Decentralised Distributed Massive MIMO

Junbo Zhao

Doctor of Philosophy

Physics, Engineering and Technology
University of York

September 2023

Abstract

In this thesis, decentralised distributed massive multiple-input multiple-output (DD-MaMIMO) is considered for providing high spectral efficiency (SE) per user. In the DD-MaMIMO system, a large number of access points (APs) within a coordination region are connected to an edge processing unit (EPU) via fronthaul links, serving the users within a service region. Initially, we investigate a DD-MaMIMO system with perfect fronthaul links and assume that the processing takes place in the EPU. To demonstrate the improved SE, we compare our proposed architecture to cell-free MaMIMO. Furthermore, we discuss the scalability of DD-MaMIMO and give its definition. Secondly, we extend our research to the limited-capacity fronthaul links which is essential in practice. To model the limited-capacity fronthaul links, we adopt the Bussgang decomposition to express the quantisation. We propose two strategies for obtaining channel state information (CSI): estimate-and-quantise (EQ) and quantise-and-estimate (QE). Particularly, in the QE scheme, we derive the closed-form expressions of Bussgang decomposition coefficients for the non-Gaussian distribution input of the quantiser, as the elements of pilots follow complex Gaussian distribution. Both CSI acquisition strategies are analysed with respect to the mean square error (MSE) of channel estimation. Finally, we explore the processing which happens at the AP which is the local estimation in DD-MaMIMO. Here, two approaches are exploited for data decoding at the EPU: simply averaging decoding and large scale fading decoding. We further compare the local estimation scheme with the decentralised processing scheme. The scalability is also discussed as the channel estimation and data detection happens at the AP.

Table of Contents

List of Figures

List of Tables

Acknowledgements

First and foremost, I would like to express my sincere gratitude to my supervisor, Prof. Alister Burr, for his attentive guidance and valuable discussions during my time pursuing a doctoral degree. I have greatly benefited from his extensive knowledge and his academic rigour. The accomplishment of this study and the completion of this thesis would not have been feasible without his continuous supports, constructive suggestions and encouragement. I would also like to extend my thanks to my thesis advisor, Dr. Youngwook Ko, for his insightful discussions on my research.

Special thanks go to all my colleagues in the communications research group who provide me with a warm and friendly working atmosphere. I would also like to express my appreciation to all the staff and friends at the University of York for their kind support during the hard and challenging pandemic period. In particular, I would like to thank Rahul, Tianyuan, Huiting, Tian, Yunlong, Naijia, Zhihao, Zeliang and Hezhihan for their sincere concern and emotional support throughout my academic journey.

Finally and most importantly, I would like to express my deepest, most heartfelt gratitude to my parents for their unconditional support and selfless love. Their sacrifices and unwavering trust in me have encouraged me to pursue my dream.

Declaration

I declare that this thesis is a presentation of original work and I am the sole author. This work has not previously been presented for an award at this, or any other, University. All sources are acknowledged as References.

Chapter 1

Introduction

1.1 Overview and Motivation

Mobile communication networks evolve with each passing decade. As of 2020, the fifth generation (5G) communication system has been standardised and commercially launched. Even though enhancements to the 5G new radio (NR) and 5G advanced are still in progress, research on the sixth generation (6G) has already commenced, aiming to meet the future demands anticipated for the 2030s. Nowadays, the pursuit of higher data rates is no longer the sole objective for the next generation mobile communication network. The goal has shifted towards the establishment of a ubiquitous intelligent communication system that enables seamless connectivity amongst humans, devices, sensors and machines. Additionally, it also aims to merge the boundaries between the physical and digital worlds, transforming communications methods and improving life quality. This leads to many novel concepts and emerging techniques such as Artificial Intelligence (AI), Internet of Things (IoT), Internet of Everything (IoE), Extended Reality (XR), self-driving cars and telemedicine.

To realise this aspiration, the future mobile communication network faces numerous challenges. According to the framework for International Mobile Telecommunications-2030 (IMT-2030) proposed by International Telecommunication Unit Radiocommunication Sector (ITU-R), 6G needs to adapt to an ultra-dense network, potentially accommodating ten to a hundred million devices per square kilometre [1]. At the same time, the peak data rate is anticipated to reach one Terabit per second (Tbit/s) [2-8], which is a 10 to 100-fold increase over 5G. As the number of accessed devices increases, spectrum resources will become scarce. Therefore, carrying out a communication system with higher spectral efficiency is necessary. On the other hand, the increased number of accessed devices and data rate also leads to a substantial energy consumption. Thus, improving energy efficiency should be considered to reduce both costs and carbon emissions per bit of transmitted data. Furthermore, in order to cater to

various demands, the future network should provide lower latency, increased reliability, and enhanced security. In response to these challenges, ITU-R has categorised 6G services into six usage scenarios [1]. The difference from 5G usage scenarios is that the communication capabilities, such as data rate, latency, reliability and connection density, are enhanced. Furthermore, the new usage scenarios for coverage extension and service extension are introduced. These scenarios are summarised as follows:

- **Immersive Communication** is an extension of the enhanced mobile broadband (eMBB) scenario [1, 2] applied in 5G, which aims to facilitate a higher data rate access to the network and a greater area traffic capacity [1, 3]. This supports the applications such as augmented reality (AR), virtual reality (VR) and wireless holographic communication [4-6].
- **Massive Communication** focuses on the ultra-dense network, which builds upon the scenario of massive machine-type communication (mMTC) [1, 2], serving a larger number of accessed devices.
- **Hyper Reliable and Low-Latency Communication** (HRLLC) in 6G targets applications that demand sensitive reliability, latency and availability, with envisioned 100 microsecond latency and availability of $1 - 10^{-7}$ [1, 3]. This advances the ultra-reliable low-latency communication (URLLC) scenario, supporting the use cases such as remote medical surgery, smart grid, and automatic vehicles [1-3].
- **Ubiquitous Connectivity**, introduced as a new usage scenario, emphasises extensive coverage and enhanced mobility [1, 9]. It promotes the integration of terrestrial and satellite-aerial networks and aims for velocities up to 1000 kilometres per hour in airline systems [1, 3] .
- **Integrated AI and Communication** represents a new application scenario in which services beyond communications assisted by AI are provided [1, 9]. It enhances radio communications in terms of physical layer signal processing and resource management, by exploiting machine learning [7, 10].
- **Integrated Sensing and Communication** aims to unify wireless signal sensing and communication within a single system for their mutual benefits [11]. It enables services beyond communications. For example, an autonomous vehicle can simultaneously sense its surroundings using radar and communicate with other vehicles. Additionally, it also facilitates high-precision localisation for

obtaining more accurate beamforming when tracking the channel state information (CSI) [12].

Motivated by some of these usage scenarios and challenges, many technologies have been proposed for satisfying an ultra-dense network with low latency and high data rate. One among others is MaMIMO where the base station is equipped with a large number of antennas, serving many user terminals in the same time-frequency resource [13]. This technique plays a key role in 5G [14, 15], and continues to hold significant potential in enhancing spectral and energy efficiency for 6G. Furthermore, to mitigate the intercellular interference, the concept of cell-free MaMIMO has been proposed, in which MaMIMO is deployed in a distributed manner, simultaneously serving many users [14, 16]. This approach increases the coverage and spectral efficiency (SE) which also can be a promising technique for 5G and beyond [17]. Many works on cell-free MaMIMO have been researched [14-30], however there is still much more to be done. Recently, an initiative to incorporate virtualisation, intelligence and open interface within a RAN has been proposed, known as the open radio access network (Open RAN) [31, 32]. This allows different vendors to use a unified interconnection standard. Additionally, in Open RAN, hardware is decoupled from software which provides greater flexibility and efficiency in network deployments. To fulfil these requirements, a scalable network needs to be considered with flexible implementation of signal processing. The related topics about adapting cell-free MaMIMO for Open RAN have not previously been investigated.

In [30], the edge effects and latency in cell-free MaMIMO has been discussed. To overcome these disadvantages, a new architecture was proposed which combines the Fog RAN (F-RAN) and cell-free MaMIMO with a defined coordination region [30]. However, the analysis in [30] is based on the assumption of channel hardening which is a phenomenon where the variations of channel gain decrease if the number of antennas at the base station is very large. This is not strict because the phenomenon of harden channel vanishes in cell-free MaMIMO when each base station is only equipped one or few antennas[19]. Therefore, motivated by these challenges, we extend the research on this architecture without assuming channel hardening, and provide a thorough study with limited-capacity fronthaul.

1.2 Contributions

The contributions of this thesis are summarised as follows:

In Chapter 3:

- We redefine the concept of Fog Massive MIMO [30] as Decentralised Distributed MaMIMO (DD-MaMIMO) in order to distinguish the definition of "fog" with [33] and its general use in fog computing respect. We explicitly emphasize that the processing occurs at the network edge, which is decentralised in a distributed MaMIMO system. To mitigate the cluster-edge effect, the coordination region among the adjacent network clusters is defined, which also enables the system to be scalable.

- A novel pilot assignment algorithm is proposed, adapted for the practical implementation in DD-MaMIMO, which takes into account the users in the overlapped coordination regions.

- A general expression of uplink achievable SE for DD-MaMIMO is derived based on the knowledge of channel estimation and data detection at the EPU without assuming channel hardening. The expression of the signal-to-interference-plus-noise ratio (SINR) is given.

In Chapter 4:

- We consider the limited-capacity fronthaul links in DD-MaMIMO, where we use a scalar uniform quantiser to transfer the analogue signals to the digital signals. We exploit the Bussgang decomposition to model the non-linear quantiser.

- We consider two strategies for obtaining the channel state information (CSI) at the EPU: i) the quantised version of the received pilots at the APs is available at the EPU, and ii) the quantised version of the estimated channels at the APs is known at the EPU. In the first strategy, we adopt two types of pilot sequences which are complex Gaussian distributed pilots and binary pilots. The performances for both strategies and both types of pilots are evaluated by calculating the mean squared error (MSE) of channel estimation.

- We analyse the effect of the complex non-Gaussian distribution input of the quantiser when the elements of pilot sequences follow a complex Gaussian distribution. To find the optimum quantisation step interval, we derive the probability distribution function (pdf) of the real part of this complex non-Gaussian distribution. The pdf of the imaginary part is the identical to that of the real part.

- Based on the pdf, we derive closed-form expressions for two linear coefficients of the Bussgang decomposition and calculate the optimum quantisation step interval for the use of complex Gaussian distributed pilots.

- Considering the quantisation distortion that affects the CSI acquisition and data transmission, we derive an achievable SE expression for the uplink DD-MaMIMO with minimum mean square error (MMSE) data detection processed at the EPU. The expression of the signal-to-interference-and-distortion-plus-noise ratio (SIDNR) is provided.

In Chapter 5:

- Similarly, we compare the local estimation, where the signal processing takes place at the APs, with decentralised processing in DD-MaMIMO assuming the perfect fronthaul. The results show that the local estimation case has a worse performance than the latter. However, with the increase of the number of antennas for each AP, the discrepancy reduces.

- We consider two final decoding methods at the EPU for the local estimation case: i) large scale fading decoding employs the channel statistics at the EPU, and ii) simply averaging the estimated data from all APs within the coordination region at the EPU.

- We derive an achievable uplink SE expression for the local estimation case with respect to two decoding methods. The different SINR expressions are given.

- The scalability of the local estimation in DD-MaMIMO is discussed and it is feasible.

- According to the different signal processing locations in DD-MaMIMO, we analyse that our system can be a promising architecture for performing in Open RAN.

Conference Papers

A. Burr, S. Islam, J. Zhao, and M. Bashar, "Cell-free Massive MIMO with multi-antenna access points and user terminals," in 2020 54th Asilomar Conference on Signals, Systems, and Computers, ieeexplore.ieee.org, Nov. 2020, pp. 821–825.

Journal Papers

There is one paper titled "Decentralised Distributed Massive MIMO" that is prepared to submit to IEEE Transactions on Wireless Communications before the 29th February 2024. Another paper related to the quantisation in Decentralised Distributed Massive MIMO is to be finished and could be submitted before May 2024.

1.3 Thesis Outline

The thesis is organised as follows. Chapter 2 provides a literature review with respect to the subsequent chapters. First, we describe the evolution of the wireless cellular network, indicating key techniques and relevant services developed in each generation. Then, the various RAN architectures along with physical entities are discussed. We next give the basics of MaMIMO, which can be deployed in single-cell, multi-cell or cell-free schemes. In the context of cell-free MaMIMO, we introduce fundamental concepts of channel estimation and uplink data transmission. Finally, we present the basic principle of the uniform quantiser and provide the method for finding its optimal step size.

In Chapter 3, a DD-MaMIMO architecture with unlimited-capacity fronthaul links is introduced. We first define the DD-MaMIMO system, which can significantly eliminate the cluster-edge effect observed in cell-free MaMIMO by leveraging the coordination region. Then, channel estimation is considered. Furthermore, we propose a pilot allocation algorithm that considers the overlapped coordination region in DD-MaMIMO, where the neighbouring clusters are deployed. We next discuss data

transmission and data detection without assuming channel hardening on the uplink. Given the general form of detected data, we derive an achievable SE expression in terms of SINR. Subsequently, a scalability issue in cell-free MaMIMO is mentioned. Following this, we explain how DD-MaMIMO can resolve this issue. Finally, numerical results are provided to validate our research and evaluate the performance of our proposed system.

Chapter 4 explores the scenario of limited-capacity fronthaul links in DD-MaMIMO, where a uniform quantiser is employed at the AP. We begin the discussion with the CSI acquisition, which can be obtained through two schemes. The pilots can either be quantised first and then sent to the EPU for channel estimation, or they can be used to estimate channels first, followed by quantisation. Subsequently, we express the quantised estimated channel adopting two different CSI acquisition strategies. Based on that, we derive the closed-form expression for calculating the optimum quantisation step interval. Furthermore, non-Gaussian distribution for the input of quantiser is analysed for finding the correct optimum quantisation step interval when the pilots follow the complex Gaussian distribution. Moreover, we also utilise binary pilot sequences to avoid the issue of non-Gaussian distribution in quantisation. After acquiring the CSI, we analyse the data transmission considering the limited-capacity fronthaul. This leads us to derive an achievable uplink SE expression for the signal-to-interference-and-distortion-plus-noise ratio (SIDNR). Finally, numerical results are provided which present the performance of two CSI acquisition schemes and the overall system with different number of quantisation bits.

In Chapter 5, we introduce the local estimation in DD-MaMIMO. Similar to Chapter 3, we first outline the basic concepts of the system model where the channel estimation and data detection occur at the AP. Subsequently, we calculate the channel estimation. This is then followed by the data detection, in which the AP only uses its own locally estimated channels. Next, the locally detected data are sent to the EPU for final decoding, which can be implemented by two methods. In the first scheme, we consider large scale fading decoding, where channel statistics are exploited due to the CSIs being unknown at the EPU. In the second scheme, the estimated data from all APs within the coordination region is simply averaged at the EPU. According to both decoding schemes, the expressions of SE with SINR are given, respectively. Before evaluating the system performance, the scalability is discussed. Finally, we give the numerical

results which shows the local estimation in DD-MaMIMO can also be a promising scheme. This validates our discussion of different locations for signal processing.

In Chapter 6, we provide a comprehensive conclusion of this thesis and discuss the potential challenges and research in the future.

1.4 Notation

In this sub-chapter, some important notations used in this thesis will be stated. Any specific notation will be introduced when it appears at the first time. The upper-case boldface Roman letter represents a matrix whereas the lower-case boldface denotes a column vector. The real and complex matrix with the size of $M \times K$ are represented by $\mathbb{R}^{M\times K}$ and $\mathbb{C}^{M\times K}$, respectively. The $M \times M$ identity matrix is $\boldsymbol{I}_{M\times M}$. The random variables are written as RVs. The complex Gaussian distribution of a random variable with mean μ and variance σ^2 is represented by $\mathcal{CN}(\mu, \sigma^2)$. For the real and imaginary part of a complex number, we use $\mathrm{Re}(\cdot)$ and $\mathrm{Im}(\cdot)$ to represent, respectively. We denote the expectation of a random variable as $\mathbb{E}\{\cdot\}$. The complex conjugate is represented by $(\cdot)^*$ and the Hermitian or transpose conjugate is $(\cdot)^H$. The m-th element of a vector is $[\cdot]_m$. We use $\|\cdot\|$ to denote the norm or magnitude of a vector. We adopt $diag(\cdot)$ for a diagonal matrix with on-diagonal elements taken from a vector. The notation arc represents the contour of a curve and Res denotes the residue. In thesis, many subscripts are used to define the different variable names and indices, respectively. The subscripts “m” , “n” and “k” denote indices. Other subscripts such as “u” in “y” (y_u) or “p” in “ τ ” (τ_p) are referred to variable names considered as entities. Consequently, $y_{u,mnk}$ represents the variable y_u with the indices m, n and k.

Chapter 2

Background

This chapter aims to provide a literature review for the subsequent chapters. Firstly, we intend to discuss the evolution of wireless cellular networks, indicating the characteristics a system needs to serve future wireless communication networks. We will then study various radio access networks (RANs) from 1G through to 5G and beyond. This will provide fundamental knowledge of the entities in the wireless network architecture. The emerging technique of massive multiple-input multiple-output (MaMIMO) will also be presented. Furthermore, a comprehensive introduction to cell-free MaMIMO will be given. Finally, we will discuss the basics of quantisation.

2.1 History of Wireless Cellular Network

Since the first generation (1G) wireless cellular network was proposed in 1980s, four decades have passed. The wireless network has experienced evolution five times with respect to the second, third, fourth and fifth generation (2G, 3G, 4G and 5G) networks. In 1G the analogue signal was used to convey the information of the voice in the wireless network, however it is replaced by the digital cellular networks in 2G where the data service such as Short Messaging Service (SMS) was introduced. In the 1990s, the Global System for Mobile Communications (GSM) as a 2G standard was developed by the European Telecommunications Standards Institute (ETSI). Then, the air interface standards for 3G, such as Wideband Code-Division Multiple Access (W-CDMA), CDMA2000 and Time-Division Synchronization CDMA (TD-SCDMA), were proposed by the Third Generation Partnership Project (3GPP). In the 3G era, the focus shifted towards data services, including mobile Internet access, web browsing, video calls and mobile television [5, 34]. Subsequently, 4G was standardised as Long Term Evolution (LTE) which applied Orthogonal Frequency Division Multiplexing (OFDM), Multiple-Input and Multiple-Output (MIMO) techniques and all-Internet Protocol (IP) flat architecture to achieve higher data rates and Quality of Service (QoS) [35]. In 2014, a paper introduced the perspective of 5G mobile networks, highlighting

the ultra-densification, millimetre wave (mmWave) and MaMIMO as key technologies for the development of 5G [36]. Since then, the conceptual framework of 5G has continued to evolve [5]. In 2015, 3GPP initiated a study on the 5G New Radio (NR) as a standard for 5G, which supported a wide range of services, including virtual reality (VR), augmented reality (AR), high-definition (HD) video streaming and automatic driving [3, 5]. After the first commercial launch of 5G network in 2019, the global deployment of 5G has been progressively advancing. However, as the rapid growth of mobile traffic, 5G may not fulfil the demands of future emerging technologies which require massive interconnectivity with highly diverse devices [7]. Thus, the academia and industry have begun to study the visions, challenges, technologies, applications, and service requirements for satisfying the future 6G networks in 2030 [3, 7, 17, 37-40].

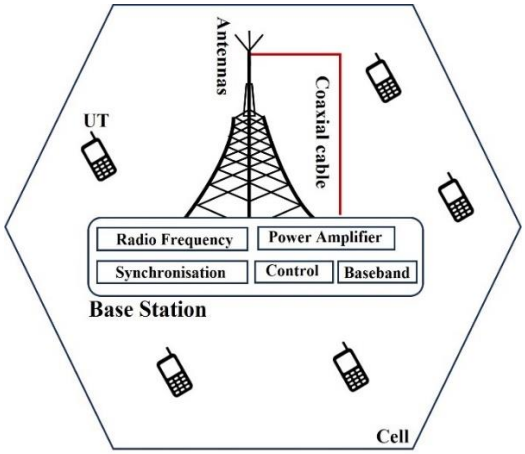
5G communication services have been categorised into three aspects with regard to the enhanced Mobile Broadband (eMBB), massive Machine-Type Communications (mMTC) and Ultra-Reliable Low-Latency Communications (URLLC) [41]. Inheriting these features, 6G services are complemented and characterised in the different application scenarios. For example, the paper [5] classified the application scenarios into five respects: eMBB Plus (eMBB-Plus), Big Communications (BigCom), Secure URLLC (SURLLC), Three-Dimensional Integrated Communications (3D-InteCom) and Unconventional Data Communications (UCDC). The author in [3] proposed the ubiquitous mobile broadband (uMBB) to emphasise the enhanced on-board communications and ubiquitous connection ability, while the ubiquitous mobile ultra-broadband (uMUB) as a new key performance indicator (KPI) of 6G was included in [7]. Merging the mMTC and URLLC, the massive URLLC (mURLLC) service was summarised in [38] which provides a scalable URLLC with massive connectivity. Although these papers had different definitions about the 6G services, they all mentioned the new architecture of the RAN should be explored and MaMIMO is still the key technique for adapting the ultra-dense 6G networks.

2.2 Radio Access Network

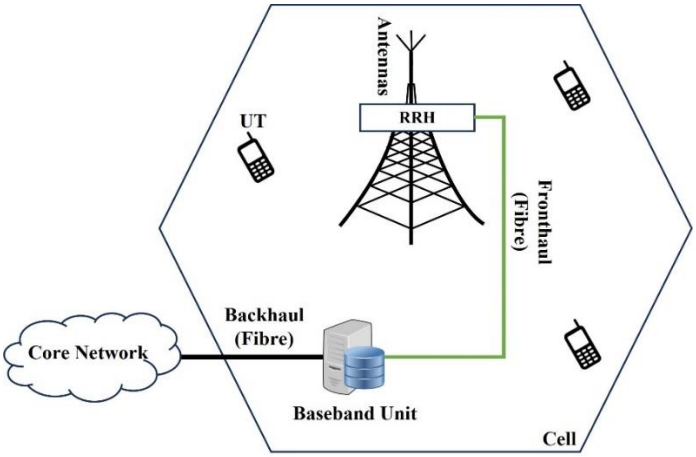
A RAN is a major component of a mobile telecommunications system that enables devices such as mobile phones, computers, and other user terminals (UTs) to access the core network via a radio link. Fundamentally, a RAN architecture consists of two

essential entities: the Radio Unit (RU) and the Baseband Unit (BBU). The RU, acting as a transceiver, includes antennas, amplifiers, analogue-to-digital converters (ADCs) and digital-to-analogue converters (DACs), transmitting and receiving the radio signals. On the other hand, the BBU processes baseband signals, performing functions such as error control, coding, decoding, modulation and demodulation. In traditional RAN, the RU and BBU are integrated into the Base Station (BS), an architecture example of which is given in Fig. 2.1(a). This design was popular in 1G and 2G, where the RU and BBU were installed in a room (or box) beneath the BS tower, connecting to the antennas on top via coaxial cable [41]. With the advent of 3G and 4G, the RU and BBU became separated as illustrated in Fig. 2.1(b). The RU, also referred to as the Remote Radio Head (RRH), is placed at the top of the tower [42], while the BBU can be deployed at a more accessible site, potentially up to 40 km away from the RRH [43]. The link between the RRH and BBU, known as the fronthaul, can be implemented using either optical fibre or microwave [41, 43]. Besides, signals are conveyed to the core network from the BBUs through the backhaul link, which is typically fibre-based. Later, for reducing the power consumption and adapting to the increasing amount of user data, the BBUs are aggregated into a BBU pool which is centralised and virtualised in 4G and beyond. This architecture given in Fig. 2.1(c) is called Centralised RAN or Cloud RAN (C-RAN) [44].

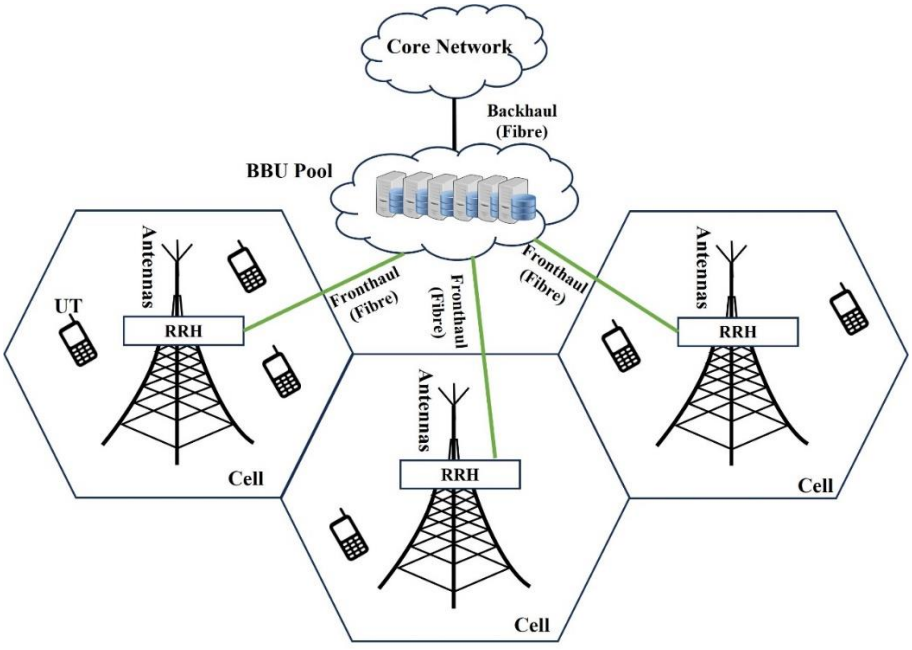
In C-RAN, a BBU pool is shared across multiple cells to serve up to tens of RRHs, leveraging cloud technology to enhance computational capabilities. This approach also mitigates the costs associated with deploying a BS in every individual cell. However, since the fronthaul in C-RAN carries the quantised signals, the capacity must be many times the user data rate [30]. In addition, all processing takes place in the BBU pool leading to a high computational complexity. These disadvantages may result in significant delay which cannot fulfil the requirement of the ultra-reliable low-latency communications (URLLC) in 5G mobile communication service standards [45]. Therefore, a new paradigm called Fog RAN (F-RAN) has been proposed in which the core idea is to move the signal processing back from the “cloud” such as the remote BBU to the network edge, which is nearer to the Access Points (APs) [30]. The related research for the F-RAN has been discussed in [46-48] which indicates the problem described above in the C-RAN has been overcome by alleviating the burdens at edge processing unit (EPU), and hence the latency will be reduced. The architecture of F-RAN is shown in Fig. 2.2. Furthermore, in conventional C-RAN and F-RAN systems,



(a) Traditional BS



(b) BS with RRH



(c) C-RAN with RRHs

Fig. 2.1 Evolution of RAN architecture.

each AP (RRH) only serves its own cell, which can cause the intercellular interference, especially for cell-edge users. Thus, the need for coordination techniques between APs should be considered. For instance, the network coordination was discussed in [49], while paper [50] studied multi-cell MIMO cooperative networks. This approach has also been mentioned as network MIMO in [51]. Additionally, another coordination technique, known as coordinated multipoint (CoMP), proposed in [52], improves both the system's average spectral efficiency (SE) and the data rates at the cell edge.

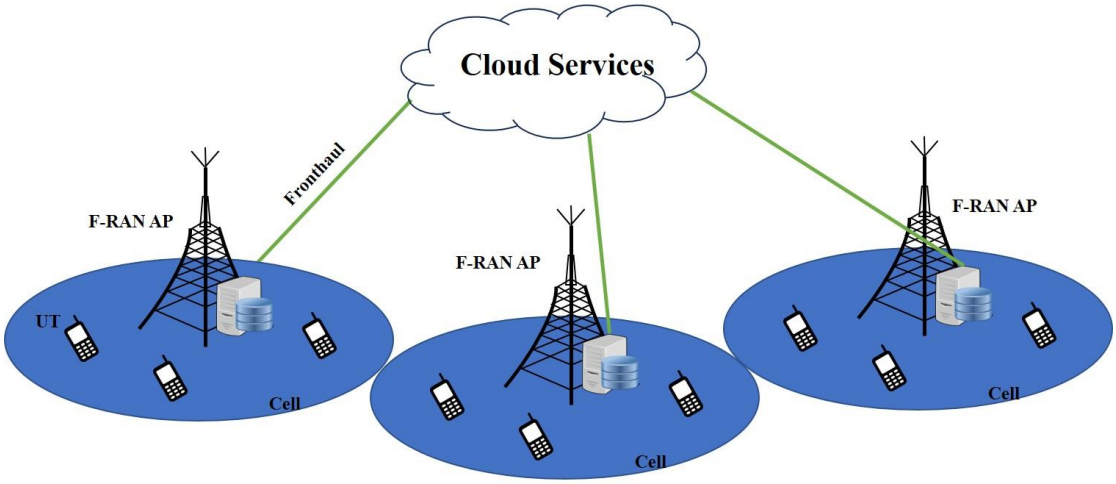


Fig. 2.2 An architecture of Fog RAN system.

More recently, a new concept known as Open RAN, also referred as to O-RAN in some papers, has been introduced by O-RAN Alliance to build upon the principles of openness and intelligence for 5G beyond and 6G [31, 32]. The architecture of Open RAN disaggregates the BS into three functional units: the Central Unit (CU), the Distributed Unit (DU), and the Radio Unit (RU) [53], where the function of the physical layer is split by 3GPP into different options for the DU and RU [53]. Specifically, the CU implements the higher layers of the protocol stack, such as the Radio Resource Control protocol, which is associated with connection establishment and release [54], [55]. In split 7.2x [53], the RU is responsible for the precoding, Fast Fourier Transform (FFT), beamforming, ADC/DAC and radio frequency (RF) functions. Meanwhile, the DU handles the remaining physical layer functions (i.e., modulation, scrambling, part of precoding, etc.), as well as the second layer functions, including Medium Access Control (MAC) and Radio Link Control (RLC). If the DU is separated from the RU and aggregated with the CU in a cloud, such a scheme corresponds to C-RAN. On the other hand, F-RAN relates to the scenario in which the DU and the RU are co-located. Thus,

the physical architecture of the Open RAN can be implemented in a manner similar to C-RAN and F-RAN, depending on the specific functional split options. Furthermore, the fronthaul load must be considered due to the digital transmission between the RU and the DU. To alleviate this burden, compression or quantisation techniques can be applied to remove redundant or less important information. In Open RAN, six fronthaul compression techniques have been released by O-RAN Alliance and European Telecommunications Standards Institute (ETSI) [56]. However, this thesis will primarily discuss the application of quantisation, rather than these compression methods. Additionally, while the RAN Intelligent Controller (RIC) is a key innovation [55] for controlling and optimising RAN functions in Open RAN, this new component largely concerns the higher layer. Therefore, we will not delve into the functionality of RIC in this thesis. Further discussion about how our approach aligns with the physical layer of the Open RAN architecture will be presented in the subsequent chapters.

2.3 Massive MIMO Technique

MaMIMO is a multi-user MIMO (MU-MIMO) technique in which the BS, equipped with significantly more antennas than UTs, serves users simultaneously. The number of antennas for each BS can reach into tens or even hundreds, which is considerably more than the amount used in MU-MIMO. This leads to two notable properties: channel hardening and favourable propagation. The terminology “channel hardening” was coined in a 2004 paper [57], revealing that as the number of antennas at the BS increases, the instantaneous data rate tends towards its average. This phenomenon occurs because enhanced spatial diversity can significantly mitigate the effects of small-scale fading [58], which typically generates the random fluctuation in the channel gain. As a result, it creates a “hardened” channel where the channel fading becomes deterministic. Favourable propagation was defined in [58] that the channel vectors between the BS and UTs are mutually orthogonal in MaMIMO. With this property, the maximum sum-capacity can be realised using simple linear processing [58, 59].

MaMIMO, initially proposed in [13], has been recognised as a key technique to meet the rapidly growing demand for data rates in 5G wireless communications [14], [15] and anticipated future 6G networks. Besides the previously mentioned benefits, MaMIMO can be more efficient in a time-division duplexing (TDD) mode by leveraging channel reciprocity [58]. In [60], the author investigated the application of

MaMIMO in TDD operation under various propagation environments. The study concluded that, due to the channel hardening or the use of a blind channel estimation scheme, there is no need for downlink pilots to obtain the channel state information (CSI) [60]. However, in a frequency-division duplexing (FDD) mode, different frequency bands are used for uplink and downlink [61], resulting in different channels. Therefore, acquiring the CSI of the downlink requires the UTs to feed the estimation information back to the BS on the uplink, which proportionally increases the transmission overhead with respect to the number of BS antennas.

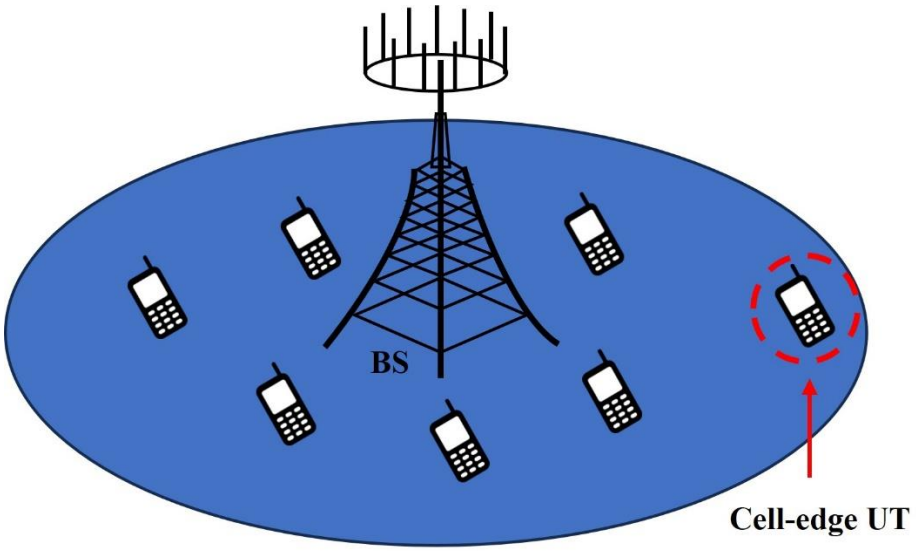


Fig. 2.3 A single-cell MaMIMO system.

Conventional MaMIMO is implemented with the co-located antennas at the BS, either within a single cell or across multiple cells. In the single-cell scenario, a multiple-antenna BS, located at the network centre, communicates with distributed single-antenna UTs throughout the entire cell. The cell-edge UTs experience poor performance due to the high pathloss being far from the BS. This scheme is not realistic in the scenario where all UTs in the network are served by only one BS. The system model is demonstrated in Fig. 2.3. In contrast, multi-cell MaMIMO can be realised by dividing the whole network into several cells, each possessing one BS at the centre, exclusively serving the UTs within its respective cell. The corresponding architecture is presented in Fig. 2.4. This scheme, however, raises the issue of inter-cell interference, especially for the cell-edge UTs at which the signals received from the adjacent BSs have similar power. To address this issue, BSs can jointly serve UTs across multiple cells by employing a cooperative network [49-51] and CoMP technique [52]. These can

be implemented either by interconnecting the BSs or by linking all BSs to the CPU via a fronthaul. However, the network coverage in co-located MaMIMO may not be very large due to the path loss, which leads to the limited data rate for UTs at the cell edge.

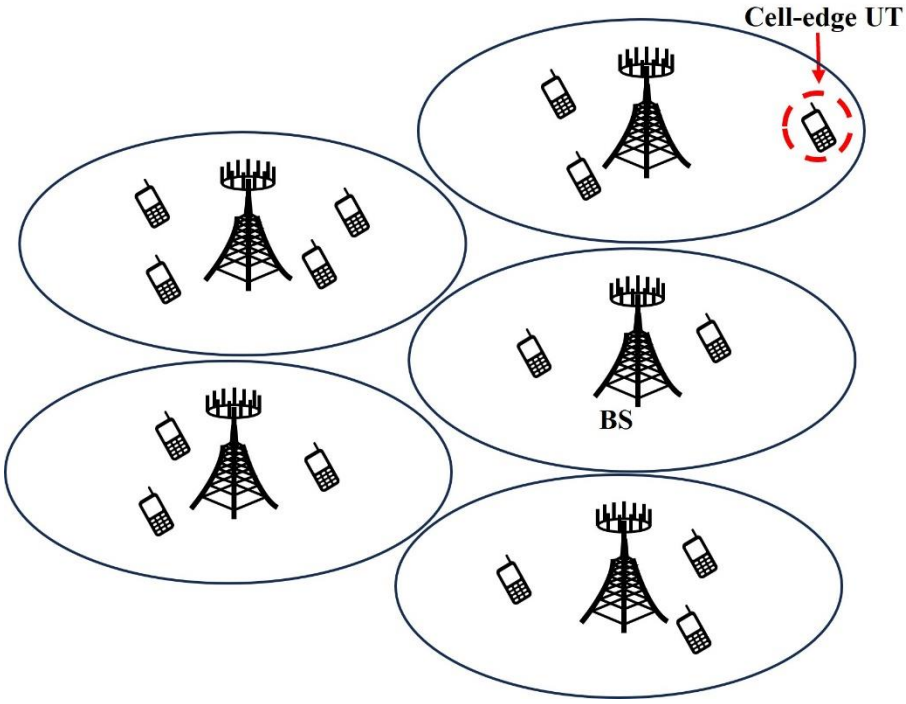


Fig. 2.4 A multi-cell MaMIMO system.

2.4 Cell-free Massive MIMO

In MaMIMO, apart from the co-located strategy, the antennas can also be deployed in a distributed manner [62]. This scheme, known as distributed MaMIMO or cell-free MaMIMO [14, 16], involves a large number of APs simultaneously serving a smaller number of single-antenna UTs. All APs are connected to the CPU via links, defined as "backhaul" by Ngo [16], a term equivalent to "fronthaul" as used in C-RAN and F-RAN. Note that in this thesis, we will refer to these links as fronthaul connections. Since each UT is served by multiple APs, the concept of a "cell" or "cell boundary" is abolished, thereby eliminating the cell-edge effect. Furthermore, distributing APs across the network significantly enhances the overall coverage of connectivity. Then, we consider an uplink transmission cell-free MaMIMO system where there are M APs and K UTs, each equipped with a single antenna, randomly distributed in the network, which is illustrated in Fig. 2.5.The channel coefficient from the k-th UT to the m-th AP is given by:

$$g_{mk} = \beta_{mk}^{1/2} h_{mk} \tag{2.1}$$

where $h_{mk} \sim \mathcal{CN}(0,1)$ is the small-scale fading coefficient and β_{mk} denotes the large-scale fading coefficient which only relates to the shadowing and the distance between the UT and AP [63]. In real life, the large-scale fading varies much slower than small-scale fading when the UT moves with slow speed and short distance [64]. Therefore, the large-scale fading is assumed to be known at the APs or the CPU by simply averaging the power of received signals over a long period of time.

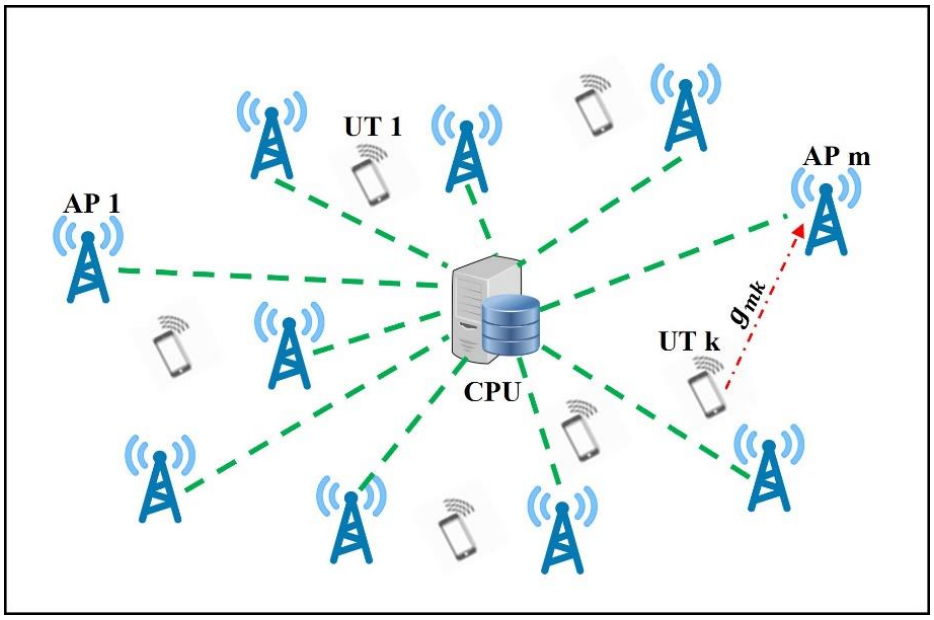


Fig. 2.5 The architecture of cell-free MaMIMO (green dash line denotes "backhaul" [16], red dash line with arrow represents the channel, and g_{mk} is the channel coefficient from the k-th UT to the m-th AP).

2.4.1 Channel Estimation in Cell-free MaMIMO

For obtaining the CSI, UTs can send a set of known training data to the APs on the uplink. Then, the channel is estimated at the local AP by using the combining technique. Alternatively, the received signals can be collected at the CPU via the fronthaul links from all APs and are exploited to estimate the channel. For the downlink CSI, the channel reciprocity can be used in cell-free MaMIMO with the TDD operation. Note that throughout this thesis the TDD mode is adopted. Let $\boldsymbol{\varphi}_k \in \mathbb{C}^{\tau_p \times 1}$, where $\|\boldsymbol{\varphi}_k\|^2 = \tau_p$, be the pilot sequence allocated to the k-th UT. The subscript "p" refers to the pilot transmission. Then, the received pilot signal $\boldsymbol{y}_p$ at the m-th AP is given by:

$$\boldsymbol{y}_{p,m} = \sum_{k=1}^{K} g_{mk}\boldsymbol{\varphi}_k^T + \boldsymbol{z}_{p,m} \tag{2.2}$$

where $\boldsymbol{z}_{p,m} \in \mathbb{C}^{1\times\tau_p}$ is the noise $\boldsymbol{z}_p$ at the m-th AP whose elements are independent and identically distributed (i.i.d) $\mathcal{CN}(0,\sigma_z^2)$ random variables (RVs). $(\cdot)^T$ denotes the transpose. Based on the received pilots, $\check{y}_{p,mk}$ is given by calculating the projection $\boldsymbol{y}_{p,m}$ onto $\boldsymbol{\varphi}_k^*$ [14, 16] and scaling by the factor τ_p, where $(\cdot)^*$ represents the complex conjugate.

$$\begin{aligned}\check{y}_{p,mk} &= \frac{1}{\tau_p}\boldsymbol{y}_{p,m}\boldsymbol{\varphi}_k^* \\ &= g_{mk} + \frac{1}{\tau_p}\sum_{k'\neq k}^{K} g_{mk'}\boldsymbol{\varphi}_{k'}^T\boldsymbol{\varphi}_k^* + \boldsymbol{z}_{p,m}\boldsymbol{\varphi}_k^*\end{aligned} \tag{2.3}$$

Although (2.3) can be used as the estimated channel, it is not sufficient when some pilot sequences are not pairwise orthogonal. Hence, the minimum mean square error (MMSE) estimation can be adopted:

$$\hat{g}_{mk} = \frac{\mathbb{E}\{g_{mk}\check{y}_{p,mk}^*\}}{\mathbb{E}\left\{\left|\check{y}_{p,mk}\right|^2\right\}}\check{y}_{p,mk} = c_{mk}\check{y}_{p,mk} \tag{2.4}$$

where the MMSE weight c_{mk} is:

$$c_{mk} = \frac{\beta_{mk}}{\beta_{mk} + \frac{1}{\tau_p^2}\sum_{k'\neq k}^{K}\beta_{mk'}\mathbb{E}\left\{\left|\boldsymbol{\varphi}_{k'}^T\boldsymbol{\varphi}_k^*\right|^2\right\} + \frac{\sigma_z^2}{\tau_p}} \tag{2.5}$$

If $K < \tau_p$, then all pilot sequences are pairwise orthogonal that $\boldsymbol{\varphi}_{k'}^T\boldsymbol{\varphi}_k^* = 0$ as $k' \neq k$, thus the second term in (2.3) disappears. On the contrary, if $K \geq \tau_p$ the pilot sequences will be re-used which results in the pilot contamination effect. In addition, the MSE of the channel estimation is calculated by:

$$\begin{aligned}&\sigma_e^2 = \mathbb{E}\{|\tilde{g}_{mk}|^2\} = \mathbb{E}\{|g_{mk} - \hat{g}_{mk}|^2\} \\ &= \mathbb{E}\{(g_{mk} - \hat{g}_{mk})(g_{mk}^* - \hat{g}_{mk}^*)\} \\ &= \mathbb{E}\{g_{mk}^2 + \hat{g}_{mk}^2 - (\hat{g}_{mk}^* + \tilde{g}_{mk}^*)\hat{g}_{mk} - (\hat{g}_{mk} + \tilde{g}_{mk})\hat{g}_{mk}^*\} \\ &= \mathbb{E}\{g_{mk}^2 - \hat{g}_{mk}^2\}\end{aligned} \tag{2.6}$$

$$= (1 - c_{mk})\beta_{mk}$$

where we assume $\tilde{g}_{mk}$ is uncorrelated with $\hat{g}_{mk}$ due to the use of MMSE estimation.

2.4.2 Uplink Data Transmission in Cell-free MaMIMO

On the uplink, all UTs send the data x, where the variance of the data $\sigma_x^2 = 1$, to the m-th AP. The received signal y_u at the m-th AP is presented as

$$y_{u,m} = \sum_{k=1}^{K} g_{mk} x_k + z_{u,m} \tag{2.7}$$

where $z_{u,m}$ is the noise z_u at the m-th AP. The subscript "u" refers to the uplink transmission. In order to recover the data, the data detection can be done at the AP by applying the combination technique. As stated in [14], we introduce the maximum-ratio combining (MRC) to estimate data. First, the received signal at the m-th AP is multiplied by the MRC weight $\hat{g}_{mk}^*$. Then, the weighted signals from all APs are sent to the CPU for detection via the fronthaul links. Other combining techniques such as zero forcing (ZF) and MMSE also can be used.

$$\begin{aligned} r_{u,k} &= \sum_{m=1}^{M} \hat{g}_{mk}^* \, y_{u,m} \\ &= \sum_{m=1}^{M} \sum_{k'=1}^{K} \hat{g}_{mk}^* g_{mk'} x_{k'} + \sum_{m=1}^{M} \hat{g}_{mk}^* z_{u,m} \end{aligned} \tag{2.8}$$

where r_u denotes the received signal on the uplink at the CPU and the subscript "k" is an index which indicates the transmitted data from the k-th user x_k can be detected from the signal $r_{u,k}$.

2.5 The Basics of Quantisation

On the uplink of RAN, the signals are collected at the receiver (AP) and transferred to the CPU by fibre-based connections. This connection, named fronthaul, was first defined in C-RAN and it is also known as backhaul in cell-free MaMIMO. In this processing, quantisation is applied since the transferred signal must be digital which is

discrete in practice. In this sub-chapter, we will introduce the Bussgang decomposition to analyse the non-linear quantisation.

The Bussgang decomposition as the result of Bussgang's theorem has been reported in [65], however the extension of the explicit decomposition application was researched in communications [66] and in the Massive MIMO field by [67-69]. The original Bussgang theorem states that the crosscorrelation of two Gaussian distributed signals, one of which undergoes the nonlinear amplitude distortion, is identical to the cross-correlation of the two initial signals scaled by a factor [65]. Further, if we consider a non-linear system where there is only one input Gaussian distributed signal, the crosscorrelation of input and output will be proportional to the autocorrelation of the input signal. Based on these contents, let x be a real Gaussian distributed signal and the output of a non-linear quantizer $y(x)$ can be expressed by a linear Bussgang decomposition function:

$$y(x) = \alpha x + \delta \tag{2.9}$$

where α is the proportionality coefficient which has been described above, and δ denotes the quantisation error or distortion, which is uncorrelated with the input signal x. The linear coefficient is given by:

$$\alpha = \frac{\mathbb{E}\{xy(x)\}}{\mathbb{E}\{x^2\}} = \frac{1}{P_x}\int_{-\infty}^{\infty} xy(x) f_X(x) dx \tag{2.10}$$

where $P_x = \mathbb{E}\{|x|^2\} = \mathbb{E}\{x^2\}$ is the power of x and the absolute value can be omitted as x is assumed to be a real number, and $f_X(x)$ is the probability distribution function (pdf) of x. Then, it is useful to define the factor γ as the ratio between the output power and the input power of the quantiser [66]:

$$\gamma = \frac{\mathbb{E}\{|y(x)|^2\}}{\mathbb{E}\{x^2\}} = \frac{1}{P_x}\int_{-\infty}^{\infty} |y(x)|^2 f_X(x) dx \tag{2.11}$$

Thus, the power of the distortion can be represented by:

$$\begin{aligned} \mathbb{E}\{|\delta|^2\} &= \mathbb{E}\{|y(x) - \alpha x|^2\} \\ &\overset{(a)}{=} \mathbb{E}\{|y(x)|^2\} - \alpha^2 P_x \\ &= (\gamma - \alpha^2) P_x \end{aligned} \tag{2.12}$$

where step (a) is true due to the uncorrelation between δ and x. We note that the power of the distortion should be equal to or larger than zero in practice, hence $\gamma \geq \alpha^2$. Then, for analysing the performance of the quantisation, the signal-to-distortion and noise ratio (SDNR) of the quantiser is defined as:

$$SDNR = \frac{\mathbb{E}\{|\alpha x|^2\}}{\mathbb{E}\{|\delta|^2\}} = \frac{\alpha^2 P_x}{(\gamma - \alpha^2)P_x} = \frac{\alpha^2}{\gamma - \alpha^2} \tag{2.13}$$

where the numerator of (2.13) is the power of the desired signal in (2.9). To achieve the best performance of quantisation, we need to reduce the effect of distortion, in other words, the SDNR should be maximised. According to [28, 70], the mid-rise uniform quantiser $y(x)$ is given by:

$$y(x) = \begin{cases} -\frac{L-1}{2}\Delta & \mathrm{x} \leq -\left(\frac{L}{2} - 1\right)\Delta \\ \left(l + \frac{1}{2}\right)\Delta & l\Delta < \mathrm{x} \leq (l+1)\Delta, l = -\frac{L}{2} + 1 \ldots \frac{L}{2} - 2 \\ \frac{L-1}{2}\Delta & x > \left(\frac{L}{2} - 1\right)\Delta \end{cases} \tag{2.14}$$

where $L = 2^b$ is the number of quantisation steps, b is the number of quantisation bits and Δ refers to the quantisation step size or quantisation step interval. An example of the uniform quantiser with $L = 8$ steps and step interval $\Delta = 1$ is shown in Fig. 2.6.

Next, we substitute (2.14) into (2.10) and (2.11), assuming the input signal $x \sim \mathcal{N}(0, \sigma_x^2)$, and have the results as:

$$\alpha = \frac{\Delta}{\sqrt{2\pi}\sigma_x}\left(\sum_{l=1}^{\frac{L}{2}-1} 2e^{-\frac{l^2\Delta^2}{2\sigma_x^2}} + 1\right) \tag{2.15}$$

$$\gamma = \frac{\Delta^2}{\sigma_x^2}\left(4\sum_{l=1}^{\frac{L}{2}-1} lQ\left(\frac{l\Delta}{\sigma_x}\right) + \frac{1}{4}\right) \tag{2.16}$$

where $Q(\cdot)$ denotes the *Q-function.* The proof is omitted in this thesis since it was already provided by [28]. Finally, we replace α and γ in (2.13) by (2.15) and (2.16), and solve the SDNR maximisation problem by finding the optimum step interval.

$$\begin{aligned}\Delta_{opt} &= \arg\max_{\Delta}\left(\frac{\alpha^2}{\gamma-\alpha^2}\right) = \arg\max_{\Delta}\left(\frac{1}{\frac{\gamma}{\alpha^2}-1}\right)\\ &= \arg\max_{\Delta}\left(\frac{\alpha^2}{\gamma}\right)\\ &= \arg\max_{\Delta}\left(\frac{\frac{\Delta^2}{2\pi\sigma_x^2}\left(\sum_{l=1}^{\frac{L}{2}-1} 2e^{-\frac{l^2\Delta^2}{2\sigma_x^2}}+1\right)^2}{\frac{\Delta^2}{\sigma_x^2}\left(4\sum_{l=1}^{\frac{L}{2}-1} lQ\left(\frac{l\Delta}{\sigma_x}\right)+\frac{1}{4}\right)}\right)\\ &= \arg\max_{\Delta}\left(\frac{\left(\sum_{l=1}^{\frac{L}{2}-1} 2e^{-\frac{l^2\Delta^2}{2\sigma_x^2}}+1\right)^2}{4\sum_{l=1}^{\frac{L}{2}-1} lQ\left(\frac{l\Delta}{\sigma_x}\right)+\frac{1}{4}}\right)\end{aligned} \tag{2.17}$$

Note that on the second equality of (2.17), we also have the constraint $\gamma \geq \alpha^2$.

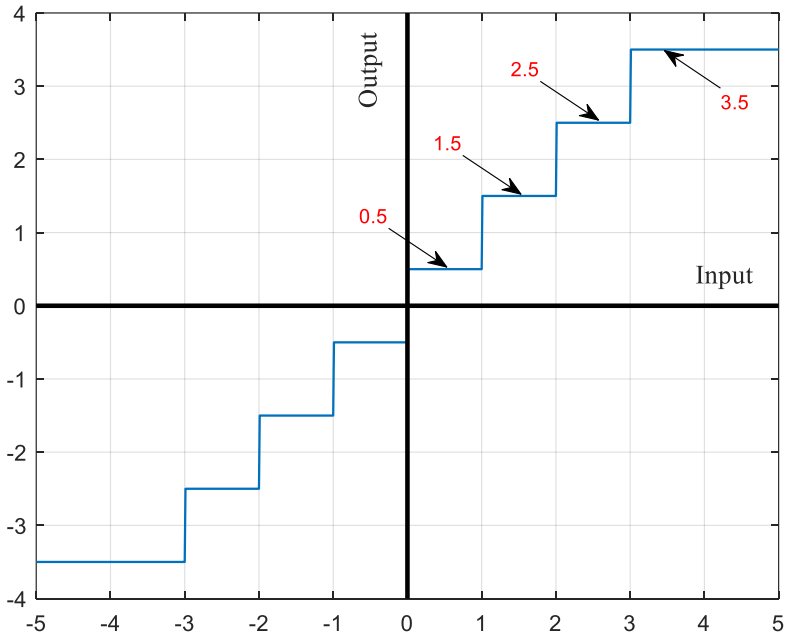


Fig. 2.6 Quantisation function of 8 steps mid-rise uniform quantiser with the step interval is one (The quantiser is symmetric with respect to the origin).

In Fig. 2.7, an example of finding the optimum step interval for maximising the SDNR with 4 quantisation steps is illustrated, where the input of the quantiser follows the standard normal distribution and we vary the step interval by the range from 0 to

1.5, to show the corresponding values α, γ, α^2/γ and SDNR. In Table 2.1, we provide a comprehensive list of the linear coefficient α and $\gamma - \alpha^2$ with respect to the optimum step interval for quantisation bits up to 12, under the standard normal distribution input. This table is generated by exploring the quantisation step interval Δ to maximise the SDNR using a tool such as Mathematica.

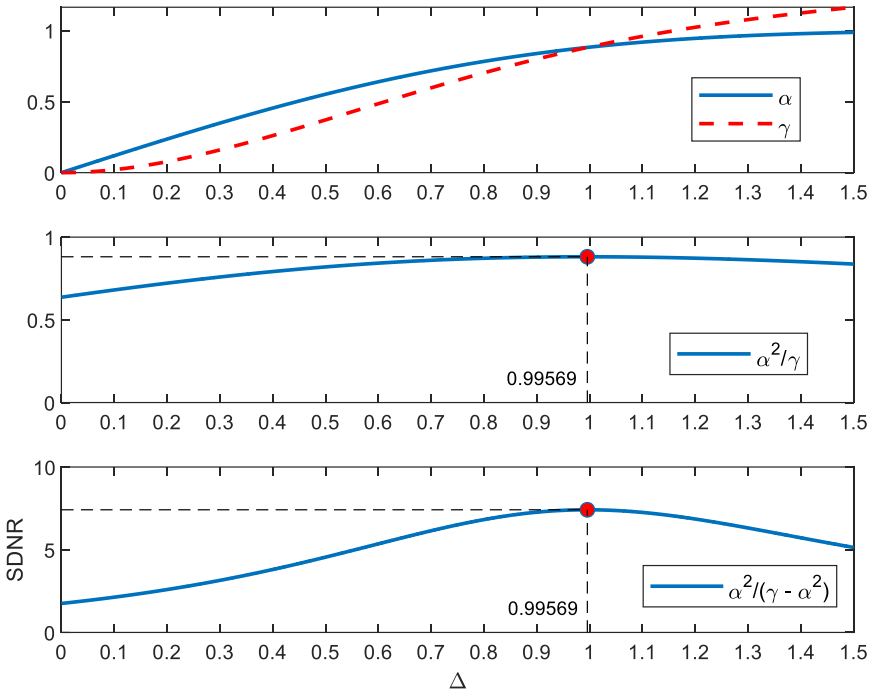


Fig. 2.7 An example of the optimum step interval by maximising SDNR, $L = 4$, the input of quantiser follows the standard normal distribution where the variance is one.

Table 2.1 Optimum step interval and Bussgang parameters

No. of bits, b	Δ_{opt}	α	$\gamma - \alpha^2$
2	0.995687	0.881154	0.104722
3	0.586019	0.96256	0.0360379
4	0.335201	0.988457	0.0114096
5	0.188139	0.996505	0.00348299
6	0.104063	0.99896	0.00103896
7	0.0568677	0.999696	0.00030424
8	0.0307624	0.999912	0.0000876785
9	0.016499	0.999975	0.0000249184
10	0.00878546	0.999993	6.99696×10^{-6}
11	0.00464984	0.999998	1.94441×10^{-6}
12	0.00244841	0.999999	5.35536×10^{-7}

Chapter 3

Decentralised Distributed Massive MIMO

3.1 Introduction

The demand for higher data rate is expected to increase exponentially in the next decade. To cater to the huge number of wireless devices requiring low latency and high spectral efficiency (SE), an architecture with adaptive deployment and related signal processing needs to be considered. Cell-free massive multiple-input multiple-output (MaMIMO), as discussed in the literature, has been identified as a promising approach that improves cell-edge services and increases per-user capacity by exploiting distributed MIMO technique. However, challenges such as the cluster-edge effect, network scalability and latency exist, which limit the system's performance. While solutions for cell-free MaMIMO have been proposed, they may be complicated or inflexible to implement. We will discuss these issues in detail in the subsequent sub-chapters. Therefore, the fundamental idea of this thesis, which is distinguished from the concepts of cell-free MaMIMO or cloud radio access network (C-RAN), is to explicitly define where the processing occurs, and to move it to the network edge from a cloud or central location.

In an earlier paper [30] it has been shown that users at the cluster edge of the cell-free MaMIMO network do not perform well, and the signal-to-interference ratio (SIR) might be negative. The reason is that the access points (APs) close to the edge of the adjacent network clusters are uncoordinated with the current cluster, resulting in inter-cluster interference. To solve this problem, we define adjacent entities, referred to as coordination regions, which naturally overlap. This means that APs near the edge of a cluster may be coordinated by more than one edge processing unit (EPU). Leveraging the features of the technologies mentioned in Chapter 2, we combine the decentralised signal processing technique with cell-free MaMIMO and define the coordination regions, proposing a new paradigm called "Decentralised Distributed Massive MIMO"

(DD-MaMIMO). This system is the same as “Fog Massive MIMO” in [30]. We chose to change the name because the terminology of “Fog Massive MIMO” was not widely used five years ago. However, in recent years, Fog Massive MIMO has been used in [33] to indicate a network with a “fog”-like density of remote radio heads (RRHs), where UTs autonomously connect to the convenient RRHs, which is different from the usage in [30]. Additionally, the term “fog” is often associated with fog/edge computing, which concerns caching within the network. Furthermore, the terms “decentralised cell-free MaMIMO” [26] and “decentralised network” in [23] have been used to describe techniques where channel estimation and data detection are locally processed at the AP, which is different from our definition. Therefore, we redefine “fog” that the signal processing locates at the network edge referring to as the decentralisation and combine it with cell-free MaMIMO which is distributed.

A rudimentary performance analysis for the uplink DD-MaMIMO was discussed in [30]. This analysis adopted least squares (LS) channel estimation with maximum ratio combining (MRC) for data detection. It also exploited the assumption of channel hardening, which might not be valid, particularly when the number of antennas at the APs is fewer or the path-loss exponent is greater [19]. Therefore, in this chapter, we will discard the assumption, mainly taking account of the general case in which each AP is equipped with either a single antenna or just a few antennas. We will employ both MRC and minimum mean square error (MMSE) techniques. To ensure a fair and reasonable comparison, we will maintain the same level of density for the AP and user terminal (UT), and apply the same propagation model as [24], which provided a comprehensive analysis of cell-free MaMIMO, to DD-MaMIMO. The power control will not be considered in this thesis; instead, we will assign equal power to all transmitted UTs.

In the subsequent sub-chapters, we will firstly introduce the system model of DD-MaMIMO, giving the initial setup parameters. Then, channel estimation processed in the decentralised entity which is the EPU will be described, by collecting the received pilot signals from all APs within the coordination region via the unlimited capacity fronthaul links. To fit with the practice, we will next propose a pilot allocation scheme which concerns the areas consisting of six adjacent clusters and the current cluster. Furthermore, we will discuss the data detection which also occurs at the EPU. Based on the different estimation methods, a general expression for calculating the achievable uplink SE is presented. Additionally, we will show that our DD-MaMIMO adapts to

the scalable network and does not consume extra overhead. Finally, we will give some numerical results and provide a comprehensive analysis.

3.2 System Model

In this sub-chapter, we consider a DD-MaMIMO system with M_{coor} randomly distributed APs in the coordination region serving K_{serv} UTs in the service region, which was originally defined in [30], see Fig. 3.1. The network as a whole contains many EPUs (there are two illustrated in Fig. 3.1), each at the centre of a coordination region (which is a circle area of radius r_{coor}) within which all APs are coordinated by and connected to the EPU. The coordination regions have an overlap between the adjacent EPUs, which indicates that an AP might be jointly coordinated by more than one EPU. According to the definition, the UTs served by the AP within this overlapping area can send their data to any EPU that serves this overlapping coordination region. However, in this sub-chapter we define a hexagon as the service region for each EPU such that the data from all served UTs in the hexagonal area are processed by this EPU.

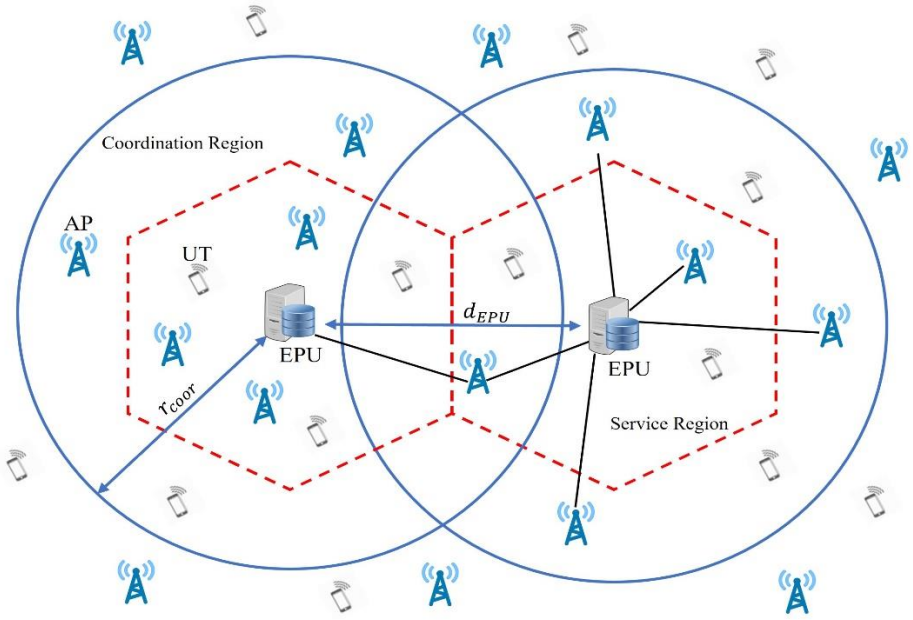


Fig. 3.1 The architecture of DD-MaMIMO, including EPUs with their service regions (red dash hexagon) and coordination regions (blue solid circle), APs with fronthaul links (black solid line), and UTs.

The key difference is that in the conventional cell-free MaMIMO [14], all APs connected to one CPU influence the size of the serving area. But EPUs involving

multiple APs in Fig3.1 produce the size of the serving area. We assume the distance between the EPUs (the centre of the coordination region) represented by d_{EPU}, is 300 m here, which adapts to the ultra-dense network. Besides, in this chapter we do not discuss the details of the fronthaul which connects the AP to the EPU, only assuming it is error-free and offers no bandwidth limitation. To analyse the performance of the system, we initially assume a scenario in which the UTs with only a single antenna and APs equipped with multiple antennas are uniformly distributed in the network area with the density of ρ_u and ρ_A, respectively. Moreover, we assume $\boldsymbol{g}_{mk}$ represents the flat-fading channel coefficients between the k-th UT in the service region of a given EPU and the m-th AP with N_r antennas in the coordination area. Here the gain vector $\boldsymbol{g}_{mk} = \left[g_{m1k}\ g_{m2k} \dots g_{mN_rk}\right]^T$ is given by:

$$\boldsymbol{g}_{mk} = \beta_{mk}^{1/2}\boldsymbol{h}_{mk} \tag{3.1}$$

where $\boldsymbol{h}_{mk} = \left[h_{m1k}\ h_{m2k} \dots h_{mN_rk}\right]^T$ represents the small-scale Rayleigh fading of the elements, as independent and identically distributed (i.i.d) complex Gaussian $\mathcal{CN}(0,1)$ random variables (RVs). We assume $\boldsymbol{h}_{mk}$ remains the same over the period of coherence interval. β_{mk} denotes the large-scale fading including the one-slope path loss applying the 3GPP Urban Microcell Model, as follows [24, 71]:

$$\beta_{mk}[\text{dB}] = -22.7 - 26\log_{10}\left(\frac{f_c}{1\ \text{GHz}}\right) - 36.7\log_{10}\left(\frac{d_{mk}}{1\ \text{m}}\right) + F_{mk} \tag{3.2}$$

where f_c is the carrier frequency (assumed to be 2 GHz), d_{mk} is the distance between the m-th AP and the k-th user, and $F_{mk} \sim \mathcal{CN}(0,16)$ represents the shadow fading. In this thesis, we assume that the shadow fading coefficients from one AP to all UEs are uncorrelated. Furthermore, the correlation among the channels will not be discussed. Note that the large-scale fading coefficients of multiple antennas at the same AP are equal because they are at the same distance from the UT, and we assume they are also affected by shadow fading in the same way.

As in [14, 20, 30], we adopt time-division duplex (TDD) operation for the uplink and downlink transmission. The coherence interval τ_c is defined as the period over which the channel is approximately constant. Based on the feature of TDD which is the

channel reciprocity, we can obtain the downlink channel knowledge from the uplink channel estimation, thus the pilots do not need to be transmitted on the downlink. The coherence interval is then divided into three phases: uplink pilot sequence (τ_p), uplink data transmission (τ_u), and downlink data transmission (τ_d), i.e., $\tau_c = \tau_p + \tau_u + \tau_d$. The pilot length should be less than the coherence interval, and cannot be too long which will decrease the efficiency of data transmission. In this chapter, we focus only on the uplink but unlike [26] and [72] where $\tau_p < K$ (the number of all UTs) in practical cell-free MaMIMO, we assume the pilot length is larger than the number of UTs in the coordination region ($\tau_p \geq K_{coor}$), ensuring the random pilots for all users in the coordination region are fully orthogonal. That is because in the ultra-dense network, there is not a large number of UTs to be served by a single EPU. For example, to support holographic video streaming within the ultra-dense network, the computational complexity can be reduced by separating the users into many EPUs. Then, each EPU only server a few users which also increases the data rate and reduces the latency. In addition, K_{coor} can be limited by reducing the coordination region and service region. The area of coordination region and service region are given by: $S_{coor} = \pi r_{coor}^2$ and $S_{serv} = \frac{1}{2} \times \frac{d_{EPU}}{\sqrt{3}} \times \frac{d_{EPU}}{2} \times 6 = \frac{\sqrt{3}}{2} d_{EPU}^2$, respectively, hence on average $K_{coor} = \rho_u \pi r_{coor}^2$ and $K_{serv} = \frac{\sqrt{3}}{2} \rho_u d_{EPU}^2$.

3.3 Channel Estimation

To estimate the channel coefficients, we assume that β_{mk} is known by all APs and EPUs. We suppose that the pilot sequence for the k-th user is $\boldsymbol{\varphi}_k \in \mathbb{C}^{\tau_p \times 1}$ with $\|\boldsymbol{\varphi}_k\|^2 = \tau_p$, and the pilots outside coordinated region are randomly assigned. The detail of pilot sequence will be discussed in the sub-chapter 4.3. Note that, as mentioned in the system model sub-chapter, the UTs in the coordination region are assigned fully orthogonal pilots. Here, the term “randomly assigned” refers to the pilot assignment chosen from that set of pilot sequences randomly. This means that interfering users outside the coordination region might transmit the same pilots as the served users. The received pilot matrix $\boldsymbol{Y}_p$ at the m-th AP $\boldsymbol{Y}_{p,m} \in \mathbb{C}^{N_r \times \tau_p}$ then is given by:

$$\boldsymbol{Y}_{p,m} = \sum_{k=1}^{K_{serv}} \boldsymbol{g}_{mk}\boldsymbol{\varphi}_k^T + \sum_{k=1}^{K_{int}} \boldsymbol{g}_{i,mk}\boldsymbol{\varphi}_{i,k}^T + \boldsymbol{Z}_{p,m} \tag{3.3}$$

where the matrix $\boldsymbol{Z}_{p,m} \in \mathbb{C}^{N_r \times \tau_p}$ is the noise at the m-th AP, the elements of $\boldsymbol{Z}_{p,m}$ are i.i.d $\mathcal{CN}(0, \sigma_z^2)$ RVs, $\boldsymbol{g}_{i,mk}$ denotes the channel coefficient from interfering users not coordinated by this EPU, defined by the same form as (3.1), and $\boldsymbol{\varphi}_{i,k}$ represents the interfering pilot sequence for the k-th interfering UT. Note that the subscripts "p" and "i" refer to the pilot transmission and interference, respectively. They are not indices. To obtain a good channel state information (CSI) between the k-th UT and the m-th AP, we first multiply the received pilots $\boldsymbol{y}_{p,mn}$, which is the n-th row vector of $\boldsymbol{Y}_{p,m}$, by the conjugate of the pilot sequence $\boldsymbol{\varphi}_k^*$. We then take the LS estimate of the channel by finding $\breve{g}_{mnk}$ to minimise the squared deviation of $\left\|\boldsymbol{y}_{p,mn}\,\boldsymbol{\varphi}_k^* - \tau_p \breve{g}_{mnk}\right\|^2$. Since the smallest value is zero, the LS estimation is given by:

$$\breve{g}_{mnk} = \frac{1}{\tau_p}\boldsymbol{y}_{p,mn}\,\boldsymbol{\varphi}_k^* = g_{mnk} + \frac{1}{\tau_p}\sum_{k'=1}^{K_{int}} g_{i,mnk'}\boldsymbol{\varphi}_{i,k'}^T\boldsymbol{\varphi}_k^* + \frac{1}{\tau_p}\boldsymbol{z}_{p,mn}\boldsymbol{\varphi}_k^* \tag{3.4}$$

where $\boldsymbol{z}_{p,mn}$ is the n-th row vector of the noise matrix. This estimated value has a good performance, if all interfering pilots are orthogonal to the pilots within the coordination region. However, this is not reasonable and practical in real life due to the shortage of orthogonal pilots. Hence, the MMSE channel estimation can be used to improve the estimation performance. The MMSE weight c_{mnk}, applied for the k-th UT to the n-th antenna of the m-th AP, is obtained by equating the first derivative of MSE $\bar{\epsilon}_{mnk} = \mathbb{E}\{|g_{mnk} - c_{mnk}\breve{g}_{mnk}|^2\}$ with respect to c_{mnk} to zero.

$$\begin{aligned} 0 &= \frac{\partial \bar{\epsilon}_{mnk}}{\partial c_{mnk}} \\ &= \frac{\partial}{\partial c_{mnk}}\left(\mathbb{E}\{|g_{mnk}|^2\} - \mathbb{E}\{g_{mnk}c_{mnk}^*\breve{g}_{mnk}^*\} - \mathbb{E}\{g_{mnk}^*c_{mnk}\breve{g}_{mnk}\}\right. \\ &\qquad \left. + \mathbb{E}\{|c_{mnk}\breve{g}_{mnk}|^2\}\right) \\ &= -\mathbb{E}\{g_{mnk}^*\breve{g}_{mnk}\} + c_{mnk}^*\mathbb{E}\{|\breve{g}_{mnk}|^2\} \end{aligned} \tag{3.5}$$

Thus, the MMSE weight is

$$
\begin{aligned}
c_{mnk} &\overset{(a)}{=} \frac{\mathbb{E}\{g_{mnk}\breve{g}_{mnk}^{*}\}}{\mathbb{E}\{|\breve{g}_{mnk}|^2\}} \\
&= \frac{\mathbb{E}\{|g_{mnk}|^2\}}{\mathbb{E}\left\{|g_{mnk}|^2 + \frac{1}{\tau_p^2}\sum_{k'=1}^{K_{int}}\left|g_{i,mnk'}\boldsymbol{\varphi}_{i,k'}^{T}\boldsymbol{\varphi}_{k}^{*}\right|^2 + \frac{1}{\tau_p^2}\left|\boldsymbol{z}_{p,mn}\boldsymbol{\varphi}_{k}^{*}\right|^2\right\}} \\
&= \frac{\beta_{mnk}}{\beta_{mnk} + \frac{1}{\tau_p^2}\sum_{k'=1}^{K_{int}}\beta_{i,mnk'}\,\mathbb{E}\left\{\left|\boldsymbol{\varphi}_{i,k'}^{T}\boldsymbol{\varphi}_{k}^{*}\right|^2\right\} + \frac{\sigma_z^2}{\tau_p}} \\
&\overset{(b)}{=} \frac{\beta_{mnk}}{\beta_{mnk} + \frac{1}{\tau_p}\sum_{k'=1}^{K_{int}}\beta_{i,mnk'} + \frac{\sigma_z^2}{\tau_p}}
\end{aligned}
\tag{3.6}
$$

where we substitute (3.4) into step (a) and assume $\mathbb{E}\left\{\left|\boldsymbol{\varphi}_{i,k'}^{T}\boldsymbol{\varphi}_{k}^{*}\right|^2\right\} = \tau_p$ in step (b). Furthermore, the terms in (3.4) are uncorrelated with each other. Therefore, the MMSE channel estimation is given by

$$
\hat{g}_{mnk} = c_{mk}\breve{g}_{mnk} = \frac{\beta_{mk}\breve{g}_{mnk}}{\beta_{mk} + \frac{1}{\tau_p}\sum_{k'=1}^{K_{int}}\beta_{i,mk'} + \frac{\sigma_z^2}{\tau_p}} \tag{3.7}
$$

Note that c_{mnk} depends only on the large-scale fading coefficients β_{mnk} and $\beta_{i,mnk}$, which remain the same value for the different antennas at the same AP. Thus, we can replace c_{mnk} by c_{mk}. The variance of (3.7) is:

$$
\begin{aligned}
\mathbb{E}\{|\hat{g}_{mnk}|^2\} &= \mathbb{E}\left\{\left|\frac{\beta_{mk}\breve{g}_{mnk}}{\beta_{mk} + \frac{1}{\tau_p}\sum_{k'=1}^{K_{int}}\beta_{i,mk'} + \frac{\sigma_z^2}{\tau_p}}\right|^2\right\} \\
&= \left(\frac{\beta_{mk}}{\beta_{mk} + \frac{1}{\tau_p}\sum_{k'=1}^{K_{int}}\beta_{i,mk'} + \frac{\sigma_z^2}{\tau_p}}\right)^2 \mathbb{E}\{|\breve{g}_{mnk}|^2\} \\
&= \frac{\beta_{mk}^2}{\beta_{mk} + \frac{1}{\tau_p}\sum_{k'=1}^{K_{int}}\beta_{i,mk'} + \frac{\sigma_z^2}{\tau_p}} = c_{mk}\beta_{mk}
\end{aligned}
\tag{3.8}
$$

3.4 Pilot Assignment

In practice, DD-MaMIMO is implemented with a configuration considering the neighbouring clusters as shown in Fig. 3.2, where the current cluster (hexagon) is

surrounded by six neighbours. Each cluster owns an EPU (0-6) at its centre and its coordination region may overlap with others. Hence, the pilot assignment needs to be considered to avoid pilot reuse within the coordination region. For instance, in a simple scenario, the coordination region is overlapped by at most three adjacent clusters. Let us consider that the pilot P3 is located in the intersection area coordinated among the EPU 0, EPU 3 and EPU 4. This pilot cannot be reused in these corresponding coordination areas. However, it can be reused in the area served by EPU 6, excluding the area of overlap that is coordinated by EPU 0. Based on this idea, we develop a pilot allocation algorithm for a user's first time accessing the DD-MaMIMO system, as described in Algorithm 3.1. Without loss of generality, we define any arbitrary EPU as the *x*-th EPU and the user as the *k*-th UT. Furthermore, in the subsequent sub-chapter, we will compare this proposed pilot allocation with a completely random pilot allocation scheme. For the latter, we only consider if a pilot is used within the coordination area of the current EPU, which serves the *k*-th user. The users outside this coordination area are viewed as interfering users, and the pilots are assigned completely randomly to them.

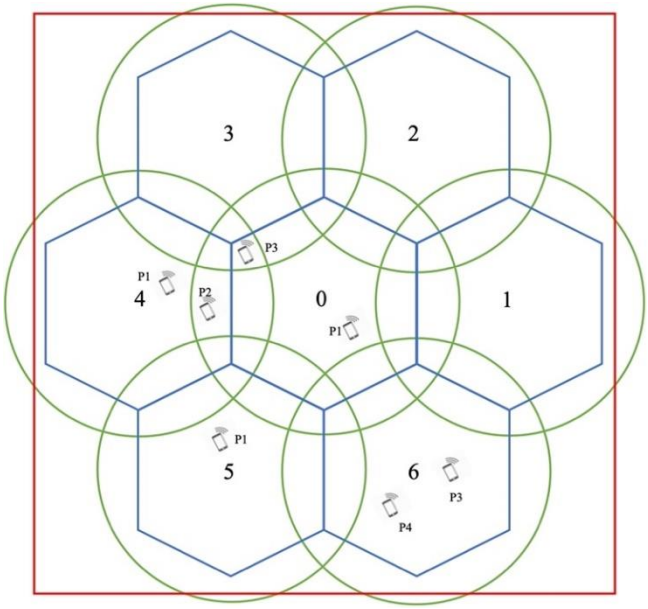


Fig. 3.2 The wrap-around DD-MaMIMO network for pilot assignment, including the current cluster with EPU 0 at centre with six surrounding clusters (with EPU 1-6). P1-P4 denotes the different pilot sequence. The red square is simulated area.

Algorithm 3.1 Pilot Allocation

Define: The coordination area of the general *x*-th EPU is $\mathcal{A}_{coor,x}$ and the service area of the *x*-th EPU is $\mathcal{A}_{serv,x}$.

The position of the general k-th user is UT_k.
$\boldsymbol{\mathcal{P}}$ denotes the total pilot sequences with the length of τ_p.
$\boldsymbol{\mathcal{P}}_x$ denotes the set of pilot sequences which were already assigned in $\mathcal{A}_{coor,x}$.
$\boldsymbol{\mathcal{P}}_{ex}$ denotes the set of pilot sequences which will be excluded.
$\boldsymbol{\mathcal{P}}_{in}$ denotes the set of pilot sequences which will be included.

1: Locate the k-th user in its EPU service area:
if $UT_k \in \mathcal{A}_{serv,x}$ **then**
2: Find the EPU coordination areas in which the k-th user lies:
if $UT_k \in \mathcal{A}_{coor,y}$ **then**
3: $\boldsymbol{\mathcal{P}}_{ex} = \boldsymbol{\mathcal{P}}_y$;
4: $\boldsymbol{\mathcal{P}}_{ex} = \boldsymbol{\mathcal{P}}_x \cup \boldsymbol{\mathcal{P}}_{ex}$;
5: **repeat** step 2 to step 4
6: **until** all coordination areas are found
7: Assign the pilot p to k-th user randomly selected from:
$\boldsymbol{\mathcal{P}}_{in} = \{p \mid p \in \boldsymbol{\mathcal{P}} \,\&\, p \notin \boldsymbol{\mathcal{P}}_{ex}\}$;
8: $\boldsymbol{\mathcal{P}}_x = \boldsymbol{\mathcal{P}}_x \cup p$;
9: **else**
10: **if** UT_k is only in $\mathcal{A}_{serv,x}$ **then**
11: $\boldsymbol{\mathcal{P}}_{ex} = \boldsymbol{\mathcal{P}}_x$;
12: **repeat** step 7 and step 8
13: **end**
14: **end**
15: **end**

3.5 Data Transmission

In this sub-chapter, we will discuss the uplink data transmission and the data detection in DD-MaMIMO. Since each EPU serves a coordination region in DD-MaMIMO, the uplink data transmission is considered only at the APs within this coordination region. This contrasts with the canonical cell-free MaMIMO where all UTs transmit data to all APs across the entire network. The signal received at the APs within the coordination region can be represented by a $M_{coor}N_r \times \tau_u$ size matrix on the uplink:

$$\boldsymbol{Y}_u = \boldsymbol{G}\boldsymbol{X} + \boldsymbol{G}_{int}\boldsymbol{X}_{int} + \boldsymbol{Z} \tag{3.9}$$

where $\boldsymbol{G} \in \mathbb{C}^{M_{coor}N_r \times K_{serv}}$ denotes the channel matrix from the UTs in the service region, $\boldsymbol{G}_{int} \in \mathbb{C}^{M_{coor}N_r \times K_{int}}$ is the channel matrix from the interfering users. The transmitted data, containing τ_u symbols, from the service area and outside the coordination region are represented by the matrix $\boldsymbol{X} \in \mathbb{R}^{K_{serv} \times \tau_u}$ and the matrix $\boldsymbol{X}_{int} \in \mathbb{R}^{K_{int} \times \tau_u}$, respectively. In this thesis, we assume the variance of data σ_x^2 is the same as

the interference $\sigma_{x_{int}}^2$. Finally, $\boldsymbol{Z} \in \mathbb{C}^{M_{coor}N_r \times \tau_u}$ is uncorrelated complex Gaussian noise at the received antennas, whose elements are i.i.d $\mathcal{CN}(0, \sigma_z^2)$ RVs. Here, we assume all data symbols are independent, so we can consider each column vector of the data symbol matrix independently, and rewrite (3.9) in vector form:

$$\boldsymbol{y}_u = \sum_{k=1}^{K_{serv}} \boldsymbol{g}_k x_k + \sum_{k=1}^{K_{int}} \boldsymbol{g}_{i,k} x_{i,k} + \boldsymbol{z} \tag{3.10}$$

where $\boldsymbol{g}_k = \left[\boldsymbol{g}_{1k}^T \, \boldsymbol{g}_{2k}^T \dots \boldsymbol{g}_{M_{coor}k}^T\right]^T$, $\boldsymbol{g}_{i,k} = \left[\boldsymbol{g}_{i,1k}^T \, \boldsymbol{g}_{i,2k}^T \dots \boldsymbol{g}_{i,M_{coor}k}^T\right]^T$ and $\boldsymbol{z}$ is the column vector of $\boldsymbol{Z}$.

Then, the received signals at all APs within the coordination region are transferred via the fronthaul links to the EPU, in which the data is recovered by utilising the combining techniques. The detected data can be represented by weighting the received signals in a general form:

$$\begin{aligned} \hat{x}_k &= \boldsymbol{w}_k \boldsymbol{y}_u \\ &= \boldsymbol{w}_k \hat{\boldsymbol{g}}_k x_k + \boldsymbol{w}_k \tilde{\boldsymbol{g}}_k x_k + \sum_{k' \neq k}^{K_{serv}} \boldsymbol{w}_k \boldsymbol{g}_{k'} x_{k'} + \sum_{k'=1}^{K_{int}} \boldsymbol{w}_k \boldsymbol{g}_{i,k'} x_{i,k'} + \boldsymbol{w}_k \boldsymbol{z} \end{aligned} \tag{3.11}$$

where the weight vector $\boldsymbol{w}_k \in \mathbb{C}^{1 \times M_{coor}N_r}$ can be used to describe various combining techniques. The channel vector $\boldsymbol{g}_k$ can be expanded as $\hat{\boldsymbol{g}}_k + \tilde{\boldsymbol{g}}_k$. The term $\hat{\boldsymbol{g}}_k = \left[\hat{\boldsymbol{g}}_{1k}^T \, \hat{\boldsymbol{g}}_{2k}^T \dots \hat{\boldsymbol{g}}_{M_{coor}k}^T\right]^T$, where $\hat{\boldsymbol{g}}_{mk} = \left[\hat{g}_{m1k} \, \hat{g}_{m2k} \dots \hat{g}_{mN_rk}\right]^T$ represents the estimated channel vector between the k-th UT and the N_r antennas of the m-th AP within the coordination area. Meanwhile, $\tilde{\boldsymbol{g}}_k$ denotes the channel estimation error vector, which has the same form as $\hat{\boldsymbol{g}}_k$. Furthermore, the first two terms in the expanded equality (3.11) are desired signals where $\boldsymbol{w}_k \hat{\boldsymbol{g}}_k x_k$ is the signal over the estimated channel and $\boldsymbol{w}_k \tilde{\boldsymbol{g}}_k x_k$ is the signal over the unknown channel. The third term represents intra-service area interference, while the fourth term denotes interference from users outside the coordination region. The last term is noise. Then, the achievable SE for the k-th UT can be given by using the standard capacity lower bounds which has the same form as [24, 73], as shown in the following proposition.

Proposition 3.1. *Suppose that K_{serv} UTs are served by M_{coor} APs where each AP is equipped with N_r antennas in DD-MaMIMO. The estimation of the channel and data*

are processed at the EPU with unlimited fronthaul. An achievable SE for the k-th UT is

$$SE_k = \frac{\tau_u}{\tau_p + \tau_u} \mathbb{E}\{\log_2(1 + SINR_k)\} \tag{3.12}$$

where the factor $\tau_u/(\tau_p + \tau_u)$ *is the fraction of uplink transmission which is used by data transmission, and the expectation is with respect to the channel estimates. The instantaneous signal-to-interference-plus-noise ratio (SINR) for the k-th UT is given by:*

$$SINR_k = \frac{|\boldsymbol{w}_k \widehat{\boldsymbol{g}}_k|^2}{\boldsymbol{w}_k \left(\sum_{k' \neq k}^{K_{serv}} \widehat{\boldsymbol{g}}_{k'} \widehat{\boldsymbol{g}}_{k'}^H + \sum_{k'=1}^{K_{serv}} \boldsymbol{C}_{\tilde{g}\tilde{g},k'} + \sum_{k'=1}^{K_{int}} \boldsymbol{C}_{g_{int}g_{int},k'} + \frac{\sigma_z^2}{\sigma_x^2} \boldsymbol{I}_{M_{coor}N_r} \right) \boldsymbol{w}_k^H} \tag{3.13}$$

where the covariance matrix of channel estimation error for the k-th UT is expressed as $\boldsymbol{C}_{\tilde{g}\tilde{g},k}$*, and* $\boldsymbol{C}_{g_{int}g_{int},k}$ *denotes the covariance matrix of channel coefficients for the k-th uncoordinated UT.*

Proof: In [73, Th. 4.1], the achievable SE of a cellular system with the discrete memoryless interference channel can be expressed by computing the lower bound of its ergodic capacity, if the channel response is a realization of a RV. In (3.11) the input of such a discrete memoryless interference channel model corresponds to $x = x_k$, the random channel response is denoted by $h = \boldsymbol{w}_k \widehat{\boldsymbol{g}}_k$ with the random realisation $u = \{\widehat{\boldsymbol{g}}_l\}$ where l is arbitrary value, the output is $y = \boldsymbol{w}_k \boldsymbol{y}_u$ and the interference term is represented by:

$$v = \boldsymbol{w}_k \tilde{\boldsymbol{g}}_k x_k + \sum_{k' \neq k}^{K_{serv}} \boldsymbol{w}_k \boldsymbol{g}_{k'} x_{k'} + \sum_{k'=1}^{K_{int}} \boldsymbol{w}_k \boldsymbol{g}_{i,k'} x_{i,k'} + \boldsymbol{w}_k \boldsymbol{z} \tag{3.14}$$

Note that the original definitions of the terms x, h, u, y and v mentioned above were stated in [73] for describing a general discrete memoryless interference channel model. The terms x_k, $x_{i,k}$ and the elements in $\boldsymbol{z}$ are zero mean, and they are independent of the realisations of the channel estimates in (3.14). So, the interference term v is also zero mean. Furthermore, since the weight vector $\boldsymbol{w}_k$ is a function of the channel estimates, the conditional variance of v given $\{\widehat{\boldsymbol{g}}_l\}$ with arbitrary value of l can be calculated by:

$$
\begin{aligned}
&\mathbb{E}\{|v|^2|u=\{\hat{\boldsymbol{g}}_l\}\} \\
&\overset{(a)}{=} \mathbb{E}\{|x_k|^2\}\mathbb{E}\{|\boldsymbol{w}_k\tilde{\boldsymbol{g}}_k|^2|\{\hat{\boldsymbol{g}}_l\}\} + \sum_{k'\neq k}^{K_{serv}} \mathbb{E}\{|x_{k'}|^2\}\mathbb{E}\{|\boldsymbol{w}_k\boldsymbol{g}_{k'}|^2|\{\hat{\boldsymbol{g}}_l\}\} \\
&\qquad + \sum_{k'\neq k}^{K_{serv}} \mathbb{E}\{|x_{k'}|^2\}\mathbb{E}\{|\boldsymbol{w}_k\boldsymbol{g}_{k'}|^2|\{\hat{\boldsymbol{g}}_l\}\} \\
&\qquad + \sum_{k'=1}^{K_{int}} \mathbb{E}\left\{|x_{i,k'}|^2\right\}\mathbb{E}\left\{|\boldsymbol{w}_k\boldsymbol{g}_{i,k'}|^2\middle|\{\hat{\boldsymbol{g}}_l\}\right\} + \mathbb{E}\{|\boldsymbol{w}_k\boldsymbol{z}|^2|\{\hat{\boldsymbol{g}}_l\}\} \\
&\overset{(b)}{=} \sigma_x^2\boldsymbol{w}_k\boldsymbol{C}_{\tilde{g}\tilde{g},k}\boldsymbol{w}_k^H + \sigma_x^2\boldsymbol{w}_k\left(\sum_{k'\neq k}^{K_{serv}} \hat{\boldsymbol{g}}_{k'}\hat{\boldsymbol{g}}_{k'}^H + \boldsymbol{C}_{\tilde{g}\tilde{g},k'}\right)\boldsymbol{w}_k^H \\
&\qquad + \sigma_x^2\boldsymbol{w}_k\sum_{k'=1}^{K_{int}} \boldsymbol{C}_{g_{int}g_{int},k'}\,\boldsymbol{w}_k^H + \sigma_z^2\boldsymbol{w}_k\boldsymbol{I}_{M_{coor}N_r}\boldsymbol{w}_k^H \\
&= \sigma_x^2\boldsymbol{w}_k\left(\sum_{k'\neq k}^{K_{serv}} \hat{\boldsymbol{g}}_{k'}\hat{\boldsymbol{g}}_{k'}^H + \sum_{k'=1}^{K_{serv}} \boldsymbol{C}_{\tilde{g}\tilde{g},k'} + \sum_{k'=1}^{K_{int}} \boldsymbol{C}_{g_{int}g_{int},k'} + \frac{\sigma_z^2}{\sigma_x^2}\boldsymbol{I}_{M_{coor}N_r}\right)\boldsymbol{w}_k^H
\end{aligned}
\tag{3.15}
$$

where step (a) exploits the fact that the terms in (3.14) are uncorrelated with each other, due to the independence of signals x_k, $x_{i,k}$ and the independence among signals, channels and noise. Next, step (b) follows $\mathbb{E}\{|\boldsymbol{w}_k\boldsymbol{g}_{k'}|^2|\{\hat{\boldsymbol{g}}_l\}\} = \boldsymbol{w}_k(\hat{\boldsymbol{g}}_{k'}\hat{\boldsymbol{g}}_{k'}^H + \boldsymbol{C}_{\tilde{g}\tilde{g},k'})\boldsymbol{w}_k^H$ for arbitrary l, because the estimated channel is uncorrelated with the channel estimation error. Additionally, the interference term v is uncorrelated with the input signal x given $\{\hat{\boldsymbol{g}}_l\}$, which is achieved by:

$$
\mathbb{E}\{xv|\{\hat{\boldsymbol{g}}_l\}\} = \mathbb{E}\{|x_k|^2\}\mathbb{E}\{\boldsymbol{w}_k\tilde{\boldsymbol{g}}_k|\{\hat{\boldsymbol{g}}_l\}\} = 0 \tag{3.16}
$$

where the first equality is satisfied since the input signal is independent of all terms of interference (3.14), except the term $\boldsymbol{w}_k\tilde{\boldsymbol{g}}_k x_k$. Then, the second equality is realised because the channel estimation error $\tilde{\boldsymbol{g}}_k$ has zero mean and it is uncorrelated with the channel estimates. Then, the ergodic channel capacity is lower bounded as:

$$
C \geq \mathbb{E}\left\{\log_2\left(1 + \frac{\mathbb{E}\{|x|^2\}|h|^2}{\mathbb{E}\{|v|^2|\{\hat{\boldsymbol{g}}_l\}\}}\right)\right\} \tag{3.17}
$$

where $\mathbb{E}\{|x|^2\} = \sigma_x^2$ and $|h|^2 = |\boldsymbol{w}_k\hat{\boldsymbol{g}}_k|^2$. Finally, we take the fraction of effective

uplink data transmission in one coherence block, which finishes the proof.

Note that $SINR_k$ in (3.12) is instantaneous which is different from the conventional SINR, since it includes the instantaneous estimated channel, the covariance of channel estimation error and the covariance of interfering channels. It needs a period to calculate these covariance matrices. This means that the term $SINR_k$ cannot be measured in a given coherence block. Therefore, we can obtain the related covariance values within a finite realisation and then adopt that to a new independent realisation with respect to a $SINR_k$ RV.

Any combining techniques can be applied to obtain the weight $\boldsymbol{w}_k$. A simple combining technique which has low computational complexity is MRC with $\boldsymbol{w}_k = \widehat{\boldsymbol{g}}_k^H$. Other estimators such as zero-forcing (ZF) also can be selected in the EPU. Moreover, we can also use the MMSE estimation vector to maximise (3.13) as follows.

Corollary 3.1. *The MMSE combining vector for maximising the instantaneous SINR in (3.13) is given by*

$$\boldsymbol{w}_k = \widehat{\boldsymbol{g}}_k^H \left[\sum_{k'=1}^{K_{serv}} \left(\widehat{\boldsymbol{g}}_{k'} \widehat{\boldsymbol{g}}_{k'}^H + \boldsymbol{C}_{\tilde{g}\tilde{g},k'} \right) + \sum_{k'=1}^{K_{int}} \boldsymbol{C}_{g_{int}g_{int},k'} + \frac{\sigma_z^2}{\sigma_x^2} \boldsymbol{I}_{M_{coor}N_r} \right]^{-1} \tag{3.18}$$

which leads to the maximum value

$$\begin{aligned} SINR_k = \widehat{\boldsymbol{g}}_k^H \Bigg(& \sum_{k' \neq k}^{K_{serv}} \widehat{\boldsymbol{g}}_{k'} \widehat{\boldsymbol{g}}_{k'}^H + \sum_{k'=1}^{K_{serv}} \boldsymbol{C}_{\tilde{g}\tilde{g},k'} + \sum_{k'=1}^{K_{int}} \boldsymbol{C}_{g_{int}g_{int},k'} \\ & + \frac{\sigma_z^2}{\sigma_x^2} \boldsymbol{I}_{M_{coor}N_r} \Bigg)^{-1} \widehat{\boldsymbol{g}}_k \end{aligned} \tag{3.19}$$

Proof: We define the MSE of the estimated data as $\bar{\epsilon}_k = \mathbb{E}\{|x_k - \boldsymbol{w}_k \boldsymbol{y}_u|^2\}$. To find the maximum instantaneous SINR, we need to minimise the MSE. Let the first derivative of $\bar{\epsilon}_k$ equal zero with respect to $\boldsymbol{w}_k$:

$$\begin{aligned} 0 = \frac{\partial \bar{\epsilon}_k}{\partial \boldsymbol{w}_k} &= \frac{\partial \mathbb{E}\{|x_k - \boldsymbol{w}_k \boldsymbol{y}_u|^2\}}{\partial \boldsymbol{w}_k} \\ &= \frac{\partial}{\partial \boldsymbol{w}_k} \left(\mathbb{E}\{|x_k|^2\} - \mathbb{E}\{x_k \boldsymbol{y}_u^H \boldsymbol{w}_k^H\} - \mathbb{E}\{x_k^* \boldsymbol{w}_k \boldsymbol{y}_u\} + \mathbb{E}\{|\boldsymbol{w}_k \boldsymbol{y}_u|^2\} \right) \end{aligned} \tag{3.20}$$

$$= -\mathbb{E}\{x_k^* \boldsymbol{y}_u\} + \mathbb{E}\{|\boldsymbol{y}_u|^2\}\boldsymbol{w}_k^H$$

Then, the MMSE combining vector for detecting *k*-th user's data is obtained:

$$\boldsymbol{w}_k = \mathbb{E}\{x_k \boldsymbol{y}_u^H\}\mathbb{E}\{|\boldsymbol{y}_u|^2\}^{-1} \tag{3.21}$$

where the expression can be expanded as (3.18), because x_k is only correlated with $\boldsymbol{g}_k x_k$ in (3.10) and the terms in (3.10) are uncorrelated with each other. Then, we substitute (3.18) into (3.13) and replace the term $\sum_{k'\neq k}^{K_{serv}} \hat{\boldsymbol{g}}_{k'}\hat{\boldsymbol{g}}_{k'}^H + \sum_{k'=1}^{K_{serv}} \boldsymbol{C}_{\tilde{g}\tilde{g},k'} + \sum_{k'=1}^{K_{int}} \boldsymbol{C}_{g_{int}g_{int},k'} + \frac{\sigma_z^2}{\sigma_x^2}\boldsymbol{I}_{M_{coor}N_r}$ by $\boldsymbol{\Theta}$

$$\begin{aligned} SINR_k &= \frac{|\boldsymbol{w}_k\hat{\boldsymbol{g}}_k|^2}{\boldsymbol{w}_k\boldsymbol{\Theta}\boldsymbol{w}_k^H} = \frac{\boldsymbol{w}_k\hat{\boldsymbol{g}}_k\hat{\boldsymbol{g}}_k^H\boldsymbol{w}_k^H}{\boldsymbol{w}_k\boldsymbol{\Theta}\boldsymbol{w}_k^H} \\ &= \frac{\hat{\boldsymbol{g}}_k^H(\hat{\boldsymbol{g}}_k\hat{\boldsymbol{g}}_k^H + \boldsymbol{\Theta})^{-1}\hat{\boldsymbol{g}}_k\hat{\boldsymbol{g}}_k^H(\hat{\boldsymbol{g}}_k\hat{\boldsymbol{g}}_k^H + \boldsymbol{\Theta})^{-1}\hat{\boldsymbol{g}}_k}{\hat{\boldsymbol{g}}_k^H(\hat{\boldsymbol{g}}_k\hat{\boldsymbol{g}}_k^H + \boldsymbol{\Theta})^{-1}\boldsymbol{\Theta}(\hat{\boldsymbol{g}}_k\hat{\boldsymbol{g}}_k^H + \boldsymbol{\Theta})^{-1}\hat{\boldsymbol{g}}_k} \\ &\overset{(a)}{=} \frac{\hat{\boldsymbol{g}}_k^H\boldsymbol{\Theta}^{-1}\hat{\boldsymbol{g}}_k\hat{\boldsymbol{g}}_k^H\boldsymbol{\Theta}^{-1}\hat{\boldsymbol{g}}_k}{\hat{\boldsymbol{g}}_k^H\boldsymbol{\Theta}^{-1}\boldsymbol{\Theta}\boldsymbol{\Theta}^{-1}\hat{\boldsymbol{g}}_k} \\ &= \hat{\boldsymbol{g}}_k^H\boldsymbol{\Theta}^{-1}\hat{\boldsymbol{g}}_k \\ &= \hat{\boldsymbol{g}}_k^H\left(\sum_{k'\neq k}^{K_{serv}}\hat{\boldsymbol{g}}_{k'}\hat{\boldsymbol{g}}_{k'}^H + \sum_{k'=1}^{K_{serv}}\boldsymbol{C}_{\tilde{g}\tilde{g},k'} + \sum_{k'=1}^{K_{int}}\boldsymbol{C}_{g_{int}g_{int},k'} + \frac{\sigma_z^2}{\sigma_x^2}\boldsymbol{I}_{M_{coor}N_r}\right)^{-1}\hat{\boldsymbol{g}}_k \end{aligned} \tag{3.22}$$

where step (a) follows the Lemma: $(\boldsymbol{A}+\boldsymbol{BCD})^{-1} = \boldsymbol{A}^{-1} - \boldsymbol{A}^{-1}\boldsymbol{B}(\boldsymbol{DA}^{-1}\boldsymbol{B} + \boldsymbol{C}^{-1})^{-1}\boldsymbol{DA}^{-1}$, and here, we assume $\boldsymbol{A} = \boldsymbol{\Theta}, \boldsymbol{B} = \hat{\boldsymbol{g}}_k, \boldsymbol{C} = \boldsymbol{I}$ and $\boldsymbol{D} = \hat{\boldsymbol{g}}_k^H$, then we have:

$$\begin{aligned} (\boldsymbol{\Theta} + \hat{\boldsymbol{g}}_k\hat{\boldsymbol{g}}_k^H)^{-1} &= \boldsymbol{\Theta}^{-1} - \boldsymbol{\Theta}^{-1}\hat{\boldsymbol{g}}_k(\hat{\boldsymbol{g}}_k^H\boldsymbol{\Theta}^{-1}\hat{\boldsymbol{g}}_k + \boldsymbol{I})^{-1}\hat{\boldsymbol{g}}_k^H\boldsymbol{\Theta}^{-1} \\ &= \boldsymbol{\Theta}^{-1} - \frac{1}{\hat{\boldsymbol{g}}_k^H\boldsymbol{\Theta}^{-1}\hat{\boldsymbol{g}}_k + 1}\boldsymbol{\Theta}^{-1}\hat{\boldsymbol{g}}_k\hat{\boldsymbol{g}}_k^H\boldsymbol{\Theta}^{-1} \end{aligned} \tag{3.23}$$

where $\hat{\boldsymbol{g}}_k^H\boldsymbol{\Theta}^{-1}\hat{\boldsymbol{g}}_k$ is a scalar. Then, we calculate the expression by left multiplying by $\hat{\boldsymbol{g}}_k^H$:

$$\begin{aligned} \hat{\boldsymbol{g}}_k^H(\boldsymbol{\Theta} + \hat{\boldsymbol{g}}_k\hat{\boldsymbol{g}}_k^H)^{-1} &= \hat{\boldsymbol{g}}_k^H\boldsymbol{\Theta}^{-1} - \frac{\hat{\boldsymbol{g}}_k^H\boldsymbol{\Theta}^{-1}\hat{\boldsymbol{g}}_k\hat{\boldsymbol{g}}_k^H\boldsymbol{\Theta}^{-1}}{\hat{\boldsymbol{g}}_k^H\boldsymbol{\Theta}^{-1}\hat{\boldsymbol{g}}_k + 1} \\ &= \frac{(\hat{\boldsymbol{g}}_k^H\boldsymbol{\Theta}^{-1}\hat{\boldsymbol{g}}_k + 1)\hat{\boldsymbol{g}}_k^H\boldsymbol{\Theta}^{-1} - \hat{\boldsymbol{g}}_k^H\boldsymbol{\Theta}^{-1}\hat{\boldsymbol{g}}_k\hat{\boldsymbol{g}}_k^H\boldsymbol{\Theta}^{-1}}{\hat{\boldsymbol{g}}_k^H\boldsymbol{\Theta}^{-1}\hat{\boldsymbol{g}}_k + 1} \end{aligned} \tag{3.24}$$

$$= \frac{\widehat{\boldsymbol{g}}_k^H \boldsymbol{\Theta}^{-1}}{\widehat{\boldsymbol{g}}_k^H \boldsymbol{\Theta}^{-1} \widehat{\boldsymbol{g}}_k + 1}$$

The proof is completed.

If the MMSE estimator is used to estimate channels, the variance of the channel estimation error for the *k*-th UT to the *n*-th antenna of the *m*-th AP is:

$$\begin{aligned}
\sigma_e^2 &= \mathbb{E}\{|\tilde{g}_{mnk}|^2\} = \mathbb{E}\{(g_{mnk} - \hat{g}_{mnk})(g_{mnk} - \hat{g}_{mnk})^*\} \\
&\quad = \mathbb{E}\{|g_{mnk}|^2 - \hat{g}_{mnk} g_{mnk}^* - g_{mnk} \hat{g}_{mnk}^* + |\hat{g}_{mnk}|^2\} \\
&= \beta_{mnk} - \mathbb{E}\{\hat{g}_{mnk}(\hat{g}_{mnk}^* + \tilde{g}_{mnk}^*)\} - \mathbb{E}\{(\hat{g}_{mnk} + \tilde{g}_{mnk})\hat{g}_{mnk}^*\} \\
&\quad + \mathbb{E}\{|\hat{g}_{mnk}|^2\} \\
&\overset{(a)}{=} \beta_{mnk} - \mathbb{E}\{|\hat{g}_{mnk}|^2\} = \beta_{mnk} - c_{mnk}\beta_{mnk} \\
&= (1 - c_{mk})\beta_{mk}
\end{aligned} \tag{3.25}$$

where the estimated channel is assumed to be uncorrelated with corresponding channel estimation error in step (a). Then, the covariance of channel estimation error is:

$$\begin{aligned}
\boldsymbol{C}_{\tilde{g}\tilde{g},k} &= \mathbb{E}\{\tilde{\boldsymbol{g}}_k \tilde{\boldsymbol{g}}_k^H\} \\
&= \begin{bmatrix} (1 - c_{11k})\beta_{11k} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & (1 - c_{M_{coor}N_r k})\beta_{M_{coor}N_r k} \end{bmatrix} \\
&= diag\{(1 - c_{11k})\beta_{11k}, \ldots, (1 - c_{1N_r k})\beta_{1N_r k}, \ldots, (1 \\
&\qquad - c_{M_{coor}1k})\beta_{M_{coor}1k}, \ldots, (1 - c_{M_{coor}N_r k})\beta_{M_{coor}N_r k}\} \\
&= diag\Big\{\underbrace{(1 - c_{1k})\beta_{1k} \ \ldots \ (1 - c_{1k})\beta_{1k}}_{N_r} \ \ldots \ (1 - c_{M_{coor}k})\beta_{M_{coor}k} \ \ldots \ (1 \\
&\qquad - c_{M_{coor}k})\beta_{M_{coor}k}\Big\}
\end{aligned} \tag{3.26}$$

where $diag$ denotes diagonal matrix. Since we assume the channels are uncorrelated with each other at the same AP with multiple antennas, the off-diagonal elements in (3.26) are zero. The covariance matrix of interfering channels is:

$$\begin{aligned}
&\boldsymbol{C}_{g_{int}g_{int},k} = \mathbb{E}\{\boldsymbol{g}_{i,k}\boldsymbol{g}_{i,k}^H\} \\
&= diag\{\beta_{i,11k}, \beta_{i,12k}, \ldots \beta_{i,1N_r k}, \ldots \beta_{i,M_{coor}1k}, \ldots \beta_{i,M_{coor}N_r k}\}
\end{aligned} \tag{3.27}$$

$$= diag\left\{\underbrace{\beta_{i,1k} \ \dots \ \beta_{i,1k}}_{N_r} \ \dots \ \underbrace{\beta_{i,M_{coor}k} \ \dots \ \beta_{i,M_{coor}k}}_{N_r}\right\}$$

3.6 Scalable Network

In the conventional cell-free MaMIMO system described in Chapter 2, the network where all APs connect to one CPU simultaneously to serve all UTs is not scalable. The reason is that if the network coverage tends to infinity, the number of fronthaul links and the fronthaul length will be infinity which cannot be implemented in real life. Furthermore, it is impractical for a user to be served by all APs. This kind of scalability issue has been mentioned in [22, 23] where the definition of scalability was given that the computational complexity of per AP remains finite when the number of UTs in the whole network tends to infinity. In addition, to solve the pilot contamination caused by non-orthogonal pilots, the greedy pilot assignment algorithm should be adopted which requires iterative re-computation at the CPU [14, 40]. This means the large scale fading gain information for all APs needs to be known at the CPU in order to find the strongest pilot contamination and reassign a proper pilot. Hence, it is not scalable and unrealistic, especially in the ultra-dense low-latency dynamic mobility network. The author in [40] has also mentioned the scalable challenge that the computational complexity and latency will be high, and the fronthaul capacity might increase when the coordination level among the APs increases.

There are several approaches to realise the scalable networks. The method in [23] is to select a dynamic cooperation cluster where a small number of APs serve one specific UT within that cluster, assuming that each AP serves no more than one UT per pilot sequence. Furthermore, the partial MMSE (P-MMSE) and the local P-MMSE (LP-MMSE) combining techniques are used for data detection in [23], ensuring the scalability of the network. However, these combining techniques sacrifice a part of ability for mitigating the interference caused by other users. The alternative simple method is to divide the cell-free MaMIMO network into separate regions served by independent CPUs. However, this results in the edge interference between the adjacent network clusters. In [22], the author discusses this simple method and assumes multiple CPUs are interconnected. But, the issue with this specific approach is that it is not clear exactly how the CPUs would cooperate. For overcoming these disadvantages, DD-

MaMIMO is considered as a new scalable network system in this thesis. To explain why DD-MaMIMO is a scalable network, we propose the definition below which is adapted from the one originally given in [23].

Definition 3.1 (Scalability): *A DD-MaMIMO network is scalable if the following tasks have finite complexity and resource requirements for each AP and EPU as the number of UTs in the whole network K tends to infinity.*

1. *Signal processing for channel estimation;*
2. *Signal processing for data reception and transmission;*
3. *Fronthaul signalling for CSI sharing and data;*
4. *Power control optimisation (not applied in this thesis).*

In DD-MaMIMO, all APs within the coordination region send the received pilots to the EPU for channel estimation, a process that does not take place at the AP. Moreover, the EPU only needs to estimate the channels from K_{serv} UTs located within the service area to each AP in the coordination region, both of which are finite regardless of K. Consequently, the complexity of the channel estimation is finite for each EPU, as $K \to \infty$. For the uplink data detection, the EPU computes $\{\boldsymbol{w}_k \boldsymbol{y}_u : k = 1, \dots, K_{serv}\}$ in (3.11) using K_{serv} different combining vectors $\{\boldsymbol{w}_k : k = 1, \dots, K_{serv}\}$, which has finite computational complexity with varying combining techniques. Note that for the MMSE combining vectors in (3.18), K_{int} different covariance matrices of interfering channels $\{\boldsymbol{C}_{g_{int}g_{int},k} : k = 1, \dots, K_{int}\}$ are added together. It might grow unboundedly if K were infinity. However, $\boldsymbol{C}_{g_{int}g_{int},k}$, given in (3.27) consisting of the large scale fading coefficients for the interfering channels, only depends on the distance between the UT and AP, which leads to a tiny value for distant UTs. Therefore, we can set a threshold for UT-AP distance and ignore any UTs that exceed it, and the complexity remains finite. Similarly, for the downlink transmission, the EPU creates transmitted signals for each AP, comprising the precoding vectors and data signals intended for K_{serv} UTs, which also has finite complexity. Over the fronthaul link, each AP in the coordination region needs to forward τ_p length received pilots and τ_u length data signals to an EPU, and in turn receives the K_{serv} downlink data signals. Hence, the fronthaul signalling is finite. Finally, in this thesis, we do not apply the power control to DD-MaMIMO but our numerical results in the next sub-chapter show that it can nevertheless achieve a

better performance than cell-free MaMIMO with power control. So, the fourth task in Definition 3.1 vanishes. Overall, the DD-MaMIMO is scalable.

3.7 System Performance

In this sub-chapter, we will show the numerical results for the uplink performance of DD-MaMIMO by setting up the simulation. We will first introduce the initial setup parameters. Then, we will compare the DD-MaMIMO with cell-free MaMIMO by using different combining techniques and applying either single or multiple antennas. Furthermore, according to the definition of DD-MaMIMO, we will analyse the results with different radii of coordination region.

3.7.1 Simulation Setup

A DD-MaMIMO system with a hexagonal service area and a circular coordination area is considered. We assume that the inter-EPU distance is $300\ m$ and simulate the whole system in a square area with side length $1\ km$. To give a reasonable comparison, we adopt the UT density $\rho_u = 40/\mathrm{km}^2$ and AP density $\rho_A = 400/\mathrm{km}^2$ in general cases, which are similar to the parameters adopted in [24]. The one-slope path loss model is given by (3.2) with the parameters referred to above. All UTs transmit with the power $\sigma_s^2 = 100$ mW, the noise power is $\sigma_n^2 = -96$ dBm, the uplink data transmission $\tau_u = 190$ and the pilot length is $\tau_p = 10$ in this thesis.

Moreover, the UTs and APs are randomly allocated within the simulation area, and the number of both follows the random Poisson distribution. The locations for the UTs and APs depend on the geometric boundary of the service region, coordination region, and the region outside coordination region. In the multiple-antenna case, we provide simulation results where the number of antennas per AP are 2, 4 and 8. For analysing the proposed pilot allocation, stated in Algorithm 3.1, the neighbouring network clusters are considered which follows Fig. 3.2.

3.7.2 Numerical Results

First, we revisit [30] in which the channel hardening was assumed. We highlight that the network cluster edge effect occurs in cell-free MaMIMO by presenting the SIRs for various UTs. Fig. 3.3 provides a scatter plot where each UT is represented by a point,

with various colours denoting differing levels of SIR. From the results, we can observe that the worst SIRs, which are negative, occur at the edge of the hexagonal cluster. The authors in [17] present the comparison of the downlink achievable data rates for the cell-free MaMIMO and collocated MaMIMO, which indicates the benefits of cell-free MaMIMO. But it also shows that the users at the cluster edge have a lower data rate.

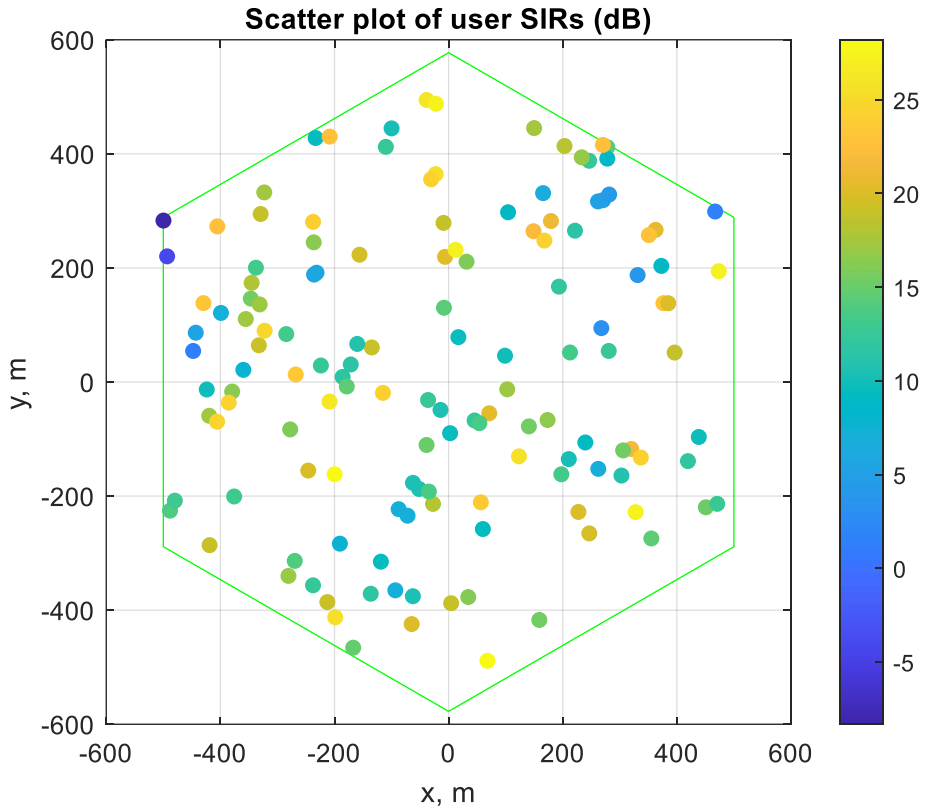


Fig. 3.3 Scatter plot of UTs served by one CPU in cell-free MaMIMO, presenting their SIRs (taken from [30, Fig. 2]).

We next evaluate our proposed DD-MaMIMO architecture. Fig. 3.4 shows the cumulative distribution function (CDF) of achievable uplink SE by applying the MRC data detection and MMSE channel estimation with various radii of coordination region. At the 90% likely SE points, all the cases have the insufficient performance which only reaches 1 or 1.5 bit/s/Hz per user. We can see that there is an improvement in average as the radius of coordination area increases, especially for the 400 m radius case. But the increase is not large when the radius increases from 282 m to 400 m. The poor processing leads to a limited SE performance even if the number of serving APs increases.

Note that the minimum coordination region has radius 174 m, where this radius is

equivalent to the distance from circle centre to the hexagon vertices. In this thesis, we mainly focus on the case of radius 282 m where the area of coordination region is a quarter of the simulation area. In this case the number of interfering pilots, 30, is the same as for the cell-free MaMIMO system in [24], where the 40 users are evenly divided between four square areas which are quarters of the simulation area, and each uses the same set of 10 pilots. So, the results are comparable to the radius 282 m case. Moreover, the "Cell-free" in the plot represents the case where the coordination area is the same as the service area. This also corresponds to a cell-free Massive MIMO system in which the network is divided into separate regions served by their own CPU.

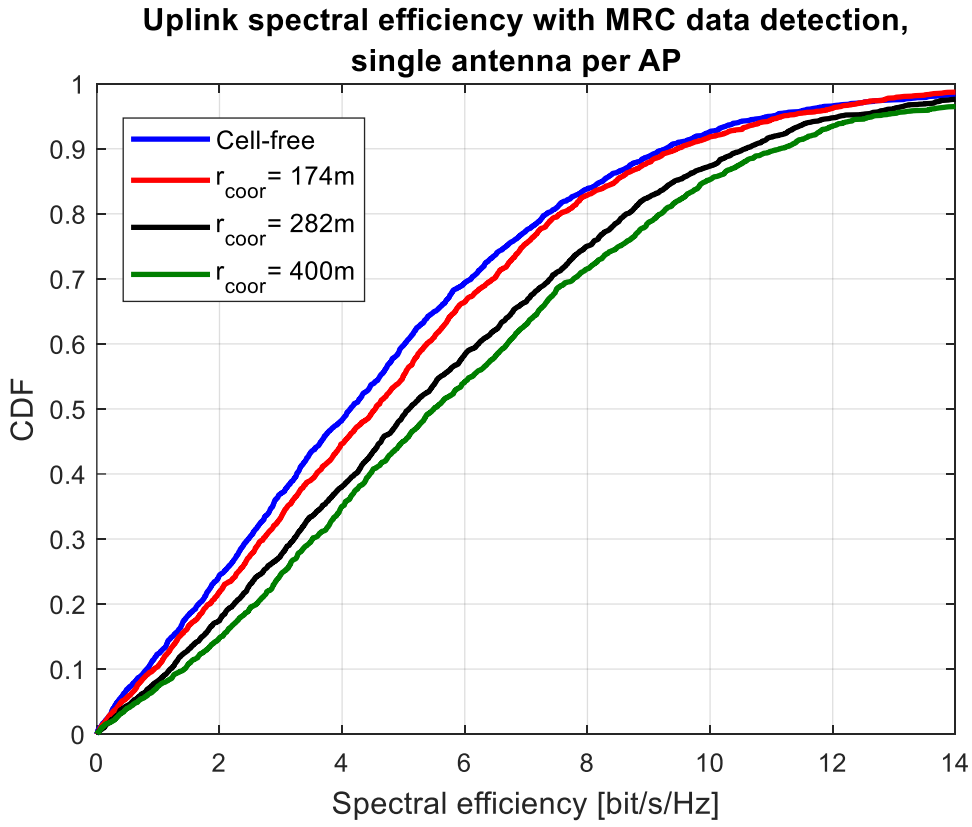


Fig. 3.4 CDF of uplink SE for MRC data detection and MMSE channel estimation, $d_{EPU} = 300\ m$, single antenna per AP, for a range of coordination radii. For "Cell-free" the coordination region is coincident with the service region, which is also the same as cell-free MaMIMO.

In Fig. 3.5, using the MMSE estimator for channel estimation and data detection, the performance is presented which again shows the CDF of uplink SE with various coordination region radii. Moreover, in this case the 90% likelihood SE increases significantly as the coordination radius increases, where there is an increase of 4 bit/s/Hz comparing the 400 m radius case with the cell-free system. This shows that

when the radius increases, more desired signal power will be collected, but the interference will decrease. Note that although the SE improves as the coordination radius grows, there is a necessary condition, which should keep K_{coor} less than the pilot length τ_p. That is also one of the key idea of DD-MaMIMO that states the pilots assigned to the users within the coordination region are fully orthogonal, which largely reduces the interference from the pilot contamination. As the discussion in Fig. 3.4, this is a fair comparison between the cell-free MaMIMO given in [24] and DD-MaMIMO with 282 m coordination radius in Fig. 3.5. The same size of simulation area and the same number of APs and UTs are used. We also consider the same number of interfering users and the same pilot length. The results show that DD-MaMIMO has a better SE than the cell-free MaMIMO for average (8.9 bits/s/Hz in Fig. 3.5 compared to 8.2 bits/s/Hz in [24]). This is because the users performance at the cluster edge are improved.

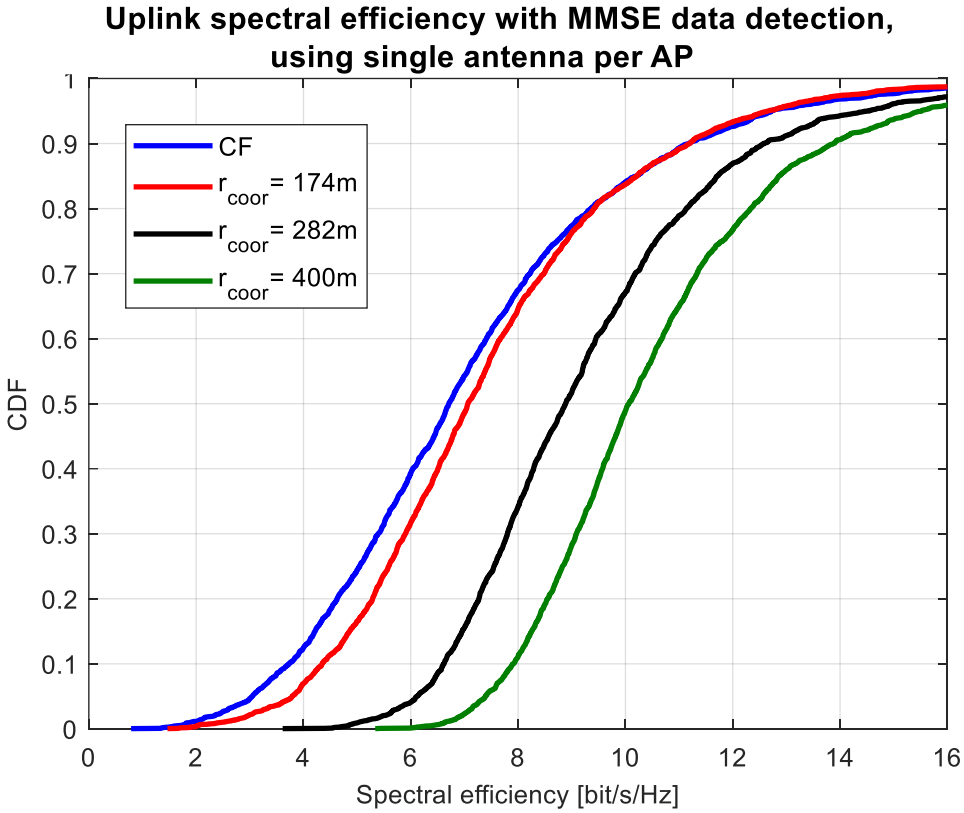


Fig. 3.5 CDF of uplink SE for MMSE data detection and channel estimation, $d_{EPU} = 300\ m$, single antenna per AP, for a range of coordination radii.

Then, Fig. 3.6 considers two cases with the same antenna density level of $400/km^2$ while varying the number of antennas per AP. The scenario with a larger number of

APs, each equipped with a single antenna, outperforms the case with a quarter of the number of APs, each equipped with four antennas. This result arises due to the sparser coverage in situations with fewer APs, which in turn increases the average distance between the UT and AP. This increase in distance subsequently diminishes the power of received signal.

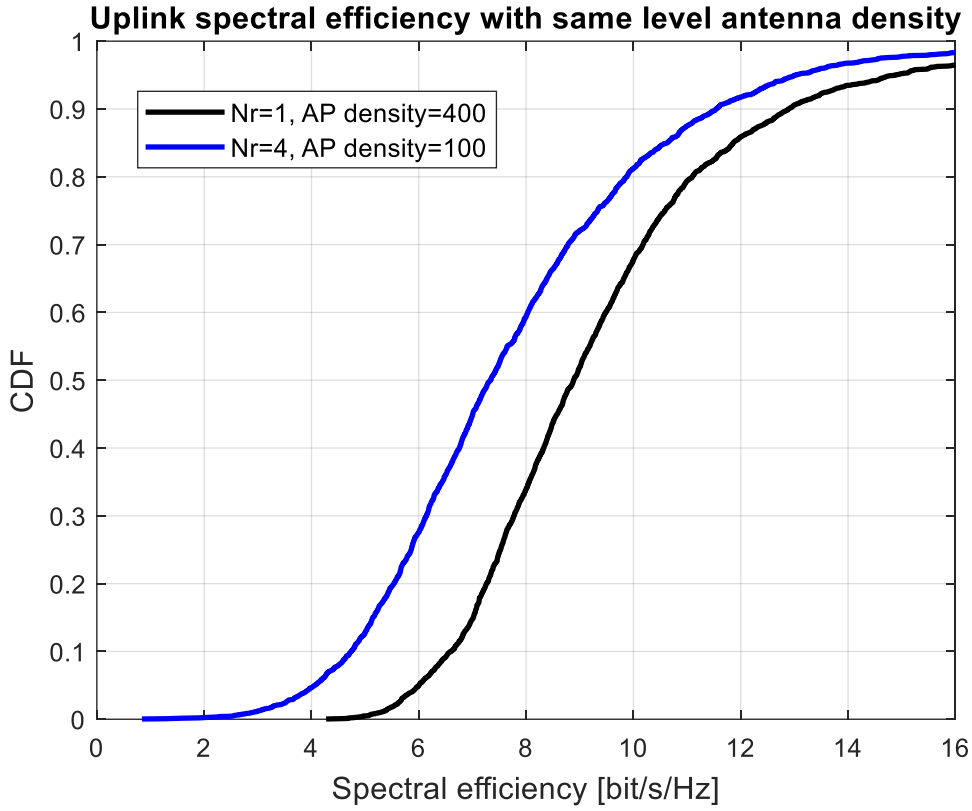


Fig. 3.6 CDF of uplink SE for MMSE data detection and channel estimation, $d_{EPU} = 300\ m$, with same density level of antennas but different No. of antennas per AP. The coordination radius is 282 m.

Fig. 3.7 compares MRC data detection with MMSE data detection using the same scheme of channel estimation (MMSE). It shows that using the MMSE estimator for detecting data offers a better SE performance, approximately a 4 bit/s/Hz increase on average and 5.5 bit/s/Hz for 90% of the users. This is comparable with the "level 4" cooperation in [24, Fig. 2(a)] and [24, Fig. 3], respectively, though the SE using MRC detection drops slightly due to the different deployment of the APs and UTs. In [24], the APs were deployed on a square grid and the UTs were evenly assigned between the four quarters of the simulation area, where the number of APs and UTs were fixed as 400 and 40. However, in our DD-MaMIMO simulation scenario, the number of APs

and UTs follows the Poisson distribution with the same density level. Hence the number of APs in the coordination region may be less and the SE is reduced.

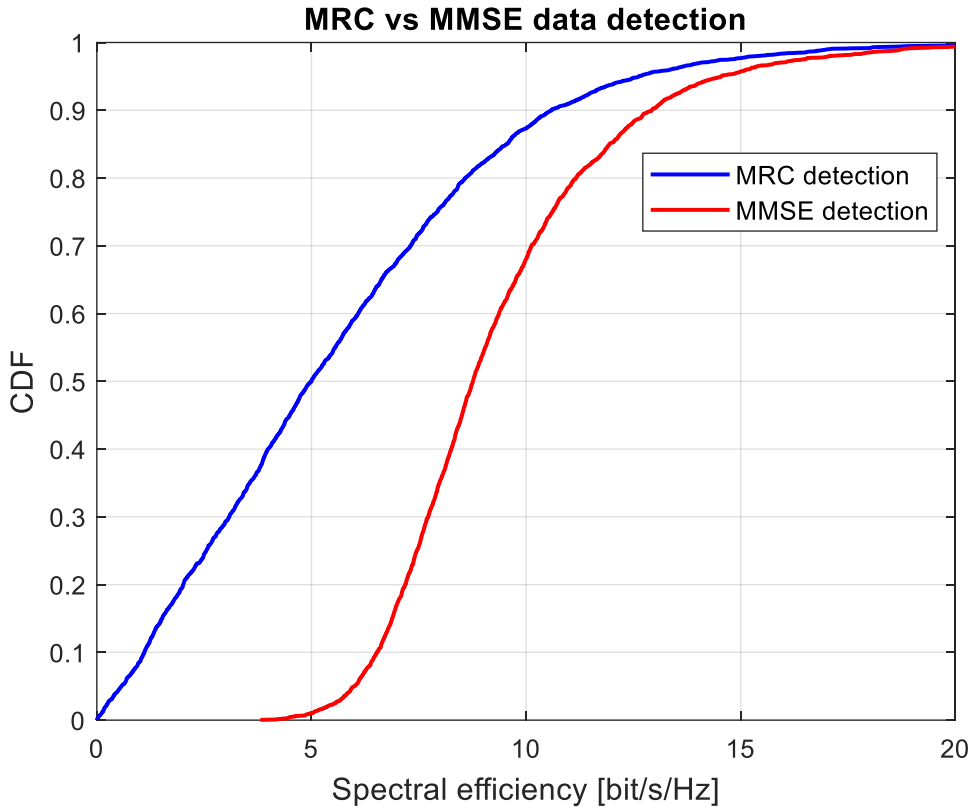


Fig. 3.7 CDF of uplink SE for the comparison between MRC (blue) and MMSE (red) data detection, $d_{EPU} = 300\ m$, single antenna per AP, and the coordination radius is 282 m.

The scenario of multiple antennas per AP is considered and presented in Fig. 3.8. We apply the same setup of the APs/UTs density and the coordination radius to the different multiple-antenna cases. The figure shows the increase of the number of antennas per AP leads to an obvious SE increase. It makes sense that the desired signal gain received at the AP is stronger. However, the requirements of the fronthaul capacity and computational complexity at the EPU will also be increased. In the multiple-antenna case, $N_r\tau_p$ received pilots and $N_r\tau_u$ data for one coherence block needs to be sent to EPU via each fronthaul link. For least square channel estimation, the EPU needs to compute $\tau_p + 1$ scalar multiplications and $\tau_p - 1$ scalar additions in (3.4). Then, three multiplications and $K_{int} + 1$ additions are performed in (3.6) for calculating one MMSE estimation weight. Finally, for obtaining all MMSE channel estimates, $M_{coor}N_rK_{serv}(\tau_p + 5)$ scalar multiplications and $M_{coor}N_rK_{serv}(\tau_p + K_{int})$ scalar

additions are computed in the EPU. Significantly, if MMSE estimation is used for data detection, a $M_{coor}N_r \times M_{coor}N_r$ matrix needs to be inverted in (3.18) which has the computational complexity $\mathcal{O}(M_{coor}^3 N_r^3)$, where $\mathcal{O}(\cdot)$ denotes big O notation.

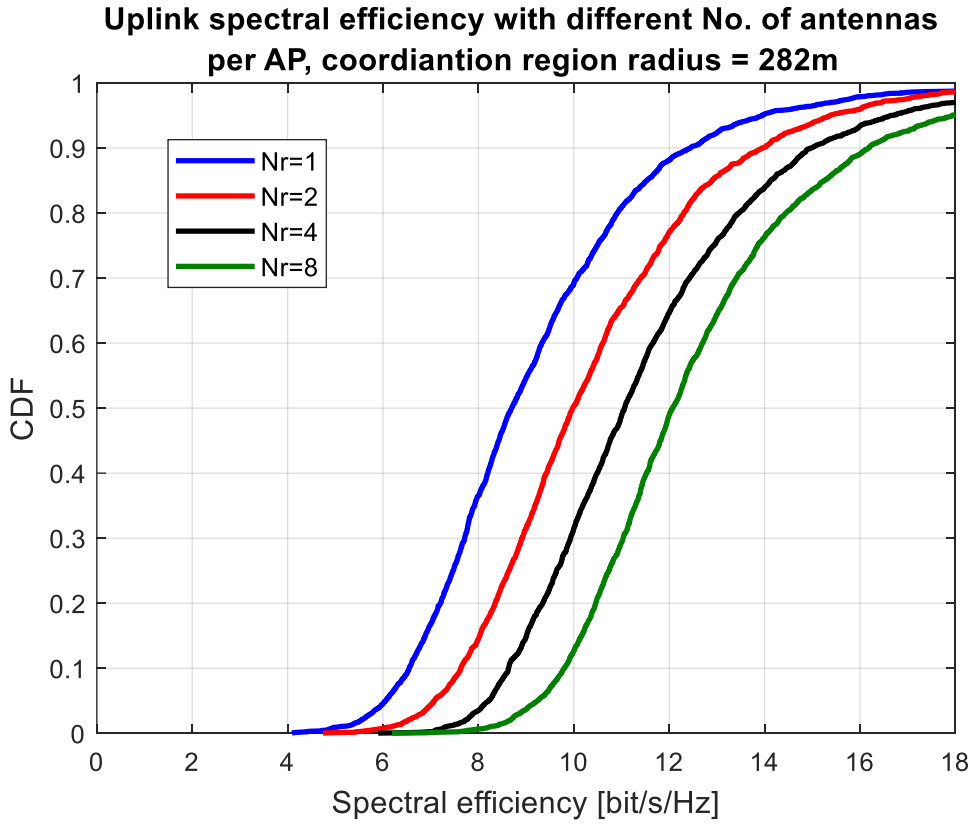


Fig. 3.8 CDF of uplink SE for the MMSE data detection in respect of the different number of antennas per AP, $d_{EPU} = 300\ m$ and the coordination radius is 282 m.

To evaluate the computational complexity, we run the simulation in MATLAB R2023a with a laptop processor (12th Gen Intel® Core™ i9-12900H). In particular, we calculate the average execution time for inverting the $M_{coor}N_r \times M_{coor}N_r$ matrix in (3.18). This could be different from the practical wireless communications system in which the hardware may have different processor and the optimisation can be applied in the software. In Fig. 3.9, numerical results are obtained by varying the radius of coordination region while keeping the same number of antennas per AP. The results show that as the coordination region radius increases, the time required to invert the matrix for calculating the MMSE combining weight grows significantly. We observe that the average execution time for a coordination region with a 400-metre radius is approximately 64 times that of one with a 200-metre radius, when the number of antennas per AP is 4. This is reasonable, as the number of APs within the coordination

region is quadrupled, leading to a 64-fold increase in computational complexity, according to $\mathcal{O}(M_{coor}^3 N_r^3)$. Note that we adopt Poisson distribution to generate the random number of APs and use the function "inv" in MATLAB which may cause the imprecise results. We can observe that the slope is different for the cases with fewer antennas or 174 m radius. Furthermore, for a given coordination region radius, the execution time rises as the number of antennas per AP increases. Specifically, in a scenario with a 230-metre coordination region radius, each AP with 16 antennas has an average execution time approximately 1000 times longer than each AP with just a single antenna.

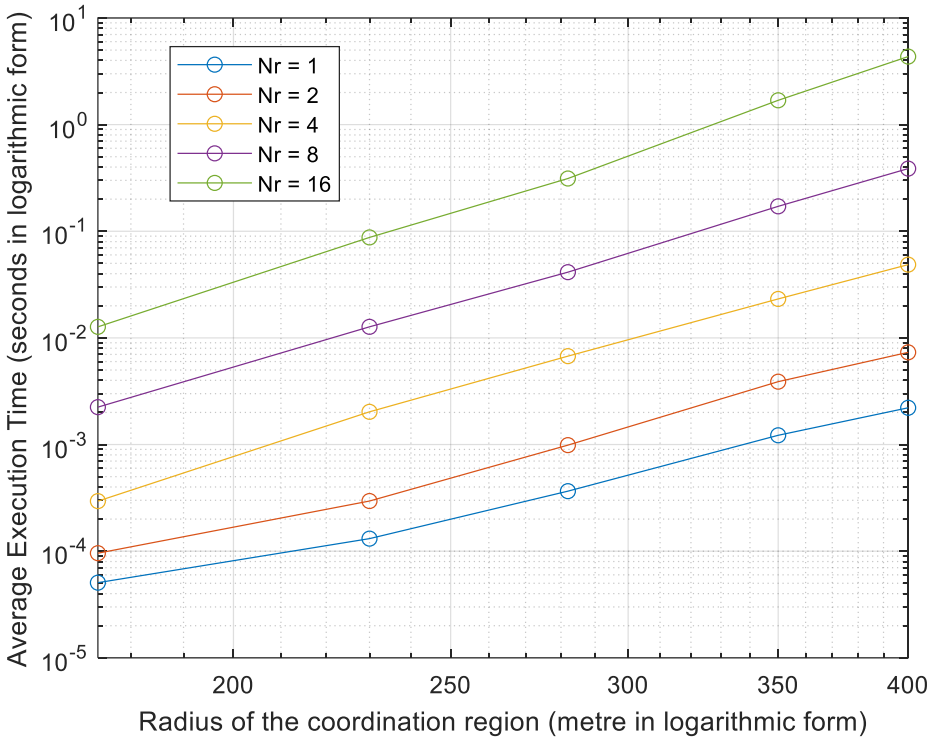


Fig. 3.9 Complexity analysis for the average execution time of inverting a matrix which is used to calculate the MMSE combining weight. The radius of the coordination region varies from 174 m to 400 m. "N_r" denotes the number of antennas per AP.

Finally, we consider the proposed pilot algorithm which is given by Algorithm 3.1 and the completely random pilot allocation which the pilots assigned to the interfering users outside of the coordination region are random. It means the interfering pilots may be the same as those of the served users. Fig. 3.10 gives the comparison of the different allocation methods in which the same setup is adopted but the coordination radius is 282 m, here. The figure shows that there is not too much difference between two

schemes. We need to notice that random pilot assignment may not be practical for the whole system, because it does not ensure that the same pilot is not reused within any coordination region, just not within the central one. This could affect the system performance due to the pilot contamination. However, the results from Fig. 3.10 prove that the system performance does not become worse by using completely random pilot allocation for users outside coordination region in DD-MaMIMO.

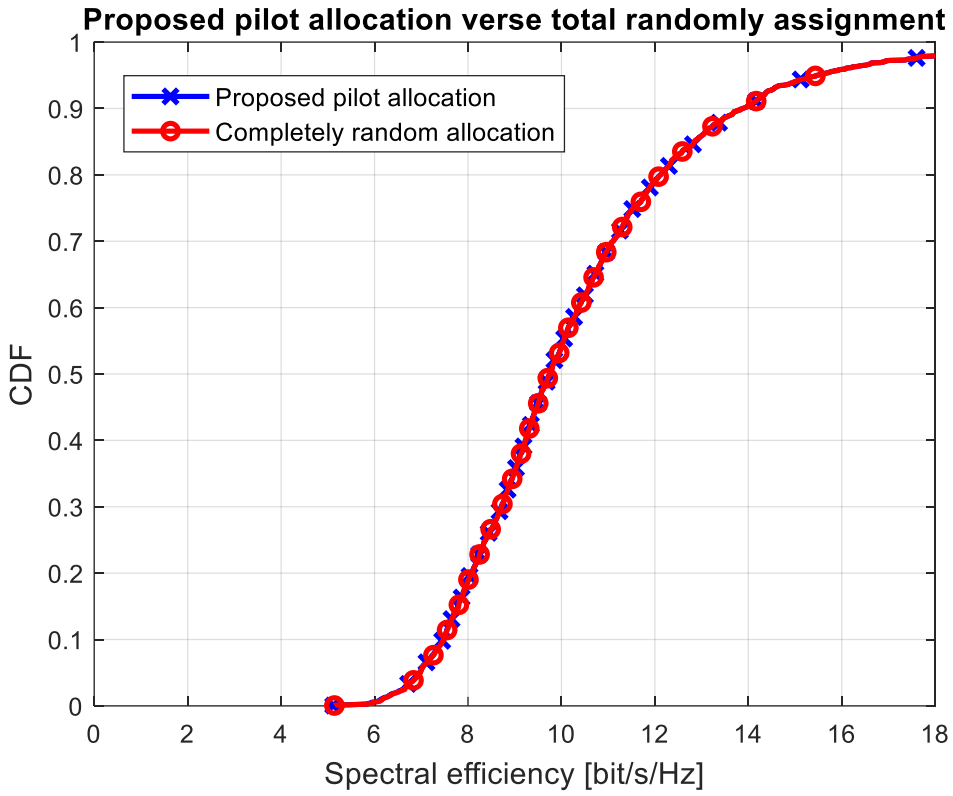


Fig. 3.10 CDF of uplink SE for different pilot allocation, the proposed pilot allocation (blue line with cross mark) and total randomly assignment (red line with round mark), $d_{EPU} = 300\ m$ and the coordination radius is 282 m.

3.8 Summary

In this chapter, we have considered a new scalable DD-MaMIMO architecture which has the potential to fulfil the ultra-dense low-latency and massive devices access requirements in 6G and future wireless network. We have exploited MMSE channel estimation and applied either MRC or MMSE to detect data for DD-MaMIMO. To evaluate the system performance, the SE expression has been proposed without assuming channel hardening. Numerical results have been provided to demonstrate that our DD-MaMIMO outperforms cell-free MaMIMO and that the edge effect can be

significantly mitigated. In addition, the results confirm that the use of MMSE estimation can achieve a better SE than MRC, even though only a single antenna is used on each AP. Furthermore, the simulations for the multiple-antenna AP and increasing the coordination radius have been provided, which indicate that more antennas per AP and larger coordination radius performed better. We finally have proposed a pilot allocation algorithm for adapting the DD-MaMIMO in practice, though the results showed that it does not have any difference compared to the completely random pilot allocation. Overall, DD-MaMIMO can be a promising system.

Chapter 4

Quantisation in Decentralised Distributed Massive MIMO

4.1 Introduction

Quantisation is necessary for communications system as Chapter 2 stated that the analogue signal has to be converted to a digital signal before sending it via the fronthaul. This is because the fronthaul which connects the access points (APs) to the edge processing unit (EPU) by the optical fibre has finite capacity in practice. Hence, the signal processing in the decentralised distributed massive multiple-input multiple-output (DD-MaMIMO) system with limited-capacity fronthaul needs to be considered. To answer this, we will research the following issues in this chapter: where the quantisation happens and how the quantisation is processed. To begin with a simple case, we mainly focus on the DD-MaMIMO system when only a single antenna is available at the AP in the subsequent sub-chapters.

Based on the research of channel state information (CSI) acquisition in [25], we also consider two strategies for channel estimation. First, all the APs quantise the received pilot sequences and then send them to the EPU for channel estimation. Here, we assume the EPU knows the information of the large scale fading and the pilots allocated to the user terminals (UTs). Another strategy is that the channel is estimated locally at the AP and the quantised version of the channel estimation is sent to the EPU. This answers where the quantisation happens for the purposes of the channel estimation. For analysing the process of quantisation, we adopt the simple mid-rise uniform quantiser and exploit the Bussgang decomposition to model this non-linear system.

There have been some papers that researched the quantisation in the cell-free MaMIMO by using the Bussgang decomposition. For instance, in [25, 28] the basic analysis of the quantisation using Bussgang decomposition for channel estimation based on complex random pilots is studied. However, the authors did not give a strict

discussion about the condition of the original Bussgang theorem [65]which requires the input of the quantiser to be Gaussian distributed. The author in [26] indicates that the central limit theorem can be exploited to approximate the input signal of quantiser to the Gaussian distribution. However, the characteristic of the input signal, which consists of the sum of many products of the complex random Gaussian distributed variables, was not researched in practice. Moreover, the central limit theorem is only applicable when the number of random variables drawn from a near-Gaussian distribution is sufficiently large, which may not hold true in the DD-MaMIMO system, particularly in rural areas with few UTs. Thus, it is imperative to research the actual scenario where the input of the quantiser has a non-Gaussian distribution. In the subsequent sub-chapters, we will derive the probability density function (pdf) of the non-Gaussian distribution and employ it to calculate the Bussgang decomposition parameters and determine the optimum step interval in DD-MaMIMO. Often, people adopt complex random pilot sequences to estimate the channel in the MaMIMO [73] and the cell-free system [14], which gives rise to the Bussgang decomposition issue mentioned above. Apart from giving the true distribution, there is an alternative approach which uses binary pilot sequences. In this study, we also investigate this binary pilot-based channel estimation method, and the results reveal that there is no significant difference in performance.

Next, we introduce the quantisation of the data transmission on the uplink of DD-MaMIMO, in which the transmitted data are quantised at the APs before being sent to the EPU for data detection. To evaluate the performance of the system, we derive an expression for the signal-to-interference-and-distortion-plus-noise ratio (SIDNR) and the formulae for the achievable ergodic capacity. Finally numerical results are given for the proposed scheme.

4.2 System Model

We assume there are K_{serv} UTs and M_{coor} APs with single antenna in the DD-MaMIMO system. The received uplink signal at the m-th AP is given by:

$$x_m = \sum_{k=1}^{K_{serv}} g_{mk} s_k + \sum_{k=1}^{K_{int}} g_{i,mk} s_{i,k} + n_m \tag{4.1}$$

$$= \sum_{k=1}^{K_{serv}} h_{mk}\beta_{mk}^{1/2} s_k + \sum_{k=1}^{K_{int}} h_{i,mk}\beta_{i,mk}^{1/2} s_{i,k} + n_m$$

where g_{mk} is the channel coefficient from the k-th served UT to the m-th AP and $g_{i,mk}$ is the interfering channel coefficient from the k-th UT outside the coordination region. s_k and $s_{i,k}$ are the transmit signals from the k-th user and the k-th interfering user, respectively. n_m denotes the noise at the m-th AP. $h_{mk} \sim \mathcal{CN}(0,1)$ represents the Rayleigh fading channel coefficient and β_{mk} is the large-scale fading with a path loss. $h_{i,mk}$ and $\beta_{i,mk}$ are the Rayleigh and large scale fading coefficients from the interfering UTs. The variance of the received uplink signal is:

$$\sigma_x^2 = \mathbb{E}\{|x_m|^2\} = \sum_{k=1}^{K_{serv}} \beta_{mk}\, \sigma_s^2 + \sum_{k=1}^{K_{int}} \beta_{i,mk}\, \sigma_{i,s}^2 + \sigma_n^2 \tag{4.2}$$

where $\sigma_s^2, \sigma_{i,s}^2$ and σ_n^2 are the variances of transmitted signal, interfering signal and noise, respectively.

4.3 Channel Estimation

In Chapter 3, we discussed the channel estimation in the DD-MaMIMO system assuming perfect fronthaul transmission. In this sub-chapter, we will present the definition and derivation of the channel estimation when the fronthaul capacity is limited. There are two CSI acquisition strategies that we will consider. The first approach, known as the *quantise-and-estimate* (QE) scheme [25], involves quantising the received pilots from all APs and sending them to the EPU for channel estimation. This approach is suitable for the case where EPU has high computing ability or the fronthaul link capacity is small. The reason is in QE scheme, each AP only needs to quantise τ_p length pilots but the channel estimation is processed at the EPU. The other is the *estimate-and-quantise* (EQ) scheme [25] where the channels are estimated by the local APs and the quantised CSI are then collected by the EPU. This approach is applied to the scenario in which the fronthaul link capacity is big or the low computational complexity is required at the EPU. This is because each AP needs to quantise all estimated channels from all UTs within the service area and the EPU does not need to estimate channels.

We assume that the k-th UT transmits a length τ_p pilot sequence $\boldsymbol{\varphi}_k \in \mathbb{C}^{\tau_p \times 1}$, where all the pilots are orthogonal with each other and $\boldsymbol{\varphi}_k^H \boldsymbol{\varphi}_{k'} = \tau_p$. The random complex Gaussian pilots are considered because of their varying amplitudes, which benefits the quantisation in terms of distortion. In quantisation, if the input signals of quantiser have the same amplitude, the correlation of the quantisation distortion will be strong. We also apply the real-valued Hadamard pilots in the sub-subchapter 4.3.4. The results and related effects will be discussed in the subsequent sub-chapters. Then, we have the received pilots at the m-th AP:

$$\boldsymbol{y}_{p,m} = \sum_{k=1}^{K_{serv}} g_{mk}\, \boldsymbol{\varphi}_k^T + \sum_{k=1}^{K_{int}} g_{i,mk}\, \boldsymbol{\varphi}_{i,k}^T + \boldsymbol{z}_{p,m} \tag{4.3}$$

where the vector $\boldsymbol{\varphi}_{i,k}$ denotes the pilot sequence transmitted by the k-th interfering UT which may not be orthogonal to the served user pilots, and the vector $\boldsymbol{z}_{p,m}$ represents the additive noise at the m-th AP. The elements of $\boldsymbol{z}_{p,m}$ follows $\mathcal{CN}(0, \sigma_z^2)$.

In many papers related to cell-free MaMIMO [14, 16, 24, 72], complex pilot sequences are utilised, and these pilots are pairwise orthogonal. However, the method for generating such a set of pilot sequences is not mentioned. To address this, we provide following details: These complex pilot sequences can be generated by exploiting either the singular value decomposition (SVD) of a given complex square matrix or the eigen decomposition of a Hermitian matrix. The unitary matrix obtained from such decomposition can subsequently be used as the pilot sequences. Furthermore, we offer a concise analysis of the distribution of the pilot sequence elements. First, we assume that the elements in the decomposed complex square matrix follow the complex Gaussian distribution. Meanwhile, the Hermitian matrix can be represented as a product of this complex square matrix and its Hermitian.

In Fig. 4.1, we compare the distribution of the real part of the pilot sequence elements, using numerical results, to the standard normal distribution. Two vertical axis forms are applied, here. Fig. 4.1(a) with linear vertical axis shows a minor difference above the vertical value of 0.5 on the top. On the contrary, an obvious discrepancy is observed in the tails of the numerical distribution results in Fig. 4.1(b). This is because the size of square matrix used for decomposition is 10-by-10 which is quite small, here, leading to limitations in the tail values. If we increase the matrix size to 50-by-50, the corresponding results are shown in Fig. 4.2 where the tails are closer to the theory. From

the results, we can assume the elements of the pilots follow the complex Gaussian distribution sufficiently closely. This assumption will be discussed further in the following sub-chapters in order to solve a general issue of quantisation.

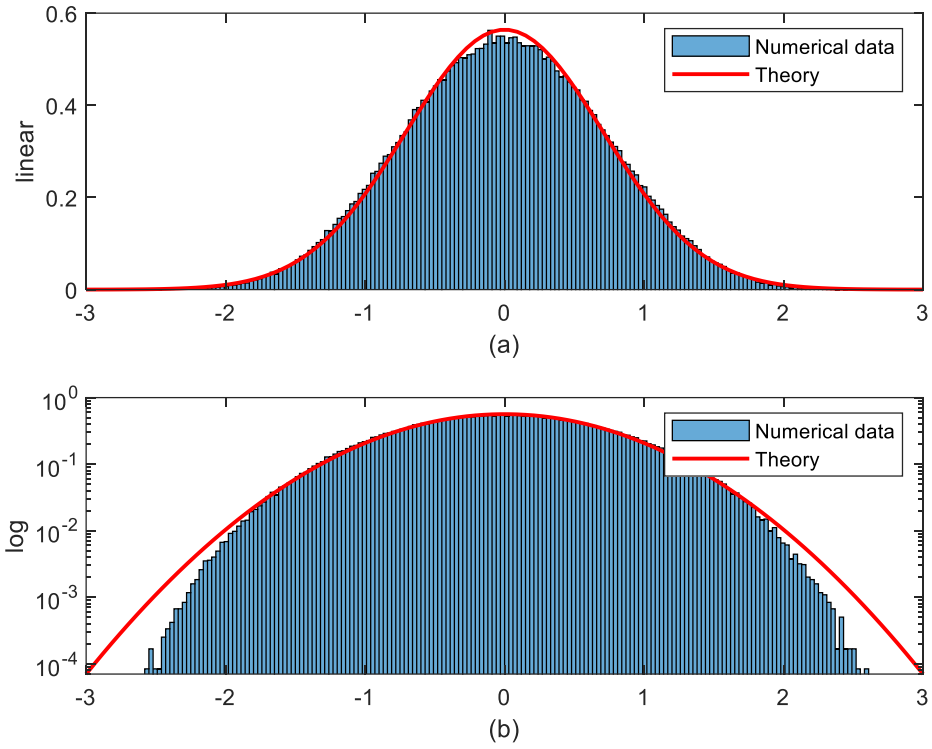


Fig. 4.1 The pdf for the real Gaussian distribution with zero mean and variance 0.5 (red solid line) versus the histogram of numerical results obtained from the pilots. (a) gives linear results of vertical axis and (b) is logarithmic. The decomposed square matrix is 10-by-10.

4.3.1 Estimate-and-Quantise

Based on the definition of the EQ scheme, the channel coefficient is estimated first, given by (3.7), at the AP and its quantised version $\hat{g}_{mk}^{eq}$ is then sent to the EPU. We use the Bussgang decomposition (2.1) to represent the quantised estimated channel as:

$$\hat{g}_{mk}^{eq} = \alpha \hat{g}_{mk} + \delta_{mk}^{eq} \tag{4.4}$$

where α denotes the linear coefficient of the Bussgang decomposition which can be calculated by (2.7), and δ_{mk}^{eq} is the quantisation distortion. In this case, the optimum quantisation step interval can be easily obtained from Table 2.1 by multiplying the standard deviation of the input which is the estimated channel, here. This process

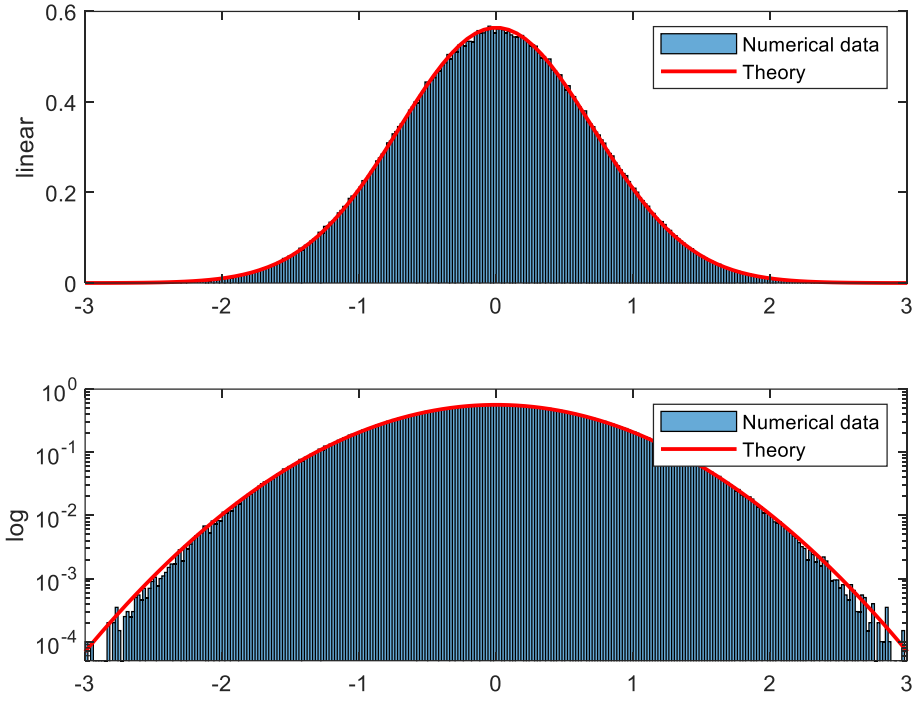


Fig. 4.2 The pdf for the real Gaussian distribution with zero mean and variance 0.5 (red solid line) versus the histogram of numerical results obtained from the pilots. The decomposed square matrix is 50-by-50.

normalises the input of the quantiser. To give the general optimum quantisation step interval we rewrite (2.9) as:

$$\begin{aligned}\Delta_{opt} &= \arg\max_{\Delta}\left(\frac{\left(\sum_{l=1}^{\frac{L}{2}-1} 2e^{-\frac{l^2\Delta^2}{2\sigma_x^2}}+1\right)^2}{4\sum_{l=1}^{\frac{L}{2}-1} lQ\left(\frac{l\Delta}{\sigma_x}\right)+\frac{1}{4}}\right)\\ &\overset{(a)}{=} \arg\max_{\Delta=\sigma_x\Delta'}\left(\frac{\left(\sum_{l=1}^{\frac{L}{2}-1} 2e^{-\frac{l^2\Delta'^2}{2}}+1\right)^2}{4\sum_{l=1}^{\frac{L}{2}-1} lQ\left(l\Delta'\right)+\frac{1}{4}}\right)\\ &= \sigma_x \arg\max_{\Delta'}\left(\frac{\left(\sum_{l=1}^{\frac{L}{2}-1} 2e^{-\frac{l^2\Delta'^2}{2}}+1\right)^2}{4\sum_{l=1}^{\frac{L}{2}-1} lQ(l\Delta')+\frac{1}{4}}\right)\end{aligned} \tag{4.5}$$

where we substitute the general quantisation step interval Δ with $\sigma_x\Delta'$ in step (a), where Δ' represents the step interval under the standard normal distribution input. Consequently, the general optimum quantisation step interval can be achieved by scaling the optimum step interval of the normalised input. In order to inspect the quality of the quantised channel estimation, it is advantageous to calculate the mean square error (MSE) of the channel estimation, which is given as:

$$\begin{aligned}
\sigma_{e_{m,k}}^2 &= \mathbb{E}\left\{\left|\hat{g}_{mk}^{eq} - g_{mk}\right|^2\right\} \\
&= \mathbb{E}\{g_{mk}g_{mk}^* + \hat{g}_{mk}^{eq}\,\hat{g}_{mk}^{eq*} - g_{mk}\hat{g}_{mk}^{eq*} - \hat{g}_{mk}^{eq}g_{mk}^*\} \\
&\overset{(a)}{=} \beta_{mk} + \mathbb{E}\Big\{\left(\alpha\hat{g}_{mk} + \delta_{mk}^{eq}\right)\left(\alpha\hat{g}_{mk} + \delta_{mk}^{eq}\right)^* - g_{mk}\left(\alpha\hat{g}_{mk} + \delta_{mk}^{eq}\right)^* \\
&\qquad - \left(\alpha\hat{g}_{mk} + \delta_{mk}^{eq}\right)g_{mk}^*\Big\} \\
&\overset{(b)}{=} \beta_{mk} + \mathbb{E}\Big\{\alpha^2|\hat{g}_{mk}|^2 + \left|\delta_{mk}^{eq}\right|^2 - (\hat{g}_{mk} + \tilde{g}_{mk})\left(\alpha\hat{g}_{mk} + \delta_{mk}^{eq}\right)^* \\
&\qquad - \left(\alpha\hat{g}_{mk} + \delta_{mk}^{eq}\right)(\hat{g}_{mk} + \tilde{g}_{mk})^*\Big\} \\
&\overset{(c)}{=} \beta_{mk} + \alpha^2\mathbb{E}\left\{|\hat{g}_{mk}|^2\right\} + \mathbb{E}\left\{\left|\delta_{mk}^{eq}\right|^2\right\} - 2\alpha\mathbb{E}\left\{|\hat{g}_{mk}|^2\right\} \\
&= \beta_{mk} + (\gamma - 2\alpha)\mathbb{E}\{|\hat{g}_{mk}|^2\} \\
&= (1 + (\gamma - 2\alpha)c_{mk})\beta_{mk}
\end{aligned} \tag{4.6}$$

where $\hat{g}_{mk}^{eq}$ is replaced by (4.4) in step (a). Based on the Bussgang decomposition, the distortion δ_{mk}^{eq} is uncorrelated with the input estimated channel $\hat{g}_{mk}$ and the channel coefficient is expanded by $\hat{g}_{mk} + \tilde{g}_{mk}$, which achieves step (b). Due to the usage of minimum mean square error (MMSE) for channel estimation, $\hat{g}_{mk}$ is uncorrelated with the estimation error and hence $\tilde{g}_{mk}$, and δ_{mk}^{eq} are also uncorrelated. This completes the step (c). The expectation term $\mathbb{E}\{|\hat{g}_{mk}|^2\}$ has already been given in (3.8).

4.3.2 Quantise-and-Estimate

Unlike the EQ scheme, the received pilots are first quantised at the APs and then delivered to the EPU for channel estimation in the QE scheme. In this case, we quantise the received pilots $\boldsymbol{y}_{p,m}$ from (4.3), and the quantised signal can be expressed using the Bussgang decomposition as:

$$\boldsymbol{y}_{p,m}^q = \alpha\boldsymbol{y}_{p,m} + \boldsymbol{\delta}_m^{qe} \tag{4.7}$$

where $\boldsymbol{\delta}_m^{qe}$ is the row vector of the quantisation distortion with τ_p elements, at the m-th AP. Then, by exploiting the information from the quantised received pilots, the channel estimation is completed at the EPU. Here, we assume that the channels among the APs are uncorrelated, hence we can perform least square (LS) estimation first from each AP separately which is shown as:

$$\begin{aligned}\breve{g}_{p,mk}^{q} &= \frac{1}{\tau_p}\boldsymbol{y}_{p,m}^{q}\boldsymbol{\varphi}_k^{*} \\ &= \alpha\left(g_{mk} + \frac{1}{\tau_p}\sum_{k'=1}^{K_{int}} g_{i,mk'}\,\boldsymbol{\varphi}_{i,k'}^{T}\boldsymbol{\varphi}_k^{*} + \frac{1}{\tau_p}\boldsymbol{z}_{p,m}\boldsymbol{\varphi}_k^{*}\right) + \frac{1}{\tau_p}\boldsymbol{\delta}_m^{qe}\boldsymbol{\varphi}_k^{*}\end{aligned} \tag{4.8}$$

We next use the MMSE to estimate the channel in the QE scheme, which is given by:

$$\hat{g}_{mk}^{qe} = c_{mk}^{qe}\breve{g}_{p,mk}^{q} \tag{4.9}$$

where the MMSE weight c_{mk}^{qe} is:

$$c_{mk}^{qe} = \frac{\mathbb{E}\{g_{mk}\breve{g}_{p,mk}^{q}{}^{*}\}}{\mathbb{E}\left\{\left|\breve{g}_{p,mk}^{q}\right|^2\right\}} \tag{4.10}$$

where

$$\mathbb{E}\{g_{mk}\breve{g}_{p,mk}^{q}{}^{*}\} = \alpha\beta_{mk} \tag{4.11}$$

$$\begin{aligned}\mathbb{E}\left\{\left|\breve{g}_{p,mk}^{q}\right|^2\right\} &\overset{(a)}{=} \alpha^2\beta_{mk} + \frac{\alpha^2}{\tau_p^2}\sum_{k'=1}^{K_{int}}\beta_{i,mk'}\,\mathbb{E}\left\{\left|\boldsymbol{\varphi}_{i,k'}^{T}\boldsymbol{\varphi}_k^{*}\right|^2\right\} + \frac{\alpha^2}{\tau_p}\sigma_z^2 + \frac{1}{\tau_p}\sigma_{\delta_m}^2 \\ &\overset{(b)}{=} \alpha^2\beta_{mk} + \frac{\alpha^2}{\tau_p}\sum_{k'=1}^{K_{int}}\beta_{i,mk'} + \frac{\alpha^2}{\tau_p}\sigma_z^2 + \frac{1}{\tau_p}\sigma_{\delta_m}^2\end{aligned} \tag{4.12}$$

where σ_z^2 and $\sigma_{\delta_m}^2$ are the variance of the noise and distortion. Step (a) is satisfied because the distortion term in (4.8) is uncorrelated with the rest. In addition, it is assumed that the quantisation distortion is uncorrelated among the different pilot symbol periods at the same AP using the pilots defined in Chapter 3. This assumption is expressed as $\mathbb{E}\{\delta_{mi}\delta_{mj}^{*}\} = 0,\ i \neq j$ where δ_{mi} and δ_{mj} are the i-th and j-th element of $\boldsymbol{\delta}_m$, respectively. Further, we assume $\mathbb{E}\left\{\left|\boldsymbol{\varphi}_{i,k'}^{T}\boldsymbol{\varphi}_k^{*}\right|^2\right\} = \tau_p$ which fulfils step (b).

Following (2.4), the variance of non-linear distortion $\sigma_{\delta_m}^2$ for the Bussgang decomposition of the received pilot sequences at the m-th AP is given by:

$$\sigma_{\delta_m}^2 = (\gamma - \alpha^2)\sigma_{y_{p,m}}^2 \tag{4.13}$$

where $\sigma_{y_{p,m}}^2$ is the variance of the received pilot sequences at the m-th AP (it is the element in one symbol period):

$$\sigma_{y_{p,m}}^2 = \mathbb{E}\{|y_{p,mi}|^2, i = 1, \dots, \tau_p\} = \sum_{k=1}^{K_{serv}} \beta_{mk} + \sum_{k=1}^{K_{int}} \beta_{i,mk} + \sigma_z^2 \tag{4.14}$$

Substitute (4.13) and (4.14) to (4.12):

$$\begin{aligned} \mathbb{E}\left\{\left|\breve{g}_{p,mk}^{q}\right|^2\right\} = \alpha^2 &\left(\beta_{mk} + \frac{1}{\tau_p}\sum_{k'=1}^{K_{int}} \beta_{i,mk'} + \frac{1}{\tau_p}\sigma_z^2\right) \\ &+ \frac{(\gamma - \alpha^2)}{\tau_p}\left(\sum_{k'=1}^{K_{serv}} \beta_{mk'} + \sum_{k'=1}^{K_{int}} \beta_{i,mk'} + \sigma_z^2\right) \end{aligned} \tag{4.15}$$

Then the MMSE channel estimation weight can be represented as:

$$c_{mk}^{qe} = \frac{\alpha\beta_{mk}}{\alpha^2\beta_{mk} + \frac{\gamma - \alpha^2}{\tau_p}\sum_{k'=1}^{K_{serv}} \beta_{mk'} + \frac{\gamma}{\tau_p}\left(\sum_{k'=1}^{K_{int}} \beta_{i,mk'} + \sigma_z^2\right)} \tag{4.16}$$

The MSE of the channel estimation in the QE scheme is derived as:

$$\begin{aligned} \sigma_{e_{m,k}}^2 &= \mathbb{E}\left\{\left|\hat{g}_{mk}^{qe} - g_{mk}\right|^2\right\} = \mathbb{E}\left\{\left(\hat{g}_{mk}^{qe} - g_{mk}\right)\left(\hat{g}_{mk}^{qe\,*} - g_{mk}^{*}\right)\right\} \\ &= \mathbb{E}\left\{\hat{g}_{mk}^{qe}\hat{g}_{mk}^{qe\,*} - g_{mk}\hat{g}_{mk}^{qe\,*} - \hat{g}_{mk}^{qe}g_{mk}^{*} + g_{mk}g_{mk}^{*}\right\} \\ &= \mathbb{E}\left\{\hat{g}_{mk}^{qe}\hat{g}_{mk}^{qe\,*} - \left(\tilde{g}_{mk} + \hat{g}_{mk}^{qe}\right)\hat{g}_{mk}^{qe\,*} - \hat{g}_{mk}^{qe}\left(\tilde{g}_{mk}^{*} + \hat{g}_{mk}^{qe\,*}\right)\right. \\ &\qquad \left. + g_{mk}g_{mk}^{*}\right\} \\ &= \mathbb{E}\left\{g_{mk}g_{mk}^{*} - \hat{g}_{mk}^{qe}\hat{g}_{mk}^{qe\,*}\right\} \\ &= \beta_{mk} - c_{mk}^{qe\,2}\mathbb{E}\left\{\left|\breve{g}_{p,mk}^{q}\right|^2\right\} \\ &= \beta_{mk} - c_{mk}^{qe\,2}\left(\alpha^2\left(\beta_{mk} + \frac{1}{\tau_p}\sum_{k'=1}^{K_{int}} \beta_{i,mk'} + \frac{1}{\tau_p}\sigma_z^2\right) + \frac{\sigma_{\delta_m}^2}{\tau_p}\right) \end{aligned} \tag{4.17}$$

$$= \left(1 - \alpha c_{mk}^{qe}\right)\beta_{mk}$$

Here, the channel estimation error $\tilde{g}_{mk}$ is uncorrelated with the estimated channel $\hat{g}_{mk}^{qe}$.

4.3.3 Non-Gaussian distribution

According to the definition of the Bussgang decomposition [65], if the input of the quantiser follows the Gaussian distribution, the Bussgang parameters have results such as (2.7) and (2.8). However, by observing (4.7) which defines the Bussgang decomposition of the received pilots in QE form at the AP, we know that the input signal (4.3) is the sum of many products of two Gaussian distributed variables plus the noise. Thus, we cannot simply exploit the results (2.7) and (2.8) to find the optimum quantisation interval. In this sub-subchapter, we will derive the Bussgang parameters from the definitions (2.2) and (2.3).

Initially, we explore the pdf of the sum of many products of two Gaussian distributed variables. To begin, a basis is considered where there exists only one product. In order to derive the pdf of the product of two real Gaussian distributed variables, the characteristic function of this product is determined. The following Lemma states the characteristic function.

Lemma 4.1. *The characteristic function for the product Z of two real Gaussian distributed variables X and Y, with zero mean and the variance* $\sigma_1^2\sigma_2^2$.

$$\varphi_Z(t) = \mathcal{M}_Z(it) = \mathcal{M}_{XY}(it) = \frac{1}{\sqrt{1 + \sigma_1^2\sigma_2^2 t^2}} \tag{4.18}$$

where i *denotes the imaginary unit and* $\mathcal{M}(\cdot)$ *represents the moment-generating function.*

Proof: First, we think about the product of two Gaussian distributed variables where $X \sim \mathcal{N}(0, \sigma_1^2)$, $Y \sim \mathcal{N}(0, \sigma_2^2)$ and $Z = XY$. The moment-generating function is given by:

$$
\begin{aligned}
&\mathcal{M}_{XY}(t) = \mathbb{E}\{e^{xyt}\} \\
&= \frac{1}{\sqrt{2\pi}\sigma_1}\frac{1}{\sqrt{2\pi}\sigma_2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-\frac{x^2}{2\sigma_1^2}} e^{-\frac{y^2}{2\sigma_2^2}} e^{xyt}\, dx\, dy \\
&= \frac{1}{\sqrt{2\pi}\sigma_2}\int_{-\infty}^{\infty} e^{-\frac{y^2}{2\sigma_2^2}} dy \int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}\sigma_1} e^{-\frac{\left(x-\sigma_1^2 yt\right)^2+\sigma_1^4 y^2 t^2}{2\sigma_1^2}} dx \\
&= \frac{1}{\sqrt{2\pi}\sigma_2}\int_{-\infty}^{\infty} e^{-\frac{\left(1-\sigma_1^2\sigma_2^2 t^2\right)y^2}{2\sigma_2^2}} dy \\
&\quad = \frac{1}{\sqrt{2\pi}\sigma_2}\int_{-\infty}^{\infty} e^{-\frac{\left(1-\sigma_1^2\sigma_2^2 t^2\right)y^2}{2\sigma_2^2}} dy \\
&\quad = \frac{1}{\sqrt{2\pi}\sigma_2}\sqrt{2\pi}\frac{\sigma_2}{\sqrt{1-\sigma_1^2\sigma_2^2 t^2}} = \frac{1}{\sqrt{1-\sigma_1^2\sigma_2^2 t^2}}
\end{aligned}
\tag{4.19}
$$

Then, the characteristic function for the product of two real Gaussian distributed variables is proved by substituting the variable t with it.

Theorem 4.1. *The probability density function for the product Z of two real Gaussian distributed variables, with zero mean and the variance $\sigma_1^2\sigma_2^2$.*

$$
f_Z(z) = \frac{K_0\left(\frac{|z|}{\sigma_1\sigma_2}\right)}{\pi\sigma_1\sigma_2} \tag{4.20}
$$

where $K_0(\cdot)$ is modified Bessel Function of the second kind in zero order.

Proof: By using the inverse theorem of characteristic function to Lemma 4.1, the pdf of the product Z is expressed as:

$$
\begin{aligned}
f_Z(z) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-itz}\varphi_Z(t)\, dt \\
&= \frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{\cos tz}{\sqrt{1+\sigma_1^2\sigma_2^2 t^2}} dt \\
&= \frac{K_0\left(\frac{|z|}{\sigma_1\sigma_2}\right)}{\pi\sigma_1\sigma_2}
\end{aligned}
\tag{4.21}
$$

where the last line follows the equation (3.754.2) in [74], which completes the proof.

In Theorem 4.1, the pdf only depends on the product of σ_1 and σ_2. Fig. 4.3 shows the pdf of the product of two real Gaussian distributions with different values of σ_1 and σ_2. It also indicates that changing the standard deviation of the product only varies the scale of the pdf.

We next discuss the scenario involving the summation of multiple products. An assumption is made that $\{Z_k\} = \{Z_1, Z_2, \dots, Z_n\}$ comprises independent and identically distributed (i.i.d) random variables, each resulting from the product of two real Gaussian distributed variables. Employing the same approach as Theorem 4.1, the characteristic function of the sum of these variables $Z = \sum_{k=1}^{n} Z_k$ is calculated and the pdf is defined in the following Theorem, based on [75].

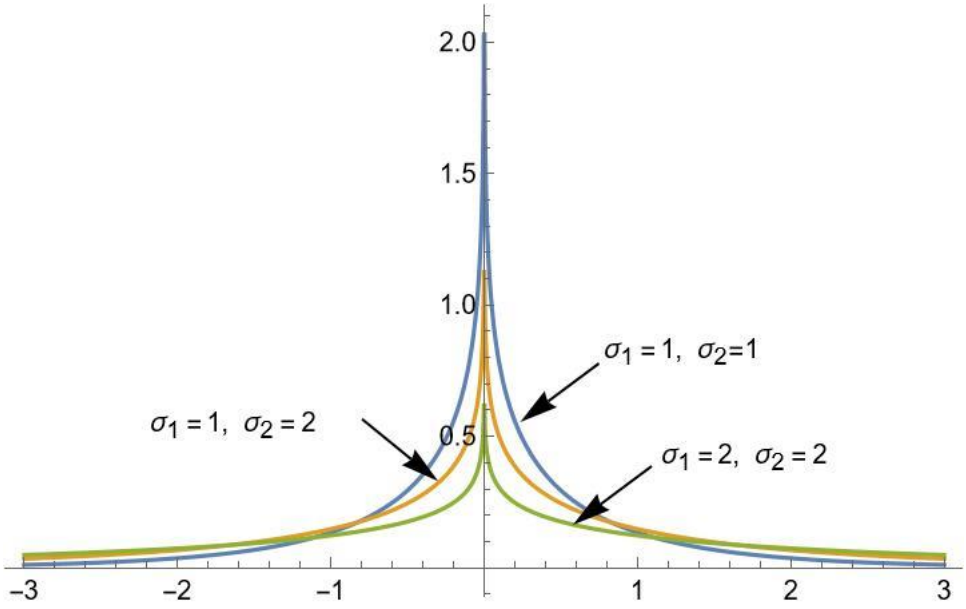


Fig. 4.3 An example: the pdf of the product of two real Gaussian distributions with different standard deviations.

Theorem 4.2. *The probability density function for the sum of n independent and identically distributed products of two real Gaussian distributions, with zero mean and the variance $\sigma_1^2\sigma_2^2$ for each product.*

$$f_Z(z) = \frac{\beta^{\frac{n+1}{2}}|z|^{\frac{n-1}{2}}2^{\frac{1-n}{2}}}{\sqrt{\pi}\Gamma\left(\frac{n}{2}\right)} K_{\frac{1-n}{2}}(\beta|z|) \tag{4.22}$$

where $\beta = (\sigma_1\sigma_2)^{-1}$, $\Gamma(\cdot)$ is the gamma function, and $K_0(\cdot)$ represents modified Bessel Function of the second kind in zero order.

Proof: Applying the Lemma 4.1, the characteristic function of the sum of n products $\{Z_1,\ Z_2, \dots ,\ Z_n\}$ is given by:

$$\varphi_Z(t) = \mathbb{E}\{e^{itz}\} = \mathbb{E}\{e^{it\sum_{k=1}^{n} z_k}\} = \{\mathbb{E}(e^{itz_k})\}^n = (1+\sigma_1^2\sigma_2^2t^2)^{-\frac{n}{2}} \tag{4.23}$$

Using the inverse theorem to obtain the pdf:

$$\begin{aligned} f_Z(z) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-izt}[(1-i\sigma_1\sigma_2t)(1+i\sigma_1\sigma_2t)]^{-\frac{n}{2}}dt \\ &\overset{(a)}{=} \frac{\beta^n}{2\pi}\int_{-\infty}^{\infty} e^{-izt}[(\beta-it)(\beta+it)]^{-\frac{n}{2}}\,dt \\ &\overset{(b)}{=} \frac{\beta^{\frac{n}{2}}|z|^{\frac{n}{2}-1}}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)}\mathcal{W}_{0,\frac{1-n}{2}}(2\beta|z|) \\ &= \frac{\beta^{\frac{n}{2}}|z|^{\frac{n}{2}-1}}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)}\sqrt{\frac{2\beta|z|}{\pi}}K_{\frac{1-n}{2}}(\beta|z|) \\ &= \frac{\beta^{\frac{n+1}{2}}|z|^{\frac{n-1}{2}}2^{\frac{1-n}{2}}}{\sqrt{\pi}\Gamma\left(\frac{n}{2}\right)}K_{\frac{1-n}{2}}(\beta|z|) \end{aligned} \tag{4.24}$$

where $\mathcal{W}_{u,v}(\cdot)$ denotes the Whittaker function which can be converted to the modified Bessel function of the second kind by the fact $\mathcal{W}_{0,v}(z) = \sqrt{z/\pi}K_v(z/2)$. Following equation (3.384.9) in [74], we simplify the equation (a) to (b) in (4.24) which completes the proof.

To facilitate comparison with the standard normal distribution input case in the original Bussgang decomposition, it is desirable to standardise the pdf (4.22). This involves normalising the variable z in the pdf such that the variance of z equals one. Since z consists of the sum of n products, we assume $X_k \sim \mathcal{N}\left(0,1/\sqrt{n}\right)$, and $Y_k \sim \mathcal{N}\left(0,1/\sqrt{n}\right)$. In this way, the variance of the sum of n products $Z = \sum_{k=1}^{n} X_kY_k$ is identical to one. Then, we substitute $\sigma_1\sigma_2 = 1/\sqrt{n}$ into (4.22) and obtain the standardised version:

$$f_Z(z) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-izt}[(1-i\sigma_1\sigma_2t)(1+i\sigma_1\sigma_2t)]^{-\frac{n}{2}}dt \tag{4.25}$$

$$= \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-izt}\left[\left(1-\frac{it}{\sqrt{n}}\right)\left(1+\frac{it}{\sqrt{n}}\right)\right]^{-\frac{n}{2}} dt$$

$$= \frac{n^{\frac{n}{2}}2^{\frac{1-n}{2}}\left|\frac{z}{\sqrt{n}}\right|^{\frac{n-1}{2}}}{\sqrt{\pi}\Gamma\left(\frac{n}{2}\right)}\mathcal{K}_{\frac{1-n}{2}}\left(n\left|\frac{z}{\sqrt{n}}\right|\right)$$

where the last line follows the same calculation as (3.384.9) in [74]. From the results, we can observe that the variable is normalised as $z/\sqrt{n}$.

In the case of complex variables, assuming that $X = \mathrm{Re}(X) + i\,\mathrm{Im}(X)$ and $Y = \mathrm{Re}(Y) + i\,\mathrm{Im}(Y)$ are the independent complex Gaussian distributions, the product $Z = XY$ is then expanded to four product variables as:

$$\begin{aligned} Z &= XY \\ &= \mathrm{Re}(X)\mathrm{Re}(Y) - \mathrm{Im}(X)\mathrm{Im}(Y) + i(\mathrm{Re}(X)\mathrm{Im}(Y) + \mathrm{Im}(X)\mathrm{Re}(Y)) \end{aligned} \tag{4.26}$$

where the variances of each product are the same. Therefore, for the complex Gaussian distribution, the number of product variables in the real or imaginary part is twice that of the real Gaussian distribution case. Fig. 4.4 presents a comparison of the pdf for the standard normal distribution and the standardised pdf for the sum of two products i.e. setting $n = 2$ in (4.25). Note that these products are from the real part of a single complex signal.

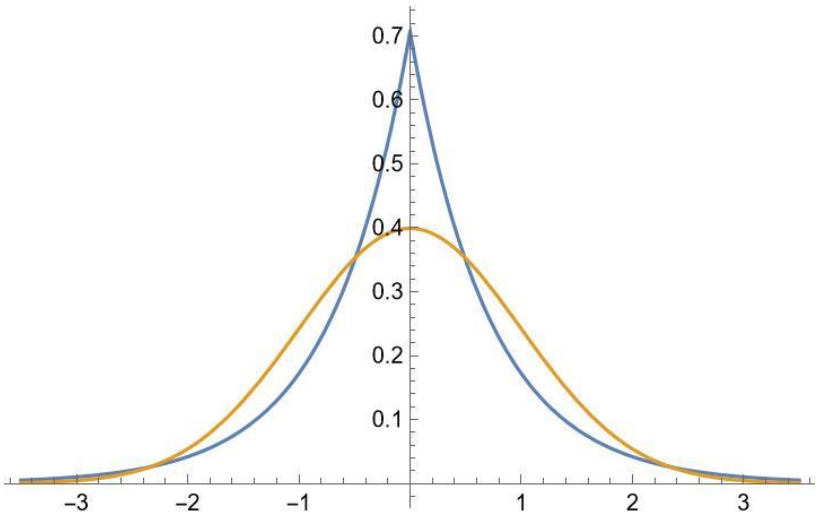


Fig. 4.4 The pdf of standard normal distribution (orange) versus the standardised pdf of sum of two product variables (blue) where setting $n = 2$ in (4.25).

From Fig. 4.4, it is clear that the pdf for the non-Gaussian case exhibits a marked difference from the standard normal distribution particularly when the number of product variables is very low. It is complicated to give a closed-form expression of α and γ in the case where all products are i.i.d variables. But it is easy to calculate the results from a tool such as Mathematica. In the following tables, we present the optimum step interval and Bussgang parameters for the two scenarios. Table 4.1 shows the optimum step interval and Bussgang parameters when using 2-bit quantisation with varying the numbers of product variables. Table 4.2 displays the results obtained by changing the number of quantisation bits while maintaining 2 product variables.

Table 4.1 Optimum step interval and Bussgang parameters for 2-bit quantisation with different number of products (identical distribution)

No. of products, n	Δ_{opt}	α	$\gamma - \alpha^2$
2	1.087393	0.803698	0.157768
4	1.064446	0.839825	0.134519
6	1.047761	0.852429	0.125794
8	1.036895	0.858987	0.121128
10	1.029531	0.86306	0.118188
20	1.013169	0.871665	0.111865
30	1.007381	0.874714	0.109589
50	1.002701	0.877232	0.107696
100	0.9991864	0.8791703	0.10622987

By observing results from Table 4.1 and comparing with the result of 2-bit quantisation from Table 2.1, the optimum step interval of the Non-Gaussian distribution case approaches $\Delta_{opt} = 0.995687$ which is the result for the Gaussian distribution case, as the number of products increases. So, a very large number of products can provide the same outcome as Gaussian distribution case. This also indicates the fact that the Bussgang decomposition cannot be applied directly when the input of the quantiser is the non-Gaussian distribution, especially with very small number of variables.

Table 4.2 Optimum step interval and Bussgang parameters for the sum of two real products with different number of quantisation bits (identical distribution)

No. of bits, b	Δ_{opt}	α	$\gamma - \alpha^2$
2	1.087393	0.803698	0.157768

3	0.730933	0.928252	0.0666
4	0.460995	0.974649	0.024708
5	0.279989	0.991287	0.008637
6	0.165681	0.997087	0.00290494
7	0.0960989	0.999051	0.000947697
8	0.0548443	0.999699	0.000301329
9	0.03088278	0.999906	0.000093722
10	0.017195504	0.999971	0.000028602
11	0.00948389	0.999991	8.58738×10^{-6}
12	0.00518861	0.999997	2.54207×10^{-6}

Comparing Table 4.2 with Table 2.1, it is evident that there exists a considerable discrepancy of over ten percent in the optimum step interval, when the number of quantisation bits surpasses 3. This contrast becomes more obvious once the number of quantisation bits increases to 11, with a significant fifty percent difference. There is also a notable increase of distortion power between two cases.

Since it is hard to derive the pdf for the sum of multiple i.i.d products plus a Gaussian distribution, we will not discuss that scenario in this thesis. But, in practice, the variances of complex product variables are not the same. For instance, the complex products in (4.3) $g_{mk}\boldsymbol{\varphi}_k^T$ have different variances when k varies because of the different path loss. Hence, it is interesting to discuss the non-Gaussian distribution input case in quantisation with non-identical variables. We assume $\{Z_k\} = \{Z_1,\ Z_2, \dots,\ Z_n\}$, n is even, are $n/2$ pairs of real independent and identically distributed product variables, where the variances differ among the $n/2$ pairs. Note that the variables in each pair have identical variance since they are obtained from the real part of one complex product variable.

To give a simple example, we firstly research the scenario with one complex product variable where there are two real products in both the real and the imaginary parts. Then, without loss of generality, we discuss the case when the number of the complex non-Gaussian distributed variables is larger than one. The following theorem states the result in two scenarios.

Theorem 4.3. *The pdf of the sum of n/2 pairs of real independent and identically distributed product variables with zero mean and the variance b_m for the m-th pair, where n is even. Here, the product variable denotes the product of two real Gaussian distributed variables.*

When $n = 2$ which means there is only one pair of real product variables $\{Z_1,\ Z_2\}$, the pdf of $z = z_1 + z_2$ is given as:

$$f_Z(z) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{e^{-itz}}{1+b_1t^2}dt = \frac{e^{-\frac{|z|}{\sqrt{b_1}}}}{2\sqrt{b_1}}\ , b_1 > 0 \tag{4.27}$$

where b_1 is the variance of each product variable.

When $n > 2$, the pdf of $z = \sum_{k=1}^{n} z_k$ is given by:

$$\begin{aligned} f_Z(z) &= \frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{e^{-itz}}{\prod_{m=1}^{\frac{n}{2}}(1+b_mt^2)}dt \\ &= \frac{1}{2}\sum_{m=1}^{n/2}\frac{b_m^{\frac{n-3}{2}}e^{-\frac{|z|}{\sqrt{b_m}}}}{\prod_{m'\neq m}^{n/2}(b_m - b_{m'})}\ , b_m \neq b_{m'} \end{aligned} \tag{4.28}$$

where b_m is the variance of m-th pair of product variables.

Proof: The case when there is only two real product variables (4.27) is proved in Appendix 4.A. The $n > 2$ case (4.28) is proved in Appendix 4.B.

Further, we consider the more realistic pdf in practice, which includes the sum of many products of two real Gaussian distributed variables plus a real noise variable which follows the Gaussian distribution. Based on Theorem 4.3, we obtain the pdf of y which consists of the real part of $n/2$ complex products and a noise variable in the following theorem.

Theorem 4.4. *The pdf of the sum of n/2 pairs of real independent and identically distributed product variables with zero mean and the variance b_m for the m-th pair, plus a real Gaussian distributed variable $x \sim \mathcal{N}(0, \sigma_x^2)$, where n is even.*

When $n = 2$, the pdf of $y = z + x$, where $z = z_1 + z_2$, is given by:

$$f_Y(y)=\frac{1}{4\sqrt{b_1}}\left(e^{-\frac{y}{\sqrt{b_1}}+\frac{\sigma_x^2}{2b_1}}\operatorname{erfc}\left(\frac{-\sqrt{b_1}y+\sigma_x^2}{\sqrt{2\sigma_x^2 b_1}}\right)+e^{\frac{y}{\sqrt{b_1}}+\frac{\sigma_x^2}{2b_1}}\operatorname{erfc}\left(\frac{\sqrt{b_1}y+\sigma_x^2}{\sqrt{2\sigma_x^2 b_1}}\right)\right) \tag{4.29}$$

where b_1 is the variance of each product variable, σ_x^2 denotes the variance of the real Gaussian distributed variable and $\operatorname{erfc}(\cdot)$ *represents the complementary error function.*

When $n>2$, the pdf of $y=z+x$, where $z=\sum_{k=1}^{n} z_k$, is given by:

$$f_Y(y)=\frac{1}{4}\sum_{m=1}^{\frac{n}{2}}\frac{b_m^{\frac{n-3}{2}}}{\prod_{m'\neq m}^{\frac{n}{2}}(b_m-b_{m'})}\left(e^{-\frac{y}{\sqrt{b_m}}+\frac{\sigma_x^2}{2b_m}}\operatorname{erfc}\left(\frac{-\sqrt{b_m}y+\sigma_x^2}{\sqrt{2\sigma_x^2 b_m}}\right)+e^{\frac{y}{\sqrt{b_m}}+\frac{\sigma_x^2}{2b_m}}\operatorname{erfc}\left(\frac{\sqrt{b_m}y+\sigma_x^2}{\sqrt{2\sigma_x^2 b_m}}\right)\right) \tag{4.30}$$

where b_m is the variance of the m-th pair of product variables.

Proof: We first prove (4.29) by exploiting the pdf of z in Theorem 4.3. Then, the pdf of $y=z+x$ is calculated by the convolution:

$$\begin{aligned}
f_Y(y)&=\int_{-\infty}^{\infty} f_X(x)\,f_Z(y-x)dx\\
&=\int_{-\infty}^{\infty}\frac{1}{\sigma_x\sqrt{2\pi}}e^{-\frac{x^2}{2\sigma_x^2}}\frac{e^{-\frac{|y-x|}{\sqrt{b_1}}}}{2\sqrt{b_1}}dx\\
&=\frac{1}{2\sigma_x\sqrt{2\pi b_1}}\left(\int_{-\infty}^{y}e^{-\frac{x^2}{2\sigma_x^2}}e^{-\frac{y-x}{\sqrt{b_1}}}dx+\int_{y}^{\infty}e^{-\frac{x^2}{2\sigma_x^2}}e^{-\frac{x-y}{\sqrt{b_1}}}dx\right)\\
&=\frac{1}{2\sigma_x\sqrt{2\pi b_1}}\left(e^{-\frac{y}{\sqrt{b_1}}+\frac{\sigma_x^2}{2b_1}}\int_{-\infty}^{y}e^{-\frac{\left(x-\frac{\sigma_x^2}{\sqrt{b_1}}\right)^2}{2\sigma_x^2}}dx+e^{\frac{y}{\sqrt{b_1}}+\frac{\sigma_x^2}{2b_1}}\int_{y}^{\infty}e^{-\frac{\left(x+\frac{\sigma_x^2}{\sqrt{b_1}}\right)^2}{2\sigma_x^2}}dx\right)\\
&=\frac{1}{4\sqrt{b_1}}\left(e^{-\frac{y}{\sqrt{b_1}}+\frac{\sigma_x^2}{2b_1}}\operatorname{erfc}\left(\frac{-\sqrt{b_1}y+\sigma_x^2}{\sqrt{2\sigma_x^2 b_1}}\right)+e^{\frac{y}{\sqrt{b_1}}+\frac{\sigma_x^2}{2b_1}}\operatorname{erfc}\left(\frac{\sqrt{b_1}y+\sigma_x^2}{\sqrt{2\sigma_x^2 b_1}}\right)\right)
\end{aligned} \tag{4.31}$$

which completes the proof. When $n > 2$, the proof is given in Appendix 4.C.

Fig. 4.5 presents a comparison between the theoretical pdf given by (4.29) and the histogram of simulated results, assuming the variances $b_1 = 0.1$ and $\sigma_x^2 = 0.01$. It shows that the histogram closely matches the theoretical formula, indicating the validity of the pdf expression for the given set of parameters. Fig. 4.5(a) presents the linear results of the vertical axis, while the logarithm results are given in Fig. 4.5(b). It is obvious that the tails of the numerical data agrees with the theory in both cases, even though there are some small discrepancies in the logarithmical results because of the calculation of the tiny values. This outcome indicates that the pilot assumed as the Gaussian distribution in Chapter 3 is not impacted by the difference of the distribution tails between Fig. 4.1 and Gaussian distribution. Thus, Theorem 4.4 is applicable in practical scenarios.

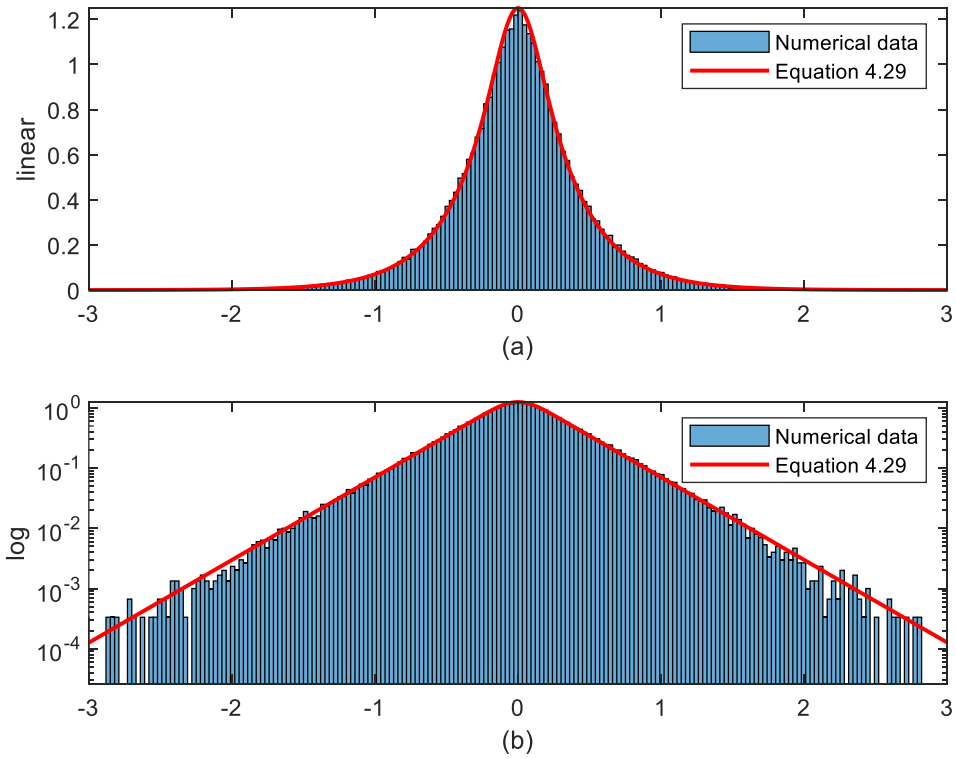


Fig. 4.5 The pdf for the sum of two real product variables plus a Gaussian distributed variable (eq. 4.29) versus the histogram of numerical results that follow the same distribution. (a) gives linear results of vertical axis and (b) is logarithm.

Next, we can apply the pdf to calculate the Bussgang parameters α and γ. Observing the definition of α (2.2), it includes the cross-correlation between the input

signal x and the output $y(x)$. If we adopt the mid-rise uniform quantiser (2.6), the cross-correlation can be expressed by solving the general indefinite integral as the following Lemma.

Lemma 4.2. *The indefinite integral of the following combination of the complementary error function with exponentials and powers, where the variable x and the coefficients a, b, c and d are real-valued,* $a, b \neq 0$*:*

$$\begin{aligned}\int xe^{bx+d}\mathrm{erfc}(ax+c)\,dx \\ = \left(\frac{1}{b^2}-\frac{1}{2a^2}+\frac{c}{ab}\right)e^{\frac{b^2}{4a^2}-\frac{bc}{a}+d}\mathrm{erfc}\left(ax+c-\frac{b}{2a}\right) \\ +\frac{bx-1}{b^2}e^{bx+d}\mathrm{erfc}(ax+c)-\frac{1}{ab\sqrt{\pi}}e^{-(ax+c)^2+bx+d}\end{aligned} \tag{4.32}$$

Proof: Please refer to Appendix 4.D.

Similarly, γ can be calculated by solving the indefinite integral in the following Lemma.

Lemma 4.3. *The indefinite integral of the following combination of the complementary error function with exponentials where the variable x and the coefficients a, b, c and d are real-valued,* $a, b \neq 0$*:*

$$\begin{aligned}\int e^{bx+d}\mathrm{erfc}(ax+c)\,dx \\ =\frac{1}{b}\left(e^{bx+d}\mathrm{erfc}(ax+c)-e^{\frac{b^2}{4a^2}-\frac{bc}{a}+d}\mathrm{erfc}\left(ax+c-\frac{b}{2a}\right)\right)\end{aligned} \tag{4.33}$$

Proof: Please refer to Appendix 4.D.

By introducing Lemma 4.2, Lemma 4.3 and Theorem 4.4, Bussgang's parameters α and γ can be expressed by the closed-form formulae, which is stated in the following theorem.

Theorem 4.5. *The closed-form Bussgang parameters α and γ are given, when the input signal is the sum of n/2 pairs of real independent and identically distributed*

product variables with zero mean and the variance b_m for the m-th pair, plus a real Gaussian distributed variable $x \sim \mathcal{N}(0, \sigma_x^2)$, where n is even.

When $n = 2$:

$$\alpha = \frac{1}{P_x}\int_{-\infty}^{\infty} xy(x) f_X(x)dx \tag{4.34}$$

$$= \frac{\Delta}{2\sigma_x^2}\left[\sum_{l=1}^{\frac{L}{2}-1}\left(\left(l\Delta + \sqrt{b_1}\right)e^{-\frac{l\Delta}{\sqrt{b_1}}+\frac{\sigma_n^2}{2b_1}}\operatorname{erfc}\left(-\frac{l\Delta}{\sqrt{2}\sigma_n}+\frac{\sigma_n}{\sqrt{2b_1}}\right)\right.\right.$$

$$\left. - \left(l\Delta - \sqrt{b_1}\right)e^{\frac{l\Delta}{\sqrt{b_1}}+\frac{\sigma_n^2}{2b_1}}\operatorname{erfc}\left(\frac{l\Delta}{\sqrt{2}\sigma_n}+\frac{\sigma_n}{\sqrt{2b_1}}\right) + 2\sigma_n\sqrt{\frac{2}{\pi}}e^{-\frac{l^2\Delta^2}{2\sigma_n^2}}\right)$$

$$\left. + \left(\sqrt{b_1}e^{\frac{\sigma_n^2}{2b_1}}\operatorname{erfc}\left(\frac{\sigma_n}{\sqrt{2b_1}}\right) + \sigma_n\sqrt{\frac{2}{\pi}}\right)\right]$$

and

$$\gamma = \frac{1}{P_x}\int_{-\infty}^{\infty}|y(x)|^2 f_X(x)dx$$

$$= \frac{\Delta^2}{\sigma_x^2}\left[\sum_{l=1}^{\frac{L}{2}-1} l\left(e^{-\frac{l\Delta}{\sqrt{b_1}}+\frac{\sigma_n^2}{2b_1}}\operatorname{erfc}\left(-\frac{l\Delta}{\sqrt{2}\sigma_n}+\frac{\sigma_n}{\sqrt{2b_1}}\right) - \operatorname{erfc}\left(-\frac{l\Delta}{\sqrt{2}\sigma_n}\right)\right.\right. \tag{4.35}$$

$$\left. - e^{\frac{l\Delta}{\sqrt{b_1}}+\frac{\sigma_n^2}{2b_1}}\operatorname{erfc}\left(\frac{l\Delta}{\sqrt{2}\sigma_n}+\frac{\sigma_n}{\sqrt{2b_1}}\right) + \operatorname{erfc}\left(\frac{l\Delta}{\sqrt{2}\sigma_n}\right)\right)$$

$$\left. + \left(\frac{L-1}{2}\right)^2\right]$$

when $n > 2$

$$\alpha = \frac{\Delta}{2\sigma_x^2}\sum_{m=1}^{\frac{n}{2}} \frac{b_m^{\frac{n}{2}-1}}{\prod_{m'\neq m}^{\frac{n}{2}}(b_m - b_{m'})}\Bigg[\sum_{l=1}^{\frac{L}{2}-1}\Bigg((l\Delta + \sqrt{b_m})e^{-\frac{l\Delta}{\sqrt{b_m}}+\frac{\sigma_n^2}{2b_m}}\operatorname{erfc}\left(-\frac{l\Delta}{\sqrt{2}\sigma_n}+\frac{\sigma_n}{\sqrt{2b_m}}\right) - \left(l\Delta - \sqrt{b_m}\right)e^{\frac{l\Delta}{\sqrt{b_m}}+\frac{\sigma_n^2}{2b_m}}\operatorname{erfc}\left(\frac{l\Delta}{\sqrt{2}\sigma_n}+\frac{\sigma_n}{\sqrt{2b_m}}\right) + 2\sigma_n\sqrt{\frac{2}{\pi}}e^{-\frac{l^2\Delta^2}{2\sigma_n^2}}\Bigg) + \left(\sqrt{b_m}e^{\frac{\sigma_n^2}{2b_m}}\operatorname{erfc}\left(\frac{\sigma_n}{\sqrt{2b_m}}\right) + \sigma_n\sqrt{\frac{2}{\pi}}\right)\Bigg] \tag{4.36}$$

and

$$\gamma = \frac{\Delta^2}{\sigma_x^2}\sum_{m=1}^{\frac{n}{2}} \frac{b_m^{\frac{n}{2}-1}}{\prod_{m'\neq m}^{\frac{n}{2}}(b_m - b_{m'})}\Bigg[\sum_{l=1}^{\frac{L}{2}-1} l\Bigg(e^{-\frac{l\Delta}{\sqrt{b_m}}+\frac{\sigma_n^2}{2b_m}}\operatorname{erfc}\left(-\frac{l\Delta}{\sqrt{2}\sigma_n}+\frac{\sigma_n}{\sqrt{2b_m}}\right) - \operatorname{erfc}\left(-\frac{l\Delta}{\sqrt{2}\sigma_n}\right) - e^{\frac{l\Delta}{\sqrt{b_m}}+\frac{\sigma_n^2}{2b_m}}\operatorname{erfc}\left(\frac{l\Delta}{\sqrt{2}\sigma_n}+\frac{\sigma_n}{\sqrt{2b_m}}\right) + \operatorname{erfc}\left(\frac{l\Delta}{\sqrt{2}\sigma_n}\right)\Bigg) + \left(\frac{L-1}{2}\right)^2\Bigg] \tag{4.37}$$

Proof: Please refer to Appendix 4.E.

We can then calculate the optimum step interval of the quantiser by finding the maximum SDNR. Due to the lack of space, the expression is omitted but it can be obtained applying Theorem 4.5 to (2.5). The maximised SDNR in (2.5) can be simplified as:

$$\begin{aligned}\Delta_{opt} &= \arg\max_{\Delta}\left(\frac{\alpha^2}{\gamma-\alpha^2}\right) = \arg\max_{\Delta}\left(\frac{1}{\frac{\gamma}{\alpha^2}-1}\right)\\ &= \arg\max_{\Delta}\left(\frac{\alpha^2}{\gamma}\right)\end{aligned} \tag{4.38}$$

Note that there are two scenarios of the parameters α and γ, thus the optimum step interval has two different forms. Since the SDNR includes the exponential function and complementary function of Δ, it is possible to derivate the SDNR with respect to Δ. But, we will not discuss that in this thesis.

4.3.4 Binary pilot sequences

As the aforementioned discussion, if the pilot sequences follow the complex Gaussian distribution, the input signal of the quantiser is not Gaussian distribution in QE form. Hence, we analyse the realistic scenario in the last sub-subchapter. Apart from using in-phase and quadrature pilots, applying the binary pilot sequences can be an alternative approach in QE form, which enables the input signal to be the Gaussian distribution. In this sub-subchapter, we investigate the binary pilots scenario which Hadamard matrix is exploited whose elements are either $+1$ or -1 and the rows are mutually orthogonal:

$$\boldsymbol{\varphi}_H \boldsymbol{\varphi}_H^T = \tau_p \boldsymbol{I}_{\tau_p} \tag{4.39}$$

where $\boldsymbol{\varphi}_H$ is Hadamard matrix and $\boldsymbol{I}$ denotes the identity matrix. Replacing the complex pilot sequences by $\boldsymbol{\varphi}_H$, we rewrite the Bussgang decomposition of the received pilots (4.7) as:

$$\boldsymbol{y}_{p,m}^q = \alpha \left(\sum_{k=1}^{K_{serv}} g_{mk}\, \boldsymbol{\varphi}_{H,k} + \sum_{k=1}^{K_{int}} g_{i,mk}\, \boldsymbol{\varphi}_{Hi,k} + \boldsymbol{z}_{p,m} \right) + \boldsymbol{\delta}_m^{qe} \tag{4.40}$$

where $\boldsymbol{\varphi}_{H,k}$ is the pilot sequence which is the k-th row vector the Hadamard matrix, transmitted by the k-th UT and $\boldsymbol{\varphi}_{Hi,k}$ denotes the transmitted pilot sequence from the k-th interfering UT, which is a row vector randomly selected from the Hadamard matrix. The rest terms were already defined in the previous sub-chapters. If the elements in the distortion term $\boldsymbol{\delta}_m^{qe}$ are uncorrelated with each other, the covariance of the distortion is given:

$$\mathbb{E}\left\{ \boldsymbol{\delta}_m^{qe\,H} \boldsymbol{\delta}_m^{qe} \right\} = diag\overbrace{\left\{ \sigma_{\delta_m}^2, \sigma_{\delta_m}^2, \ldots, \sigma_{\delta_m}^2 \right\}}^{\tau_p} \tag{4.41}$$

where $diag\{\cdot\}$ denotes the diagonal matrix in which the entries outside the main diagonal are all zero and $\sigma_{\delta_m}^2$ is the variance of the distortion which is defined in (4.13).

As the pilot sequence is binary, each symbol of the received pilots at the same AP is composed of many channel coefficients added and subtracted together, along with the noise. The same set of channel coefficients is employed for the computation of the different symbols, which establishes a relationship among the received pilots at the same AP. This interrelation subsequently leads to a correlation among the elements of the quantisation distortion at the given AP. In particular, when a certain channel coefficient is much stronger than the others, it can dominate the relationship and produce a noticeable correlation among the elements of the quantisation distortion. Such correlation issues can influence the theoretical calculation of the MMSE estimation weights as defined in (4.10).

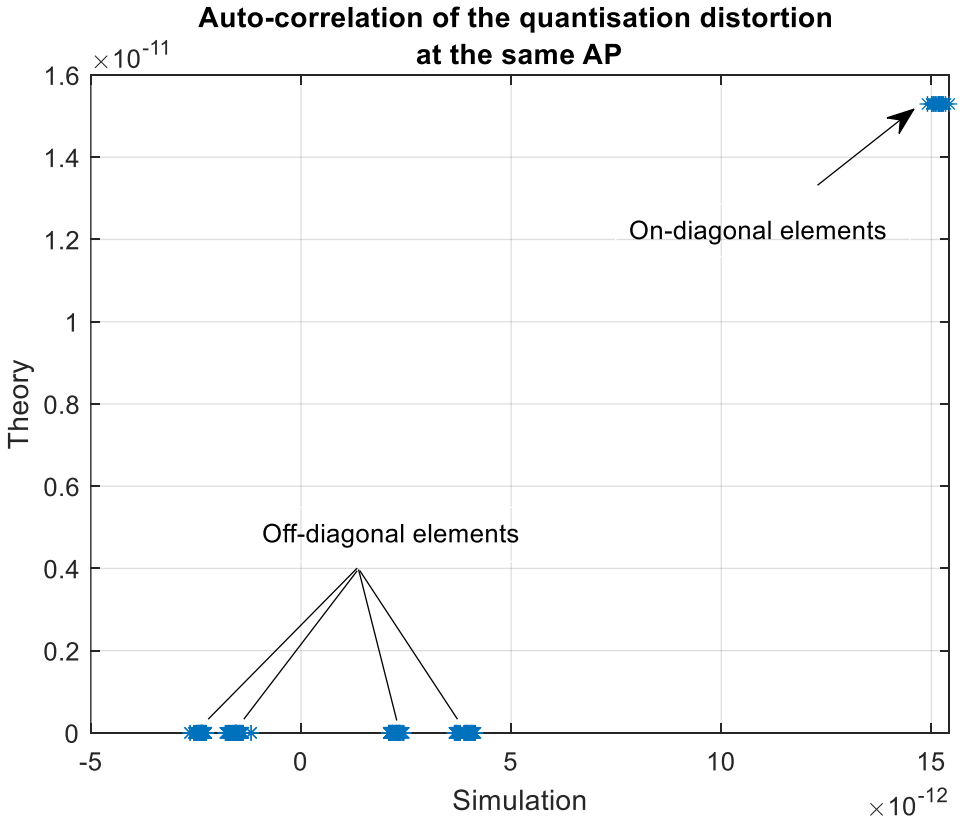


Fig. 4.6 The comparison between the simulated results and theoretical results of the auto-correlation of the quantisation distortion at the same AP.

Fig. 4.6 presents the auto-correlation of the quantisation distortion at one AP, where the horizontal axis denotes the simulated results and the vertical axis corresponds to the theoretical values derived from (4.41). In the simulation, the quantisation distortion is collected from one AP to compute its auto-correlation where the entities are marked with stars in Fig. 4.6. In the top right section of the figure, it is evident that the

on-diagonal elements of the simulation align with the theoretical predictions. However, the simulated off-diagonal elements are not zero, and their order of magnitude does not significantly differ from that of the on-diagonal elements. It proves that the different symbols of the quantisation distortion at the same AP are correlated.

Fig. 4.7 illustrates the comparison between the simulation (4.10) and the theory (4.16) of the MMSE estimation weights for channel estimation with the respective values denoted as stars. The figure reveals numerous discrepancies between the simulated and theoretical values, which indicate the theory may not be rigorous. The correlation of the quantisation distortion also impacts the theoretical calculation of the MSE for channel estimation. A more comprehensive discussion on the MSE of channel estimation will be provided in a subsequent sub-chapter.

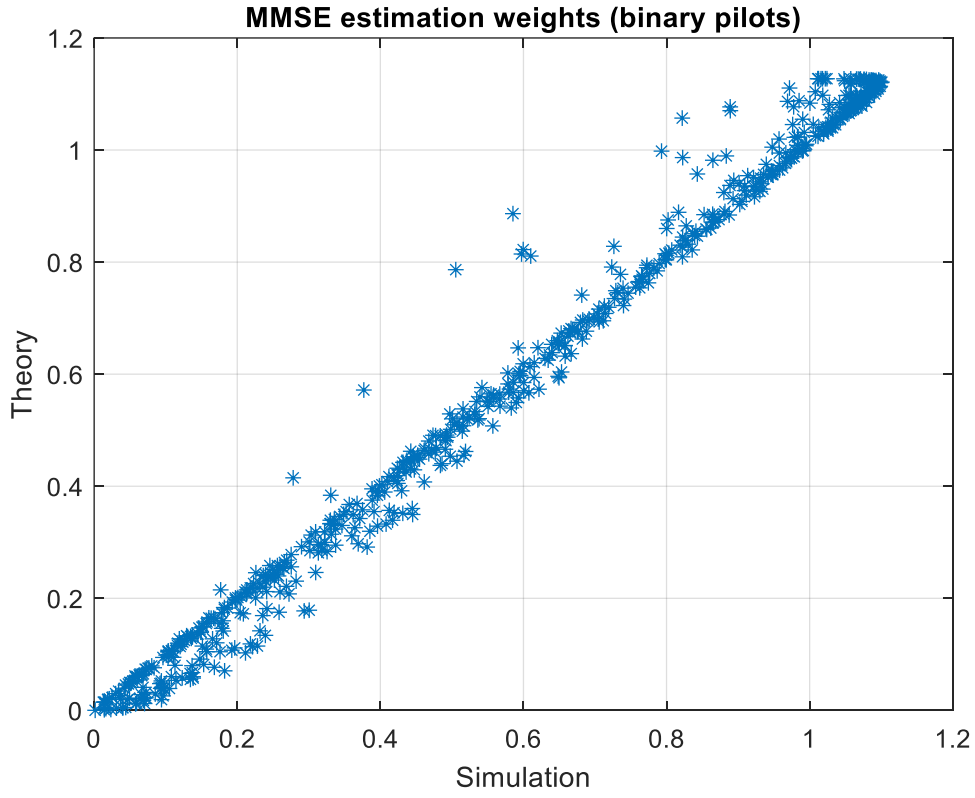


Fig. 4.7 The comparison between the simulated results and theoretical results of the MMSE estimation weights while using binary pilots.

4.4 Data Detection

After obtaining the CSI from all served UTs at the EPU, we proceed to explore the uplink data transmission under limited fronthaul constraints in DD-MaMIMO. As

discussed in the channel estimation sub-chapter, the signal can either be collected at the local AP and transmitted its quantised version to the EPU for estimation, or it can be estimated locally and then sent the quantised CSI to the EPU. In contrast to the channel estimation, this sub-chapter primarily focuses on the scenario where the transmitted data is quantised at the APs by a few bits and then sent to the EPU for detection.

On the uplink, all UTs send the data simultaneously to the M_{coor} APs where each AP equips N_r antennas and the received signals at APs can be expressed by a matrix:

$$\boldsymbol{Y}_u = \boldsymbol{G}\boldsymbol{X} + \boldsymbol{G}_{int}\boldsymbol{X}_{int} + \boldsymbol{Z} \tag{4.42}$$

where $\boldsymbol{G} \in \mathbb{C}^{M_{coor}N_r \times K_{serv}}$ and $\boldsymbol{G}_{int} \in \mathbb{C}^{M_{coor}N_r \times K_{int}}$ are channel coefficient matrices from served UTs and interfering UTs to the APs in the coordination region, separately. $\boldsymbol{X} \in \mathbb{C}^{K_{serv} \times \tau_d}$ and $\boldsymbol{X}_{int} \in \mathbb{C}^{K_{int} \times \tau_d}$ are transmitted signals. $\boldsymbol{Z}$ denotes the noise at the multiple antennas of each AP. Here we assume all data symbols are independent, so (4.42) can be rewritten in terms of column vector. Then, the received signals are quantised and it can be expressed by the Bussgang decomposition:

$$\boldsymbol{y}_u^q = \alpha \boldsymbol{y_u} + \boldsymbol{\delta} = \alpha \left(\sum_{k=1}^{K_{serv}} \boldsymbol{g}_k x_k + \sum_{k=1}^{K_{int}} \boldsymbol{g}_{i,k} x_{i,k} + \boldsymbol{z} \right) + \boldsymbol{\delta} \tag{4.43}$$

where $\boldsymbol{y_u}$ is the vector form of the received signals, and $\boldsymbol{g}_k = \left[\boldsymbol{g}_{1k}^T \, \boldsymbol{g}_{2k}^T \dots \boldsymbol{g}_{M_{coor}k}^T\right]^T$ is the k-th column vector of $\boldsymbol{G}$ in which $\boldsymbol{g}_{mk} = \left[g_{m1k} \, g_{m2k} \dots g_{mN_rk}\right]^T$ are the channel coefficients from the k-th UT to the N_r antennas of the m-th AP. $\boldsymbol{g}_{i,k}$ denotes the channel coefficients for the interfering UT, which has the same structure as $\boldsymbol{g}_k$. x_k and $x_{i,k}$ represent the transmitted data from the k-th UT and the k-th interfering UT, both with the variance σ_x^2, respectively. In this sub-chapter, we assume that binary phase-shift keying (BPSK) is applied to the data modulation, resulting in binary transmitted data with $\sigma_x^2 = 1$. $\boldsymbol{z}$ is the noise with the variance of each element σ_z^2 and $\boldsymbol{\delta}$ denotes the quantisation distortion. Depending on different CSI acquisition strategies (4.4) and (4.9), the channel coefficient $\boldsymbol{g}_k$ is expressed as:

$$\boldsymbol{g}_k = \widehat{\boldsymbol{g}}_k + \widetilde{\boldsymbol{g}}_k, \text{where } [\widehat{\boldsymbol{g}}_k]_{mnk} = \hat{g}_{mk}^{eq} \; or \; [\widehat{\boldsymbol{g}}_k]_{mnk} = \hat{g}_{mk}^{qe} \tag{4.44}$$

where $\tilde{\boldsymbol{g}}_k$ denotes the channel estimation error. To detect the transmitted data, the L-MMSE estimator is applied at the EPU in this thesis and the MMSE weight $\boldsymbol{w}_k^q$ for detecting the data transmitted by the k-th UT can be expressed by:

$$\boldsymbol{w}_k^q = \mathbb{E}\{x_k \boldsymbol{y}_u^{qH}\}\mathbb{E}\{\boldsymbol{y}_u^q \boldsymbol{y}_u^{qH}\}^{-1} \tag{4.45}$$

where

$$\mathbb{E}\{x_k \boldsymbol{y}_u^{qH}\} = \alpha\sigma_x^2 \hat{\boldsymbol{g}}_k^H \tag{4.46}$$

and

$$\begin{aligned}\mathbb{E}\left\{\left|\boldsymbol{y}_u^q\right|^2\right\} = \alpha^2\sigma_x^2 \Bigg(&\sum_{k'=1}^{K_{serv}} \left(\hat{\boldsymbol{g}}_{k'}\hat{\boldsymbol{g}}_{k'}^H + \boldsymbol{C}_{\tilde{g}\tilde{g},k'}\right) + \sum_{k'=1}^{K_{int}} \boldsymbol{C}_{g_{int}g_{int},k'} \\ &+ \frac{\sigma_z^2}{\sigma_x^2}\boldsymbol{I}_{M_{coor}N_r}\Bigg) + \boldsymbol{C}_\delta\end{aligned} \tag{4.47}$$

where $\boldsymbol{C}_{g_{int}g_{int},k}$ is the covariance matrix of interfering channel coefficients defined in (3.27). $\boldsymbol{C}_{\tilde{g}\tilde{g},k}$ denotes the covariance matrix of the channel estimation errors given by:

$$\boldsymbol{C}_{\tilde{g}\tilde{g},k} = \mathbb{E}\{\tilde{\boldsymbol{g}}_k \tilde{\boldsymbol{g}}_k^H\} = diag\left(\sigma_{e_{i,k}}^2, i = 1, \dots, M_{coor}N_r\right) \tag{4.48}$$

where the expression of $\sigma_{e_i}^2$ depends on the different quantisation schemes (4.6) or (4.17). $\boldsymbol{C}_\delta$ represents the covariance matrix of the quantisation distortion:

$$\boldsymbol{C}_\delta = \mathbb{E}\{\boldsymbol{\delta}\boldsymbol{\delta}^H\} = diag(\ \mathbb{E}\{|\delta_i|^2\}, i = 1, \dots, M_{coor}N_r\ ) \tag{4.49}$$

where δ_i is the i-th element of $\boldsymbol{\delta}$ and δ_i for the different AP is uncorrelated with each other. Exploiting (2.4) we rewrite (4.49) as:

$$\begin{aligned}\boldsymbol{C}_\delta &= (\gamma - \alpha^2)\boldsymbol{C}_{y_u} \\ &= (\gamma - \alpha^2)\left(\sigma_x^2 \sum_{k'=1}^{K_{serv}} \boldsymbol{C}_{gg,k'} + \sigma_x^2 \sum_{k'=1}^{K_{int}} \boldsymbol{C}_{g_{int}g_{int},k'} + \sigma_z^2 \boldsymbol{I}_{M_{coor}N_r}\right)\end{aligned} \tag{4.50}$$

where $\boldsymbol{C}_{y_u}$ is the covariance of the received signals at APs and $\boldsymbol{C}_{gg,k'} = \mathbb{E}\{\boldsymbol{g}_{k'}\boldsymbol{g}_{k'}^H\} = diag(\beta_{mk'}, m = 1, \dots, M_{coor}N_r)$. Then, we substitute (4.46), (4.47) and (4.50) into (4.45):

$$\boldsymbol{w}_k^q = \alpha\widehat{\boldsymbol{g}}_k^H \left(\alpha^2 \sum_{k'=1}^{K_{serv}} \left(\widehat{\boldsymbol{g}}_{k'}\widehat{\boldsymbol{g}}_{k'}^H + \boldsymbol{C}_{\tilde{g}\tilde{g},k'}\right) + \gamma \left(\sum_{k'=1}^{K_{int}} \boldsymbol{C}_{g_{int}g_{int},k'} + \frac{\sigma_z^2}{\sigma_x^2}\boldsymbol{I}_{M_{coor}N_r} \right) + (\gamma - \alpha^2) \sum_{k'=1}^{K_{serv}} \boldsymbol{C}_{gg,k'} \right)^{-1} \tag{4.51}$$

Using the MMSE weights (4.51), we then obtain the detected data transmitted by the *k*-th UT as:

$$\begin{aligned} \hat{x}_k &= \boldsymbol{w}_k^q \boldsymbol{y}_u^q \\ &= \boldsymbol{w}_k^q \left(\alpha \left(\sum_{k=1}^{K_{serv}} \boldsymbol{g}_k x_k + \sum_{k=1}^{K_{int}} \boldsymbol{g}_{i,k} x_{i,k} + \boldsymbol{z} \right) + \boldsymbol{\delta} \right) \\ &= \alpha \boldsymbol{w}_k^q \left(\widehat{\boldsymbol{g}}_k x_k + \widetilde{\boldsymbol{g}}_k x_k + \sum_{k' \neq k}^{K_{serv}} \boldsymbol{g}_{k'} x_{k'} + \sum_{k'=1}^{K_{int}} \boldsymbol{g}_{i,k'} x_{i,k'} + \boldsymbol{z} \right) + \boldsymbol{w}_k^q \boldsymbol{\delta} \end{aligned} \tag{4.52}$$

4.5 Achievable Spectral Efficiency

In this sub-chapter, we will research the spectral efficiency (SE) in the DD-MaMIMO system when the fronthaul is limited. By using the MMSE detection, the components of the estimated data (4.52) can be categorised as: the desired signal over estimated channel is $\alpha \boldsymbol{w}_k^q \widehat{\boldsymbol{g}}_k x_k$; the desired signal over unknown channel is $\alpha \boldsymbol{w}_k^q \widetilde{\boldsymbol{g}}_k x_k$; the terms $\alpha \boldsymbol{w}_k^q \sum_{k' \neq k}^{K_{serv}} \boldsymbol{g}_{k'} x_{k'}$ and $\alpha \boldsymbol{w}_k^q \sum_{k' \neq k}^{K_{int}} \boldsymbol{g}_{i,k'} x_{i,k'}$ are the interference from the intra service region and inter service region, respectively; $\alpha \boldsymbol{w}_k^q \boldsymbol{z}$ denotes the noise, and $\boldsymbol{w}_k^q \boldsymbol{\delta}$ represents the distortion. Thus, we can formulate the achievable SE in the following proposition.

Proposition 4.1. *Suppose that K_{serv} UTs are served by M_{coor} APs where each AP equips N_r antennas in DD-MaMIMO with limited fronthaul. If the MMSE estimator is used to estimate channel and detect data for all UTs, an achievable SE of k-th UT is*

$$SE_k = \frac{\tau_u}{\tau_p + \tau_u} \mathbb{E}\{\log_2(1 + SIDNR_k)\} \tag{4.53}$$

where the factor $\tau_u/(\tau_p + \tau_u)$ *is the fraction of the uplink transmission which is used by data transmission. The instantaneous signal-to-interference-and-distortion-plus-noise ratio* ($SIDNR_k$) *for the k-th UT is given by*

$$SIDNR_k = \frac{\left|\boldsymbol{w}_k^q \widehat{\boldsymbol{g}}_k\right|^2}{\boldsymbol{w}_k^q \left(\begin{matrix} \sum_{k' \neq k}^{K_{serv}} \widehat{\boldsymbol{g}}_{k'} \widehat{\boldsymbol{g}}_{k'}^H + \sum_{k'=1}^{K_{serv}} \boldsymbol{C}_{\tilde{g}\tilde{g},k'} + \sum_{k'=1}^{K_{int}} \boldsymbol{C}_{g_{int}g_{int},k'} \\ + \frac{\sigma_z^2}{\sigma_x^2} \boldsymbol{I}_{M_{coor}N_r} + \frac{1}{\alpha^2 \sigma_x^2} \boldsymbol{C}_{\delta} \end{matrix} \right) \boldsymbol{w}_k^{qH}} \tag{4.54}$$

$$\overset{(a)}{=} \alpha^2 \widehat{\boldsymbol{g}}_k^H \left(\begin{matrix} \alpha^2 \sum_{k' \neq k}^{K_{serv}} \widehat{\boldsymbol{g}}_{k'} \widehat{\boldsymbol{g}}_{k'}^H + \alpha^2 \sum_{k'=1}^{K_{serv}} \boldsymbol{C}_{\tilde{g}\tilde{g},k'} + \gamma \sum_{k'=1}^{K_{int}} \boldsymbol{C}_{g_{int}g_{int},k'} \\ + \gamma \frac{\sigma_z^2}{\sigma_x^2} \boldsymbol{I}_{M_{coor}N_r} + (\gamma - \alpha^2) \sum_{k'=1}^{K_{serv}} \boldsymbol{C}_{gg,k'} \end{matrix} \right)^{-1} \widehat{\boldsymbol{g}}_k$$

and the expectation is with respect to the channel estimates $\widehat{\boldsymbol{g}}_k$.

Proof: The sketch of the proof is the same as the proof in [73, Th. 4.1], where the expression of the achievable SE is derived by giving the lower bound of the ergodic capacity of the discrete memoryless interference channel. The input of the channel is $x = \alpha x_k$, the output is $y = \hat{x}_k$, the channel response is given by $h = \boldsymbol{w}_k^q \widehat{\boldsymbol{g}}_k$, the random realisation $u = \{\widehat{\boldsymbol{g}}_l\}$ with arbitrary values of l affects the interference variance, and the interference v is the sum of the terms (4.52) excluding $\alpha \boldsymbol{w}_k^q \widehat{\boldsymbol{g}}_k x_k$:

$$v = \alpha \boldsymbol{w}_k^q \left(\widetilde{\boldsymbol{g}}_k x_k + \sum_{k' \neq k}^{K_{serv}} \boldsymbol{g}_{k'} x_{k'} + \sum_{k'=1}^{K_{int}} \boldsymbol{g}_{i,k'} x_{i,k'} + \boldsymbol{z} \right) + \boldsymbol{w}_k^q \boldsymbol{\delta} \tag{4.55}$$

Given the realisation h and u, the interference is conditionally uncorrelated with the input:

$$\mathbb{E}\{xv|h,u\} = \mathbb{E}\{xv|\{\widehat{\boldsymbol{g}}_l\}\} = 0 \tag{4.56}$$

Furthermore, the terms of the interference v are uncorrelated with each other. The step (a) in (4.54) follows the same derivation as (3.22).

4.6 Performance Analysis

In this sub-chapter, we provide some numerical results for quantisation in the DD-MaMIMO system. The same path loss model and parameters given in Chapter 3 are utilised for the simulation. All the simulations are based on the case where each AP only equips one antenna. We first evaluate the MSE of channel estimation in the QE form, considering different distribution analyses for the input signals of the quantiser.

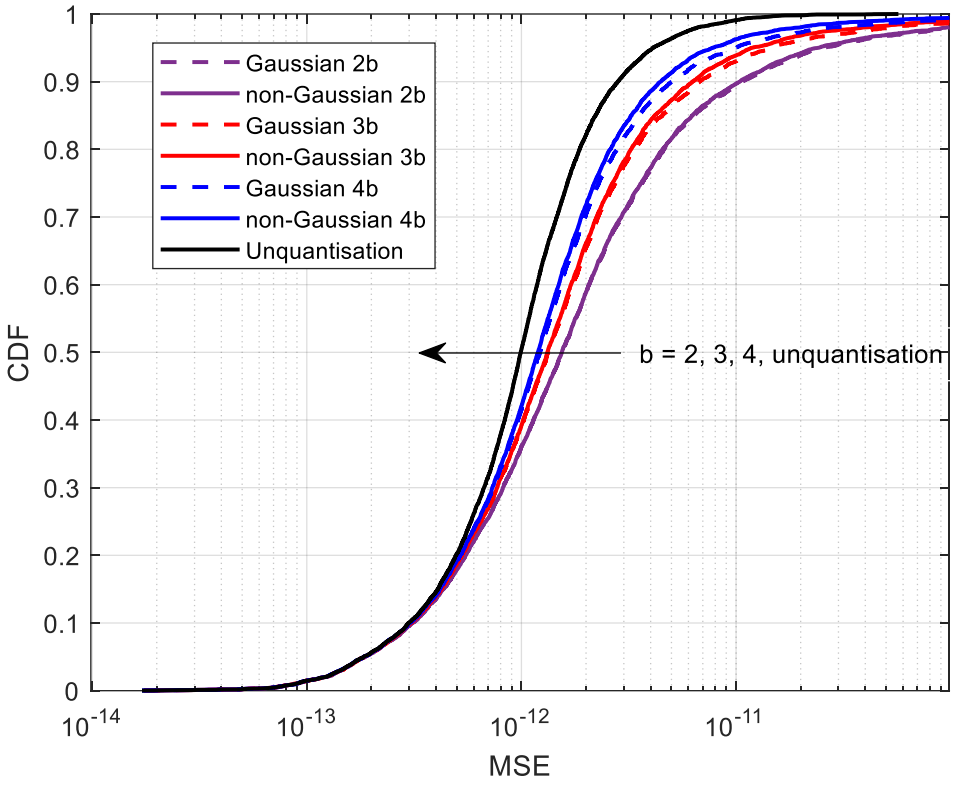


Fig. 4.8 The CDF of the MSE of channel estimation with a range of quantisation bits, based on different distributions of the input of the quantiser.

Fig. 4.8 shows the CDF of the MSE of channel estimation which is obtained in (4.17) using the Bussgang decomposition. Here, we assume the coordination region radius is 282 m in the DD-MaMIMO system. The legends "Gaussian 2b" and "non-Gaussian 2b" denote the input signals follow the Gaussian distribution with 2-bit quantisation and the true (non-Gaussian) distribution derived from Theorem 4.4 with 2 quantisation bits, respectively. Note that since the large scale fading coefficients tend to -15 order of magnitude in our simulation, the channel gain is very small and also the value of MSE of channel estimation. In comparison to the Gaussian distribution case (dash line), the MSE of channel estimation is slightly smaller in the non-Gaussian case (solid line),

which is noticeable in the 4-bit quantisation scenario. The results indicate that the channel estimation, acquired by applying Theorem 4.4 to the Bussgang decomposition, exhibits a modest improvement, albeit not significant.

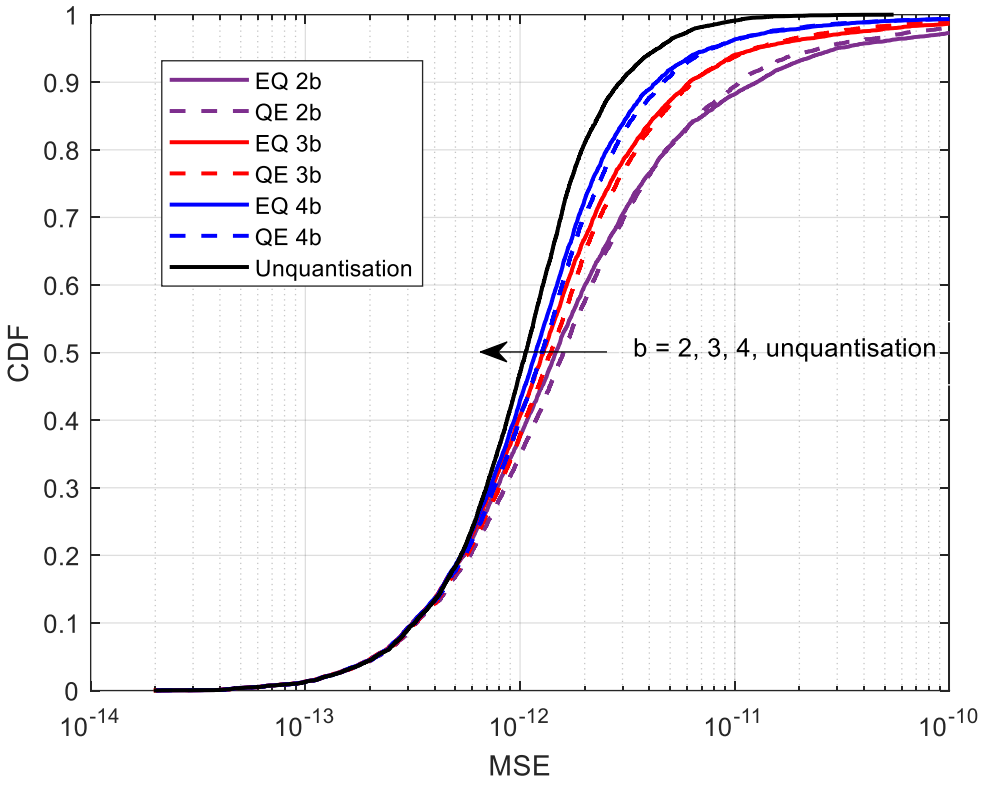


Fig. 4.9 The CDF of the MSE of channel estimation applying the different CSI acquisition strategies with a range of quantisation bits.

Next, we analyse the MSE of channel estimation by adopting different CSI acquisition strategies which are presented in (4.6) and (4.17). The same parameters are used for the simulation. The simulation results are shown in Fig. 4.9 where the legend "EQ" denotes the estimate-and-quantise form and "QE" is the quantise-and-estimate form. The results indicate that the QE form (dash line) has a poorer MSE than the EQ scheme in the most of cases, though the discrepancies are not substantial. Fig 4.8 presents the results of the MSE of channel estimation using the complex pilots and binary pilots in the QE form with 2-bit quantisation. In this simulation, we adopt the Hadamard matrix to generate binary pilots (4.39) which the matrix order must be 1, 2 or a multiple of 4, prompting us to adjust the length of the pilot sequences from 10 to 8. As we discussed in the sub-subchapter 4.3.4, the expression of the theoretical MMSE weight c_{mk}^{qe} (4.16) in the binary pilot case is not strict. Hence, in Fig. 4.10, we utilise

the simulated and the theoretical MMSE weights, respectively, to calculate the MSE for both types of pilots. The legend "Simu-weight" represents that the MMSE weights and the MSE are simulated using the Monte Carlo method, while "Theo-weight" denotes that the MSE (4.17) is computed by exploiting the theoretical MMSE weights (4.16). It is shown that the simulation using complex pilots (blue dash line with cross) match the theory (red solid line). Nevertheless, there is a small gap between the simulation (black dash line with circle) and the theory (green solid line) in the binary pilot case. This is because the correlation between the quantisation distortions at the same AP is ignored in theory when the binary pilots are used.

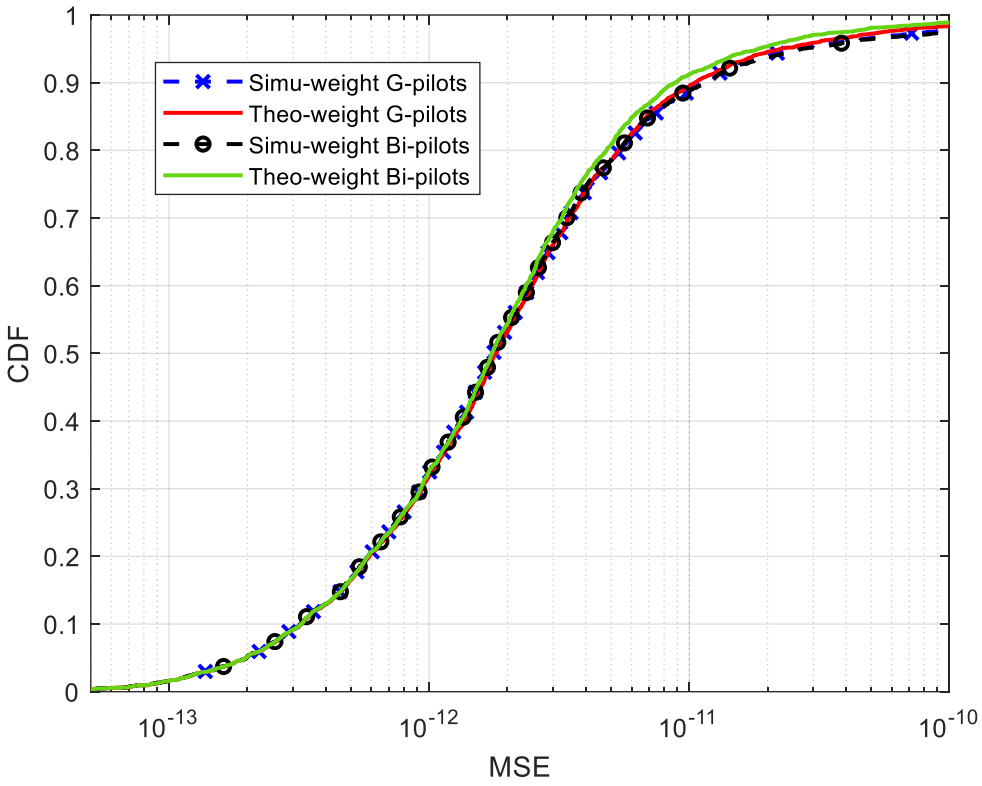


Fig. 4.10 The CDF of the MSE of channel estimation using the different pilots with 2-bit quantisation.

After the discussion of the MSE of channel estimation using different pilots, we know that there is no significant difference in obtaining the CSI in the two scenarios. Then, the overall performance, represented by the CDF of the SE, is shown in Fig. 4.11 where "Unq" denotes the unquantisation, "b" is the quantisation bit and "Bi" represents the binary. Here, the same quantisation bits for the channel estimation are used to quantise the data. We can observe that the results of SE overlap in the ideal case where

the signal is not quantised, however the binary pilot case outperforms when the quantisation is introduced. In our simulation, the theoretical MMSE weights are adopted for channel estimation in both pilot cases, leading to the error of the MSE of channel estimation in binary pilot case. As a result, the simulation of the binary pilot case shows a deviation from the complex pilot case.

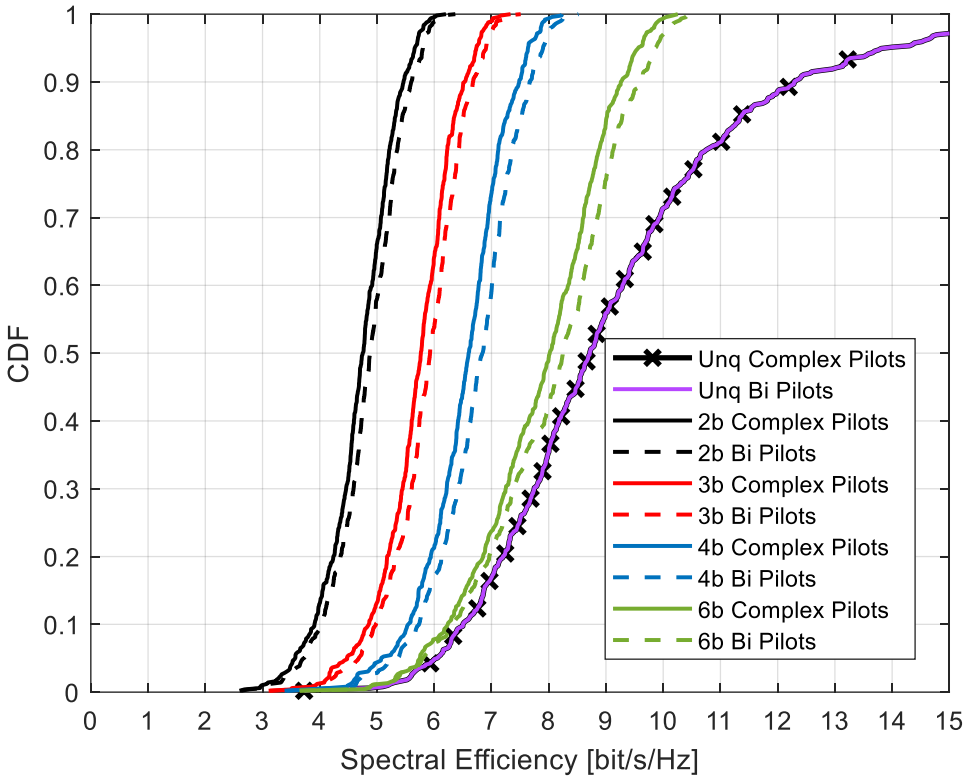


Fig. 4.11 The CDF of the SE using different pilots with a range of quantisation bits.

Applying different CSI acquisition strategies, the SE is then presented in Fig. 4.12 with varying numbers of quantisation bits. We adopt the same number of quantisation bits to quantise the pilots and data in this simulation. The results demonstrate that the QE scheme performs better than the EQ scheme when fewer quantisation bits are used. This is because each AP only needs to send the quantised pilots to the CPU in the QE form, the fronthaul load is lower compared to the EQ form, where each AP needs to send a quantised version of the estimated channel for each UT. Furthermore, as the number of quantisation bits increases, the performance gap between the two strategies narrows, and the results for both cases begin to coverage towards the ideal scenario. In the simulation of the QE form, we exploit Theorem 4.5 and (4.38) to resolve the maximisation problem by seeking the optimum step interval. The range of the step

interval needs to be decided at the outset of the simulation, which might not provide precise results. Therefore, the decrease in the performance gap may be caused by the imperfect optimum step interval finding for channel estimation in the QE form.

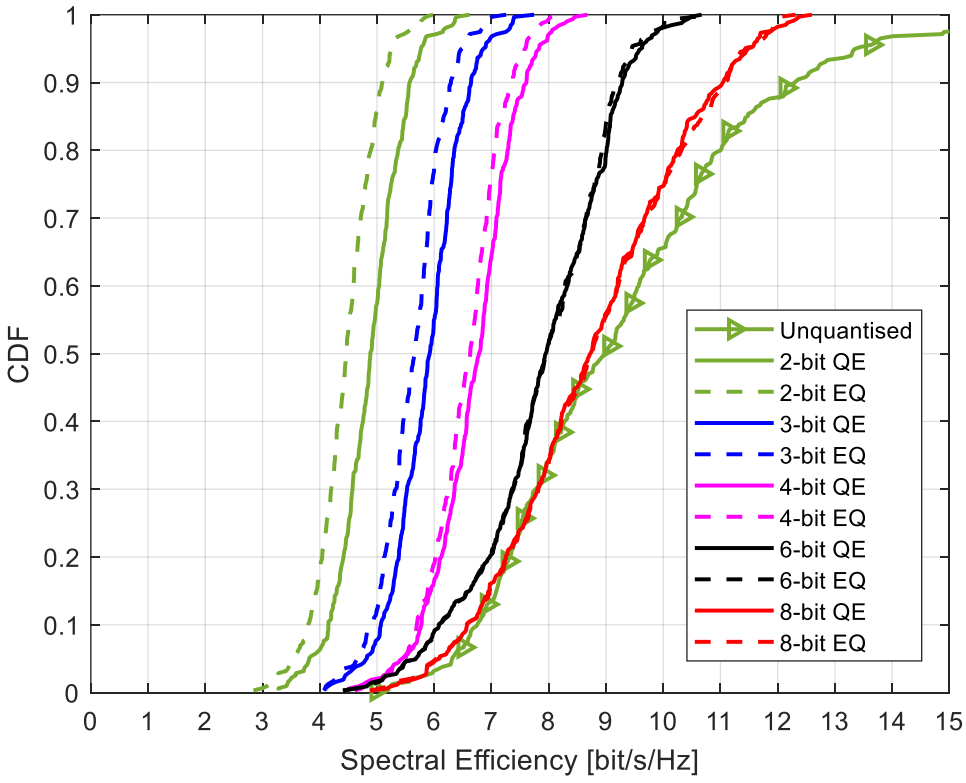


Fig. 4.12 The CDF of the SE using different CSI acquisition strategies with various quantisation bits.

4.7 Summary

In this chapter, we have presented the quantisation in the DD-MaMIMO system. The general concept was elaborated which the digital signals were transferred to the EPU for processing. To analyse the output of the quantiser, Bussgang decomposition was studied for characterising this non-linear system. We have taken account of two strategies which are estimate-and-quantise (EQ) scheme and quantise-and-estimate (EQ) scheme for obtaining CSI at the EPU. The expression of the MSE of channel estimation was derived in both cases and the performance was evaluated in the simulation. The results showed that there is no significant difference of the MSE between two strategies. Further, we have had a comprehensive research about the input distribution of the quantiser when the pilots followed the complex Gaussian distribution in the QE form.

We have proposed the theorem to express this pdf and applied that to find the optimum step interval using Bussgang decomposition. The numerical results showed that the MSE decreased adopting the proposed theorem. Moreover, the binary pilots were exploited in the QE form to avoid the issue of distribution. We have presented the performance of the binary pilot case where the MSE had no difference from the complex pilot case in the simulation. At last, we have provided the quantisation in the data transmission on the uplink of DD-MaMIMO. The achievable SE per user was derived with respect to the MMSE estimation of the channel and data, given a limited-capacity fronthaul.

The simulation results showed that the QE form with the MMSE data detection had a better SE than the EQ form in the DD-MaMIMO system. But, the disadvantage is all the estimation is processed at the EPU which leads to the high computation complexity, and the latency may be increased if the fronthaul links have a bad quality of service. In the EQ form, the computation burden at the EPU can be deducted by processing the channel estimation at the AP where only the quantised version of the estimated channel are transferred to the EPU. Overall, the trade off between the SE and the computation complexity depends on the specific requirements of the system, including the factors such as fronthaul capacity, latency requirements, computational resources and the nature of the application being supported.

4.8 Appendix

Appendix 4.A: Proof of Theorem 4.3 ($n = 2$)

First, we prove the case when there are only two real product variables ($n = 2$). The characteristic function of the sum of Z_1 and Z_2 based on Lemma 4.1 is given by:

$$\begin{aligned}\varphi_Z(t) &= \mathbb{E}\{e^{itz}\} = \mathbb{E}\{e^{itz_1}e^{itz_2}\} \\ &= (1 + b_1t^2)^{-\frac{1}{2}}(1 + b_1t^2)^{-\frac{1}{2}} \\ &= (1 + b_1t^2)^{-1}\end{aligned} \tag{4.57}$$

By using the inverse theorem of characteristic function and solving the integral, the pdf (4.27) is then calculated. Here, we can solve the integral by exploiting the residue theorem. We first discuss the case $z < 0$ and the contour C goes from $-R$ to R where

$R > 1/\sqrt{b_1}$ on the real axis and then counterclockwise along a semicircle centred at origin from R to $-R$. The integral (4.27) can be expressed by:

$$\begin{aligned}
f_Z(z) &= \frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{e^{-itz}}{1+b_1t^2}dt = \lim_{R\to\infty}\frac{1}{2\pi}\int_{-R}^{R}\frac{e^{-itz}}{1+b_1t^2}dt \\
&= \lim_{R\to\infty}\frac{1}{2\pi}\left(\oint_C \frac{e^{-itz}}{1+b_1t^2}dt - \int_{arc}\frac{e^{-itz}}{1+b_1t^2}dt\right) \\
&= \lim_{R\to\infty}\frac{1}{2\pi}\left(\oint_C \frac{e^{-itz}}{b_1\left(t+i/\sqrt{b_1}\right)\left(t-i/\sqrt{b_1}\right)}dt - \int_{arc}\frac{e^{-itz}}{1+b_1t^2}dt\right) \\
&= \frac{1}{2\pi}\left(2\pi i\cdot Res_{t=\frac{i}{\sqrt{b_1}}}\frac{e^{-itz}}{1+b_1t^2} - \lim_{R\to\infty}\int_{arc}\frac{e^{-itz}}{1+b_1t^2}dt\right) \\
&\overset{(a)}{=} \frac{1}{2\pi}\left(2\pi i\cdot\frac{e^{\frac{z}{\sqrt{b_1}}}}{2\sqrt{b_1}i} - \lim_{R\to\infty}\int_{arc}\frac{e^{-itz}}{1+b_1t^2}dt\right) \\
&= \frac{e^{\frac{z}{\sqrt{b_1}}}}{2\sqrt{b_1}}
\end{aligned} \tag{4.58}$$

where arc represents the contour of the semicircle, Res denotes the residue and the integral in step (a) is solved by:

$$\begin{aligned}
&\lim_{R\to\infty}\left|\int_{arc}\frac{e^{-itz}}{1+b_1t^2}dt\right| \\
&\overset{(a)}{=} \lim_{R\to\infty}\left|\int_0^{\pi}\frac{e^{-izRe^{i\varphi}}}{1+b_1(Re^{i\varphi})^2}iRe^{i\varphi}d\varphi\right| \\
&\overset{(b)}{\leq} \lim_{R\to\infty}\pi R\cdot\sup_{\varphi\in[0,\pi]}\left|\frac{e^{-izRe^{i\varphi}}e^{i\varphi}}{1+b_1(Re^{i\varphi})^2}\right| \\
&= \lim_{R\to\infty}\pi R\cdot\sup_{\varphi\in[0,\pi]}\frac{\left|e^{-izR(\cos\varphi+i\cdot\sin\varphi)+i\varphi}\right|}{\left|1+b_1R^2(\cos 2\varphi+i\cdot\sin 2\varphi)\right|} \\
&\overset{(c)}{=} \lim_{R\to\infty}\pi R\cdot\sup_{\varphi\in[0,\pi]}\frac{\left|e^{i(\varphi-zR\cos\varphi)+zR\sin\varphi}\right|}{\sqrt{(1+b_1R^2\cos 2\varphi)^2+b_1^2R^4\sin^2 2\varphi}} \\
&= \lim_{R\to\infty}\pi R\cdot\sup_{\varphi\in[0,\pi]}\frac{\left|e^{zR\sin\varphi}\right|}{\sqrt{b_1^2R^4+2b_1R^2\cos 2\varphi+1}}
\end{aligned} \tag{4.59}$$

$$\overset{(d)}{\leq} \lim_{R\to\infty} \pi R \cdot \sup_{\varphi\in[0,\pi]} \frac{1}{\sqrt{b_1^2R^4 + 2b_1R^2\cos 2\varphi + 1}}$$

$$\leq \lim_{R\to\infty} \pi R \cdot \frac{1}{\sqrt{b_1^2R^4 - 2b_1R^2 + 1}} = \lim_{R\to\infty} \pi R \cdot \frac{1}{b_1R^2 - 1} = 0$$

where step (a) expresses t in polar form: $t = Re^{i\varphi}$, and the absolute value of i is equal to one. The argument sup denotes the supremum in step (b). The inequality is satisfied because of the property of the integral by which we can multiply the maximum value of the integrand by the length of the integral trajectory. Step (c) includes the term $\left|e^{i(\varphi - zR\cos\varphi)}\right|$ of which the absolute value is one. Since we assume $z < 0$ and $R\sin\varphi \geq 0$ when $\varphi \in [0,\pi]$, $\left|e^{zR\sin\varphi}\right| \leq 1$ achieves step (d). Finally, the term $2b_1R^2\cos 2\varphi$ has the minimum value $-2b_1R^2$ which gives the result of the supremum, and (4.58) is proved.

Next, we research the case $z > 0$ and the contour C^- goes from R to $-R$ where $R > 1/\sqrt{b_1}$ on the real axis and then counterclockwise along a semicircle centred at origin from $-R$ to R. Then, we have a similar calculation as (4.58):

$$\begin{aligned}
f_Z(z) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{e^{-itz}}{1 + b_1t^2}\,dt \\
&= \lim_{R\to\infty} \frac{1}{2\pi}\int_{-R}^{R} \frac{e^{-itz}}{1 + b_1t^2}\,dt \\
&= \lim_{R\to\infty} -\frac{1}{2\pi}\left(\oint_{C^-} \frac{e^{-itz}}{1 + b_1t^2}\,dt - \int_{arc^-} \frac{e^{-itz}}{1 + b_1t^2}\,dt\right) \\
&= -\frac{1}{2\pi}\left(2\pi i \cdot Res_{t=-\frac{i}{\sqrt{b_1}}} \frac{e^{-itz}}{1 + b_1t^2} - \lim_{R\to\infty}\int_{arc^-} \frac{e^{-itz}}{1 + b_1t^2}\,dt\right) \\
&= -\frac{1}{2\pi}\cdot 2\pi i \cdot \left(-\frac{e^{-\frac{z}{\sqrt{b_1}}}}{2\sqrt{b_1}i}\right) \\
&= \frac{e^{-\frac{z}{\sqrt{b_1}}}}{2\sqrt{b_1}}
\end{aligned} \tag{4.60}$$

According to the integral contour, we have the relationship $\oint_{C^-} dt = \int_{arc^-} dt + \int_R^{-R} dt$, where arc^- represents the semicircle from $-R$ to R with the angle π to 2π. The integral on the arc^- is estimated:

$$\lim_{R\to\infty}\left|\int_{arc^-}\frac{e^{-itz}}{1+b_1t^2}dt\right| = \lim_{R\to\infty}\left|\int_{\pi}^{2\pi}\frac{e^{-izRe^{i\varphi}}}{1+b_1R^2e^{2i\varphi}}iRe^{i\varphi}d\varphi\right| \tag{4.61}$$

$$= \lim_{R\to\infty}\pi R\cdot\sup_{\varphi\in[\pi,2\pi]}\left|\frac{e^{zR\sin\varphi}}{1+b_1R^2e^{2i\varphi}}\right| \le \lim_{R\to\infty}\pi R\cdot\sup_{\varphi\in[\pi,2\pi]}\left|\frac{1}{1+b_1R^2e^{2i\varphi}}\right|$$

$$\le \lim_{R\to\infty}\pi R\cdot\frac{1}{b_1R^2-1} = 0$$

where $z > 0$ and $R\sin\varphi \le 0$ when $\varphi \in [\pi, 2\pi]$, hence $e^{zR\sin\varphi} \le 1$. If $z = 0$, we can easily obtain the result of the integral as $\pi/\sqrt{b_1}$ and thus

$$f_Z(z) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{e^{-itz}}{1+b_1t^2}dt = \frac{e^{-\frac{|z|}{\sqrt{b_1}}}}{2\sqrt{b_1}} \tag{4.62}$$

We combine the results of (4.58), (4.60) and (4.62), which completes the proof.

Appendix 4.B: Proof of Theorem 4.3 ($n > 2$)

The characteristic function of the sum of the set $Z = \sum_{k=1}^{n} Z_k$ is given by Lemma 4.1:

$$\varphi_Z(t) = \mathbb{E}\{e^{itz}\} \tag{4.63}$$

$$= \mathbb{E}\{e^{it\sum_{k=1}^{n}z_k}\} = \mathbb{E}\left\{\prod_{k=1}^{n}e^{itz_k}\right\}$$

$$= \frac{1}{\prod_{m=1}^{n/2}(1+b_mt^2)}$$

Using the inverse theorem of the characteristic function:

$$f_Z(z) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{e^{-itz}}{\prod_{m=1}^{n/2}(1+b_mt^2)}dt \tag{4.64}$$

To solve this integral, we adopt the residue theorem that has the same steps as (4.58). First, we present the case $z < 0$ and the contour C goes from $-R$ to R where $R > \max 1/\sqrt{b_m}$ on the real axis and then counterclockwise along a semicircle centred at origin from R to $-R$. The integral (4.64) can be expressed by:

$$\begin{aligned} f_Z(z) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{e^{-itz}}{\prod_{m=1}^{\frac{n}{2}}(1+b_m t^2)}\,dt \\ &= \lim_{R\to\infty}\frac{1}{2\pi}\int_{-R}^{R} \frac{e^{-itz}}{\prod_{m=1}^{\frac{n}{2}}(1+b_m t^2)}\,dt \\ &= \lim_{R\to\infty}\frac{1}{2\pi}\left(\oint_C \frac{e^{-itz}}{\prod_{m=1}^{\frac{n}{2}}(1+b_m t^2)}\,dt - \int_{arc} \frac{e^{-itz}}{\prod_{m=1}^{\frac{n}{2}}(1+b_m t^2)}\,dt\right) \\ &= \frac{1}{2\pi}\left(2\pi i \cdot \sum_{m=1}^{\frac{n}{2}} Res_{t=\frac{i}{\sqrt{b_m}}} \frac{e^{-itz}}{\prod_{m=1}^{\frac{n}{2}}(1+b_m t^2)}\right) \\ &= i \cdot \sum_{m=1}^{\frac{n}{2}} \frac{e^{\frac{z}{\sqrt{b_m}}}}{b_m\left(i/\sqrt{b_m}+i/\sqrt{b_m}\right)\cdot\prod_{m'\neq m}^{\frac{n}{2}}\left(1-\frac{b_{m'}}{b_m}\right)} \\ &= \frac{1}{2}\sum_{m=1}^{n/2} \frac{b_m^{\frac{n-3}{2}} e^{\frac{z}{\sqrt{b_m}}}}{\prod_{m'\neq m}^{n/2}(b_m - b_{m'})}\,, b_m \neq b_{m'} \end{aligned} \tag{4.65}$$

where arc represents the semicircle from $-R$ to R with the angle from 0 to π and the integral on arc has the result zero when R tends to infinity. This can be proved in the same approach as (4.59). Similarly, if $z > 0$, we obtain the result as:

$$\begin{aligned} f_Z(z) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{e^{-itz}}{\prod_{m=1}^{\frac{n}{2}}(1+b_m t^2)}\,dt \\ &= \frac{1}{2}\sum_{m=1}^{n/2} \frac{b_m^{\frac{n-3}{2}} e^{-\frac{z}{\sqrt{b_m}}}}{\prod_{m'\neq m}^{n/2}(b_m - b_{m'})}\,, b_m \neq b_{m'} \end{aligned} \tag{4.66}$$

and when $z = 0$

$$
\begin{aligned}
f_Z(0) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{1}{\prod_{m=1}^{\frac{n}{2}}(1+b_m t^2)}\,dt \\
&= \frac{1}{2}\sum_{m=1}^{n/2} \frac{b_m^{\frac{n-3}{2}}}{\prod_{m'\neq m}^{n/2}(b_m - b_{m'})}\,, b_m \neq b_{m'}
\end{aligned}
\tag{4.67}
$$

Thus, we combine the results (4.65), (4.66) and (4.67) which completes the proof.

Appendix 4.C: Proof of Theorem 4.4 ($n > 2$)

Introducing the Lemma 4.1 which gives the pdf of z, the pdf of $y = z + x$ can be expressed by:

$$
\begin{aligned}
f_Y(y) &= \int_{-\infty}^{\infty} f_X(x)\, f_Z(y-x)dx \\
&= \int_{-\infty}^{\infty} \frac{1}{\sigma_x\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma_x^2}} \frac{1}{2}\sum_{m=1}^{\frac{n}{2}} \frac{b_m^{\frac{n-3}{2}} e^{-\frac{|y-x|}{\sqrt{b_m}}}}{\prod_{m'\neq m}^{\frac{n}{2}}(b_m - b_{m'})}\,dx \\
&= \frac{1}{2\sigma_x\sqrt{2\pi}} \sum_{m=1}^{\frac{n}{2}} \frac{b_m^{\frac{n-3}{2}}}{\prod_{m'\neq m}^{\frac{n}{2}}(b_m - b_{m'})} \int_{-\infty}^{\infty} e^{-\frac{x^2}{2\sigma_x^2}} e^{-\frac{|y-x|}{\sqrt{b_m}}}\,dx \\
&= \frac{1}{2\sigma_x\sqrt{2\pi}} \sum_{m=1}^{\frac{n}{2}} \frac{b_m^{\frac{n-3}{2}}}{\prod_{m'\neq m}^{\frac{n}{2}}(b_m - b_{m'})} I_m
\end{aligned}
\tag{4.68}
$$

where

$$
\begin{aligned}
I_m &= \int_{-\infty}^{\infty} e^{-\frac{x^2}{2\sigma_x^2}} e^{-\frac{|y-x|}{\sqrt{b_m}}}\,dx \\
&= \int_{-\infty}^{y} e^{-\frac{x^2}{2\sigma_x^2} - \frac{y-x}{\sqrt{b_m}}}\,dx + \int_{y}^{\infty} e^{-\frac{x^2}{2\sigma_x^2} + \frac{y-x}{\sqrt{b_m}}}\,dx \\
&= \int_{-\infty}^{y} e^{-\frac{x^2 - \frac{2\sigma_x^2}{\sqrt{b_m}}x + \frac{2\sigma_x^2}{\sqrt{b_m}}y}{2\sigma_x^2}}\,dx + \int_{y}^{\infty} e^{-\frac{x^2 + \frac{2\sigma_x^2}{\sqrt{b_m}}x - \frac{2\sigma_x^2}{\sqrt{b_m}}y}{2\sigma_x^2}}\,dx
\end{aligned}
\tag{4.69}
$$

$$= \int_{-\infty}^{y} e^{-\frac{\left(x-\frac{\sigma_x^2}{\sqrt{b_m}}\right)^2}{2\sigma_x^2}} e^{-\frac{y}{\sqrt{b_m}}+\frac{\sigma_x^2}{2b_m}} dx + \int_{y}^{\infty} e^{-\frac{\left(x+\frac{\sigma_x^2}{\sqrt{b_m}}\right)^2}{2\sigma_x^2}} e^{\frac{y}{\sqrt{b_m}}+\frac{\sigma_x^2}{2b_m}} dx$$

$$= e^{-\frac{y}{\sqrt{b_m}}+\frac{\sigma_x^2}{2b_m}} \sqrt{\frac{\pi\sigma_x^2}{2}} \operatorname{erfc}\left(\frac{-\sqrt{b_m}y+\sigma_x^2}{\sqrt{2b_m\sigma_x^2}}\right) + e^{\frac{y}{\sqrt{b_m}}+\frac{\sigma_x^2}{2b_m}} \sqrt{\frac{\pi\sigma_x^2}{2}} \operatorname{erfc}\left(\frac{\sqrt{b_m}y+\sigma_x^2}{\sqrt{2b_m\sigma_x^2}}\right)$$

Then, we substitute (4.69) into (4.68):

$$f_Y(y) = \frac{1}{4}\sum_{m=1}^{\frac{n}{2}} \frac{b_m^{\frac{n-3}{2}}}{\prod_{m'\neq m}^{\frac{n}{2}}(b_m - b_{m'})} \left(e^{-\frac{y}{\sqrt{b_m}}+\frac{\sigma_x^2}{2b_m}} \operatorname{erfc}\left(\frac{-\sqrt{b_m}y+\sigma_x^2}{\sqrt{2\sigma_x^2 b_m}}\right) + e^{\frac{y}{\sqrt{b_m}}+\frac{\sigma_x^2}{2b_m}} \operatorname{erfc}\left(\frac{\sqrt{b_m}y+\sigma_x^2}{\sqrt{2\sigma_x^2 b_m}}\right) \right) \tag{4.70}$$

which completes the proof.

Appendix 4.D: Proof of Lemma 4.2 and Lemma 4.3

$$\begin{aligned} &\int x e^{bx+d} \operatorname{erfc}(ax+c)\, dx \\ &= \frac{1}{b} e^{bx+d} x \operatorname{erfc}(ax+c) - \frac{1}{b}\int e^{bx+d}\, d\big(x \cdot \operatorname{erfc}(ax+c)\big) \\ &= \frac{1}{b} e^{bx+d} x \operatorname{erfc}(ax+c) - \frac{1}{b}\int e^{bx+d} \operatorname{erfc}(ax+c)\, dx \\ &\qquad - \frac{1}{b}\int e^{bx+d} x\, d\big(\operatorname{erfc}(ax+c)\big) \end{aligned} \tag{4.71}$$

where

$$\begin{aligned} &\int e^{bx+d} \operatorname{erfc}(ax+c)\, dx \\ &= \frac{1}{b} e^{bx+d} \operatorname{erfc}(ax+c) - \frac{1}{b}\int e^{bx+d}\, d\big(\operatorname{erfc}(ax+c)\big) \\ &= \frac{1}{b} e^{bx+d} \operatorname{erfc}(ax+c) - \frac{1}{b}\int e^{bx+d} \left(-\frac{2a}{\sqrt{\pi}}\right) e^{-(ax+c)^2} dx \\ &= \frac{1}{b} e^{bx+d} \operatorname{erfc}(ax+c) + \frac{2a}{b\sqrt{\pi}} \int e^{-a^2x^2-2acx+bx-c^2+d}\, dx \end{aligned} \tag{4.72}$$

$$
\begin{aligned}
&= \frac{1}{b} e^{bx+d} \operatorname{erfc}(ax+c) + \frac{2a}{b\sqrt{\pi}} \int e^{-\left(ax+c-\frac{b}{2a}\right)^2 - \frac{bc}{a} + \frac{b^2}{4a^2} + d} \, dx \\
&= \frac{1}{b} e^{bx+d} \operatorname{erfc}(ax+c) + \frac{2a}{b\sqrt{\pi}} e^{-\frac{bc}{a} + \frac{b^2}{4a^2} + d} \cdot \left(-\frac{\sqrt{\pi}}{2a}\right) \operatorname{erfc}\left(ax+c-\frac{b}{2a}\right) \\
&= \frac{1}{b} \left(e^{bx+d} \operatorname{erfc}(ax+c) - e^{\frac{b^2}{4a^2} - \frac{bc}{a} + d} \operatorname{erfc}\left(ax+c-\frac{b}{2a}\right) \right)
\end{aligned}
$$

which completes the proof of Lemma 4.3.

$$
\begin{aligned}
&\int e^{bx+d} x \, d\big(\operatorname{erfc}(ax+c)\big) \\
&= -\frac{2a}{\sqrt{\pi}} \int e^{bx+d} x e^{-(ax+c)^2} \, dx \\
&= -\frac{2a}{\sqrt{\pi}} \left(-\frac{1}{2a^2} e^{-a^2x^2} e^{bx+d-2acx-c^2} + \frac{1}{2a^2} \int e^{-a^2x^2} \, de^{bx+d-2acx-c^2} \right) \\
&= \frac{1}{a\sqrt{\pi}} \left(e^{-(ax+c)^2+bx-d} - \int e^{-a^2x^2} (b-2ac) \, e^{bx+d-2acx-c^2} dx \right) \\
&= \frac{1}{a\sqrt{\pi}} \left(e^{-(ax+c)^2+bx-d} - (b-2ac) \int e^{-\left(ax+c-\frac{b}{2a}\right)^2 - \frac{bc}{a} + \frac{b^2}{4a^2} + d} \, dx \right) \\
&= \frac{1}{a\sqrt{\pi}} \left(e^{-(ax+c)^2+bx-d} + \frac{(b-2ac)\sqrt{\pi}}{2a} e^{\frac{b^2}{4a^2} - \frac{bc}{a} + d} \operatorname{erfc}\left(ax+c-\frac{b}{2a}\right) \right) \\
&= \frac{1}{a\sqrt{\pi}} e^{-(ax+c)^2+bx-d} + \frac{(b-2ac)}{2a^2} e^{\frac{b^2}{4a^2} - \frac{bc}{a} + d} \operatorname{erfc}\left(ax+c-\frac{b}{2a}\right)
\end{aligned}
\tag{4.73}
$$

Then, we substitute (4.72) and (4.73) into (4.71):

$$
\begin{aligned}
&\int x e^{bx+d} \operatorname{erfc}(ax+c) \, dx \\
&= \frac{1}{b} e^{bx+d} x \operatorname{erfc}(ax+c) - \frac{1}{b} \int e^{bx+d} \operatorname{erfc}(ax+c) \, dx \\
&\qquad\qquad - \frac{1}{b} \int e^{bx+d} x \, d\big(\operatorname{erfc}(ax+c)\big) \\
&= \frac{1}{b} e^{bx+d} x \operatorname{erfc}(ax+c) \\
&\qquad\qquad - \frac{1}{b^2} \left(e^{bx+d} \operatorname{erfc}(ax+c) - e^{\frac{b^2}{4a^2} - \frac{bc}{a} + d} \operatorname{erfc}\left(ax+c-\frac{b}{2a}\right) \right)
\end{aligned}
\tag{4.74}
$$

$$-\frac{1}{ab\sqrt{\pi}}e^{-(ax+c)^2+bx+d}-\frac{(b-2ac)}{2a^2b}e^{\frac{b^2}{4a^2}-\frac{bc}{a}+d}\operatorname{erfc}\left(ax+c-\frac{b}{2a}\right)$$
$$=\left(\frac{1}{b^2}-\frac{1}{2a^2}+\frac{c}{ab}\right)e^{\frac{b^2}{4a^2}-\frac{bc}{a}+d}\operatorname{erfc}\left(ax+c-\frac{b}{2a}\right)+\frac{bx-1}{b^2}e^{bx+d}\operatorname{erfc}(ax+c)$$
$$-\frac{1}{ab\sqrt{\pi}}e^{-(ax+c)^2+bx+d}$$

which completes the proof of Lemma 4.2.

Appendix 4.E: Proof of Theorem 4.5

We first prove the case when $n = 2$:

$$\begin{aligned}
\alpha &= \frac{1}{P_x}\int_{-\infty}^{\infty} xy(x)f_X(x)dx \\
&= \frac{1}{\sigma_x^2}\int_{-\infty}^{\left(-\frac{L}{2}+1\right)\Delta} x\left(-\frac{L-1}{2}\right)\Delta\frac{1}{4\sqrt{b_1}}\left(e^{-\frac{x}{\sqrt{b_1}}+\frac{\sigma_n^2}{2b_1}}\operatorname{erfc}\left(\frac{-\sqrt{b_1}x+\sigma_n^2}{\sqrt{2\sigma_n^2 b_1}}\right)\right. \\
&\qquad \left. + e^{\frac{x}{\sqrt{b_1}}+\frac{\sigma_n^2}{2b_1}}\operatorname{erfc}\left(\frac{\sqrt{b_1}x+\sigma_n^2}{\sqrt{2\sigma_n^2 b_1}}\right)\right)dx \\
&\quad + \frac{1}{\sigma_x^2}\sum_{l=-\frac{L}{2}+1}^{\frac{L}{2}-2}\int_{l\Delta}^{(l+1)\Delta} x\left(l+\frac{1}{2}\right)\Delta\frac{1}{4\sqrt{b_1}}\left(e^{-\frac{x}{\sqrt{b_1}}+\frac{\sigma_n^2}{2b_1}}\operatorname{erfc}\left(\frac{-\sqrt{b_1}x+\sigma_n^2}{\sqrt{2\sigma_n^2 b_1}}\right)\right. \\
&\qquad \left. + e^{\frac{x}{\sqrt{b_1}}+\frac{\sigma_n^2}{2b_1}}\operatorname{erfc}\left(\frac{\sqrt{b_1}x+\sigma_n^2}{\sqrt{2\sigma_n^2 b_1}}\right)\right)dx \\
&\quad + \frac{1}{\sigma_x^2}\int_{\left(\frac{L}{2}-1\right)\Delta}^{\infty} x\frac{L-1}{2}\Delta\frac{1}{4\sqrt{b_1}}\left(e^{-\frac{x}{\sqrt{b_1}}+\frac{\sigma_n^2}{2b_1}}\operatorname{erfc}\left(\frac{-\sqrt{b_1}x+\sigma_n^2}{\sqrt{2\sigma_n^2 b_1}}\right)\right. \\
&\qquad \left. + e^{\frac{x}{\sqrt{b_1}}+\frac{\sigma_n^2}{2b_1}}\operatorname{erfc}\left(\frac{\sqrt{b_1}x+\sigma_n^2}{\sqrt{2\sigma_n^2 b_1}}\right)\right)dx
\end{aligned} \tag{4.75}$$

$$= \sum_{l=-\frac{L}{2}+1}^{\frac{L}{2}-2} \frac{\left(l+\frac{1}{2}\right)\Delta}{4\sigma_x^2\sqrt{b_1}} \Bigg[\left(\sqrt{b_1}(l+1)\Delta - b_1\right) e^{\frac{l+1}{\sqrt{b_1}}\Delta + \frac{\sigma_n^2}{2b_1}} \operatorname{erfc}\left(\frac{l+1}{\sqrt{2}\sigma_n}\Delta + \frac{\sigma_n}{\sqrt{2b_1}}\right)$$

$$- \left(\sqrt{b_1}(l+1)\Delta + b_1\right) e^{-\frac{l+1}{\sqrt{b_1}}\Delta + \frac{\sigma_n^2}{2b_1}} \operatorname{erfc}\left(-\frac{l+1}{\sqrt{2}\sigma_n}\Delta + \frac{\sigma_n}{\sqrt{2b_1}}\right)$$

$$- 2\sigma_n \sqrt{\frac{2b_1}{\pi}} e^{-\frac{(l+1)^2\Delta^2}{2\sigma_n^2}} \Bigg]$$

$$+ \sum_{l=-\frac{L}{2}+1}^{\frac{L}{2}-2} \frac{\left(l+\frac{1}{2}\right)\Delta}{4\sigma_x^2\sqrt{b_1}} \Bigg[\left(\sqrt{b_1}l\Delta + b_1\right) e^{-\frac{l\Delta}{\sqrt{b_1}} + \frac{\sigma_n^2}{2b_1}} \operatorname{erfc}\left(-\frac{l\Delta}{\sqrt{2}\sigma_n} + \frac{\sigma_n}{\sqrt{2b_1}}\right)$$

$$- \left(\sqrt{b_1}l\Delta - b_1\right) e^{\frac{l}{\sqrt{b_1}}\Delta + \frac{\sigma_n^2}{2b_1}} \operatorname{erfc}\left(\frac{l\Delta}{\sqrt{2}\sigma_n} + \frac{\sigma_n}{\sqrt{2b_1}}\right) + 2\sigma_n \sqrt{\frac{2b_1}{\pi}} e^{-\frac{l^2\Delta^2}{2\sigma_n^2}} \Bigg]$$

$$+ \frac{(L-1)\Delta}{4\sigma_x^2\sqrt{b_1}} \Bigg[\left(\frac{L-2}{2}\Delta\sqrt{b_1} + b_1\right) e^{-\frac{L-2}{2\sqrt{b_1}}\Delta + \frac{\sigma_n^2}{2b_1}} \operatorname{erfc}\left(-\frac{L-2}{2\sqrt{2}\sigma_n}\Delta + \frac{\sigma_n}{2b_1}\right)$$

$$- \left(\frac{L-2}{2}\Delta\sqrt{b_1} - b_1\right) e^{\frac{L-2}{2\sqrt{b_1}}\Delta + \frac{\sigma_n^2}{2b_1}} \operatorname{erfc}\left(\frac{L-2}{2\sqrt{2}\sigma_n}\Delta + \frac{\sigma_n}{2b_1}\right)$$

$$+ 2\sigma_n \sqrt{\frac{2b_1}{\pi}} e^{-\frac{(L-2)^2\Delta^2}{8\sigma_n^2}} \Bigg]$$

$$= \sum_{l'=-\frac{L}{2}+1}^{\frac{L}{2}-1} \frac{\left(l'-\frac{1}{2}\right)\Delta}{4\sigma_x^2\sqrt{b_1}} \Bigg[\left(\sqrt{b_1}l'\Delta - b_1\right) e^{\frac{l'}{\sqrt{b_1}}\Delta + \frac{\sigma_n^2}{2b_1}} \operatorname{erfc}\left(\frac{l'}{\sqrt{2}\sigma_n}\Delta + \frac{\sigma_n}{\sqrt{2b_1}}\right)$$

$$- \left(\sqrt{b_1}l'\Delta + b_1\right) e^{-\frac{l'}{\sqrt{b_1}}\Delta + \frac{\sigma_n^2}{2b_1}} \operatorname{erfc}\left(-\frac{l'}{\sqrt{2}\sigma_n}\Delta + \frac{\sigma_n}{\sqrt{2b_1}}\right)$$

$$- 2\sigma_n \sqrt{\frac{2b_1}{\pi}} e^{-\frac{l'^2\Delta^2}{2\sigma_n^2}} \Bigg]$$

$$+\sum_{l=-\frac{L}{2}+1}^{\frac{L}{2}-1}\frac{\left(l+\frac{1}{2}\right)\Delta}{4\sigma_x^2\sqrt{b_1}}\left[\left(\sqrt{b_1}l\Delta+b_1\right)e^{-\frac{l\Delta}{\sqrt{b_1}}+\frac{\sigma_n^2}{2b_1}}\operatorname{erfc}\left(-\frac{l\Delta}{\sqrt{2}\sigma_n}+\frac{\sigma_n}{\sqrt{2b_1}}\right)\right.$$

$$\left.-\left(\sqrt{b_1}l\Delta-b_1\right)e^{\frac{l}{\sqrt{b_1}}\Delta+\frac{\sigma_n^2}{2b_1}}\operatorname{erfc}\left(\frac{l\Delta}{\sqrt{2}\sigma_n}+\frac{\sigma_n}{\sqrt{2b_1}}\right)+2\sigma_n\sqrt{\frac{2b_1}{\pi}}e^{-\frac{l^2\Delta^2}{2\sigma_n^2}}\right]$$

$$=\sum_{l=-\frac{L}{2}+1}^{\frac{L}{2}-1}\frac{\Delta}{4\sigma_x^2\sqrt{b_1}}\left[\left(\sqrt{b_1}l\Delta+b_1\right)e^{-\frac{l\Delta}{\sqrt{b_1}}+\frac{\sigma_n^2}{2b_1}}\operatorname{erfc}\left(-\frac{l\Delta}{\sqrt{2}\sigma_n}+\frac{\sigma_n}{\sqrt{2b_1}}\right)\right.$$

$$\left.-\left(\sqrt{b_1}l\Delta-b_1\right)e^{\frac{l\Delta}{\sqrt{b_1}}+\frac{\sigma_n^2}{2b_1}}\operatorname{erfc}\left(\frac{l\Delta}{\sqrt{2}\sigma_n}+\frac{\sigma_n}{\sqrt{2b_1}}\right)+2\sigma_n\sqrt{\frac{2b_1}{\pi}}e^{-\frac{l^2\Delta^2}{2\sigma_n^2}}\right]$$

$$=\frac{\Delta}{2\sigma_x^2}\sum_{l=1}^{\frac{L}{2}-1}\left[\left(l\Delta+\sqrt{b_1}\right)e^{-\frac{l\Delta}{\sqrt{b_1}}+\frac{\sigma_n^2}{2b_1}}\operatorname{erfc}\left(-\frac{l\Delta}{\sqrt{2}\sigma_n}+\frac{\sigma_n}{\sqrt{2b_1}}\right)\right.$$

$$\left.-\left(l\Delta-\sqrt{b_1}\right)e^{\frac{l}{\sqrt{b_1}}\Delta+\frac{\sigma_n^2}{2b_1}}\operatorname{erfc}\left(\frac{l\Delta}{\sqrt{2}\sigma_n}+\frac{\sigma_n}{\sqrt{2b_1}}\right)+2\sigma_n\sqrt{\frac{2}{\pi}}e^{-\frac{l^2\Delta^2}{2\sigma_n^2}}\right]$$

$$+\frac{\Delta}{2\sigma_x^2}\left(\sqrt{b_1}e^{\frac{\sigma_n^2}{2b_1}}\operatorname{erfc}\left(\frac{\sigma_n}{\sqrt{2b_1}}\right)+\sigma_n\sqrt{\frac{2}{\pi}}\right)$$

where $l'=l+1$. Then, we calculate γ:

$$\begin{aligned}\gamma&=\frac{1}{P_x}\int_{-\infty}^{\infty}|y(x)|^2f_X(x)dx\\&=\frac{1}{\sigma_x^2}\int_{-\infty}^{-\left(\frac{L}{2}-1\right)\Delta}\left(\frac{L-1}{2}\right)^2\Delta^2\frac{1}{4\sqrt{b_1}}\left(e^{-\frac{x}{\sqrt{b_1}}+\frac{\sigma_n^2}{2b_1}}\operatorname{erfc}\left(\frac{-\sqrt{b_1}x+\sigma_n^2}{\sqrt{2\sigma_n^2b_1}}\right)\right.\\&\qquad\left.+e^{\frac{x}{\sqrt{b_1}}+\frac{\sigma_n^2}{2b_1}}\operatorname{erfc}\left(\frac{\sqrt{b_1}x+\sigma_n^2}{\sqrt{2\sigma_n^2b_1}}\right)\right)dx\end{aligned}\tag{4.76}$$

$$+\frac{1}{\sigma_x^2}\sum_{l=-\frac{L}{2}+1}^{\frac{L}{2}-2}\int_{l\Delta}^{(l+1)\Delta}\left(l+\frac{1}{2}\right)^2\Delta^2\frac{1}{4\sqrt{b_1}}\left(e^{-\frac{x}{\sqrt{b_1}}+\frac{\sigma_n^2}{2b_1}}\operatorname{erfc}\left(\frac{-\sqrt{b_1}x+\sigma_n^2}{\sqrt{2\sigma_n^2 b_1}}\right)\right.$$

$$\left.+e^{\frac{x}{\sqrt{b_1}}+\frac{\sigma_n^2}{2b_1}}\operatorname{erfc}\left(\frac{\sqrt{b_1}x+\sigma_n^2}{\sqrt{2\sigma_n^2 b_1}}\right)\right)dx$$

$$+\frac{1}{\sigma_x^2}\int_{\left(\frac{L}{2}-1\right)\Delta}^{\infty}\left(\frac{L-1}{2}\right)^2\Delta^2\frac{1}{4\sqrt{b_1}}\left(e^{-\frac{x}{\sqrt{b_1}}+\frac{\sigma_n^2}{2b_1}}\operatorname{erfc}\left(\frac{-\sqrt{b_1}x+\sigma_n^2}{\sqrt{2\sigma_n^2 b_1}}\right)\right.$$

$$\left.+e^{\frac{x}{\sqrt{b_1}}+\frac{\sigma_n^2}{2b_1}}\operatorname{erfc}\left(\frac{\sqrt{b_1}x+\sigma_n^2}{\sqrt{2\sigma_n^2 b_1}}\right)\right)dx$$

$$=\frac{\Delta^2}{4\sigma_x^2}\sum_{l'=-\frac{L}{2}+1}^{\frac{L}{2}-1}\left(l'-\frac{1}{2}\right)^2\left[e^{\frac{l'\Delta}{\sqrt{b_1}}+\frac{\sigma_n^2}{2b_1}}\operatorname{erfc}\left(\frac{l'\Delta}{\sqrt{2}\sigma_n}+\frac{\sigma_n}{\sqrt{2b_1}}\right)-\operatorname{erfc}\left(\frac{l'\Delta}{\sqrt{2}\sigma_n}\right)\right.$$

$$\left.-e^{-\frac{l'\Delta}{\sqrt{b_1}}+\frac{\sigma_n^2}{2b_1}}\operatorname{erfc}\left(-\frac{l'\Delta}{\sqrt{2}\sigma_n}+\frac{\sigma_n}{\sqrt{2b_1}}\right)+\operatorname{erfc}\left(-\frac{l'\Delta}{\sqrt{2}\sigma_n}\right)\right]$$

$$-\frac{\Delta^2}{4\sigma_x^2}\sum_{l=-\frac{L}{2}+1}^{\frac{L}{2}-1}\left(l+\frac{1}{2}\right)^2\left[e^{\frac{l\Delta}{\sqrt{b_1}}+\frac{\sigma_n^2}{2b_1}}\operatorname{erfc}\left(\frac{l\Delta}{\sqrt{2}\sigma_n}+\frac{\sigma_n}{\sqrt{2b_1}}\right)-\operatorname{erfc}\left(\frac{l\Delta}{\sqrt{2}\sigma_n}\right)\right.$$

$$\left.-e^{-\frac{l\Delta}{\sqrt{b_1}}+\frac{\sigma_n^2}{2b_1}}\operatorname{erfc}\left(-\frac{l\Delta}{\sqrt{2}\sigma_n}+\frac{\sigma_n}{\sqrt{2b_1}}\right)+\operatorname{erfc}\left(-\frac{l\Delta}{\sqrt{2}\sigma_n}\right)\right]+\frac{\Delta^2}{\sigma_x^2}\left(\frac{L-1}{2}\right)^2$$

$$=\frac{\Delta^2}{4\sigma_x^2}\sum_{l=-\frac{L}{2}+1}^{\frac{L}{2}-1}2l\left[e^{-\frac{l\Delta}{\sqrt{b_1}}+\frac{\sigma_n^2}{2b_1}}\operatorname{erfc}\left(-\frac{l\Delta}{\sqrt{2}\sigma_n}+\frac{\sigma_n}{\sqrt{2b_1}}\right)-\operatorname{erfc}\left(-\frac{l\Delta}{\sqrt{2}\sigma_n}\right)\right.$$

$$\left.-e^{\frac{l\Delta}{\sqrt{b_1}}+\frac{\sigma_n^2}{2b_1}}\operatorname{erfc}\left(\frac{l\Delta}{\sqrt{2}\sigma_n}+\frac{\sigma_n}{\sqrt{2b_1}}\right)+\operatorname{erfc}\left(\frac{l\Delta}{\sqrt{2}\sigma_n}\right)\right]+\frac{\Delta^2}{\sigma_x^2}\left(\frac{L-1}{2}\right)^2$$

$$=\frac{\Delta^2}{\sigma_x^2}\sum_{l=1}^{\frac{L}{2}-1}l\left[e^{-\frac{l\Delta}{\sqrt{b_1}}+\frac{\sigma_n^2}{2b_1}}\operatorname{erfc}\left(-\frac{l\Delta}{\sqrt{2}\sigma_n}+\frac{\sigma_n}{\sqrt{2b_1}}\right)-\operatorname{erfc}\left(-\frac{l\Delta}{\sqrt{2}\sigma_n}\right)\right.$$

$$\left.-e^{\frac{l\Delta}{\sqrt{b_1}}+\frac{\sigma_n^2}{2b_1}}\operatorname{erfc}\left(\frac{l\Delta}{\sqrt{2}\sigma_n}+\frac{\sigma_n}{\sqrt{2b_1}}\right)+\operatorname{erfc}\left(\frac{l\Delta}{\sqrt{2}\sigma_n}\right)\right]+\frac{\Delta^2}{\sigma_x^2}\left(\frac{L-1}{2}\right)^2$$

where $l'=l+1$. Since the pdf when $n>2$ has the similar form as the case $n=2$ in

Theorem 4.4, the scenario of $n > 2$ in Theorem 4.5 can be proved in the same method as (4.75) and (4.76) by introducing Lemma 4.2 and Lemma 4.3. Then, the proof is completed.

Chapter 5

Local Estimation in Decentralised Distributed Massive MIMO

5.1 Introduction

Considering where the signal processing is performed has become a notable and significant topic of discussion within the context of the physical layer of radio access network (RAN). To fulfil the demand of sixth generation (6G) and the future wireless network, the location of signal processing needs to be investigated thoroughly. As we have stated in Chapter 3, all channel state information (CSI) and data are sent to the edge processing unit (EPU) for signal processing in decentralised distributed massive multiple-input multiple-output (DD-MaMIMO) which achieves a very high spectral efficiency (SE). However, the computational complexity for the EPU may dramatically increase when the number of served user terminals (UTs) grows. Furthermore, the performance is also influenced by the capacity of fronthaul which has been discussed in Chapter 4. In order to reduce the computational burden of EPU, we can move the signal processing to the access point (AP) in which the channel and data can be locally estimated and then transmitted to the EPU for the final decoding. Note that people often use either the term “local” or “decentralised” to indicate the signal processing which is happened at the AP in cell-free MaMIMO. We hence define this scheme as local estimation in DD-MaMIMO to distinguish it from the definition of “decentralised” given in Chapter 3.

Related work on local estimation has been widely studied for uplink transmission in cell-free MaMIMO. In [14, 21], minimum mean square error (MMSE) channel estimation and conjugate beamforming, also known as maximum ratio combining (MRC) or matched filter, were applied at the APs. Papers [76, 77] researched the use of full-pilot zero-forcing (FZF) and partial full-pilot zero-forcing (PFZF) combining techniques at the AP, which presented a SE improvement compared to MRC. In [24], an MMSE estimator was adopted to estimate the channel and data at the APs with

different decoding methods: large scale fading decoding and directly averaging the estimated data sent by all APs. For both decoding schemes, the results in [24] showed that the use of local MMSE estimation outperformed the MRC technique. Additionally, for the local estimation with large scale fading decoding in cell-free MaMIMO, if the local number of antennas in the system remains constant, a configuration which contains fewer APs with multiple antennas can achieve a higher SE than the case with more APs equipped with single antenna. This is because using multiple antennas increases spatial diversity improving the data detection at each AP. In contrast, the scenario involving a large number of single-antenna APs performs better in DD-MaMIMO with decentralised processing at the EPU. Furthermore, the local estimation scheme also fits the architecture of Open RAN, where the physical layer function of the baseband unit (BBU) in the conventional cloud RAN (C-RAN) is split between radio unit (RU) and distributed unit (DU) according to different options [53, 54].

Therefore, in this chapter we will engage in a comprehensive discussion of local estimation in DD-MaMIMO, assuming perfect fronthaul links. Firstly, the system model and channel estimation will be introduced: these are similar to the concepts presented in Chapter 3, except for the location of channel estimation. Following that, we will explore local data detection, considering the estimation weights with different combining techniques. Given the estimated data acquired at the APs, two decoding schemes for data recovery will then be discussed. Furthermore, we will derive a general expression that can calculate SE in the local estimation system for arbitrary combinations. Finally, we will provide numerical results generated by the simulation. The results demonstrates that local estimation employing large scale fading decoding can achieve higher SE than when applying simple decoding in DD-MaMIMO. Besides, compared to the decentralised processing at the EPU, the local estimation perform a worse SE, especially for the single-antenna AP case. However, for a large number of antennas case, the discrepancy is not huge.

5.2 System Model

As stated in Chapter 3, we also assume in this chapter that there are K_{serv} UTs with a single antenna and M_{coor} APs equipped N_r antennas each in the DD-MaMIMO system. The received signal $x_{m,n}$ at the *n*-th antenna of the *m*-th AP on the uplink is given by:

$$x_{m,n} = \sum_{k=1}^{K_{serv}} g_{mnk} s_k + \sum_{k=1}^{K_{int}} g_{i,mnk} s_{i,k} + z_{m,n} = \sum_{k=1}^{K_{serv}} h_{mnk} \beta_{mk}^{1/2} s_k + \sum_{k=1}^{K_{int}} h_{i,mnk} \beta_{i,mk}^{1/2} s_{i,k} + z_{m,n} \quad (5.1)$$

where g_{mnk} and $g_{i,mnk}$ are the channel coefficients, s_k and $s_{i,k}$ denote the transmit signals sent by served UTs and interfering UTs, respectively, $z_{m,n}$ is the noise at the antenna, h_{mnk} and $h_{i,mnk}$ represent the Rayleigh fading coefficients and β_{mk} and $\beta_{i,mk}$ are large scale fading coefficients. All parameters were defined in Chapter 3. This scalar received signal (5.1) can also be written in the vector form which is given in (3.10).

5.3 Channel Estimation

For the channel estimation the pilot sequences are sent from K_{serv} UTs to each antenna of the APs. Then, exploiting the received signals the channel is estimated at the AP which constitutes the local estimation. We consider the same method as Chapter 3 for channel estimation where the received pilots are given:

$$\boldsymbol{y}_{p,mn} = \sum_{k=1}^{K_{serv}} g_{mnk} \boldsymbol{\varphi}_k^T + \sum_{k=1}^{K_{int}} g_{i,mnk} \boldsymbol{\varphi}_{i,k}^T + \boldsymbol{z}_{p,mn} \quad (5.2)$$

where $\boldsymbol{\varphi}_k \in \mathbb{C}^{\tau_p \times 1}$ is one pilot sequence transmitted by the *k*-th UT, $\boldsymbol{\varphi}_{i,k} \in \mathbb{C}^{\tau_p \times 1}$ is the pilot sequence sent by the interfering UT and $\boldsymbol{z}_{p,mn} \in \mathbb{C}^{1 \times \tau_p}$ represents the noise at the *n*-th antenna of the *m*-th AP, in which the elements of the vector follow $\mathcal{CN}(0, \sigma_z^2)$. Here, we also assume that the channel coefficients among the N_r antennas at the same AP are uncorrelated. Thus, we are able to apply least square (LS) estimation to the channel and obtain the estimated channel by weighting the LS outcomes. This is calculated by first projecting $\boldsymbol{y}_{p,mn}$ onto $\boldsymbol{\varphi}_k^*$ and scaling it by $1/\tau_p$, followed by weighting the corresponding result with a weight c_{mnk} that minimises the MSE. The channel estimate has the same form as (3.7):

$$\hat{g}_{mnk} = c_{mnk} \frac{1}{\tau_p} \boldsymbol{y}_{p,mn} \boldsymbol{\varphi}_k^* \quad (5.3)$$

$$= \left(\frac{\beta_{mnk}}{\beta_{mnk} + \frac{1}{\tau_p^2} \sum_{k'=1}^{K_{int}} \beta_{i,mnk'} \mathbb{E}\left\{ \left| \boldsymbol{\varphi}_{i,k'}^T \boldsymbol{\varphi}_k^* \right|^2 \right\} + \frac{\sigma_z^2}{\tau_p}} \right) \frac{1}{\tau_p} \boldsymbol{y}_{p,mn} \boldsymbol{\varphi}_k^*$$

where we assume that the variance of the term $\boldsymbol{\varphi}_{i,k'}^T \boldsymbol{\varphi}_k^*$ is τ_p and the MMSE weight c_{mnk} is defined in (3.6) where the large scale fading coefficient β_{mnk} is determined by the positions of the AP and UT. Hence, c_{mnk} is the same for all $n = 1, \dots, N_r$. Due to the local estimation, the received pilot sequences are not required at the EPU which reduces by $\tau_p N_r M_{coor}$ complex scalars the data to be sent from the APs to the EPU via the fronthaul in each coherence block. The MSE of the channel estimation is:

$$\begin{aligned} \sigma_{e_{mnk}}^2 &= \mathbb{E}\{|\tilde{g}_{mnk}|^2\} = \mathbb{E}\{|g_{mnk} - \hat{g}_{mnk}|^2\} \\ &= (1 - c_{mnk})\beta_{mnk} \\ &= (1 - c_{mk})\beta_{mk} \end{aligned} \tag{5.4}$$

The proof is the same as (3.25). Here, the channel estimation error $\tilde{g}_{mnk}$ is uncorrelated with the estimated channel $\hat{g}_{mnk}$ because of the use of MMSE.

5.4 Data Detection and Decoding

Instead of forwarding the channel estimates and data to the EPU, the known CSI for all served UTs at each AP can be utilised to locally estimate the data at the respective AP. After the initial data detection at the AP, the results are then sent to the EPU for the final decoding. In this sub-chapter, we mainly discuss the MR combining and the MMSE estimation for computing the local estimation of the data. First, the uplink received data at the *m*-th AP can be expressed by:

$$\boldsymbol{y}_{u,m} = \sum_{k=1}^{K_{serv}} \boldsymbol{g}_{mk} x_k + \sum_{k=1}^{K_{int}} \boldsymbol{g}_{i,mk} x_{i,k} + \boldsymbol{z}_m \tag{5.5}$$

where the channel coefficients for the K_{serv} UTs to the N_r antennas of the *m*-th AP $\boldsymbol{g}_{mk} = \left[g_{m1k}\ g_{m2k} \cdots g_{mN_rk} \right]^T$, the interfering channel coefficients are $\boldsymbol{g}_{i,mk} = \left[g_{i,m1k}\ g_{i,m2k} \cdots g_{i,mN_rk} \right]^T$, $\boldsymbol{z}_m \in \mathbb{C}^{N_r \times 1}$ denotes the receiver noise at the antennas with independent $\mathcal{CN}(0, \sigma_z^2)$ entities. Next, we assume that the local combining vector is

$\boldsymbol{v}_{mk} \in \mathbb{C}^{1\times N_r}$ for weighting the data at the *m*-th AP, transmitted from the *k*-th UT. Then, the local estimation of x_k at the *m*-th AP is given by:

$$\check{x}_{mk} = \boldsymbol{v}_{mk}\boldsymbol{y}_{u,m} = \sum_{k'=1}^{K_{serv}} \boldsymbol{v}_{mk}\boldsymbol{g}_{mk'}x_{k'} + \sum_{k'=1}^{K_{int}} \boldsymbol{v}_{mk}\boldsymbol{g}_{i,mk'}x_{i,k'} + \boldsymbol{v}_{mk}\boldsymbol{z}_m \tag{5.6}$$

Based on the different selection of the combining techniques, the combining vector can either be simple MRC with $\boldsymbol{v}_{mk} = \hat{\boldsymbol{g}}_{mk}^H$ or MMSE where the weight vector can be derived by computing the MSE of the estimated data which is a conditional expectation $\mathbb{E}\left\{\left|x_k - \boldsymbol{v}_{mk}\boldsymbol{y}_{u,m}\right|^2 \middle| \{\hat{\boldsymbol{g}}_{mk}\}\right\}$ and equating its first derivative with respect to $\boldsymbol{v}_{mk}$ to zero. The MSE ϵ_k is calculated by:

$$\begin{aligned}\epsilon_k &= \mathbb{E}\left\{\left|x_k - \boldsymbol{v}_{mk}\boldsymbol{y}_{u,m}\right|^2 \middle| \{\hat{\boldsymbol{g}}_{mk}\}\right\} \\ &= \mathbb{E}\{|x_k|^2\} - \mathbb{E}\{x_k^*\boldsymbol{v}_{mk}\boldsymbol{y}_{u,m}\} - \mathbb{E}\{x_k\boldsymbol{y}_{u,m}^H\boldsymbol{v}_{mk}^H\} + \boldsymbol{v}_{mk}\mathbb{E}\left\{\left|\boldsymbol{y}_{u,m}\right|^2\right\}\boldsymbol{v}_{mk}^H\end{aligned} \tag{5.7}$$

Then let the first derivative $\partial\epsilon/\partial\boldsymbol{v}_{mk} = 0$:

$$\begin{aligned}\frac{\partial\epsilon_k}{\partial\boldsymbol{v}_{mk}} &= -\mathbb{E}\{x_k^*\boldsymbol{y}_{u,m}\} + \left\{\left|\boldsymbol{y}_{u,m}\right|^2\right\}\boldsymbol{v}_{mk}^H \\ &= \left(\sigma_x^2\sum_{k'=1}^{K_{serv}}\left(\hat{\boldsymbol{g}}_{mk'}\hat{\boldsymbol{g}}_{mk'}^H + \boldsymbol{C}_{\tilde{g}\tilde{g},mk'}\right) + \sigma_x^2\sum_{k'=1}^{K_{int}}\boldsymbol{C}_{g_{int}g_{int},mk'}\right. \\ &\qquad \left. + \sigma_z^2\boldsymbol{I}_{N_r}\right)\boldsymbol{v}_{mk}^H - \sigma_x^2\hat{\boldsymbol{g}}_{mk} = 0\end{aligned} \tag{5.8}$$

Then, we can obtain the result:

$$\begin{aligned}\boldsymbol{v}_{mk}^H = &\left(\sigma_x^2\sum_{k'=1}^{K_{serv}}\left(\hat{\boldsymbol{g}}_{mk'}\hat{\boldsymbol{g}}_{mk'}^H + \boldsymbol{C}_{\tilde{g}\tilde{g},mk'}\right) + \sigma_x^2\sum_{k'=1}^{K_{int}}\boldsymbol{C}_{g_{int}g_{int},mk'}\right. \\ &\left. + \sigma_z^2\boldsymbol{I}_{N_r}\right)^{-1}\sigma_x^2\hat{\boldsymbol{g}}_{mk}\end{aligned} \tag{5.9}$$

Since the entities in the inverse matrix of (5.9) are Hermitian matrices, the Hermitian of $\boldsymbol{v}_{mk}^H$ is given by:

$$\boldsymbol{v}_{mk} = \widehat{\boldsymbol{g}}_{mk}^{H}\left(\sum_{k'=1}^{K_{serv}}\left(\widehat{\boldsymbol{g}}_{mk'}\widehat{\boldsymbol{g}}_{mk'}^{H} + \boldsymbol{C}_{\tilde{g}\tilde{g},mk'}\right) + \sum_{k'=1}^{K_{int}}\boldsymbol{C}_{g_{int}g_{int},mk'} + \frac{\sigma_z^2}{\sigma_x^2}\boldsymbol{I}_{N_r}\right)^{-1} \tag{5.10}$$

where the covariance matrix of the channel estimation error is:

$$\begin{aligned}\boldsymbol{C}_{\tilde{g}\tilde{g},mk} &= \mathbb{E}\{\tilde{\boldsymbol{g}}_{mk}\tilde{\boldsymbol{g}}_{mk}^{H}\} = diag\left(\sigma_{e_{m1k}}^2, \dots, \sigma_{e_{mN_rk}}^2\right)\\ &= (1 - c_{mk})\beta_{mk}\boldsymbol{I}_{N_r}\end{aligned} \tag{5.11}$$

and the covariance matrix of the interfering channel coefficients is given by:

$$\boldsymbol{C}_{g_{int}g_{int},mk} = \mathbb{E}\{\boldsymbol{g}_{i,mk}\boldsymbol{g}_{i,mk}^{H}\} = diag\left(\underbrace{\beta_{i,mk}, \dots, \beta_{i,mk}}_{N_r}\right) = \beta_{i,mk}\boldsymbol{I}_{N_r} \tag{5.12}$$

Compared to the MMSE combining at the EPU, the benefit of local MMSE estimation is that the size of the inverse matrix is $N_r \times N_r$ which is much smaller than the inverse matrix of (3.18). Furthermore, we notice that when $N_r = 1$, (5.10) is not equal to the MRC weight but is a scaled version:

$$v_{mk} = \left(\sum_{k'=1}^{K_{serv}}\left(\hat{g}_{mk'}\hat{g}_{mk'}^{*} + (1 - c_{mk'})\beta_{mk'}\right) + \sum_{k'=1}^{K_{int}}\beta_{i,mk'} + \frac{\sigma_z^2}{\sigma_x^2}\right)^{-1}\hat{g}_{mk}^{*} \tag{5.13}$$

Subsequently, the local estimated data $\check{x}_{mk}$ at the different APs are sent to the EPU together for decoding the data transmitted by the k-th UT. In the EPU, the linear combination adopting the weights w_{mk} is applied. The recovered data is then:

$$\begin{aligned}\hat{x}_k &= \sum_{m=1}^{M_{coor}} w_{mk}\check{x}_{mk}\\ &= \sum_{m=1}^{M_{coor}} w_{mk}\left(\sum_{k'=1}^{K_{serv}}\boldsymbol{v}_{mk}\boldsymbol{g}_{mk'}x_{k'} + \sum_{k'=1}^{K_{int}}\boldsymbol{v}_{mk}\boldsymbol{g}_{i,mk'}x_{i,k'} + \boldsymbol{v}_{mk}\boldsymbol{z}_m\right)\end{aligned} \tag{5.14}$$

If we define the vector form of the local estimated data as $\check{\boldsymbol{x}}_k = \left[\check{x}_{1k} \ \dots \ \check{x}_{M_{coor}k}\right]^T$, the local weighted channel $\boldsymbol{\gamma}_{kk'} = \left[\boldsymbol{v}_{1k}\boldsymbol{g}_{1k'} \ \dots \ \boldsymbol{v}_{M_{coor}k}\boldsymbol{g}_{M_{coor}k'}\right]^T$, and the local

weighted interfering channel $\boldsymbol{\gamma}_{i,kk'} = \left[\boldsymbol{v}_{1k}\boldsymbol{g}_{i,1k'} \ \ldots \ \boldsymbol{v}_{M_{coor}k}\boldsymbol{g}_{i,M_{coor}k'}\right]^T$, (5.14) can be rewritten as:

$$\hat{x}_k = \boldsymbol{w}_k\breve{\boldsymbol{x}}_k = \sum_{k'=1}^{K_{serv}} \boldsymbol{w}_k\boldsymbol{\gamma}_{kk'}x_{k'} + \sum_{k'=1}^{K_{int}} \boldsymbol{w}_k\boldsymbol{\gamma}_{i,kk'}x_{i,k'} + \boldsymbol{w}_k\boldsymbol{z}'_k \tag{5.15}$$

where $\boldsymbol{w}_k = \left[w_{1k} \ldots w_{M_{coor}k}\right]$ and $\boldsymbol{z}'_k = \left[\boldsymbol{v}_{1k}\boldsymbol{z}_1 \ldots \boldsymbol{v}_{M_{coor}k}\boldsymbol{z}_{M_{coor}}\right]^T$.

Although the CSI is unknown at the EPU, we notice that its statistics can be exploited. They may be obtained by averaging the effective channel $\boldsymbol{w}_k\boldsymbol{\gamma}_{kk'}$ within a finite coherence block. Therefore, the achievable SE is stated in the following.

Proposition 5.1. *Suppose that K_{serv} UTs are served by M_{coor} APs where each AP is equipped with N_r antennas in DD-MaMIMO. The estimations of the channel and data are processed at the local and decoding the data at the EPU. An achievable SE for the k-th UT is*

$$SE_k = \frac{\tau_u}{\tau_p + \tau_u}\log_2(1 + SIDNR_k) \tag{5.16}$$

where the factor $\tau_u/(\tau_p + \tau_u)$ is the fraction of the uplink transmission used for sending the data. The signal-to-interference-plus-noise ratio (SINR) for the k-th UT is given by

$$SINR_k = \frac{|\boldsymbol{w}_k\mathbb{E}\{\boldsymbol{\gamma}_{kk}\}|^2}{\sum_{k'=1}^{K_{serv}}\mathbb{E}\{|\boldsymbol{w}_k\boldsymbol{\gamma}_{kk'}|^2\} - |\boldsymbol{w}_k\mathbb{E}\{\boldsymbol{\gamma}_{kk}\}|^2 + \sum_{k'=1}^{K_{int}}\mathbb{E}\left\{\left|\boldsymbol{w}_k\boldsymbol{\gamma}_{i,kk'}\right|^2\right\} + \frac{\sigma_z^2}{\sigma_x^2}\boldsymbol{w}_k\boldsymbol{D}_k\boldsymbol{w}_k^H} \tag{5.17}$$

where $\boldsymbol{D}_k = diag\left(\mathbb{E}\{\|\boldsymbol{v}_{1k}\|^2\}, \mathbb{E}\{\|\boldsymbol{v}_{2k}\|^2\}, \ldots, \mathbb{E}\left\{\left\|\boldsymbol{v}_{M_{coor}k}\right\|^2\right\}\right) \in \mathbb{C}^{M_{coor}\times M_{coor}}$.

Proof: The proof follows the same steps as in [73, Corollary 1.3], where the achievable SE of a cellular system is derived by calculating the lower bound of the capacity of the discrete memoryless interference channel, if the channel response is deterministic. Similarly, the signal in (5.15) matches this discrete memoryless interference channel, where the input $x = x_k$, the deterministic channel response $h = \boldsymbol{w}_k\mathbb{E}\{\boldsymbol{\gamma}_{kk}\}$, the output $y = \boldsymbol{w}_k\breve{\boldsymbol{x}}_k$ and the interference term is

$$v = \sum_{k'=1}^{K_{serv}} \boldsymbol{w}_k \boldsymbol{\gamma}_{kk'} x_{k'} - \boldsymbol{w}_k \mathbb{E}\{\boldsymbol{\gamma}_{kk}\} x_k + \sum_{k'=1}^{K_{int}} \boldsymbol{w}_k \boldsymbol{\gamma}_{i,kk'} x_{i,k'} + \boldsymbol{w}_k \boldsymbol{z}'_k \tag{5.18}$$

We should note that all terms x, h, y and v are stated for the discrete memoryless interference channel model which follow the same definition as [73]. To utilise the statistical channel response, the channel hardening [57] is assumed which approximates $\boldsymbol{w}_k \boldsymbol{\gamma}_{kk}$ to its mean value $\boldsymbol{w}_k \mathbb{E}\{\boldsymbol{\gamma}_{kk}\}$ [24]. However, in [19] the author indicated that when the number of antennas is small, the channel hardening is not necessarily established. In such a scenario, the achievable SE will be underestimated. The interference term is obtained by adding and subtracting the desired signal $\boldsymbol{w}_k \mathbb{E}\{\boldsymbol{\gamma}_{kk}\} x_k$ from $\hat{x}_k$. Then, since the terms $\{x_k : k = 1, \dots, K_{serv}\}$, $\{x_{i,k} : k = 1, \dots, K_{int}\}$ and the elements in $\boldsymbol{z}'_k$ are zero mean, the interference term has zero mean $\mathbb{E}\{v\} = 0$. Furthermore, the interference term is uncorrelated with the input:

$$\mathbb{E}\{xv\} = \underbrace{\mathbb{E}\{\boldsymbol{w}_k \boldsymbol{\gamma}_{kk} - \boldsymbol{w}_k \mathbb{E}\{\boldsymbol{\gamma}_{kk}\}\}}_{=0} \mathbb{E}\{|x_k|^2\} = 0 \tag{5.19}$$

Therefore, the channel capacity C is lower bounded as:

$$C \geq \log_2\left(1 + \frac{\mathbb{E}\{|x|^2\}|h|^2}{\mathbb{E}\{|v|^2\}}\right) \tag{5.20}$$

where $\mathbb{E}\{|x|^2\} = \sigma_x^2$, $|h|^2 = |\boldsymbol{w}_k \mathbb{E}\{\boldsymbol{\gamma}_{kk}\}|^2$ and the variance of the interference term is:

$$\begin{aligned} \mathbb{E}\{|v|^2\} = \sigma_x^2 \sum_{k'=1}^{K_{serv}} \mathbb{E}\{|\boldsymbol{w}_k \boldsymbol{\gamma}_{kk'}|^2\} - \sigma_x^2 |\boldsymbol{w}_k \mathbb{E}\{\boldsymbol{\gamma}_{kk}\}|^2 \\ + \sigma_x^2 \sum_{k'=1}^{K_{int}} \mathbb{E}\left\{\left|\boldsymbol{w}_k \boldsymbol{\gamma}_{i,kk'}\right|^2\right\} + \sigma_z^2 \boldsymbol{w}_k \boldsymbol{D}_k \boldsymbol{w}_k^H \end{aligned} \tag{5.21}$$

where the terms in (5.18) are uncorrelated with each other. Then, by substituting (5.21) and the expressions for $\mathbb{E}\{|x|^2\}$ and $|h|^2$ into (5.20), and by considering the effective transmission fraction, we complete the proof.

5.4.1 Large Scale Fading Decoding

The large scale fading decoding which was initially known as pilot contamination postcoding in [78] was proposed to mitigate the interference for cell-free MaMIMO in [79]. It is an approach in which only large scale fading coefficients are applied to maximise the SE. Note that since the statistics of the local processing parameters are known at the EPU in DD-MaMIMO, we can find the combining weight to maximise the SE. This large scale fading decoding approach was also used for "*Level 3*" decoding in [24].

Corollary 5.1. *The combining vector for maximising the SE in Proposition 5.1 is given by*

$$\boldsymbol{w}_k = \mathbb{E}\{\boldsymbol{\gamma}_{kk}^H\}\left(\sum_{k'=1}^{K_{serv}} \mathbb{E}\{\boldsymbol{\gamma}_{kk'}\boldsymbol{\gamma}_{kk'}^H\} + \sum_{k'=1}^{K_{int}} \mathbb{E}\{\boldsymbol{\gamma}_{i,kk'}\boldsymbol{\gamma}_{i,kk'}^H\} + \frac{\sigma_z^2}{\sigma_x^2}\boldsymbol{D}_k\right)^{-1} \tag{5.22}$$

which gives rise to the maximum SINR value

$$\begin{aligned} SINR_k = \mathbb{E}\{\boldsymbol{\gamma}_{kk}^H\}\Bigg(&\sum_{k'=1}^{K_{serv}} \mathbb{E}\{\boldsymbol{\gamma}_{kk'}\boldsymbol{\gamma}_{kk'}^H\} - \mathbb{E}\{\boldsymbol{\gamma}_{kk}\}\mathbb{E}\{\boldsymbol{\gamma}_{kk}^H\} + \sum_{k'=1}^{K_{int}} \mathbb{E}\{\boldsymbol{\gamma}_{i,kk'}\boldsymbol{\gamma}_{i,kk'}^H\} \\ &+ \frac{\sigma_z^2}{\sigma_x^2}\boldsymbol{D}_k\Bigg)^{-1} \mathbb{E}\{\boldsymbol{\gamma}_{kk}\} \end{aligned} \tag{5.23}$$

Proof: For maximising the SINR, we first calculate the MSE $\epsilon_{d,k}$ between the data and the decoded data and then set the first derivative of the MSE with respect to $\boldsymbol{w}_k$ to zero.

$$\begin{aligned} \epsilon_{d,k} &= \mathbb{E}\{|x_k - \boldsymbol{w}_k\breve{\boldsymbol{x}}_k|^2\} \\ &= \mathbb{E}\{|x_k|^2\} + \mathbb{E}\{|\boldsymbol{w}_k\breve{\boldsymbol{x}}_k|^2\} - \mathbb{E}\{x_k\breve{\boldsymbol{x}}_k^H\boldsymbol{w}_k^H\} - \mathbb{E}\{x_k^*\boldsymbol{w}_k\breve{\boldsymbol{x}}_k\} \\ &= \sigma_x^2 + \boldsymbol{w}_k\mathbb{E}\{|\breve{\boldsymbol{x}}_k|^2\}\boldsymbol{w}_k^H - \mathbb{E}\{x_k\breve{\boldsymbol{x}}_k^H\}\boldsymbol{w}_k^H - \boldsymbol{w}_k\mathbb{E}\{x_k^*\breve{\boldsymbol{x}}_k\} \end{aligned} \tag{5.24}$$

Then, its first derivative is:

$$\frac{\partial \epsilon_{d,k}}{\partial \boldsymbol{w}_k} = \mathbb{E}\{|\breve{\boldsymbol{x}}_k|^2\}\boldsymbol{w}_k^H - \mathbb{E}\{x_k^*\breve{\boldsymbol{x}}_k\} \tag{5.25}$$

$$= \left(\sigma_x^2 \sum_{k'=1}^{K_{serv}} \mathbb{E}\{\boldsymbol{\gamma}_{kk'} \boldsymbol{\gamma}_{kk'}^H\} + \sigma_x^2 \sum_{k'=1}^{K_{int}} \mathbb{E}\{\boldsymbol{\gamma}_{i,kk'} \boldsymbol{\gamma}_{i,kk'}^H\} + \sigma_z^2 \boldsymbol{D}_k \right) \boldsymbol{w}_k^H - \sigma_x^2 \mathbb{E}\{\boldsymbol{\gamma}_{kk}\} = 0$$

Finally, we obtain the result:

$$\boldsymbol{w}_k^H = \left(\sigma_x^2 \sum_{k'=1}^{K_{serv}} \mathbb{E}\{\boldsymbol{\gamma}_{kk'} \boldsymbol{\gamma}_{kk'}^H\} + \sigma_x^2 \sum_{k'=1}^{K_{int}} \mathbb{E}\{\boldsymbol{\gamma}_{i,kk'} \boldsymbol{\gamma}_{i,kk'}^H\} + \sigma_z^2 \boldsymbol{D}_k \right)^{-1} \mathbb{E}\{\boldsymbol{\gamma}_{kk}\} \quad (5.26)$$

Note that all terms in the inverse matrix in (5.26) are Hermitian matrices, thus the Hermitian of (5.26) is equivalent to (5.22). Moreover, substituting (5.22) into (5.17), $SINR_k$ can be simplified as (5.23) which follows the same method as (3.22).

While the large scale fading decoding was adopted in [24], the authors did not explicitly explain why the use of combining vectors, which are composed of statistical parameters for data recovery, is considered the relevant decoding method. Hence, we attempt to answer this in two respects. Firstly, the statistical parameters within the combining vectors are deterministic at the EPU which only relates to the large scale fading coefficients. Secondly, it can be simply proved that the final combining vectors at the EPU only depend on the large scale fading coefficients in the case where local MRC data detection is applied. If each AP has a single antenna and uses local MRC estimation where the local estimation weight is $v_{mk} = \hat{g}_{mk}^*$, the statistical parameters in (5.22) are:

$$\begin{aligned}
&\mathbb{E}\{\boldsymbol{\gamma}_{kk}^H\} = \left[\mathbb{E}\{g_{1k}^* v_{1k}^*\} \dots \mathbb{E}\{g_{M_{coor}k}^* v_{M_{coor}k}^*\}\right] \\
&= \left[\mathbb{E}\{(\hat{g}_{1k}^* + \tilde{g}_{1k}^*)\hat{g}_{1k}\} \dots \mathbb{E}\{(\hat{g}_{M_{coor}k}^* + \tilde{g}_{M_{coor}k}^*)\hat{g}_{M_{coor}k}\}\right] \\
&= \left[\mathbb{E}\{\hat{g}_{1k}^* \hat{g}_{1k}\} \dots \mathbb{E}\{\hat{g}_{M_{coor}k}^* \hat{g}_{M_{coor}k}\}\right] \\
&= \left[c_{1k}\beta_{1k} \dots c_{M_{coor}k}\beta_{M_{coor}k}\right] \\
&= \left[\frac{\tau_p \beta_{1k}^2}{\tau_p \beta_{1k} + \sum_{k'=1}^{K_{int}} \beta_{i,1k'} + \sigma_z^2} \dots \frac{\tau_p \beta_{M_{coor}k}^2}{\tau_p \beta_{M_{coor}k} + \sum_{k'=1}^{K_{int}} \beta_{i,\mathrm{M_{coor}}k'} + \sigma_z^2}\right]
\end{aligned} \quad (5.27)$$

The term $\mathbb{E}\{\boldsymbol{\gamma}_{kk'} \boldsymbol{\gamma}_{kk'}^H\}$ in (5.22) can be expressed by:

$$\mathbb{E}\{\boldsymbol{\gamma}_{kk'}\boldsymbol{\gamma}_{kk'}^{H}\} = \begin{bmatrix} \mathbb{E}\{v_{1k}g_{1k'}g_{1k'}^{*}v_{1k}^{*}\} & \cdots & \mathbb{E}\left\{v_{1k}g_{1k'}g_{M_{coor}k'}^{*}v_{M_{coor}k}^{*}\right\} \\ \vdots & \ddots & \vdots \\ \mathbb{E}\{v_{M_{coor}k}g_{M_{coor}k'}g_{1k'}^{*}v_{1k}^{*}\} & \cdots & \mathbb{E}\left\{v_{M_{coor}k}g_{M_{coor}k'}g_{M_{coor}k'}^{*}v_{M_{coor}k}^{*}\right\} \end{bmatrix} \quad (5.28)$$

The on-diagonal elements in (5.28) are calculated (for an example), when $k' \neq k$:

$$\begin{aligned} &\mathbb{E}\{v_{mk}g_{mk'}g_{mk'}^{*}v_{mk}^{*}\} = \mathbb{E}\{\hat{g}_{mk}^{*}g_{mk'}g_{mk'}^{*}\hat{g}_{mk}\} \\ &= \beta_{mk'}\mathbb{E}\{\hat{g}_{mk}^{*}\hat{g}_{mk}\} \\ &= \beta_{mk'}c_{mk}\beta_{mk} \\ &= \frac{\tau_p\beta_{mk}^2\beta_{mk'}}{\tau_p\beta_{mk} + \sum_{k''=1}^{K_{int}}\beta_{i,\mathrm{m}k''} + \sigma_z^2} \end{aligned} \quad (5.29)$$

The on-diagonal elements for the case $k' = k$:

$$\begin{aligned} \mathbb{E}\{v_{mk}g_{mk}g_{mk}^{*}v_{mk}^{*}\} &= \mathbb{E}\{\hat{g}_{mk}^{*}(\hat{g}_{mk} + \tilde{g}_{mk})(\hat{g}_{mk}^{*} + \tilde{g}_{mk}^{*})\hat{g}_{mk}\} \\ &= \mathbb{E}\{|\hat{g}_{mk}|^4 + |\hat{g}_{mk}|^2|\tilde{g}_{mk}|^2\} \\ &\overset{(a)}{=} 2c_{mk}^2\beta_{mk}^2 + c_{mk}\beta_{mk}(1 - c_{mk})\beta_{mk} \\ &= c_{mk}^2\beta_{mk}^2 + c_{mk}\beta_{mk} \end{aligned} \quad (5.30)$$

where $\mathbb{E}\{|\hat{g}_{mk}|^4\} = 2c_{mk}^2\beta_{mk}^2$ in step (a) follows the fact given in [14, Eq. 60], but the proof was not provided. In Appendix 5.A, we give a rigorous proof and compare the theoretical results with the simulation which shows the original expression in (5.30) is not strict. However, it does not significantly affect the overall results because of the tiny discrepancy, and the term $\mathbb{E}\{|\hat{g}_{mk}|^4\}$ is only related to the large scale fading coefficients. Next, the off-diagonal elements in (5.28) are calculated when $k' = k$ and $m' \neq m$:

$$\begin{aligned} &\mathbb{E}\{v_{mk}g_{mk'}g_{m'k'}^{*}v_{m'k}^{*}\} = \mathbb{E}\{\hat{g}_{mk}^{*}g_{mk}g_{m'k}^{*}\hat{g}_{m'k}\} \\ &= \mathbb{E}\{|\hat{g}_{mk}|^2|\hat{g}_{m'k}|^2\} \\ &= \frac{\tau_p^2\beta_{mk}^2\beta_{m'k}^2}{\left(\tau_p\beta_{mk} + \sum_{k'=1}^{K_{int}}\beta_{i,\mathrm{m}k'} + \sigma_z^2\right)\left(\tau_p\beta_{m'k} + \sum_{k'=1}^{K_{int}}\beta_{i,\mathrm{m}'k'} + \sigma_z^2\right)} \end{aligned} \quad (5.31)$$

When $k' \neq k$, the expectation is zero. Then, the term $\mathbb{E}\{\boldsymbol{\gamma}_{i,kk'}\boldsymbol{\gamma}_{i,kk'}^{H}\}$ in (5.22) can be expressed by:

$$\mathbb{E}\{\boldsymbol{\gamma}_{i,kk'}\boldsymbol{\gamma}_{i,kk'}^{H}\} = \begin{bmatrix} \mathbb{E}\{v_{1k}g_{i,1k'}g_{i,1k'}^{*}v_{1k}^{*}\} & \cdots & \mathbb{E}\{v_{1k}g_{i,1k'}g_{i,M_{coor}k'}^{*}v_{M_{coor}k}^{*}\} \\ \vdots & \ddots & \vdots \\ \mathbb{E}\{v_{M_{coor}k}g_{i,M_{coor}k'}g_{i,1k'}^{*}v_{1k}^{*}\} & \cdots & \mathbb{E}\{v_{M_{coor}k}g_{i,M_{coor}k'}g_{i,M_{coor}k'}^{*}v_{M_{coor}k}^{*}\} \end{bmatrix} \quad (5.32)$$

The on-diagonal elements in (5.32) are:

$$\begin{aligned} \mathbb{E}\{v_{mk}g_{i,mk'}g_{i,mk'}^{*}v_{mk}^{*}\} &= \mathbb{E}\{\hat{g}_{mk}^{*}g_{i,mk'}g_{i,mk'}^{*}\hat{g}_{mk}\} \\ &= \frac{\tau_p\beta_{mk}^2\beta_{i,mk'}}{\tau_p\beta_{mk} + \sum_{k'=1}^{K_{int}}\beta_{i,mk'} + \sigma_z^2} \end{aligned} \quad (5.33)$$

Due to the uncorrelation among the different interfering channels, the off-diagonal elements in (5.32) are zero. Finally, the term $\boldsymbol{D}_k$ can be calculated

$$\begin{aligned} \boldsymbol{D}_k &= diag\left(\mathbb{E}\{\|\boldsymbol{v}_{1k}\|^2\}, \mathbb{E}\{\|\boldsymbol{v}_{2k}\|^2\}, \ldots, \mathbb{E}\left\{\left\|\boldsymbol{v}_{M_{coor}k}\right\|^2\right\}\right) \\ &= diag\left(\mathbb{E}\{|\hat{g}_{1k}|^2\}, \ldots, \mathbb{E}\left\{\left|\hat{g}_{M_{coor}k}\right|^2\right\}\right) \\ &= diag\left(\frac{\tau_p\beta_{1k}^2}{\tau_p\beta_{1k} + \sum_{k'=1}^{K_{int}}\beta_{i,1k'} + \sigma_z^2}, \ldots, \frac{\tau_p\beta_{M_{coor}k}^2}{\tau_p\beta_{M_{coor}k} + \sum_{k'=1}^{K_{int}}\beta_{i,\mathrm{M_{coor}}k'} + \sigma_z^2}\right) \end{aligned} \quad (5.34)$$

Substituting (5.27), (5.29), (5.30), (5.31), (5.33) and (5.34) into (5.22), this proves that the combining scalar in the single-antenna AP with local MRC estimation case only depends on the large scale fading coefficients. Subsequently, a sketch of a derivation of the combining vector for the multiple-antenna AP with local MRC estimation is given in Appendix 5.B. However, it is hard to derive the closed-form expression for the combining vector with local MMSE estimation. So, we leave this for future work.

5.4.2 Simple Decoding

In the large scale fading decoding, the local estimated data $\check{x}_{mk}$ for all UTs, which include $(\tau_c - \tau_p)K_{serv}$ complex values, are sent to the EPU from each AP in each coherence block. In addition, the knowledge of $\mathbb{E}\{\boldsymbol{\gamma}_{kk}^{H}\}$, $\sum_{k'=1}^{K_{serv}}\mathbb{E}\{\boldsymbol{\gamma}_{kk'}\boldsymbol{\gamma}_{kk'}^{H}\}$, $\sum_{k'=1}^{K_{int}}\mathbb{E}\{\boldsymbol{\gamma}_{i,kk'}\boldsymbol{\gamma}_{i,kk'}^{H}\}$ and $\boldsymbol{D}_k$ are also required at the EPU. Hence, it needs M_{coor} complex values for $\mathbb{E}\{\boldsymbol{\gamma}_{kk}^{H}\}$, $K_{serv}M_{coor}^2$ complex values for $\sum_{k'=1}^{K_{serv}}\mathbb{E}\{\boldsymbol{\gamma}_{kk'}\boldsymbol{\gamma}_{kk'}^{H}\}$, $K_{int}M_{coor}^2$ complex values for $\sum_{k'=1}^{K_{int}}\mathbb{E}\{\boldsymbol{\gamma}_{i,kk'}\boldsymbol{\gamma}_{i,kk'}^{H}\}$ and M_{coor} real value for $\boldsymbol{D}_k$ to be

sent for joint decoding of the *k*-th UT. This scheme optimises the processing which provides the highest SE in the local estimation of DD-MaMIMO, but the demand of the statistical parameters will be increased if the number of APs and UTs grows which leads to inaccurate decoding when the statistics vary with time [24].

To solve this issue, we propose an alternative decoding method in which the local estimated data is simply decoded by averaging at the EPU. This is computed by:

$$\hat{x}_k = \sum_{m=1}^{M_{coor}} w_{mk} \breve{x}_{mk} = \frac{1}{M_{coor}} \sum_{m=1}^{M_{coor}} \breve{x}_{mk} \tag{5.35}$$

where the local estimated data $\breve{x}_{mk}$ is acquired by (5.6). This also can be expressed in the same form as (5.15) if we define the vector of combining weight $\boldsymbol{w}_k = [1/M_{coor}, \dots, 1/M_{coor}]$.

Corollary 5.2. *Adopting the simple decoding at the EPU with the local estimation in DD-MaMIMO, the effective SINR for calculating SE in Proposition 5.1 can be derived from*

$$SINR_k = \frac{\left|\sum_{m=1}^{M_{coor}} \mathbb{E}\{\boldsymbol{v}_{mk}\boldsymbol{g}_{mk}\}\right|^2}{\begin{matrix}\sum_{k'=1}^{K_{serv}} \mathbb{E}\left\{\left|\sum_{m=1}^{M_{coor}} \boldsymbol{v}_{mk}\boldsymbol{g}_{mk'}\right|^2\right\} - \left|\sum_{m=1}^{M_{coor}} \mathbb{E}\{\boldsymbol{v}_{mk}\boldsymbol{g}_{mk}\}\right|^2 \\ + \sum_{k'=1}^{K_{int}} \mathbb{E}\left\{\left|\sum_{m=1}^{M_{coor}} \boldsymbol{v}_{mk}\boldsymbol{g}_{i,mk'}\right|^2\right\} + \frac{\sigma_z^2}{\sigma_x^2}\sum_{m=1}^{M_{coor}} \mathbb{E}\{\|\boldsymbol{v}_{mk}\|^2\}\end{matrix}} \tag{5.36}$$

where $\boldsymbol{v}_{mk}$ is the local data estimation weight which can be any combining, such as MRC or MMSE.

Proof: The proof follows the same steps as the proof of Proposition 5.1.

In this simple decoding scenario, the EPU does not require the statistics of the local processing parameters, thereby reducing the overhead of fronthaul. When compared to cell-free MaMIMO, a notable advantage of simple decoding with local estimation in DD-MaMIMO is that the weaker APs, potentially located at the network edge, do not significantly impair the decoding process. According to (5.35), the final decoding in the cell-free MaMIMO system can be impacted by averaging the poor estimated data from a weaker AP. However, in DD-MaMIMO, only APs within the coordination region serve the UTs within the service region, thereby mitigating this effect.

5.5 Scalable Network

In sub-chapter 3.6, we defined scalability for a DD-MaMIMO system in which the channel estimation and data detection are processed in the EPU. Applying this definition, we will discuss scalability for local estimation in DD-MaMIMO in this sub-chapter. First, each AP only needs to estimate $K_{serv}N_r$ channels, even as the number of UTs over the whole network $K \to \infty$. We assume the AP knows large scale fading coefficients from all UTs to itself, which can be obtained by averaging the power of received signals over a finite time period. Note that the UTs far outside the coordination region can be ignored due to their tiny values. Hence, the number of effective UTs corresponding to the system is finite. Second, for the uplink data detection, each AP needs to compute $\{\boldsymbol{v}_{mk}\boldsymbol{y}_{u,m}: k = 1, \dots, K_{serv}\}$ in (5.6) exploiting K_{serv} combining vectors $\{\boldsymbol{v}_{mk}: k = 1, \dots, K_{serv}\}$, which has finite computational complexity as $K \to \infty$. Similarly, on the downlink, each AP only creates the transmitted signals for K_{serv} UTs within the service area with respect to the same coordination region. Then, each APs forwards K_{serv} locally detected data to the EPU for the final decoding through the fronthaul link. If we use simple decoding method, no extra fronthaul signalling is required. Alternatively, the statistical parameters need to be sent to EPU for the large scale fading decoding. In (5.22), the combining vector $\boldsymbol{w}_k$ is calculated by leveraging the expectation of $\{\boldsymbol{\gamma}_{kk'}\boldsymbol{\gamma}_{kk'}^H: k' = 1, \dots, K_{serv}\}$, $\{\boldsymbol{\gamma}_{i,kk'}\boldsymbol{\gamma}_{i,kk'}^H: k' = 1, \dots, K_{int}\}$, $\boldsymbol{\gamma}_{kk}^H$ and $\boldsymbol{D}_k$. These statistics can be obtained by averaging the power of locally detected data $|\check{x}_{mk}|^2$ and the power of $|\check{x}_{mk}\check{x}_{mk}^H|^2$ over a finite time period. Therefore, it satisfies the third task in Definition 3.1. Finally, the concept of scalability of power control is not applicable to the local estimation in DD-MaMIMO, due to the absence of power control.

5.6 System Performance

In this sub-chapter, we will provide the numerical results for analysing the system performance of local estimation in DD-MaMIMO. We adopt the same parameters as Chapter 3, which includes the path loss model, the large scale fading coefficient with the uncorrelated shadowing, the simulation areas, the density of the UTs and APs, and the distance between the neighbouring EPUs.

In Fig. 5.1 we show the CDF of SE for local estimation in DD-MaMIMO with

different radii of the coordination region. MMSE estimation is used for the channel estimation and data detection at the single-antenna APs, and the large scale fading decoding method detailed in sub-subchapter 5.4.1 is exploited. The result indicates that the DD-MaMIMO system can achieve higher SE than cell-free MaMIMO, especially as the radius of the coordination region expands. At the 90% likely SE points which correspond to the 0.1 value on the vertical axis, DD-MaMIMO with a 400m radius outperforms the cell-free case by a factor of 2, which is a substantial enhancement. In [24, Fig. 2(a)], the author presented the CDF of the SE for "Level 3" decoding in cell-free MaMIMO which applied the local MMSE estimation and large scale fading decoding. To give a comparable result in DD-MaMIMO, we select a coordination region with a radius of 282 metres. This configuration maintains system-level equivalence, where 30 UTs reuse the $\tau_p = 10$ pilot sequences. Comparing with the SE in [24, Fig. 2(a)] where it achieves 3.3 bit/s/Hz for 90% UTs, the SE in the comparable DD-MaMIMO system reaches 3.5 bit/s/Hz. Furthermore, for 50% UTs DD-MaMIMO can obtain 5.2 bit/s/Hz however it only has 4.4 bit/s/Hz in [24]. Besides, the most important difference is that DD-MaMIMO is scalable.

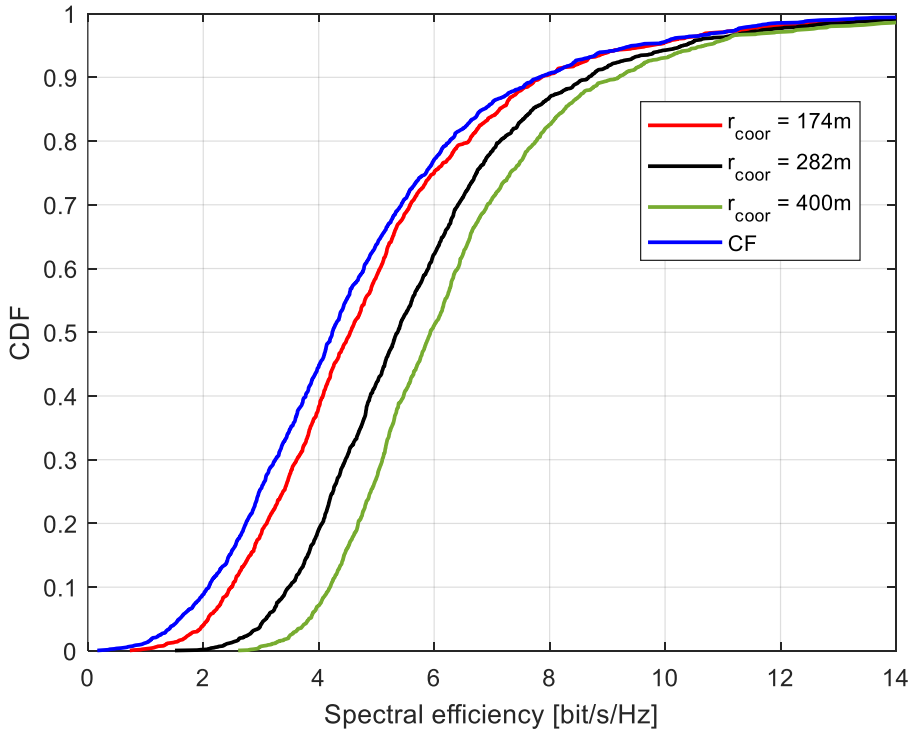


Fig. 5.1 The CDF of SE in DD-MaMIMO with varying radii of the coordination region, adopting local MMSE estimation for the channel and data at the APs with large scale fading decoding at the EPU (the single-antenna AP case). CF denotes cell-free MaMIMO. r_{coor} is the radius of the coordination region.

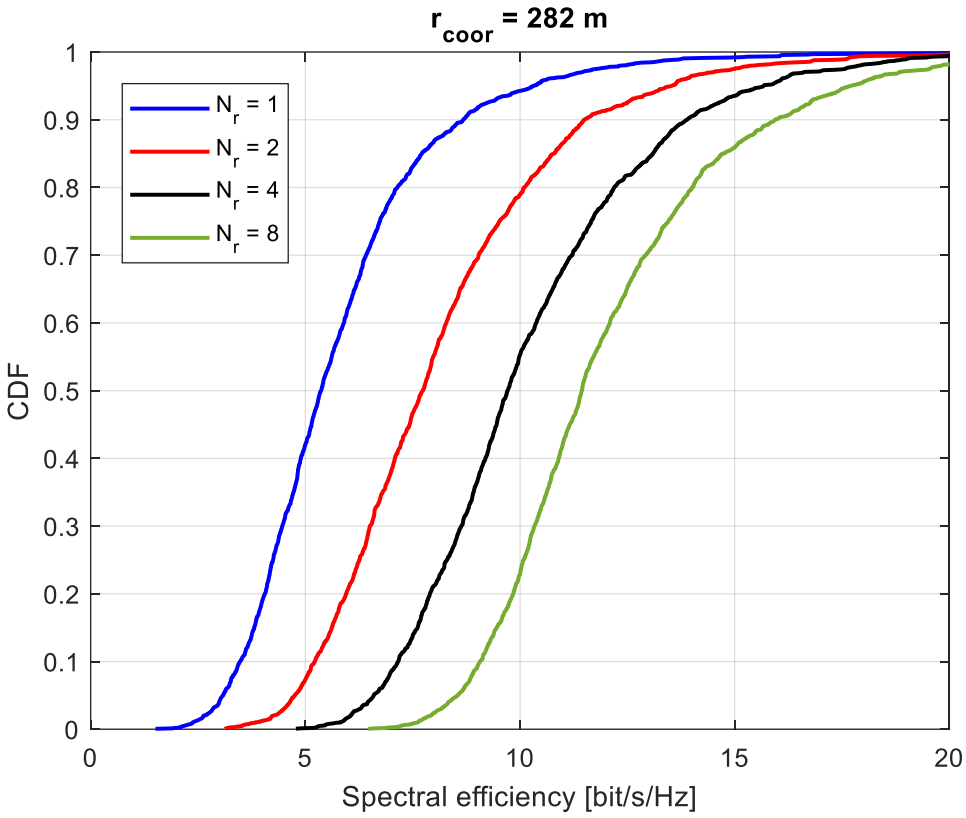


Fig. 5.2 The CDF of SE in DD-MaMIMO with different number of antennas N_r for each AP, adopting local MMSE estimation for the channel and data at the APs with large scale fading decoding at the EPU (the radius of the coordination region is 282 m).

Fig. 5.2 presents the CDF of SE with different number of antennas for each AP in DD-MaMIMO. It is obvious that the SE increases with the number of antennas. At the 90% likely SE points, the performance improves by more than a factor of 2 when the number of antennas is 8, compared to the single-antenna scenario. However, the local computation complexity will be dramatically increased.

Then, we compare the performance using the local estimation in DD-MaMIMO with the case applying the centralised estimation which is at the EPU to DD-MaMIMO. Fig. 5.3 shows that when the channel and the data are estimated at the EPU, the system can obtain the best SE because of the coordination among the APs. Moreover, adopting the large scale fading decoding with local MMSE estimation achieves higher SE than the case where the local estimated data is simply decoded by averaging at the EPU. Especially, in the multiple-antenna AP scenario $N_r = 4$, the former has a significant improvement around 7.3 bit/s/Hz for 90% UTs, compared to the simple decoding 4 bit/s/Hz for 90% UTs. Furthermore, we also compare the single-antenna case with [24], the results indicate that our system performs better.

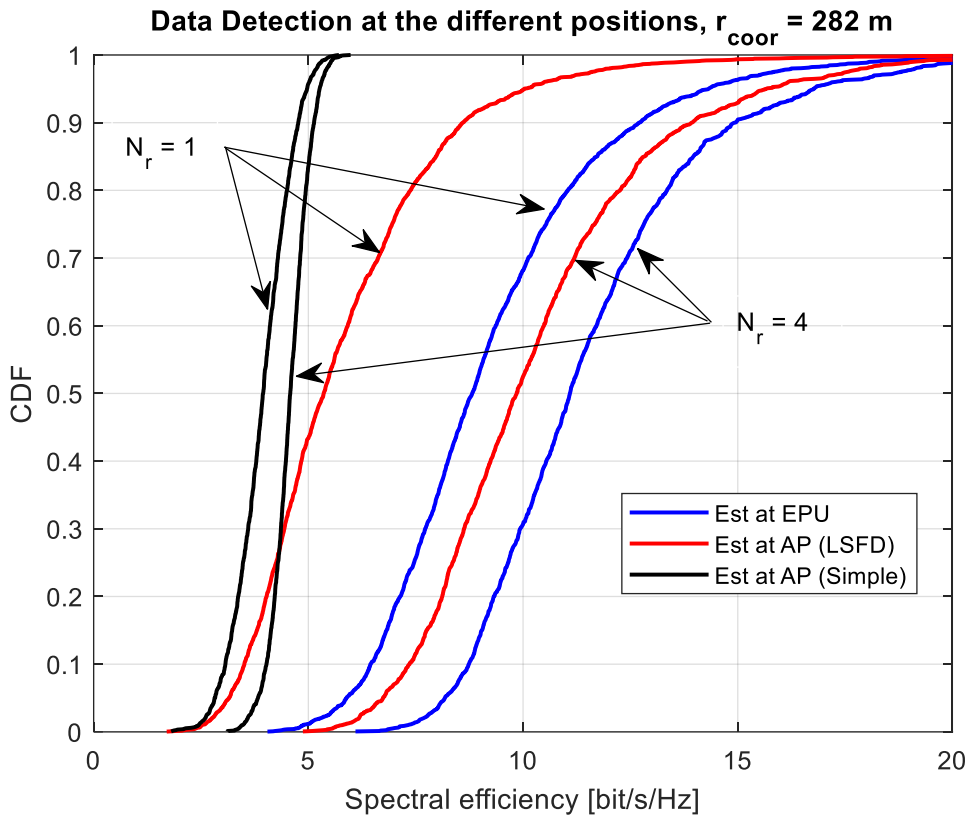


Fig. 5.3 The CDF of SE in DD-MaMIMO with the data detection at the different positions and different number of antennas N_r for each AP. The local MMSE estimation is used. LSFD denotes large scale fading decoding. Simple refers to the simple decoding. Est means estimate.

As we discussed in sub-chapter 5.4, when the number of antennas for each AP is one the local MMSE estimation is different from the MRC. Hence, we adopt the different schemes of the local data estimation and final decoding for the DD-MaMIMO system and the performance is given in Fig. 5.4. It shows that the use of local MRC estimation to detect data is worse than MMSE estimation, even if there is only one antenna for each AP. The large scale fading decoding gives the better performance than the simple decoding for both local estimation schemes.

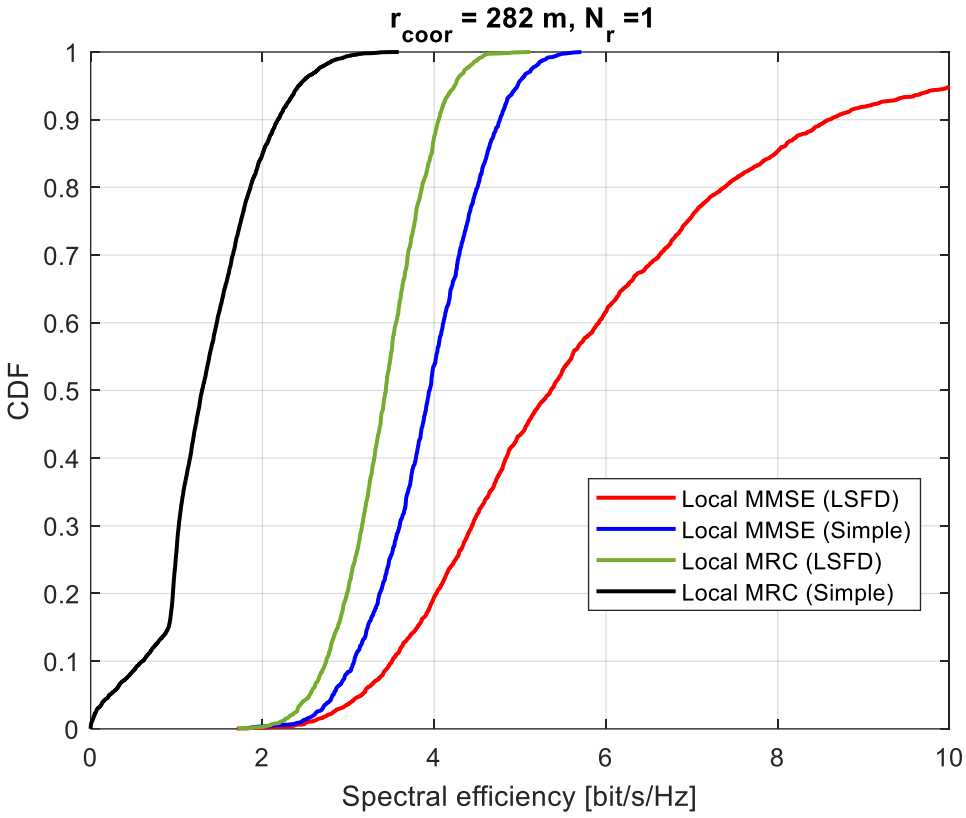


Fig. 5.4 The CDF of SE in DD-MaMIMO with the different local estimation and final decoding schemes. Local MMSE/MRC denotes the local data estimation method. LSFD/Simple represents the decoding scheme.

5.7 Summary

In this chapter, we have investigated the local estimation in DD-MaMIMO where the signal processing occurs at the APs. To obtain the CSI locally, the MMSE channel estimation was explored: this involves the same calculations as the centralised version. We then discussed the local MRC and local MMSE data detection which are processed by exploiting the CSI acquired only at the related AP. Our simulations reveal that MMSE data detection can be implemented even with a single-antenna AP, and it performs better than the MRC case. Subsequently, we considered two decoding methods—large scale fading decoding and simple decoding—to recover the data at the EPU. The combining vectors for both approaches were provided. In the large scale fading decoding scenario, we derived the closed-form expression for the combining vectors with local MRC data estimation. However, it was hard to obtain the closed-form expression applying MMSE data detection due to the statistical CSI needed for calculating the combing vectors, which includes an inverse matrix that is not easy to

solve. In the simple decoding case, the combining vectors were given with respect to the averaging factors which exclude the statistics of CSI, hence the computation complexity for the EPU is greatly decreased.

Finally, we have presented numerical results to evaluate the performance of local estimation in DD-MaMIMO. The simulation indicated that shifting the signal processing to the AP led to a loss of SE, especially in the single-antenna AP case with simple decoding. However, this reduced the computational burden on the EPU. Additionally, we observed that the performance in the multiple-antenna AP scenario, where the local estimation and large scale fading decoding are implemented, does not significantly differ from centralised estimation in DD-MaMIMO. Overall, the selection of signal processing in different positions is a trade-off. Both schemes can be implemented in Open RAN and fulfil the demands of 6G.

5.8 Appendix

Appendix 5.A: The proof of the term $\mathbb{E}\{|\hat{g}_{mk}|^4\}$

In [14, Eq. 60], the author mentioned the fact $\mathbb{E}\{|\hat{g}_{mk}|^4\} = 2\gamma^2 = 2c_{mk}^2\beta_{mk}^2$ but the detail was not provided and referenced. Hence, we first give the probable reason and then provide a strict proof. Since the term $\hat{g}_{mk}$ is the estimation of g_{mk}, the estimated channel can follow the same type of distribution as the channel coefficient. We know that $g_{mk} = \beta_{mk}^{1/2} h_{mk}$ in which $h_{mk} \sim \mathcal{CN}(0,1)$, hence $g_{mk} \sim \mathcal{CN}(0, \beta_{mk})$. Further, the term $\mathbb{E}\{|g_{mk}|^4\} = 2\beta_{mk}^2$ which is because:

$$\mathbb{E}\{|g_{mk}|^4\} = \beta_{mk}^2 \mathbb{E}\{|h_{mk}|^4\} = \beta_{mk}^2 \mathbb{E}\left\{\left|\frac{\sqrt{2}}{2} h'_{mk}\right|^4\right\} = 2\beta_{mk}^2 \tag{5.37}$$

where $h'_{mk} \sim \mathcal{CN}(0,2)$ and $|h'_{mk}|^2 \sim \mathcal{X}^2(2)$ in which $\mathcal{X}^2(k)$ denotes the chi-squared distribution with k degrees of the freedom. Then, $\mathbb{E}\{|h'_{mk}|^4\} = \mathbb{E}^2\{|h'_{mk}|^2\} + \sigma^2$ where the variance σ^2 of $|h'_{mk}|^2$ equals 4 and the mean is 2. Therefore $\mathbb{E}\{|h'_{mk}|^4\} = 8$, leading us to (5.38). Given the variance $\mathbb{E}\{|\hat{g}_{mk}|^2\} = c_{mk}\beta_{mk}$, we have the result $\mathbb{E}\{|\hat{g}_{mk}|^4\} = 2c_{mk}^2\beta_{mk}^2$ with the same calculation as (5.37).

Next, we give a rigorous proof by applying the definition of $\hat{g}_{mk}$ which is expressed by (3.6).

$$
\begin{aligned}
&\mathbb{E}\{|\hat{g}_{mk}|^4\} = c_{mk}^4 \mathbb{E}\{|\breve{g}_{mk}|^4\} \\
&= c_{mk}^4 \mathbb{E}\left\{\left| g_{mk} + \frac{1}{\tau_p}\sum_{k'=1}^{K_{int}} g_{i,mk'} \boldsymbol{\varphi}_{i,k'}^T \boldsymbol{\varphi}_k^* + \frac{1}{\tau_p} \boldsymbol{z}_{p,m} \boldsymbol{\varphi}_k^* \right|^4 \right\} \\
&= c_{mk}^4 \mathbb{E}\left\{ |g_{mk}|^4 + 4|g_{mk}|^2 \frac{1}{\tau_p^2} \sum_{k'=1}^{K_{int}} \left| g_{i,mk'} \boldsymbol{\varphi}_{i,k'}^T \boldsymbol{\varphi}_k^* \right|^2 \right. \\
&\qquad + 4|g_{mk}|^2 \frac{1}{\tau_p^2} \left| \boldsymbol{z}_{p,m} \boldsymbol{\varphi}_k^* \right|^2 + \frac{1}{\tau_p^4} \sum_{k'=1}^{K_{int}} \left| g_{i,mk'} \right|^4 \left| \boldsymbol{\varphi}_{i,k'}^T \boldsymbol{\varphi}_k^* \right|^4 \\
&\qquad + \frac{2}{\tau_p^4} \sum_{k'=1}^{K_{int}} \sum_{k'' \neq k'}^{K_{int}} \left| g_{i,mk'} \right|^2 \left| \boldsymbol{\varphi}_{i,k'}^T \boldsymbol{\varphi}_k^* \right|^2 \left| g_{i,mk''} \right|^2 \left| \boldsymbol{\varphi}_{i,k''}^T \boldsymbol{\varphi}_k^* \right|^2 \\
&\qquad \left. + \frac{4}{\tau_p^4} \sum_{k'=1}^{K_{int}} \left| g_{i,mk'} \boldsymbol{\varphi}_{i,k'}^T \boldsymbol{\varphi}_k^* \right|^2 \left| \boldsymbol{z}_{p,m} \boldsymbol{\varphi}_k^* \right|^2 + \frac{1}{\tau_p^4} \left| \boldsymbol{z}_{p,m} \boldsymbol{\varphi}_k^* \right|^4 \right\} \qquad (5.38) \\
&\overset{(a)}{=} c_{mk}^4 \left(2\beta_{mk}^2 + \frac{4}{\tau_p} \beta_{mk} \sum_{k'=1}^{K_{int}} \beta_{i,mk'} + \frac{4}{\tau_p} \beta_{mk} \sigma_z^2 + \frac{2}{\tau_p^4} \sum_{k'=1}^{K_{int}} \beta_{i,mk'}^2 \tau_p^3 \right. \\
&\qquad \left. + \frac{2}{\tau_p^4} \sum_{k'=1}^{K_{int}} \sum_{k'' \neq k'}^{K_{int}} \beta_{i,mk'} \beta_{i,mk''} \tau_p^2 + \frac{4}{\tau_p^4} \sum_{k'=1}^{K_{int}} \beta_{i,mk'} \tau_p^2 \sigma_z^2 + \frac{2}{\tau_p^4} \sigma_z^4 \tau_p^2 \right) \\
&= 2c_{mk}^4 \left(\beta_{mk}^2 + \frac{2}{\tau_p} \beta_{mk} \sum_{k'=1}^{K_{int}} \beta_{i,mk'} + \frac{2\sigma_z^2}{\tau_p} \beta_{mk} + \frac{1}{\tau_p} \sum_{k'=1}^{K_{int}} \beta_{i,mk'}^2 \right. \\
&\qquad \left. + \frac{1}{\tau_p^2} \sum_{k'=1}^{K_{int}} \sum_{k'' \neq k'}^{K_{int}} \beta_{i,mk'} \beta_{i,mk''} + \frac{2\sigma_z^2}{\tau_p^2} \sum_{k'=1}^{K_{int}} \beta_{i,mk'} + \frac{\sigma_z^4}{\tau_p^2} \right)
\end{aligned}
$$

where in step (a) we assume that $\mathbb{E}\left\{\left|\boldsymbol{\varphi}_{i,k'}^T \boldsymbol{\varphi}_k^*\right|^4\right\} = \tau_p^3$. Moreover, we notice that the result is not equivalent to $2c_{mk}^2 \beta_{mk}^2$. This is because the difference happens in the term $\mathbb{E}\left\{\left|\frac{1}{\tau_p}\sum_{k'=1}^{K_{int}} g_{i,mk'} \boldsymbol{\varphi}_{i,k'}^T \boldsymbol{\varphi}_k^*\right|^4\right\}$ in (5.38), which is calculated by:

$$
\mathbb{E}\left\{\left|\frac{1}{\tau_p}\sum_{k'=1}^{K_{int}} g_{i,mk'} \boldsymbol{\varphi}_{i,k'}^T \boldsymbol{\varphi}_k^*\right|^4\right\} = \frac{2}{\tau_p}\sum_{k'=1}^{K_{int}} \beta_{i,mk'}^2 + \frac{2}{\tau_p^2}\sum_{k'=1}^{K_{int}}\sum_{k'' \neq k'}^{K_{int}} \beta_{i,mk'}\beta_{i,mk''} \qquad (5.39)
$$

We expand the result $2c_{mk}^2\beta_{mk}^2$ and find that all terms are the same as (5.38) except the term $\frac{2}{\tau_p^2}\sum_{k'=1}^{K_{int}}\sum_{k''=1}^{K_{int}}\beta_{i,mk'}\beta_{i,mk''}$ which is different from (5.39). To evaluate this effect of the discrepancy, we compare this expanded term and (5.39) with the simulated result, respectively. In addition, we also make a comparison between the simulation and the theory of (5.38), plus the original assumption $2c_{mk}^2\beta_{mk}^2$. Fig. 5.5(a) shows that the theoretical result of (5.39) is twice $\frac{2}{\tau_p^2}\sum_{k'=1}^{K_{int}}\sum_{k''=1}^{K_{int}}\beta_{i,mk'}\beta_{i,mk''}$ but the value is tiny. Besides, the simulation matches what we expect in the figure. From Fig. 5.5(b), we can conclude that the original fact mentioned in [14] where $\mathbb{E}\{|\hat{g}_{mk}|^4\} = 2c_{mk}^2\beta_{mk}^2$ is not strict and the closed-form expression (5.38) is the solution.

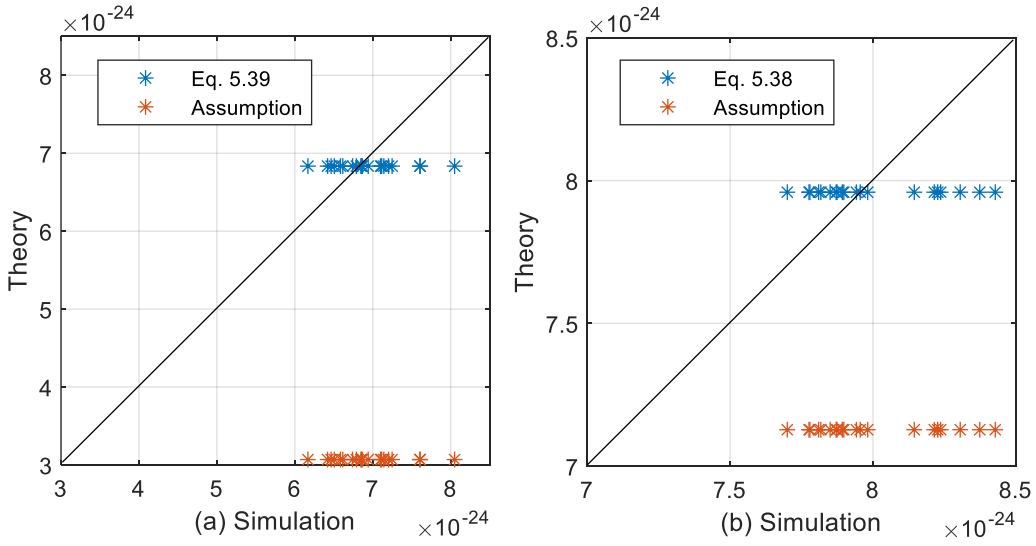


Fig. 5.5 (a): The simulated results of (5.39) versus the theory (blue star) and the assumption (red star). (b): The simulated results of (5.38) versus the theory (blue star) and the assumption $2c_{mk}^2\beta_{mk}^2$ (red star).

Appendix 5.B: The derivation of the combining vector for the multiple-antenna AP with local MRC estimation

In this scenario, the local estimation vector $\boldsymbol{v}_{mk} = \hat{\boldsymbol{g}}_{mk}^H$

$$\begin{aligned}
&\mathbb{E}\{\boldsymbol{\gamma}_{kk}^H\} = \left[\mathbb{E}\{\boldsymbol{g}_{1k}^H \boldsymbol{v}_{1k}^H\} \dots \mathbb{E}\{\boldsymbol{g}_{M_{coor}k}^H \boldsymbol{v}_{M_{coor}k}^H\}\right] \\
&= \left[\mathbb{E}\{(\hat{\boldsymbol{g}}_{1k}^H + \tilde{\boldsymbol{g}}_{1k}^H)\hat{\boldsymbol{g}}_{1k}\} \dots \mathbb{E}\{(\hat{\boldsymbol{g}}_{M_{coor}k}^H + \tilde{\boldsymbol{g}}_{M_{coor}k}^H)\hat{\boldsymbol{g}}_{M_{coor}k}\}\right] \\
&= \left[\mathbb{E}\{\hat{\boldsymbol{g}}_{1k}^H \hat{\boldsymbol{g}}_{1k}\} \dots \mathbb{E}\{\hat{\boldsymbol{g}}_{M_{coor}k}^H \hat{\boldsymbol{g}}_{M_{coor}k}\}\right] \\
&= \left[\sum_{n=1}^{N_r} \mathbb{E}\{\hat{g}_{1nk}^* \hat{g}_{1nk}\} \dots \sum_{n=1}^{N_r} \mathbb{E}\{\hat{g}_{M_{coor}nk}^* \hat{g}_{M_{coor}nk}\}\right] \\
&= \left[\sum_{n=1}^{N_r} c_{1nk}\beta_{1nk} \dots \sum_{n=1}^{N_r} c_{M_{coor}nk}\beta_{M_{coor}nk}\right] \\
&= \left[\frac{\tau_p N_r \beta_{1k}^2}{\tau_p \beta_{1k} + \sum_{k'=1}^{K_{int}} \beta_{i,1k'} + \sigma_z^2} \dots \frac{\tau_p N_r \beta_{M_{coor}k}^2}{\tau_p \beta_{M_{coor}k} + \sum_{k'=1}^{K_{int}} \beta_{i,\mathrm{M_{coor}}k'} + \sigma_z^2}\right]
\end{aligned} \tag{5.40}$$

The term $\mathbb{E}\{\boldsymbol{\gamma}_{kk'}\boldsymbol{\gamma}_{kk'}^H\}$ in (5.22) can be expressed by:

$$\begin{aligned}
&\mathbb{E}\{\boldsymbol{\gamma}_{kk'}\boldsymbol{\gamma}_{kk'}^H\} \\
&= \begin{bmatrix} \mathbb{E}\{\boldsymbol{v}_{1k}\boldsymbol{g}_{1k'}\boldsymbol{g}_{1k'}^H\boldsymbol{v}_{1k}^H\} & \cdots & \mathbb{E}\left\{\boldsymbol{v}_{1k}\boldsymbol{g}_{1k'}\boldsymbol{g}_{M_{coor}k'}^H\boldsymbol{v}_{M_{coor}k}^H\right\} \\ \vdots & \ddots & \vdots \\ \mathbb{E}\{\boldsymbol{v}_{M_{coor}k}\boldsymbol{g}_{M_{coor}k'}\boldsymbol{g}_{1k'}^H\boldsymbol{v}_{1k}^H\} & \cdots & \mathbb{E}\left\{\boldsymbol{v}_{M_{coor}k}\boldsymbol{g}_{M_{coor}k'}\boldsymbol{g}_{M_{coor}k'}^H\boldsymbol{v}_{M_{coor}k}^H\right\} \end{bmatrix}
\end{aligned} \tag{5.41}$$

The on-diagonal elements in (5.41) are calculated, when $k' \neq k$:

$$\begin{aligned}
&\mathbb{E}\{\boldsymbol{v}_{mk}\boldsymbol{g}_{mk'}\boldsymbol{g}_{mk'}^H\boldsymbol{v}_{mk}^H\} = \mathbb{E}\{\hat{\boldsymbol{g}}_{mk}^H\boldsymbol{g}_{mk'}\boldsymbol{g}_{mk'}^H\hat{\boldsymbol{g}}_{mk}\} \\
&= \mathbb{E}\left\{\sum_{n=1}^{N_r} \hat{g}_{mnk}^* g_{mnk'} \sum_{n'=1}^{N_r} g_{mn'k'}^* \hat{g}_{mn'k}\right\} \\
&= \sum_{n=1}^{N_r} \mathbb{E}\{\hat{g}_{mnk}^* g_{mnk'} g_{mnk'}^* \hat{g}_{mnk}\} \\
&= N_r\left(\frac{\tau_p \beta_{mk}^2 \beta_{mk'}}{\tau_p \beta_{mk} + \sum_{k''=1}^{K_{int}} \beta_{i,mk''} + \sigma_z^2}\right)
\end{aligned} \tag{5.42}$$

The on-diagonal elements for the case $k' = k$:

$$\begin{aligned}
&\mathbb{E}\{\boldsymbol{v}_{mk}\boldsymbol{g}_{mk}\boldsymbol{g}_{mk}^H\boldsymbol{v}_{mk}^H\} = \mathbb{E}\{\hat{\boldsymbol{g}}_{mk}^H(\hat{\boldsymbol{g}}_{mk}\hat{\boldsymbol{g}}_{mk}^H + \tilde{\boldsymbol{g}}_{mk}\tilde{\boldsymbol{g}}_{mk}^H)\hat{\boldsymbol{g}}_{mk}\} \\
&= \mathbb{E}\{\hat{\boldsymbol{g}}_{mk}^H\,\hat{\boldsymbol{g}}_{mk}\hat{\boldsymbol{g}}_{mk}^H\hat{\boldsymbol{g}}_{mk}\} + \mathbb{E}\{\hat{\boldsymbol{g}}_{mk}^H\,\tilde{\boldsymbol{g}}_{mk}\tilde{\boldsymbol{g}}_{mk}^H\hat{\boldsymbol{g}}_{mk}\}
\end{aligned} \tag{5.43}$$

$$
\begin{aligned}
&= \mathbb{E}\left\{\sum_{n=1}^{N_r}|\hat{g}_{mnk}|^2\sum_{n'=1}^{N_r}|\hat{g}_{mn'k}|^2\right\} + \mathbb{E}\left\{\sum_{n=1}^{N_r}\hat{g}^*_{mnk}\tilde{g}_{mnk}\sum_{n'=1}^{N_r}\tilde{g}^*_{mn'k}\hat{g}_{mn'k}\right\} \\
&= \sum_{n=1}^{N_r}\mathbb{E}\{|\hat{g}_{mnk}|^4\} + \sum_{n=1}^{N_r}\sum_{n'\neq n}^{N_r}\mathbb{E}\{|\hat{g}_{mnk}|^2|\hat{g}_{mn'k}|^2\} + \sum_{n=1}^{N_r}\mathbb{E}\{|\hat{g}_{mnk}|^2|\tilde{g}_{mnk}|^2\} \\
&\overset{(a)}{=} 2N_r c_{mk}^4\left(\beta_{mk}^2 + \frac{2}{\tau_p}\beta_{mk}\sum_{k'=1}^{K_{int}}\beta_{i,mk'} + \frac{2\sigma_z^2}{\tau_p}\beta_{mk} + \frac{1}{\tau_p}\sum_{k'=1}^{K_{int}}\beta_{i,mk'}^2\right. \\
&\qquad\left. + \frac{1}{\tau_p^2}\sum_{k'=1}^{K_{int}}\sum_{k''\neq k'}^{K_{int}}\beta_{i,mk'}\beta_{i,mk''} + \frac{2\sigma_z^2}{\tau_p^2}\sum_{k'=1}^{K_{int}}\beta_{i,mk'} + \frac{\sigma_z^4}{\tau_p^2}\right) \\
&\qquad + (N_r^2 - N_r)c_{mk}^2\beta_{mk}^2 + N_r c_{mk}\beta_{mk}(1 - c_{mk})\beta_{mk} \\
&= 2N_r c_{mk}^4\left(\beta_{mk}^2 + \frac{2}{\tau_p}\beta_{mk}\sum_{k'=1}^{K_{int}}\beta_{i,mk'} + \frac{2\sigma_z^2}{\tau_p}\beta_{mk} + \frac{1}{\tau_p}\sum_{k'=1}^{K_{int}}\beta_{i,mk'}^2\right. \\
&\qquad\left. + \frac{1}{\tau_p^2}\sum_{k'=1}^{K_{int}}\sum_{k''\neq k'}^{K_{int}}\beta_{i,mk'}\beta_{i,mk''} + \frac{2\sigma_z^2}{\tau_p^2}\sum_{k'=1}^{K_{int}}\beta_{i,mk'} + \frac{\sigma_z^4}{\tau_p^2}\right) \\
&\qquad + (N_r^2 - 2N_r)c_{mk}^2\beta_{mk}^2 + N_r c_{mk}\beta_{mk}^2
\end{aligned}
$$

where we substitute (5.38) into step (a). The off-diagonal elements in (5.41) are calculated, when $k' = k$ and $m' \neq m$:

$$
\begin{aligned}
&\mathbb{E}\{\boldsymbol{v}_{mk}\boldsymbol{g}_{mk'}\boldsymbol{g}_{m'k'}^H\boldsymbol{v}_{m'k}^H\} = \mathbb{E}\{\hat{\boldsymbol{g}}_{mk}^H\boldsymbol{g}_{mk}\boldsymbol{g}_{m'k}^H\hat{\boldsymbol{g}}_{m'k}\} \\
&= \mathbb{E}\left\{\sum_{n=1}^{N_r}\hat{g}^*_{mnk}g_{mnk}\sum_{n'=1}^{N_r}g^*_{m'n'k}\hat{g}_{m'n'k}\right\} \\
&= \mathbb{E}\left\{\sum_{n=1}^{N_r}|\hat{g}_{mnk}|^2\sum_{n'=1}^{N_r}|\hat{g}_{m'n'k}|^2\right\} \\
&= N_r^2 c_{mk}\beta_{mk}c_{m'k}\beta_{m'k}
\end{aligned}
\tag{5.44}
$$

If $k' \neq k$, the off-diagonal elements in (5.41) are zero. Then, the term $\mathbb{E}\{\boldsymbol{\gamma}_{i,kk'}\boldsymbol{\gamma}_{i,kk'}^H\}$ in (5.22) can be expressed by:

$$\mathbb{E}\{\boldsymbol{\gamma}_{i,kk'}\boldsymbol{\gamma}_{i,kk'}^H\} \tag{5.45}$$
$$= \begin{bmatrix} \mathbb{E}\{\boldsymbol{v}_{1k}\boldsymbol{g}_{i,1k'}\boldsymbol{g}_{i,1k'}^H\boldsymbol{v}_{1k}^H\} & \cdots & \mathbb{E}\left\{\boldsymbol{v}_{1k}\boldsymbol{g}_{i,1k'}\boldsymbol{g}_{i,M_{coor}k'}^H\boldsymbol{v}_{M_{coor}k}^H\right\} \\ \vdots & \ddots & \vdots \\ \mathbb{E}\{\boldsymbol{v}_{M_{coor}k}\boldsymbol{g}_{i,M_{coor}k'}\boldsymbol{g}_{i,1k'}^H\boldsymbol{v}_{1k}^H\} & \cdots & \mathbb{E}\left\{\boldsymbol{v}_{M_{coor}k}\boldsymbol{g}_{i,M_{coor}k'}\boldsymbol{g}_{i,M_{coor}k'}^H\boldsymbol{v}_{M_{coor}k}^H\right\} \end{bmatrix}$$

The on-diagonal elements can be expressed by:

$$\begin{aligned} &\mathbb{E}\{\boldsymbol{v}_{mk}\boldsymbol{g}_{i,mk'}\boldsymbol{g}_{i,mk'}^H\boldsymbol{v}_{mk}^H\} = \mathbb{E}\{\hat{\boldsymbol{g}}_{mk}^H\boldsymbol{g}_{i,mk'}\boldsymbol{g}_{i,mk'}^H\hat{\boldsymbol{g}}_{mk}\} \\ &= \mathbb{E}\left\{\sum_{n=1}^{N_r}\hat{g}_{mnk}^* g_{i,mnk'}\sum_{n'=1}^{N_r} g_{i,mn'k'}^*\hat{g}_{mn'k}\right\} \\ &= \sum_{n=1}^{N_r}\mathbb{E}\left\{|\hat{g}_{mnk}|^2\left|g_{i,mnk'}\right|^2\right\} \\ &= N_r c_{mk}\beta_{mk}\beta_{i,mk'} \end{aligned} \tag{5.46}$$

The off-diagonal elements is derived when $m' \neq m$:

$$\begin{aligned} &\mathbb{E}\{\boldsymbol{v}_{mk}\boldsymbol{g}_{i,mk'}\boldsymbol{g}_{i,m'k'}^H\boldsymbol{v}_{m'k}^H\} = \mathbb{E}\{\hat{\boldsymbol{g}}_{mk}^H\boldsymbol{g}_{i,mk'}\boldsymbol{g}_{i,m'k'}^H\hat{\boldsymbol{g}}_{m'k}\} \\ &= \mathbb{E}\left\{\sum_{n=1}^{N_r}\hat{g}_{mnk}^* g_{i,mnk'}\sum_{n'=1}^{N_r} g_{i,m'n'k'}^*\hat{g}_{m'n'k}\right\} = 0 \end{aligned} \tag{5.47}$$

The term $\boldsymbol{D}_k$ is computed by:

$$\begin{aligned} &\boldsymbol{D}_k = diag\left(\mathbb{E}\{\|\boldsymbol{v}_{1k}\|^2\}, \mathbb{E}\{\|\boldsymbol{v}_{2k}\|^2\}, \dots, \mathbb{E}\left\{\left\|\boldsymbol{v}_{M_{coor}k}\right\|^2\right\}\right) \\ &= diag\left(\mathbb{E}\{\hat{\boldsymbol{g}}_{1k}^H\hat{\boldsymbol{g}}_{1k}\}, \dots, \mathbb{E}\{\hat{\boldsymbol{g}}_{M_{coor}k}^H\hat{\boldsymbol{g}}_{M_{coor}k}\}\right) \\ &= diag\left(N_r c_{1k}\beta_{1k}, \dots, N_r c_{M_{coor}k}\beta_{M_{coor}k}\right) \end{aligned} \tag{5.48}$$

Finally, the closed-form combining vector for large scale fading decoding with local MRC in the multiple-antenna AP case is derived by substituting (5.40), (5.42), (5.43), (5.44), (5.46), (5.47) and (5.48) into (5.22). From the results, this combining vector only depends on the large scale fading coefficients.

Chapter 6

Conclusion and Future Research

6.1 Conclusion

In this thesis, we have studied decentralised distributed massive multiple-input multiple-output (DD-MaMIMO), exploring its dimensions with regard to varying processing locations and fronthaul capacity assumptions. Our comprehensive analysis has demonstrated the potential of DD-MaMIMO in augmenting spectral efficiency (SE) and satisfying the demands of six generation (6G) communication networks. For a clearer understanding of DD-MaMIMO, we summarise each chapter as follows.

In Chapter 2, literature review has been presented that primarily introduces various architectures of radio access networks (RANs), the MaMIMO technique with its diverse deployment approaches, and the general basics of optimum uniform quantisation. This review suggests that cell-free MaMIMO, as an approach to perform MaMIMO, can be a promising technique for fifth generation (5G) network and beyond. However, cell-free MaMIMO encounters challenges such as the network edge effect, the quest for low latency, and the need for a flexible yet scalable implementation in Open RAN. Guided by these insights, the primary objective of this thesis is to design a network that solves these issues and fulfils the requirements of the next generation.

To address these challenges, we have proposed the concept of DD-MaMIMO in Chapter 3 and compared it to cell-free MaMIMO. Our studies have demonstrated that the network edge effect can be mitigated by appropriately defining the coordination region. Furthermore, moving processing back to the network edge can reduce the fronthaul load and computational complexity at the central processing unit (CPU), resulting in the decreased latency. For a whole system evaluation, we have investigated the performance of DD-MaMIMO with unlimited-capacity fronthaul on the uplink, without resorting to channel hardening assumption. Numerical results have been provided to demonstrate that DD-MaMIMO outperforms cell-free MaMIMO, especially as the radius of coordination region increases. In additional, our findings

have revealed that applying minimum mean square error (MMSE) estimator for data detection is more effective than using maximum ratio combining (MRC), even if each access point (AP) is equipped with only a single antenna. Moreover, we have discussed that DD-MaMIMO is a scalable network and offers flexible implementation in Open RAN. Finally, an innovative pilot allocation algorithm tailored for neighbouring networks has been proposed, taking into account users located in overlapping areas across many coordination regions.

In general, optical fibre is used to convey the digital signals between the APs and the EPU which requires quantisation to convert the analogue signal to the digital. Therefore, in Chapter 4, we have researched DD-MaMIMO with a limited-capacity fronthaul. For obtaining the channel state information (CSI), we have considered two strategies: quantise-and-estimate (QE) and estimate-and-quantise (EQ). The performance for both strategies has been analysed by exploiting Bussgang decomposition. In particular, in the QE form, the non-Gaussian distribution of the input signal of quantiser has been thoroughly studied. The probability density function (pdf) for this non-Gaussian distributed input signal has been derived. Then, we have exploited this pdf to further derive the closed-form expressions in order to find the optimum quantisation step interval when the elements of pilot sequences follows the complex Gaussian distribution. It can achieve a smaller MSE of channel estimation than using the optimum quantisation step interval with Gaussian distributed input signal assumption. We have also explored the use of binary pilot sequences (plus and minus one) in the QE form. The performance did not show a very large difference between the two types of pilots. Finally, we have compared numerical results for different CSI acquisition strategies with different number of quantisation bits. It has shown that 2-bit quantisation with QE strategy can outperform the SE of cell-free MaMIMO with ideal fronthaul for 90% likely users.

In Chapter 5, to mitigate the computational complexity at the EPU, we have investigated local estimation at the AP within DD-MaMIMO when using an unlimited-capacity fronthaul. With this approach, the fronthaul load diminishes since there is no need to transmit CSI to the EPU. We have explored two decoding strategies: large scale fading decoding and the straightforward method of averaging locally estimated data at the EPU. It has been observed that using large scale fading decoding yields superior results. However, this requires the channel statistics to be known at the EPU. We have also contrasted this local estimation scheme with the decentralised processing scheme.

Our numerical evaluations have indicated that the performance gap between the local estimation using large scale fading decoding and the decentralised case is not significant, if each AP is equipped with a large number of antennas. This phenomenon exhibits the adaptability of DD-MaMIMO with various processing locations to the flexible deployment requirements of Open RAN, meeting the demands of forthcoming wireless network.

6.2 Future Research

We recognise that this thesis has left some prospective challenges open for further investigate. The future research directions could be categorised into short-term and long-term objectives.

The short-term objectives:

- We adopted one-slope path loss model for the large scale fading channels in this thesis, in order to provide a reasonable comparison with previous research of cell-free MaMIMO in [24]. This may not be very suitable for an ultra-dense network in the future. This is because in the conventional cellular system, the signal-to-interference ratio (SIR) remains constant as the cell size scales. However, this becomes invalid, since the interference increases faster than signal power, when the cell size decreases further. Therefore, we could apply the multi-slope path loss model to DD-MaMIMO with an ultra-dense network.

- In Chapter 3, we observed that the performance for the SE improves with the increase of the coordination radius in DD-MaMIMO. The reason is, in a larger coordination region, the number of serving APs is greater, resulting in the rise of signal power and the decrease of interference. However, the radius of the coordination region cannot be grown infinitely. It should have a range limitation with respect to the length of pilot sequences to ensure that pilots assigned to all user terminals (UTs) within coordination region are completely orthogonal with each other. Furthermore, if the length of pilots is less than the number of UTs within the coordination region, optimal pilot allocation methods such as greedy pilot allocation needs to be considered to mitigate the pilot contamination. We

could make a comparison for both cases to find whether there is a significant discrepancy. Based on this result, a trade-off could be evaluated.

- In Chapter 4, we used the true pdf of the input signal of the quantiser to determine the optimal quantisation step interval. However, the MSE of channel performance did not show a significant improvement, compared to when we assumed a Gaussian distribution for the input signal. This outcome might happen when the number of users, including the interfering ones, is sufficiently large to align with a normal distribution by leveraging the central limit theorem. This phenomenon could be researched for a more comprehensive understanding. Moreover, the approach introduced in this chapter could be particularly relevant to rural areas where the number of devices accessing the network is limited.

- We evaluated the performance of quantisation in Chapter 4 with the least 2 bits, but 1-bit quantisation also could be explored for the worst quantisation resolution circumstance. This might be suitable for local estimation in DD-MaMIMO as stated in Chapter 5 where a large fronthaul capacity is not required.

The long-term objectives:

- In this thesis, we investigated the performance of DD-MaMIMO in terms of uncorrelated Rayleigh fading channel. But, in the realistic scenario, the channel between different antennas at one AP in MaMIMO system could be correlated. Though this would not affect the single-antenna AP case in this thesis, it could degrade the system performance with regard to the AP equipped a large number of antennas. For example, the comparison between the uncorrelated and correlated Rayleigh fading channel for cell-free MaMIMO was discussed in [80], indicating the effect of spatial correlation. Additionally, in [14], the correlated shadowing also caused the performance degradation. Therefore, more thorough research on the performance of DD-MaMIMO taking account of correlated channels and shadowing could be done in the future.

- We focused on the flat fading channel in this thesis which may not be realistic. In practice, due to reflections, refractions or scattering from objectives, the AP receives the signal via multiple paths. This multipath propagation can make the

channel appear frequency-selective. To cope with this multipath effect, orthogonal frequency-division multiplexing (OFDM) can be applied. Therefore, the research of using OFDM in DD-MaMIMO could be developed in the future.

- To improve system performance, power control should be implemented in DD-MaMIMO. Furthermore, it is important to discuss the scalability aspects of the power control.

- In this thesis, we derived an expression for the achievable SE referencing the upper bound of Shannon capacity, which represents a theoretical value. However, in real life, the selected modulation determines how close a practical communications system can approach this theoretical limit. Therefore, taking account of modulation could be an avenue for future research. Moreover, in the scenario with low resolution, especially when using only one or two quantisation bits, an adaptive modulation scheme should be chosen to strike a balance between data transmission efficiency and accuracy.

- We investigated two distinct types of pilots in QE form. While numerical results for both types were quite similar, there remains an opportunity to delve into the specific application scenarios where each pilot type might be best suited.

- In addition to applying quantisation, the compressed sensing is another emergent technique applicable to the limited-capacity fronthaul link. By using compressed sensing, the transmitted data can be sparsely represented which reduces the transmission bits and the power consumption [81]. However, the optimisation problems often needs to be solved in compressed sensing when the signals are decoded, which may increase the computational complexity [81]. Hence, the comparison for the application of quantisation and compressed sensing can be researched. The challenge of the compressed sensing can be solved in the future.

- In Chapter 5, we only introduced local estimation in DD-MaMIMO with ideal fronthaul, which is not realistic. Therefore, quantisation for this local estimation case could be developed in the future work. There are several points needed to be considered, such as how to express the locally estimated data in a

quantisation form with two different decoding methods, how to quantise the channel statistics for large scale fading decoding, and if the original Bussgang decomposition could be directly used to the locally estimated data.

- As stated in [7], artificial intelligence (AI) or machine learning (ML) could be applied in 6G for improving signal processing. For example, AI/ML can be used to predict more accurate CSI with rapidly changing channel conditions, which becomes significant especially for the millimetre wave and terahertz (THz) frequency. Though we did not discuss this in this thesis, it could be considered for DD-MaMIMO in the future.

Glossary

3GPP	Third Generation Partnership Project
5G	Fifth Generation
6G	Sixth Generation
AI	Artificial Intelligence
AP	Access Point
AR	Augmented Reality
BBU	Baseband Unit
BS	Base Station
CDF	Cumulative Distribution Function
CoMP	Coordinated Multipoint
CPU	Central Processing Unit
C-RAN	Cloud Radio Access Network
CSI	Channel State Information
CU	Central Unit
DD-MaMIMO	Decentralised Distributed Massive MIMO
DU	Distributed Unit
eMBB	Enhanced Mobile Broadband
EPU	Edge Processing Unit
EQ	Estimate-and-Quantise
FDD	Frequency-Division Duplexing
F-RAN	Fog Radio Access Network
GSM	Global System for Mobile Communications
HD	High Definition
HRLLC	Hyper Reliable and Low-Latency Communication
i.i.d	Independent and Identically Distributed
IMT	International Mobile Telecommunications
IoE	Internet of Everything
IoT	Internet of Things

IP	Internet Protocol
ITU	International Telecommunication Unit
LS	Least Square
MaMIMO	Massive Multiple-Input Multiple-Output
MMSE	Minimum Mean Square Error
mMTC	Massive Machine-Type Communication
mmWave	Millimetre Wave
MRC	Maximum-Ratio Combining
MU-MIMO	Multi-user MIMO
NR	New Radio
OFDM	Orthogonal Frequency Division Multiplexing
QE	Quantise-and-Estimate
QoS	Quality of Service
RAN	Radio Access Network
RIC	Radio Access Network Intelligent Controller
RRH	Remote Radio Head
RU	Radio Unit
RV	Random Variable
SDNR	Signal-to-Distortion Ratio
SE	Spectral Efficiency
SIDNR	Signal-to-Interference-and-Distortion-plus-Noise Ratio
SINR	Signal-to-Interference-plus-Noise Ratio
SVD	Singular Value Decomposition
TDD	Time-Division Duplexing
uMBB	Ubiquitous Mobile Broadband
URLLC	Ultra-Reliable Low-Latency Communication
UT	User Terminal
VR	Virtual Reality
XR	Extended Reality
ZF	Zero-Forcing

Reference

[1] *The ITU-R Framework For IMT-2030*, ITU-R Working Party 5D, Jul. 2023. [Online]. Available: https://www.itu.int/en/ITU-R/study-groups/rsg5/rwp5d/imt-2030/Documents/IMT-2030%20Framework_WP%205D%20Management%20Team.pdf [Accessed: 4 August 2023].

[2] *IMT vision—framework and overall objectives of the future development of IMT for 2020 and beyond*, ITU-R M.2083-0, Sep. 2015. [Online]. Available: https://www.itu.int/dms_pubrec/itu-r/rec/m/R-REC-M.2083-0-201509-I!!PDF-E.pdf [Accessed: 4 August 2023].

[3] W. Jiang, B. Han, M. A. Habibi, and H. D. Schotten, "The Road Towards 6G: A Comprehensive Survey," IEEE Open Journal of the Communications Society, vol. 2, pp. 334–366, 2021.

[4] S. Chen, Y.-C. Liang, S. Sun, S. Kang, W. Cheng, and M. Peng, "Vision, Requirements, and Technology Trend of 6G: How to Tackle the Challenges of System Coverage, Capacity, User Data-Rate and Movement Speed," IEEE Wirel. Commun., vol. 27, no. 2, pp. 218–228, Apr. 2020.

[5] S. Dang, O. Amin, B. Shihada, and M.-S. Alouini, "What should 6G be?," arXiv [cs.NI], Jul. 29, 2020. doi: 10.36227/techrxiv.10247726.

[6] H. Tataria, M. Shafi, A. F. Molisch, M. Dohler, H. Sjöland, and F. Tufvesson, "6G Wireless Systems: Vision, Requirements, Challenges, Insights, and Opportunities," Proc. IEEE, vol. 109, no. 7, pp. 1166–1199, Jul. 2021

[7] M. Z. Chowdhury, M. Shahjalal, S. Ahmed, and Y. M. Jang, "6G Wireless Communication Systems: Applications, Requirements, Technologies, Challenges,

and Research Directions," IEEE Open Journal of the Communications Society, vol. 1, pp. 957–975, 2020.

[8] A. Weissberger. "Summary of ITU-R Workshop on 'IMT for 2030 and beyond' (aka '6G')". 20 Jun 2022. [Blog entry]. *IEEE Communications Society Technology Blog.* Available: https://techblog.comsoc.org/2022/06/20/summary-of-itu-r-workshop-on-imt-for-2030-and-beyond-aka-6g/ [Accessed: 4 August 2023]

[9] Huawei. Y. Chen, P. Zhu, and W. Tong. (2023, Jun). *ITU-R WP5D Completed the Recommendation Framework for IMT-2030 (Global 6G Vision)*. HuaweiTech. [Online]. Available: https://www.huawei.com/uk/huaweitech/future-technologies/itu-r-wp5d-completed-recommendation-framework-imt-2030 [Accessed: 4 August 2023].

[10] N. Kato, B. Mao, F. Tang, Y. Kawamoto, and J. Liu, "Ten Challenges in Advancing Machine Learning Technologies toward 6G," IEEE Wirel. Commun., vol. 27, no. 3, pp. 96–103, Jun. 2020.

[11] F. Liu et al., "Integrated Sensing and Communications: Toward Dual-Functional Wireless Networks for 6G and Beyond," IEEE J. Sel. Areas Commun., vol. 40, no. 6, pp. 1728–1767, Jun. 2022.

[12] D. K. Pin Tan et al., "Integrated Sensing and Communication in 6G: Motivations, Use Cases, Requirements, Challenges and Future Directions," in 2021 1st IEEE International Online Symposium on Joint Communications & Sensing (JC&S), ieeexplore.ieee.org, Feb. 2021, pp. 1–6.

[13] T. L. Marzetta, "Noncooperative Cellular Wireless with Unlimited Numbers of Base Station Antennas," IEEE Trans. Wireless Commun., vol. 9, no. 11, pp. 3590–3600, Nov. 2010.

[14] 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.

[15] A. Burr, S. Islam, J. Zhao, and M. Bashar, “Cell-free Massive MIMO with multi-antenna access points and user terminals,” in 2020 54th Asilomar Conference on Signals, Systems, and Computers, Nov. 2020, pp. 821–825.

[16] H. Q. Ngo, A. Ashikhmin, H. Yang, E. G. Larsson, and T. L. Marzetta, “Cell-Free Massive MIMO: Uniformly great service for everyone,” in 2015 IEEE 16th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC), Jun. 2015, pp. 201–205.

[17] M. Matthaiou, O. Yurduseven, H. Q. Ngo, D. Morales-Jimenez, S. L. Cotton, and V. F. Fusco, “The Road to 6G: Ten Physical Layer Challenges for Communications Engineers,” IEEE Commun. Mag., vol. 59, no. 1, pp. 64–69, Jan. 2021.

[18] Z. Chen and E. Björnson, “Channel Hardening and Favorable Propagation in Cell-Free Massive MIMO With Stochastic Geometry,” IEEE Trans. Commun., vol. 66, no. 11, pp. 5205–5219, Nov. 2018.

[19] Z. Chen and E. Bjoernson, “Can We Rely on Channel Hardening in Cell-Free Massive MIMO?,” in 2017 IEEE Globecom Workshops (GC Wkshps), Dec. 2017, pp. 1–6.

[20] G. Interdonato, E. Björnson, H. Quoc Ngo, P. Frenger, and E. G. Larsson, “Ubiquitous cell-free Massive MIMO communications,” Eurasip J. Wirel. Commun. Network., vol. 2019, no. 1, p. 197, Aug. 2019.

[21] M. Bashar, K. Cumanan, A. G. Burr, M. Debbah, and H. Q. Ngo, “On the Uplink Max–Min SINR of Cell-Free Massive MIMO Systems,” IEEE Trans. Wireless Commun., vol. 18, no. 4, pp. 2021–2036, Apr. 2019.

[22] G. Interdonato, P. Frenger, and E. G. Larsson, “Scalability Aspects of Cell-Free Massive MIMO,” in ICC 2019 - 2019 IEEE International Conference on Communications (ICC), ieeexplore.ieee.org, May 2019, pp. 1–6.

[23] E. Björnson and L. Sanguinetti, “Scalable Cell-Free Massive MIMO Systems,” IEEE Trans. Commun., vol. 68, no. 7, pp. 4247–4261, Jul. 2020.

[24] E. Björnson and L. Sanguinetti, "Making Cell-Free Massive MIMO Competitive With MMSE Processing and Centralized Implementation," IEEE Trans. Wireless Commun., vol. 19, no. 1, pp. 77–90, Jan. 2020.

[25] D. Maryopi, M. Bashar, and A. Burr, "On the Uplink Throughput of Zero Forcing in Cell-Free Massive MIMO With Coarse Quantization," IEEE Trans. Veh. Technol., vol. 68, no. 7, pp. 7220–7224, Jul. 2019.

[26] M. Bashar et al., "Uplink Spectral and Energy Efficiency of Cell-Free Massive MIMO With Optimal Uniform Quantization," IEEE Trans. Commun., vol. 69, no. 1, pp. 223–245, Jan. 2021.

[27] M. Bashar, K. Cumanan, A. G. Burr, H. Q. Ngo, E. G. Larsson, and P. Xiao, "Energy Efficiency of the Cell-Free Massive MIMO Uplink With Optimal Uniform Quantization," IEEE Transactions on Green Communications and Networking, vol. 3, no. 4, pp. 971–987, Dec. 2019.

[28] A. Burr, M. Bashar, and D. Maryopi, "Cooperative access networks: Optimum fronthaul quantization in distributed massive MIMO and cloud RAN-invited paper," 2018 IEEE 87th Vehicular, 2018, [Online]. Available: https://ieeexplore.ieee.org/abstract/document/8417560/

[29] A. Burr and D. Maryopi, "On the modelling of coarse vector quantization in distributed massive mimo," 2021 IEEE Statistical Signal Processing, 2021, [Online]. Available: https://ieeexplore.ieee.org/abstract/document/9513792/

[30] A. Burr, M. Bashar, and D. Maryopi, "Ultra-dense radio access networks for smart cities: Cloud-RAN, Fog-RAN and 'cell-free' massive MIMO," 4th International Workshop of CorNer: Communications for Networked Smart Cities, PIMRC , Bologna, September 2018. Available: https://arxiv.org/abs/1811.11077

[31] S. Niknam et al., "Intelligent O-RAN for Beyond 5G and 6G Wireless Networks," in 2022 IEEE Globecom Workshops (GC Wkshps), Dec. 2022, pp. 215–220.

[32] L. Gavrilovska, V. Rakovic, and D. Denkovski, "From cloud RAN to open RAN," Wirel. Pers. Commun., vol. 113, no. 3, pp. 1523–1539, Aug. 2020.

[33] O. Y. Bursalioglu, G. Caire, R. K. Mungara, H. C. Papadopoulos, and C. Wang, "Fog Massive MIMO: A User-Centric Seamless Hot-Spot Architecture," IEEE Trans. Wireless Commun., vol. 18, no. 1, pp. 559–574, Jan. 2019.

[34] S. Chen, J. Zhao, and Y. Peng, "The development of TD-SCDMA 3G to TD-LTE-advanced 4G from 1998 to 2013," IEEE Wirel. Commun., vol. 21, no. 6, pp. 167–176, Dec. 2014.

[35] I. F. Akyildiz, D. M. Gutierrez-Estevez, and E. C. Reyes, "The evolution to 4G cellular systems: LTE-Advanced," Physical Communication, vol. 3, no. 4, pp. 217–244, Dec. 2010.

[36] J. G. Andrews et al., "What Will 5G Be?," IEEE J. Sel. Areas Commun., vol. 32, no. 6, pp. 1065–1082, Jun. 2014.

[37] Y. L. Lee, D. Qin, L. C. Wang, and G. H. Sim, "6G Massive Radio Access Networks: Key Applications, Requirements and Challenges," IEEE Open Journal of Vehicular Technology, 2020.

[38] W. Saad, M. Bennis, and M. Chen, "A Vision of 6G Wireless Systems: Applications, Trends, Technologies, and Open Research Problems," IEEE Netw., vol. 34, no. 3, pp. 134–142, May 2020.

[39] M. Giordani, M. Polese, M. Mezzavilla, S. Rangan, and M. Zorzi, "Toward 6G Networks: Use Cases and Technologies," IEEE Commun. Mag., vol. 58, no. 3, pp. 55–61, Mar. 2020.

[40] N. Rajatheva, I. Atzeni, E. Bjornson, and A. Bourdoux, "White Paper on Broadband Connectivity in 6G," arXiv preprint arXiv, 2020, [Online]. Available: https://arxiv.org/abs/2004.14247

[41] 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. 70371–70421, 2019.

[42] S. K. Singh, R. Singh, and B. Kumbhani, "The Evolution of Radio Access Network Towards Open-RAN: Challenges and Opportunities," in 2020 IEEE Wireless

Communications and Networking Conference Workshops (WCNCW), Apr. 2020, pp. 1–6.

[43] A. Checko et al., "Cloud RAN for Mobile Networks—A Technology Overview," IEEE Communications Surveys Tutorials, vol. 17, no. 1, pp. 405–426, 2015.

[44] "C-RAN: the road towards green RAN," China Mobile Research Institute, Beijing, China, Oct. 2011, Tech. Rep.

[45] C.-P. Li, J. Jiang, W. Chen, T. Ji, and J. Smee, "5G ultra-reliable and low-latency systems design," in 2017 European Conference on Networks and Communications (EuCNC), Jun. 2017, pp. 1–5.

[46] M. Peng, S. Yan, K. Zhang, and C. Wang, "Fog-computing-based radio access networks: issues and challenges," IEEE Netw., vol. 30, no. 4, pp. 46–53, Jul. 2016.

[47] Y.-Y. Shih, W.-H. Chung, A.-C. Pang, T.-C. Chiu, and H.-Y. Wei, "Enabling Low-Latency Applications in Fog-Radio Access Networks," IEEE Netw., vol. 31, no. 1, pp. 52–58, Jan. 2017.

[48] Y.-J. Ku et al., "5G Radio Access Network Design with the Fog Paradigm: Confluence of Communications and Computing," IEEE Commun. Mag., vol. 55, no. 4, pp. 46–52, Apr. 2017.

[49] M. K. Karakayali, G. J. Foschini, and R. A. Valenzuela, "Network coordination for spectrally efficient communications in cellular systems," IEEE Wirel. Commun., vol. 13, no. 4, pp. 56–61, Aug. 2006.

[50] D. Gesbert, S. Hanly, H. Huang, S. Shamai Shitz, O. Simeone, and W. Yu, "Multi-Cell MIMO Cooperative Networks: A New Look at Interference," IEEE J. Sel. Areas Commun., vol. 28, no. 9, pp. 1380–1408, Dec. 2010.

[51] S. Venkatesan, A. Lozano, and R. Valenzuela, "Network MIMO: Overcoming Intercell Interference in Indoor Wireless Systems," in 2007 Conference Record of the Forty-First Asilomar Conference on Signals, Systems and Computers, Nov. 2007, pp. 83–87.

[52] R. Irmer et al., “Coordinated multipoint: Concepts, performance, and field trial results,” IEEE Commun. Mag., vol. 49, no. 2, pp. 102–111, Feb. 2011.

[53] “Study on new radio access technology: Radio access architecture and interfaces,” 3rd Generation Partnership Project (3GPP), Technical Specification (TS) 38.801, April 2017, version 14.0.0.

[54] M. Polese, L. Bonati, S. D’Oro, S. Basagni, and T. Melodia, “Understanding O-RAN: Architecture, Interfaces, Algorithms, Security, and Research Challenges,” IEEE Communications Surveys & Tutorials, vol. 25, no. 2, pp. 1376–1411, Secondquarter 2023.

[55] L. Bonati, S. D’Oro, M. Polese, S. Basagni, and T. Melodia, “Intelligence and Learning in O-RAN for Data-Driven NextG Cellular Networks,” IEEE Commun. Mag., vol. 59, no. 10, pp. 21–27, Oct. 2021.

[56] O-RAN Alliance, “Control, User and Synchronization Plane Specification,” O-RAN Fronthaul Working Group, O-RAN-WG4.CUS.0-v07.02, Sep. 2022.

[57] B. M. Hochwald, T. L. Marzetta, and V. Tarokh, “Multiple-antenna channel hardening and its implications for rate feedback and scheduling,” IEEE Trans. Inf. Theory, vol. 50, no. 9, pp. 1893–1909, Sep. 2004.

[58] T. L. Marzetta, E. G. Larsson, H. Yang, and H. Q. Ngo, *Fundamentals of Massive MIMO*, Cambridge University Press, 2016.

[59] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, “Aspects of favorable propagation in Massive MIMO,” in 2014 22nd European Signal Processing Conference (EUSIPCO), Sep. 2014, pp. 76–80.

[60] H. Q. Ngo and E. G. Larsson, “No Downlink Pilots Are Needed in TDD Massive MIMO,” IEEE Trans. Wireless Commun., vol. 16, no. 5, pp. 2921–2935, May 2017.

[61] E. Björnson, E. G. Larsson, and T. L. Marzetta, “Massive MIMO: ten myths and one critical question,” IEEE Commun. Mag., vol. 54, no. 2, pp. 114–123, Feb. 2016.

[62] E. G. Larsson, O. Edfors, F. Tufvesson, and T. L. Marzetta, “Massive MIMO for next generation wireless systems,” IEEE Commun. Mag., vol. 52, no. 2, pp. 186–195, Feb. 2014.

[63] A. Ashikhmin, L. Li, and T. L. Marzetta, “Interference Reduction in Multi-Cell Massive MIMO Systems With Large-Scale Fading Precoding,” IEEE Trans. Inf. Theory, vol. 64, no. 9, pp. 6340–6361, Sep. 2018.

[64] T. S. Rappaport, Wireless Communications: Principles and Practice, Englewood Cliffs, NJ, USA: Prentice-Hall, 2002.

[65] J. Bussgang, “Crosscorrelation Functions of Amplitude-Distorted Gaussian Signals,” RLE Technical Reports, vol. 216, 1952.

[66] P. Zillmann and G. R. Fettweis, “On the capacity of multicarrier transmission over nonlinear channels,” in 2005 IEEE 61st Vehicular Technology Conference, May 2005, vol. 2, pp. 1148–1152 Vol. 2.

[67] E. Björnson, J. Hoydis, M. Kountouris, and M. Debbah, “Massive MIMO Systems With Non-Ideal Hardware: Energy Efficiency, Estimation, and Capacity Limits,” IEEE Trans. Inf. Theory, vol. 60, no. 11, pp. 7112–7139, Nov. 2014.

[68] M. Fozooni, M. Matthaiou, E. Bjornson, and T. Q. Duong, “Performance Limits of MIMO Systems with Nonlinear Power Amplifiers,” in 2015 IEEE Global Communications Conference (GLOBECOM), Dec. 2015, pp. 1–7.

[69] I. Iofedov and D. Wulich, “MIMO–OFDM With Nonlinear Power Amplifiers,” IEEE Trans. Commun., vol. 63, no. 12, pp. 4894–4904, Dec. 2015.

[70] P. Zillmann, “Relationship Between Two Distortion Measures for Memoryless Nonlinear Systems,” IEEE Signal Process. Lett., vol. 17, no. 11, pp. 917–920, Nov. 2010.

[71] Further Advancements for E-UTRA Physical Layer Aspects (Release 9), document TS 36.814, 3GPP, Mar. 2017.

[72] H. Q. Ngo, L. Tran, T. Q. Duong, M. Matthaiou, and E. G. Larsson, "On the Total Energy Efficiency of Cell-Free Massive MIMO," IEEE Transactions on Green Communications and Networking, vol. 2, no. 1, pp. 25–39, Mar. 2018.

[73] E. Björnson, J. Hoydis, and L. Sanguinetti, "Massive MIMO networks: Spectral, energy, and hardware efficiency," in Foundations and Trends in Signal Processing, vol. 11, nos. 3–4. 2017, pp. 154–655.

[74] I. S. Gradshteyn and I. M. Ryzhik, Table of Integral, Series and Products, 7th ed., A Jeffrey and D. Zwillinger, Eds. Amsterdam, The Netherlands: Elsevier, 2007

[75] S. Nadarajah and T. K. Pogány, 'On the distribution of the product of correlated normal random variables', C. R. Math., vol. 354, no. 2, pp. 201–204, Feb. 2016.

[76] G. Interdonato, M. Karlsson, E. Björnson, and E. G. Larsson, "Local Partial Zero-Forcing Precoding for Cell-Free Massive MIMO," IEEE Trans. Wireless Commun., pp. 1–1, 2020.

[77] J. Zhang, J. Zhang, E. Björnson, and B. Ai, "Local Partial Zero-Forcing Combining for Cell-Free Massive MIMO Systems," IEEE Trans. Commun., vol. 69, no. 12, pp. 8459–8473, Dec. 2021.

[78] A. Ashikhmin and T. Marzetta, "Pilot contamination precoding in multi-cell large scale antenna systems," in 2012 IEEE International Symposium on Information Theory Proceedings, Jul. 2012, pp. 1137–1141.

[79] E. Nayebi, A. Ashikhmin, T. L. Marzetta, and B. D. Rao, "Performance of cell-free massive MIMO systems with MMSE and LSFD receivers," in 2016 50th Asilomar Conference on Signals, Systems and Computers, Nov. 2016, pp. 203–207.

[80] W. Fan, J. Zhang, E. Bjornson, S. Chen, and Z. Zhong, "Performance Analysis of Cell-Free Massive MIMO Over Spatially Correlated Fading Channels," in ICC 2019 - 2019 IEEE International Conference on Communications (ICC), ieeexplore.ieee.org, May 2019, pp. 1–6.

[81] Z. Han, H. Li, and W. Yin, Compressive Sensing for Wireless Networks. Cambridge University Press, 2013.