algorithm based on BCD and SCA methods was proposed to solve both problems iteratively.

To illustrate the optimization process, numerical results were provided in this chapter, including the path change of UAVs' placement and the achievable minimum rate within each iteration. It can be observed that our proposed design can significantly increase the minimum achievable rate compared to the two benchmark schemes. Furthermore, the impact of the number of UEs and the transmit power of UAVs on the achievable rate was also investigated.

Chapter 7

Conclusion and Future Work

In this thesis, we introduced innovative architectures and optimization designs for CRAN with massive MIMO technology enabled in the wireless fronthaul link. The proposed joint cooperative approach for the wireless fronthaul and access links was developed for both downlink and uplink transmissions. Specifically, this thesis addressed the challenges of optimizing varying capacity constraints of the wireless fronthaul link while considering multivariate compression noise and interference that naturally exists in wireless channels. Furthermore, this thesis presented other designs with lower computational complexity and compared them to the joint design. These included the independent point-to-point compression strategy and separate optimization designs for wireless fronthaul and access links. To summarize this thesis, we now conclude the primary contributions and outcomes of each chapter.

7.1 Conclusion

In Chapter 1, we first provided a concise introduction to the developmental history of RAN and proceeded to present ongoing research in cloud-based RAN architecture, alongside its potential future applications. Given the increasing attention to CRAN architecture

in recent years, we focused our research on addressing several relevant questions and challenges. Now, at the end of this thesis, we aim to address those challenges posed in Chapter 1, based on our research in this thesis:

- **Limited Wireless Fronthaul Capacity:** Our proposed designs extensively utilized massive MIMO technology at the fronthaul link. Numerical results from our comprehensive analysis have demonstrated that the deployment of large-scale antenna arrays at the CU can substantially enhance fronthaul capacity without requiring an increase in total transmit power. It also showed the future potential to increase data throughput for the access link by employing more transmitting antennas at the RRHs.
- **Fast and Low-cost Deployment and Mobility:** In our proposed models, we considered the least functionality left at the RRHs, thereby making them lighter and less expensive, while fully centralizing most of the protocols at the CU. The wireless connectivity of the fronthaul link is of utmost importance in our designs, as it serves as a crucial factor in enhancing the mobility and portability of RRHs.
- **Imperfect Channel State Information:** We have thoroughly considered both perfect instantaneous CSI and stochastic CSI, and also proposed corresponding designs in order to maximize the sum-rate and minimize the total power consumption for both scenarios. The lack of CSI knowledge led to the difficulty in designing precoding matrices, but through our proposed designs, it was still possible to increase the network ergodic sum-rate across all coherence blocks.
- **Support for Aerial Communication:** We have proposed a novel model by employing UAVs as the flying RRHs embedded in CRAN architecture with a massive MIMO fronthaul connection. UAVs' high mobility and rapid deployment make them valuable partners for terrestrial communication networks, as they provide air-to-ground LoS channels and extensive coverage. Moreover, our work has demonstrated that jointly optimizing the placement of UAVs, precoding matrices, and compression schemes can significantly increase the achievable data throughput of the whole network.

In Chapter 2, we presented an overview of the necessary background information on the relevant technologies and concepts used in this thesis. This was followed by an extensive review of the latest research in the field, providing a comprehensive understanding of the current state-of-the-art.

In Chapter 3, we focused on the downlink transmission of the CRAN architecture and proposed massive MIMO employed for enhancing the fronthaul networks. Perfect instantaneous CSI of the access link was assumed for designing quantization schemes, precoding matrices, and transmit power allocation. Numerical simulations are conducted to demonstrate the superiority of the proposed designs and algorithms compared to the benchmarks. In this chapter, we have thoroughly discussed and compared various solvers in terms of their performance on the convex problems we proposed. Specifically, we evaluated their feasibility, computational complexity, and convergence behavior. Numerical results showed that employing joint compression and adequate number of antennas at the CU can lead to a significant increase in the achievable sum-rate and a superior reduction in power consumption.

In Chapter 4, we concentrated on the scenario where only stochastic CSI is available at the CU. Adopting the same architecture as used in Chapter 3, we aimed to optimize system SE and EE by considering ergodic rates as the objective functions and constraints, respectively. Through several simulations, we conducted a thorough analysis of the impact of key system parameters on the achievable SE and EE. These parameters included the number of antennas at the CU, the number of UEs, and power budgets at both the CU and RRHs. Compared to the system model with the assumption of perfect instantaneous CSI, as studied in Chapter 3, the lack of CSI knowledge results in a considerable decrease in the achievable sum-rates, as expected. However, through the joint design of both links, it is possible to efficiently increase the ergodic sum-rates and decrease the required power consumption. Furthermore, the increasing the number of antennas at the CU brings more improvement.

In Chapter 5, we considered the uplink transmission in a massive MIMO CRAN with

the CF scheme implemented at the RRHs to alleviate the burden of the fronthaul link. Two conventional strategies, namely ZF and MRC, were utilized in the beamforming designs for the fronthaul link, where ZF has generally outperformed MRC, as expected. The impact of different numbers of antennas at the CU was investigated and demonstrated that massive MIMO has the potential to significantly increase the achievable sum-rate.

In Chapter 6, we proposed to use UAVs as flying RRHs for establishing the connection between the UEs and a CU equipped with a large-scale antenna array. The placement of UAVs, quantization noise matrices, and transmit power allocation were jointly optimized aiming to maximize the minimum rate among all UEs. The proposed algorithm divided the resulting highly complex optimization problem into two sub-problems and demonstrated superior performance in terms of fast convergence speed and increased achievable rates in numerical simulations.

7.2 Future Work

Despite the significant progress we have made in this thesis, there are still several potential scenarios that can be extended for practical applications. These extensions can be explored in our future work.

- **Estimation of CSI:** The lack of CSI knowledge leads to a wide gap between the achievable rate and theoretical maximum capacity. To improve system SE and EE, some traditional channel estimation methods can be implemented such as *minimum mean square error* (MMSE). This can be achieved by having UEs send training pilots to the RRHs in the uplink transmission [101]. The downlink channels can be estimated based on channel reciprocity in TDD mode. However, the received pilots at the RRHs must be compressed and forwarded to the CU for processing and computation in order to achieve cooperation among all RRHs and mitigate interference in the access link. This can lead to difficulties in system design and requires further investigation into estimation error, quantization error, and resource

allocation.

- **Hybrid Terrestrial and Non-terrestrial CRANs:** We have studied both terrestrial communications for uplink and downlink transmissions in Chapters 3-5, and also non-terrestrial communication with multi-UAV-enabled CRANs in Chapter 6. To further increase the communication network performance, the large-scale cooperation enabled by hybrid communication networks is of practical interest for effectively serving ground UEs distributed over a wide-range area. It can be achieved by leveraging the capability of high-performance CU in massive MIMO CRANs. However, designing a dynamic activation scheme to select the most suitable servers (terrestrial RRHs and/or UAVs) for each UE is a challenging task that needs further investigation in order to mitigate interference among all units to further increase system SE and EE.

Appendix

In this appendix, we prove Theorem 3.3.1. By following the framework used in [57, Chapter 9], the Lagrangian of problem (3.25) is given by

$$\begin{aligned}\mathcal{L} =& \sum\nolimits_{j=1}^{N_{\mathrm{U}}} \mathrm{tr}\left(\bar{\mathbf{U}}_j\right) + \mathrm{tr}\left(\mathbf{Q}\right) + \sum\nolimits_{i=1}^{N_{\mathrm{R}}} p_i - \sum\nolimits_{j=1}^{N_{\mathrm{U}}} \mathrm{tr}\left(\mathbf{F}_j \bar{\mathbf{U}}_j\right) - \mathrm{tr}\left(\mathbf{T}\mathbf{Q}\right) \\ &+ \sum\nolimits_{m\in\mathcal{N}_{\mathcal{S}}} \eta_m \left(\varphi_m^{\mathrm{ub}}(\{\bar{\mathbf{U}}_j\}, \mathbf{Q}) - \sum\nolimits_{i\in\mathcal{S}_m} C_{\mathrm{fr},i}^{\mathrm{lb}}(\mathbf{p})\right) \\ &- \sum\nolimits_{k=1}^{N_{\mathrm{U}}} \lambda_k \left(C_{ac,k}^{\mathrm{lb}}(\{\bar{\mathbf{U}}_j\}, \mathbf{Q}) - \gamma_k\right) + \theta \left(\sum\nolimits_{i=1}^{N_{\mathrm{R}}} p_i - P_C\right) \\ &+ \sum\nolimits_{i=1}^{N_{\mathrm{R}}} \rho_i \left(\sum\nolimits_{j=1}^{N_{\mathrm{U}}} \mathrm{tr}\left(\mathbf{\Gamma}_i^H \bar{\mathbf{U}}_j \mathbf{\Gamma}_i\right) + \mathrm{tr}\left(\mathbf{Q}_{ii}\right) - P_{R,i}\right), \end{aligned} \tag{A.1}$$

where $\{\eta_m\}$, $\{\lambda_k\}$, θ, and $\{\rho_i\}$, $\{\mathbf{F}_j\}$, and $\mathbf{T}$ are the Lagrange multipliers corresponding to constraints of problem (3.25), $\{\bar{\mathbf{U}}_j\} \succeq \mathbf{0}$, and $\mathbf{Q} \succeq \mathbf{0}$, respectively. Since we are interested in the rank of $\bar{\mathbf{U}}_j$, we can reexpress the Lagrangian as a function of $\bar{\mathbf{U}}_j$ and relevant Lagrange multipliers as follows

$$\begin{aligned}\mathcal{L} =& \sum_{j=1}^{N_{\mathrm{U}}} \mathrm{tr}\left(\bar{\mathbf{U}}_j\right) - \sum_{j=1}^{N_{\mathrm{U}}} \mathrm{tr}\left(\mathbf{F}_j \bar{\mathbf{U}}_j\right) + \sum_{i=1}^{N_{\mathrm{R}}} \sum_{j=1}^{N_{\mathrm{U}}} \rho_i \mathrm{tr}\left(\mathbf{\Gamma}_i^H \bar{\mathbf{U}}_j \mathbf{\Gamma}_i\right) + \Psi \\ &+ \sum_{m\in\mathcal{N}_{\mathcal{S}}} \sum_{i\in\mathcal{S}_m} \frac{\eta_m}{\ln(2)} \times \mathrm{tr}\left(\left(\sum_{j=1}^{N_{\mathrm{U}}} \mathbf{\Gamma}_i^H \tilde{\bar{\mathbf{U}}}_j \mathbf{\Gamma}_i + \tilde{\mathbf{Q}}_{ii}\right)^{-1} \sum_{j=1}^{N_{\mathrm{U}}} \mathbf{\Gamma}_i^H \bar{\mathbf{U}}_j \mathbf{\Gamma}_i\right) \\ &+ \frac{1}{\ln(2)} \mathrm{tr}\left(\left(\mathbf{G}_k \left(\sum_{j\neq k}^{N_{\mathrm{U}}} \tilde{\bar{\mathbf{U}}}_j + \tilde{\mathbf{Q}}\right) \mathbf{G}_k^H + \sigma_U^2 \mathbf{I}\right)^{-1} \times \sum_{j\neq k}^{N_{\mathrm{U}}} \mathbf{G}_k \bar{\mathbf{U}}_j \mathbf{G}_k^H\right)\right) \\ &- \sum_{k=1}^{N_{\mathrm{U}}} \lambda_k \left(\log\det\left(\mathbf{G}_k \left(\sum_{j=1}^{N_{\mathrm{U}}} \bar{\mathbf{U}}_j + \mathbf{Q}\right) \mathbf{G}_k^H + \sigma_U^2 \mathbf{I}\right)\right), \end{aligned} \tag{A.2}$$

where Ψ contains all the remaining terms of the Lagrangian in (A.1) that are independent of $\bar{\mathbf{U}}_j$. Since problem (3.25) is convex and satisfies Slater's condition, Karush-Kuhn-

Tucker (KKT) conditions provide necessary and sufficient conditions for optimality. The KKT conditions relevant to our derivation are

$$\frac{\partial \mathcal{L}}{\partial \bar{\mathbf{U}}_\ell^\star} = \mathbf{0},\ \ell \in N_{\mathrm{U}}, \tag{A.3}$$

$$\mathbf{F}_\ell^\star \bar{\mathbf{U}}_\ell = \mathbf{0},\ \ell \in N_{\mathrm{U}}. \tag{A.4}$$

From (A.3), we have

$$\begin{aligned}
\frac{\partial \mathcal{L}}{\partial \bar{\mathbf{U}}_\ell} &= \mathbf{I} + \sum\nolimits_{i=1}^{N_{\mathrm{R}}} \rho_i \mathbf{\Gamma}_i^* \mathbf{\Gamma}_i^T - \mathbf{F}_\ell^T \\
&+ \sum\nolimits_{m \in \mathcal{N}_{\mathcal{S}}} \sum\nolimits_{i \in \mathcal{S}_m} \frac{\eta_m}{\ln(2)} \mathbf{\Gamma}_i^* \left(\sum\nolimits_{j=1}^{N_{\mathrm{U}}} \mathbf{\Gamma}_i^H \tilde{\bar{\mathbf{U}}}_j \mathbf{\Gamma}_i + \tilde{\mathbf{Q}}_{ii} \right)^{-T} \mathbf{\Gamma}_i^T \\
&- \sum\nolimits_{k=1}^{N_{\mathrm{U}}} \frac{\lambda_k}{\ln(2)} \mathbf{G}_k^T \left(\mathbf{G}_k \left(\sum\nolimits_{j=1}^{N_{\mathrm{U}}} \bar{\mathbf{U}}_j + \mathbf{Q} \right) \mathbf{G}_k^H + \sigma_U^2 \mathbf{I} \right)^{-T} \mathbf{G}_k^* \\
&+ \sum\nolimits_{k \neq \ell}^{N_{\mathrm{U}}} \frac{\lambda_k}{\ln(2)} \mathbf{G}_k^T \left(\mathbf{G}_k \left(\sum\nolimits_{j \neq k}^{N_{\mathrm{U}}} \tilde{\bar{\mathbf{U}}}_j + \tilde{\mathbf{Q}} \right) \mathbf{G}_k^H + \sigma_U^2 \mathbf{I} \right)^{-T} \mathbf{G}_k^* \\
&= 0.
\end{aligned} \tag{A.5}$$

Hence

$$\begin{aligned}
\mathbf{F}_\ell^\star &= \mathbf{I} + \sum\nolimits_{i=1}^{N_{\mathrm{R}}} \rho_i^\star \mathbf{\Gamma}_i \mathbf{\Gamma}_i^H \\
&+ \sum\nolimits_{m \in \mathcal{N}_{\mathcal{S}}} \sum\nolimits_{i \in \mathcal{S}_m} \frac{\eta_m^\star}{\ln(2)} \mathbf{\Gamma}_i \left(\sum\nolimits_{j=1}^{N_{\mathrm{U}}} \mathbf{\Gamma}_i^H \tilde{\bar{\mathbf{U}}}_j \mathbf{\Gamma}_i + \tilde{\mathbf{Q}}_{ii} \right)^{-1} \mathbf{\Gamma}_i^H \\
&+ \sum\nolimits_{k \neq \ell}^{N_{\mathrm{U}}} \frac{\lambda_k^\star}{\ln(2)} \mathbf{G}_k^H \left(\mathbf{G}_k \left(\sum\nolimits_{j \neq k}^{N_{\mathrm{U}}} \tilde{\bar{\mathbf{U}}}_j + \tilde{\mathbf{Q}} \right) \mathbf{G}_k^H + \sigma_U^2 \mathbf{I} \right)^{-1} \mathbf{G}_k \\
&- \sum\nolimits_{k=1}^{N_{\mathrm{U}}} \frac{\lambda_k^\star}{\ln(2)} \mathbf{G}_k^H \left(\mathbf{G}_k \left(\sum\nolimits_{j=1}^{N_{\mathrm{U}}} \bar{\mathbf{U}}_j^\star + \mathbf{Q}^\star \right) \mathbf{G}_k^H + \sigma_U^2 \mathbf{I} \right)^{-1} \mathbf{G}_k.
\end{aligned} \tag{A.6}$$

Let

$$\mathbf{\Phi}_k = \mathbf{G}_k \left(\sum\nolimits_{j \neq k}^{N_{\mathrm{U}}} \tilde{\bar{\mathbf{U}}}_j + \tilde{\mathbf{Q}} \right) \mathbf{G}_k^H + \sigma_U^2 \mathbf{I} \tag{A.7}$$

$$\mathbf{\Omega}_k = \mathbf{G}_k \left(\sum\nolimits_{j=1}^{N_{\mathrm{U}}} \bar{\mathbf{U}}_j^\star + \mathbf{Q}^\star \right) \mathbf{G}_k^H + \sigma_U^2 \mathbf{I}. \tag{A.8}$$

Hence, we have

$$\begin{aligned}\mathbf{F}_\ell =&\mathbf{I}+\sum\nolimits_{i=1}^{N_{\mathrm{R}}}\rho_i^\star\mathbf{\Gamma}_i\mathbf{\Gamma}_i^H+\sum\nolimits_{k\neq\ell}^{N_{\mathrm{U}}}\frac{\lambda_k^\star}{\ln(2)}\mathbf{G}_k^H\left(\mathbf{\Phi}_k^{-1}-\mathbf{\Omega}_k^{-1}\right)\mathbf{G}_k\\&+\sum\nolimits_{m\in\mathcal{N}_{\mathcal{S}}}\sum\nolimits_{i\in\mathcal{S}_m}\frac{\eta_m^\star}{\ln(2)}\mathbf{\Gamma}_i\left(\sum\nolimits_{j=1}^{N_{\mathrm{U}}}\mathbf{\Gamma}_i^H\tilde{\bar{\mathbf{U}}}_j\mathbf{\Gamma}_i+\tilde{\mathbf{Q}}_{ii}\right)^{-1}\mathbf{\Gamma}_i^H\\&-\frac{\lambda_\ell^\star}{\ln(2)}\mathbf{G}_\ell^H\mathbf{\Omega}_\ell^{-1}\mathbf{G}_\ell.\end{aligned}\tag{A.9}$$

As shown in [102, Theorem 1], the MM algorithm is guaranteed to converge to a stationary point of the original non-convex problem. Therefore, after convergence, we have: $\tilde{\bar{\mathbf{U}}}_j = \bar{\mathbf{U}}_j^\star$ and $\tilde{\mathbf{Q}} = \mathbf{Q}^\star$. Hence, from (A.7) and (A.8), we obtain $\mathbf{\Omega}_k = \mathbf{\Phi}_k + \mathbf{G}_k\bar{\mathbf{U}}_k\mathbf{G}_k^H$. Since $\bar{\mathbf{U}}_k \succeq \mathbf{0}$, then $\mathbf{G}_k\bar{\mathbf{U}}_k\mathbf{G}_k^H \succeq \mathbf{0}$ leading to $\mathbf{\Omega}_k \succeq \mathbf{\Phi}_k$; Consequently, $\mathbf{\Phi}_k^{-1} \succeq \mathbf{\Omega}_k^{-1}$. Since $\lambda_k \geq 0$, this results in $\sum_{k\neq\ell}^{N_{\mathrm{U}}}\frac{\lambda_k^\star}{\ln(2)}\mathbf{G}_k^H\left(\mathbf{\Phi}_k^{-1}-\mathbf{\Omega}_k^{-1}\right)\mathbf{G}_k \succeq \mathbf{0}$. Let

$$\begin{aligned}\mathbf{\Upsilon} =&\mathbf{I}+\sum\nolimits_{i=1}^{N_{\mathrm{R}}}\rho_i^\star\mathbf{\Gamma}_i\mathbf{\Gamma}_i^H+\sum\nolimits_{k\neq\ell}^{N_{\mathrm{U}}}\frac{\lambda_k^\star}{\ln(2)}\mathbf{G}_k^H\left(\mathbf{\Phi}_k^{-1}-\mathbf{\Omega}_k^{-1}\right)\mathbf{G}_k\\&+\sum\nolimits_{m\in\mathcal{N}_{\mathcal{S}}}\sum\nolimits_{i\in\mathcal{S}_m}\frac{\eta_m^\star}{\ln(2)}\mathbf{\Gamma}_i\left(\sum\nolimits_{j=1}^{N_{\mathrm{U}}}\mathbf{\Gamma}_i^H\tilde{\bar{\mathbf{U}}}_j\mathbf{\Gamma}_i+\tilde{\mathbf{Q}}_{ii}\right)^{-1}\mathbf{\Gamma}_i^H.\end{aligned}\tag{A.10}$$

Clearly, $\mathrm{rank}(\mathbf{\Upsilon}) = NN_{\mathrm{R}}$ since $\mathbf{\Upsilon}$ is the sum of an identity matrix of size $NN_{\mathrm{R}} \times NN_{\mathrm{R}}$ and positive semidefinite matrices.

On the other hand, since $\mathrm{rank}(\boldsymbol{P}_\ell) = K$ (full rank) and $\mathrm{rank}(\mathbf{G}_\ell) = \min(K, NN_{\mathrm{R}})$, then $\mathrm{rank}\left(\frac{\lambda_\ell^\star}{\ln(2)}\mathbf{G}_\ell^H\mathbf{\Omega}_\ell^{-1}\mathbf{G}_\ell\right) = \mathrm{rank}(\mathbf{G}_\ell^H\mathbf{\Omega}_\ell^{-\frac{1}{2}}\mathbf{\Omega}_\ell^{-\frac{1}{2}}\mathbf{G}_\ell) = \mathrm{rank}(\mathbf{G}_\ell) = \min(K, NN_{\mathrm{R}})$. Thus, we have

$$\begin{aligned}\mathrm{rank}(\mathbf{F}_\ell) &= \mathrm{rank}(\mathbf{\Upsilon} - \frac{\lambda_\ell^\star}{\ln(2)}\mathbf{G}_\ell^H\mathbf{\Omega}_\ell^{-1}\mathbf{G}_\ell)\\&\geq \mathrm{rank}(\mathbf{\Upsilon}) - \mathrm{rank}(\frac{\lambda_\ell^\star}{\ln(2)}\mathbf{G}_\ell^H\mathbf{\Omega}_\ell^{-1}\mathbf{G}_\ell)\\&= NN_{\mathrm{R}} - \min(K, NN_{\mathrm{R}}).\end{aligned}\tag{A.11}$$

where made use of $\mathrm{rank}(\mathbf{A}-\mathbf{B}) \geq \mathrm{rank}(\mathbf{A}) - \mathrm{rank}(\mathbf{B})$ [103].

From (A.4), we have

$$\begin{aligned}\mathrm{rank}(\bar{\mathbf{U}}_\ell^\star) &\leq \mathrm{Nullity}(\mathbf{F}_\ell^\star) = NN_{\mathrm{R}} - \mathrm{rank}(\mathbf{F}_\ell^\star)\\&\leq \min(K, NN_{\mathrm{R}}),\end{aligned}\tag{A.12}$$

where $\mathrm{Nullity}(\mathbf{A})$ denotes the dimension of the null space of $\mathbf{A}$. For the special case of $K = 1$ and from (A.12), we have $\mathrm{rank}(\bar{\mathbf{U}}_\ell^\star) \leq 1$. Also, since $\bar{\mathbf{U}}_\ell^\star \neq \mathbf{0}$, i.e., $\mathrm{rank}(\bar{\mathbf{U}}_\ell^\star) \geq 1$, this implies that $\mathrm{rank}(\bar{\mathbf{U}}_\ell^\star) = 1$ and thus concludes the proof.

Bibliography

[1] GSMA, "The Mobile Economy 2023," Oct. 2023. [Online]. Available: https://www.gsma.com/mobileeconomy/#trends

[2] ——, "The State of Mobile Internet Connectivity 2022," Oct. 2022. [Online]. Available: https://www.gsma.com/r/somic/

[3] ITU-R, "Minimum requirements related to technical performance for IMT-2020 radio interface(s)." Report ITU-R M.2410-0, Nov. 2017. [Online]. Available: https://www.itu.int/pub/R-REP-M.2410-2017

[4] International Telecommunication Union (ITU), "Global Connectivity Report 2022," 2022. [Online]. Available: https://www.itu.int/itu-d/reports/statistics/global-connectivity-report-2022/

[5] C. Bai, P. Dallasega, G. Orzes, and J. Sarkis, "Industry 4.0 technologies assessment: A sustainability perspective," *International Journal of Production Economics*, vol. 229, p. 107776, Nov. 2020.

[6] G. A. Akpakwu, B. J. Silva, G. P. Hancke, and A. M. Abu-Mahfouz, "A survey on 5G networks for the internet of things: Communication technologies and challenges," *IEEE Access*, vol. 6, pp. 3619–3647, 2018.

[7] M. Agiwal, A. Roy, and N. Saxena, "Next generation 5G wireless networks: A comprehensive survey," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 3, pp. 1617–1655, 2016.

[8] 3GPP, "Study on scenarios and requirements for next generation access technologies (release 17)." 3GPP TR 38.913 V17.0.0, Mar. 2022. [Online]. Avail-

able: https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=2996

[9] Cisco, "5G Security Innovation with Cisco," *White paper*, Jun. 2018.

[10] J. L. Burbank, J. Andrusenko, J. S. Everett, and W. T. Kasch, "Second-generation (2G) cellular communications," in *Wireless Networking: Understanding Internetworking Challenges*, 2013, pp. 250–365.

[11] O. Tipmongkolsilp, S. Zaghloul, and A. Jukan, "The evolution of cellular backhaul technologies: Current issues and future trends," *IEEE Commun. Surveys Tuts.*, vol. 13, no. 1, pp. 97–113, 2011.

[12] 3GPP, "Release 15 description; summary of Rel-15 work items (release 15)." 3GPP TR 21.915 V15.0.0, Sep. 2019. [Online]. Available: https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3389

[13] ——, "Issue 05: the 5G standard." 3GPP Highlights, Nov. 2022. [Online]. Available: https://www.3gpp.org/images/newsletters/3GPP_Highlights_Issue_5_WEB_opt1.pdf

[14] W. Ejaz, S. K. Sharma, S. Saadat, M. Naeem, A. Anpalagan, and N. A. Chughtai, "A comprehensive survey on resource allocation for CRAN in 5G and beyond networks," *Journal of Network and Computer Applications*, vol. 160, p. 102638, Jun. 2020.

[15] S. Park, C.-B. Chae, and S. Bahk, "Large-scale antenna operation in heterogeneous cloud radio access networks: A partial centralization approach," *IEEE Wireless Commun.*, vol. 22, no. 3, pp. 32–40, Jun. 2015.

[16] China Mobile Research Institute, "C-RAN: the road towards green RAN," *White paper, Version 2.5*, Oct. 2011.

[17] T. Q. S. Quek, M. Peng, O. Simeone, and W. Yu, *Cloud radio access networks: Principles, technologies, and applications*. Cambridge University Press, 2017.

[18] J. G. Andrews, "Seven ways that HetNets are a cellular paradigm shift," *IEEE Commun. Mag.*, vol. 51, no. 3, pp. 136–144, Mar. 2013.

[19] H. Q. Ngo, A. Ashikhmin, H. Yang, E. G. Larsson, and T. L. Marzetta, "Cell-free massive MIMO versus small cells," *IEEE Trans. Wireless Commun.*, vol. 16, no. 3,

pp. 1834–1850, Mar. 2017.

[20] Q. Wu, S. Zhang, B. Zheng, C. You, and R. Zhang, "Intelligent reflecting surface-aided wireless communications: A tutorial," *IEEE Trans. Commun.*, vol. 69, no. 5, pp. 3313–3351, May 2021.

[21] Y. Zhang, X. He, C. Zhong, L. Meng, and Z. Zhang, "Fronthaul compression and beamforming optimization for uplink C-RAN with intelligent reflecting surface-enhanced wireless fronthauling," *IEEE Commun. Lett.*, vol. 25, no. 6, pp. 1979–1983, Jun. 2021.

[22] B. Paul, A. R. Chiriyath, and D. W. Bliss, "Survey of RF communications and sensing convergence research," *IEEE Access*, vol. 5, pp. 252–270, 2017.

[23] ETSI, "Cloud RAN and MEC: a perfect pairing," *White paper No. 23*, Feb. 2018.

[24] T. T. Vu, D. T. Ngo, M. N. Dao, S. Durrani, D. H. N. Nguyen, and R. H. Middleton, "Energy efficiency maximization for downlink cloud radio access networks with data sharing and data compression," *IEEE Trans. Wireless Commun.*, vol. 17, no. 8, pp. 4955–4970, Aug. 2018.

[25] J. Tang, W. P. Tay, T. Q. S. Quek, and B. Liang, "System cost minimization in cloud RAN with limited fronthaul capacity," *IEEE Trans. Wireless Commun.*, vol. 16, no. 5, pp. 3371–3384, May 2017.

[26] B. B.s and S. Azeem, "A survey on increasing the capacity of 5G fronthaul systems using RoF," *Optical Fiber Technology*, vol. 74, p. 103078, Dec. 2022.

[27] M. Peng, C. Wang, V. Lau, and H. V. Poor, "Fronthaul-constrained cloud radio access networks: Insights and challenges," *IEEE Wireless Commun.*, vol. 22, no. 2, pp. 152–160, Apr. 2015.

[28] S.-H. Park, C. Song, and K.-J. Lee, "Inter-cluster design of wireless fronthaul and access links for the downlink of C-RAN," *IEEE Wireless Commun. Lett.*, vol. 6, no. 2, pp. 270–273, Apr. 2017.

[29] B. Hu, C. Hua, C. Chen, and X. Guan, "Joint beamformer design for wireless fronthaul and access links in C-RANs," *IEEE Trans. Wireless Commun.*, vol. 17, no. 5, pp. 2869–2881, May 2018.

[30] Z.-q. Luo, W.-k. Ma, A. M.-c. So, Y. Ye, and S. Zhang, "Semidefinite relaxation

of quadratic optimization problems," *IEEE Signal Processing Magazine*, vol. 27, no. 3, pp. 20–34, May 2010.

[31] P. Rost, C. J. Bernardos, A. D. Domenico, M. D. Girolamo, M. Lalam, A. Maeder, D. Sabella, and D. Wübben, "Cloud technologies for flexible 5G radio access networks," *IEEE Commun. Mag.*, vol. 52, no. 5, pp. 68–76, May 2014.

[32] C.-L. I, C. Rowell, S. Han, Z. Xu, G. Li, and Z. Pan, "Toward green and soft: A 5G perspective," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 66–73, Feb. 2014.

[33] 3GPP, "Study on new radio access technology: Radio access architecture and interfaces (release 14)." 3GPP TR 38.801 V14.0.0, Mar. 2017. [Online]. Available: https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3056

[34] L. M. P. Larsen, A. Checko, and H. L. Christiansen, "A survey of the functional splits proposed for 5G mobile crosshaul networks," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 1, pp. 146–172, 2019.

[35] A. Checko, H. L. Christiansen, Y. Yan, L. Scolari, G. Kardaras, M. S. Berger, and L. Dittmann, "Cloud RAN for mobile networks—a technology overview," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 1, pp. 405–426, 2015.

[36] M. F. Hossain, A. U. Mahin, T. Debnath, F. B. Mosharrof, and K. Z. Islam, "Recent research in cloud radio access network (C-RAN) for 5G cellular systems - A survey," *Journal of Network and Computer Applications*, vol. 139, pp. 31–48, Aug. 2019.

[37] A. El Gamal and Y.-H. Kim, *Network Information Theory*. Cambridge University Press, 2011.

[38] J. Kim, S.-H. Park, O. Simeone, I. Lee, and S. Shamai Shitz, "Joint design of fronthauling and hybrid beamforming for downlink C-RAN systems," *IEEE Trans. Commun.*, vol. 67, no. 6, pp. 4423–4434, Jun. 2019.

[39] S.-H. Park, O. Simeone, O. Sahin, and S. Shamai, "Joint precoding and multivariate backhaul compression for the downlink of cloud radio access networks," *IEEE Trans. Signal Process.*, vol. 61, no. 22, pp. 5646–5658, Nov. 2013.

[40] S.-H. Park, O. Simeone, O. Sahin, and S. Shamai Shitz, "Fronthaul compression for cloud radio access networks: Signal processing advances inspired by network

information theory," *IEEE Signal Processing Magazine*, vol. 31, no. 6, pp. 69–79, Nov. 2014.

[41] Y. Zhou and W. Yu, "Fronthaul compression and transmit beamforming optimization for multi-antenna uplink C-RAN," *IEEE Trans. Signal Process.*, vol. 64, no. 16, pp. 4138–4151, Aug. 2016.

[42] T. L. Marzetta, E. G. Larsson, H. Yang, and H. Q. Ngo, *Fundamentals of massive MIMO*. Cambridge University Press, 2016. [Online]. Available: https://doi.org/10.1017/CBO9781316799895

[43] E. Björnson, J. Hoydis, and L. Sanguinetti, "Massive MIMO networks: Spectral, energy, and hardware efficiency," *Foundations and Trends® in Signal Processing*, vol. 11, no. 3-4, pp. 154–655, Nov 2017. [Online]. Available: http://dx.doi.org/10.1561/2000000093

[44] J. Chen, "When does asymptotic orthogonality exist for very large arrays?" in *2013 IEEE Global Communications Conference (GLOBECOM)*, Dec. 2013, pp. 4146–4150.

[45] L. Lu, G. Y. Li, A. L. Swindlehurst, A. Ashikhmin, and R. Zhang, "An overview of massive MIMO: Benefits and challenges," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 742–758, Oct. 2014.

[46] K. P. Valavanis and G. J. Vachtsevanos, *Handbook of unmanned aerial vehicles*. Springer, 2015.

[47] Y. Zeng, Q. Wu, and R. Zhang, "Accessing from the sky: A tutorial on UAV communications for 5G and beyond," *Proceedings of the IEEE*, vol. 107, no. 12, pp. 2327–2375, Dec. 2019.

[48] R. I. Bor-Yaliniz, A. El-Keyi, and H. Yanikomeroglu, "Efficient 3-D placement of an aerial base station in next generation cellular networks," in *2016 IEEE International Conference on Communications (ICC)*, May 2016, pp. 1–5.

[49] M. Mozaffari, W. Saad, M. Bennis, and M. Debbah, "Efficient deployment of multiple unmanned aerial vehicles for optimal wireless coverage," *IEEE Commun. Lett.*, vol. 20, no. 8, pp. 1647–1650, Aug. 2016.

[50] E. Yanmaz, "Connectivity versus area coverage in unmanned aerial vehicle net-

works," in *2012 IEEE International Conference on Communications (ICC)*, Jun. 2012, pp. 719–723.

[51] S. Hayat, E. Yanmaz, and R. Muzaffar, "Survey on unmanned aerial vehicle networks for civil applications: A communications viewpoint," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 4, pp. 2624–2661, 2016.

[52] Y. Zeng, R. Zhang, and T. J. Lim, "Wireless communications with unmanned aerial vehicles: Opportunities and challenges," *IEEE Commun. Mag.*, vol. 54, no. 5, pp. 36–42, May 2016.

[53] M. M. Azari, F. Rosas, K.-C. Chen, and S. Pollin, "Ultra reliable UAV communication using altitude and cooperation diversity," *IEEE Trans. Commun.*, vol. 66, no. 1, pp. 330–344, Jan. 2018.

[54] L. Liu, S. Zhang, and R. Zhang, "Multi-beam UAV communication in cellular uplink: Cooperative interference cancellation and sum-rate maximization," *IEEE Trans. Wireless Commun.*, vol. 18, no. 10, pp. 4679–4691, Oct. 2019.

[55] Y. Zeng and R. Zhang, "Energy-efficient UAV communication with trajectory optimization," *IEEE Trans. Wireless Commun.*, vol. 16, no. 6, pp. 3747–3760, Jun. 2017.

[56] Q. Wu, Y. Zeng, and R. Zhang, "Joint trajectory and communication design for multi-UAV enabled wireless networks," *IEEE Trans. Wireless Commun.*, vol. 17, no. 3, pp. 2109–2121, Mar. 2018.

[57] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.

[58] M. A. Habibi, M. Nasimi, B. Han, and H. D. Schotten, "A comprehensive survey of RAN architectures toward 5G mobile communication system," *IEEE Access*, vol. 7, pp. 70 371–70 421, 2019.

[59] G. E. Gonçalves, G. L. Santos, L. Ferreira, É. d. S. Rocha, L. M. F. de Souza, A. L. C. Moreira, J. Kelner, and D. Sadok, "Flying to the clouds: The evolution of the 5G radio access networks," in *The Cloud-to-Thing Continuum: Opportunities and Challenges in Cloud, Fog and Edge Computing*, T. Lynn, J. G. Mooney, B. Lee, and P. T. Endo, Eds. Cham: Springer International Publishing, 2020, pp. 41–60.

[60] M. Peng, Y. Li, J. Jiang, J. Li, and C. Wang, "Heterogeneous cloud radio access networks: A new perspective for enhancing spectral and energy efficiencies," *IEEE Wireless Commun.*, vol. 21, no. 6, pp. 126–135, Dec. 2014.

[61] I. A. Alimi, A. L. Teixeira, and P. P. Monteiro, "Toward an efficient C-RAN optical fronthaul for the future networks: A tutorial on technologies, requirements, challenges, and solutions," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 1, pp. 708–769, 2018.

[62] M. Jaber, M. A. Imran, R. Tafazolli, and A. Tukmanov, "5G backhaul challenges and emerging research directions: A survey," *IEEE Access*, vol. 4, pp. 1743–1766, 2016.

[63] 3GPP, "Coordinated multi-point operation for lte physical layer aspects (release 11)." 3GPP TR 36.819 V11.2.0, Sep. 2013. [Online]. Available: https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=2498

[64] A. Liu and V. K. N. Lau, "Joint power and antenna selection optimization for energy-efficient large distributed MIMO networks," in *2012 IEEE International Conference on Communication Systems (ICCS)*, Nov. 2012, pp. 230–234.

[65] S. Mahboob, C. Mahapatra, and V. C. Leung, "Energy-efficient multiuser MIMO downlink transmissions in massively distributed antenna systems with predefined capacity constraints," in *2012 Seventh International Conference on Broadband, Wireless Computing, Communication and Applications*, Nov. 2012, pp. 208–211.

[66] J. Lu, J. Zhang, S. Yang, S. Cai, and Y. Ni, "Joint precoder and decoder design in downlink multi-user MIMO C-RAN with imperfect CSI," *Physical Communication*, vol. 48, p. 101406, Oct. 2021.

[67] S.-H. Park, K.-J. Lee, C. Song, and I. Lee, "Joint design of fronthaul and access links for C-RAN with wireless fronthauling," *IEEE Signal Process. Lett.*, vol. 23, no. 11, pp. 1657–1661, Nov. 2016.

[68] L. Liu and R. Zhang, "Optimized uplink transmission in multi-antenna C-RAN with spatial compression and forward," *IEEE Trans. Signal Process.*, vol. 63, no. 19, pp. 5083–5095, Oct. 2015.

[69] S.-H. Park, O. Simeone, O. Sahin, and S. Shamai, "Joint decompression and decoding for cloud radio access networks," *IEEE Signal Process. Lett.*, vol. 20, no. 5, pp. 503–506, May 2013.

[70] Y. Zhou and W. Yu, "Optimized backhaul compression for uplink cloud radio access network," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1295–1307, Jun. 2014.

[71] ——, "Optimized beamforming and backhaul compression for uplink MIMO cloud radio access networks," in *2014 IEEE Globecom Workshops (GC Wkshps)*, Dec. 2014, pp. 1493–1498.

[72] Y. Zhou, Y. Xu, J. Chen, and W. Yu, "Optimality of gaussian fronthaul compression for uplink MIMO cloud radio access networks," in *2015 IEEE International Symposium on Information Theory (ISIT)*, Jun. 2015, pp. 2241–2245.

[73] Y. Zhou, Y. Xu, W. Yu, and J. Chen, "On the optimal fronthaul compression and decoding strategies for uplink cloud radio access networks," *IEEE Transactions on Information Theory*, vol. 62, no. 12, pp. 7402–7418, Dec. 2016.

[74] S.-H. Park, K.-J. Lee, C. Song, and I. Lee, "Compressed cooperative reception for the uplink of C-RAN with wireless fronthaul," in *2017 International Symposium on Wireless Communication Systems (ISWCS)*, Aug. 2017, pp. 211–215.

[75] T. X. Vu, H. D. Nguyen, and T. Q. S. Quek, "Adaptive compression and joint detection for fronthaul uplinks in cloud radio access networks," *IEEE Trans. Commun.*, vol. 63, no. 11, pp. 4565–4575, Nov. 2015.

[76] T. X. Vu, H. D. Nguyen, T. Q. S. Quek, and S. Sun, "Adaptive cloud radio access networks: Compression and optimization," *IEEE Trans. Signal Process.*, vol. 65, no. 1, pp. 228–241, Jan. 2017.

[77] A. C. Cirik, O. Taghizadeh, L. Lampe, and R. Mathar, "Fronthaul compression and precoding design for full-duplex cloud radio access network," *IEEE Syst. J.*, vol. 13, no. 2, pp. 1113–1124, Jun. 2019.

[78] F. E. Kadan and A. Ö. Yılmaz, "A theoretical performance bound for joint beamformer design of wireless fronthaul and access links in downlink C-RAN," *IEEE Trans. Wireless Commun.*, vol. 21, no. 4, pp. 2177–2192, Apr. 2022.

[79] S. Park, C.-B. Chae, and S. Bahk, "Before/after precoded massive MIMO in cloud

radio access networks," in *2013 IEEE International Conference on Communications Workshops (ICC)*, Jun. 2013, pp. 169–173.

[80] L. Liu, S. Zhang, and R. Zhang, "CoMP in the sky: UAV placement and movement optimization for multi-user communications," *IEEE Trans. Commun.*, vol. 67, no. 8, pp. 5645–5658, Aug. 2019.

[81] S. Roth, A. Kariminezhad, and A. Sezgin, "Base-stations up in the air: Multi-UAV trajectory control for min-rate maximization in uplink C-RAN," in *ICC 2019 - 2019 IEEE International Conference on Communications (ICC)*, May 2019, pp. 1–6.

[82] X. Li, C. Pan, C. Zhang, C. He, and K. Wang, "Data rate maximization in UAV-assisted C-RAN," *IEEE Wireless Commun. Lett.*, vol. 9, no. 12, pp. 2163–2167, Dec. 2020.

[83] L. Li, T.-H. Chang, and S. Cai, "UAV positioning and power control for two-way wireless relaying," *IEEE Trans. Wireless Commun.*, vol. 19, no. 2, pp. 1008–1024, Feb. 2020.

[84] H. Yang and T. L. Marzetta, "Massive MIMO with max-min power control in line-of-sight propagation environment," *IEEE Trans. Commun.*, vol. 65, no. 11, pp. 4685–4693, Nov. 2017.

[85] C. Pan, H. Zhu, N. J. Gomes, and J. Wang, "Joint precoding and RRH selection for user-centric green MIMO C-RAN," *IEEE Trans. Wireless Commun.*, vol. 16, no. 5, pp. 2891–2906, May 2017.

[86] ——, "Joint user selection and energy minimization for ultra-dense multi-channel C-RAN with incomplete CSI," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 8, pp. 1809–1824, Aug. 2017.

[87] P. Chandhar, D. Danev, and E. G. Larsson, "Massive MIMO for communications with drone swarms," *IEEE Trans. Wireless Commun.*, vol. 17, no. 3, pp. 1604–1629, Mar. 2018.

[88] J. Kim, J. Kim, and S.-H. Park, "Joint design of power control and SIC decoding order for max-min fairness optimization in uplink NOMA systems," in *2021 International Conference on Information Networking (ICOIN)*. Jeju Island, Korea (South): IEEE, Jan. 2021, pp. 339–342.

[89] D. Maryopi, Y. Huang, and A. Ikhlef, "Sum-rate maximization in uplink CRAN with a massive MIMO fronthaul," in *2021 IEEE Globecom Workshops (GC Wkshps)*, Dec. 2021, pp. 1–6.

[90] Y. Huang and A. Ikhlef, "Joint optimization of wireless fronthaul and access links in CRAN with a massive MIMO central unit," in *ICC 2022 - IEEE International Conference on Communications*, May 2022, pp. 1906–1911.

[91] M. Grant and S. Boyd, "The CVX users' guide, release 2.2," Jan. 2020. [Online]. Available: http://cvxr.com/cvx/doc/CVX.pdf

[92] Y. Sun, P. Babu, and D. P. Palomar, "Majorization-minimization algorithms in signal processing, communications, and machine learning," *IEEE Trans. Signal Process.*, vol. 65, no. 3, pp. 794–816, 2017.

[93] H. S. Dhillon, R. K. Ganti, and J. G. Andrews, "Modeling non-uniform ue distributions in downlink cellular networks," *IEEE Wireless Commun. Lett.*, vol. 2, no. 3, pp. 339–342, 2013.

[94] 3GPP, *Study on channel model for frequencies from 0.5 to 100 GHz*. 3rd Generation Partnership Project (3GPP), TR 38.901 V16.1.0, Jan. 2020.

[95] K. Toh, M. Todd, and R. Tutuncu, "SDPT3 — a Matlab software package for semidefinite programming," *Optimization Methods and Software*, vol. 11, p. 545–581, 1999.

[96] R. Tutuncu, K. Toh, and M. Todd, "Solving semidefinite-quadratic-linear programs using SDPT3," *Mathematical Programming Ser. B*, vol. 95, p. 189–217, 2003.

[97] J. F. Sturm, "Using sedumi 1.02, a MATLAB toolbox for optimization over symmetric cones," *Optimization Methods and Software*, vol. 11, no. 1-4, pp. 625–653, 1999. [Online]. Available: https://doi.org/10.1080/10556789908805766

[98] J. Kang, O. Simeone, J. Kang, and S. Shamai, "Fronthaul compression and precoding design for C-RANs over ergodic fading channels," *IEEE Trans. Veh. Technol.*, vol. 65, no. 7, pp. 5022–5032, Jul. 2016.

[99] A. Liu and V. K. N. Lau, "Impact of CSI knowledge on the codebook-based hybrid beamforming in massive MIMO," *IEEE Trans. Signal Process.*, vol. 64, no. 24, pp. 6545–6556, Dec. 2016.

[100] 3GPP, "Study on channel model for frequencies from 0.5 to 100 GHz," Dec. 2019.

[101] D. Han, J. Park, S.-H. Park, and N. Lee, "Sparse joint transmission for cloud radio access networks with limited fronthaul capacity," *IEEE Trans. Wireless Commun.*, vol. 21, no. 5, pp. 3395–3408, May 2022.

[102] M. Hong, Q. Li, and Y.-F. Liu, "Decomposition by successive convex approximation: A unifying approach for linear transceiver design in heterogeneous networks," *IEEE Trans. Wireless Commun.*, vol. 15, no. 2, pp. 1377–1392, Feb. 2016.

[103] J. E. Gentle, *Matrix Algebra: Theory, Computations, and Applications in Statistics*. Springer Nature, 2007.